

ISSN 1333-5693



Vol. 3, broj 78, 2002.

Ruđer

tematski dvobroj:
informatika i računarstvo
u znanstvenom radu



optička mreža * klaster * Grid * e-Science * aplikacije * usluge

U ovom broju:

L. Budin:

IKT u znanosti3

K. Skala:

Optička mreža4

M. Stipčević:

Klaster računalo
na IRB-u6

D. Krišto:

Optimizacija rada na
grozdu8

K. Skala:

Grid i e-Science11

K. Vlahoviček:

Bioinformatika15

K. Skala, N. Pavković:

Farma klastera16

M. Mihaljević:

Grid primjene18

Z. Šojat:

Nanoračunarstvo20

J. Meštrović:

Baza podataka i
znanja23

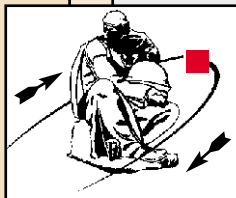
R. Mileta:

Usluge
znanstvenicima24

M. Pranjić

Računalna sigurnost . .26

Suradnja, projekti . . .27



Na naslovnici:

**Donirani optički kabel
i institutski grozd**

Cijenjeno čitateljstvo !

U skoroj budućnosti, društvo će biti vođeno znanjem, koristit će se cjeloživotno obrazovanje i uspostaviti će se novi model življenja i rada, a to će se najprije i odnositi na znanstveni rad. Prepoznaje se pojava e-Sciencea, kao aplikacijska nadgradnja Grida. Taj izraziti generički razvoj očit je u prikazu dekadske skale dinamike razvoja računarstva kako slijedi:

- 1950. otkriće tranzistora,
(e-tranzistor),
- 1960. ostvarenje integriranih sklopva,
(e-chip),
- 1970. razvoj mikroprocesora, μP
(e-micro processor),
- 1980. osobno računalo, PC,
(e-personal computer),
- 1990. pojava Interneta, WEB
(e-mail, e-business, e-learning...),
- 2000. razvoj Grida,
(e-Science),
- 2010. elektroničko društvo,
(e-society).

Gornji prikaz ukazuje na nelinearnu dinamiku razvoja formaliziranog Mooreovim zakonom koji čini stratešku okosnicu u planiranju svakog razvojno-investicijskog programa, sve dok se ne pojavi nanoračunarstvo, o čemu

suradnik Računalnog centra piše u ovom dvo-broju. Prema tom zakonu udvostručenje brzine rada procesora odvija se svakih 18 mjeseci, a brzina prijenosa putem računalnih mreža, svakih 9 mjeseci. To je dovelo do novih mogućnosti jer su se brzina rada PC računala i brzina prijenosa podataka preko mreže izjednačile i tako su ostvareni preduvjeti za intenzivan razvoj razvedenih distribuiranih sustava (klaster, grid...).

Te činjenice nas navode na stratešku odluku po kojoj, kao i do sada, Institut Ruđer Bošković mora biti među prvima koji u znanstvenom radu primjenjuje najsuvremenije informatičke i računalne tehnologije i metode, kako bi se u globalnim okvirima osigurala primjerena znanstvena produkcija u multidisciplinarnom i multiinstitucionalnom radu. Razmatrajući te teze na Ruđeru je pokrenut razvojni proces (klaster, optička kampus mreža, nove usluge, CRO GRID inicijativa..), čiji se djelomični rezultati očituju i kroz sadržaj ovog tematskog dvobroja. Težimo praćenju značajnih promjena koje se uočavaju na prekretnici tisućljeća, u dobroj namjeri za daljnji napredak i prosperitet. Ovaj dvobroj želi biti ilustracija te težnje i djelovanja, po svemu sudeći nužnog strateškog opredjeljenja.

Karolj Skala
urednik tematskog broja

Došli u Institut tijekom srpnja i kolovoza 2002.:

Ines-Ana Biljan; dr. sc. Ivančica Bogdanović-Radović; Barbara Mikac dipl. inž. biologije; Luka Mikelić dipl. inž. kemije; Ivana Rajčan dr. medicine; Milan Stupar; Helena Štorek dipl. inž. kemije; Margarita Vuković dipl. ecc.; dr. sc. Igor Weber.

Otišli iz Instituta tijekom srpnja i kolovoza 2002.:

Adrijana Culak; Katja Džepina, dipl. inž. kemije; Maja Gašparić dipl. inž. medicinske biokemije; Liana Hoelbling; Igor Jurak dipl. inž. biologije; Vjekoslava Starčević; dr. sc. Natko Urli; Boris Vetnić.

Izbori u zvanja tijekom srpnja i kolovoza 2002.

mlađi asistent: Goran Landek, Ivana Rajčan, Igor Tomažić, Dijana Žilić.

asistent: Sandra Horvat, Mauro Štifanić.

viši asistent: Maša Katić, Tamara Stipčević, Hrvoje Štefančić, Marijana Vinković.

znanstveni suradnik: Ivančica Bogdanović-Radović, Krunoslav Brčić-Kostić, Irena Ciglenečki-Jušić, Marijana Erk, Dragica Fuks, Blaženka Gašparović, Nadica Ivošević, Neven Soić, Mario Stipčević, Senka Terzić.

viši znanstveni suradnik: Bartolo Ozretić.

Disertacije izradene u Institutu i obranjene tijekom srpnja i kolovoza 2002.

Željko Jakšić: Razvoj i primjena brze mikro metode određivanja i praćenja oštećenja DNA u škrgama dagnje *Mytilus galloprovincialis* (Lamarck 1819), voditelj R. Batel, obrana (Rovinj) 18. 07. 2002.

Maja Sikirić: Utjecaj poveznika na fizičko-kemijska

svojstva asimetričnih dimernih surfaktanata, voditeljica N. Filipović-Vinceković, obrana 16. 07. 2002.

Hrvoje Štefančić: Hijerarhija vremena života hadrona s teškim kvarkovima, voditelj B. Guberina, obrana 01. 07. 2002.

Amir Dubravić: Nova metoda u analizi biokinetičkog sustava, voditelj A. Šantić, obrana 17. 07. 2002.

Magistarski radovi izrađeni u Institutu i obranjeni tijekom srpnja i kolovoza 2002.

Sandra Horvat: Laserski sustav ALMY za precizno određivanje položaja detektora tragova čestica, voditelj K. Kadija, obrana 03. 07. 2002.

Vesna Janicki: Modeliranje profila loma tankih slojeva, voditelji H. Zorc i A. Dulčić, obrana 16. 07. 2002.

Diplomski radovi izrađeni u Institutu i obranjeni tijekom srpnja i kolovoza 2002.

Sonja Bec: Oblikovanje objektno-relacijske baze podataka, mentor Z. Skočir (izvanredni prof. FER-a Zagreb), voditelji I. Ružić i J. Pecar-Ilić, obrana 10. 07. 2002.

Željka Maglica: Organizacija 18S - 25S rDNA kromatina u stanicama korjenovog vrška hrasta lužnjaka (*Quercus robur* L.) i hrasta kitnjaka (*Q. petraea* Liebl.), voditelj N. Ljubešić, obrana 12. 07. 2002.

Maja Šepelj: Priprava 4-[(4-deciloksifenil)oksikarbonil]fenil-2-(4-oktiloksifenil)acetata kao potencijalnog dopanta za feroelektrične materijale, voditeljica A. Lesac, obrana 10. 07. 2002.

Informacijska i komunikacijska tehnologija u znanosti

piše: **Leo Budin,**
FER

Nakon drugog svjetskog rata zbili su se neki značajni događaji koji su odredili pojavu i razvitak informacijske i komunikacijske tehnologije. Prvo, objelodanjena su istraživanja i konstrukcije računalnih naprava koje su za vrijeme rata čuvane kao vojna tajna. Početkom pedesetih godina na tržištu su se pojavila prva računala. To su bile naprave sastavljene pretežno od elektronskih cijevi a zauzimale su prostor cijelih soba i trošile na stotine kilovatsati električne energije dnevno. Drugo, krajem četrdesetih godina izumljen je tranzistor. Pokazalo se da tranzistor može, uz mnogo manju potrošnju energije, a i mnogo pouzdanije, nadomjestiti elektronsku cijev. Daljnjim razvojem poluvodičke tehnologije nastala je nova mikroelektronička tehnološka grana. Današnji mikroelektronički sklopovi sadrže do nekoliko milijuna tranzistora smještenih na pločicu silicija od nekoliko kvadratnih centimetara. Treće, istovremeno s razvojem mikroelektronike i računalne tehnologije razvijalo se i područje telekomunikacija. Od jednostavnih telefonskih centrala, povezanih žičanim kabelima, i relativno izoliranih radiodifuznih sustava razvio se današnji svjetski sustav za prijenos glasa, slike i teksta. Tim se sustavom do nedavno međusobno izolirana računala međusobno povezuju u multimedijske računalne mreže.

Za spregu mikroelektronike, računalne tehnologije i komunikacija upotrebljavaju se u zadnjih nekoliko godina dva skupna naziva. Jedan od njih je naziv informacijska tehnologija koji se obilježava kraticom IT, a drugi informacijska i komunikacijska tehnologija (s kraticom ICT od engleskog Information and Communication Technology), koji je rašireniji u europskom okruženju. Svoje je attribute ta tehnologija dobila zbog toga što omogućuje prihvata, pohranjivanje, prijenos, preradu i jednostavnu uporabu svih vrsta informacija.

U znanosti se računala upotrebljavaju od njihovih prvih dana. Moglo bi se ustvrditi da su znanstvene potrebe, posebice one usmjerene na vojne primjene, početno bile i glavni motiv inovativnih poduhvata u razvitku te tehnologije. Iako su u okruženju tržišnog gospodarstva dominantan udio tržišta postupno zauzele poslovne primjene i osobna računala, napredak u području računanja za potrebe znanstvenih istraživanja slijedio je tehnološki razvitak. Tijekom godina izgrađivana su specijalizirana super računala s posebnim naglaskom na postizanju što veće računalne snage. U zadnjih su desetak godina naponi u tom smjeru objedinjeni u jednu novu granu računarstva visokih performansi (High-Performance Computing). Osnovna obilježja te grane računarstva je uvođenje raznolike vrste sklopovskog paralelizma, počevši od čvrsto povezanih višeprocorskih sustava (tightly-coupled multiprocessor systems), preko računalnih grozdova (clusters), pa do raspodijeljenih računalnih mreža (computer networks) i računalnih spletova (computer grids). Programska potpora (middleware) ovakvih sustava mora, s jedne strane, osigurati djelotvornu uporabu svih njihovih dijelova i, s druge strane, omogućiti što jednostavnije odvijanje različitih primjenskih programa. Primjenski programi moraju biti zasnovani na paraleliziranim algoritmiima i pripremljeni kao paralelni sustav zadataka prilagođen arhitekturi pojedinih sustava.

Podsjetimo se da se znanstveniku uobičajena paradigma djelovanja pri istraživanju nekog fenomena sastoji od slijedećih koraka: postavljanja hipoteze, razrade modela i predviđanja njegovog ponašanja, postavljanja eksperimenta i prikupljanja podataka iz tog eksperimenta te analize rezultata. Kada se ustanovi neslaganje između predviđenog ponašanja modela i rezultata eksperimenta ti se koraci moraju iterirati. Prve uporabe računala najčešće su se svodile na predviđanje ponašanja modela (simuliranje) pri čemu su veliki naponi ulagani za izgradnju robusnih algoritma zasnovanih na numeričkoj matematici. Povećanjem računalne snage sklopovlja i razvitkom prikladnih arhitektura računala te razradom programske potpore omogućeno je svojevrsno automatsko vođenje eksperimenata i prikupljanje podataka, predočavanje (vizualizacija) rezultata i pomoć u njihovoj interpretaciji. Svaki znanstvenik

u svom polju djelovanja može najbolje posvjedočiti kako mu računalo podupire vlastita istraživanja.

Današnje stanje dobro ilustrira pregled koji su sačinili C. J. Holland i J. Grosh opisujući modernizaciju računalnih resursa Ministarstva obrane SAD-a namijenjenih istraživačkoj djelatnosti (High-Performance Computing Modernization Program). Oni navode da

je zbog bolje organizacije i boljeg razumijevanja potreba obavljena podjela na deset računskih polja (computational technology areas) i to: Climate/weather/ocean modeling and simulation, Computational chemistry and material science, Computational electromagnetics and acoustics, Computational electronics and nanoelectronics, Computational fluid dynamics, Computational structural mechanics, Environmental quality modeling and simulation, Forces modeling and simulation, Integrated modeling and test environment, Signal/image processing.

Atribut computational koji se pojavljuje ispred konvencionalnih naziva disciplina ukazuje na bitno promijenjeni metodološki pristup svakoj od njih. U prijevodu na hrvatski jezik bilo bi prikladno taj atribut prevesti u računski. Tako bi se dobili nazivi Računska kemija i Računska elektronika. Eventualni nazivi Računalna kemija ili Računalna elektronika bi se, naime, mogli protumačiti kao kemija koja se bavi kemijskim procesima u računalima (što možda nema smisla) odnosno kao elektronika koja razmatra elektroničke sklopove unutar računala (što ima smisla, ali može biti pogrešno). Naziv computational technology mogao bi se prevesti u računska tehnologija.

Još jednom treba naglasiti da se, za razliku od konvencionalnih stolnih računala, gdje se ubrzavanje računanja postiže naprosto ubrzavanjem rada procesora, u nakupini procesora, spremničkih modula, prospojnih putova ili umreženih računala povećanje računalne snage može bitno unaprijediti samo uz pažljivo pripremljenu programsku podršku prilagođenu svakoj pojedinačnoj arhitekturi.

U već spomenutom projektu američkog Ministarstva obrane planirane se dvije aktivnosti koje omogućuju lakše snalaženje u računskoj infrastrukturi. Prva od aktivnosti (pod nazivom Program Environment and Training) omogućit će stvaranje mreže konzultanata te uspostavljanje centara za izobrazbu usklađenih s potrebama pojedinog od deset računskih polja. Druga je aktivnost (Common HPC Software Support Initiative) usmjerena na projekte koji će obuhvatiti razradu algoritama i pripremu programa prilagođenih pojedinim računskim strukturama.

Mi bi u Hrvatskoj svakako trebali slijediti ovakve i slične zamisli i prilagoditi ih našim uvjetima i potrebama. Pritom bi trebalo objediniti sve postojeće znanje i ekspertizu. U strateškom dokumentu Informacijska i komunikacijska tehnologija - Hrvatska u 21. stoljeću Vlade Republike Hrvatske nalazi se potpora za takvo djelovanje gdje se u okviru nekoliko preporuka ističe povezanost znanosti i informacijske i komunikacijske tehnologije. Između ostalog jedna od ponuđenih aktivnosti u okviru preporuke broj 16 glasi: "Potaknuti temeljna istraživanja usmjerena prema novoj informacijskoj i komunikacijskoj infrastrukturi zasnovanoj na radikalno novim načelima distribuiranog računarstva i komunikacija kako bi ona postala sveprisutna, pokretna i skalabilna te kako bi jamčila kvalitetu usluge."



piše: Karolj Skala,
IRB ZE/RC

Optička mreža kao okosnica razvoja Instituta



U suvremenom informatičkom okruženju znanje je dostupno svima, bez vremenskih, prostornih ili socioekonomskih ograničenja. Na taj se način ostvaruje napredak i pokretanje novih primjena računala i računalnih mreža u gotovo svim pogledima ljudskog djelovanja. Računala, mreže i komunikacije tehnologije su stvaranja novog znanja izuzetno značajnog pokretača napretka.

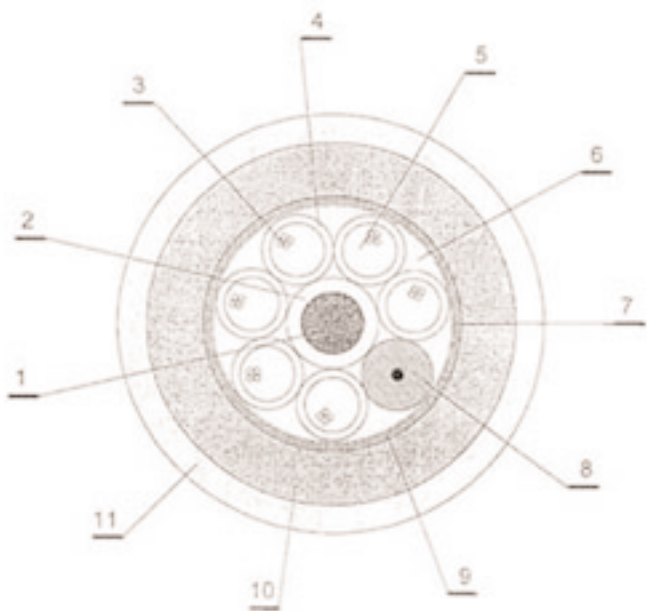
U tehnološkom smislu svi oblici prijenosa tradicionalnih oblika komunikacija (glas, slika ili razmjena podataka) konvergiraju prema jednom prijenosnom protokolu a to je Internet Protokol (IP). Istovremeno pojavljuje se izrazita konvergencija IP aplikacija i optičkih komunikacijskih tehnologija.

Institutsku računalnu mrežu projektirali smo na načelima sveobuhvatnih optičkih mreža (optical to desktop) i nove računalne paradigme iz koje proizlazi Grid infrastruktura. Kod projektiranja smo primijenili nove tehnologije i metodologije u provedbi s Institutskim znanstvenim i stručnim resursima, čime smo izvršili tehničku i financijsku optimalizaciju cijele investicije. U tom smislu mreža ima značajke koje će omogućiti dugoročne potrebe, bez obzira što se brzina protoka informacija po mrežama udvostručuje svakih devet mjeseci. Tvornica Elka je prema našoj konstrukciji tehnološki oblikovala, proizvodno realizirala i donirala 4772 m namjenskog (custom)

optičkog kabla kojim smo povezali 21 objekt unutar Instituta i tako stvorili kampus optičku okosnicu. Brzina prijenosa informacija u multi modnim (MM) optičkim vlaknima iznosi 1 Gb, a u postavljenim optičkim vodovima će se prijelaskom na jednomodnu (singl mode, SM) aktivnu opremu povećati do Tb u FDM modu prijenosa. Tako je na fizikalnoj razini infrastruktura postavljena za duže vrijeme.

Osim veće brzine prijenosa, postignuta je i paralelna redundancija (50 % više niti) kako bi se trajno zadovoljilo svim konfiguracijskim i topološkim zahtjevima koji mogu nastati kod budućih Grid i P2P (peer to peer) aplikacija.

Kod projektiranja smo primijenili nove tehnologije i metodologije u provedbi s Institutskim znanstvenim i stručnim resursima, te smo tako izvršili tehničku i financijsku optimalizaciju cijele investicije.



**Presjek hibridnog MM/SM optičkog kabla:
(namjenski razvijen u suradnji IRB - ELKA)**

1-nemetalna nosiva jezgra, 2-Pe plašt, 3-svjetlovodno vlakno, 4-cjevčica, 5-punilo cjevčice, 6-punilo jezgre kabla, 7-traka, 8-ispuna, 9-aramidna vlakna, 10-PE plašt, 11-PA plašt

Kod optičkog magistralnog razvoda osim navedenog tehničkog dobitka ostvarena je i značajna financijska ušteda temeljem doniranog optičkog kabla. Kod aktivne opreme Cisco System Inc je našem mrežnom projektu odobrio poseban akademski popust, u iznosu od 39 % što iznosi uštedu od oko 800 000 kuna.

