Computing Grid Access with Federated Identity

Dave Dykstra, dwd@fnal.gov
Fermilab contributers: Mine Altunay, Dennis Box, Ken Herner, Tanya Levshina, Jeny Teheran

NSF Cybersecurity Summit
August 17, 2016

Fermilab Open Science Grid

Outline

- Introduction & motivation
- Background
  - Grid security & job management
  - InCommon, CILogon, and SAML ECP
  - MyProxy
- Details of the Federated Identity/Grid integration
- Status
- Related Work
- Security considerations
- Bonus: pilot/payload isolation
- Conclusions

Introduction

- Open Science Grid (OSG)
  - NSF- and DOE-funded
  - Collaboration between over 100 independent sites supplying High Throughput Computing (HTC)
    - OSG does not own the computers, commodity hardware
    - Also about 100 Virtual Organizations (VOs) and separately about 100 individual Principal Investigators (PIs)
    - Continually changing and growing
    - Now expanding to commercial clouds & portions of HPC systems
  - Grown to 100 million CPU hours/month end of 2015
    - 10%-20% used opportunistically
- Fermilab is one of the major entry points

Introduction

- Grid security is heavily based on X.509 certificates
  - Very important for its distributed multiple-owner nature
- Managing certificates by hand is often an impediment for grid users that are not tech savvy
  - Especially each year as certificates expire
- Fermilab has a grid job submission system (Jobsub) that hides certificates from users
  - The certificate management piece had shortcomings, however

Motivations for change

- The shortcomings are
  - It only works with Fermilab Kerberos
    - Inconvenient challenge for remote collaborators
  - It requires running our own Kerberos Certificate Authority (KCA)
    - Expensive to maintain
    - Losing software support later this year
- Jobsub also supports manually-maintained certs, but we don’t want to lose automation
- We want to modernize to Federated Identity and so not require everyone to have local login

