
A ClusterGrid bróker rendszere

Stefán Péter
Szalai Ferenc
Vitéz Gábor



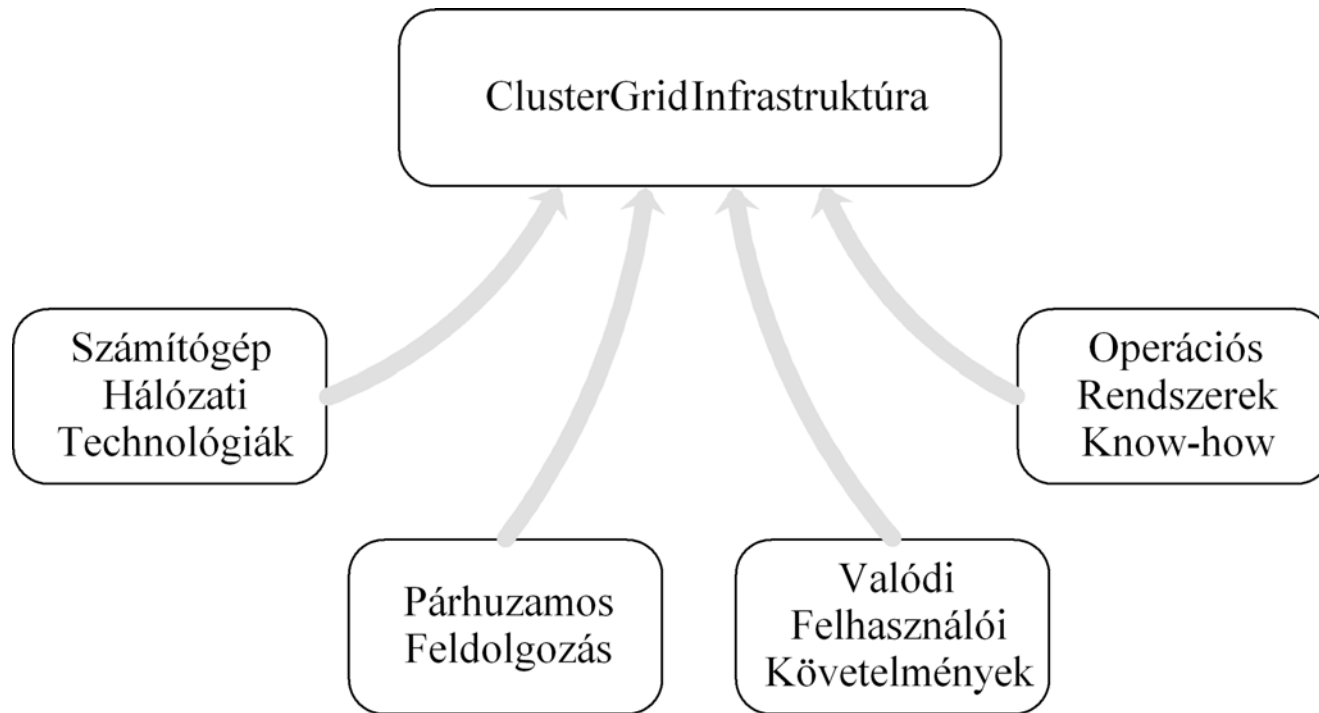
Tervezett témák

- A gridről általában, a ClusterGrid rendszerről röviden.
 - A grid réteg feladatai.
 - Erőforrás bróker rendszer. Miért van szükség rá?
 - A ClusterGrid monitoring rendszere. Miért van szükség rá?
 - 2004. novemberi statisztikák, felhasználók.
 - 2005. évi tervezett fejlesztések.
 - Összefoglalás.
-

A grid fogalma

- A grid fogalma sokféleképp definiálható:
 - Nyílt szolgáltatások halmaza,
 - Együttműködést támogató eszköz (tág határok),
 - Erőforrások (eszközök, számítási-, tárolókapacitás) megosztása.
- A grid kutatás és szolgáltatás szintű fejlesztés kiemelt szerepe:
 - Hatást gyakorol a kutatói (akadémia + cégek K+F tevékenysége) versenyképességre,
 - Ez közvetve az ország versenyképességére.

