

A08.02 Sluttrapport HALv2-prosjektet

Versjon 0.90 2010-02-28
Godkjent av prosjekteier og USITs ledermøte.

Innhold

- [A08.02 Sluttrapport HALv2-prosjektet](#)
 - [1 Etablering og gjennomføring av HALv2-prosjektet](#)
 - [1.1 Formål](#)
 - [1.2 Resultatmål og leveranser](#)
 - [1.3 Milepæler](#)
 - [1.4 Oppstart](#)
 - [1.5 Omfang og ressurser](#)
 - [1.6 Endring i prosjektplaner underveis](#)
 - [1.7 Risiko](#)
 - [1.8 Timeregnskap](#)
 - [2 Prosjektarbeid](#)
 - [3 Samarbeid](#)
 - [3.1 Samarbeid innad i prosjektet](#)
 - [3.2 Samarbeid med USIT](#)
 - [3.3 Samarbeid med eksterne](#)
 - [4 Leveranser](#)
 - [4.1 Interessentanalyse](#)
 - [4.2 Oversikt over tjenester](#)
 - [4.3 HA-løsninger i markedet og USITs kravspesifikasjon til slike løsninger](#)
 - [4.3.1 Trender og produkter innen høytilgjengelighet](#)
 - [4.3.1.1 Applikasjoner med innebygget HA-funksjonalitet](#)
 - [4.3.1.2 Virtualisering](#)
 - [4.3.2 Oversikt over HA-løsninger](#)
 - [4.3.3 Kravspesifikasjon](#)
 - [4.4 Avstemming mot kravspesifikasjon](#)
 - [4.4.1 Avstemming av krav \(dvs ikke ønsker\)](#)
 - [4.4.2 Avstemming av ønsker](#)
 - [4.4.3 Tilleggstester](#)
 - [4.5 Utvalg av løsninger for test](#)
 - [4.6 Gjennomføring av tester](#)
 - [4.6.1 DataGuard](#)
 - [4.6.2 RHCS på RHEL5](#)
 - [4.6.3 RHCS på RHEL6](#)
 - [4.6.4 Nagios](#)
 - [4.7 Oppsummering fra testene](#)
 - [4.8 Samarbeid med leverandøren](#)
 - [5 Konklusjon](#)
 - [6 Anbefalinger](#)
 - [7 Veien videre](#)
 - [8 Appendix](#)

1 Etablering og gjennomføring av HALv2-prosjektet

1.1 Formål

På USITs ledermøte den [25. mai 2010](#) ble det vedtatt å planlegge oppstart av HALv2-prosjektet. Prosjektets formål er å komme med en anbefaling til teknologivalg slik at USIT kan tilby stabile, sikre og høytilgjengelige tjenester som kjører på industristandard maskinvare med Linux som operativsystem.

Bakgrunnen for prosjektet er at en av dagens foretrukne teknologier, *ServiceGuard* for Linux, blir [desupportert fra leverandøren](#) i utgangen av april 2014.

1.2 Resultatmål og leveranser

Fra mandatet til prosjektet:

Med utgangspunkt i en interessentanalyse skal det gjøres en avgrensning av hvilke typer tjenester prosjektet skal omfatte. For de valgte tjenestene, skal det lages en kravspesifikasjon til en høytilgjengelighetsløsning. Dersom det er hensiktsmessig, kan prosjektet stille ulike krav til høytilgjengelighetsløsningen basert på en gruppering av tjenestene.

Alternativene skal vurderes opp mot funksjonalitet, kompleksitet, innføringkostnader, driftskostnader, levetid og sikkerhet.

Prosjektet kan anbefale ulike høytilgjengelighetsløsninger basert på en gruppering av tjenestene. USITs lastbalanseringsløsning skal inngå som en av de løsningene som evalueres.

1.3 Milepæler

Prosjektet skulle inneholde følgende milepæler:

- **M01** Når mandat er godkjent av USITs ledermøte.
- **M02** Når USITs ledermøte har godkjent oppstart av prosjektet.
- **M03** Når interessentanalyse er godkjent av prosjekteier.
- **M04** Når omfanget av tjenester prosjektet skal omfatte er godkjent av prosjekteier.
- **M05** Når kravspesifikasjonene er godkjent av prosjekteier.
- **M06** Når et utvalg kandidatteknologier er definert og godkjent av prosjekteier.
- **M07** Når testing av kandidatteknologier er godkjent av prosjekteier.
- **M08** Når anbefaling og sluttrapport er godkjent av prosjekteier.

1.4 Oppstart

På ledermøte 7. september 2010 [vedtok](#) ledermøtet oppstart av prosjektet.

1.5 Omfang og ressurser

Prosjektet ble opprinnelig estimert til i underkant av 800 timer, og var planlagt avsluttet i midten av januar. Det ble inngått ressursavtale for følgende personer:

- Anders Odberg (SAPP/BSD)
- Mikael Dalsgard (SAPP/BSD)
- Trond Hasle Amundsen (SAPP/BSD)
- Morten Werner Forsbring (SAPP/BSD)
- Ellef Thurman (SAPP/DBA)
- Gunnar Steira Mikkelsen (SAPP/DBA) (prosjektleder)

1.6 Endring i prosjektplaner underveis

Før jul ble det klart at vi måtte justere prosjektplanen. Årsaken var at oppsett av testmiljøet tok mer tid enn planlagt og at det var mer tidkrevende enn forutsatt å få inn konsulenthjelp til kvalitetssikring. På dette tidspunktet hadde prosjekteier allerede bestemt at prosjektet i utgangspunktet kun skulle teste den samme HA-løsningen på to ulike versjoner av RedHat Linux, dvs ikke teste to ulike HA-løsninger slik som opprinnelig planlagt. Denne beslutningen bidro til å redusere forsinkelsen noe siden vi slapp å bygge opp kompetanse på to ulike HA-løsninger.

Prosjektet endte opp med følgende endringer i prosjektplanen:

- Milepæl 7 er i ny plan estimert til å bli ferdig 22.02, opprinnelig skulle den være ferdig 7.1.
- Sluttrapporten vil ikke bli levert før ca 10.03, opprinnelig skulle dette skje 27.01.

Det ville vært uheldig om valget av fremtidig HA-løsning for USIT ble forsinket dersom nye tjenester med behov for høytilgjengelighet brått måtte settes opp. Det er imidlertid ikke kjent for prosjektet at dette var en reell problemstilling. Vi tror derfor ikke at vår forsinkelse innebar noen målbar risiko for USIT eller våre kunder.

For mer detaljer om prosjektplanen, se følgende Appendix:

- [Appendix 4: Mandat](#)
- [Appendix 5: Milepæler](#)
- [Appendix 6: Prosjektplan](#)

1.7 Risiko

Tidlig i prosjektet ble det gjort en risikoanalyse.

"Mangelfull testing utført av eksterne ressurser ifbm test av tjenester på HA-løsningen" ble antatt å være det desidert største risikoen. Vi planla å redusere denne risikoen ved å utarbeide detaljerte testplaner og testskjema. Vi ønsket å involvere de aktuelle miljøene tidlig slik at det ble enklere for dem å sette av ressurser til testingen samtidig som de ville få bedre forståelse av hva vi ønsket å få testet.

Andre risiki vi vurderte som betydelige dreide seg om komplikasjoner rundt anskaffelse av maskin- og programvare til testingen. Vi forsøkte derfor å legge inn tilstrekkelig slakk i prosjektplanen i denne fasen slik at tidsplanen skulle holde.

I ettertid ser vi at vi unngikk den største risikoen fordi vi valgte å gjøre testene med tjenester som vi hadde kompetanse på innad i prosjektet. Vi ser også at vi ikke la inn nok slakk i prosjektplanen til oppsett av testmiljøet, i desember måtte vi som tidligere nevnt replanlegge milepæl 7 og 8 nettopp pga forsinkelser på dette området.

Vi fikk også forsinkelser fordi det tok lang tid å få leid inn konsulenthjelp, og dette var en risiko vi ikke forutså i risikoanalysen.

Se [Appendix 3](#) for mer detaljer om risikoanalysen.

1.8 Timeregnskap

I revidert prosjektplan var prosjektet estimert til 817 timer. Dette estimatet er for prosjektarbeid og er ikke inkludert prosjektadministrasjon.

Aktuelt timeforbruk har vært

Oppgave	Timer
Prosjektarbeid	770,85

Prosjektledelse	195,75
-----------------	--------

I dette regnskapet mangler ca 10 timer fra en prosjektdeltaker.

Vi ser at vi har brukt ca 26 timer mindre enn estimert på prosjektarbeid.

Frem til testingen av HA-løsninger startet i januar, lå timeforbruket ca 180 timer under det antall timer prosjektet hadde produsert. Vi har mao. brukt langt flere timer enn estimert i siste del av prosjektet da vi

- satte opp testmiljøene
- gjennomførte tester
- dokumenterte resultatene fra testene
- skrev sluttrapport.

Hovedsaklig var det under oppsett av testmiljø at estimatene sprakk.

Timeforbruket for prosjektadministrasjon er som forventet for å drive et prosjekt av denne størrelsen over 6 måneder.

2 Prosjektarbeid

HALv2-prosjektet ble kjørt iht USITs prosjektstandard og dokumentert på nettstedet

<https://www.usit.uio.no/prosjekter/halv2>

I august ble det avholdt et mini-seminar ("kickoff") på 2 ganger 4 timer hvor prosjektleder og de som senere ble medlemmer i prosjektet gjennomførte

- interessentanalyse
- aktivitetsplanlegging

Minst like viktig var det at disse to dagene markerte en oppstart av prosjektet, og bidro til å bygge en følelse av prosjektets "eksistens" blant deltakerne. Rent praktisk leide vi rom i Forskningsparken og USIT spanderte lunch på deltakerne - et populært tiltak!

USITs prosjektstandard har ikke hatt noen mal for interessentanalyse. HALv2-prosjektet utviklet derfor en modell løst basert på NAVs maler for interessentanalyse (se [Appendix A1 for vår mal](#) for interessentanalyse og [Appendix A2 for forklaring av metoden](#)).

Vi anbefaler at HALv2 sin mal og metode tas inn i USITs prosjektstandard.

Milepæler og aktiviteter ble ført inn i ansvarskart. Prosjektet var nøye med å holde dokumentasjonen oppdatert underveis i prosjektet, og vi gjorde oss selv avhengige av å holde planverket oppdatert ved å linke inn alle leveranser fra aktivitetsplanene.

Det ble lagt stor vekt på at alle aktiviteter skulle lede til en skriftlig leveranse og at det skulle skrives referater fra alle møter. Dette ble gjort for å gjøre prosessen mer etterrettelig, men også for å unngå "omkamper". Prosjektleder la også stor vekt på å gjøre det enkelt for prosjektdeltakerne å få oversikt over hva som ble forventet produsert av resultater til enhver tid; rent konkret ble det utarbeidet ukentlige lister med oppgaver som var forventet utført neste 7 dager. Dette fungerte godt.

3 Samarbeid

3.1 Samarbeid innad i prosjektet

Samarbeidet innad i prosjektet har fungert meget bra, og deltakerne i prosjektet har klart å prioritere prosjektarbeidet i en travel hverdag. Stemningen i prosjektet har også vært meget god. Det viste seg at prosjektdeltakerne tilsammen hadde den riktige blandingen av tung teknisk kompetanse og kunnskap om tjenester på USIT.

Tre uker i prosjektet må fremheves fordi det oppsto en dynamikk som kom overraskende på oss.

Mens vi utførte tester satt vi fire stykker sammen i et møterom. Prosjektleder sin bærbare PC var koblet til projektor hvor vi viste status-informasjon fra test-miljøet. De andre satt med egne bærbare PCer som var koblet til test-clusteret og andre verktøy. Vi gjennomførte testplanen og kunne sammen via projektoren få oversikt over hvilke konsekvenser testene fikk for clusteret vårt.

Denne måten å jobbe på viste seg å være svært effektiv. Når det oppsto problemer kom det alltid et innspill fra en i rommet som raskt ble diskutert og ev. prøvd ut. Litt ulik bakgrunn og erfaring gjorde at forskjellige personer bidro med løsninger på ulike problemer. Vi kunne også raskt ta beslutninger og komme videre.

Ved å jobbe på denne måten brukte vi bare noen dager på å gjennomføre krevende tester. Vi hadde på forhånd avtalt å begrense oss til økter på maksimalt 4 timer hver dag, og avsluttet konsekvent når tiden var ute. Vi jobbet neste dag på samme måte, og slik holdt vi på til testene var ferdige.

De som var involvert opplevde denne testingen som intenst og egentlig ganske slitsomt, men det var samtidig gøy, lærerikt og sosialt.

I SAPP jobber man ofte individuelt med arbeidsoppgaver, og vi tror at endel oppgaver kan løses langt raskere og gi leveranser av høyere kvalitet dersom man for konkrete og krevende oppgaver setter ned små lag av folk med god kjemi og riktig kompetanse. Slike lag må da jobbe med høy intensitet over få dager slik at man utnytter "momentet" som oppstår av laginnsatsen.

3.2 Samarbeid med USIT

Samarbeidet med resten av USIT fungerte bra.

Gruppelederne for DBA og BSD var enkle å samarbeide med rundt ressurstilgang. Vi opplevde et godt samarbeid med Nett-drift som måtte gjøre endringer i nett-oppsettet for at vi skulle tilfredsstille de krav som clusterprogramvaren stilte. De som jobber med PostgreSQL bidro med ressurser til riktig tid selv på kort varsel.

Prosjektet samarbeidet også med andre grupper og ressurspersoner på USIT for å lære mer om tjenester som har eller kan ha behov for høytilgjenglighet. Dette samarbeidet opplevde vi fungerte godt, og vi fikk veldig nyttige innspill.

3.3 Samarbeid med eksterne

Da vi skulle danne oss et bilde av aktuelle kandidatteknologier, tok vi kontakt med noen av de andre universitetene (UiB, NTNU, UiT og UiA). Vi fikk tilbakemeldinger, men vi opplevde at de ikke var i samme situasjon som oss og dermed ikke hadde konkrete anbefalinger som var relevante. På den annen side opplevde vi genuin interesse, og flere ytret ønsket om å bli informert om konklusjonene fra prosjektet.

Vi postet også en epost på sage-members-listen ([SAGE](#) sin epostliste hvor medlemmer diskuterer alt innen systemadministrasjon) uten å få noen svar med nyttig informasjon.

4 Leveranser

4.1 Interessentanalyse

Prosjekteier ønsket at det tidlig i prosjektet skulle gjøres en interessentanalyse, og dette gjennomførte vi i Milepæl 3.

Det opplevdes som nyttig å gjennomføre interessentanalysen. Vi fikk bekreftet at de viktigste interessentene allerede var involvert i prosjektet gjennom eierskap til prosjektet eller ved deltagelse, og ble oppmerksom på en interessent vi ikke hadde tenkt på i utgangspunktet som det var viktig å involvere i deler av prosjektgjennomføringen. Vi laget også en kommunikasjonsplan med bakgrunn av interessentanalysen.

Innholdet i interessentanalysen er ikke tatt med i sluttrapporten fordi den kun er ment som et internt verktøy for prosjektet.

4.2 Oversikt over tjenester

Med bakgrunn i interessentanalysen gikk vi i Milepæl 4 videre og tok fatt på oppgaven med å få en oversikt over hvilke tjenester vi skulle ta hensyn til når vi skulle lage kravspesifikasjon for en fremtidig høytilgjengelighetsløsning (HA-løsning). Vi kategoriserte tjenestene som

- tjenester som i dag kjører i HA-miljø
- tjenester som i dag ikke kjører i HA-miljø men som har behov for å kjøre i et HA-miljø nå eller i nær fremtid
- fremtidige tjenester som kan ha behov for HA-miljø

Vi startet med tjenester som i dag kjører i et HA-miljø på ulike plattformer og innledet dialog med ulike miljøer for å identifisere flere tjenester.

For en liste over de tjenester som vi valgte å ta med, se

- [Appendix 7 Tjenester vi vil ta hensyn til i kravspesifikasjonen](#)

4.3 HA-løsninger i markedet og USITs kravspesifikasjon til slike løsninger

4.3.1 Trender og produkter innen høytilgjengelighet

Vi startet Milepæl 5 med å gå ut bredt for å få en oversikt over trender i markedet rundt høytilgjengelighet, og for å få en oversikt over aktuelle produkter. Vi dokumenterte resultatene i en tabell, men det opplevdes som vanskelig for prosjektdeltakerne å få prioritert denne type undersøkende virksomhet. Flere i prosjektet leverte få timer i denne fasen, og vi endte nok opp med en produktliste mer enn en analyse av trender i markedet.

Vi ble overrasket over hvor mange løsninger som eksisterte, men også over at det var få som var aktuelle for oss.

Det er to tydelige trender i markedet

- applikasjoner med innebygget HA-funksjonalitet
- virtualisering

4.3.1.1 Applikasjoner med innebygget HA-funksjonalitet

Flere applikasjoner som kjører på dagens ServiceGuard-løsning på USIT har fått innebygget HA-funksjonalitet, dette gjelder f.eks

- Databaser
 - Oracle RAC
 - PostgreSQL
 - MySQL
- IMAP-server
 - Cyrus ("[Murder](#)")

- Applikasjonstjenere
 - Jboss
 - Resin

Dette kan redusere behovet for en sentralisert HA-løsning som ServiceGuard. Når HA-funksjonaliteten legges i applikasjoner på denne måten, er det mulig å "skreddersy" implementasjonen slik at funksjonaliteten for brukerne av applikasjonene blir bedre enn med en sentralisert HA-løsning. Eksempler på en forbedret funksjonalitet er sterkt reduserte failover-tider eller mulighet for skalering over flere noder.

Samtidig øker kompleksiteten i applikasjonene når de også inneholder HA-funksjonalitet, og likeledes øker kravene til kompetanse i organisasjonen. For en organisasjon som USIT hvor vi drifter mange applikasjoner, kan stordriftsfordelene av en sentralisert HA-løsninger overstige gevinsten ved å ta i bruk egen HA-funksjonalitet innebygget i applikasjonene.

Miljøene rundt drift av MySQL og PostgreSQL har gitt uttrykk for at de ønsker en sentralisert HA-løsning og ikke benytte seg av HA-funksjonaliteten i applikasjonen, mens miljøene rundt Jboss, Resin og Cyrus har ønsket å se på HA-funksjonalitet i selve applikasjonen.

4.3.1.2 Virtualisering

Det er liten tvil om at "virtualisering" er kommet inn i produktivitetsfasen på Gartners såkalte "Hype kurve". USIT har valgt å benytte VMware som virtualiseringsløsning, og denne teknologien er i bruk på en rekke områder på USIT.

Samtidig er det stor skepsis mot virtualisering i noen miljøer, og spesielt blant de som drifter databaser. Det hevdes at virtualisering og databaser ikke nødvendigvis er en god kombinasjon, man kan f.eks ikke benyttes seg av vanlige "overbooknings-teknikker" fordi hele databaseinstansen gjerne er i aktiv bruk samtidig. Databaser er også avhengig av kjapp og forutsigbar I/O-ytelse, noe som har vært en utfordring på virtuelle maskiner. Dette har imidlertid blitt bedre i senere versjoner av VMware og maskinvare (både servere og lagringsløsninger), men vi har i regi av prosjektet ikke fått testet dette.

I tillegg til at VMware ikke er støttet som plattform for Oracle-databaser, er det blant de som drifter databaser lite ønske om å kjøre MySQL eller PostgreSQL på VMware.

For andre typer applikasjoner ser vi at virtualisering med fordel kan benyttes i stedet for en tradisjonell HA-løsning. Et eksempel er UiO's egenutviklede system for brukeradministrasjon, *Cerebrum*, som sannsynligvis ville fungere minst like godt på en virtuell maskin som i et tradisjonelt failover-cluster.

4.3.2 Oversikt over HA-løsninger

Vi laget en oversikt over tilgjengelige HA-løsninger og samlet inn informasjon om disse. I tillegg hadde vi et møte med selskapet [Redpill Linpro](#) som gav nyttig informasjon om løsningene de hadde erfaring med. Løsningene vi fant inkluderte alt fra hardware-basert clustering vha. blades til *opensource*-rammeverk for å bygge HA-løsninger. Flere av løsningene var tradisjonelle failover-løsninger på samme måte som ServiceGuard. Videre fant vi flere løsninger for replikering av data mellom noder.

For en komplett liste over de løsningene vi fant i denne fasen av prosjektet, se

- [Appendix 8 Oversikt over løsninger og konsepter for HA](#)

4.3.3 Kravspesifikasjon

Neste skritt var å lage en kravspesifikasjon for HA-løsningen. Først definerte vi generelle krav til en HA-løsning. Vi baserte disse kravene på egen erfaring og kunnskap om eksisterende HA-løsninger.

Deretter gikk vi gjennom alle tjenestene som i dag kjører på ServiceGuard og definerte hvilke krav disse stiller til HA-løsningen. Det samme gjorde vi med tjenester som kjører på Oracle RAC.

Til slutt gikk vi gjennom kravene fra de tjenestene vi i forrige milepæl bestemte at vi skulle ta hensyn til. Etter nærmere analyse viste det seg at alle kravene fra disse applikasjonene var kommet med i allerede formulerte krav. Unntaket var den fremtidige søketjenesten hvor det var uaktuelt å vurdere eller benytte applikasjoner uten egen HA-funksjonalitet, så denne tjenesten var det heller ikke nødvendig å jobbe videre med.

Hele kravspesifikasjonen er gjengitt her:

- [Appendix 9 Kravspesifikasjon til HA-løsning](#)

4.4 Avstemming mot kravspesifikasjon

I Milepæl 6 skulle vi avstemme HA-løsninger mot kravspesifikasjonen.

For en fullstendig oversikt over alle HA-løsningene vi vurderte, se

- [Appendix 10 Liste over HA-teknologier](#)

Målet var å komme ned i 6 HA-løsninger som vi skulle avstemme mot kravspesifikasjonen. Vi jobbet i flere omganger og etterhvert kunne vi eliminere flere løsninger slik at vi sto igjen med disse:

- Oracle DataGuard
- Oracle Grid (som er cluster løsningen Oracle RAC bygger på)
- RedHat Cluster Suite (RHCS)
- SteelEye LifeKeeper
- Veritas Cluster Server (VCS)
- VMware HA

Vi fordelte arbeidet blant prosjektdeltakerne som avstemte funksjonaliteten i løsningen mot kravene. I denne forbindelse ble det avholdt webinar med selskapet *Nordic Mind* i Danmark (kompetanse- og support-senter for SteelEye Lifekeeper i Skandinavia, Baltikum og [CIS](#)) som var svært behjelpelig med å svare på spørsmål rundt SteelEye Lifekeeper. Like god nytte hadde vi av et webinar med eksperter fra *RedHat* om RHCS. Vi fikk også veldig god og grundig informasjon om VCS fra Kåre Juvkam hos *Symantec*.

Siden HA-løsningene

- Oracle DataGuard
- Oracle Grid
- VMware HA

allerede er tatt bruk av USIT, var det ikke nødvendig å teste disse. Disse er heller ikke generelle HA-løsninger. Derfor vurderte vi kun følgende løsninger mot kravspesifikasjonen med henblikk på testing:

- RHCS
- SteelEye Lifekeeper
- VCS

4.4.1 Avstemming av krav (dvs ikke ønsker)

Alle disse tre generelle HA-løsningene tilfredsstilte våre generelle krav til HA-løsninger, men med følgende merknader:

- Noen av USITs tjenester som i dag kjører på Oracle RAC stiller spesielle krav til HA-løsningen. Disse kravene kunne ingen av disse tre løsningene tilfredsstille. Det betyr at Oracle RAC ikke uten videre kan erstattes av disse tre generelle HA-løsningene.
- SAP stiller endel egne krav til HA-løsninger, men alle de tre løsningene støtter SAP. Det er også støttet å

kjøre SAP på Oracle RAC.

Ut fra kravene i kravspesifikasjonen var det ingen av de tre løsningene som ble diskvalifisert.

4.4.2 Avstemming av *ønsker*

Kravspesifikasjonen inneholder også *ønsker*. Her var det litt mer variasjon i hvordan disse ble tilfredsstilt, og videre vurdering ble påvirket av hvordan vi *vektla* de ulike ønskene.

Følgende *ønsker* ble tillagt spesiell vekt av prosjektet

- [SUP06](#) : Færrest mulig leverandører (kun RHCS innfrir denne)
- [SUP12](#): HA-løsningen er open source (kun RCHS innfrir denne)

Følgende *ønske* er tillagt spesiell vekt av prosjekteier.

- [LIC01](#): Kostnaden for løsningen ved anskaffelse og drift bør ikke være vesentlig høyere enn for dagens SG-løsninger, også for eksterne kunder som ikke er omfattet av UiO sine lisenser med reduserte priser for akademiske institusjoner. (RHCS er vesentlig billigere enn de andre HA-løsningene)

Hovedkonklusjon etter mapping av produkter mot våre *krav* og *ønsker*:

- Alle tre HA-løsninger innfrir våre **krav**, og det er ikke mulig å kunne si at den ene løsningen bedre enn den andre før vi har testet dem.
- Når det gjelder **ønsker**, har RHCS et fortrinn spesielt fordi prosjekteier legger vekt på pris.
- Ingen av løsningene har noe fortrinn når det gjelder å kjøre SAP.
- Løsninger som i dag er avhengige av RAC-funksjonalitet, må fortsatt kjøre på RAC.

4.4.3 Tilleggstester

I tillegg gikk prosjektet inn for å gjøre følgende tester hvis tiden strakk til:

- Teste failover mellom primær og standby database med Dataguard. Testene gjøres på eksisterende maskiner
- Gjøre enkle ytelsetester slik at vi evt kan sammenligne med VMware HA på de samme maskinene etterpå.

For detaljer rundt avstemming av HA-løsningene mot våre krav, se

- [Appendix 11 Mapping av HA-produkter mot kravspesifikasjon](#)
- [Appendix 12 Mapping av Oracle dataguard mot kravspesifikasjon - også active dataguard opsjonen](#)
- [Appendix 13 Mapping av Oracle Clusterware \(Oracle Grid\) mot kravspesifikasjon](#)
- [Appendix 14 Mapping av RHCS mot Kravspesifikasjon](#)
- [Appendix 15 Mapping av SteelEye LifeKeeper mot kravspesifikasjon](#)
- [Appendix 16 Mapping av Veritas Cluster Server \(VCS\) mot kravspesifikasjon](#)
- [Appendix 17 Mapping av VMwares HA-løsning mot Kravspesifikasjon](#)
- [Appendix 23 Dokumentere valg av HA-løsning for test](#)

4.5 Utvalg av løsninger for test

Prosjektet hentet inn pris på de forskjellige løsningene siden kravet [LIC01](#) definerer et *ønske* om at "*prisen for den nye løsningen ikke skal være vesentlig høyere enn dagens SG-løsninger*"

Etter dialog med leverandørene kom vi frem til "budsjettpriser" som visste at RHCS var vesentlig billigere enn VCS og SteelEye Lifekeeper.

Prosjekteier bestemte derfor følgende:

- RHCS skal testes.
- Kun dersom RHCS ikke fungerer godt nok, skal vi vurdere om VCS eller LifeKeeper skal testes.

I HALv2-prosjektet hadde vi opprinnelig planlagt å teste flere ulike HA-løsninger. Deretter skulle vi sammenligne resultatene og tilslutt velge den beste. Når det så ble bestemt at kun RHCS skulle testes for å se om det fungerte godt nok, innebar det kun små endringer i prosjektplanverket, mens det for prosjektdeltakerne føltes som en reell kursdreining.

4.6 Gjennomføring av tester

Følgende tester ble gjennomført:

- Vi testet RedHat Cluster Suite (RHCS) mot kravspesifikasjon
- Vi tester failover mellom primær- og standby-database for Oracle DataGuard

For mer informasjon om DataGuard, RHCS, bakgrunn for testene og en oppsummering av testresultatene, se

- [Appendix 22: Oppsummering av tester](#)

4.6.1 DataGuard

Testene viser at DataGuard *kan* brukes i en HA-løsning hvor en klient uten rekonfigurasjon kan koble til enten primær eller standby-database alt etter hvilken som er åpen for spørringer.

Dersom noden som kjører primær-databasen er nede, får man "TCP-heng" på ca 1 minutt ved påkobling fra klienten. For mange tjenester er et slikt heng på alle forespørsler mot databasen uakseptabelt. DataGuard har ingen egne løsninger for å håndtere dette, men man kan selvsagt selv lage en løsning hvor man setter opp ip-nummeret på en annen node.

Vi gjorde en *manuell* switch fra primær- til standby-database. Oracle har egne løsninger for hvordan dette kan automatiseres (Oracle Data Guard Fast-Start Failover) men dette er ikke noe vi har testet på USIT. Vi er mao. avhengig av manuelle operasjoner for å gjøre en failover uten slike ekstra-produkter.

4.6.2 RHCS på RHEL5

RHCS på RHEL5 fungerte bra, og vi kunne kjøre både PostgreSQL, Oracle og vår egenutviklede tjener ("*Halv2-tjener*") på løsningen. Vi hadde noen få tilfeller hvor vi fikk inkonsistent oppførsel i clusteret, og vi løste dette ved å reboote maskinen som hadde problemer. Vi kom aldri til bunns i hva som forårsaket inkonsistensen.

Vi noterte oss noen avvik fra forventede resultater, men ingen av avvikene er kritiske.

4.6.3 RHCS på RHEL6

RHCS på RHEL6 fungerte bra, og vi kunne kjøre både PostgreSQL, Oracle og vår egenutviklede tjener ("*Halv2-tjener*") på løsningen. Logging var merkbart bedre enn på RHEL5, og funksjonaliteten har utviklet seg og virker på flere områder mer gjennomarbeidet.

Vi noterte oss noen avvik fra forventede resultater, men ingen av avvikene er kritiske.

4.6.4 Nagios

Vi gjorde noen innledende tester av en tilsynelatende lovende Nagios-plugin for RHCS. Selv under tester hvor

RHCS åpenbart var helt eller delvis nede, gav Nagios feilaktig positiv status for tilstanden i RHCS. Kort fortalt fungerte denne plugin svært dårlig.

Prosjektet har konkludert med at vi må skrive vår egen Nagios-plugin for RHCS, og vi mener at dette skal kunne gjøres på ca 8 timer siden informasjonen fra MIB'en ikke er avskrekkende å forholde seg til.

4.7 Oppsummering fra testene

Ut fra testene trakk vi følgende konklusjoner:

- RedHat Cluster Suite (RHCS) på RHEL5 og RHEL6 tilfredsstiller våre krav godt nok i henhold til vår testplan.
- Oracle DataGuard kan også brukes til å lage failover-løsninger for Oracle-databaser, men det er ikke trivielt å lage løsninger hvor klientene ikke vil oppleve tregere responstid ved driftsforstyrrelser.
- Dersom RHCS tas i bruk, bør USIT utvikle en egen Nagios-plugin for RHCS.

Den offisielle dokumentasjonen for RHCS på både RHEL5 og RHEL6 var tildels mangelfull, mens vi på mindre offisielle forum og wiki'er ofte fant veldig nyttig informasjon.

4.8 Samarbeid med leverandøren

Vi leide inn konsulenthjelp fra leverandøren RedHat for å

- kvalitetssikre arbeidet med oppsett av test-clusteret
- få svar på spørsmål underveis i testingen.

Mens vi satte opp test-clusteret, sendte vi inn spørsmål til konsulenten og hadde dialog både på epost og i telefonkonferanser. Dette fungerte veldig bra, og vi opplevde dette som nyttig fordi vi hadde mange spørsmål mens vi satte opp RHCS for første gang.

Før vi startet med gjennomkjøring av testplanen, kom konsulenten hit til USIT. Han gjennomgikk hele oppsettet vårt for å sikre at det etterlevde RedHat sin "*best practice*" for RHCS. Vi opplevde det som betryggende å gjøre denne gjennomgangen. Vi gjennomførte også innledende tester med konsulenten tilstede, og mye tid ble spart fordi han bidro til at vi sammen raskt klarte å løse problemer.

Etter at konsulenten dro, fortsatte vi testene. Men nå opplevde vi ikke å få raske og gode svar; uvisst av hvilken grunn.

Vi har i løpet av de siste månedene vært i kontakt med flere tekniske ressurser i RedHat, og vi har fått et godt inntrykk av det faglige nivået og støtten de kan tilby.

RedHat var også behjelpelig med å skaffe test-lisenser for RHCS.

5 Konklusjon

RedHat Cluster Suite (RHCS) tilfredstilte alle *krav* i kravspesifikasjonen, og var også den HA-løsningen som best tilfredstilte de *ønskene* vi la spesielt vekt på. Vi gjennomførte testplanen vår for RHCS på RedHat Enterprise Linux (RHEL) 5 og RHEL6, og begge versjonene passerte testene.

Vi fikk noen avvik fra forventede resultater, men ingen av disse avvikene var kritiske.

RHCS på RHEL6 virker mer gjennomarbeidet enn RHCS på RHEL5, og deler av funksjonaliteten er bedre.

RedHat gir inntrykk av at de satser på produktet RHCS, og de har skissert flere forbedringer som er forventet i fremtidige versjoner.

Under testingen av RHCS, fikk vi et godt inntrykk av det faglige nivået og støtten som RedHat kunne tilby oss.

Ut fra dette konkluderer vi:

Prosjektet anbefaler RHCS på RHEL6 som fremtidig HA-løsning for USIT.

USIT kjører allerede HA-løsninger som *Oracle RAC* og *VMware HA*, men vi ønsket også å teste hvordan *Oracle DataGuard* egner seg som HA-løsning. Siden DataGuard ikke har noen innebygget failover mellom primær- og standby-database og heller ikke har noen mekanisme for å unngå TCP-heng når en node er nede, synes vi ikke Oracle DataGuard fungerer som en erstatning for en generell HA-løsning som f.eks ServiceGuard.

Vi konkluderer dermed:

Prosjektet mener at Oracle DataGuard kan brukes som disaster recovery-løsning for Oracle databaser, men ikke som en generell HA-løsning for Oracle. Da bør heller RHCS eller RAC benyttes.

6 Anbefalinger

USIT kjører tjenester på en rekke forskjellige typer HA-løsninger. Hver type HA-løsning vi drifter innebærer en kostnad i form av lisensutgifter og/eller kompetanse. Derfor bør USIT

- fase ut HA-løsninger som er lite i bruk dersom det er hensiktsmessig
- bruke den "billigste" tilgjengelige løsningen som likevel tilfredsstillende krav til oppetid som stilles

Det kan være komplekst å prissette hva som er en "billig" løsning, men vi tenker oss følgende tabell hvor kompetanse er vektlagt mer enn lisenskostnader

Kostnad	HA-løsning	Kommentar
0	Ingen HA	Enkeltstående maskin
1	Lastbalansering	USIT får snart en ny lastbalanseringsløsning
2	Virtualisering	F.eks VMware HA
3	Replikeringsløsning (manuell failover)	F.eks Oracle DataGuard • Ikke automatisk "failover" og gir mer nedetid, men krever ikke delt lagring.
4	RHCS	RHCS på RHEL6 • Automatisk failover av tjenester • Kort nedetid ved failover (typisk et par minutter for en middels stor Oracle-database)
5	HA-funksjonalitet innebygget i applikasjonen	F.eks "Oracle RAC" eller "Murder" for IMAP • Tilbyr best funksjonalitet men koster mer i ◦ kompetanse

- | |
|--|
| lisenskostnader (gjelder kun kommersielle produkter)
• Mulig å sette opp løsninger slik at man har neglisjerbar nedetid ved "failover". |
|--|

Eksempel:

Når man skal legge ned dagens ServiceGuard-clustre må man vurdere hver tjeneste ut fra tabellen ovenfor. En gitt tjeneste antas å inneholde en Oracle-database med historiske data og vi antar videre at det er akseptabelt for kundene å ha en nedetid på opp til en halv arbeidsdag. Tjenesten behøver derfor ikke RHCS eller RAC og kan klare seg med en DataGuard-installasjon. Av supportmessige årsaker kan vi ikke benytte virtualisering, og lastbalansering passer ikke for databaser. Det er ikke et alternativ å kjøre helt uten HA fordi databasen må være oppe etter en halv arbeidsdag. Derfor bør man i utgangspunktet velge DataGuard for denne databasen, eller dokumentere hvorfor det er hensiktsmessig å velge en dyrere løsning.

Vi anbefaler at USIT lager en prosedyre for valg av riktig HA-løsning for tjenester. I prosedyren må man veie krav til oppetid og sikkerhet opp mot kostnaden ved ulike HA-løsninger.

7 Veien videre

Dersom USIT velger å implementere RHCS som erstatning for ServiceGuard, anbefaler vi at USIT etablerer et prosjekt for å sette opp den første tjenesten på RHCS.

Dette prosjektet bør startes etter at prosjektet for valg av fremtidig lagringsløsning ("[VAFFEL-prosjektet](#)") har konkludert. På den måten kan USIT implementere RHCS på neste generasjons lagringsløsning. I tillegg til å flytte en tjeneste til RHCS, er det naturlig at prosjektet lager driftsdokumentasjon for RHCS slik at påfølgende installasjoner av RHCS kan gjøres som linjeaktiviteter. Videre bør prosjektet gå opp grensene for myndighet og ansvar for de forskjellige komponentene av et RHCS-cluster.

Det bør også lages en Nagios-plugin for RHCS.

Deltakerne til dette fremtidige prosjektet bør inkludere deltakere fra både HALv2- og VAFFEL-prosjektet.

Vi anbefaler også at det dannes en løs gruppe (ikke en tjenestegruppe som vil være "overkill") for drift av RHCS. Her bør ressurspersoner for oppsett og drift av RHCS være med, og det bør finnes egne e-postlister og chatroom for gruppen. På denne måten blir det enklere å bygge opp et miljø rundt RHCS.

Det bør gjennomføres opplæring for de som skal drifte RHCS eller tjenester på RHCS. Vi mener en skreddersydd workshop kan være mer nyttig enn et generelt kurs, og anbefaler at en slik workshop utarbeides etter USITs ønsker av en kursleverandør som har erfaring med å kjøre kurs i RHCS. En leverandør som tilbyr aktuelle kurs er *Redpill Linpro*.

8 Appendix

- [Appendix 1 Mal Interessentanalyse](#)
- [Appendix 2 HALv2 Metode for interessentanalyse](#)
- [Appendix 3 Forenklet risikoanalyse for HALv2-prosjektet](#)
- [Appendix 4 Mandat](#)
- [Appendix 5 Milepæler](#)
- [Appendix 6 Prosjektplan](#)
- [Appendix 7 Tjenester vi vil ta hensyn til i kravspesifikasjonen](#)
- [Appendix 8 Oversikt over løsninger og konsepter for HA](#)
- [Appendix 9 Kravspesifikasjon til HA-løsning](#)
- [Appendix 10 Liste over HA-teknologier](#)

- [Appendix 11 Mapping av HA-produkter mot kravspesifikasjon](#)
- [Appendix 12 Mapping av Oracle dataguard mot kravspesifikasjon - også active dataguard opsjonen](#)
- [Appendix 13 Mapping av Oracle Clusterware \(Oracle Grid\) mot kravspesifikasjon](#)
- [Appendix 14 Mapping av RHCS mot Kravspesifikasjon](#)
- [Appendix 15 Mapping av SteelEye LifeKeeper mot kravspesifikasjon](#)
- [Appendix 16 Mapping av Veritas Cluster Server \(VCS\) mot kravspesifikasjon](#)
- [Appendix 17 Mapping av VMwares HA-løsning mot Kravspesifikasjon](#)
- [Appendix 18 Testplan for HA-løsning](#)
- [Appendix 19 Testing av Oracle Dataguard](#)
- [Appendix 20 Kravspesifikasjon for Nagios-plugin RHCS](#)
- [Appendix 21 Testresultater RHCS](#)
- [Appendix 22 Oppsummering av tester](#)
- [Appendix 23 Dokumentere valg av HA-løsning for test](#)
- [Appendix 24 Testresultater RHCS på RHEL6](#)
- [Appendix 25 Oppsett av pakken rhcsoracle - Noen notater](#)
- [Appendix 26 Infrastruktur](#)
- [Appendix 27 Konfigurasjon av OS før man lager RHCS services](#)
- [Appendix 28 Failoverdomains](#)
- [Appendix 29 Endelig utgave av cluster.conf på RHEL5](#)
- [Appendix 30 Oppsett av Oracle-service i RHCS på RHEL5](#)
- [Appendix 31 Oppsett av PostgreSQL-tjeneste i RHCS på RHEL5](#)
- [Appendix 32 Sette opp "HALv2-daemon" i RHCS på RHEL5](#)
- [Appendix 33 Notater rundt oppsett av RHCS på RHEL6](#)
- [Appendix 34 Oppsett av Oracle-service i RHCS på RHEL6](#)
- [Appendix 35 Oppsett av PostgreSQL-tjeneste i RHCS på RHEL6](#)

Appendix 4 Prosjektmandat for HALv2

(prosjektets navn kan bli endret når prosjektet etableres sammen med prosjektdeltagerene)

Mandat versjon: 1.0
Godkjent av USITs ledermøte: [2010-05-25](#)
Oppdragsgiver: Knut Borge
Prosjekteier ved USIT: Knut Borge
Prosjektleder: Gunnar Steira Mikkelsen

Andre aktører

<ingent>

Formål

Prosjektets formål er å komme med en anbefaling til teknologivalg slik at USIT kan tilby stabile, sikre og høytilgjengelige tjenester som kjører på industristandard maskinvare med Linux operativsystem.

Bakgrunnen til prosjektet er at en av dagens foretrukne teknologier, ServiceGuard for Linux, om noen år blir desupportert fra leverandøren.

Rapportering

Statusmøte hver 14. dag med prosjekteier

Resultatmål og leveranser

Prosjektets hovedleveranse er en anbefaling til teknologivalg slik at USIT kan tilby stabile, sikre og høytilgjengelige tjenester som kjører på industristandard maskinvare med Linux operativsystem.

Med utgangspunkt i en interessentanalyse skal det gjøres en avgrensning av hvilke typer tjenester prosjektet skal omfatte. For de valgte tjenestene, skal det lages en kravspesifikasjon til en høytilgjengelighetsløsning. Dersom det er hensiktsmessig, kan prosjektet stille ulike krav til høytilgjengelighetsløsningen basert på en gruppering av tjenestene.

Alternativene skal vurderes opp mot funksjonalitet, kompleksitet, innføringskostnader, driftskostnader, levetid og sikkerhet.

Prosjektet kan anbefale ulike høytilgjengelighetsløsninger basert på en gruppering av tjenestene. USITs lastbalanseringsløsning skal inngå som en av de løsningene som evalueres.

Rammer

Startdato: 7/9 2010
Sluttdato: 10/3 2011
Timeverk: 817 (prosjektadministrasjon ikke medregnet)
Direkte kostnader: <ikke estimert>
CT-koder: Halv2-Prosjektarbeid og Halv2-Administrasjon

Hovedmilpæler

M01 Når mandat er godkjent av USITs ledermøte.

M02 Når USITs ledermøte har godkjent oppstart av prosjektet

M03 Når interessentanalyse er godkjent av PE.

M04 Når omfanget av tjenester prosjektet skal omfatte er godkjent av PE.

M05 Når kravspesifikasjon er godkjent av PE.

M06 Når et utvalg kandidatteknologier er definert og godkjent av PE.

M07 Når testing av kandidatteknologier er godkjent av PE.

M08 Når anbefaling og sluttrapport er godkjent av PE.

*) PE er Prosjekteier

Appendix 5: Milepæler

Milepælsplan og ansvarskart

Versjon: 1.0

- **U** - Utfører arbeidet
 - **B** - Hovedbeslutning
 - **b** - Delbeslutning
 - **a** - Framdriftsansvar
 - **A** - Kompetanseoverføring
 - **R** - Må rådspørres
 - **I** - Må informeres
 - **r** - Kan tilkalles for diskusjon
- **PE** - Prosjekteier
 - **PL** - Prosjektleder
 - **PG** - Prosjektgruppe
 - **LM** - Ledermøte

Planlagt ferdig dato	Faktisk ferdig dato	Status	Avhengig av	Milepæl nr	Milpælbeskrivelse	PE	PL	PG	LM
	25.05.10	Ferdig		M01	Når mandat er godkjent av USITs ledermøte	b	Ua		B
07.09.10	07.09.10	Ferdig	M01	M02	Når USITs ledermøte har godkjent oppstart av prosjektet		Ua	U	B
02.09.10	02.09.10	Ferdig	M02	M03	Når interessentanalyse er godkjent av PE	B	Ua	U	
10.09.10	20.09.10	Ferdig	M03	M04	Når hvilke av tjenester prosjektet skal omfatte er godkjent av PE	B	Ua	U	
01.10.10	04.10.10	Ferdig	M04	M05	Når kravspesifikasjon er godkjent av PE	B	Ua	U	
01.11.10	03.11.10	Ferdig	M05	M06	Når et utvalg kandidatteknologier er definert og godkjent av PE	B	Ua	U	
22.02.11	04.02.11	Ferdig	M06	M07	Når testing av kandidatteknologier er godkjent av PE.	B	Ua	U	
17.03.11		Påbegynt	M07	M08	Når anbefaling og sluttrapport er godkjent av PE	B	Ua	U	

Appendix 6 Prosjektplan

Dette er en mal for prosjektplan på USIT. Se [forklaring til feltene i prosjektguiden](#).

Formål

Prosjektets formål er å komme med en anbefaling til teknologivalg slik at USIT kan tilby stabile, sikre og høytilgjengelige tjenester som kjører på industristandard maskinvare med Linux operativsystem.

Bakgrunnen til prosjektet er at en av dagens foretrukne teknologier, ServiceGuard for Linux, om noen år blir desupportert fra leverandøren.

Mandat

Se [egen side](#).

Ressurser

Ressursavtale ble utarbeidet i september. (ikke tatt med som Appendix i sluttrapporten)

Prosjektorganisasjon

Oppdragsgiver og prosjekteier er Knut Borge (SAPP) og prosjektleder er Gunnar Steira Mikkelsen (SAPP-DBA). Deltakerne i prosjektet er

<fylles ut når endelig avklart>

Leveranser

Prosjektets hovedleveranse er en anbefaling til teknologivalg slik at USIT kan tilby stabile, sikre og høytilgjengelige tjenester som kjører på industristandard maskinvare med Linux operativsystem.

Med utgangspunkt i en interessentanalyse skal det gjøres en avgrensning av hvilke typer tjenester prosjektet skal omfatte. For de valgte tjenestene, skal det lages en kravspesifikasjon til en høytilgjengelighetsløsning. Dersom det er hensiktsmessig, kan prosjektet stille ulike krav til høytilgjengelighetsløsningen basert på en gruppering av tjenestene.

