Overview of activities of the Application Scalability and Performance (ASAP) group

John Dennis (dennis@ucar.edu)

Computational & Information Systems Laboratory

CISL

NCAR

Application Scalability and Performance Group

- **CISL Research group with 3 focuses**
  - Scalability Applications (Scale)
  - Accelerators and Micro-processor performance (Accel)
  - Workflow and I/O (WIO)
- **Staff:**
  - Allison Baker (WIO, Scale)
  - John Dennis (Accel, Scale, WIO)
  - Ben Jamroz (Accel, Scale)
  - Youngsung Kim (Accel)
  - Sheri Mickelson (WIO)
  - Kevin Paul (WIO)
  - Srinath Vadlamani (Accel)
  - Haiying Xu (WIO)

Outline

- Data-compression
- CESM on Xeon Phi
- Performance Enhancement Methodology
- Conclusions

Data compression

<u>Components</u> Compression: $X \Longrightarrow C$

Reconstruction: $C \Longrightarrow \tilde{X}$

<u>Types</u> **Lossless** (no info is lost) : $X = \tilde{X}$

Lossy (some loss of info) : $X \sim \tilde{X}$

<u>Example</u> 8-byte → 4-byte

Original: T = 290.1234567890123

Reconstructed: T = 290.1234500000000

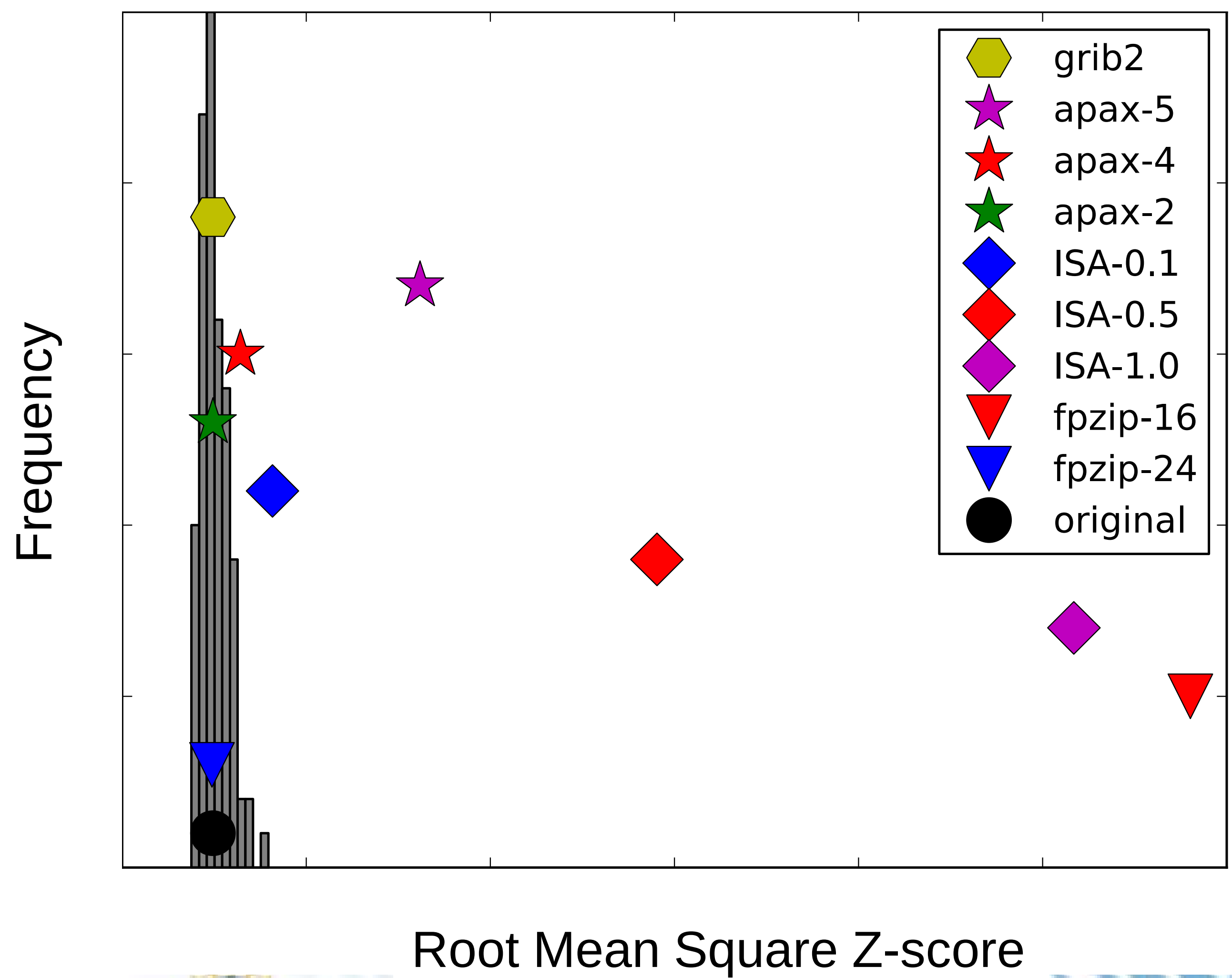
<u>Evaluate $\tilde{X}$ in the context of an ensemble</u>:

NEW

- 101 one-year CESM runs
- Double-precision perturbation in initial ATM temperature
- Creates an "accepted" distribution
- 1-deg atmosphere model: 170 variables

Is the difference "climate-changing ….

