

---

# GRID – Concepte de baza

---

Anca Hangan  
Universitatea Tehnica din Cluj-Napoca

---

# Cuprins

- Introducere
- Definitie
- Concepte de baza
- Arhitectura Grid
- Middleware pentru Grid

# Introducere

- **Stiinta si inginerie:**
  - Probleme complexe
  - **Volum** foarte mare de date
  - **Timp** de executie ridicat
  - **Colaborare** intre institutii diferite
- **Problemele tipice:**
  - Volumul (Petabytes) si locatia datelor
  - Volumul calculelor realizate (ani de timp CPU)
- **Exemple:**
  - Determinarea nivelului de ozon din atmosfera din imagini satelitare (100GB/zi date stocate la diferite institutii)
  - Simularea interactiunilor biochimice intre molecule
  - Probleme de modelare în domeniul ingineriei
  - Realizarea predictiilor meteorologice
  - etc.

---

# Introducere (cont.)

## ■ Supercalculatoare

- ❑ Avantaje: putere de calcul si capacitate de stocare a datelor, posibile optimizari
- ❑ Dezavantaje: costuri ridicate, dificultati de acces si utilizare in comun, dimensiuni mari

## ■ Grid – supercalculatorul virtual

- ❑ Avantaje: utilizarea in comun a resurselor, colaborare intre institutii, distributie geografica, descentralizare
- ❑ Dezavantaje: management, securitate

# Exemple de aplicatii

- **SETI@home** ([setiathome.berkeley.edu](http://setiathome.berkeley.edu)) – utilizarea puterii de calcul a calculatoarelor personale conectate la retea
- **Protein Data Bank** ([www.rcsb.org/pdb/](http://www.rcsb.org/pdb/)) - resurse si unelte pentru cautare si realizare de rapoarte, pentru studiul proteinelor.
- **myGrid** ([www.mygrid.info](http://www.mygrid.info)) – resurse si date în domeniul bio-informaticii
- **MCell** ([www.mcell.cnl.salk.edu](http://www.mcell.cnl.salk.edu)) – simulator de micro-fiziologie celulara
- **DAME** Distributed Aircraft Maintenance Environment ([www.cs.york.ac.uk/dame/](http://www.cs.york.ac.uk/dame/)) – sistem distribuit de diagnoza pentru motoarele vehiculelor aeriene

---

# Grid - definitie

- Infrastructura care integreaza resurse distribuite geografic (putere de calcul, retele de comunicatie, capacitate de stocare a datelor, informatie, etc.) cu scopul de a furniza o platforma virtuala, colaborativa, utilizata pentru calcul si managementul datelor

---

# Concepte de baza

- Partajarea resurselor
- Securitatea accesului
- Utilizarea eficienta a resurselor
- Transparenta topologiei
- Standarde deschise
- Organizatii virtuale

---

# Partajarea resurselor

- Acces la resurse diverse: putere de calcul, spatiu de stocare a datelor, date, software, calculatoare, senzori
- Detinute de mai multe persoane sau institutii
- Conditii impuse pentru utilizarea resurselor
- Problema : integrarea resurselor partajate intr-o platforma unitara



---

# Securitatea accesului

- Resurse “de incredere”
- Utilizatori “de incredere”
- Securitate: politici de acces, mecanisme de autentificare si autorizare
- Procesul de autentificare a utilizatorului are loc o singura data (single sign-on)
- Mecanisme de delegare a autoritatii

---

# Utilizarea eficienta a resurselor

- Timpul de asteptare pentru o resursa trebuie sa fie cat mai mic
- Mecanisme de alocare optima a resurselor
- Mecanisme de planificare a executiei proceselor lansate in Grid
- Mecanisme de rezervare a resurselor

---

# Transparența topologiei

- Grid – supercalculator virtual, resursa unică
- Utilizatorii nu îi cunosc topologia
- Mecanisme automate de distribuire a proceselor în Grid
- Asigurarea calitatii serviciilor de comunicație și a toleranței la defecte

