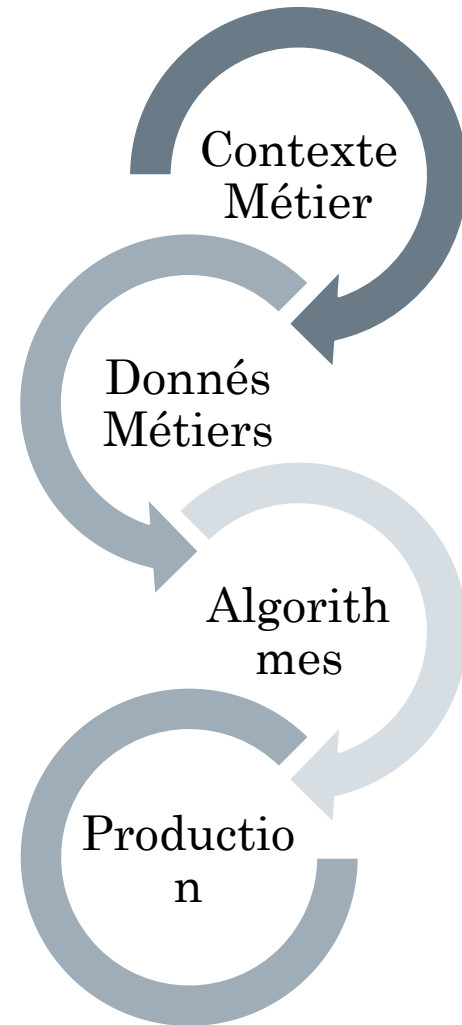


# MARIA

Maîtrise Augmentée du Réseau par l'Intelligence Artificielle

Pierre Rebours

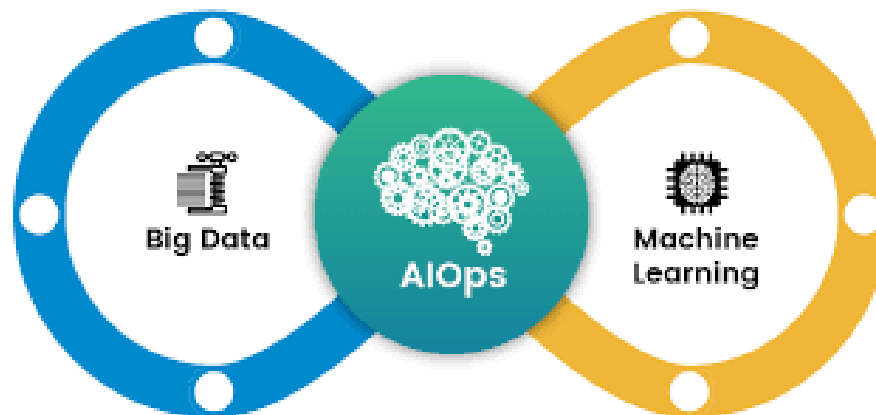
# Objectifs



# Contexte



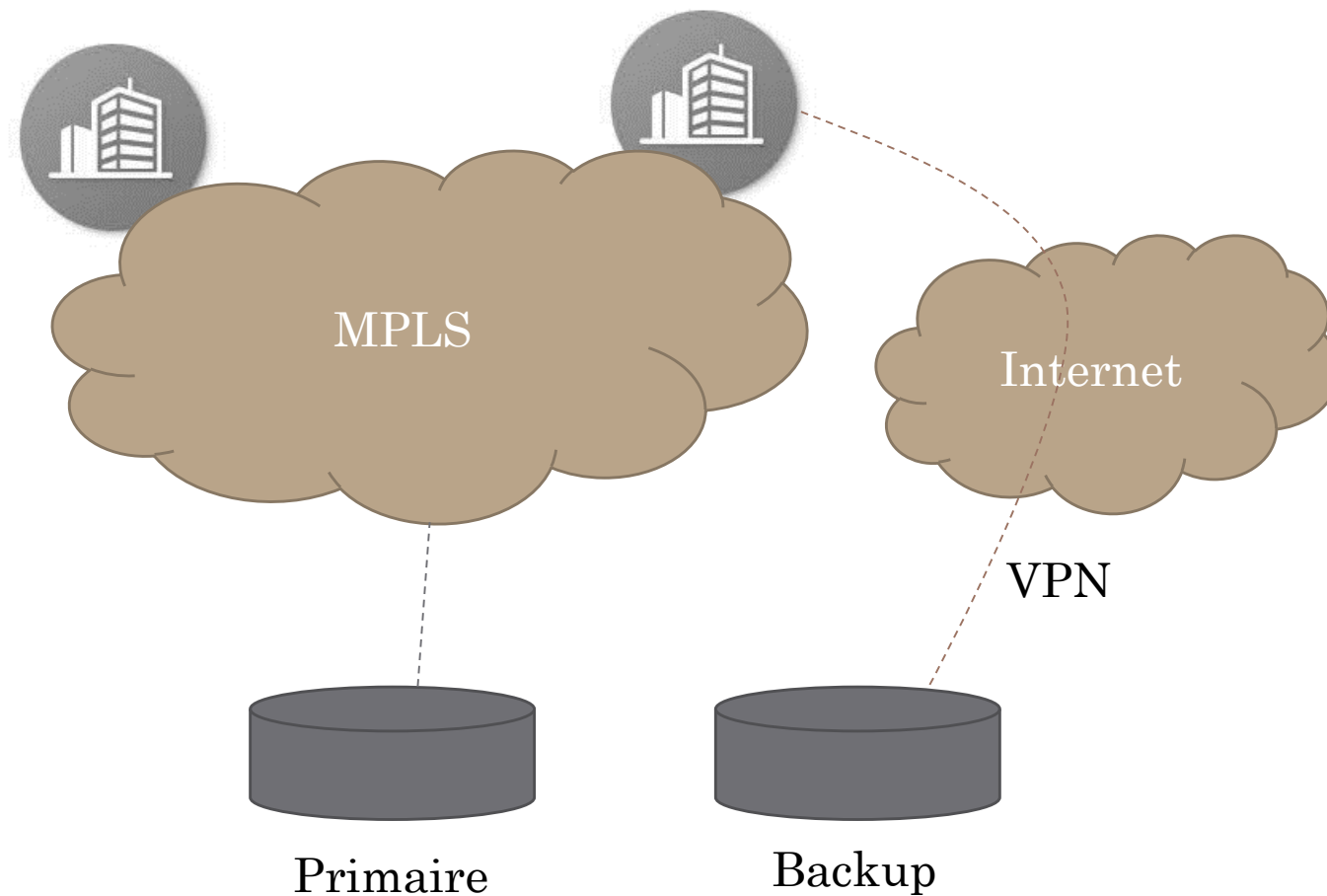
- Supervision des systèmes informatiques
- Les 3V: Volume / Velocity / Variety
- Problème: Difficulté d'interpréter



# Les données



- Disponibilité
- Performance
- Consommation
- Type de flux



# Prédiction de Saturation



M . A . R . I . A

Méthode Augmentée du Réseau par Intelligence Artificielle

**NUMBER OF SITES**  
204

**213 025€**  
Saturation impact on the business for the rest of the day

**83.00%**  
Prediction Accuracy (Last Week)

Settings +

## Predictions

Overview

Map

Opening Hours

Show 10 entries

Search:

