

A grid technológia alkalmazásának lehetőségei a mezőgazdaságban és a vidékfejlesztésben

Herdon Miklós, Salga Péter – Debreceni Egyetem, Gazdasági- és Agrárinformatikai Tanszék

Kivonat

A mezőgazdasági döntések és a vidékfejlesztési igények sokszor különböző típusú adatok feldolgozását, hozzáférését és integrálását teszik szükségessé, mely jelentős számítógépes kapacitást igényel. A mezőgazdaságban ilyen kérdések lehetnek többek között, hogy milyen vetőmagot vessünk és milyen területen, mikor öntözzük és milyen mértékben, permetezzük-e vagy sem, mikor takarítsuk be a termést stb. A másik területen, a vidékfejlesztésben pedig az egyik legfontosabb igény az információk és ismeretek megosztása, és ezen adatok jogosultsághoz kötött gyors és egyszerű elérése.

Az egyre szélesebb körben elterjedő grid technológia egy jó megoldást kínálhat az adatok feldolgozására és összegyűjtésére a különböző alkalmazásokban, pl. a döntéstámogatásban, tudásbázis építésben vagy a monitorozó rendszerekben.

A grid képes egyesíteni a földrajzilag és szervezetenként különálló számítógépes kapacitásokat. Nemcsak a processzorokat és a tároló egységeket, hanem a kommunikációs rendszereket, valós idejű adatforrásokat, mérőeszközöket és az együttműködő szakembereket és kutatókat is. Képes maximalizálni az adat és alkalmazáshasználatot centralizáció nélkül, egyszerűvé teszi a fenntartást és segítségével dinamikusan integrálhatóak a különböző helyről származó adatok.

Prezentációnkban szeretnénk bemutatni a grid technológia alapvető fogalmait - az alkalmazásokat hangsúlyozva - a mi értelmezésünkben. Bemutatunk néhány alkalmazást, amely segítséget nyújthat a mezőgazdasági döntésekben és vidékfejlesztési igények kielégítésében. Az AgModel projekt keretén belül a nyílt forráskódú MetBroker, DemBroker, ChizuBroker, SoilBroker programok képesek összegyűjteni a meteorológiai-, tengerszint feletti magassági-, térképészeti- és talajadatokat egy közös virtuális adatbázisba. Szeretnénk bemutatni a tanszékünkön telepített klasztert, mely terveink szerint a NorduGrid-hez csatlakozik az ARC middleware segítségével, a tesztelt programozási környezeteket és az alkalmazásfejlesztési terveinket.

Döntéstámogató rendszerek a mezőgazdaságban és az élelmiszeriparban

A mezőgazdasági döntések meghozásához rendkívül összetett információkra van szükség, amelyek különböző forrásokból szerezhetők be és figyelembe kell venni, hogy sztochasztikus rendszerben és rendszerkörnyezetben kell a döntés-előkészítési vizsgálatokat elvégezni a döntést meghozni. Egy nagyobb gazdaságban stratégiai döntéshelyzetek lehetnek például, hogy milyen növényt, milyen területen termesszünk, milyen vetőmagot használjunk, milyen természetstechnológiát alkalmazzunk, milyen kockázatokra kell felkészülni (fagykár, aszály, stb.).

A számtalan operatív döntéshelyzetre példa lehet még, hogy egy váratlan kártevő megjelenése esetén milyen védekezési módszert használjunk.

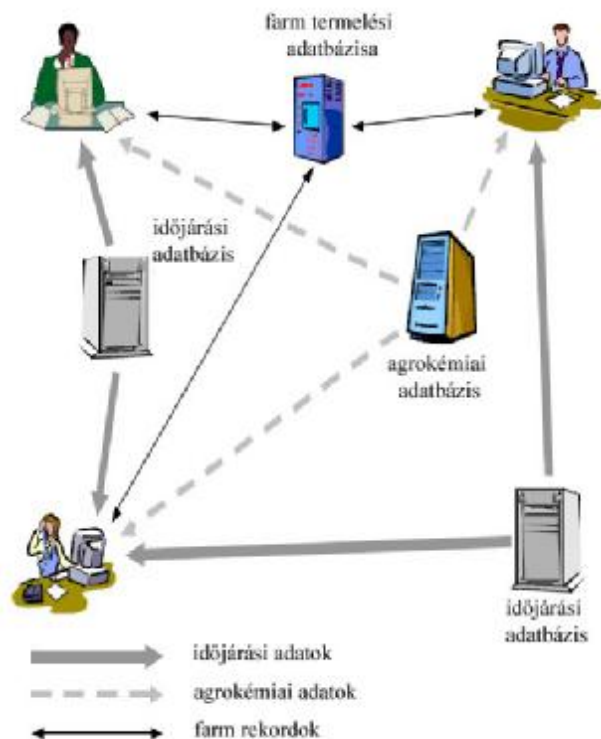
A döntésekhez szükséges adatok megtalálhatóak a különféle mezőgazdasági, meteorológiai, földrajzi, talajtani és szakmai adatbázisokban. Viszont a megfelelő információ nehezen nyerhető ki ezekből az adatbázisokból, illetve az adatintegráció az egyetlen mód arra, hogy a hatékonyságot növelni lehessen. Máskor az adatokon alapuló mintaelemzéses becslés szolgáltatja a legfontosabb információt, ehhez viszont komoly számítási kapacitás szükséges. Az adatok áramlásának egy további fontos gátja a kereskedelmi verseny, amely következtében a gazdálkodók megtévesztő adatokat szolgáltatnak a hivataloknak.

