Projektisuunnitelma

Makromolekyylikompleksien rekonstruktioprosessin rinnakkaistaminen Beowulf-ympäristössä

Antti Perälä

Ilkka Koskenniemi

Matti Kääriäinen

Peter Lamberg

Taneli Mielikäinen

Valtteri Rahkonen

Helsinki 27. tammikuuta 2000

Ohjelmistotuotantoprojekti: Projektisuunnitelma

Ohjaaja: Satu Sihvo

Asiakkaat:

Janne Ravantti

professori Dennis Bamford

Vastuuhenkilö: Juha Taina

HELSINGIN YLIOPISTO

Tietojenkäsittelytieteen laitos

Sisältö

Kuvat

1 Kuvaus

1.1 Makromolekyylikompleksien rekonstruktioprosessin rinnakkaistaminen Beowulf-ympäristössä

Makromolekyylikompleksien rekonstruktioprosessin rinnakkaistaminen Beowulf-ympäristössä ohjelmistotuotantoprojektin tehtävänä on toteuttaa ohjelma, joka määrittää makromolekyylikompleksien orientaatiot niitä esittävissä kaksiulotteisissa kuvissa. Ohjelman tavoitteena on löytää paras mahdollinen kulma jokaiselle makromolekyylin kuvalle. Ohjelma on suunniteltu apuvälineeksi määritettäessä esimerkiksi elektronimikroskoopin pyyhkäisykuvassa olevien virusten (makromolekyylien) orientaatiot. [Lanc98]

1.2 Toiminta

Ohjelma saa syötteenään joukon makromolekyylikompleksin kaksiulotteisia kuvia, sekä saman makromolekyylikompleksin kolmiulotteisen mallin. Kolmiulotteisesta mallista muodostetaan kaksiulotteinen projektio aina jokaiselle kulmalle, jossa kuvien yhteensopivuutta verrataan. Kunkin kuvan yhteensopivuuksista eri kulmista muodostettuihin kaksiulotteisiin projektioihin, voidaan laskea makromolekyylikompleksin orientaatio kuvassa.

1.3 Projektin ympäristö

Koska projektioiden muodostus, ja sen vertailu kuviin vaatii paljon laskentatehoa, laskenta hajautetaan `Beowulf`-klusteriin [Beowulf] käyttäen `MPI`-kirjastoa (<u>M</u>essage-<u>P</u>assing <u>I</u>nterface [MPI]).

Beowulf-ympäristöön voidaan lisätä prosessoreita laskennan vaativuuden mukaan, jopa laskennan aikana. Ympäristön eri prosessorit voivat olla ominaisuuksiltaan heterogeenisiä.

MPI-kirjasto tarjoaa korkean tason palveluita, joita käyttämällä ohjelmasta on tarkoitus tehdä riittävän skaalautuva ja riippumaton Beowulf-ympäristön rakenteesta.

2 Toteutus

2.1 Projektin vaiheet

Projekti toteutetaan vesiputousmallia käyttäen. Mallin mukaisesti projekti on jaettu neljään vaiheeseen. Vaiheet erottuvat selvästi Gantt-kaaviosta (s. 6)

Vaihe 1. Vaatimusanalyysi Vaatimusanalyysin tuloksena on määrittelydokumentti. Määrittelydokumentti kuvaa yksityiskohtaisesti järjestelmäna vaatimukset ja rajoitteet. Vaatimukset rajataan tarkasti välttämättömiin ja lisäominaisuuksiin, jotka toteutetaan, jos jää aikaa. Määrittelydokumenttin tekemiseen on varattu puolitoista viikkoa ja se valmistuu 4.2.

Vaihe 2. Suunnittelu Suunnitteludokumenttissa tehdään määrittelydokumenttiin pohjautuen neljä suunnitelmaa: tietosuunnitelma, arkkitehtuurisuunnitelma, liittymäsuunnitelma ja moduulisuunnitelma. Aikaa suunnitteludokumentin tekoon on projektin luonteesta johtuen varattu kahdeksan viikkoa, sillä suurin työ projektissa sisältyy algoritmeihin tutustumiseen ja sopivien algoritmien valitsemiseen. Algoritmien toteuttaminen ei vaadi suuria määriä koodia. Suunnitteludokumentti valmistuu 31.3.

