USC Viterbi
School of Engineering

Productivity Metrics for HPC WPL Applications

Information Sciences Institute, Computational Sciences Division

ISI
Information Sciences Institute
USC

12 December 2006

Bob Lucas
rflucas@isi.edu

USC Viterbi School of Engineering

Application Session Speakers

ISI Information Sciences Institute

Bruce Hendrickson

Graph-based Informatics

Mike Heroux

Petra Object Model

Kaylan Perumualla

Parallel Discrete Events

Jeffery Vetter

How Are Applications Evolving?

USC

USC Viterbi School of Engineering

Themes

ISI Information Sciences Institute

Workshop: HPC Languages

Session: “New” Applications

My Talk: Productivity

USC

USC Viterbi School of Engineering

Where's the "value" in HPC?

ISI Information Sciences Institute

More value in application S/W than in H/W

- GM spends more on S/W than systems
- DOE NNSA ASC spends O(10%) on H/W

S/W outlives H/W

Each generation of HPC abandons earlier apps.

USC

USC Viterbi School of Engineering

"New" Application Thoughts

ISI Information Sciences Institute

Some are so old we've forgotten them

e.g., NASTRAN, TCAD, etc.

Old, but unfamiliar to this community

e.g., Forces modeling

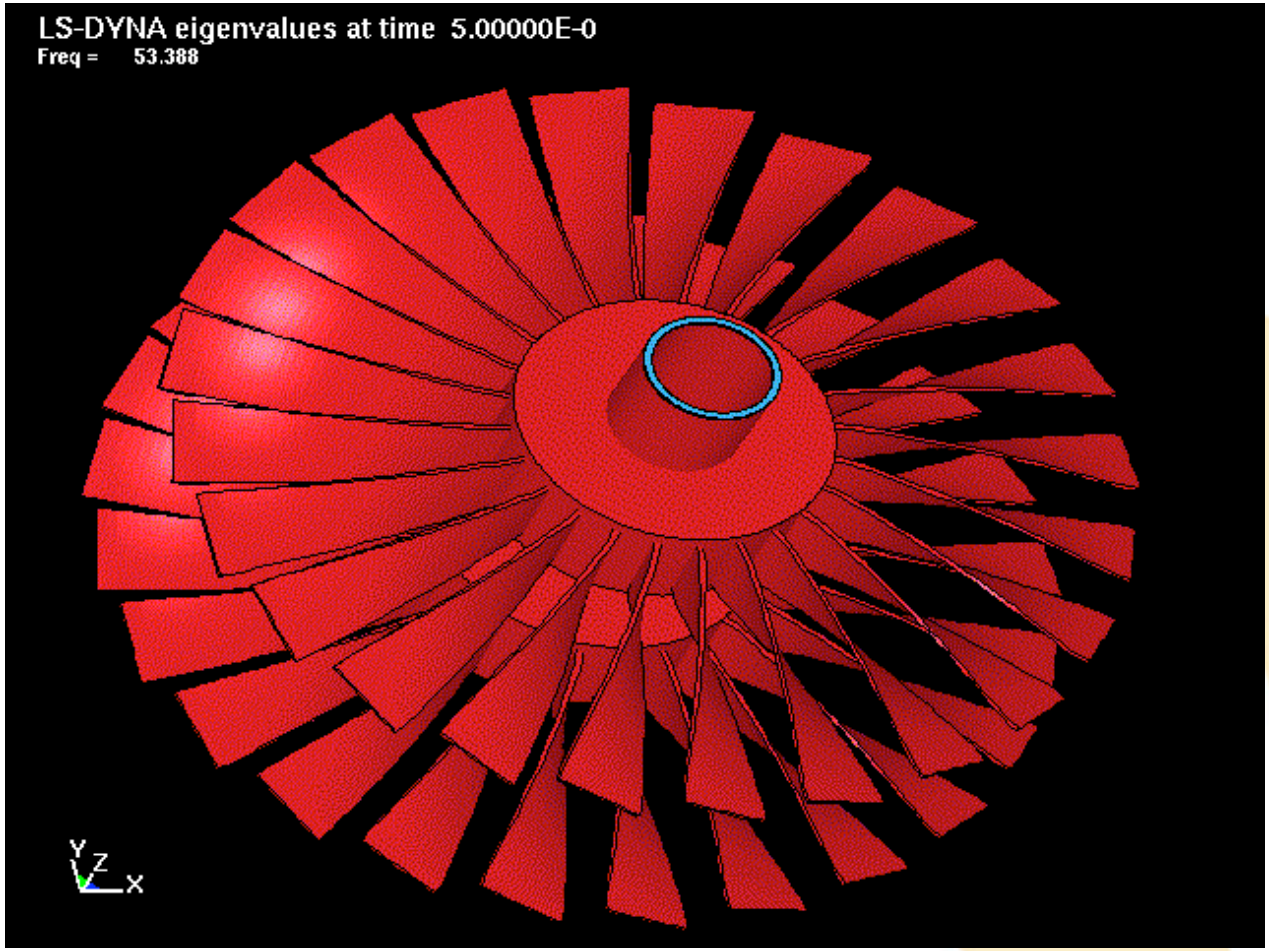
Truly new ones

Wait for the next speakers

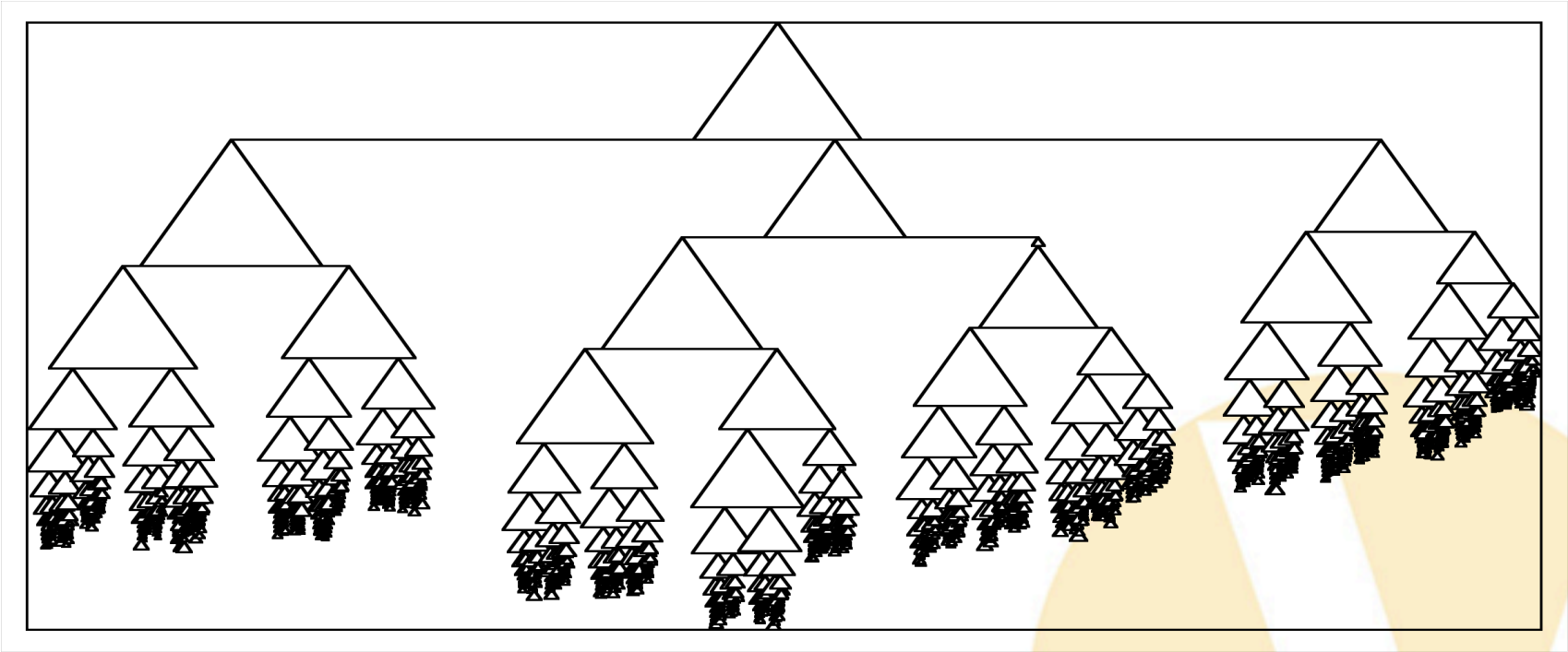
USC

Vibration (i.e., NASTRAN): Oldie but Goodie

Jet engine impeller, modeled with LS-DYNA



"Hood" Elimination Tree



Each frontal matrix's triangle scaled by operations required to factor it. How do you load balance this on distributed memory?

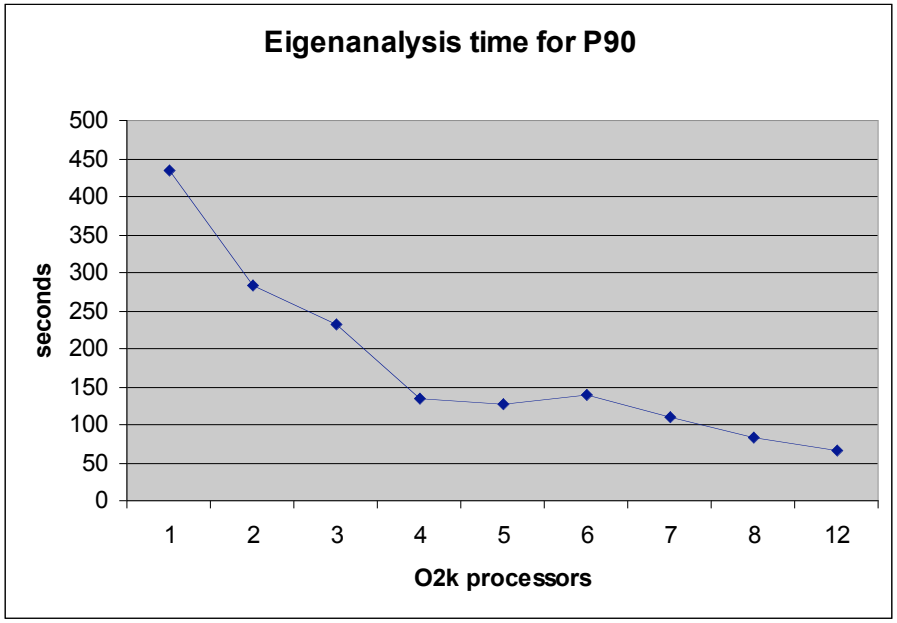
Compounded in Lanczos Algorithm

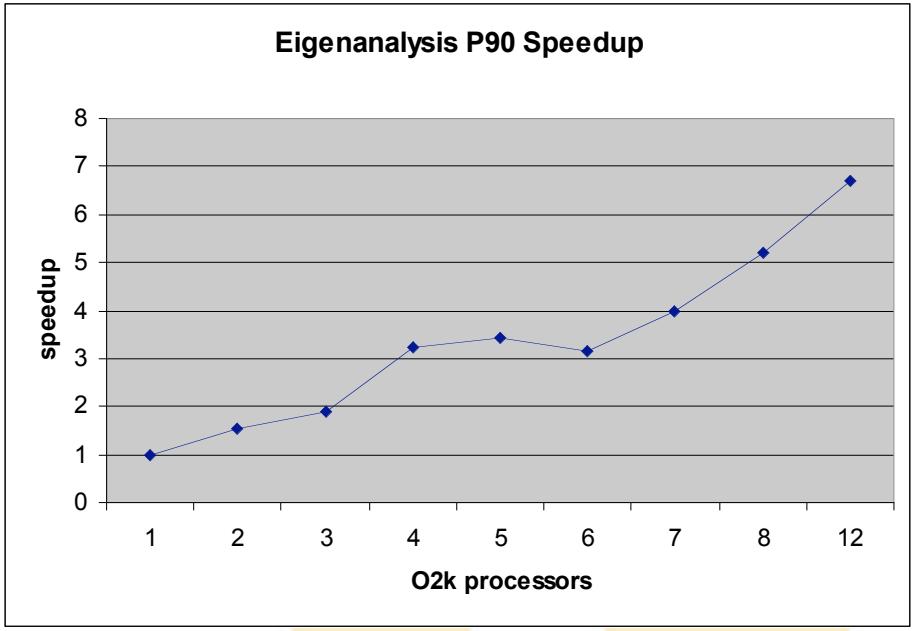
- Block shift invert Lanczos algorithm
  - 1/3 each of factor, solve, DGEMM
  - Optimal partitioning different
    - NP-complete for factor and solve
- Curse of the "MPI" distributed model
  - One distribution reduces communication

