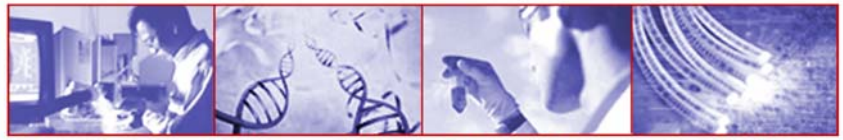
euchinagrid



Interconnection & Interoperability of Grids between Europe & China

Interoperability between EGEE gLite and CNGrid GOS

Yaodong CHENG

IHEP, Chinese Academy of Sciences

ISGC 2008

FP6-2004-Infrastructures-6-SSA-026634

Information Society and Media

Outline

- Major issues of interoperability between different grid infrastructures
- Status of our work in interoperability
- Gateway-based interoperability

Use of Grid

Write problem-solving code

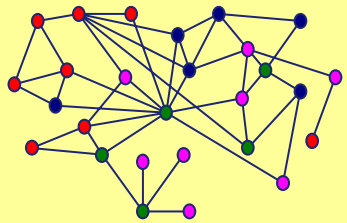
"Adapt" to middleware

Submit to Grid

security

middleware

Publish

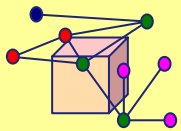


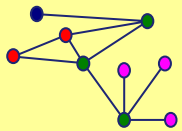
Stage data

Accounting

Dispatch to resources

Steering and visualisation





Select resources

Common issues in interoperability

- Job Description Language
- Job Submission
- Resource Discovery
- Resource Selection
- Data Staging
- Cross-domain Security

Job Description Languages

- Specify the job to run and how it will run
  - Different systems have their own job description languages
  - Choose to use the same description language or to do conversion
    - JSDL to JDL and JDL to JSDL conversion have been done in the gateway component in EUChinaGrid project
    - JSDL is a preferred job descriptor language, adopted by OGSA-BES

Condor	Complex almost programming language (ClassAds)
CNGrid GOS	Job Submission Description Language (JSDL)
EGEE gLite	Variation on the Condor ClassAds language (JDL)

Job Submission

- The way of submitting jobs to the Grid:
  - Different systems have different job submission mechanisms
  - In EUChinaGrid Project, we support interface similar to OGSA-BES and plan to provide a complete implementation of OGSA-BES;

Condor	Command line, Web Service, port, Standard DRMAA
CNGrid GOS	Portal , Web Service
EGEE gLite	Command line, API, (Some) Web Service

Resource Discovery

- Find availability of resources
  - Having a good knowledge of the current state of the resources helps in resource selection
- Three different schemas are widely used:
  - Glue Schema used by OSG, EGEE and Teragrid, mapped to LDAP, XML and the relational model and CNGrid GOS will support GLUE schema in the upcoming version
  - ARC schema used from NDGF
  - CIM schema used by NAREGI
- Use the same schema or perform necessary conversion for interoperability

Condor	Resources publish themselves to the scheduler
CNGrid GOS	Resource register themselves to router service
EGEE gLite	Resources publish themselves to an information service that the WMS can query

Resource Selection

- Select the best resources to run the job
  - Ensure that each job is placed on the most proper resource
  - A big problem for interoperability
    - Difficult to determine whether the received batch job should be dispatch to other grid middlewares or not
    - Usually resource selection is the core component of grid middleware and difficult to modify for interoperability

Condor	Jobs and resources are “matched” together. Jobs will be launched when an idle resource matching the requirements is found
CNGrid GOS	Meta Schedule choose resource according to some predefine condition
EGEE gLite	Workload Management Services are used to select the best CE to run the job

Data Staging

- Getting the data into and out of the resources
  - Data Staging interoperability focuses on the following fields
    - Point to point movement of data between storage in different grids
      - For example: Grid-ftp interoperability or OGSA-ByteIO
    - Usage of managed resources and their APIs (SRM, SRB):
      - For example: SRM interoperability

Condor	Jobs are given a virtual file space with read and write operations being passed back to the submission node
CNGrid GOS	Using FTP or HTTP as underlying transport protocols
EGEE gLite	Jobs can be staged out or provided by streams. Storage elements can hold files

Security

- Three security issues involved in grid environment
  - Authentication
    - How do we positively identify users and resources?
  - Authorisation
    - How to do the authorization operation?
  - Accounting
    - How to do the accounting operation?