RMSZ ensemble test



CAM Data Compression

For each variable (170), choose highest compression rate (CR) such that:

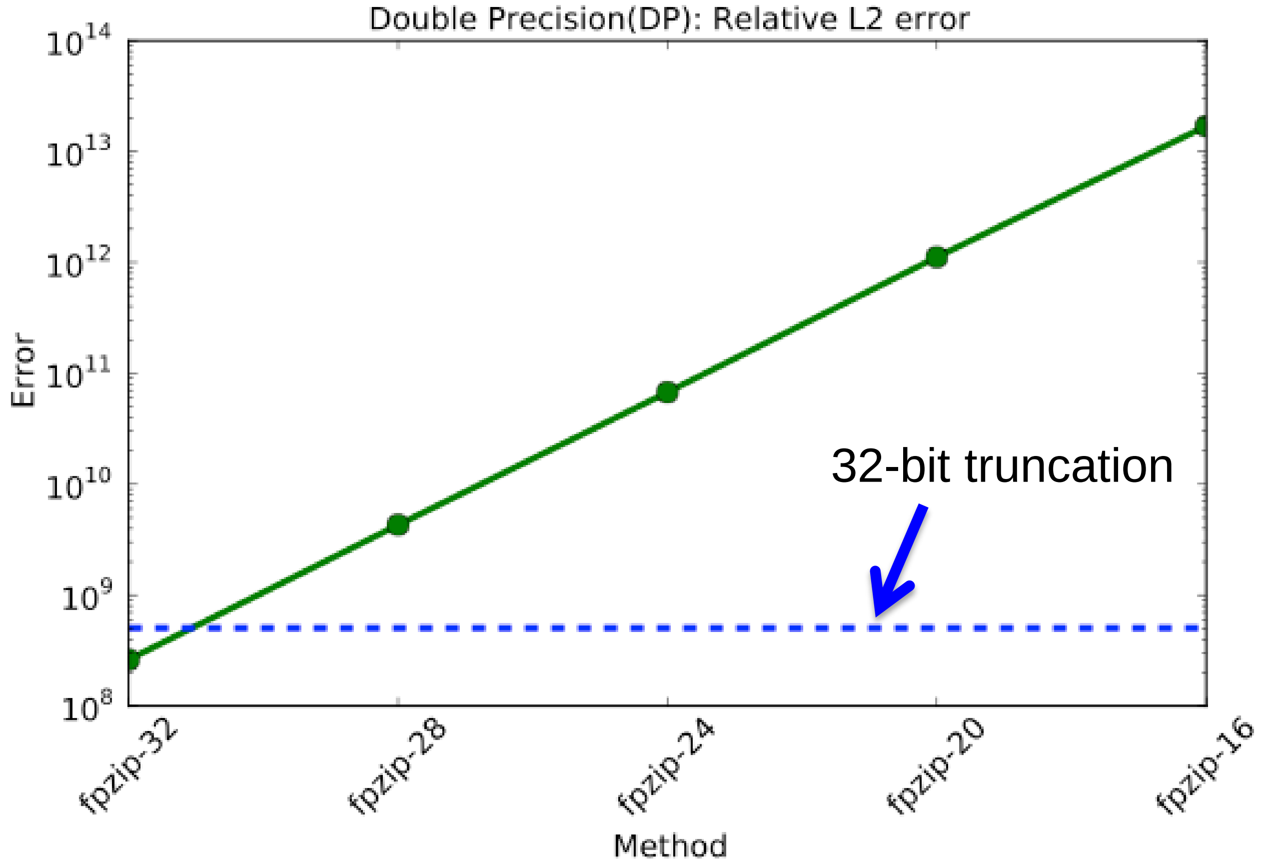
- RMSZ-ensemble test ✔
- Max-error ensemble test ✔
- RMSZ-bias test ✔
- Correlation coefficient test ✔

CR	GRIB2	ISABELA	fpzip	APAX
average	.37	.42	.18	.29
best	.03	.20	.02	.06
worst	.86	.77	.68	.80

POP Data Compression

- **How to create an ensemble (?)**
- **Derived vector quantities: Temperature tendency**
- **Compare internal calculation with reconstructed**
  - 64-bit output
  - 32-bit output
  - Compressed with fpzip

Relative L2 error: Temperature tendency



Computational & Information Systems Laboratory

CISL

NCAR

Data-Compression: Ongoing work:

- **Evaluating new compression algorithms**
  - I. Horenko, W. Sawyer (CSCS) temporal compression
  - K. Sato, N. Sasaki, et al (Titech) wavelet based method
  - L. Gomez (ANL), preconditioner based
  - P. Lindstrom (LLNL), revised fpzip
  - S. Liu, X. Huang, et al (Tsinghua U.)

- **Blind evaluation of data-compression**
  - Large ensemble project: 30 member ensemble (www.cesm.ucar.edu/experiments/cesm1.1/LE/)
  - Add 3 additional members: 1-2 of which have been compressed
  - Can climate scientists tell which member was compressed?

