

Données ouvertes, smart-city, HPC et doubles numériques

François Bodin



Introduction

- On assiste à l'émergence de projets de doubles digitaux des métropoles
 - Modèles numériques et qualitatifs & données
 - Capteurs et réseaux
- Aide à la décision, anticipation des problèmes
 - Ecosystèmes/modèles systémiques (e.g. *smart-city*)
 - Recherche de modèles de villes durables
- Fondé sur la base de données libres mais pas seulement
 - Peu de projets peuvent être mis en place uniquement sur la base de données ouvertes
 - La centralisation des données est généralement / techniquement impossible
- Gérer l'hétérogénéité et la pluridisciplinarité est au cœur de ces projets

« Doubles numériques »

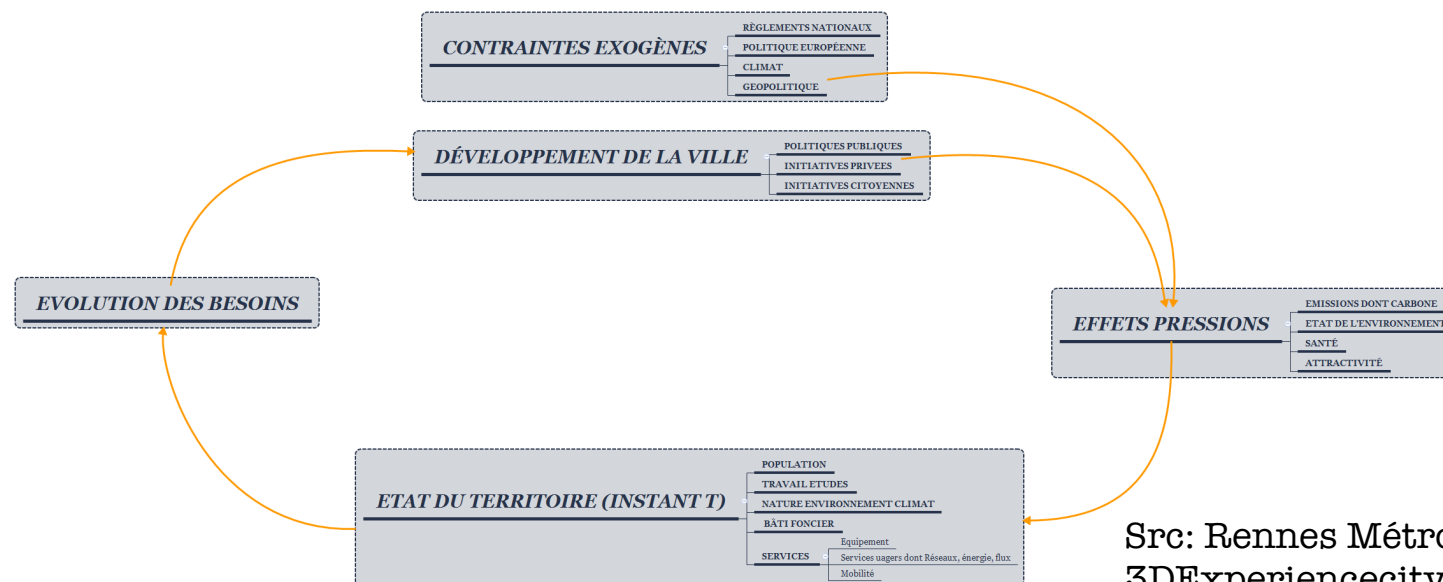
- Représentation numérique d'une partie du monde réel
 - Capture des relations spatiaux-temporels
 - Informe et prédit
- Plateformes de tests virtuels
 - Permettre des expérimentations
 - Alimentées par les données de l'IoT
- Modèles numériques et qualitatifs
 - CFD, modèles économiques, d'acteurs, ...

Plan de la présentation

- Modèle systémique pour la ville
- Mise en œuvre d'un double numérique
- Deux illustrations concrètes
 - Le projet AQMO
 - Le projet Flux Campus
- Quelques challenges et leurs approches

