CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

Hewlett Packard Enterprise

University of BRISTOL

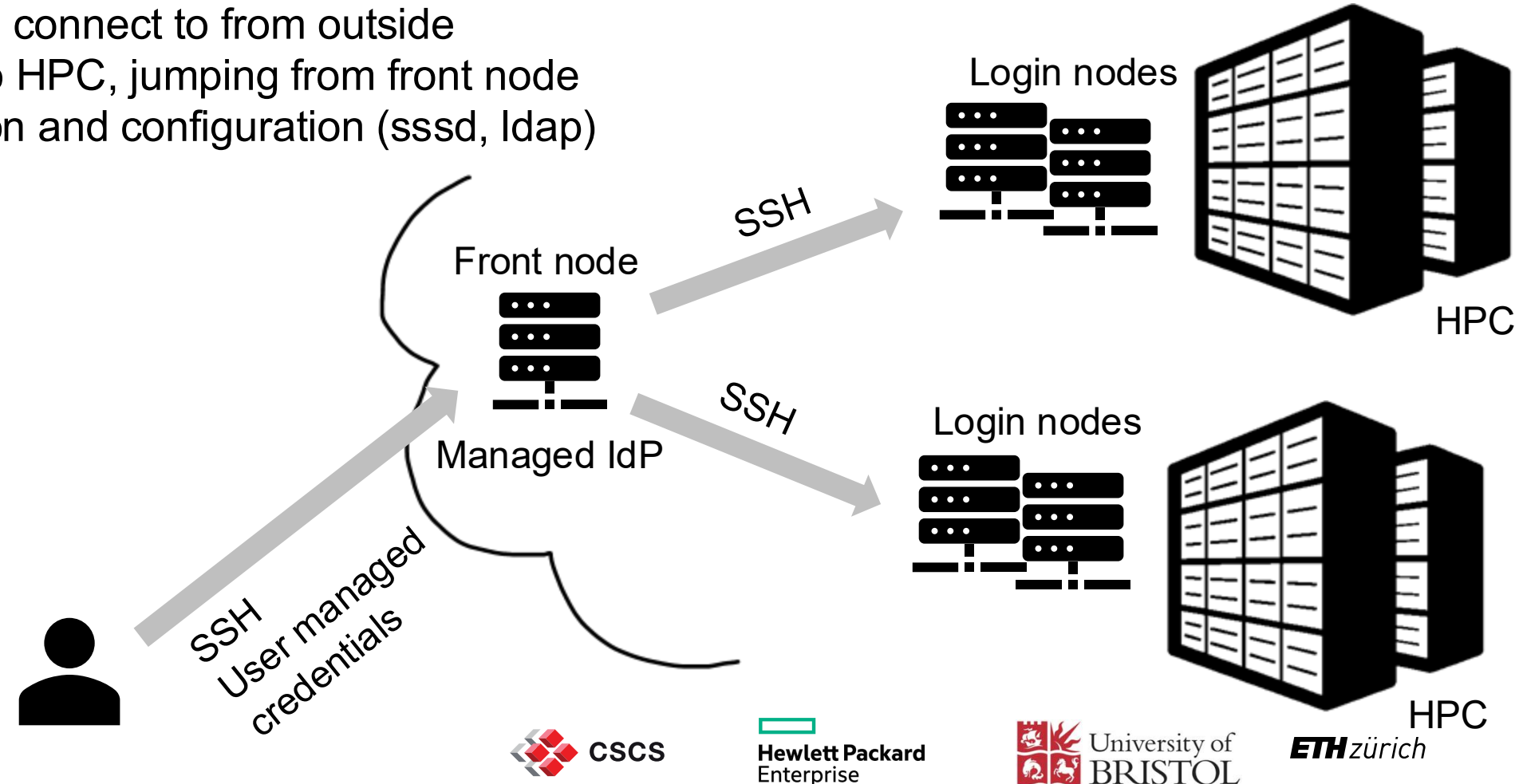
ETH zürich

CUG25 BOF: Rethinking Interactive HPC Resource Access

Maxime Martinasso – CSCS / ETH Zurich
Sadaf Alam – University of Bristol / BriCS
Isa Wazirzada – HPE