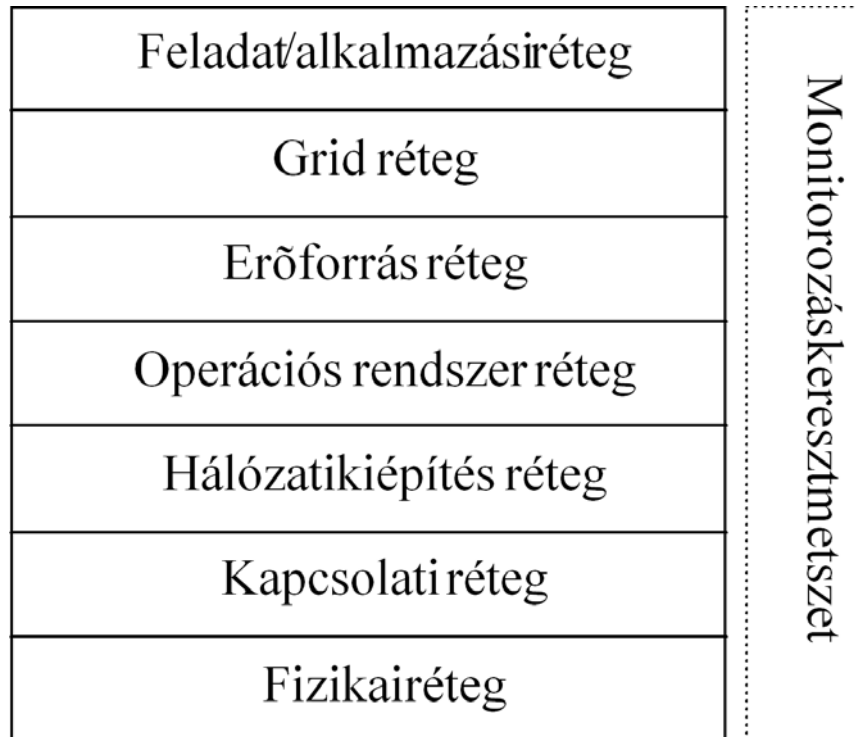
A grid fogalma



A ClusterGrid-ről röviden

- 2002-ben indult számítási infrastruktúra fejlesztési project.
 - Célja egy országos számítási erőforrás hálózat kialakítása, túlnyomó többségben desktop PC laborokból. (De!) Szuperszámítógépek, fürtök bekapcsolása is napirenden van.
 - A desktop laborok sajátosságai:
 - Nappal: office környezetben használják (Windows),
 - Éjjel: cluster környezetben működnek (Linux).
 - ClusterGrid architektúra: egy lehetséges megoldást nyújt a gépek összekapcsolására.
-