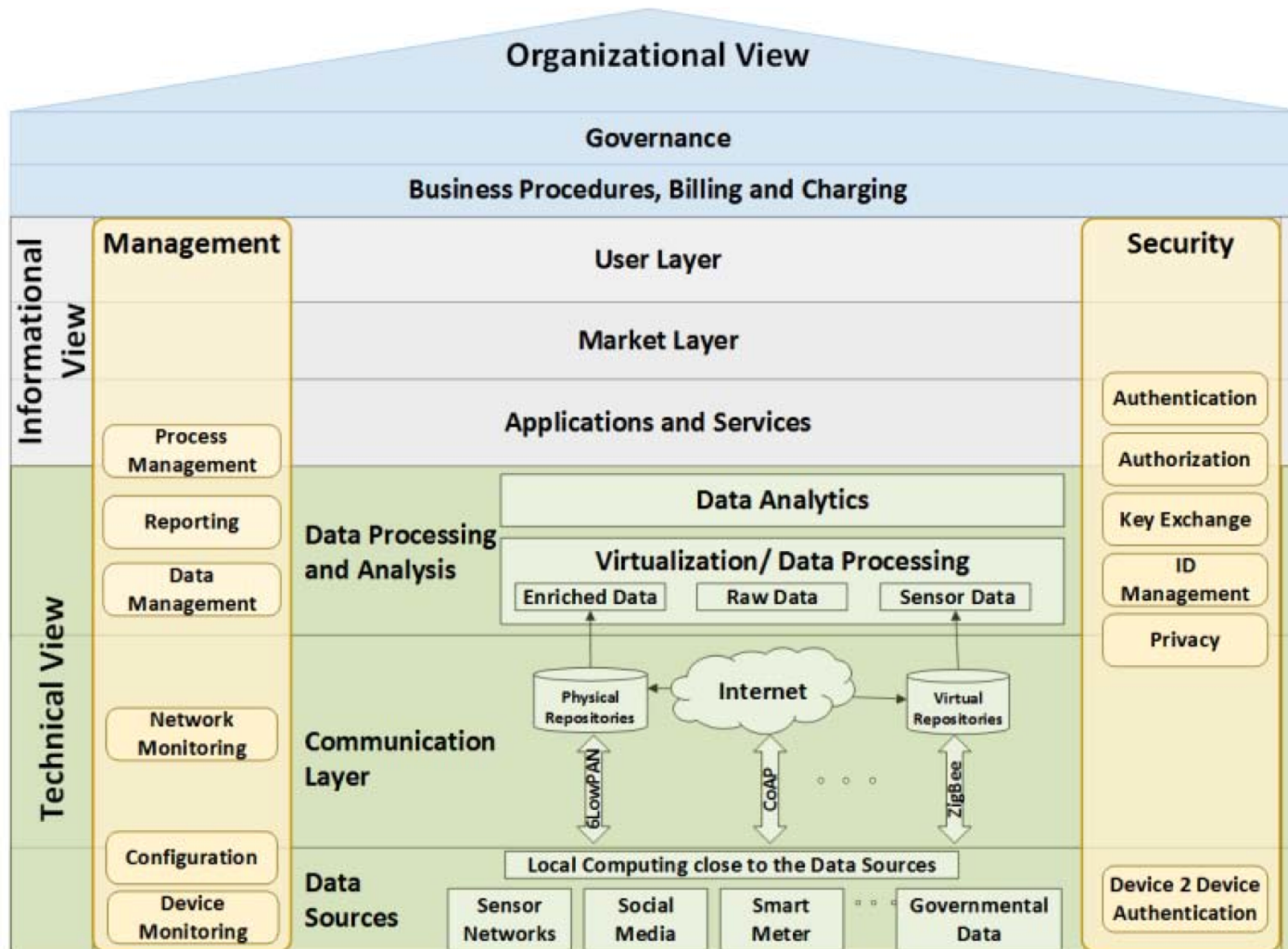
Modèle systémique pour la ville

- Ici la composition d'un ensemble de doubles numériques
 - Comprendre et analyser les interactions entre sous-systèmes
