

A Hungrid virtuális szervezet szolgáltatásai

A Hungrid, a magyar NGI bölcsője

Szeberényi Ágnes, Hernáth Szabocs KFKI RMKI
<szagi@iit.bme.hu>, <hernath@mail.kfki.hu>

Kivonat

. A Hungrid Virtuális Organizáció az EGEE egyetlen, hivatalos magyar virtuális organizációja. A szervezet létrehozásakor az volt a fő cél, hogy a magyar tudományos kutatói közösség számára egy gyűjtő szervezetet hozzunk létre, mely lehetővé teszi, hogy az EGEE projekt keretében kifejlesztett és üzemeltett hazai Grid infrastruktúrát a hazai kutatói közösség az EGEE projekt virtuális szervezeteitől eltérő kutatási célok elérése érdekében is használhassa. Ez első 24 órás Grid szolgáltatás Magyarországon. [1] A rendszerben a szekvenciális és paralel programok futtatása mellett a fájlok/programok tárolására, regisztrációjára is lehetőség nyílik. A létrehozott Hungrid egy olyan virtuális organizáció, mely kiemelt szerepet játszhat a közeljövőben történő EGEE- EGI/NGI modellváltásban. A Hungrid célja további magyar VO-k megalapításának támogatása.

1. Bevezetés (Mi a Grid?)

Manapság a kutatók, tudósok egyre bonyolultabb problémákkal szembesülnek, melyeknek megoldása jóval többet igényel egy számológéppel használatánál. Kutatás számítógép nélkül szinte lehetetlen. Gyakran előfordul, hogy egy vagy két CPU nem elég egy számítás elvégzéséhez. Egyre gyakoribb, hogy egy számítógépes klaszter, de még a speciális célú szuperszámítógépek sem elegendők a számításokhoz, amelyeket a tudósok el akarnak végezni. Ennek áthidalására jöttek létre a Gridok.

A Grid a szuperszámítástechnika egyik fejlesztési iránya, és egyben alapfogalma. Határos a klaszterekkel és az önálló szuperszámítógépekkel. A markáns különbség a fizikai kiterjedésben lehet: egy szuperszámítógép néhány szekrény méretével ér fel, egy klaszter egy gépteremben elfér, viszont egy kisebb Grid kialakításához is legalább országos méretű összefogás kell. Az esetek többségében a Grid lokális csomópontjait képező állomások (site) a nyilvános interneten keresztül kommunikálnak egymással, bár egyre több a dedikált illetve akadémiai hálózatra csatlakozó csomópont a mai Grid-implementációkban.

A Grid egy olyan szolgáltatás, melynek lényege a számítógép erőforrások és az adattárolási kapacitás megosztása az Interneten. A Grid több mint számítógépek közötti egyszerű kommunikáció, végső célja az, hogy a számítógépek globális hálózatát egyetlen óriási számítógépes erőforrássá alakítsa.

Funkcionálisan többféle Gridet különböztetünk meg: [2]

- Számítási Grid (beleértve a CPU-idő gyűjtő Grideket is), amely a nagy számítási-idő igényű feladatok végrehajtására koncentrál. Heterogén, elosztott számítási erőforrásokból álló infrastruktúra.
- Adat Gridek, vagy más néven olyan rendszerek, amelyek nagy mennyiségű adat felügyelt elosztására és menedzselésére képesek. Heterogén, elosztott adatkezelése erőforrásokból álló infrastruktúra.
- Speciális erőforrásokat tartalmazó Gridek, amelyhez egy saját erőforrás tartozik (például teleszkóp), és ahol a Gridet arra használják, hogy távolról vezéreljék azt illetve feldolgozzák az általa szolgáltatott adatokat.