Security

- **Protect underlying resources**
  - **Authentication and Authorisation are key points**
  - **Need to develop a level of trust for both users and the resource owners**
  - **Cross-domain security is a big challenge. We just made a first simple approach in EUChinaGrid Project.**

Condor	**Uses public key infrastructure x509 & Proxy**
CNGrid GOS	**Uses public key infrastructure x509 & Proxy**
EGEE gLite	**Uses public key infrastructure x509 & Proxy + Annotations on the certificates**

Overview of our work in interoperability

- Our major work
  - Design of a flexible gateway and proposal to a generic design for more complex scenarios
    - Use SEDA model as the task process tool
    - Use IoC model as the configuration and assembly tool
  - CNGrid GOS JobManager Framework extension
  - GLite LCG-CE JobManager Framework extension
- Works achieved and going on
  - First implementation of a testbed in IHEP (CAS) and in Catania (INFN)
    - running stably for about three months
    - Processed more than 1,500 batch jobs (including both GOS to GLite and GLite to GOS)
  - Focusing on data interoperability

Role of Gateway

- A logical component
  - Interface conversion
  - Function mapping
- Support the following features
  - Transparent to end users of different grid infrastructures
  - Easy to extend
  - Concurrency and high throughout
  - Standalone deployment or integrated underlying grid middleware

