# The Future of Performance Engineering in HPC

**Jeffrey S. Vetter**

**Oak Ridge National Lab and Georgia Tech**

**SciDAC Performance Engineering Research Institute**

**INCITE Performance Evaluation and Analysis Consortium (PEAC) End Station**

http://ft.ornl.gov

vetter@ornl.gov

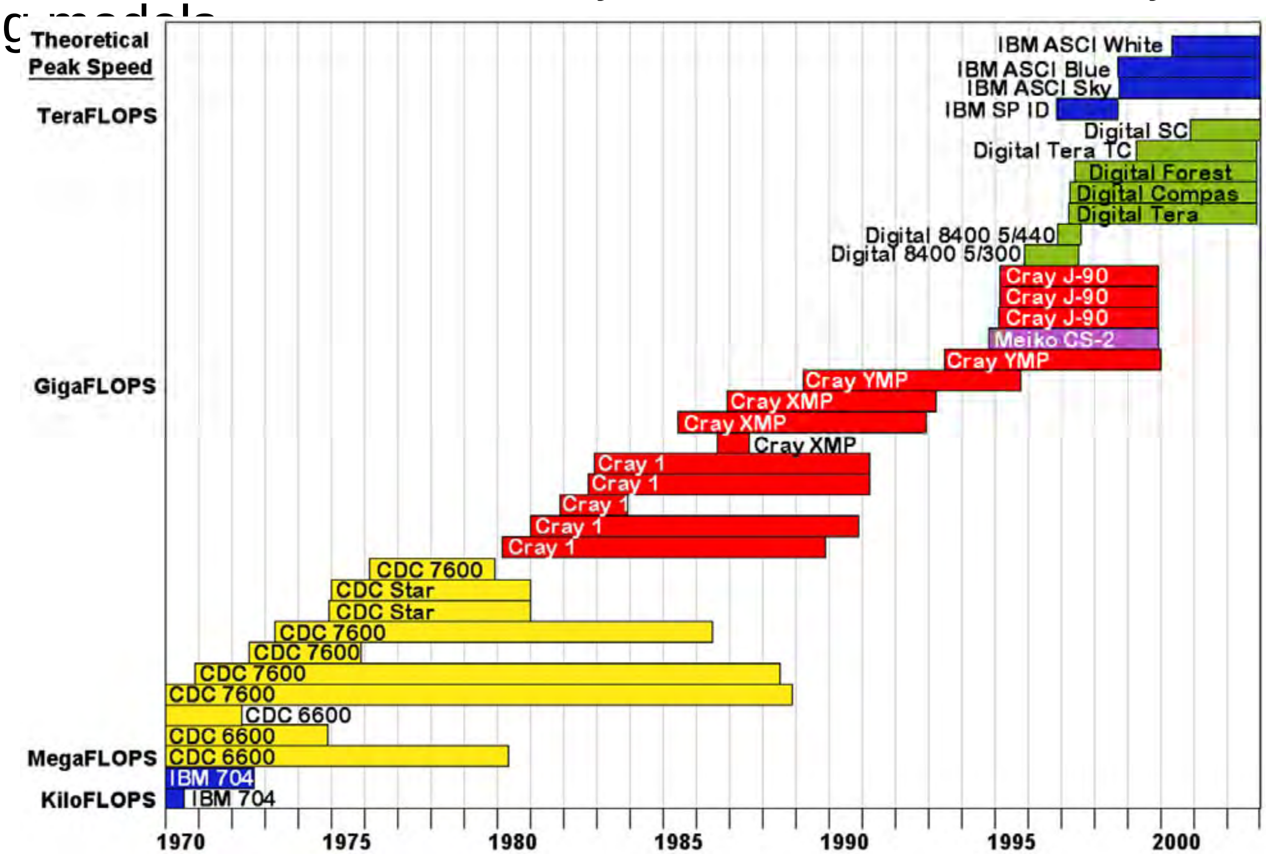
# Talk Highlights

- Significant trends emerging in HPC
  - Architectural complexity: multicore, heterogeneity, power mgt
  - Scale
  - Application complexity: multiphase, multiscale, multiresoution
- Taken together, these trends can create a widening gap between expected and realized performance
- Performance engineering is critical to address this gap
  - Measurement
  - Prediction
  - Optimization
- <u>Some</u> performance engineering solutions that help to close this gap
  - Engagement
    - Frequent interaction between applications teams and performance experts
  - Tools
    - Instrumentation, collection, and analysis tools for measurement
  - Automatic optimization
    - Static and dynamic optimization of applications and libraries
    - Integrate of performance engineering into application/system lifecycle
  - Feedback to architects and system software designers

# Years of Prosperity

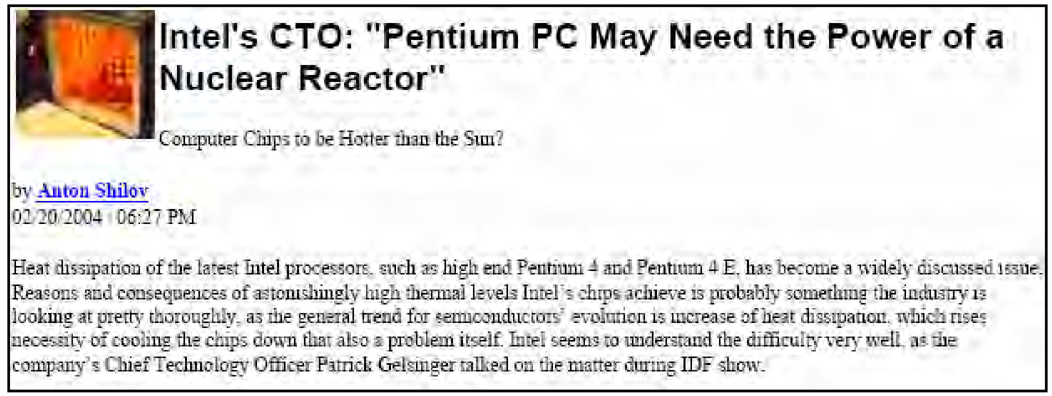
- Increasing large-scale parallelism
- Increasing number of transistors
- Increasing clock speed
- Stable programming models and languages

LLNL System Lifetimes. M. McCoy



# ‘New’ constraints for architectures

- Power
- Heat / thermal envelope
- Signaling
- Packaging
- Instruction level parallelism
- Memory, I/O, interconnect latency and bandwidth
- Market trends favor ‘good enough’ computing – *Economist*



**Intel's CTO: "Pentium PC May Need the Power of a Nuclear Reactor"**

Computer Chips to be Hotter than the Sun?