    -or-

  - Multiple distributions reduce load-balance

Distributed Lanczos Scaling





I need a new model to get this to Terascale

Start with shared memory

- Facilitates reassigning ops w/o moving data

Evolve from Fortran/C

- Can't abandon billions of dollars of S/W

USC Viterbi School of Engineering

Another "New" Application

ISI Information Sciences Institute

UNITED STATES JOINT FORCES COMMAND

USJFCOM

Joint Experimentation

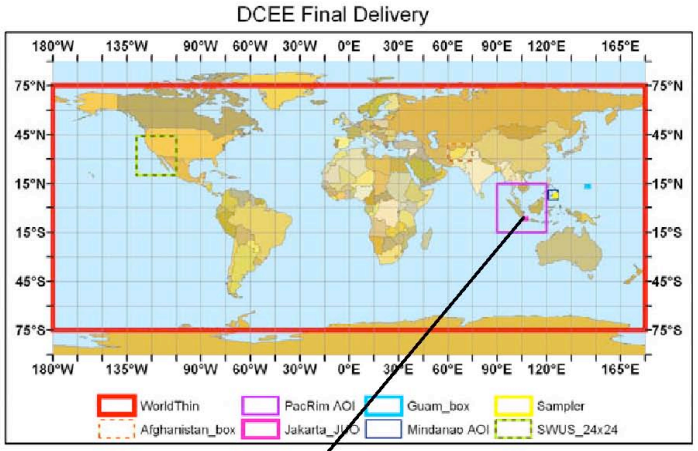
Urban Resolve 2015

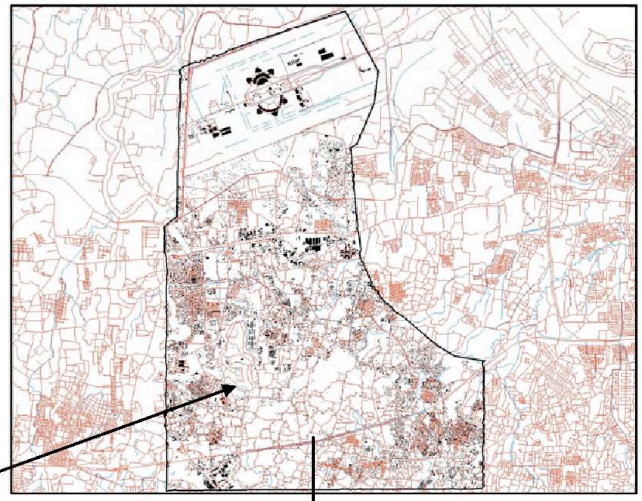
Simulated experimentation is over a decade old

