



I-SITE
CLERMONT
Clermont Auvergne Project

CEBA

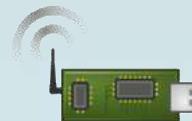
Cloud Environnemental au Bénéfice de l'Agriculture

Jérémie Mezhoud – Francis Ogereau



Les données à (in)gérer

Capteurs



“ temps réel ”

Fichiers



Base de données



Quoi, quand, où

Données spatiales, GPS



Image, audio, vidéo



Besoins

- (1) Collecter/Ingérer** (fichiers, flux...)
- (2) Stocker** (données)
- (3) Indexer** (sources de données, jeux de données)
- (4) Publier** (jeux de données, données)
- (5) Retrouver** (par mot clé, par position, par valeur,...)
- (6) Visualiser/Restituer** (jeux de données, graphes,...)
- (7) Analyser/Transformer**
- (8) Sécuriser** (accès limité, accès public)

Éléments du CEBA

- | | |
|--------------------------|---|
| • Data Lake | (1)(2) |
| • Moteur d'indexation | (3)(4)(5)(6)(7) |
| • Catalogue de données | (4)(5)(6) |
| • Outil de visualisation | (5)(6) |
| • Base de données | (2)(7) |
| • Système de fichiers | (2)

(1) Collecter/Ingérer
(2) Stocker
(3) Indexer
(4) Publier
(5) Retrouver
(6) Visualiser/Restituer
(7) Analyser/Transformer
(8) Sécuriser |

