

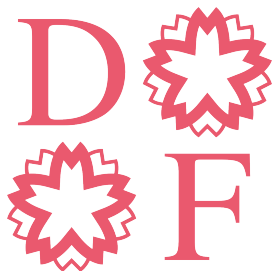
Remettre internet sur la place publique

L'entre-hébergement avec Garage

Vincent Giraud

Pas Sage En Seine 2023

Vendredi 16 juin 2023



<https://deuxfleurs.fr>

Deuxfleurs est une association militant en faveur d'un internet plus convivial, avec une organisation et des rapports de force repensés.

Nous faisons partie du CHATONS¹ depuis avril 2022.



1. Collectif des Hébergeurs Alternatifs, Transparents, Ouverts, Neutres et Solidaires

Les ressources accessibles par internet reposent aujourd'hui essentiellement dans des centres de données.



Cette concentration répond aux logiques industrielles et commerciales solidement implantées dans le domaine du numérique.

Cet état de fait entraîne d'abord un contrôle monumental sur les ressources en question.

2.

Cet état de fait entraîne d'abord un contrôle monumental sur les ressources en question.

Pour les entreprises et groupes privés qui peuvent les valoriser notamment :

- via un ciblage publicitaire
- pour entraîner des intelligences artificielles
- pour maintenir leur position concurrentielle

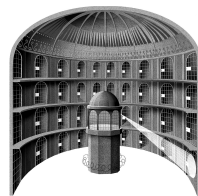
2.

Cet état de fait entraîne d'abord un contrôle monumental sur les ressources en question.

Pour les entreprises et groupes privés qui peuvent les valoriser notamment :

- via un ciblage publicitaire
- pour entraîner des intelligences artificielles
- pour maintenir leur position concurrentielle

Pour les États, qui y trouvent des points d'entrée privilégiés où piocher dans le cadre de surveillance de masse²



2. Cf. les révélations d'Edward Snowden, en particulier ici le programme PRISM

Dans l'industrie, la part des ressources humaines interagissant avec les centres de données est minimisée autant que possible.

Mais l'invisibilisation des ressources matérielles touche d'abord la société publique³ :

3. Sondage réalisé par Wakefield Research pour Citrix, sur 1000 États-Uniens

Dans l'industrie, la part des ressources humaines interagissant avec les centres de données est minimisée autant que possible.

Mais l'invisibilisation des ressources matérielles touche d'abord la société publique³ :

- 16% des États-Uniens savent réellement ce que signifie le *cloud*

3. Sondage réalisé par Wakefield Research pour Citrix, sur 1000 États-Uniens

Dans l'industrie, la part des ressources humaines interagissant avec les centres de données est minimisée autant que possible.

Mais l'invisibilisation des ressources matérielles touche d'abord la société publique³ :

- 16% des États-Uniens savent réellement ce que signifie le *cloud*
- 22% des États-Uniens font semblant de savoir ce dont il s'agit au quotidien

3. Sondage réalisé par Wakefield Research pour Citrix, sur 1000 États-Uniens

Dans l'industrie, la part des ressources humaines interagissant avec les centres de données est minimisée autant que possible.

Mais l'invisibilisation des ressources matérielles touche d'abord la société publique³ :

- 16% des États-Uniens savent réellement ce que signifie le *cloud*
- 22% des États-Uniens font semblant de savoir ce dont il s'agit au quotidien
- 29% des États-Uniens pensent que le *cloud* utilise des nuages

3. Sondage réalisé par Wakefield Research pour Citrix, sur 1000 États-Uniens

Dans l'industrie, la part des ressources humaines interagissant avec les centres de données est minimisée autant que possible.

Mais l'invisibilisation des ressources matérielles touche d'abord la société publique³ :

- 16% des États-Uniens savent réellement ce que signifie le *cloud*
- 22% des États-Uniens font semblant de savoir ce dont il s'agit au quotidien
- 29% des États-Uniens pensent que le *cloud* utilise des nuages
- 51% des États-Uniens pensent qu'une mauvaise météo influe sur ses performances

3. Sondage réalisé par Wakefield Research pour Citrix, sur 1000 États-Uniens

Dans l'industrie, la part des ressources humaines interagissant avec les centres de données est minimisée autant que possible.

Mais l'invisibilisation des ressources matérielles touche d'abord la société publique³ :

- 16% des États-Uniens savent réellement ce que signifie le *cloud*
- 22% des États-Uniens font semblant de savoir ce dont il s'agit au quotidien
- 29% des États-Uniens pensent que le *cloud* utilise des nuages
- 51% des États-Uniens pensent qu'une mauvaise météo influe sur ses performances

Cette abstraction du matériel derrière le concept d'internet est déterminant dans les comportements de chacun.

3. Sondage réalisé par Wakefield Research pour Citrix, sur 1000 États-Uniens

La mutation d'internet à travers le prisme commercial divise les acteurs en deux :
les vendeurs, et les clients.

La définition industrielle des valeurs rend extrêmement difficile à l'usager de percevoir la structure profonde des moyens sociaux. (...) Il a du mal à imaginer que l'on puisse gagner en rendement social ce que l'on perd en rentabilité industrielle.

Ivan Illich dans *La convivialité*, 1973

La Civilisation des Machines a besoin, sous peine de mort, d'écouler l'énorme production de sa machinerie et elle utilise dans ce but (...) des machines à bourrer le crâne.

Le danger n'est pas dans la multiplication des machines, mais dans le nombre sans cesse croissant d'hommes habitués, dès leur enfance, à ne désirer que ce que les machines peuvent donner.

