

Pour intégrer les données de météo France, j'ai dû réorganiser la base de données, qui prend plus de place qu'actuellement. Cette nouvelle structure permettra d'ajouter des mesures à tout moment.

Actuellement la base de données climato fait 10Go, et il reste 15Gb de disponible sur le disque.

J'ai chargé sur mon mac nos données, ainsi que tout l'historique de toutes les données météo France sur la réunion, et la base de données fait 120Gb. Les tests de performance sont tout à fait corrects pour permettre une utilisation interactive via Grafana, à condition de faire des requêtes sur les bonnes tables.

Par contre le serveur actuel dina, ne permet pas de monter cette base de données, ni de l'exploiter. Il faudrait un serveur plus puissant, avec plus de capacité.

Point d'avancement sur le projet :

=====

Le projet comporte plusieurs sous-projets :

- Gestion de WeeWX sur les Raspberry. Pour le moment les Raspberry envoient chaque nuit un backup de la base de données de WeeWX. Un petit nombre peut envoyer un fichier de mesure (appelé fichier JSON) toutes les 5 mn. Il faut prévoir une mise à jour de l'ensemble des Raspberry pour avoir cet envoi généralisé, ainsi que de changer la carte SD qui se détériore avec le temps.
- L'interface utilisateur : Se fait via Grafana, qui peut soit afficher des « Dashboard » préparés, soit être utilisé en mode interactif par des utilisateurs plus avancés, pour des demandes spécifiques.
- La partie serveur, qui centralise les données venant de :
 - Backup WeeWX quand seulement cela est disponible
 - Fichiers JSON envoyés toutes les 5 mn par WeeWX
 - Fichiers CSV de météo France, qui contient plus de mesures horaires, et regroupés par deux années. Au 31/12/24 ce fichier fera environ 500Mb, ce qui limite à une mise à jour journalière, la nuit. Les dates des valeurs extrêmes ne se trouvent que dans ces fichiers. Aussi concernant l'historique, il sera possible de ne charger que les données quotidiennes, pour limiter la taille de la base de données.
 - API (interface de programmation) de météo France, qui donne les mesures principales de la station. La fréquence peut être de 6mn, ou par heure. Vu le volume seules les données horaires seront utilisées. Cette interface a plusieurs inconvénients :
 - Actuellement seules les données du jour sont disponibles. En cas de problème de météo France ou de notre serveur, il faudra une procédure de reprise via les fichiers CSV.
 - Il y a moins de mesures disponibles via ces APIs

Au niveau des Raspberry, nous sommes en train d'évaluer la dernière version de WeeWX, et de générer une nouvelle image SD qui sera utilisable lors de la mise à jour, ou remise en

route post Belal. Ce qui bloque c'est surtout le manque de ressource humaine pour aller sur chacun des sites du réseau.

Au niveau de l'interface, de nombreux Dashboard ont été réalisés. Il faudra les adapter à la structure de la nouvelle base de données. Ce seront les utilisateurs qui vont faire remonter des demandes de nouveaux Dashboard.

Au niveau du serveur, cela est un peu plus compliqué :

- Pour le moment les données des stations WeeWX sont mises à jour chaque nuit à partir des backups venant des Raspberry
- Le chargement à partir des fichiers JSON n'a pas été activé. Il sera réactivé sur le serveur actuel dans les semaines à venir pour les stations envoyant ces fichiers.
- Le chargement de fichiers CSV des mesures horaires fonctionne. Il reste à automatiser le download du fichier à partir du site de météo France. Il restera à coder deux choses :
 - o Gestion de reprise de données pour les stations mises à jour par API, à la suite d'un incident
 - o Chargement partiel de données (pour compléter les données non disponibles via l'API météo France)
- Coté API, je suis actuellement sur le sujet, et cela devrait être terminé rapidement.

Il me semble qu'il faudra ajouter une fonction d'export automatique permettant de partager nos données avec météo France, et probablement InfoClimat.

Actuellement dina, le serveur actuel arrive à gérer l'ancienne version de la base de données, ainsi qu'une petite base de test de la nouvelle structure, permettant de commencer l'adaptation de Grafana. La nouvelle base ne pourra être mise en ligne que sur un serveur plus puissant, dont voici les caractéristiques nécessaires :

- Processeur plus récent. Le chargement des données sur un mac récent (processeur M2) a duré plus de 8 jours. La CPU était le facteur bloquant.
- RAM : 64 Gb me paraît être le minimum pour faire tourner MySQL (backup WeeWX), PostgreSQL/TimeScaleDB, et Grafana.
- Stockage : 2 disques SSD de 512Gb en RAID, avec éventuellement un disque dur 2Gb pour sauvegarde locale (je n'ai pas encore regardé comment les sauvegardes étaient gérées).