

CSC:n laiteympäristön kehityssuunnitelma (investointi yhteensä 1 510 000 €)

CSC:n laiteympäristön kehittymistä ohjaa kokonaisarkkitehtuurityö, jossa huomioidaan julkisen hallinnon kokonaisarkkitehtuurityön periaatteet ja palveluihin liittyvät viitearkkitehtuurit. Sopimuksen kohdissa 3.1–3.4 esitetyt palvelukokonaisuudet on rakennettu tässä kuvattujen järjestelmäkokonaisuuksien päälle. Alla kuvatut tekniset arkkitehtuurit ja niiden kuvaukset ovat jatkuvan kehityksen kohteena. Laiteympäristöjä kehitetään KTPO:n ja asiakasorganisaatioiden tarpeista lähtien.

Tässä liitteessä esitettävät kuvaukset tiivistävät laiteympäristöjen tilanteen ja kehittämissuunnitelmat vuonna 2015.

1. Tutkimuksen tietoverkko Funet

Palvelun kuvaus

Funet-verkko tarjoaa tietoliikenneyhteydet suomalaisille korkeakouluille ja tutkimuslaitoksille (Funet-jäsenet). Funet tarjoaa jäsenilleen luotettavat, ruuhkattomat ja huippunopeat yhteydet muihin Funet-verkkoon kytkeytyviin organisaatioihin, kansainvälisiin tutkimusverkkoihin ja yleiseen Internetiin. Yhteyksien lisäksi Funet-verkossa on käytettävissä kaikki verkon tärkeimmät peruspalvelut kuten nimipalvelu ja aikapalvelut.

Funet tarjoaa myös lisäpalveluina myytäviä valopolkupalveluita, joiden avulla Funet-jäsenorganisaatiot pystyvät tekemään suuritehoista tiedonsiirtoa ja yhdistämään kampuksiaan ja toimipisteitään. Funet tarjoaa luotettavat yhteydet myös kaikille CSC:n tarjoamille palveluille ja CSC:n Kajaanin datakeskukselle.

Funet-verkon nykyinen tekninen toteutus

Funet-runkoverkon perustana on valokuituverkko, joka yhdistää merkittävimmät korkeakoulupaikkakunnat. Valokuitukaapeleissa tehtävä optinen tiedonsiirto tehdään Funetin optisen DWDM-siirtojärjestelmän avulla.

DWDM-siirtojärjestelmä tekee mahdolliseksi useiden rinnakkaisten suurteho-yhteyksien toteuttamisen yhteen valokuitupariin. DWDM-siirtojärjestelmää tarvitaan, jotta valokuitukaapeleiden välityksellä pystytään tekemään paikkakuntien välistä tiedonsiirtoa virheettää ja tehokkaasti.

Funetin DWDM-siirtojärjestelmäkokonaisuus muodostaa rengasmaisen DWDM-siirtojärjestelmäverkon, joka toimii perustana koko Funet-verkon sisäiselle tiedonsiirrolle Suomessa. DWDM-siirtojärjestelmää käytetään Funet-reititinverkon runkoyhteyksien ja muiden suuritehoisten yhteyksien toteuttamiseen (valopolkupalvelut).

Funet-reititinverkon avulla toteutetaan Funet-verkon sisäiset IP-yhteydet, yhteydet kansainvälisiin tutkimusverkkoihin ja yleiseen Internetiin. Reititinverkko koostuu suurimmilla korkeakoulupaikkakunnilla sijaitsevista IP-reititinlaitteista, joihin kaikki Funet-jäsenorganisaatiot kytkeytyvät joko paikkakuntien sisäisten valokuituyhteyksien tai Funetin DWDM-siirtojärjestelmäverkon välityksellä. Reititinlaitteet on kytketty toisiinsa suurikapasiteettisilla runkoyhteyksillä, jotka on toteutettu Funetin DWDM-siirtojärjestelmän avulla. Reititinverkko kytkeytyy

kansainvälisiin tutkimusverkkoihin ja yleiseen Internetiin NORDUnetin ja suomalaisten yhdysliikennepisteiden (FICIX, TREX) välityksellä. Funetin reititinverkko muodostaa suurikapasiteettisen ja vikasietoisen verkon, jonka välityksellä kaikki päivittäinen tutkimus- ja opetuskäyttö sekä Internetin käyttö on mahdollista.

Funet-verkko vaatii toimiakseen muutamia välttämättömiä peruspalveluita, kuten nimipalvelun (DNS) ja aikapalvelut. Funet-verkon peruspalvelut tarjotaan maksutta myös Funet-jäsenorganisaatioiden käyttöön.

Nimipalvelun avulla muunnetaan verkossa käytettävät nimet IP-osoitteiksi ja päinvastoin. Nimipalvelun toiminta on välttämätöntä verkon palveluiden toimintaa varten. Aikapalveluiden avulla vastaavasti huolehditaan siitä, että verkon laitteet ja palvelimet käyttävät toiminnassaan samaa, yhtenäistä kellonaikaa. Yhtenäinen kellonaika on välttämättömyys esimerkiksi vikatilanteiden selvittämistä varten.

Funet-verkon päivittäinen operointi ja valvonta vaatii lisäksi erilaisia hallinta-, valvonta- ja raportointijärjestelmiä. Näiden järjestelmien avulla hallitaan koko Funet-verkon useille paikkakunnille sijoittuvaa järjestelmäkokonaisuutta. Järjestelmien avulla valvotaan verkon ja laitteiden tilaa ja havaitaan vikatilanteet. Raportointijärjestelmien avulla tuotetaan muun muassa Funet-verkon käyttötilastoja.

Tavoitetilan mukaiset investoinnit

Jotta nykyinen Funet-verkko ja sen palvelut pystyvät täyttämään tavoitetilan mukaiset odotukset, Funet-verkkoon tehdään laitehankintoja vuosina 2015–2016 seuraavan suunnitelman mukaisesti (suluissa olevat summat kohdistuvat vuodelle 2015):

Runkoverkon kapasiteetin kasvattaminen kasvavien liikennemäärien vuoksi (800 000 €).

Funet-verkossa tehtävän tiedonsiirron määrä kasvaa tasaisesti, noin 30–40 % vuosittain. Jotta liikennemäärien kasvu ei aiheuta palvelunlaadun heikentymistä tai verkon ruuhkaantumista, Funet-verkon kapasiteettia on kasvatettava ajoissa, ennen kuin liikennemäärät ylittävät verkon nykyisen välityskyvyn.

Arvioimme, että suurimpien kaupunkien väliset liikennemäärät ovat ylittämässä runkoverkon yhteyksien nykyisen kapasiteetin vuosien 2015 ja 2016 aikana. Näin ollen Funet-verkossa kasvatetaan runkoverkon yhteyksien kapasiteettia vaiheittain lähivuosien aikana, jotta verkko ei ruuhkautuisi. Kapasiteetin kasvattamisen ohessa huomioidaan myös verkon toimintavarmuuden ylläpito ja kehittäminen.

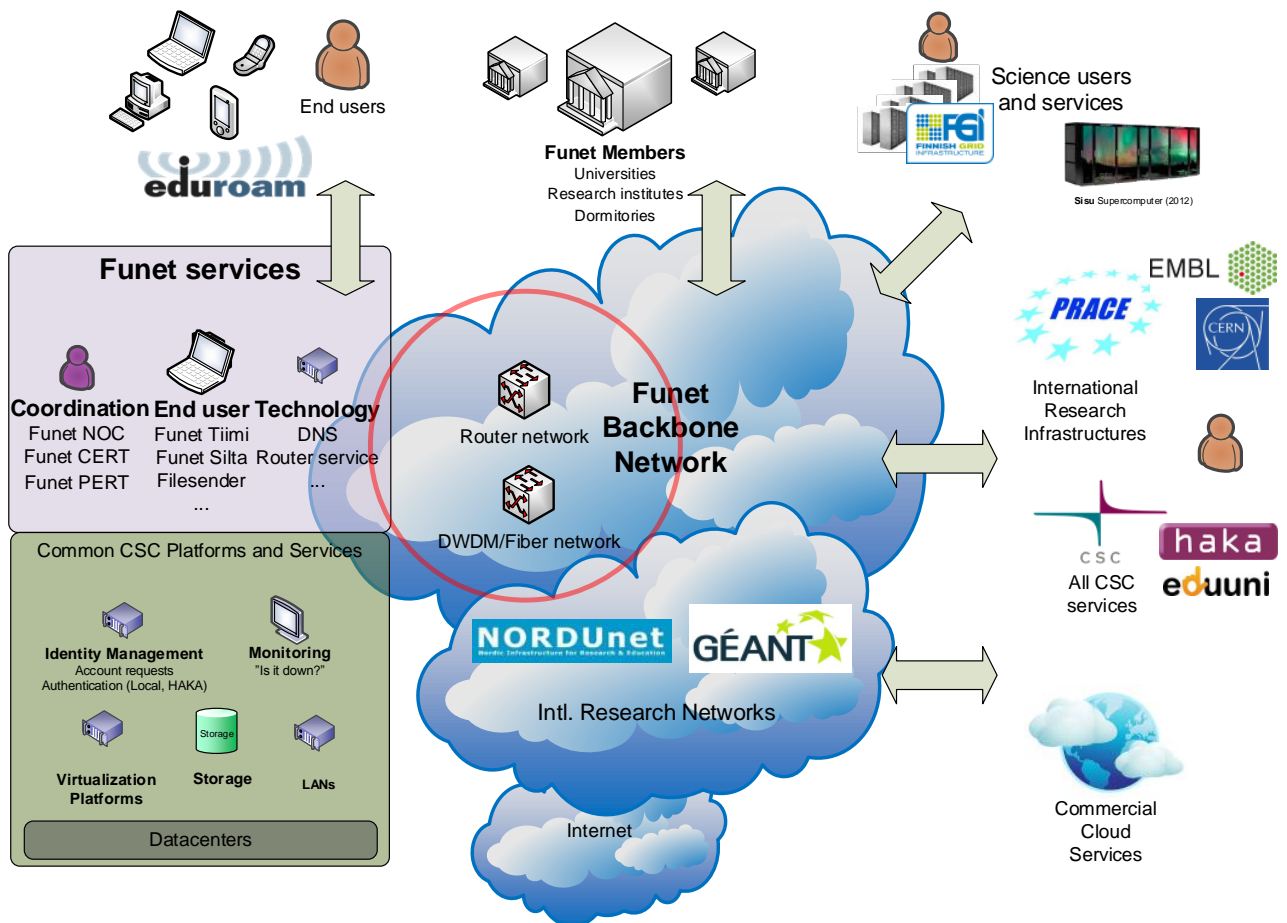
Verkon kapasiteetin kasvattaminen tapahtuu hankkimalla Funet-verkon nykyisiin reititinlaitteisiin ja optisiin DWDM-siirtojärjestelmiin uusia tai tehokkaampia laitteistokomponentteja niihin osiin verkkoa, joissa lisäkapasiteettia tarvitaan.

