Lessons from the WIP and Vision for the Future

IS-ENES3 Kickoff Meeting

V. Balaji

NOAA/GFDL and Princeton University
Visiting Scientist, LSCE and IPSL, Paris

10 January 2019

Outline

1. Current status of ESGF and CMIP6
   - CMIP6 design and timeline
   - Global data infrastructure
   - WIP Documents on Infrastructure Needs
   - Current status

2. Vision for the future
   - Trends in data technology
   - New approaches to data
   - Learning

3. Ideas and challenges

Outline

1. Current status of ESGF and CMIP6
   - CMIP6 design and timeline
   - Global data infrastructure
   - WIP Documents on Infrastructure Needs
   - Current status

2. Vision for the future
   - Trends in data technology
   - New approaches to data
   - Learning

3. Ideas and challenges

Fast forward to today...

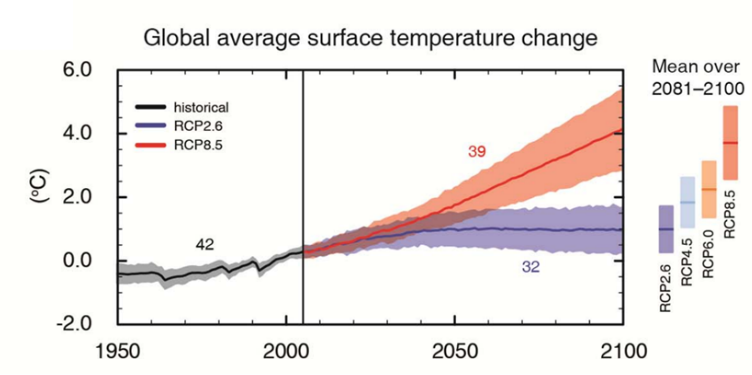
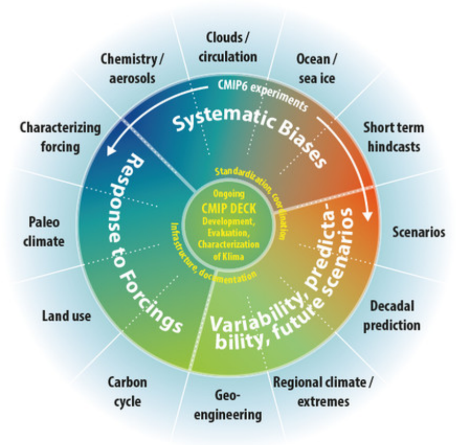


Figure SPM.7 from the IPCC AR5 Report. 20th century warming cannot be explained without greenhouse gas forcings.

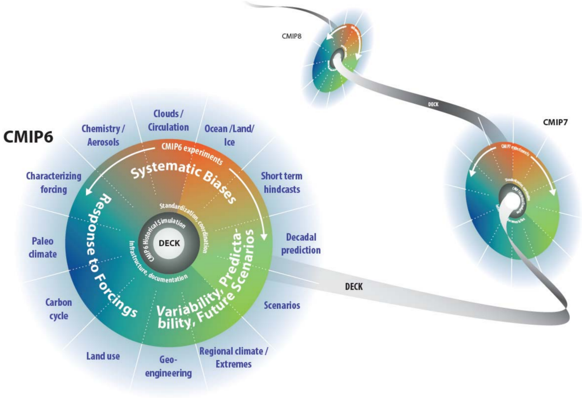
CMIP6 design: DECK and MIPs



DECK experiments form the core; many specialized MIPs for smaller communities, some 24 of which have been endorsed by CMIP panel. Figure courtesy Meehl et al (*Eos* 2014).

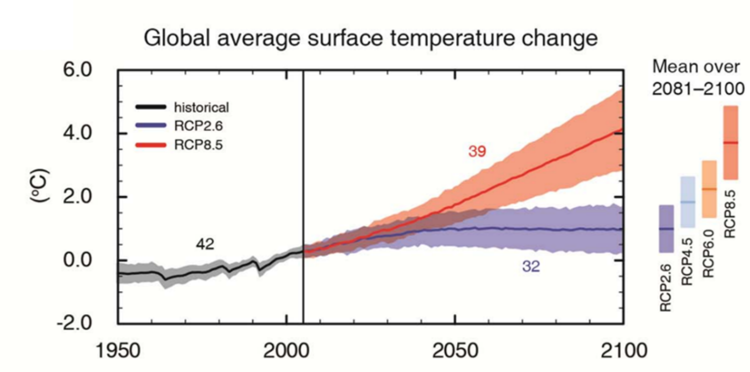
CMIP evolution

DECK is designed to evolve slowly or not at all.