Az élelmiszeriparban a többi iparághoz hasonló döntéshelyzetek mellett (árképzés, erőforrás-ütemezés, diszpozíciók, logisztikai feladatok), kiemelt fontosságúak a a minél magasabb termékminőséget, termékbiztonságot segítő minőségbiztosítási rendszerhez kapcsolódó feladatok. A beszállítói lánc menedzselése nehezen megoldható feladat egy hagyományos vállalatirányítási rendszer számára, mivel itt is jelentős tényező az adatok nem megfelelő szolgáltatása és az adatintegrációs problémák széles köre.

A globalizáció és az erőforrások szétforgácsolódása miatt a problémák csak növekednek, az informatikai technológiák fejlesztése viszont segítséget nyújthat az élelmiszeripar és a mezőgazdaság ezekhez hasonló problémáinak megoldásához, illetve csakis informatikai támogatással tehető a döntéstámogatás hatékony eszközzé a mezőgazdaságban (**1. ábra**)

A mezőgazdasági adatok jellemzői:

- szétszórt, kis területre vonatkozó, gyakran átfedő adatok;
- heterogén adatbázisok;
- a tulajdonosi jogok nagyon különbözőek lehetnek;
- a centralizáció politikai és gazdasági nehézségekbe ütközik.



1. ábra A mezőgazdasági döntés informatikai támogatása [4]

A fejlődő grid technológia kézenfekvő megoldás lehet ezekre a problémákra, hiszen természetesen módon szolgáltat

- egységes adathozzáférést az alkalmazások és felhasználók számára;
- nagy háttértároló-kapacitást;
- számítási kapacitást az optimalizáló feladatok számára (neurális hálózatok, genetikus algoritmusok);
- jogosultságkezelést;
- számlázó rendszert az adatszolgáltatók és felhasználók részére;
- műszerek, mérőeszközök megosztási lehetőségét.

Képzelnünk el egy rendszert, ahol a farmer vagy egy, a mezőgazdasági területre telepített, wireless eszköz digitális felvételeket készít egy búzatábláról. Ezek a fotók elemzésre kerülnek egy mintaelmézésen alapuló rendszerben, összehasonlítva több millió, adatbázisban tárolt képpel. A GPS koordináták alapján a meghatározhatóak a földrajzi adatok (tengerszint feletti magasság, lejtésszög), a talaj minőségi adatai, az adott terület meteorológiai adatai (elmúlt hetek időjárása, és ami várható), a gazdasági növény növekedési modelljének paraméterei, a betegségekre, kórokozókra vonatkozó információk. És mindezek alapján a rendszer javaslatot ad a gazdának az öntözésre, permetezésre, betakarításra és egyéb fontos operatív feladatokra vonatkozólag.

A vidékfejlesztés informatikai problémái

Az Európai Unió egyik stratégiai célja, hogy „2010-re Európa legyen a legversenyképesebb és legdinamikusabban tudás alapú gazdaság a világon, amely képes a fenntartható fejlődésre, több és jobb munkahely létesítésével és nagyobb

társadalmi összefogás segítségével”. Ennek kapcsán számos konzorcium pályázati célja egy olyan technológiai platform kialakítása, amely

- nyílt (bárki feltölthet az információkat),
- magas fokú interakciót valósít meg (pl. intelligens hozzáférést biztosít az egyént érintő információkhoz),
- adatintegrációs és adatbányászati eszköztárral rendelkezik,
- rejtett adatfeldolgozással biztosítja az információ elérését (intézményfüggetlen jogosultságkezelés).

