

Suomi ottaa vastaan LHC-tietoa 10 gigabitin sekuntivauhdilla

6.10.2008 08:08 — Kalevi Nikulainen

Fysiikan tutkimuslaitos HIP on yhdessä Tieteen tietotekniikan keskuksen CSC:n kanssa toteuttanut suomalaisten hiukkasfysiikan tutkijoiden käyttöön Cernistä tulevan LHC-tiedon laskentajärjestelmän, joka on osa pohjoismaista NDGF-laskentaresurssia.

HIP on hankkinut CSC:ltä ylläpidon Suomen osuuteen hajautetusta Tier-1-resurssista. Tier-2-resurssi on pääosin myös CSC:llä, HIP:n ja Jyväskylän yliopiston paikallisten resurssien lisäksi.

Tällä hetkellä ALICE- ja CMS-kokeiden tutkijoiden käytettävissä on noin 700 laskentaydintä, 170 teratavua levytilaa ja noin 70 teratavua nauhatilaa. Tämän lisäksi Suomi sai LHC:n tiedonsiirtoa varten ensimmäisenä Pohjoismaana suoran 10 gigabit/s:n (Gbps) valopolkuyhteyden.

NDGF palvelee WLCG:n lisäksi myös muita tieteenaloja kuten bioinformatiikkaa ja geotieteitä.

Pohjoismaiset laskentakeskukset

Nordic DataGrid Facility (NDGF) koostuu pohjoismaisista laskentakeskuksista Norjassa, Ruotsissa, Suomessa ja Tanskassa, aloittaa. CSC:n mukaan se aloittaa nyt Cernissä syyskuussa käynnistetyn Large Hadron Collider (LHC)-törmäytin tuottaman datan varastoimisen ja käsittelyn.

NDGF on yksi yhdestätoista Tier-1-resurssista eli pääresurssista, joilla LHC:n tieteellinen työ tehdään. Tier-1-keskukset hallinnoivat yhdessä Cernissä vuosittain kertyvää 15 miljoonan gigatavun datamäärää ja joutuvat käsittelemään vieläkin suurempia raakadatasta johdettuja datamääriä. NDGF:n osuus tästä maailmanlaajuisesta tehtävästä on noin kuusi prosenttia, mikä on ennätyksellistä pohjoismaiselta tieteelliseltä yhteisöltä, kertoo CSC.

Maailmanlaajuisen kokeellisen hiukkasfysiikan yhteisön

keräämiä tiedon tallennus- ja varastointiresursseja kutsutaan nimellä Worldwide LHC Computing Grid (WLCG). WLCG käyttää eturintaman tietotekniikan menetelmiä keräämään yhteen eri maissa olevat resurssit. Grid-teknologia on niin WLCG-hankkeen kuin NDGF-hankkeen perusta.

Muista Tier-1-laskentaresursseista poiketen NDGF on itsessään hajautettu resurssi, joka yhdistää yhdeksän laskentakeskusta neljässä Pohjoismaassa.

LHC-laskenta ja grid

LHC-tiedeyhteisön etsimiä vastauksia ei ole käsillä ilmeisessä muodossa. Uuden energia-alueen protoni-protoni-törmäyksissä syntyy useita uusia alkeishiukkasia, ja niiden jättämistä jäljistä neljä suurta koeasemaa ALICE, ATLAS, CMS ja LHCb ottavat eräänlaisia digitaalisia kuvia.

Uuden hiukkasen tai ilmiön havaitseminen vaatii näiden törmäysten jälkien analysoimisen valtavalla määrällä tietokoneita. Aikaisemmat hiukkasfysiikan kokeet toteutettiin lähinnä Cerniin paikallisesti sijoitetuilla laskentaresursseilla. LHC-aikakaudella tämä ei ole enää mahdollista.

Tarvittavan laskenta- ja tallennuskapasiteettiin keräämiseksi WLCG käyttää grid-teknologiaa. Grid on laskentaresurssien tai muiden palvelujen muodostama verkko, joka tarjoaa tallennus- ja laskentapalveluja niitä tarvitseville virtuaaliorganisaatioille, hieman vastaavasti kuin sähkövoimalat tuottavat sähköverkossa kuluttajien käyttämän sähkön.

Markkinoilla on erilaisia grid-ratkaisuja. Pohjoismaissa suosituin grid-ohjelmisto on avointa lähdekoodia käyttävä Advanced Resource Connector, jonka pohjoismainen Nordugrid yhteisö on kehittänyt Linux- ja Unix-käyttöjärjestelmille. Hajautettua tallennusta varten käytössä on Euroopan ja USA:n kiihdytinlaboratorioissa kehitetty dCache-ohjelmisto, joka on erikoistunut valtavien levy- ja nauha-arkistojen käsittelyyn.

<http://www.digitoday.fi/data/2008/10/06/suomi-ottaa-vastaan-lhc-tietoa-10-gigabitin-sekuntivauhdilla/200825949/66>