Isambard
IAM and Cybersecurity

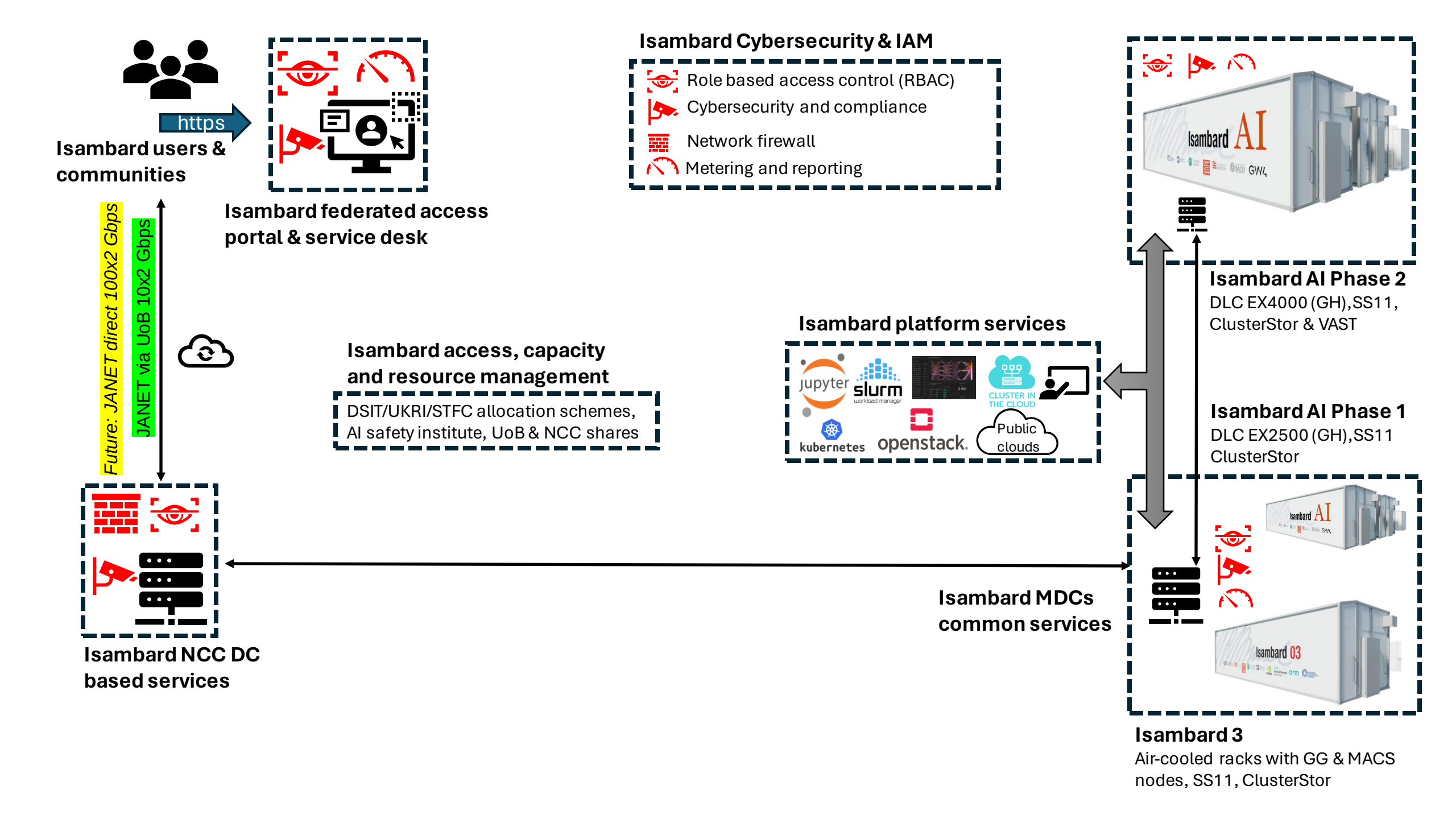
Matt Williams

Sadaf Alam

University of BRISTOL

The Isambard environment

- Large national resource
  - Isambard-AI: ~5k Grace-Hopper GPUs
  - Isambard 3: ~50k ARM CPU cores
- Isambard-AI is part of AIRR, along with Dawn at Cambridge
- Federated identities and access
  - Users will bring their institutional identities
- SSH
  - Slurm
  - Containerised workflows
- HTTPS
  - JupyterHub
  - Other services in future