A vidékfejlesztés egyik kulcskérdése az ilyen tudáscentrumok kiépítésének lehetősége, de a hagyományos kliens-szerver technológián alapuló rendszerek még elméletben sem képesek ilyen feladatok megvalósítására igen nagy méretekben. Ráadásul a centralizált adattárolás, adatkezelés felveti az információk elfogult kezelésének lehetőségét, ezért mindenképpen kerülendő.

A grid technológia ebben az esetben is megoldást nyújthat a problémákra, a következő tulajdonságaival:

- centralizáció nélküli adat- és alkalmazáskezelés,
- egyszerű fenntartás,
- adatok és alkalmazások gyors frissítése,
- új adatok gyors kezelése,
- dinamikus és flexibilis adatintegráció,
- lehetőség mesterséges intelligencia alapú technológiák alkalmazására.

Mindemellett a vidékfejlesztésben még nem jellemzőek a grid technológián alapuló projektek. Ennek az lehet az oka, hogy ebben az esetben a hangsúly a griden futó speciális adatbányászati és jogosultságkezelő alkalmazásokon lenne, amelyek komoly kihívást jelentenek a mesterséges intelligencia kutatóinak.

Mezőgazdasági grid projektek és elosztott rendszerek

AgModel

Az ázsiai, elsősorban Japán és Korea területén működő, *AgModel* projekt hordozható mezőgazdasági modellek és fejlesztések felgyorsításához nyújt eszközt. A hordozhatóság ebben az esetben azt jelenti, hogy a modellek könnyen futtathatóak bármely országban, kapcsolódva az ottani adatbázisokhoz. A megoldás lényege az intelligens központok kialakítása, amelyek konzisztensen kezelik le a különböző adatbázisokat. A mezőgazdasági alkalmazások ezekhez a központokhoz férnek hozzá elosztott megközelítéssel. A legtöbb szoftver nyílt forráskódú. A rendszer több részből áll, ezek a következők:

MetBroker: Az időjárási felvételek, adatok fontosak sok mezőgazdasági modell számára, de ezek az információk csak több adatbázisból gyűjthetők össze maradéktalanul, ráadásul állandó frissítésre van szükség. A *MetBroker* feladata, hogy állandó kapcsolatot tartson fenn az Internet-alapú időjárás-adatbázisokkal 8 országban, és biztosítsa a folyamatos frissítést.

DemBroker: Konzisztens hozzáférést biztosít a digitális tengerszint feletti magassági adatokhoz, amelyek szintén különböző formátumban, különböző

adatbázisokban vannak tárolva. Globális felbontása 1 km, míg Japán területén ugyanez 50 m.

ResourceServer: Az alkalmazások többnyelvűségét biztosítja.

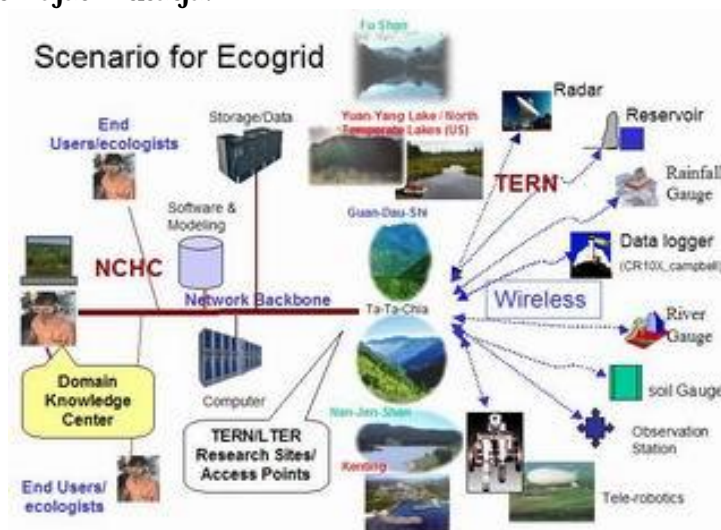
ChizuBroker: Online, raszteres térképadatbázisokhoz nyújt hozzáférési lehetőséget, amelyek az infrastruktúráról, térbeli referenciákról nyújtanak információkat a mezőgazdasági döntésekhez.

Agricultural Computing Grid (ACG)

Elsősorban a mezőgazdasági együttműködések támogatására fejlesztették az USA-ban. Használatával növelni kívánják az agráriumban dolgozók hatékonyságát, új lehetőségeket szeretnének teremteni az agrár közösségnek együttműködésével.

