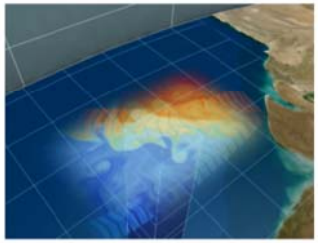
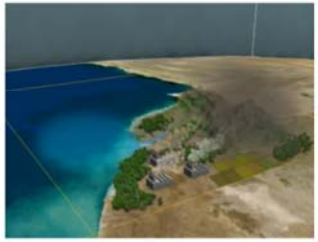
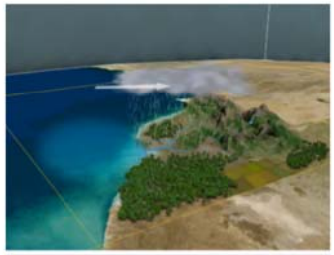
Climate model worflows at the IPSL Climate Modeling Center

Sébastien Denvil, Arnaud Caubel, Nicolas Carenton, Anne Cozic, Jean-Louis Dufresne, Marie-Alice Foujols, Josefine Ghattas, Mark Greenslade, Sonia Labetoulle, Jérôme Raciazek et al.

Laboratories and funders

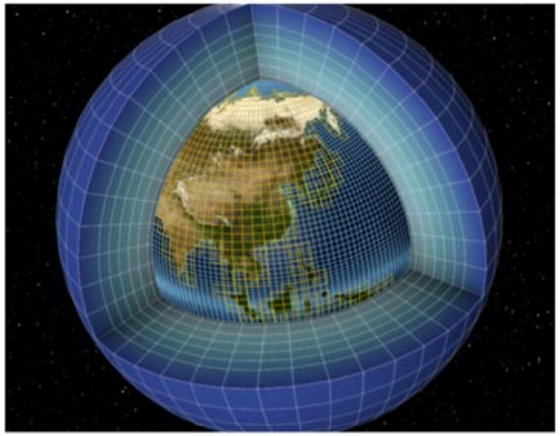
IPSL Earth System Model

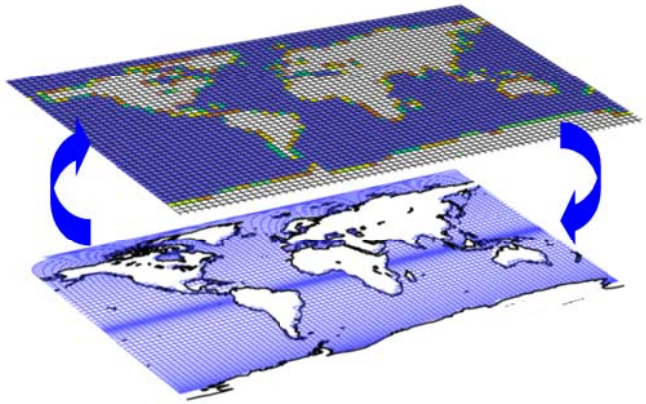
INCA / REPROBUS
(chimie atmosphérique)
(aérosol)

LMDZ
(atmosphère)

ORCHIDEE
(surfaces continentales)
(végétation)



OASIS
(coupleur)



OPA
(océan)

LIM
(glace de mer)

PISCES
(biogéochimie marine)

NEMO

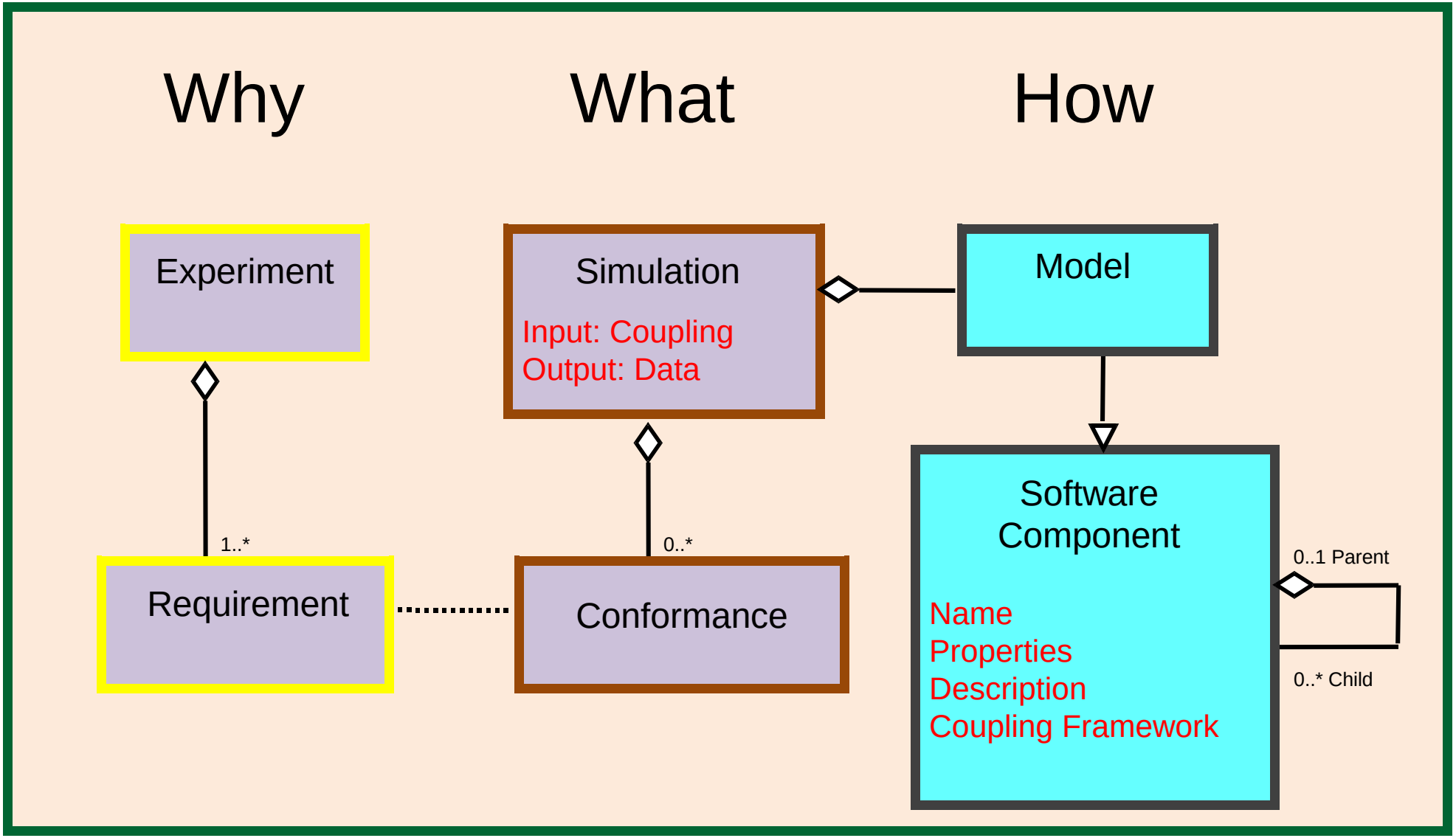
Institut Pierre Simon Laplace
Sciences de l'environnement

IPCC AR5 variable counts

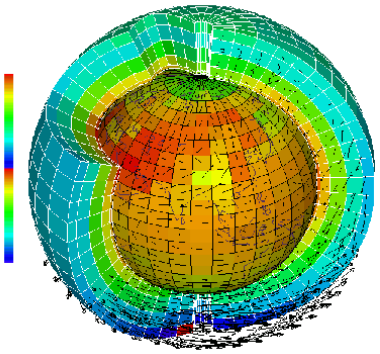
	1 hour	3 hour	6 hour	daily	month	annual	**totals**
aerosol	0	0	0	0	81	0	**81**
atmosphere	75	101	9	86	184	0	**455**
land	0	3	0	2	59	0	**64**
land ice	0	0	0	2	13	0	**15**
ocean	0	1	0	3	116	0	**120**
biogeochemistry	0	0	0	0	88	71	**159**
sea ice	0	0	0	4	47	0	**51**
totals	**75**	**105**	**9**	**97**	**588**	**71**	**945**

A climate simulation

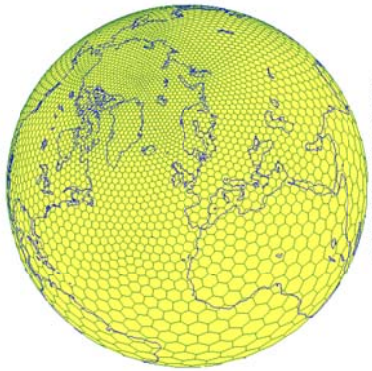
http://earthsystemcog.org/projects/es-doc-models/



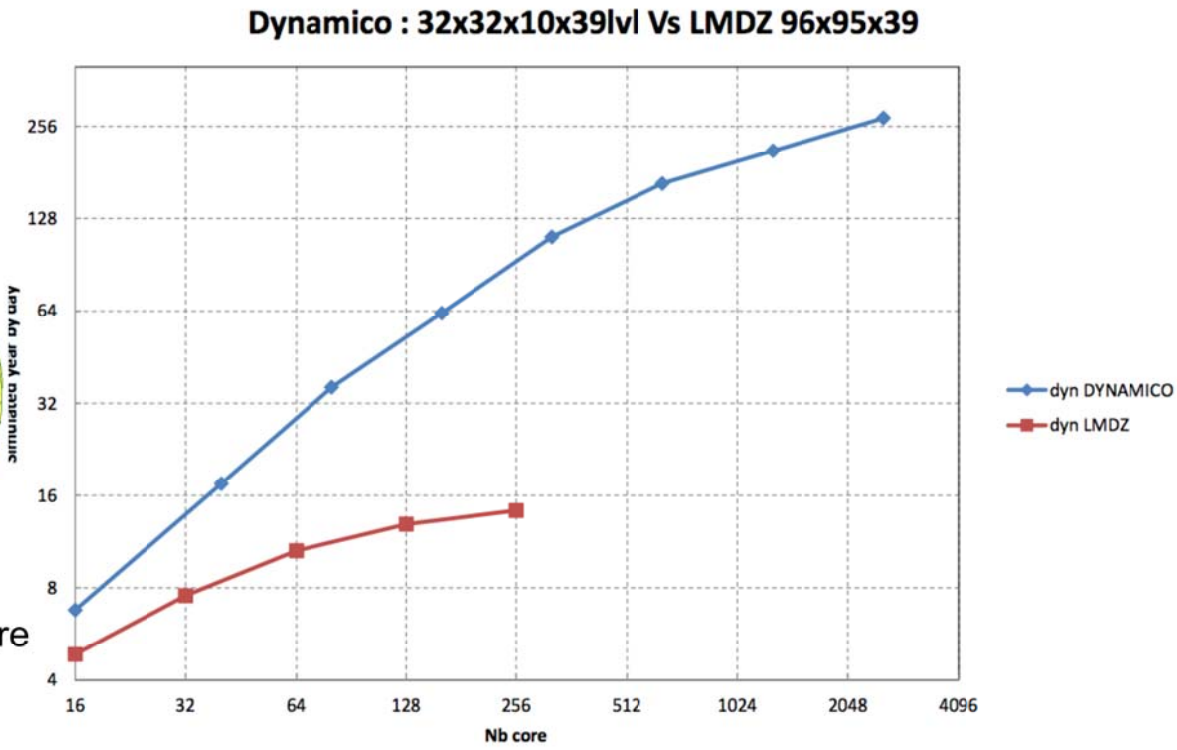
Next generation model performance

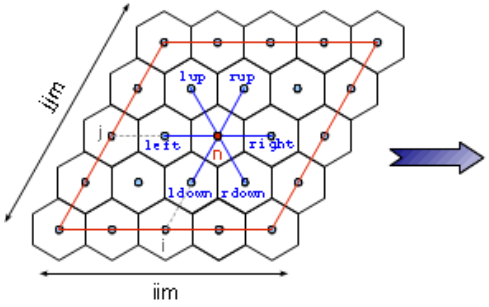


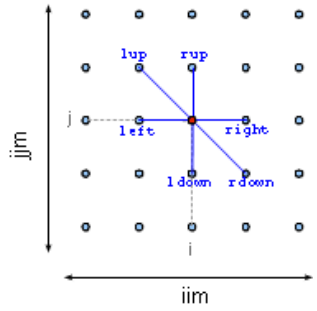
LMD-Z lon-lat core



DYNAMICO icosahedral core







degrés	nb cœurs	année/j	Mh/siècle	
3	320	**110**	0,0077	*Measured*
1	1280	**20**	0,15	
1	5120	**55**	0,22	
½	5120	*10*	1,2	*Extrapolated*
½	11520	*18*	1,5	
½	20480	*28*	1,8	
¼	20480	*5*	10	
¼	46080	*8*	14	

