



# Globus Workshop



## Introduction To Globus 5

### **GridKa School 2011**

**7. September 2011, Karlsruhe Germany**

Florian Zrenner ([zrenner@lrz.de](mailto:zrenner@lrz.de))

Adrian Colesa ([adrian.colesa@cs.utcluj.ro](mailto:adrian.colesa@cs.utcluj.ro))

# Overview

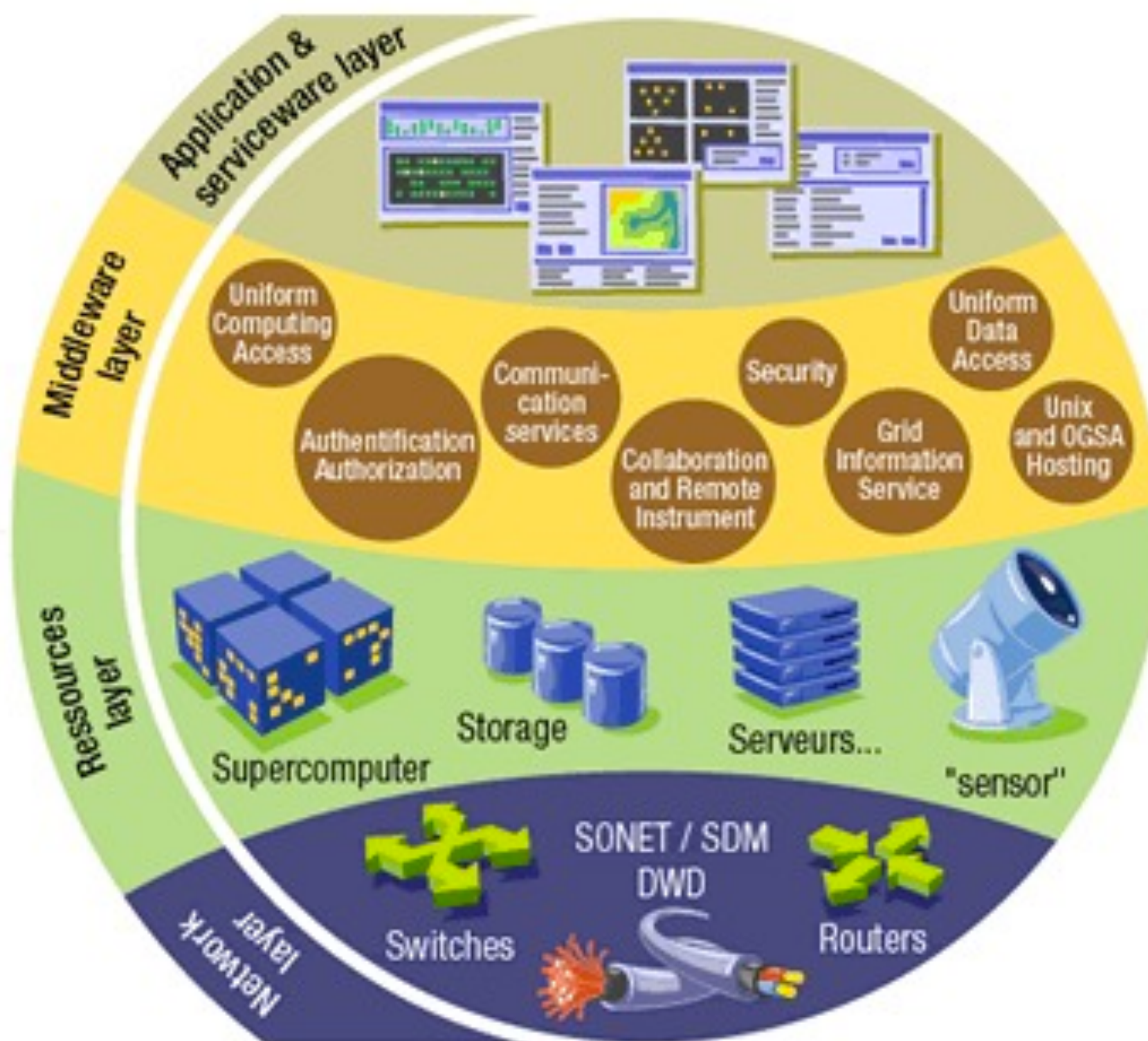
2

- **General**
  - Grid, Globus Toolkit Overview
  - Authentication & Authorisation (A&A)
    - PKI, Certificates
    - GSI, Proxy Certificates, MyProxy
- **GT5**
  - Interactive Access
    - GSI-OpenSSH
    - Clients
  - Job Submission
    - GRAM5
  - Data Transfer
    - Globus GridFTP
  - Data movement service
    - Globusonline

# Grid Architecture

# Grid Architecture

3



# It's About Grid Computing

4

- Resource sharing
  - Distributed computing
  - Computing sites
- Secure access
  - Trust between resource providers and users
- Grid-Software (Middleware): How to provide and access „Resources“
  - Abstraction layer

# Globus Organization 1/2

5

- Globus Alliance
  - International community to drive the development of Globus
- Globus Toolkit
  - Set of tools for building Grid systems and applications
  - Open source
  - Developed worldwide

# Globus Organization 2/2

6

- IGE (Initiative for Globus In Europe)
  - Coordination of European Globus activities
  - Introduce adjustments critical for Europe into Globus code base
  - Act as Globus service provider for European Grids like DEISA, PRACE, and EGI
  - Training, promotion, and documentation
  - Organize Globus Europe conference and Globus community forum
  - Bundle European input to Globus

# IGE Globus Repository

7

- Installation packages of Globus Toolkit provided by the IGE project
- How to use the repository
  - Import Key
  - Add local configuration (YUM/APT available)
  - Use your software management tool to install globus software
    - Instructions: <http://www.ige-project.eu/downloads>



# Overview

8

- **General**
  - Grid, Globus Toolkit Overview
  - Authentication & Authorisation (A&A)
    - PKI, Certificates
    - GSI, Proxy Certificates, MyProxy
- **GT5**
  - Interactive Access
    - GSI-OpenSSH
    - Clients
  - Job Submission
    - GRAM5
  - Data Transfer
    - Globus GridFTP
  - Data movement service
    - Globusonline

# Authentication And Authorisation (AA)

9

- Authentication is the process of
  - Verifying that s.b./s.th. is who he claims to be
  - Identify a user or a resource
- Authorisation
  - Give permission to perform certain operations or access specific resources

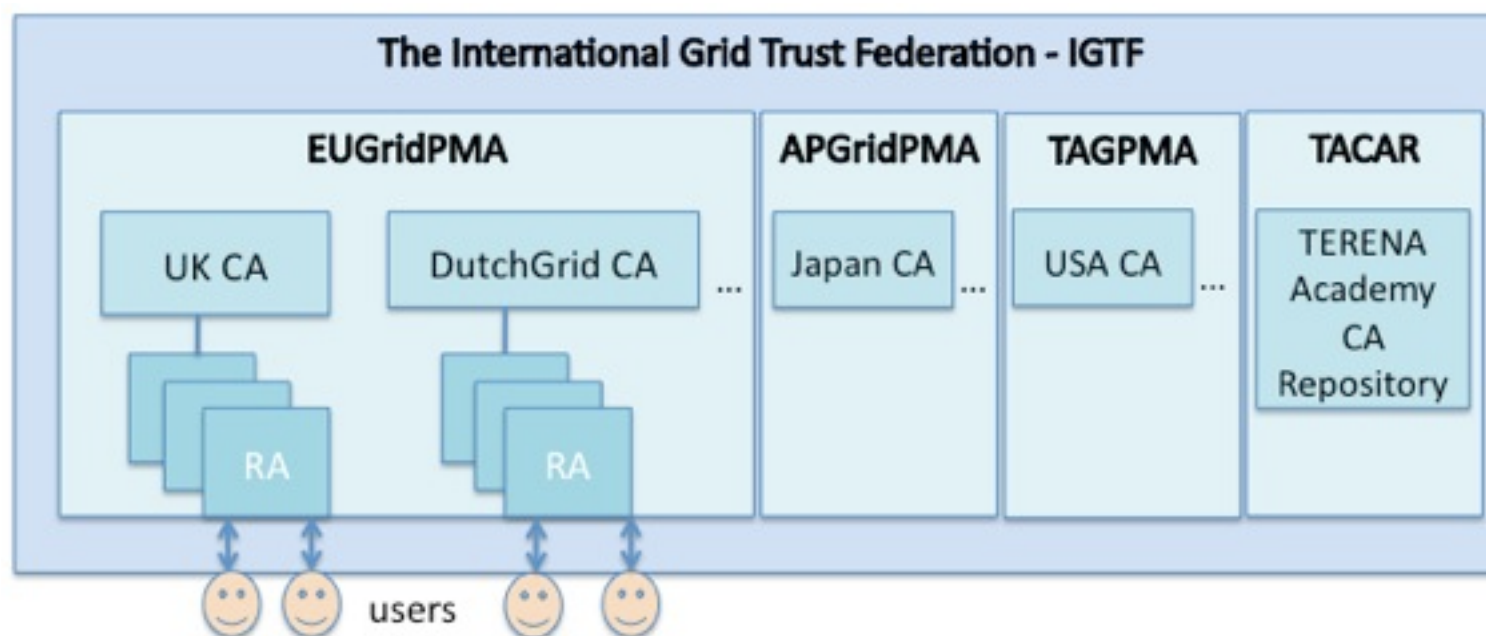
# Authentication With Certificates

10

- Based on Public Key Infrastructure (PKI)
- Certificates contain, e.g.:
  - „Distinguished Name – unique Grid id for user/service
    - Example: "/C=DE/O=GridGermany/OU=Leibniz/CN=Your Name"
  - Public Key
- Certificate Authority (CA)
  - Trusted 3rd party that confirms identity and issues certificate
  - Using a CA means you trust that this CA verified person/host after common rules
- Certificate Authority Certificates
  - Both sides must have CA certificates that they trust
  - Certificate is used to authorise user and resource provider against each other

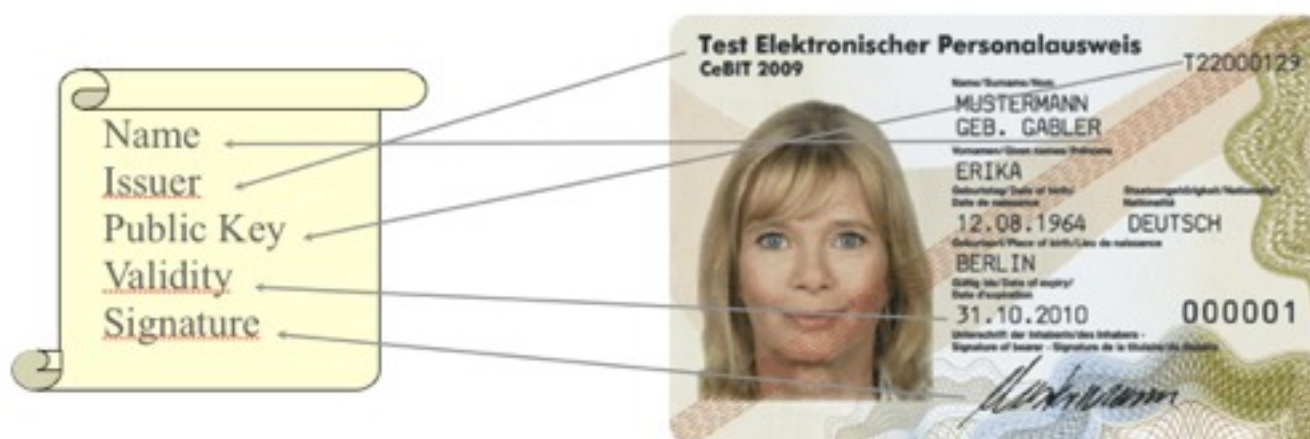
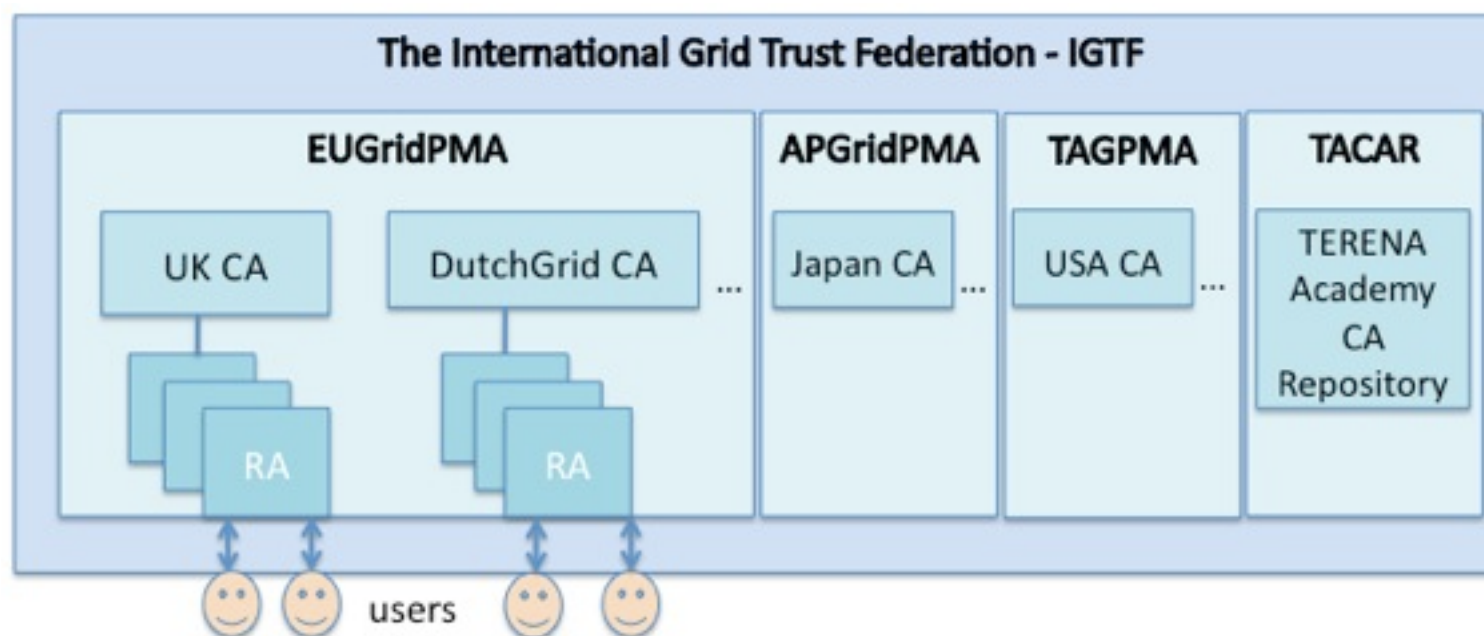
# Grid Security Infrastructure (GSI)

- Allows to identify a person to be authorized by a resource provider without previous communication



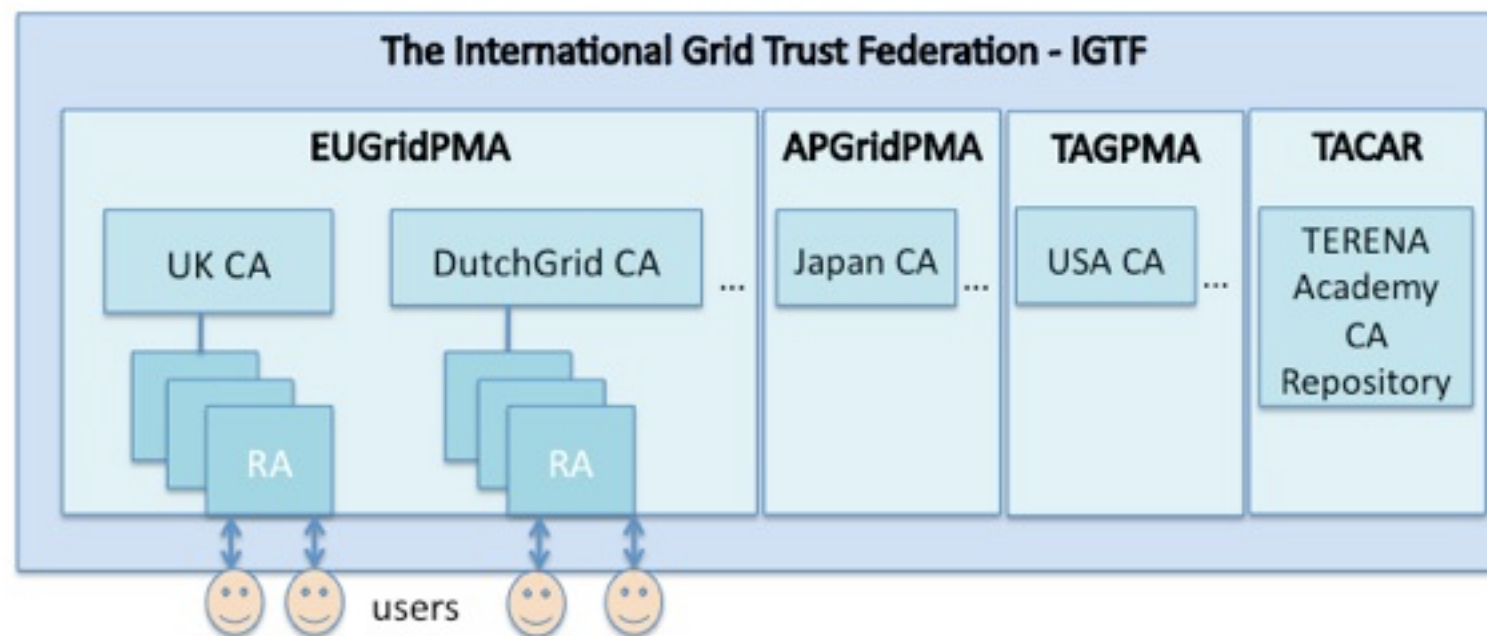
# Grid Security Infrastructure (GSI)

- Allows to identify a person to be authorized by a resource provider without previous communication



# Grid Security Infrastructure (GSI)

- Allows to identify a person to be authorized by a resource provider without previous communication

