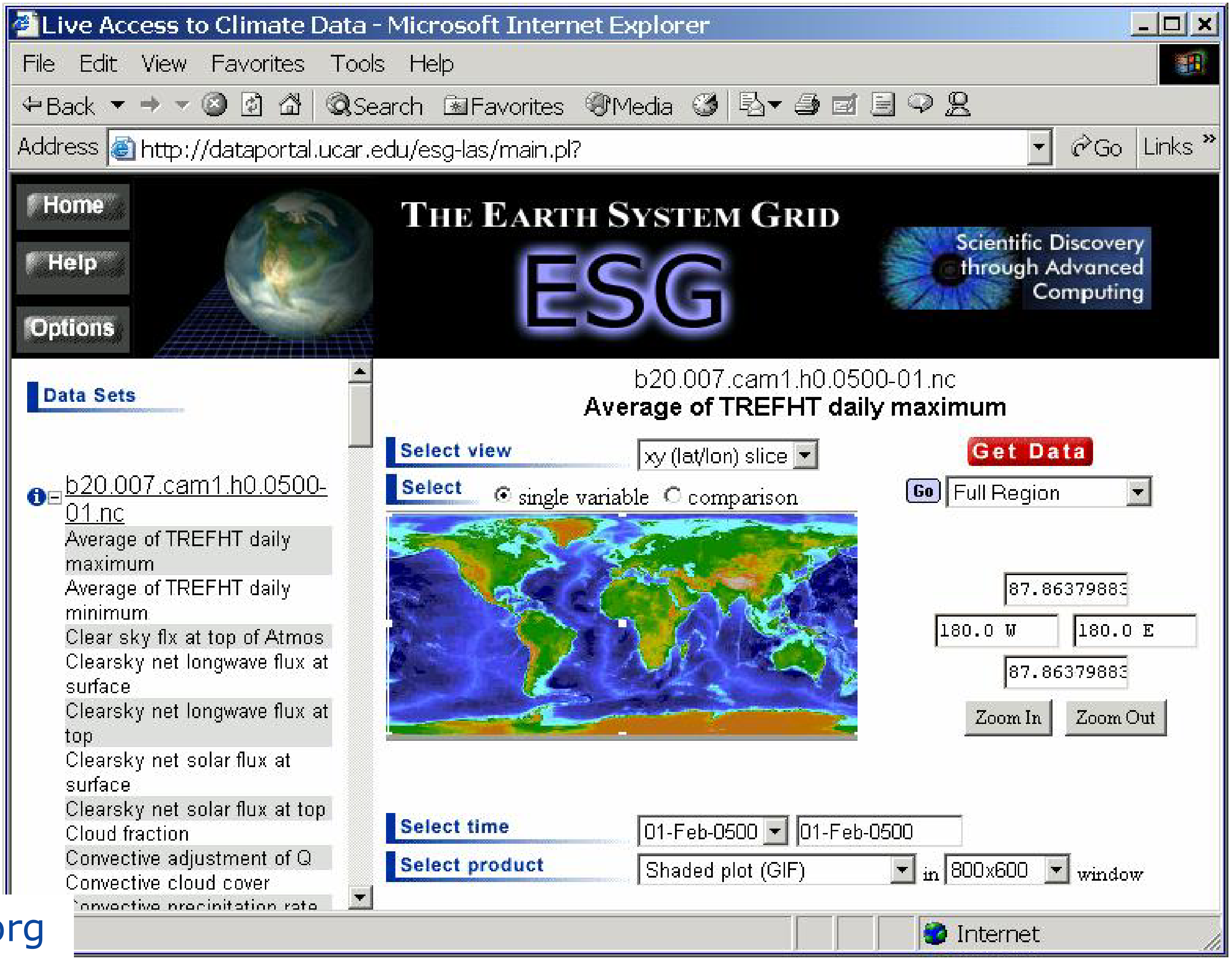
(Reusable) Portal-based Authorization Solution for the Earth System Grid SciDAC Project

Veronika Nefedova, Frank Siebenlist
Argonne National Laboratory

With: Rachana Ananthakrishnan, Ian Foster, Laura Pearlman, Von Welch

DOE Earth System Grid

Goal: address technical obstacles to the sharing & analysis of high-volume data from advanced earth system models


Live Access to Climate Data - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Go Links
Home
Help
Options
The Earth System Grid
ESG
Scientific Discovery through Advanced Computing
Get Data
Go
Zoom In
Zoom Out
Internet

Address: http://dataportal.ucar.edu/esg-las/main.pl?