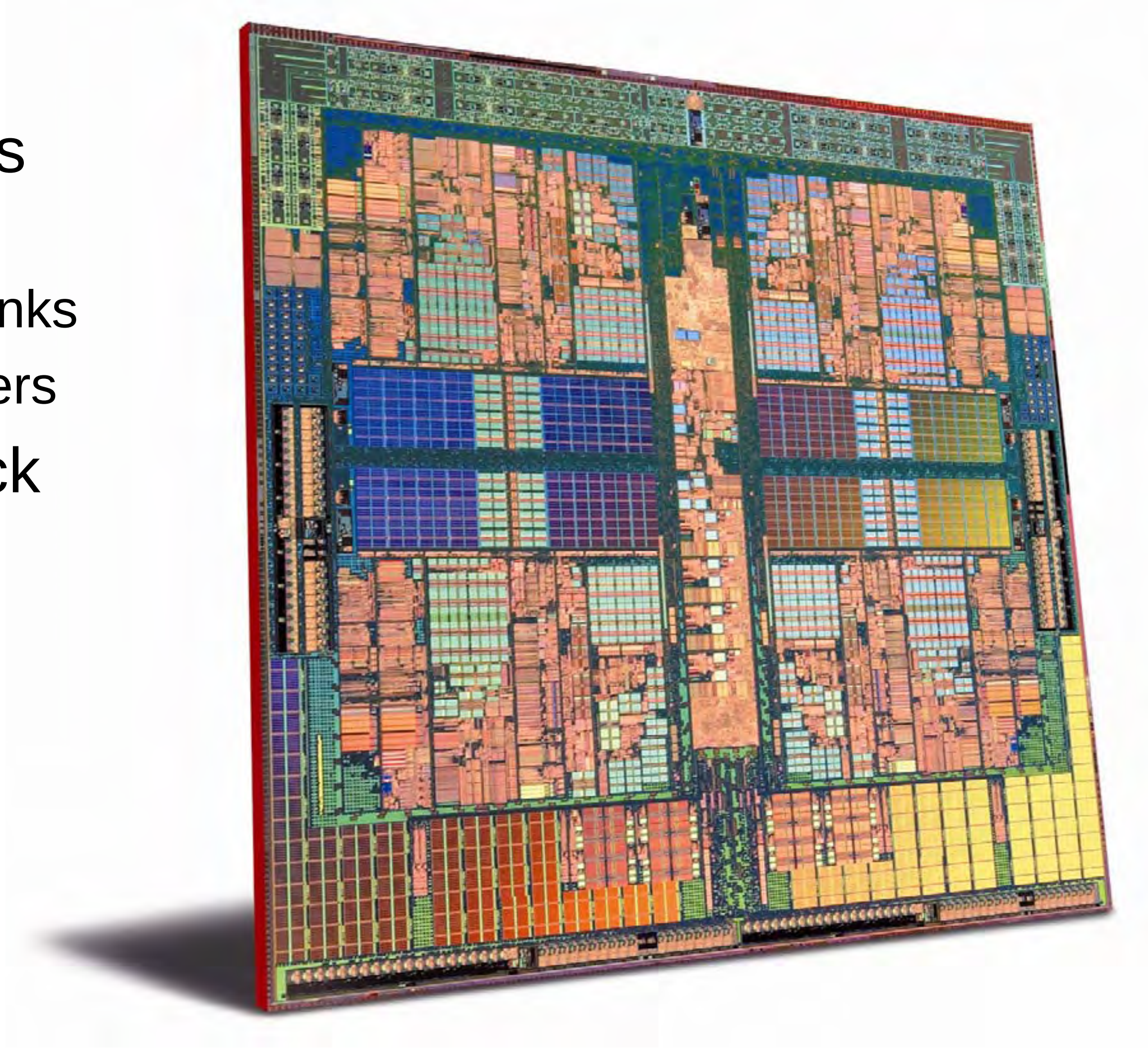
by Anton Shilov
02/20/2004 | 06:27 PM

Heat dissipation of the latest Intel processors, such as high end Pentium 4 and Pentium 4 E, has become a widely discussed issue. Reasons and consequences of astonishingly high thermal levels Intel's chips achieve is probably something the industry is looking at pretty thoroughly, as the general trend for semiconductors' evolution is increase of heat dissipation, which rises necessity of cooling the chips down that also a problem itself. Intel seems to understand the difficulty very well, as the company's Chief Technology Officer Patrick Gelsinger talked on the matter during IDF show.

# Architectural Complexity – Multicore

## AMD quad-core due on Sept 10

- 4 cores
- Enhanced FPUs
- Shared resources
  - L3 cache
  - Hypertransport links
  - Memory controllers
- Independent clock frequencies

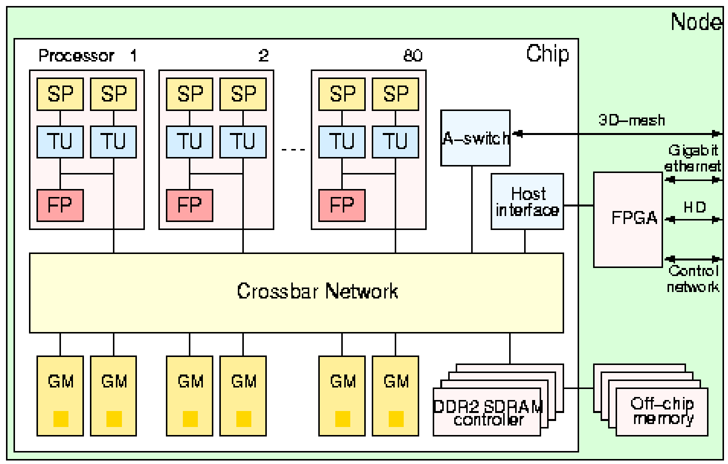


# Architectural Complexity – Multicore

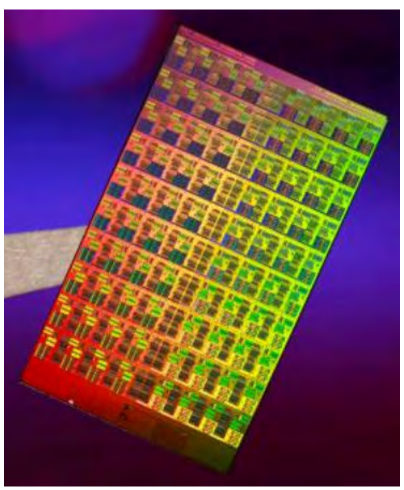
## Core count is easy to increase. Resource contention is a challenge!