site_code	daily_turnover	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00
FR045	€118,170	99.65%	22.38%	93.21%	86.52%	86.60%	88.56%	38.50%	62.33%	77.20%
FR112	€198,454	74.01%	89.65%	95.35%	44.21%	32.82%	89.86%	93.00%	79.65%	29.66%
FR078	€206,451	61.07%	93.70%	62.97%	74.85%	90.96%	85.69%	56.01%	73.36%	21.06%
FR230	€139,444	97.90%	33.53%	86.71%	62.49%	88.73%	66.51%	35.62%	53.36%	86.16%
FR070	€179,249	4.01%	77.22%	74.81%	55.76%	67.15%	95.71%	84.06%	66.63%	82.05%
FR181	€31,477	82.25%	12.34%	85.57%	78.92%	74.13%	14.14%	73.03%	84.75%	94.62%
FR116	€204,700	25.56%	97.34%	87.09%	4.66%	59.85%	93.86%	77.94%	48.49%	98.10%
FR046	€167,556	86.00%	20.64%	97.24%	56.49%	56.80%	67.11%	65.06%	86.45%	47.04%
FR100	€116,013	70.08%	82.77%	14.47%	87.00%	65.27%	83.22%	30.62%	63.80%	81.31%
FR104	€192,019	65.52%	86.16%	72.64%	83.14%	18.71%	64.40%	60.92%	76.84%	49.80%

Showing 1 to 10 of 204 entries

Previous 1 2 3 4 5 ... 21 Next

Display the predictions for the current opening hours for the site.

## Analysis

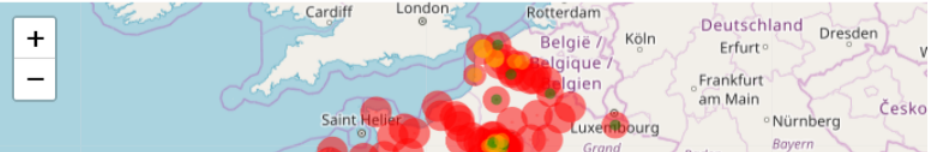
### Diagnostics

Category	Name	Impacted Sites
Production	Saturation Risk	82 sites
Continuous Improvement	Behavior Segmentation	45 sites

Map

Diagram

Overview



Global Overview

Site Overview

Predictions

Outliers

COS Configuration **live**

Plateau Issues **live**

Previous Forward

Slider  
07:00:00 12:00:00 21:00:00

Date  
2017-07-18

Powered by



M . A . R . I . A  
Machine Augmentée du Réseau par Intelligence Artificielle

# Prediction (FR063)

Settings +

Introduction -

TODO

## Prediction

Legend	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00
actual / predicted	0%	0%	0%	0%	0%	76%	73%	52%	72%	23%	14%	7%	15%	14%

Display the upcoming prediction alongside the past predictions.

## Assess

Explore model

**SENSITIVITY**

**No saturation encountered**

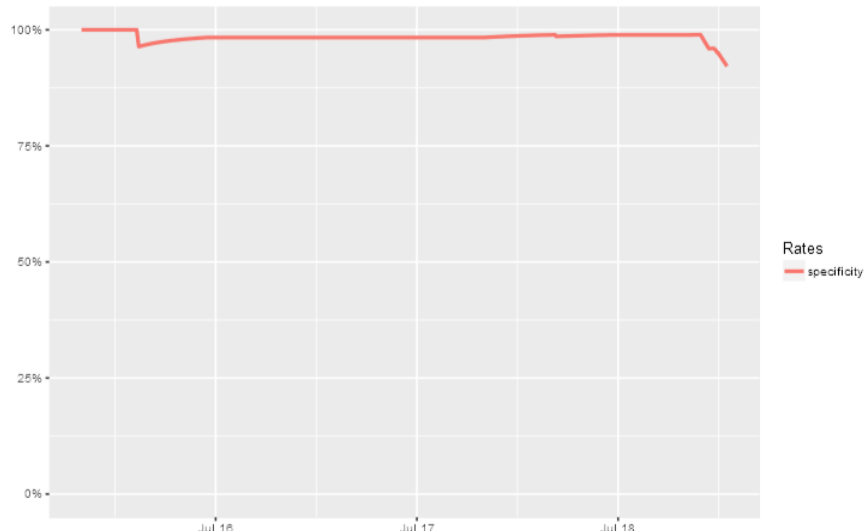
Proportion of positives correctly identified (true positive rate). TP: 0/FN: 0

**SPECIFICITY**

**92.16%**

Proportion of negatives correctly identified (true negative rate). TN: 388/FP: 33

Horizon Matrix View Data Prediction Overtime



## Understand & Act

Recommendations Variable Importance Variable Details

### Understand

- 82% of the pass saturaton were due on traffic on the D3 COS
- 32% of the payload is related to internet traffic.
- Long-trend analysis and short-term topology likely to cause saturation: check the variable importance of the model.

