

Maintenance du Cluster CDS

Gilles Landais

27 aout 2009

Table des matières

1	Supervision	2
2	Architecture	3
3	Le serveur d'administration	4
4	Serveurs frontaux	5
5	Les serveurs d'exécutions	6
6	Taches d'administrations	7
7	Récapitulatif ALI	8

Le cluster de calcul du CDS est utilisé par les projets :

- **VizieR** : interrogation des grands catalogues (2MASS, SDSS, GSC...)
- **Aladin** : sextractor, AladinBatch, xmatch ...

Note : lors d'un reboot complet du cluster (i.e. : tous les noeuds). Il faudra **redémarrer en premier le serveur d'administration** (en raison de la présence d'un serveur dhcp).

1 Supervision

superviseur système de l'ensemble du cluster :

<http://wali1.u-strasbg.fr/ganglia/>

superviseur applicatif du cluster :

<http://wali1.u-strasbg.fr/perl/index.pl>

http://wali1.u-strasbg.fr/perl/workers2.pl?refresh=10&show_ip=unchecked&show_items=&mali_name=sali1&mali_port=6061

Log applicative sur wali1 /var/local/ali

2 Architecture

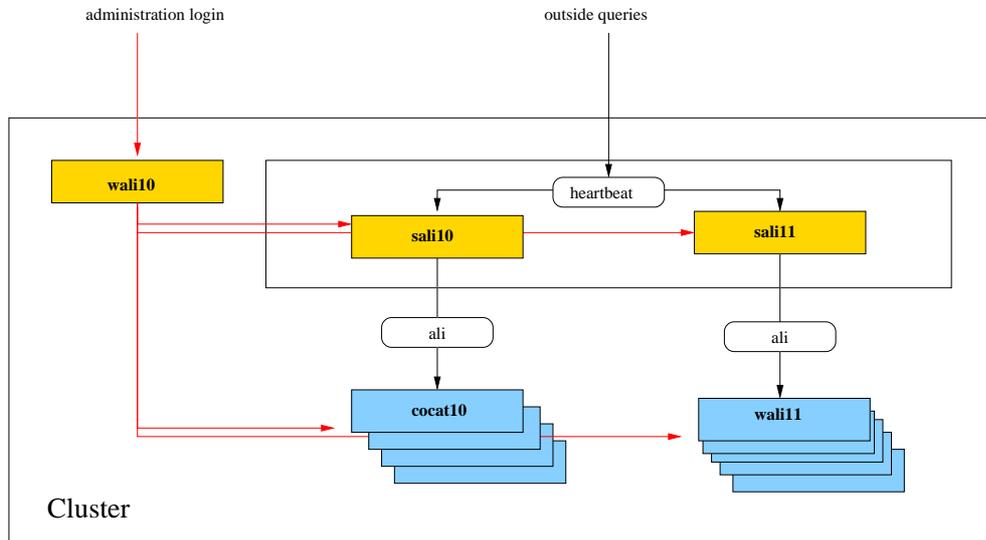


FIG. 1 – Architecture du cluster

Composition du cluster :

- 2 serveurs frontaux (sali10 et sali11) qui reçoivent et redirigent les requêtes extérieurs vers les serveurs du cluster qui les exécuteront.
- 9 serveurs d'exécutions (cocat1x pour VizieR et wali1X pour Aladin) qui exécutent les requêtes
- 1 serveur d'administration (wali1) qui gère les serveurs de l'ensemble du cluster. Il est aussi utilisé pour l'administration et les interventions inhérente au cluster.

Note : Seuls les serveurs frontaux et le serveur d'administration sont accessibles de l'extérieur.

Le cluster repose sur les éléments :

- **ALI** (développé par le CDS) gère le parallélisme, l'équilibre de charge et l'ordonnancement des tâches.
- **heartbeat** (de la suite Linux HA) qui assigne une unique adresse ip/dns aux 2 serveurs frontaux. Il opère la bascule d'un serveur à l'autre lorsqu'il détecte un dysfonctionnement.

3 Le serveur d'administration

Réseau :

```
dns  wali1
ip    130.79.129.202
```

services :

Nom	Démarrage/arret	Fichiers de config	
ali_logger	/etc/init.d/lali_service	/etc/ali/	service de log ALI
dhcp	/etc/init.d/dhcp3-server	/etc/dhcp3	serveur DHCP du reseau local
apache	/etc/init.d/apache	/etc/apache2	pour ganglia

Maintenance :

user	password	
alice	alice1	compte utilisateur services ALI
root	aqwzsl	

Tâches de maintenances :

- **les log ali** : sous /var/local/ali
Les fichiers sont mis à jours par le service ali_logger. Quant le service est arrêté les logs se feront localement sous les répertoires /var/log/ali/local des différents noeuds du cluster.
- **réinitialisation de l'image d'un noeud** :
La réinitialisation d'un noeud se fait à partir de son image stockée sur le serveur d'administration. En cas de defaillance d'un noeud, vous pouvez désactiver un noeud (voir ci-dessous), ou le cas échéant arrêter le noeud (shutdown).
- **action sur un worker**
On utilise ali_admin depuis le serveur d'administration pour :
 - désactiver un worker. Le worker (ali_worker) termine ces jobs courant mais ne prends plus de nouveaux jobs.
ali_admin -m -p worker -s nodename (où nodename = cocat1X ou wali1X)
(la même commande exécutée une seconde fois remet le worker en état normal)
 - tuer un worker (ex : ali_admin -k -p worker -s cocat10)
 - tuer un job (ex : ali_admin -k -n pid -p worker -s cocat10)