Georges Bernanos dans *La France contre les robots*, 1945

Récupérer les infrastructures

Tout porte à croire que ramener les fondations d'internet au sein de structures à taille plus humaines et aux organisations plus conviviales aiderait à :

Tout porte à croire que ramener les fondations d'internet au sein de structures à taille plus humaines et aux organisations plus conviviales aiderait à :

protéger les utilisateur·rices

Tout porte à croire que ramener les fondations d'internet au sein de structures à taille plus humaines et aux organisations plus conviviales aiderait à :

protéger les utilisateur·rices

tout en les responsabilisant

Tout porte à croire que ramener les fondations d'internet au sein de structures à taille plus humaines et aux organisations plus conviviales aiderait à :

protéger les utilisateur·rices

tout en les responsabilisant

Cette ligne directrice peut
s'appliquer sur bien des axes
problématiques du
numérique aujourd'hui :

- l'écologie
- la modération
- l'archivage
- les dépendances
- etc.

Dans cette optique, auto-héberger ses services et ses données présente de nombreux avantages :

- On récupère la souveraineté sur ses données
- On gagne en vie privée
- On gagne en liberté
- On est responsabilisé·e face à ses besoins
- On exploite réellement l'énergie consommée par ses serveurs

Dans cette optique, auto-héberger ses services et ses données présente de nombreux avantages :

- On récupère la souveraineté sur ses données
- On En hiver, si on se chauffe électriquement avec un thermostat automatique, l'énergie utilisée par les serveurs est gratuite !
- On gagne en liberté
- On est responsabilisé-e face à ses besoins
- On exploite réellement l'énergie consommée par ses serveurs

Mais cela recouvre également bien des contraintes...

- On repose sur une connexion internet pour particulier
- Un certain savoir-faire et moultes compétences sont requis
- Assurer la résilience de ses services est difficile et chronophage
- Bien sauvegarder ses données n'est pas évident

Mais cela recouvre également bien des contraintes...

- On repose sur une connexion internet pour particulier
- Un certain savoir-faire et moultes compétences sont requis
- Assurer la résilience de ses services est difficile et chronophage
- Bien sauvegarder ses données n'est pas évident

Sauvegarder pour se parer contre les pannes matérielles est une chose...
Sauvegarder pour se parer contre les cambriolages et les incendies en est une autre !

Sauvegarder pour se parer contre les pannes matérielles est une chose...
Sauvegarder pour se parer contre les cambriolages et les incendies en est une autre !

Répartir géographiquement ses données devient alors nécessaire.

On a vu récemment se développer au sein du CHATONS la notion d'**entre-hébergement** : dans le cadre du collectif, il s'agit de partager ses volumes de données entre hébergeurs.

En plus de renforcer l'intégrité des sauvegardes, on va améliorer la disponibilité pendant les coupures de liaison internet, de courant, ou pendant les déménagements d'administrateur·rices par exemple.

Pour assurer la confidentialité, on peut chiffrer les données au niveau applicatif.

Dans le cadre de l'administration de services en ligne, les systèmes de fichiers impliquent certaines difficultés.

Le standard S3 apporte des facilités. Il réduit le stockage à un paradigme de clé-valeur basé essentiellement sur deux opérations seulement : lire ou écrire une clé.

Dans le cadre de l'administration de services en ligne, les systèmes de fichiers impliquent certaines difficultés.

Le standard S3 apporte des facilités. Il réduit le stockage à un paradigme de clé-valeur basé essentiellement sur deux opérations seulement : lire ou écrire une clé.

- Il permet de s'affranchir de contraintes présentes avec le standard POSIX, notamment sur les nommages et la concurrence
- Il est pensé pour l'accès à distance
- Il s'intègre avec le protocole HTTP
- Il abstrait toute la couche système

Historiquement, S3 a été conçu pour être «*malloc()* sur internet»⁴.

4. <https://aws.amazon.com/fr/blogs/aws/amazon-s3-path-deprecation-plan-the-rest-of-the-story/>

Historiquement, S3 a été conçu pour être «*malloc() sur internet*»⁴.

Très pratique pour l'industrie, ce paradigme participe cependant à la centralisation et l'invisibilisation citées précédemment.

4. <https://aws.amazon.com/fr/blogs/aws/amazon-s3-path-deprecation-plan-the-rest-of-the-story/>

Garage

Garage essaye de fournir les facilités d'utilisation de S3, mais en conservant une infrastructure accessible derrière.

Il s'agit d'un logiciel libre permettant de distribuer un service S3 sur diverses machines éloignées.

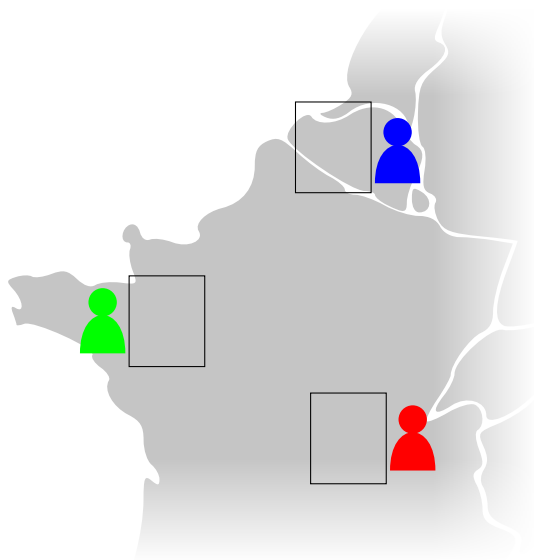


Garage

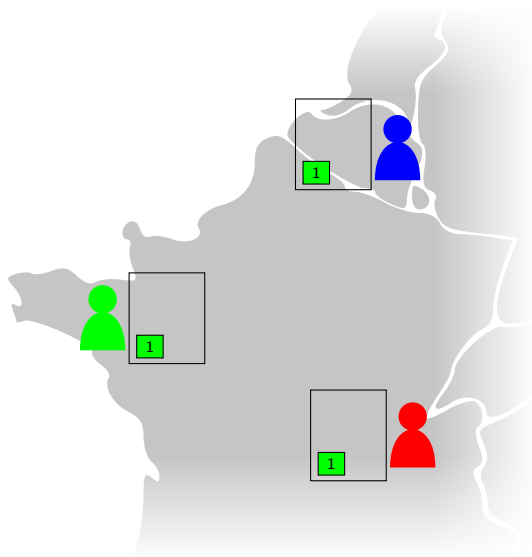
<https://garagehq.deuxfleurs.fr/>

Garage va prendre en compte les zones géographiques au moment de répliquer les données.

