UMD Roadmap & Sustainability

e-infrastructure

Software Sustainability

1. Functional classification

    –   What functionality/capability does the software provide?

2. Deployment classification

    –   Who covers the costs for deployment and administration of the software/service?

3. Maintenance classification

    –   Who covers the costs for software/service development and maintenance?

Functional classification

1. Security   - Authentication, Authorization, etc.

2. Information   - Discovery, Model, Messaging

3. Operations   - Accounting, Monitoring

4. Storage   - Management, Access, Transfer, Scheduling

5. Data   - Databases, Metadata

6. Compute   - Job exec, Parallel Job, Scheduling, etc.

7. Virtualisation - VM Management, Images, etc.

8. Instrumentation - Remote instrumentation

9. Clients   - Client Tools and APIs

Deployment classification

- **Core Capability**
  - Critical to the operation of EGI's production infrastructure
  - Deployed by EGI (as a global service?)

- **Resource Capability**
  - Exposes EGI resources to all EGI user communities
  - Deployed & managed by EGI resource centres

- **Community Capability**
  - Critical to a particular user community
  - Deployed & managed by the pertinent user community

- **User Capability**
  - Individual users need this capability
  - Deployed and used by individual users

Maintenance classification

- Generic software
  - POTS software (e.g. Oracle, Apache httpd)
  - No coupling to EGI

- Custom infrastructure
  - Originates within EGI, driven by requirements to sustain the EGI production infrastructure
  - Maintenance by select providers mandated by EGI

- User community
  - Domain-specific software
  - Maintenance covered by community specific sources

