Common Instrument Middleware Architecture (CIMA)

Donald F. (Rick) McMullen
Knowledge Acquisition and Projection Laboratory,
Pervasive Technology Labs at Indiana University
mcmullen@indiana.edu

Net-Ready Sensors Workshop
Oak Ridge National Laboratory
August 3, 2006

pervasivetechnologylabs
AT INDIANA UNIVERSITY

V4.0 1-August-2006

NMI

CIMA Project Goals

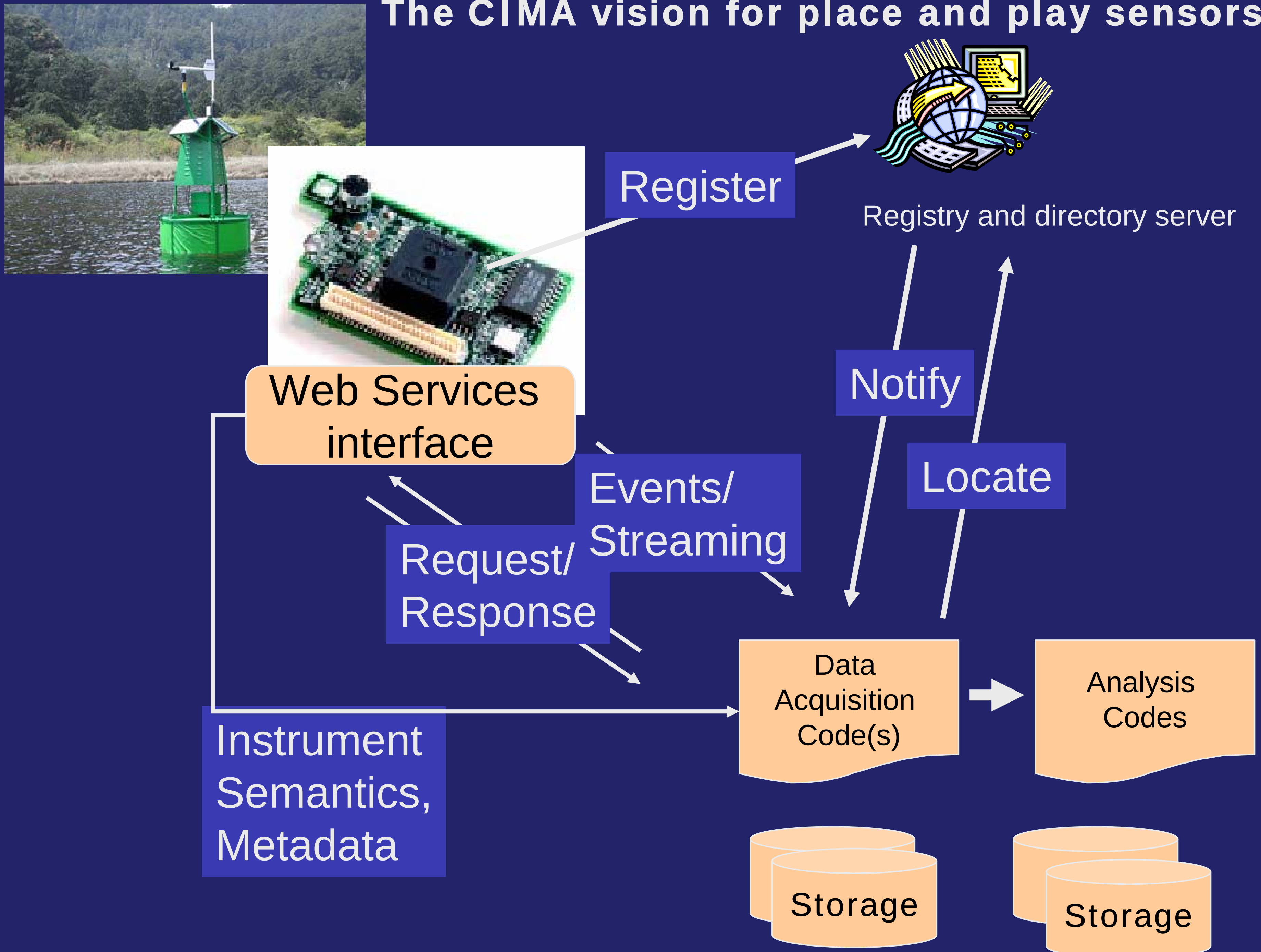
Project supported by the NSF Middleware Initiative to:

- Integrate instruments and sensors as real-time data sources into grid computing environments through a Service Oriented Architecture
  - Improve accessibility and throughput in instrumentation investments
  - Promote sharing across institutions and disciplines
- Develop a methodology for describing instrument capabilities and functions and embedding these in the hardware to improve flexibility and lifetime of data acquisition and analysis applications
  - Towards a Semantic Web for instruments and sensors?
- Move production of metadata as close to instruments as possible and facilitate the automatic production of metadata
  - Improve data management, provenance and reuse

CIMA Components

- Service architecture
  - Instrument representative (IR) code (device independent)
  - Drivers (Plug-ins, device dependent)
  - Service life cycle, high level protocol, communications, self description, discovery, security
  - Proxy service and embeddable code
  - Client coding practices
- Communications protocol
  - Transport neutral
  - Standards based
  - Support for synchronous and asynchronous interactions between client and IR
- Instrument and sensor description
  - Ontology-based with static and dynamic information
  - Clients should be able to build functional model from description
  - Description instance development parallels plug-in development
  - Ontology should be extensible as a community effort

