

# Aligning traces for performance evaluation



Todd Mytkowicz (University of Colorado)

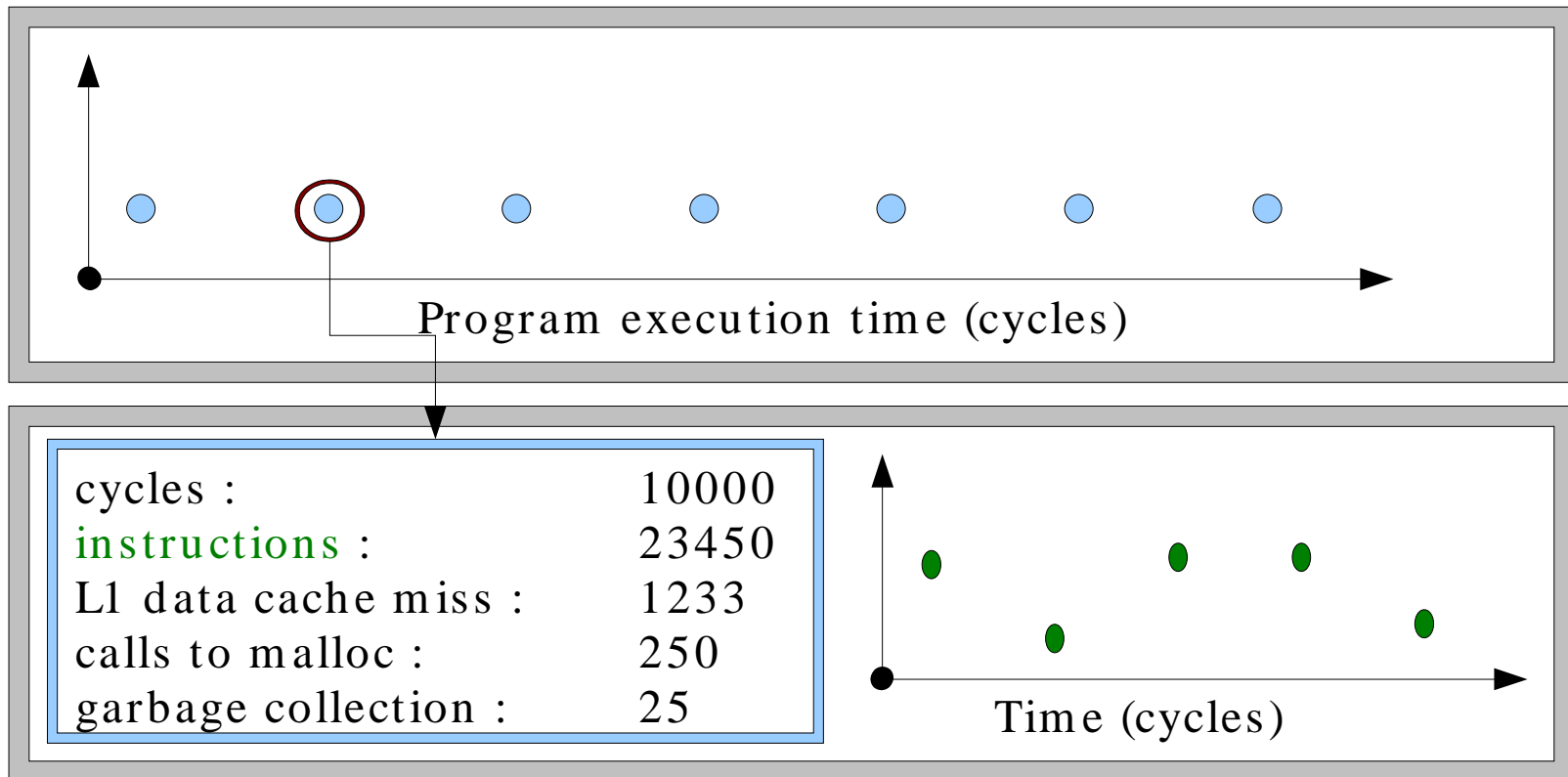
Amer Diwan (University of Colorado)

Peter F. Sweeney (IBM Research)

Matthias Hauswirth (University of Lugano)

# What is a trace

- Time series of trace records
  - each trace record can contain multiple metrics
  - time-varying rather than aggregate performance

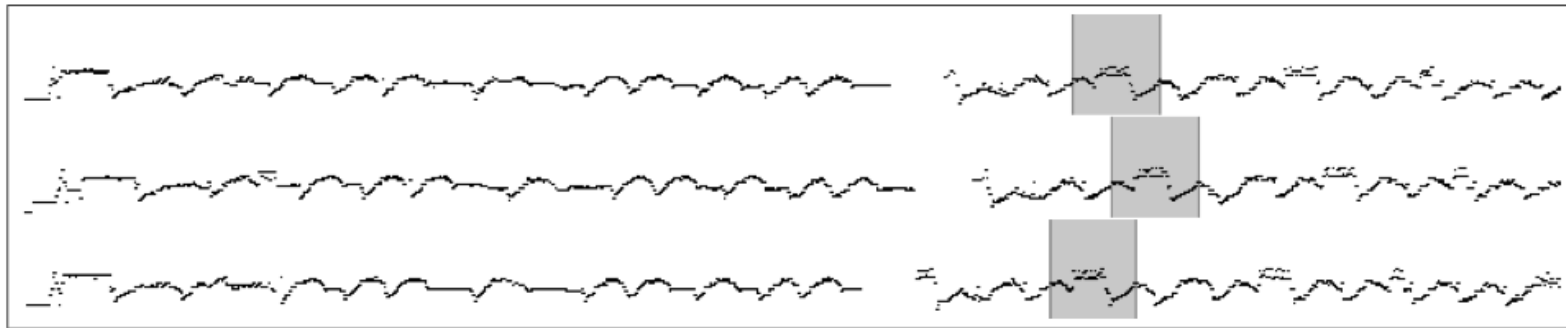


# Multiple runs for multiple metrics

- We would like to capture more metrics than we can in a single run
  - hardware issues
    - register limitations
  - software issues
    - measurement perturbation
- We need multiple runs

