

DEZVOLTAREA SITE-ULUI RO-01-ICI CA NOD ÎN INFRASTRUCTURA NAȚIONALĂ GRID PENTRU CERCETARE

ing. Gabriel Neagu

gneagu@ici.ro

ing. Alexandru Stanciu

alex@ici.ro

Institutul Național de Cercetare-Dezvoltare în Informatică – ICI București

Rezumat: La nivel european e-Infrastructurile reprezintă un pilon esențial pentru implementarea conceptului ERA (*European Research Area*). Principalele componente ale acestor infrastructuri sunt rețeaua pentru educație și cercetare și rețeaua de centre de resurse Grid. Atenția acordată pe plan național dezvoltării acestor componente ale e-Infrastructurilor este confirmată de Raportul pe 2007 al Comitetului Român pentru infrastructurile cercetării, care include în lista de propuneri pentru proiecte prioritare investiții în domeniul Tehnologiei Informației și Comunicațiilor proiectele privind rețeaua de comunicații de mare viteză pentru educație și cercetare și, respectiv, infrastructura națională Grid pentru cercetare.

Programul operațional sectorial pentru creșterea competitivității economice (POS-CCE) oferă un substanțial suport financiar pentru dezvoltarea acestor proiecte prin apelul de proiecte organizat în anul 2008. ICI București se numără printre beneficiarii acestui suport, propunerea de proiect depusă la acest apel fiind aprobată spre finanțare.

Lucrarea prezintă contextul european de dezvoltare a e-Infrastructurilor, evoluția și stadiul actual al infrastructurii de centre de resurse Grid la nivel național, obiectivele POS-CCE relevante pentru susținerea acestui domeniu. Ca studiu de caz, este prezentat succint proiectul menționat mai sus, finanțat din fonduri structurale, care va implementa un site de mare putere pentru infrastructura națională Grid.

Cuvinte cheie: e-Infrastructură, EGEE, SEE-GRID, EGI, infrastructura națională Grid, fonduri structurale, POS-CCE, site-ul Grid RO-01-ICI.

Abstract: At the European level e-Infrastructures are considered a vital pillar for the implementation of the ERA (*European Research Area*) concept. The main components of these infrastructures are the network for education and research and the Grid infrastructures. At the national level a confirmation of the attention paid to the development of these e-Infrastructure components is provided by the 2007 Report of the Romanian Committee for Research Infrastructures, where projects regarding high-speed communication network for education and research, and national Grid infrastructure for research are included in the list of proposals for priority investment projects in the Information and Communications Technologies domain.

Sectorial Operational Programme “*Increase of Economic Competitiveness*” (SOP IEC) has been providing a consistent financial support for the development of these projects through the call organized in 2008. ICI Bucharest is among beneficiary organizations of this support as its project proposal submitted to this call was accepted for funding.

This paper presents the European context for e-Infrastructures development, the evolution and current status of Grid resource centers at the national level, the relevant objectives of the SOP IEC programme regarding the support of this domain. As a cases study, the paper provides also a short description of the above mentioned structural funds project, which will implement a high capacity site for the national Grid infrastructure.

Keywords: e-Infrastructure, EGEE, SEE-GRID, EGI, national Grid infrastructure, structural funds, SOP IEC, RO-01-ICI Grid site.

1. Introducere

Spațiul European de Cercetare (ERA – European Research Area) reprezintă unul din elementele de bază ale Strategiei de la Lisabona pentru creștere economică și ocuparea forței de muncă, care își propune să construiască în Europa cea mai competitivă și dinamică economie bazată pe cunoaștere la nivel mondial. Conform [1], principalele caracteristici ale ERA sunt:

- un flux adecvat de cercetători competenți, cu un nivel ridicat de mobilitate între instituții, discipline și țări;
- infrastructuri de cercetare de nivel internațional, integrate, interconectate și accesibile echipelor de cercetare din toată Europa și din lume, în special datorită noilor generații de infrastructuri de comunicații electronice;

- instituții de excelență în cercetare, angajate în cooperare și în parteneriate public-privat, „comunitățile virtuale de cercetare”, specializate în domeniile interdisciplinare și care atrag o masă critică de resurse;
- un schimb de cunoștințe eficient, în special între cercetarea publică și industrie, precum și cu publicul larg;
- programe și priorități de cercetare bine coordonate, inclusiv un volum semnificativ de investiții în cercetarea publică, programate în comun la nivel european, implicând priorități, implementare și evaluare comune.

Sub aspect financiar, principalul suport pentru implementarea acestui concept este oferit de Programul Cadru 7, Programul Cadru pentru competitivitate și inovare, programele pentru educație și instruire, fondurile structurale și cele de coeziune pentru convergență și competitivitate.

Din punct de vedere tehnic, în implementarea ERA un rol esențial revine e-Infrastructurilor, al căror potențial în accelerarea utilizării ICT în cercetare se referă la constituirea spațiului comun de informare la nivel european, la consolidarea cercetării în domeniul ICT și implementarea rezultatelor sale, la promovarea societății informaționale accesibile pentru toți.

În mod evident, contribuția cercetătorilor din România la realizarea obiectivelor ERA și valorificarea avantajelor pe care acest concept le oferă depinde în mare măsură de compatibilizarea orientărilor de dezvoltare și a nivelului de performanță al e-Infrastructurii naționale cu cele existente pe plan european. Lucrarea de față prezintă un proiect dedicat acestui deziderat, aflat în derulare la ICI București.

