Introduction to HPC and parallel architectures: software

Mike Ashworth

Head of Technology Division
Scientific Computing Department
and
STFC Hartree Centre
STFC Daresbury Laboratory

mike.ashworth@stfc.ac.uk

Overview

Hardware and software:
from evolution to revolution

STFC's Scientific Computing Department

STFC Hartree Centre: future developments to meet future challenges

Opportunities for Discussion

Overview

Hardware and software:
from evolution to revolution

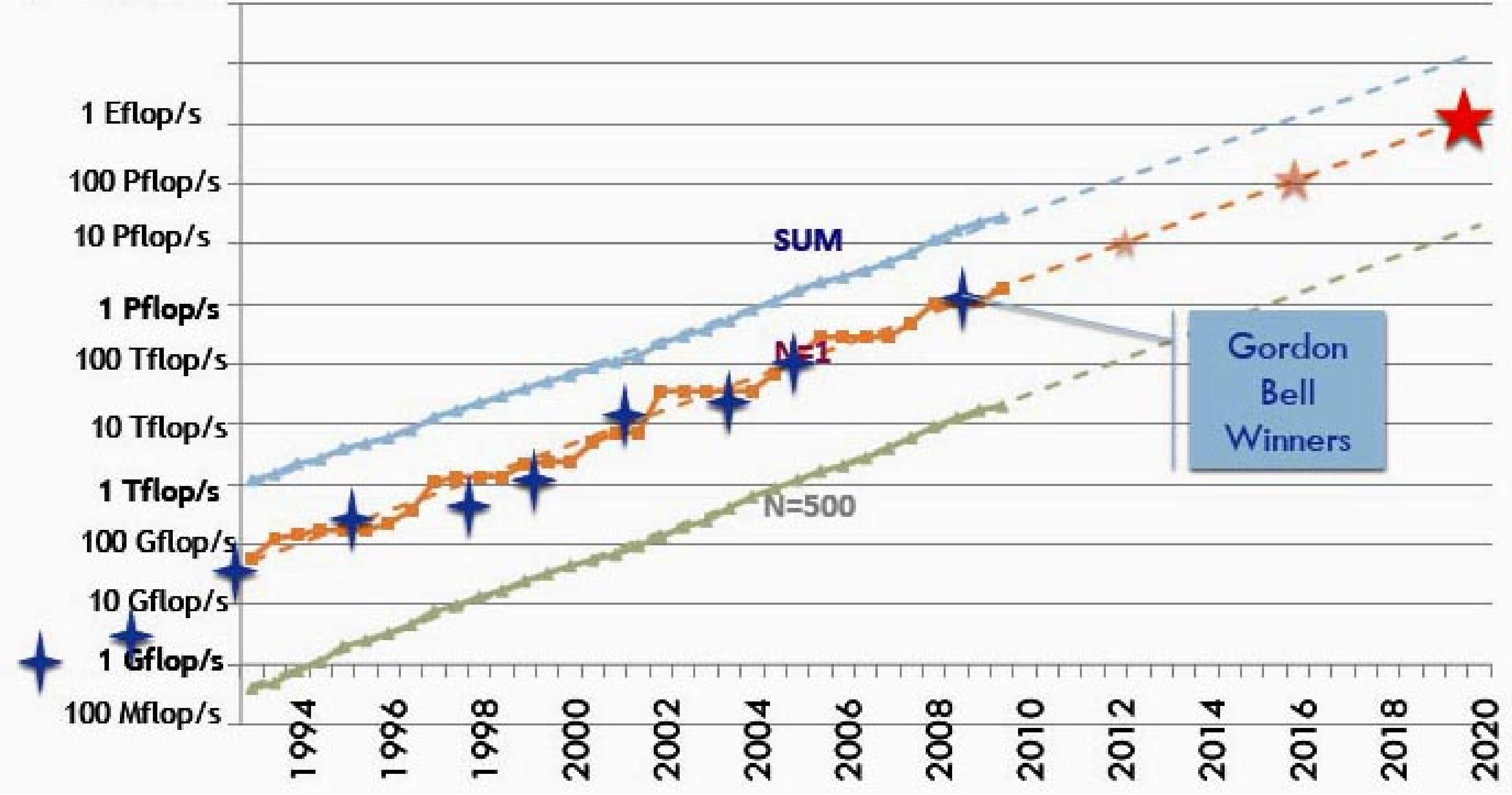
STFC's Scientific Computing Department

STFC Hartree Centre: future developments to meet future challenges

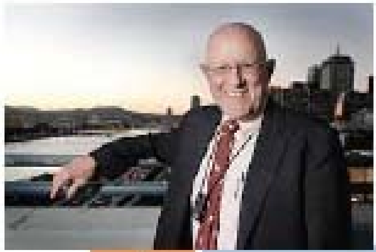
Opportunities for Discussion

TOP500 Global Trends

from Jack Dongarra, IESP



from Jack Dongarra, IESP



Looking at the Gordon Bell Prize

(Recognize outstanding achievement in high-performance computing applications and encourage development of parallel processing)

- 1 GFlop/s; 1988; Cray Y-MP; 8 Processors
  - Static finite element analysis

- 1 TFlop/s; 1998; Cray T3E; 1024 Processors
  - Modeling of metallic magnet atoms, using a variation of the locally self-consistent multiple scattering method.
- 1 PFlop/s; 2008; Cray XT5; 1.5×10^{5} Processors
  - Superconductive materials

- 1 EFlop/s; ~2018; ?; 1×10^{7} Processors (10^{9} threads)

Limits Reached?

