Sixth Framework Programme

IST-5-033437

Chemomentum

Chemomentum
UNICORE 6 based infrastructure for complex applications in science and technology

Bernd Schuller
Forschungszentrum Jülich

Forschungszentrum Jülich
in der Helmholtz-Gemeinschaft

Collaborators

- Piotr Bala, Krzysztof Benedyczak, Al. Nowinski
- Bastian Demuth, Henning Mersch, Valentina Huber
- Uko Maran, Sulev Sild
- Hartmut Mix, Katharina Rasch, Robert Schoene
- Mathilde Romberg, Vitaliy Ostropytskyy
- Kim Baldridge, Wibke Sudholt
- Emilio Benfenati, Mosé Casalegno

- Johann Lauthier, Nadège Piclin, Marco Pintore
- Enrico del Grosso

Outline

- Project overview
- Workflow
- VO management
- Clients
- Summary and outlook

Technical project vision

Chemomentum

- Support for workflow-centric, complex applications
- Data and knowledge management
- Targeting a wide range of end-users with widely varying IT and Grid expertise
- Easy to use, install, configure, administrate
- Generic system, applicable in a wide range of application domains

Scenario: REACH

„Registration, Evaluation, and Authorization of CHemicals“

EC Legislation, valid from June 2007

30000 substances (volume > 1 ton/year) have to be comprehensively tested for possible risks to human health and the environment,

resulting in....

Cost: **3.5 billion Euro + 3.9 million lab animals**

E.Benfenati, M. Casalegno (Inst. Mario Negri)

„In silico“ methods for REACH

Chemomentum

Can this cost be reduced?

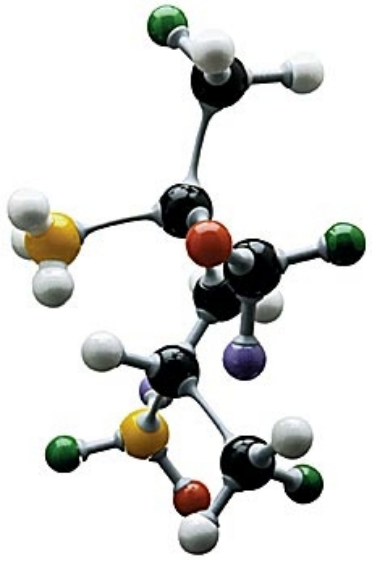
YES.

The use of „in silico“ methods such as **QSAR**, could save 1.3-1.9 million lab animals

E.Benfenati, M. Casalegno (Inst. Mario Negri)

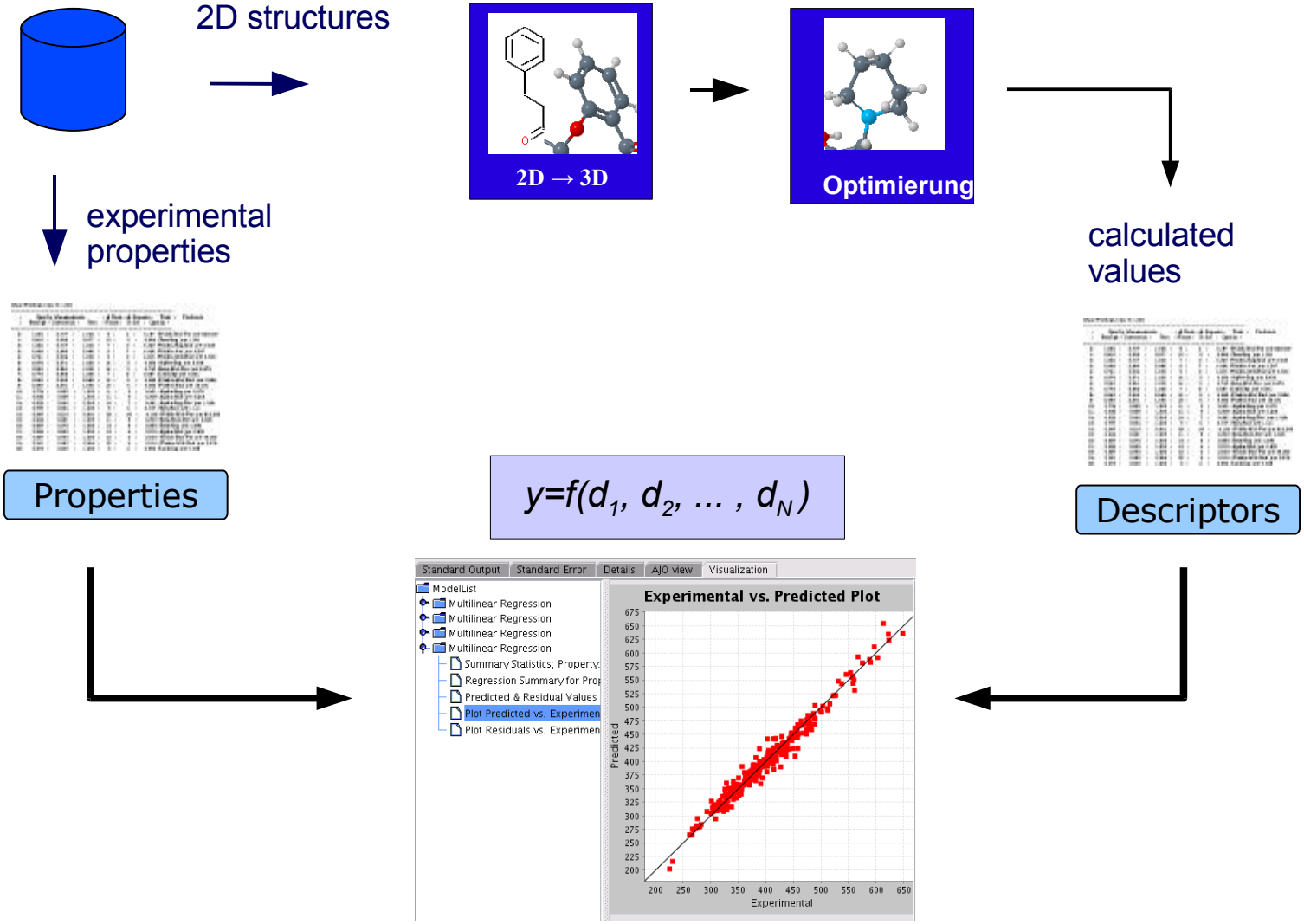
Quantitative Structure-Activity Relationship

