



HAL
open science

Rapport annuel du centre de données sismologiques national Résif-DC pour l'année 2022

Jonathan Schaeffer, Helle A. Pedersen

► **To cite this version:**

Jonathan Schaeffer, Helle A. Pedersen. Rapport annuel du centre de données sismologiques national Résif-DC pour l'année 2022. INSU, CNRS. 2023. hal-03994791v3

HAL Id: hal-03994791

<https://hal.science/hal-03994791v3>

Submitted on 30 Aug 2023

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.



Distributed under a Creative Commons Attribution 4.0 International License

Rapport annuel des services de Résif-DC

Jonathan Schaeffer 

Helle Pedersen 

Équipe Résif-DC <sismo-help@resif.fr>

24 août 2023

Table des matières

1	Résumé	2
2	Disponibilité des services	3
2.1	Indicateurs	3
3	Hébergement des données	4
4	Services de distribution de donnée	6
4.1	Obtenir le flux des données temps réel seedlink	6
4.2	Obtenir des données fdsnws-dataselect	7
4.3	Obtenir des métadonnées fdsnws-station	11
4.4	Récupérer des jeux de données par rsync	12
4.5	Connaître la disponibilité fdsnws-availability	14
4.6	Métriques qualité EIDA eidaws-wfcatalog	15
4.7	Consulter des séries temporelles prétraitées timeseries et timeseriesplot	16
4.8	Consulter la qualité des données avec des PSD	17
4.9	Obtenir des données d'événement pré-assemblés assembleddata	17
4.10	Obtenir des données Large-N au format miniSEED ph5-dataselect	19
4.11	Consulter la disponibilité des données Large-N ph5-availability	20
4.12	Autres webservice	21
5	Services d'intégration et de gestion des données	22
5.1	Création de nouveaux réseaux temporaires	22
5.2	Soumettre de la donnée et de la métadonnée validées ResifDataTransfer	23
5.3	Connaître l'état d'une transaction d'intégration (ws transaction)	25
5.4	Lister les fichiers orphelins (ws orphanfile)	26
6	Autres services	27
6.1	Gestion des DOI	27
6.2	Assistance aux utilisateurs	28
6.3	Portail web de RESIF-DC seismology.resif.fr	29
6.4	Sauvegarde de données scientifiques	29
6.5	Consulter les statistiques sur les données hébergées	30
7	Impact environnemental	31
7.1	Infrastructure	31
7.2	Transfert de données	33
7.3	Transports	33

Ce document est à usage interne du centre de données Résif-DC. Cependant les informations et figures ne sont pas confidentielles et peuvent être réutilisées sous les termes de la licence Creative Commons Attribution 4.0.

1 Résumé

Les indicateurs pour l'année 2022 montrent une excellente disponibilité des services critiques.

La quantité de données hébergée continue de croître avec une tendance linéaire.

La tendance d'usage des principaux webservice se confirme, `fdsnws-station` a connu une forte augmentation en nombre de requêtes. Le nouveau webservice `resifws-seedpsd` est référencé dans ce document, mais sa mise en production a été réalisée au mois d'octobre 2022.

2 Disponibilité des services

2.1 Indicateurs

Disponibilité des services d'accès aux données/métadonnées.

En 2022, nous avons planifié un arrêt au mois de décembre en minimisant l'interruption des services. Le taux de disponibilité est mesuré par un service externe (uptime robot) et montre une excellente disponibilité globale des services.

Table 1 – taux de disponibilité des services pour l'année 2022

service	disponibilité	commentaire
ringserver	99.992	Coupure de 22 minutes en juin
fdsnws-station	99.913	3 coupures dont la plus longue a duré 2h
fdsnws-dataselect	99.961	6 coupures dont 5 de moins de 10 minutes
fdsnws-availability	99.888	
fdsnws-event	99.963	Service hébergé au RENASS

— 0.05% équivaut environ à 4h d'arrêt

— 1% équivaut à 3.5 jours d'arrêt

3 Hébergement des données

La quantité de données hébergées pour chaque année. On fait apparaître les réseaux temporaires et permanents ainsi que la donnée brute et validée.

En janvier 2023, le centre de données :

- héberge 92.0 téraoctets.
- distribue de la données pour 13 réseaux permanents, 79 réseaux temporaires.
- référence 3967 stations et 21941 canaux.

Table 2 – Données hébergées pour chaque année de la donnée. Exemple : Nous hébergeons 2829 GB de donnée brute concernant l'année 2020.

year	raw (GB)	valid. perm. (GB)	valid. temp. (GB)	perm. network	temp. network
2003	0	60	184	3	5
2004	0	77	148	3	5
2005	0	126	74	3	3
2006	0	153	48	4	2
2007	0	180	603	4	6
2008	0	192	1417	5	10
2009	0	384	533	5	11
2010	0	871	888	6	9
2011	0	2060	1437	9	10
2012	14	2597	2006	11	8
2013	13	3729	2257	11	6
2014	218	4565	2025	11	10
2015	224	4984	1137	11	12
2016	183	5435	1432	11	12
2017	94	5777	2141	11	11
2018	710	5541	6122	11	17
2019	2023	5147	2287	11	18
2020	1113	5467	1076	11	15
2021	1458	5403	1858	11	17
2022	1513	4854	859	13	14

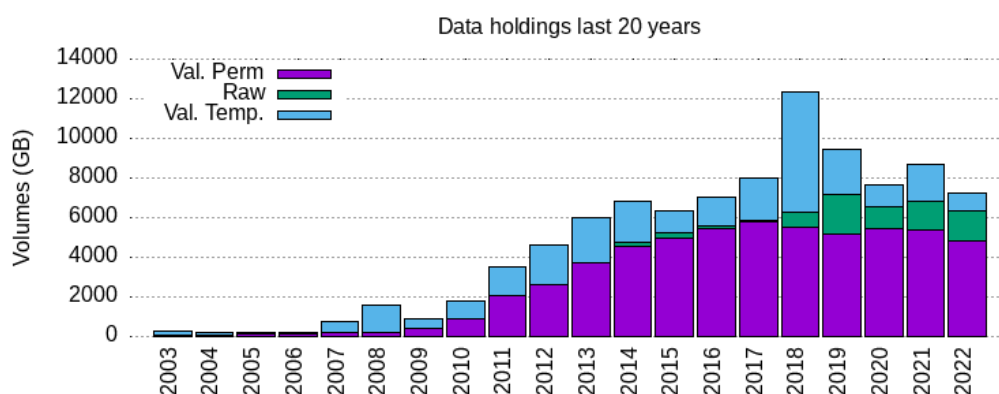


Figure 1 – Hébergement des données regroupées par année. Représentation graphique tu tableau 2

Table 3 – Volumétrie cumulée par année des données hébergées

year	cumulated raw (GB)	validated permanent (GB)	validated temporary (GB)
2003	0	341	435
2004	0	418	583
2005	0	544	657
2006	0	698	705
2007	0	878	1308
2008	0	1070	2725
2009	0	1454	3256
2010	0	2325	4140
2011	0	4385	5577
2012	14	6982	7583
2013	25	10711	9840
2014	243	15275	11881
2015	467	20260	13018
2016	649	25695	14449
2017	744	31472	16590
2018	1443	37036	22712
2019	2921	42598	25008
2020	3845	48163	26088
2021	5283	53573	27976
2022	6668	58933	29503

