Rights Management in Globus Data Services

Ann Chervenak, ISI/USC

Bill Allcock, ANL/UC

the globus alliance
www.globus.org

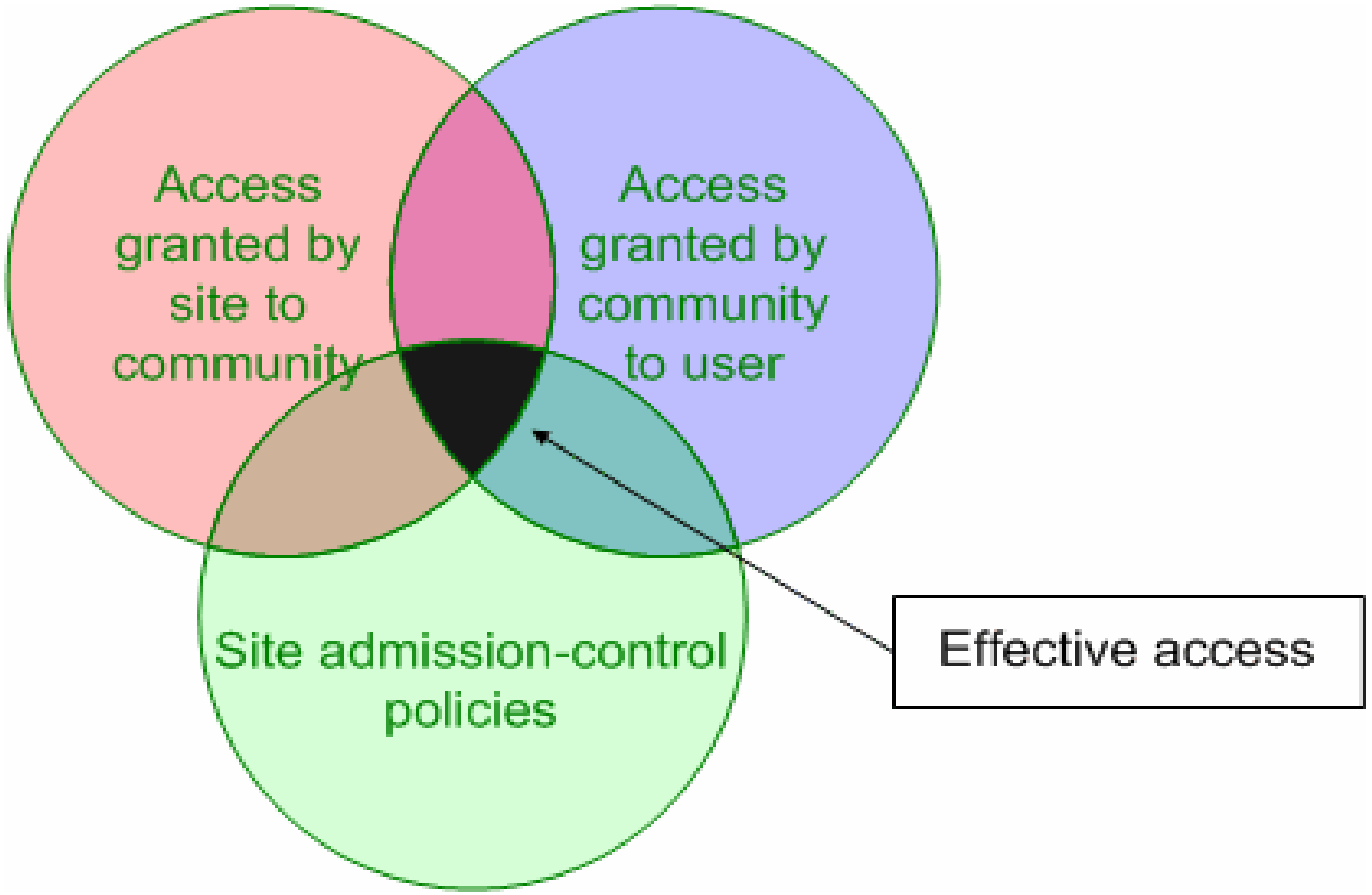
Outline

- **Brief discussion of Globus Authorization in General**
- Authorization in GridFTP and related tools
- Authorization in the Globus Replica Location Service and related tools.

