UKHEC

UK HEC Collaboration

Martyn Guest and Robert Allan
Computational Science and Engineering Department, Daresbury Laboratory

Alan Simpson and David Henty
Edinburgh Parallel Computing Centre

Terry Hewitt
Manchester Research Centre for Computational Science

UKHEC

Background

In January 2000 EPSRC funded a 3 year core activity to track and disseminate information on international activities in computer architectures, software and programming tools and to promote good programming practice for the HEC community via workshops, seminars and mentoring.

Information is provided in a Newsletter and via the Web site:

http://www.ukhec.ac.uk

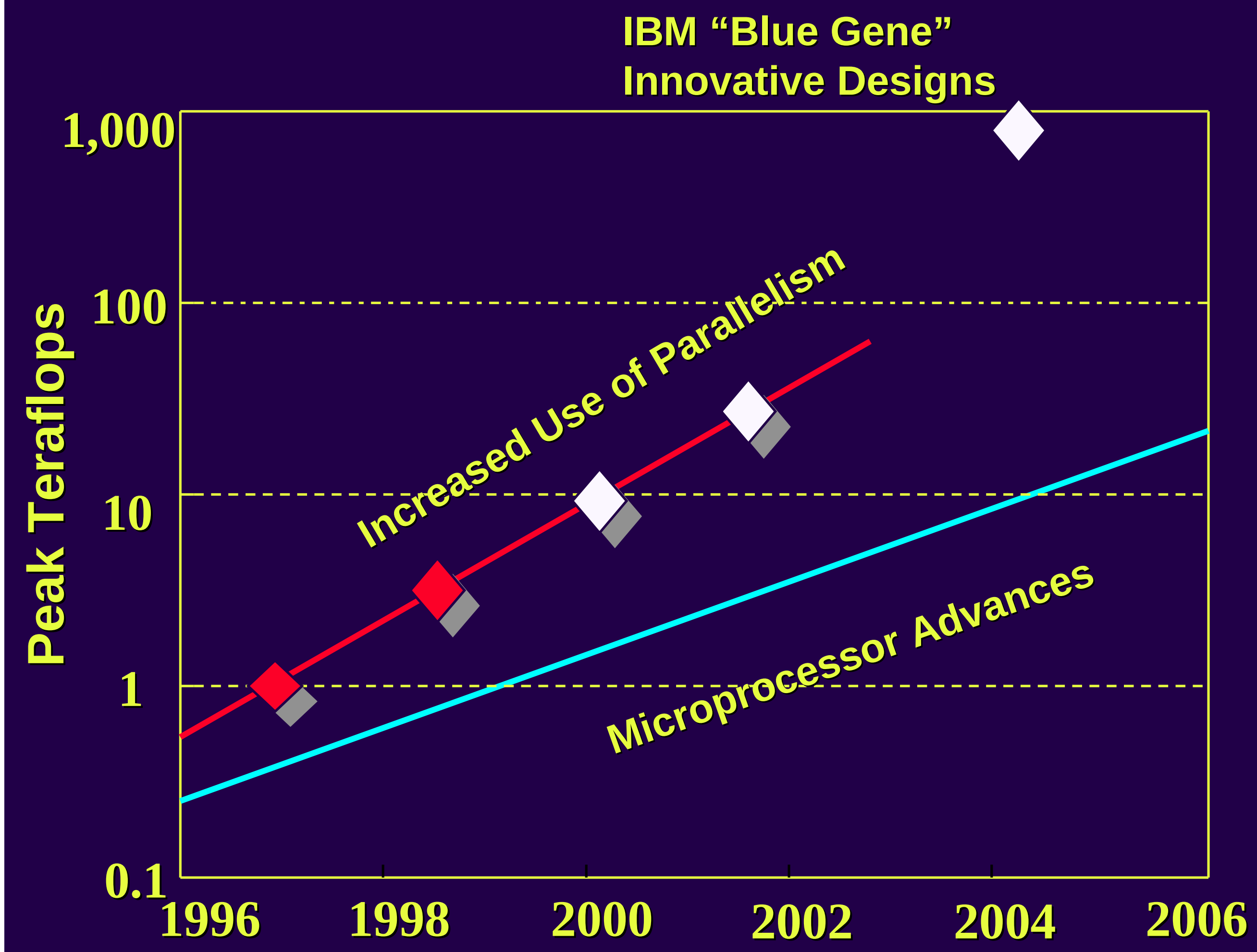
The "UKHEC Collaboration" involves staff from

- CLRC's Daresbury Laboratory (2 FTEs),
- Edinburgh Parallel Computing Centre (EPCC, 2 FTEs), and
- Manchester Research Centre for Computational Science (MRCCS, 1 FTE)

A strategic collaboration between the main UK centres offering nation-wide academic computing support.

