

# Activity Instance Schema

Activity Use Cases & Requirements

Thursday, 05 June, 2008

Barcelona, ES

# OGF IPR Policies Apply



- "I acknowledge that participation in this meeting is subject to the OGF Intellectual Property Policy."
- Intellectual Property Notices Note Well: All statements related to the activities of the OGF and addressed to the OGF are subject to all provisions of Appendix B of GFD-C.1, which grants to the OGF and its participants certain licenses and rights in such statements. Such statements include verbal statements in OGF meetings, as well as written and electronic communications made at any time or place, which are addressed to:
  - the OGF plenary session,
  - any OGF working group or portion thereof,
  - the OGF Board of Directors, the GFSG, or any member thereof on behalf of the OGF,
  - the ADCOM, or any member thereof on behalf of the ADCOM,
  - any OGF mailing list, including any group list, or any other list functioning under OGF auspices,
  - the OGF Editor or the document authoring and review process
- Statements made outside of a OGF meeting, mailing list or other function, that are clearly not intended to be input to an OGF activity, group or function, are not subject to these provisions.
- Excerpt from Appendix B of GFD-C.1: "Where the OGF knows of rights, or claimed rights, the OGF secretariat shall attempt to obtain from the claimant of such rights, a written assurance that upon approval by the GFSG of the relevant OGF document(s), any party will be able to obtain the right to implement, use and distribute the technology or works when implementing, using or distributing technology based upon the specific specification(s) under openly specified, reasonable, non-discriminatory terms. The working group or research group proposing the use of the technology with respect to which the proprietary rights are claimed may assist the OGF secretariat in this effort. The results of this procedure shall not affect advancement of document, except that the GFSG may defer approval where a delay may facilitate the obtaining of such assurances. The results will, however, be recorded by the OGF Secretariat, and made available. The GFSG may also direct that a summary of the results be included in any GFD published containing the specification."
- OGF Intellectual Property Policies are adapted from the IETF Intellectual Property Policies that support the Internet Standards Process.

# Content

---

- Objectives
- Status Quo of Activity Discussion
- Use Cases
- Initial Requirements
- Next Steps

# Objectives of this Session

---



- Understand the use cases collected so far
- Collect more use cases from the audience
- Discuss requirements derived from the use cases
- Plans until OGF 24