1024…


160…

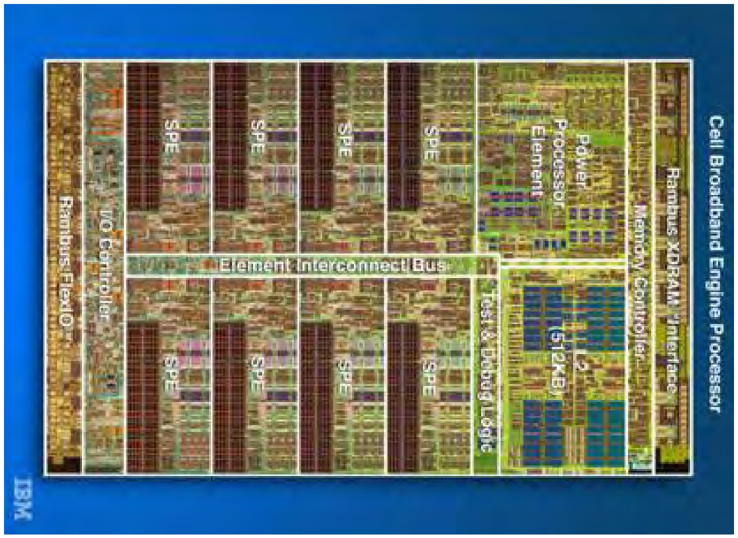


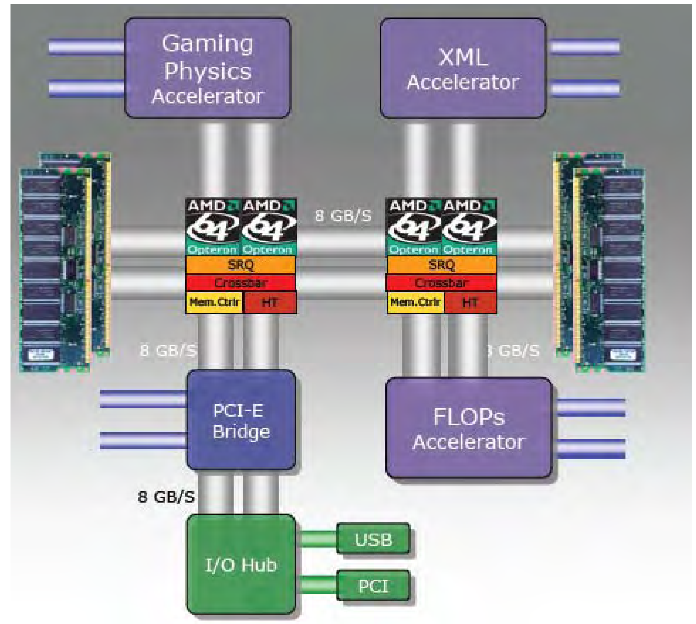

80…

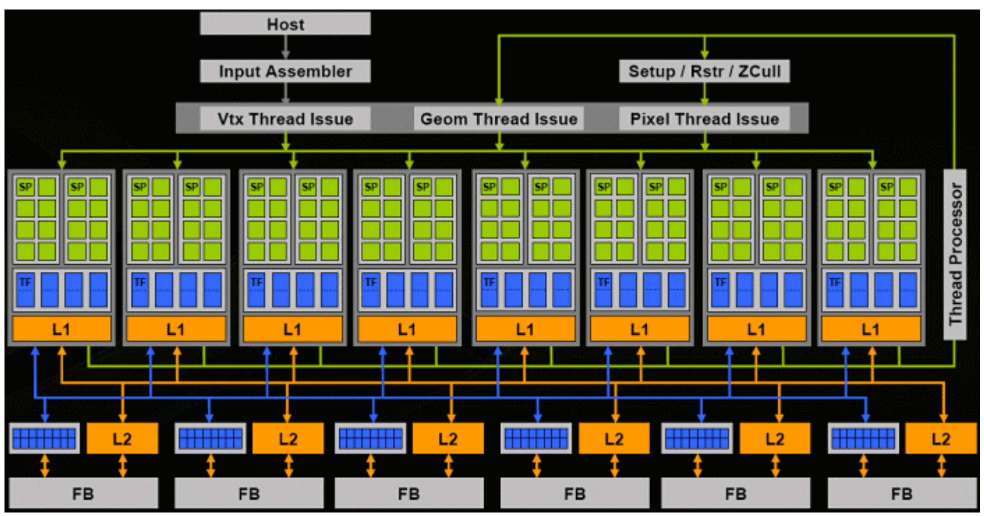


# Architectural Complexity – Heterogeneity, Specialization

**Architectures target specific workloads: games, graphics, business, encryption, media**

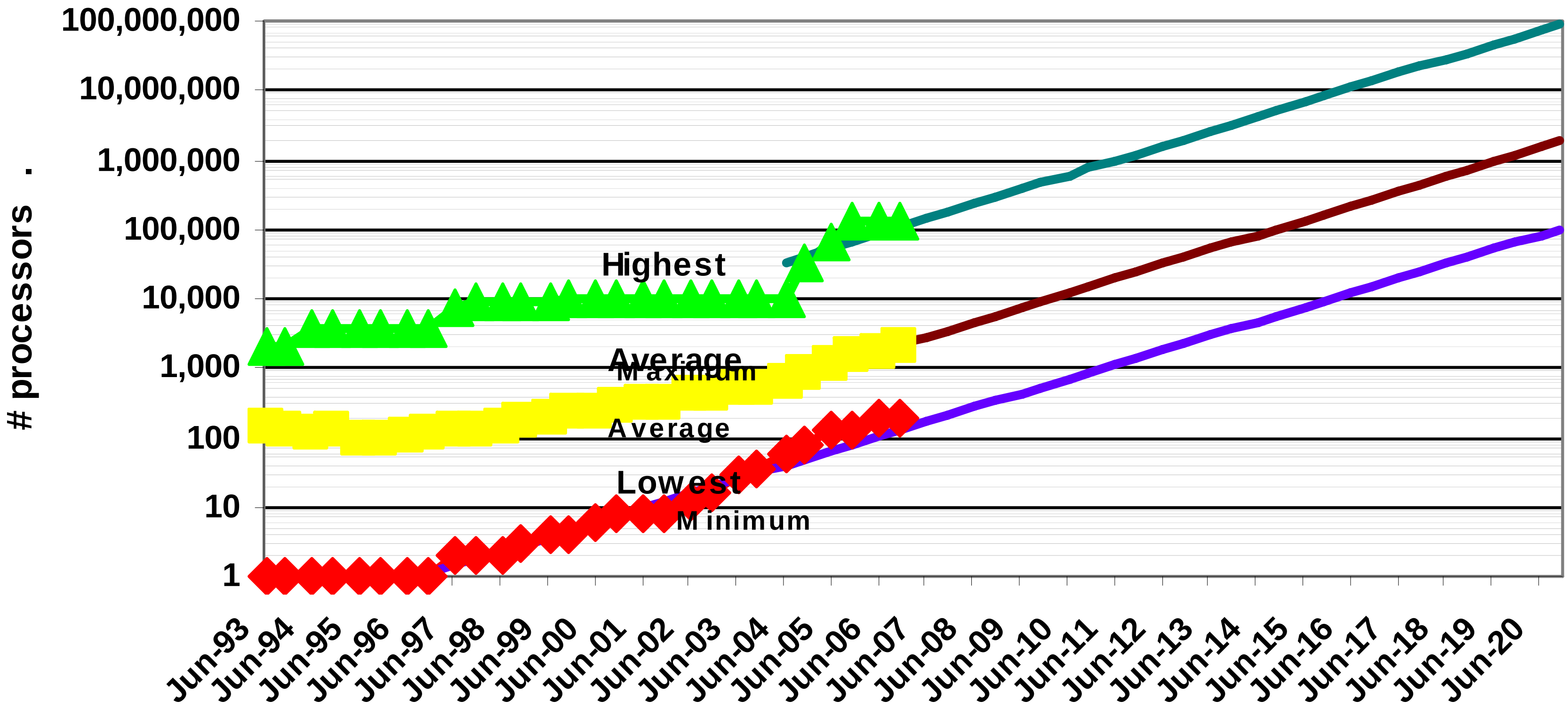