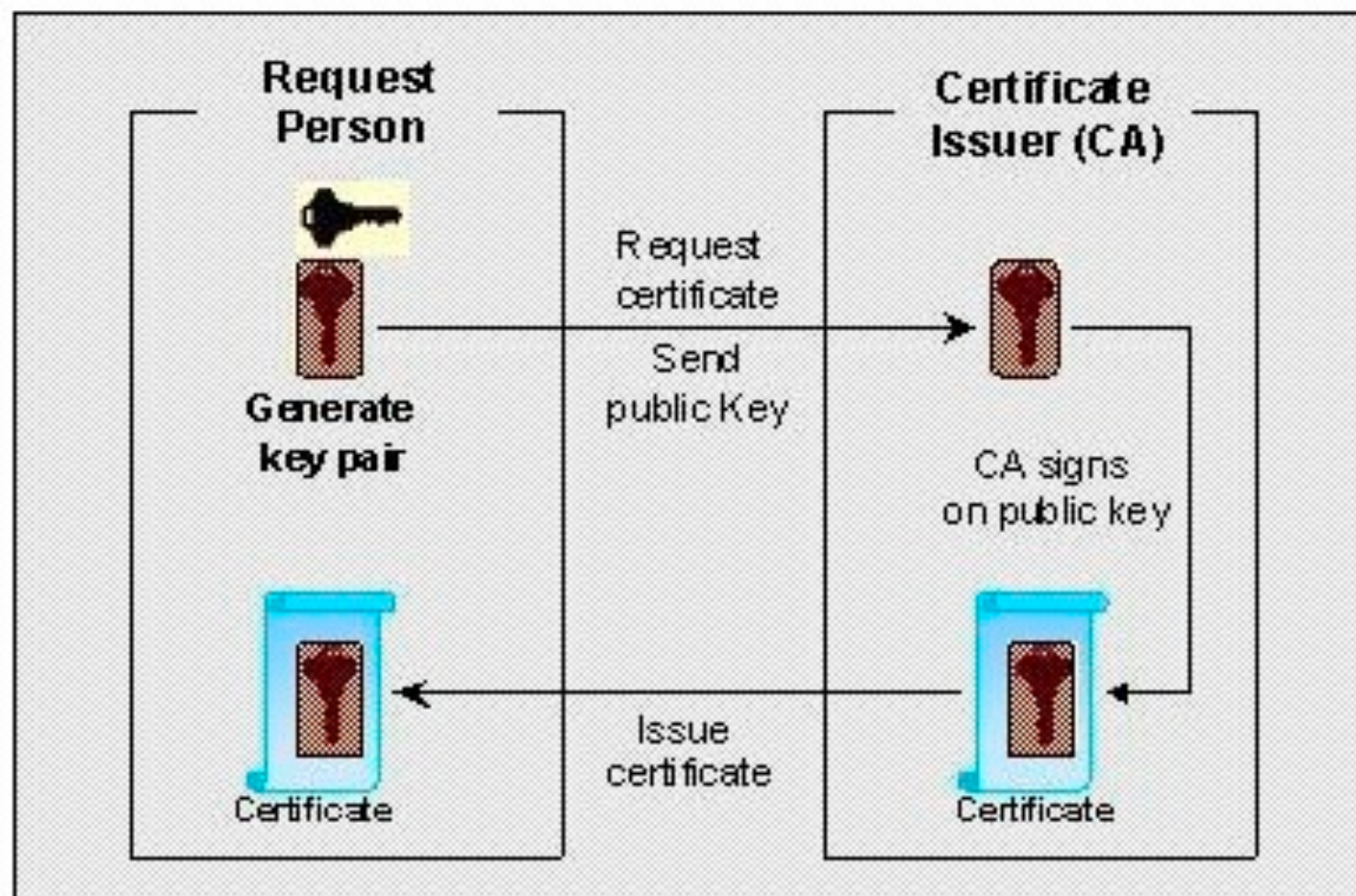


- Certificate Key must be secret: only store local
- If compromised: Revoke certificate immediately (Revoke Lists)



# Sign A Certificate

12



# Authorization With Certificates

13

- Grid user needs to be authorized to access remote site
- Authorization in GSI via grid-mapfile
  - Mapping of global DN to a local system account
- Format: Textfile (“DN“ local Account)  
Example entry:
  - “/C=DE/O=GridGermany/OU=Leibniz/CN=Your Name“ lrz28230
- All GSI-Services use grid-mapfile



# Certificate Based Access To Resources

14



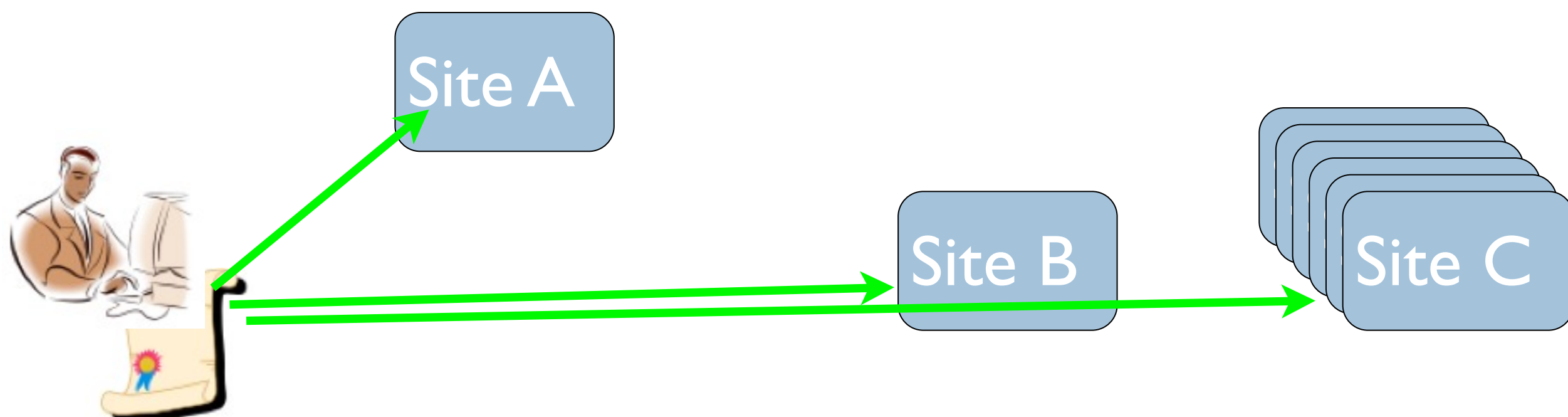
Site A

Site B

Site C

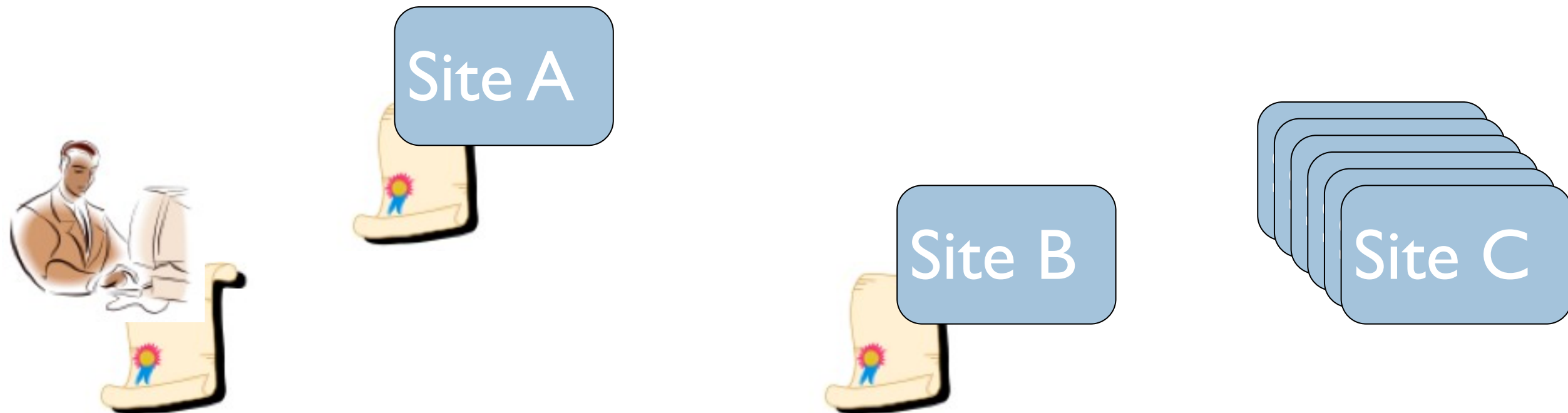
# Certificate Based Access To Resources

14



# Certificate Based Access To Resources

14



# Proxy Delegation

15

- Resource can delegate a proxy for its access to further resources (Delegation)



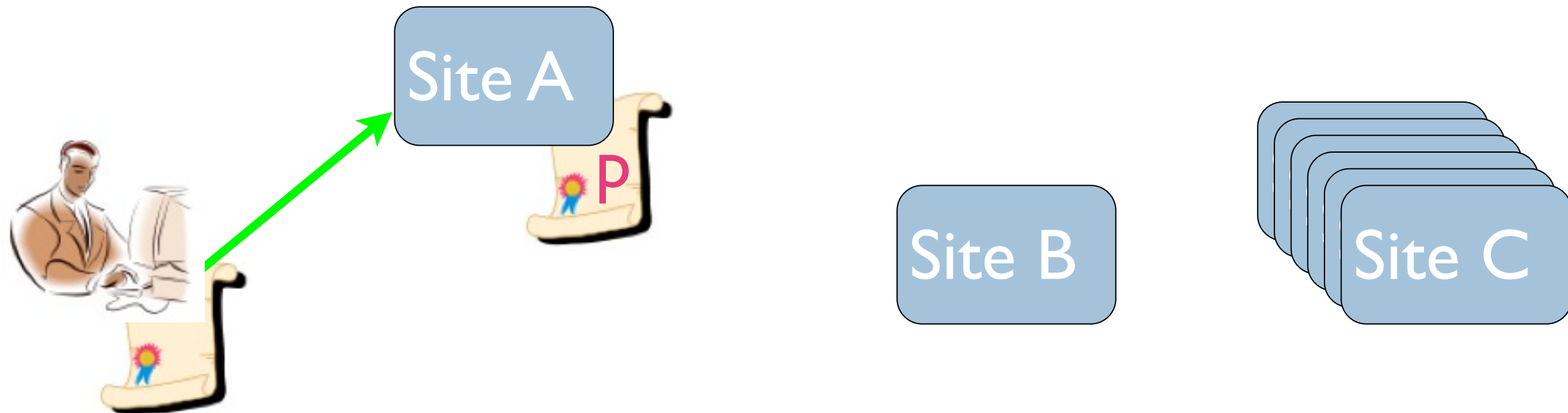
Site A

Site B

Site C

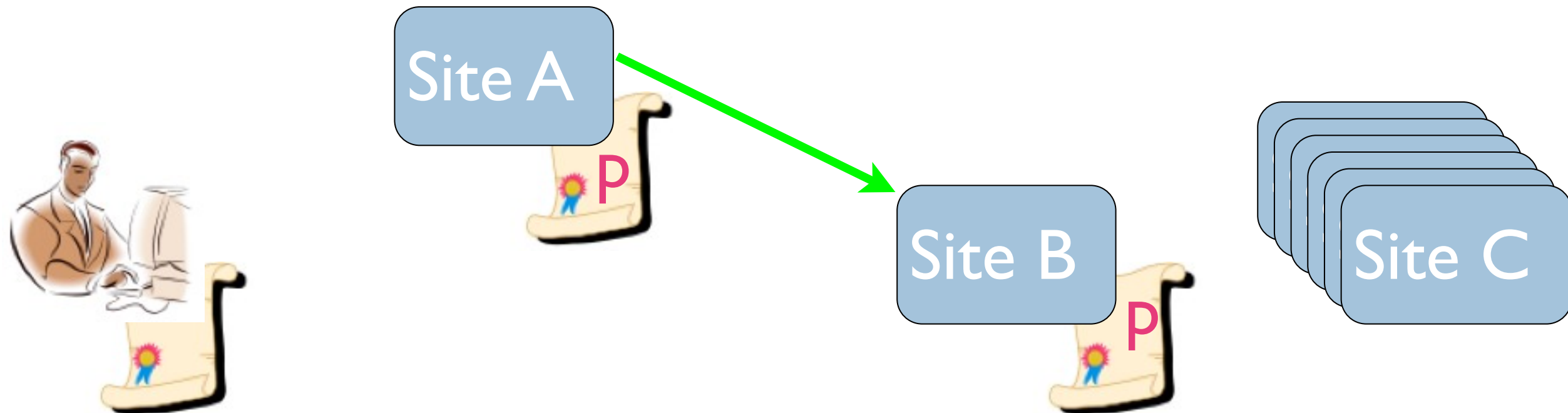
# Proxy Delegation

- Resource can delegate a proxy for its access to further resources (Delegation)



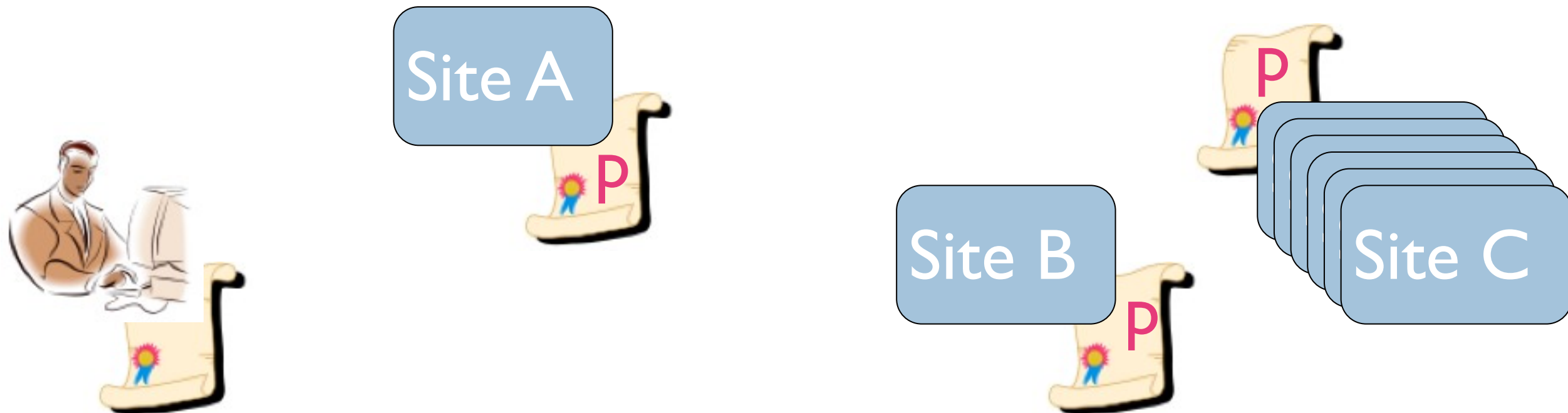
# Proxy Delegation

- Resource can delegate a proxy for its access to further resources (Delegation)



# Proxy Delegation

- Resource can delegate a proxy for its access to further resources (Delegation)



# Proxy Certificates

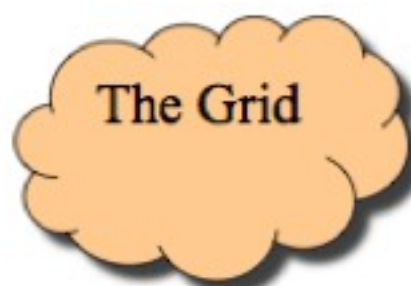
16

- Proxy Certificates
  - Backbone of trust delegation
  - Security gain:
    - No password has to be transmitted
    - Limited life time of the proxy certificate
    - Limited capability
  - Generated from user certificate key pair
    - Signed with your normal private key
  - Proxy certificate **consists** of
    - User normal public certificate
    - Newly generated proxy private key – without password
- Single sign-on: Login only once
  - Only type your password once (for your private key)
- Used by Globus services



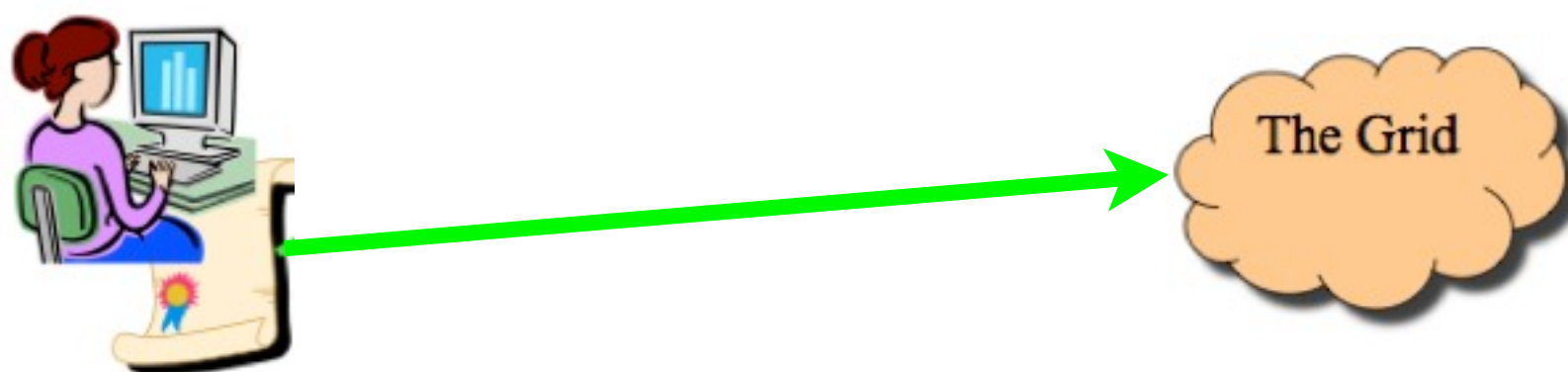
# MyProxy - Credential Repository

17



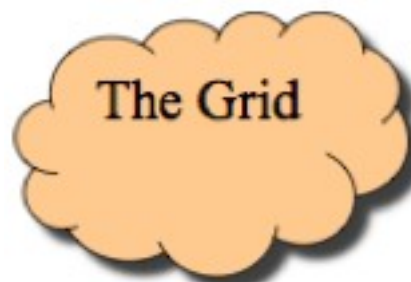
# MyProxy - Credential Repository

17



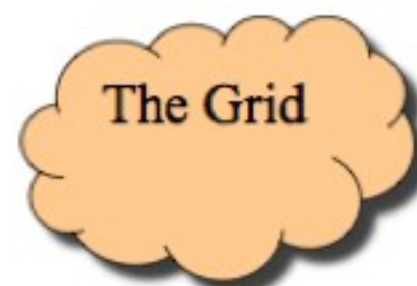
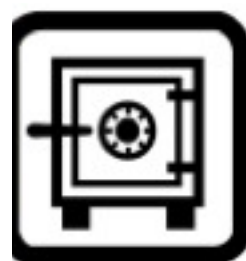
# MyProxy - Credential Repository