- establish *statistical correlations* between *structural descriptors* and *experimental properties* of chemical substances



Name	Type	Value
Total number of atoms	constitutional	38.0
Number of C atoms	constitutional	14.0
Number of H atoms	constitutional	20.0
Average Complementary Informatio...	topological	2.23675377
Bonding Information content (order 1)	topological	21.8037696
Average Bonding Information conte...	topological	0.57378341
Positively Charged Surface Area (Ze...	charge-related	336.40959
Positively Charged Partial Surface A...	charge-related	0.691687523
Positively Charged Part of Charged ...	charge-related	7.70101071

QSAR: workflow

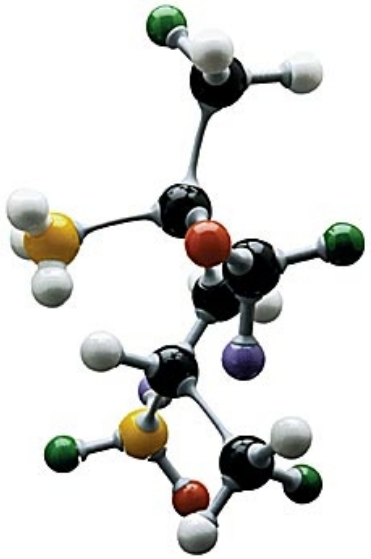


QSAR for risk assessment

QSAR could be used in the frame of REACH ...

Challenges of the available QSAR tools, e.g.

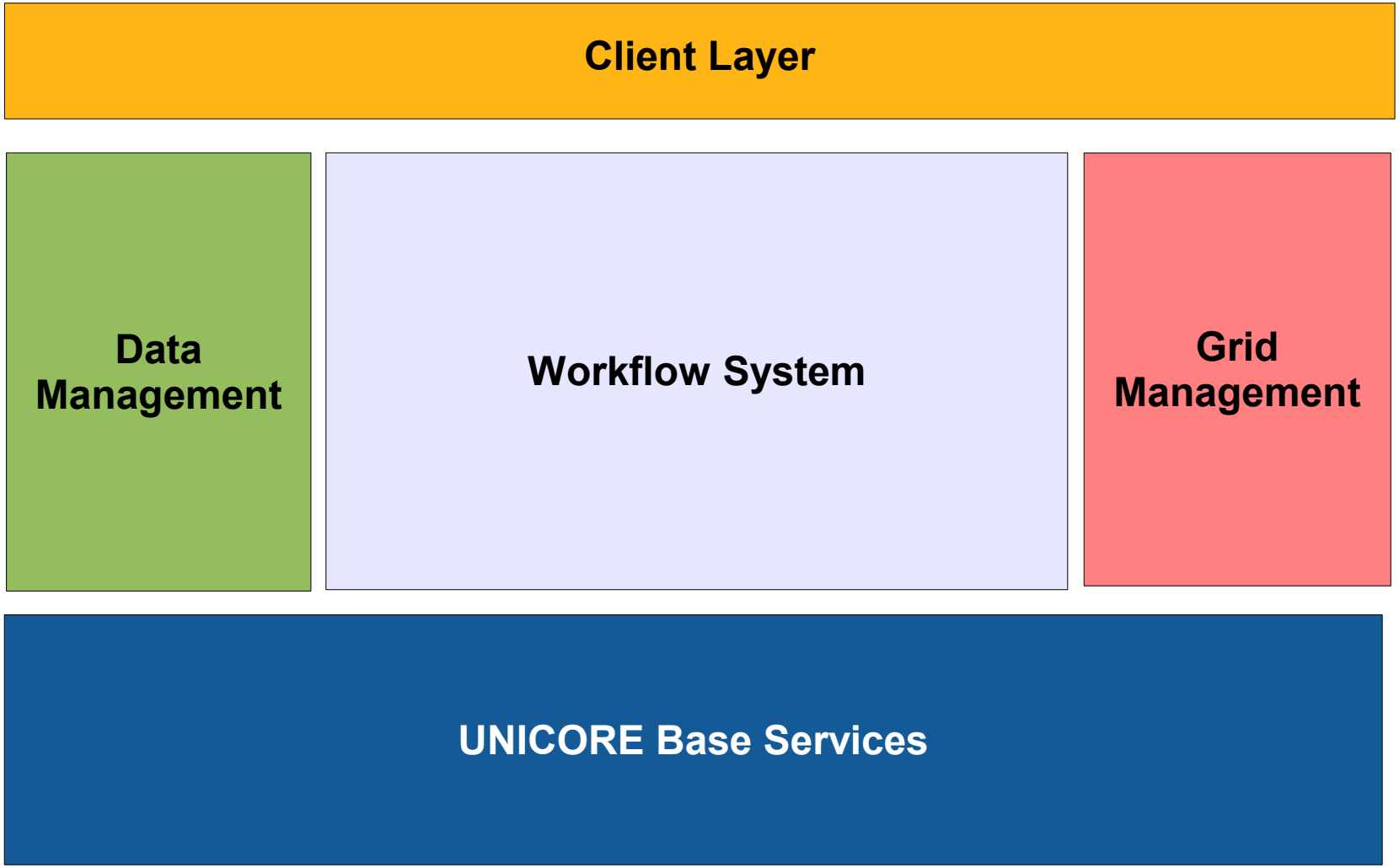
- Reproducability: Different users, different results?
- Scalability: Deal with large data sets
- Automation: Currently workflow is user-controlled
- Standardisation: No generally accepted procedures
- User friendliness: Expert knowledge needed
- Provenance: How exactly was the model produced?