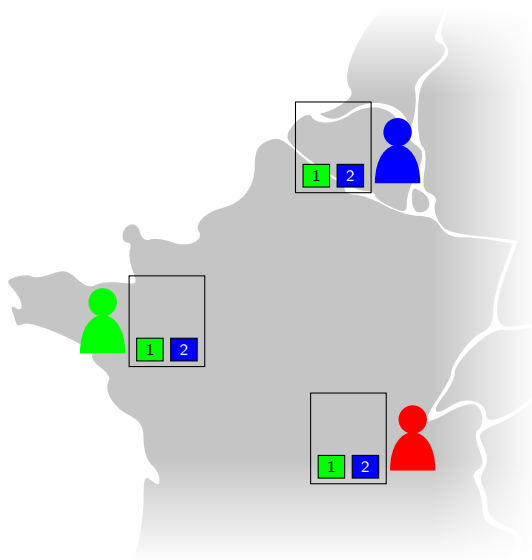
```
alex@io:~$ docker exec -ti garage /garage status
==== HEALTHY NODES ====
ID                Hostname  Address                               Tags                                Zone    Capacity
7d50f042280fea98  io       [2a01:e0a:5e4:1d0::57]:3901         [io,jupiter]                       jupiter 20
d9b5959e58a3ab8c  drosera  [2a01:e0a:260:b5b0::4]:3901         [drosera,atuin]                     atuin   20
966dfc7ed8049744  datura   [2a01:e0a:260:b5b0::2]:3901         [datura,atuin]                      atuin   10
8cf284e7df17d0fd  celeri   [2a06:a004:3025:1::33]:3901         [celeri,neptune]                    neptune 5
156d0f7a88b1e328  digitale [2a01:e0a:260:b5b0::3]:3901         [digitale,atuin]                    atuin   10
5fcb3b6e39db3dcb  concomb [2a06:a004:3025:1::31]:3901         [concomb,neptune]                   neptune 5
a717e5b618267806  courgette [2a06:a004:3025:1::32]:3901        [courgette,neptune]                 neptune 5
alex@io:~$
```

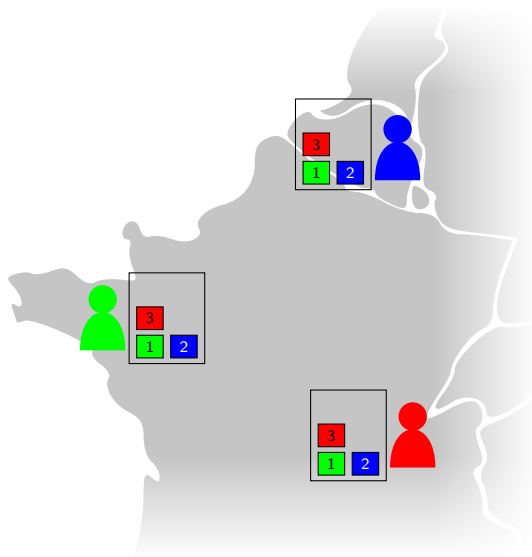
Chaque objet est dupliqué sur plusieurs zones différentes.



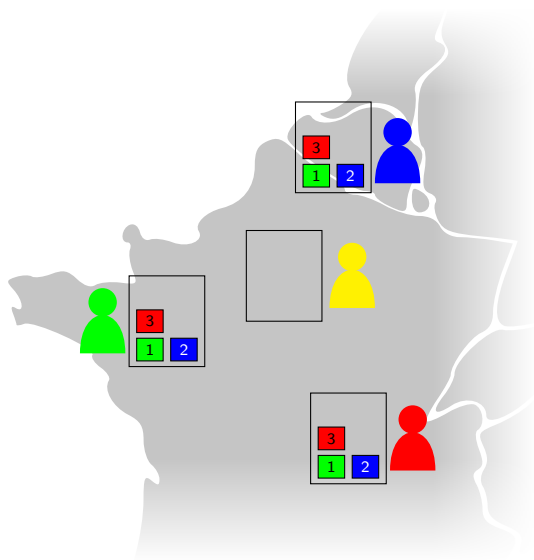
Chaque objet est dupliqué sur plusieurs zones différentes.



Chaque objet est dupliqué sur plusieurs zones différentes.

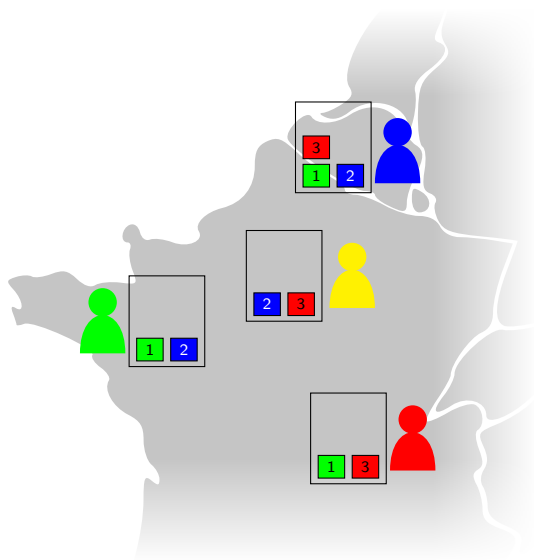


Chaque objet est dupliqué sur plusieurs zones différentes.