Data Sets

b20.007.cam1.h0.0500-01.nc
- Average of TREFHT daily maximum
- Average of TREFHT daily minimum
- Clear sky flx at top of Atmos
- Clearsky net longwave flux at surface
- Clearsky net longwave flux at top
- Clearsky net solar flux at surface
- Clearsky net solar flux at top
- Cloud fraction
- Convective adjustment of Q
- Convective cloud cover
- Convective precipitation rate

b20.007.cam1.h0.0500-01.nc
Average of TREFHT daily maximum

Select view: xy (lat/lon) slice

Select: ● single variable ○ comparison

Full Region

87.86379883
180.0 W 180.0 E
87.86379883

Select time: 01-Feb-0500 01-Feb-0500

Select product: Shaded plot (GIF) in 800x600 window

www.earthsystemgrid.org

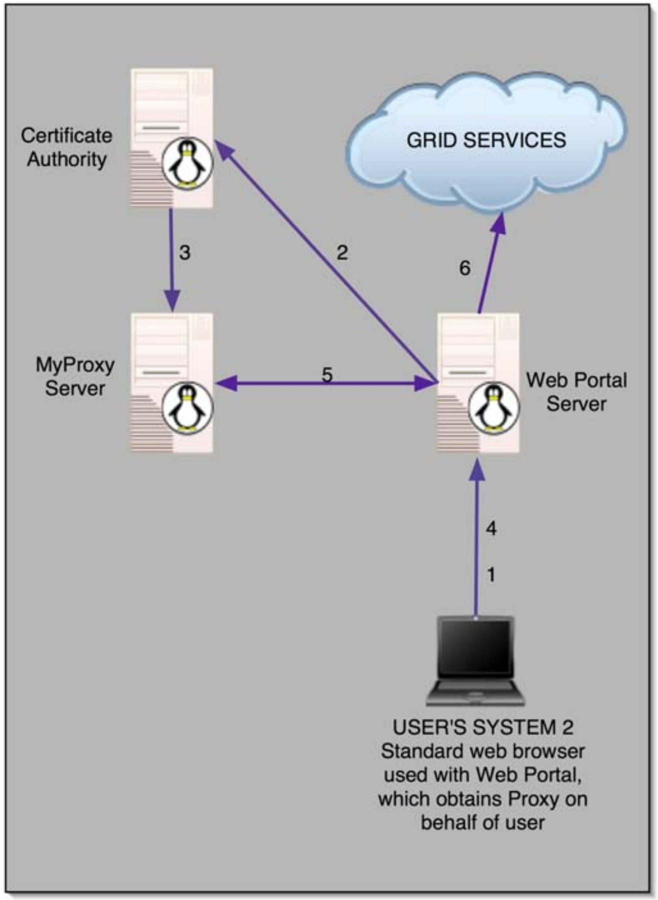
ESG Authorization requirements

- Access to most data requires that the name of the requesting user be logged.
- Access to some private data is restricted to specific users.
- Some data is located on mass storage systems to which access is restricted to users with approved PKI credentials.
- Some data is located on HPSS storage behind GridFTP server
- Some data is located on disk storage behind HTTPS server.