The CIMA vision for place and play sensors

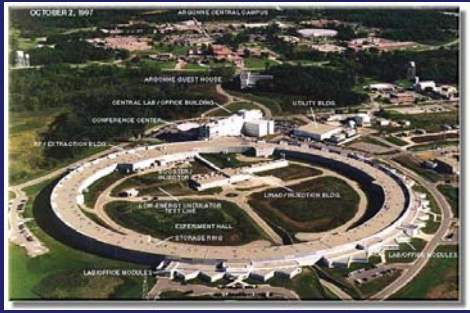


Design Considerations

- Prefer IP-accessible devices (IPv4 or v6) but support bridging to other buses and interconnects such as ZigBee, Motes RF protocol, RF modems, CANBus, tin cans and string
- Use “well known” transport protocols
  - SOAP over HTTP/S (doc/literal encoding)
- But support alternatives and extensions
  - Additional WS-X layered specs
  - Alternative message transports: JMS, NaradaBrokering, Antelope, IBM MQ, Jabber, etc.
  - Unreliable or intermittent connectivity (DTN2, message persistence in above messaging services)
- Prefer “proxy” approach to CIMA service implementation for richer capabilities set and less impact on existing instruments
  - Service is implemented on a separate PC or controller that is directly or indirectly interfaced to the instrument
- But support embedded systems through C++/gSoap version
- Location and cooperative use of instruments at the semantic level
- Integration with emerging Grid Computing standards for SOA (primarily WS-RF)

CIMA Reference Implementation Applications

- Synchrotron X-Ray crystallography
  - Argonne APS ChemMatCARS & DND-CAT
  - Also lab systems through CrystalGrid (global network of crystallography labs)

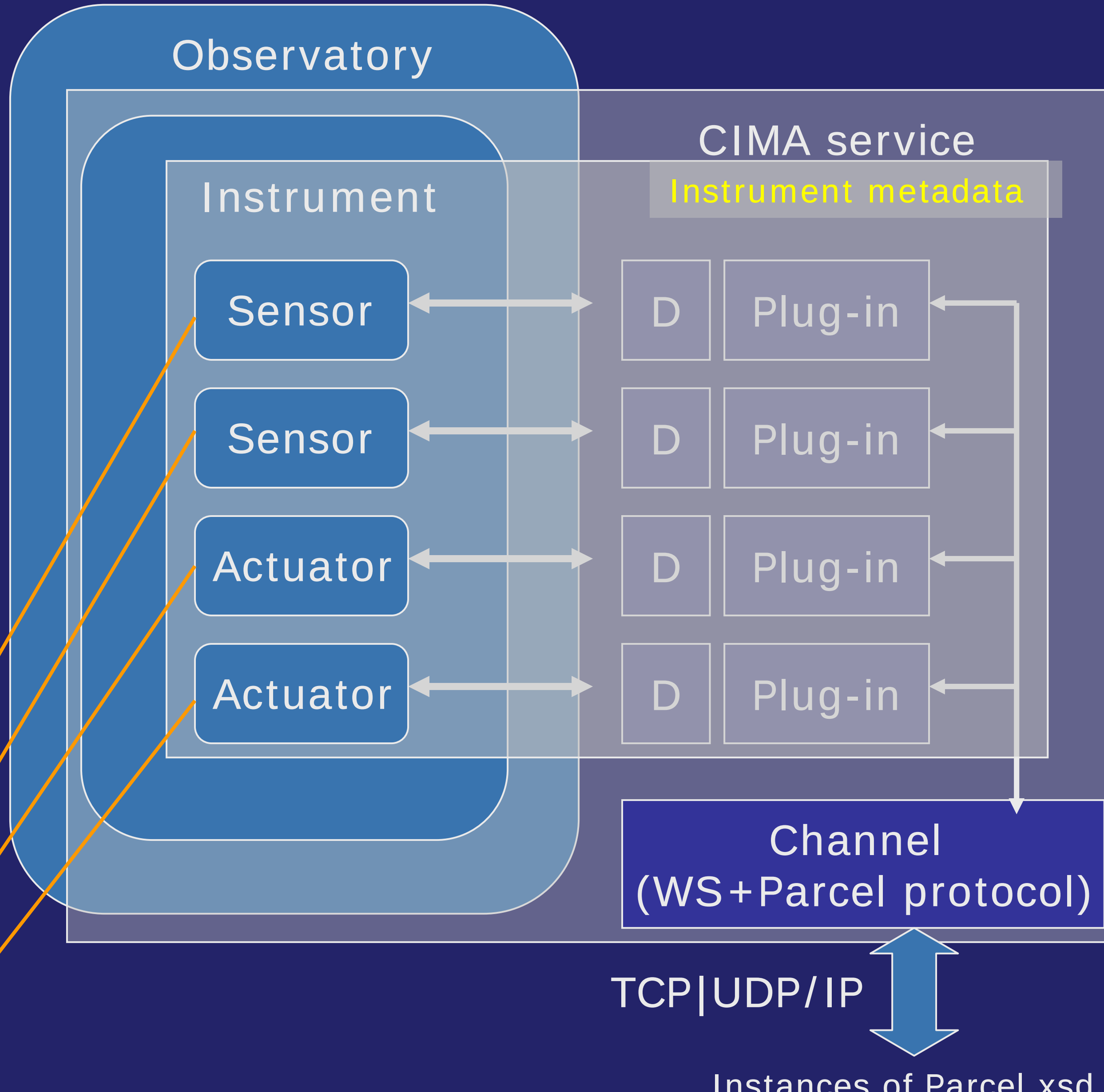


- TOF-MS
  - Identification of proteins and other macromolecules

- Robotic telescopes
  - Star variability
  - Looking for killer asteroids?
- Sensor networks
  - Ecological observation (LTER lake buoys)
  - Crossbow MOTE sensor package

Top-level CIMA Instrument Model

Components
- Optics and imaging
- Interior and exterior conditions
- Positioning
- Dome control

