



Crank Web Interface

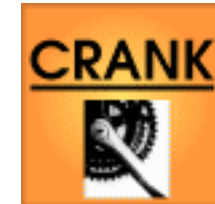
Remote automated structure solution

Reint Boer Iwema

Leiden University

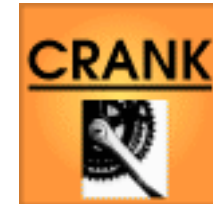
MAX-INF2 – 29 November 2007

Contents



- Crank
 - Automated macromolecular structure determination
 - Two interfaces: command line and CCP4i
 - Adding third interface: web-based (CWI)
- Benefits
- How does it work?
 - Stepwise set-up
 - Leiden Grid Infrastructure
 - Communication between CWI and LGI
- Where are we now
- What's next?

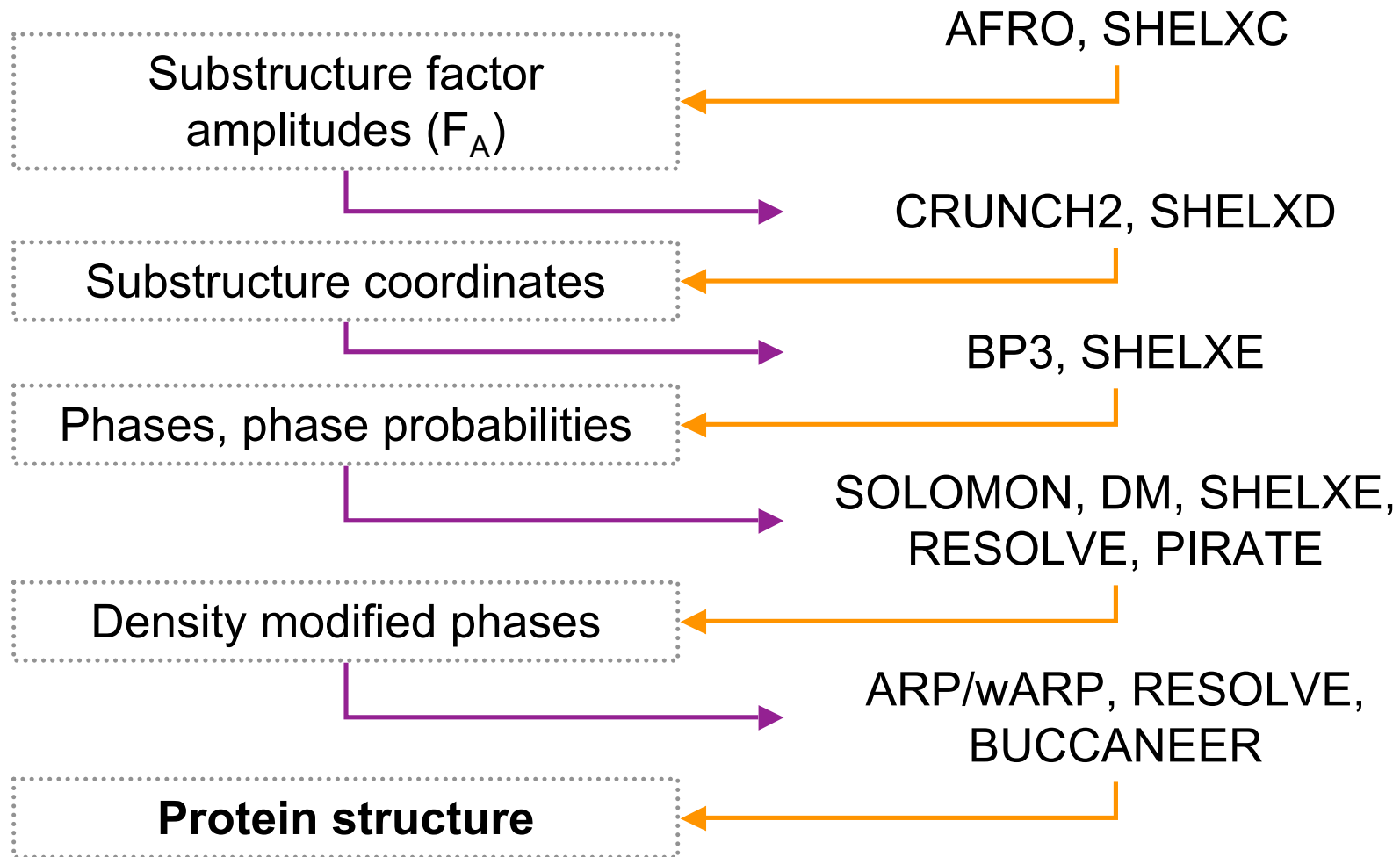
Crank



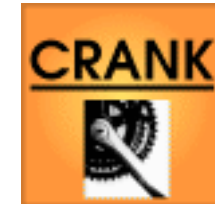
- Automated macromolecular structure determination for SAD, SIRAS, MAD, and MAD+native experiments
- Interfaces with various crystallographic programs

www.bfsc.leidenuniv.nl/software/crank

Crank : programs

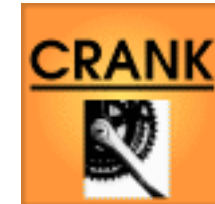


Crank / BP3 : many test cases



- 33 test cases on SAD, MAD, SIRAS data show robustness and speed
 - www.bfsc.leidenuniv.nl/software/crank/tests.html
- Currently testing JCSG-datasets as well
 - www.jcsg.org
- Published results show in 8 out of 10 real SAD cases, BP3 produced better results over SHARP, MLPHARE or SOLVE
 - Ness, de Graaff, Abrahams, Pannu (2004) *Structure*, **12**, 1753-1761

Crank : how to get it



- Download from BFSC or CCP4 (pre-release zone)
 - Suitable for SAD, SIRAS, MAD and MAD+native experiments
 - Supported platforms: Linux, Mac OS X (Intel, PPC)
 - CRUNCH2, BP3 and most of Crank is GPL
- Version 1.2.0 is available: *much* improved over earlier versions
- Interfaces to BP3 exist in AutoRickshaw (EMBL Hamburg pipeline) and MIFit

Crank : two interfaces



CRANK

Help

Title

MTZ in Full path.. Browse View

MTZ out PROJECT Browse View