Nova mrežna infrastruktura koja se implementira na Ruđeru, temelji se na AVVID arhitekturi (Architecture for Voice, Video and Integrated Data). Aktivni komunikacijski uređaj čini inteligentni preklopnik serije Catalyst 6500, dok su na rubovima preklopnici iz nove serije Catalyst 2950. On omogućuje prijelazak prema konvergentnoj IP mreži s jednim prijenosnim protokolom i s implementiranjem čitavog niza novih aplikacija. QoS (Quality of Service) mehanizmi, uz veliku propusnost cijelog sustava, osiguravaju određenu razinu servisa na komunikacijskoj mreži, kao zahtjev IP aplikacija. Takva je bazična infrastruktura nužna kod primjene svih mrežnih aplikacija, multimedijalnih mrežnih servisa (IP telefonije, Video konferencije, Video broadcasting, Video on demand), ali i integracije klasične telefonije (VoIP - Voice over IP), e-learninga i Linux/Microsoft Windows i Office paketa koji su standardi za operativni sustav i poslovnu aplikaciju na svakom računalu. Nova će mreža imati daleko veći infrastrukturni značaj od klasičnog poimanja mreže i mrežnih funkcionalnosti. Dugoročno će se riješiti integralna komunikacijska infrastruktura; od klasičnih primjena kao što je telefonija, do naprednih Grid sustava.



Upravljanje kampus LAN mrežom

Upravljanje kampus mrežom Instituta na aplikacijskoj razini osigurava se pomoću SNMP-a (Simple Network Management Protocola). On je dio TCP/IP-a (Transmission Control Protocol/Internet Protocola), a omogućuje administraciju mreže u smislu nadzora, upravljanja mrežnim resursima, značajkama i funkcionalnostima. Istovremeno vrši samoprilagodbene, statističke i obradne funkcije, pamti i knjiži mrežne upravljačke procedure, otkriva mrežne zastoje i optimalizira promet i plan za daljnji razvoj mreže. Nadgradnja programske podrške preko EMS-a (Element Manager Software) omogućuje uvođenje novih IP naprednih usluga s unapređenom razinom servisa. Na taj će se način u Računalnom centru uvesti čitav niz novih usluga i ekspertne podrške znanstvenicima.

Stvaranje virtualnih LAN-ova

Implementacijom virtualnih mreža VLAN (Virtual Local Area Network) novi će mrežni sustav u organizacijsko - logističkom pogledu imati velike mogućnosti. Programski se virtualna mreža postavlja iznad postojeće mreže i zajedno s njom čini dio mreže koji se može programirati kao da komuniciraju poput odvojenih mrežnih dijelova. Tako se umjesto fizičkog spajanja vrši logičko spajanje (preko mrežnog management programa). VLAN-ovi će se moći postaviti na razinu ustrojbenih jedinica (zavodi, laboratoriji, grupe...) ili poslovnih funkcionalnih cjelina (projekti, djelatnosti...) stvarajući logički povezane istaknute radne skupine (workgroups). Tako npr. ako se određene radne skupine nalaze unutar nekoliko lokacija, one će moći imati isti pristup svojim lokalnim resursima kao da se nalaze na istom mjestu. Dobiveni virtualni segmenti mreže imat će obilježja promjenljivosti, autonomnosti, funkcionalno-



snimio: Josip UhrI

Računalni centar je osposobljen za testiranje računalnih mreža najviše kategorije i standarda

sti i sigurnosti prema posebnim zahtjevima. Stvaranje i održavanje virtualnih mreža bit će nova usluga Računalnog centra. Prednosti mrežne arhitekture s virtualnim LAN-ovima se očituju u povećanju kakvoće usluga (Quality of Service), povećanju upravljivosti (manageability), optimalizaciji mrežnog prometa (network tuning), fleksibilnosti fizičke topologije i dodatnom povećanju lokalne sigurnosti.



Optička mreža je izvedena tehnologijom varenja optičkih niti

IP telefonija

Digitalna telefonija zamijenila je analognu zbog dodatnih funkcionalnosti koje nudi. Danas se u svijetu odvija migracija digitalne telefonije prema IP telefoniji upravo zbog niza još većih dodatnih mogućnosti. Primjena IP telefonije u velikim tvrtkama, povećava skalabilnost, raspoloživost i "inteligentnost sustava", dajući konvergentnu infrastrukturu koja omogućuje veću produktivnost u organizaciji. Osnovna prednost uvođenja IP telefonije sastoji se u tome što korisnici na nove lokacije ne moraju uvoditi telefonske centrale, nego koristeći vlastitu mrežnu infrastrukturu dobivaju sve telefonske servise. Pomoću IP telefonije povećava se fleksibilnost i brzina implementacije telefonije i novih usluga na svojim lokacijama. Najjednostavnije objašnjenje IP telefonije bilo bi da su to telefonski razgovori koji se uspostavljaju kroz IP mrežu uporabom IP mrežnog protokola. Ta definicija ne obuhvaća tehničko rješenje unutar kojeg postoji mnogo zahtjeva za kvalitetom usluga, kao i faze prelaska na IP telefoniju. Tehnološki gledano u procesu migracije postoji jedna međufaza, a to jest VoIP

nastavak na str. 14

piše: **Mario Stipčević,**
IRB EF/suradnik RC-a



Grubo
rečeno,
klaster
(grozd) je
skup računala
koja prema
konačnom
korisniku
djeluju kao
jedno.

Klaster računalo na IRB-u

Moderna znanstvena istraživanja sve više ovise o mogućnostima modeliranja i simuliranja složenih sistema i izvođenja opsežnih izračuna na snažnim računalima. Modeliranja atomskih jezgara u nuklearnoj fizici i modeliranje dugih isprepletenih molekularnih lanaca makrolidnih antibiotika, simulacije interakcija čestica u velikim detektorima eksperimentalne fizike čestica ili izračuni u Lattice QCD teoriji, sve su to samo neki primjeri gdje je računalo postalo jedno od osnovnih sredstava istraživanja, baš kao što su to u nekim drugim primjerima epruveta, elektronski mikroskop ili akcelerator.

Stoga ne čudi da i znanstvena istraživanja na Institutu Ruđer Bošković sve jače ovise o mogućnostima korištenja snažnih računala. Znanstveni rad određenog broja laboratorija i grupa već bitno ovisi o raspoloživoj procesorskoj snazi, u čemu za sada prednjače kvantni kemičari, fizičari i elektroničari. Taj novi trend ukazuje na rastuću potrebu za definiranjem adekvatnog odgovora.

Pilot projekt

U odboru za računala i komunikacije (ORK, <http://www.irb.hr/ork>) razmatrali smo to pitanje na nekoliko sjednica održanih u svibnju i lipnju 2001. i donijeli smo odluku da se ide na gradnju grozdastog super-računala koje bi se u pokusnoj (pilot) fazi sastojalo od pet dvo-procesorskih PC računala. Ukupni su troškovi procijenjeni na 100.000 kuna, a trajanje projekta na 6 mjeseci. Odlukom tadašnjeg ravnatelja M. Boranića odlučeno je da će Institut pokriti 60 posto troškova za izgradnju pilot klastera, dok će zainteresirane grupe (teme) podijeliti na jednake dijelove preostalih 40 posto troškova. U prvom momentu prijavilo se osamnaest zainteresiranih grupa, a do kraja pilot projekta prijavile su se ukupno 42 teme. Nakon zatvaranja polazne financijske konstrukcije, ORK mi je povjerio zadatak da okupim i vodim Grupu za izgradnju

klastera (GZIK), koja počela je s radom 15.06.2001., istom kada je isporučena i prva oprema za klaster, a kao takva završila je s radom 31.12.2001. Do tog trenutka grozd je predan na upravljanje Računalnom centru. Grupu ZIK sačinjavalo je 8 mladih znanstvenika s Instituta i dva člana RC-a: dr. T. Antičić (EF), dr. D. Babić (FK), dr. Z. Glasovac (OKB), N. Kmetić (RC), dr. B. Kovačević (OKB), dr. A. Maksimović (EL), N. Pavković (RC), mr. B. Perić (KM), dr. T. Šmuc (EL) i dr. M. Stipčević (EF). Veliku pomoć tijekom cijele pilot faze imali smo i od voditelja Računalnog centra dr. K. Skale. Valja istaći i posebnu podršku koju nam je dao Odbor za prirodne računarske znanosti Znanstvenog vijeća.

Što je grozdasto računalo?

Grubo rečeno, grozd (klaster) je skup računala koja prema konačnom korisniku djeluju kao jedno. Ideja je da korisnik ne treba znati koliko stvarno ima računala niti kako se zovu. Za njega cijeli grozd ima samo jedno ime i njegovo korisničko okruženje (vlastite datoteke, raspoloživi programi i programske biblioteke) invarijantno je na bilo kakve hardverske promjene u grozdu. U našoj koncepciji grozd se sastoji od jednog "frontalnog" računala na kojeg se korisnici mogu logirati (grozd.irb.hr) i "radnog" dijela kojeg sačinjavaju svi ostali kompjuteri u grozdu. Sistem rada je takav da korisnici svoje "poslove" šalju putem sistema PBS u red za izvršavanje. PBS uzima poslove iz reda jedan po jedan i smješta ih na slobodne procesore. Pri tom procesori se opterećuju s najviše jednim poslom istovremeno jer se to pokazalo optimalnim. Treba reći da PBS ne određuje redosljed poslova u redu čekanja niti taj redosljed apsolutno odgovara redosljedu pristizanja poslova u red. Naime MAUI "scheduler" se brine da redosljed izvršavanja poslova bude pravičan, tj. da promatrano u duljem vremenu svi korisnici ostvare odgovarajući udjel utrošenog CPU vremena, a to se postiže

korištenjem tzv. "fair share" formule. Za sada "odgovarajući" udjel znači "jednaki" jer su sve grupe dale istu svotu novca za gradnju klastera, no u budućnosti se to može mijenjati proporcionalno eventualnim dodatnim ulaganjima od strane pojedinih grupa. Korištenjem PBS-a i MAUI-a riješen jedan od glavnih prigovora na način rada prethodnih kompjutera u RC-u kada je bilo moguće da jedna grupa korisnika bombardiranjem stroja zahtjevima za izvršavanje poslova potpuno preuzme kontrolu nad računalom ili izazove njegovo zagušenje (a najčešće oboje !).

Daljnje financiranje grozda

U Hrvatskoj se malo što održava: obično se izgradi pa pusti da propadne. Takva je bila i sudbina svih akademskih klastera do danas u Hrvata. Mi smo htjeli pošto-poto izbjeći upadanje u istu zamku, a to je značilo postaviti korištenje grozda na zdrave ekonomske noge. Najlogičniji način jest da se naplaćuje utrošeno CPU vrijeme na klasteru. Cijena je određena tako da se uz broj raspoloživih sati godišnje i predviđeno iskorištenje radnog dijela od 90-ak posto dođe, uz određeno sudjelovanje Instituta, do sredstava dovoljnih za tehnološku obnovu grozda. Rast grozda, međutim, prepušten je dodatnim donacijama zainteresiranih grupa. Time smo htjeli otkloniti moguće spekulacije oko toga da se grozd na račun Instituta gradi za pojedine povlaštene grupe. Ova je strategija već polučila nevjerojatno dobre rezultate. U ožujku ove godine došlo je do ujedinjavanja sredstava i napora za nabavku klastera za grupu dr. S. Sabljica, tako da je sredstvima koja bi bila dovoljna za kupnju komercijalnog 12-procesorskog kompjutera nabavljeno čak dvostruko više računala, a ona su sklopljena u klaster u našem Računalnom centru. Nadalje, prof. Z. Maksić postigao je s MZT-om dogovor o sufinanciranju u jednakom omjeru dodatne opreme za grozd. Dogovor je vrijedan 160.000 kuna, odnosno otprilike 16 procesora frekvencije 1.8 GHz po sadašnjim cijenama. U ovom trenutku je već dovršen dio klastera koji se odnosi na dogovor sa dr. S. Sabljicom i sada grozd broji ukupno 34 procesora u radnom dijelu (8 x 1 GHz + 26 x 1.8 GHz). Kada se realizira i dogovor s ministarstvom, grozd će preći veličinu od 50 procesora što više i nije mali klaster. Ukupna vrijednost cijelog klastera time će se popesti na 650.000 kuna. Za usporedbu, prvi po veličini slijedeći akademski klaster nalazi se na SRCE-u i broji 11 procesora od 1 GHz. Nota bene, i taj je klaster izgrađen u bitnome snagama IRB-a, a projekt CERN-Datagrid u okviru kojeg je sagrađen vodio je s našeg Instituta, dr. Tome Antičić, član bivše grupe ZIK, a programsku implementaciju je izvršio Nikola Pavković. (Veličinu/snagu klastera bi zapravo trebalo izražavati u GHz. Tako naš klaster sada ima snagu od 55 GHz, a onaj u SRCE-u 11. Realan omjer snaga IRB/SRCE je 5.)

Perspektive

Kada smo krajem prošle godine privodili kraju pilot klaster, osobno nisam ni sanjao da će se stvari s njegovim daljnjim rastom tako "oteti kontroli" u pozitivnom smislu. Kad se samo sjetim da u ovo doba prije godinu dana za računanje nismo imali u RC-u ništa osim posrnulog faust-a koji je 3 puta sporiji od notebooka na kojem ovo pišem! Uspjeh gradnje klastera na IRB-u koji je rezultat

jedne šire i uglavnom dobro koordinirane inicijative većeg broja ljudi, dobra je podloga za daljnje posebne dogovore s Ministarstvom, nadajmo se promjenu njihova stava prema financiranju kompjuterske i mrežne opreme (što se već pomalo osjeća) kao i za otvaranje mogućnosti prijave raznih tehnoloških i informatičkih projekata. Tu svakako treba istaći i projekt CRO-GRID za uključenje u europski DATA GRID koji se koncipira kao nadgradnja klastera na IRB-u. Takvi projekti trebali bi postati značajnom sa stavnicom daljnjeg razvoja RC-a i privlačenja većeg broja mladih tehničkih kadrova od kojih bi one najbolje pokušali zadržati/zaposliti.

Zaključak

Ovaj je grozd prema mojem znanju prvi u Hrvatskoj kojeg je izvela i financirala jedna znanstvena institucija pa u tom smislu predstavlja pionirski iskorak. No da ne bismo upali u pretjeranu samohvalu treba reći da smo pri gradnji ovog



Pilot klaster u fazi rađanja

klastera uglavnom pokupili već postojeća rješenja i samo ih lukavo uklopili u jednu cjelinu kakva nama najbolje odgovara. Besprijekorno funkcioniranje grozda daje naslutiti da ovo rješenje predstavlja dobru osnovu za daljnje tehničko usavršavanje, pa ako se sam klaster i tehničke aktivnosti vezane za njega nastave dalje razvijati, vjerojatno ćemo se jednog dana moći pohvaliti nekim originalnim doprinosom našeg Instituta i na tom području.

Kao zaključak želim istaći da je ovaj klaster nastao kao rezultat potrebe za znanstvenim računanjem na Institutu i da će on postojati i rasti samo u skladu s tim potrebama. Da nije tako zar ne bi bilo logično da se ovaj klaster nalazi recimo na FER-u, umjesto da sada kvalitetni FER-ovci grade i održavaju klaster na Ruđeru? Razmišljajmo o tome i njegujemo ono iz čega izvire naša snaga - kvalitetnu i suvremenu znanost.

Dodatne informacije o projektu i samom grozdu izvolite naći na <http://grozd.irb.hr>.

piše: **Domagoj Krišto,**
IRB RC



Iskustva i mogućnosti u optimizaciji rada na grozdu

Ovaj članak opisuje batch-sustav grozdastog računalnog sustava i iznosi statističke podatke o korištenju resursa grozdastog računalnog sustava (kraće grozd). Isto tako, ovaj članak daje savjete

za korištenje grozda čije poštivanje će dovesti do boljeg iskorištenja njegovih resursa, a samim time i do ubrzanja izvršavanja poslova na njemu. Na kraju se nalazi opis sustava za udaljeno praćenje trenutne iskorištenosti grozda.

Sustav za upravljanje poslovima

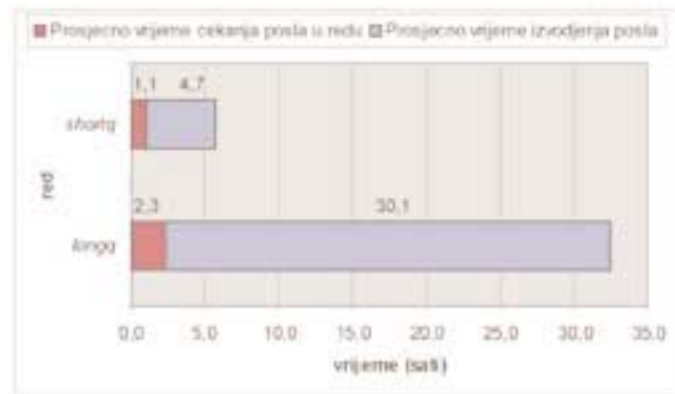
Sustav za upravljanje poslovima se na engleskom jeziku zove batch system. Kako u hrvatskom jeziku još nema naziva za taj sustav, u ovom tekstu će se koristiti naziv batch-sustav. Dakle, batch-sustav je sustav koji zaprima zahtjeve korisnika za pokretanjem poslova (engl. job). Posao se, zatim, stavlja u red (engl. queue) za izvršenje u kojem čeka da se izvrše poslovi koji su u red došli prije njega. Kada posao dođe na prvo mjesto u redu (pred početak izvršavanja), onda mora čekati sve dok se ne oslobode potrebni računalni resursi za izvođenje tog posla. U posebnim slučajevima, posao može biti izvršen i preko reda. Tako nešto se može dogoditi ako za poslove koji se nalaze ispred u redu nema dovoljno računalnih resursa i ne će ih biti u vremenu koje je duže od vremena izvršenja posla za koji ima dovoljno resursa. Takav postupak nije nepošten prema poslovima koji su u red ušli prije, zato što se oni zbog nedostatka resursa ionako nisu mogli izvršiti. Dapače, takav postupak povećava iskorištenost grozda.

Kada grozd ne bi imao batch-sustav, poslovi bi se pokretali ne koordinirano, bez ikakvog reda što bi dovelo do preopterećenja čitavog sustava, bila bi smanjena učinkovitost sustava i bila bi vrlo učestala rušenja operacijskog sustava. Dakle, prvenstvena uloga batch-sustava je organiziranje izvršavanja poslova na grozdu na način da njegovo funkcioniranje ne bude ugroženo. Takvim sustavom se, također, omogućava izbor procesora na kojemu se posao želi izvršiti.

Još jedna važna uloga batch-sustava je određivanje prioriteta izvođenja poslova. Naime, kada sustav koristi više korisnika s jednakim pravima, ali s nejednakim zahtjevima, potrebno je uspostaviti sustav prioriteta koji će povremenim korisnicima i onim korisnicima koji imaju manje zahtjeve za resursima omogućiti brži dolazak na red za izvršavanje poslova. Pri izboru posla za izvršavanje, batch-sustav može uzimati u obzir nedavnu potrošnju svih korisnika čiji poslovi čekaju i izabrati posao korisnika koji je utrošio najmanje računalnog vremena. Različiti batch-sustavi imaju i znatno složenije sustave za

određivanje prioriteta koji uzimaju u obzir i niz drugih faktora. Grozd prvenstveno služi za izvršavanje velikih poslova koji se u pravilu izvršavaju nekoliko dana i traže puno računalnih resursa. No na takvom sustavu izvršavaju se i manji poslovi kojima izvršavanje traje znatno manje i iziskuje manje računalnih resursa. Zato, na grozdu koji je razvijen na Institutu Ruđer Bošković, postoje dva reda za čekanje na izvršavanje. Prvi, longq, služi za čekanje na izvršavanje poslova koji se dugo izvršavaju (nekoliko dana) i koji zahtijevaju puno računalnih resursa. U takvom redu poslovi koji vremenski traju kratko (nekoliko sati) mogu čekati i znatno duže od svog vremena izvršavanja. Zato, da bi se omogućila pravednija raspodjela računalnih resursa i vremena korištenja sustava, postoji i red shortq koji služi za izvršavanje poslova koji vremenski traju kratko.