Capitolul 2 trece în revistă contextul european actual privind e-Infrastructurile de cercetare, cu accent pe documentele programatice care fundamentează o dezvoltare durabilă a domeniului și pe organismele responsabile la nivel european pentru formularea de politici și recomandări privind evoluția acestui domeniu. A doua parte a capitolului este dedicată prezentării stadiului actual al infrastructurii Grid pentru cercetare la nivel european, pe exemplul oferit de proiectele europene EGEE și SEE-GRID, la care participă și consorții de organizații din România.

Capitolul 3 descrie succint evoluția și stadiul actual al infrastructurii Grid la nivel național, documentele cu caracter strategic care au stat la baza acestui proces, contribuția consorțiului RoGrid, coordonat de ICI București.

Capitolul 4 este dedicat prezentării site-ului RO-01-ICI, care funcționează în regim de producție din anul 2004, în infrastructurile EGEE și SEE-GRID. Suportul de finanțare a dezvoltării acestui site la nivelul actual de cerințe pentru un site de infrastructură națională, obiectivele proiectului respectiv, planul de activități și potențialii beneficiari ai rezultatelor proiectului fac obiectul capitolului 5.

În final sunt prezentate concluziile lucrării și lista de referințe bibliografice.

2. Contextul european

2.1 e-Infrastructuri

Conceptul „e-Infrastructuri” a fost lansat în 2003 pentru a marca viziunea referitoare la următoarea generație de infrastructuri ICT trans-naționale de cercetare la nivel european. În consecință, termenul „e-Infrastructuri” se referă la un nou mediu de cercetare, în care toți cercetătorii au acces partajat la facilități de natură științifică unice sau distribuite, indiferent de natura sau locația acestora.

Integrând tehnologii informaționale și de comunicații de vârf, ca de exemplu rețelele de bandă largă și mobile, calculul Grid, tehnologii pentru aplicații și servicii orientate Internet și Web, e-Infrastructurile joacă un rol strategic în schimbarea modului de derulare a activității de

cercetare în diverse domenii științifice și ingineresti, dar și în afaceri.

Cele trei niveluri principale ale e-Infrastructurii – resurse de calcul și memorare, comunicații și servicii – asigură suportul pentru implementarea unor forme virtuale de colaborare, ca de exemplu centrele virtuale de excelență sau laboratoarele virtuale de cercetare.

Din bugetul Programului FP7 pentru infrastructuri de cercetare (1.715 MEuro), 33 % revine componentei e-Infrastructurii (572 MEuro).

În recenta Comunicare a Comisiei Europene către Parlamentul European, Consiliu, Comitetul Economic și Social European și Comitetul Regiunilor, intitulată „*Infrastructuri ICT pentru e-știință*” [2], e-infrastructurile sunt definite ca un mediu al resurselor cercetării care include rețele, grid-uri și middleware, resurse de calcul, bănci de lucrări experimentale, colecții de date, instrumente și sprijin operațional pentru colaborarea mondială virtuală în domeniul cercetării.

Acțiunile propuse de acest document statelor membre ale Uniunii Europene pentru asigurarea dezvoltării strategice a acestui domeniu vizează: consolidarea supremației mondiale a rețelei GÉANT, structurarea cadrului de dezvoltare a infrastructurii Grid pentru e-știință, ameliorarea accesului la informațiile științifice, construirea unei noi generații de instalații de calcul intensiv, găzduirea comunităților mondiale virtuale de cercetare. În domeniul Grid, aceste acțiuni se referă la consolidarea și dezvoltarea în continuare a inițiativelor naționale Grid (NGI) ca bază pentru o strategie europeană reînnoită, precum și la tranziția spre noi modele de gestionare a grid-urilor de e-știință europene, bazate pe interoperabilitate, pentru a servi unei game largi de domenii de cercetare.

În ceea ce privește politicile de dezvoltare a e-Infrastructurilor, rolul esențial revine *e-Infrastructure Reflection Group (e-IRG)* și *European Strategy Forum on Research Infrastructures (ESFRI)*.

Misiunea e-IRG este să faciliteze evoluția către o e-Infrastructură europeană de utilitate generală. e-IRG (www.e-irg.eu) este compus din delegați desemnați de ministerele de cercetare din țările membre și are ca principal obiectiv formularea de politici și recomandări privind evoluția acestui domeniu. Documentul intitulat *Cartea albă pe 2009* [3], publicat recent de e-IRG, evidențiază următoarele subiecte de actualitate pentru dezvoltarea acestui domeniu:

- colaborarea globală: suportul oferit comunităților de utilizatori din diverse domenii științifice în cadrul unor modele de colaborare mai fluide și al lipsei unor structuri organizatorice clar definite;
- educație și instruirea în utilizarea infrastructurii: recomandări și acțiuni pentru armonizarea și standardizarea la nivel ERA a competențelor și specializărilor în calculul distribuit, în concordanță cu procesul Bologna;
- Grid și „Cloud Computing”: studierea oportunităților oferite de integrarea celor două paradigme de calcul și a posibilităților de aplicare pentru domeniul „Cloud Computing” a rezultatelor obținute până în prezent de inițiativele Grid;
- abordarea holistică a problemei securității: stabilirea unor structuri inter-organizaționale pentru managementul incidentelor de securitate prin servicii și soluții reactive și proactive care să acopere toate nivelurile e-Infrastructurii;
- virtualizarea: integrarea tehnologiilor de virtualizare în e-Infrastructurile de cercetare și încurajarea adoptării unor standarde deschise în domeniu;
- instrumentar la distanță: facilitarea accesului comunităților de utilizatori la diverse componente ale infrastructurii de cercetare, aflate la distanță, integrarea cu middleware-ul Grid și elaborarea standardelor aferente;
- sustenabilitatea e-Infrastructurii: analiza modelelor de organizare și de finanțare adoptate de furnizorii și utilizatorii de resurse.

