WISDOM-II, status of preparation

Nicolas Jacq

LPC of Clermont-Ferrand, CNRS/IN2P3

Life Science session, EGEE'06

Geneva, 26.09.2006

www.eu-egee.org

Information Society

- **Several laboratories have expressed an interest in proposing targets for a second computing challenge against neglected diseases.**

- **Several grid projects have expressed interest in contributing to the WISDOM initiative by providing computing resources.**

- **A rough estimation of the needed resources is about 500 CPU years and about 4 terabytes storage.**

- **Date: from October, 1st to December, 15th**

Enabling Grids for E-sciencE

Organism	Target	Partner	Country	Status
P. *falciparum*	GST	U. of Pretoria	South-Africa	Ready
P. *falciparum*	DHFR	U. di Modena e Reggio Emilia	Italia	Under preparation
P. *vivax*	DHFR	U. of Los Andes	Venezuela	Under preparation
Plasmodium/ plant/mamal	Tubulin	CEA, Acamba project	France	Accepted

- **Other prospected targets**
  - P. falciparum, DNA-polymerase – University of Glasgow, United Kingdom
  - Leishmania, transketolase – University of Glasgow, United Kingdom
  - Leishmania, MAP-Kinase1 - Bernhard Nocht Institute for Tropical Medicine, Germany

- **Known targets**
  - DHFR from P. *falciparum*
    - Known target of the Chloroquine drug
    - Chloroquine now inefficient against drug-resistance parasites due to DHFR mutations
    - Search for new drugs
  - DHFR from P. v*ivax*
    - Plasmodium Vivax responsible for malaria in South America
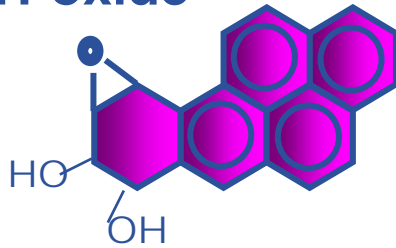   - *Malaria case in Corsica (France) during the summer*

- **New targets**
  - Tubulin from P. *falciparum*
    - Tubulin involved in cell replication, potential target for cancer treatment
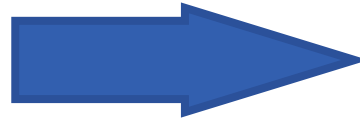   - Comparative docking of human and P. *falciparum*. tubulins
  - GST from P. *falciparum*
    - Protein involved in parasite detoxification
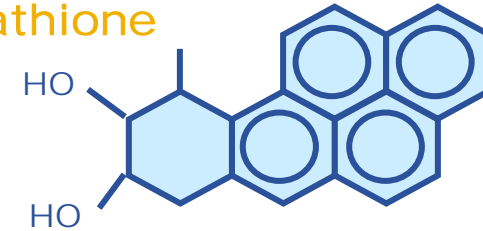   - Search for inhibitors

PAH-oxide

Glutathione

GST
+ Glutathione

DNA Reactive

Inactive

- **Role: Involved in detoxification phase in Plasmodium**

- **Features:**
  - Co-crystallized structure available in PDB with GTX
  - Dimer
  - Binding site known
  - 4 inhibitors known

- **Problem: present in almost all organisms (human…)**

- **BioSolveIT make graciously available FlexX docking software for all the data challenge**

- **Difference with previous data challenges:**
  - More targets but less structures and parameters
  - FlexX is much faster than Autodock
  - Much higher virtual screening throughput expected
    - Present record from WISDOM-I: 11,6 docking computations per second

- **Room for additional deployment within the 500 CPU years requested**
  - Docking of new targets
  - Grid-enabled molecular dynamics
    - SCAI Fraunhofer
    - Reranking of WISDOM-I results (BioInfoGRID, University of Modena)

- **3,000 licenses will be available during the data challenge**

- **FLEXlm license server is maintained in SCAI Fraunhofer**

- **A second server will be installed in SCAI Fraunhofer**

- **Ports 23,200 and 23,201 required to be open on each Worker Node (ports already open for Globus)**

