



# **BiBiServ 2 (*in the cloud*)**

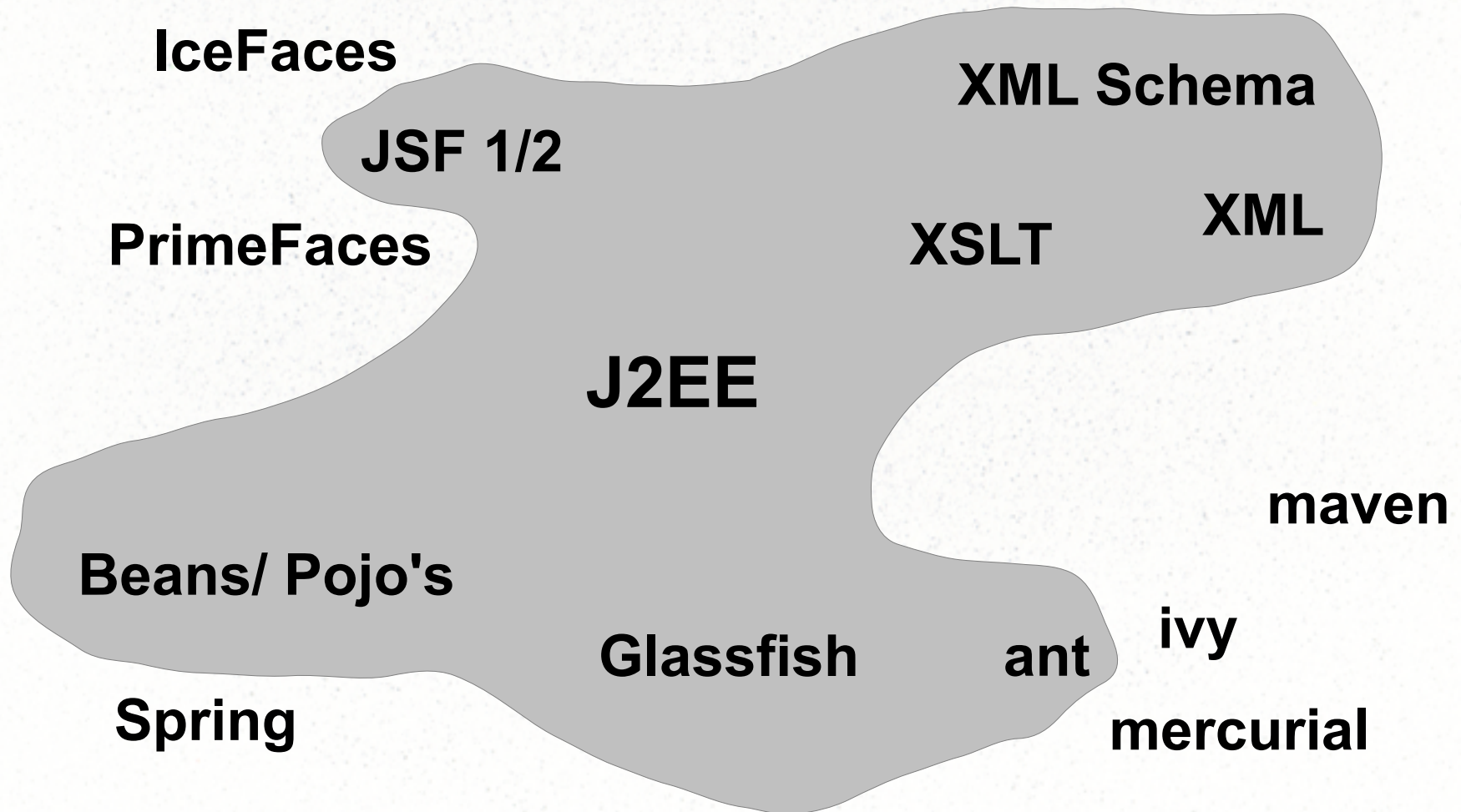
Jan Krüger

AGPI Seminar  
12.12.2011

# Overview

- Technical terms and definitions
- Historic (current) approach
- BiBiServ2 aims
- BiBiServ2 Component Overview & LiveDemo
- AmazonWebServices
- BiBiServ in the cloud – instantbibi
- What's missing ? Near future work to do ...