- Other
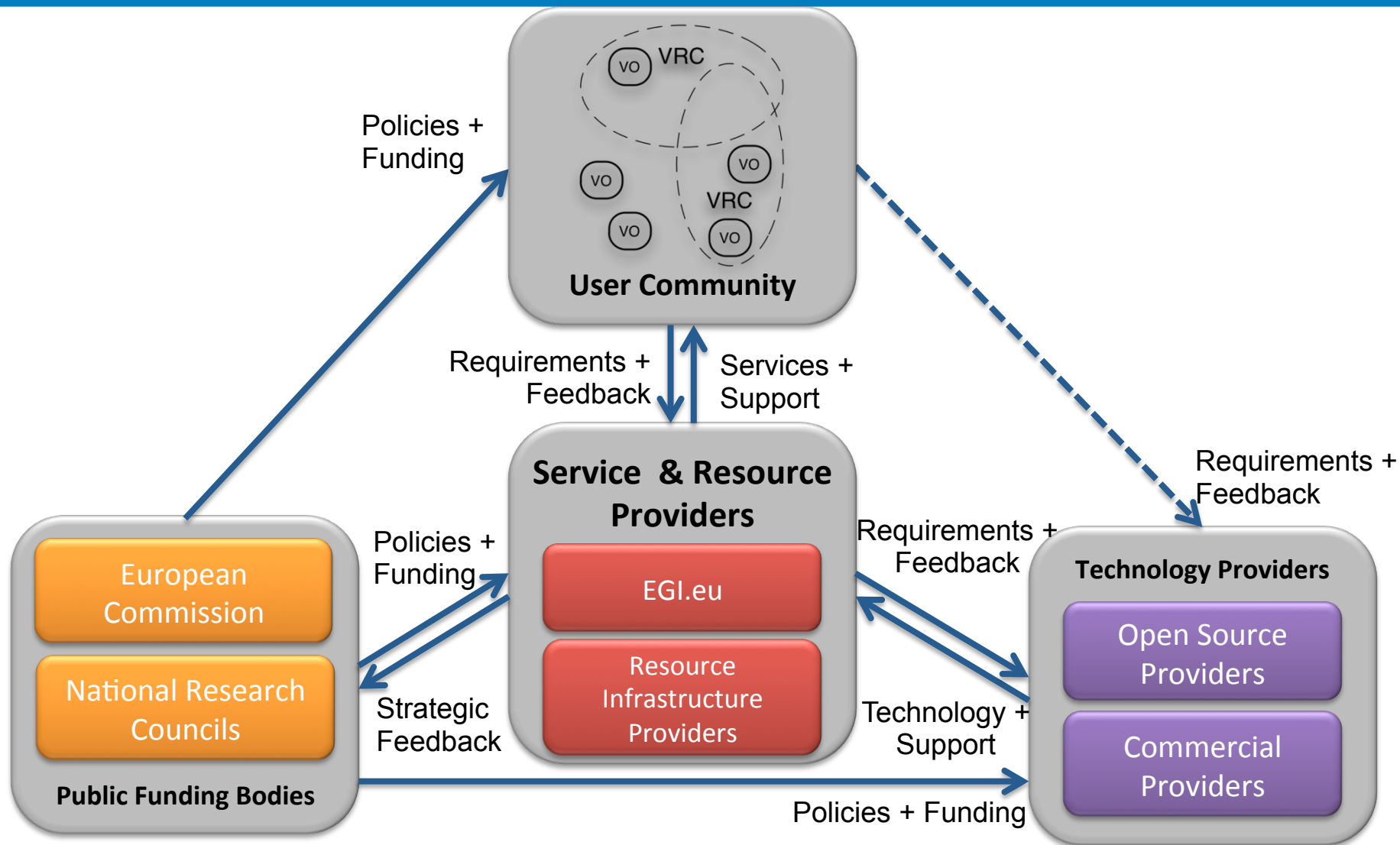 - No clear categorisation of this software is possible
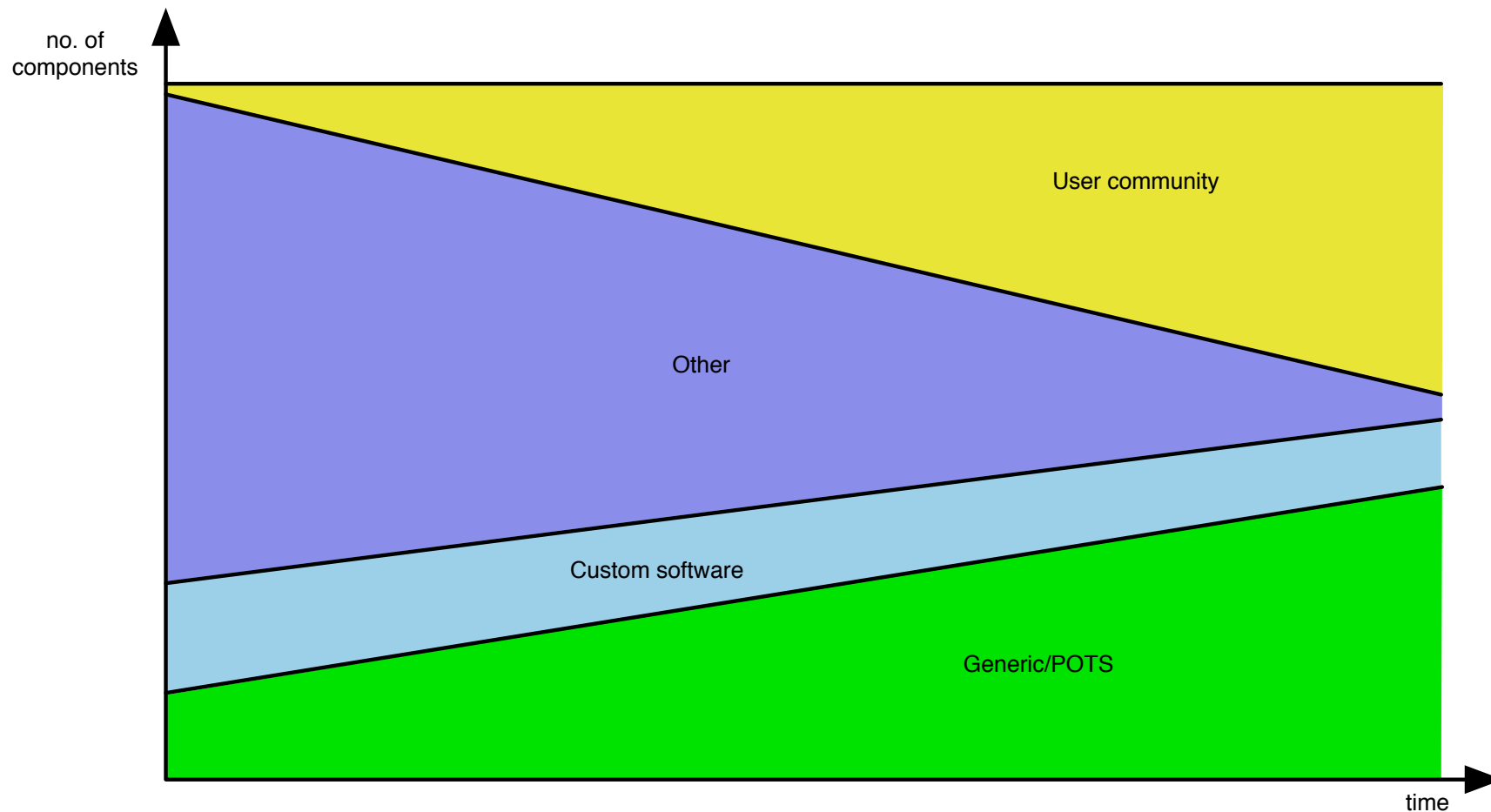
Maintenance affinities

