

A HunGrid bemutatása és alkalmazása levegőszennyezés előrejelzésére¹

Patvarczki József*, Debreczeni Gergely, Lovas Róbert*, Lagzi István***, Kacsuk Péter*, Turányi Tamás*****

*MTA SZTAKI

{patvarcz, kacsuk}@sztaki.hu

**MTA KFKI RMKI

dgergo@rmki.kfki.hu

***ELTE

lagzi@vuk.chem.elte.hu, turanyi@garfield.chem.elte.hu

1. Bevezetés

Ez a cikk az EGEE projekt keretén belül létrehozott Magyar Virtuális Szervezetet, a HunGrid-et mutatja be. A cikk röviden ismerteti a jelenlegi EGEE Grid infrastruktúrát, ami több mint 9000 processzorával jelenleg a világ legnagyobb Grid rendszere. Bemutatásra kerülnek az EGEE Grid middleware (LCG-2) legfontosabb komponensei. A cikk külön kitér a virtuális organizációk (VO) szerepére és szervezési módjukra, majd részletesen ismerteti a magyar VO kialakításának módját, a HunGrid felépítését, ill. a résztvevő intézmények infrastruktúráját.

A HunGrid nem egyszerűen az LCG-2 egyik virtuális organizációja, hanem több is annál, mivel olyan új elemeket tartalmaz, amik az eredeti LCG-2 rendszerben nem szerepelnek. Ilyen bővítő elemek a P-GRADE Grid portál és a Mercury Grid monitor. A P-GRADE Grid Portál a HunGrid WEB alapú belépési pontját teremti meg, melyen keresztül kényelmesen, az amúgy hosszú és bonyolult szöveges parancsok megtanulása és alkalmazása nélkül is használhatóvá válik a HunGrid. A portál lehetővé teszi komplex workflow alkalmazások kényelmes definiálását és futtatását a HunGriden. A Mercury Grid monitor segítségével a párhuzamos Grid alkalmazások processz-szintű monitorozása valósulhat meg.

A cikk ismerteti a potenciális felhasználók számára a HunGrid alkalmazásba vételének lépéseit, beleértve a szükséges Grid tanúsítványok megszerzésének módját is. A HunGrid dinamikusan bővíthető bármely egyetemi, vagy akadémiai erőforrással. A cikk leírja a csatlakozni kívánó intézmények számára a csatlakozás módját is. Az aktuális tervnek megfelelően, egy általános kép bemutatásán keresztül szeretne információt nyújtani a jövőbeli HunGrid felépítéséről, megteremtve a teljes magyarországi LCG-2 alapú Gridet, amely nyitott mindazok számára, akik tudományos kutatómunkát vagy éppen oktatási tevékenységet folytatnak.

A GRID rendszerek ebben az évtizedben egyre inkább népszerűvé válnak a természettudomány területén, melyekben nagyszámú heterogén erőforrást köthetnek össze, hogy komplex problémákat oldjanak meg velük. Az egyik hazai projekt – a „Kémiai Grid és alkalmazása légszennyezettség előrejelzésére” – egyik legfontosabb célja, hogy megvizsgálja a Gridet, mint nagy számítási kapacitást biztosító infrastruktúrát és gyakorlati megoldásokat találjon a kémia területén.

Az MTA SZTAKI egy olyan fejlesztőeszköz családot dolgozott ki, a P-GRADE-et, a P-GRADE portált és a MERCURY-t, melyek magas szintű grafikus megközelítésük segítségével hatékony támogatást nyújtanak a szekvenciális és a már elkészített alkalmazások újraterveléséhez, továbbá az alkalmazások teljesítményvizsgálatához és végrehajtásához akár párhuzamos, akár GRID platformokon. Újdonság, hogy komplex programokat hajthatunk végre különböző GRID-eken, mivel ezek az eszközök támogatják az ún.

¹ A cikkben leírt eredmények az EU által támogatott EGEE es SEEGRID projekt, valamint az OMF-00580/2003 (IKTA5-137), az IHM 4671/1/2003, az OTKA T 042459 és az OTKA D048673 szerződésszámú projektek keretében készültek.

workflow-k létrehozását. A bemutatott eszközök a magyar egyetemek és akadémiai intézetek számára elérhetőek, így könnyen párhuzamosíthatják a nagy számítási igényekkel rendelkező szekvenciális szimulációikat, majd GRID rendszereken futtathatják azokat, mint pl. az új HUNGRID-en.

Az MTA SZTAKI és az ELTE Fizika-Kémiai Tanszékének közös eredményeként, a P-GRADE-et, a P-GRADE portált és a MERCURY monitorozó rendszert már sikeresen alkalmazták a kémiai Grid projektben, hogy Griden futtathatóvá tegyenek egy már korábban reakció-diffúzió-advekciónál rendszerekhez készített kémiai szimulátort. A kifejlesztett alkalmazás a P-GRADE portálon keresztül érhető el, ahol minden egyes összetevő (párhuzamos vagy szekvenciális job) képes együttműködni a Gridben a P-GRADE workflow koncepciójára alapozva, hogy hatékony légszennyezettség előrejelzést biztosíthasson, pl. radioaktív nuklidok terjedése esetén.

Továbbá a cikkben röviden bemutatjuk a reakció-diffúzió-advekciónál rendszerek alapjait, valamint részletesen ismertetjük ezek szimulációját a P-GRADE programozási környezet és a P-GRADE portál segítségével; kezdve a tervezéstől a teljesítményanalízisen át a végrehajtási fázisig. Továbbá beszámolunk a HUNGRID-en történt kísérleti futtatások eredményeiről is demonstrálva a P-GRADE portál hatékony felhasználási lehetőségeit.

2. EGEE és LCG-2

Az EGEE [1] (Enabling Grid for E-sciencE) az Európai Unió 6. keretprogramja keretében támogatott legnagyobb Grid projekt. A jelenleg két évre, több mint 30 millió Euróval támogatott program célja a Grid-technológia legújabb eredményeinek felhasználásával létrehozni egy olyan Grid szolgáltatást, amely a napi 24 órájában rendelkezésre áll. Először az európai kutatás meghatározott területei számára, majd szélesebb körben, és általában a kutatás-fejlesztés, később az ipari-szolgáltatási alkalmazások számára is.

Az EGEE első időszakban két kiválasztott mintaprojektet támogat: a nagyenergiájú fizikai kutatások (LHC kísérletek), valamint a biomedikai problémák megoldása kapcsán a biofizikát és az egészségügyet.

Az EGEE-ben 27 ország 70 kutatóintézete vesz részt, ezen intézmények 12 federációra oszlanak. A közép-európai federációban többek között a Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutatóintézetét (MTA SZTAKI) és a KFKI RMKI vesz részt Magyarországról.

A világ legnagyobb részecskegyorsítója van épülőben Genf közelében, a francia svájci határon lévő CERN-ben. Az LHC-nak (Large Hadron Collidernek) nevezett gyorsító 4 nemzetközi együttműködés keretében 2007-re készül el, majd elkészülte után mintegy 6000 fizikus és mérnök fogja a világ számos egyeteméről és kutatóintézetéből a tényleges kísérleteket végezni.

Az LHC kísérletek különlegesen nagy számítástechnikai igényeket támasztanak, körülbelül 12-14 PetaByte-nyi mérési adat kerül majd begyűjtésre évenként, mely adattömeg feldolgozásához a mai leggyorsabb PC processzorokat figyelembe véve majd kb. 200 ezerre lenne szükség. A CERN tervei alapján az adatfeldolgozáshoz egy olyan komplex rendszerre van szükség, melynek közel kétharmada az egész világon ún. Regionális centrumokban többek között Európában, Amerikában és Ázsiában lesz elosztva. Az 2.1. ábra az LCG Grid már jelenleg is működő ilyen erőforrásait mutatja.

Az LHC-hoz szükséges számítástechnikai rendszer, mint globális Grid kerül kialakításra azzal a céllal, hogy a földrajzilag szétszórt erőforrásokat egyetlen virtuális egészévé integrálja.

A feladat megoldása igen sok területen komoly kihívást jelent. Ilyen területek például a tudományos feladatok elosztott hálózaton való megoldása, az ehhez szükséges Grid middleware kidolgozása, objektum orientált adatbázisok kezelése, valamint egy ekkora Grid mindennapos menedzsmentje.

A kutatás és fejlesztés tudományos intézetek és ipari partnerek bevonásával a CERN által koordinált projekt keretében történik. Az LHC Computing Grid rövidítéséből LCG-nek keresztelt projekt szervesen integrálódik a nagyobb európai nemzeti és regionális Grid kezdeményezésekkel, és szorosan együtt fog működni olyan projektekkel, amelyek a Grid technológia nagyterjedésű hálózati alkalmazásával foglalkoznak. Ilyen európai projektek az Európai Unió által szponzorált GEANT, Datagrid és DataTAG, Amerikában pedig a National Science Foundation és Department of Energy által támogatott GriPhyN, Globus, iVDGL és PPDG.

A CERN 2003 első felében létrehozta a bevezető jellegű első Gridjét, az LCG-1 -et, azzal a céllal, hogy megbízható szolgáltatást nyújtson az LHC kísérletekben dolgozók részére. Az LCG erőforrások használatához legalább egy Virtuális Szervezetnek (Virtual Organization) a tagjának kell lenni. Az LCG több VO-t is definiál. A VO egy olyan szerveződés, amely egy közös tevékenység köré egyesít felhasználókat, intézményeket és erőforrásokat.