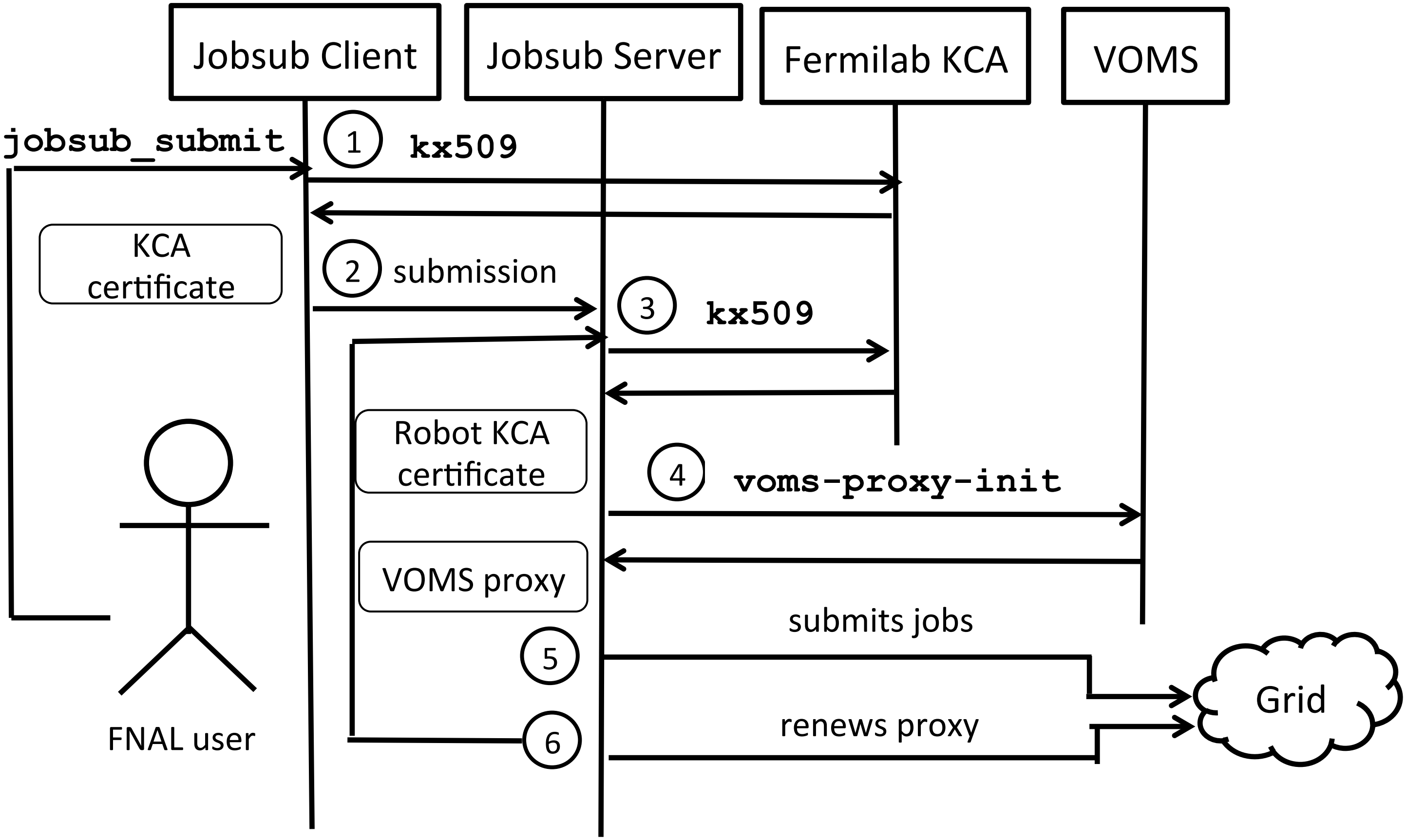
Background – grid security

- Grid users tracked in Virtual Organizations (VOs)
- User certificate Distinguished Names (DNs) registered in Virtual Organization Membership Service (VOMS) servers
  - Cryptographically adds VO info to proxy certificate
- VOMS proxy certs are sent with jobs
  - primarily to access storage
  - usually short-lived to limit their use if stolen and in case user's VO membership is revoked
- Grid User Mapping Service (GUMS) servers additionally used at OSG grid sites to map certs to access rights

Background – grid job management

- Grid job management typically uses two layers
  - Pilot Workflow Management System (e.g. GlideinWMS) provides uniform global queue
  - Grid job submission system (e.g. Jobsub) feeds the global queue
- End users interact with the job submission system
  - System responsible for renewing users' VOMS proxy certificates for long-lived jobs
  - Old Jobsub maintained extra "Robot" kerberos credentials for every potential user in order to get new KCA certs to make new VOMS proxies
    - DNs derived from user's, separately registered in VOMS

Background - old Jobsub submit flow



Background – InCommon, CILogon, ECP

- InCommon Federation
  - Internet2's identity federation for education & research
- CILogon
  - InCommon's X.509 Certificate Authority (CA) service
  - The CA we use is CILogon Basic CA
- InCommon primarily used for web authentication, but CILogon also supports SAML 2.0's protocol for non-web browser environments
  - Enhanced Client or Proxy (ECP)
  - Does not require javascript or web forms support
  - Option in Shibboleth Identity Provider (IdP)