Deployment scope	Generic	Custom	User community	Other
Core	+++	++	--	----
Resource	++		+++	----
Community	+		+++	--
User			+++	-

Security Capabilities

- **Authentication**
  - PKI/X.509
  - Shibboleth/OpenID?
  - SAML for users?
- **Authorisation**
  - Very diverse
  - ARGUS emerges as core candidate?
  - SAML/XACML based!
- **Attribute Authority**
  - VOMS is core candidate
    - Oracle integration issues!
  - SAML/RFC3281 based!

Security	Maintenance			
Deployment Scope	Generic	Custom	User	Other
Core				
Resource	AuthN	Attr.A / Cred. Mgmt / AuthZ		
Community				
User				
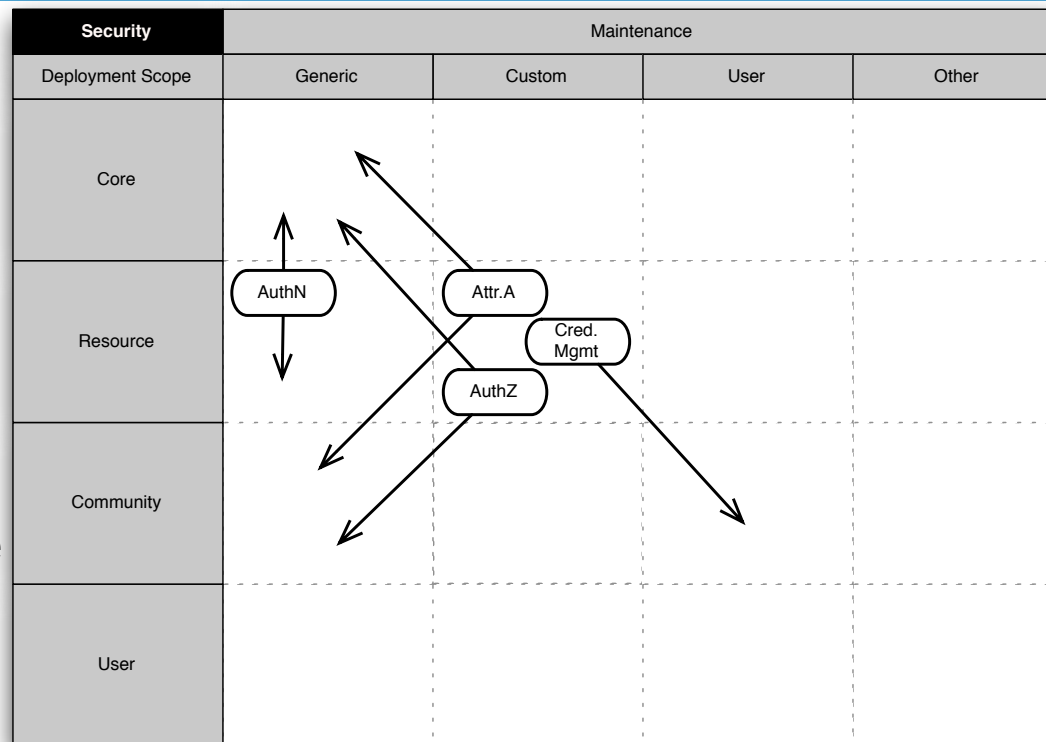
- **Credential Management**
  - MyProxy is a mature solution
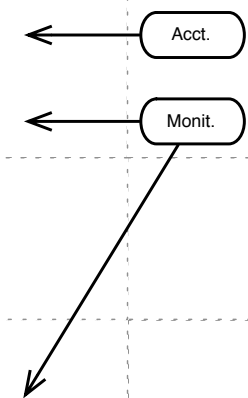
Information Capabilities