Alternativene skal vurderes opp mot funksjonalitet, kompleksitet, innføringskostnader, driftskostnader, levetid og sikkerhet.

Prosjektet kan anbefale ulike høytilgjengelighetsløsninger basert på en gruppering av tjenestene. USITs lastbalanseringsløsning skal inngå som en av de løsningene som evalueres.

Aktiviteter

Aktivitetene i prosjektet finner du under [milepælene](#) i prosjektet.

Rapporteringsrutiner

Statusmøte hver 14. dag med prosjekteier etter at prosjektet er vedtatt startet opp. Mer sjelden rapportering nå i initieringsfasen.

Risikoanalyse

En [forenklet risikoanalyse](#) ble gjort i september og funnene ble tatt i betrakning da aktivitetsplanen ble utarbeidet.

Avvikshåndtering

Mindre avvik tas opp med prosjekteier på statusmøte slik at man kan bestemme hvordan avviket skal håndteres. Større avvik tas opp med prosjekteier utenom de fastlagte statusmøtene dersom det er lenge til neste møte.

Overlevering

Ingen spesiell overlevering er nødvendig for dette prosjektet.

Appendix 3 Forenklet risikoanalyse for HALv2-prosjektet

Trinn	1. Hvilke uønskede hendelser eller tilstander kan oppstå	2. Hva er sannsynligheten for at hendelsen skal oppstå? 1=Lite sannsynlig 2=Mindre sannsynlig 3=Sannsynlig 4=Meget sannsynlig 5=Svært sannsynlig	3. Hva er konsekvensen for prosjektet av hendelsen 1=ufarlig 2=farlig 3=kritisk 4=meget kritisk 5=katastrofalt	4. Hva blir risikoen regnet ut som produktet av sannsynlighet og konsekvens	5. Hvilke tiltak settes i verk?
-------	--	---	---	---	---------------------------------

Risikoanalyse				
Hendelse/Tilstand	Sannsynlighet (1-5) (S)	Konsekvens (1-5) (K)	Risiko (S*K)	Tiltak
Ingen HA-løsning tilfredstiller kravene i kravspesifikasjon	2	2	4	• Grundig forarbeid og analyse av HA-løsninger. • Hvis hendelsen inntreffer, må kravene i kravspesifikasjonen revurderes.
Prosjektdeltakere får ikke frigjort tid til prosjektarbeid	3	3	9	• Innarbeide slack i planen slik at prosjektet ikke blir for sårbart. • Ta kontakt med PE straks forsinkelser oppstår.
Får ikke tak i HW til testing tidnok	3	3	9	• Slack i prosjektplan slik at noen ukers leveringstid ikke får store konsekvenser. • Bestille HW så tidlig som mulig • Bestille fra en leverandør hvor vi har god erfarings med forutsibare leveringstider.
Får ikke tak i SW eller lisenser tidnok	2	3	6	• Slack i prosjektplan slik at noen ukers leveringstid ikke får store konsekvenser. • Bestille SW så tidlig som mulig

SW eller HW feil og ventetid på support fra ekstern leverandør	4	2	8	• Feilmelde tidlig dersom problemer oppstår • Bruke leverandøren aktivt
Sertifiseringsmatrise for HW og SW stack skaper forsinkelser pga tilgjengelighet av produkter	3	3	9	• Bygge inn litt slack i prosjektplanen i aktiviteter hvor HA-løsningen skal testes
Mangelfull testing utført av eksterne ressurser ifbm test av tjenester på HA-løsningen.	5 1)	4	20	• Involvere PE for å få motivert eksterne ressurser til å bruke tid på testing • Informere tidlig de miljøene som er aktuelle for testing slik at de har mulighet til å prioritere dette. • Utvikle og dokumentere detaljerte testskjema • Planlegge og dokumentere testene som skal gjøres grundig
Mangelfull testing utført av prosjektetgruppen	1	4	4	• Utvikle og dokumentere detaljerte testskjema • Planlegge og dokumentere testene som skal gjøres grundig
Mangelfull input fra interessenter ifbm krav til HA-løsning	2	2	4	Ingen spesielle 2)
Forsinkelser med å få beslutninger fra PE	2	4	8	• Avtale møter med Knut i god tid • Avtale stedfortreder ved fravær

1) Det er sannsynlig at eksterne ressurser ikke har mulighet til å sette av nok tid til testing, og sannsynlig at HALv2-prosjektet undervurderer informasjonsbehovet som er nødvendig for å teste løsningen grundig nok.

2) Prosjektets medlemmer har selv solide kunnskaper om clusterløsningene og hvordan de driftes i dag, så dette ventes å kompensere for denne usikkerheten.

Sammendrag

Matrisen viser at vi kun har en risiko som peker seg ut, det er

- Mangelfull testing utført av eksterne ressurser ifbm test av tjenester på HA-løsningen.

Det høye påslaget i "Sannsynlighet" er ikke begrunnet i en nedvurdering av andre miljø, men heller i at det erfaringsmessig fort kan oppstå ressursproblemer hos den eneste ressursen når testingen skal prioriteres.

Prosjektet må redusere denne risikoen ved å utarbeide meget detaljerte testplaner og testskjema, og forsikre seg om at kommunikasjonen med de som skal teste er god nok slik at de har forstått hva prosjektet forventer.

De andres store risikoene dreier seg om komplikasjoner rundt anskaffelse av HW og SW til testingen. Prosjektplanen må derfor legge inn tilstrekkelig slack i denne fasen slik at tidsplanen ikke sprekker.

Appendix 1 : Mal Interessentanalyse

Av: Gunnar S. Mikkelsen, USIT
Versjon 1.1

Trinn	1. Hvem stiller krav eller har forventninger til prosjektet	2. Beskriv hva Interessenten bidrar med?	3. Vurderer interessentens interesse for prosjektet (Høy/Middels/Lav)	4. Vurderer interessentens innflydelse for prosjektet (Høy/Middels/Lav)	5. Beskriv interessentens forventninger og ”suksesskriterier”	6. Beskriv hvordan interessenten skal håndteres
-------	---	--	---	---	---	---

For hjelp og hint til utfylling av skjemaet, se [egen dokumentasjon](#).

Interessentanalyse					
Interessent (I)	I's bidrag til prosjektet	I's interesse for prosjektet (Høy/Middels/Lav)	I's innflytelse på prosjektet/resultatet (Høy/Middels/Lav)	I's forventninger og suksess-kriterier	Tiltak

Appendix 2 : HALv2 Metode for interessentanalyse

Interessent: Aktør som påvirkes av prosjekt-gjennomføringen eller prosjekt-resultatet.

Eksempler på interessent

Her er noen eksempler på interessenter.

- Oppdragsgiver
- System/produkt-brukere
- Konsulenter/rådgivere
- Myndigheter
- Leverandører
- Samarbeidspartnere
- Interne Kunder
- Eksterne kunder (f.eks mhp lisenskostnader)
- Linjeledelse
- Ansatte i andre grupper/seksjoner
- Andre avdelinger som leverer oss tjenester
- Prosjektledelse og prosjektdeltagere

1. Sjekkliste for å identifisere interessentene

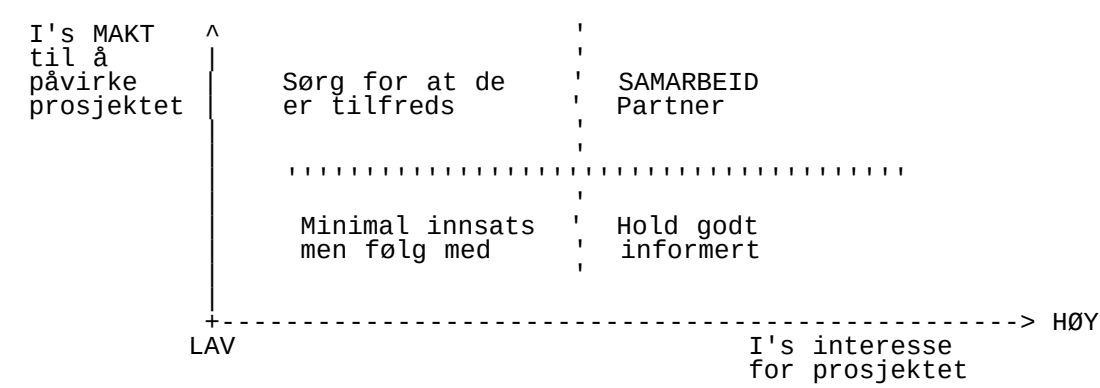
- Hvem er oppdragsgiver?
- Hvem skal bruke resultatet?
- Hvem skal akseptere resultatet?
- Hvem betaler for arbeidet og resultatet?
- Hvem skal akseptere at aktiviteten utføres og måten den utføres på?
- Hvem berøres (får utbytte, taper, får endrede forhold etc) av resultatet?
- Hvem leverer innsatsfaktorer, kunnskap, ressurser?

2. Gjøre seg kjent med interessentene (posisjonsanalyse)

- Hvilke krav, ønsker, behov og forventninger til prosjektet har interessenten?
- Innstilling til prosjektet?
- Hvilken posisjon og betydning har interessenten?
- Hvilke bidrag / motstand har de til prosjektet?
- Opplevd bidrags / belønningsbalanse?
- Hvilke syn har de på de andre interessentene?
- Hvilke sammenfallende og motstridende interesser ser vi?
- Hvilken adferd kan forventes, spesielt om de ikke får det som de vil?
- Ta stilling til hvordan den enkelte interessent skal håndteres

3. Velge strategier mot interessenter

HØY



- **Type A: Liten interesse, liten makt:** Strategien er å følge med (monitoring) i interessentens situasjon. Iverkesette tilbake bare hvis interessentens mulighet for å påvirke prosjektet endrer seg
- **Type B: Stor interesse, liten makt:** Strategien blir å holde interessenten godt informert. Følge med på om mulighetene for makt øker
- **Type C: Liten interesse, stor makt:** Strategien blir å holde interessenten vel tilfreds.
- **Type D: Stor interesse, stor makt:** Strategien er samarbeid. Det beste er å utvikle en eller annen form for partnerskap

Appendix 7 Tjenester vi vil ta hensyn til i kravspesifikasjonen

Author: Anders Odberg / Morten Werner Forsbring
Contact: halv2-core@usit.uio.no
Date: 2010-09-17

Innhold:

- [1 Tjenester vi vil ta hensyn til i kravspesifikasjonen](#)
 - [1.1 Løsninger som kjøres i et HA-miljø idag](#)
 - [1.2 Tjenester som idag ikke kjøres i et HA-miljø](#)
 - [1.3 Fremtidige tjenester med behov for å kjøre i et HA-miljø](#)

Dette er leveransedokument i aktivitet A04.05 : "Dokumentere de tjenester vi ønsker å se på videre". Dette blir de tjenestene vi vil ta hensyn til når vi utarbeider kravspesifikasjon i M05.

Forutsetning: Vi antar at USIT også i fremtiden vil tilby en form for lastbalansering som er like god eller bedre enn det som tilbys via LVS-hotellene våre idag. Vi har gått gjennom tjenestene som idag kjøres via LVS-hotellet og gjort en kjapp sjekk på om disse ønsker en annen type HA enn det LVS-hotellet kan gi de. De tjenestene som lever lykkelig på LVS-hotellet og hvor systemeier ikke har ytret ønske om en annen løsning tar vi ikke med videre når vi skal begynne å arbeide med en kravspesifikasjon (de vil ikke være nevnt i dette dokumentet).

1 Tjenester vi vil ta hensyn til i kravspesifikasjonen

Listen under baserer seg på resultatene fra aktivitetene [A04.01](#), [A04.02](#), [A04.03](#) og [A04.04](#) samt en avklaringssamtale med prosjekteier Knut Borge.

1.1 Løsninger som kjøres i et HA-miljø idag

- Alle løsningene vi idag kjører på ServiceGuard-cluster
 - Cerebrum/Egenutviklet
 - IMAP/Egenutviklet
 - MySQL
 - Oracle
 - PostgreSQL
 - SAP
- Alle løsningene vi idag kjører på Oracle RAC
 - Database for Atlas 3D Cern
 - Database for Oracle Infrastruktur OAS
 - Database for Oracle eBusiness Suite
 - Database for Vortex

1.2 Tjenester som idag ikke kjøres i et HA-miljø

- Filsynkronisering (her kan vi ta kontakt med Jasmina/Arne Meyer/Øystein for mer informasjon)
- JBoss- og Resin-baserte løsninger (her snakker vi med w3d-core@usit.uio.no)
- Nagios og snmp.uio.no (vurdere senere om vi skal/kan ta hensyn til slike overvåkingstjenester)
- Vortex (det kommer informasjon fra vortex-core@ til halv2-core@ etterhvert med mer informasjon)

1.3 Fremtidige tjenester med behov for å kjøre i et HA-miljø

Hvis vi ikke vet hvordan disse tjenestene fungerer, kan vi naturlig nok ikke ta mye hensyn til disse under utarbeidelse av kravspesifikasjonen.

- Bundlede applikasjoner (hovedsakelig på web-relaterte applikasjoner bundlet med applikasjonstjener (f.eks. servlet engine og webserver) som vanskelig lar seg integrere i GAPs allerede redundante driftsmiljø)
- Cristin (dette er Frida-NG, mulig vi bør ta kontakt med SUAF for å få mer detaljer om hvordan de ser for seg løsningen og evt. kunne påvirke de slik at de lager en løsning som lar seg HA-ifisere)
- Kerberos-server
- Nytt kalendersystem
- Søkjetjenesten skal fornyes / erstattes i 2011

Appendix 8 Oversikt over løsninger og konsepter for HA

Bare fiks på oppsettet om dere ikke liker det :)

Lagt inn av	Kategori	Produktnavn/beskrivelse Konsept	URL	Kommentarer (fjernet i sluttrapporten)
Gunnar		EMC Autostart	http://www.emc.com/products/detail/software/autostart.htm	
Ellef/Gunnar	Oracle	DBvisit	http://www.avisit.co.nz/	
Gunnar	Oracle	Generell oversikt over HA-løsninger fra Oracle for databaser	Oversikt	
Gunnar	Oracle	Oracle Cloud	Oversikt	
Gunnar	Oracle	Oracle Maximum Availability Architecture - MAA	MAA-oversikt	
Gunnar	Oracle	Oracle VM	Oversikt	
Ellef/Gunnar	Oracle	Oracle RAC OneNode	• Offisiell beskrivelse • Datasheet (PDF) • Technical Overview (PDF) • Blogg 1 2	
Gunnar	Oracle	Oracle DataGuard	Offisiell side	
Gunnar	Oracle	Oracle Active DataGuard	Offisiell side	
Ellef/Gunnar	Oracle	Oracle Clusterware	Offisiell side	
Ellef/Gunnar		Oracle Solaris Cluster	Offisiell side	
Ellef/Gunnar		RedHat Cluster Suite	• Overview • Migrasjon fra ServiceGuard til RCHS • www.redhat.com/docs/manuals/csgfs/	
Ellef/Gunnar		Linux HA (heartbeat)	http://www.linux-ha.org/ http://www.drbd.org/ Webcast drbd, mysql og litt om heartbeat	
Gunnar		Pacemaker	http://www.clusterlabs.org/	
Ellef/Gunnar		OpenAIS	http://openais.org	
		OpenSAFfire	http://goahead.com	
Ellef		Steeleye lifekeeper	http://www.steeleye.com/	
Rafael/Gunnar		Generelt om HA og Postgres	Dokument skrevet av Rafael	
		Veritas Cluster	http://www.symantec.com/business/cluster-server Migrere fra ServiceGuard til Veritas Cluster Server Oracle og VCS	
Gunnar		Service Availability Forum	offisiell website og wikipedia-artikkel	
Gunnar		VMWare HA	VmWare HA HA Deepdive - presentasjon	
Gunnar		IEEE TFCC-HA	Taskforce on Cluster Computing - HA	
Gunnar		Remus - HA via Async	Whitepaper	

		VM replication		
Gunnar	MySQL	Implementing High Availability in MySQL	Tutorial	
Ellef		Microsoft Hyper-V		
Ellef		Xen		
Gunnar	Postgres		different-replication-solutions.html	
Werner		Anycast	Anycast på Wikipedia	
JareleB		Egenera PAM	http://www.egenera.com/	

Appendix 9 Kravspesifikasjon til HA-løsning

Versjon 1.0 2010-10-01

Nedenfor er ulike krav og ønsker til HA-løsning for USIT listet opp.

Kode	Beskrivelse	Krav	Ønske
	Administrasjon, drift og driftsmiljø		
ADM01	Det må være støtte "rolling upgrade" av operativsystemet på nodene i HA-løsningen.	x	
ADM02	Nodene i HA-løsningen må kunne oppdateres fortløpende med systemoppgraderinger uten at vi mister support på HA-løsningen.		x
ADM03	Det er ønskelig at leverandøren av HA-løsningen raskt støtter nye Linux-kjerner. Forklaring: Vi ønsker ikke å havne i en situasjon hvor vi må velge om vi vil kjøre med en <i>supportert</i> Linux-kjerne eller en <i>nyere kjerne som har fikset viktige sikkerhetshull</i> .		x
ADM04	HA-løsningen må ikke være "for kompleks" å sette opp og å drifte		x
ADM05	Nye versjoner av HA-løsningen er mest mulig bakoverkompatibel slik at oppgraderinger ikke bryter eksisterende konfigurasjon		x
ADM06	Programvaren i HA-løsningen kan oppgraderes "rolling", dvs node for node uten at alle tjenester må tas ned	x	
ADM07	Det er mulighet til å skripte kommunikasjonen mot HA-systemet (ikke avhengig av GUI) for å kunne automatisere driftsmiljøet rundt løsningen.		x
ADM08	Tjenester i løsningen kan oppgraderes "rolling" slik at oppgraderingen ikke gir nedetid på løsningen. Eksempel: I RAC kan Oracle-installasjonen oppgraderes rolling på node for node mens tjenesten fortsatt er tilgjengelig på de andre nodene.	x 3)	x
	Applikasjonssupport		
APP01	Det er supportert å kjøre Oracle på HA-løsningen	x	
APP02	Det finnes ferdig Oracle-rammeverk tilpasset HA-løsningen		x
APP03	Postgres kan kjøre på løsningen	x	
APP04	Det finnes ferdig Postgres-rammeverk tilpasset HA-løsningen		x
APP05	Det er supportert/mulig (avhengig av hvilken MySQL løsning vi velger) å kjøre MySQL på løsningen	x	
APP06	Det finnes ferdig MySQL-rammeverk tilpasset HA-løsningen		x
APP07	Det er supportert å kjøre SAP på HA-løsningen.	x	
APP08	Det finnes ferdig SAP-rammeverk tilpasset HA-løsningen	x 1)	
APP09	IMAP kan kjøre på løsningen	x	

APP10	Det finnes ferdig IMAP-rammeverk tilpasset HA-løsningen		x
APP11	HA-løsningen har støtte for <i>fast application notification</i> (FAN) failover tilkoplingsteknologi	x 3)	x
	Forklaring: Vortex bruker dette for å redusere nedetid når en node eller databaseinstanse feiler		
	Failover og flytting av tjenester		
FAIL01	Ved flytting av en tjeneste fra en node til en annen i HA-løsningen er nedetiden for tjenesten omtrent det samme som tiden det tar å stoppe og starte ressursene i tjenesten på tilsvarende maskinvare utenfor HA-løsningen.	x	
FAIL02	Det må være mulig å flytte en tjeneste manuelt fra en node til en annen.	x	
FAIL03	Det er ønskelig at man gjøre en planlagt flytting til en annen node uten nedetid på tjenesten (live migrering).		x
FAIL04	Feil på maskinvare i HA-løsningen skal ikke gi langvarig nedetid for brukere av tjenesten	x	
FAIL05	Feil på maskinvare i HA-løsningen skal ikke gi nedetid for brukere av tjenesten		x
FAIL06	Feil på maskinvare i HA-løsningen skal ikke kreve manuelle inngrep for å gjøre tjenesten tilgjengelig for brukeren.	x	
FAIL07	Brukere av tjenesten skal ikke måtte endre i konfigurasjon på klienten etter at en tjeneste er flyttet til en annen node i HA-løsningen.	x	
	HW- og SW-support og leverandørforhold		
SUP01	HA-løsningen skal kunne kjøre på maskinvare fra flere leverandører Eksempel: Ikke bare på Dell-maskiner av en gitt type		x
SUP02	Nodene i HA-løsningen kan være virtuelle		x
SUP03	HA-løsningen skal kunne kjøre på ett eller flere operativsystemer som støttes av USIT i dag.	x	
SUP04	HA-løsningen bør kunne kjøre på på de nyeste versjonene av de operativsystemer som støttes av USIT i dag.		x
SUP05	HA løsningen må ha support fra leverandør(er) eller tredjepart(er) i alle ledd.	x	
SUP06	Det er ønskelig med færrest mulig forskjellige leverandører involvert i den totale løsningen		x
	Eksempel på uønsket situasjon: Oracle-SW, HP-lagring, RedHat-OS, Veritas-cluster, Dell servere		
SUP07	Løsningen må være supportert på de versjoner av kommersiell software som er aktuelt for USIT å kjøre	x	
	Eksempel: DBA-gruppen migrerer nå databaser til Oracle 11.2.0 som må være støttet.		
SUP08	USIT må ikke miste support på HA-løsningen dersom HA-løsningen er satt opp av USIT, eller er om konfigurasjonen er endret av USIT etter installasjon og initiell konfigurasjon.	x	
SUP09	Det finnes referansekunder for HA-løsningen som kjører i en driftssituasjon som kan sammenlignes med vår egen.	x	
SUP10	HA-løsningen skal inngå i et aktivt prosjekt eller være i aktiv utvikling hos en "solid og	x	

	levedyktig" leverandør.		
SUP11	Det finnes aktive epost-lister og/eller webforum for produktet		x
SUP12	HA-løsningen er open source		x
SUP13	HA-løsningen er dokumentert.	x	
SUP14	HA-løsningen er godt dokumentert.		x
	Konfigurasjon av tjenester		
CONF01	Det må være mulig å konfigurere hvilke ressurser som skal være knyttet til en tjeneste Eksempler: filsystemer, ip-adresser og prosesser.	x	
CONF02	Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt. Eksempel på egen betingelse: Når filsystemene til pakken er utilgjengelige skal tjeneste tas ned og ikke flyttes til en annen node	x	
CONF03	Det oppleves som enkelt å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt. Eksempel på egen betingelse: Når filsystemene til pakken er utilgjengelige skal tjeneste tas ned og ikke flyttes til en annen node		x
CONF04	Det må være mulig å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.	x	
CONF05	Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned. Eksempel 1: Innstillinger rundt hvilke andre noder en tjeneste kan kjøre på bør kunne endres uten at pakken må tas ned. Eksempel 2: Hvor mange ganger HA-løsningen forsøker å avmontere et filsystem bør være mulig å justere uten at tjenesten må tas ned (gjelder HA-løsninger med aktiv/passiv oppsett hvor dette er en aktuell problemstilling)		x
CONF06	Det oppleves som enkelt å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.		x
CONF07	Det må være mulig å spesifisere hvilke noder en tjeneste skal kunne gå på, og rekkefølgen for hvordan tjenesten skal gjøre failover til andre noder.	x	
CONF08	Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder Eksempel: Tjeneste X skal ikke gå på samme node som tjeneste Y", eller "Tjeneste X og tjeneste Y må alltid gå på samme node".	x 1)	x
CONF09	HA-løsningen kan konfigureres slik at den restarter ressurser i en tjeneste som den har detektert at ikke lenge kjører	x	
	Lagring og filsystem	x	
STOR01	Løsningen må kunne benytte seg av delt lagring, enten via et SAN (FC eller iSCSI) eller nettverksfilssystem (NFS).	x	
STOR02	Det er ønskelig at løsningen kan støtte et cluster-filsystem som tillater samtidig skriving		x

	fra flere noder.		
STOR03	Løsningen bør være supportert på den lagrings-infrastruktur som USIT allerede støtter (SAN, fc-switcher, lagringsprodukter) så vi slipper å kjøpe inn egen lagring til løsningen.		x
	Overvåkning og feilsøk		
MON01	Det må enkelt å få ut info om status for HA-løsningen, dets noder og dets tjenester	x	
MON02	Logging må være tilstrekkelig slik at det er mulig å feilsøke ved eventuelle problemer.	x	
MON03	Det er ønskelig at HA-løsningen kan sende sine logger fra alle noder til et sentralisert sted Eksempel: I ServiceGuard blir pakkeloggene liggende på den noden hvor pakken kjørte til enhver tid.		x
MON04	HA-løsningen må la seg overvåke av Nagios	x	
MON05	Oracle-databaser som kjører i HA-løsningen må la seg overvåke av GridControl.	x	
MON06	Det må være mulig å få generert en alarm dersom en ressurs i tjenesten ikke kjører	x	
	Nettverk		
NET01	HA-løsningen støtter IPv6		x
	Priser og lisenser		
LIC01	Kostnaden for løsningen ved anskaffelse og drift bør ikke være vesentlig høyere enn for dagens SG-løsninger, også for eksterne kunder som ikke er omfattet av UiO sin site-lisens.		x
	Redundanse og robusthet		
ROB01	HA-løsningen skal være en "moden" teknologi idet det tas i bruk av USIT Eksempel: Vi unngår produkter som er i ".0"-versjon, dvs første release av en gitt versjon.	x	
ROB02	Løsningen må kunne settes opp og være supportert med multiple paths til både nettverk og lagring for å unngå single-point-of-failure.	x	
ROB03	Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).	x	
ROB04	Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på to noder samtidig).	x	
ROB05	Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing)	x	
ROB06	Det er ønskelig med mest mulig isolasjon mellom forskjellige tjenester som går på samme node i løsningen. Eksempel: Et sikkerhetshull i en tjeneste på en node bør ikke gjøre det mulig å kompromitere andre tjenester på samme node.		x

ROB07	Nodene i HA-løsningen kan stå i forskjellige nett.		x
	Skalerbarhet		
SCA01	Det må være mulig å melde noder inn i og ut av HA-løsningen uten at det påvirker tjenestene i løsningen (med unntak av tjenester som går på noder som skal fjernes).	x	
SCA02	Det må være mulig å definere nye tjenester i løsningen uten at det krever nedetid på andre tjenester i løsningen.	x	
SCA03	Det er ønskelig at løsningen gjør det mulig å skalere tjenester ut over kapasiteten til en enkelt node Forklaring: Tjenester kan kjøre på flere noder samtidig (aktiv-aktiv)	x 2)	x
SCA04	HA-løsningen må støtte opp til 16 noder i løsningen	x	
SCA05	HA-løsningen støtter over 16 noder i løsningen		x
	Ytelse og ressurser		
PERF01	Tjenester på HA-løsningen må ikke ha nevneverdig dårligere ytelse enn om tjenesten kjørte på tilsvarende maskinvare utenfor en HA-løsning.	x	

Fotnoter:

- 1) SAP har behov for å definere avhengighet mellom pakker. Ingen andre løsninger vi kjører på SG har dette kravet.
- 2) Noen Oracle-databaser har behov for å kunne skalere ut over kapasiteten til en node, f.eks Atlas 3D. Dette gjelder bare et fåtall databaser så dette er ikke et generelt krav for alle tjenester.
- 3) Dette kravet gjelder kun for noen av Oracle-databasene som i dag kjører på RAC.

Appendix 10 Liste over HA-teknologier

Status	Produktnavn/beskrivelse Konsept	Linker	Kommentarer / Begrunnelse (fjernet fra sluttrapporten)
Utgår	EMC Autostart	http://www.emc.com/products/detail/software/autostart.htm	
Utgår	DBvisit	http://www.avisit.co.nz/	
Utgår	Oracle VM	Oversikt	
Utgår	Oracle RAC OneNode	Offisiell beskrivelse Datasheet (PDF) Technical Overview (PDF) - Blogg 1 2	
Kandidat	Oracle DataGuard	Offisiell side	
Kandidat	Oracle Active DataGuard	Offisiell side	
Kandidat	Oracle Clusterware	Offisiell side	
Utgår	Oracle Solaris Cluster	Offisiell side	
Utgår	Fujitsu PrimeCluster	http://www.fujitsu.com/global/services/computing/server/unix/optionalsw/PRMPWR_pclproducts.html	
Kandidat	RedHat Cluster Suite	Overview Migrasjon fra ServiceGuard til RCHS www.redhat.com/docs/manuals/csgfs/ http://sources.redhat.com/cluster/wiki/FAQ http://sources.redhat.com/cluster/wiki/CategoryHowTo http://news.gmane.org/gmane.linux.redhat.cluster	
Utgår	Linux HA (heartbeat)	http://www.linux-ha.org/ http://www.drbd.org/ Webcast drbd, mysql og litt om heartbeat	
Utgår	OpenAIS	http://openais.org	
Utgår	OpenSAFfire	http://goahead.com	
Kandidat	Steeleye lifekeeper	http://www.steeleye.com/	
Kandidat	Veritas Cluster	http://www.symantec.com/business/cluster-server Migrere fra ServiceGuard til Veritas Cluster Server Oracle og VCS	
Kandidat	VMware HA / VMware FT	VmWare HA HA Deepdive - presentasjon VMware FT (Fault Tolerance)	
Utgår	Microsoft Hyper-V		
Utgår	Xen		
Utgår	Anycast	Anycast på Wikipedia	
Utgår	Egenera PAM	http://www.egenera.com/	
Utgår	IBM Tivoli System Automation (TSA)	http://www-01.ibm.com/software/tivoli/products/sys-auto-multi/	
Utgår	NEC Express-Cluster	http://www.necam.com/entsw/ExpressCluster/	
Utgår	SUSE Linux Enterprise High Availability Extension	http://www.novell.com/products/highavailability/	

ADM08	oppgraderes "rolling" slik at oppgraderingen ikke gir nedetid på løsningen. Eksempel: I RAC kan Oracle-installasjonen oppgraderes rolling på node for node mens tjenesten fortsatt er tilgjengelig på de andre nodene.	Krav/ Ønske 3)	Nei	Nei	Nei	Nei	Nei	Nei
	Applikasjonssupport							
APP01	Det er supportert å kjøre Oracle på HA-løsningen	Krav	Ja	Ja	Ja	Ja	Ja	Nei
APP02	Det finnes ferdig Oracle-rammeverk tilpasset HA-løsningen	Ønske	Ja	Ja	Ja	Ja	Ja	N/A
APP03	Postgres kan kjøre på løsningen	Krav	Nei	Nei	Ja	Ja	Ja	Delvis
APP04	Det finnes ferdig Postgres-rammeverk tilpasset HA-løsningen	Ønske	Nei	Nei	Ja	Ja	Nei	N/A
APP05	Det er supportert/mulig (avhengig av hvilken MySQL løsning vi velger) å kjøre MySQL på løsningen	Krav	Nei	Nei	Ja	Ja	Ja	Delvis
APP06	Det finnes ferdig MySQL-rammeverk tilpasset HA-løsningen	Ønske	Nei	Nei	Ja	Ja	Ja	N/A
APP07	Det er supportert å kjøre SAP på HA-løsningen.	Krav	Ja	Ja	Ja	Ja	Ja	Ukjent
APP08	Det finnes ferdig SAP-rammeverk tilpasset HA-løsningen	Krav 1)	N/A	Ja	Ja	Ja	Ja	N/A
APP09	IMAP kan kjøre på løsningen	Krav	Nei	Nei	Ja	Ja	Ja	Delvis
APP10	Det finnes ferdig IMAP-rammeverk tilpasset HA-løsningen	Ønske	Nei	Nei	Nei	Nei	Nei	N/A
APP11	HA-løsningen har støtte for <i>fast application notification</i> (FAN) failover tilkoplingsteknologi Forklaring: Vortex bruker dette for å redusere nedetid når en node eller databaseinstanse feiler	Krav/ Ønske 3)	Ja	Ja	Nei	Nei	Nei	Nei
	Failover og flytting av tjenester							
FAIL01	Ved flytting av en tjeneste fra en node til en annen i HA-løsningen er nedetiden for tjenesten omtrent det samme som tiden det tar å stoppe og starte ressursene i	Krav	Ja	Ja	Ja	Ja	Ja	Ja

	tjenesten på tilsvarende maskinvare utenfor HA-løsningen.							
FAIL02	Det må være mulig å flytte en tjeneste manuelt fra en node til en annen.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
FAIL03	Det er ønskelig at man gjøre en planlagt flytting til en annen node uten nedetid på tjenesten (live migrering).	Ønske	Delvis	Ja	Nei	Nei	Nei	Ja
FAIL04	Feil på maskinvare i HA-løsningen skal ikke gi langvarig nedetid for brukere av tjenesten	Krav	Ja	Ja	Ja	Ja	Ja	Ja
FAIL05	Feil på maskinvare i HA-løsningen skal ikke gi nedetid for brukere av tjenesten	Ønske	Ja	Ja	Nei	Nei	Nei	Ja
FAIL06	Feil på maskinvare i HA-løsningen skal ikke kreve manuelle inngrep for å gjøre tjenesten tilgjengelig for brukeren.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
FAIL07	Brukere av tjenesten skal ikke måtte endre i konfigurasjon på klienten etter at en tjeneste er flyttet til en annen node i HA-løsningen.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
	HW- og SW-support og leverandørforhold							
SUP01	HA-løsningen skal kunne kjøre på maskinvare fra flere leverandører Eksempel: Ikke bare på Dell-maskiner av en gitt type	Ønske	Ja	Ja	Ja	Ja	Ja	Ja
SUP02	Nodene i HA-løsningen kan være virtuelle	Ønske	Ja	Delvis	Ja	Ja	Ja	N/A
SUP03	HA-løsningen skal kunne kjøre på ett eller flere operativsystemer som støttes av USIT i dag.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
SUP04	HA-løsningen bør kunne kjøre på på de nyeste versjonene av de operativsystemer som støttes av USIT i dag.	Ønske	Ja	Ja	Ja	Ja	Ja	Ja
SUP05	HA løsningen må ha support fra leverandør(er) eller tredjepart(er) i alle ledd.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
SUP06	Det er ønskelig med færrest mulig forskjellige leverandører involvert i den totale løsningen Eksempel på uønsket situasjon: Oracle-	Ønske	Ja (for noder med Oracle)	Ja (for noder med Oracle)	Ja	Nei	Nei	Nei

	SW, HP-lagring, RedHat-OS, Veritas-cluster, Dell servere							
SUP07	Løsningen må være supportert på de versjoner av kommersiell software som er aktuelt for USIT å kjøre Eksempel: DBA-gruppen migrererer nå databaser til Oracle 11.2.0 som må være støttet.	Krav	Delvis	Delvis	Ja trolig	Ja	Ja	Ukjent
SUP08	USIT må ikke miste support på HA-løsningen dersom HA-løsningen er satt opp av USIT, eller er om konfigurasjonen er endret av USIT etter installasjon og initiell konfigurasjon.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
SUP09	Det finnes referanse sekunder for HA-løsningen som kjører i en driftssituasjon som kan sammenlignes med vår egen.	Krav	Ukjent	Ja	Ja	Ja	Ja	Ja
SUP10	HA-løsningen skal inngå i et aktivt prosjekt eller være i aktiv utvikling hos en "solid og levedyktig" leverandør.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
SUP11	Det finnes aktive epost-lister og/eller webforum for produktet	Ønske	Ja	Ja	Ja	Nei	Ja	Ja
SUP12	HA-løsningen er open source	Ønske	Nei	Nei	Ja	Nei	Nei	Nei
SUP13	HA-løsningen er dokumentert.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
SUP14	HA-løsningen er godt dokumentert.	Ønske	Ja	Ja	Ja	Ja	Ja	Ja
	Konfigurasjon av tjenester							
CONF01	Det må være mulig å konfigurere hvilke ressurser som skal være knyttet til en tjeneste Eksempler: filsystemer, ip-adresser og prosesser.	Krav	Nei	Ja	Ja	Ja	Ja	Ja
CONF02	Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt. Eksempel på egen betingelse: Når filsystemene til pakken er utilgjengelige skal tjeneste tas ned og ikke flyttes til en annen node	Krav	Nei	Nei	Ja	Ja	Ja	Ja
	Det oppleves som enkelt å kunne spesifisere egne betingelser som							

CONF03	må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt. Eksempel på egen betingelse: Når filsystemene til pakken er utilgjengelige skal tjeneste tas ned og ikke flyttes til en annen node	Ønske	Nei	Nei	Ukjent	Trolig	Trolig	Ja
CONF04	Det må være mulig å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.	Krav	Nei	Ja	Ja	Ja	Ja	Ja
CONF05	Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned. Eksempel 1: Innstillinger rundt hvilke andre noder en tjeneste kan kjøre på bør kunne endres uten at pakken må tas ned. Eksempel 2: Hvor mange ganger HA-løsningen forsøker å avmontere et filsystem bør være mulig å justere uten at tjenesten må tas ned (gjelder HA-løsninger med aktiv/passiv oppsett hvor dette er en aktuell problemstilling)	Ønske	Nei	Nei	Ja	Ja	Ja	Ja
CONF06	Det oppleves som enkelt å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.	Ønske	Nei	Nei	Ukjent	Ja	Ukjent	Ja
CONF07	Det må være mulig å spesifisere hvilke noder en tjeneste skal kunne gå på, og rekkefølgen for hvordan tjenesten skal gjøre failover til andre noder.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
CONF08	Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder Eksempel: Tjeneste X skal ikke gå på samme node som tjeneste Y", eller "Tjeneste X og tjeneste Y må alltid gå på samme node".	Krav/ Ønske 1)	Nei	Nei	Delvis	Ja	Ja	Ja
CONF09	HA-løsningen kan konfigureres slik at den restarter ressurser i en tjeneste som den har detektert at ikke lenge kjører	Krav	Nei	Nei	Ja	Ja	Ja	Nei
	Lagring og filsystem	Krav						
	Løsningen må kunne benytte seg							

STOR01	av delt lagring, enten via et SAN (FC eller iSCSI) eller nettverksfilssystem (NFS).	Krav	Ja	Ja	Ja	Ja	Ja	Ja
STOR02	Det er ønskelig at løsningen kan støtte et cluster-filsystem som tillater samtidig skriving fra flere noder.	Ønske	Nei	Ja	Ja	Nei	Ja	Ja
STOR03	Løsningen bør være supportert på den lagrings-infrastruktur som USIT allerede støtter (SAN, fc-switcher, lagringsprodukter) så vi slipper å kjøpe inn egen lagring til løsningen.	Ønske	Ja	Ja	Ja	Ja	Ja	Ja
	Overvåkning og feilsøk							
MON01	Det må enkelt å få ut info om status for HA-løsningen, dets noder og dets tjenester	Krav	Ja	Ja	Ja	Ja	Ja	Ja
MON02	Logging må være tilstrekkelig slik at det er mulig å feilsøke ved eventuelle problemer.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
MON03	Det er ønskelig at HA-løsningen kan sende sine logger fra alle noder til et sentralisert sted Eksempel: I ServiceGuard blir pakkeloggene liggende på den noden hvor pakken kjørte til enhver tid.	Ønske	Nei	Nei	Ja	Ja	Ja	Ja
MON04	HA-løsningen må la seg overvåke av Nagios	Krav	Ja	Ja	Ja	Trolig	Trolig	Ja
MON05	Oracle-databaser som kjører i HA-løsningen må la seg overvåke av GridControl.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
MON06	Det må være mulig å få generert en alarm dersom en ressurs i tjenesten ikke kjører	Krav	Ja	Ja	Ja	Ja	Ja	Ja
	Nettverk							
NET01	HA-løsningen støtter IPv6	Ønske	Ja	Ja	Nei	Nei	Ja	Ja
	Priser og lisenser							
LIC01	Kostnaden for løsningen ved anskaffelse og drift bør ikke være vesentlig høyere enn for dagens SG-løsninger, også for eksterne kunder som ikke er omfattet av	Ønske	Ja	Nei	Ukjent	Trolig	Ukjent	Nei

	UiO sin site-lisens.							
	Redundanse og robusthet							
ROB01	HA-løsningen skal være en "moden" teknologi idet det tas i bruk av USIT Eksempel: Vi unngår produkter som er i ".0"-versjon, dvs første release av en gitt versjon.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
ROB02	Løsningen må kunne settes opp og være supportert med multiple paths til både nettverk og lagring for å unngå single-point-of-failure.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
ROB03	Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).	Krav	Ja	N/A (Ja)	Ja	Ja	Ja	Ja
ROB04	Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på to noder samtidig).	Krav	Ja	N/A (Ja)	Ja	Ja	Ja	Ja
ROB05	Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing)	Krav	N/A	N/A (Ja)	Ja	Ja	Ja	Nei
ROB06	Det er ønskelig med mest mulig isolasjon mellom forskjellige tjenester som går på samme node i løsningen. Eksempel: Et sikkerhetshull i en tjeneste på en node bør ikke gjøre det mulig å kompromitere andre tjenester på samme node.	Ønske	Nei	Nei	Nei	Ukjent	Nei	Ja
ROB07	Nodene i HA-løsningen kan stå i forskjellige nett.	Ønske	Ja	Ja	Ja	Ja		Ja
	Skalerbarhet							
SCA01	Det må være mulig å melde noder inn i og ut av HA-løsningen uten at det påvirker tjenestene i løsningen (med unntak av tjenester som går på noder som skal fjernes).	Krav	Nei	Ja	Ja	Ja	Ja	Ja

SCA02	Det må være mulig å definere nye tjenester i løsningen uten at det krever nedetid på andre tjenester i løsningen.	Krav	Ja	Ja	Ja	Ja	Ja	Ja
SCA03	Det er ønskelig at løsningen gjør det mulig å skalere tjenester ut over kapasiteten til en enkelt node Forklaring: Tjenester kan kjøre på flere noder samtidig (aktiv-aktiv)	Krav/ Ønske 2)	Delvis	Ja	Nei	Nei	Nei	Nei
SCA04	HA-løsningen må støtte opp til 16 noder i løsningen	Krav	Ja	Ja	Ja	Ja	Ja	Ja
SCA05	HA-løsningen støtter over 16 noder i løsningen	Ønske	Ja	Trolig	Nei	Ja	Ja	Ja
	Ytelse og ressurser							
PERF01	Tjenester på HA-løsningen må ikke ha nevneverdig dårligere ytelse enn om tjenesten kjørte på tilsvarende maskinvare utenfor en HA-løsning.	Krav	Ja	Ja	Ja	Ja	Ja	Delvis

Fotnoter:

- 1) SAP har behov for å definere avhengighet mellom pakker. Ingen andre løsninger vi kjører på SG har dette kravet.
- 2) Noen Oracle-databaser har behov for å kunne skalere ut over kapasiteten til en node, f.eks Atlas 3D. Dette gjelder bare et fåtall databaser så dette er ikke et generelt krav for alle tjenester.
- 3) Dette kravet gjelder kun for noen av Oracle-databasene som i dag kjører på RAC.

Appendix 12 Mapping av Oracle dataguard mot kravspesifikasjon - også active dataguard opsjonen

Administrasjon, drift og driftsmiljø

-  ADM01 Det må være støtte "rolling upgrade" av operativsystemet på nodene i HA-løsningen.
 - Ja, man kan faile over mellom standby og primary mens OS oppdateres; http://download.oracle.com/docs/cd/E11882_01/server.112/e17022/concepts.htm#i1039416
-  ADM02 Nodene i HA-løsningen må kunne oppdateres fortløpende med systemoppgraderinger uten at vi mister support på HA-løsningen.
 - Ja, sertifisert i hht [my oracle support - certification tab](#)
-  ADM03 Det er ønskelig at leverandøren av HA-løsningen raskt støtter nye Linux-kjerner.
 - Ja, som vist i ADM02, så støttes feks. RHEL 5, så om RHEL 5 får en ny kjerne er det fortsatt RHEL 5.
-  ADM04 HA-løsningen må ikke være "for kompleks" å sette opp og å drifte
 - Noe - litt jobb å sette opp, men enklere å drifte med dgmgrl. http://download.oracle.com/docs/cd/E11882_01/server.112/e17022/toc.htm
-  ADM05 Nye versjoner av HA-løsningen er mest mulig bakoverkompatibel slik at oppgraderinger ikke bryter eksisterende konfigurasjon
 - Nei - må ha samme versjon på primary og standby ved bruk av physical standby. Logisk standby gir denne muligheten, men er umodent.
http://download.oracle.com/docs/cd/E11882_01/server.112/e17022/upgrades.htm#CHDBJAGG
-  ADM06 Programvaren i HA-løsningen kan oppgraderes "rolling", dvs node for node uten at alle tjenester må tas ned
 - Nei, men det kan gjøres med en transient logical standby -
http://download.oracle.com/docs/cd/B28359_01/server.111/b28282/schedule_outage.htm#BABGBADE
-  ADM07 Det er mulighet til å skripte kommunikasjonen mot HA-systemet (ikke avhengig av GUI) for å kunne automatisere driftsmiljøet rundt løsningen.
 - Ja, mot dgmgrl - http://download.oracle.com/docs/cd/E11882_01/server.112/e17023/toc.htm
-  ADM08 Tjenester i løsningen kan oppgraderes "rolling" slik at oppgraderingen ikke gir nedetid på løsningen.
 - Nei, ikke praktisk siden standby er en kopi av primary så bør begge oppgraderes samtidig.

Applikasjonssupport

-  APP01 Det er supportert å kjøre Oracle på HA-løsningen
 - Ja
-  APP02 Det finnes ferdig Oracle-rammeverk tilpasset HA-løsningen
 - Ja
-  APP03 Postgres kan kjøre på løsningen
 - Nei
-  APP04 Det finnes ferdig Postgres-rammeverk tilpasset HA-løsningen

- Nei
-  APP05 Det er supportert/mulig (avhengig av hvilken MySQL løsning vi velger) å kjøre MySQL på løsningen
 - Nei
-  APP06 Det finnes ferdig MySQL-rammeverk tilpasset HA-løsningen
 - Nei
-  APP07 Det er supportert å kjøre SAP på HA-løsningen.
 - Ja - databasen kan kjøre på standby. <http://www.oracle.com/newsletters/sap/products/dataguard.html>
-  APP08 Det finnes ferdig SAP-rammeverk tilpasset HA-løsningen
 - Nei, det er ingen spesielle utover databasekravene. SAP er uvitende om at oracle kjører på dataguard eller vanlig instanse. Kun databasen blir håndtert av DataGuard.
-  APP09 IMAP kan kjøre på løsningen
 - Nei
-  APP10 Det finnes ferdig IMAP-rammeverk tilpasset HA-løsningen
 - Nei
-  APP11 HA-løsningen har støtte for fast application notification (FAN) failover tilkoplingsteknologi
 - Ja - http://download.oracle.com/docs/cd/E14072_01/server.112/e10595/restart001.htm

Failover og flytting av tjenester

-  FAIL01 Ved flytting av en tjeneste fra en node til en annen i HA-løsningen er nedetiden for tjenesten omtrent det samme som tiden det tar å stoppe og starte ressursene i tjenesten på tilsvarende maskinvare utenfor HA-løsningen.
 - Ja - man bytter over med dgmgrl
-  FAIL02 Det må være mulig å flytte en tjeneste manuelt fra en node til en annen.
 - Ja - med dgmgrl
-  FAIL03 Det er ønskelig at man gjøre en planlagt flytting til en annen node uten nedetid på tjenesten (live migrering).
 - Ja - med dgmgrl - vanligvis vil det medføre nedetid, men med tilpassede klienter og FAN teknologi kan det gjøres uten merkbar nedetid - http://download.oracle.com/docs/cd/E11882_01/server.112/e17120/restart002.htm#ADMIN13192
-  FAIL04 Feil på maskinvare i HA-løsningen skal ikke gi langvarig nedetid for brukere av tjenesten.
 - Ja - man kan konfigurere dgmgrl til å automatisk bytte til standby, slik at standby blir primary - http://download.oracle.com/docs/cd/E11882_01/server.112/e17023/sofo.htm#i1027843
-  FAIL05 Feil på maskinvare i HA-løsningen skal ikke gi nedetid for brukere av tjenesten
 - Ja - se FAIL04
-  FAIL06 Feil på maskinvare i HA-løsningen skal ikke kreve manuelle inngrep for å gjøre tjenesten tilgjengelig for brukeren.
 - Ja
-  FAIL07 Brukere av tjenesten skal ikke måtte endre i konfigurasjon på klienten etter at en tjeneste er flyttet til en annen node i HA-løsningen.

