



From zero to espSRC

Presented by Manuel Parra on behalf of the
Spanish SKA Regional Centre Team

Swiss SKA Days



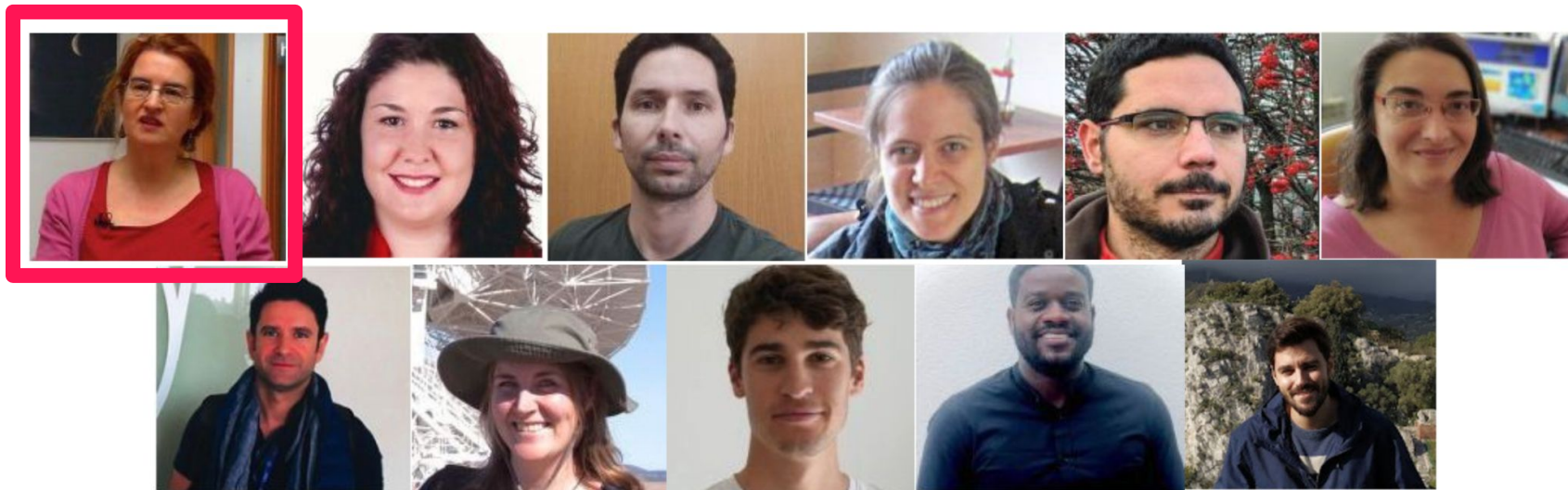
The Spanish SKA Regional Centre

- The Spanish SKA Regional Centre (espSRC) is hosted at the **Instituto de Astrofísica de Andalucía (IAA-CSIC)** – Spanish National Research Council.
- The espSRC team and its core computing infrastructure are co-located at the IAA-CSIC Computing Centre.
- Based in Granada, Spain.
- A multidisciplinary team of currently 11 members.