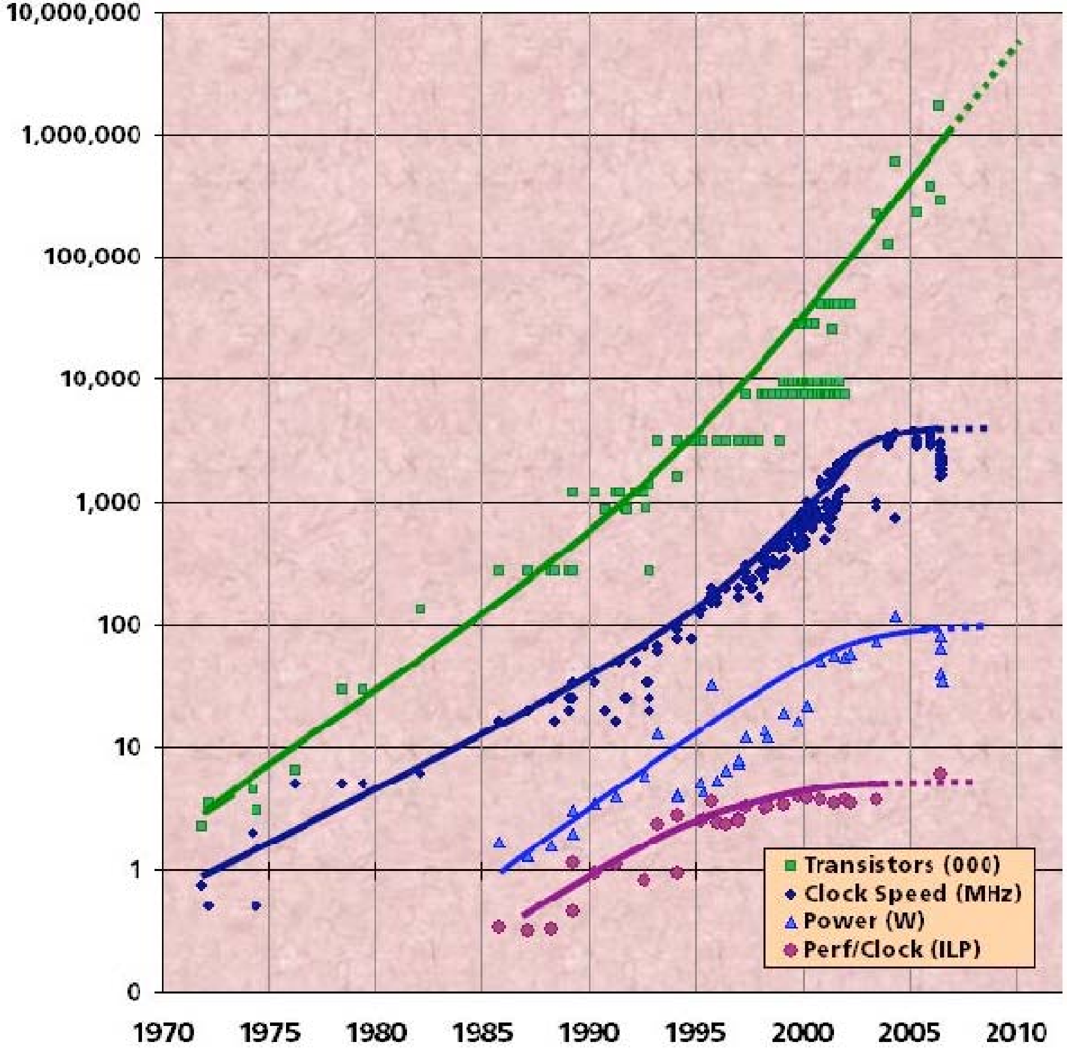
"It would appear that we have reached the limits of what is possible to achieve with computer technology, although one should be careful with such statements, as they tend to sound pretty silly in five years"

John von Neumann, 1949

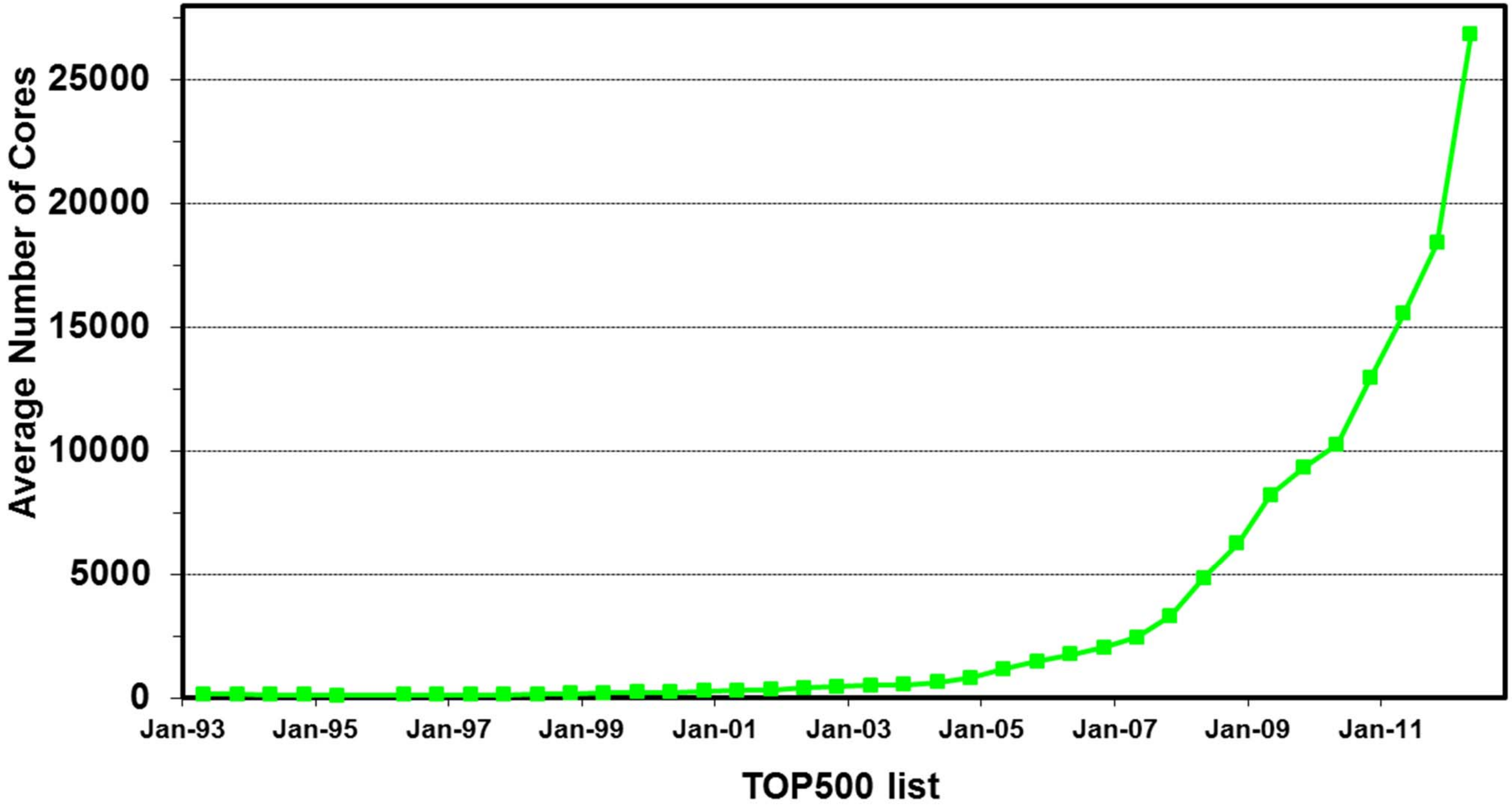
Technology Challenges: The Power Wall

- **Transistor density is still increasing**
- **Clock frequency is not due to power density constraints**
- **Cores per chip is increasing, multi-core CPUs (currently 8-16) and GPUs (500)**
- **Little further scope for instruction level parallelism**

