

Suomi valmis LHC-datalle

25.6.2008 14:56 — Tuomas Linnake

tuomas.linnake@digitoday.fi

Cernin LHC-hiukkastörmäytin muodostaa ennätysellisen tietoteknisen haasteen, josta Suomikin saa osansa purtavakseen, tieteen tietotekniikan keskus CSC korostaa. Suomen pitäisi kuitenkin olla nyt valmis vastaanottamaan massoittain kiihdyttimen tuottamaa dataa.

Kiihdyttimen avulla tehtävien hiukkaskokeiden tarkoituksena on tuottaa uutta tietoa aineen perusrakenteista. Kokeiden on arvioitu tuottavan mittausdataa noin viisitoista petatavua (15 miljoonaa gigatavua) vuodessa. Data hajautetaan tietoverkkojen kautta eri laskentakeskuksille ympäri maailmaa.

Cernissä pidettiin laskentaharjoitus helmi- ja toukokuussa. Siihen liittyneessä CSA08-harjoituksessa simuloitiin CMS-koeaseman (Compact Muon Solenoid) tuottaman tiedon tallentamista ja käsittelyä ensimmäisen kolmen kuukauden aikana.

Harjoituksessa lähetettiin päivässä keskimäärin 100 000 työtä, jotka jakaantuivat niin sanotun Tier-mallin mukaan alemman tason keskuksille ajettaviksi. Simuloitua CMS-raaka- ja testidataa siirrettiin enimmillään 1700 megatavun verran sekunnissa. CMS-kokeessa siirrettiin dataa kaikkiin laskentakeskuksiin yhteensä 3,6 petatavua toukokuun harjoituksen aikana.

CSC:n mukaan CSA08 oli suomalaisille erityisen merkittävä, koska harjoituksessa pystyttiin ensimmäisen kerran lähettämään CMS-kokeen analyysitöitä Cernin tukeman niin sanotun gLite-grid-väliohjelmiston sekä Yhdysvalloista Condor-järjestelmän kautta Suomeen NorduGrid ARC -väliohjelmistolla suoritettavaksi. Ennen harjoitusta tämä ei ollut mahdollista.

Sovituksen merkittävyyttä suomalaisille CMS-tutkijoille voidaankin CSC:n mielestä verrata tilanteeseen, missä muualta maailmasta ei voitaisi lähettää postia Suomeen. Laskentatöiden lähettämisen onnistuminen oli siten tärkeä edistysaskel.

Pientä viilattavaa

Mutta paljastiko harjoitus mitään, mikä Suomen päässä vaatisi

vielä hieromista ennen tositoimien alkamista loppukesällä tai alkusyksyllä?

- Kriittiset komponentit toimivat. Mutta meillä on esimerkiksi monitoroinnin ja tilastojen keräämisen kanssa vielä tiettyjä haasteita. Ne eivät ole aivan niin hyvissä kantimissa kuin voisivat olla, projektipäällikkö **Antti Pirinen** Fysiikan tutkimuslaitos HIPistä kertoo.

- Se ei kuitenkaan estä fysiikkalaskentaa. Harjoitus ei paljastanut mitään, mitä emme olisi tienneet, vaan ennemminkin sen, että homma toimii riittävin osin, Pirinen korostaa. Hän ei näe mitään syytä, miksi ongelmia ei ehdittäisi korjata ajoissa.

Pirisen mukaan haasteet liittyvät lähinnä byrokratiaan, siihen, saadaanko hienoja PowerPoint-esityksiä tehtyä.

Dataa 11:een keskukseen

Cernissä käytettävä laskentamalli toimii siten, että Cern itse toimii niin sanottuna 0-tason keskuksena (Tier-0), josta raakadata ja esikäsitelty data hajautetaan yhteentoista eri maissa sijaitsevaan 1-tason keskukseen (Tier-1).

Tier-1-keskukset tallentavat datan, käsittelevät sitä sekä tarjoavat datapalveluja sekä Cernille että pienemmille 2-tason keskuksille. Kakkoskeskukset tarjoavat levytilaa sekä edellytykset simulaatioille ja data-analyysille.

Nordic Data Grid Facility (NDGF) tarjoaa yhden pohjoismaisen Tier-1-tason resurssin, joka on hajautettu Suomeen, Ruotsiin, Norjaan ja Tanskaan. Tieteen tietotekniikan keskus CSC ylläpitää tästä Suomen osuutta. Mukana on myös HIP.

Tämän lisäksi Suomessa on myös Tier-2-tason resurssi, joka on toteutettu HIPin ja CSC:n yhteistyönä.

<http://www.digitoday.fi/data/2008/06/25/suomi-valmis-lhc-datalle/200817025/66>