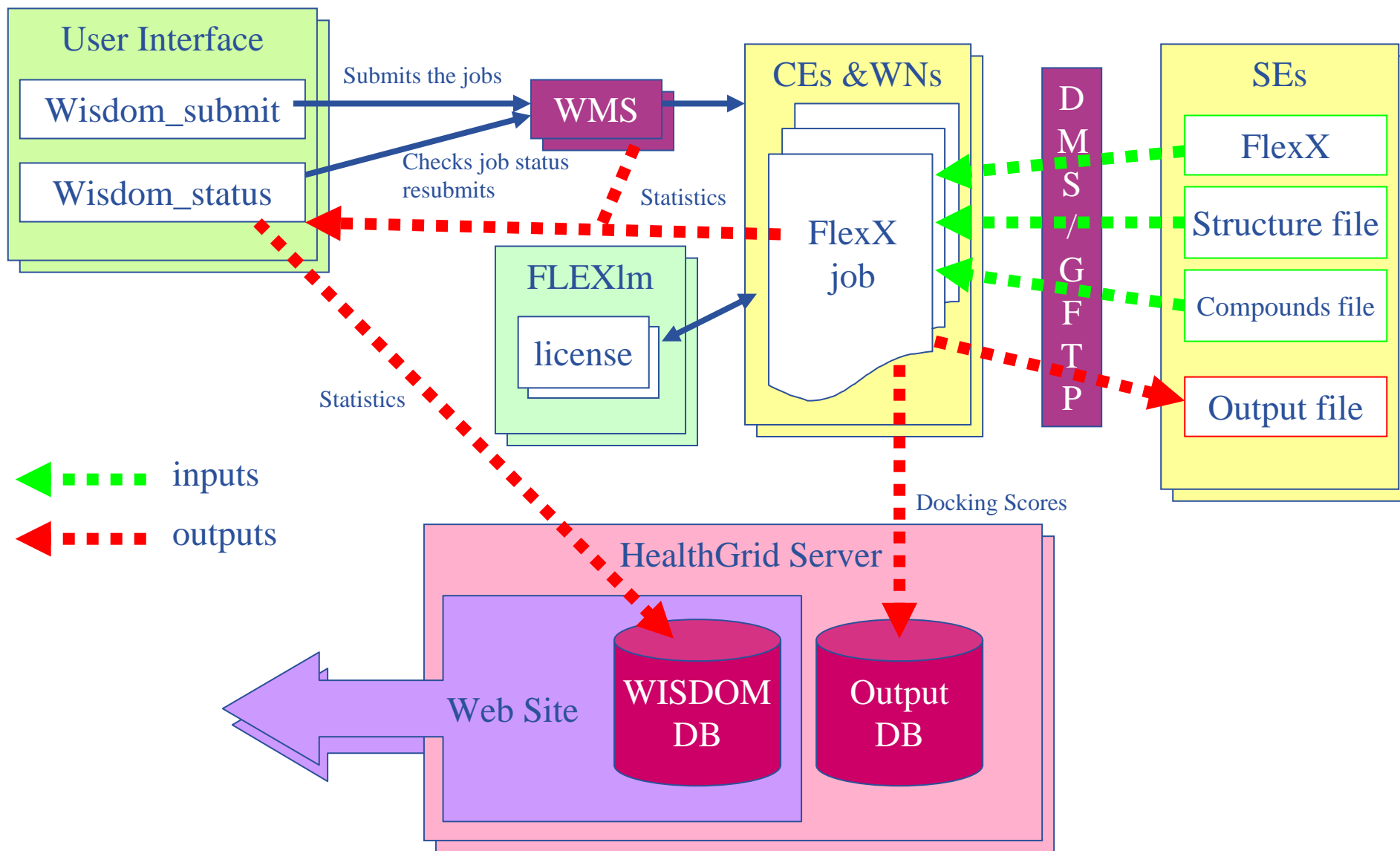- **Tests on all grid nodes are required before the data challenge**

- **2 structures**
  - Chain A and B of the dimer

- **2 Flexx parameters**
  - Place particle: on/off
  - Max overlap volume: 2.5/5.0
  - No crystal water

- **4.3 million compounds from ZINC**
  - Random approach

- **FlexX software**
  - 30s by docking

- **4 different instances with 2400 jobs of 15h**

Infrastructure/ project	Experiment operator	Target to be deployed	CPUs contribution
EGEE	LPC, Embrace	All	Available: ~8000 CPUs Free CPUs: ~2000
BioInfoGRID	BioInfoGRID	All	?, included in Biomed
TWGrid	TWGrid	All	300, included in Biomed
Auvergrid	LPC	GST	500
EELA	UPV, ULA	DHFR *vivax*	500 (CIEMAT)
EUMedGRID	INFN ?	DHFR *falciparum*	?
EUChinaGRID	INFN ?	Tubulin	?

- **2 main scripts, running in parallel: wisdom_submit, wisdom_status**

- **No more input and output sandboxes in jobs in order to avoid RB overload or to prevent RB crash**

- **Job JDL and scripts are generated just before any submission in order to take CE/RB black list or job submission frequency modifications into account**

- **Dynamic insertions of docking scores and statistics in relational databases which allow real-time visualisation**

eGee

User Interface

Wisdom_submit

Wisdom_status

Submits the jobs

Checks job status
resubmits

Statistics

WMS

CEs &WNs

FlexX
job

FLEXlm

license

Statistics

D
M
S
/
G
F
T
P

SEs

FlexX

Structure file

Compounds file

Output file

inputs

outputs

Docking Scores

HealthGrid Server

Web Site

WISDOM
DB

Output
DB

- **October, 1: GST target on EGEE-biomed and Auvergrid**

- **October, 15: DHFR vivax on EGEE-biomed and EELA**
- **October, 15: DHFR falciparum on EGEE-biomed and EUMedGRID**

- **November, 1: Tubulin on EGEE-biomed and EUChinaGRID**

- **November, 1: MD (SCAI Fraunhofer)**
- **November, 15: MD (BioInfoGRID, University of Modena)**

Credits

Academia Sinica
BioSolveIT
CNR-ITB
CNRS
CEA
Healthgrid
IN2P3
LPC
SCAI Fraunhofer
Università di Modena e Reggio Emilia
Université Blaise Pascal
University of Pretoria
University of Los Andes

WISDOM
Initiative for grid-enabled drug discovery
against neglected and emergent diseases

Auvergrid
Accamba project
BioInfoGRID
Conseil Regional d'Auvergne
EGEE
Embrace
EUChinaGRID
EUMedGRID
European Union
Information and Media Technology
Share
TWGrid