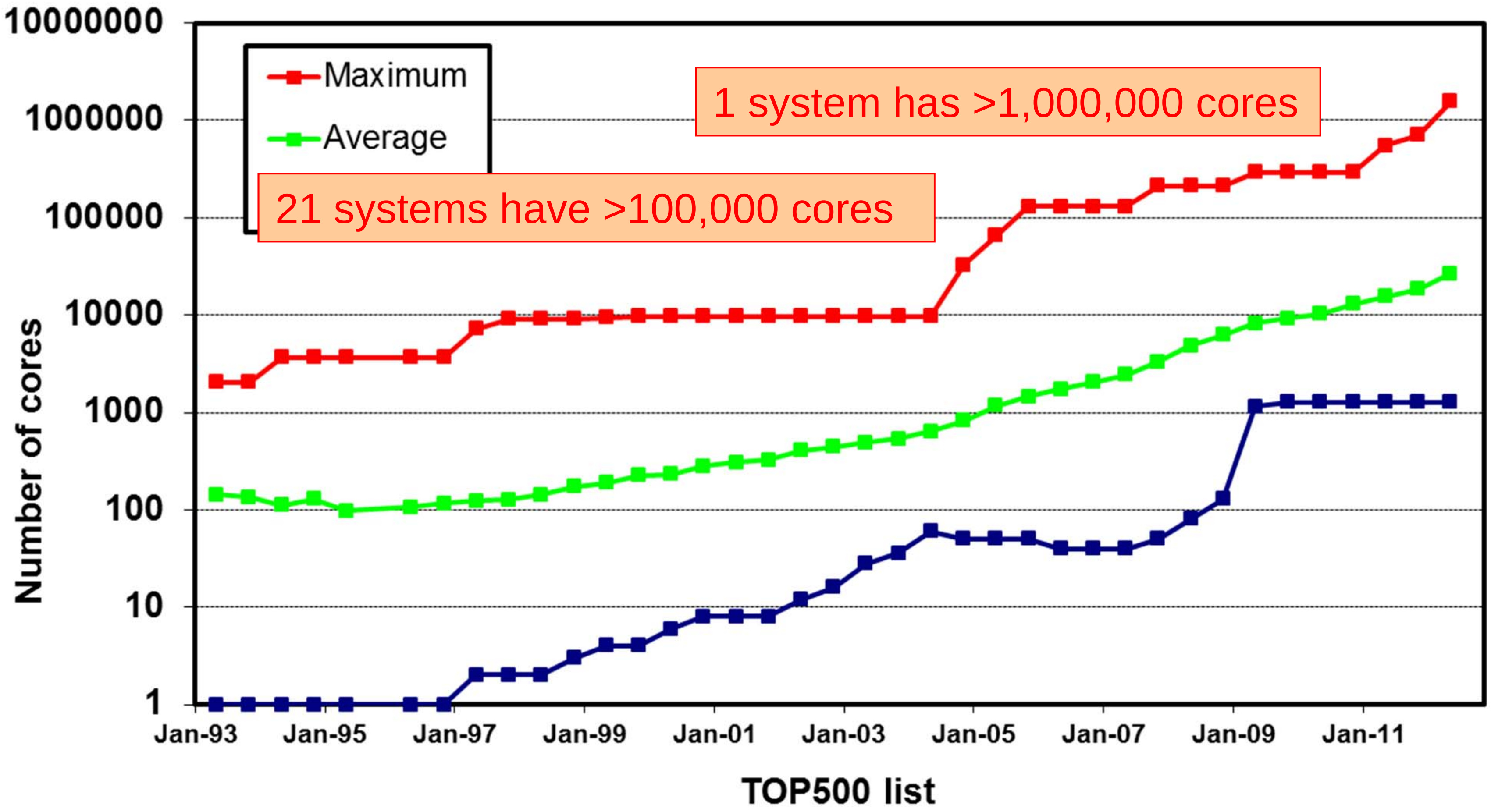
Source: Intel, Microsoft (Sutter) and Stanford (Olukotun, Hammond)



Increase in the number of cores



Increase in the number of cores



Exascale Reports

ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems

Peter Kogge, Editor & Study Lead
Keren Bergman
Shekhar Borkar
Dan Campbell
William Carlson
William Dally
Monty Denneau
Paul Franzon
William Harrod
Kerry Hill
Jon Hiller
Sherman Karp
Stephen Keckler
Dean Klein
Robert Lucas
Mark Richards
Al Scarpelli
Steven Scott
Allan Snavely
Thomas Sterling
R. Stanley Williams
Katherine Yelick

September 28, 2008

This work was sponsored by DARPA IPTO in the
as Program Manager; AFRL contract number FA8
interest of scientific and technical information exc
Government's approval or disapproval of its ideas

NOT

Using Government drawings, specifications, or ot
purpose other than Government procurement does
The fact that the Government formulated or suppl
does not license the holder or any other person or
manufacture, use, or sell any patented invention th

APPROVED FOR PUBLIC RELEASE, DISTRIB

ExaScale Software Study: Software Challenges in Extreme Scale Systems

Saman Amarasinghe
Dan Campbell
William Carlson
Andrew Chien
William Dally
Elmootazbellah Elnohazy
Mary Hall
Robert Harrison
William Harrod
Kerry Hill
Jon Hiller
Sherman Karp
Charles Koelbel
David Koester
Peter Kogge
John Levesque
Daniel Reed
Vivek Sarkar, Editor & Study Lead
Robert Schreiber
Mark Richards
Al Scarpelli
John Shalf
Allan Snavely
Thomas Sterling