ESG Authorization Requirements

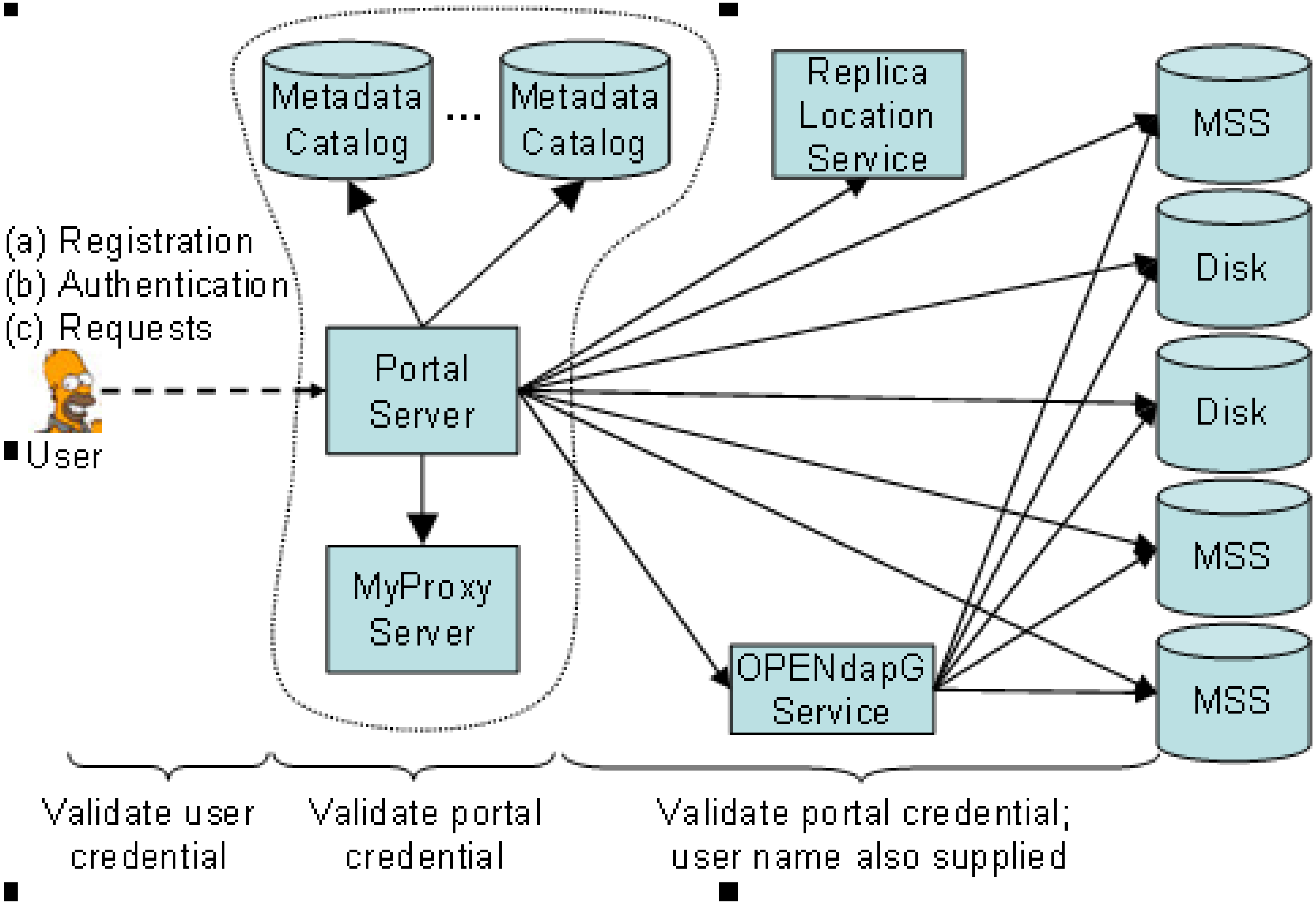
- Access control for data accessed via portal
  - Group-based control to data and metadata
- Variety of data return modalities, e.g.:
  - From portal as intermediary to servers
  - Directly from GridFTP server
- Credentials of a variety of qualities
  - Higher quality via formal CA (personal review)
  - Lower quality via Web (email verification)
- Easy to use Web sign on
  - MyProxy as credential repository
- GSI credentials for GridFTP server access