Poslovi se pokreću tako da postoje skripta u koja se napišu zahtjevi za resursima i naredbe za pokretanje poslova. To mogu biti nazivi procesora na kojima se želi izvršiti posao, potrebna memorija, potrebno vrijeme za izvršenje, red za izvršavanje i sl. Ako se ne zada neki od podataka, onda se uzima njegova podrazumijevana vrijednost. Podrazumijevani red za izvršavanje posla je longq. Za potrebnu memoriju u redu longq podrazumijevana vrijednost je 400 MB, a u redu shortq je 100 MB. Vremensko ograničenje za izvršavanje posla u redu longq je 168



slika 1. Vrijeme čekanja posla u redu i izvođenja posla

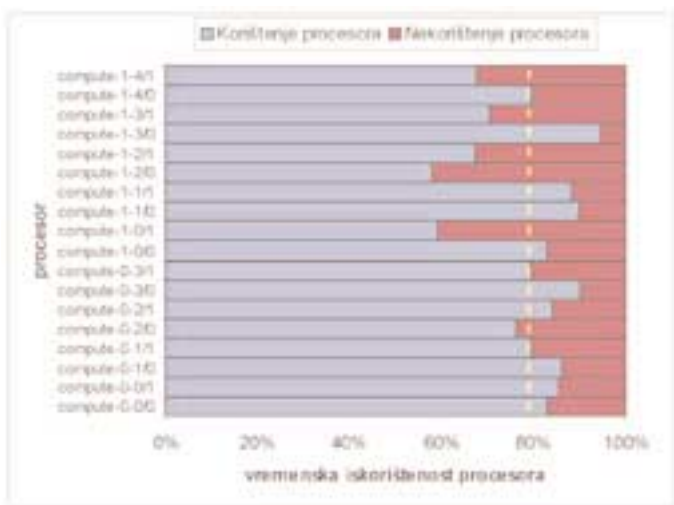
sati (7 dana) i to je podrazumijevana vrijednost. Za red shortq to je 8 sati. Ukoliko se posao ne izvede u vremenu u koje je određeno za izvršenje, onda posao biva prisilno prekinut. Isto se događa i s poslovima koji pokušaju trošiti više resursa no što im je dodijeljeno na korištenje.

Podaci o korištenju grozda

Podaci koji su prikazani u ovom poglavlju odnose se na poslove koji su završili svoj rad tijekom lipnja 2002. Slika 1. prikazuje prosječno vrijeme čekanja poslova po redovima i prosječno vri-

jeme izvođenje poslova po redovima. Vrijeme čekanja u redu shortq je 1,1 sat. To se možda ne čini puno, ali u usporedbi s vremenom izvođenja, to je puno (23,00 %). U redu longq vrijeme čekanja je duže (2,3 sata), ali u usporedbi s vremenom izvršavanja to je relativno malo (7,79 %).

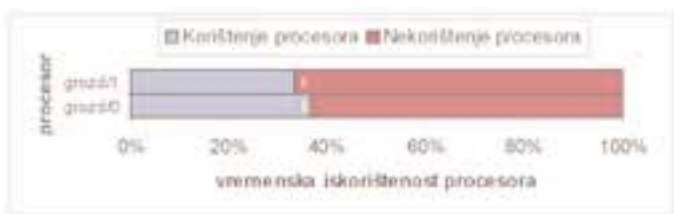
Slika 2. prikazuje vremensku iskorištenost procesora radnih čvorova grozda. Žuta crta predstavlja prosječnu vremensku iskorištenost procesora koja iznosi 79,0 %. Slika 3. prikazuje vremensku iskorištenost procesora glavnog (frontalnog) čvora. Njihova prosječna vremenska iskorištenost je 35,0 %. Iskorištenost procesora glavnog čvora je vrlo mala. Naime, to je čvor koji upravlja sustavom i poslovi koji se izvršavaju na njegovim procesorima imaju najmanji prioritet izvršavanja. Zato se na tome čvoru izvršavaju poslovi iz reda shortq. To djelomično



slika 2. Vremenska iskorištenost procesora radnih čvorova

objašnjava i zašto je vrijeme čekanja poslova u tom redu dugo u odnosu na vrijeme izvršavanja.

Slika 4. prikazuje koliko se za poslove zauzima memorije i koliko se te memorije stvarno koristi. Sustav koji se brine da poslovi koji koriste više memorije no što su dobili na korište-

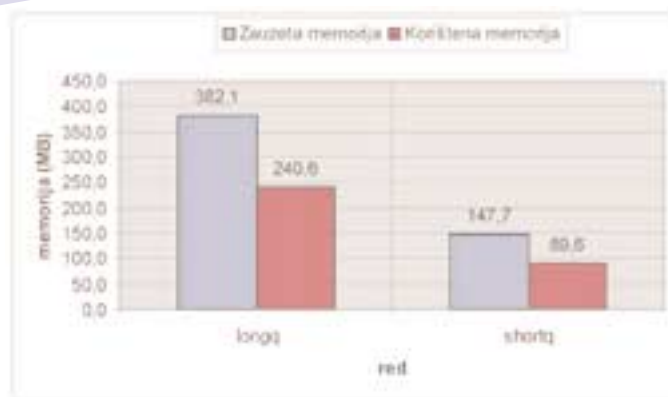


slika 3. Vremenska iskorištenost procesora glavnog čvora

nje, budu prisilno prekinuti ima pogrešku i dopušta daljnje izvršavanje čak i ako posao koristi više memorije no što je dobio na korištenje. Slika 5. prikazuje koliko se zauzima memorije i koliko se memorije stvarno koristi za poslove koji nisu koristili više memorije no što im je dano na korištenje. Postotak korištenja dodijeljene memorije u redu longq je 50,1 %, a u redu shortq je 44,4 %.

Savjeti korisnicima za bolje iskorištenje sustava

Iz prikazanih statističkih podataka se može vidjeti da iskorištenost sustava nije najbolja. Glavni razlog je što korisnici



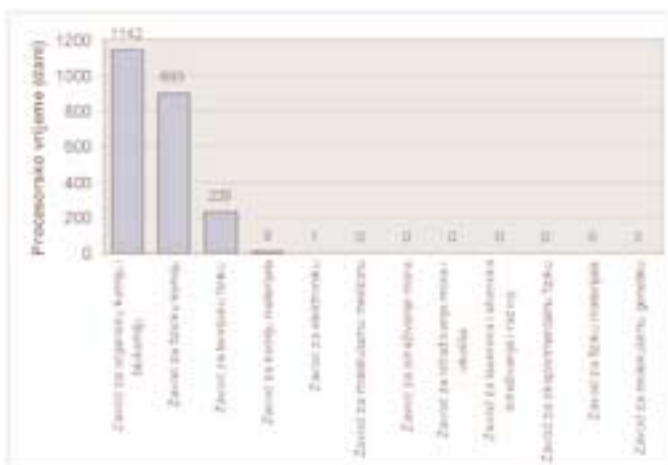
slika 4. Zauzetost memorije

postavljaju prevelike zahtjeve za resursima prilikom pokretanja poslova na grozdu. Korisnici to rade iz više razloga. Prvi je bojazan da će njihov posao biti prekinut ako pokuša iskoristiti više resursa no što mu je dodijeljeno. Drugi je da korisnici



slika 5. Zauzetost memorije (bez poslova koji su koristili više memorije no što su je dobili na korištenje)

uopće ne postavljaju zahtjeve pa se oni postavljaju na podrazumijevanu vrijednost koja nije mala zato da se poslovi ipak ne bi morali prisilno prekidati. I treći razlog je da korisnici često ne



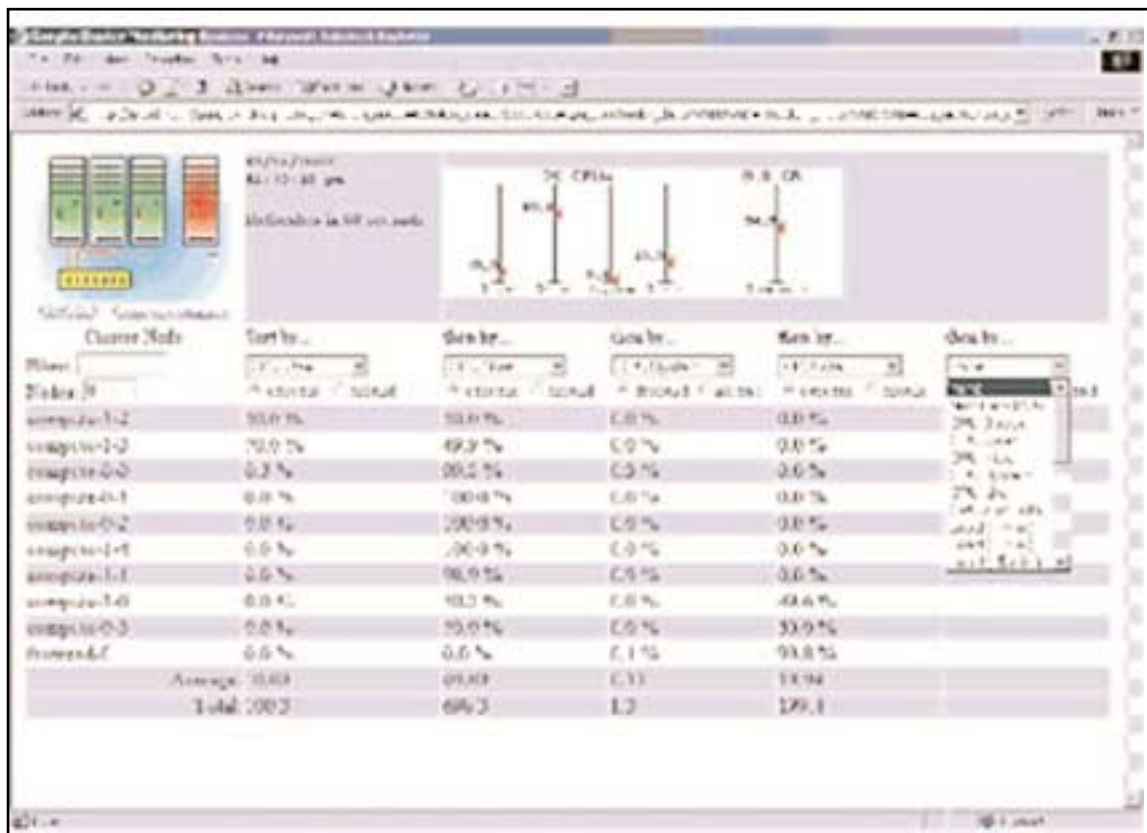
slika 6. Potrošnja procesorskog vremena po zavodima

poznaju dovoljno svoju aplikaciju i teško im je procijeniti koliko će biti potrebno resursa za izvođenje njezinih poslova. Da bi se poboljšala iskorištenost sustava i skratilo vrijeme čekanja u redu za čekanje na izvršavanje, korisnici bi trebali bolje upoznati svoje aplikacije i postavljati zahtjeve za resursima u skladu s potrebama njihovih aplikacija. No to nije uvijek lako. Zato korisnici mogu za probu pokrenuti izvođenje posla s većim zahtjevima za resursima i nakon što se posao izvede trebaju pogledati koliko je resursa posao stvarno koristio. Svaki sljedeći put kada se taj posao pokreće, od sustava treba zahtijevati onoliko resursa koliko je posao potrošio u probnom radu. Da se ne bi dogodilo da posao pokuša koristiti više resursa, dobro je zahtijevati malo više resursa od onih koje je posao potrošio u probnom radu (recimo 5 do 6 posto).

Udaljeno praćenje grozda

Prilikom izgradnje računalnog grozda primijetilo se da korisnici nemaju jednostavan način za praćenje njegove trenutne iskorištenosti. Podaci o trenutnoj iskorištenosti sustava mogu se pronaći i na samome grozdu, ali većini korisnika takav način je traženja podataka nije jednostavan.

Zato je razvijen sustav za udaljeno praćenje trenutne iskorištenosti grozda preko Interneta pomoću nekog od mrežnih pretraživača (Internet Explorer, Netscape Communicator, Opera i dr.). Ovaj sustav se nalazi na internetskoj adresi <http://grozd.irb.hr/ganglia/index.php>. Slika 7. prikazuje taj sustav. U gornjem dijelu se vide grafovi koji prikazuju neke



slika 7. Programski sustav za udaljeno praćenje trenutne iskorištenosti grozda

Nije dobro tražiti više procesora za izvođenje posla koji nije napravljen za višeprocorski rad jer će se tada koristiti samo jedan procesor. U tom slučaju se može povećati čekanje u redu za čekanje na izvođenje jer se čeka da svi "potrebni" procesori budu dostupni. Isto tako, procesori koji se ne koriste možda trebaju nekom drugom poslu, a samo vrijeme izvođenja procesora se ne će povećati.

Slika 6. prikazuje potrošnju procesorskog vremena po zavodima (vremena su izražena u danima i mjerena su od kraja listopada 2001. do kraja lipnja 2002.). Vidi se da su grozd najviše koristili Zavod za organsku kemiju, Zavod za fizičku kemiju i Zavod za teorijsku fiziku. Osim tih zavoda grozd su koristili i Zavod za kemiju materijala i Zavod za elektroniku. Ostali zavodi do sada nisu koristili grozd. Vjerojatno će djelatnici tih zavoda početi koristiti grozd kada uvide sve mogućnosti koje ima jedan tako moćan sustav.

memoriji, prosječnom broju poslova koji su se izvršavali u zadnjih nekoliko minuta (opterećenost), ukupnom broju poslova itd.

Zaključak

Ovaj članak daje pregled dosadašnjeg korištenja grozdastog računalnog grozda. U članku su dani savjeti za još bolje iskorištenje grozda. Njihovo poštivanje će biti od posebne važnosti kada sustav bude koristio još veći broj korisnika. Kako broj korisnika bude rastao, bez obzira na poboljšano iskorištenje grozda, bit će nužno dodati nove čvorove u grozd i tako dodatno povećati moć sustava.

Daljnji razvoj grozda se odvija u smjeru proširenja, optimizacije rada i tehničkog usavršavanja u sklopu odobrenog informatičkog projekta. Taj projekt predviđa razvoj monitoringa i upravljanja složenog superračunalnog sustava.

e-Science

(Znanstveni rad na GRID platformi)

piše: Karolj Skala,
IRB ZE/RC



Rast računarstva

Tehnološki rast zorno je prikazan kao vrijeme unutar kojeg se brzina ili kapacitet udvostručuje. To je u slobodnim tržišnim okolnostima ekvivalentno vremenskom periodu unutar kojega se cijena prepolovljuje. Generička informatička tehnologija koja se intenzivno razvija, mjeri to vrijeme u mjesecima. Brzina rada računalnih mreža se udvostručuje svakih 9 mjeseci, kapacitet se memorijskih jedinica u 12 mjeseci udvostručuje, a računalna je moć (snaga) svakih 18 mjeseci dvostruko veća, što znači da je godišnji faktor rasta 1.5. To čini eksponencijalni rast s različitim vremenskim konstantama koje proizlazi iz Mooreova zakona a on kazuje da se svakih 18 mjeseci broj integriranih tranzistora u procesorima udvostručuje. Taj se zakon temelji na činjenici da je od početka razvoja procesora (1975. godine, Intel 8080) broj tranzistora od 4 500 narastao na sadašnjih 20 milijuna (Pentium IV procesor). Temeljem spomenutih zakonitosti može se predvidjeti da će za slijednih pet godina brzina mreže narasti na 10 Gb/s, a PC računala će imati snagu veću od 32 GIPS (Giga Instructions Per Second), s memorijom od 16 GB i diskovima od 2 TB. Superračunala će biti paralelni strojevi od 10 000 procesora. Dvostruko veća brzina rasta mrežnih komunikacija u odnosu na računala dovodi do novih načela, metoda i procesa na globalnoj razini u cjelosti i implicira spajanje resursa i ljudi u novo suradničko okruženje. Taj rast neminovno otvara mogućnost i nužnost sveobuhvatne mrežnoracionalne resursne integracije koji zovemo Grid sustavom.

Konstatacija stanja

Zahtjevne potrebe mnogih znanstvenih i tehničkih istraživanja u pogledu računarske obrade i izračuna sa složenim algoritima je nemoguće provesti na PC-ima, radnim stanicama ili brzim server računalima u smislenom vremenskom trajanju unatoč snažnom razvoju. To proizlazi iz sve kompleksnijih analiza i struktura složenih procesa koje se mogu pratiti, obrađivati i razvijati jedino računalno i podatkovno intenzivnim obradama. Bez obzira na eksponencijalni rast računalne tehnologije, potrebe su veće i one se zadovoljavaju novim paralelnim i distribuiranim načinom spajanja računala. To dovodi do spajanja više procesora u paralelan rad (tako nastaju superračunala), koja imaju veliku ulogu kod ubrzanog razvoja znanosti i tehnologije uz primjenu analitičkih i simulacijskih metoda, te oblikovanja i vizualizacije složenih modela u raznim područjima. Zahtjevnost i mogućnost računalne moći i memorijskog kapaciteta neprekidno rastu, pa današnja superračunala od nekoliko GIPS-a samo djelomično zadovoljavaju računarne potrebe znanstvenih istraživanja.

Tendencija razvoja

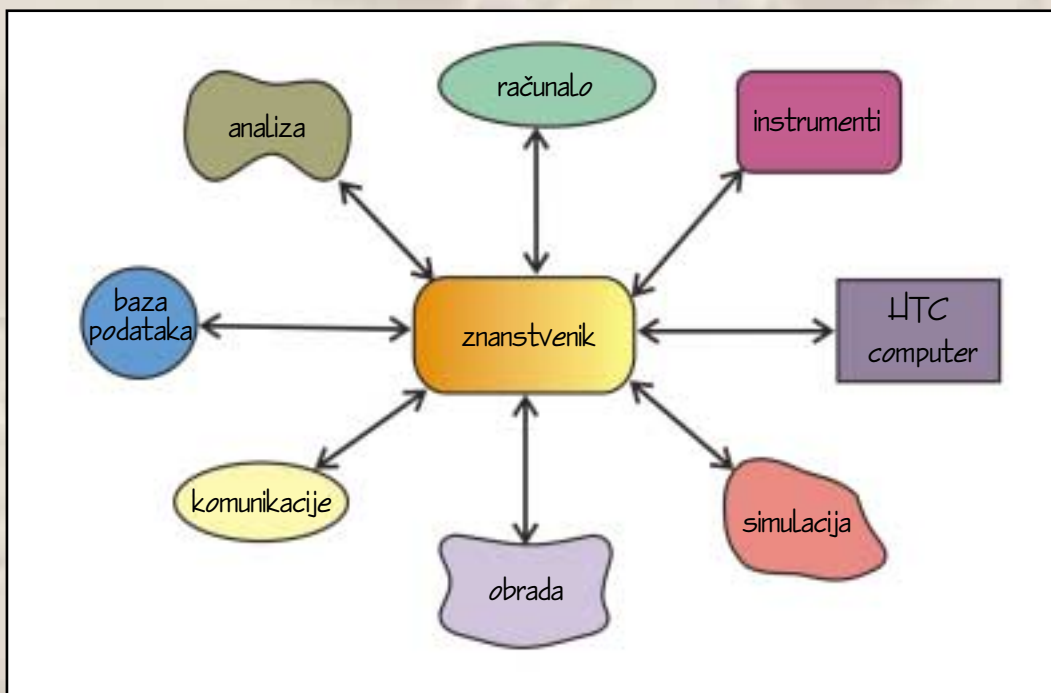
Radi rješavanja navedenih problema američki su znanstvenici 1995. godine u sklopu projekta I - WAY započeli razvoj mrežnog povezivanja superračunala, što je rezultiralo World Wide Grid-om (WWG) ili Grid sustavom¹. Na prekretnici tisućljeća u cijelom se svijetu ulažu značajna sredstva u izgradnju nove grid paradigme koja ima posebno značenje u razvoju znanstvenog rada. To je nova računalnomrežna tehnologija koja ne prenosi samo podatke i informacije, nego dijeli resurse pomoću brze mreže. Paralelni razvoj gigabitne (Gb) mreže i GHz računala stvorio je uvjete za uspostavljanje mrežno distribuiranog računarstva. Otvara se mogućnost nove generacije računarstva čija je moć TIPS-a (Tera Instruction Per Second) s petabajtnim podacima u mrežnom spoju iznad Gb brzine. Grid projekti su uglavnom financirani iz proračunskih sredstava jer predstavljaju generičnu infrastrukturnu tehnologiju.² Zahvaljujući sukladnom preklapanju tehnoloških trendova i znanstvenorazvojnih stremljenja i potreba, razvija se i uspostavlja Grid sustav, koji stvara novu međunarodnu znanstvenu infrastrukturu s tehnologijom koja označava strategiju razvoja znanosti u 21. stoljeću. Značajne su znanstvene sredine prihvatile Grid tehnologiju kao osnovu za budući znanstveni rad perspektivnog opstanka i prosperiteta.