 - Planification, pilotage, concertation



Src: Rennes Métropole
3DExperiencecity Virtual Rennes

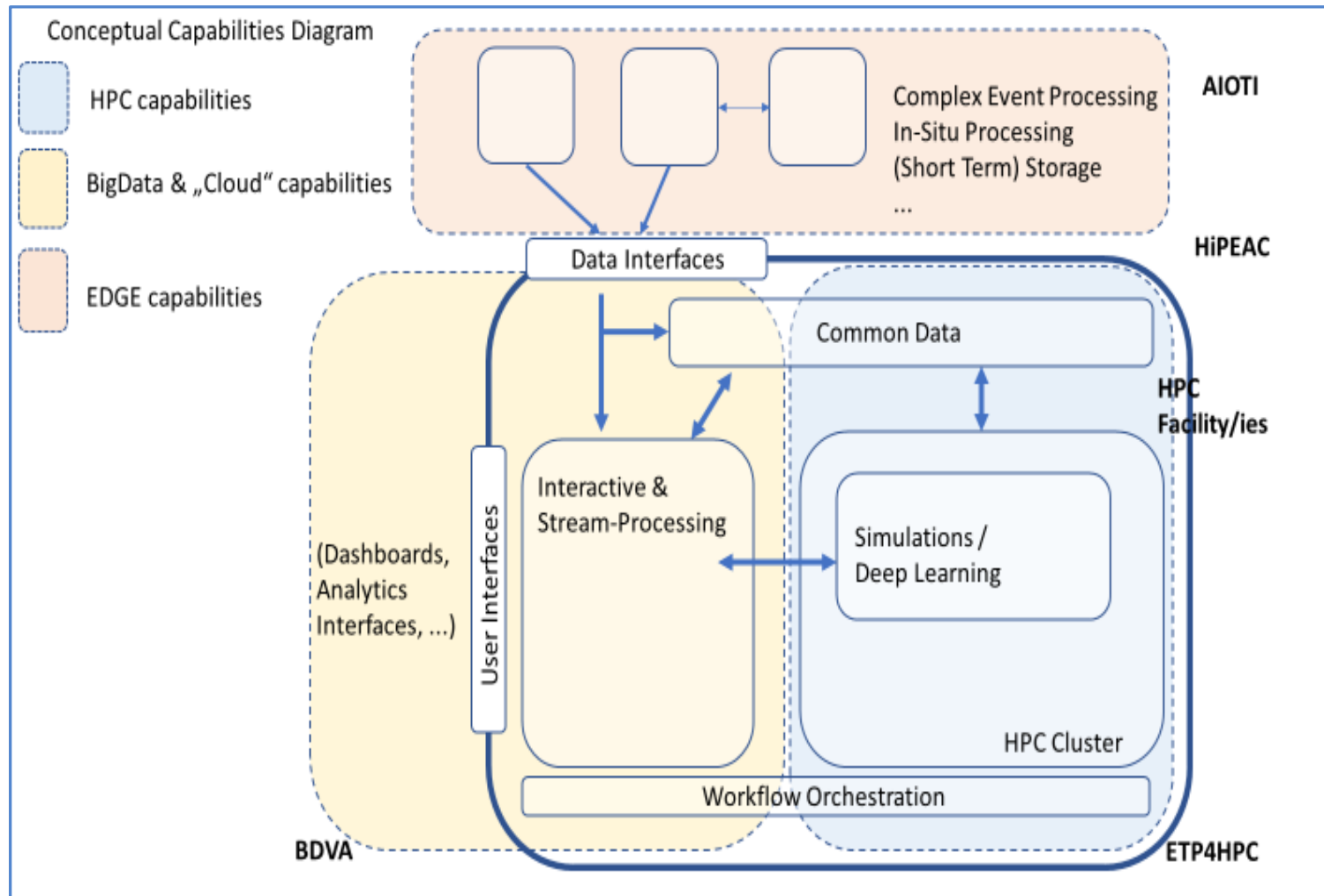
Vue globale *partielle* des infrastructures pour les smart-city



Ou est le HPC?