September 14, 2009

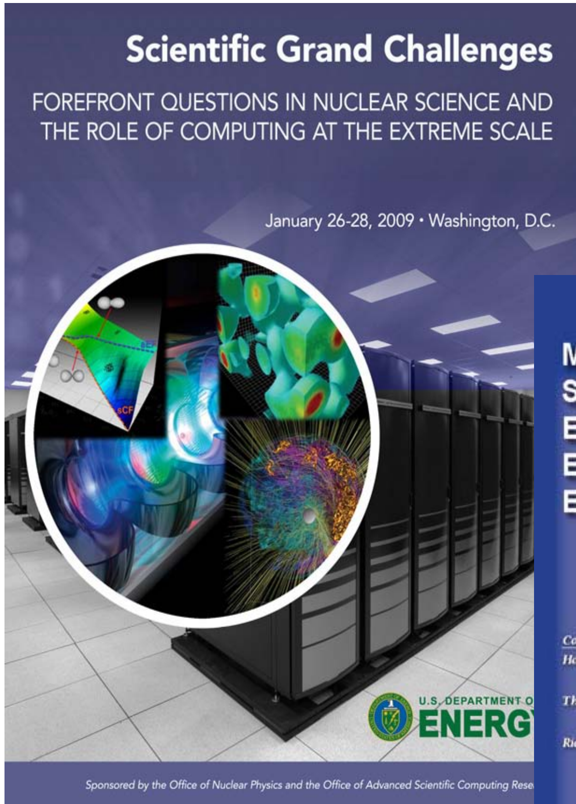
This work was sponsored by DARPA IPTO in the ExaScale Computing Study with Dr. William Harrod as Program Manager; AFRL contract number FA8650-07-C-7724. This report is published in the interest of scientific and technical information exchange and its publication does not constitute the Government's approval or disapproval of its ideas or findings

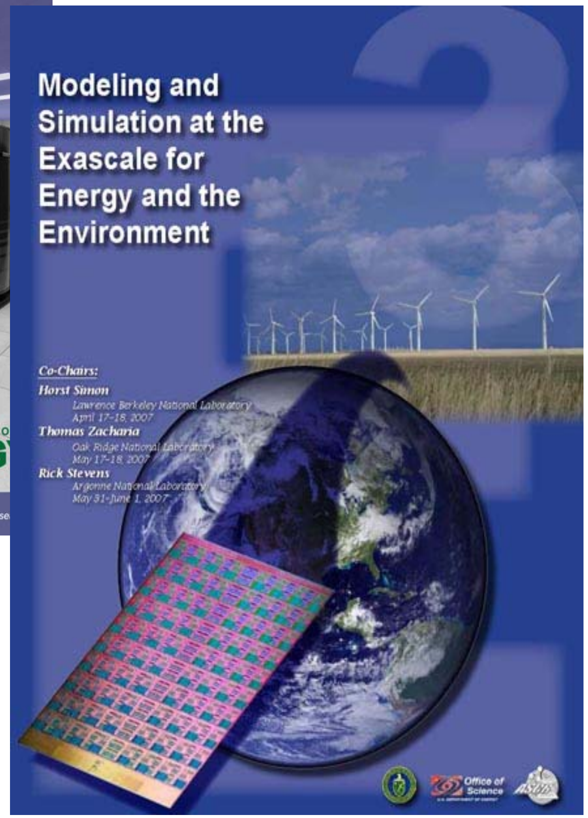
NOTICE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

APPROVED FOR PUBLIC RELEASE, DISTRIBUTION UNLIMITED.

1





Science & Technology Facilities Council

DARPA Exascale Computing Report

- Exascale systems achievable by 2017
- Little increase in clock rate
- 10M-100M processor cores
- Heterogeneous (1:100 full instruction set to lightweight cores)
- 3D packaging with memory close to cores
- Mixed optical plus electrical interconnect
- 10PB-100PB memory
- Fault tolerance with support in hardware and software

ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems

Peter Kogge, Editor & Study Lead
Keren Bergman
Shekhar Borkar
Dan Campbell
William Carlson
William Dally
Monty Denneau
Paul Franzon
William Harrod
Kerry Hill
Jon Hiller
Sherman Karp
Stephen Keckler
Dean Klein
Robert Lucas
Mark Richards
Al Scarpelli
Steven Scott
Allan Snavely
Thomas Sterling
R. Stanley Williams
Katherine Yelick

September 28, 2008

This work was sponsored by DARPA IPTO in the ExaScale Computing Study with Dr. William Harrod as Program Manager; AFRL contract number FA8650-07-C-7724. This report is published in the interest of scientific and technical information exchange and its publication does not constitute the Government's approval or disapproval of its ideas or findings

NOTICE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

APPROVED FOR PUBLIC RELEASE, DISTRIBUTION UNLIMITED.

For total system power of 10-20 MW, need 50-100 GF/W cf. Green500 #1 Nov12 2.5 GF/W

Exascale Software issues have been examined by IESP
www.exascale.org

The Changing Balance of Systems

The huge increase in power efficiency required for Exascale will lead to Petascale systems using only 100kW

System	(HPCx 2002)	ale (Jaguar 2009)	Exascale (DARPA strawman)
# of nodes	160	18,688	223,872
# cores/ node	8	12	742
# of cores	1280	224,256	166,113,024
# racks	40	284	583
Total Mem (TB)	1.28	300	3,580
Disk (TB)	18	600	3,580
Tape (TB)	35	10,000	3,580,000
Peak (Petaflop/s)	0.0067	2.33	1000
Total Power (MW)	0.5	7.0	68
Gflops/W	0.013	0.33	14.73
Bytes/Flop	0.5	0.2	0.0036

100M Threads?

"I know how to make 4 horses pull a cart –
I don't know how to make 1024 chickens do it"

Enrico Clementi, mid1980s?

We are now "pulling carts" with 1,572,864 fruit flies!