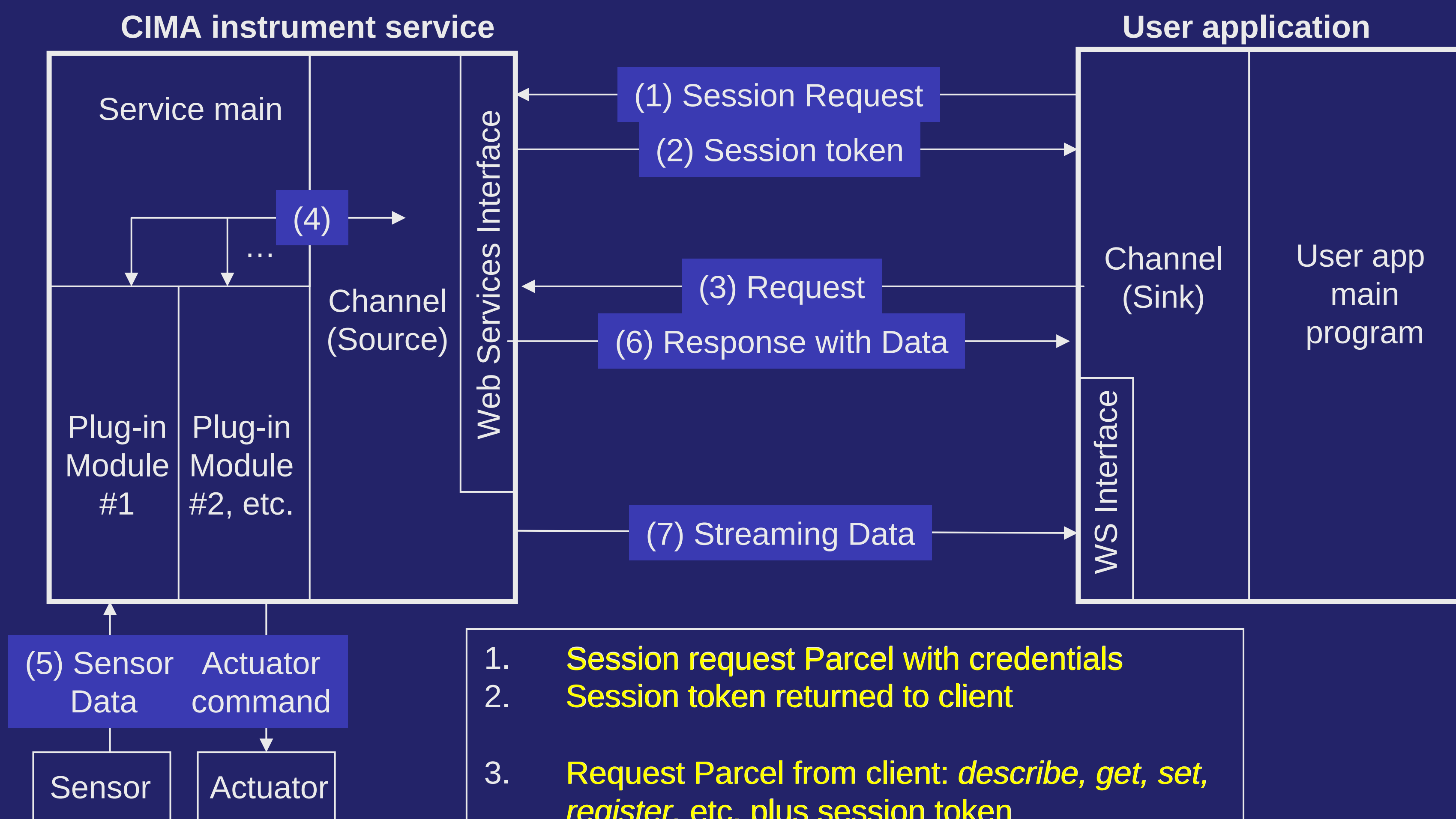


Instrument description and metadata

- Physical and logical attributes of the hardware and data produced by it
- Maps sensors and actuators defined in Plug-ins to the semantics of control functions and data outputs of the instrument
- Applications can query the instrument description to build an operational model of the instrument on the fly (W3C RDQL, Seaborne)
- Description instance based on ontology which can be implemented as OWL-DL (possibly plain XML as well)

Instrument description (con't)

- Ontology in OWL-DL
  - Description logic form allows for consistency checking at the semantic level, future application in machine reasoning (at least predicate logic)
- Provides classes and properties for
  - Identifying the instrument or sensor
  - Physical location (*e.g.* geospatial coords.)
  - Service location in communications space (URI)
  - Instrument characteristics
    - Names of data produced (observables, phenomena, units, derived or base, derivation formulae as λ/functional programs)
    - Shapes and types (scalar, array, integer, floating point, etc. using XML data types)
    - Response model (accuracy, resolution, dynamic range, sampling frequency)
    - How to acquire the data (WSDL+Parcel <-> observable mapping)



1. Session request Parcel with credentials
2. Session token returned to client

3. Request Parcel from client: *describe, get, set, register,* etc. plus session token
4. Channel calls appropriate plug-in for request type and data source
5. Plug-in retrieves data or runs actuator
6. Response Parcel is returned to client (data or result code)

7. If Client *register*s for event or streaming data Service calls client periodically or when data is available (timer or event-driven from plug-in)

CIMA Channel Protocol

- Simple Web Services interface: One endpoint at instrument, zero or one at client, both are String type
- SOAP payload is a *Parcel* XML document or fragment; REST-like protocol using document oriented SOAP messages (*Parcels*): *describe, register, get, set*
- Additional layered specifications can be added as needed, *e.g.* WS-Security, WS-Addressing ...
- Protocol can be extended by versioning the parcel schema without modification to older clients or instrument services
- Clients can request data on multiple streams at different rates; One instrument service supports multiple clients
- Performance can be augmented through use of Binary XML or gzip compression; non-HTTP/non-IP transports also possible