Éléments du CEBA

Data lake

COLLECTER

TRANSFORMER, RESTITUER, ANALYSER

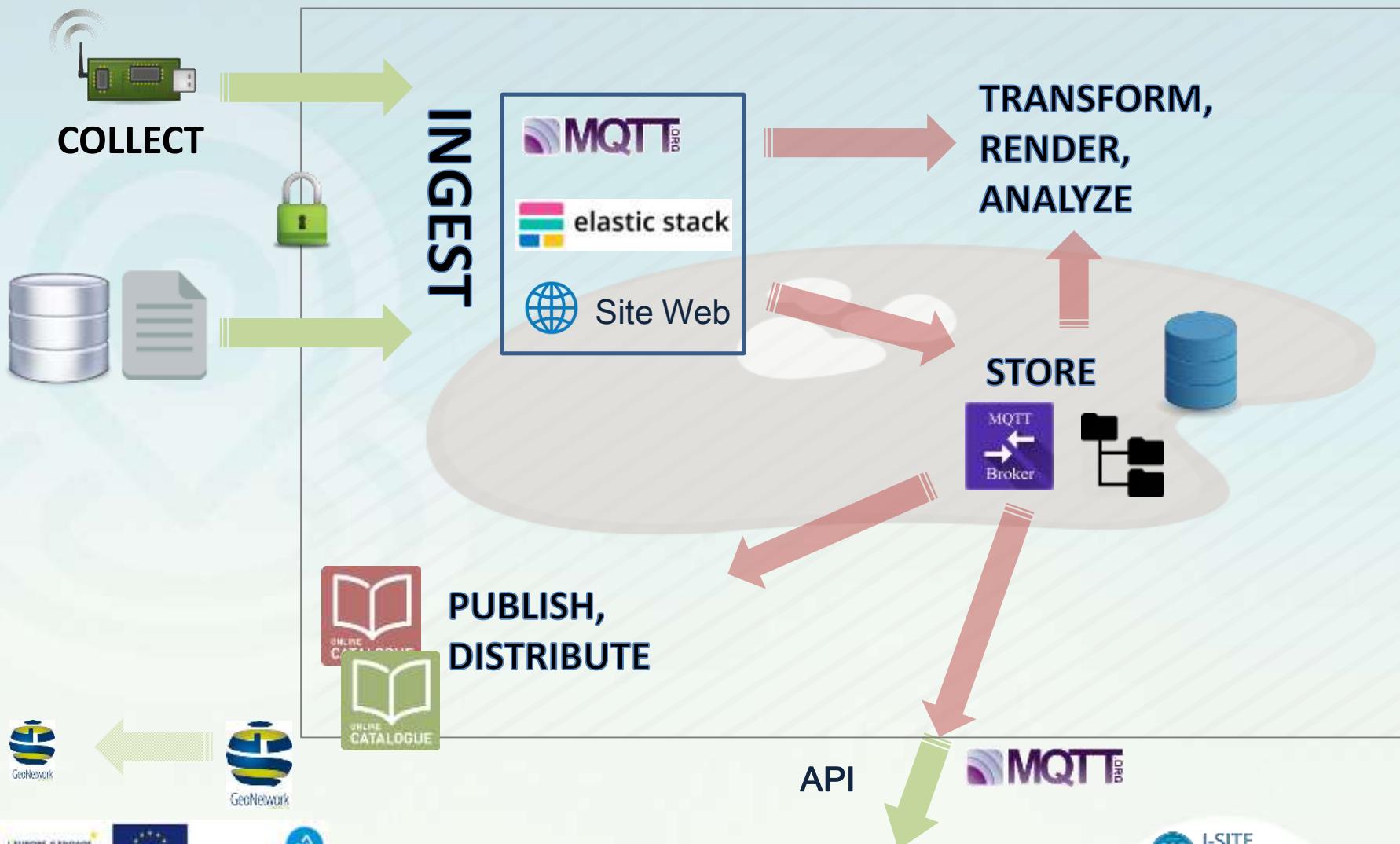


INGERER → STOCKER

PUBLIER, DISTRIBUER

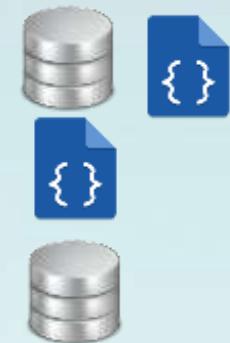
Éléments du CEBA

Big picture



Élément du CEBA par l'exemple

- Réseau de capteurs sans fils
 - Réseau libelium (bd) (json)
 - Réseau ConnecSens (json)
- Base de données (bd)
- Site web
- Catalogue de données
- Métadonnées



Réseau de capteurs sans fils

Réseau Libelium

Description des données



Base de données MySQL

Column	Type
id	int(11)
name	varchar(45)
description	varchar(45)
id_ascii	varchar(45)
units	varchar(45)
value	int(11)



sensor
TCB
HUMB
PA
TCC
BAT
PLV1
PLV2
PLV3
ANE
WV
IN_TEMP
RAD
TIME
MAC
ACC
DATE
GPS
ALT

Flux



//



↓



logstash

Requête SQL
Synchronisation (5 min)

Table	25000
# @timestamp	2018-12-10T16:12:44.000Z
# @version	1
# _id	325482
# _index	mtlsensors
# _score	2
# _type	sensorDataAsp
# frame_number	31
# id	325482
# id_secret	400578826
# id_wasp	RA002
# sensor	BAT
# timestamp	2018-12-10 16:12:44
# value	97



elasticsearch



kibana

Création d'alias pour faciliter les requêtes/visualisation

```
POST /_aliases
{
  "actions": [
    {
      "add": {
        "index": "mtlsensors",
        "alias": "mtlsensorstcalias",
        "filter": { "term": { "sensor.keyword": "TCB" } }
      }
    }
  ]
}
```

Réseau de capteurs sans fils

Réseau ConnecSens

Description des données

```
object {39}
  data-illuminance-unit : lux
  txInfo-codeRate : 4/5
  fPort : 2
  data-node-timestampUTC : 2018-10-02 07:09:43
  txInfo-adr : false
  servertimestampUTC : 2018-10-02T10:54:21.402434
  txInfo-dataRate-bandwidth : 125
  data-node-batteryVoltage-alarmIsLow : false
  txInfo-dataRate-modulation : LORA
  data-DataChannel : 1
  rxInfo-longitude : 0
  fCnt : 0
  deviceName : HE31294
  applicationName : ConnecSens-Nodes
  rxInfo-name : Kerlink01
  rxInfo-rssi : -26
  devEUI : 434e535301e31294
  data-illuminance : 352
  applicationID : 1
  rxInfo-mac : 7276ff00000000f4c
  data-node-batteryVoltage : 3.92
  data-node-batteryVoltage-unit : V
  rxInfo-latitude : 0
  rxInfo-loRaSNR : 0
  txInfo-dataRate-spreadFactor : 7
  rxInfo-altitude : 0
  txInfo-frequency : 868300000
  data-temperature-alarmL : false
  data-atmosphericPressure-unit : hPa
  data-airHumidity-unit : %
  data-temperature-unit : °C
  data-atmosphericPressure : 967.76
  data-acceleration-unit : g
  data-acceleration-axisZ : -0.271
  data-temperature-alarmH : false
  data-airHumidity : 36
  data-temperature : 19.4
  data-acceleration-axisY : 0.021
  data-acceleration-axisX : -0.896
```