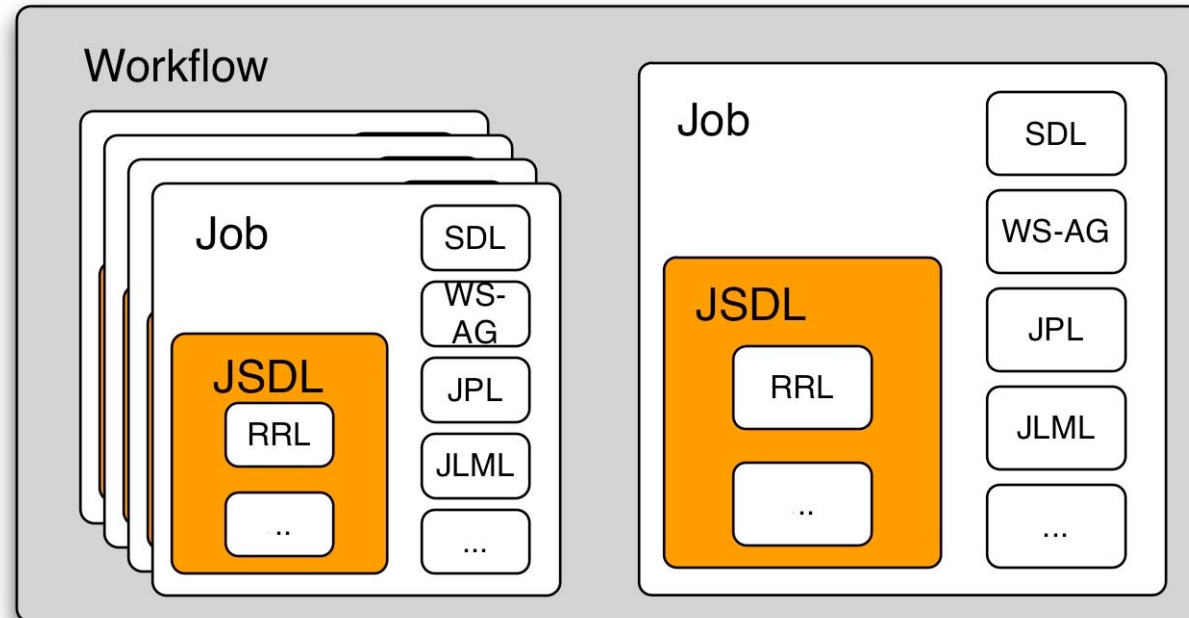
# Problem Domain - the Activity view

---



- Activity concept is broader than just compute (as e.g. currently seen by BES)
- Think about business scenarios using SLAs
  - Modelling SLA lifecycle
  - Creation and Monitoring of SLAs
  - Evaluation
- SLA = “Business Activity”
  - Comprises compute, but also ...
  - ... data, accounting info, provenance, lifecycle, etc.
- This information has to be captured and related to an activity

# Problem Domain - the JSDL view



JSDL comprises a core Resource Request Language and exists within an environment of other languages like Scheduling Description Language, Job Policy Language, etc. --> more requirements, dependencies, etc. wrt to an Activity