EcoGrid

Három intézet együttműködéseként jött létre: a „National Center for High-Performance Computing” (NCHC), a „Taiwan Ecological Research Network” (TERN) és a „Taiwan Forestry Research Institute” (TFRI). A projekt célja, hogy grid-alapú informatikai rendszert hozzon létre a hosszú távú ökológiai kutatás segítésére, és hogy integrálja a szétszórt erőforrásokat (számítógépek, vezetékes és vezeték nélküli hálózatok, adatbázisok, megfigyelő kamerák és szenzorok). A munka nemzetközi méretűre bővült azzal, hogy a megfigyelő rendszert kiterjesztették a Wisconsin állambeli Trout Lake területére. A **2. ábra** az ökológiai grid vázlatos sémáját mutatja.

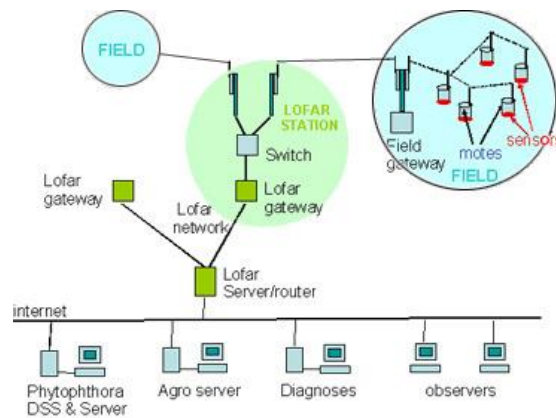


2. ábra Az EcoGrid munkafolyamata [15]

LOFAR

A Holland kormány alapította projekt (Low Frequency Array telescope – LOFAR) egy nagy tudományos műszernek tekinthető, amely több tízezer antennát tartalmaz egy nagy területű szenzor-hálózatban, ami egy klaszter erőforrásait használja a számításaihoz. Teleszkópként is használható, de egyik legfontosabb

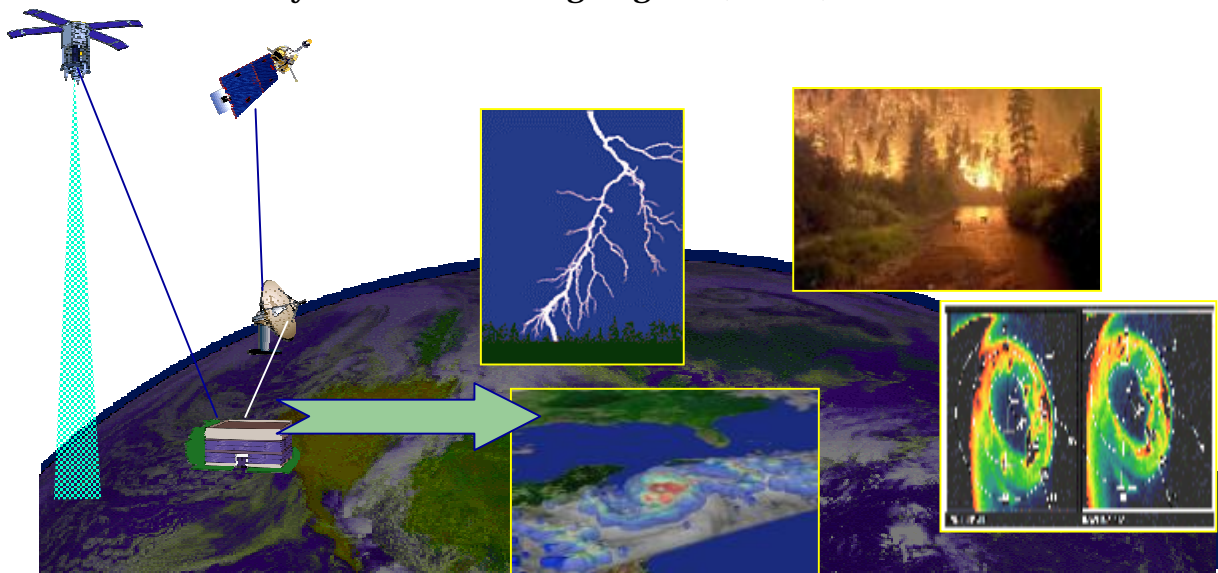
alkalmazási területe mégis a precíziós mezőgazdaság. Az egyik kutatott terület a Phytophthora gombás növényi betegség elleni védekezés, ahol hasznosnak bizonyult a pontosan beállított mikroklíma. A szenzorhálózat detektálja a relatív páratartalmat, megvilágítást, a légköri nyomást és a talaj kémiai összetételét, amelyek a legfontosabb faktorok a mikroklímában és a mért adatokat tárolja, majd azok alapján komplex döntéseket hoz a mikroklíma egyensúlyban tartására (3. ábra).