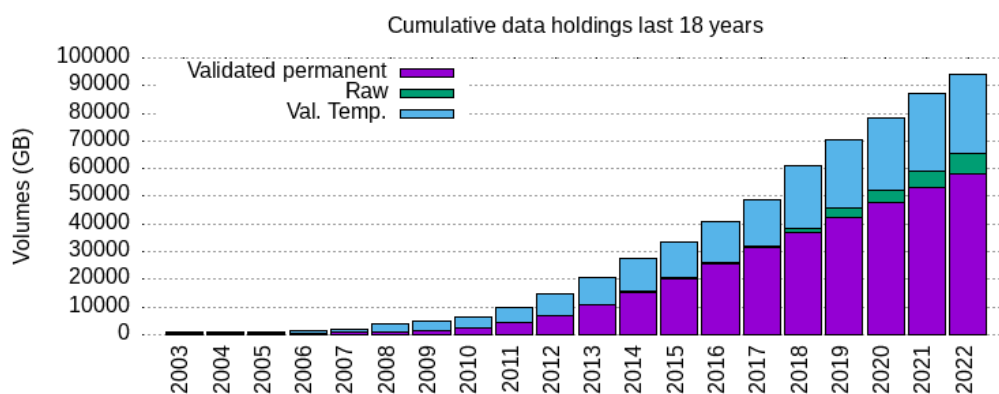


Figure 2 – Volumétrie cumulée par année de toutes les données hébergées. Graph du tableau 3

4 Services de distribution de donnée

4.1 Obtenir le flux des données temps réel seedlink

Le service temps réel permet de récupérer un flux de données au format miniSEED. Documentation utilisateur : <https://seismology.resif.fr/real-time-seedlink/>

Indicateurs

Table 4 – Nombre d'IP unique et volume de données livrées par année

year	unique IP	volume (TB)
2013	50	4.7
2014	175	23.8
2015	279	35.6
2016	340	40.3
2017	367	40.5
2018	479	36.3
2019	1591	27.4
2020	1234	30.7
2021	867	27.8
2022	634	38.0

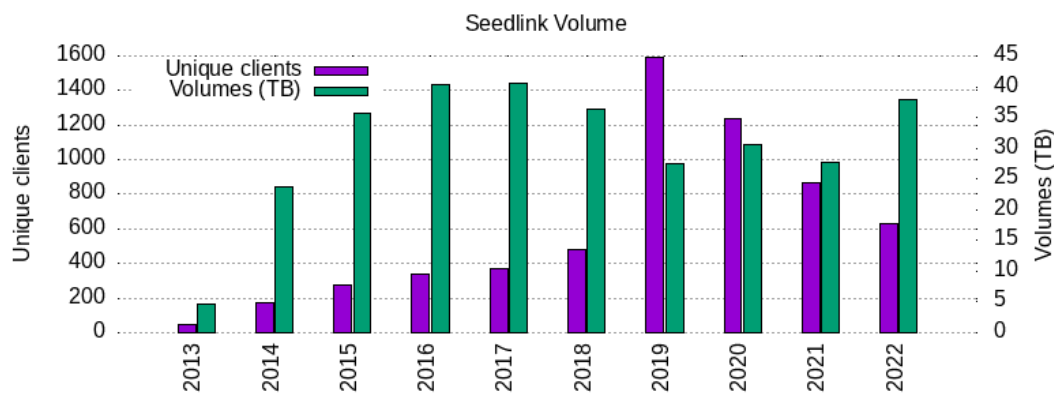


Figure 3 – Nombre d'IP unique et volume de données livrées par année. Graph dur tableau 4

4.2 Obtenir des données fdsnws-dataselect

Le webservice FDSN dataselect permet d'obtenir les données de manière standard, au format miniSEED. Documentation utilisateur : <http://ws.resif.fr/fdsnws/dataselect/1>

Indicateurs

Table 5 – Décompte des requêtes au webservice dataselect en million. “total requests” = “ok requests” + “no data” + “errors”; “ok requests” = “ok query methods” + “ok other methods” “errors” : Tous les codes retours de type 400 ou 500

year	total req	ok req	no data	bad req	internal errors	ok query	ok other
2019	28.98	15.44	8.70	4.74	0.10	12.90	2.54
2020	80.70	47.05	17.06	16.49	0.10	33.72	13.34
2021	78.58	38.68	22.19	16.68	1.02	26.25	12.43
2022	78.35	39.13	28.46	10.36	0.39	33.46	5.67

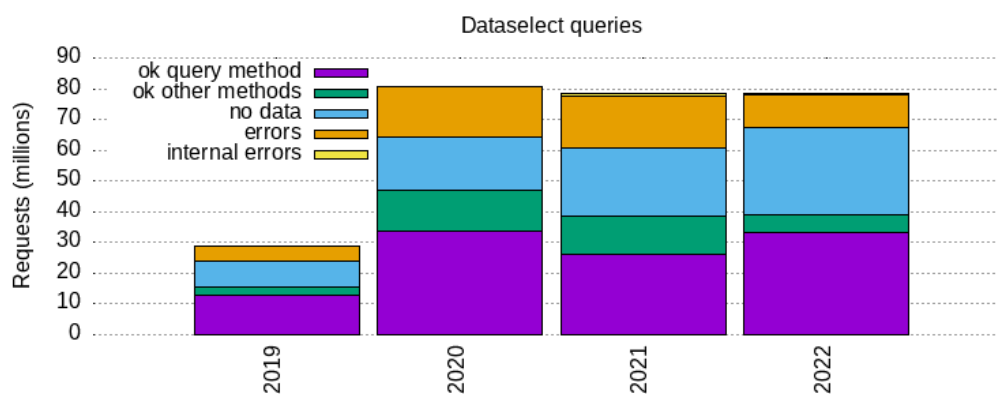


Figure 4 – Graph du tableau 5.