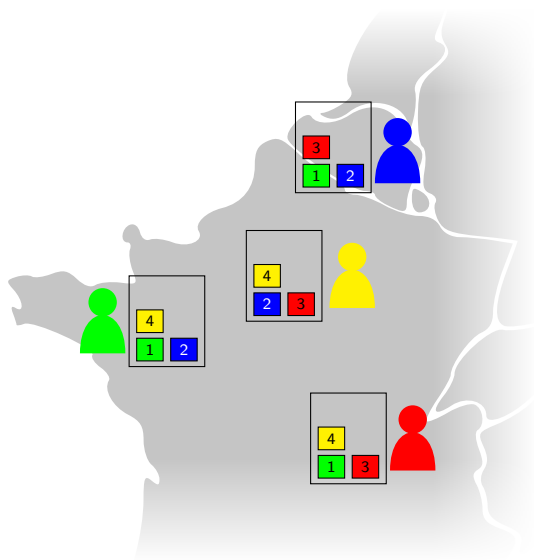
Chaque objet est dupliqué sur plusieurs zones différentes.

Lorsqu'un-e nouvel-le hébergeur-euse rejoint le réseau, la charge se voit équilibrée.



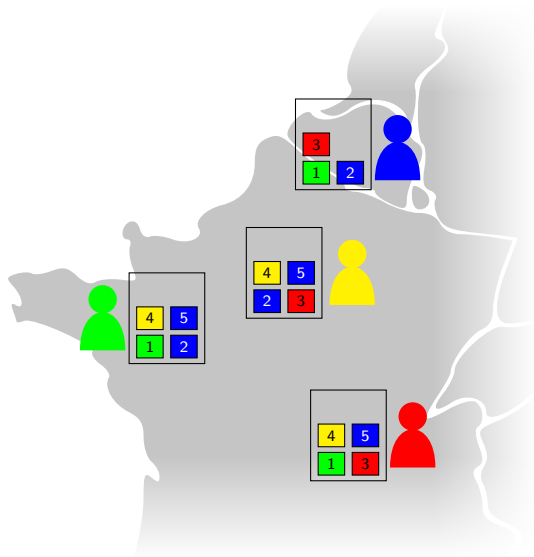
Chaque objet est dupliqué sur plusieurs zones différentes.

Lorsqu'un-e nouvel-le hébergeur-euse rejoint le réseau, la charge se voit équilibrée.



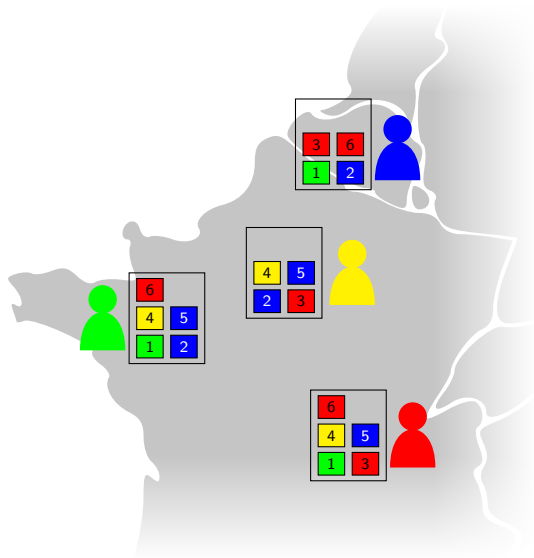
Chaque objet est dupliqué sur plusieurs zones différentes.

Lorsqu'un-e nouvel-le hébergeur-euse rejoint le réseau, la charge se voit équilibrée.



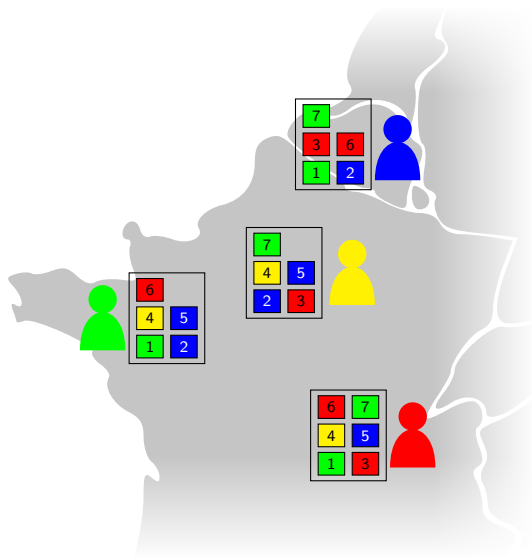
Chaque objet est dupliqué sur plusieurs zones différentes.

Lorsqu'un-e nouvel-le hébergeur-euse rejoint le réseau, la charge se voit équilibrée.



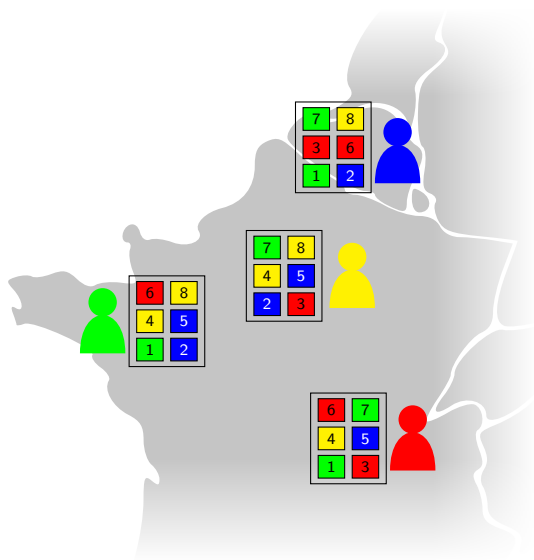
Chaque objet est dupliqué sur plusieurs zones différentes.

Lorsqu'un-e nouvel-le hébergeur-euse rejoint le réseau, la charge se voit équilibrée.



