



I-SITE
CLERMONT
Clermont Auvergne Project

CEBA

Cloud Environnemental au Bénéfice de l'Agriculture

Echanges 09/02/2021
TerraForma

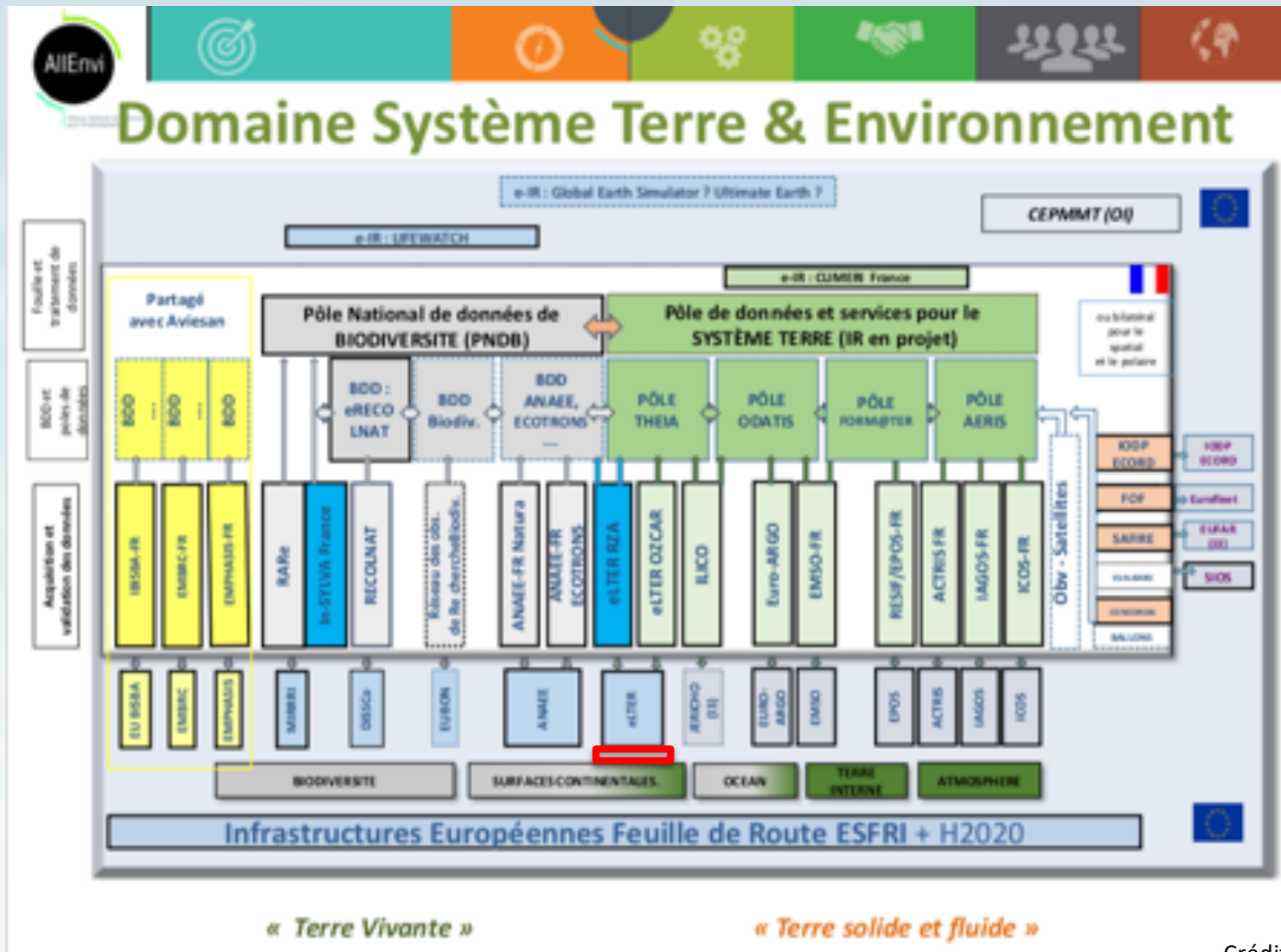


Qu'est ce que le CEBA ?