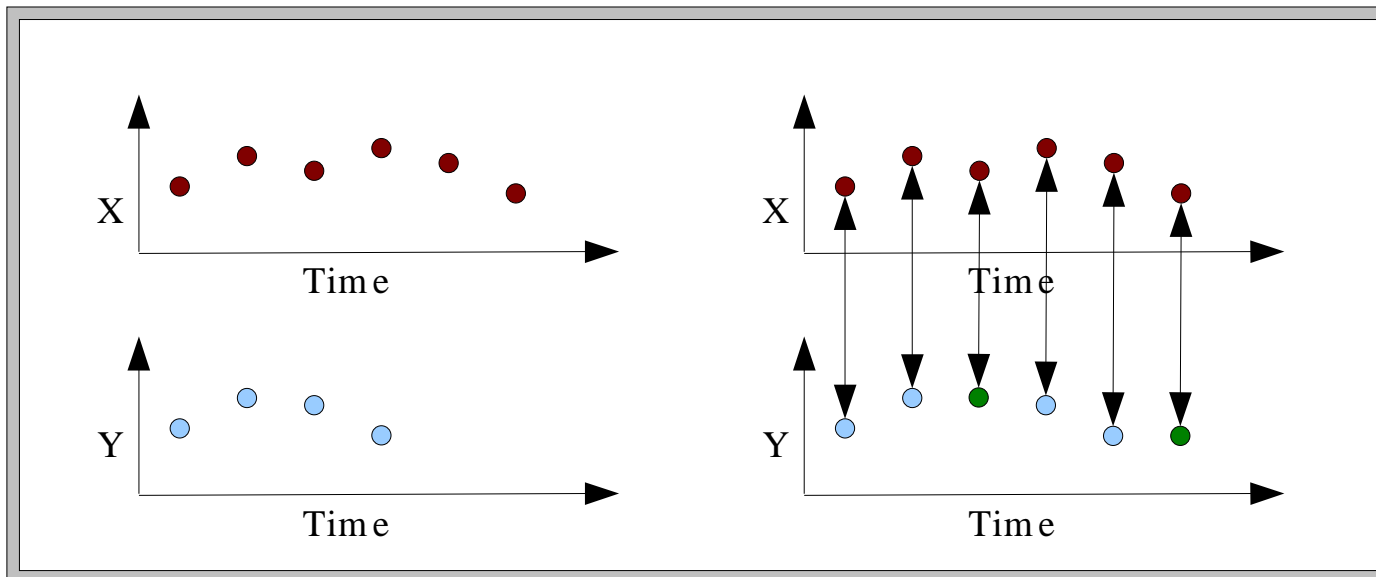
# Trace alignment problem

- How do we align trace records from multiple runs?
  - We do not have a one to one correspondence
    - System and measurement perturbation



# Full Uniform Scaling

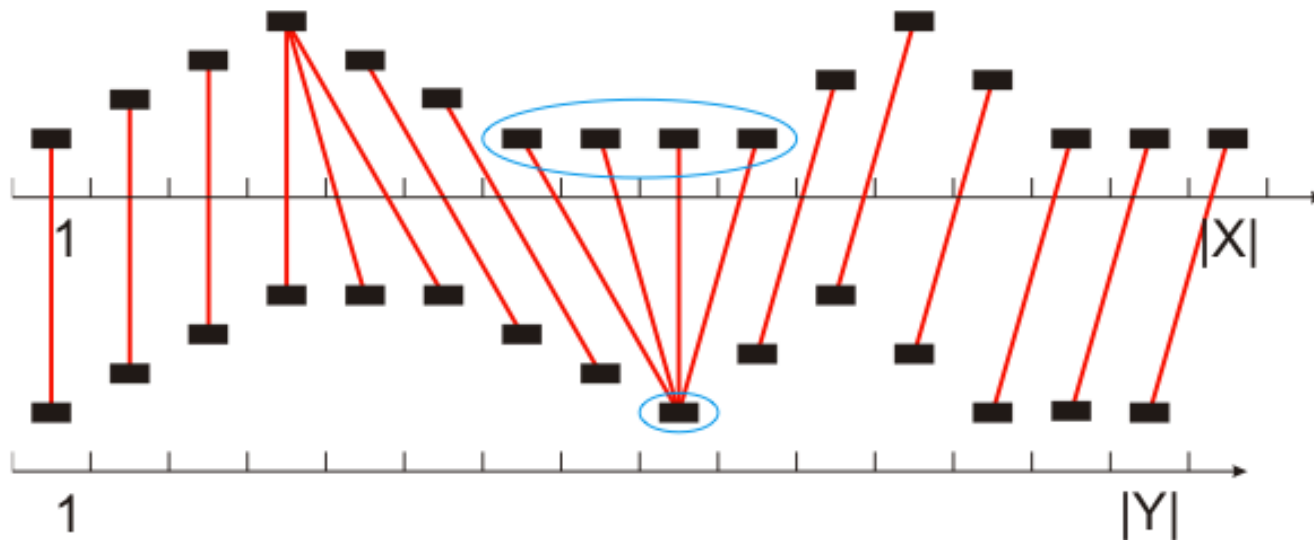
- Uniformly scale the shorter sequence to match the length of the longer one [E. Keogh]
  - Insert repeated points in shorter sequence
- Naive to fluctuations in time series



