The WLCG Journey at CSCS: from Piz Daint to Alps

Riccardo Di Maria
Swiss National Supercomputing Centre
ETH Zurich
Lugano, Switzerland
riccardo.dimaria@cscs.ch

Dino Conciatore
Swiss National Supercomputing Centre
ETH Zurich
Lugano, Switzerland
dino.conciatore@cscs.ch

Miguel Gila
Swiss National Supercomputing Centre
ETH Zurich
Lugano, Switzerland
miguel.gila@cscs.ch

Giuseppe Lo Re
Swiss National Supercomputing Centre
ETH Zurich
Lugano, Switzerland
giuseppe.lore@cscs.ch

Elia Oggian
Swiss National Supercomputing Centre
ETH Zurich
Lugano, Switzerland
elia.oggian@cscs.ch

Dario Petrusic
Swiss National Supercomputing Centre
ETH Zurich
Lugano, Switzerland
dario.petrusic@cscs.ch

***Abstract*—The Swiss National Supercomputing Centre (CSCS), in close collaboration with the Swiss Institute for Particle Physics, provides the Worldwide LHC Computing Grid (WLCG) project with cutting-edge HPC and HTC resources. These are reachable through a number of Computing Elements that, along with a Storage Element, characterise CSCS as a Tier-2 Grid site. The current flagship system, an HPE Cray XC named Piz Daint, has been the platform where all the computing requirements for the Tier-2 have been met for the last 6 years. With the commissioning of the future flagship infrastructure, an HPE Cray EX referred to as Alps, CSCS is gradually moving the computational resources to the new environment. The Centre has been investing heavily in the concept of Infrastructure as Code and it is embracing the multi-tenancy paradigm for its infrastructure. As a result, the project leverages modern approaches and technologies borrowed from the cloud to perform a complete re-design of the service. The goal of this contribution is to describe the journey, design choices, and challenges encountered along the way to implement the new WLCG platform, which is also being profited from by other projects such as the Cherenkov Telescope Array.**

I. Introduction

The Swiss National Supercomputing Centre (CSCS) has been successfully operating the HPE Cray XC system named Piz Daint as its flagship supercomputer for various customers and types of workloads. A new era has now begun with Alps, an HPE Cray EX system, for which the multi-tenancy paradigm has been comprehensively endorsed by the Centre through an Infrastructure as Code (IaC) approach. As a consequence, challenges are presented not only by future clients and partnerships, but also by existing use cases to be migrated and adapted to the new architecture. However, re-designing services by leveraging modern cloud technologies and approaches allows CSCS to achieve the goal, improving flexibility throughout the quest while moving the computational resources from Piz Daint to the new environment.

The Worldwide LHC Computing Grid (WLCG) [1] is a global collaboration of computer centres that provides data storage and computational resources for the experiments at the European Organization for Nuclear Research (CERN). In collaboration with the Swiss Institute for Particle Physics (CHiPP) and as one of the Tier-2 Grid sites, CSCS provides WLCG with advanced High Throughput Computing (HTC) and High Performance Computing (HPC) resources. The HTC workflows from WLCG have unique characteristics and it is widely acknowledged that they are not ideal for traditional HPC resources. Nonetheless, attaining a suitable equilibrium in this scenario represents also a valuable benchmark for the Alps system, which can demonstrate its elasticity and showcase its potential. Furthermore, this use case represents an example of off-loading of micro-services from HPC resources since the front-end services are located outside the Alps system, which is consequently entirely exploited for computing.

The infrastructure architecture is presented in the following section, focusing on the modern cloud approaches and technologies used to enable efficient WLCG data processing. Subsequently, the design of the new WLCG platform is described, along with its usage and leveraging of the multi-tenancy paradigm to meet the unique needs and demands of the Large Hadron Collider (LHC) [2] experiments.

