

Open Source În Proiectul Gisheo



Universitatea de Vest din Timișoara

Despre Proiect

- Scopul proiectului:
 - Crearea unei platforme de eLearning pentru GIS folosind capabilitățile GRID
- Parteneri:
 - Universitatea de Vest din Timișoara
 - Agenția Spațială Română
 - Institutul National de Cercetari Aerospatiale „ELIE CARAFOLI”
 - Universitatea Tehnică Cluj-Napoca

Despre Proiect

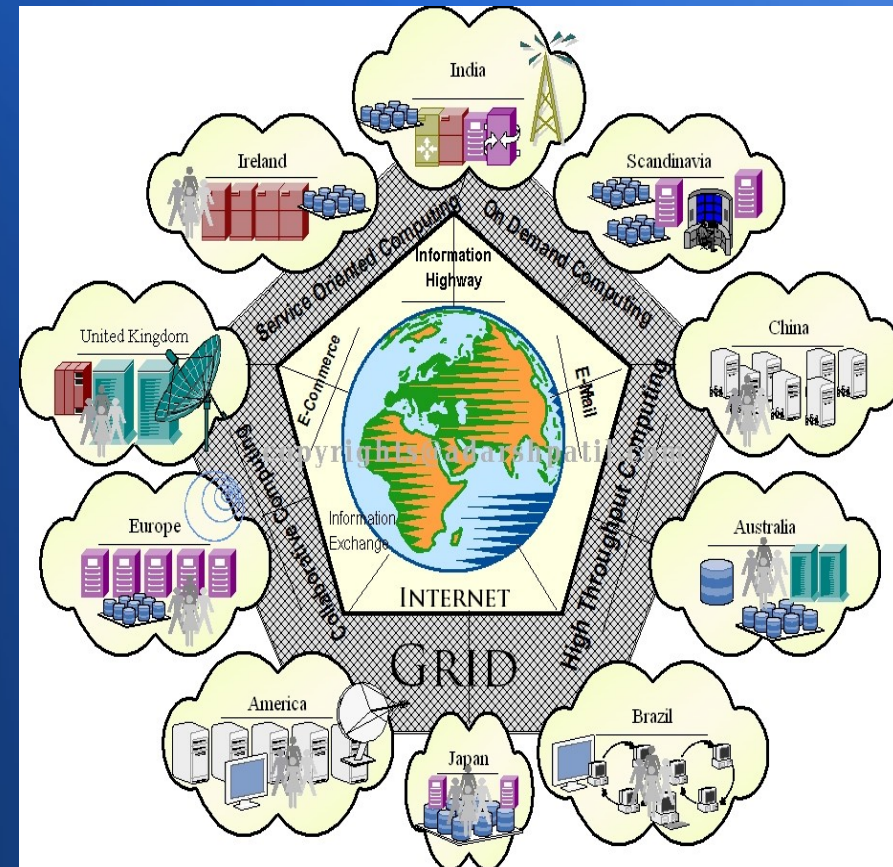
- Finanțare:
 - Proiectul este finanțat de Agenția Spațială Europeană prin intermediul programului PECS
 - Programul PECS este un program de suport al țărilor aderente la EU după 2004, cu scopul aderării la ESA

Despre Proiect

- Infrastructura:
 - 4 Clustere interconectate:
 - UVT
 - UTCN
 - INCAS
 - ROSA
 - Middleware: Condor:
 - 1 nod central
 - 3 grupuri de noduri de lucru: UVT, UTCN și INCAS (offline)
 - 2 noduri de stocare: UVT și INCAS

Despre GRID (în general)

