High Performance Computing

Model Data Fusion and Earth System Modelling

www.cawcr.gov.au

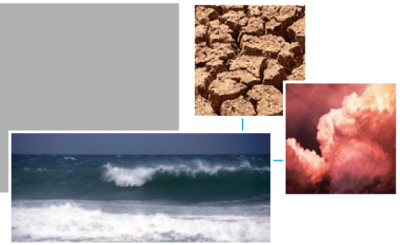
Tim F. Pugh
CAWCR Senior IT Officer
Australian Bureau of Meteorology
11 May 2010

Australian Government
Bureau of Meteorology

The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology

CSIRO

Questions

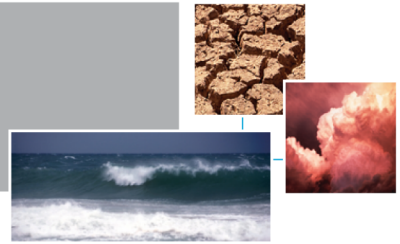


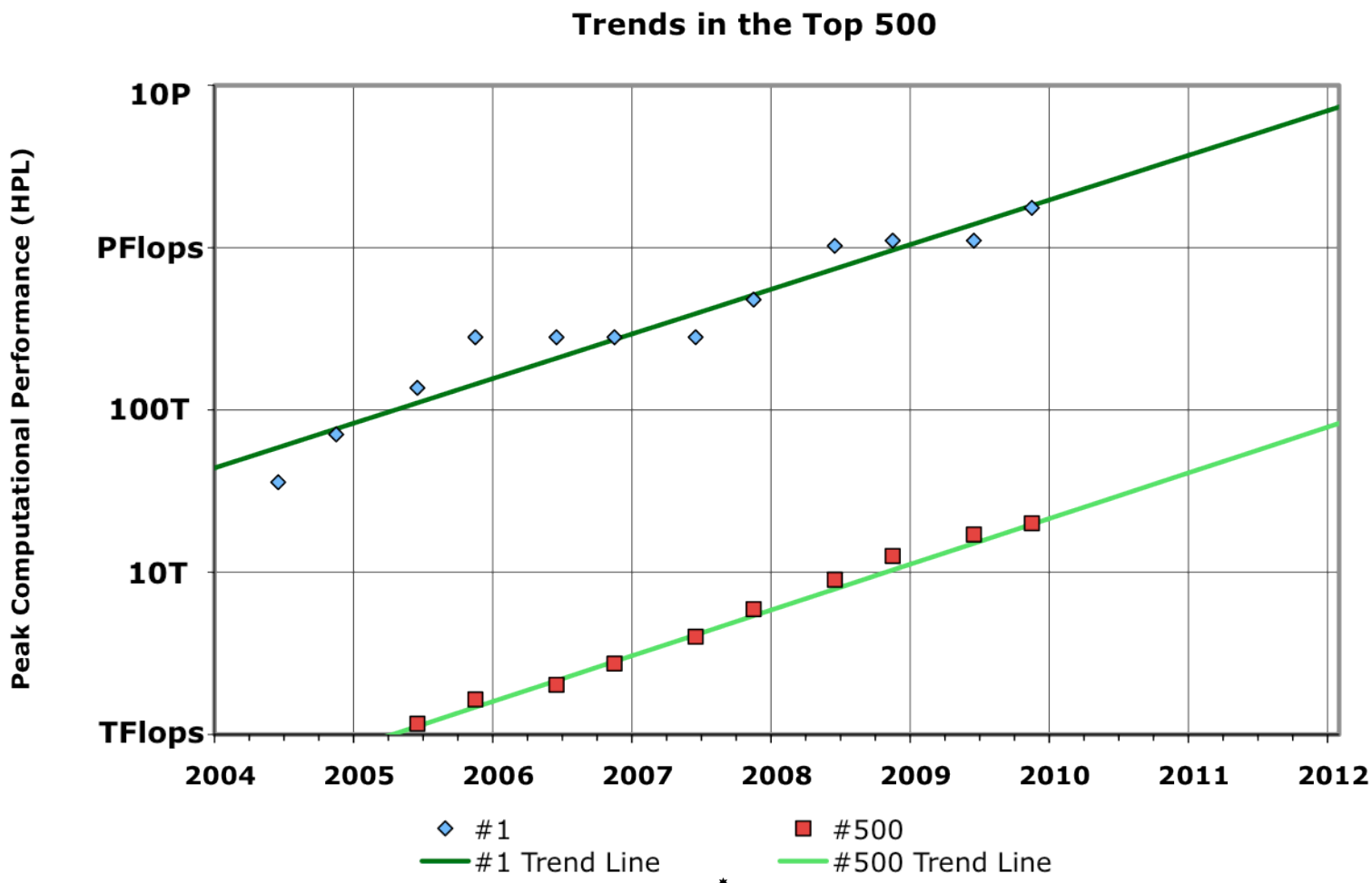
- What are the current and expected platforms that the vendors are delivering / promising now?