Courtesy Thomas Dubos LMD/IPSL and Yann Meurdesoif LSCE/IPSL

Institut Pierre Simon Laplace
Sciences de l'environnement

To keep in mind

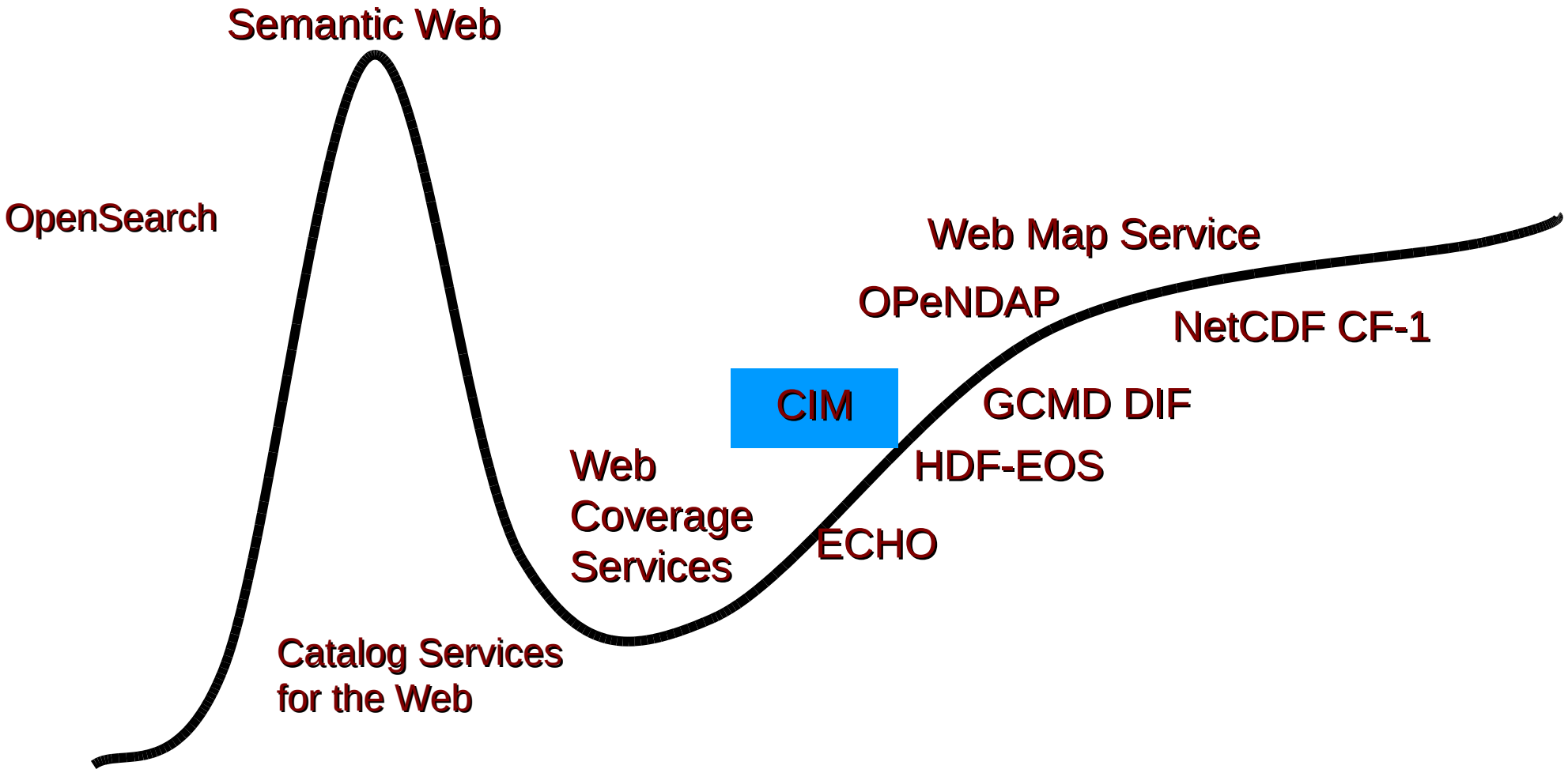
"the potential to interpret, compare and reuse climate information results is strongly related to the quality of their description"

But metadata alone won't get us there !

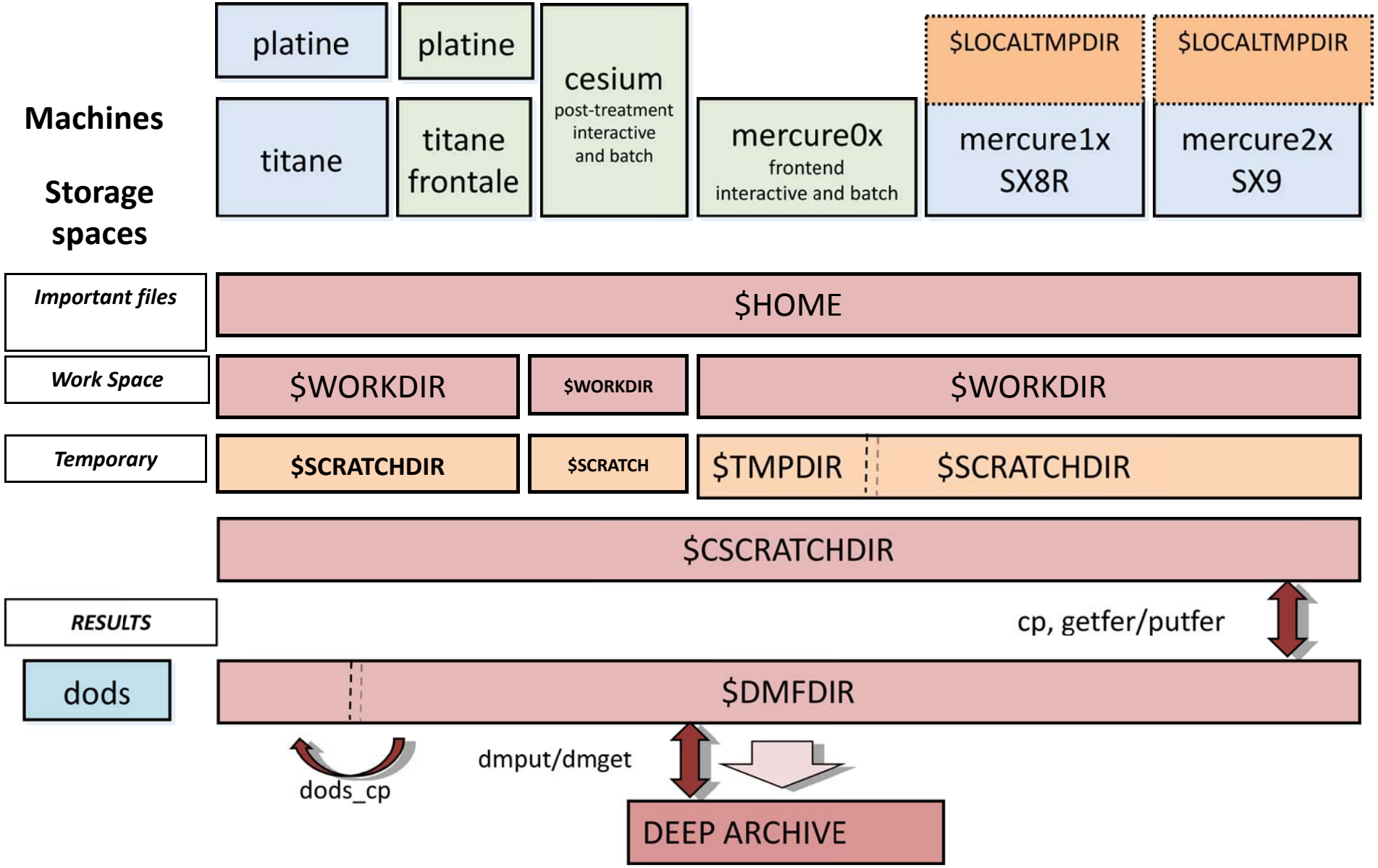
Computation useless if results cannot be stored/distributed/read

Institut Pierre Simon Laplace

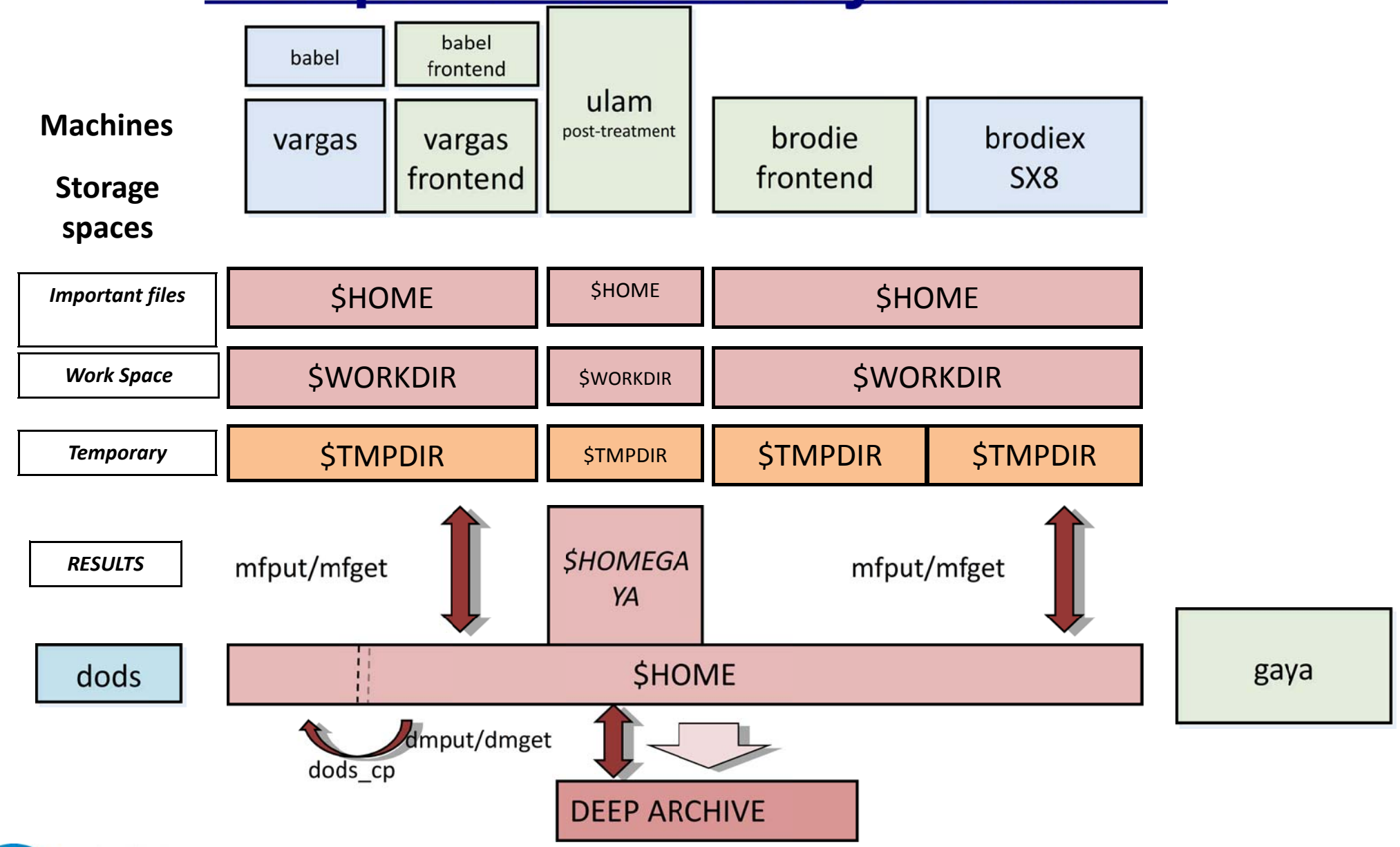
Hype .vs. Reality
Interoperability in Earth Sciences