2.1. ábra Az LCG Grid jelenlegi erőforrásai (forrás: <http://lcg.web.cern.ch/LCG>)

Egy VO-hoz tartozó intézmények és egyének között a VO erőforrások, fájlok, alkalmazások biztonságosan és koordináltan megoszthatók. Az LCG-1-ben elérhető VO-k a következők:

- ALICE: A Large Ion Collider Experiment; <http://alice.web.cern.ch/Alice/AliceNew>,
- ATLAS: A Toroidal LHC ApparatuS; <http://atlas.web.cern.ch/Atlas>,
- CMS: The Compact Muon Solenoid; <http://cmsinfo.cern.ch/Welcome.html>,
- LHCb: The Large Hadron Collider beauty experiment; <http://lhcb.web.cern.ch/lhcb>.

Az ALICE keretében együttműködő felek célja, hogy egy nehéz-ion detektor segítségével tanulmányozzák a részecskék reakcióit az LHC által biztosított energiaszinteken. Az ATLAS tulajdonképpen egy részecske fizikai kísérlet, amely a természetes alapanyagokat kutatja különböző anyagokban, a világegyetemet kialakító és formáló erők tanulmányozása mellett. A CMS keretében magas energiaszinteken új fizikai koncepciókat, elméleteket szeretnének bebizonyítani, melyeket később a telekommunikáció és a számítástechnika félvezetőkön alapuló világában tudnak majd hasznosítani. Az LHCB olyan nagy precizitású mérések elvégzését tűzte ki célul, amelyek a radioaktív sugárzásokat és bomlási folyamatokat vizsgálják.

Érzelhető, hogy a különböző VO-k által felölelt alkalmazás területek skálája igen széles és komplex. A feldolgozott adatok szolgáltatásával szemben támasztott követelmények igen nagyok mind a számítási erőforrások, tároló kapacitások, mind az adat elérését és a rendszert üzemeltető szakemberektől tekintetében. Az igényeknek megfelelő speciális Grid Middleware LCG szoftver elrejti a felhasználó elől a rendszer valódi komplexitását, és egy egységes eszként tünteti fel a földrajzilag szétszórt erőforrásokat. Az LCG-1 szolgáltatás az első ún. adat intenzív produkciós Grid. Az LCG új verziója 2004-ben vált elérhetővé LCG-2 néven. Az LCG-2 teljes mértékben az LCG-1-re épül, kiegészítve azt még több elérhető erőforrással és funkcionalitással, eleget téve az LHC 2004. év adatfeldolgozási követelményeinek.

2.1. Az LCG-2 felépítése

Az LCG-2 szintén VO-k szerveződéséből áll, ezért a használatához legalább egy VO-hoz kell tartozni. Ezen kívül az erőforrások használatához a felhasználónak rendelkeznie kell egy digitális X.509 alapú Grid igazolvánnyal, amelyet az LCG-ben jegyzett CA hatóság állíthat ki.

Az LCG biztonsági rétege (Grid Security Infrastructure, azaz GSI) biztonságos azonosítást és kommunikációt tesz lehetővé az egyes erőforrások között a nyílt kulcsú titkosítás, az X.509 certifikátum-alapú azonosítás [2] és a Secure Sockets Layer (SSL) kommunikációs protokoll segítségével.

Az LCG támogat még más, nem LHC által létrehozott VO-kat is, úgy mint a Baar, D0, H1, Zeus, Biomed, ESR (lásd https://lcg-registrar.cern.ch/virtual_organization.html).

2.1.1. LCG-2 User Interface (UI)

Az LCG-2 Grid belépési pontja. Ezen a gépen történik az LCG felhasználók account-jainak és a jogosítványainak tárolása. Ez tulajdonképpen egy kaput biztosít a Grid szolgáltatásainak használatára felé. A felhasználóknak a UI gépnél a hivatalos LCG-s bizonyítványukkal kell azonosítani magukat. Ezzel az alábbi szolgáltatások válnak számukra elérhetővé: job-ok beküldése (job submit), leállítás, keletkezett eredmények megtekintése, job-ok állapotlekerdezése, valamint a felhasználó által támasztott követelményeknek megfelelő erőforrások listázása, fájlok másolása.

2.1.2. Computing Element (CE)

A CE tulajdonképpen egy job végrehajtási erőforrás, azaz ütemezési sort definiál egy Grides erőforrás vagy erőforrás-halmaz – egy site – számára. Itt célszerű megjegyezni, hogy egy site-on belül nem csak egy, hanem több különböző ütemezési sor is definiálva lehet, mely sorok akár különböző CE-khez is tartozhatnak. Minden LCG-2 site-nak legalább egy CE-vel rendelkeznie kell. Minden CE homogén, ún. worker node-ok halmazára és a Grid Gate-re (GG) épül. Ez a homogenitás többek között a processzorra és memóriára értendő. Worker node-oknak nevezzük a tulajdonképpeni végrehajtó node-okat, GG-nek pedig a front-end node-ot, amelynek nemcsak belülről, de kívülről is elérhetőnek kell lennie. A GG-n többek között

a Globus gatekeeper, a Globus GRAM (Globus Resource Allocation Manager), valamint a lokális erőforrás menedzser (Local Resource Management System - LRMS) is fut. Az LCG-2-ben a támogatott LRMS az ún. Portable Batch System, azaz a PBS [3], a Load Sharing Facility [4] és a Condor [5].

A GG feladatai közé tartozik a job-ok fogadása és átadása a worker node-oknak végrehajtásra.

2.1.3. Storage Element (SE)

Az SE nagy tárolókapacitást nyújtó szolgáltatás, amely egyrészt egységes hozzáférést biztosít a Gridben tárolt adatokhoz, továbbá vezérli a különböző ún. mass storage rendszereket (MMS). Mindegyik LCG-2 site-nál legalább egy SE található. Az SE-en többek között a GSIFTP szerver fut, amely rendelkezik az FTP minden alap funkciójával kiegészítve azt a GSI biztonsági réteggel, ezzel lehetővé téve a biztonságos fájl műveleteket.

Az LCG-2 legutóbbi verziójában megtalálható az ún. Storage Resource Manager (SRM), amely dinamikusan menedzseli az SE-ek által tárolt adatokat, ezáltal könnyítve meg az adatok elérését.

2.2. Az LCG-2 információs rendszere

Az LCG-2 Grid információs rendszerének felépítését a 2.2. ábra mutatja. Az információs rendszer a GLOBUS MDS (*Monitoring and Discovery Service*) alapra épül. Az MDS-ben található GRIS, illetve GIIS az LCG-2 Gridben szintén megtalálható. Mindegyik CE-n, illetve SE-n egy olyan GRIS (*Grid Resource Information Servers*) fut, amely az adott CE-ről illetve SE-ről tartalmaz információkat. Mindegyik *site*-on belül, valamelyik SE-n fut egy *Site GIIS* (*Grid Index Information Server*). A *Site GIIS* a *site*-on levő összes CE és SE GRIS-ekben eltárolt információit tartalmazza. A *Site GIIS* kétpercenként lekérdezi az alárendelt GRIS-eket, így frissítve az adatbázisát.

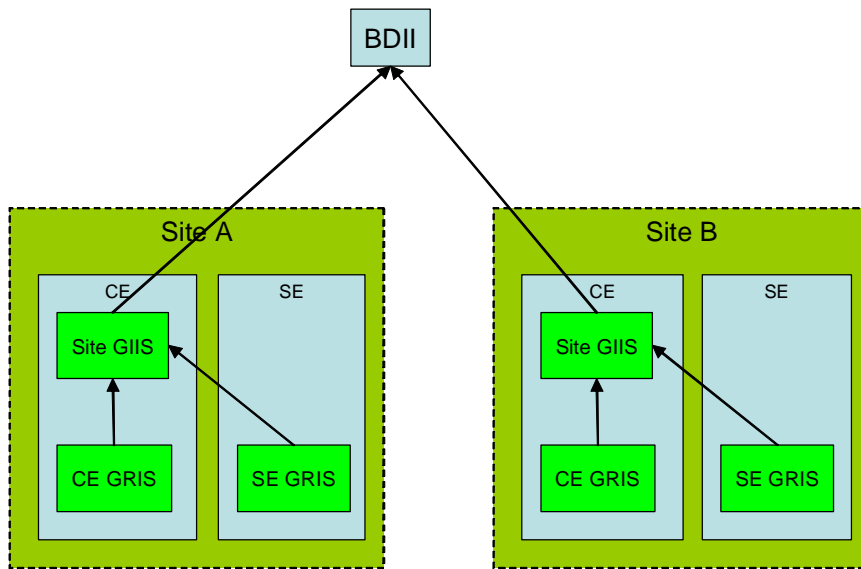
A GIIS-ek felett egy BDII (*Berkeley DB Information Index*) szolgáltatás (adatbázis) áll. Általában minden nagyobb *site* tartalmaz egy-egy ilyen adatbázist. A BDII szerver kétpercenként lekérdezi az alárendelt *Site GIIS* szervereket, és ezek alapján frissíti a BDII adatbázisát. A legtöbb BDII szolgáltatás úgy kerül konfigurálásra, hogy egy webszerveren található konfigurációs fájl alapján dolgozik. A konfigurációs fájl *site*-ok listáját tartalmazza. Így mindegyik BDII csak azokról a *site*-okról tárol információkat, amelyek szerepelnek annak konfigurációs fájljában.

A BDII-ktől a Grid felhasználóknak és különböző Grid szolgáltatásoknak (pl. erőforrás bróker) lehetőségük van információkat lekérdezni. Amennyiben friss információkra van szükség, célszerű közvetlenül a *Site GIIS*-ekhez fordulni.