Ja - men krever tilpasset oppsett -

http://download.oracle.com/docs/cd/E11882_01/server.112/e17023/sofo.htm#BCGFFFBG

HW- og SW-support og leverandørforhold

-  SUP01 HA-løsningen skal kunne kjøre på maskinvare fra flere leverandører
 - Ja - se my oracle support - [certification](#)
-  SUP02 Nodene i HA-løsningen kan være virtuelle
 - Ja - men bare på oracle VM
-  SUP03 HA-løsningen skal kunne kjøre på ett eller flere operativsystemer som støttes av USIT i dag.
 - Ja - RHEL 5u2 - [certification](#)
-  SUP04 HA-løsningen bør kunne kjøre på de nyeste versjonene av de operativsystemer som støttes av USIT i dag.
 - Ja - se SUP03
-  SUP05 HA løsningen må ha support fra leverandør(er) eller tredjepart(er) i alle ledd.
 - Ja - se SUP03
-  SUP06 Det er ønskelig med færrest mulig forskjellige leverandører involvert i den totale løsningen
 - Ja - leverandør er oracle
-  SUP07 Løsningen må være supportert på de versjoner av kommersiell software som er aktuelt for USIT å kjøre
 - Ja - Oracle 10.2.0 og 11.2.0 databaser
 - Nei - støtter i seg selv ikke noe mer enn databasen - ikke SAP (NetWeaver 7.01, NetWeaver PI 7.10, ERP 6.0, SolMan 7.01)
-  SUP08 USIT må ikke miste support på HA-løsningen dersom HA-løsningen er satt opp av USIT, eller er om konfigurasjonen er endret av USIT etter installasjon og initiell konfigurasjon.
 - Ja - det er meningen å kunne konfigurere oracle dataguard selv
-  SUP09 Det finnes referansekunder for HA-løsningen som kjører i en driftssituasjon som kan sammenlignes med vår egen.
 - Nei
-  SUP10 HA-løsningen skal inngå i et aktivt prosjekt eller være i aktiv utvikling hos en "solid og levedyktig" leverandør.
 - Ja - oracle har en ok økonomi - http://www.nasdaq.com/asp/quotes_reports.asp?symbol=orcl&selected=orcl
-  SUP11 Det finnes aktive epost-lister og/eller webforum for produktet.
 - Ja - feks: <http://forums.oracle.com/>, <http://asktom.oracle.com/pls/apex/f?p=100:1:0>
-  SUP12 HA-løsningen er open source.
 - Nei
-  SUP13 HA-løsningen er dokumentert.
 - Ja - <http://www.oracle.com/technetwork/database/features/availability/dataguardoverview-083155.html>
-  SUP14 HA-løsningen er godt dokumentert.

Ja - se SUP13 - men ofte man må bruke support for å finne ut av ting fordi det ikke alltid fungerer som dokumentert, eller det er ikke dokumentert.

Konfigurasjon av tjenester

-  CONF01 Det må være mulig å konfigurere hvilke ressurser som skal være knyttet til en tjeneste
 - Nei - støtter bare database
-  CONF02 Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.
 - Nei - dataguard er mer disaster recovery enn high availability.
-  CONF03 Det oppleves som enkelt å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.
 - Nei - Kan ikke spesifisere egne betingelser på en enkel måte.
-  CONF04 Det må være mulig å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.
 - Nei - Innenfor databasen kan man lage triggere, men ikke støtte/mening å starte andre kommandoer enn databasen.
-  CONF05 Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned.
 - Nei - dataguard støtter ikke annet enn databasen
-  CONF06 Det oppleves som enkelt å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.
 - Nei
-  CONF07 Det må være mulig å spesifisere hvilke noder en tjeneste skal kunne gå på, og rekkefølgen for hvordan tjenesten skal gjøre failover til andre noder.
 - Ja - fra primary til n standby databaser.
-  CONF08 Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder
 - Nei
-  CONF09 HA-løsningen kan konfigureres slik at den restarter ressurser i en tjeneste som den har detektert at ikke lenge kjører.
 - Nei , unntatt noen til oracle
selv: http://download.oracle.com/docs/cd/E14072_01/server.112/e10595/restart001.htm

Lagring og filsystem

-  STOR01 Løsningen må kunne benytte seg av delt lagring, enten via et SAN (FC eller iSCSI) eller nettverksfilssystem (NFS).
 - Ja - kan bruke ASM
-  STOR02 Det er ønskelig at løsningen kan støtte et cluster-filssystem som tillater samtidig skriving fra flere noder.

Nei - det er ikke hensikten med dataguard - da bruker man rac.

-  STOR03 Løsningen bør være supportert på den lagrings-infrastruktur som USIT allerede støtter (SAN, fc-switcher, lagringsprodukter) så vi slipper å kjøpe inn egen lagring til løsningen.
 - Ja

Overvåkning og feilsøk

-  MON01 Det må enkelt å få ut info om status for HA-løsningen, dets noder og dets tjenester.
 - Ja - dgmgrl
-  MON02 Logging må være tilstrekkelig slik at det er mulig å feilsøke ved eventuelle problemer.
 - Ja - alert logger til databaseinstansene, og [dgmgrl](#)
-  MON03 Det er ønskelig at HA-løsningen kan sende sine logger fra alle noder til et sentralisert sted
 - Nei
-  MON04 HA-løsningen må la seg overvåke av Nagios
 - Ja - på samme måte som det gjøres i dag, via Gridcontrol
-  MON05 Oracle-databaser som kjører i HA-løsningen må la seg overvåke av GridControl.
 - Ja
-  MON06 Det må være mulig å få generert en alarm dersom en ressurs i tjenesten ikke kjører
 - Ja - via gridcontrol

Nettverk

-  NET01 HA-løsningen støtter IPv6
 - Ja - http://download.oracle.com/docs/cd/E11882_01/network.112/e10836/layers.htm#CACGCEAG

Priser og lisenser

-  LIC01 Kostnaden for løsningen ved anskaffelse og drift bør ikke være vesentlig høyere enn for dagens SG-løsninger, også for eksterne kunder som ikke er omfattet av UiO sin site-lisens.
 - Ja - har allerede dataguard inne

Redundanse og robusthet

-  ROB01 HA-løsningen skal være en "moden" teknologi idet det tas i bruk av USIT
 - Ja. Har vært gjennom flere oracle versjoner.
-  ROB02 Løsningen må kunne settes opp og være supportert med multiple paths til både nettverk og lagring for å unngå single-point-of-failure.
 - Ja, støtter seg på at linux støtter dette.
-  ROB03 Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).
 - Ja, dgmgrl, og standby blir satt i mount/read-only



-  ROB04 Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på to noder samtidig).
 - Ja - den er ment å være aktiv på 2 steder samtidig. Man deler ikke ressurser, men har to replikerte enheter.
-  ROB05 Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing).
 - Nei - det er ikke et cluster. Men mulig å ta vekk noder, men da må man rydde opp, feks. fjerne standby_log parametre -> krever restart(?)
-  ROB06 Det er ønskelig med mest mulig isolasjon mellom forskjellige tjenester som går på samme node i løsningen.
 - Nei, kjører som samme bruker
-  ROB07 Nodene i HA-løsningen kan stå i forskjellige nett.
 - Ja - dataguard lagt for Distaster Recovery

Skalerbarhet

-  SCA01 Det må være mulig å melde noder inn i og ut av HA-løsningen uten at det påvirker tjenestene i løsningen (med unntak av tjenester som går på noder som skal fjernes).
 - Nei - påvirker primary, må endre parametre, og den må sende redo til nye noder
-  SCA02 Det må være mulig å definere nye tjenester i løsningen uten at det krever nedetid på andre tjenester i løsningen.
 - Ja
-  SCA03 Det er ønskelig at løsningen gjør det mulig å skalere tjenester ut over kapasiteten til en enkelt node
 - Ja - man kan kjøre opp standby i read only og således ta en pause i redo apply'en, for read only, evt kjøre active dataguard - som er read only mens redo apply'es.
-  SCA04 HA-løsningen må støtte opp til 16 noder i løsningen.
 - Ja - opptil 31 destinasjoner å sende redo
http://download.oracle.com/docs/cd/E11882_01/server.112/e17110/initparams121.htm#REFRN10086
-  SCA05 HA-løsningen støtter over 16 noder i løsningen.
 - Ja - se SCA04

Ytelse og ressurser

-  PERF01 Tjenester på HA-løsningen må ikke ha nevneverdig dårligere ytelse enn om tjenesten kjørte på tilsvarende maskinvare utenfor en HA-løsning.
 - Ja - men noe mindre ytelse pga nettverksbelastning og cpu for å sende redo logger til standby database(r). Men standby databaser kan også benyttes til read only spørringer.

Appendix 13 Mapping av Oracle Clusterware (Oracle Grid) mot kravspesifikasjon

Oracle RAC 11.2

Administrasjon, drift og driftsmiljø

- ADM01 Det må være støtte "rolling upgrade" av operativsystemet på nodene i HA-løsningen.
 - Ja, men ytelsen vil da være tilsvarende redusert når en node er nede.
- ADM02 Nodene i HA-løsningen må kunne oppdateres forløpende med systemoppgraderinger uten at vi mister support på HA-løsningen.
 - Ja, sertifisert i hht [my oracle support - certification tab](#)
- ADM03 Det er ønskelig at leverandøren av HA-løsningen raskt støtter nye Linux-kjerner.
 - Ja, som vist i ADM02, så støttes feks. RHEL 5, så om RHEL 5 får en ny kjerne er det fortsatt RHEL 5.
- ADM04 HA-løsningen må ikke være "for kompleks" å sette opp og å drifte
 - Nei - litt jobb å sette opp, men enklere å drifte ettersom man kan bla. patche rullende, en node av gangen.
- ADM05 Nye versjoner av HA-løsningen er mest mulig bakoverkompatibel slik at oppgraderinger ikke bryter eksisterende konfigurasjon
 - Nei - må ha samme versjon på nodene. Kan ikke ha 10.2 clusterware på en node og 11.2 grid clusterware på en
- ADM06 Programvaren i HA-løsningen kan oppgraderes "rolling", dvs node for node uten at alle tjenester må tas ned
 - Ja, clusterware kan oppgraderes rolling - men ikke databasen - ID 338706.1
- ADM07 Det er mulighet til å skripte kommunikasjonen mot HA-systemet (ikke avhengig av GUI) for å kunne automatisere driftsmiljøet rundt løsningen.
 - Ja, srvctl er kommandolinje verktøyet
- ADM08 Tjenester i løsningen kan oppgraderes "rolling" slik at oppgraderingen ikke gir nedetid på løsningen.
 - Nei, det er fortsatt en database i bunn som må oppgraderes.

Applikasjonssupport

- APP01 Det er supportert å kjøre Oracle på HA-løsningen
 - Ja
- APP02 Det finnes ferdig Oracle-rammeverk tilpasset HA-løsningen
 - Ja
- APP03 Postgres kan kjøre på løsningen
 - Nei
- APP04 Det finnes ferdig Postgres-rammeverk tilpasset HA-løsningen
 - Nei
- APP05 Det er supportert/mulig (avhengig av hvilken MySQL løsning vi velger) å kjøre MySQL på løsningen
 - Nei
- APP06 Det finnes ferdig MySQL-rammeverk tilpasset HA-løsningen
 - Nei
- APP07 Det er supportert å kjøre SAP på HA-løsningen.
 - Ja, man kan kjøre databasen til SAP på rac
- APP08 Det finnes ferdig SAP-rammeverk tilpasset HA-løsningen
 - Ja, SAP har et eget opplegg i RAC sammenheng. I tillegg har de sitt eget 'verktøy', SAPCTL for styring av dette.

Det er 2 tjenester i SAP som er helt vesentlige for applikasjonen. Dette er Enqueue og Message serverene. Dersom disse 2 av en eller annen grunn går ned, må hele applikasjonen restartes. I RAC har SAP lagt opp til et master-slave system som sørger for en vesentlig bedre overlevelsessevne for applikasjonen da slaven kan overta som master som følge av at den er et speil av status til master.

SAP legger også inn servicene til applikasjonen(e) inn i RAC, se side 74 og utover i <https://www.usit.uio.no/tj-grupper/sapuio-drift/dokumentasjon/driftdokumenter/ConfigurationOfSAPNetWeaverForOracleGridInfrastructure11.2withOracleRealApplicationClusters11gRelease2.pdf>

- APP09 IMAP kan kjøre på løsningen
 - Nei
- APP10 Det finnes ferdig IMAP-rammeverk tilpasset HA-løsningen
 - Nei



-  APP11 HA-løsningen har støtte for fast application notification (FAN) failover tilkoplingsteknologi
 - Ja - http://download.oracle.com/docs/cd/E11882_01/server.112/e17120/restart002.htm#ADMIN13194

Failover og flytting av tjenester

-  FAIL01 Ved flytting av en tjeneste fra en node til en annen i HA-løsningen er nedetiden for tjenesten omtrent det samme som tiden det tar å stoppe og starte ressursene i tjenesten på tilsvarende maskinvare utenfor HA-løsningen.
 - Ja - eller umerkbar om feks. FAN brukes
-  FAIL02 Det må være mulig å flytte en tjeneste manuelt fra en node til en annen.
 - Ja - tjenesten er oppe på begge samtidig
-  FAIL03 Det er ønskelig at man gjøre en planlagt flytting til en annen node uten nedetid på tjenesten (live migrering).
 - Ja
-  FAIL04 Feil på maskinvare i HA-løsningen skal ikke gi langvarig nedetid for brukere av tjenesten.
 - Ja - annen node vil kjøre, men totalt sett blir det en nodes bortfalls dårligere ytelse
-  FAIL05 Feil på maskinvare i HA-løsningen skal ikke gi nedetid for brukere av tjenesten
 - Ja - se FAIL04
-  FAIL06 Feil på maskinvare i HA-løsningen skal ikke kreve manuelle inngrep for å gjøre tjenesten tilgjengelig for brukeren.
 - Ja - med SCAN vil klientene kople seg på overlevende noder - http://download.oracle.com/docs/cd/E11882_01/install.112/e17214/undrstnd.htm#BEIFHHBE
-  FAIL07 Brukere av tjenesten skal ikke måtte endre i konfigurasjon på klienten etter at en tjeneste er flyttet til en annen node i HA-løsningen.
 - Ja - men krever 11gr2 klient - se FAIL06

HW- og SW-support og leverandørforhold

-  SUP01 HA-løsningen skal kunne kjøre på maskinvare fra flere leverandører
 - Ja - se my oracle support - [certification](#)
-  SUP02 Nodene i HA-løsningen kan være virtuelle
 - Ja - men bare på oracle VM
-  SUP03 HA-løsningen skal kunne kjøre på ett eller flere operativsystemer som støttes av USIT i dag.
 - Ja - RHEL 5u2 - [certification](#)
-  SUP04 HA-løsningen bør kunne kjøre på de nyeste versjonene av de operativsystemer som støttes av USIT i dag.
 - Ja - se SUP03
-  SUP05 HA løsningen må ha support fra leverandør(er) eller tredjepart(er) i alle ledd.
 - Ja - se SUP03
-  SUP06 Det er ønskelig med færrest mulig forskjellige leverandører involvert i den totale løsningen
 - Ja - leverandør er oracle
-  SUP07 Løsningen må være supportert på de versjoner av kommersiell software som er aktuelt for USIT å kjøre
 - Ja - Oracle 11.2 databaser
 - Nei - støtter i seg selv ikke noe mer enn databasen - ikke SAP (NetWeaver 7.01, NetWeaver PI 7.10, ERP 6.0, SolMan 7.01)
-  SUP08 USIT må ikke miste support på HA-løsningen dersom HA-løsningen er satt opp av USIT, eller er om konfigurasjonen er endret av USIT etter installasjon og initiell konfigurasjon.
 - Ja - det er meningen å kunne konfigurere oracle rac selv
-  SUP09 Det finnes referanse sekunder for HA-løsningen som kjører i en driftssituasjon som kan sammenlignes med vår egen.
 - Ja, vi vet at UiB, UiT og flere har konsolidert på RAC.
-  SUP10 HA-løsningen skal inngå i et aktivt prosjekt eller være i aktiv utvikling hos en "solid og levedyktig" leverandør.
 - Ja - oracle har en ok økonomi - http://www.nasdaq.com/asp/quotes_reports.asp?symbol=orcl&selected=orcl
-  SUP11 Det finnes aktive epost-lister og/eller webforum for produktet.
 - Ja - feks: <http://forums.oracle.com/>, <http://asktom.oracle.com/pls/apex/f?p=100:1:0>
-  SUP12 HA-løsningen er open source.
 - Nei
-  SUP13 HA-løsningen er dokumentert.
 - Ja - http://www.oracle.com/pls/db112/portal.portal_db?selected=16&frame=
-  SUP14 HA-løsningen er godt dokumentert.
 - Ja - se SUP13 - men ofte man må bruke support for å finne ut av ting fordi det ikke alltid fungerer som dokumentert, eller det er ikke dokumentert.

Konfigurasjon av tjenester

-  CONF01 Det må være mulig å konfigurere hvilke ressurser som skal være knyttet til en tjeneste
 - Ja - støtter i utgangspunktet databasen, men har også mulighet til å la clusterware starte andre ressurser - http://download.oracle.com/docs/cd/E11882_01/rac.112/e16794/crschp.htm#CWADD91297
-  CONF02 Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.
 - Nei - har ikke funnet dokumentasjon på dette, annet enn at man kan skrive det selv.
-  CONF03 Det oppleves som enkelt å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.
 - Nei - Kan ikke spesifisere egne betingelser på en enkel måte.
-  CONF04 Det må være mulig å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.
 - Ja - Innenfor databasen kan man lage triggere, og i CRS kan man gjøre som vist i CONF01.
-  CONF05 Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned.
 - Nei - ikke ut over det som er vanlig for en database
-  CONF06 Det oppleves som enkelt å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.
 - Nei
-  CONF07 Det må være mulig å spesifisere hvilke noder en tjeneste skal kunne gå på, og rekkefølgen for hvordan tjenesten skal gjøre failover til andre noder.
 - Ja - men man bestemmer ikke rekkefølge da alle noder er aktive.
-  CONF08 Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder
 - Nei
-  CONF09 HA-løsningen kan konfigureres slik at den restarter ressurser i en tjeneste som den har detektert at ikke lenge kjører.
 - Nei , unntatt noen til oracle selv: http://download.oracle.com/docs/cd/E14072_01/server.112/e10595/restart001.htm

Lagring og filsystem

-  STOR01 Løsningen må kunne benytte seg av delt lagring, enten via et SAN (FC eller iSCSI) eller nettverksfilssystem (NFS).
 - Ja - kan bruke ASM
-  STOR02 Det er ønskelig at løsningen kan støtte et cluster-filssystem som tillater samtidig skriving fra flere noder.
 - Ja - ACFS - ASM som filsystem - http://download.oracle.com/docs/cd/E11882_01/server.112/e16102/asmfilesystem.htm#OSTMG30000
-  STOR03 Løsningen bør være supportert på den lagrings-infrastruktur som USIT allerede støtter (SAN, fc-switcher, lagringsprodukter) så vi slipper å kjøpe inn egen lagring til løsningen.
 - Ja

Overvåkning og feilsøk

-  MON01 Det må enkelt å få ut info om status for HA-løsningen, dets noder og dets tjenester.
 - Ja - srvctl - http://download.oracle.com/docs/cd/E11882_01/server.112/e17120/restart004.htm#ADMIN12726
-  MON02 Logging må være tilstrekkelig slik at det er mulig å feilsøke ved eventuelle problemer.
 - Ja - alert logger til databaseinstansene, og cluster-komponentene - http://download.oracle.com/docs/cd/E11882_01/rac.112/e16794/troubleshoot.htm#CWADD91091
-  MON03 Det er ønskelig at HA-løsningen kan sende sine logger fra alle noder til et sentralisert sted
 - Nei
-  MON04 HA-løsningen må la seg overvåke av Nagios
 - Ja - på samme måte som det gjøres i dag, via Gridcontrol
-  MON05 Oracle-databaser som kjører i HA-løsningen må la seg overvåke av GridControl.
 - Ja
-  MON06 Det må være mulig å få generert en alarm dersom en ressurs i tjenesten ikke kjører
 - Ja - via gridcontrol

Nettverk

-  NET01 HA-løsningen støtter IPv6
 - Ja - http://download.oracle.com/docs/cd/E11882_01/network.112/e10836/layers.htm#CACGCEAG

Priser og lisenser

-  LIC01 Kostnaden for løsningen ved anskaffelse og drift bør ikke være vesentlig høyere enn for dagens SG-løsninger, også for eksterne kunder som ikke er omfattet av UiO sin site-lisens.
 - Nei - egen lisens for RAC i EE. Følger med i Standard Edition

Redundanse og robusthet

-  ROB01 HA-løsningen skal være en "moden" teknologi idet det tas i bruk av USIT
 - Ja. Har vært gjennom flere oracle versjoner.
-  ROB02 Løsningen må kunne settes opp og være supportert med multiple paths til både nettverk og lagring for å unngå single-point-of-failure.
 - Ja, støtter seg på at linux støtter dette.
-  ROB03 Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).
 - Ja - den er ment å være aktiv på flere noder samtidig.
-  ROB04 Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på to noder samtidig).
 - Ja - den er ment å være aktiv på flere noder samtidig.
-  ROB05 Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing).
 - Ja, man skal kunne ta ut og inn noder transparent - uten at klienter påvirkes. Feks slik: http://download.oracle.com/docs/cd/E11882_01/rac.112/e16795/clonerac.htm#CHDGFHCA
-  ROB06 Det er ønskelig med mest mulig isolasjon mellom forskjellige tjenester som går på samme node i løsningen.
 - Nei, kjører som samme bruker
-  ROB07 Nodene i HA-løsningen kan stå i forskjellige nett.
 - Ja - men de må ha delt lagring og nettverksforbindelse. RAC extended clusters - http://download.oracle.com/docs/cd/E11882_01/server.112/e17157/architectures.htm#HAQVW220

Skalerbarhet

-  SCA01 Det må være mulig å melde noder inn i og ut av HA-løsningen uten at det påvirker tjenestene i løsningen (med unntak av tjenester som går på noder som skal fjernes).
 - Ja, se ROB05
-  SCA02 Det må være mulig å definere nye tjenester i løsningen uten at det krever nedetid på andre tjenester i løsningen.
 - Ja
-  SCA03 Det er ønskelig at løsningen gjør det mulig å skalere tjenester ut over kapasiteten til en enkelt node
 - Ja - man skalerer disk med å legge til disker, cpu ved å legge til noder, nettverk ved å legge til nettverkskort.
-  SCA04 HA-løsningen må støtte opp til 16 noder i løsningen.
 - Ja
-  SCA05 HA-løsningen støtter over 16 noder i løsningen.
 - Ja - men ingen dokumentasjon på dette er funnet hittil.

Ytelse og ressurser

-  PERF01 Tjenester på HA-løsningen må ikke ha nevneverdig dårligere ytelse enn om tjenesten kjørte på tilsvarende maskinvare utenfor en HA-løsning.
 - Ja - har bedre ytelse pga aktiv-aktiv cluster

Appendix 14 Mapping av RHCS mot Kravspesifikasjon

Administrasjon, drift og driftsmiljø

- ADM01 Det må være støtte "rolling upgrade" av operativsystemet på nodene i HA-løsningen.
 - Ja, dette kan gjøres innenfor samme major-versjon.^[1]
- ADM02 Nodene i HA-løsningen må kunne oppdateres fortløpende med systemoppgraderinger uten at vi mister support på HA-løsningen.
 - Ja. Såfremt det som oppdateres er støttet av RHEL så er det også støttet av RHCS.
- ADM03 Det er ønskelig at leverandøren av HA-løsningen raskt støtter nye Linux-kjerner.
 - Ja. RH er selv leverandør av nye kjerner, og RHCS er automatisk støttet på nye (offisielle) kjerner.
- ADM04 HA-løsningen må ikke være "for kompleks" å sette opp og å drifte
 - Ja. Mye av konseptet ligner på ServiceGuard's måte å gjøre det på, slik at det vil være kjent for oss.
- ADM05 Nye versjoner av HA-løsningen er mest mulig bakoverkompatibel slik at oppgraderinger ikke bryter eksisterende konfigurasjon
 - Ja, konfigurasjon er bakoverkompatibelt innenfor samme major-versjon.^[1]
- ADM06 Programvaren i HA-løsningen kan oppgraderes "rolling", dvs node for node uten at alle tjenester må tas ned
 - Ja.
- ADM07 Det er mulighet til å skripte kommunikasjonen mot HA-systemet (ikke avhengig av GUI) for å kunne automatisere driftsmiljøet rundt løsningen.
 - Ja, det finnes CLI-verktøy både for styring, konfigurering og status.
- ADM08 Tjenester i løsningen kan oppgraderes "rolling" slik at oppgraderingen ikke gir nedetid på løsningen.
 - Nei. Dette krever at cluster-løsningen støtter "aktiv-aktiv"-konfigurasjon. Det gjør ikke RHCS.

Applikasjonssupport

- APP01 Det er supportert å kjøre Oracle på HA-løsningen
 - "Ja". RH mener det ikke skal være forskjell mellom å kjøre under RHCS og på *rent* RH-OS. Må verifiseres?
- APP02 Det finnes ferdig Oracle-rammeverk tilpasset HA-løsningen
 - Ja. Men i RHCS 5 ser det ut til kun å være støtte for 10g.^[2]
- APP03 Postgres kan kjøre på løsningen
 - Ja.
- APP04 Det finnes ferdig Postgres-rammeverk tilpasset HA-løsningen
 - Ja. I RHCS 5 er det agent for PostgreSQL 8.
- APP05 Det er supportert/mulig (avhengig av hvilken MySQL løsning vi velger) å kjøre MySQL på løsningen
 - Ja.

- APP06 Det finnes ferdig MySQL-rammeverk tilpasset HA-løsningen
 - Ja.
- APP07 Det er supportert å kjøre SAP på HA-løsningen.
 - Det er mange kunder på SAP idag. Hvordan SAP selv stiller seg til support på disse løsningene er foreløpig usikkert.
- APP08 Det finnes ferdig SAP-rammeverk tilpasset HA-løsningen
 - Ja.
- APP09 IMAP kan kjøre på løsningen
 - Ja.
- APP10 Det finnes ferdig IMAP-rammeverk tilpasset HA-løsningen
 - Nei.
- APP11 HA-løsningen har støtte for fast application notification (FAN) failover tilkoplingsteknologi
 - Nei. **Basert på syensing**; dette krever at både cluster-system og applikasjoner har implementert en felles protokoll. Så vidt vi kjenner til er det ingen slik idag mellom applikasjoner fra ulike leverandører.

Failover og flytting av tjenester

- FAIL01 Ved flytting av en tjeneste fra en node til en annen i HA-løsningen er nedetiden for tjenesten omtrent det samme som tiden det tar å stoppe og starte ressursene i tjenesten på tilsvarende maskinvare utenfor HA-løsningen.
 - Ja, tror vi.^[5]
- FAIL02 Det må være mulig å flytte en tjeneste manuelt fra en node til en annen.
 - Ja.
- FAIL03 Det er ønskelig at man gjøre en planlagt flytting til en annen node uten nedetid på tjenesten (live migrering).
 - Nei, ikke støttet.^[1]
- FAIL04 Feil på maskinvare i HA-løsningen skal ikke gi langvarig nedetid for brukere av tjenesten.
 - Ja (dvs. feil på maskinvare gir **ikke** langvarig nedetid forutsatt at *fencing* er konfigurert skikkelig og fungerer som det skal).
- FAIL05 Feil på maskinvare i HA-løsningen skal ikke gi nedetid for brukere av tjenesten
 - Nei (dvs: det blir nedetid da tjenesten flyttes til annen node).
- FAIL06 Feil på maskinvare i HA-løsningen skal ikke kreve manuelle inngrep for å gjøre tjenesten tilgjengelig for brukeren.
 - Ja (krever **ikke** manuelle inngrep).^[3]
- FAIL07 Brukere av tjenesten skal ikke måtte endre i konfigurasjon på klienten etter at en tjeneste er flyttet til en annen node i HA-løsningen.
 - Ja (skal **ikke** måtte endre i konfigurasjonen).^[3]

HW- og SW-support og leverandørforhold

- SUP01 HA-løsningen skal kunne kjøre på maskinvare fra flere leverandører
 - Ja. All maskinvare støttet i RHEL er også støttet i RHCS.
- SUP02 Nodene i HA-løsningen kan være virtuelle
 - Ja, tror vi.^[6]
- SUP03 HA-løsningen skal kunne kjøre på ett eller flere operativsystemer som støttes av USIT i dag.
 - Ja. Den kjører på RHEL.
- SUP04 HA-løsningen bør kunne kjøre på de nyeste versjonene av de operativsystemer som støttes av USIT i dag.
 - Ja. RHCS kan alltid kjøres på nyeste versjon av RHEL.
- SUP05 HA løsningen må ha support fra leverandør(er) eller tredjepart(er) i alle ledd.
 - "Ja". Er noe usikre ang. applikasjonslaget, spesielt vedrørende SAP.
- SUP06 Det er ønskelig med færrest mulig forskjellige leverandører involvert i den totale løsningen
 - Ja. I alle fall samme leverandør på operativsystem som cluster-løsning.
- SUP07 Løsningen må være supportert på de versjoner av kommersiell software som er aktuelt for USIT å kjøre
 - Ja/Nei Oracle 10.2.0 og 11.2.0 databaser
 - 10.x er dokumentert støttet i RHCS 5. Usikkert om det gjelder alle minor-versjoner.
 - Ikke 11.x i RHCS 5.
 - Ja/Nei SAP (NetWeaver 7.01, NetWeaver PI 7.10, ERP 6.0, SolMan 7.01)
 - Ikke spesifisert nøyaktig i tilgjengelige lister.^[2]
- SUP08 USIT må ikke miste support på HA-løsningen dersom HA-løsningen er satt opp av USIT, eller er om konfigurasjonen er endret av USIT etter installasjon og initiell konfigurasjon.
 - Ja. Dog, dersom agenter er endret må evt. problemer kunne reproduseres med standard-agent(er) før support kan påregnes.
- SUP09 Det finnes referansekunder for HA-løsningen som kjører i en driftssituasjon som kan sammenlignes med vår egen.
 - Ja, se <https://www.usit.uio.no/prosjekter/halv2/referater/notater-webinar-redhat-2010-10-06.txt>
- SUP10 HA-løsningen skal inngå i et aktivt prosjekt eller være i aktiv utvikling hos en "solid og levedyktig" leverandør.
 - Ja.
- SUP11 Det finnes aktive epost-lister og/eller webforum for produktet.
 - Ja. Redhat selv hoster bl.a. en aktiv mailingliste "*Linux-cluster*".
- SUP12 HA-løsningen er open source.
 - Ja.
- SUP13 HA-løsningen er dokumentert.
 - Ja.
- SUP14 HA-løsningen er godt dokumentert.
 - Ja.

Konfigurasjon av tjenester

- CONF01 Det må være mulig å konfigurere hvilke ressurser som skal være knyttet til en tjeneste
 - Ja.
- CONF02 Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.
 - Ja, i det minste ved bruk av **skript-ressursen**.^[5]
- CONF03 Det oppleves som enkelt å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.
 - Vanskelig å si uten egen erfaring.
- CONF04 Det må være mulig å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.
 - Ja.
- CONF05 Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned.
 - Ja, tror vi.^[5]
- CONF06 Det oppleves som enkelt å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.
 - Vanskelig å si uten egen erfaring.
- CONF07 Det må være mulig å spesifisere hvilke noder en tjeneste skal kunne gå på, og rekkefølgen for hvordan tjenesten skal gjøre failover til andre noder.
 - Ja. Ikke like avanserte muligheter som i ServiceGuard, men noder kan ha prioritet slik at det det er mulig å styre hvilke noder tjenesten feiler over til.
- CONF08 Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder
 - "Ja". Dette må i så fall skje i agenten.^[4]
- CONF09 HA-løsningen kan konfigureres slik at den restarter ressurser i en tjeneste som den har detektert at ikke lenge kjører.
 - Ja. Se f.eks. <http://sources.redhat.com/cluster/wiki/ServicePolicies>

Lagring og filsystem

- STOR01 Løsningen må kunne benytte seg av delt lagring, enten via et SAN (FC eller iSCSI) eller nettverksfilssystem (NFS).
 - Ja.
- STOR02 Det er ønskelig at løsningen kan støtte et cluster-filsystem som tillater samtidig skriving fra flere noder.
 - Ja. RH tilbyr eget cluster-filsystem: *GFS*.
- STOR03 Løsningen bør være supportert på den lagrings-infrastruktur som USIT allerede støtter (SAN, fc-switcher, lagringsprodukter) så vi slipper å kjøpe inn egen lagring til løsningen.
 - Ja. Alt som støttes av RHEL støttes også på RHCS.

Overvåkning og feilsøk

- MON01 Det må enkelt å få ut info om status for HA-løsningen, dets noder og dets tjenester.
 - Ja. Enten via *Conga* (GUI) eller *clustat* (CLI).
- MON02 Logging må være tilstrekkelig slik at det er mulig å feilsøke ved eventuelle problemer.
 - Ja, tror vi.^[5]
- MON03 Det er ønskelig at HA-løsningen kan sende sine logger fra alle noder til et sentralisert sted
 - Ja. Kan logge via *syslog*.
- MON04 HA-løsningen må la seg overvåke av Nagios
 - Ja. Ingen offisielt støttede metoder/plugins fra RedHat, men det bør være enkelt å implementere:
 - Det finnes flere tredjeparts-plugins til Nagios f.eks. på [Nagios Exchange](#) (kategori "*Clustering and High-Availability*")
 - RHCS inkluderer (enkel) SNMP-agent med egen MIB
 - Det bør være enkelt å skrive egne plugins som f.eks. kommuniserer via *clustat*
- MON05 Oracle-databaser som kjører i HA-løsningen må la seg overvåke av GridControl.
 - Ja, tror vi.^[5] eller ^[6] Det bør fungere som ordinær Oracle-instans.
- MON06 Det må være mulig å få generert en alarm dersom en ressurs i tjenesten ikke kjører
 - Ja. Det er mulig å definere monitor-skript, hvilke kan settes opp til å generere alarmer slik man selv vil.

Nettverk

- NET01 HA-løsningen støtter IPv6
 - Nei, ikke i RHCS 5. Det er i [Release Notes](#) for versjon 6 bemerket "**Enhanced support for Internet Protocol version 6 (IPv6)**"

Priser og lisenser

- LIC01 Kostnaden for løsningen ved anskaffelse og drift bør ikke være vesentlig høyere enn for dagens SG-løsninger, også for eksterne kunder som ikke er omfattet av UiO sin site-lisens.
 - Ukjent. Priser foreligger ennå ikke for RHCS 6, som troligvis er den som er aktuell for UiO.

Redundanse og robusthet

- ROB01 HA-løsningen skal være en "moden" teknologi idet det tas i bruk av USIT
 - Ja. Ikke like moden som ServiceGuard, men produktet har vært tilgjengelig og under utvikling i lengre tid.
- ROB02 Løsningen må kunne settes opp og være supportert med multiple paths til både nettverk og lagring for å unngå single-point-of-failure.
 - Ja. Støttes det av RHEL støttes det av RHCS.
- ROB03 Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).
 - Ja. Dog kreves det egen erfaring før det er mulig å si noe om hvor pålitelige disse mekanismene er.

- ROB04 Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på to noder samtidig).
 - Ja. Samme kommentar som ovenfor.
- ROB05 Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing).
 - Ja. Samme kommentar som ovenfor.
- ROB06 Det er ønskelig med mest mulig isolasjon mellom forskjellige tjenester som går på samme node i løsningen.
 - Nei. Det er ikke fullstendig isolasjon som f.eks. i virtuelle maskinvare-domener eller "*jails*". Dersom en tjeneste har problemer kan det påvirke andre.
- ROB07 Nodene i HA-løsningen kan stå i forskjellige nett.
 - Ja. Men det anbefales fra leverandør sterkt å ikke ha dette, da det kan bli problematisk i forhold til bl.a. fencing.

Skalerbarhet

- SCA01 Det må være mulig å melde noder inn i og ut av HA-løsningen uten at det påvirker tjenestene i løsningen (med unntak av tjenester som går på noder som skal fjernes).
 - Ja. Noder kan legges til live.
- SCA02 Det må være mulig å definere nye tjenester i løsningen uten at det krever nedetid på andre tjenester i løsningen.
 - Ja.
- SCA03 Det er ønskelig at løsningen gjør det mulig å skalere tjenester ut over kapasiteten til en enkelt node
 - Nei. Ikke noe i RHCS som tilrettelegger for dette.
- SCA04 HA-løsningen må støtte opp til 16 noder i løsningen.
 - Ja. RHCS støtter opp til 16 noder.
- SCA05 HA-løsningen støtter over 16 noder i løsningen.
 - Nei. RHCS støtter maksimalt 16 noder.

Ytelse og ressurser

- PERF01 Tjenester på HA-løsningen må ikke ha nevneverdig dårligere ytelse enn om tjenesten kjørte på tilsvarende maskinvare utenfor en HA-løsning.
 - Ja. Dvs. i teorien skal det ikke være merkbart dårligere ytelse enn om tjenesten kjører på "*rent*" OS. Dette forutsetter dog at evt. monitor-skript m.v. ikke utfører operasjoner som påvirker ytelsen i særlig grad.

[1] Bekreftet av RH-representanter på Webinar 7/10 2010

[2] RedHat-personell skal sende liste over offisielt støttede agenter (*rammeverk*). Ennå ikke mottatt 13/10 2010.

[3] Selvsagt forutsatt at tjenesten er designet fornuftig.

[4] Uttalt av representanter for Redhat på Webinar i forbindelse med spørsmål rundt SAP og RHCS.

[5] Må ha egen erfaring for å bekrefte.

[6] Må verifiseres av leverandør.

Appendix 15 Mapping av SteelEye LifeKeeper mot kravspesifikasjon

Administrasjon, drift og driftsmiljø

* ADM01 Det må være støtte "rolling upgrade" av operativsystemet på nodene i HA-løsningen.

Ja

MS: Hvis I med ”rolling upgrade” mener fx yum update, eller lign. så ja, SteelEye støtter dette.

* ADM02 Nodene i HA-løsningen må kunne oppdateres fortløpende med systemoppgraderinger uten at vi mister support på HA-løsningen.

Ja

MS: SteelEye støtter dette.

* ADM03 Det er ønskelig at leverandøren av HA-løsningen raskt støtter nye Linux-kjerner.

Ja - minor releases

MS: SteelEye er ikke avhengig af specifik linuxkerne versjon. Der er dog nogle enkelte moduler der kræver dette. Fx NFS, MultiPath hvor man skal kompilere kernemodul.

* ADM04 HA-løsningen må ikke være "for kompleks" å sette opp og å drifte

Ja

MS: SteelEye er MEGET let å sette opp og drifte. Jeg vil på torsdag vise daglig drift av MySQL HA – hvis vi kan nå det kan jeg evt. også vise installasjon/konfigurasjon av løsningen.

* ADM05 Nye versjoner av HA-løsningen er mest mulig bakoverkompatibel slik at oppgraderinger ikke bryter eksisterende konfigurasjon

Ja – men bare mens man oppgraderer http://www.steeleye.com/downloads/Support/Documentation/Linux/7_2/LKInstall.pdf

MS: ALLE SteelEye-løsninger er bakoverkompatibel.

* ADM06 Programvaren i HA-løsningen kan oppgraderes "rolling", dvs node for node uten at alle tjenester må tas ned

Ja

MS: Ved oppgradering av SteelEye vil man typisk lave failover av tjenester til secondary node, og så oppgradere primary. Derefter failover tilbake til primary og derefter oppgradere secondary. SteelEye kan settes til automatisk at gjøre dette.

* ADM07 Det er mulighet til å skripte kommunikasjonen mot HA-systemet (ikke avhengig av GUI) for å kunne automatisere driftsmiljøet rundt løsningen.

Ja

MS: Alt hvad man kan i GUI, kan gøres fra CLI/scriptes – og mere til ?

* ADM08 Tjenester i løsningen kan oppgraderes "rolling" slik at oppgraderingen ikke gir nedetid på løsningen.

Ja

MS: SteelEye kan også håndtere dette.. og det er dokumenteret.

Applikasjonssupport

* APP01 Det er supportert å kjøre Oracle på HA-løsningen

Ja

MS: Ja, SteelEye er Oracle-partner og støtter Oracle v9-v11

* APP02 Det finnes ferdig Oracle-rammeverk tilpasset HA-løsningen

Ja

MS: JA

* APP03 Postgres kan kjøre på løsningen

Ja

MS: JA – både PostgreSQL og Postgres Plus (EnterpriseDB)

* APP04 Det finnes ferdig Postgres-rammeverk tilpasset HA-løsningen

Ja

MS: JA

* APP05 Det er supportert/mulig (avhengig av hvilken MySQL løsning vi velger) å kjøre MySQL på løsningen

Ja

MS: JA

* APP06 Det finnes ferdig MySQL-rammeverk tilpasset HA-løsningen

Ja

MS: JA

* APP07 Det er supportert å kjøre SAP på HA-løsningen.

Ja

MS: JA. SteelEye var den første kommersielle HA-løsningen for SAP.

* APP08 Det finnes ferdig SAP-rammeverk tilpasset HA-løsningen

Ja

MS: JA

* APP09 IMAP kan kjøre på løsningen

Ja

MS: JA

* APP10 Det finnes ferdig IMAP-rammeverk tilpasset HA-løsningen

Nei

MS: Understøttes via scripts

* APP11 HA-løsningen har støtte for fast application notification (FAN) failover tilkoplingsteknologi

Nei

MS: Nej. LifeKeeper benytter ikke FAN, da dette er en RAC feature. LifeKeeper bruker API til å monitorere Oracle, samt OS, filsystemer og ip-adresser.

Failover og flytting av tjenester

* FAIL01 Ved flytting av en tjeneste fra en node til en annen i HA-løsningen er nedetiden for tjenesten omtrent det samme som tiden det tar å stoppe og starte ressursene i tjenesten på tilsvarende maskinvare utenfor HA-løsningen.

Ja

MS: JA. Avhenger selvfølgelig af netverkk og om man kjører datareplikering (indbygget i SteelEye) mellem noderne.

* FAIL02 Det må være mulig å flytte en tjeneste manuelt fra en node til en annen.

Ja

MS: JA, Ganske simpelt.

* FAIL03 Det er ønskelig at man gjøre en planlagt flytting til en annen node uten nedetid på tjenesten (live migrering).

Nei

MS: Det er ikke muligt å lave live migrering på applikasjonsniveau. Hvis man kører virtuelt, støtter SteelEye dog livemigrering av virtuelle maskiner.

* FAIL04 Feil på maskinvare i HA-løsningen skal ikke gi langvarig nedetid for brukere av tjenesten.

Ja

MS: Man kan med SteelEye finjustere connection timeout-værdier, og derved minimere nedetid ved hardware feil.

* FAIL05 Feil på maskinvare i HA-løsningen skal ikke gi nedetid for brukere av tjenesten

Nei

MS: Se tidligere svar.

* FAIL06 Feil på maskinvare i HA-løsningen skal ikke kreve manuelle inngrep for å gjøre tjenesten tilgjengelig for brukeren.

Ja

MS: SteelEye støtter automatisk failover (+ evt. fallback)

* FAIL07 Brukere av tjenesten skal ikke måtte endre i konfigurasjon på klienten etter at en tjeneste er flyttet til en annen node i HA-løsningen.

Ja

MS: SteelEye støtter dette.

HW- og SW-support og leverandørforhold

* SUP01 HA-løsningen skal kunne kjøre på maskinvare fra flere leverandører

Ja

MS: JA

* SUP02 Nodene i HA-løsningen kan være virtuelle

Ja

MS: Ja, SteelEye støtter alle scenarier – fx P2P, P2V, V2P, V2V failover (P = Physical, V = Virtual)

* SUP03 HA-løsningen skal kunne kjøre på ett eller flere operativsystemer som støttes av USIT i dag.

Ja

MS: SteelEye støtter: CentOS, SLES, OEL og RHEL

* SUP04 HA-løsningen bør kunne kjøre på de nyeste versjonene av de operativsystemer som støttes av USIT i dag.

Ja

MS: SteelEye arbeider tætt sammen med fx Red Hat og Microsoft om å støtte kommende versjoner av operativsystemerne. Dvs. SteelEye også støtter RHEL6 når denne kommer.

* SUP05 HA løsningen må ha support fra leverandør(er) eller tredjepart(er) i alle ledd.

Ja

MS: Der er fuld support både fra SteelEye (US/UK) og Nordicmind (DK/FIN)

* SUP06 Det er ønskelig med færrest mulig forskjellige leverandører involvert i den totale løsningen

Delvis

MS: SteelEye og Nordicmind vil være de eneste der er involvert.

* SUP07 Løsningen må være supportert på de versjoner av kommersiell software som er aktuelt for USIT å kjøre

o Ja/Nei Oracle 10.2.0 og 11.2.0 databaser

o Ja/Nei SAP (NetWeaver 7.01, NetWeaver PI 7.10, ERP 6.0, SolMan 7.01)

Ja

MS: SteelEye støtter de nevnte programvarer/versjoner.

