

Compte-rendu HEPiX d'automne 2018 Barcelone, 8-12 octobre 2018



- <https://indico.cern.ch/e/hepixon-autumn2018>
- Nouveau record de nombre de participants : 137 !
- Les présentations par track (69 au total) :
 - Site reports (16 +3)
 - End users IT services & OS (7 +3) Sébastien
 - Grid, cloud and virtualization & OS (3 -4) Sébastien
 - IT Facilities and business continuity (4 -5) Sébastien
 - Computing and batch systems (9 -2) Jean-Michel
 - Security and networking (11 +1) Jean-Michel
 - Storage and filesystems (11 +2) Pierre-François
 - Basic IT services (4 -2) Pierre-François
- News du board et prochains HEPiX



Retour HEPiX

End-User IT Services

- ▶ « Promotion » de l'Open Source : présentation
- ▶ **But** : étudier la possibilité de se séparer des solutions logicielles propriétaires et commerciales
 - Volonté de fournir les mêmes services à tout les utilisateurs du CERN
 - Prix de ces solutions en augmentation (licence par utilisateur)
 - Car nombre d'utilisateurs en augmentation
 - Problème d'emprisonnement dans un (éco-)système
 - Garder le contrôle des données
- ▶ Procédé
 - Remplacer « élément par élément » en partant des « core softwares »
 - Mais pas une solution par une solution équivalente
 - Adresse les besoins (plutôt que les solutions technos)
 - Favorise l'indépendance des éléments
 - Favorise les API standards ou ouverts
- ▶ Produits étudiés (ou en cours)
 - Service mail, Adobe Acrobat, Sharepoint, ...
 - « Suite » LaTeX, OfficeOnline, au-dessus de CERNBox/OwnCloud

▶ « PaaS pour les applications web » : présentation

▶ Objectifs

- Web
 - Fournir flexibilité et isolation (entre sites)
- Applications tierce partie (Mattermost, GitLab, JIRA, Discourse, Rundeck, Jenkins, ...)
 - Déploiement facilitée
- Mise à la disposition de « templates »
 - Isolation (pour chaque instance), provision « self-service », délégation des droits admin.

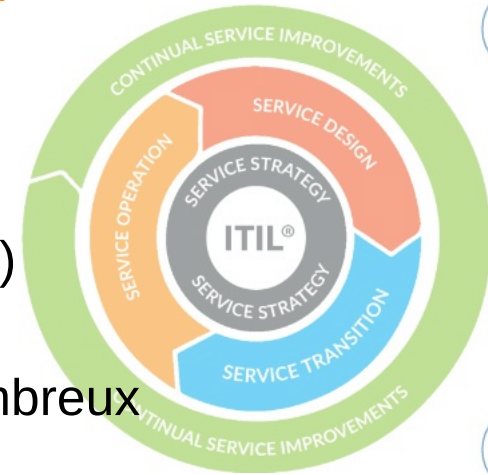
▶ Techno

- Basé sur RH OpenShift (a.k.a. OKD)
- Haute disponibilité : HA, monitoring, data backup, certificat TLS, ...
- Intégration complète dans l'environnement CERN : SSO, stockage, GitLab...



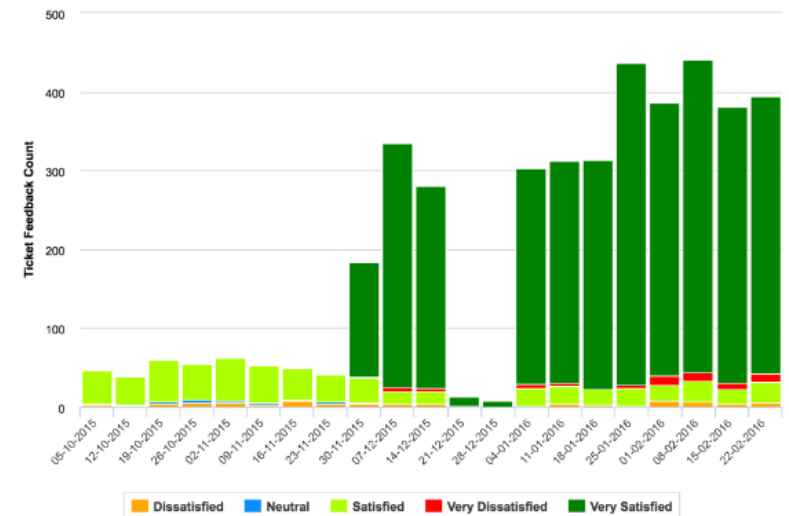
▶ Rationalisation et simplification de la gestion de service : présentation

- ▶ Nécessaire pour la mise à l'échelle des services
 - Simplification de la mise en production
 - Rationalisation (pas de séparation entre services IT, RH, sites, ...)
- ▶ Définition de KPI (indicateurs de performances) simples et peu nombreux



- ▶ Expérience utilisateur améliorée
 - Quelque soit le canal utilisé, même procédure dans le traitement de la demande
 - Plus de retours et retours beaucoup plus positifs !
 - 4 choix possibles
(très content, content, insatisfait, très insatisfait)

- ▶ Disponibilité des services
 - Code 3 couleurs (Ok, dégradé, down)



- ▶ (Plein) d'autres cas en cours : CMDB, portail utilisateur, RGPD, SLAs, ...