UKHEC

Dramatic Advances in Computing



MICROPROCESSORS

2 X increase in microprocessor speeds every 18-24 months ("Moores Law")

PARALLELISM

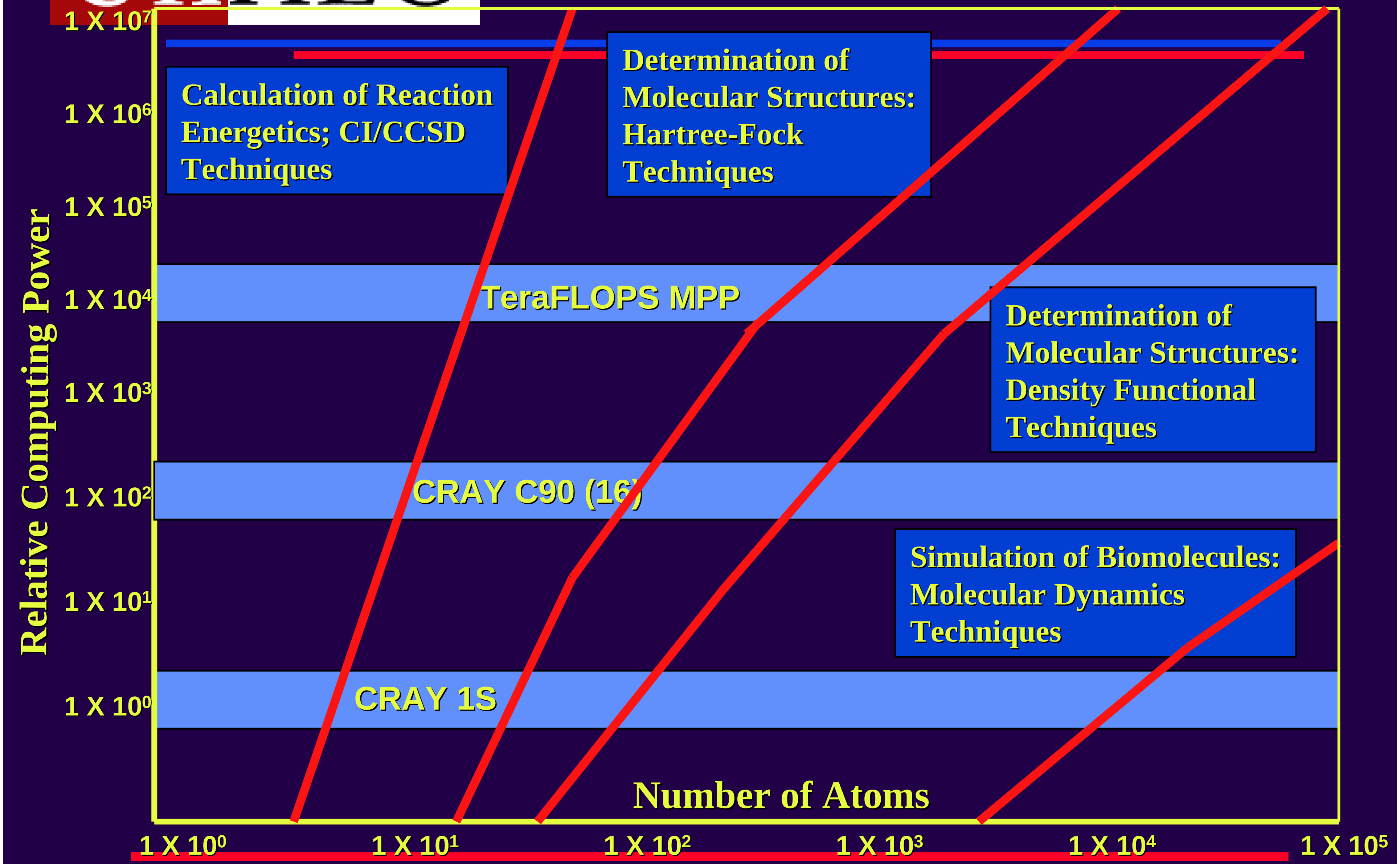
More and More processors being used on single problem