* SUP08 USIT må ikke miste support på HA-løsningen dersom HA-løsningen er satt opp av USIT, eller er om konfigurasjonen er endret av USIT etter installasjon og initiell konfigurasjon.

Ja

MS: SteelEye er utviklet med enkelt brugerflade som mål, så kunderne ofte selv kan installere/konfigurere uden hjelp fra leverandør/partner. SteelEye (og Nordicmind) støtter 100% dette.

* SUP09 Det finnes referanse kunder for HA-løsningen som kjører i en driftssituasjon som kan sammenlignes med vår egen.

Ja - University of Texas

MS: Vi har både kunder der kører SAP og Oracle – og universiteter (dog ikke i Norden) – og kunder med mange nodes/cluster (500+). Kunder spænder vidt inden for bl.a. televirksomheder, tv, olieselskaper, offentlige institutioner og andre store private kunder.

* SUP10 HA-løsningen skal inngå i et aktivt prosjekt eller være i aktiv utvikling hos en "solid og levedyktig" leverandør.

Ja

MS: Programvaren er utviklet av utviklingsteam hos ATT/NCR i 1991, og firmaet SteelEye blev dannet i 1999. Nordicmind har eksisteret siden 2002, er i dag SteelEye’s Support og Competence Center for Norden. Vi har kontorer i Finland og Danmark, og har forhandlere og partnere i over 15 forskjellige lande i Nordeuropa og CIS.

* SUP11 Det finnes aktive epost-lister og/eller webforum for produktet.

Nei

MS: Nej, men der er hurtig support.

* SUP12 HA-løsningen er open source.

Nei

MS: Nej, løsningen er dog meget fleksibel, så man kan skræddersy til ens behov.

* SUP13 HA-løsningen er dokumentert.

Ja

MS: JA

* SUP14 HA-løsningen er godt dokumentert.

Ja

MS: JA

Konfigurasjon av tjenester

* CONF01 Det må være mulig å konfigurere hvilke ressurser som skal være knyttet til en tjeneste

Ja

MS: Hvis dere med ”ressourser” mener fx SAN, LVM, ip-adresse, osv. – så JA.

* CONF02 Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.

Ja - via resource hierarchies
http://www.steeleye.com/downloads/Support/Documentation/Linux/7_1/OnlineProductManual/productmanual/maintaining_resource_hierarchy.htm

* CONF03 Det oppleves som enkelt å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.

Mja – Se CONF 02

* CONF04 Det må være mulig å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.

Ja

MS: Dette er muligt. Det er også muligt å lave speciel monitoreringsscripts, som sjekker tjenesten på en speciel måde, hvis dette ønskes.

* CONF05 Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned.

Ja

MS: Alle indstillinger er globale – on-the-fly. Krever ikke genstart av tjenste eller lignende.

* CONF06 Det oppleves som enkelt å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.

Ja

MS: Kan gøres fra GUI – eller man kan lave scripts (perl, bash, etc.)

* CONF07 Det må være mulig å spesifisere hvilke noder en tjeneste skal kunne gå på, og rekkefølgen for hvordan tjenesten skal gjøre failover til andre noder.

Ja

MS: Hver node/tejeneste kan prioriteres, så man kan vælge hvor tjenesten gjøre failover til andre noder.

* CONF08 Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder

Ja – se CONF02

MS: JA

* CONF09 HA-løsningen kan konfigureres slik at den restarter ressurser i en tjeneste som den har detektert at ikke lenge kjører.

Ja

MS: JA, SteelEye understøtter ”local recovery”, hvor en ressource/tjeneste genstartes hvis den ikke kjøres lenger – eller der er sket en feil så den ikke kjører. Da vil SteelEye udbedre dette og genstarte ressourcen/tjenesten.

Lagring og filsystem

* STOR01 Løsningen må kunne benytte seg av delt lagring, enten via et SAN (FC eller iSCSI) eller nettverksfilssystem (NFS).

Ja

MS: SteelEye støtter SANs, NFS/NAS eller datareplikering (indbygget i løsningen)

* STOR02 Det er ønskelig at løsningen kan støtte et cluster-filssystem som tillater samtidig skriving fra flere noder.

Nei

MS: LifeKeeper støtter ikke GFS. da LifeKeeper sikrer at kun én server/applikasjon kan skrive til disk for å undgå bl.a. inkonsistens.

* STOR03 Løsningen bør være supportert på den lagrings-infrastruktur som USIT allerede støtter (SAN, fc-switcher, lagringsprodukter) så vi slipper å kjøpe inn egen lagring til løsningen.

Ja

MS: Generelt støtter SteelEye storageprodukter hvis operativsystemet støtter dem. Se evt. en liste her:

http://www.steeleye.com/Linux_Product_Support_Matrix_86.htm

Overvåkning og feilsøk

* MON01 Det må enkelt å få ut info om status for HA-løsningen, dets noder og dets tjenester.

Ja

MS: Det kan man via GUI, scripts og logfiler

* MON02 Logging må være tilstrekkelig slik at det er mulig å feilsøke ved eventuelle problemer.

Ja – det ble skrevet ut info ved failover demonstrasjonen vi så på.

MS: Der er logging på flere forskjellige debug-levels

* MON03 Det er ønskelig at HA-løsningen kan sende sine logger fra alle noder til et sentralisert sted

Ja

MS: Man kan fra GUI se LifeKeeper-logs fra alle noder.

* MON04 HA-løsningen må la seg overvåke av Nagios

Ja - antagelig

MS: Nagios er rimelig fleksibel, så det burde kunne lade sig gjøre.

* MON05 Oracle-databaser som kjører i HA-løsningen må la seg overvåke av GridControl.

Ja

MS: LifeKeeper overvåker selv Oracle, så der er ikke krav om GridControl. Men da LifeKeeper ikke kræver special Oracle-konfiguration for å lave failover osv. da burde dere kunne bruke GridControl samtidig.

* MON06 Det må være mulig å få generert en alarm dersom en ressurs i tjenesten ikke kjører

Ja

MS: Det er muligt å konfigurere alarmer på forskjellige niveauer – fx mail, sms, m.m.

Nettverk

* NET01 HA-løsningen støtter IPv6

Nei

MS: Nej – men dette skulle komme i starten av 2011.

Priser og lisenser

* LIC01 Kostnaden for løsningen ved anskaffelse og drift bør ikke være vesentlig høyere enn for dagens SG-løsninger, også for eksterne kunder som ikke er omfattet av UiO sin site-lisens.

mulig

MS: Jeg er sikker på at vi kan følge med.

Redundanse og robusthet

* ROB01 HA-løsningen skal være en "moden" teknologi idet det tas i bruk av USIT

Ja

MS: Løsningen blev oprindelig udviklet av ATT/NCR i starten av 90'erne til å lave failover på telefonsentraler i USA. Så løsningen er gjennomtestet og afprøvet.

* ROB02 Løsningen må kunne settes opp og være supportert med multiple paths til både nettverk og lagring for å unngå single-point-of-failure.

Ja

MS: Vi anbefaler alltid multipath.

* ROB03 Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).

Ja, men må testes. Har også mulighet for Quorum – se side 38 http://www.steeleye.com/downloads/Support/Documentation/Linux/7_2/LKInstall.pdf

MS: SteelEye er en ressource-driven cluster løsning. Dvs den bruker ikke quorum device. Split-brain er et kjent fænomen i et cluster, og SteelEye har pålitelige mekanismer for dette.

* ROB04 Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på

to noder samtidig).

Ja

MS: En tjeneste/ressource kan kun være aktiv på én node med SteelEye (Active/Passive). Det er dog muligt å køre Active-Active løsninger også.

* ROB05 Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing).

Ja

MS: Nodes udmeldes meget let fra clusteret via GUI/CLI/Script.

* ROB06 Det er ønskelig med mest mulig isolasjon mellom forskjellige tjenester som går på samme node i løsningen.

Nei - root som kjører selve pakkene, men i pakkene kan applikasjoner forskjellige brukere kjøre.

MS: Forklar nærmere..

* ROB07 Nodene i HA-løsningen kan stå i forskjellige nett.

Ja

MS: Ja, SteelEye støtter både LAN og WAN (ikke samme subnet)

Skalerbarhet

* SCA01 Det må være mulig å melde noder inn i og ut av HA-løsningen uten at det påvirker tjenestene i løsningen (med unntak av tjenester som går på noder som skal fjernes).

Ja

MS: Nodes indmeldes meget let fra clusteret via GUI/CLI/Script.

* SCA02 Det må være mulig å definere nye tjenester i løsningen uten at det krever nedetid på andre tjenester i løsningen.

Ja

MS: Nodes udmeldes meget let fra clusteret via GUI/CLI/Script.

* SCA03 Det er ønskelig at løsningen gjør det mulig å skalere tjenester ut over kapasiteten til en enkelt node

Nei, men har replikering innebygd

* SCA04 HA-løsningen må støtte opp til 16 noder i løsningen.

Ja

MS: JA

* SCA05 HA-løsningen støtter over 16 noder i løsningen.

Ja

MS: JA, Op til 32 noder

Ytelse og ressurser

* PERF01 Tjenester på HA-løsningen må ikke ha nevneverdig dårligere ytelse enn om tjenesten kjørte på tilsvarende maskinvare utenfor en HA-løsning.

Ja

MS: SteelEye giver ikke dårligere ytelse (avhænger av konfigurasjon).

Appendix 16 Mapping av Veritas Cluster Server (VCS) mot kravspesifikasjon.

Administrasjon, drift og driftsmiljø

- ADM01 Det må være støtte "rolling upgrade" av operativsystemet på nodene i HA-løsningen.
 - o Ja. Man failer over pakker til andre noder mens OS på en node oppgraderes.
- ADM02 Nodene i HA-løsningen må kunne oppdateres fortløpende med systemoppgraderinger uten at vi mister support på HA-løsningen.
 - o Ja. Gitt at minimumskrav fra VCS er tilfredstilt, så er oppgradering innen samme major-versjon støttet.
 - o Ref. http://sfdoccentral.symantec.com/sf/5.1/linux/pdf/vcs_notes.pdf
- ADM03 Det er ønskelig at leverandøren av HA-løsningen raskt støtter nye Linux-kjerner.
 - o Ja. Gitt at minimumskrav fra VCS er tilfredstilt, så er oppgradering innen samme major-versjon støttet.
 - o Ref. http://sfdoccentral.symantec.com/sf/5.1/linux/pdf/vcs_notes.pdf
- ADM04 HA-løsningen må ikke være "for kompleks" å sette opp og å drifte
 - o Ja. Ut fra dokumentasjonen ser det greit ut. Det ser også ut til å ligne mye på tankegangen fra ServiceGuard, og det antas derfor å være enkelt for oss å sette seg inn i.
- ADM05 Nye versjoner av HA-løsningen er mest mulig bakoverkompatibel slik at oppgraderinger ikke bryter eksisterende konfigurasjon
 - o Ja. Det hender at nye versjoner innebærer oppdatering av konfig, men det håndteres da som en del av oppgraderingen. Hvis det ikke er påkrevd med ny konfigurasjon, så er det opp til oss om vi ønsker det for å få med ny funksjonalitet.
- ADM06 Programvaren i HA-løsningen kan oppgraderes "rolling", dvs node for node uten at alle tjenester må tas ned
 - o Ja. Man failer over pakker til andre noder mens en node oppgraderes. Trenger ikke reboot node for å oppgradere VCS (mellom visse versjoner).
 - o Ref. http://sfdoccentral.symantec.com/sf/5.1/linux/pdf/vcs_install.pdf
- ADM07 Det er mulighet til å skripte kommunikasjonen mot HA-systemet (ikke avhengig av GUI) for å kunne automatisere driftsmiljøet rundt løsningen.
 - o Ja. Har både GUI og CLI.
- ADM08 Tjenester i løsningen kan oppgraderes "rolling" slik at oppgraderingen ikke gir nedetid på løsningen.

o Nei, tjenesten generelt kun kjører på en node.

Applikasjonssupport

- APP01 Det er supportert å kjøre Oracle på HA-løsningen
 - o Ja.
- APP02 Det finnes ferdig Oracle-rammeverk tilpasset HA-løsningen
 - o Ja.
 - o Ref. ftp://ftp.entsupport.symantec.com/pub/support/documentation/vcs_oracle_agent_51_lin.pdf
- APP03 Postgres kan kjøre på løsningen
 - o Ja.
 - o Ref. <http://www.postgresql.org/about/press/faq>
- APP04 Det finnes ferdig Postgres-rammeverk tilpasset HA-løsningen
 - o Nei. Men kan lage egne agenter fra en generell application agent.
- APP05 Det er supportert/mulig (avhengig av hvilken MySQL løsning vi velger) å kjøre MySQL på løsningen
 - o Ja.
- APP06 Det finnes ferdig MySQL-rammeverk tilpasset HA-løsningen
 - o Ja.
 - o Ref. <https://sort.symantec.com/agents/detail/3108>
- APP07 Det er supportert å kjøre SAP på HA-løsningen.
 - o Ja.
- APP08 Det finnes ferdig SAP-rammeverk tilpasset HA-løsningen
 - o Ja. Er dette et eget produkt?
 - o Ref. <https://sort.symantec.com/agents>
- APP09 IMAP kan kjøre på løsningen
 - o Ja.
- APP10 Det finnes ferdig IMAP-rammeverk tilpasset HA-løsningen

- o Nei. Må lage selv.
- APP11 HA-løsningen har støtte for fast application notification (FAN) failover tilkoplingsteknologi
 - o Nei i utgangspunktet ikke, men det er mulig å lage selv.

Failover og flytting av tjenester

- FAIL01 Ved flytting av en tjeneste fra en node til en annen i HA-løsningen er nedetiden for tjenesten omtrent det samme som tiden det tar å stoppe og starte ressursene i tjenesten på tilsvarende maskinvare utenfor HA-løsningen.
 - o Ja. Antar dette er omtrent som for ServiceGuard, men vi vet ikke sikkert før vi har prøvd.
- FAIL02 Det må være mulig å flytte en tjeneste manuelt fra en node til en annen.
 - o Ja. Med GUI eller CLI.
- FAIL03 Det er ønskelig at man gjøre en planlagt flytting til en annen node uten nedetid på tjenesten (live migrering).
 - o Nei.
- FAIL04 Feil på maskinvare i HA-løsningen skal ikke gi langvarig nedetid for brukere av tjenesten.
 - o Ja. Antar dette er omtrent som for ServiceGuard, men vi vet ikke sikkert før vi har prøvd.
- FAIL05 Feil på maskinvare i HA-løsningen skal ikke gi nedetid for brukere av tjenesten
 - o Nei. Det blir noe nedetid for restart. Antar dette er omtrent som for ServiceGuard, men vi vet ikke sikkert før vi har prøvd.
- FAIL06 Feil på maskinvare i HA-løsningen skal ikke kreve manuelle inngrep for å gjøre tjenesten tilgjengelig for brukeren.
 - o Ja. Failover skjer automatisk.
- FAIL07 Brukere av tjenesten skal ikke måtte endre i konfigurasjon på klienten etter at en tjeneste er flyttet til en annen node i HA-løsningen.
 - o Ja.

HW- og SW-support og leverandørforhold

- SUP01 HA-løsningen skal kunne kjøre på maskinvare fra flere leverandører
 - o Ja.
- SUP02 Nodene i HA-løsningen kan være virtuelle
 - o Ja.
 - o Ref. http://eval.symantec.com/mktginfo/enterprise/other_resources/ent-vcs_5_for_vmware_faq.en-us.pdf

- SUP03 HA-løsningen skal kunne kjøre på ett eller flere operativsystemer som støttes av USIT i dag.
 - o Ja. RHEL-5 er supportet.
 - o Ref. http://sfdoccentral.symantec.com/sf/5.1/linux/pdf/vcs_notes.pdf
- SUP04 HA-løsningen bør kunne kjøre på de nyeste versjonene av de operativsystemer som støttes av USIT i dag.
 - o Ja.
 - o Ref. http://sfdoccentral.symantec.com/sf/5.1/linux/pdf/vcs_notes.pdf
- SUP05 HA løsningen må ha support fra leverandør(er) eller tredjepart(er) i alle ledd.
 - o Ja. Kan få support fra Symantec eller tredjepart.
- SUP06 Det er ønskelig med færrest mulig forskjellige leverandører involvert i den totale løsningen
 - o Nei. VCS kommer fra en leverandør som ikke leverer andre deler av HW/OS/SW, så det blir en ny leverandør.
- SUP07 Løsningen må være supportert på de versjoner av kommersiell software som er aktuelt for USIT å kjøre
 - o Ja, støtter Oracle 10g og 11gR2
 - o Ref. <http://www.symantec.com/connect/sfha51>
 - o Ja støtter SAP (ERP og Netweaver på "våre" versjoner o Ref. <https://sort.symantec.com/agents>
- SUP08 USIT må ikke miste support på HA-løsningen dersom HA-løsningen er satt opp av USIT, eller er om konfigurasjonen er endret av USIT etter installasjon og initiell konfigurasjon.
 - o Ja. Verifisert av Juvkam hos Symantec 2010-10-19
- SUP09 Det finnes referansekunder for HA-løsningen som kjører i en driftssituasjon som kan sammenlignes med vår egen.
 - o Ja. Store kunder i Norge som kjører VCS på RHEL, og også kunder som har migrert fra SG til VCS. Vi har fått navn.
- SUP10 HA-løsningen skal inngå i et aktivt prosjekt eller være i aktiv utvikling hos en "solid og levedyktig" leverandør.
 - o Ja. Stadig utvikling, aktiv leverandør.
- SUP11 Det finnes aktive epost-lister og/eller webforum for produktet.
 - o Ja. <http://www.symantec.com/connect/> og <http://mailman.eng.auburn.edu/mailman/listinfo/veritas-ha>
- SUP12 HA-løsningen er open source.
 - o Nei.
- SUP13 HA-løsningen er dokumentert.

o Ja.

- SUP14 HA-løsningen er godt dokumentert.

o Ja. Veldig imponerende dokumentasjon.

Konfigurasjon av tjenester

- CONF01 Det må være mulig å konfigurere hvilke ressurser som skal være knyttet til en tjeneste

o Ja.

- CONF02 Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.

o Ja, kan lage egne betingelser for monitorering i agenter.

- CONF03 Det oppleves som enkelt å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.

o Tror det, men vi vet ikke sikkert før vi har forsøkt.

- CONF04 Det må være mulig å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.

o Ja. Gjøres enklest ved monitorering av prosessen i servicegruppen.

- CONF05 Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned.

o Ja. Ref. mail fra Sveum hos Symantec.

- CONF06 Det oppleves som enkelt å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.

o Uvisst. Vet ikke før vi har forsøkt.

- CONF07 Det må være mulig å spesifisere hvilke noder en tjeneste skal kunne gå på, og rekkefølgen for hvordan tjenesten skal gjøre failover til andre noder.

o Ja.

o Ref. [ftp://ftp.entsupport.symantec.com/pub/support/documentation/vcs_admin_51_lin.pdf](http://ftp.entsupport.symantec.com/pub/support/documentation/vcs_admin_51_lin.pdf)

- CONF08 Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder

o Ja.

o Ref. [ftp://ftp.entsupport.symantec.com/pub/support/documentation/vcs_admin_51_lin.pdf](http://ftp.entsupport.symantec.com/pub/support/documentation/vcs_admin_51_lin.pdf) (kapittel 12)

- CONF09 HA-løsningen kan konfigureres slik at den restarter ressurser i en tjeneste som den har detektert at ikke lenger kjører.

o Ja. Ref. mail fra Sveum hos Symantec.

Lagring og filsystem

- STOR01 Løsningen må kunne benytte seg av delt lagring, enten via et SAN (FC eller iSCSI) eller nettverksfilssystem (NFS).
 - o Ja.
- STOR02 Det er ønskelig at løsningen kan støtte et cluster-filsystem som tillater samtidig skriving fra flere noder.
 - o Ja. Veritas Cluster File System.
- STOR03 Løsningen bør være supportert på den lagrings-infrastruktur som USIT allerede støtter (SAN, fc-switcher, lagringsprodukter) så vi slipper å kjøpe inn egen lagring til løsningen.
 - o Ja, EVA8k er støttet.
 - o Ref. http://www.symantec.com/business/support/resources/sites/BUSINESS/content/live/TECHNICAL_SOLUTION/74000/TECH74012/en_US/330441.pdf

Overvåkning og feilsøk

- MON01 Det må enkelt å få ut info om status for HA-løsningen, dets noder og dets tjenester.
 - o Ja. "hasys" CLI eller Java GUI.
- MON02 Logging må være tilstrekkelig slik at det er mulig å feilsøke ved eventuelle problemer.
 - o Ja.
- MON03 Det er ønskelig at HA-løsningen kan sende sine logger fra alle noder til et sentralisert sted
 - o Ja. Både med syslog og Veritas Operation Manager konsollet.
- MON04 HA-løsningen må la seg overvåke av Nagios
 - o Trolig, støtter SNMP2 trap (men ikke get). Finnes ikke ferdig Nagios-plugin, men trolig enkelt å lage selv.
- MON05 Oracle-databaser som kjører i HA-løsningen må la seg overvåke av GridControl.
 - o Ja.
 - o Ref. <http://www.oracle.com/technetwork/oem/extensions/index-100685.html>
- MON06 Det må være mulig å få generert en alarm dersom en ressurs i tjenesten ikke kjører
 - o Ja. Enten til logger, eller snmp/smtp fra Notify Agenten og Veritas Operation Manager konsollet.

Nettverk

- NET01 HA-løsningen støtter IPv6
 - o Ja, full støtte fra versjon 5.1.

Priser og lisenser

- LIC01 Kostnaden for løsningen ved anskaffelse og drift bør ikke være vesentlig høyere enn for dagens SG-løsninger, også for eksterne kunder som ikke er omfattet av UiO sin site-lisens.
 - o Uvisst. Gunnar svarer når han har fått pris

Redundanse og robusthet

- ROB01 HA-løsningen skal være en "moden" teknologi idet det tas i bruk av USIT
 - o Ja. VCS har vært på markedet lenge, også for Linux.
- ROB02 Løsningen må kunne settes opp og være supportert med multiple paths til både nettverk og lagring for å unngå single-point-of-failure.
 - o Ja. Man kan bruke vanlig Linux multipath eller "Dynamic Multipathing" fra produktet Veritas Storage Foundation. Ref email fra Juvkam 21/10-10
- ROB03 Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).
 - o Ja. SCSI3 persistent reservation samt Coordination Point Server.
 - o Ref. ftp://ftp.entsupport.symantec.com/pub/support/documentation/vcs_admin_51_lin.pdf (Kap 9 og 10).
- ROB04 Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på to noder samtidig).
 - o Ja. SCSI3 persistent reservation samt Coordination Point Server.
 - o Ref. ftp://ftp.entsupport.symantec.com/pub/support/documentation/vcs_admin_51_lin.pdf (Kap 9 og 10).
- ROB05 Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing).
 - o Ja. Bruker i utgangspunktet SCSI3 Persistent Reservation eller Coordination Point Server. Det er mulig å sette opp power fencing, men det er ikke vanlig i VCS.
- ROB06 Det er ønskelig med mest mulig isolasjon mellom forskjellige tjenester som går på samme node i løsningen.
 - o Nei.

- ROB07 Nodene i HA-løsningen kan stå i forskjellige nett.

o Ja. Ref. mail fra Sveum hos Symantec.

Skalerbarhet

- SCA01 Det må være mulig å melde noder inn i og ut av HA-løsningen uten at det påvirker tjenestene i løsningen (med unntak av tjenester som går på noder som skal fjernes).

o Ja.

- SCA02 Det må være mulig å definere nye tjenester i løsningen uten at det krever nedetid på andre tjenester i løsningen.

o Ja.

- SCA03 Det er ønskelig at løsningen gjør det mulig å skalere tjenester ut over kapasiteten til en enkelt node

o Nei.

- SCA04 HA-løsningen må støtte opp til 16 noder i løsningen.

o Ja, støtter 32 noder.

- SCA05 HA-løsningen støtter over 16 noder i løsningen.

o Ja, støtter 32 noder.

Ytelse og ressurser

- PERF01 Tjenester på HA-løsningen må ikke ha nevneverdig dårligere ytelse enn om tjenesten kjørte på tilsvarende maskinvare utenfor en HA-løsning.

o Ja. Antar dette er omtrent som for ServiceGuard, men vi vet ikke sikkert før vi har prøvd.

Appendix 17 Mapping av VMwares HA-løsning mot Kravspesifikasjon

Administrasjon, drift og driftsmiljø

- ADM01 Det må være støtte "rolling upgrade" av operativsystemet på nodene i HA-løsningen.
 - Ja, de virtuelle maskinene flyttes i fart (uten nedetid) for patching av de fysiske maskinene i clusteret. Patching av OS på de virtuelle maskinene gjøres på vanlig måte.
- ADM02 Nodene i HA-løsningen må kunne oppdateres forløpende med systemoppgraderinger uten at vi mister support på HA-løsningen.
 - Ja.
- ADM03 Det er ønskelig at leverandøren av HA-løsningen raskt støtter nye Linux-kjerner.
 - Ja, støtten for RHEL går kun på major-versjoner.
- ADM04 HA-løsningen må ikke være "for kompleks" å sette opp og å drifte
 - Vi har allerede opplegg for VMware HA. [VMware FT](#) er ikke kompleks og kan skrues av/på i fart (også behovsbasert).
- ADM05 Nye versjoner av HA-løsningen er mest mulig bakoverkompatibel slik at oppgraderinger ikke bryter eksisterende konfigurasjon
 - Ja.
- ADM06 Programvaren i HA-løsningen kan oppgraderes "rolling", dvs node for node uten at alle tjenester må tas ned
 - Ja.
- ADM07 Det er mulighet til å skripte kommunikasjonen mot HA-systemet (ikke avhengig av GUI) for å kunne automatisere driftsmiljøet rundt løsningen.
 - Ja.
- ADM08 Tjenester i løsningen kan oppgraderes "rolling" slik at oppgraderingen ikke gir nedetid på løsningen.
 - Nei, på samme måte som for frittstående fysiske maskiner. Oppgradering av VMware-produktene kan gjøres i fart uten nedetid.

¹ [VMware FT](#) har foreløpig en begrensning på 1vCPU på den virtuelle maskinen.

Applikasjonssupport

- APP01 Det er supportert å kjøre Oracle på HA-løsningen
 - Tjaei, Oracle forbeholder seg retten til å kreve bug reprodusert på fysisk hardware hvis de har mistanke om at bugen kan skyldes det at maskinvaren er virtuell.
- APP02 Det finnes ferdig Oracle-rammeverk tilpasset HA-løsningen
 - Ikke relevant, samme som for frittstående fysiske maskiner.
- APP03 Postgres kan kjøre på løsningen
 - Ja, kan være issues med I/O (noe VMware jobber med, og med ny (og raskere) lagringsløsninger og nyere versjoner av VMware vil dette bli bedre.
- APP04 Det finnes ferdig Postgres-rammeverk tilpasset HA-løsningen
 - Ikke relevant, samme som for frittstående fysiske maskiner.

- APP05 Det er supportert/mulig (avhengig av hvilken MySQL løsning vi velger) å kjøre MySQL på løsningen
 - Ja, kan være issues med I/O (noe VMware jobber med, og med ny (og raskere) lagringsløsninger og nyere versjoner av VMware vil dette bli bedre.
- APP06 Det finnes ferdig MySQL-rammeverk tilpasset HA-løsningen
 - Ikke relevant, samme som for frittstående fysiske maskiner.
- APP07 Det er supportert å kjøre SAP på HA-løsningen.
 - Vet ikke .. må sjekkes (FIXME).
- APP08 Det finnes ferdig SAP-rammeverk tilpasset HA-løsningen
 - Ja, samme som for frittstående fysiske maskiner.
- APP09 IMAP kan kjøre på løsningen
 - Ja, kan være issues med I/O (noe VMware jobber med, og med ny (og raskere) lagringsløsninger og nyere versjoner av VMware vil dette bli bedre.
- APP10 Det finnes ferdig IMAP-rammeverk tilpasset HA-løsningen
 - Ja, samme som for frittstående fysiske maskiner.
- APP11 HA-løsningen har støtte for fast application notification (FAN) failover tilkoplingsteknologi
 - Nei, samme som for frittstående fysiske maskiner.

Failover og flytting av tjenester

- FAIL01 Ved flytting av en tjeneste fra en node til en annen i HA-løsningen er nedetiden for tjenesten omtrent det samme som tiden det tar å stoppe og starte ressursene i tjenesten på tilsvarende maskinvare utenfor HA-løsningen.
 - Ok - det er ikke nedetid.
- FAIL02 Det må være mulig å flytte en tjeneste manuelt fra en node til en annen.
 - Ja.
- FAIL03 Det er ønskelig at man gjøre en planlagt flytting til en annen node uten nedetid på tjenesten (live migrering).
 - Ja.
- FAIL04 Feil på maskinvare i HA-løsningen skal ikke gi langvarig nedetid for brukere av tjenesten.
 - Ok - så lang tid det tar å reboote en virtuell maskin og starte de relevante tjenestene. Ved bruk av [VMware FT](#)¹ er det mikroskopisk nedetid (et par ms) ved feil på en av nodene som kjører den virtuelle maskinen.
- FAIL05 Feil på maskinvare i HA-løsningen skal ikke gi nedetid for brukere av tjenesten
 - Ok - så lang tid det tar å reboote en virtuell maskin og starte de relevante tjenestene. Ved bruk av [VMware FT](#)¹ er det mikroskopisk nedetid (et par ms) ved feil på en av nodene som kjører den virtuelle maskinen.
- FAIL06 Feil på maskinvare i HA-løsningen skal ikke kreve manuelle inngrep for å gjøre tjenesten tilgjengelig for brukeren.
 - Ok.
- FAIL07 Brukere av tjenesten skal ikke måtte endre i konfigurasjon på klienten etter at en tjeneste er flyttet til en annen node i HA-løsningen.

Ok.

¹ [VMware FT](#) har foreløpig en begrensning på 1vCPU på den virtuelle maskinen.

HW- og SW-support og leverandørforhold

- SUP01 HA-løsningen skal kunne kjøre på maskinvare fra flere leverandører
 - Ja.
- SUP02 Nodene i HA-løsningen kan være virtuelle
 - Nei, man trenger fysiske maskiner for å kjøre VMware. :)
- SUP03 HA-løsningen skal kunne kjøre på ett eller flere operativsystemer som støttes av USIT i dag.
 - Ja, USIT støtter både VMware og RHEL.
- SUP04 HA-løsningen bør kunne kjøre på de nyeste versjonene av de operativsystemer som støttes av USIT i dag.
 - Ja.
- SUP05 HA-løsningen må ha support fra leverandør(er) eller tredjepart(er) i alle ledd.
 - Ja, eneste issue er Oracle (se punkt APP01)
- SUP06 Det er ønskelig med færrest mulig forskjellige leverandører involvert i den totale løsningen
 - Ok, VMware blir en ekstra leverandør av "clusterløsningen". Lisenser kjøper vi gjennom Dell eller HP.
- SUP07 Løsningen må være supportert på de versjoner av kommersiell software som er aktuelt for USIT å kjøre
 - Oracle 10.2.0 og 11.2.0 databaser: tjæi, (se punkt APP01).
 - SAP (NetWeaver 7.01, NetWeaver PI 7.10, ERP 6.0, SolMan 7.01): vet ikke .. må sjekkes (FIXME).
- SUP08 USIT må ikke miste support på HA-løsningen dersom HA-løsningen er satt opp av USIT, eller er om konfigurasjonen er endret av USIT etter installasjon og initiell konfigurasjon.
 - Ok.
- SUP09 Det finnes referansekunder for HA-løsningen som kjører i en driftssituasjon som kan sammenlignes med vår egen.
 - Ja.
- SUP10 HA-løsningen skal inngå i et aktivt prosjekt eller være i aktiv utvikling hos en "solid og levedyktig" leverandør.
 - Ja.
- SUP11 Det finnes aktive epost-lister og/eller webforum for produktet.
 - Ja.
- SUP12 HA-løsningen er open source.
 - Nei.
- SUP13 HA-løsningen er dokumentert.
 - Ja.
- SUP14 HA-løsningen er godt dokumentert.
 - Ja.

Konfigurasjon av tjenester

- CONF01 Det må være mulig å konfigurere hvilke ressurser som skal være knyttet til en tjeneste
 - Ja (mye av det man tenker på her defineres i OS som hvis man kjører på en frittstående fysisk maskin).
- CONF02 Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.
 - Ja, VMware har endel slike muligheter.
- CONF03 Det oppleves som enkelt å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.
 - Ja.
- CONF04 Det må være mulig å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.
 - Ja, dette gjøres i OSet. Kan kanskje være vanskelig å få kjørt script/kommandoer ifm. start/stopp av en virtuell maskin (må evt. sjekkes).
- CONF05 Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned.
 - Ja, gjøres i OS på vanlig måte. Noen endringer av den virtuelle maskinen kan ikke gjøres i fart.
- CONF06 Det oppleves som enkelt å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.
 - Ja, på samme måte som frittstående fysisk maskin.
- CONF07 Det må være mulig å spesifisere hvilke noder en tjeneste skal kunne gå på, og rekkefølgen for hvordan tjenesten skal gjøre failover til andre noder.
 - Ja og nei, en virtuell maskin kan gå på alle nodene i et cluster som har de ressursene (f.eks. vlan eller lagringsløsning) den virtuelle maskinen er avhengig av. Det er mulig å prioritere ressurser mellom virtuelle maskiner. VMware har også noe som kalles DRS (distributed resource scheduler) som sørger for å flytte virtuelle maskiner mellom fysiske maskiner for å spre lasten.
- CONF08 Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder
 - Ja, men ettersom dette er virtuelle maskiner snakker vi kun om avhengigheter mellom disse (og evt. at en virtuell maskin må eller ikke må kjøre på samme fysiske maskin som en annen).
- CONF09 HA-løsningen kan konfigureres slik at den restarter ressurser i en tjeneste som den har detektert at ikke lenge kjører.
 - Nei, VMware har ikke noe forhold til hvilke applikasjoner/tjenester som kjører inne i en virtuell maskin.

Lagring og filsystem

- STOR01 Løsningen må kunne benytte seg av delt lagring, enten via et SAN (FC eller iSCSI) eller nettverksfilssystem (NFS).
 - Ja - FC, iSCSI og NFS er støttet.
- STOR02 Det er ønskelig at løsningen kan støtte et cluster-filsystem som tillater samtidig skriving fra flere noder.
 - Ja - for FC og iSCSI brukes VMFS3 (VMwares egne cluster-filsystem) og NFS er jo et delt filsystem.
- STOR03 Løsningen bør være supportert på den lagrings-infrastruktur som USIT allerede støtter (SAN, fc-

- switcher, lagringsprodukter) så vi slipper å kjøpe inn egen lagring til løsningen.
- Ja.

Overvåkning og feilsøk

- MON01 Det må enkelt å få ut info om status for HA-løsningen, dets noder og dets tjenester.
 - Ja.
- MON02 Logging må være tilstrekkelig slik at det er mulig å feilsøke ved eventuelle problemer.
 - Ja.
- MON03 Det er ønskelig at HA-løsningen kan sende sine logger fra alle noder til et sentralisert sted
 - Ja, blant annet syslog og evt. epost ved alarmer internt i VMware.
- MON04 HA-løsningen må la seg overvåke av Nagios
 - Tja, det går - men er nok ikke trivielt. De virtuelle maskinene overvåkes av Nagios på normalt vis.
- MON05 Oracle-databaser som kjører i HA-løsningen må la seg overvåke av GridControl.
 - Ja, som på frittstående fysiske maskiner.
- MON06 Det må være mulig å få generert en alarm dersom en ressurs i tjenesten ikke kjører
 - Ja, kan gjøres med Nagios som for normale fysiske og virtuelle maskiner.

Nettverk

- NET01 HA-løsningen støtter IPv6
 - Ja.

Priser og lisenser

- LIC01 Kostnaden for løsningen ved anskaffelse og drift bør ikke være vesentlig høyere enn for dagens SG-løsninger, også for eksterne kunder som ikke er omfattet av UiO sin site-lisens.
 - ServiceGuard er veldig billig, ca 8000 inkludert tre års support.. VMware-versjonen vi bruker idag ligger på ~35.000,- eks. mva. per 2-socket-maskin, men det er nok ønskelig å gå opp til neste nivå hvis det blir utstrakt bruk av HA (trolig opp mot ~40.000,- eks. mva. per 2-socket-maskin). Oracle-lisenser for Trofast er p.t pr "named users". Generelt må man betale pr cpu for alle noder clusteret dersom man bruker VmWare, men vi har mao ikke lisens pr cpu for Trofast i dag så lisenskostnad for Oracle blir det samme om man bruker VmWare.

Redundanse og robusthet

- ROB01 HA-løsningen skal være en "moden" teknologi idet det tas i bruk av USIT
 - Ja.
- ROB02 Løsningen må kunne settes opp og være supportert med multiple paths til både nettverk og lagring for å unngå single-point-of-failure.
 - Ja.
- ROB03 Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).

- Ja.
- ROB04 Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på to noder samtidig).
 - Ja.
- ROB05 Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing).
 - Nei. Vi har ikke klart å finne noen måter clusteret kan konfigureres til å skru av noder som blir isolert med IPMI, iLO eller DRAC. Det som er mulig er å skru på hva en ESX-boks skal gjøre med VMene sine når den blir isolert (den kan da skru dem av og de vil bli tatt opp igjen på andre ESX-servere).
- ROB06 Det er ønskelig med mest mulig isolasjon mellom forskjellige tjenester som går på samme node i løsningen.
 - Ja (egen instans av OS per tjeneste).
- ROB07 Nodene i HA-løsningen kan stå i forskjellige nett.
 - Ja, (tror de må ha en felles IP-adresse som svarer på ping)

Skalerbarhet

- SCA01 Det må være mulig å melde noder inn i og ut av HA-løsningen uten at det påvirker tjenestene i løsningen (med unntak av tjenester som går på noder som skal fjernes).
 - Ja.
- SCA02 Det må være mulig å definere nye tjenester i løsningen uten at det krever nedetid på andre tjenester i løsningen.
 - Ja.
- SCA03 Det er ønskelig at løsningen gjør det mulig å skalere tjenester ut over kapasiteten til en enkelt node
 - Nei.
- SCA04 HA-løsningen må støtte opp til 16 noder i løsningen.
 - Ja.
- SCA05 HA-løsningen støtter over 16 noder i løsningen.
 - Ja.

Ytelse og ressurser

- PERF01 Tjenester på HA-løsningen må ikke ha nevneverdig dårligere ytelse enn om tjenesten kjørte på tilsvarende maskinvare utenfor en HA-løsning.
 - Tjaei, bør sjekke ut med noen uavhengige hva som er dagens kurs. Tidligere har det vært snakk om 5-10%, men det kommer ann på hvilke ressurser man snakker om.

Appendix 23 Dokumentere valg av HA-løsning for test

Kravspesifikasjon og hvilke utvalg av HA-løsninger som skal vurderes

Hovedleveransen i Milepæl 5 var at prosjekteier godkjent en kravliste for HA-løsninger til de tjenester som prosjektet omfatter:

- [Appendix 9 Kravspesifikasjon til HA-løsning](#)

I Milepæl 6 laget vi en oversikt over HA-teknologier

- [Appendix 10 Liste over HA-teknologier](#)

som vi mente var aktuelle.

Vi studerte så disse HA-løsningene, og på prosjektmøter 14. og 21. oktober ble mange løsninger utelukket fra videre vurdering. Det er dokumentert i oversikten over HA-løsninger *hvorfor* løsningene ble utelukket.

På prosjektmøte den 21. oktober bestemte vi at følgende løsninger skulle mappes mot kravlisten:

- Oracle DataGuard
- Oracle Grid
- RedHat Cluster Suite (RHCS)
- SteelEye
- Veritas Cluster Server (VCS)
- VmWare HA

Mappingen ble gjort i et eget dokument for hver løsning, og resultatet ble sammenfalt i en krav-produkt-matrise.

[Appendix 11 Mapping av HA-produkter mot kravspesifikasjon](#)

(Klikk på produktnavnet i tabellheaderen for å se detaljmappingen)

Gjennomgang av krav-produktmastrisen.

På prosjektmøte den 28. oktober gikk vi gjennom krav-produktmatrisen for å evaluere de ulike HA-løsningene.

Scope

Siden HA-løsningene

- Oracle DataGuard
- Oracle Grid
- VmWare HA

allerede tatt bruk av USIT for spesifikke tjenester, er det ikke nødvendig å teste disse. Dette er heller ikke generelle HA-løsninger. Derfor vurderer vi kun følgende løsninger mot kravspesifikasjonen når vi skal vurdere hvilke vi skal teste:

- RHCS
- SteelEye Lifekeeper
- VCS

Generelle krav

Nedenfor følger en oversikt over de **generelle krav** som ikke ble innfridd av en eller flere av HA-løsningene. Krav som alle HA-løsningene innfridde er ikke tatt med.

Kode	Beskrivelse	Krav/ Ønske	RHCS	SteelEye	VCS	Kommentar
SUP07	Løsningen må være supportert på de versjoner av kommersiell software som er aktuelt for USIT å kjøre Eksempel: DBA-gruppen migrerer nå databaser til Oracle 11.2.0 som må være støttet.	Krav	Ja trolig	Ja	Ja	RHCS i RHEL5 ser ikke ut til å støtte Oracle 11gR2 direkte. Dette kan fikses ved å bruke modulen for Oracle 10gR2.
MON04	HA-løsningen må la seg overvåke av Nagios	Krav	Ja	Trolig	Trolig	Det finnes en ferdig modul til RHCS. Man kan lage egen modul til SteelEye og VCS.

Konklusjon: Alle de tre HA-løsningene støtter samtlige krav som stilles, men

- det må gjøres en tilpassing for å kunne kjøre Oracle 11.2.0 på RHCS
- vi må lage en egen modul for Nagios på samme måte som vi gjorde for ServiceGuard

*Det er mao ikke mulig å utelukke noen av HA-løsningene pga **kravene** i kravspesifikasjonen.*

Krav for tjenester som i dag kjører på Oracle RAC

Tabellen nedenfor krav som er spesifikke for applikasjoner som i dag kjører på Oracle RAC, og hvor en eller flere HA-løsninger ikke kunne innfri kravene.

Kode	Beskrivelse	Krav/ Ønske	RHCS	SteelEye	VCS	Kommentar
ADM08	Tjenester i løsningen kan oppgraderes "rolling" slik at oppgraderingen ikke gir nedetid på løsningen. Eksempel: I RAC kan Oracle-installasjonen oppgraderes rolling på node for node mens tjenesten fortsatt er tilgjengelig på de andre nodene.	Krav fra noen tjenester på RAC	Nei	Nei	Nei	Bare Oracle Grid med RAC tilfredstiller dette. Løsninger som krever dette må

						fortsette å kjøre på RAC.
APP11	HA-løsningen har støtte for <i>fast application notification</i> (FAN) failover tilkoplingsteknologi Forklaring: Vortex bruker dette for å redusere nedetid når en node eller databaseinstanse feiler	Krav for Vortex som kjører på RAC	Nei	Nei	Nei	FAN støttes kun av RAC out-of-the-box. Løsninger som krever dette må fortsette å kjøre på RAC.
SCA03	Det er ønskelig at løsningen gjør det mulig å skalere tjenester ut over kapasiteten til en enkelt node Forklaring: Tjenester kan kjøre på flere noder samtidig (aktiv-aktiv)	Krav for Atlas 3D som kjører på RAC	Nei	Nei	Nei	Bare Oracle Grid med RAC tilfredstiller dette. Løsninger som krever dette må fortsette å kjøre på RAC.

Konklusjon: Ingen av de tre HA-løsningene innfrir disse kravene. Tjenester som har disse kravene må kjøre på Oracle RAC som i dag.

Ingen av HA-løsningene har noen fortrinn når det gjelder disse kravene.

Krav fra SAP

Nedenfor ser vi krav som gjelder applikasjonen SAP og som ikke ble innfridd av en eller alle HA-løsningene:

Kode	Beskrivelse	Krav/Ønske	RHCS	SteelEye	VCS	Kommentar
CONF08	Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder Eksempel: Tjeneste X skal ikke gå på samme node som tjeneste Y", eller "Tjeneste X og tjeneste Y må alltid gå på samme node".	Krav for SAP	Delvis	Ja	Ja	Sannsynligvis vil SAP i fremtiden kjøres på RAC, så dette er ikke kritisk.

Konklusjon: Alle tre HA-løsningene kan kjøre SAP, men de har implementert slik avhengigheter i ulik grad.

Ingen av HA-løsningene har noen fortrinn når det gjelder disse kravene.

Generelle ønsker

Nedenfor følger en oversikt over de **ønsker** som ikke ble tilfredstilt av en eller flere av HA-løsningene :

Kode	Beskrivelse	Krav/	RHCS	SteelEye	VCS	Kommentar
------	-------------	-------	----------------------	--------------------------	---------------------	-----------

		Ønske				
APP04	Det finnes ferdig Postgres-rammeverk tilpasset HA-løsningen	Ønske	Ja	Ja	Nei	Man kan lage eget rammeverk for Postgress på VCS
APP10	Det finnes ferdig IMAP-rammeverk tilpasset HA-løsningen	Ønske	Nei	Nei	Nei	Kan lage rammeverk selv i alle tre løsningene.
FAIL05	Feil på maskinvare i HA-løsningen skal ikke gi nedetid for brukere av tjenesten	Ønske	Nei	Nei	Nei	
SUP06	Det er ønskelig med færrest mulig forskjellige leverandører involvert i den totale løsningen Eksempel på uønsket situasjon: Oracle-SW, HP-lagring, RedHat-OS, Veritas-cluster, Dell servere	Ønske	Ja	Nei	Nei	RHCS innfrir dette fordi vi har standardisert på RHEL.
SUP11	Det finnes aktive epost-lister og/eller webforum for produktet	Ønske	Ja	Nei	Ja	
SUP12	HA-løsningen er open source	Ønske	Ja	Nei	Nei	
CONF03	Det oppleves som enkelt å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt. Eksempel på egen betingelse: Når filsystemene til pakken er utilgjengelige skal tjeneste tas ned og ikke flyttes til en annen node	Ønske	Ukjent	Trolig	Trolig	
CONF06	Det oppleves som enkelt å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.	Ønske	Ukjent	Ja	Ukjent	
STOR02	Det er ønskelig at løsningen kan støtte et cluster-filsystem som tillater samtidig skriving fra flere noder.	Ønske	Ja	Nei	Ja	Vi kommer neppe til å benytte clusterfilsystem.
NET01	HA-løsningen støtter IPv6	Ønske	Nei	Nei	Ja	Vi kommer ikke til å bruke IPv6 på lang tid ennå.
LIC01	Kostnaden for løsningen ved anskaffelse og drift bør ikke være vesentlig høyere enn for dagens SG-løsninger, også for eksterne kunder som ikke er omfattet av UiO sin site-lisens.	Ønske	Trolig	Ukjent	Ukjent	Dersom prisene varierer veldig vil dette kunne få betydning for valg av HA-løsning
ROB06	Det er ønskelig med mest mulig isolasjon mellom forskjellige tjenester som går på samme node i løsningen. Eksempel: Et sikkerhetshull i en tjeneste på en node bør ikke gjøre det mulig å kompromitere andre tjenester på samme node.	Ønske	Nei	Ukjent	Nei	Kun med VmWare er dette ønsket greit å innfri.
SCA05	HA-løsningen støtter over 16 noder i løsningen	Ønske	Nei	Ja	Ja	

--	--	--	--	--	--	--	--	--	--

Konklusjon:

Ikke alle ønskene er like viktige.

