AAI need of the DCIs in Life Sciences

Amsterdam, September 14, 2010

Johan Montagnat – CNRS

Life Sciences VRC

Security requirements in Life Sciences

- Preserve patient privacy
  - Data protection
- Protect copyrighted data processing tools
  - Specific application services protection
- Protect sensitive activity (competing industries, e.g. pharmacy companies)
  - Activity traces protection

A&A requirements

- The (user) community is completely technology agnostic!
  - The function implemented is all that matters

- Data protection
  - Authentication: patients, data owners, data users...
  - Authorization: complex authorization chain (patient -> radiologist -> hospital data manager -> physicians -> data user)
  - Data encryption: metadata (encryption keys) access control

A&A requirements

- Specific application services protection
  - Authentication: service users
  - Authorization: service developers -> service providers -> service users

- Activity traces protection
  - Authentication: monitoring service users
  - Authorization: users -> monitoring service administrators -> monitoring service users

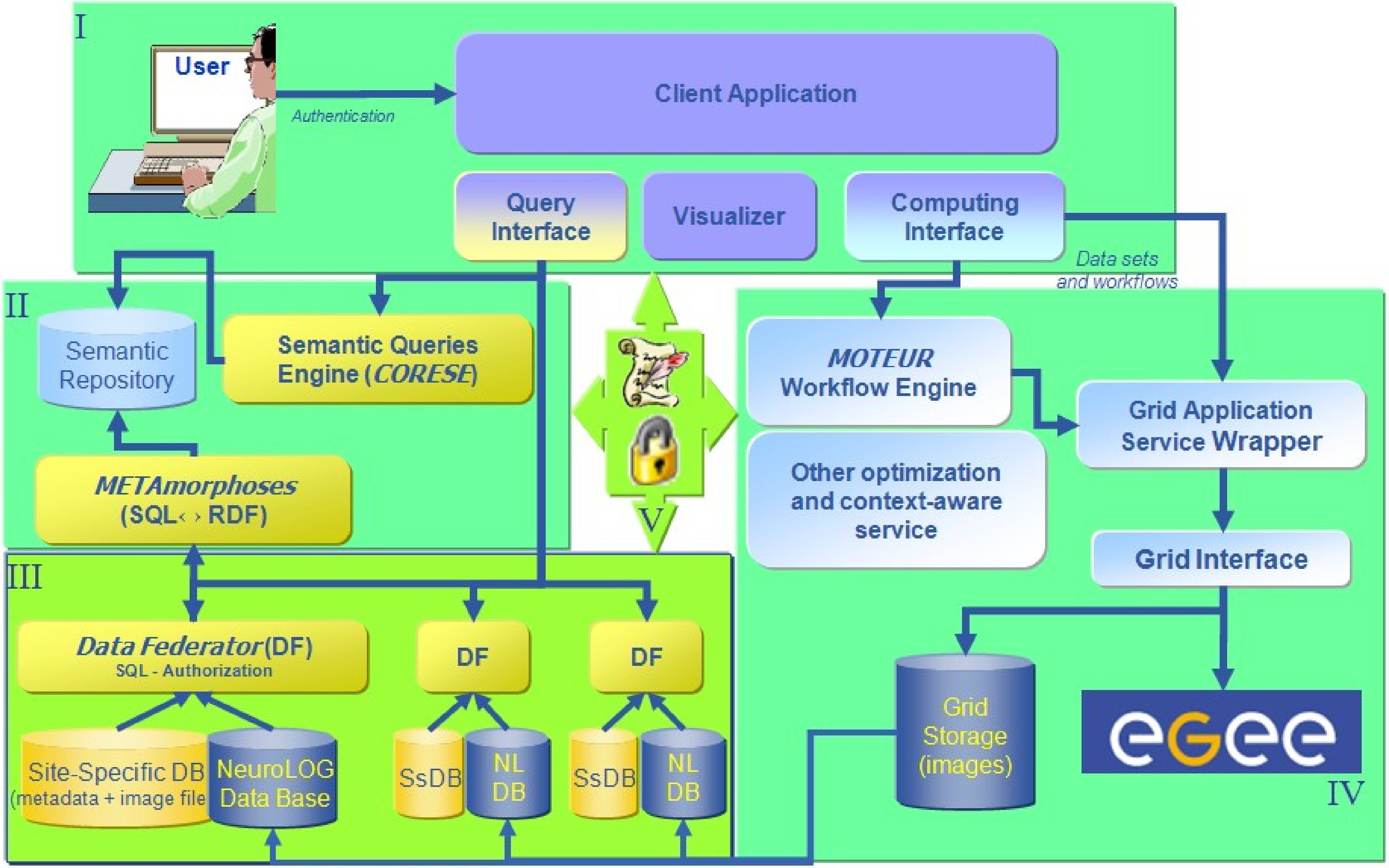
A&A requirements

- Summary
  - Authentication: individuals identification + roles
  - Authorization: shared responsibility, complex chains of roles and authorization delegation

NeuroLOG middleware example

- Neurosciences specific middleware to bridge local resources (medical image data stores) and grid resources
  - Multiple credentials + mapping
- Preserving legacy environments
  - No one-fit-all solution
- Multiple sources of data
  - Image files + associated relational metadata + extracted semantic data
  - Multiple (collaborating) data management services