Compute/Store/Analysis TGCC

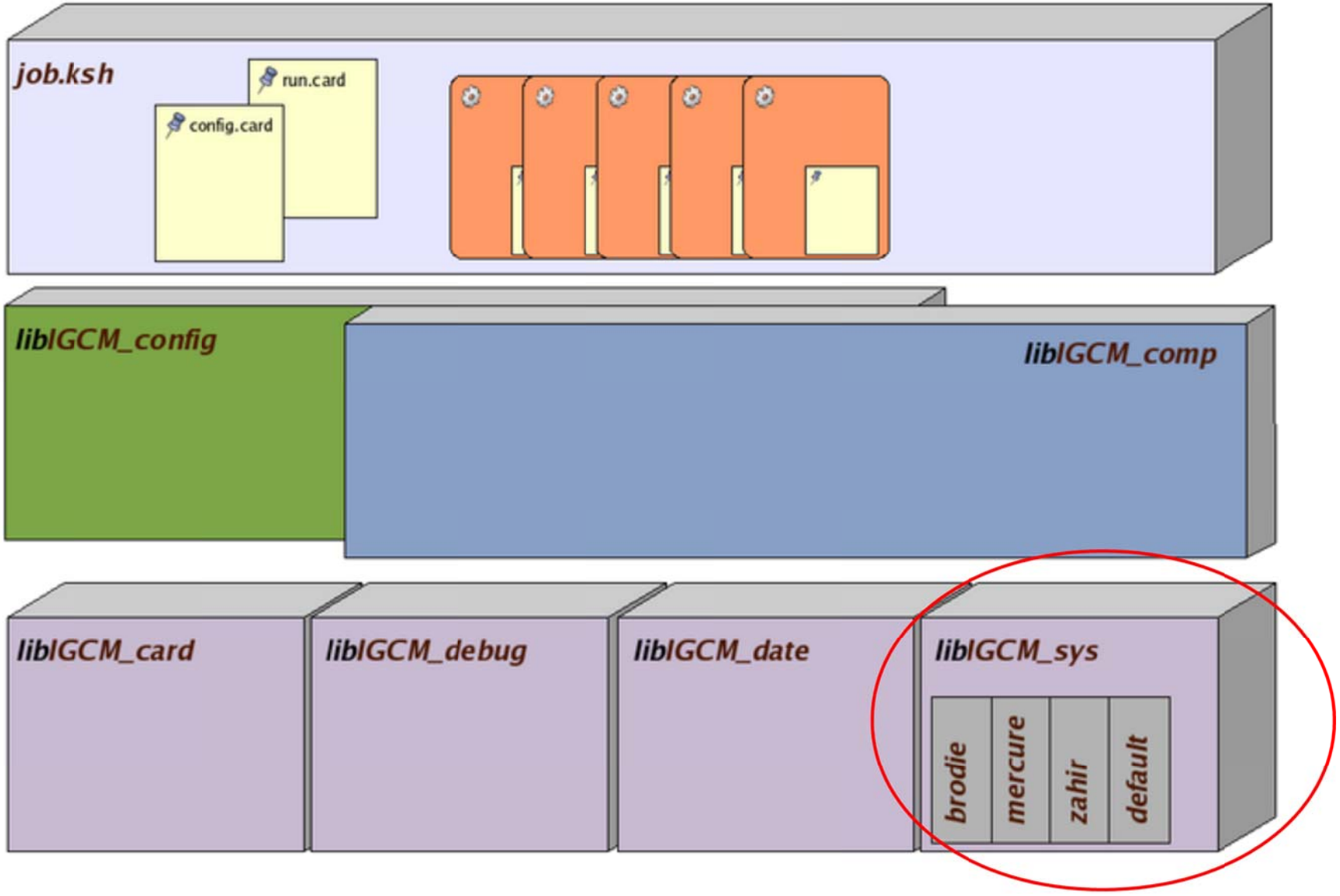


Compute/Store/Analysis IDRIS

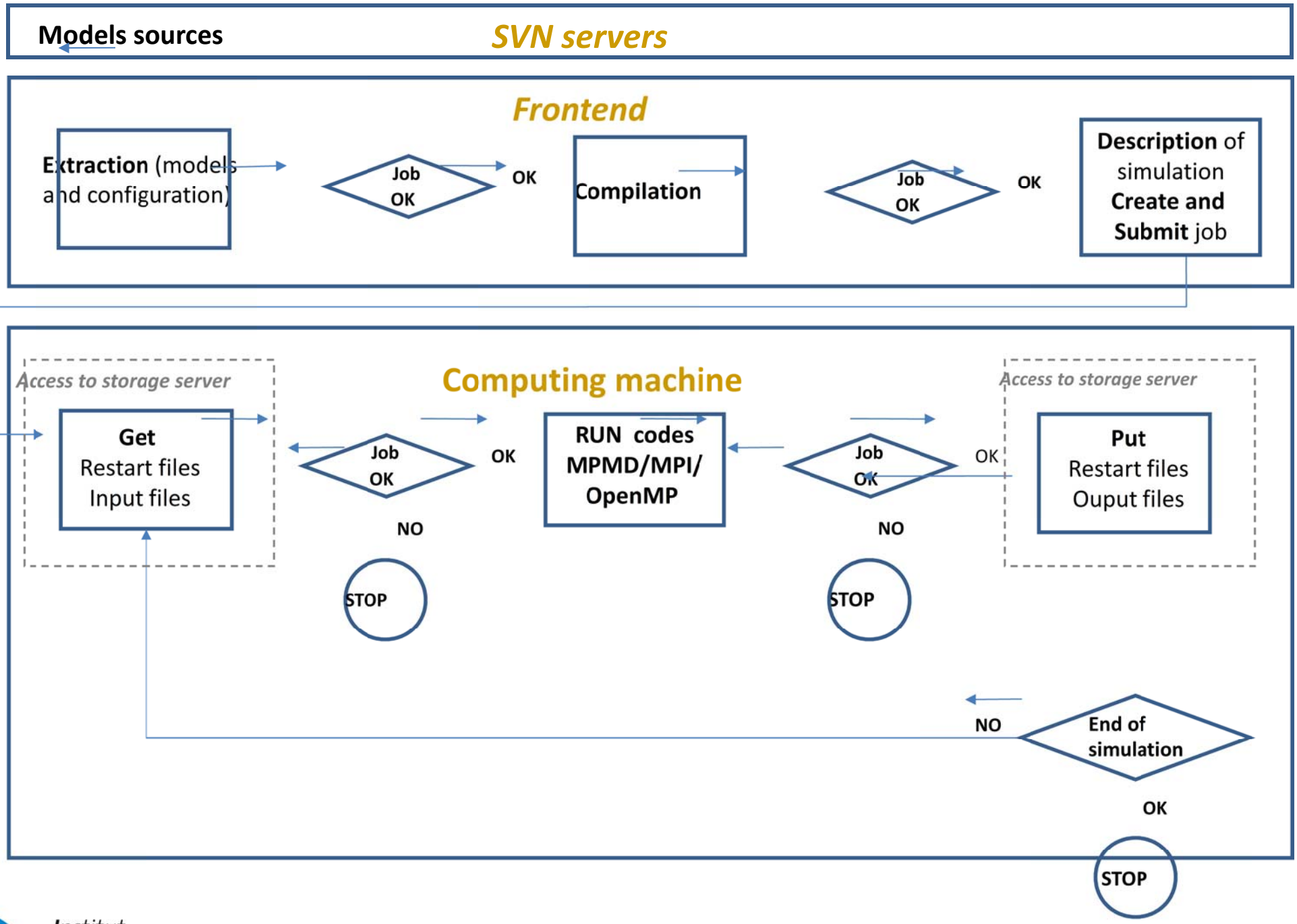


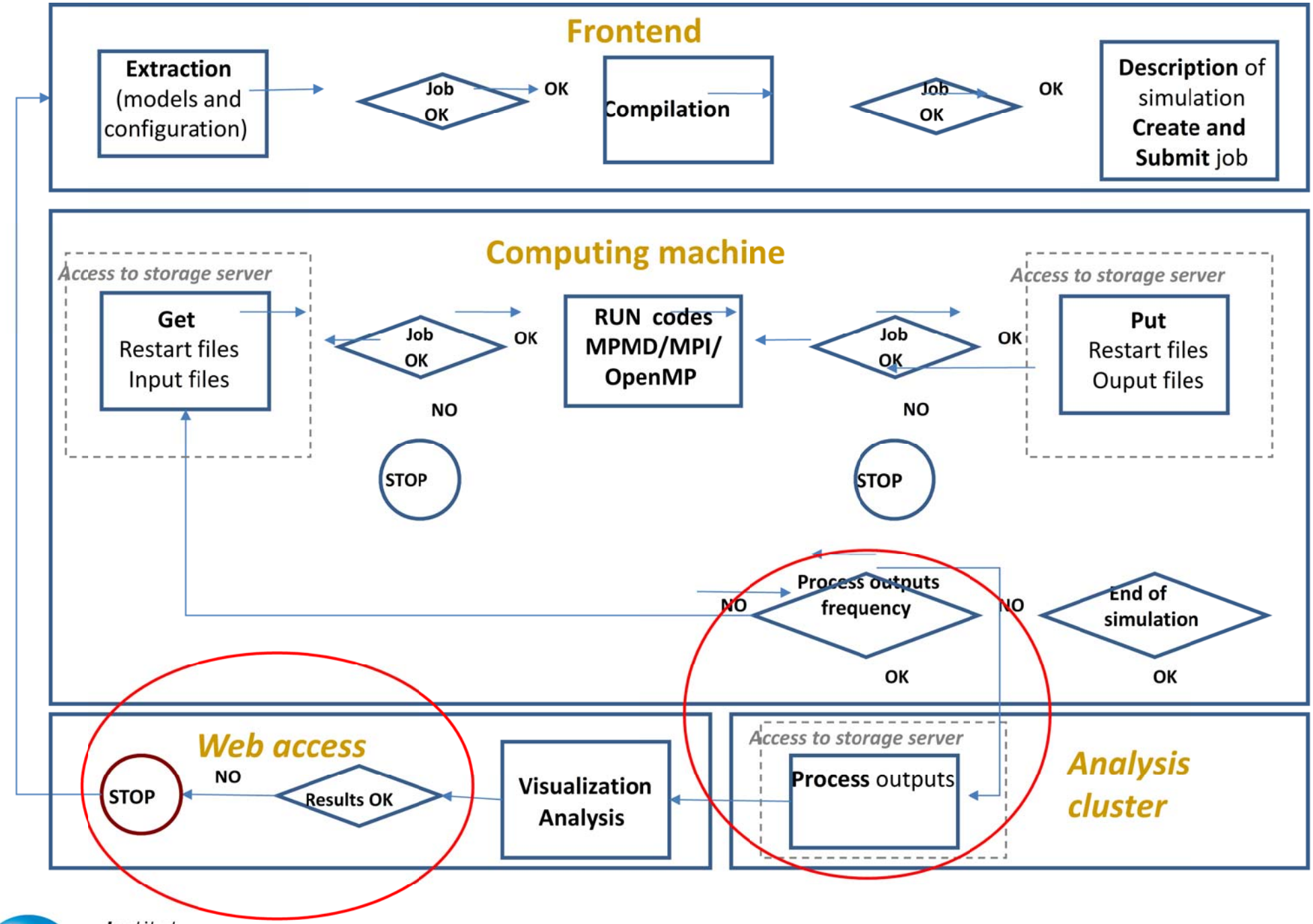
Institut Pierre Simon Laplace

Simulation Control Environment



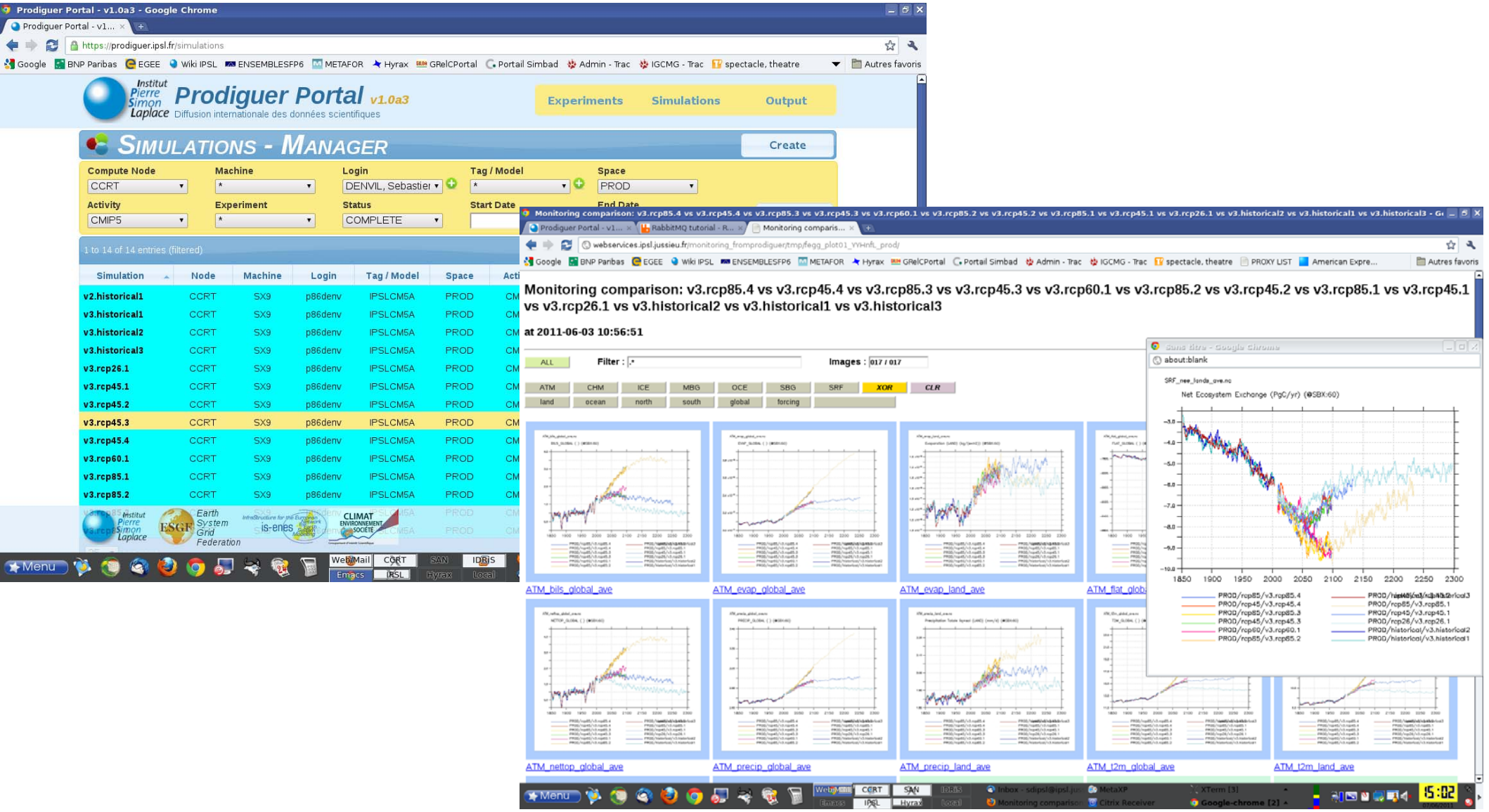
Institut Pierre Simon Laplace


Models sources
SVN servers
Frontend
Extraction (models and configuration)
Job OK
OK
Compilation
Job OK
OK
Description of simulation Create and Submit job
Computing machine
Access to storage server
Get
Restart files
Input files
Job OK
OK
NO
STOP
RUN codes MPMD/MPI/ OpenMP
Job OK
OK
NO
STOP
Access to storage server
Put
Restart files
Ouput files
NO
End of simulation
OK
STOP



Institut Pierre Simon Laplace

Dashboard and data access



Trusting web service

Trusting

Release: 0.60

Check reliability of configurations on the SX9 from the CCRT.

Ask for support
Discover other services

Configuration IPSLCM5_v2

2010-06-07T06:10

IOIPSL/src	`svn checkout -r HEAD http://forge.ipsl.jussieu.fr/igcmg/svn/IOIPSL/tags/v2_1_9/src IOIPSL/src`
ORCHIDEE	`cvs -d :pserver:sechiba@cvs.ipsl.jussieu.fr:/home/ssipsl/CVSREP checkout -r orchidee_1_9_4_2 ORCHIDEE`
OASIS3	`cvs -d :pserver:anonymous@cvs.ipsl.jussieu.fr:/home/ioipsl/CVSROOT checkout -r ipslcm5a -d prism OASIS3`
LMDZ4	`svn checkout -r 1329 http://svn.lmd.jussieu.fr/LMDZ/LMDZ4/trunk LMDZ4`
IPSLCM5	`svn checkout -r HEAD http://forge.ipsl.jussieu.fr/igcmg/svn/CONFIG/IPSLCM/IPSLCM5/branches/IPSLCM5_v2 IPSLCM5`
libIGCM	`svn checkout -r 246 http://forge.ipsl.jussieu.fr/libigcm/svn/trunk/libIGCM libIGCM`
NEMO	`svn checkout -r HEAD http://forge.ipsl.jussieu.fr/nemo/svn/tags/nemo_v3_2/NEMO`
UTIL	`svn checkout -r HEAD http://forge.ipsl.jussieu.fr/nemo/svn/tags/nemo_v3_2/UTIL`
XMLF90	`svn checkout -r 54 http://forge.ipsl.jussieu.fr/ioserver/svn/XMLF90`
XMLIO_SERVER	`svn checkout -r 54 http://forge.ipsl.jussieu.fr/ioserver/svn/XMLIO_SERVER/trunk XMLIO_SERVER`

Trusting log

Date	Status	Step	Comments	C++	F90	MPI	CROSSKIT	NETCDF	IOIPSL/src	ORCHIDEE	OASIS3	LMDZ4	IPSLCM5	libIGCM	NEMO	UTIL