INNOVATIVE DESIGNS

Processors-in- memory

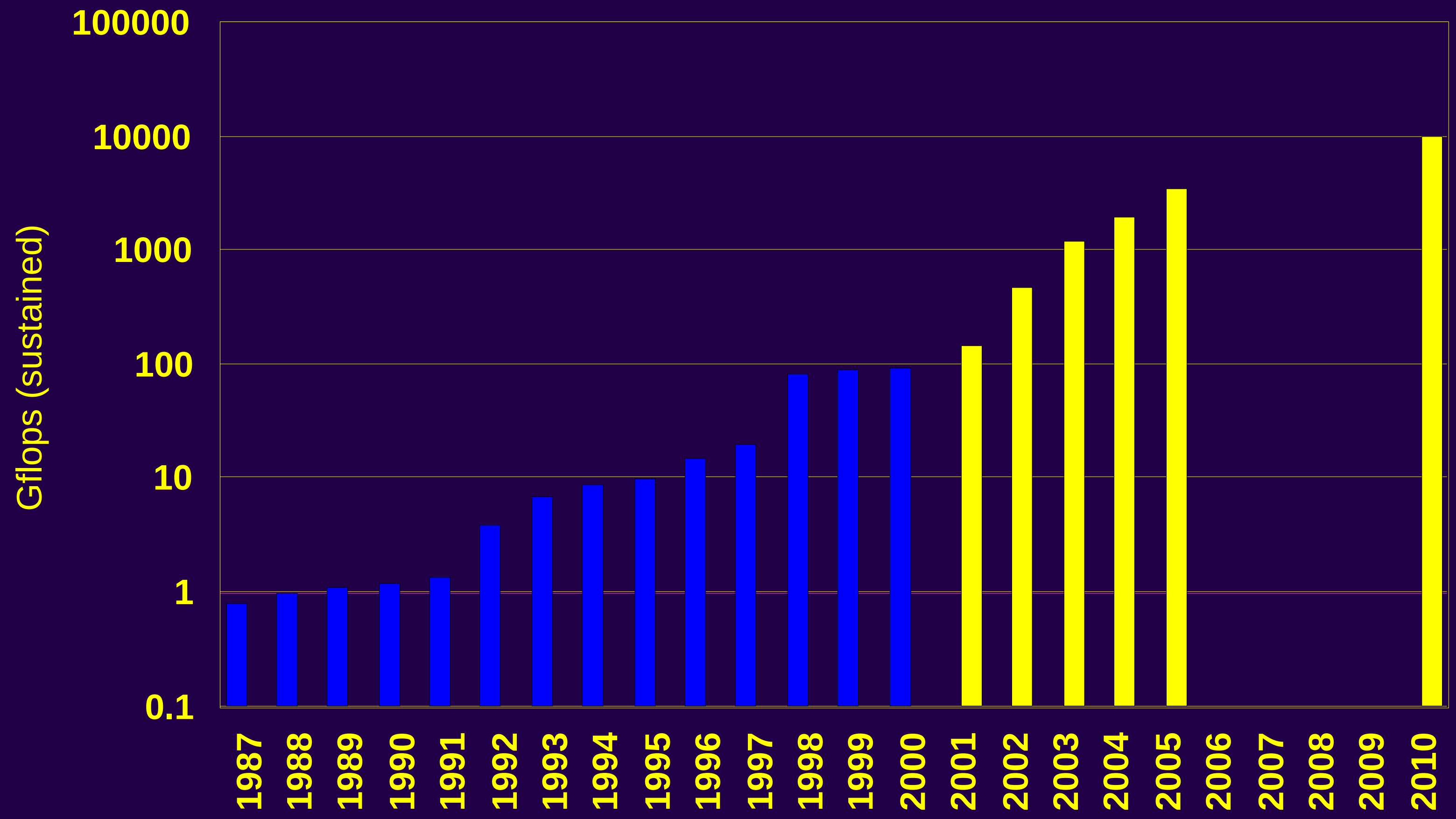
UKHEC

Scaling of Molecular Computations



Performance Trends

(<u>Sustained</u> performance for atmospheric and ocean models)
Collated set of measured (blue) and required (yellow) performance capabilities
(ECMWF, NCAR, UK Met, DWD, NCEP, etc.)