II. The Infrastructure

A. Rancher

The Kubernetes (K8s) management tool exploited by CSCS to deploy and run containerised applications at scale is Rancher by SUSE [3]. It provides the Centre with a centralised platform to manage multiple K8s clusters, simplifies the deployment process, and streamlines the administration of applications and micro-services. Rancher offers a wide range of features including multi-tenancy, role-based access control, monitoring, and security to ensure that K8s-based applications are secure, reliable, and easily scalable. The K8s clusters at CSCS are managed across different environments, which in this context are "Alps", "bare-metal", and "virtual" infrastructure through Harvester [4]. The former is characterised by a K8s cluster deployed on cutting-edge HPC resources, namely Alps compute nodes. The bare-metal environment is an ideal choice for resource-intensive applications that require maximum performance without misspending Alps compute.

The latter is designed to ease the deployment of multiple clusters in a virtual environment through Harvester, and is the focus of this article as an example of micro-services off-loading from HPC resources. The environments and their main characteristics with respect to CSCS infrastructure are summarised in Fig. 1.

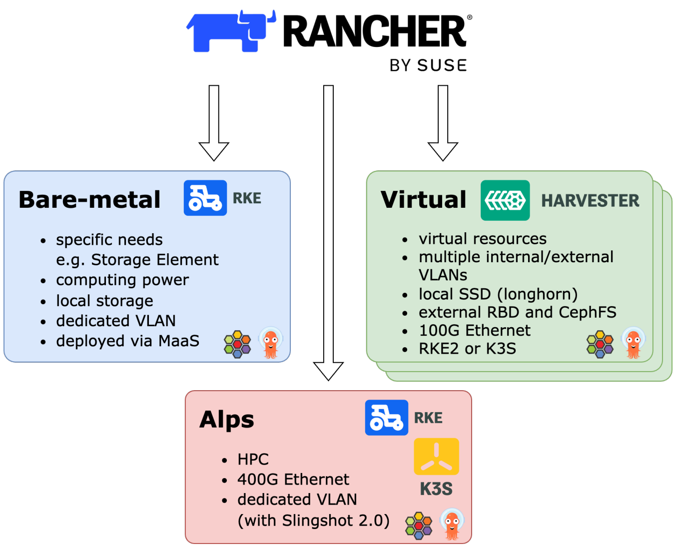

Fig. 1. Environments hosting Rancher-managed K8s clusters.

Rancher-managed clusters are created using Cilium [5] for the container network interface, a flexible networking solution that provides enhanced security, observability, and performance for micro-services-based applications. It leverages extended Berkeley Packet Filter (eBPF) [6] technology to offer transparent visibility and control of network traffic between services, enabling fine-grained policy enforcement and network segmentation. Cilium also supports multiple deployment models, including bare-metal, virtual machines, and cloud environments, providing greater flexibility in deploying and scaling applications.

Rancher is installed via RKE2 [3] through Ansible on 3 dedicated servers in High Availability (HA) mode, each being Intel dual-socket 12-core with 128 GB RAM and provisioned with Metal-as-a-Service (MaaS) by Canonical [7].

B. Harvester

Harvester by SUSE is a hyper-converged infrastructure and open-source virtualisation solution that simplifies the management of K8s clusters, whose master and worker nodes are virtual machines (VMs). The integration with Rancher provides CSCS with a unified management platform, allowing administrators and users to easily create, manage, and orchestrate their clusters.

Harvester is installed via iPXE boot through the network on 8 dedicated servers in HA mode, referred to as "Diablons" machines, each being AMD EPYC 64-core with 512 GB RAM and 8 TB NVMe local storage. The installation procedure is detailed in Fig. 2: a DHCP request is performed by Diablons nodes and an offer is made by the management plane; the iPXE boot file is fetched along with the Harvester ISO and configuration; the installation is initiated on each server and finalised with a reboot.

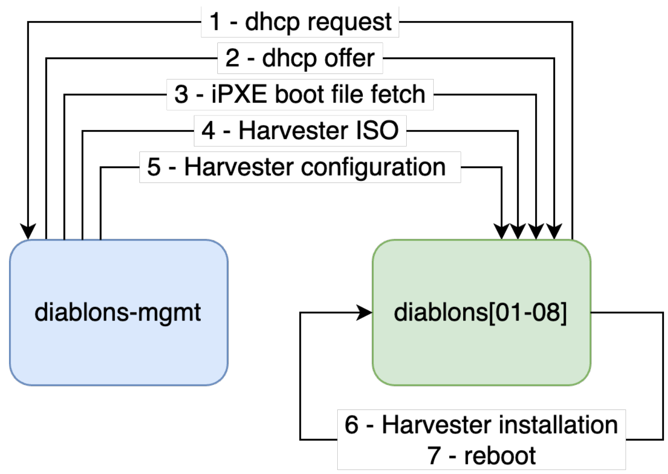

Fig. 2. Harvester installation procedure on Diablons servers.

