

FORUM HPC XIII

28/11/2023

SUPPORT HPC :
AUDIT, PORTAGE ET
OPTIMISATION DE CODES

www.thalesgroup.com



Support HPC : Audit, portage et optimisation de codes

▣ 6 personnes à temps plein

▣ Types d'activités :

- Audit de code (analyses, étude de faisabilité)
- Optimisation (multiprocessing, dask , etc)
- Développement / Portage (C/C++, Fortran, Python) (IHM et Cœur de calcul)
- Possibilité de travailler sur des activités DR

▣ Les + :

- Connaissance de l'environnement CNES et de son cluster de calcul
- Un cadre contractuel allégé et des échanges réguliers

▣ Prestations de 1 à 12 mois sur bons de commandes (contacter L-SIS-poleHPC@cnes.fr)

OPEN | THALES GROUP LIMITED DISTRIBUTION | THALES GROUP CONFIDENTIAL | THALES GROUP SECRET

REF xxxxxxxxxxxx rev xxx – date Name of the company / Template: 87211168-DOC-GRP-FR-007

Ce document ne peut être reproduit, modifié, adapté, publié, traduit, de quelque manière que ce soit, en tout ou en partie, ou divulgué à un tiers sans l'accord écrit préalable de Thales © 2023 THALES. Tous droits réservés.

Support HPC : Audit, portage et optimisation de codes

▣ Exemples d'activités : Services sur le centre de Calcul

▸ Datalake

- Mise en place des environnements (création buckets, users, règles d'accès, ...)
- Adaptation du protocole S3 au sein du centre de calcul,
- Documentation, Support Utilisateur et Projet

▸ Datalabs

- Développements, évolutions
- Migration TREX
- Support

▸ Metrologie

- Rapports projets, dashboards

▸ Kubernetes

- Documentation et Support

▸ Amalthee

- Evolutions, développement, support

Support HPC : Audit, portage et optimisation de codes

▣ Exemples : Audit CARS

▸ Outil de traitement 3D d'images satellites :

- Génération de Modèles Numériques de Surface (DSM) par photogrammétrie, à partir d'images satellites acquises à différents points de vue



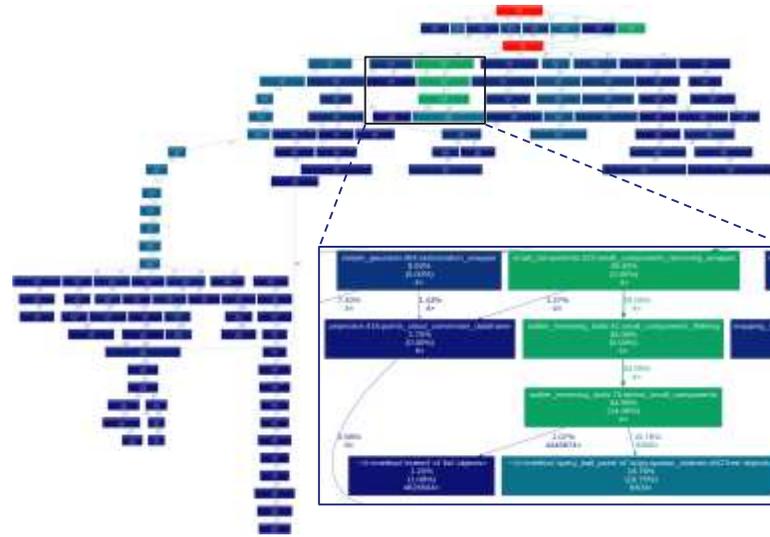
- Parallélisé avec Dask sur PBS
- Problèmes de performance lors de traitements massifs

