

H P C - Big Data

High Performance Computing and Big Data for climate research

- Jens Weismüller, LRZ -

Challenges and Benefits of Supercomputing in ClimEx

- 84,000,000 core-hours
≈ 2,500 years on a modern Laptop



Image: redjar (Jared C. Benedict), Wikipedia



Challenges and Benefits of Supercomputing in ClimEx

- 84,000,000 core-hours
≈ 2,500 years on a modern Laptop
- What do we need instead?



Image: redjar (Jared C. Benedict), Wikipedia

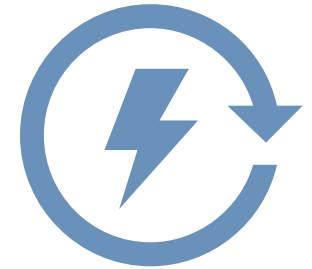




Image: Markus Müller, LRZ

Challenges and Benefits of Supercomputing in ClimEx

Total energy consumption: ~1,25 GWh



Run the wind turbine
in Fröttmaning for
about nine months



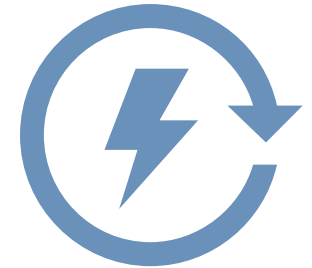
Power ~350
households
for one year



~300 people flying
from Munich to
Montreal and back



Climate simulation means massive energy consumption!
How do we use energy as efficiently as possible?



Innovative cooling technology

- Temperature record high Munich: 37.1°C in August 2003
- SuperMUC: 40°C cooling water
- No active cooling



Challenges and Benefits of Supercomputing in ClimEx

The ClimEx dataset consists of ~400 TB. How much is that?



Image: Asim18 / Wikipedia

Challenges and Benefits of Supercomputing in ClimEx

The ClimEx dataset consists
of ~400 TB. How much is that?

- 50 modern harddisks



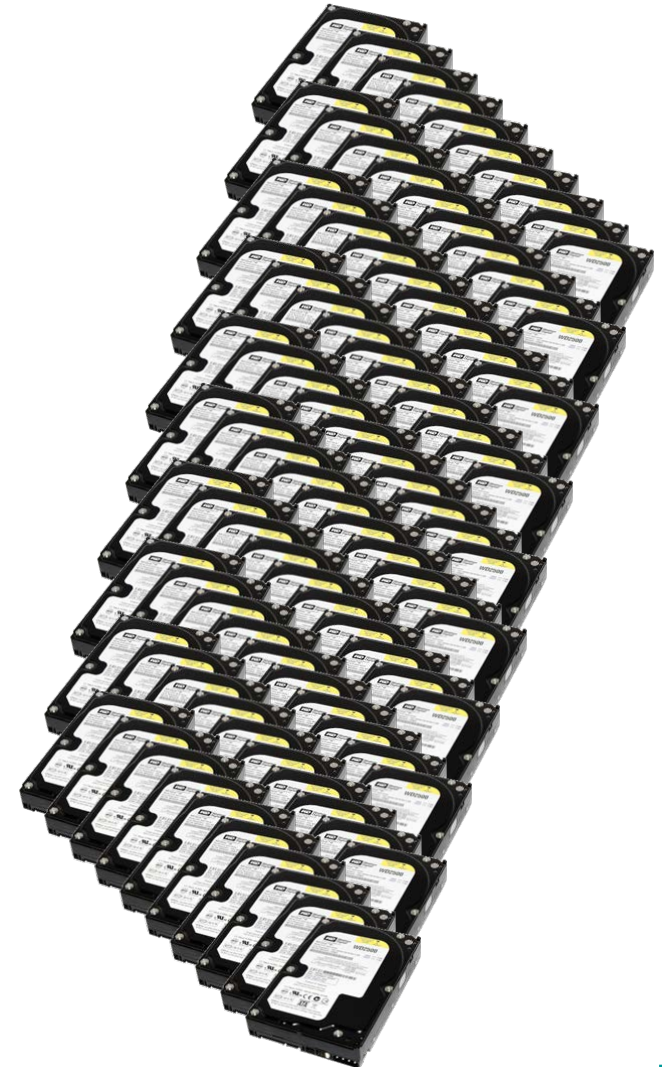
Challenges and Benefits of Supercomputing in ClimEx