Overview

- Climate model verification
- Data compression
- Workflow
- Many-core & CESM
- Performance enhancement methodology
- KGEN

Current Status of CESM on Phi

- **Workaround provided for Intel compiler bug**

  -mP2OPT_hpo_matrix_opt_framework=0

  - Fixed in Intel compiler 14 update 3
- **Verification Status:**
  - CAM5, ideal physics, ne16 (2 deg) 3 members passed
  - CAM5, full physics, ne30 (1 deg) 1 member passed
    - 6/122 global mean tests failed 4% outside range
    - Used ~4500 core hours

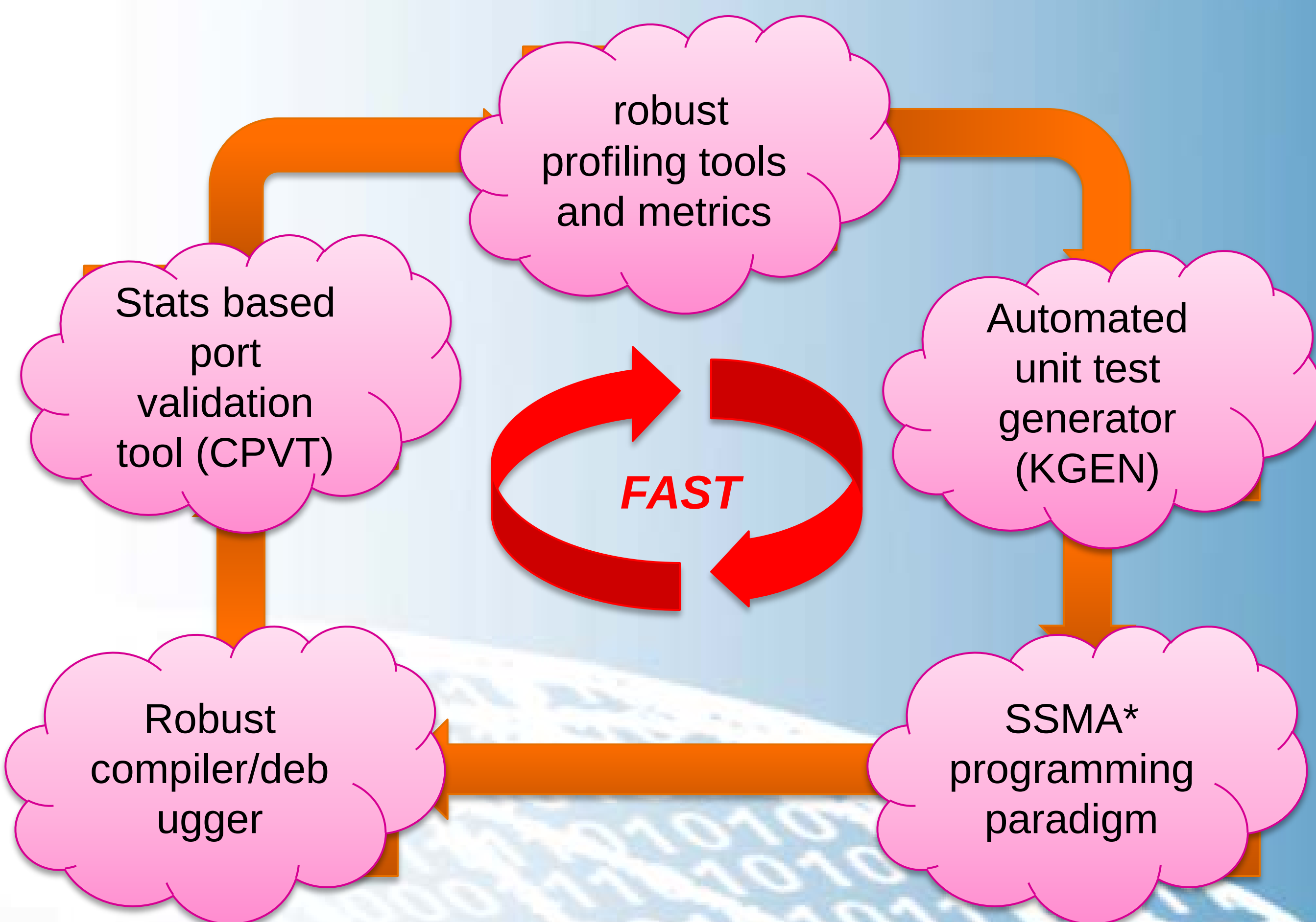
Current Status of CESM on Xeon Phi (con't)

- **Requires small number of changes to CESM code based to address Xeon Phi compiler**
- **Upcoming CESM development tag**
  - Stampede (TACC)
  - Babbage (NERSC)
  - Pronghorn (NCAR)

CESM performance on Phi

- **Configuration:**
  - CESM, full physics, ne16 (2 degree)
  - Native mode
  - 2 nodes of Stampede
    - Host: 4 sockets, 8 cores per socket
    - Phi: 2 KNC cards
  - Simulation rate (without I/O): 29%