### Communicate

- A migration projet is starting in 2 hours. Please contact the project team to mitigate the situation.
- The IT local team is reachable on the following [phone line](#).

### Act

- [Access](#) the netflow for this site.
- [Access](#) the Insight for this site.
- [Access](#) the high-level reporting.
- [Launch](#) the in-depth diagnostics tools for the primary link.

**939€**

Impact on the business for the day

**939€**

Predicted Impact on the business for the rest of the day

Global Overview

Site Overview

Predictions

Overview

Model

Data Explorer

Outliers

COS Configuration **live**

Plateau Issues **live**

Site

FR063

Previous Forward

Slider

07:00:00 12:00:00 21:00:00

07:00:00 11:00:00 15:00:00 19:00:00

Date

2017-07-18

2017-07-18 12:00:00

Powered by



# Prediction (FR063)

**Settings**

- Staggered Display
- Show Past Predictions

**Display mode**

Probability

**Introduction**

TODO

## Prediction

Legend	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00
actual	0%	0%	0%	0%	0%	0%	100%							
1h 0min 0s	20%	35%	53%	35%	68%	76%	62%	52%						
2h 0min 0s		31%	51%	27%	45%	67%	73%		68%					
3h 0min 0s			54%	57%	37%	32%	52%			27%				
4h 0min 0s				66%	61%	34%	57%				14%			
5h 0min 0s					42%	28%	77%					15%		
6h 0min 0s						32%	66%						14%	
7h 0min 0s							60%							22%

Display the upcoming prediction alongside the past predictions.

## Assess

Explore model

**SENSITIVITY**  
**100.00%**  
Proportion of positives correctly identified (true positive rate). TP: 17/FN: 0

**SPECIFICITY**  
**90.65%**  
Proportion of negatives correctly identified (true negative rate). TN: 388/FP: 40

Horizon | Matrix View | Data | Prediction Overtime

## Understand & Act

Recommendations | Variable Importance | Variable Details

**Understand**

- 82% of the pass saturaton were due on traffic on the D3 COS
- 32% of the payload is related to internet traffic.
- Long-trend analysis and short-term topology likely to cause saturation: check the variable importance of the model.

**Communicate**

- A migration projet is starting in 2 hours. Please contact the project team to mitigate the situation.
- The IT local team is reachable on the following [phone line](#).

**Act**

Global Overview

Site Overview

Predictions

Overview

Model

Data Explorer

Outliers

COS Configuration live

Plateau Issues live

Site

FR063

Previous | Forward

Slider

07:00:00 | 14:00:00 | 21:00:00

Date

2017-07-18

2017-07-18 14:00:00

Powered by





M . A . R . I . A

Machine Augmentée du Réseau par Intelligence Artificielle

# Prediction (FR063)

### Settings

- Staggered Display
- Show Past Predictions

### Display mode

Probability

### Introduction

TODO

## Prediction

Legend	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00
actual	0%	0%	0%	0%	0%									
1h 0min 0s	9%	5%	12%	27%	37%	47%								
2h 0min 0s		11%	12%	26%	40%		39%							
3h 0min 0s			19%	25%	37%			30%						
4h 0min 0s				26%	28%				35%					
5h 0min 0s					24%					29%				
6h 0min 0s											38%			
7h 0min 0s												41%		
8h 0min 0s													21%	
9h 0min 0s														26%

Display the upcoming prediction alongside the past predictions.