- What are the critical bottlenecks that must be addressed and balanced
  - Processors: Moore's Law continue?
  - Memory wall (latency and bandwidth)
  - System interconnects
  - Storage (disk and tape)
  - data management systems
  - Wide area networks

- How would you plan to make use of it in, say, 5 years?

Top500 List



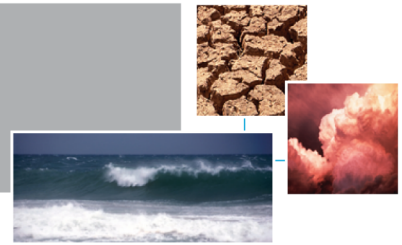


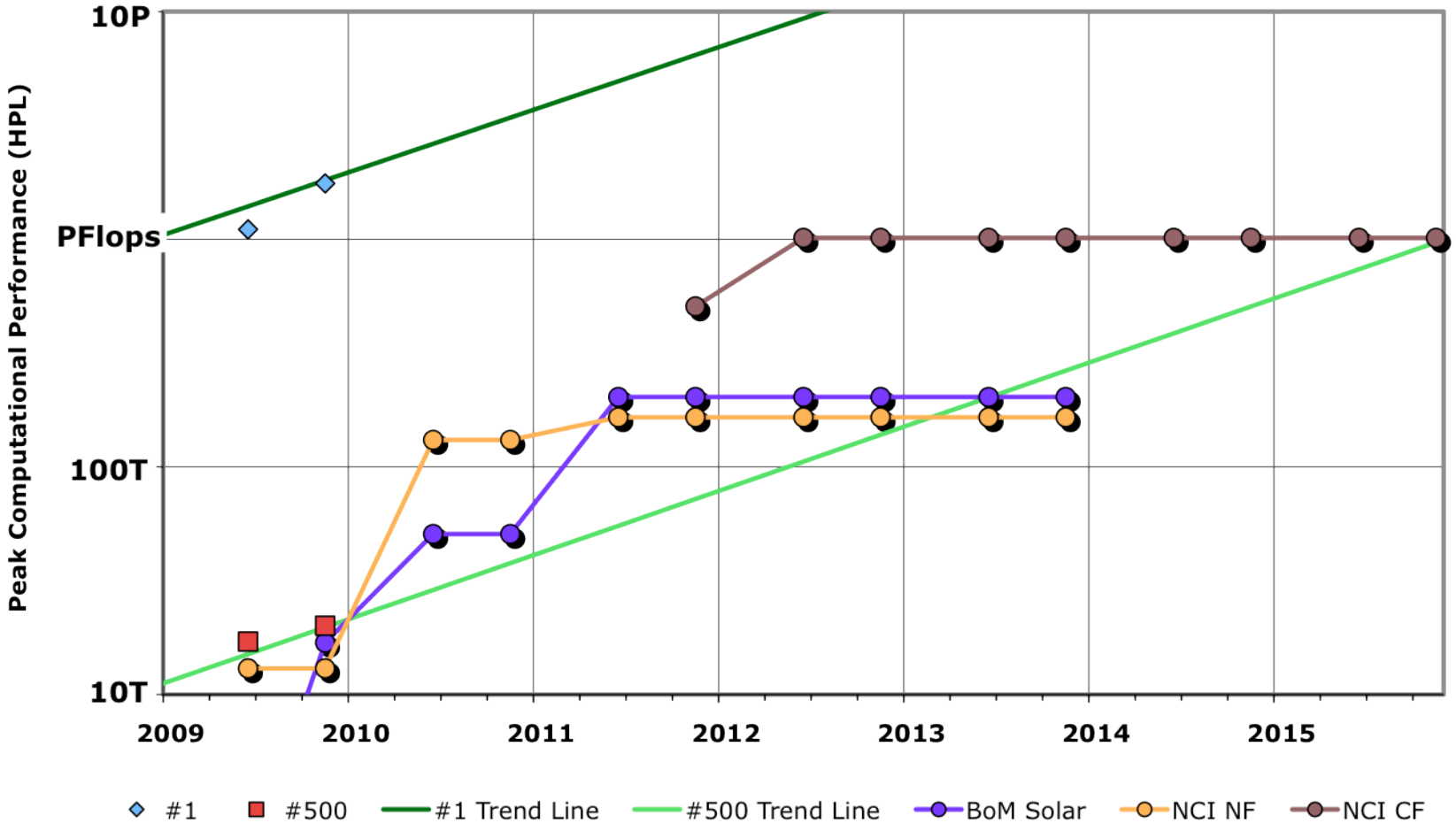
Australian Government
Bureau of Meteorology

