



Mappy-ing Open Source

25/09/2014, Postgresql sessions



Audrey Malherbe

**Responsable developpement
plateforme cartographique**

audrey.malherbe@mappy.com

 [@AudreyMalherbe](https://twitter.com/AudreyMalherbe)

m  ppy

Mappy, une marque de SoLocal Group

Spécialiste du déplacement et des services locaux sur Internet, tablettes et mobiles



+ de 10 millions
d'utilisateurs mensuels sur
Internet, tablettes et mobiles

27 millions
de visites / mois
(web et mobiles)

367 millions
d'itinéraires calculés en 2012

4 millions
de professionnels référencés

**Retour sur le projet de refonte
open source du service de
production des images de la
plateforme LBS Mappy**

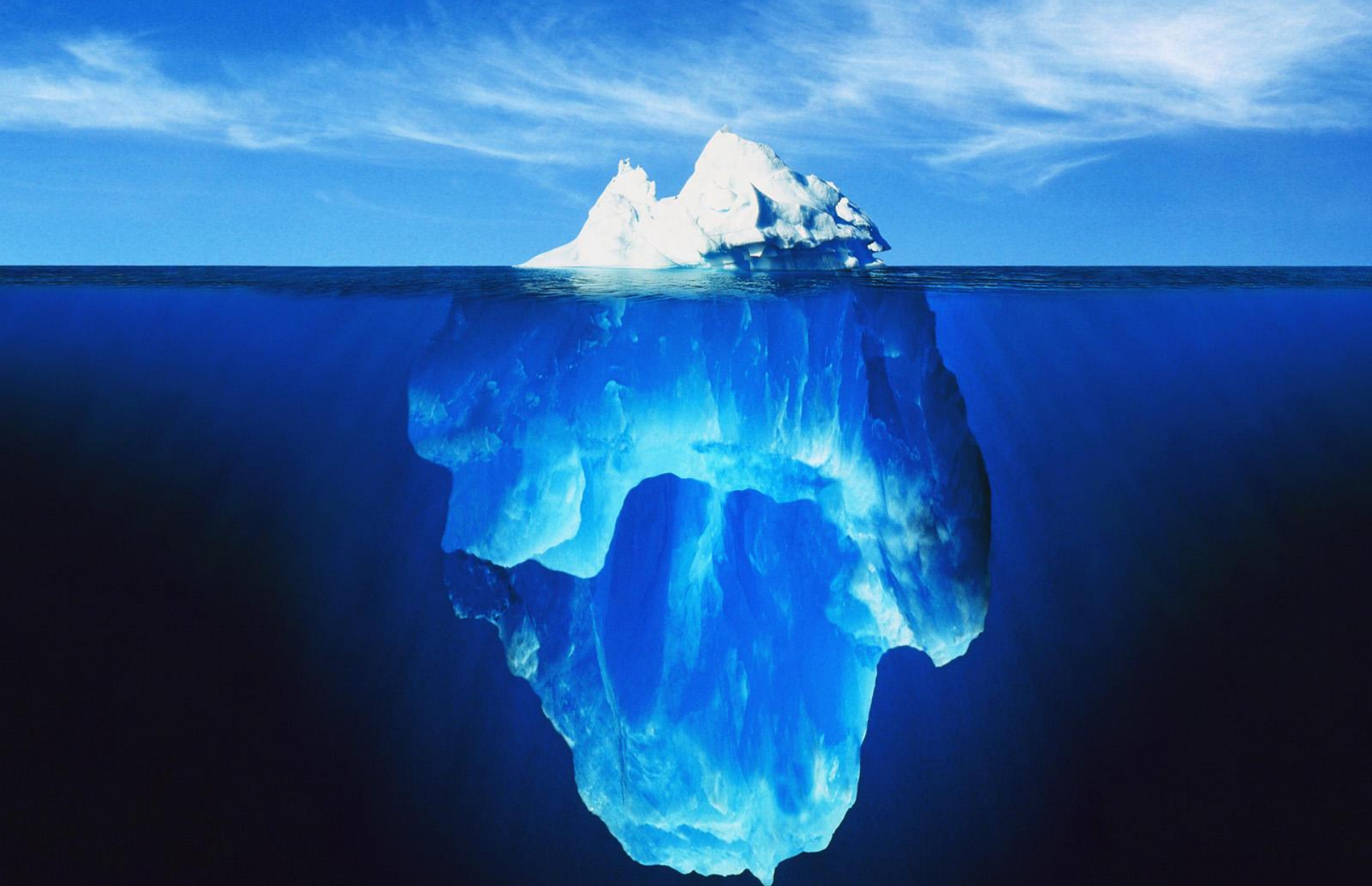
Plan

Les enjeux

Notre route

Les retours

Les prochaines étapes



Les enjeux

Les enjeux – point de depart

L'historique

- Production de plans depuis 16 ans avec des données TomTom/AND
- Carte Distribuée sur l'ensemble des supports Mappy/PagesJaunes
- Chaine de production « made in Mappy » maitrisée mais vieillissante

