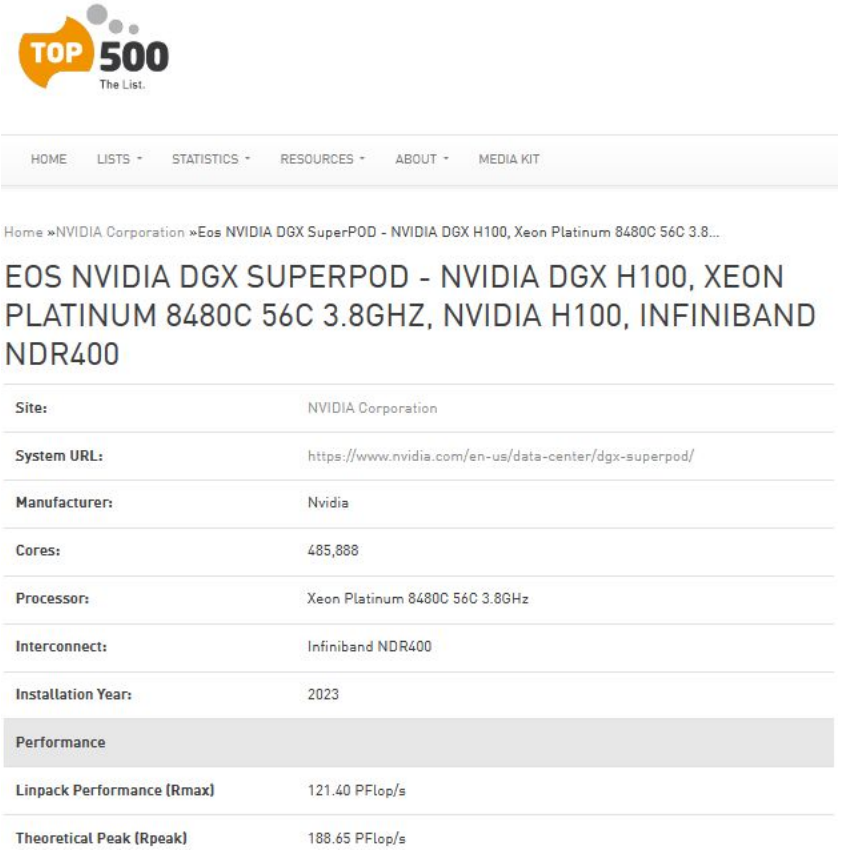
Comprehensive Federated Authentication for AI/HPC Infrastructure

Jonathan Calmels | FOSDEM'25 / 1–2 February, 2025

Applied Systems at NVIDIA

What do we do?

- We bring up the next-gen supercomputers for AI at scale
  - Eos, DGXH100, 2023, #9 in Nov 2023
  - pre-Eos, DGXH100, 2023, #14 in May 2023
  - Selene, DGX A100, #5 in 2020
  - Circe, DGX2H, #61 in 2018
- We enable large scale clusters for internal users and customers.
- We work on new features and advances in the Deep Learning/AI world (e.g. MLPerf, LLMs).

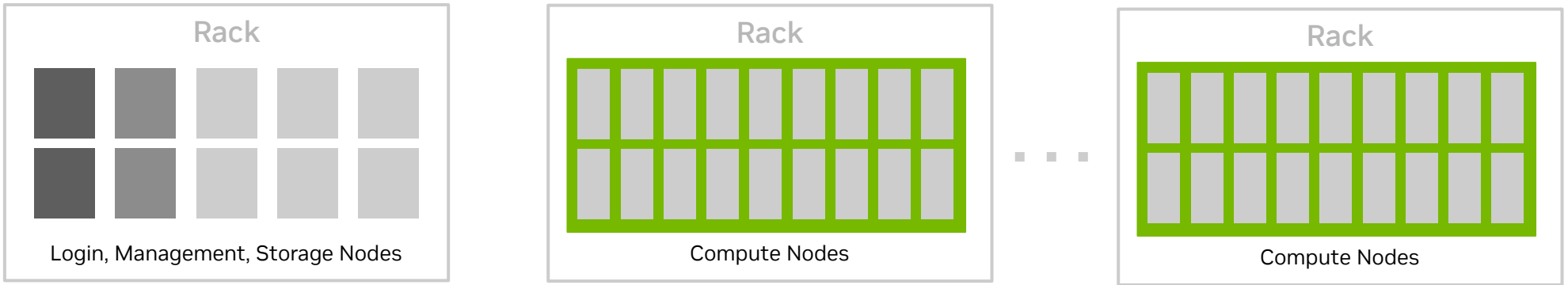


List of large language models

Megatron-Turing NLG	October 2021[28]	Microsoft and Nvidia	530[29]	338.6 billion tokens[29]	38000[30]	Trained for 3 months on over 2000 A100 GPUs on the NVIDIA Selene Supercomputer, for over 3 million GPU-hours.[30]

AI/HPC Infrastructure overview

GB200 SuperPod architecture



NVLink Memory Fabric

Infiniband Fabric Storage (Lustre)

Infiniband Fabric Compute

Ethernet Management

Ethernet Compute & Storage (NFS)

Authenticating users end-to-end

The case for Kerberos and FreeIPA

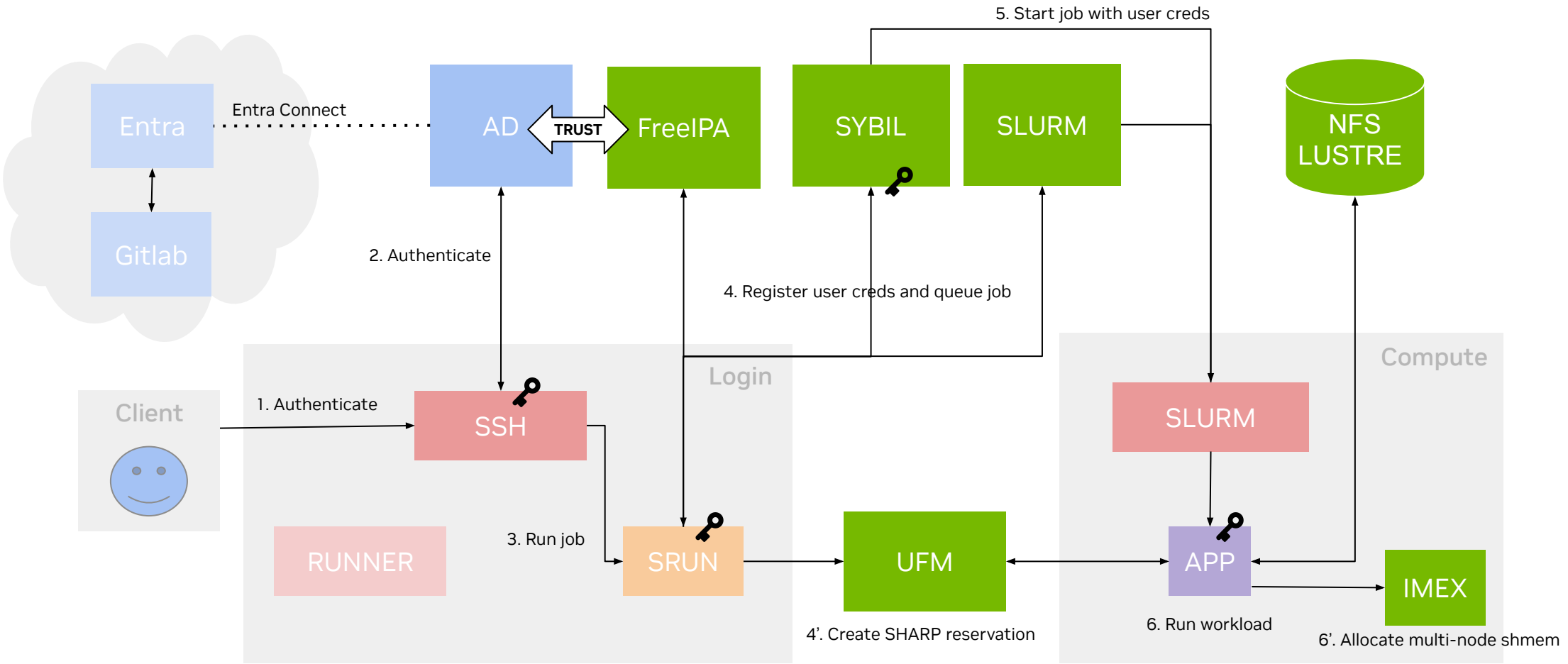
Challenge: Our security posture requires strong e2e authentication federated by corporate infrastructure