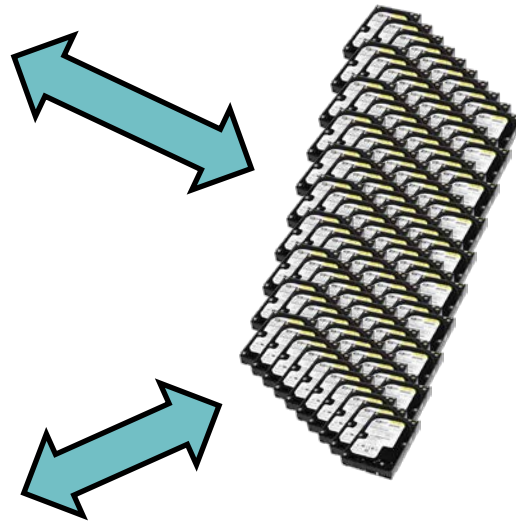
The ClimEx dataset consists
of ~400 TB. How much is that?

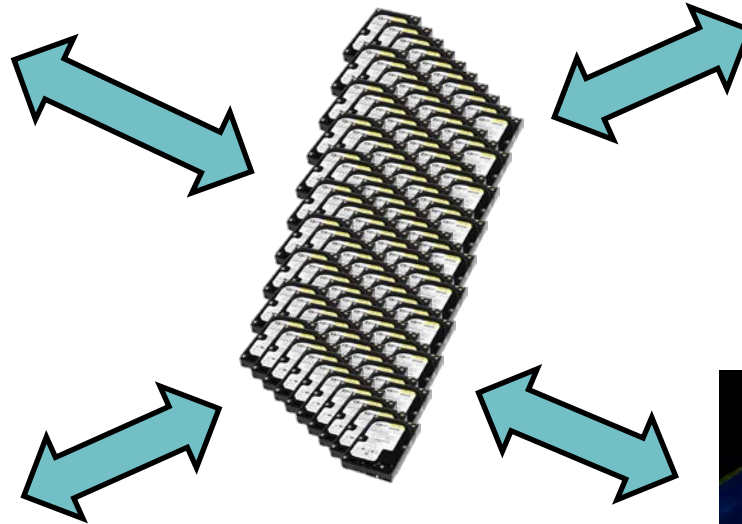
- 50 modern harddisks

But we also need a backup...

- 100 modern harddisk?