Support HPC : Audit, portage et optimisation de codes

Exemples : Audit CARS

Analyse du code

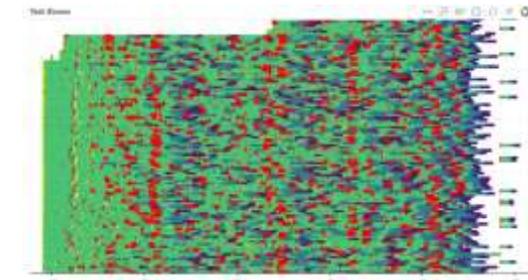
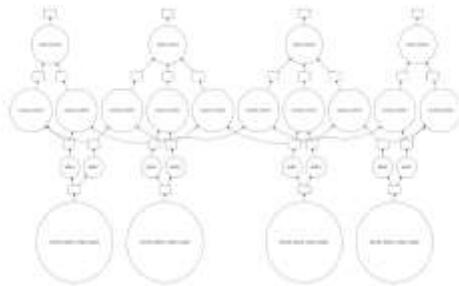
- Séquentiel
 - › CPU / MEM
 - › I/O
 - › Point durs du code
 - › Graphe d'exécution
 - › Identification ligne à ligne
- Distribué
 - › Étude de la parallélisation des tâches avec dask



```

Line  kfp  Time for 0.0 - 5 Time  Line Contents
-----
0      0
1      0
2      0
3      0
4      0
5      0
6      0
7      0
8      0
9      0
10     0
11     0
12     0
13     0
14     0
15     0
16     0
17     0
18     0
19     0
20     0
21     0
22     0
23     0
24     0
25     0
26     0
27     0
28     0
29     0
30     0
31     0
32     0
33     0
34     0
35     0