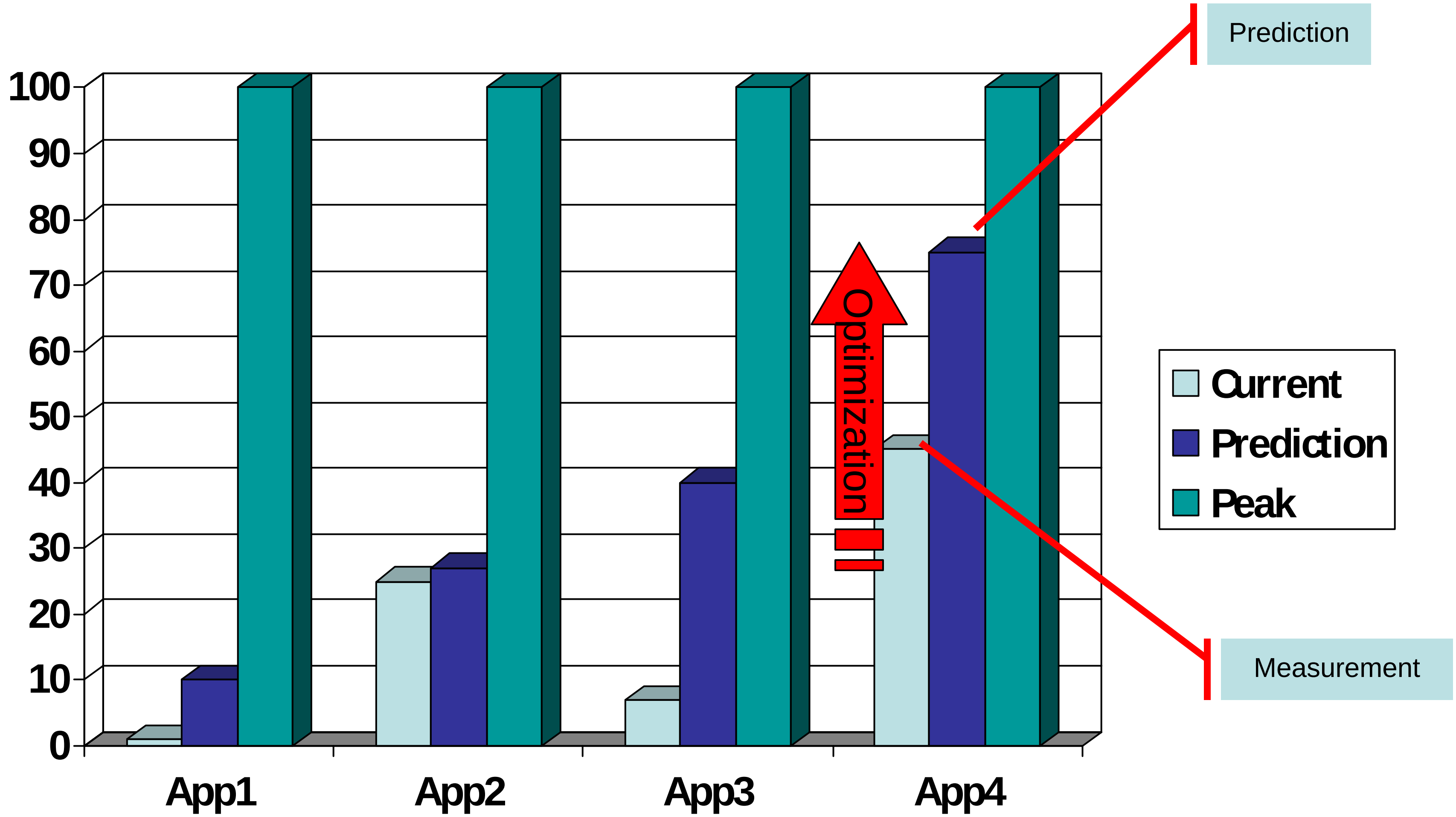



# System Scale

## Interconnect design and cost limits system scale

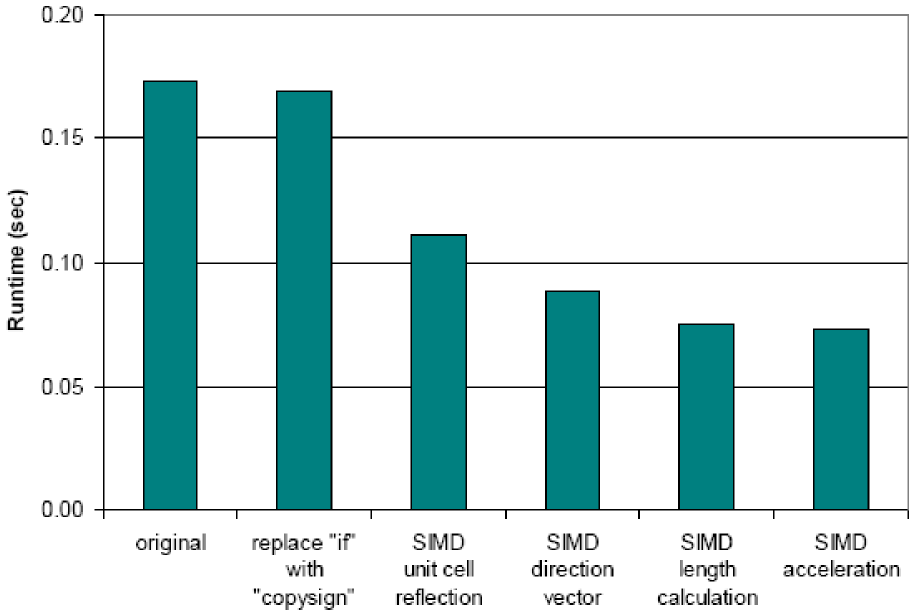


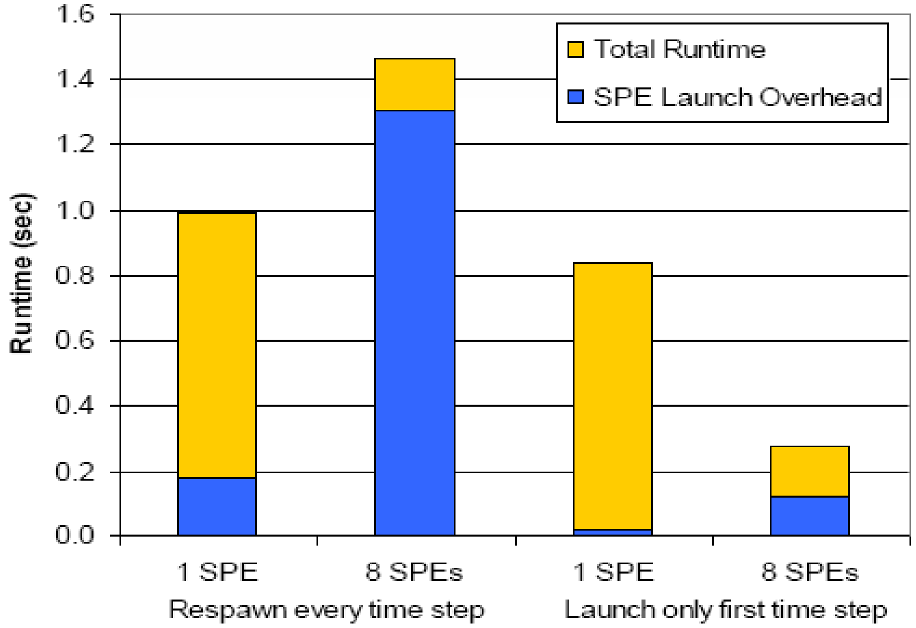
# Performance Engineering encompasses Measurement, Prediction, and Optimization