Towards an Open Data based ICT Reference Architecture for Smart Cities, Src: <https://ieeexplore.ieee.org/document/8046284/>

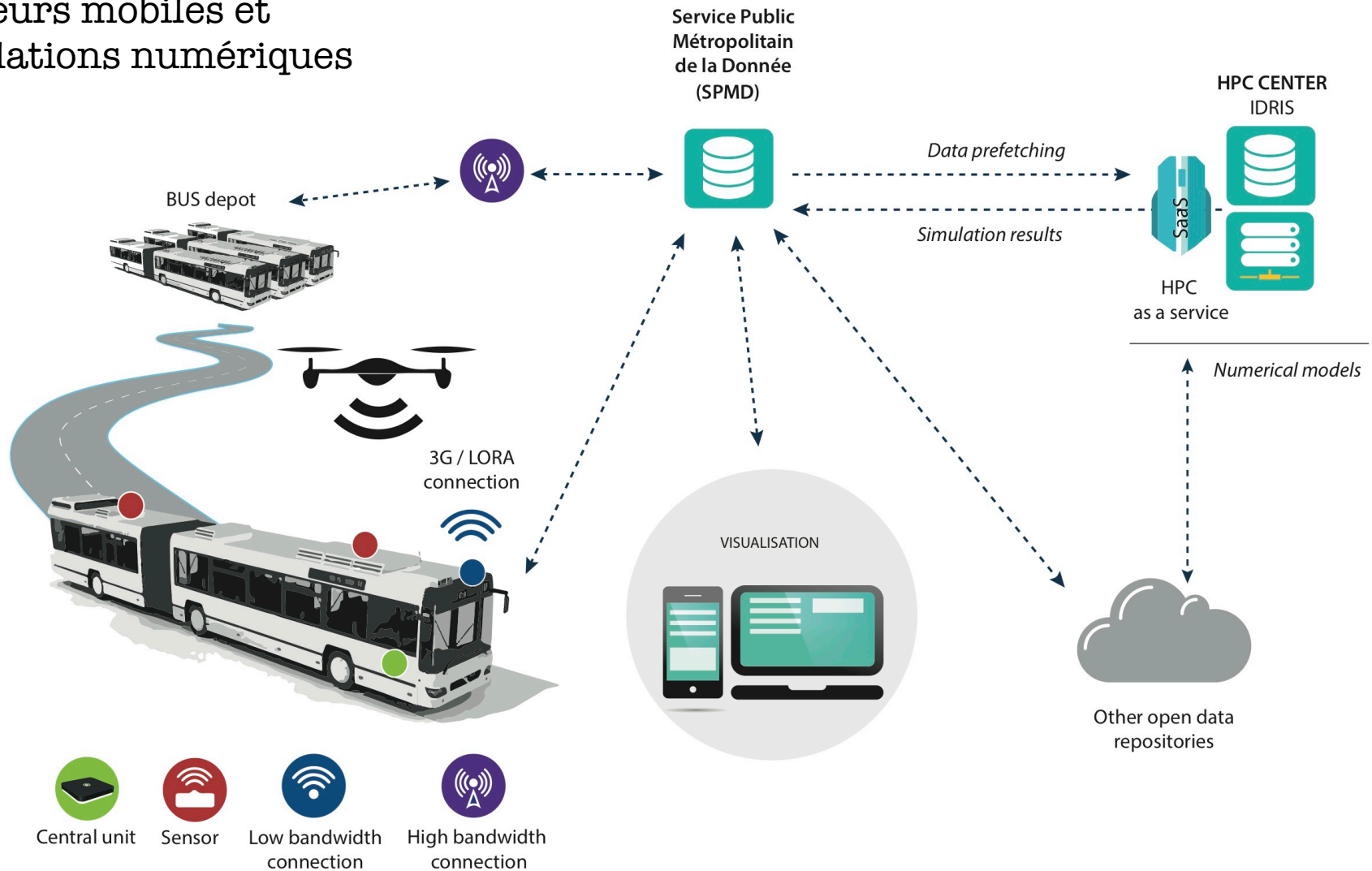
Mise en œuvre d'un double numérique



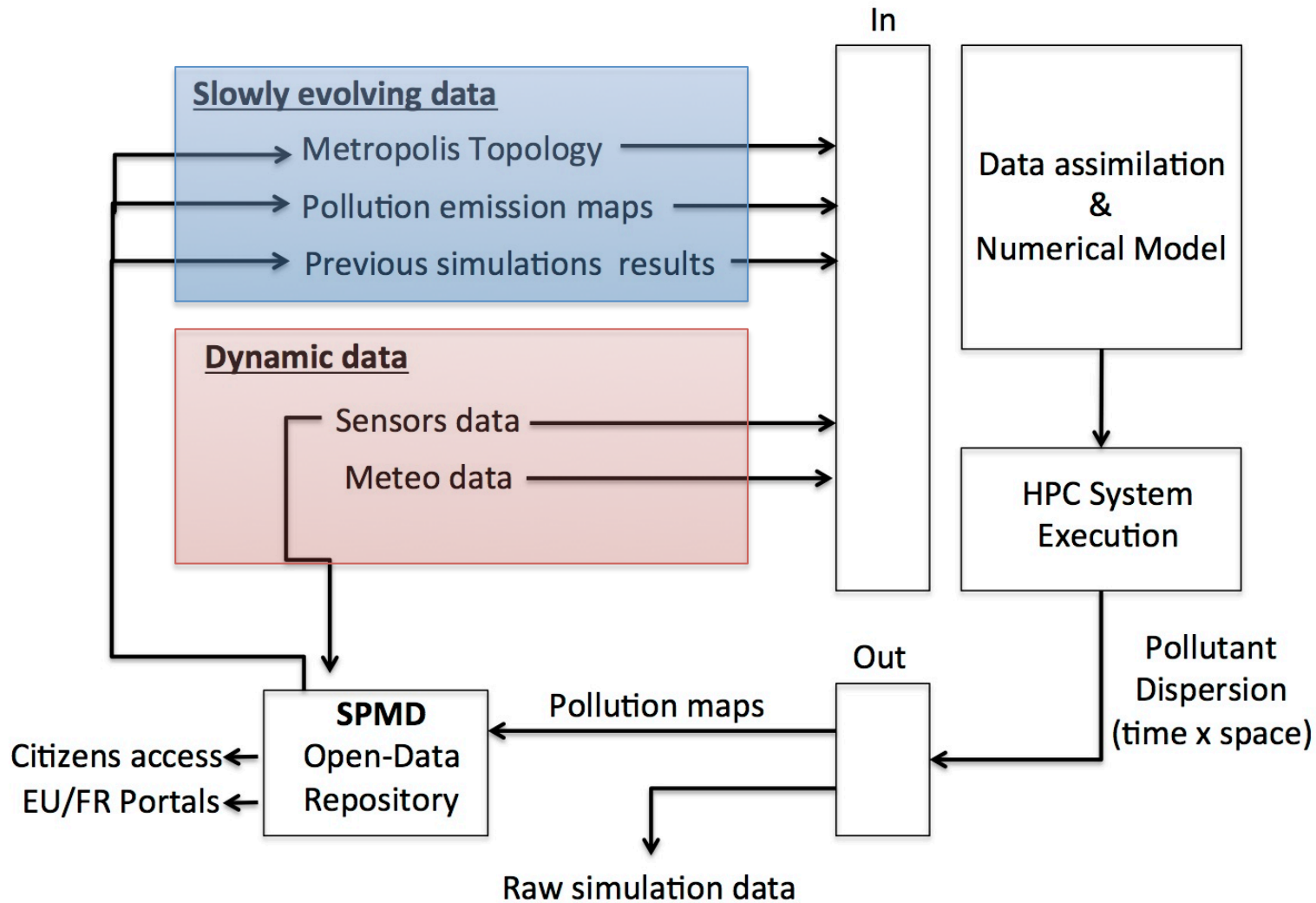
Src: Jens Krueger ITWM Faunhofer, BDVA/ETP4HPC/EXDCI2