# Technical terms & definitions



# BiBiServ

- Apache WebServer + Perl/CGI
- J2EE Server + Axis WebServices
- Template based
- BiBiTools Libraries for Perl/Java
- (S|O)GE Support

# BiBiServ (advantages)

- Stable infrastructure
- Test environment
- (Relative) Simple
  - no web session
  - tools are independent from each other
  - static page generation
- Well known UI

# BiBiServ (disadvantages)

user point of view

- Layout sometimes follows CI  $\Leftrightarrow$  Which CI ?
- Sometimes bad error handling
- Often bad (or non existing) documentation
- Limited input data types
- Different meaning/handling for same data type (e.g. fasta)
- Bad tested form evaluation (mostly for complex tools)
- Often missing of (good) examples

# BiBiServ (disadvantages)

developer point of view

- Need a lot of understanding of web development
- Different technologies (perl/html + ajax + java)
- Two different server = Two error sources
- Test environment only internally available
- Developers hate/fears/love programming effort
  - Parameter and Input Validation
  - Error handling
- Nearly impossible to install test environment on a laptop
- Live debugging is impossible

# BiBiServ (disadvantages)

admin point of view

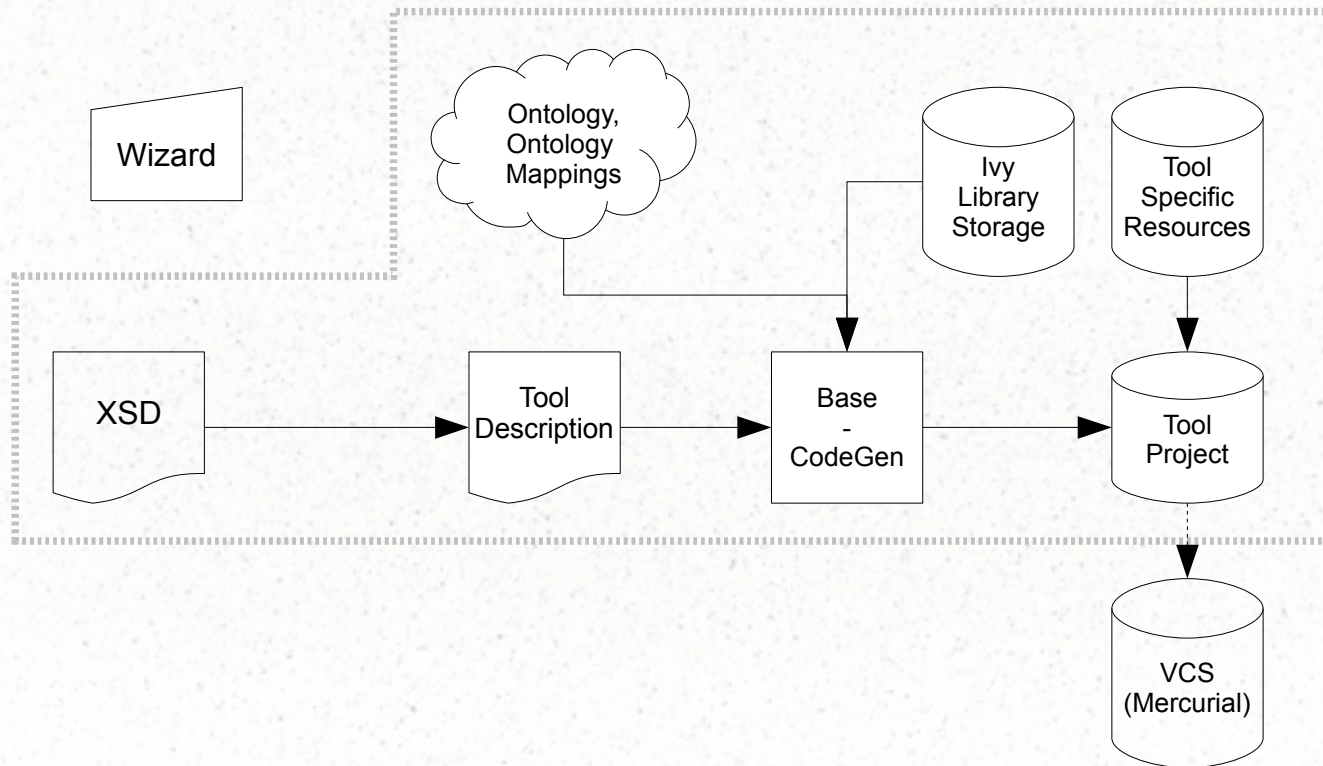
- 2 (4) server + database to maintain
- No dedicated database
- platform depended binaries (on webserver)
- Nearly impossible to install test environment on a laptop
- Webservice : axis 1.x based (= non standard data types, broken wsdl), SOAP::Lite
- Mostly no documentation of tools
- Porting to a new platform is very time consuming