Luonteeltaan verkon kapasiteetin päivittäminen on kasvaneeseen palvelutarpeeseen ja toimintaympäristön muutokseen (mm. pilvipalveluiden käytön lisääntyminen, yhteispohjoismaiset palveluhankinnat) vastaamista. Kasvanutta kapasiteettia voidaan hyödyntää myös aiempaa joustavampien valopolku-palveluiden toteutuksessa.

Edellä kuvatut investoinnit ovat osa Funet-verkon vuosittaista päivityssykliä, jolla ylläpidetään verkon nykyistä palvelutasoa ja taataan verkon ruuhkattomuus liikennemäärien noustessa. Investointien lykkääminen myöhemmäksi näkyisi Funet-verkkoon liittyville organisaatioille verkon heikentyneenä palvelutasona (ruuhkautuminen) ja toimintavarmuuden heikentymisenä. Myös tulevana vuosina Funet-verkkoon kohdistuvat investointitarpeet kasvaisivat suuremmaksi.

Investoinnit on suunniteltu kirjoitushetkellä tiedossa olevan parhaan tarvearvion mukaisesti. Laitehankintojen tarkempi kohdistaminen verkossa maantieteellisesti tehdään hankintahetkellä tiedossa olevan kapasiteetti- ja palvelutarpeen mukaisesti. Laitehankinnat tehdään aina suunnitelmallisesti niin, että hankinnassa huomioidaan koko Funet-verkon toimintakyky ja kapasiteettitarve, jotta varmistetaan verkon ruuhkattomuus ja toimintavarmuus myös jatkossa.

Oheisessa kaaviokuvassa kerrotaan mihin osiin Funet-verkon arkkitehtuuria tehtävät investoinnit kohdistuvat. Investoinnin kohteena oleva osa arkkitehtuurista on korostettu kuvassa olevalla punaisella ympyrällä.



Kuva 1. Funet-verkon arkkitehtuuri ja investoinnin kohde (merkitty punaisella ympyrällä).

2. Virtualisointipalvelu

Palvelun kuvaus

CSC:n virtualisointipalvelu tarjoaa palvelinisännöintiä jaetulta alustalta tavoitteenaan korkea saatavuus, skaalautuvuus, tietoturva ja kustannustehokkuus. Palvelu on tarkoitettu pääasiassa KTPO-palveluille ja toimitetaan yleensä osana suurempaa CSC:n tuottamaa palvelukokonaisuutta. Palvelu on tarkoitettu Funet-verkon yli tapahtuvaan tuotantokäyttöön. Tarvittaessa palvelun maantieteellinen kahdennus Espoo–Kajaani-välillä on mahdollista tuottaen mahdollisimman katkeamattoman saatavuuden. CSC:n tuottamat muut palvelut ovat liitettävissä mahdollisimman saumattomasti virtualisointiympäristössä tuotettuihin palveluihin. Näitä ovat mm. HAKA-autentikointi, arkistointi/pitkäaikaissäilytys, Funet-lisäpalvelut sekä eritasoiset levypalvelut.

Palvelun kehityssuunnitelma

Korkea saatavuus, skaalautuvuus, tietoturvallisuus ja kustannustehokkuus ovat palvelun tavoitteita. Vaikka palvelu on kypsyyssasteeltaan suhteellisen kehittynyt, se on kriittisyytensä johdosta jatkuvan kehityksen kohteena. Kehityssuunnitelmassa on virtuaalipalvelinten replikointi eri datakeskusten välillä, itsepalvelutoiminnallisuuden kartoittaminen, asiakaspalvelinten haavoittuvuuksien automaattinen havainnointi sekä palvelun sovittaminen korotetun tietoturvatason saavuttamiseksi. Jatkovaa seurantaa ja kehitystä tehdään myös kustannustehokkuuden lisäämiseksi mm. turhien resurssivarausten havaitseminen ja poisto sekä palvelun muu optimointi.

Laitearkkitehtuuri

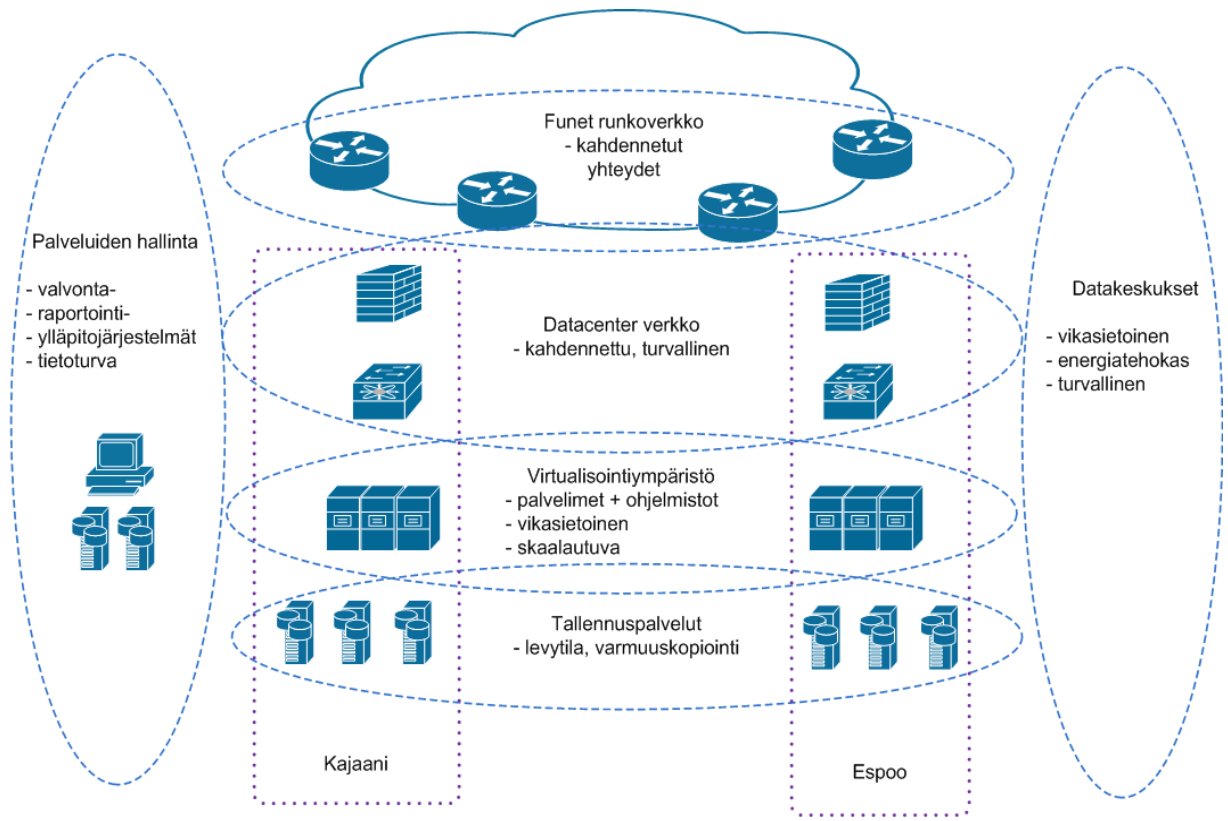
Virtualisointipalvelu muodostuu 18 palvelimesta, jotka ovat hajautettu Espooseen ja Kajaaniin.

Virtualisointikerros on toteutettu VMWare hypervisor -ohjelmistolla. Hallintapalvelu tuotetaan erillisellä palvelimella VMWare vCenter -ohjelmistolla. Käyttöastetta ja kuormitusta valvotaan jatkuvasti.

Levy-, varmuuskopiointi-, verkko- ja konosalipalveluina käytetään CSC:n omia tuotantopalveluita. Valvonnassa, geneerisissä ylläpitojärjestelmissä ja rekistereissä tukeudutaan CSC:n keskitettyyn palveluiden hallintaan.

Palvelun käyttö

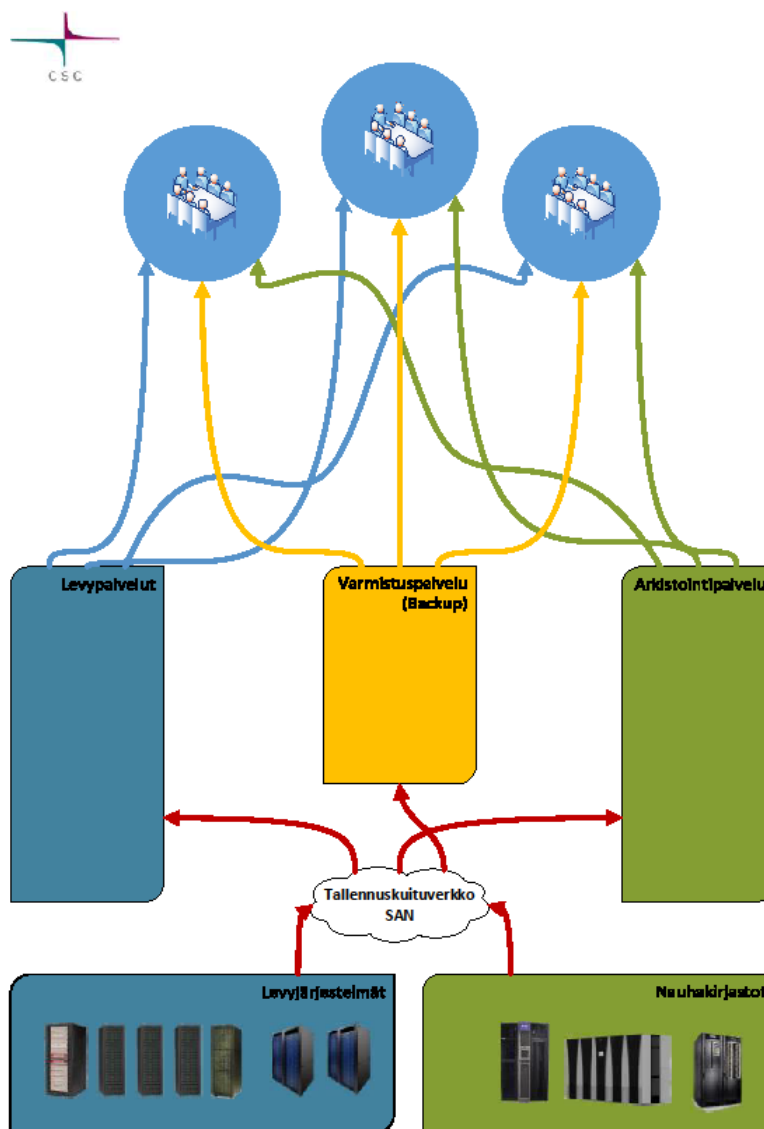
Ympäristön dynaamisuuden ja skaalautuvuuden ansiosta palvelun käyttöaste on helppo pitää korkeana ja laajentaa ympäristöä vain tarpeen mukaan. Palvelun suurimpia käyttäjiä ovat mm. eri tietovarantohankkeet (KTPO DW Infra), KDK-PAS, IDA-palvelu, KATA ja tietoaaineistopalvelut.