UKHEC

HPC and Scientific Opportunity in the Electronic Era

Report of the Trends and Opportunities Panel (P.J. Durham, March 2001)

- **From Molecules to Matter – Electronic Structure**
  - **The Structure of Proteins**
  - **Materials Design**
  - **Heterogeneous Catalysis**
  - **Environmental and Atmospheric Science**
  - **Rational Drug Design**
  - **New areas which will require exceptional computer resources:**
    - **Bridging Length and Time Scales**
    - **Quantum Computation**
    - **Thermodynamics**
- **From Molecules to Cells and beyond – Computational Biology**
- **From Eddies to Aircraft - Fluid Dynamics**
- **From Oceans to the Earth - Environmental Modelling**
- **From the Earth to the Solar System – Solar Plasma Physics**

"Simulation of whole systems, and not just system components"

UKHEC

Scientific Software Infrastructure

One of the Major Software Challenges

Peak Performance is skyrocketing

- In past 10 years, peak performance has increased 100x; in next 5+ years, it will increase 1000x

but ...

- Efficiency has declined from 40-50% on the vector supercomputers of 1990s to as little as 5-10% on parallel supercomputers of today and may decrease further on future machines

Research challenge is software

- Scientific codes to model and simulate physical processes and systems
- Computing and mathematics software to enable use of advanced computers for scientific applications
- Continuing challenge as computer architectures undergo fundamental changes: *Algorithms that scale to thousands-millions processors*

DOE SCIDAC Hardware Infrastructure

A Future Model for the UK ?

- **Flagship Computing Facility**
  - **To provide robust, high-end computing resources for all research programs**
- **Topical Computing Facilities**
  - **To provide the most effective and efficient computing resources for a selected set of scientific applications**
  - **To serve as a focal point for a scientific research community as it adapts to new computing technologies**
- **Experimental Computing Facilities**
  - **To assess new computing technologies for scientific applications**

UKHEC

Why Topical Facilities ?

Variation in Scientific Application Needs*

Code	Application	Time (TFLOPS-HRS)	Memory (TBYTES)	Storage (TBYTES)	Node I/O (MBYTES/S)
Cactus	Astrophysics	300	1.8	20	5
ARPS	Weather	25	0.25	16	18
MILC	Particle Physics	10,000	0.2	1	3
PPM	Turbulent Flow	500	0.5	54	6
PUPI	Liquids	150	0.1	0.2	3
ASPCG	Fluid Dynamics	5,000	0.5	50	3
ENZO	Galaxies	1,000	0.9	10	12
	Variation	**400x**	**18x**	**100x**	**6x**

** From "High-level Application Resource Characterization," NSF/PACI National Computational Science Alliance, May 2000.*

UKHEC

Cluster Computing:
Where Do Clusters Fit?

- Gather (unused) resources
- System SW manages resources
- System SW adds value
- 10% - 20% overhead is OK
- Resources drive applications
- Time to completion is not critical
- Time-shared
- Commercial: PopularPower, United Devices, Centrata, ProcessTree, Applied Meta, etc.

- Bounded set of resources
- Apps grow to consume all cycles
- Application manages resources
- System SW gets in the way
- 5% overhead is maximum
- Apps drive purchase of equipment
- Real-time constraints
- Space-shared