Cloud Environnemental

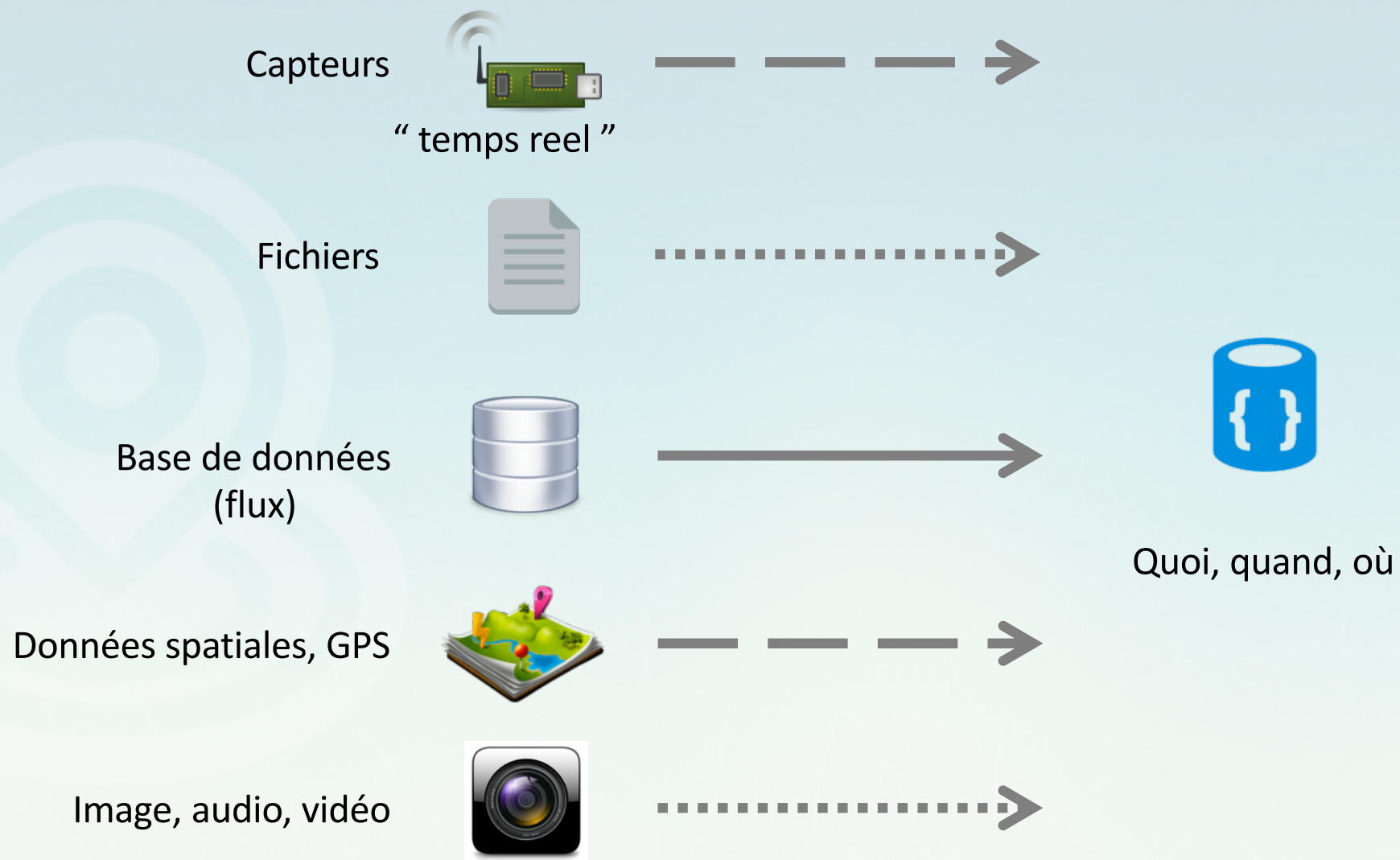
- Délivrable du Challenge 1 de l'I-Site de Clermont : Agro-écosystèmes durables dans un contexte de changement global
- Une ambition importante : Création d'un « grand » **observatoire de l'environnement** en Auvergne, unique en Europe.
- Des objectifs scientifiques à fort impact :
 - ➡ – Concevoir un **environnement numérique** interconnecté valorisant les données environnementales existantes ou à acquérir.
 - ➡ – **Automatiser le monitoring** de données à distance par la mise au point de nouveaux systèmes de **capteurs sans fil**.
 - Accroître notre **compréhension des compartiments environnementaux** face au changement global (évolutions, interactions).
 - Comprendre **l'impact de ces compartiments sur les agroécosystèmes** et réciproquement.
 - Conceptualiser les **interrelations entre les compartiments** pour comprendre, modéliser et prédire le comportement des agroécosystèmes.
 - Accompagner **l'aide à la décision** (agriculteurs, entreprises, collectivités territoriales).

Le contexte national



Crédits Philippe Mauguin

CEBA - Les données à (in)gérer



CEBA – concept global

COLLECTER

TRANSFORMER, VISUALISER, ANALYSER

INGERER → STOCKER

PUBLIER, DISTRIBUER

CEBA - Infrastructure

Ingérer ●

Stocker ●

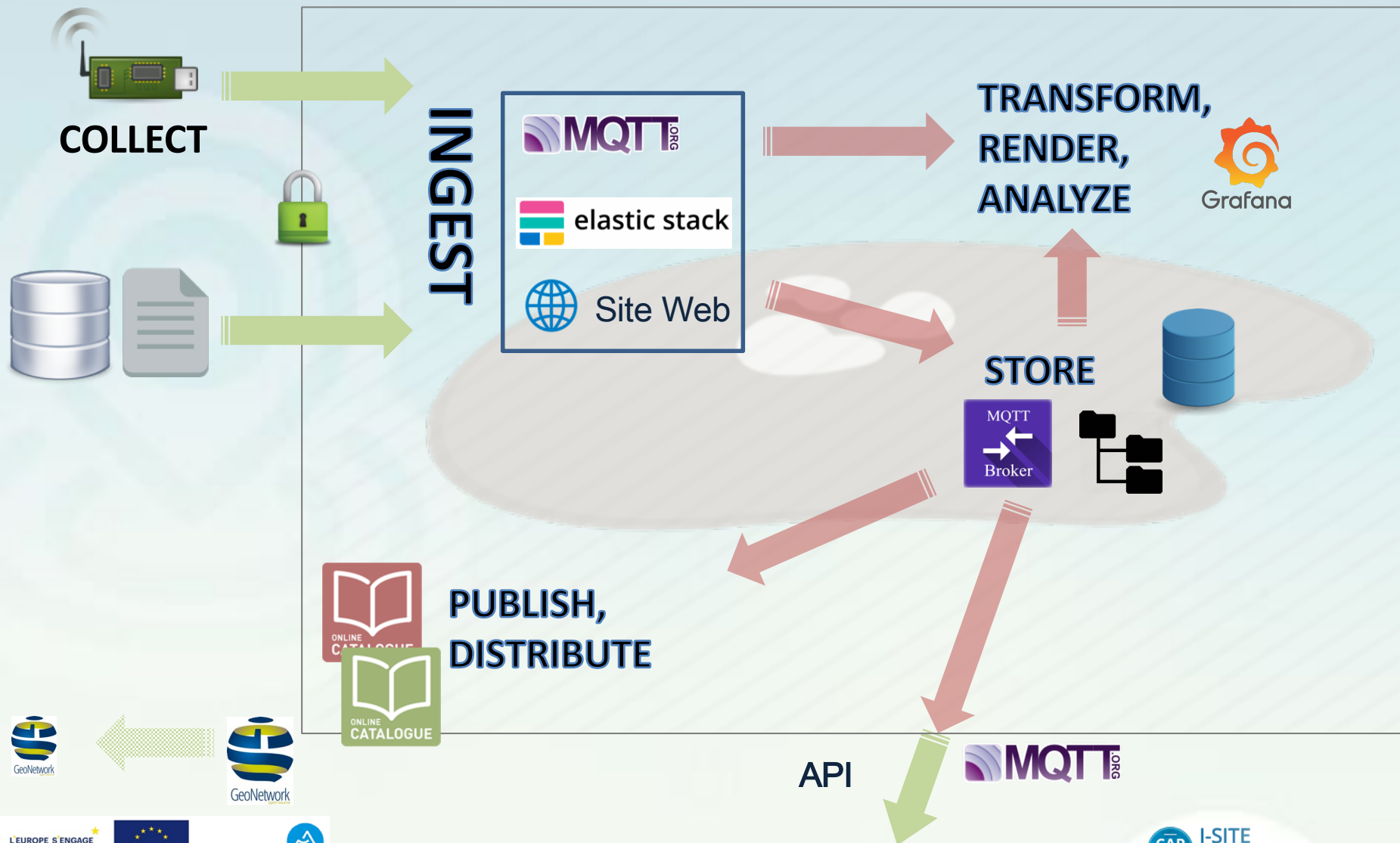
Visualisation /
restitution/
consultation ●

