

Centre de Calcul
de l'Institut National de Physique Nucléaire
et de Physique des Particules



Informatique Décisionnelle au CC-IN2P3

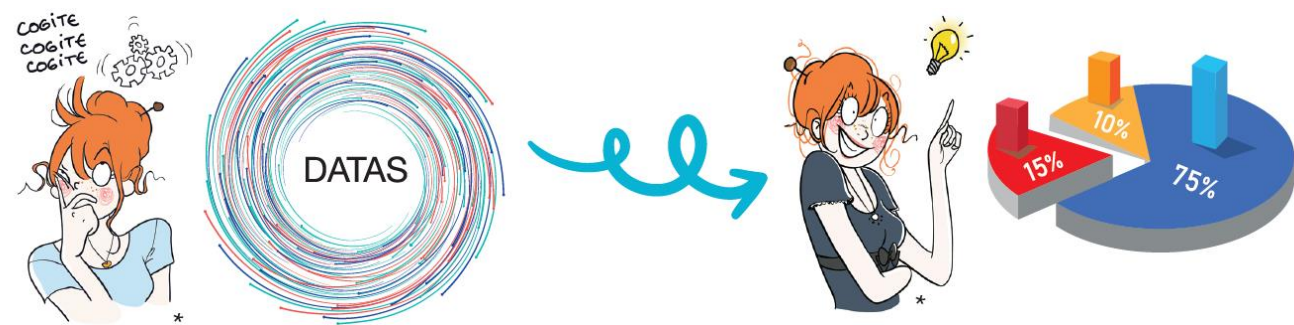
CATHALA Dominique & EVESQUE Cécile



■ **CONTEXTE**
CC-IN2P3 = services, stockage, calcul

■ **ENJEUX**
Faciliter + Améliorer le suivi des activités

■ **OBJECTIFS**
Décisionnel = centralise + métriques + visuels



■ Project décisionnel

- Naissance en 2010

■ Démarche d'évolution

- Initiée en 2016
- Interactif
- Libre
- Mutualisation des services

■ Démarche progressive

- Visualisation 2018