GT4 Security

- Public-key-based authentication
- Extensible authorization framework based on Web services standards
  - SAML-based authorization callout
    - As specified in GGF OGSA-Authz WG
  - Integrated policy decision engine
    - XACML policy language, per-operation policies, pluggable
- Credential management service
  - MyProxy (One time password support)
- Community Authorization Service
- Standalone Delegation Service
- SimpleCA: Online credential generation
- PERMIS: Authorization service callout

Effective Policy Governing Access Within A Collaboration



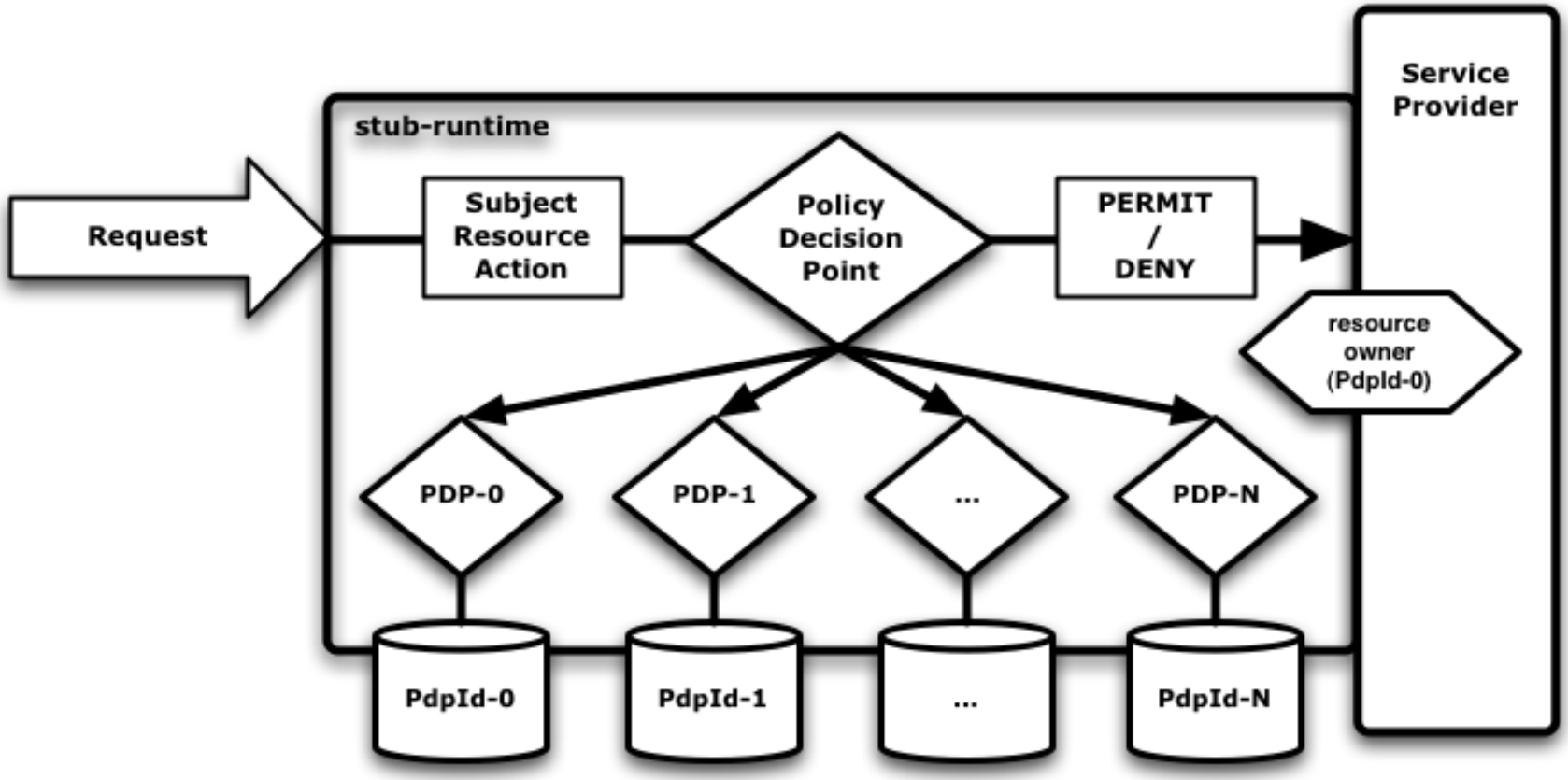
the globus alliance
www.globus.org

GT's Assertion Processing "Problem"