Partager ●

Site Web	● ● ● ●		
Outils d'ingestion	●		 beats  logstash
Moteur d'indexation	● ● ●		 elasticsearch
Catalogue de données	● ● ●	 GeoNetwork	 Lucene
Outils de visualisation	● ●	 Grafana	 kibana
Bases de données	● ●	 PostgreSQL	
Système de fichiers	●		

Éléments du CEBA

Big picture



Le réalisé en cours de validation

- Nature des données
 - a. Réseau de capteurs
 - b. Limesurvey
 - c. Fichier
 - d. Fiche de métadonnées
- Stockage
 - Quoi : Fichiers, flux
 - Comment : Fichiers, base de données, indexes
- Restitution
 - Graphique (a)
 - Brute (a b c)
 - Complète / Sélection (a)
 - Donnée/métadonnée
- Partage
 - Site web
 - Projet
 - Catalogue



- Restitution
 - Flux « quasi » temps réel



ETNA

Statut Connexion à la base de données ok ✓

Accéder à l'interface Grafana :
Grafana

Récupérer toutes les données :

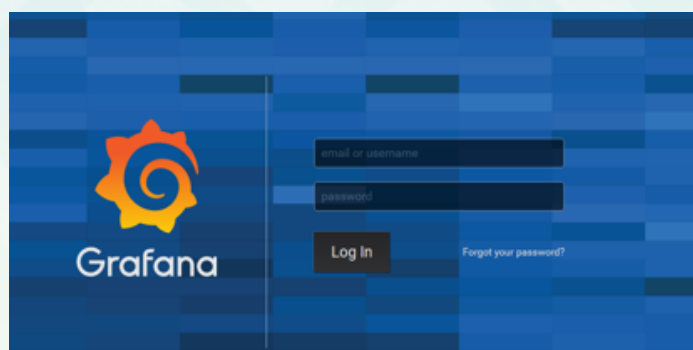
Sélectionner une plage temporelle :
Date début Date Fin Envoyer

Récupérer les données d'une mesure :

Préciser la mesure :
Sélectionner une plage temporelle :
Date début Date Fin Envoyer

Récupérer les données d'un devEUI :