The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology

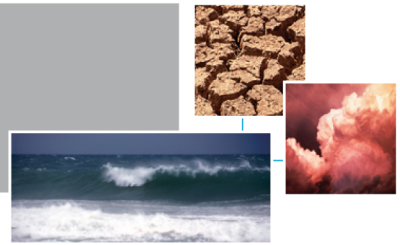
CSIRO

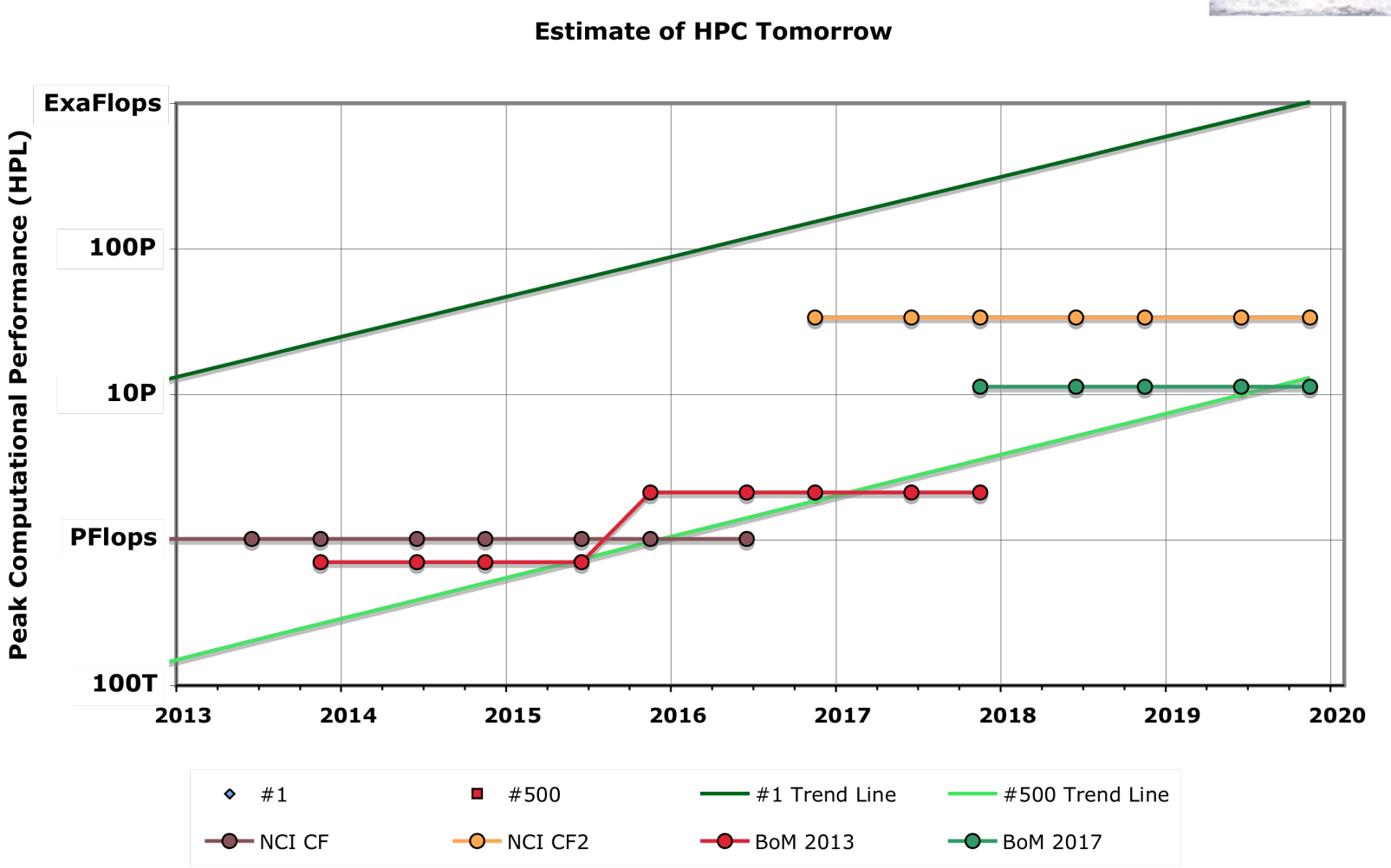
HPC Capability Today



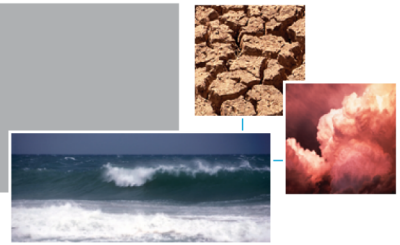


Estimate of HPC Tomorrow



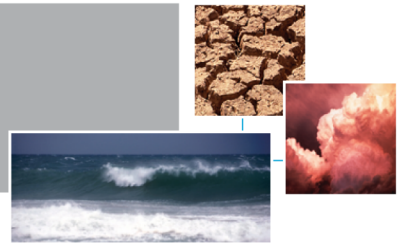


Processor Trends



- **Components of Moore's Law**
  - Pipeline depth (gates/clock)
  - Instruction Level Parallelism (ILP)
  - Clock frequency
  - Compute cores
- **Problems with increasing frequency/cores and reducing feature size indefinitely:**
  - Power proportional to clock frequency and number of cores.
