

Architecture d'accès à des ressources de Calcul et de Données

*Stage réalisé à l'Observatoire Astronomique
de Strasbourg*

Du 1^{er} mars au 31 août

Tuteur : Schaaff André

Étudiant : Pestel Cyril

Plan

1. Présentation du laboratoire
2. Analyse de la situation
3. Solutions adoptées
4. Intégration des services
5. Bilan et évolutions possibles

Plan

1. Présentation du laboratoire
2. Analyse de la situation
3. Solutions adoptées
4. Intégration des services
5. Bilan et évolutions possibles

Présentation du laboratoire

- ❖ CDS (*Centre de Données astronomiques de Strasbourg*)
- ❖ Collecte et distribue des données astronomiques
- ❖ Principaux services : *Simbad, VizieR, Aladin*

Plan

1. Présentation du laboratoire
2. Analyse de la situation
 - Existant
 - Objectifs
 - Bénéfices attendus
3. Solutions adoptées
4. Intégration des services
5. Bilan et évolutions possibles

Analyse de la situation

❖ Existant

- En 2003
- Cluster de 5 machines
 - ◆ 3 Intel P4 2,4 Ghz
 - ◆ 2 bi-proc Intel Xeon 3 Ghz

Analyse de la situation

❖ Objectifs

- Refonte du cluster
- Amélioration de la gestion (ajout de nouveaux nœuds, de services supplémentaires, ...)
- Meilleure résistance aux pannes

Analyse de la situation

❖ Bénéfices attendus

- Amélioration des performances
- Meilleur maintien face à une charge élevée
- Simplification de l'administration

Plan

1. Présentation du laboratoire
2. Analyse de la situation
3. Solutions adoptées
 - ✦ Pour l'administration
 - ✦ Pour la Haute Disponibilité
4. Intégration des services
5. Bilan et évolutions possibles

Solutions adoptées

❖ Administration

⊕ DCC (Debian Cluster Components)

- ❖ Basé sur Debian
- ❖ Installe et configure un ensemble de machines
- ❖ Automatisation de procédure