Input Amplitudes Setup experiment SAD

Input protein sequence

SEQ in Full path.. Browse View

DNA/RNA present

Crystal # 1

Native Substructure atom Se Number of substructure atoms per monomer 2

Dataset : 1 Type SAD Anomalous

Data collected at CuK α wavelength

f' -7.6 f'' 4.0

FP+	<input type="text" value="F_peak(+)"/>	<input type="text"/>	SIGFP+	<input type="text" value="SIGF_peak(+)"/>	<input type="text"/>
FP-	<input type="text" value="F_peak(-)"/>	<input type="text"/>	SIGFP-	<input type="text" value="SIGF_peak(-)"/>	<input type="text"/>

Derived parameters

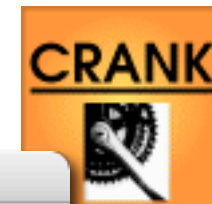
Experimental Pipeline

Start the pipeline with Substructure detection and end with Model building

Pipeline :

Display individual program options

Crank : two interfaces



```
Terminal — bash — 100x30
#####
### CCP4 5.0: gcx_120          version 0.6      :      ##
#####
User: reint  Run date: 25/11/2007 Run time: 22:43:21

Please reference: Collaborative Computational Project, Number 4. 1994.
"The CCP4 Suite: Programs for Protein Crystallography". Acta Cryst. D50, 760-763.
as well as any specific reference in the program write-up.

                GCX
Generate a crank xml file to run crank via scripts
Developmental Version
              7 April 2006

      http://www.bfsc.leidenuniv.nl/software/

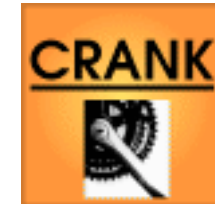
Xtal DER1
  Data line--- Xtal DER1
ATOM Se
  Data line--- ATOM Se
NUMB 2
  Data line--- NUMB 2
DNAME PEAK
  Data line--- DNAME PEAK
COLUMN F+=F_peak(+) SF+=SIGF_peak(+) F-=F_peak(-) SF-=SIGF_peak(-)
  Data line--- COLUMN F+=F_peak(+) SF+=SIGF_peak(+) F-=F_peak(-) SF-=SIGF_peak(-)
FORM Se FP=-7.6 FPP=4
  Data line--- FORM Se FP=-7.6 FPP=4
END
```