ESFRI (<http://cordis.europa.eu/esfri/>) este tot un forum al reprezentanților țărilor membre UE plus un reprezentant al Comisiei Europene, cu rolul de a promova o abordare coerentă în domeniul deciziilor de politică a dezvoltării infrastructurilor de cercetare la nivel european și de a acționa ca incubator pentru negocierile internaționale privind inițiative concrete în acest domeniu. ESFRI elaborează și foia de parcurs europeană în domeniul noilor infrastructuri de cercetare.

2.2 Infrastructuri Grid pentru cercetare la nivel european

Primul proiect reprezentativ la nivel european dedicat infrastructurii Grid a fost proiectul FP5 *DataGrid* (<http://eu-datagrid.web.cern.ch/eu-datagrid/>), derulat în perioada 2001-2003 de un consorțiu format din 6 parteneri, sub coordonarea CERN Geneva. În continuarea acestui proiect, începând cu anul 2004, cadrul de dezvoltare a infrastructurii Grid la nivel european a fost asigurat de proiectul european EGEE – “*Enabling Grids for E-science*” (www.egee.org), finanțat parțial prin programele FP5 și FP6 ale Uniunii Europene.

Principalele activități derulate în cadrul proiectului EGEE se referă la: dezvoltarea de middleware, operarea infrastructurii Grid la nivel european, migrare aplicații în mediul Grid și asistență tehnică pentru acest gen de activitate, suportul organizațiilor virtuale Grid dedicate diverselor comunități de utilizatori, training, diseminare și colaborare internațională. La etapa curentă a proiectului (mai 2008 – aprilie 2010) participă aproximativ 140 de organizații din 33 de țări. România este reprezentată printr-un consorțiu de 7 organizații membre ale RoGrid, coordonat de ICI București. Nivelul actual de dezvoltare al infrastructurii EGEE este ilustrat de următorii parametri [4]: 140.000 de core-uri, 260 de site-uri, 25PB memorie disc, 39 PB memorie bandă, 12 milioane job-uri pe lună (cu o creștere de 45% în ultimul an), aproximativ 13.000 de utilizatori. Domeniul fizicii energiilor înalte reprezintă cel mai important contribuitor și utilizator al acestei infrastructuri, dar și alte domenii au o dinamică ridicată de creștere: astronomia și astrofizica (53 aplicații, 20 VO-uri și 373 utilizatori), științele vieții (46 aplicații, 9 VO-uri și 379 utilizatori), chimia computațională (12 aplicații, 4 VO-uri și 347 utilizatori), științele Pământului (15 aplicații, 7 VO-uri, 142 utilizatori).

EGEE are relații de colaborare cu 27 de proiecte Grid din Europa și din alte zone, între care un important grup îl formează proiectele regionale (26 VO-uri, 1658 de utilizatori). Între acestea se numără și inițiativa regională din Europa de sud-est SEE-GRID, care are ca obiectiv dezvoltarea unei infrastructuri regionale Grid pe baza transferului de tehnologie și de soluții de administrare din proiectul EGEE. În prezent, această inițiativă este materializată prin proiectul FP7 SEE-GRID-SCI – “*SEE-GRID eInfrastructure for regional eScience*” (www.see-grid-sci.eu), al cărui consorțiu include 44 de organizații din 14 țări din regiune. Prin prisma obiectivului menționat, rolul principal revine țărilor care participă și la proiectul EGEE: Bulgaria, Grecia, România, Serbia și Turcia. În prezent infrastructura SEE-GRID include 40 de site-uri de producție, 1086 core-uri dedicate exclusiv acestei infrastructuri și 288 TB capacitate de stocare. Pe lângă oferirea de servicii de infrastructură Grid în regim de producție, acest proiect mai include două activități prioritare [5]: (a) constituirea de comunități internaționale de utilizatori ai acestor servicii în domeniile seismologie, meteorologie, protecția mediului și oferirea de aplicații pentru aceste domenii, (b) sprijinirea inițiativelor naționale Grid (NGI-uri) din țările reprezentate în consorțiul proiectului, pentru aderarea la noul model de organizare și funcționare a infrastructurii Grid la nivel european. În concepția acestui proiect, NGI reprezintă un consorțiu deschis, format din entități juridice sau o entitate juridică acționând în numele acestora, care, în beneficiul comunității academice și de cercetare, coordonează, promovează și implementează activități Grid la nivel național, cu accent pe administrarea și operarea infrastructurii Grid, în conformitate cu strategia națională sau cu programul național de cercetare în acest domeniu.