Parcel data structure

- Web Services interfaces are very simple: receive and send *Parcel* XML docs
- Protocol is in the Parcel schema, not in the WSDL so Web Services is not a requirement
- Parcels are small XML documents that contain
  - Header information about the requested operation and reply destination, and
  - Body containing request parameters, reply detail (actuator status or sensor data information, metadata)
- Producer-> intermediary-> consumer model needs “routable” structures that intermediaries can forward without completely parsing
- Suitable for resource constrained wireless sensor networks and source routing protocols

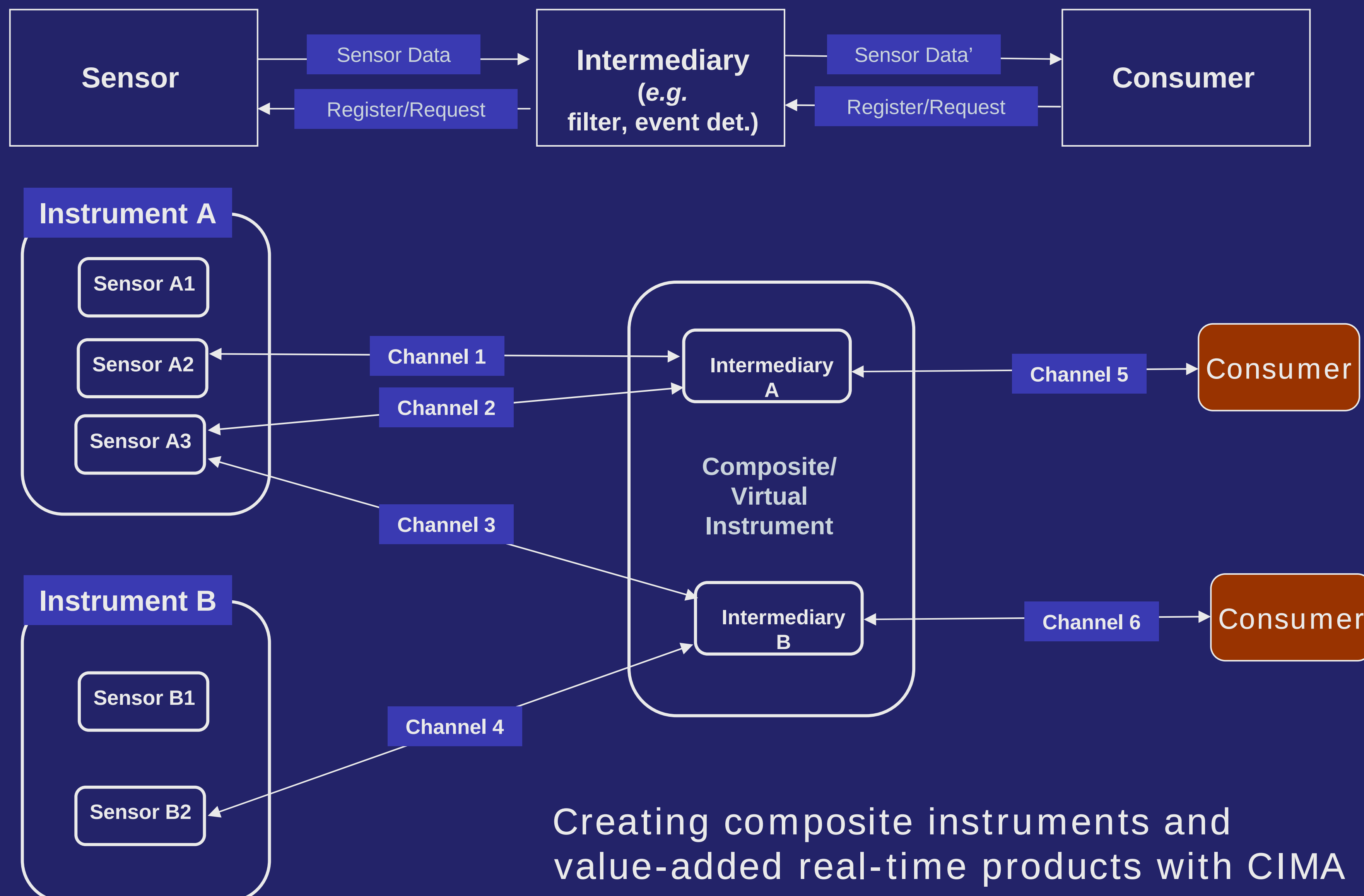
Security (SOAP/HTTP)

- Line level security through SSL (HTTPS)
- Additional security available through WS-Security or other SOAP message encryption schemes
- Works through firewalls that allow HTTP or HTTPS
- Access control
  - Authentication by user-provided credentials and pluggable authN modules to set up sessions
  - Output filtering by IP address or range
  - External role or capabilities-based authorization using SAML and Shibboleth or PERMIS

Transport

- Web Services does not define a transport layer for SOAP messages
- WSDL defines a transport binding per port
- Common bindings: http/s, JMS, Narada Brokering, TCP sockets
- Intermediaries can be used as:
  - message converters such as CIMA Channel SOAP to RSS or CAP
  - transport converters, *e.g.* https->NaradaBrokering

