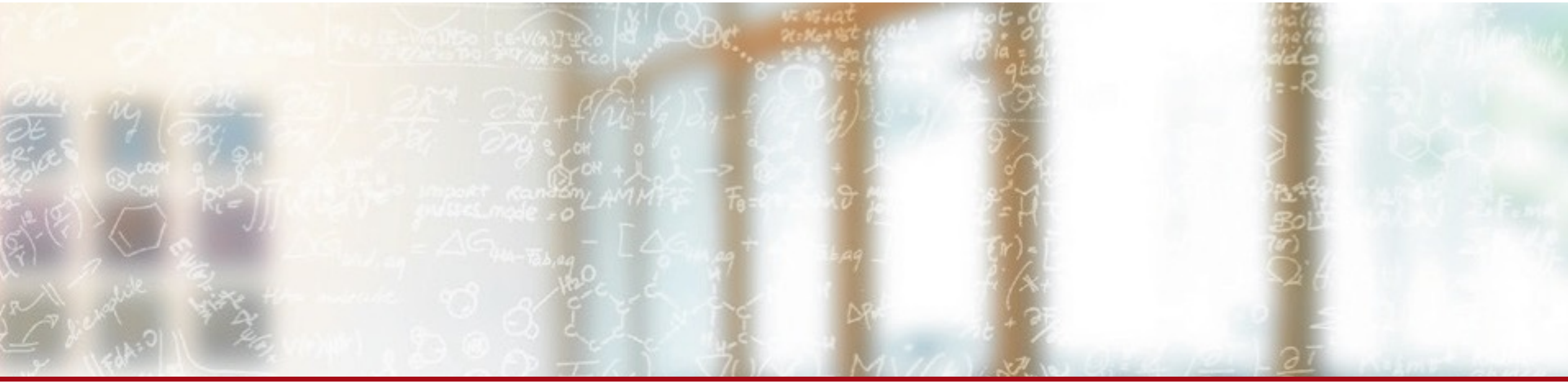
CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich



Replacing POSIX/Linux authentication, towards a new Identity and Access Management for HPC

Maxime Martinasso, CSCS
hpc-ch forum
07 Oct 2021

Identity and Access Management (IAM) concepts

- **Authentication**
  - Verifies that an entity is who/what is claiming a digital identity
- **Authorization**
  - Defines what operation an entity can perform
- Roles
  - Group of operations associated to identities
  - Easier management of authorization
- Other concepts:
  - Role-based access control
  - Federation
  - Delegation
  - Protocols (Oauth2, OIDC, SAML)

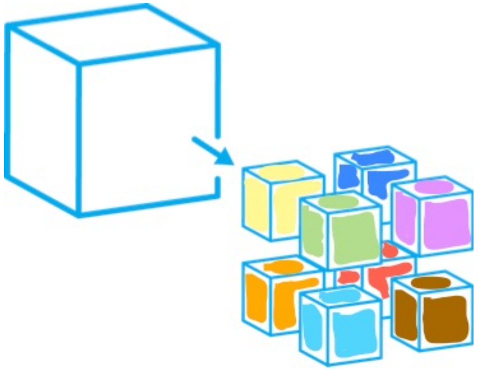
Existing and future IAM technologies

- Origin
  - Growth of Internet users accessing commercial services
  - Needs of a Single-Sign On capability to avoid explosion of identities
  - Definition of an ISO standard

- Architecture
  - IAM services that are deployed using an IAM tools
    - RedHat Keycloak
    - FreeIPA
  - Cloud service to manage identities
    - Cloud players: AWS, Google, Azure, Oracle,…
    - Specific companies
    - Pay by the number of users

- Future
  - Zero-trust – verify at every step
  - Self-sovereign identity – only users own their identity data
  - Bio/Behavior authentication – your body is your password

Why do HPC needs a better IAM?



Heterogeneity of hardware and services

- Authentication: HPC needs higher security
  - 2^{nd} factor authentication, bio authentication
  - Manage SSH keys lifetime
- Authorization: HPC needs a finer role control
  - Special roles: principal investigator, different levels of admin
  - Delegate roles to other users
- Automated interoperability among services
  - Breaking HPC monolithic integration into micro services
    - Micro services need access control
  - APIs use IAM protocols
- Federations
  - One identity per person not per person/service
  - Reduce business/legal logic of managing identities

Linux user management / POSIX file permission

- Linux user management
  - Authentication done with a username/password
  - One user id number (uid) per user
  - Several group id numbers (gid) per user
  - Two roles: user or root