# Dynamic time warping

- Dynamic programming technique that aligns trace records
  - Resulting alignment minimizes DTWError

$$DTWError = \sum_{k=1} |x_i - y_i|$$



# When is an alignment good enough?

- ... when it allows a performance analyst to correlate metrics within a trace as effectively as across traces

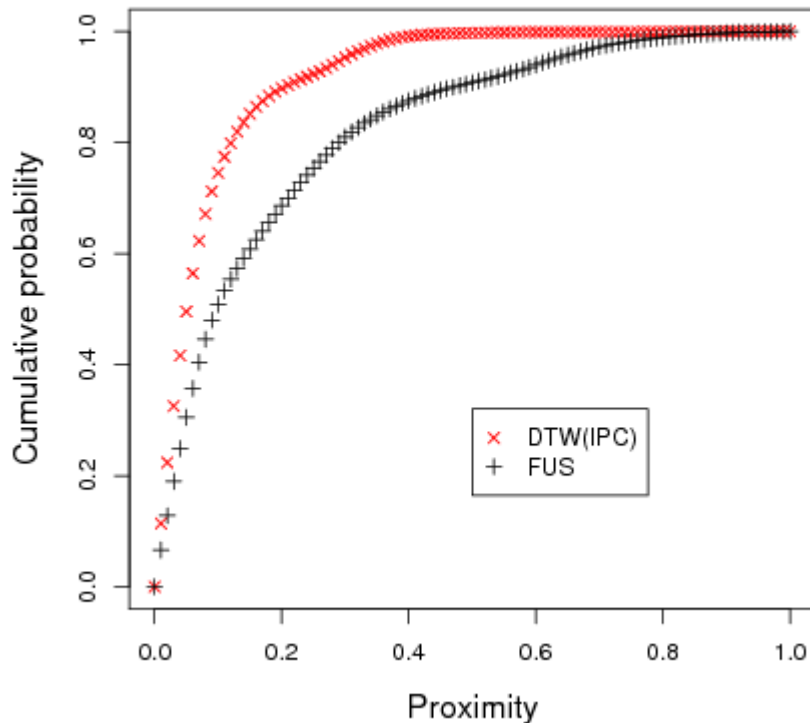
# Experimental methods

- POWER4
  - 8 hardware events
- JikesRVM
  - Extended to collect software metrics
- SpecJVM98
  - 5 benchmarks
  - 40 traces
  - 42 metrics per trace