17



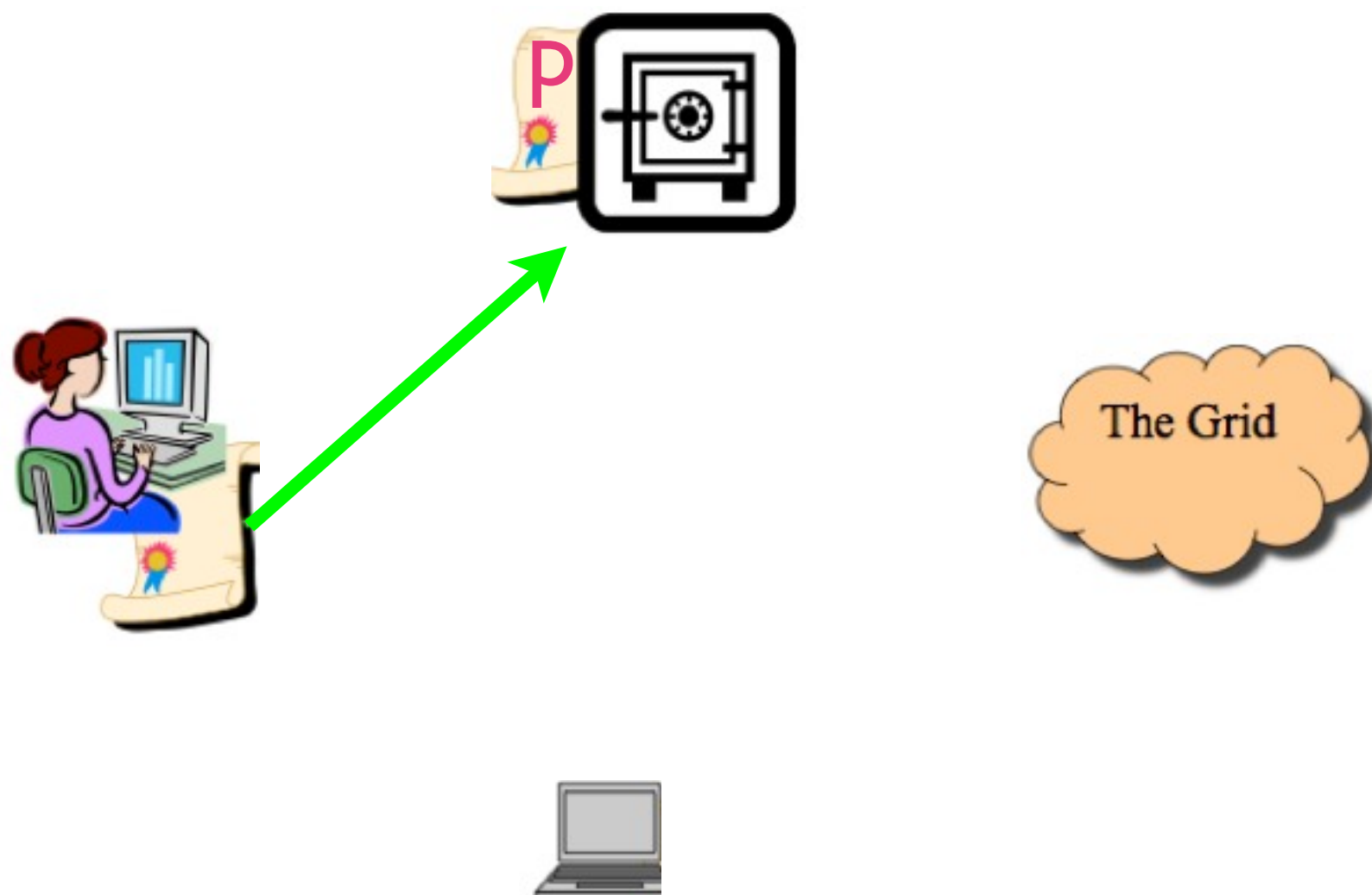
# MyProxy - Credential Repository

17



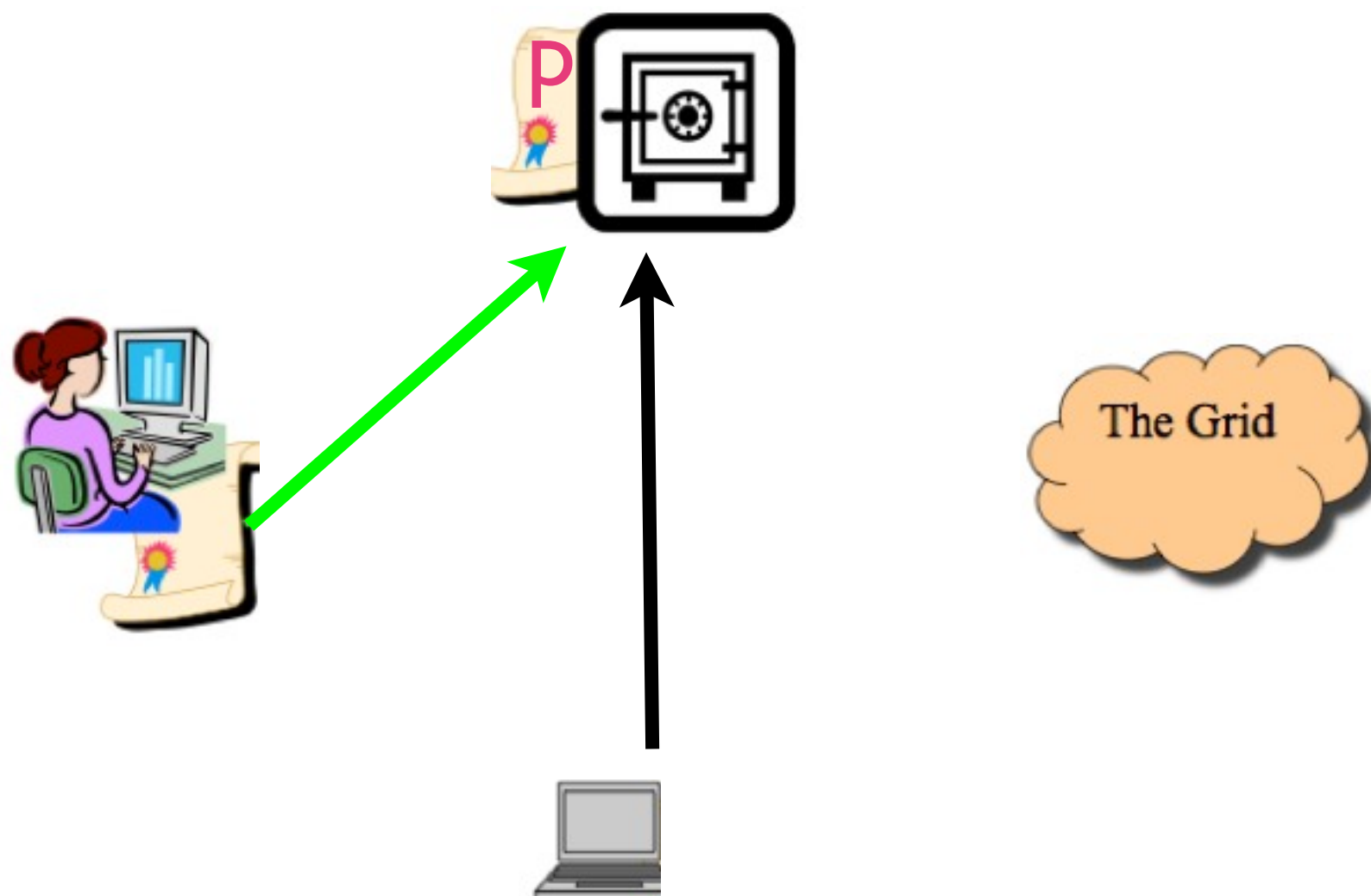
# MyProxy - Credential Repository

17



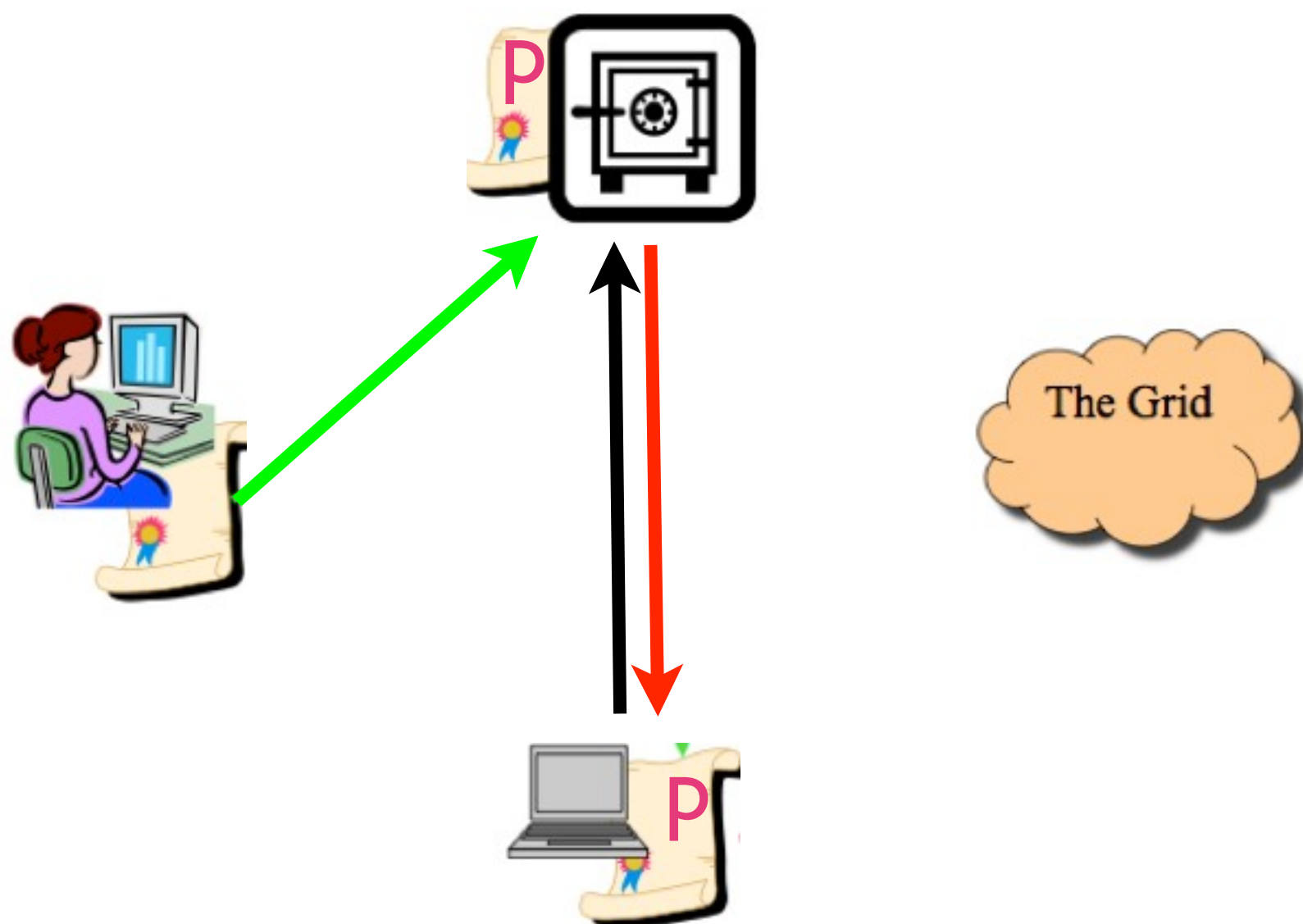
# MyProxy - Credential Repository

17



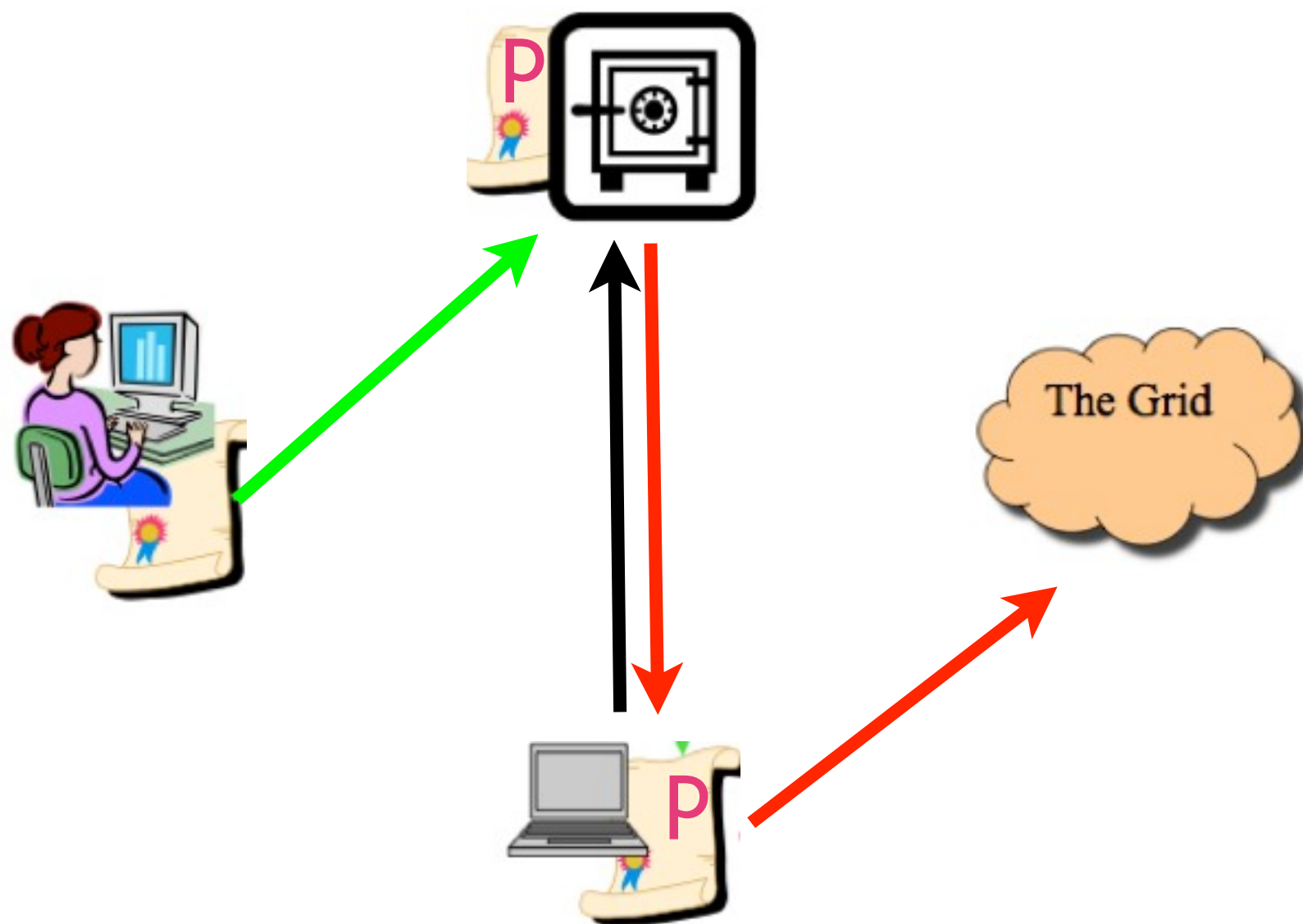
# MyProxy - Credential Repository

17



# MyProxy - Credential Repository

17





# Short Lived Credential Service (SLCS)

18

- Alternative for long-lived certificate
  - Without visiting a RA
- Institutes/companies already checked your id
- User is **authenticated by home institute** via web browser with username and password and gets a short-lived certificate
  - Only valid for a short period of time (e.g. one week)

# Short Lived Credential Service (SLCS)

18



SLCS  
(DFN)

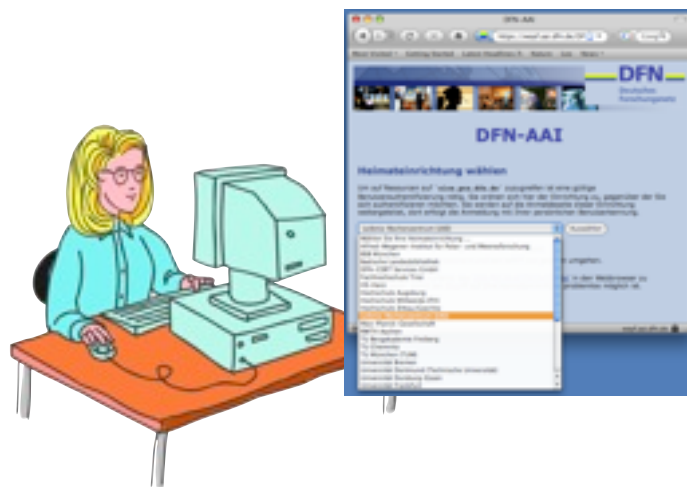
# Short Lived Credential Service (SLCS)

18

Identity Provider (RWTH)

Identity Provider (LRZ)

Identity Provider (MPG)



SLCS  
(DFN)

# Short Lived Credential Service (SLCS)

18

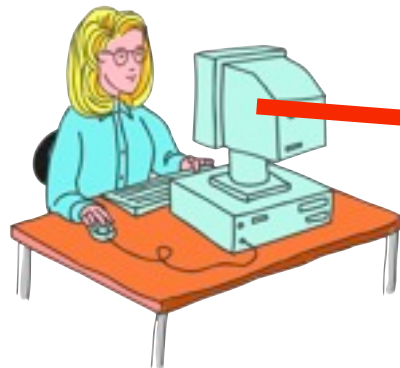
Identity Provider (RWTH)



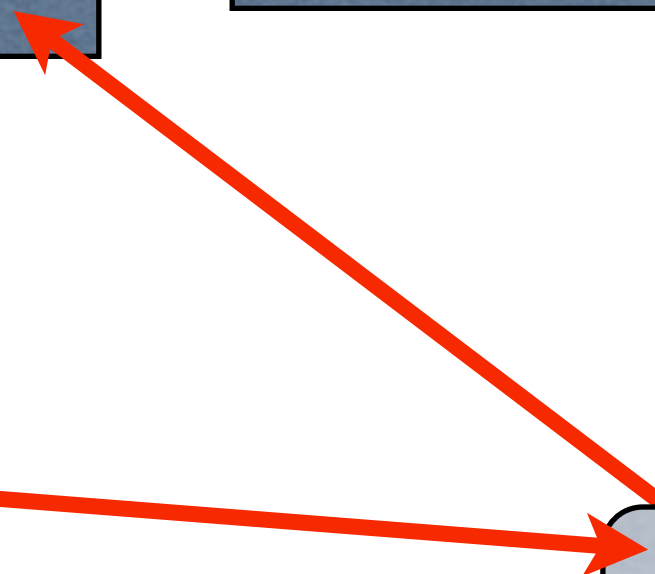
Shibboleth.