The Harvester management network on Diablons is in HA mode at 25 Gb/s and it is isolated through a tagged virtual local area network (VLAN) from the network serving the K8s clusters. The Harvester VLAN network is at 100 Gb/s and similarly in HA mode, using LACP (in IEEE 802.3ad). Each Rancher-managed cluster is created on a dedicated VLAN whose characteristics depend on the requirements and needs of the client operating the hosted applications and services. Fig. 3 pictures the Harvester networks with respect to the client clusters created and managed by Rancher. As the nodes of a K8s cluster are VMs on the Harvester layer, each physical server can host multiple K8s nodes from different client clusters, which are therefore spread across several Diablons machines. Moreover, the ease of expanding the Diablons physical cluster by adding "new-host[01-02]" is highlighted by Fig. 3, for which the Harvester installation procedure stands as stated above.

C. ArgoCD

The automation of continuous integration and continuous deployment (CI/CD) procedures for services and applications is crucial during operations at scale, and CSCS profits from open-source tools, e.g. ArgoCD [8], to define and manage the application deployment life-cycle. ArgoCD provides a web-based user interface and a command-line interface for administrating and monitoring the application deployment process on Rancher-managed K8s clusters, adopting a GitOps approach. Git repositories are used as the source of truth, and the tool can automatically detect changes to the application configuration or source code and deploy the changes to the target environment, being Alps, bare-metal, or virtual infrastructure through Harvester.

As ArgoCD eases and streamlines the process of deploying applications, CSCS delivers K8s clusters to clients with key