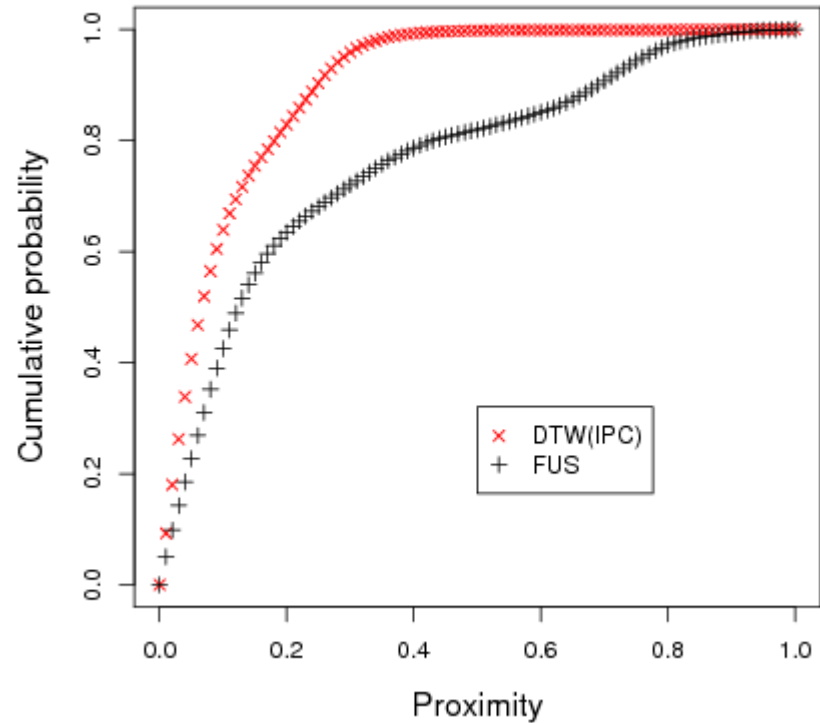
# Proximity: DTW and FUS

- DTW(IPC) lets us correlate across traces more accurately than FUS

Jess

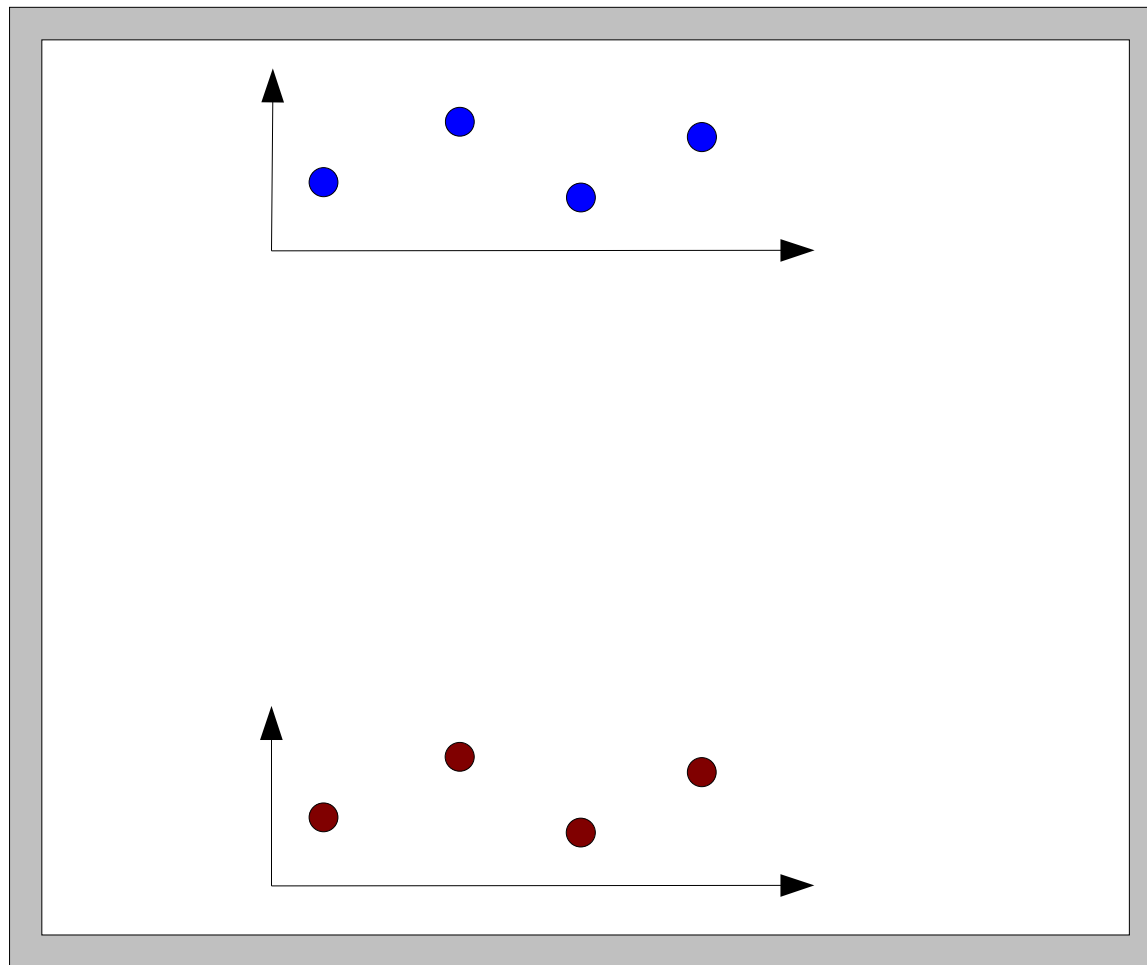


Raytrace



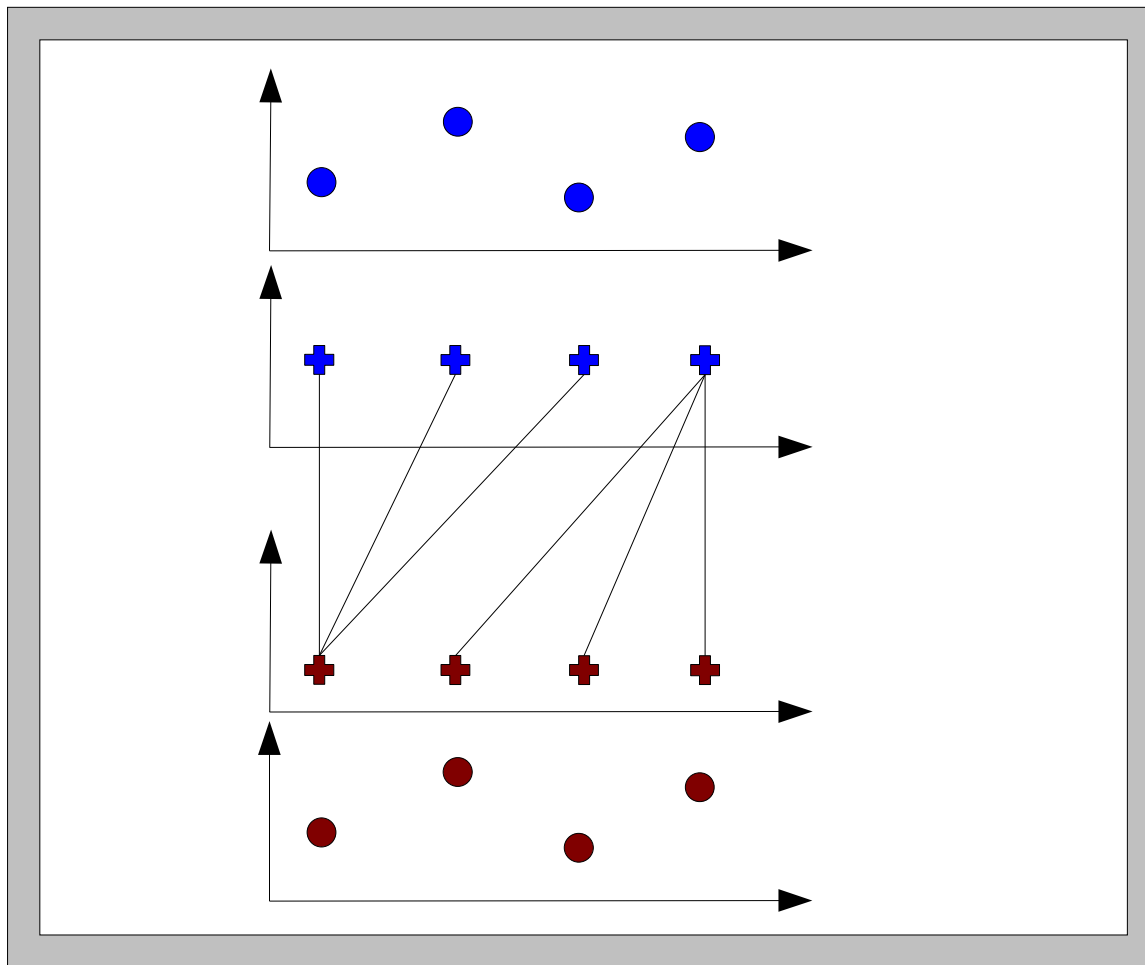
# Properties of a good alignment metric

- Alignment metric must reflect transitions of correlation metrics



# Properties of a good alignment metric

- Alignment metric must reflect transitions of correlation metrics



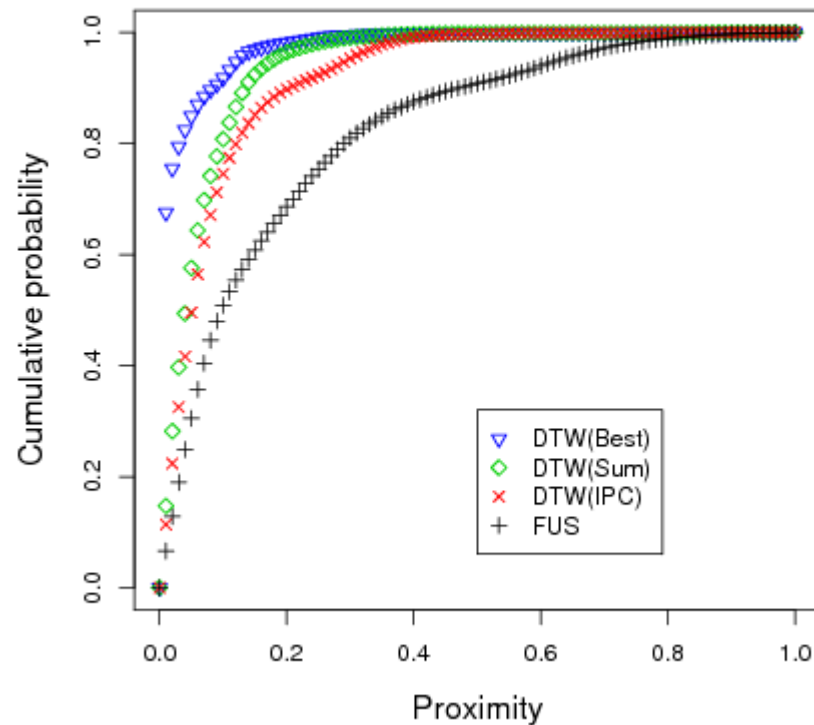