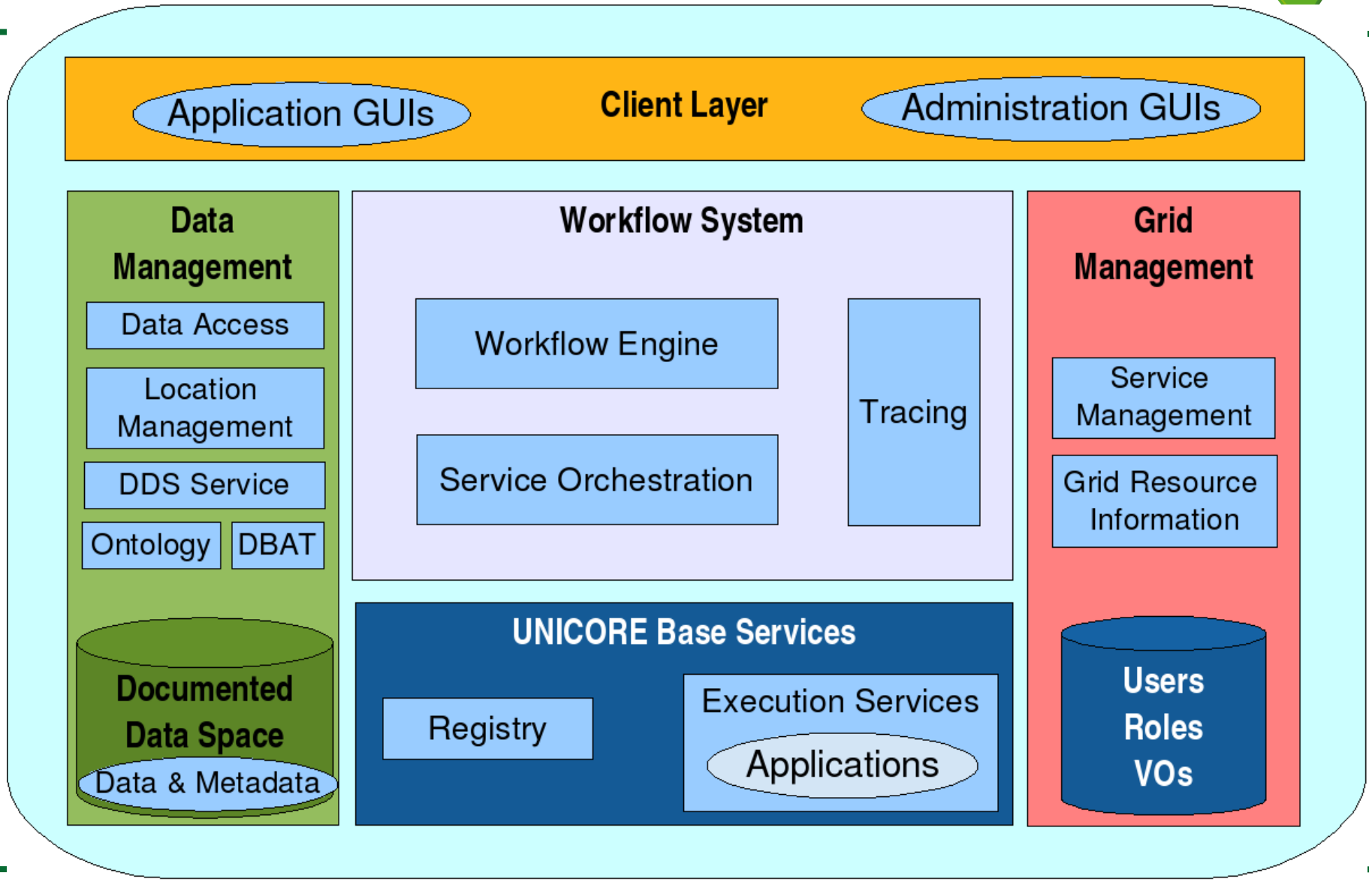
Chemomentum offers solutions to these challenges!

E.Benfenati, M. Casalegno (Inst. Mario Negri)

Architecture: the big picture



Architecture: detailed view



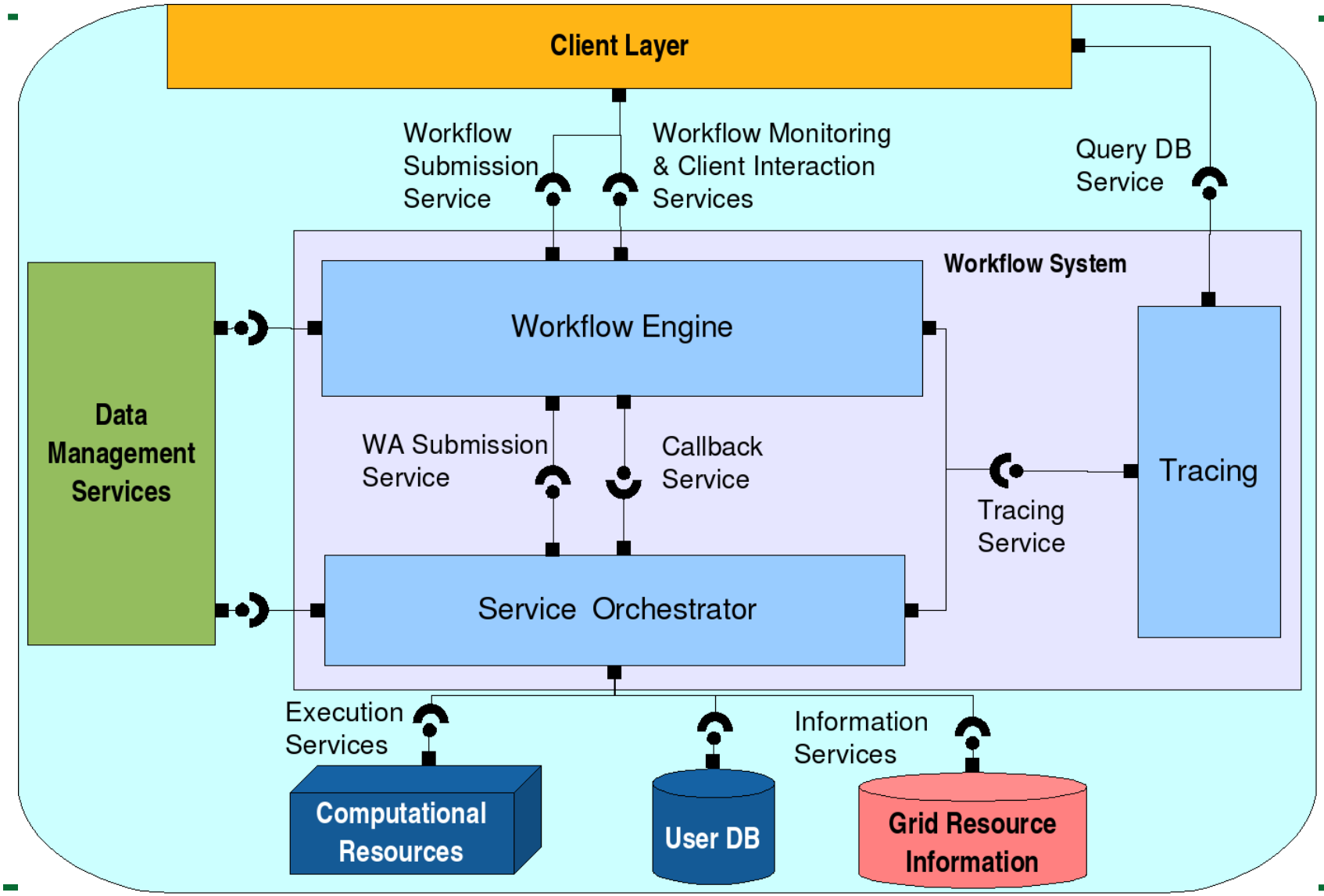
Some highlights ...