Performance Enhancement Methodology: a virtuous cycle for code improvement

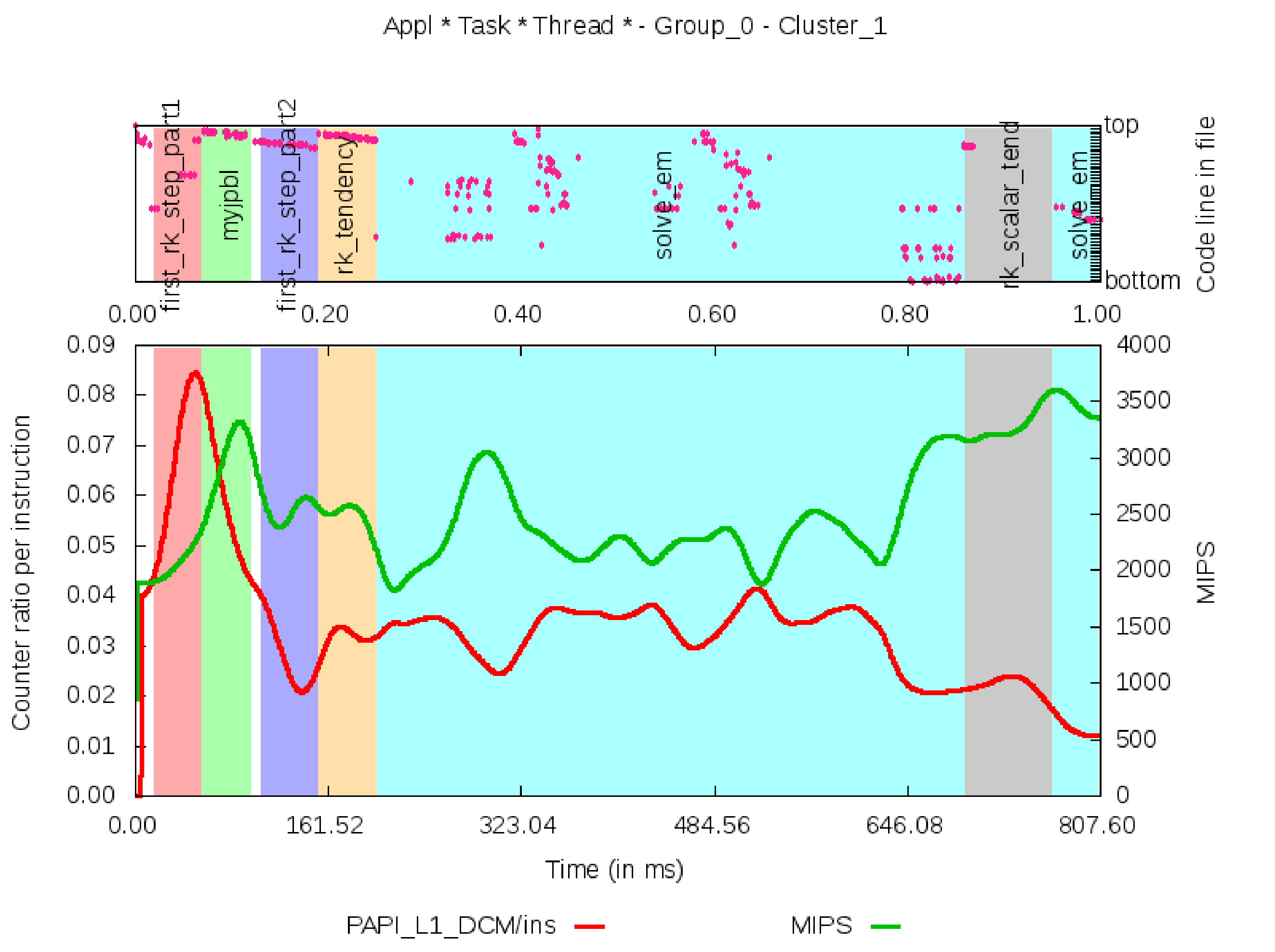


*single source multiple architecture

Identifying optimization targets with Extrae and Paraver.

- **Extrae tracing tool developed at Barcelona Supercomputer Center**
  - H. Servat, H. Labart, J. Gimenez
  - Automatic performance identification is a BSC research project.
  - Produces a time series of communication & hardware counter events.
- **Paraver is the visualizer that also performs statistical analysis.**
  - There are clustering techniques which uses a folding concept plus the research identification process to create “synthetic” traces with fewer samples.
- **Has been applied to multiple codes**
  - CESM
  - WRF
  - PORT: standalone RRTMG driver
  - HOMME: dynamical core of CAM-SE

Performance Analysis tools: Antarctic Mesoscale Prediction System (AMPS)



KGEN

- **Extracts a Fortran subprogram by placing “OpenMP-like” directives.**

EX:
```
!KGEN BEGIN
 SUBROUTINE sub
…
END SUBROUTINE sub
```

- **Scans source files to collect parameters, hierarchy of derived types, and stack of subprogram calls, and then generates following files:**
  1. Kernel template file
  2. Modified input source file for data generation

KGEN Usage Example

HOMME* Source

```
derivative_mod.F90
…
!KGEN BEGIN
!KGEN
SAVE(filename=foo.dat;counter_at=100;mpi_rank_at=0)
FUNCTION vlaplace_sphere_wk(…) result(laplace)
…
END FUNCTION
…
```