- VOMS/Permis/X509/Shibboleth/SAML/Kerberos identity/attribute assertions
- XACML/SAML/CAS/XCAP/Permis/ProxyCert authorization assertions
- Assertions can be pushed by client, pulled from service, or locally available
- Policy decision engines can be local and/or remote
- Delegation of Rights is required "feature" implemented through many different means

GT-runtime has to mix and match all policy information and decisions in a consistent manner…

GT Authorization Framework



the globus alliance
www.globus.org

GT's Authorization Processing Model (1)

- Use of a Policy Decision Point (PDP) abstraction that conceptually resembles the one defined for XACML.
  - Normalized request context and decision format
  - Modeled PDP as black box authorization decision oracle
- After validation, map all attribute assertions to XACML Request Context Attribute format
- Create mechanism-specific PDP instances for each authorization assertion and call-out service
- The end result is a set of PDP instances where the different mechanisms are abstracted behind the common PDP interface.

the globus alliance
www.globus.org

GT's Authorization Processing Model (2)

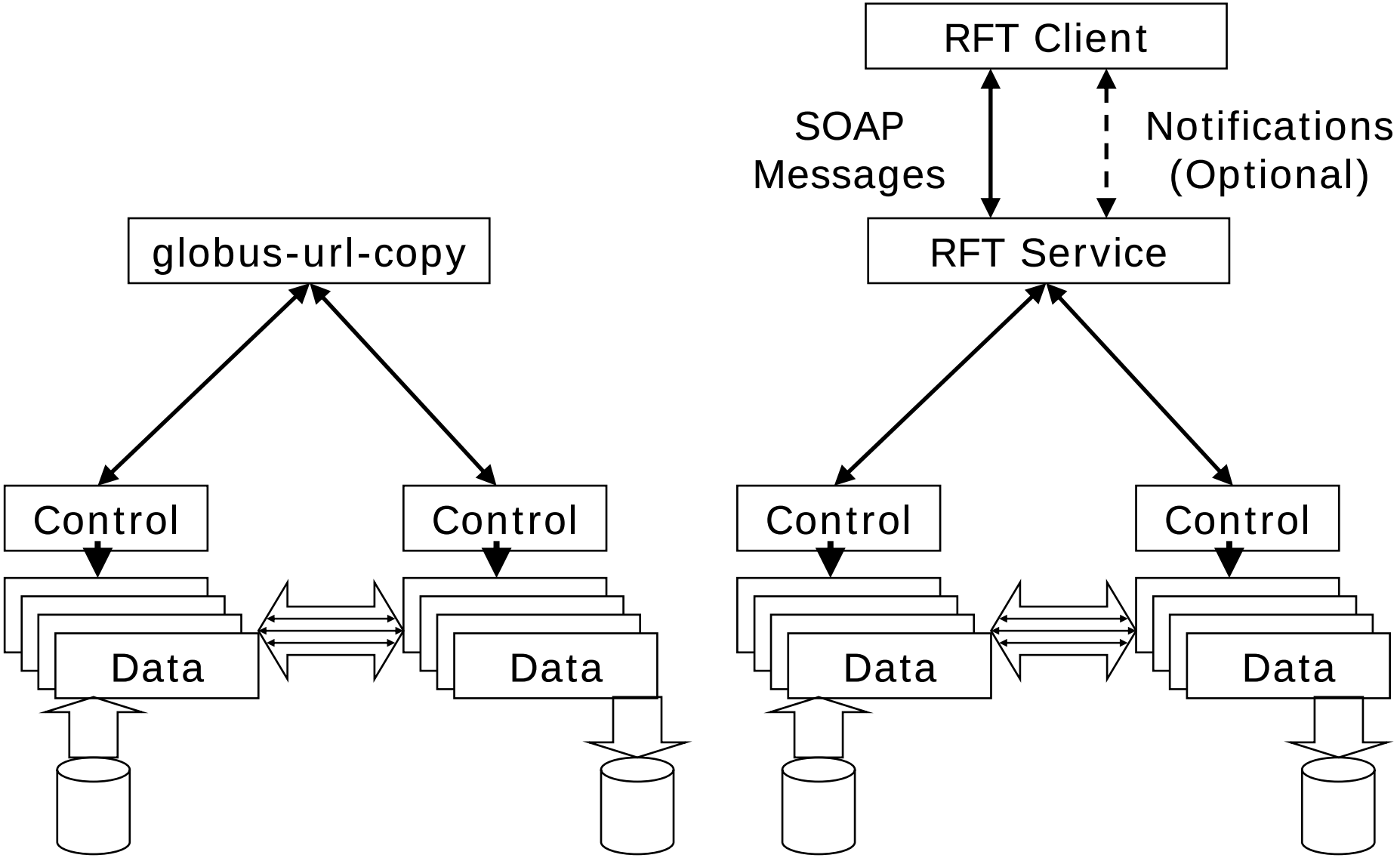
- The Master-PDP orchestrates the querying of each applicable PDP instance for authorization decisions.
- Pre-defined combination rules determine how the different results from the PDP instances are to be combined to yield a single decision.
- The Master-PDP is to find delegation decision chains by asking the individual PDP instances whether the issuer has delegated administrative rights to other subjects.
- the Master-PDP can determine authorization decisions based on delegated rights without explicit support from the native policy language evaluators.