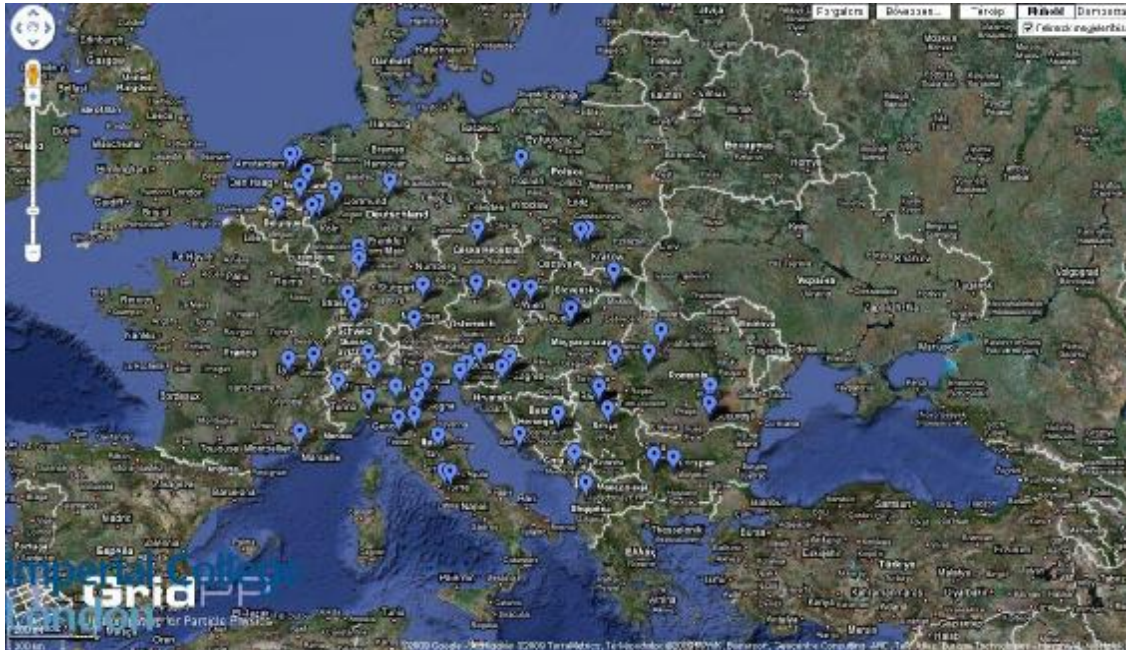
A Gridek egyik legfontosabb kérdése a biztonság: a legtöbb biztonsági kérdést az veti fel, hogy a hozzáférést hálózati szinten nem lehet - és nem szabad - korlátozni, hiszen a szolgáltatás egyik lényeges tulajdonsága, hogy az interneten bárholnan elérhető legyen.

2. EGEE projekt és a Virtuális Szervezetek

A 2004-ben induló EGEE (Enabling Grids for E-science) projekt az EU által támogatott legnagyobb grides projekt. A három fázisú projektnek jelenleg a harmadik, utolsó fázisában vagyunk. A projekt eredeti céljai között volt egy olyan grid infrastruktúra létrehozása, mely 24/7 rendelkezésre állású.

A projekt előző fázisai során kialakításra került ez a 24/7 rendelkezésre állású Grid infrastruktúra, mely meghatározó infrastruktúrává vált számos európai kutatói közösség számára. Az EGEE gridje mára a legelterjedtebb és legszélesebb körben használt multidiszciplináris Grid infrastruktúrává nőtte ki magát a világon. Jelenleg közel 50 ország 140 intézete vesz részt az együttműködésben, ezzel a mintegy 10 ezer felhasználójának 80 ezres CPU számítási kapacitást, illetve 20 petabyte ($5 \cdot 10^{15}$ byte) tárolási kapacitást biztosítva.

Az EGEE utolsó fázisának célja, hogy továbbra is egy világszínvonalú, koherens és megbízható európai grid infrastruktúrát biztosítson, illetve további felhasználói köröket vonjon be. Ezen utolsó fázis további célkitűzése, hogy alapot jelentsen a grid további, hosszútávú, hasonló színvonalú fenntartásához, az EGEE projekt befejeződése után is az EGI-n keresztül (Európai Grid Iniciatíva).