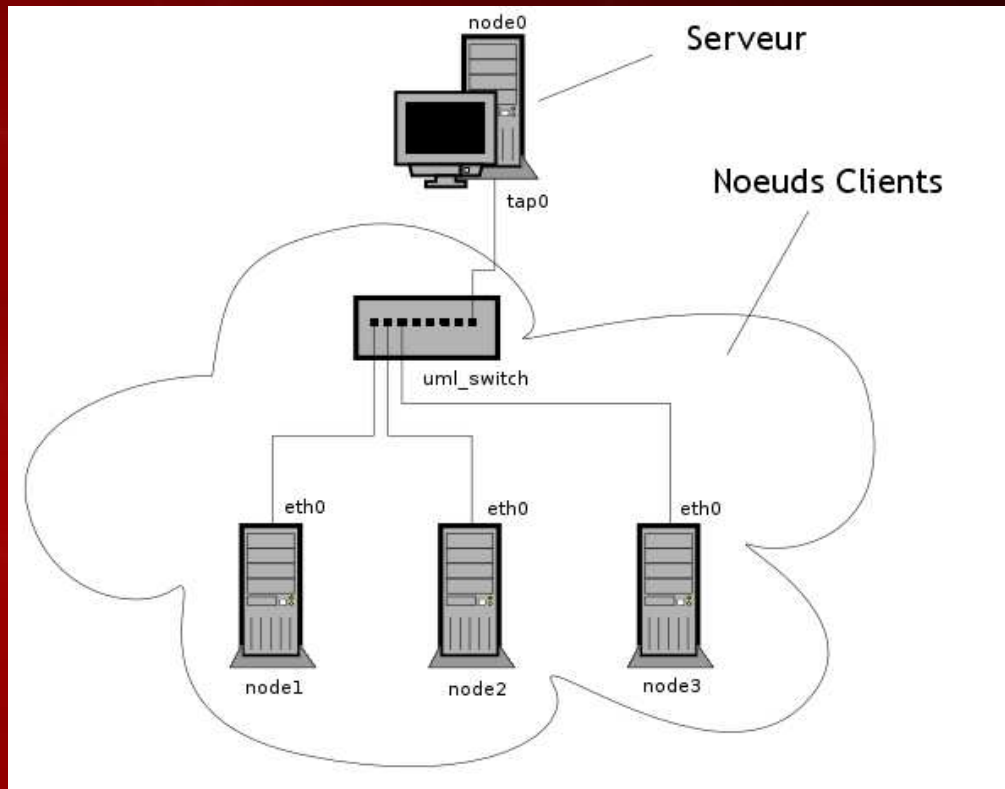
⊕ Personnalisation des paquets Debian

- ❖ Plus simple à utiliser
- ❖ Plus rapide à l'installation



Solutions adoptées

• Architecture Clients / Serveur



Architecture d'accès à des
ressources de Calcul et de
Données

Solutions adoptées

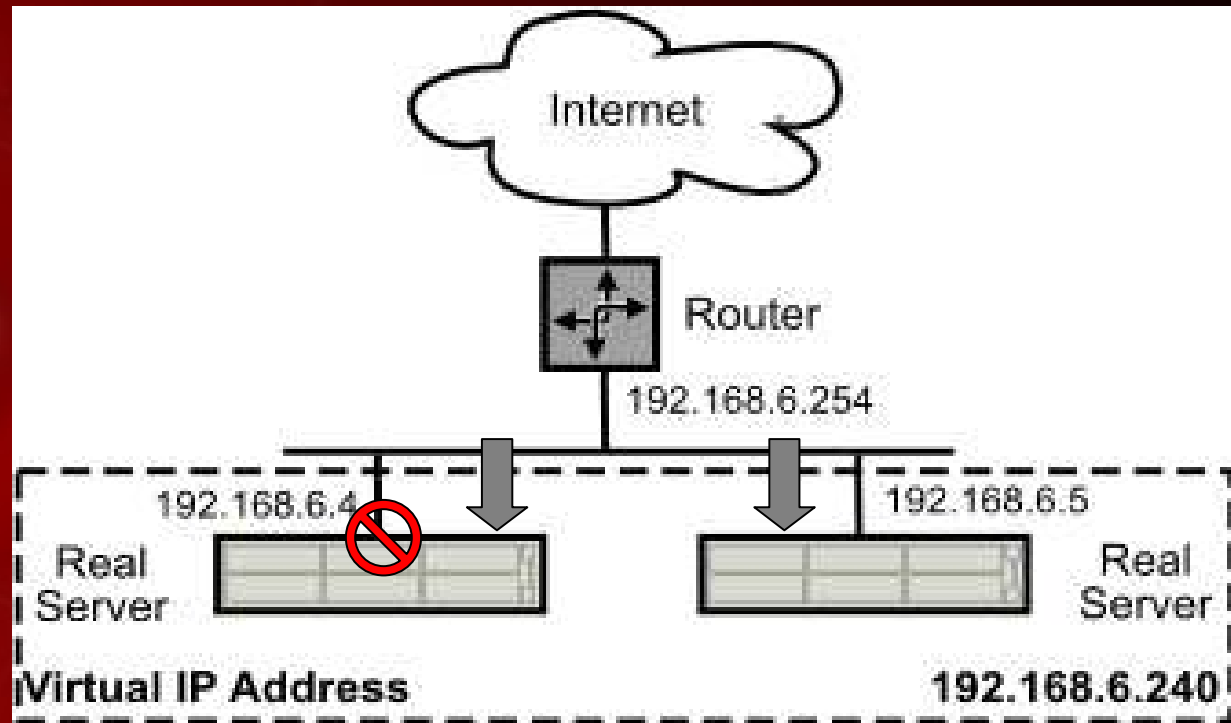
✚ Outils :

