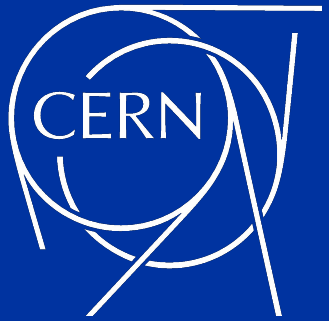


CERN Cloud Infrastructure status update

HEPiX Spring 2021

Patrycja Ewa Gorniak (CERN Cloud Infrastructure Team)
patrycja.ewa.gorniak@cern.ch

18.03.2021


Cloud resources
Used
357.7 K cores
Available
321.5 K cores
Used
1016.2 TiB RAM
Available
1.1 PiB RAM
Used
9.6 PiB disk
Available
16.3 PiB disk
Openstack services stats
Users
3446
Projects
4608
VMs
34919
Magnum clusters
643
Hypervisors
7960
Images
4092
Baremetal Nodes
9223
Volumes
6983
Volume size
3.11 PiB
Fileshares
2463
Fileshares size
665 TiB
Resource overview by time
VMs created/deleted (last 7days)
VMs created
VMs deleted
Shared cells availability (last 30days)
ratio
Total VMs
Active VMs
Average VM boot time
p50 without DNS Avg: 1.3 min
p99 without DNS Avg: 5.9 min
p50 with DNS Avg: 9.1 min
p99 with DNS Avg: 16.2 min
VM changes
Difference
Hypervisors
Total HVs
Projects and users
Projects
Users
Magnum clusters
dcos Current: 7
kubernetes Current: 592
mesos Current: 2
swarm-mode Current: 43
Cores
Total
Used
Memory
Total
Used
Disk
Total
Used
Baremetal Nodes
Total

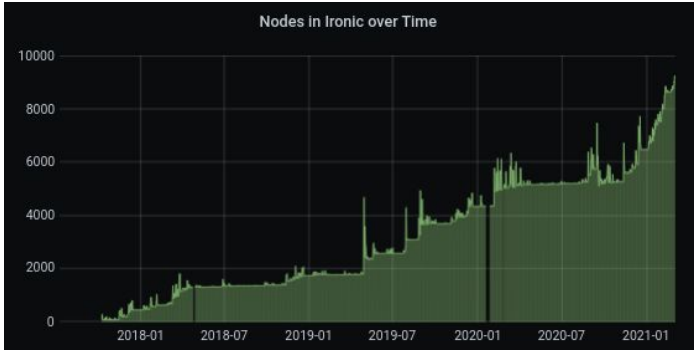
Since last HEPiX...

- ➔ OpenStack upgrades
- ➔ Integration with CERN identity management
- ➔ Ironic enrollment of production machines
- ➔ GPUs as managed resources in OpenStack (quotas, vGPUs)
- ➔ More SDN (tenant networks, virtual routing)
- ➔ CentOS 8
- ➔ Continue hardware replacement
- ➔ Block storage availability zones

Bare Metal Provisioning & Life-cycle management

- **Based on OpenStack Ironic:** https://ironicbaremetal.org/

  - **Same ‘creation’ API as virtual machines**
  - **Managing 9’000 physical nodes in CERN IT**
  - **Aiming at the full cycle:**
    auto-registration, validation, stress testing, benchmarking, provisioning, repairs, retirement
  - **Integration with Inventory (OpenDCIM)**



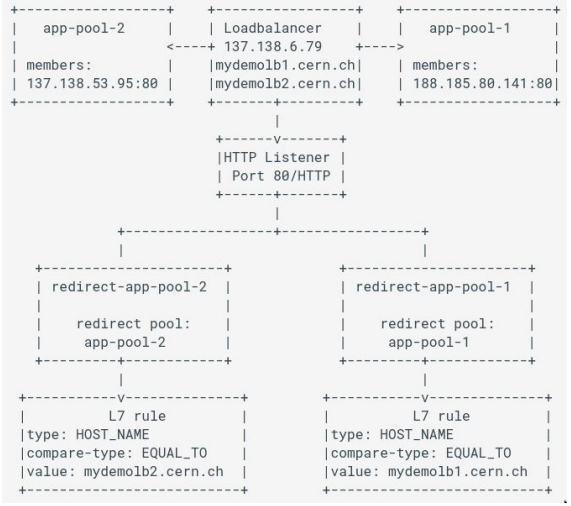
- **Transparent adoption of in-production nodes ongoing**
  - https://techblog.web.cern.ch/techblog

- **Shifting to physical batch using k8s or Terraform (CHEP talk)**
  https://indico.cern.ch/event/773049/contributions/3473820/attachments/1937858/3212037/Managing_the_CERN_Batch_System_with_Kubernetes.pdf

- **Working on Redfish for better console access**

Software Defined Networking (SDN) and LBaaS

- **IP Based Load Balancing as a Service (LBaaS) in production for almost a year**
  - https://clouddocs.web.cern.ch/networking/load_balancing.html
- **Over 200 LB instances, some backing critical services**
  - Windows Terminal Services for Cryo, Vacuum
- **Recently added support for L7 policies** →
  - And ability to apply ACLs to LB instances
  - https://indico.cern.ch/event/976468/
- **New region with advanced SDN functionality**
  - Private networks, Floating IPs, ...



