



HAL
open science

Bilan et perspectives pour le centre de données sismologiques de Résif (noeud B)

Grégory Arneodo, Catherine Péquegnat, Helle A. Pedersen, Pierre Volcke

► **To cite this version:**

Grégory Arneodo, Catherine Péquegnat, Helle A. Pedersen, Pierre Volcke. Bilan et perspectives pour le centre de données sismologiques de Résif (noeud B). 3èmes Rencontres scientifiques et techniques RESIF, Oct 2017, Saint-Jean-De-Mont (85160), France. 2017. hal-02166055

HAL Id: hal-02166055

<https://hal.science/hal-02166055v1>

Submitted on 26 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Résumé



Le centre de données sismologiques RESIF (dit aussi 'noeud B', DOI:10.17616/R37Q06) est hébergé par l'Université Grenoble Alpes et co-exploité par l'UMS GRICAD (<https://gricad.univ-grenoble-alpes.fr/>) et le centre de données de l'OSUG. Les services de données RESIF sont développés, déployés et maintenus par l'équipe RESIF d'ISTerre.

Le prototypage du centre s'est fait sur des moyens propres hébergés dès 2011 par la DSI de l'université de Grenoble. Les données et services ont été progressivement migrés sur les infrastructures de GRICAD : SPRING (pour le réseau), SUMMER (pour le stockage) et WINTER (pour la virtualisation). La fin de cette migration (à l'été 2017) correspond au passage en exploitation du noeud B, qui est composé à ce jour de 4 machines physiques et de 22 machines virtuelles.

Les codes développés pour le noeud B sont accessibles sur les dépôts GIT : <https://github.com/resif> et <https://gricad-gitlab.univ-grenoble-alpes.fr/OSUG/RESIF>.

RESIF-DC est l'un des 11 noeuds de l'archive virtuelle européenne EIDA (<http://www.orfeus-eu.org/data/eida/>). Il est l'un des 19 centres de données mondiaux exposant des données et des métadonnées selon les standards de l'International Federation of Digital Seismograph Networks (FDSN). A ce jour, RESIF-DC donne accès aux données de 11 réseaux permanents et 50 réseaux temporaires, pour 2019 stations et 13563 channels. La volumétrie des données validées est d'environ 50 To.

Infrastructures, monitoring

Informations et contact : <https://gricad.univ-grenoble-alpes.fr/catalogue-services/stockage-summer>



L'infrastructure, basée sur les technologies du constructeur NetApp, est répartie sur trois centres de données situés sur le campus de Saint Martin d'Hères. Elle peut accueillir des volumétries pouvant aller à plusieurs Péta octets. L'infrastructure est techniquement construite pour pouvoir assurer un taux de disponibilité de 99,99%. Ce qui représente environ 1h d'indisponibilité par année. En 2016, la performance était de 95.000 IOPS partagées. Tous les équipements sont redondés (double contrôleur, double alimentation, double attachement réseau, disques de secours,...)

- Snapshot : jusqu'à 255 copies de snapshot par volume.
- Réplication asynchrone dans un délai d'1h sur un second site.
- Sauvegarde sur site distant : journalière sur 30 jours pour un taux de variation du volume de 20% maximum
- Restauration : accès iSCSI et NFS en lecture seule sur le site de sauvegarde
- Extensibilité de la volumétrie à la demande

Figures 1 & 2 : monitoring des espaces RESIF sur SUMMER adaptation d'un outil NetApp par le service informatique d'ISTerre

SUMMER (stockage)

WINTER (virtualisation)

Informations et contact : <https://gricad.univ-grenoble-alpes.fr/catalogue-services/virtualisation-winter>

Infrastructure VMware VSAN Stretched Cluster. Le choix d'une plateforme de type "hyperconvergente", le stockage des données sur disques flash, ainsi qu'un réseau très haut-débit et fortement maillé permettent d'obtenir un niveau de résilience élevé.

- Nombre de vCPU (Processeurs/Coeurs) : De 1 à 8
- RAM : De 1 à 16 Go
- Stockage (Flash SSD) : De 5 à 250 Go

La sauvegarde est assurée à l'aide du logiciel VEEAM Backup & Replication. Entre 18h et minuit, deux sauvegardes sont réalisées par deux serveurs de sauvegarde différents, chacun étant situé dans des Data Centers différents. Ces serveurs gardent chacun 7 points de restauration en local.