Command Line (CLI) interactive access methods (traditional)

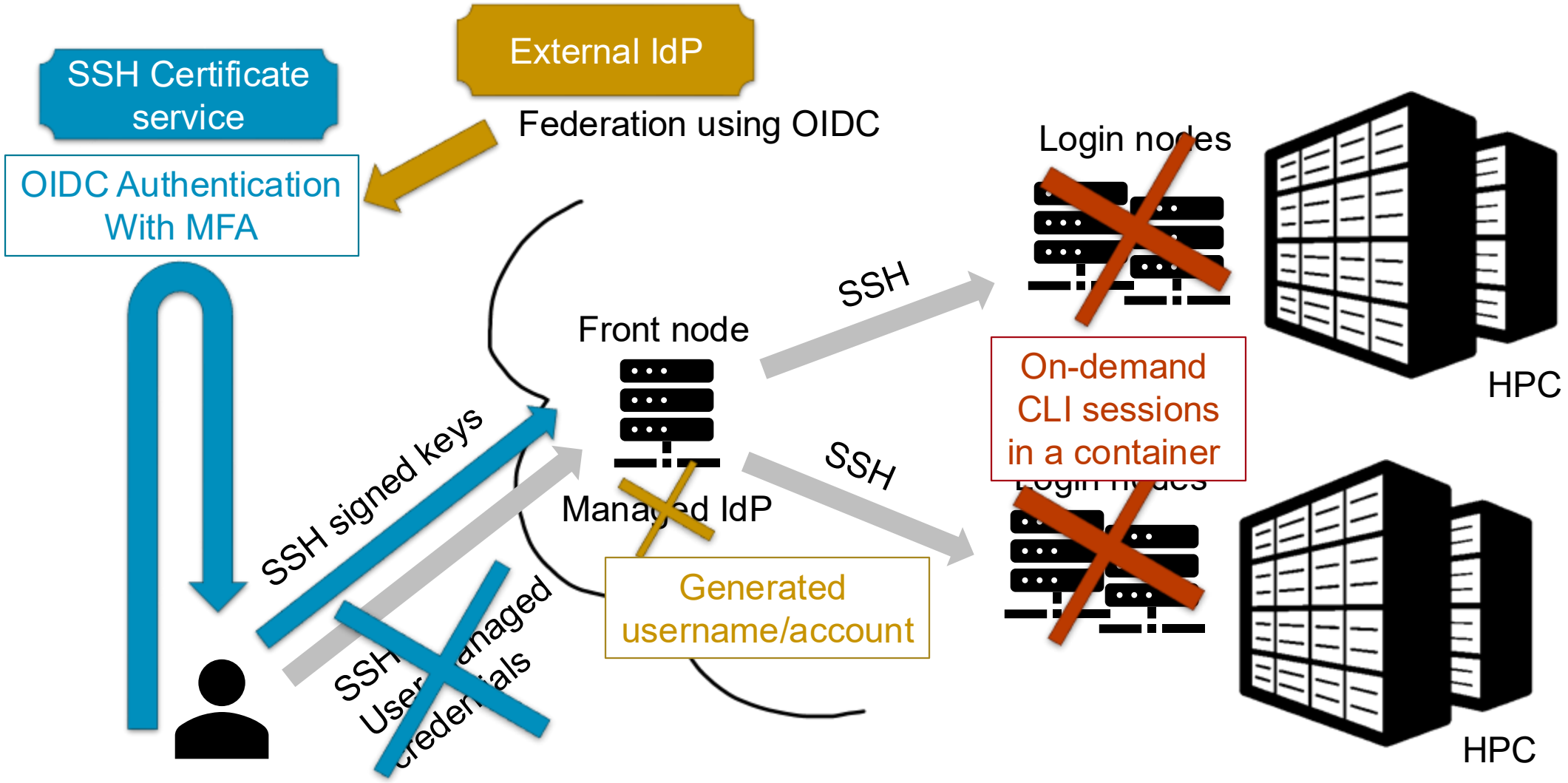
- Interactive access CLI needs
  - Identity provider (IdP) to provide a POSIX IAM (Unix account)
  - A front node to connect to from outside
  - Login nodes to HPC, jumping from front node
  - SSH installation and configuration (sssd, ldap)