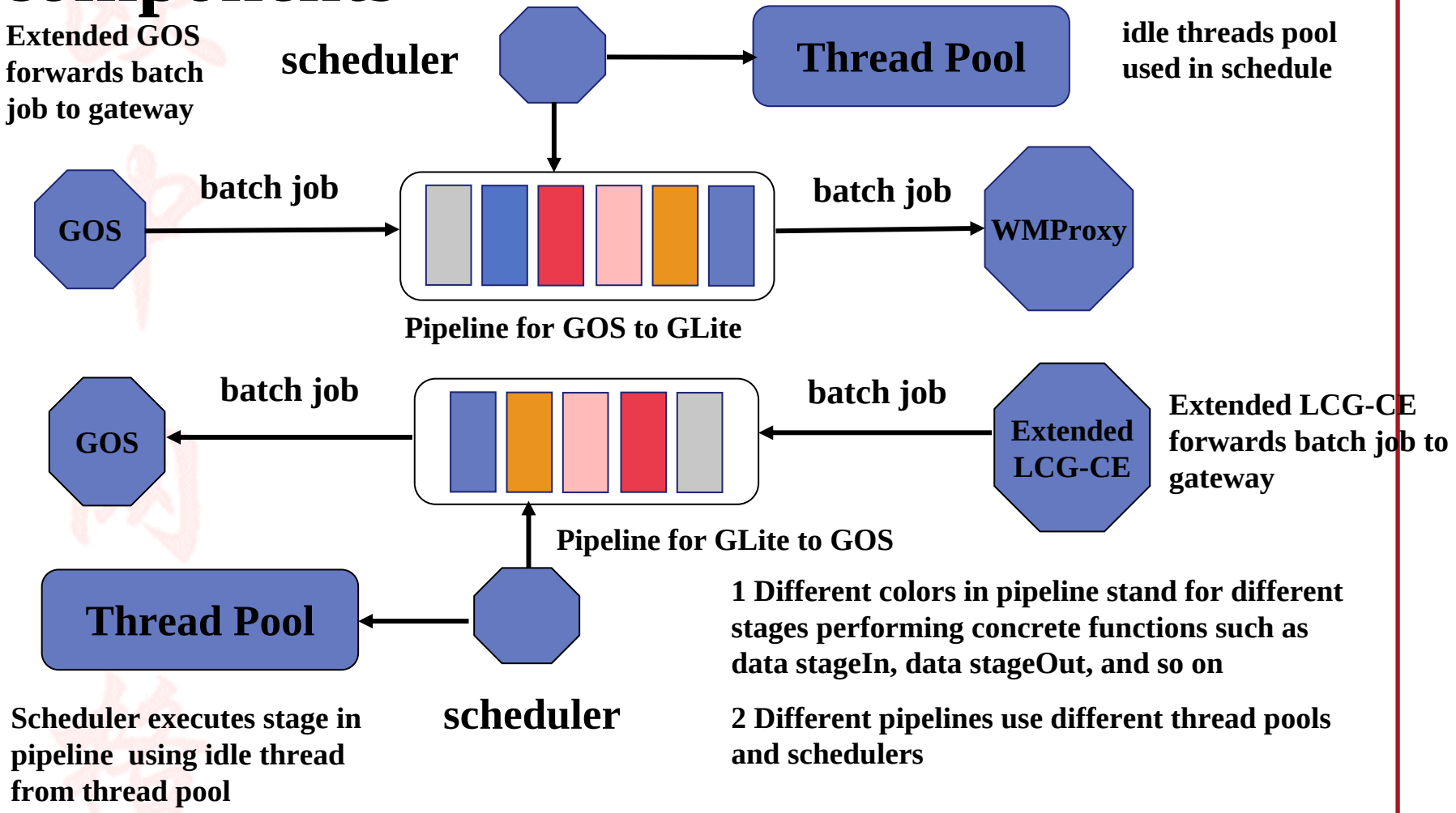
Gateway design

- Our Gateway design heavily depends on SEDA and IoC models
  - SEDA model
    - SEDA--Staged Event Driven Architecture
    - Firstly proposed by Matt Welsh, David Culler, and Eric Brewer of UC Berkeley
      - Support massive concurrency, high throughout
      - Simplify the construction of well-conditioned Internet services
    - In our design, process is divided into independent basic stages of different pipelines for different purpose such as GLite-to-GOS batch job forwarding, and so on
  - IoC model
    - IoC--Inversion of Control
      - Provide loose coupling among different modules and allow easy reuse of basic modules
      - Assemble new module easily and quickly
    - In our design, HiveMind 1.1 released under LGPL license is used as IoC container

Core components of our gateway

- Core components of our gateway
  - Pipelines for different purposes
    - Composed of different basic processing stages
    - Used for different purposes such as forwarding batch jobs from GOS to GLite and vice verse
  - Scheduler
    - Execute processing stages at fixed rate
    - One to one mapping between Pipeline and scheduler
  - Threads pool
    - Improves performance
    - One to one mapping between Thread pool and scheduler
- Processing stages in the same pipeline perform different concrete functions such as StageIn, StageOut, and so on