3. ábra LOFAR Phytophthora projekt [13]

ADaM

Az ADaM egy alkalmazás, amely már számos grid projektben bizonyított (NASA IPG, TeraGrid), de klaszterekben vagy akár asztali gépeken is futtatható. Segítségével lehetőség nyílik automatikus mintafelismerésre, anomáliák felfedezésére on-line vizsgálati adatokban. A rendszert különböző területeken, így a mezőgazdaságban is használják, és a döntéstámogatás számára szállít ismereteket adatbányászati eszközök segítségével (4. ábra).



4. ábra ADaM [10]

Grid-Enabled Teleimmersive Spatial Decision Support System (TIDSS)

A projekt a térinformatikai rendszerekben felhalmozódott adatokat nyeri ki speciálisan az agrárökonómia területéhez tartozó vízminőség-ellenőrzési feladatok megoldásához. Ez igen nagy számítási igénnyel jár, ezért van szükség elosztott programozásra. A TIDSS tulajdonképpen egy virtuális 4-dimenziós rendszer, ahol a döntéshozó egy virtuális valóságban tud mozogni térben és időben, vizsgálva a vízminőségi adatokat.

Karshaka információs rendszer

Egy integrált agrár-információs rendszer, melyet elosztott programozással kívánnak létrehozni. A rendszer feladata lenne az információ-megosztás mellett, a farmerek és az egyetemek közötti tudás alapú kapcsolat erősítése

Milyen grid kell egy agrár-szakembernek?

A kérdést első közelítésben könnyű megválaszolni, és a válasz nem kötődik az agrár-szektorhoz: olyat, amit használni tud. Ilyen rendszer természetesen nagyon sok van, de ha a feltétel az, hogy: olyan rendszer, amit felhasználóbarát módon, hosszútávon, frusztrációk nélkül, hatékonyan tud használni, akkor a kép már sokkal árnyaltabb.

Rengeteg pénzt elköltenek elosztott programozásra, lassan bekerül a köztudatba a grid fogalma, kontinenseken átívelő együttműködések szerveződnek, és mégsem látni, a boldog, gridet használó felhasználók tömegeit. Pedig szükség van a gridre, de lehet, hogy a jelenleg működő rendszerek használhatóságával van baj?

Nagyon kevés az elfogadott szabvány ezen a területen, sőt még magát a grid fogalmát is többféleképpen definiálják. Kis és nagy cégek használják ezt a szót, „hozzábiggyesztve” szoftverük nevéhez, marketing eszközként, és kutatási intézmények indítanak kliens-szerver architektúra épülő projekteket grid néven, hogy a bűvös szó hatására megnyíljanak a pénzforrások. Számos igazi grid projekt ezzel szemben abbamarad az érdeklődés hiánya miatt (pl. Original Grid Forum Applications csoport).

Az alkalmazásokkal talán az a legnagyobb probléma, hogy nem elegendő egyszerűen csak a kódot párhuzamos feldolgozású klaszterekre implementálni, ennél több problémával találjuk szembe magunkat. De valószínűleg az idő majd megoldja ezeket a gondokat. A probléma gyökerei mélyebbre nyúlnak. Igen nehéz az alap-infrastruktúrát kialakítani, és használatba venni a gridet. A telepítés sokszor nagyon „izzadságszagú”, kevés dinamikus információt tudunk kinyerni a rendszerből, nem kapunk jóslatot arra, hogy mikor fog befejeződni a job-om futása, stb.

A griddel szemben megfogalmazott fontosabb követelmények: ne legyen centralizált, stabilitás, megbízhatóság, hordozhatóság, felhasználóbarátság. Legyen egyszerű a használata, de elég robusztus ahhoz, hogy „bolond-biztos” legyen, és adjon lehetőséget a professzionális használatra is. Az agrár- és vidékfejlesztési alkalmazásokban kiemelt fontosságú a kifinomult jogosultságkezelés, és ezen a területen elsősorban az adat-gridek kialakítása a fontosabb.