A ClusterGrid-ről röviden



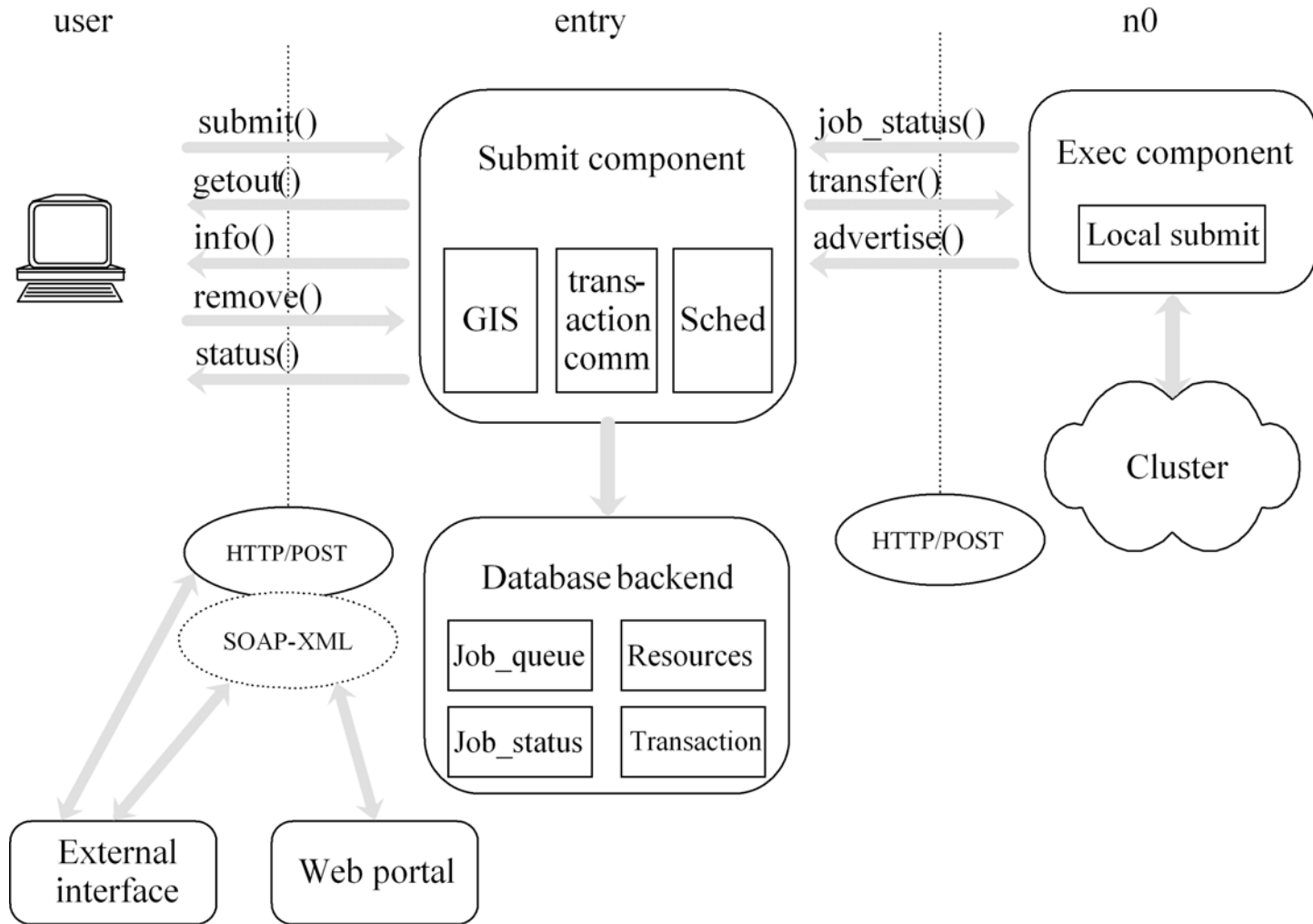
A grid réteg feladatai

- A grid réteg absztrakciós szint az erőforrások fölött.
 - Főbb feladata: Erőforrások szolgáltatás szintű összekapcsolása (több szolgáltatásra kiterjed, pl. file átvitel, autentikáció, információs kapcsolat, stb.)
 - Web fölött futó rendszerek (WSRF) vs. klasszikus szolgáltatás alapú rendszerek (Globus TK. 3.0).
 - Klasszikus grid funkciók: GRAM (feladat futtató rendszer), MDS (információs rendszer), GridFTP (adatátviteli rendszer).
 - Hogy valósulnak meg ezek a ClusterGrid bróker rendszerében?
-

Az erőforrás bróker

- HTTPS fölött működik, PHP nyelven íródott alkalmazás. (Egyportos alkalmazás).
 - Megvalósítja a klasszikus grid rendszerek fontosabb funkcióit, és néhány innovatív elemet is tartalmaz:
 - Runtime környezet (jobdir),
 - Dinamikusan leképzett környezetben (jogokkal futó) felhasználói feladatok,
 - Elosztott belépési pont környezet.
 - Jelenlegi verzió: 0.3.x.
-