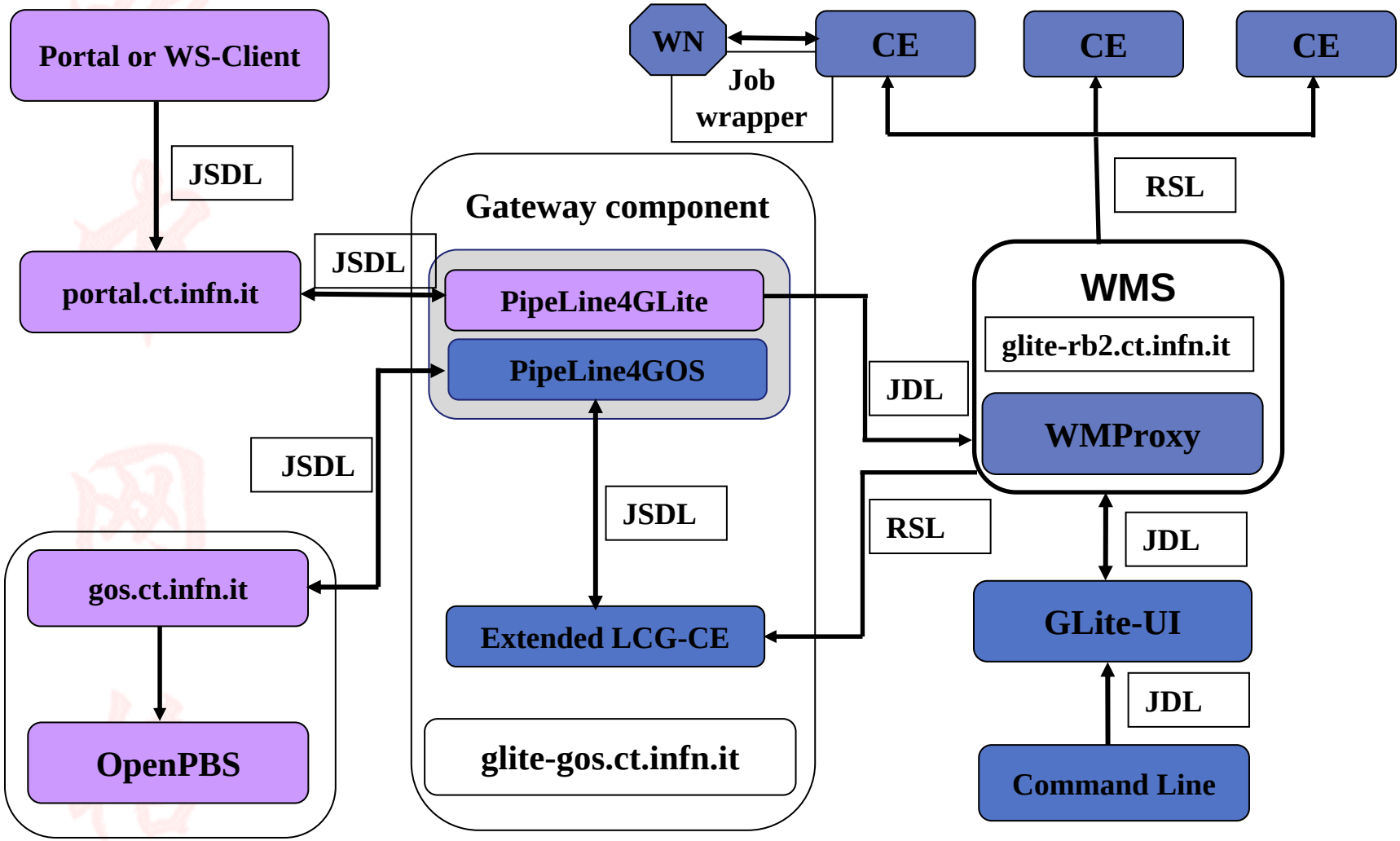
Detailed description of gateway components