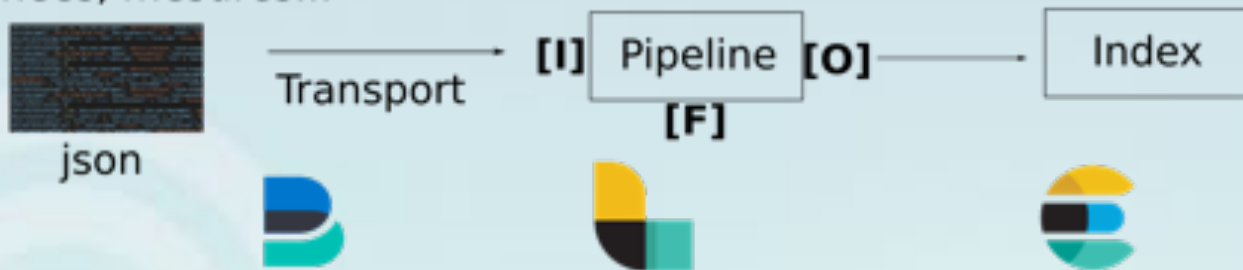
Préciser le devEUI : Envoyer



Éléments du CEBA

Traitement de flux de capteurs

Données, mesures...



```
type: log

# Change to true to enable this input configuration.
enabled: true

# Paths that should be crawled and fetched. Glob based paths.
paths:
  - /_show/
  - /api/status/other

# Run keep_alive every: true
# Run add_server_log: true

# Fields:
# log_type: database-audit

# Run keep_alive every: true
# Run add_server_log: true

processors:
  add_location:
    format: abbreviation
```

```
input {
  beats {
    port => 5042
    ssl => true
    ssl_certificate => "/etc/logstash/
    ssl_key => "/etc/logstash/logstash
    ssl_verify_mode => "force_peer"
    ssl_certificate_authorities => [
  ]
}
}
filter {
  if [type] == "log" {
    grok {
      match => { "message" => "%{SYSLOG}"
      add_field => [ "received_at", ""
      add_field => [ "received_from", ""
      remove_field => ["message"]
    }
    syslog_pri {
    }
    date {
      match => { "syslog_timestamp", "%Y-%m-%d %H:%M:%S"
    }
  }
}
}
output {
  if [fields][log_type] == "message" {
    elasticsearch {
      hosts => ["192.168.1.100:9200"]
      index => "cebalora-001-message-%{type}"
    }
  }
  if [fields][log_type] == "test-data" {
    elasticsearch {
      hosts => ["192.168.1.100:9200"]
      index => "cebalora-001-test-data-%{type}"
    }
  }
  if [fields][log_type] == "data-lora" {
    elasticsearch {
      hosts => ["192.168.1.100:9200"]
      index => "cebalora-001-data-lora-%{type}"
    }
  }
}
```

Time Filter field name: `json.server.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)	
Q Filter Apply					
Name	Type	Format	Searchable	Aggregatable	
@timestamp	date		●	●	
@version	string		●		
@version.keyword	string		●	●	
_id	string		●	●	
_index	string		●	●	
_score	number				
_source	string				

Équipe

- Jean-Claude Bailly
- Alexandre Claude
- Gilles Mailhot
- Jérémie Mezhoud
- Francis Ogereau
- David Sarramia
- Antoine Mahul
- David Grimbichler

Infrastructure

Suite elastic & architecture

Coordination scientifique

Base de données

Web, data-catalog

Coordination & Suite elastic

Mésocentre

Mésocentre

Retours d'expérience

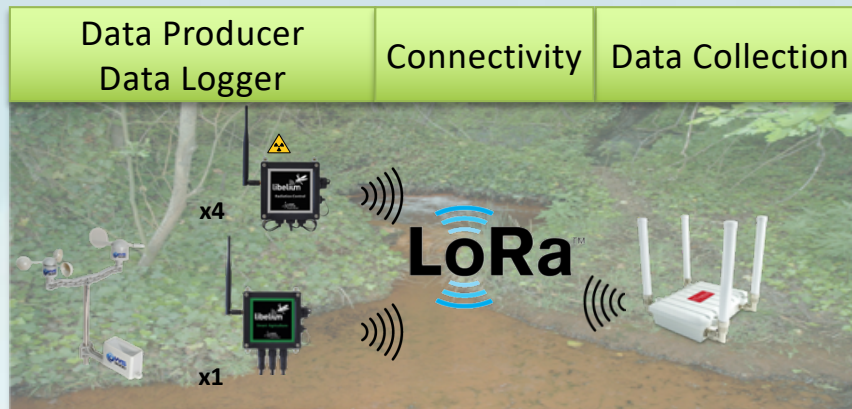
Exemples

- Réseau de capteurs sans fils
 - Mesures météo & radioactivité (bd) 
 - Nœud sans fil (json) 
- Zatu
 - Position des arbres (csv) 
 - Mesures de niveau de radioactivité (csv) 
- Bouée Aydat
 - Mesures (txt) 

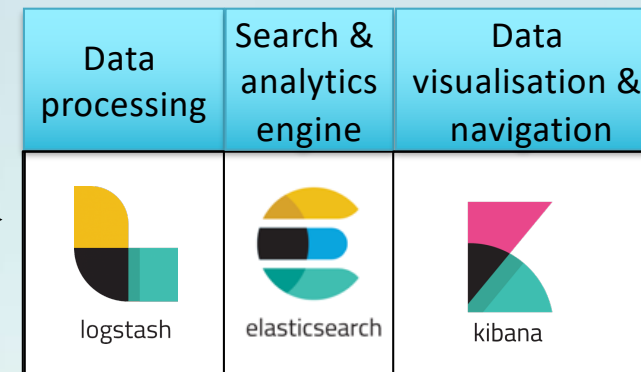
Réseau de capteurs sans fils



Réseau de capteurs V1



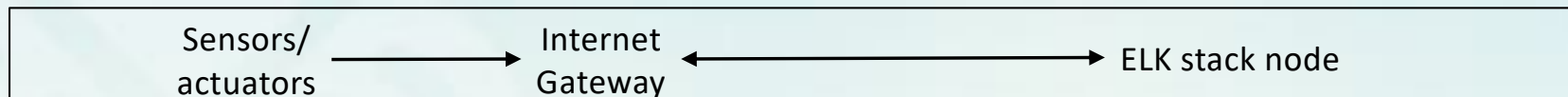
Mesocentre Auvergne



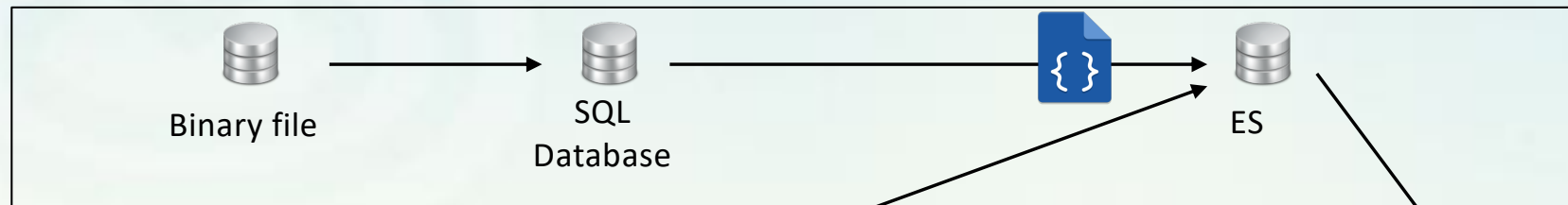
End User



Network



Storage



Fichiers

Third party data



Prélèvements, mesure



Réseau de capteurs sans fils

Description des données v1



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	JSON
@timestamp	December 10th 2018, 16:12:44.000
@version	1
_id	325482
_index	mtlsensors
_score	2
_type	sensorDatawasp
frame_number	31
id	325,482
id_secret	400578826
id_wasp	RAD02
sensor	BAT
timestamp	2018-12-10 16:12:44
value	97