Explore model

## Assess

### SENSITIVITY

**No saturation encountered**

Proportion of positives correctly identified (true positive rate). TP: 0/FN: 0

### SPECIFICITY

**98.76%**

Proportion of negatives correctly identified (true negative rate). TN: 238/FP: 3

## Understand & Act

Recommendations

Variable Importance

Variable Details

### Understand

- 82% of the pass saturaton were due on traffic on the D3 COS
- 32% of the payload is related to internet traffic.
- Long-trend analysis and short-term topology likely to cause saturation: check the variable importance of the model.

### Communicate

Global Overview

Site Overview

Predictions

Overview

Model

Data Explorer

Outliers

COS Configuration

live

Plateau Issues

live

Site

FR063

Previous

Forward

Slider

07:00:00 12:00:00 21:00:00

07:00:00 11:00:00 15:00:00 19:00:00

Date

2017-07-17

2017-07-17 12:00:00

Powered by



M. A. R. I. A

Machine Augmentée du Réseau par Intelligence Artificielle

Global Overview

Site Overview

Predictions

Overview

Model

Data Explorer

Outliers

COS Configuration live

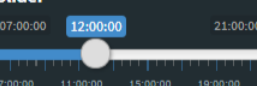
Plateau Issues live

Site

FR063

Previous Forward

Slider



Date

2017-07-17

2017-07-17 12:00:00

Powered by

**SENSITIVITY**  
**96.47%**  
 Proportion of positives correctly identified (true positive rate). TP: 301/FN: 11

**SPECIFICITY**  
**93.05%**  
 Proportion of negatives correctly identified (true negative rate). TN: 1887/FP: 141

## Experiment (FR063)

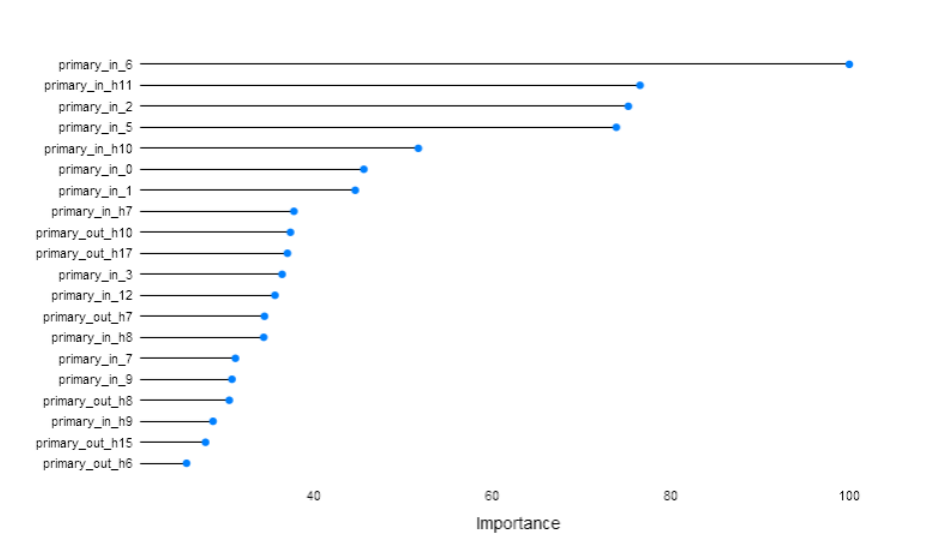
Settings +

### Model Selection

```
List of 4
$ site_code: chr "FR063"
$ period :List of 2
..$ from: Date[1:1], format: "2017-05-01"
..$ to : Date[1:1], format: "2017-07-31"
$ data :List of 4
..$ nb_observation_in_past: int 36
..$ features : chr [1:2] "primary_in" "primary_out"
..$ heatmap : logi TRUE
..$ future_def :List of 2
.. ..$ start_obs : num 12
.. ..$ length_obs: num 12
$ model :List of 6
..$ control_type : chr "none"
..$ summaryFunction: chr "multiple_summary_function"
..$ metric : chr "Kappa"
..$ sampling_when : chr "resampling"
..$ sampling_method: chr "down"
..$ method : chr "rf"
```