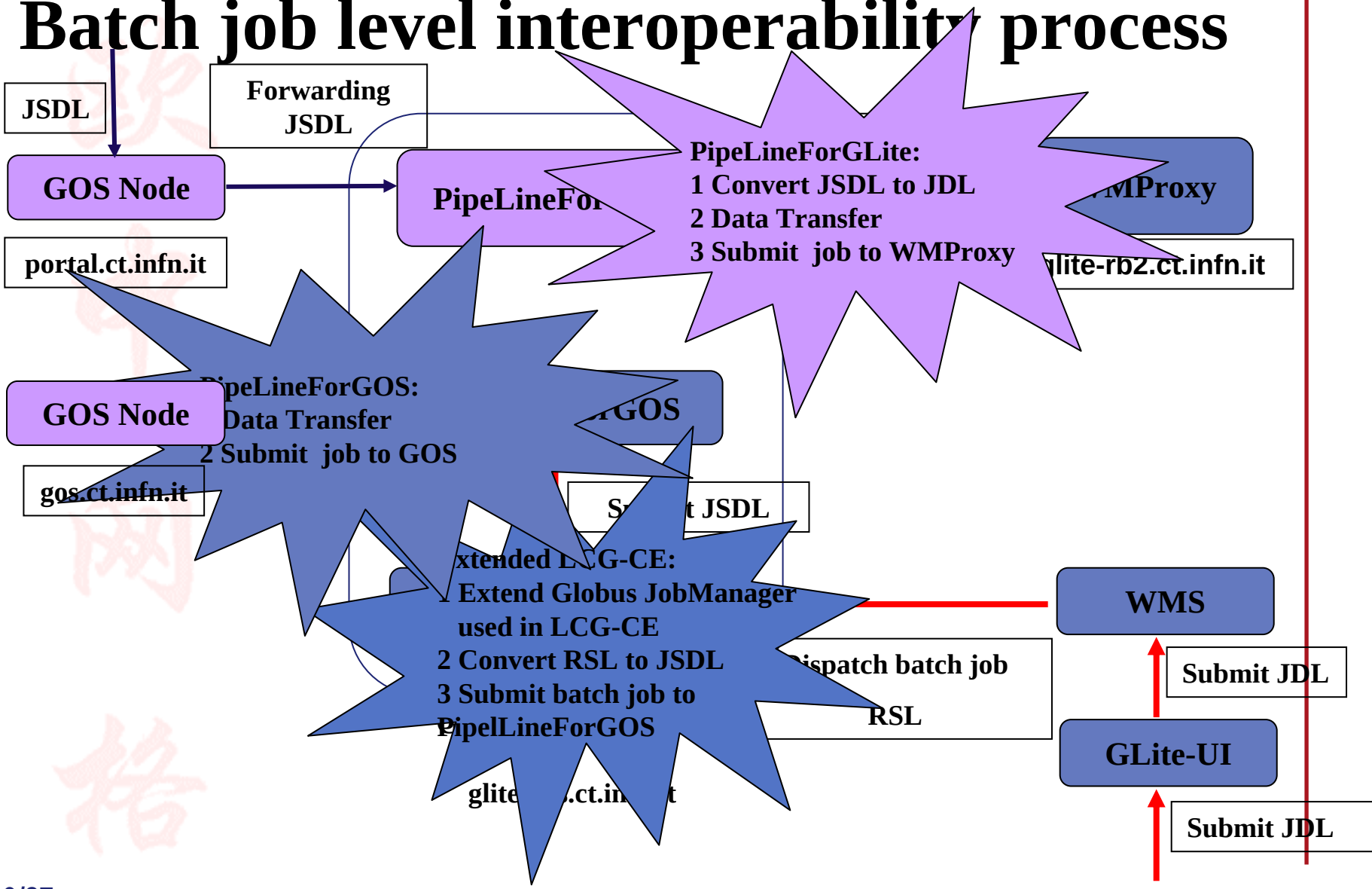
Batch job level interoperability

- Extend JobManager in both GLite and GOS
  - Extend GLite LCG-CE JobManager Framework
    - LCG-CE JobManager Framework is closely coupled with resource scheduling mechanism of GLite
    - Relatively difficult to extend, cost a lot time
  - Provide Broker plugin for GOS JobManager framework
- Sandbox mode data transfer
- A fast approach for cross-domain security scenario