elasticsearch



kibana

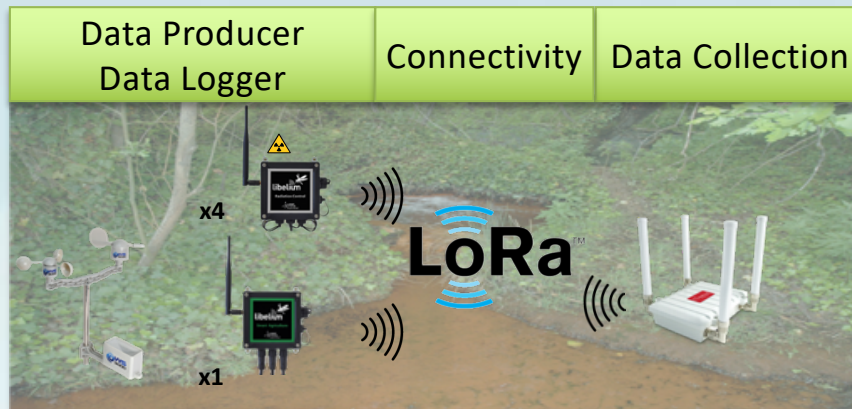
Création d'alias pour
faciliter les
requêtes/visualisation

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

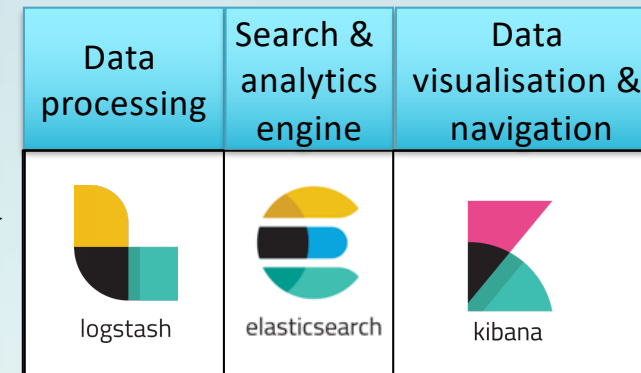
Réseau de capteurs sans fils



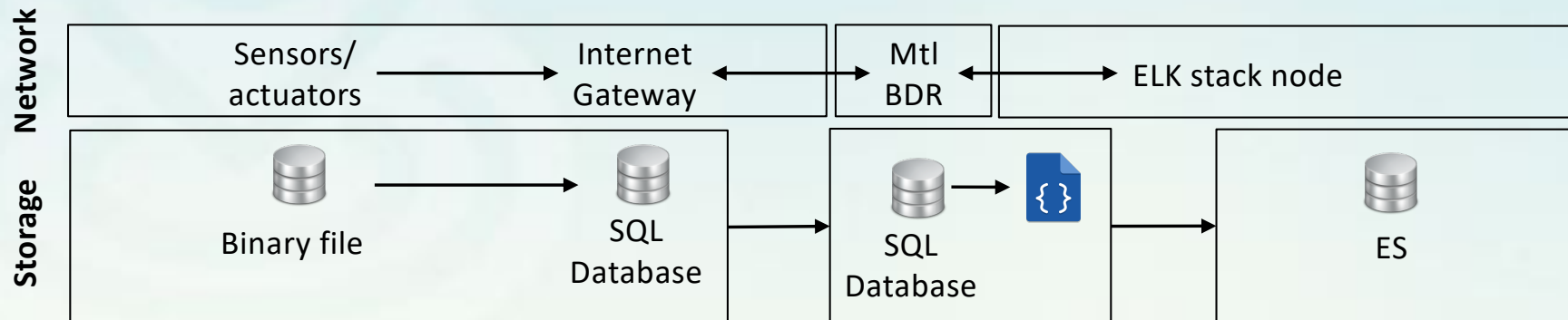
Réseau de capteurs V2



Mesocentre Auvergne



End User



Serveur Montluçon

Réseau de capteurs sans fils

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 : 7276ff000b030f4c
  data-node-batteryVoltage : 3.92
  data-node-batteryVoltage-unit : V
  rxInfo-latitude : 0
  rxInfo-loRaSNR : 9
  txInfo-dataRate-spreadFactor : 7
  rxInfo-altitude : 0
  txInfo-frequency : 868300000
  data-temperature-alarm : 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

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-rssi": -26, "devEUI":
  "434e535301e31294", "data-illuminance": 340, "applicationI
  mac": "7276ff000b030f4c", "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
r _id	14
r _index	Flattened_4_1
# _score	1
r _type	type4
r applicationID	1
r applicationName	ConnecSens-Nodes
# data-DataChannel	1
# data-airHumidity	40.5
r data-airHumidity-unit	%
# data-atmosphericPressure	982.14
r data-atmosphericPressure-unit	hPa
# data-illuminance	251
r data-illuminance-unit	lux



logstash



elasticsearch



kibana

ZATU - arbres

Positionnement des arbres sur lesquels un prélèvement a été réalisé

Description des données

```
3.552079,46.010464,1,upstream
3.552032,46.010206,6,upstream
3.551967,46.009947,7,upstream
3.551369,46.008169,2,downstream
3.55073,46.007882,3,downstream
3.550334,46.007557,5,downstream
```