Grid kao mreža resursa

Grid nije alternativa nego sljedeća generacija Interneta. On je set dodatnih protokola i usluga izgrađen uz protokole i usluge Interneta s ciljem da pruži podršku u izradi i upotrebi komunikacijski, informacijski i infrastrukturno obogaćenom okruženju. Grid je mrežni izvor slobodnih i raspoloživih resursa. To nije sustav koji omogućava neograničen pristup resursima, on naprotiv omogućava kontrolirano dijeljenje i uporabu resursa. Vlasnici resursa tržišno će valorizirati i ograničiti pristup korisnicima, pa zato Grid arhitektura mora sadržavati protokole za kontrolu razmjene informacija i resursa. Grid zahtjeva distributivni operativni sustav koji nazivamo posrednikom (middleware), a on čini programski međusloj između korisnika i resursa. Posrednik bi trebao omogućiti registraciju, autentikaciju i autorizaciju između korisnika aplikacije. Grid programska podrška se mora instalirati na svaki sustav koji u

¹ Koncept podjele distribucije resursa potiče još od 1965. U MIT-u je Fernando Corbató osmislio Multics operacijski sustav koji je težio da informatički resurs postane sličan energetskom i ostalim resursima. U nastavku, 1968. Licklider i Taylor anticipiraju Grid viziju u radu *The Computer as a Communications Device*.

² Projekti se mogu pogledati na adresi: <http://www.mcs.anl.gov/~foster/grid-projects>.



Slika 1: Klasičan model rada znanstvenika (Client-Server model)

Gridu sudjeluje. Zahvaljujući posredniku korisnik ne bi trebao znati gdje se nalaze resursi ili usluge koje koristi, već samo koliko to košta (u slučaju naplaćivanja). Posrednik bi trebao omogućiti praćenje i evidentiranje korisnikovih aktivnosti radi naplaćivanja usluga. Osim toga, praćenje aktivnosti na mreži trebalo bi iskoristiti i za povećanje sigurnosti sustava. U sklopu posrednika treba biti i mehanizam očuvanja privatnosti koji bi brinuo o zaštiti podataka o uslugama i korisnicima.

Zbog velike fizičke i administrativne heterogenosti Grida tradicionalne su transparentije nedostižne, a čini se da je moguće postići dogovor vezan za standarde i protokole. Predložena arhitektura je zato otvorene prirode, definira kompaktni i minimalni set protokola potrebnih izvoru resursa da bi se nalazio na Gridu. Osim toga, pruža okvir unutar kojeg je moguće specificirati više načina ponašanja. Grid kao novi sustav zahtjeva nove modele programiranja. Programiranje unutar Grida predstavlja izazov kojeg nema u radu s jednim ili više paralelnih računala, kao što je npr. višestruka administrativna domena. Kao i u drugim sustavima, apstraktnost i sažimanje mogu reducirati kompleksnost i poboljšati pouzdanost, pa je poželjno da aplikacija ima niz apstrakcija viših nivoa. Stotine, tisuće ili čak milijun procesora povezanih u raspodijeljenu računalnu arhitekturu, bit će dostupni putem virtualne strukture. Oni će predstavljati značajan izvor računalnokomunikacijske snage (moći), ukoliko će biti jednostavni za korištenje.

e-Science tehnologija

Odvijanje znanstvenog rada u klasičnom smislu podrazumijeva da znanstvenik mora osigurati ljudske i materijalne resurse, eksperimentalnu bazu i ostale preduvjete, slika 1.

Svaki znanstvenik osmišljava i osigurava oskudne resurse s nedorečenom programskom podrškom ili onom koja nije optimalna. Ovaj model u provedbi nema elemente kompatibilizacije, optimalizacije i mrežne uklopljenosti u smislu moguće mrežne integracije rada.

Uvođenje Grid tehnologije stvara se računalna moć, tj. optimalizacija resursa i primjena novih metoda i modela u znanstvenom radu. Kako bi znanstvenici lakše pristupili resur-

sima, aplikacijama i rezultatima, razvija se Grid portal kao korisničko sučelje koje se nadovezuje na sučelje posrednika. Preko Grid portala znanstvenik može pretraživati i pregledavati resurse, pronaći nove aplikacije, pokretati vlastite aplikacije, pretraživati baze podataka i znanje na Gridu. U napredne Grid aplikacije podrazumijeva povezivanje složene i skupe znanstvene instrumentacije preko Grida, čime se omogućava rad u okruženju virtualnog laboratorija (e-Lab). Skupini znanstvenika širom svijeta omogućava se pristup i skupni eksperimentalni rad na skupim i ekskluzivnim znanstvenim instrumentalnim resursima, izgrađuje se suradnja na rješavanju znanstvenih problema na najbolji način. Znanstvenici mogu dijeliti instrumente, koristiti resurse Grida, generirati i uspoređivati rezultate i biti neprestano u suradnji s obzirom da se koncept e-Laba zasniva na neovisnosti o geografskoj lokaciji. Isto tako, znanstvenici mogu pokrenuti određeno mjerenje distribuirano na jednoj lokaciji, a obradu vršiti na drugoj. Osnovni resursi Grida su istovremeno i resursi virtualnog laboratorija.

Procesni resursi:

- računalni poslužitelji za izvršavanje aplikacija izravno ili preko portala Grida,
- sophisticirani znanstveni instrumentarij povezan u Gridu (distribuirana osjetila, astronomski teleskop, satelitski akvizicijski sustavi, elektronski mikroskop, NMR...),
- mrežna infrastruktura za Grid potrebe (propusnost, konfigurabilnost, latencija..).

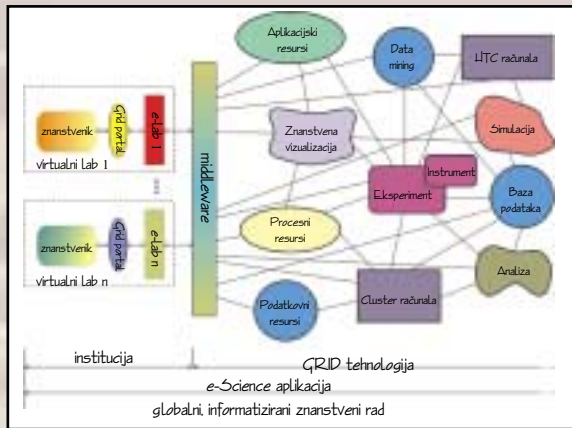
Data resursi:

- raspodijeljene baze podataka,
- baze znanja.

Aplikacijski resursi:

- multimedijska suradnička podrška (slika, video, zvuk),
- napredna aplikacijska programska podrška (data mining, knowledge management ..).

Primjenom Grid platforme u znanstvenom radu ostvaruje se e-znanost (e-Science) koja omogućuje cjeloviti (end to end) proces od e-Laba do publikacije. Ona stvara visok stupanj informatizacije u znanstvenom radu. Podržava obradu i znanstvenu vizualizaciju različitih tipova podataka i baza. Ostvaruje nove metode u znanstvenom radu na globalnoj razini, slika 2.



Slika 2: e-Science - znanstveni rad na Grid platformi.

(Povećana slika nalazi se na stražnjoj stranici korica).

ili tekućih eksperimentalnih procesa, tako i događanja. Razvija se sustavna obrada distribuiranih podataka u virtualnom globalnom memorijskom prostoru.

Baze podataka i znanja

Znanstveni problemi zahtijevaju analizu ili sintezu velikih skupova podataka. Spajanjem mnogih stručno-tematski usmjerenih baza podataka stvaraju se metabaze. Otvara se mogućnost inteligentnog pretraživanja, ekstrakcije i obrade na velikim nizovima podataka. Razvija se izuzetno važno područje pretrage i otkrivanja podataka i znanja (Grid data mining).

Napredna računalna instrumentacija

Sofisticirani znanstveni instrumenti temelje se na računarskoj podršci s mrežnim sučeljem. Ti uređaji služe kao generatori ogromne količine nizova podataka koji se analiziraju u odgođenom ili kvazi realnom vremenu. Time se za vrijeme odvijanja eksperimenta omogućuje inteligentno djelovanje na tijek mjernog postupka.

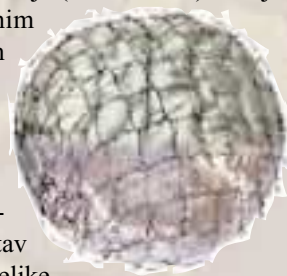
e-Lab (virtualni laboratoriji)

Virtualni laboratoriji su okruženja koja osim razmjene i obrade podataka, omogućavaju suradnju na daljinu u realnom vremenu. Formalizacija problema i rješavanje na daljinu, obrada i konzultacija vode do virtualne kompresije prostora. Omogućavaju formiranje distribuiranih članova projektnih timova na globalnoj razini. Sofisticirana 3D znanstvena vizualizacija omogućit će praćenje i intelektualno djelovanje na radne i stvaralačke procese u realnom vremenu. U oskudnim materijalnim uvjetima rješenje se nazire u globaliziranim projektima koji se mogu ostvariti preko distribuiranih virtualnih laboratorija primjenom daljinske suradnje. Treba stvoriti preduvjete da znanstvenici nakon prve faze informatizacije (računala, operacijski sustavi, lokalna i globalna mreža) imaju uvjete za daljinsku znanstvenoistraživačku suradnju u realnom i odgođenom vremenu. Treba omogućiti virtualnu kompresiju prostora u globaliziranom, znanstvenoistraživačkom svijetu. Na taj bi se način usporio proces odljeva mozgova i bili bi osigurani uvjeti za učinkovitu oplodnju intelektualnog kapitala, zadržavanjem znanstvenika u Hrvatskoj s aktivnim znanstvenim djelovanjem uz pomoć Grida i e-Science aplikacije.

Ne mogu se predvidjeti svi oblici ove nove generične promjene. Ipak, već sad je jasno da će nova informatička i računalna tehnologija u znanstvenom radu, gospodarstvu, pa tako i u cijelom nacionalnom i globalnom društvu imati ogroman značaj. U tim okolnostima treba sagledati utjecaj (impact) nove tehnologije na narode različitih statusa i mogućnosti. U tom pogledu prepoznaje se povoljnost u nailazećim tehnologijama, u integraciji intelektualnog kapitala za budući razvoj Hrvatske.

Utjecaj na gospodarski razvoj

Grid tehnologija je izuzetno zanimljiva i za komercijalnu uporabu. Rast elektroničkog poslovanja (e-Businessa) zahtijeva novu operabilnost na heterogenim platformama i raznim mrežnim okruženjima. Visoka kakvoća servisa u dinamičkim virtualnim gospodarskim sredinama jednako je važna kao u znanosti i inženjerskom razvoju. Gospodarski subjekti počinju uvoditi Grid sustav lokalno unutar korporacije, radi velike



Učinak e-Sciencea

Učinak grida u znanstvenom radu očituje se u distribuiranom računarstvu. Izgradnjom povezane farme clustera omogućuje se intenzivno računanje i stvaranje velike baze podataka, te se ostvaruje individualno i kolektivno napredovanje u znanju.

Istaknuti znanstveni rezultati se danas ostvaruju primjenom superračunala (klastera), a sutra će to biti primjena Grid sustava i e-Science aplikacije.

Uvodi se 3C (Computing Communication Collaboration) integracija, što dovodi do tehnologije teleprisutnosti (teleimmersion). To je radni sustav u obliku virtualne realnosti (e-Lab) u okolnostima Grid infrastrukture. On predstavlja kompleksnu simbiozu tehnologije i aplikacije iz koje proizlazi infrastrukturna suradnja na daljinu (distribuirani design, nova znanstvena vizualizacija, suradničko inženjerstvo itd.). Znanstveni instituti postat će generatori podataka iz inteligentnih znanstvenih instrumenata, nizova sofisticiranih senzora, robotiziranih eksperimentalnih aktuatora itd., iz kojih će se na globalnoj razini stvarati znanje unutar distribuiranih grid čvorova. Nastajat će metabaze podataka i znanja te će oživjeti globalna funkcionalna integracija u znanosti i znanstvenom radu. Grid pruža nove metode i modele rada u znanstvenoistraživačkom radu, te omogućuje napredne e-Science aplikacije.

Znanstveni Grid portali

To su web okruženja s dohvatljivim metodama i alatima za rješavanje problema iz određenog područja znanosti. Tu se nalaze programi i dokumentacije koje se mogu pokretati na udaljenim strojevima unutar Grid sustava. Portali se razvijaju po osnovnim tematskim područjima kao npr. za molekularnu biologiju, astronomiju, kemiju i ostala polja znanosti.

Superračunalni resursi

Zahvaljujući Grid sustavu radne stanice, klasteri i superračunala će biti dostupna znanstvenicima za obavljanje poslova u znanstvenom radu. Postoji veliki računalni resurs (više desetina tisuća računala) koji svakim danom raste i kojeg se po potrebi može koristiti za zahtjevne znanstvene izračune.

Memorijski prostori

Razvoj distribuiranog višepetabajtnog globalnog memorijskog prostora za znanost ima veliki značaj, kako za podatke minulih

uštede na račun optimalnog korištenja postojećih resursa. Uz pomoć internog Grida Intel Corp je uštedio 500 milijuna \$, zato se javlja veliki interes i težnja prema konvergenciji i integraciji Grid tehnologije s prethodno pokrenutim razvojem Web servisa (Web services). Web servis ne podrazumijeva web site, pre-

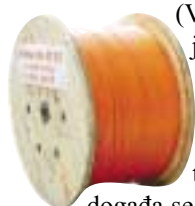
*U globaliziranom,
znanstvenoistraživačkom svijetu, treba
omogućiti virtualnu kompresiju prostora.
Na taj bi se način usporio proces odljeva
mozgova i bili bi osigurani uvjeti za
učinkovitu oplodnju intelektualnog
kapitala, zadržavanjem znanstvenika u
Hrvatskoj s aktivnim znanstvenim
djelovanjem pomoću Grida i e-Science
aplikacije.*

traživače (browsere) i protokole, već sučelje radi komuniciranja daljinskim procesima, što predstavlja novi Internet servis. Web servis postavlja programe za daljinske servise na webu, odnosno pokušava preko Interneta približiti poduzeća i kupce koristeći posebne programske podrške i standardne protokole;

XML, SOAP, WDSL i UDDI za opis, identifikaciju i komunikaciju podataka webom. Mnoge komercijalne tvrtke (IBM, Microsoft, NET, WebSphere i Sun) rade intenzivan razvoj na Web servisu. Istovremeno, tržišno napredne tvrtke počele su investirati i razvijati Grid aplikacije. Glavna interesna borba se prelama na činjenici što su Web servisi na Windows platformi, dok je Grid na Linux open source platformi. Veliki korak u tom smjeru je postignut suradnjom Globus tima i IBM-a na stvaranju OGSA (Open Grid Service Architecture), s ciljem da postave standarde spajanjem prednosti Grid sustava i Web servisa. Ovaj razvoj je koordiniran preko Global Grid Forum koji stvara standard za OGSA arhitekturu. Za očekivati je da će se ta konvergencija završiti u stapanju i nadvladavanju open source koncepcije, kao zdrave globalne programske orijentacije. Zato je orijentacija na Linux neupitno dobar izbor.

Perspektiva razvoja

Posebno treba istaknuti korisnost CRO GRID projektne inicijative Akademije tehničkih znanosti Hrvatske, a na poticaj Instituta Ruđer Bošković u pogledu akceleriranja i udomljenja svjetskih generičnih tehnologija. Važnost projekta ogleda se u razvoju domaćih kadrova i znanja, kao i stvaranju preduvjeta za integraciju u međunarodnu Grid mrežu. U suprotnom će doći do zaostajanja u znanosti i gospodarstvu, jer se vrhunski rezultati danas temelje na primjeni superračunala, a uskoro će to biti na primjeni grid tehnologije i e-Science aplikacije. Stoga je zalet prema World Wide Gridu i e-Science aplikaciji temeljna generička strategija informatičkog razvoja Instituta, kao i cijele Hrvatske.



nastavak sa str. 5

(Voice over IP), koji je zbog smanjenih izdataka za javnu telefoniju danas u velikoj ekspanziji. VoIP je našao primjenu u povezivanju klasičnih PBX-ica putem data mreže integracije govor (voice) i podatkovnog (data) prometa, čime se osigurava znatna ušteda. Enkapsulacija voice prometa u IP pakete događa se na prijelazu iz PBX-ica u data mrežu, a vrše je VoIP gateway-i. IP telefonija je otišla još korak dalje, i uvela IP protokol u telefonski aparat. Tako se cijela komunikacija između oba telefonska aparata odvija uporabom IP protokola. Robusnost, mogućnost reguliranja kvalitete usluge (QoS), rasprostranjenost i sve ostale prednosti IP protokola zadržane su u podatkovnom prometu. Njima je pridodan i cijeli niz funkcija radi postizanja visoke kvalitete prijenosa govora. Komutaciju poziva i rezervaciju resursa u toj mreži obavlja IP telefonska centrala (npr. Cisco CallManager). Zbog svoje primarne namjene i visokih standarda koje postavlja današnja javna telefonija, IP telefonija zahtjeva visoku raspoloživost podatkovne mreže. Uporabom IP telefonije robusnost, mogućnost reguliranja kvalitete usluge (QoS), rasprostranjenost i sve ostale prednosti IP protokola su zadržane. Postizanje visoke kvalitete prijenosa govora osigurava se nizom dodatnih funkcija. Komutaciju poziva i rezervaciju resursa u takvoj mreži obavlja IP telefonska centrala. IP telefonija je potpuno otvorena prema operativnom sustavu, bazama podataka (SQL) i prema proizvodima koji se koriste kao client uređaji (telefonski primopredajnici). Redundantnost i raspoloživost takvih rješenja već odavno nije upitna, a standardi su sve izraženiji. Već postoje aplikacije koje će IPAQ (Compaq Handheld uređaj) moći spojiti na 802.11b WLAN (bežični LAN baziran na 2.4 GHz frekvencijskom području) i pretvoriti ga u bežični IP telefon, temeljen na standardiziranom protokolu.

Video aplikacije

Postoje značajne primjene cijelog niza novih video aplikacija baziranih na IP protokolu. Velika propusnost mrežnog sustava i QoS mehanizama omogućuju profesionalno komuniciranje pomoću videokonferencijskih sustava, ostvarujući pritom istovremenu virtualnu prisutnost na više mjesta, klikom na H.323 video klijentu. Multicast i unicast mehanizmi preklapanja omogućuju i druge video aplikacije pomoću IP protokola i vrlo jednostavnog implementiranja aplikacija, kao što su e-learning, odnosno e-Science. Osnova tih aplikacija su Video on Demand i Video Broadcast, odnosno Video Multicast aplikacije koje se zajedničkim imenom nazivaju IP/TV aplikacijama.

IP/TV aplikacije su dio "content networking" aplikacija koje su obično gotova rješenja za davanje video usluga preko IP mrežnih servisa. Takve IP/TV aplikacije obično podržavaju "live video", "scheduled video" "video on demand (VoD)", sinhroniziranje prezentacije s video slikom i široku paletu video management funkcija. Unutar Instituta na strateški važnim mjestima moguće je primijeniti video sigurnosnu zaštitu na on-line ili off-line modu pomoću prepoznavanja pokreta.