the globus alliance
www.globus.org

Outline

- Brief discussion of Globus Authorization in General
- **Authorization in GridFTP and related tools**
- Authorization in the Globus Replica Location Service and related tools.

One Slide Overview of GridFTP



What is the problem?

- MxN problem
  - Too many accounts, not scalable
- Site wants to authorize by community, but must be able to identify an individual if there is a problem
- Process still runs under a UID and file system permissions must be honored.
- Requirement for authorization based not only identity, but a role or an attribute.
- Scalability
  - Millions of file, 1000s of users...

Initial Observations

- RFT authorization is NOT data access authz
  - No data access authorization, but right to run (like a job)
- GridFTP is more than read and write
  - Many people use it as a “remote shell”
  - Listings, file/dir creation and deletion, permissions, etc.
- GridFTP is authorization agnostic.
  - The protocol, nor the Globus implementation of it, has a specified authorization mechanism.

What is provided from Globus

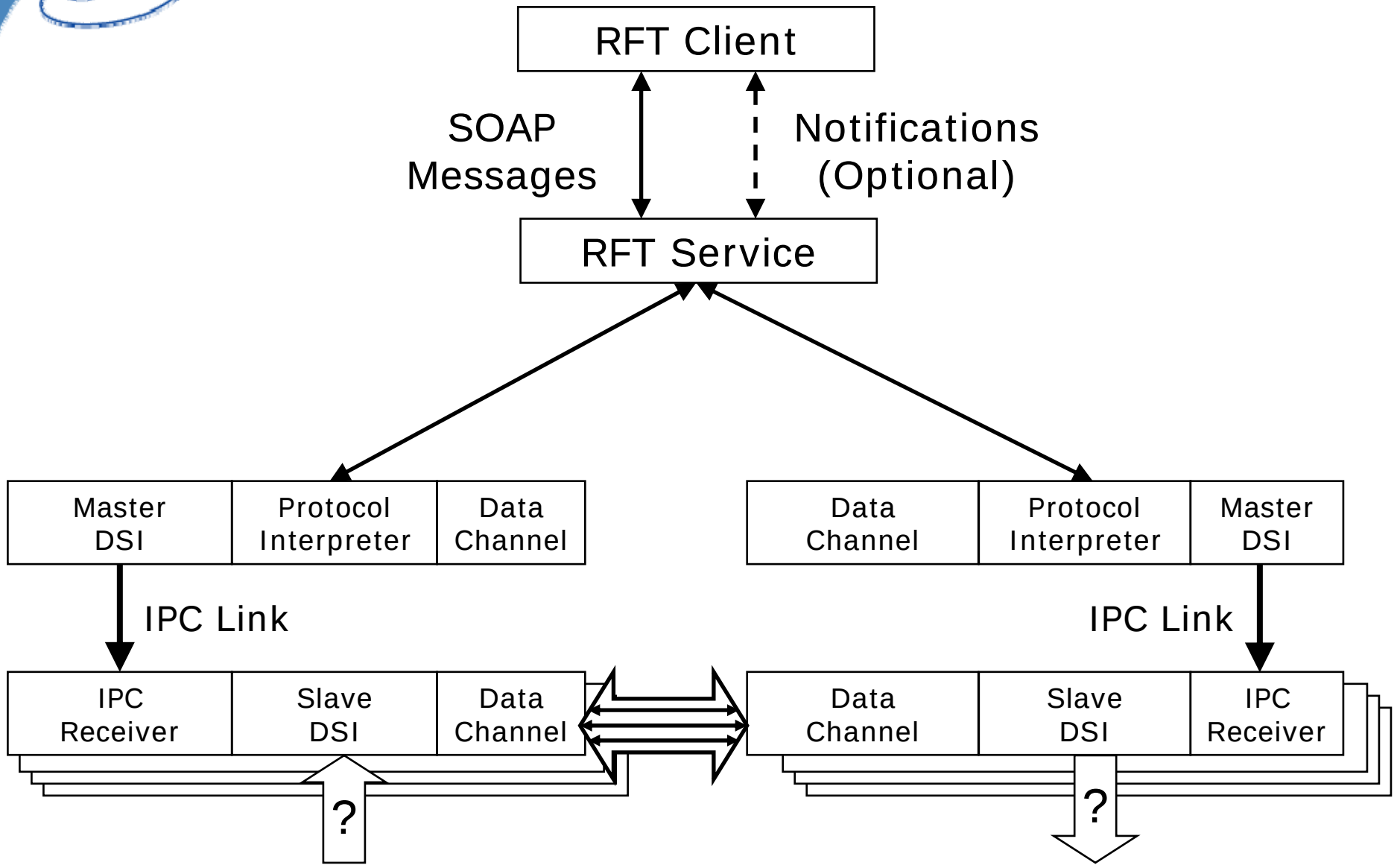
- Essentially, we use OS file system permissions
- PI typically runs as root during connection establishment
  - Uses host certificate for authentication
- To determine what account to map the user to a security callout is used
  - CAS enabled
    - globus_gss_assist_map_and_authorize()
  - Else
    - globus_gss_assist_gridmap
    - globus_gss_assist_userok (specific account requested)

