

Mettre en place un miroir public

Comment faire un Distrib-coffee

Olivier Thauvin

7 juillet 2010

Plan

- 1 Généralités
- 2 Définition du besoin
- 3 Le Matériel
- 4 Coté Soft
- 5 Les miroirs
- 6 Des statistiques

Plan - Généralités

- 1 Généralités
- 2 Définition du besoin
- 3 Le Matériel
- 4 Coté Soft
- 5 Les miroirs
- 6 Des statistiques

Définition

Service de fichiers

Un serveur accessible pour distribuer des fichiers de manière anonyme et gratuite

- distributions Linux/BSD/...
- logiciels
- documentations (RFC, ...)

Définition

Service de fichiers

Un serveur accessible pour distribuer des fichiers de manière anonyme et gratuite

- distributions Linux/BSD/...
- logiciels
- documentations (RFC, ...)

Pourquoi ?

- accélérer les installations
- diffuser (aussi) vos propres fichiers

Les risques ?

Ce que vous risquez

- votre serveur va devenir important

Les risques ?

Ce que vous risquez

- votre serveur va devenir important
- il va tomber en panne

Les risques ?

Ce que vous risquez

- votre serveur va devenir important
- il va tomber en panne
- sous la pression des utilisateurs, vous essayez de le réparer toute la nuit

Les risques ?

Ce que vous risquez

- votre serveur va devenir important
- il va tomber en panne
- sous la pression des utilisateurs, vous essayez de le réparer toute la nuit
 - votre femme vous quitte
 - vous êtes viré car le service n'est pas assuré

Les risques ?

Ce que vous risquez

- votre serveur va devenir important
- il va tomber en panne
- sous la pression des utilisateurs, vous essayez de le réparer toute la nuit
 - votre femme vous quitte
 - vous êtes viré car le service n'est pas assuré
- vous finissez célibataire, ruiné et honteux !

Les risques ?

Ce que vous risquez

- votre serveur va devenir important
- il va tomber en panne
- sous la pression des utilisateurs, vous essayez de le réparer toute la nuit
 - votre femme vous quitte
 - vous êtes viré car le service n'est pas assuré
- vous finissez célibataire, ruiné et honteux !

Avez-vous mérité cela ?

Plan - Définition du besoin

- 1 Généralités
- 2 Définition du besoin**
- 3 Le Matériel
- 4 Coté Soft
- 5 Les miroirs
- 6 Des statistiques

Rappels !

Unités de mesure

- 1 octet fait 8 bits,
donc 100 mbits fait 13 107 200 octets (12,5 Mo),
donc 1 gbit fait 134 217 728 octets (128 Mo)

Rappels !

Unités de mesure

- 1 octet fait 8 bits,
donc 100 mbits fait 13 107 200 octets (12,5 Mo),
donc 1 gbit fait 134 217 728 octets (128 Mo)
- 1 kilo fait 1024,
donc 1 Go fait 1 073 741 824 octets,
donc 1 To fait 1 099 511 627 780 octets

Rappels !

Unités de mesure

- 1 octet fait 8 bits,
donc 100 mbits fait 13 107 200 octets (12,5 Mo),
donc 1 gbit fait 134 217 728 octets (128 Mo)
- 1 kilo fait 1024,
donc 1 Go fait 1 073 741 824 octets,
donc 1 To fait 1 099 511 627 780 octets
- un disque de 1 To fait 10^{12} octets,
donc ce disque fait 931,32 Go

Portée du service

Serveur privé

- réservé à des personnes / réseaux identifiés
- charge prévisible

Ne posent généralement pas de problèmes particuliers.

Portée du service

Serveur privé

- réservé à des personnes / réseaux identifiés
- charge prévisible

Ne posent généralement pas de problèmes particuliers.

Serveur public

- visiteurs inconnus
 - concernant leur identité
 - concernant leur nombre
- risque de surcharge
- risque d'attaque (DoS, piratage)

Contenus diffusés

Quels contenus

- Systèmes d'exploitation
 - mises à jour
 - systèmes complets, images cd/dvd
- Sources de logiciels
- Fichiers vidéos/son (légaux !!!)

Contenus diffusés

Quels contenus

- Systèmes d'exploitation
 - mises à jour
 - systèmes complets, images cd/dvd
- Sources de logiciels
- Fichiers vidéos/son (légaux !!!)