Crank : third interface



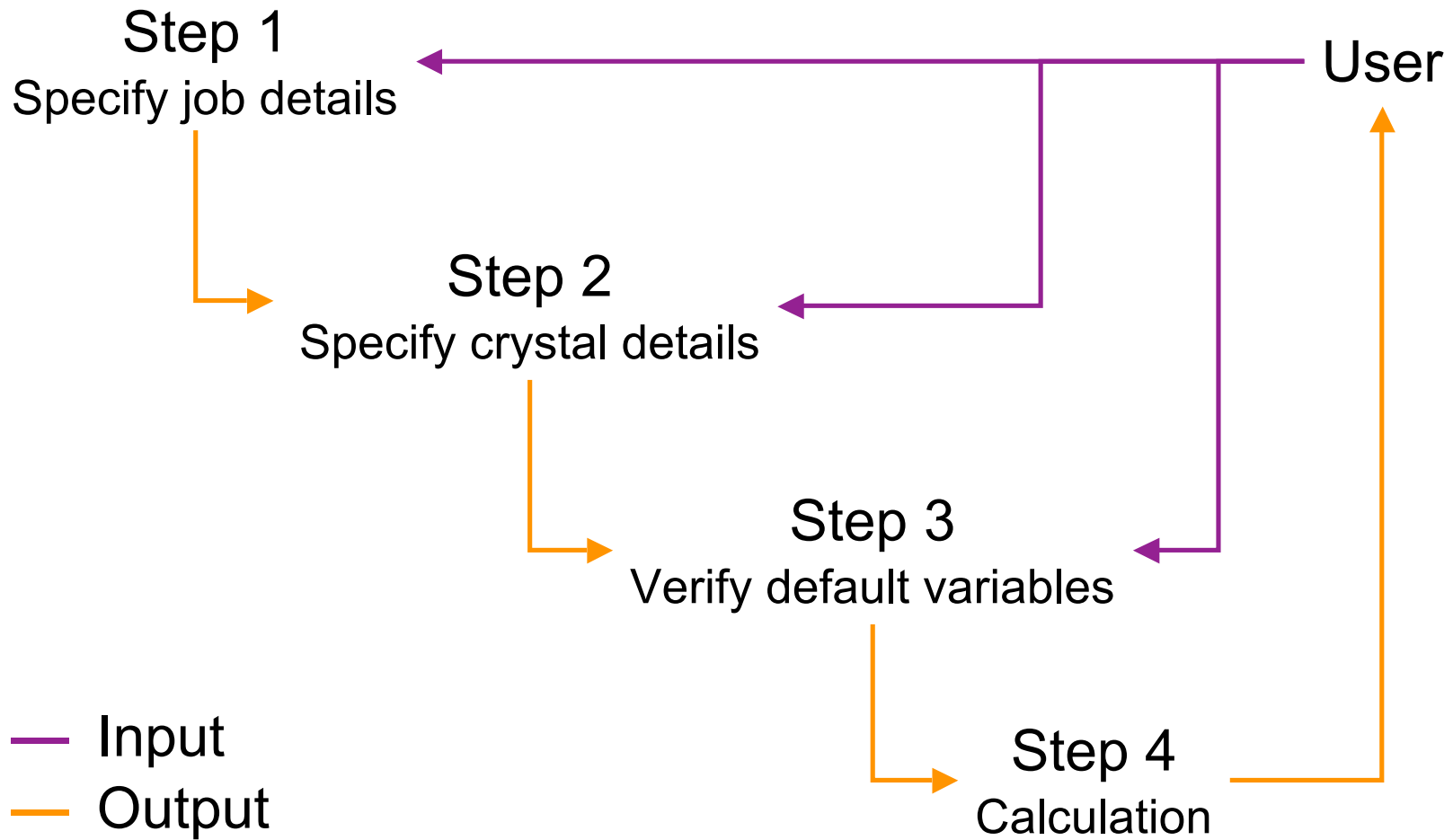
- A third interface to Crank has been added
 - Interface Crank to computing cluster for remote job submission
 - Works like CCP4i interface