36     0
37     0
38     0
39     0
40     0
41     0
42     0
43     0
44     0
45     0
46     0
47     0
48     0
49     0
50     0
51     0
52     0
53     0
54     0
55     0
56     0
57     0
58     0
59     0
60     0
61     0
62     0
63     0
64     0
65     0
66     0
67     0
68     0
69     0
70     0
71     0
72     0
73     0
74     0
75     0
76     0
77     0
78     0
79     0
80     0
81     0
82     0
83     0
84     0
85     0
86     0
87     0
88     0
89     0
90     0
91     0
92     0
93     0
94     0
95     0
96     0
97     0
98     0
99     0
100    0

```



Support HPC : Audit, portage et optimisation de codes

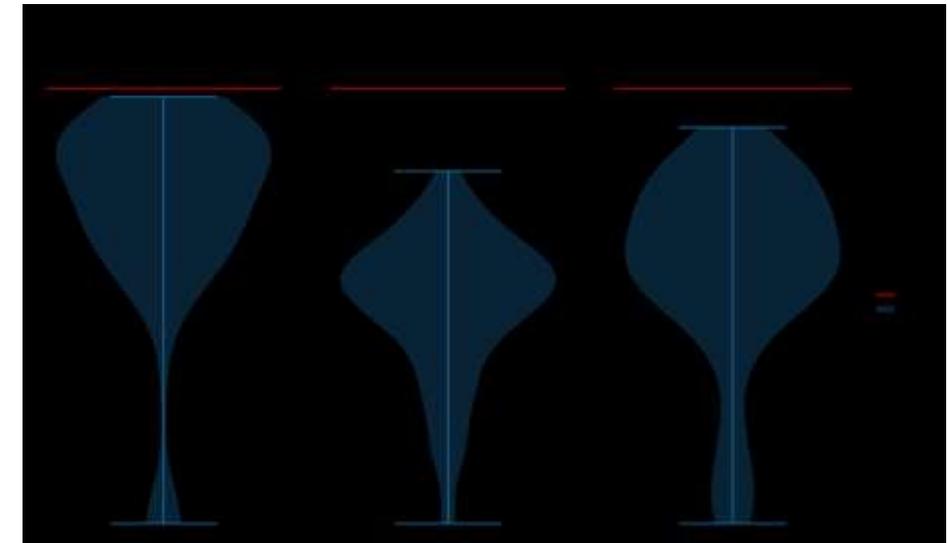
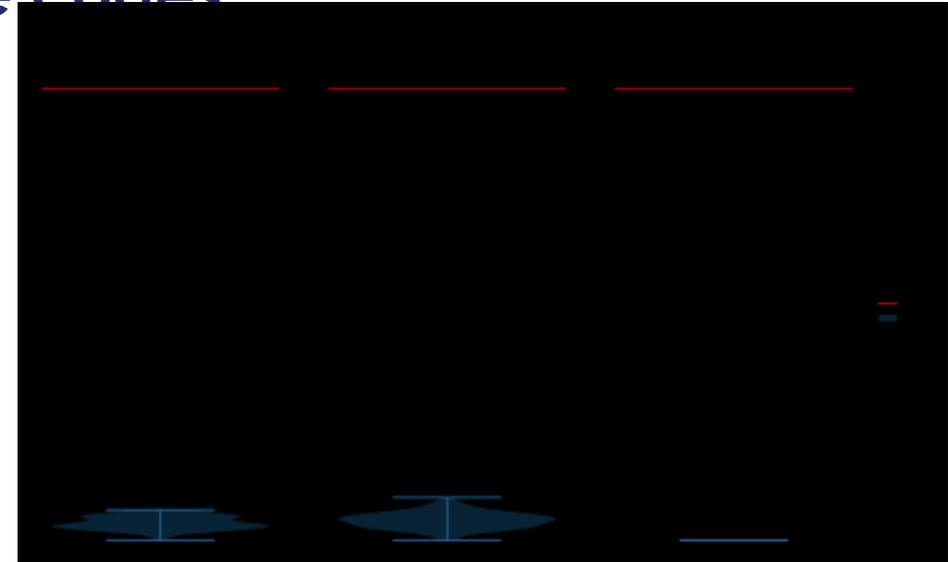
► Exemples : Audit CARS