# Picking a good alignment metric

- Given metrics of interest, B and C:
  - IPC
    - Manually pick a reflective metric
  - Sum
    - Pick the metric that is most highly correlated with B and C, respectively
  - Best
    - Pick metric that gives the smallest difference between correlation within a trace and correlation across traces

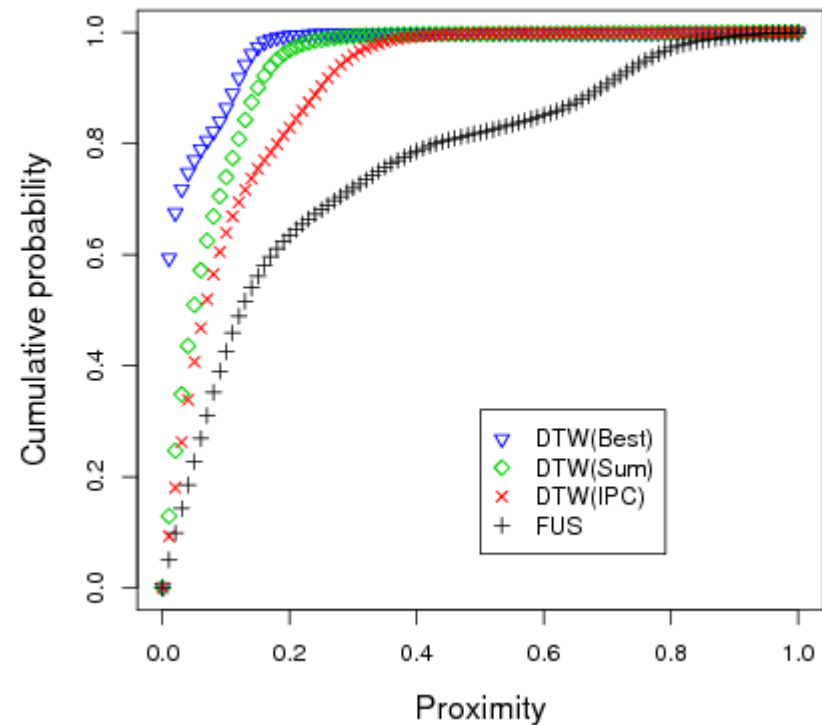
# Smart Alignment Metric

- DTW(Sum) allows us to reason more accurately than DTW(IPC)

