



Grid aprēķinu vide

teorija • metodes • aprēķini



GRIDTĪKLS UN TĀ IZMANTOŠANA

Dokumenta faila vārds:	LG-nodevums-gridtiklsunizmantosana.doc
Aktivitāte:	3.aktivitāte "Lietojumprogrammu un lielu datu apjomu izmantošanas risinājumu izstrāde Grid vidē"
Projekta numurs:	VPD1/ERAF/CFLA/05/APK/2.5.1./000055/027
Organizācija:	Latvijas Universitātes aģentūra "Latvijas Universitātes Matemātikas un informātikas institūts"
Autori:	Bruno Martuzāns, Baiba Kaškina

Anotācija:

Šajā darbā tiek aprakstīta gridtīkla būtība un tā izmantošanas iespējas un pieredze. Tiek aplūkoti dažādi varianti gridtīkla izpratnē un uztverē. Tiek izskaidrota gridtīkla vieta Eiropas pētniecības telpas veidošanā un e-zinātnes attīstībā. Tiek izanalizēti vairāki Eiropas ietvarprogrammās izstrādātie projekti, kas veltīti gridtīklu pielietošanai liela apjoma aprēķinu veikšanā, molekulu inženierijā u.c. Tāpat šajā dokumentā tiek aplūkoti gridtīklu projekti, kas veltīti liela apjoma datu uzkrāšanai un apstrādei. Nobeigumā tiek piedāvāts to iespējamo gridtīkla pielietojumu apskats, kuri varētu attīstīties Latvijā.



Saturs

1. VISPĀRĪGA INFORMĀCIJA UN IEVADS GRIDTĪKLU APRĒĶINOS	3
1.1. GRIDTĪKLA DEFINĪCIJA.....	3
1.2. LOKĀLĀ TĪKLA IDEJAS TĀLĀKATTĪSTĪBA.....	3
2. GRIDTĪKLI EIROPĀ	5
2.1. EIROPAS ZINĀTNES STRATĒGIJA.....	5
2.2. EIROPAS LĪMEŅA GRIDTĪKLU VEIDOŠANA	6
3. PAŠREIZĒJIE GRIDTĪKLA PIELIETOJUMI.....	7
3.1. GRIDTĪKLA IZMANTOŠANA APRĒĶINOS	7
3.1.1. <i>Saules aktivitātes pētījumi</i>	7
3.1.2. <i>Projekts CrossGrid</i>	8
3.1.3. <i>Projekts GEMSS</i>	9
3.1.4. <i>Molekulu inženierija</i>	9
3.2. DATU GRIDTĪKLS UN TĀ IZMANTOŠANA.....	10
3.2.1. <i>Grid kā liela datu glabātuve</i>	10
3.2.2. <i>Lielākais strādājošais gridtīkls pasaulē</i>	11
4. GRIDTĪKLA PERSPEKTĪVAS LATVIJĀ	12

Šodien, pateicoties Gridtīklu tehnoloģijas attīstībai, ko virza Eiropas datorzinātnieku jaunā paaudze, mēs varam gaidīt jaunu tehnoloģisku revolūciju.

V.Redinga, Eiropas informācijas sabiedrības komisāre

1. VISPĀRĪGA INFORMĀCIJA UN IEVADS GRIDTĪKLU APRĒĶINOS

1.1. GRIDTĪKLA DEFINĪCIJA

Pasaulē ir daudzas gridtīklu definīcijas, no kurām ir grūti izvēlēties labāko. Definīciju ir tik daudz tāpēc, ka katrs ar gridtīklu saprot kaut ko citu, jo gridtīkls ir nākotne, un nākotni katrs iztēlojas savādāk, un tā katram arī būs citāda. Gridtīkla būtību var saprast, salīdzinot to ar globālo tīmekli. Ja tīmeklis nodrošina katram neierobežotas iespējas iegūt informāciju, kas glabājas internetam pieslēgtajos datoros, gridtīkls nodrošina neierobežotas iespējas izmantot citus datoru resursus – spēju uzglabāt datus un veikt aprēķinus.

Izmantojot gridtīkla pakalpojumus, cilvēks var uzglabāt praktiski neierobežotu datu apjomu vai veikt neierobežota apjoma aprēķinus. Ļoti raksturīga gridtīkla pakalpojumiem ir paralelizācija, t.i., iespēja vienu uzdevumu sadalīt vairākās daļās un tās vienlaicīgi izpildīt uz daudziem datoriem, tādējādi paātrinot kopējo uzdevuma izpildi. Tāpat gridtīklam ļoti piemēroti daudzvariantu uzdevumi, kur jārisina viens un tas pats uzdevums daudzas reizes ar dažādiem ievadāmiem parametriem.