Testbed in Catania, INFN



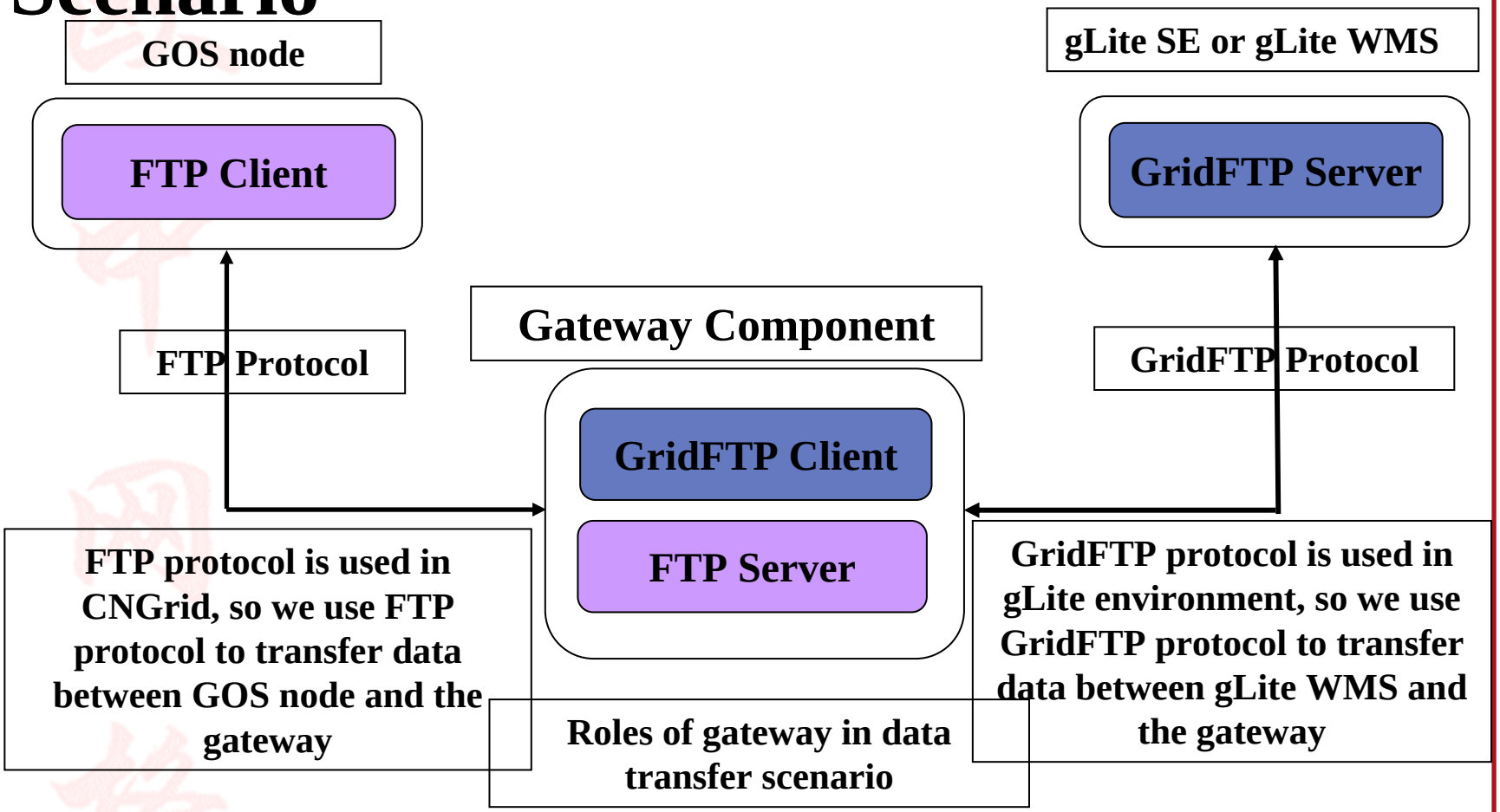
Batch job level interoperability process