2010-06-07T06:10	OK	Code is reliable	-	087	400	8.0.4	18.1/3	3.6.1	HEAD 740	orchidee_1_9_4_2 2010-05-27T10:30:35	ipslcm5a 2010-03-18T15:16:18	1329	HEAD 951	246	HEAD 1779	HEAD 1773
2010-06-06T06:10	OK	Code is reliable	-	087	400	8.0.4	18.1/3	3.6.1	HEAD 740	orchidee_1_9_4_2 2010-05-27T10:30:35	ipslcm5a 2010-03-18T15:16:18	1329	HEAD 951	246	HEAD 1779	HEAD 1773
2010-06-05T06:10	OK	Code is reliable	-	087	400	8.0.4	18.1/3	3.6.1	HEAD 740	orchidee_1_9_4_2 2010-05-27T10:30:35	ipslcm5a 2010-03-18T15:16:18	1329	HEAD 951	246	HEAD 1779	HEAD 1773
2010-06-04T06:10	OK	Code is reliable	-	087	400	8.0.4	18.1/3	3.6.1	HEAD 740	orchidee_1_9_4_2 2010-05-27T10:30:35	ipslcm5a 2010-03-18T15:16:18	1329	HEAD 951	246	HEAD 1779	HEAD 1773

Sciences de l'environnement
Institut Pierre Simon Laplace

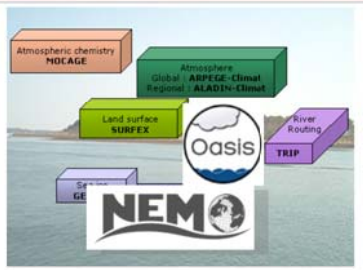
ANR MN2013 CONVERGENCE

T0 : management

T1 : platform

ensemble of tools
different configurations
different resolution
set of simulations
set of diagnostics
assessment

IPSL implementation



GAME-CERFACS implementation

T2 : towards a high-resolution coupled model

- Improving coupled model parallelism in terms of computing and memory
- Managing efficiently input and restart files
- Integrating parallel interpolation mechanisms in XIOS
- Parallel component coupling

T3 : runtime environments

- Process assignment
- Optimization, Load balancing
- Climate Simulations Supervision

T4 : Big Data management and analytics

- XIOS implemented within project models
- XIOS a bridge towards standardisation
- Data and metadata services
- Big Data Analytics

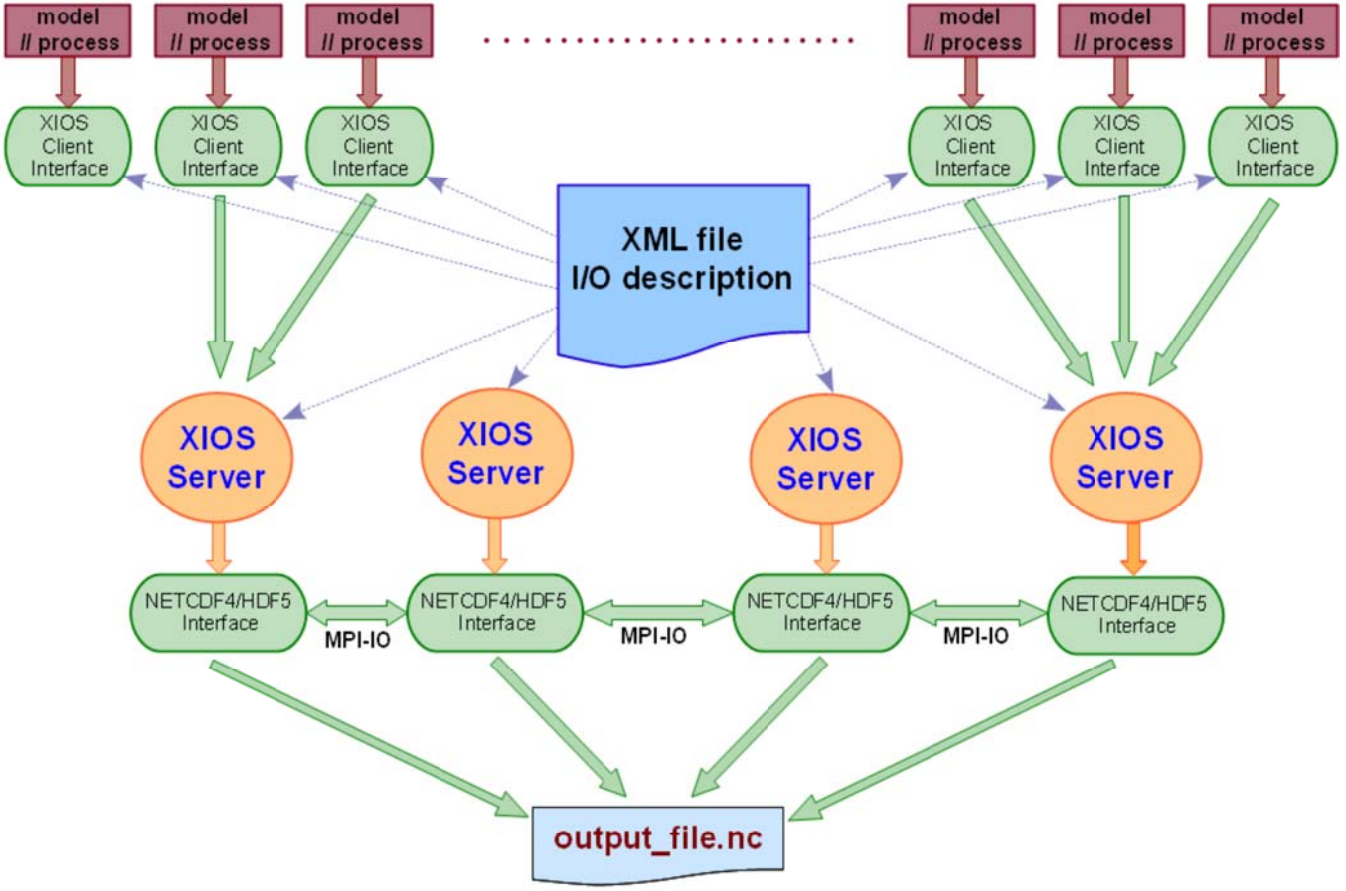
T5 : CliMAF: a framework for climate models evaluation

- General driver and upstream user interface
- Services layer
- Visualization tools
- Evaluation and monitoring diagnostics

Task 1 : XIOS included in IPSL-CM and CNRM-CM components

Objective : Having XIOS as our primary software to generate our data. Having a common piece of software to achieve this important task will have a lot of benefits on the long run (synergy speaking).