Bieži gridtīklu salīdzina ar elektrības piegādes tīklu. Gridtīkls nodrošina iespēju veikt vajadzīgā apjoma aprēķinus un saglabāt jebkura apjoma datus jebkurā vietā, pieslēdzoties šim tīklam ar jebkuru datoru, nebēdājot un neinteresējoties par to, kurā vietā šie aprēķini patiesībā tiks veikti, kur šie dati tiks saglabāti, pa kādiem vadiem un uz kuriem attiecīgā informācija tiks pārsūtīta.

Tas ir tāpat kā elektrības tīklā, kurš piegādā strāvu jebkuram elektriskajam aparātam un neviens neinteresējas, kurā spēkstacijā šī strāva ir sākotnēji radusies. Nevar nepateikt arī to, ka par strāvas patēriņu ir jāmaksā. Arī gridtīkla izmantošana bez maksas nav un nebūs, taču šobrīd tā izmantošanas izmaksas sedz dažādi nacionāli un Eiropas mēroga projekti.

1.2. LOKĀLĀ TĪKLA IDEJAS TĀLĀKATTĪSTĪBA

Gridtīkla ideja sasaucas ar tām idejām, kas jau no paša datortīklu sākuma tika izteiktas un diskutētas. Jau tad aprēķini uz attāliem datoriem notika, un arī informāciju varēja glabāt ne tikai uz savas organizācijas lieldatora diskem. Arī lokālie tīkli tika radīti tieši šādu un līdzīgu vajadzību dēļ. Tāpēc gridtīklu var arī uztvert kā globālu lokālo tīklu. Bet tas viss nemazina gridtīkla lomu – ja kaut kas izaug liels, tas kļūst spēcīgs, tas var izdarīt daudz un kļūst vajadzīgs daudziem.



Gridtīkls tiek pieminēts tik bieži un tik dažādos gadījumos, ka tam ir ne tikai daudzas definīcijas, bet arī vairāki nosaukumi. Dažas kompānijas, kas uzsāk vai paredz piedāvāt gridtīkla komerciālus pakalpojumus, izdomā arī jaunus, skaistākus nosaukumus šādiem pakalpojumiem. Protams, „liels lokāls tīkls” vai „viegli izmantojams klusters”, vai „dalītie aprēķini” neskan tik labi kā gridtīkls, toties „aprēķini no padebešiem” skan labāk un poētiskāk. Aprēķini no padebešiem, kā šeit tiek tulkots termins „*cloud computing*” (acīmredzamais tulkojums – “aprēķini no mākoņiem” tika atmests kā politiski tendēts), parasti tiek skaidroti kā tādi aprēķini, par kuriem lietotāji nezina, kur tie notiek un kāda mašīna tos izpilda, tāpēc viņiem liekas, ka šī aprēķinu iespēja nāk gluži kā no padebešiem.

Eiropas Pētniecības telpa nav paredzēta tikai zinātniekiem vai modernajai industrijai, tā ir domāta ikvienam un katram eiropietim.