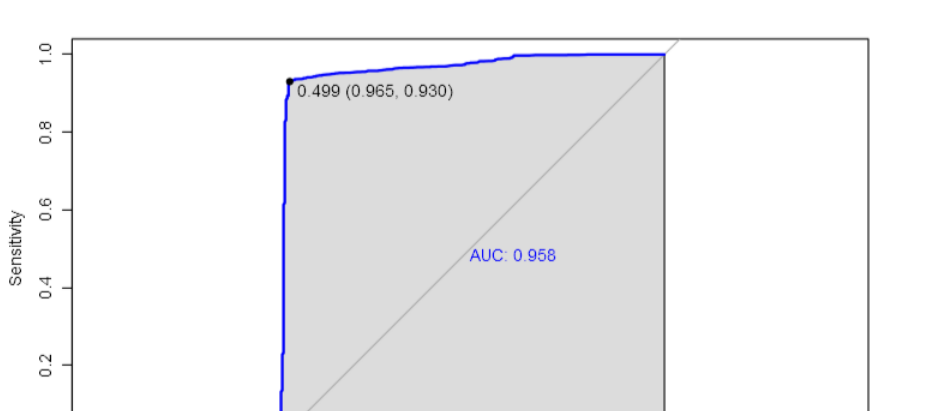
Describe the set-up of the experiment

### Variable Importance

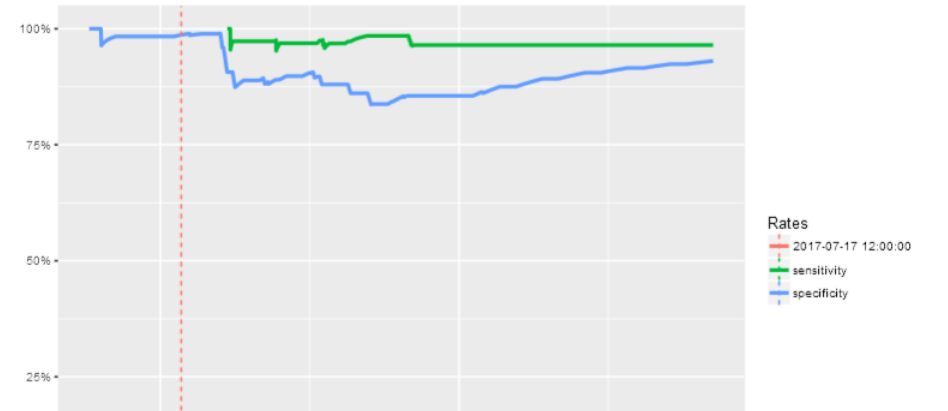


List the variables that impact the most the model

### Test Evaluation



### Test Overtime



# Prédiction de Saturation

- Objectif: Prédire si un site sera saturée à plus de 90% de sa capacité
- 1 modèle par site et par segment horaire
- Features: Données récentes et tendance historique sur de la consommation de bande passante mais aussi sur des signaux moins directs
- Techniques: resampling / métrique de sélection Kappa
- Packages: caret, rf