- IPCC Assessment Reports are snapshots of the “state of the science”, but not directly linked to CMIP.

Multi-model ensembles for climate projection



- Critically depends on software, metadata, and data standards: the Earth System Grid Federation (ESGF) archive and software stack, which includes many non-ESGF stacks (ES-DOC, ..)
- Key technical issues like replication, versioning, subsetting, QC, citation.

The global data infrastructure underpinning MIPs

- MIPs, and in general any science involving cross-model comparisons, critically depend on the global data infrastructure – the “vast machine” (Edwards 2010) – making this sort of data-sharing possible.
- Infrastructure should not be a research project.
- Infrastructure should be treated as such by the national and international research agencies, but it is instead funded piecemeal, as a soft-money afterthought. This places the system at risk (NRC 2012: “A National Strategy for Advancing Climate Modeling”, ISENES-2 Infrastructure Strategy document, 2012.)

Role of WGCM and its infrastructure panel

- Provide scientific guidance and requirements for the GDI; exert greater influence over its design and features.
- Provide standards governance allowing for orderly evolution of standards.
- Provide design templates (e.g CMOR extensions) for groups designing MIPs and work to ensure their conformance to standards.
- Work with academies and publishers to require adequate data citation and recognition for data providers.
- Intercede with national agencies to provision data infrastructure with adequate and stable long-term funding.

WIP: The WGCM Infrastructure Panel formed 2014

- Chaired by V. Balaji (Princeton/GFDL) and K. Taylor (PCMDI).
- Strategy to develop a series of “position papers” on global data infrastructure and its interaction with the scientific design of experiments. These will be presented to WGCM annual meeting.
- Close involvement of the WIP and CMIP panel (e.g. joint papers)
- Interest from other WCRP working groups (WGSIP, WGNE)
- Covers not only ESGF requirements but also other tools: ESDOC, CMOR, CF Conventions, ..
- a blend of computer and climate scientists representing data centers and modeling groups: rotating membership with overlapping 2-year cycles

ESGF, a global network of compute and data nodes

Figure courtesy IPSL.

WIP Position Papers

- Requirements for global data infrastructure, Balaji et al (2018).
- https://earthsystemcog.org/projects/wip/

- CDNOT Terms of Reference: a charter for the CMIP6 Data Node Operations Team.
- CMIP6 Global Attributes, DRS, Filenames, Directory Structure, and CVs: conventions and controlled vocabularies for consistent naming of files and variables.
- CMIP6 Persistent Identifiers Implementation Plan: a system of identifying and citing datasets used in studies, at a fine grain.
- CMIP6 Replication and Versioning: a system for ensuring reliable and verifiable replication; tracking of dataset versions, retractions and errata.

WIP Position Papers

https://earthsystemcog.org/projects/wip/

- CMIP6 Quality Assurance: systems for ensuring data compliance with rules and conventions listed above.
- CMIP6 Data Citation and Long Term Archival: a system for generating Document Object Identifies (DOIs) to ensure long-term data curation.
- CMIP6 Licensing and Access Control: terms of use and licenses to use data.
- CMIP6 ESGF Publication Requirements: linking WIP specifications to the ESGF software stack, conventions that software developers can build against.
- Errata System for CMIP6: a system for tracking and discovery of reported errata in CMIP6.

IPCC Timeline

All dates in red are official dates from IPCC plenary in Nairobi, 2016-04, and IPCC XC meeting 2016-05-19.

- 2022-09: AR6 Synthesis Report
- 2021-02: WG1 Report Approved
- 2020-12: Final WG1 Draft and SP goes to inter-governmental review
- 2020-06: 4th Lead Author meeting
- 2020-02: Post 3rd LA meeting, second-order draft sent out for expert review. Any citations here will have to have been submitted for peer review by this date, accepted by 2020-09. First-order draft can use pre-citation material.
- 2019-07: Data in public domain.
- 2019-08: Third Lead Author Meeting.
- 2019-04: Second Draft Expert Review
- 2019-01: Second Lead Author Meeting.

Earlier special reports (1.5C, cryosphere, land) not based on CMIP6.