CIMA in processing pipelines



Creating composite instruments and value-added real-time products with CIMA

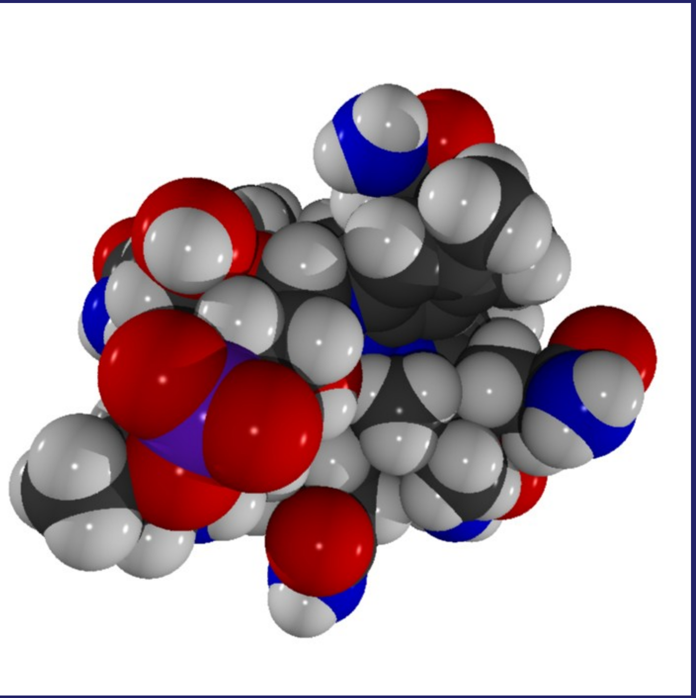
Data flow and service orchestration

- CIMA currently provides its own support for data flow with manual composition
- Consumer applications can be components in a data flow graph
- Data flow + resource allocation = workflow
  - Composition schemes (e.g. BPEL4WS) can be used to build workflows and bind these to resources
  - Some useful workflow testbed tools: Kepler, Triana (WS-capable), and Taverna (fully WS-based)
  - Persistent workflows captured as BPEL
  - Portlets (portal components) as user interfaces to workflow graphs

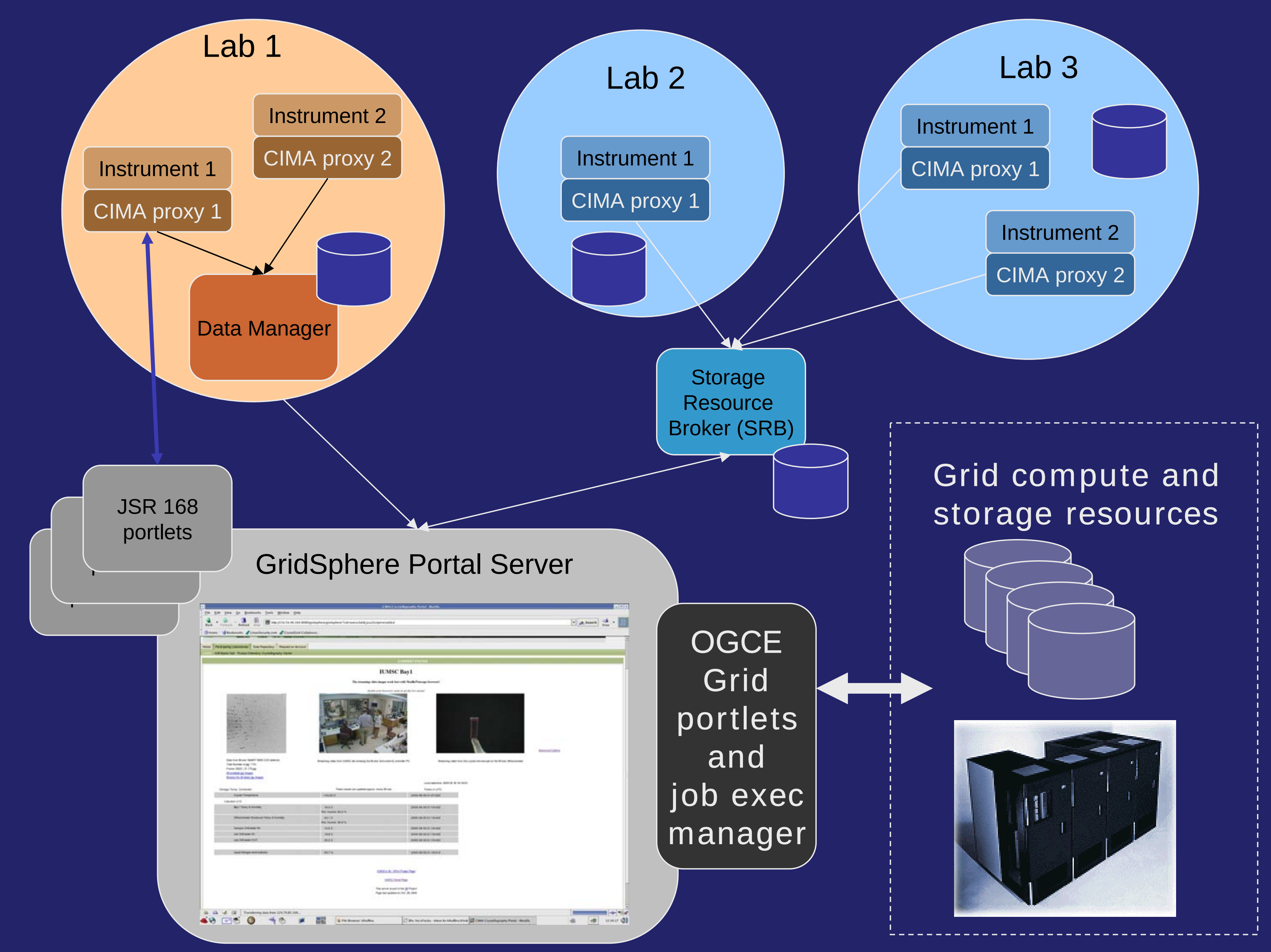
Applications

X-ray Diffraction Crystallography

- X-ray diffraction crystallography is an important technique for determining the 3-D structure of both large and small molecules
- Major instrumentation facilities at synchrotron X-ray sources in the US, Canada, EU, Australia and Asia
- Remote access improves throughput and data quality
- Sharing of individual instruments at smaller labs with differing capabilities