Flux



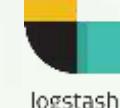
Fichiers (au mesocentre)

Json Elasticsearch

```
1 {"index":{"_id":"1"}}
2 {"data-illuminance-unit": "lux", "txInfo-codeRate": "4/5", "fp
  node-timestampUTC": "2018-10-02 07:10:13", "txInfo-adr": f
  servertimestampUTC": "2018-10-02T10:54:21.402434", "txInf
  bandwidth": 125, "data-node-batteryVoltage-alarmIsLow": fa
  dataRate-modulation": "LORA", "data-DataChannel": 1, "rxIn
  "fCnt": 0, "deviceName": "HE31294", "applicationName": "Co
  "rxInfo-name": "Kerlink01", "rxInfo-ssai": -26, "devEUI": "
  434e535301e31294", "data-illuminance": 340, "applicationI
  mac": "7276ff00000000f4c", "data-node-batteryVoltage": 3.92
  batteryVoltage-unit": "V", "rxInfo-latitude": 0, "rxInfo-l
  "txInfo-dataRate-spreadFactor": 7, "rxInfo-altitude": 0, "
  868300000}
```

```
3 {"index":{"_id":"2"}}
```

Table	JSON
_id	Q Q D * 34
_index	Q Q D * flattened_4_1
_score	Q Q D * 1
_type	Q Q D * type4
applicationID	Q Q D * 1
applicationName	Q Q D * ConnecSens-Nodes
data-DataChannel	Q Q D * 1
data-airHumidity	Q Q D * 40.5
data-airHumidity-unit	Q Q D * %
data-atmosphericPressure	Q Q D * 962.14
data-atmosphericPressure-unit	Q Q D * hPa
data-illuminance	Q Q D * 251
data-illuminance-unit	Q Q D * lux



I-SITE
CLERMONT
Clement Auvergne Project

Réseau de capteurs sans fils

Monitoring/Dashboard

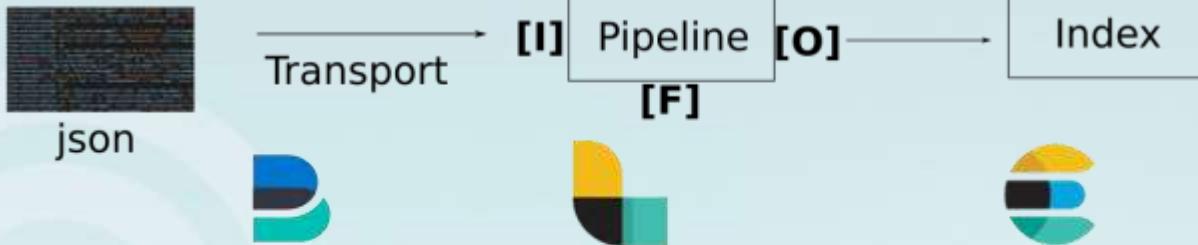
Grafana / Kibana



Réseau de capteurs sans fils

Traitement de flux de données capteurs

Données, mesures...



```
tail -qg  
# Change to true to enable this input configuration.  
input {  
    file {  
        path => "/var/log/test.log"  
        type => "log"  
        type => "datastore"  
        type => "datastore-test"  
        type => "datastore-log"  
        type => "datastore-raw"  
        type => "datastore-raw-test"  
        type => "datastore-raw-log"  
        type => "datastore-raw-test-log"  
    }  
}
```

```
input {  
    beats {  
        port => 5042  
        ssl => true  
        ssl_certificate => "/etc/logstash/cert.pem"  
        ssl_key => "/etc/logstash/logstash.key"  
        ssl_verify_mode => "force_peer"  
        ssl_certificateAuthorities => [...]  
    }  
}  
filter {  
    if [type] == "log" {  
        grok {  
            match => { "message" => "%{SYSLOG}" }  
            add_field => [ "received_at", "%{@timestamp}" ]  
            add_field => [ "received_from", "%{host}" ]  
            remove_field => [ "message" ]  
        }  
        syslog_pri {}  
        date {  
            match => [ "syslog_timestamp", "%{@timestamp}" ]  
        }  
    }  
}  
output {  
    if [fields][log_type] == "message" {  
        elasticsearch {  
            hosts => ["192.168.1.92:9200"]  
            index => "cebalora01-data-%{@_id}"  
        }  
    }  
    if [fields][log_type] == "test-data" {  
        elasticsearch {  
            hosts => ["192.168.1.92:9200"]  
            index => "cebalora01-test-data-%{@_id}"  
        }  
    }  
    if [fields][log_type] == "datastore" {  
        elasticsearch {  
            hosts => ["192.168.1.92:9200"]  
            index => "cebalora01-datastore-%{@_id}"  
        }  
    }  
}
```