PURSE: Portal-based User Registration Service

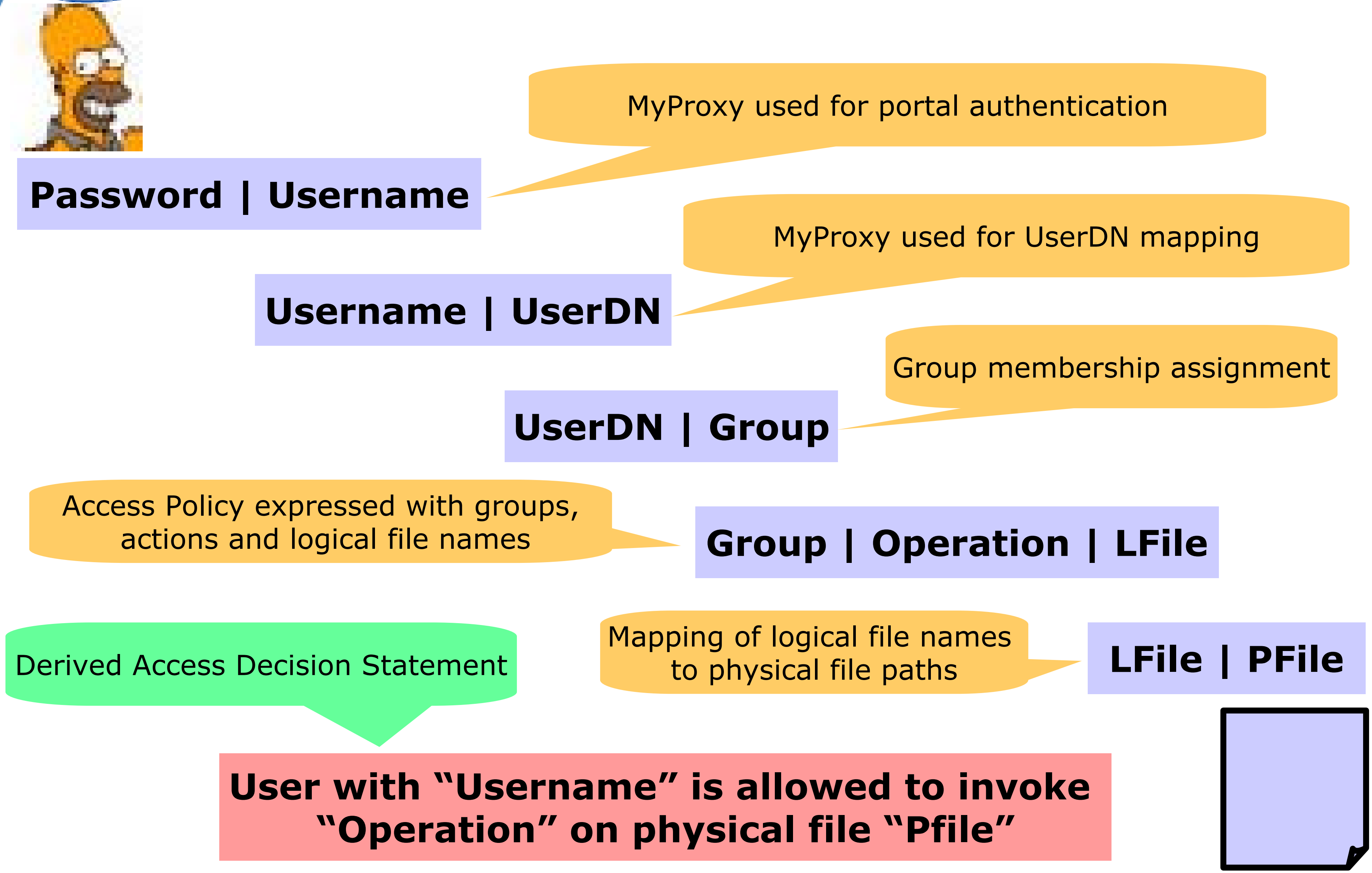


- Portal extensions (CGI scripts) that automate user registration requests.
    - Solicits basic data from user.
    - Generates cert request from ESG CA (implemented with “simple CA” from GT).
    - Admin interface allows CA admin to accept/reject request.
    - Generates a certificate and stores in MyProxy service.
    - Gives user ID/password for MyProxy.
- Benefits
    - Users never have to deal with certificates.
    - Portal can get user cert from MyProxy when needed.
    - Database is populated with user data.
    - Users are assigned to one or several user groups (with different data access permissions)