Le futur

- Repositionnement de la carte pour les nouveaux besoins : la carte est un média et doit servir de support à des contenus.
- Volonté de se tourner vers le monde de l'open source
- Préparation des défis techniques et fonctionnels à venir (cartes vectorielles, HD, reliefs, etc.)

BACK OFFICE

PRODUCTION

ORACLE[®]
SPATIAL



Stack Mappy
Moteur
Templates



Stack Mappy
Cache



API AJAX
Mappy

SDK Mobile
Mappy



Données
géographiques

Non standard

Format/Sortie propriétaire
12 niveaux de zoom

Processus de préparation

Privilégie le pré-calcul
10 jours de traitement

Suite logicielle

Comportement maîtrisé
Code peu vivant
Evolutions très coûteuses

Performances non négociables

6Md req/mois

Les enjeux - les objectifs

- Proposer un design de carte permettant de mettre en valeur du contenu
- Se synchroniser avec la sortie de la nouvelle version du site Mappy
- Réduire les temps de pré-calcul
- Combler les faiblesses de la plateforme et garder ses forces
- Rationaliser
- Se standardiser (API, 20 niveaux de zoom, x2 entre chaque niveau)
- Préparer l'avenir (montée en charge et nouveaux services)

**Take a Deep Breath.
Start Again.**



Temps de préparation réduit
2j (vectoriel) + 2j (relief)

Chaîne en production

Contribution Mapnik

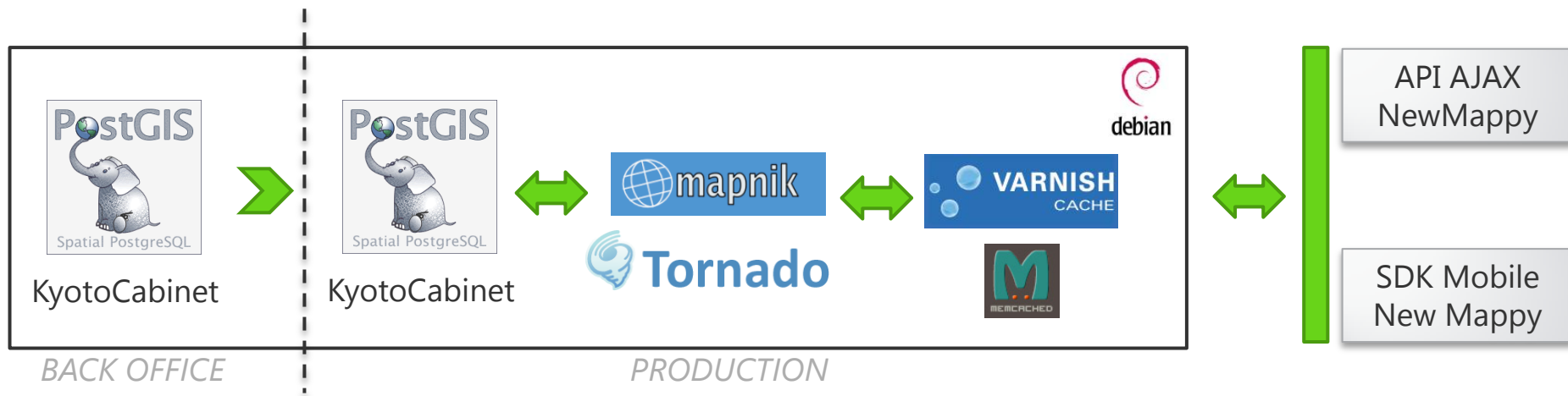
Standard

Nombre de lignes de code
optimisé

Code vivant

Baisse coût de la plateforme

Données
géographiques



Notre route

An aerial photograph of a wide, meandering river flowing through a valley. The river is light-colored, likely due to sediment, and winds in large, sweeping loops across the landscape. The surrounding terrain is a mix of dark green forests and brownish, rocky slopes. In the background, a range of rugged mountains stretches across the horizon under a sky filled with heavy, grey clouds. The lighting suggests late afternoon or early morning, with long shadows and a soft glow on the landscape.