Table 6 – Nombre de requêtes réussies (millions) sur la donnée, nombre d'IP uniques et volumétrie en téraoctets. Ces chiffres concernent uniquement les requêtes réussies.

year	requests (millions)	unique IP	volume (TB)
2013	0.00	6	0.0
2014	0.86	533	8.5
2015	3.01	603	7.7
2016	4.42	322	5.4
2017	8.93	638	14.2
2018	21.33	1238	50.1
2019	12.90	1595	26.0
2020	33.72	2645	51.3
2021	26.25	1448	85.5
2022	33.46	3887	114.4

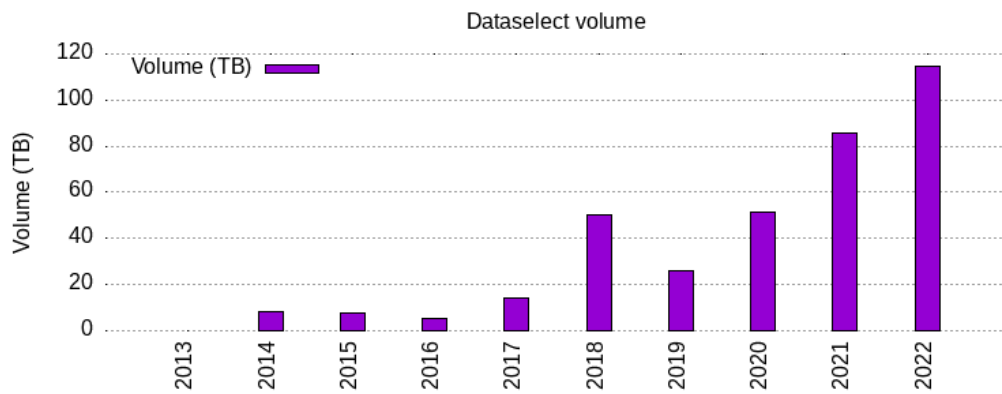


Figure 5 – Volume de données livrées par protocole dataselect. Graph du tableau 6.

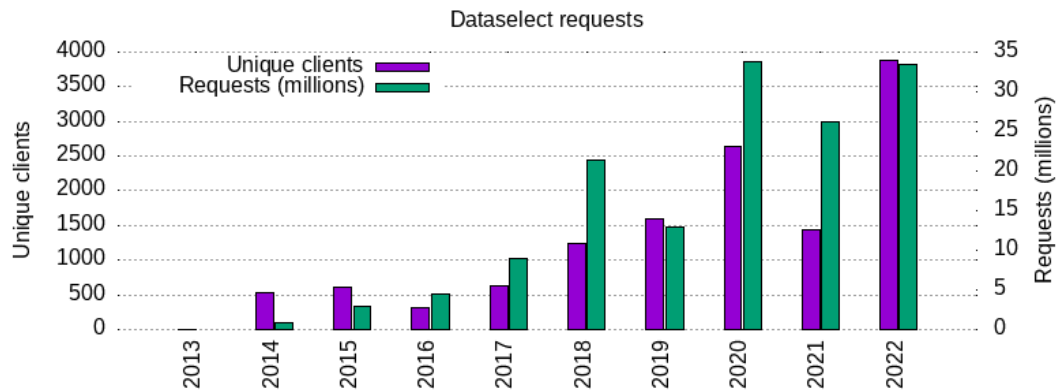


Figure 6 – Nombre de clients uniques et nombre de requêtes. Graph du tableau 6.

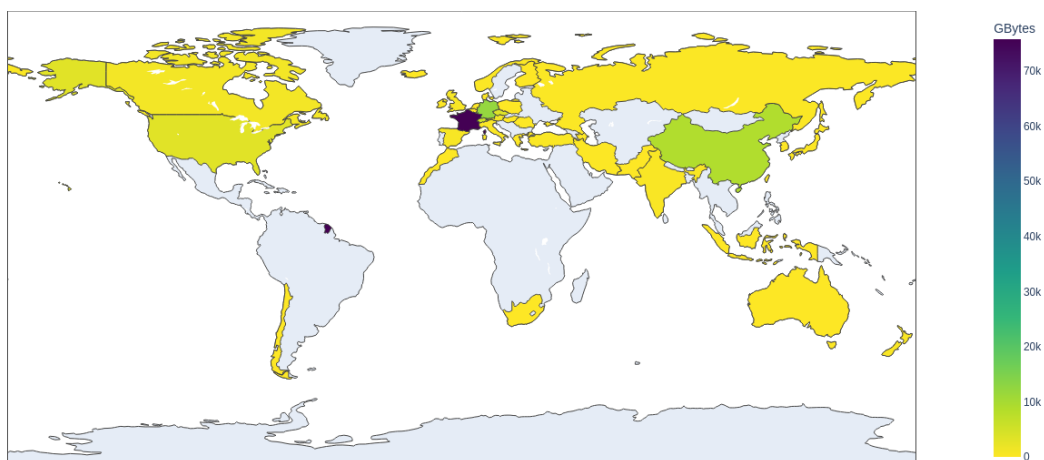


Figure 7 – Volume de données distribuées dans chaque pays pour l'année 2022

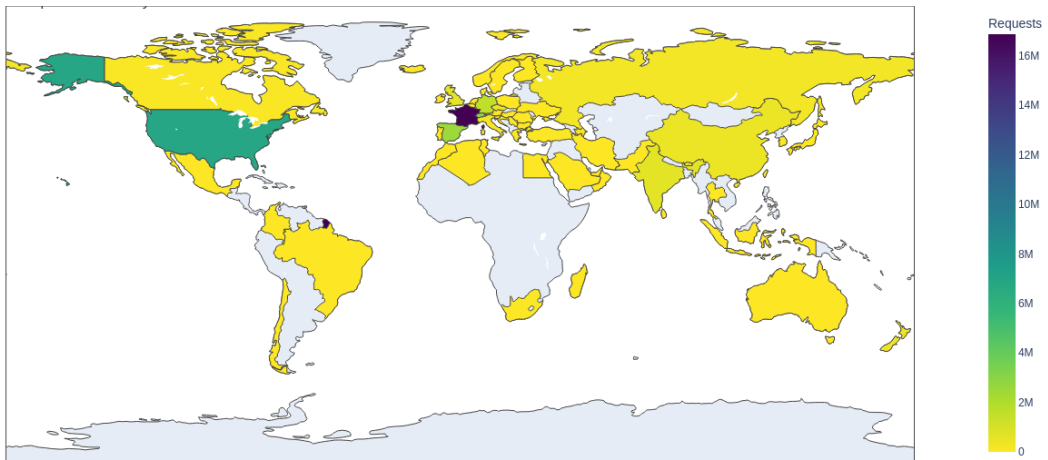


Figure 8 – Origine des requêtes par pays pour l'année 2022

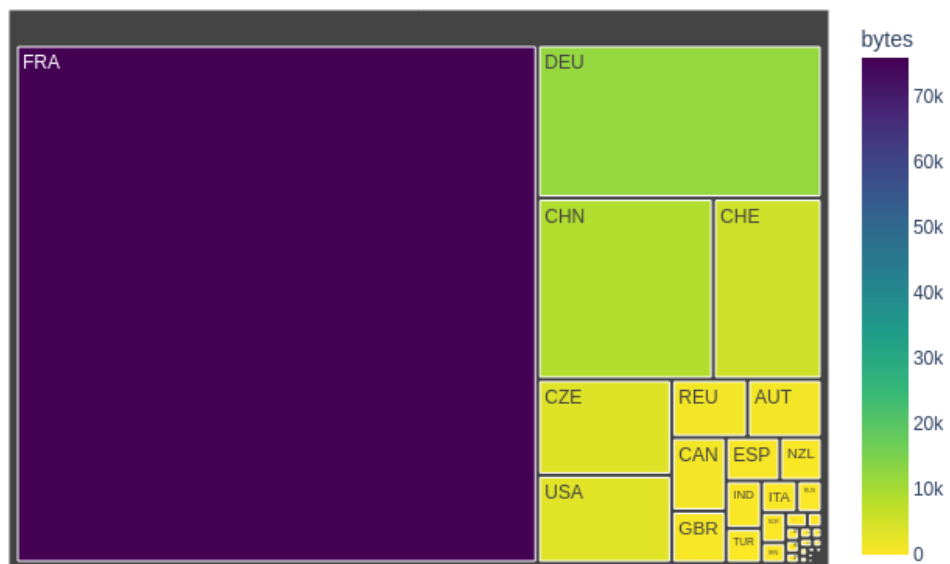


Figure 9 – Répartition de la quantité de données livrées en 2022 par pays



Figure 10 – Répartition des requêtes dataselect en 2022 par pays

4.3 Obtenir des métadonnées fdsnws-station

Le webservice station permet aux usagers de consulter les métadonnées. Documentation : <http://ws.resif.fr/fdsnws/station/1>

