Power Signatures of High-Performance Computing Workloads

Jacob Combs
Jolie Nazor
Rachelle Thysell
Fabian Santiago
Matthew Hardwick
Lowell Olson
Suzanne Rivoire

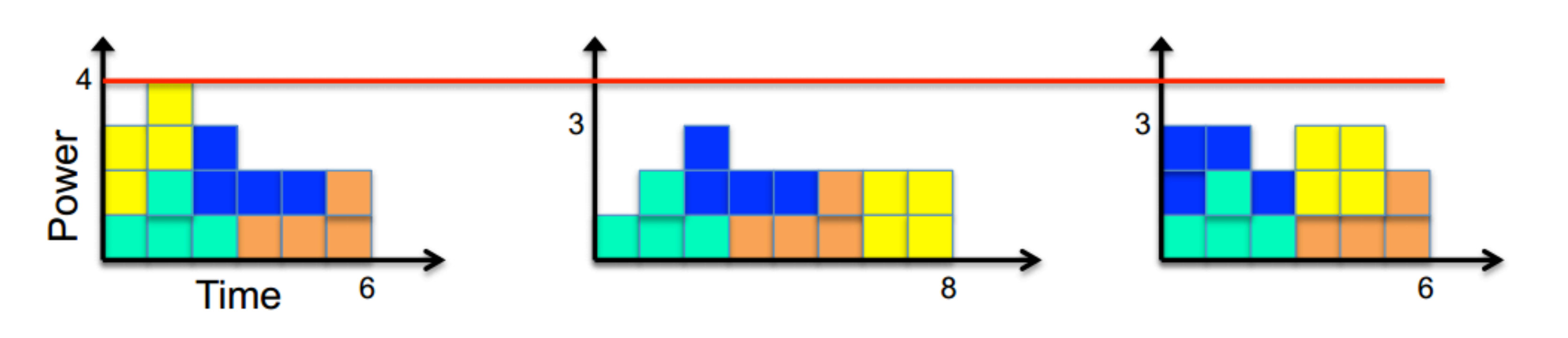
Chung-Hsing Hsu
Stephen W. Poole

SONOMA STATE UNIVERSITY

OAK RIDGE National Laboratory

Motivation

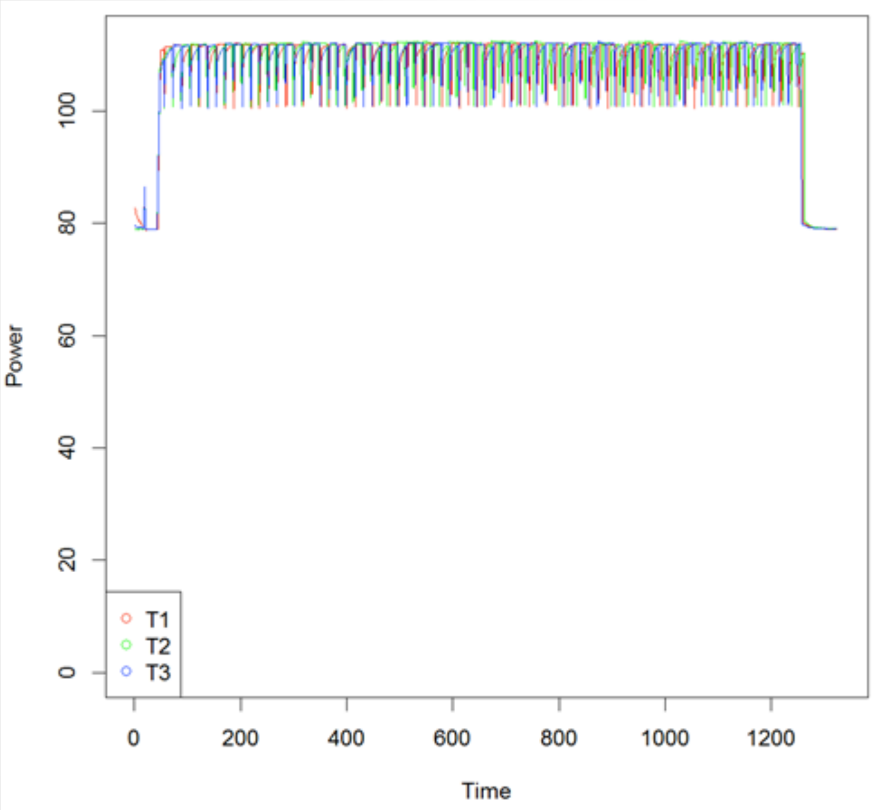
- Job scheduling as a Tetris game



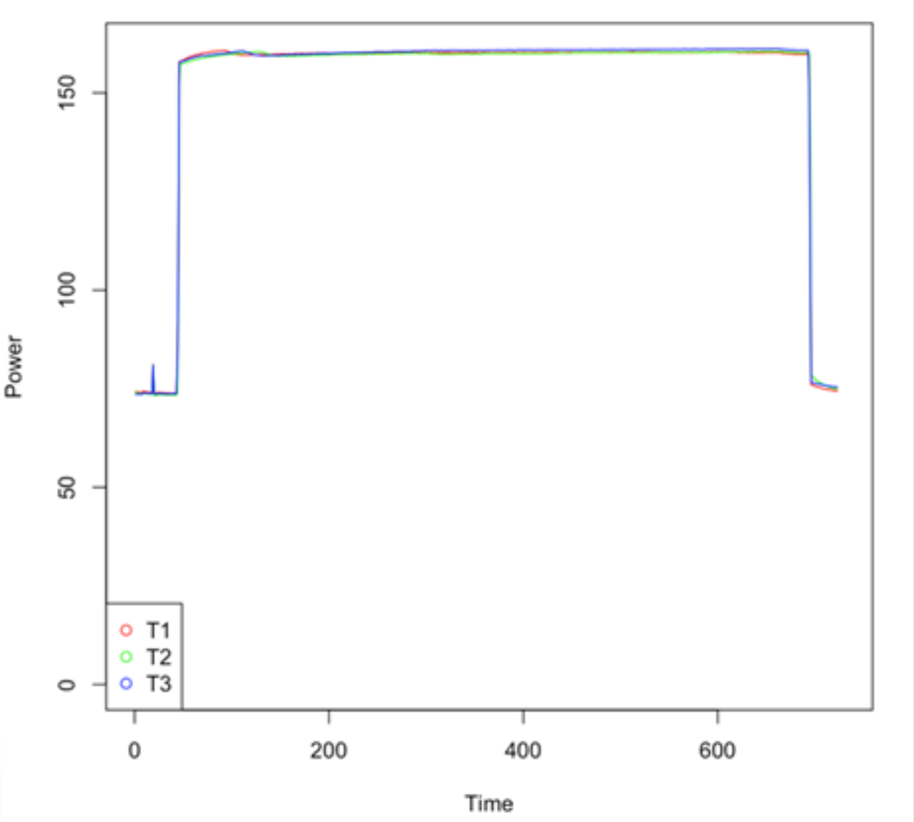
- Driven by power usage patterns.
  Can we:
  - Associate a pattern with each application?
  - Enhance scheduler with pattern information?

Motivation

- Qualitative patterns in applications' traces



FFT



CUBLAS

Talk Outline

- Research questions
- What is a power signature?
- Methodology:
    - Signature validation
    - Experimental setup
- Results
- Current and future work

Research Questions

- Can we summarize HPC workloads' power behavior into distinctive signatures?

- Is such a signature consistent across
  - runs?
  - input data?
  - hardware configurations?
  - hardware platforms?

- How well (quantitatively) does a signature distinguish a workload?

What is a power signature?

A. The trace itself: vector of power measurements.

B. Statistical summary of the trace

Time-series-based Signature

How do we quantify the difference between two traces?

1. Mean Squared Difference (MSD)
   - Match power observations pairwise, and take MSD
   - Traces must be same length