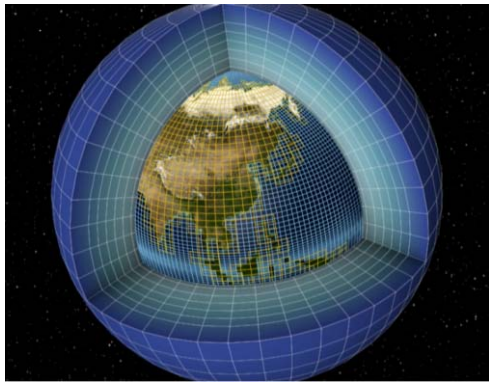
- IPSL-CM : LMDz, NEMO, ORCHIDEE, INCA, REPROBUS
- CNRM-CM : ARPEGE, NEMO, GELATO, SURFEX
-



Institut Pierre Simon Laplace

Task 2 : XIOS a bridge towards standardization

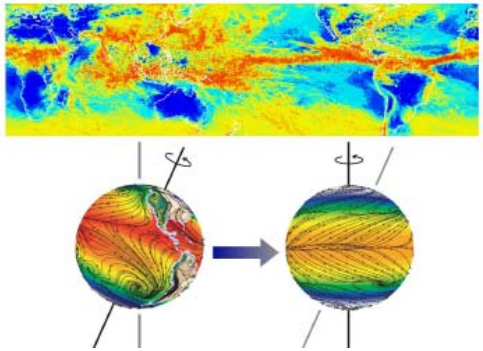
Objective : Output format, structure and description will (1.) conform to the Climate and Forecast (CF) convention (2.) conform to CMIP controlled vocabulary and (3.) conform to CIM ontology describing simulation and model documentation.



Write « standard » results

Ingest in data & metadata services

Analyse

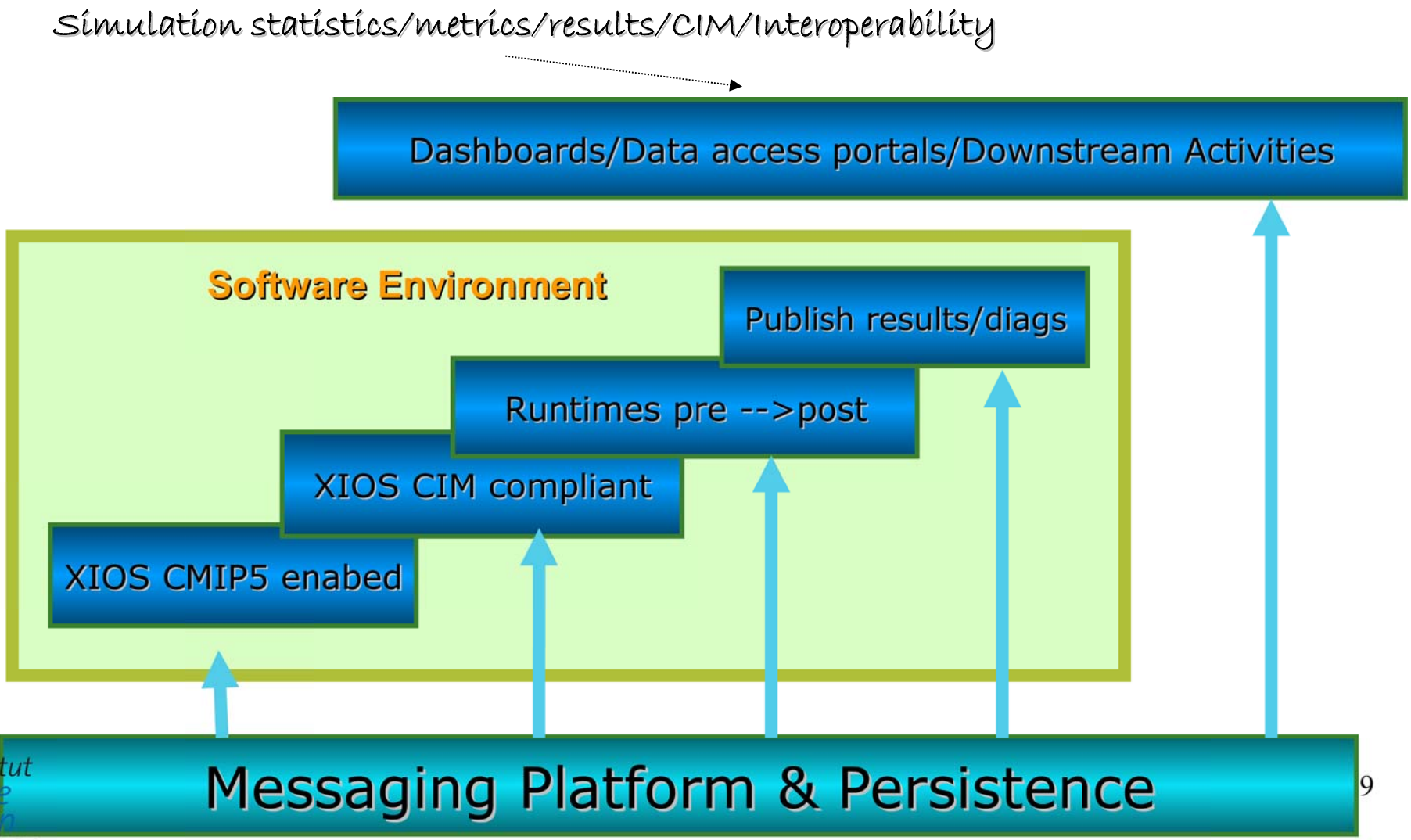


Necessary condition to enable easier and faster systematic ingestion of outputs by data services developed in Task 4.3, and ensure a high level of documentation, provenance, standardization and reuse.

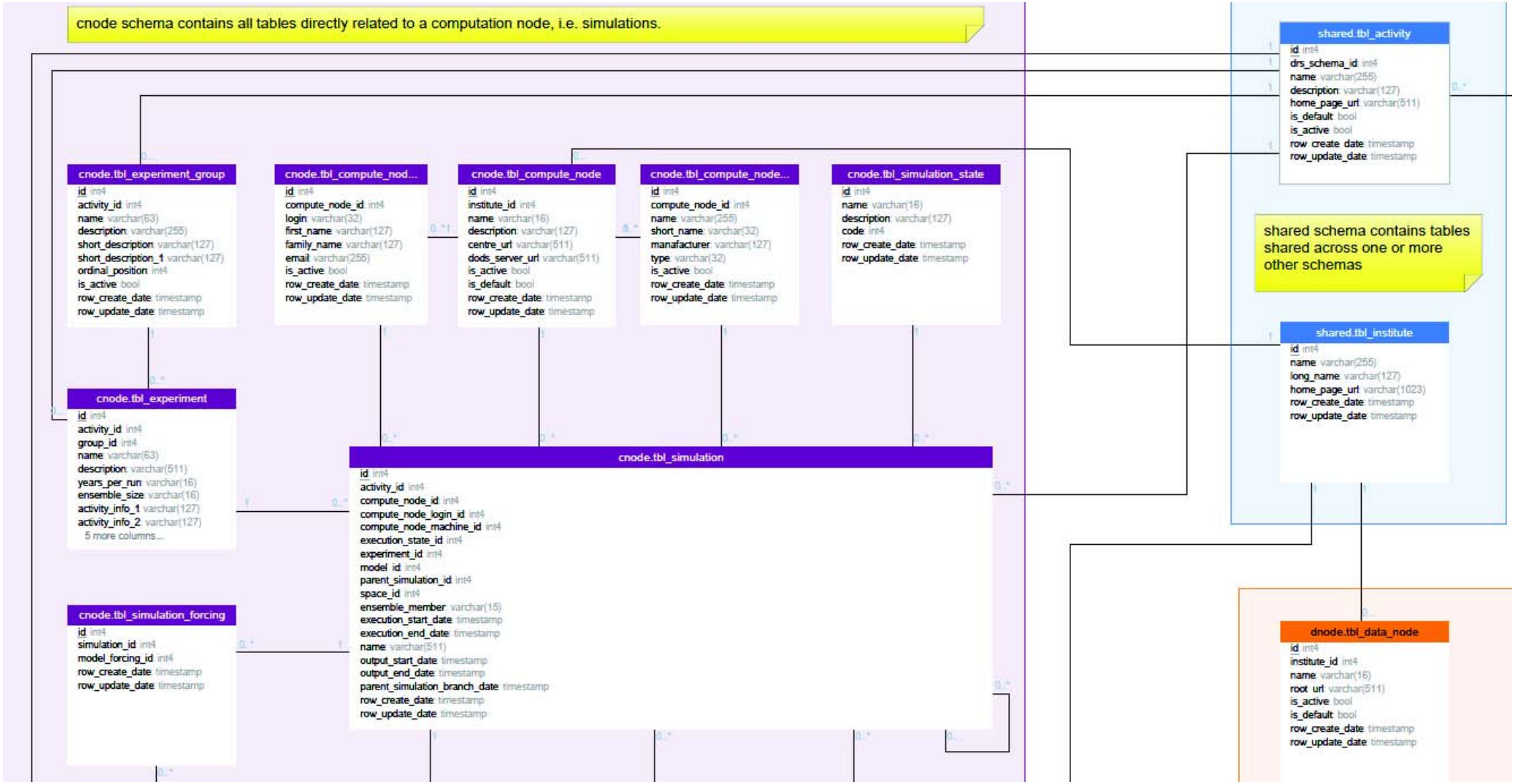
Institut Pierre Simon Laplace

Task 3 : Data & Metadata Services

Objective : This environment will be our communication bus, accessible through RESTful interfaces and will be an integrative middle layer on which various data services will be built.