Vaihe 3. Toteutus Toteutusvaiheen tuloksena saadaan toteutusdokumentti, joka pitää sisällään tuotetun ohjelmakoodin, sekä mahdolliset muutokset suunnitteludokumenttiin. Optimaalisimmillaan toteutusdokumentti on vain ohjelmalistaus. Toteutusvaiheeseen on varattu neljä viikkoa, valmistuen 28.4.

Vaihe 4. Testaus Testausta tehdään jokaisessa osavaiheessa, ei pelkästään toteutuksen jälkeen. Lopullisen tuotteen testaussuunnitelman tekeminen aloitetaan hyvissä ajoin ennen tuotteen valmistumista, heti suunnitteludokumentin valmistumisen jälkeen. Tuloksena saadaan testaussuunnitelma ja testauksen tulokset, jotka kerätään testausdokumenttiin. Testausdokumentti valmistuu 10.5.

2.2 Määräajat

Projektien osien valmistumispäivät on asetettu seuraavasti:

Projektisuunnitelma 27.1.2000

Määrittelydokumentti 4.2.2000

Suunnitteludokumentti 31.3.2000

Toteutusdokumentti 28.4.2000

Käyttöohje 10.5.2000

Ylläpitodokumentti 10.5.2000

Testausdokumentti 10.5.2000

Loppuraportti 10.5.2000

2.3 Työtunnit

Projektin aikatauluarvio työtunteina on 2040 (työntekijöitä 6 * tunteja viikossa 20 * viikkoja 17). Kukin projektin jäsen seuraa säännöllisesti suorittamiaan työtunteja lomakkeella. Työtuntien kirjaaminen tehdään projektin etenemisen seurannan helpottamiseksi ja loppuraporttia silmällä pitäen.

3 Työnjako

3.1 Infrastruktuuri

Peter Lambergin ja Valtteri Rahkosen vastuulla on hajautukseen käytettävään kirjastoon (`MPI`-kirjasto) tutustuminen ja ohjelman hajautuksen suunnittelu.

3.2 Algoritmit

Taneli Mielikäinen, Matti Kääriäinen, Ilkka Koskenniemi ja Antti Perfumedälä vastaavat projektioiden laskemisesta makromolekyylin kolmiulotteisesta mallista sekä saatujen projektioiden vertailusta makromolekyylin kaksiulotteisiin elektronimikroskooppikuviin. Tämän työryhmän ensimmäinen tavoite on tutustua eri algoritmivaihtoehtoihin ja määrittää niiden asettamat vaatimukset muulle järjestelmälle.

3.3 Dokumentointi

Kaikki ryhmän jäsenet osallistuvat dokumenttien tuotantoon. Ohjelmointi-osuudet määritellään suunnitteludokumentin pohjalta.

3.4 Projektipäällikkö

Projektipäällikkönä toimii aluksi Taneli Mielikäinen. Projektipäällikköä vaihdetaan projektin puolivälissä.