Noul model menționat se bazează pe colaborarea între Inițiativa Europeană Grid (EGI) și NGI-urile din țările participante, urmând să înlocuiască modelul EGEE, după finalizarea acestui proiect, în aprilie 2010. Detalierea acestui model și elaborarea graficului de tranziție EGEE –

EGI sunt asigurate de proiectul EGI_DS – “*EGI Design Study*” (web.eu-egi.eu). Conform documentului [6], scopul global al acestui model îl reprezintă asigurarea sustenabilității pe termen lung a infrastructurilor Grid din Europa, prin federalizarea acestora într-o infrastructură Grid multi-națională și multi-disciplinară la nivel european.

3. Stadiul actual al infrastructurii naționale Grid

Într-o primă perioadă, dezvoltarea rețelei de centre de resurse Grid în țara noastră a avut la bază participarea organizațiilor academice și de cercetare la proiecte naționale sau internaționale de profil, cu aportul principal al organizațiilor membre ale consorțiului RoGrid: I.C.I. București – coordinator, Universitatea Politehnică București (UPB), Institutul Național de Cercetare – Dezvoltare pentru Fizică și Inginerie Nucleară “Horia Hulubei” (IFIN-HH) București, Institutul Național de Cercetări Aerospațiale (INCAS) București, Universitatea București (UB), Universitatea Tehnică din Cluj-Napoca (UTCN), Universitatea de Vest din Timișoara (UVT).

Un rol important în maturizarea acestor centre și în adaptarea lor la cerințele de funcționare în regim de producție (24/7) l-a avut participarea în infrastructuri Grid la nivel european, neutre din punct de vedere al domeniului de utilizare (EGEE, SEE-GRID) sau dedicate experimentelor CERN în domeniul fizicii energiilor înalte.

O primă încercare de formulare a unor principii și obiective unitare de dezvoltare pentru acest domeniu a reprezentat-o propunerea “*Strategia națională privind dezvoltarea Grid în România* (decembrie 2002), proiect elaborat de consorțiul RoGrid (ICI, UPB, IFIN-HH, INCAS, UB) [7]. Obiectivul general al acestei strategii l-a reprezentat organizarea, dezvoltarea și punerea în exploatare a unei infrastructuri Grid în România, în vederea măririi accesibilității, fiabilității, scalabilității, securității și eficienței utilizării aplicațiilor științifice și comerciale complexe. Obiectivele specifice complementare vizau: organizarea și dezvoltarea la nivel național a unei infrastructuri de înaltă performanță, compatibilă cu rețele de tip Grid în curs de implementare în Europa și SUA; creșterea potențialului științific și a capacităților existente în domeniul software aplicativ, prin integrarea într-un sistem global Grid; promovarea în societatea românească a noilor tehnologii specifice Grid, orientarea mediului academic, de cercetare și afaceri către noul mediu virtual global. Formularea acestor obiective a avut la bază următoarele principii: realizarea unei infrastructuri deschise; asigurarea unui echilibru între investiția în infrastructura hardware și software specifică și dezvoltarea de software aplicativ; valorificarea producției proprii de software științific, promovarea activă a rezultatelor cercetării; asigurarea condițiilor de penetrare și utilizare a infrastructurii RoGrid în domenii ale vieții economice și sociale.

În martie 2006 a fost publicat “*Planul strategic pentru Proiectul Pilot de implementare a infrastructurii naționale Grid*”, având la bază conceptul „e-Infrastructuri” și recomandările formulate în strategia din 2002 [8]. Documentul a fost elaborat de un „task force” național aflat în coordonarea Autorității Naționale pentru Cercetare Științifică (ANCS), cu aportul principal al consorțiului RoGrid și al RoEdu.

În conformitate cu acest plan strategic, Raportul pe anul 2007 al Comitetului Român pentru Infrastructurile Cercetării [9] a inclus în lista de propuneri pentru proiecte de investiții prioritare, la capitolul “Tehnologia Informației și Comunicații”, proiectele *Rețeaua de comunicații de mare viteză pentru educație și cercetare* și *Infrastructura națională Grid pentru cercetare (RoReGI)*. Conform acestui document, proiectul RoReGI are ca obiectiv organizarea, implementarea și operarea infrastructurii Grid pentru cercetare în România, dedicată dezvoltării și utilizării aplicațiilor științifice de mare complexitate, prin asigurarea accesului sigur al comunităților virtuale de utilizatori la resursele de calcul și de memorie de care au nevoie, în condiții de fiabilitate, scalabilitate, securitate și eficiență în funcționare. Infrastructura Grid va include centre de resurse (site-uri) Grid de capacitate mare și medie, inclusiv software specializat (middleware), conectate prin rețeaua de comunicații de mare viteză pentru educație și cercetare. Proiectul este prevăzut să se desfășoare în trei etape:

a) instalarea site-urilor de mare capacitate în localitățile în care sunt plasate nodurile principale ale rețelei RoEduNet de mare viteză, pentru a asigura interconectarea acestora la 10Gbps;

b) implementarea serviciilor de management operațional al infrastructurii la nivel național;

c) extinderea infrastructurii (etapă opțională, dimensionată în funcție de nivelul de încărcare al primului tronson), prin instalarea de site-uri de capacitate medie, în corelare cu dezvoltarea rețelei de mare viteză și cu dinamica cerințelor de utilizare.