▶ Indico : présentation

- Ré-écriture du code pour la version 2
- Version actuelle 2.1
- À venir : réservation de salles, revue des proceedings de conférence, internationalisation

▶ CERN Linux : présentation

- Le CERN participe à des SIG au sein de CentOS 7
 - Cloud (OpenStack), stockage (CEPH), virtualisation (Kubernetes), « software collection »
- Support des versions :
 - SLC6 → 11/2020, CentOS7 → 04/2024
- Intégration GitLab/Koji
 - Permet de construire paquets et images docker (accessible à l'utilisateur)
- CentOS8 ? RHEL8 en beta cette année

Retour HEPiX

Grid, Cloud and virtualisation

▶ Extending local computing facilities using Helix Nebula Science Cloud : présentation

- 2 fournisseurs de cloud toujours présents dans la phase pilote actuelle :
 - T-System (OpenStack), RHEA (CloudStack)
- 3 cas à évaluer : MAGIC, CTA et HTCondor (nouveau !)
- Conclusion 29/11/2018@CERN

▶ A Data Lake Prototype for WLCG : présentation

- HL-LHC (ainsi que la plupart des communautés scientifiques) voient leur besoins en stockage bien supérieurs à ce qu'elles pourront obtenir
- Les concepts de gestion de données doivent être revus pour optimiser le coût :
 - Nombre de réplica mieux gérer (locaux vs globaux)
 - « File workflow » : meilleure prise en compte de la data popularity
 - QoS : prise en compte des performances voulues & du coût du média support
- Prototype « eulake » pour PoC
- Systèmes de gestion des données de CMS et ATLAS intégrés
- Intégration des framework de HammerCloud
- Mise en place de différents WGs : access, distribution, storage class QoS

Retour HEPiX

IT Facilities and Business Continuity

- ▶ Cost and system performance modelling in WLCG and HSF : présentation
 - Nécessité de mieux comprendre l'utilisation des ressources afin d'optimiser l'efficacité de leur utilisation (HL-LHC)
 - Mieux comprendre les workloads, l'utilisation des ressources et les coûts
 - Définition des métriques pour caractériser les workloads et mesurer les performances d'utilisation des ressources
 - Définition d'un framework pour modéliser les modèles de calcul des expériences (et les besoins associés)
 - Estimation des coûts des sites
 - Travail exploratoire sur divers aspects en cours
 - Collaboration avec d'autres WGs (DOMA)

Track Network & Security

Cern

- Network config management
- CERN Campus and Technical networks upgrade
- WiFi @ CERN
- IoT @ CERN

Software defined networks (SDN)

- Network Functions Virtualisation Working Group
- IHEP SDN

Autres

- IPV6 working group
- OSG Network update

Networking @ CERN (1)

Network config management

- 400 routeurs et 4000 switchs multi-vendeurs
- Solution actuelle CFMGR (BD+scripts perl)
- Recherche d'une solution open-source python
- Association d'outils autours de N.A.P.A.L.M.

CERN Campus and Technical networks update

- 2 réseaux distincts : Campus et Technical
- Des besoins différents : Campus=nombre connexions, Technical=résilience
- Solutions (différentes) Juniper
- Campagnes d'upgrade du réseau Technical durant les log shutdowns, upgrade Campus entre les deux

Networking @ CERN (2)

WiFi @ CERN

- WiSE : WiFi Service Enhancement project
- ~200 bâtiments et outdoor, Meyrin et Preveessin
- Solution ARUBA : centralisation et redondance
- Presque terminé : 3700 bornes, 300 switches, 7000 utilisateurs/jour

IoT @ CERN

- Des appareils très divers, connexions diverses
- Outdoor wireless : cellulaire LPWAN : LoRA
- Indoor wireless : Bluetooth, Zigbee, WiFi
- Indoor wired : Tunnelling to controllers
- La sécurité doit rester la priorité !