1. ábra: Az EGEE gridje Európában

Az E-scienceE közmű nyílt és elosztott kollaborációt biztosít felhasználói számára. A projekt során kialakított infrastruktúra fő elemei közé tartoznak:

- a VO modell,
- autentikáció és autorizáció
- gLite köztesréteg

2.1 A VO-modell

A Grid rendszert különböző felhasználói csoportok használhatják, ezeket VO-nak (Virtual Organization, virtuális szervezetek) nevezzük. A Virtuális Szervezetek azonos kutatási tevékenységben érdekelt felhasználókat fognak össze. Ezek a szervezetek egyenként köthetnek megállapodást az állomásokkal, a felhasználni kívánt erőforrások tekintetében. Ilyen értelemben az állomások a szolgáltatók, a VO-k megrendelik a szolgáltatást, és a VO felhasználói dolgoznak a csoportjuk számára lefoglalt erőforrásokon, az előre meghatározott jogok, és jogosultságok szerint.

A VO-k számos feladatot látnak el, felhasználó menedzsment és támogatás, autorizáció, erőforrás koordináció, szoftvereszközök fejlesztése és telepítése. Az EGEE projekt keretén belül több mint 200 Virtuális Szervezetet tartanak nyilván, körülbelül 10 000 felhasználóval.

2.2 Autentikáció és autorizáció

Az egee-ben a felhasználók azonosítása digitális tanúsítványokon keresztül történik, ez elengedhetetlen fontos háttér infrastruktúrája a gridnek. Az EGEE projekt megteremtette a globális e-science közmű CA szolgáltatás háttérét.

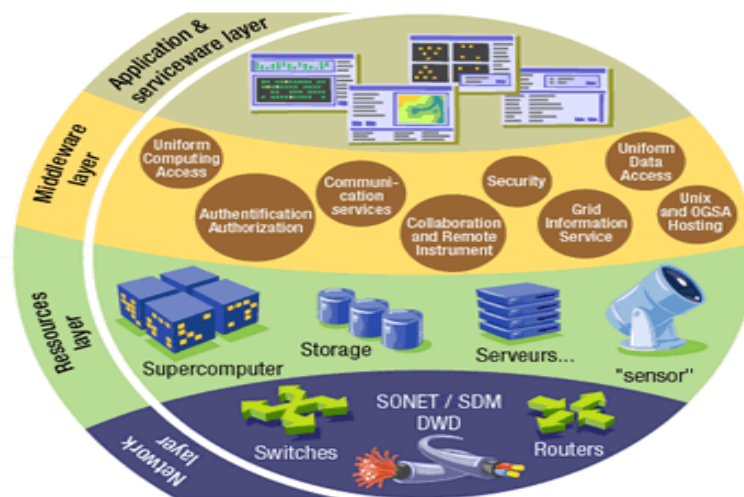
2.3 gLite köztesréteg

Az EGEE projekt keretében kifejlesztett infrastruktúra egyik legfőbb eleme közé tartozik a gLite köztesréteg és az azáltal biztosított alapvető és magasabb szintű szolgáltatások.

A Grid architektúra rétegekkel modellezhető. Míg a magasabb rétegek felhasználó centrikusak, addig az alacsonyabb rétegek hardver centrikusak. A legalsó réteg, a hálózati réteg biztosítja a kapcsolatot a Grid erőforrásaival. E felett helyezkedik el az erőforrás réteg, ami a Grid részét képező tényleges erőforrásokból áll. Ezek lehetnek számítógépek, tároló rendszerek, elektronikus adatkatalógusok és még érzékelők is (pl. teleszkópok), melyeket közvetlenül be lehet kötni a hálózatba. [4]

A köztesréteg nyújtja azokat az eszközöket, amelyek lehetővé teszik különböző elemeknek (szerverek, tárolók, hálózatok stb.), hogy az egységesített Grid környezet részét képezhessék. A köztesréteg egy felületet biztosít az erőforrások és alkalmazások között. A Grid köztesrétege olyan alapvető és magasabb szintű szolgáltatásokat biztosít, mint az erőforrásokhoz való hatékony és biztonságos hozzáférés, optimális erőforrás kihasználtság, hitelesítés különböző site-ok felé, monitorozás, stb.