2. Dynamic Time Warping (DTW)
   - Identifies similarities of two time series
   - Accounts for offsets and differences in periodic frequency

Feature-based Signature

What features are useful?

- Basic statistics:
  - 2-vector: < Maximum, Median >
  - (Divide each by trace's minimum power)
  - Call this **MaxMed**
- More involved statistics that have been found useful in time-series clustering:
  - Standard Deviation + 11 other features
  - Augmented with **MaxMed**, call this **stat14**.

Signature Validation

- Clustering: “optimally” partition a set of traces

- Classification: automatically identify the label (e.g. workload) of a trace

Signature Validation: Clustering

- Input:
  - Data points (traces)
  - Notion of distance (signature)
- Output: Partition

Algorithms:

- kmeans: centroid-based clustering
- dbscan: density-based clustering
- hclust: hierarchical clustering
  - dendrograms

Signature Validation: Clustering

Our signature is good if the partition is good.
How do we know a partition is good?

1. Look at the partition qualitatively:
   Are workloads grouped together?

2. Quantitatively compare partition to some "ideal" reference.
   - Example ideal reference: grouped by workload

Signature Validation: Classification

Algorithm: Random forest

Leave-one-out accuracy measures a signature’s utility

Bonus: Variable importance measures

Experimental Setup

255 power traces from 13 benchmarks.