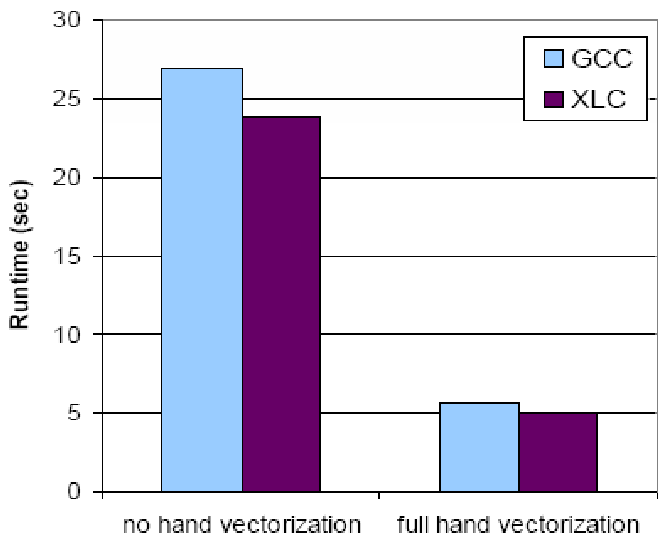


# STI Cell Demonstrates these Sensitivities

## GPUs, FPGAs, and other devices are similar

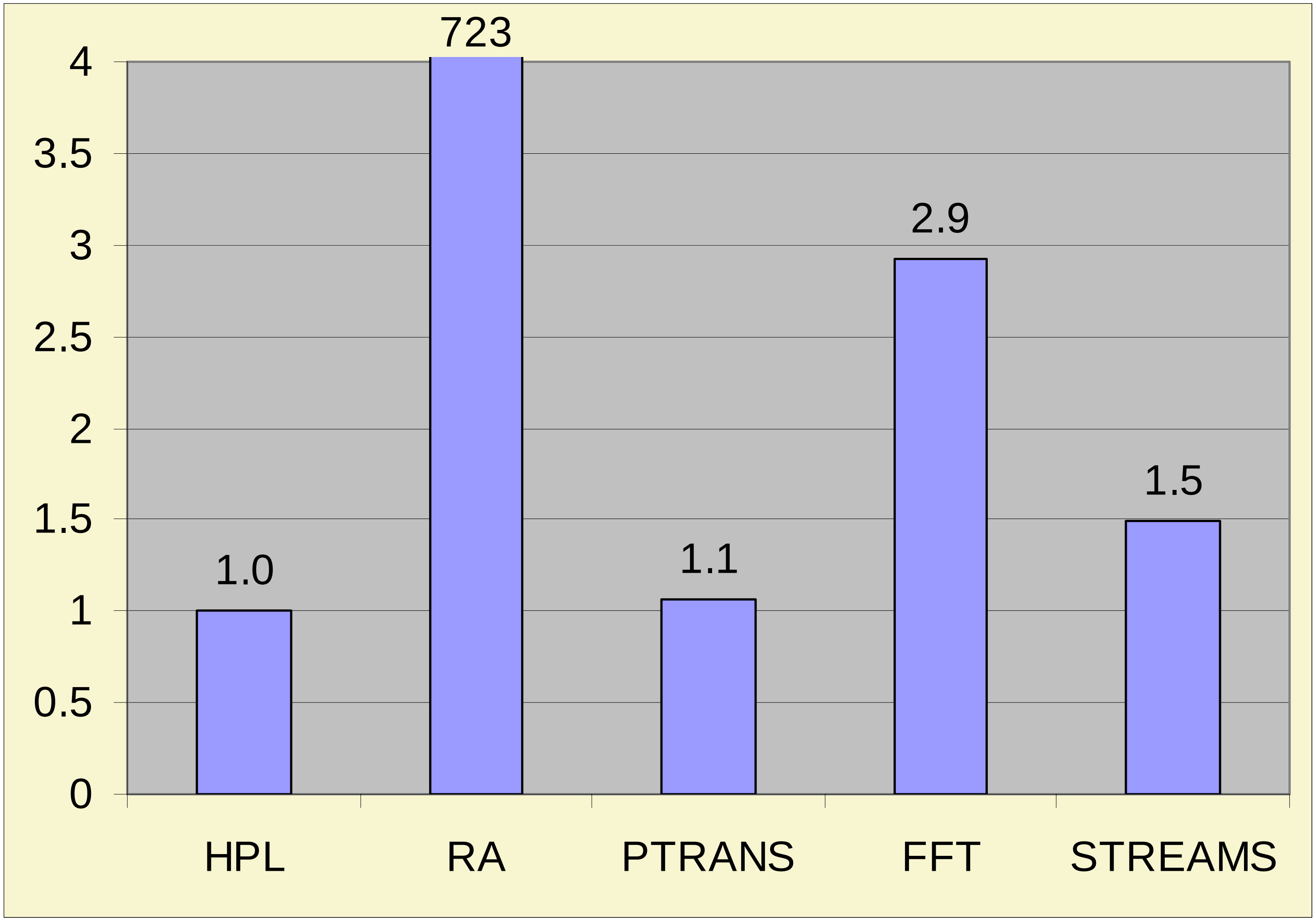


# HPC Challenge Benchmarks Demonstrate these Issues

## HPCC on Cray X1 – Baseline v. Optimized



# Application Diversity

- Multi-phase, multi-scale applications present challenges in performance engineering
  - Multiple languages
  - Multiple phases of physics, chemistry
  - Adaptive meshing, multigrid solvers, etc
- Applications teams know this best!

# Application Diversity

## Dwarfs illustrate some dimensions of this diversity

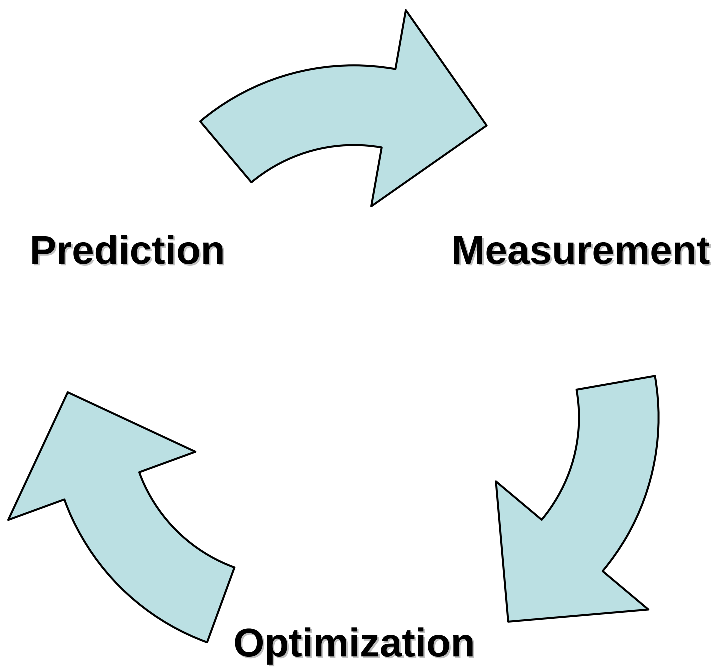
| Science Domain | Code | Structured Grids | Unstructured Grids | FFT | Dense Linear Algebra | Sparse Linear Algebra | Particles | Monte Carlo |
|---|---|---|---|---|---|---|---|---|
| *Accelerator Physics* | *T3P* | | X | | | X | | |
| *Astrophysics* | *CHIMERA* | X | | | X | X | X | |
| *Astrophysics* | *VULCAN/2D* | | X | | X | | | |
| *Biology* | *LAMMPS* | | | X | | | X | |
| *Chemistry* | *MADNESS* | | X | | X | | | |
| *Chemistry* | *NWCHEM* | | | X | X | | | |
| *Chemistry* | *OReTran* | X | | X | X | | | |
| *Climate* | *CAM* | X | | X | | | X | |
| *Climate* | *POP/CICE* | X | | | | X | X | |
| *Climate* | *MITgcm* | X | | | | X | X | |
| *Combustion* | *S3D* | X | | | | | | |
| *Fusion* | *AORSA* | X | | X | X | | | |
| *Fusion* | *GTC* | X | | | | X | X | X |
| *Fusion* | *GYRO* | X | | X | X | X | | |
| *Geophysics* | *PFLOTRAN* | X | X | | | X | | |
| *Materials Science* | *QMC/DCA* | | | | X | | | X |
| *Materials Science* | *QBOX* | | | X | X | | X | |
| *Nanoscience* | *CASINO* | | | | | | X | X |
| *Nanoscience* | *LSMS* | X | | | X | | | |
| *Nuclear Energy* | *NEWTRNX* | | X | | X | X | | |
| *Nuclear Physics* | *CCSD* | | | | X | | | |
| *QCD* | *MILC* | X | | | | | | X |