---

# Standarde deschise

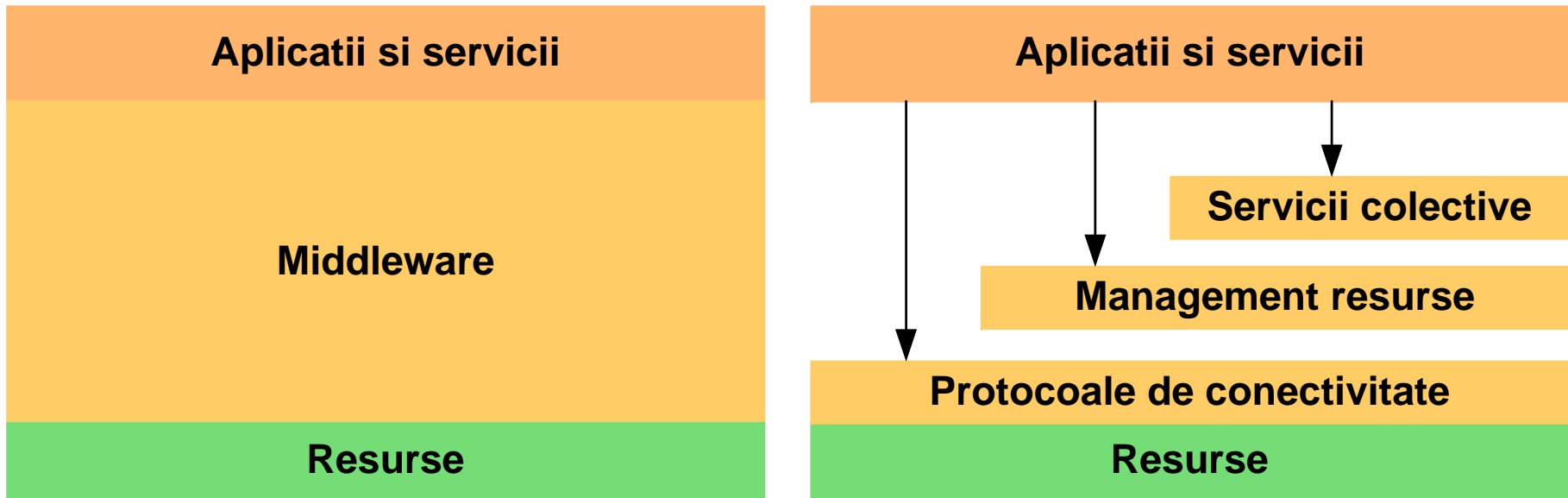
- O multime de proiecte Grid
- Numitor comun in privinta tehnologiilor de baza
- Standarde deschise
- Autoritate de standardizare: Open Grid Forum
- OGSA (Open Grid Services Architecture) considerat referinta pentru viitoarele proiecte de dezvoltare bazate pe Grid

# Organizatii virtuale

- Grid = colaborare intre comunitati
- Partajare a resurselor intr-un mediu colaborativ
- Multime formata din indivizi si/sau institutii pentru care s-a definit un set de reguli de partajare a resurselor
- Grupuri diferite:
  - Acelasi scop
  - Necesita mai multe resurse decat cele pe care le detin
  - Necesita un mediu care sa faciliteze colaborarea de la distanta

# Arhitectura Grid

- Arhitectura multinivel
- Mai multe nivele middleware



---

# Nivelul Middleware

- Mecanisme automate care organizeaza și integreaza resursele pentru a forma Grid-ul
- Nivelul protocoalelor de conectivitate
  - Protocoalele de comunicare specifice
  - Protocoale de autentificare specifice
- Nivelul de management al resurselor:
  - Protocoale de informare, utilizate pentru obtinerea de informații despre structura și starea resurselor
  - Protocoale de management, utilizate pentru negocierea accesului la resurse

# Nivelul Middleware (cont.)

- Nivelul serviciilor colective:
  - Servicii de monitorizare și diagnostic a resurselor
  - Servere de autorizare pentru impunerea politicilor de acces la resurse în cadrul unei comunități
  - Servicii de descoperire a resurselor
  - Servicii de replicare a datelor
  - Servicii pentru menținerea directoarelor de resurse
- Protocoalele și serviciile de la nivelul middleware sunt accesate de aplicațiile Grid



---

# Middleware pentru Grid

- Prima generatie:
  - Mai buna utilizare a puterii de calcul si a resurselor accesibile prin reseaua de comunicatie
  - Ex: Condor
- A doua generatie:
  - Resurse distribuite geografic
  - Domenii de administrare distincte
  - Organizatii virtuale
  - Securitate
  - Ex: Globus (GT2, GT3)
- A treia generatie:
  - Orientat inspre servicii
  - Ex: Globus (GT4), OGSA-DAI, gLite

---

# Intrebari?

---