For noen *ønsker* kan man selv lage en løsning som vil innfri ønsket selv om det ikke finnes en ferdig løsning. For andre *ønsker* er det vanskelig å gi svar før vi har testet HA-løsningen.

Følgende *ønsker* er tillagt spesiell vekt av prosjektet

- SUP06 : Færrest mulig leverandører (kun RHCS innfrir denne)
- SUP12: HA-løsningen er open source (kun RCHS innfrir denne)

Følgende *ønske* er tillagt spesiell vekt av prosjekteier.

- LIC01: Kostnaden for løsningen ved anskaffelse og drift... (RHCS er billigere enn de andre)

Vi ser at RHCS trekker det lengste strået når det gjelder å innfri ønsker. Spesielt når det gjelder LIC01/pris vil dette kunne være utslagsgivende.

Hovedkonklusjon etter mapping av produkter mot krav

- *Alle tre HA-løsninger innfrir våre **krav**, og det er ikke mulig å kunne si at den ene løsningen bedre enn den annen før vi har testet dem.*
- *Når det gjelder **ønsker**, har RHCS et fortrinn spesielt fordi prosjekteier legger vekt på LIC01.*
- *Ingen av løsningen har noe fortrinn når det gjelder å kjøre SAP.*
- *Løsninger som i dag er avhenige av RAC-funksjonalitet, må fortsatt kjøre på RAC.*

Tilleggstester

I tillegg går prosjektet inn for å gjøre følgende tester:"

- Teste failover mellom primær og standby database med Dataguard. Testene gjøres på eksisterende maskiner og vil ikke ta mye ressurser
- Gjøre enkle ytelsetester slik at vi evt kan sammenligne med VmWare HA på de samme maskinene etterpå.

Priser på HA-løsninger

Som avsnittet ovenfor viser, tilfredstiller alle HA-løsningene våre krav. Av den grunn vil *ønsker* i kravspesifikasjonen også kunne spille avgjørende rolle.

Prosjektet har derfor hentet inn pris på de forskjellige løsningene siden kravet [LIC01](#) definerer et *ønske* om at "*prisen for den nye løsningen ikke skal være vesentlig høyere enn dagens SG-løsninger*"

Etter dialog med leverandørene har vi kommet frem til "budsjettpriser":

Det var stor forskjell i prisene.

Knut legger vekt på [LIC01](#) (pris på løsningen) og ønsker at vi verifiserer alle prisene med sammenlignbare supportlevel.

Dersom RHCS etter en verifikasjon fortsatt er så mye billigere, ønsker prosjekteier at vi først tester RHCS. Kun

dersom RHCS viser seg ikke å fungere godt nok, skal vi teste VCS og SteelEye Lifekeeper.

En verifikasjonen av prisene viser at prisnivået som presentert ovenfor blir stående.

Konklusjon på A06.08

På bakgrunn av mapping av HA-løsningene mot kravlisten og prosjekteier sin vektlegging av [LIC01](#)/pris på HA-løsningen, er konklusjonen:

- **RHCS skal testes.**
- **Kun dersom RHCS ikke fungerer godt nok, skal vi vurdere om VCS eller LifeKeeper skal testes.**
- **I tillegg gjør vi tester av ytelse på RHCS for senere å kunne teste om VmWare gir sammenlignbare ytelse.**
- **Vi tester også failover for Oracle DataGuard på eksisterende HW og Dataguard-oppsett**

Aktivitetene for M7 ("Når testing av kandidatteknologier er godkjent av PE") må omarbeides noe til å reflektere disse valgene. Dette er noe som gjøres intern i prosjektgruppen.

Appendix 22 Oppsummering av tester

Innhold

- [Hva vi ønsker å teste](#)
- [Presentasjon av testplaner](#)
- [Gjennomgang av testresultater](#)
- [Konklusjon](#)

Hva vi ønsker å teste

I dette kapittelet tar vi for oss hva vi ønsker å teste på de forskjellige løsningene.

Dataguard

Med Oracle Dataguard har man kort fortalt mulighet til å replikere en database ("primærdatabasen") realtime slik at man har en oppdatert kopi ("standby-databasen") som man senere kan switch'e over til dersom primærdatabasen skulle feile. En slik switch hvor en standbydatabase bytter rolle til primærdatabase kan automatiseres, men i våre tester gjorde vi denne switch'en manuelt.

For mer informasjon om DataGuard kan man ta en titt her:

<http://www.oracle.com/technetwork/database/features/availability/dataguardoverview-083155.html>

USIT har i lengre tid benyttet Dataguard til forskjellige formål, men ikke i et failover-oppsett. Vi ønsket derfor å teste hvordan man kan sette opp Oracle-klienter til automatisk å benytte seg av standbydatabasen dersom primærdatabasen skulle feile.

DataGuard inngår i våre allerede innkjøpte lisenser for Oracle, og er således "gratis" å ta i bruk.

RHCS

RHCS er en tradisjonell failover-løsning hvor tjenester automatisk blir relokert til en annen node i clusteret dersom tjenesten ikke kan kjøre på sin opprinnelige node. Oppsettet til et RHCS-cluster er definert i xml-filen `/etc/cluster/cluster.conf` og består grovt sett av følgende:

- hvilke noder som inngår i clusteret
- hvordan de skal "*fences*"
 - "*fencing*" er en mekanisme for å "uskadeliggjøre" en node i clusteret slik at den ikke kan forårsake datakorupsjon. RedHat anbefaler generelt at man bruker såkalt "powerfencing" hvor man slår strømmen av og på ("powercycle") for en node. Vi har valgt å gjøre dette ved å bruke IPMI til iDRAC6-kortet på noden
 - noder må "*fences*" før RHCS relokerer tjenester fra en node til en annen ved feilsituasjoner på opprinnelig node.
- *failoverdomains*
 - definerer ulike grupperinger av noder som tjenestene kan kjøre på. Kan f.eks inkludere alle nodene i clusteret, eller et subsett av nodene.
- Tjenester ("*services*") som skal kjøre i clusteret

En tjeneste består av en eller flere ressurser som kan defineres som søsken (uavhengige av hverandre) eller i et

foreldre-barn forhold (barn er avhengig av at foreldre-ressursen fungerer). Det finnes mange ulike typer ressurser som følger med RHCS, f.eks ressurser for LVM volumgruppe og logiske volum, filsystemer, ip-nr, Oracle-databaser, PostgreSQL-databaser, SAP, Apache web-tjener osv. Det finnes også en generell "script-ressurs" hvor man bare angir path til et "init-script" for tjenesten. På den måten man kan bruke "script-ressursen" til å starte "hva som helst" av programmer i RHCS.

Eksempel på en tjeneste: Oracle database

Oracle-servicen *rhcsoracle* i clusteret vårt består av følgende ressurser som settes opp i et "foreldre-barn forhold" fordi det er avhengigheter mellom ressursene:

- Ip-nr
- Volumgruppe
- Filsystem
- Oracle-database
- Script-ressurs for oppdatert Oracle Managment Agent (for GridControl)

Testene for RHCS ble kjørt på en presumptivt stabil utgave av RHCS på RHEL5.5, og på en ny versjon på RHEL6.0.

Hensikten med testene er å verifisere at

- HA-løsningen innfrir de krav vi stiller
- HA-løsningen støtter de tjenestene vi skal kjøre på en god måte

I testplanen er det derfor *ikke* et primært mål å verifisere at HW fungerer eller at all kabling er korrekt utført i oppsettet slik man tester produksjonsmiljøer før oppstart.

Det er også viktig å være klar over at RHCS og det fleste tilsvarende systemer er designet til å håndtere en feil om gangen.

Nagios

På samme måte som USIT i sin tid laget en Nagios-modul for ServiceGuard for Linux (SGLX) for å få bedre overvåkning, ønsker vi å ha Nagios-alarmer fra RHCS. Vi planla følgende fremgangsmåte:

- Lage kravspesifikasjon for Nagios
- Evaluere og teste mulige Nagios-moduler fra Nagios-exchange.
- Kjøre Nagios-modul på nodene mens vi testet RHCS

Ytelsestester

Pga ressursmangel ble det ikke utført ytelsestester mot diskløsningen. Dette må evt gjøres i regi av tjenestegruppen for virtualisering.

Presentasjon av testplaner

Dataguard

For testing av av DataGuard benyttet vi eksisterende maskinvare. Testplanen og resultatene befinner seg her:

- [Appendix 19 Testing av Oracle Dataguard](#)

Her ble det definert en *database-service* som bestod av både primær- og standby-database. Vi brukte `sqlplus` for å

koble til denne database-servicen. Når primær-databasen var oppe og standby-databasen var i såkalt "apply mode" (hvor den legger på oppdateringer som blir replikert fra primary-databasen) er forventningen at man med `sqlplus` skal koble opp mot database-servicen og uten forsinkelser bli tilkoblet primærdatabasen.

Som testplanen viser, gjør man deretter en *switchover* slik at standbydatabasen blir en ny *primær* (dvs at man bruker "kopien" i stedet for originalen). Da er forventningen at man skal slippe å endre i oppsettet på klienten og likevel kunne bli tilkoblet korrekt database. Under gitte forutsetninger er det også forventet at man ikke skal oppleve noe forsinkelse ved påkobling.

RHCS

I dagens ServiceGuard for Linux (SGLX) kjører vi med quorum-server. Hvordan *fencing* og quorum løses i andre HA-løsninger er nytt for oss og derfor har vi valgt å legge spesielt vekt på testing av dette.

Vi har derfor delt testplanen i tre deler:

- funksjonelle tester med utgangspunkt i kravspesifikasjonen unntatt robusthetstester
- test av robusthet mhp fencing og quorum.
- destruktive tester med utgangspunkt i en systemskisse

Testplanen befinner seg her:

- [Appendix 18 Testplan for HA-løsning](#)

Vi skal lage tre services i clusteret som inneholder hhv

- Oracle-database og GridControl agent
- Postgres-database
- egenutviklet tjener ("Halv2-Server") som svarer på en port og som skriver til en fil.

Nagios

Vi laget en enkel kravspesifikasjon for en Nagios-modul for RHCS som kan leses her:

- [Appendix 20 Kravspesifikasjon for Nagios-modul RHCS](#)

Etter videre evaluering ønsket vi i praksis bare alarm i følgende tre tilfeller:

Krav#	Melding i Nagios når...	Krav/Ønske
N03	en ressurs i en service ikke kjører: F.eks: "Ressurs ZZZ kjører ikke lenger i servicen XXX i clusteret CCC"	Krav
N04	en service er nede: Feks: "Service XXX er nede i clusteret CCC"	Krav
N06	en node ikke er med i clusteret: F. eks: "Node YYY er ikke lenger med i clusteret CCC"	Krav

rpm-pakken `cluster-snmpp` inneholder bla `redhat-cluster MIB`, så denne må installeres før man kan teste en Nagios-modul.

Gjennomgang av testresultater

Dataguard

Resultatene fra testing av Dataguard kan leseren finne her:

- [Appendix 19 Testing av Oracle Dataguard](#)

Her vil man se om Oracle-klienten uten å endre i sin konfigurasjon automatisk kan benytte seg av både primær- og standby-databasen. Så lenge klienten får svar fra ip-nummeret som er konfigurert for tilkobling til primær-databasen, vil også pålogging til den tidligere standbydatabasen skje uten målbar forsinkelse. Dersom maskinen som kjørte primær-databasen ikke svarer på ip-nummeret, får man et vanlig tcp-heng på ca 1 minutt før klienten forsøker å koble til standbydatabasen. Denne forsinkelsen er ikke akseptabel i alle typer systemer.

Vi gjorde en manuell switch fra primær- til standby-database. Oracle har egne løsninger for hvordan dette kan automatiseres (Oracle Data Guard Fast-Start Failover) men dette er ikke noe vi har testet på USIT.

Testene viser at DataGuard kan brukes i en HA-løsning, men at man må håndtere dette med tcp_heng og automatisk switch fra primary database.

RHCS på RHEL5

Testresultatene for testing av RHCS på RHEL5 finner man her:

- [Appendix 21 Testresultater RHCS](#)

På forhånd var vi usikker på hva vi egentlig forventet som resultat i noen av testene, noe som jo skyldes at vi på forhånd ikke hadde noen erfaring med RHCS. I noen av testene måtte vi også modifisere oppsettet slik at vi fikk et resultat som svarte til det vi krever av løsningen.

RHCS på RHEL5 fungerte bra, og vi kunne kjøre både PostgreSQL, Oracle og vår egenutviklede tjener ("Halv2-tjener") på løsningen. Vi hadde noen få tilfeller hvor vi fikk inkonsistent oppførsel i clusteret, og vi løste dette ved å reboote maskinen som hadde problemer. Vi kom aldri til bunns i hva som forårsaket inkonsistensen.

Vi noterte oss følgende avvik fra forventede resultater:

Test:	FU02 Programvaren i HA-løsningen kan oppgraderes "rolling", dvs node for node uten at alle tjenester må tas ned
Avvik:	Overskrevne ressursscript
Diskusjon:	Siden vi ikke fikk de originale scriptene for oppstart av Oracle og PostgreSQL til å fungere, laget vi tilpassede versjoner. Ved oppgradering av RHCS blir våre endringer overskrevet. Dersom vi også i fremtiden skal bruke modifiserte script, må vi ta opp med RedHat hvordan dette kan løses

Test:	FU08 Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.
-------	--

Avvik:	Ikke støttet "out of the box"
Diskusjon:	I RHCS er det ikke supportert "out of the box" at det kan stilles egne betingelser for at en service skal starte eller kjøre på en node, og hva som skal skje dersom betingelsen ikke er oppfylt. RedHat anbefaler å håndtere betingelser ved å legge inn en script-ressurs i tjenesten som returnerer gitte feilkoder ut fra om betingelsen er tilfredsstilt. Denne ressursen settes opp som <i>parent</i> av resten av ressursene i tjenesten. Vi kan ikke styre om tjeneste skal flytte til annen node eller restartes ut fra en gitt betingelse slik som formulert i kravet siden failover-egenskapene er satt <u>en gang for alle på servicen</u> og ikke kan endres mens man kjører. Det kravet vi tester her (CONF02) egentlig er ment å formulere er at ikke <u>alle</u> ressurser som feiler skal måtte medføre at servicen flyttes eller restartes, og dette lar seg enkelt implementere i RHCS ved å endre i ressursscriptene.

Test:	FU10 Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned.
Avvik:	Restart av ressurser og tjenester uten varsel
Diskusjon:	Alle endringer i clusteroppsettet på <i>service-</i> eller <i>ressurs-</i>nivå gir restart av det nivået man har endret på og de som er avhengige at denne ressursen. <i>Man får ingen forvarsel om disse restartene når man gjør endringer.</i> Dette ble verifisert med at vi gjorde endringer med tre supporterte metoder: • editere rett i <code>cluster.conf</code> og kjøre <code>ccs_tool update</code> • gjøre endringer i <i>Conga</i> (webbasert administrasjonsverktøy for RHCS) • gjøre endringer i <code>system-cluster-conf</code>. (X-basert administrasjonsverktøy for RHCS) Dette verktøyet kan ikke brukes fordi det ikke engang kjenner igjen oracle-ressursen i <code>cluster.conf</code>. Endringer i <i>failoverdomain</i> eller <i>fence-metode</i> for noder gir ingen restarter. Kravet vi tester her (CONF05) er i kravspesifikasjonen formulert som et ønske og ikke et krav.

Test:	FU12 :Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder
Avvik:	Begrenset støtte for avhengigheter mellom services
Diskusjon:	RHCS støtter trivelles dependencies "out of the box" (via <code>depend</code> og <code>depend_mode</code> for servicen). Se DOC-26981 (hos RedHat). Det ser ut som om mer avanserte avhengigheter kan implementeres med såkalte <i>EventScript</i>. Støtte for dependencies iht kravspesifikasjonen er et "Ønske" og ikke et absolutt krav, og kun aktuelt for SAP som hvor det finnes et eget resource-kit hvor dependencies er implementert.

Test:	DE04 Bortfall all fc for en node
Avvik:	Manuelt inngrep ved bortfall av all fc mot disk
Diskusjon:	Ved bortfall av all fibre-channel connectivity for en node, blir servicene på noden tatt ned og status satt til <i>"failed"</i> . Dermed blir ikke servicene flyttet til neste node og vi får nedetid for brukeren. Man

må selv manuelt fjerne lvm-tags og starte opp tjenestene på nytt.

Årsaken er at noden uten *fc* ikke klarer å fjerne lvm-tag'ene på volumgruppen. Løsningen er å finne i RHEL5.6 eller RHEL6 hvor man kan kjøre HA-LVM sammen med *clvmd*. Dette har vi ikke testet, og generelt ønsker vi ikke å benytte *clvmd*.

Denne feilen er ingen "single failure" men heller to feil som inntreffer samtidig, så denne problemstillingen er derfor ikke regnet som kritisk.

Test:	DE14 Node failure DE18 Bortfall av HB-nett/Fencingnett og nettaksess til iDRAC6
Avvik:	Manuelt inngrep når IDRAC6-kort ikke er tilgjengelig
Diskusjon:	I de to testene hvor iDRAC6-kortet er utilgjengelig fordi • nettkabelen til iDRAC6-kortet er koblet fra • strømmen på noden er koblet fra vil ikke de andre nodene klare å <i>fence</i> noden dersom man <u>har fencing over IPMI</u> som <u>eneste</u> fencing-metode. Siden RHCS legger til grunn en konservativ tilnærming, insisteres det på at noden må fences før services kan startes på andre noder. Resultatet er dermed nedetid for brukerne. En måte å løse dette på er å innføre en ny fencingmetode som heter <i>manual fence</i>. Dette gir systemadministratoren mulighet til å fortelle clusteret at det er trygt å betrakte noden som fenced. Siden vi ønsker at RHCS skal forsøke å fence noden flere ganger vha IPMI før den venter på at vi skal gjøre en manuell fencing, legger vi inn tre forsøk på IPMI-fencing for hver node FØR den faller tilbake til manuell fencing. Dette ser slik ut i <code>cluster.conf</code> (eksempel fra node1 i clusteret) <pre><clusternode name="halv2-test01-hb" nodeid="1" votes="1"> <fence> <method name="1"> <device name="halv2-test01-ipmi"/> </method> <method name="2"> <device name="halv2-test01-ipmi"/> </method> <method name="3"> <device name="halv2-test01-ipmi"/> </method> <method name="4"> <device name="manualfence" nodename="halv2-test01-hb"/> </method> </fence> </clusternode></pre> Det kommer da en melding i syslog med instruksjoner om hvordan man kan kjøre en kommando for å fortelle clusteret at noden kan betraktes som fenced, og dette gjøres fra den samme noden som først forsøkte å utføre fencingen via IPMI: <pre># fence_ack_manual -n <node></pre> Når noden er merket som fenced, blir servicene relokert til en annen node, og kundene vil igjen oppleve tjenesten som tilgjengelig. Skal man ha <u>automatisk failover</u> av services i denne situasjonen må man legge inn en ny sekundær fencing-metode, feks • strømmeskinne fencing • scsi persistent reservation

	• fc fencing Som tertiær fencing kan man da legge inn manuell fencing. Denne feilen er ingen "single failure" men heller to feil som inntreffer samtidig, så denne problemstillingen er derfor ikke regnet som kritisk.
--	---

RHCS på RHEL6

Testresultatene for testing av RHCS på RHEL6 finner man her:

- [Appendix 24 Testresultater RHCS på RHEL6](#)

RHCS på RHEL6 fungerte bra, og vi kunne kjøre både PostgreSQL, Oracle og vår egenutviklede tjener "Halv2-tjener" på løsningen. Logging var merkbart bedre enn på RHEL5, og funksjonaliteten har utviklet seg og virker på flere områder mer gjennomtenkt.

Vi noterte oss følgende avvik fra forventede resultater:

Test:	FU02 Programvaren i HA-løsningen kan oppgraderes "rolling", dvs node for node uten at alle tjenester må tas ned
Avvik:	Overskrevne ressursscript
Diskusjon:	Ved oppgradering av RHCS, blir ressursscriptene overskrevet og evt lokale tilpassinger overskrevet. Under testene på RHEL6 kjørte vi med umodifiserte script siden originalscriptene fungerte, mens vi på RHEL5 kjørte med modifiserte script. Dersom vi i fremtiden kjører på modifiserte script, må ta opp med RedHat hvordan dette kan løses.

Test:	FU08 Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt.
Avvik:	Ikke støttet "out of the box"
Diskusjon:	I RHCS er det ikke supportert "out of the box" at det kan stilles egne betingelser for at en pakke skal starte eller kjøre på en node, og hva som skal skje dersom betingelsen ikke er oppfylt. RedHat anbefaler å håndtere betingelser ved å legge inn en script-ressurs i tjenesten som returnerer gitte feilkoder ut fra om betingelsen er tilfredsstilt. Denne ressursen settes opp som <i>parent</i> av resten av ressursene i tjenesten. Vi kan ikke styre om tjeneste skal flytte til annen node eller restartes ut fra en gitt betingelse slik som formulert i kravet siden failover-egenskapene er satt <u>en gang for alle på servicen</u> og ikke kan endres mens man kjører. Det kravet vi tester her (CONF02) egentlig er ment å formulere er at ikke alle ressurser som feiler skal måtte medføre at servicen flyttes eller restartes, og dette lar seg enkelt implementere i RHCS ved å endre i ressursscriptene.

Test:	FU10 Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned.

Avvik:	Restart av ressurser og tjenester uten varsel
Diskusjon:	Alle endringer i clusteroppsettet på <i>service-</i> eller <i>ressurs-</i>nivå gir restart av det nivået man har endret på og de som er avhengige at denne ressursen. <i>Man får ingen forvarsel om disse restartene når man gjør endringer.</i> Dette ble verifisert med at vi gjorde endringer med to supporterte metoder: • editere rett i <code>cluster.conf</code> og kjøre <code>cman_tool version -r</code> • gjøre endringer i <i>Conga</i> (webbasert administrasjonsverktøy for RHCS) Endringer i <i>failoverdomain</i> eller <i>fence-metode</i> for noder gir ingen restarter. Kravet vi tester her (CONF05) er i kravspesifikasjonen formulert som et ønske og ikke et krav.

Test:	FU16 Logging må være tilstrekkelig slik at det er mulig å feilsøke ved eventuelle problemer
Avvik:	Noen mangler rundt logging.
Diskusjon:	RHCS under RHEL6 har bedre logging enn i RHEL5, f.eks gir loggene i <code>/var/log/cluster/rgmanager.log</code> veldig god informasjon om tilstandene til ressursene. Det virker som om enkeltkomponentene i RHCS logger tilfredsstillende, men det mangler logging og <i>santitychecks</i> på mer overordnede strukturer. F.eks gis det ingen gode logger på alvorlige feil i <code>cluster.conf</code> som gjør at clusteret faktisk ikke fungerer. Eksempler på slike feil er duplikate <code>nodeid</code> på noder i clusteret eller <i>fence-metoder</i> som ikke finnes.

Test:	FU12 :Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder
Avvik:	Begrenset støtte for avhengigheter mellom services
Diskusjon:	RHCS støtter trivelte dependencies "out of the box" (via <code>depend</code> og <code>depend_mode</code> for servicen). Se DOC-26981 (hos RedHat). Det ser ut som om mer avanserte avhengigheter kan implementeres med EventScript. Støtte for dependencies iht kravspesifikasjonen er et "Ønske" og ikke et absolutt krav, og kun aktuelt for SAP som hvor det finnes et eget resource-kit hvor dependencies er implementert.

Test:	DE04 Bortfall all fc for en node
Avvik:	Manuelt inngrep ved bortfall av all fc mot disk
Diskusjon:	Ved bortfall av all fibre-channel connectivity for en node, blir servicene på noden tatt ned og status blir satt til "failed". Dermed blir ikke servicene flyttet til neste node og vi får nedetid for kunden. Man må selv manuelt fjerne <code>lvm-tags</code> og starte opp tjenestene. Årsaken er at noden uten fc ikke klarer å fjerne <code>lvm-tag</code>'ene på volumgruppen. Løsningen i RHEL5.6 eller RHEL6 hvor man kan kjøre HA-LVM sammen med <code>clvmd</code> har vi ikke testet. Denne feilen er ingen "single failure" men heller to feil som inntreffer samtidig, så denne problemstillingen er derfor ikke regnet som kritisk.

Test:	DE14 Node failure
-------	-----------------------------------

	DE18 Bortfall av HB-nett/Fencingnett og nettaksess til iDRAC6
Avvik:	Manuelt inngrep når IDRAC6-kort ikke er tilgjengelig
Diskusjon:	Dersom en node blir utsatt for en av to hendelser: 1. blir strømløs 2. mister heartbeatnett og nett til iDRAC6 managmentkort vil ikke fencing av noden vi IPMI fungere. Her er et eksempel fra test DE18 som innbærer situasjonen beskrevet i 2) ovenfor. Node 1 forsøker å fence node 2: <pre>Jan 26 10:08:20 halv2-test01 fenced[3262]: fencing node halv2-test02-hb Jan 26 10:08:40 halv2-test01 fenced[3262]: fence halv2-test02-hb dev 0.0 agent fence_ipmilan result: error from agent Jan 26 10:08:40 halv2-test01 fenced[3262]: fence halv2-test02-hb failed</pre> Dette går vil "evig tid". Servicene vil forbli på noden fordi resten av nodene ikke kan være sikre på at den har forlatt clusteret; disse nodene insisterer på å fence noden. RedHat har imidlertid lagt opp til at man kan gjøre en manuell fencing, dvs at vi som operatører kan fortelle clusteret at noden ikke lenger kan skrive til disk og dermed forårsake korrupsjon. Kjøring av manuell fencing ser slik ut: <pre>[root@halv2-test01 cluster]# fence_ack_manual halv2-test02-hb About to override fencing for halv2-test02-hb. Improper use of this command can cause severe file system damage. Continue [NO/absolutely]? absolutely Done</pre> Noden betraktes nå som nede og servicene flyttes til neste node. Merk at denne manuelle fencingen er støttet for alle typer fencing-metoder, og man trenger ikke legge inn en egen fencing-metode for manuell fencing slik som under RHEL5. Oppsettet for en node ser dermed slikt ut: <pre><clusternode name="halv2-test01-hb" nodeid="1" votes="1"> <fence> <method name="1"> <device name="halv2-test01-ipmi"/> </method> </fence> </clusternode></pre> Vi ser altså at disse situasjonene <u>ikke gir automatisk failover</u> og dermed nedetid på tjenestene. Skal man ha automatisk failover av services i denne situasjonen må man legge inn en ny sekundær fencing-metode, feks • strømmeskinne fencing • scsi persistent reservation • fc fencing Denne feilen er ingen "single failure" men heller to feil som inntreffer samtidig, så denne problemstillingen er derfor ikke regnet som kritisk.

Vi testet ingen Nagios-moduler grundig siden initielle tester gav helt åpenbart feil resultater.

Prosjektet har konkludert med at vi må skrive vår egen Nagios-modul for RHCS, og vi mener at dette skal kunne gjøres på ca 8 timer siden informasjonen fra MIB'en ikke er avskrekkende å forholde seg til.

Konklusjon

RedHat Cluster Suite (RHCS) på RHEL5 og RHEL6 tilfredsstiller våre krav godt nok i henhold til vår testplan.

Oracle DataGuard kan også brukes til å lage failover-løsninger for Oracle-databaser, men det er ikke trivielt å lage løsninger hvor klientene ikke vil oppleve tregere responsstid ved driftsforstyrrelser.

Dersom RHCS tas i bruk, bør USIT utvikle en egen Nagios-modul for RHCS.

Appendix 18 Testplan for HA-løsning

Versjon 2.0 steira 2011-01-13

Tidligere versjon: [Versjon 1.21 steira 2011-01-07](#)

Etter workshop med RedHat oppdaget vi at noen av beskrivelsene og forventningene var feil, og denne versjon 2.x av testplanen er inkludert disse endringene.

Tredelt testplan

Hensikten med testene er å verifisere at

- HA-løsningen innfrir de krav som vi stiller
- HA-løsningen støtter de tjenestene vi skal kjøre på en god måte

I testplanen er det derfor *ikke* et primært mål å verifisere at HW fungerer eller at all kabling er korrekt utført i oppsettet. Tester for å verifisere et konkret oppsett må gjøres på det konkrete oppsettet før systemet settes i drift.

I dagens SGLX kjører vi med quorum-server. Hvordan fencing og quorum løses i andre HA-løsninger er nytt for oss og derfor har vi valgt å legge spesielt vekt på testing av dette.

Vi har derfor delt testplanen i tre deler:

- funksjonelle tester med utgangspunkt i kravspesifikasjonen unntatt robusthesttester
- test av robusthet mhp fencing og quorum.
- destruktive tester med utgangspunkt i systemskisse

Vi skal lage tre pakker i clusteret:

- pakke med Oracle-database og GridControl agent
- pakke med Postgres-database
- pakke med egenutviklet tjener ("Halv2-Server") som svarer på en port og som skriver til en fil.

Testene er tenkt utført hovedsaklig med Oracle-pakken. Det må lages et testscript som kjøres fra en sqlplus-sesjon på en maskin som står utenfor clusteret. Testscriptet skal gjøre lesing og skriving i databasen.

Funksjonelle tester

I kolonnen "Krav" er det link til konkrete krav i [kravspesifikasjonen](#) som danner grunnlag for den konkrete testen.

Test#	Krav	Hvordan teste	Forventet resultat	Forventet Logging/Alarm	Faktisk resultat	Faktisk Logging/Alarm	Avvik	Status/Dato
FU1	ADM01 : Det må være støtte "rolling upgrade" av operativsystemet på nodene i HA-løsningen.	Flytt "pakke" bort fra node. Oppgrader kjerne (eller nedgrader dersom ny kjerne ikke tilgjengelig). Reboot server. Flytt "pakke" tilbake til node. Ref RedHat Knowledgebase artikkel 43210.	Node skal kunne joine cluster etter reboot. Pakke skal kunne overføres til node. Kjørende pakker på andre noder skal ikke påvirkes.	Logg på de andre nodene i clusteret om at nodene har gått inn og ut av clusteret				
FU2	ADM06 : Programvaren i HA-løsningen kan oppgraderes "rolling", dvs node for node uten at alle tjenester må tas ned	Flytt "pakke" bort fra node. Oppgraderer RCHS Flytt "pakke" tilbake til node.	Node skal kunne joine cluster etter oppgradering av RHCS på node. Pakke skal kunne starte på node. Kjørende pakker på andre noder skal ikke påvirkes.	Logg på de andre nodene i clusteret om at noden har gått inn og ut av clusteret				
	FAIL01 : Ved flytting av en tjeneste fra en node til en annen i HA-løsningen	Flytte "pakke" fra en node til en annen som følger:	Montering av filsystemer og oppstart av tjeneste skal ta	N/A				

FU3	er nedetiden for tjenesten omtrent det samme som tidendet tar å stoppe og starte ressursene i tjenesten på tilsvarende maskinvare utenfor HA-løsningen.	Stopp Oracle-pakken pakken: c lusvcadm -s rhcsoracle Start pakken på en annen node: c lusvcadm -e rhcsoracle -m <node>	"normal" tid. <u>For Oracle tjeneste:</u> Det skal være mulighet for å koble til databasen med sqlplus straks listener og instanse kjører.					
FU4	FAIL05 Feil på maskinvare i HA-løsningen skal ikke gi nedetid for brukere av tjenesten	<Gjøres under destruktive tester>						
FU5	FAIL06 Feil på maskinvare i HA-løsingen skal ikke kreve manuelle inngrep for å gjøre tjenesten tilgjengelig for brukeren.	<Gjøres under destruktive tester>						
FU6	FAIL07 Brukere av tjenesten skal ikke måtte endre i konfigurasjon på klienten etter at en tjeneste er flyttet til en annen node i HA-løsningen.	<kan besvares når FAIL01 er testet fordi det innebærer flytting av "pakke">	Klient skal uten modifikasjon kunne koble til tjeneste før og etter "pakke"-flytting.	N/A				
FU7	CONF01 Det må være mulig å konfigurere hvilke ressurser som skal være knyttet til en tjeneste Eksempler: filsystemer, ip-adresser og prosesser.	<kan besvares når vi har laget pakke for Oracle, Postgres og "HA2-service".>	CONF01	N/A				
FU8	CONF02 Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt. Eksempel på egen betingelse: Når filsystemene til pakken er utilgjengelige skal tjeneste tas ned og ikke flyttes til en annen node	<lage "pakke" med ekstra betingelse i forhold til toolkit'et. F.eks ikke kjøre tjeneste om load > 10>	CONF02	Det må være mulig å implementere logging om at pakken ikke blir startet med beskrivelse av <i>hvorfor</i>				
FU9	CONF04 Det må være mulig å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.	Legge inn oppstart av GridControl-agent i Oracle-"pakke".	Agenten startes når "pakken" startes og avsluttes når "pakken" stoppes.	Vi må kunne legge inn skriving til logg om at GridControl agent har blitt startet eller at den er blitt stoppet.				
FU10	CONF05 Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned.	(Hvilke parametre vi kunne teste med ble lagt til i versjon 2.x av denne planen) Endre force_umount fra 1 til 0 i fs-ressursen til en service.		N/A				
FU11	CONF07 Det må være mulig å spesifisere hvilke noder en tjeneste skal kunne gå på, og rekkefølgen for hvordan tjenesten skal gjøre failover til andre noder.	Lage pakke med Oracle tjeneste med angitt failover-rekkefølge. Ta ned noden og verifiser at pakken feiler over til korrekt node.	CONF07	N/A				

FU12	CONF08 Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder	Legg inn at Postgres-pakken ikke skal kjøre på samme node som Oracle-pakken, men på samme node som "HALv2-server"-pakken.		Loggmelding om at en avhengighet var årsaken til at pakken ble startet på en gitt node.				
FU13	CONF09 HA-løsningen kan konfigureres slik at den restarter ressurser i en tjeneste som den har detektert at ikke lenge kjører	Kill alle oracle-prosessene til Oracle-tjeneste.	Oracle-instanse blir startet opp igjen og klient kan koble til. Oracle blir ikke forsøkt startet opp flere ganger dersom oppstart feilet.k	Det må logges at ressursen er nede og at den ble restartet. Status for restarten må også logges.				
FU14	FAKE01 (ikke i kravspesikasjon) HA-løsningen kan automatisk faile tilbake "pakker" når noden igjen er tilgjengelig etter en failover.	Konfigure pakke for failback. Ta strøm av node som kjører pakken. Verfisiere at "pakke" flytter til annen node. Ta opp node igjen. Verfisiere at "pakke" flytter tilbake til original node.	<som beskrevet i FAKE01>	Det må logges at pakken ble flyttet, og <i>hvorfor</i> den ble flyttet. Det må fremgå at det var en automatisk failback som var årsaken til flyttingen.				
FU15	MON01 Det må enkelt å få ut info om status for HA-løsningen, dets noder og dets tjenester	Kjøre CLI-kommando for oversikt over noder og "pakker". Åpne GUI og vise status over clusteret.	Oversiktene fra kommandoene gir "tilstrekkelig" informasjon om clusteret.	N/A				
FU16	MON02 Logging må være tilstrekkelig slik at det er mulig å feilsøke ved eventuelle problemer.	<Besvare etter destruktive tester med eksempler>	MON02	N/A				
FU17	MON03 Det er ønskelig at HA-løsningen kan sende sine logger fra alle noder til et sentralisert sted	<Besvare etter destruktive tester med eksempler>	MON03	N/A				
FU18	MON04 HA-løsningen må la seg overvåke av Nagios	Finne eksisterende modul eller lage ny modul som overvåker clusteret.	Nagios alarm når • <se forventet alarm for alle testene>	N/A				
FU19	MON05 Oracle-databaser som kjører i HA-løsningen må la seg overvåke av GridControl.	<besvares når CONF04 er utført>	MON05	N/A				
FU20	MON06 Det må være mulig å få generert en alarm dersom en ressurs i tjenesten ikke kjører	<besvares når MON04> er utført>	MON06	Alarm i Nagios				
FU21	SCA01 Det må være mulig å melde noder inn i og ut av HA-løsningen uten at det påvirker tjenestene i løsningen (med unntak av tjenester som går på noder som skal fjernes).	Både for 2 og 3 node. Også teste om antall noder påvirker fencing; må hele cluster ned for å legge til flere noder i fencing-mekanisme?	SCA01	Logg fra de andre nodene om at en node går inn og ut av clusteret.				

	SCA02	Lage en ny pakke.	SCA02	N/A				
FU22	Det må være mulig å definere nye tjenester i løsningen uten at det krever nedetid på andre tjenester i løsningen.							

Test av robusthet; fencing og quorum

Hensikten med disse testene er å avgjøre hvilke typer quorum og fencing som fungerer best og som er enkle å drifte.

Kanhende vil noen metoder for fencing bli utelukket før testing dersom vi leser dokumentasjon og konkluderer med atat metodene er uhensiktsmessige.

Testene skal gjøre for både 2 og 3 noder i clusteret.

Gunnar lager utkast til mail for å sende til RH for å spørre om

- gode måter å gjøre quorum og fencing
- hva de fleste kunder gjør

Vi må også finne svar på om lvm2-tagging er i bruk av toolkitene.

To node cluster:

Test#	Krav	Hvordan teste	Antall ganger gj.ført	Forventet resultat OS	Forventet resultat HA-løsning	Forventet Logging/Alarm	Faktisk resultat OS	Faktisk resultat HA-løsning	Faktisk Logging/Alarm	Avvik	Status/Dato
RO01	ROB03 : Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).	Dra ut alle nettverkskabler til HB-nettet. Kill prosessen til clusterkommunikasjonen: aisexec			Noden skal bli fenced og pakken skal ikke kunne starte på en annen node før noden er nede.						
RO02	ROB04 : Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på to noder samtidig).	a) Monter et av filsystemene som er montert opp i en pakken på nodeX manuelt på nodeY. b) Ta opp en pakke på nodeY som allerede kjører på nodeX. ~~c) Irrelevant test c) i versjon 1.x er fjernet i versjon 2.x av testplanen~~		a) Filsystemet lar seg ikke montere pga lvm2-tags	b) HA-løsningen vil ikke la pakken starte på nodeY og man får feilmelding om at den kjører på nodeX	a) Ingen b) Ingen c) Ingen	c)				
RO03	ROB05 : Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i	<besvares når ROB03 er utført fordi fencing>									

	clusteret (fencing)										
--	------------------------	--	--	--	--	--	--	--	--	--	--

Tre node cluster

Test#	Krav	Hvordan teste	Antall ganger gj.ført	Forventet resultat OS	Forventet resultat HA-løsning	Forventet Logging/Alarm	Faktisk resultat OS	Faktisk resultat HA-løsning	Faktisk Logging/Alarm	Avvik	Status/Dato
RO04	ROB03 : Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).	Dra ut alle nettverkskabler til HB-nettet. Kill prosessen til clusterkommunikasjonen aisexec			Noden skal bli fenced og pakken skal ikke kunne starte på en annen node før noden er nede.						
RO05	ROB04 : Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på to noder samtidig).	a) Monter et av filsystemene som er montert opp i en pakken på nodeX manuelt på nodeY. b) Ta opp en pakke på nodeY som allerede kjører på nodeX. c) Irrelevant test c) i versjon 1.x er fjernet i versjon 2.x av testplanen		a) Filsystemet lar seg ikke montere pga lvm2-tags .	b) HA-løsningen vil ikke la pakken starte på nodeY og man får feilmelding om at den kjører på nodeX	a) Ingen b) Ingen					
RO06	ROB05 : Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing)	<besvares når ROB03 er utført fordi fencing>									

Destruktive tester

Tester som innebærer reboot eller å ta strømmen av fc-switcher kan ikke utføres fordi vi ikke kjører på dedikerte fc-switcher uten annen trafikk.

Test#	Hendelse	Hvordan teste	Antall ganger gj. ført	Forventet resultat OS	Forventet resultat HA-løsning	Forventet Logging/Alarm	Faktisk resultat OS	Faktisk resultat HA-løsning	Faktisk Logging/Alarm	Avvik	Status/Dato
	FC										
	Bortfall av en redundant fc-	frakoble fc-kabel på node		fc-failover	Ikke påvirket	Logg bortfall av fc-path.					

DE01	path på node					Nagios-alarm bortfall av fc- path.					
~~DE02~~	-	reboot av fc- switch	-	Muligens Ikke påvirket fordi fc- link beholdes under reboot.	Ikke påvirket	Muligens ingenting?	-	-	-	-	-
DE03	-	ta strøm fra fc- switch	-	fc- failover	Ikke påvirket	Logg bortfall av fc-path. Nagios-alarm bortfall av fc- path.	-	-	-	-	-
DE04	Bortfall all fc for en node	frakoble begge fc-kabler på node Verifiser at pakkeflytting skyldes bortfall av device.		fc-failure	Pakke flyttes til annen node	Logg bortfall av fc-pather Nagios-alarm bortfall av fc- pather Nagios-alarm om at pakke er flyttet					
DE05	Bortfall all fc i clusteret	frakoble begge fc-kabler på alle nodene		fc-failure	Alle pakker i cluster tas ned.	Logg bortfall av fc-pather Nagios-alarm bortfall av fc- pather Nagios-alarm om at pakker er nede.					
~~DE06~~	-	reboot av begge fc-switchene	-	Muligens Ikke påvirket fordi fc- link beholdes under reboot.	Muligens Ikke påvirket fordi fc-link beholdes under reboot.	Muligens ingenting?	-	-	-	-	-
~~DE07~~	-	ta strøm fra begge fc-switchene	-	fc-failure	Alle pakker i cluster tas ned.	Logg bortfall av alle fc-pather Nagios-alarm bortfall av alle fc-pather Nagios-alarm om at pakker er nede.	-	-	-	-	-
~~DE08~~	-	ta ned diskarray	-	fc-failure	Alle pakker i cluster tas ned.	Logg bortfall av alle fc-path. Nagios-alarm bortfall av alle fc-pather Nagios-alarm om at pakker er nede.	-	-	-	-	-
	Nett										

	HearBeat=HB										
DE09	Bortfall av en redundant nettverkspath i HB-nett /Fencing nett	Trekke ut nettverkskabel		failover på bonding hvis aktiv path.	Ingen	Logg om nettpath som nede. Nagios alarm om nettpath som er nede					
DE10	Bortfall av en redundant nettverkspath i public-nett	Trekke ut nettverkskabel		failover på bonding hvis aktiv path.	Ingen	Logg om nettpath som er nede. Nagios alarm om nettpath som er nede.					
DE11	Bortfall av HB-nett/Fencing nett (iDRAC6 kort fortsatt tilgjengelig)	Trekke ut alle kablene fra noden for HB-nett/Fencing nett		Melding om at nett er nede.	<Vi vet ikke> ??? Dersom noden kjører pakke så må noden fences før pakkene kan overføres til en annen node.	Logg om nettpather som nede. Nagios alarm om nettpath som er ende. Logg fra RHCS om at HB-nett er nede. Nagios Alarm om at HB-nett er nede.					
DE18	Bortfall av HB-nett/Fencingnett og nettaksess til iDRAC6	Trekke ut alle kablene fra noden for HB-nett/Fencing nett Trekke ut nettverkskabel til iDRAC6		Melding om at nett er nede.	<Vi vet ikke> ??? Dersom noden kjører pakke så må noden fences før pakkene kan overføres til en annen node.	Logg om nettpather som nede. Nagios alarm om nettpath som er ende. Logg fra RHCS om at HB-nett er nede. Nagios Alarm om at HB-nett er nede.					
DE12	Bortfall av public nett	Trekke ut begge kablene til public-nettet.		Melding i syslog	Pakker skal bli tatt ned og flyttet til annen noden fordi ip-nr i pakken ikke lenger vil fungere.	Logg om nettpath som nede. Nagios alarm om at noden ikke svarer på ping. Logg om at pakker flyttes.					
~~DE13~~	~~Bortfall av alle nett-på noden~~	~~Trekke ut alle nettverkskablene.~~	~~-~~	~~Melding i syslog.~~	~~<Vi vet ikke> ??? Dersom noden kjører pakke så må noden fences før pakkene kan overføres til en annen node.~~	~~Logg om nettpather som nede. Andre noder logger at noden er utilgjengelig og at den tilslutt blir fenced. Nagios alarm om at maskinen er nede (etter at den er fenced) Nagios alarm~~	~~-Redundant test, se DE11 og DE12.~~	~~-~~	~~-~~	~~-~~	Utgår

						om at noden er blitt fenced					
	Node										
DE14	Node failure	Frakoble alle strømkabler			Alle "pakker" som kjører på noden flyttes automatisk til neste node i cluster iht config Tjenesten er tilgjengelig igjen når "pakkene" er ferdig med oppstart.	Andre noder i clusteret logger at noden er nede. Nagiosalarm om at noden ute av clusteret.					
DE15		Kjøre shutdown på noden			Som ovenfor	Som ovenfor					
DE16		Generere panic på noden <Trond finne ut hvordan>			Som ovenfor	Som ovenfor					
	iLo / DRAC										
DE17	Oppdatering eller reset av iLO/DRAC	Hva skjer dersom vi oppdaterer firmware eller resetter på iLO/DRA? Dette kan også simuleres ved å reboot/resette iDRAC6 på noden			Forventer at intet skjer selv om vi bruker fencing via iLO/DRAC	N/A					

Appendix 19 Testing av oracle dataguard

Testoppsett:

Primary database GLADTST på maskin nodeprimary.

Standby database GLADSBY1 på maskin nodestandby.