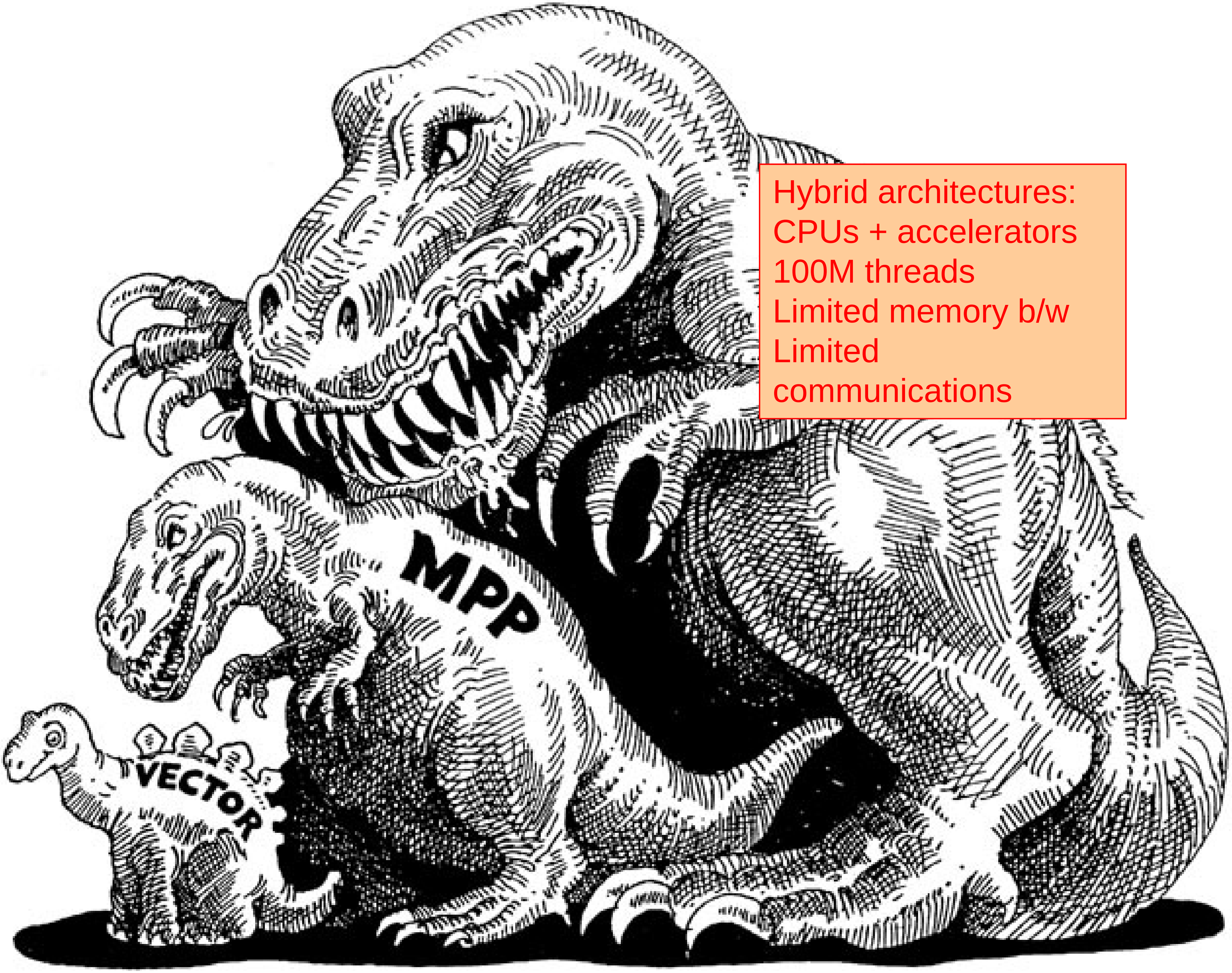
The Exascale software challenge

The Software Challenge of Petascale and Exascale

- Huge growth in no. of threads of execution (10^6, 10^7, ...)
- Restricted memory per thread (~1GB per core and dropping)
- Limited bandwidth to memory (bytes per flop)
- Limited interconnect (bytes per flop)
- Complex software ecosystem (OS, runtime, compilers, libraries, tools etc.)
- Software evolution much slower then hardware
- Software revolution is very expensive

Software issues are constrained by power:

- **Flops are 'free'**
- **Data storage & movement are becoming limited**



Hybrid architectures:
CPUs + accelerators
100M threads
Limited memory b/w
Limited
communications

Multi-Disciplinary Challenges

- Revolutionary changes in hardware are required to reach Exascale performance (within reasonable cost, power etc)
- Greater integration required in the software stack (e.g. OS, compilers, libraries, applications) means co-design
- Scientific opportunities to exploit Exascale require multi-scale, multi-physics applications
- Multi-disciplinary teams are required to address the technological and scientific challenges
  - Application scientists, computational scientists, computer scientists, systems software & hardware technologists

Practical Software Evolution and Revolution

- **Systems will have multi-CPU nodes (~32 cores) with or without accelerators (~200 cores)**
- **Evolutionary development of existing codes**
  - Use hybrid MPI/OpenMP to increase concurrency, while maintaining computational intensity and load balance
  - Use PGAS languages (e.g. Co-Array Fortran) to speed-up communications and overlap with computation
  - OpenAcc and OpenMP standards are emerging targeting accelerators
- **Revolutionary development**
  - New codes and new algorithms targeting many-core memory-light architectures

Overview

Hardware and software:
from evolution to revolution

STFC's Scientific Computing Department

STFC Hartree Centre: future developments to meet future challenges

Opportunities for Discussion

Science & Technology Facilities Council

Scientific Computing Department

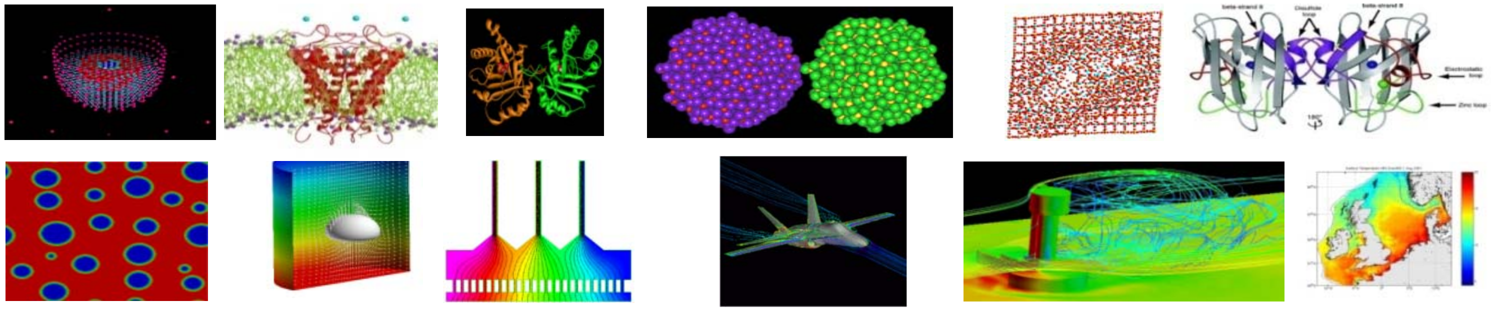
Director: Adrian Wander
Appointed 24th July 2012

Major funded activities

- 160 staff supporting over 7500 users
- Applications development and support
- Compute and data facilities and services
- Research: over 100 publications per annum
- Deliver over 3500 training days per annum
- Systems administration, data services, high-performance computing, numerical analysis & software engineering.

