



Institut
**Pierre
Simon
Laplace**



**IPSL climate data management backed by EUDAT
services**

<http://www.ipsl.fr>
eudat@ipsl.jussieu.fr

Agenda

- IPSL presentation
- Data & calculation
- Eudat cooperation



Institut Pierre Simon Laplace

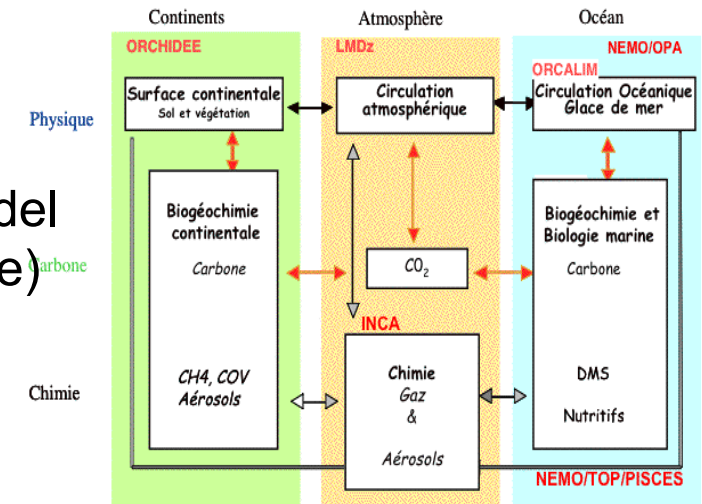
- Research federation – founded in 1991
- Gather 8 laboratories
- More than 1250 scientists and engineers

What science at the IPSL ?

- Main research topic : develop science and understanding of Earth's global environment
 - understand dynamical, chemical & biological processes in the oceans and in the atmosphere
 - understand the natural climate variability at various scales
 - understand impacts of human activities
 - themes related to solar system and other planets

What is IPSL – Institut Pierre Simon Laplace ?

- Development of generic tools :
 - models for climate modelisation
 - development of an integrated Earth system model (atmosphere, oceans, cryosphere and lithosphere)
- Observation data
 - gathering data from various instrumental tools
 - Earth satellite data (IASI – GOSAT)
- Distribute data
 - TDS - ESG Data node
 - IPSL is an IPCC Data distribution node



Just one example among many others ...

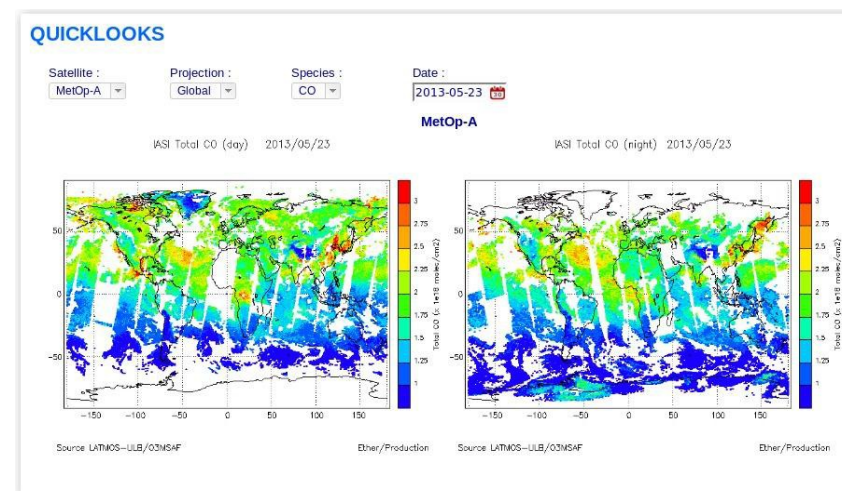
FROM ...



- IASI satellite data acquisition since 1997 from MetOp A-B
- clouds, temperature, ozone, CH₄, CO
- 30 TB/y

TO ...

- multicriteria quick access tool
- daily maps
- goals : weather forecast, climate change, pollution prediction



So what do we need ?