Indicateurs

Table 7 – Nombre de requêtes réussies (millions), nombre de clients uniques, volume des données envoyées

year	all requests (millions)	ok requests	unique IP	volume (GO)
2019	11.2	9.0	2619	1222.4
2020	21.9	16.0	5502	600.0
2021	43.2	36.0	7888	1161.1
2022	89.2	81.0	10605	2880.6

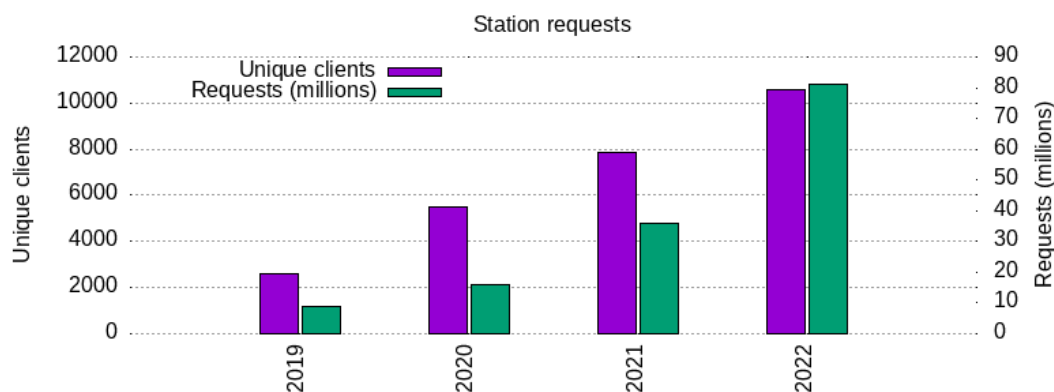


Figure 11 – Représentation graphique du tableau 7

4.4 Récupérer des jeux de données par rsync

Données récupérées par le portail rsync concernant des réseaux.

Indicateurs

Table 8 – Données envoyées par protocole rsync, nombre de requêtes et nombre de clients uniques.

year	Sent (GB)	requests	clients
2015	0	1	1
2016	200	12	2
2017	368	47585	14
2018	3736	166562	23
2019	5684	175245	49
2020	1383	19034	19
2021	8	1987	4
2022	1485	2052	12

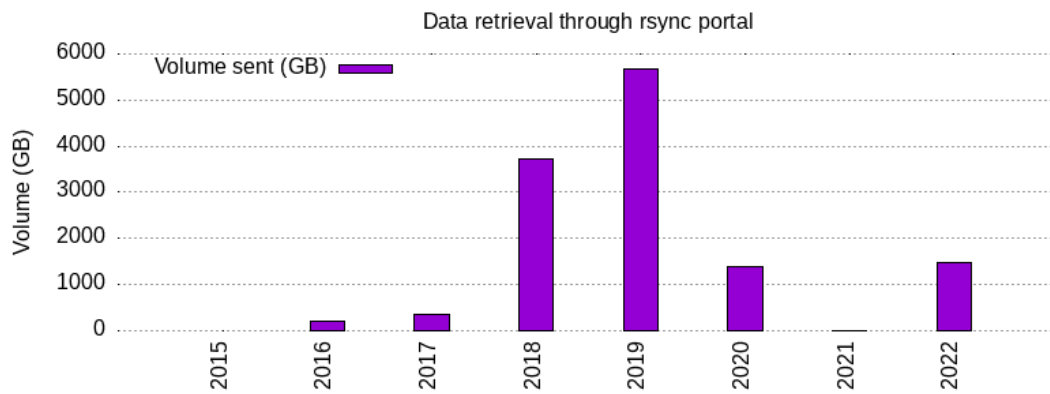


Figure 12 – Volumes envoyés. Représentation de la première colonne du tableau 8

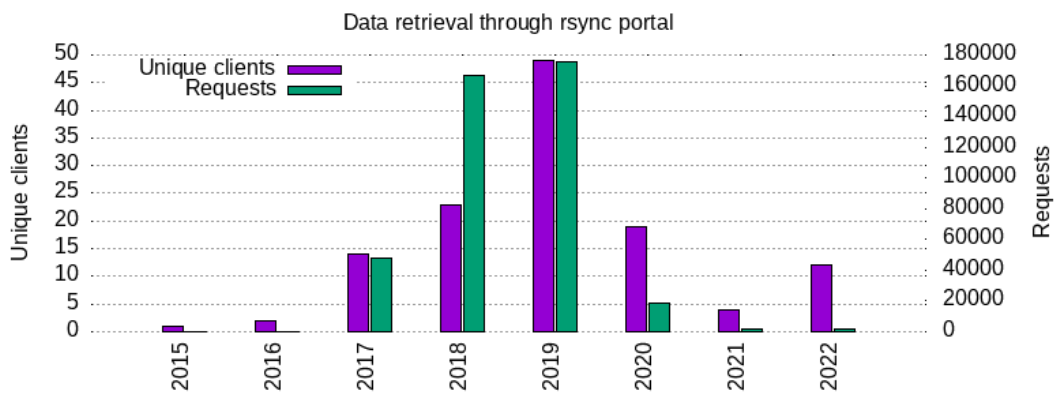


Figure 13 – Nombre de clients et de requêtes. Représentation du tableau 8

4.5 Connaître la disponibilité fdsnws-availability

Webservice permettant de consulter la disponibilité des données. Documentation : <http://ws.resif.fr/fdsnws/availability/1>

Indicateurs

Table 9 – Nombre de requêtes et d'IP uniques pour le webservice availability.

year	requests	unique IP
2019	2057	125
2020	8832	291
2021	44133	1288
2022	69846	2267