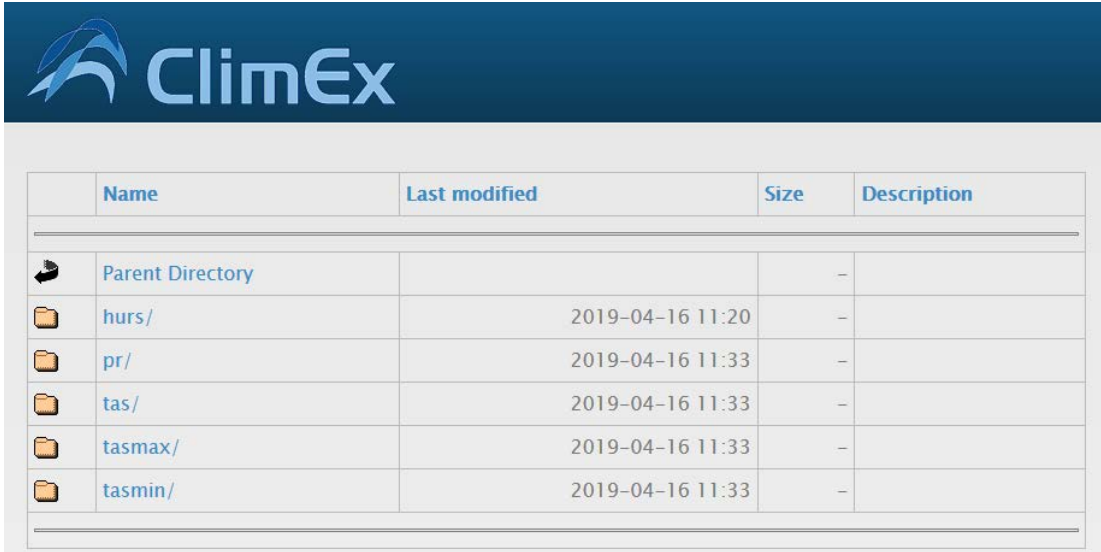












www.climex-project.org/en/data-access

- Web-Interface for exploring the European data

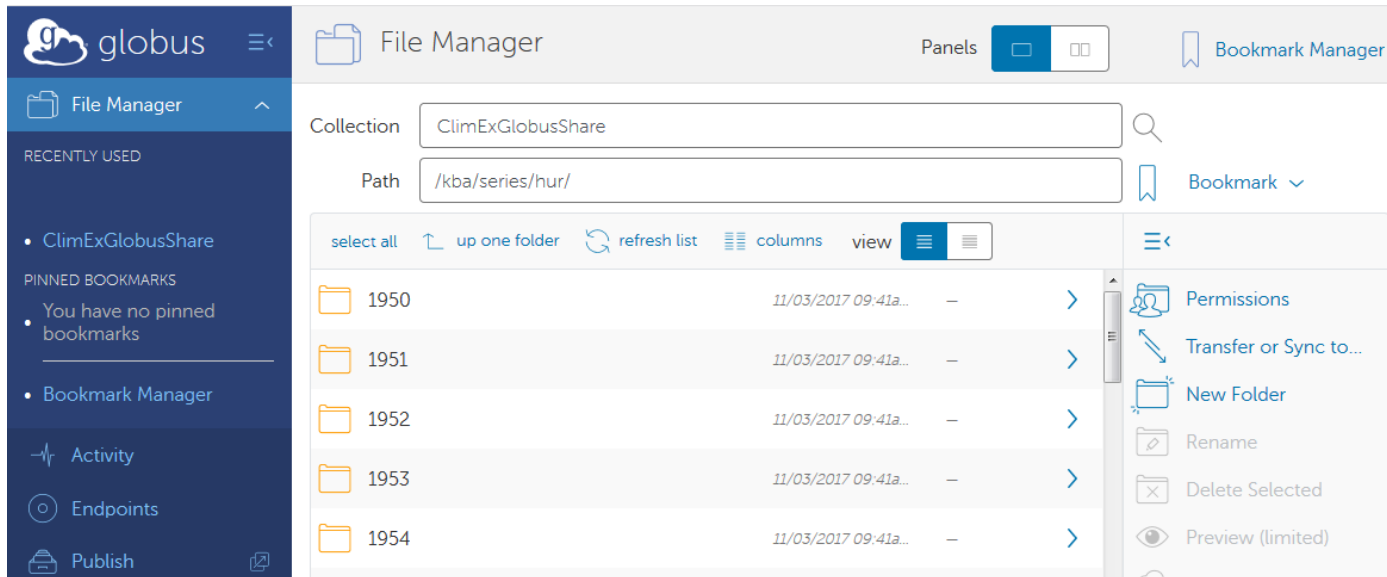


The screenshot shows the ClimEx web interface with a dark blue header containing the logo and name. Below the header is a table listing directory contents. The table has four columns: Name, Last modified, Size, and Description. The rows include a 'Parent Directory' link and five subdirectories: hurs/, pr/, tas/, tasmax/, and tasmin/. Each subdirectory row shows a folder icon, the name, the last modified date and time (2019-04-16 11:33), a size of '-', and an empty description.

	Name	Last modified	Size	Description
	Parent Directory		-	
	hurs/	2019-04-16 11:20	-	
	pr/	2019-04-16 11:33	-	
	tas/	2019-04-16 11:33	-	
	tasmax/	2019-04-16 11:33	-	
	tasmin/	2019-04-16 11:33	-	

www.climex-project.org/en/data-access

- GLOBUS access for transferring large amounts of the European data already in place



The screenshot displays the Globus File Manager interface. The left sidebar shows the 'globus' logo and navigation options like 'File Manager', 'RECENTLY USED', 'PINNED BOOKMARKS', 'Activity', 'Endpoints', and 'Publish'. The main area is titled 'File Manager' and shows the 'Collection' as 'ClimExGlobusShare' and the 'Path' as '/kba/series/hur/'. Below this, a table lists folders for the years 1950 through 1954, each with a date of '11/03/2017 09:41a...' and a right-pointing arrow. A context menu is open on the right, showing options such as 'Permissions', 'Transfer or Sync to...', 'New Folder', 'Rename', 'Delete Selected', and 'Preview (limited)'.