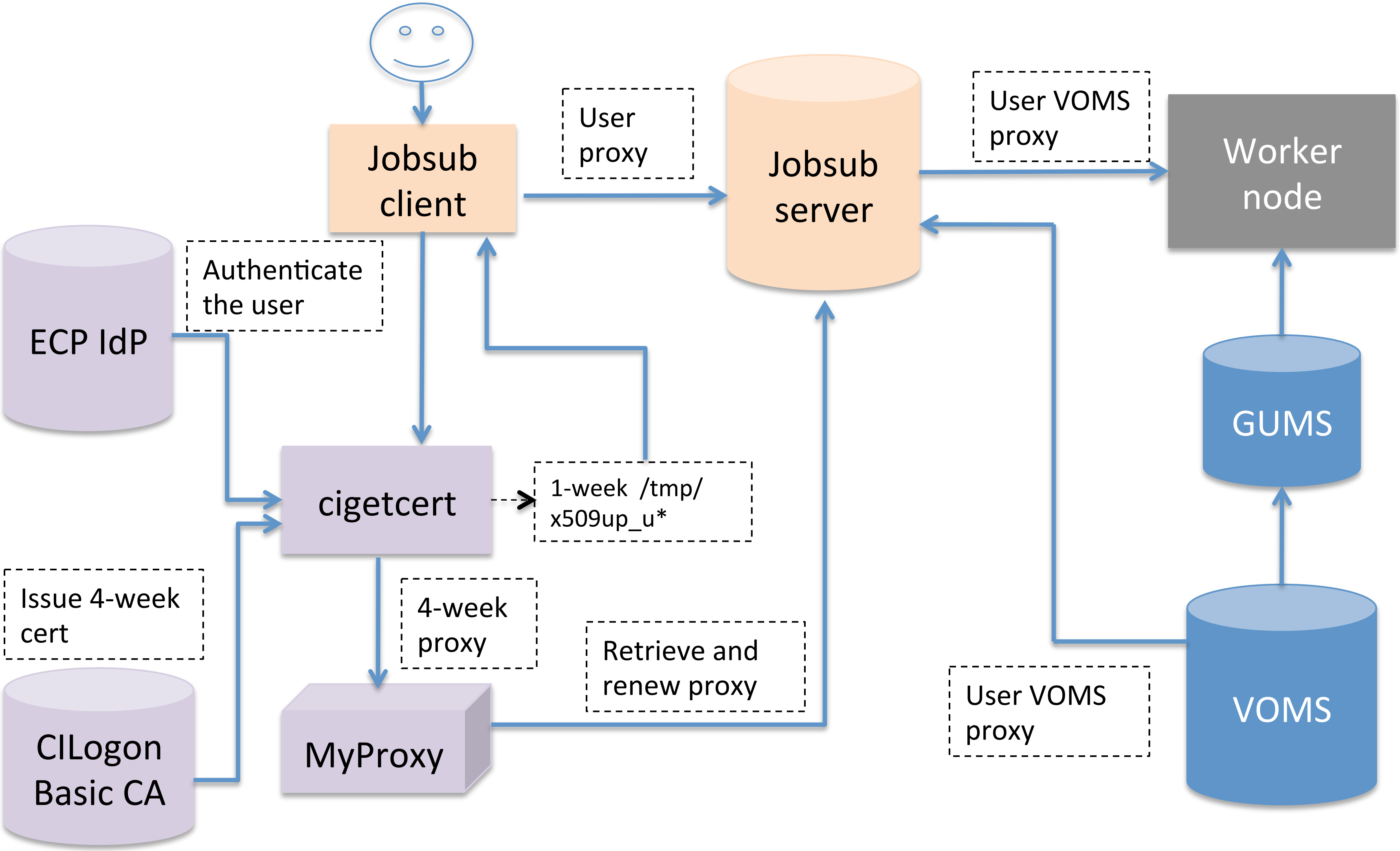
Background - MyProxy

- MyProxy is a secure server for storage of proxy X.509 certificates
  - Software available from NCSA
  - Has many controls over who can access the proxies