Az erőforrásokról (jelen esetben a *site*-ok) az adatokat valamennyi szinten (GRIS, GIIS, BDII) a GLUE sémában (*Grid Laboratory for a Uniform Environment Schema*) kerülnek eltárolásra [6]. A séma a CE-k, illetve SE-k attribútumait és azok lehetséges értékeit definiálja. Az információs rendszer valamennyi szinten az LDAP protokoll segítségével biztosítja az információk elérését.

2.3. Az LCG-2 adat-menedzsmentje

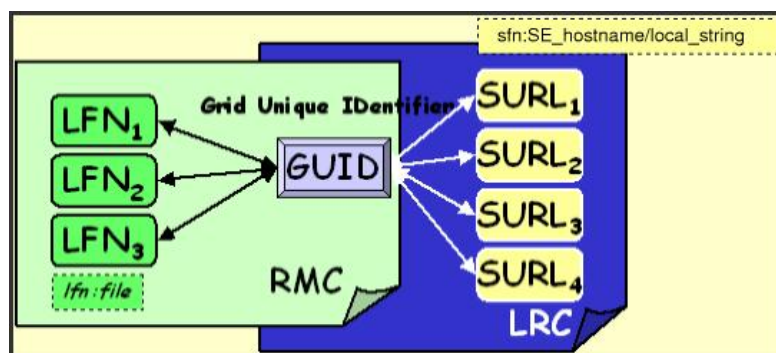
Az adat menedzsment réteg egy része a European DataGrid projekt [7] Replica Management (RMS) rendszerére épül, a másik része az LCG adat-menedzser programjára. GRID környezetben az adat fájlokat replikálhatjuk, azaz másolat készülhet róluk. A felhasználónak ill. az alkalmazásoknak nem szükséges tudniuk, hogy ténylegesen hol is helyezkednek el az adatok, ugyanis ők az adatok logikai nevét használják, mivel az adat-menedzsment szolgáltatás lesz a felelős a kért adatok névfeloldásáról.



2.2. ábra Az LCG-2 információs rendszerének sémája

A Gridben különböző azonosítókkal hivatkozhatunk a fájljainkra, úgy mint a Grid Unique Identifier (GUID), Logical Fájl Name (LFN), Storage URL (SURL) és a Transport URL (TURL). A GUID és az LFN kizárólag az eredeti fájlokra hivatkozik a replikált adatokra nem, és nem mond semmit a tényleges elérhetőségük. Ezzel szemben a SURL és a TURL a replikált adatok fizikai helyéről szolgáltat információt.

A fájlok azonosítása a hozzájuk tartozó egyéni GUID azonosító által történik. Ez az azonosító a fájl létrejöttkor automatikusan generálódik és rendelődik az adathoz, az egész Gridben egy egyéni azonosítót biztosítva ezzel a fájl számára. Az összes fájl replikációnak ugyanaz az azonosítója lesz, mint amit az eredeti fájl kapott a létrehozásakor.

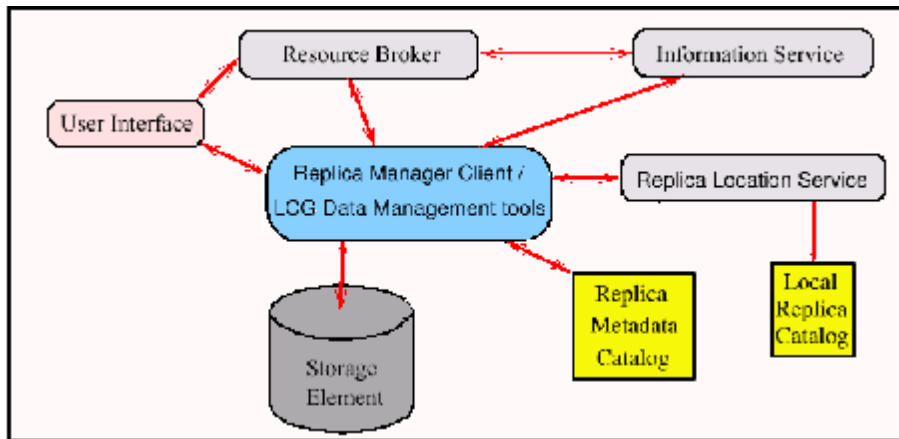


2.3. ábra Az LCG-2 által használt különböző fájl-azonosítók, forrás [8]

Az RMS a SURL azonosító használata által találja meg az egyes replikák fizikai helyét, továbbá ugyanezt az azonosítót használják a Storage Element-ek is a fájlok megtalálásához.

A TURL azonosító azokat a szükséges információkat nyújtja a replikákról, amelyek azok azonosítására szolgálnak. Szolgáltatja a szükséges host nevet, elérési utat, a protokollt és az eléréshez szükséges portot is ugyanúgy, mint egy hagyományos URL esetén. Így az egyes adatok könnyen elérhetőek és használhatóak. A 2.3. ábra az LCG-2-ben a fájlokra történő különböző hivatkozásokat mutatja be.

Maga az RMS két fő részből áll: Replica Location Service (RLS) és Replica Metadata Catalog (RMC). Az RLS felelős azokért az információkért, amelyek az adatok tényleges fizikai elhelyezkedését mondják meg. Ezek az információk különböző katalógusokba ún. Local Replica Catalog-ba (LRC) lesznek összegyűjtve, mely katalógusok Virtuális Szervezeteként tárolják a replika adatokat. Az RMC tárolja a GUID-k és a logikai fájl néven (LFN) létrehozott állományok közötti összerendeléseket, továbbá a metainformációkat, pl. méretek, dátumok és jogosultságok.



2.4. ábra A Replica-menedzser kapcsolatai a Grid egyéb komponenseivel, forrás [8]

Az utolsó bemutatni kívánt menedzsmenttel kapcsolatos szolgáltatás a Replica Manager. Az adat menedzsmentért felelős Replica Manager tulajdonképpen egy egyszerű interfészt valósít meg az RMS és a felhasználók ill. az egyéb szolgáltatások között (pl. Resource Broker) az 2.4. ábra. Az LCG-2-ben a Replica Manager integrálva lett az User Interface-szel.

2.4. Az LCG-2 job-menedzsmentje

LCG-2-ben a Workload-menedzsment (Workload Management System, azaz WMS) rendszer felelős a job-ok rendszerbe történő beküldéséért és a megfelelő Computing Element-hez történő irányításáért, figyelembe véve az egyes job-követelményeket és az elérhető szabad erőforrásokat. Ezért a WMS-nek rendelkeznie kell az ún. BDII és az RLS-által szolgáltatott információkkal. A Berkeley DB Information Index (BDII) az LCG-2 Gridet alkotó site-okról tartalmaz friss információkat (pl. elérhetőség, terheltség, éppen futó jobok száma). Az erőforrás bróker, Resource Broker (RB), az LCG-2-nek az a gépe, ahol a WMS szolgáltatások tulajdonképpen futnak. Ezek a szolgáltatások a következők:

Network Server (NS), amely fogadja az UI-től bejövő kéréseket, és támogatásokat nyújt a job vezérléséhez szükséges funkciókhoz,

Workload Manager, amely a job-menedzsment rendszer magja,

Match-Maker, vagy más néven Resource Broker, melynek feladata a job követelményeit kielégítő legjobb erőforrás megtalálása és kiválasztása,

Job Adapter, amely létrehozza a job számára szükséges egyéb környezeti feltételeket, mielőtt átadja a job-ot a Job Control (JCS) szolgáltatásnak,

Job Control Service (JCS), amely végrehatja a job menedzseléséhez szükséges műveleteket (job submit, job removal).

Az RB gépen futó további szolgáltatás az ún. LB (Logging and Bookkeeping service), amely nyilvántartja az összes job-menedzseléssel kapcsolatos Grid-eseményt.

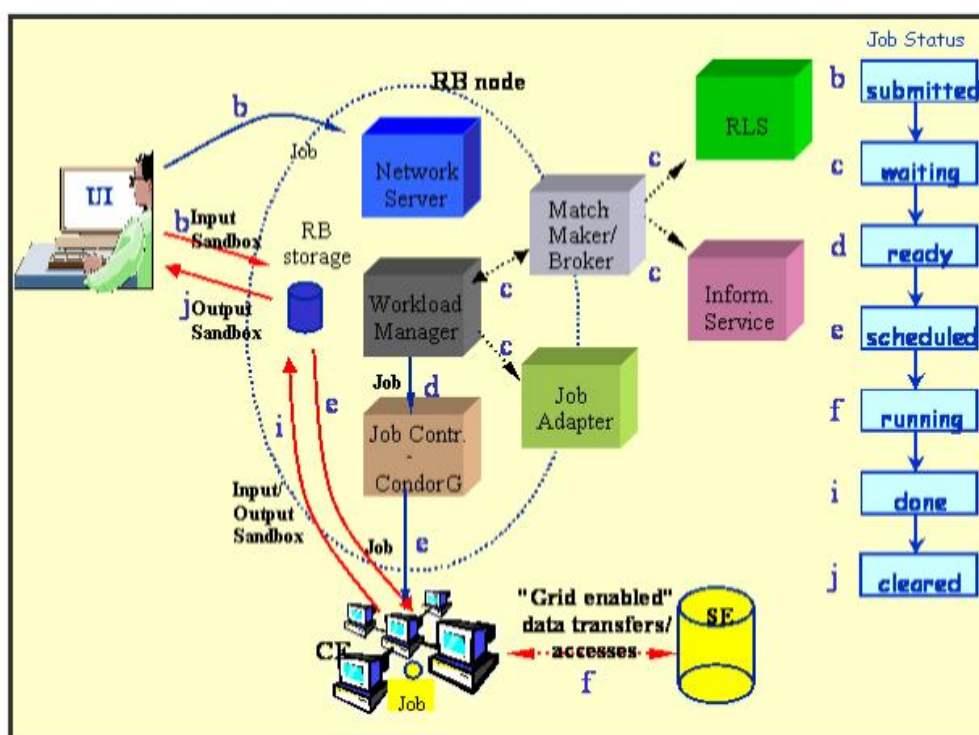
Az LCG-2 komponensei kapcsán meg kell még említeni a Proxy szervert (PS) is. Amikor egy felhasználó hozzáférést kap a Grid erőforrásainak és szolgáltatásainak használatához, akkor a számára kiállított

igazolvány segítségével létrehoz egy ún. proxy-t. A proxy egy ideiglenesen érvényes igazolvány, amely az érvényességi ideje lejárt esetén érvényét veszti. A felhasználó egy proxy legenerálása után ennek a segítségével tudja az LCG-2 Gridet használni és job-okat végrehajtásra távoli site-okra küldeni. Azonban, ha a felhasználó proxy-ja a job futásának befejezése előtt lejár, akkor a job futása nem folytatható. Ezt megelőzendő a WMS segítségével lehetőség van a használatban lévő proxy-k érvényességi idejének használat közbeni meghosszabbítására.