the globus alliance
www.globus.org

All cases are basically the same

- Just as in web services slides, three pieces of information are provided
  - Security context (DN plus optional assertions. Essentially an opaque structure)
  - The resource to be acted on
  - The action being request
- A boolean response determines action
- Failure propagation is an issue

A little more detail



GridFTP can't have a single system

- The system behind the Data Storage Interface can be complex, with its own authorization system
  - HPSS uses Kerberos and LDAP lookups
  - SRB maps DN to SRB credential
- Different VOs will have their own system.
- So far, this has been sufficient

Odds and Ends...

- Web services version of GridFTP
  - Reproduce the PDPs of Java container?
- NFSv4 integration of DN into file attributes sounds interesting as a solution to the dynamic account problem
  - Doesn’t address roles / attributes
- UID does provide other useful functionality
  - Quotas for instance
- Authorization of Connections
  - Have a prototype developed with UWis / Condor
  - Proposal in to improve

the globus alliance
www.globus.org

Outline

- Brief discussion of Globus Authorization in General
- Authorization in GridFTP and related tools
- **Authorization in the Globus Replica Location Service and related tools.**

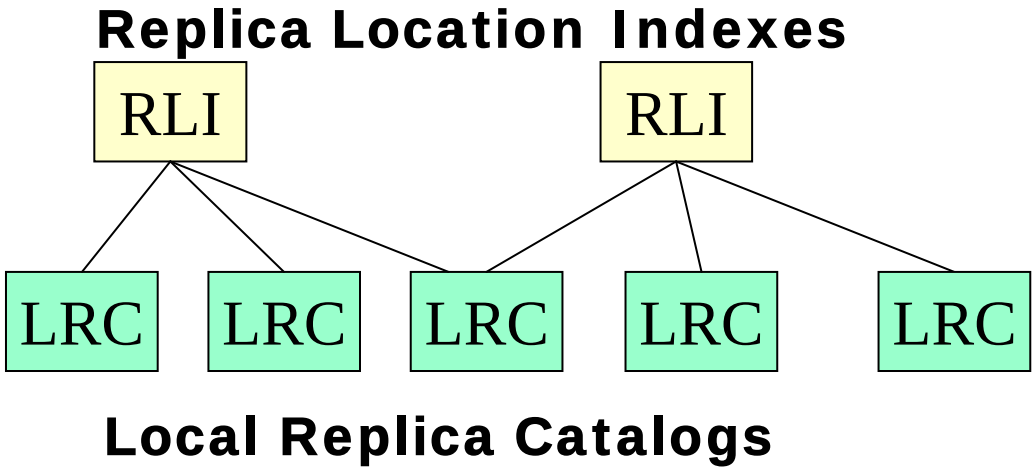
the globus alliance
www.globus.org

Globus Replica Location Service