Behind the scene



OBJECTIVE

Real time simulation monitoring across all computation centres

Sciences de l'environnement
Institut Pierre Simon Laplace

Web Service API: Message Queues

Rabbit MQ

http://www.rabbitmq.com/

Durable Message Queues
AMQP : Advanced Message Queue Protocol

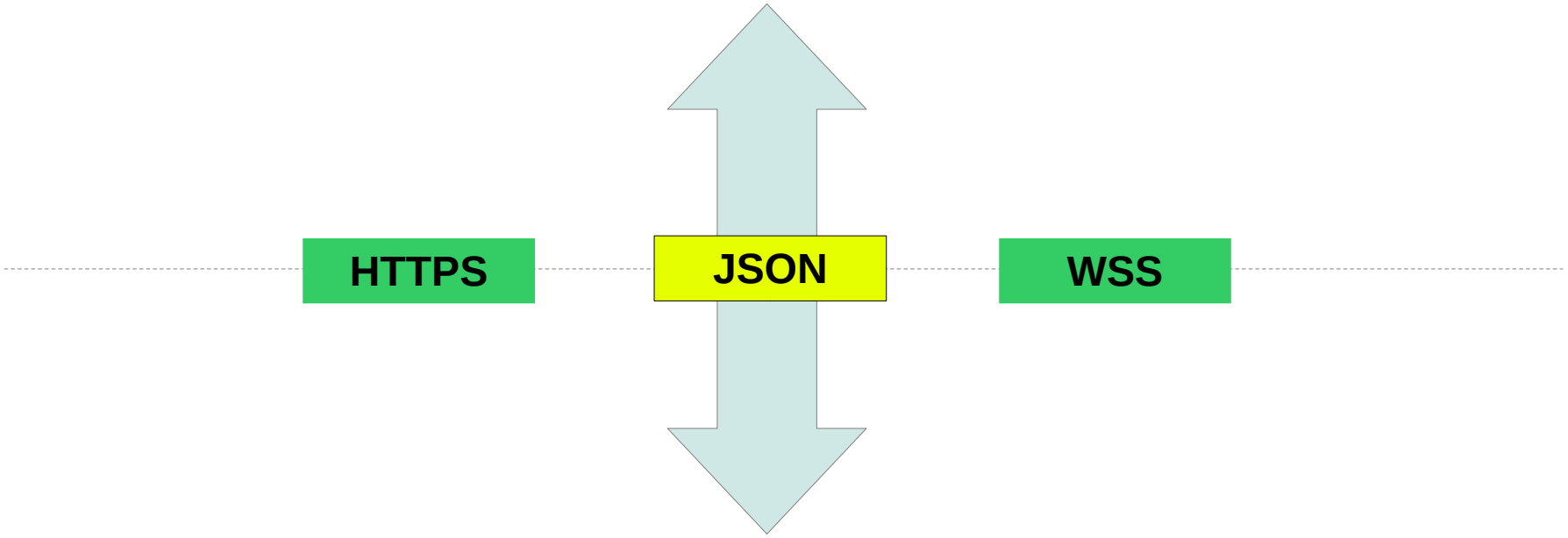
open source message broker

robust
Powerful
surprisingly simple to use

Institut Pierre Simon Laplace
Sciences de l'environnement

HOW

Web Service API

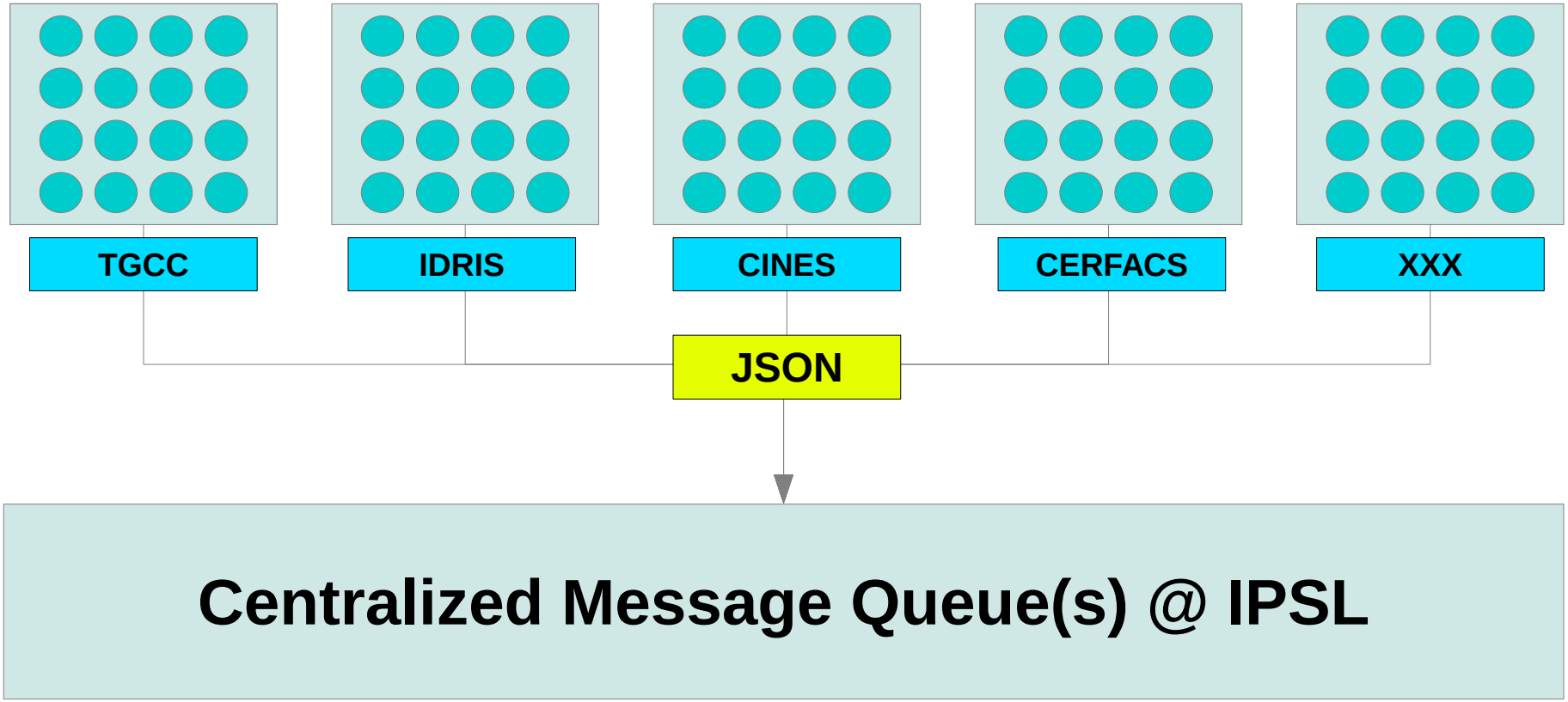


Single Page Javascript App

Sciences de l'environnement
Institut Pierre Simon Laplace

Web Service API

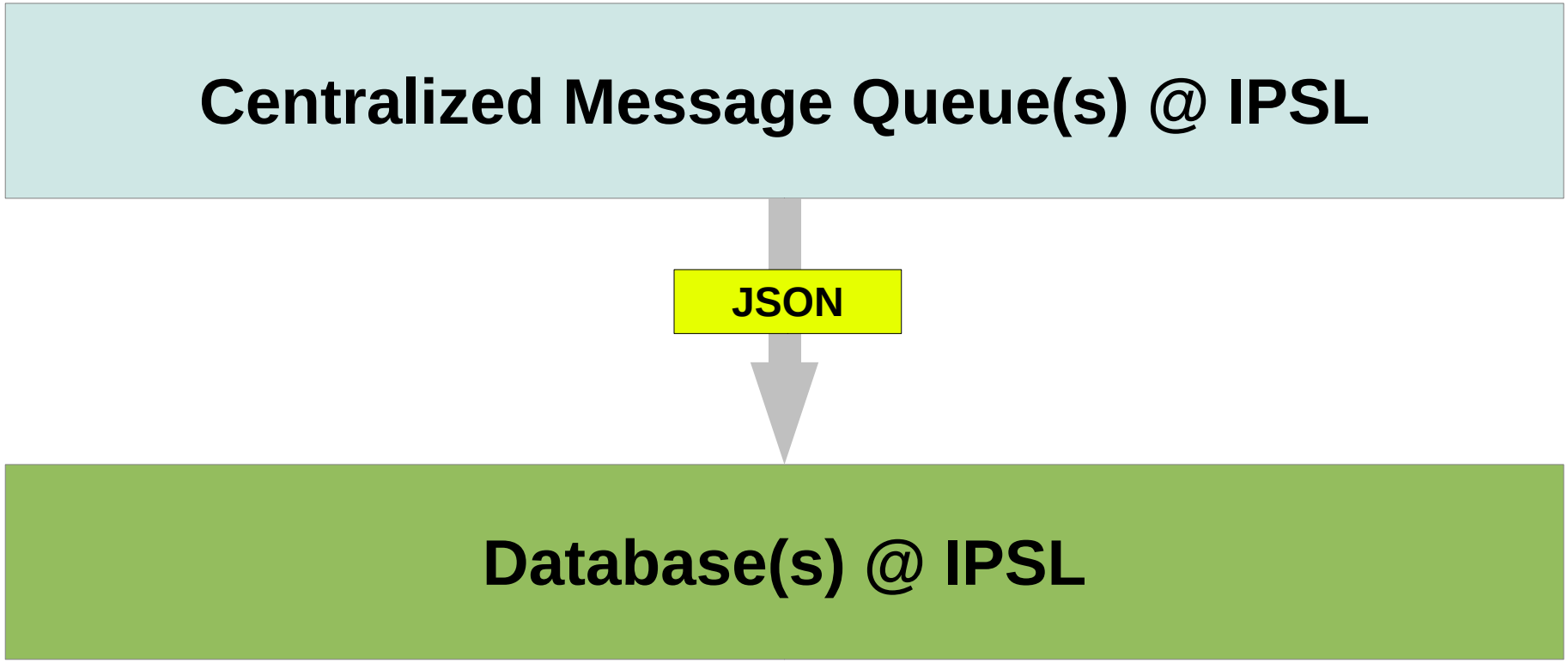
Step 1: Centralize simulation messages



Sciences de l'environnement
Institut Pierre Simon Laplace

Web Service API

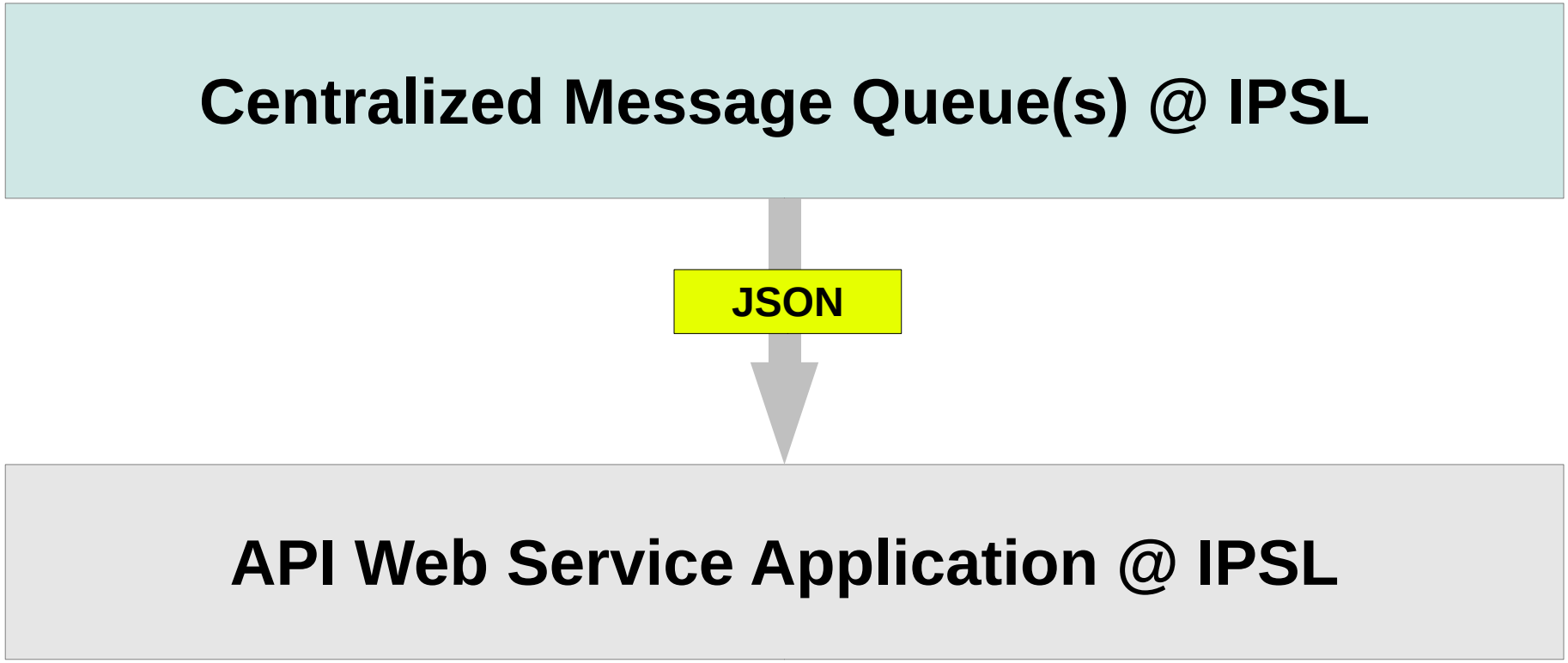
Step 2: Persist simulation messages



Sciences de l'environnement
Institut Pierre Simon Laplace

Web Service API

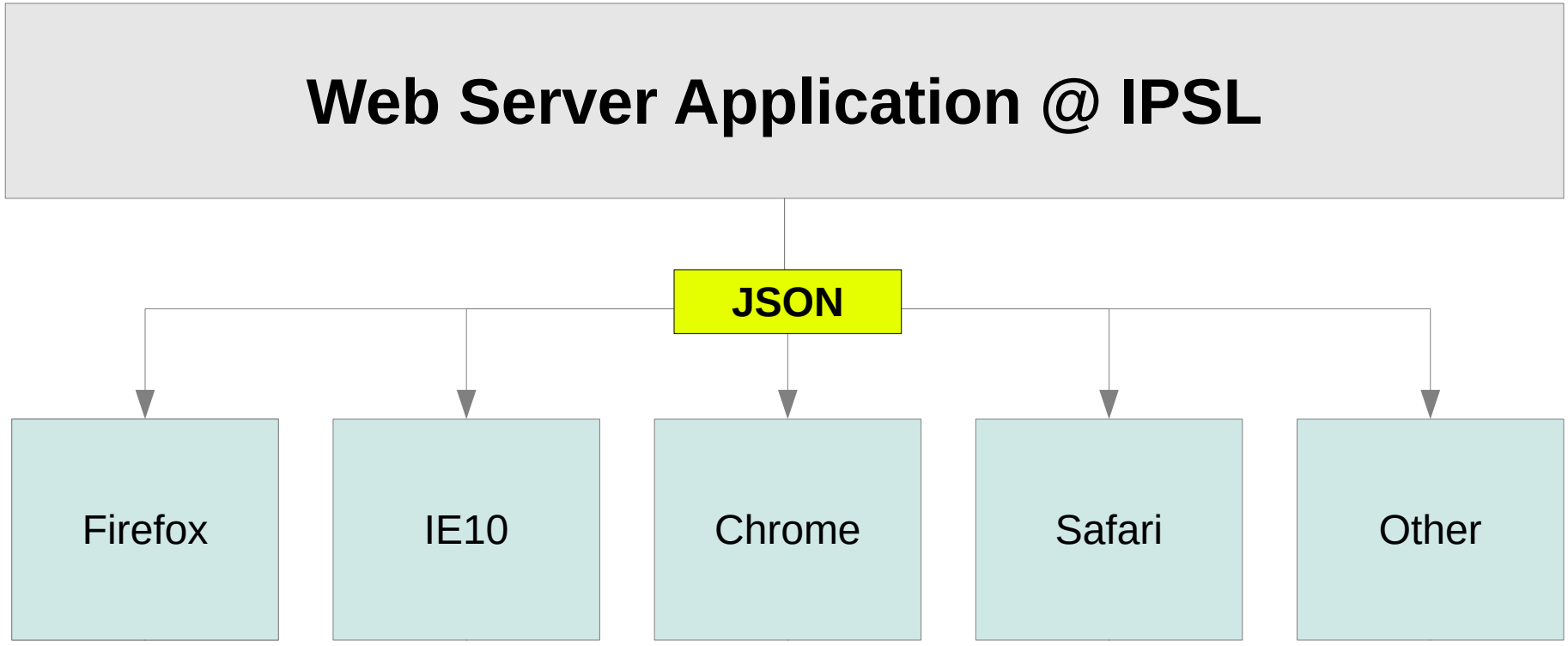
Step 3: Notify API web service application