Flux



logstash

Fichiers

Table	JSON
t _id	AwPgaqx09t70v2bKNvBq
t _index	tree_rophin_info_v1
# _score	1
t _type	tree
location	{ "lon": "3.552079", "lat": "46.010464" }
t position	upstream
t tree_number	1



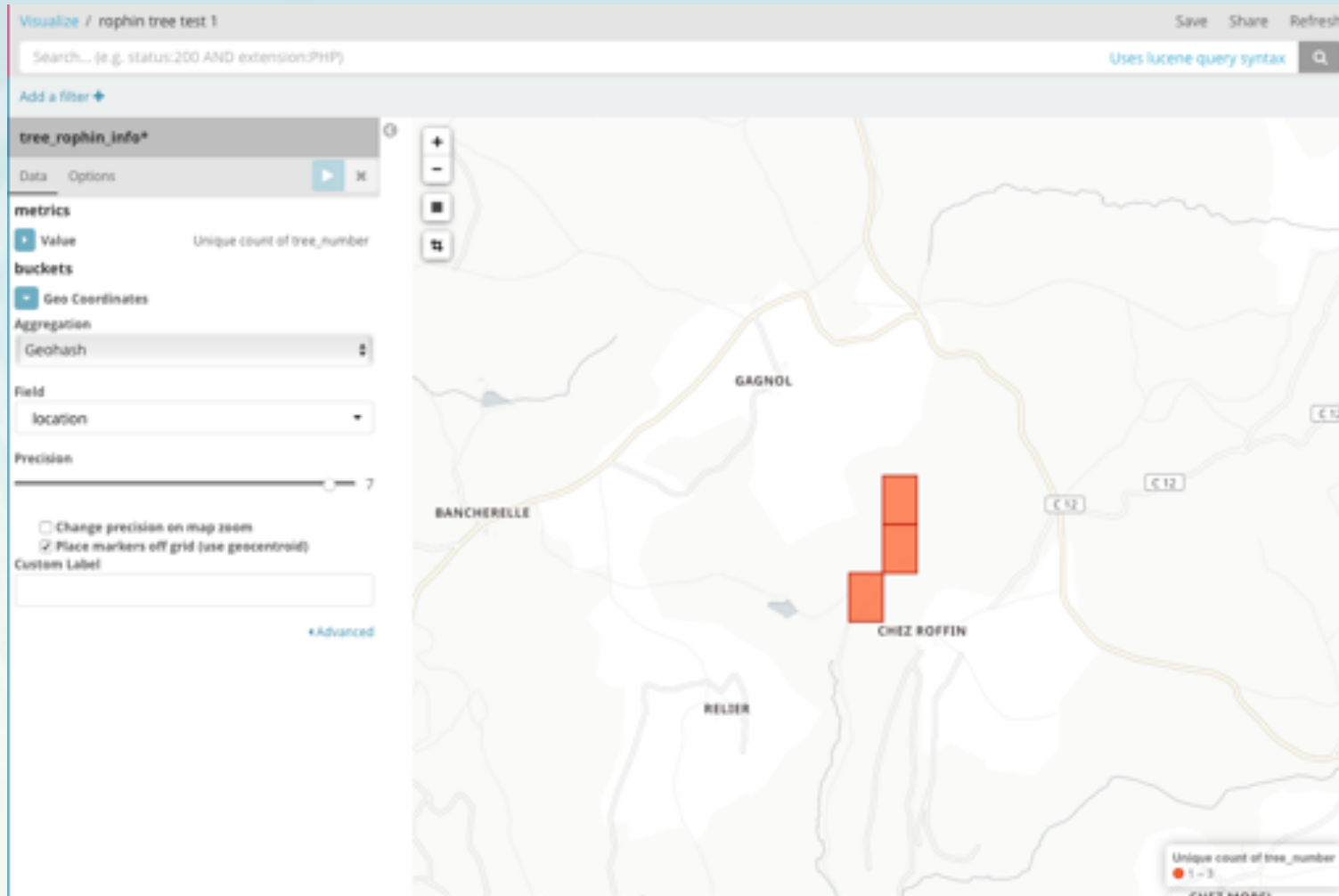
elasticsearch



kibana

ZATU - arbres

Positionnement des arbres sur lesquels un prélèvement a été réalisé



ZATU - radioactivité

Description des données

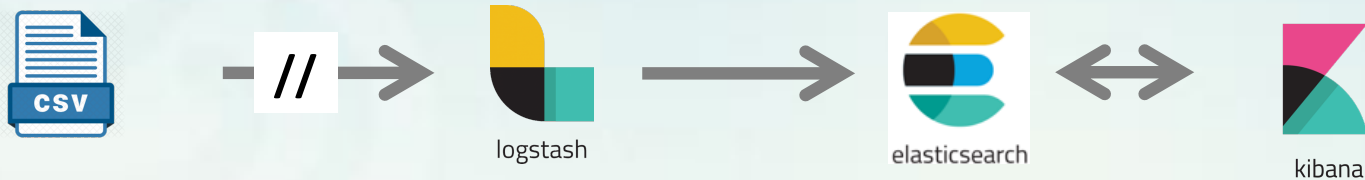
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Latitude GPS	Longitude GPS	Date	Fuseau horaire	Utilisateur	Mesure	Unité	Dose cumulée	Unité dose cumulée	Coefficient de calibration	Source de calibration	Date de calibration	Validité de calibration (Mois)	Nom de la sonde	N° de série de la sonde	Information détecteur	Firmware	N° de série du Colibri
2	0.000000	0.000000	2017/06/15 10h:43m:28s	GMT +1		0.000000258974	Sv/h	0.000000000359	Sv	0.012200µSv/h/c/s	Cs137	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