cebalora01-data*

Time Filter field name: @source.timestampUTC

This page lists every field in the cebalora01-data* index and the field's associated type as recorded by Elasticsearch. To change a field type, use the Elasticsearch Mapping API.

Fields (171)	Scripted fields (0)	Source filters (0)		
<input type="text"/> Filter		A		
Name	Type	Format	Search...	Aggregate
@timestamp	date		•	•
@version	string		•	
@version.keyword	string		•	•
_id	string		•	•
_index	string		•	•
_score	number			
_source	_source			

Base de données

- Suivi en amont
- Création d'une base avec POSTGRES pour le stockage de données JSON
 - Données récupérées par capteurs
- Intégration via un script bash
 - Ligne par ligne
 - Lancement manuel (Automatisation en cours)

Base de données

Exemples

- Récupérer la liste des sites par fichier JSON
- Requête:

```
select
    json_file.filename,
    json_row.data ->> 'applicationName' as
    nom
from
    json_row, json_file
where
    json_row.json_file_id = json_file.id
group by
    json_file.filename, nom
;
```

Base de données

Exemples

- Récupérer la liste des sites par fichier JSON
- Résultat:

	filename text	nom text
1	data20190401.json	Cezeaux
2	data20190401.json	Jardin-IRSTEA
3	data20190401.json	Montoldre
4	data20190401.json	Radon-Sensors
5	data20190402.json	Cezeaux
6	data20190402.json	Jardin-IRSTEA
7	data20190402.json	Montoldre
8	data20190402.json	Radon-Sensors
9	data20190403.json	Cezeaux
10	data20190403.json	Jardin-IRSTEA
11	data20190403.json	Montoldre
12	data20190403.json	Radon-Sensors
13	data20190404.json	Cezeaux
14	data20190404.json	Jardin-IRSTEA
15	data20190404.json	Montoldre
16	data20190404.json	Radon-Sensors

Base de données

Exemples

- Récupérer la liste des Dev-EUI par application-name
- Résultat:

	nom text	deveui text
1	Cezeaux	434e535305e36210
2	Cezeaux	434e535305e31297
3	Cezeaux	434e535305e31273
4	Cezeaux	434e535305e36250
5	Cezeaux	434e535305e36226
6	Cezeaux	434e535305e31233
7	Cezeaux	434e535305e36244
8	Cezeaux	434e535305507403
9	Cezeaux	434e535305e31263
10	Cezeaux	434e535305e31291
11	Cezeaux	434e535305e31279
12	Cezeaux	434e535305e31231
13	Cezeaux	434e535305e31230
14	Cezeaux	434e535305e31277
15	ETNA	434e535302e36236
16	ETNA	434e535302e36241

Base de données

Exemples

- Afficher l'évolution du nombre de trames par devEUI par jour
- Résultat:

	deveui text	dateserveur date	nbtrames bigint
1	434e535300e31208	2019-04-01	24
2	434e535300e31216	2019-04-01	861
3	434e535300e31235	2019-04-01	240
4	434e535301e36231	2019-04-01	6
5	434e535303e36199	2019-04-01	154
6	434e535303e36212	2019-04-01	154
7	434e535305e31230	2019-04-01	232
8	434e535305e31231	2019-04-01	230
9	434e535305e31233	2019-04-01	232
10	434e535305e31273	2019-04-01	232
11	434e535305e31277	2019-04-01	232
12	434e535305e31279	2019-04-01	232
13	434e535305e31291	2019-04-01	80
14	434e535305e31297	2019-04-01	170
15	434e535300e31208	2019-04-02	24
16	434e535300e31216	2019-04-02	1105

Site Web

On peut:

- Voir projet/expérience
- Rechercher par mot clé
- Voir des zones géographiques

Un utilisateur peut créer:

- un projet
- Un site
- Une expérience

Éditer le Projet

INFORMATIONS

Nom du projet	REseau de Suivi des Eaux en Auvergne	Status	Actif
Date de création du projet	01/01/2014	Organisme porteur	Observatoire de Physique du Globe de Clermont
			Site web
			http://wwwobs.univ-bpclermont.fr/SO/auvernat

IMAGE

Image: Browse

DESCRIPTION DU PROJET

Résumé:

Ce projet, initialement appelé RESEAU, devenu AUVER-WATCH, est prévu pour un fonctionnement de 2014 à 2024 et s'attache à caractériser les masses d'eau dans l'espace et le temps.