TNS entry GLADSMASK.uio.no opprettet slik:

```
GLADSMASK.uio.no=
(DESCRIPTION_LIST=(LOAD_BALANCE=off)(FAILOVER=on)
(DESCRIPTION=(
ADDRESS_LIST=(ADDRESS=(PROTOCOL=TCP)(HOST=nodeprimary.uio.no)(PORT=1531)))
(CONNECT_DATA=(SERVICE_NAME=GLADTST.uio.no)))
(DESCRIPTION=(
ADDRESS_LIST=(ADDRESS=(PROTOCOL=TCP)(HOST=nodestandby.uio.no)(PORT=1521)))
(CONNECT_DATA=(SERVICE_NAME=GLADSBY1.uio.no))))
```

Følgende tester ble laget

Nr	Test	Forventet resultat	Faktisk resultat
1	GLADTST og -SBY1 er oppe. Sqlplus system@gladsmask.uio.no	Ok	Som forventet
2	GLADTST er standby, GLADSBY1 er ny primary etter switchover. Sqlplus system@gladsmask.uio.no	Feilmelding, ok på 2 forsøk	Som forventet
3	Som 2, men GLADTST instansen er nede.	Man går rett inn på GLADSBY1	Som forventet
4	Som 2, men noden GLADTST kjører på er nede (simulert ved å skru av pakken stoltstst)	TCP heng mot GLADTST, deretter ok mot GLADSBY1	Som forventet - 1 min heng
5			

Oppsett ble gjennomført via GRIDCONTROL innlogget som SYSMAN, og SYS AS SYSDBA mot instansene, og mot ORACLE bruker med passord på nodene nodestandby og nodeprimary.

Om dette skal brukes må man konfigurere automatisk failover med feks. observer - verktøyet.

Output fra testing

```
DGMGRL> show configuration
Configuration - GLADTST.uio.no
  Protection Mode: MaxPerformance
  Databases:
    GLADTST - Primary database
    GLADSBY1 - Physical standby database

Fast-Start Failover: DISABLED

Configuration Status:
SUCCESS

-----
testuser@nodestandby ~ $ rs system@gladtst
SQL*Plus: Release 10.2.0.3.0 - Production on Tue Dec 14 12:59:09 2010
Copyright (c) 1982, 2006, Oracle. All Rights Reserved.
Enter password: *****
Connected to:
```

Oracle Database 11g Enterprise Edition Release 11.2.0.2.0 - 64bit Production
With the Partitioning, OLAP, Data Mining and Real Application Testing options

```
SQL> create table lfctest(a number);  
  
Table created.  
  
SQL> insert into lfctest values( 1);  
  
1 row created.  
  
SQL> commit;  
  
Commit complete.  
  
SQL> select * from lfctest;
```

```
-----  
A  
-----  
1
```

SQL>

Switcher så over til gladsby1 med å klikke på switchover

```
SYS@GLADSBY1> conn system  
Enter password:  
Connected.
```

```
GLOBAL_NAME  
-----  
GLADTST.UIO.NO
```

```
SYSTEM@GLADSBY1> select * from lfctest;
```

```
-----  
A  
-----  
1
```

Testcase 1

```
-----  
GLADSBY1@nodestandby:~[12] time sqlplus system@gladsmask  
  
SQL*Plus: Release 11.2.0.2.0 Production on Wed Dec 15 13:08:08 2010  
  
Copyright (c) 1982, 2010, Oracle. All rights reserved.  
  
Enter password:
```

Connected to:
Oracle Database 11g Enterprise Edition Release 11.2.0.2.0 - 64bit Production
With the Partitioning, OLAP, Data Mining and Real Application Testing options

```
GLOBAL_NAME  
-----  
GLADTST.UIO.NO
```

```
SYSTEM@gladsmask> Disconnected from Oracle Database 11g Enterprise Edition Release 11.2.0.2.0 -  
64bit Production  
With the Partitioning, OLAP, Data Mining and Real Application Testing options
```

```
real      0m4.661s  
user      0m0.029s  
sys       0m0.010s  
GLADSBY1@nodestandby:~[13]
```

Testcase 2 - har tatt switchover til GLADSBY1:

```
-----  
GLADSBY1@nodestandby:~[60] sqlplus system@gladsmask  
  
SQL*Plus: Release 11.2.0.2.0 Production on Tue Dec 14 14:54:53 2010  
  
Copyright (c) 1982, 2010, Oracle. All rights reserved.  
  
Enter password:  
ERROR:
```

ORA-01033: ORACLE initialization or shutdown in progress
Process ID: 0
Session ID: 0 Serial number: 0

Enter user-name: system
Enter password:

Connected to:
Oracle Database 11g Enterprise Edition Release 11.2.0.2.0 - 64bit Production
With the Partitioning, OLAP, Data Mining and Real Application Testing options

GLOBAL_NAME

GLADTST.UIO.NO

SYSTEM@GLADSBY1>

Dette er reproducerbart. Dvs prøver den alltid nr.1 først.

Testcase 3

GLADSBY1@nodestandby:~[6] sqlplus system@gladsmask

SQL*Plus: Release 11.2.0.2.0 Production on Wed Dec 15 10:55:48 2010

Copyright (c) 1982, 2010, Oracle. All rights reserved.

Enter password:

Connected to:
Oracle Database 11g Enterprise Edition Release 11.2.0.2.0 - 64bit Production
With the Partitioning, OLAP, Data Mining and Real Application Testing options

GLOBAL_NAME

GLADTST.UIO.NO

SYSTEM@gladsmask>

Testcase 4

GLADSBY1@nodestandby:~[11] time sqlplus system@gladsmask

SQL*Plus: Release 11.2.0.2.0 Production on Wed Dec 15 11:19:51 2010

Copyright (c) 1982, 2010, Oracle. All rights reserved.

Enter password:

Connected to:
Oracle Database 11g Enterprise Edition Release 11.2.0.2.0 - 64bit Production
With the Partitioning, OLAP, Data Mining and Real Application Testing options

GLOBAL_NAME

GLADTST.UIO.NO

SYSTEM@gladsmask> Disconnected from Oracle Database 11g Enterprise Edition Release 11.2.0.2.0 -
64bit Production
With the Partitioning, OLAP, Data Mining and Real Application Testing options

real 1m2.316s
user 0m0.020s
sys 0m0.016s
GLADSBY1@nodestandby:~[12]

Appendix 20 Kravspesifikasjon for Nagios-modul RHCS

Versjon 1.0 steira 2010-12-17

Innledning

Kravene til Nagios-modulen er avledet fra "forventet resultat" fra vår testplan

- [Appendix 18 Testplan for HA-løsning](#)

Krav til Nagiosmodul

Ordet "Pakke" nedenfor erstattes med det ordet vi velger å bruke for services i RHCS.

Krav#	Melding i Nagios når...	Krav/Ønske
N01	en service blir stoppet på en node: "Service XXX ble stoppet på node YYY i clusteret CCC"	Utgår
N02	en service blir startet til en node: "Service XXX er startet på node YYY i clusteret CCC"	Utgår
N03	en ressurs i en service ikke kjører: F.eks: "Ressurs ZZZ kjører ikke lenger i servicen XXX i clusteret CCC"	Krav
N04	en service er nede: Feks: "Service XXX er nede i clusteret CCC"	Krav
N05	en service kjører på annen node en primary: "Service XXX kjører på noden YYY som ikke er primary node for servicen i clusteret CCC"	Utgår ¹
N06	en node ikke er med i clusteret: F. eks: "Node YYY er ikke lenger med i clusteret CCC"	Krav

Fotnote

¹⁾ ServiceGuard sin Nagios modul hadde mulighet for å sette op monitorering av pakker som ikke kjørte på primærnode, men man erfarte at dette lage mye støy så denne funksjonaliteten har aldri vært tatt i bruk. Men ønsker imidlertid at services i RHCS sender email når pakker stopper og starter slik at man om nødvendig selv kan undersøke hva som var årsaken til pakkeflyttingen.

Appendix 21 Testresultater RHCS

V 1.2
Basert på versjon 1.21 av A07.03 Testplan.

Tredelt testplan

Hensikten med testene er å verifisere at

- HA-løsningen innfrir de krav som vi stiller
- HA-løsningen støtter de tjenestene vi skal kjøre på en god måte

I testplanen er det derfor *ikke* et primært mål å verifisere at HW fungerer eller at all kabling er korrekt utført i oppsettet. Tester for å verifisere et konkret oppsett må gjøres på det konkrete oppsettet før systemet settes i drift.

I dagens SGLX kjører vi med quorum-server. Hvordan fencing og quorum løses i andre HA-løsninger er nytt for oss og derfor har vi valgt å legge spesielt vekt på testing av dette.

Vi har derfor delt testplanen i tre deler:

- funksjonelle tester med utgangspunkt i kravspesifikasjonen unntatt robusthetstester
- test av robusthet mhp fencing og quorum.
- destruktive tester med utgangspunkt i systemskisse

Vi skal lage tre pakker i clusteret:

- pakke med Oracle-database og GridControl agent
- pakke med Postgres-database
- pakke med egenutviklet tjener ("Halv2-Server") som svarer på en port og som skriver til en fil.

Testene er tenkt utført hovedsaklig med Oracle-pakken. Det må lages et testscript som kjøres fra en sqlplus-sesjon på en maskin som står utenfor clusteret. Testscriptet skal gjøre lesing og skriving i databasen.

Funksjonelle tester

I kolonnen "Krav" er det link til konkrete krav i [kravspesifikasjonen](#) som danner grunnlag for den konkrete testen.

Test#	Krav	Hvordan teste	Forventet resultat	Forventet Logging/Alarm	Faktisk resultat	Faktisk Logging/Alarm	Avvik	Status/Dato
FU1	ADM01: Det må være støtte "rolling upgrade" av operativsystemet på nodene i HA-løsningen.	Planlagt test: Flytt "pakke" bort fra node. Oppgrader kjerne (eller nedgrader dersom ny kjerne ikke tilgjengelig). Reboot server. Flytt "pakke" tilbake til node. Ref RedHat Knowledgebase artikkel 43210. Gjennomført test: Vi flyttet rhcsmisc pakken fra node1 til node3. Så kjørter vi "yum -y update" på node1 som gav oppgradering fra RHEL5.5 til RHEL5.6	Node skal kunne joine cluster etter reboot. Pakke skal kunne overføres til node. Kjørende pakker på andre noder skal ikke påvirkes.	Logg på de andre nodene i clusteret om at nodene har gått inn og ut av clusteret	Som forventet.	N/A		OK 14/01
FU2	ADM06: Programvaren i HA-løsningen kan oppgraderes "rolling", dvs node for node uten at alle tjenester må tas ned	Planlagt test: Flytt "pakke" bort fra node. Oppgraderer RCHS Flytt "pakke" tilbake til node. Gjennomført test: Vi flyttet rhcsmisc pakken fra node1 til node3. Så kjørter vi "yum -y update" på node1 som faktisk gav RHEL5.6 og ny clusterprogramvare.	Node skal kunne joine cluster etter oppgradering av RHCS på node. Pakke skal kunne starte på node. Kjørende pakker på andre noder skal ikke påvirkes.	Logg på de andre nodene i clusteret om at noden har gått inn og ut av clusteret	Som forventet.	N/A	Oppgraderte ressursscript	OK 14/01
	FAIL01: Ved flytting av en	Flytte "pakke" fra en node til en annen som følger:	Montering av filsystemer og oppstart av	N/A	Som forventet			OK 10/1

FU3	tjeneste fra en node til en annen i HA-løsningen er nedetiden for tjenesten omtrent det samme som tiden det tar å stoppe og starte ressursene i tjenesten på tilsvarende maskinvare utenfor HA-løsningen.	Stopp Oracle-pakken pakken: <pre>clusvcadm -s rhcsoracle</pre> Start pakken på en annen node: <pre>clusvcadm -e rhcsoracle -m <node></pre>	tjeneste skal ta "normal" tid. <u>For Oracle tjeneste:</u> Det skal være mulighet for å koble til databasen med sqlplus straks listener og instanse kjører.					
FU4	FAIL05 Feil på maskinvare i HA-løsningen skal ikke gi nedetid for brukere av tjenesten	<Gjøres under destruktive tester>						
FU5	FAIL06 Feil på maskinvare i HA-løsningen skal ikke kreve manuelle inngrep for å gjøre tjenesten tilgjengelig for brukeren.	<Gjøres under destruktive tester>						
FU6	FAIL07 Brukere av tjenesten skal ikke måtte endre i konfigurasjon på klienten etter at en tjeneste er flyttet til en annen node i HA-løsningen.	<kan besvares når FAIL01 er testet fordi det innebærer flytting av "pakke">	Klient skal uten modifikasjon kunne koble til tjeneste før og etter "pakke"-flytting.	N/A	Som forventet			OK 10/1
FU7	CONF01 Det må være mulig å konfigurere hvilke ressurser som skal være knyttet til en tjeneste Eksempler: filesystemer, ip-adresser og prosesser.	<kan besvares når vi har laget pakke for Oracle, Postgres og "HA2-service".	CONF01	N/A	Som forventet			OK 10/1
FU8	CONF02 Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt. Eksempel på egen betingelse: Når filesystemene til pakken er utilgjengelige skal tjeneste tas ned og ikke flyttes til en annen node	<u>Planlagt test:</u> <lage "pakke" med ekstra betingelse i forhold til toolkit'et. F.eks ikke kjøre tjeneste om load > 10 <u>Gjennomført test:</u> Vi laget et script rhcsmisc1-server-req.sh som kun gav returkode 0 dersom filen /etc/uscit/run-half2-server eksisterte når det ble kalt opp med argumentet status eller start. Dersom denne filen ikke finnes, gir dette returkode 150 . Argumentet stop gir alltid returkode 0. I servicen rhcsmisc ble det laget en script-ressurs med scriptet rhcsmisc1-server-req.sh som er helt øverst av alle ressursene til servicen i cluster.conf, dvs at alle andre ressurser er avhengig av at dette scriptet kjører og returnerer feilkode 0. Da kan man under oppstart av servicen (da kalles scriptet opp med start) og under kjøring av servicen (da kalles scriptet opp med status) kontrollere hele servicen hva	CONF02	Det må være mulig å implementere logging om at pakken ikke blir startet med beskrivelse av <i>hvorfor</i>	Denne funksjonaliteten er ikke støttet out-of-the box. <u>Kravet er derfor ikke tilfredstillt slik det står formulert,</u> men etter diskusjon i gruppen har vi funnet at kravet er <i>feilformulert</i>. Det kravet egentlig formulerer er at ikke alle ressurser som feiler skal måtte medføre at servicen flyttes eller restartes, og dette lar seg enkelt implementere i RHCS ved å endre i ressursscriptene.	Som forventet.	Ikke støttet "out of the box"	OK med forbehold om at kravets ordlyd var upresist. 13/1

		filen /etc/usit/run-halv2-server						
FU9	CONF04 Det må være mulig å gjøre tilpassinger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.	Legge inn oppstart av GridControl-agent i Oracle-"pakke".	Agenten startes når "pakken" startes og avsluttes når "pakken" stoppes.	Vi må kunne legge inn skriving til logg om at GridControl agent har blitt startet eller at den er blitt stoppet.	Som forventet	Implementerte ikke skriving til logg om oppstart av agent, men dette er enkelt å få til.		OK 10/1
FU10	CONF05 Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned.	I servicen rhcsmisc:Endre force_umount fra 1 til 0 i fs-ressursen.		N/A	Ikke som forventet.		Restart av ressurser og tjenester uten forvarsel	FAILED 11/1. (men CONF05 der et <i>ønske</i> og ikke et <i>krav</i>)
FU11	CONF07 Det må være mulig å spesifisere hvilke noder en tjeneste skal kunne gå på, og rekkefølgen for hvordan tjenesten skal gjøre failover til andre noder.	Lage pakke med Oracle tjeneste med angitt failover-rekkefølge. Ta ned noden og verifiser at pakken feiler over til korrekt node.	CONF07	N/A	Som forventet.			12/02 OK
FU12	CONF08 Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder	Legg inn at Postgres-pakken ikke skal kjøre på samme node som Oracle-pakken, men på samme node som "HALv2-server"-pakken.		Loggmelding om at en avhengighet var årsaken til at pakken ble startet på en gitt node.	Delvis innfridd.	Som forventet	Begrenset støtte, avanserte avhengigheter må lages selv. Men SAP er støttet!	OK 11/01
FU13	CONF09 HA-løsningen kan konfigureres slik at den restarter ressurser i en tjeneste som den har detektert at ikke lenge kjører	Kill alle oracle-prosessene til Oracle-tjeneste.	Oracle-innstanse blir startet opp igjen og klient kan koble til. Oracle blir ikke forsøkt startet opp flere ganger dersom oppstart feilet.k	Det må logges at ressursen er nede og at den ble restartet. Status for restarten må også logges.	Som forventet, Oracle-scriptet restarter instansen. Det blir logget at instansen blir restartet, men ikke at den var nede: Jan 10 14:31:18 halv2-test02 oracledb.sh.USIT: Restarting Oracle Database...			OK 10/1
FU14	FAKE01 (ikke i kravspesikasjon) HA-løsningen kan automatisk faile tilbake "pakker" når noden igjen er tilgjengelig etter en failover.	Konfigure failoverdomain for failback. Ta shutdown av node som kjører pakken. Verfisiere at "pakke" flytter til annen node. Ta opp node igjen. Verfisiere at "pakke" flytter tilbake til original node.	<som beskrevet i FAKE01>	Det må logges at pakken ble flyttet, og <i>hvorfor</i> den ble flyttet. Det må fremgå at det var en automatisk failback som var årsaken til flyttingen.	Som forventet, nofailback=0 gir dette.			OK 10/1
FU15	MON01 Det må enkelt å få ut info om status for HA-løsningen, dets noder og dets tjenester	Kjøre CLI-kommando for oversikt over noder og "pakker". Åpne GUI og vise status over clusteret.	Oversiktene fra kommandoene gir "tilstrekkelig" informasjon om clusteret.	N/A	Som forventet			OK 10/1
FU16	MON02 Logging må være tilstrekkelig slik at det er mulig å feilsøke ved eventuelle problemer.	<Besvare etter destruktive tester med eksempler>	MON02	N/A				OK 11/1

FU17	MON03 Det er ønskelig at HA-løsningen kan sende sine logger fra alle noder til et sentralisert sted	<Besvare etter destruktive tester med eksempler>	MON03	N/A	Som forventet. Det logges til syslog og dermed kan dette sendes videre.			10/1 OK
FU18	MON04 HA-løsningen må la seg overvåke av Nagios	Finne eksisterende modul eller lage ny modul som overvåker clusteret.	Nagios alarm når <se forventet alarm for alle testene> 2011-01-11 : Ferdige nagios-moduler er ikke gode nok. Må lage egen	N/A	Som forventet. Det er mulig å overvåke fordi vi får fornuftig info fra snmp og evt clustat. Eksisternede Nagios-moduler gir informasjon om RHCS, men den er ikke god nok så vi må nok lage egen modul.			OK 11/1
FU19	MON05 Oracle-databaser som kjører i HA-løsningen må la seg overvåke av GridControl.	<besvares når CONF04 er utført>	MON05	N/A	Som forventet.			OK 10/1
FU20	MON06 Det må være mulig å få generert en alarm dersom en ressurs i tjenesten ikke kjører	<besvares når MON04> er utført> <Sjekke med RedHat hvordan tilstand til ressurser i service overvåkes. Alarm kan generes fra init-scriptets status-del>	MON06	Alarm i Nagios	Som forventet. Generering av alarm lages i scriptet til ressursen, ikke støtte for dette direkte. Oracle-ressursscriptet har kode for restart av oracle i "status"-delen av scriptet, og dette viser at man kan lage alarmer eller logging på denne måten.			OK 10/1
FU21	SCA01 Det må være mulig å melde noder inn i og ut av HA-løsningen uten at det påvirker tjenestene i løsningen (med unntak av tjenester som går på noder som skal fjernes).	Både for 2 og 3 node. Også teste om antall noder påvirker fencing; må hele cluster ned for å legge til flere noder i fencing-mekanisme?	SCA01	Logg fra de andre nodene om at en node går inn og ut av clusteret.	Det er ikke mulig å øke antall noder fra 2 til 3 noder mens clusteret er oppe siden 2-node cluster er en egen type cluster med andre quorum-regler. Vi trenger derfor 4 noder for å få testet dette. Testen UTGÅR.	-	-	UTGÅR
FU22	SCA02 Det må være mulig å definere nye tjenester i løsningen uten at det krever nedetid på andre tjenester i løsningen.	Lage en ny pakke.	SCA02	N/A	Som forventet men med komplikasjoner:		Komplikasjoner	OK 11/1

Test av robusthet; fencing og quorum

Hensikten med disse testene er å avgjøre hvilke typer quorum og fencing som fungerer best og som er enkle å drifte.

Kanhende vil noen metoder for fencing bli utelukket før testing dersom vi leser dokumentasjon og konkluderer med at metodene er uhensiktsmessige.

Testene skal gjøre for både 2 og 3 noder i clusteret.

Gunnar lager utkast til mail for å sende til RH for å spørre om

- gode måter å gjøre quorum og fencing
- hva de fleste kunder gjør

Vi må også finne svar på om 2-tagging er i bruk av toolkitene.

To node cluster:

Overgang fra tre noder til to noder.

- 1) Ta backup av cluster.conf (cluster.conf-3-node)
- 2) /site/bin/stop-node på alle noder for å melde node ut av cluster
- 3) Fjerne node1 fra cluster.conf (husk failoverdomains)
- 4) Legg inn parameteren two_node=1
- 5) Start opp cluster på gjenstående noder /site/bin/start-node
- 6) Dist ny config med ccs_tool_update
- 7) Node1 får status "Offline, Estranged". Reboote clusteret.

FIXME: Doc scriptene et sted

Test#	Krav	Hvordan teste	Antall ganger gj.ført	Forventet resultat OS	Forventet resultat HA-løsning	Forventet Logging/Alarm	Faktisk resultat OS	Faktisk resultat HA-løsning	Faktisk Logging/Alarm	Avvik	Status/Dato
RO01	ROB03: Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).	a) Dra ut alle nettverkskabler til HB-nettet. b) Kill prosessen til clusterkommunikasjonen aisexec c) Gi noden en panic echo c > /proc/sysrq-trigger			Noden skal bli fenced og pakken skal ikke kunne starte på en annen node før noden er nede.		.	a) Som forventet b) Som forventet c) Som forventet			OK 14/01
2	ROB04: Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på to noder samtidig).	a) Monter et av filsystemene som er montert opp i en pakken på nodeX manuelt på nodeY. b) Ta opp en pakke på nodeY som allerede kjører på nodeX.	-	a) Filsystemet lar seg ikke montere pga lvm2-tags	b) HA-løsningen vil ikke la pakken starte på nodeY og man får feilmelding om at den kjører på nodeX	a) Ingen b) Ingen	-	-	-Testet for 3-node cluster	-	
RO03	ROB05: Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing)	<besvares når ROB03 er utført fordi fencing>							JA.		OK 14/01

Tre node cluster

Test#	Krav	Hvordan teste	Antall ganger gj.ført	Forventet resultat OS	Forventet resultat HA-løsning	Forventet Logging/Alarm	Faktisk resultat OS	Faktisk resultat HA-løsning	Faktisk Logging/Alarm	Avvik	Status/Dato
	ROB03: Løsningen må ha pålitelige	a) Dra ut alle nettverkskabler til HB-nettet.			Noden skal bli fenced og pakken skal ikke			a) Allerede testet i destruktive			OK 11/1

RO04	mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).	b) Kill prosessene til clusterkommunikasjonen aisexec			kunne starte på en annen node før noden er nede.			tester DE11 b) Som forventet.			
RO05	ROB04: Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på to noder samtidig).	a) Monter et av filsystemene som er montert opp i en pakken på nodeX manuelt på nodeY. b) Ta opp en pakke på nodeY som allerede kjører på nodeX.		a) Filsystemet har seg ikke montere pga lvm2-tags	b) HA-løsningen vil ikke la pakken starte på nodeY og man får feilmelding om at den kjører på nodeX	a) Ingen b) Ingen	a) Som forventet fordi vi ikke får aktivert volumgruppen, men vi kan jo selv sette tags og omgå dette om vi vil. b) Som forventet: Får meldingen "Service is already running"	Som forventet.			OK 11/1
RO06	ROB05: Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing)	<besvares når ROB03 er utført fordi fencing>									OK 11/1

Destruktive tester

Tester som innebærer reboot eller å ta strømmen av fc-switcher kan ikke utføres fordi vi ikke kjører på dedikerte fc-switcher uten annen trafikk.

Test#	Hendelse	Hvordan teste	Antall ganger gj. ført	Forventet resultat OS	Forventet resultat HA-løsning	Forventet Logging/Alarm	Faktisk resultat OS	Faktisk resultat HA-løsning	Faktisk Logging/Alarm	Avvik	Status/Dato
	FC										
DE01	Bortfall av en redundant fc-path på node	frakoble fc-kabel på node		fc-failover	Ikke påvirket	Logg bortfall av fc-path. Nagios-alarm bortfall av fc-path.	Som forventet	Som forventet			OK 10/1
DE02	-	reboot av fc-switch	-	Muligens Ikke påvirket fordi fc-link beholdes under reboot.	Ikke påvirket	Muligens ingenting?	-	-	-	-	-
DE03	-	ta strøm fra fc-switch	-	fc-failover	Ikke påvirket	Logg bortfall av fc-path. Nagios-alarm bortfall av fc-path.	-	-	-	-	-
	Bortfall all fc for en node	frakoble begge fc-kabler på		fc-failure	Pakke flyttes til	Logg bortfall av fc-pather	Som forventet.	Servicene blir satt i "failed" og	Som forventet,	Manuelt inngrep	OK 11/1

DE04		node Verifiser at pakkeflytting skyldes bortfall av device.			annen node	Nagios-alarm bortfall av fc-pather Nagios-alarm om at pakke er flyttet		blir ikke flyttet til annen node. Manuelt inngrep er nødvendig ved bortfall av all fc. Denne feilen er ingen "single failure" men to feil på en gang, og er ikke regnet som kritisk. Setter derfor testen til OK.	men vi testet ikke med Nagios-modul.	nødvendig ved bortfall av all fc.	
DE05	Bortfall all fc i clusteret	frakoble begge fc-kabler på alle nodene	-	fc-failure	Alle pakker i cluster-tas ned.	Logg bortfall av fc-pather Nagios-alarm bortfall av fc-pather Nagios-alarm om at pakker er nede.	Ikke aktuell å utføre siden DE04 ikke gav failover av pakker.	-	-		-OK 11/1
DE06	-	reboot av begge fc-switchene	-	Muligens Ikke påvirket fordi fc-link beholdes under reboot.	Muligens Ikke påvirket fordi fc-link beholdes under reboot.	Muligens ingenting?	-	-	-	-	-
DE07	-	ta strøm fra begge fc-switchene	-	fc-failure	Alle pakker i cluster-tas ned.	Logg bortfall av alle fc-pather Nagios-alarm bortfall av alle fc-pather Nagios-alarm om at pakker er nede.	-	-	-	-	-
DE08	-	ta ned diskarray	-	fc-failure	Alle pakker i cluster-tas ned.	Logg bortfall av alle fc-path. Nagios-alarm bortfall av alle fc-pather Nagios-alarm om at pakker er nede.	-	-	-	-	-
	Nett										
	HearBeat=HB										
DE09	Bortfall av en redundant nettopath i HB-nett /Fencing nett	Trekke ut nettopathskabel		failover på bonding hvis aktiv path.	Ingen	Logg om nettopath som nede. Nagios alarm om nettopath som er nede	Som forventet	Som forventet			OK 10/1
DE10	Bortfall av en redundant nettopath i public-nett	Trekke ut nettopathskabel		failover på bonding hvis aktiv path.	Ingen	Logg om nettopath som er nede. Nagios alarm om nettopath som er nede.	Som forventet	Som forventet			OK 10/1
	Bortfall av HB-	Trekke ut alle		Melding	<Vi vet	Logg om	Som	Som forventet			OK 10/1

DE11	nett fra node (iDRAC6 fortsatt tilgjengelig)	kablene fra noden for HB-nett/Fencing nett		om at nett er nede.	ikke> ??? Dersom noden kjører pakke så må noden fences før pakkene kan overføres til en annen node.	nettpather som nede. Nagios alarm om nettpath som er ende. Logg fra RHCS om at HB-nett er nede. Nagios Alarm om at HB-nett er nede.	forventet				
DE18	Bortfall av HB-nett/Fencingnett og nettaksess til iDRAC6	Trekke ut alle kablene fra noden for HB-nett/Fencing nett Trekke ut nettverkskabel til iDRAC6		Melding om at nett er nede.	<Vi vet ikke> ??? Dersom noden kjører pakke så må noden fences før pakkene kan overføres til en annen node.	Logg om nettpather som nede. Nagios alarm om nettpath som er ende. Logg fra RHCS om at HB-nett er nede. Nagios Alarm om at HB-nett er nede.	Som forventet.	Avvik: Nye masternode forsøker forgjeves å fence noden over IPMI "til evig tid". Måtte legge inn ekstra fencingmetode "fence_manual" og vi la inn 3 forsøk på IPMI fencing før manual fencing ble kjørt. Se Manuelt inngrep når IDRAC6-kort ikke er tilgjengelig Med innføring av manuell fencing : Restulatet er som forventet:	Som forventet.	Manuelt inngrep når IDRAC6-kort ikke er tilgjengelig	OK 11/1
DE12	Bortfall av public nett	Trekke ut begge kablene til public-nettet.		Melding i syslog	Pakker skal bli tatt ned og flyttet til annen noden fordi ip-nr i pakken ikke lenger vil fungere.	Logg om nettpath som nede. Nagios alarm om at noden ikke svarer på ping. Logg om at pakker flyttes.	Som forventet.	Som forventet			OK 11/1
DE13	Bortfall av alle nett på noden	Trekke ut alle nettverkskablene.	-	Melding i syslog.	<Vi vet ikke> ??? Dersom noden kjører pakke så må noden fences før pakkene kan overføres til en annen node.	Logg om nettpather som nede. Andre noder logger at noden er utilgjengelig og at den tilslutt blir fenced. Nagios alarm om at maskinen er nede (etter at den er fenced) Nagios-alarm om at noden er blitt fenced	Redunant test, se DE11 og DE12	-	-	-	-
	Node										
	Node failure	Frakoble alle strømkabler		N/A	Alle "pakker" som kjører på noden	a) Andre noder i clusteret logger at noden er nede.	N/A	Se DE18 .	a) Som forventet b) Ikke	Se DE18 .	OK 11/1

DE14					flyttes automatisk til neste node i cluster iht config Tjenesten er tilgjengelig igjen når "pakkene" er ferdig med oppstart.	b) Nagiosalarm om at noden ute av clusteret.			testet		
DE15		Kjøre shutdown på noden			Som ovenfor	Som ovenfor	Som forventet	a) Som forventet b) Ikke testet fordi vi ikke har ren nagios modul			OK 11/1
DE16		Generere panic på noden echo c > /proc/sysrq-trigger (echo h gir info om hvilke ting man kan legge inn her)			Som ovenfor	Som ovenfor	Som forventet	Som forventet			OK 11/1
	iLo / DRAC										
DE17	Oppdatering eller reset av iLO/DRAC	Hva skjer dersom vi oppdaterer firmware eller resetter på iLO/DRA? Dette kan også simuleres ved å reboot/resette iDRAC6 på noden			Forventer at intet skjer selv om vi bruker fencing via iLO/DRAC	N/A	Som forventet	N/A			OK 11/1

Oversikt over Avvik:

FU02 Overskrevne ressursscript

Ved oppgradering av RHCS, blir ressurscriptene overskrevet. Det betyr at våre lokale tilpassinger blir nullstilt, og dermed fungerer ikke f.eks oppstart av Oracle.

Mens testene ble gjort laget, brukte vi tilpassede versjoner av scriptene ved at vi laget linker som dette:

```
halv2-test01 cluster# ls -l /usr/share/cluster/oracledb.sh
lrwxrwxrwx 1 root root 16 Jan 14 11:14 /usr/share/cluster/oracledb.sh -> oracledb.sh.USIT
```

Ved oppgradering blir denne linken overskrevet. Vi må ta opp med RedHat hvordan dette kan løses. Evt må vi bruke den såkalte "script-ressursen" og ikke oracle-ressursen til alle formål siden man da har mulighet til å legge inn egne script i clusteroppsettet.

FU08 Ikke støttet "out of the box"

I RHCS er det ikke supportert "out of the box" at det kan stilles egne betingelser for at en pakke skal starte eller kjøre på en node, og hva som ska skje dersom betingelsen ikke er oppfylt.

Vi laget et script rhcsmisc1-server-req.sh som kun gav returkode 0 dersom filen /etc/usit/run-halv2-server eksisterte når det ble kalt opp med argumentet status eller start. Dersom denne filen ikke finnes, gir dette returkode 150. Argumentet stop gir alltid returkode 0.

I servicen rhcsmisc ble det laget en script-ressurs med scriptet rhcsmisc1-server-req.sh som er helt øverst av alle ressursene til servicen i cluster.conf, dvs at alle andre ressurser er avhengig av at dette scriptet kjører og returnerer feilkode 0. Da kan man under oppstart av servicen (da kalles scriptet opp med start) og under kjøring av servicen (da kalles scriptet opp med status) kontrollere hele servicen hva filen /etc/usit/run-halv2-server

Vi kan styre om en service skal starte eller kjøre ved å legge inn en scriptagent tidlig i servicen. Da kan man styre hvilke status som returneres ut fra gitte kriterier, og dermed styre hvordan hele servicen skal fungere.

Vi brukte logger til å skrive meldinger til syslog slik at vi kunne se betingelsen fungerte:

```
Jan 13 13:30:55 halv2-test03 logger: status: Requirement not met: File /etc/usit/run-halv2-server does not exist.
Jan 13 13:30:55 halv2-test03 clurgmgrd: [9055]: <err> script:halv2-server-req: status of /etc/cluster-script/rhcsmisc1-server-req.sh
failed (returned 150)
```

Kode i `cluster.conf`:

```
<script file="/etc/cluster-script/rhcsmisc1-server-req.sh" name="halv2-server-req">
RESTEN AV RESSURSENE
</script>
```

Vi kan **ikke** styre om service skal flytte til annen node eller restarter ut fra en gitt betingelse slik som formulert i kravet siden failover-egenskapene er satt en gang for alle på servicen og ikke kan endres mens man kjører.

FU10 Restart av ressurser og tjenester uten varsel

Alle endringer på service eller ressurs-nivå gir restart av det nivået man har endret på og de som er avhengige at denne ressursen.

Man får ingen forvarsel om disse restartene når man gjør endringer.

Dette ble verifisert med at vi gjorde endringer med tre offisielle metoder:

- editere rett i `cluster.conf` og kjøre `ccs_tool update`
- gjøre endringer i *Conga*
- gjøre endringer i `system-cluster.conf`. Dette verktøyet kan ikke brukes fordi det ikke engang kjenner igjen oracle-ressursen.

Endringer i failoverdomain eller fence-metode gir ingen restart.

FU12 Begrenset støtte for avhengigheter mellom services

RHCS støtter trivelles dependencies "out of the box" (via `depend` og `depend_mode` for servicen). Se DOC-26981 (hos [RedHat](#) eller [lokal kopi](#))Det ser ut som om mer avanserte avhengigheter kan implementeres med EventScript.

Side 28 i

<https://www.usit.uio.no/prosjekter/halv2/arbeidsdokumenter/rhcs/rhcs-dokumentasjon/Migrating%20SGLX%20cluster%20to%20RHCS%20Cluster-RH%20Version-final.pdf>

Se også

<http://sources.redhat.com/cluster/wiki/EventScripting>

Støtte for dependencies er et "Ønske" og ikke et absolutt krav, og kun aktuelt for SAP som hvor det finnes et eget resource-kit hvor dependencies er implementert.

Gjorde en enkel test: `rhcsoracle` skal være avhengig av `rhcsmisc1`.

Fra `cluster.conf`: `depend="service:rhcsmisc1" depend_mode="hard"`

Legger inn avhengighet emens `rhcsmisc` er nede: Vi får ikke melding i syslog om at oracle ikke kan starte, men den starter automatisk når `rhcsmisc` er oppe. Når begge er oppe og vi tar ned `rhcsmisc`, blir også `rhscoracle` tatt ned og vi får følgende melding:

```
Jan 12 14:25:17 halv2-test01 clurgmgrd[6998]: <warning> Stopping service service
:rhcsoracle: Dependency missing
Jan 12 14:25:17 halv2-test01 clurgmgrd[6998]: <notice> Stopping service service:
rhcsoracle
```

FU22 Komplikasjoner

Det oppstod noen da vi fjernet en service for å reteste at vi kan legge til nye service i fart.

Ved første test forsvant ikke servicen fra halv2-test03. Vi restartet `rgmanager` på denne noden, og da forvant servicen slik den skulle. Dersom man restarter `rgmanager` på en node med services blir disse restartet.

Ved gjentatte tester fungerte fjerning og å legge til services uten at andre services ble påvirket. Vi fjernet service via *Conga* og ved å editere `cluster.conf`.

Vi vet ikke hvorfor det ikke fungerte å fjerne pakke i første test.

Å legge til pakker fungerte hver gang.

DE04 Manuelt inngrep ved bortfall av all fc mot disk

Ved bortfall av all fibrechannel connectivity for en node, blir servicene på noden tatt ned og status blir satt til "failed". Dermed blir ikke servicene flyttet til neste node og vi får nedetid for kunden. Man må selv manuelt fjerne `lvm-tags` og starte opp tjenestene.

Årsaken er at noden uten `fc` ikke klarer å fjerne `lvm-tag`'ene på volumgruppen. Løsningen er å finne i RHEL5.6 eller RHEL6 hvor man kan kjøre HA-LVM sammen med `clvmd`. Dette har vi ikke testet, og generelt ønsker vi ikke å benytte `clvmd`.

Se også [Meldinger syslog](#)

Denne feilen er ingen "single failure" men to feil på en gang, og er ikke regnet som kritisk. Setter derfor testen til OK.

DE14 og DE18: Manuelt inngrep når IDRAC6-kort ikke er tilgjengelig

I de to testene hvor iDRAC6-kortet er utilgjengelig, f.eks fordi

- nettkabelen til iDRAC6-kortet er koblet fra
- strømmen på noden er koblet fra

vil ikke de andre nodene klare å *fence* noden dersom man kun har fencing over IPMI som eneste fencing-metode. Siden RHCS legger til grunn en konservativ tilnærming, insisteres det på at noden må fences før services kan tas opp på andre noder. Resultatet er dermed nedetid for brukerne. Det er ikke trivielt å løse dette i RHCS uten å legge til en ny fencingmetode som heter *manual_fence*. Dette gir systemadministratoren mulighet til å fortelle clusteret at det er trygt å betrakte noden som fenced. Siden vi ønsker at RHCS skal forsøke å fence noden flere ganger vha IPMI før den venter på at vi skal gjøre en manuell fencing, legger vi inn tre forsøk på IPMI-fencing for hver node FØR den faller tilbake til manuell fencing. Dette ser slik ut i `cluster.conf` (eksempel fra node1 i clusteret)

```
<clusternode name="halv2-test01-hb" nodeid="1" votes="1">
  <fence>
    <method name="1">
      <device name="halv2-test01-ipmi"/>
    </method>
    <method name="2">
      <device name="halv2-test01-ipmi"/>
    </method>
    <method name="3">
      <device name="halv2-test01-ipmi"/>
    </method>
    <method name="4">
      <device name="manualfence" nodename="halv2-test01-hb"/>
    </method>
  </fence>
</clusternode>
```

Det kommer da en melding i syslog med instruksjoner om hvordan man kan kjøre en kommando for å fortelle clusteret at noden kan betraktes som fenced, og dette gjøres fra den samme noden som først forsøkte å utføre fencingen via IPMI:

```
# fence_ack_manual -n <node>
```

Når noden er merket som fenced, blir servicene relokert til en annen node, og kundene vil igjen oppleve tjenesten som tilgjengelig.

Skal man ha automatisk failover av services i denne situasjonen kan må man legge inn en ny sekundær fencing-metode, feks

- strømmeskinne fencing
- scsi persistent reservation
- fc fencing

Som tertiær fencing kan man da legge inn manuell fencing.

Denne feilen er ingen "single failure" men to feil på en gang, og er ikke regnet som kritisk. Setter derfor testen til OK.

Appendix 24 Testresultater RHCS på RHEL6

Tredelt testplan

Hensikten med testene er å verifisere at

- HA-løsningen innfrir de krav som vi stiller
- HA-løsningen støtter de tjenestene vi skal kjøre på en god måte

I testplanen er det derfor *ikke* et primært mål å verifisere at HW fungerer eller at all kabling er korrekt utført i oppsettet. Tester for å verifisere et konkret oppsett må gjøres på det konkrete oppsettet før systemet settes i drift.

I dagens SGLX kjører vi med quorum-server. Hvordan fencing og quorum løses i andre HA-løsninger er nytt for oss og derfor har vi valgt å legge spesielt vekt på testing av dette.

Vi har derfor delt testplanen i tre deler:

- funksjonelle tester med utgangspunkt i kravspesifikasjonen unntatt robusthetstester
- test av robusthet mhp fencing og quorum.
- destruktive tester med utgangspunkt i systemskisse

Vi skal lage tre pakker i clusteret:

- pakke med Oracle-database og GridControl agent
- pakke med Postgres-database
- pakke med egenutviklet tjener ("Halv2-Server") som svarer på en port og som skriver til en fil.

Testene er tenkt utført hovedsaklig med Oracle-pakken. Det må lages et testscript som kjøres fra en sqlplus-sesjon på en maskin som står utenfor clusteret. Testscriptet skal gjøre lesing og skriving i databasen.

Funksjonelle tester

I kolonnen "Krav" er det link til konkrete krav i [kravspesifikasjonen](#) som danner grunnlag for den konkrete testen.