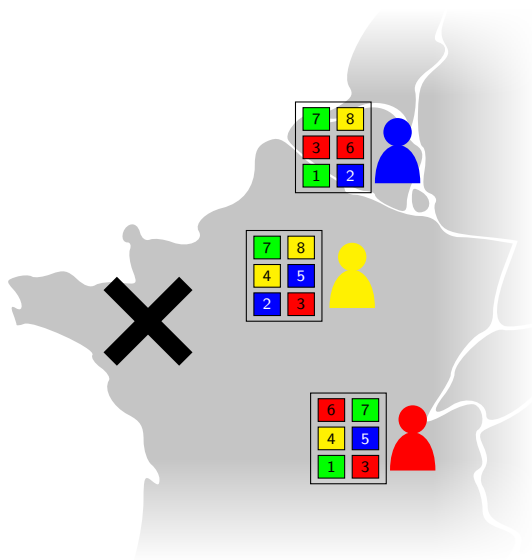
Chaque objet est dupliqué sur plusieurs zones différentes.

Lorsqu'un-e nouvel-le hébergeur-euse rejoint le réseau, la charge se voit équilibrée.



Chaque objet est dupliqué sur plusieurs zones différentes.

Lorsqu'un-e nouvel-le hébergeur-euse rejoint le réseau, la charge se voit équilibrée.



Chaque objet est dupliqué sur plusieurs zones différentes.

Lorsqu'un-e nouvel-le hébergeur-euse rejoint le réseau, la charge se voit équilibrée.

Si une zone devient indisponible, les autres continuent d'assurer le service.

Dans le cadre du programme *Horizon 2020* de l'Union Européenne, nous avons reçu une subvention de la part de l'initiative NGI Pointer⁵.



Nous avons ainsi pu financer le développement de Garage pendant 1 an.

5. Next Generation Internet Program for Open Internet Renovation

De par nos valeurs, nous avons attribué la licence AGPL version 3 à Garage, notamment afin qu'il reste parmi les biens communs.



Nous avons décidé d'écrire Garage à l'aide du langage Rust, afin d'obtenir une compilation vers des binaires natifs et efficaces.



Ce choix permet également de bénéficier des avantages reconnus de Rust en termes de sécurité.

Garage peut ainsi être performant sur des machines limitées. Les prérequis sont minimales : n'importe quelle machine avec un processeur qui a moins d'une décennie, 1 gigaoctet de mémoire vive, et 16 gigaoctets de stockage suffit.

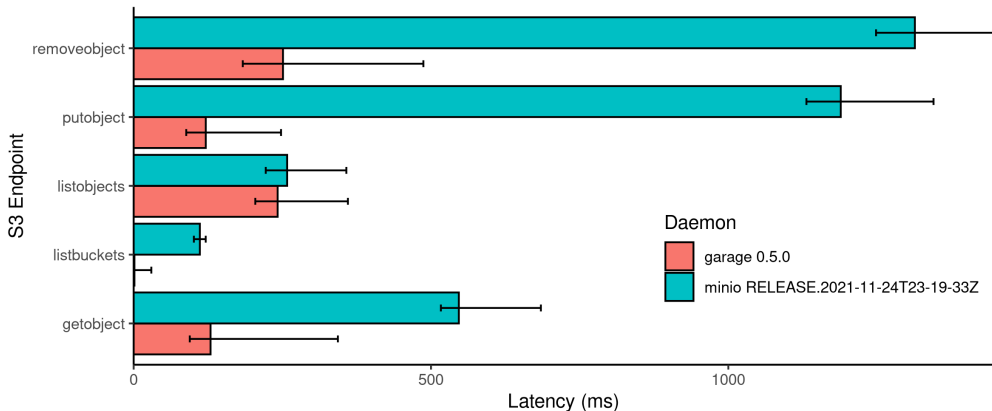
Cet aspect est déterminant : il permet en effet d'héberger sur du matériel acheté d'occasion, pour réduire l'impact écologique de nos infrastructures.

S3 endpoint latency in a simulated geo-distributed cluster

100 measurements, 5 nodes, 100ms RTT + 20ms jitter between nodes

no contention: latency is due to intra-cluster communications

colored bar = mean latency, error bar = min and max

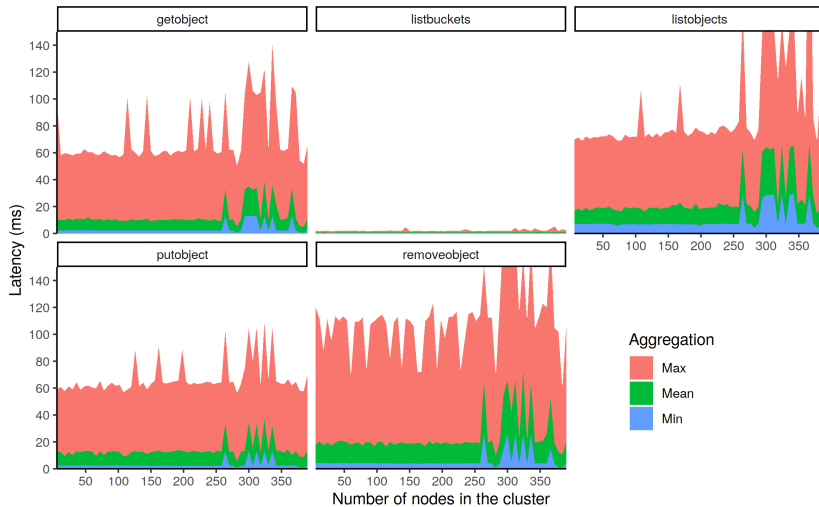


Get the code to reproduce this graph at <https://git.deuxfleurs.fr/quentin/benchmarks>

Garage response time on common S3 endpoints with various cluster size

ran on 6 physical nodes on Grid5000 with multiple instances on each node
spread on 3 zones: Lyon (nova), Rennes (paravance) and Nantes (economie)

few contention, latency is mainly due to RPC communications except on biggest values



Get the code to reproduce this graph at <https://git.deuxfleurs.fr/quentin/benchmarks>

Puisqu'il suit le standard S3, beaucoup de services populaires sont par conséquence compatibles avec Garage :



[**matrix**]



Et comme souvent avec S3, on peut assimiler un bucket à un site, et utiliser le serveur pour héberger des sites web statiques.