Vers qui ?

- Utilisateurs/Sys admin.
- Developeurs/contributeurs
- D'autres miroirs

Exemples de taille

Distributions (en Go)

	totalité		une version	
Mandriva	1258.90	2010.0	110.18	
Fedora	620.89	12	89.95	
Ubuntu	318.50	-	-	

Questions :

La vie, l'univers et le reste

- quelle capacité ?
- quelle bande passante ?
- quelle fiabilité du service ?

Questions :

La vie, l'univers et le reste

- quelle capacité ?
- quelle bande passante ?
- quelle fiabilité du service ?

- le temps que vous pouvez y accorder
- le coût en €

Plan - Le Matériel

- 1 Généralités
- 2 Définition du besoin
- 3 Le Matériel**
- 4 Coté Soft
- 5 Les miroirs
- 6 Des statistiques

Plan - Le Matériel

- 3** Le Matériel
 - infrastructure nécessaire
 - La machine
 - Stockage
 - Résumé

Hébergement

Les classiques

- réseau électrique ondulé
- salle climatisée
- une solution d'accès distant est un plus (ipmi, kvm ip)

Hébergement

Les classiques

- réseau électrique ondulé
- salle climatisée
- une solution d'accès distant est un plus (ipmi, kvm ip)

connection du site à internet

- en download
 - 20 mbits peuvent suffire
 - pour des mises à jour correcte : 100 mbits
- en upload (serveur public)
 - 100 mbits minimum, gigabit recommandé

Plan - Le Matériel

- 3** Le Matériel
 - infrastructure nécessaire
 - La machine**
 - Stockage
 - Résumé

Configuration

Puissance

- 2 ou 4 CPU, peu importe la puissance

Configuration

Puissance

- 2 ou 4 CPU, peu importe la puissance
- RAM : 4 Go minimum (faire du cache)

Configuration

Puissance

- 2 ou 4 CPU, peu importe la puissance
- RAM : 4 Go minimum (faire du cache)
- disque (pour l'OS) :
 - 20 Go (système + swap)
 - 30 Go pour les logs

Configuration

Puissance

- 2 ou 4 CPU, peu importe la puissance
- RAM : 4 Go minimum (faire du cache)
- disque (pour l'OS) :
 - 20 Go (système + swap)
 - 30 Go pour les logs

Sécurité

- alimentations redondantes
- garantie/maintenance constructeur (3 voire 5 ans)

Configuration

Puissance

- 2 ou 4 CPU, peu importe la puissance
- RAM : 4 Go minimum (faire du cache)
- disque (pour l'OS) :
 - 20 Go (système + swap)
 - 30 Go pour les logs

Sécurité

- alimentations redondantes
- garantie/maintenance constructeur (3 voire 5 ans)
- un système de sauvegarde du système

Plan - Le Matériel

- 3** Le Matériel
 - infrastructure nécessaire
 - La machine
 - Stockage**
 - Résumé

Besoins :

Performance

- peu d'écritures, beaucoup de lectures
- peu de très gros débits
- beaucoup d'accès concurrents

Besoins :

Performance

- peu d'écritures, beaucoup de lectures
- peu de très gros débits
- beaucoup d'accès concurrents

Taille : prévoir large

- espace utile \approx 80% espace brut
- les distributions grossissent
(Mdv 2007.0 : 29Go, Mdv 2010.0 : 48Go)
- répondre aux futurs besoins

Solutions de stockage

Disques internes

- disques SATA ou SAS
- **Raid** hard ou soft

Solutions de stockage

Disques internes

- disques SATA ou SAS
- **Raid** hard ou soft

Solutions externes / Baies autonomes

Conn.	type disque	commentaires
SCSI	IDE	vieilles baies
SCSI	SATA	bon compromis qualité/prix
SAS	SAS	très couteux
F.C.	SATA	performant, voir le prix de la carte
F.C.	SAS	le top mais cher !

Exemple de baies

Baies de distrib-coffee



- haut : 8 x 250Go, Raid 5
- bas : 16 x 1To, Raid 6

Faire intervenir le RAID

RAID hard ou soft ?

Pourquoi pas du soft :

- peu couteux sur les CPUs actuels
- indépendant du matériel

Faire intervenir le RAID

RAID hard ou soft ?