AQMO: mesure de la qualité de l'air

Capteurs mobiles et simulations numériques

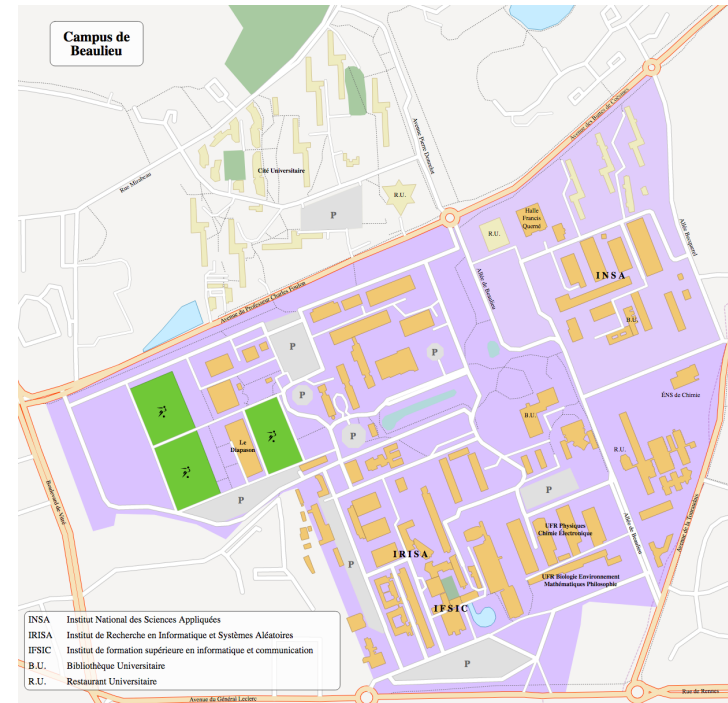


Flux de données dans AQMO

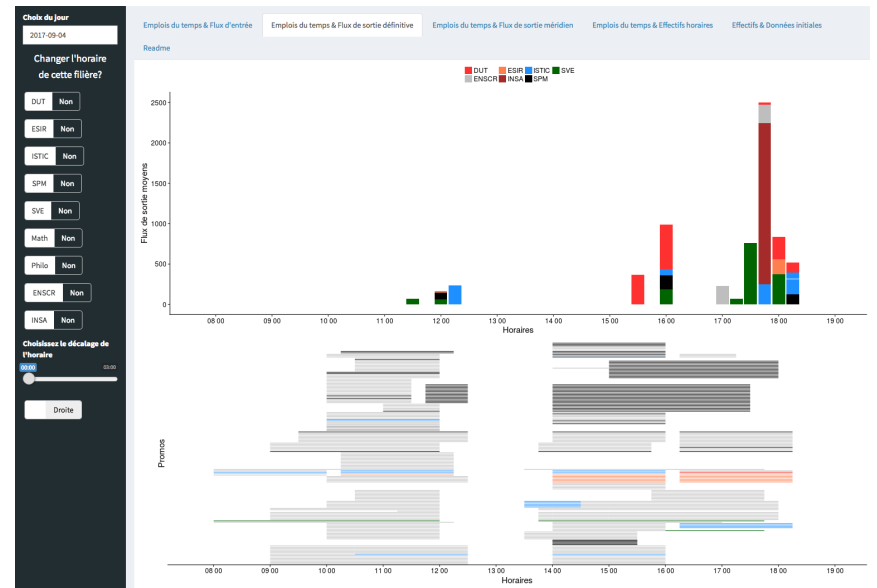
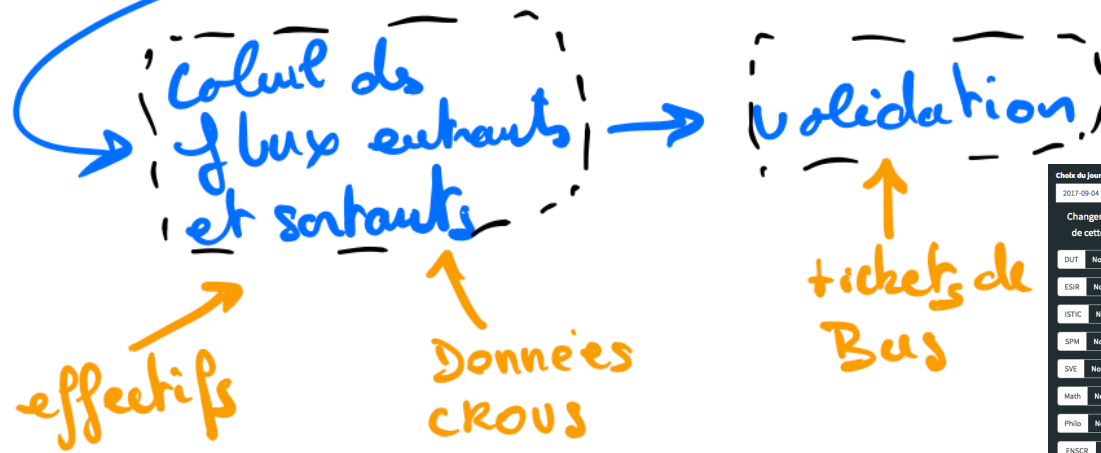
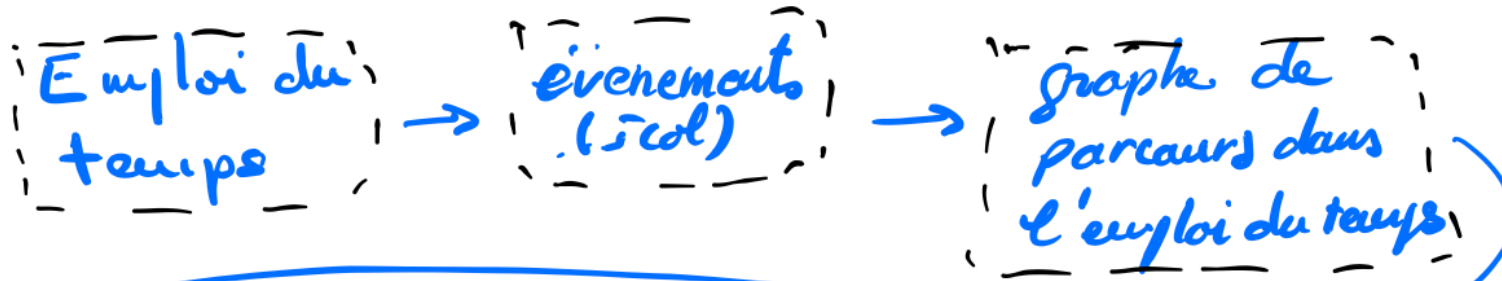


Flux Campus

- Un exemple à petite échelle d'un double numérique du campus universitaire de Beaulieu
 - Exemple à petite échelle d'un écosystème
- Le campus de Beaulieu
 - Homogénéité territoriale et fonctionnelle
- Projet Flux Campus
 - Analyser l'impact des emplois du temps étudiants sur les transports et la restauration
 - Gestion des pics d'affluence
- Etude menée par L. Lesoil avec le *bureau des temps* de la métropole rennaise



Flux Campus



Quelques challenges

- Challenge #1: La mesure
- Challenge #2: Accès aux données
- Challenge #3: La validation
- Challenge #4: Infrastructure matérielle et logicielle
- Challenge #5: Pluridisciplinarité
- Challenge #6: Partage des ressources
- Challenge #7: Transport des données
- Challenge #8: Archivage / Historique

Challenge #1: La mesure

- Capteurs ≠ mesures
 - De nombreuses technologies allant des instruments scientifiques pointus à des capteurs faibles coûts
- Peu de capteurs coûteux ou beaucoup de capteurs peu coûteux?
 - Instrument AASQA → ~ 25 k€
 - Capteur pm10, pm2.5 → ~ 25 € / 400 €
- La question de la mesure est centrale
 - Comment calibrer?
 - Comment valider?