UKHEC

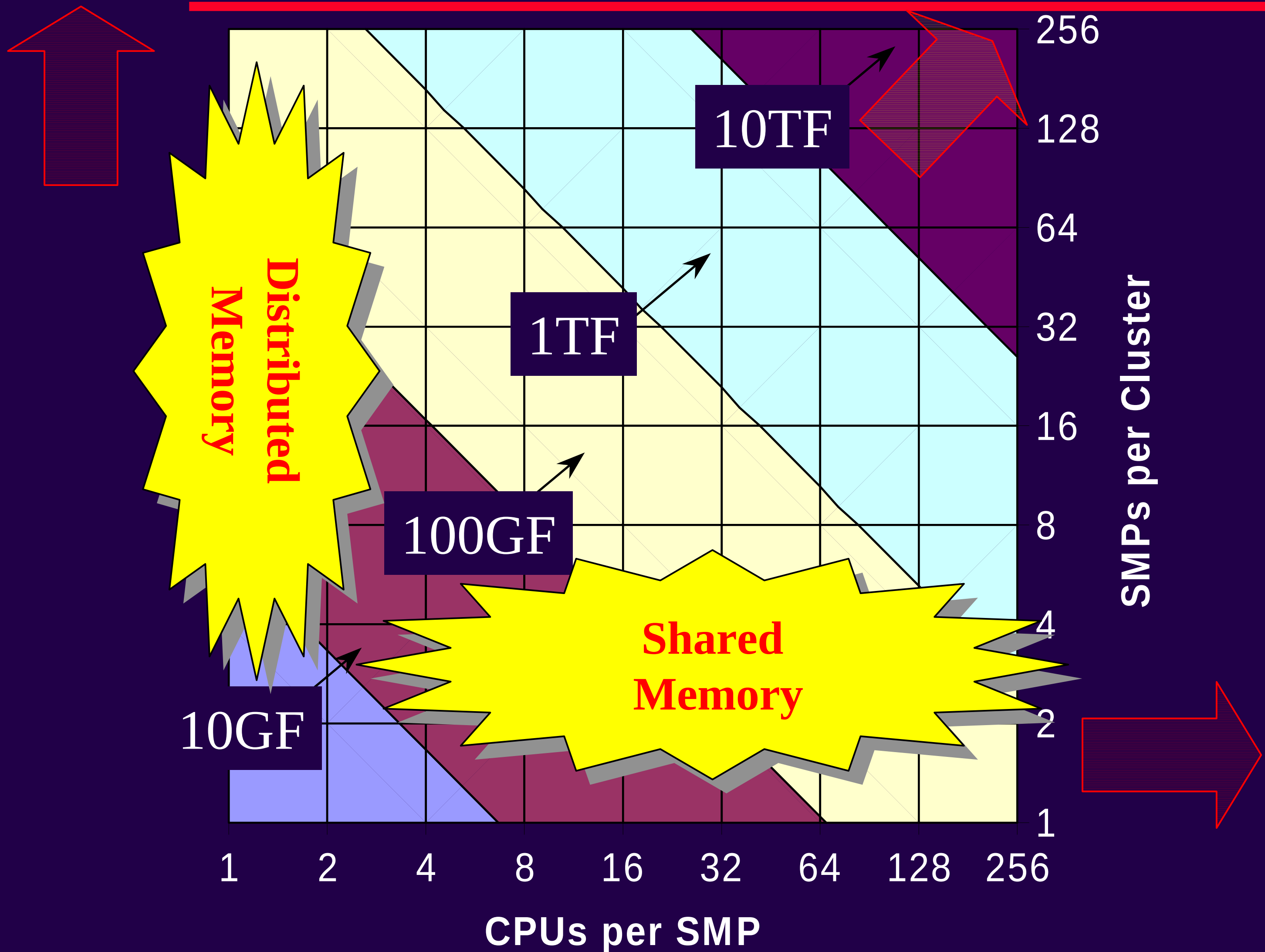
Key Topics

- ***Hardware: SMP/DMM (ASCI) architectures, clusters (Beowulf)***
- **Software Development and QA Tools: faster development, maintenance and exploitation**
- ***Languages (Java, C++, Fortran90): ease of use and performance***
- **Optimisation: need highest performance - can tools help ?**
- **Visualisation and VR: demonstrate VR capabilities for scientists**
- **Data Management: demonstrate fast storage and access for science apps.**
- ***Grid computing environment: evaluation and use of Globus, coordination of e-Science activities between centres***
- **Standards: portability and longer code lifetime**
- ***OpenMP/ MPI Programming: optimise for heterogeneous architectures***

Six highlights will now be given

UKHEC

Highlight 1:
Parallel Computer Hardware Evaluation



UKHEC

Highlight 2:
Application Development

More robust, portable and easily maintained applications

- **Revision control systems**
- **Makefiles and the compilation process**
- **Debugging, profiling and browsing tools, e.g. TotalView, Vampir**
- **Software testing**
- **Portability, standards, ease of use and maintenance**
- **Software design and QA tools - e.g. UML, Assure, Forecheck, VAST**
- **Object-oriented programming - C++ and Java vs. Fortran95**
- **Estimation of requirements and risk management**

PDRA workshops

Technical reports

Training course: *"Practical Software Development for Scientists and Engineers",* Edinburgh 27/2-1/3/2001