ESG data access control



ESG Authorization Model



Password | Username

MyProxy used for portal authentication

Username | UserDN

MyProxy used for UserDN mapping

UserDN | Group

Group membership assignment

Access Policy expressed with groups, actions and logical file names

Group | Operation | LFile

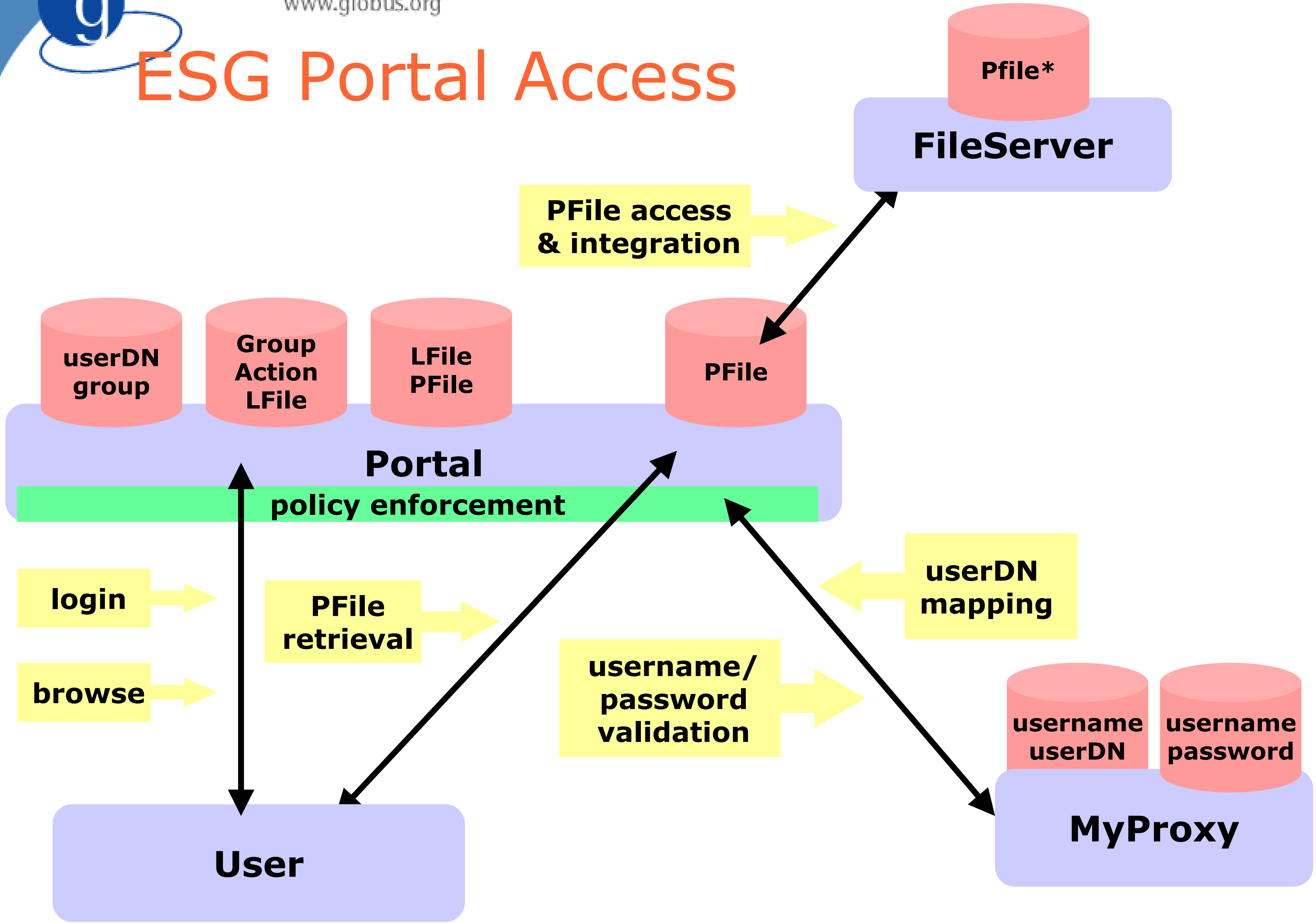
Mapping of logical file names to physical file paths

LFile | PFile

Derived Access Decision Statement

User with “Username” is allowed to invoke “Operation” on physical file “Pfile”

