

PostgreSQL à la DGFiP

PGSession du 17 novembre 2021



PostgreSQL à la DGFIP

- 1. L'arrivée de PostgreSQL**
- 2. Les moyens mis en œuvre**
- 3. La confection des socles**
- 4. Les difficultés rencontrées**
- 5. Les challenges actuels**
- 6. L'avenir de PostgreSQL**

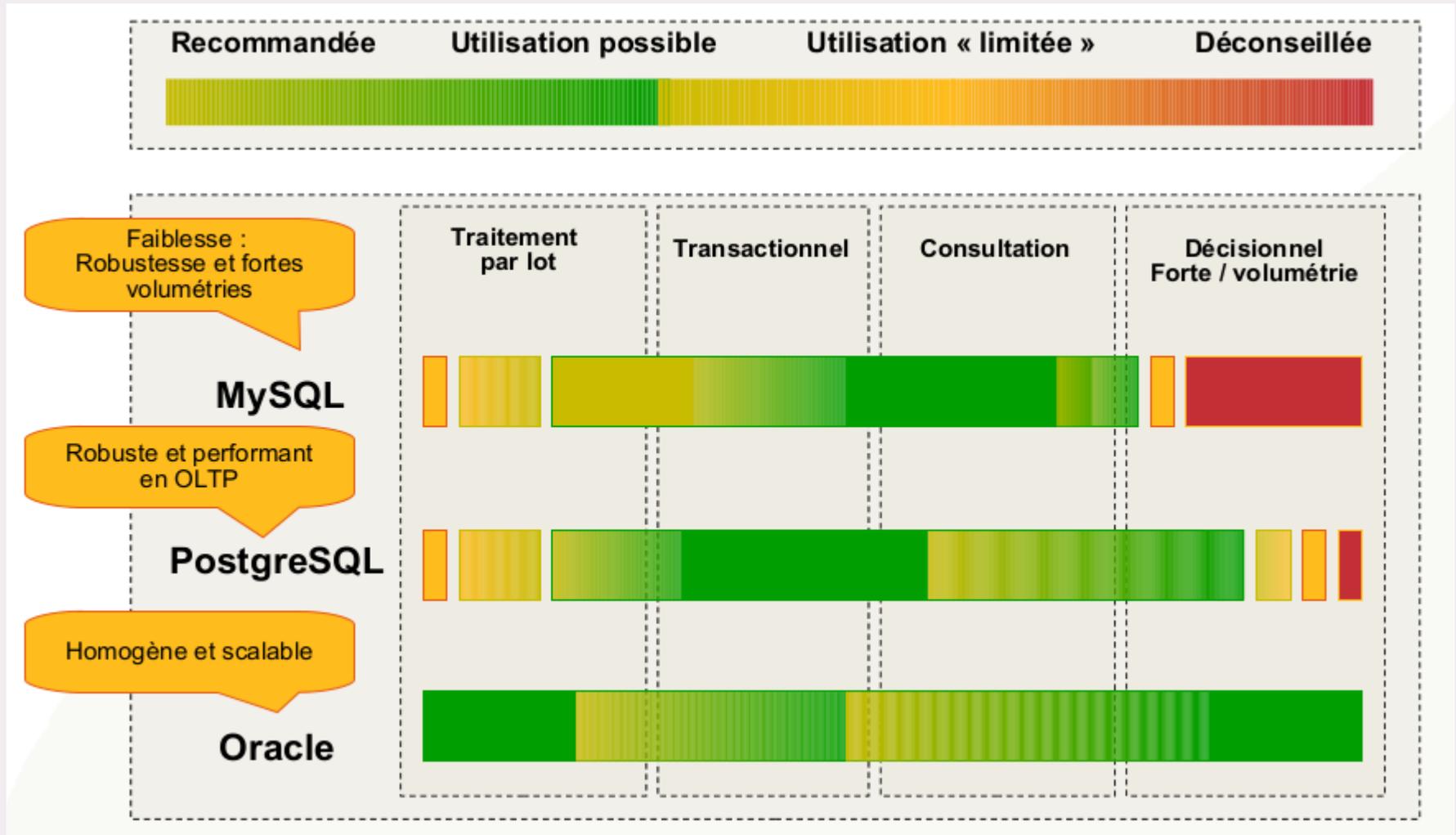
L'arrivée de PostgreSQL à la DGFiP

- **L'engagement de l'état français et de ses administrations à utiliser, à promouvoir et à contribuer aux outils libres**
 - Un axe fondamental de la politique des administrations françaises et des grands groupes français consiste à garder la maîtrise de leurs systèmes d'information
 - Afin d'assurer sa souveraineté et de limiter les coûts, les administrations ont décidé d'utiliser, de promouvoir et de contribuer aux produits libres