# Clustering de sites

07:00:00 11:00:00 15:00:00 19:00:00

Date

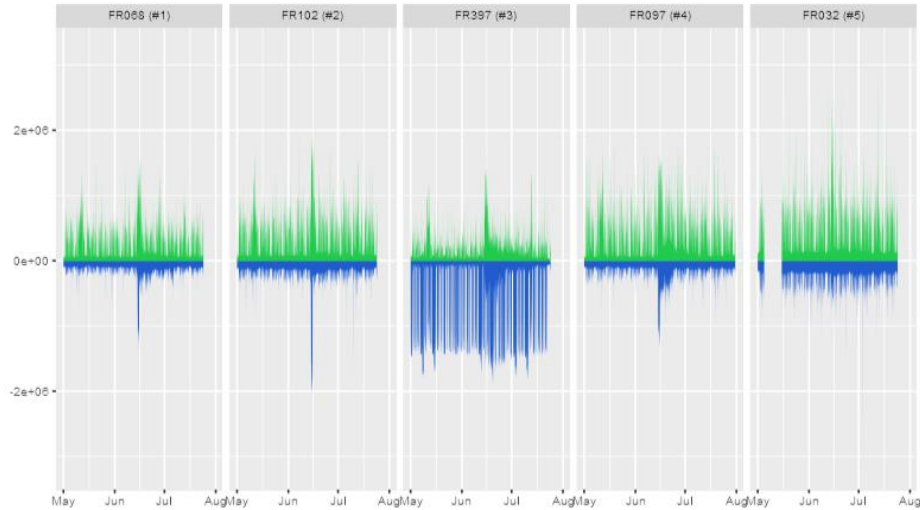
2017-07-17

2017-07-17 12:00:00

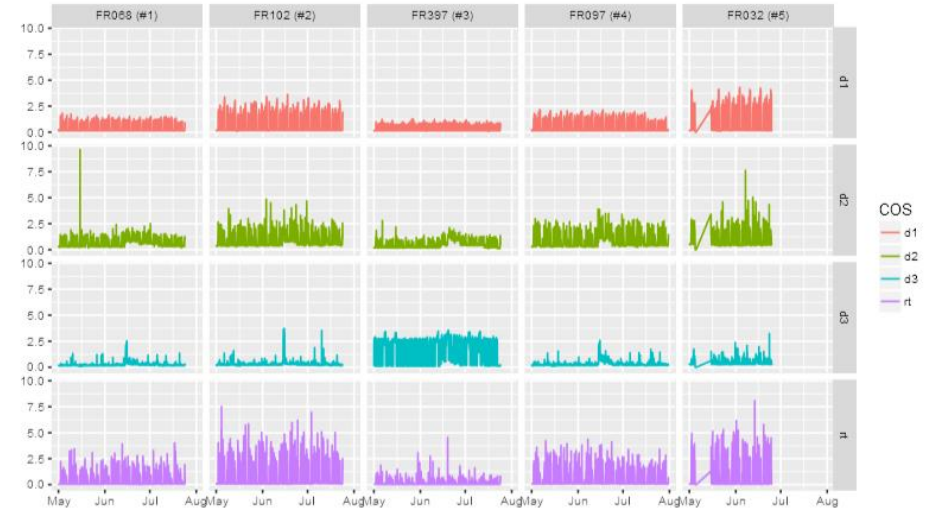
Powered by 

## Centroids

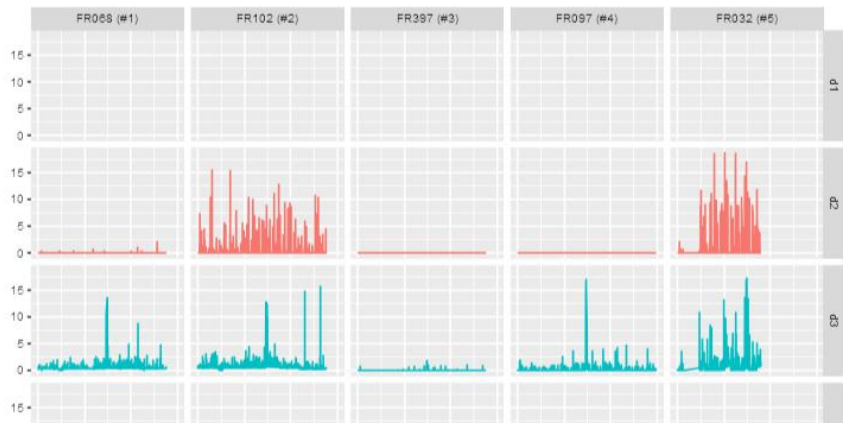
### traffic\_primary



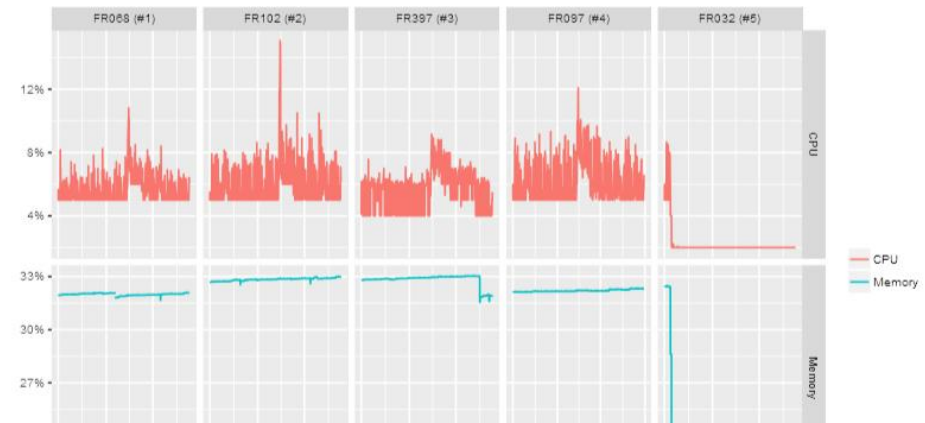
### cos\_consumption



### cos\_drop



### load\_primary



# Détection de Plateaux

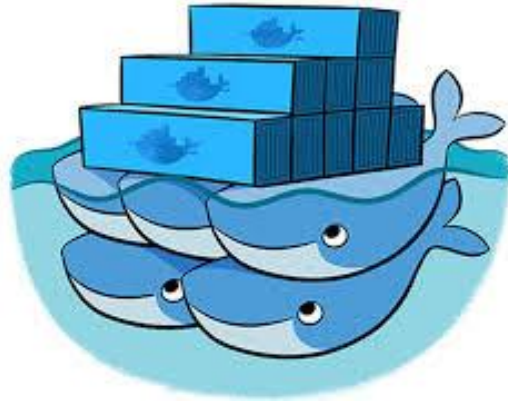
The dashboard displays the following information:

- Site Count:** 10 (Number of sites with detected plateaux)
- Plateau Count:** 216 (Number of detected plateaux)
- Introduction:** Identify plateau issues on the bandwidth measurement. A plateau is identified when the site is not 100% saturated, but the throughput seems to be throttled. The detection is based on minimizing the standard deviation before and after the current value. The lower the standard deviation is, the more defined the plateau is.