Principalele direcții de impact ale implementării infrastructurii naționale Grid se referă la: valorificarea experienței existente în centrele active de competență în domeniu și reutilizarea rezultatelor obținute în colaborarea națională și internațională în domeniu; dezvoltarea și valorificarea potențialului la nivel național în cercetarea-dezvoltarea de aplicații dependente de resurse importante de prelucrare și / sau memorare date; concentrarea resurselor financiare la nivelul organizațiilor beneficiare către rezolvarea problemelor specifice domeniului și mai puțin către upgradarea continuă și administrarea infrastructurii TI; constituirea la nivel național a comunităților virtuale de utilizatori ai resurselor RoReGI; creșterea vizibilității și accesibilității ofertei de aplicații în domeniile beneficiare ale infrastructurii Grid; asigurarea premiselor de participare activă la colaborarea științifică internațională în aceste domenii.

În prezent, în nomenclatorul la nivel național sunt înregistrate 16 site-uri Grid, numărul acestora reflectând interesul în creștere din marile centre universitare și de cercetare pentru utilizarea tehnologiilor Grid, nu numai în domenii cu tradiție din acest punct de vedere (fizica nucleară), dar și în modelare matematică, biomedicină, chimie, protecția mediului, seismologie.

11 site-uri funcționează în regim de producție: RO-01-ICI, RO-02-NIPNE, RO-03-UPB, NIHAM (alias RO-04-NIPNE), RO-07-NIPNE, RO-08-UVT, RO-09-UTCN, RO-11-NIPNE, RO-13-ISS (Institutul pentru Știința Spațiului București), RO-14-ITIM (Institutul Național de Cercetare – Dezvoltare pentru Tehnologii Izotopice și Moleculare Cluj-Napoca), RO-15-NIPNE. Aceste site-uri sunt active în infrastructura EGEE (11) și SEE-GRID-SCI (4). Celelalte 5 site-uri sunt în diverse stadii de instalare / operaționalizare: RO-12-ICI, RO-05-INCAS, RO-06-UB, RO-10-TUIASI (Universitatea Tehnică „Gh. Asachi” Iași), RO-16-UAIC (Universitatea „A.I. Cuza” Iași).

4. Site-ul RO-01-ICI

Prima versiune a site-ului Grid RO-01-ICI a fost realizată la sfârșitul anului 2003, pentru susținerea activității colectivului de cercetare în domeniul Grid, domeniu promovat prin strategia de cercetare – dezvoltare a institutului, în concordanță cu obiectivele Programului național de cercetare și ale programelor cadru europene FP6 și FP7. Configurația inițială (4 procesoare), realizată prin efort investițional al institutului, a fost în continuare upgradată, în limita bugetelor de achiziții echipamente ale proiectelor naționale de cercetare derulare de colectivul de specialitate din ICI București: *“IComGrid - Infrastructură pilot de comunicații pentru Grid”* (contract RELANSIN nr. 2003/2004, derulat în perioada 2004-2006), *„GridMOSI - Organizație virtuală în tehnologie Grid pentru modelare, simulare și optimizare de înaltă performanță”* (contract CEEEX nr. 95/2005, derulat în perioada 2005-2008), *“PEGAF – Platformă experimentală Grid pentru dezvoltarea de aplicații orientate pe fluxuri de activități cu alocarea dinamică a resurselor”* (contract PN II nr.11064/2007, aflat în derulare). Aceste proiecte au asigurat și continuitatea activităților de întreținere și administrarea a site-ului.

Începând cu anul 2004, colectivul de cercetare a participat la proiectele europene EGEE și SEE-GRID, fapt care a permis acumularea unei experiențe importante în activitatea de administrare resurse și infrastructuri Grid. În luna mai 2004, site-ul a fost certificat în infrastructura regională, iar în noiembrie 2004 a fost primul site din România înregistrat în infrastructura EGEE, intrând în regimul de producție implementat în acest proiect: funcționare 24 de ore pe zi, 7 zile pe săptămână, reacție rapidă la incidente în funcționare, asigurarea

continuă a compatibilității cu platforma software utilizată în infrastructura EGEE (sistem de operare și middleware).

Pe baza acestei experiențe, în cadrul proiectului CEEX GridMOSI, coordonat de ICI București, s-a reușit realizarea unei infrastructuri Grid funcționale incluzând 5 site-uri, cu nivel ridicat de autonomie față de infrastructurile externe, ceea ce a demonstrat fezabilitatea implementării în viitor a unei infrastructuri naționale Grid. De asemenea, în cadrul acestui proiect a fost instalată și înregistrată în EGEE prima organizație virtuală din țara noastră, dedicată domeniului modelării și simulării de înaltă performanță – Gridmosi.ici.ro.



Figura 1. Configurația actuală a site-ului RO-01-ICI

Din punct de vedere hardware clusterul include în prezent 9 servere (figura 1):

- un server Supermicro cu două procesoare Intel Quad-Core Xeon cu frecvența 2,67 GHz, 16 GB RAM;
- 4 dual Xeon cu frecvența de 2.4 GHz, 2 GB RAM;
- 4 servere Pentium 4 cu frecvența de 3 GHz, la 1 GB RAM.

Celelalte servere din componența site-ului RO-01-ICI sunt formate din computere Pentium 4 cu frecvența de 3 GHz și 2 GB RAM. Clusterul de calcul poate atinge un maxim de 9 Gflops, conform cu benchmark-ul HPL care este aplicat supercalculatoarelor din top 500. Puterea de procesare evaluată după criteriul SpecInt2000 are valoarea 13.000.

