

CSAN : Comprehensive Software Archive Network

Un projet Decalog - ComputeOps

Martin Souchal

Octobre 2022

IN2P3



1. DECALOG

ComputeOps

2. Organisation DevOps

3. Ressources pour le DevOps

Gitlab

Projet SingularityHub

Projet CSAN

DECALOG

10 ans pour gagner un facteur 10 dans le logiciel.

- Master projet IN2P3 : pas vocation à avoir de lourdes activités et objectifs en propre, mais plutôt de favoriser et d'accueillir les projets de R&D transverses IN2P3 qui s'inscrivent dans ses thématiques.
- 2 projets en cours :
 1. <https://reprises.in2p3.fr/> Reprises (calcul reproductible et optimisation)
 2. ComputeOps (portabilité du calcul)

- **étudier** les technologies de conteneurisation logicielles sur du matériel de calcul hétérogène (CPUs à instructions vectorielles, GPUs, FPGAs,...).
- **étudier** les possibilités offertes par les technologies de conteneurisations logicielles dans une optique de portabilité et de reproductibilité des codes.
- **étudier** et **diffuser** les bénéfices des pratiques “DevOps” dans le milieu de la recherche

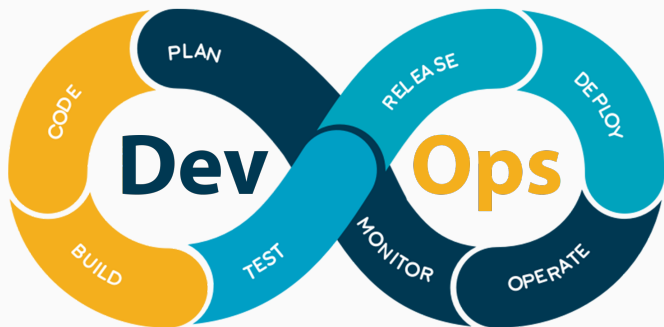
- **Participants** : labos IN2P3 (APC, IJCLAB, LPC, LPNHE, LLR, IPHC, CC), extérieurs (IAS, INRAE, CEA, Universities, INS2I, CNES) et secteur privé (Sylabs, INTEL).
- **Comparer** les différentes technologies de conteneurs (docker, rocket, lxd, udocker, singularity, shifter, apptainer) : effet sur les performances, la vectorisation, accès aux accélérateurs de calcul, sécurité, facilité d'administration et d'utilisation. Y'a-t-il un avantage décisif à dédaigner l'outil dominant (Docker) et à se tourner vers les alternatives "HPC" (Apptainer, Singularity, Shifter,...) ?
- **Comparer** les différentes technologies d'orchestrateurs (Kubernetes, Nomad...)
- **Étudier** l'inter-opérabilité des technologies. Notamment les images et les fichiers de reconstruction d'image. Par exemple, est-ce qu'un utilisateur pourrait développer sur son poste en Docker, puis déployer sur centre de calcul en Singularity ?
- **Valider** la compatibilité des conteneurs avec la grille.

- **Animation** et coordination d'un réseau inter-instituts (conseil et tests d'infrastructures clouds/HTC/HPC)
- **Veille technologique** (formation, ateliers, workshops, publications...)
 1. Workshops in 2018 and 2020
 2. Posters and conference proceedings at CHEP 2018, 2019 and 2020
 3. Singularity pool at IJ (Oct 2018).
 4. Singularity training at SBAC-PAD (Sept. 2018).
 5. JDEV 2017: Training transport its parallel applications with LXD and Singularity containers.
 6. Docker/Singularity at FACe (Jan. 2018), Docker Workshops at APC during summer 2020
 7. Containers for Computing - AI DevTalks INRIA (July 13, 2021)
 8. Containers for Computing - UST4HPC Training (January 20, 2021)
 9. Kubernetes for Computing - IN2P3 CC Kubernetes Day (February 19, 2020)