- ◆ **SIS** : Installation d'images systèmes sur une machine du réseau
- ◆ **C3** : Utilitaire d'aide à la maintenance
- ◆ **MPICH** et **LAM/MPI** : Librairies MPI
- ◆ **NSS/PAM/SLAPD** : Serveur LDAP avec des extensions
- ◆ **Ganglia** : Moniteur de surveillance des nœuds
- ◆ **Shorewall** : Pare-feu évolué

Solutions adoptées

❖ Haute Disponibilité

• Heartbeat



Architecture d'accès à des
ressources de Calcul et de
Données

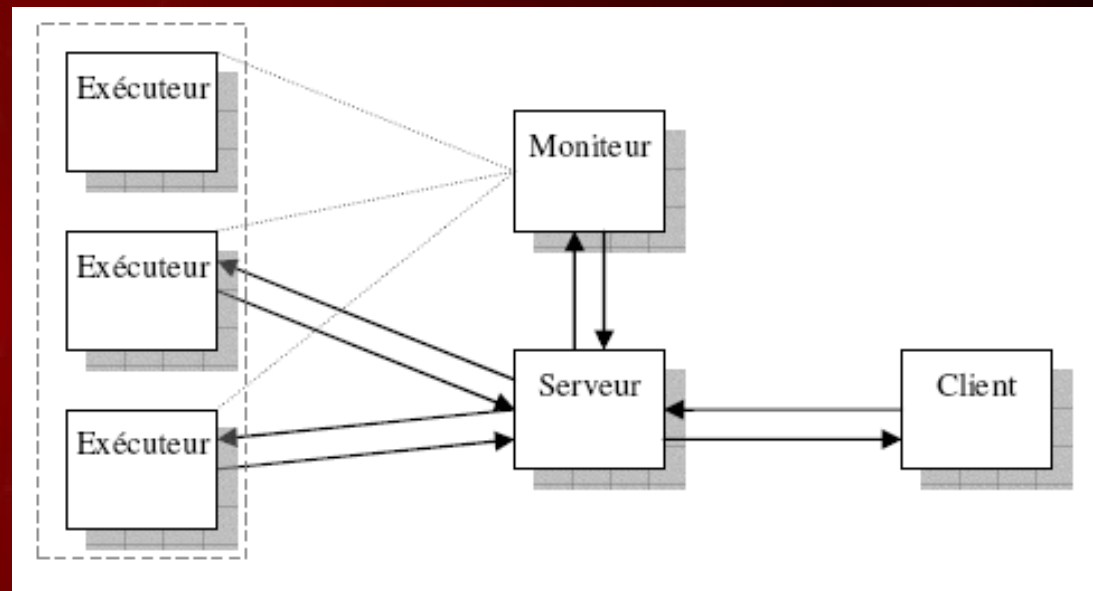
Plan

1. Présentation du laboratoire
2. Analyse de la situation
3. Solutions adoptées
4. Intégration des services
 - Ali
 - Aladin
 - AïDA
 - Workflow
5. Bilan et évolutions possibles

