

BioinfoGRID Plug-in to Soaplab

The way how to integrate GRID

Martin Senger (martin.senger@gmail.com)

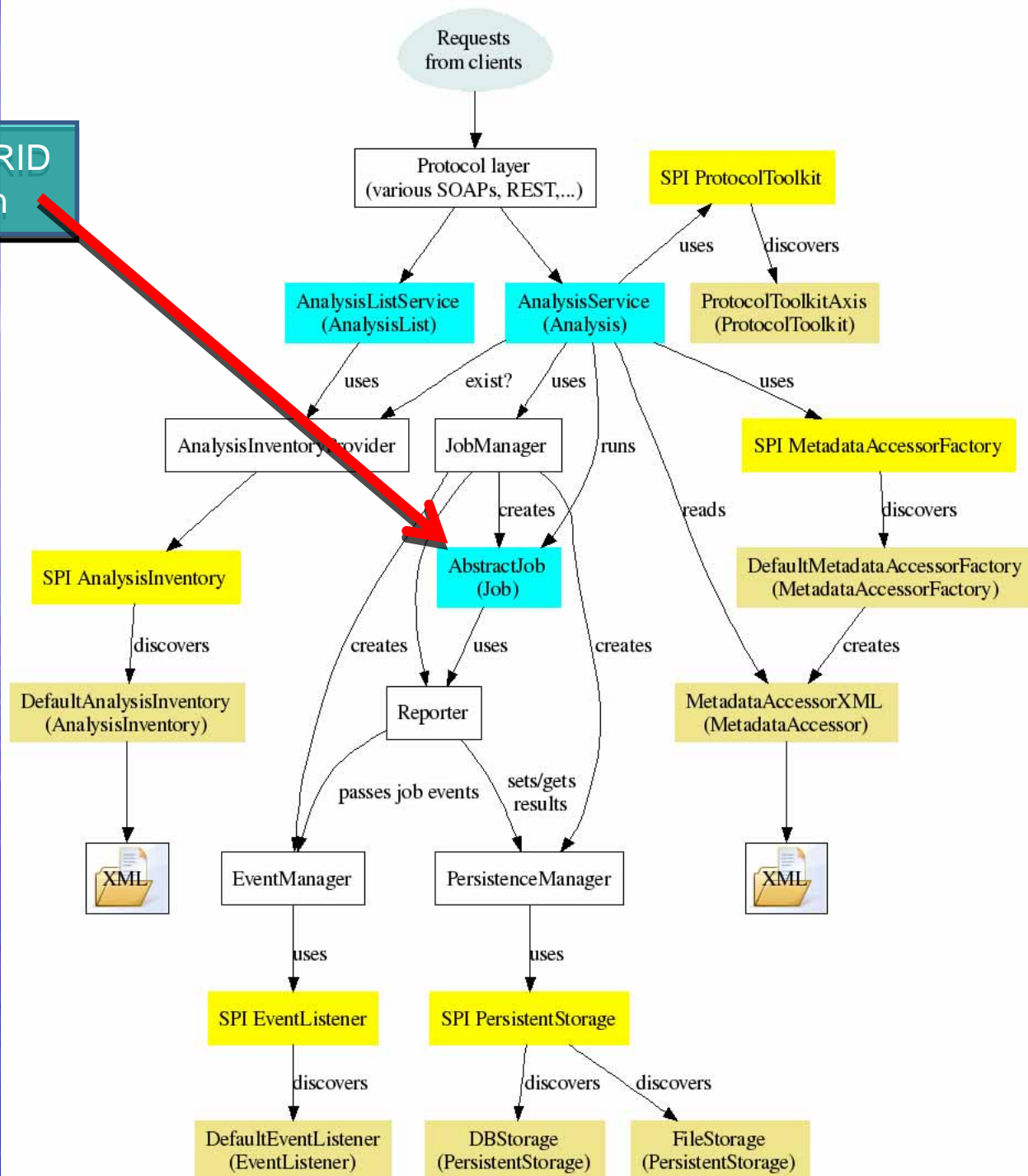
Motivation

- To re-use existing components
- To allow access to GRID from existing clients
 - Taverna
- To hide GRID complexity
 - to the end-users (as usual)
 - but also to service providers (hard)
- Main goal: to make GRID usable in the bioinformatics domains

What is Soaplab[2]

- A general framework for creating Web Services. What do they do?
 - they start and control various analysis tools (hundreds of them)
 - they go and fetch data from various web sites
 - **they start and control GRID jobs**
- A unified interface
 - which makes it easier to combine with various clients
- Services are driven by metadata
 - which makes the infrastructure more flexible

BioinfoGRID
Plug-in



Soaplab Architecture

BioinfoGRID Plug-in Features

- Fully configurable
- It generates JDL ("a GRID language") from user real inputs on-the-fly
- It has access to Soaplab2 features
 - results in permanent storage
 - event handlers mechanism
 - configurable by metadata
 - selection of protocols for user access

Resources

- Soaplab2
 - Open source project hosted at SourceForge
 - <http://soaplab.sourceforge.net/soaplab2/>
- BioinfoGRID
 - Open source project hosted at SourceForge
 - <http://soaplab.sourceforge.net/bioinfoGRID-plugin/>

Wedding Cake

Clients
(Taverna, W3H,
command-line,...)

Soaplab2 Framework

BioinfoGRID Plug-In

Gilda UI PlugAndPlay

GRID
(Hic sunt leones)

Bioinformatics
tools

Good and Bad News (1)

- UIPlugAndPlay Toolkit
 - Good: it works
 - Bad:
 - platform dependent (only Linux)
 - packaging makes it difficult to use
 - under-documented
 - Unknown:
 - is it still maintained?
 - should we replace it by something else? [WMProxy?]

Good and Bad News (2)

- GRID (globus, condor, glite, *whatever...*)
 - Good:
 - it is an excellent idea
 - enthusiastic and helpful community
 - Bad:
 - “an easy access” does not really work
 - it is hard to configure it (in order to make it work)
 - Unknown:
 - how does it cope with the real bioinformatics tools?

Good and Bad news (3)

- Soaplab2 infrastructure needs to be finalized
 - easy deployment
 - modern metadata generator
 - GUI for service providers
- BioinfoGRID Plug-in may be re-implemented
 - different GRID Java toolkit?
- Biological applications need to be plugged-in
 - **this is the Holy Grail**

Conclusions

- a need for dedicated GRID nodes
 - at least for the time of development
- get recommendations from this workshop