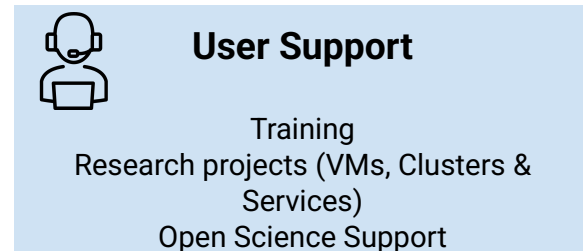
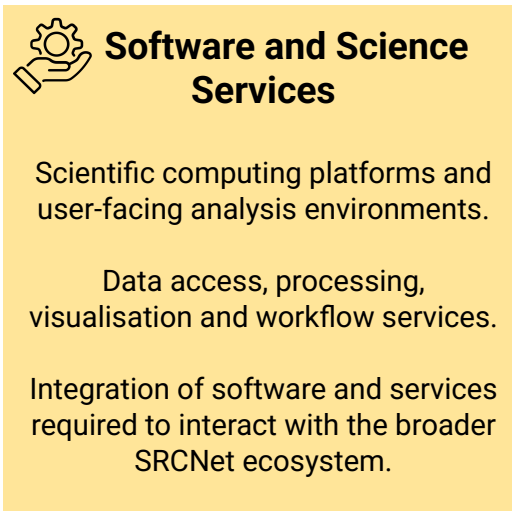
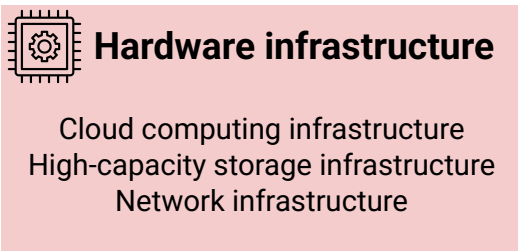
The espSRC combines expertise in **astronomy, scientific computing, distributed systems, software engineering, cloud infrastructure, data management and infrastructure operations.**



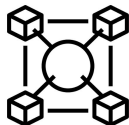


The espSRC Infrastructure: **Beyond a Computing Cluster**

espSRC is conceived not simply as a computing facility, but as an integrated environment combining **infrastructure**, **services**, **expertise** and **user support**.



Interoperability with **SRCNet nodes**





The espSRC Infrastructure: **Beyond a Computing Cluster**

A two-sided contribution

SRCNet:

The espSRC is a node within the SRCNet providing computing and storage resources, and science services to the SRCNet distributed infrastructure.

SKA preparatory science projects for the Spanish scientific community:

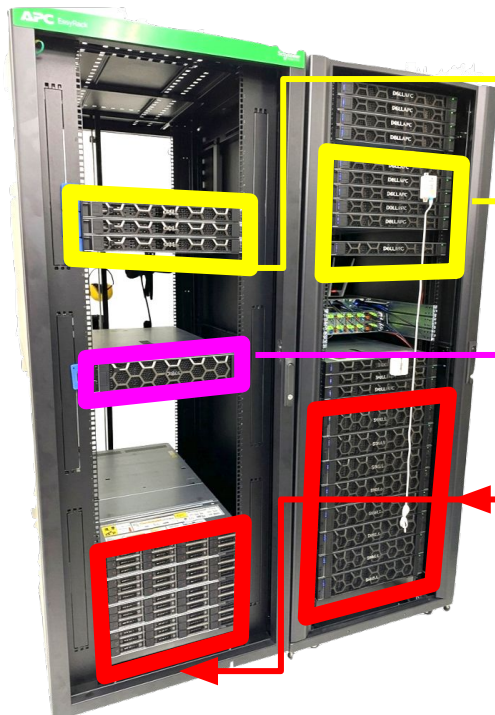
The espSRC provides dedicated resources to development, training, and research projects.

Currently more than 100 different projects registered and running.

→ Apertif, HCG-MeerKAT, GASKAP, LOFAR, Galactic Center



The espSRC Infrastructure: **Hardware**



- 10 nodes (9 CPU + 1 GPU)
 - Intel: 6 nodes with 40 cores (80 w/HT)
 - AMD: 3 nodes with 64 cores (x2)
 - Nvidia L40
- 720 CPU cores and 8.3 TBytes of RAM
 - Intel: 9.6 GB Ram per core / 25 GB per core
 - AMD: 13GB RAM per core
 - GPU: 8GB per cpu core
- ~ 3 PB of RAW data storage
- 100 Gbits Local Network connection (Computing and Storage)
- 10 Gbits Shared Internet connection