Test#	Krav	Hvordan teste	Forventet resultat	Forventet Logging/Alarm	Faktisk resultat	Faktisk Logging/Alarm	Avvik	Status/Dato
FU1	ADM01: Det må være støtte for "rolling upgrade" av operativsystemet på nodene i HA-løsningen.	<u>Planlagt test:</u> Flytt "pakke" bort fra node. Oppgrader kjerne (eller nedgrader dersom ny kjerne ikke tilgjengelig). Reboot server. Flytt "pakke" tilbake til node. Ref RedHat Knowledgebase artikkel 43210.	Node skal kunne joine cluster etter reboot. Pakke skal kunne overføres til node. Kjørende pakker på andre noder skal ikke påvirkes.	Logg på de andre nodene i clusteret om at nodene har gått inn og ut av clusteret	Som forventet	Som forventet		OK 26/1
FU2	ADM06: Programvaren i HA-løsningen kan oppgraderes "rolling", dvs node for node uten at alle tjenester må tas ned	<u>Planlagt test:</u> Flytt "pakke" bort fra node. Oppgraderer RHCS Flytt "pakke" tilbake til node.	Node skal kunne joine cluster etter oppgradering av RHCS på node. Pakke skal kunne starte på node. Kjørende pakker på andre noder skal ikke påvirkes.	Logg på de andre nodene i clusteret om at noden har gått inn og ut av clusteret	Ikke testet fordi det ikke er kommet nye versjoner på RHEL6	N/A	Da vi testet på RHEL5 fikk vi et avvik: Oppgraderte ressurscript	Utgår fordi det ikke er kommet nye versjoner på RHEL6
	FAIL01: Ved flytting av en tjeneste fra en node til en annen i HA-løsningen er nedetiden for tjenesten omtrent det	Flytte "pakke" fra en node til en annen som følger:	Montering av filsystemer og oppstart av tjeneste skal ta "normal" tid. <u>For Oracle</u>	N/A	Som forventet Det er mulig å koble til oracle straks databaseinstanse har registrert seg	N/A		OK 27/1

FU3	samme som tiden det tar å stoppe og starte ressursene i tjenesten på tilsvarende maskinvare utenfor HA-løsningen.		tjeneste: Det skal være mulighet for å koble til databasen med sqlplus straks listener og instanse kjører.		hos lytteprossen etter ca 1 minutt. For å fikse dette må oracledb.sh endres slik at lytteprosess endres først.			
FU4	FAIL04 Feil på maskinvare i HA-løsningen skal ikke gi langvarig nedetid for brukere av tjenesten	<Gjøres under destruktive tester>		N/A	Som forventet	N/A		OK 27/1
FU5	FAIL06 Feil på maskinvare i HA-løsingen skal ikke kreve manuelle inngrep for å gjøre tjenesten tilgjengelig for brukeren.	<Gjøres under destruktive tester>		N/A	Som forventet	N/A		OK 27/1
FU6	FAIL07 Brukere av tjenesten skal ikke måtte endre i konfigurasjon på klienten etter at en tjeneste er flyttet til en annen node i HA-løsningen.	<kan besvares når FAIL01 er testet fordi det innebærer flytting av "pakke")	Klient skal uten modifikasjon kunne koble til tjeneste før og etter "pakke"-flytting.	N/A	Som forventet	N/A		OK 27/1
FU7	CONF01 Det må være mulig å konfigurere hvilke ressurser som skal være knyttet til en tjeneste Eksempler: filsystemer, ip-adresser og prosesser.	<kan besvares når vi har laget pakke for Oracle, Postgres og "HA2-service".	CONF01	N/A	Som forventet	N/A		OK 26/1
FU8	CONF02 Det må være mulig å kunne spesifisere egne betingelser som må være oppfylt for at tjenesten skal kunne gå på en node, og hva som skal skje når disse betingelsene ikke er oppfylt. Eksempel på egen betingelse: Når filsystemene til pakken er utilgjengelige skal tjeneste tas ned og ikke flyttes til en annen node	Planlagt test: <lage "pakke" med ekstra betingelse i forhold til toolkit'et. F.eks ikke kjøre tjeneste om load > 10 Gjennomført test: Vi laget et script som kun gav returkode 0 dersom en fil eksisterte når det ble kalt opp med argumentet status eller start. I servicen rhcsmisc ble det laget en script-ressurs med scriptet rhcsmisc1-server-req.sh som er helt øverst av alle ressursene til servicen i cluster.conf, dvs at alle andre ressurser er avhengig av at dette scriptet. Se Ikke støttet "out of the box" for detaljer	CONF02	Det må være mulig å implementere logging om at pakken ikke blir startet med beskrivelse av <i>hvorfor</i>	Kravet er ikke tilfredsstillt slik det står formulert , men etter diskusjon i gruppen har vi funnet at kravet er <i>feilformulert</i> . Det kravet egentlig formulerer er at ikke alle ressurser som feiler skal måtte medføre at servicen flyttes eller restartes, og dette lar seg enkelt implementere i RHCS ved å endre i ressursscriptene.	Som forventet. Kan implementeres i script-ressurs.	Ikke støttet "out of the box"	OK 26/1
FU9	CONF04 Det må være mulig å gjøre tilpasninger i hvilke script/kommandoer som skal kjøres når en tjeneste startes/stoppes på en node.	Legge inn oppstart av GridControl-agent i Oracle-"pakke". Testet ved å legge inn requirement-scriptet i FU8.	Agenten startes når "pakken" startes og avsluttes når "pakken" stoppes.	Vi må kunne legge inn skrivning til logg om at GridControl agent har blitt startet eller at den er blitt stoppet.	Som forventet	Implementerte ikke skrivning til logg om oppstart av agent, men dette er enkelt å få til.		OK 26/1
	CONF05	I servicen rhcsmisc:Endre force_umount fra 1 til 0 i fs-		N/A	Se avvik i egen kolonne.	N/A	Restart av ressurser og	FAILED. 26/1

FU10	Det bør være mulig å endre parametre som styrer innstillinger rundt oppstart og avslutning av tjenester i HA-løsninger uten at tjenesten må tas ned.	ressursen.					tjenester uten forvarsel	(en CONF05 er et ØNSKE, ikke KRAV)
FU11	CONF07 Det må være mulig å spesifisere hvilke noder en tjeneste skal kunne gå på, og rekkefølgen for hvordan tjenesten skal gjøre failover til andre noder.	Lage pakke med Oracle tjeneste med angitt failover-rekkefølge. Ta ned noden og verifisere at pakken feiler over til korrekt node.	CONF07	N/A	Som forventet	N/A		OK 26/1
FU12	CONF08 Det kan spesifiseres avhengigheter mellom tjenester og mellom tjenester og noder	Legg inn at Postgres-pakken ikke skal kjøre på samme node som Oracle-pakken, men på samme node som "HALv2-server"-pakken. Kjørt test: Gjorde oracle-service avhengig av halv2-tjenseren ved å sette depend="service:rhcsmisc1" depend_mode="hard"	N/A	Loggmelding om at en avhengighet var årsaken til at pakken ble startet på en gitt node.	N/A	Som forventet. Jan 27 11:33:56 rgmanager Stopping service:rhcsoracle: Dependency missing	Begrenset støtte, avanserte avhengigheter må lages selv. Men SAP er støttet!	OK 27/1
FU13	CONF09 HA-løsningen kan konfigureres slik at den restarter ressurser i en tjeneste som den har detektert at ikke lenge kjører	Kill alle oracle-prosessene til Oracle-tjeneste.	Oracle-instanse blir startet opp igjen og klient kan koble til. Oracle blir ikke forsøkt startet opp flere ganger dersom oppstart feilet.k	Det må logges at ressursen er nede og at den ble restartet. Status for restarten må også logges.	Som forventet	Som forventet: Det blir logget at instansen blir restartet, men ikke at den var nede: Jan 27 12:32:17 halv2-test01 oracledb.sh: Restarting Oracle Database...		OK 27/1
FU14	FAKE01 (ikke i kravspesikasjon) HA-løsningen kan automatisk faile tilbake "pakker" når noden igjen er tilgjengelig etter en failover.	Konfigure failoverdomain for failback. Ta shutdown av node som kjører pakken. Verfisiere at "pakke" flytter til annen node. Ta opp node igjen. Verfisiere at "pakke" flytter tilbake til original node.	<som beskrevet i FAKE01>	Det må logges at pakken ble flyttet, og <i>hvorfor</i> den ble flyttet. Det må fremgå at det var en automatisk failback som var årsaken til flyttingen.	Som forventet	Som forventet. Syslog: Jan 26 13:40:47 halv2-test02 rgmanager[4392]: Relocating service:rhcsmisc1 to better node halv2-test03-hb		OK 26/1
FU15	MON01 Det må enkelt å få ut info om status for HA-løsningen, dets noder og dets tjenester	Kjøre CLI-kommando for oversikt over noder og "pakker". Åpne GUI og vise status over clusteret.	Oversiktene fra kommandoene gir "tilstrekkelig" informasjon om clusteret.	N/A	Som forventet	N/A		OK 26/1
FU16	MON02 Logging må være tilstrekkelig slik at det er mulig å feilsøke ved eventuelle problemer.	<Besvare etter destruktive tester med eksempler>	N/A	MON02	N/A	Noen mangler i logging.	Noen mangler i logging.	Delvis OK 26/1
FU17	MON03 Det er ønskelig at HA-løsningen kan sende sine logger fra alle noder til et sentralisert sted	<Besvare etter destruktive tester med eksempler>	N/A	MON03	N/A	Ja, det logges jo bla jo syslog.		OK 26/1
	MON04	Finne eksisterende modul eller lage ny modul som	N/A	Kravspesifikasjon for Nagios.	N/A	rpm-pakken cluster-snmp inneholder bla		OK 27/1

FU18	HA-løsningen må la seg overvåke av Nagios	overvåker clusteret.		Ferdige nagios-moduler er ikke gode nok. Må lage egen		redhat-cluster MiB Det er mulig å overvåke fordi vi får fornuftig info fra snmp og evt c lustat. Eksisternede Nagios-moduler gir informasjon om RHCS, men den er ikke god nok så vi må nok lage egen modul.		
FU19	MON05 Oracle-databaser som kjører i HA-løsningen må la seg overvåke av GridControl.	<besvares når CONF04 er utført>	MON05	N/A	Som forventet. Vi kan kjøre egne programmer, men selve GC-agenten (OMA) er ikke testet på RHEL6	N/A		OK 27/1
FU20	MON06 Det må være mulig å få generert en alarm dersom en ressurs i tjenesten ikke kjører	<besvares når MON04> er utført>	MON06	Alarm i Nagios	Som forventet. Samme som for test av RHEL5: Generering av alarm lages i scriptet til ressursen, ikke støtte for dette direkte. Oracle-ressursscriptet har kode for restart av oracle i "status"-delen av scriptet, og dette viser at man kan lage alarmer eller logging på denne måten.	Ikke testet.		OK 27/1
FU21	SCA01 Det må være mulig å melde noder inn i og ut av HA-løsningen uten at det påvirker tjenestene i løsningen (med unntak av tjenester som går på noder som skal fjernes).	Både for 2 og 3 node. Også teste om antall noder påvirker fencing, må hele cluster ned for å legge til flere noder i fencing-mekanisme?	SCA01	Logg fra de andre nodene om at en node går inn og ut av clusteret.	Det er ikke mulig å øke antall noder fra 2 til 3 noder mens clusteret er oppe siden 2-node cluster er en egen type cluster med andre quorum-regler. Vi trenger derfor 4 noder for å få testet dette. Testen ~~UTGÅR~~ .	-	-	Utgår fordi vi trenger 4 noder for å testet dette
FU22	SCA02 Det må være mulig å definere nye tjenester i løsningen uten at det krever nedetid på andre tjenester i løsningen.	Lage en ny pakke.	SCA02	N/A	Som forventet	N/A		OK 27/1

Test av robusthet; fencing og quorum

Hensikten med disse testene er å avgjøre hvilke typer quorum og fencing som fungerer best og som er enkle å drifte.

Kanhende vil noen metoder for fencing bli utelukket før testing dersom vi leser dokumentasjon og konkluderer med atat metodene er uhensiktsmessige.

Testene skal gjøre for både 2 og 3 noder i clusteret.

Gunnar lager utkast til mail for å sende til RH for å spørre om

- gode måter å gjøre quorum og fencing
- hva de fleste kunder gjør

Vi må også finne svar på om 2-tagging er i bruk av toolkitene.

To node cluster:

Overgang fra tre noder til to noder.

- 1) Ta backup av cluster.conf (cluster.conf-3-node)
- 2) Stop node3 som skal ut clustersoftware: servive rgmanager stop && service cman stop
- 3) Fjerne node3 fra cluster.conf (husk failoverdomains)
- 4) Legg inn parameteren for two_node: <cman two_node="1" expected_votes="1"/>
- 5) Distribuer ny config
- 6) Restart clustersoftware på de andre nodene
- 7) Kjør cman_tool nodes for å verifiser nodene i clusteret :)

FIXME: Doc scriptene et sted

Test#	Krav	Hvordan teste	Antall ganger gj.ført	Forventet resultat OS	Forventet resultat HA-løsning	Forventet Logging/Alarm	Faktisk resultat OS	Faktisk resultat HA-løsning	Faktisk Logging/Alarm	Avvik	Status/Dato
RO01	ROB03 : Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).	a) Dra ut alle nettverkskabler til HB-nettet. b) Kill prosessen til clusterkommunikasjonen aisexec c) Gi noden en panic echo c > /proc/sysrq-trigger	1	N/A	Noden skal bli fenced og pakken skal ikke kunne starte på en annen node før noden er nede.	<ikke spesifisert>	N/A	a) Som forventet ~~b) Testen utgår da~~ aisexec kjører ikke på RHEL6. ~~Spekulasjon:~~ Denne kommunikasjonen gjøres av en kjernemodul og ikke et program. Testen utgår. c) Som forventet	a) Syslog: Jan 27 13:18:52 halv2-test01 fenced[23041]: fencing node halv2-test02-hb		OK 27/1
2	ROB04 : Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip- adresser eller filsystemer som er aktive på to noder samtidig).	a) Monter et av filsystemene som er montert opp i en pakke på nodeX manuelt på nodeY. b) Ta opp en pakke på nodeY som allerede kjører på nodeX.	-	a) Filsystemet lar seg ikke montere pga lvm2-tags	b) HA-løsningen vil ikke la pakken starte på nodeY og man får feilmelding om at den kjører på nodeX	a) Ingen b) Ingen	-	-	-Testet for 3-node-cluster		Utgår fordi vi har testet dette for 3-node-cluster.
RO03	ROB05 : Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing)	<besvares når ROB03 er utført fordi fencing>									OK 27/1

Tre node cluster

Test#	Krav	Hvordan teste	Antall ganger gj.ført	Forventet resultat OS	Forventet resultat HA-løsning	Forventet Logging/Alarm	Faktisk resultat OS	Faktisk resultat HA-løsning	Faktisk Logging/Alarm	Avvik	Status/Dato
RO04	ROB03: Løsningen må ha pålitelige mekanismer for å hindre at en tjeneste kan bli forsøkt startet på flere noder samtidig (Quorum / split-brain).	a) Dra ut alle nettverkskabler til HB-nettet. b) Kill prosessene til clusterkommunikasjonen aisexec		-	Noden skal bli fenced og pakken skal ikke kunne starte på en annen node før noden er nede.	-	a) HB-kabler ble tatt ut i test DE11 og vi testet ikke igjen. b) Testen utgår da aisexec kjører ikke på RHEL6. Spekulasjon: Denne kommunikasjonen gjøres av en kjernemodul og ikke et program. Testen utgår.				Utgår 26/1 fordi dette allerede er testet i DE11 og fordi aisexec ikke finnes på RHEL6
RO05	ROB04: Løsningen må ha pålitelige mekanismer for å hindre at ressurser i tjenesten kan være aktive på flere noder samtidig (e.g. ip-adresser eller filsystemer som er aktive på to noder samtidig).	a) Monter et av filsystemene som er montert opp i en pakke på nodeX manuelt på nodeY. b) Ta opp en pakke på nodeY som allerede kjører på nodeX.	1	a) Filsystemet lar seg ikke montere pga lvm2-tags	b) HA-løsningen vil ikke la pakken starte på nodeY og man får feilmelding om at den kjører på nodeX	a) Ingen b) Ingen	a) Som forventet	b) Som forventet	Som forventet		OK 26/1
RO06	ROB05: Løsningen må ha pålitelige mekanismer for å ta ned noder som ikke lenger får være med i clusteret (fencing)	<besvares når ROB03 er utført fordi fencing>									OK 26/1

Destruktive tester

Tester som innebærer reboot eller å ta strømmen av fc-switcher kan ikke utføres fordi vi ikke kjører på dedikerte fc-switcher uten annen trafikk.

Test#	Hendelse	Hvordan teste	Antall ganger gj. ført	Forventet resultat OS	Forventet resultat HA-løsning	Forventet Logging/Alarm	Faktisk resultat OS	Faktisk resultat HA-løsning	Faktisk Logging/Alarm	Avvik	Status/Dato
	FC										
DE01	Bortfall av en redundant fc-path på node	frakoble fc-kabel på node Test utført 25/01 11:31:15		fc-failover	Ikke påvirket	Logg bortfall av fc-path. Nagios-alarm bortfall av fc-path.	Som forventet.	Som forventet.	Som forventet. (men USIT har ingen Nagios plugin for nye multipath-format på meldinger)		OK 25/1

DE02		reboot av fc-switch		Muligens Ikke påvirket fordi fc-link beholdes under reboot.	Ikke påvirket	Muligens ingenting?					Utgår fordi vi ikke har fc-dedikerte switcher eller diskarray.
DE03	-	ta strøm fra fc-switch	-	fc-failover	Ikke påvirket	Logg bortfall av fc-path. Nagios-alarm bortfall av fc-path.					Utgår fordi vi ikke har fc-dedikerte switcher eller diskarray.
DE04	Bortfall all fc for en node	frakoble begge fc-kabler på node Verifiser at pakkeflytting skyldes bortfall av device. Test utført: Jan 25 13:34:32		fc-failure	Pakke flyttes til annen node	Logg bortfall av fc-pather Nagios-alarm bortfall av fc-pather Nagios-alarm om at pakke er flyttet	Som forventet.	Servicene blir satt i "failed" og blir ikke flyttet til annen node. Manuelt inngrep er nødvendig ved bortfall av all fc. Denne feilen er ingen "single failure" men to feil på en gang, og er ikke regnet som kritisk. Setter derfor testen til OK.	Som forventet, men vi har ingen nagios-modul på RedHat6.	Manuelt inngrep nødvendig ved bortfall av all fc.	OK 25/1
DE05	Bortfall all fc i clusteret	frakoble begge fc-kabler på alle nodene	-	fc-failure	Alle pakker i cluster tas ned.	Logg bortfall av fc-pather Nagios-alarm bortfall av fc-pather Nagios-alarm om at pakker er nede.	Ikke aktuell å utføre siden DE04 ikke gav failover av pakker.	-	-		Utgår, testen er ikke aktuell å utføre siden DE04 ikke gav failover av pakker.
DE06	-	reboot av begge fc-switchene	-	Muligens Ikke påvirket fordi fc-link beholdes under reboot.	Muligens Ikke påvirket fordi fc-link beholdes under reboot.	Muligens ingenting?	-	-	-	-	Utgår fordi vi ikke har fc-dedikerte switcher eller diskarray.
DE07	-	ta strøm fra begge fc-switchene	-	fc-failure	Alle pakker i cluster tas ned.	Logg bortfall av alle fc-pather Nagios-alarm bortfall av alle fc-pather Nagios-alarm om at pakker er nede.	-	-	-	-	Utgår fordi vi ikke har fc-dedikerte switcher eller diskarray.
DE08	-	ta ned diskarray	-	fc-failure	Alle pakker i cluster tas ned.	Logg bortfall av alle fc-path. Nagios-alarm bortfall av alle fc-pather Nagios-alarm om at pakker er nede.	-	-	-	-	Utgår fordi vi ikke har fc-dedikerte switcher eller diskarray.

	Nett										
	HearBeat=HB										
DE09	Bortfall av en nettverkspath i HB-nett /Fencing nett	Trekke ut nettverkskabel Test utført: Jan 25 13:43:13		failover på bonding hvis aktiv path.	Ingen	Logg om nettpath som nede. Nagios alarm om nettpath som er nede	Som forventet	Som forventet	Som forventet, men vi har ingen nagios-modul på RedHat6.		OK 25/1
DE10	Bortfall av en redundant nettverkspath i public-nett	Trekke ut nettverkskabel Test utført: Jan 25 13:46:18		failover på bonding hvis aktiv path.	Ingen	Logg om nettpath som er nede. Nagios alarm om nettpath som er nede.	Som forventet	Som forventet	Som forventet, men vi har ingen nagios-modul på RedHat6.		OK 25/1
DE11	Bortfall av HB-nett fra node (iDRAC6 fortsatt tilgjengelig)	Trekke ut alle kablene fra noden for HB-nett/Fencing nett Test gjort: Jan 26 09:49:41 halv2-test01 kernel: bnx2: eth1 NIC Copper Link is Down Jan 26 09:49:46 halv2-test01 kernel: bnx2: eth5 NIC Copper Link is Down		Melding om at nett er nede.	Dersom noden kjører pakke så må noden fences før pakkene kan overføres til en annen node.	Logg om nettpather som nede. Nagios alarm om nettpath som er ende. Logg fra RHCS om at HB-nett er nede. Nagios Alarm om at HB-nett er nede.	Som forventet.	Jan 26 09:50:12 halv2-test02 fenced[3210]: fencing node halv2-test01-hb Jan 26 09:50:26 halv2-test02 fenced[3210]: fence halv2-test01-hb success	Som forventet, men vi har ingen nagios-modul på RedHat6.		OK 26/1
DE18	Bortfall av HB-nett/Fencingnett og nettaksess til iDRAC6	Trekke ut alle kablene fra noden for HB-nett/Fencing nett Trekke ut nettverkskabel til iDRAC6 Test gjort: Jan 26 10:07:39 halv2-test02 kernel: bnx2: eth1 NIC Copper Link is Down Jan 26 10:07:48 halv2-test02 kernel: bnx2: eth5 NIC Copper Link is Down		Melding om at nett er nede.	<Vi vet ikke> ??? Dersom noden kjører pakke så må noden fences før pakkene kan overføres til en annen node.	Logg om nettpather som nede. Nagios alarm om nettpath som er ende. Logg fra RHCS om at HB-nett er nede. Nagios Alarm om at HB-nett er nede.	Som forventet.	manuelt inngrep nødvendig	Som forventet, men vi har ingen nagios-modul på RedHat6.	manuelt inngrep nødvendig	26/1 OK
DE12	Bortfall av public nett	Trekke ut begge kablene til public-nettet. Test gjort: Jan 26 11:11:48 halv2-test02 kernel: bnx2: eth0 NIC Copper Link is Down Jan 26 11:11:48 halv2-test02 kernel: bnx2: eth4 NIC Copper Link is Down		Melding i syslog	Pakker skal bli tatt ned og flyttet til annen noden fordi ip-nr i pakken ikke lenger vil fungere.	Logg om nettpath som nede. Nagios alarm om at noden ikke svarer på ping. Logg om at pakker flyttes.	Som forventet	Som forventet	Som forventet, men vi har ingen nagios-modul på RedHat6.		OK 26/1
	Bortfall av alle nett på noden	Trekke ut alle nettverkskablene.	-	Melding i syslog.	<Vi vet ikke> ??? Dersom noden kjører	Logg om nettpather som nede. Andre noder	-Redunant test, se DE11 og DE12	-	-	-	Utgår fordi testen er redundant, se DE11 og DE12

DE13					pakke så må noden fences før pakkene kan overføres til en annen node.	logger at noden er utilgjengelig og at den tilslutt blir fenced. Nagios-alarm om at maskinen er nede (etter at den er fenced) Nagios-alarm om at noden er blitt fenced					
	Node										
DE14	Node failure	Frakoble alle strømkabler På node 2 etter at node 1 mistet strøm: Jan 26 11:19:41 halv2-test02 corosync[3171]: [TOTEM] A processor failed, forming new configuration.			Alle "pakker" som kjører på noden flyttes automatisk til neste node i cluster iht config Tjenesten er tilgjengelig igjen når "pakkene" er ferdig med oppstart.	a) Andre noder i clusteret logger at noden er nede. b) Nagiosalarm om at noden ute av clusteret.		Se DE18	Se DE18	Se DE18	OK 26/1
DE15		Kjøre shutdown på noden			Som ovenfor	Som ovenfor	Som forventet	Som forventet	Som forventet		OK 26/1
DE16		Generere panic på noden echo c > /proc/sysrq-trigger (echo h gir info om hvilke ting man kan legge inn her)			Som ovenfor	Som ovenfor	Som forventet	Som forventet	Som forventet		OK 26/1
	iLo / DRAC										
DE17	Oppdatering eller reset av iLO/DRAC	Hva skjer dersom vi oppdaterer firmware eller resetter på iLO/DRA? Dette kan også simuleres ved å reboot/resette iDRAC6 på noden [root@halv2-test01 ~]# racadm racreset			Forventer at intet skjer selv om vi bruker fencing via iLO/DRAC	N/A	Som forventet	Som forventet	Som forventet		OK 26/1

Oversikt over Avvik:

FU02 Overskrevne ressurscript

Ved oppgradering av RHCS, blir ressurscriptene overskrevet. Det betyr at evt lokale tilpassinger blir nullstilt. Under testene på RHEL6 kjørte vi med umodifiserte script siden disse fungerte, mens vi på RHEL5 kjørte med modifiserte script.

Mens testene ble gjort på RHEL5, brukte vi tilpassede versjoner av scriptene ved at vi laget linker som dette:

```
halv2-test01 cluster# ls -l /usr/share/cluster/oracledb.sh
lrwxrwxrwx 1 root root 16 Jan 14 11:14 /usr/share/cluster/oracledb.sh -> oracledb.sh.USIT
```

Ved oppgradering blir denne linken overskrevet. Dersom vi også i fremtiden ikke kan bruke de originale scriptene, må ta opp med RedHat hvordan dette kan løses. Evt må vi bruke den såkalte "script-ressursen" og ikke oracle-ressursen til alle formål siden man da har mulighet til å legge inn egne script i clusteroppsettet.

FU08 Ikke støttet "out of the box"

I RHCS er det ikke supportert "out of the box" at det kan stilles egne betingelser for at en pakke skal starte eller kjøre på en node, og hva som ska skje dersom betingelsen ikke er oppfylt.

Vi laget et script rhcsmisc1-server-req.sh som kun gav returkode 0 dersom filen /etc/usit/run-halv2-server eksisterte når det ble kalt opp med argumentet status eller start. Dersom denne filen ikke finnes, gir dette returkode 150. Argumentet stop gir alltid returkode 0.

I servicen rhcsmisc ble det laget en script-ressurs med scriptet rhcsmisc1-server-req.sh som er helt øverst av alle ressursene til servicen i cluster.conf, dvs at alle andre ressurser er avhengig av at dette scriptet kjører og returnerer feilkode 0. Da kan man under oppstart av servicen (da kalles scriptet opp med start) og under kjøring av servicen (da kalles scriptet opp med status) kontrollere hele servicen hva filen /etc/usit/run-halv2-server

Vi kan styre om en service skal starte eller kjøre ved å legge inn en scriptagent tidlig i servicen. Da kan man styre hvilke status som returneres ut fra gitte kriterier, og dermed styre hvordan hele servicen skal fungere.

Vi brukte logger til å skrive meldinger til syslog slik at vi kunne se betingelsen fungerte:

Betingelses-filen finnes:

```
Jan 26 13:01:46 halv2-test01 rgmanager[28689]: Executing /etc/cluster-script/rhcsmisc1-server-req.sh start
Jan 26 13:01:46 halv2-test01 logger: start: Requirement is met: File /etc/usit/run-halv2-server actually does exist.
```

Betingelses-filen finnes ikke:

```
Jan 26 13:00:04 halv2-test01 logger: status: Requirement not met: File /etc/usit/run-halv2-server does not exist.
Jan 26 13:00:04 halv2-test01 rgmanager[27435]: script:halv2-server-req: status of /etc/cluster-script/rhcsmisc1-server-req.sh failed
(returned 150)
```

Når vi fjerner filen blir service tatt ned og får status "recoverable".

Melding i syslog:

```
Jan 26 13:04:35 halv2-test01 rgmanager[986]: Executing /etc/cluster-script/rhcsmisc1-server-req.sh stop
Jan 26 13:04:35 halv2-test01 rgmanager[4345]: Service service:rhcsmisc1 is recovering
Jan 26 13:04:37 halv2-test01 rgmanager[4345]: #2: Service service:rhcsmisc1 returned failure code. Last Owner: halv2-test02-hb
Jan 26 13:04:37 halv2-test01 rgmanager[4345]: #4: Administrator intervention required.
```

Funksjonaliteten rundt slike betingelser er mao ikke støttet "out-of-the box".

Kravet er derfor ikke tilfredstillt slik det står formulert, men etter diskusjon i gruppen har vi funnet at kravet er *feilformulert*.

Det kravet egentlig formulerer er at ikke alle ressurser som feiler skal måtte medføre at servicen flyttes eller restartes, og dette lar seg enkelt implementere i RHCS ved å endre i ressursscriptene.

Kode i cluster.conf for å legge inn denne ekstra ressursen med betingelse:

```
<script file="/etc/cluster-script/rhcsmisc1-server-req.sh" name="halv2-server-req">
RESTEN AV RESSURSENE
</script>
```

Vi kan **ikke** styre om service skal flytte til annen node eller restartes ut fra en gitt betingelse slik som formulert i kravet siden failover-egenskapene er satt en gang for alle på servicen og ikke kan endres mens man kjører.

FU10 Restart av ressurser og tjenester uten varsel

Alle endringer på service eller ressurs-nivå gir restart av det nivået man har endret på og de som er avhengige at denne ressursen.

Man får ingen forvarsel om disse restartene når man gjør endringer.

Dette ble verifisert med at vi gjorde endringer med tre offisielle metoder:

- editere rett i cluster.conf og kjøre cman_tool version -r
- gjøre endringer i Conga

Endringer i failoverdomain eller fence-metode gir ingen restart.

Noen mangler rundt logging.

RHCS under RHEL6 har bedre logging enn i RHEL5, f.eks gir loggene i /var/log/cluster/rgmanager.log veldig god informasjon om tilstandene til ressursene. Det virker som om enkeltkomponentene i RHCS logger tilfredstillende, men det mangler logging og sanitychecks på mer overordnede strukturer. F.eks gis det ingen gode logger på alvorlige feil i cluster.conf som gjør at clusteret faktisk ikke fungerer. Eksempler på slike feil er duplikate nodeid på noder i clusteret eller fence-metoder som ikke finnes.

FU12 Begrenset støtte for avhengigheter mellom services

RHCS støtter trivelle dependencies "out of the box" (via depend og depend_mode for servicen). Se DOC-26981 (hos RedHat eller lokal kopi)Det ser ut som om mer avanserte avhengigheter kan implementeres med EventScript.

Se også

<http://sources.redhat.com/cluster/wiki/EventScripting>

Støtte for dependencies er et "Ønske" og ikke et absolutt krav, og kun aktuelt for SAP som hvor det finnes et eget resource-kit hvor dependencies er implementert.

Gjorde en enkel test: rhcsoracle skal være avhengig av rhcsmisc1.

Fra cluster.conf: depend="service:rhcsmisc1" depend_mode="hard"

Legger inn avhengighet emens rhcsmisc er nede: Vi får ikke melding i syslog om at oracle ikke kan starte, men den starter automatisk når rhcsmisc er oppe. Når begge er oppe og vi tar ned rhcsmisc, blir også rhscoracle tatt ned og vi får følgende melding:

```
Jan 12 14:25:17 halv2-test01 clurgmgrd[6998]: <warning> Stopping service service
:rhcsoracle: Dependency missing
Jan 12 14:25:17 halv2-test01 clurgmgrd[6998]: <notice> Stopping service service:
rhcsoracle
```

DE04 Manuelt inngrep ved bortfall av all fc mot disk

Ved bortfall av all fibrechannel connectivity for en node, blir servicene på noden tatt ned og status blir satt til "failed". Dermed blir ikke servicene flyttet til neste node og vi får nedetid for kunden. Man må selv manuelt fjerne lvm-tags og starte opp tjenestene.

Årsaken er at noden uten fc ikke klarer å fjerne lvm-tag'ene på volumgruppen. Løsningen i RHEL5.6 eller RHEL6 hvor man kan kjøre HA-LVM sammen med clvmd. Dette har vi ikke testet, og generelt ønsker vi ikke å benytte clvmd.

Denne feilen er ingen "single failure" men to feil på en gang, og er ikke regnet som kritisk. Setter derfor testen til OK.

DE14 og DE18 Node mister power eller HB-nett/iDRAC6 nett - manuelt inngrep nødvendig

Dersom en node blir utsatt for en av to hendelser:

- 1. blir strømløs
- 2. mister heartbeatnett og nett til iDRAC6 managmentkort

vil ikke fencing av noden vi IPMI fungere.

Her er et eksempele fra test DE18 som innbærer situasjonen beskrevet i 2) ovenfor.

Node 1 forsøker å fence node 2:

```
Jan 26 10:08:20 halv2-test01 fenced[3262]: fencing node halv2-test02-hb
Jan 26 10:08:40 halv2-test01 fenced[3262]: fence halv2-test02-hb dev 0.0 agent
```

Dette går vil "evig tid". Servicene vil forbli på noden fordi resten av nodene ikke kan være sikre på at den har forlatt clusteret; disse nodene insisterer på å fence noden. RedHat har imidlertid lagt opp til at man kan gjøre en manuell fencing, dvs at vi som operatører kan fortelle clusteret at noden ikke lenger kan skrive til disk og dermed forårsake korupsjon.

Kjøring av manuell fencing ser slik ut:

```
[root@halv2-test01 cluster]# fence_ack_manual halv2-test02-hb
About to override fencing for halv2-test02-hb.
Improper use of this command can cause severe file system damage.

Continue [NO/absolutely]? absolutely
Done
```

Noden betraktes nå som nede og servicene flyttes til neste node.

Vi ser altså at disse situasjonene ikke gir automatisk failover og dermed nedetid på tjenestene.

Skal man ha automatisk failover av services i denne sitasjonen kan man legge inn en ny sekundær fencing-metode, feks

- strømmeskinne fencing
- scsi persistent reservation
- fc fencing

Denne feilen er ingen "single failure" men to feil som inntreffer på en gang, og er ikke regnet som kritisk. Setter derfor testen til OK.

=====

Appendix 25 Oppsett av pakken rhcsoracle - Noen notater

=====

Runde 1 : Få opp en service med filsystemer og ip-nr

Først i Conga:

Definerte en service "rhcsoracle" og deretter en ressurs for IP (fant ny IP fra bofh) og en for en volumgruppe (LVM) samt et ext3 filsystem som barn av denne igjen. VG og filsystem definerte i OS fra før.

Valgte så å starte tjenesten, hvilket ikke ga noen feilmelding og tjenestens navn ble nå listet i grønt.

Viste seg dog at det ikke var startet (som vist via "clustat").

Manuell start av tjenesten med "clusvcadm -e rhcsoracle" ga melding om feilet oppstart.

Leting i loggene ga forholdsvis klar beskjed om at det måtte aktiveres for LVM-tagger!

Enblet LVM-tags ved å definere følgende i /etc/lvm/lvm.conf (på alle noder):

globalt:
/etc/lvm/lvm.conf

I seksjonen "activation":

```
volume_list=["@halv2-test01.uio.no", "@halv2-test01-hb1"]
```

Denne siste må også inn her og ikke i en egen fil da vi har root (/) på LVM og kun lvm.conf tas med i initrd av normale rutiner.

NB: Husk altså å kjøre "mkinitrd" etter endringene!

Først brukte vi kun hostnavnet (tilsvarende "\$(uname -n)") men det viste seg at RHCS vil sette tag lik det hostnavn som brukes for heartbeat-nettet (og uten domene-del), dvs. i vårt tilfelle "halv2-test0[123]-hb1". RHCS sanity-sjekk sjekker at det hostnavn (den taggen) den vil bruke er lovlig/inne i volum-listen i lvm.conf.

Tilleggsfunn: RHCS sanity-sjekk ved enabling av tjenester inkluderer også å sjekke om initrd er nyere enn lvm.conf!

Satte deretter tag på "internvg" så denne kan aktiveres ved boot:
vgchange --addtag \$(uname -n) internvg (merk: vi bruker kun ordinært hostnavn som tag på interne volumgruppen)
Bekreftet med "vgs -o +tags".

Laget deretter ny initrd-fil som inkluderer vår nye "lvm.conf":
mkinitrd -v -f /boot/initrd-<kjerne>.img <kjerne>

Deretter var det mulig å aktivere tjenesten ("clusvcadm -e rhcsoracle").
Filsystemet ble montert der det skulle, og bond0-grensesnittet fikk IP-adressen lagt til.

Merknad 1: ip-adressen i servicen sees kun med "ip addr list", dvs ikke med "ifconfig".

Merknad 2: Man ser ingen lvm-tags på volumgruppen til services. Det er derimot satt lvm-tags på de logiske volumene. Disse kan man se med

```
# lvs -o +lv_tags
```

Merknad 3: Også IP-nr må inngå i en parent-relasjon til Oracle-instansen siden den må være på plass for lytteprosessen kan startes.

Appendix 26 Infrastruktur

I forbindelse med uttesting av RedHat Cluster Suite i Halv2-prosjektet er det satt opp en test-rigg med 3 noder i rack M22,12. Dette dokumentet beskriver konfigurasjon og øvrig oppsett.

Switcher

- To switcher tilgjengelig kun for oss: sw-248-72 og sw-248-200
- Portene #1-10 er satt i VLAN566 (database-test)
- Portene #11-20 er satt i VLAN386 (Heartbeat og management-nett)

Nettverksswitchdata

Navn	VLAN566	VLAN386	1	3	5	11	12	13	15	17	19	24
sw-248-72	1-10	11-20	halv2-test01#eth0	halv2-test02#eth0	halv2-test03#eth0	halv2-test01-lc	halv2-test02-lc	halv2-test03-lc	halv2-test01#eth1	halv2-test02#eth1	halv2-test03#eth1	uplink
sw-248-200	1-10	11-20	halv2-test01#eth4	halv2-test02#eth4	halv2-test03#eth4				halv2-test01#eth5	halv2-test02#eth5	halv2-test03#eth5	uplink

Servere

- 3 stk. Dell PowerEdge R710
 - halv2-test01
 - halv2-test02
 - halv2-test03
- 4 interne nettverksgrensesnitt eth0-3 (logisk to stk. dual-port) samt et ekstra kort med 2 NIC eth4-5.
- eth0 og eth4 bondet i bond0, disse er kablet i hver sin switch (0 til -72, 4 til -200) og satt i VLAN566.
- eth1 og eth5 bondet i bond1, disse er kablet i hver sin switch (1 til -72, 5 til -200) og satt i VLAN386 - dette er heartbeat-nett (*cluster communication network*)
- iDRAC-kort er kablet opp i et management-nett som også vil være fencing-nettverk slik:
 - Kablet i sw-248-72. Kan diskuteres om det burde være variert?
 - Portene er satt i VLAN386 som er heartbeat-nettet.

Serverdata

Servernavn	Public nett	iDRAC6	heartbeat-nett
halv2-test01	129.240.147.151 - bond0 (eth0/eth4)	129.240.197.183	129.240.197.180 - bond1 (eth1/eth5)
halv2-test02	129.240.147.152 - bond0 (eth0/eth4)	129.240.197.184	129.240.197.181 - bond1 (eth1/eth5)
halv2-test03	129.240.147.153 - bond0 (eth0/eth4)	129.240.197.185	129.240.197.182 - bond1 (eth1/eth5)

Alt utstyr plassert i rack M22,12.

Nettoversikt:

129.240.147.128/25: VLAN566 (database-test)
129.240.197.176/28: VLAN386 (heartbeat- og management-nett)

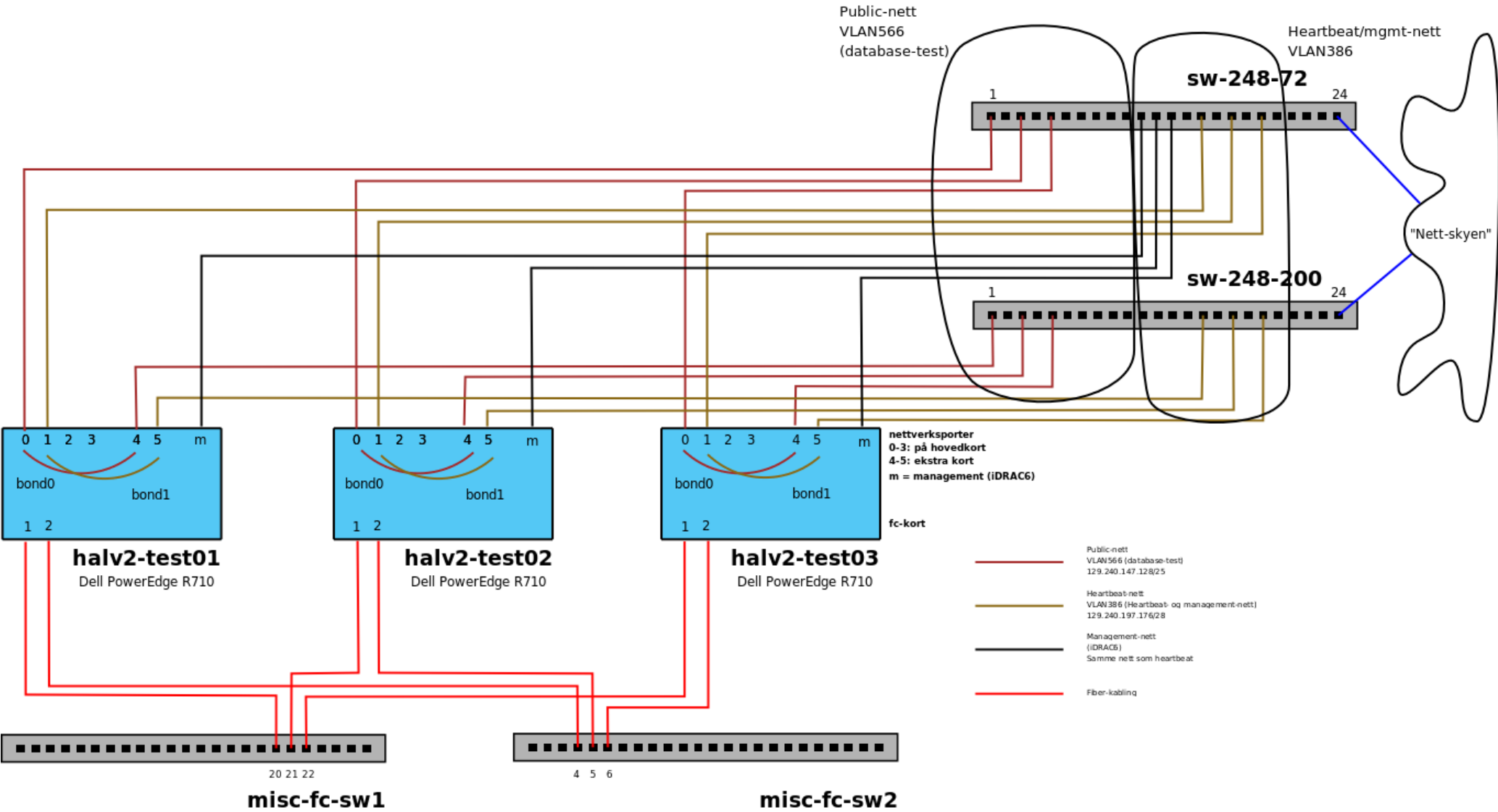
FC-switcher

Vår forbindelse til SAN (foreløpig **misc-eva01**) går via to FC-switcher. Disse er ikke eksklusivt tilgjengelig for prosjektet.

FC-switchdata

Navn	Rack	halv2-test01#fc1	halv2-test01#fc2	halv2-test02#fc1	halv2-test02#fc2	halv2-test03#fc1	halv2-test03#fc2
misc-fc-sw-1	M23,12	20		21		22	
misc-fc-sw-2	M22,12		4		5		6

Systemskisse



Generated on: 2011-01-04 13:43 UTC.

Appendix 27 Konfigurasjon av OS før man lager RHCS services

Versjon 1.0 steira 2011-01-14

Dette dokumentet beskriver kort hvilke endringer man må gjøre i OS'et før man lager RHCS pakker i et RHCS-cluster.

Dokumentet er en huskeliste mer enn en oppskrift for hvordan alle endringer utføres.

Innhold:

- [Oppsett av lvm2 tags \(HA-LVM\)](#)
- [Ikke bruke "primary device" på bondede nettverksgrensesnitt](#)
- [Diskusjon rundt multipath.conf parametre](#)
- [Kjerneparametre](#)

Oppsett av lvm2 tags (HA-LVM)

Dersom man skal lage RHCS-services som inneholder ressurser av typen vg (volumgruppe) og fs (filsystem), må man sørge for å endre i lvm-oppsettet på noden slik at lvm2-tags er støttet. RedHat kaller dette oppsettet for *HA-LVM*, men det er i virkeligheten bare lvm2 tags.

I `lvm.conf` må legge inn nodenavnene til noden slik den er lagt inn i clusteret. Vi bruker hostnavnet som representerer ip-nummeret som er på heartbeat-nettet i RHCS (du ser hvilket nodenavn som er brukt av RHCS ved å kjøre kommandoen `clustat`)

1) Endring av lvm.conf

`lvm.conf` må endres på alle noder i clusteret. Her er et eksempel på alle endringer vi gjorde på node1 i clusteret:

Globalt:

```
tags { hosttags = 1 }
```

I seksjonen *"devices"*:

```
filter = [ "a/dev/mpath/.*/", "a/dev/cciss/.*/", "a/dev/sda[0-9]/", "r/.*/" ]
```

..

```
types = [ "device-mapper", 1]
```

I seksjonen *"activation"*:

```
volume_list=["@halv2-test01.uio.no", "@halv2-test01-hb1"]
```

Denne siste må også inn her og ikke i en egen fil da vi har root (/) på LVM og kun `lvm.conf` tas med i `initrd` av

normale rutiner.

RHCS sanity-sjekk ved enabling av tjenester inkluderer også å sjekke om initrd er nyere enn `lvm.conf`!

2) Sette tags på intervvg

Settderetter tag på "*internvg*" så denne kan aktiveres ved boot:

```
# vgchange --addtag $(uname -n) internvg
```

Merk at vi bruker kun ordinært hostnavn som tag på interne volumgruppen.

Bekreft at tag er satt med "`vgs -o +tags`".

3) Lage ny initrd

Lag deretter ny initrd-fil som inkluderer vår nye "`lvm.conf`":

```
# mkinitrd -v -f /boot/initrd-<kjerne>.img<kjerne>
```

Ikke bruke "primary device" på bondede nettverksgrensesnitt

Dersom man setter en device som inngår i en bondet interface som *primary*, forsøker RHEL å faile tilbake til denne devicen straks den får link igjen etter et evt bortfall. I tester av RHCS så at dette var uheldig fordi link var oppe før kommuniasjonen på interfacet fungerte, og dermed ble noden fenced av RHCS fordi den ikke var tilgjengelig over heartbeatnettet for de andre nodene over dette nettet.

Løsningen er enten på justere på noen timeouts i RHCS, eller å fjerne primary slik at nettrafikken går over et gitt nettverksinterface til det er utilgjengelig uten å faile over til primary så fort som mulig.

Dette ble endret slik for alle bondede interfaces (bond0 og bond1), her vist ved et eksempel fra node1 sin bond1:

Før endring

```
BONDING_OPTS="mode=1 miimon=100 primary=eth1"
```

Etter endring:

```
BONDING_OPTS="mode=1 miimon=100"
```

Diskusjon rundt multipath.conf parametre

Når en node mister all tilgang til disk aksessert over fc (f.eks ved at man frakobler alle fc-kabler på noden) vil man fra OS'et i lang tid tilsynelatende kunne skrive til filer og tilogmed gjøre "`ls`" i filsystemer som egentlig er frakoblet.

Man kan korte ned tiden det tar før OS'et oppdager at diskene er borte ved å justere på parameteren `no_path_retry` i `/etc/multipath.conf`.

Fra mansiden:

`no_path_retry` Specify the number of retries until disable queueing, or fail for immediate failure (no queueing), queue for never stop queueing. Default is 0.

RedHat anbefaler at man bruker den verdien som er anbefalt av leverandøren av lagringsløsningen, og mot HP EVA8k er verdien 12 anbefalt. Dersom man setter parameteren `no_path_retry FAIL` vil filsystemet bli utilgjengelig "straks".

Vi velger å kjøre med HP sin anbefaling.

Kjerneparametre

Dersom noden i clusteret går tom for minne, ønsker vi ikke at tilfeldige prosesser skal bli drept. Derfor anbefaler RedHat at man gjør en panic på noden dersom den går tom for minne.

Sett følgende i `sysctl.conf`:

```
vm.panic_on_oom = 1
```

Appendix 28 Failoverdomains

Korrekt oppsett av failoverdomain er viktig. Her er dokumentasjon av dette fra RedHat:

<http://sources.redhat.com/cluster/wiki/FailoverDomains>

På USIT ønsker vi å kunne definere hvilke noder en service skal gå på, men vi ønsker ikke at en service skal flytte tilbake til sin opprinnelig node dersom denne plutselig blir tilgjengelig igjen. Dette vil jo gi helt utforutsigbare nedetider.

Derfor setter vi gjerne opp våre failoverdomain som følger:

```
restricted="1">
    <failoverdomain name="domain-node1-2-3" nofailback="1" ordered="1"
        <failoverdomainnode name="halv2-test01-hb" priority="1"/>
        <failoverdomainnode name="halv2-test02-hb" priority="2"/>
        <failoverdomainnode name="halv2-test03-hb" priority="3"/>
    </failoverdomain>
```

Appendix 29 Endelig utgave av cluster.conf på RHEL5

Innledning

I løpet av prosjektets workshop med RedHat konsulent for kvalitetsikring av vårt RHCS-oppsett slik at vi opererte i forhold til RedHat sin best practice, ble det gjort noen endringer i cluster.conf. I dette dokumentet blir endelig utgave av cluster.conf presentert, og det blir gitt noen kommentarer om hvorfor verdiene er satt slik de er.

cluster.conf

RHCS-clusteret er definert med

- tre noder
 - halv2-test01
 - halv2-test02
 - halv2-test03
- To fencingmetoder
 - 3 forsøk med fencing via ipmi
 - manuell fencing
- to failoverdomains
 - en med alle tre noder
 - en med to av nodene
- tre tjenester
 - Egenutviklet "HALv2-tjener"
 - Postgres-database
 - Oracle-database.