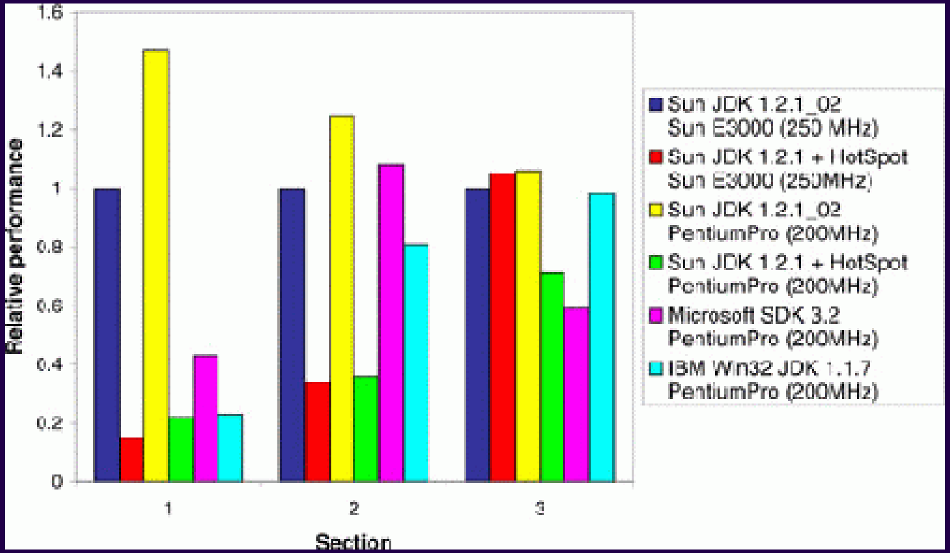
Highlight 3:
Languages and Paradigms

Choosing the best way to write codes

- **Java: Java Grande and JASPA benchmarks**
- **C++**
- **Fortran 95**
- **OpenMP and mixed mode**
- **MPI-2, SHMEM, Lapi, GATools**

Technical Reports

Workshops



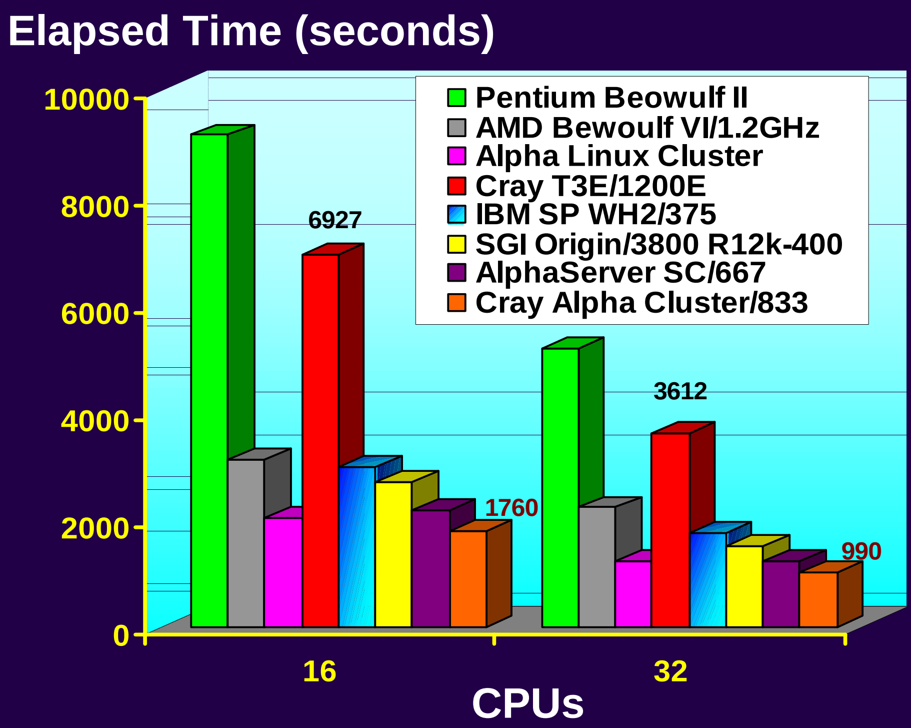
Relative performance of selected Java execution environments on the three sections of the Java Grande benchmark suite.