Major science themes and capabilities

- Expertise across the length and time scales from processes occurring inside atoms to environmental modelling



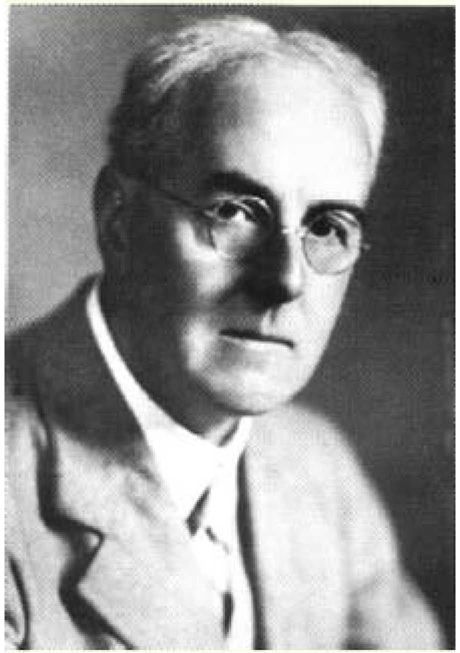
Computational Science
- The Third Pillar of Science

Computational Models can Augment Observations, especially where experiments are difficult:

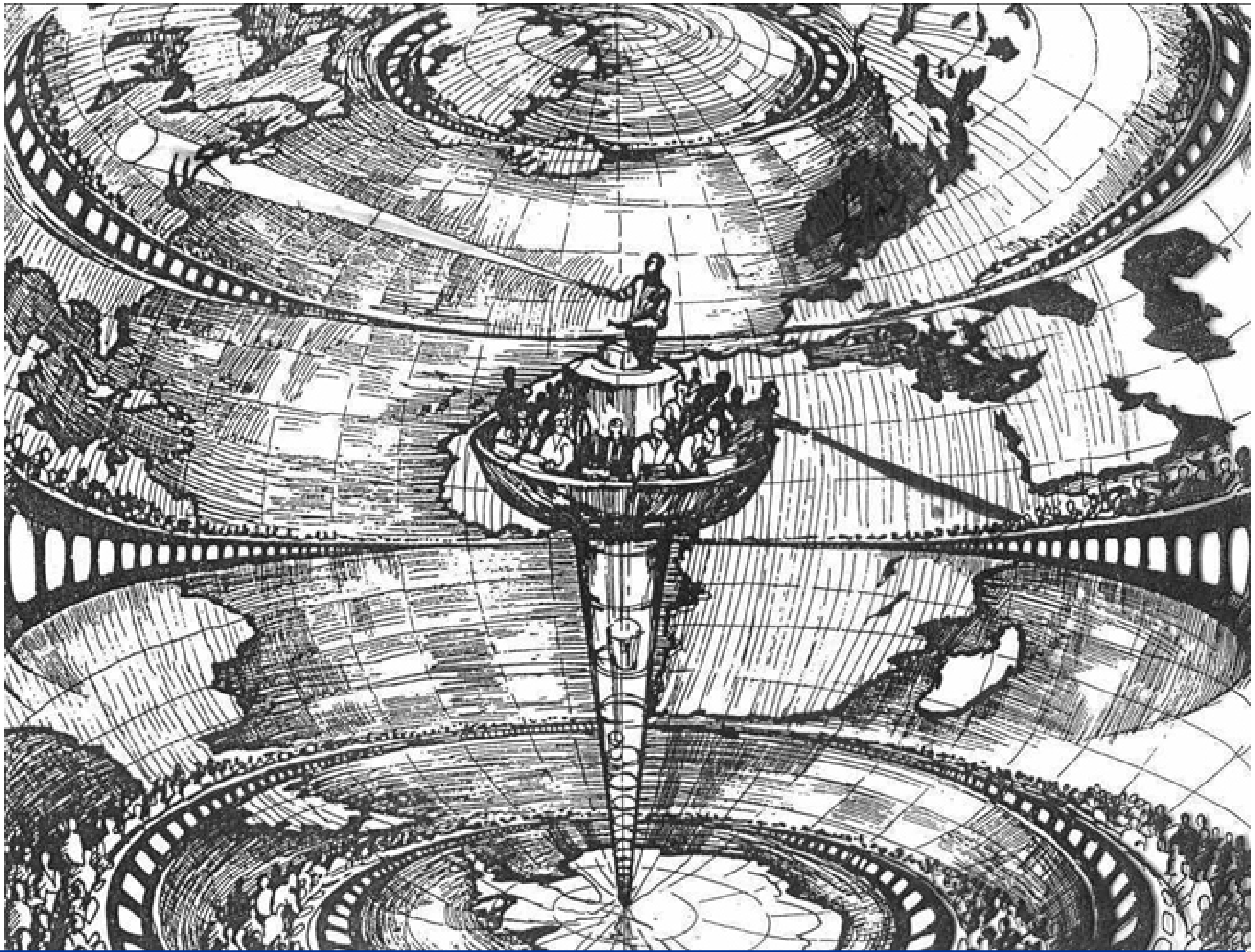
Too big	e.g. ocean currents
Too small	e.g. electronic devices
Too fast	e.g. chemical reactions
Too slow	e.g. tectonic plates
Infrequent	e.g. supernovae, Big Bang
Too dangerous	e.g. hurricanes, volcanoes
Too personal	e.g. human & animal testing
Inaccessible	e.g. earth's core
Too expensive	e.g. crash testing, wind tunnels

Even where experiments are easy, comparison with computer models is a very good test of observations and theory