Capacitate de stocare este asigurată de un sistem disk array Promise Vtrak cu interfață SCSI.

Ultimul upgrade a fost implementat în august 2008, pe suportul proiectului PN II PEGAF și a inclus server-ul Supermicro, un switch (distribuitor) KVM pentru facilitarea administrării site-ului, un server HP ProLiant DL140 G3 5130, 2x1GB, pentru implementarea unei soluții IDS (Intrusion Detection System) pentru întregul domeniu din care face parte site-ul Grid și care este administrat independent de rețeaua ICI.

Sistemul de operare este Scientific Linux 4.x. Varianta de middleware este gLite 3.1, cu următoarele componente:

- testbed001.grid.ici.ro: *Computing Element*, interfața clusterului format din cele 9 servere (în total 20 cores) către infrastructura Grid.
- ca jobmanager este utilizat “pbs” ceea ce înseamnă că directoarele / home din Computing Element (CE) și Worker Node (WN) sunt comune, fiind exportate de pe CE pe WN-uri, acest lucru făcându-se pentru a utiliza MPI;
- testbed002.grid.ici.ro: *Storage Element* instalat cu variantă de middleware DPM_mysql;
- testbed003.grid.ici.ro: *User interface*. De pe acest sistem se utilizează comenzile de logare în mediul Grid, de trimitere și execuție a job-urilor, replicarea fișierelor, interogarea sistemului de informații, comenzi specifice mediului grid;
- testbed004.grid.ici.ro: *Monitoring Node*. Îndeplinește funcția de server de monitorizare pentru site în sistemul R-GMA, și de asemenea îndeplinește și funcția de nod colector de pentru clientul de accounting APEL;
- testbed005.grid.ici.ro: *Berkeley Database Information Index (BDII)* și *Resource Broker(RB)*.

Serviciul BDII are rolul de a menține informațiile cât mai exacte referitoare la resursele grid disponibile. Formatul în care se păstrează aceste informații este ldap.

Clusterul RO-01-ICI este administrat cu ajutorul aplicației Torque, o variantă îmbunătățită a mult mai popularului OpenPBS, în plus folosindu-se și scheduler-ul Maui.

În sfârșit, în baza responsabilităților ce revin ICI București în administrarea infrastructurilor specifice diverselor proiecte la care institutul participă în calitate de coordonator, pe site-ul RO-01-ICI sunt instalate și administrate servicii a căror disponibilitate este esențială pentru funcționarea acestor infrastructuri (de exemplu serviciul de Helpdesk în cazul infrastructurilor SEE-GRID și EGEE-SEE) sau pentru unele din organizațiile virtuale pe care le suportă acest site (Gridmosi.ici.ro și SEE-GRID).

Site-ul funcționează într-o incintă fără dotări speciale, cu excepția climatizării de tip birou, a unor UPS-uri de capacitate mică-medie și a unui sistem antiefracție.

5. Grant-ul POS-CCE 2.2.3 nr. 223

5.1 Suportul de finanțare

Finanțarea grantului de dezvoltare a site-ului RO-01-ICI este asigurată de Programul operațional sectorial „Creșterea competitivității economice” (POS-CCE), Axa prioritară 2 „Creșterea competitivității economice prin cercetare-dezvoltare și inovare” (POS CCE – CD), Domeniul major de intervenție 2.2 „Investiții în infrastructura de cercetare – dezvoltare - inovare”, Operațiunea 2.2.3: *Dezvoltarea unor rețele de centre C-D, coordonate la nivel național și racordate la rețele europene și internaționale de profil (GRID, GÉANT)* [10].

POS-CCE este unul dintre cele șapte programe operaționale sectoriale care reprezintă instrumente pentru realizarea priorităților trasate prin Cadrul Național Strategic de Referință (CNSR) și prin Planul Național de Dezvoltare (PND) 2007 – 2013, având ca obiectiv creșterea productivității întreprinderilor românești și reducerea decalajelor față de productivitatea medie la nivelul UE.

Obiectivul specific al axei prioritare 2 îl constituie creșterea capacității de C-D, stimularea cooperării dintre instituțiile de CDI și întreprinderi, creșterea accesului întreprinderilor la CDI.

Operațiunea 2.2.3 urmărește creșterea implicării cercetătorilor români în rețele internaționale de cercetare foarte specializate de tip Grid, cu implicații importante în dezvoltarea viitoare a științei și tehnologiei, precum și creșterea capacității rețelei de cercetare și educație RoEduNet, spre a se apropia de standardele rețelei GÉANT.

5.1 Obiective

La nivelul infrastructurilor internaționale menționate, după o perioadă de dezvoltare extensivă a acestora, în care s-a încurajat apariția a cât mai multor site-uri, pe măsura creșterii numărului de organizații virtuale și a gradului de utilizare a serviciilor acestor infrastructuri, au crescut corespunzător și cerințele față de dimensiunile și performanțele site-urilor componente. Proiectul GridMOSI a evidențiat și el limitările de capacitate ale site-ului RO-01-ICI pentru anumite clase de probleme. De asemenea, experiența colaborării cu elaboratorii de aplicații în cadrul acestui proiect a evidențiat interesul pentru facilități suplimentare celor de putere de calcul și capacitate de memorie, de tipul componentelor software necesare dezvoltării și execuției aplicațiilor științifice.