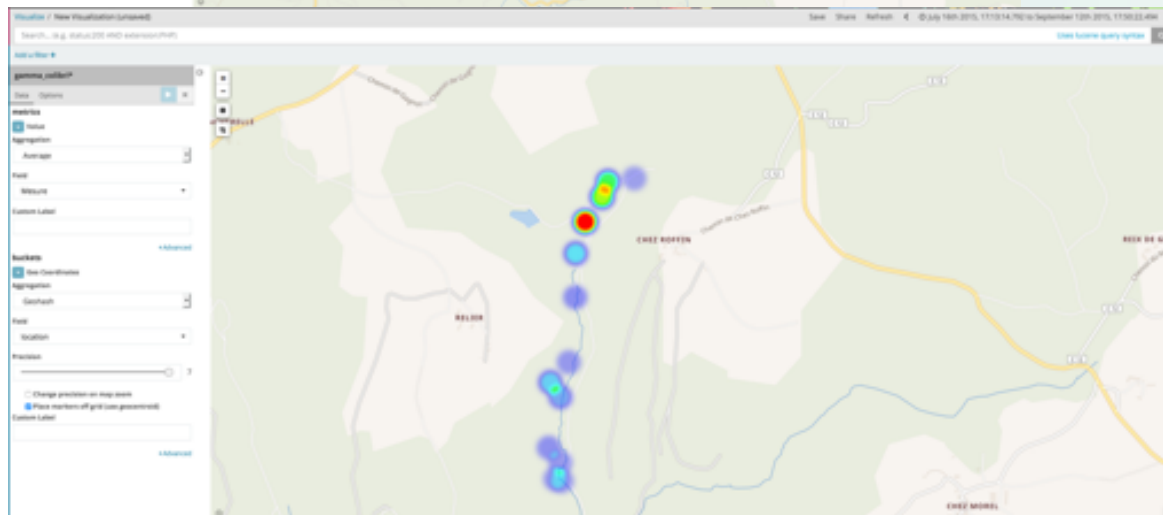
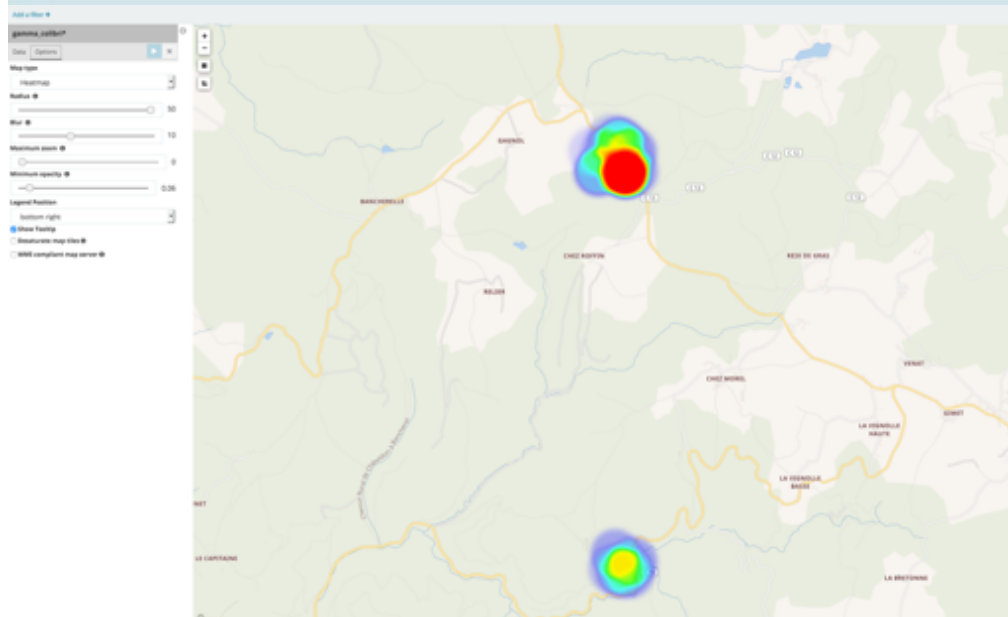
Fichiers



ZATU - radioactivité

Mesures provenant d'un Colibri (csv) :
Positionnement GPS de chaque mesure
Niveau de radioactivité

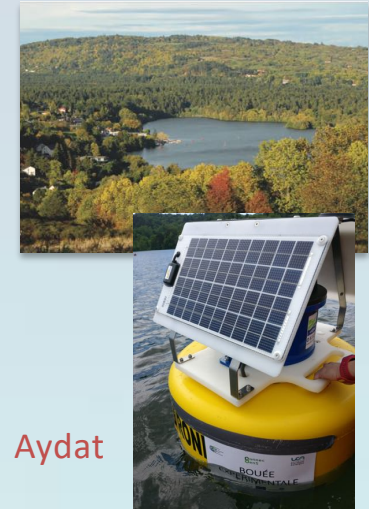
(couleur = niveau relatif)



GET gamma_colibri_rophin/_search

```
{
  "query": {
    "bool": {
      "must": {
        "match_all": {}
      },
      "filter": {
        "geo_polygon": {
          "location": {
            "points": [
              [3.5459891613159016, 46.01295812457846],
              [3.5592929180297688, 46.01295812457846],
              [3.5592929180297688, 46.002704419144706],
              [3.5459891613159016, 46.002704419144706],
              [3.5459891613159016, 46.01295812457846]
            ]
          }
        }
      }
    }
  }
}
```