- Workflows
  - Workflow system architecture
  - Service orchestrator and brokering
- VO management service
- Chemomentum client layer

Workflow system: architectural goals

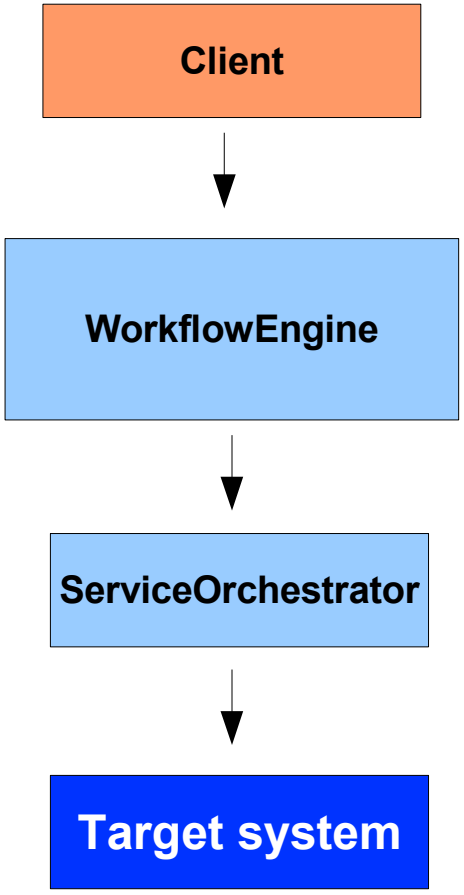
- Scalability and performance
  - Many target systems, applications parallelisable
- Support for domain-specific workflow languages
- Clear architecture
  - Separate process description and enactment
  - Keep workflow engine free of infrastructural concerns (e.g. guaranteed job execution)

Workflow system: details



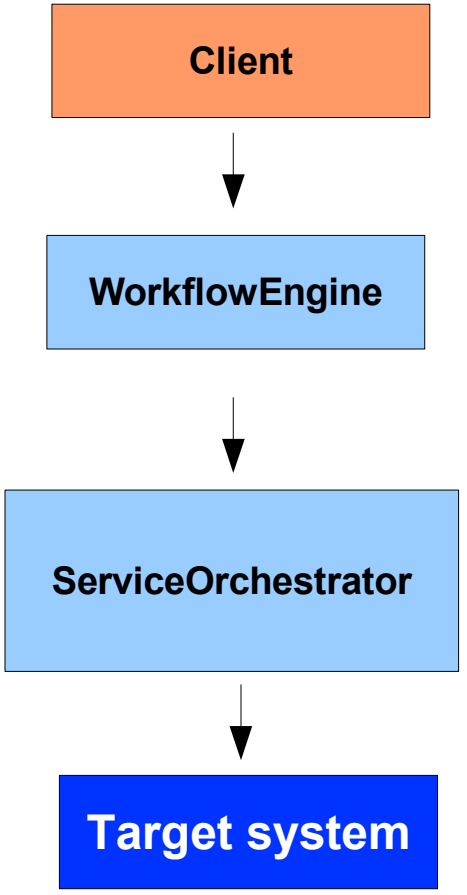
Workflow engine

- Supports domain-specific workflows
  - Used by clients
  - Pluggable
- Internal process description
  - XPDL 1.0 industry standard
- Constructs „atomic“ units of work
  - „Work assignments“
  - Not bound to target systems
  - Use „logical names“ for files