What we needed:

- Authn/Authz with Posix accounts tied to Corporate identities (i.e. Microsoft AD)
- Unified solution across our services:
  - SSH, Slurm, Gitlab CI/CD, NFS, Lustre, UFM (SHARP), IMEX (Multinode NVLink)
- On-prem and/or Cloud authentication
- Short-lived credentials with rotation
- Doesn't require client-side modification or custom tooling
- Multi-factor & SSO capable
- Easily auditable
- User friendly
- Centralized solution with minimal DC footprint
- Scalable to various clusters

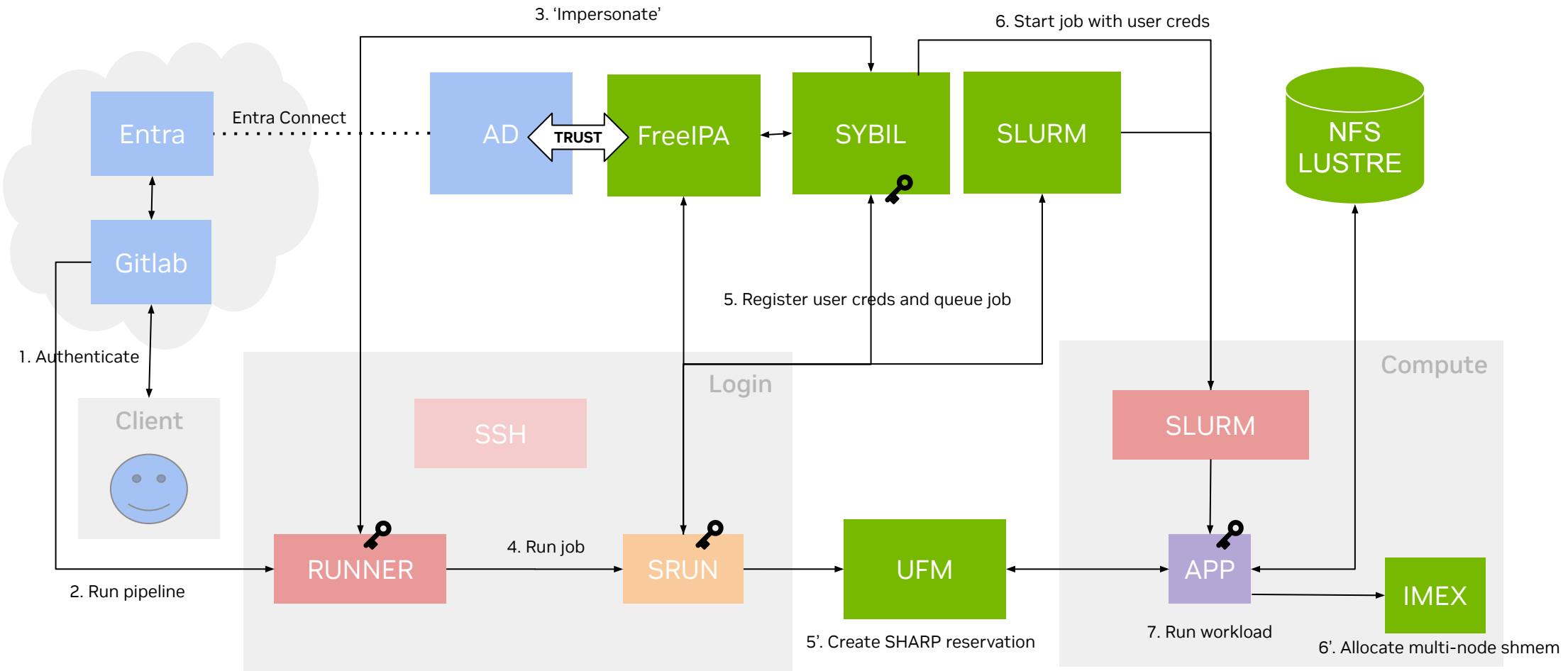
Authentication Architecture

High level overview of user job submission

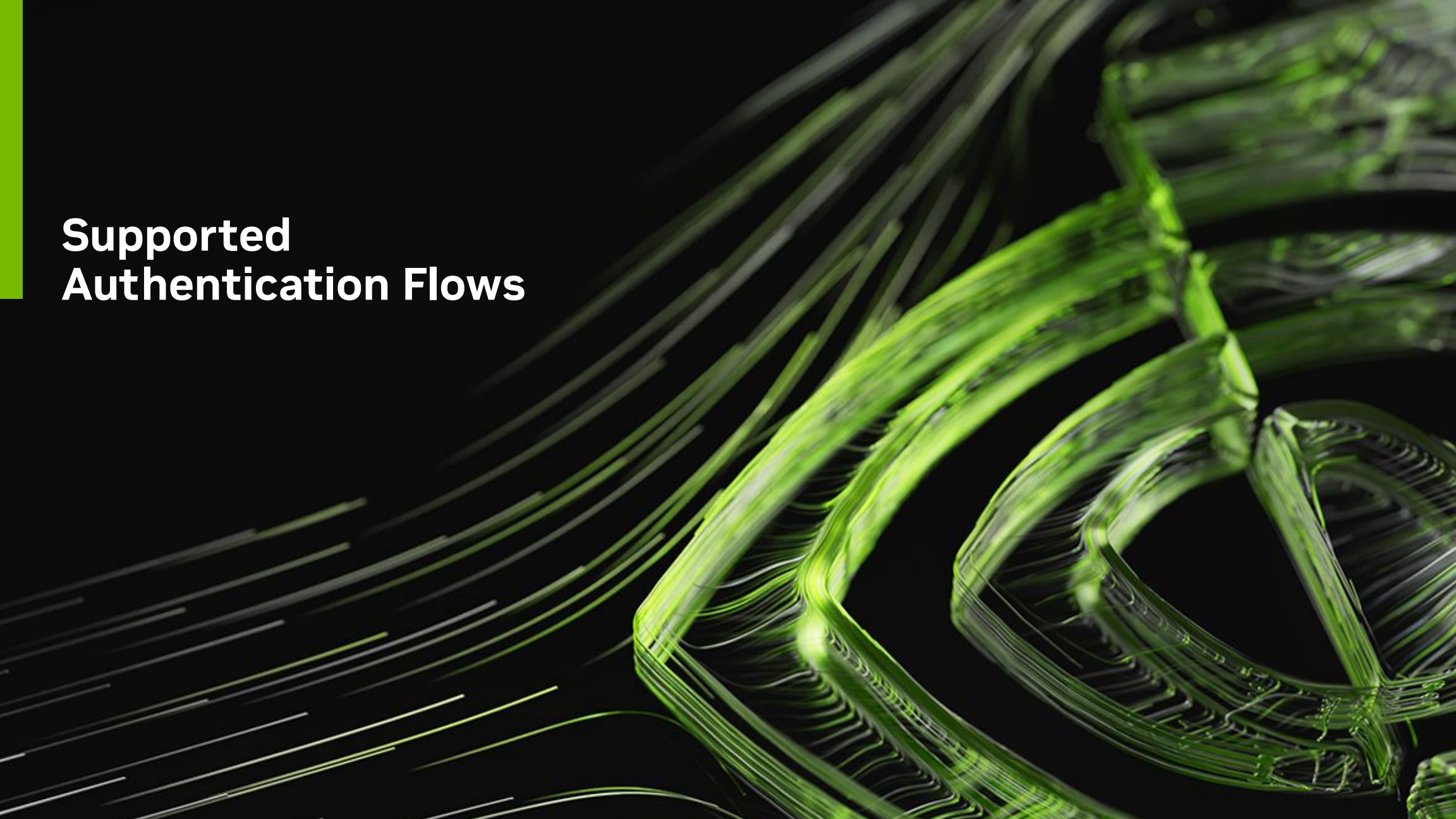


Authentication Architecture

High level overview of CI workflow