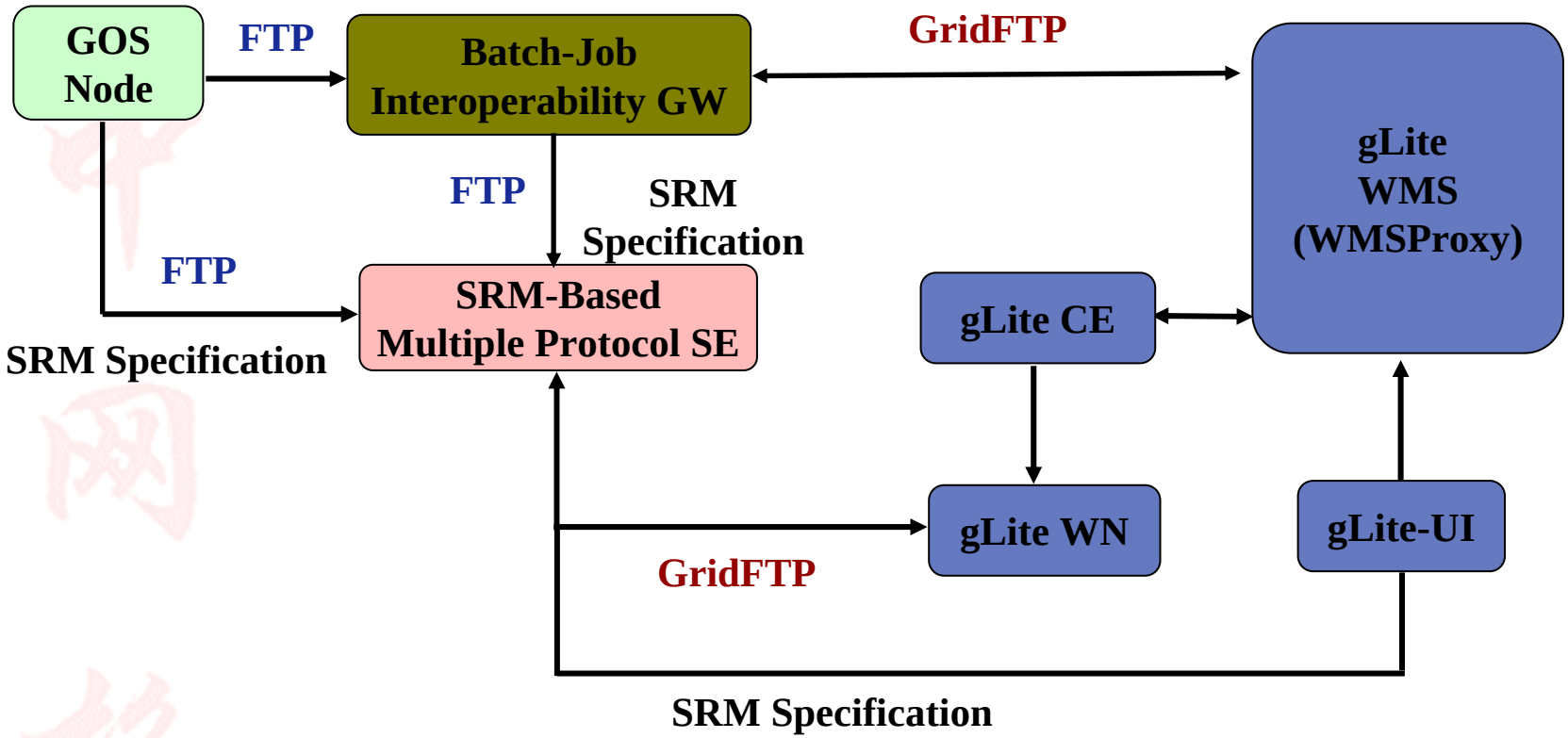
Data Transfer

- Data transfer between CNGrid and EGEE supports two different modes:
  - Sandbox-Based Data Transfer, for small scale data transfer:
    - All data transfer operations pass the batch job gateway
    - Batch job gateway acts as data transfer center and has two different roles at the same time
      - GridFTP client: Gateway can upload/download necessary data to/from GLite WMS;
      - FTP server: GOS node can upload/download data to/from gateway component;
  - SRM-Based Data Transfer, for large scale data transfer:
    - There is a separate data interoperability gateway which supports SRM specification and can be interacted through multiple protocols including GridFTP and FTP;
    - CNGrid GOS/EGEE gLite interact directly with data interoperability gateway;
      - gLite WN upload/download data files using gridFTP protocol;
      - CNGrid GOS and Batch job interoperability Gateway upload/download using SRM Specification which is based on FTP protocol;