Challenge #1: La mesure

- Flux campus
 - La fréquentation des bus sur le campus de Beaulieu (Keolis)
- AQMO
 - Comparaison des capteurs avec ceux de références (Air Breizh)

Challenge #2: Accès aux données

- La logistique des données dépend de nombreux paramètres
 - Localisation et volume
 - Droit d'accès, *privacy*
 - Transports (e.g. CDN)
 - Pérennité des accès, des données
 - Performance globale du système
 - Compromis calculs vs communications
- Données hétérogènes
 - Beaucoup de données libres
 - Mais pas seulement, multiples propriétaires
 - Souvent éparses et de qualité variée

Challenge #2: Accès aux données

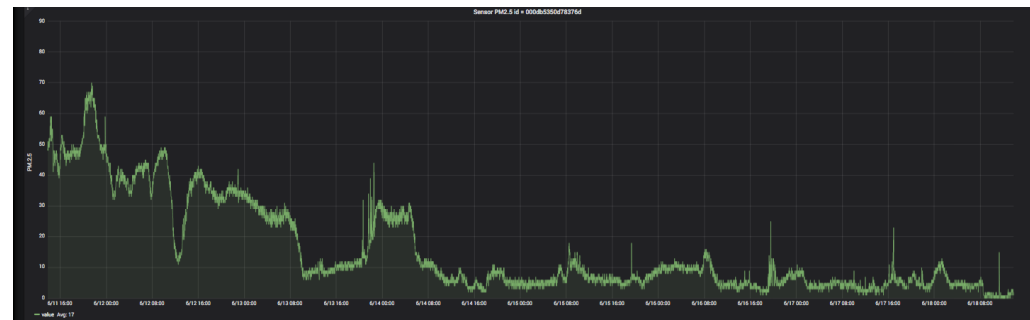
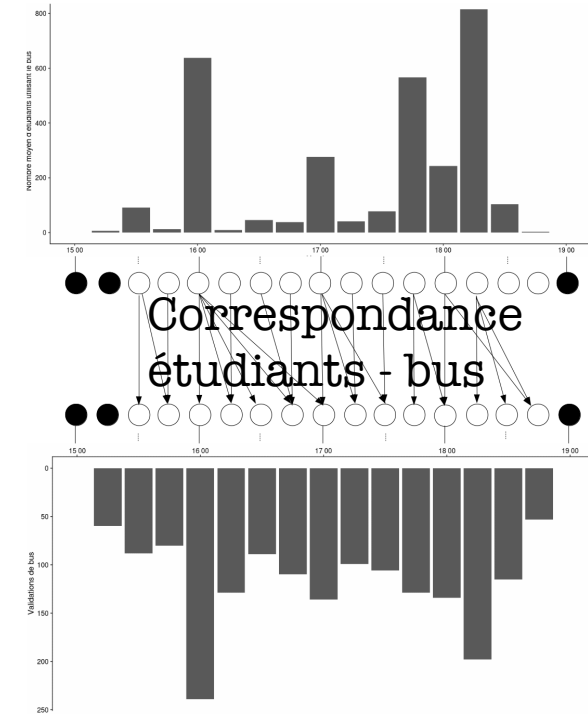
- AQMO
 - API ouverte, composants techniques mixtes
 - Service Métropolitain de la Donnée de Rennes Métropole
 - Modèle métropolitain 3D
- Flux Campus
 - Echanges de fichiers
 - CROUS, UR1, Keolis
 - Données partiellement publiques

Challenge #3: La validation

- La stratégie de validation doit être conçue en amont de la construction du double numérique
- La question de la reproductibilité des résultats doit aussi être évaluée

Challenge #3: La validation

- Flux campus
 - Validation des tickets de bus
- AQMO
 - Comparaison des capteurs avec ceux de référence d'Air Breizh
 - Travaux CFD