Current status

- 45 registered institutions, 102 models, 287 experiments, see CVs on Github.
- Multiple working versions of CMIP6 Data Request, current version 01.00.29.
- CVs on Github are sole source for all conventions (DREQ, DRS, CVs, CF) and verification tools (PrePARE, CMOR, ESGF Publisher)
- Expected data volume ∼20 PB (compressed netCDF).
- input4MIPs, obs4MIPS, etc. also hosted on ESGF.
- CDNOT has stress-tested the system with a series of data challenges from publication, to replication, to replication at scale.
- 8 models now are submitting DECK runs.
- Issues: Globus, model documentation (ES-DOC), server-side capabilities, ...

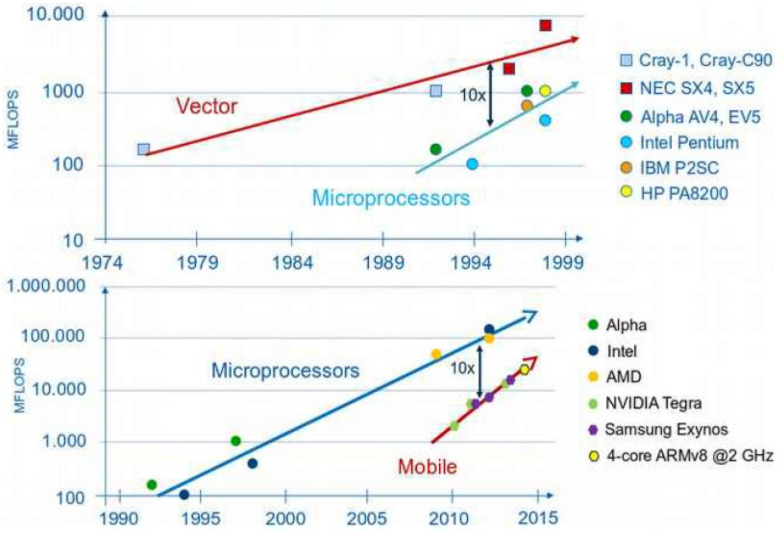
Outline

1. Current status of ESGF and CMIP6
   - CMIP6 design and timeline
   - Global data infrastructure
   - WIP Documents on Infrastructure Needs
   - Current status
2. Vision for the future
   - Trends in data technology
   - New approaches to data
   - Learning
3. Ideas and challenges

Scaling the problem

- Model analysis is performed on a 6D dataset (x, y, z, t, v, n)
  - NX,NY $\sim$ 1000 today... 10^4 at $\mathcal{O}$(1 km)
  - NZ $\sim$ 100
  - NT $\sim$ 10^6 (e.g 100y of daily data)
  - NV $\sim$ 1000
  - NENS $\sim$ 1000 (multi-model, multi-parameter, multi-IC)
- That's (conservatively) 10^{21} bytes. In comparison, GFDL's (tape!) archive is 120 PB, ECMWF 240 PB ($\sim$ 10^{17} bytes).
- Done hierarchically: analyze responses to changes at single-parameter, single-component, single-model, multi-model
- Feedback from outer cycles of model development is not readily available to earlier stages.

The inexorable triumph of commodity computing

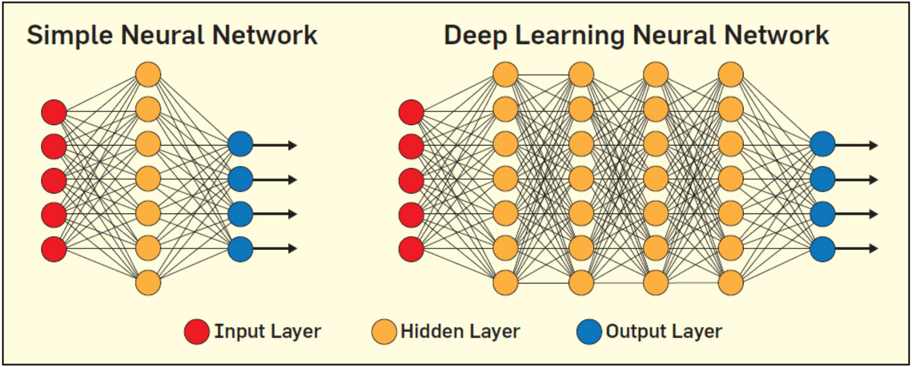


From *The Platform*, Hemsoth (2015).