2.4.1. Szolgáltatások együttműködése egy job végrehajtása esetén

Ebben a részben röviden áttekintjük egy job végrehajtásának menetét, és a végrehajtásban érintett komponenseket. A 2.5. ábra ezt a folyamatsorozatot és komponenshalmazt mutatja.

Miután a rendszer használatához szükséges hosszú távú igazolvány a rendelkezésünkre áll, az abból generálható rövid távú proxy segítségével azonosíthatjuk magunkat a Grid UI gépénél.



2.5. ábra Egy job végrehajtása az LCG-2 rendszerben (forrás [8])

A felhasználó job-ját az UI-ról küldi (submit) a rendszerbe, ahonnan az a WMS segítségével a megfelelő számítási erőforrásokra kerül. A felhasználó egy job leíró fájl segítségével specifikálhatja, mely fájlokat szeretné a job-jával együtt kezelni (input/output fájlok). Ezeknek a mozgatni kívánt fájloknak a job leíró fájl egy ún. Input Sandbox-ot definiál. A job feladásának ténye adminisztrálásra kerül az LB-ben, és a job státusza SUBMITTED lesz. Ezek után a WMS megkeresi a job számára legmegfelelőbb CE-et. Ehhez a Resource Broker (Match Maker) lekérdezi a BDII-től az egyes számítási és tárolási erőforrásokat, valamint lekérdezi az RLS-től a felhasználni kívánt adatok helyét. A tényleges feladás után a job státusza WAIT-re vált, és az esemény rögzítődik az LB-ben.

A WMS Job Adapter felkészíti a job-ot a futtatásra, létrehozva egy ún. wrapper scriptet, amelyet átad a JCS-nek minden egyéb paraméterrel együtt. Ezek után a job státusza READY lesz.

A kiválasztott CE-n futó Globus Gatekeeper fogadja a kérést és végrehajtásra továbbküldi a job-ot a helyi erőforrás menedzsernek (LRMS). A job státusza SCHEDULED-re vált.

Az LRMS menedzseli a job végrehajtását az elérhető lokális worker node-okon. A felhasználó fájlljai a végrehajtó WN-ra másolódnak az RB-ről. Majd a job státusza RUNNING-ra vált.

Ha a futást biztosító site erőforrásai kiesnének, a job automatikusan egy új CE-re lesz elküldve, ezzel egy időben a job státusza ABORTED-re vált.

Amennyiben a job futása hiba nélkül fejeződik be, akkor a job által létrehozott kimeneti (azaz azon kimeneti fájlok, amelyeket a felhasználó meghatározott az Output Sandbox segítségével) vissza lesznek szállítva az RB-re. A job státusza DONE-ra változik. Ezzel párhuzamosan természetesen minden esemény rögzítésre kerül.

A felhasználó ezek után a job kimeneti fájlokhoz a WMS segítségével férhet hozzá és hozhatja azokat vissza a helyi gépére. Az eredmények letöltése után a job státusza CLEARED-re áll.

Az eddigiek összegzéseként megállapítható, hogy az LCG-2 a legnagyobb második generációs Grid, (eddig 82 összekapcsolt site, 7269 CPU) melyet széles körben használnak a fizikai tudományok különböző területein. Magyarországról az MTA KFKI Rézszecke- és Magfizikai Kutatóintézet [9] tartozik az LCG-2 site-ok közé, azonban a SEE-GRID projekt [10] keretén belül a SZTAKI-nál is alkalmazásra került az LCG-2 Grid middleware, melynek eredményeképpen az LCG-2 Grid rövidesen újabb magyar site-okkal fog bővülni.

3. HUNGRID

A HunGrid az RMKI, a SZTAKI és az ELTE által -az EGEE projekt keretén belül- újonnan létrehozott Magyar Virtuális Szervezet.

A HunGrid célja hogy összegyűjtse mindazokat a magyar LCG felhasználókat, akik még nem tagjai már létező támogatott virtuális szervezetnek ill. már tagjai, de újabb virtuális szervezetet szeretnének támogatni. A létrehozott HunGrid egy olyan virtuális szervezet, amely nyitva áll bárki előtt, aki támogatni és használni szeretné az LHC Grid-et akár tudományos, akár oktatási célokra.

Ahhoz, hogy a HunGrid Virtuális Szervezet tagjává váljunk egy regisztrációs eljárást kell végigjárnunk. Legelső lépésként érvényes tanúsítvánnyal kell rendelkezünk, amelyet valamely ismert és elfogadott tanúsítvány kiállító hatóság írt alá. (<http://marianne.in2p3.fr/datagrid/ca/ca-table-ca.html>).

3.1. Digitális felhasználói tanúsítvány igénylése

Felhasználói tanúsítvány a KFKI RMKI CA oldalon tudunk igényelni (<http://pki.kfki.hu>). Az igényléshez először is mindenképpen ajánlott elolvasni a KFKI RMKI CA működési szabályait rögzítő dokumentumot. Szükséges továbbá egy SSL-et támogató böngészőprogram (IE, Mozilla/Netscape és Opera frissebb változatai biztosan működnek). Kell a foglalkoztató intézmény (egyetem, kutatóintézet) részéről egy igazolás a munkaviszonyról vagy a szóban forgó kutatási programban való részvételről. A regisztráláshoz szükséges lépéseket a <http://pki.kfki.hu/pki/pub/userreq.hu.html> címen érhetjük el.

3.2. Digitális szerver tanúsítvány igénylése

Ahhoz hogy az LCG-2 rendszerünk csatlakozzon ill. regisztrálódjon az EGEE-ben vagy a CERN teszt zónájához valamint hogy egyáltalán az LCG-2 alacsony szintű Globus alapú rétegét használni tudjunk szerver tanúsítványra van szükségünk. Ez azt jelenti, ha gatekeeper-t (LCG-2 esetén a UI gépen) vagy

GridFTP szervert szeretnénk futtatni szükségünk van a tanúsítványra. A gatekeeper a Globus csomag része, amely fogadja az egyes jobok-at és foglalkozik a job és a tulajdonosának azonosításával.

A regisztráláshoz szükséges lépéseket a <http://pki.kfki.hu/pki/pub/serverreq.hu.html> címen érhetjük el.

3.3. A virtuális organizációk szervezése

Miután a digitális felhasználói tanúsítványok birtokában vagyunk job-jaink LCG-2 rendszeren történő futtatásához tagjának kell legyünk valamelyik virtuális szervezetnek. Egy ilyen szervezet olyan felhasználók köre akik ugyan olyan jellegű projekten dolgoznak és ugyanazokat az alkalmazásokat használják. Magyarországon a HunGrid teremti meg ezt a lehetőséget magába foglalva olyan különálló felhasználói csoportokat mint a kémikusok és fizikusok. Természetesen a felhasználók számának növekedésével új szervezetek létrehozása válik szükségessé, úgy mint kémikus virtuális szervezet vagy éppen a fizikusoké. A HunGrid-be tartozó LCG-2 erőforrások helyileg adhatnak prioritást alkalmazásaiknak, letiltva ezzel más külső alkalmazások végrehajtását erőforrásaikon. A szervezetek kialakítása szükségessé teszi a szervezet saját Resource Broker és BDII LCG-2 komponensek létrehozását, amelyek segítségével a saját szervezetünk menedzselését végezhetjük.

Tehát a Grid felhasználónak kérelemmel kell fordulnia valamelyik virtuális organizáció menedzseréhez, aki intézkedik a felvételtől. A HunGrid esetében ezt a <http://www.lcg.kfki.hu> honlapján elérhető "user registration" név alatt tehetjük meg ha rendelkezünk már érvényes felhasználói tanúsítvánnyal.

Új virtuális szervezet létrehozása esetén azonos érdeklődési körön belüli felhasználókkal kell rendelkezünk pl. tudományos közösség a kémia, biológia, fizika területekről és át kell menjünk néhány szükséges lépésen (<http://grid-deployment.web.cern.ch/grid-deployment/cgi-bin/index.cgi?var=gis/vo-deploy>) ahhoz, hogy az új virtuális szervezetünk be legyen jegyezve az EGEE-be. Röviden a következő alaplépéseket kell megtenni:

1. Saját virtuális szervezet szerver üzemeltetése és regisztrációs szolgáltatás nyújtása.
2. RLS (Replica Location Service) üzemeltetése vagy más olyan tag felkeresése, aki RLS-t üzemeltetne.
3. Megfelelő mennyiségű rendelkezésre álló erőforrás nyújtása, melyek a létrehozni kívánt virtuális szervezetet támogatják.