Kuva 2. Virtualisointipalvelu, yleisen tason arkkitehtuurikuva.

3. CSC:n tallennusinfrastruktuurin yleiskuvaus

CSC:n tarjoama ja ylläpitämä tallennusinfrastruktuuri (kuva alla) sisältää useita eri tallennuspalveluja sekä ministeriön että muiden asiakkaiden datan hallintaan joko reaaliaikaiseen työskentelyyn tai pitkäaikaiseen säilytykseen. Teknisessä mielessä keskeiset tallennuspalvelut ovat levypalvelut, arkistointi ja varmistuspalvelu. Kaikkien näiden osalta joudutaan miettimään miten palvelu voidaan tuottaa optimaalisesti kunkin asiakkaan tarpeita silmällä pitäen minimoimalla päällekkäiset investoinnit ja hyödyntämällä CSC:n asiantuntija-resurssit ja osaaminen maksimaalisesti. Asiakkailla on yksilölliset tarpeensa niin suorituskyvyn, kapasiteetin kuin palvelun luotettavuuden suhteen. Alla kuvataan lyhyesti palvelujen nykytila ja CSC:n keskeiset periaatteet niiden hallinnoinnissa.



Kuva 3. CSC:n tallennusympäristön yleiskuvaus.

4. Levypalvelut

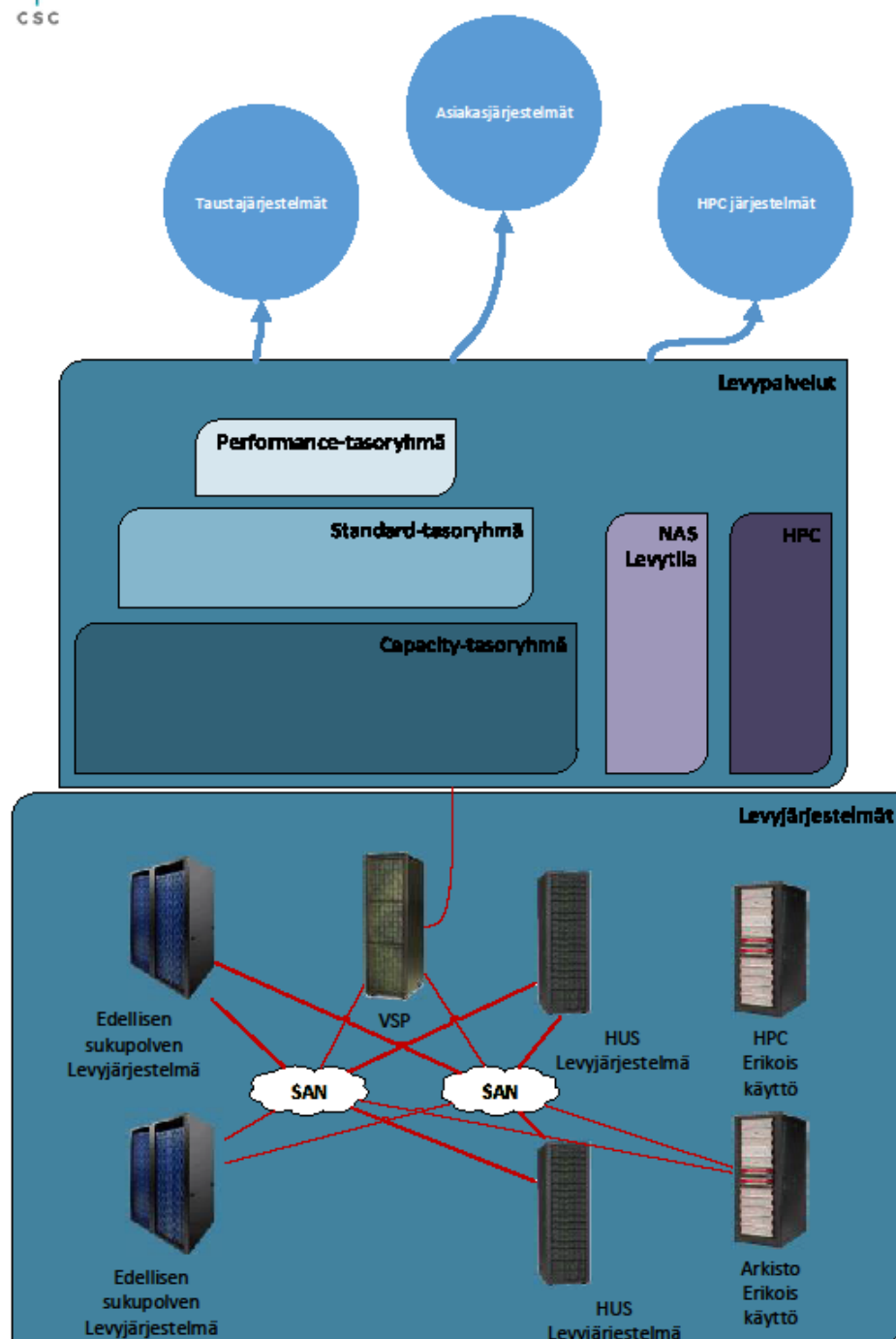
Levypalveluilla tarjotaan tietojärjestelmille työtilaa datan reaaliaikaiseen käsittelyyn, lyhytaikaiseen ja keskipitkään varastointiin. Levypalvelujen suunnittelussa, kehittämisessä ja toteutuksessa huomioidaan mm. seuraavat asiat:

- Luotettavuuden pitää olla sopimuksen mukainen, virhemarginaali on hyvin pieni ja vikatilanteissa vaikutukset syvällisiä ja laajoja.
- Kapasiteettia pitää pystyä lisäämään nopeasti ja dynaamisesti: levytilan loppuessa asiakkaalle tarjottava palvelu kaatuu kokonaan.
- Kapasiteetin allokointi asiakkaalle on käytännössä pysyvä, sitä ei voida helposti ottaa takaisin sitä tarvitsevalle (levyinvestoinnin yleiskäytettävyys usein huono). Tätä pyritään tehostamaan virtualisointitekniologioilla.
- Tapauskohtaisesti vaadittuun suorituskykyyn ei välttämättä päästä tarjoamalla vain tarvittava tallennuskapasiteetti sillä levyjärjestelmien suorituskyky riippuu viimekädessä pyörivien levyjen määrästä.
- Levyjärjestelmien ikääntyessä niiden luotettavuus heikkenee ja data vaarantuu. Näiden järjestelmien elinkaaren hallinta on hyvin tärkeää ja vaativaa. Uudet korvaavat järjestelmät on hankittava ja otettava tuotantoon kuukausia ennen kuin vanhan järjestelmän osalta voidaan – usein pitkäaikainenkin migraatioprosessi – aloittaa. Tästä tulee väistämättä päällekkäisiä järjestelmäkustannuksia ja usein myös haittaa asiakkaille.
- Levyjärjestelmät ovat monimutkaisia ja edellyttävät laajaa osaamista, joka on hyvin valmistaja- ja merkkikohtaista. CSC:n henkilöstö on aktiivisesti koulutettua ja pyrimme varmistamaan parhaan osaamistason.

Keskeiset käytetyt tekniikat ovat SAN (Storage Area Network) sekä NAS (Network Attached Storage). SAN-tekniikkaa käytetään lähinnä levytilan näyttämiseen esimerkiksi virtuaalialustojen palvelimille, NAS/HCP-palvelua tarjoaville laitteille, varmistuspalvelimille, Espoossa sijaitseville superlaskenta-palveluille sekä suoraan kaikille muille palvelimille, joiden tarvitsee nähdä levytila block-tasolla. SAN-storagea käyttäköseen palvelin on kytkettävä kiinni tallennusverkon kytkimeen. SAN-verkossa tarjottavassa levytilassa on kolme eri tasoryhmää. Erona ryhmillä on käytettävien RAID-ryhmien tasot ja levytyypit:

- Capacity-tasoryhmä on niitä palveluita varten, jotka tarvitsevat paljon kapasiteettia, mutta eivät niin paljon nopeutta. Capacity-ryhmän levytyyppinä on tällä hetkellä 3 TB 7.2 krpm SAS7k-levyt, joiden suojaustaso on RAID6 (8+2).
- Standard-tasoryhmä on niitä palveluita varten, jotka tarvitsevat paljon kapasiteettia, mutta myös kohtalaisesti nopeutta. Standard-tason levytyyppi on myöskin 3 TB 7.2 krpm SAS7K, mutta suojaustaso RAID10 (8+8).
- Performance-ryhmä on niitä varten, jotka tarvitsevat todella paljon nopeutta levyiltä. Performance-tasolla käytettävä levytyyppi on 900 GB 10 k SAS-levyt.

Seuraavassa kuvassa on esitetty keskeisiltä osin CSC:n levypalvelu-infrastruktuuri.



Kuva 4. Levypalvelujen keskeinen infrastruktuuri.

Levypalvelujen lyhyen aikavälin kehityssuunnitelma (2014–2015)

Vuonna 2014 alkupuolella on keskitytty SAN-levypalvelujen aggressiiviseen migraatioon vanhoista AMS-2500-järjestelmistä uudemmalle HUS-150-alustalle. Tähän mennessä noin puolet vanhasta laitteistosta on ajettu alas ja tällä on saavutettu merkittäviä kustannussäästöjä. Uuden sukupolven laitteet ovat merkittävästi tiheämpiä kapasiteetillaan, ja tämä tuo huomattavat säästöt

laitetila- ja sähkökustannuksissa. Laitteistojen elinkaaren lopussa myös huoltokustannusten suhteellinen osuus kustannuksista kasvaa verrattuna uuteen korvaavaan kapasiteettiin. Lisäksi uudet laitteet hyödyntävät paremmin VSP-virtualisointia, jonka kautta lähes kaikki kapasiteetti voidaan nyt tarjota. Loput käytössä olevat AMS-2500-laitteistot tullaan ajamaan alas vuoden 2015 kuluessa.

Osana migraatiota olemme ottamassa käyttöön uuden tuotteistetun ”premium standard” -levytason. Kyseinen taso yhdistää suurien NL-SAS-levyjen edullisen kapasiteetin ja flash-levyn suorituskyvyn. Uuden tason hinta/suorituskyky-suhteen odotetaan olevan riittävä, että sitä voidaan käyttää pääasiallisena tallennusratkaisuna SAN-ympäristössä.

2014 CSC on panostanut vahvasti pilvitallennusratkaisujen kehitykseen. Valittu tekniikka on Ceph FS, ja tällä hetkellä käytössä on merkittävän kokoinen testiympäristö. Ensimmäiset tuotantoon menevät asiakasympäristöt tulevat olemaan pilviympäristöt cPouta ja ePouta.