Pourquoi pas du soft :

- peu couteux sur les CPUs actuels
- indépendant du matériel
- un RAID full hard reste plus souple et performant
- **fuyez les RAID semi hard**

Faire intervenir le RAID

RAID hard ou soft ?

Pourquoi pas du soft :

- peu couteux sur les CPUs actuels
- indépendant du matériel
- un RAID full hard reste plus souple et performant
- **fuyez les RAID semi hard**

Quel RAID ?

RAID 0 Stripping Vous plaisantez ?

Faire intervenir le RAID

RAID hard ou soft ?

Pourquoi pas du soft :

- peu couteux sur les CPUs actuels
- indépendant du matériel
- un RAID full hard reste plus souple et performant
- **fuyez les RAID semi hard**

Quel RAID ?

RAID 1	Mirroring	Cher
RAID 0+1	Mirror + Stripe	Excellent mais cher
RAID 5	Parité	Bon compromis
RAID 6	2 parités	Bon compromis, plus lourd

Plan - Le Matériel

- 3** Le Matériel
 - infrastructure nécessaire
 - La machine
 - Stockage
 - Résumé**

Prix

Tarifs :

Machine	500 - 2500 €
Stockage	300 - 20000 €

Prix

Tarifs :

Machine	500 - 2500 €
Stockage	300 - 20000 €

Distrib-coffee

bi-quad 2ghz, 32Go RAM	1500 €
Stockage :	
Baie 8x500Go SCSI	3000 €
Baie 16x1To SCSI	5000 €

Prix

Tarifs :

Machine	500 - 2500 €
Stockage	300 - 20000 €

Distrib-coffee

bi-quad 2ghz, 32Go RAM	1500 €
Stockage :	
Baie 8x500Go SCSI	3000 €
Baie 16x1To SCSI	5000 €

Nos prix sont exprimés hors taxes

Plan - Coté Soft

- 1 Généralités
- 2 Définition du besoin
- 3 Le Matériel
- 4 Coté Soft**
- 5 Les miroirs
- 6 Des statistiques

Plan - Côté Soft

4 Côté Soft

■ L'OS

■ Services

- http
- ftp
- Rsync
- Autres

■ QOS

■ Visibilité

Quel Os ?

Unix

Préférez un système 64bits.

Quel Os ?



Do not feed the
troll.

Unix

Préférez un système 64bits.

- Linux
- BSD
- autres...

Quel Os ?



Do not feed the troll.

Unix

Préférez un système 64bits.

- Linux
- BSD
- autres...

Windows ?

Les retours d'expérience m'intéressent !

Quel Os ?



Do not feed the troll.

Unix

Préférez un système 64bits.

- Linux
- BSD
- autres...

Windows ?

Les retours d'expérience m'intéressent !

Distrib-coffee

Linux Mandriva 2010.0 x86_64

Partitionnement 1/2

Pour le système

Prévoir :

- swap (4Go max)
- /var séparé (pensez aux logs !)

Partitionnement 1/2

Pour le système

Prévoir :

- swap (4Go max)
- /var séparé (pensez aux logs !)

Stockage

Utilisation de LVM (linux) :

- facilite la gestion de gros volumes
- extension sur plusieurs disques

Partitionnement 2/2

Conseils

- un seul système de fichiers (évite les déplacements coûteux entre disques)
- ne pas exposer la racine mais un sous répertoire

Partitionnement 2/2

Conseils

- un seul système de fichiers (évite les déplacements coûteux entre disques)
- ne pas exposer la racine mais un sous répertoire

```
/dev/vg/lv          /mnt/stockage auto  
/mnt/stockage/pub /var/ftp/pub  bind
```


Partitionnement 2/2

Conseils

- un seul système de fichiers (évite les déplacements coûteux entre disques)
- ne pas exposer la racine mais un sous répertoire
`/dev/vg/lv` `/mnt/stockage auto`
`/mnt/stockage/pub` `/var/ftp/pub` `bind`
- système supportant un grand nombre de fichiers

Partitionnement 2/2

Conseils

- un seul système de fichiers (évite les déplacements coûteux entre disques)
- ne pas exposer la racine mais un sous répertoire
`/dev/vg/lv` `/mnt/stockage auto`
`/mnt/stockage/pub` `/var/ftp/pub` `bind`
- système supportant un grand nombre de fichiers
- redimensionnement (à chaud)