US Exascale Roadmap

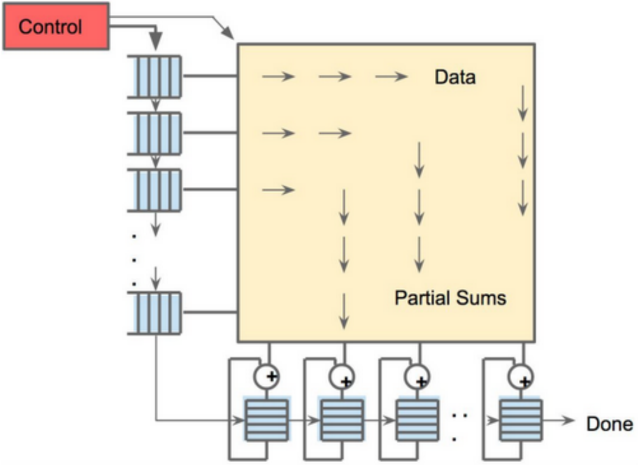
Courtesy Exascale Computing Project.

Deep Learning



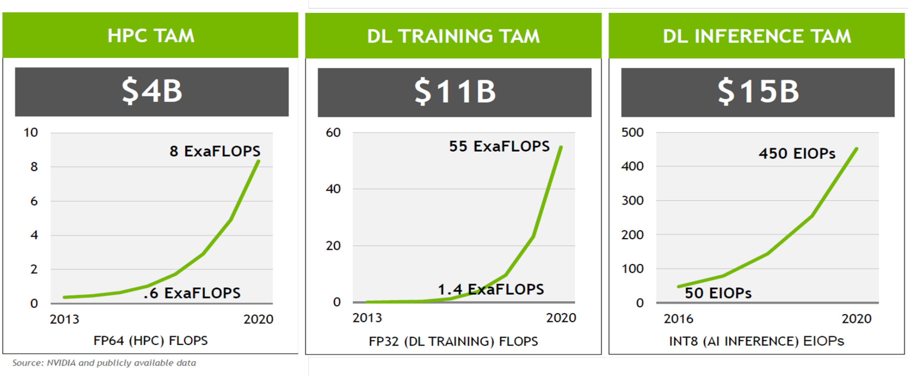
From Edwards (2018), ACM.

Google TPU (Tensor Processing Unit)



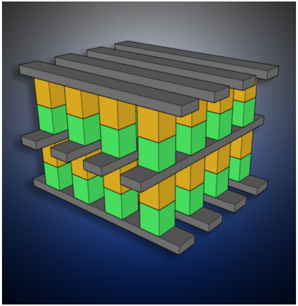
Hardware pipelining of steps in matrix-multiply. Figure courtesy Google.

ML is subverting the HPC market



Courtesy NVidia, via Seeking Alpha.

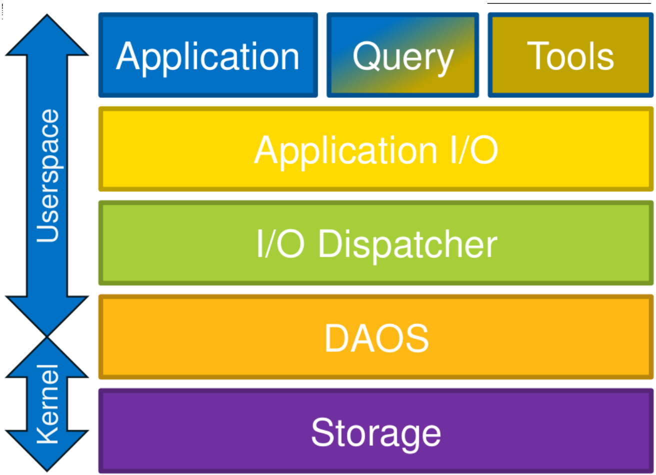
Energy cost of data movement



- To fit 10^{18} op/s within a 1 MW power budget, an operation should be 1 pJ: data movement is ∼10 pJ to main memory; ∼100 pJ on network!
- New technologies (NAND flash, 3DXpoint) reduce this, but by introducing so much parallelism that “POSIX files” become a dubious proposition

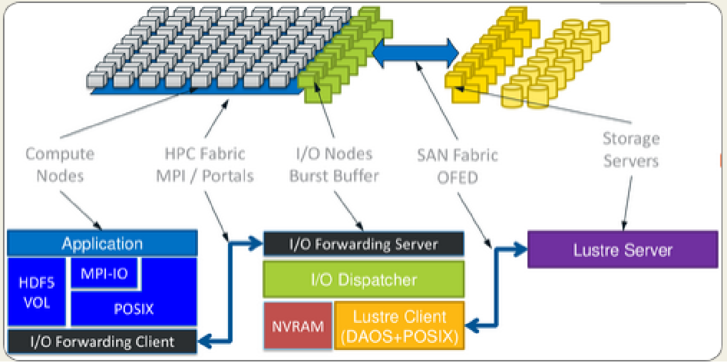
Figure courtesy Intel.