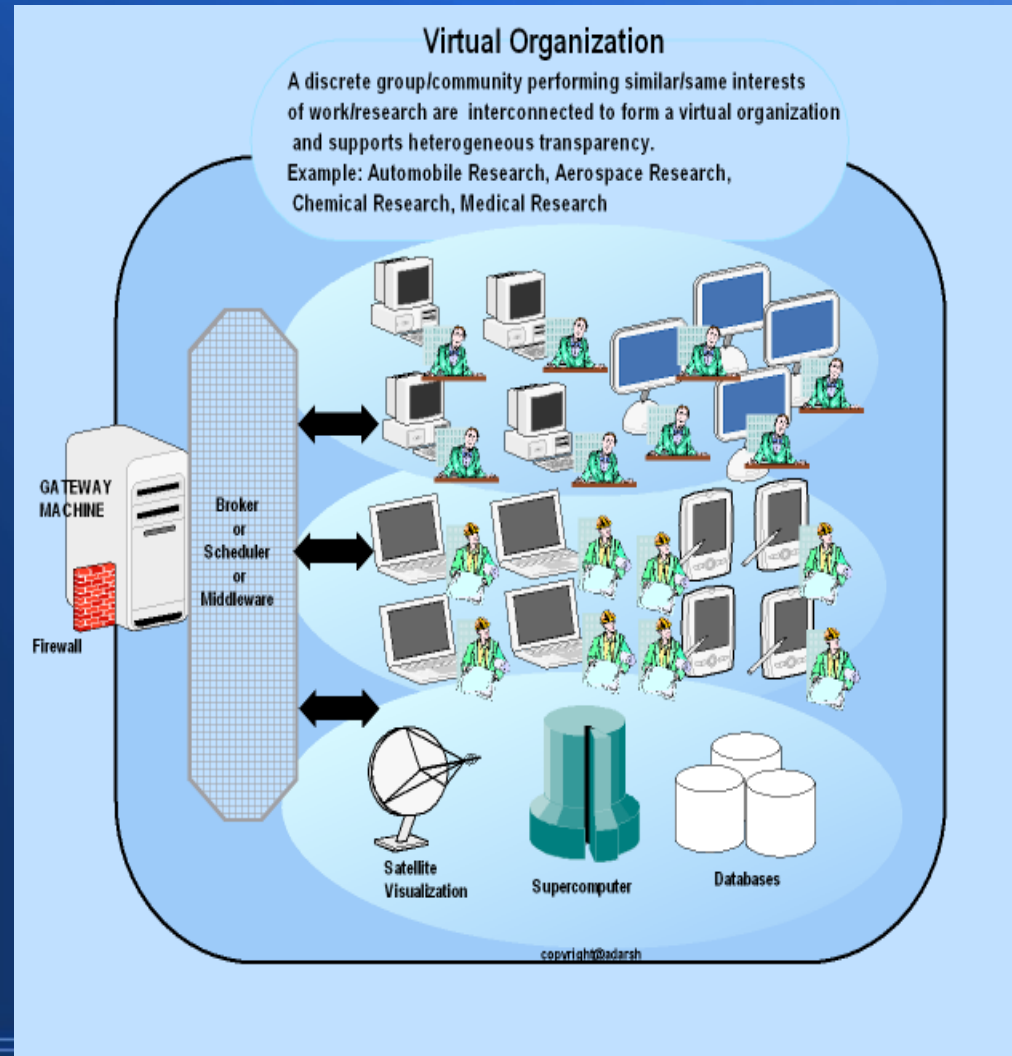
- Este o colecție eterogenă de resurse (calculatoare, stocare, dispozitive, servicii, etc) care sunt geografic distribuite, cu conținut dinamic
- GRID Computing
 - OGSA (<http://www.globus.org/ogsa>)



Despre GRID (în general)