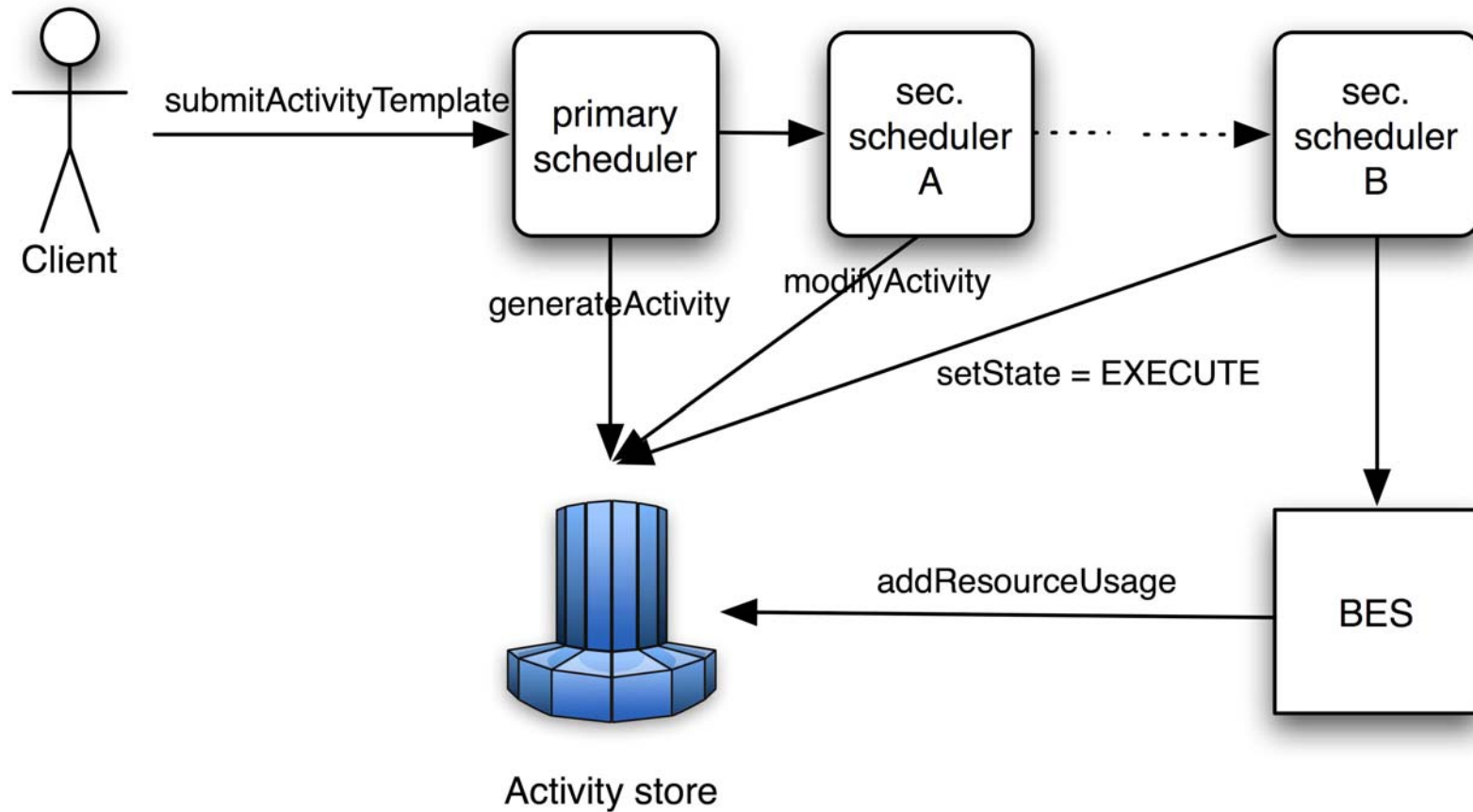
# Progress since OGF 22

---



- Initial NextGRID schema accepted as foundation of an Activity schema
- Use case collection and discussion
  - Four use cases for now
  - One still in the pipe
  - More from the community?
- Initial requirements list generated
  - To be discussed here
  - First step: decide what is in/out of scope