► Analyses des ressources utilisées par les workers dask

- Aggregation des ressources requises / utilisées :
 - › Parallélisation interne aux workers inexistante → quantité de CPUs/mémoire
 - › Walltime de 10h pour contourner → latence élevée dans le scheduling des workers
 - › 4 CPU alloués par worker – 1 seul utilisé

► Conclusion

- Code bien conçu : pas d'amélioration significative à y apporter
- Amélioration de l'usage des ressources cluster :
 - › Walltime : pas de perte de temps en scheduling
 - › CPU : On quadruple l'efficacité de la parallélisation

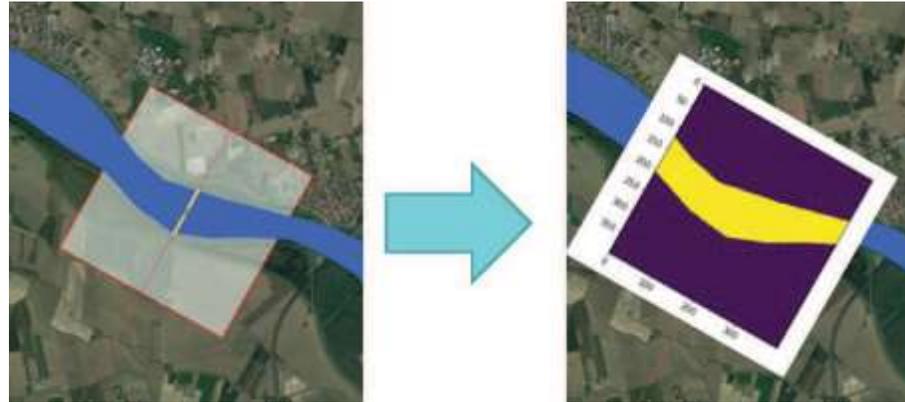


Support HPC : Audit, portage et optimisation de codes

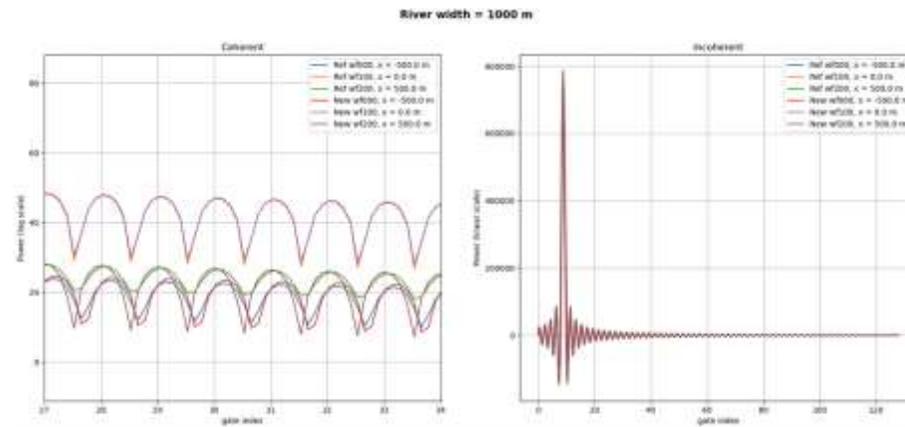
Exemples : Portage Nadir Hydro

Etat des lieux :

- Génération de masque d'eau
 - › Contour d'eau
 - › Orbite du satellite
 - › Généré manuellement via QGIS



- Simulations du signal altimétrique
 - › Code C
 - › Librairie C SHP embarquée



Support HPC : Audit, portage et optimisation de codes

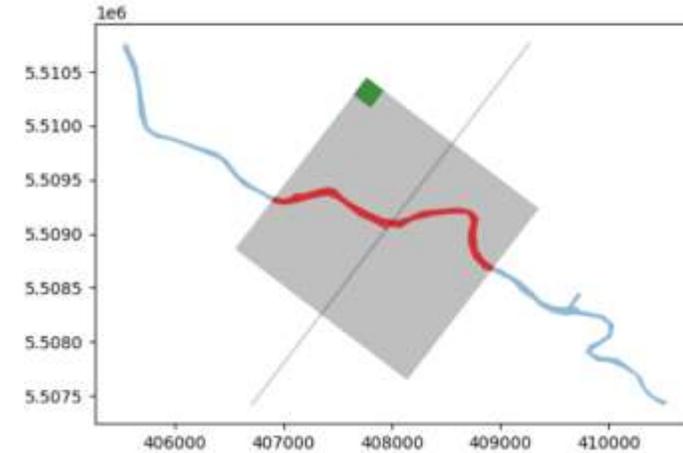
Exemples : Portage Nadir Hydro

‣ Génération du masque d'eau

- Implémentation python
 - Numpy / geopandas / shapely

‣ Simulations du signal altimétrique

- Etat des lieux du code existant :
 - Structure, dépendances, I/O, Langage, Librairies, Performance
- Propositions d'améliorations
- Implémentation python (amélioration performances)



Description	Temps d'exécution
Code C	2min30s
Python numpy amélioré	1min20s

‣ Ajout exécution distribuée sur les ressources de HAL/TREX (Dask, PBS/SLURM)

‣ Validation numérique des résultats

Support HPC : Audit, portage et optimisation de codes

► Exemples : POC OpenEO

▸ Openeo = framework mise à disposition API unifiée :

– Interface cliente

- Accès à des catalogues de données
- Travail sur des métadonnées
- Lancement de traitements sur le serveur