The espSRC Infrastructure: **Cloud Platform**

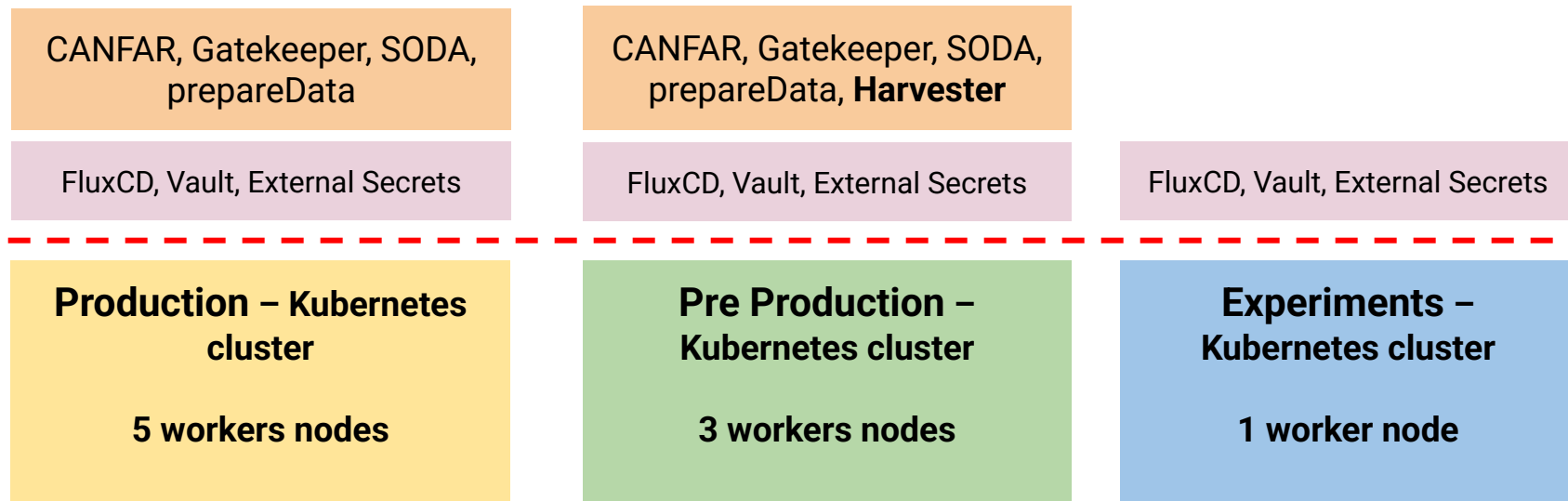


- Computing nodes, Storage and Networks are orchestrated by **OpenStack**
 - **VMs and Clusters**
 - Storage managed by **Ceph Reef 18.2.8**
 - **SSD and HDD pools**
 - **Erasure Code and Replica-2**
 - **Blocks storage and CephFS**
 - 3 Controller nodes
 - 3 Monitoring nodes
- Fault Tolerant



The espSRC: **Node Architecture within the SRCNet**

↪ Services orchestration on Kubernetes – Clusters on OpenStack VMs





The espSRC Node Architecture in SRCNet

↪ Services on VMs

IVOA Registry at espSRC (2nd
in sweSRC)

Networking:
perfSONAR VM

Image registry VM (Harbor)

Monitoring Services VM:
Prometheus, Zabbix and
Grafana

SLURM Cluster

espSRC Rucio Storage
Element – WebDav

HAProxy + SSL certs



The espSRC Node Architecture in SRCNet

↪ Services deployment & CI/CD procedure

CANFAR
Gatekeeper
Preparedata
Soda, vault,
external secrets

Services orchestration **on Kubernetes**

- Deployment **with FluxCD** and code stored in **SRCNet repository**:
<https://gitlab.com/ska-telescope/src/deployments/espsrc>
 - Production Cluster
 - Pre Production Cluster

Rucio RSE
WebDav, perfSonar,
Monitoring services,
Harbor, SLURM
cluster, IVOA
registry, HAProxy

Services **on VMs**

- Deployment based on Ansible scripts



The espSRC Node Architecture in SRCNet

↪ Services deployment & CI/CD procedure



Once an update/change is communicated through the service Slack channel or by Operations, this protocol is followed:

- 1) **Experiments / Test Cluster** → Test and validate deployment viability.
- 2) **Pre-Production** → Validate functional & operational readiness.
- 3) **Production** → Final release to live environment.