Benefits



- Focus on the science
 - No need to set-up and administer crystallographic software
- Faster computation
- Leiden programs are “free”
 - If user verifies ARP/wARP, CCP4, SHELX license, user can run those suites
- Remote job submission has shown its merits
 - Web-based ARP/wARP / Auto-Rickshaw

How does it work



How does it work : step 1



Crank Web Interface

https://localhost:8443/crank/ Google

Job options

Title

MTZ in

Input Setup experiment

Input sequence

Sequence in

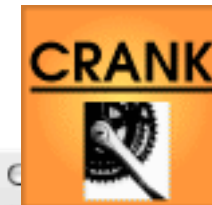
DNA/RNA present

Start pipeline with and end with

Pipeline

I do have the necessary licenses to run the specified pipeline.

How does it work : step 2



Crank Web Interface

https://localhost:8443/crank/index. Google

MTZ-file out: **gere_MAD_nat_output.mtz**

Sequence-file in: **gere_MAD_nat.pir**

Crystal #1

Native

Substructure atom Number of substructure atoms per monomer

Dataset 1: Type Anomalous

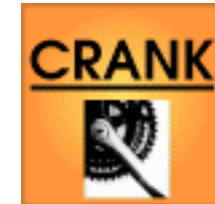
Data collected at CuK α wavelength

f' f''