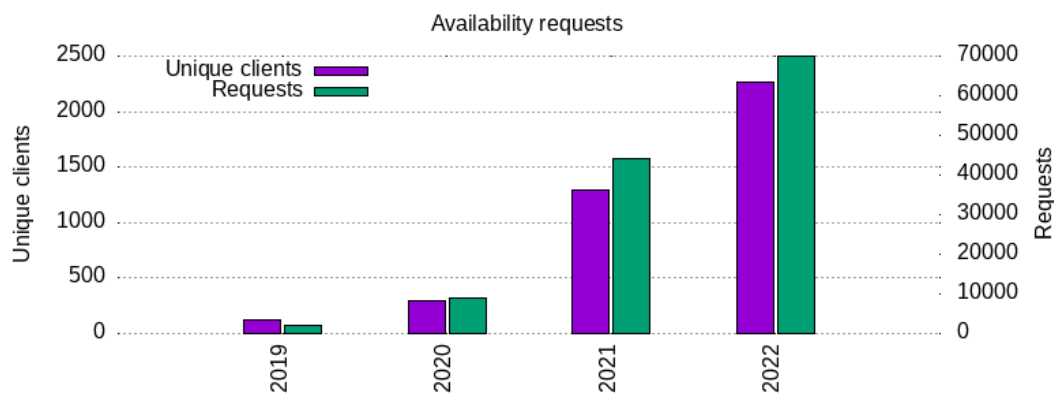


Figure 14 – Représentation du tableau 9

4.6 Métriques qualité EIDA eidaws-wfcatalog

Indicateurs

Table 10 – Nombre de requêtes et d'IP uniques pour le webservice wfcatalog.

year	requests	unique IP
2019	60351	82
2020	213785	379
2021	487573	520
2022	170726	2060

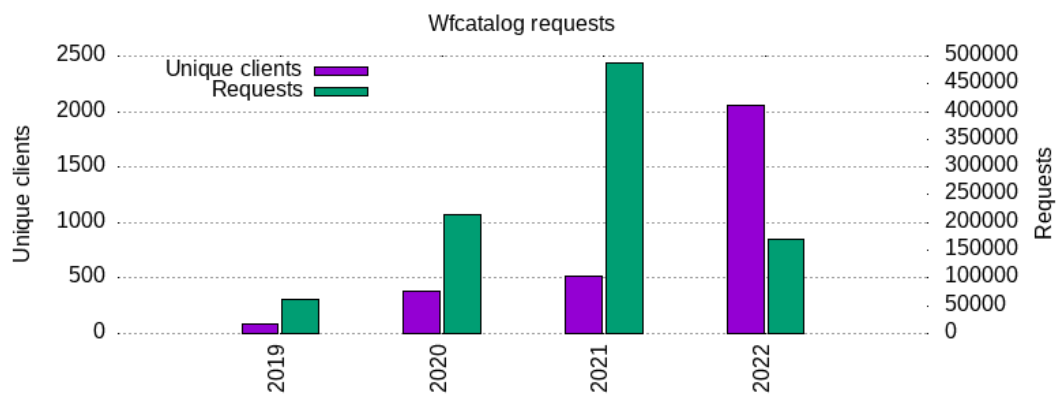


Figure 15 – Représentation du tableau 10

4.7 Consulter des séries temporelles prétraitées timeseries et timeseriesplot

Ce service permet de consulter les séries temporelles sous différents formats, en particulier sous forme graphique. Documentation : <https://ws.resif.fr/resifws/timeseries/> et <http://ws.resif.fr/resifws/timeseriesplot>

Indicateurs

Table 11 – Nombre de requêtes et IP uniques sur les webservices timeseries et timeseriesplot.

Year	Requests	unique IP
2019	24681	929
2020	42936	901
2021	8052413	1314
2022	11127434	1497



Figure 16 – Représentation du tableau 11

4.8 Consulter la qualité des données avec des PSD

Ce service a été ouvert aux utilisateurs en octobre 2022 : <https://ws.resif.fr/resifws/seedpsd/1/>

indicateurs

Table 12 – Nombre de requêtes et IP uniques sur le webservice seedpsd

Year	Requests	unique IP
2022	288	56

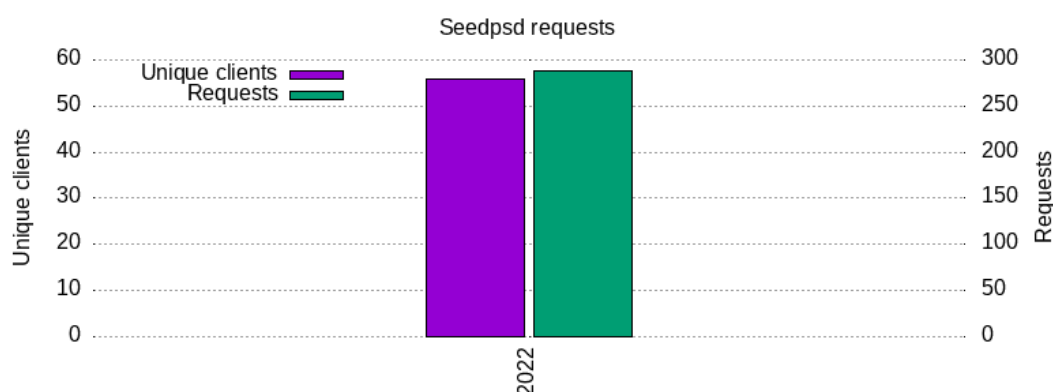


Figure 17 – Représentation graphique du tableau 12

4.9 Obtenir des données d'événement pré-assemblés assembleddata

Permet d'obtenir des données d'événements pré-assemblés à partir des données du RAP. Documentation : <http://ws.resif.fr/resifsi/assembleddata/1>

Indicateurs

Table 13 – Nombre de requêtes et IP uniques sur le webservice assembleddata

Année	Requêtes	unique IP	volume (GO)
2019	56	26	0.3
2020	331	120	0.1
2021	66	18	0.1
2022	1496	19	1.2