Isambard Cybersecurity & IAM
Role based access control (RBAC)
Cybersecurity and compliance
Network firewall
Metering and reporting
https
Isambard users & communities
Isambard federated access portal & service desk
Future: JANET direct 100x2 Gbps
JANET via UoB 10x2 Gbps
Isambard access, capacity and resource management
DSIT/UKRI/STFC allocation schemes, AI safety institute, UoB & NCC shares
Isambard platform services
jupyter
slurm
workload manager
CLUSTER IN THE CLOUD
kubernetes
openstack.
Public clouds
Isambard AI
Isambard AI Phase 2
DLC EX4000 (GH),SS11, ClusterStor & VAST
Isambard AI Phase 1
DLC EX2500 (GH),SS11 ClusterStor
Isambard NCC DC based services
Isambard MDCs common services
Isambard 03
Isambard 3
Air-cooled racks with GG & MACS nodes, SS11, ClusterStor

IAM - users

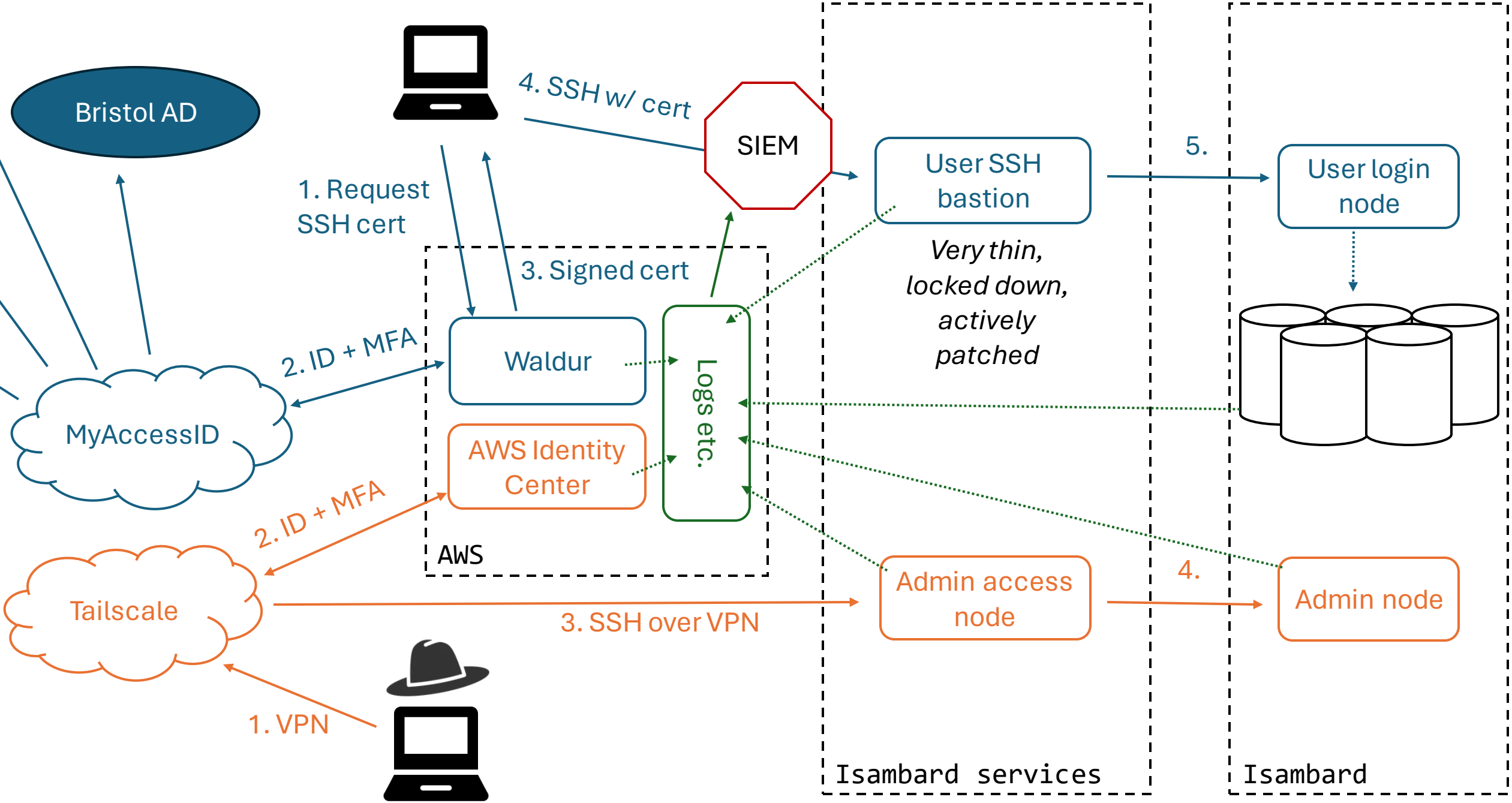
- Access to system is granted based on projects
- Waldur: Central source of user truth
  - Provides Authorisation via OIDC
  - Manages projects, groups and roles
  - Integrates with MyAccessID for Identity and Authentication (MFA)
- SSH connection
  - An SSH key signing CA gets authorisation from Waldur (via OIDC)
  - Signs a short-lived SSH certificate
- HTTPS
  - Connects to Waldur via standard OIDC