Software defined networks (1)

Network Functions Virtualisation working group

- Phase exploratoire : étudier les cas d'usage en intra-site et en multi-site pour optimiser l'accès au stockage de bout-en-bout.
- 2 projets pilotes au CERN :
 - Tungsten fabrik : test multi-stack
 - Datacenter interconnect : SDN multi-site
- Optimisation de l'accès au stockage
 - Configuration réseau piloté par les applications
 - AmLight SDN project pour LSST

Software defined networks (2)

IHEP SDN

- Besoins : allocation adresses, DHCP, contrôle accès,
- Comparaison 2 solutions : Huawei ou Ruijie
- Plans : remplacement des équipements et full SDN
- Besoin de plus d'outils de monitoring pour interroger les contrôleurs

IPV6, OSG, Perfsonar

Update from the HEPiX IPV6 working group

- Tests FTS sur les T1s : OK sauf quelques sites
- Déploiement IPV6 sur les T2 (stockage dual stack fin 2018) :
 - France dans le peloton de tête
 - 41 % fait (oct 2018), très peu de sites posant problème

OSG/WLCG Networking update

- WLCG Network Throughput working group
 - Coordinate network performance incidents
 - WLCG Perfsonar Infrastructure
 - Perfsonar 4.1 released (Aug 2018), Drop SL6 support
 - 288 Perfsonar hosts
 - Nouveau Dashboard Perfsonar sur Grafana au CERN
 - A venir : Perfsonar 4.2 (Q1-2019), Perfsonar 4.3 (Q3-2019)

Security

Computer security

Personal Data protection@ CERN

Framework for OpenScience CyberSecurity

Computer security

Revue des principales vulnérabilités 2018

- Vulnérabilités du microcode (spectre, meltdown and al.)
- Vulnérabilités BMC management (Quanta)
- Phishing visant les universités
- Divulgations de mot de passe (Facebook, Google) : système d'alerte (envoi de mail aux contacts sécurité des sites)

Conclusion

- Capacité à corriger rapidement et automatique les vulnérabilités
- Log des activités
- Choisir la sécurité face à la facilité

Openscience CyberSecurity

Le CERT de la NSF ?

Un référentiel de bonnes pratiques

Des outils

Des alertes

Track Computing and Batch Services

CPU, performance, benchmarking

- AMD EPYC
- Benchmarking WG update
- Test de la « scalabilité » OpenMP en userland

Batch systems

- Too much beer
- Report on HT-Condor workshop

Autres

- Data analysis as a service
- RHIC Central computing workshop
- PDSF migration to Cori
- T0 reconstruction@ PizDaint CSCS

CPU, performances, benchmarking

AMD EPYC

- Jusqu'à 32 coeurs physiques (64 logiques)
- Comparé au Intel Skylake-SP
- Performances similaire mais AMD-EPYC intéressant économiquement (HEPSPEC06/Dollar)
- A venir en 2019 : AMD-EPYC2 et Intel Xeon Cascade-Lake

Benchmarking working group update

- 2 axes :
 - Fast-benchmark
 - nouvelle métrique (pledges) pour remplacer HEPSEPC06
- Janvier 2018 : Etude de l'impact des correctifs pour Spectre et Meltdown : $\pm 3 \%$
- Fast benchmark DB12 ok mais pas assez fiable pour une référence
- Etude de SPEC CPU2017
- Travail en cours avec des « workloads » réels des expériences

Batch systems

Retour sur le workshop HT-Condor

- Sept. 2018 au RAL, UK. 54 participants, 38 contributions
- Tutoriaux, conseils, développeurs présents
- Retours d'expérience
- Sujet important : l'accounting

Too much Beer

- Lancement de containers via HT-Condor sur les serveurs de disques EOS
- Definition de limite pour ne pas impacter le storage
- MC production (2500 cores par Atlas, 1300 par LHCb)
- Projet d'intégration dans un cloud Kubernetes

Computing (2)

Report on the workshop on central computing in Photon Science

- RHIC
- Participation de BNL, SLAC, LBL, Argonne, STFC, Diamond, DESY
- Répondre aux challenges sur les données et le calcul

Migration PDSF vers Cori

- Solution pour importer les données CVMFS dans un cluster type HPC (4 DVS servers : I/O forwarding et cache)
- SLURM

Computing (2)

Activités T0 dans le cluster PizDaint CSCS

- Un autre exemple d'utilisation de cluster HPC
- Environnement Linux restreint
- Pas de disque local, pas de swap, 2GB/core
- CVMFS préchargé dans GPFS et exposé aux workers via DVS
- Utilisation de ARC-CE, ARC-datamovers
- Efforts d'intégration coûteux et adaptation nécessaire

Workflows automation @CERN IT with CERNMegabus (Zhechka Toteva CERN)

- CERNMEGABUS: SERVICE DE MESSAGES (ACTIVEMQ)
- AMÉLIORER LA RÉSILIENCE ET LA MISE À L'ÉCHELLE D'OUTILS DE GESTION DE SYSTÈME.
- ÉTUDE DE LA GESTION DE COUPURE ÉLECTRIQUE (ACTIONS AUTOMATIQUES EN CAS DE MESSAGE 'POWER CUT' DURANT LES 20MIN DE COURANT SECOURS)