Description:

Auverwatch

L'hydrologie au sens large est définie par ses auteurs comme la discipline consacrée à l'étude du cycle de l'eau pris sous tous les angles physiques, physico-chimiques et biologiques, et à l'étude des interactions entre le cycle de l'eau et les activités humaines dans ce milieu. Ses applications sont multiples tant du point de vue de la gestion de l'environnement que de son implication dans les activités humaines : production d'eau potable ou industrielle, géotechnique, production d'énergie... Les pressions qui pèsent ainsi sur les ressources en eau sont de plus en plus fortes, principalement en raison des activités humaines -notamment urbanisation, la croissance démographique, l'augmentation du niveau de vie, la consommation croissante pour les ressources en eau et la pollution. De plus, le changement climatique et les variations des conditions naturelles viennent aggraver ces pressions. Dans ce contexte, l'enjeu majeur est de parvenir à une gestion du cycle de l'eau conforme aux exigences d'un

Site Web

Un utilisateur peut:

- Créer un jeu de données
 - Metadonnée/ Thesaurus



Site Web CEBA

Import de Métadonnées

...
...

Informations obligatoires

Status
completed

Titre
Mesures d'activités radiologiques de source

Résumé
Mesure par spectrométrie de l'activité en radioéléments d'origine naturelle (chaînes de décroissance de

Ressource type
dataset

Type de représentation spatiale
vector

Date de début
31 / 12 / 2018

Date de fin
03 / 07 / 2019

Nom de l'étendue spatiale
Lachaux

Catégories de sujets
environment, geoscientificinformation

Catalogue de données

- Fonctions principales
 - Dépôt
 - Outils de recherche
 - Association de métadonnées
 - Classement thématique
- Fonctions Optionnelles
 - Exposition en OAI-PMH
 - Embargo / Diffusion restreinte
 - Statistiques de consultation

Catalogue de données

- Intérêt du catalogage :
 - Echange avec les partenaires extérieurs
 - Visibilité interne des données disponibles
 - Thesaurus
- Inconvénients
 - Métadonnées à renseigner
 - Ne traite pas les flux continus

Catalogue de données

Jeu de données

backTohome

< Précédent

Suivant >

Télécharger

Mode affichage

Mesures d'activités radiologiques des sources minérales ZATU

Mesures par spectrométrie gamma de l'activité en radioéléments d'origine naturelle (chaines de décroissance de l'Uranium238 et du Thorium 232) dans l'eau et les sédiments des sources minérales radioactives en Auvergne. Contributeurs : LPC

completed

À propos de cette ressource

Catégories

Jeux de données

Environnement

Informations géoscientifiques

GEMET - INSPIRE themes, version 1.0

- ressources minérales
- lieux de production et sites industriels
- installations de suivi environnemental

GEMET

- eau de source
- radioactivité naturelle
- rayonnement gamma
- dose équivalente

Langue

- fra

Identificateur de ressource unique

- ZATU_a96c0982-d71f-452e-8c4c-fc41322f8ae8

Contraintes légales

This work is licensed under a Creative Commons Attribution 4.0 License (CC BY SA 4.0, <https://creativecommons.org/licenses/by-sa/4.0/>).

Aucune évaluation

Voir tous les commentaires

Ajouter votre commentaire

Étendue spatiale

Étendue spatiale

Étendue temporelle

Date de création

2018-12-31

Date de révision

2019-07-03

Fourni par



Mis à jour :

il y a 2 mois

Partager



Métadonnée

Pour les jeux de données

- Données décrivant un jeu de données
- Minimum: Quoi, Quand et Où
- Flux de données de capteurs
 - Métadonnées standardisées et potentiellement extraites automatiquement
- Ajout manuel
 - Liste de métadonnées épurée
 - Indexation

Métadonnées

Proposition

- Fiche de métadonnées épurée grâce au Réseau des ZA (90 % INSPIRE)
- Jeu de données moissonnable et disponible depuis l'extérieur
 - Title, Abstract, Start_date, End_date, Inspire_themes...
- Utilisation de thésaurus (Envthes, Gemet...)