- **Information Model**
  - Converge to GLUE 2.0
  - Extension where required
  - EGI core more conservative

- **Information Discovery**
  - EGI core more conservative
  - Start with BDII
  - Explore existing generic solutions

- **Messaging**
  - Stable ApacheMQ
  - Migration to AMQP, STOMP?

Information	Maintenance			
Deployment Scope	Generic	Custom	User	Other
Core	Model / Disco very / Messa ging			
Resource	Model / Disco very			
Community				
User				

  - EGI core and communities may diverge!

Operations Capabilities
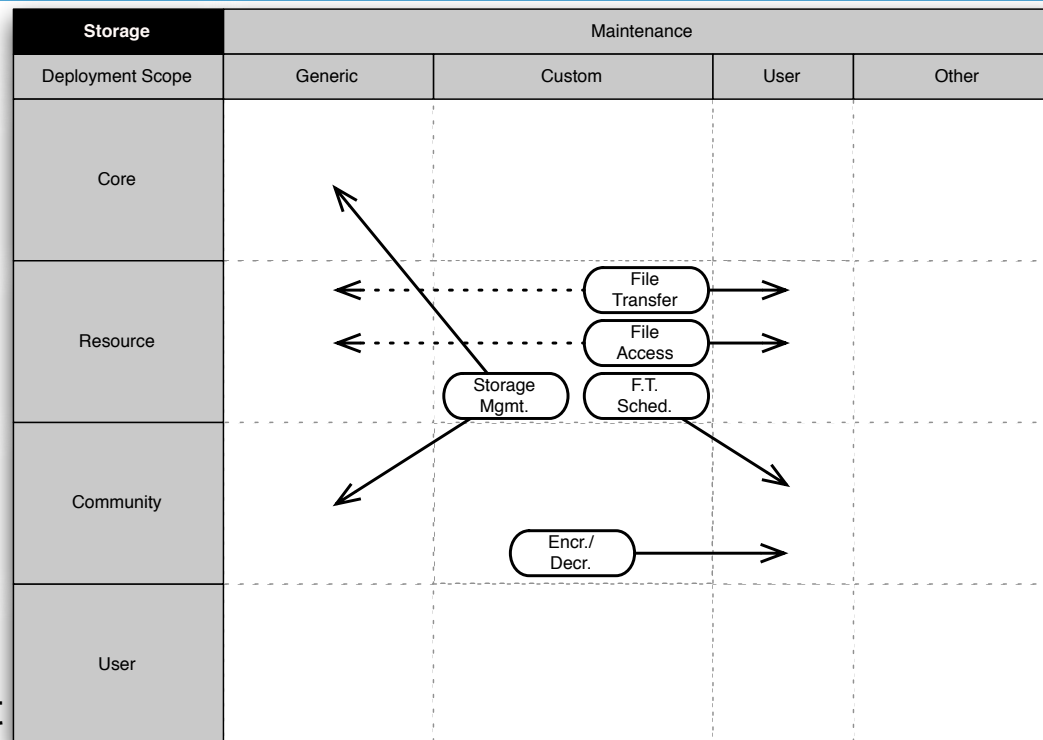
- **Accounting**
  - Based on generic DB
  - Custom reporting and portals
  - Need to migrate to configurable generic solutions