Zbog opsežnosti suvremene mrežne arhitekture, tehnologije i mogućnosti koje oni pružaju, na ovom ograničenom prostoru nije moguće dati detaljan prikaz, stoga je članak samo ilustrativne prirode. No, nova je mreža zacijelo okosnica novog razvoja informatizacije i uvođenja čitavog niza novih usluga, koje će Računalni centar osigurati znanstvenim radnicima Instituta, ukoliko nastavi započeti model intenzivnog razvoja, oslanjajući se pri tome na unutarnji znanstveno-stručni intelektualni kapital.

piše: **Kristian Ulahoviček,**
ICGEB Trst

Bioinformatika i računalna biologija



U posljednjih dvadesetak godina svjedoci smo eksplozivnog rasta biotehnologije i molekularne biologije. Usavršavanje i sve veća primjena metoda molekularne biologije doveli su do mnogostrukog povećanja našeg znanja o životnim i staničnim procesima na molekularnoj razini. Između ostalog, tehnike utvrđivanja primarnih sljedova proteina i nukleinskih kiselina (tzv. "sekvenciranje") omogućile su nam da pročitamo što je zapisano u nositelju genske informacije - DNA. Istodobno, količina i raznolikost podataka nastala kao proizvod istraživanja molekularne biologije narasla je do stupnja kada je pohranu i obradu postalo moguće obavljati isključivo automatski - spajem informacijske tehnologije, računalnih znanosti i biologije nastaje nova znanstvena disciplina: bioinformatika.

Nakon ulaska u "eru genoma", pokrenula se lavina biološke informacije u obliku primarnih sljedova nastalih sekvenciranjem cjelokupnih genoma. Stoga je prvotna zadaća bioinformatike bila organizirati biološko znanje u banke podataka, te omogućiti jednostavnu pohranu novih kao i dohvat već postojećih podataka. Najveće banke primarnih sljedova nukleinskih kiselina su GenBank, EMBL i DDBJ a nalaze se redom u Nacionalnom centru za biotehnološku informaciju (NCBI, www.ncbi.nlm.nih.gov), Europskom bioinformatičkom institutu (EBI, www.ebi.ac.uk) i Japanskom Nacionalnom institutu za genetiku (NIG, www.nig.ac.jp). Podaci pristigli iz laboratorija prikupljaju se u geografski najbližem centru i dnevno se razmjenjuju i međusobno sinhroniziraju. Trenutno se nukleotidne banke podataka udvostručuju svakih 15 mjeseci a predviđanja govore da će se taj period skraćivati. Najvažnija banka proteinskih sljedova, SwissProt (www.expasy.org), nalazi se u Švicarskom institutu za bioinformatiku i jedna je od najbolje održavanih bioloških baza znanja, dok su modeli prostornih struktura makromolekula pohranjeni u Protein Data Bank (PDB, www.rcsb.org).

Usporedno s razvojem računalne infrastrukture za organizaciju podataka, javlja se potreba za razvojem metoda za analizu, pretraživanje i povezivanje raznolike biološke informacije, koje zajednički dobiva naziv računalna biologija. Neke od metoda računalne biologije su:

- pronalazak gena u nukleotidnim slijedovima,
- pretraživanje po sličnosti (similarity search),
- sravnjenja (alignments),
- višestruka sravnjenja i filogenetika,
- predviđanja funkcije proteina,
- predviđanja sekundarne strukture proteina,
- istraživanje i predviđanje prostornih struktura bioloških makromolekula, poznate pod nazivom strukturalna bioinformatika.

Ulaskom u "post-genomsko razdoblje" molekularne biologije, zanimanje računalne biologije također se pomiče prema povezivanju biološke informacije i analizi molekularnih procesa na staničnoj razini - funkcijska genomika (određivanje funkcije svih gena u organizmu), transkriptomika (analiza regulacije trenutno aktivnih staničnih gena) i proteomika (određivanje strukture i osobina svih staničnih proteina i analiza njihove međusobne ovisnosti i međudjelovanja) samo su neke od kovanica nastalih u posljednjih pet godina u kojima računalna biologija ima ključnu ulogu rizničara i tumača podataka. Razumijevanje staničnih procesa će kao krajnji cilj imati i bolji uvid u mehanizme mnogih bolesti kao i načine na koji ih liječiti ili zaustaviti.

Bioinformatika u većini slučajeva traži snažna računala, kako u pogledu količine skladišnog prostora, tako i u pogledu računalne snage. Farme računala ili računalni clusteri su posebno zanimljivi za primjenu u računalnoj biologiji zbog činjenice da je većinu zahtjevnih računskih operacija u bioinformatici moguće izvoditi paralelno (npr. pretraživanje po sličnosti ili sravnjenja - jedna usporedba sljedova ne ovisi o drugoj) i na taj ih način mnogostruko ubrzati. Najsnažnija računala glavnih centara za pohranu bioloških baza podataka, poput već spomenutih EBI, NCBI ili glavnog centra u Projektu ljudskog genoma, Instituta Sanger (www.sanger.ac.uk) su upravo farme računala koje danonočno uspoređuju sekvence, pretražuju banke literaturnih podataka ili rade na sastavljanju ljudskog i ostalih genoma.

Primjena bioinformatike i računalne biologije¹ postala je sveprisutna u svakodnevnom praktičnom istraživanju. Danas je rutinska stvar pretražiti literaturne izvore u potrazi za informacijama o genu kojeg istražujemo, pretražiti banke nukleotidnih ili proteinskih sljedova da bi utvrdili postoji li slijed sličan onome kojeg smo upravo sekvencirali, prevesti slijed nukleotida u slijed aminokiselina, pokušati predvidjeti sekundarnu strukturu ili vidjeti postoji li već utvrđena prostorna struktura proteina sličnog našem. S druge strane, sve je više znanstvenika koji odabiru bioinformatiku i računalnu biologiju kao osnovno polje istraživanja, pokušavajući razviti nove metode za analizu bioloških podataka, integrirati raznolikost i kompleksnu informaciju na razini cjelokupnog genoma i doprinijeti spoznajama o živom svijetu.

¹ Opširnije pogledati članak u glasilu 'Ruđer', broj 2, veljča 2001., Bioinformatika i Internet

pišu: Karoli Skala,
Nikola Pavković,
IRB RC



Osvrt na stanje

Od otkrivanja mikro-procesora odvijaju se znanstvena istraživanja i razvoj računala, korištenjem postignute razine tehničkog i tehnološkog stanja s ciljem povećanja računarske moći (High Throughout Computing) i povećanja kvalitete računanja (High Performance Computing). Razvoj

suvremenog superračunarstva ide u smjeru spajanja više računala u paralelan rad, što dovodi do nastajanja klastera te spajanja više klastera u farmu klastera (klaster grid). To ima veliku ulogu kod razvoja znanosti i tehnologije kod primjene analitičkih i simulacijskih metoda, te oblikovanja složenih procesa i modela.

Prije deset godina, superračunala su se prvenstveno koristila u istraživačkim institutima i na sveučilištima. Danas se preko 50 % superračunala nalazi u gospodarskim primjenama i uočljiva je tendencija intenzivnog širenja i primjene klaster superračunala. Uočljiva je sve veća primjena znanstvenog istraživanja u gospodarstvu i ostalim djelatnostima u društvu. Upravo u toj činjenici leži značaj budućeg Grid sustava. To je razlog zbog čega sve razvijene zemlje ulažu velika sredstva u organiziranu izgradnju nacionalnih računarskih resursa pod Grid okruženjem. U razvijenim zemljama svaki značajniji institut ili fakultet, kao i veći gospodarski subjekt opremljen je klaster superračunalom. U tim zemljama nacionalni razvojni projekti nadmašili su superračunala i umjesto njih se intenzivno otvaraju veliki projekti koji su usmjereni na povezivanje klastera, gradeći pri tome nacionalni Grid sustav. To je nova računalno-mrežna tehnologija koja ne prenosi samo podatke i informacije, nego dijeli resurse (klustere) uz pomoć brze mreže. Suština ove tehnologije se svodi na povezivanje različitih klaster-računala s brzom mrežom, kako bi se tvorilo novo jedinstveno virtualno superračunalo. Očekuje se da će Grid u širem smislu primjene imati značajniju ulogu na polju obrade informacija nego što je Web ima na dobavi informacija. Pretpostavlja se da će društveno razvojni učinak umreženih klastera pomoću posrednika (middleware) biti izuzetno velik, te da će predstavljati sveobuhvatnu distribuiranu računalno-komunikacijsku mrežnu tehnologiju sličnu onoj koju električna mreža predstavlja kod energetike (distribuirana energetika).

Integracija u farmu klastera

Računalni resursi u globalnim razmjerama su postali golemi. Više od 400 milijuna PC računala na svijetu (koja su pojedinačno u prosjeku jača od superračunala 90-ih godina) koriste

se od 5 do 10 % vremena. Uporaba Grid tehnologije omogućuje iskorištavanje računalnih resursa i do 80 %, čime se stvaraju enormne računalne moći na globalnoj razini. U svijetu i u našem radnom okolišu postoji velik broj računala koje preko dana koriste studenti, učenici, polaznici tečajeva i sl., međutim tijekom noći ta računala stoje neiskorištena. Obzirom da se uglavnom radi o računalima relativno velike procesorske moći, te da se u takvim učionicama nalazi i po nekoliko desetaka računala, uočili smo da je tu zapravo neiskorišten veliki potencijal koji takve 'skupine' računala mogu pružiti.

Ti računalni resursi mogu poslužiti za obavljanje računarskih usluga za akademsku i znanstvenu zajednicu.

Budući da su brza računala skupa, često je za neki tim znanstvenika vrlo teško priskrbiti sredstva za izgradnju klastera dostatnog kapaciteta, pa se postavlja potreba organizirane usluge mrežnog računarstva. Distribuirano računarstvo, povezivanje klastera u farme klastera na način da se iskoriste postojeći računalni resursi u smjeru stvaranja uvjeta za oživljavanje velikog virtualnog superračunalnog sustava je pravi izazov i učinkovito strateško opredjeljenje. Razvoj ide u smjeru stvaranja Grid okruženja s programskom kontrolom posrednika (middleware) u svrhu dosega visokoorganiziranog mrežno distribuiranog računarstva.

Distribuirana računala određene arhitekture i programske podrške postupno se razvijaju u smjeru samoodržavajućih sustava. Pod tim se podrazumijeva automatizirana, inteligentna i samoprilagodna funkcionalnost sustava. Klasteri i farme klastera koji predstavljaju visoko učinkovite računalne sustave moraju imati hardverske i programske pretpostavke za prilagodljivo samovođenje (adaptable self mangement), što podrazumijeva brzu detekciju i operacionalizaciju funkcionalnog problema i automatsku prilagodbu novonastalom stanju.

Klaster učionice

Multifunkcionalna računalna učionica je programski vođena tako da računala preko noći rade kao jedinstveno klaster računalo pod Linuxom te se izvan radnog vremena 'uključuje' kao dodatni resurs u postojeći klaster na IRB-u (<http://clu->

Farma klastera
povezivanjem slotova

klastera obodnih resursa

ster.irb.hr/) prihvaćajući zahtjeve za obradom koji se trenutno nalaze u redu čekanja.

Ujutro se računala vraćaju u 'normalan' mod rada, tj. služe kao nastavno sredstvo pod Windows operacijskim sustavom.

Postojeća programska infrastruktura na klasteru grozd.irb.hr može se rekonfigurirati tako da batch sustav skupine računala iz učionice automatski učini 'raspoloživim' za pokretanje poslova čim se ona podigne u Linux operacijskom sustavu na kraju radnog vremena.

Opis izvedbe farme klastera

Trenutni računalni resurs lociran na Institutu Ruđer Bošković kao računalni klaster grozd.irb.hr sastoji se od 5 računala sa po 2 x 1 GHz procesora i 1 GB radne memorije, 10 računala sa po 2 x 1.5 GHz procesora i 1 GB radne memorije, u postupku je instalacija još 8 računala sa po 2 x 2 GHz procesora i po 2 GB radne memorije. Kako bismo mogli steći uvid u 'veličinu' resursa, uvest ćemo jedinicu GHz*cpu, koja predstavlja količinu raspoloživog računalnog resursa na jednom procesoru brzine 1 GHz. Izražavajući se u toj jedinici, računalni klaster grozd.irb.hr ima 'snagu' od oko 80 GHz*cpu. U ovom razmatranju nismo uzeli u obzir količinu radne memorije raspoložive na svakom od radnih čvorova, već smo radne čvorove razmatrali kao 'čistu' procesorsku snagu. Ovaj klaster je namijenjen samo za izvršavanje računanja u sklopu znanstvenih projekata na Institutu Ruđer Bošković, te je kao takav raspoloživ 24 sata dnevno. Kod klaster-učionica resurs nije dostupan 24 sata dnevno, njegova dostupnost uvjetovana nekom vrstom 'radnog vremena' takvog klastera.

Kao poseban razvojni projekt¹ izgradili smo klaster u sklopu računalne učionice na Grafičkom Fakultetu u Zagrebu. Tu je na raspolaganju resurs od 10 GHz*cpu. Tako smo instalirali multifunkcionalnu računalnu učionicu koja u radne dane od 7:00 do 19:00 sati služi za obrazovanje pod Windows operacijskim sustavom, a preko vikenda i noću ona se prelaze u Linux klaster. Rješenje automatizacije prelazaka iz jednog u drugi mod rada postignuto je paralelnim implementiranjem određenih 'open source' alata i vlastitim programskim razvojem određenog dijela sistemskog software-a. Kao dodatni produkt, razvijena je i tehnologija polu-automatskog administriranja Windows operacijskog sustava na računalima u učionici, što uvelike olakšava rad sistem inženjeru zaduženom za učionicu u navedenoj ustavnosti. Tako dobiveni

klaster je CARNet mrežom spojen na klaster u Institutu. Razmatrajući tako stvorenu farmu klastera (inicijalni grid) kao distribuirani računalni sustav, može se reći da je klaster lociran na Institutu Ruđer Bošković na neki način 'glavni' klaster. Naime, njegov resurs je konstantan, dostupan 24 sata. On iznosi, već prije spomenutih 80 GHz*cpu. Kada se, međutim, noću 'probude' i računala na Grafičkom fakultetu povećava se raspoloživi resurs za 10 GHz*cpu-a.

Međutim, javlja se problem distribucije posla prema dislociranom klasteru, što nije nimalo jednostavan postupak, i trenutno u svijetu nema univerzalnog i potpuno funkcionalnog rješenja koje bi takve dislocirane resurse u potpunosti ujedinilo u jedinstveni računalni grid.

Postoji nekoliko pristupa rješavanju tog problema, no svi pokušaji su još relativno nedorečeni da bi se od njih očekivalo ozbiljan i stabilan režim rada. Trenutno radimo na implementaciji Globus programskog paketa koji ima učinkovite alate za dijeljenje resursa, te pokušavamo dohvatiti Silver meta-scheduler kao krovni scheduler, kako bi integrirali sve raspoložive računalne resurse u jedan funkcionalan grid sustav. Dodatna poteškoća nastaje zbog uskog grla (2 Mb) CARNet mreže na ulazu Grafičkog fakulteta što je presporo da bi preko njega korisnički home direktoriji bili distribuirani po klasteru, pa se koriste stage-in i stage-out skripte za predavanje datasetova, input datoteka i dobivanje rezultata. U svrhu povećanja brzine prijenosa podataka između dvaju klastera (GF i IRB) izgrađen je bežični mikrovalni link kojim se postiže znatno veća brzina.

Bežični LAN/MAN u sustavu povezivanja klastera

Bežični LAN/MAN (mikrovalni ili laserski) ima posebno značenje u sustavu urbanih mreža (Metropolitan Area Networks) kao ekstenzija i nadopuna grid mreže. Služi za umrežavanje računalnih resursa u grid sustav na onim udaljenostima gdje nije moguće ili nije isplativo postavljanje optičkih kablova za mrežnu komunikaciju, kao i u sredinama sa nezadovoljavajućom ili nepostojećom (neizvedivom) mrežnom infrastrukturom.

Za ostvarivanje bežične komunikacije između Grafičkog fakulteta i vodotornja IRB koristimo mikrovalnu komunikacijsku vezu. Na toj relaciji je jako dobra optička vidljivost između objekata. Kod povezivanja tih dviju lokacije postavljen je po jedan most (bridge) na svaku stranu. Bridge ima dva linka, jedan od njih je 10BaseT za lokalnu mrežu, a na drugi se spaja antena.

U slučaju kada se povezuju dva fizički odvojena lokalna klaster LAN-a, tj. klaster računala, tada se bridgeovi konfiguriraju kao "point-to-point" uređaji, što znači da samo oni komuniciraju međusobno. U tom slučaju se bridgeovi ponašaju kao Ethernet ponavljači (repeater). Moguće je napraviti i "point-to-multi-point" vezu što može poslužiti kada je klaster smješten na više fizičkih lokacija, a potrebno je ostvariti komunikaciju svih čvorova sa frontalnim čvorom. Za spajanje dvaju objekata koriste se helikoidalne usmjerene antene koje usmjeravaju mikrovalnu energiju. Bežična mreža između GF-a i vodotornja IRB-a je ostvarena na 2.4 GHz teoretskom brzinom pri-

nastavak na str. 22

¹ Članovi projektne skupine su: K. Skala, N. Pavković, V. Klein, T. Skala, G. Topić, Z. Šojat, E. Wendling.

piše: Miroslav Mihaljević,
IRB RC

Grid aplikacije u znanstvenom radu



Uvod

Od pedesetih godina dvadesetog stoljeća do današnjih dana snaga stolnih računala povećala se približno milijun puta. Međutim, napredak računala i računalne tehnike još uvijek nije dovoljno velik za rješavanje mnogih složenih znanstvenih problema. Spremanje velikih količina podataka (npr.

spremanje mamografskih snimki) i izračunavanje složenih proračuna (npr. eksperimenti koji uključuju promatranje razbijanja čestica) još uvijek predstavljaju kamen spoticanja kod korištenja računalne tehnologije u znanstvenim istraživanjima.

Postoje dva pristupa rješavanju ovog problema. Prvi jest izgradnja superračunala koje će moći spremati velike količine podataka i obavljati velik broj računalnih operacija u sekundi. Zbog iznimne skupoće i nepraktičnosti ovog pristupa danas se sve više pristupa drugom rješenju - spajanju velikog broja stolnih računala u jedinstveni sustav tzv. grozd računala i njihovo povezivanje u zajedničku, snažnu mrežu računala - Grid.

Problem upravljanja podacima

Izumom prvog tranzistora i ne tako davne 1948. godine otvorilo se čitavo novo područje u razvoju elektronike, kao i čitav niz znanstvenih problema koje je trebalo riješiti uporabom nove tehnologije. Kako je tekao napredak elektronike, računala i računalne tehnike tako se pojavljivalo sve više problema različitih znanstvenih profila koji su zahtijevali uporabu računala. Problemi su rasli od jednostavnih upravljačkih mehanizama do složenih sustava koji upravljaju velikim bazama podataka.

Usprkos brzom razvoju računalne tehnike i tehnologije još uvijek postoji niz znanstvenih problema koje je izuzetno teško riješiti čak i uz primjenu najsnažnijih računala današnjice.

Na primjer, detektori čestica u CERN-u (European Laboratory for Particle Physics) do 2005. godine proizvodit će nekoliko petabajtova podataka godišnje, što je približno milijun puta više od mogućnosti spremanja podataka današnjih stolnih računala. Pretraživanje i analiziranje ove velike količine podataka zahtjeva aplikacije koje koriste približno 20 teraflopsa po sekundi. Najsnažnija superračunala današnjice mogu obrađivati približno 3 teraflopsa po sekundi, što ukazuje na potrebu za postizanjem veće računalne snage koju nije moguće postići samo s jednim računalom.

Podijeli pa vladaj

Jedno od rješenja nedostatka računalne snage jest raspodjeljivanje poslova na više računala. Drugim riječima, za

postizanje dovoljne razine računalne snage prišlo se grupiranju velikog broja samostalnih računala u tzv. grozd računala (engl. cluster). Ovaj pristup prvi put je primijenjen ranih 1980-ih godina, a danas predstavlja standardnu praksu svih velikih centara za razvoj superračunala i znanstvenih institucija. Najbrža superračunala današnjice predstavljaju grozdovi, poput sustava ASCI White System od 8000 procesora u Lawrence Livermore National Laboratory u Kaliforniji (SAD).