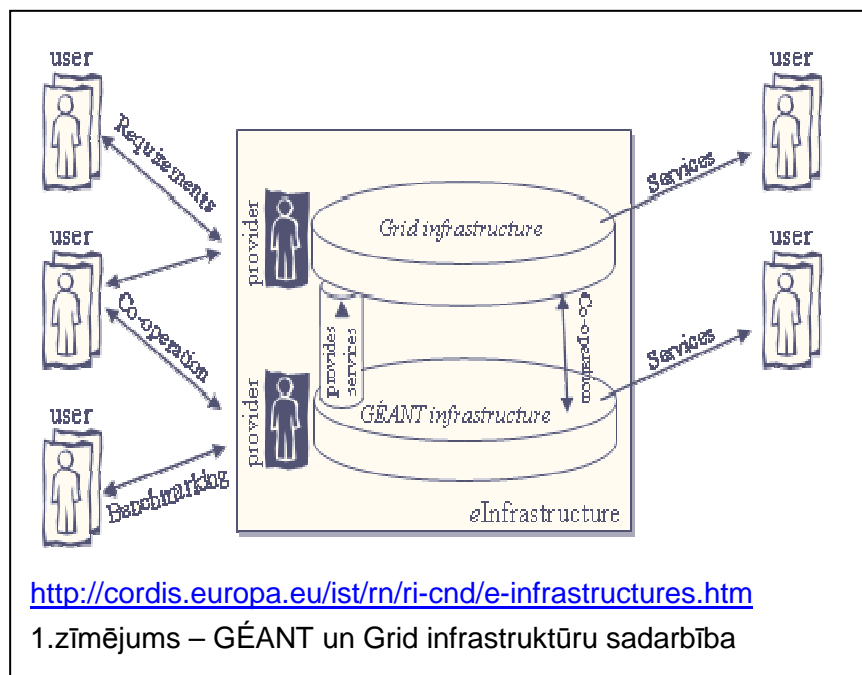
ES zinātnes komisārs J.Potočņiks

2. GRIDTĪKLI EIROPĀ

2.1. EIROPAS ZINĀTNES STRATĒGIJA

Eiropas politikas veidotājiem nav mazāko šaubu, ka Eiropas tālākā attīstība un tās vieta pasaulē ir būtiski atkarīga no zinātnes attīstības. Jau 2000. gadā Eiropas Savienība nolēma radīt **Eiropas Pētniecības telpu** (*European Research Area*), kas ļautu nodrošināt visas Eiropas zinātnieku sadarbību kopīgu pētniecisko mērķu sasniegšanā un palīdzētu nostiprināt Eiropas zinātnes vadošo lomu globālajā mērogā.

Eiropas Pētniecības telpas radīšanā kā būtisks faktors tiek uzsvērta e-infrastruktūras veidošana, kurai savukārt ir jā sastāv no divām pamatkomponentēm – Eiropas gigabitu akadēmiskā tīkla GÉANT un Eiropas līmeņa gridtīkla infrastruktūras. Šo abi komponentu loma un sadarbība ir parādīta 1. zīmējumā.



Akadēmisko tīklu GÉANT Latvijā veido LU MII speciālisti un, pateicoties IZM atbalstam, tas aktīvi attīstās. Jāpiebilst, ka tikai GÉANT tīkls spēj nodrošināt zinātniekus ar citiem modernākajiem tīkla pakalpojumiem, ne tikai gridtīkla izmantošanu.

Lai nodrošinātu Latvijas zinātniekus ar visām e-infrastruktūras iespējām, LU MII speciālisti tika iesaistīti arī Baltijas gridtīkla veidošanas projektā BalticGrid (2005-2008), kas izveidoja vienotu Baltijas valstu gridtīklu ar kopīgu politiku, nodrošināja plašas zinātnieku sadarbības iespējas un radīja bāzi tālākai attīstībai. Gridtīkla infrastruktūras veidošanai tieši Latvijā ir

veltīts projekts “GRID aprēķinu vide: teorija, metodes, pielietojumi”, kura ietvaros ir izstrādāts arī šis materiāls.

2.2. EIROPAS LĪMEŅA GRIDTĪKLU VEIDOŠANA

Gan Latvijas, gan Baltijas gridtīkli gūs pilnīgu piepildījumu tikai tad, ja tie iekļausies visas Eiropas līmeņa tīklos. Šie tīkli varētu tikt pieslēgti tīklam EGEE (*Enabling Grids for E-science*), kas ir vislielākais Eiropas gridtīkls, kurš tika izveidots 5. un 6. ietvarprogrammu projektos. Šajos projektos izstrādāts visas Eiropas vienota gridtīkla veidošanas plāns, kas paredz gridtīkla plašu izmantošanu e-zinātnē (e-science), t.i., zinātnisko atklājumu veikšanā galvenokārt ar aprēķinu palīdzību. EGEE projekta iniciators ir Eiropas kodolpētījumu centrs CERN, un šis projekts sākotnēji lielā mērā bija orientēts uz šī institūta vadībā būvēto Lielo hadronu kolaidēru (LHC), kas pilnībā tiks palaists 2008. gadā un spēs iegūt tik daudz elementāro daļiņu fizikas datu, ka to apstrādei būs nepieciešami datori visā Eiropā un vēl arī citur pasaulē. Pašlaik šī kolaidēra datu apstrādei tiek veidots atsevišķs LHC aprēķinu gridtīkls.

Tuvākos gados Eiropas gridtīkla struktūrā gan var notikt izmaiņas, jo 7. ietvarprogrammā tika apstiprināts projekts “Eiropas Gridtīkla iniciatīva, organizācijas pētījums” (European Grid Initiative Design Study), saīsināti EGI, kurā ir jāizstrādā jauna viseiropas gridtīkla koncepcija un organizatoriskais modelis. Šādam tīklam ir jāspēj attīstīties līdzsvaroti un ir jābūt spējīgam pastāvēt ilgstoši.