- POSIX file permission
  - **R**ead, **W**rite and e**X**ecute
  - Classes
    - For a user owning a file
    - For a group associated to that file
    - For anyone else
  - Other execution modes
    - set user id, set group id
    - Impersonate users/groups at execution

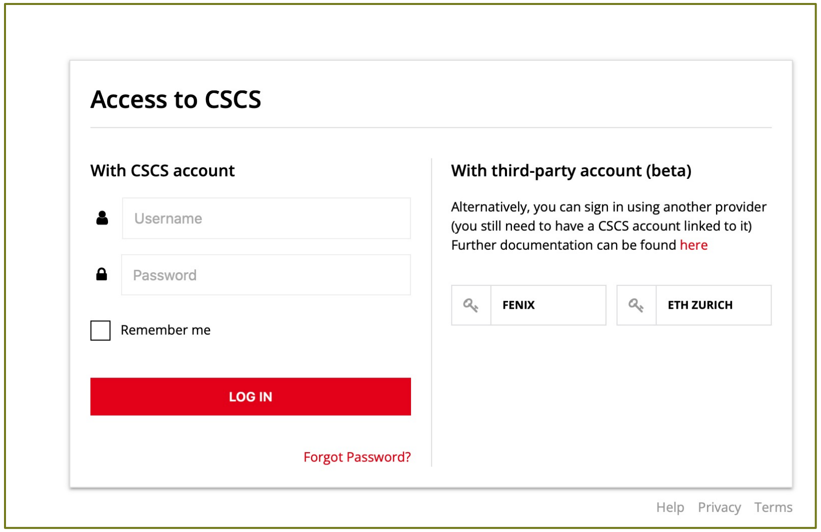
Technology from other needs

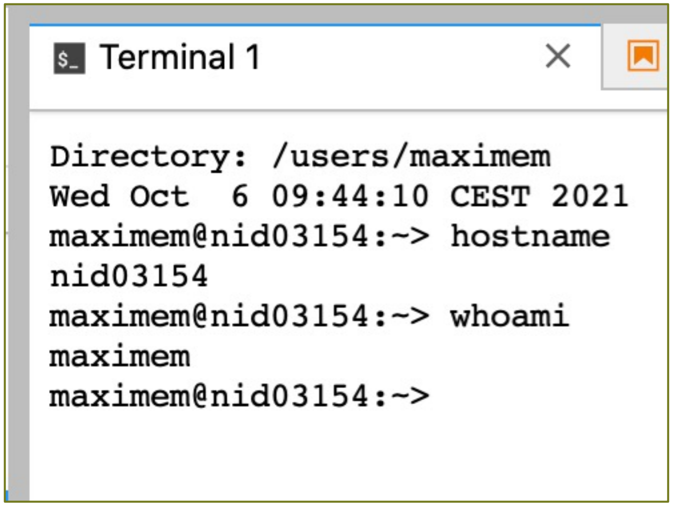
For HPC IAM:
SSH access is the key and the problem!

- Use for secure remote access to systems
  - Execution of remote commands or interactive login
  - Two ways to access
    - A user enters credentials (login and password)
    - A user generates a key pair
  - Often one single non-web point of entry to systems
- Integration of new authentication capabilities?
  - Higher level of authentication?
  - New form of authentication proof?
- How SSH keys are being managed?
  - How keys are being created?
  - Where keys are being stored?
  - One key for all doors?
  - Well, it all depends on the users…

New authentication for HPC - Web Terminal

- Using the web browser to access systems
  - A terminal in the browser
  - Many solution exists: OpenOnDemand, ShellInABox,…
- Using IAM to authenticate before to access the Terminal
- Example: CSCS interactive service (jupyter)
  - Limited to a job allocation



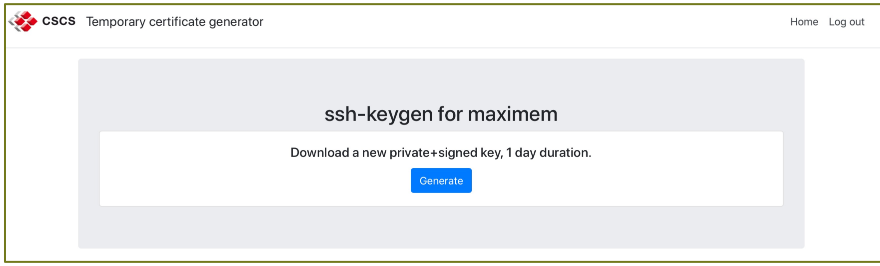


New authentication for HPC - Terminal

- From a terminal with a web browser pop-up
  - Is this feasible/manageable for all users?
- From a terminal
  - Use of a special binary on the client side
    - Doesn't use ssh command but install a new program to do ssh
  - Use of a special binary on the server side
    - Configure ssh to use an extra command
  - Often limited to provide a 2nd factor authentication, federation?
- What about SSH keys?
  - If a user can create keys, then the new authentication is bypassed
  - If it is forbidden to create keys, then many automated workflows will break