Sandbox-Based Data Transfer Scenario



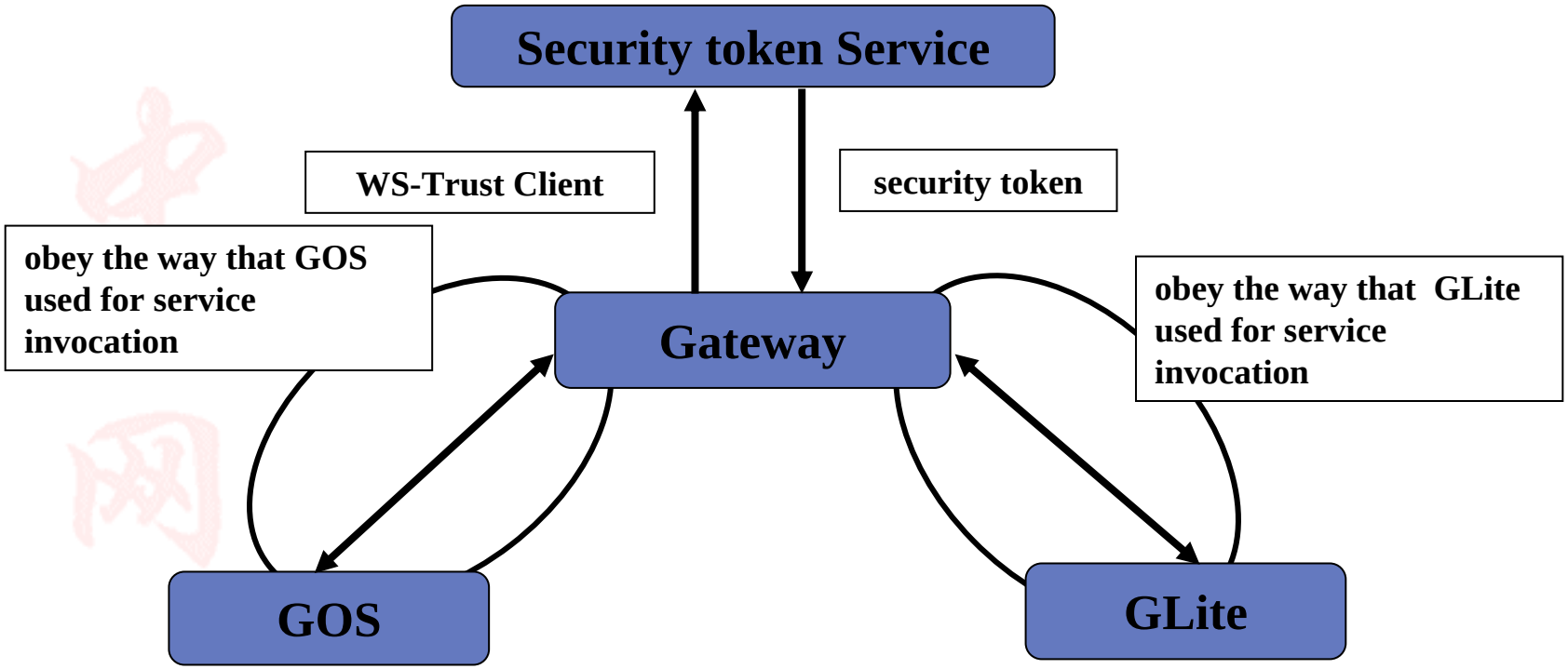
SRM-Based Data Transfer Scenario



Security issues

- A first simple approach for cross-domain security
  - Some users in GOS and GLite are predefined for interoperability purpose, a static approach
    - Requests from GOS to GLite use predefined voms proxy
    - Requests from GLite to GOS use predefined name
  - User Management module is designed to keep mapping of relationships
- Security token service
  - Used to keep, distribute, exchange and verify security tokens between GOS and GLite and provide dynamic approach
  - MyProxy Server is currently used to store temporary security token. Plan to replace it with newly developed security token service

Role of security token service



Roles of security token service in cross-domain security scenario

What to do next

- Cross-domain Security
  - Security token service based token distribution in cross-domain scenario:
    - More generic solution for cross-domain security token distribution
    - Comply with WS-Trust specification
- Comply with some work of OGF GIN Group
  - OGSA-BES
  - SRM
- Support real grid application interoperability between CNGrid and EGEE
  - POSIX (normal) application is supported now
  - Choose from applications supported by EUChinaGrid project

Thanks to Yongjian WANG, Diego Scardaci, Bingheng YAN, Gang CHEN, Giuseppe Andronico and other people involved in EUChinaGrid for their contribution to this work!

Thanks for your attentions

Any Questions?