GPUs

- **Centrally managed GPU offering for over a year - T4s and V100(S)**

  https://clouddocs.web.cern.ch/gpu/README.html

- **Integration with multiple CERN IT systems: VMs, K8s, Batch, Notebooks, ...**
- **Recently added support for Virtual GPUs**
  - Relying on T4s, up to 4 virtual slots per physical GPU
  - Based on time sharing, no physical partitioning like MIG on A100s
  - Requires a license setup on the nodes, we provide recipes to users

Kubernetes

- **Continued growth: over 600 clusters, 3000 nodes**
- **Functionality since last update**
  - Support for Kubernetes 1.19 and 1.20
  - OIDC integration: auth/authz against clusters using CERN SSO

    Support for linked CERN groups/roles to RBAC rules
- **Automated setup of Virtual GPUs, requiring custom change to Nvidia operator**
- **CERN container webinars:**

  https://www.youtube.com/channel/UCvFftdXaeIJLua2wVoKWz6A

Improving Kubernetes storage

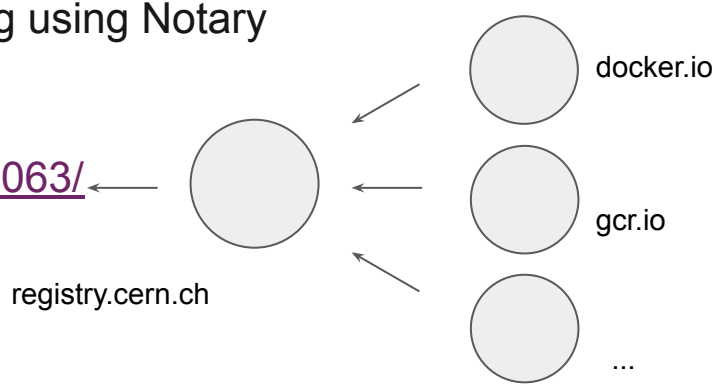
- **Kubernetes Storage at CERN - current status**
- **Issues and Next Steps**
  - Snapshots in CephFS
  - OAuth2 Token Based Access
  - Object Storage APIs

https://techblog.web.cern.ch/techblog/post/kubernetes-storage-next-steps/

Container Registry

- **Relying on a Harbor deployment for over 2 years to host Helm charts**
- **Recently expanding its usage to include other OCI artifacts**
  - Container Images, ML Models, …
- **Key Functionality (over the existing GitLab Registry)**
  - Vulnerability Scanning (CVE), Image Signing using Notary
  - Proxy Caches, Registry Replication

https://indico.cern.ch/event/995485/contributions/4258063/



OpenStack Placement Extraction

- **Placement API service introduced in the “Newton” release**
  - Required in any OpenStack Nova deployment since then
- **Placement tracks the resource provider inventories and usages (for example: CPU, Memory, Disk, ...)**
- **Since Stein release, Placement is an independent project**
- **At CERN, Placement was extracted from the 3 main regions**
  - Small downtime required to copy the DB
- **Placement was then upgraded to the “Train” release and the service moved to CentOS 8**

CentOS 8

- **Moved APIs to CentOS 8 to install newest releases of OpenStack**
- **Hypervisors still in CentOS 7, paused until further decisions are made**
- **Details:** https://indico.cern.ch/event/995485/contributions/4256466/