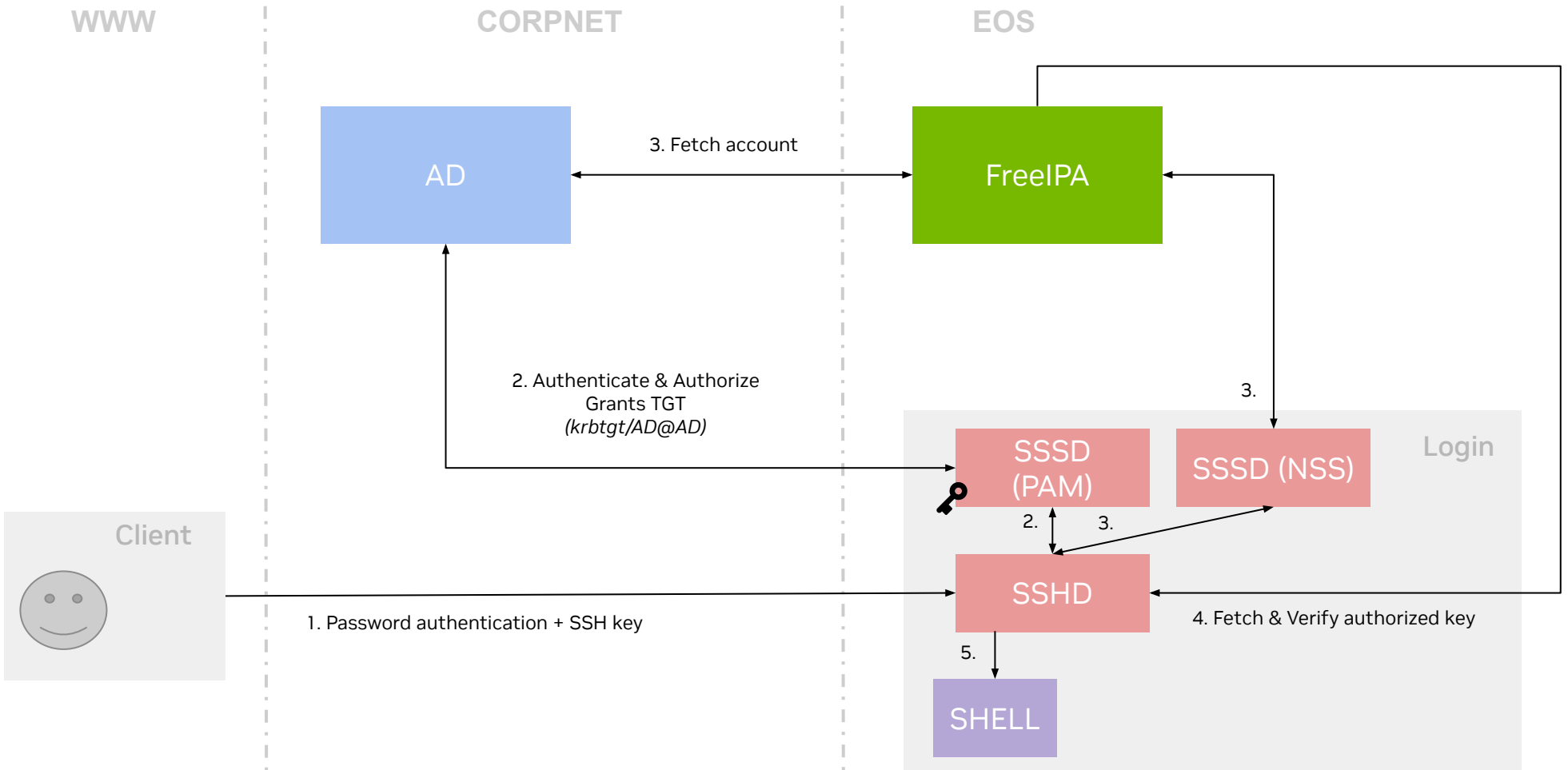


Supported Authentication Flows



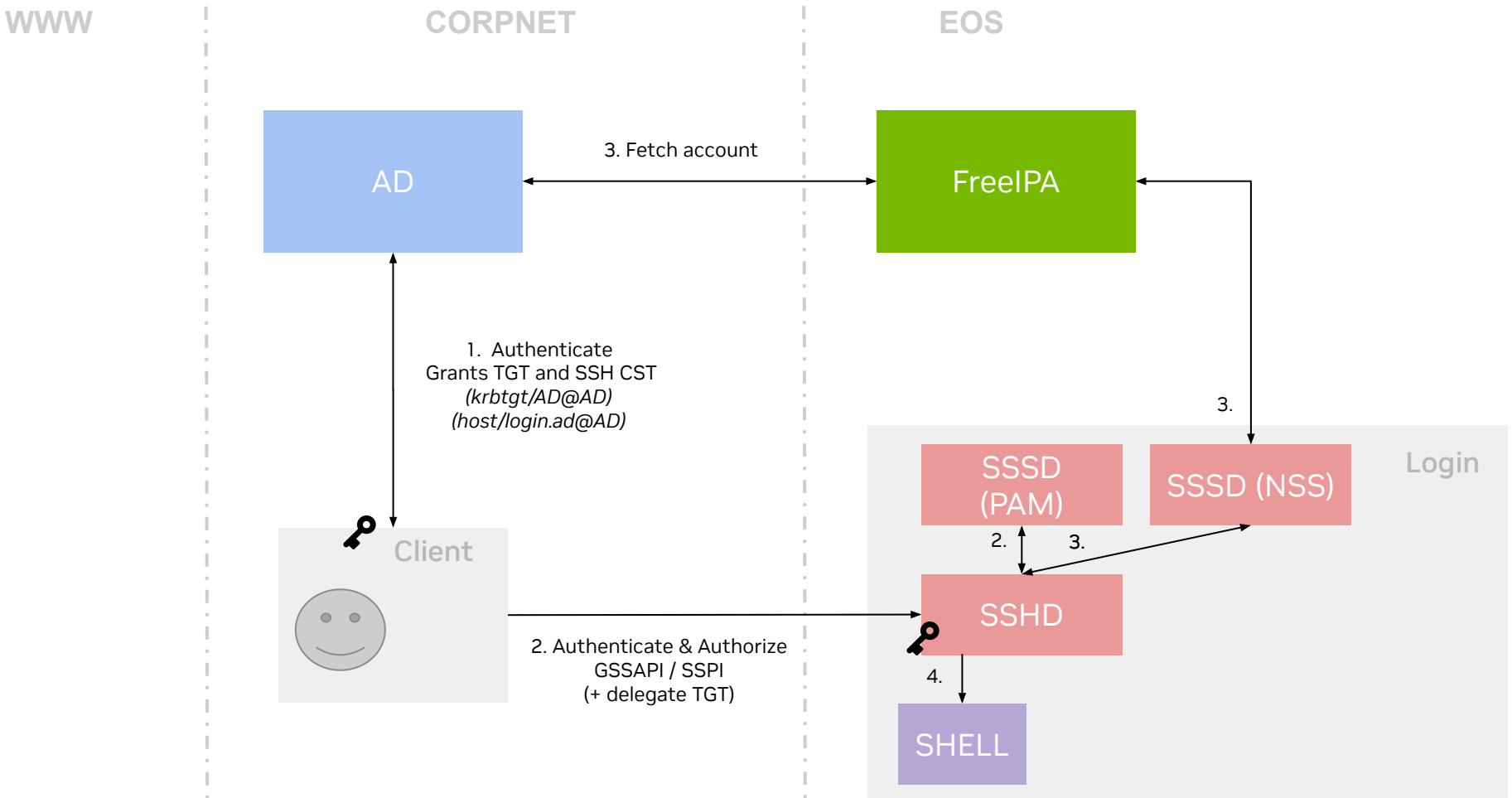
Password Authentication with 2FA

Corporate password + SSH key



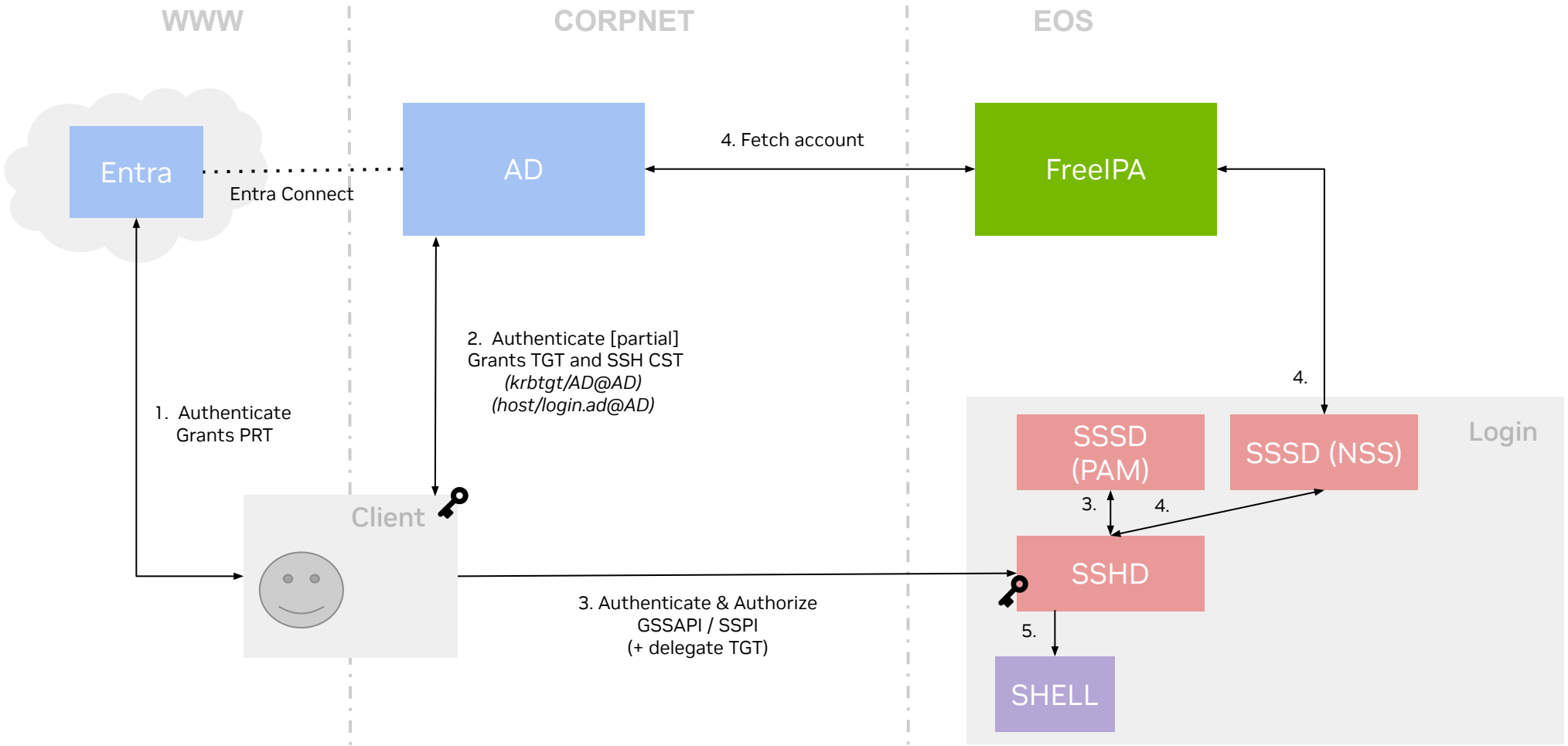
AD Single Sign On

Domain-joined with AD ticket



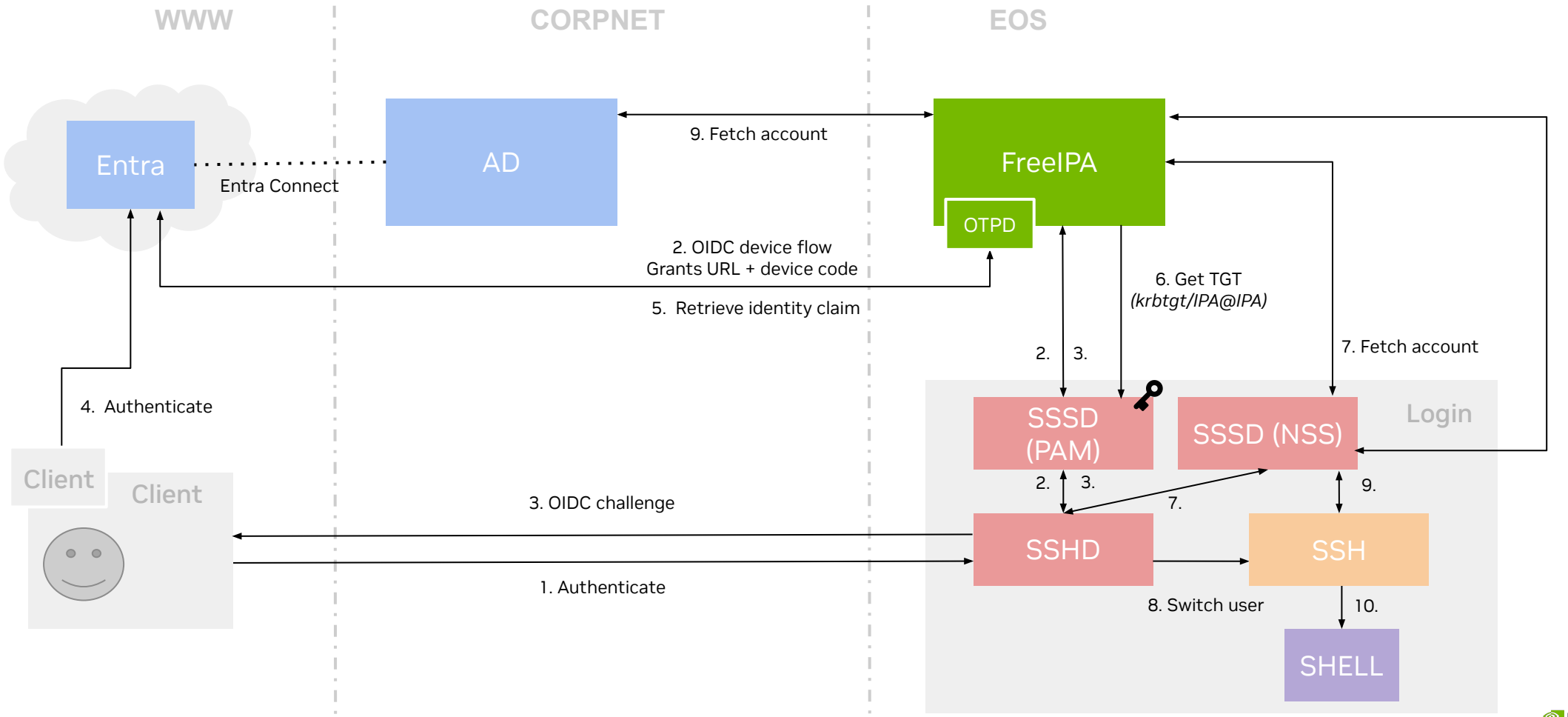
Entra Cloud Kerberos Trust (WHfB)