Intégration des services

❖ Ali : répartiteur de charge

- En place depuis 3 ans
- Réactif

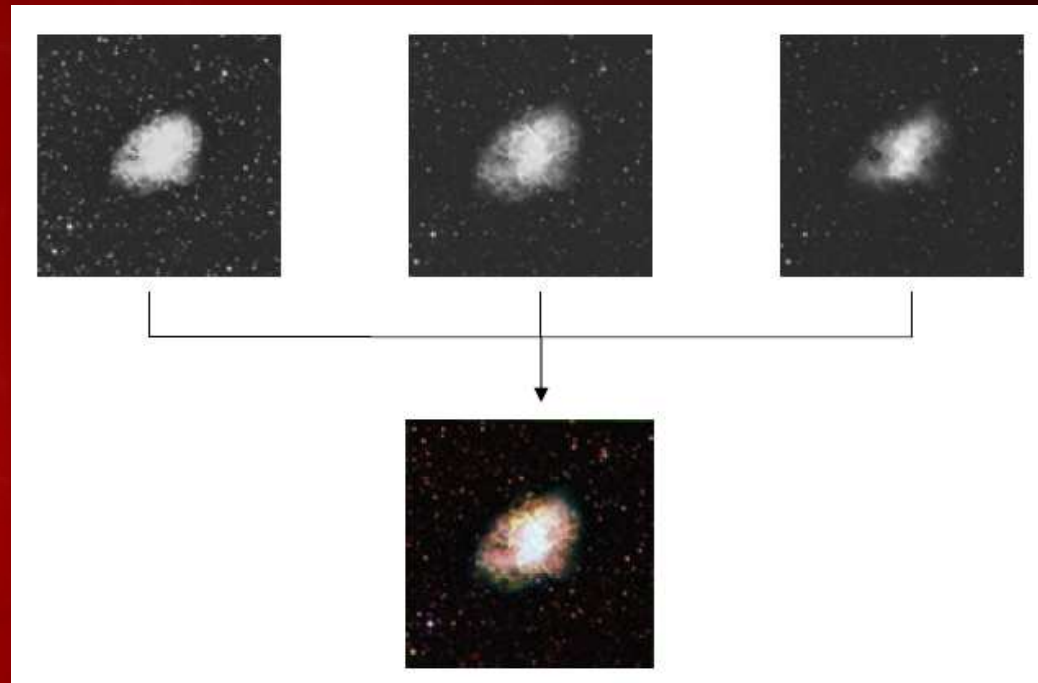


Architecture d'accès à des ressources de Calcul et de Données

Intégration des services

❖ Aladin

• Coloration RGB



Architecture d'accès à des
ressources de Calcul et de
Données

Intégration des services

Localisation de sources (objet astronomique)

```
# 1 NUMBER      Running object number
# 2 MAG_ISO     Isophotal magnitude [mag]
# 3 MAGERR_ISO  RMS error for isophotal magnitude [mag]
# 4 X_IMAGE     object position along x [pixel]
# 5 Y_IMAGE     object position along y [pixel]
# 6 ALPHA_J2000 Right ascension of barycenter (J2000) [deg]
# 7 DELTA_J2000 Declination of barycenter (J2000) [deg]
# 8 A_WORLD     Profile RMS along major axis (world units) [deg]
# 9 B_WORLD     Profile RMS along minor axis (world units) [deg]
# 10 THETA_WORLD Position angle (ccw/world-x) [deg]
# 11 FLAGS      Extraction flags
# 12 CLASS_STAR S/G classifier output
# 13           1 -7.9761 0.0055 256.173 345.802 13.4897848 +42.6645817 1.294875
# 14           2 -7.4532 0.0030 325.355 298.883 11.1841173 +41.4682230 0.3076304
# 15           3 -6.5831 0.0091 488.532 204.895 6.0498366 +38.9019189 0.3732468
# 16           4 -5.8055 0.0134 122.464 69.307 17.5974007 +35.6951906 0.2220612
# 17           5 -6.4558 0.0113 112.304 108.888 17.9680208 +36.6660735 0.5141431
# 18           6 -6.8623 0.0089 42.796 110.897 20.1177235 +36.6171153 0.4007519
# 19           7 -4.7349 0.0271 447.537 30.087 7.7202747 +34.6509259 0.1418599
# 20           8 -4.8464 0.0244 141.073 26.322 16.9774792 +34.6486083 0.1856625
# 21           9 -3.6233 0.0439 149.390 6.947 16.7058481 +34.1767346 0.1032758
# 22          10 -4.6290 0.0270 116.059 24.841 17.7291956 +34.5885929 0.1536183
# 23          11 -4.6401 0.0267 123.899 11.771 17.4758084 +34.2737689 0.121874
# 24          12 -4.3480 0.0300 358.904 31.671 10.3866784 +34.7882701 0.1371353
# 25          13 -3.9992 0.0372 228.023 5.477 14.3442881 +34.1783551 0.1473174
# 26          14 -4.5015 0.0294 264.743 23.606 13.2389150 +34.6290196 0.2331842
# 27          15 -4.9884 0.0231 248.059 16.326 13.7429189 +34.4490879 0.209314
# 28          16 -5.6699 0.0162 303.545 8.196 12.0739439 +34.2393553 0.2182117
```