Jess



Raytrace

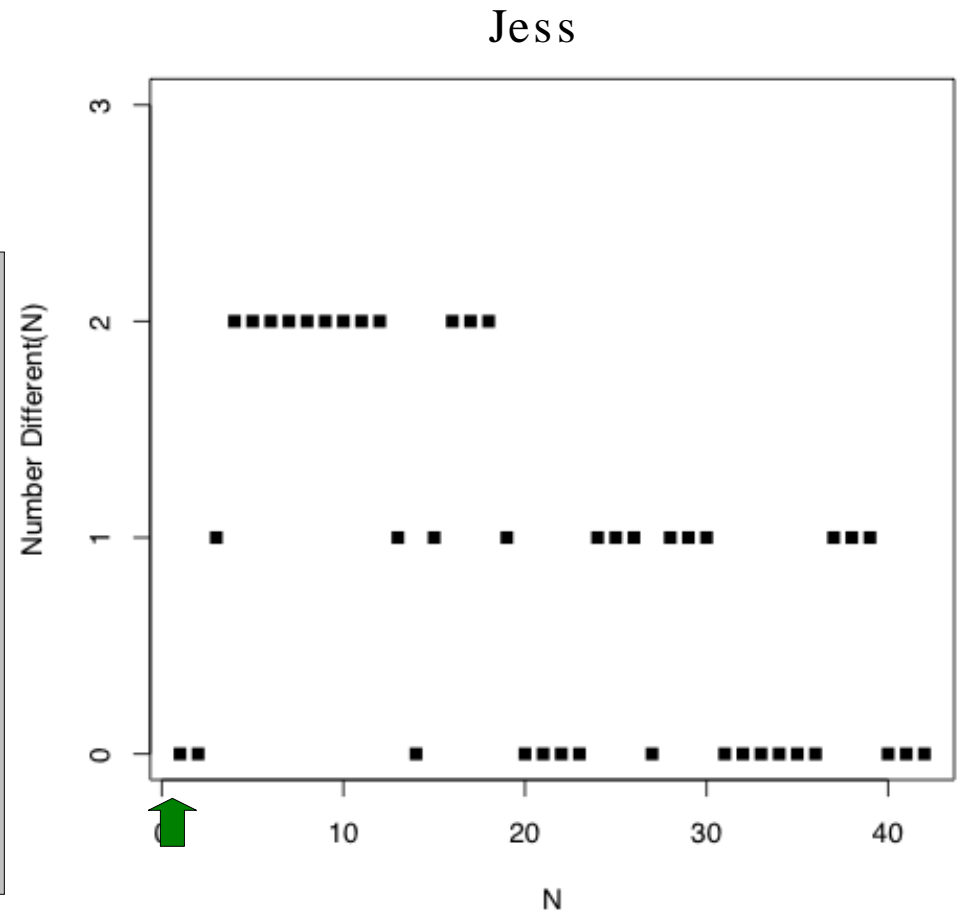


# Ranking

- DTW(Sum) orders metrics of interest within trace almost as well as out of trace

Example Data for Jess

N	within trace rank	across trace rank
1	A	A

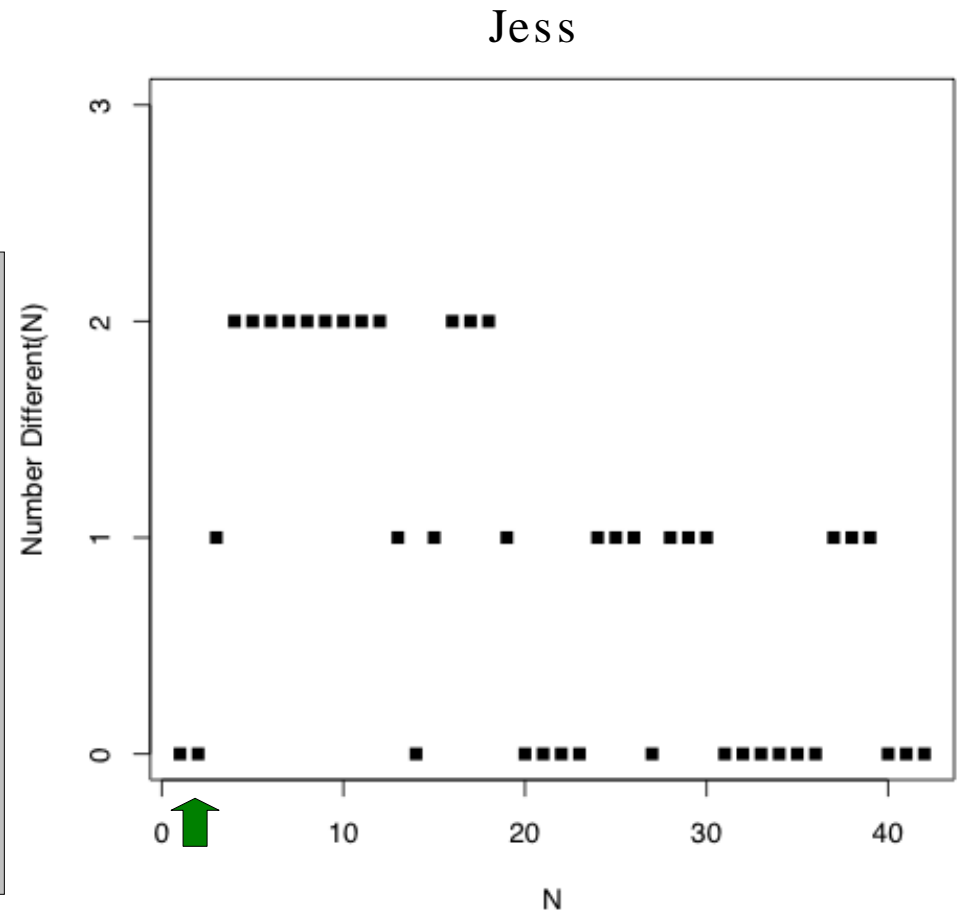


# Ranking

- DTW(Sum) orders metrics of interest within trace almost as well as out of trace