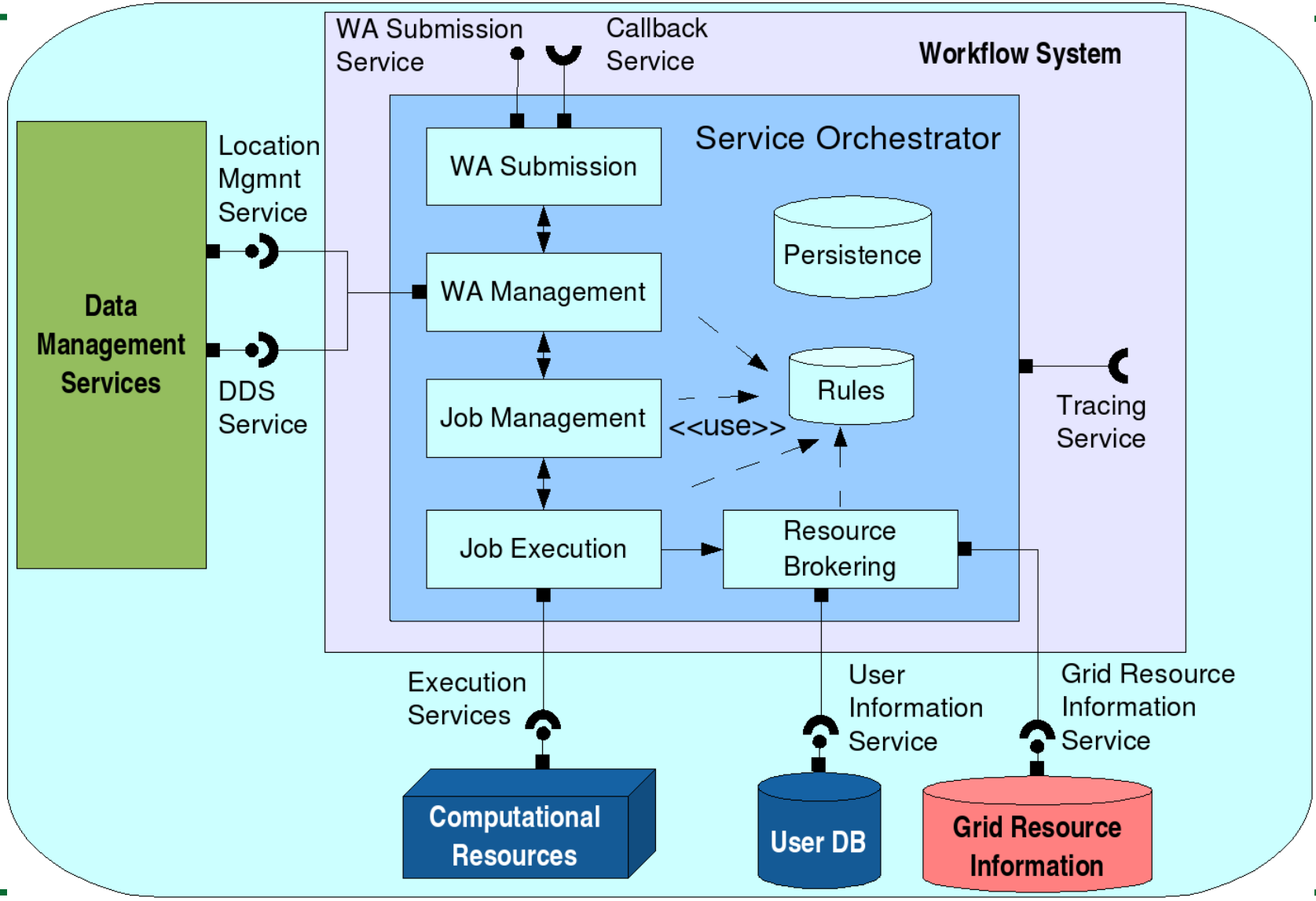


Service Orchestrator

- Execute atomic units of work (work assignments)
- Select and assign resources
- Resolve logical filenames to physical locations
- Store results to permanent storage



Service Orchestrator



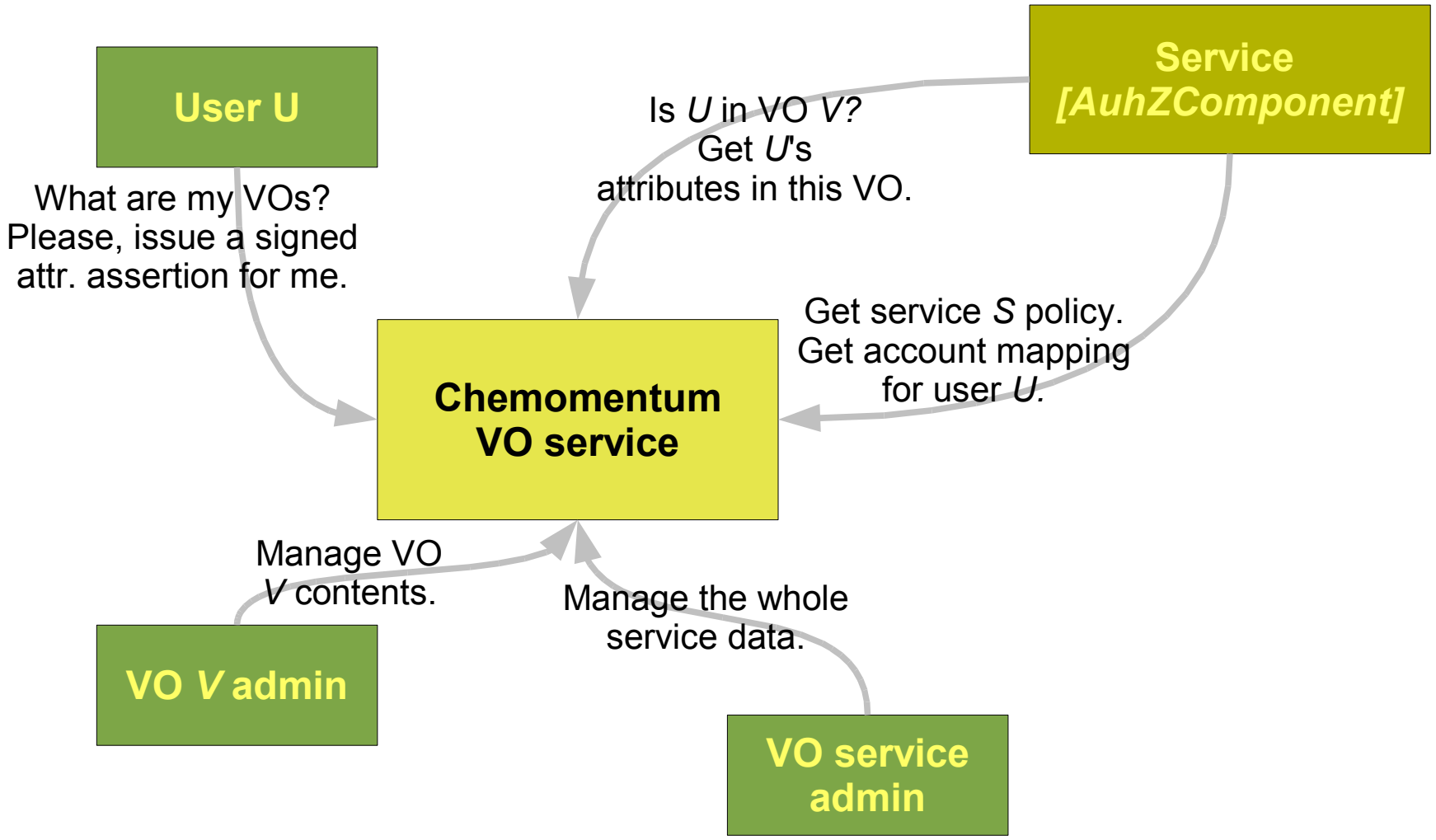
Service Orchestrator: Brokering

- Brokering:
  - Use live, dynamic information (where available)
    - System load
    - Job turnaround times (collected by the SO)
  - Static information
    - Installed applications, user rights, ...
- Grid Information Service
  - stores data relevant for brokering
- Brokering strategies are configurable (round robin, least turnaround time, ...)