UKHEC

Highlight 4:
Optimisation and Performance

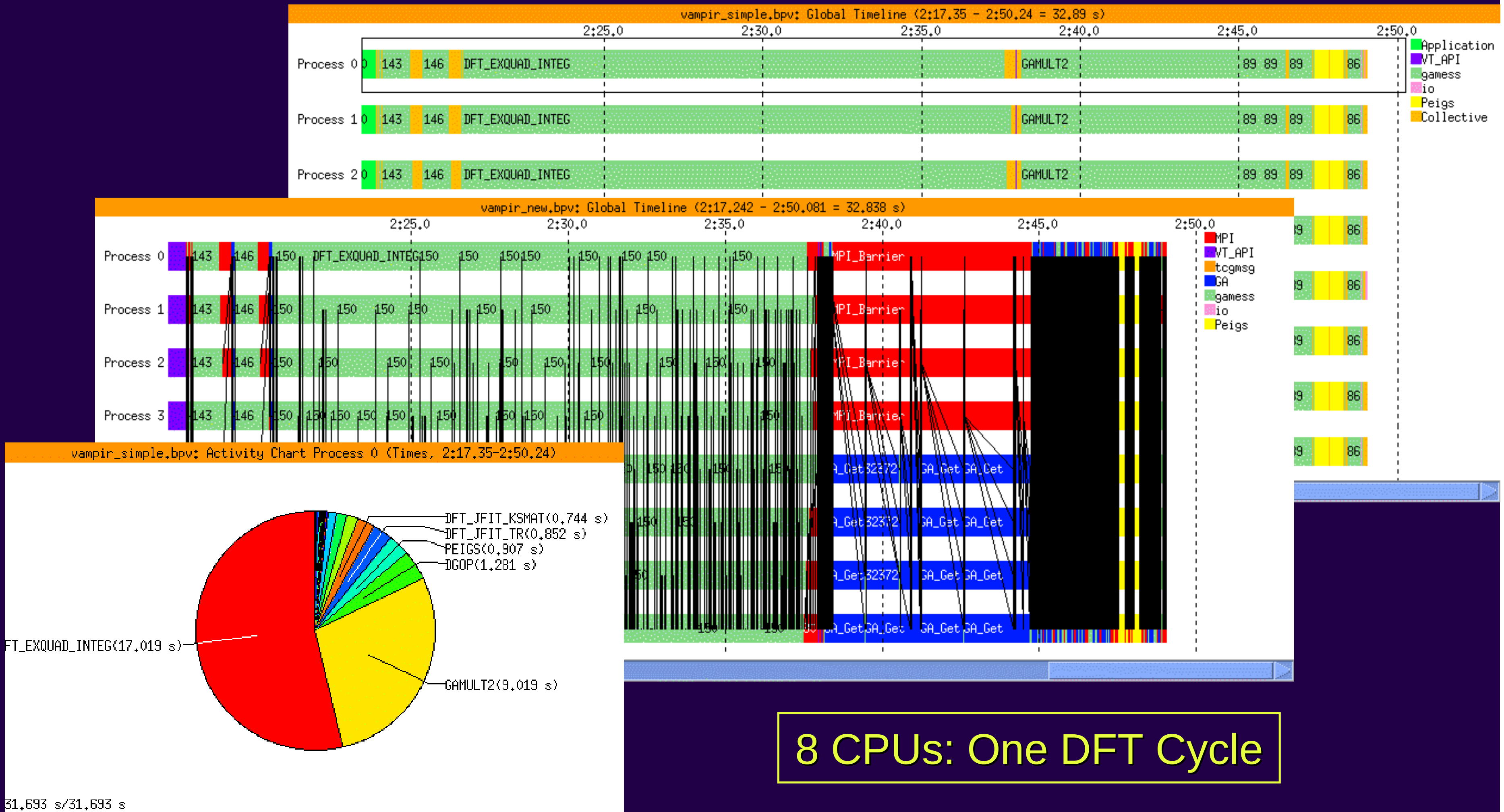
Application Profiling and benchmarking to improve performance

- **Architectural constraints**
  - **SMP vs. vector vs. MPP**
- **Coding constraints**
  - **language, paradigm (MPI vs. OpenMP)**
- **Cost effectiveness:**
  - **capability vs. capacity**
  - **proprietary vs. commodity**
- **Profiling tools**
  - **Vampir, GuideView, Paraver etc.**
- **Performance Modelling**