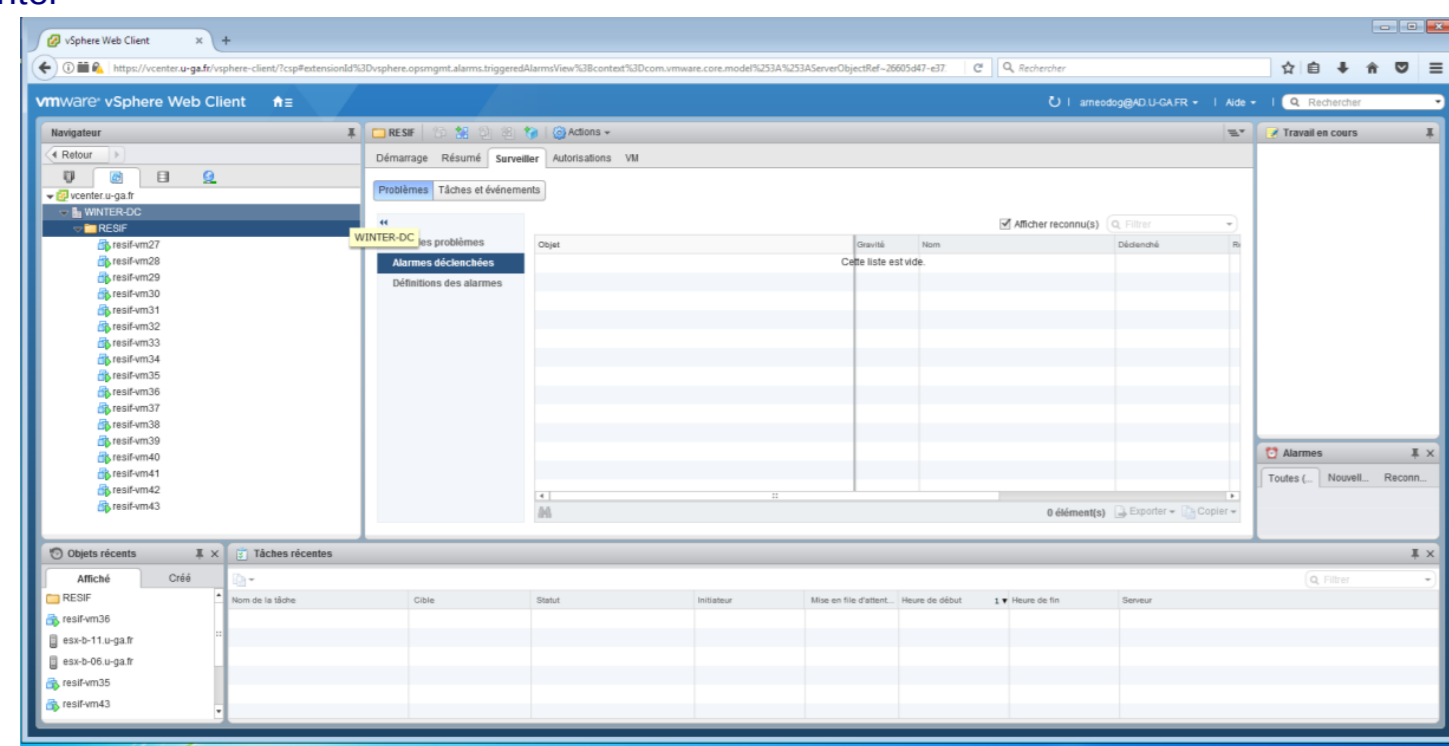
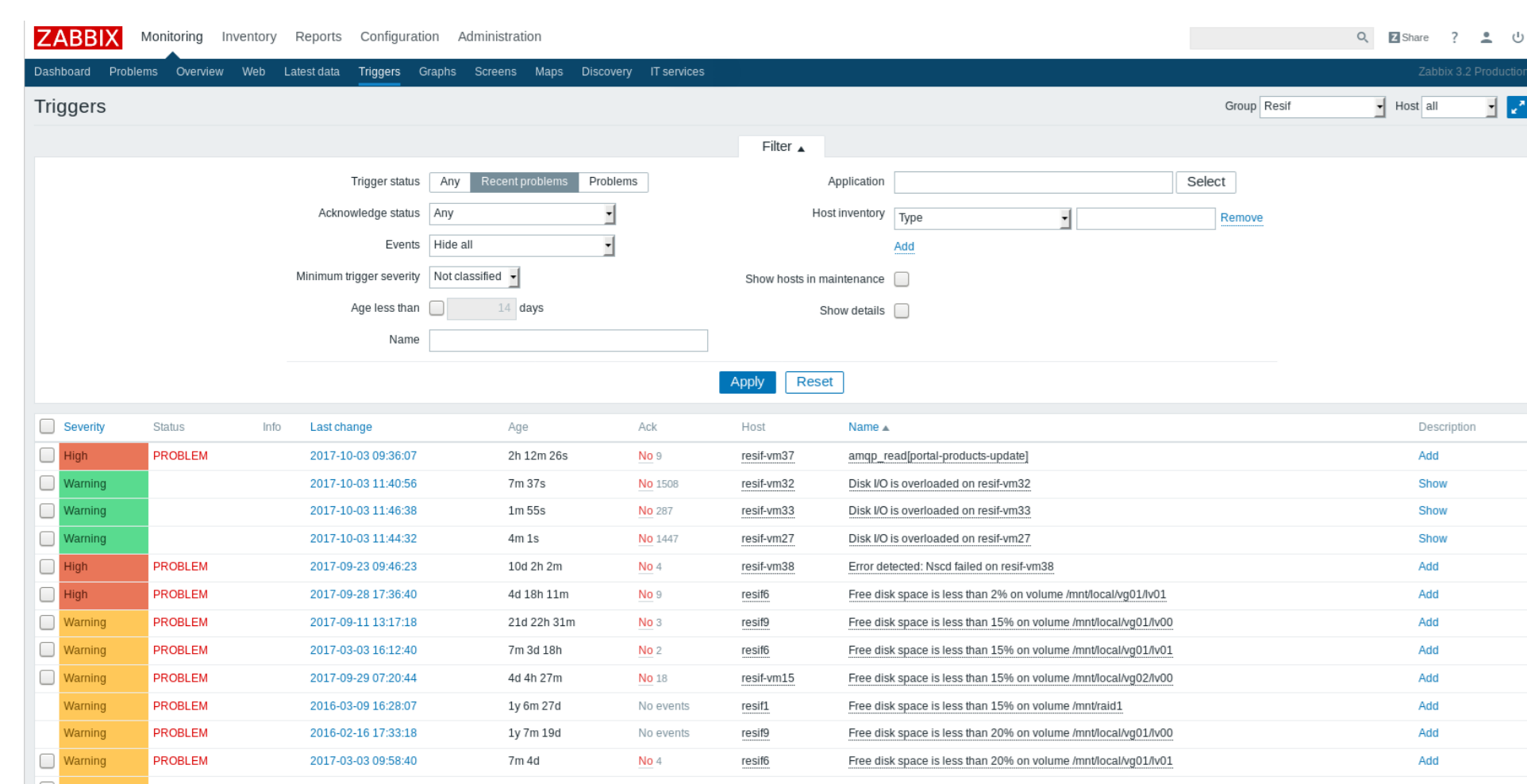