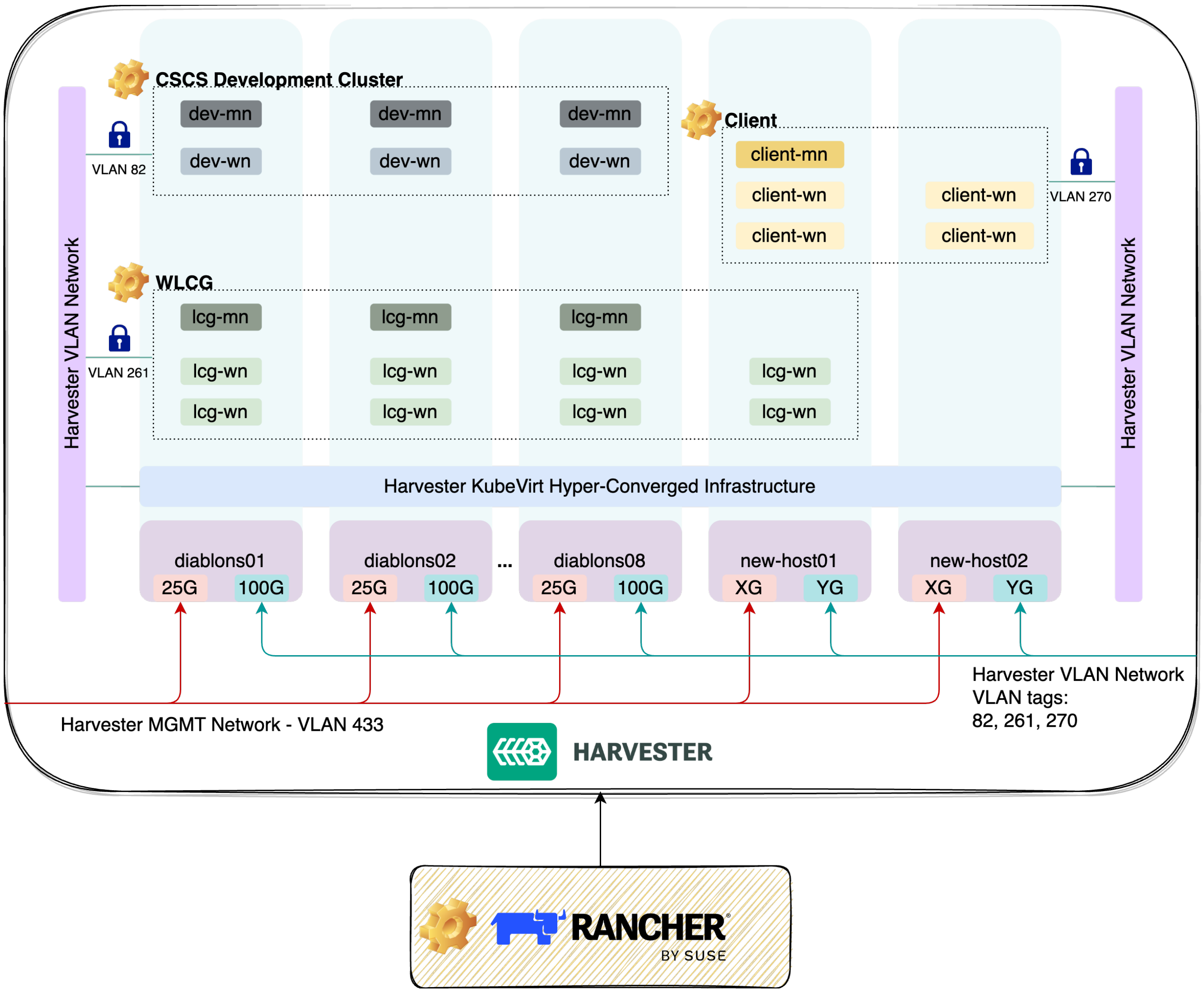

Fig. 3. Rancher-managed K8s clusters deployed on Harvester virtual environment.

features enabled and several services installed and configured, as follows:

- Metricbeat and Filebeat [9], lightweight data shippers, are used for collecting and sending data to centralised CSCS Elasticsearch [9], such as metrics, logs, and network data from the cluster;
- backup of etcd, a distributed key-value store designed to contain K8s cluster state data, is performed to an S3 bucket for disaster recovery and data protection;
- Velero [10], a tool for backing up and restoring K8s resources and persistent volumes, allows users to create backup schedules, manage backups, and restore resources to a previous state;
- ingress NGINX, a controller handling ingress traffic, enables users to define and manage HTTP(S) routes to the cluster, allowing for load balancing and routing of traffic to the appropriate services;
- CephFS and Rados Block Device (RBD) [11] container storage interface drivers enable clusters to access external storage systems, being used as persistent volumes;
- cert-manager [12] controller is used for managing SSL/TLS certificates, in order for users to automatically obtain, renew, and revoke certificates from certificate authorities;
- MetalLB [13] exposes services externally to the cluster by configuring a pool of IP addresses exploited to load balance the traffic;
- external-DNS automatically creates and manages DNS records for K8s services and ingresses.

Moreover, clusters are configured with external-secrets and integrated with the central HashiCorp Vault [14] instance improving their management by CSCS. The former is a K8s controller used for managing secrets stored outside of the cluster, Vault in this context.

D. Alps

The commissioning of the future flagship machine, an HPE Cray EX named Alps, is characterised by a phased installation and expansion of diverse types of CPU and GPU hardware. In this context, the focus is on the computing resources reserved to WLCG: 100 AMD EPYC "Rome" dual-socket 64-core nodes with 256 GB RAM. Additionally, 4 nodes are utilised for testing and development purposes.

The Alps infrastructure is logically divided into a management plane and a managed plane. The HPE Cray core components are containerised and deployed on the management plane through the Cray System Management software suite which integrates Kubernetes. The concept of virtual cluster (vCluster) results in the joint of the dedicated part of the management plane that administers the compute resources via a specific K8s-namespace and the related portion of the managed plane.

A software-defined infrastructure as foundation of the IaC approach, jointly with the features offered by the Cray Programming Environment software suite, allows the Centre to implement multiple instances of the job scheduling system. An instance of Slurm is deployed on the dedicated K8s-namespace of the management plane per vCluster, and it is responsible for handling job scheduling and resource allocation on the compute nodes in the managed plane, enabling users to seamlessly manage traditional HPC workloads alongside containerised workflows.