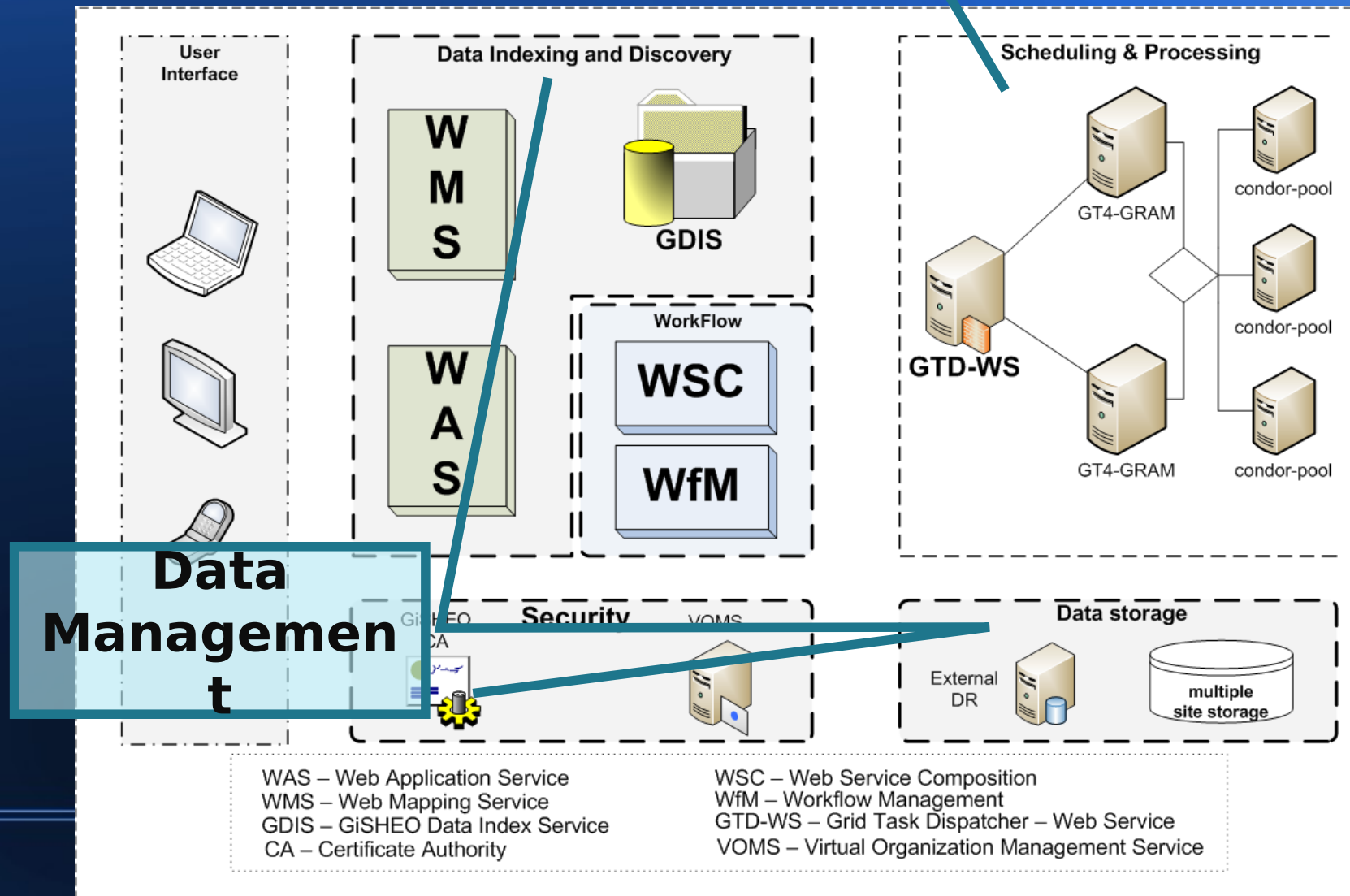
Organizație Virtuală

- Persoane care colaborează partajând date într-un mod organizat
- Grup de indivizi sau instituții care partajează resurse computaționale pentru un țel comun
- Infrastructură de control al accesului la resurse pe baza unor credențiale



Arhitectura Gisheo

Processing



Țeluri Inițiale

- Țeluri
 - Utilizarea „serviciilor” GRID pentru fundamentul proiectului.
 - Integrarea a diverse proiecte realizate/finanțate de ESA (BEAM, BEAT).
 - Folosirea exclusivă de proiecte „Open Source”.
 - Oferirea de servicii de procesare în timp real