Hybrid-joined with Entra token + AD ticket



Entra MFA

Cloud-joined with OIDC Web auth (external IdP)



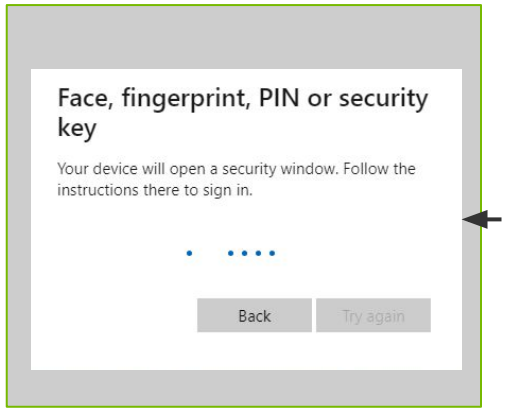
User experience

The different authentication flows

Authentication can be done on-prem or in the Cloud, with or without SSO.

On login, Kerberos tickets are granted transparently and prove one's identity across services

Cloud authentication can be done on any device, through any mean supported by Entra
(e.g. WHfB, FIDO2)




```
# Password authentication + SSH key
$ ssh jcalmels@cluster.domain.lan

(jcalmels@cluster.domain.lan) Password: ********

# AD Single Sign On (or) Entra Cloud Kerberos Trust
$ ssh -K jcalmels@cluster.domain.lan

# Entra MFA
$ ssh jcalmels-mfa@cluster.domain.lan

Authenticate with PIN XXXXXXXX at https://microsoft.com/devicelogin and
press ENTER.

# List tickets
$ klist

Ticket cache: KEYRING:persistent:2000437969:krb_ccache_vr0czrg
Default principal: jcalmels@CLUSTER.DOMAIN.LAN

Valid starting     Expires            Service principal

11/02/23 04:17:55  11/02/23 14:17:30
nfs/nfs.cluster.domain.lan@CLUSTER.DOMAIN.LAN
        renew until 11/09/23 03:17:30

11/02/23 04:17:30  11/02/23 14:17:30
sybil/ipa.cluster.domain.lan@CLUSTER.DOMAIN.LAN
        renew until 11/09/23 03:17:30

11/02/23 04:17:30  11/02/23 14:17:30
krbtgt/CLUSTER.DOMAIN.LAN@CLUSTER.DOMAIN.LAN
        renew until 11/09/23 03:17:30

# Run job
$ srun --container-image ubuntu --container-mount-home bash
```