USC

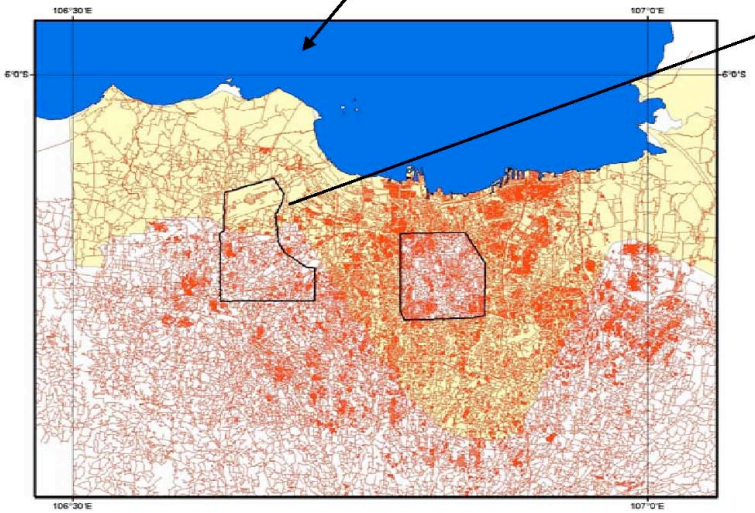
Terrain Box

Simulate World-Wide Terrain





• Geo-Specific Terrain in Western AOI



•1.8 million Buildings & Structures

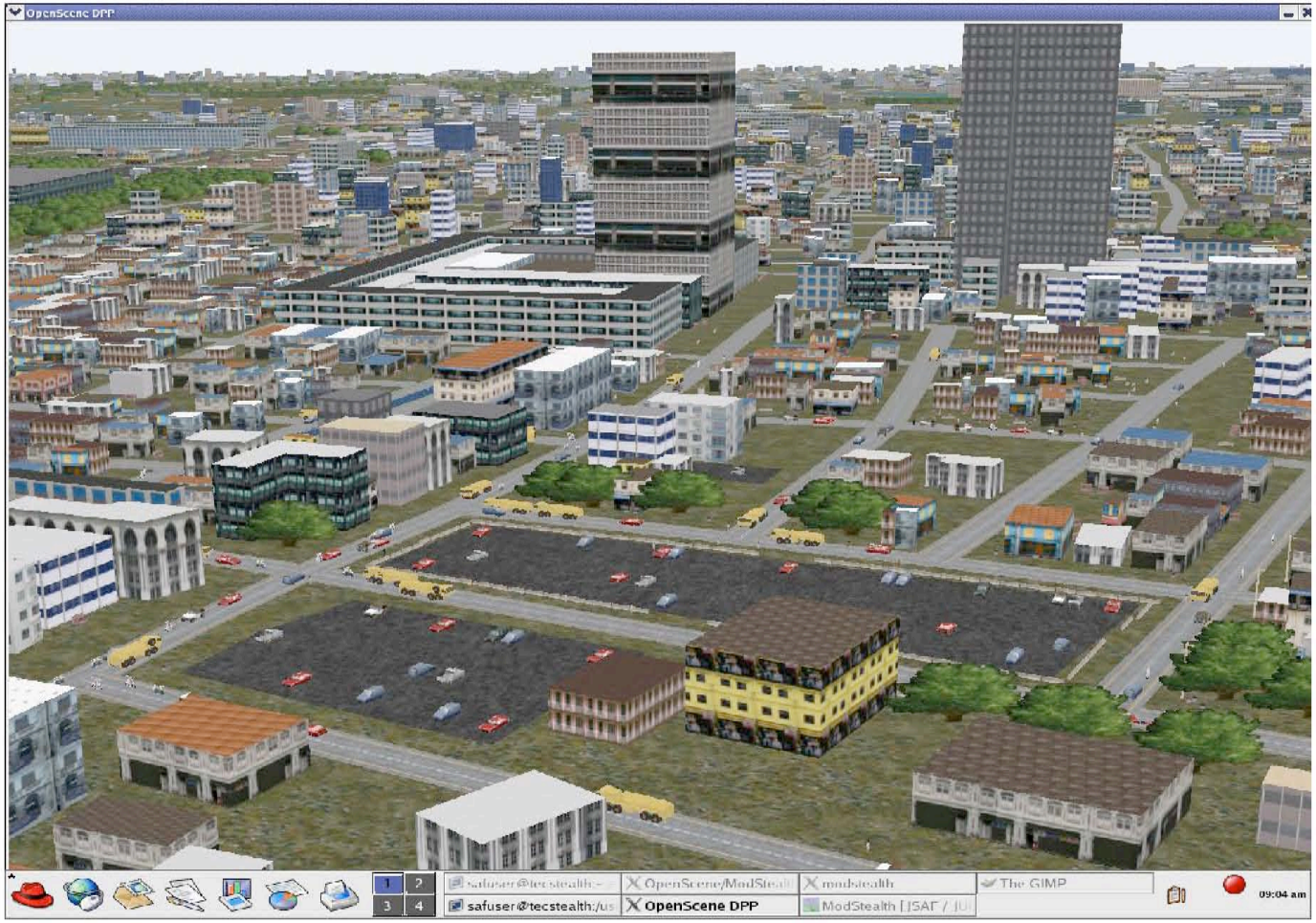
• 47,000 Buildings & Structures in AOI

USC Viterbi School of Engineering

View of a City

ISI Information Sciences Institute

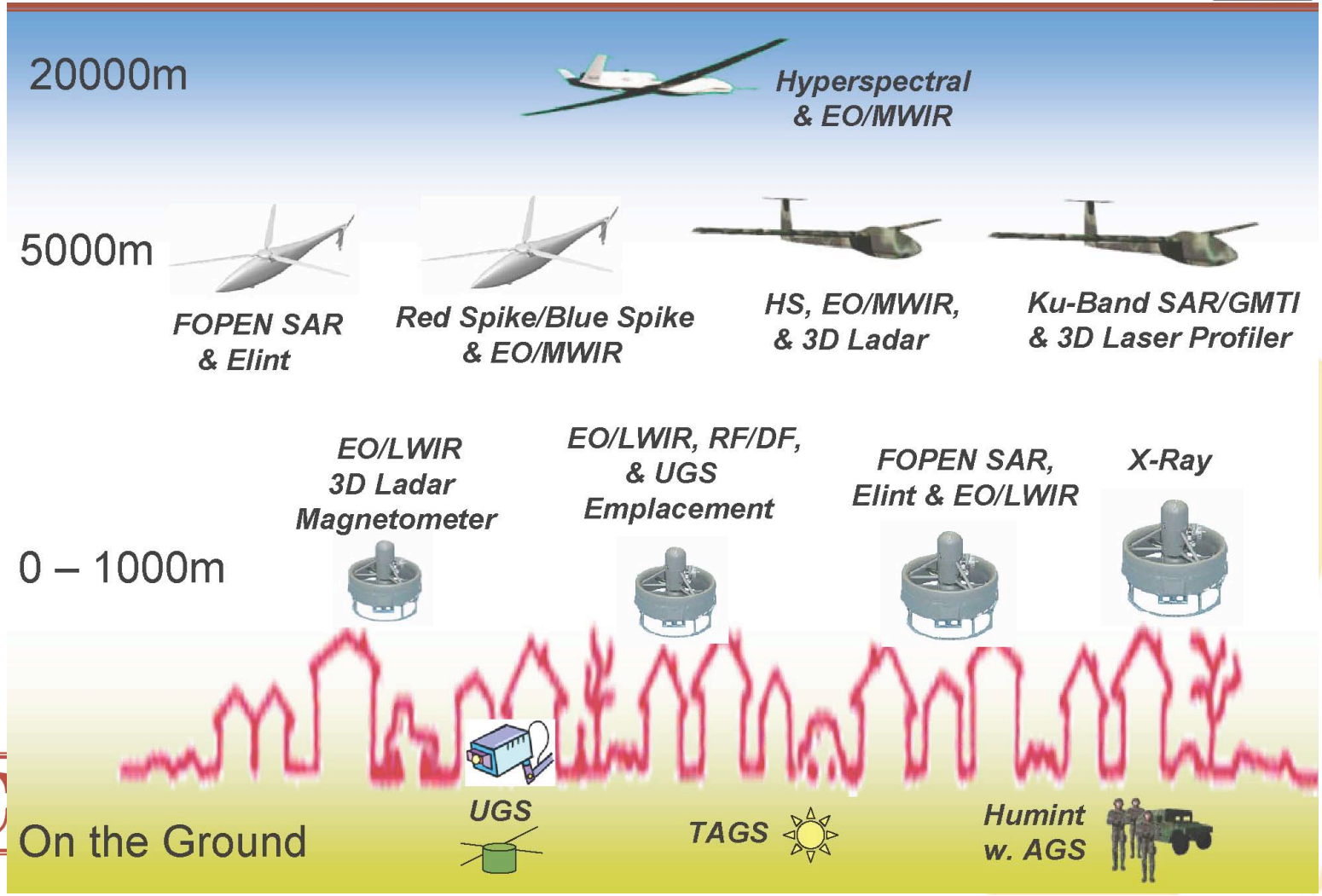
STEALTH View-Urban Area



USC

Experimental Sensor Architecture

Sensor/Platform Architecture

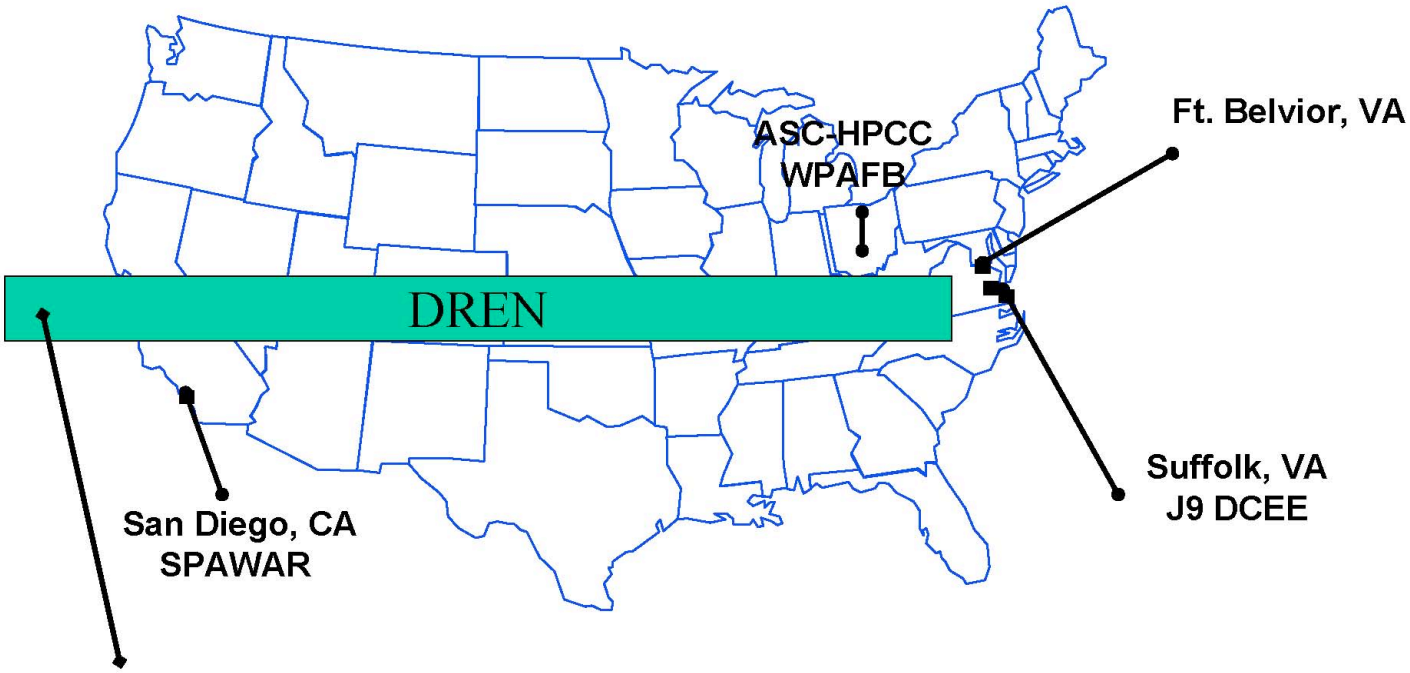


USC Viterbi School of Engineering

Distributed Over a "Grid"

Information Sciences Institute

JUO Participants Today



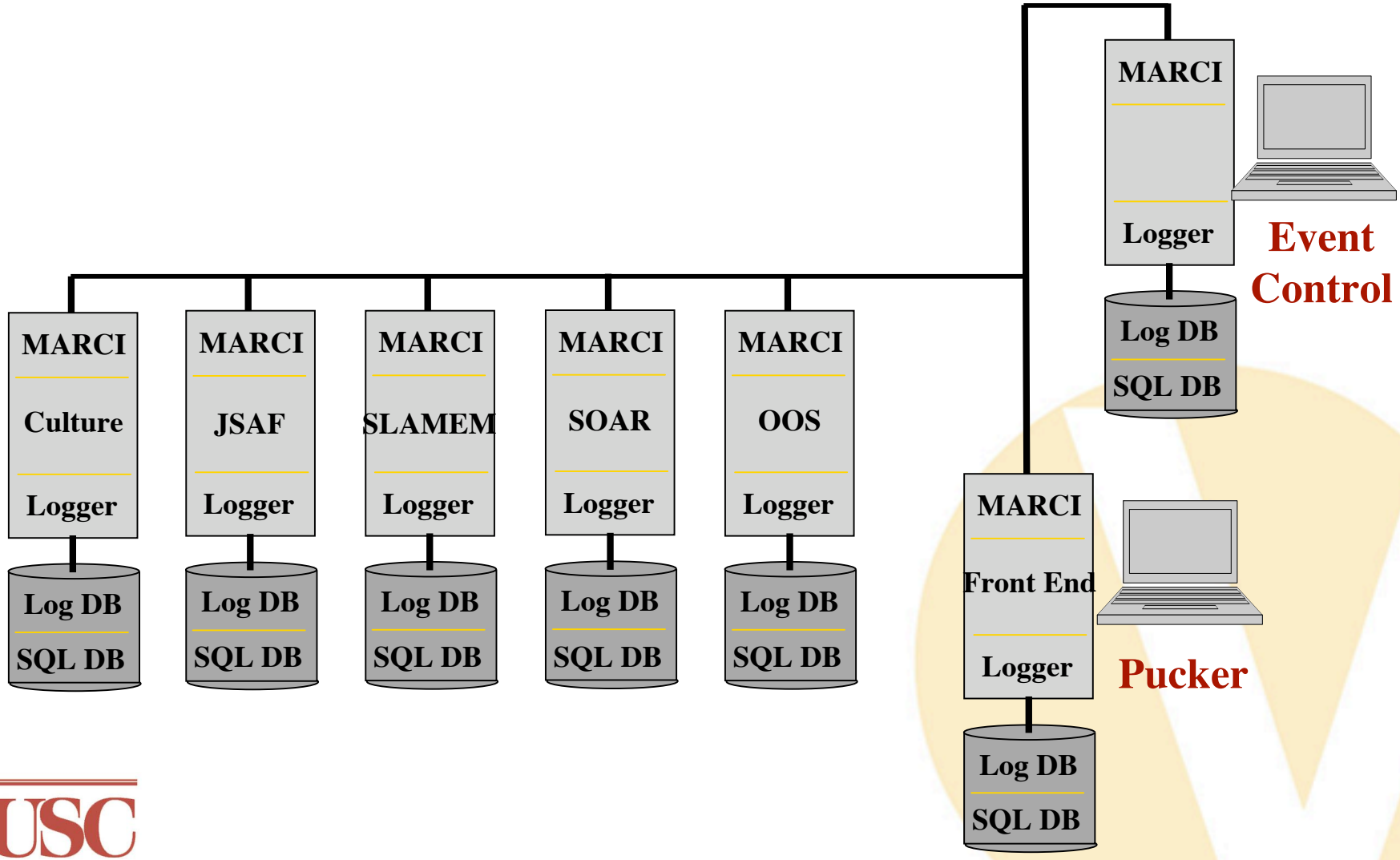
Up to 140 Mb/S available
Typically use @ 35Mb/S

USC

USC Viterbi School of Engineering

Heterogeneous Computing

ISI Information Sciences Institute



USC

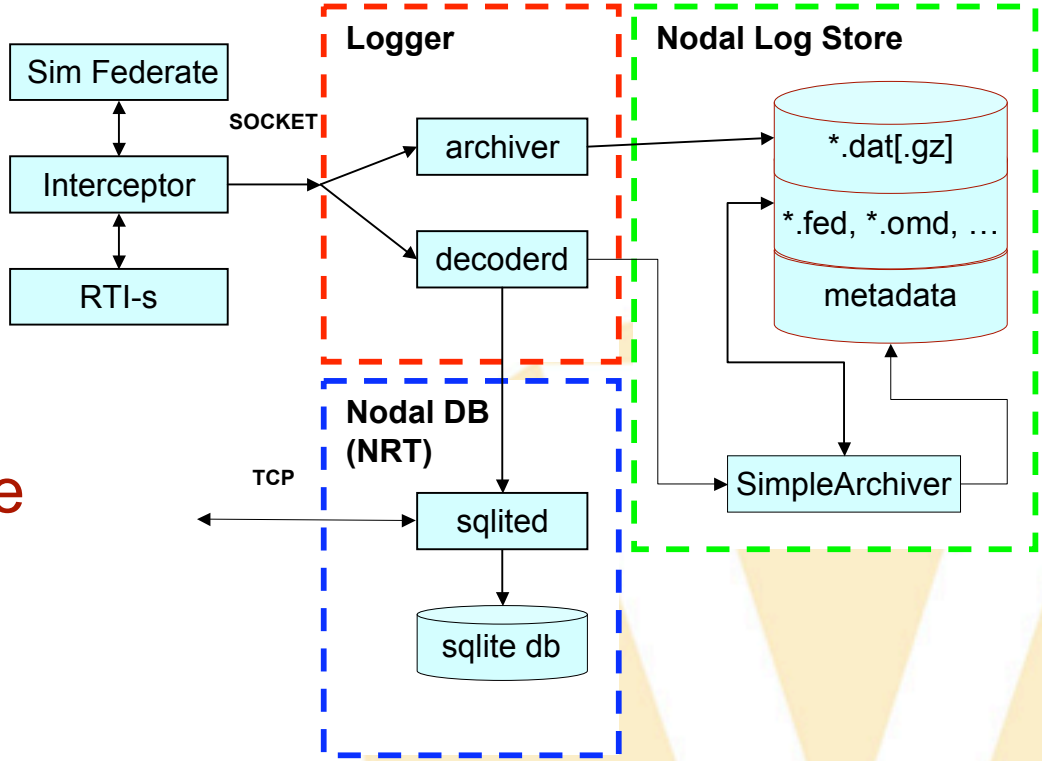
Data Intensive Too

Fully distributed logging at point of generation

- **RTI Interceptor** captures simulation events
- **Archiver** stores events to local disk
- **Decoderd** stores events to local relational database