ESG Portal Access



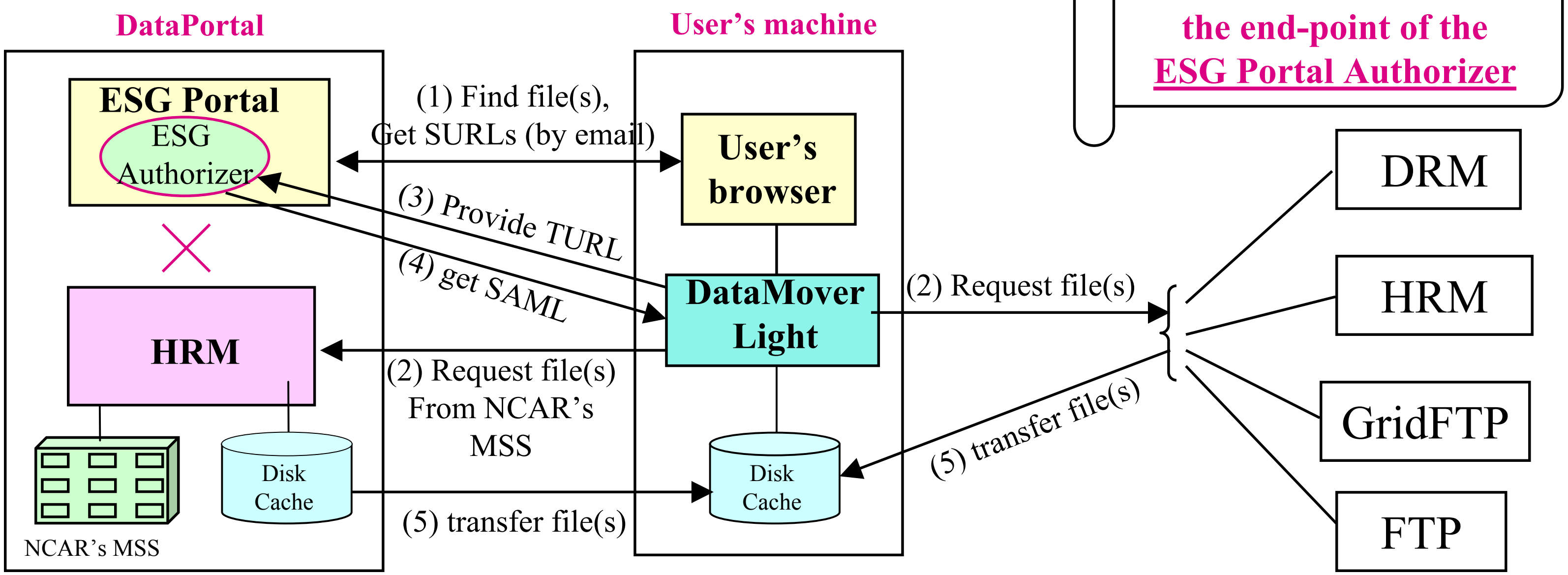
ESG External GridFTP Access

- User browses portal to identify file(s)
- Portal returns
  - Physical file location (URL)
  - SAML assertion in CAS format: "User can invoke requested operation on file(s)"
- User:
  - Obtains proxy-certificate from MyProxy
  - Embeds SAML-assertion in proxy-cert
  - Uses GridFTP client to retrieve physical file(s) from CAS-enabled GridFTP server

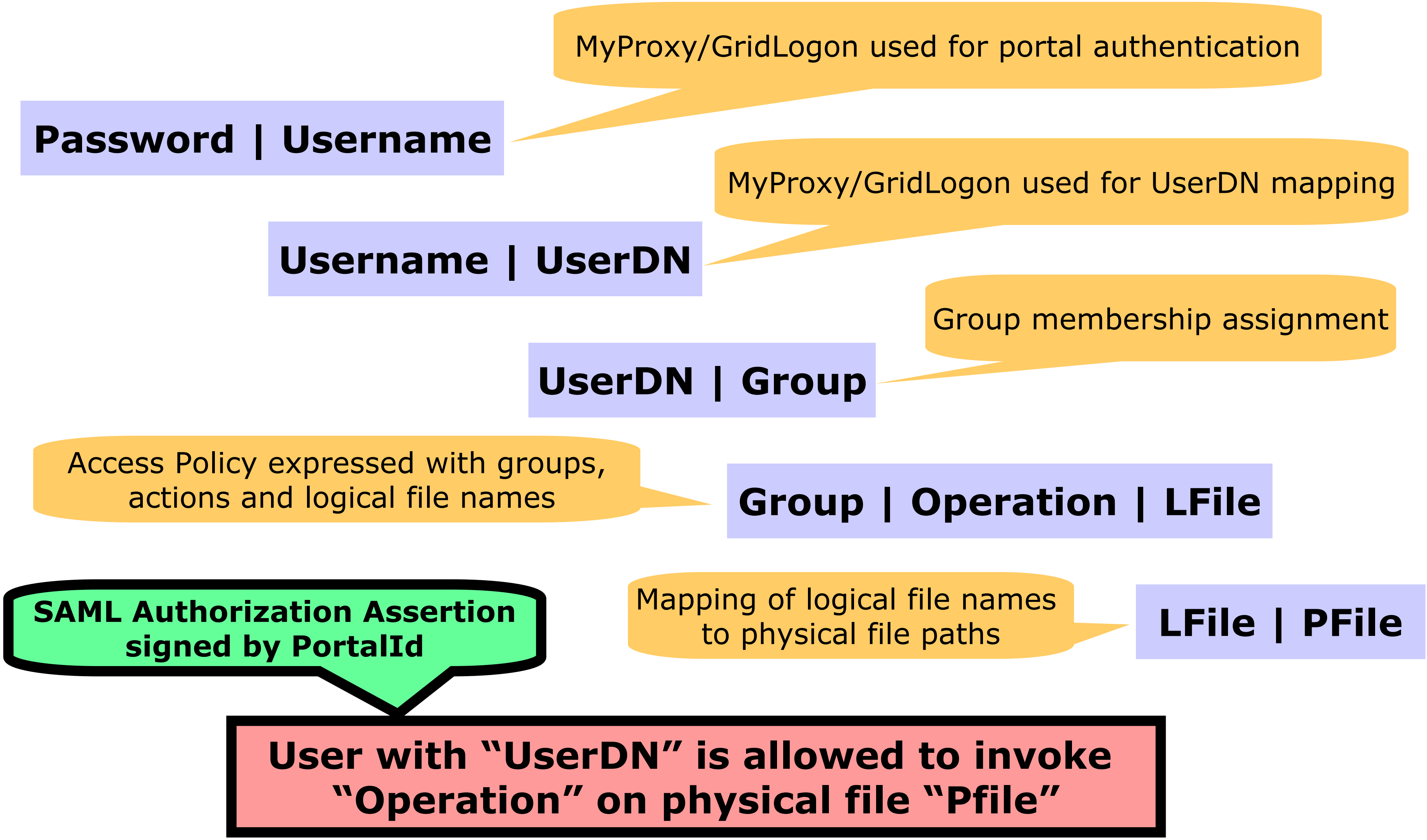
Download with DML from any source locations