Proiectul are ca obiectiv general adaptarea site-ului RO-01-ICI la aceste cerințe. Obiectivele specifice se referă la:

- adaptarea condițiilor de funcționare (climatizare, alimentare cu energie electrică, securitate acces, PSI) la specificul regimului de producție în care funcționează acest site;
- upgradarea clusterului de calcul de înaltă performanță până la 1 Tflops și creșterea capacității de stocare la 20 TB;
- elaborarea unui model experimental pentru un sistem bazat pe cunoștințe, dedicat asistării / automatizării activităților de monitorizare și control pentru operarea resurselor infrastructurii Grid;
- pentru componenta infrastructurii Grid dedicată nivelului de aplicații și servicii pentru comunitățile științifice se are în vedere achiziționarea licențelor pentru versiunea cluster a produsului MATLAB, precum și instalarea unor aplicații open-source de interes larg.

Având în vedere că probabilitatea ca un site să satisfacă cerințele diverselor cereri de execuție lansate în infrastructura Grid este direct proporțională cu resursele oferite de site, rezultă că noua versiune a RO-01-ICI va avea un potențial mult mai mare de execuție a job-urilor din infrastructurile la care participă.

Un obiectiv de interes mai larg îl constituie valorificarea suportului pe care existența acestui centru de resurse performant îl poate oferi pentru demararea de cercetări pe tematici de actualitate în infrastructura sistemelor distribuite, complementare problematicii Grid, ca de exemplu „Cloud Computing” ([11], [12]), „Autonomic Computing” [13].

5.2 Clase de activități

Pentru implementarea acestor obiective, planul de realizare al proiectului include următoarele clase de activități:

- amenajare incintă site Grid:
 - achiziționarea serviciilor de proiectare tehnică a lucrărilor de amenajare și de instalare utilități (alimentare electrică, climatizare, securitate acces, PSI, comunicații);
 - achiziționarea lucrărilor de amenajare incintă, inclusiv alimentare electrică, securitate acces, PSI.
- achiziționarea și instalarea de active corporale:
 - echipamente de calcul, stocare date și comunicații;
 - soluție de back-up pentru alimentare electrică;
 - echipamente de climatizare.
- achiziționarea de active fixe necorporale - licențe de produse software dedicate diverselor comunități științifice, cu accent pe licența MATLAB pentru configurație cluster, împreună cu diverse toolkit-uri dedicate (calcul distribuit, controlul sistemelor);

- punerea în funcțiune a noii versiuni a site-ului RO-01-ICI și certificarea sa ca site de producție;
- realizarea și instalarea de software specific:
 - realizare prototip (analiza, proiectarea, implementarea și punerea în funcțiune) pentru sistemul de monitorizare și asistare inteligentă a administrării resurselor Grid, dedicat middleware-ului gLite și produselor software ce asigură funcționarea infrastructurii Grid;
 - instalarea middleware-ului gLite folosit în cadrul proiectelor EGEE și SEE-GRID, și a serviciilor esențiale pentru funcționarea unui VO cum ar fi VOMS - Virtual Organization Management Service, BDII - Berkely Database Information Index, WMS – Workload Management Service, LFC - LCG File Catalog, MyProxy;
 - instalarea de aplicații open-source.
- activități de diseminare: prezentarea proiectului la manifestări științifice interne și internaționale, publicații, site-ul proiectului;
- activități de informare și publicitate privind proiectul: anunțuri de presă, editare postere, pliante și CD-uri promoționale;
- managementul de proiect: planificarea detaliată a activităților și subactivităților, precizarea clară a sarcinilor pe fiecare membru al echipei de implementare, urmărirea realizării planului de implementare a proiectului, legătura permanentă cu monitorul proiectului, raportarea trimestrială către Organismul Intermediar a stadiului proiectului;
- achiziție servicii de audit final proiect.

5.3 Beneficiari potențiali

Beneficiarii direcți sunt colectivele de cercetare care participă la organizațiile virtuale susținute de site-ul RO-01-ICI în noua sa versiune, precum și instituțiile de cercetare cărora le aparțin. Aceste colective includ elaboratorii de aplicații Grid incluse în oferta organizațiilor virtuale respective, utilizatorii acestor aplicații sau utilizatorii interesați exclusiv în accesul la resursele Grid pentru testarea unor aplicații proprii.

O altă categorie de beneficiari direcți sunt administratorii de site-uri asistate pe probleme operaționale de administratorul principal al site-ului RO-01-ICI, care pot beneficia, dacă este cazul, de experiența în implementarea și operarea unui site de acest nivel de performanță.

Beneficiarii indirecti sunt alte colective de cercetare din instituțiile de mai sus sau din instituții partenere, dacă realizarea proiectelor de cercetare în care sunt angajați depinde de accesul la resursele și serviciile oferite de site-ul RO-01-ICI. Resursele sunt de natură hard (puterea de calcul și de capacitatea de stocare a datelor) sau soft (software licențiat sau open-source pus la dispoziția dezvoltatorilor de aplicații, aplicațiile instalate pe site care aparțin diverselor organizații virtuale susținute de site). Serviciile sunt de natură soft, instalate și administrate pe site, ca de exemplu: componenta de middleware User interface, suportul pentru MPI – Message Passing Interface, serviciul Helpdesk de asistare pe bază de tichete a utilizatorilor în rezolvarea unor incidente, serviciile specifice pentru funcționarea organizațiilor virtuale menționate anterior. La acestea se adaugă serviciile generale de asistare a potențialilor utilizatori pe suport web. De exemplu, portalul www.gridmosi.ro oferă informații privind procedura de obținere a certificatului digital și de înregistrare în organizația virtuală gridmosi.ici.ro, precum și acces la aplicațiile înregistrate în această organizație virtuală.