Binary DB for R/T queries

SQL DB for after action



USC Viterbi School of Engineering

Compute Bottlenecks

ISI Information Sciences Institute

- Geometry
  - Determine Line-of-sight
  - Route planning
  - Collision detection
- Artificial Intelligence
  - Model human behavior
  - Learn and adapt
- Analysis
  - Database queries

USC

USC Viterbi School of Engineering

Themes

ISI Information Sciences Institute

Workshop: HPC Languages

Session: New Applications

Supposed to talk about: Productivity

USC

HPCS Notion of Productivity

Subjective Values: Site-specific

Objective Measurable System Properties

Project Utility Job Utility

System-level Utilization Job-level Utilization

System Availability

Computational Resources

$$P = \frac{U_{sys} E_{util} E_{adm} A_{sys} E_{job} R}{C_{sys}^{initial} + c_{sys}^{annual} T + c_{adm} n_{adm} T + N c_{proj} \bar{t}_{proj}}$$

System Costs

System Lifetime

Admin Costs

Development Costs

USC Viterbi School of Engineering

How Do We Increase It?

ISI Information Sciences Institute

- Reduce the denominator
  - Minimize H/W cost
    - Beowulf model
  - Reduce S/W development cost
    - HPCS development time goal
- Increase the numerator
  - HPCS execution time goal

USC

USC Viterbi School of Engineering

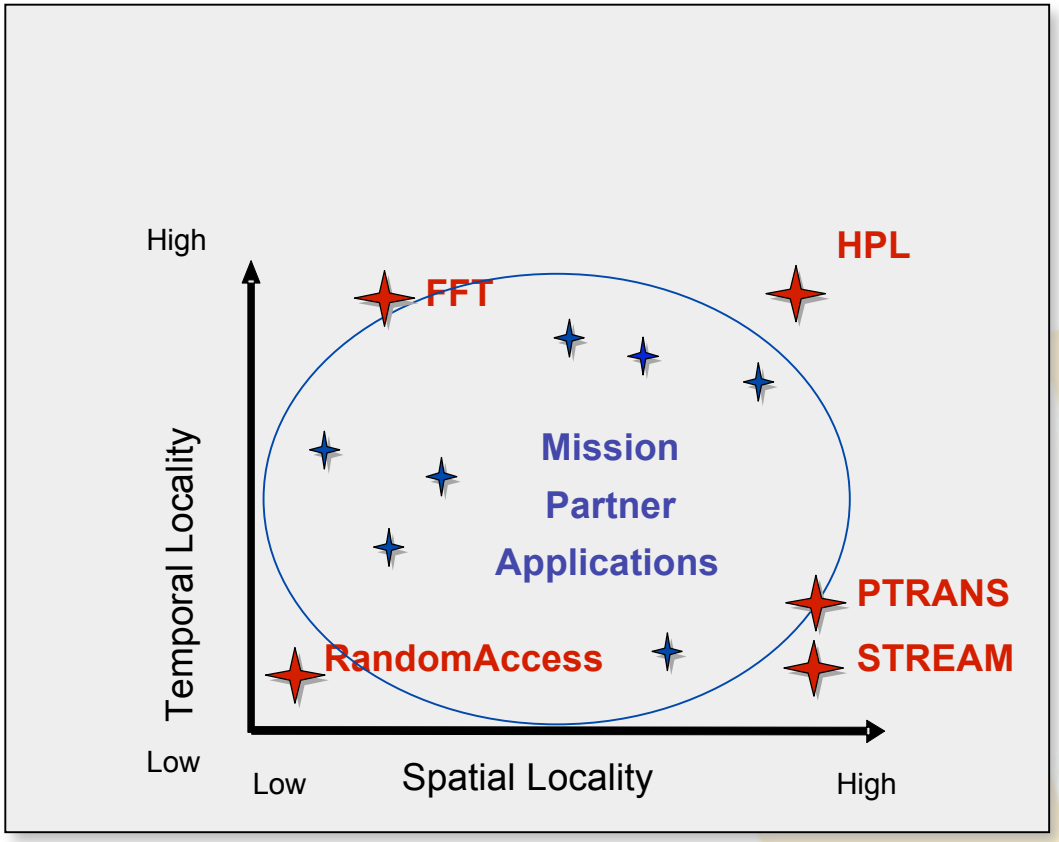
How Do We Measure It?

ISI Information Sciences Institute

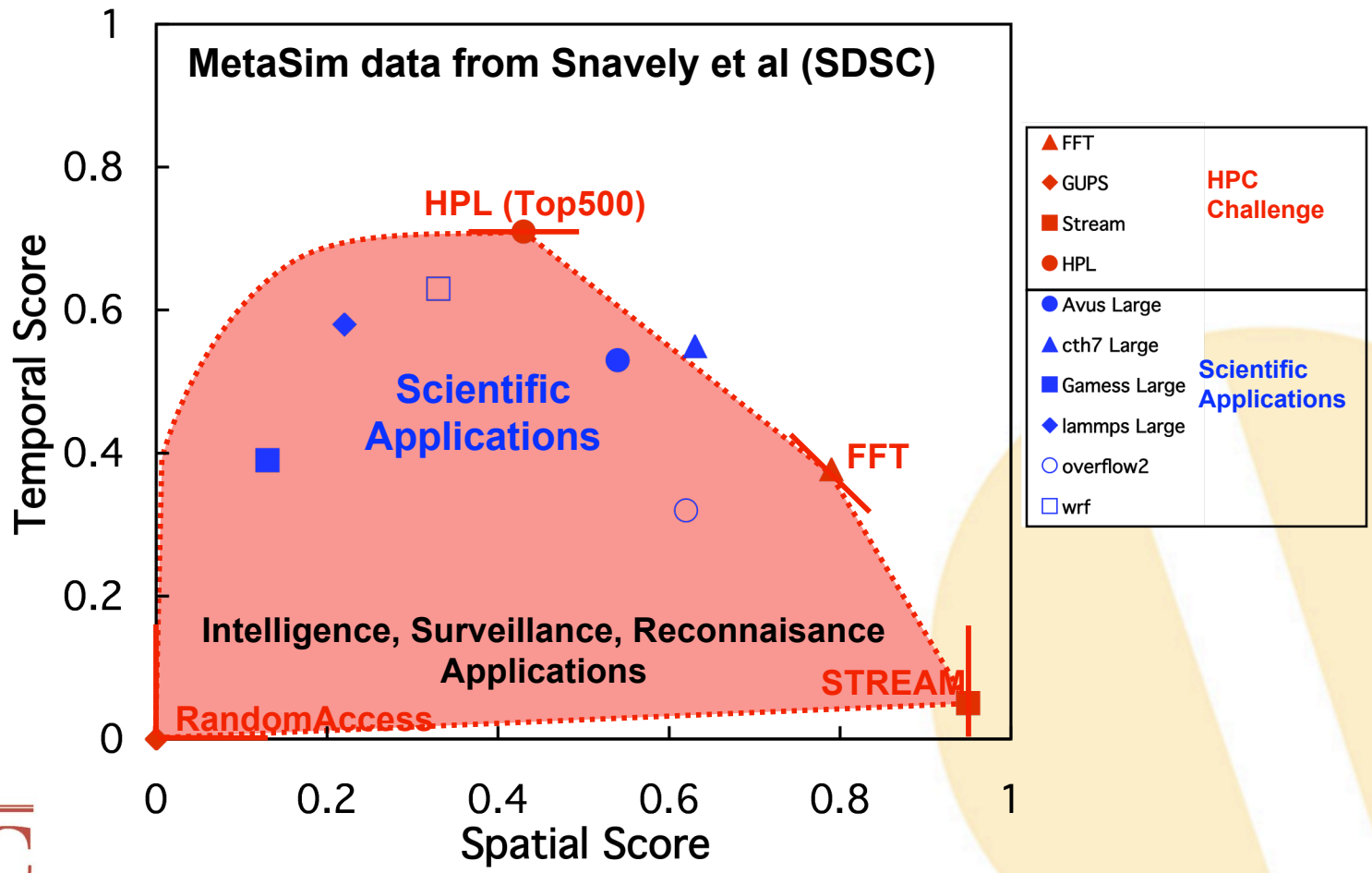
- Execution Time
  - Relatively easy …
    - Undoubtedy with benchmarks
    - Ideally count the applications
    - Measure their run time
- Development Time
  - Much harder
    - Counting SLOCs is inadequate
    - No other widely used measure