- Intégration 2023
- Stockage 2025

■ Equipe

- Dominique Cathala
- Cécile Evesque
- Sylvain Reynaud

■ Compétences

- Data analyst
- Intégrateur d'application
- Développeur

■ Visualisation

- SSRS



■ Visualisation

- Tableau

■ Stockage

- MSS



■ Stockage

- PostgreSQL

■ ETL

- SSIS



■ ETL

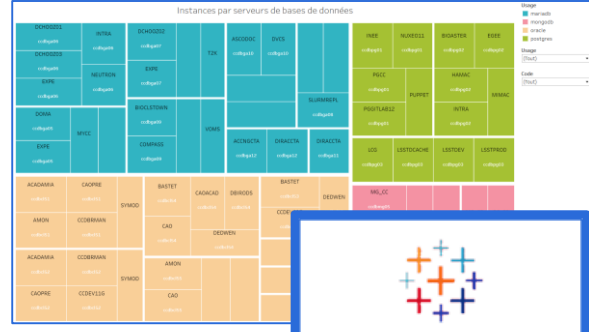
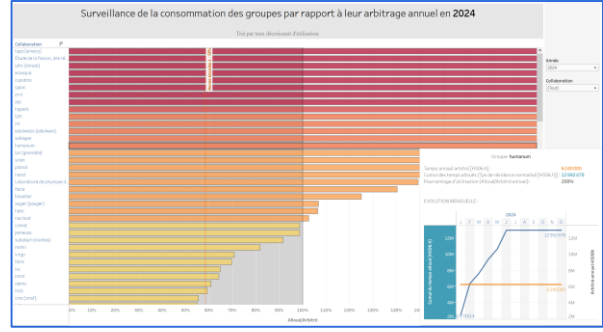
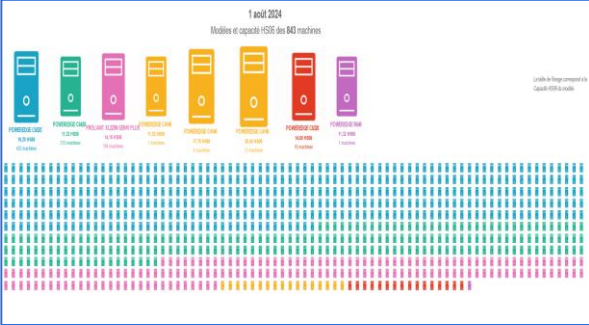
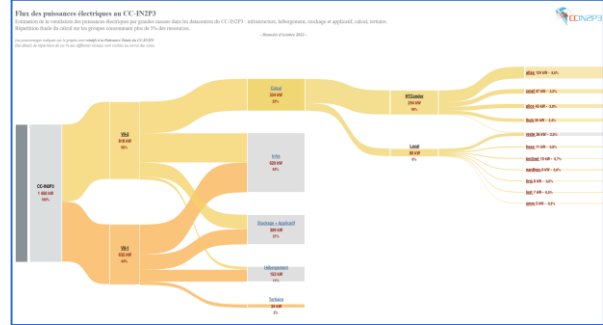
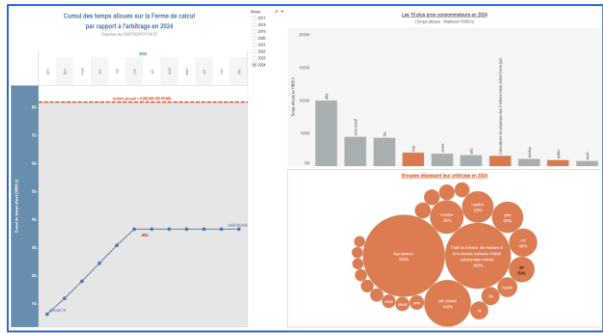
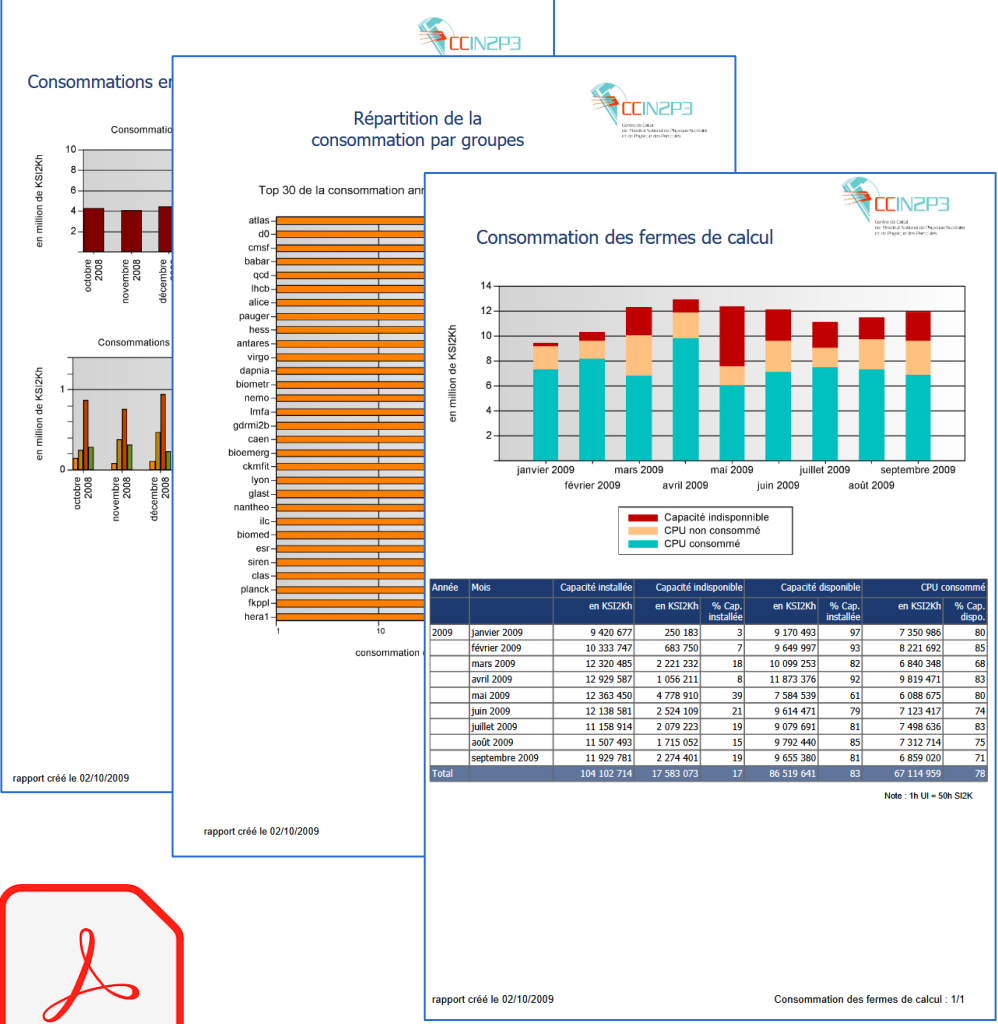
- Spark
- Airflow
- Scala, Python