Levypalvelujen pitkän aikavälin kehityssuunnitelma (2015–2017):

- Levypalvelujen laajentaminen Kajaanin datakeskukseen.
- Entistä vikasietoisempien HA-levypalvelujen kehitys (Espoo-Kajaani failover).
- 2016 päättyvien DATA11-puitesopimusten korvaavien sopimusten kilpailutukset.

5. Varmistuspalvelu (Backup)

Palvelun kuvaus

Varmistuspalvelun tehtävänä on suojata tieto tuotantojärjestelmien vikatilanteiden ja asiakkaan itsensä aiheuttamien virheiden varalta. Tyypilliset tilanteet, joihin varaudutaan ovat:

- Äkillinen laiterikko. Tämä on ”disaster”-tilanne, jossa fyysisen vian takia suuri levyalue menetetään kokonaan.
- Tiedostojen tai tiedostojärjestelmän korruptio.
 - Fyysinen korruptio; laitteistovian tai ohjelmistovirheen takia tiedostoihin tulee muutoksia, ts. bitit kääntyvät ympäri. Korruptiota voi olla vaikea havaita ja se voi kehittyä pitkän ajan kuluessa.
 - Looginen korruptio; väärän toimintatavan tai ohjelmistovirheen takia tiedostoihin tulee ei-toivottuja muutoksia, esimerkiksi virheelliset käyttöoikeudet, aikaleimat tms.
- Käyttäjän virhe. Tyypillisimmillään on epähuomiossa tuhottu tiedosto, joka halutaan palauttaa.

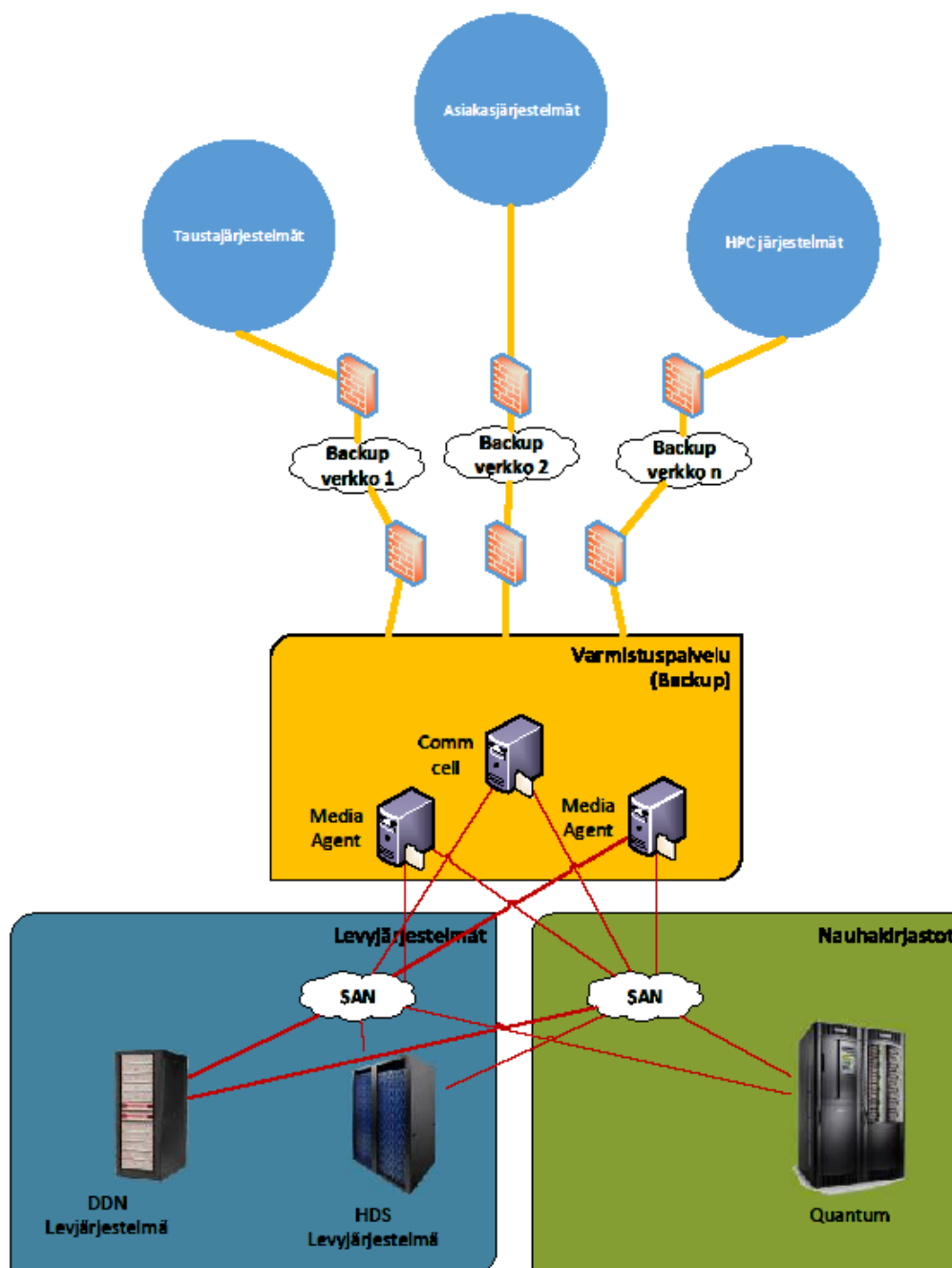
Tämänhetkisen varmistuspolitiikan mukaan CSC säästää muutokset varmistettavista levyalueista 12 viikon ajalta. Kerran viikossa otetaan täysi varmistus ja muina päivinä vain muutokset edelliseen täyteen varmistukseen. Tavanomainen palautustarve on joko täysi palautus uusimmasta varmistuksesta, tai yksittäinen vanhempi tiedosto.

Varmistuspalvelun ”kapasiteetti” on käytännössä rajaton, sillä nauhaa voidaan aina ostaa lisää, mutta ongelma on se, kuinka paljon dataa voidaan 24 tunnin kuluessa ottaa sisään ennen seuraavan varmistusjakson alkua. Mitä enemmän

varmistettavia palvelimia ja mitä tiheämmin varmistettaviin data-alueisiin tehdään muutoksia, sitä enemmän dataa joudutaan palvelimilta siirtämään tallennuspalveluun yhden vuorokauden aikana. Tätä on hyvin vaikea ennakoida, sillä käytännössä yksi tallennusasiakkuus voi "villiintyessään" aiheuttaa niin suuria datamäärän muutoksia, että muut asiakkuudet jäävät sen jalkoihin.

Varmistukset tehdään Hitachi Data Protection Suite (Simpana CommVault) – ohjelmistolla, joka nykyisellään on lisensioitu 630 varmistettavalle palvelimelle. Varmistusjärjestelmän suorituskykyä säädetään sekä varmistuspalvelimien lukumäärällä, jakamalla asiakkuuksia niiden kesken mahdollisimman optimaalisesti, että hankkimalla rinnakkaisia nauhureita robottiin datan vastaanottoa varten. Tallennuksesta nauhamedialle vastaa Quantumin 800 nauhan kirjasto, joka on varusteltu kahdeksalla lto5-nauhurilla. Nauhurien suorituskyky ilman kompressiota on noin 100 TB kirjoitusta tai lukua vuorokaudessa ja kirjaston kapasiteetti 1,2 PB ilman kompressiota. Tämänhetkisen varmistuksen koko on jo noin 80 TB (täysi varmistus) ja koko varmistuskierto käyttää nauhakirjaston tilaa yli 12-kertaisen määrän. Varmistusohjelmisto tekee varmistukset ensin levyille. Tämän lisäksi ohjelmisto käyttää hyväkseen dedupliointia; varmistukset pakataan jo varmistettavilla palvelimilla, tämä vähentää varmistuksista aiheutuvaa verkon kuormitusta ja tarvittavan välivaraston käytettävän levyn määrää.

Arvio suorituskyvyn riittävydestä tehdään keskipitkän ajan trendianalyysillä. Satsaukset backup-palveluun tulevat jatkumaan. Tämän lisäksi joudumme tarkkaan miettimään, mitä dataa ylipäänsä voidaan varmistaa. Backup-järjestelmän ylläpito on työlästä ja siihen tarvittava teknologia kallista, mutta toipumistilanteissa (disaster recovery) sen rooli on täysin keskeinen.



Kuva 5. Tallennusympäristön yleiskuvaus.

Varmistuspalvelujen lyhyen aikavälin kehittämissuunnitelma (2014–2015)

Järjestelmän nykyinen kapasiteetti pystyy vastaamaan 2014 luonnolliseen kasvuun. 04/2014 tukisopimusta varmistusohjelmistosta jatkettiin kolmella vuodella 31.3.2017 asti, samalla lisenssimäärää kasvatettiin 150 kappaleella. CSC:llä on räätälöity tukisopimus, jonka puitteissa voidaan käyttää kaikkia varmistusohjelmiston ominaisuuksia, jotka erikseen hankittuna olisivat huomattavasti kalliimpia. Ainoa rajoittava tekijä on varmistettavien palvelimien (fyysinen tai virtuaalinen) kokonaismäärä. Tämä mahdollistaa ohjelmiston helpon

laajentamisen ja hyvin laajan tuen erityyppisille varmistustarpeille, kuten eri käyttöjärjestelmät, tietokannat, sähköpostipalvelimet jne.

Varmistuspalvelujen pitkän aikavälin kehittämissuunnitelma (2015–2017)

Levyjärjestelmiin ja palvelimiin verrattuna nauha-arkistot ovat erittäin pitkäaikaisia investointeja, arkiston tyypillinen käyttöikä on 10–15 vuotta, olettaen että sitä päivitetään elinkaaren aikana. Yleensä päivityksen yhteydessä vaihdetaan nauhurit uudempaan sukupolveen. Uusi nauhurisukupolvi pystyy tyypillisesti lukemaan yhtä tai useampaa edellistä nauhamediaa mikä helpottaa migraatiota. Lisäksi itse arkistorobotit ovat modulaarisia, joten niiden nauhakapasiteettia ja nauhurimäärää voidaan yleensä kasvattaa joustavasti.

Nykyiset nauhurit ovat LTO5-nauhureita. LTO5-nauhan natiivikapasiteetti on 1.5 TB ja suorituskyky noin 140 MB/s. Kirjastomme kapasiteetin laajentaminen tulee ajankohtaiseksi 2015. LOT6-tekniikka on ollut saatavilla vuodesta 2012, ja LOT7-tekniikka tulee markkinoille arviolta 2015–2016. LTO-standardi tukee kahta aiempaa sukupolvea, eli vielä LTO7-nauhureilla voitaisiin suorittaa palautuksia LTO5-medialta.