Identity Provider (LRZ)

Identity Provider (MPG)



SLCS  
(DFN)



# Short Lived Credential Service (SLCS)

18

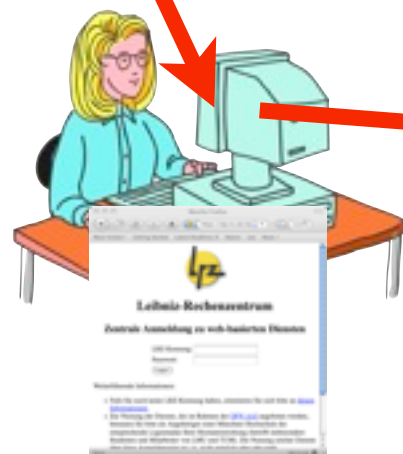
Identity Provider (RWTH)



Shibboleth.

Identity Provider (LRZ)

Identity Provider (MPG)



SLCS  
(DFN)



# Short Lived Credential Service (SLCS)

18

Identity Provider (RWTH)

Identity Provider (LRZ)

Identity Provider (MPG)



SLCS  
(DFN)

# Short Lived Credential Service (SLCS)

18

Identity Provider (RWTH)

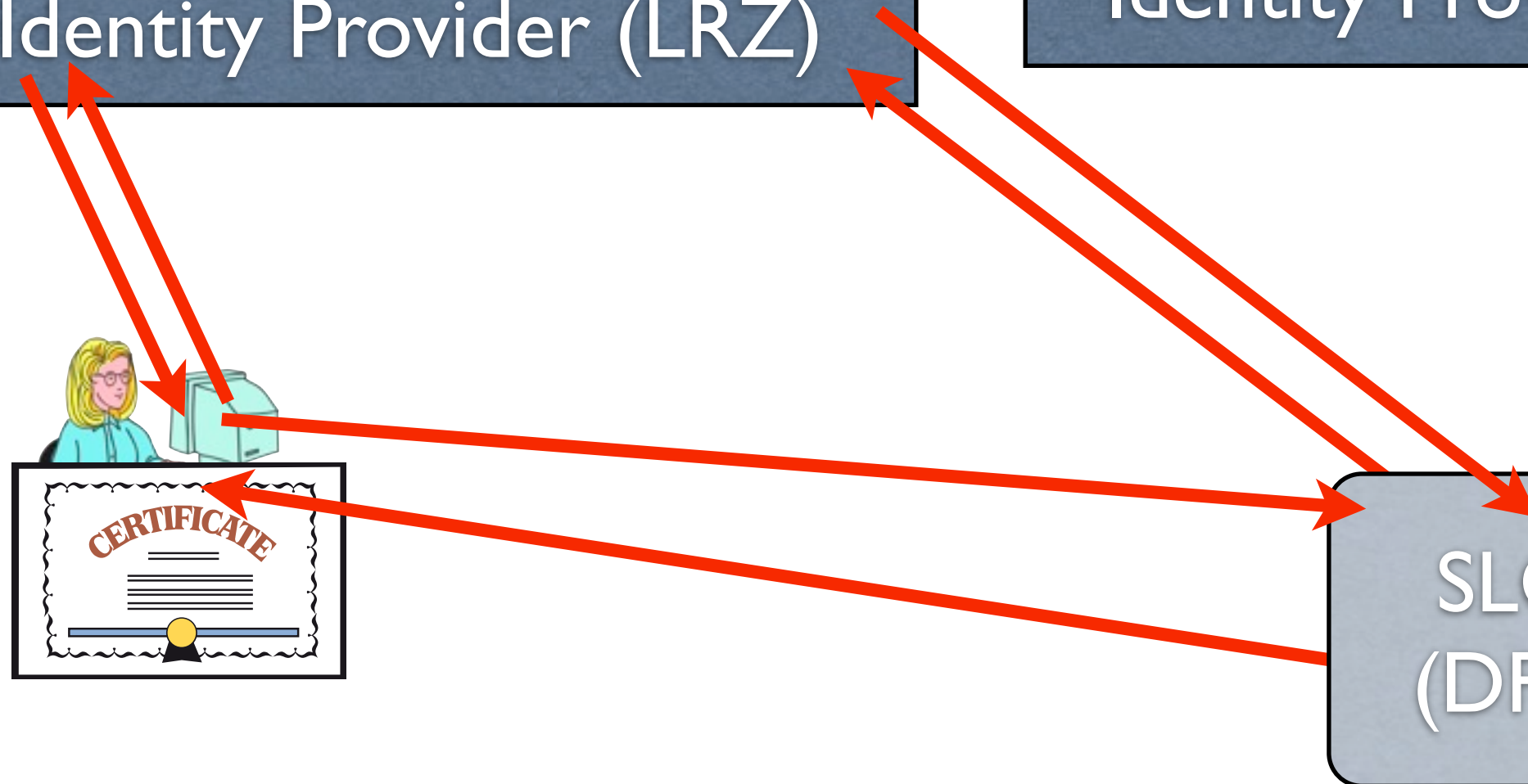
- <https://slcs.pca.dfn.de/gridshib-ca/>

Identity Provider (LRZ)

Identity Provider (MPG)



SLCS  
(DFN)



# Overview

19

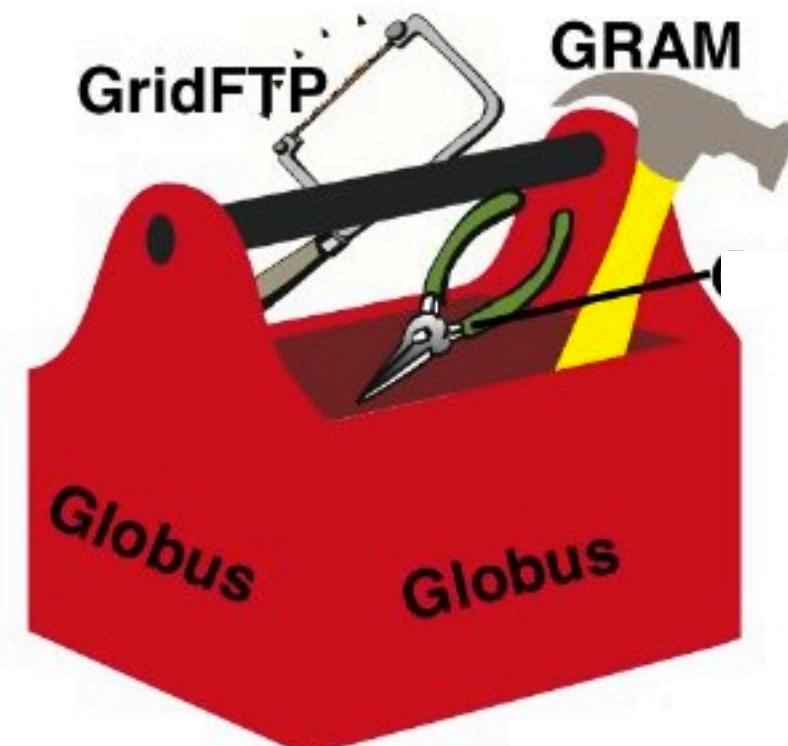
- **General**
  - Grid, Globus Toolkit Overview
  - Authentication & Authorisation (A&A)
    - PKI, Certificates
    - GSI, Proxy Certificates, MyProxy
- **GT5**
  - Interactive Access
    - GSI-OpenSSH
    - Clients
  - Job Submission
    - GRAM5
  - Data Transfer
    - Globus GridFTP
  - Data movement service
    - Globusonline



# Globus Toolkit

20

- **GSI**
  - The backbone
- **GSI-ssh**
  - Secure access
- **VOMS**
  - VO membership service
- **OGSA-DAI**
  - Data integration
- **GridFTP**
  - Super fast data transfer
- **GRAM**
  - Job submission framework



# What's New In GT5?

21

- Job submission is now GRAM5
  - Compatible with Globus v. 2.x. NOT with v. 4.x Web Services GRAM
- There is no Web Services interface (Java container) any more
- New GridFTP features
  - Resumeable file transfer
  - Compatible with older versions.  
No Reliable File Transfer (rft) anymore
- Globusonline
  - SAS file transfer
- GSI-SSH and MyProxy are compatible with older versions

# Interactive Access Overview

22

- GSI-OpenSSH
- Clients
- Login to a remote site

# GSI-Enabled OpenSSH Server

23

- GSI-OpenSSH is a modified version of OpenSSH
- Added support for GSI authentication and credential forwarding (delegation)
- Provides a single sign-on remote login

# Setup the GSI-Enabled OpenSSH Server

24

- Acquire a host certificate for the GSI-SSHD host.
- Authorise users you want to be able to connect with GSI SSH
  - grid-mapfile
- Configure and run the GSI-SSH daemon
  - Optional: Allowing only GSI authentication

# GSI-Enabled OpenSSH Clients

25

gsissh	gsissh-term
Native shell tool	Java (+ Java-Webstart)
Needs to install (a subset of) GT	No installation Easy use cross platforms
As your user interface	Java look and feel
Scriptable (Shell scripts)	Locked in

# Client: GSI-Enabled OpenSSH Client

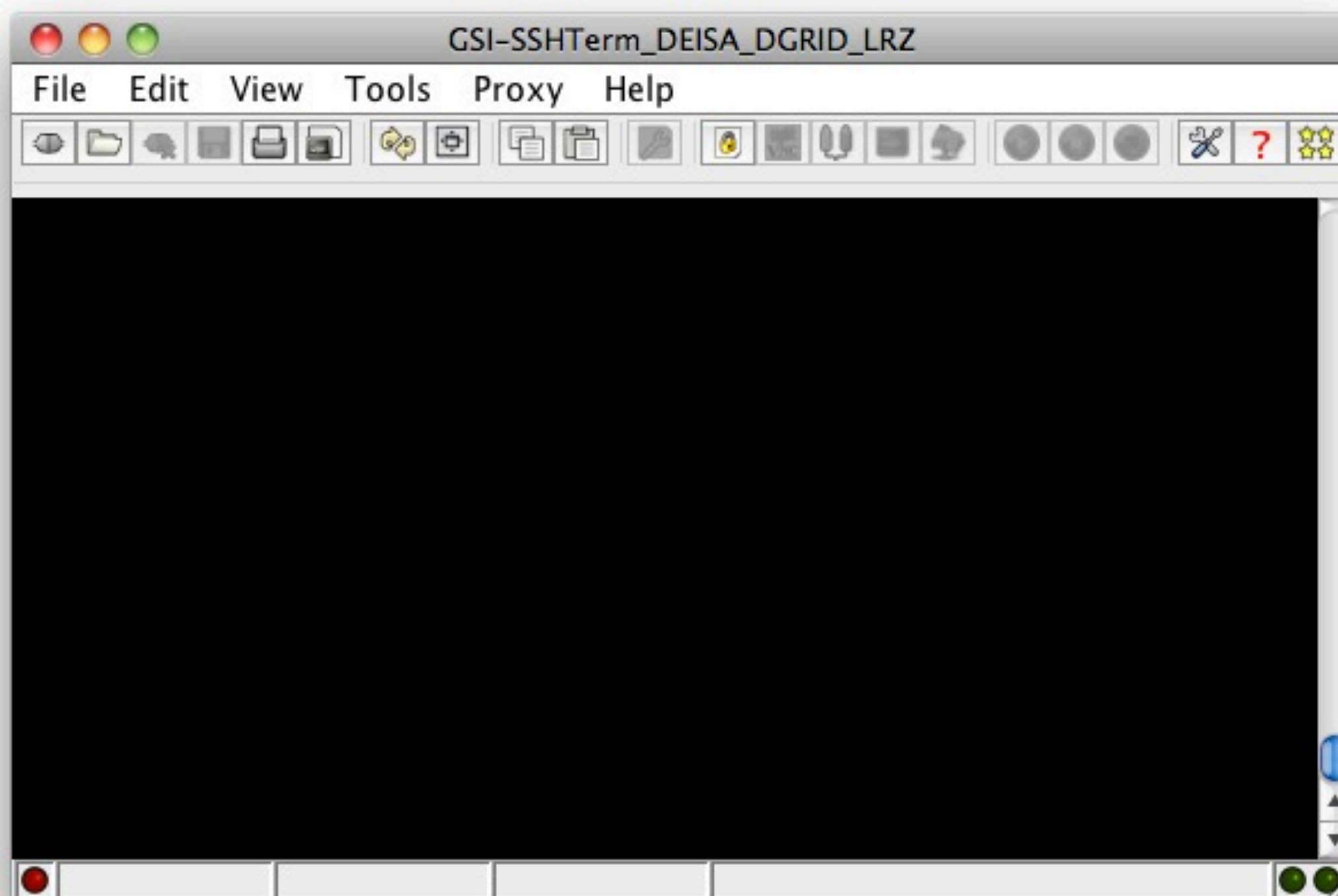
26

- Creating proxy credentials:
  - `grid-proxy-init`
- Information about your created proxy:
  - `grid-proxy-info`
- Login:
  - `gsissh host`
- Delete your proxy:
  - `grid-proxy-destroy`

# Client: GSISSH-Term

27

- GSISSH-Term login to HLRB-II (operated at LRZ):