Mnoge znanstvene institucije posjeduju grozdove računala čija je cijena puno manja od jednog superračunala, a koji pružaju dovoljno računalne snage za analiziranje i spremanje velikih količina podataka. Usprkos prednostima koje pruža izgradnja grozda, grozd i dalje ostaje računalni sustav koji se nalazi na mjestu izgradnje, unutar jedne institucije i ograničen je brojnim tehničkim, financijskim i političkim ograničenjima. Na primjer, spomenuti ASCI White System koštao je 110 milijuna USD i zahtijevao je izgradnju dodatne zgrade za njegov smještaj. Zbog navedenih razloga, svega mali broj institucija i znanstvenih ustanova može osigurati sredstva za izgradnju velikih grozdova računala.

Internet računarstvo

Paralelno s razvojem računala razvijale su se i komunikacijske tehnologije. Razvoj komunikacija omogućio je razmatranje i usvajanje koncepta distribuiranog računarstva putem Interneta. Diljem svijeta nalazi se preko 400 milijuna osobnih računala od kojih su mnoga jednako snažna kao i superračunala s početka devedesetih godina 20. st. Većinu vremena mnoga računala su u stanju mirovanja (engl. idle). Svaka velika tvrtka ili institucija ima na stotine ili tisuće ovakvih računala koja većinu dana (npr. noću) ne rade ništa. Internet računarstvo upravo teži iskorištavanju velike računalne snage koja se nalazi u tim računalima.

Mogućnost korištenja snage računala koje je u stanju mirovanja poznata je već duže vremena. 1985. godine Miron Livny je proveo istraživanje kojim je pokazao kako su radne stanice najčešće u stanju mirovanja, čekajući izvršavanje nekog novog zadatka. Zadaci prilikom svojeg izvršavanja zahtijevaju određeni broj računalnih ciklusa (taktova procesora, proračuna u aritmetičko-logičkoj jedinici, računalne memorije itd.). Nakon što se zadatak na računalu obavi do kraja, računalno prelazi u stanje mirovanja i proizvodi tzv. prazne cikluse, čekajući na novi zadatak. Ideja jest da se u takvom slučaju računalu ipak zada neki zadatak kako bi se ono korisno upotrijebilo.

Najpoznatiji primjer Internet računarstva jest SETI@home projekt Davida Andersona. SETI projekt bavi se analizom radio signala koji se uz pomoć odgovarajućih radio teleskopa prikupljaju iz svemira s ciljem otkrivanja života u svemiru. Veliki problem jest analiziranje snimljenih zapisa zbog njihove veličine. Kao rješenje izabrano je Internet računarstvo.

Većina korisnika Interneta posjeduje neki od screensavera koji se uključuju nakon što korisnik nije nekoliko minuta radio s računalom. Screensaveri najčešće prikazuju slike, simuliraju putovanje kroz svemir i sl. Primarna uloga im je smanjenje potrošnje električne energije i produljenje životnog vijeka monitora. Međutim, sa SETI projektom screensaver i korisnikovo

računalo postaju dio velikog znanstvenog projekta. Svaki put kada se pokrene SETI screensaver program putem Interneta pristupa se bazi podataka radio signala, dohvaća se određeni radio zapis i započinje njegova obrada na korisnikovom računaru. Obrada traje samo dok je uključen screensaver i ne utječe na svakodnevni korisnikov rad na računaru. Po završetku obrade radio zapisa rezultati se prosljeđuju putem Interneta SETI centru, dohvaća se novi zapis i cijeli postupak se ponavlja. SETI projekt danas predstavlja najveći distribuirani sustav na području Internet računarstva koji obuhvaća oko pola milijuna kućnih računala koji isporučuju 1000 CPU godina rada dnevno, što predstavlja najbrži, iako specijalizirani računalni sustav svijeta.

Kakav potencijal predstavlja Internet računarstvo za znanstvenike danas? Na prvi pogled čini se da su znanstvenici Internet računarstvom dobili iznimno snažan računalni sustav za rješavanje znanstvenih problema. Sve što je potrebno napraviti jest pretočiti znanstveni problem u oblik pogodan za izvršavanje na kućnim računalima i uvjeriti javnost (ili Internet tvrtku) u važnost izvršavanja njihovog projekta i korištenja "spavajuće" računalne snage.

Grid

Međutim, pravi potencijal krije se u povezivanju pristupa grozda i Internet računarstva. Grozdovi predstavljaju iznimno snažna računala, ali su još uvijek koncentrirani oko pojedinih, dovoljno bogatih i moćnih institucija. Internet računarstvo ograničeno je na stolna računala korisnika Interneta, koje nije uvijek lako i jednostavno privoljeti na trošenje vremena i novaca na sudjelovanje u znanstvenim projektima. Internet računarstvo predstavlja samo prvi korak prema nečem mnogo značajnijem - mogućnosti da institucije i pojedinci zajedno povezuju i dijele resurse kako bi postigli više ciljeve. Danas se znanost sve više temelji na suradnji znanstvenih institucija različitih profila i nije neobično da se pojedine institucije povezuju s drugim institucijama, tvrtkama i državnim ustanovama. Tehnologije poput Interneta i e-maila omogućuju povezivanje i zajednički rad različitih institucija. Ali, što ako te institucije ujedine podatke, računala, senzore i ostale resurse s ciljem stvaranja jednog velikog virtualnog laboratorija? Rješenje jest uporaba tzv. Grid tehnologija koje pružaju protokole, usluge i razvojne alate za ostvarivanje jednog takvog virtualnog laboratorija, omogućavajući kontrolirano dijeljenje resursa s ciljem ostvarivanja znanstvene suradnje.

Grid koncept po prvi puta primijenjen je 1995. godine u I-WAY eksperimentu, kojim je brzom mrežom u kratkom vremenskom roku povezano 17 različitih siteova diljem sjeverne Amerike. Do današnjih dana pokrenut je velik broj različitih projekata s ciljem istraživanja Grid tehnologija. Na primjer, National Technology Grid i NASA-in Information Power Grid koji su prvenstveno namijenjeni za istraživanja u NASA-i i na američkim sveučilištima. U međunarodnoj suradnji razvijeni su European Data Grid, Particle Physics Data Grid, i Grid Physics Network koji su prvenstveno namijenjeni istraživanjima na području fizike.

Grid sustav se prvenstveno temelji na već postojećim tehnologijama poput TCP/IP protokola. Međutim, još uvijek nisu definirani protokoli i standardi za komunikaciju i povezivanje grozdova koji se nalaze diljem svijeta. Ipak, intenzivan rad na tom području rezultirao je izgradnjom open source aplikacija koje pružaju osnovne usluge za izgradnju Grid aplikacija, poput sigurnosnih mehanizama, otkrivanja resursa, upravljanja resursi-

ma i pristupu podacima.

Grid u Hrvatskoj

Na inicijativu CERN-a u Europi je pokrenut projekt European Data Grid čiji je cilj uspostavljanje računarske mreže resursa za analizu podataka dobivenih znanstvenim istraživanjima. Glavna svrha ovog projekta jest efikasna podjela velikog broja podataka preko postojeće mrežne arhitekture. Ideja projekta jest izgradnja centara za testiranje smještenih u raznim znanstvenim institucijama diljem Europe, koji će izvršavati aplikacije i obrađivati podatke sa različitih znanstvenih područja poput fizike visokih energija, biologije i meteorologije.

Cern-ov Data Grid projekt su uključene europske i američke znanstvene institucije pa tako i Institut Ruđer Bošković (Zagreb), Prirodoslovni Matematički Fakultet (Zagreb), Prirodoslovni Matematički Fakultet (Split) i Fakultet elektrotehnike, strojarstva i brodogradnje (Split). Kristeći stečeno iskustvo pri izgradnji grozd računala na Institutu Ruđer Bošković, napravljen je jedan manji klaster u SRCU koji se koristi za europski Data Grid project.

Na Institutu Ruđer Bošković, izgrađen je grozd računala namijenjen izvođenju najsloženijih znanstvenih aplikacija. Grozd se sastoji od desetak međusobno povezanih računala čime je postignuta značajna računalna snaga koja se očituje u brzini obrade podataka i mogućnostima njihovog spremanja. Trenutno se na grozdu izvršava nekoliko znanstvenih aplikacija različitih profila, i u velikoj mjeri je iskorišten.

Aplikacije koje trenutno postoje pružaju mnoge mogućnosti: analiziranje vibracija, analiziranje problema iz područja termokemije, različiti kemijski proračuni i programi za modeliranje i proučavanje strukture atoma. Međutim, grozd se može koristiti za sve vrste znanstvenih istraživanja i znanstvenih disciplina - za astrofiziku, meteorologiju, fiziku materijala, turbulentna gibanja, hidrodinamiku, biologiju, proučavanje i stvaranje novih lijekova itd. Poprilično velik, a neiskorišten potencijal za korištenje Grida, posjeduju istraživanja na području molekularne biologije. Istraživanja na području ljudskih gena i proučavanje DNA strukture i otkrivanja odgovarajućih bioloških funkcija u DNA slijedu zahtijevaju obradu velikih količina informacija i složena, dugotrajna pretraživanja i uspoređivanja pojedinih dijelova DNA zavojnice. Ovakva istraživanja zahtijevaju iznimnu računalnu snagu koja se može postići korištenjem grozda računala. Grozd računala razvija dovoljnu računalnu snagu za izvođenje najsloženijih računarskih algoritama i pohranjivanje enormne količine podataka.

Velika prednost grozda računala u odnosu na pojedina super-računala jest njihova jednostavna, efikasna i brza nadogradnja. Kupnjom dodatnih računala i njihovim uključivanjem u grozd raste procesna moć, skladišni prostor i raste snaga čitavog sustava. Samo uključivanje novih računala danas je iznimno jednostavno zahvaljujući moćnoj programskoj podršci.

Međutim, pravi potencijal distribuiranog računarstva i računalnih grozdova krije se u povezivanju grozdova računala koji postoje u drugim institucijama i znanstvenim ustanovama diljem Hrvatske zajedno sa grozdom Instituta Ruđer Bošković. U Hrvatskoj je na inicijativu Akademije tehničkih znanosti a na poticaj Instituta Ruđer Bošković predložen složen tehnološki projekt pod imenom CRO GRID koji predviđa spajanje IRB-SRCE-FESB grozd računala u Grid sustav. Stvaranjem Grida u Hrvatskoj, znanstvene institucije došle bi do golemog računalnog potencijala.

nastavak na str. 26

piše: Zorislav Šojat,
vanjski suradnik RC-a



Nanoračunarstvo i prirodno distribuirani paralelizam

Ne bi bilo pretjerano smjelo tvrditi da u kompjutoristici već poneko vrijeme više gledamo u postignuto, i samo na tome, ponekad i nekritično, gradimo svoje vizije, nego li što gledamo u ono što

dolazi, našom voljom i željom, ili po prirodi stvari. Gotovo zdravo za gotovo mnogi ljudi, pa i stručnjaci, uzimaju vrtoglavu brzinu porasta snage računanja, veličine raznih vrsta strojnoga pamćenja i opće brzine obrade osnovnih strojnih naredbi. Programe svih vrsta i za sve namjene to, poneki bolje, poneki lošije, obilno iskorištavaju. Kakvoća prikaza i brzina obrade veoma složenih problema postaje dostupna za prikaz mnogih kompleksnih procesa u stvarnome vremenu. I budućnost, sa sve bržim, boljim i sposobnijim strojevima, izgleda sve bliže i bliže, i mnogima podpunno izvjesno.

Ipak, često nas naše zadovoljstvo postignutim dovodi i do neugodnih prepreka na putu. Nama suvremeni, gotovo isključivi, pristup pisanju programa za strojeve jest uporaba jednostaznoga Turingova stroja, tj. uporaba strojnih naredbi koja slijede jedna za drugom, a obrađuju podatke isto tako, jedan za drugim. Takav ustroj stroja četveroslovnom se engleskom kraticom zove SISD, 'Single Instruction, Single Data', "Jedna naredba, jedan podatak". Takav pristup izradi procesnih jedinica dovodi do potrebe da se ukupnost svih radnji koje stroj mora obaviti da bi se primjerice neka obrada izvela, te prikazala na monitoru (uključujući, dakako, i oslikavanje svake pojedine točkice na ekranu), izvodi u jednome jedinome pažljivo isprepletenom nizu naredbi, a svaka naredba izvodi jednu i samo jednu (veoma elementarnu) unarnu ili binarnu (veoma rijetko plenarniju) operaciju na podatku. Uobičajeno se to svodi na pribrajanje, množenje s dva i ispitivanje jednakosti, te varijacije tih operacija.

Da bismo postigli dovoljne brzine obrade potrebite za takav pristup izvođenju svih, od najosnovnijih do najsloženijih, poslova, brzina se obradnih podsustava iz godine u godinu naglo povećava. Dok je još od sredine 70-tih pa sve do početka 80-tih godina prošloga stoljeća bilo potpuno prihvatljivo imati SISD procesor s brzinom izvedbe od 2 do 16 mikrosekundi po naredbi, s taktnom frekvencijom od 1 MHz, početkom ovoga tisućljeća taktne se frekvencije uobičajenih prihvatljivih strojeva kreću već i do 2,5 GHz i više. To znači da bi procesor iste arhitekture iz zadnje četvrtine prošloga stoljeća sada izvodio, s taktnom frekvencijom od 2 GHz, instrukcije u roku od 1 do 8 nanosekundi, tj. 2×10^3 brže (dvije tisuće puta brže).

Već gotovo cijelo jedno stoljeće svijet poznaje brzinu svjetlosti kao gornji limes brzine ijedne materijalne čestice, kako je to

lijepo Einstein uobličio u svojoj teoriji relativnosti. Ta brzina, koju rabimo kao kućnu riječ kada svjetlosnom brzinom nešto obavimo, ali i kao važnu mjernu jedinicu u proučavanju svemira, nevjerojatnih je 300.000 kilometara u samo jednoj sekundi. Ekvator svjetlost obiđe odprilike punih 7 i po puta u samo jednoj sekundi, do našega Mjeseca treba joj prosječno jedna sekunda, do dalekoga Sunca poprilično manje od desetak minuta. A koliko li ta ista "svjetlost" prijeđe za vrijeme izvođenje jednoga (punovalnoga) takta od 2GHz? Zastrašujuće malo - svega 15 centimetara!

Na naše veliko zaprepaštenje iz gornjega proizlazi da smo u manje od tridesetak godina kraja prošloga stoljeća ubrzali ritam rada naših uobičajenih strojeva za 2000 puta, a ubrzanje od još samo 15 (!) puta, dakle ubrzanje od samo 0.75% u odnosu na tridesetogodišnje postignuto ubrzanje, dovodi nas do situacije da svjetlost, tj. signal unutar procesora u idealnoj situaciji, prijeđe svega 1 cm. To se događa na 30 GHz taktne frekvencije. Ni to nije tako strašno, ali suvremeni su procesori uistinu veličine veće od 1 cm s kraja na kraj.

12. lipnja 2002. u časopisu Nature objavljen je niz članaka s rezultatima istraživanja koja su omogućila izvedbu jednoatomske i dvoatomske tranzistora. Iako u ovome trenutku, kako se napominje u jednom od članaka iz tog broja časopisa, jednoatomske i jednomolekularne tranzistori nisu rivali silicijevim tranzistorima, razvoj je njihove primjene na pragu ulaska kroz velika vrata.

Dakako i jednoatomske i jednomolekularne tranzistori moraju imati ulazno/izlazne mogućnosti povezivanja sa svojom sudjelujućom okolinom, te njihova uporaba zahtijeva njihovu specifičnu prostornu povezanost (za sada preko zlatnih sondi, a u budućnosti i atomskim provodnicima) s ostatkom pomoću njih projektiranoga uređaja. Gore navedeni jednoatomske tranzistori sastoji se, primjerice, od atoma kobalta okruženoga atomima ugljika i vodika, vezanih na okupu pyridinom.

Možemo reći da je red veličine jezgre atoma 10^{-15} metara, a cijeloga atoma oko 10^{-11} . Zaokruživši, možemo uzeti da bi trebalo negdje dvije stotine milijuna atoma jedan do drugoga za nanožicu od jednoga centimetra. Koliko ćemo se atomskim i molekularnim nanotranzistorima približiti prirodnim granicama brzine obrade teško je uistinu predvidjeti, jer mjesta za ubrzanje potpuno novim pristupima uistinu ima, posebice izradom više ne plošno, već prostorno raspoređenih komponenti obradnih i drugih mikro- (tj. nano-) jedinica. Dakako pri tome se javljaju svi mogući problemi, od zagrijavanja pa sve do problema potrebnoga prostora za organizaciju atomskih ili molekularnih tranzistora u uređaj koji nešto radi, a da pri tome nipošto ne zaboravimo da se logička vrata bilo koje vrsti u kompjutorskome procesoru ne mogu izvesti jednom komponentom. No također je važno sjetiti se da i idealan atomski kompjutor mora na neki način komunicirati sa svojom okolinom, prvenstveno sa svojom memorijom. 256 MB (tj. 268.435.456 Byte-a) memori-

je početkom ovoga tisućljeća nije mnogo niti za kućni stroj. Kada bi svaki atom gore navedene veličine mogao preuzeti ulogu jednoga tranzistora, s izravnim međuvezama, a za pamćenja samo jednoga bita potrebna su najmanje dva (uobičajeni flip/flop s križnim povezivanjem), za pamćenje tih 256MB bila bi potrebna količina atoma koja bi se poredala u niz dugačak čak 16 cm. Čitaoc će sam uvidjeti u kakvim se problemima kompjutoristika nalazi početkom 21. tisućljeća.

Potpuno drukčiji pristup pojmu obrade obavijesti postoji u tzv. "kvantnih kompjutora". Kvantni kompjutori, koji se početkom ovoga tisućljeća počinju pojavljivati u laboratorijima eksperimentalnih fizičara, no ne kao oruđe, već kao izum, uistinu preskaču velik dijapazon ideja klasične kompjutoristike, uistinu zasnivajući svoju moć na preciznoj kontroli pojedinih spinova. S toga razloga i naziv kvantni kompjutor. Već i u samome načinu zapisa postoji osnovna bitna razlika. Kvantni kompjutori ne rade s bitovima, dakle dvojičanim oznakama, 0 ili 1, + ili -, već s tzv. qubitima, tj. "quantum bit"-ovima, koji su posebni po tome što osim očitih položaja za 0 i 1, tj. + i -, imaju čitav beskonačan niz međustanja.

Još je 1972. Feynman rekao da je vjerojatno pravilan način obrade mnoštva podataka iz kvantne mehanike, obrada u stroju koji se i sam ponaša po tim istim načelima. Kako je područje kvantnoga računanja u potpunosti u začetku, ne samo postojanjem strojeva, već i samim svojim principima, jer radi se o procesiranju zasnovanome na međudnosima beskonačnih brojeva mogućih stanja pojedinih memorijskih jedinki, čime se uistinu ti kompjutori svrstavaju u, mogli bismo reći, digitalizirane analogne kompjutore, broj je razvijenih algoritama koji daju upotrebljive rezultate izuzetno ograničen, a jedan od najvažnijih programa je program izračunavanja najmanjega zajedničkoga nazivnika. Interesantno je primijetiti da za već postojeće algoritme proračuni ne traju vremenski eksponencijalno, već polinomno.

I kamo dalje?

Prije nego li pokušamo odgovoriti na jedan od mogućih načina na ovo vječno pitanje, ovdje i sada vezano na mogućnosti obrade, te unosa i vizualizacije sve veće količine podataka, sa sve složenijim matematičkim proračunima, kako se znanje o finesima postojanja ovakovoga materijalnoga svijeta sve više produbljuje, sjetimo se mozga.