În raport cu administratorii de site-uri ca beneficiari direcți, beneficiarii indirecti sunt colectivele de cercetare care utilizează resursele site-urilor respective.

6. Concluzii

La nivel european dezvoltarea e-Infrastructurilor este considerată un obiectiv de importanță strategică al politicilor de cercetare și inovare pentru ca Europa să facă față provocărilor din următorii 10-15 ani. Conform [2], elementele cheie ale unei strategii reînnoite pentru a răspunde acestor provocări sunt:

- *e-știința*, Europa urmând să devină un centru de excelență pentru e-știință, bazat pe consolidarea capacităților de cercetare în domeniul calculului de înaltă performanță;
- *e-infrastructurile*, ca platforme de cercetare cu funcționare permanentă, sustenabile pe termen lung prin coordonarea eforturilor la nivel național și al UE;
- *inovarea*, bazată în mod esențial pe transferul expertizei în domenii din afara științei (e-sănătatea, e-guvernarea, e-învățarea) și pe utilizarea e-infrastructurilor ca platforme eficiente de experimentare tehnologică pe scară largă (de exemplu Internetul viitorului, programe masiv paralele, „Living labs”).

Dezvoltarea componentelor specifice e-Infrastructurii la nivel național – rețeaua de mare viteză, infrastructura națională Grid (RoReGI) – reprezintă obiective prioritare ale investițiilor în infrastructurile cercetării.

Programul operațional sectorial „Creșterea Competitivității Economice”, prin Axa prioritară 2 „Creșterea competitivității economice prin cercetare-dezvoltare și inovare” (POS CCE – CD), Domeniul major de intervenție 2.2 „Investiții în infrastructura de cercetare – dezvoltare - inovare”, oferă un important suport financiar pentru realizarea acestor componente de infrastructură. Având în vedere experiența acumulată în administrarea, operarea și utilizarea infrastructurilor Grid, colectivul de cercetare din ICI este interesat să participe la acest efort. Proiectul acceptat spre finanțare în cadrul operațiunii 2.2.3 “Dezvoltarea unor rețele de centre C-D, coordonate la nivel național și racordate la rețele europene și internaționale de profil (GRID, GÉANT)” va permite materializarea acestei intenții și valorificarea contribuției pe care ICI o poate avea la dezvoltarea infrastructurii RoReGI.

Prin parametrii de putere de calcul și capacitate de stocare propuși a fi atinși prin acest proiect, site-ul RO-01-ICI se plasează la acest moment între primele 5 site-uri din infrastructura SEE-GRID și între primele 30% din site-urile infrastructurii EGEE.

REFERINȚE

1. **CCE:** GREEN PAPER - The European Research Area: New Perspectives. Brussels, 4.4.2007, COM(2007) 161 final.
2. **CEC:** Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - ICT infrastructures for e-Science. Bruxelles, 5.3.2009 COM(2009) 108 final.
3. **e-IRG:** White Paper 2009.
4. **JONES, B.:** EGEE current status. EGEE 09 Conference, Barcelona, 21-25.09.2009.
5. **NEAGU, G.:** SEE-GRID-SCI project overview and its application offer. SEE-GRID-SCI national dissemination event, UVT, 26.09.2009.
6. **EGI_DS:** EGI Blueprint proposal. Deliverable D4.4, 12/09/2008.
7. **RoGrid:** „Strategia privind dezvoltarea GRID în România”. Proiect Orizont 2000, oct-dec. 2002.
8. **ANCS:** Planul strategic pentru Proiectul Pilot de implementare a infrastructurii naționale Grid. Task Force Grid, martie 2006.

9. **C.R.I.C.:** Raport privind infrastructurile de cercetare în România. Raport C.R.I.C. pe 2007, ANCS , ianuarie 2008.
10. **ANCS:** Ghidul solicitantului pentru dezvoltarea unor rețele de centre C-D, coordonate la nivel național și racordate la rețelele europene și internaționale de profil (Grid, GEANT). Anexa 2 la Decizia Președintelui ANCS nr. 9686 / 26.06.2008.
11. **BÉGIN, M.-E., B. JONES, J. CASEY, E. LAURE, F. GREY, C. LOOMIS, R. KUBLI:** An EGEE comparative study: Grids and Clouds - evolution or revolution ?, EGEE III project Report, 30.05.2008.
12. **FOSTER, I., Y. ZHAO, I. RAICU, S. LU.:** Cloud Computing and Grid Computing 360-Degree Compared. În: Grid Computing Environments Workshop, 2008. GCE'08, 2008, pp. 1–10.
13. **GMAC'09:** Workshop Grids Meets Autonomic Computing. Associated with the 6th International Conference on Autonomic Computing (ICAC'09), Barcelona, 15-19.06.2009.