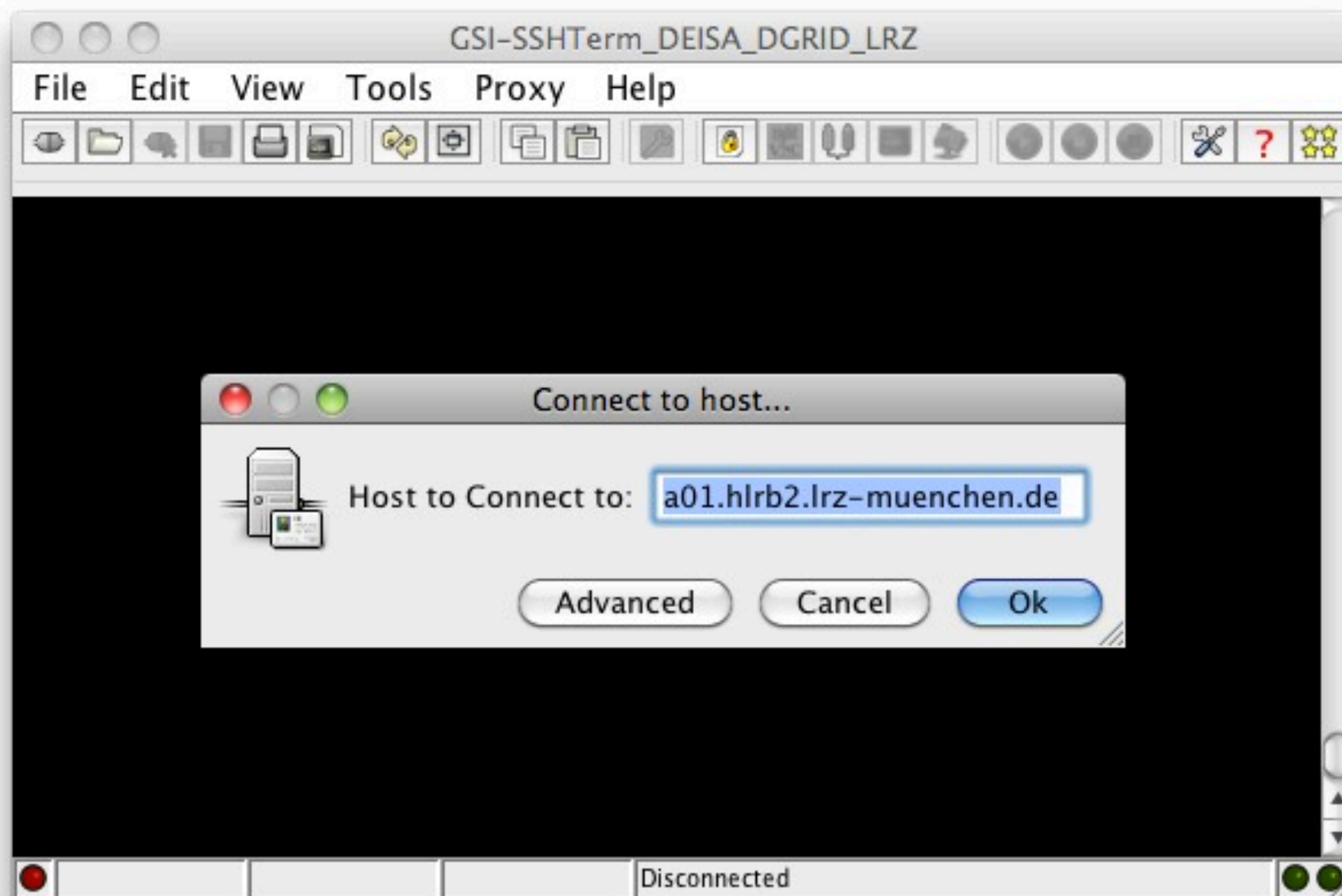




# Client: GSISSTerm

27

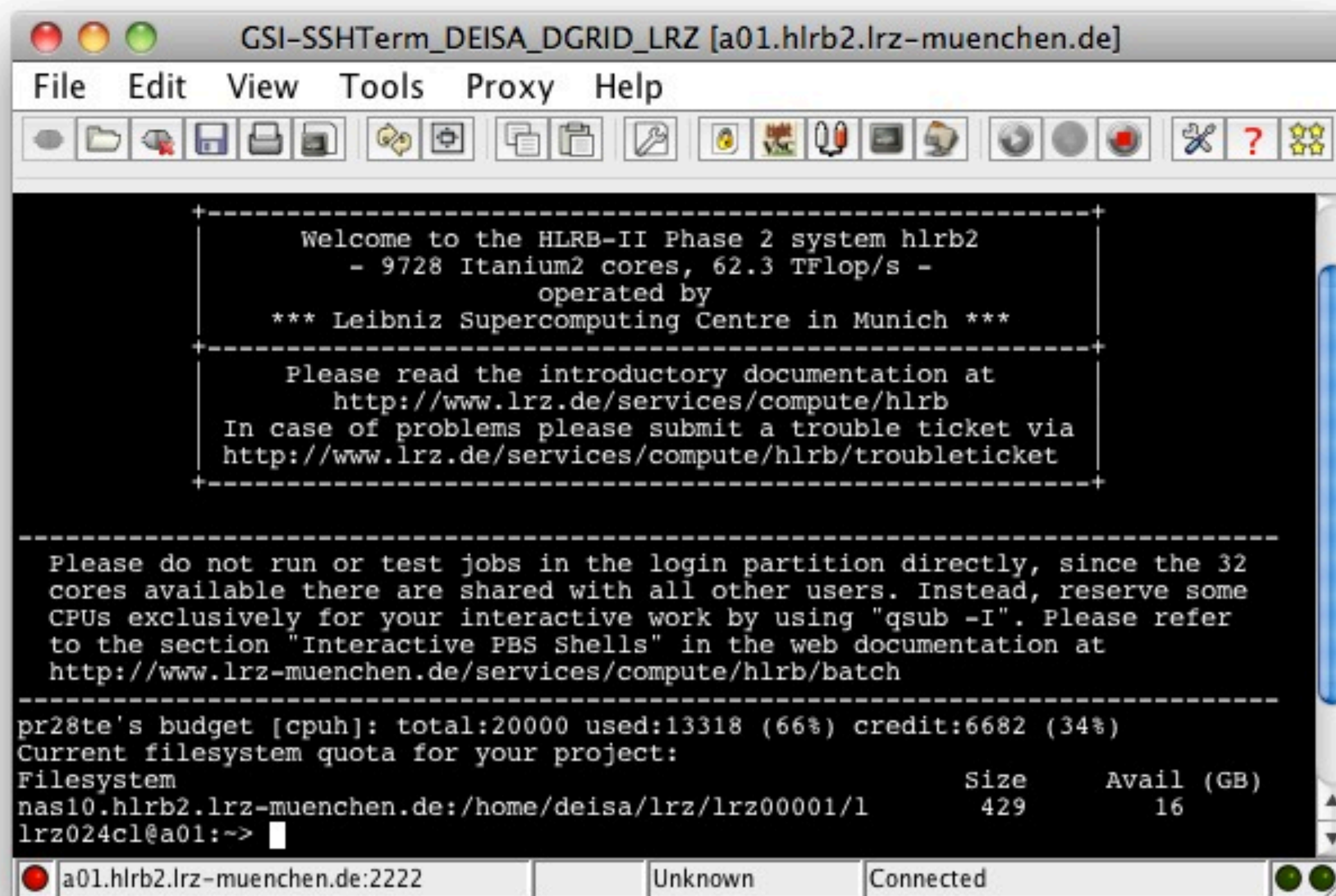
- GSISSTerm login to HLRB-II (operated at LRZ):



# Client: GSISSTerm

27

- GSISSTerm login to HLRB-II (operated at LRZ):



GSISSTerm\_DEISA\_DGRID\_LRZ [a01.hlrb2.lrz-muenchen.de]

File Edit View Tools Proxy Help

```

+-----+
|       Welcome to the HLRB-II Phase 2 system hlrb2       |
|       - 9728 Itanium2 cores, 62.3 TFlop/s -             |
|               operated by                               |
| *** Leibniz Supercomputing Centre in Munich ***         |
+-----+
| Please read the introductory documentation at           |
|   http://www.lrz.de/services/compute/hlrb               |
| In case of problems please submit a trouble ticket via |
|   http://www.lrz.de/services/compute/hlrb/troubleticket |
+-----+

Please do not run or test jobs in the login partition directly, since the 32
cores available there are shared with all other users. Instead, reserve some
CPUs exclusively for your interactive work by using "qsub -I". Please refer
to the section "Interactive PBS Shells" in the web documentation at
http://www.lrz-muenchen.de/services/compute/hlrb/batch

-----
pr28te's budget [cpuh]: total:20000 used:13318 (66%) credit:6682 (34%)
Current filesystem quota for your project:
Filesystem                               Size      Avail (GB)
nas10.hlrb2.lrz-muenchen.de:/home/deisa/lrz/lrz00001/1  429        16
lrz024cl@a01:~>
  
```

a01.hlrb2.lrz-muenchen.de:2222      Unknown      Connected

# Overview

28

- **General**
  - Grid, Globus Toolkit Overview
  - Authentication & Authorisation (A&A)
    - PKI, Certificates
    - GSI, Proxy Certificates, MyProxy
- **GT5**
  - Interactive Access
    - GSI-OpenSSH
    - Clients
  - Job Submission
    - GRAM5
  - Data Transfer
    - Globus GridFTP
  - Data movement service
    - Globusonline

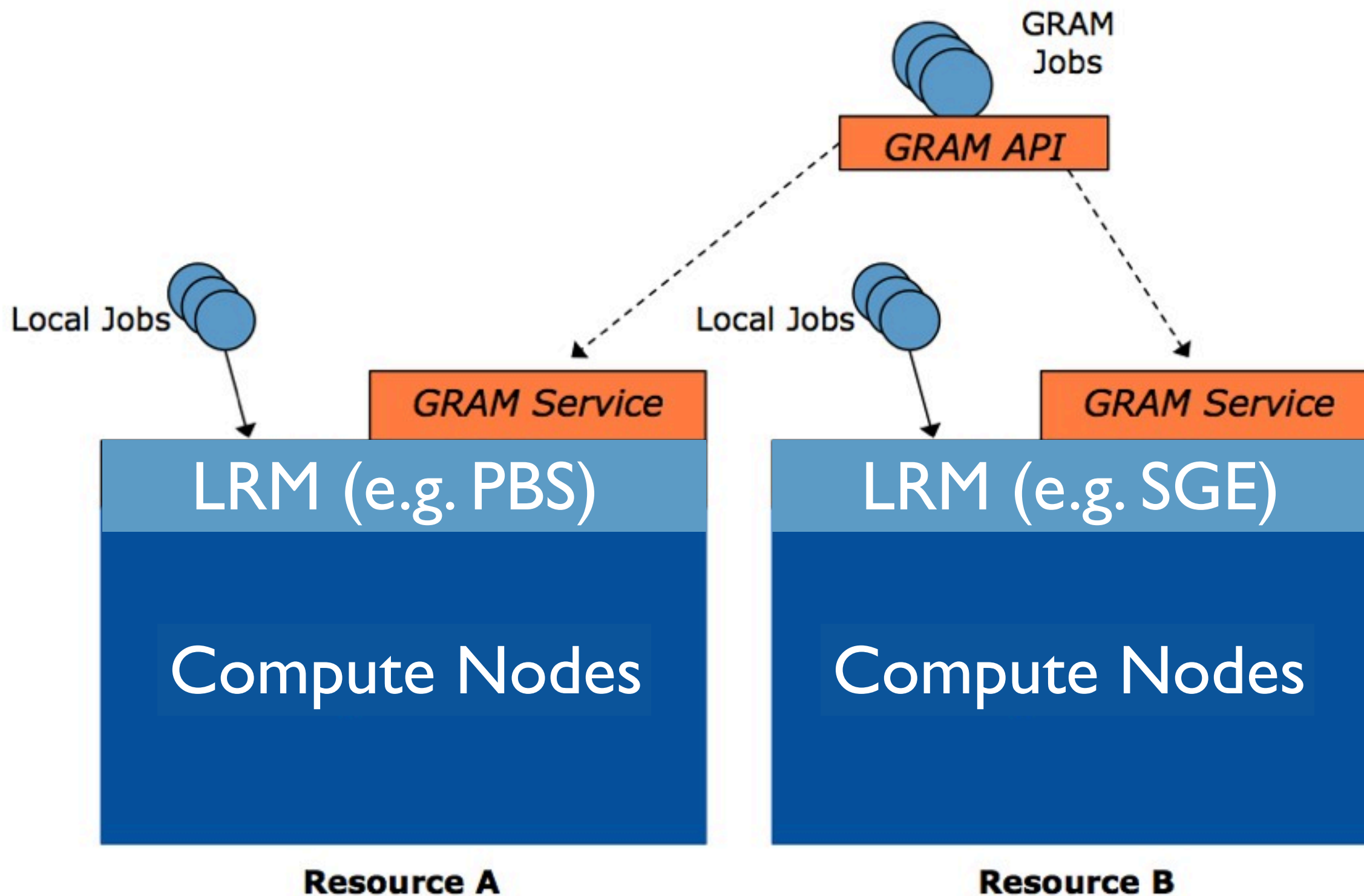
# Grid Job Management Goals

29

- User sends his Job via middleware
- Grid Middleware provides common interface for different Local Resource Management Systems (LRMS)
- **Functionality**
  - Certificate based A&A
  - Stage files to/from resource
  - Initiate execution of job process(es)
  - Monitor execution
  - Signal important state changes to client

# GRAM On A Site

30



# Job Execution Management

31

- Globus Resource Allocation Manager (GRAM5)
  - GRAM is a Globus Toolkit component for grid job submission
  - Interfaces to many batch systems:
    - PBS/Torque, LSF, SGE
    - LoadLeveler (LL) – IGE contribution!
- GRAM is a unifying remote interface to Resource Managers
- GRAM provides stateful job control
  - Asynchronous monitoring and control
  - Remote credential management
  - Remote file staging and file cleanup



# GRAM5: Interfacing The System

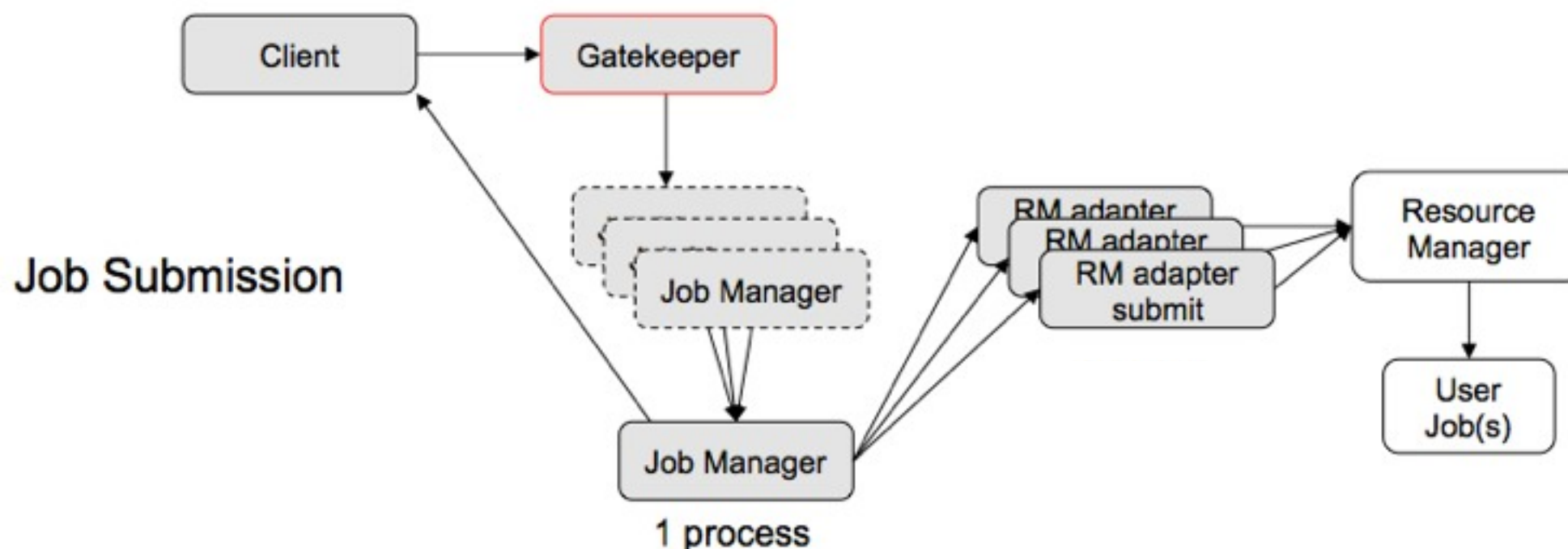
32

- User submits a Globus job using
  - Globus client commands
  - optionally a job script (in Resource Specification Language (RSL))
  - submitting returns a „contact“-URL for his job
- User interaction via „Contact“-URL
  - query the job status
  - cancel the job
- Globus Server
  - Translate your job script for the specific LRMS
  - Use native LRMS commands to submit the job
  - Globus will check the job state using LRMS log file

# GRAM5 Components

33

- **Gatekeeper**
  - Authentication
  - Starts job management service (on request)
- **Job Manager**
  - Processes job requests and coordinates file transfer
  - One process per user per LRSM
- **Job Manager Script (RM adapter submit)**
  - Interacts with LRMS and does the file transfer





# GRAM Client Interfaces

34

- Globus's (job related) command line commands:
  - `globus-job-submit`
  - `globus-job-status`
  - `globus-job-get-output`
  - `globus-job-clean`
- Application Programming Interface (API) for C and JAVA

# Overview

35

- General
  - Grid, Globus Toolkit Overview
  - Authentication & Authorisation (A&A)
    - PKI, Certificates
    - GSI, Proxy Certificates, MyProxy
- GT5
  - Interactive Access
    - GSI-OpenSSH
    - Clients
  - Job Submission
    - GRAM5
  - Data Transfer
    - Globus GridFTP
  - Data movement service
    - Globusonline

# What is GridFTP?

36

- High-performance, reliable data transfer protocol optimized for high-bandwidth wide area networks
- Based on FTP protocol - defines extensions for high-performance operation and security
  - Authenticate control and data channels with GSI
- Standardized through Open Grid Forum (OGF)
- GridFTP is the OGF recommended data movement protocol

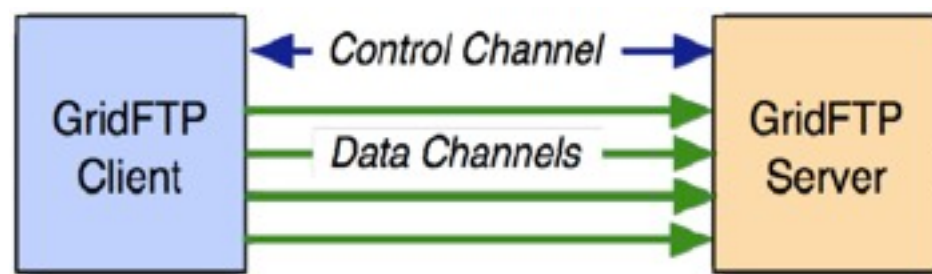
# Understanding GridFTP

37

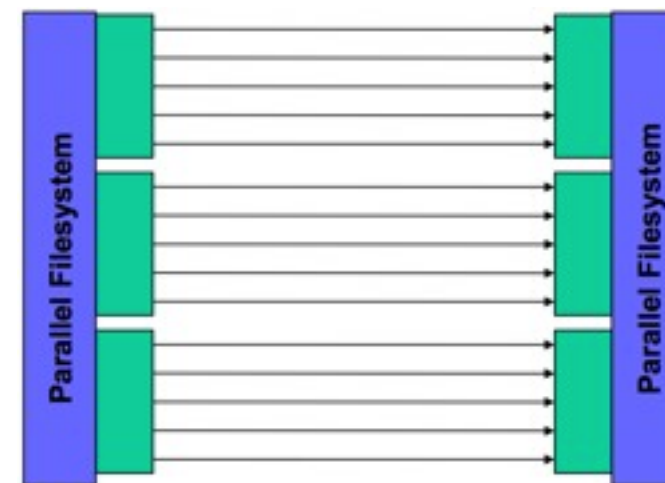
- Two channel protocol like FTP
- Control Channel
  - Command/Response
  - Used to establish data channels
  - Basic file system operations eg. mkdir, delete etc
- Data channel
  - Pathway where file is transferred

# GridFTP Performance Options

- Adjustable buffer size of data channels (-tcp-bs)
- Parallel TCP streams (-p)