- User downloads DataMover-light (must contain GridFTP and GSI)
- User goes to portal, select files
- Portal does not get any file s
- Portal sends email to user
  - Contains a text file of files to be moved
  - Contains instructions and lifetime
- DML contacts source SRMs to get TURL (for GridFTP it is not necessary)
- DML Contact "ESG authorizer", provides it with LFN & TURL
- DML gets back SAML with long lifetime (days) for each file
- DML invokes GridFTP
- DML automatically releases files after it moves them

Note:
The datamover.txt file will contain a header with the end-point of the ESG Portal Authorizer



Authorization Assertions



MyProxy/GridLogon used for portal authentication

Password | Username

MyProxy/GridLogon used for UserDN mapping

Username | UserDN

Group membership assignment

UserDN | Group

Access Policy expressed with groups, actions and logical file names

Group | Operation | LFile

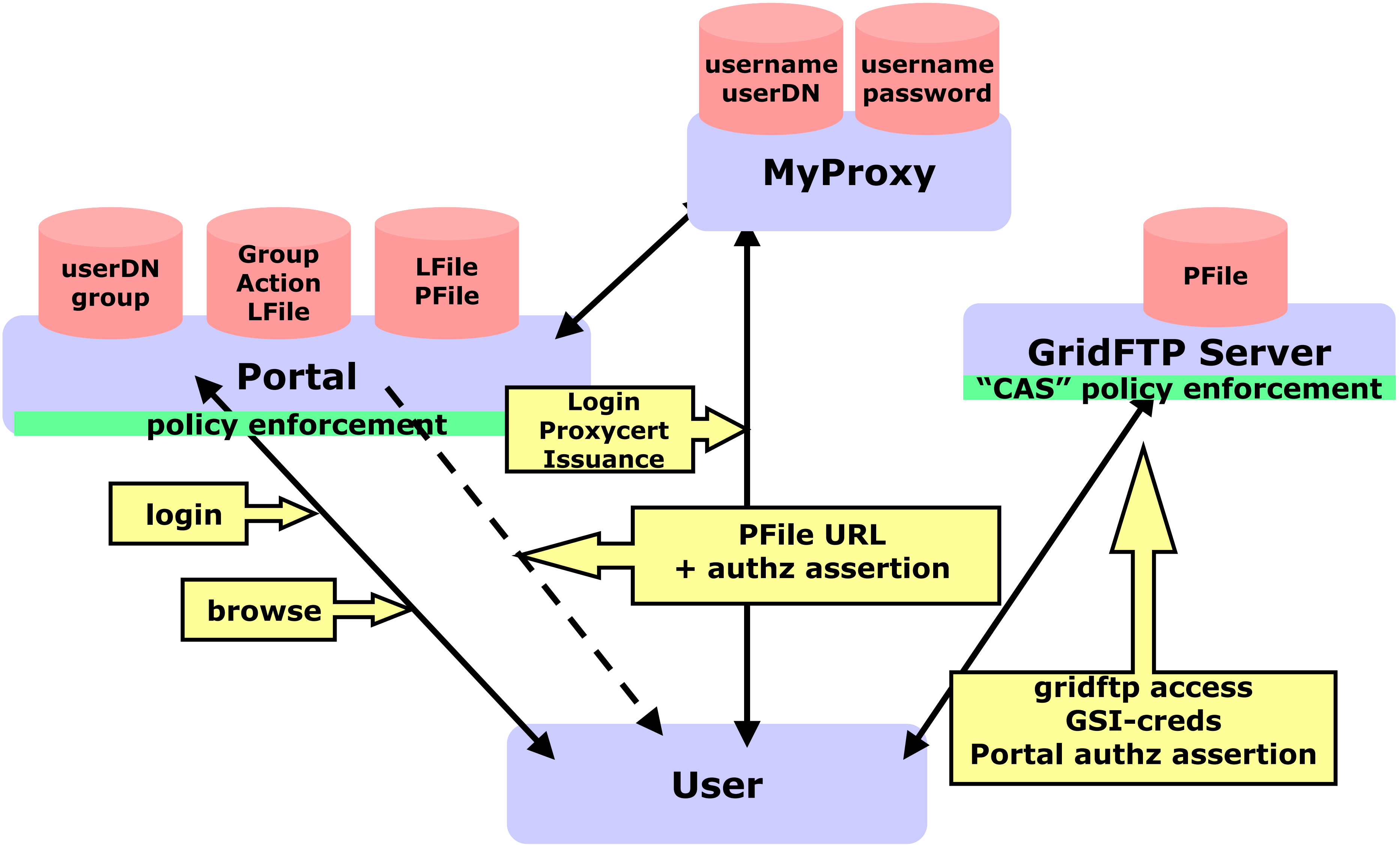
Mapping of logical file names to physical file paths