A teljes regisztrációhoz szükséges lépéseket az alábbi dokumentum tartalmazza: https://edms.cern.ch/file/503245/2/VO_Registration.doc

3.4. A HunGrid jelenlegi infrastruktúrája

A HunGrid jelenlegi tagjai az RMKI KFKI, MTA SZTAKI és jelenleg kialakítás alatt áll az MTA KKKI-nál a rendszer. Az RMKI KFKI infrastruktúrájára jellemző főbb paramétereket az 1-es és a 2-es táblázat mutatja be.

Server:	Elonex Disk Server
Mem:	1 GB
Proc:	Dual Intel Xeon (2.0 GHz)
Op.r.:	Red Hat 7.3
Disk:	20 GB (system) + 3.4 TB (data)

1. táblázat: Tároló kapacitás

Group0:	25 számítógép	Group1:	50 számítógép
Proc.:	Dual AMD ATHLON MP 2000+	Proc:	AMD ATHLON XP 1900+
Mem.:	1 GB	Mem:	512 MB
Disk:	60 GB	Disk:	120 GB
Op. r.:	Red Hat 7.3 (Valhalla)	Op. r.:	Red Hat 7.3 (Valhalla)

2. táblázat: Számítási kapacitás

Az MTA SZTAKI infrastruktúrájára jellemző főbb paramétereket a 3. táblázat mutatja.

Group0:	1 számítógép
Mem.:	1 GB
Proc.:	Intel Pentium 4 CPU (3.00 GHz)
Op.r.:	Scientific Linux 3.03
Disk:	20 GB (system) + 100 GB (data)

3. táblázat: SE,CE,UI

Group1:	12 számítógép
Mem.:	256 MB
Proc.:	Pentium III Katamai Dual (500 MHz)
Op.r.:	Scientific Linux 3.03
Disk:	10 GB (system)

4. táblázat: Számítási kapacitás

3.5. HunGrid jövőkép

A jelenlegi tagokon kívül a HunGrid több résztvevővel bővül. Párhuzamosan folynak a munkálatok az ELTE, MTA Kémiai kutatóközpont –mind Budapest mind Veszprém-, Szegedi Tudományegyetem valamint a Miskolci Egyetem keretein belül.

A Kémiai kutatóközpont HunGrid-es leendő erőforrásait az 5. és a 6. táblázat szemlélteti.

Group0:	1 számítógép
Mem.:	1 GB
Proc.:	Intel Pentium 4 CPU (3.00 GHz)
Op.r.:	Scientific Linux 3.03
Disk:	20 GB (system) + 100 GB (data)

5a. táblázat: Budapest: SE,CE,UI

Group0:	4 számítógép
Mem.:	1 GB
Proc.:	Intel Pentium 4 CPU (3.00 GHz)
Op.r.:	Scientific Linux 3.03
Disk:	20 GB (system) + 100 GB (data)

5b. táblázat: Budapest számítási kapacitás

Group0:	1 számítógép
Mem.:	256 MB
Proc.:	Intel Pentium 2 CPU (400 MHz)
Op.r.:	Scientific Linux 3.03
Disk:	8 GB (system+data)

6a. táblázat: Veszprém: SE,CE,UI

Group0:	2 számítógép
Mem.:	512 MB
Proc.:	Intel Pentium 3 dual CPU (450 MHz)
Op.r.:	Scientific Linux 3.03
Disk:	40 GB (system+data)

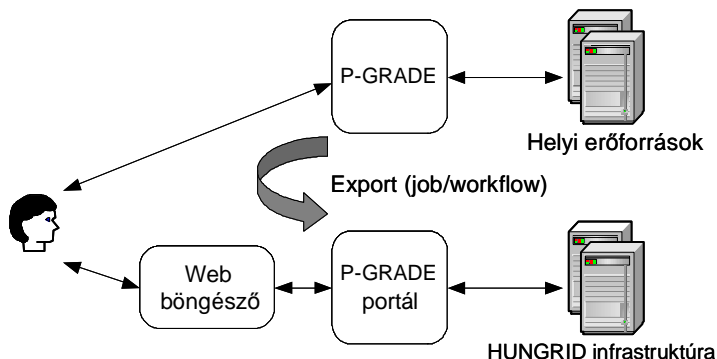
6b. táblázat: Veszprém számítási kapacitás

A HunGrid egy operatív magyarországi LCG-2 alapú Grid infrastruktúrát teremt meg a nemzetközi Grid erőfeszítésekkel összhangban, és a tudományos kutatómunkára vagy oktatásra használható egységes Grid nyitott lesz minden csatlakozni vágyó partner számára.

4. P-GRADE portál

A 2004-ben véget ért SzuperGrid projekt [11] egyik legjelentősebb eredménye a P-GRADE portál [12] kifejlesztése volt, ami a Magyar Szuperszámítógép Grid távoli, grafikus felületen keresztüli elérését teszi lehetővé. A portál segítségével ún. workflow-ba szervezett szekvenciális és párhuzamos job-okat lehet monitorozott módon, automatizálva végrehajtani.

A P-GRADE portál [12] grafikus felülete egy korábbi, szintén magas szintű alkalmazásfejlesztő környezet, a P-GRADE [13] koncepcióját követi. A tendenciák és a felhasználói tapasztalatok mind azt mutatták, hogy a számítási Gridok használata nem kényelmes ún. vastag klienseken keresztül a körülményes installálás és a folyamatosan változó Grid technológiákhoz való igazodás miatt. Ezen nehézségek leküzdése miatt került kialakításra a HUNGRID-hez is egy Web alapú hozzáférést biztosító P-GRADE portál változat, amely átveszi és az aktuális igények figyelembevételével kiegészíti a P-GRADE által nyújtott alapszintű támogatást az adott Grid infrastruktúra eléréséhez. A P-GRADE így továbbra is megmaradhat alkalmazásfejlesztő környezetnek, de minden olyan teendő, amely során a felhasználónak a Grid infrastruktúrát el kell érnie (az alkalmazás üzemserű használata közben) immár célszerűen a P-GRADE portálon keresztül történhet. A P-GRADE környezet és a P-GRADE portál közötti szerepmegosztást a 6. Ábra vázolja.



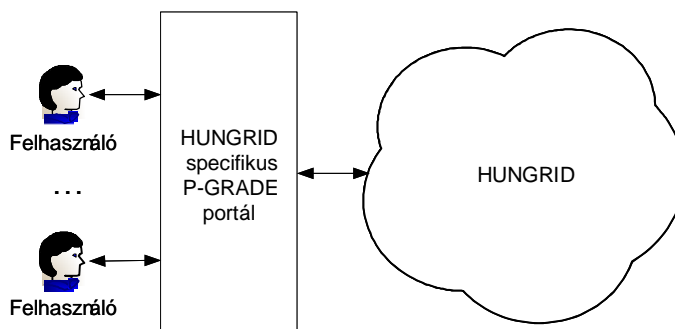
6. Ábra: A P-GRADE és a HUNGRID-et támogató P-GRADE portál szerepe

Amíg a hagyományos P-GRADE használható programfejlesztésre, a programok fordítására, hibakeresésre és javítására, addig a portál változat a job-ok végrehajtásához, Griden futtatható workflow-ba szervezéséhez és teljesítményanalíziséhez [14] ad magas szintű támogatást. A P-GRADE-et közvetlenül, a P-GRADE portált közvetve, egy Web böngésző segítségével lehet használni, de a két eszköz között lehetősége van a felhasználónak a P-GRADE exportálási funkcióira támaszkodva komplett workflow-kat, illetve párhuzamos alkalmazásokat átvinni a P-GRADE portál alá.

A SzuperGrid felhasználóktól érkezett kedvező visszajelzések alapján világossá vált, hogy a HUNGRID felhasználói körének szélesítését megcélzó munkáknak is jelentős lökést adhat az alacsony szintű részleteket Web alapú grafikus felülettel elfedő portál. Mivel a magyar SzuperGrid és a HUNGRID hasonló GLOBUS 2 alapú [15] infrastruktúrák, ezért első lépésben egy dedikált P-GRADE portált alapvető módosítások nélkül illeszthettük hozzá a HUNGRID-hez. Ezáltal a P-GRADE portál az LCG-2 alacsony GLOBUS 2 alapú rétegével működik együtt, WEB alapú belépési pontot teremtve és elérhetővé

téve a HunGrid felhasználók számára az LCG-2 erőforrásokat. Ennek megfelelően a HunGrid P-GRADE portál kialakításra került: <http://hgportal.hpc.sztaki.hu:8080/gridsphere/gridsphere>.

A közeljövőben tervezzük, a HUNGRID, jórészt LCG-2 specifikus portleiteinek kidolgozását, hasonlóan a KlaszterGRID-hez [19] történt illesztéshez megvalósítva ezzel az LCG-2 felső rétegével való illesztést. A koncepciót a 7. Ábra szemlélteti:



7. Ábra: A P-GRADE portál szerepe a HUNGRID-ben

4.1. A Mercury Grid monitor

A Mercury Monitor [16] feladata a grid-teljesítmény felügyelete: ellenőrző adatokat biztosít metrika szerint ábrázolva, mind pull- mind push-típusú hozzáférési szemantika szerint, ezenkívül támogatja a vezérlőkön keresztüli irányítást. Általános, kiterjeszhető és skálázható módon támogatja olyan grid elemek felügyeletét, mint például erőforrások és alkalmazások. Kivitelezésekor fő szempont volt a modularitás, különös hangsúlyt fektetve az egyszerűsége, hatékonyságra, szállíthatóságra és a felügyelt rendszer minél kisebb fokú terhelésére.