IAM - admins

- Completely separate access route to users'
- Admin accounts are managed inside AWS
- Roles are granted to give access to different resources
  - Principle of least privilege for admin roles
- Access is granted via a VPN (WireGuard)
  - Authenticated (MFA) against AWS Identity Center
- Possibly using signed SSH certificates as well

Security, monitoring and alerting

- Primary focus now is reducing incident blast radius, reducing impact and a reliable audit trail.
- Logs and events from all systems will be ingested into an off-site security centre.
  - This will feed into the University's SIEM for the network kill switch
- Role/policy based systems with short-lived access for admins


Bristol AD
4. SSH w/ cert
SIEM
User SSH bastion
5.
User login node
1. Request SSH cert
3. Signed cert
Very thin, locked down, actively patched
2. ID + MFA
Waldur
Logs etc.
MyAccessID
AWS Identity Center
AWS
2. ID + MFA
Tailscale
Admin access node
4.
Admin node
3. SSH over VPN
1. VPN
Isambard services
Isambard

The near future

- Bringing the service online in a secure manner
  - Designing a new system and architecture
  - Scaling up while understand any unique challenges
- Working with other organisations and groups in the UK to share understanding and resources
  - Identifying attacks nationally
  - Responding to new breaches/threats/vulnerabilities
- Compliance

(ARM based Isambard 1, 2) Isambard 3 in 2023/4 (no data centre and Isambard PI Simon MS with a non-dedicated GW4 team)

2016–July 2023

Isambard AI procurement

Sep.–Oct. 2023

Isambard hiring and site groundwork started

Dec. 2023

Service configuration and hardening for AI users

Apr. 2024

Aug. 2023

UK govt feedback on Bristol's Isambard AI proposal

Nov. 2023

AI Safety Summit

Mar. 2024

Modular data centre (POD) and Isambard AI phase 1 installed + a new team in ~3 months. POD and supercomputing delivery & install in < 2 weeks

DSIT AI RR Access Principles Concept Proposal

- Highlights
  - Model oriented towards AI applications
  - Aligned with strategic, high impact projects
    - Research (applied foundation models, AI safety and understanding, large scale AI)
    - Academic (CDT-Centre for Doctoral Training on AI)
    - Startups (TBD)
  - Peer reviewed allocation (similar to HPC—PIs and team vetted)
  - Credits for allocated grants (PIs and team vetted)
  - End station managed by AI experts responsible for selection & vetting
- Infrastructure requirements
  - Containerisation e.g. running Nvidia GPU containers without friction
    - Need for software supply chain and runtime by end users
  - Capacity and resource management and scheduling
    - Batch, interactive (JupyterHub) and reservation-based options

Grant Funding Obligations (Isambard AI and 3)

- Background
  - Isambard 3 funded by UKRI/EPSRC early 2023
  - Isambard AI funded by DSIT/UKRI/STFC end of 2023
- Isambard 3 (classic HPC like Isambard 1 and 2)
  - General Data Protection Regulation, the Data Protection Act 2018, the Bribery, Act 2010, the Fraud Act 2006, the Equality Act 2010 and the Modern Slavery Act 2015, National Security and Investment (NSI) Act 2021, UK Subsidy Control Framework, unlawful State Aid
  - Due Diligence must be undertaken in line with UKRI's Principles on Trusted Research and Innovation: https://www.ukri.org/about-us/policies-standards-and-data/good-research-resource-hub/trusted-research-and-innovation
  - Adequate business continuity plans in place to ensure minimum operational interruptions
  - Adopt the principles, standards and good practice for public engagement with research set out in the 2010 Concordat for Engaging the Public with Research: www.ukri.org/about-us/policies-standards-and-data/good-research-resource- hub/public-engagement/
- Isambard AI (new UK national AI RR)
  - Quite a bit of overlap with Isambard 3 with some changes
  - Explicitly defined “authorized user”—a person approved by DSIT/UKRI/STFC or Bristol in accordance with the Access Principles
  - Explicitly defined Access Principles and concepts:
    - Access, eligibility and allocation of time will be managed by UKRI who have discretion as to use and prioritisation.
    - Access is subject to agreeing the University of Bristol’s Isambard-AI access terms in place from time to time.