# Observations

- Take together, these three trends have the potential for creating a widening gap between expected and realized performance

- Performance engineering is critical to address this gap
  - Measurement
  - Prediction
  - Optimization

- We must feed this information back to architects and system software designers



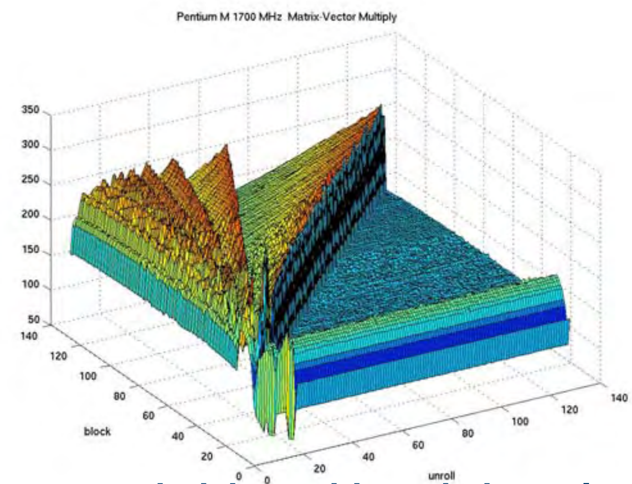
# Some Performance Engineering Solutions

- Engagement
  - Frequent interaction between applications teams and performance experts
- Tools
  - Instrumentation, collection, and analysis tools for measurement
- Automatic optimization
  - Static and dynamic optimization of applications and libraries
  - Integrate of performance engineering into application/system lifecycle
- Feedback to architects and system software designers

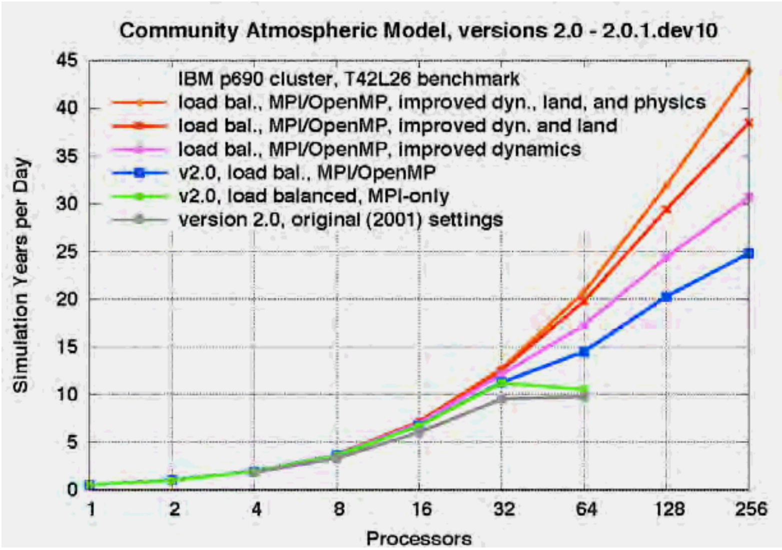
# Engagement: SciDAC PERI

- Application Engagement
  - Work directly with DOE computational scientists
  - Ensure successful performance porting of scientific software
  - Focus PERI research on real problems
- Application Liaisons
  - Build long-term personal relationships with PERI researchers and scientific code teams
- Tiger Teams
  - Focus on DOE's highest priorities
    - SciDAC-2
    - INCITE

See www.peri-scidac.org for more info.