- **Monitoring**
  - Nagios framework
  - Nagios plugins: custom!
  - Interface: Nagios plugin specification
  - Sustainability pressure on plugin development/availability

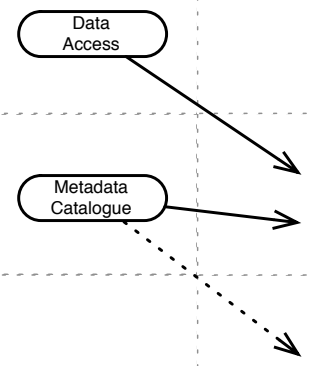| Security	Maintenance			
Deployment Scope	Generic	Custom	User	Other
Core	← Acct. ← Monit.			
Resource				
Community				
User				

Storage Capabilities

- **File Transfer**
  - Globus GridFTP
  - Parallelised HTTP?

- **File Access**
  - POSIX for local access
  - WebDAV for remote access
  - DPM, StoRM in UMD
  - dCache not assessed yet

- **Storage Management**
  - SRMv2 based management
  - Remains distributed, per resource
  - Centralised storage management?

- **File Transfer Scheduling**
  - FTS reloaded?
  - Must interface with File Access and File Transfer

Storage	Maintenance			
Deployment Scope	Generic	Custom	User	Other
Core				
Resource		File Transfer / File Access / Storage Mgmt. / F.T. Sched.		
Community		Encr./Decr.		
User				

- **Encryption/Decryption**
  - Key management is community specific
  - Should be transparent to the resources!

Date(base) Capabilities

- **Data Access**
  - WS-DAI based?
  - Implemented by storage nodes?
  - Data structure is domain specific!

- **Metadata catalogue**
  - Communities may share data
  - But the data taxonomy may differ!
  - AMGA, LFC – complementary?
    - LFC/ Oracle integration issues!

Data	Maintenance			
Deployment Scope	Generic	Custom	User	Other
Core				
Resource		Data Access		
Community		Metadata Catalogue		
User				

Compute Capabilities

- **Job Execution**
  - OGSA-BES, JSDL, HPC Basic Profile
  - Emi Execution Service?
  - CREAM, GRAM5
  - Boundary to LRMS?
  - Worker node concept too static?

- **Parallel Job**
  - Tight coupling
    - MPI, OpenMP
    - CREAM, GRAM5
  - Loose coupling/asynch
    - High-throughput messaging?

- **Job Scheduling**
  - WMS – twice rejected from UMD!
  - Gridway as alternative?