4 Serveurs frontaux

Réseau :

Serveur	Reseau	Accès commun (Heartbeat)
sali10	local : 10.0.0.50 ext : 130.79.129.204	dns : sali1 ip : 130.79.129.161
sali11	local : 10.0.0.51 ext : 130.79.129.203	

Services :

Nom	Démarrage/arrêt	Fichier de config	
ali_server (*)	/etc/init.d/sali_service	/etc/ali	repartition des jobs ALI
ali_monitor (*)	/etc/init.d/mali_service	/etc/ali	monitoring des noeuds ALI
aserver (*)	/etc/init.d/rcatd	/usr/local/lib/rpc.rcats	service RPC pour VizieR
mon	/etc/init.d/mon	/etc/mon	monitor les services ali
heartbeat	/etc/init.d/heartbeat	/etc/ha.d/haresources /etc/ha.d/ha.cf	bascule des machines sali1x

(*) : Les services ALI et aserver sont lancés uniquement sur la machine désigné par Heartbeat (i.e. : sali1 qui est soit sali10, soit sali11)

Maintenance :

user	password	
alice	alice1	compte utilisateur
wali	-	services ALI
root	par ssh depuis le serveur d'administration	

Note : lorsque « mon » détecte l'absence d'un service ALI, le service heartbeat est redémarré et les serveurs frontaux basculent.

Pour forcer la bascule d'un serveur frontal à l'autre : arreter le service ali_server ou ali_monitor

5 Les serveurs d'exécutions

Réseau :

Uniquement accessible dans le réseau local (ssh via le serveur d'administration)

Services :

Nom	Démarrage/arrêt	Fichiers de config	
ali_worker	/etc/init.d/wali_service	/etc/ali/	exécutions des jobs ALI

Maintenance :

user	password	
alice	alice1	compte utilisateur services ALI
root	par ssh depuis le serveur d'administration	

6 Taches d'administrations

Note : lors d'un reboot complet du cluster (i.e. : tous les noeuds). Il faudra **redémarrer en premier le serveur d'administration** (en raison de la présence d'un serveur dhcp).

Reboot d'un serveur frontal bloqué	Cela peut être du à un mauvais arrêt du service heartbeat. Dans ce cas tuez les processus de heartbeat.
Plus aucune requête ne s'effectue	Arreter ali_server sur le serveur frontal pour forcer la bascule entre serveurs frontaux
Log plein sur le serveur d'administration	Arreter ali_logger
Processus ali_logger trop gourmand en cpu	Arreter ali_logger
Worker défectueux	Arrêter le service ali_worker sur le noeud. (tout service est représenté sur au moins 2 noeuds, le service reste donc disponible au niveau global (voir superviseur))

Comportement en cas de reboot :

sali1x	Si le noeud est celui designe pas sali1, heartbeat bascule les services et l'adresse ip sur le second noeud. Les jobs en cours sont perdus. Sinon, reboot normal
wali1	Le superviseur est indisponible durant le temps du reboot Les logs ALI sont locaux (/var/log/ali/local)
cocat1x wali1x	Les services du noeuds sont indisponibles. Toutefois, chaque service étant dupliqué sur au moins 2 noeuds, le service reste disponible au niveau global

En cas de reboot, vérifier que le service ALI soit toujours disponible une fois celui-ci terminé.

Test : tester ALI en utilisant VizieR :

<http://vizier.u-strasbg.fr/local/viz-bin/VizieR-4?-source=2246>

L'url VizieR ci-dessus est en mode bavard. Vérifier que le texte suivant apparait (fin de page) :

```
#--- generated by cocat1X for Ali
```

Où X=1,..4

7 Récapitulatif ALI

Server	ali_server	ali_monitor	ali_worker	ali_logger
wali1	-	-	-	X
sali1x	X	X	-	-
cocat1x	-	-	X	-
wali11, ..., wali15	-	-	X	-

Remarque : l'ensemble des fichiers de configurations ALI listés ci-dessous peuvent se trouver sur tous les serveurs. Il ne sont cependant nécessaire que sur les serveurs mentionnés dans le tableau ci-dessous.

Fichier	Description	Serveurs
cali.conf	pour ali_sumitter	sali1
sali.conf	pour ali_server	sali1
sali_auth_clients	clients autorisés pour ali_server	sali1
mali.conf	pour ali_monitor	sali1, wali10
wali.conf	pour ali_worker	cocat1x, sali11, ..., sali14
wali_auth_clients	clients autorisés pour ali_worker	cocat1x, sali11, ..., sali14
wali_avail_items	liste des applications disponible du worker	cocat1x, sali11, ..., sali14
lali.conf	pour ali_logger	wali10