Tehnologii Utilizate

- Execuție:
 - Condor HTC (<http://www.cs.wisc.edu/condor/>)
- Manipulare:
 - GDAL (<http://www.gdal.org/>)
 - GRASS (<http://grass.osgeo.org/>)
 - OpenCV (<http://opencv.willowgarage.com/wiki/>)
- Stocare:
 - Apache HDFS (<http://hadoop.apache.org/hdfs/>)

Tehnologii Utilizate

- Indexare:
 - PostgreSQL (<http://www.postgresql.org/>)
 - PostGIS (<http://postgis.refrations.net/>)
- Distribuție:
 - Apache Mina (<http://mina.apache.org/>)
 - Mina FTP Server
 - Mapnik (<http://mapnik.org/>)
 - TileCache (<http://tilecache.org/>)

Tehnologii Utilizate

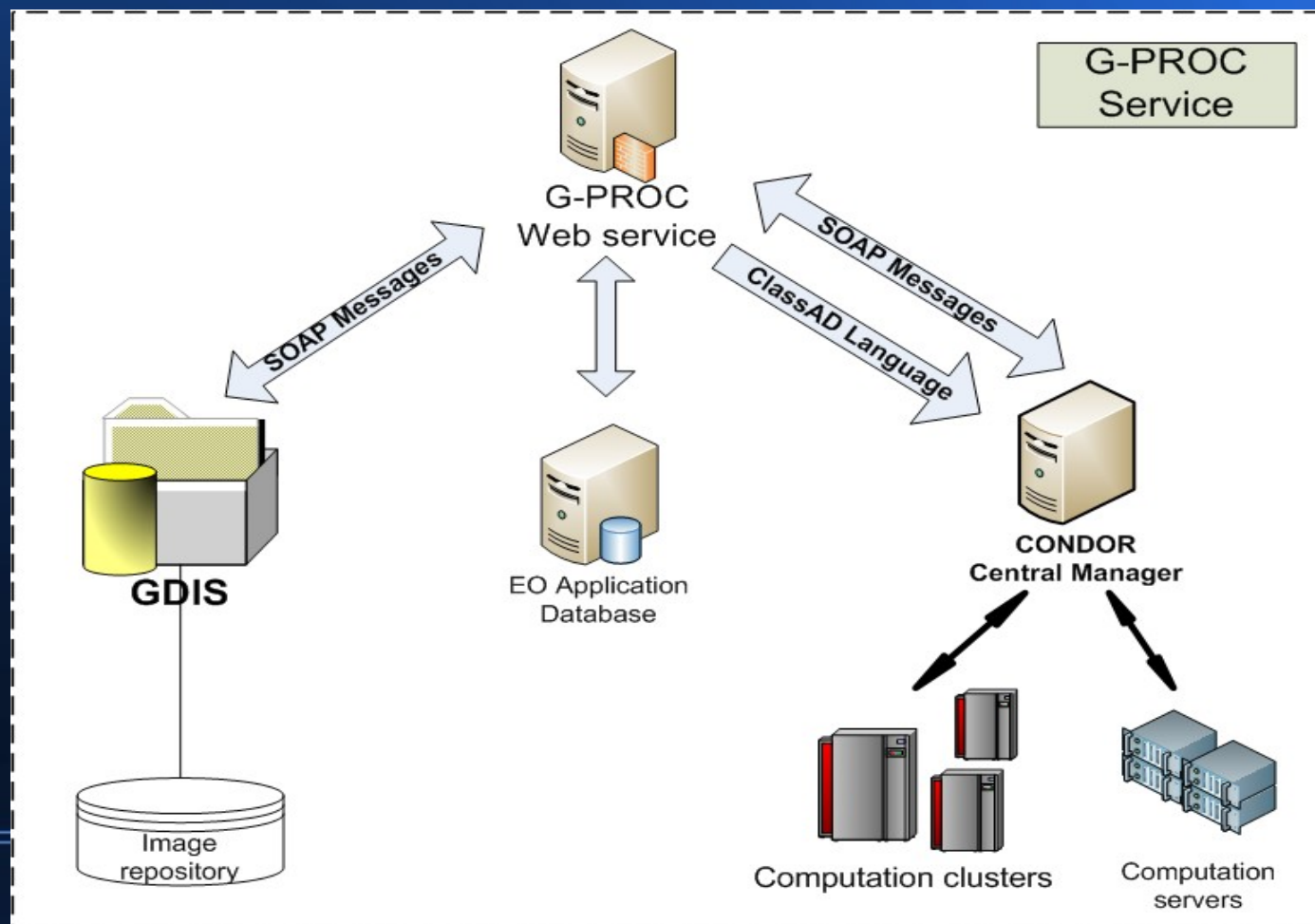
- Prezentare
 - OpenLayers (<http://openlayers.org/>)
 - Google Web Toolkit (<http://code.google.com/intl/ro/webtoolkit/>)