- (Baseline)
- SystemBurn*:
  - FFT1D
  - FFT2D
  - TILT
  - DGEMM
  - GUPS
  - SCUBLAS
  - DGEMM+SCUBLAS
- Synthetic: Power Model Calibration**
- Sort
- Prime95
- Graph500
- Stream
- Linpack-CBLAS

** Rivoire et al, Hot Power, 2008

* Josh Lothian et al., ORNL Technical Report, 2013

Experimental Setup

	S1 (RR)	S2 (OC)	S3 (LC)	S4 (RF)
CPU	AMD Athlon 65 X2 4800+ @ 2.5 GHz	Intel Core i5-750 @ 2.67GHz	Intel Core i5-750 @ 2.67GHz	Intel Core i7-3770 @ 3.40GHz
RAM	4 GB	8 GB	8 GB	8 GB
GPU	GeForce 9800gt	GeForce GTX 285	GeForce GTX 650 Ti 1GB	GeForce GTX 670 2GB
Power	115–195 W	120–226 W	85–252 W	74–309 W

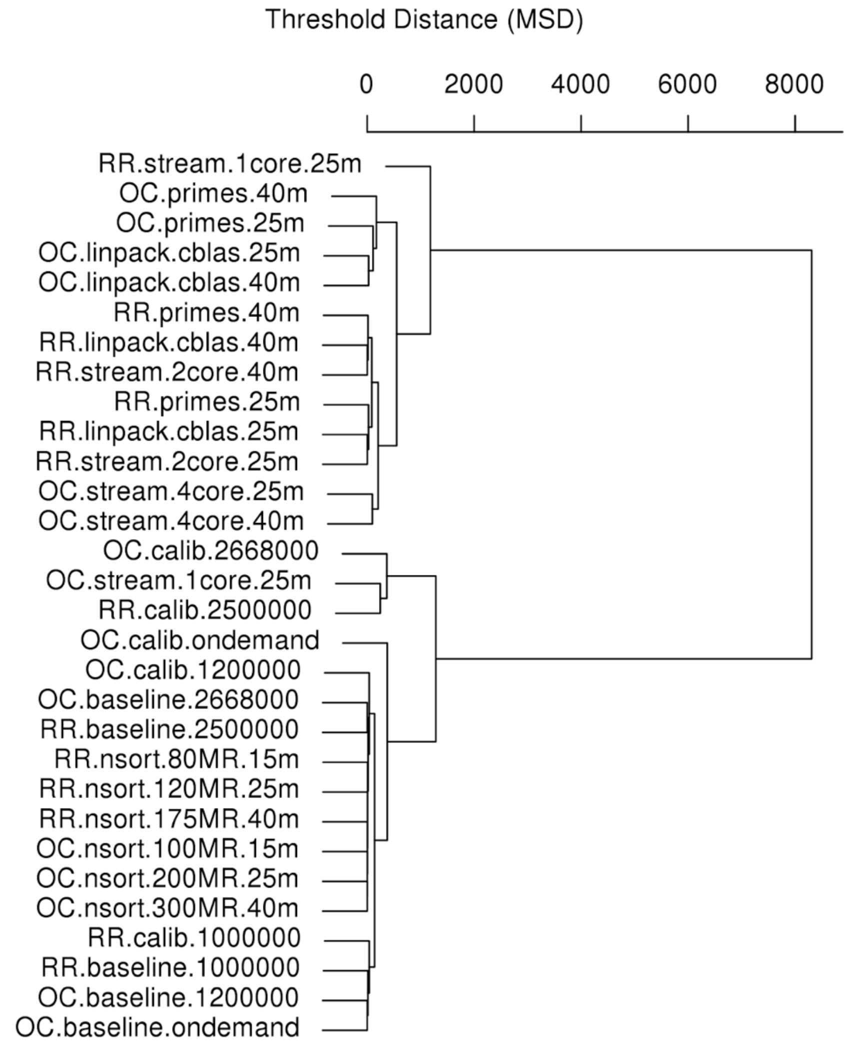
Watts Up? Pro power meter reports power consumption once per second.