3. PAŠREIZĒJIE GRIDTĪKLA PIELIETOJUMI

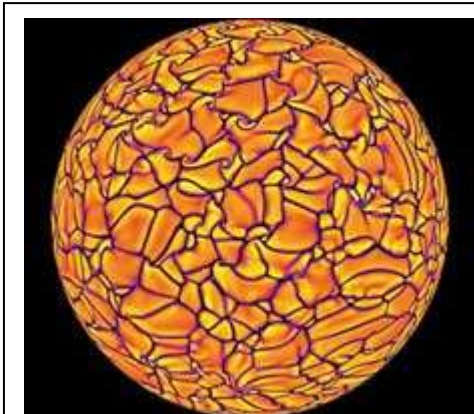
Gridtīklu visvairāk pielieto zinātniskajos pētījumos. Šajā virzienā visperspektīvāk ir izmantot šī tīkla iespējas matemātiskajā vai statistiskajā modelēšanā, kas tagad tiek saukta par e-zinātni. Kādu fizikālu parādību, procesus vai tehnisku iekārtu apraksta ar matemātiskiem vienādojumiem un tad risina šos vienādojumus. Var arī sarežģītus procesus noreducēt uz vienkāršāku atsevišķu procesu apvienojumu (piemēram, balstoties uz „pirmajiem principiem”) un tad šos atsevišķos procesus atsevišķi modelēt. Tieši šāda veida pielietojumiem tika izstrādāti pirmie datoru. Tā kā pati ideja ir ļoti veca, bet zinātnieki vienmēr sūdzējās, ka datoru aprēķinu jauda nav pietiekoša. Nav šaubu, ka gridtīkli varēs apmierināt lielu daļu no aprēķinu jaudas vajadzībām, kaut arī var viegli paredzēt, ka zinātnieku prasības atkal pieaugs un būs jāveido jaudīgākas mašīnas un tīkli, vai arī jāmeklē principiāli jaunas pieejas.

3.1. GRIDTĪKLA IZMANTOŠANA APRĒĶINOS

Liela apjoma aprēķini ir nepieciešami daudzās zinātnes un tehnikas jomās, un iespējai tos izpildīt ļoti bieži ir būtiska nozīme. Daudzi Eiropas Savienības atbalstītie projekti jau sākot no 5. ietvarprogrammas veltī lielu uzmanību gridtīkla izmantošanai aprēķinos. Dažas no problēmām, kur gridtīkls ir izrādījies būtisks, ir aprakstītas zemāk. Jāpiebilst, ka šādos aprēķinos tiek izmantoti gan personālo datoru, gan superdatoru gridtīkli.

3.1.1. Saules aktivitātes pētījumi

Pirmais aprēķinu piemērs gan nav no Eiropas, bet no Amerikas, kur gridtīkla izmantošana jau sen tiek uzskatīta par būtisku komponenti zinātniskajā darbībā. Viens no aktīvākajiem



2.zīmējums – uzliesmojumi
Saules fotosfērā

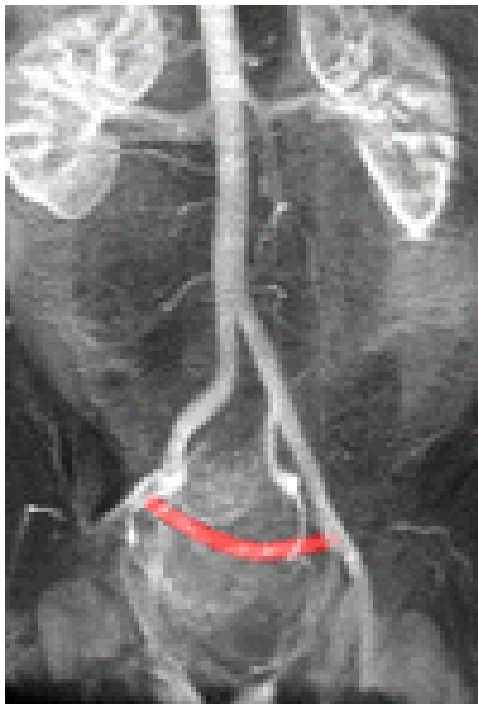
gridtīkliem TeraGrid, kas savieno 11 superdatorus, sevi dēvē par “atvērtu zinātnisko atklājumu infrastruktūru”, tā liekot manīt, ka zinātnisko atklājumu veikšana tagad tiek automatizēta. Cik lielā mērā tas tā ir, var katrs spriest pats, bet šis gridtīkls tiešām ļauj izpildīt aprēķinus, kas agrāk nebija iespējami.

Tā Kolorado Universitātes pētnieki, izmantojot šī gridtīkla iespējas, modelēja saules vielas plūsmas dinamiku ar mērķi noskaidrot kā rodas zvaigžņu magnētiskais lauks (zvaigžņu dinamo), kas ir ļoti veca, interesanta un grūta problēma. Jāpiebilst, ka saules vielas plūsmas pētījumi ir pamatā arī saules aktivitātes pētījumiem, kuri ļauj prognozēt, piemēram, uzliesmojumus saules fotosfērā, kas ietekmē radiokomunikāciju uz zemes vai pat var

izsaukt slimību saasinājumus cilvēkiem, it sevišķi kosmonautiem, kā arī ietekmēt lidmašīnu apkalpes un pasažierus.