LTO6:n ja LTO7:n suorituskyvyt ovat 2.5 TB/160 MB/s ja 4.5 TB/300 MB/s. Lisäksi LTO6-sukupolvesta lähtien tiedon pakkaustehokkuus kasvaa kahdesta 2.5-kertaiseksi. Näin ollen nykyisen kirjaston kapasiteettia voidaan kasvattaa yli nelinkertaiseksi lisäämättä nauhureiden tai nauhapaikkojen määrää. On kuitenkin huomioitava, että kapasiteetti kasvaa jatkuvasti nopeammin kuin kirjoitus- tai lukunopeus. Tämän takia on varauduttava myös välivarastona käytettävän levyn lisäykseen, jotta varmistusikkuna säilyy riittävän lyhyenä.

Varmistusjärjestelmän käyttämässä nauhakirjastossa on valmiiksi asennettua kapasiteettia joka voidaan ottaa käyttöön lisenssipäivityksellä. Näin nykyisen kirjaston hallinnoimaa mediamäärää voidaan nostaa helposti noin 50 % (800->1176). Yhdessä nauhuriteknologian päivitysten kanssa tämä mahdollistaa kokonaiskapasiteetin kasvatuksen moninkertaiseksi.

6. Arkistointipalvelu

Palvelun kuvaus

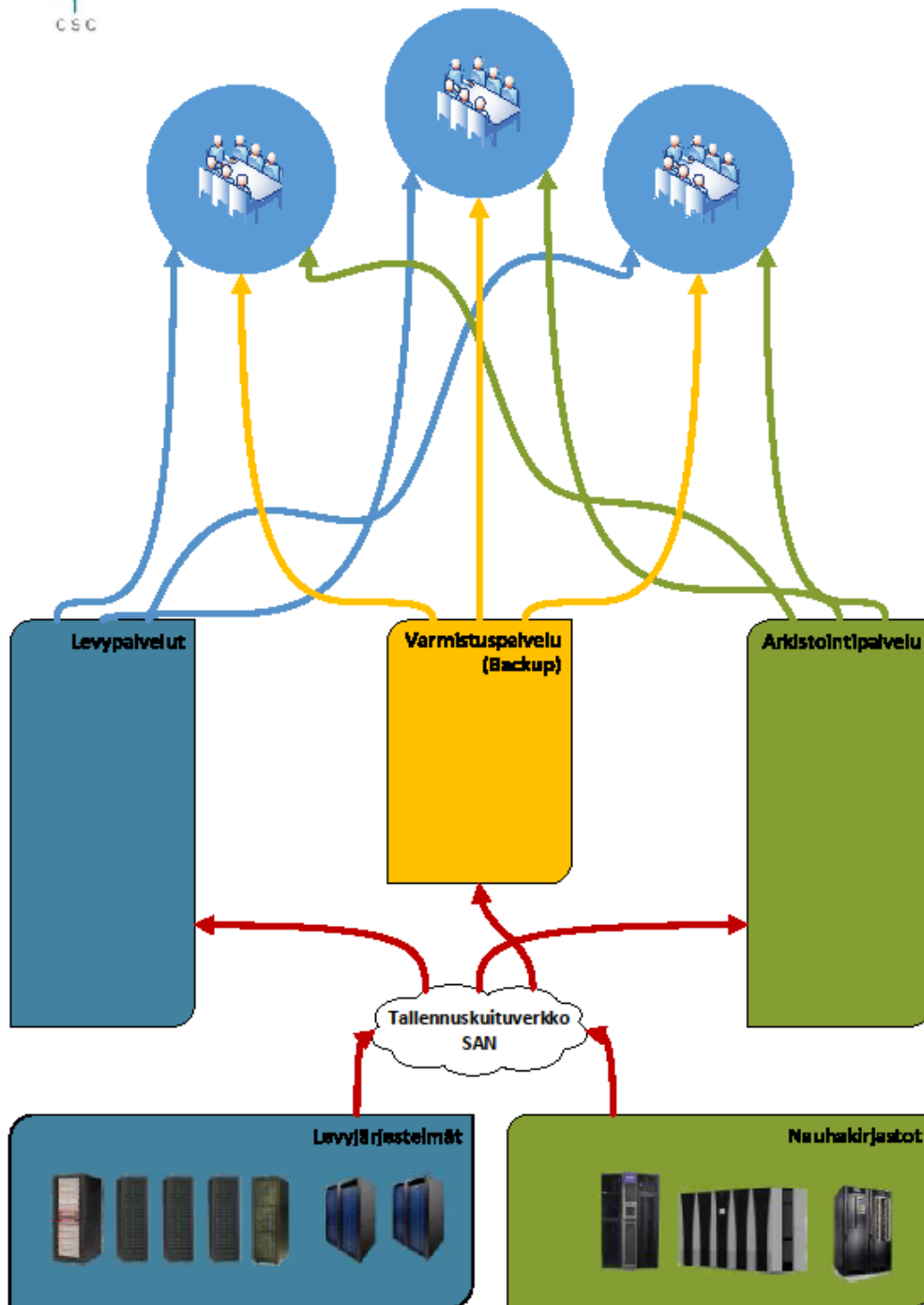
Arkistointi mahdollistaa suurten tietomäärien tallentamisen taloudellisesti. Haittana on levymuistiin verrattuna hitaampi tiedon palautus. Arkistopalvelu hyödyntää automaattista nauharobottia ja suurta levytilaa. Kapasiteettia on käytettävissä 4 petatavua ja sitä voidaan laajentaa aina useisiin satoihin petatavuihin saakka. Tällä hetkellä tallennuskapasiteetista on käytetty noin 1 petatavu.

Arkistopalvelu käyttää ohjelmistona SGI:n DMF-ohjelmistoa, joka tarjoaa automaattisen hierarkkisen tiedostopalvelun (Hierarchical Storage Management - HSM). Palvelua käytetään verkon kautta tallettamalla sinne tiedostoja ja lukemalla sieltä tiedostoja takaisin. CSC:n arkistopalvelu tukeutuu kuuteen DMF-palvelinkoneeseen, joiden taustalla toimii nauhakirjastona Spectra Logicin T-Finity viidellä kehikolla. Arkistopalvelun levyvälimuistina toimii DataDirect Networks SFA10K-X-levyjärjestelmä.

Kun tiedosto viedään arkistopalveluun, se tallennetaan ensin levyvälimuistiin, josta se aikanaan kopioidaan nauhalle. Tiedostoista tehdään oletusarvoisesti

kaksi kopiota. Haettaessa arkistojärjestelmään talletettua dataa järjestelmä siirtää sen ensin välimuistiin levyille, josta edelleen käyttäjälle. Tiedosto jää levyille, kunnes levytilaa tarvitaan jonkun muun tiedoston käsittelyyn. Levy toimii välimuistina sekä haku- että tallennusoperaatioissa. Välimuistia siivotaan siten, että vanhempien pitempään käyttämättä olleiden tiedostojen levykopio poistetaan, jolloin kapasiteettia on tarjolla usein käytettävälle datalle. Jos tiedosto on arkistoitu, eikä sitä löydy välimuistista, on haku-aika yleensä huomattavasti pidempi. Aikaa kuluu siihen, että tiedosto kopioidaan nauhakirjastosta välimuistiin. Tätä aikaa voi arvioida siten, että nauhan haku kirjastosta ja asettaminen asemaan kestää yleensä alle minuutin. Sen jälkeen nauha kelataan oikeaan kohtaan. Tähän kuluu aikaa muutamista sekunneista kymmeneen sekunteihin. Tämän jälkeen tiedosto kopioidaan verkon yli.

Arkistointipalvelu on edelleen tallennusteknologian kehityksestä huolimatta ylivoimaisesti energiatehokkain tapa tallentaa suuria ja vähän käytettyjä datamääriä pitempiä aikoja. Esimerkiksi digitoituun muotoon saatettujen analogisten audio- ja videotallenteiden tapauksessa verkkopalveluissa yleisesti käytettävissä oleva kopio voi olla useaa kertaluokkaa pienempi tallenne kuin suurella tarkkuudella digitoitu originaali, jota tarvitaan huomattavasti harvemmin kuin käyttökopiota. Arkistointiympäristön yleiskuvaus on kuvassa 6.



Kuva 6. Arkistointiympäristön yleiskuvaus.

Arkistopalvelun lyhyen aikavälin kehityssuunnitelma (2014-2015)

Arkistopalveluun ei ollut suunnitteilla merkittäviä päivityksiä vuodelle 2014. 11/2013 asennettiin kahden kehikon laajennus pitkäaikaissäilytysprojektin tarpeita varten. Vuonna 2014 suunnitteilla oli pieni laajennus DDN SFA10k

levyjärjestelmään, mikä palvelisi sekä varmistuspalvelun, että arkistopalvelun käyttäjiä. Lisäksi DMF-ohjelmiston lisensiointia oli kasvatettava vastaamaan ulkoisten asiakkaiden tarpeita.

Arkistopalvelun pitkän aikavälin kehityssuunnitelma (2015–2017)

Nykyinen järjestelmä on mitoitettu siten että pystymme toteuttamaan palvelusitoumukset tämän hetkisille asiakkaille 2017 asti. Vuosille 2015–2016 on suunniteltu arkistointi-infrastruktuurin laajennus Kajaaniin datakeskukseen, mikäli asiakastarpeet (esim. pitkäaikaissäilytys, varmistuspalvelu) kasvavat merkittävästi ja vaativat maantieteellistä hajautusta. Takarajana suunnitelma toteutukselle voidaan pitää 2016 jolloin puitesopimuksemme arkistopalvelun laitteistotoimittajien kanssa umpeutuu.

Suunnitteilla on myös laajentaa DMF-ohjelmiston käyttöä ja liittää se superlaskentaympäristön levyjärjestelmiin, siten että nyt käytössä olevat, erittäin suorituskykyiset lustre-tiedostojärjestelmät olisivat käytännössä rajattomia. Näin ollen käytettävää tilaa ei pitäisi siivota tai rajoittaa käyttäjäkohtaisesti, vanhemmat tiedostot voitaisiin siirtää automaattisesti halvemmalle medialle (nauha). Tämä lisäsi helppokäyttöisyyttä ja mahdollistaisi helpomman levyjärjestelmien migraation. DMF ohjelmiston tuki lustre-tiedostojärjestelmälle on tulossa vuoden 2014 aikana.