  - Control of fab process increasingly difficult as feature size reduced. Yields drop.
  - Reliability inversely decreases with increased gates
  - Reliability reduced with higher power densities
- ~~**Moore's Law increasingly depending on frequency**~~
- **Moore's Law increasingly depending on cores**

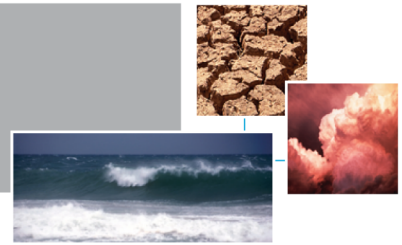
Processor Trends (cont.)



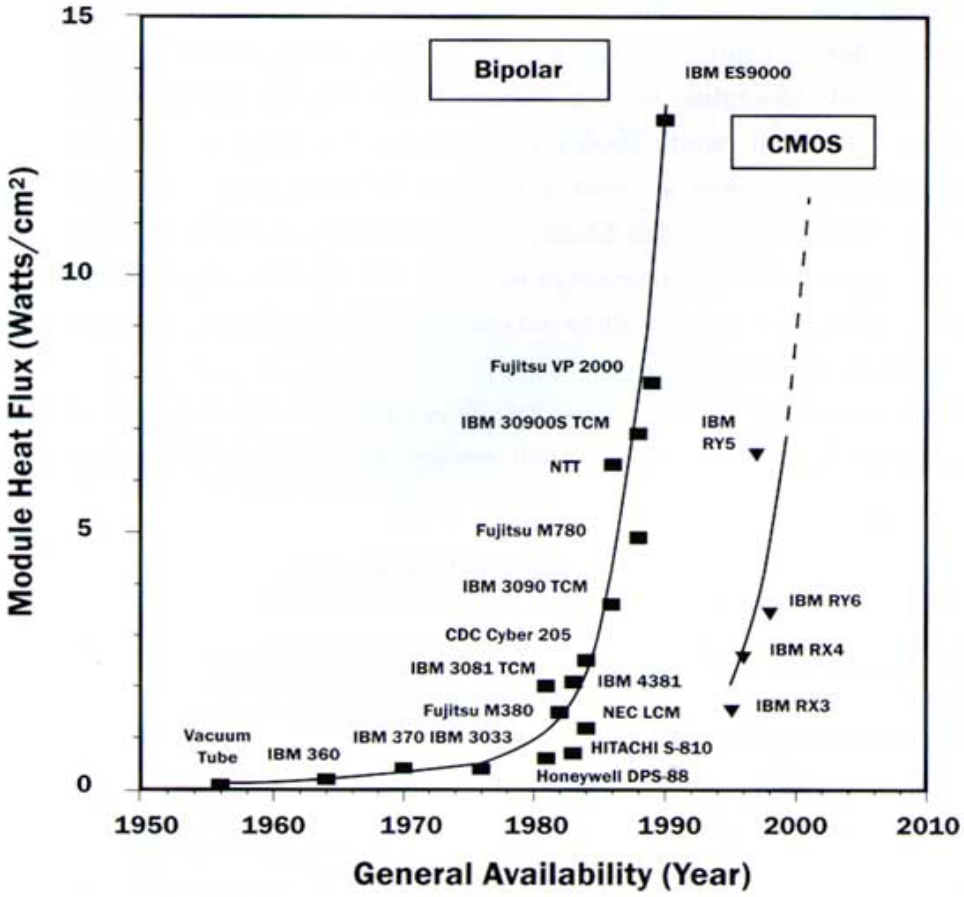
- **End of Moore's Law in 2020?**
  - Can't afford the power budgets
  - End of CMOS
  - 32 um -> 22 um -> 15um -> 10 um
  - Approaching a few silicon atoms layer for transistors
  - Power leakage goes up
  - Voltage is pushed down -> loss of reliability