Blurring distinction between memory and filesystem



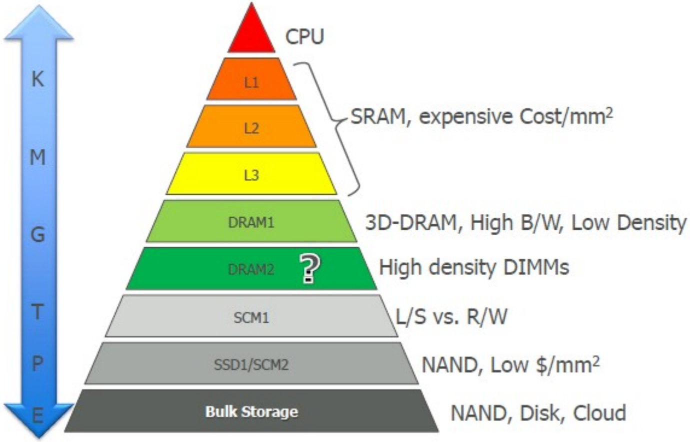
Hemsoth, 2014.

NVRAM as primary storage



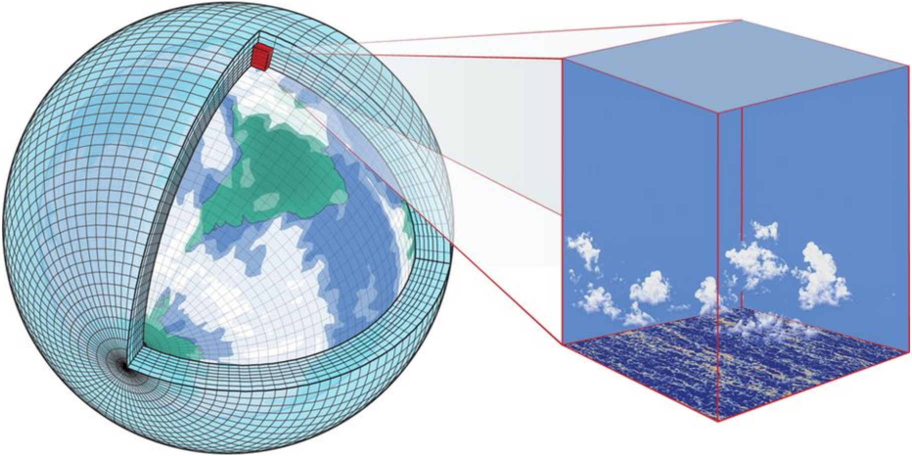
Hemsoth, 2014.

Deep memory and storage hierarchy



From Morgan (2015), *The Next Platform*.

Caltech/MIT Earth Machine



From Schneider et al 2017.

Pangeo: a creative attack on the problem

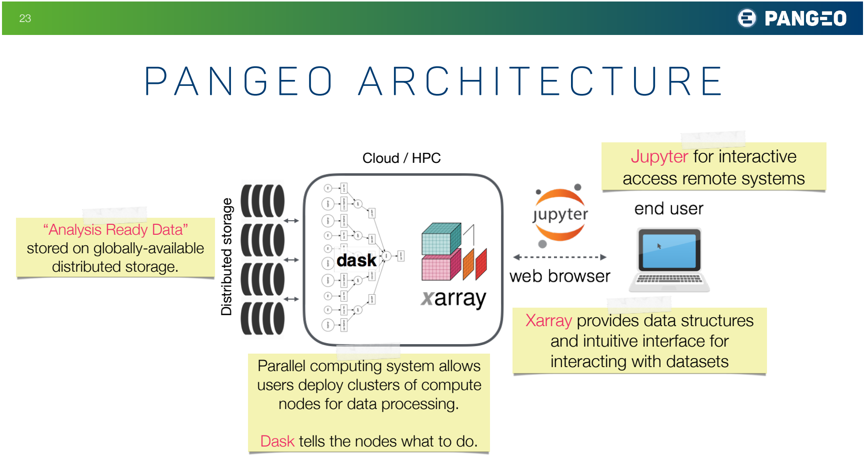


Figure courtesy Ryan Abernathey, Columbia.