Partitionnement 2/2

Conseils

- un seul système de fichiers (évite les déplacements coûteux entre disques)
- ne pas exposer la racine mais un sous répertoire
`/dev/vg/lv` `/mnt/stockage auto`
`/mnt/stockage/pub` `/var/ftp/pub` `bind`
- système supportant un grand nombre de fichiers
- redimensionnement (à chaud)
- **journalisé**

Partitionnement 2/2

Conseils

- un seul système de fichiers (évite les déplacements coûteux entre disques)
- ne pas exposer la racine mais un sous répertoire
`/dev/vg/lv` `/mnt/stockage auto`
`/mnt/stockage/pub` `/var/ftp/pub` `bind`
- système supportant un grand nombre de fichiers
- redimensionnement (à chaud)
- **journalisé**
- fiable et maintenu

Partitionnement 2/2

Conseils

- un seul système de fichiers (évite les déplacements coûteux entre disques)
- ne pas exposer la racine mais un sous répertoire
`/dev/vg/lv` `/mnt/stockage auto`
`/mnt/stockage/pub` `/var/ftp/pub` `bind`
- système supportant un grand nombre de fichiers
- redimensionnement (à chaud)
- **journalisé**
- fiable et maintenu

Donc ext3, ext4, xfs (linux).

Plan - Coté Soft

4 Coté Soft

- L'OS

- Services

 - http

 - ftp

 - Rsync

 - Autres

- QOS

- Visibilité

Le web

Contraintes

- protocole mal adapté à du téléchargement pur
- beaucoup de connexions
- des connexions longues (temps de téléchargement)

Le web

Contraintes

- protocole mal adapté à du téléchargement pur
- beaucoup de connexions
- des connexions longues (temps de téléchargement)

Apache ?

- souple, nombreux supports (via modules)
- configuration parfois peu aisée
- besoin de *restart* qui arrête toutes les connexions en cours
- gourmand en RAM

Apache : configuration

Pensez à :

- adapter le nombre de processus autorisés (`MaxClients`)
- autoriser sur l'arborescence le suivi des liens (`SymLinksIfOwnerMatch`)

Répertoire des miroirs

```
<Directory /var/ftp/pub>
  order allow,deny
  Allow from all
  Options +Indexes +MultiViews +SymLinksIfOwnerMatch
</Directory>

Alias /pub /var/ftp/pub
```

Téléchargement intelligent ?

Problème des mass downloader

- plusieurs connections (jusque 30 par téléchargement) simultanées et répétitives
- des accès (ouverture de fichiers) répétitifs et concurrents

Téléchargement intelligent ?

Problème des mass downloader

- plusieurs connections (jusque 30 par téléchargement) simultanées et répétitives
- des accès (ouverture de fichiers) répétitifs et concurrents

Solution :

Pour apache : `mod-limitipconn`, renvoie une erreur 503 en cas de connections multiples

- permet de limiter seulement une arborescence
- ne détecte pas bien la fin des connections : \
- efficace

Apache mod-limitipconn : configuration

```
<Directory /var/ftp/pub>
  [...]
  <IfModule mod_limitipconn.c>
    MaxConnPerIP 5
    ErrorDocument 503 "5 download max allowed."
  </IfModule>
</Directory>
```

File Transfert Protocol

Les problèmes :

- `ls` récursif : à interdire
- penser à autoriser le `ftp` `passif`

File Transfert Protocol

Les problèmes :

- `ls` récursif : à interdire
- penser à autoriser le `ftp` passif

Quel serveur ?

- `proftpd`

File Transfert Protocol

Les problèmes :

- `ls` récursif : à interdire
- penser à autoriser le `ftp` passif

Quel serveur ?

- `proftpd`
- `vsftpd`

File Transfert Protocol

Les problèmes :

- `ls` récursif : à interdire
- penser à autoriser le `ftp` passif

Quel serveur ?

- `proftpd`
- `vsftpd`
- `pureftpd`
- ...