A struktúra legmagasabb rétege az alkalmazás réteg, ami magában foglalja az összes felhasználói alkalmazást (tudományos, mérnöki, üzleti, pénzügyi), valamint a portálokat és fejlesztési eszköztárakat is, melyek az alkalmazásokat támogatják. Ez a felhasználók által is látott szint.



2. ábra: A rétegek elhelyezkedése

Ezek után tisztán látható, hogy az EGEE projekt keretében a világ legnagyobb méretű multitudományos grid infrastruktúráját sikerült létrehozni.

4. A Hungrid magyar virtuális szervezet (VO) [5]

A Hungrid az RMKI, a SZTAKI és az ELTE által – az EGEE projekt keretén belül – 2005 júliusában létrehozott egyetlen, magyar, általános célú magyar virtuális szervezet.

A Hungrid lehetővé teszi, hogy az EGEE projekt keretében kifejlesztett és üzemeltetett hazai Grid infrastruktúrát a hazai kutatói közösség az EGEE projekt virtuális szervezeteitől eltérő kutatási célok elérése érdekében is használhassa.

A VO megalapításának célja egy olyan magyarországi tudományos Grid megteremtése, kialakítása, amelyet a tudományos élet legkülönbözőbb szakterületein munkálkodó kutatók, oktatók és diákok minél szélesebb rétegei használhatnak.

Megalapításakor a VO korlátozott erőforrásokkal és szolgáltatásokkal korlátozott számú felhasználó számára volt elérhető. A kiindulási cél egy olyan magyar gyűjtőszervezet létrehozása, amelynek feladata a Grid használatának elősegítése volt a nagyenergiás fizika területén túl, mindezt 24/7 rendelkezésre állással.

A Hungrid VO jelenleg több hazai intézmény (KFKI-RMKI, BME, ELTE, NIIFI, SZTAKI) Grid erőforrásait fogja össze és teszi elérhetővé olyan hazai kutatói közösségek számára, melyek nagy számítási és/vagy tárolási kapacitást igényelnek, vagy csak éppen ki akarják próbálni a Grid rendszerek nyújtotta lehetőségeket. [6]

	BME	ELTE	NIIF	KFKI RMKI	SZTAKI
CPU	25	25	10	50	
Storage	20 GB	700 GB	30 GB	5 TB	
Egyéb	UI	UI	CA	UI, WMS, LCG, Proxy, CA/RA,	Portál
				Hungrid honlap	támogatás