A Mercury Grid monitor segítségével a párhuzamos Grid alkalmazások processz-szintű monitorozása valósulhat meg, tehát megvalósulhat az LCG-2 erőforrásokon futtatott alkalmazások workflow és processz szintű monitorozása egyaránt.

5. A HUNGRID kémiai célú felhasználása

Számítógépes modellezés alkalmazása kémiai problémák vizsgálatára az utóbbi években általánosan használt és elfogadott módszerré vált. A kémiai reakciók leírása egyrészt a molekulák belső szerkezetének, valamint az atomok illetve molekulák kölcsönhatásainak részletes meghatározását, másrészt pedig a reagáló atomok, molekulák mozgásának következtében megfigyelhető makroszkopikus jelenségek megfelelő kezelését igényli [17]. A molekulaszerkezet meghatározása az úgynevezett elektron-Schrödinger-egyenlet megoldásával, az atomok és molekulák mozgásának leírása pedig a klasszikus vagy kvantum-mechanikai mozgásegyenletek megoldásával, majd a statisztikus mechanika alkalmazásával történik. A reakciók lefolyásáról így kapott adatokat a sok reakcióból álló makroszkopikus rendszerek modellezésére használják [18]. A gyakorlatban a kémiai reakciók nem elkülönülten, hanem a folyadékok vagy gázok áramlásával és az anyagok diffúziójával csatoltan hatnak. Sok fontos jelenség, mint például a szennyezőanyagok terjedése és kémiai átalakulásai a levegőben csak csatolt reakció-diffúzió-advekciónal modellezésével értelmezhetők. Mindezek az eljárások rendkívül számításgényesek. Míg a molekulaszerkezet meghatározására szolgáló úgynevezett kvantumkémiai módszerek jól programozhatók úgy, hogy a különböző felhasználók jól tudják ugyanazt a programot használni és ezért kereskedelmi

szoftverek állnak rendelkezésre, addig a molekulák mozgását és az összetett reakciórendszerek leírását sokkal kevésbé lehet univerzális programokkal megoldani. Ennek megfelelően az ilyen problémák megoldásához egyedi fejlesztésű algoritmusokat és programokat használnak és a hatékony programozás legtöbbször a fejlesztő-felhasználó feladata. Az ebbe a problémakörbe tartozó programok számításgénye messze felülmúlja a szokásos munkaállomások által nyújtott számítási kapacitást. Az OM által támogatott „Kémiai grid kialakítása és alkalmazása a légszennyezés rövid és hosszú távú előrejelzésében” IKTA-5 projektben [22] többek között a levegőszennyezés szimulációjával foglalkoznak a partnerek, amihez reakció-diffúzió-advekción rendszerek számítógépes modellezésére alkalmas programok adaptálása szükséges a Gridre.

A fejlesztés Linux klasztereken történt, mivel tapasztalataink alapján a fejlesztési idő alatt ezek alkalmazása a leggazdaságosabb. A tényleges futtatások túlnyomórészt Grid-en történnek, a rövid távú légszennyezetség előrejelzőt kivéve – ez utóbbi az OMSZ szuperszámítógépén operatív módon kerül alkalmazásra. Ez azért fontos, mert a párhuzamosított, nagy számításgényű modellek tényleges futtatásához már a klaszterek sebessége sem elég és így ezeket már egyszerre több klaszteren, azaz egy Grid rendszerben – kisebb alkalmazások és szigorú futási időkorlátok esetében szuperszámítógépen - érdemes futtatni. A klaszterről a Grid rendszerre történő áttérés megvalósításában döntő szerep jutott a projektben a P-GRADE és P-GRADE portál rendszereknek, melyek egyrészt támogatják a párhuzamos programok hordozhatóságát, másrészt hatékony, magas szintű fejlesztőkörnyezetet biztosítanak az eltérő architektúrákon szükséges teljesítményhangoláshoz.

5.1. Levegőszennyezés előrejelző alkalmazás

A projekt keretében az ELTE FKT és MTA SZTAKI közösen kidolgozott egy programot a reakció-diffúzió-advekción egyenletek megoldása, passzív nyomanyagok (radioaktív nuklidok) terjedésének szimulációjára. A reakció-diffúzió-advekción egyenletek matematikailag másodrendű parciális differenciálegyenletek, amelyeknek az általános alakja:

$$\frac{\partial c_i}{\partial t} = \nabla(D_i \nabla c_i) - \mathbf{u} \nabla c_i + R_i, \quad 1)$$

ahol t az idő, c_i az i -dik anyagfajta koncentrációja, D_i az i -dik anyagfajta turbulens diffúziós együtthatója, R_i a kémiai reakciókat leíró tag, \mathbf{u} az advekción (szél-) mezőt leíró vektor. ∇ a Nabla operátor.

Ezekhez az egyenletekhez az adott feladattól függő kezdeti- és peremfeltételeket kell megadnunk. Legtöbb esetben a reakció-diffúzió-advekción egyenleteket, hasonlóan az egyszerűbb felépítésű reakció-diffúzió egyenletekhez, nem lehet analitikusan megoldani. Az egyenletek megoldása numerikusan történik és a numerikus megoldás egyik hatékony módszere az ún. "method of lines". Ilyenkor a reakció-diffúzió-advekción egyenletek megoldásához a szimulálandó tér diszkretizációja szükséges, amelynek során a fizikai teret (jelen esetben a síkot) cellákra (rácshálózatra) bontjuk, majd a tér mindkét térbeli változó irányában megvizsgáljuk a reakción, advekción és a diffúzió okozta koncentrációváltozásokat. A párhuzamosított algoritmus lényege, hogy a tartományt felbontjuk egyenlő részekre és az egyes processzek egymástól függetlenül számítják az adott tartományban végbemenő koncentrációváltozásokat a differenciálegyenletek segítségével.

A diffúzió és advekción rövid kölcsönhatási távolsága miatt csak a szomszédos cellákban lejátszódó folyamatok hatnak egymásra, ami elősegíti a probléma futtatását többprocesszoros rendszereken. A diffúzió és az advekción miatt minden időlépésben a szomszédos tartományok határainak koncentrációvektorát át kell küldeni a szomszédos részeknek.

Az eredmények felhasználásának lehetőségei köze tartozik, hogy magasan elhelyezkedő pontforrások esetén a légszennyeződések terjedésének modellezésére gyakran alkalmazható a Gauss-féle egyenlet (Gauss-

modell). Ez az egyenlet határesetként levezethető a (1) kontinuitási egyenletből, néhány egyszerűsítő feltétel alkalmazásával, amelyek a következők. A meteorológiai helyzet stacionárius, a forrást pontszerűnek tekintjük, amelynek kibocsátása folytonos és időben állandó, a földfelszín sík és a szennyezőanyagoknak nincs ülepedésük, a szélmezőnek csak az egyik ((x) irányú) komponense nem nulla, az adott irányban csak advekciónak van, turbulens diffúzió nincs, a kémiai átalakulásokat figyelmen kívül hagyjuk.

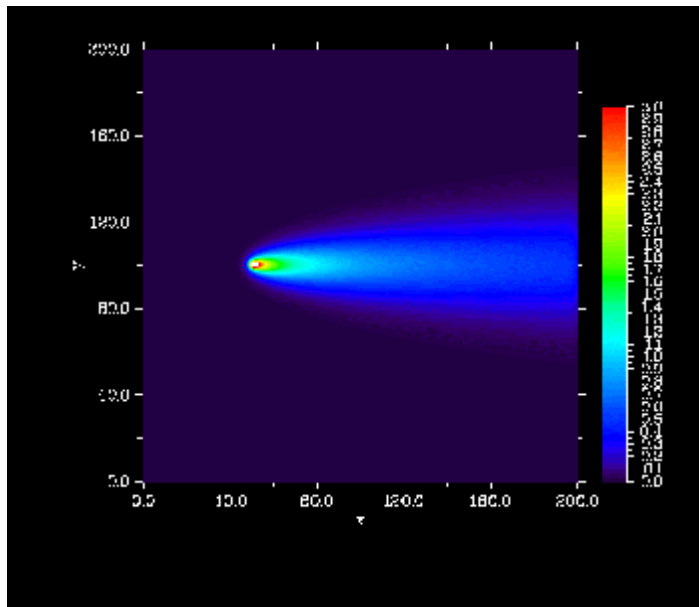
A fenti feltételek teljesülése esetén az (1) egyenlet megoldása a következő:

$$\bar{c}(x, y, z) = \frac{Q}{2ps_y s_z V_x} \exp\left(-\frac{1}{2}\left(\frac{y}{s_y}\right)^2\right) \exp\left(-\frac{1}{2}\left(\frac{z-h_e}{s_z}\right)^2\right) + \exp\left(-\frac{1}{2}\left(\frac{z+h_e}{s_z}\right)^2\right) \quad 2)$$

ahol Q a kibocsátott szennyező anyag mennyiségét, h_e az effektív kéménymagasságot jelöli és s_y és s_z szóródási paraméterek, amelyek a pontforrástól való távolság és a légköri stabilitás függvényében empirikus úton meghatározhatók. A Gauss-egyenletet (2) használva számítható a pontforrásból származó légszennyező koncentrációjának térbeli eloszlása. Az egyszerűsítő feltételezések miatt a Gauss modell alkalmazása korlátozott, összetettebb rendszerek esetén a részletesebb és flexibilisebb, Euler-típusú terjedési modellek alkalmazhatóak.