**Production – Kubernetes
cluster**

5 workers nodes

**Pre Production – Kubernetes
cluster**

3 workers nodes

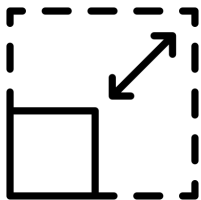
**Experiments – Kubernetes
cluster**

1 worker node



The espSRC Node Architecture in SRCNet

↪ Scaling infrastructure – Computing and Storage



Compute Scaling (Kubernetes)

- **VM Flavor Templates:** Kubernetes cluster nodes scale up/down via pre-configured Virtual Machine templates.
- **Ansible Automation:** Cluster scale-up & down are executed using automated Ansible playbook scripts.

Storage Scaling (CephFS)

- **Shared CephFS Mounts:** All cluster nodes share access to RSE storage (350 TB) as well as home & project volumes (200 TB).
- **Dynamic Quota Adjustment:** Scaling is instantaneous and performed in real-time by modifying shared resource quotas.
- Storage volume adjustments require no cluster restart and are immediately visible across all nodes.



The espSRC Node Architecture in SRCNet

↪ Monitoring our infrastructure



Core Observability Stack

- **Exporters:** Node Exporter & Blackbox Exporter gather system & synthetic endpoint metrics.
- **Prometheus & Grafana:** Centralized metrics collection, time-series storage, and visual analytics.
- **Full Coverage:** Monitors both Kubernetes cluster infrastructure and standalone VMs.



Alerting & Notifications



- **Zabbix Engine:** Automated fault detection and health trigger evaluation.
- **Slack Integration:** Real-time event notifications and problem alerts routed directly to team channels.

Dedicated Dashboards

- **Node Health:** Real-time tracking of CPU/RAM consumption, network throughput, and storage I/O.
- **Service Health:** Monitors service uptime/availability, response latency, and SSL cert validity.

Daily Health Inspection

- **Routine Audits:** Daily active review of all Grafana dashboards by the operations team.
- **E2E Test Suite:** Continuous verification through automated End-to-End service tests.



