

X/Stra - Le stockage à l'IRMA

Alexis Palaticky

18 juin 2024



Introduction - Présentation de l'Institut et du service informatique

- ▶ UMR Université de Strasbourg - CNRS
- ▶ 180 personnes
- ▶ 7 équipes de recherche
- ▶ Le service informatique :
 - Administration système et réseau : 1 IR, 1 IE
 - Calcul scientifique et développement : 3 IR



Introduction - Type de données produites

- ▶ Les codes sources des projets
- ▶ Les résultats des calculs
- ▶ Les publications scientifiques
- ▶ Les données des utilisateurs
- ▶ Les données administratives



Introduction - Plan de la présentation

- ▶ Infrastructure matérielle du stockage
- ▶ Protocoles et applications
- ▶ La sécurisation des données
- ▶ Conclusion



Infrastructure matérielle du stockage

Volumétrie Totale de **650 To**

- ▶ Serveur de fichiers : 43 To
- ▶ Cluster de virtualisation : 30 To
- ▶ Serveur de sauvegarde : 256 To
- ▶ Les clusters de Calcul : 275 To
- ▶ Serveur de logs : 44 To



Le serveur de fichiers **43 To**

- ▶ Dell PowerEdge R7425 (AMD EPYC)
- ▶ Red Hat Enterprise Linux 8.9 (Oopta)

Infrastructure matérielle du stockage

Le cluster de virtualisation **30 To** (VMware)

- ▶ 4 noeuds ESXi Dell PowerEdge R7425 (AMD EPYC)
- ▶ Attachement SAS/ISCSI à une baie Dell SCV 3020



Le serveur de sauvegarde **256 To**

- ▶ Dell PowerEdge R7515 (AMD EPYC) (128 To).
- ▶ Attachement SAS à une Baie Dell Storage MD1400 (128 To).
- ▶ Red Hat Enterprise Linux 8.9 (Oopta)

Infrastructure matérielle du stockage

Le cluster de calcul gaya **205 To**

- ▶ Frontale : Dell PowerEdge R7525 (2 proc AMD EPYC 7552 2,1 Ghz 48 coeurs) (30 To NVME).
- ▶ Baie Dell Storage MD1400 (175 To) (Attachement SAS).
- ▶ 6 noeuds Dell Power Edge R6525 (2 proc AMD 7713 2 Ghz, 64 coeurs)
- ▶ Ubuntu 22.04 (Jammy)



Infrastructure matérielle du stockage

Le cluster de calcul atlas **70 To** (en fin de vie)

- ▶ Frontale : HP Proliant DL 585 G7 (4 proc AMD Opteron 6386 SE 2,7 Ghz total de 64 coeurs)
- ▶ Baie HP MDS 600 (70 To) (Attachement SAS).
- ▶ 6 noeuds Dell PowerEdge R730 Server (2 proc Intel Xeon CPU E5-2680 2,5 Ghz total de 48 coeurs)
- ▶ Ubuntu 16.04 (Xenial)



Infrastructure matérielle du stockage

Le serveur de logs **44 To**

- ▶ Dell PowerEdge R620 (Intel(R) Xeon)
- ▶ Attachement SAS Baie stockage Dell PowerVault MD 1200 (44 To).
- ▶ Ubuntu 22.04 (Jammy)

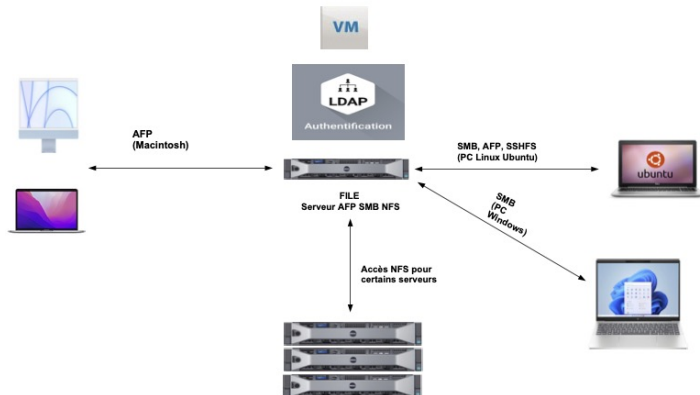


Le serveur de fichiers est accessible par :

- ▶ NFS pour partager le stockage avec les autres serveurs
- ▶ AFP (Apple Filing Protocole) (Pour les Macintosh et les PC sous Ubuntu)
- ▶ SMB (Server Message Block) (Pour les postes windows)
- ▶ SSHFS pour les postes sous Linux

Les mêmes données accessibles par ces différents protocoles

Accès aux données IRMA



Les applications :

- ▶ La forge gitlab
- ▶ Le serveur owncloud
- ▶ La plateforme de gestion de données Girder



- ▶ Machine virtuelle VMware
- ▶ Dépôts des projets et des codes (500 projets)
- ▶ Plus de 100 utilisateurs actifs (collaborations extérieures)
- ▶ Montage du serveur de fichiers pour stockage des artifacts
- ▶ Multiples fonctionnalités : GitlabPages (52 sites), tickets, stockage LFS, wiki...



- ▶ Machine virtuelle VMware
- ▶ Accès via https
- ▶ Client de synchronisation (Mac OS, Linux, Windows)
- ▶ Accès par le protocole WebDav
- ▶ Stockage des données sur le serveur de fichier (NFS)

Applications



Girder

Data management platform

- ▶ Machine virtuelle VMware
- ▶ Accès via https
- ▶ Accès au serveur de calcul par NFS



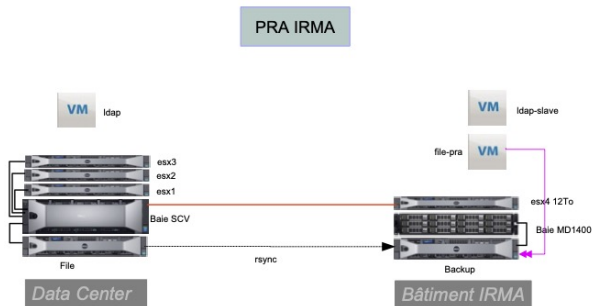
La sauvegarde :

- ▶ Veritas Netbackup
 - Une sauvegarde complète par mois (rétention 3 mois)
 - Une sauvegarde incrémentale par nuit (rétention 2 semaines)
 - Une sauvegarde cumulative par semaine (rétention 6 semaines)
- ▶ Time Machine pour certains Macintosh
- ▶ Synchronisation des /home quotidienne sur le serveur de backup (PRA).

Le Plan de Reprise d'Activité :

- ▶ 1 serveur LDAP de secours (slave)
- ▶ 1 serveur AFP de secours (file-pra qui monte les /home sur le serveur backup)
- ▶ 1 serveur ESX pouvant supporter nos VM les plus importantes

Accès aux données IRMA



Protection du réseau et supervision :

- ▶ Filtrage du réseau (Cluster Firewall Stormshield SN710)
- ▶ Supervision et monitoring des espaces disques
 - Cacti pour suivre l'évolution des espaces disques
 - Nagios pour alerter lorsqu'un espace de stockage s'approche de la saturation ou est en défaut

Conclusion

- ▶ Un système de stockage permettant aux utilisateurs d'accéder à leurs données depuis leurs postes clients ou à partir d'un serveur.
- ▶ Un stockage des données sécurisé.
- ▶ Les mêmes données accessibles via différents protocoles
- ▶ Quelles évolutions pour le stockage de demain ?