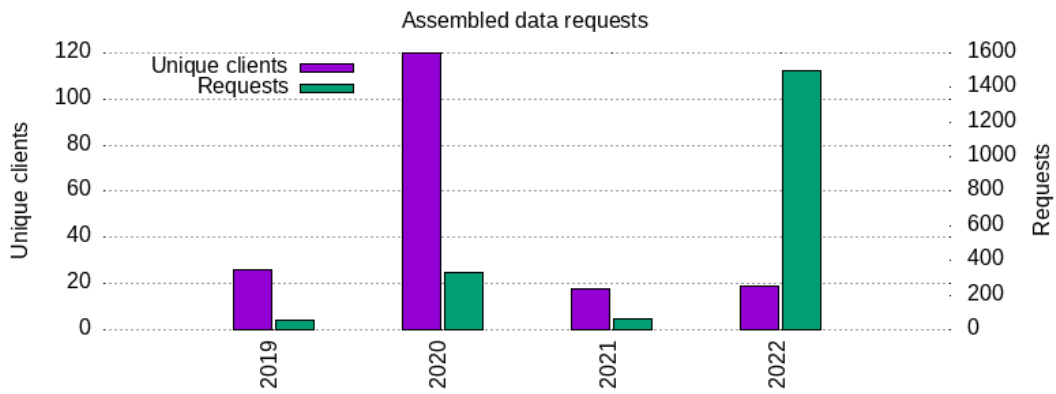


Figure 18 – Représentation du tableau 13

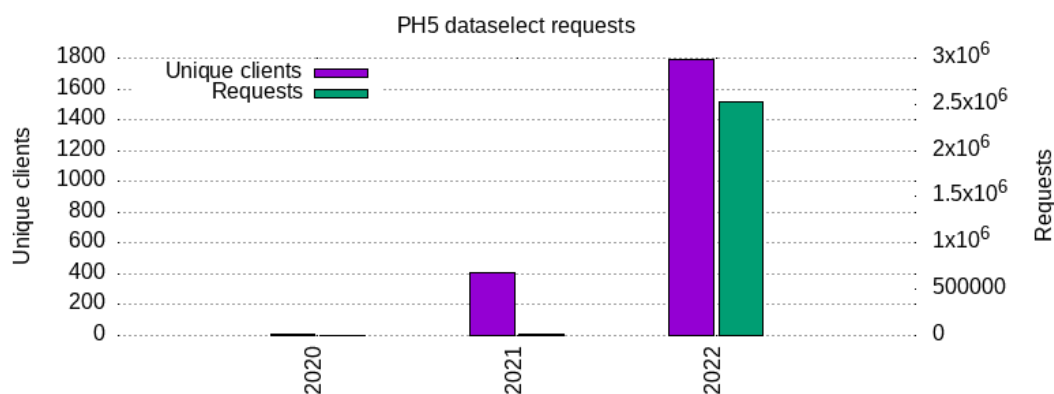
4.10 Obtenir des données Large-N au format miniSEED ph5-dataselect

Ce service permet de distribuer en miniSEED la donnée stockée au format PH5. Documentation : <http://ws-ph5.resif.fr/fdsnws/ph5-dataselect/1>

Note: le service est déclaré dans obspy depuis sa version 1.3.0 en mars 2022.

Indicateurs

Année	requêtes	unique IP	volume (GO)
2020	6	5	0.0
2021	17807	412	48.1
2022	2525270	1794	37.6



4.11 Consulter la disponibilité des données Large-N ph5-availability

Documentation : <https://ws.resif.fr/resifws/ph5-availability/1/>

Indicateurs

Table 14 – requêtes au webservice ph5-availability

Année	requêtes	unique IP	volume (GO)
2020	9	7	0.0
2021	1082	298	0.0
2022	6356	546	0.0

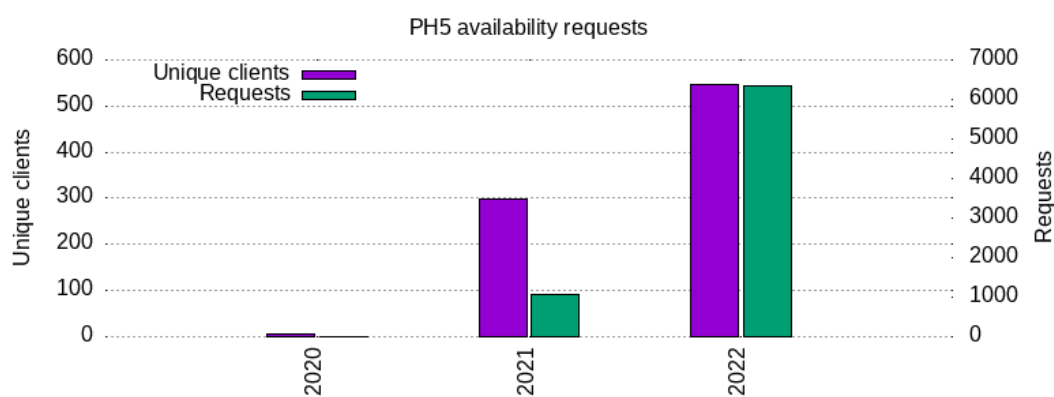


Figure 19 – Graph du tableau 14

4.12 Autres webservices

- resp : <http://ws.resif.fr/resifws/resp/1>
- sacpz : <http://ws.resif.fr/resifws/sacpz/1>
- evalresp : <http://ws/resif.fr/resifws/evalresp/1>

Indicateurs

Table 15 – Indicateurs pour le webservice resp

Année	requêtes	unique IP
2020	69103	30
2021	14625	409
2022	23536	605

Table 16 – Indicateurs pour le webservice sacpz

Année	requêtes	unique IP
2020	3	3
2021	1738	384
2022	5972	546

Table 17 – Indicateurs pour le webservice evalresp

Année	requêtes	unique IP
2020	3	3
2021	2423	460
2022	4749	744

5 Services d'intégration et de gestion des données

5.1 Création de nouveaux réseaux temporaires

Nombre de nouveaux réseaux temporaires ajoutés au centre de données chaque année.

Indicateurs

Le tableau 18 montre la répartition des réseaux temporaire par date de début déclaré à la FDSN. La répartition par année n'est pas exacte car elle ne correspond pas à sa date de création au centre de données.

Table 18 – Nombre de réseaux temporaires déclarés par année

year	count
1998	1
1999	1
2000	1
2001	3
2003	4
2004	2
2007	5
2008	5
2009	5
2010	4
2011	2
2012	1
2013	1
2014	5
2015	6
2016	4
2017	3
2018	12
2019	7
2020	5
2021	6
2022	3

5.2 Soumettre de la donnée et de la métadonnée validées ResifDataTransfer

Ce service permet aux producteurs de donnée et de métadonnée de soumettre leurs produits au centre de données. Documentation : https://gitlab.com/resif/resif-data-transfer/-/blob/master/MANUAL_fr.md

Indicateurs

Les tableaux 19, 20, 21 montrent les historiques d'indicateurs pour ce service.
Les figures 20, 21 et 22 reprennent les chiffres sous forme de graphs.

Table 19 – indicateurs pour les transactions de données miniSEED

Année	transactions mSEED	volume (GO)
2014	700	11638
2015	1592	11168
2016	2186	13969
2017	2601	12955
2018	1467	17001
2019	1966	17951
2020	1333	7730
2021	1976	10264
2022	3146	14680