The WLCG production vCluster is referred to as "Mont Fort", whereas the test instance is named "Mont Gele", and both vClusters have no login nodes as in standard HPC scenarios but solely compute nodes as common practice for HTC environments. Due to the requirement of 2 GB RAM per core for WLCG jobs, hyper-threading is disabled by allocating 128 CPUs per node, and Slurm reservations are utilised to manage the specific loads from the different LHC experiments. Although several Configuration Framework Service layers are shared with HPC-standard vClusters within Alps, the last stage is tailored to WLCG needs and requirements. Furthermore, a dedicated VLAN is planned to host the Mont Fort vCluster after upgrading to Slingshot 2.0 and testing the feasibility. As a consequence, a common network is shared between the vCluster and the homonymous Rancher-managed K8s cluster hosting the front-end services which is described in Section III-A.

Finally, it should be emphasised that a WLCG workload consists typically of single-core or multi-core jobs but excludes MPI tasks. As a consequence, the Slurm instances dedicated to Mont Fort and Mont Gele have been configured and tuned to support this scenario.

E. Storage

The storage backend solution currently implemented for WLCG is established on Ceph [11]. This is a software-defined storage system and it runs on CSCS commodity hardware. It provides the foundation to build a "data lake", where multiple methods of data access are used depending on the data workflow. Ceph is based on Rados, an object store layer which handles the storage resources, and it is characterised by 3 data management applications providing different interfaces to user data, namely CephFS, RBD, and RadosGateway.

Network traffic between Alps and the external Ceph storage backend is via Ethernet. CephFS is used to implement a shared area, */scratch/shared*, for both compute nodes and front-end services, whereas NVMe block devices are provided through RBD for the local area of each compute node. As a result, the compute node performs the bulk of IO operations on an independent */scratch/local* file system, allowing the metadata performance scale up with the number of compute nodes and hence avoiding bottlenecks typical with shared HPC file systems. Moreover, potential issues caused by problematic jobs are confined to the nodes computing those specific tasks. As of today, the Ceph cluster devoted to WLCG is composed of 51 storage servers delivering 530 TiB and 22 PiB of usable NVMe and HDD capacity, respectively.

F. Monitoring

The complexity and scale of the systems described in this contribution affect the monitoring and accounting data which are to be collected and organised. The management of this vast amount of data is a challenge that CSCS mitigated by introducing a K8s cluster dedicated to dynamically deploy data collection and analysis stacks comprising Elastic Stack [9], Kafka [15], and Grafana [16]. As these are exploited for both internal and external customers' use cases, the solution proved to be crucial since it provides CSCS with correlation of events and meaningful insights from event-related data. Bridging the gap between the computation workload and resources status enables failure diagnosis, telemetry and effective collection of accounting data. Additionally, the integrated environment from data collection to visualisation enables internal and external users to produce their own dashboards and monitoring displays, tailored to their data analysis needs.

III. CSCS as WLCG Tier-2 Site

A Tier-2 site in WLCG is a computing centre that provides computing and storage resources to support data analysis for scientific research on physics observed by the LHC experiments. CSCS, in close collaboration with CHiPP, supports 3 experiments, namely ATLAS [17], CMS [18], and LHCb [19], and is responsible for processing and analysing data, along with Monte Carlo simulations, and for providing the results to Tier-1 sites. A subset of the data are stored at CSCS site and are easily accessible to researchers exploiting the Grid, providing a bridge between the LHC experiments and the local scientists. Furthermore, data are shared through the WLCG network infrastructure with other Tier-1 and Tier-2 sites as part of the experiments' Collaboration.

A WLCG Tier-2 site is characterised by 2 fundamental components: a Computing Element (CE) and a Storage Element (SE). The following sections aim at describing these elements and their implementation, being the interface between the Grid and HPC/HTC resources at CSCS.

A. Compute Element

The CE at CSCS is based on NorduGrid Advanced Resource Connector (ARC) [20] middleware, a front-end for WLCG workloads that acts as an edge service to the Slurm-based batch system, enabling access to Alps computing resources via a secure common Grid layer. The authorisation relies on X.509 proxy certificates and JWT tokens within the concept of Virtual Organisations (VOs).