Example Data for Jess

N	within trace rank	across trace rank
1	A	A
2	F	F

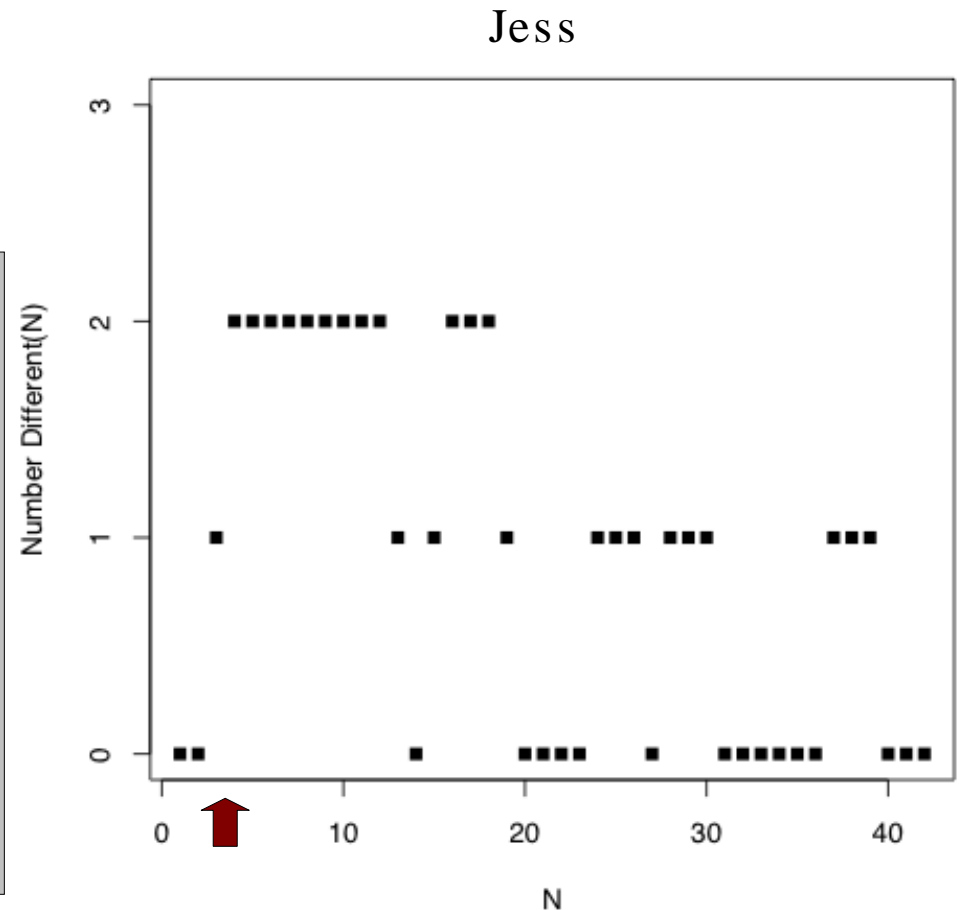


# Ranking

- DTW(Sum) orders metrics of interest within trace almost as well as out of trace

Example Data for Jess

N	within trace rank	across trace rank
1	A	A
2	F	F
3	B	C

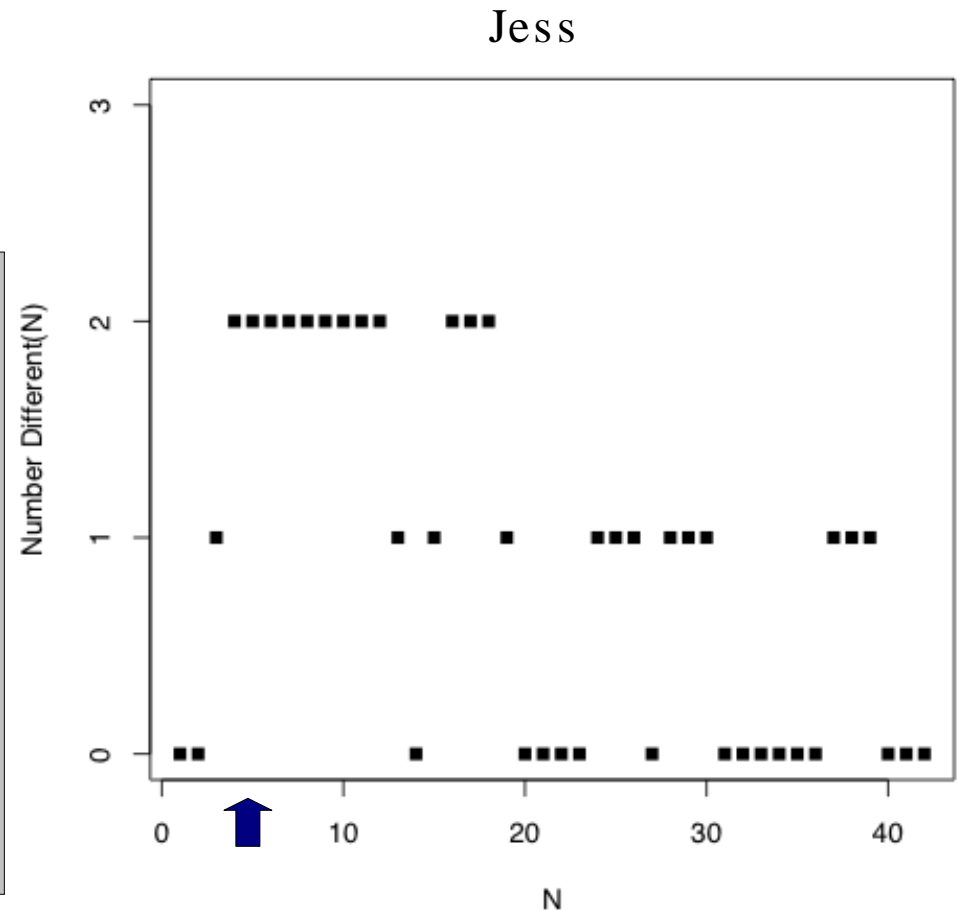


# Ranking

- DTW(Sum) orders metrics of interest within trace almost as well as out of trace