- Overview Table:**

Site	Interface	Type	Plateau Count	Deviation
FR-CC-07	Interface: WAN	out	47	
FR-CC-06	Interface: WAN	in	46	
FR-CC-07	Interface: WAN	in	41	
FR-CC-05	Interface: WAN	in	27	
FR-CC-02	Interface: WAN	out	25	
FR-CC-02	Interface: WAN	in	18	
FR-CS0011	Interface: WAN	in	3	
FR-CC-03	Interface: WAN	in	3	
FR-CC-01	Interface: WAN	in	2	
FR-CS0021	Interface: WAN	in	2	
- Time-series Graph:** Shows bandwidth (bytes\_in and bytes\_out) over time from April 18 to May 18. A plateau is visible around April 28-30.
- Calendar Heat Map:** Shows the spread of issues over time for 2018.

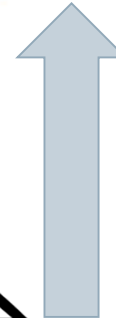
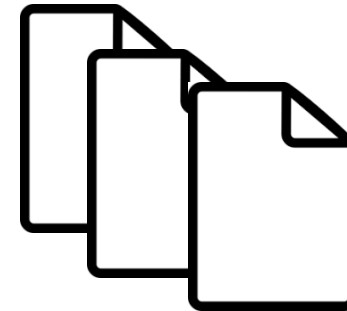
# Architecture

UI



Extrait

Calcul



Fichiers RDS

# Conclusion

Capacité de R à

1. définir des algorithmes
2. à restituer les résultats



*À la recherche de  
développeurs R/Shiny*

**Pierre Rebours**  
**[prebours@maltem.com](mailto:prebours@maltem.com)**  
**07 85 03 91 53**