Installation simple en 5 minutes

L'infrastructure de Deuxfleurs

Quid de son exploitation chez Deuxfleurs ?

Nous bénéficions d'une grande répartition géographique de nos membres.

Et nous visons une offre publique avec un nombre d'utilisateurs pour l'instant relativement modeste.

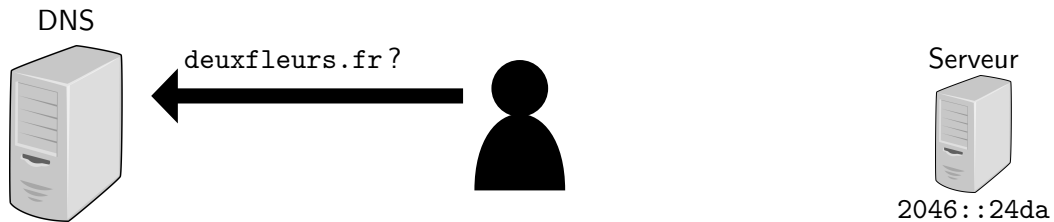
DNS

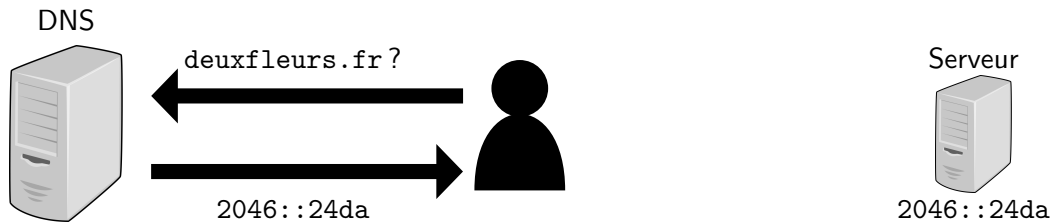


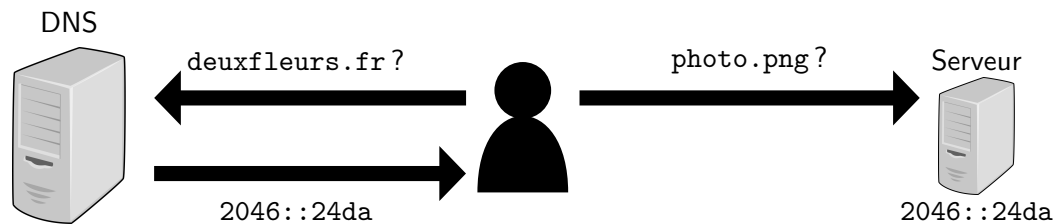
Serveur

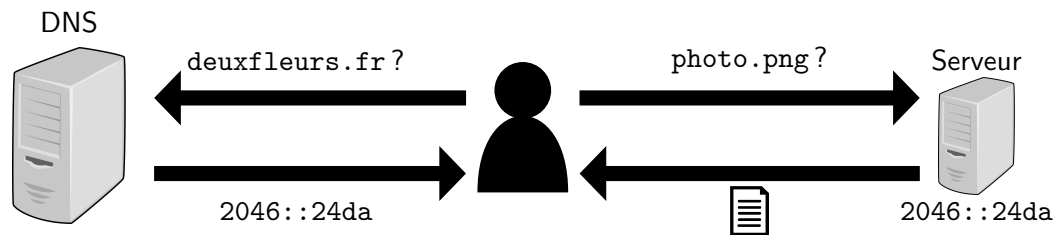


2046::24da











Serveur 1



2ac3::d3dc

Serveur 2

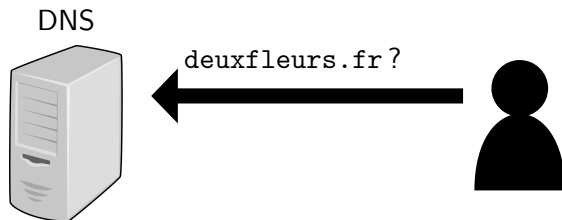


2046::24da

Serveur 3



262d::16bf



Serveur 1



2ac3::d3dc

Serveur 2

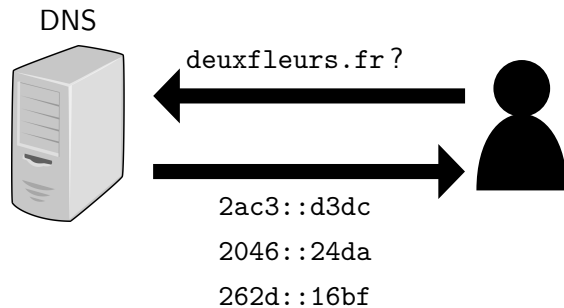


2046::24da

Serveur 3



262d::16bf



Serveur 1



2ac3::d3dc

Serveur 2

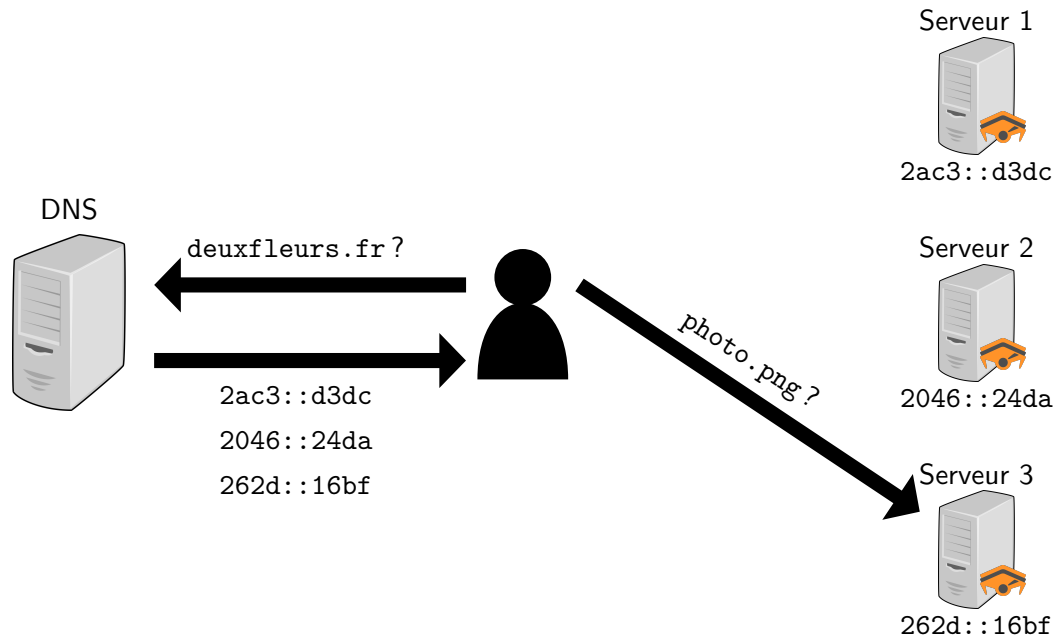


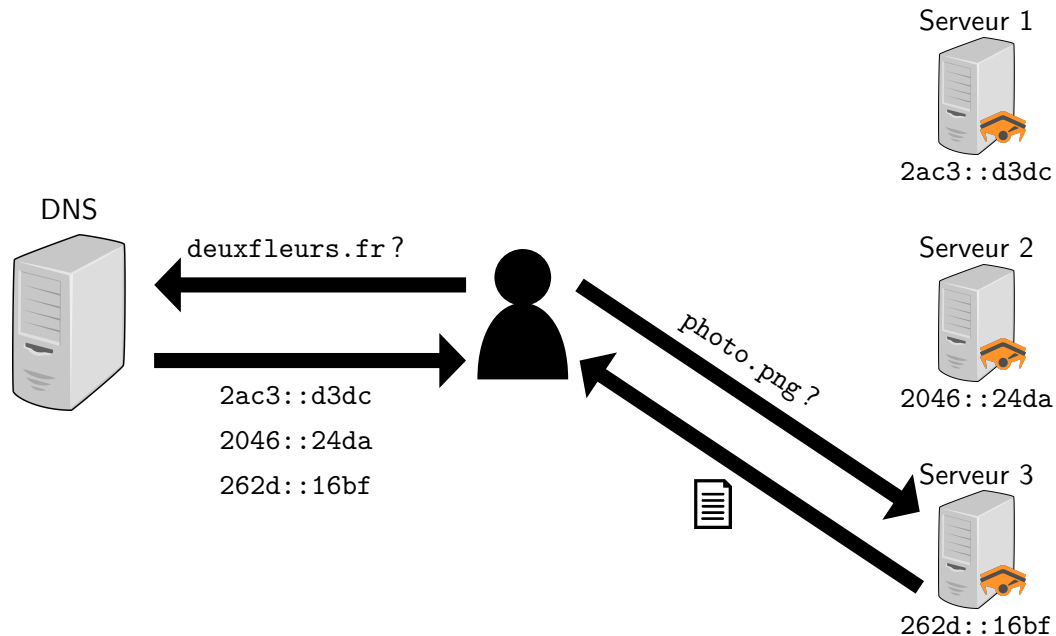
2046::24da

Serveur 3



262d::16bf





Serveur 1



2ac3::d3dc

Serveur 2



2046::24da

Serveur 3



262d::16bf

DNS



Serveur 1



2ac3::d3dc

Serveur 2



2046::24da

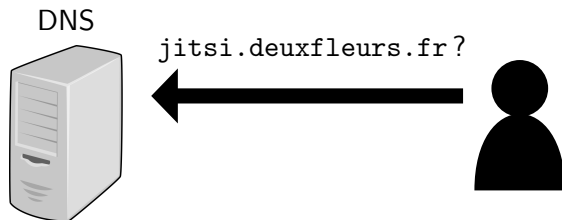