Slik ser dagens versjon av cluster.conf ut:

```
<?xml version="1.0"?>
<cluster alias="rhcs-test" config_version="103" name="rhcs-test">
  <fence_daemon clean_start="0" post_fail_delay="0" post_join_delay="3"/>
  <clusternodes>
    <clusternode name="halv2-test01-hb" nodeid="1" votes="1">
      <fence>
        <method name="1">
          <device name="halv2-test01-ipmi"/>
        </method>
        <method name="2">
          <device name="halv2-test01-ipmi"/>
        </method>
        <method name="3">
          <device name="halv2-test01-ipmi"/>
        </method>
        <method name="4">
          <device name="manualfence" nodename="halv2-test01-
hb"/>
        </method>
      </fence>
    </clusternode>
    <clusternode name="halv2-test02-hb" nodeid="2" votes="1">
      <fence>
        <method name="1">
          <device name="halv2-test02-ipmi"/>
        </method>
        <method name="2">
          <device name="halv2-test02-ipmi"/>
        </method>
      </fence>
    </clusternode>
  </clusternodes>
</cluster>
```

```

        </method>
        <method name="3">
            <device name="halv2-test02-ipmi"/>
        </method>
        <method name="4">
            <device name="manualfence" nodename="halv2-test02-
hb"/>
        </method>
    </fence>
</clusternode>
<clusternode name="halv2-test03-hb" nodeid="3" votes="1">
    <fence>
        <method name="1">
            <device name="halv2-test03-ipmi"/>
        </method>
        <method name="2">
            <device name="halv2-test03-ipmi"/>
        </method>
        <method name="3">
            <device name="halv2-test03-ipmi"/>
        </method>
        <method name="4">
            <device name="manualfence" nodename="halv2-test03-
hb"/>
        </method>
    </fence>
</clusternode>
</clusternodes>
<cman log_facility="local4" log_level="7"/>
<fencedevices>
    <fencedevice agent="fence_ipmilan" ipaddr="129.240.197.183" login="aaaaa"
name="halv2-test01-ipmi" passwd="****"/>
    <fencedevice agent="fence_ipmilan" ipaddr="129.240.197.184" login="aaaaa"
name="halv2-test02-ipmi" passwd="****"/>
    <fencedevice agent="fence_ipmilan" ipaddr="129.240.197.185" login="aaaaa"
name="halv2-test03-ipmi" passwd="****"/>
    <fencedevice agent="fence_manual" name="manualfence"/>
</fencedevices>
<rm log_facility="local4" log_level="7">
    <failoverdomains>
        <failoverdomain name="domain-node1-2" nofailback="1" ordered="1"
restricted="1">
            <failoverdomainnode name="halv2-test01-hb" priority="1"/>
            <failoverdomainnode name="halv2-test02-hb" priority="2"/>
        </failoverdomain>
        <failoverdomain name="domain-node1-2-3" nofailback="1" ordered="1"
restricted="1">
            <failoverdomainnode name="halv2-test01-hb" priority="1"/>
            <failoverdomainnode name="halv2-test02-hb" priority="3"/>
            <failoverdomainnode name="halv2-test03-hb" priority="2"/>
        </failoverdomain>
    </failoverdomains>
    <resources>
        <fs device="/dev/vg_dbpg1_pgxx/dbpg1" force_fsck="0" force_unmount="1"
fsid="20098" fstype="ext3" mountpoint="/usit/dbpg1" name="pgfs" self_fence="1"/>
        <lvm lv_name="dbpg1" name="pglvm" vg_name="vg_dbpg1_pgxx"/>
    </resources>
    <service autostart="0" domain="domain-node1-2" exclusive="0" name="rhcspg"
recovery="relocate">
        <ip address="129.240.147.155" monitor_link="1">
            <lvm ref="pglvm">
                <fs ref="pgfs">
                    <postgres-8
config_file="/usit/dbpg1/pg_cluster01/data/postgresql.conf" name="DBPG1" postmaster_options="-D
/usit/dbpg1/pg_cluster01/data" postmaster_user="postgres" shutdown_wait="10"/>
                </fs>
            </lvm>
        </ip>
    </service>
    <service autostart="1" domain="domain-node1-2-3" exclusive="0"
name="rhcsoracle" recovery="relocate">
        <ip address="129.240.147.154" monitor_link="1">
            <lvm lv_name="dbora1" name="RHCSORACLE_VG"
vg_name="vg_dbora1_oraxx">
                <fs device="/dev/vg_dbora1_oraxx/dbora1"
force_fsck="0" force_unmount="1" fsid="58742" fstype="ext3" mountpoint="/usit/dbora1"
name="RHCSORACLE_FS" self_fence="1">
                    <oracledb
home="/opt/ora/oracle/product/11.2.0/dbhome_1" listener_name="listener_RHCSTST" name="RHCSTST"
type="10g" user="oracle">
                        <script file="/etc/cluster-
script/rhcsoracle-agent.sh" name="agent_rhcsoracle"/>
                    </oracledb>
                </fs>
            </lvm>
        </ip>
    </service>

```

```

    </service>
    <service autostart="1" domain="domain-node1-2-3" exclusive="0"
name="rhcsmisc1" recovery="relocate">
        <ip address="129.240.147.156" monitor_link="1">
            <lvm lv_name="misc1" name="misc1" vg_name="vg_misc1">
                <fs device="/dev/vg_misc1/misc1" force_fsck="0"
force_unmount="1" fsid="4671" fstype="ext3" mountpoint="/usit/misc1" name="misc1"
self_fence="1">
                    <script file="/etc/cluster-script/rhcsmisc1-
server.sh" name="halv2-server"/>
                </fs>
            </lvm>
        </ip>
    </service>
</rm>
</cluster>
```

Kommentarer til oppsettet

Oppsett av fencing

Dersom en node mister strømmen, vil andre noder ikke kunne fence noden via iDRAC6-kortet og servicene blir ikke failet over. Derfor er det innført en manuell fence-metode som gir mulighet for at operatør kan fortelle de andre nodene at en node kan betraktes som nede. Instruksjoner om hvordan dette kan gjøres blir skrevet til syslog.

For i størst mulig grad å legge opp til at fencing skjer automatisk, har vi lagt inn tre forsøk på å fence over IPMI først.

Dette er satt opp slik (eksempel fra node2 i clusteret):

```

<clusternode name="halv2-test02-hb" nodeid="2" votes="1">
    <fence>
        <method name="1">
            <device name="halv2-test02-ipmi"/>
        </method>
        <method name="2">
            <device name="halv2-test02-ipmi"/>
        </method>
        <method name="3">
            <device name="halv2-test02-ipmi"/>
        </method>
        <method name="4">
            <device name="manualfence" nodename="halv2-test02-
hb"/>
        </method>
    </fence>
</clusternode>
```

Avmontnering av filsystemer

RedHat anbefaler at noden fencer seg selv dersom RHCS ikke klarer å avmontere et filsystem. Dette er satt opp slik:

```

<fs device="/dev/vg_dbora1_oraxx/dbora1" force_fsck="0" force_unmount="1" fsid="58742"
fstype="ext3" mountpoint="/usit/dbora1" name="RHCSORACLE_FS" self_fence="1">
```

Oppsett av failoverdomains

Vi ønsker at services

- automatisk skal overføres til en bestemt node dersom den noden som kjører servicen feiler
- ikke automatisk blir overført tilbake til opprinnelig node når denne er tilgjengelig

Oppsett av failover-domains er forklart nærmere i dette dokumentet:

- [Appendix 28 Failoverdomains](#)

Oppsett av tjenester

Oppsett av servicene er dokumentert i egne leveranser:

HALv2-tjener: [Appendix 32 Sette opp "HALv2-daemon" i RHCS på RHEL5](#)
PostgreSQL: [Appendix 31 Oppsett av PostgreSQL-tjeneste i RHCS på RHEL5](#)
Oracle : [Appendix 30 Oppsett av Oracle-service i RHCS på RHEL5](#)

Extended logging

Mens man tester et cluster er det en fordel å kjøre med mer logging enn default. Dette ble satt opp slik:

```
<cman log_facility="local4" log_level="7"/>
```

Se DOC-5953 for mer detaljer om dette (hos [Redhat](#))

Appendix 30 Oppsett av Oracle-service i RHCS på RHEL5

Innhold

- [Innledning](#)
- [Oppsett av en Oracle Service](#)
- [Oppstart av Oracle Database](#)
- [Oppstart av Oracle Managment Agent](#)

Innledning

Som nevnt nedenfor, inneholder en Oracle-service flere typer ressurser. Vi startet med å sette opp en basal pakke med ip-nr, en volumgruppe og et filsystem som ble montert. For å få dette til å fungere måtte vi ble sette om lvm-tags på nodene. Vi gjorde noen notater fra dette arbeidet

[Appendix 25 Oppsett av pakken rhcsoracle - Noen notater](#)

I RHCS følger det med en resource agent for oppstart av Oracle-databaser i

```
/usr/shar/cluster/oracledb.sh
```

Dette scriptet har endel begrensninger og ulemper fordi det

- forutsetter at `listener.ora` ligger i `$ORACLE_HOME/network/admin` (noe som betyr at man bare kan ha en `listener.ora`)
- starter opp `isqlplus` og *db console*
- starter lytteprosess etter at databasen er startet opp (gir 1 minutt ekstra nedetid for klienter pga dynamisk registrering fra db til lytteprosess)

Oppsett av en Oracle-service

Oracle-servicen *rhcsoracle* i clusteret vårt består av følgende ressurser som settes opp i et parent-child relationship fordi det er avhengigheter mellom ressursene:

- Ip-nr
- Volumgruppe
- Filsystem
- Oracle-database
- Oracle Managment Agent (for GridControl)

Dette kan settes opp i *Conga*, men der er det ikke mulig å sette opp parameteren `listener_name`. Derfor må man endre manuelt i `cluster.conf`.

Slik ser definisjonen av pakken ut i `/etc/cluster/cluster.conf`:

```
<service autostart="1" domain="domain-node1-2-3" exclusive="0" name="rhcsoracle"
recovery="relocate">
  <ip address="129.240.147.154" monitor_link="1">
    <lvm lv_name="dbora1" name="RHCSORACLE_VG" vg_name="vg_dbora1_oraxx">
```

```
<fs device="/dev/vg_dbora1_oraxx/dbora1" force_fsck="0" force_unmount="1" fsid="58742"
fstype="ext3" mountpoint="/usit/dbora1" name="RHCSORACLE_FS" self_fence="0">
<oracledb home="/opt/ora/oracle/product/11.2.0/dbhome_1"
listener_name="LISTENER_RHCSTST" name="RHCSTST" type="10g" user="oracle"/>
<script file="/usit/dbora1/bin/agent_rhcstest"
name="agent_rhcsoracle"/>
</fs>
</lvm>
</ip>
</service>
```

Oppstart av Oracle Database

Endringer gjort av USIT

Scriptet

```
/usr/shar/cluster/oracledb.sh.USIT
```

er laget med følgende enringer i forhold til original-scriptet:

- Oppstart av isqlplus og *db console* er fjernet
- Environment-variable ORACLE_BASE blir satt (hardkoding) fordi Oracle-instansen klager når denne mangler
- Environment-variable TNS_ADMIN blir satt dersom det finnes egen network-katalog under admin/\$ORACLE_SID
- Starter lytteprosess **før** database-instanse

Dette scriptet tas i bruk ved å gjøre følgende på alle noder i clusteret:

```
# cd /usr/share/cluster
# mv oracledb.sh{,.ORG}
# ln -s oracledb.sh.USIT oracledb.sh
```

Tilpassinger som må gjøres før evt produksjonssetting av RHCS

Det bør gjøres noen tilpassinger før scriptet kan kjøres i reell produksjon:

Meldinger ved oppstart av pakke

Det må vurderes om scriptet skal sende email om at oracle blir startet.

Hva skal skje når en ressurs i service er nede?

Man kan konfigurere hva som skal skje når en ressurs (f.eks lytteprosessen eller databasen) i en service er nede

- *relocate* - flytte til en annen node
- *restart* - restarte på samme node
- kombinasjon - f.eks *restart* en gang og deretter *relocate*

Vi som DBA'er har behov for å kunne utføre planlagt restart av databaseinstansene og lytteprosessene, f.eks for å

sette parametre som ikke kan endres i fart eller når man legger inn oppgraderinger. Da ønsker vi ikke at RHCS skal ta ned hele servicen og flytte den til annen node, spesielt ikke dersom det går flere databaser i samme service.

Under SGLX kjørte vi med modifiserte database-script som bare sendte email dersom noen av de monitorerte prosessene feilet. Nyere versjoner av Oracle-toolkit'et for SGLX gir mulighet for å sette et flagg som suspenderer monitorering, og dette kan man bruke når man gjør vedlikehold på databasen.

USIT må bestemme seg for hvordan vi skal håndtere dette under RHCS.

Støtte for flere databaser i samme service

Det må verifiseres at det går fint å kjøre flere database-ressurser under samme service. Den modifiserte koden er lagt opp til å støtte dette, men dette er ikke testet.

Diff mellom originalt og USIT-modifisert oracledb.sh script

```
# cd /usr/share/cluster
# diff -u oracledb.sh.ORG oracledb.sh.USIT
--- oracledb.sh.ORG      2010-12-15 15:41:34.000000000 +0100
+++ oracledb.sh.USIT    2010-12-20 11:06:45.000000000 +0100
@@ -65,7 +65,7 @@
#     to the Oracle user and restart.  Oracle needs to run as the Oracle
#     user, not as root.
#
-#[ -n "$ORACLE_USER" ] || ORACLE_USER=oracle
+[ -n "$ORACLE_USER" ] || ORACLE_USER=oracle

#
# 2. Oracle home.  This is set up during the installation phase of Oracle.
@@ -73,12 +73,12 @@
#     you intend to use as the mount point for your Oracle Infrastructure
#     service.
#
-#[ -n "$ORACLE_HOME" ] || ORACLE_HOME=/mnt/oracle/home
+[ -n "$ORACLE_HOME" ] || ORACLE_HOME=/opt/ora/oracle/product/11.2.0/dbhome_1

#
# 3. This is your SID.  This is set up during oracle installation as well.
#
-#[ -n "$ORACLE_SID" ] || ORACLE_SID=orcl
+[ -n "$ORACLE_SID" ] || ORACLE_SID=RHCSTST

#
# 4. The oracle user probably doesn't have the permission to write to
@@ -91,7 +91,7 @@
#
# 5. Type of Oracle Database.  Currently supported: 10g 10g-iAS(untested!)
#
-[-n "$ORACLE_TYPE" ] || ORACLE_TYPE="base-em"
+[-n "$ORACLE_TYPE" ] || ORACLE_TYPE="10g"

#
# 6. Oracle virtual hostname.  This is the hostname you gave Oracle during
@@ -100,13 +100,35 @@
#[ -n "$ORACLE_HOSTNAME" ] || ORACLE_HOSTNAME=svc0.foo.test.com
```

Full versjon av USIT-modifisert oracledb.sh script

Her er en fullstendig versjon av det USIT-modifiserte oracledb-scriptet:

```
#!/bin/bash
#
```

```
# Copyright 2003-2004, 2006-2008 Red Hat, Inc.
#
# Author(s):
#   Hardy Merrill
#   Lon Hohberger
#   Michael Moon
#
# This program is Open Source software.  You may modify and/or redistribute
# it pursuant to the terms of the Open Software License version 2.1, which
# is available from the following URL and is included herein by reference:
#
#   http://opensource.org/licenses/osl-2.1.php
#
# chkconfig: 345 99 01
# description: Service script for starting/stopping      \
#               Oracle(R) Database 10g on                 \
#               Red Hat Enterprise Linux 5
#
# NOTES:
#
# (1) You can comment out the LOCKFILE declaration below.  This will prevent
# the need for this script to access anything outside of the ORACLE_HOME
# path.
#
# (2) You MUST customize ORACLE_USER, ORACLE_HOME, ORACLE_SID, and
# ORACLE_HOSTNAME to match your installation if not running from within
# rgmanager.
#
# (3) Do NOT place this script in shared storage; place it in ORACLE_USER's
# home directory in non-clustered environments and /usr/share/cluster
# in rgmanager/Red Hat cluster environments.
#
# Oracle is a registered trademark of Oracle Corporation.
# Oracle9i is a trademark of Oracle Corporation.
# Oracle10g is a trademark of Oracle Corporation.
# All other trademarks are property of their respective owners.
#
```

Oppstart av Oracle Managment Agent

Oracle Managment Agent blir startet av RHCS via en script-ressurs i `cluster.conf`. Dette scriptet må tilfredstille endel krav til init-script. Fra dokumentasjonen "[Migrating an HP Serviceguard for Linux \(SGLX\) cluster to Red Hat Cluster Suite \(RHCS\)](#)":

The resource script RA has to comply with the LSB specification which is described in more detail at http://refspecs.freestandards.org/LSB_2.0.1/LSB-Core/LSB-Core/iniscriptact.html. If the rules are not followed, especially the return codes for start, stop, and monitor functions; the resource manager can behave inconsistently.

Siden RedHat følger LSB, har vi benyttet følgende mal:

<https://fedoraproject.org/wiki/Packaging:SysVInitScript>

Slik ser scriptet `/usit/dbora1/bin/agent_rhcstest` ut for oppstart Oracle Managment Agent:

```
#!/bin/sh
#
#
#
# chkconfig:
# description:
#
### BEGIN INIT INFO
# Provides:
# Required-Start:
# Required-Stop:
```

```
# Should-Start:
# Should-Stop:
# Default-Start:
# Default-Stop:
# Short-Description:
# Description:
#### END INIT INFO

# Source function library.
. /etc/rc.d/init.d/functions

exec="/site/etc/init.d/oracle-oma.sh"
prog="agent_rhcsoracle"
arg="/usit/dbora1/ora_drift/agent/agent"

[ -e /etc/sysconfig/$prog ] && . /etc/sysconfig/$prog

lockfile=/var/lock/subsys/$prog

start() {
    [ -x $exec ] || exit 5
    echo -n "Starting $prog: "
    daemon $exec start $arg
    retval=$?
    echo
    [ $retval -eq 0 ] && touch $lockfile
    return $retval
}
```

Dette scriptet kaller `/site/etc/init.d/oracle-oma.sh` som tradisjonelt har vært brukt til å starte opp Oracle Managment Agent

- fra `/site/etc/rc` under oppstart av maskiner
- fra SGLX-pakker for oppstart av OMA for pakker

Scriptet `oracle-oma.sh` har som første argument `"start|stop|status"` og som andre argument *kan* man angi `ORACLE_HOME` til agenten som skal startes opp. Scriptet `/usit/dbora1/bin/agent_rhctest` som kalles opp fra RHCS tar ikke slike ekstra argumenter, så derfor er `ORACLE_HOME` for agenten lagt inn som et argument i kallet til `oracle-oma.sh`.

Slik ser `oracle-oma.sh` ut:

```
#!/bin/sh
#
#
#
# chkconfig:
# description:

#### BEGIN INIT INFO
# Provides:
# Required-Start:
# Required-Stop:
# Should-Start:
# Should-Stop:
# Default-Start:
# Default-Stop:
# Short-Description:
# Description:
#### END INIT INFO

# Source function library.
. /etc/rc.d/init.d/functions

exec="/site/etc/init.d/oracle-oma.sh"
prog="agent_rhcsoracle"
arg="/usit/dbora1/ora_drift/agent/agent"

[ -e /etc/sysconfig/$prog ] && . /etc/sysconfig/$prog
```

```
lockfile=/var/lock/subsys/$prog

start() {
    [ -x $exec ] || exit 5
    echo -n "Starting $prog: "
    daemon $exec start $arg
    retval=$?
    echo
    [ $retval -eq 0 ] && touch $lockfile
    return $retval
}
```

Det er verdt å notere seg at `oracle-oma.sh` må gjennomgå slik at det gir riktige returkoder og oppfører seg i henhold til hva som forventes av LSB-kompatible script.

Appendix 31 Oppsett av PostgreSQL-tjeneste i RHCS på RHEL5

I RHCS følger det med en resource agent for oppstart av PostgreSQL-databaser i

```
/usr/shar/cluster/postgres-8.sh
```

Dette scriptet er i utgangspunktet laget for å bruke med PostgreSQL 8.x, men kan uten problemer tilpasses for bruk også med PostgreSQL 9.x

Skriptet er ennå ikke gått gjennom i detalj i forhold til om det kan brukes for å starte PostgreSQL-baser slik det er vanlig på USIT. Men det er testet for kjøring av en enkel, ordinær singel instans. I den forbindelsen viser det seg at det er nødvendig med en enkel endring (se nedenfor), men forøvrig fungerer den som det er.

Oppsett av en PostgreSQL-service

Postgres-servicen *rhcspg* i clusteret vårt består av følgende ressurser som settes opp i et parent-child relationship fordi det er avhengigheter mellom ressursene:

- Ip-nr
- Volumgruppe
- Filsystem
- PostgreSQL-instans

Dette kan settes opp i *Conga*, eller manuelt i `cluster.conf`. For å teste separate ressurs-definisjoner som så refereres av tjenesten, er *LVM* og *filsystem*-ressursene beskrevet utenfor tjeneste-seksjonen, og deretter dratt inn dit ved å bruke *ref*-opsjonen.

Slik ser definisjonen av tjenesten ut i `/etc/cluster/cluster.conf`:

```
<resources>
  <fs device="/dev/vg_dbpg1_pgxx/dbpg1" force_fsck="0" force_unmount="1" fsid="20098"
fstype="ext3" mountpoint="/usit/dbpg1" name="pgfs" self_fence="0"/>
  <lvm lv_name="dbpg1" name="pglvm" vg_name="vg_dbpg1_pgxx"/>
</resources>
<service autostart="0" domain="domain-node1-2" exclusive="0" name="rhcspg" recovery="relocate">
  <ip address="129.240.147.155" monitor_link="1">
    <lvm ref="pglvm">
      <fs ref="pgfs">
        <postgres-8 config_file="/usit/dbpg1/pg_cluster01/data/postgresql.conf"
name="DBPG1" postmaster_options="-D /usit/dbpg1/pg_cluster01/data" postmaster_user="postgres"
shutdown_wait="10"/>
      </fs>
    </lvm>
  </ip>
</service>
```

Oppstart av PostgreSQL

Endringer gjort av USIT

Scriptet

```
/usr/shar/cluster/postgres-8.sh.USIT
```

er laget med følgende enring i forhold til original-scriptet:

- Redefinert PSQL_POSTMASTER (til "/usr/pgsql-9.0/bin/postmaster") da dette ikke er del av meta-dataene

Dette scriptet tas i bruk ved å gjøre følgende på alle noder i clusteret:

```
# cd /usr/share/cluster
# mv postgres-8.sh{,.ORG}
# ln -s postgres-8.sh.USIT postgres-8.sh
```

Tilpassinger som må gjøres før evt produksjonssetting av RHCS

Det bør gjøres noen tilpassinger før scriptet kan kjøres i reell produksjon:

Meldinger ved oppstart av pakke

Scriptet må endres, slik at det sender en e-post ved tilfeller der pakken går ned eller blir startet opp.

Hva skal skje når en ressurs i service er nede?

Man kan konfigurere hva som skal skje når en ressurs (f.eks lytteprosessen eller databasen) i en service er nede

- *relocate* - flytte til en annen node
- *restart* - restarte på samme node
- kombinasjon - f.eks *restart* en gang og deretter *relocate*

USIT må bestemme seg for hvordan vi skal håndtere dette under RHCS.

Støtte for flere databaser i samme service

Selv om det tilsynelatende ser ut til at scriptet støtter kjøring av flere postgres-pakker på samme maskin, så må det verifiseres at det går fint å kjøre flere database-ressurser under samme service. Skriptet er foreløpig ikke testet for dette.

Ved oppstart av en pakke, så lager scriptet en kopi av postgresql.conf som ligger under /usit/#PAKKENAVN#/pg_cluster01/data/postgresql.conf, som brukes som config-fil for PostgreSQL. Denne kopien lagres under et filpath som inneholder pakkenavnet.

Diff mellom originalt og USIT-modifisert oracledb.sh script

```
--- postgres-8.sh.ORG      2010-12-21 11:33:19.0000000000 +0100
+++ postgres-8.sh        2010-12-22 12:03:01.0000000000 +0100
@@ -32,7 +32,7 @@
. $(dirname $0)/utils/messages.sh
. $(dirname $0)/utils/ra-skelet.sh

-declare PSQL_POSTMASTER="/usr/bin/postmaster"
+declare PSQL_POSTMASTER="/usr/pgsql-9.0/bin/postmaster"
declare PSQL_pid_file="\generate_name_for_pid_file"
declare PSQL_conf_dir="\generate_name_for_conf_dir"
declare PSQL_gen_config_file="$PSQL_conf_dir/postgresql.conf"
```

Full versjon av USIT-modifisert oracledb.sh script

Siden det kun er en så enkel endring av skriptet som må til vises ikke fullstendig listing av skriptet. Det henvises istedet til

```
/usr/share/cluster/postgres-8.sh
```

på en hvilken som helst RHCS 5.5-installasjon.

Appendix 32 Sette opp "HALv2-daemon" i RHCS på RHEL5

Selve HALv2-daemonen

Det er laget en pl-modul som

- åpner socket på angitt ip-nr og port
- gjør echo på tekst som skrives til socket og logger teksten.
- logger timestamp regelmessig

Her er koden til HALv2-daemonen:

```
#!/usr/bin/perl

package Halv2Server;

use strict;
use warnings;
use base qw(Net::Server::Fork);
use Getopt::Long;
use threads;

#-----
# Main program
#-----

# Options with default values
my %opt = ( 'log'      => '/tmp/halv2-server01.log',
            'ip'       => '*',
            'port'     => 9999,
            'user'     => undef,
            'group'    => undef,
            'help'     => 0,
            );

# Get options
GetOptions('l|log=s'      => \$opt{log},
           'i|ip=s'      => \$opt{ip},
           'p|port=i'    => \$opt{port},
           'u|user=s'    => \$opt{user},
           'g|group=s'   => \$opt{group},
           'h|help'      => \$opt{help},
           ) or exit 1;
```

Oppsett av service for HALv2-daemonen

HALv2-daemonen settes opp som en service i RHCS med følgende ressurser i et parent-child relationship:

- IP-nr
- LVM
- Filsystem
- Script Agent (som kaller opp et LSB-kompatibelt init-script)

Slik ser dette ut i /etc/cluster/cluster.conf:

```
<service autostart="1" domain="domain-node1-2-3" exclusive="0" max_restarts="1"
name="rhcsmisc1" recovery="restart" restart_expire_time="10">
  <ip address="129.240.147.156" monitor_link="1">
    <lvm lv_name="misc1" name="misc1" vg_name="vg_misc1">
      <fs device="/dev/vg_misc1/misc1" force_fsck="0" force_unmount="1" fsid="4671"
fstype="ext3" mountpoint="/usit/misc1" name="misc1" self_fence="0">
        <script file="/usit/misc1/halv2-server01" name="halv2-server"/>
      </fs>
    </lvm>
  </ip>
</service>
```

Scriptet /usit/misc1/halv2 - server01

Dette scriptet må tilfredstille endel krav til init-script. Fra dokumentasjonen ["Migrating an HP Serviceguard for Linux \(SGLX\) cluster to Red Hat Cluster Suite \(RHCS\)"](#):

The resource script RA has to comply with the LSB specification which is described in more detail at http://refspecs.freestandards.org/LSB_2.0.1/LSB-Core/LSB-Core/iniscriptact.html. If the rules are not followed, especially the return codes for start, stop, and monitor functions; the resource manager can behave inconsistently.

Siden RedHat følger LSB, har vi benyttet følgende mal:

<https://fedoraproject.org/wiki/Packaging:SysVInitScript>

Scriptet for HALv2-daemonen ser slik ut:

```
#!/bin/sh
#
#
# chkconfig: 345 57 23
# description:

# Source function library.
. /etc/rc.d/init.d/functions

exec="/usit/misc1/halv2-server01.pl"
arg="-l /usit/misc1/log -i 129.240.147.156 -u nobody -g nobody"
prog="halv2-server01.pl"

[ -e /etc/sysconfig/$prog ] && . /etc/sysconfig/$prog

lockfile=/var/lock/subsys/$prog

start() {
  [ -x $exec ] || exit 5
  echo -n $"Starting $prog: "
  daemon $exec $arg &
  retval=$?
}
```

Appendix 33 Notater rundt oppsett av RHCS på RHEL6

Oppsett av Cluster og services vha Conga

Momenter rundt oppsett av RHCS på RHEL6:

- Installasjonsguide for RHCS inneholder ikke lenger kjøring av `luci_admin init` lenger
- Navn på cluster måtte inneholde et tall. Navnet på clusteret ble derfor `rhctest1`
- Innlegging av `fence-devices` er forskjellig.
 - Man legger inn en metode og gir den `name=1`
 - Etter velger man instanse
- Oppsett av filsystemressurs `fs`
 - Hva er "use quick status checks" i setup av `fs` (ført inn i "Løse tråder")
- Man kan ikke bruke samme navn på flere ressurser i samme clusteroppsett, da blir de overskrevet senere.
- Oppdatering av `cluster.conf` skjer nå med `cman_tool version -r`
- Innlegging av `ip` ressurser
 - Ikke angi nettmasker, og dersom du oppgir det må du bruke CIDR-syntaks (i.e. `/xx`) for nettmasker.
- Innlegging av logisk volum ressurser
 - logiske volumer har `self_fence` opsjon slik som `fs` ressursen som medfører at noden blir rebootet dersom man ikke klarer å deaktivere volumgruppen.
- det fungerte **ikke** å legge inn logging fra `cman` ved å legge inn `<cman log_facility="local4" log_level="7"/>`
- `force_unmount="1"` for filsystem (ressurs av typen `fs`) blir overskrevet av Conga dersom man editrer `fs`-ressursen der.
- `fuser` (pga `force_unmount="1"` i `fs`-ressursen) fungerer ikke, problemer med SELinux:

```
Jan 21 15:15:32 halv2-test02 setroubleshoot: SELinux is preventing
/sbin/fuser "sigkill" access . For complete SELinux messages. run sealert
-l ef12445f-ebb2-420f-887d-2c4e77a89a90
```

Vi har feilmeldt dette <https://access.redhat.com/support/cases/00411039>

Anbefalt metode (fra David Juran) for å fikse dette manuelt: <https://access.redhat.com/kb/docs/DOC-10106>

- Det ligger veldig nyttige logger i `/var/log/cluster` og `/var/log/luci`
- Fencing-metoden `fence_manual` er ikke lenger supporter, jfr <https://access.redhat.com/kb/docs/DOC-30003>. Når man har behov for manuell fencing, bruker man kommandoen `fence_ack_manual` som nå fungerer mot alle typer `fence-agenter`. Man kjører denne kommandoen på den noden som forsøker å fence en node. Dette er dokumentert her: <https://access.redhat.com/kb/docs/DOC-43212>. Dermed har vi i RHCS for RHEL5 endret i oppsett rundt fencing for en node:
 - RHEL5: Vi lar `fence_impilan` få tre sjanser, så bruker vi manuell fencing:

```
<clusternode name="halv2-test02-hb" nodeid="2" votes="1">
  <fence>
    <method name="1">
      <device name="halv2-test02-ipmi"/>
    </method>
    <method name="2">
      <device name="halv2-test02-ipmi"/>
    </method>
    <method name="3">
      <device name="halv2-test02-ipmi"/>
    </method>
  </fence>
</clusternode>
```

```

        <method name="4">
            <device name="manualfence" nodename="halv2-test02-
hb"/>
        </method>
    </fence>
</clusternode>

```

- o RHEL6: Vi bruker fence_impilan som vil forsøke å fence en node til "evig tid" (man vi kan bruke kommandoen fence_ack_manual)

```

<clusternode name="halv2-test01-hb" nodeid="1" votes="1">
    <fence>
        <method name="1">
            <device name="halv2-test01-ipmi"/>
        </method>
    </fence>
</clusternode>

```

Endelig versjon av /etc/cluster/cluster.conf for RHCS på RHEL6

Her er endelig versjon av /etc/cluster/cluster.conf på RHEL6:

```

<?xml version="1.0"?>
<cluster config_version="50" name="rhctest01">
    <fence_daemon clean_start="0" post_fail_delay="0" post_join_delay="3"/>
    <clusternodes>
        <clusternode name="halv2-test01-hb" nodeid="1" votes="1">
            <fence>
                <method name="1">
                    <device name="halv2-test01-ipmi"/>
                </method>
            </fence>
        </clusternode>
        <clusternode name="halv2-test02-hb" nodeid="2" votes="1">
            <fence>
                <method name="1">
                    <device name="halv2-test02-ipmi"/>
                </method>
            </fence>
        </clusternode>
        <clusternode name="halv2-test03-hb" nodeid="3" votes="1">
            <fence>
                <method name="1">
                    <device name="halv2-test03-ipmi"/>
                </method>
            </fence>
        </clusternode>
    </clusternodes>
    <cman/>
    <fencedevices>
        <fencedevice agent="fence_ipmilan" auth="md5" ipaddr="129.240.zzz.zzz" login="yyy"
name="halv2-test01-ipmi" passwd="xxx"/>
        <fencedevice agent="fence_ipmilan" auth="md5" ipaddr="129.240.zzz.sss" login="yyy"
name="halv2-test02-ipmi" passwd="xxx"/>
        <fencedevice agent="fence_ipmilan" auth="md5" ipaddr="129.240.zzz.ttt" login="yyy"
name="halv2-test03-ipmi" passwd="xxx"/>
    </fencedevices>
    <rm log_facility="local4" log_level="7">
        <failoverdomains>
            <failoverdomain name="domain-node1-2-3" nofailback="1" ordered="1" restricted="1">
                <failoverdomainnode name="halv2-test01-hb" priority="1"/>
                <failoverdomainnode name="halv2-test02-hb" priority="2"/>
                <failoverdomainnode name="halv2-test03-hb" priority="3"/>
            </failoverdomain>
            <failoverdomain name="domain-node3-2-1" nofailback="0" ordered="1" restricted="1">
                <failoverdomainnode name="halv2-test01-hb" priority="3"/>
                <failoverdomainnode name="halv2-test02-hb" priority="2"/>
                <failoverdomainnode name="halv2-test03-hb" priority="1"/>
            </failoverdomain>
        </failoverdomains>
        <resources>
            <ip address="129.240.147.156" monitor_link="on" sleeptime="10"/>
            <lvm lv_name="misc1" name="lv_misc1" self_fence="on" vg_name="vg_misc1"/>
            <script file="/etc/cluster-script/rhcsmisc1-server.sh" name="halv2-server"/>
            <fs device="/dev/vg_misc1/misc1" force_unmount="1" fsid="4671"
mountpoint="/usit/misc1" name="fs_misc1" self_fence="on"/>
        </resources>
        <service autostart="0" domain="domain-node3-2-1" exclusive="0" name="rhcspg"

```

```
recovery="relocate">
  <ip address="129.240.147.155" monitor_link="1">
    <lvm ref="pglvm">
      <fs ref="pgfs">
        <postgres-8 config_file="/usit/dbpg1/pg_cluster01/data/postgresql.conf"
name="DBPG1" postmaster_options="-D /usit/dbpg1/pg_cluster01/data -r
/tmp/postgresql_server.log" postmaster_user="postgres" shutdown_wait="10"/>
      </fs>
    </lvm>
  </ip>
</service>
<service autostart="1" domain="domain-node3-2-1" exclusive="0" name="rhcsmisc1"
recovery="relocate">
  <ip ref="129.240.147.156">
    <lvm ref="lv_misc1">
      <fs ref="fs_misc1">
        <script ref="halv2-server"/>
      </fs>
    </lvm>
  </ip>
</service>
<service autostart="1" domain="domain-node1-2-3" exclusive="0" name="rhcsoracle"
recovery="relocate">
  <ip address="129.240.147.154" monitor_link="1" sleeptime="10">
    <lvm lv_name="dbora1" name="RHCSORACLE_VG" self_fence="1"
vg_name="vg_dbora1_oraxx">
      <fs device="/dev/vg_dbora1_oraxx/dbora1" force_unmount="1" fsid="58742"
mountpoint="/usit/dbora1" name="RHCSORACLE_FS" self_fence="0">
        <oracledb home="/opt/ora/oracle/product/11.2.0/dbhome_1"
listener_name="listener_RHCSTST" name="RHCSTST" type="base" user="oracle"/>
      </fs>
    </lvm>
  </ip>
</service>
</rm>
</cluster>
```

Appendix 34 Oppsett av Oracle-service i RHCS på RHEL6

Innhold

- [Oppsett av en Oracle Service](#)
- [Oppstart av Oracle Database](#)
- [Oppstart av Oracle Managment Agent](#)
- [Hva skal skje når en ressurs i service er nede?](#)

Oppsett av en Oracle-service

Oracle-servicen *rhcsoracle* i clusteret vårt består av følgende ressurser som settes opp i et parent-child relationship fordi det er avhengigheter mellom ressursene:

- Ip-nr
- Volumgruppe
- Filsystem
- Oracle-database
- Oracle Managment Agent (for GridControl, ikke testet for RHEL6, kun på RHEL5)

Dette kan settes opp i *Conga* siden det under RHEL6 nå også er mulig å sette opp parameteren `listener_name`.

Slik ser definisjonen av pakken ut i `/etc/cluster/cluster.conf`:

```
<service autostart="1" domain="domain-node1-2-3" exclusive="0" name="rhcsoracle"
recovery="relocate">
  <ip address="129.240.147.154" monitor_link="1">
    <lvm lv_name="dbora1" name="RHCSORACLE_VG" vg_name="vg_dbora1_oraxx">
      <fs device="/dev/vg_dbora1_oraxx/dbora1" force_fsck="0" force_unmount="1" fsid="58742"
fstype="ext3" mountpoint="/usit/dbora1" name="RHCSORACLE_FS" self_fence="0">
        <oracledb home="/opt/ora/oracle/product/11.2.0/dbhome_1"
listener_name="LISTENER_RHCSTST" name="RHCSTST" type="base" user="oracle"/>
      </fs>
    </lvm>
  </ip>
</service>
```

Ved å sette `type="base"` ovenfor blir kun databaseinstansen og lytteprosessen startet. Det er andre settinger som f.eks starter opp `isqlplus`, men det er ikke noe vi på USIT benytter.

Oppstart av Oracle Database

Endringer gjort av USIT i forhold til default oppsett fra RedHat

Scriptet for oppstart av Oracle er

```
/usr/share/cluster/oracledb.sh
```

Dette scriptet blir overskrevet når man oppdaterer RHCS, så det er ønskelig å bruke scriptet umodifisert. Vi møtte følgende utfordringer med dette scriptet

- krav om skriverettighet til katalogen `/var/log/cluster`.
- Environment-variable `ORACLE_BASE` blir ikke satt

- Environment-variable `TNS_ADMIN` blir ikke satt
- Lytteprosess startes **etter** database-instanse

Dersom vi ønsker endringer i scriptet, bør vi teste disse grundig og gi koden til RedHat. Da er det mulig at våre endringer blir tatt med i en fremtidige versjon av scriptet.

De endringer vi måtte gjøre i forhold til default RedHat-oppsett er gjengitt i dette kapittelet.

Krav om skriverettighet til katalogen `/var/log/cluster`

Scriptet til logg-filer i katalogen `/var/log/cluster`. Brukeren *oracle* har normalt ikke skrivrettighet til denne katalogen.

Her er det to løsninger:

- Gi brukeren *oracle* rettighet til å skrive til denne katalogen
- Endre i scriptet slik at det logger til alternativ lokasjon.

Vi ønsket å teste `oracledb.sh` i umodifisert form, og som nødløsning endret vi gruppen på `/var/log/cluster` til *oracle* og gav *group* skrivetilgang. Dette skalerer selvsagt ikke fordi andre services gjerne kjører som andre brukere som også har samme behov for skrivetilgang.

Senere implementerte vi tilgangene med ACL'er, og dette ble gjort som følger:

```
# setfacl -m u:oracle:rwx /var/log/cluster; setfacl -m d:u:oracle:rwx /var/log/cluster
```

Environment-variable `ORACLE_BASE` blir ikke satt

Under oppstart av Oracle-databasen kommer det warnings i *alert-loggen* om at environment-variabelen `ORACLE_BASE` ikke er satt, men databasen fortsetter oppstarten og benytter den verdien som tidligere har vært angitt.

Man kan løse dette ved å modifisere scriptet og legge til denne variabelen, men vi har valgt å kjøre uten denne endringen siden databasen tilsynelatende fungerer.

Environment-variable `TNS_ADMIN` blir ikke satt

Det er ingen mulighet for å sette environment-variabelen `TNS_ADMIN` i scriptet. Når lytteprosessen skal startes opp, blir det antatt at konfigurasjonsfilen `listener.ora` befinner seg i default-lokasjonen `$ORACLE_HOME/network/admin`. I USIT's standaroppsett har vi på alle databaser som er aktuelle å kjøre i RHCS-clustre en egen `listener.ora` pr database, og lokasjonen til denne filen er alltid angitt med environment-variabelen `TNS_ADMIN`. Det er ingen mulighet til å sette denne variabelen med default-oppsettet til RedHat.

Dersom man ønsker å kjøre med umodifisert script, kan man løse dette på to måter:

1. Dersom man kun har *en* database i clusteret
 - Lag en softlink fra `$ORACLE_HOME/network/admin/listener.ora` til `$TNS_ADMIN/listener.ora`
2. Dersom man har *flere* databaser i clusteret
 - Lag en sentral `listener.ora` som inneholde definisjonene til alle lytteprosessene i clusteret

Under testene av RHCS løste vi problemet med å metode 1).
Å innføre metode 2) innebærer at DBA-gruppen sitt standaroppsett må endres.

Starter lytteprosess etter database-instanse

Scriptet starter lytteprosessen *før* databaseinstansen. Som en del av DBA-gruppens standardoppsett blir lytteprosessene satt opp med såkalt dynamisk registrering hvor databaseinstansene selv registrerer seg hos lytteprosessen. Denne registreringen skjer hvert minutt. Når lytteprosessen ikke er tilgjengelig ved oppstart av database-instansen, vil den ikke ha noen lytteprosess å registrere seg hos. Dermed vil klienter som forsøker å koble på oppleve at nedetiden blir opp til ett minutt lengere enn nødvendig.

Det er ikke aktuelt å endre DBA-gruppens policy på dette området, så dersom man må endre på dette bør det gjøres ved å endre i scriptet. Vi kjørte testene uten å gjøre denne endringen.

Andre tilpassinger som kan gjøres før evt produksjonssetting av RHCS

Det er andre tilpassinger som kan gjøre før scriptet kan kjøres i reell produksjon:

- Det må vurderes om scriptet skal sende email om at databasen er blitt startet.
- Det må verifiseres at det går fint å kjøre flere database-ressurser under samme service. Den modifiserte koden er lagt opp til å støtte dette, men dette er ikke testet.

Oppstart av Oracle Managment Agent

Det er verdt å notere seg at oppstartscript for Oracle Managment Agent ("Gridcontrol-agenten") `oracle-oma.sh` må gjennomgå slik at det gir riktige returkoder og oppfører seg i henhold til hva som forventes av LSB-kompatible script.

Hva skal skje når en ressurs i service er nede?

Man kan konfigurere hva som skal skje når en ressurs (f.eks lytteprosessen eller databasen) i en service er nede

- *relocate* - flytte til en annen node
- *restart* - restarte på samme node
- kombinasjon - f.eks *restart* en gang og deretter *relocate*

Vi som DBA'er har behov for å kunne utføre planlagt restart av databaseinstansene og lytteprosessene, f.eks for å sette parametre som ikke kan endres i fart eller når man legger inn oppgraderinger. Da ønsker vi ikke at RHCS skal ta ned hele servicen og flytte den til annen node, spesielt ikke dersom det går flere databaser i samme service.

I RHCS setter man opp en "*freeze*" av services slik at statussjekkene for en service ikke lenger kjører. På den måten vil ikke RCHS "oppdage" at services er nede, og man kan gjøre sitt vedlikehold uten at RHCS griper inn.

Eksempel på å starte en freeze:

```
# clusvcadm -Z rhcsoracle
```

Eksempel på å avslutte en freeze:

```
# clusvcadm -U rhcsoracle
```

Det kommer meldinger i loggfilen `/var/log/cluster/rgmanager.log` om at freeze er enablet for en service.

Appendix 35 Oppsett av PostgreSQL-tjeneste i RHCS på RHEL6

I RHCS følger det med en resource agent for oppstart av PostgreSQL-databaser i

```
/usr/shar/cluster/postgres-8.sh
```

Dette skriptet virker å være bra tilpasset kjøring med de PostgreSQL-pakker som følger med som del av RedHat 6-distribusjonen. Det er også disse pakkene det er bestemt å benytte under denne testen.

Skriptet er ennå ikke gått gjennom i detalj i forhold til om det kan brukes for å starte PostgreSQL-baser slik det er vanlig på USIT. Men det er testet for kjøring av en enkel, ordinær singel instans. Konklusjonen er at i forhold til RHCS så er det ikke behov for noen tilpassninger, dog fungerer ikke PostgreSQL med den SELinux-konfigurasjonen som er default i RHEL 6.

Oppsett av en PostgreSQL-service

Postgres-servicen *rhcspg* i clusteret vårt består av følgende ressurser som settes opp i et parent-child relationship fordi det er avhengigheter mellom ressursene:

- IP-nr
- Volumgruppe
- Filsystem
- PostgreSQL-instans

Dette kan settes opp i *Conga*, eller manuelt i `cluster.conf`. På samme måte som under test av RHCS 5 er det her benyttet separate ressurs-definisjoner som så refereres av tjenesten (med unntak av IP-ressursen).

Slik ser definisjonen av tjenesten ut i `/etc/cluster/cluster.conf`:

```
<resources>
  <fs device="/dev/vg_dbpg1_pgxx/dbpg1" force_unmount="1" mountpoint="/usit/dbpg1"
name="pgfs" self_fence="0"/>
  <lvm lv_name="dbpg1" name="pglvm" vg_name="vg_dbpg1_pgxx"/>
</resources>
<service autostart="0" domain="domain-node1-2" exclusive="0" name="rhcspg" recovery="relocate">
  <ip address="129.240.147.155" monitor_link="1">
    <lvm ref="pglvm">
      <fs ref="pgfs">
        <postgres-8 config_file="/usit/dbpg1/pg_cluster01/data/postgresql.conf"
name="DBPG1" postmaster_options="-D /usit/dbpg1/pg_cluster01/data -r
/tmp/postgresql_server.log" postmaster_user="postgres" shutdown_wait="10"/>
      </fs>
    </lvm>
  </ip>
</service>
```

Oppstart av PostgreSQL

Behov for endringer

Siden det ble benyttet database levert som del av distribusjonen, var det ikke nødvendig å endre noe i skript-ressurs eller annen RHCS-komponent. Oppsett og kjøring av enkel, standard pg-instans fungerte *"out-of-the-box"*!

Det er her kun vurdert om PostgreSQL i prinsippet enkelt kan kjøres under RHCS, det er altså ikke gjort en fullskala-test elelr kjørt med samme miljø og behov som USIT har i dagens produksjon. Det vil derfor være nødvendig å gjøre de samme vurderinger som ble anført etter testen under RHCS 5. Det henvises til [Appendix 31](#)

for flere detaljer rundt dette.

SELinux

Det er i utgangspunktet tenkt at alle RHEL 6-maskiner kjører SELinux. Dette (SELinux) er også noe som følger med RHEL. Det var derfor forventet at PostgreSQL - levert av RedHat - skulle kjøre fint under RHCS selv om SELinux var aktivert. Dette viste seg å ikke stemme, basen klarte aldri å starte opp så lenge SELinux var aktivert. F.eks. alle forsøk på å aksessere filer bl.a. under */var/cluster* (der den genererte konfigurasjonen lå) eller */usit[/dbpg]* (monteringspunktet for SAN-området) ble blokkert. Det ble såpass omfattende å åpne opp for det som trengs (spesielt da det kun ble logget en om gangen) at testen til slutt ble kjørt med SELinux deaktivert.

Dette er selvsagt noe som må vurderes før eventuell produksjonssetting.