Optimizing arithmetic kernels



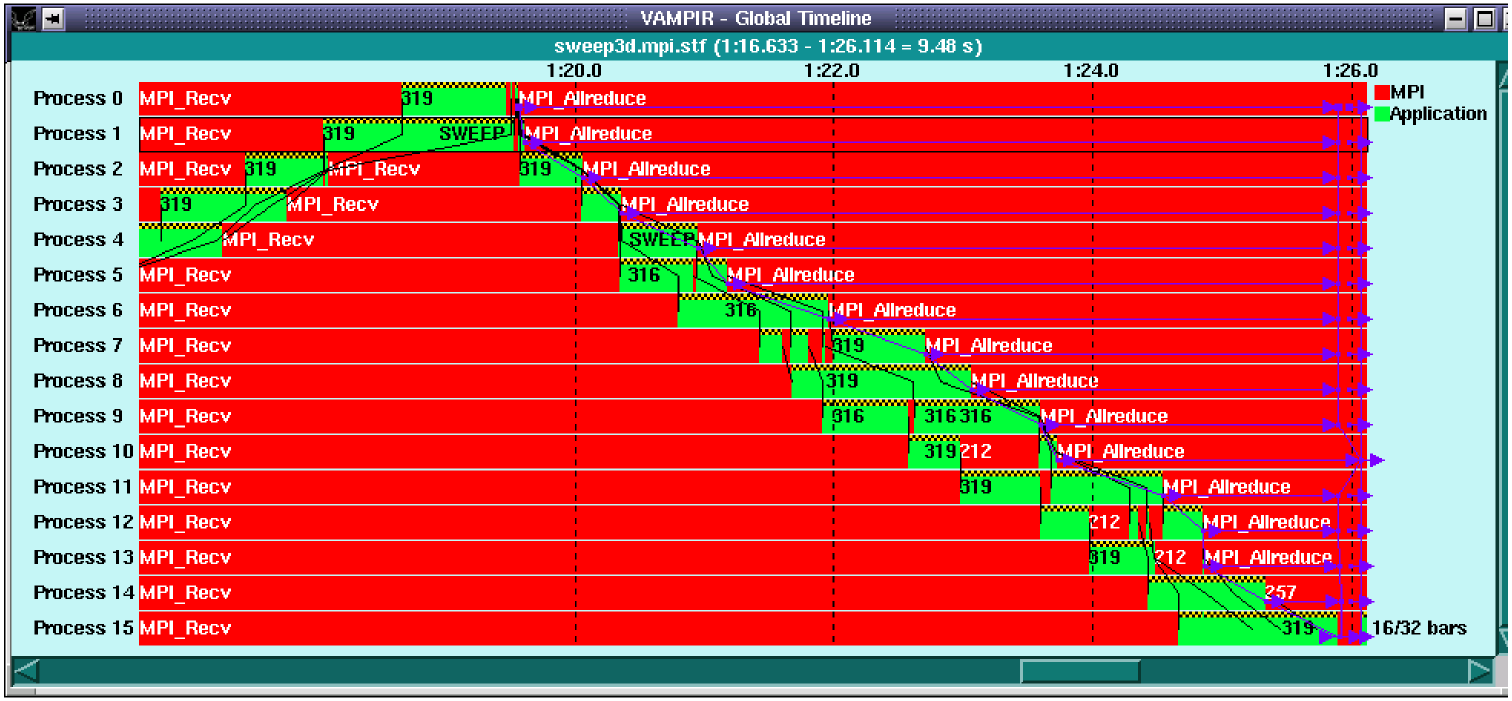

Maximizing scientific throughput

# Tools: Software Development Tools for Petascale Computing

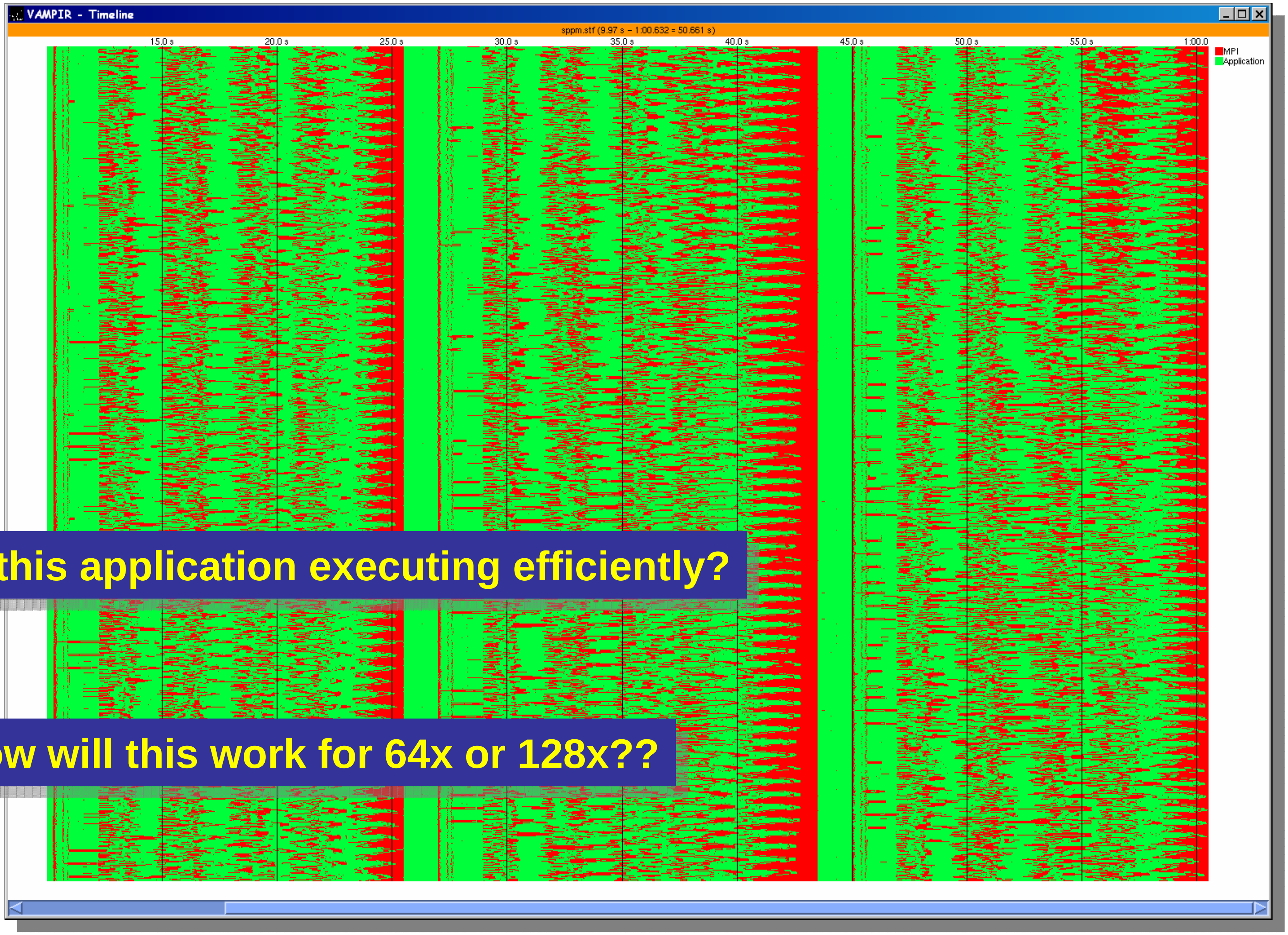
- Assembled ~60 experts in software development tools to identify challenges for Petascale
- See Fred Johnson's presentation

# Tools: Traditional Performance Analysis of Communication Operations

- Many MPI tools use tracing
  - Produces very detailed information about communication activity
  - Illustrates dependencies among tasks