Az erőforrás bróker



Az erőforrás bróker – GIS

- A submit komponens és az exec komponens is tartalmazza a részeit.
 - Megfelel a klasszikus MDS-nek.
 - A grid információs rendszer (GIS) alapvetően egy kétszintű hierarchikus rendszer, melyben az erőforrások státuszriportot küldenek a belépési pontoknak.
 - Nyelvfüggetlen hirdetési mechanizmus.
 - Default: 5-percenként frissített információ.
 - Információit a helyi erőforrás menedzsertől, illetve speciális külső mechanizmusból kapja (Fontos! Lehetőség nyílik absztrakt erőforrások definíciójára).
-

Az erőforrás bróker – GIS

```
> clgr_info
```

ResourceID	FreeNode	NodeNum	FreeDisk
https://n0.elte-testlab.grid/	0	2	5991
https://n0.mie-lab1.grid/	12	16	3036
https://n0.sztaki-test.grid/	0	0	994

Az erőforrás bróker – TSS

- A tranzakciós alrendszer minden komponensben megtalálható alapelem.
- Nagyjából a GRAM-nak és GridFTP feladatait végzi.
- Tehermentesíti a bróker rendszer különböző részeit a kommunikációs tevékenységtől.
- CURL alapokon nyugszik, a feladatot egy önálló osztály végzi.
- Aszinkron és szinkron kommunikációt tesz lehetővé.
- Az üzeneteket fix formátumú tranzakciókká alakítja, majd végrehajtja azokat.
- A tranzakció atomi művelet (nyugta v. hiba a végén).

Az erőforrás bróker – TSS

- Az általános tranzakció formátum:

```
CLGR_TRANSACTION#<timestamp>#<transaction_id>#<resource_id>#<type>#<grid_id>#<argument>
```

- A <type> mező értéke az alábbi lehet: submit, remove, result, info.

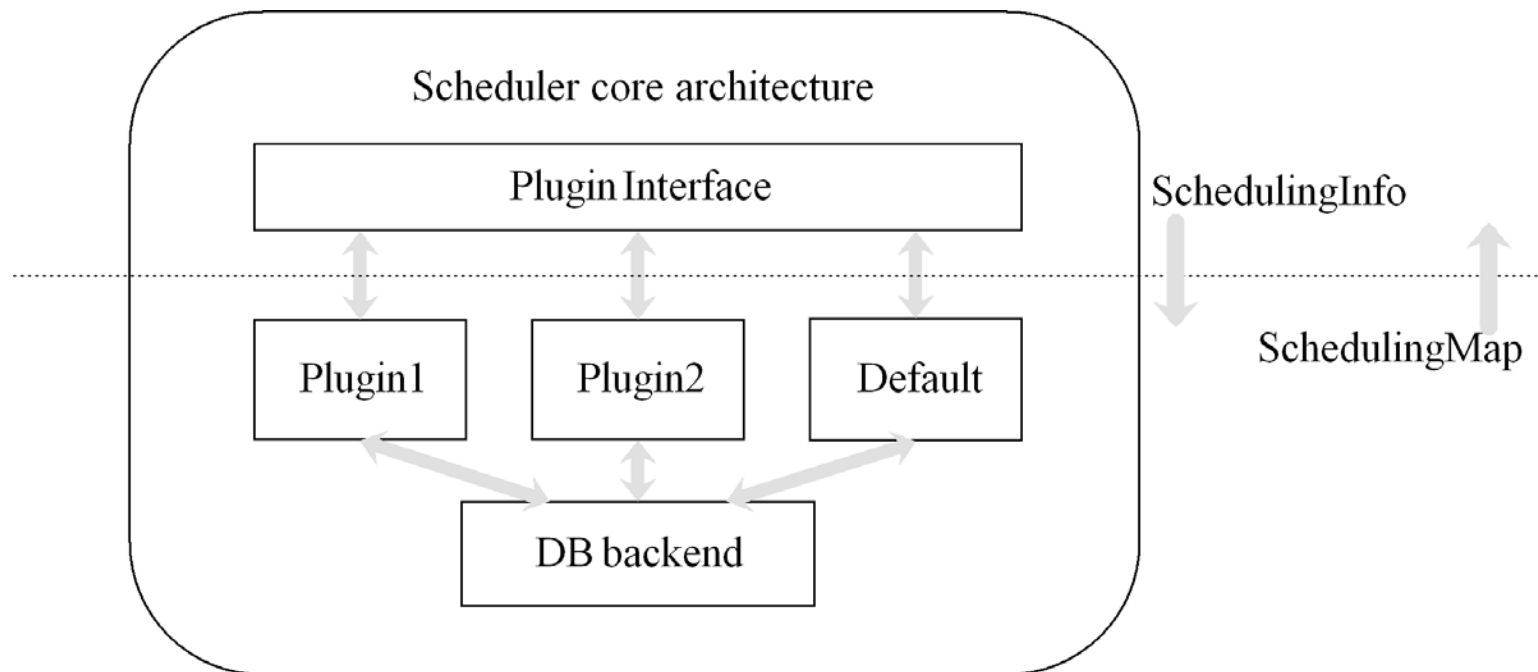
- A hibaüzenet formátuma:

```
CLGR_ERROR#<timestamp>#<transaction_id>#<error_code>#<error_message>
```