Clustering Results

- OCRR data
  - n=30
  - 6 workloads (different input configurations)
- Algorithm: hclust
- Signature: raw trace
- Distance: MSD

2-clustering:

- Top: Stream, Prime95, Linpack-CBLAS (CPU-intensive)
- Bottom: Calib, Baseline, Sort

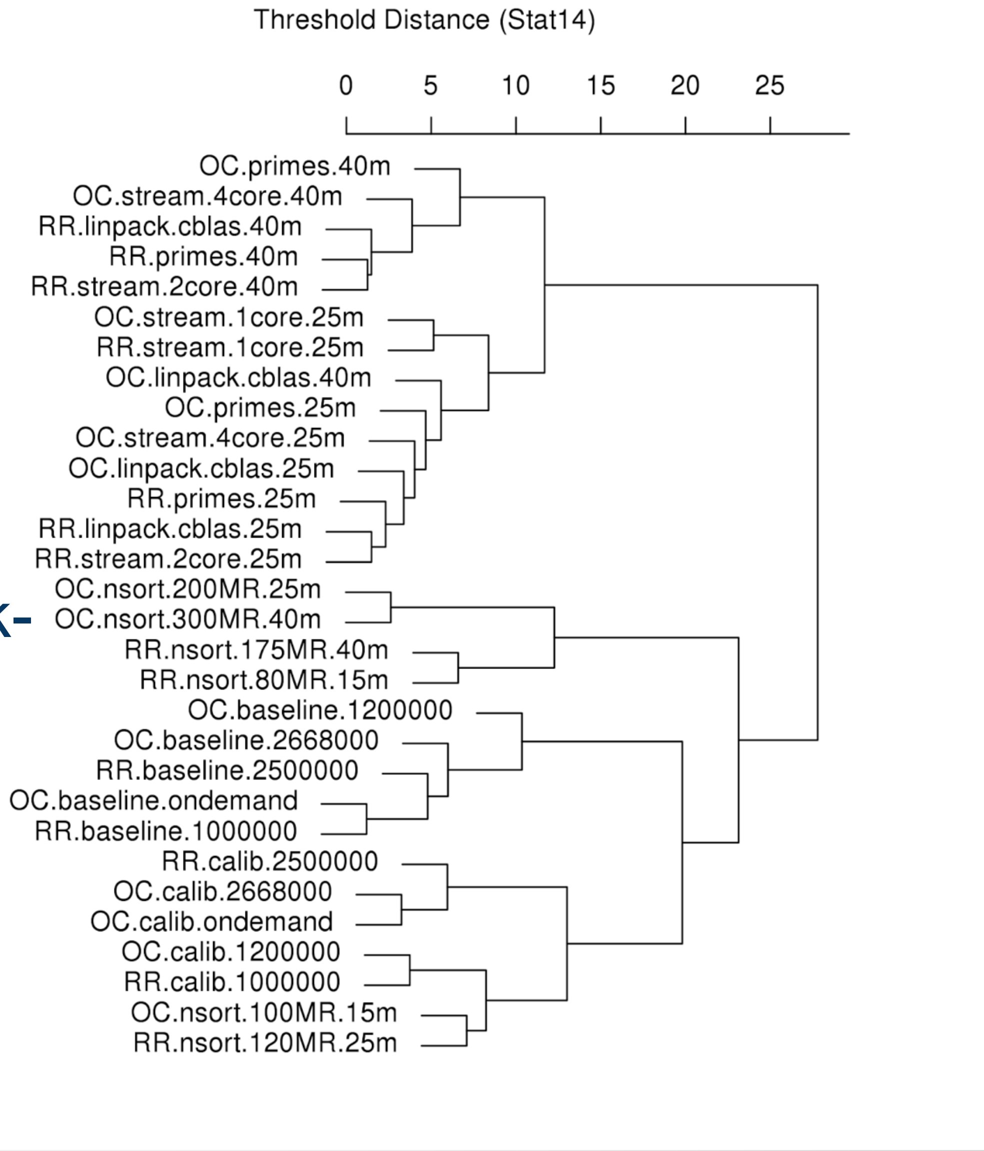


Clustering Results

- OCRR data
  - n=30
  - 6 workloads (different input configurations)
- Algorithm: hclust
- Signature: **stat14**
- Distance: Manhattan

4-clustering:

- Stream, Prime95, Linpack-CBLAS
- Sort
- Baseline
- Calib



Clustering Metric

Ideal clustering: by workload.