# Use Case I - Scheduling.



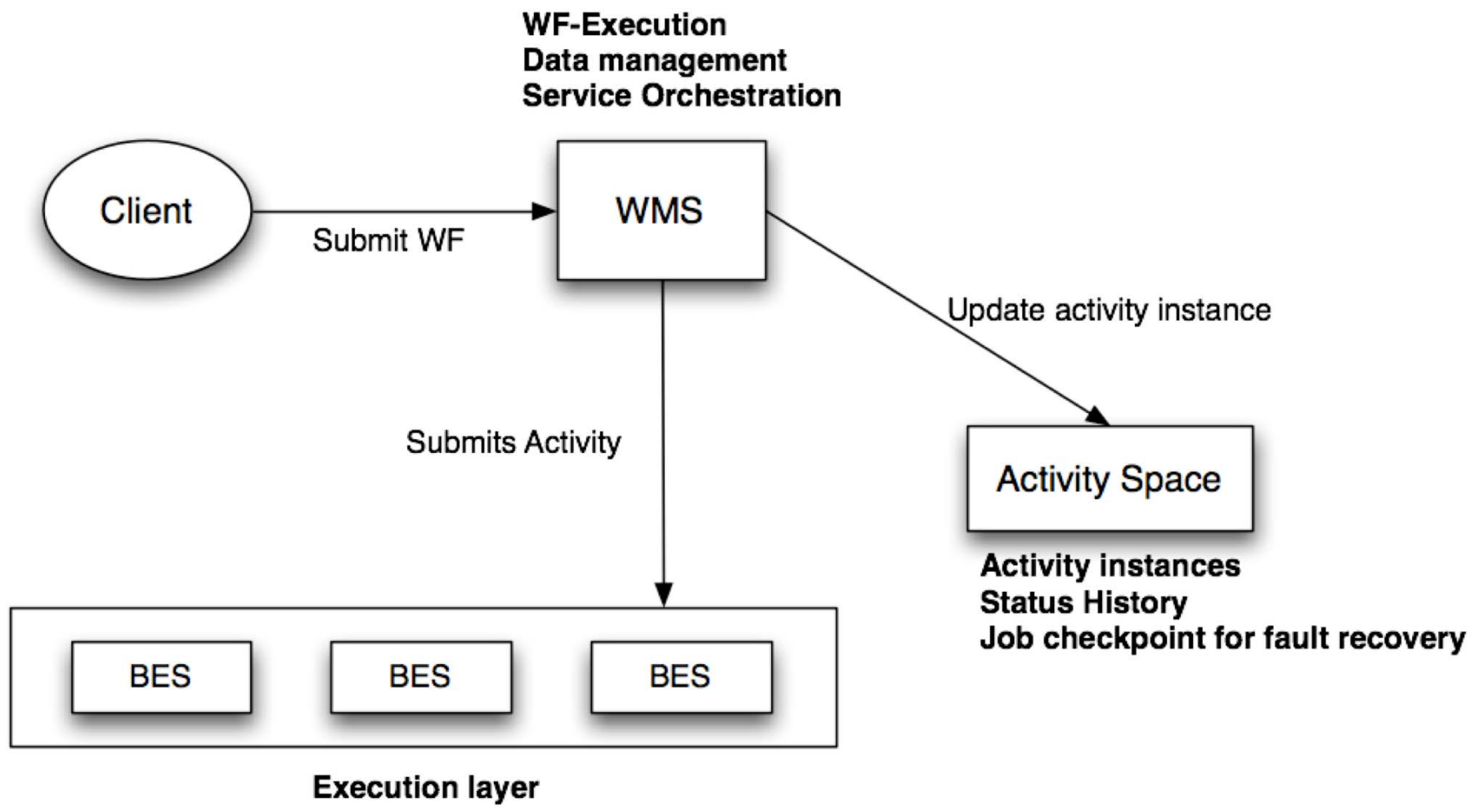
# Use Case II - Job lifecycle tracking

---



- Purpose
  - to track a job (and its attributes) throughout the entire life cycle of a job
- Information tracked
  - Point of submission (and the unmodified parameters).
  - Any state change that occur.
  - Information about the activity being forwarded from one scheduler to another.
  - A report of resources consumed by the job both.
  - Final “accounting” record.
- Functional requirements
  - Access to activity information during/after lifetime

# Use Case III - Workflow fault detection



# Use case IV - Job History

---



- Record historical job info for estimating future job executions
- Administrators need this for troubleshooting, decision making, capacity planning, etc.
- Customers can use this info to see how/if things work properly
- Info can also be used to simulate new job flows based in existing ones
- Combination of Activity and JSDL schemata to support interoperability

# Requirements

---



See document ...

# Next steps

---

- Collect & finalise more use cases & requirements
- Decide which requirements are in scope
- Start adapting the initial NextGRID activity schema
- AOB?

# Full Copyright Notice

---



Copyright (C) Open Grid Forum (2007, 2008). All Rights Reserved.

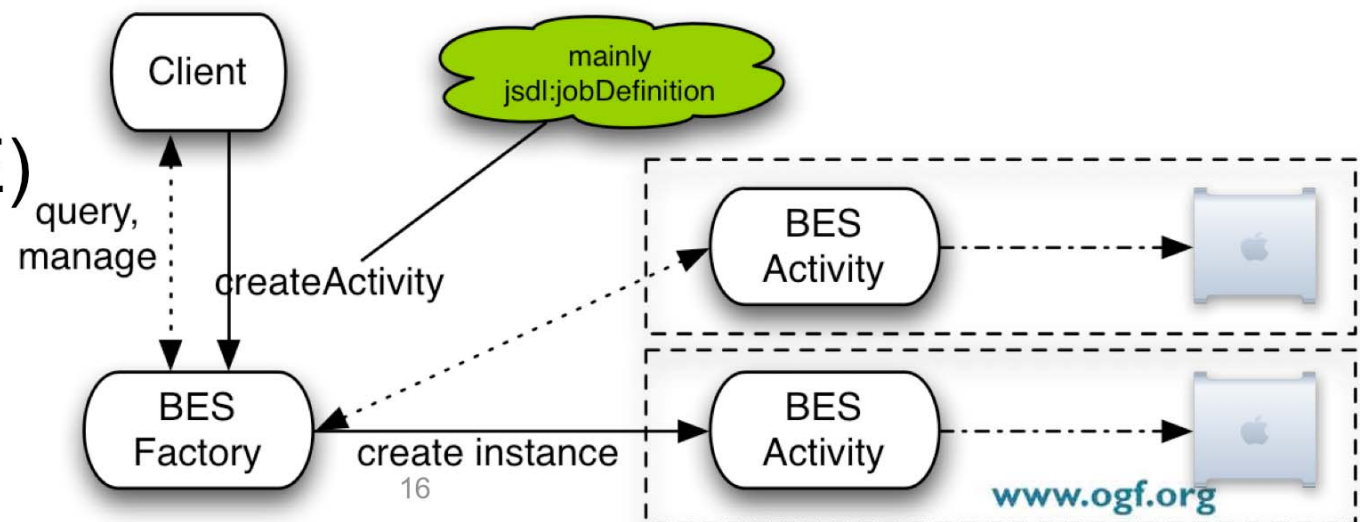
This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works.

The limited permissions granted above are perpetual and will not be revoked by the OGF or its successors or assignees.