VO service: goals

- Serve as Attribute and Policy Information Point, to support authorization process,
- Support authentication and new style (UNICORE 6) trust delegation by acting as identity format mapper,
- Provide flexible VO service access policy, to allow for distributed, multidomain management and usage of VO service's contents.

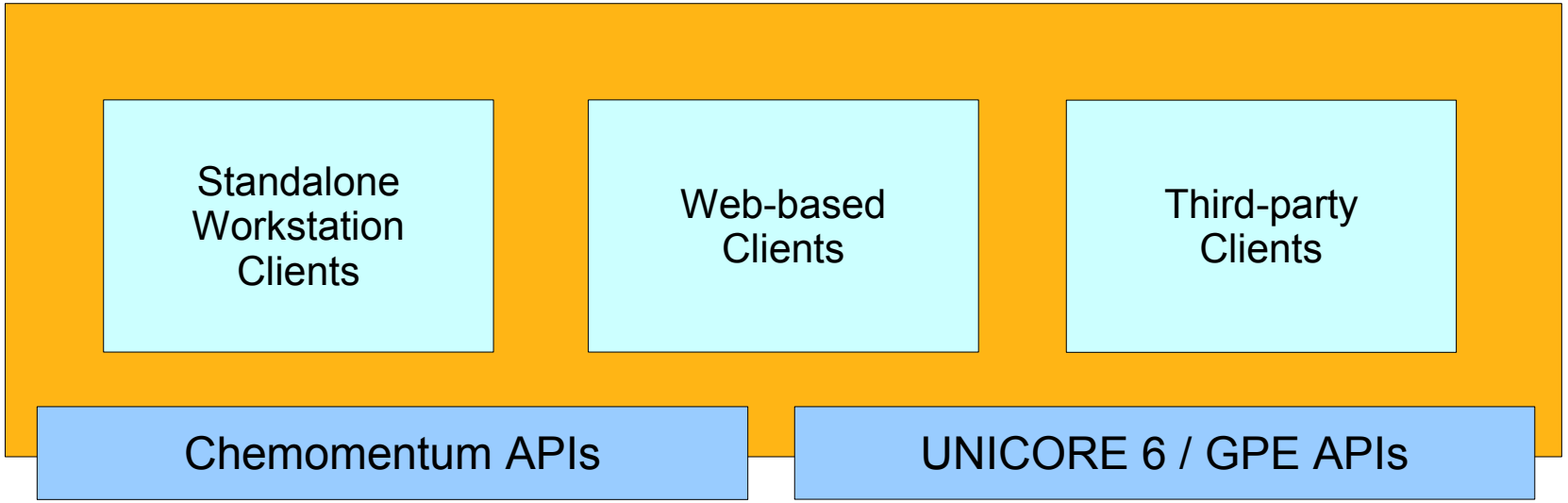
Use cases


User U
Service
[AuhZComponent]
What are my VOs?
Please, issue a signed
attr. assertion for me.
Is U in VO V?
Get U's
attributes in this VO.
Get service S policy.
Get account mapping
for user U.
Chemomentum
VO service
Manage VO
V contents.
Manage the whole
service data.
VO V admin
VO service
admin

VO service highlights

- SAML v2 protocol is used as interface for all query operations relevant from VO service contents consumer point of view
- Different formats of identities fully supported (i.e. with parsing/validation): X509 certificates and chains, DNs, email&password
- Pure Java, possibility to deploy without *any* extra software (using embedded DB)
- Support for querying the *full* VO service history
- Can be used as a drop-in replacement for the UNICORE 6 XUUDB