■ Tout **PROPRITAIRE**



■ Majoritairement **LIBRE**





STATIQUES

DYNAMIQUES



■ Les « + »

- Mutualisation des services
 - ETL
 - Bases de données
- Redondance des compétences
- Recentrage de notre métier
 - La donnée



■ Les « - »

- Dépendance d'un service
- Multiplication des outils
- Continuité et pérennité

Pour en savoir plus ...

... RDV
À l'espace poster

LE DÉCISIONNEL AU CENTRE DE CALCUL DE L'IN2P3





CONTEXTE
Le CC-IN2P3 conçoit et opère un ensemble de services, en particulier un système de stockage de masse et de moyens de traitement de grands volumes de données. L'ensemble des services produit un flux continu de données très hétérogènes.

ENJEUX
Faciliter l'analyse des données pour améliorer le suivi des activités du CC-IN2P3.

OBJECTIFS
La plateforme Décisionnelle centralise, agrège et archive les métriques les plus pertinentes pour permettre des analyses fines à travers des visuels.

Comprendre et analyser nos données métier complexes et hétérogènes



SOURCES
Sources données
- Multi-sources de données
- Données hétérogènes
- Fichiers, Bdd, API, ...
FICHIERS
LAVOISIER
BASES DE DONNÉES

ETL
Extract Transform Load
- Extraire les données
- Transformer les données
- Charger les données
Technologies
SPARK Scala
AIRFLOW
open source

Data Warehouse
PostgreSQL
- Recueil de données
- Sources variées hétérogènes
- Stockage PostgreSQL
Data Warehouse
Operational Data Store
open source

Data Analyse
TableauSoftware
- Tableaux de bord interactifs
- Identification de KPI
- Analyse de corrélations
PROPRIÉTAIRE

Diffusion
Portail Web
- 2500 utilisateurs externes
- 100 utilisateurs internes
Multi support
- Affichage adaptatif
- Multi-résolution

EXEMPLES
Flux de puissances électriques
Suivi des licences CATIA par labo.
Suivi des ressources de CLOUD

Cycle de vie de la création d'une visualisation décisionnelle
COLLECTER → TRANSFORMER → STOCKER → ANALYSER → RESTITUER

ETL	TABLEAU SOFTWARE	DATA VISUALISATION	CHIFFRES
<ul style="list-style-type: none">- Spark/Scala Traitement en parallèle de gros volume- Airflow Planificateur de tâches- GitLab/ Intégration Continue Gestion de cycle de vie de développement logiciel	<ul style="list-style-type: none">- Tableau serveur Linux- Tableau Desktop Windows et Mac- Interopérabilité des sources multiples- Optimisation par des Extraits / de la mise en cache	<ul style="list-style-type: none">- Interactivité- Filtres multiples- Affichage dynamique- Exploration des données- Solution attrayante et ludique	<ul style="list-style-type: none">- 3,5 millions de données traitées par mois- 40 sources de données certifiées- 80 tableaux de bords interactifs- 2000 utilisateurs environ

Contact
🏠 <https://cc.in2p3.fr>
✉ decisionmaster@cc.in2p3.fr

Equipe Applications
CC-IN2P3 CNRS
21 Avenue Pierre de Coubertin
69627 Villeurbanne Cedex

* Crédits
Illustrations par Nathalie Jomard
<https://www.nathaliejomard.com/>

Merci