- CIMA used to develop federation of labs for remote access and facility sharing
- Participating facilities
  - Active: US - APS ChemMatCARS, IU, IU biology, Purdue, Uminn, AU - USydney, James Cook University
  - Joining: US - CWRU, UW-Madison, AU - UQueensland, UK - Soton


Lab 1
Instrument 2
CIMA proxy 2
Instrument 1
CIMA proxy 1
Data Manager
Lab 2
Instrument 1
CIMA proxy 1
Lab 3
Instrument 1
CIMA proxy 1
Instrument 2
CIMA proxy 2
Storage Resource Broker (SRB)
JSR 168 portlets
GridSphere Portal Server
OGCE Grid portlets and job exec manager
Grid compute and storage resources

CIMA Crystallography Portal - Mozilla

File Edit View Go Bookmarks Tools Window Help

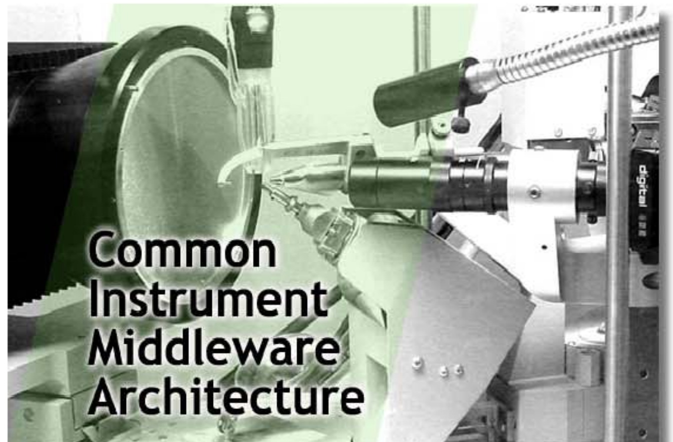
Back Forward Reload Stop Search Print

Home Bookmarks LinuxSecurity.com CrystalGrid Collaborat... WJ-NT104 MAIN PAGE Apache Tomcat/5.0.2... CIMA Crystallography... CIMA Mass Spectrom...

Home | IUMSC | ChemMatCARS | Purdue | Myers Hall | Request an Account

Home News About Contact

- IUMSC
- ChemMat-CARS Univ. of Chicago at APS
- Purdue Crystallography Center
- IUB Myers Hall
- Other Collaborating Laboratories
- CSAF University of Sydney, AU
- NCS Southampton, UK
- X-ray Cryst. Lab Univ. of Minnesota
- Other CIMA Application Portals



This portal demonstrates some of the functionality that will be available to the crystallographers collecting data in participating laboratories (left hand side menu). Researcher can follow the progress of an experiment using a simple web browser connected to the Internet. The IUMSC instrument is being used as a test system for the Instrument Middleware project, see *http://www.instrumentmiddleware.org*

The work being described on these pages are supported by NSF Cooperative Agreement OCI-0330568 and MRI CDA-0116050

Login

User Name

Password

Remember my login

Login

Forget your password?

powered by gridsphere

Done

[File Brows | Re: update | CIMA Cryst | The GIMP | Layers | Tool Optior | SC05_thea | Collaborati | Feb14_IUN | CIMA_crys

18:11:28

CIMA Crystallography Portal - Mozilla

File Edit View Go Bookmarks Tools Window Help

Back Forward Reload Stop Search Print

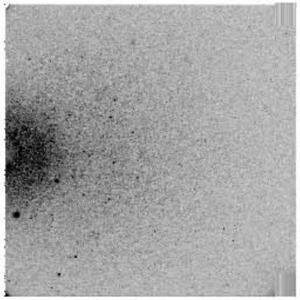
Home Bookmarks LinuxSecurity.com CrystalGrid Collaborat... WJ-NT104 MAIN PAGE Apache Tomcat/5.0.2... CIMA Crystallography... CIMA Mass Spectrom...

Home | IUMSC | ChemMatCARS | Purdue | Myers Hall | Request an Account

IUMSC Current Status   Current Sample   Previous Samples   About This Lab

IUMSC Bay1

Disable your browser's cache to get the live stream!



Advanced Options

Data from Bruker SMART 6000 CCD detector
Total Number of jpg: 7251
Frame: 0531911.498.jpg
All available jpg images
Browse the 20 latest jpg images