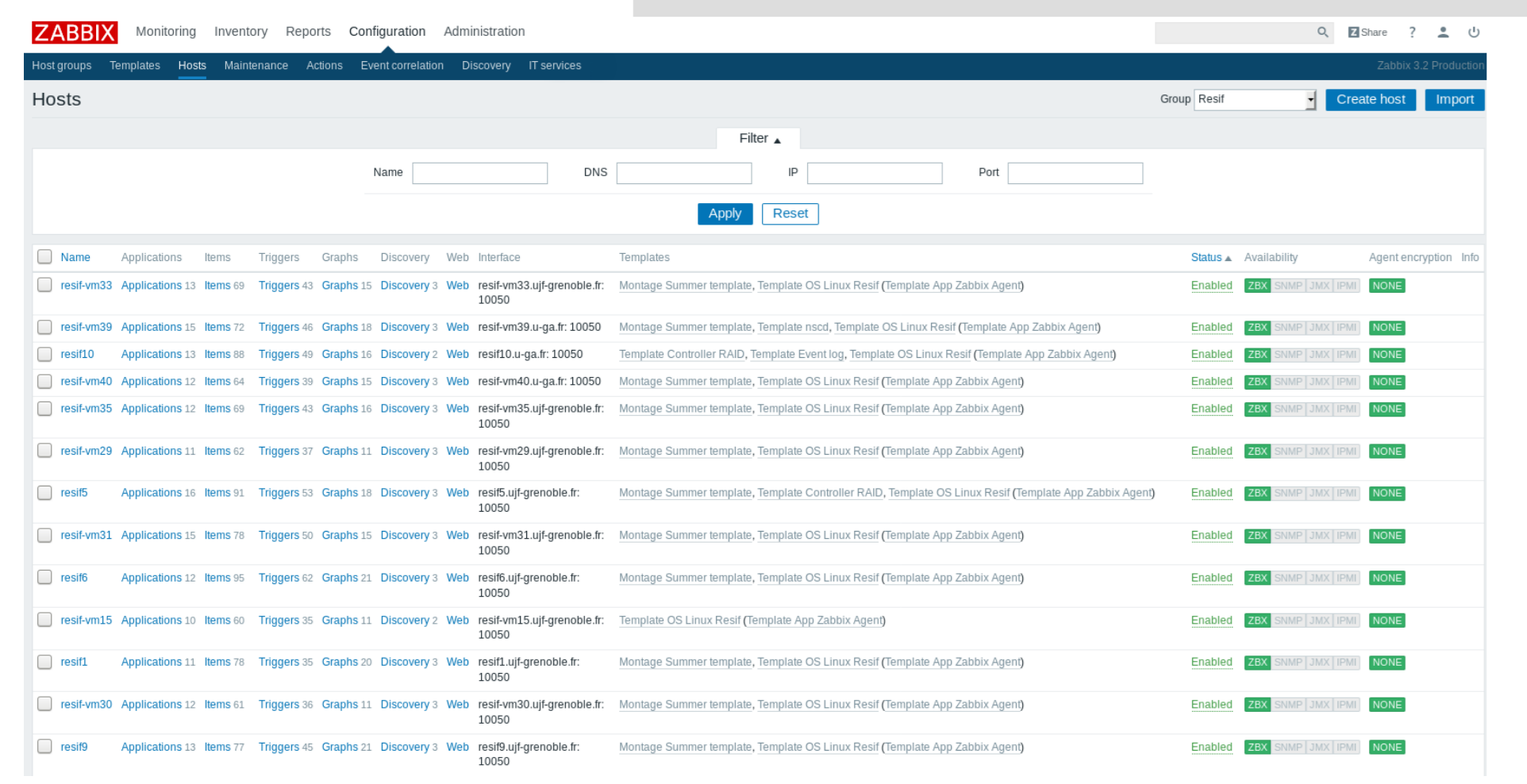