KGEN

Kernel template

```
program kgen_kernel_vlaplace_sphere_wk
…
Kgen_laplace = vlaplace_sphere_wk(…)
…
CONTAINS
…
FUNCTION vlaplace_sphere_wk(…)
result(laplace)
…
END FUNCTION
…
```

KGEN

Modified input source file

```
Modified derivative_mod.F90
…
FUNCTION vlaplace_sphere_wk(…)
result(laplace)
…
WRITE(unit=n) varA
…
WRITE(unit=n) laplace
…
END FUNCTION
…
```

Execute HOMME

Runtime data

vlaplace_wk.dat.100.0

Kernel

HOMME*: HIGH-ORDER METHODS MODELING ENVIRONMENT

Conclusions:

- Impact of lossy compression not distinguishable from natural variability
- CESM data compressed by as much as **5:1**
- Blind evaluation of data-compression through large ensemble project
- CESM port to Xeon Phi verified
- CESM on Xeon Phi: 29% of Sandybridge
- Developed a Performance Enhancement methodology

A.H. Baker, et al

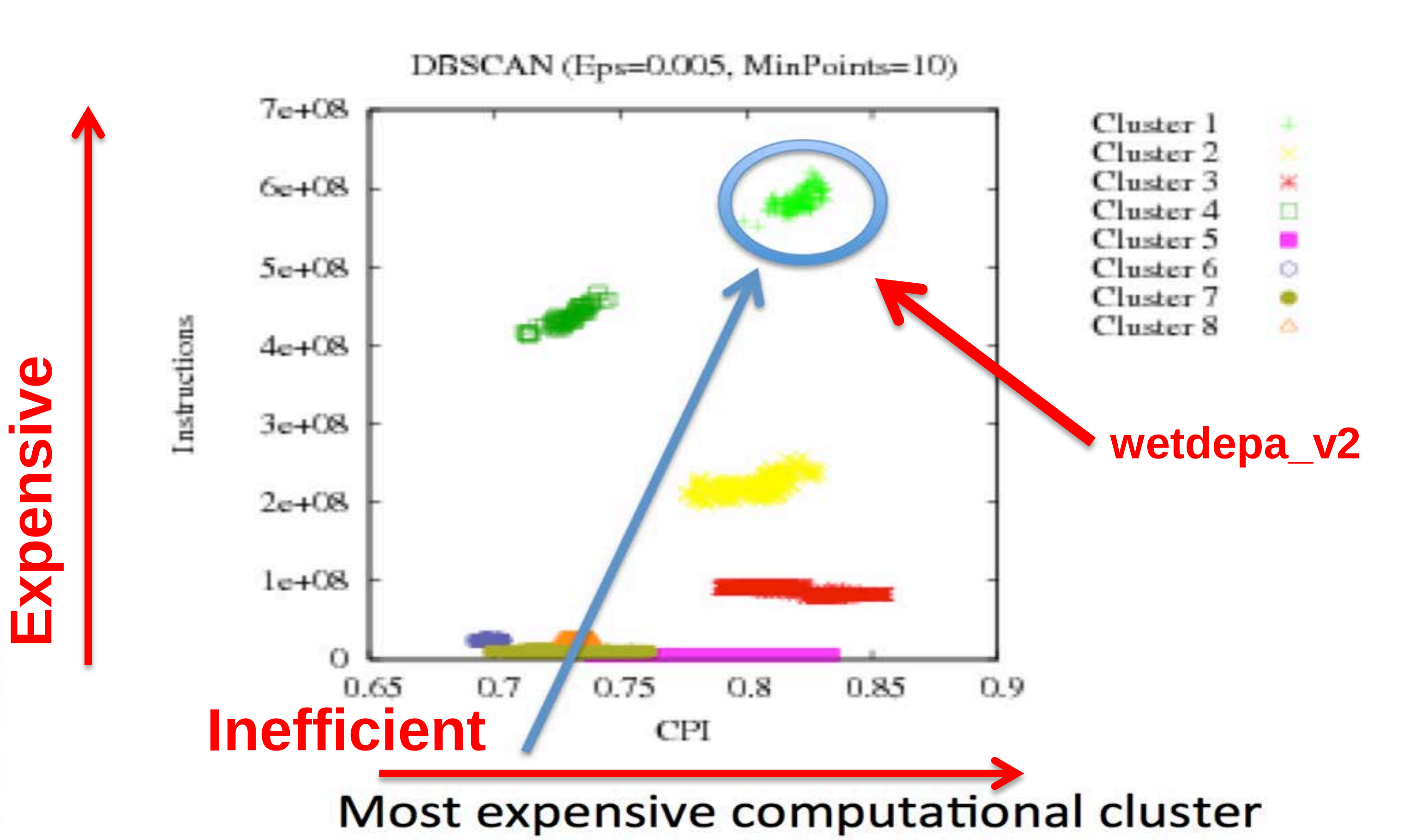
"A Methodology for Evaluating the Impact of Data Compression on Climate Simulation Data."
Proc. of the 23rd International ACM Symposium on High Performance Parallel and Distributed Computing (HPDC14), Vancouver, CA, June 2014 (*to appear*).

Computational & Information Systems Laboratory

Questions?