The ARC-CE is deployed on a Rancher-managed K8s cluster on Diablons, named "Mont Fort", which creates a unique logical entity with the homonymous vCluster on Alps. A HA implementation of ARC-CE serves the LHC experiments, resulting in 2 instances dedicated to the ATLAS experiment and 2 for the CMS and LHCb experiments. However, there is a single common Slurm queue on the vCluster for the entire WLCG workload. A dedicated VLAN is exploited to ensure isolation with other CSCS services and network fine-tuning is performed to allow connections only with relevant applications. Mont Fort hosts additional services critical for WLCG workflows, such as an identity and access management component based on Go-lang LDAP Authentication (GLAuth) for user and account mapping, and CernVM File System (CVMFS) [21] for job-context provisioning. CVMFS is in fact exploited to fetch images, lightweight in comparison with HPC-standard, executed in nested containers on Alps compute nodes. A REST daemon is also deployed on the cluster, pivotal for the communication with the Slurm control daemon running on the management plane of Alps. An additional instance of ARC, namely "Mont Gele" with its vCluster counterpart, is deployed for testing purposes with respect to features integration, as well as improvement and re-design of the CE.

Physicists submit computational tasks to CSCS resources through the ARC-CE service using the extended Resource Specification Language, and input/output data movement is handled on the server side decreasing the impact on the compute resources. The management information of the job, considered in this context as experiment agnostic, is stored in an RBD, whereas the session of each job, together with its cache, relies on a CephFS volume shared with each compute node in Alps. The job obtains its context by accessing the CSCS CVMFS instance, and moves its session from the shared CephFS volume to the local scratch RBD, exploited during the execution of the task. The output is eventually moved back to the shared volume and provided to ARC-CE to be sent to the user. The LDAP mechanism provides consistency for the user mapping and accounting among the logical components of Mont Fort, ensuring the reinforcement of Slurm reservations for each experiment based on the pledge computing resources agreed upon with the VOs. The WLCG job submission schema on Mont Fort is illustrated in Fig. 4.

B. Storage Element

The SE component of CSCS Tier-2 Grid site is based on dCache [22] distributed storage system on Ceph. This middleware encloses all resources in a single namespace and hence on one file system, and it relies on components,