Streaming video from IUMSC lab showing the Bruker instrument & controller PC

Streaming video from the crystal microscope on the Bruker diffractometer

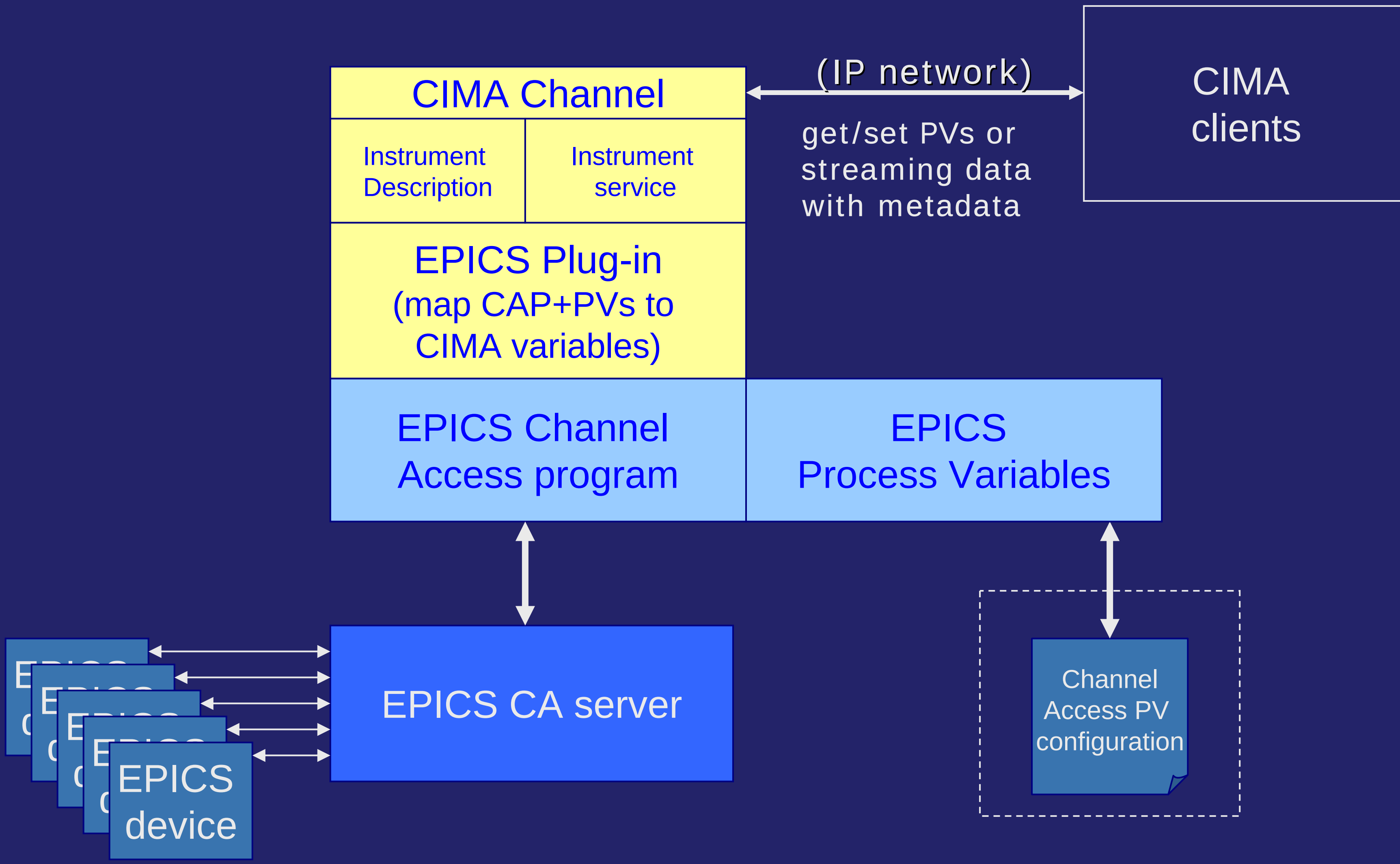
Local date/time: 2005-11-09 18:15:19

Omega Temp. Controller	These values are updated approx. every 60 sec.	Times in UTC
Crystal Temperature	-146.20 C	2005-11-09 23:14:36Z
LabJack U12		
Bay1 Temp. & Humidity:	19.7 C Rel. Humid. 36.9 %	2005-11-09 23:14:50Z
Diffractometer Enclosure Temp. & Humidity:	25.8 C Rel. Humid. 23.5 %	2005-11-09 23:14:51Z
Campus Chill water IN:	10.5 C	2005-11-09 23:15:13Z
Lab Chill water IN:	19.3 C	2005-11-09 23:15:13Z
Lab Chill water OUT:	25.7 C	2005-11-09 23:15:13Z
Liquid Nitrogen level indicator	35.4 %	2005-11-09 23:15:13Z

Transferring data from 129.79.85.100...