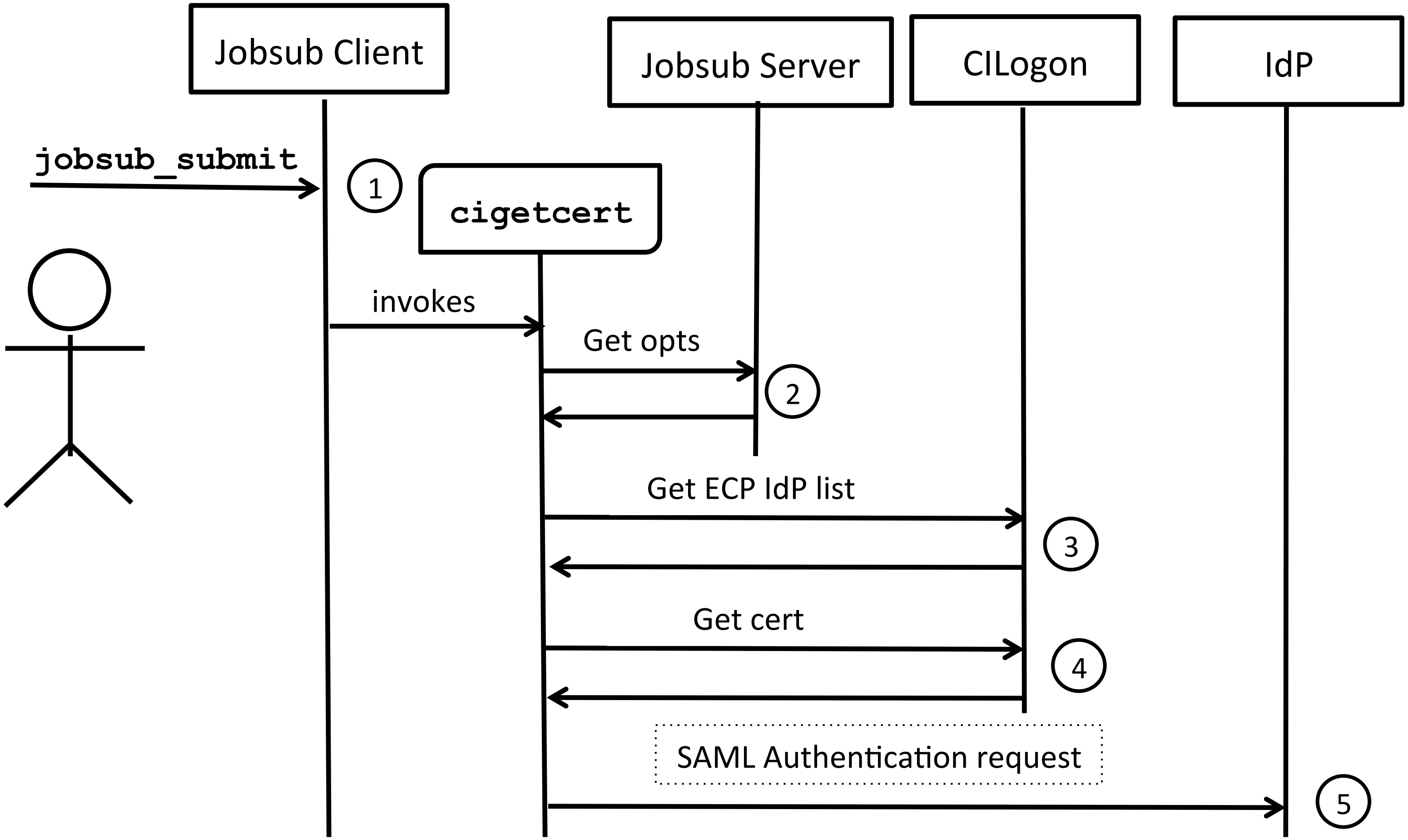
Basic grid+federated identity plan

- Make use of existing InCommon CILogon Basic CA and existing federated identity service
- Write new **`cigetcert`** command line tool to get certs
  - Generic tool, not Fermilab-specific
  - Authenticate with Kerberos or username/password
  - Get 4 week certificate from CILogon, store 1 week proxy on local disk and 4 week proxy in MyProxy, unencrypted
    - Complies with International Grid Trust Foundation (IGTF) rules
- Change **`jobsub_submit`** to attempt to use **`cigetcert`** with Kerberos, and if that fails, tell user to run it to enter "Services" password
  - Keep commands that prompt for password to minimum
- Change Jobsub server to renew proxies out of MyProxy
- Automatically register all new user DNs in VOMS (as old ones are)