Command Line (CLI) interactive access issues

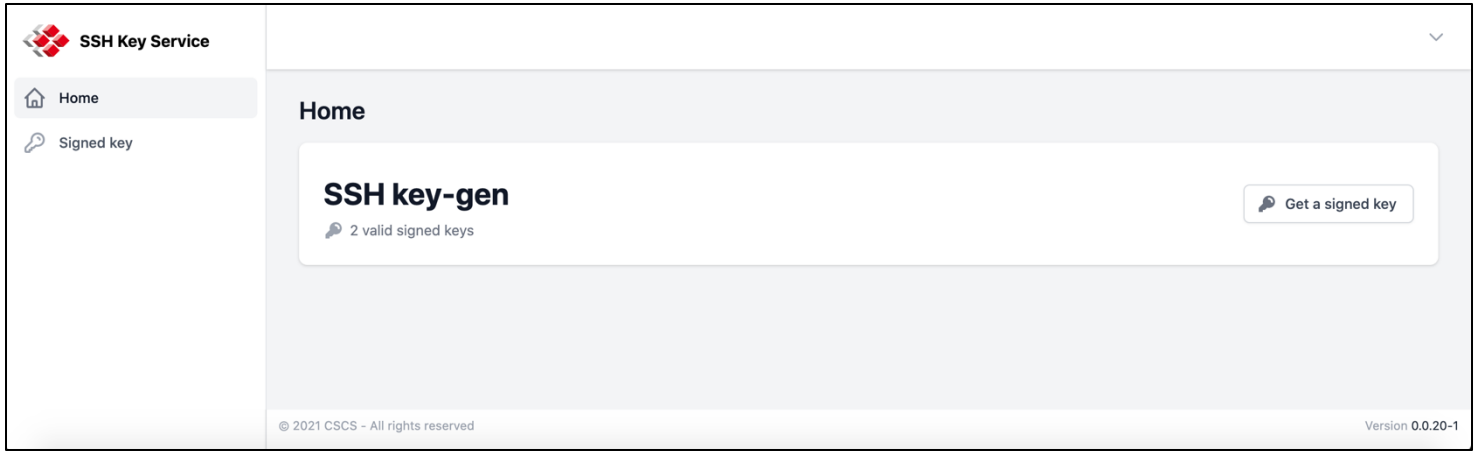
- Security burden
  - User credentials stealing: user-managed keys has proven to be dangerous
  - Privilege escalation: users have access to front and login nodes with potentially vulnerability
  - Security needs to harden, monitor login nodes, alert system, response…
- Shared resources
  - Login nodes are shared, no confidentiality, interferences among users
  - Layer of front and login nodes, users need to jump, it has a cost
  - Require multiple load-balanced login nodes
  - Insure high availability of login nodes, no login nodes = no interactive access
- Limiting further extensions:
  - SSH is used for programmatic remote execution instead of API calls
  - SSH is used for tunnelling/port forwarding instead of a proxy service
  - Difficult integration of SSH technology with AAI standards (OIDC, MFA)
  - No new evolution of SSH standard with new capability (only updates)

CSCS Hewlett Packard Enterprise University of BRISTOL ETH zürich


SSH Certificate service
External IdP
Federation using OIDC
OIDC Authentication With MFA
Login nodes
HPC
SSH
Front node
On-demand CLI sessions in a container
SSH signed keys
Managed IdP
SSH
Generated username/account
HPC

SSH service at CSCS

- Authentication with a token
  - MFA
- Generate a SSH key pair
- Keys are signed by a CA
- Limited to 24h
- Only method to use SSH
- Service with a web-facing API
- Web and CLI access