LFile | PFile

SAML Authorization Assertion signed by PortalId

User with “UserDN” is allowed to invoke “Operation” on physical file “Pfile”

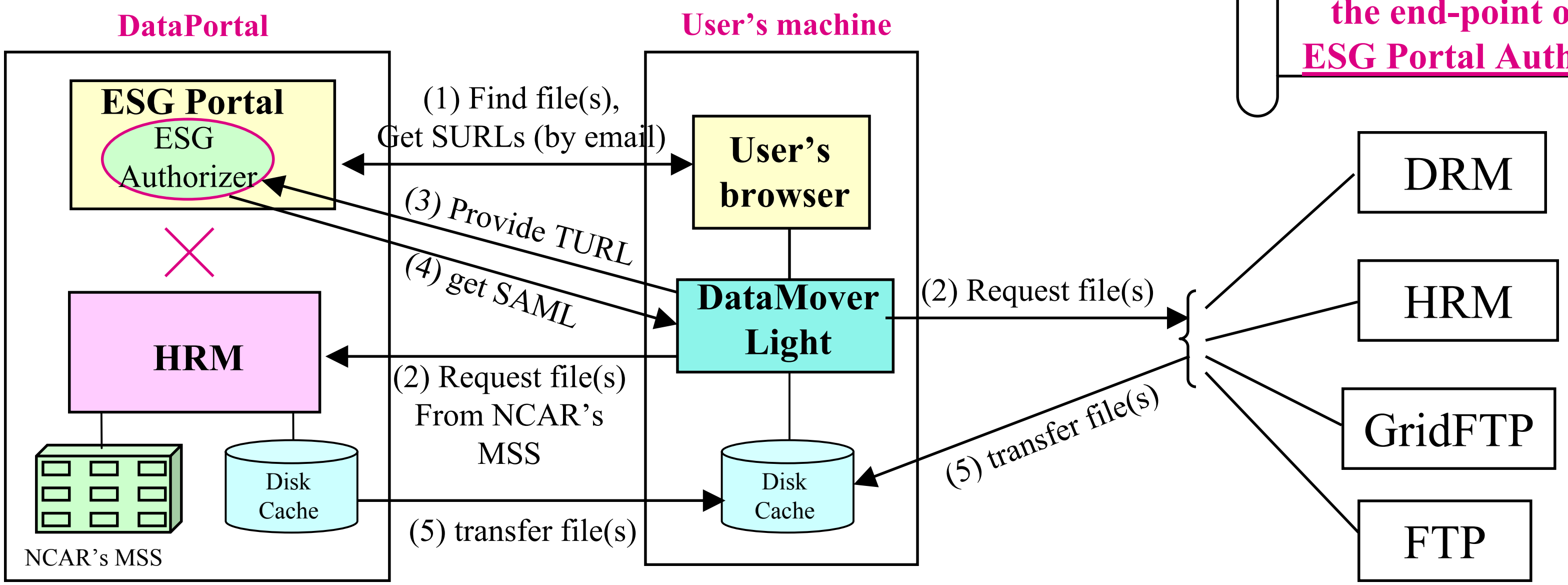
ESG External GridFTP Retrieval



SAML over GridFTP control channel

- •User goes to portal, select files
- •Portal does <u>not</u> get any files
- •Portal sends email to user
- •DML contacts source SRMs to get TURL (not necessary for GridFTP)
- •DML Contact "ESG authorizer", provides it with LFN & TURL
- •DML gets back SAML with long lifetime (days) for each file
- •DML invokes GridFTP
- •DML automatically releases files after it moves them

Note:
The datamover.txt file will contain a header with the end-point of the <u>ESG Portal Authorizer</u>



Reuse of Fabric & Plumbing from Community Auth. Service (CAS)

- ESG-Portal uses no CAS server but generates its own authorization statements
  - Statements are domain specific
- Same assertion format as CAS
  - Standard “SAML” assertion signed by PortalId
- User deploys CAS-enabled GridFTP client
  - Deploys identical GSI creds and proxy-certs
- Site uses CAS-enabled GridFTP server
- Remote site trusts Portal (instead of CAS)
- Portal makes access control decisions

Summary

- ESG work builds on GSI & CAS technology to address specific ESG requirements
  - Ease of use
  - Group-based access control to distributed data
- Yields two broadly useful tools that should find further application in Grid projects
  - PURSE: Portal-based User Reg. System
  - Service access authorization assertions
- Next steps to be defined with ESG & others
  - E.g., access control for arbitrary services