



TEXAS ADVANCED COMPUTING CENTER

WWW.TACC.UTEXAS.EDU



TEXAS
The University of Texas at Austin

Profiling Tools for KNL

Intel Xeon Phi Users Group

ISC 2016

Frankfurt, 2016-06-22

PRESENTED BY:

Carlos Rosales

carlos@tacc.utexas.edu

Acknowledgements

Antonio Gomez-Iglesias

- Principal Perfexpert developer,
REMORA developer

Todd Evans

- Principal TACC Stats developer
Contributions to portability
improvements in REMORA IO



National Science Foundation Grant ACI-1134872

This talk is a shameless plug

- Because I will be focusing on tools developed by TACC
- But I also hope to spark a serious discussion
- Profiling support for KNL is reasonably advanced
 - Perf system
 - Intel tools: VTune, Advisor
- But it is a complex processor and many HPC users find regular profiling overwhelming
- This increased complexity trend in HPC is unlikely to change in the near future

The need for a user-centric view

- Researchers prioritize their time differently than staff and developers
 - 5% of their time is **NOT** worth a 5% performance increase. Or 10%. Or 15%.
- If a tool is difficult to use, they simply **WILL NOT** use it
- Difficult to use is a relative term...
 - Data collection
 - Data processing
 - Data interpretation
- We believe there is a need to focus on
 - Collection simplicity
 - Well designed results summaries

Do not underestimate the value of “simple” data

- Hardware is not getting simpler:
 - More elements to consider
 - More data to collect
 - More sources of problems
- But users do not know things like
 - Their memory footprint and their level of locality
 - Their maximum IO rates
 - Their vectorization quality
 - The potential parallelism in their code

REMORA

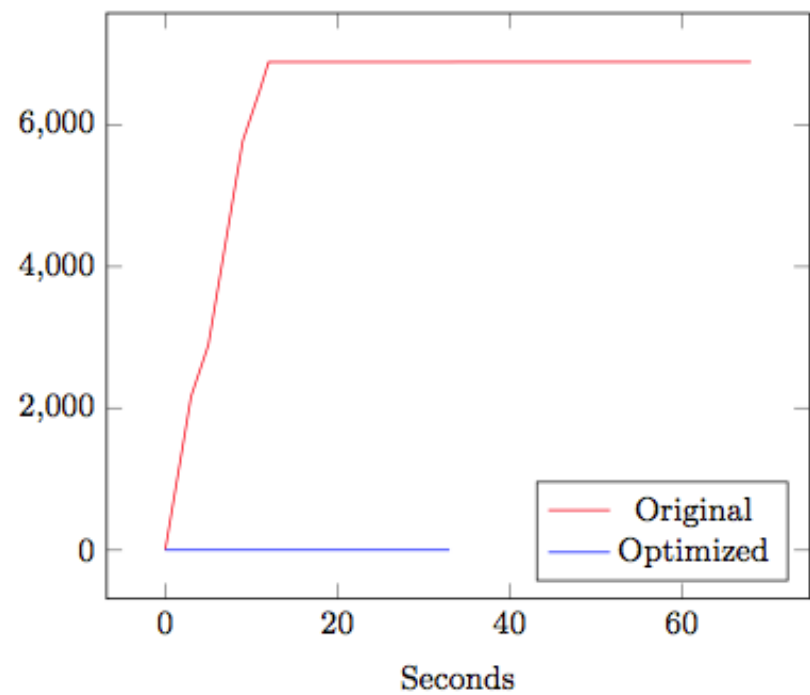
- REsource MOnitoring for Remote Applications
- Monitoring tool for hardware utilization
- Extremely portable and extensible
- Runs and installs in user space
- Measures CPU, memory, IO (Lustre, DVFS), NUMA, network topology
- Provides simplified summary of utilization
- KNL Highlights
 - Monitor MCDRAM / DDR4 utilization using NUMA stats
 - Determine when MCDRAM memory is exceeded
- Coming: resource utilization balance summary, MPI comms

Remora – example output

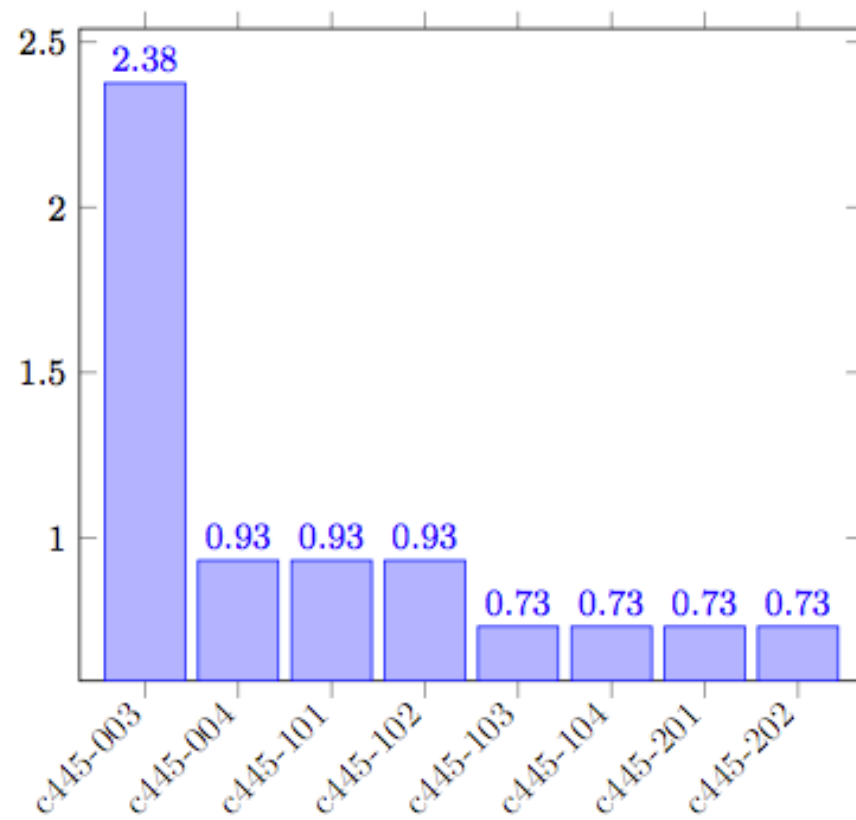
```
===== REMORA SUMMARY =====  
Max Memory Used Per Node      : 3.77 GB  
Total Elapsed Time            : 0d 0h 0m 11s 721ms  
-----  
Max IO Load / scratch         :      310 IOPS      2.3 RD(MB/S)    12.7 WR(MB/S)  
Max IO Load / work            :         0 IOPS       0 RD(MB/S)       0 WR(MB/S)  
Max IO Load / home1           :         0 IOPS      --- RD(MB/S)    --- WR(MB/S)  
=====
```

```
Sampling Period                : 2 seconds  
Complete Report Data           : /scratch/01157/carlos/lbm_bench/bin/remora_171685  
=====
```