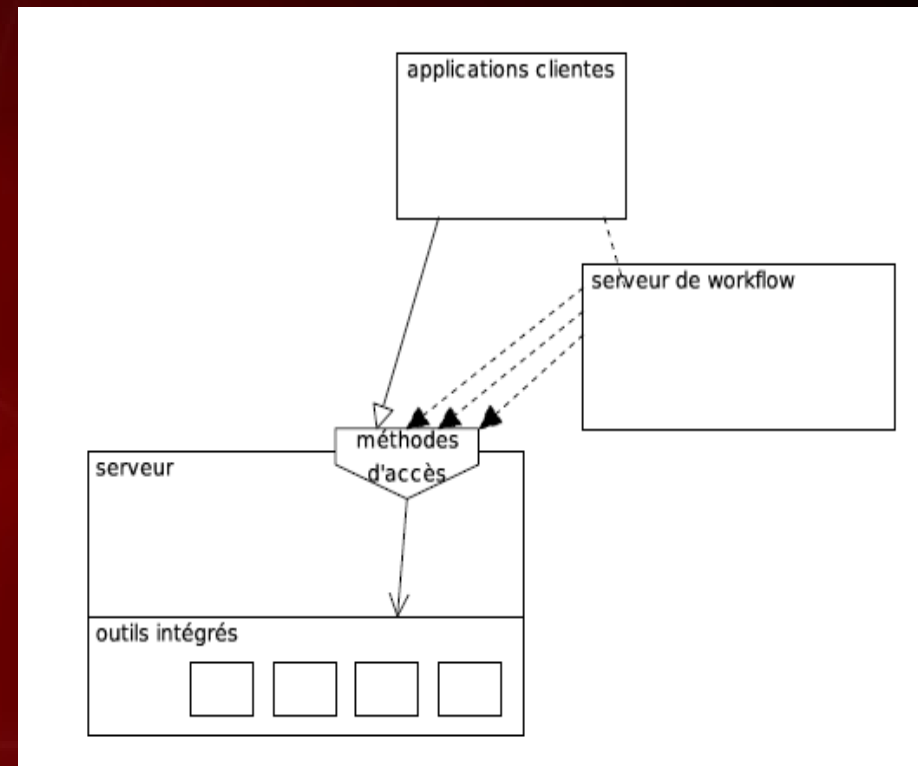
Architecture d'accès à des
ressources de Calcul et de
Données

Intégration des services

❖ AIDA (Astronomical Image processing Distribution Architecture)

