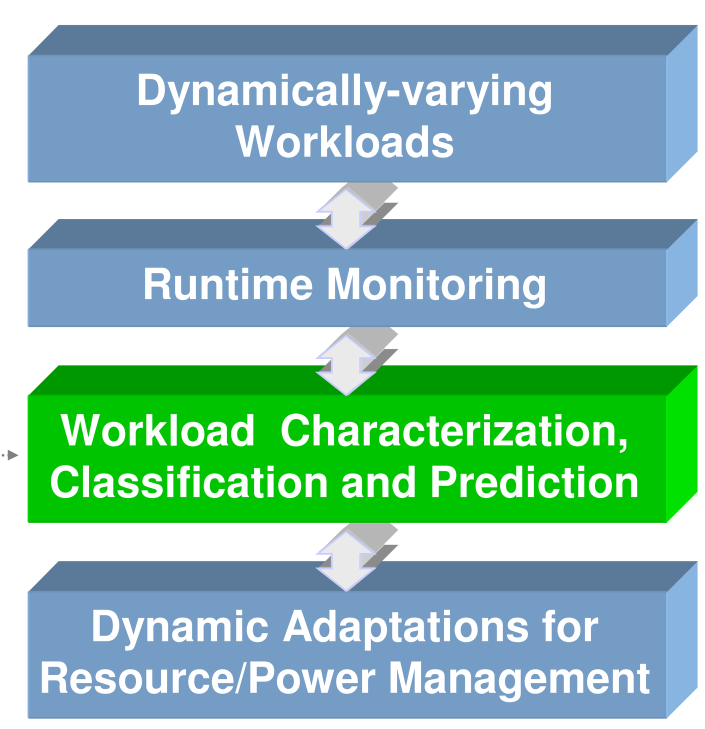
IBM Research
T.J. Watson Research Center, NY

Program Behavior Prediction Using a Statistical Metric Model

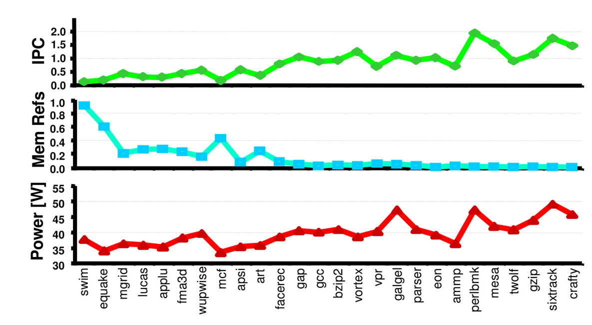
Ruhi Sarikaya, Canturk Isci and Alper Buyuktosunoglu