Serveur 3



262d::16bf

DNS





Serveur 1



2ac3::d3dc

Serveur 2

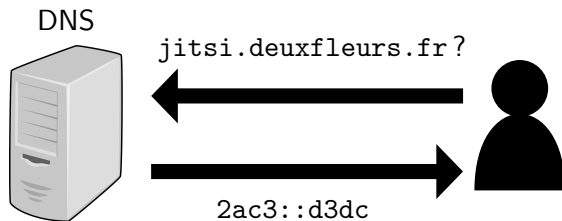


2046::24da

Serveur 3



262d::16bf



Serveur 1



2ac3::d3dc

Serveur 2

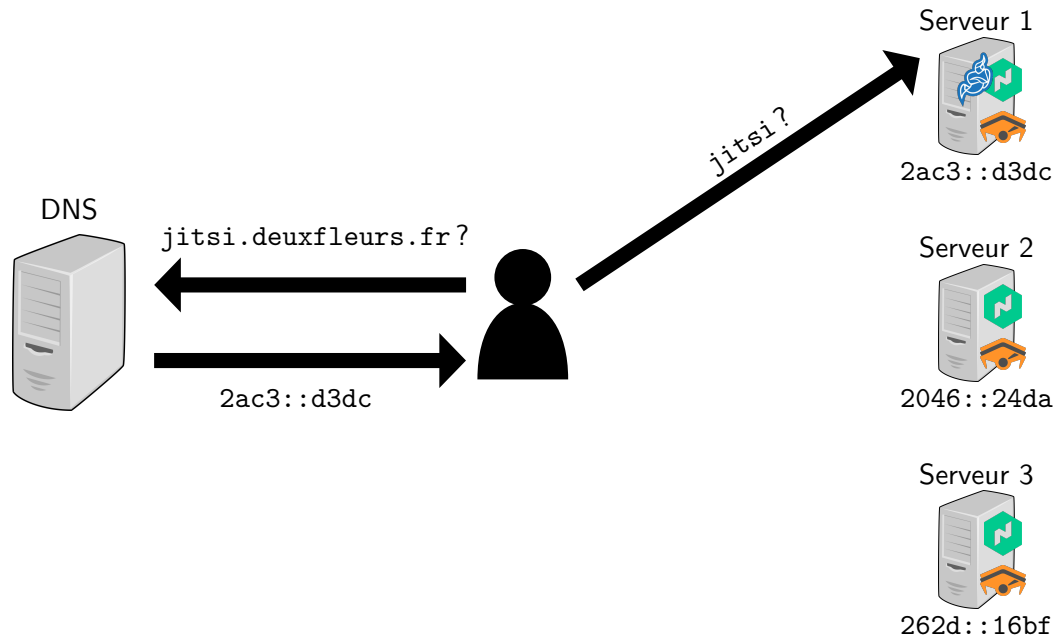


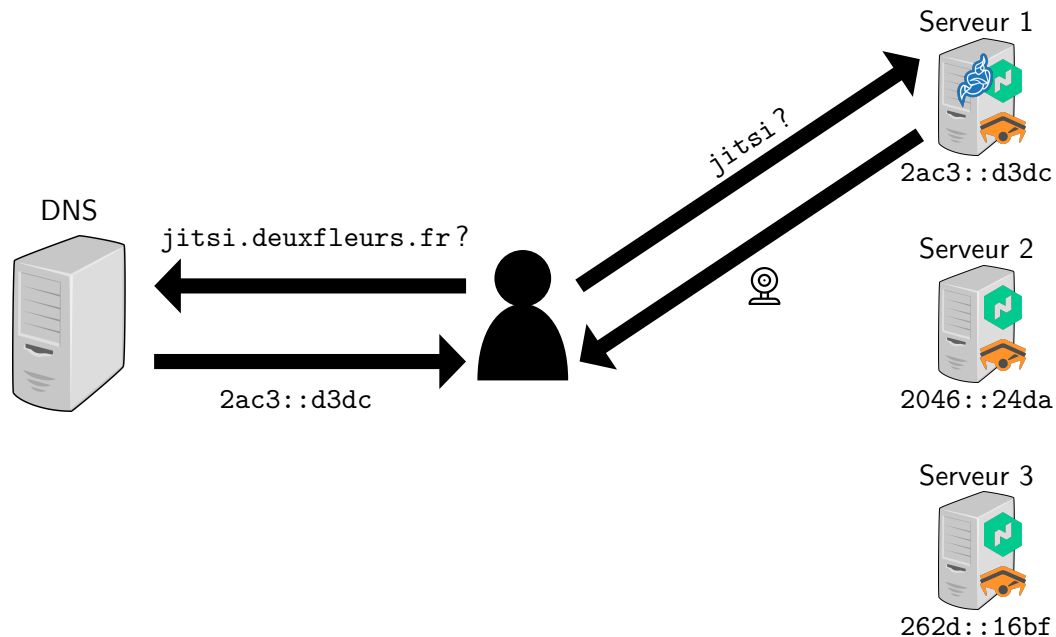
2046::24da

Serveur 3



262d::16bf







En pratique, nos serveurs ne sont effectivement que des machines achetées d'occasion (très souvent des anciens ordinateurs destinés à la bureautique en entreprise).

Pour faciliter la reproduction d'un environnement connu, NixOS est installé sur nos machines.

Cela nous permet de mieux contrôler l'état logiciel de notre parc, et de potentiellement réduire le ticket d'entrée en terme de savoir pour participer à l'infrastructure.

Nous fonctionnons beaucoup derrière des routeurs grands publics : Livebox, Freebox, Bbox...

Pour s'accommoder des réseaux qu'on trouve derrière des routeurs pour particuliers, nous nous aidons de notre logiciel Diplonat⁶.

6. <https://git.deuxfleurs.fr/Deuxfleurs/diplonat>

Le futur