- **L'arrivée de PostgreSQL**
 - Les travaux autour de la base de données PostgreSQL ont démarré officiellement en 2007 à la DGFiP, avec plusieurs études chargées d'étudier l'état de l'art en matière de base de données relationnelles. Etude réalisée sur le premier marché de support logiciel libre de la DGI (2005-2008) (Capgemini, Linagora, Bull)
 - Trois SGBD ont été étudiées : MySQL / PostgreSQL / Ingres
 - Démarche générale en trois étapes : Etude théorique, Prototypage fonctionnel, Prototypage technique & Tests



Résultat de l'étude des alternatives libres 2007/2008



PostgreSQL à la DGFIP

1. L'arrivée de PostgreSQL
2. Les moyens mis en œuvre
3. La confection des socles
4. Les difficultés rencontrées
5. Les challenges actuels
6. L'avenir de PostgreSQL

Les moyens de mise en œuvre à la DGFIP (1/2)

- **Les moyens humains**

- Des équipes dédiées pour :

- Réalisation d'études de veille
 - Recueil des besoins auprès des projets
 - Réalisation d'un projet pilote
 - Conduite du changement
 - Formation des MOE et des exploitants
 - Confection des socles PostgreSQL
 - Apport du support auprès des projets sur les socles
 - Mise en place de la supervision
 - Mise en place des sauvegardes et des restaurations
 - Assistance à la mise en place de la haute disponibilité (notamment la réplication)
 - Accompagnement des projets pour la migration d'Oracle vers PostgreSQL



Les moyens de mise en œuvre à la DGFIP (2/2)

- **Les moyens matériels**

- L'utilisation des environnements virtualisés (OpenNebula, OpenStack, etc)
- L'utilisation des serveurs physiques pour des besoins de performances
- L'utilisation de baies de disques

- **L'expertise PostgreSQL**

- Le marché d'expertise PostgreSQL au niveau interministériel a été remporté par DALIBO (et OSLANDIA)



PostgreSQL à la DGFIP

1. L'arrivée de PostgreSQL
2. Les moyens mis en œuvre
3. La confection des socles
4. Les difficultés rencontrées
5. Les challenges actuels
6. L'avenir de PostgreSQL

La confection des socles techniques (1/4)

- **Chronologie des socles à la DGFIP**
 - Fin 2009 : premier socle PostgreSQL v8.1.4 sur HP-UX dédiés aux outils HP natifs, suivi d'une mise à jour à la version 8.3.8
 - A partir de 2010 : début des socles PostgreSQL v8.3.x sur socles LINUX dédiés aux projets de la DGFIP
 - De 2010 à 2018 : évolutions des socles PostgreSQL utilisant les versions 9.1, 9.4, 9.5 et 9.6 avec recompilations des paquets spécifiques à la DGFIP
 - A partir de 2018 : les socles de versions 10 et 12 utilisent les paquets communautaires
 - En 2022 : futur socle prévu avec la version cible 14 de PostgreSQL



La confection des socles techniques (2/4)

- **Composition des socles de la DGFIP**
 - Le noyau SGBDR complet (serveur, client, extensions « contrib », etc)
 - Des outils spécifiques DGFIP (gestion et administration des instances, audit)
 - La « Boîte-A-Outils » comprenant des outils d'analyse, de performance, de partitionnement, de chargements, de gestion des connexions, etc
 - Les « kits », un ensemble d'outils regroupés par fonctionnalités telles la migration (montées de versions, Oracle vers PostgreSQL), la réplication (physique et logique), la cartographie (« postgis »)
 - D'autres outils sont disponibles pour évaluation avant passage en production
 - Les socles fournissent également la documentation (procédures d'installation, de mise en œuvre, d'exploitation, fiches de description, bonnes pratiques, etc)
 - Des paliers correctifs (versions mineures, bugs, évolutions) sont disponibles périodiquement



La confection des socles techniques (3/4)

- **Processus de développement des socles de la DGFIP**
 - Recueil des besoins
 - Spécifications techniques
 - Développement et intégration
 - Recette et validation
 - Publication en production