Lewis Fry Richardson
Pioneer of Numerical Weather Prediction



L.F. Richardson
1881-1953

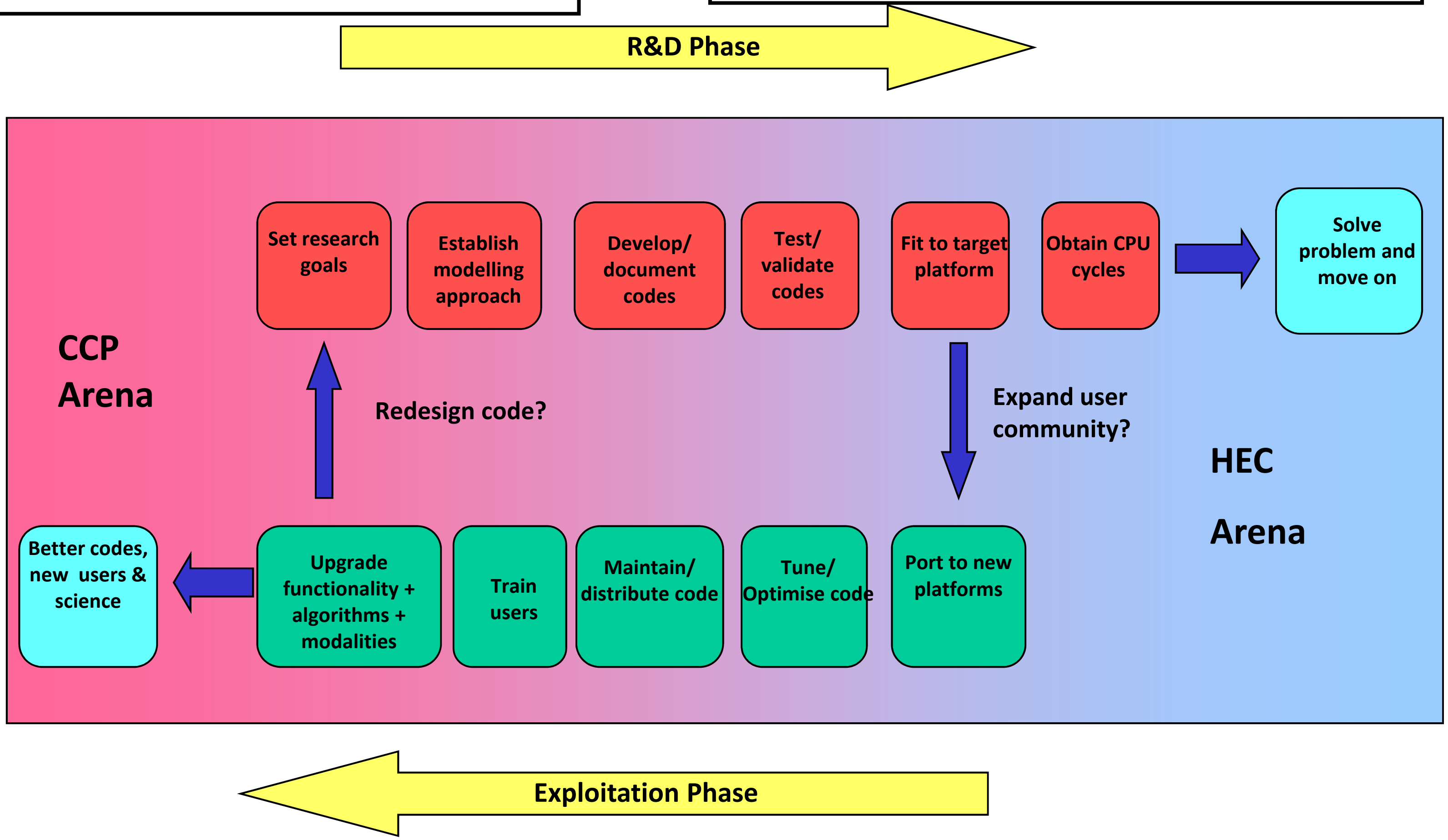


"A myriad computers are at work upon the weather of the part of the map where each sits, but each computer attends only to one equation or part of an equation. The work of each region is coordinated by an official of higher rank." (Richardson, 1922)

The Software Life-Cycle

Competitive research funding – flagship grants, HPC development calls

HECToR dCSE support – optimising implementation for current national services



Collaborative research funding – CSED Service Level Agreement/ Hartree Centre/ Hartree Centre

Opportunities

Political Opportunity	• Demonstrate growth through economic and societal impact from investments in HPC
Business Opportunity	• Engage industry in HPC simulation for competitive advantage • Exploit multi-core
Scientific Opportunity	• Build multi-scale, multi-physics coupled apps • Tackle complex Grand Challenge problems
Technical Opportunity	• Exploit new Petascale and Exascale architectures • Adapt to multi-core and hybrid architectures

Overview

Hardware and software:
from evolution to revolution

STFC's Scientific Computing Department

STFC Hartree Centre: future developments to meet future challenges

Opportunities for Discussion

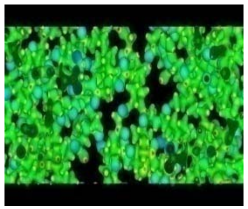
Science & Technology
Facilities Council

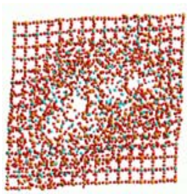
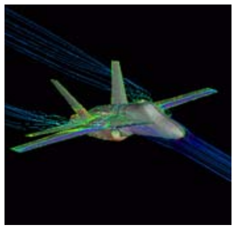
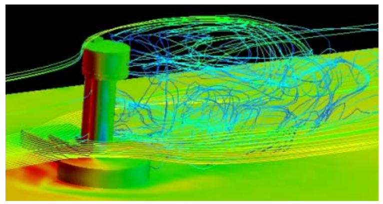
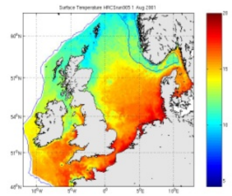
Hartree Centre Mission

Hartree Centre at the STFC Daresbury Science and Innovation Campus will be an International Centre of Excellence for Computational Science and Engineering.

It will bring together academic, government and industry communities and focus on multi-disciplinary, multi-scale, efficient and effective computation.