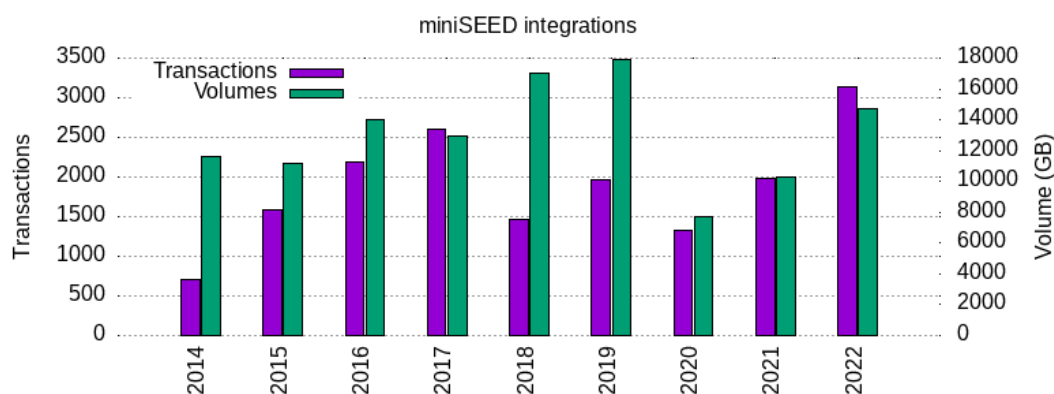


Figure 20 – Transactions miniSEED, selon le tableau 19

Table 20 – indicateurs pour les transactions de données PH5

Année	transactions PH5	volume (GO)
2020	10	3549
2021	2	424
2022	3	507

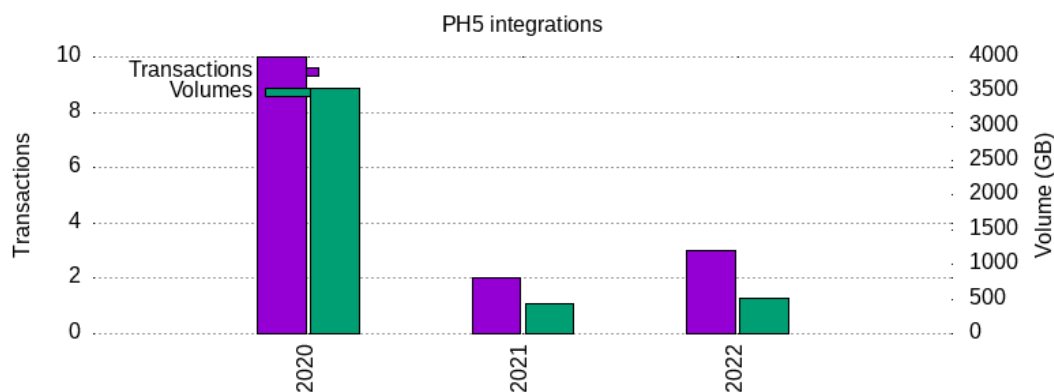


Figure 21 – Transactions PH5, selon le tableau 20

Table 21 – indicateurs d'intégration de métadonnées (stationXML et dataless)

Année	transactions métadonnée
2015	975
2016	1550
2017	918
2018	2161
2019	2698
2020	2824
2021	1067
2022	1058

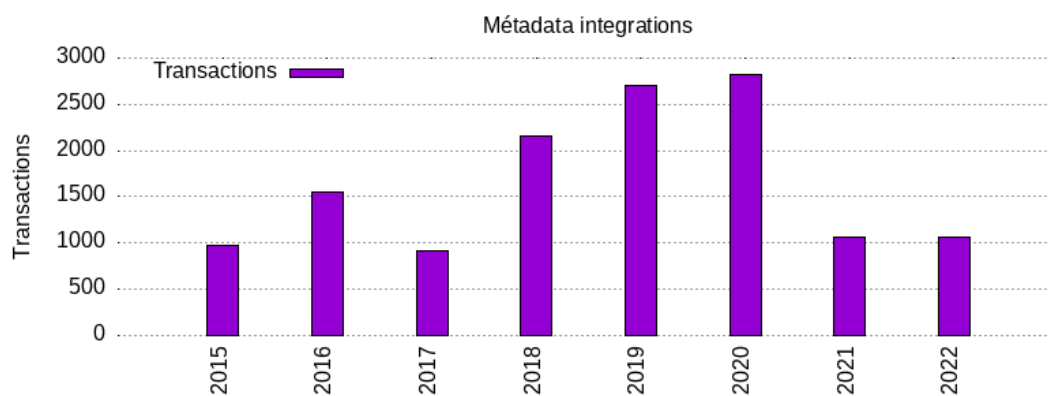


Figure 22 – Transactions de métadonnées, selon le tableau 21

5.3 Connaître l'état d'une transaction d'intégration (ws transaction)

L'état d'une transaction peut être consulté :

- par requête sur le module rsync TRANSACTION_XML
- par requête au webservice transaction.

Indicateurs

Table 22 – Nombre de téléchargement sur le module rsync TRANSACTION_XML

Year	requests	clients
2022	19318	6

Table 23 – Nombre d'accès au webservice transaction

Année	requêtes	unique IP
2019	88	11
2020	695	70
2021	1070	219
2022	768	192

5.4 Lister les fichiers orphelins (ws orphanfile)

Les fichiers orphelins ne sont décrits par aucune métadonnée. En conséquence de quoi, ils ne peuvent être distribués par le centre de données. Ce service permet aux producteurs de données et de métadonnées d'obtenir une liste des fichiers orphelins.

Documentation : <http://ws.resif.fr/resifsi/orphanfile/1>

Indicateurs

Table 24 – Nombre d'accès au webservice orphanfile

Année	requêtes	unique IP
2019	30	7
2020	157	22

6 Autres services

6.1 Gestion des DOI

Résif-DC assure un service d'enregistrement et de maintenance des DOI pour les réseaux sismologiques.

Documentation : Pour enregistrer un DOI, il faut en faire la demande par ticket mailto: sismo-help@resif.fr

Indicateurs

Table 25 – Nombre de DOI créés par an

year	DOIs
2022	6
2021	10
2020	5
2019	4
2018	26
2017	2
2016	5
2015	8
2014	7

6.2 Assistance aux utilisateurs

Plusieurs interfaces d'assistance existent en fonction du périmètres et du statut de l'utilisateur

— Usersupport EIDA

— Helpdesk Résif-DC (par mail <mailto:sismo-help@resif.fr>). Ce helpdesk a migré en juin 2021 vers la plateforme <https://gitlab.com/resif/sismo-help>

Les 164 demandes se répartissent en : 102 pour GLPI et 62 pour gitlab

Indicateurs

Le tableau 26 reprend l'historique des demandes soumises au helpdesk.

Table 26 – Nombre de demandes extérieures par an sur le helpdesk UGA

Year	Tickets
2021	164
2020	138
2019	139
2018	140

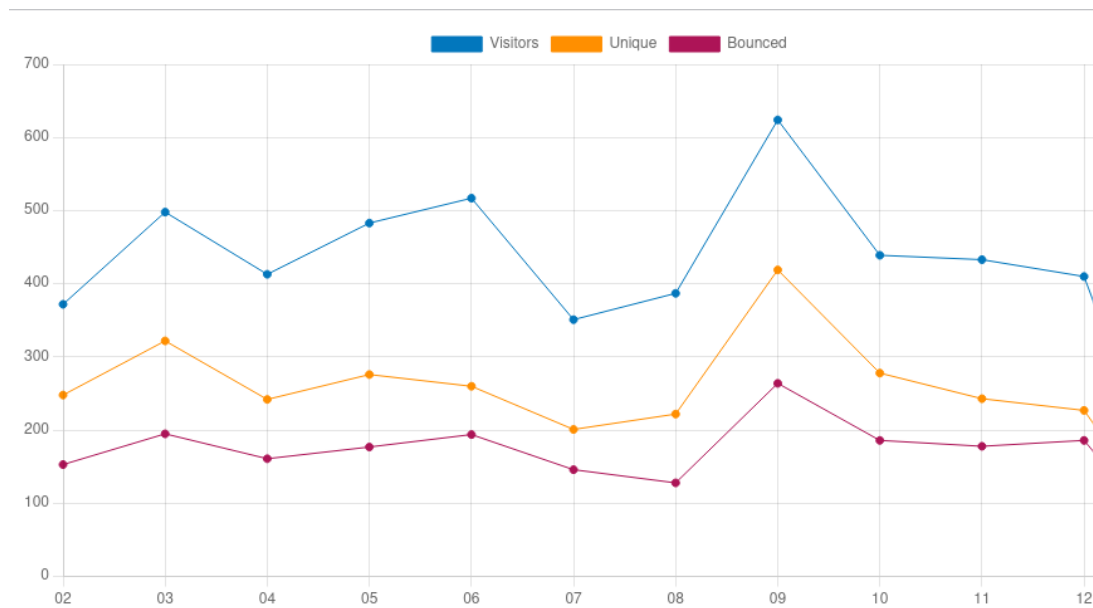
Le tableau 27 reprend l'historique des demandes soumises au nouveau servicedesk :