3.1.2. Projekts CrossGrid

Viens no pirmajiem projektiem ar vislielāko interesi par liela apjoma aprēķiniem bija 5. ietvarprogrammas projekts CrossGrid. Šajā projektā tika aplūkotas problēmas četros zinātnes virzienos. Pirmais tika veltīts plūdu monitorēšanai uz upēm, protams, pievēršot lielāku vērību tām problēmām, kas rodas, pārplūstot lielajām Viduseiropas upēm. Ņemot vērā sniega un lietus dinamiku, ūdens noteci, upju hidrauliku un citus faktorus, izdodas modelēt plūdus un prognozēt to rašanos un tālāku attīstību pilsētu un lauku rajonos, pie tam arī gadījumos, ja notiek kāda kataklizma, piemēram, kāds dambis tiek pārrauts. Cits virziens tika veltīts atmosfēras un jūras piesārņojuma attīstības modelēšanai, kā arī datizrači meteoroloģisku datu datubāzēs. Tāpat šajā projektā tika pētītas iespējas veikt lielā hadronu kolaidierā iegūto datu apstrādi.



3.zīmējums – asinsvadu šuntēšana

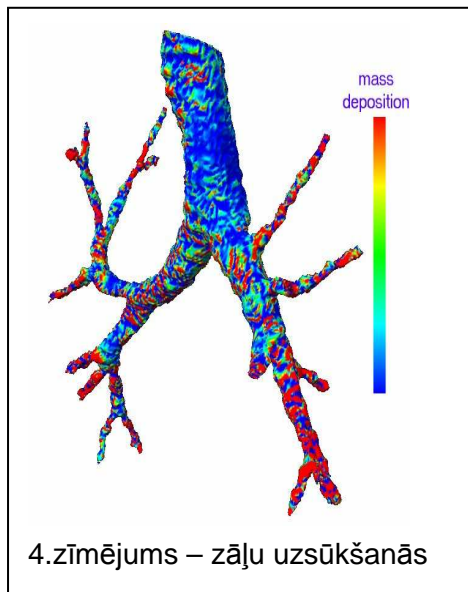
Interesanta ir problemātika, kas saistīta ar asinsvadu ķirurģiju. Gadījumos, kad rodas nepieciešamība veikt arteriālo šuntēšanu, projektā izstrādātā sistēma dod ķirurgam iespēju izvēlēties pacientam labāko operācijas variantu. Izmantojot reālos pacienta asinsvadu magnetiskās rezonanses izmeklēšanas vai datortomogrāfijas attēlus, ķirurgs interaktīvi izvietojamo šuntu (būtībā virtuālu artēriju) un sistēma aprēķina asins plūsmas dinamiku, nosakot panākto uzlabojumu asins apgādē un izvērtējot riska faktorus, kas var rasties, piemēram, papildu plūsmas virpuļu dēļ.

Šis darbs prasa liela apjoma sarežģītu aprēķinu veikšanu, jo asinsvadu plūsma tiek aprēķināta sarežģītas konfigurācijas trīsdimensionālā sistēmā, un pie tam klasiskā

hidrodinamika šeit nav pielietojama – asinis nav Ņūtona šķidrums. Papildu sarežģījumi rodas arī tad, ja tiek ņemta vērā artēriju sienīņu elastība un pašas plūsmas pulsējošais raksturs.

3.1.3. Projekts GEMSS

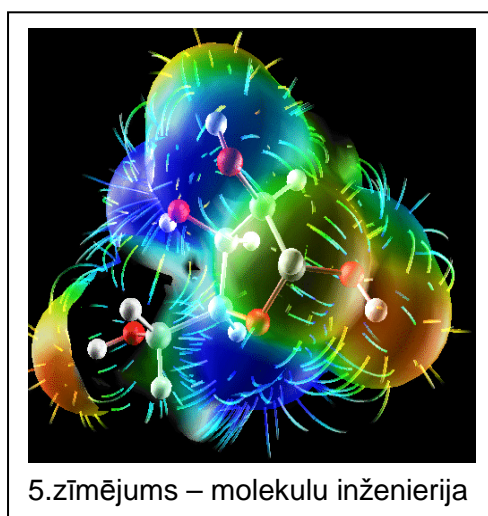
Projekts GEMSS (Grid-Enabled Medical Simulation Services) izstrādāja modelēšanas un attēlu apstrādes starpprogrammatūru gridtīkliem, kuru varētu plaši izmantot medicīniskos pielietojumos. Pielietošanas iespējas tika novērtētas testa problēmu risināšanā, kuras galvenokārt ir saistītas ar invazīvās iejaukšanās plānošanu. Piemēram, tika izstrādāta pirmsoperācijas plānošanas sistēma žokļa un sejas ķirurģijā ar iespēju apskatīt uz datora ekrāna ķirurģiskās iejaukšanās sekas vēl pirms operācijas. Cits testa pielietojums ļauj novērtēt smadzeņu izvietojanos galvaskausā pēc nopietnas neuroķirurģiskas operācijas. Līdzīgi cits testa pielietojums pētī iespēju novērtēt jonizējošā izstarojuma kūļu intensitāti smadzeņu audzēju apstarošanas plānošanā, ņemot vērā dažādo izstarojuma izkliedi un absorbciju dažādos audos.