▸ Mise en place d'un backend OpenEO au CNES

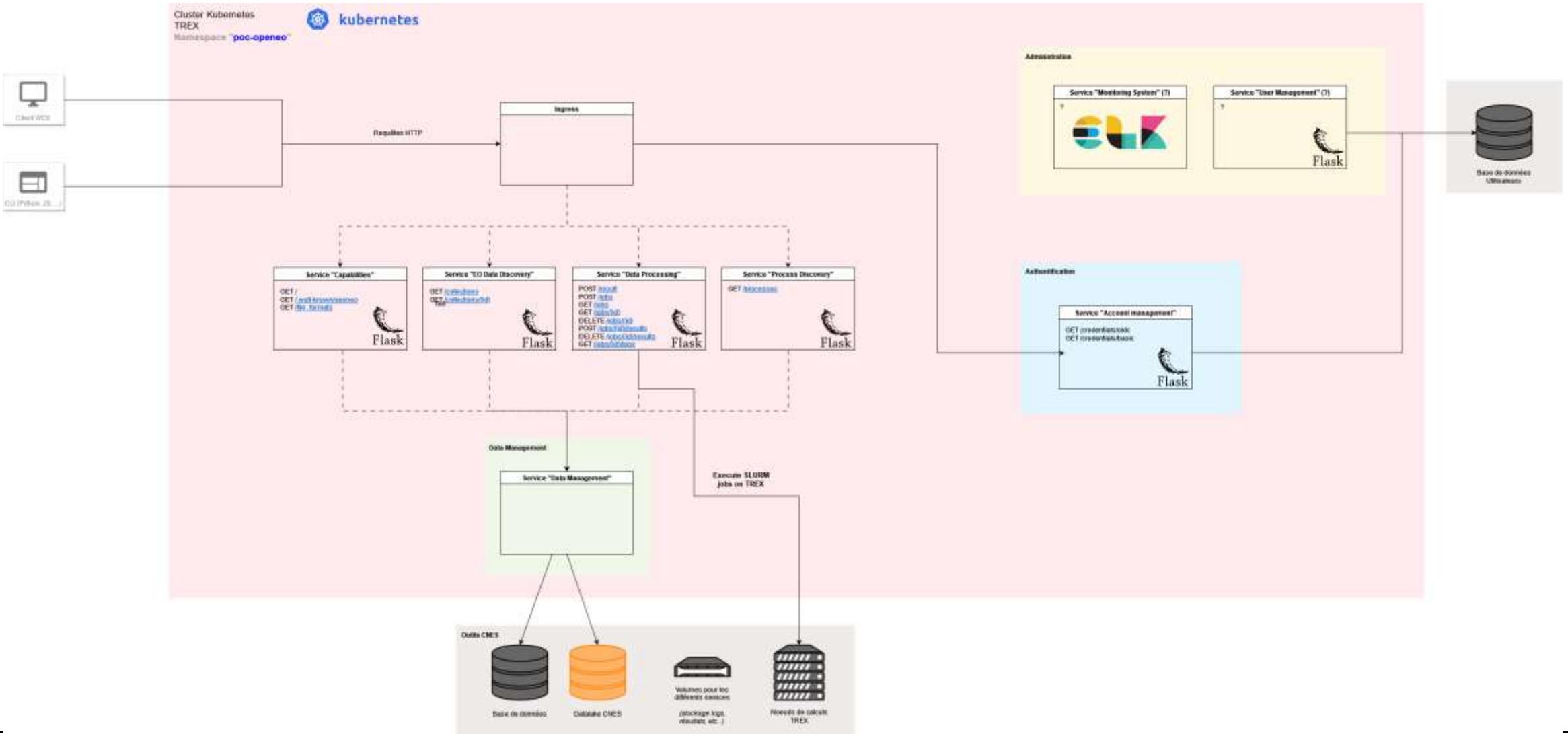
– Accédant à :

- S3 Datalake
- BDD Postgres
- Catalogue REGARD
- Des ressources de calcul

– Projet Beta-testeur de Kubernetes sur TREX

- Chaîne de traitement sous forme de microservices Kubernetes
- Scalabilité des microservices possible

Support HPC : Audit, portage et optimisation de codes



OPEN | THALES GROUP LIMITED DISTRIBUTION | THALES GROUP CONFIDENTIAL | THALES GROUP SECRET

REF xxxxxxxxxxxx rev xxx – date Name of the company / Template: 87211168-DOC-GRP-FR-007

Ce document ne peut être reproduit, modifié, adapté, publié, traduit, de quelque manière que ce soit, en tout ou en partie, ou divulgué à un tiers sans l'accord écrit préalable de Thales © 2023 THALES. Tous droits réservés.