Geocatalog of French LTSERs: metada.data-za.org



Conclusions

- **Architecture de préproduction**
 - Projet /experience pour créer des jeux de données → geonetwork
- **Le management des données :**
 - Réseau de capteurs sans fil
 - 5 réseaux en ligne
 - Stockage dans une base de donnée JSON (travail en cours)
 - Fichiers (testé depuis geonetwork)
 - Base de données (travail en cours)
- **Management des métadonnées (automatique)**
- **Solutions Open source**

Collaborations / Échanges

- Régional
 - ZATU
 - ConnecSens
 - LIT
 - CRAIG
 - EFOA
- National
 - Réseau des ZA
- International
 - Ravioli / LMV / INGV (Italie)

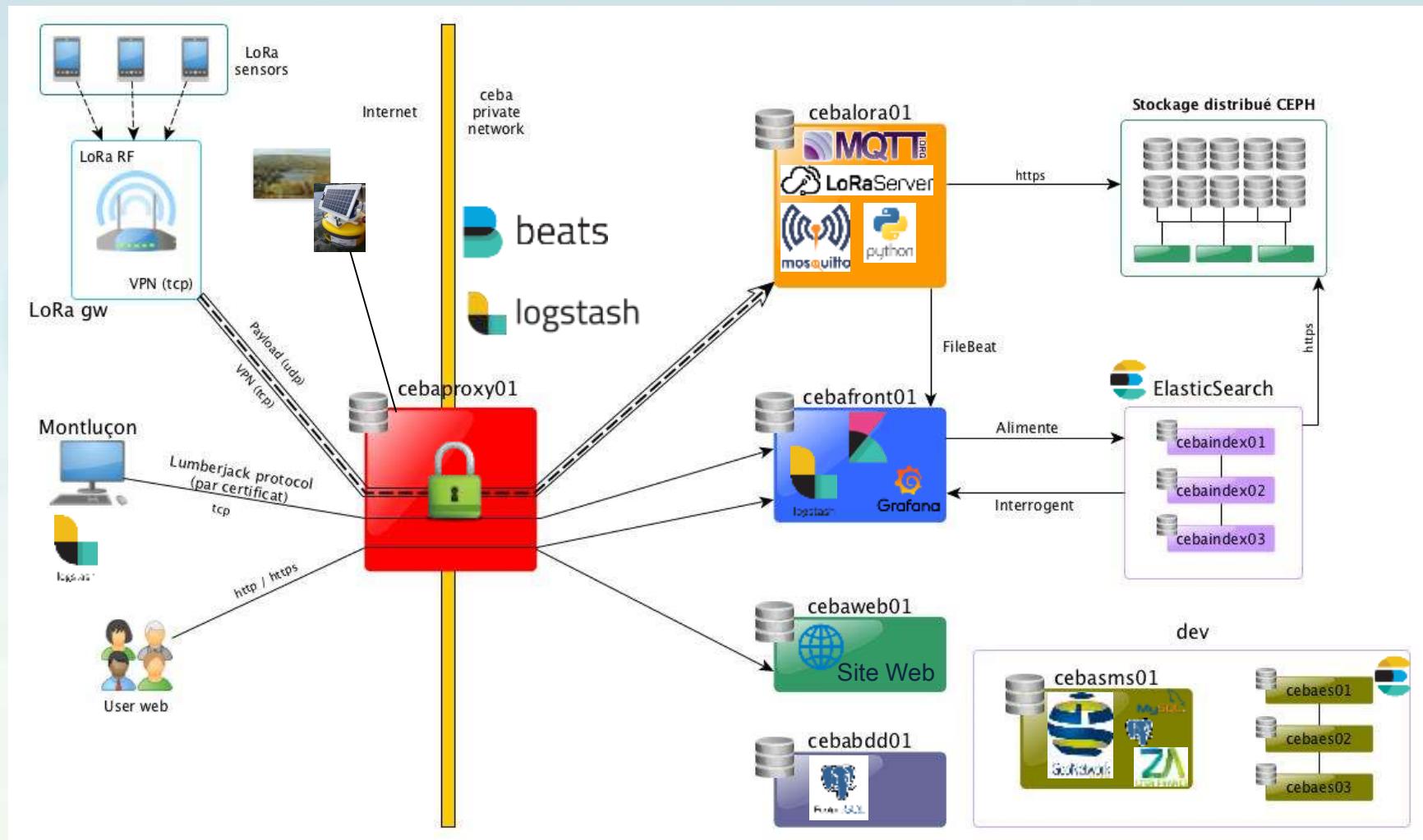
Équipe

- **Vincent Breton** Coordination
- **Alexandre Claude** Suite elastic & architecture
- **Gilles Mailhot** Coordination
- **Jérémy Mezhoud** Base de données
- **Francis Ogereau** Web, data-catalog
- **David Saramia** Suite elastic
- **Estelle Théveniaud** Geolab
- **Antoine Mahul** Mésocentre
- **David Grimbichler** Mésocentre