Notre route - Demarrage

- Niveau hétérogène de l'équipe sur les technologies open source

- Se former, se faire aider
- Apporter la culture de l'open source

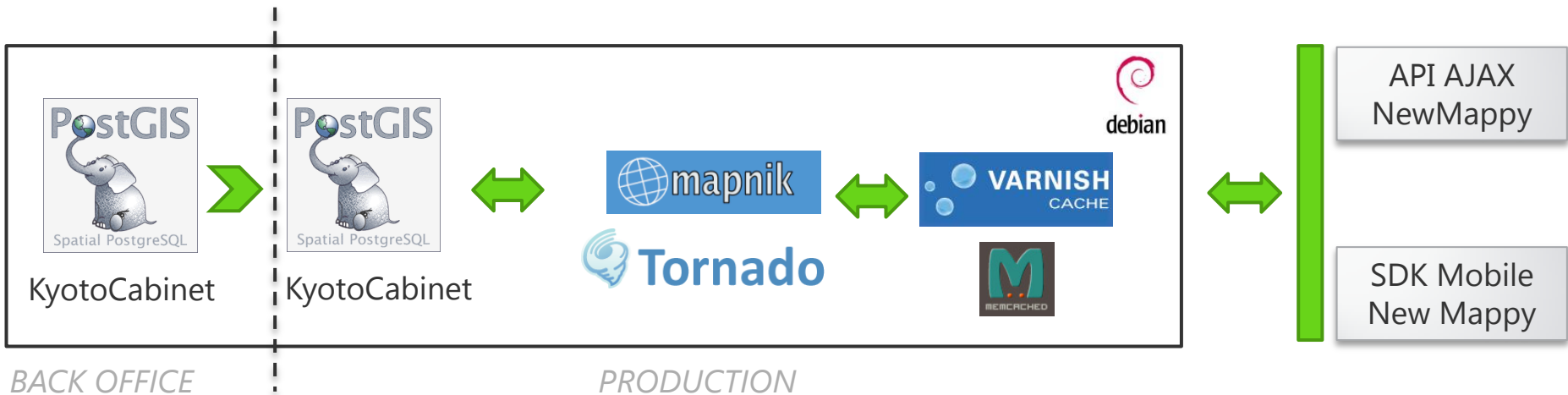
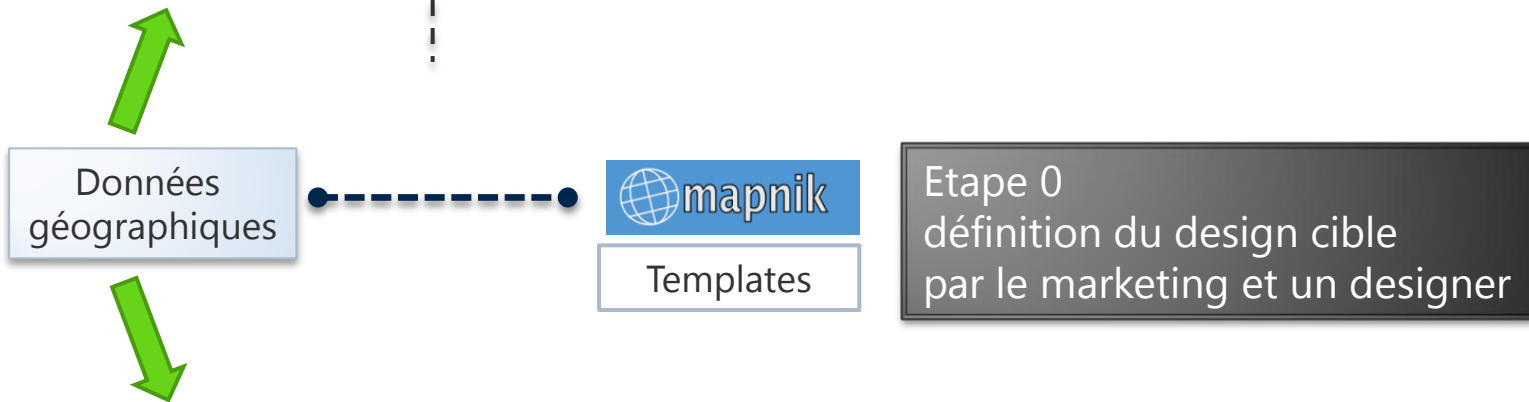
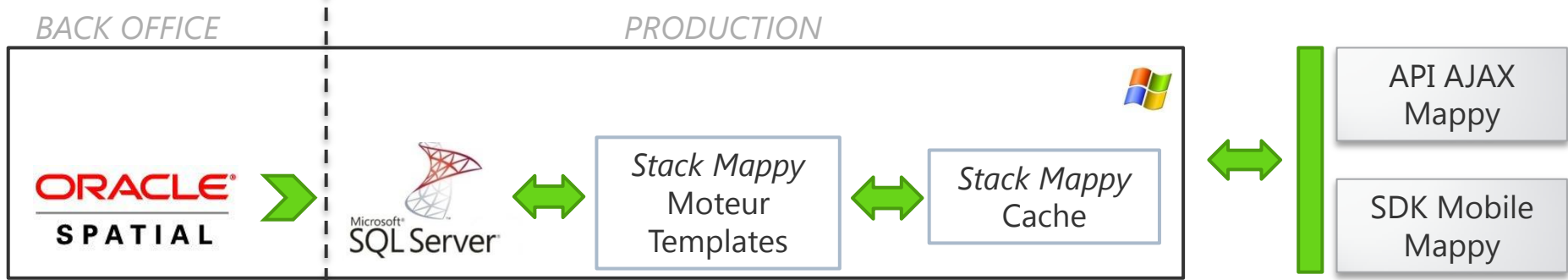