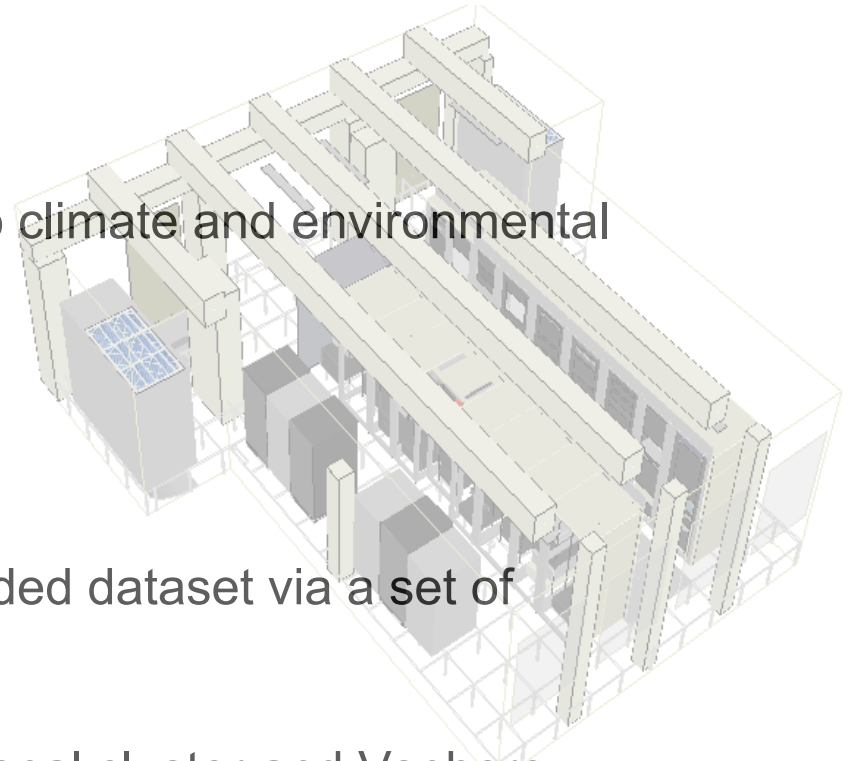
- Data storage
- HPC (close to data and users)
- Data distribution

With what tool ?

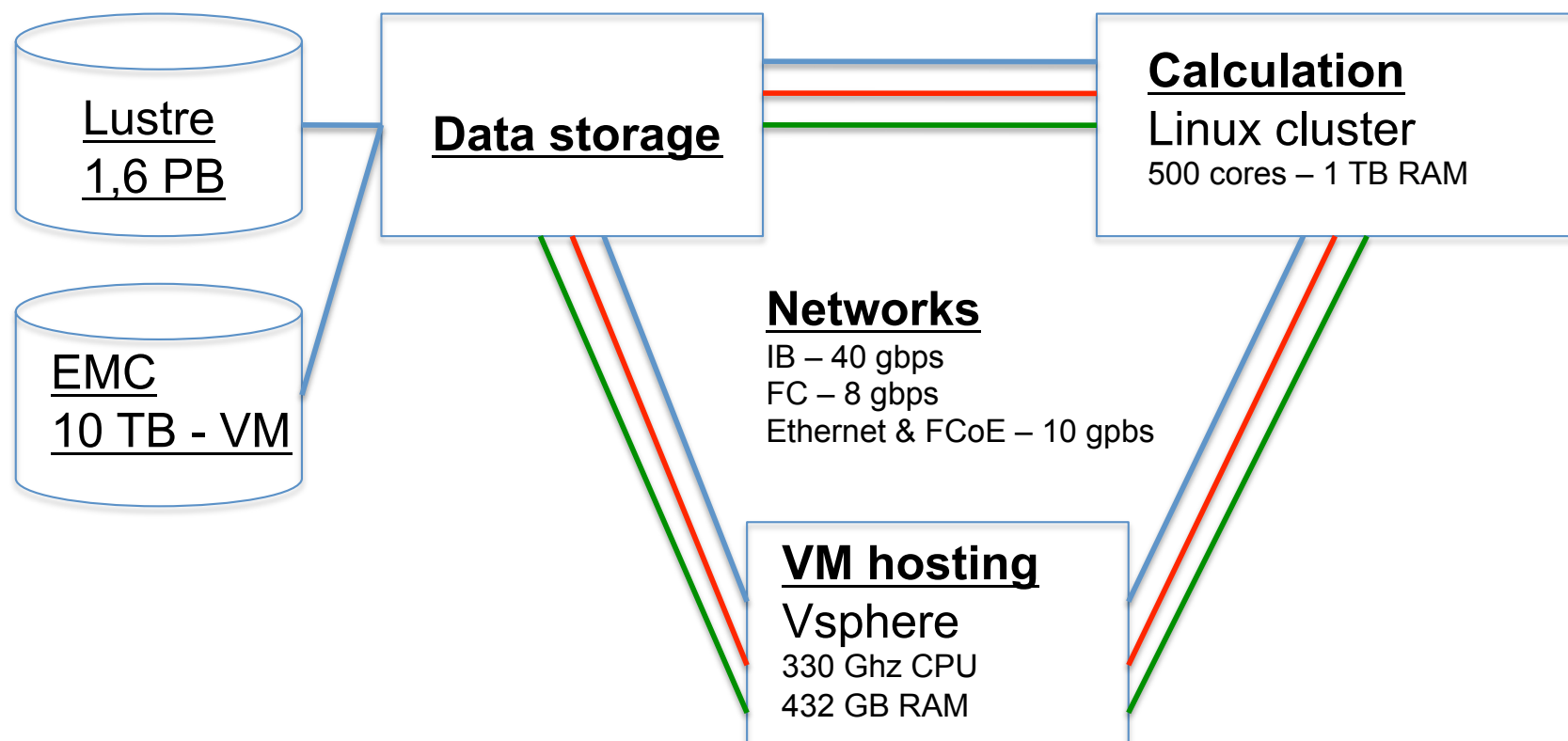
- 2 data clusters (<http://mesocentre.ipsl.fr>)
- CICLAD and CLIMSERV

