

PAUL SCHERRER INSTITUT



WIR SCHAFFEN WISSEN - HEUTE FÜR
MORGEN

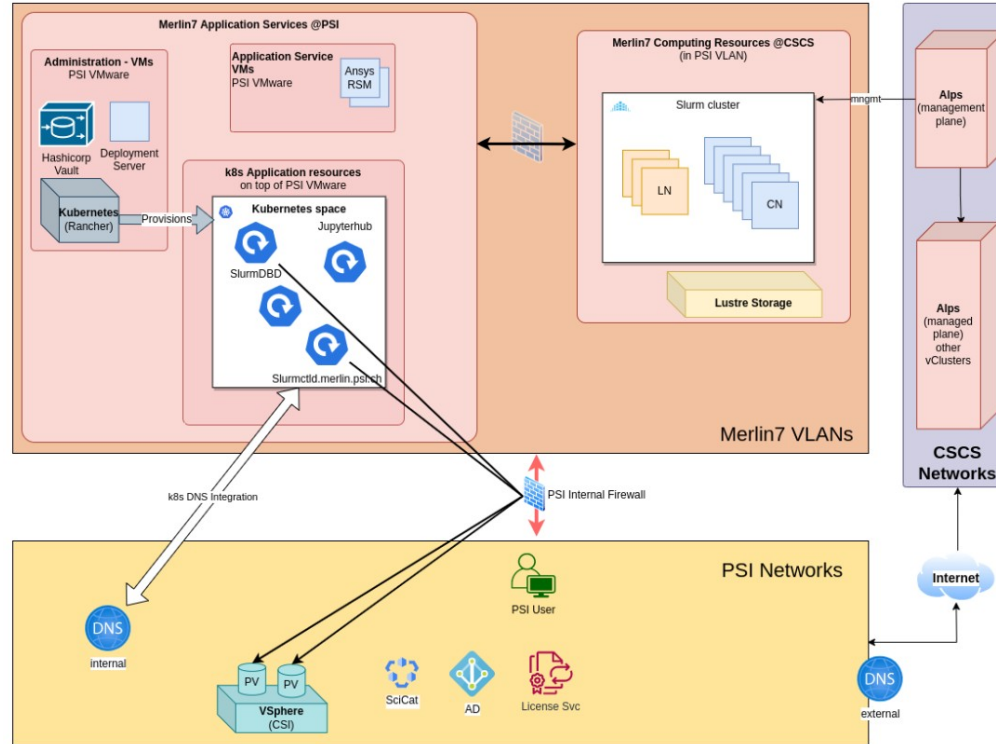
Marc Caubet Serrabou

HPCE AWI Group Update 2024-04-30

Infrastructure: Network, Security & VMs



Merlin7 Network Setup



Solution for running Cluster services at PSI: form a Merlin7 network zone composed of CSCS compute /storage resources and PSI service infrastructure

- minimal FW barriers between cluster services and compute nodes / storage
- must accept more restrictive FW barriers between PSI internal net and Merlin7 application services and compute nodes

21. 11. 2023 D. Feichtinger / M. Caubet

- **Access to PSI NetOps (<https://netops.psi.ch/>) to manage existing firewall rules, providing a good overview and easy management:**
 - Existing rules can be **edited**, will open a new internal ticket to the Network team.
 - Existing rules can be (immediately) **revoked**.
 - **New rules** have to be requested as usual (**PSI Service NOW** or **network@psi.ch**)

ID	Customer	Description	Source	Source User	Destination	Ports	Contacts	Last update	Expiry date	Action
ID3184078			172.30.100.0/24 - 172.30.102.178 (merlin7-cn0001) 172.30.136.0/22 + 172.30.102.0/24	any	any (to zone Internet)	udp/7000 udp/7002 udp/7003 udp/7005	achim.gseel@psi.ch merlin-admins@lists.psi.ch	27.03.2024 15:15	09.07.2024	[?] Request Update [X] Revoke
ID258696			172.30.100.0/24 - 172.30.102.178 (merlin7-cn0001) 172.30.136.0/22 + 172.30.102.0/24	any	129.129.160.132 (winlic03)	tcp/1055 tcp/27000	achim.gseel@psi.ch merlin-admins@lists.psi.ch	27.03.2024 10:44	04.07.2024	[?] Request Update [X] Revoke
ID2313568			129.129.0.0/16	any	10.129.160.96 (mcaubet-cluster) 10.129.160.98 (mcaubet01) 10.129.160.99 (mcaubet02)	tcp/4000	marc.caubet@psi.ch	26.04.2024 17:59	26.04.2025	[?] Request Update [X] Revoke