- **Comparatif** des technologies de conteneurs pour le calcul IN2P3
- **Services** à la communauté
 1. Singularity Hub pour la Recherche
 2. CSAN ("marketplace" pour le milieu de la recherche avec des codes disponibles et optimisés pour les différents mésocentres/clouds)
 3. Conteneurisation d'applications pilotes

Organisation DevOps

*Le devops — ou DevOps (selon la graphie habituellement utilisée en langue anglaise) — est un mouvement en ingénierie informatique et une pratique technique visant à l'unification du développement logiciel (dev) et de l'administration des infrastructures informatiques (ops), notamment l'administration système.
Ref <https://fr.wikipedia.org/wiki/Devops>*



Ressources pour le DevOps

L'instance Gitlab de l'IN2P3 propose toutes les ressources nécessaires pour travailler en mode DevOps :

- Hébergement de codes sources
- Travail collaboratif
- Fonctionnalités intégration continue et livraison continue
- Runners disponibles (partagés ou privés)
- Ouvert largement à l'ESR



Origine du projet :

- Organisation du calcul en mésocentre, mutualisation des ressources
- Diminution des emplois en ASR
- Open-science
- Convergence des orchestrateurs de conteneur et des gestionnaires de job
- Présence de singularity dans la grande majorité des mésocentres

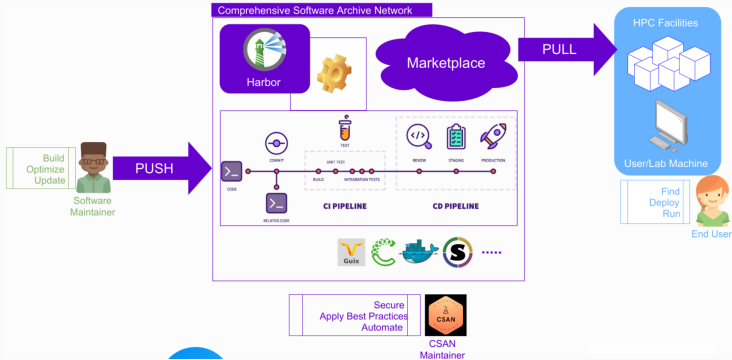


→Catalogue de conteneurs Singularity prêts à l'emploi

- L'initiative du projet CSAN (Comprehensive Software Archive Network) est de proposer à la communauté ESRI un **catalogue** d'applications scientifiques open-source prêtes à l'emploi et optimisées pour les centres de calcul nationaux et les mésocentres.
- La plateforme CSAN permettra aux **auteurs** de logiciels d'y déposer les versions de leurs sources (ou de pointer vers des dépôts versionnés)
- Via des méthodes d'**intégration** et de **développement** continu, rendre accessible et installable sans effort sur différents systèmes d'exploitation et au plus grand nombre d'utilisateurs et d'infrastructure possible les codes

- Ce catalogue d'application sera consultable par le biais d'un **portail web** et d'une **API ouverte** pour l'intégration dans les différents mésocentres.
- infrastructure **multi-site** entre l'IPHC à Strasbourg et le mésocentre de Montpellier
- Instances **féderables**
- Prêt à s'intégrer au projet **Software Heritage**

CSAN Workflow



- POC réalisé à l'APC en 2021
- Développements effectués pour adapter l'outil Harbor a Singularity
- Présentation aux JCADs et JRES 2021 et 2022 : bonne réception du public
- Finalisation de la première version en production en cours
- Infrastructure testée et validée

- **Force** : sujet de R&D très attractif
- **Faiblesses** : départ de plusieurs membres qui n'ont pas été renouvelés.
- **Opportunités** : le projet CSAN s'intègre dans la démarche de Science Ouverte avec une connexion directe avec Software Heritage
- **Menaces** : la volatilité des solutions de conteneurs (Singularity racheté par Sylabs , fork de la communauté open source avec Apptainer) demande un suivi constant de R&D

Merci beaucoup!

GitLab: ComputeOps

Site: `https://si-apc.
pages.in2p3.
fr/
face-website/
computeops/`