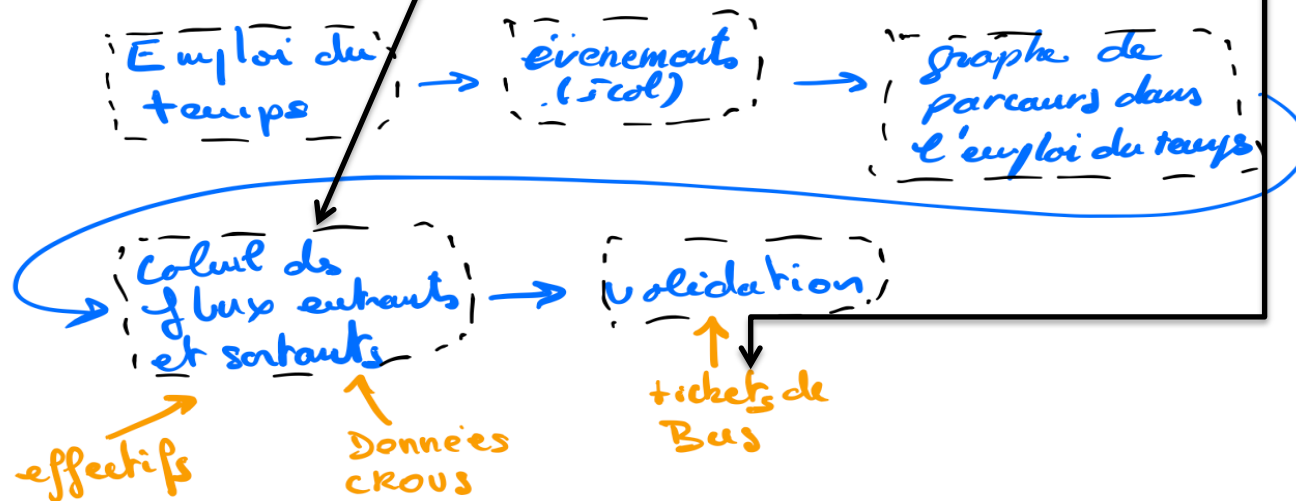
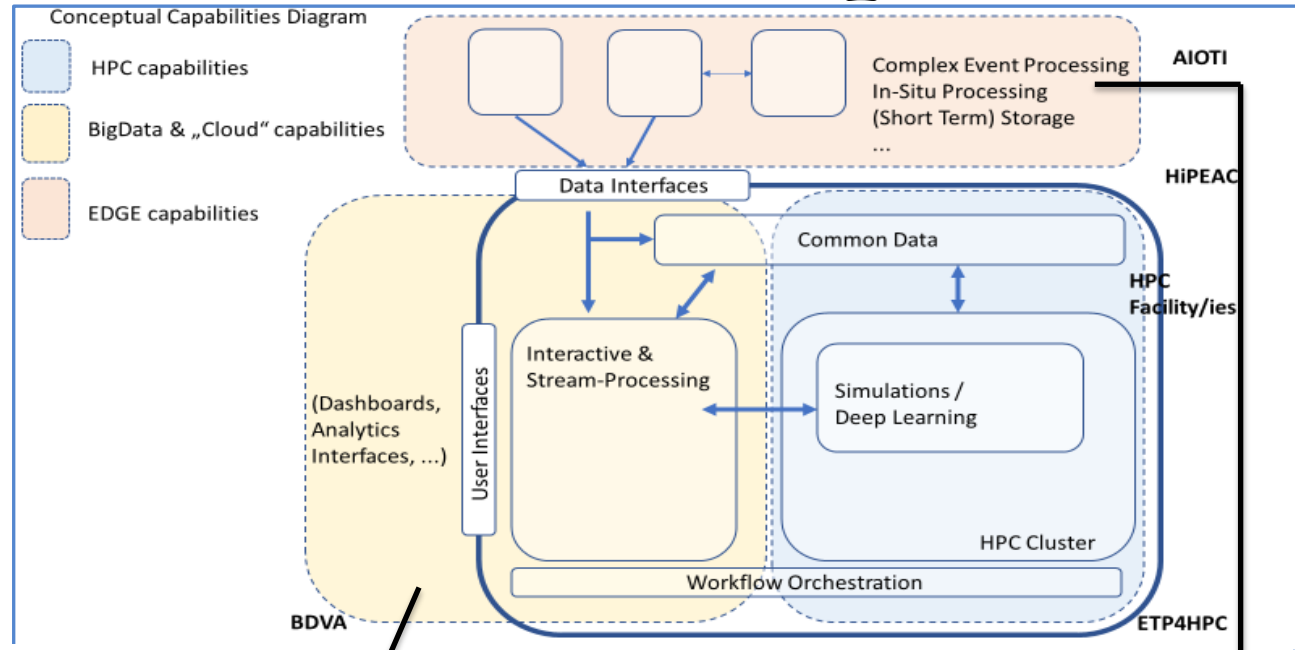


Capteur LS-113 (PM2.5), corrélation 0.77

Challenge #4: Infrastructure matérielle et logicielle

- Intégration de workflows complexes sur des architectures hétérogènes
 - Edge-computing
 - Analyse de données
 - Calcul numérique HPC
 - Containers (e.g. Docker)
 - Outils d'orchestration
- Interopérabilité
 - Modèle des utilisateurs
 - Modèle de ressources
 - Gestion transversale de l'adressage et des métadonnées