- CGI

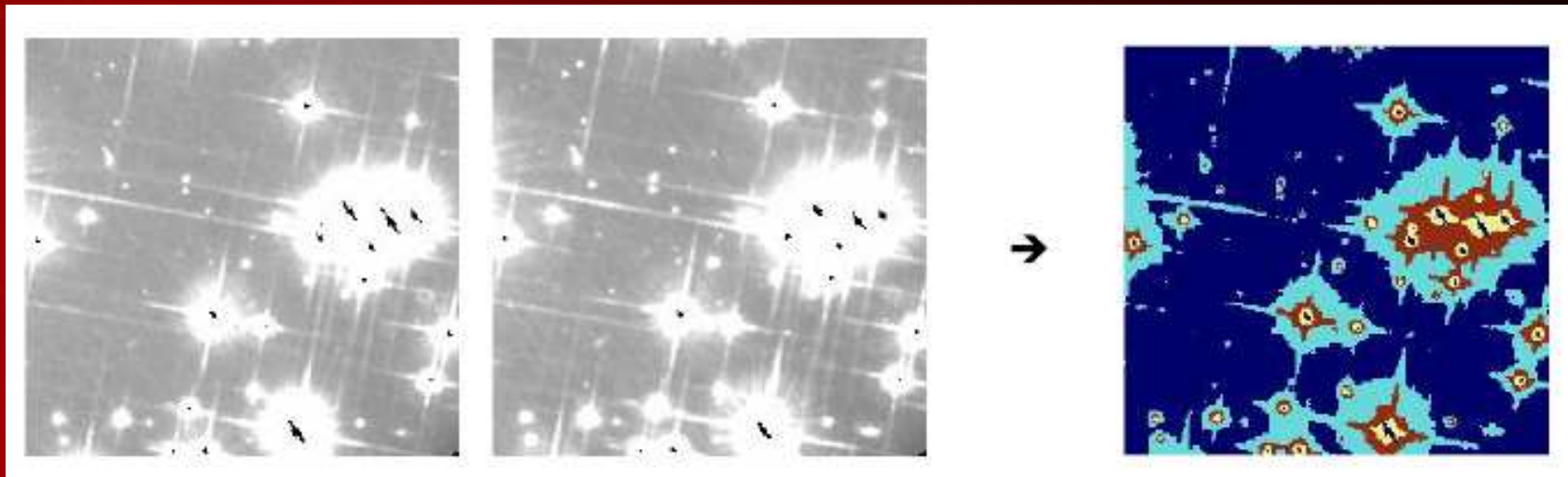
- Uniformise l'accès à des outils



Intégration des services

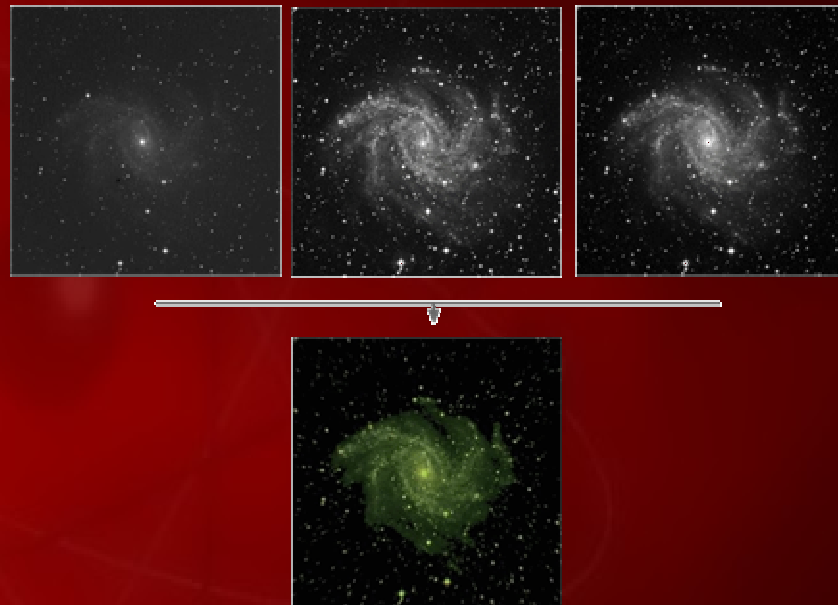
✚ 2 outils

◆ Carte de segmentation



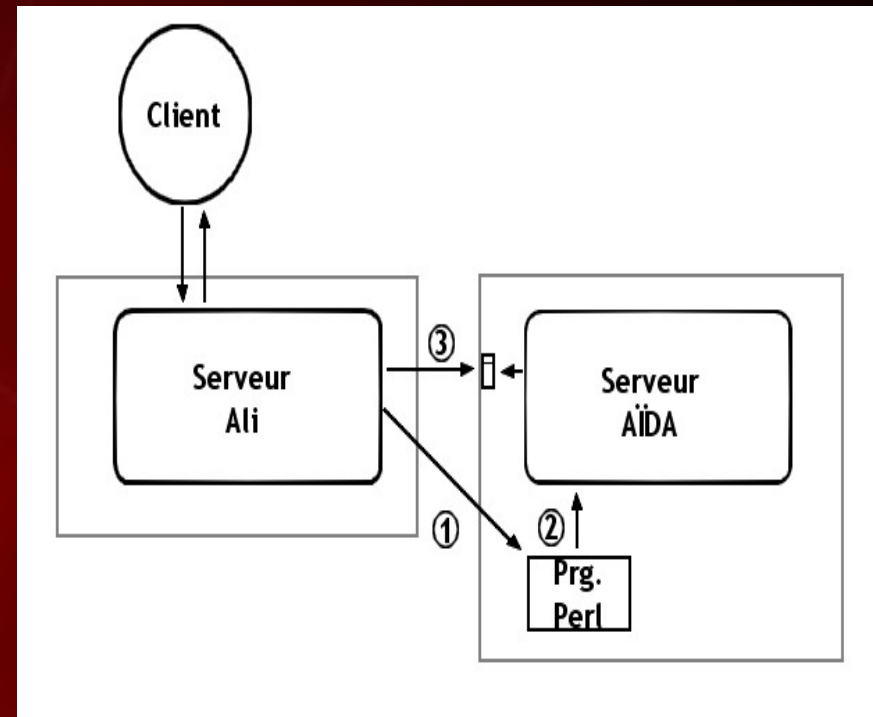
Intégration des services