Merci de votre attention

Éléments du CEBA

Architecture technique générale



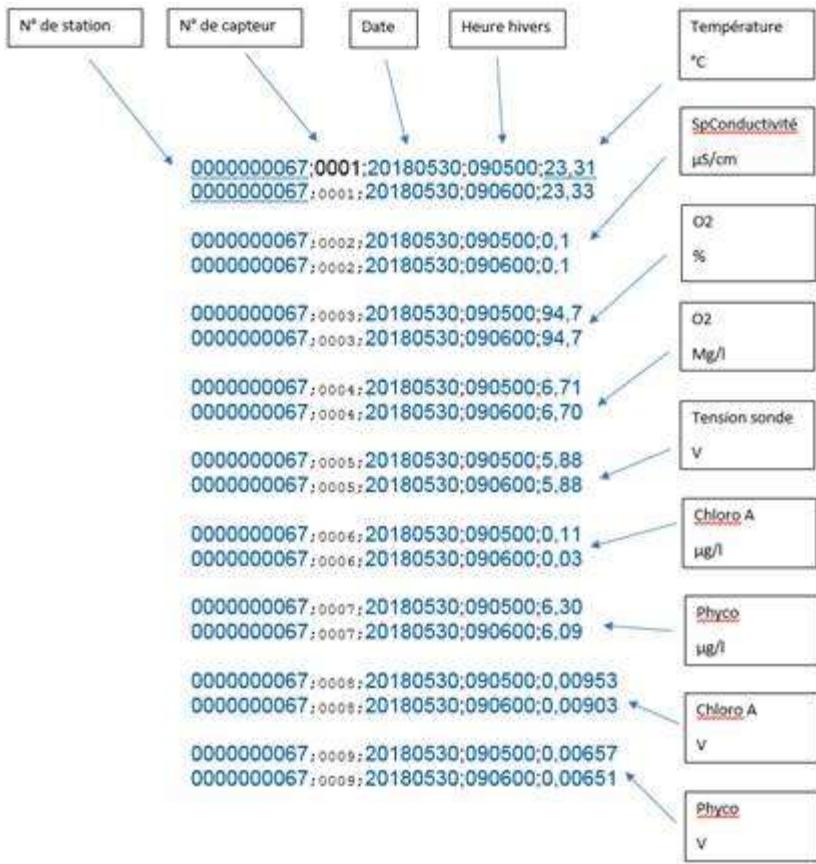
Éléments du CEBA

Traitement de flux de capteurs

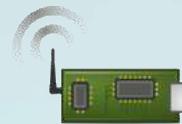
Aydat - bouée



Description des données



Flux



Capteurs



Serveur

Table _INDEX	
o _timestamp	2018-05-21T14:00:00.000Z
t _version	1
s _id	0000000067-21v0f6fk3u6
_index	bouee_aydat_v1_full
# _score	-
# _type	bouee_aydat
: dateheure	20180521
: heurehiver	140000
: host	192.168.1.100
: message	0000000067:0004:20180530:090500:9,01
: measures	9,01
: numCapteur	0004
: numStation	0000000067
: path	/Volumes/OSFILE/documents/PROJET/LPC/radioecologie/elasticsearch/tests/boueeaydat/test_index_018-05-21_10000.dat