Sciences de l'environnement
Institut Pierre Simon Laplace

Web Service API

Step 4: Broadcast to connected web browsers

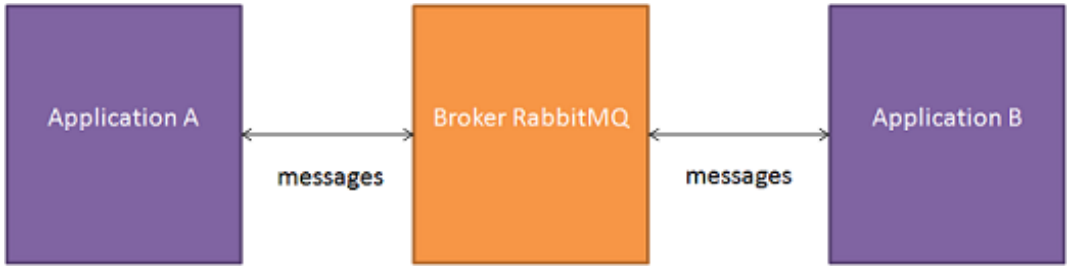


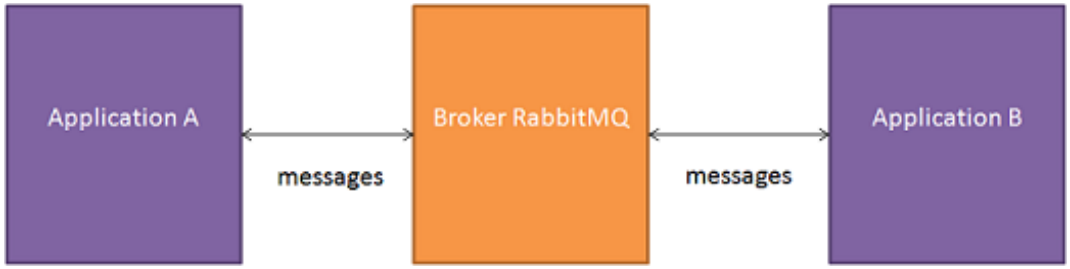
Sciences de l'environnement
Institut Pierre Simon Laplace

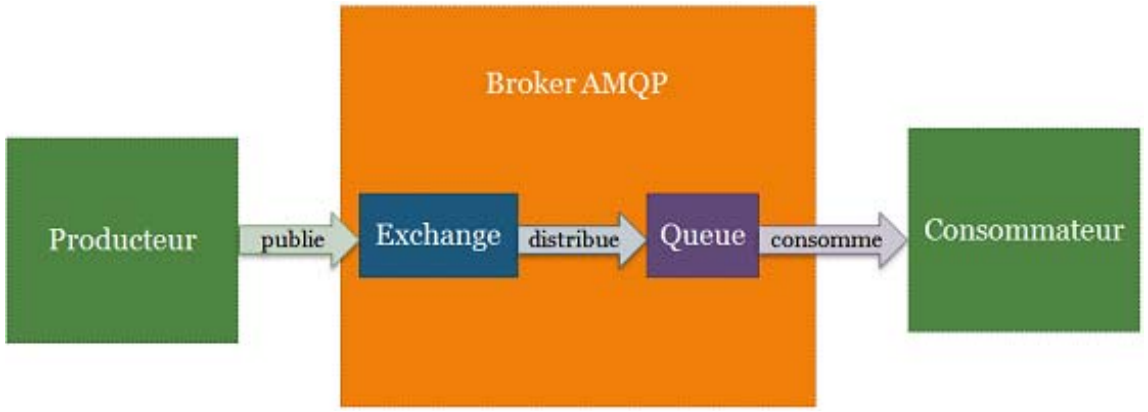
Web Service API: System Constraints

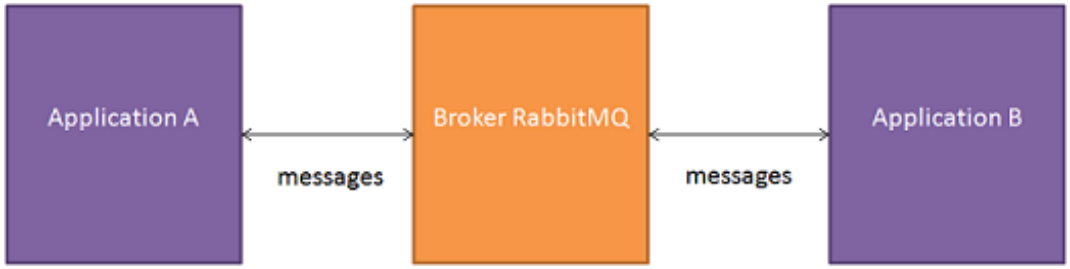
Secure
Fault Tolerant
Distributed
Reliable
Extensible
Loosely Coupled
Lossless
Real-time

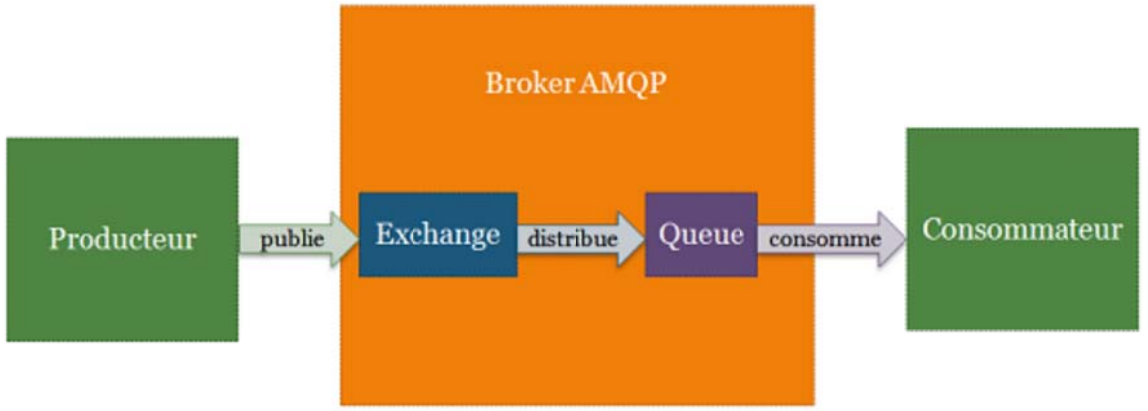
Sciences de l'environnement
Institut Pierre Simon Laplace