FP+ SIGFP+

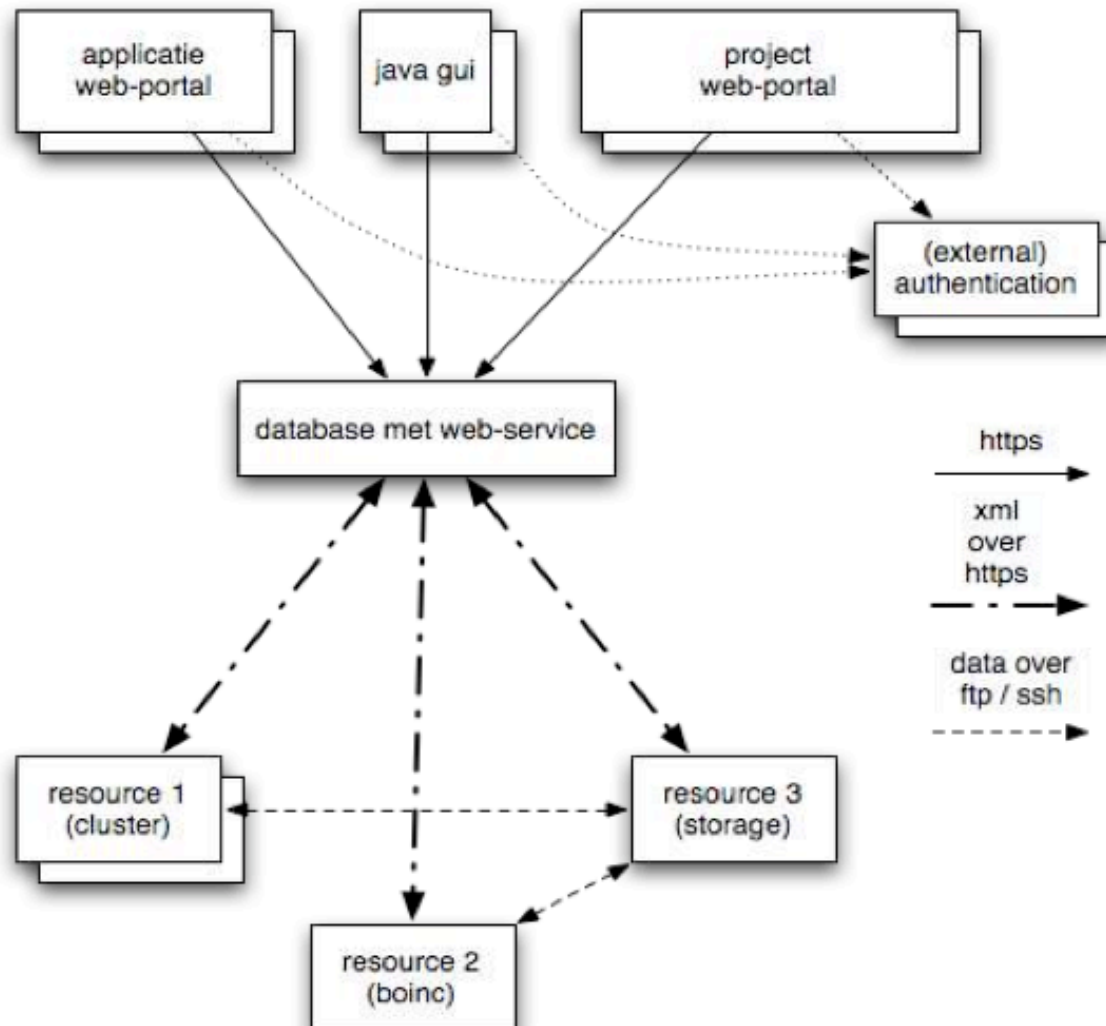
FP- SIGIP-

LGI (Leiden Grid Infrastructure)



- Lightweight grid middleware
 - Communication between job submission (front-end) and computing grid (backend)
 - Distribution of jobs to various resources
 - Use of standards: HTTPS and XML
- Make programs more intuitive and accessible for most common tasks
- Provide scalability and continuity without interference with the user

LGI (Leiden Grid Infrastructure)



LGI (Leiden Grid Infrastructure)



https://localhost:8443/lgi/basic_interface/basic_interface_job_state.php

https://localhost:8443/lgi/basic_interface/ Google

Leiden Grid Infrastructure basic interface at 29 Nov 2007 7:44 UTC

Project:	cyttron
This project server:	https://localhost:8443/lgi
Project master server:	https://localhost:8443/lgi
User:	reint@home
Groups:	reint@home
Application:	any
State:	any
Start index:	0
Index limit:	64
Number of jobs listed:	1

job id	job state	target resources	application	time stamp
103	finished	reint@localhost	crank	27 Nov 2007 7:56 UTC

[Show project server list](#)
[Show project resource list](#)
[Submit a job](#)
[Go to main menu](#)

LGI (Leiden Grid Infrastructure)



https://localhost:8443/lgi/basic_interface/bas...p?job_id=103&groups=reint@home&project=cyttron