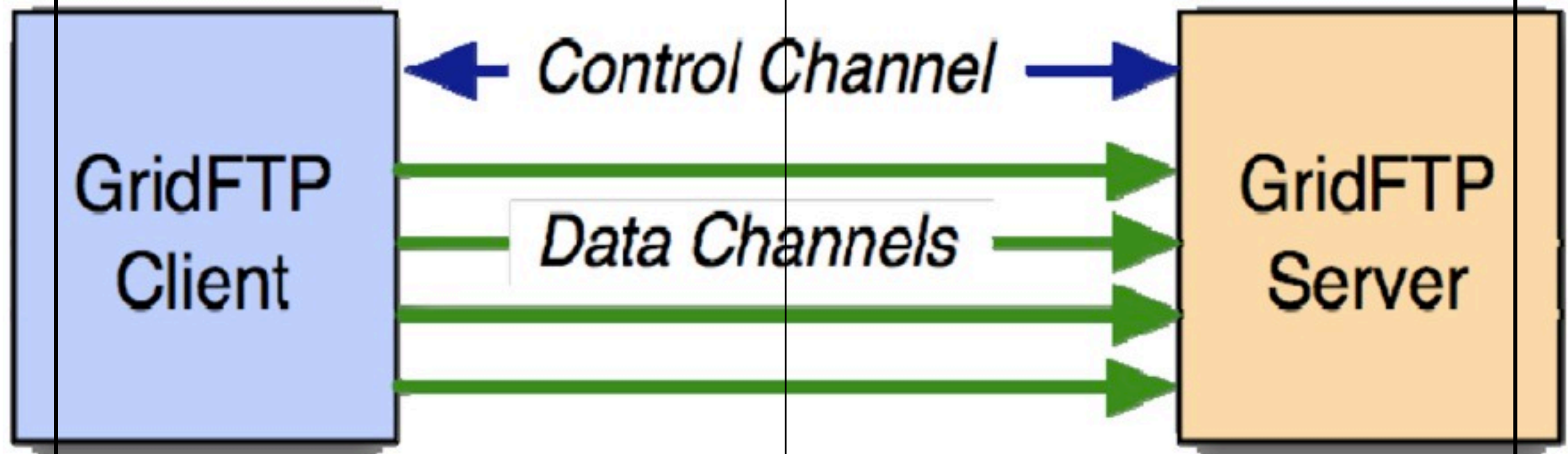
- Striped GridFTP
  - multiple network endpoints for the transfer of the same file



# GridFTP Performance Options

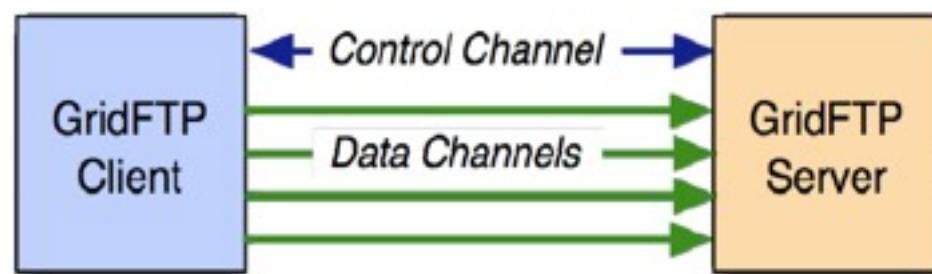
38

- Adjustable buffer size of data channels (-tcp-bs)
- Parallel TCP streams (-p)
- Striped GridFTP
  - multiple network endpoints for the transfer of the same file

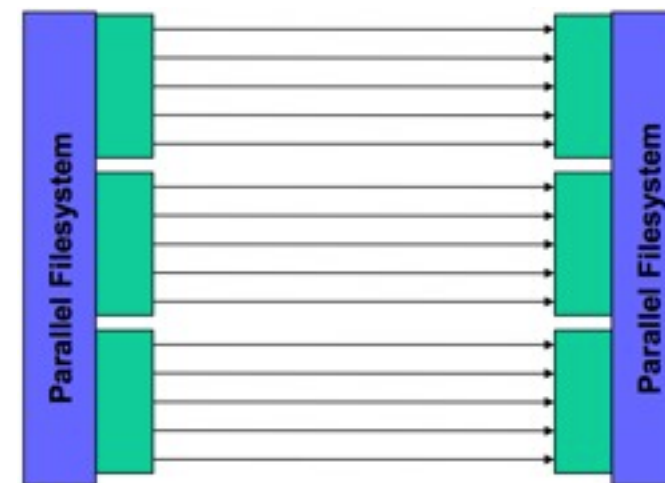


# GridFTP Performance Options

- Adjustable buffer size of data channels (-tcp-bs)
- Parallel TCP streams (-p)



- Striped GridFTP
  - multiple network endpoints for the transfer of the same file



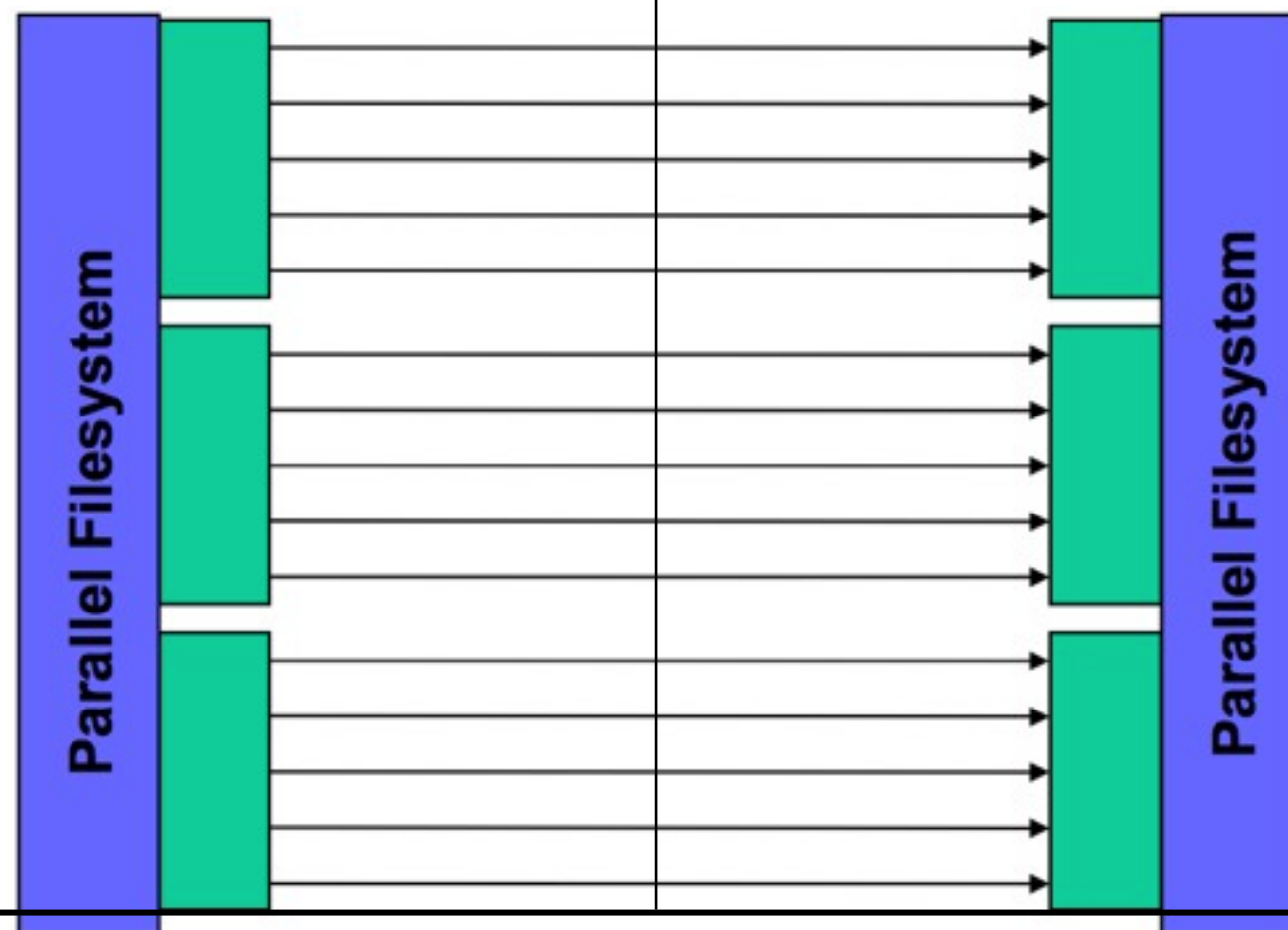
# GridFTP Performance Options

38

- Adjustable buffer size of data channels (-tcp-bs)
- Parallel TCP streams (-p)

## Striped GridFTP

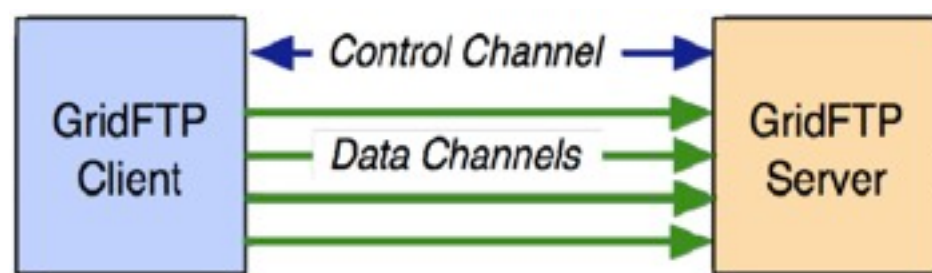
- multiple network endpoints for the transfer of the same file



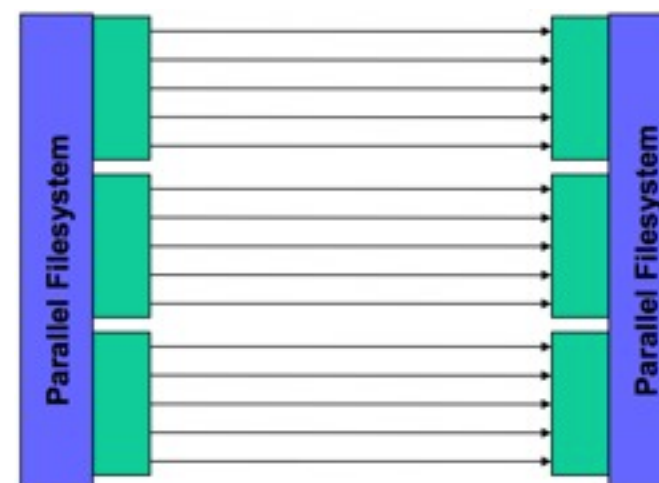


# GridFTP Performance Options

- Adjustable buffer size of data channels (-tcp-bs)
- Parallel TCP streams (-p)



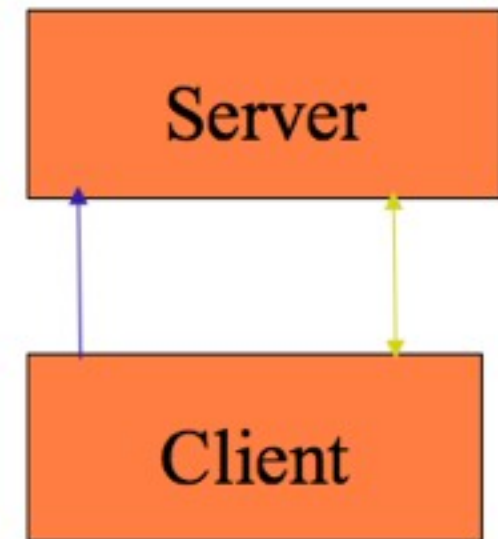
- Striped GridFTP
  - multiple network endpoints for the transfer of the same file



# GridFTP's Third Party Transfers

39

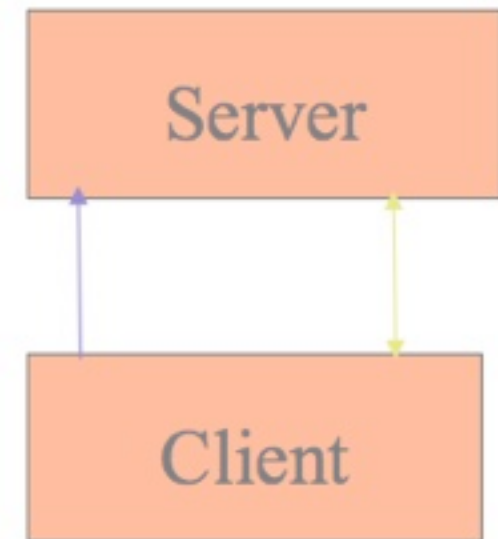
- Two party transfer
  - The client **connects** to the server
  - Information is exchanged to establish the **DC**
  - A file is transferred over the **DC**



# GridFTP's Third Party Transfers

39

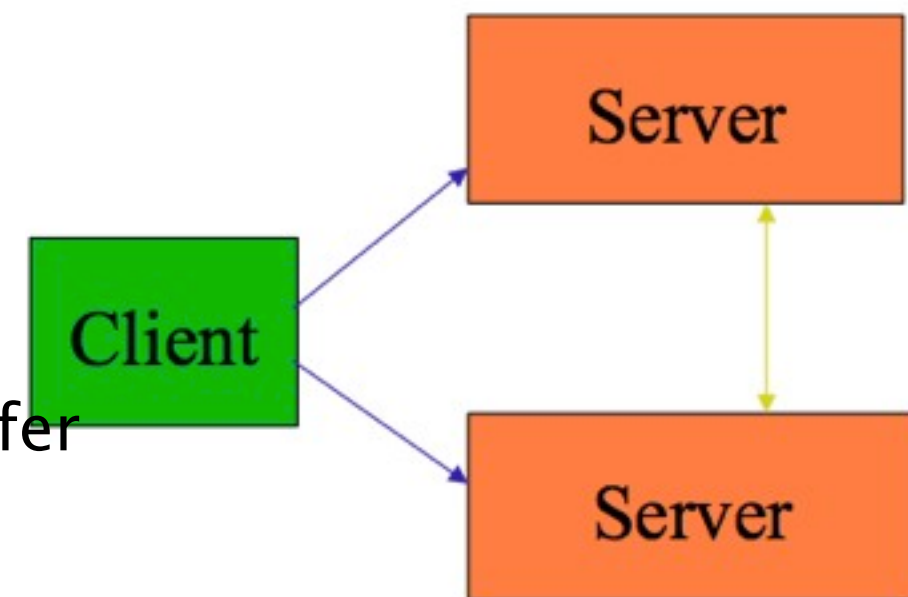
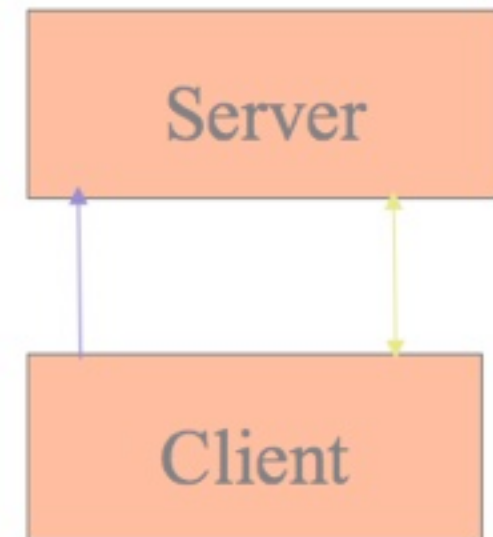
- Two party transfer
  - The client **connects** to the server
  - Information is exchanged to establish the **DC**
  - A file is transferred over the **DC**



# GridFTP's Third Party Transfers

39

- Two party transfer
  - The client **connects** to the server
  - Information is exchanged to establish the **DC**
  - A file is transferred over the **DC**
  
- Third party transfer
  - Client initiates data transfer **between 2 servers**
  - Information is routed through the client to establish DC between the two servers.
  - Data flows directly **between servers**
  - Client is notified by each server when the transfer is complete



# Globus-url-copy

40

- Command line client
  - Scriptable
- Commonly used client for GridFTP
- Syntax overview
  - `globus-url-copy [options] sourceURL destinationURL`
  - `globus-url-copy gsiftp://host/foo file:///tmp/bar`
- URL
  - `protocol://[user@][host]/path`
  - [host] can be IP address, localhost, DNS name

# Overview

41

- **General**
  - Grid, Globus Toolkit Overview
  - Authentication & Authorisation (A&A)
    - PKI, Certificates
    - GSI, Proxy Certificates, MyProxy
- **GT5**
  - Interactive Access
    - GSI-OpenSSH
    - Clients
  - Job Submission
    - GRAM5
  - Data Transfer
    - Globus GridFTP
  - Data movement service
    - Globusonline