Vēl viens testa pielietojums ir ar pētniecisku ievirzi – tiek pētīta ieelpojamo zāļu uzsūkšanās reālās plaušās, atkarībā no inhalatora konstrukcijas, zāļu sastāva, pilienu lieluma u.c. (skat. 4. zīm.). Arī šie procesi

prasa sarežģītu liela apjoma aprēķinu izpildi, jo tiek rēķināta trīsdimensionāla šķidrums plūsma sarežģītas konfigurācijas apgabalā.

3.1.4. Molekulu inženierija



Eiropas zinātnieki lielas cerības liek uz gridtīkla izmantošanu organisko un it sevišķi bioorganisko molekulāro procesu modelēšanā.

Projekts Bio-GRID piedāvā pieejas portālu biomolekulārajai modelēšanai, lai nodrošinātu speciālistiem liela apjoma aprēķinu izmantošanas iespējas, izmantojot interaktīvu vizuālu interfeisu.

Projekts OpenMolGrid veido gridtīkla infrastruktūru, kas ļautu konstruēt molekulas ar iepriekš noteiktām īpašībām. Molekulu inženieris no dažādiem molekulu fragmentiem izveido jaunu molekulu kandidātus, kuriem būtu jābūt noteiktām īpašībām. Aprēķinu ceļā tiek noteikts, vai šādas īpašības jaunajai molekulai būs. Šādi aprēķini ir komplicēti un var tikt veikti tikai gridtīklos. Var paredzēt, ka nākotnē jaunu zāļu

projektēšana un to iedarbības pārbaude uz cilvēka šūnās esošām ķīmiskām vielām notiks aprēķinu ceļā. Dažādu molekulu savstarpēja iedarbība ir ļoti nozīmīga arī visos ģenētiskās informācijas pārneses procesos, gēnu ekspresijā u.tml., un arī šo iedarbību var aprēķināt, izmantojot gridtīklus.

3.2. DATU GRIDTĪKLS UN TĀ IZMANTOŠANA

Akadēmisko lietotāju aprindās gridtīklu uztver galvenokārt kā kolektīvas lietošanas aprēķinu vidi, kā nākotnes globālu infrastruktūru, kura gala rezultātā visus pasaules datorus apvienos vienā patiešām milzīgā Superdatorā, kurš izpildīs visus dotajā brīdī pasaulē veicamos aprēķinus vienlaicīgi uz visiem šajā brīdī ieslēgtajiem pasaules datoriem, saglabās aprēķinu rezultātus drošā vietā un paziņos zinātniekiem, ka viss ir izdarīts. Pagaidām vēl mazāk izplatīta ir uztvere, ka gridtīkls varētu būt datu glabāšanas un apstrādes vide, kur glabājas visi pasaulē glabājamie dati, kur tie tiek apstrādāti un vajadzīgā informācija tiek nekavējoties piegādāta pieprasītājam.

3.2.1. Grid kā liela datu glabātuve

Protams, zinātnieki veic arī tādus pētījumus, kas rada lielus datu masīvus, kuri ir jā saglabā un jāapstrādā. Pirmkārt, esošie un iegūtie zinātniskie dati ir jā saglabā, ar tiem tūlīt ir jāstrādā tālāk un tie ir jāpadara visiem cilvēkiem vai vismaz visiem zinātniekiem pieejami. Šo uzdevumu veiks digitālās bibliotēkas un datu repozitoriji, kas jāveido gan nacionālajā, gan Eiropas līmenī, un tiem ir jā satur zinātniskie raksti, grāmatas, multimediju un e-izglītības materiāli u.c.

Dati Eiropas digitālās bibliotēkās tiks ievietoti dažādās valodās un to pilnvērtīga izmantošana prasīs izstrādāt tādas tulkošanas programmas, kuru tulkojumus varēs saprast un tiem ticēt. Tas, protams, nebūs iespējams bez semantiskās analīzes, kuras sarežģītība strauji aug līdz ar analizējamā teksta apjomu un atkal šeit var palīdzēt tikai gridtīkli.

Visi dati, kurus iegūst dažādos novērojumos, piemēram, radioastronomiskie novērojumi, saules novērojumi, zemes atmosfēras un jūras plūsmu novērojumi ar satelītiem, atmosfēras ozona monitoringa dati, un vēl daudzi citi tiek saglabāti un ir jāpadara plaši pieejami. Pie tam ir jāņem vērā, ka datu apjoms pieaug par vairākiem terabaitiem katru dienu un līdz ar to šo datu apstrāde kļūst netriviāla un arī saglabāšana kļūst par nopietnu problēmu, it sevišķi, ja tiek glabāta grafiskā informācija.