Info-theoretic measure of partition similarity:
Adjusted Normalized Mutual Information

(Derived from NMI)

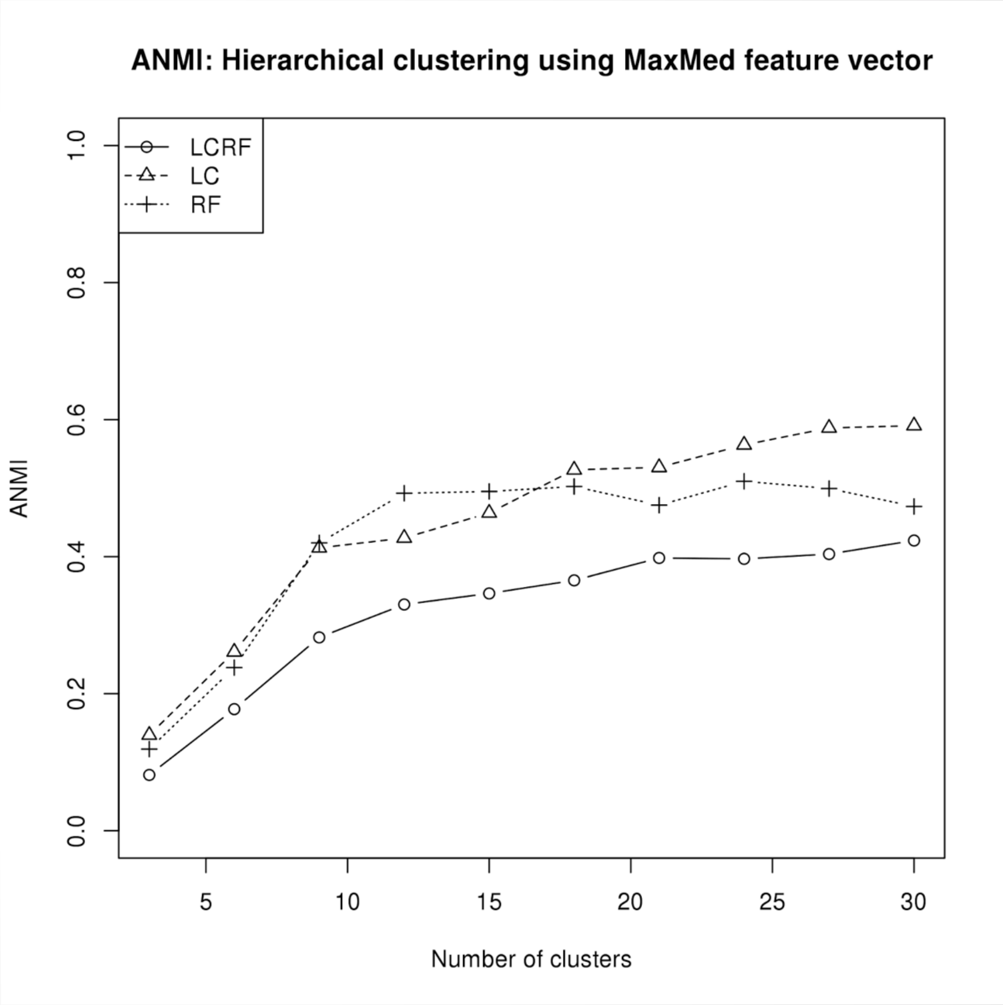
- NMI = (Mutual information) / (Joint entropy)
- NMI is between 0 (worst) and 1 (best)

- Expected ANMI of two random partitions is 0.

Clustering Results

- Data:
    - LCRF (n=225)
    - LC (n=111)
    - RF (n=114)
- Algorithm: hclust
- Signature: **MaxMed**