Figure 3 Gestion de l'environnement de virtualisation RESIF sur WINTER

	NON-CRITIQUE	NORMAL	CRITIQUE
DATACENTER SÉCURISÉ	•	•	•
STOCKAGE FLASH HAUTE-PERFORMANCE	•	•	•
COPIE SYNCHRONE VERS UN AUTRE DATACENTER	•	•	•
REDÉMARRAGE AUTOMATIQUE DE LA VM	•	•	•
SAUVEGARDE	(en option)	(en option)	(en option)

Figure 4 Profil des VM sur WINTER

ZABBIX Monitoring des infrastructures et des applications



Zabbix est une solution de supervision open source. Ce logiciel permet de surveiller les statuts de différents services, systèmes et réseaux. Un peu similaire à Nagios, Zabbix utilise un système de templates pour définir des hôtes à superviser, ce qui le rend très pratique pour la définition de grosses infrastructures. Nous supervisons ainsi l'ensemble des 26 machines RESIF (virtuelles et physiques)

Figures 5 & 6 Monitoring des infrastructures et services RESIF avec zabbix.

Architecture fonctionnelle du centre de données

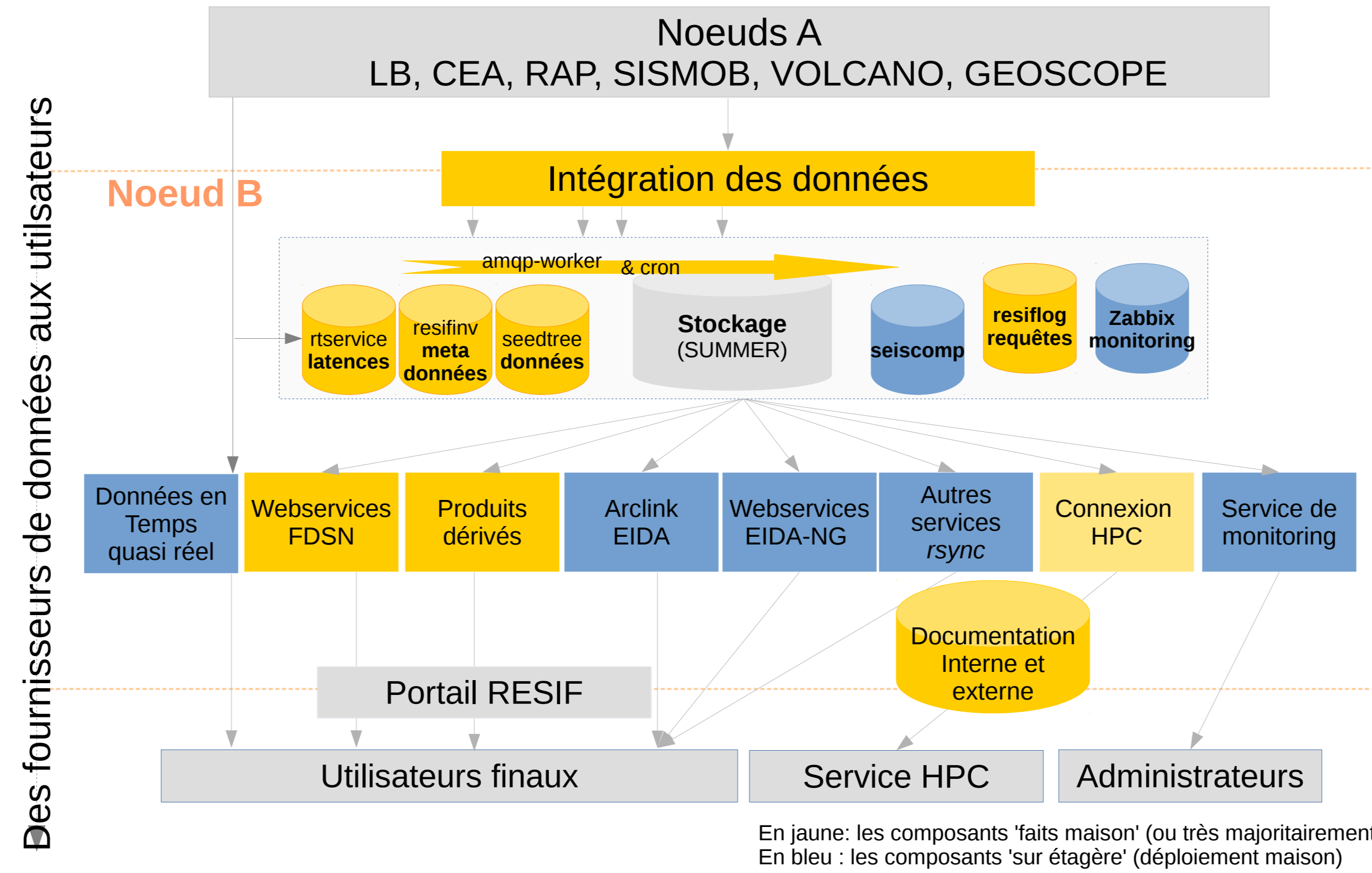
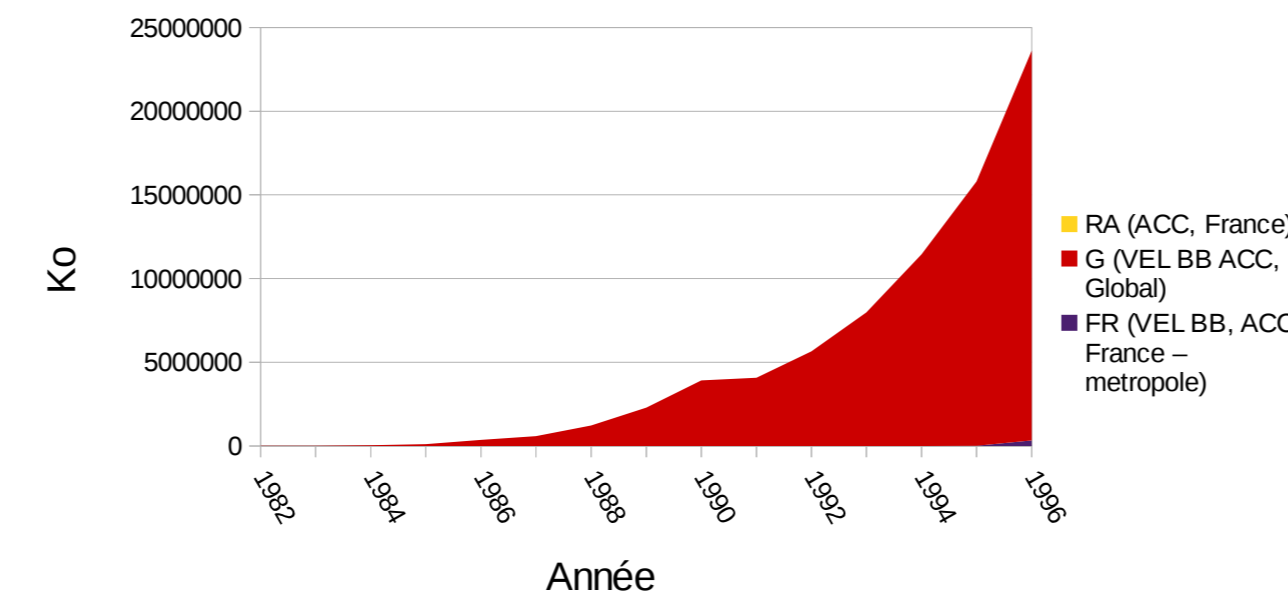