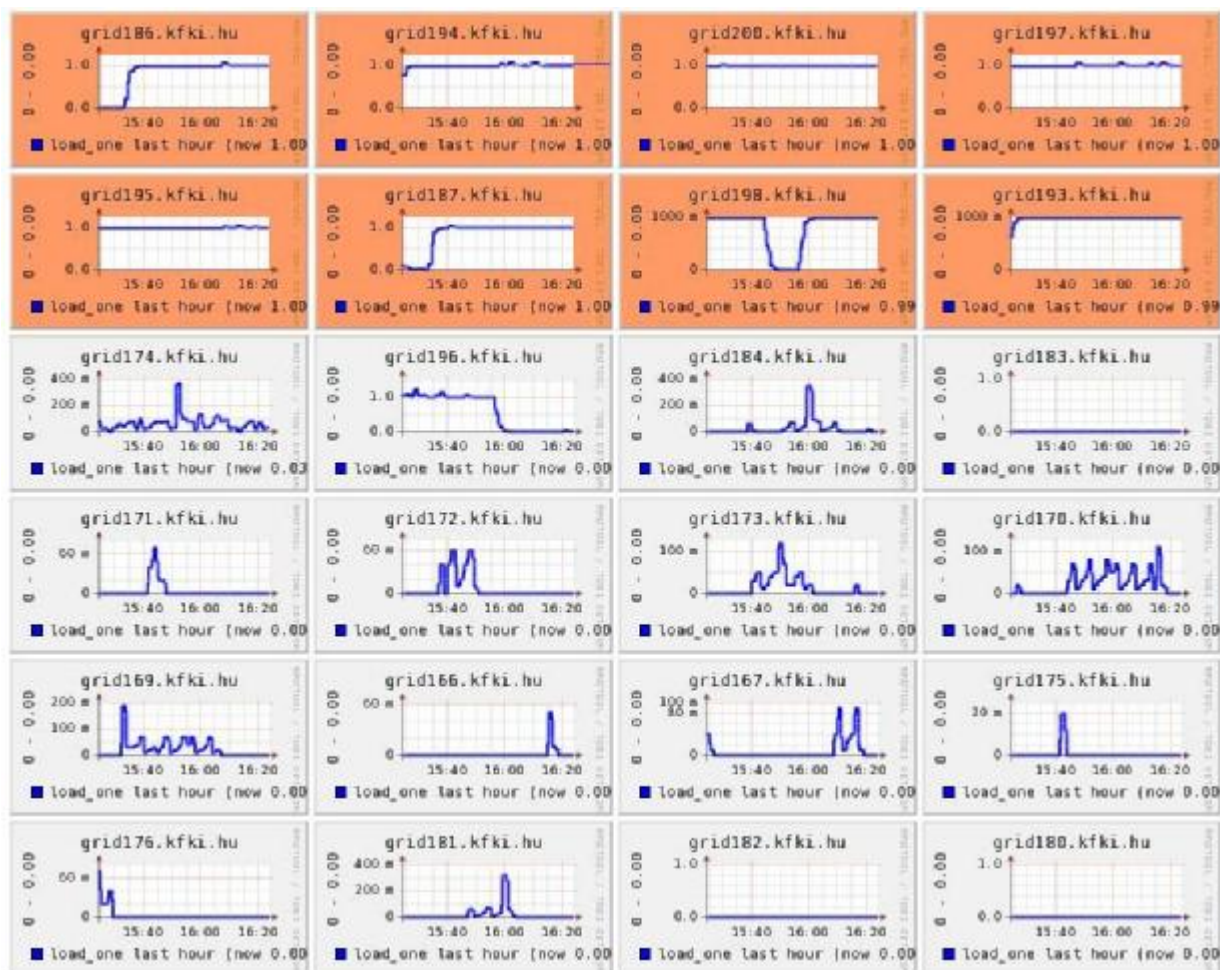
3. ábra: A Hungrid VO-t támogató intézmények erőforrásai

A Hungrid az első 24 órás Grid szolgáltatás Magyarországon. Szekvenciális és paralel programok futtatása mellett a fájlok/programok tárolása, regisztrációja, webes felhasználói felület és részletes információs rendszer is megtalálható a szolgáltatások között.

A Hungrid VO-t támogató intézmények a gLite middleware segítségével nem csak Hungrid VO-t szolgálják ki, hanem az EGEE, és a NIIFI estében a NorduGrid VO-it is. [7] A Hungrid VO-ban igénybe vehető alapszolgáltatások a következők [8]:

- Felhasználói felület (User Interface): A felhasználó Gridhez való kapcsolódási pontja.
- Fájl katalógus (LCG File Catalog): A Griden megtalálható fájlok és replikák tárolásához.
- Számítási egységek (Computing Element): a számítási kapacitást nyújtó erőforrás, egy homogén klaszter, amely egy sorban várakoztatja a beérkezett végrehajtandó programokat.
- Tárolási egységek (Storage Element): uniform hozzáférést tesz lehetővé az adattárolók számára.
- Információs rendszer: Összegyűjti az SE és CE-ek állapotát (BDII)
- Naplózás és könyvelés (Logging and Bookkeeping services): A feladatok könyvelési információinak tárolása
- A VOMS szerver a Hungrid virtuális szervezet tagjainak nyilvántartására.
- Munkavégző egységek (Worker Node): itt futnak a job-ok. Hungrid számára.

Az infrastruktúránk a közelmúltban bővült egy WMS szerverrel, mellyel kiváltottuk a korábbi erőforrás brókert (RB). Az elmúlt közel 4 év során jelentős erőforrás bővítésen ment keresztül a VO. Jelenleg a Hungrid-es WN-ek 32 bites MD Athlon(tm) XP 2500+, 2 GHz-es gépek.