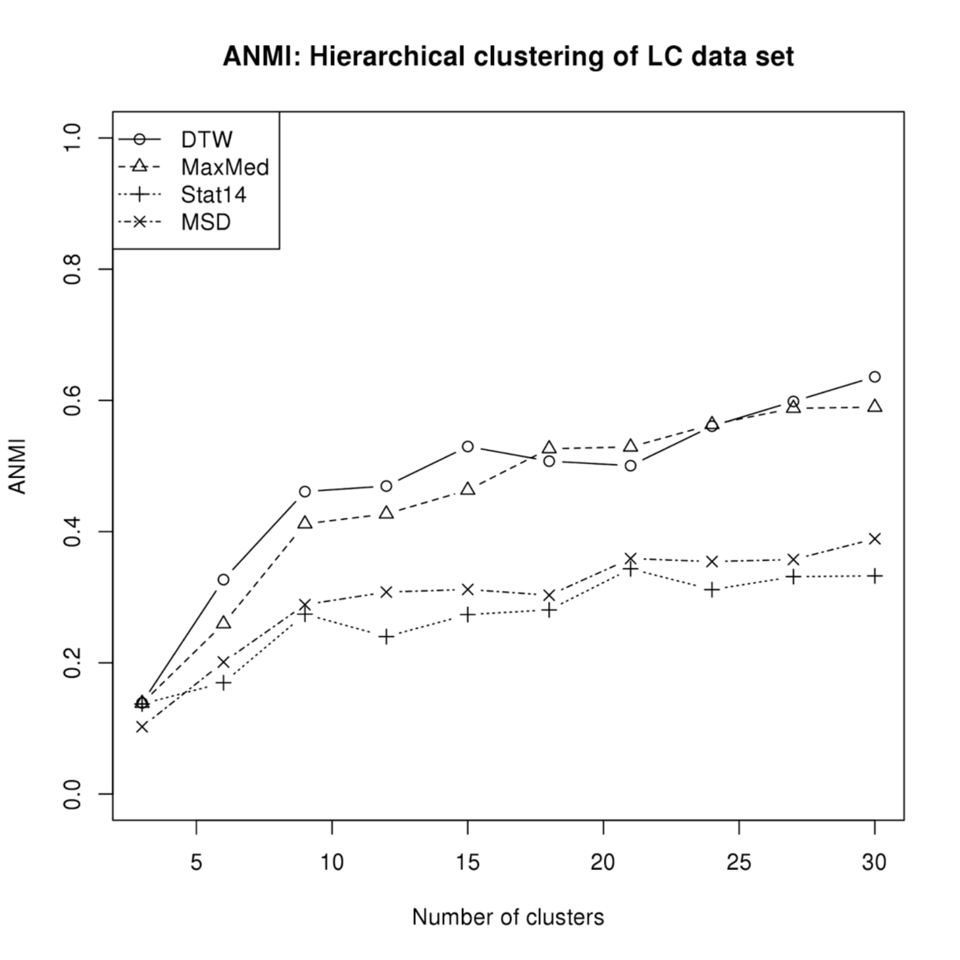
Signatures may be more consistent *within* hardware platform



Clustering Results

- Data: LC (n=111)
- Algorithm: hclust

MaxMed and **DTW** signature methods are more effective than **Stat14** and **MSD**



Classification Results

- Trained a random forest classifier on LCRF data (n=225)
- Using **MaxMed** or **Stat14** yields leave-one-out accuracy >80%