- Reliability
  - Ability to prevent failure
- Resilience
  - Ability to recover from failure

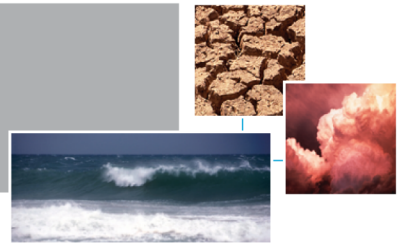
The CMOS Power Problem is at Hand



- **Between 2000 and 2009, max chip power will have increased more than 100%**
- **Heat flux will have more than doubled**
- **The main culprits are increasing clock frequencies, additional cores, and decreasing feature sizes**
  - **Power (Watts) = C V^2 f**
  - **Heat flux = power/area**

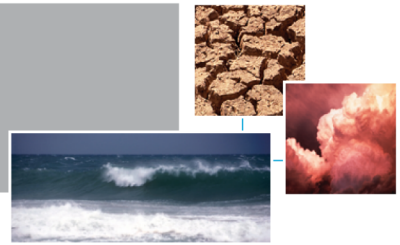


CMOS Power Perspective



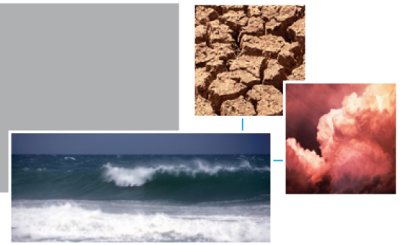
- CMOS situation is similar to what happened to Bipolar in the early 1990s.
  - Difference: there is no ready replacement for CMOS on the horizon.
  - As feature size decreases, static power (leakage) will become comparable to dynamic (switching) power consumption. (Loss of reliability)
  - Costs of facility infrastructure for providing power to and dissipating heat from supercomputers can no longer be neglected.

Processor Trends (through 2012)



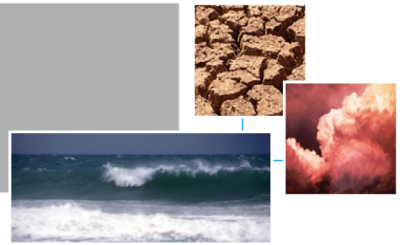
- Multi-Threaded architectures
- Multi-Core Processors
- O(3GHz) clock speeds
- Large/Huge L2, L3 caches
- Media/vector processing extensions
- Static and dynamic power management

Processor Trend Conclusions



- Trends in processor architecture
  - Moore's Law will slow
  - Even so, the memory wall (latency measured in clocks) will continue to increase
  - Level of system integration on a chip will increase
    - Multi-core on a chip
    - Memory controller on a chip
  - Power consumption issues will increasingly constrain design/market
  - Commodity vs custom chip battle will continue
- All Trends are interrelated and will impact future designs.