Aydat - bouée

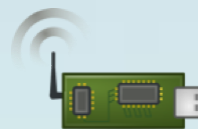


Description des données



N° de station	N° de capteur	Date	Heure Hiver	Température °C
0000000067	0004	20180630	090600	23.31
0000000067	0004	20180630	090600	23.33
0000000067	0004	20180630	090600	0.1
0000000067	0004	20180630	090600	0.1
0000000067	0004	20180630	090600	94.7
0000000067	0004	20180630	090600	94.7
0000000067	0004	20180630	090600	6.71
0000000067	0004	20180630	090600	6.79
0000000067	0004	20180630	090600	5.88
0000000067	0004	20180630	090600	5.88
0000000067	0004	20180630	090600	0.11
0000000067	0004	20180630	090600	0.03
0000000067	0004	20180630	090600	6.30
0000000067	0004	20180630	090600	6.09
0000000067	0004	20180630	090600	0.00953
0000000067	0004	20180630	090600	0.00903
0000000067	0004	20180630	090600	0.00657
0000000067	0004	20180630	090600	0.00651

Flux



Capteurs

Aydat



Serveur



Mesocentre Auvergne



Table	2506	View surrounding documents	View single document
@timestamp	June 21st 2018, 14:00:00.000		
@version	1		
_id	AwqnfC--21v016fk3A67		
_index	bouee_aydat_v1_full		
_score	-		
_type	bouee_aydat		
dateMesure	20180621		
heureHiver	140000		
host	lMac-de-dasarrram-2.local		
message	0000000067;0004;20180621;140000;9.01		
# mesures	9.01		
numCapteur	0004		
numstation	0000000067		
path	/Volumes/DSFILES/Documents/PROJET/LPC/radioecologie/elasticsearch/tests/boueeaydat/test_sed/2018-06-21_130111.dat		



logstash



elasticsearch



CAP 20-25
I-SITE CLERMONT
Clermont Auvergne Project

Indexation

data nodes

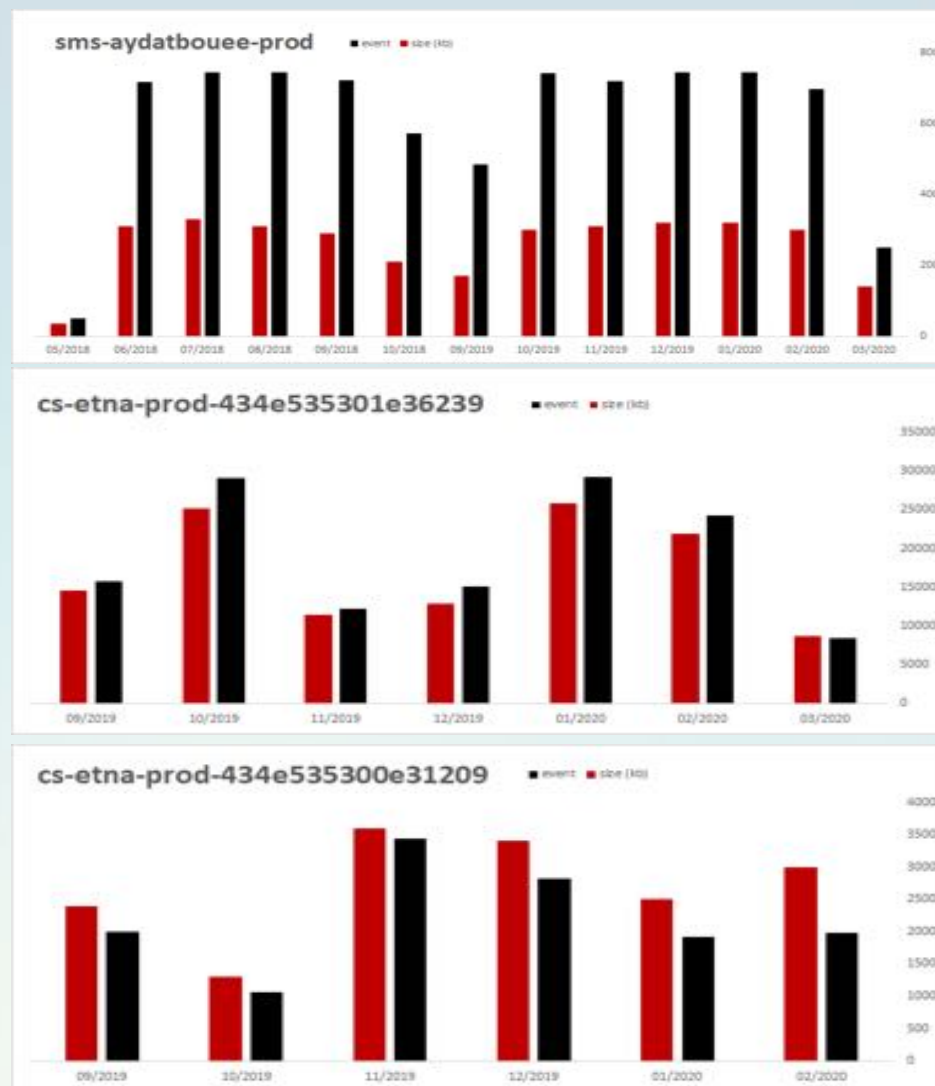
disk.total	disk.used	disk.available	used_percent
3.9gb	514.3mb	3.4gb	12.59
3.9gb	518.4mb	3.4gb	12.69

Stats

- 133 (88) index
- 954997 documents
- 28008+28049 opérations d'indexation
- 25991+27098 temps d'indexation (ms)

Usage

- mensuel
- Proto_Projet_statut_tech_Date
- **Dashboard « temps réel » - Assistance au terrain, réactivité**



Ce qui est important

- Expliquer la donnée
 - On ne connaît pas
 - Donnée autodescriptive
- Savoir ce que l'on veut faire
 - Science et/ou Ingénierie
 - Traitement (manque de données)
 - Standards (interopérabilité infra externe)
 - Partager/publier/pousser
- Connaître la production de donnée
 - Fréquence
 - Volume
 - Evolution

Merci de votre attention