Innophi



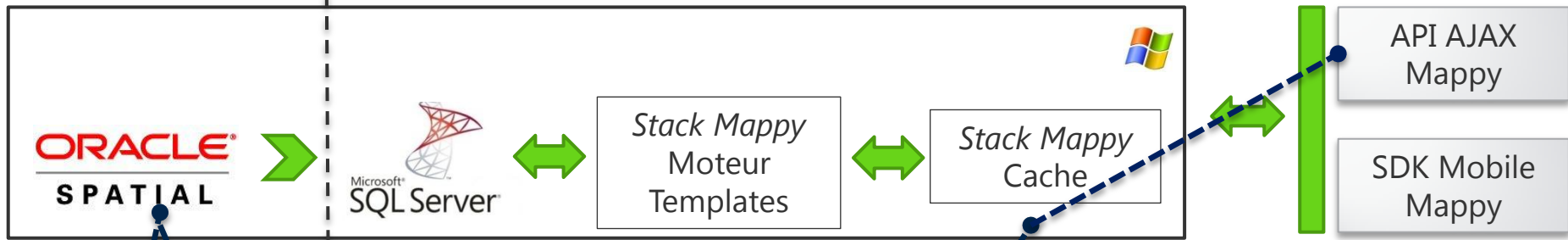
- Découpage du projet

- Migrer brique par brique pour une prise en main progressive
- Manger notre propre soupe
- Tester le plus vite possible



BACK OFFICE

PRODUCTION

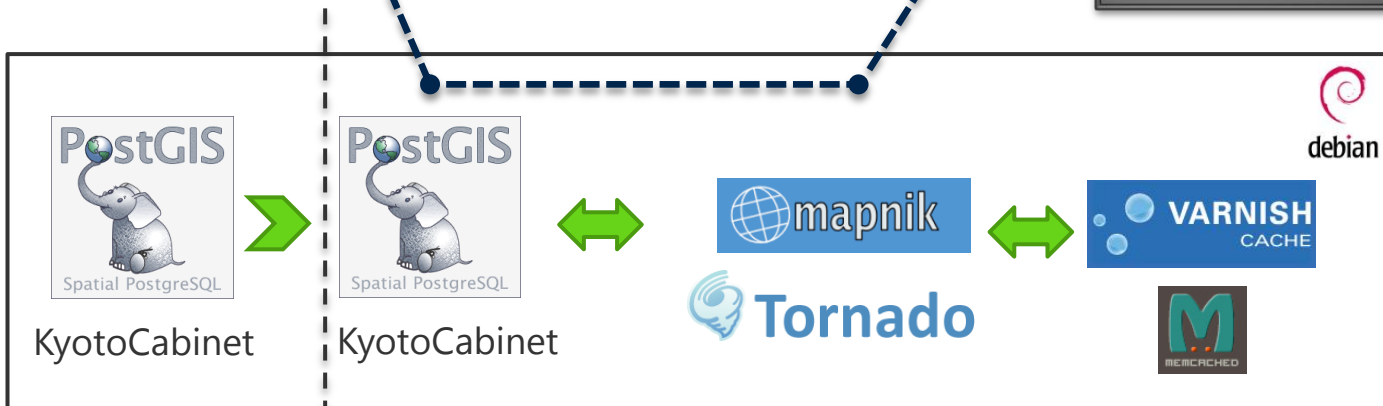


API AJAX Mappy

SDK Mobile Mappy

Données géographiques

Etape 1
Migration de la BDD et du moteur en restant dans le contexte historique
Travail sur les performances BDD