# Globusonline Service

42

- Globusonline manages third party transfers

globusonline



# Globusonline Service

42

- Globusonline manages third party transfers



globusonline

# Globusonline Service

42

- Globusonline manages third party transfers



globusonline

# Globusonline Service

42

- Globusonline manages third party transfers

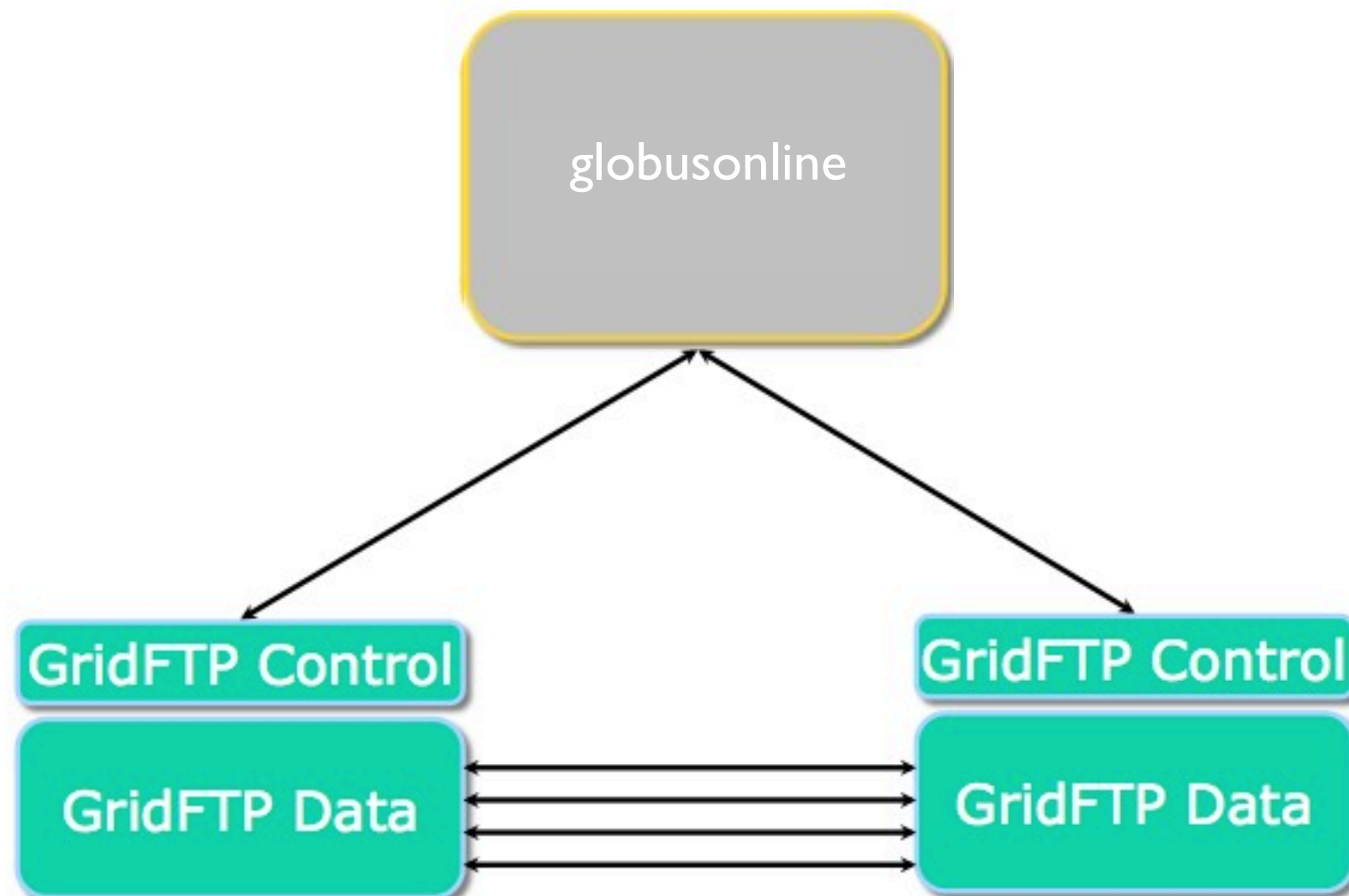


globusonline

# Globusonline Service

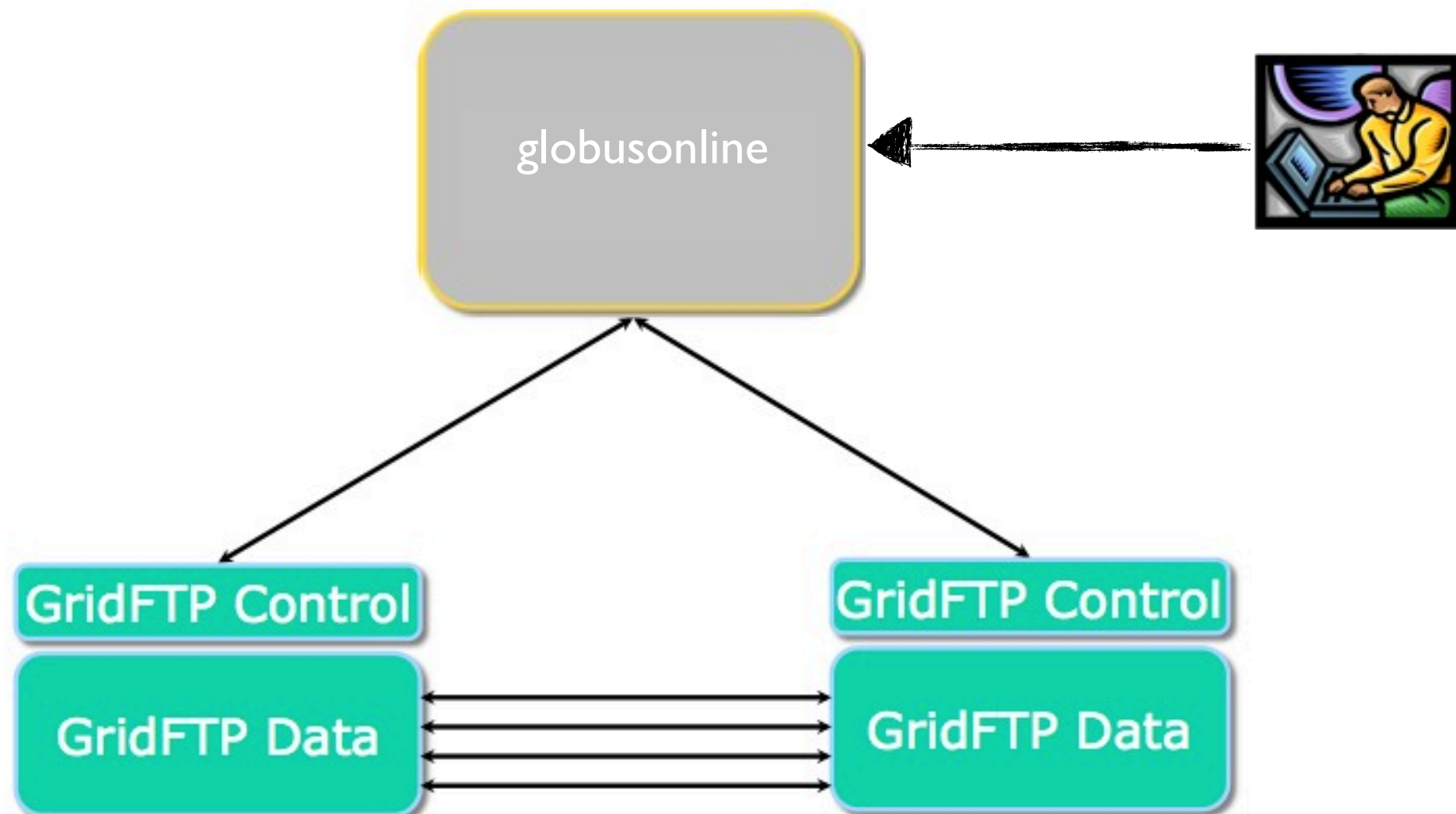
42

- Globusonline manages third party transfers



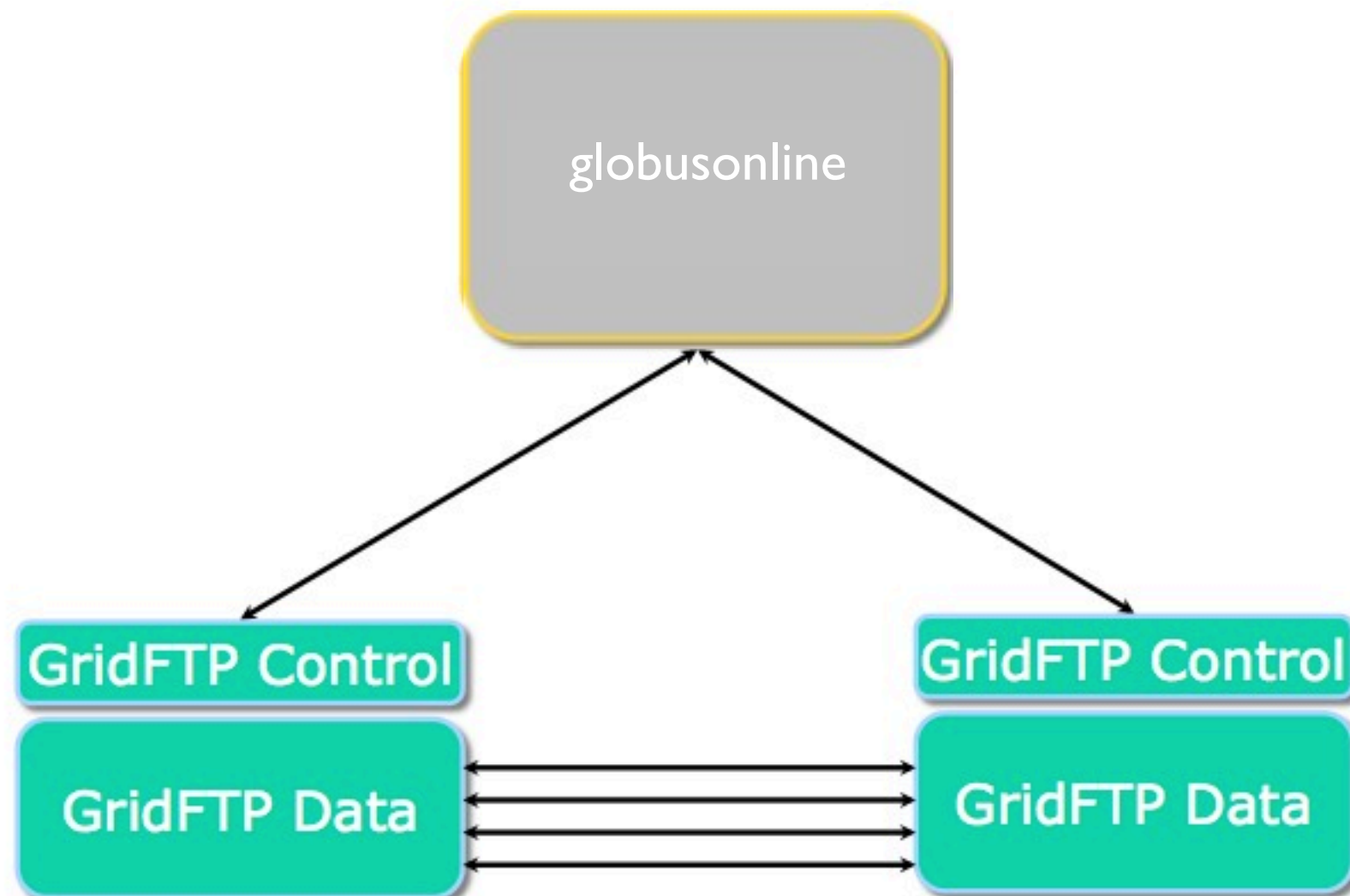
# Globusonline Service

- Globusonline manages third party transfers



# Globusonline Service

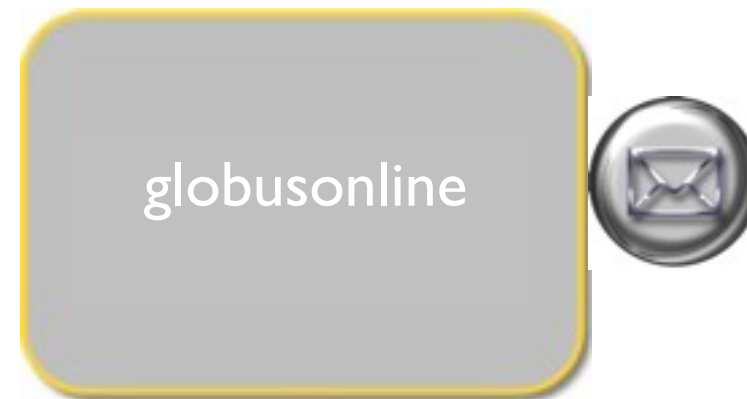
- Globusonline manages third party transfers



# Globusonline Service

42

- Globusonline manages third party transfers





# Globusonline Service

- Globusonline manages third party transfers



# Globusonline Service

43

- Globusonline is SaaS Data Movement
- CLI version
  - scriptable
- Web application
  - GUI
  - Only webbrowser needed – access world wide