Figure 7 architecture fonctionnelle du centre de données

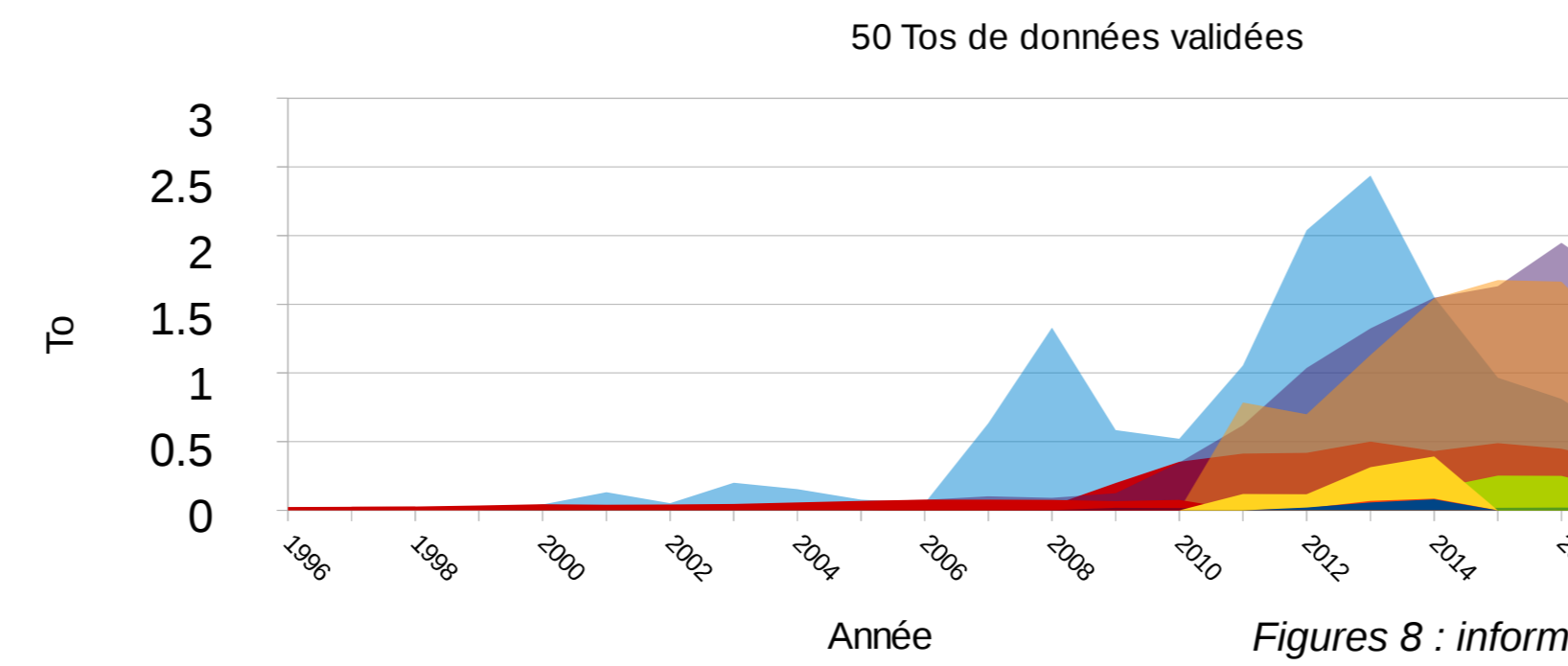
Contenu

Volumétrie dans l'archive RESIF (<=1996)



11 réseaux permanents et 50 réseaux temporaires, pour 2019 stations et 13563 channels

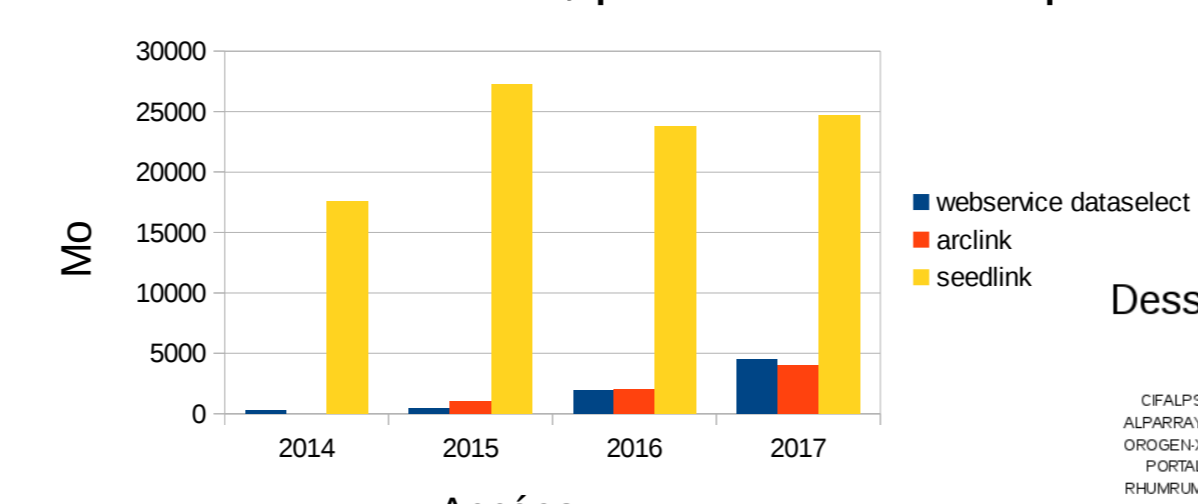
Volumétrie de l'archive RESIF (>= 1996)