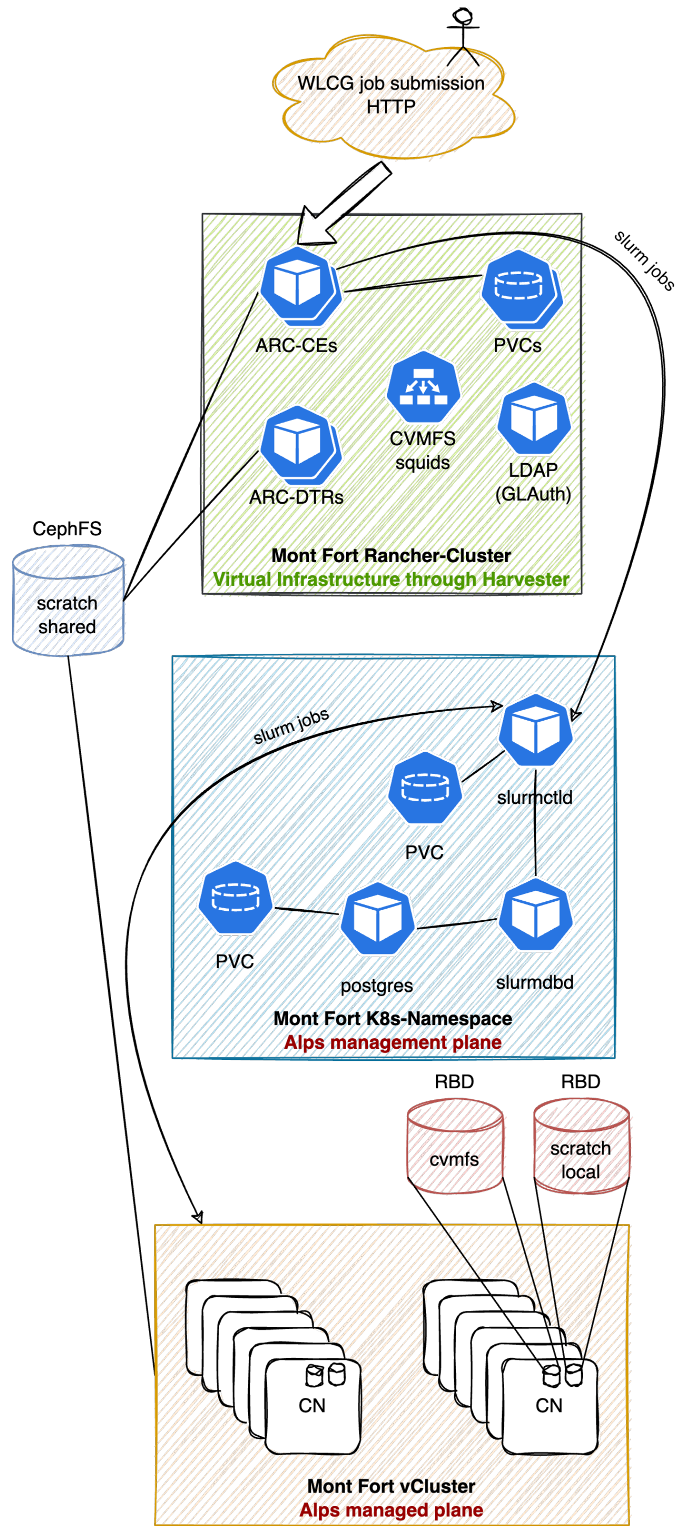

Fig. 4. WLCG job submission schema on Mont Fort.

such as PNFS module that guarantees the functionality of the underlying file system, or the *pools* that encapsulate the available physical disk space. Subsidiary tasks of dCache are carried out by *cells*, of which gPlazma, the service responsible for user authorisation, is an example.

The entire deployment is based on K8s on a dedicated VLAN and managed by Rancher via a bare-metal environment. As a matter of fact, the SE use case is better suited when the K8s cluster runs on physical commodity hardware tuned to accommodate the intensive load arriving from clients as WLCG. Connectivity and transfers between the SE and Alps, or other sites' SEs, is eased by public IPs offered to all CSCS services. The access to the dCache service is direct from users or from the workload management system of an experiment, as well as indirect through the jobs running on Alps compute nodes via ARC-CE. At a Tier-2 Grid site, researchers generally submit jobs to the CE that is close to the SE storing the data to be accessed. Among the various deployments on the Grid, a pioneering implementation of dCache on K8s, pictured on Fig. 5, aims at harmonising within the on-premise Cloud-like infrastructure that characterises the new Alps.

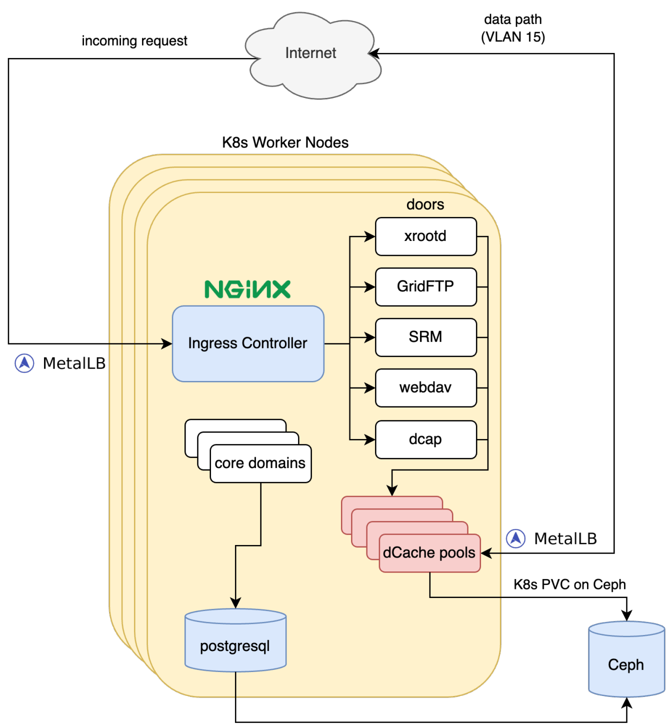

Fig. 5. Layout of the dCache K8s deployment for the SE.