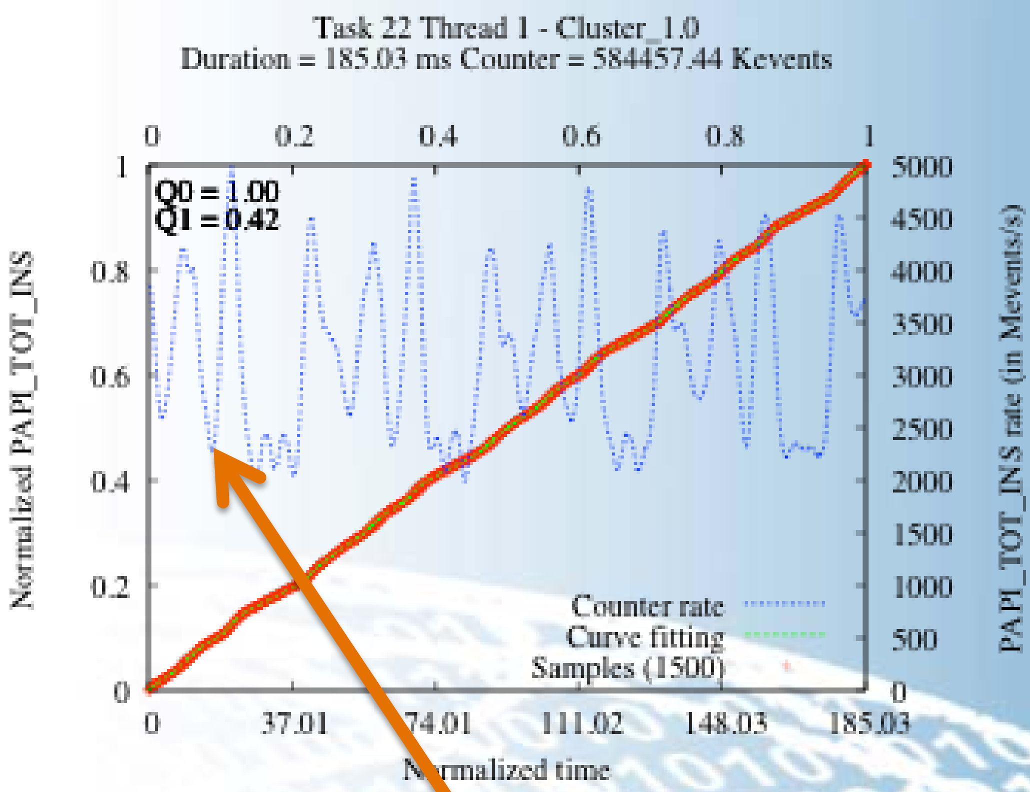
John Dennis (dennis@ucar.edu)

BSC tools helped us to find high priority sections that are expensive *and* inefficient.



- Result of an Extrae trace of CESM on Yellowstone.
- Similar to exclusive execution time.