Ļoti daudz informācijas apstrādes un saglabāšanas problēmu piegādā bioinformātika, it sevišķi, ja grib atrast un izpētīt ģenētisko informāciju, kas nosaka cilvēka iespējamās saslimšanas. Raksturīgais skaitlis šeit ir 3,2 miljardi, kas ir nukleoīdu skaits cilvēka DNS molekulā. Brīvi operēt ar šāda apjoma informāciju nespēj pat dators, ja tas ir tikai viens. Svarīgs uzdevums ir veidot datu bāzes ar iegūtajiem ģenētiskajiem un klīniskajiem datiem, padarīt tās pieejamas pētniekiem un izstrādāt to pētniecības metodes.

Daudz informācijas uzkrāj medicīna, it sevišķi attēlu formā. Eiropā zinātnieki aktīvi strādā pie projekta MammoGrid+, kas veic mammogrāfisko attēlu uzkrāšanu un apstrādi ar gridtīkla metodēm. Cits piemērs ir 2007. gada maijā izziņotā IBM sistēma Grid Medical Archive Solution (GMAS), kas kopā ar glabāšanu piedāvā arī programmatūru, serverus un pakalpojumus. GMAS ir komerciāls produkts, kas vismaz pagaidām nepretendē uz globāliem mērogiem; tas tiek piedāvāts slimnīcām un medicīnas pētniecības iestādēm, lai varētu nodrošināt medicīnisko attēlu un citas pacientu informācijas glabāšanu ilgu laiku un tās saņemšanu pēc ārstniecības iestāžu pieprasījuma.

Datu uzkrāšanas un apstrādes problēmu risināšanā gridtīkliem ir lielas perspektīvas. Piemēram, ASV lielveikalu tīkls Wal-Mart ir uzsācis plašu radioiezīmju (RFID) izmantošanu savā preču plūsmas organizācijā. Katra prece tiks apgādāta ar radioiezīmi, un automātiski lasītāji reģistrēs šīs preces ceļu no noliktavas uz veikala plauktu, tālāk pircēja somā un

beidzot gar lasīšanas iekārtu pie izejas, kura nolasīs visu somas saturu un automātiski atskaitīs atbilstošo naudas summu no maksājumu kartes. Protams, informācija par katra pircēja pirkumiem tiks uzkrāta, un šī pircēja vēlmes un pirkšanas ieradumi tiks izpētīti. Lai kā tas viss arī izskatītos no privātuma saglabāšanas viedokļa, šo datu apstrādi pa visu lielveikalu tīklu nevar uzskatīt par vienkāršu uzdevumu, kas būtu veicams uz viena lokāla tīkla bez darbu sadalīšanas paralēli veicamos uzdevumos. Tiek prognozēts, ka katru darba dienu tiks uzkrāti 7 terabaiti dienā, tātad 2,5 petabaiti gadā, kas ir nozīmīgs apjoms.

3.2.2. Lielākais strādājošais gridtīkls pasaulē

Gridtīkli, turklāt globālā mēroga, ir izveidoti vairāki un ir pieejami visiem interneta lietotājiem. Vislabāk pazīstamais tīkls ir Google kompānijas radītā interneta informācijas meklēšanas infrastruktūra. Katru pieprasījumu Google meklētājam apstrādā kāds no serveriem, kas sadala šī pieprasījuma izpildi daudziem šajā brīdī mazāk noslogotiem datoriem, savāc ziņas atpakaļ no šiem datoriem, sagrupē atrasto informāciju, ņemot vērā iepriekš reģistrētās klienta intereses, un izsniedz to klientam. Apmēram tā strādā šis meklētājs. Pati kompānija nesniedz informāciju par savas infrastruktūras pamatmodeļi, nerunājot jau par detaļām. Tas ļauj speciālistiem trenēties dažādos minējumos, ka pašlaik kompānijai pieder ap 30 datu centriem visā pasaulē, kuros ir izvietoti daži desmiti tūkstošu serveru klasteru un pāri par pusmiljonam personālo datoru. Ļoti būtiski ir tas, ka Google informācijas gridtīkls ir pieejams no jebkura interneta pieslēguma pasaulē, t.i., tas ir patiesi globāls.

Tieši šī, gridtīklam raksturīgā pieeja izmantot nevis vienu vai divus lielus superdatorus un glabāt informāciju uz cietajiem diskkiem, bet gan daudzus datorus, kuri varētu informāciju ātri apstrādāt galvenokārt operatīvajā atmiņā, būtiski palīdzēja Google kompānijai uzvarēt konkurentus interneta meklētāju nozarē un izrauties priekšgalā. Galu galā tādu datoru, kas varētu strādāt ar vairāk kā simts miljonu meklēšanas pieprasījumu dienā, nodrošinot datu pārraidi 3-4 petabiti sekundē, nemaz nav. Taču Google infrastruktūra to panāk. Un jāatzīst, Google ar savu informācijas meklēšanas darbu galā tiek labi un attīsta vēl arī citus pakalpojumus, kuri izmanto līdzīgu pieeju.