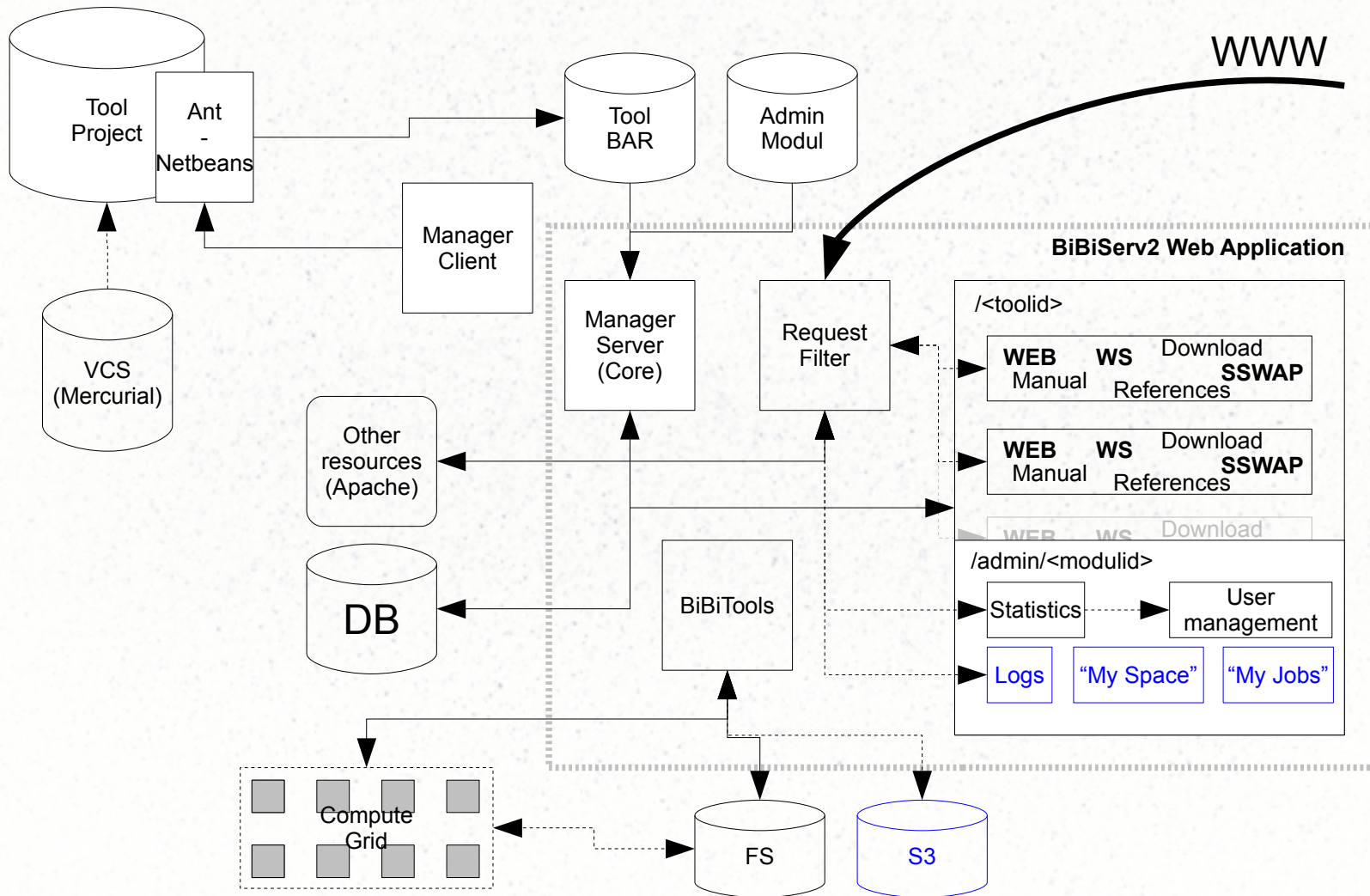
# BiBiServ2 aims

- Global layout /CI
- Well documented (manual, forms, errors), detailed tool description
- Persistent URL, long term support, different states of publication
- Modern, up-to-date UI
- Input data follows data type not format
- Developer :
  - Fast publishing with minimal programming effort
  - Content/Functionality “can” modified manually
- Platform independent server structure
- Easy server setup for local development

# BiBiServ2 overview (1)



# BiBiServ2 overview (2)



# Amazon Web Service



- AWS : set of online services
- Elastic Compute Cloud (EC2)
  - Various instance types from XS to XXXL
  - various AMI's
- Simple Storage Service (S3)
  - Key-value storage
  - Fast & independent from instances
  - various interfaces (SOAP, REST, API, Shell, ...)

# BiBiServ (in the Cloud)

bioinformatic server limitation

- Problems
  - huge (and fast growing) data set
  - limited bandwidth
  - Limited CPU time, Limited memory per job
- Solution
  - Move the server to the data instead the data to server

# BiBiServ (in the Cloud)

InstantBibi

- set up an application server is time consuming
  - Download needed resources
  - Configuration of Server/Database
  - Start server and deploy tools, extensions, ...
- 1<sup>st</sup> developed to full-fill one of BiBiServ2 aims → easy setup of a development system
- Extended for setup a end-user system, including example apps/bins for different platforms
- Run on any Unix based system with installed Java and ant build tool

# BiBiServ (in the Cloud)

InstantBibi

- Check java & ant version
- Get instantbibi
- Call `ant instant`
  - Get depended resources
  - Set up and configure domain && database
  - Start domain
- Call `ant install.google`
  - Get precompiled guugle resources
  - Install BAR and system depended tool binaries

# What's missing ? (1)

Future work to do ...

- improve Layout
- support a more “neutral” layout for non bibiserv usage
- support for cloud fs (e.g. Amazon S3)
- extend/rewrite MyBiBiServ area
- GUI/Module Manager
- Wizard
- enhanced statistics
- “MyJobs”
- “MySpace”