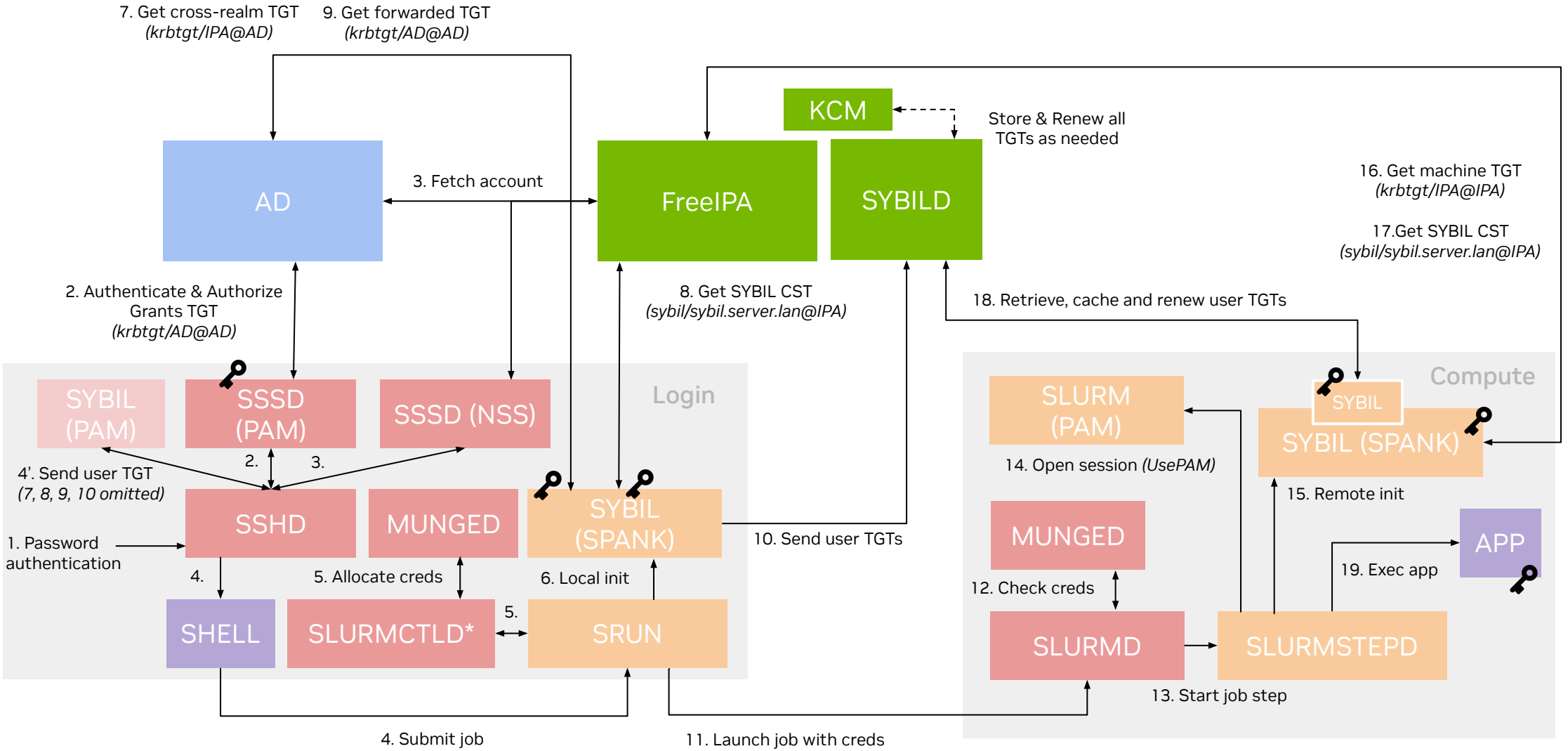
Sybil for Batch/CI Workflows

Sybil

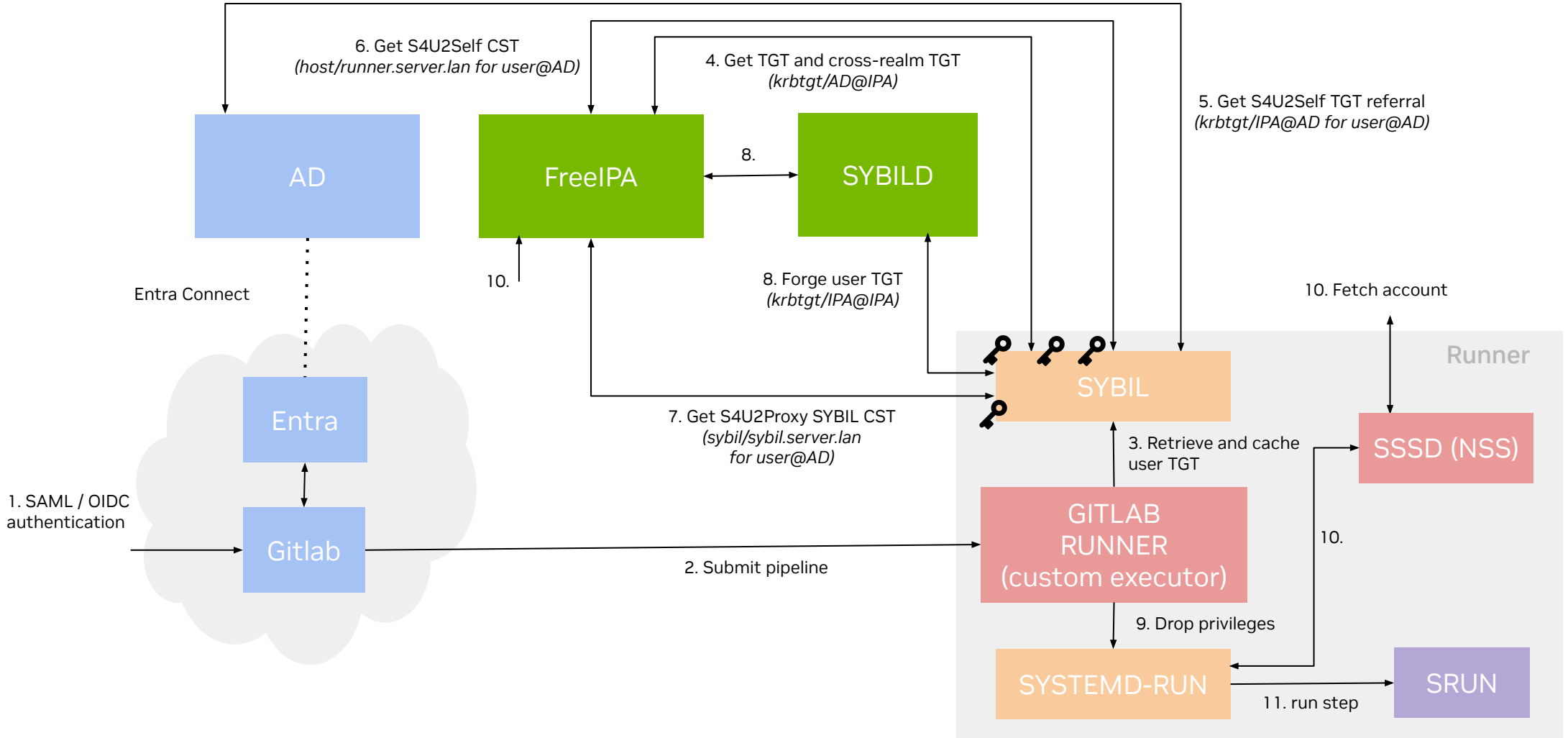
Tools for Kerberos ticket delegation and impersonation

- Small project developed in C/Rust: https://github.com/NVIDIA/sybil consisting of
  - Server to be hosted alongside FreeIPA
  - CLI to delegate/impersonate user credentials
  - Slurm plugin to manage credentials lifecycle with batch workflows
- Improve on https://github.com/cea-hpc/auks
- Adds support for impersonation through protocol transition (S4U) and TGT forgery
- Full documentation is still WIP

Slurm infrastructure



CI/CD authentication



Conclusions & Future Work

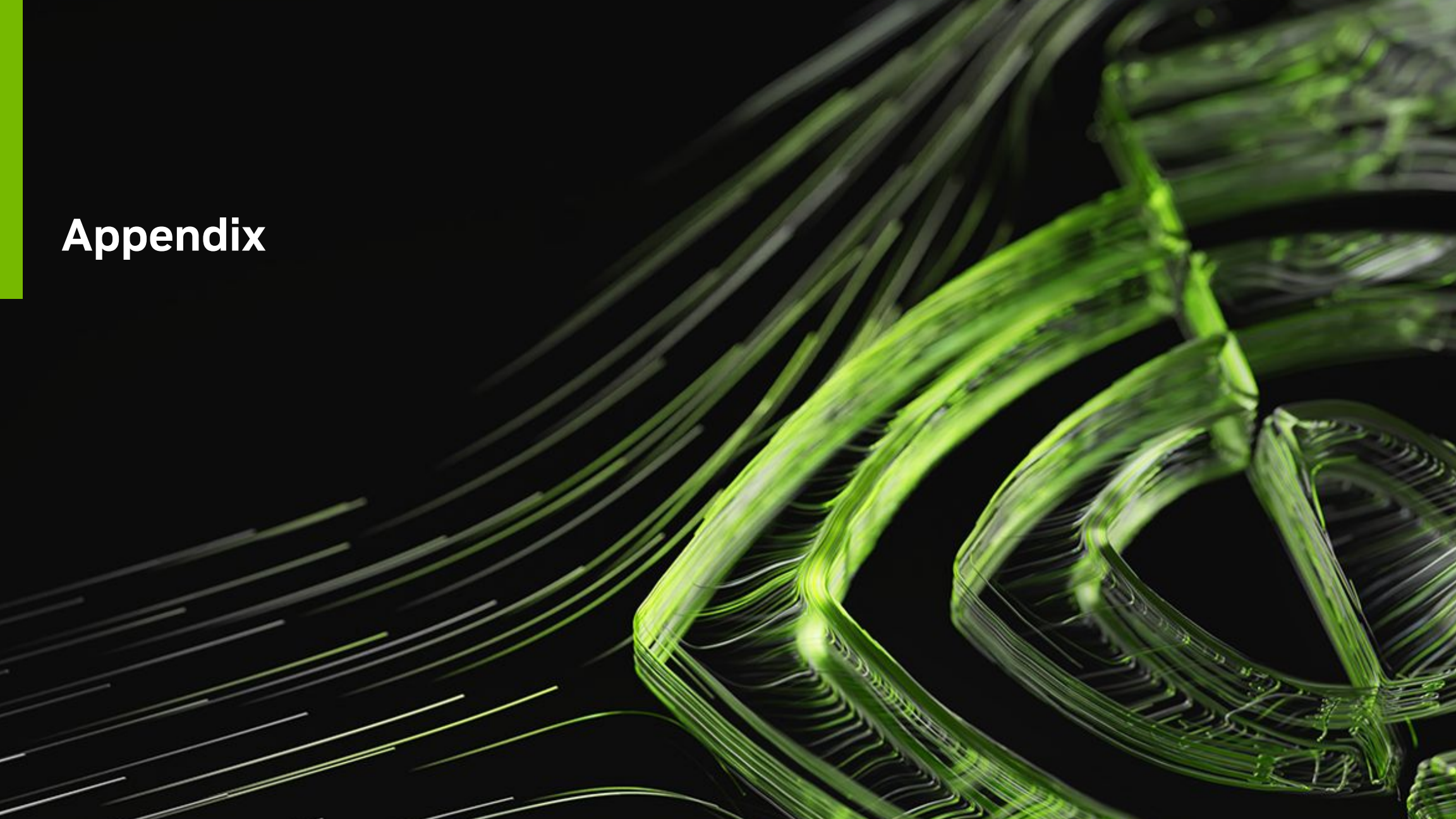
- FreeIPA bridges the gap between Corporate identities, Cloud authentication and traditional Posix accounts
- We could implement complete end-to-end authentication across all cluster resources
- We developed Sybil to help address some scenarios that couldn't be easily supported with standard Kerberos, such as Batch and CI workflows