API AJAX NewMappy

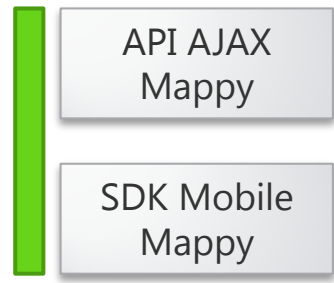
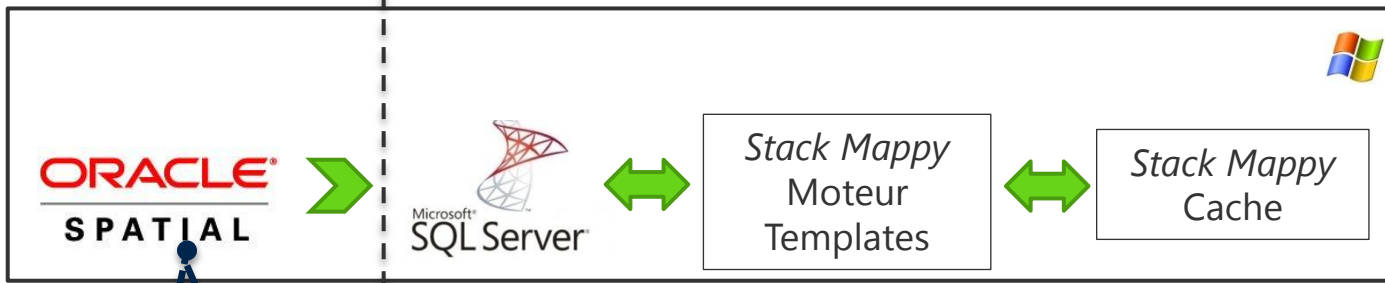
SDK Mobile New Mappy

BACK OFFICE

PRODUCTION

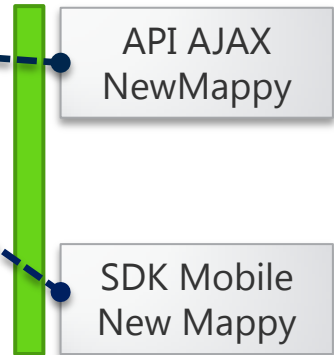
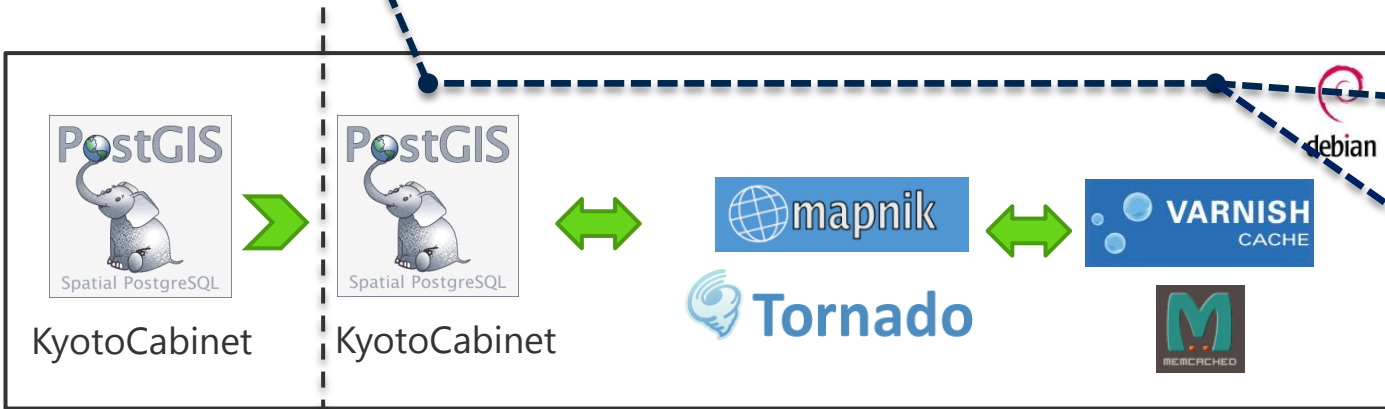
BACK OFFICE