# What's missing ? (1)

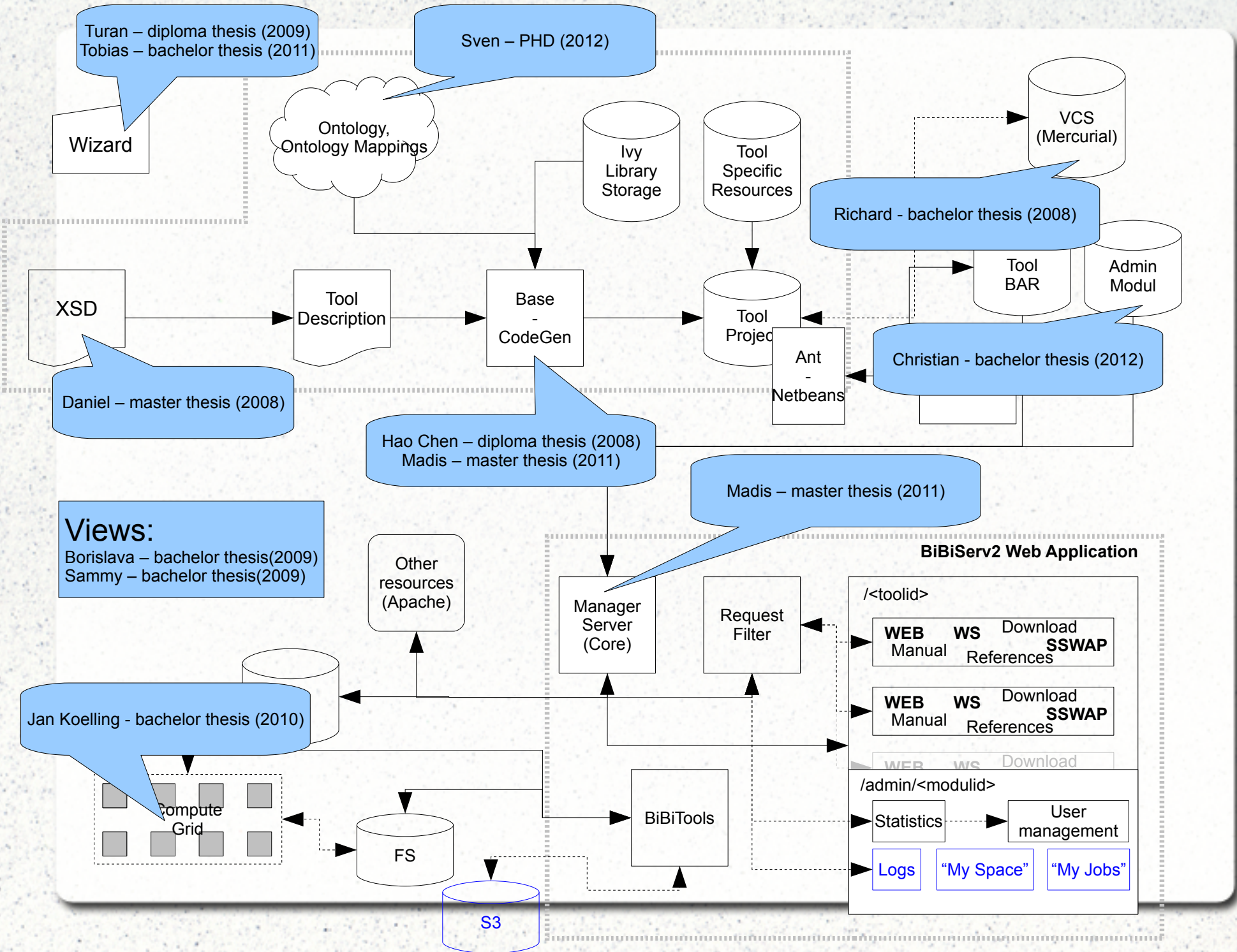
Future work to do ...


- improve Layout
- support a more “neutral” layout for non bibiserv usage
- support for cloud fs (e.g. Amazon S3)
- extend/rewrite MyBiBiServ area
- GUI/Module Manager
- Wizard
- enhanced statistics
- “MyJobs”
- “MySpace”

# What's missing ? (2)


Future work to do ...

- Views
  - Sequence Alignments – Raeda
  - Repeats – RepVis
  - RNA secondary structure – PseudoViewer Korea
  - RNA secondary structure – RNAMovies
- Runtime estimation





Thanks for your attention!