https://localhost:8443/lgi/basic_interface/basic_int

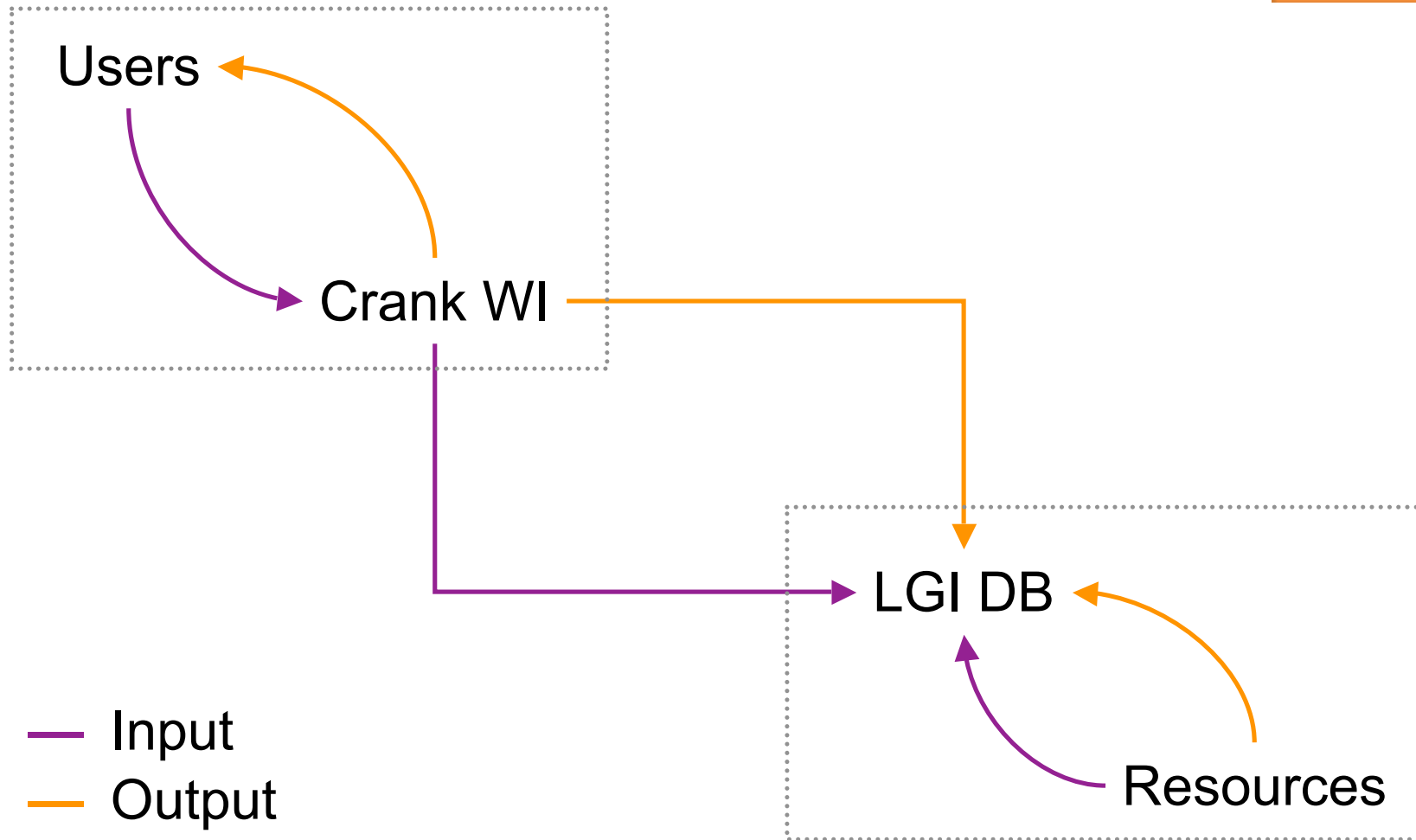
Leiden Grid Infrastructure basic interface at 29 Nov 2007 9:15 UTC

Project:	cyttron
This project server:	https://localhost:8443/lgi
Project master server:	https://localhost:8443/lgi
User:	reint@home
Groups:	reint@home

Job details

Job ID:	103
Application:	crank
State:	finished
State time stamp:	27 Nov 2007 7:56 UTC
Owners:	reint@home
Read access:	reint@home, cyttron, bfsc
Target resources:	reint@localhost
Job specifics:	
Input:	<code><crank_lgi><input> /Applications/MAMP/htdocs/crank/temp/reint/0f70d0d225c1ad23d67dbf84fb748ef6/ </input><basename> gere_MAD_nat_out </basename><web_user> reint </web_user><web_url> localhost </web_url></crank_lgi> <crank lei><input></code>

Communication CWI \leftrightarrow LGI



Where are we now



- Simple job submission with default variables
- Calculation on BFSC computing cluster
- Send results back to user

What's next



- Finalise step 3
- User/session management
- Interface with Cyttron Visualisation Platform (CVP)
- A lot of testing ...

Acknowledgements



- Raj Pannu, Irakli Sikharulidze, Pavol Skubak, Roel Bolsius, Jan Pieter Abrahams (BFSC)
- Mark Somers, Hugo Meiland (LGI)
- Others
 - All dataset contributors
 - Garib Murshudov, George Sheldrick, Victor Lamzin, Tassos Perrakis, Tom Terwilliger, Kevin Cowtan, Paul Emsley and their co-workers



Universiteit Leiden