In a complex architecture such as the one for the CE and SE, it is crucial to establish a CI/CD mechanism that orchestrates, monitors, and eases the development of the services. ArgoCD seamlessly integrates and operates in this multi-cluster scenario, in synergy with Rancher, Harvester, and the SE bare-metal environment. The management of the K8s resources via Git repositories further facilitates the integration of various instances of the same service per usage, such as production and development instances. Moreover, the deployments of multiple ARC-CE and dCache services tailored to the needs and requirements of the clients, i.e. WLCG and Cherenkov Telescope Array [23], are by far eased by relying on the aforementioned CI/CD tool.

IV. Conclusions

The CSCS Tier-2 Grid site has been playing a leading role in the commissioning of the future flagship Alps infrastructure as it can be considered a benchmarking exercise in the migration of computational resources from Piz Daint. It profits from the newly adopted multi-tenancy paradigm, a pillar of the IaC approach that characterises the new system. Furthermore, it has undergone several upgrades and enhancements to improve its performance, scalability, and reliability by leveraging modern cloud approaches and technologies. However, Mont Fort should be considered as a common HTC Grid site deployed on non-traditional HPC resources within the WLCG domain.

Relying on IaC-based implementation of Alps vClusters and of Rancher-managed K8s-clusters, such as dCache for the Storage Element and ARC for the CEs, allows CSCS to scale the infrastructure dynamically and according to the changing requirements of the customers.

The pioneering Kubernetes-isation of the CE and SE components, fundamental to the functioning of a Grid site, is considered a successful exemplification of the aforesaid concepts, and it improves the flexibility with respect to the evolving needs of the scientific community. Moreover, the off-loading of these front-end services from Alps decreases the impact of costly compute resources and optimises the support for scientific research performed on data produced at the LHC and collected by ATLAS, CMS, and LHCb experiments. Conclusively, Alps infrastructure represents the future for data analysis also in the physics domain, being CSCS Grid site a crucial element of the global scientific community's efforts in their quest for new discoveries.

Acknowledgment

Authors acknowledge support from the Swiss National Science Foundation and the Swiss Institute for Particle Physics.

References

[1] I. Bird, K. Bos, N. Brook, D. Duellmann, C. Eck, I. Fisk, D. Foster, B. Gibbard, C. Grandi, F. Grey, et al. LHC Computing Grid: Technical Design Report. CERN, Geneva, 2005. CERN-LHCC-2005-024, LCG-TDR-001.
[2] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock. LHC Design Report. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. CERN-2004-003-V-1.
[3] Rancher by SUSE, https://rancher.com/
[4] Harvester by SUSE, https://harvesterhci.io/
[5] Cilium, https://cilium.io/
[6] eBPF, https://ebpf.io/
[7] MaaS by Canonical, https://maas.io/
[8] ArgoCD, https://argo-cd.readthedocs.io/en/stable/
[9] Elastic, https://www.elastic.co/
[10] Velero, https://velero.io/
[11] Ceph, https://ceph.io/
[12] cert-manager, https://cert-manager.io/
[13] MetalLB, https://metallb.universe.tf/
[14] HashiCorp Vault, https://www.hashicorp.com/products/vault

[15] Apache Kafka, https://kafka.apache.org/
[16] Grafana, https://grafana.com/
[17] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. JINST, 3:S08003, 2008.
[18] CMS Collaboration. The CMS Experiment at the CERN LHC. JINST, 3:S08004, 2008.
[19] LHCb Collaboration. The LHCb Detector at the LHC. JINST, 3:S08005, 2008.
[20] NorduGrid ARC, https://www.nordugrid.org/arc/
[21] P. Buncic et al. CernVM – a virtual software appliance for LHC applications. 2010 J. Phys.: Conf. Ser. 219 042003.
[22] P. Fuhrmann, V. Gülzow. dCache, Storage System for the Future. Euro-Par 2006 Parallel Processing, pp 1106–1113.
[23] The CTA Consortium, M. Actis, G. Agnetta, et al. Design concepts for the Cherenkov Telescope Array CTA: an advanced facility for ground-based high-energy gamma-ray astronomy. Exp Astron 32, 193–316 (2011).