PRODUCTION



Etape 2
 Migration du cache
 Nouvelles APIs
 Ajout du relief/Carte HD
 Travail sur les performances
 et les configurations Cache/Service/APIs

Données géographiques

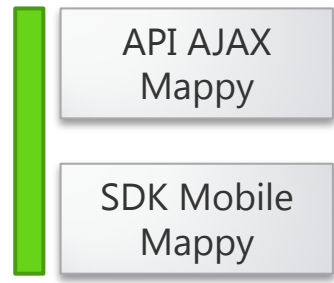
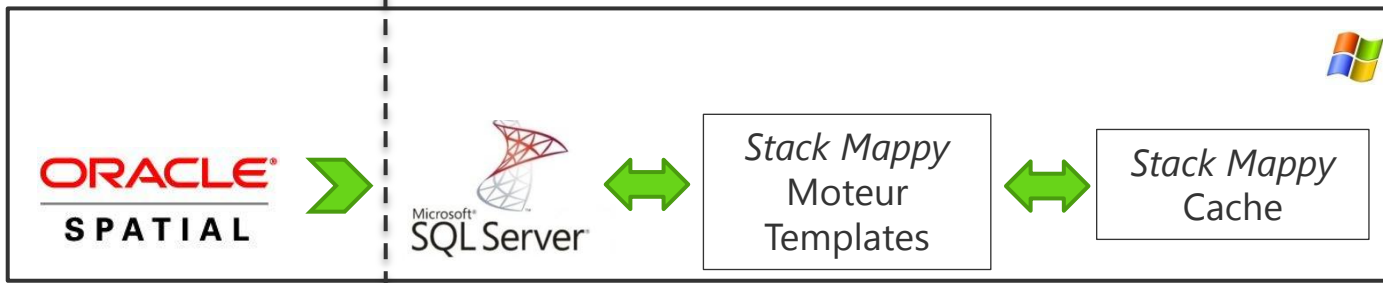


BACK OFFICE

PRODUCTION

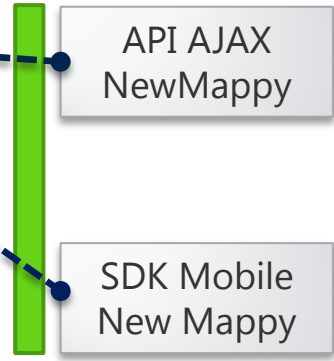
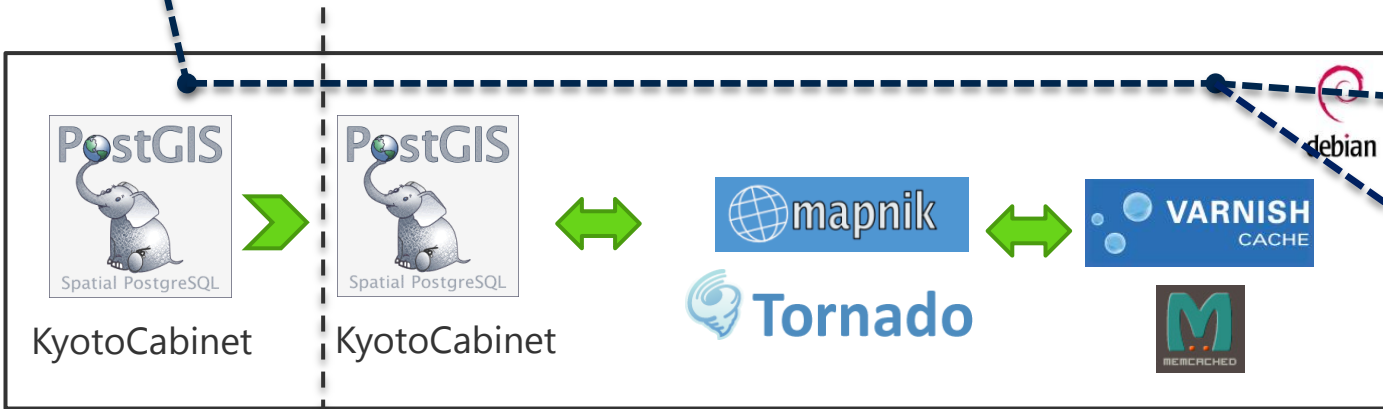
BACK OFFICE

PRODUCTION



Etape 3
Rationalisation des BDDs
Montée en charge progressive
Optimisation de la chaine