**BiBiServ 2 (*in the cloud*)**

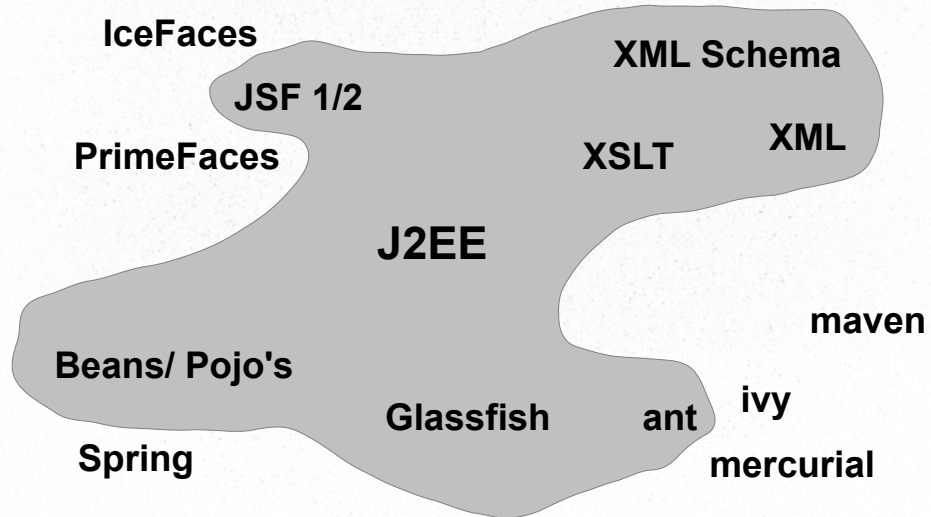
Jan Krüger

AGPI Seminar  
12.12.2011

# Overview

- Technical terms and definitions
- Historic (current) approach
- BiBiServ2 aims
- BiBiServ2 Component Overview & LiveDemo
- AmazonWebServices
- BiBiServ in the cloud – instantbibi
- What's missing ? Near future work to do ...

## Technical terms & definitions



## **BiBiServ**

- Apache WebServer + Perl/CGI
- J2EE Server + Axis WebServices
- Template based
- BiBiTools Libraries for Perl/Java
- (S|O)GE Support



## **BiBiServ (advantages)**

- Stable infrastructure
- Test environment
- (Relative) Simple
  - no web session
  - tools are independent from each other
  - static page generation
- Well known UI

## **BiBiServ (disadvantages)**

user point of view

- Layout sometimes follows CI  $\Leftrightarrow$  Which CI ?
- Sometimes bad error handling
- Often bad (or non existing) documentation
- Limited input data types
- Different meaning/handling for same data type (e.g. fasta)
- Bad tested form evaluation (mostly for complex tools)
- Often missing of (good) examples

## **BiBiServ (disadvantages)**

developer point of view

- Need a lot of understanding of web development
- Different technologies (perl/html + ajax + java)
- Two different server = Two error sources
- Test environment only internally available
- Developers hate/fears/love programming effort
  - Parameter and Input Validation
  - Error handling
- Nearly impossible to install test environment on a laptop
- Live debugging is impossible

## **BiBiServ (disadvantages)**

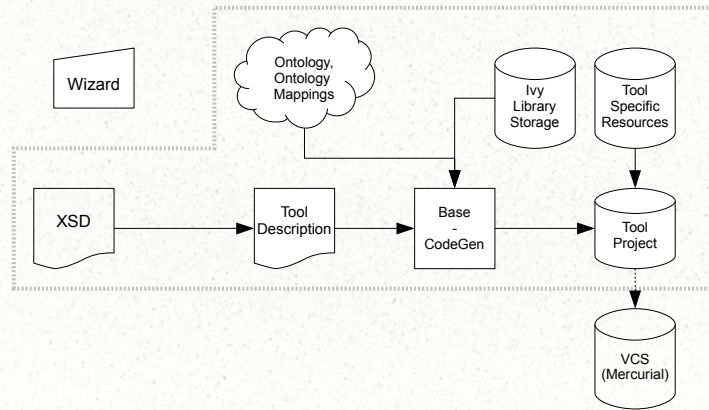
admin point of view

- 2 (4) server + database to maintain
- No dedicated database
- platform depended binaries (on webserver)
- Nearly impossible to install test environment on a laptop
- Webservice : axis 1.x based (= non standard data types, broken wsdl), SOAP::Lite
- Mostly no documentation of tools
- Porting to a new platform is very time consuming