Table 27 – Demandes par année sur le servicedesk gitlab.com

Year	Tickets
2022	89

6.3 Portail web de RESIF-DC seismology.resif.fr

Le portail web permet d'accéder à de la documentation, des informations riches et des produits dérivés calculés au centre de données. <https://seismology.resif.fr>



Indicateurs

Consultations annuelles du site web

6.4 Sauvegarde de données scientifiques

Certaines données, en particulier issues du RENAG sont sauvegardées au centre de données.

Indicateur

Volumétrie sauvegardée : 15To

6.5 Consulter les statistiques sur les données hébergées

Le webservice statistiques permet d'exposer aux utilisateurs des chiffres concernant les volumes de données et le nombre de requêtes. Documentation : <http://ws.resif.fr/resifws/statistics/1>

Indicateurs

Table 28 – Nombre de requêtes et IP uniques sur les webservices statistiques

Year	Requests	unique IP
2020	1456	39
2021	54767	719
2022	23875	770



Figure 23 – Représentation du tableau 28

7 Impact environnemental

Répartition globale :

Table 29 – kg eqCO₂ pour les différentes activités Résif

Année	Serveurs	Serveurs Virt.	Stockage	Distrib. données	Transports	Total
2021	899	302	11370	161	2700	15432
2022	444	420	14136	210	2016	17226

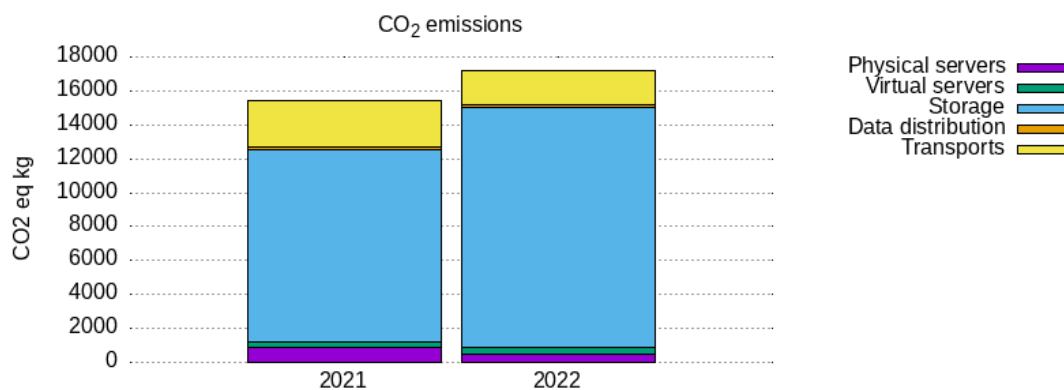


Figure 24 – CO₂ emissions per year.

7.1 Infrastructure

Espaces de stockage

Le stockage SUMMER est évalué à 37,9g de CO₂ équivalent par GigaOctet (selon l'étude EcolInfo/Gricad 2021).

Nous réservons actuellement 200TB pour la partie "Centre de Données" et 100TB pour les nœuds A RAP et SISMOB.

Table 30 – Impact énergétique et environnemental du stockage par année

Année	Volume total (Go)	CO ₂ eq (kg)	Consommation (kWh)
2021	300000	11370.	105.27778
2022	373000	14136.7	130.89537

Serveurs physiques

La consommation électrique moyenne des serveurs physiques est calculée à partir des mesures sur leurs interfaces de contrôle. Cette mesure est fournie par le serveur en grandeur cumulée depuis une date de départ et doit donc être moyennée sur une année.

Nous appliquons un PUE de 1.4 correspondant à la valeur estimée dans l'étude EcolInfo/Gricad.

Serveurs virtuels

WINTER est composé d'un ensemble de serveurs physiques hébergés dans plusieurs data centres au PUE moyen de 1.4. La part de consommation imputée à Résif-DC est évaluée par la proportion de facturation.

Impact carbone de Résif-DC en 2022

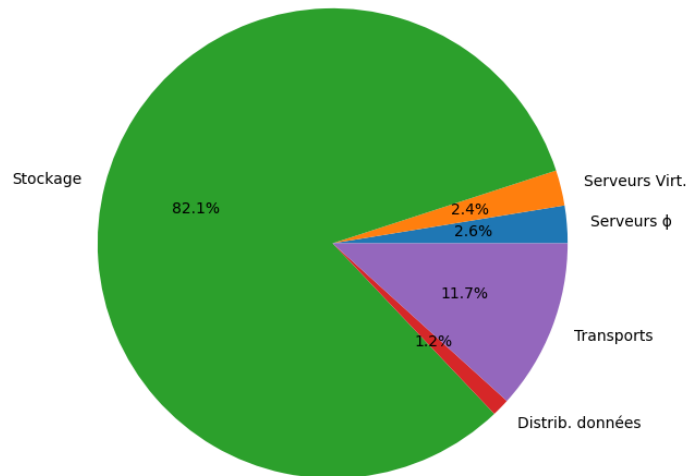
Figure 25 – CO₂ emissions for last year

Table 31 – Consommation moyenne annuelle des serveurs physiques, hébergement compris

Année	kWh	CO ₂ eq
2021	8323	899
2022	4118	444

Année	Nombre de VM	eq CO ₂ (kg)
2021	18	302.4
2022	25	420.

7.2 Transfert de données

Selon l'estimation RENATER/EcoInfo, Nous avons transféré 80To, donc l'impact carbone est de :

Nous utilisons les hypothèses décrites dans <https://ecoinfo.cnrs.fr/2020/12/04/quelle-est-lempreint>
Avec pour mesure : 1.4g eqCO₂ pour 1Go

year	data transfered (Go)	CO ₂ eq (kg)
2021	115000	161.
2022	150000	210.

7.3 Transports

L'équipe des ingénieurs Résif se déplace occasionnellement en réunions. La politique est de prendre le moins possible l'avion

year	Transports for work	Transports for meetings
2021	2644	120
2022	2000	13