Next steps:

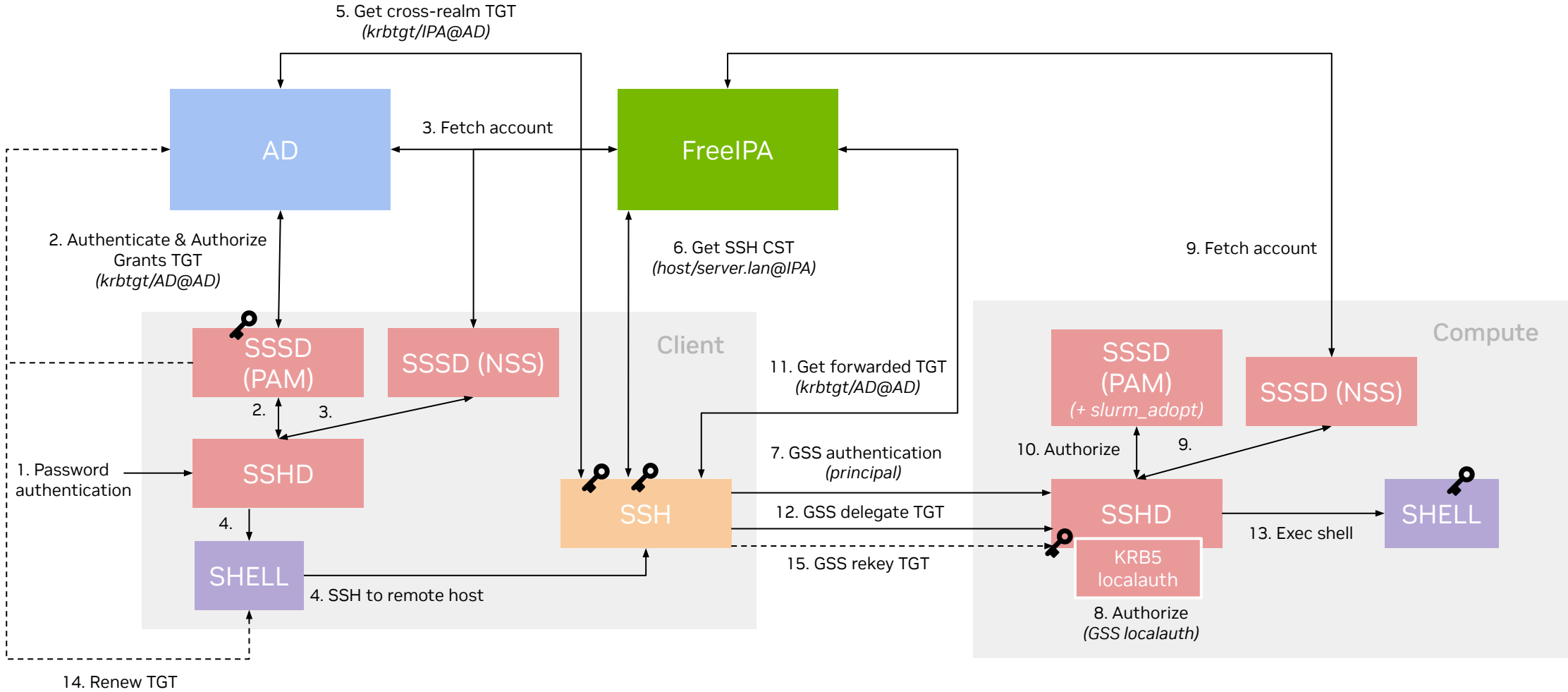
- Look into Keberizing Slurm authentication
- Look into Kerberizing MPI authentication (Pmix)
- Look into replacing local SSH for MFA with systemd run0