Deployment of compute nodes for the WLCG with Cobbler, Ansible and Salt (Damien François UC Louvain)

- WHICH IS WHAT FOR WHAT

The new CERN Authentication and Authorization Paolo Tedesco CERN

- ACTUELLEMENT MÉLANGE DE AD, LDAP, SSO, CERTIFICATS ...
- OBJECTIF : PROPOSER UN SCHÉMA UNIFORME : JETONS SSO OU WLCG TOKENS ET FACILITER ACCÈS DES NON-CERN VIA LES FÉDÉRATIONS D'IDENTITÉ

Storage at CERN (Cristian Contescu CERN)

- MISE EN PLACE DE CEPHFS EN REMPLACEMENT DE NFS

LHC Long Shutdown 2 and database changes (Ludovico Caldarra CERN)

- IMPACT OF VENDOR RELEASE CHANGES ON STORAGE SYSTEMS AND EXPERIMENTS TIMELINES: EXAMPLE OF ORACLE DATABASES AT CERN

Latest developments of the CERN Data Management Tools (Edward Karavakis CERN)

- RENOUELEMENT ET ADAPTATION
- FTS / DPM/ XROOTD / EOS

Updates on ATLAS Data Carousel R&D (Xin Zhao BNL)

- TESTS DE PERFORMANCE D'EXECUTION DE RUN TEST DIRECTEMENT DEPUIS LES BANDES

The OSiRIS Project: A Multi-Institutional Ceph Storage Infrastructure (Shawn Mc Kee University of Michigan)

- CLUSTER CEPH RÉPARTIT SUR 3 SITES DANS LA RÉGION DES GRANDS LACS
- INTERCONNEXION 100 GBP/S, NVME POUR CACHES BLUESTORE ET TIERING

LTO experiences at CERN (German Cancio Melia CERN)

- ALGORITHME POUR OPTIMISER LE POSITIONNEMENT DE LA TÊTE DE LECTURE DES DÉROULEURS LTO

Backup Infrastructure at CERN (Aurelien Gounon CERN)

- OVERVIEW 'NEW' THEMES BOTH IN INFRASTRUCTURE AND IN USER REQUIREMENTS
- FUTUR: EOS

CERN Tape Archive initial deployments (CTA) (Julien Leduc CERN)

- TAPE + EOS = (LOVE)
- TEST AVEC LES RUNS ALICE IONS LOURDS PUIS ÉTENDUS À TOUTES LES EXPÉRIENCES
- OPTIMISATION DES PERFORMANCES PAR INSTANCES EN ÉTALANT LA CHARGE SUR PLUSIEURS SERVEURS DE METADATA
- PARTICIPATION AUX TESTS AVEC ATLAS DATA CAROUSEL

BOF (BIRD OF A FEATHER)

- 7 PRÉSENTATIONS SUR RÉSULTATS DE TESTS AVEC ATLAS DATA CAROUSEL
- CC IN2P3 +++

3 PRÉSENTATIONS DES SPONSORS

FUJIFILM: DEVELOPMENT OF TAPE TECHNOLOGY AND CHALLENGES TO OVERCOME

- PRÉSENTATION DE LA NOUVELLE GÉNÉRATION DE BANDES 3592

AcceStor: RAID is dead

- OPTIMISATION DU STOCKAGE FLASH

Spectra: Future-Looking Data Storage for Peak Performance in HPC Environments

- DO LESS OR DO IT FASTER (HOW?)

Retour HEPiX

HEPiX Activities and Future meetings

AAI - BoF Session

“Scheduled spontaneously” after the CERN AA talk

- More than 30 participants
- Multiple sites revamping their authentication system
- Keycloak (identity and access mgmt.) being investigated in several institutes
- Legacy applications require "token translation"
- Lots of needs - support both HEP and Photoscience communities
- Granularity required but not always easy to define given the diverse users

HEPiX AAI Working group created

- Kick-off meeting on 9th of November <https://indico.cern.ch/event/769924/>

HEPiX Activities: Working Groups

- Active
 - Migration to IPv6
 - CPU benchmarking
 - Batch system monitoring
 - Network function virtualisation
 - Tracking of technology and market evolutions (tech watch)
 - SIG on AAI
- Completed (selection)
 - Data preservation
 - Configuration management
 - Storage
 - Virtualisation
 - Unix environment scripts
 - AFS
 - Security

See you in San Diego / CA (USA)!



25 – 29 March 2019



[HEPiX d'automne](#) : Nikhef / Amsterdam 14 - 18 octobre

- HSF/WLCG/OSG workshop
 - 18-22 mars (semaine précédent HEPiX) @Jefferson Laboratory
- HTCondor week
 - 24-27 septembre (TBC) à Ispra (Italie) au Joint Research Center of the European Commission