Az erőforrás bróker – Scheduler

- A submit komponens része.
 - Feladata:
 - Erőforrások és feladatok összerendelése.
 - Az összerendelt feladatok kiosztása (dispatching).
 - Elosztott ütemezés fontossága (több ütemező dolgozhat ugyanazon az adatbázison).
 - A dispatching folyamatba be lett építve egy elfogadás-visszautasítás mechanizmus.
 - Két komponens: core és nyelvfüggetlen modulok.
 - Szabványos felületei vannak (XML).
-

Az erőforrás bróker – Scheduler



Az erőforrás bróker – LS

- Az exec komponens része.
 - SGE és Condor a jelenleg támogatott platformok. (Hamarosan fork() környezet is lesz.)
 - Igénybe vesz két fontos, a brókertől kvázi független alkalmazás-szintű szolgáltatást:
 - LibNSS
 - IDRegister.
-

Az erőforrás bróker – DB Backend

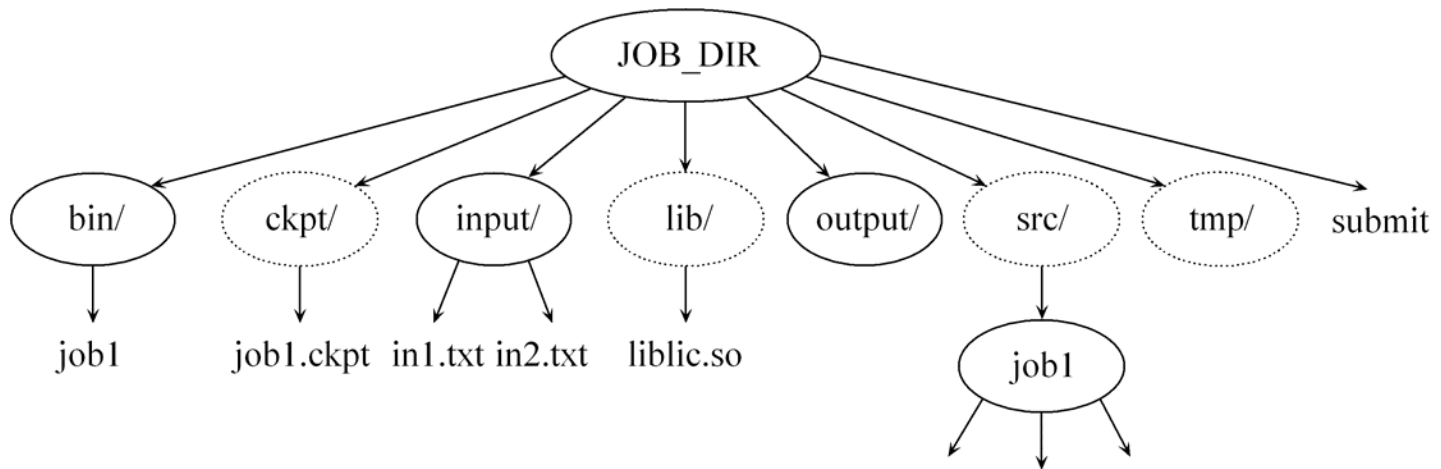
- A submit komponens által gyűjtött és feldolgozott adatokat tárolja.
 - Jelenleg centralizált, de könnyen HA cluster-esíthető (MySQL NDB), illetve elosztható (izolált belépési pontok).
 - A táblák normalizáltak – a tranzakciók atomi jellege, illetve a táblák szerkezete elvileg anomália-mentesítik az adatbázist.
 - Az egyes táblákat a különböző komponensek használják.
-