NVIDIA

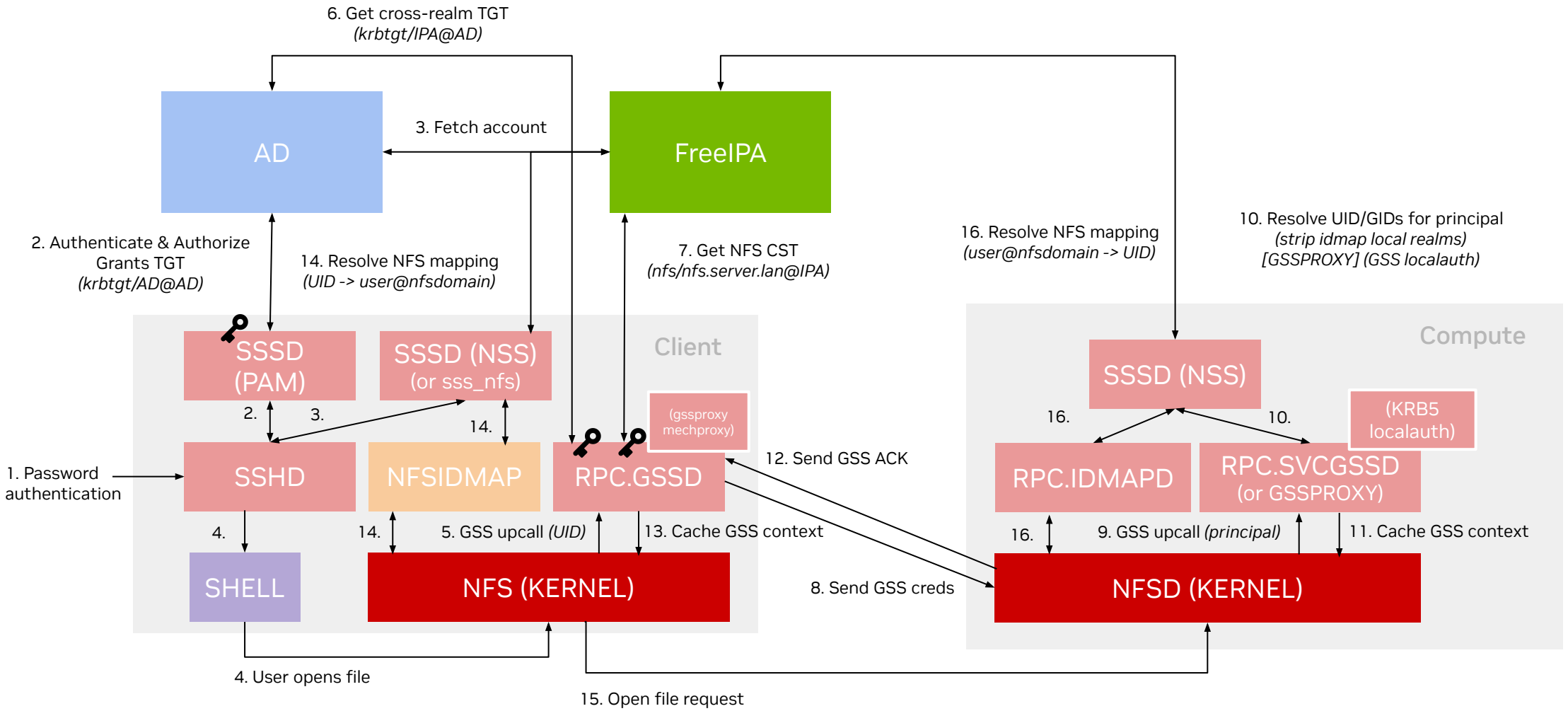
Appendix



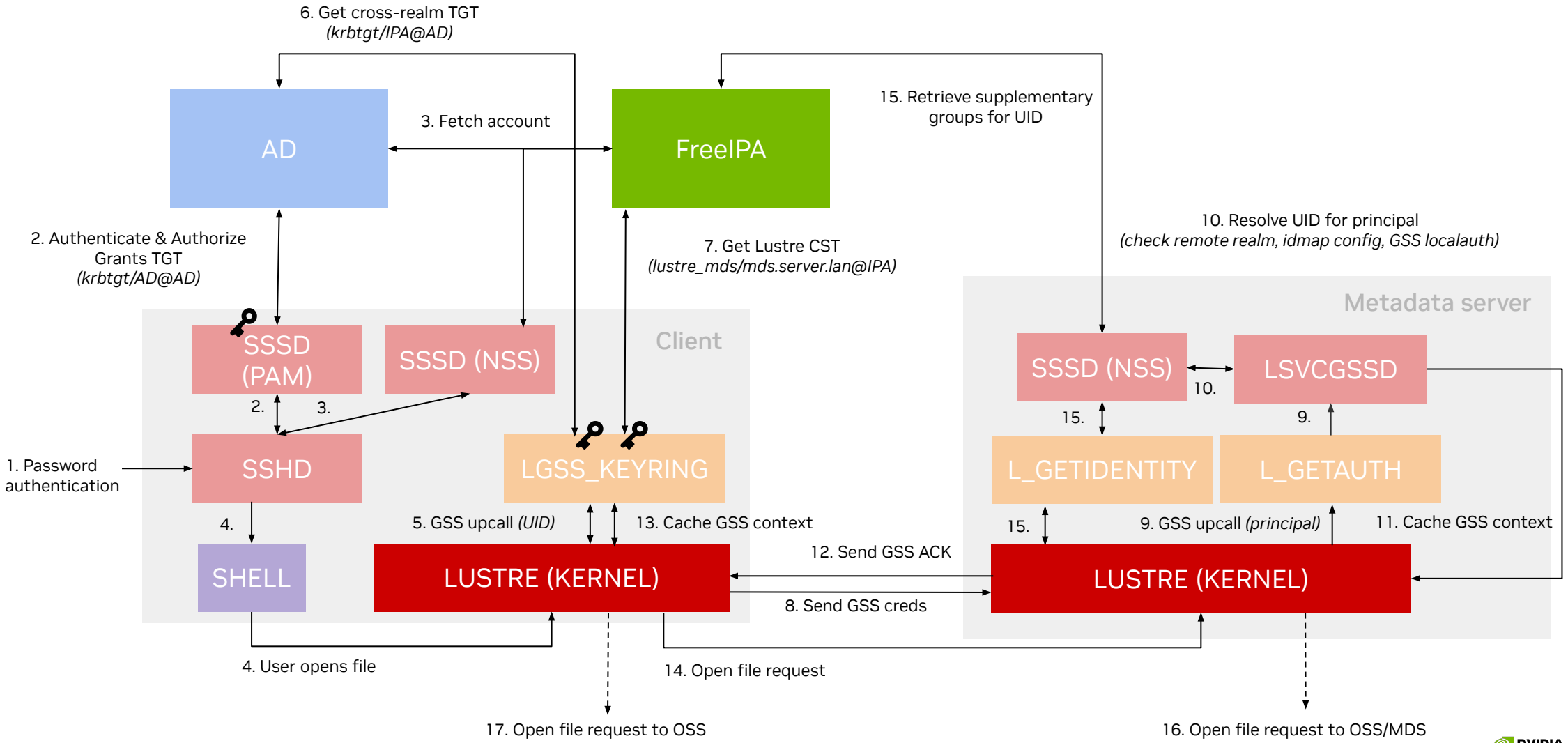
SSH Authentication



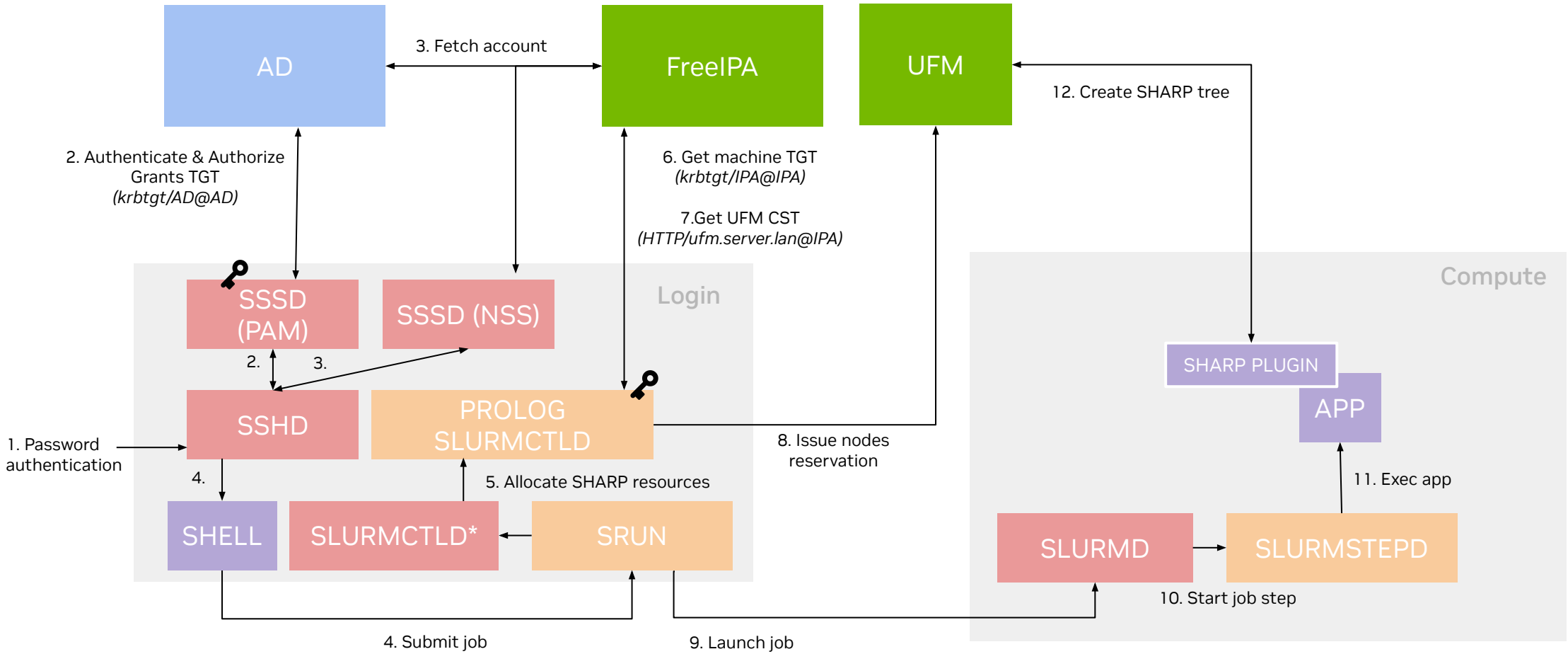
NFS Authentication



Lustre Authentication



SHARP infrastructure



Multi-Node NVLink Infrastructure