3.5 Projektisihteeri

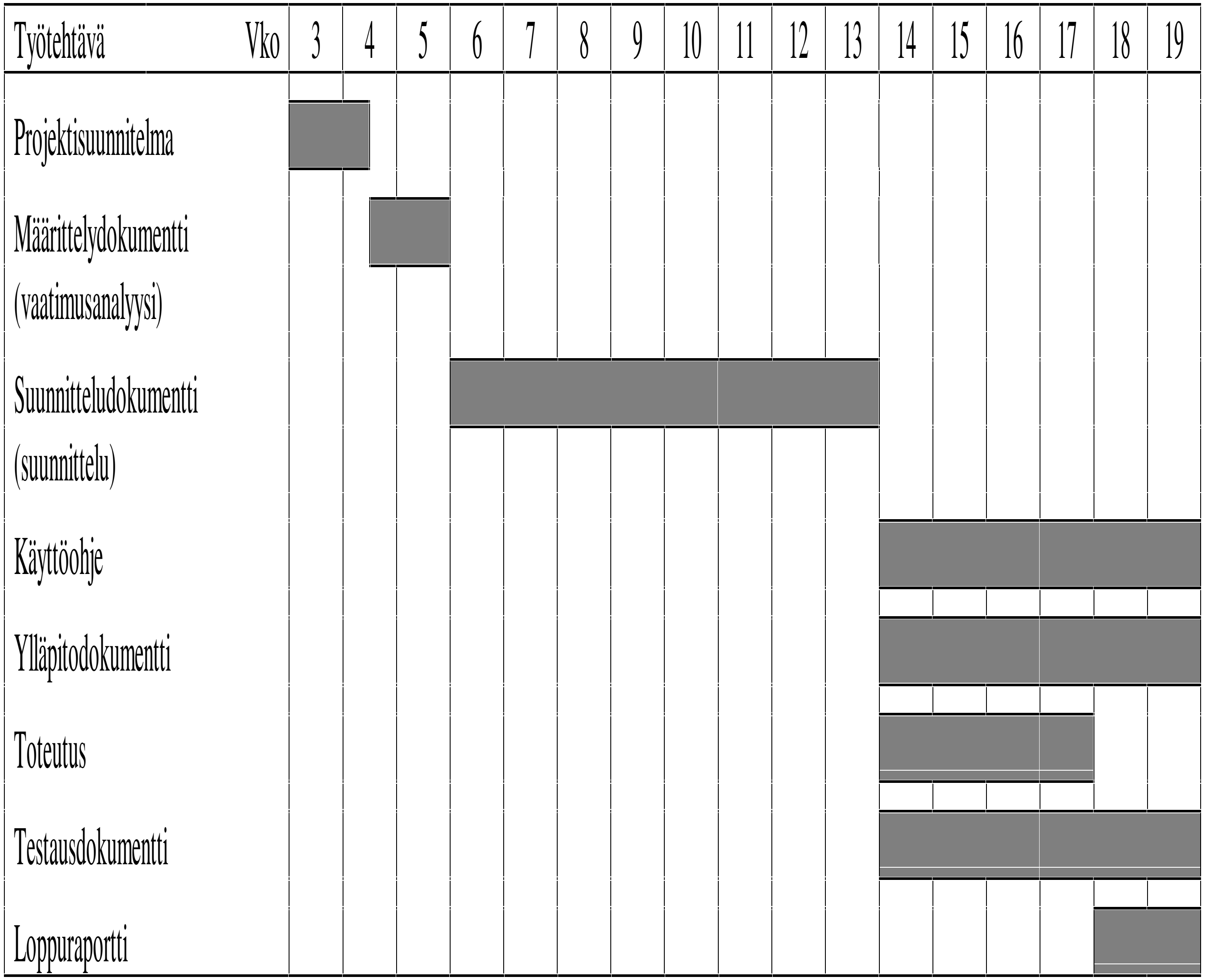
Projektin sihteerinä toimii Antti Perälä.

Lähteet

[Lanc98] Cristopher J.Lanczycki, Calvin A. Johnson, Benes L. Trus, James F. Conway, Alasdair C. Steven, Robert L. Martino
US National National Institutes of Health
Parallel Computing Strategies for Determining Viral Capsid Structure by Cryoelectron Microscopy
IEEE Computational Science & Engineering, Huhti-Touko 1998, s. 76-79

[Beowulf] http://www.beowulf.org/
Daniel Ridge, Donald Becker, Phillip Merkey, Thomas Sterling, Becker, Phillip Merkey
Beowulf: Harnessing the Power of Parallelism in a Pile-of-PCs
Proceedings, IEEE Aerospace, 1997

[MPI] http://www.mpi-forum.org/index.html
http://www-unix.mcs.anl.gov/mpi/



Kuva 1: Gantt-kaavio