Operations

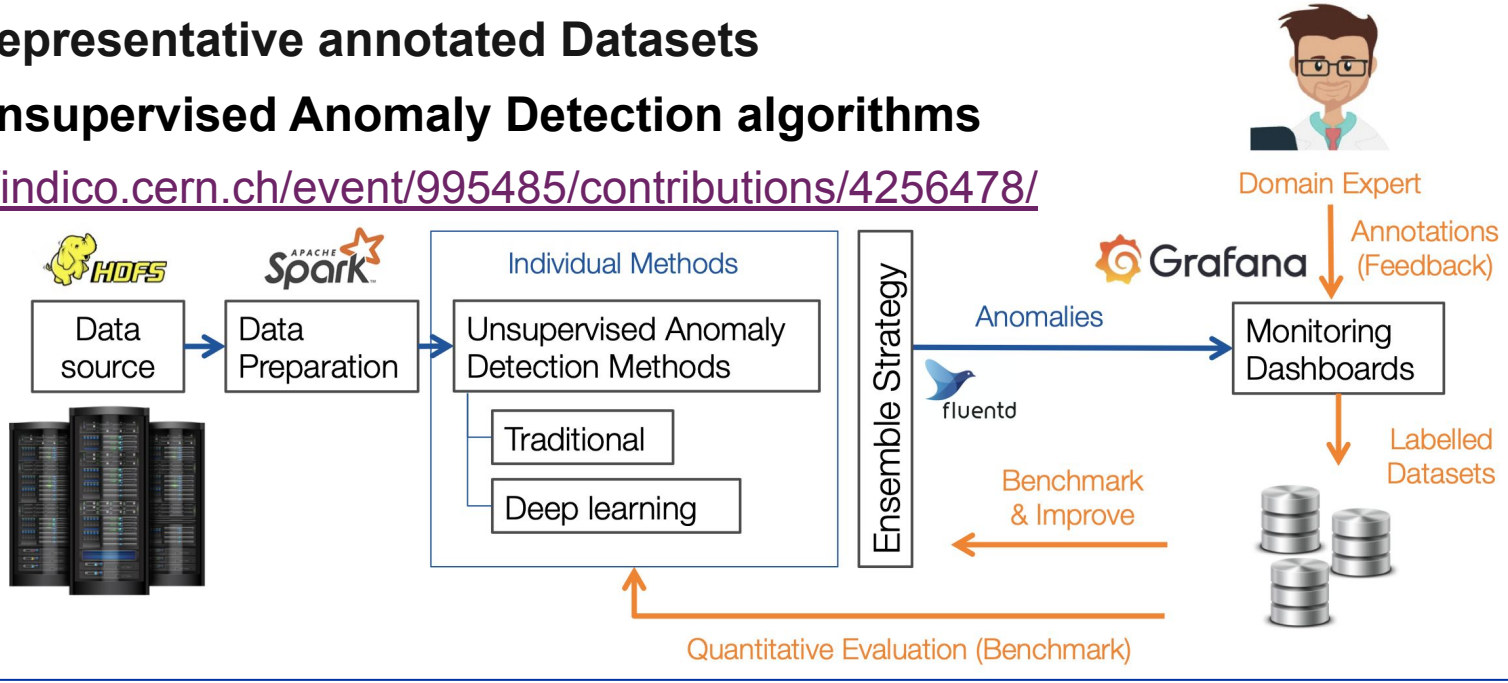
- ➔ **Few incidents affecting Cloud Service noted during last 6 months**
- ➔ **Live migration, improvement of the repair team ops**
- ➔ **High-Availability Database (HA DB)**

  https://indico.cern.ch/event/898285/contributions/4034158/

Anomaly Detection

- ➔ **Anomaly Detection Pipeline with Feedback**
- ➔ **Representative annotated Datasets**
- ➔ **Unsupervised Anomaly Detection algorithms**

https://indico.cern.ch/event/995485/contributions/4256478/



Future

- **GPUs as managed resources in OpenStack (quotas)**
- **More SDN (tenant networks, virtual routing)**
- **Ironic auto-registration of new physical servers**
- **Block storage availability zones**
- **In the process to containerize some services**
  dedicated webinar: https://www.youtube.com/watch?v=ViROtsY2hXU
- **Anomaly detection investigations** Anomaly Detection in the CERN Cloud Infrastructure

https://indico.cern.ch/event/995485/contributions/4256478/

Useful links

- **Cloud docs**
  https://clouddocs.web.cern.ch
- **Tech blog**
  https://techblog.web.cern.ch/techblog/
- **Container Webinars**
  https://www.youtube.com/channel/UCvFftdXaeIJLua2wVoKWz6A
- **Grafana - OpenStack Overview page**
  https://monit-grafana.cern.ch/d/000000024/cern-openstack-overview?orgId=3&refresh=15m

Questions?


jan.van.eldik@cern.ch
Jose Castro Leon
Spyridon Trigazis
Matteo Paltenghi
dejangolubovic
arne.wiebalck@cern.ch
domenico.giordano@cern.ch
Belmiro Moreira
jay
patrycja.ewa.gorniak@cern.ch
hamza.zafar@cern.ch
manos.bagakis
António Paulo
thomas.george.hartland@cern.ch
dioguerra
Dimitris
Ricardo Rocha

home.cern