5.1.1. Helyi erőforrásokat használó verzió

Egy pontforrásból származó passzív anyag terjedését mutatja a 8. Ábra, A pontforrás az (50.0, 100.0) koordinátájú pontban található. A szimuláció során felhasznált bemeneti paraméterek (2d.inpl fájl): $h = 1000$ m (forrás magassága), $\Delta t = 900$ s (időlépcső), $D = 50$ m²/s (kibocsátás), $u = 5$ m/s (szélmező vízszintes komponense), $v = 0$ m/s (szélmező függőleges komponense).



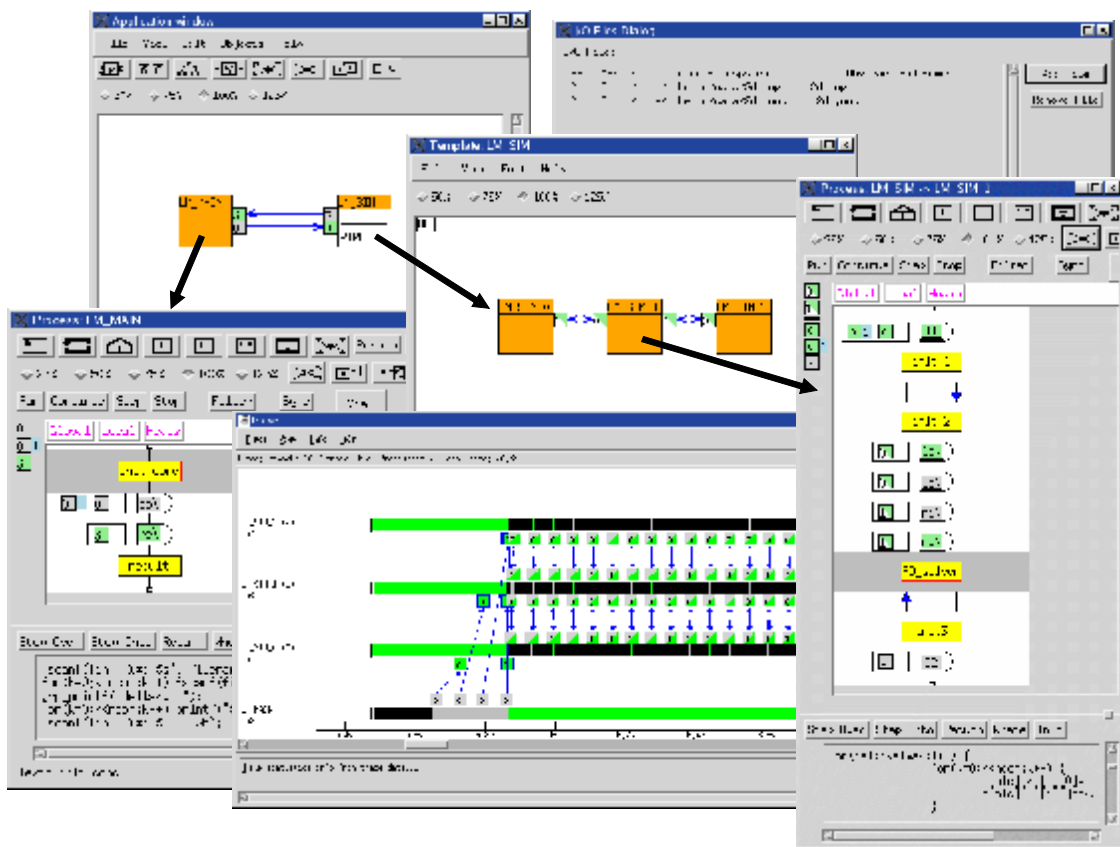
8. Ábra: A szimuláció eredménye *dislin* programcsomag segítségével vizualizálva

Ahogy azt a 9. Ábra is mutatja a P-GRADE-en belül az *I/O Files Dialog* ablakban, a kimeneti értékeket a *2d1.out* fájlba írja az LM_MAIN processz, amit később a *dislin* programcsomaggal vizualizálhatunk (ld. 8. Ábra).

A folyamatok legfelső (alkalmazás) szintű kommunikációs topológiáját a 9. Ábra, *Application window* mutatja. Az LM_MAIN processz felelős a bemeneti paraméterek beolvasásáért, a kezdeti értékadásokért, az információ szétszétlásáért a szimulációt végző processzek között a két 0-ával jelölt kommunikációs port és a közöttük kialakított kommunikációs csatorna segítségével. Később az eredmények összegyűjtését is ez a folyamat végzi a 3-sal jelölt kommunikációs portok és a közöttük definiált kommunikációs csatorna segítségével, majd a megfelelő kimeneti fájlba írja a szimuláció eredményét. A LM_SIM egy skálázható PIPE kommunikációs topológia, ami a már említett szimulációs lépések elvégzésért felelős.

Az LM_MAIN processzen belül (9. Ábra, *Process: LM_MAIN*) a kezdeti deklarációkat és értékadásokat követően definiáljuk a kezdeti feltételeket (*init cond*), majd egy multicast üzenettel elküldi a szimulációt végző processzeknek a szükséges információkat (*co*), majd egy másik kollektív kommunikáció segítségével (*re*) begyűjti a részeredményeket, végül az eredmények értékelése és kiírása következik (*result*). A szerkesztőablak alján kezdeti feltételek meghatározásához szükséges kód egy rövid részlete látható.

A PIPE szervezésekor (9. Ábra, *Template: LM_SIM*) minden egyes processz a fizikai tér egy szegmensét jelöli, amelyek minden időlépcsőben kommunikálnak egymással, átadva a határoló részekben a koncentrációvektorokat. Ezért szükséges a kétirányú kommunikációs csatorna és a ki/bemeneti kommunikációs portok alkalmazása ezen a szinten.



9. Ábra: A levegőminőségi alkalmazás klaszter verziója P-GRADE alatt

A reakció-diffúzió-advekció egyenletet megoldó egyik processz szerkezetét a 9. Ábra, *Process: LM_SIM* → *LM_SIM_1* mutatja. A kezdeti inicializálás után (üzenetek fogadása az *II* –en keresztül, majd *init_1* szekvenciális ikon) az adott szegmensre vonatkozó kezdeti feltételekkel ciklusba szervezve oldjuk meg a problémát. Minden időlépcsőben meghatározzuk a határfeltételeket (*init_2*), a szegmensek közötti határon a koncentráció vektorokat átadjuk a szomszédoknak (*le* és *re* kommunikációs akciók), majd következik a transzport és a kémiai reakció tag numerikus megoldása a bevezetőben ismertetett módszerek alapján (*RD_solver*). A szerkesztőablak alján a numerikus megoldó kód egy rövid részlete látható. A ciklus végén a folyamat visszaküldi a részeredményeket az *LM_MAIN* processznek.

A párhuzamos program teljesítményanalízise 200 iterációval és a numerikus szimuláció 200×200 felbontású négyzetgrácson történt az MTA SZTAKI intézeti klaszterén. A 9. Ábra, *PROVE* ablaka az 5,5 és 6,8 másodperc közötti időszakot mutatja, ahol jól megfigyelhető, ahogy a kezdeti értékek átadása és a szimulációs lépések egymás után történnek.

5.1.2. A HUNGRID infrastruktúrán futó változat

A már korábban ismertetett reakció-diffúzió-advekció hatásokat szimuláló kémia program alkalmas arra, hogy hatékonyan végrehajthassuk a HUNGRID-en is a P-GRADE portál segítségével. A HUNGRID verzió alapötlete, hogy egyszerre több szimulációt (ún. job-ot, mint pl. *HG_LM1*, *HG_LM2* és *HG_LM3*) hajtunk végre párhuzamosan a HUNGRID különböző távoli erőforrásain, különböző bemeneti paraméterekkel vizsgálva a reakció-diffúzió-advekció hatásait (ld. 10. Ábra, *Workflow editor: HUNGRID_LM*).

A különböző bemeneti paramétereket egy külön job készíti elő, az *HG_INIT*, a felhasználó által megadott paraméterfájlok alapján. Mivel ez a feladat viszonylag kevés erőforrást igényel elég, ha egy hagyományos szekvenciális program végzi el ezt a munkát. A futás következő fázisában az előkészített fájlok automatikusan átmásolásra kerülnek a szimulációt végző job-ok erőforrásaira.

A szimulációt végző kód a P-GRADE-del készült MPI fordítási opció, job és monitorozási üzemmód kiválasztásával. Jelen példában az *HG_LM1* és *HG_LM3* job-hoz a szimulációs program 5 hosszúságú PIPE-ot tartalmazó verziója tartozik, a harmadik job kisebb 3 hosszúságú PIPE-pal rendelkezik. Mivel a P-GRADE program tartalmaz a PIPE-on kívül még egy felhasználói processzt és egy a rendszer által létrehozott (de a felhasználó elől elrejtett) adminisztrációs processzt is, így 7 szerepel a „*Process Number*” mezőben az *HG_LM1* párhuzamos MPI job-hoz tartozó dialógusablakban (ld. 10. Ábra, *HG_LM1 properties*). A csatlakozó be- és kimeneti fájlok változatlanul *2d.inp* és a *2d1.out*, ahogy azt már az előző fejezetben ismertettük. Továbbá a dialógusablakban beállításra került, hogy a párhuzamos változat a KFKI RMKI erőforrásán fusson a HUNGRID-en belül, a *-n* és *-m* parancssori argumentumok pedig jelzik, hogy a program most nem lokális üzemmódban fut a P-GRADE felügyelete alatt.