◆ Visualisation TSL (Teinte Saturation Lumière)



Intégration des services

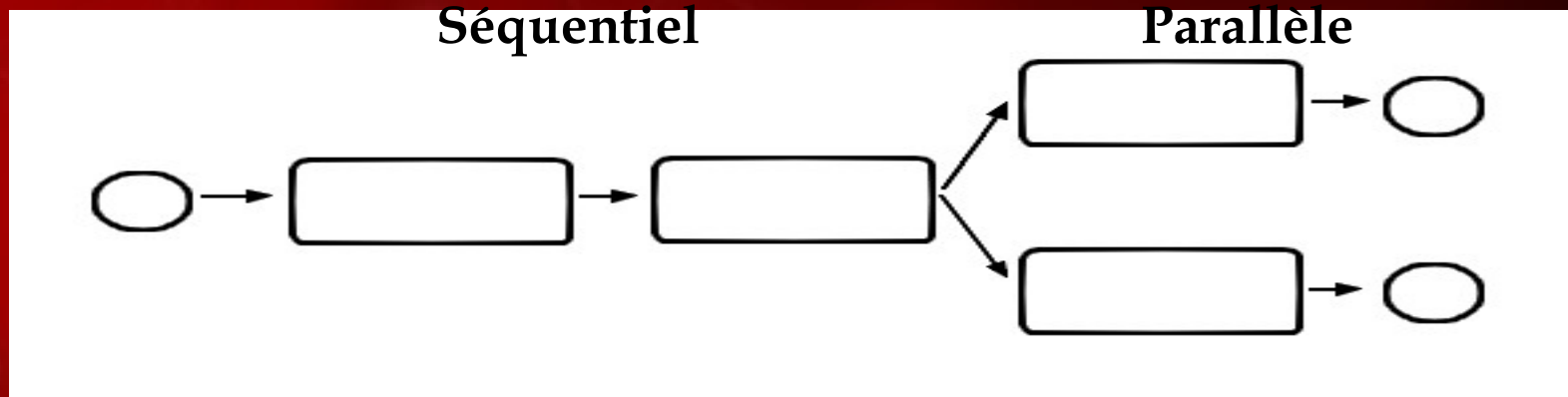
1. Ali exécute un script PERL qui émet la requête HTTP
2. AïDA lance le traitement
3. Ali récupère le fichier résultant