Client layer

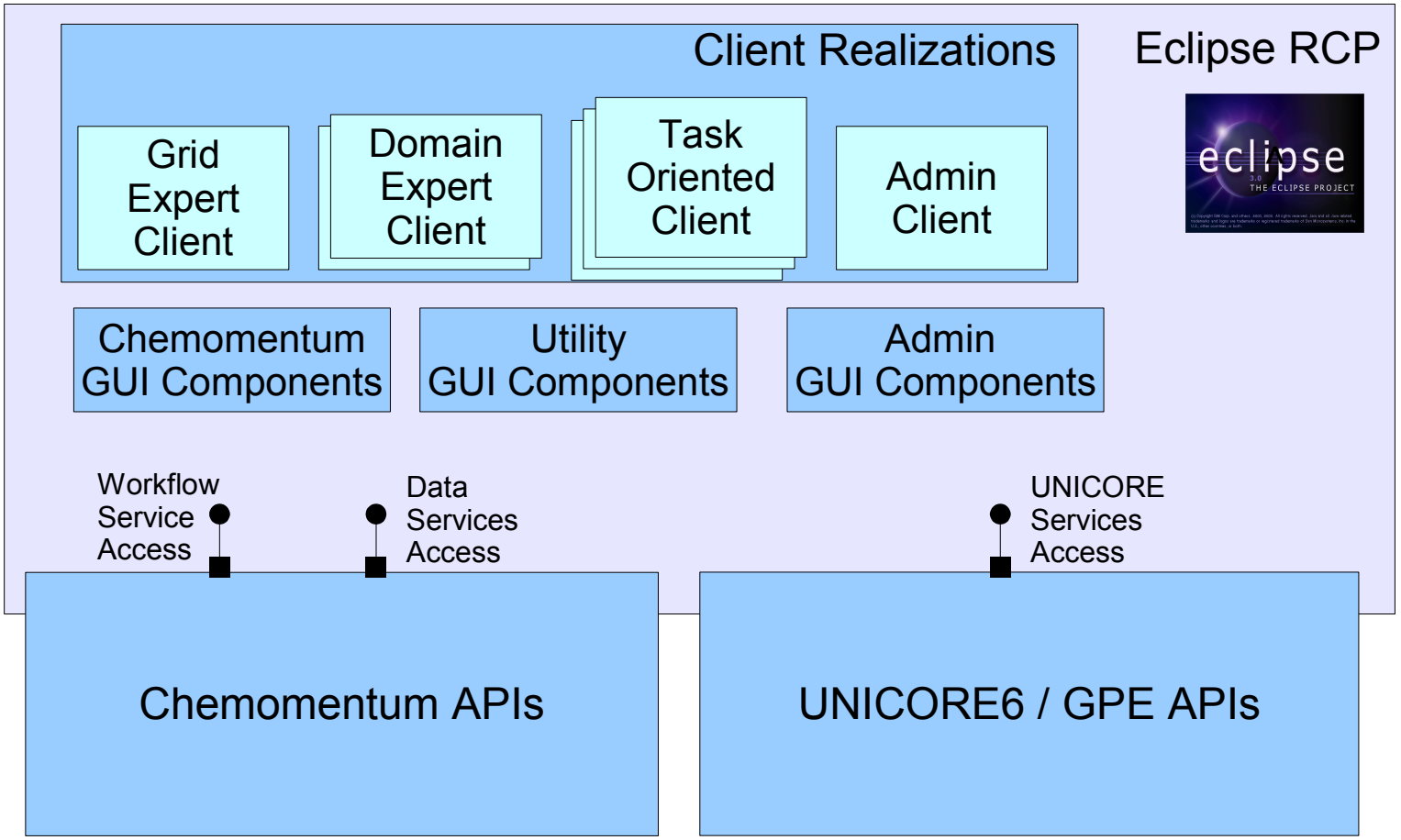


- Two main tracks
  - Standalone workstation clients
  - Web-based clients
- Side result: Grid-enabling third-party applications

Client layer: Goals

- Support standard UNICORE 6 application interfaces (GridBeans)
- Support different client configurations for different user groups (expert users, novice users)
- Easy-to-use, state of the art software