General Purpose GPU

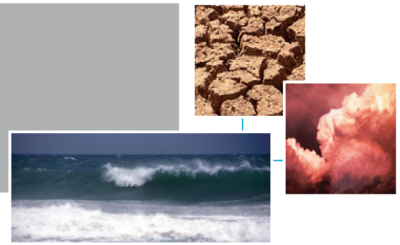


- Today's general purpose CPU's per-core performance has not increased in performance significantly
  - CPU multi-core chip performance increasing --> many-core chips
  - Increasing the number of processors / cores has diminishing value

- Multi-threaded, many-core chips for graphics processing are not general purpose, but they have significant performance advantages over general purpose CPUs

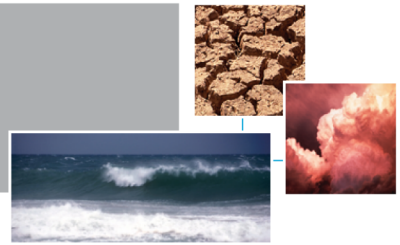
- Peak CPU performance (4 core) Today
  - *~11GFlops per core, 20-30W*
  - *~44GFlops per chip, 80-120W*
- Peak GPU performance Today
  - *930Gflops, 150-300W*

General Purpose GPU (cont.)



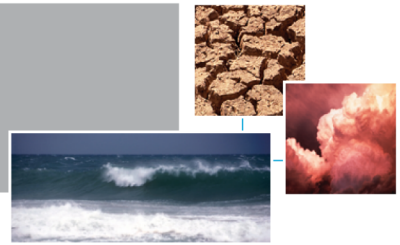
- What's the catch?
  - Achieving good performance may require significant modifications and/or restructuring
  - Significant time required to transfer data in memory between the GPU and CPU
  - Portability across different accelerator technologies may be challenging
  - Many-core, GPU architectures are quite different
  - OpenCL, CUDA may not be sufficient
  - Limited development tools

Preparing for the Future



- Scalability Projects
  - Prepare for the possibility on running codes on (more) massively parallel computers (BlueGene)
- Optimization Projects
  - Improve efficiency of codes on scalar CPU architectures
    - Access codes running at 5% of peak today, desire target of 10% or more.
- Code Portability Projects
  - Prepare for the possibility on running codes on new processor technology and new languages (GPGPU, FPGA)
- Collaborative effort with other groups
  - Access to code and benchmarks for community and HPC help
  - Exploration of new technology benefits and costs

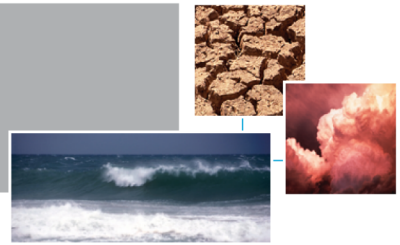
Future Systems



- Enhance our capacity for computing (100x by 2020)
  - Greater computing resources for ensemble model and assimilating data
  - Greater memory resources and performance
    - 3GB memory per Gflop of sustained computing
  - Greater storage resources and performance
    - Global parallel file systems (Lustre) scale-out in storage and bandwidth
    - Flash technology improves I/O bandwidth and Iops

- Enhanced capability for computing depends on
  - Whether the application can scale with growing number of cores?
    - Improvement in communications, and overlaps in computations
    - Hybrid models using MPI-OpenMP
    - Application I/O, parallel I/O or parallel data streams
  - Whether the application needs greater single processor capability?
    - Possibly processors, new coding techniques and languages, or
    - Wait for better tools