Extended Performance Analysis: VAMPIR

GAMESS-UK / Zeolite fragment



8 CPUs: One DFT Cycle

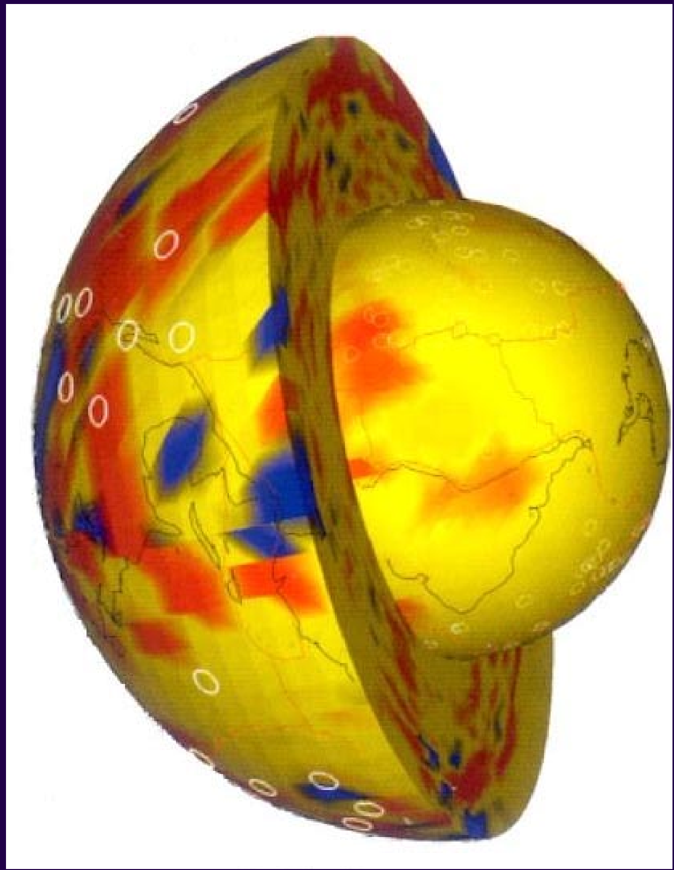
UKHEC

Highlight 5:
Visualisation for Scientific Computing

Analysis of scientific data and new insight

- **Immersive visual environments**
- **multi-dimensional representations**
- **comparing simulated and real data**
- **data analysis tools**
- **interactive steering**

Technical reports and demonstrations, e.g. Terra project
Training course: *"Visualisation and Virtual Reality for Scientific Applications"*, Manchester 20-21/9/2001



Earth's Mantle and the Terra Project

UKHEC

Highlight 6:
Computational Grid

Enable flexible and easy access to local and national facilities

- Single point of login
- Internet security standards and authentication
- remote access to local and national facilities
- integration of compute resources, data centres and experimental instruments
- development of problem solving environments

Middleware: Globus; Condor; UNICORE; SRB

Message-passing: MPICH-G; PACX-MPI

Resource Allocation Managers: LoadLeveller, PBS, LSF, NQE, etc.

"Collaboratory" and VR Software: COVISE

- International meta-computing demonstrations, e.g. SC'2000
- Consultancy and technical input into many e-Science projects
- Partners all established e-Science centres and collaborate in UK Grid Support Centre and Access Grid
- Training course under development for late 2001

UKHEC

The Future:
HPC Systems - Access and Efficiency

HPC(X) and UK e-Science Environment

1) Continue and Extend Technology watch brief
   - **evaluate new methodology, architectures and programming standards (working with tools fora, international research groups and vendors)**
   - **target new communities (e.g. other RCs)**

2) The Computational Grid and HEC Applications
   - ***outreach: extend Grid-based applications to <u>High End</u> facilities***
   - ***investigate distributed computing algorithms to provide highly scalable "loosely coupled" methods for HEC systems***

3) Work with Application Developers
   - **extend dissemination brief with *more focus on applications* by working directly with code developers to evaluate new techniques**

Grid computing, scalable algorithms, tools, data management, software development, languages, visualisation and VR.

UKHEC

Projected Deliverables

Technology Watch

- Produce reports and organise seminars (annual seminar)
- Technology demonstrations for JREI etc. - extend role of DL MEW
- Closer interaction with H/W and S/W companies and research groups

The Computational Grid and HEC Applications

- Coordinate porting from desktop to top-end facilities via the Grid, extend role of HPCI, promote use of HPC facilities
- Prototype meta-computing application demonstrators
- Virtual Reality Centre - provide access to VR equipment and tools

Work with Application Developers

- *work directly with code developers to evaluate new techniques in practice*
- encourage uptake of design, testing, debugging and profiling tools
- Provide support for science-driven PDRAs (CCPs + HPC(X) + CSAR)
- Extend role to other organisations - AWE, Met. Office, ECMWF, other sites ..
- Coordinate training via collaborative working and AccessGrid activities