CICLAD focus

- facility supported by IPSL and dedicated to climate and environmental community
- designed with a 'bottom-up' rationale
- technical platform delivering
 - petascale data storage
 - HPC cluster
 - easy access for curated and value-added dataset via a set of dedicated virtual machines
- *Originality* : lustre data cluster, computationnal cluster and Vsphere virtualization cluster converge on an Infiniband network

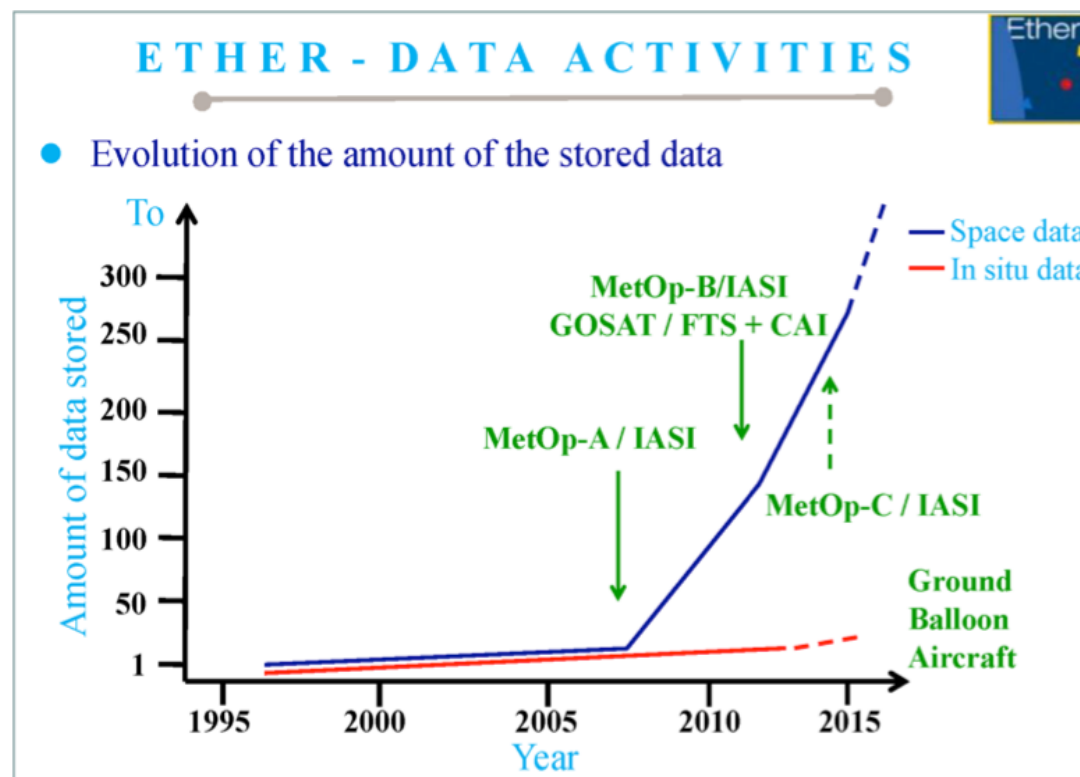


Focus on CICALAD data cluster metrics



Data trend ?

- CMIPdata at IPSL
 - CMIP 3 = 35 To
 - CMIP 5 = 450 To
- Increasing model resolution
- Example :
Atmospheric chemistry data



- Need for secure a subset of our scientific data (primary storage on disks)
- Eudat call for collaboration project
- 2 dataset categories
 - Observation – very costly or non reproducible observation dataset (60 TB)
 - Dataset requiring durable archiving - (450 TB)
- Data format
 - modeling results datasets (HDF5)
 - observation datasets (ASCII, « project-specific » binaire)
 - satellite datasets (ASCII, BUFR, HDF5)

- Life before Eudat
 - « local » archives on LTO tapes
 - time consuming
 - quite unsafe (archived data in the same room as disks)
 - ... archiving data is a full time job
 - ... let (thanks to) specialists (to) do that job !
- Eudat collaboration
 - support of B2SAFE service to replicate our data
- First step into Eudat services

MERCI !

- Thanks to EUDAT to help IPSL for this project
(Johannes Reetz)
- Thanks to CINES (Stephane Coutin – Marion Massol)
 - kick off meeting was held on 11 Apr 2014
 - we hope starting first data transfers at summer 2014