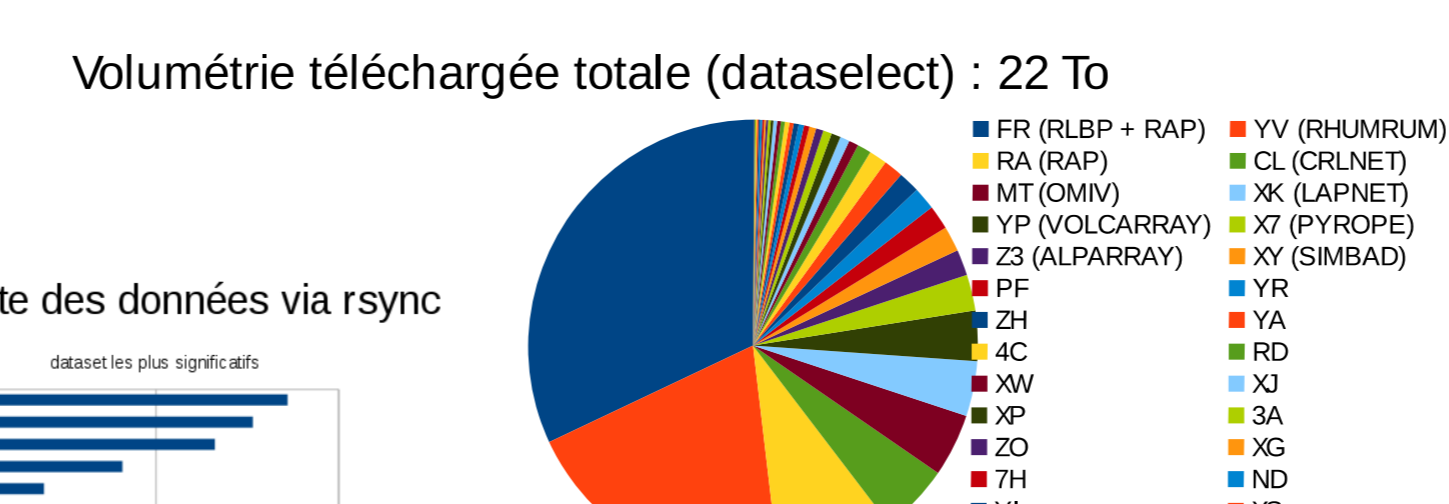
Figures 8 : information sur les données (octobre 2017)

Desserte des données

Volumétrie desservie, par moteur de requêtes



Volumétrie téléchargée par réseau (datasetselect)



Figures 9 : informations sur la desserte des données (octobre 2017)

Services d'accès aux données

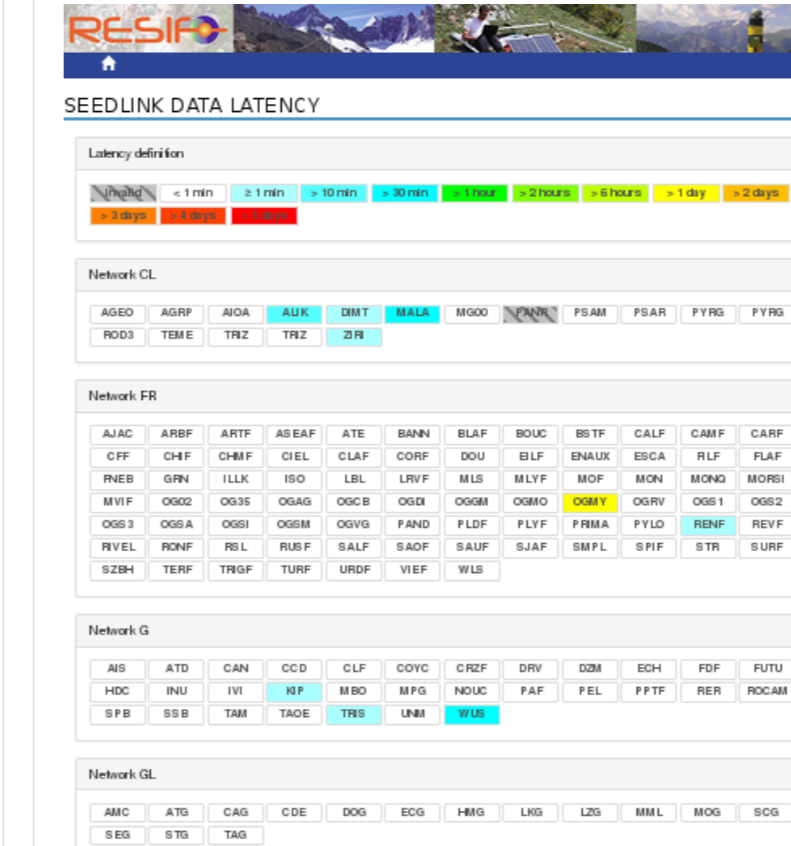


Figure 10. Latences en sortie pour les données temps réel

Portail d'accès aux données <http://seismology.resif.fr>

Accéder aux données en temps quasi réel
Connexion au serveur rtserve.resif.fr (port 18000)

Accéder aux données validées (continues ou déclanchées, publiques ou privées) : utiliser les WS station & dataset <http://ws.resif.fr>