Allocation Workflows and Responsibilities

- **Grant allocation process**. Remains UKRI and DSIT competence, like Isambard 2. This issues a unique grant ID.
- **Resource access registration process for Isambard via PI and an allocated grant.** Isambard team sets up allocation with validated grant ID e.g. compute quota for a project with a PI. Note that each project has a unique ID, PI, start/end date, capacity, description, users (PI can invite users). This is similar to Isambard 2.
- **New user registration and user link to project**. A user can request to join a project (not preferred) or a PI can invite a user to join his/her project. An invited user will follow a workflow where they register with their institutional ID and the PI confirms that they are indeed the invited user because they'll be sharing his/her quota or allocated capacity. These workflows exist for Isambard 2.
- **Operational monitoring and alerts**. We can define and enforce policies for token refresh timelines and reregistration as needed. All service access follow multi-factor authentication. Details to be agreed with the funders including any need for geolocation monitoring.

Adopting a risk-based approach

- Isambard DRI assets
  - Hardware, software by suppliers incl. MDC/POD
  - Vendor specific config
- Threats
  - Unauthorised access and usage
  - Service interruptions/regression (ensure minimum operational interruptions)
- Vulnerabilities
  - Resiliency and redundancy of HPC services (power interruptions)
  - Denial of service due to legit or unlegit load on shared access point, like ssh gateways/bastion
- …

Security policy framework (policies)

- Existing docs covering Isambard 2 service, SAFE, JISC as a baseline:
  - https://gw4-isambard.github.io/docs/policies/terms.html
  - https://gw4-isambard.github.io/docs/policies/privacy.html
  - https://www.archer2.ac.uk/about/policies/tandc.html
  - https://community.jisc.ac.uk/library/acceptable-use-policy
  - https://www.jisc.ac.uk/uk-federation (bottom of the page: Our UK Access Management Federation service is included within the scope of our ISO9001 and ISO27001 certificates)
  - https://www.archer2.ac.uk/about/policies/safe_acceptable_use_policy.html
  - https://www.archer2.ac.uk/about/policies/safe_privacy_policy.html
- Additional relevant documentation for MyAccessID, Waldur & Puhuri:
  - https://wiki.geant.org/display/MyAccessID/Policies
  - If needed, we can refer to Puhuri, which is based on Waldur, user and project management tool:
    - https://puhuri.io/privacy-policy
    - https://wiki.geant.org/display/MyAccessID/Service+Operations+Security+Policy
    - https://www.lumi-supercomputer.eu/privacy-notice-for-processing-of-user-data-on-the-lumi-service/

Work in progress

- Security operations (incident response)
- Security operations (asset monitoring & management)
- Exercises and pen testing
- Training and documentation
  - Security awareness for technical and non-technical staff and users
- Relationship with other communities
- UKRI DRI cybersecurity task force led
  - IRIS CSIRT
  - EGI
  - …

References

1) Use of Grant Proposal & Training Grant information addendum: www.ukri.org/apply-for-funding/before-you-apply/your-responsibilities-if-you-get-funding/meeting-ukri-terms-and-conditions-for-funding/

2) UKRI Privacy Notice: www.ukri.org/about-us/privacy-notice/

3) UKRI Grant Terms and Conditions web page: www.ukri.org/apply-for- funding/before-you-apply/your-responsibilities-if-you-get-funding/meeting-ukri- terms-and-conditions-for-funding/

4) UK Subsidy Control Framework: See UK Government guidance 'Complying with the UK's international obligations on subsidy control: guidance for public authorities'

5) State Aid: Articles 107 to 109 of Section 2, Title VII, of the Common Rules on Competition, Taxation and Approximation of Laws, Consolidated versions of the Treaty on European Union and the Treaty on the Functioning of the European Union (2008/C 115/01)

9) UKRI Due Diligence Guidance and Questionnaire: www.ukri.org/about-us/policies-standards-and-data/good-research-resource-hub/equitable-partnerships/

10) UKRI Trusted Research and Innovation Principles: https://www.ukri.org/about-us/policies-standards-and-data/good-research-resource-hub/trusted-research-and-innovation/

11) Concordat for Engaging the Public with Research: www.ukri.org/about-us/policies-standards-and-data/good-research-resource-hub/public-engagement/

13) Policy and Guidelines on Governance of Good Research Conduct: www.ukri.org/publications/ukri-policy-on-the-governance-of-good-research-practice/