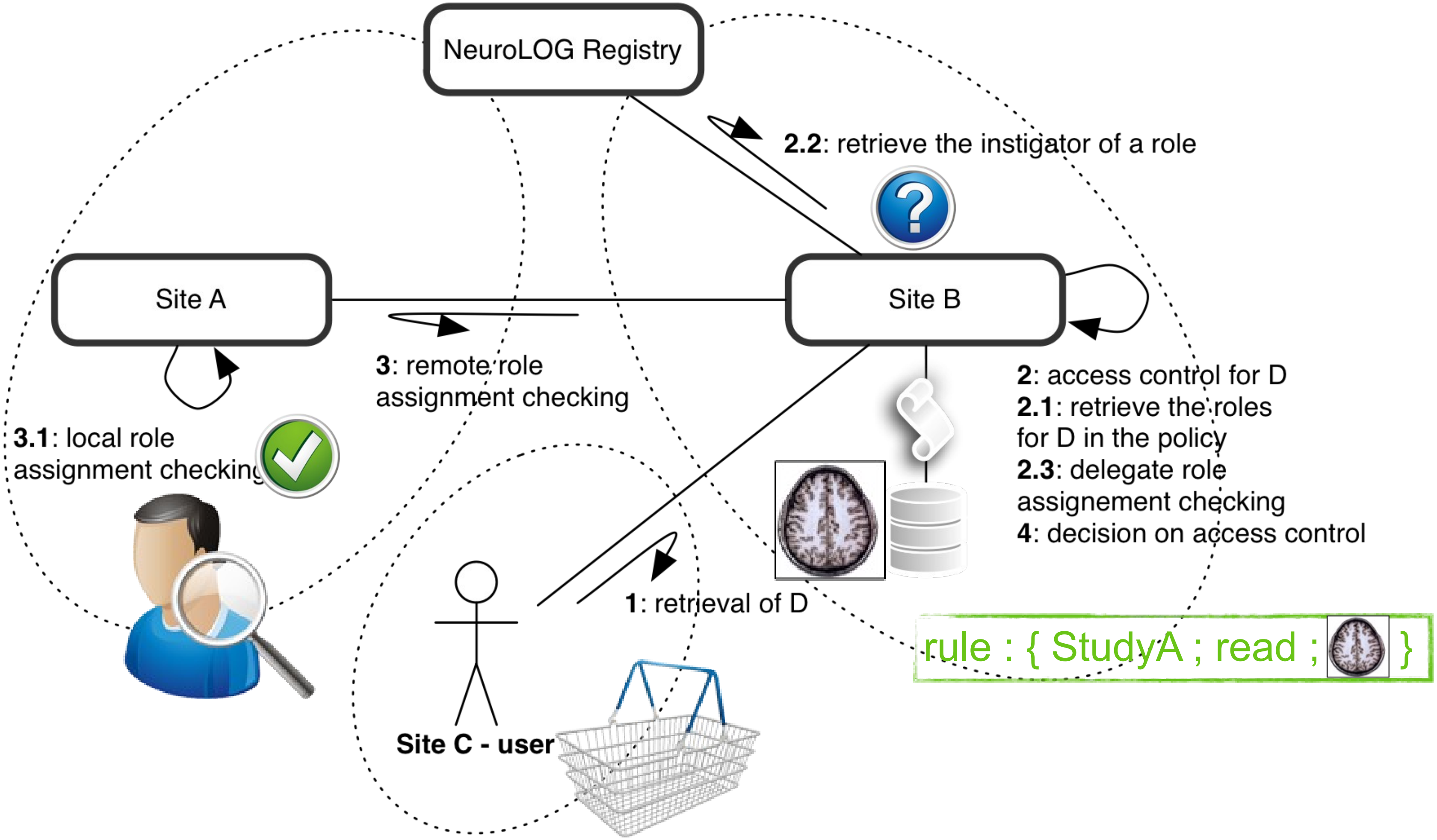
NeuroLOG middleware architecture



NeuroLOG security architecture

- Reference: A. Gaignard et al. HealthGrid'09
- Distributed access control with prevailing local site policies
  - Data owner control data access
  - No global administrator for the overall platform
- Decentralized RBAC-based access control policy
  - Data and service invocation protection
  - Support multi-centric data / service federation
  - Sites independence

Distributed access control


NeuroLOG Registry
2.2: retrieve the instigator of a role
Site A
Site B
3: remote role assignment checking
2: access control for D
2.1: retrieve the roles for D in the policy
2.3: delegate role assignement checking
4: decision on access control
3.1: local role assignment checking
1: retrieval of D
Site C - user
rule : { StudyA ; read ; }

Implementation

- Authentication
  - X509 certificates, java API to manage PKI
  - HTTPS protocol for WS middleware services
    - Apache Tomcat container configured with ciphered and mutual authenticated communication channels
- Access control
  - Extended RBAC with a database backend
  - SSL identity of users retrieved at runtime by Tomcat WS container
  - Instrumented application services

Lessons learnt

- The deployment environment has an impact on the software development
  - Code specific to the Tomcat container
- Access control to relational data is very challenging
  - No solution for fine-grained access control with multiple federated heterogeneous RDMS
- Semantic data tooling not providing data access control concerns

CPS smart-cards

- CPS = smart-cards for Health Professionals identification in France