Traditional data storage

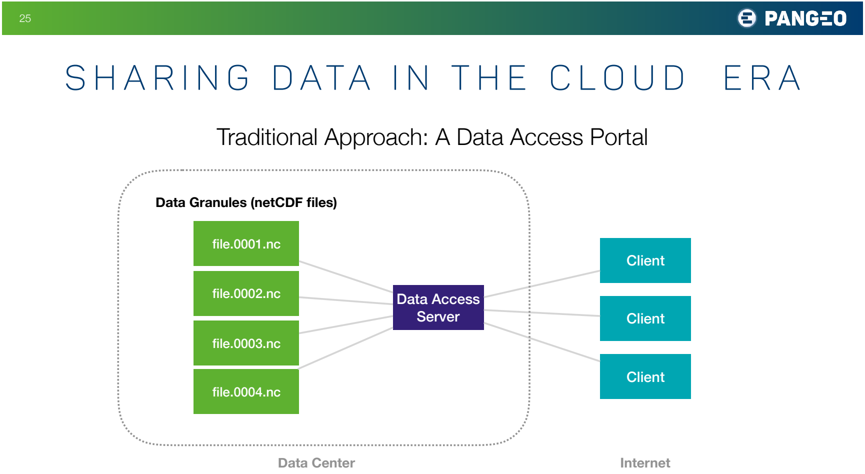
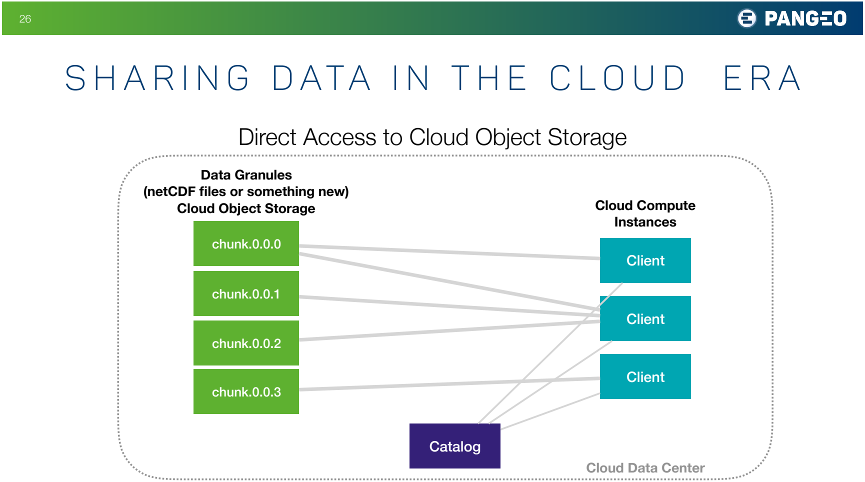


Figure courtesy Ryan Abernathey, Columbia.

... versus object store



But which hyperslab of (x, y, z, t, v, n) should go in a chunk?

Outline

1. Current status of ESGF and CMIP6
   - CMIP6 design and timeline
   - Global data infrastructure
   - WIP Documents on Infrastructure Needs
   - Current status
2. Vision for the future
   - Trends in data technology
   - New approaches to data
   - Learning
3. Ideas and challenges

Solve the problem across the data lifecycle

- Computing: end of Dennard scaling
  - Minimum capability for weather: 0.5 SYPD
  - Minimum capacity for validating a seasonal forecast model: 10^4 SY
  - Typical load for a CMIP model: 10^4 SY (tip of the model development iceberg!)
  - Science needs: both GCRM (HR, short duration) and paleoclimate (LR, long) models: single infrastructure for model hierarchy: idealized to comprehensive; algorithmic (conservation, ...)
- Data: serial workflow: simulate, postprocess, analyze.
  - Consequence: simulation to analysis in $\sim$years
  - Consequence: data harmonization in $\sim$ months
  - Consequence: data replication of PB-scale datasets
  - Potential breakthrough: demonstrate an analysis infrastructure that processes the 6D (x, y, z, t, v, n) dataset in real time as models run.
- How to decarbonize our science? See CPMIP (JPSY)
- Tension between robustness and reliability and keeping up with technology evolution (containers, cloud, ML): let science lead the way!