- **DNS management through QIP (<https://qip.psi.ch/qip>)**
 - HPCE can **manually manage DNS names** (ALPS, Virtual Machines) → *No RestFUL API enabled.*
 - **Missing Dynamic DNS**, mainly needed for K8s.

IP Address	Object Name	Domain	Object Class	Status	MAC Address	Description	Client Id	Network Status	Client Fingerprint
172.30.102.100	morgana-w-01	psi.ch	Virtualized Server	Static		RKE K8s cluster			
172.30.102.101	morgana-w-02	psi.ch	Virtualized Server	Static		RKE K8s cluster			
172.30.102.102	morgana-w-03	psi.ch	Virtualized Server	Static		RKE K8s cluster: DNS aliases set			
172.30.102.103	morgana-w-04	psi.ch	Virtualized Server	Static		RKE K8s cluster			
172.30.102.104	morgana-w-05	psi.ch	Virtualized Server	Static		RKE K8s cluster			
172.30.102.105	morgana-w-06	psi.ch	Virtualized Server	Static		RKE K8s cluster			
172.30.102.106	morgana-w-07	psi.ch	Virtualized Server	Static		RKE K8s cluster			
172.30.102.107				Unused					
172.30.102.108				Unused					
172.30.102.109				Unused					
172.30.102.110	morgana-rb-01	psi.ch	Load Balancer	Static		MeshLB for RKE: slurmcd			
172.30.102.111	morgana-rb-02	psi.ch	Load Balancer	Static		MeshLB for RKE: slurmcd			
172.30.102.112	morgana-rb-03	psi.ch	Load Balancer	Static		MeshLB for RKE: slurmbd			
172.30.102.113	morgana-rb-04	psi.ch	Load Balancer	Static		MeshLB for RKE: slurmbd			
172.30.102.114	morgana-rb-05	psi.ch	Load Balancer	Static		MeshLB for RKE: slurmcd TEST			

- **Need for Automatic Certificate Management Environment (ACME) SSL Certificates** to simplify certificates renewal
 - K8s (Harbor, Rancher), Hashicorp Vault, JupyterHUB, etc.

- **Dedicated cluster for TransAlps:**
 - **Admin rights** to manage *Virtual Machines*
 - VM creation, deletion and update
 - Snapshots management (default by any VM owner)
 - **Admin rights** for *Datastore*
 - VMs disks/files management
 - Uploading ISO files possible (i.e. OpenSuse Installation Image)
- **Virtual Machines:**
 - Most services installed with **Central Linux Infrastructure (RHEL-based + Puppet)**
 - *Optional* - Extra provisioning with Ansible
 - **Examples:** Morgana K8s, Hashicorp Vault, Syslog/Journald server, MariaDB
 - Services closer to the cluster based on **OpenSuse15.4 / 15.5 + Ansible**
 - **Generic templates** installed → first network setup required + ready to run Ansible
 - **Examples:** Slurm core services, Test (VM-based) compute nodes and login nodes
 - **Repositories centrally managed** by the Linux Team:
 - Repository snapshots/tags
 - Other repositories included (currently NoMachine, CUDA)
 - Current repositories:
 - <https://repos.psi.ch/opensuse-leap15.3/> (to decommission)
 - <https://repos.psi.ch/opensuse-leap15.4/>
 - <https://repos.psi.ch/opensuse-leap15.5/>
 - **Known Issues** → Can't reproduce 100% the nodes in ALPS, but to have similar configurations:
 - Different distros (OpenSuse vs SLES), different repositories.
 - **Examples:** mounting Lustre with OpenSource client, HPC packages from OpenSuse, no Slingshot, no GPU.