Rsync (serveur)

Utilisations

- avant tout synchronisation d'arborescence
- téléchargement de fichier (certaines distributions)

Rsync (serveur)

Utilisations

- avant tout synchronisation d'arborescence
- téléchargement de fichier (certaines distributions)
- reprise de téléchargement long (y compris gros fichiers)
- gourmand en accès disque (coté serveur et client)

Rsync (serveur)

Utilisations

- avant tout synchronisation d'arborescence
- téléchargement de fichier (certaines distributions)
- reprise de téléchargement long (y compris gros fichiers)
- gourmand en accès disque (coté serveur et client)

Limiter rsync

- désactiver la compression (`-z`) : très souvent inutile (fichiers déjà compressés)

Rsync (serveur)

Utilisations

- avant tout synchronisation d'arborescence
- téléchargement de fichier (certaines distributions)
- reprise de téléchargement long (y compris gros fichiers)
- gourmand en accès disque (coté serveur et client)

Limiter rsync

- désactiver la compression ($-z$) : très souvent inutile (fichiers déjà compressés)
- désactiver le recalcul des *checksum* ($-c$) qui force coté serveur la relecture des fichiers

Soyez inventif

Des exemples

- p2p (`bittorent`) : très lourd sur les firewall et parfois interdit

Soyez inventif

Des exemples

- p2p (`bittorent`) : très lourd sur les firewall et parfois interdit
- smb/cifs (`samba`) : très utile, mais reste en **interne**

Soyez inventif

Des exemples

- p2p (`bittorent`) : très lourd sur les firewall et parfois interdit
- smb/cifs (`samba`) : très utile, mais reste en **interne**
- appletalk (`netatalk`) : de moins en moins utile, en **interne** seulement

Soyez inventif

Des exemples

- p2p (`bittorent`) : très lourd sur les firewall et parfois interdit
- smb/cifs (`samba`) : très utile, mais reste en **interne**
- appletalk (`netatalk`) : de moins en moins utile, en **interne** seulement
- nfs : très utile sur les UNIX, en **interne** et seulement en **lecture seule**

Plan - Côté Soft

4 Côté Soft

- L'OS
- Services
 - http
 - ftp
 - Rsync
 - Autres
- QOS
- Visibilité

QOS : pourquoi ?

Selon Philippe

De la qualité qui se résume à réduire la bande passante

QOS : pourquoi ?

Selon Philippe

De la qualité qui se résume à réduire la bande passante

Raisons :

- éviter la surcharge de connection du site
- favoriser les réseaux amis (ou l'entreprise)
- favoriser certaines méthodes de téléchargement
- rendre équitable les téléchargements

QOS, mise en pratique (linux)

Méthodes :

- `tc` : filtres et qos
- `netfilter` pour les filtres, `tc` pour la qos

Les filtres `tc` sont subtilement puissants, on préfère utiliser `netfilter` (`iptables`).

QOS, mise en pratique (linux)

Méthodes :

- `tc` : filtres et qos
- `netfilter` pour les filtres, `tc` pour la qos

Les filtres `tc` sont subtilement puissants, on préfère utiliser `netfilter` (`iptables`).

Quelles QOS ?

- contrôle de la bande passante : `htb`
- régulation :
 - `pfifo` pour les très gros débits
 - `cfq` pour égaliser les flux (`ftp`, `http`)

QOS, exemples

iptables

```
iptables -t mangle -A OUTPUT -p tcp -m tcp \  
  --sport 80 -j MARK --set-mark 122
```

tc

```
tc qdisc add dev eth0 root handle 1: htb r2q 1500  
tc class add dev eth0 parent 1:1 classid 1:122 htb \  
  rate 26mbit ceil 60mbit prio 1 burst 50mbit cburst 100kb  
  
tc qdisc add dev eth0 parent 1:122 handle 122: sfq perturb 5  
  
# égalise par source/destination et non par connection:  
tc filter add dev eth0 protocol ip parent 122: handle 1221 \  
  flow hash keys src,dst divisor 1024
```

Plan - Coté Soft

4 Coté Soft

- L'OS
- Services
 - http
 - ftp
 - Rsync
 - Autres
- QOS
- **Visibilité**

Ce que les gens verront

Page Web principale :

- un contact (un mail)
- hébergeur du serveur (votre société)
- comment accéder aux miroirs
- informations légales
- les recommandations d'usage

Sur chaque service

- rsync : motd
- ftp : welcome

Plan - Les miroirs

- 1 Généralités
- 2 Définition du besoin
- 3 Le Matériel
- 4 Coté Soft
- 5 Les miroirs**
- 6 Des statistiques

Mise en place

Les distributions

- difficulté à trouver les informations
- politiques différentes (outils, options des outils)
- multitudes de contacts (un par distribution)
- communication inappropriée et/ou surnuméraire

Les (quelques) progrès

- Mirror Brain

Mes conseils

rsync !