Flux Campus



AQMO

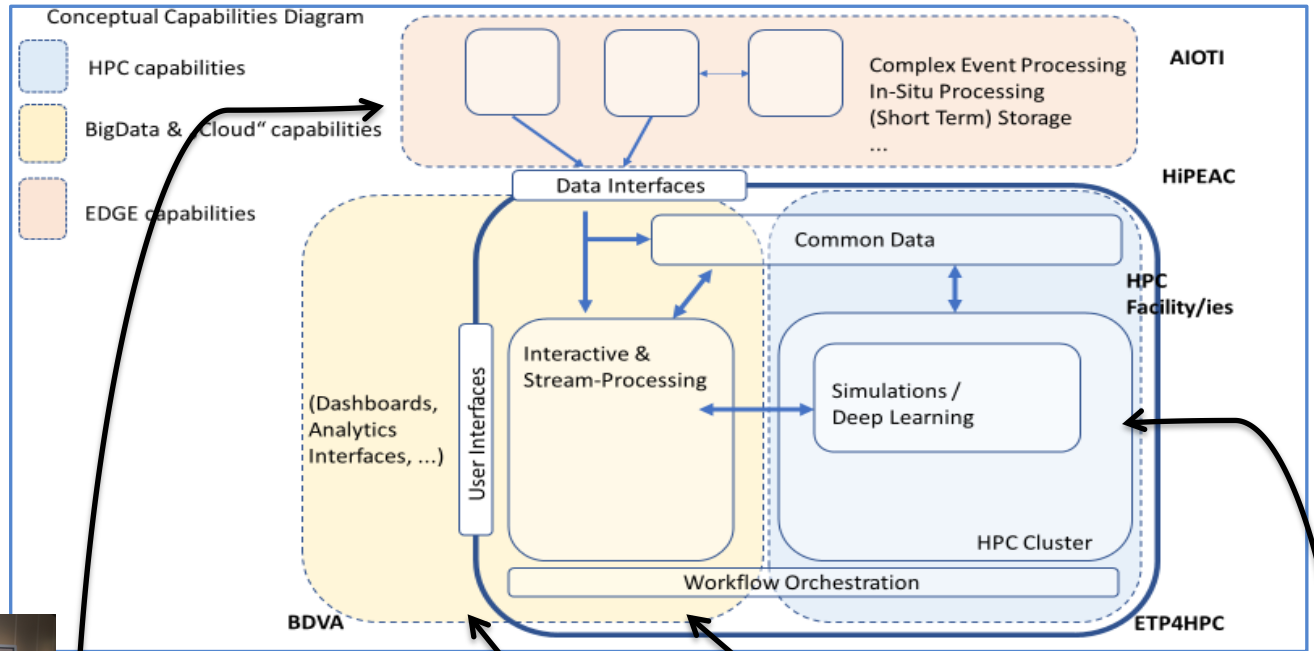
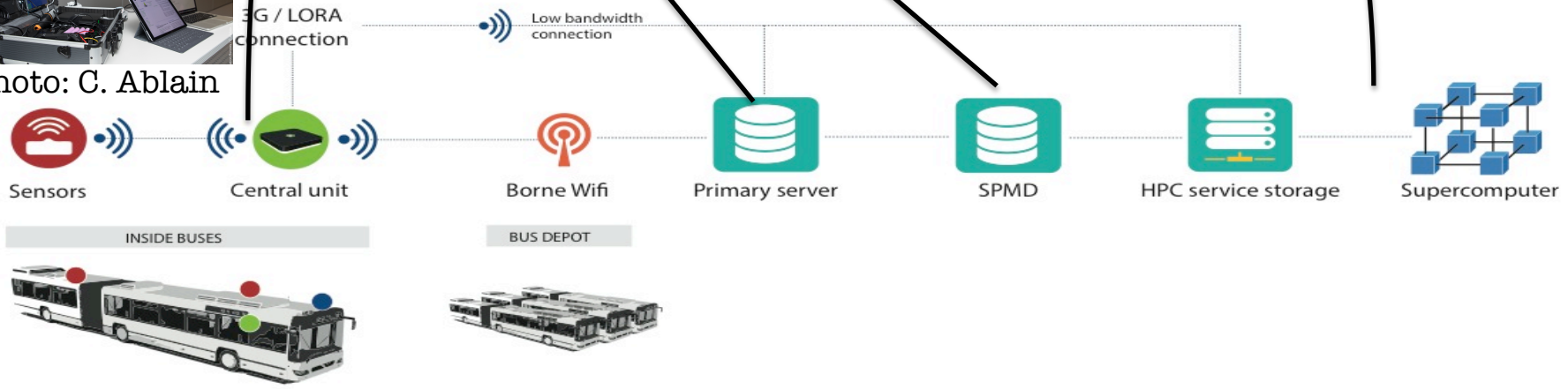


Photo: C. Ablain



Challenge #5:

Pluridisciplinarité

- Un double numérique est conçu par une équipe pluridisciplinaire
 - Informatique
 - Electronique
 - Data science
 - HPC
 - SHS
 - Métropole
 - etc.

Challenge #5: Pluridisciplinarité

- Flux campus
 - Informatique « web »
 - Analyse de données
 - Bureau des temps de Rennes Métropole
- AQMO
 - Spécialistes calcul numérique, HPC
 - Spécialistes systèmes d'informations
 - Spécialistes run-time, réseaux, ...
 - Spécialistes mesure de la qualité de l'air
 - Exploitant réseaux de bus

Challenge #6: Partage des ressources

- L'ensemble des ressources nécessaires ne sont pas coordonnées / gérées par un acteur unique
 - e.g. Cloud hybride, capteurs, moyens de calculs, réseaux de communications
 - *Multi-tenant*
 - Plusieurs organisations partagent les ressources de manière transversale
- L'ensemble des parties doit convenir des modalités de gestion des ressources et des données
 - Identification des utilisateurs
 - Gestion des droits d'accès
 - Interfaces
 - Programmation des composants

Challenge #6: Partage des ressources

- Flux campus
 - Pas un problème ici
- AQMO
 - API d'accès aux services et données ouvertes
 - Composants ouverts ou fermés
 - SPMD centralise les données
 - API pour l'usage du HPC as a service

Challenge #7:

Transport des données

- Assurer un transport des données efficace et adapté aux contraintes
 - Hétérogénéité des réseaux
 - Internet / 4G / Réseaux IoT / ...
- Contrôle des *tuyaux*
 - Pour permettre l'optimisation de la logistique des données
 - *Content Delivery Networks* (CDN)

Challenge #7: Transport des données

- Flux campus
 - Pas un problème
- AQMO
 - Lora, 4G, Wifi, Internet
 - Coordination métropolitaine
 - Pré-chargement des données pour la simulation et la visualisation

Challenge #8:

Archivage / Historique

- Les données d'hier sont les bases de la validation des modèles de maintenant
 - Doit inclure des versions « *containered* » des logiciels
 - Doit favoriser la réutilisation (documentation et format rigoureux)
 - Enjeu fort autour des métadonnées
- Gouvernance des données
 - Résoudre « Quand et quoi effacer »
 - Coût important et croissant avec le volume / l'augmentation du cycle de vie

Challenge #8: Archivage / Historique

- Flux campus, AQMO
 - Actuellement non traité

Conclusion

- Les plateformes expérimentales sont sur le chemin critique de la recherche dans ce domaine
 - La collaboration avec une métropole est incontournable
 - La place du HPC est appelée à grandir
- La validation des modèles restent un problème central
- Peu (pas assez) de données (ouvertes) sont disponibles
- Domaine en forte évolution
 - **Fog Computing Conceptual Model**
 - *Fog computing is a layered model for enabling ubiquitous access to a shared continuum of scalable computing resources (src: NIST 2018)*
 - Données à l'échelle du « **Yotta** » en vue