Power and Cores

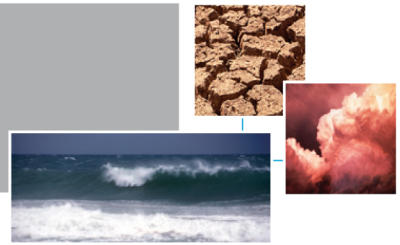


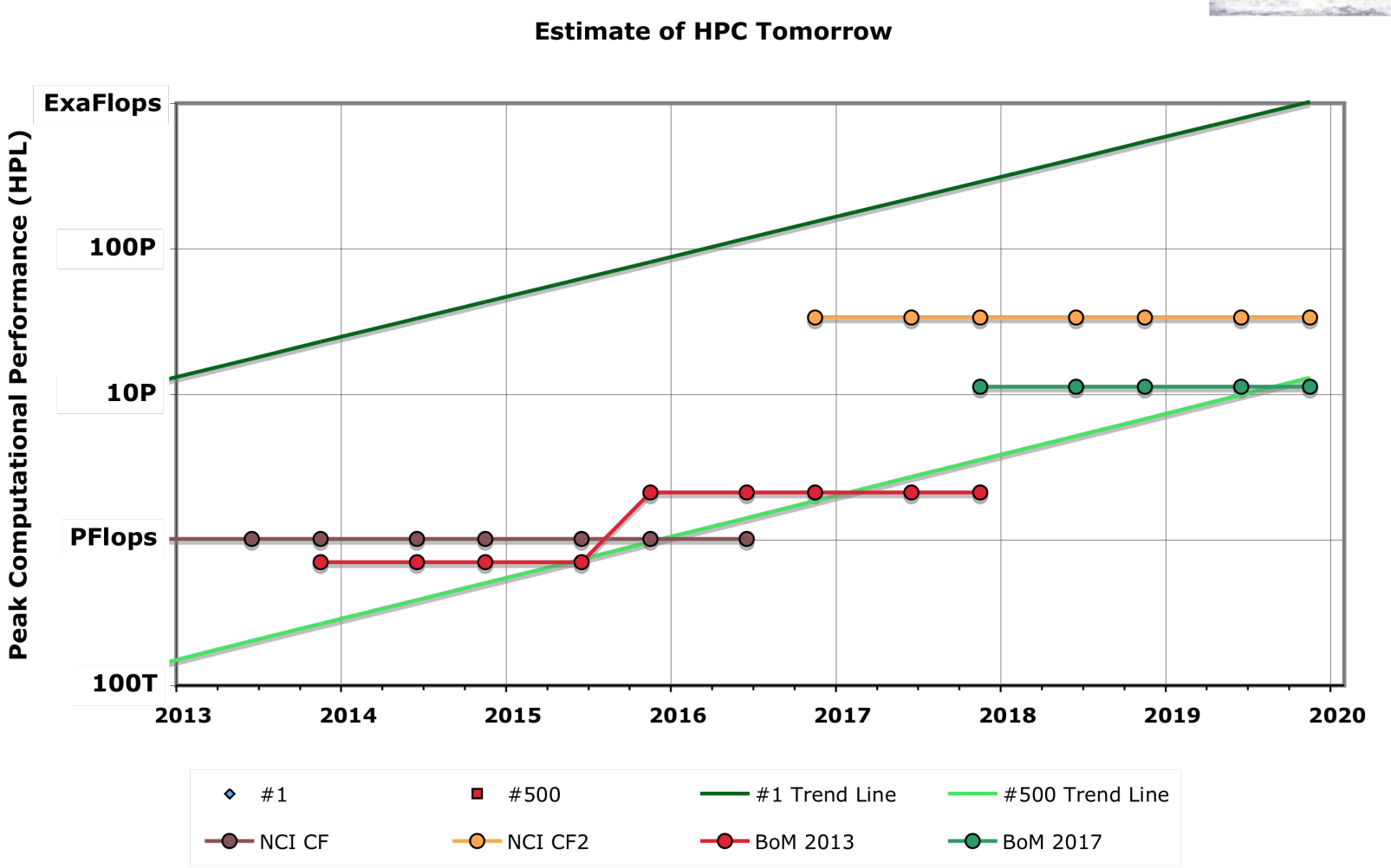
- Today - Processor Counts and System Power
  - TeraScale is >100 cores (~10 KW)
  - PetaScale is >90,000 cores (~10 MW)
  - ExaScale is >90,000,000 cores (>300 MW)

- Tomorrow – Exascale Processor Counts and System Power
  - TeraScale is >1 cores (1 KW?)
  - PetaScale is >1,000 cores (1 MW?)
  - ExaScale is >100,000 cores (100 MW?)

- Is Exascale achievable?
  - Not today. Issues with system power and application scalability (cores)
  - In 2020, yes if…Achieve affordable power budgets
    - Achieve balanced system design
    - Improve application scalability

Estimate of HPC Tomorrow





Australian Government
Bureau of Meteorology

The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology

CSIRO

Tim F. Pugh
High Performance Computing

Phone: 03 9669 4345
Email: t.pugh@bom.gov.au
Web: www.cawcr.gov.au

Thank you

www.cawcr.gov.au