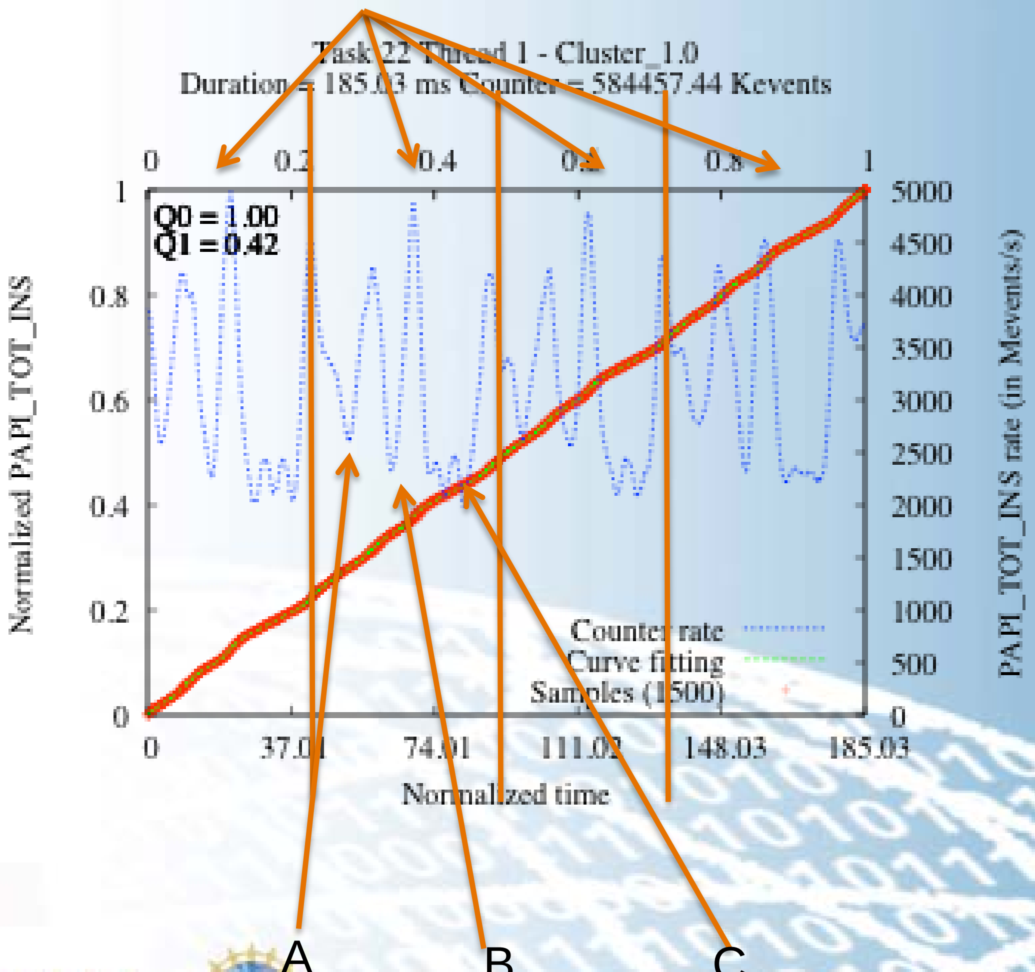
Total Instructions: Cluster 1



Notice drops in Instruction rates

Total Instructions: Cluster 1

4 cycles in Cluster 1



Underperforming subroutines Cluster 1

- **Group A:**
  - conden: 2.7%
  - compute_uwshcu: 3.3%
  - rtrnmc: 1.75%
- **Group B:**
  - micro_mg_tend: 1.36% (1.73%)
  - wetdepa_v2: 2.5% ← Focus effort on one subroutine
- **Group C:**
  - reftra_sw: 1.71%
  - spcvmc_sw: 1.21%
  - vrtqdr_sw 1.43%

Optimizing (vectorizing) wetdepa_v2

- **Consists of a double nested loop**
  - Very long ~400 lines
  - Unnecessary branches with inhibit vectorization
- **Restructuring wetdepa_v2**
  - Break up long loop to simplify vectorization
  - Promote scalar to vector temporaries
  - Common expression elimination

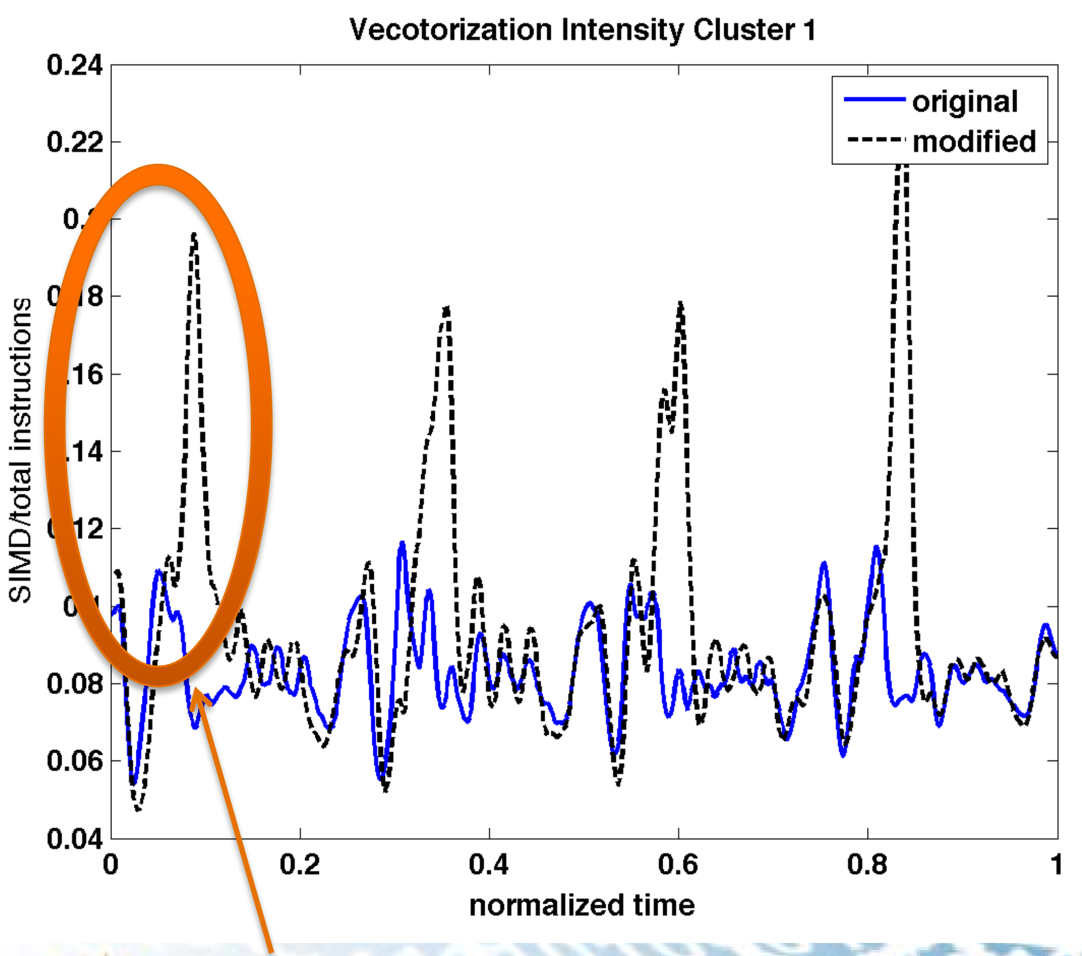
Now use what we've learned from dg_kernel to speed up wetdepa_v2… to get fast and right!