USC

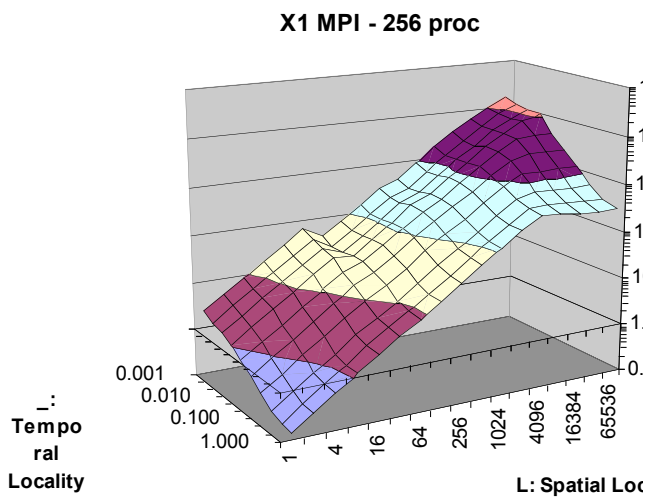
Original HPCS Spatial/Temporal Notion

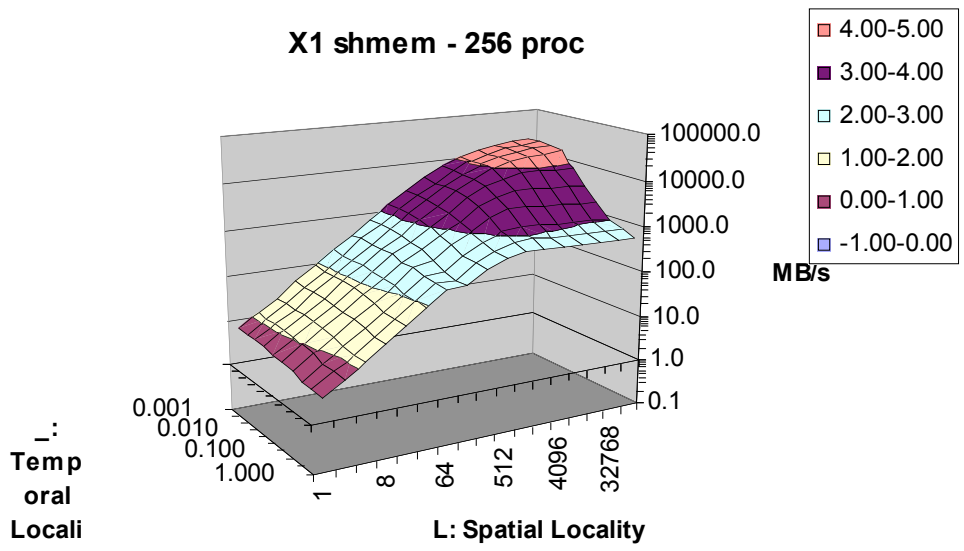


Current HPCS Spatial/Temporal Notion



Performance Surfaces





Computed with Strohmaier's APEX-Map

Both H/W and S/W impact throughput

MPI vs. SHMEM above

USC Viterbi School of Engineering

www.hpcchallenge.org

ISI Information Sciences Institute

HPCCHALLENGE HPCS

HPCCHALLENGE HPCS

- Home
- News
- Publications
- FAQ
- Links
- Committee
- Partners
- Awards

HPC Challenge Award Competition

The DARPA High Productivity Computing Systems (HPCS) Program and HPCWire are pleased to announce the annual HPC Challenge Award Competition. The goal of the competition is to focus the HPC community's attention on developing a broad set of HPC hardware and HPC software capabilities that are necessary to productively use HPC systems. The core of the HPC Challenge Award Competition is the HPC Challenge benchmark suite developed at the University of Tennessee under the DARPA HPCS program with inputs from a wide range of organizations from around the world (see http://icl.cs.utk.edu/hpcc/).
The Competition will focus on four of the most challenging benchmarks in the suite:

- Global HPL
- Global RandomAccess
- EP STREAM (Triad) per system
- Global FFT

There will be two classes of awards.

Class 1: Best Performance (4 awards) Best performance on a base or optimized run submitted to the HPC Challenge website. The benchmarks to be judged are: Global HPL, Global RandomAccess, EP STREAM (Triad) per system and Global FFT. The prize will be $500 plus a certificate for the best of each. Entries will be considered for award if they are submitted before November 13, 2006.

Class 2: Most Productivity Most "elegant" implementation of **three** or more of the HPC Challenge benchmarks with special emphasis being placed on: Global HPL, Global RandomAccess, EP STREAM (Triad) per system and Global FFT. This award would be weighted 50% on performance and 50% on code elegance, clarity, and size. Both will be determined by an evaluation committee. For this award, the implementer must submit to hpcc-awards @cs.utk.edu (by October 27th, 2006) a short description of:

- the implementation description,
- the performance achieved,
- lines-of-code,
- and the actual source code of their implementation.

See also:

- Class 2 Official Specification
- Class 2 QuickRef Document
- Class 2 Additional Rules
- FAQ

The evaluation committee will select a set of finalists who will be invited to give a short presentation during the HPC Challenge Award BOF at SC06. This presentation will be judged by the evaluation committee to select the winner. The prize will be $1500 plus a certificate for this award and may be split among the "best" entries. The awards will be presented at the HPC Challenge Award BOF at SC06 and the prizes are sponsored by HPCWire.

The following are the winners of the 2006 HPC Challenge Class 1 Awards:

G-HPL	Achieved	System	Affiliation	Submitter
1st place	259 Tflop/s	IBM BG/L	DOE/NNSA/LLNL	Tom Spelce
1st runner up	67 Tflop/s	IBM BG/L	IBM T.J. Watson	John Gunnels
2nd runner up	57 Tflop/s	IBM p5-575	LLNL	Charles Grassl
G-RandomAccess	**Achieved**	**System**	**Affiliation**	**Submitter**
1st place	35 GUPS	IBM BG/L	DOE/NNSA/LLNL	Tom Spelce
1st runner up	17 GUPS	IBM BG/L	IBM T.J. Watson	John Gunnels
2nd runner up	10 GUPS	Cray XT3 Dual	ORNL	Jeff Larkin
G-FFT	**Achieved**	**System**	**Affiliation**	**Submitter**
1st place	2311 Gflop/s	IBM BG/L	DOE/NNSA/LLNL	Tom Spelce
1st runner up	1122 Gflop/s	Cray XT3 Dual	ORNL	Jeff Larkin
2nd runner up	1118 Gflop/s	Cray XT3	SNL	Courtenay Vaughan
EP-STREAM-Triad (system)	**Achieved**	**System**	**Affiliation**	**Submitter**
1st place	160 TB/s	IBM BG/L	DOE/NNSA/LLNL	Tom Spelce
1st runner up	55 TB/s	IBM p5-575	LLNL	Charles Grassl
2nd runner up	43 TB/s	Cray XT3	SNL	Courtenay Vaughan

The following are the winners of the 2006 HPC Challenge Class 2 Awards:

Award	Recipient	Affiliation	Language	PDF
Best Overall Productivity	Bradley Kuszmaul	MIT CSAIL	Cilk	PDF
Best Productivity in Performance	Calin Cascaval	IBM	UPC	PDF
Best Productivity and Elegance	Cleve Moler	MathWorks	Parallel MATLAB	PDF
Best Student Paper	Vadim Guzev	Russian People Friendship University	MC#	PPT
Honorable Mention	Brad Chamberlain	Cray Inc.	Chapel	PDF

USC Viterbi School of Engineering

Measure Real Codes

ISI Information Sciences Institute

Top Crunch: track the aggregate performance trends of high performance computer systems and engineering software

The TopCrunch project was initiated to track the aggregate performance trends of high performance computer systems and engineering software. Instead of using a synthetic benchmark, actual engineering software applications are used with real data and are run on high performance computer systems. The data are available for download in the form of data files for our current software suite. With time, we expect to track the evolution of delivered performance as a function of enhnacements in both software algorithms and hardware. The results of the benchmarks are available as submitted, and may be searched by data, code name, and year. Summaries and overall rankings are posted twice per year following the precedent set by TOP500.

Latest Downloads

Upload Date	Problem Name	Code Version
03/07/06	car2car	LS-DYNA
02/16/06	Parallel_Implicit	ANSYS, Inc.
10/01/03	3 Vehicle Collision	LS-DYNA

Latest Results

Upload Date	Problem	Code Version
12/06/06	3 Vehicle Collision	LS-DYNA
12/06/06	3 Vehicle Collision	LS-DYNA
12/06/06	3 Vehicle Collision	LS-DYNA

News

- *May 25, 2006* - Benchmark results can now be exported as XML or tab delimited files.
- *September 1, 2005* - First dual core processor benchmark results now available.
- *April 2005* - ANSYS, Inc. has contributed a parallel implicit finite element benchmark problem to Top Crunch. It is available in the downloads section.

USC

USC Viterbi School of Engineering

Top Crunch Results

ISI Information Sciences Institute

TOP CRUNCH

FAQ NEWS RESULTS DOWNLOADS PLOT SERIES

BENCHMARK RESULTS

Submit New Result

Benchmark Result(s) by Code Name: LS-DYNA
Result(s) Found: 87

- *Click on problem name, e.g. "3 Vehicle Collision", for additional benchmark details*
- *Click on column header to sort results*

Vendor/ Submitter Org.	Computer/ Interconnect	Processor	#Nodes x #Processors per Node x #Cores Per Processor = Total #CPU	Time (Sec)	Benchmark Problem	Submission Date
CRAY Inc./ CRAY Inc.	CRAY XD1/ RapidArray	AMD Dual Core Opteron 2.2 GHZ	64 x 2 x 2 = **256**	1696	3 Vehicle Collision	09/02/2005
Rackable Systems/ AMD	Emerald/ PathScale InfiniPath / Silverstorm InfiniBand switch	AMD Dual-Core Opteron Model 275 (2.2GHz)	64 x 2 x 2 = **256**	1988	3 Vehicle Collision	11/04/2005
IBM/IBM	eserver p5 575/eserver HPS	POWER5, 1.9 GHz	32 x 8 x 1 = **256**	2058	3 Vehicle Collision	08/19/2005
CRAY Inc./ CRAY Inc.	CRAY XD1/ RapidArray	AMD Dual Core Opteron 2.2 GHZ	64 x 2 x 1 = **128**	2135	3 Vehicle Collision	09/14/2005
CRAY Inc./ CRAY Inc.	CRAY XD1/ RapidArray	AMD Dual Core Opteron 2.2 GHZ	32 x 2 x 2 = **128**	2416	3 Vehicle Collision	09/02/2005

HP/HP	Opteron CP4000/ Topspin InfiniBand	AMD Opteron 2.6 GHz DL145	1 x 2 x 1 = **2**	81810	3 Vehicle Collision	08/25/2005
HP/HP	Opteron CP4000/ Voltaire	Dual Core Opteron 2.2 GHz DL145	1 x 2 x 1 = **2**	89555	3 Vehicle Collision	11/15/2005
HP/HP	Opteron CP4000/ Myrinet	AMD Opteron 2.4GHz DL145	1 x 2 x 1 = **2**	90271	3 Vehicle Collision	07/28/2005
CRAY Inc./ CRAY Inc.	CRAY XD1/ RapidArray	AMD Opteron 2.4 GHZ	1 x 2 x 1 = **2**	92228	3 Vehicle Collision	05/12/2005
HP/HP	Opteron CP4000/ Voltaire	Dual Core Opteron 2.2 GHz	1 x 1 x 2 = **2**	99578	3 Vehicle Collision	11/15/2005
HP/HP	Opteron CP4000/ InfiniBand Voltaire/ InfiniBand Voltaire	AMD Opteron 2.2GHz DL145	1 x 2 x 1 = **2**	104384	3 Vehicle Collision	01/10/2005

Search Benchmark Results by:

File Name: --- All Files ---

Code Name: --- All Codes ---

Total #CPU --- All ---

Year: select

Send Results to: Screen

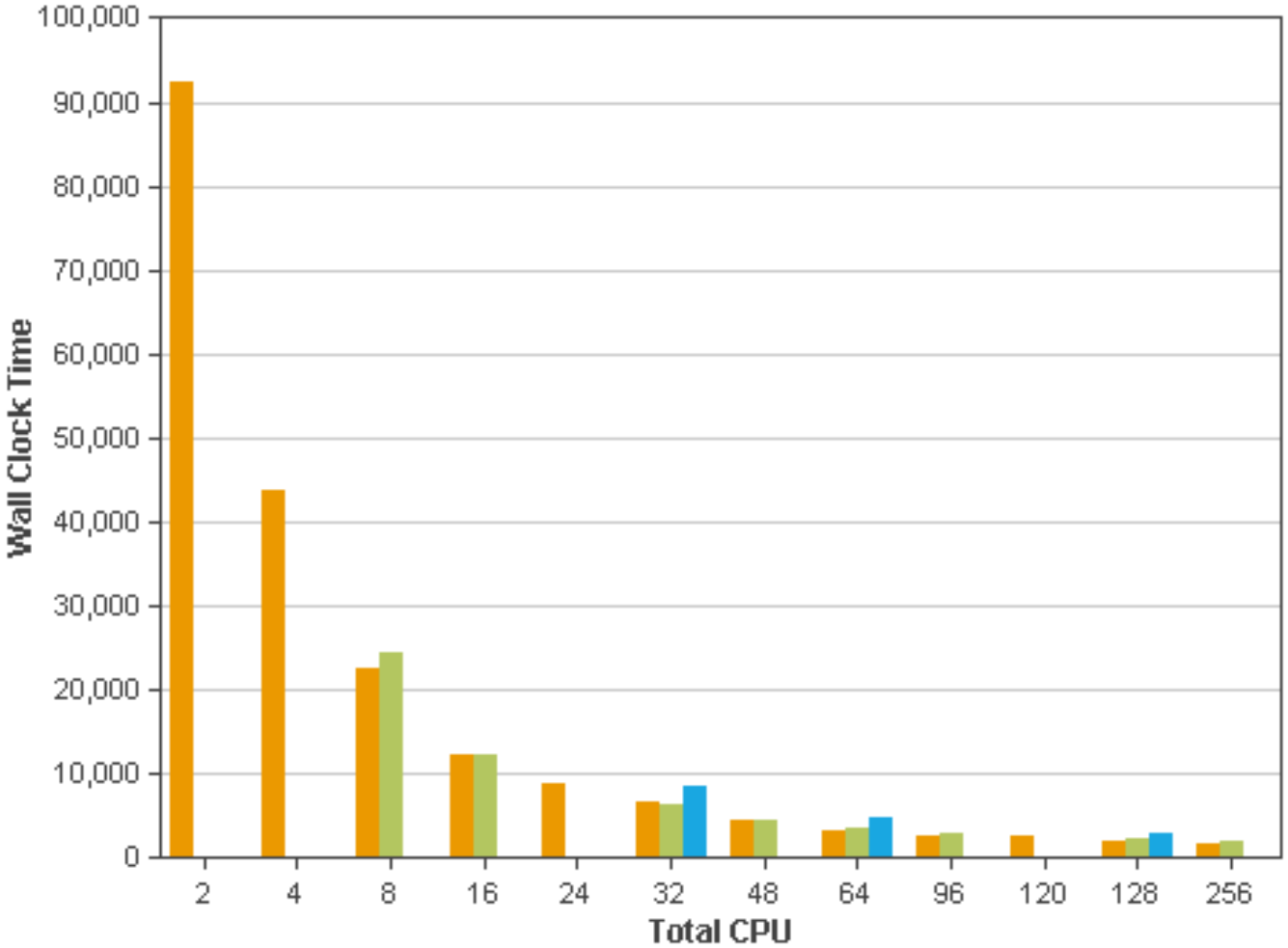
Search results may be printed to the screen as an HTML file, an XML file, or a tab delimited file that can be saved and imported to a spread sheet application such as MS Excel.

Search Reset

Administrator | Home | Webmaster

Supported by DARPA HPCS

Top Crunch Results (cont.)



USC Viterbi School of Engineering

TopCrunch Emulates TOP500

ISI Information Sciences Institute

Domain experts specify the problem

Let vendors and users run the code

Collect results and publish on the Web

Track performance vs. time

- "Linpack" has four decades of data

We need to get more applications

- More "TopSomethings"

USC

USC Viterbi School of Engineering

Execution on Future Systems

ISI Information Sciences Institute

- Success predicting individual processors
  - UCSD/SDSC
- Success predicting scaled systems
  - LANL
- Research Problem
  - Modeling applications and predicting performance on non-existent H/W

USC

USC Viterbi School of Engineering

New Language Thoughts

Information Sciences Institute

- I'm a shared memory bigot
  - Suffered 20 years and 6 months of "MPI"
- I like UPC a lot better than "MPI"
  - But its only a small step forward
  - Minimal extension of C
    - Two-level memory hierarchy
- Evolve from familiar languages
  - Retain old code
  - Retain old programmers

USC

USC Viterbi School of Engineering

Productivity Impact?

ISI Information Sciences Institute

- Anecdotes aren't enough
  - People at Sandia like "MPI"???
- SLOC is an poor measure
  - What Churchill said of democracy
- "Time to solution" hard to define
  - Most codes are slowly evolving
- How can leadership choose?
  - Shared memory vs. maximum Flop/s

USC

USC Viterbi School of Engineering

Measuring Development Time Productivity

ISI Information Sciences Institute

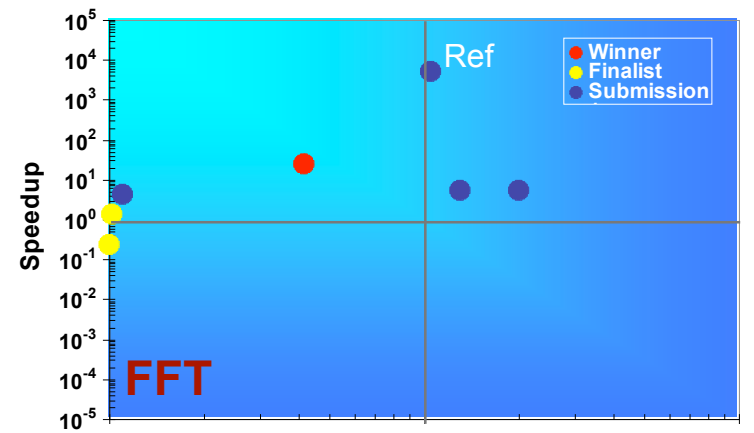
- HPCS Challenge Type 2 Awards
- Code Team Surveys
- Experiment with human subjects
  - Mainly classrooms so far
- Scalable Synthetic Compact Applications
  - Allow bigger experiments

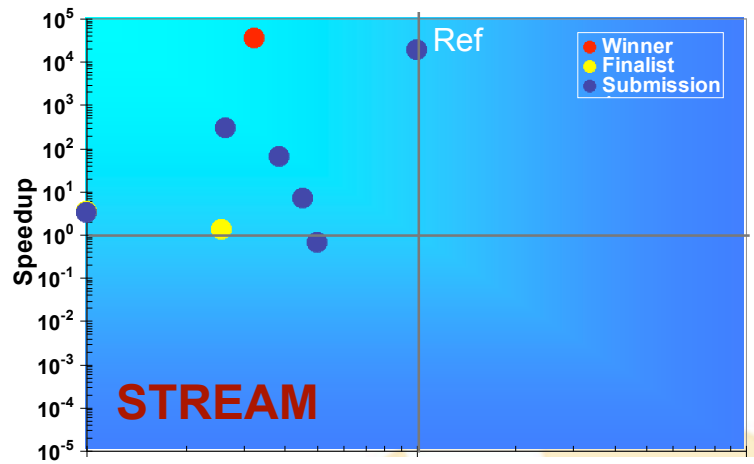
USC

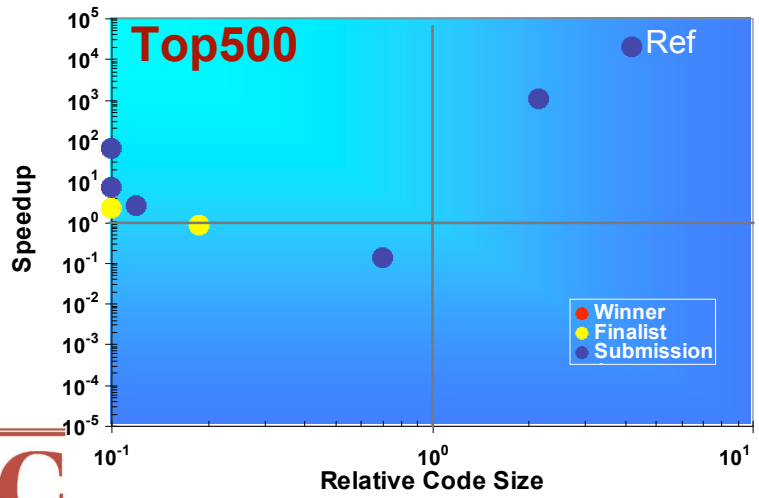
USC Viterbi School of Engineering

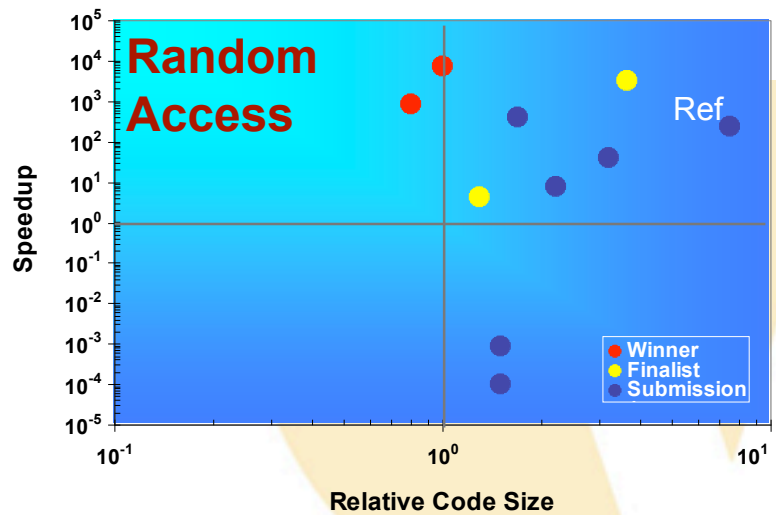
HPC Challenge Beauty Contest

ISI Information Sciences Institute









USC

USC Viterbi School of Engineering

Survey Real Code Teams

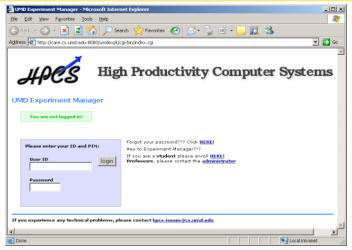
Information Sciences Institute

An obvious thought is to go ask people

	Falcon	Hawk	Condor	Eagle
Identify Project	√	√	√	√
Negotiate study with sponsor	√	√	√	√
Complete pre-interview questionnaire	√	√	√	√
Analyze questionnaire and plan on-site interview	√	√	√	√
Conduct on-site interview	√	√	√	√
Analyze on-site interview and integrate with questionnaire	√	√	√	√
Conduct follow-up to resolve unanswered questions	√	√	√	√
Write report	√	√	√	√
Present/publish report	√	√	√	

USC

Human Subject Experiments



Local server: capture data

UMD Server care: store data

UMD server hpcs+care: analyze data

Technician

Install

Upload

Professor

Create a course

Monitor registration

Write/run code

Student

Sign up for account/key

Manual online logs questionnaire

HPC Machine

Umdinst + Hackystat Sensors

Local Log

Upload

UM Experiment Manager (EM) + Hackystat Server

UM Admin

Master DB

Sanitized data

HPDBugBase (defect data)