Az erőforrás bróker – UI

- A felhasználó által elérhető szolgáltatásokat fogja össze.
 - Ehhez a kvázi-szabványos (GGF, W3 ajánlások) felülethez kapcsolódnak a kliens programok (CGI parancsok, portálok).
 - A felület a jelenlegi POST mechanizmuson túl ellátható SOAP interfésszel, mely külső kapcsolati felületként is működik.
 - Standard X509 autentikációs lehetőséggel is rendelkezik.
-

Az erőforrás bróker – UI

- Mi történik egy feladat submitálásakor?



Monitorozás

- A grid implementáció teljes vertikumát érinti (alacsony szintektől a magas szintekig).
 - Mit monitorozhatunk?
 - Alkalmazás monitorozás.
 - Szolgáltatás monitorozás.
 - A régi rendszer: patch-elt mon.
-

Monitorozás

- Régi rendszer problémái:
 - ❑ A mon lassú, sok gépidőt fogyaszt, lehetetlen nagy átfogó, valós idejű lekérdezéseket csinálni.
 - ❑ A lekérdezése nehézkes, nehezen integrálható más rendszerekbe.
 - ❑ Adatgyűjtés és monitorozás nincs integrálva, nincs standard interface az adatgyűjtéshez.
 - ❑ Emiatt a grid weathermap erősen „hack-elt”.
-

Monitorozás

- Új rendszer jellemzői:
 - Hierarchikus felépítésű rendszer.
 - Egy belépési pontról (entry/service) minden lekérdezhető.
 - Jóval gyorsabb és egyszerűbb (fontos cél!!!), mint más rendszerek.
 - Több monitorozó gépet hierarchikusan össze lehet fűzni.
-

Monitorozás

- Két komponensből áll:
 - Mérést végző rész,
 - Kiértékelő rész.
 - A mérést végző rész szolgáltatja a mérés leírását és a mérés eredményét.
 - A kiértékelő komponens megvizsgálja, hogy a mért érték megfelel-e előre rögzített kívánalmainknak és ennek megfelelően egy logikai értéket ad eredményül.
-

Monitorozás

- Az attribútumok hálózaton keresztül a megfelelő kliens (CLI + web) segítségével lekérdezhetők.
- A monitorokhoz attribútumok vannak rendelve, melyek alapján keresni lehet:

```
host=n0, domain=testlab.grid, service=ssh
```

- Erre illeszkedő logikai részekkel lehet keresni például az összes helyi kiszolgálón.
 - A keresési tér tetszőlegesen, akár egyetlen hostra is beszűkíthető.
-

Monitorozás

- Integrált időjárás-térkép.
 - Integrált RRD kezelés.
 - Egyszerű érték-grafikon ábrák.
 - Több érték-grafikon ábrák Igazi Szivárvány színskálával.
 - Több érték összegzett diagrammok.
 - Egy host összes típusú grafikonok.
-

Továbbfejlesztési elképzelések

- Minél több operációs rendszeren működőképessé tenni.
 - További szolgáltatások bevonása.
 - Cizelláltabb ütemezők írása.
 - XML/SOAP interfész elkészítése.
 - Különböző grid rendszerek összekapcsolása (Nordugrid ARC).
-

Továbbfejlesztési elképzelések

- Job fordítás támogatása.
 - GIS API alkalmazásokhoz, szolgáltatásokhoz.
 - Valószínűsíthető futási idő becslése, számlázási alrendszer.
 - Virtuális szervezetek kialakítása.
 - A rendszer biztonságának fokozása, a felhasználók rendszerbe vetett bizalmának további erősítése.
-

Köszönjük a figyelmet!

<http://www.clustergrid.iif.hu>

grid-tech@niif.hu