Miután mindhárom job elkészítette a munkáját, a *HG_TIFF* job felelős azért, hogy a szimulációs eredmények könnyen értelmezhető grafikus reprezentációját előállítsa (8. Ábra). Mivel ez a feladat is a szimulációhoz képest kevesebb erőforrást igényel, ezért a képgeneráló program is egy meglévő szekvenciális program, ami a *dislin* szabad forráskódú vizualizációs programcsomagra építve végzi el a feladatát.

The screenshot displays the P-Grade portal interface. At the top, there is a navigation bar with 'Workflow Manager' and 'Job Properties' tabs. The 'Workflow Manager' tab shows a table of workflows:

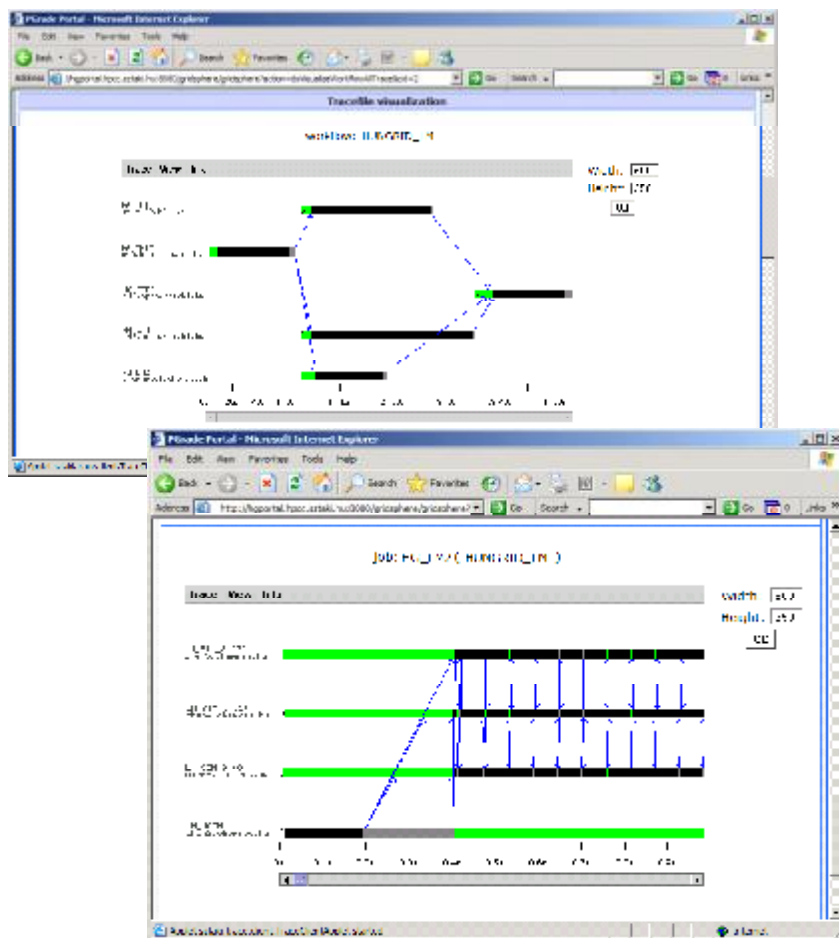
Workflow	Job	Hostname	Status	Last Comp. P	Workflow	Action
HUNGRID_01						
HG_INIT	hg0.hpc.sztak.hu	submit				
HG_INIT	hg0.hpc.sztak.hu	submit				
HG_INIT	hg0.hpc.sztak.hu	submit				
HG_LMD	hg0.hpc.sztak.hu	submit				
HG_TIFF	hg0.hpc.sztak.hu	submit				

The 'Workflow Editor' window shows a flowchart with the following jobs: HG_INIT, HG_LMD, HG_LMD, HG_TIFF, and HG_TIFF. The 'Job Properties' dialog box is open, showing details for the 'HG_INIT' job, including its name, type, and job description.

10. Ábra: A levegőminőségi alkalmazás HUNGRID-en futó változata P-GRADE portál alatt

A kémiai szimulációs program HUNGRID verziójának lefutása jól nyomon követhető a portál által szolgáltatott idődiagram alapján (ld. 11. Ábra). Megfigyelhető, hogy az *HG_INIT* job megkapta a felhasználó által beállított bemeneti fájlokat, majd lefutása után mindhárom szimulációs job megkapta a szükséges bemeneti paramétereket. Végül a *HG_TIFF* job sikeresen feldolgozta a szimulációs job-ok eredményeit. Az eredmények tömörített ZIP fájl formátumban a felhasználó rendelkezésére állnak; mind a közbenső részeredmények, mind az eredmények grafikus reprezentációja.

A párhuzamosan végrehajtott szimulációs job-ok esetén on-line módon is meg tudjuk vizsgálni az egyes processzek lefutását a távoli Grid erőforrásokon is, ahogy azt az alsó ablak mutatja (11. Ábra) az MTA KKKI HUNGRID-be már kísérletileg integrált klasztere esetén. Megfigyelhető, hogy a HUNGRID-en futó alkalmazás hasonlóan jó teljesítménymutatókkal rendelkezik, mint az előző fejezetben bemutatott lokális erőforrást használó verzió.



11. Ábra: Teljesítmény-vizualizáció a HUNGRID-en

6. Összefoglalás

Ez a cikk az EGEE projekt keretén belül létrehozott Magyar Virtuális Szervezetet a HunGrid-et mutatta be, valamint egy konkrét alkalmazásával foglalkozott. Mivel a Grid egyre fontosabb szerepet kap a katasztrófák következtében kialakult szituációk kezelésében [20][21], a HUNGRID-en sikeresen futtatott levegőkémiai szimulációs kísérletek ehhez az erőfeszítésekhez nyújthatnak hatékony segítséget.

Többek között összefoglalta még a cikk az EGEE Grid middleware az LCG-2 legfontosabb komponenseit, kitért a HunGrid-ben lévő intézmények infrastruktúrájának ismertetésére, a virtuális szervezetek szervezési módjára. A cikk foglalkozott a HunGrid WEB alapú belépési pontját megteremtő P-GRADE portállal és a hozzá kapcsolódó monitorozást végző Mercury Monitor-ral.

Összefoglalva, a magyarországi LCG-2 alapú Grid megteremtésével létrejött egy új, nemzetközi viszonylatokban is korszerű, felhasználóbarát, mindenki számára nyitott és operatív magyar Grid infrastruktúra.

7. Rerenciák

- [1] EGEE, www.eu-egee.org
- [2] X.509 certificate-alapú azonosítás: <http://www.networksorcery.com/enp/data/x509.htm>
- [3] PBS job management: <http://www.openpbs.org>
- [4] Load Sharing Facility, LSF job management: <http://accl.grc.nasa.gov/lfs>
- [5] Condor job management: <http://www.cs.wisc.edu/condor>.
- [6] Glue schema: <http://www.cnaf.infn.it/~sergio/datatag/glue/v11/CE/index.htm>
- [7] European DataGrid projekt <http://eu-datagrid.web.cern.ch/eu-datagrid>
- [8] LCG-2 User Guide http://egee.itep.ru/User_Guide.html
- [9] SEE-GRID projekt www.see-grid.org
- [10] MTA KFKI Részecske- és Magfizikai Kutatóintézet LCG-2 site <http://www.lcg.kfki.hu>
- [11] P. Kacsuk: Hungarian Supercomputing Grid, Proc. of ICCS'2002, Amsterdam. Springer Verlag, Part II, pp. 671-678, 2002
- [12] R. Lovas, G. Dózsa, P. Kacsuk, N. Podhorszki, D. Drótos: Workflow Support for Complex Grid Applications: Integrated and Portal Solutions, In: Grid Computing – Second European AcrossGrids Conference, AxGrids 2004, Nicosia, Cyprus, Lecture Notes in Computer Science, Vol. 3165, pp. 129-138, Springer-Verlag, 2004
- [13] Kacsuk, P., Dozsa, G., Kovacs, J., Lovas, R., Podhorszki, N., Balaton, Z., Gombas, G.: PGRADE: a Grid Programming Environment. Journal of Grid Computing Volume 1, Issue 2, 2003, pp. 171-197
- [14] Balaton, Z., Gombas, G.: Resource and Job Monitoring in the Grid. Proceedings of EuroPar' 2003 Conference, Klagenfurt, Austria, pp. 404-411
- [15] Globus Toolkit, www.globus.org/toolkit
- [16] Mercury Monitor, <http://www.lpds.sztaki.hu/mercury>
- [17] Bencsura, A., Lendvay, Gy.: Parallelization of reaction dynamics codes using P-GRADE: a case study. Computational Science and Its Applications, ICCSA 2004, LNCS, Vol. 3044, pp. 290-299
- [18] Lovas, R., Kacsuk, P., Lagzi, I., Turanyi, T.: Unified development solution for cluster and grid computing and its application in chemistry. Computational Science and Its Applications, ICCSA 2004, LNCS, Vol. 3044, pp. 226-235
- [19] P. Stefán: The Hungarian ClusterGrid Project, Proc. of MIPRO'2003, Opatija, 2003
- [20] Ladislav Hluchy, Viet Tran, Branislav Simo, Ondrej Habala, Jan Astalos, Emil Gatial: Flood Forecasting in CrossGrid project, Proceedings of 2nd European Across Grids Conference, Nicosia, Cyprus, 2004
<http://grid.ucy.ac.cy/axgrids04/AxGrids/Papers/E00-1408001255.pdf>
- [21] Next Generation Grid report: ftp://ftp.cordis.lu/pub/ist/docs/ngg2_eg_final.pdf
- [22] <http://www.lpds.sztaki.hu/chemistrygrid>