- **A Replica Location Service (RLS)** is a distributed registry that records the locations of data copies and allows replica discovery
  - RLS maintains mappings between *logical* identifiers and *target names*
  - Must perform and scale well: support hundreds of millions of objects, hundreds of clients
- E.g., LIGO (Laser Interferometer Gravitational Wave Observatory) Project
  - RLS servers at 10 sites
  - Maintain associations between 6 million+ logical file names & 120 million physical file locations

RLS Features

- Local Replica Catalogs (LRCs) contain consistent information about logical-to-target mappings



- Replica Location Index (RLI) nodes aggregate information about one or more LRCs

- LRCs use soft state update mechanisms to inform RLIs about their state: relaxed consistency of index

- Optional compression of state updates reduces communication, CPU and storage overheads

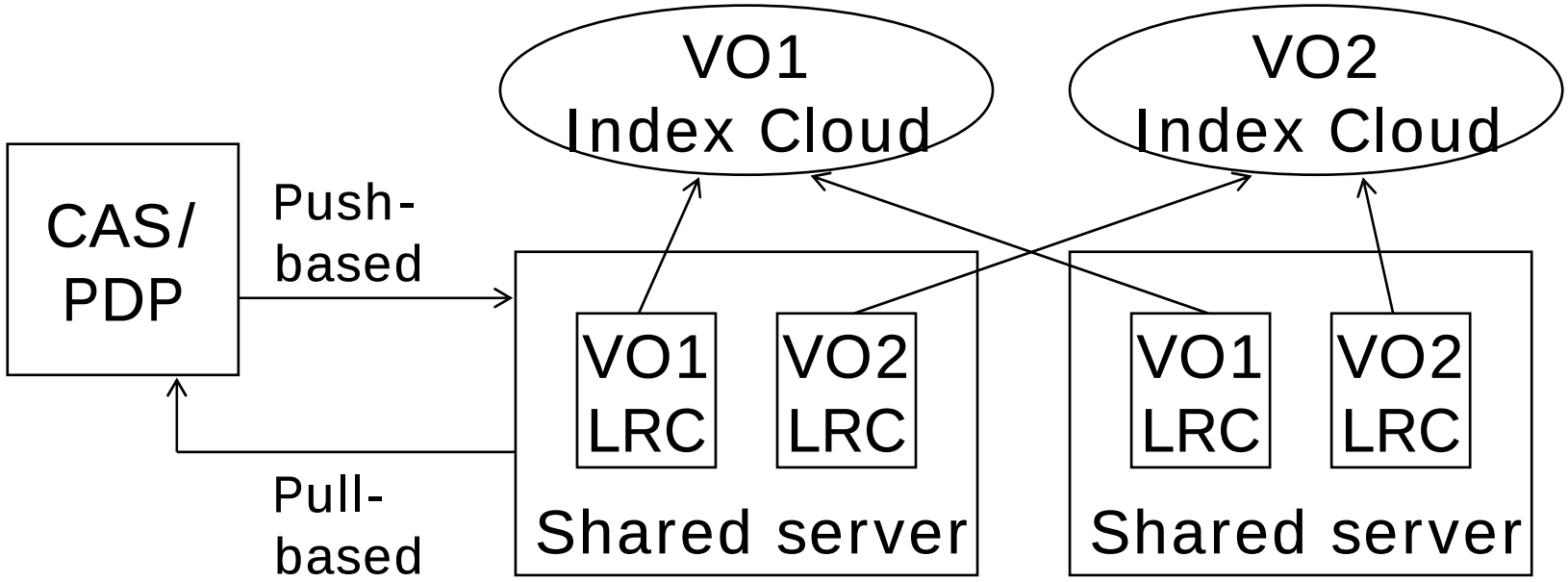
the globus alliance
www.globus.org

Current Authorization in RLS

- Gridmap files and ACLs
- Allow users to associate read, write permissions with users
- Emphasis on *scalability* and *performance*
- Problems:
  - No finer grained access control
  - Undesirable for VOs to share RLS servers, since a writer in one VO can alter entries in another
- How to provide additional functionality without destroying performance for power users?
  - Need to register and discover millions of files quickly

