



Grid-hankkeita ja tulevaisuuden näkymiä

Arto Teräs

arto.teras@csc.fi

Tietotekniikan professoritapaaminen CSC:llä 20.12.2004



Sisällys

- Grid pähkinänkuoressa
- CSC:n grid-hankkeita
 - Haka-luottamusverkosto
 - Materiaalitutkimuksen grid (M-grid) ja NorduGrid-yhteistyö
 - EU-projektit: DEISA, ENACTS, EMBRACE
 - European Southern Observatory -hanke (ESO)
- Tietojenkäsittelytiede ja grid-ympäristön haasteet



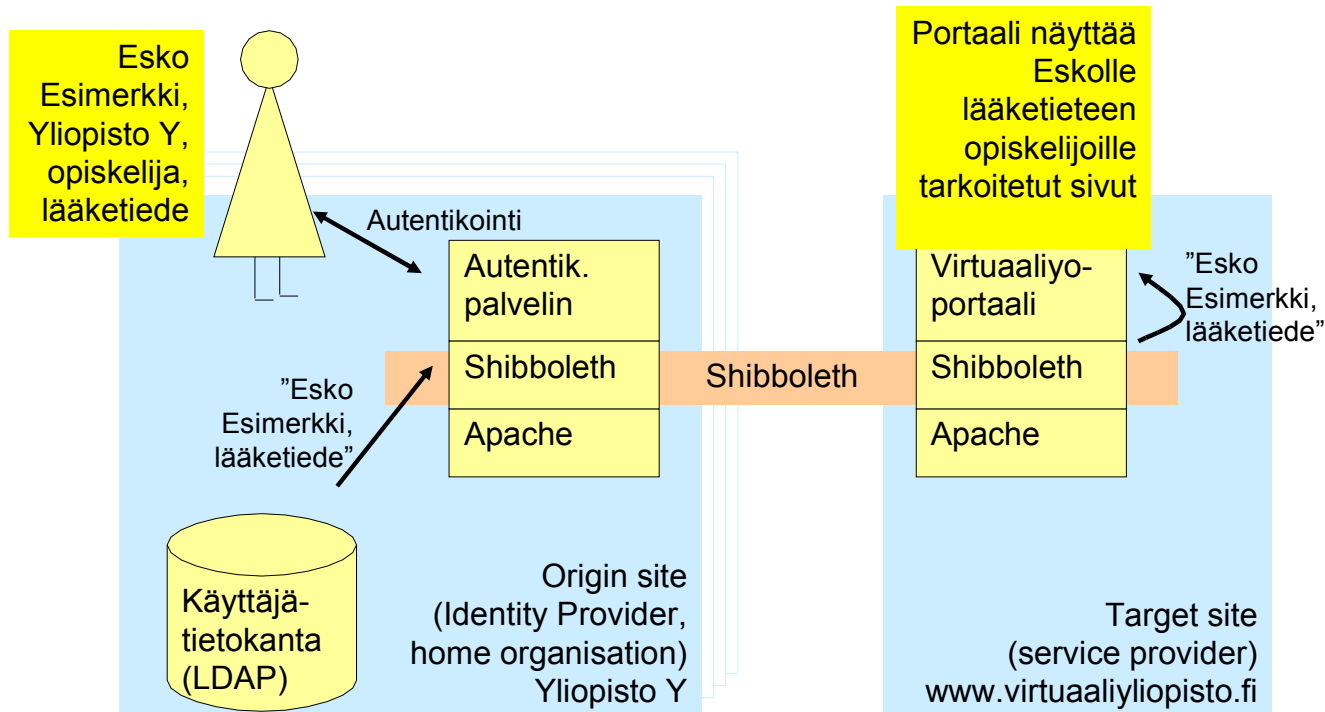
Grid pähkinäkuoressa

- Muotisana: monta käyttäjää, monta merkitystä
- Grid-tekniikoilla pyritään tarjoamaan helppo ja tietoturvallinen pääsy hajallaan sijaitseviin tieto- ja laskentaresursseihin
- Grid ylittää organisaatorajat!
 - Tietoturva- ja luottamuskysymykset
 - Aiemmin sisäisiä toimintoja (käyttäjähallinta yms.) joudutaan sovittamaan yhteen kumppanien kanssa
- Grid-yhteistyö on paljon muutakin kuin pelkkää tekniikkaa



Haka-luottamusverkosto

- Käyttäjille yksi salasana verkkopalveluihin yliopistorajojen yli, ajantasaiset tiedot suoraan kotikorkeakoulun rekistereistä





Materiaalitutkimuksen grid (M-grid)

- Suomen Akatemian rahoittama seitsemän yliopiston, Fysiikan tutkimuslaitoksen ja CSC:n yhteishanke
 - Kumppaneina fysiikan ja kemian laboratoriot, käyttäjinä samojen ryhmien tutkijat
- Yhdeksän Linux-klusteria eri puolilla Suomea
- Ensimmäinen laaja tuotantokäyttöön tarkoitettu grid-hanke Suomessa
- Tarkoitettu peräkkäisohjelmien ja helposti rinnakkaistuvien ohjelmien ajamiseen
- Grid-väliohjelmistona NorduGrid ARC





NorduGrid-yhteistyö



- Alkoi vuonna 2001
- Tuotantokäytössä oleva grid, yhteensä n. 5000 suoritinta
 - Oikeita käyttäjiä ja sovelluksia, mutta myös hyvä testialusta
- Osin ohjelmakehitystä, osin kokemusten ja resurssien jakoa
- Suomessa mukana CSC ja Fysiikan tutkimuslaitos HIP
- **Kaikille avoin**, ei vaatimuksia laiteresurssien tarjoamisesta