European Integrated Data Archive <http://www.orfeus-eu.org/eida/>

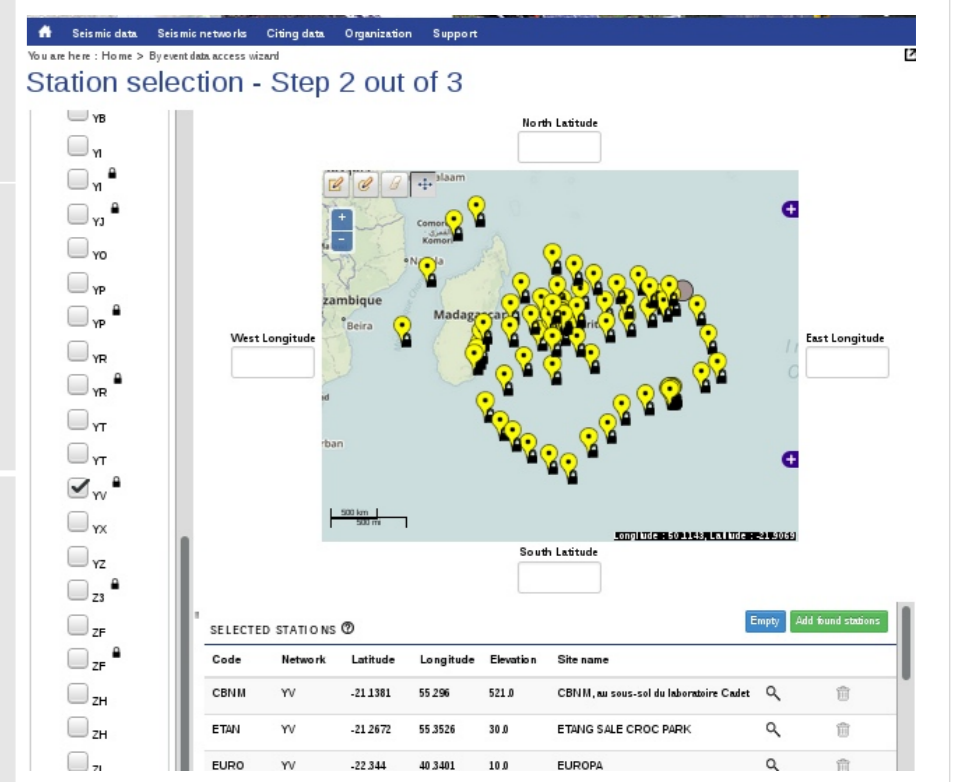


Figure 11. Accès aux données par événements sur le portail RESIF. Recherche de données pour les stations du réseau YV (RHUMRUM).

Nouveau service & accès aux données RESIF par les services fédérés

WFCatalog

Accéder aux indicateurs de qualité des données
utiliser le webservice EIDA Wfcatalog <http://ws.resif.fr/wfcatalog/alpha>

Le webservice WFCatalog (WaveForm Catalog) fournit des informations sur le contenu des formes d'ondes et des indicateurs de qualités des données (au format JSON). Le WFCatalog est un standard EIDA, décrit sur le site WEB ORFEUS. Une homogénéisation des informations fournies par WFCatalog avec les informations fournies par MUSTANG (IRIS-DC) est en cours par la FDSN.

Accéder aux données RESIF en utilisant les webservices fédérés (systèmes en cours de développement ou de normalisation)

European Integrated Data Archive Federator (en cours de développement).

Fetch script : https://www.seiscomp3.org/wiki/doc/applications/fdsnws_fetch.
`fdsnws_fetch.py -v --network=G,RD,CH,IV --channel=BHZ --starttime=2016-01-01T00:00:00 --endtime=2016-01-01T00:00:30 -o mydata.mseed`

Basé sur les tables de routage déclarées par les noeuds EIDA pour localiser les ressources et les données : http://ws.resif.fr/eida_routing.xml

IRIS DMC Fedcatalog

webservice fetch : <https://seiscode.iris.washington.edu/projects/ws-fetch-scripts/wiki>
<https://service.iris.edu/irisws/fedcatalog/1/query?cha=HZ&start=2008-09-01&end=2009-01-01>

Evolution des services (non exhaustif)

	Description du service	Calendrier	Commentaire
EIDA-NG	Authentification fédérée	2017.Q4 .. 2018.Q2	Le mécanisme d'authentification permettra l'utilisation des services fédérés de EIDA, y compris pour les données en accès privé (ex : AlpArray).
FDSN	Service Timeseriesplot	Spécification en cours À partir de 2018.Q3	Affichage graphique des formes d'ondes Voir : https://service.iris.edu/irisws/timeseriesplot/1/
	Service Availability	Spécification en cours À partir de 2018.Q3	Présentation (forme tabulaire) de la disponibilité des données Voir : http://service.iris.edu/irisws/availability/1/
	Service Timeseries	Spécification en cours À partir de 2018.Q3	Traitement à la volée des formes d'onde (déconvolution, déconvolution, conversion de format, ..). Voir : https://service.iris.edu/irisws/timeseries/1/
RESIF	Service Transaction	En cours	Accès aux compte-rendu des transactions d'intégration. A usage des noeuds A
	Service Orphanfiles	En cours	Accès aux données orphelines (sans métadonnées). A usage des noeuds A
	Service Deletefiles	Spécification par RESIF-SI à faire	Possibilité de retirer des données de l'archive. A usage des noeuds A
	Diverses statistiques d'accès, tables ou plots	Spécification par RESIF-SI à faire	Obtenir des informations sur le téléchargement de données, par réseau, station, channel. A usage des réseaux
	Métriques Densification des indicateurs de fonctionnement	ASAP	Aide à la gestion des évolutions des infras & middleware du noeud B. A usage du noeud B.

Figure 12. Quelques évolutions prévues pour le centre de données RESIF