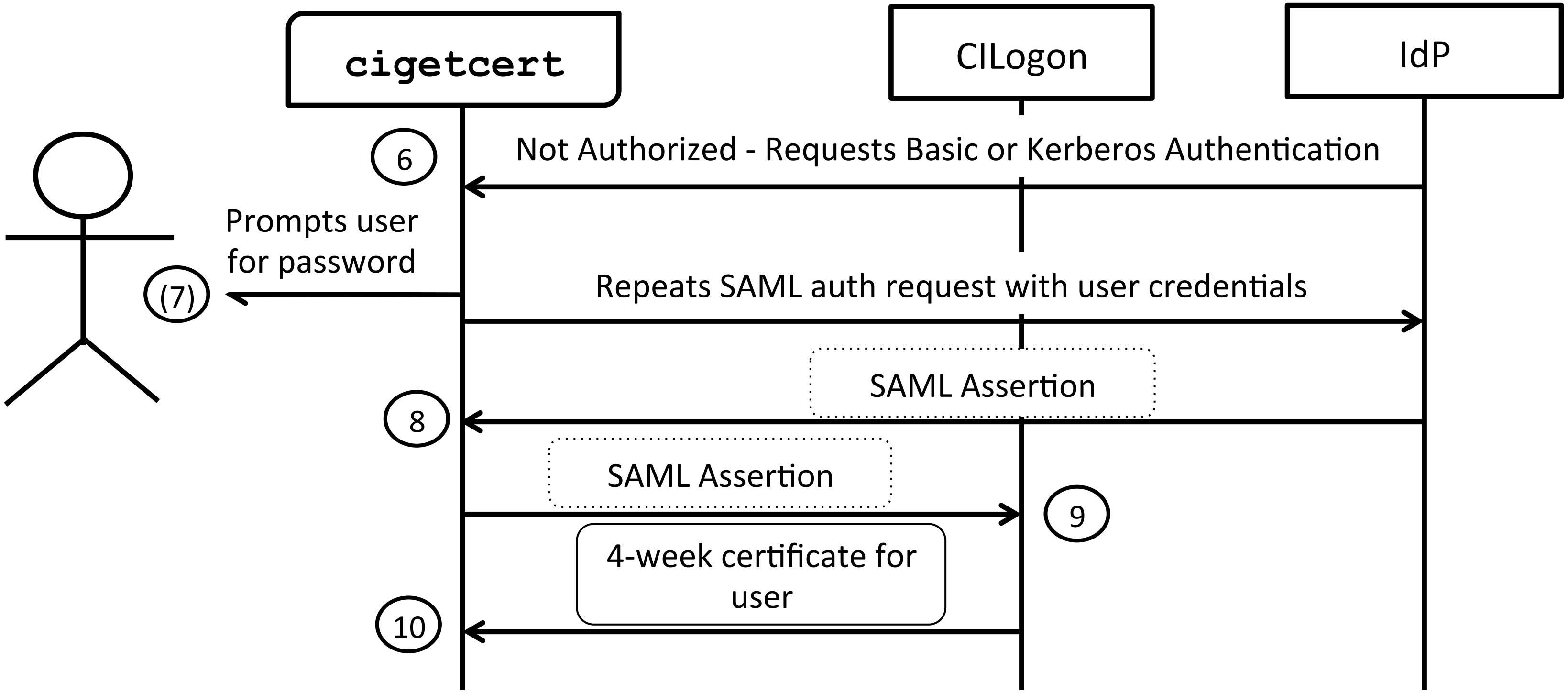
Jobsub infrastructure with CILogon



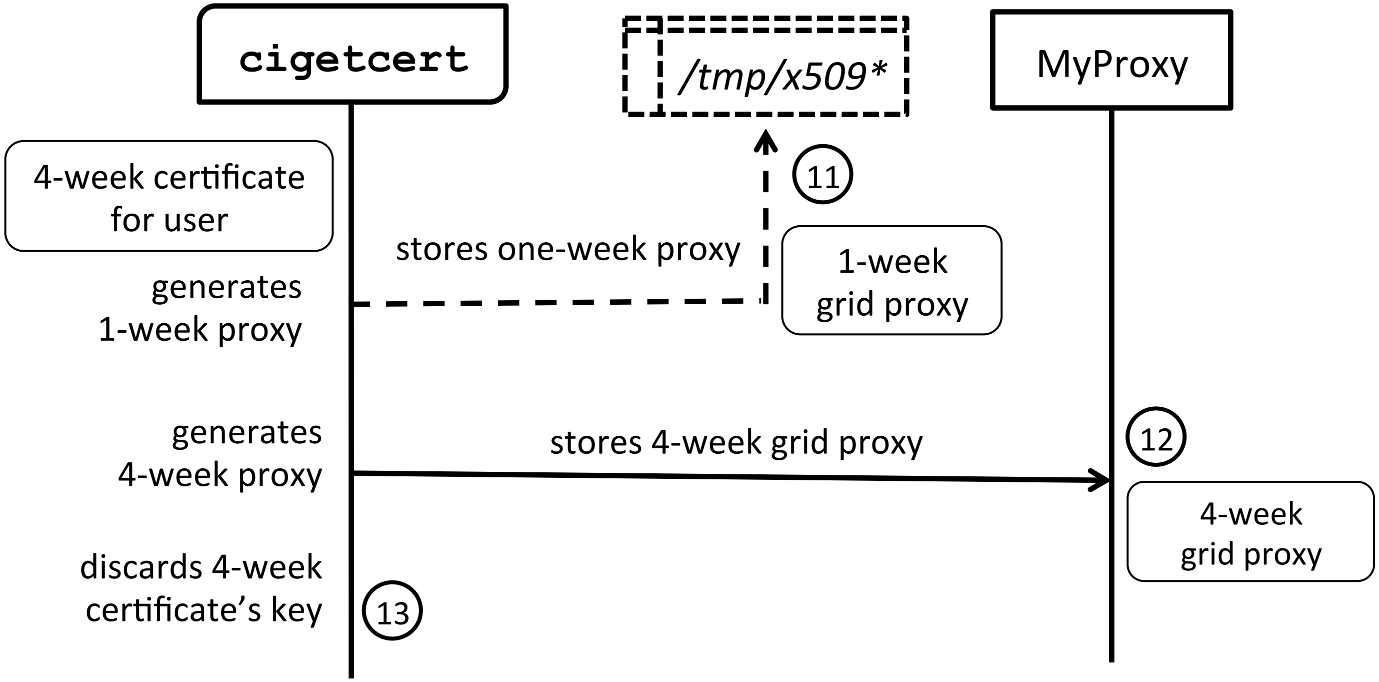
Startup


Jobsub Client
Jobsub Server
CILogon
IdP
jobsub_submit
1
cigetcert
invokes
Get opts
2
Get ECP IdP list
3
Get cert
4
SAML Authentication request
5

Getting a certificate


cigetcert
CILogon
IdP
6
Not Authorized - Requests Basic or Kerberos Authentication
Prompts user for password
(7)
Repeats SAML auth request with user credentials
SAML Assertion
8
SAML Assertion
9
4-week certificate for user
10