**TeDomum**

@tedomum@mastodon.tedomum.net



Suivre

Tout est dans les clous et super nouvelle infrastructure pour nous.

Dorénavant les vidéos peertube sont stockées sur un cluster garage (dev des copains de deuxfleurs.fr) lui même tournant sur un cluster hepto (notre distro k8s avec une prochaine release bientôt), dans le salon des membres :)

Bientôt enfin plus grand chose en datacenter ! Prochain obstacle Mastodon.

**Deuxfleurs : fabriquons un internet convivial**

Association qui œuvre en utilisant ses propres infrastructures, des outils sobres et des valeurs fortes.

deuxfleurs.fr

9 novembre 2022 à 23:10 · 🌐 · Tusky · 🔄 2 · 📄 3 · ★ 4

Garage a pris son
envol au-delà de
Deuxfleurs.

Et ceci dans tous
types de structures.

Le développement d'Aerogramme⁷ est en cours.



Il s'agit d'un serveur de stockage de courriels chiffrés.

Il est conçu pour reposer sur Garage.

7. <https://aerogramme.deuxfleurs.fr/>



Intéressé-e ?

Contactez-nous !

Par courriel : coucou@deuxfleurs.fr

Via matrix : [#forum:deuxfleurs.fr](https://matrix.to/#/#forum:deuxfleurs.fr)