Folder Name	Date	Permissions
1950	11/03/2017 09:41a...	>
1951	11/03/2017 09:41a...	>
1952	11/03/2017 09:41a...	>
1953	11/03/2017 09:41a...	>
1954	11/03/2017 09:41a...	>

Match scientific data to the viewers everyday experiences

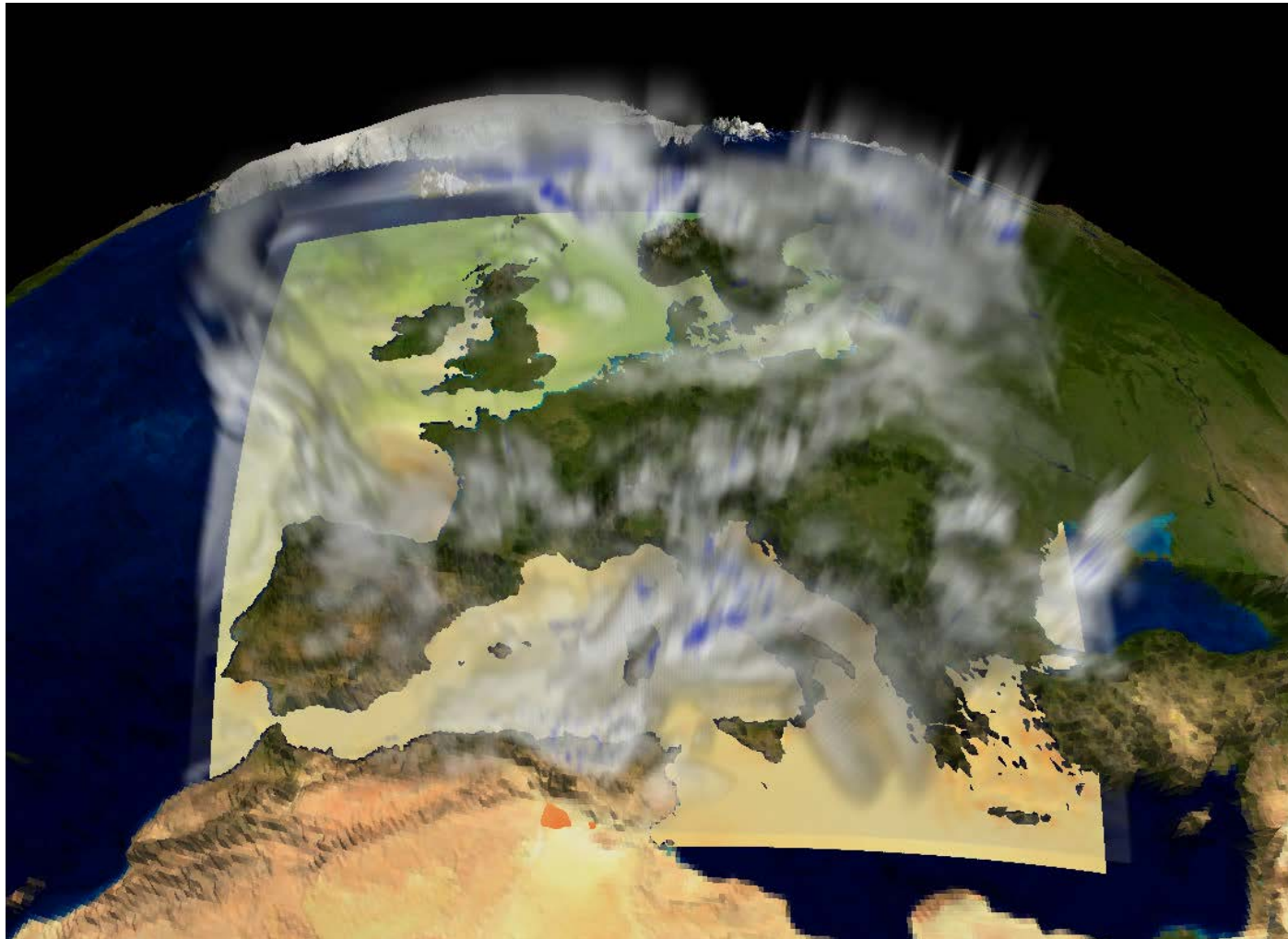




Image: LRZ



Image: Daniel Kolb, LRZ

- CRCM5 and WaSiM ported and adapted to HPC systems
- management of a huge job farming environment
- production of huge climate simulation data set
- data provisioning to the scientific community (e.g. *Wasser Zukunft Bayern*) and stakeholders in the agriculture, forestry, energy, health ... sectors
- pilot usage of new LRZ services and tools (e.g. Data Sharing, Data Science Storage)
- new data visualization approach to support the science-policy interface



www.climex-project.org/en/data-access