4. ábra: Néhány Hungrides WN terhelése

A Hungridben rejlő lehetőségek ezzel nem merülnek ki. Miután a felhasználók felmérték igényeiket, lehetőségeiket, cél, hogy ők maguk is egy közösségbe tömörüljenek, és további magyar VO-kat hozzanak létre, melyek később saját csatlakozási pontot teremtenek az EGEE infrastruktúrához. Ezen VO-k megalapításához szolgál ugródeszkaként a Hungrid.

Jelenleg a Hungrid felhasználói köre nem túl kiterjedt. A 82 felhasználóból csak 10 küldi aktívan a job-okat.



5. ábra: A Hungrid, a magyar VO

5. Hungrid szerepe a modellváltásban

Az EGEE utolsó fázisában vagyunk jelenleg. Ez 2010 közepén fog befejeződni. Ez a dátum fordulópontra lesz a grid infrastruktúra támogatásában. Megszűnik az Unió támogatása, a cél viszont ugyanaz, fenntartani a már létrehozott sikeres grid infrastruktúrát. Ezt a tervek szerint az EGI intézménye fogja koordinálni a nemzeti grid kezdeményezések szövetségén keresztül. Magyarországon a modellváltást, illetve a magyar NGI koordinálását a 2003-ban alakult MGKK (Magyar Grid Kompetencia Központ) hajthatná végre. Ez a központ azzal a céllal alakult, hogy a grid kutatási tevékenységeket, pályázatokat koordinálja és publikálja nemzeti szinten.

Kézenfekvő, hogy a tervezett magyar NGI egyik induló platformja a Hungrid legyen. Ezt több dolog indokolja; a Hungrid maga is EGEE kompatibilis, azaz rendelkezik az alapvető és magasabb szintű EGEE szolgáltatásokkal, és képes az EGEE nélküli működésre. Illetve a másik fő indok, hogy a legátfogóbb szervezet, hiszen a hazai e-science infrastruktúrában résztvevő intézményeket fogja össze.

A modellváltás után az NGI intézményének egyik feladata lesz a felhasználói kör és erőforrások bővítése.

A Hungrid kiemelkedő szerepet játszhat ezen feladatkörök ellátásában.

A felhasználói kör bővítése érdekében az egyik legkritikusabb pont a PR. Ez a legfontosabb lépése az új felhasználók bevonásának. Természetesen ehhez menedzselési feladatok is járnak. A már meglévő felhasználókat aktivizálni kell, rendszeres tagság konfirmációt kérni tőlük e-mail formájában, ezáltal az inaktív felhasználók szűrésére. Kezeleni kell a felhasználói visszajelzéseket, illetve nyilván kell tartani a kutatási eredményeket (publikációk). Az oktatásban is kulcsszerepet tölthet be a Hungrid, hiszen létrehozásakor a magyar oktatási élet támogatását is célul tűzte ki. Az accountingnak is fontos szerepe van, arra szolgál, hogy felmérjük, milyen minőségű szolgáltatást tud nyújtani az infrastruktúra. Mindezen feladatok

hatékony ellátására létrehoztunk egy folyamatosan frissülő web lapot és egy a felhasználók számára fenntartott levelezőlistát. Továbbá tervezzük felhasználói fórumok megrendezését, amelyek célja a Hungrid még szélesebb körben való elterjesztése, felhasználói körök bevonása, tapasztalatok cseréje, információ áramoltatás.

Másik feladatkör, amiben a Hungrid kulcsfontosságú lehet, az erőforrások bővítése. Az egyetlen magyar VO EGEE kompatibilis gerincszolgáltatásokkal rendelkezik, és már a mai formájában is képes az önálló működésre. Jövőbeni feladatok közé tartozik az erőforrások központi monitorozása, új erőforrások tesztelése, job accounting. Az erőforrások bővítésének leghatékonyabb formája új magyar intézmények bevonása. Az intézményi csatlakozásnak több szintje lehetséges.