Tādā vai līdzīgā veidā savus gridtīklus veido arī citas globālā datortīkla komerciālās organizācijas. Tagad tie ir arī citiem interneta informācijas gigantiem kā Yahoo, Excite vai Incomi. Protams, arī lielie interneta veikali eBay vai Amazon nespētu nodrošināt savus pakalpojumus, neizmantojot gridtīklus. Katra no Amazon veikala daudzu desmitu miljonu pircējiem veiktos pirkumus uzkrāj un analizē ne jau vienā datorā. Šādu informāciju vienā datorā diez vai var saglabāt, un tas arī nebūtu ieteicams no datu drošības un ātras apstrādes viedokļiem.

4. GRIDTĪKLA PERSPEKTĪVAS LATVIJĀ

Protams, gridtīklam Latvijā, tāpat kā citur Eiropā, ir jāattīstās. Mazāk skaidrības ir par to, kā tas varētu attīstīties un kā to varētu izmantot Latvijas zinātnieki un uzņēmēji. Un vēl, kā tas varētu ietekmēt katra cilvēka ikdienas dzīvi.

Matemātiskā modelēšana kā zinātnes nozare pēdējos gados Latvijā nav attīstījusies strauji, kaut arī vairākas zinātnieku grupas ir saglabājušās un izpilda aprēķinus ārzemju partneriem vai pat vietējām organizācijām. Zinātnes un rūpniecības attīstības tendences neļauj apgalvot, ka e-zinātnei Latvijā būtu paredzams straujš pieprasījuma kāpums, tomēr tajā iesaistās arī studenti un ieinteresētie speciālisti. Katram, kas ir sagatavojis kādu aprēķinu programmu (datoram vai savam darbam), ir izdevīgi aprēķinus veikt kaut kur citur nevis uz sava datora. Bieži to dara uz savas iestādes datora, kas ir pieejams no lokālā tīkla. Ja pieejamo aprēķina iespēju nav pietiekami, tad piekļuve gridtīklam kļūst svarīga. Nākas secināt, ka gridtīkls kļūst svarīgs tikai intensīvi strādājošiem speciālistiem, kas var un vēlas konkurēt Eiropas līmenī.

Var paredzēt, ka gridtīkls kļūs pieprasīts bioinformātikas speciālistu vidū. Ja papildīsies nepārtraukti izteiktās prognozes, ka ģenētikas sasniegumu rezultātā kļūs iespējams paredzēt daudzas konkrētā cilvēka saslimšanas, balstoties uz informāciju viņa genomā, tad būs nepieciešama regulāra liela apjoma ģenētiskās informācijas apstrāde, kas visērtāk būs izdarāma gridtīklā. Acīmredzot šīs informācijas apstrādes organizācijas – speciāli centri vai kabineti lielākajās slimnīcās pieslēgsies gridtīklam un aktīvi to izmantos.

Viena no paredzamajām gridtīkla izmantošanas nozarēm, kas būs svarīga Latvijā, ir nākamās paaudzes tīmeklis, kurā varēs veikt meklēšanu pēc satura nevis pēc noteiktiem atslēgas vārdiem. Šāda meklēšana balstās uz semantisko analīzi, kas ir pietiekoši darbietilpīgs uzdevums un var tikt veikts uz vairākiem datoriem uzreiz. Semantiskās analīzes sekmes un praktiskas izmantošanas iespējas nosaka piekļuve gridtīklam, kas var būt speciāli veidots šim nolūkam līdzīgi kā iepriekš aprakstītais Google meklēšanas gridtīkls. Semantiskās analīzes problēmas latviešu valodai Latvijas zinātnieki jau risina.

Nav šaubu, ka gridtīkli tiks plaši izmantoti dažādu drošības problēmu risināšanā. Novērošanas kameru ieraksti būs jā saglabā pietiekami ilgi un pie tam glabātavai ir jāatrodas tālu no ieraksta vietas. Gridtīkls var glabāt datus grūti atrodamā vietā pavisam citā valstī. Līdz ar radioiezīmju izplatīšanos arī Latvijā, radīsies nepieciešamība uzkrāt šo ierīču piegādāto informāciju, kas nenoliedzami būs ļoti apjomīga. Arī šeit veidosies atsevišķa industrija ar plašām iespējām iesaistīt gridtīklus.

Grid tīkla attīstībai ir iespējami daudz un dažādi scenāriji, mēs nezinām, kurš būs īstais, tomēr viens ir skaidrs, gridtīkla attīstība būs intensīva un interesanta.