Données géographiques



BACK OFFICE

PRODUCTION

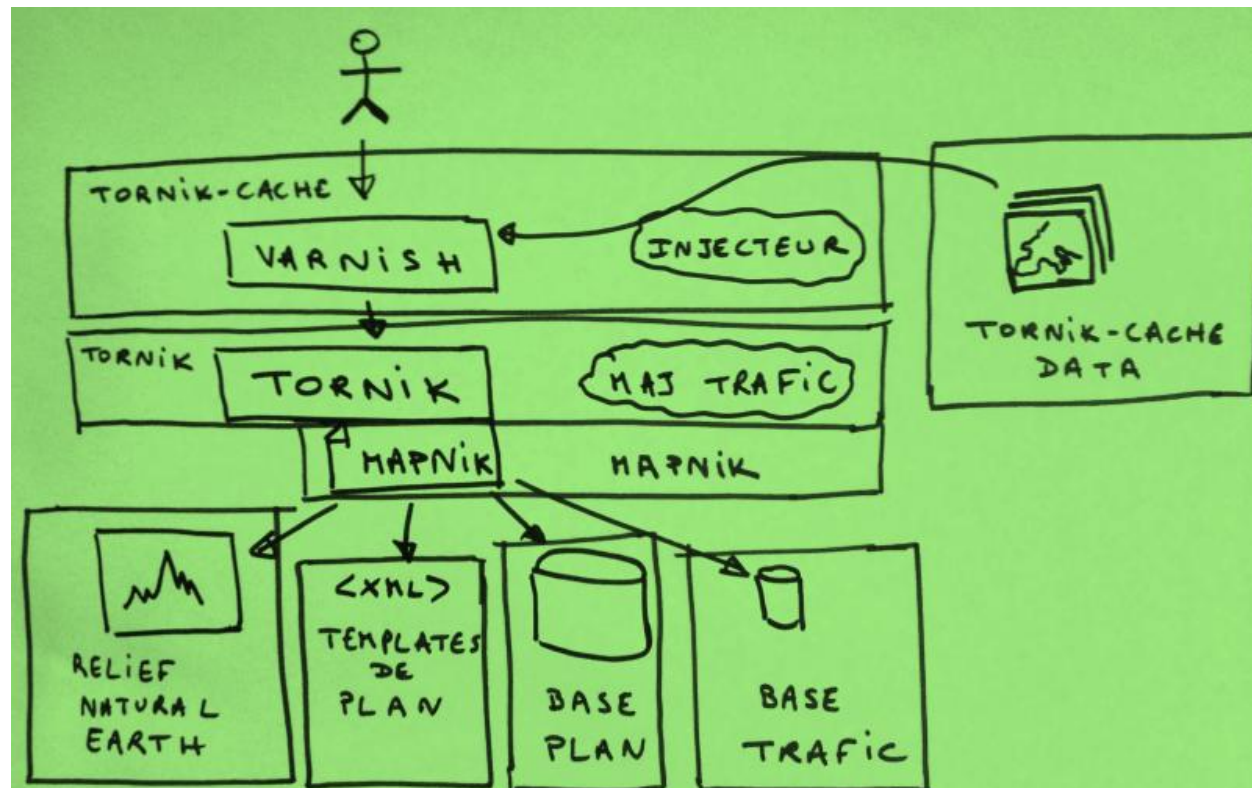
Notre route – points remarquables

- Capacité à challenger les choix historiques
- Obtention d'un meilleur consensus avec nos clients (Web/Mobile)
- Mise en place des outils versus configuration
 - prototype en moins de deux semaines
 - Configuration satisfaisante au bout de plusieurs mois
- Difficulté d'associer correctement de la méta-data aux images (zone réactive/copyright)



Notre route – choix techniques

- PostGIS : what else?
- Mapnik : validation marketing et technique
- Tornado :
 - serveur web python, déjà utilisé au sein de la DT Mappy
- Varnish :
 - performant
 - utilisé au sein de la DT Mappy



Notre route - utiliser c'est bien... ... contribuer c'est mieux!