UM Workflow Tool

Data Analysis Interfaces

SQL Queries

Sanitized DB

UM Data Analysis Environment

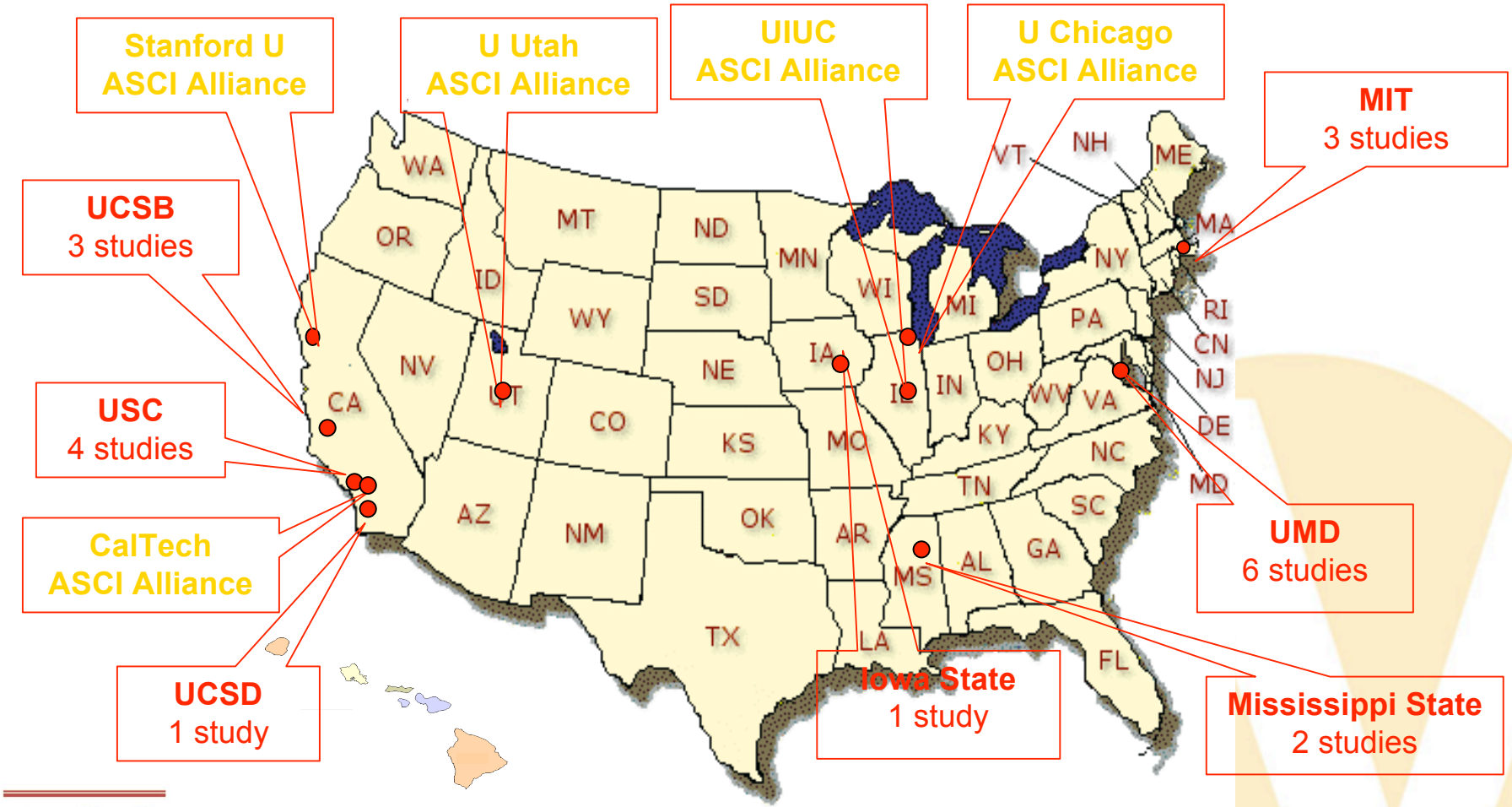
Upload Develop Analysis Tool

Download Analysis Results

USC Viterbi School of Engineering

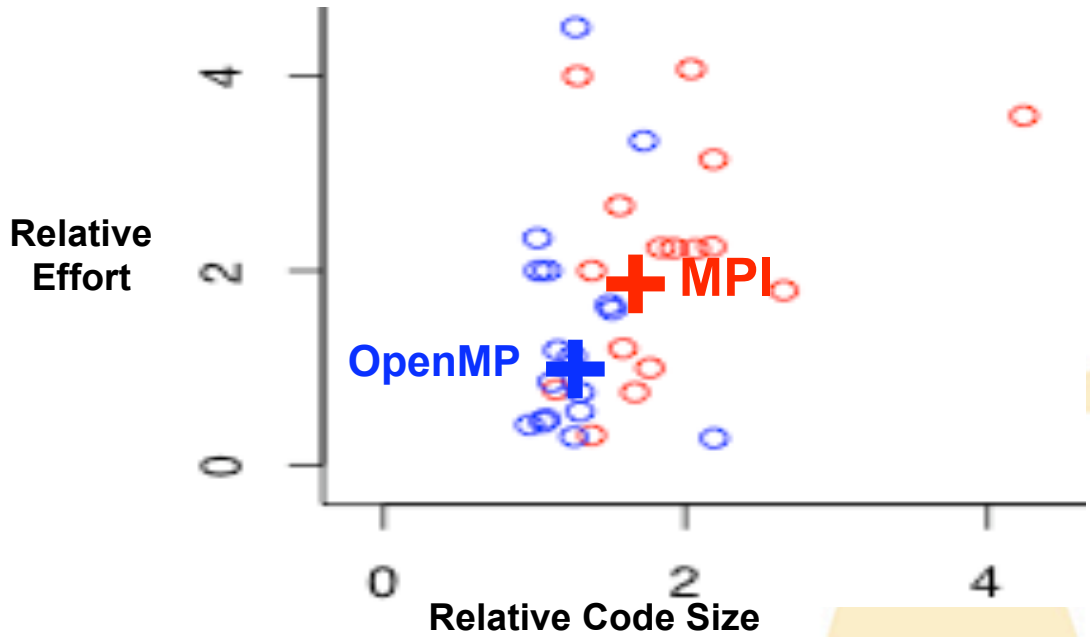
In Lots of Classrooms

ISI Information Sciences Institute



USC

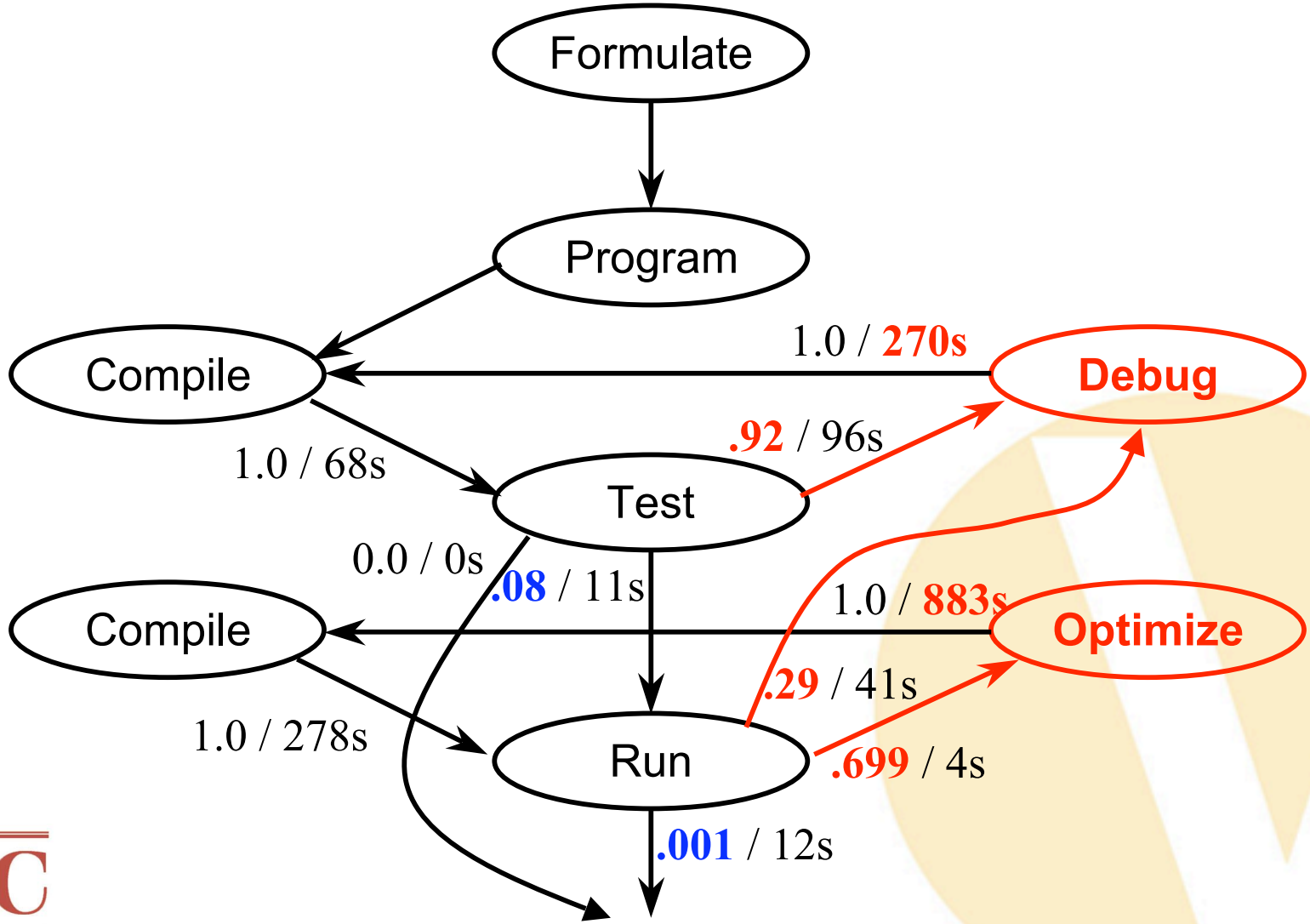
Classroom-Scale Results



Suggests OpenMP requires less effort

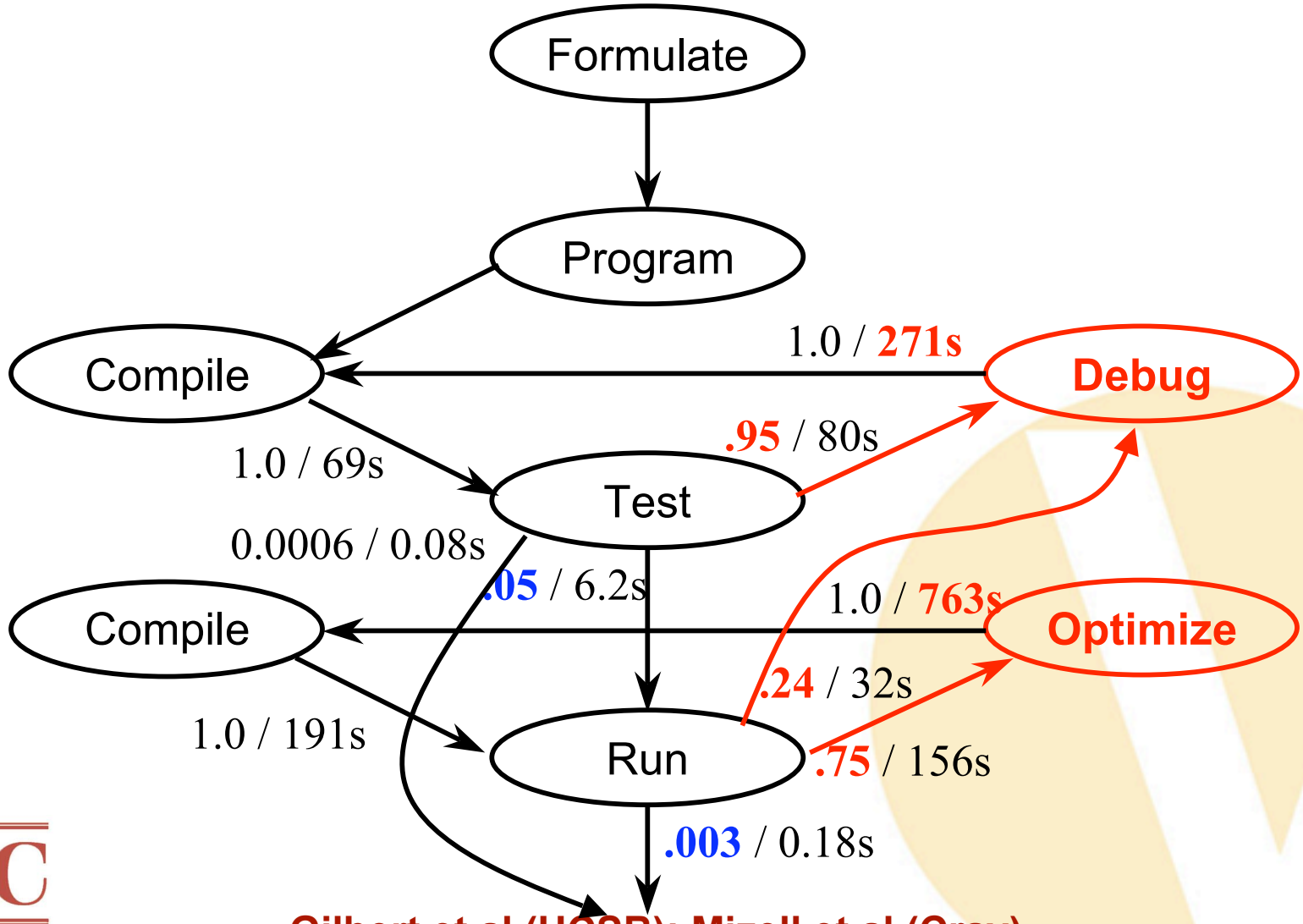
OpenMP code averaged fewer SLOC

More Classroom Results (MPI)



Gilbert et al (UCSB); Mizell et al (Cray)

More Classroom Results (UPC)



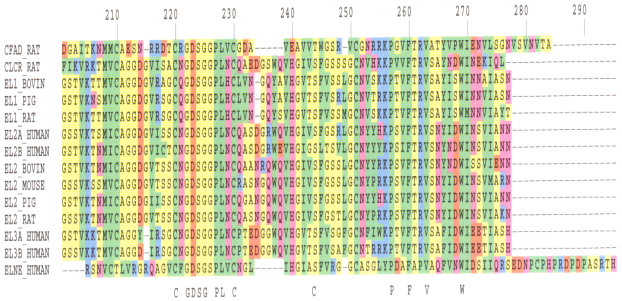
Gilbert et al (UCSB); Mizell et al (Cray)

Scalable Synthetic Compact Applications (SSCA)

USC Viterbi School of Engineering

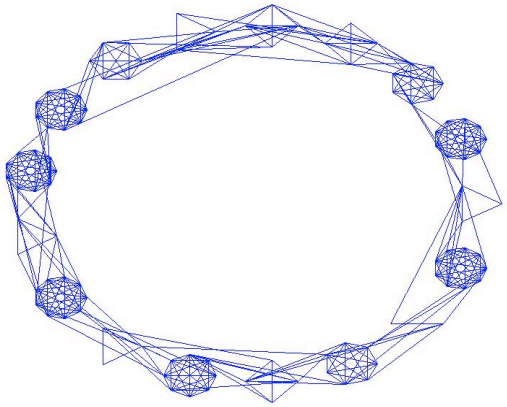
ISI Information Sciences Institute

Bioinformatics