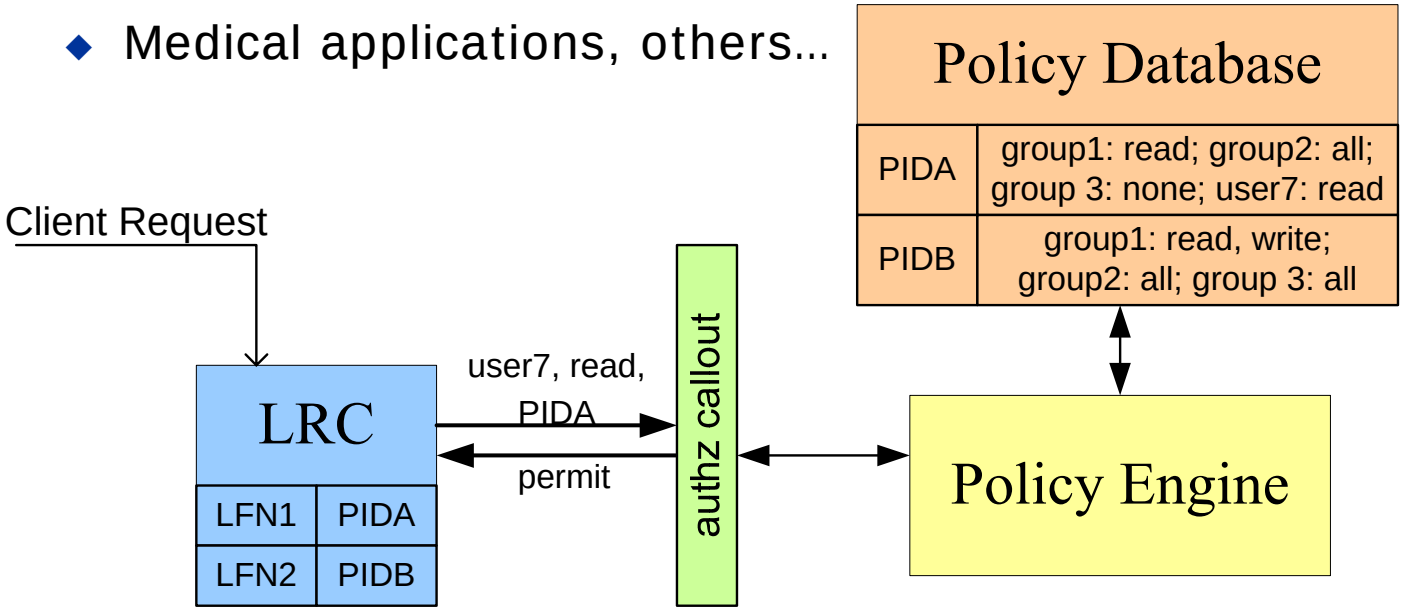
Planned Support for Multiple VOs to Coexist Safely

- Dynamic deployment of new VOs within existing LRC deployment
  - Possibly by deploying separate tables
- VO-specific index layer: lighter-weight, bitmaps in memory
- Callout to PDP for SAML authorization: identify user's VO



Longer-Term: Fine-grained Authorization

- On a per-file or group of files basis
  - LRC makes authz callout, providing operation and policy ID
  - Policy engine decides based on credentials, operations, policy DB
- Is this level of write authorization needed at the catalogs?
  - Typically a small group of privileged publishers who have broad permissions
  - Medical applications, others…



WS-RLS and DRS

- WS RLS: new Web Service interface to RLS
- Data Replication Service (DRS): combines RFT, RLS operations
  - Both WS-RF compatible interfaces
- Can make use of GT4 authorization infrastructure
  - Configure a chain of authorization mechanisms and allow plug-ins of new authorization schemes
  - Evaluate a chain of Policy Decision Points (PDPs)
  - Includes a SAML callout, mechanism for grid-mapfile authorization
  - Make authorization decisions at container level
- Support for adding custom authorization modules
  - Passes full client request message to the custom call-out
  - PDP can access the internal state of service to extract information needed to make authorization decisions