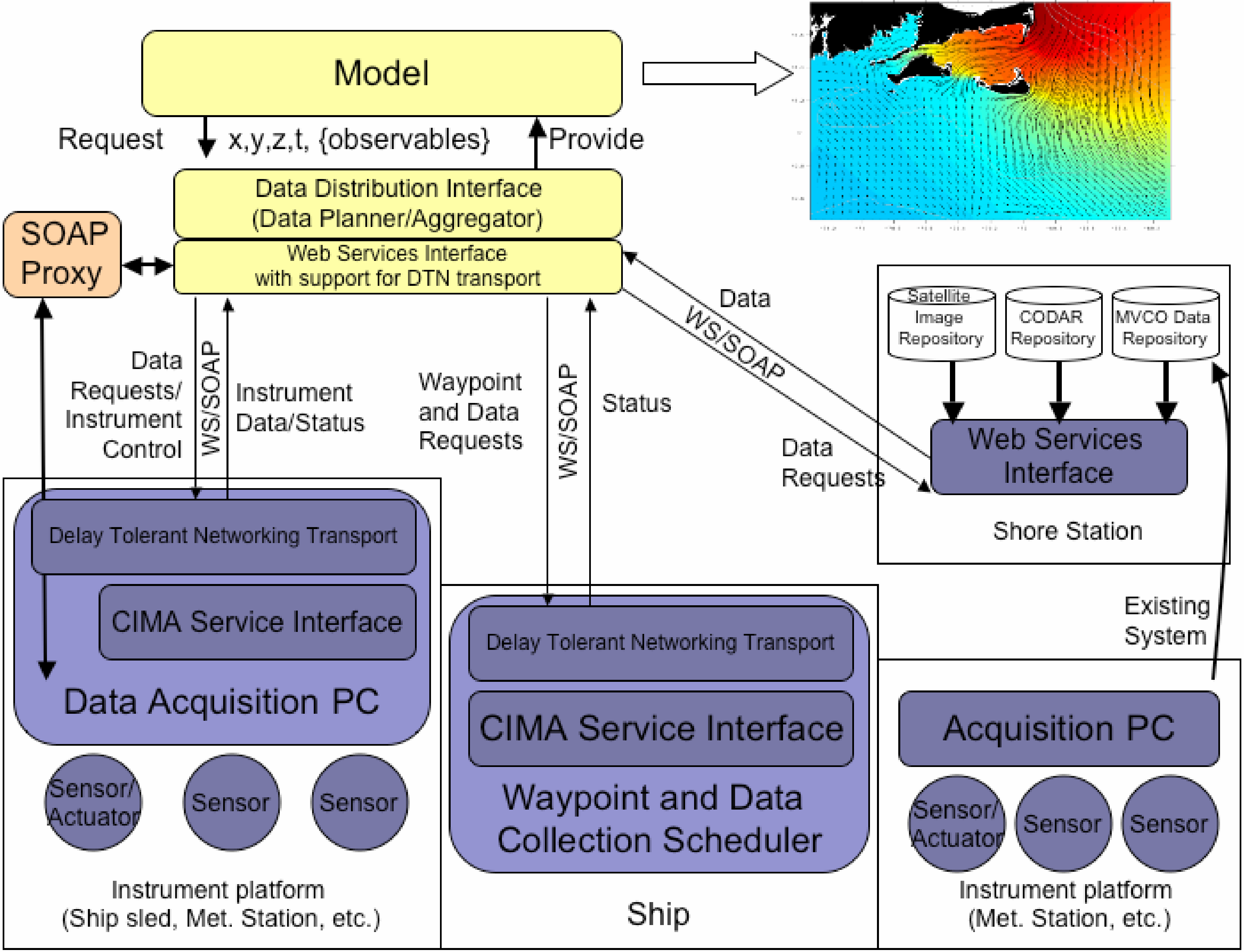
File Browser: sc05 | Re: update link - I | CIMA Crystallogra | The GIMP | Layers | Tool Options | SC05_theater_pre | 18:15:25

CIMA-EPICS Integration:
Extending accessibility of beamline control systems



Environmental sensing - LTER lake buoys

- Network of freshwater lake sensing buoys in US, EU and Asia
- Buoys collect temperature, current profiles, chemistry data
- Need to acquire data from a number of sites at different locations around the world with varied capabilities
- Need flexible, reconfigurable event detection and response in the field
- Sensor placement (expand or contract network), maintenance and calibration, data QA/QC
- Implemented as place and play sensor nodes and an agent network for event detection, network reconfiguration, QA/QC and sensor calibration



Model
Request
x,y,z,t, {observables}
Provide
Data Distribution Interface
(Data Planner/Aggregator)
Web Services Interface
with support for DTN transport
SOAP Proxy
Data Requests/ Instrument Control
WS/SOAP
Instrument Data/Status
Waypoint and Data Requests
WS/SOAP
Status
Data
WS/SOAP
Data Requests
Satellite Image Repository
CODAR Repository
MVCO Data Repository
Web Services Interface
Shore Station
Existing System
Delay Tolerant Networking Transport
CIMA Service Interface
Data Acquisition PC
Sensor/ Actuator
Sensor
Sensor
Instrument platform
(Ship sled, Met. Station, etc.)
Delay Tolerant Networking Transport
CIMA Service Interface
Waypoint and Data Collection Scheduler
Ship
Acquisition PC
Sensor/ Actuator
Sensor
Sensor
Instrument platform
(Met. Station, etc.)

Implementation Issues

- Location (UDDI, MDS4/WSRF, WS-Discovery; WFS?)
- Security (at each node, transport level)
- Authentication and authorization
  - SAML at low level using VO-level mechanisms for rights management (*e.g.* Shibboleth, PERMIS)
  - PKI/CA, Kerberos, etc. for authentication
  - Authorization by attribute or capabilities-based rights mgmt. system such as Shibboleth, PERMIS or Xpola
- Workflow
  - Capture acquisition-reduction-analysis-publishing-archiving processes as workflows with instruments at the "head-end"
  - Initial experiments underway with WS workflow systems, *e.g.* Taverna (GUI), ActiveBPEL (BPEL4WS) and WS-supporting sysems like Triana and Kepler.
  - Scheduling and authorization are major design issues

More Implementation Issues

- Aggregation and transformation of real-time data streams
  - Availability of instrument description and metadata in real time supports data fusion and real time signal processing
  - Enabled by Parcel header/content separation
  - Composite instruments (concatenation and nesting of individual CIMA services)
- Performance
  - Single readings or control commands (via SOAP)
  - Streaming
  - Remote control (bounded or known latency)

Related work

- Device capabilities
  - OpenGIS Consortium GML, sensorML - XML schema for sensor description
  - W3C Composite Capabilities/ Preferences Profile - ontology based RDF descriptions of mobile device capabilities
  - IEEE 1451.0 - layered specification for “universal” sensor access; .0 is Transducer Electronic Data Sheet (TEDS)
  - Monterey Bay Aquarium Research Institute - “plug and work” instrument puck
- Device access
  - IEEE 1451.1 - access to sensors using IP
  - Sensor Web Enablement community - Some interest in Web Services as protocol for sensor access.
  - NeesGrid Telecontrol Protocol
- Integration
  - ORNL SensorNet
  - SDSC/CALIT2 RoadNet

Thanks! Questions?

(mcmullen@indiana.edu)

Support for this work provided by the National Science Foundation is gratefully acknowledged. (SCI 0330568, DBI 0446802)

NSF Middleware Initiative: www.nsf-middleware.org

CIMA project: www.instrument-middleware.org