# Tools: Timeline with 1024 tasks



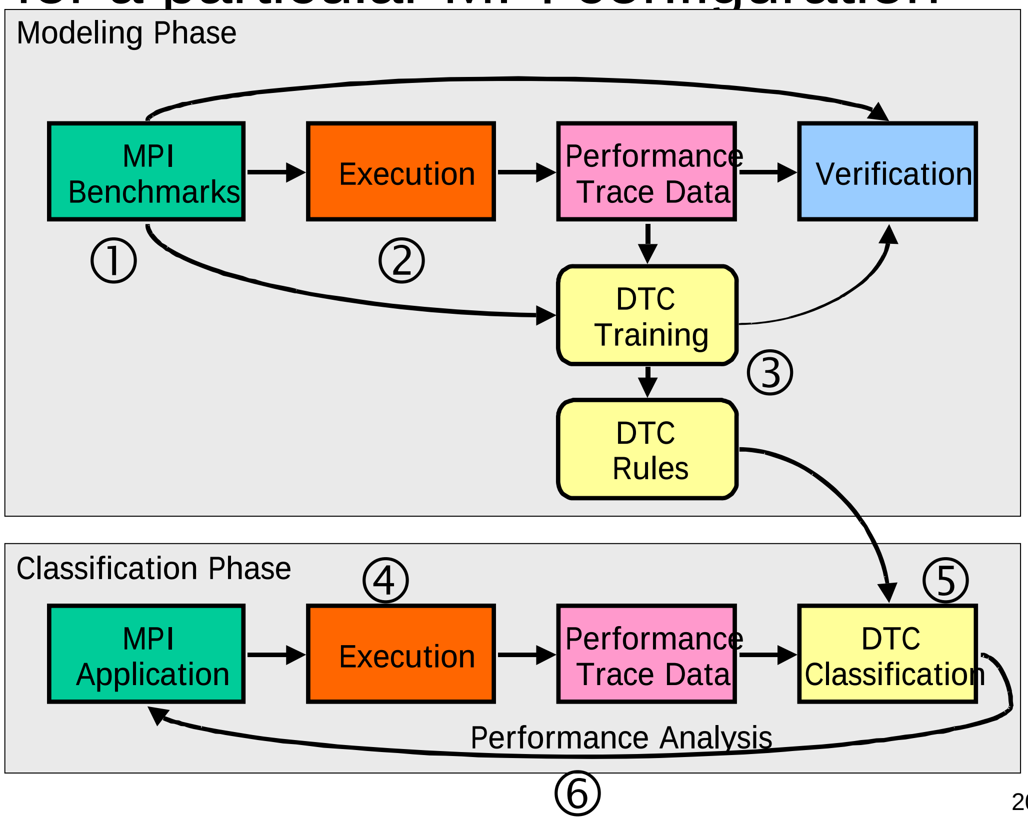

**Is this application executing efficiently?**

**How will this work for 64x or 128x??**

# Tools: Automatic Classification for Communication Performance Analysis

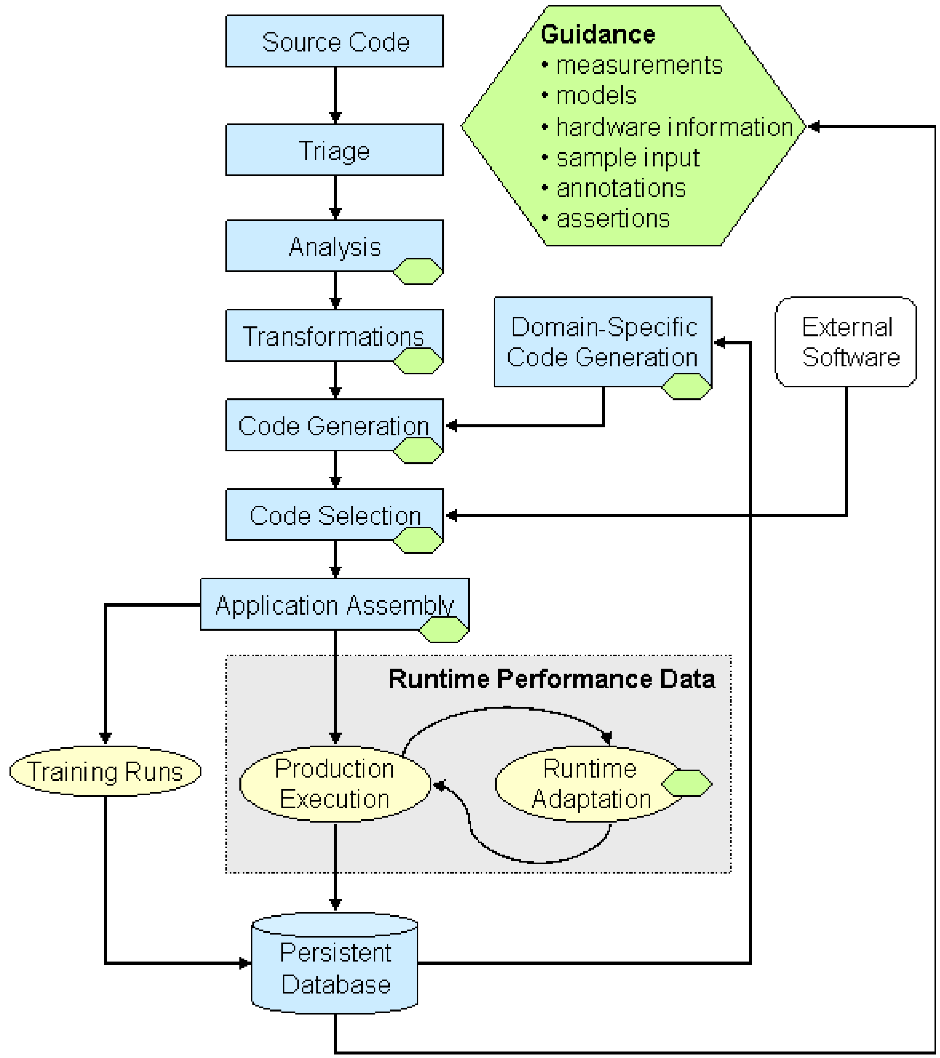
- Use decision tree classification (a supervised learning technique) to classify application's messages automatically
- Compare an application's message operations to 'normal' communication for a particular MPI configuration

- Modeling Phase (once)
  - Use benchmarks to generate decision tree
  - Both efficient and inefficient
- Classification Phase (many)
  - Execute application
  - Analyze application trace with classifier based on decision tree



# Automatic Optimization: SciDAC PERI Framework

- Long-term goals for PERI
- Automate the process of tuning software to maximize its performance
- Build upon forty years of human experience and recent success with libraries
  - PHIPAC, ATLAS, FFTW, SPIRAL, SPOOLES
- Reduce the performance portability challenge facing computational scientists
- Address the problem that performance experts are in short supply




PERI automatic tuning framework

# Automatic Optimization: Engineering Applications for Performance Throughout their Lifecycle

- Use performance assertions to verify the performance explicitly

```
pa_start(&pa, "$nFlops", PA_AEQ, "11 * %g * %g", &ym, &xm);
for (j=ys; j<ys+ym; j++) {
  for (i=xs; i<xs+xm; i++) {
    if (i == 0 || j == 0 || i == Mx-1 || j == My-1) {
      f[j][i] = x[j][i];
    } else {
      u       = x[j][i];
      uxx     = (two*u - x[j][i-1] - x[j][i+1])*hydhx;
      uyy     = (two*u - x[j-1][i] - x[j+1][i])*hxdhy;
      f[j][i] = uxx + uyy - sc*PetscExpScalar(u);
    }
  }
}
pa_end(pa);
PetscLogFlops(11*ym*xm);
```

- Expression
  - `"$nFlops", PA_AEQ, "11 * %g * %g", &ym, &xm`
  - Empirically measure number of floating point operations with instrumentation
  - Test approximate equality (±10%) to `'11 * ym * xm'` ?
- Empirical measurements verify performance model

# Many Other Performance Engineering Topics…

- Performance prediction
  - Analytical modeling
  - Simulation
  - Hybrid
  - Historical predictions
- New programming models, languages
- Reliability, correctness, fault tolerance
- IO
- Cooperation with vendors on hardware and software architecture and performance engineering support
- Etc…

# Summary

- Significant trends emerging in HPC
  - Architectural complexity: multicore, heterogeneity, power mgt
  - Scale
  - Application complexity: multiphase, multiscale, multiresoution
- Taken together, these trends can create a gap between expected and realized performance
- Performance engineering is critical to bridging this gap
  - Measurement
  - Prediction
  - Optimization
  - Feedback to architects and system software designers
- Some performance engineering solutions that help to close this gap
  - Engagement
    - Frequent interaction between applications teams and performance experts
  - Tools
    - Instrumentation, collection, and analysis tools for measurement
  - Automatic optimization
    - Static and dynamic optimization of applications and libraries
    - Integrate of performance engineering into application/system lifecycle
  - Feedback to architects and system software designers

# Acknowledgements

- Thanks to contributors, sponsors
  - ORNL Future Tech team http://ft.ornl.gov
  - SciDAC PERI http://www.peri-scidac.org
  - INCITE PEAC
  - NCCS http://nccs.gov
  - Experimental Computing Lab @ ORNL
  - Cray, AMD
- More information
  - Future Tech: Publications: http://ft.ornl.gov/pubs
- Acknowledgements
  - This research was sponsored by the Office of Mathematical, Information, and Computational Sciences, Office of Science, U.S. Department of Energy under Contract No. DE-AC05-00OR22725 with UT-Battelle, LLC. Accordingly, the U.S. Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.