Compute	Maintenance			
Deployment Scope	Generic	Custom	User	Other
Core				
Resource	Job Exec / Parall. Job / Job Sched.			
Community	Inter. Job Mgmt / Workfl.			
User				
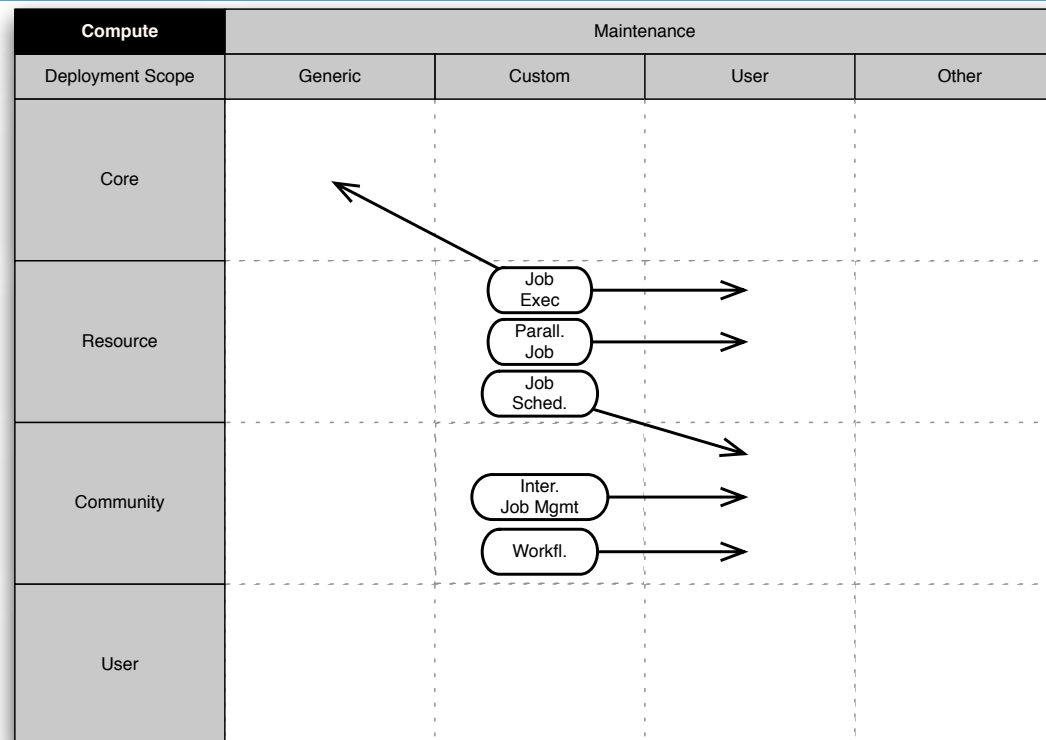
- **Workflows**
  - Very community specific! → SHIWA?

- **Interactive Job Mgmt.**
  - SSH interface suitable for anything beyond job start/stop/restart etc.

Virtualisation Capabilities

- **VM Management**
  - Services: OGF OCCI!
  - OpenNebula,
    OpenStack,
    Eucalyptus,
    VMWare vCenter etc.?
  - Clients: Portals? Shells?

- **VM Image Format**
  - DMTF OVM container
  - Different Hypervisor formats
    - Linux KVM, VMWare vDisk
    - ISO FS, etc.

- **VM Image Distribution**
  - Separate images from metadata
    - Requires permalinks for images!
  - Content delivery Network (CDN) for images

Virtualisation	Maintenance			
Deployment Scope	Generic	Custom	User	Other
Core	VM Mgmt. / VM Image Format / VM Image Distrib.			
Resource				
Community				
User				

- StratusLab Marketplace as EGI global service?
- AMGA for community specific service?

- **Client API**
  - Targets grid applications that access the Grid
  - Most effective if deployed in the infrastructure
  - SAGA portfolio available
    - Requires service plug-ins

- **Client Tools**
  - Command line client tools
  - Shells
  - EMI UI, HILA, etc.

- **Remote Instrumentation**
  - Instruments are not exposed as EGI resources
  - Instruments are often community specific tools!

Instrumentation & Clients	Maintenance			
Deployment Scope	Generic	Custom	User	Other
Core				
Resource		Client API →		
Community		Remote Instr.		
User		Client Tools →		