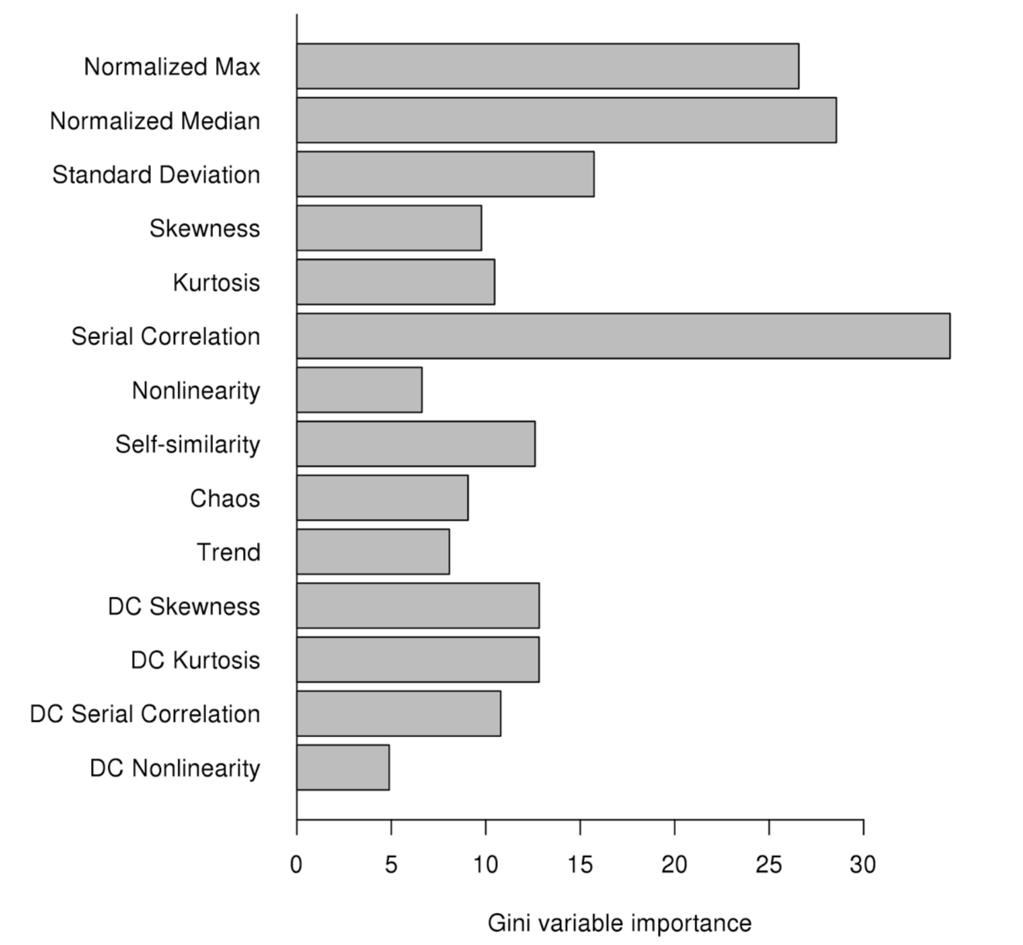
Classification Results

Gini variable importance suggests:

- **MaxMed** is a good subset of **Stat14**

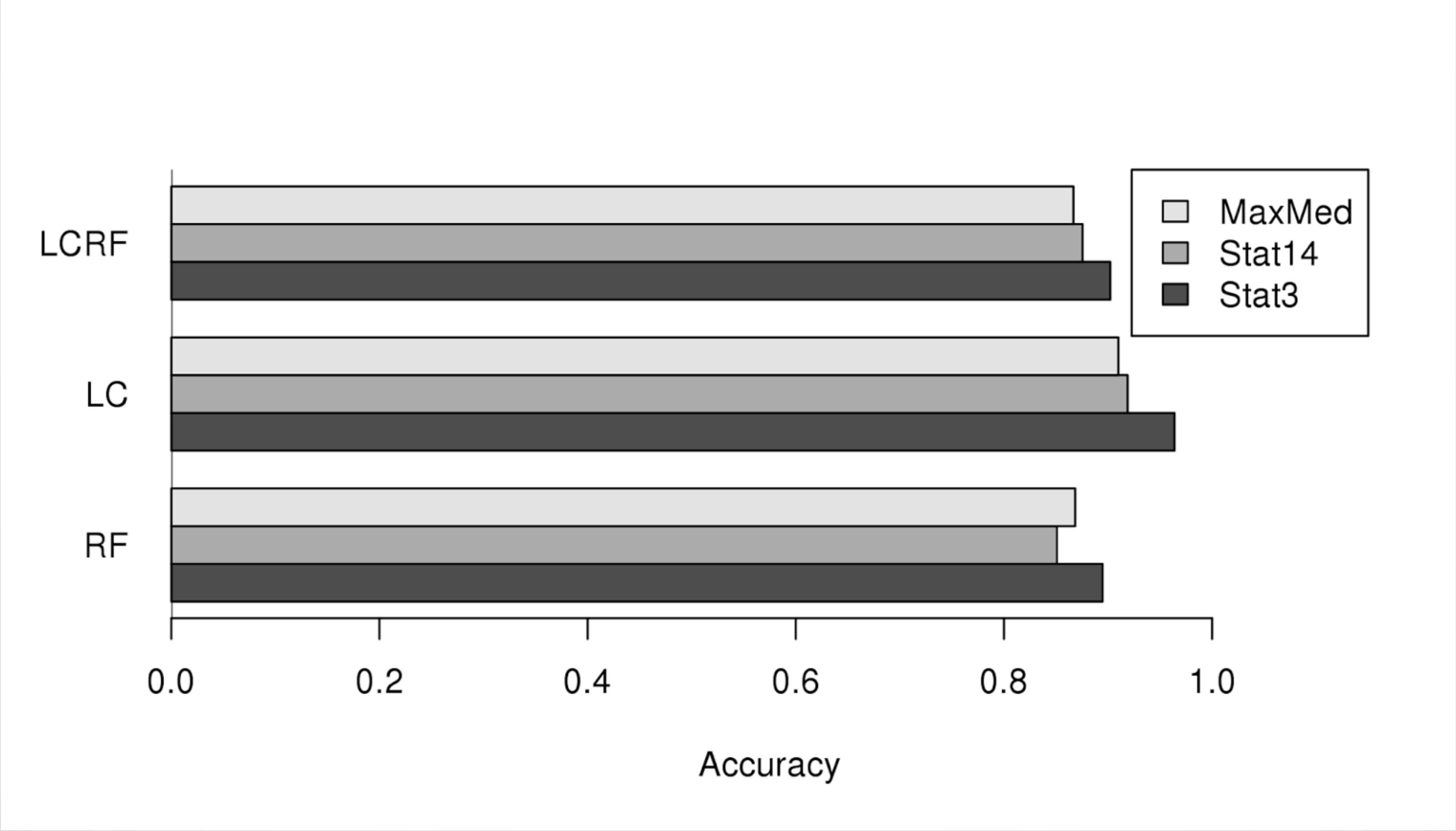
- Try **Stat3**:

< Normalized Maximum,
Normalized Median,
Serial Correlation
>



Classification Results

- **Stat3** classifier labels traces with >85% accuracy



Conclusions

- We evaluated different types of signatures:
  - Time-series-based
  - **Feature-based**
- Some workloads have unique signatures, some workloads are less easily distinguished from others.
- Signatures can distinguish workloads across hardware platforms, but are more effective given data from a single machine type.

Current and Future Work

- Expand to:
  - Heterogeneous workloads
  - MPI/distributed workloads
  - Finer-grained or coarser-grained samples

- Online workload recognition

- Workload-aware energy-efficient scheduling

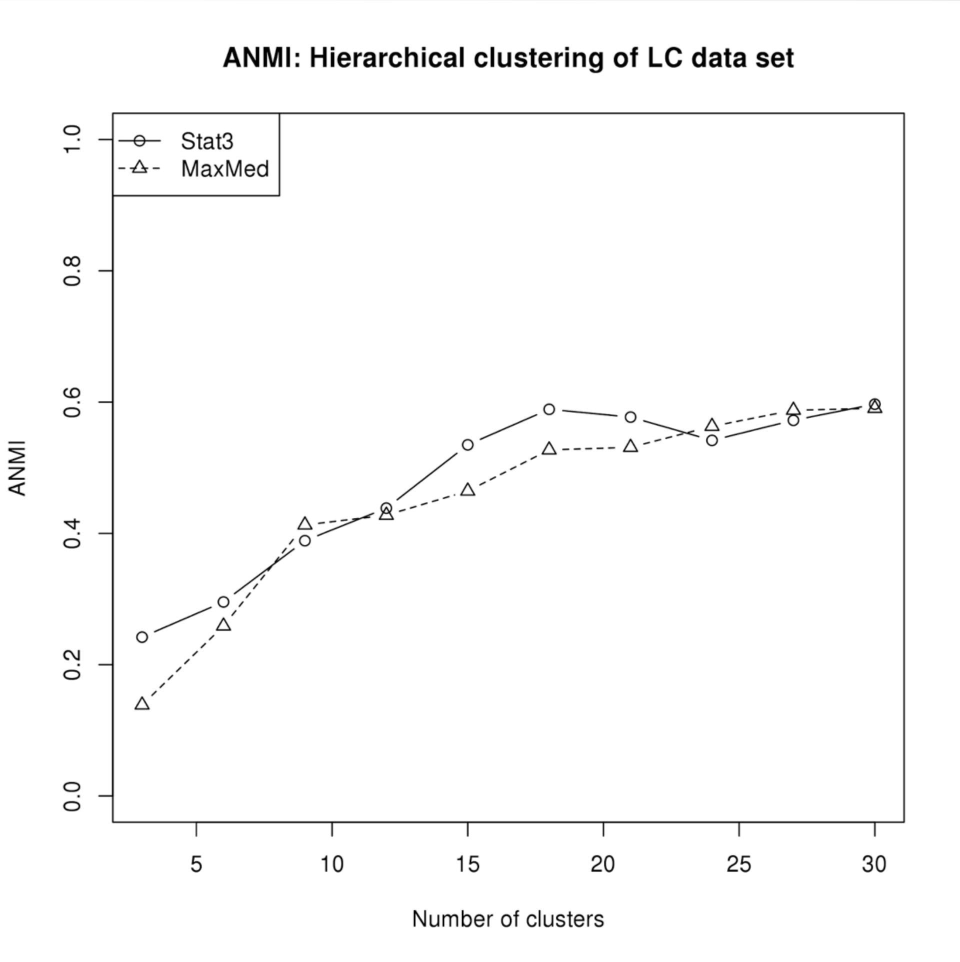
Acknowledgements

This work was supported by the United States Department of Defense (DoD) and used resources of the DoD-HPC Program at Oak Ridge National Laboratory.

Afterthought: Clustering Again

- Data: LC (n=111)
- Algorithm: hclust

Stat3 is *not* obviously better than **MaxMed** for clustering



Backup: More Clustering Results

- Data: LCRF (n=225)
- Algorithm: hclust

The result holds for multiple platforms:

MaxMed and **DTW** signature methods are more effective than **Stat14** and **MSD**