Ovaj začudan organ omogućuje takove mogućnosti obrade podataka svih vrsta, od abstrakcije, asocijacije, stvaranja i primjene prediktivnosti modela u obliku intencije, pa sve do savršene koordinacije vida, sluha i pokreta, kakove si mogućnosti u kompjutera još dugo ne možemo niti izdaleka zamisliti, uz sav naš najbolji trud. Pri tome ovdje govorimo uistinu samo o "čistim" obradnim mogućnostima mozga. Da se lakše shvati sva kompleksnost proračuna koji se u, primjerice ljudskome mozgu odvijaju, samo je potrebno zamisliti uobičajenu situaciju s prometnicom, koja se, uz statistički izuzetno malu količinu (nažalost često fatalnih) grešaka, odvija na mirijadama cesta u cijelome svijetu svakoga dana već stotinjak godina. To je običan čin pretjecanja. U trenutku donošenja odluke o pretjecanju (zamislimo uvjete ravne prometnice sa sporijim vozilom ispred i nadolazećim vozilom iz daljine), na moždanim je procesima da na temelju brzine povećanja veličine dolazećega vozila, na temelju odnosa brzine vlastitoga vozila i vozila ispred njega (koja se načelno procjenjuje brzinom promjene veličine vozila ispred), zvuka vlastitoga motora i drugih podataka o njegovoj trenutačnoj sposobnosti razvijanja ubrzanja, znanja o voznim karakteristikama svoga vozila, udaljenosti dolazećega

vozila, te mnoštva drugih važnih čimbenika, poput stanja kolnika, širine, vidljivosti za sve sudionike itd., odlučuje u veoma kratkome vremenu o mogućnosti da vlastito vozilo uđe u svoj trak prije nego se zatvori prostor prolaza. Uz to treba paziti i na druge sudionike u prometu, ispred i iza u istom, te u suprotnom smjeru, kao i o vjerojatnostima pojave nepredviđenih okolnosti. Pokušamo li taj izuzetno kompleksni proračun već i samo uobičajenim metodama sekvencijalizirati za programiranje kompjutora, naići ćemo na nevjerovatne poteškoće složenosti toga projekta. Brzina obrade istih podataka (oko, uho, ruka u odgovarajućoj strojnoj inačici) za takovu odluku u tako kompleksnome programu na 2 GHz kompjutoru nije vrijedna spomena, posebice zato što se radi o naučenim i samoučećim procesima.

I sve to mozak radi s živčanim stanicama koje nisu pretjerano sretne ako moraju raditi na taktnoj frekvenciji većoj od stotinjak herza (najveće vrijednosti uzastopnoga pobuđivanja pojedinoga neurona su oko 170 Hz, ali mnogi se procesi u mozgu sinkroniziraju na mnogo rjeđe, sve do svega desetak herza). 2 GHz je 2×10^7 puta brže od 100 Hz.

U čemu je tajna? Tajna je u paralelnoj i istovremenoj obradi, te pravilnoj raspodjeli podataka, podobradnih elemenata i sinkronizaciji prikupljanja polu-produkata sporih ali mnogih neovisnih pod-obrada.

Učeći na tim iskustvima došlo se do ideje o uporabi mnoštva kompjutora međusobno komunikacijski povezanih, koji zajedničkim paralelnim radom rješavaju isti, distribuirani problem, pri čemu svaka ta obradna jedinica radi samo jednu n-tinu cjelokupnoga posla, te se raznim sinkronizacijskim metodama ishodi tih obrada skupljaju u jedan rezultat. Takva nakupina kompjutora za paralelno procesiranje uobičajeno se naziva grozdom, u engleskome "cluster".

Posebice se paralelizacija obrade lako primjenjuje na zadatke koji su ili matematički predstavljeni vektorima, matricama i poljima, ili čiji se algoritamski obradni niz može lako dekomponirati u neovisne sastavne komponente, ili koji po svojoj prirodi ne stavljaju ograničenje na veliku količinu međusobnih podobradnih međuveza. Programi za vizualizaciju veoma se često izuzetno lako, ili barem ne pretjerano teško, prilagođuju paralelnome procesiranju.

No, iako se uistinu i bez pretjerivanja, iz svih gore navedenih razloga, nalazimo na bitnoj prekretnici u ovoj novoj, ali izuzetno važnoj i brzoučećoj grani znanosti, još uvijek ne postoji pravilno znanje za čak niti zadovoljavajuće opće programirne jezike za paralelnu obradu, kao što su to primjerice Fortran, Lisp, Pascal, Ada, Smalltalk i drugi pogodni za razne serijske algoritme. Od programirnih jezika za paralelno procesiranje koji će svakako prije ili kasnije zauzeti časno mjesto pionira, izdvajamo jezike ORCA i OCCAM (veoma zanimljiv jezik vezan uz "transpjutore"). Od starih programirnih jezika treba svakako izdvojiti APL i J Kenetha Iversona, koji su, po svojoj prirodi, matematički, te se samim time posebice odlikuju mogućnošću paralelne implementacije algoritama matematičkih modela.

Drugim riječima rečeno, jedno od područja "spasa" jest svakako paralelna obrada podataka u grozdovima. Za zagovarati je svakako posebice razvoj znanja o mogućnostima paraleliziranja, posebice automatskoga, pojedinih algoritama i postupaka, no velik će se napor na početku ovoga tisućljeća u kompjutoristici morati svakako usmjeriti na istraživanje i pronalaženje osnovnih načela koja omogućuju raspačanu obradu, na zamisao principa i izradu odgovarajućih programirnih jezika koji

omogućuju čovjeku-programeru izricanje njegovih zahtjeva na način dovoljno mu razumljiv, te na izradu radnih sustava dovoljno stabilnih i praktičnih da bi se mogla iskoristiti snaga svakoga pojedinoga grozda.

Dakako da kod povezivanja pojedinih kompjutora u grozd nema nikakvoga razloga ne povezati strojeve iz raznih generacija i raznih arhitektura. Takav heterogeni grozd omogućuje uporabu posebnih sposobnosti pojedinih strojeva, ali mu je mana da svaka skupina istovjetnih strojeva obrađuje zadatak različitom brzinom. Stoga se u posljednje vrijeme grozdovi projektiraju tako da im je relativna razlika brzina pojedinih strojeva mala.

Koliko god se mi trudili algoritme držati brzo izvedivima, potrebe suvremene znanosti, sa svojim sve složenijim i preciznijim modelima pojava koje istražuju, zahtijevaju sve jače mogućnosti obrade. Taj se problem rješava povezivanjem pojedinih, već samih po sebi moćnih grozdova, u široko raspačanu rešetku (koja se po engleskome naziva grid), u kojoj se kroz Internet, ili pak WAN (Wide Area Network), obrada pojedinoga problema upućuje na u tome trenutku najsposobniji grozd za tu obradu, ili se obrađuje istovremeno na više grozdova. Ta se rešetka, grid, poput grozda, treba vidjeti u istome smislu kao i sâm grozd, samo što su njoj obradne jedinice pojedini grozdovi, a grozdu pojedini kompjutori. Načelno su svi principi primijenjeni na stvaranje grozda izravno primijenljivi, na slijedećoj razini, i na grid.

Bitna je razlika između grozda podjednakih strojeva i grida samo u širokome spektru odnosa brzina međusobne komunikacije i čiste obradne snage pojedinih procesnih elemenata (kod grozda pojedinih kompjutora, kod grida pojedinih grozdova).

Kako je već spomenuto, kvantno je računanje, nažalost, u još ranoj fazi razvoja, jer još nije poznato niti koju je vrstu algoritama uopće moguće zamisliti, budući da je načelno istraženo i ostvareno svega nekoliko uistinu jednostavnih postupnika, tj. algoritama. Pa ipak, upravo mogućnost da se u stvarnome svijetu stalno nazočne analogne pojave, koje se u digitalnim kompjutorima moraju uzorkovati određenom grubošču, kvantnim računarom mogu ispravno predstaviti, te da je sama osnova dobivanja rezultata zasnovana na, uvjetno rečeno, bezvremenskoj komunikaciji kroz međusobne kvantno-mehaničke odnose pojedinih spinova, daje mnogo razloga da znakovito i s nadom gledamo i sudjelujemo u razvoju ovoga potpuno novoga polja načela obrade podataka.

Nažalost na svijetu, u suvremenosti početka ovoga tisućljeća

postoji svega na prste jedne do dvije ruke prebrojivih mjesta gdje je trenutačno moguće fizički ostvariti kvantne kompjutore, koji se kreću od jedno-qubitnih do nekoliko-qubitnih obradnika (tj. procesnih jedinica)

No ono što je uistinu utješno, te daje velike mogućnosti daljnje razvoja u tome smjeru, jest da su načela kvantne mehanike i funkcioniranja kvantnih kompjutora dovoljno dobro poznata, te je već do sada izrađen veći broj simulatora i emulatora kvantnih kompjutora, nego li ima samih fizičkih strojeva.¹ Upravo se u ovome otvara izuzetna mogućnost uporabe grozdova i gridova, jer je procesiranje svakoga pojedinoga qubita u emulaciji moguće izvoditi, uz pravilno i dobro programiranje samoga emulatora, na po jednome ili više strojeva u grozdu, ili na po jednome ili više grozdova u rešetki, tj. u gridu. Postoji nada da ćemo upravo ovim smjerom istraživanja riješiti neke od bitnih prepreka koje nam postavlja naše sadašnje znanje iz područja obrade informacija, a koje su, nažalost, često perpetuirane nezanimanjem i predrasudama.

Mnogi podatci govore u prilog tvrdnji da početkom trećega tisućljeća prošlostoljetni pristup kompjutoristici uskoro više neće moći pratiti potrebe novih korisnika, posebice znanstvenika. Očito je da strateško ulaganje u razvoj prvo grozdova, a onda još i više gridova, koji imaju tu prednost nad grozdovima da se u njih, osim procesnih clustera, mogu uvezivati i pojedinačni strojevi, kompjutori, mjerni instrumenti, eksperimentalna aparatura, izvedbeni i proizvodni alatni strojevi i dr., predstavlja izravno ulaganje u blisko predstojeću budućnost. Krajem 20. stoljeća premalo se ulagalo u razvoj znanja o paralelnoj obradi, jer se, naivno, činilo da je jednostavno von Neumanovsko računanje dovoljno snažna pokretačka snaga za sve izazove. Nažalost, ili na sreću, taj se pristup pokazao nedostatnim. Stoga je početkom 3. tisućljeća uistinu potrebno usmjeriti sve svoje snage u ostvarivanje međusobnoga uvezivanja što većega mogućega broja raznih grozdova i individualnih uređaja u jedinstveni grid, jer će budućnost cjelokupnoga daljnje znanstveno-tehničkoga razvoja čovječanstva ovisiti upravo o tim strateškim smjernicama, o nanoračunarstvu - kvantnome i klasičnome - i o uporabi prirodno distribuiranoga paralelizma obrade.

Stoga je za napredak znanosti i tehnike, te time cjelokupnoga čovječanstva, prijeko potrebno uvidjeti stratešku važnost usmjerenja i ulaganja u fundamentalni i aplicirani razvoj grozdova, a posebice i naročito gridova.

(Članak je objavljen u izvornom obliku na zahtjev autora)

¹ Samo kao moguća početna referenca je primjerice <http://www.dcs.ex.ac.uk/~jwallace/qcandqal.htm>, koja, za sada, daje pregled većine istraživanja kvantnoga računanja.

nastavak sa str. 17



jenosa podataka od 11 Mbit/s, uz donju graničnu brzinu prijenosa od 6 Mbit/s pri izlaznoj snazi od 50 mW. Osim mikrovalne veze u skoroj budućnosti predviđen je prelazak na lasersku vezu koja se razvija u Institutu Ruđer Bošković.

Ostvareni učinci

Korištenje slobodnih resursa za stvaranje farme klastera uz odgovarajući posrednik (middleware) predstavlja grid sustav (virtualno klaster superračunalo). On pre-

dstavlja višestruko koristan i zanimljiv razvoj primjene nove arhitekture i integracije klaster računala. Na taj način se domaća tehnologija izgradnje klaster računala, koja je razvijena u Institutu primjenjuje na neiskorištenim resursima i ostvaruje se veliko virtualno superračunalo. Tako se razvija mrežno dohvatljiv i djeljiv računalni resurs i nova usluga na visokoj tehnološkoj razini uz male materijalne troškove.

piše: **Jakov Meštrović,**
vanjski suradnik
RC-a

Baze podataka i baze znanja u razvoju

Baza podataka je skup međusobno povezanih podataka, pohranjenih zajedno bez štetne ili nepotrebne redundancije, a s ciljem omogućavanja što lakšeg korištenja tih podataka od strane različitih aplikacija. Razvojem računalnih znanosti te širenjem i zadiranjem tih istih znanosti u sve pore ljudskog društva, potreba za različitim oblicima računalne pohrane i organizacije podataka neprestano raste. U skladu sa tim, i Institut Ruđer Bošković kao jedan od najvažnijih kotačića u razvoju hrvatske znanosti svakodnevno se susreće sa problemom pohrane, organizacije i što lakšeg pretraživanja raznovrsnih podataka, a u cilju povećanja učinkovitosti djelatnika te pomicanja fokusa pri radu sa tehničkih detalja na rješavanje konkretnih problema. Tako i računalni centar Instituta zbog povećanja obima posla te sveopće informatizacije nalazi se pred problemom što lakše manipulacije sve većim i većim količinama podataka. U skladu s tim, razvijaju se nove aplikacije, osnova kojih su baze podataka. Osim osnovnog zadatka koji se pred te baze podataka postavlja, dakle pohranu i pretraživanje trenutno potrebnih i prikupljenih podataka, vrlo je bitan ispravan dizajn baza koji se, koliko je to moguće, brine o budućim potrebama te ostavlja prostora za dodatne nadogradnje i proširenja. Na taj način, stvara se solidna osnova budućeg rada kako samog Računalnog centra tako i Instituta Ruđer Bošković u cjelini.

U sklopu spomenutog razvoja bitnu ulogu ima baza podataka o resursima: zaposlenicima, opreme, računalima, inventaru, dokumenata i svim ostalim za Institut bitnim informacijama, a koja se planiraju ili trenutačno nalazi u fazi razvoja. Pošto sami podaci o djelatnicima Instituta već postoje u kadrovskoj službi, njihovo ponovno neovisno prikupljanje od strane računalnog centra bilo bi nepotrebno i neefikasno te bi predstavljalo bespotrebnu komplikaciju, stoga je pribegnuto jednostavnijem i kvalitetnijem rješenju: repliciranju dijela podataka koji je računalnom centru potreban i redovitom ažuriranju tih istih podataka. Također, pristup bazi podataka, kako onoj računalnog centra tako i centralnoj, kadrovskoj, omogućen je ovlaštenim djelatnicima, no svakome od njih omogućen je pristup i izmjena samo onih podataka koji se izravno vežu za njihovo područje djelovanja, što znači da je primjerice dio koji se tiče telefonskih brojeva moguće mijenjati samo od strane ovlaštenih djelatnika telefonske centrale, sama kadrovska služba zadužena je za osobne podatke, djelatnici računalnog centra zaduženi su za podatke o e-mail adresama te o računalima, itd. Na taj način dolazi do podjele odgovornosti za podatke, te se minimizira vjerojatnost pojave netočnih podataka. Također, postojeća kadrovska baza podataka proširuje se elementima zahvaljujući kojima joj je omogućena dvosmjerna komunikacija sa ostalim elementima buduće institutske strukture.

Osim navedenih funkcija, spomenuta baza podataka računalnog centra koja se nalazi u izgradnji predstavlja i svojevrsan katalog svih računala koji se nalaze u domeni IRB-a. Kao takva omogućava praćenje tih računala, instaliranih operativnih sustava te korisnika istih, kako lokalnih tako i onih vanjskih (za računala, odn. operativne sustave koji dozvoljavaju vanjske korisnike), a u nekoj kasnijoj fazi izgradnje, baza će omogućavati i utvrđivanje fizičke lokacije samih računala. Također, baza predstavlja temelj za izgradnju baze znanja koja će u budućnosti svakako predstavljati osnovu za što kvalitetniji i promptniji odziv djelatnika službe za održavanje Računalnog centra Instituta Ruđer Bošković na upite i zahtjeve korisnika. Ta bi baza znanja predstavljala izravnu nadogradnju na bazu djelatnika i računala, a trebala bi sadržavati sve

akcije koje su poduzete nad nekim računalom, tj. trebala bi na prikladan način enkapsulirati korisničke upite i opise problema te detaljno razrađen opis koraka koji su doveli do rješavanja problema. Zahvaljujući tako pohranjenim podacima te korištenjem web tehnologija korisnici će biti u mogućnosti sami pristupiti podacima u bazi, te bi pomoću njih mogli samostalno doći do rješenja problema koji se ponavljaju. Osim navedene uloge, pomoću ove baze moći će se pratiti svi poduzeti zahvati nad bilo kojim računalom na institutu, te će djelatnici računalnog centra na jednostavan način pristupali podacima o svim kvarovima koji su se pojavljivali, a moći će doći i do statističkih podataka o učestalosti pojedinih kvarova kod pojedinih korisnika, što bi se moglo iskoristiti za eventualnu daljnju edukaciju korisnika, a sve u svrhu minimiziranja pojavljivanja dotičnih problema.

Osim navedenog, u procesu implementacije nalazi se i cjeloviti sustav web stranica koje će na krajnje jednostavan i kvalitetan način djelatnicima računalnog centra, ali i svim ostalim djelatnicima instituta te vanjskim suradnicima omogućavati pristup svim važnim informacijama u najkraćem mogućem roku, no pored toga omogućavati će i međusobnu komunikaciju djelatnika, kako putem osobnih poruka među djelatnicima, tako i putem foruma na kojima će se moći razvijati diskusija o raznim problemima (te razmjenjivati iskustva u rješavanju istih), planovima, projektima, itd.

Problem na koji se nailazi već pri površnom pogledu na baze podataka i trenutno korištene sustave za upravljanje bazama podataka je poprilična šarolikost dotičnih sustava te platformi na kojima se dotični sustavi izvršavaju. Nastavak takve tradicije nipošto nije preporučljiv. Kako sa stanovišta znanja, tako i sa stanovišta računalnih resursa, svakako bi bilo prihvatljivija unifikacija dotičnih sustava. Pošto nije svaki sustav za upravljanje bazama podataka "dorastao" potrebama instituta, u početku bi se trebalo odlučiti za sustav koji će na kvalitetan način opsluživati institutske potrebe. Nabavkom kvalitetnog sustava za upravljanje bazama podataka riješili bi se mnogi problemi tehničke prirode, te bi se korištenjem specifičnih svojstava vezanih za dotični sustav za upravljanje bazama podataka omogućila izrada kvalitetnijih aplikacija i servisa, te bi njihova međusobna interakcija bila znatno olakšana.

Neosporno je da je put kojim je Računalni centar Instituta Ruđer Bošković krenuo, dakle put suradnje sa ostalim faktorima na Institutu, distribucije podataka te podjele odgovornosti za održavanje i točnost podataka prema sektorima rada pojedinih djelatnika ispravan put kojim će se osigurati lakše držanje koraka sa galopirajućim razvojem informacijskih znanosti, a u skladu sa razvojem potreba samog Instituta i njegovih djelatnika te relevantnih vanjskih čimbenika. No, trebalo bi razmisliti i o sveobuhvatnoj informatizaciji koja će uključiti sve već povezane djelove, kao i neke potpuno odvojene u jedan sustav koji će svojom cjelovitošću i koherentnošću omogućavati centralno administriranje i distribuiranje podataka prema potrebama, te koji će biti u mogućnosti promptno reagirati na sve eventualne promjene i opsluživati sve zahtjeve koji će se s vremenom pojaviti.

Pred razvojnim je timom Računalnog centra, kao i pred ostalim eventualnim unutarnjim razvojnim timovima i vanjskim suradnicima zahtjev da pri razmatranju i izgradnji (pod)sustava koje razvijaju uzmu u obzir i buduću "globalizaciju" svih tih sustava u jednu sveobuhvatnu, modularnu cjelinu.

piše: Ratko Mileta,
IRB RC



Računarske i informatičke usluge znanstvenicima

Kako bi znanstvenicima olakšao rad i omogućio im da kvalitetnije prezentiraju svoje rezultate Računalni centar pruža ili sprema se pružiti niz usluga na tom području.

Kao prvu istaknuti ćemo multimedijalnu uslugu koju Računalni centar pruža u dvorani III. krila. Dvorana III. krila zamišljena je kao multimedijalna dvorana koja bi trebala služiti za predavanja, skupove, omogućavati prikazivanje powerpoint prezentacija, pristup Internetu, demonstraciji softwera vezanog uz predavanje, gledanju multimedijalnih cd-rom i dvd prezentacija, prikazivanju video traka i slično. Oprema koja se nalazi na raspolaganju sastoji se od računalnog sustava, LCD projektora, digitalne kamere, razglasnog sustava i tri mikrofona. Oprema u gotovo potpunosti zadovoljava gore navedene zahtjeve i često se koristi. No kako Vas ne bi zamarali sa tehničkim detaljima i specifikacijama trenutne opreme (što možete pronaći na web stranicama Računalnog centra) pokušati ćemo Vam prezentirati buduće smjernice razvoja ove dvorane i mogućih tehnologija.