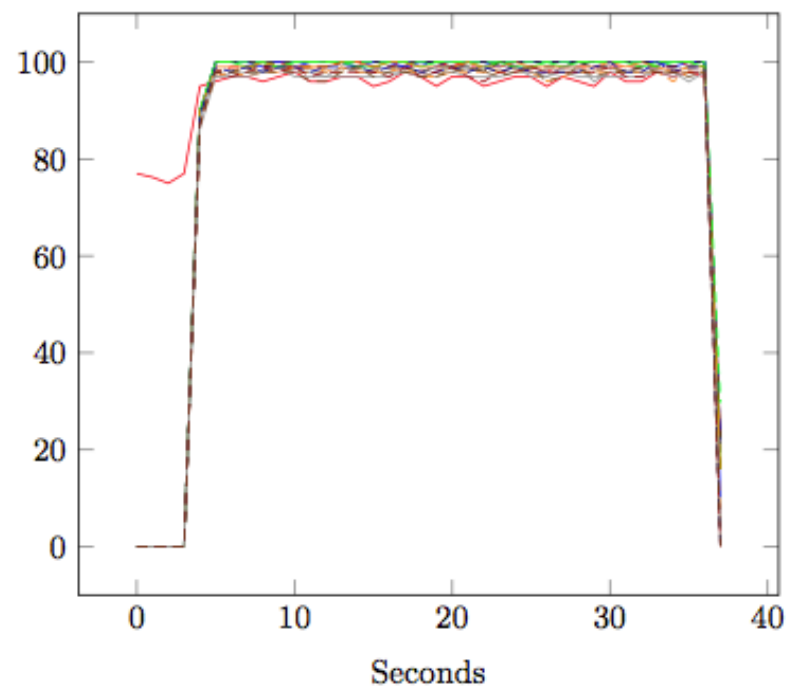
NUMA Other Node



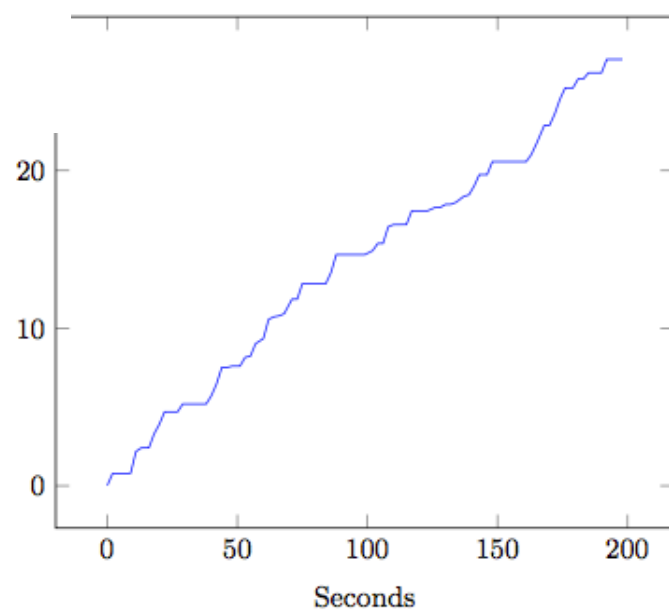
Memory Used GB



CPU Load



Memory Used (GB)



Perfexpert

- Easy to use profiles
 - Collects counter data
- Analyzes the data and creates set of performance metrics
 - Metrics presented to users in an easy to understand manner
- Forecasts optimal degree of parallelism at loop / function levels
- Analyzes vector reports to present users with alternatives in the code
- Interacts with PAPI / HPCTOOLKIT / VTUNE to collect statistics
- Initial support for KNL already implemented

Perfexpert – example output

```
Function IntegrateStressForElems in line 604 of lulesh.cc (1.20% of the total runtime)
=====
Instructions Ratio      %  0.....25.....50.....75.....100
- data accesses        9.5 [>>>>> . . . . .]

Performance Assessment  LCPI  good.....okay.....fair.....poor.....bad
* overall               3.15 [+++++]

```

- Results presented in a qualitative, easy to understand manner
- Output highlights specific code section to investigate

Some Thoughts

- Is there value in adding simplified reports to existing profiling tools (Scalasca, HPCtoolkit, VTune, TAU, ...) ?
- Do you have a tool I have not mentioned that we all should know about?

Links to repositories

REMORA

<https://github.com/TACC/remora>

PerfExpert

<https://www.tacc.utexas.edu/perfexpert>

<https://github.com/TACC/perfexpert>