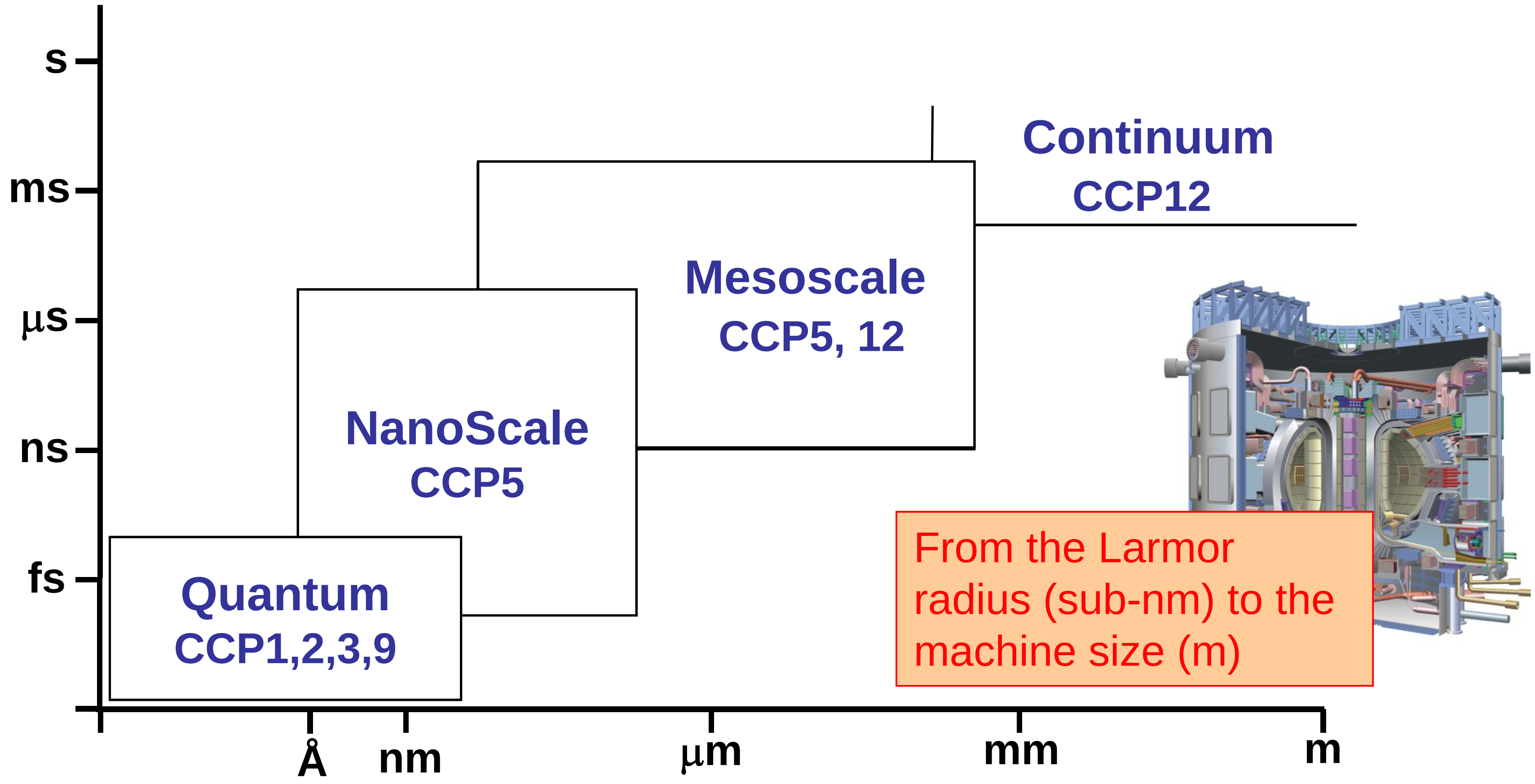
The goal is to provide a step-change in modelling capabilities for strategic themes including energy, life sciences, the environment, materials and security.



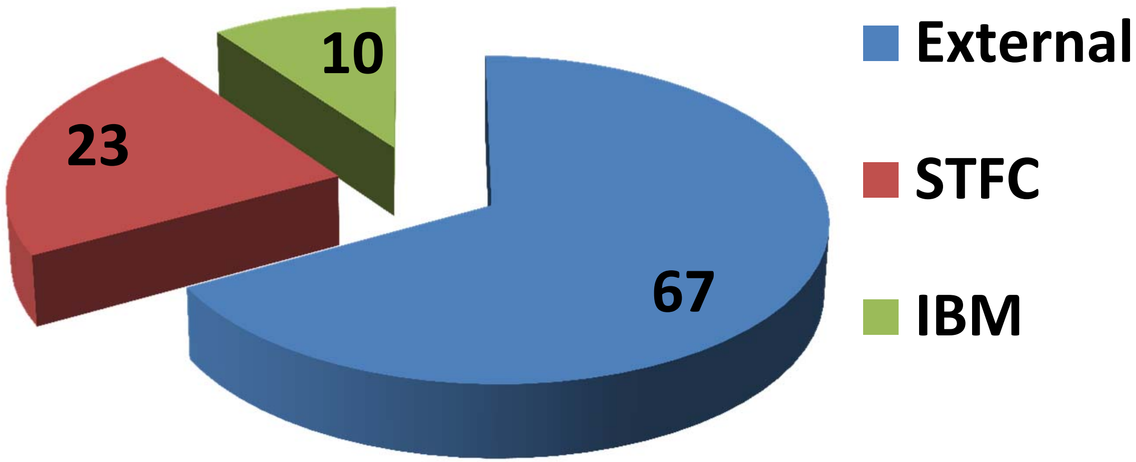





Science & Technology Facilities Council

The Challenge – Integrated Science



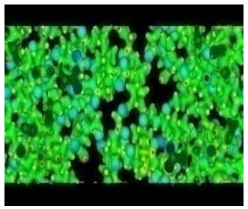
Hartree Centre Typical Project

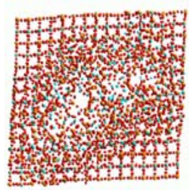
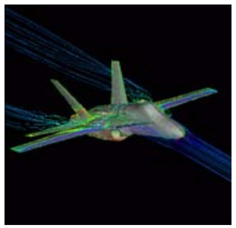
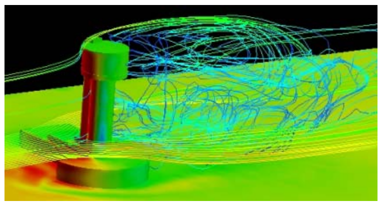
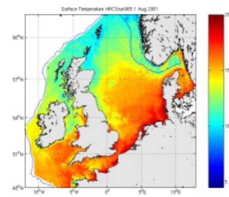


"External":	research councils, industry, government PSREs
"IBM":	also OCF, Intel, nVIDIA, Mellanox, DDN, ScaleMP
"STFC":	includes system time and staff effort

Collaborative Computational Projects

- UK national consortia
- Bring together leading theoretical and experimental research groups in thematic areas
- Tackle large-scale scientific software development
- Maintenance (long-term)
- Distribution
- Training



Collaborative Computational Projects

- CCP4 – Macromolecular crystallography
- CCP5 – Computer simulations of condensed phases
- CCP9 – Computational Electronic Structure of Condensed Matter
- CCP12 – High performance computing in engineering
- CCP-ASEArch – Algorithms and Software for Emerging Architectures
- CCPBioSim – Biomolecular simulation at the life sciences interface
- CCPN – NMR spectroscopy
- CCP-EM – Electron cryo-Microscopy
- CCPi – Tomographic Imaging
- CCPN – NMR
- CCP-NC – NMR Crystallography
- CCPQ – Quantum dynamics in Atomic, Molecular and Optical Physics

Supported by
EPSRC
BBSRC
MRC

How about a CCP in Computational Accelerator Physics ?

Conclusions

- **Revolutionary hardware changes are driving a revolution in software development**
- **Current state-of-the-art is hybrid MPI/OpenMP with the introduction of OpenAcc regions for accelerators**
- **SCD / Hartree Centre has a critical mass of ~160 staff to drive forward software developments**
- **There are many ways of collaborating:**
  - **Joint projects**
  - **CCPs and other high-end computing consortia**
  - **Hartree centre projects**

Come and talk to us!

If you have been … … thank you for listening

Mike Ashworth mike.ashworth@stfc.ac.uk

http://www.cse.stfc.ac.uk/