Execuție: Condor HTC / Middleware

- Condor HTC – platformă de calcul distribuit
 - High Throughput Computing
- Platformă modulară:
 - Condor-manager: nod de control
 - Condor-worker: nod specializat în execuția codului utilizatorilor/dezvoltatorilor
 - Condor-scheduler: componenta de planificare (de selecție a nodurilor pentru execuția unui task)

Procesare și Manipulare

- Procesare - Arhitectura Generală:



Procesare – Componente Principale

- Serviciul G-PROC:
 - Descrie sarcinile date de utilizator în „limbajul” middleware-ului utilizat
 - Planifica sarcinile pentru execuție
 - Se integrează cu planificatorul dezvoltat în proiect (OsyRIS)
 - Este accesibil ca serviciu web (SOAP)
- Workload and Resource Management:
 - condor-scheduler

Manipulare

- Majoritatea operațiilor sunt susținute direct de biblioteca GDAL.
 - Este folosit pentru a oferi access la diverse „containere” (Ex. HDF, GeoTiff, etc) într-un mod uniform.
 - Ne permite dezvoltarea de procesări în mod generic, fără a fi nevoiți să tratăm special fiecare tip de date (cel puțin teoretic :)

Manipulare

- O parte din procesări sunt „wrapere” pentru funcționalități oferite de GRASS.
 - Exemplu: r.mapcalc, etc
- OpenCV:
 - Bibliotecă pentru „Computer Vision” care s-a dovedit extrem de utilă în proiect. Mai ales datorită algoritmilor implementați de aceasta

Indexare (GDIS)

- Realizată cu ajutorul PostgreSQL și a extensiei PostGIS.
 - Indexează informațiile stocate în nodurile de stocare pe baza caracteristicilor acestora: extent, „meta date”, etc
 - Expune informațiile prin intermediul serviciilor SOAP și parțial WMS
 - Permite regăsirea eficientă a datelor

Stocare

- Datele indexate de GDIS sunt stocate folosind diverse medii:
 - Servere FTP
 - Hadoop Distributed File System:
 - Stocare distribuită (și redundantă)
 - Extrem de eficientă dacă este combinată cu un planificator conștient de amplasarea datelor.

Distribuție

- FTP:
 - Servere FTP clasice (și GridFTP)
 - Server FTP bazat pe HDFS
- WMS:
 - Bazat pe o combinație între Mapnik și TileCache
 - Combinat în general cu mai multe nivele de caching (Memcachedb și Memcache)

Prezentare

- Google Web Toolkit:
 - Platformă WEB 2.0 (buzz)
 - Folosită pentru a crea o interfață „interactivă” cu utilizatorul
- OpenLayers:
 - Client JavaScript pentru WMS/TMS

Sfârșit

Întrebări ?