Additional slides from previous sessions

# Activity - Status Quo

- Focus is on “Compute Activities”
  - Client defines Activity using JSDL
  - Activity is executed on a compute resource
- Related OGF specifications
  - OGSA-BES
  - JSDL
  - (GLUE)

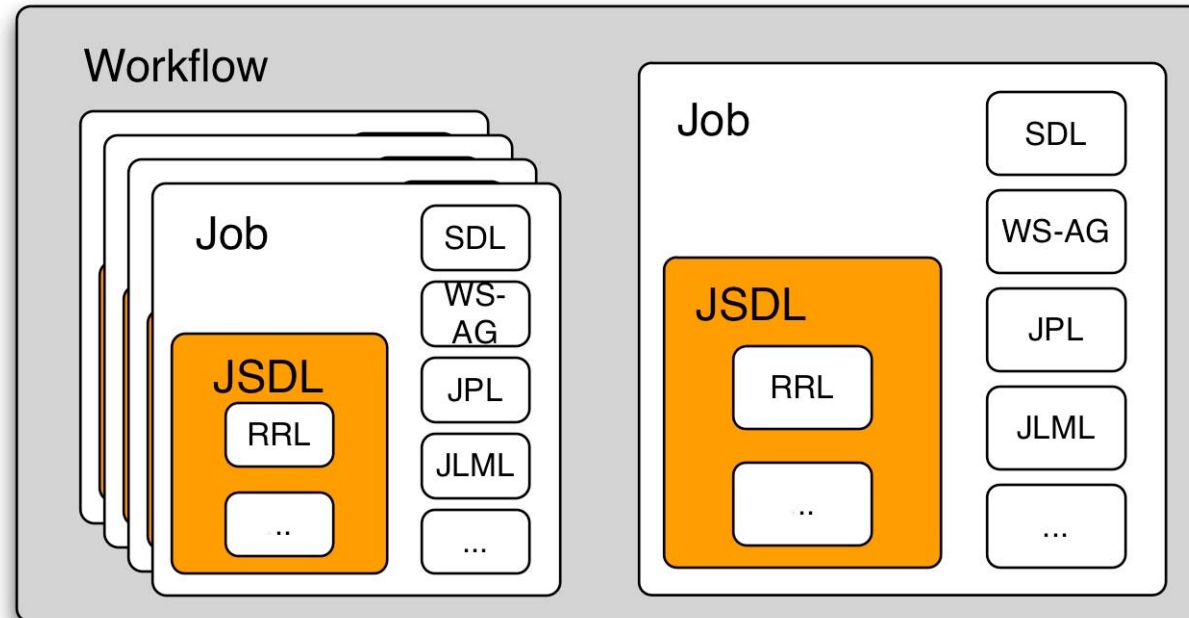


# Problem Domain - the Activity view

---

- Activity concept is broader than just compute
- Think about business scenarios using SLAs
  - Modelling SLA lifecycle
  - Creation and Monitoring of SLAs
  - Evaluation
- SLA = “Business Activity”
  - Comprises compute, but also ...
  - ... data, accounting info, provenance, lifecycle, etc.
- This information has to be captured and related to an activity

# Problem Domain - the JSDL view



JSDL comprises a core Resource Request Language and exists within an environment of other languages like Scheduling Description Language, Job Policy Language, etc. --> more requirements, dependencies, etc. wrt to an Activity

# Content

---



- Objectives of the BoF
- Status Quo of activity discussion
- Problem Domain
- **Universal Dynamic Activity Package**
  - Concept
  - Model
- Existing schema discussion

# UDAP - Activity Definition

---



- Activity:
  - A unit of work
  - A job
  - A task
  - A data processing operation
  - A data access operation
  - An application execution
  - A Web Service invocation
  - A “thing” you need to do/take care of/execute!
  - ...etc.
- From the OGSA-BES specification:
  - “...computational entities such as UNIX or Windows processes, Web Services, or parallel programs—what we call *activities*...”

# UDAP - Activity Definition

---



- We take a holistic view of an activity
- We, therefore, consider ALL that there is to know about an activity:
  - ALL of its requirements
  - ALL of its dependencies (on data and other activities)
  - ALL of its contextual information
    - Topical domain (financial markets, weather forecasting, etc.)
    - Security (who owns the activity, who is allowed to run it, etc.)
    - SLAs, QoS and other related policies
  - ALL of its monitoring information
    - Status, history, resource information, accounting, policy conformation etc.