- support des liens hard
- reprise des téléchargements partiels

rsync : les bonnes options

- `-a` : récursif, sauvegarde des permissions
- `-H` : préserve les *liens hard*

Mes conseils

rsync !

- support des liens hard
- reprise des téléchargements partiels

rsync : les bonnes options

- `-a` : récursif, sauvegarde des permissions
- `-H` : préserve les *liens hard*
- `--temp-dir` : répertoire temporaire

Mes conseils

rsync !

- support des liens hard
- reprise des téléchargements partiels

rsync : les bonnes options

- `-a` : récursif, sauvegarde des permissions
- `-H` : préserve les *liens hard*
- `--temp-dir` : répertoire temporaire
- `--partial --partial-dir` : garde les fichiers temporaires en dehors de l'arborescence

Automatiser les synchronisations

Pas de solution toute faite

Solution simple : `cron` et script shell.

Automatiser les synchronisations

Pas de solution toute faite

Solution simple : `cron` et script shell.

Il faut :

- poser un verrou pour éviter les lancements multiples

Automatiser les synchronisations

Pas de solution toute faite

Solution simple : `cron` et script shell.

Il faut :

- poser un verrou pour éviter les lancements multiples
- poser un timeout (rsync se bloque parfois)

Automatiser les synchronisations

Pas de solution toute faite

Solution simple : `cron` et script shell.

Il faut :

- poser un verrou pour éviter les lancements multiples
- poser un timeout (rsync se bloque parfois)
- un script par miroir

Automatiser les synchronisations

Pas de solution toute faite

Solution simple : `cron` et script shell.

Il faut :

- poser un verrou pour éviter les lancements multiples
- poser un timeout (rsync se bloque parfois)
- un script par miroir

MMM

Démons en `perl` :

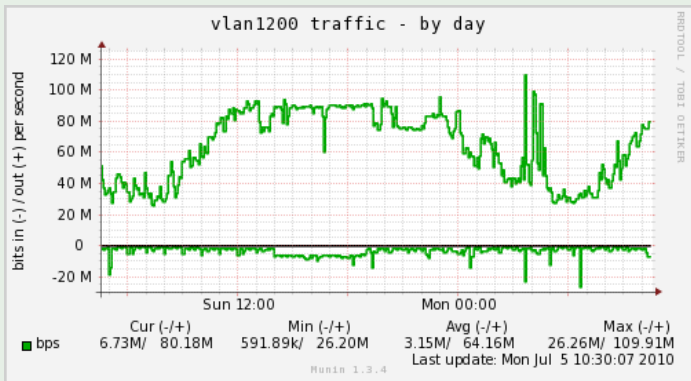
- fichier de configuration unique
- détection des erreurs
- pose de verrou

Plan - Des statistiques

- 1 Généralités
- 2 Définition du besoin
- 3 Le Matériel
- 4 Coté Soft
- 5 Les miroirs
- 6 Des statistiques**