Workload Prediction and Adaptive Management

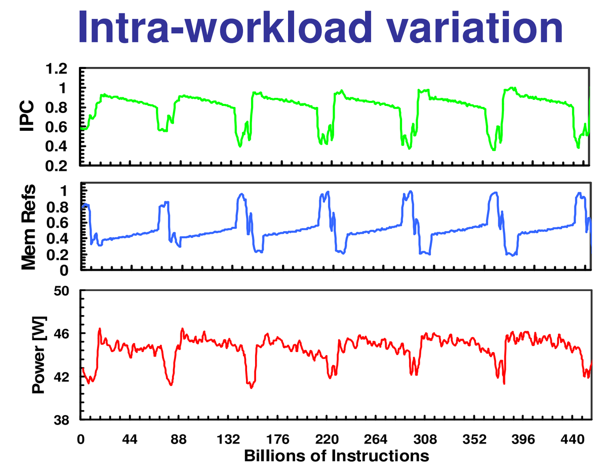


Our Work:
Understand and predict dynamic workload behavior

Inter-workload variation

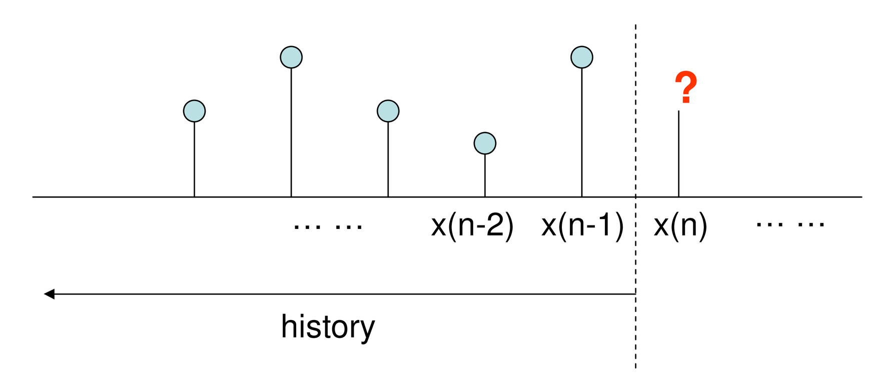


Intra-workload variation

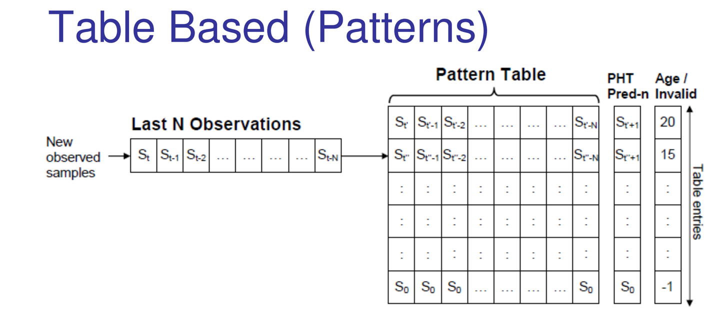


Predicting Workload Behavior

Define workload features
Determine future characteristics



Existing Techniques:
Last Value (Strawman/Reactive)
History Based (Statistical)
Table Based (Patterns)



Our Approach: Statistical Metric Modeling

- Inspired from natural language modeling
- Workload features ⇔ words in language
- Workload patterns ⇔ grammar structure
- Model workload structure at runtime
- Build metric probability distributions
- Predict future characteristics

Statistical Metric Model (SMM)

SMM Overview:

Probability distribution P(s) over sequences s:

$$s = (s_1, s_2, \ldots, s_l)$$

Ex: $P(\text{"How are you doing"}) \approx 0.001$

Difficult to compute $P(s) = P(s_1, s_2, \ldots, s_l)$

Decompose the probability instead:

$$P(s) = P(s_1) \times P(s_2 \mid s_1) \times P(s_3 \mid s_2, s_1) \times \ldots \times P(s_l \mid s_{l-1}, \ldots, s_1)$$

Ex: $P(\text{"How are you doing"}) = P(\text{"how"}) \times P(\text{"are|how"}) \times P(\text{"you|how are"}) \times P(\text{"doing|how are you"})$

Use ***n-gram*** Approximation:

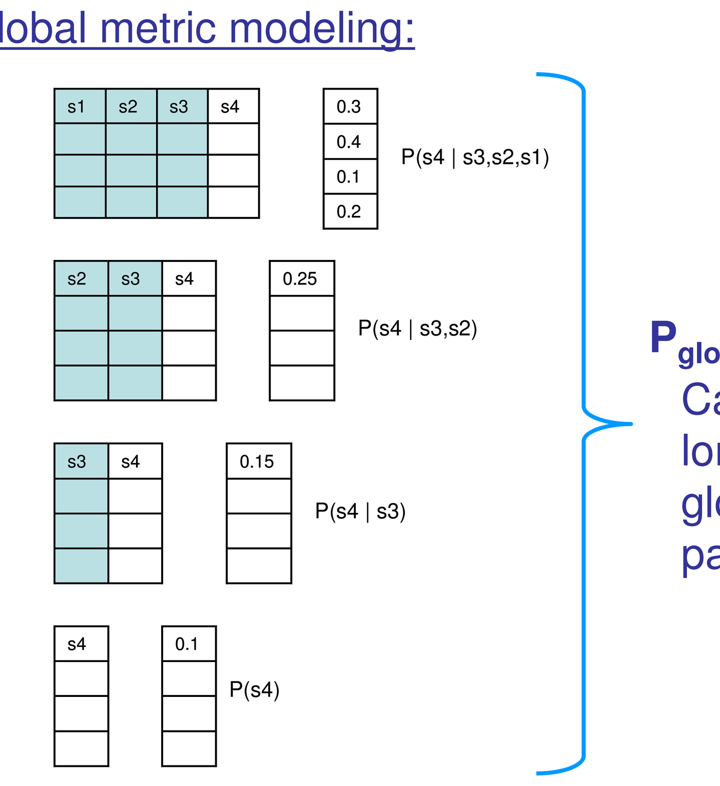
*Assume each word depends only on the previous **n** words*

$$P(s) \approx \prod_{i=1}^{l} P(s_i \mid s_{i-1}, \ldots s_{i-n+1})$$

Apply *model smoothing* to conditional distributions to compensate for data sparsity

SMM for Workload Behavior Prediction:

Global metric modeling:



P_{global} Capture long-term global patterns

Temporal metric modeling:

s1	s2	s3	…	…	sn
0.1	0.2	0.1	…	…	0.1

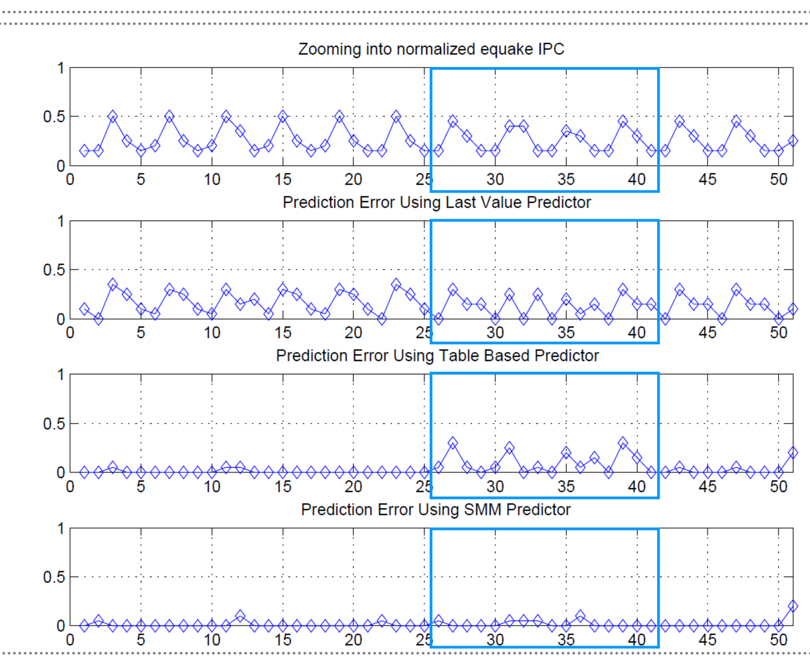
$P_{temporal}$ Capture recent behavior

Overall model:

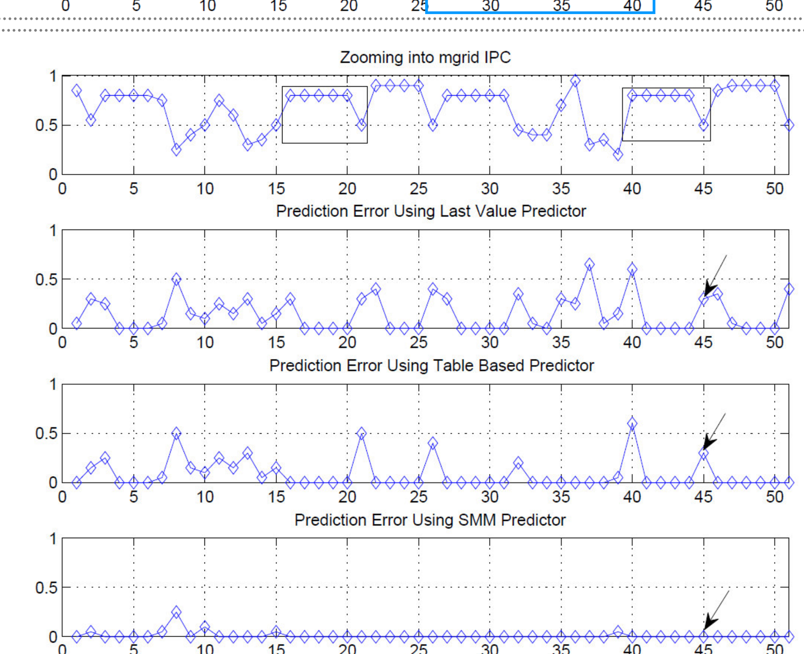
$$P_{final} = \beta_1 \cdot P_{global} + \beta_2 \cdot P_{temporal}$$

SMM Application Examples:

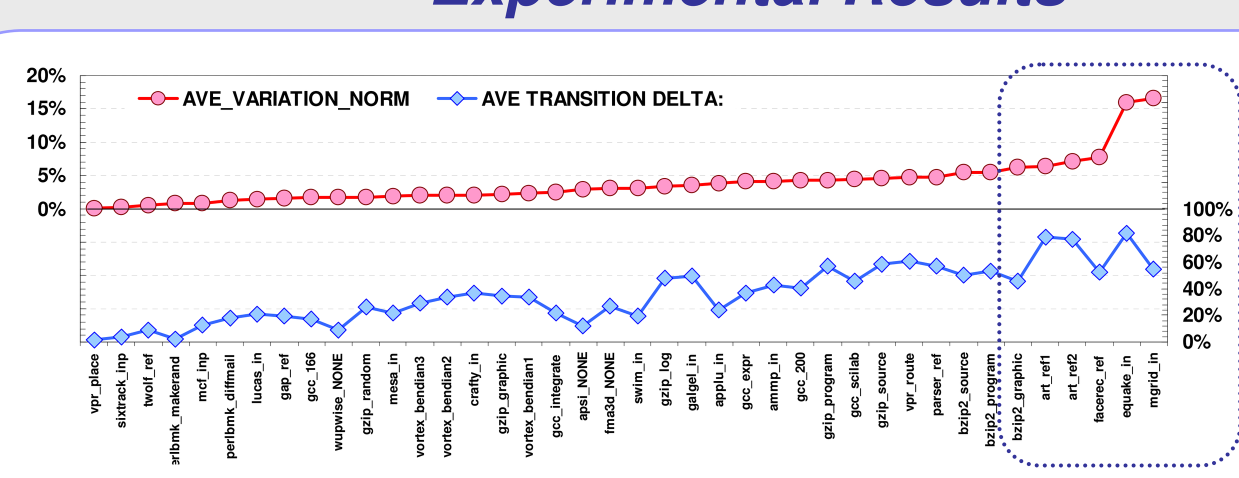
Resilient to pattern fluctuations



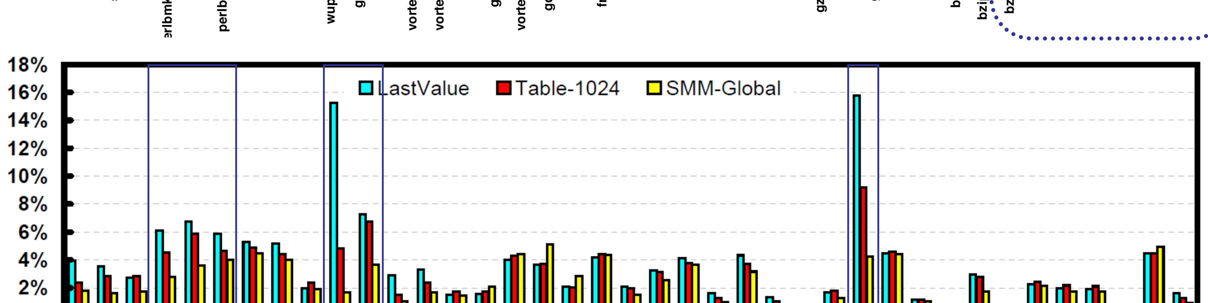
Tracks varying pattern lengths

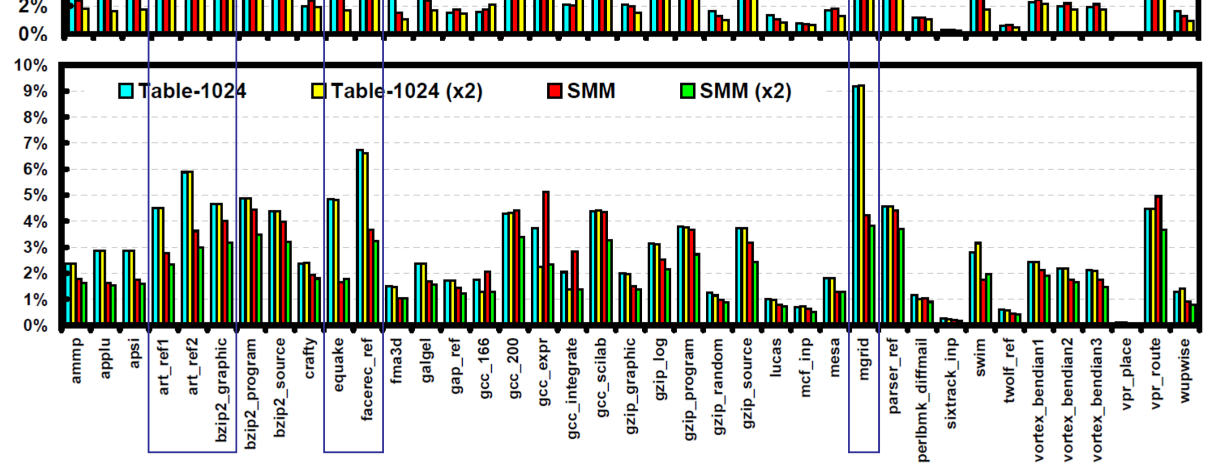


Experimental Results



More Important for SMM:
More variability
Harder to predict





Prediction Accuracy:

Across all workloads:
20% improvement

Variable workloads:
40% improvement

Repetitive runs (x2):
Additional 15% improvement across all workloads

Summary

Primary Contributions

- New workload behavior prediction approach Inspired by language modeling
- Evaluation with a comprehensive set of benchmarks and datasets
- Significant improvement in accuracy over prior approaches

4 Main SMM Strengths

- Models long-term global patterns in application behavior
- Can track and predict variable-length patterns
- Resilient to small fluctuations in workload behavior
- Adapts and improves over time, as it learns more it predicts better

Evaluation Highlights

- Improve prediction accuracy by 40% for variable workloads
- Average improvement of 20% across all benchmarks
- Additional 15% improvement with recurring workloads