Example Data for Jess

N	within trace rank	across trace rank
1	A	A
2	F	F
3	B	C
4	D	U

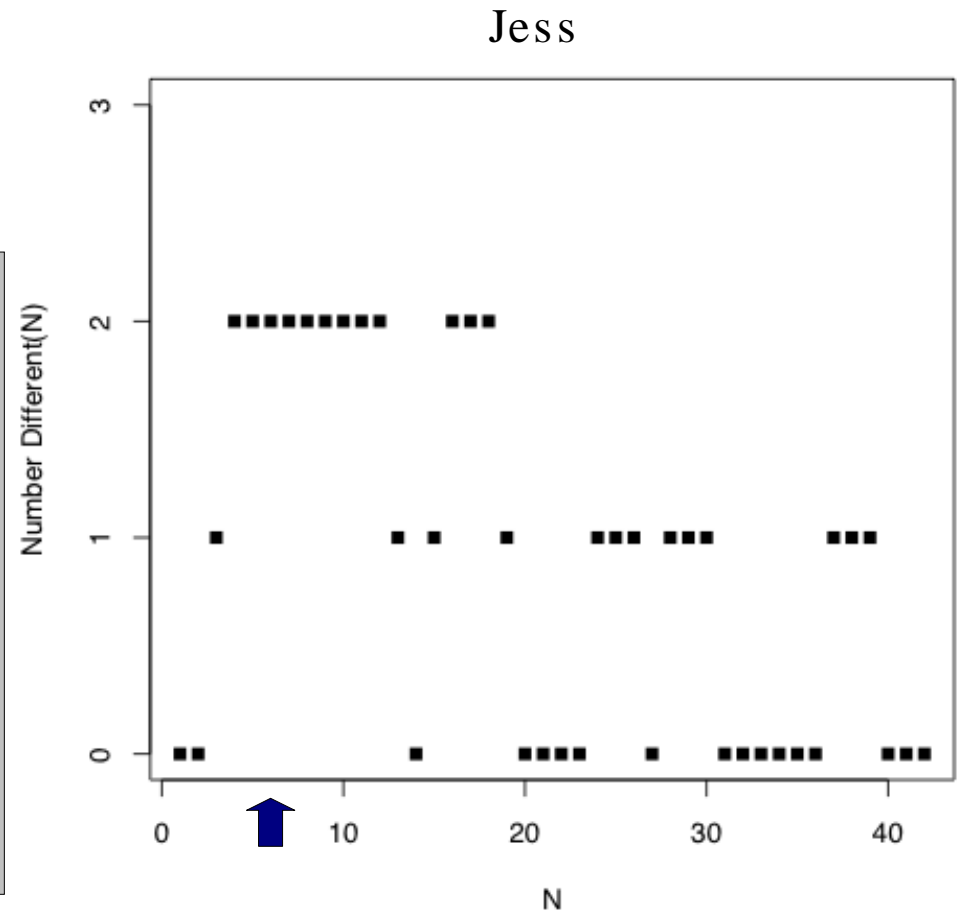


# Ranking

- DTW(Sum) orders metrics of interest within trace almost as well as out of trace

Example Data for Jess

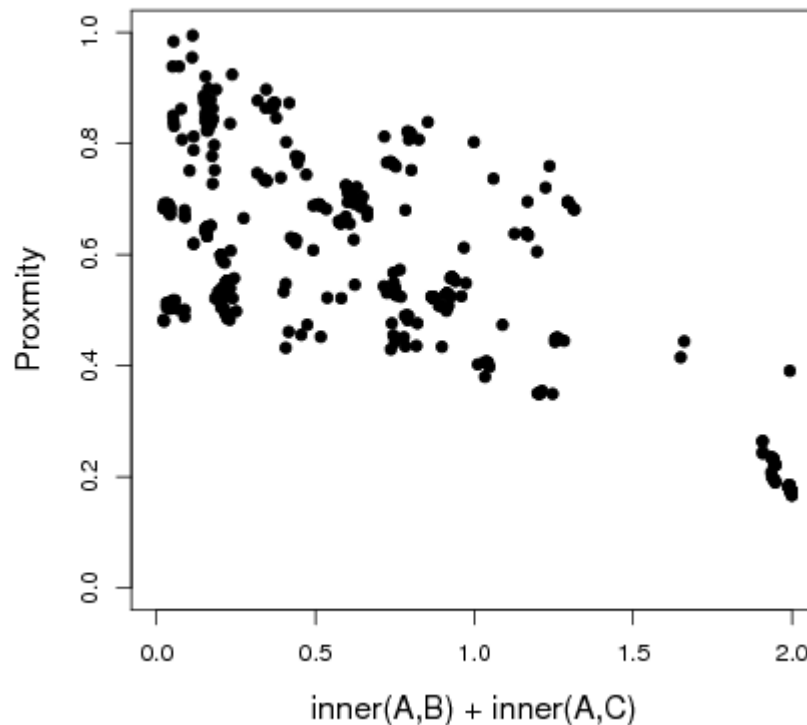
N	within trace rank	across trace rank
1	A	A
2	F	F
3	B	C
4	D	U
5	R	D



# When to trust alignment

- High inner correlation with alignment metric provides insight into aligning errors

Top 10 inner-correlated metric pairs for Jess



# Related Work

- Multiplexing hardware events
  - R. Azimi, M. Stumm, and R. W. Wisniewski. *Online performance analysis by statistical sampling of microprocessor performance counter*
  - Wiplove Mathur and Jeanine Cook. *Toward accurate performance evaluation using hardware counters*
- Dynamic Time warping
  - Matthias Hauswirth, Amer Diwan, Peter F. Sweeney, and Michael C. Mozer. *Automating vertical profiling*
- Full Uniform Scaling
  - E. Keogh. *Efficiently finding arbitrarily scaled patterns in massive time series databases*

# Conclusions

- We provide a methodology to determine when an alignment can be trusted
- Using this, we show one can correlate across traces as effectively as within a trace