DEISA

- EU-rahoitteinen, eurooppalaisten superlaskentakeskusten välinen nelivuotinen yhteistyöprojekti
- Keskittyy erityisesti vaativiin sovelluksiin superkoneympäristössä
- CSC:n vastuulla projektin tiedotus
- Alkuvaiheessa yhdistetään CSC:n ja muiden keskusten keskenään samanlaisia IBM-koneita
- Myöhemmin mukaan erilaisia konearkkitehtuureja



ENACTS ja EMBRACE

- European Network for Advanced Computing Technology for Science (ENACTS) on juuri päättymässä oleva EU-projekti, jossa tehtiin selvityksiä teknologioista ja käyttäjien tarpeista
 - CSC osallistui erityisesti käyttäjätutkimuksen tekoon, loppuraportti valmistumassa
 - Muita hyviä raportteja kotisivuilla <http://www.enacts.org>
- European Common Model for Biological Resource And Content Exchange (EMBRACE) on vuonna 2005 alkava EU-rahoitteinen grid-hanke, jossa pyritään yhdistämään erilaisia biotieteiden tietokantoja ja analyysityökaluja
 - Projektiin rekrytoidaan näillä näkymin yksi uusi työntekijä alkuvuonna 2005



ESO-hanke

- Ohjelmistohanke, jossa kehitetään massiivisten tähtitieteen tietomäärien analyysia
 - Tekniset ratkaisut vielä pitkälti avoimia, tarvitaan omaa kehitystyötä
 - Grid-teknologioita käytettäneen eri puolilla maapalloa sijaitsevien tietokantojen yhdistämiseen
- Osallistujat CSC ja Helsingin yliopiston tähtitieteen laitos
 - Suomessa kuusi työntekijää, alkaa vuoden 2005 alusta
- Suomi liittyi European Southern Observatoryyn (ESO) heinäkuussa 2004, tämä projekti on osa Suomen liittymismaksua



Uudet algoritmit grid-ympäristössä

- Maantieteellisesti hajautetussa ympäristössä valon nopeus asettaa alarajan latenssille
 - ➔ Korkeaa latenssia sietävien algoritmien merkitys kasvaa
- Grid-ympäristöt dynaamisia, resursseja tulee ja menee
 - ➔ Sekä väli- että sovellusohjelmien vikasietoisuus nousee entistä tärkeämpään rooliin
- Optimointi
 - ➔ Grid tarjoaa tunnistuksen yhdellä salasanalla, mutta miten valitaan kulloinkin paras resurssi?
 - ➔ Välimuistien toiminta ja datan hajautus: missä ja kuinka monena kopiona tieto kannattaa säilyttää?



Muita teknisiä haasteita

- Yhteensopivuus
 - Heterogeeninen grid vain korostaa kirjastoriippuvuus- ja muita käyttöläheisiä ongelmia, virtuaalikoneteknologiasta ratkaisu?
 - Halutaan yhdistää massiivisia, usein vuosia toisistaan erillään kehitettyjä ohjelmistoja
- Käyttöoikeuksien hallinta
 - Käyttäjän tunnistukseen jo vakiintuneet ratkaisut (PKI), mutta käyttöoikeuksien hallinnassa paljon tekemistä
 - Vaaditaan hienojakoista ja hallittua, mutta käyttäjän kannalta läpinäkyvästi toimivaa oikeuksien edelleenvälitystä ohjelmien välillä (ja organisaatiorajojen yli!)



Ei-teknisiä haasteita

- Käyttöliittymät
 - Nykyisissä ympäristöissä erityisesti debuggaus hankalaa
 - Tallennuskapasiteetti ei enää ole ongelma, tiedon hallinta on!
- Luottamus
 - Miten hävitetään paikan merkitys mutta taataan käyttäjälle että hänen tietonsa ovat silti tallessa ja yksityisyyssuojan piirissä?
- Yhteistyö
 - Toimiva yhteistyö on voimaa, vaikkei teknologiassa mitään niin ihmeellistä olisikaan!



Pari esimerkkiä opinnäytetöistä

- Grid VCR (Henrik Høy Karlsen, Tanska)
 - TV-kortilla varustettu Linux-tietokone, jolle voi NorduGridin kautta lähettää pyynnön nauhoittaa pätkä ohjelmaa halutulta kanavalta
 - <http://www.imada.sdu.dk/~karlsen/vcrrecord.html>
- Accounting in Grid Environments (Peter Gardfjäll, Ruotsi)
 - Käytönseurannan ja grid-”pankin” arkkitehtuurin suunnittelu ja prototyypin toteutus
 - Toimi pohjana Swegridissä vastikään käyttöönotetulle järjestelmälle (<http://www.sgas.se>)
 - <http://www.cs.umu.se/~peterg/thesis/thesis.pdf>





Kiitos!

- CSC:n grid-sivut:
 - <http://www.csc.fi/grid/>
- Tämän esityksen kalvot:
 - [http://staff.csc.fi/ajt/presentations/](http://staff.csc.fi/ajt/presentations/Grid-hankkeita_ja_tulevaisuuden_nakymia_2004-12-20.pdf)
[Grid-hankkeita_ja_tulevaisuuden_nakymia_2004-12-20.pdf](http://staff.csc.fi/ajt/presentations/Grid-hankkeita_ja_tulevaisuuden_nakymia_2004-12-20.pdf)
- Kysymyksiä?