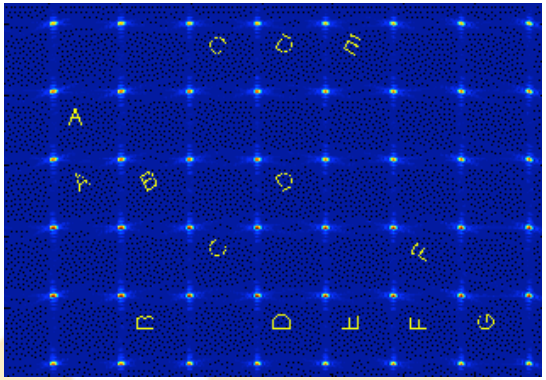


Protein Alignment

Graph Theory



Sensor and IO



Bigger than a homework assignment

Proxies for real codes

UMT2000 is another example

Small enough for coding experiments

Plan to make more of them …

USC

USC Viterbi School of Engineering

Summary

ISI Information Sciences Institute

Applications

- Go after the new ones (next talks)
- Recover some old ones too

Execution Time Productivity

- Measure benchmarks and real applications
- At least in HPCS phase 3 time-frame

Development Time Productivity

- On-going research project
- We're making progress

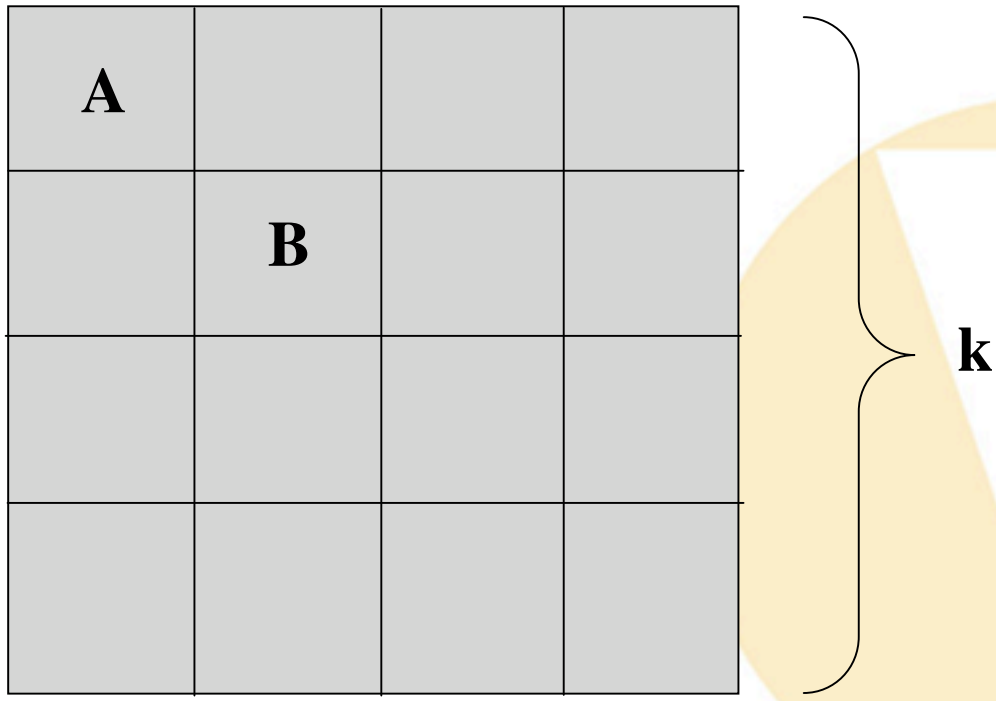
USC

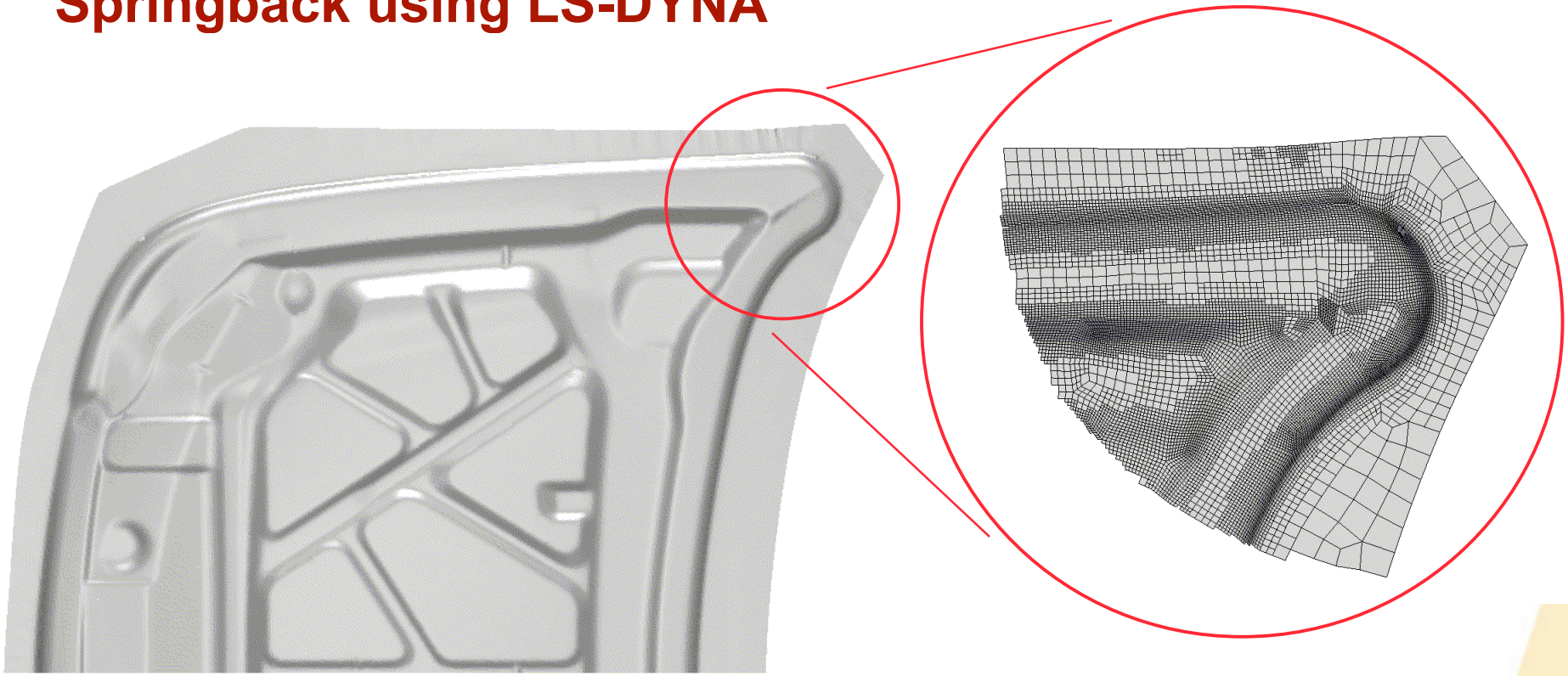
Bonus

Load Balance

Partition k-by-k grid by nested dissection

What is the ratio of LU factorization work for processors assigned to the A and B domains?



Unstructured Grid

Automotive Hood Inner Panel Springback using LS-DYNA