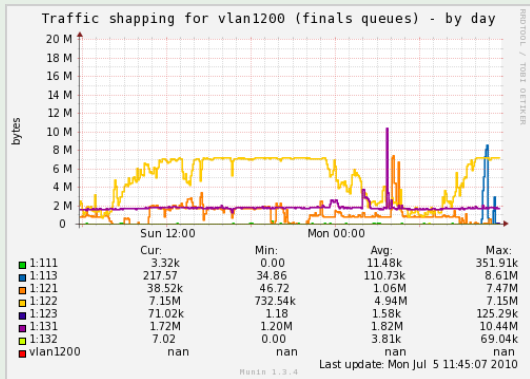
Réseau

Traffic réseau



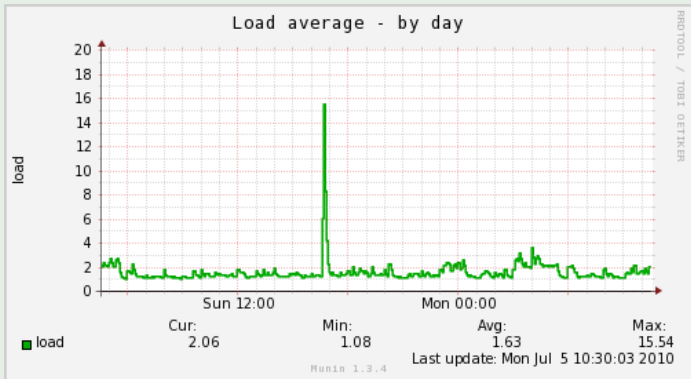
Réseau

Débits par QOS



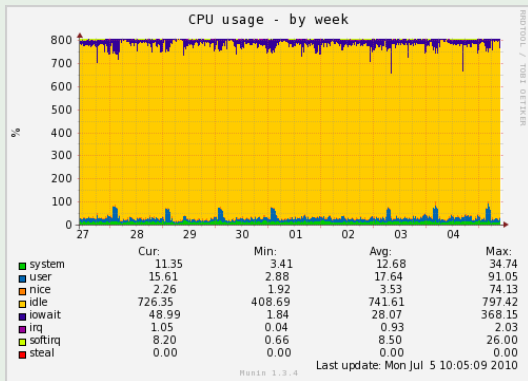
Système

Charge de la machine



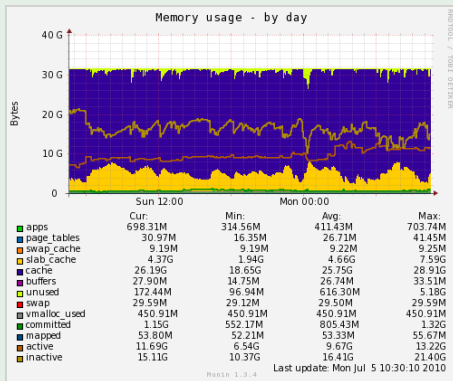
Système

Utilisation CPU



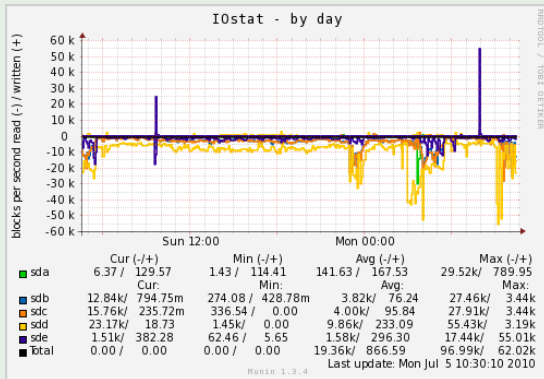
Système

Utilisation de la mémoire



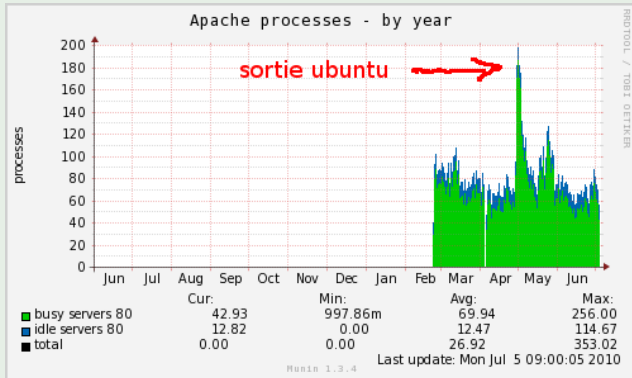
Système

I/O disques



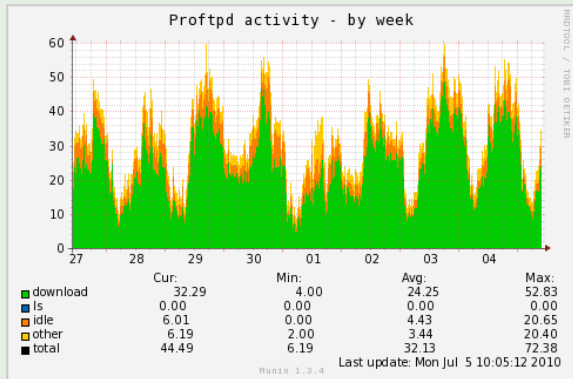
Services

Apache



Services

Proftpd



Merci

Url

`http://distrib-coffee.ipsl.jussieu.fr`

Des questions ?

`olivier.thauvin at latmos.ipsl.fr`

Mille Mercis

- Amaz
- Jehane
- DjeZaeL
- shikamaru