Storing proxies



cigetcert

/tmp/x509*

MyProxy

4-week certificate for user

generates 1-week proxy

stores one-week proxy

11

1-week grid proxy

generates 4-week proxy

stores 4-week grid proxy

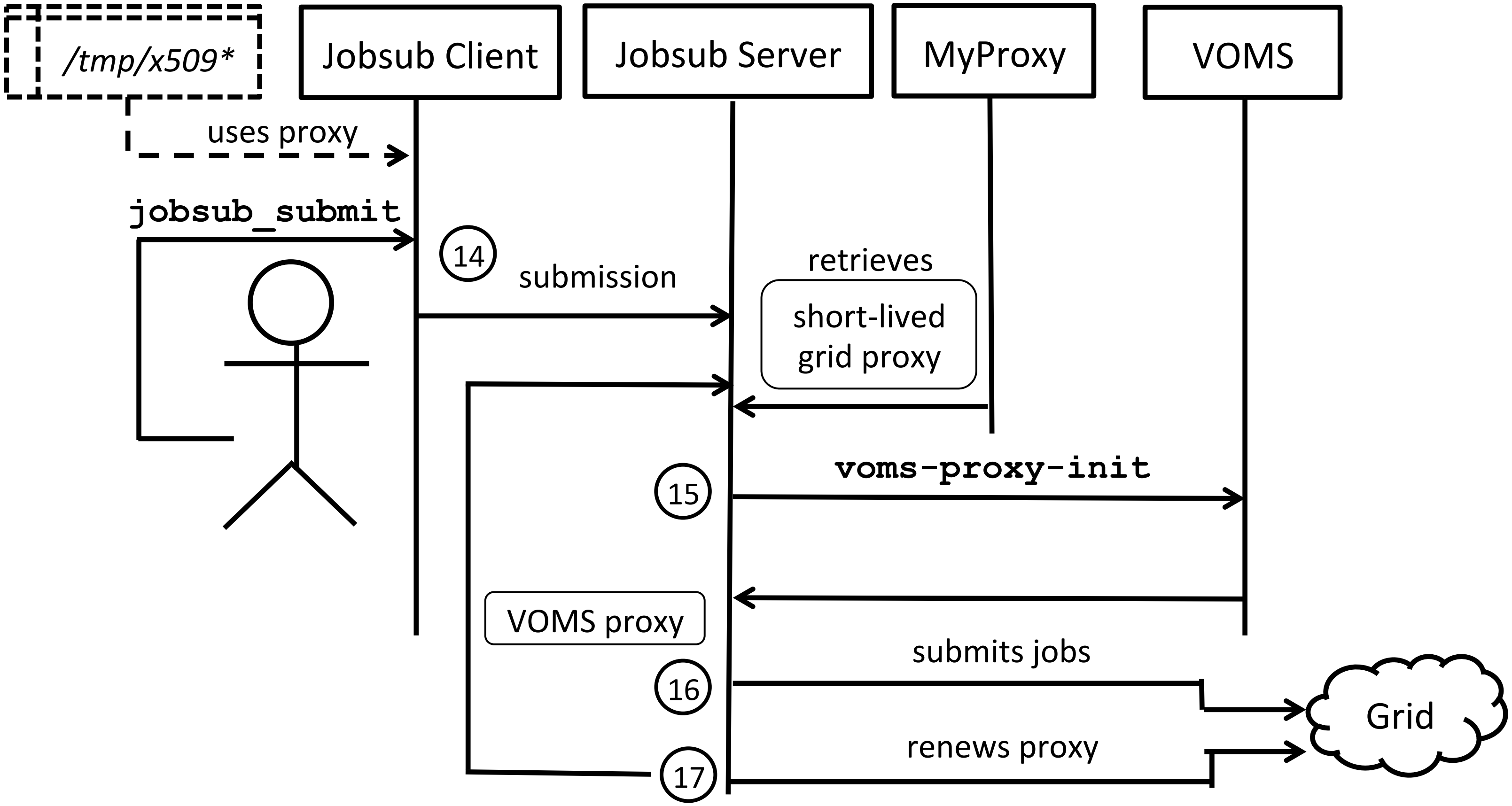
12

4-week grid proxy

discards 4-week certificate's key

13

Job submission & renewal



Status

- **cigetcert** reuses existing proxies if they still have some time until they expire, to lower CILogon/IdP load
- **cigetcert** is in production
  - Available in Scientific Linux Fermi
  - Could move into Scientific Linux if needed
- MyProxy and Jobsub changes also in production
- Most of 16 VOs transitioned, the remainder in the next two weeks
- Only Fermilab IdP supported this year
  - Phase 2 plans to add other institutions' IdPs
  - **cigetcer**t & Jobsub are ready for phase 2

Related work

- LIGO
  - Similar tool for getting a certificate with ECP
  - LIGO-specific, and without Kerberos or MyProxy support
- LTERN & DataOne
  - Use ECP, but little other published details
- ECP clients
  - https://wiki.shibboleth.net/confluence/display/CONCEPT/ECP

Security considerations

- Federated trust
  - Institutions are trusted, and verified by certs
  - If can't reach misbehaving user's institution, they can be cut off at VOMS and/or GUMS
- Limit number of command line tools that prompt for passwords
  - Don't want users to become callous about typing in their password

Bonus: pilot/payload isolation

- Pilot jobs run as an unprivileged user on worker nodes, and run payloads from different users
  - Without isolation, users could use pilot's certificate or other users' certificates, or modify pilot's logs
- The OSG's answer is to use **`glexec`**
  - Switches to separate user id based on certificate credentials
  - Setuid-root
  - Somewhat challenging to administer

Singularity

- OSG now experimenting with replacement tool **`singularity`** from LBL
  - Switches to isolated container-like namespace under same user id
  - Still setuid-root for now, but doesn't need to be on modern kernels
  - Even with setuid-root, easier on system administrators:
    - No separate user accounts to create
    - No Certifying Authority certs or CRLs to maintain
  - Becoming popular on supercomputers

Conclusions

- Certificate-free as far as user is concerned
- Easier on remote users – no need for Kerberos
- Easier on FNAL – no need for our own CA
- Easily expandable to other institutions' IdPs
- `cigetcert` available for general use with any institution that has an ECP-enabled IdP

Links

- cigetcert
  - https://github.com/fermitools/cigetcert
  - man page: https://git.io/vgcZm
- ECP
  - http://www.cilogon.org/ecp