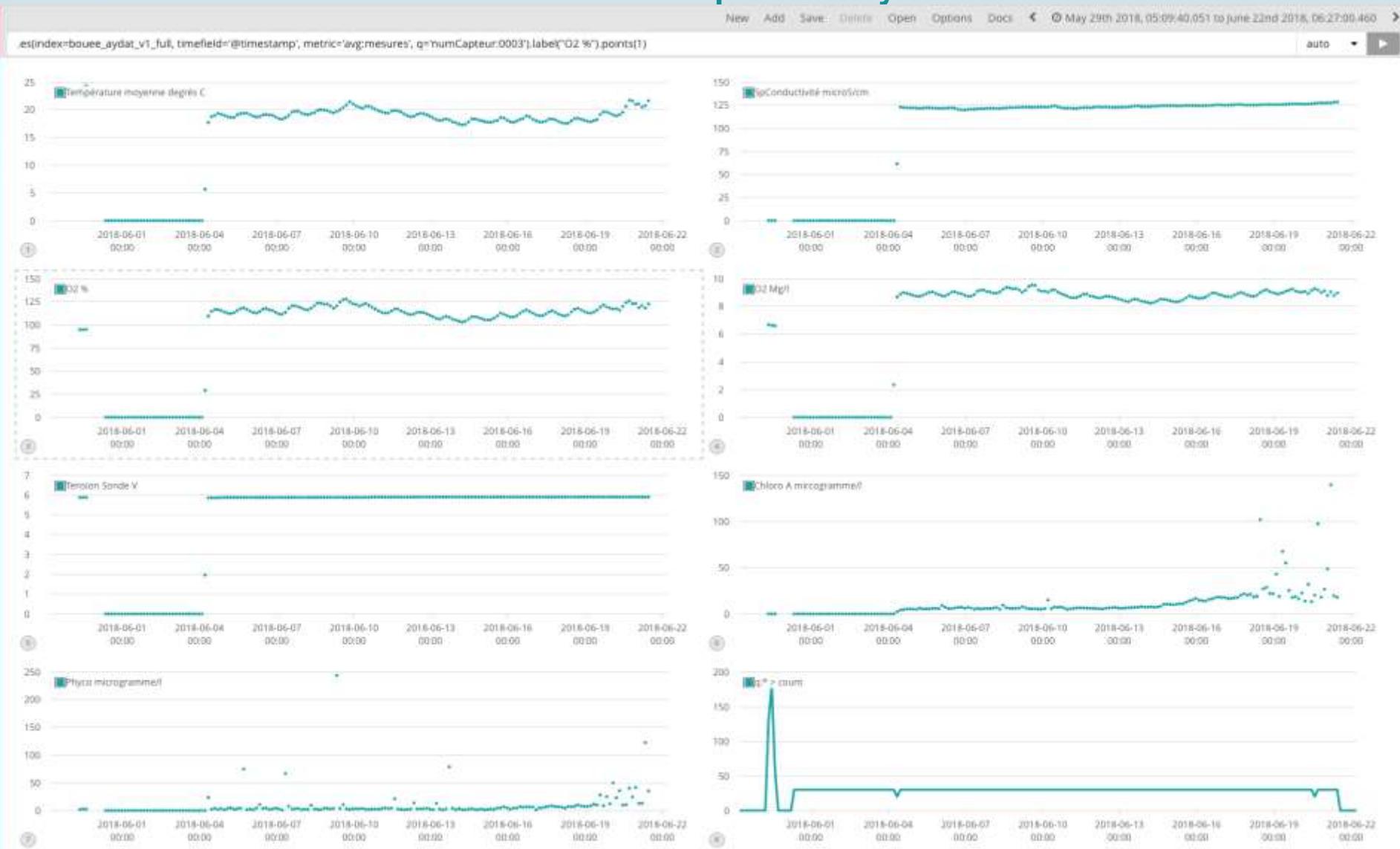


I-SITE
CLERMONT
Clement Auvergne Project

Graphique de plusieurs capteurs @ nombre de mesures

Éléments du CEBA

Traitement de flux de capteurs - Aydat - bouée



Éléments du CEBA

Traitement de flux de capteurs - offline ZATU - radioactivité

Description des données

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
Latitude GPS	Longitude GPS	Date	Fuseau horaire	Utilisateur	Mesure	Unité	Dose cumulée	Unité dose cumulée	Coefficient de calibration	Source de calibration	Date de calibration	Validité de calibration (Mois)	Nom de la sonde	N° de série de la sonde	Information détecteur	Firmware	N° de série du Colibri
2.000000	0.000000	2017/06/15 10h43m28s	GMT +1		0.000000258974	Sv/h	0.000000000359	Sv	0.012200	5e/h/c/s	2016.12.12	12.SVLD	SN0712	2688.00	13.04.16	SN0712	

Mesures provenant d'un Colibri (csv) :

- Positionnement GPS de chaque mesure
- Niveau de radioactivité



Flux

Fichiers