Support HPC : Audit, portage et optimisation de codes

► **Contactez L-SIS-poleHPC@cnes.fr !**

► **Questions?**

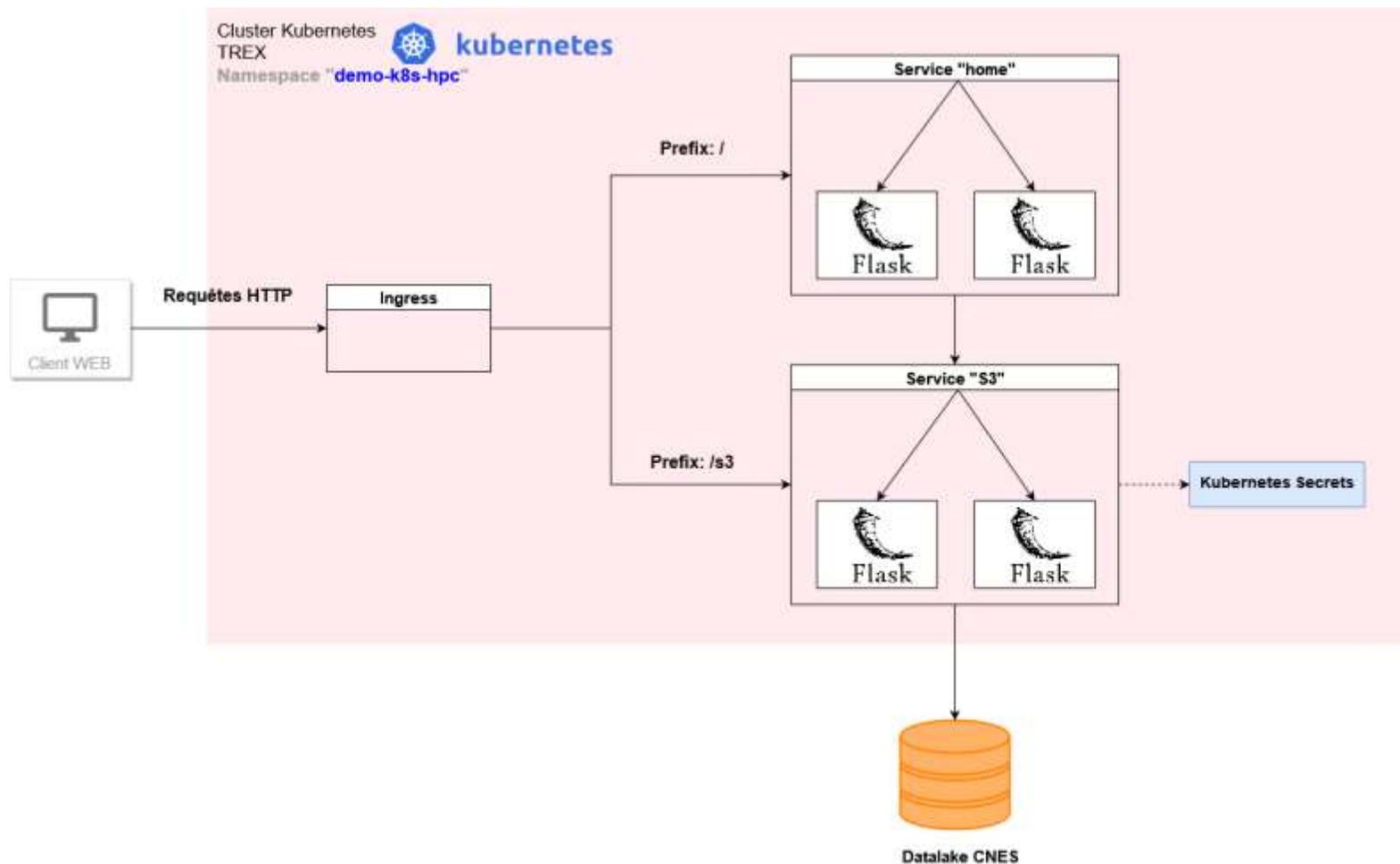
OPEN | THALES GROUP LIMITED DISTRIBUTION | THALES GROUP CONFIDENTIAL | THALES GROUP SECRET

REF xxxxxxxxxxxx rev xxx – date Name of the company / Template: 87211168-DOC-GRP-FR-007

Ce document ne peut être reproduit, modifié, adapté, publié, traduit, de quelque manière que ce soit, en tout ou en partie, ou divulgué à un tiers sans l'accord écrit préalable de Thales © 2023 THALES. Tous droits réservés.