# Globusonline With MyProxy

44

cineca  
GridFTP

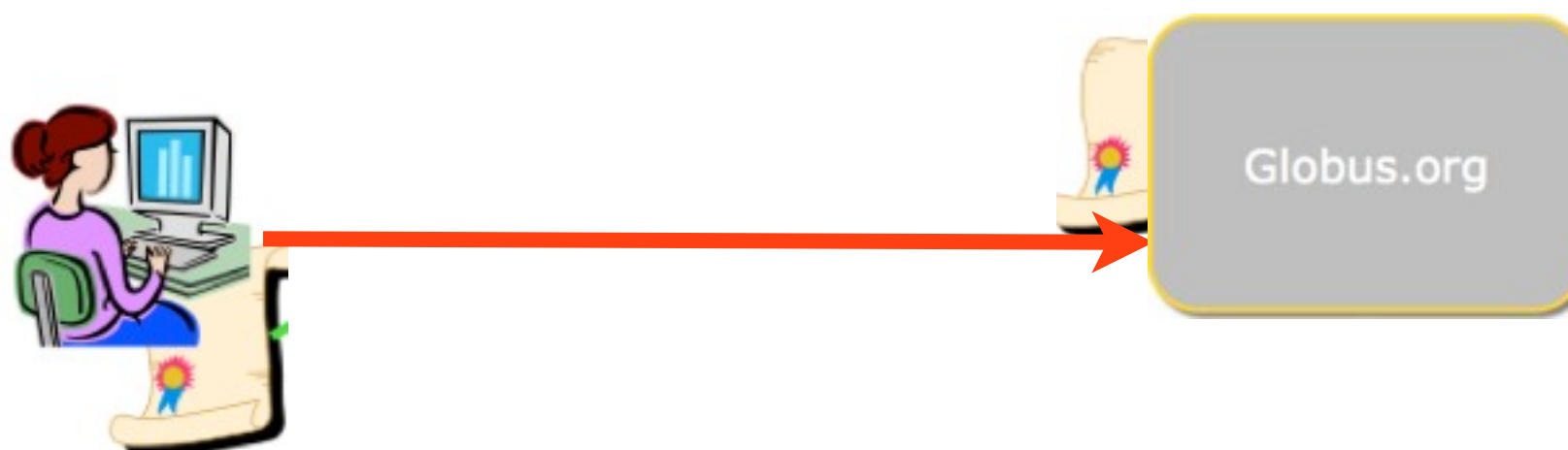


Globus.org

# Globusonline With MyProxy

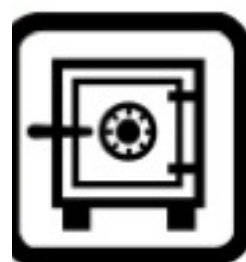
44

cineca  
GridFTP



# Globusonline With MyProxy

44



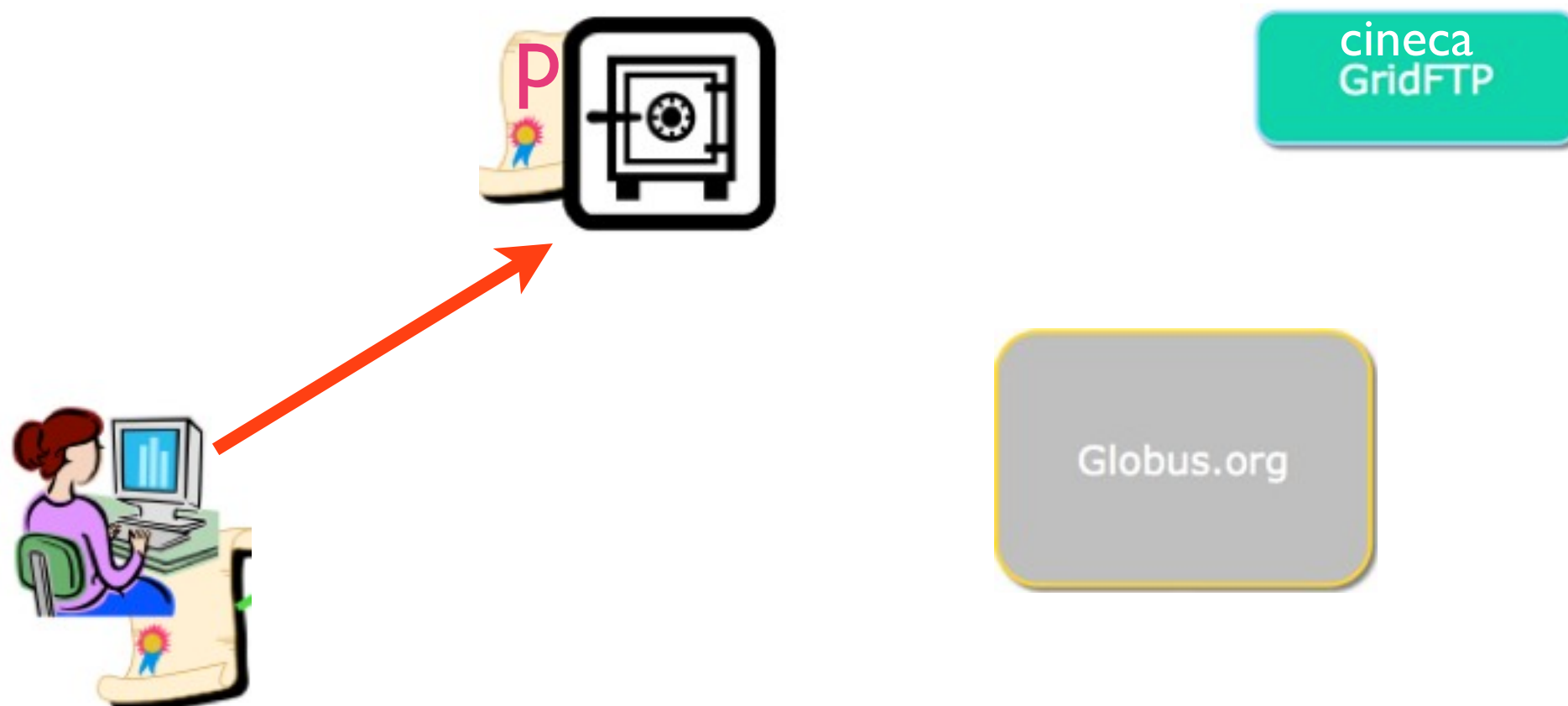
cinca  
GridFTP



Globus.org

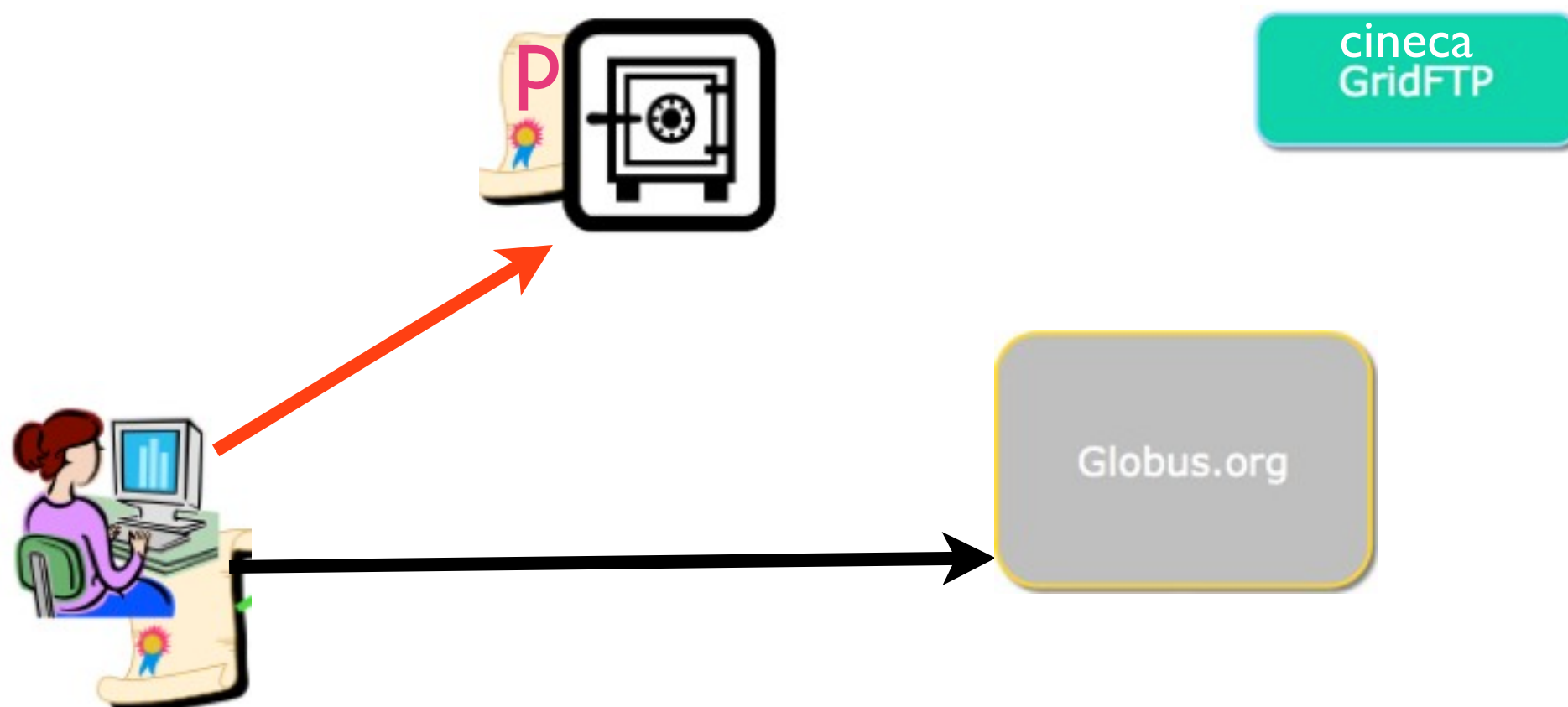
# Globusonline With MyProxy

44



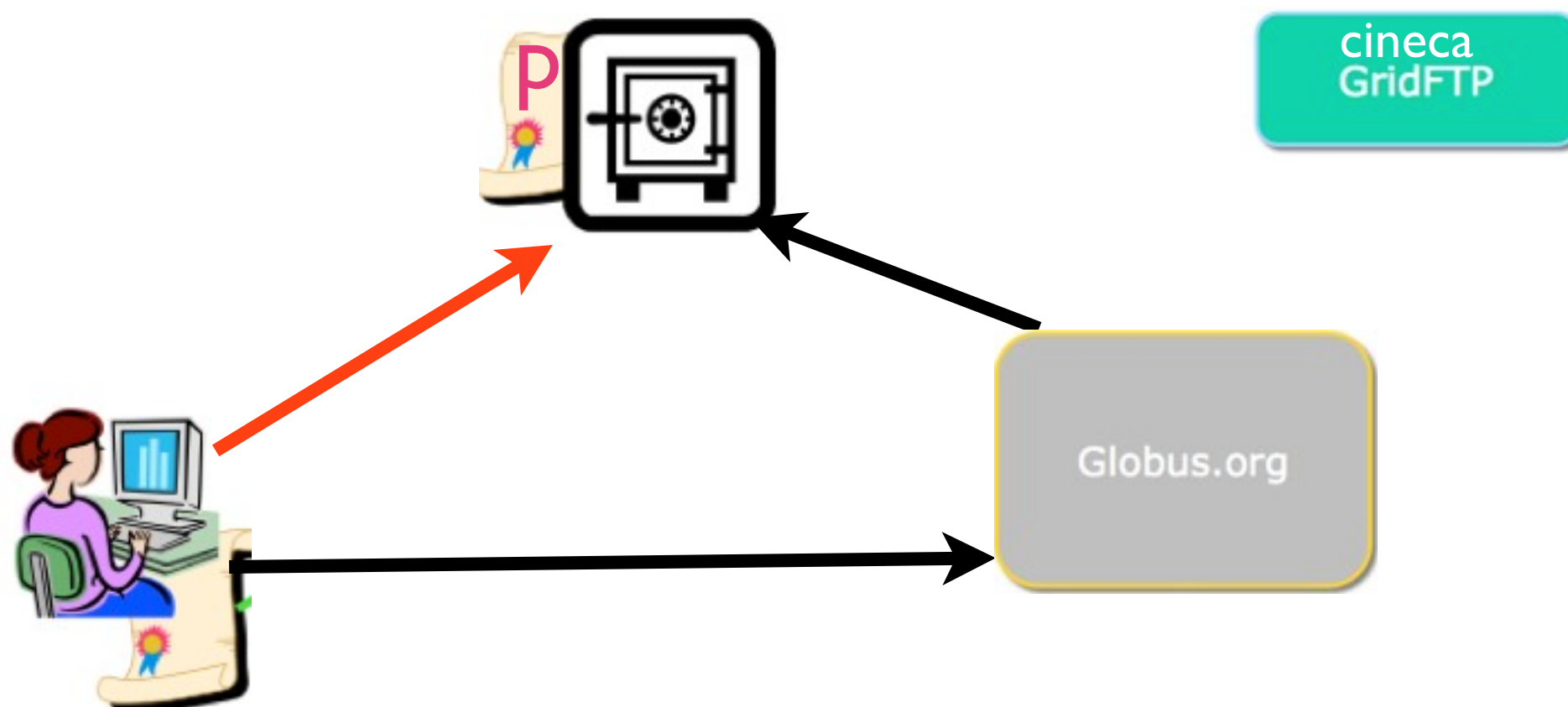
# Globusonline With MyProxy

44



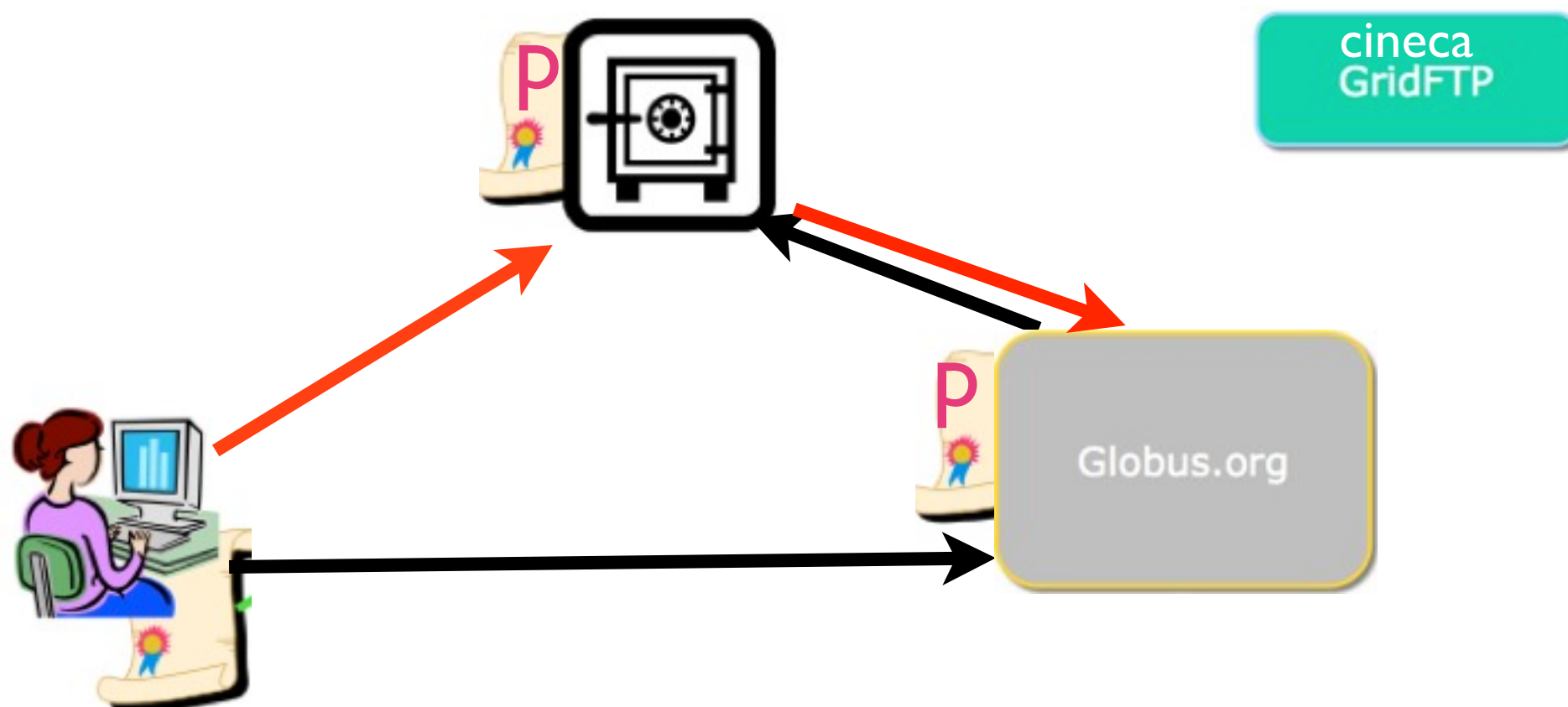


# Globusonline With MyProxy



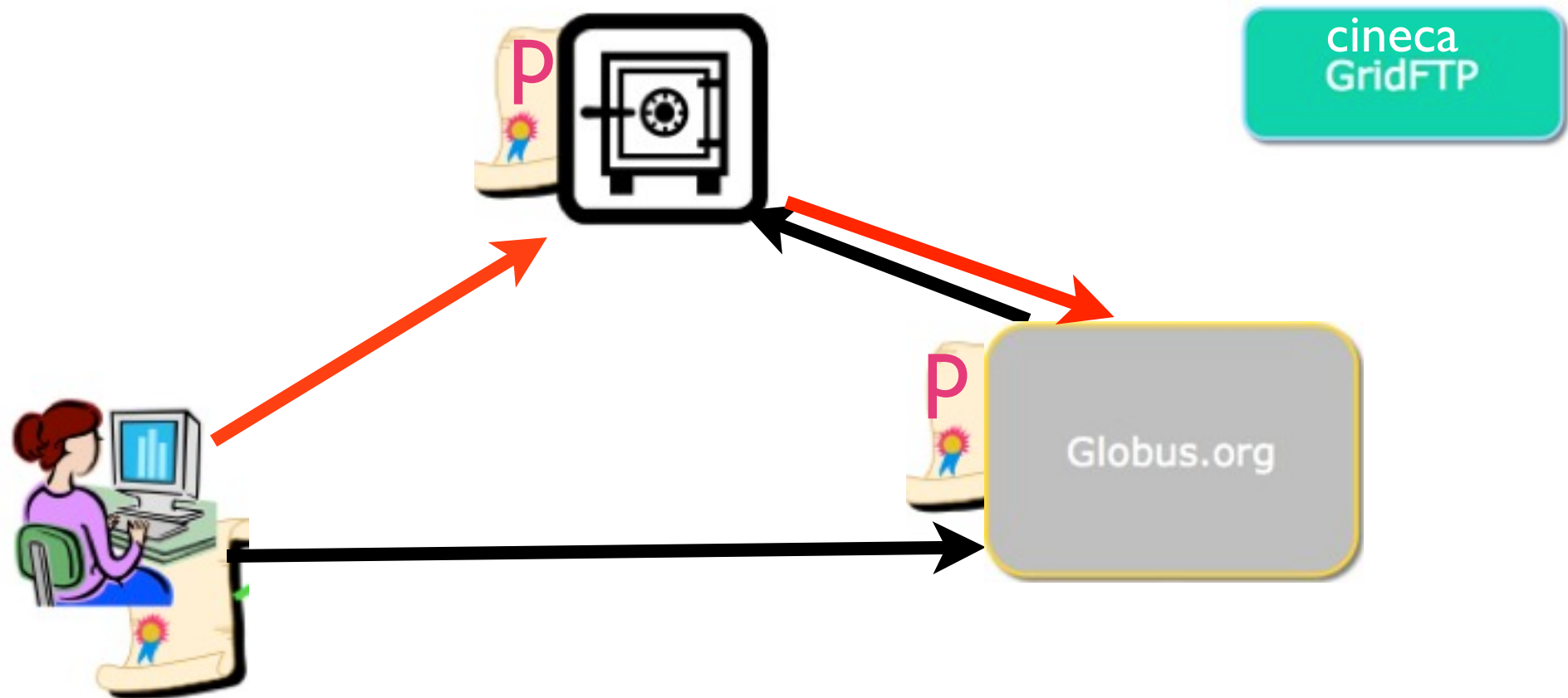
# Globusonline With MyProxy

44



# Globusonline With MyProxy

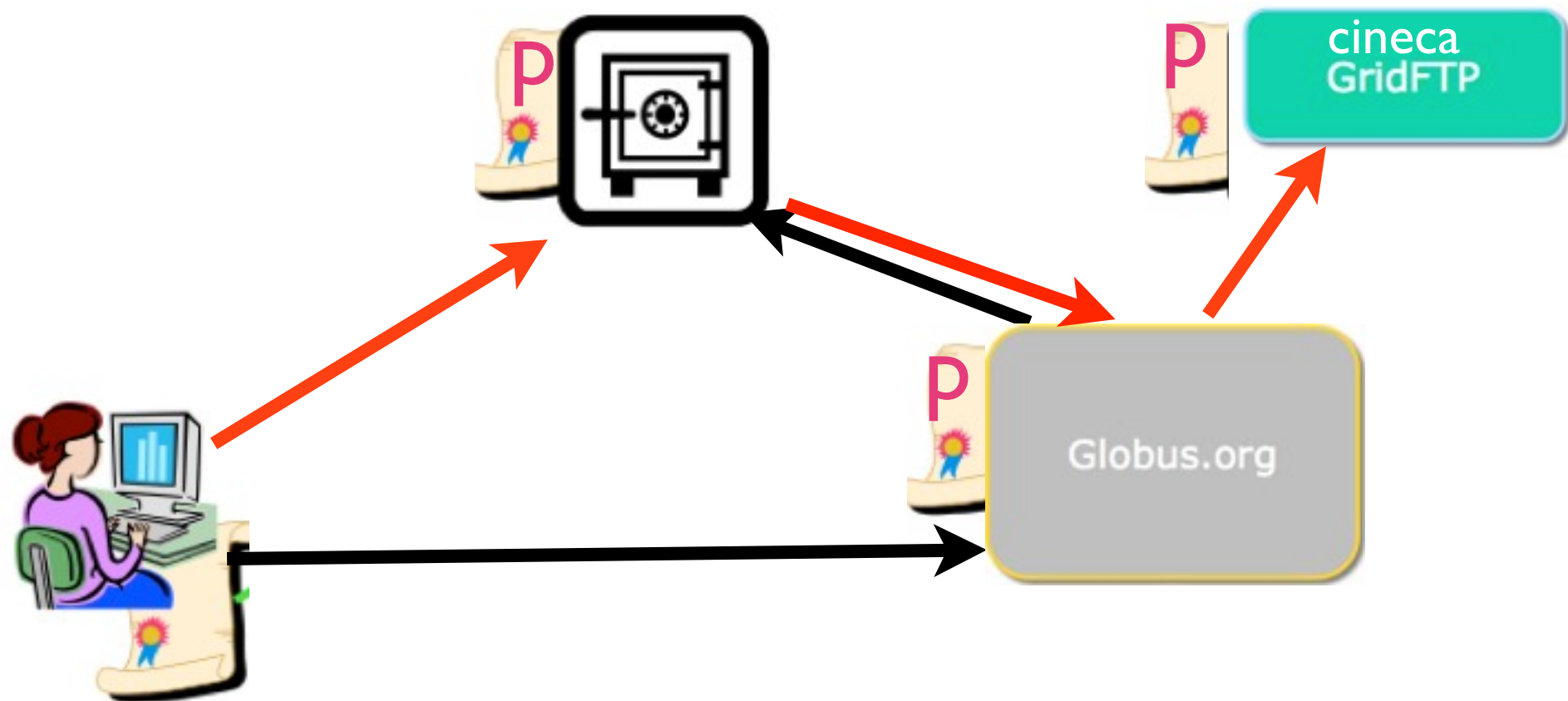
44



```
gsissh cli.globus.org endpoint-activate -m myproxy.lrz.de cineca
Welcome to globus.org
Do you want to activate using myproxy myproxy.lrz.de?
Enter Y(es, n)o: Y
Enter MyProxy pass phrase:
A credential has been received for user florian in
/tmp/x509nFXEttkoa.
Connection to cli.globusonline.org closed.
```

# Globusonline With MyProxy

44



```

gsissh cli.globus.org endpoint-activate -m myproxy.lrz.de cineca
Welcome to globus.org
Do you want to activate using myproxy myproxy.lrz.de?
Enter Y(es, n)o: Y
Enter MyProxy pass phrase:
A credential has been received for user florian in
/tmp/x509nFXEttkoa.
Connection to cli.globusonline.org closed.
  
```

# Globusonline Service

45

- Access via gsissh
  - On login you don't get a shell but
  - Text interface
- Functions
  - Manage end-points
  - Delegate your proxy
    - Direct via gsi-ssh
    - Indirect via myproxy
  - Initiate, observe, stop your transfers
  - Notifications
- API
  - REST interface

# Overview

46

- **General**
  - Grid, Globus Toolkit Overview
  - Authentication & Authorisation (A&A)
    - PKI, Certificates
    - GSI, Proxy Certificates, MyProxy
- **GT5**
  - Interactive Access
    - GSI-OpenSSH
    - Clients
  - Data Transfer
    - Globus GridFTP
  - Job Submission
    - GRAM5
  - Data movement service
    - Globusonline