- **Veille technologique**
 - Etudes techniques :
 - Evaluation de nouvelles fonctionnalités des versions majeures de PostgreSQL
 - Performances (parallélisme, partitionnement, etc)
 - Haute disponibilité utilisant PAF, RHCS, DRBD
 - Anonymisation des données
 - Sauvegarde du secondaire en réplication
 - Utilisation de « pgCrypto », de « pgPool2 », etc
 - Participations actives aux conférences et ateliers (« pgSession », « pgDay », etc)



La confection des socles techniques (4/4)

- **Liste des outils des socles de la DGFIP (non exhaustive)**
 - amcheck, check_postgres, pg_activity, pg_top, pgcenter, pg_catcheck
 - pgbadger, pgcluu
 - ora2pg
 - pg_repack
 - pglogical, repmgr
 - postgis2
 - pg_bulkload, pgloader
 - pg_partman, pg_pathman
 - pgbouncer
 - postgresql_anonymizer
 - etc



PostgreSQL à la DGFIP

1. L'arrivée de PostgreSQL
2. Les moyens mis en œuvre
3. La confection des socles
4. Les difficultés rencontrées
5. Les challenges actuels
6. L'avenir de PostgreSQL

Les difficultés rencontrées (1/2)

- **Accompagnement difficile vers PostgreSQL**

- Manque de projets sur l'utilisation d'une SGBDR libre à la place d'Oracle qui a déjà fait ses preuves
- Convaincre les projets de la fiabilité et des performances de PostgreSQL par la réalisation de quelques projets pilotes
- Démontrer la simplicité d'une migration d'Oracle vers PostgreSQL
- Conduite du changement
- Confusions Oracle/PostgreSQL :
 - Utilisation des objets en majuscules
 - Utilisation abusive des tablespaces
 - Utilisation des fonctionnalités Oracle dans PostgreSQL (recours aux modules tiers « orafce », etc)
- Contraintes applicatives :
 - Conception et modélisation des données hors périmètre
 - L'utilisation des « framework » rend les requêtes difficilement exploitables, d'où l'optimisation difficile



Les difficultés rencontrées (2/2)

- **Oracle privilégiée par quelques projets**
 - Manque de compétences sur PostgreSQL au sein de ces projets
 - Absence de fonctionnalité RAC dans PostgreSQL
 - Applications critiques

- **Contraintes en infrastructure et en sécurité**
 - En infrastructure :
 - L'utilisation des serveurs physiques limite le nombre de nœuds en haute disponibilité, etc
 - En sécurité :
 - Certains outils (Powa, pgAdmin4, etc) ne sont pas utilisables en production
 - En virtualisation (OpenNebula), certaine fonctionnalité telle que le « fencing » système n'est pas possible
 - De façon générale, la réplication « streaming » ne fonctionne pas à travers les pare-feux, etc



PostgreSQL à la DGFIP

1. L'arrivée de PostgreSQL
2. Les moyens mis en œuvre
3. La confection des socles
4. Les difficultés rencontrées
5. Les challenges actuels
6. L'avenir de PostgreSQL

Les challenges actuels

- **Le « cloud »**
 - Il existe aujourd'hui une plate-forme « cloud » interministérielle au sein de la DGFIP
- **Solution de haute disponibilité PostgreSQL dans le « cloud »**
 - Une POC est prévue avec la société DALIBO pour offrir une solution de haute disponibilité à plusieurs niveaux de résilience dans la plate-forme « cloud »



PostgreSQL à la DGFIP

1. L'arrivée de PostgreSQL
2. Les moyens mis en œuvre
3. La confection des socles
4. Les difficultés rencontrées
5. Les challenges actuels
6. L'avenir de PostgreSQL

L'avenir de PostgreSQL à la DGFIP

- **But**
 - S'affranchir d'Oracle
 - A la DGFIP, en quelques années, on est passé de quelques projets à quelques centaines sous PostgreSQL
 - Aujourd'hui, la plupart de nos applications sont sous PostgreSQL et les quelques unes utilisant Oracle sont sollicitées à migrer vers PostgreSQL
 - La DGFIP n'investira plus sur Oracle. A terme, il n'y aura plus de socle Oracle
- **Projets du Big-Data**
 - Des projets du Big-Data utilisent PostgreSQL, notamment :
 - DATALAKE interministériel de la DGFIP s'appuiera également sur PostgreSQL pour la partie des méta-données
 - DATAVIZ utilise les datamart sous PostgreSQL
- **Plate-forme « cloud »**
 - Ce serait aussi l'opportunité pour PostgreSQL, SGBDR « Open Source », de remplacer Oracle (dont les binaires n'y sont pas disponibles)



PostgreSQL à la DGFIP

Questions / Réponses

