



d4SCIENCE

AHM 2008
10 September 2008
Edinburgh

High-level e-Infrastructures the gCube Approach

Richard Connor
University of Strathclyde

- cross-discipline research collaborations
 - among autonomous and widely dispersed parties
- may rely on data/processing intensive workflows
 - e.g. 'conventional' heavy crunching of big numbers
 - e.g. synthesis and lifetime management of complex digital artefacts from cross-media sources
- may rely on high-level interactions between users and content
 - e.g. classes of structured and unstructured searches across heterogeneous sources
 - e.g. content, metadata, and annotation management
- may stretch across full life-cycle of scientific enquiry
 - from observation to literature to observation: instruments, data, information, knowledge

- hard (functional and non-functional) requirements
 - cost-effectiveness, security, autonomic adaptation to highly dynamic environments
 - beyond commodity technologies (or ad-hoc developments)
- true Virtual Research Environments (VRE) require infrastructural support
 - provision & sharing of human, hardware, and software resources (as per Grid vision)
- scientific infrastructures require holistic sharing
 - beyond cycles, storage, programs, and data (OGSA domain)
 - system, domain, and application services also a fully virtualised commodity within the infrastructure
- a tall order for infrastructures and control systems
 - dynamic sharing of services well beyond current state of the art

- layered upon lower-level ones
 - layers correspond to extensions to the scope of sharing (and historical evolutions of Grid views)
 - provide or delegate to lower layers, based on type of sharing (e.g. to execute programs over Grid meta-computer)
- controlled by equally high-level systems
 - fully service-oriented, rich in content-oriented functionality
 - open to 3rd parties as development and runtime platforms
- offer comprehensive VRE support
 - (definition) interactively select, configure, and orchestrate required services
 - (deployment) place service instances strategically and transparently across available hardware
 - (monitoring) re-deploy instances dynamically according to QoS & failures
 - (scoping) let instances operate securely across multiple VREs

- engaged in high-level infrastructure building since 2004
 - international consortium of 11 partners with partly-EU funding
- we build atop EGEE for shared hardware, data, and programs
 - and in turn GEANT for shared connections
- we offer WSRF renderings of conventional DL services
 - to collate, describe, annotate, transform, index, and search cross-media content
- we organise them in a SOA and group them into extensible runtime frameworks
 - open to 3rd party development
- we offer transparencies for adoption, service development, and service usage
 - client libraries and a dedicated container+application framework built on Globus technology (**gCore**)
- we support full VRE life-cycle management
 - with a state-of-the art autonomic system for coordinated and dynamic sharing of services: **gCube**

- an ensemble of more than 140 services architecturally distributed across three layers
 - *presentation*: mediate between users and services
 - *information*: implement content management functionality
 - *control*: implement service management functionality
- presentation through portlets for dynamic composition of portals
- content-management via stack of services...
 - rooted in unifying information model of binary relationships laid upon storage replication and distribution services
 - services higher up in stack specialise the semantics of relationships (has-a, is-part-of, is-described-by, is-annotated-by...)
- ...and runtime search framework
 - over federation of forward, inverted, and geo-spatial indices

- VRE management services
 - host service implementations in repositories
 - translate interactive VRE definitions into declarative specifications for their deployment & runtime maintenance
 - implement the specification during VRE lifetime
- security services
 - extend lower-level technologies for the authentication and authorisation of services (rather than users)
 - including transparent renewal of service credentials
- information services
 - gather information about all current aspects of the current state of the infrastructure
 - organise it into a peer-to-peer network.
- process management services
 - consume interactively defined workflows of service invocations
 - distribute the optimisation, monitoring, and execution of individual steps across the infrastructure

- in **Diligent** (2004-2007) we built a testbed infrastructure
 - tested its services with collaborations from Cultural Heritage & Environmental Monitoring
- in **D4Science** (2008-2010) we are moving to production
 - and support additional collaborations from Fishery & Aquaculture Management
- in the future we will be turning the infrastructure in the pivotal element of a knowledge ecosystem of interoperable infrastructures
 - and will expand into HEP, Earth Observation, and cross-discipline federated repositories

- ERCIM  ERCIM
- National and Kapodistrian University of Athens 
- Consiglio Nazionale delle Ricerche 
ISTITUTO DI SCIENZA E TECNOLOGIE
DELL'INFORMAZIONE "A. FAEDO"
- European Organization for Nuclear Research 
- Engineering - Ingegneria Informatica SpA 
- University of Strathclyde 
University of
Strathclyde
- Universitat Basel 
UNI
BASEL
- European Space Agency 
- The Food and Agriculture Organization of the United Nations 
- International Center for Living Aquatic Resources Management 
- 4D SOFT Számítástechnikai Kft 