Investigating alternatives

- Podmansh
  - Execute a user shell within a container when the user logs into the system
  - Easy to limit user visibility of the login node
- OIDC device authorisation flow
  - Create a URL with unique code to access from another device
- Web-based interactive environments
  - Open On-Demand, Jupyter notebooks,…
  - Heterogeneous architecture, different objectives
- Shell built using a programmatic interface
  - Solves the IAM and SSH part
  - Interactivity depends on the API capability and performance
  - Need to provide a web-enable API and ecosystem
- ContainerSSH
- OpenPubkey SSH (cloudflare, opensource now)

Introduction and Overview

UANs and UAIs can both provide interactive login access to a CSM managed HPE Cray EX system

User Access Node (UAN)

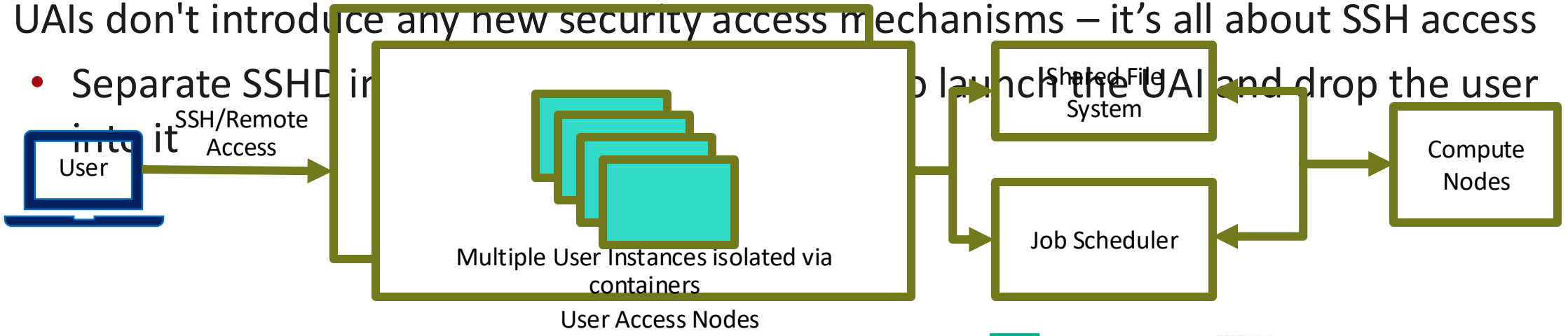
- Dedicated multi-user node (HW & SW) ideal for stable long term persistent tasks
- Can be configured to match compute nodes to ensure a common programming environment
- Restricted to a bare metal image

User Access Instances (UAI)

- Container based environment
- Flexibility to launch UAIs on physical UAN or potentially a repurposed compute node
- Disposable: environments come and go without loss of user data
- Content easily customized for specific activities using both images and volume mounts
- Can have multiple container images for different use cases
- Resource requirements easily customized for specific activities using resource specifications

Security Viewpoint: User Access Instances

- Create and delete rootless Podman container-based UAIs
- No privilege escalation users have same `uid` and `gid` in the UAI as they would on the UAN
- Non-admins cannot provide UAI images
- Admins customize UAI images to limit content, packages, and avenues users have access to
- UAIs don't introduce any new security access mechanisms – it's all about SSH access
  - Separate SSHD ir[illegible]o launch the UAI and drop the user into it



Questions to the audience

1. How can HPC centers transition away from POSIX IAM while ensuring secure and seamless data ownership models?
2. What steps are required to integrate modern authentication methods, such as biometric authentication (e.g., Face ID), into login workflows for HPC resources, and what challenges must be addressed?
3. What technical and operational considerations are needed to provide containerized login instances that support session management, user customization, and scalability?
4. What security measures are necessary to ensure the safe deployment of arbitrary containers as login sessions?
5. Which HPC centers are willing to collaborate on establishing common principles and developing shared solutions for modernizing IAM and resource access workflows?
6. What strategies can be implemented to transition users from traditional permanent resource-provider-managed identities to temporary, session-based access models? How can user adoption and satisfaction be measured during this shift?

https://www.menti.com/alvr35bnbw2e

CSCS Hewlett Packard Enterprise University of BRISTOL ETH zürich