7. Tallennuskuituverkko (Fibre Channel, FC-verkko)

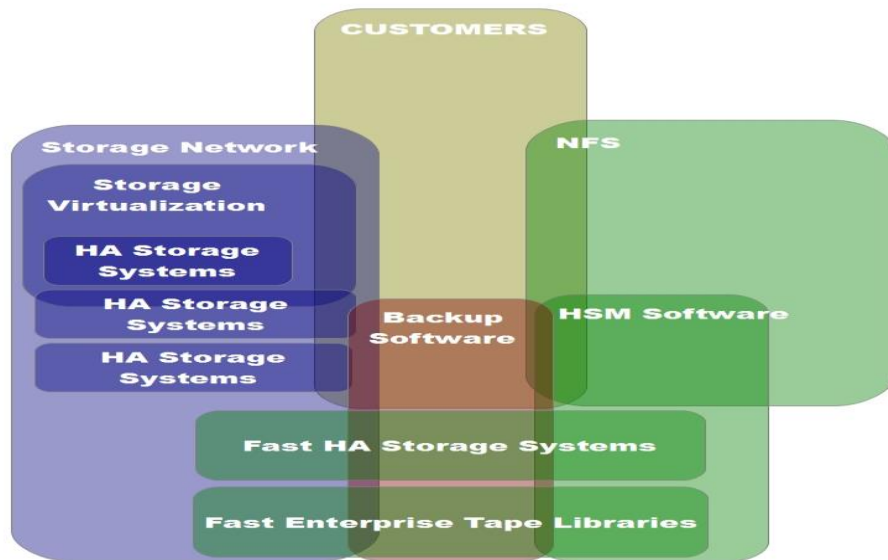
Palvelun kuvaus

Tallennuskuituverkko (FC-verkko) tuottaa palveluita kaikille tallennusalustojen toimittamille palveluille (arkisto, varmistus, levypalvelut). FC-verkon kautta ovat yhdistettynä niin levypalvelut palvelimiin kuin palvelimet tallennusjärjestelmään sekä arkistojärjestelmän sisäiset yhteydet nauhureiden välillä. Eräs suuri käyttäjä palvelulle ovat CSC:n kaikki virtuaalipalvelimet, joiden levypinta tuodaan tallennusjärjestelmästä palvelimelle FC-verkon kautta.

FC-verkko on kokonaan kahdennettu luotettavuuden takaamiseksi, etteivät esimerkiksi palvelimien levynäkymät katoa, jos toisessa verkossa tapahtuu häiriöitä. Kahdennus auttaa myös tilanteissa, joissa toisessa verkossa tarvitsee tehdä huoltoa, jolloin palvelimille ei tarvitse tehdä käyttökatkua huollon ajaksi. Kumpikin verkko on toisistaan eristetty.

Verkko käyttää Fibre Channel-teknologiaa, jolla voidaan rakentaa nopeat valokuituyhteydet palvelimien, levyjärjestelmien, nauhakirjastojen sekä nauhureiden välillä. Kytkiminä näiden kaikkien välillä toimivat Brocaden valmistamat kuitukytkimet. Brocaden Director -laitteet toimivat keskuskytkiminä, joihin on kytketty suurimmat levyjärjestelmät sekä muut kytkimet. Keskuskytkimiä on kaksi kappaletta (yksi kummassakin verkossa), molemmissa on tällä hetkellä 256 porttia. Keskuskytkimien porttimäärä on laajennettavissa aina 512 porttiin asti. Pienemmät kytkimet toimivat reunakytkiminä ja niihin kytketään pääasiassa palvelimia. Kapasiteetti kytkimen sekä palvelimen välillä on tyypillisesti uusissa kytkimissä ja palvelimissa 8Gbps. Vanhemmissa kytkimissä maksimi on 4Gbps. Reunakytkimen ja keskuskytkimen välillä kapasiteetti voi olla jopa 48Gbps. Reunakytkimiä on tällä hetkellä 13 kappaletta kummassakin verkossa, eli yhteensä 26 kytkintä, joissa on kussakin 24-48 porttia. Lisäksi arkistossa on

varmistus- ja arkistojärjestelmien käytössä kaksi erillistä 80-porttista kytkintä, jotka eivät ole kytkettynä muihin. Kytkimet on sijoiteltuna niihin palvelinhuoneisiin, missä kuituverkkoa tarvitaan. Yhteydet kytkimien ja palvelimien välillä kaapeloidaan tyypillisesti käyttäen MPO-paneeleja ja -kaapelointeja, jotka mahdollistavat suuremman tiheyden kytkinpaneeleihin kuin tavallisesti.



Kuva 7. Kaaviokuva tallennusinfrastruktuurin tulevaisuuden näkymistä ja tahtotilasta.

8. CSC:n datakeskus- ja konosalipalvelut

Palvelun kuvaus

CSC:n datakeskus ja konesalit on toteutettu kestävän kehityksen periaatteita noudattaen. Konesalit luovat edellytykset kansainvälisesti huipputasoisten tieteen, tutkimuksen, koulutun ja kulttuurin ICT-palveluille. Konesalit sijaitsevat Kajaanissa ja Espoossa.

Kajaanin datakeskus on CSC:n kustannustehokkain ja ympäristöystävällisin sali. Datakeskus on toteutettu modulaarisesti, joten se on helposti laajennettavissa. Kajaanin datakeskuksessa on käytössä ainutlaatuinen energiaa säästävä jäähdytysjärjestelmä, valtakunnan kattavin sähköjakeluverkko ja Funetin tehokkaat verkkoyhteydet. Datakeskuksen energiatehokkuuden mittari eli PUE-arvo on ilmajäähdytteisessä moduulissa 1,06 ja vesijäähdytteisessä salissa 1,28. Kajaanin datakeskus sai ISO/ IEC 27001 -sertifikaatin 2013.

Muut konesalit (3 kpl) sijaitsevat Espoossa, suurimpien korkeakoulujen sekä tutkimus- ja kulttuurilaitosten läheisyydessä. CSC tarjoaa hyvän vaihtoehdon palveluiden hajautukseen maantieteellisesti. Yksi saleista on lisäksi varustettu EMP-suojauksella, mikä tarjoaa hyvän perustan vaativimpien palveluiden

sijoituspaikaksi. Espoon datakeskukset saivat ISO/ IEC 27001 -sertifikaatin 2014.

Konesaliratkaisuja kehitetään asiakkaiden tarpeiden mukaan. Monipuolinen konesalitarjontamme mahdollistaa:

- asiakkaan itsensä ylläpitämien palveluiden sijoittamisen tiloihimme sekä virtuaalipalvelinalustoilla että asiakaskohtaisesti toteutettavilla fyysisillä palvelimilla
- konesalien vaatimustenhallinnan, suunnittelun- ja räätälöinnin asiantuntijapalvelut asiakkaan tarpeiden mukaisesti
- palvelutason joka pohjautuu JHS174 suosituksiin

Infrastruktuuripalvelut kattavat konesalipalvelut:

- Asiakkaan palvelinten ja oheislaitteiden sijoitussuunnittelu CSC:n konesaleihin käyttötarkoituksen mukaisesti
- Asiakkaan lähetysten vastaanotto ja tarkistus, siirto konesaliin, myös laitteiston poisto
- Tuotantovalmiit palvelimet ja oheislaitteet CSC:n operoimassa konesalissa
- Varmennettu sähkönsyöttö ja jäähdytys energiatehokkailla ratkaisuilla
- Uusimmat teknologiat ja ammattitaitoinen henkilöstö/palveluntuottajat
- Konesalin toiminnan kiinteistötekniinen valvonta, konesalitekniikan automaattiset hälytykset, kulunvalvonta, vartiointi ja kameravalvonta, palonilmaisimet ja automaattiset sammutusjärjestelmät
- Korkeat turvallisuusvaateet täyttävät toimintaprosessit (ISO K27001)
- Tuotannon turvaavien prosessien jatkuva ylläpito
- Laiterekisterin ylläpito ja muutoshallinta, raportointi
- Palvelutaso pohjautuu JHS174 suosituksiin

CSC:n konesaleista asiakkaillemme tarjoamme esimerkiksi tallennuspalveluita (levyratkaisut, varmistuspalvelut ja tallennusratkaisut) luotettavan Funet-verkon kautta.

Kehityshankkeita

- Uusi modulaarinen konesali rakennetaan palvelemaan korkeakoulujen tietohallintoratkaisuja. Tämä kustannus- ja energiatehokas ratkaisu on valmis syksyllä 2014. Moduuli on 85 %:sti EAKR-rahoituksella toteutettu.

9. Laskentapalvelimet

Palveluiden kuvaus

Tieteellisen laskennan palvelut korkeakouluille tarjoavat korkean suorituskyvyn laskentapalvelua (Capability Computing) vaatimaan massiivisesti rinnakkaisiin laskenta-ajoihin ja kapasiteettilaskentapalvelua (Capacity Computing) pieniin ja keskisuuriin laskenta-ajoihin. Tarjottavat palvelut muodostavat Suomen tehokkaimman ja moderneimman akateemisen tutkimuksen käyttöön tarkoitetun laskentaympäristön. Palvelut toteutetaan erittäin energia- ja kustannus-tehokkaasti. CSC:n laskentaympäristö on viime vuosina virtualisoinnin avulla kehittynyt tarjoamaan laskentapalveluita data-intensiiviseen laskentaan tai visualisointiin. Virtualisoitu, laskentaan kelpaava levykapasiteetti integroidaan

normaalin CSC:n klusterilaskentakapasiteetin kanssa. Tämä mahdollistaa käyttäjien itse rakentamat data-analyysi työvuot ja avaa CSC:n keskitetyt laskentaresurssin paremmin uusille käyttäjäryhmille.

Laskentapalvelut tarjoavat monipuolisen käyttöliittymävalikoiman, joka vastaa eritasoisten käyttäjien tarpeisiin ja eri käyttötarkoituksiin. Vaikka laskentapalvelut teknisesti muodostuvat useista eri osakokonaisuudesta, niistä muodostuu käyttökokemukseltaan mahdollisimman yhtenäinen, helppokäyttöinen ja helposti saavutettava laskentaympäristö, joka pyritään myös säilyttämään siirryttäessä laitesukupolvesta seuraavaan. Laskennan pilvipalveluiden kautta käyttäjillä tai asiakasorganisaatioilla on mahdollisuus myös luoda täysin omiin tarpeisiin räätälöity laskentaympäristö.

Palveluiden toimintavarmuus ja turvallisuus ovat huippuluokkaa ja niiden operoinnista, kehityksestä ja käyttäjätuesta vastaavien asiantuntijoiden osaaminen edustaa alan terävintä kärkeä. Palveluiden turvallisuuden ja laadun indikaattoreina ovat kansainvälisesti hyväksytyt IT-sertifioinnit. Palveluiden kehityksessä ja laitehankinnoissa osataan vastata tehokkaasti muuttuviin asiakastarpeisiin ja huomioimaan alan teknologiakehitys. Modernit ylläpitotyökalut ja -käytännöt tarjoavat perustan ketterälle palvelukehitykselle.

Pilvipalvelut

Vuoden 2014 aikana laskentapalvelut keskitettiin onnistuneesti Kajaanin datakeskuksiin kehityssuunnitelman mukaisesti: Vuori-klusteri ja Hippupalvelimet ajettiin alas ja niiden tarjoamat erikoispalvelut siirrettiin Taitoon.

Pouta-pilvipalvelu on otettu tuotantokäyttöön ja tarjoaa laaS-palvelua jossa loppukäyttäjä toteuttaa itsenäisesti oman laskentaympäristönsä.