# UDAP - Concept

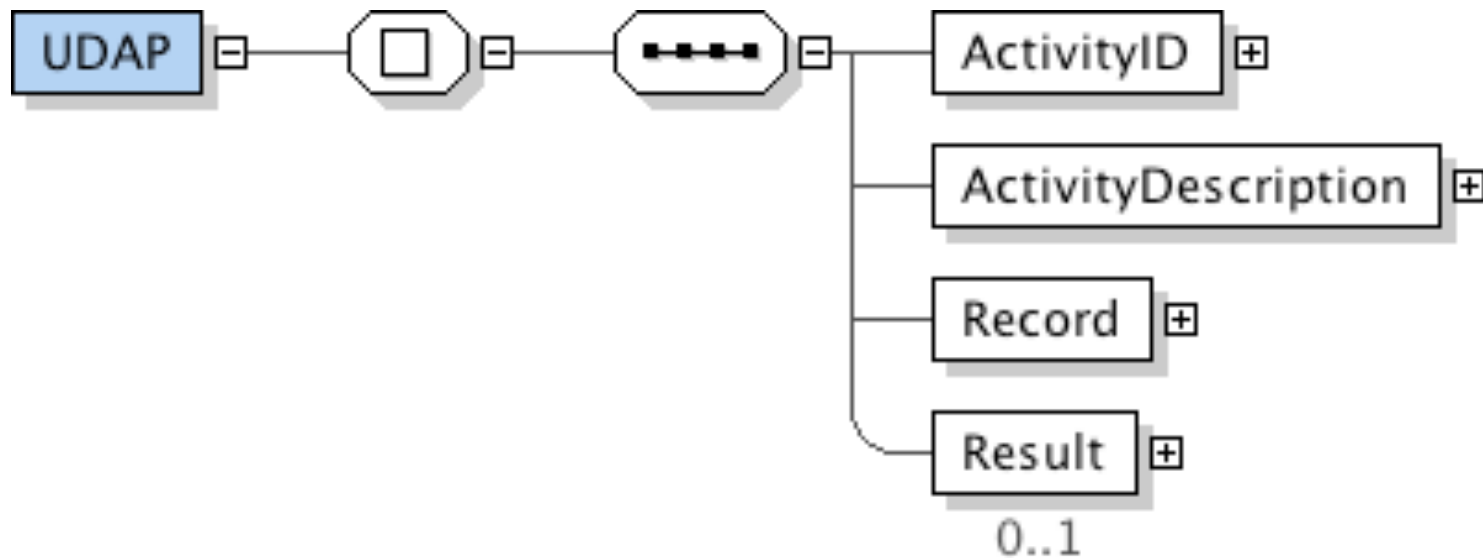
---



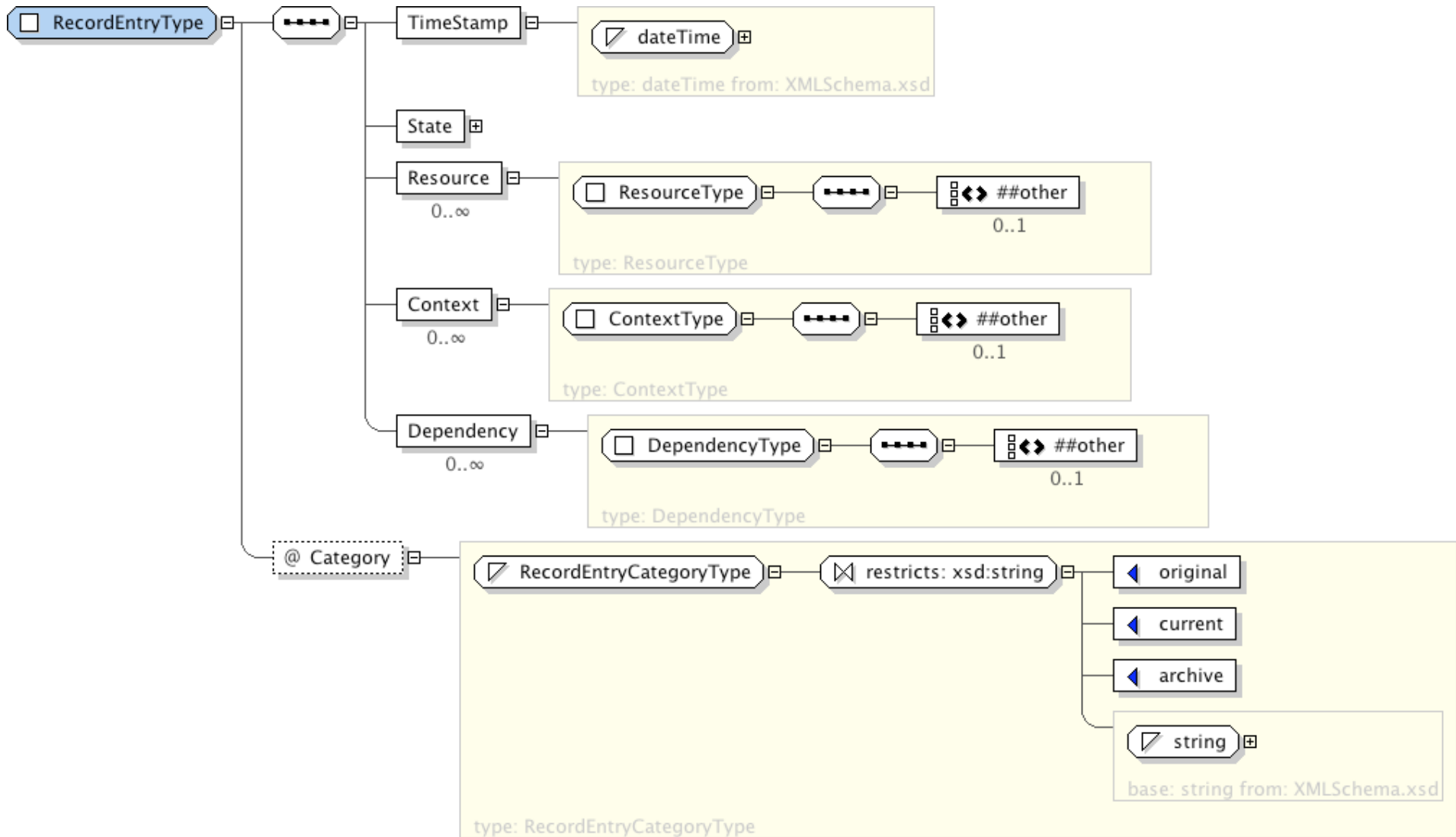
- The UDAP package can contain any information about an activity, regardless of the schema used to present that information
- The values of the activity information can be updated or appended to reflect the past, present, and future state of the activity
- All of the information associated with each activity is contained in a single package
- The information in a UDAP package is kept up-to-date for its activity, once it is submitted to and managed within a Grid