The espSRC Node Architecture in SRCNet

↔ Monitoring our infrastructure

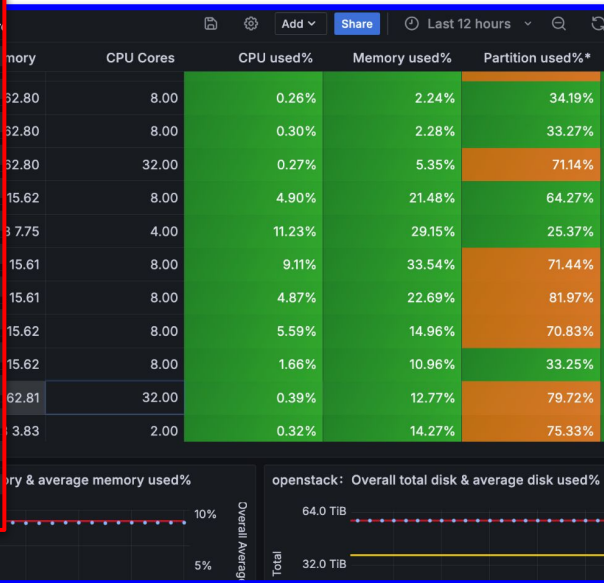


ZABBIX

HTTP Probe Overview

Service name	Instance	Status	Code	SSL	TLS Version	SSL Cert Expiry (days)	Pr
RSE	https://rucio.espsrc.iaa.csic.es	UP	200	OK	TLS 1.3	5	
Harbor	https://spsrc26.iaa.csic.es/	UP	200	OK	TLS 1.3	29	
Perfsonar	https://perfsonar.espsrc.iaa.csic.e...	UP	200	OK	TLS 1.3	31	
Cavern	https://canfar.espsrc.iaa.csic.es/c...	UP	200	OK	TLS 1.3	35	
Science Portal	https://canfar.espsrc.iaa.csic.es/s...	UP	200	OK	TLS 1.3	35	
Posix Mapper	https://canfar.espsrc.iaa.csic.es/p...	UP	200	OK	TLS 1.3	35	
Skaha	https://canfar.espsrc.iaa.csic.es/s...	UP	200	OK	TLS 1.3	35	
Monitor	https://monitor.espsrc.iaa.csic.es/	UP	200	OK	TLS 1.3	35	
Notebook	https://notebook.espsrc.iaa.csic.es/	UP	404	OK	TLS 1.3	35	
ECHO	https://echo.espsrc.iaa.csic.es/	UP	200	OK	TLS 1.3	39	
Gatekeeper	https://gatekeeper.espsrc.iaa.csic....	UP	200	OK	TLS 1.3	39	
PrepareData Gatekeeper	https://gatekeeper.espsrc.iaa.csic....	UP	401	OK	TLS 1.3	39	
SODA Gatekeeper	https://gatekeeper.espsrc.iaa.csic....	UP	401	OK	TLS 1.3	39	
Global Reg	https://spsrc27.iaa.csic.es/reg/	UP	200	OK	TLS 1.2	64	

ed Dashboards



problem alerts ro



The espSRC Node Architecture in SRCNet

↔ Improvements and changes for the next PI

Vault Replacement: Replace Vault with Kubernetes Sealed Secrets across nodes. Node changes and updates will no longer cause Vault sealing or service disruptions.



Central Log Collection: Implement an improved system for log aggregation and error capturing. Enhanced notification pipeline for faster SRCNet service issue resolution.

Single Pane: Design one customized master dashboard for at-a-glance node health. Consolidate metrics from multiple dispersed dashboards into one central view.

CI/CD Release Pipeline: Embed automated local End-to-End test suites into the release process. Verify service reliability locally before advancing through deployment stages.



The espSRC Node Architecture in SRCNet

↔ Improvements and changes for the next PI

Multi-Site Service Replication: We aim to leverage our CI/CD pipeline to deploy our service stack at a secondary location as a Proof of Concept (PoC), validating that the node architecture is site-agnostic, easily reproducible, and fully capable High Availability.





The espSRC Node Architecture in SRCNet

↪ Ideas and thoughts

Bare-Metal K8s vs. OpenStack



- **Bypassing OpenStack:** Evaluating direct bare-metal K8s deployment vs. managing a complex OpenStack virtualization layer.
- **User VM Requirements:** Local researchers prefer VM access. **Potential Solution:** Testing CANFAR interactive desktop sessions as a replacement for traditional VMs.

Multi-Instance Ceph Cluster

- **Current State:** All storage resources currently depend on one Ceph instance.
Improvement: A multi-instance Ceph to avoid single points of failure.

Multi-Cluster Scaling Model: Rather than expanding the existing cluster indefinitely, deploy separated dedicated clusters.



The espSRC Node Architecture in SRCNet

↪ Acknowledgements

Financial support from the grant PID2021-1239300B-C21 and PID2024-1558170B-I00 funded by MICIU/AEI/ 10.13039/501100011033 and by ERDF/EU

The SPSRC acknowledges financial support from the Agencia Estatal de Investigación (AEI) through the "Center of Excellence Severo Ochoa" award to the Instituto de Astrofísica de Andalucía (IAA-CSIC) (SEV-2017-0709) and from the grant CEX2021-001131-S funded by MICIU/AEI/ 10.13039/501100011033.