Jelenleg nincsen olyan grid rendszer, amely ezen feltételek mindegyikének megfelelne, mivel a gridnek még csak az alapvető funkcionalitása van meg. De mindaddig, amíg ezek a feltételek nem teljesülnek, az üzleti szférában csupán a grid nevét fogják használni, de a valódi grid rendszerek nem lesznek használatban.

Egy Magyar AgriGrid projekt indulása

A Debreceni Egyetem Agrárgazdasági és Vidékfejlesztési Intézetének Gazdasági- és Agrárinformatikai Tanszékén egy Fedora C2-es klaszter került installálásra, mely a Torque PBS alapú klaszter-menedzsment szoftvert használja. Amint a magyar indexszerver elkészül, a „certificate” birtokában csatlakozni szeretnénk a NorduGrid rendszerhez a NorduGrid ARC middleware segítségével (www.nordugrid.org). Azért esett a választásunk a NorduGridre, mert egy aktív, dinamikusan fejlődő felhasználói körrel rendelkezik, leginkább ők teljesítik a fentebb vázolt feltételeket, közvetlen kapcsolatunk van a fejlesztői team-mel, és mert mindenképpen egy nemzetközi projekthez szertettünk volna kapcsolódni.

Jelenleg a programozási környezetek tesztje folyik (MPI vs. Java alkalmazásszerverek). Egyszerűbb optimalizációs feladatok megoldására implementált algoritmusok teljesítményének, hatékonyságának, alkalmazhatóságának tesztelése folyik C++ és Java nyelvi környezetben. Az előzetes eredmények szerint a legcélravezetőbb a Java-ban írt alap kód, ami a számításigényesebb feladatokban C és MPI unitokat hív meg, és egy JBoss vagy J2EE alkalmazásszerveren fut. Szempontjaink között kiemelt fontosságú a kód portabilitása, ami a mobil eszközökön való használat szempontjából jelentős. További terveink, hogy az AgModel nyomán és szabad forráskódú megoldásaikat használva, egy olyan integrált rendszert alakítsunk ki, ami valóban hatékony segítséget jelenthet az előzőekben vázolt mezőgazdasági és vidékfejlesztési problémák megoldásában.

Hivatkozások

1. Buyya, R. ed. (1999). *High Performance Cluster Computing: Architectures and Systems*. Volume 1, Prentice Hall, Old Tappan.
2. Buyya, R. (April 2002). *Economic-based distributed resource management and scheduling for Grid computing*. PhD Thesis, Monash University, Melbourne, Australia.
3. Laurenson, M. R., Yamakawa, A., Meng, H., Kiura, T., Wang, J. and Ninomiya, S. (August 2004). *Integration of Data Broker Web Services for Agricultural Grid*. Proceedings AFITA/WCCA2004, ISBN 974-229-639-1.
4. Ninomiya, S., Laurenson, M. (March 2003). *A Grid for Efficient Decision Support in Agriculture*. International Symposium Grid Computing, Taipei, Taiwan.
5. Kónya, B. (Nov 2004). *Advanced Resource Connector (ARC) – The Grid Middleware of the NorduGrid*. Lecture Notes in Computer Science, Volume 3241, Page 10.

6. Kacsuk, P., Kónya, B., Stefán, P. (Nov 2004). *Production Grid Systems and Their Programming*. Lecture Notes in Computer Science, Volume 3241, Page 13.
7. Burger, T. W. *Trends in Distributed Computing*, <http://www.intel.com/cd/ids/developer/asmo-na/eng/95223.htm>
8. Binstock, A. *Multiprocessors, Clusters, Grids, and Parallel Computing: What's the Difference?* <http://www.intel.com/cd/ids/developer/asmo-na/eng/95581.htm>
9. (Oct 2004). *One Grid to Rule Them All*, Economist Magazine.
10. Graves, S., Redman, S. (January 2005). Focus Study: Mining on the Grid with ADaM. Grid Application Planning & Implementation, Atlanta, USA.
11. Wang, S., Bennett, D. A., Armstrong, M. P., Rajagopal, R. and Brands E. (June 2002). Using Grid-Enabled Teleimmersive Spatial Decision Support Systems (TIDSS) to Visualize Uncertainty for Water Quality Protection in Agroecosystems. Proceedings of International Conference of Geoinformatics'2002, Nanjing, P.R. China.
12. <http://www.nchc.org.tw>
13. <http://www.lofar.org/>
14. <http://www.nordugrid.org>
15. <http://ecogrid.nchc.org.tw/>