Intégration des services

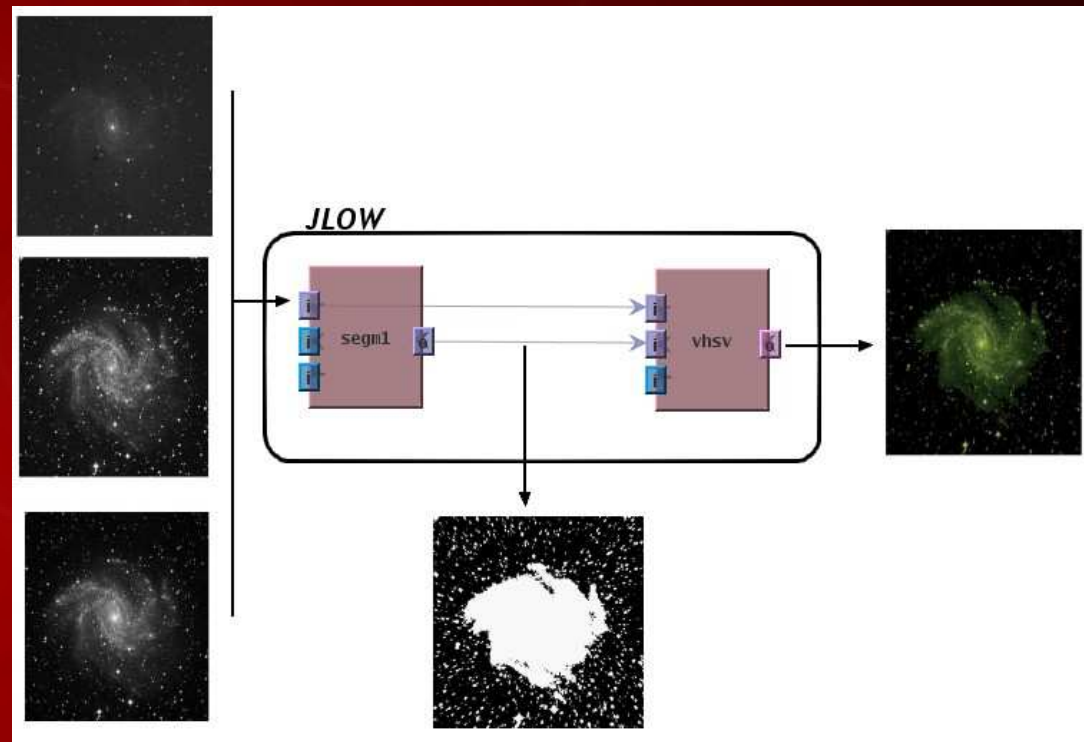
❖ *Workflow*

- Modélisation d'une chaîne de traitements
- Exécute plusieurs tâches séquentiellement ou en parallèle



Intégration des services

✦ *Jlow* (Java Library fOr Workflow)



Architecture d'accès à des
ressources de Calcul et de
Données

Plan

1. Présentation du laboratoire
2. Existant et besoins
3. Solutions utilisées
4. Intégration des services
5. Bilan et évolutions possibles

Bilan

❖ Cluster de 8 machines

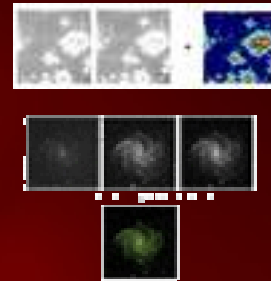
- Aisément administrable
- Disponible
- Meilleures Performances (diminution du temps de réponse de 30%)

Bilan

❖ De nouveaux services

✚ AIDA

- ❖ Segmentation
- ❖ Visualisation TSL



✚ Workflow

- ❖ Utilisation de *JLOW*
- ❖ Soumission de requêtes à *Ali*

Évolutions possibles

❖ Accès à une grille

- Compte existant chez *Grid5000*



❖ *VOspace* : espace de stockage de fichiers

Architecture d'accès à des ressources de Calcul et de Données

Tuteur : Schaaff André

Étudiant : Pestel Cyril