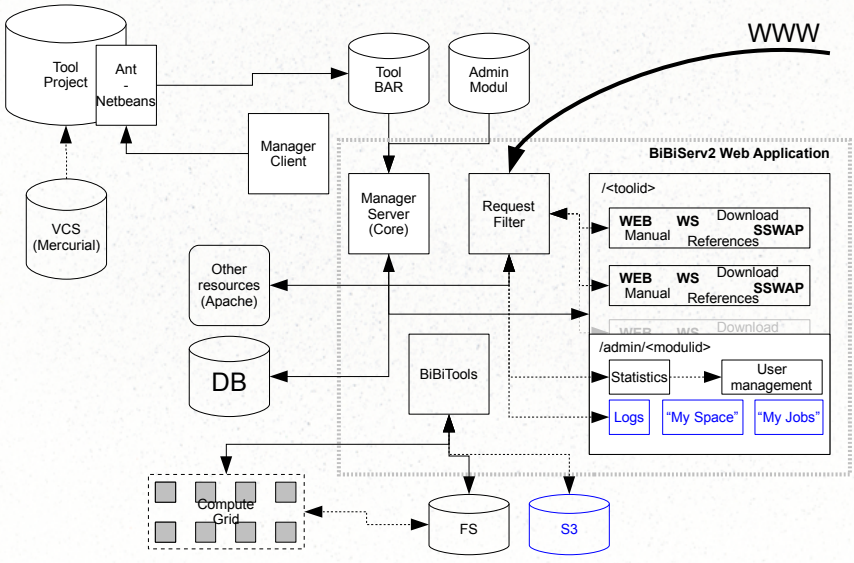
## **BiBiServ2 aims**

- Global layout /CI
- Well documented (manual, forms, errors), detailed tool description
- Persistent URL, long term support, different states of publication
- Modern, up-to-date UI
- Input data follows data type not format
- Developer :
  - Fast publishing with minimal programming effort
  - Content/Functionality “can” modified manually
- Platform independent server structure
- Easy server setup for local development

# BiBiServ2 overview (1)



# BiBiServ2 overview (2)



# Amazon Web Service



- AWS : set of online services
- Elastic Compute Cloud (EC2)
  - Various instance types from XS to XXXL
  - various AMI's
- Simple Storage Service (S3)
  - Key-value storage
  - Fast & independent from instances
  - various interfaces (SOAP, REST, API, Shell, ...)



# **BiBiServ (in the Cloud)**

bioinformatic server limitation

- **Problems**
  - huge (and fast growing) data set
  - limited bandwidth
  - Limited CPU time, Limited memory per job
- **Solution**
  - Move the server to the data instead the data to server

# **BiBiServ (in the Cloud)**

InstantBibi

- set up an application server is time consuming
  - Download needed resources
  - Configuration of Server/Database
  - Start server and deploy tools, extensions, ...
- 1<sup>st</sup> developed to full-fill one of BiBiServ2 aims → easy setup of a development system
- Extended for setup a end-user system, including example apps/bins for different platforms
- Run on any Unix based system with installed Java and ant build tool

# BiBiServ (in the Cloud)

InstantBibi

- Check java & ant version
- Get instantbibi
- Call `ant instant`
  - Get depended resources
  - Set up and configure domain && database
  - Start domain
- Call `ant install.google`
  - Get precompiled guugle resources
  - Install BAR and system depended tool binaries

# What's missing ? (1)

Future work to do ...