	Intel Phi (Intel 13.1.1)			Intel Sandybridge (Intel 13.1.2)		
	-O2	-O3	-O3 -fast	-O2	-O3	-O3 -fast
orig	42.85	41.24	3.74	3.43	3.32	0.97
mod	6.50	6.61	4.58	1.09	1.12	1.04

- wetdepa is small only ~600 lines
- Restructured branched loops + promoted scalars to vectors.
- -O3 fast for original code gave incorrect results
- 2.5% to 0.7% of code execution time = $222K savings

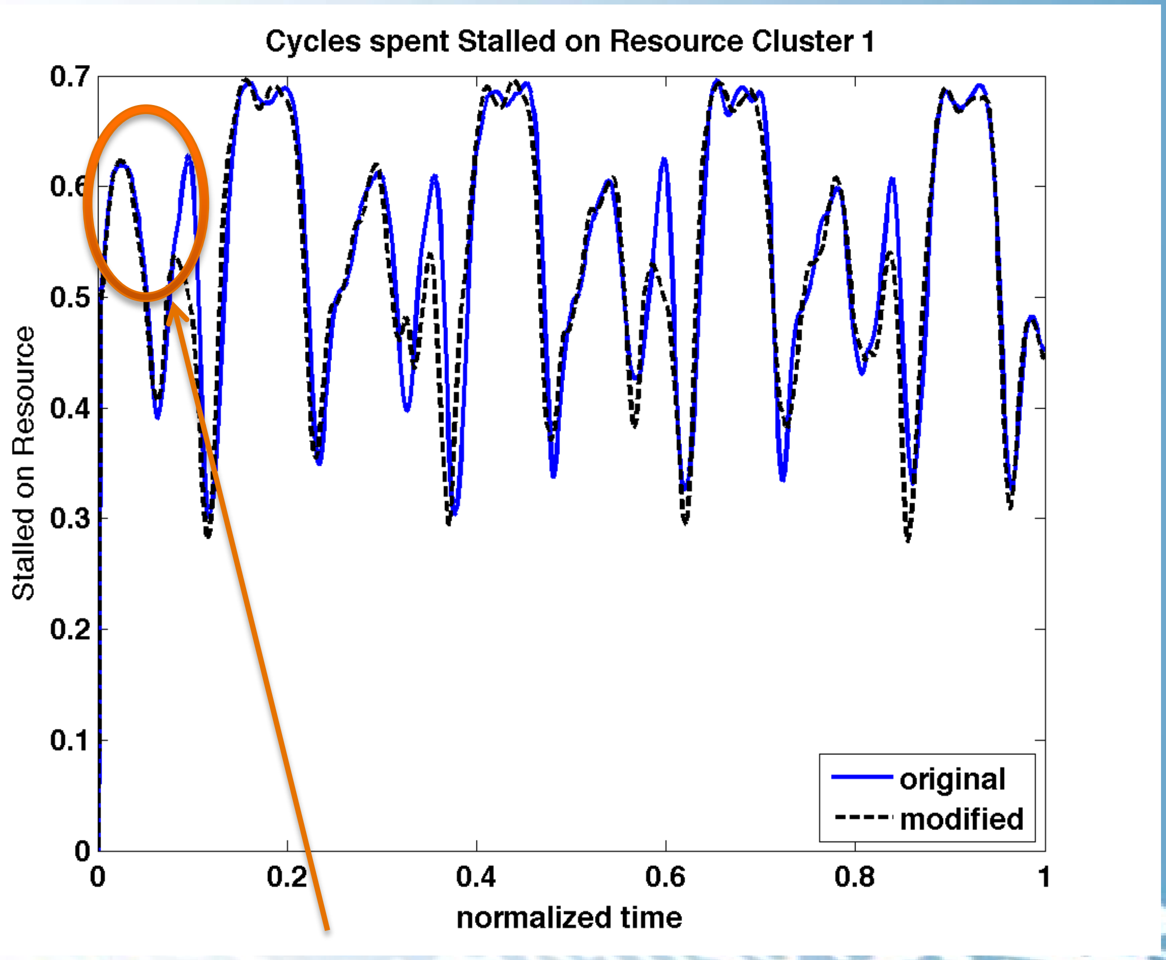
Maybe 2x is possible from code refactoring!

Vectorization Intensity Cluster #1



Increase in code vectorization

Stalls on Resources Cluster #1



Reduction in cycles stalled on resources

CESM/CAM Port-Verification Tool (CPVT)

Create ensemble E, with 101 members

Z-score: compares x_i in ensemble member m to x_i in remaining 100 ensemble members {E\m}:

$$Z_{x_i}^{m} = \frac{x_i^m - \bar{x}_i^{E\setminus m}}{\sigma_{x_i}^{E\setminus m}}$$

Root mean square Z-score for dataset X of ensemble m:

$$\mathrm{RMSZ}_X^m = \sqrt{\frac{1}{N_X}\sum_i \left(Z_{x_i}^m\right)^2}$$

Maximum error: Temperature tendency