1. Hungrid felelős:

Intézményenként olyan személy, aki maga is aktív Hungrid VO tag, segít az új felhasználóknak a kezdő lépésekben, illetve a már meglévő felhasználók egyszerűbb gyakorlati kérdéssel hozzá fordulhatnak. (pl tanúsítvány meghosszabbítás)

2. User Interface gép installálása és karbantartása:

Az olyan intézmények, amik számítógépes erőforrással is szeretnének hozzájárulni a Hungrid sikeres működéséhez, már egy, a helyi felhasználók számára telepített UI-val is megtehetik ezt.

3. Új számítási és tárolási erőforrások bekapcsolása a Hungridba:

Azok az intézmények/intézetek, amik erőforrást vagy tárolási erőforrást kívánnak felajánlani, azoknak a gLite middleware-t kell telepíteniük, és megfelelően konfigurálni CE, ill. SE erőforrásként. Az így telepített gépek erőforrásai elérhetővé válnak a Hungrid közösség számára.

4. Új erőforrások bekapcsolása az EGEE infrastuktúrájába:

A megfelelően konfigurált tesztüzemű site ellenőrzések után bekapcsolható az EGEE 7*24 órában működő produktív rendszerébe, amely azonban megfelelő szintű szolgáltatás nyújtás vállalást igényel.

A felhasználókat érintő feladatkörök ellátásához hasonlóan létrehoztunk a Hungridban részt vevő intézmények számára egy adminisztrátori listát, aminek a célja a hatékony együttműködés és a felhasználók igényeinek minél megbízhatóbb kiszolgálása.

Összefoglalás

A Hungrid az EGEE egyetlen hivatalos, magyar és általános célú virtuális szervezete. Jelen pillanatban is fejlődik, a felhasználás menete és szabályai pontosítódnak, a rendelkezésére álló erőforrás gyarapszik. Jelenlegi formájában bárki, aki a magyar kutató vagy oktatási szférában dolgozik, és egy Grid rendszerre indokoltan szüksége van, igénybe veheti.

A Hungrid virtuális szervezetben az intézmények számítási és adattárolási kapacitásaikat egy közös Grid rendszerbe tudják összefogni és magyar tudományos számítási projektek végrehajtására felhasználni.

Célunk, hogy ösztönözzük a magyar grides fejlesztéseket akkor is, ha nincs meg az önálló VO kialakításához/üzemeltetéséhez a kritikus tömeg.

Jövőre vonatkozó terveink között szerepel az EGEE-NGI modellváltásban való kulcsszerep, illetve egyfajta kapuként szolgálni új magyar VO-k és az EGEE VO-i között. A grid rendszerek oktatásban való megjelenése lehetőséget teremt a Hungrid VO oktatási szférában való felhasználásának is. [10].

Köszönetnyilvánítás:

Jelen cikkben közvetlenül felhasználtuk az EU EGEE (INFSO-RI-222667) projekt eredményeit, valamint köszönetet mondunk az OTKA NK-67974 és H07C-74153 támogatásokért.

Irodalomjegyzék:

[1] Podhorszki Norbert, Kacsuk Péter: A Hungrid Virtuális Szervezet az intézményi számítási erőforrások országos összefogására:

www.lpds.sztaki.hu/~pnorbert/pub/if05-hungrid.pdf

[2] Grid computing: http://hu.wikipedia.org/wiki/Grid_computing

[3] EGEE projekt, <http://www.eu-egee.org>

[4] GridCafe: gridcafe.eu-egee.hu

[5] Hungrid hivatalos honlapja: <http://grid.kfki.hu/hungrid>

[6] Horváth Dezső: Grid: a számítógép-hálózatok új rendszere: Magyar Tudomány, 2008/08 928. o.

[7] Hernáth Szabolcs: Hungrid: Magyarország kapuja az EGEE infrastruktúrájához: www.kfki.hu/~horvath/RIPNP-GRID/Talks/2008/Hernath_Hungrid-GRIDCAFE.pdf

[8] gLite 3 User Guide. <http://glite.web.cern.ch/>

[9] MTA KFKI RMKI honlapja: <http://www.lcg.kfki.hu>

[10] Kápolnai Richárd, Németh Dénes, Dr. Szeberényi Imre BME IIT – BME IK: Grid az oktatásban <https://nws.niif.hu/ncd2007/docs/ehu/035.pdf>