- improve Layout
- support a more “neutral” layout for non bibiserv usage
- support for cloud fs (e.g. Amazon S3)
- extend/rewrite MyBiBiServ area
- GUI/Module Manager
- Wizard
- enhanced statistics
- “MyJobs”
- “MySpace”

# What's missing ? (1)

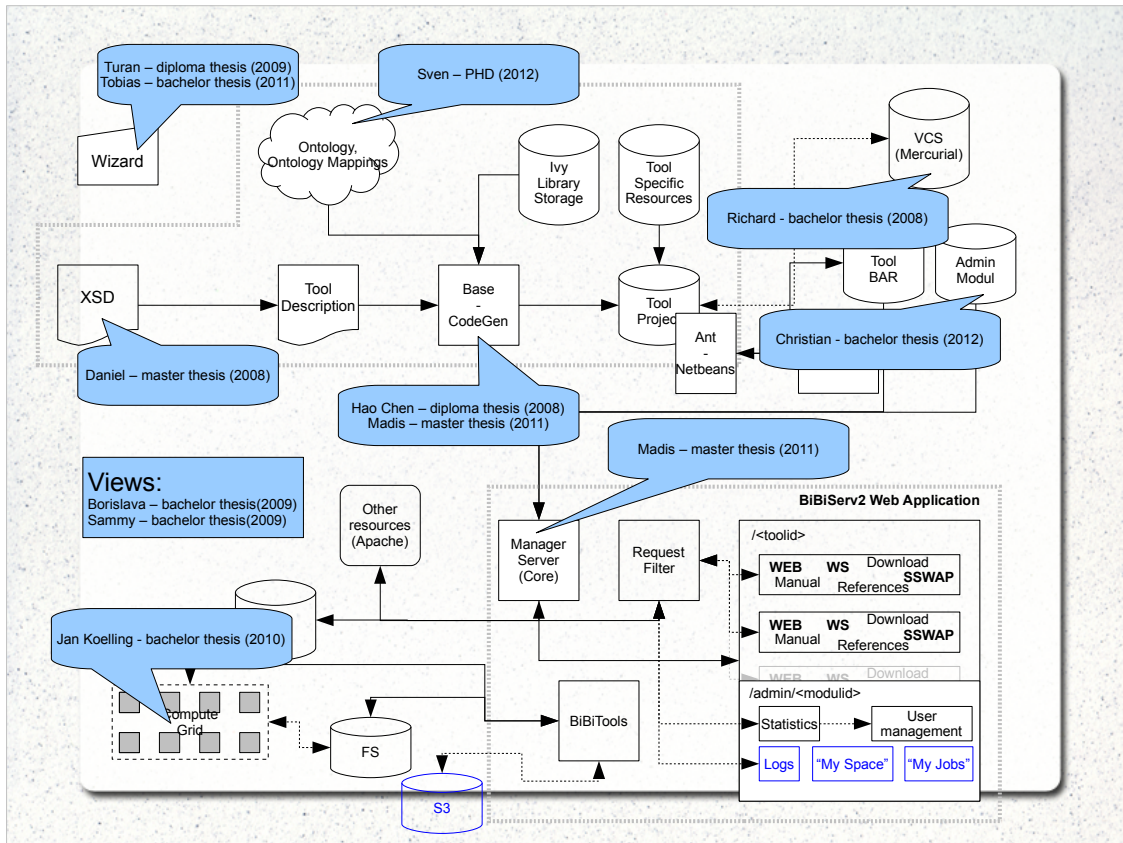
Future work to do ...

- improve Layout
- support a more “neutral” layout for non bibiserv usage
- support for cloud fs (e.g. Amazon S3)
- extend/rewrite MyBiBiServ area
- GUI/Module Manager
- Wizard
- enhanced statistics
- “MyJobs”
- “MySpace”

## What's missing ? (2)

Future work to do ...

- Views
  - Sequence Alignments – Raeda
  - Repeats – RepVis
  - RNA secondary structure – PseudoViewer Korea
  - RNA secondary structure – RNAMovies
- Runtime estimation



Thanks for your attention!