- Légitimité pour participer à un projet open source
- Mise en place d'une politique open source dans l'entreprise
- Mise en place d'une organisation technique (responsable Open Source, Github d'entreprise)
- Contribution à Mapnik : discussion en cours pour intégrer l'implémentation du parallélisme avec PostGIS (25% de gain)
- Mise à disposition de l'outillage sur GITHUB (pycnik)

Les retours



Les retours – les chiffres

- Version en production (absorbe plus de charge que l'ancienne pf)
- 14 machines :
 - Varnish : 2 / Tornik : 8 / PostGIS : 4
- Débit sans cache :
 - Max 600 req/s
 - Débit moyen en prod 170 req/s
 - Pic en prod a 240req/s
- Débit avec cache : 2500 req/s (pic)
- Cache Hit Ratio : 92% (améliorable)
- Préchauffage du cache : 2 M d'objets
- Base Postgis : 75 Go
- Base Relief : 300 Go

Les retours ... des équipes (+)

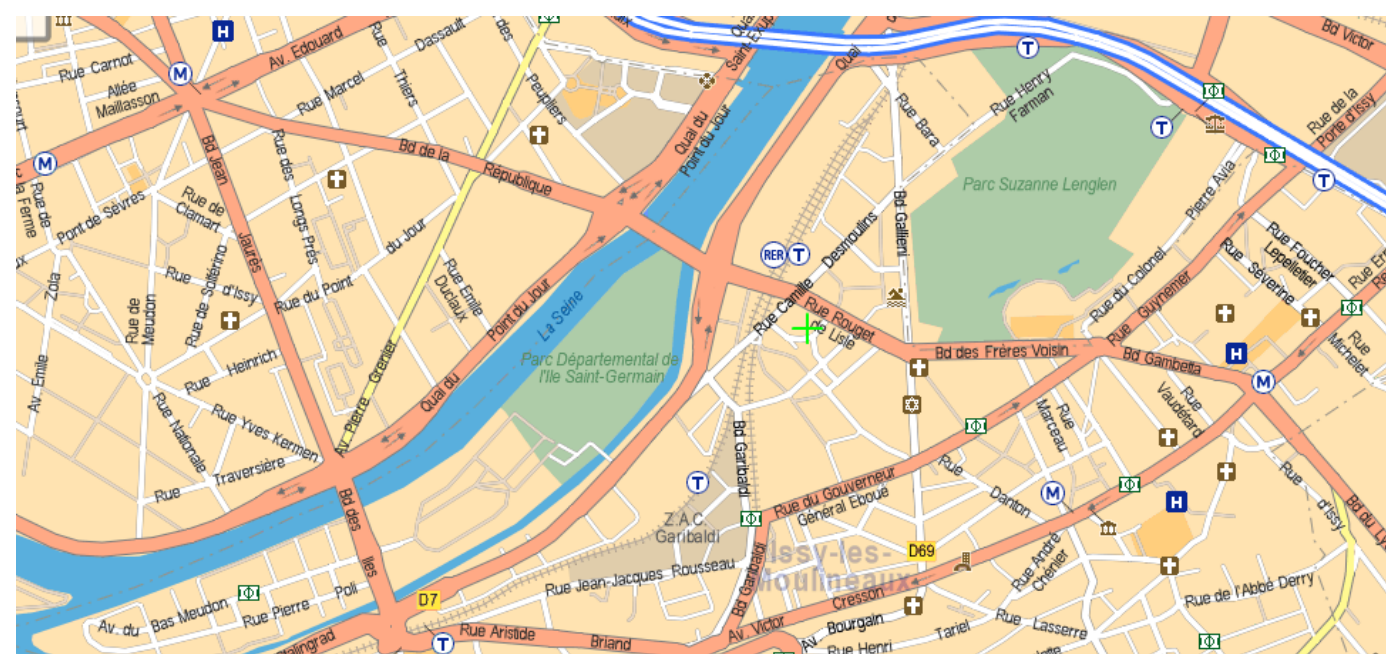
- Motivation de l'équipe
- Découpage du projet : prise en main progressive des outils
- Travail transverse au sein de Mappy
- Reprise en main approfondie de la production des cartes
- Les outils PostGIS et Mapnik
- Open source : les outils à disposition, la culture, contribuer, les contributeurs
- Plateforme scalable à moindre coût
- Le python, c'est bon!

Les retours ... des équipes (-)

- Trouver le bon curseur pour découper le projet
- Comprendre le comportement de certaines fonctionnalités sur Mapnik (Manque de documentation)
- Contribuer plus tôt sur la documentation et sur le code Mapnik
- Contribuer demande du temps (pousser un dev sur la master/délai de conversation sur un sujet)
- Améliorer notre communication technique
- Déploiement pas aussi rapide qu'espéré

Les prochaines étapes

- Sur la carte
 - Développement de nouvelles fonctionnalités (ex: Vectoriel)
 - Contribution régulière dans Mapnik
 - Migration vers PostGIS 2.X/9.X
- Plus largement
 - Appliquer cette méthode pour la refonte des autres services de la plateforme cartographique (en cours Géocodage....itinéraire à venir)



Merci



Questions