CSC:n pilvipalvelukehityksen pioneerina toimivan ELIXIR FIRI hankkeen pilottikäyttötapaukset vuosina 2011–2013 ovat osoittaneet että organisaatiotason pilvipalveluille on pysyvää kysyntää. Nämä ympäristöt ovat tiiviisti integroitua asiakkaiden omiin järjestelmiin. ELIXIRin Biomedinfra-pilvipalvelussa tehty työ ja pilottikäyttö ovat tukeneet CSC:n Pouta-pilvipalvelun kehitystä ratkaisevasti. Tarvetta olisi myös korotetulle tietoturvasolulle esimerkiksi tutkimusprojekteissa jossa käsitellään potilasdataa (kts. ePouta investointi alla).

Virtualisointi ja sen päälle rakennetut pilvipalvelut, kuten Pouta ovat mahdollistaneet CSC:n keskitettyjen laskenta- ja tallennusresurssien käytön uusille käyttäjäryhmille. Ne ovat myös osoittautuneet joustavaksi ja kustannustehokkaaksi tavaksi toteuttaa keskitetysti kapasiteettia, data-analyysijä ja informaation käsittelyä, joita kaikki organisaatiot tarvitsevat omaa erikoistumisistaan varten.

CSC tuotantoon viedyssä (community) cPouta-palvelussa kaikkia mm. tietoturvaan liittyviä vaatimuksia ei ole nykyisellään mahdollista täyttää. Nykyisen ELIXIR FIRI -rahoituksen avulla hankitun laitekannan elinkaari on tulossa päätökseensä 2015 aikana ja korvaavaa infrastruktuuria ei ole olemassa. Tämän vuoksi Pouta-pilvipalvelussa on havaittu että tarvitaan sekä tuotannossa olevaa (community) cPouta laaS -palvelua että organisaatiotasolle kohdennettua (enterprise) ePouta laaS -palvelua.

Laskenta-alustat

Sisu- ja Taito-superkoneiden 2. vaiheen päivitys moninkertaisti näiden tehon ja paransi energiatehokkuutta huomattavasti. PRACE-tutkimusinfrastruktuurin osarahoittama Bull-prototyyppi on tuotantokäytössä ja tarjoaa äärimmäisen energiatehokasta laskentakapasiteettia osana Taito-klusteria.

Nykyiset supertietokoneet ovat suunniteltu erityisesti laskentaintensiiville kuormille ja ne eivät ole optimaalisia tai kustannustehokkaita kaikkiin dataintensiivisiin (ns. Big Data) ongelmiin tällä hetkellä.

Monitorointia, raportointia ja käyttäjähallintaa on uudistettu joka mahdollistaa entistä paremman kapasiteetin hallinnan mutta kehitettävää on edelleen.

Teknologiaseuranta

Tiivis teknologiaseuranta ja yhteistyö laitetoimittajien kanssa on mahdollistanut uusimpien prosessoriteknologioiden varhaisen käyttöönoton (early access) ensimmäisien keskusten joukossa maailmassa.

Palveluiden kehityssuunnitelma

Pilvipalvelut

Pilvipalveluiden nimeämiskäytäntöjä selkeytetään: Nykyinen Pouta-palvelu brändätään nimellä cPouta (community Pouta) ja palvelunimi Pouta laajennetaan käsittämään kaikki CSC:n laskentaan suunnatut pilvipalvelut.

cPouta-pilvipalveluita kehitetään edelleen erityisesti parantamalla luotettavuutta ja lisäämällä ominaisuuksia käyttäjäpalautteen perusteella.

cPoutaa täydentämään rakennetaan ePouta (enterprise Pouta), joka on suunniteltu erityisesti käyttäjille, joilla on pitkäaikainen tarve räätälöidylle laskentaympäristölle tai joilla on poikkeuksellisen korkeat tietoturva-vaatimukset. Ensimmäiset käyttäjät uudelle palvelulle saadaan suoraan ELIXIR FIRI -hankkeen pilottikäyttäjistä.

Korkean tason integraatiotavoite on että ePouta ja cPouta ovat jatkossa samalla tavalla osa CSC:n keskitettyä kapasiteettia kuin perinteiset supertietokoneet ja -klusterit. Ajurina toimivat suurdatan (big data, esim. suomalainen referenssigenomi terveyden sovelluksia varten <http://www.sisuproject.fi>) hankkeet ja niiden tarpeet ml. oikein mitoitettu tietoturvaso.

Laskenta-alustat

Niin supertietokoneiden kuin pilvipalveluiden operoinnissa yhtenäistetään ja modernisoidaan ylläpitokäytäntöjä sekä -työkaluja. Lisäksi laskenta-alustojen testiympäristöjä kehitetään. Tämä parantaa mm. palveluiden luotettavuutta, ylläpitotyön tehokkuutta ja mahdollistaa uusien palvelualueiden toteuttamisen entistä ketterämmin ja kustannustehokkaammin.

Teknologiaseuranta

Laskentapalveluihin liittyvien teknologioiden kehitys on erittäin nopeaa ja usein hankalasti ennakoitavaa jonka vuoksi suoritetaan jatkuvaa teknologiaseurantaa.

Vuonna 2015 teknologiaseuranta keskittyy erityisesti dataintensiivisen laskentaan, pilvipalveluihin ja seuraavan superkonehankinnan kannalta oleellisimpiin teknologioihin. Mahdollisuuksien mukaan näitä teknologioita testataan myös käytännössä pienillä prototyypeillä.

Big data

Suuret tai muuten vaativat tietoaaineistot, Big data, asettavat tulevaisuudessa aivan uudenlaisia tarpeita tietotekniselle infrastruktuurille. Tieteen piiristä nousee jatkuvasti uusia mittalaitetekniikoita, jotka kykenevät tuottamaan ennen-näkemättömiä määriä dataa, muun muassa ihmisen perimästä, alkeishiukkasten ominaisuuksista tai ilmakehän tilasta. Datamäärien kasvaminen ei kuitenkaan rajoitu vain luonnontieteiden piiriin, vaan myös aivan uudet tieteenalat siirtyvät laskennallisten menetelmien käyttöön, suurten tietoaaineistojen avaamisen mahdollisuuksien johdosta. Esimerkiksi fonetiikan tutkimuksessa automaattinen äänneiden tunnistus mahdollistaa huomattavasti suurempien aineistojen eli puhekorpuksen rakentamisen. Samoin kuva- ja videopohjaisten aineistojen analyysitarpeiden nähdään kasvavan tulevaisuudessa.

CSC pyrkii vastaamaan näihin nouseviin tarpeisiin investoimalla dataintensiivisen laskennan kapasiteettiin. Kansallisena merkittävänä toimijana CSC:llä on erinomainen mahdollisuus rakentaa tarvittavia palveluja olemassa olevaa infrastruktuuria ja palvelukonsepteja hyödyntäen, näin mahdollistaen investoidun kapasiteetin nopean ja tehokkaan hyödyntämisen. CSC:n rooli Big data -infrastruktuurin kehittäjänä, sekä tärkeiden kokeilu- ja testausympäristöjen tarjoajana, on tunnistettu myös Liikenne- ja viestintäministeriön vastikään julkaisemassa Big datan hyödyntäminen -raportissa.

Järjestelmien uusiminen

Taito-klusterin ensimmäisen vaiheen kesäkuussa 2012 asennettu osio tulee käyttöikänsä päähän vuoden 2016 puolellavälissä ja vuonna 2014 asennetut Sisu- ja Taito-järjestelmien toisen vaiheen osiot vuoden 2017 aikana. Näiden järjestelmien uudistamiseen tähtäävä projekti on aloitettava viimeistään vuoden 2015 alussa. Hankintaprojektissa tulisi ottaa huomioon entistä paremmin palveluiden jatkuvuus ja elinkaarikustannukset siirryttäessä uuteen laitesukupolveen. Tässä liitteessä kuvataan laskentapalveluiden osalta ainoastaan vuoden 2015 investointitarpeet, joita ei pystytty täysimääräisesti sisällyttämään edelliseen superkone-rahoitukseen.

Tietoarkkitehtuuri

Tietovarastojen tavoitetila

Kaikilla laskentapalveluilla on yhteiset levyalueet, jotka mahdollistavat eri palveluiden saumattoman yhteiskäytön:

- Varmistettu kotihakemisto
- Työlevyalueita eri tarpeisiin (rinnakkaislaskenta, kustannustehokas levyalue pilvipalveluille, erittäin data-intensiivinen laskenta)

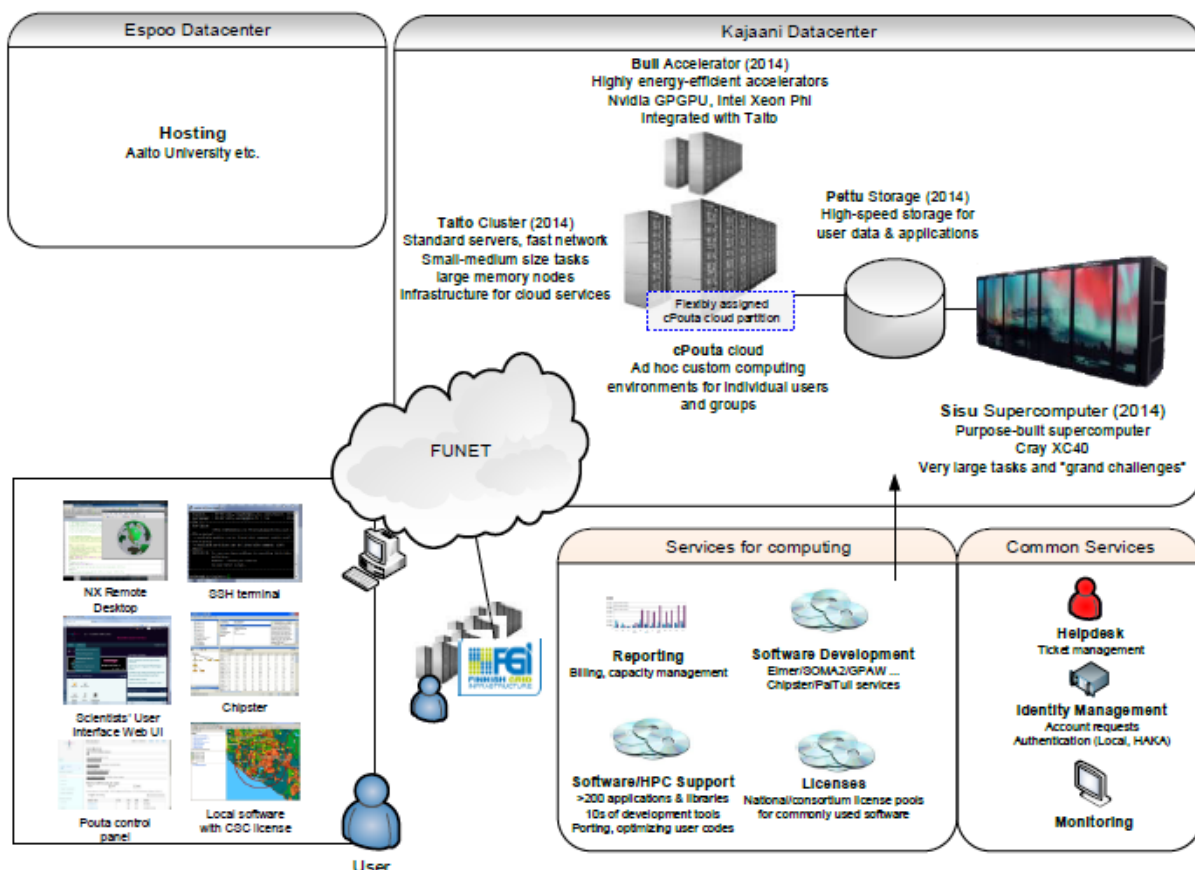
Käyttötilastoja kerätään ja analysoidaan säännöllisesti mm. tulevaisuuden laitehankintojen suunnitteluun. Myös tietoa käyttäjän ajojen energiankulutuksesta ja resurssien käytön tehokkuudesta kerätään ja raportoidaan. Käyttäjän on helppo tarkastella tätä dataa tutkijan käyttöliittymän kautta.