  - Single national CA
  - Authentication: Smart-cards with X509 certificates on board
  - Authorization: Nation-wide ID control server
  - Client-side card readers + API
- Integration on-going
  - CPS identify internally mapped on NeuroLOG identities

Questions and Answers

- 1. How are users currently authenticated
  - 1.1. which credential(s) is/are used?

    X509 (both grid users and French Health Professionals smartcards)
  - 1.2. how is the user vetting done?

    RBAC-style (NeuroLOG RBAC-based distributed authZ)

    Difficulties to set up access control for relational / semantic data stores
- 2. Is there a link to national identities? If so, how are different national identities leveraged?

  Health Professionals CPS smartcards

Questions and Answers

- 3. Which types of resources are in use and how are users authorized?
  - 3.1. Resources accessed through Grid technology: computing resources, storage, etc...

    EGI: storage resources (SRM), unequally supporting ACLs
  - 3.2. Resources accessed without Grid technology: computing resources, storage, etc...

    External data repositories (any authentication mechanism)
  - 3.3. web-based resources

    Web Services over HTTPS

Questions and Answers

- 4. Where does the project want to be in ~5 years with regards to authentication and authorization

  Homogeneous handling of AA in grid services

  Access control to relational stores

  Access contol to semantic stores

- 5. Are your users and resource owners happy with the current AAI scheme that you use?

  Scheme is irrelevant. Only functionality matters.

  Dedicated solutions often needed in Life Sciences