Workstation clients: Approach

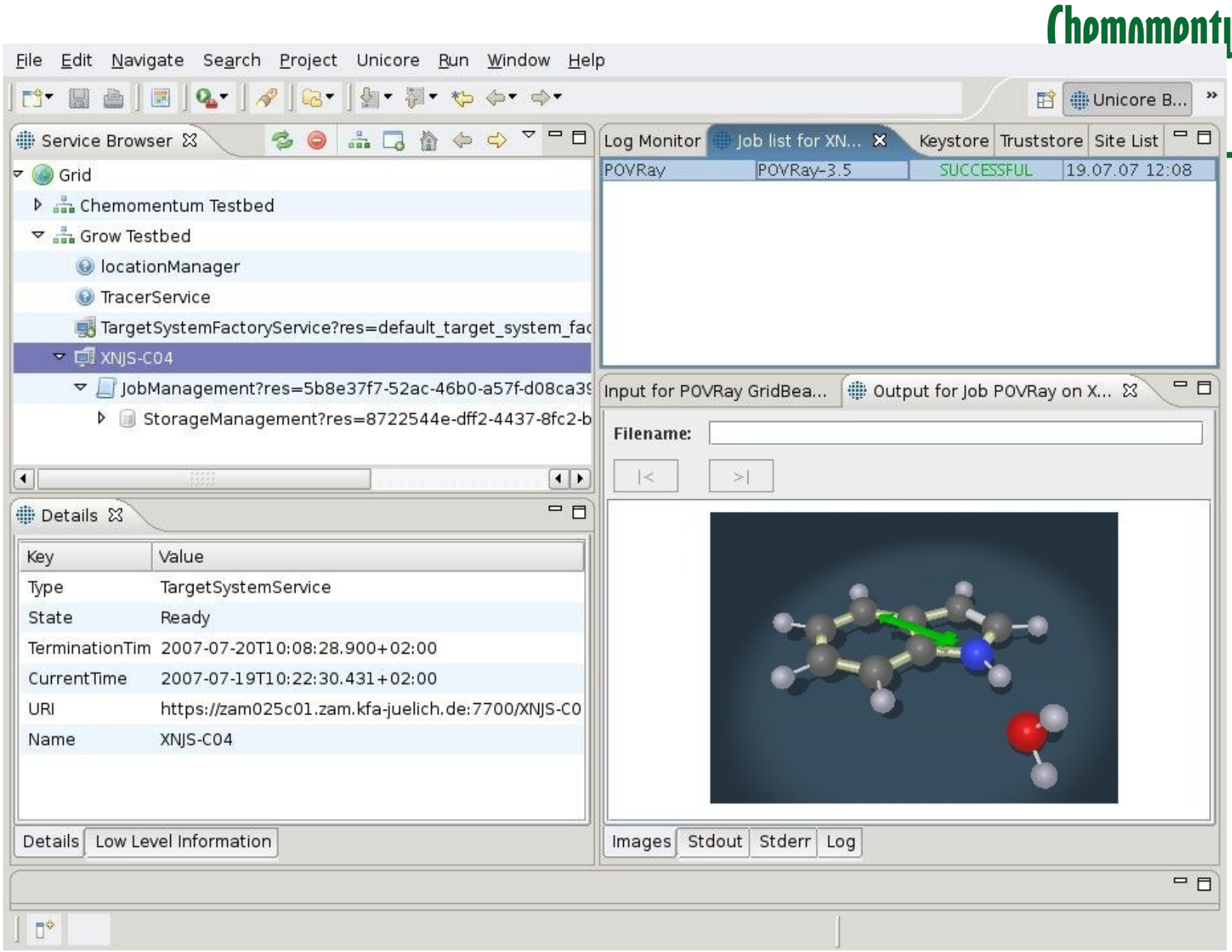


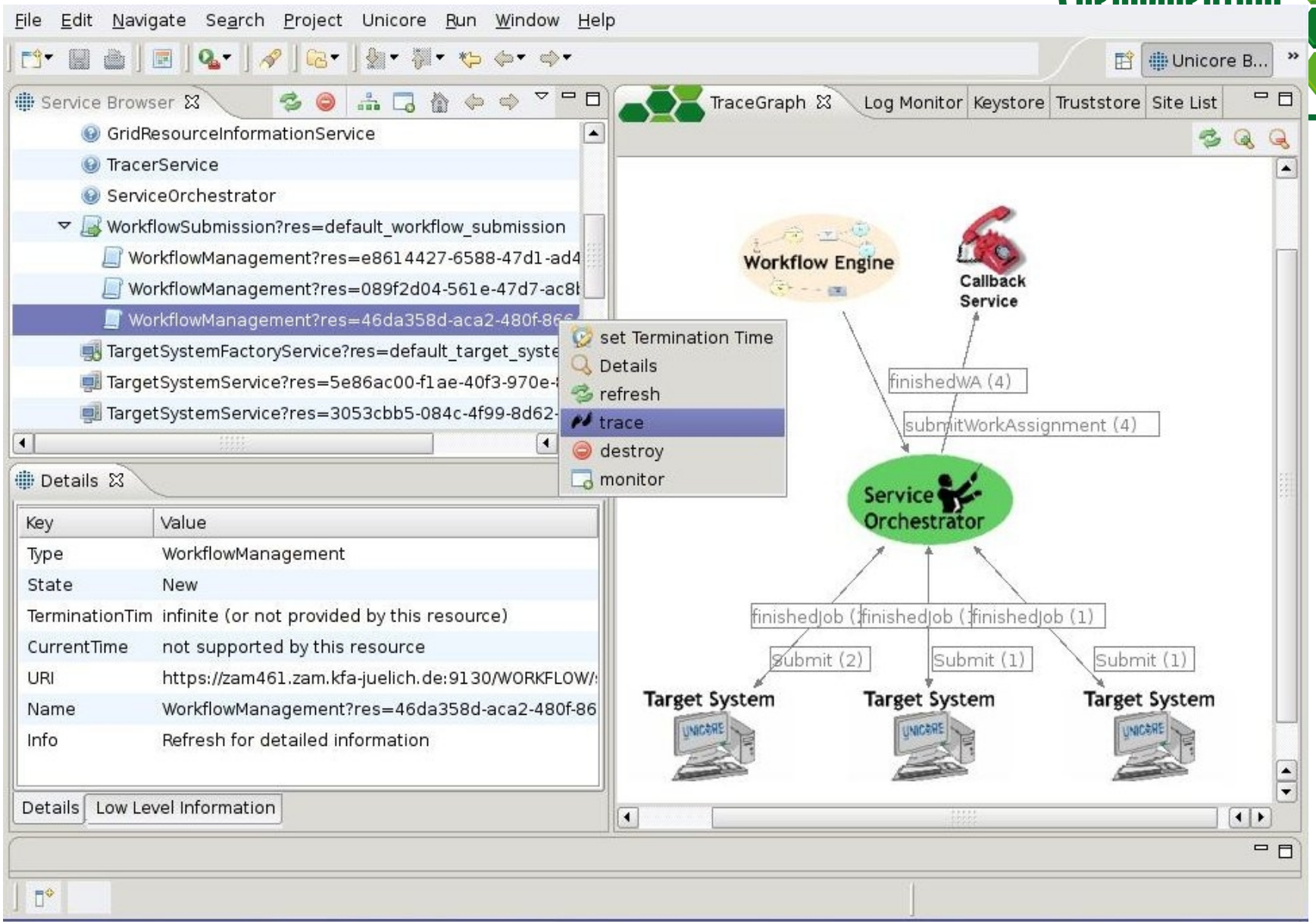
Software: Client framework

- Started close collaboration with Intel on evolving GPE / GridBeans towards Eclipse-based framework
  - UNCORE 6.1

- Many GUI components are already available

- Work in progress
  - Workflow editor based on Eclipse GEF


File Edit Navigate Search Project Unicore Run Window Help
Unicore B...
Service Browser
Grid
Chemomentum Testbed
Grow Testbed
locationManager
TracerService
TargetSystemFactoryService?res=default_target_system_fac
XNJS-C04
JobManagement?res=5b8e37f7-52ac-46b0-a57f-d08ca3
StorageManagement?res=8722544e-dff2-4437-8fc2-b
Log Monitor
Job list for XN...
Keystore
Truststore
Site List
POVRay
POVRay-3.5
SUCCESSFUL
19.07.07 12:08
Input for POVRay GridBea...
Output for Job POVRay on X...
Filename:
|<
>|
Details
Key
Value
Type
TargetSystemService
State
Ready
TerminationTim
2007-07-20T10:08:28.900+02:00
CurrentTime
2007-07-19T10:22:30.431+02:00
URI
https://zam025c01.zam.kfa-juelich.de:7700/XNJS-C0
Name
XNJS-C04
Details
Low Level Information
Images
Stdout
Stderr
Log


File Edit Navigate Search Project Unicore Run Window Help
Unicore B...
Service Browser
GridResourceInformationService
TracerService
ServiceOrchestrator
WorkflowSubmission?res=default_workflow_submission
WorkflowManagement?res=e8614427-6588-47d1-ad4
WorkflowManagement?res=089f2d04-561e-47d7-ac8
WorkflowManagement?res=46da358d-aca2-480f-86
TargetSystemFactoryService?res=default_target_syste
TargetSystemService?res=5e86ac00-f1ae-40f3-970e-
TargetSystemService?res=3053cbb5-084c-4f99-8d62-
set Termination Time
Details
refresh
trace
destroy
monitor
Details
Key
Value
Type
WorkflowManagement
State
New
TerminationTim
infinite (or not provided by this resource)
CurrentTime
not supported by this resource
URI
https://zam461.zam.kfa-juelich.de:9130/WORKFLOW/
Name
WorkflowManagement?res=46da358d-aca2-480f-86
Info
Refresh for detailed information
Details
Low Level Information
TraceGraph
Log Monitor
Keystore
Truststore
Site List
Workflow Engine
Callback Service
finishedWA (4)
submitWorkAssignment (4)
Service Orchestrator
finishedJob (1)
finishedJob (1)
finishedJob (1)
Submit (2)
Submit (1)
Submit (1)
Target System
Target System
Target System

Summary and outlook

- Chemomentum
  - UNICORE 6 based system for complex applications in science and technology
  - Workflow system (DSLs, 2-tiered architecture, flexible brokering system)
  - VO management
  - Eclipse based client framework
- First software release is due Sept. 7

Chemomentum

Thank you!