Prvi korak prema poboljšanju ove usluge svakako bi trebalo biti odvajanje dijela opreme od samog predavača i dvorane uz postavljanje fiksne instalacije za što je ide-

alna soba koja se nalazi iznad dvorane i ima sam pogled na nju. Tim korakom dvorana III dosegla bi današnji standard multimedijalnih dvorana. Taj potez donio bi višestruke koristi i znatno bi olakšao korištenje dvorane kao što bi i pridonio sigurnosti same opreme. Idući korak bi svakako trebao biti u pravcu prenošenja događanja iz dvorane III na druge lokacije (testirano prilikom 50 obljetnice IRB-a) i ostvarivanju povratne veze između predavača i auditorija na drugim lokacijama (koristi se npr. na FER-u). Korak u tom pravcu svakako će olakšati već započeta izgradnja nove računalne mreže koja će omogućiti IP/VC na području cijelog Instituta uz pomoć CISCO tehnologije (Cisco IP/VC 3510 Multipoint Control Unit (MCU), the Cisco IP/VC 3520 and IP/VC 3525 Video Conferencing Gateways, and the Cisco IP/VC 3530 Video Terminal Adapter (VTA)). Ta tehnologija omogućavati će višestrukim sudionicima na raznim lokacijama da prisustvuju istom sastanku sa interakcijom u stvarnom vremenu. Postaviti će se javni priključci i kamere za tu svrhu. IP/VC tehnologija primjenjuje se već danas na području poslovnih komunikacija i suradnje, obučavanja kadrova, telemedicine, financija i e-learninga. Nova mreža osim većih brzina omogućavati će i virtualne lanove koji će omogućiti odvajanja



LCD projektor u dvorani 3. krila



HP DesignJet 500 PS Color Plotter

nje broadcast domena čime će se postići manje zagušenje mrežnog linka unutar samog virtualnog lana. Nova mreža neće više biti djeljena (shareana) nego preklapana (switchana) što znači da svi unicast paketi više neće putovati po cijelom IP subnetu već samo na točno određeno mjesto. Samim time razina sigurnosti će biti poboljšana pošto plain tekst passwordi više neće putovati po cijeloj mreži.

Danas u inozemstvu aktualni programi znanstvenog računanja i vizualizacije odmakli su već daleko u područjima high-performance računarstva i mreža, paralelizacije i optimizacije znanstvenog koda, virtualne realnosti i okruženja, interaktivne 3D grafike, vizualizacije znanstvenih podataka, mrežno baziranih video konferencija i alata za zajednički rad, visokokvalitetnih kolor ispisa, real time video recordinga, video produciranja itd. Za potrebe svega toga koriste superkompjuteri, vrhunski serveri i radne stanice uz postojanje brzih mreža. Postoje grafičke i virtual reality usluge sa visoko kvalitetnim uvjetima rada i opremom - ekranima, uređajima za unos, grafičkim radnim stanicama, softwareom za kreiranje virtualnih okruženja, kompjutersko generiranim slikama, animacijama i interaktivnim grafičkim aplikacijama. Također se pružaju usluge za izradu slajdova, ispisa i video zapisa.



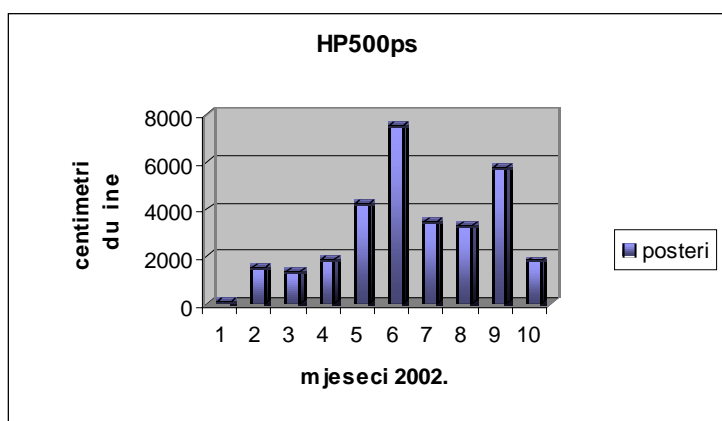
High Resolution: 2048 x 1536
Commodity: PCs, Linux, NEC projectors
Large Scale: 10'W x 7.5'H
Rear Projected
Passive Stereo (3D)

znanstvena i tehnička obrazovanja, distribuiranom računarstvu, ljudskim/računalnim sučeljima, apstraktnoj matematičkoj vizualizaciji itd. Radi se na razvoju VR (Virtual Reality) sustava koji bi podržavali kolaboracijska virtualna okruženja za dizajn, interakciju i vizualizaciju podataka. Interkonektivnost kroz brzu i naprednu mrežu omogućavati će suradnju ne samo znanstvenika međusobno već i suradnju sa obrazovnim, vladinim, industrijskim i drugim institucijama. Već se postavljaju praktični ciljevi micanja teleimerzije (sustav koji omogućuje zemljopisno dislociranim korisnicima da surađuju u simuliranom i dijeljenom virtualnom okruženju u realnom vremenu) iz laboratorija na slijedeću generaciju Interneta.

Kao drugu značajnu uslugu izdvojili bi smo uslugu printanja i plotanja koja postaje sve više popularna među znanstvenicima zbog izrazite potrebe za njom kao i vrlo povoljnim cijenama. Usluga se pruža na HP 4550dn color laserskom printeru i HP DesignJet 500ps color ploteru. Ploter je baziran na HP Thermal Inkjet tehnologiji i plota postere 1200x600dpi do 90 centimetara širine na glossy rolu. Od puštanja u pogon u ožujku 2001 godine ploter isplotao preko 700 metara dužine rola.

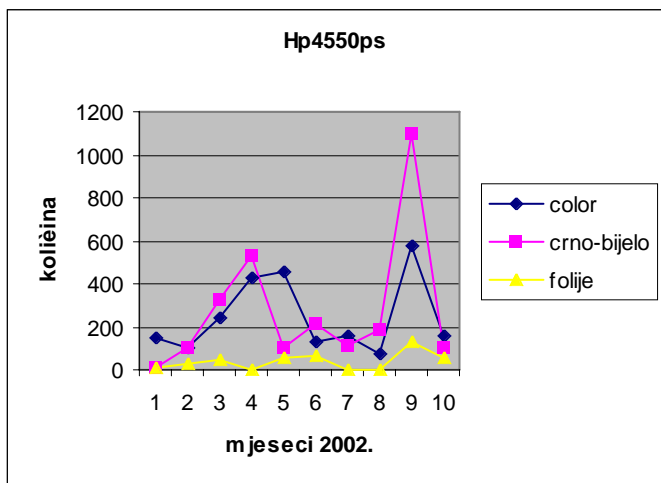
Printer HP 4550dn ispisuje na papir (duplex) i folije a brzina ispisa iznosi 16 stranica u minuti za crno bijele stranice i 4 stranice u minuti za color stranice. Ispis se vrši u rezoluciji od 600x600dpi a koristi se Image Ret tehnologija za kvalitetu boja i foto realistične ispise. Hp ColorSmartII brine se oko toniranja i kontrole boja. Oba proizvoda predstavljaju sam vrh kvalitete takvih proizvoda i idealni su za plotanje postera i printanje drugih znanstvenih radova. Gledajući u budućnost ove usluge nažalost potrebno je prvo riješiti prostorni problem pošto se ova usluga vrši u prostorijama knjižnice V. Krila što je prostorno i lokacijski neadekvatan prostor koji sputava daljnje širenje i poboljšanje kvalitete ove usluge. Planira se poboljšanje usluga nabavkom tiskarskog noža za obradu tiskanog materijala, nabavkom kvalitetnog skenera, nizom druge opreme i pružanjem dodatnih usluga kao što su digitalna obrada fotografija sa svih medija i slično što bi na kraju u budućnosti trebalo ovu uslugu zaokružiti kao uslugu digitalnog tiska i samim time je još više povezati sa multimedijalnim uslugama dvorane III.

Treća usluga vezana je uz mrežne printere i zapravo koncepcijski vrlo je jednostavna. Trenutno u praksi funkcionira u I. Krilu gdje se nalazi HP 4100dtn mrežni printer na kojeg korisnici direktno preko mreže sami printaju. Ideja i plan ove usluge je da se ona proširi na cijeli institut tako da se u svakom većem krilu nalazi mrežni printer na raspolaganju znanstvenicima uz postojanje centralnog accountinga. Omogućilo bi se printanje sa Linux i Windows računala na bilo koji javni printer naravno uz odgovarajuća ovlaštenja.



Dužina isplotanih postera po mjesecima

Gledajući u budućnost kao i na druge znanstvene i obrazovne ustanove postoji još puno tehnologija koje bi se mogle implementirati ne samo u sklopu dvorane III. već i cijelog Instituta. Riječ je o povezivanju više dvorana, učionica i konferencijskih soba u jedan sustav koji poticao interakciju a bio bi vizualno i elektronički povezan omogućavajući prijenos zvuka, slike i računalnih podataka između prostorija. Pritom se tako komplicirani sustav može izgraditi a da ne zahtjeva velik broj kadrova za upravljanje zbog sistema centralne kontrole kakve npr. proizvodi Crestron. Gledajući dalje u budućnosti danas je velik broj istraživanja usmjeren prema znanstvenoj vizualizaciji, novim metodologijama za



Broj isprintanih stranica po mjesecima

Zaključak

nastavak sa str. 19

Nove tehnologije Internet računarstva i Grida do sada su pokazale velik potencijal za provođenje znanstvenih istraživanja koja zahtijevaju veliku računalnu snagu. Daljnjim sazrijevanjem tehnologija, izgradnjom i prihvaćanjem odgovarajućih konvencija i protokola Internet računarstvo i Grid polako će postati sastavni dio velikih znanstvenih projekata u kojima će se problemi različitih profila upućivati odgovarajućim računalnim resursima: složeni proračuni izvršavat će se na moćnim super-računalima izgrađenim od povezanih grozdova računala, dok će se jednostavniji, ali ne i manje zahtjevni problemi analize podataka, izvršavati u "spavajućem" načinu rada na stolnim računalima diljem svijeta.

Literatura:

Ian Foster
Internet Computing and Emerging Grid
7.12.2002.
<http://www.nature.com/nature/webmatters/grid/grid.html>

SETI@home projekt
<http://setiathome.ssl.berkeley.edu>

Grid Physics Network project
www.griphyn.org

European Data Grid
grid.web.cern.ch/grid

Particle Physics Data Grid
www.ppdg.net

Network for Earthquake Engineering Simulation Grid
www.neesgrid.org

The Globus Project
www.globus.org

The Global Grid Forum
www.gridforum.org

Grid outreach
www.gridoutreach.org.uk

Računalna sigurnost

CARNet CERT (Computer Emergency Response Team) na svojim web stranicama redovito obavještava o ranjivostima i sigurnosnim problemima na svim operacijskim sustavima.

Ne postoji softver za koji bi se moglo reći da je 100% siguran i bez bugova. Zato je organizacija kao što je CERT od presudne važnosti za računalnu sigurnost. O ciljevima i misiji CERT-a možete pročitati na http://www.cert.hr/o_certu.html. Na njihovim web stranicama možete pronaći najnovije sigurnosne novosti, pročitati sigurnosna upozorenja, prijaviti napade na vlastiti sustav... Odnedavno su na webu mnogi dokumenti (<http://www.cert.hr/dokumenti.html>) u pdf formatu koji mogu pomoći sistemcima, ali i običnim korisnicima da bolje razumiju sigurnosne probleme, zaštite se od napada, te pravilno reagiraju u slučaju da budu napadnuti.

Ukoliko želite redovito putem e-mail-a primati sigurnosne obavijesti za jedan ili više operacijskih sustava, posjetite http://www.cert.hr/mailling_liste.php i prijavite se na neku od mailing lista.

Za svoje korisnike CARNet CERT je izdao brošuru pod nazivom "20 najčešćih računalno sigurnosnih problema". Brošuru možete besplatno downloadati sa <http://www.cert.hr/top20/index.html>.

U zadnje vrijeme smo svjedoci poplave neželjenih e-mail-ova koji upozoravaju na viruse koji ne postoje sugerirajući korisnicima da brišu određene datoteke na računalo. Obično se radi i važnim sistemskim datotekama bez kojih operacijski sustav nije u stanju normalno raditi. CERT nudi pomoć i oko tih tzv. 'hoax' poruka: <http://www.cert.hr/hoax.html>.

Osim hrvatskog, korisno je pogledati i međunarodni CERT na <http://www.cert.org/>.

Uz redovitu nadogradnju antivirusnog softvera, CERT web stranice bi trebale biti redovita referenca za sve korisnike čija su računala stalno ili povremeno spojena na Internet. Redovita nadogradnja problematičnog softvera uvelike smanjuje mogućnost da vam računalo bude napadnuto, a podaci oštećeni ili ukradeni.

I na kraju zapamtite: nesreća se ne događa nekom drugom.

Mario Pranjić

impressum:

Znanstveno glasilo
Instituta "Ruđer Bošković"
Bijenička c. 54, 10 002 Zagreb
tel: +385 (0)1 4561 111,
fax: 4560 084
e-mail: rudjer@rudjer.irb.hr
URL: <http://www.irb.hr>

Glavni urednik:
Mislav Jurin

Urednik ovog tematskog broja:
Karolj Skala

Uredništvo:
Velimir Bardek
Dunja Čukman
Koraljka Gall-Trošelj
Kata Majerski
Iva Melinščak-Zlodi
Tvrko Smital
Jadranka Stojanovski

Digitalna obrada i izvedba:
Institut Ruđer Bošković
Grafički fakultet u Zagrebu

ISSN 1333-5693
UDK 061.6:5

Tisak: Kratis d.o.o.
Izlazi mjesečno u nakladi od 600
primjeraka uz financijsku potporu
Instituta

SURADNJA

ERICSSON
Ericsson Nikola Tesla d.d.

&



&



Dugoročna suradnja na Grid tehnologiji

Potpisan je Ugovora o suradnja u znanstvenoistraživačkoj i razvojnoj djelatnosti iz područja Grid tehnologije između tvrtke Ericsson Nikola Tesla i Instituta Ruđer Bošković. Suradnjom se obuhvaćaju:

- stvaranje zajedničkih znanstveno-stručnih timova kompetentnih za rješavanje najsloženijih poslova u računarstvu i informatici,
- razvoj, primjena i testiranje programske podrške u području grid tehnologije,
- suradnja na projektima u Hrvatskoj i u inozemstvu.

Gornja suradnja će se ostvariti u znanstvenoistraživačkom radu kroz zajedničke projekte i sudjelovanje u projektima. Ugovorene strane smatraju sveukupne aktivnosti na području distribuiranog računarstva i Grid tehnologije izuzetno značajnim za njihov razvitak u Republici Hrvatskoj te za međunarodnu suradnju na polju istraživanja, razvoja, proizvodnje i primjene generičke Grid tehnologije. Polazeći od toga stavlja se poseban naglasak na razvoj novih primjena kao i testiranje (testbed) u realnim uvjetima rada u Grid okruženju za koje će se ostvariti uvjeti na IRB-u. Dugoročna suradnja se proteže i na edukaciji i razvoju know-howa na području Grid tehnologije.

Suradnja s perspektivom

Početak godine uspostavljena je suradnja IRB-a i IBM-a Hrvatska, podružnice svjetske korporacije IBM. Najintenzivniji kontakti trenutno su vezani za područje GRID tehnologija, obzirom da je IBM predvodnik u razvoju ove još mlade grane komunikacijskih tehnologija, a IRB je jedan od začetnika hrvatskih inicijativa na tom području. Zajednički interes je utoliko veći što IBM usko suraduje s akademskim ustanovama širom svijeta na razvoju GRID rješenja i jedan je od najvećih kontributora open source inicijativa.

Nakon službenog posjeta dr. Skale predstavnika Instituta IBM-ovom razvojnom laboratoriju i jednom od GRID centara u Hursleyu u Velikoj Britaniji u svibnju ove godine, iskazan je obostran interes za daljnju suradnju, te dogovorena stručna pomoć i podrška na planiranim projektima IRB-a. U sklopu tog dogovora ljetos je Syd Chapman, Istaknuti izumitelj u IBM-u, održao za zainteresirane djelatnike Instituta predavanje o trenutnom stupnju razvoja GRID tehnologije, iskustvima u drugim akademskim zajednicama, iskušeni-

ma s kojima se trenutno suočavaju oni najnapredniji, te pogledom na budućnost ove komunikacijske infrastrukture budućnosti.

Trenutno se razmatra mogućnost uključanja IRB-a u međunarodne projekte (jedan od kojih je AccessGrid, u koji je uključeno sveučilište iz Southamptona) kojima IBM daje tehnološku i stručnu podršku, a kroz koje bi IRB mogao razmjenjivati znanstvene i tehnološke potencijale s ostalim suradnicima u projektu, istovremeno otvarajući nove mogućnosti neposrednog sudjelovanja u izgradnji rješenja nužnih za funkcioniranje buduće svjetske infrastrukture GRID-a.

IBM je u svijetu dugo godina prisutan i poznat i u drugim područjima znanosti, poput genetskih, medicinskih i farmakoloških istraživanja, a istraživački laboratoriji IBM-a sami su prošle godine "proizveli" više od 3.000 patenata. U razgovorima s predstavnicima IBM-a Hrvatska ukazano je na zajednički interes suradnje i na drugim područjima djelovanja Instituta.

Popis informatičkih i računalnih projekata na IRB-u

(kronološki redoslijed)

mr. sc. Jadranka Stojanovski
Sustav znanstvenih informacija RH - tematski podsustav Prirodoslovlje, 1995 do danas

mr. sc. Jadranka Stojanovski
Hrvatska znanstvena bibliografija (CROSB), 1998 do danas

mr. sc. Jadranka Stojanovski
Centar za online baze podataka (CARNet), 1996 do danas

prof. dr. sc. Karolj Skala
Izgradnja institutskog LAN-a, MZT, 2001-2002

Ivana Pazar, dipl. bibl.
EJOL (Electronic Journal Online Library), 2001-2002

mr. sc. Jadranka Stojanovski
"Tko je tko u znanosti u Hrvatskoj?" 2001-2002

prof. dr. sc. Karolj Skala
Sustav računalne mreže s laserskim prijenosom podataka u slobodnom prostoru, 2001-2003

prof. dr. sc. Nikola Bogunović
Automatizirano otkrivanje znanja i rasuđivanje, 2002-2005, Projekt 0098023

prof. dr. sc. Nenad Trinajstić
Razvoj i primjena modela u kemiji i bioinformatici, 2002-2005, Projekt 0098034

prof. dr. sc. Karolj Skala
Development of New Grid Technologies for Advanced Scientific and Engineering Application, Institute for Software Science, University of Viena, MZT, Austrijsko-Hrvatski projekt, 2002.

dr. sc. Branko Pivac
Računalo podržana automatizacija, udaljeno upravljanje i moni-

toriranje sistema za mjerenje vodljivosti i Hall efekta, Informatički projekt MZT, 2002-138

dr. sc. Ivan Marić
Međunarodni standardi za mjerenje protoka - aktivni mrežni objekti, Informatički projekt MZT, 2002-060

prof. dr. sc. Karolj Skala
Mrežni nadzor i IC upravljanje klaster računalom, Informatički projekt MZT, 2002 021

dr. sc. Mario Stipčević
Daljinska kontrola procesa i mjerenja putem mobilne telefonije i Interneta, Informatički projekt MZT, 2002-007

dr. sc. Tomislav Šmuc
Informacijski portal EH-UH, Informatički projekt MZT, 2002-057

prof. dr. sc. Karolj Skala
CRO GRID, složeni tehnološki projektni prijedlog u razmatranju