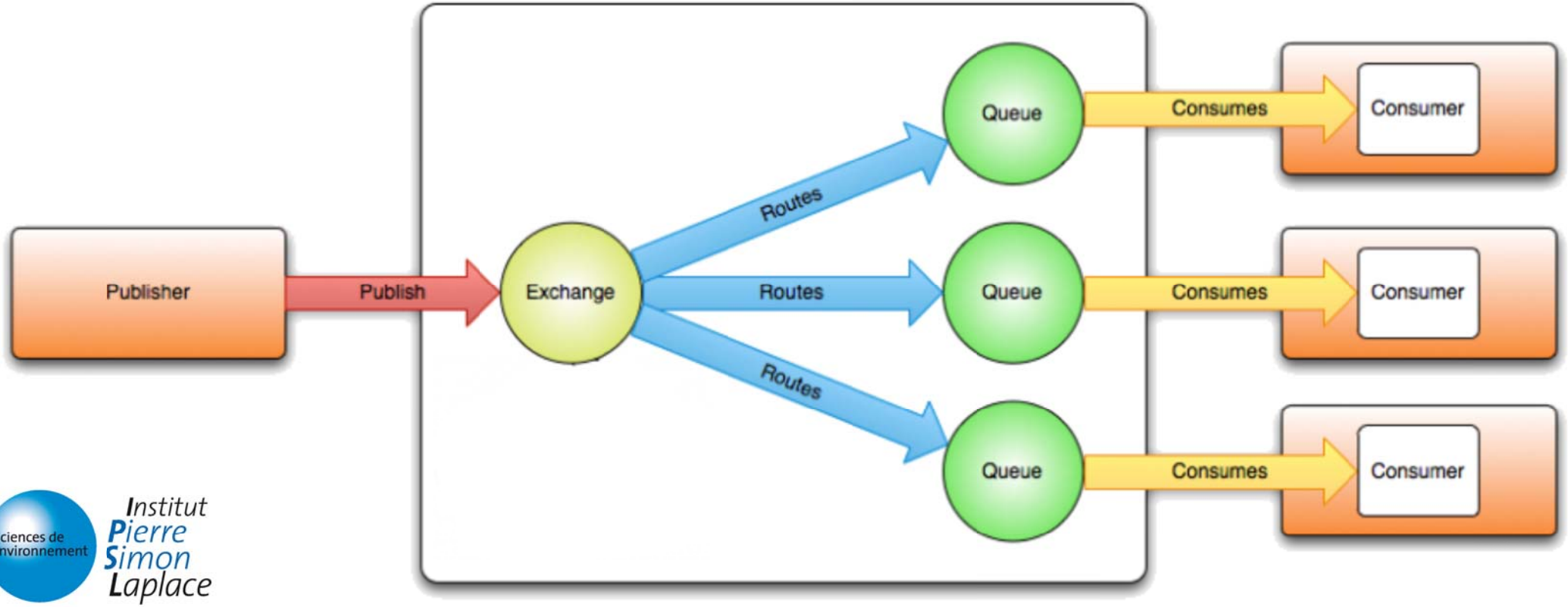

Application A
messages
Broker RabbitMQ
messages
Application B


Application A
messages
Broker RabbitMQ
messages
Application B

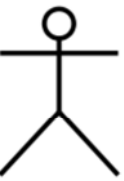

Broker AMQP
Producteur
publie
Exchange
distribue
Queue
consomme
Consommateur







Sciences de l'environnement
Institut Pierre Simon Laplace



User

RunChecker Module

In-Memory Simulation Tree

python™

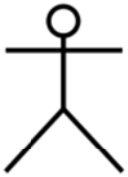
Failover Daemon

PostgreSQL

Watcher Daemon

python™

Administration Module



Administrator

Cancel, Submit

Progress, Heartbeat

ERLANG

RabbitMQ

Messaging that just works

vmware

mozilla

Mozilla Public License

Broker down

Consumer down

Manage Persistence

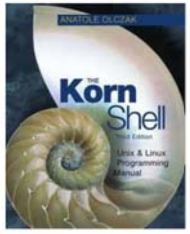
Clustering

Asynchronous send

Acknoledgement

Progress, Heartbeat

Cancel, Submit



sendMessage

C Language

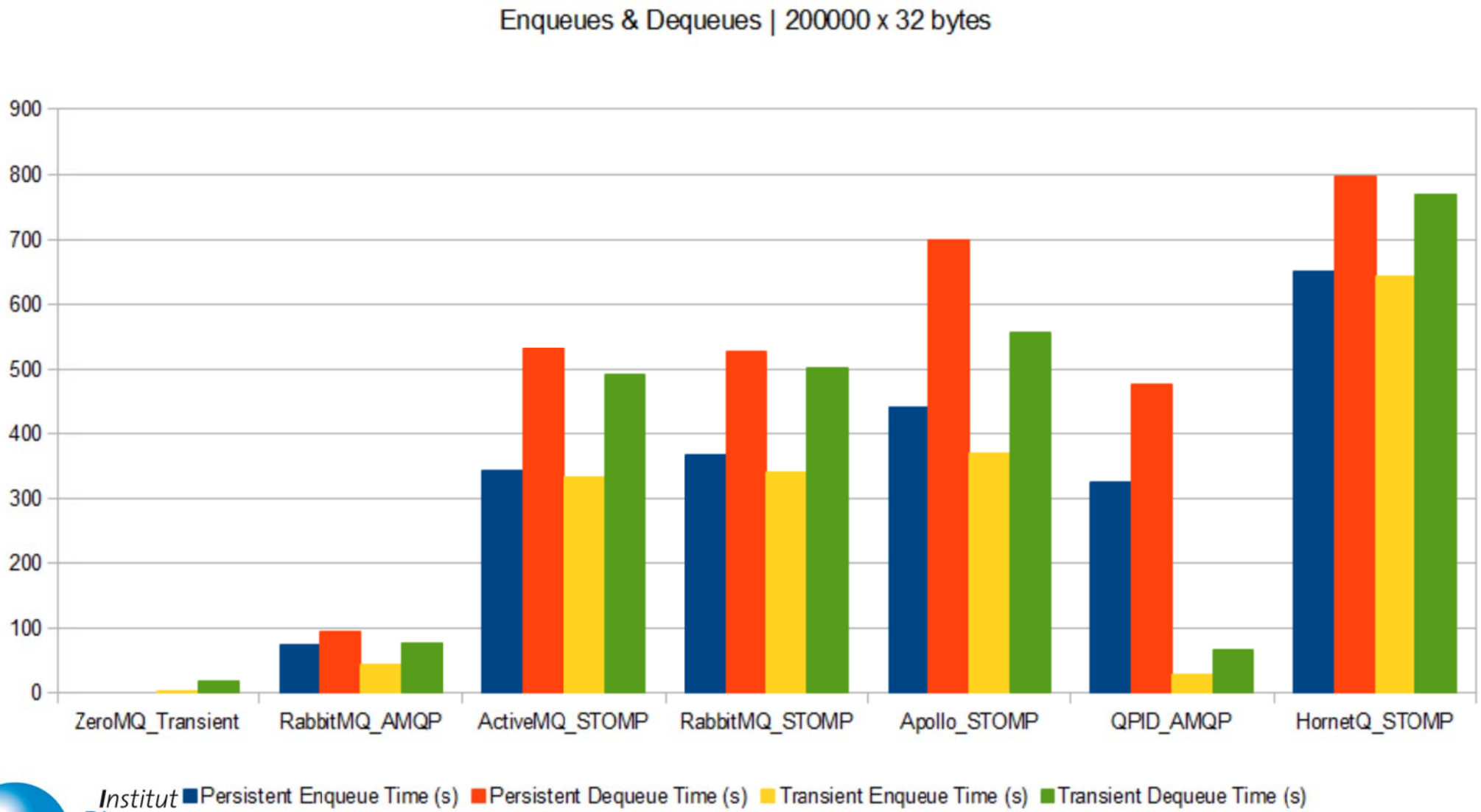
LibIGCM

receiveMessage

Sciences de l'environnement

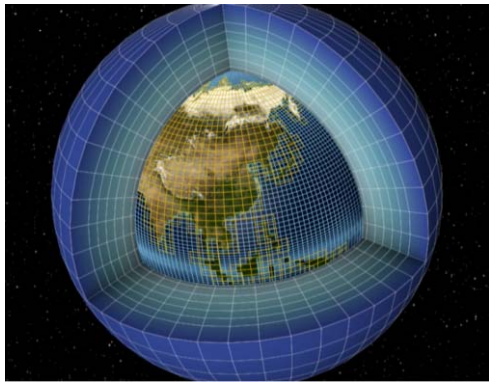
Institut Pierre Simon Laplace

Performances



Task 4 : Big Data Analytics

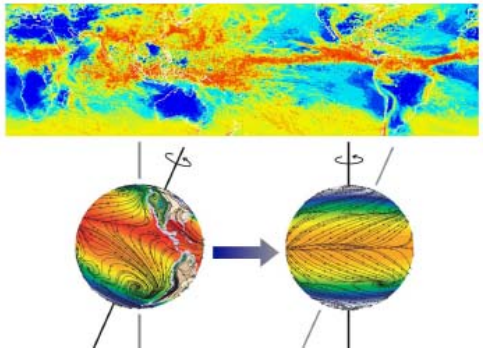
Objective : We will focus on specific development to enable server side computation capabilities at data location balanced and orchestrated by local cluster computation having high capacity link with the data location.



Write « standard » results

Ingest in data & metadata services

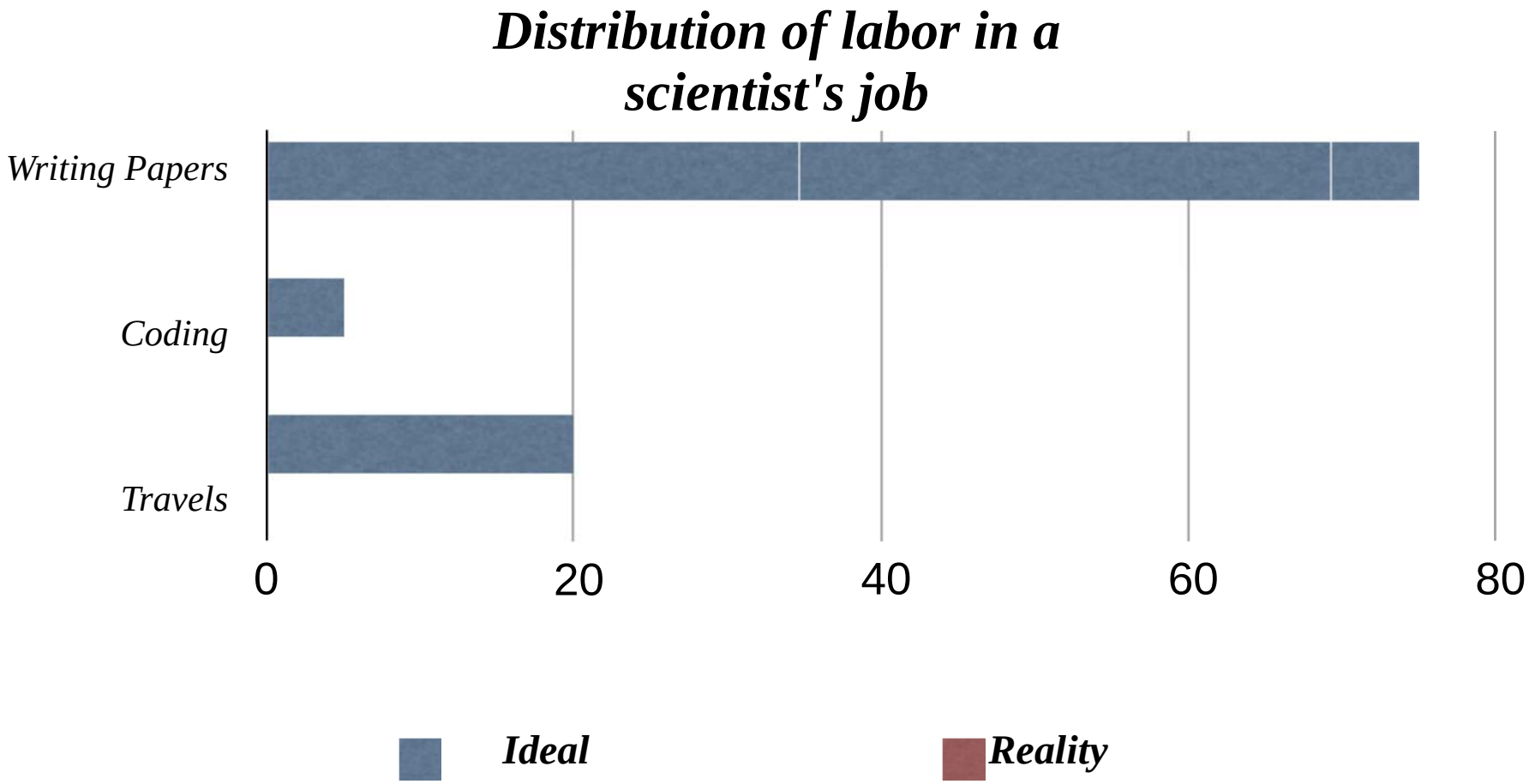
Analyse



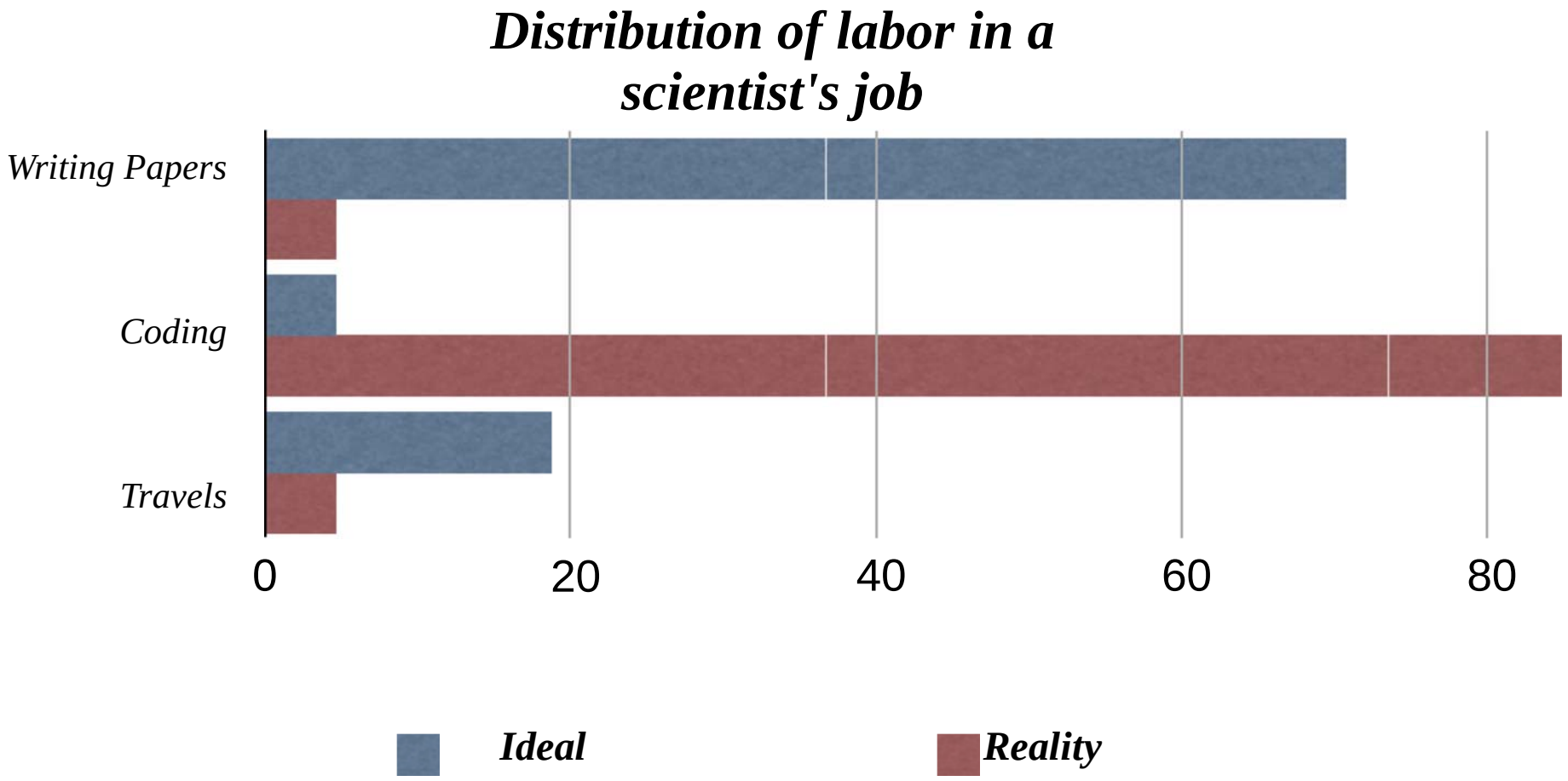
There are a wide range of technological possible choice here (WPS, Grid Engine, Cloud). **Large part** of the success here will depend on making a **wise choice**.

Institut Pierre Simon Laplace

Efficiency


Distribution of labor in a scientist's job
Writing Papers
Coding
Travels
0
20
40
60
80
Ideal
Reality

Efficiency


Distribution of labor in a scientist's job
Writing Papers
Coding
Travels
0
20
40
60
80
Ideal
Reality


Merci de votre attention