Laitearkkitehtuuri

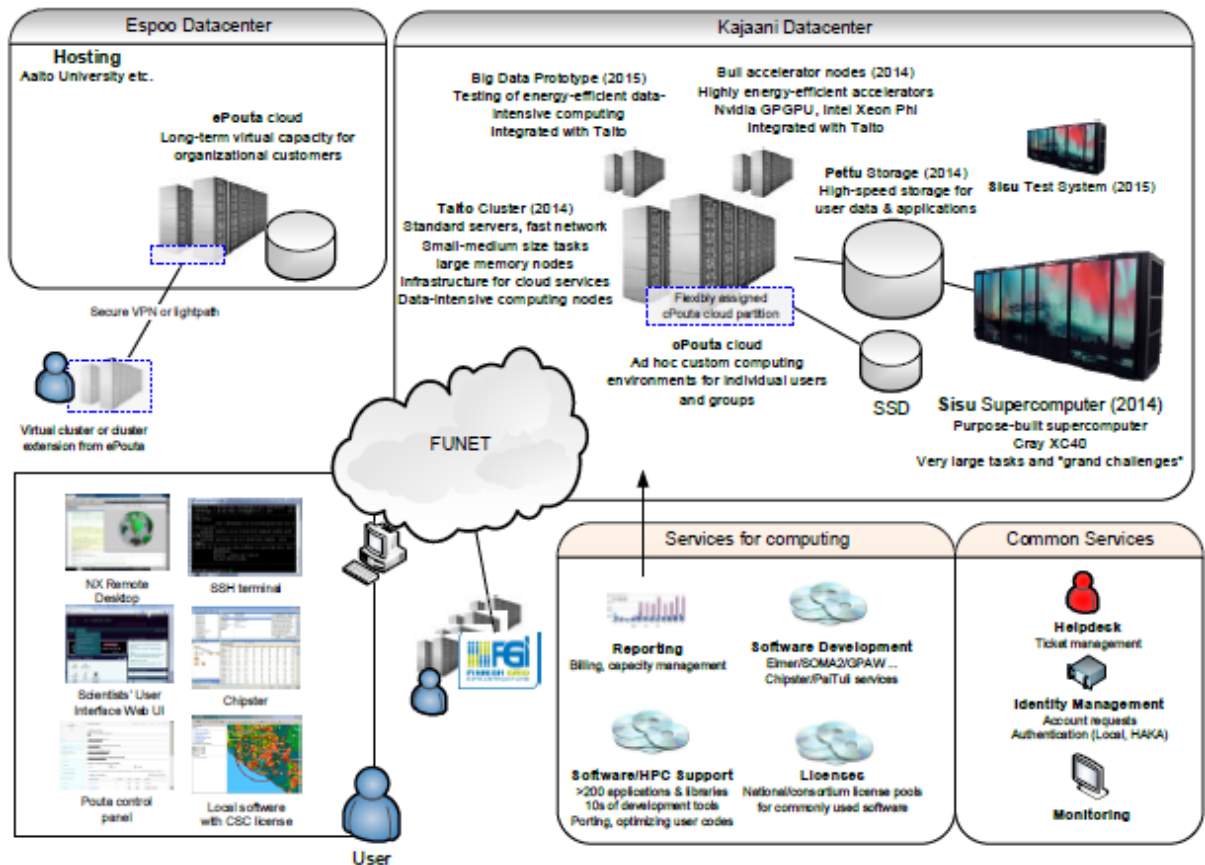
Laskentapalvelut muodostuvat nykyään seuraavista palvelimista:

- Korkean suorituskyvyn laskenta: Kajaanin datakeskuksessa sijaitseva Sisu-supertietokone (Cray XC30)
- Kapasiteetilaskenta: Kajaanin datakeskuksessa sijaitseva Taito
- Pilvilaskenta: Pouta-pilvipalvelu jonka kapasiteetti allokoidaan Taito-klusterista. Kapasiteettia voidaan skaalata helposti.
- Laskentaan käytettävä kustannustehokas tallennuskapasiteetti, jota voidaan joustavasti liittää laskentaresursseihin (ceph, object storage, cloud nfs).
- Erikoislaskenta: Taito-klusteriin integroitu erittäin energiatehokas suoranestejähdytetty Bull-klusteri, jossa on uusimman sukupolven Nvidia Tesla ja Intel Xeon Phi erikoisprosessoreita. Taito-klusterissa on myös kaksi suurimuistista (1,5TB) palvelinta

Nämä palvelimet yhdessä data-alueineen muodostavat erillisiä saarekkeita niin käyttäjä- kuin ylläpidollisesta näkökulmasta (8).



Kuva 8. Arkkitehtuurin nykytila vuoden 2014 lopussa.



Kuva 9. Arkkitehtuurin tavoitetila 2015 lopussa.

Tavoitetilan mukaiset investoinnit (sopimuksen kohdassa 3.4)

Kehityssuunnitelman toteuttamiseksi tarvitaan vuoden 2015 aikana erilaisia investointeja tukemaan niin dataintensiivisen laskennan, pilvipalveluiden kuin ylläpitokäytäntöjen kehittämiseksi. Investoinnit pyritään toteuttamaan siten, että ne integroituvat mahdollisimman tehokkaasti osaksi CSC:n olemassa olevia palvelualueita. Tällöin minimoidaan uusien investointien ylläpitoon kuluva henkilötyö ja tuetaan yhtenäisen palvelukokonaisuuden muodostumista.

a) Dataintensiivisen laskennan pilvikapasiteetti

Hankitaan erillistä dataintensiivisen laskennan kapasiteettia, joka tarjotaan nykyisen pilvipalvelun kautta. Osa kapasiteetista voidaan tarjota myös HPC-klusterin osana. Kapasiteetin painotukset poikkeavat merkittävästi nykyisestä HPC-käyttöön optimoidusta ympäristöstä. Sen avulla voidaan toteuttaa entistä suurempia ja vaativampia dataintensiivisen laskennan tehtäviä, niin että siihen käytetään tarkoituksenmukaista ja siksi huomattavasti kustannustehokkaampaa laiteympäristöä.

Investoinnin suuruus: 250 000 euroa

b) Epäsäännöllisen dataintensiivisen laskennan pilvikapasiteetti

Hankitaan erityiskapasiteettia kaikista vaativimpiin satunnaisen/epäsäännöllisen IO:n laskentatehtäviin johon perinteiset kiintolevyt soveltuvat huonosti. Pahimmillaan tällaiset työt ovat aiheuttaneet koko järjestelmän laajuisia

suorituskykyongelmia jos niitä on suoritettu tavallisella rinnakkaislevyjärjestelmällä. Todennäköinen toteutusteknologia ovat korkean suorituskyvyn SSD-massamuistit, kytkettynä riittävän tehokkaiisiin klusterin laskentasoelmuihin. Kapasiteetti voidaan tarjota saumattomasti osana HPC-klusteria.

Investoinnin suuruus: 50 000 euroa

c) Energiatehokkaan dataintensiivisen laskennan prototyyppi

Hankitaan energiatehokkaan dataintensiivisen laskennan mahdollistavaa prosessorikapasiteettia. Todennäköinen toteutus on ARM-prosessoriteknologialla. Prototyypin avulla voidaan perehtyä kapasiteetin soveltuvuuteen laajemmassa tuotantokäytössä. Evaluointi ja kokeilukäyttö toteutetaan yhteistyössä yliopistokumppaneiden kanssa.

Investoinnin suuruus: 20 000 euroa

d) Sisu-testijärjestelmä

Hankitaan Sisu-supertietokoneelle testijärjestelmä (Testing and Development System, TDS). Päivitysten ja muutosten kokeilu ennen tuotantoon ottamista pienentää muutosten aiheuttamaa riskiä huomattavasti ja mahdollistaa uusien ominaisuuksien nopeamman käyttöönoton. CSC:n omat asiantuntijat voivat ajaa myös osan testiajoistaan laitteistolla, vapauttaen Sisun tuotantojärjestelmän kapasiteettia. Osa investoinnista voidaan kattaa Crayn Sisun 1. vaiheen asennuksen viivästyisestä saatuja kompensatioita, joiden arvo on n. 102 000 euroa. Pienimmän tarkoituksenmukaisen TDS-konfiguraation kokonaishinta on n. 190 000 euroa.

Investoinnin suuruus: 88 000 euroa

e) ePouta-pilvikapasiteetti

Sensitiivisen datan käsittelyyn soveltuva pilvilaskenta- ja tallennusympäristö, joka on mahdollista kytkeä Funet valopolkuyhteyksien avulla saumattomasti asiakasorganisaatioiden kotiverkkoihin. Korotetun tietoturvasoatimuksen (Katakri III) takia järjestelmä joudutaan toteuttamaan vielä nykyteknologialla erillisenä kokonaisuutena ja se tullaan asentamaan CSC:n Espoon konesaleihin pääkaupunkiseudulle. Kajaanin datakeskuksessa ei sen nykyisellä toteutuksella voi tarjota korotetun tietoturvasoan palveluita. Tarvittavien muutosten aikataulu ja kustannukset olisivat huomattavan suuret esitettyyn laitteistoinvestointiin nähden. Asiakastarpeet sensitiivisen datan analysoinnin osalta ovat jyrkässä kasvussa. .

Erillinen korotetun tietoturvasoan pilvipalvelu on tärkeä osa CSC:n laskentapalveluiden kokonaisarkkitehtuurissa ja sitä ei ole kustannustehokasta eikä tarkoituksenmukaista rakentaa jokaisen palvelua tarvitsevan tutkijaryhmän tarpeisiin erikseen. Ilman palvelua emme voi vastata moniin biolääketieteen tutkimuspalvelu tarpeisiin, joita mm. biopankkien kokonaisarkkitehtuuri ja biotieteiden tutkimuksen kehitys (mm. henkilökohtainen lääketiede ja 2014 säädetty ELIXIR-laki) lisää.

Investoinnin suuruus: 302 000 euroa