Support HPC : Audit, portage et optimisation de codes

► Démonstration Kubernetes



OPEN | THALES GROUP LIMITED DISTRIBUTION | THALES GROUP CONFIDENTIAL | THALES GROUP SECRET

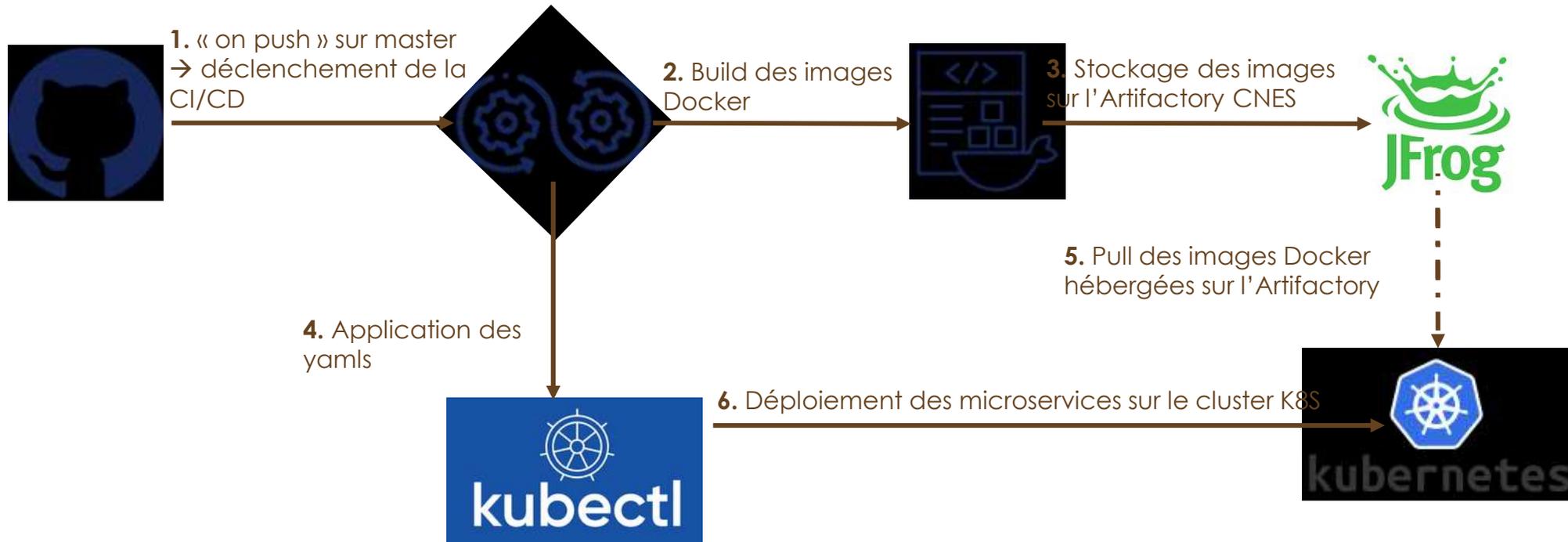
REF xxxxxxxxxxxx rev xxx – date Name of the company / Template: 87211168-DOC-GRP-FR-007

Ce document ne peut être reproduit, modifié, adapté, publié, traduit, de quelque manière que ce soit, en tout ou en partie, ou divulgué à un tiers sans l'accord écrit préalable de Thales © 2023 THALES. Tous droits réservés.

Support HPC : Audit, portage et optimisation de codes

► Démonstration Kubernetes

- ▶ Construction d'images docker contenant l'applicatif
- ▶ Déploiement facilité en yaml
- ▶ Exemple de workflow déploiement applicatif sur le cluster Kubernetes



OPEN | THALES GROUP LIMITED DISTRIBUTION | THALES GROUP CONFIDENTIAL | THALES GROUP SECRET

REF xxxxxxxxxxxx rev xxx – date Name of the company / Template: 87211168-DOC-GRP-FR-007

Ce document ne peut être reproduit, modifié, adapté, publié, traduit, de quelque manière que ce soit, en tout ou en partie, ou divulgué à un tiers sans l'accord écrit préalable de Thales © 2023 THALES. Tous droits réservés.