# UDAP Model - Overview

- ID and Description
- The Record is the core of an Activity
- Result is xsd:any



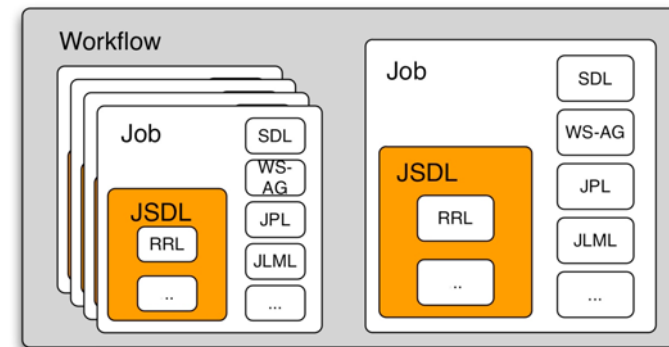
# UDAP Model - Record



# UDAP and JSDL



```
<UDAP ...>
  <ActivityID>ID1</ActivityID>
  <ActivityDescriptionDialect="JSDL">Example using JSDL</ActivityDescription>
  <Record>
    <EntryCategory="original">
      <TimeStamp>2006-05-04T18:13:51.0Z</TimeStamp>
      <State>pending</State>
      <Resource>
        <jSDL:jobDocument> ... </jSDL:jobDocument>
      </Resource>
      <Context>
        <jSDL:schedulingPolicy> ... </jSDL:schedulingPolicy>
        <jPL:somePolicy> ... </jPL:somePolicy>
      </Context>
      <Dependency>
        <wf:someWorkflowDependencies> ... </wf:someWorkflowDependencies>
      </Dependency>
    </Entry>
  </Record>
  <Result> </Result>
</UDAP>
```



# A word on Resource & Context



- “Resource” element in UDAP actually captures the *resource request*
  - Can be JSDL
  - Can be RSL, ...
- “Context” may contain the security, legal, billing, ... etc. context in which the activity is executed
  - Much broader
- Naming should be discussed ...

# We proudly present ...

---



... Steve McGough on GridSAM

... Shahbaz Memon on UNICORE

# Decisions

---

- Go for a *Activity Instance* Document schema
- Define it high-level, then profile it for specific OGF use cases
  - JSDL
  - Usage Records
  - ...
- Issues to resolve wrt UDAP schema