New authentication for HPC - SSH service

- Using only SSH keys
  - Create a web service to generate key pairs
  - Access the service after an IAM authentication
  - Use of a certificator to sign keys
  - Different scopes: lifetime, IP origin, limited to certain commands
  - Provide a script that uses the API of the web service to fetch keys
- Disable login with credentials
  - Only certified keys can be used
- Only use native SSH features



- Different scopes of keys
  - User default access to systems – short lifetime (1 day)
  - Automated workflow – mid lifetime (1 week)
  - Service accounts – IP origin, long lifetime (1 year)

What can you do with this SSH service ?

- Multi-factor authentication
  - Something you know (password)
  - Something you have (mobile phone/laptop)
  - Increase security by better proving identity of users
- Federation of identities
  - Avoid managing accounts of external identities
  - Simpler for user, one identity for multiple services
  - Legal aspect of the concept of owning a digital identity
- Easily integrate new authentication technology
  - Bio/Behavior
  - Device flow with push notification

Wait, what about authorization for HPC?

It is a very complex issue without changing key HPC components or established workflows/mentality.

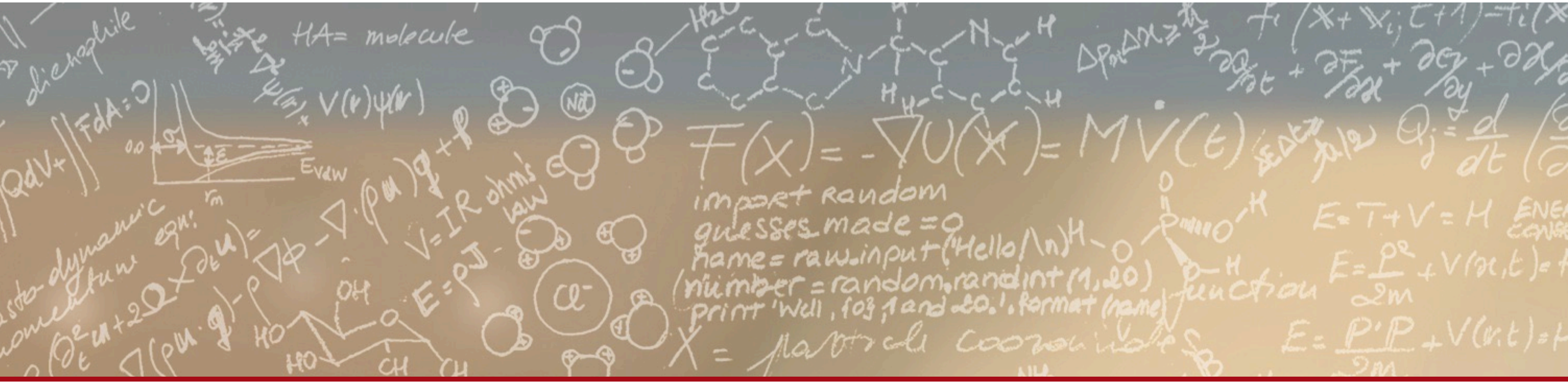
- Provide an authorization service
  - Some tools exist: OPA, Okta,…
- Integrates authorization at several HPC services
  - Is Slurm or PBSPro able to use an external autorization service?
  - Impersonating of the users, is the service authorized to do so?
- (Re)think what are roles and authorization for HPC ?
  - Can I access resources? Move data, use compute resource?
  - Can I configure resources?
- Data access permissions ruled by POSIX
  - Need to use a non-POSIX parallel file system
  - Alternatives are object storage
  - What about performance?

Conclusion

- Indentity and Access management is a great technology
  - Also for HPC!
  - Need to integrate authentication within SSH access
- IAM for HPC offers
  - Higher security: MFA
  - Federation of identity providers
  - Solve the IAM layer for programmable access of resources
  - Follow new technology (bio/behavior, device flow,…)
- A lot remains to do…
  - Common way to define SSH access across HPC centers
  - Level of trust among identity providers and identity proxies
  - Authorization layer at the service level

CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich



Thank you for your attention.