The Relevance of Long-Range Dependence in Disk Traffic and Implications for Trace Synthesis

Bo Hong, Storage Systems Research Center
Tara M. Madhyastha, Dept. of Computer Engineering

University of California, Santa Cruz

UC Santa Cruz

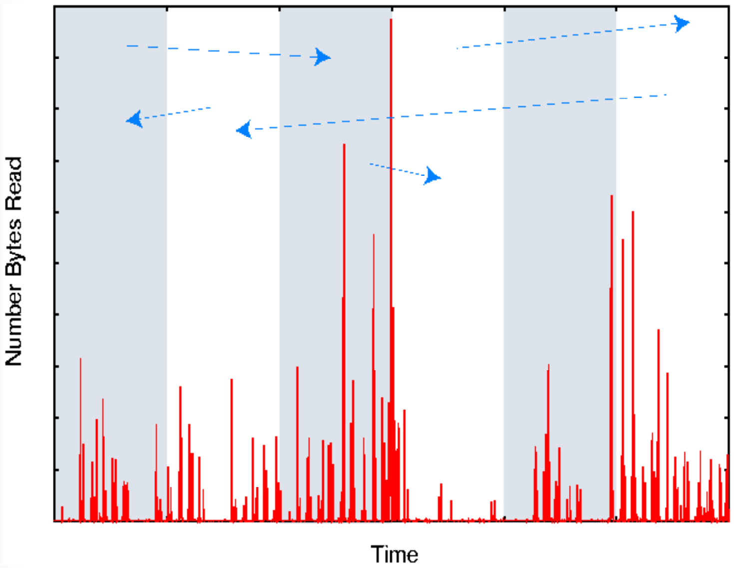
I/O Workload Modeling and Synthesis

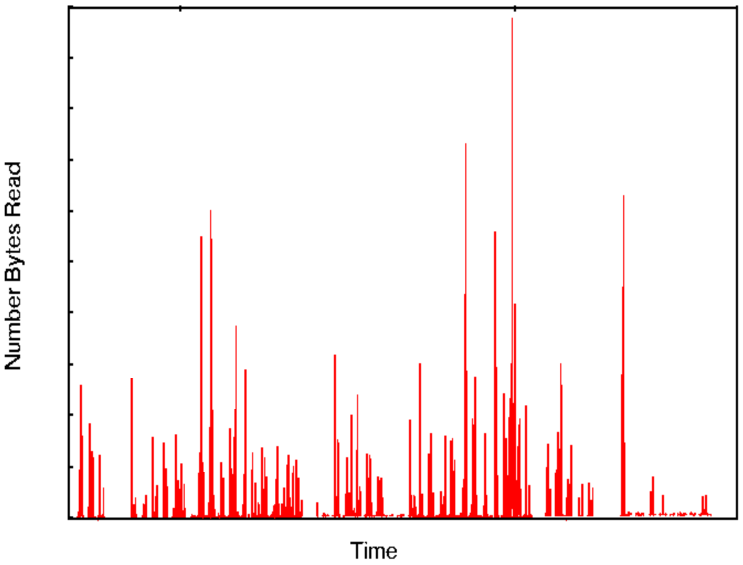
- Block-level disk I/O trace
  - Timestamp, r/w, offset, size
- Important because …
  - Real traces are difficult to obtain
  - Performance analysis and architecture of storage systems depend upon traces and simulations
- Difficult because …
  - Disk traffic is bursty
  - No consensus on what a “good” model should capture
    - We focus on performance-related characteristics
- Still an unsolved problem [Ganger]

Long-Range Dependence in Disk Traffic

- Long-range dependence
  - Non-negligible event correlations across large time spans
  - Measured by the Hurst coefficient (0.5 ~ 1.0)
- Observed at …
  - Network traffic [Leland et al.]
  - Web traffic [Crovella and Bestavros]
  - File system traffic [Gribble et al.]
  - Disk I/O traffic [Gomez and Santonja]
- Is long-range behavior really important?

Removing Long-Range Dependence



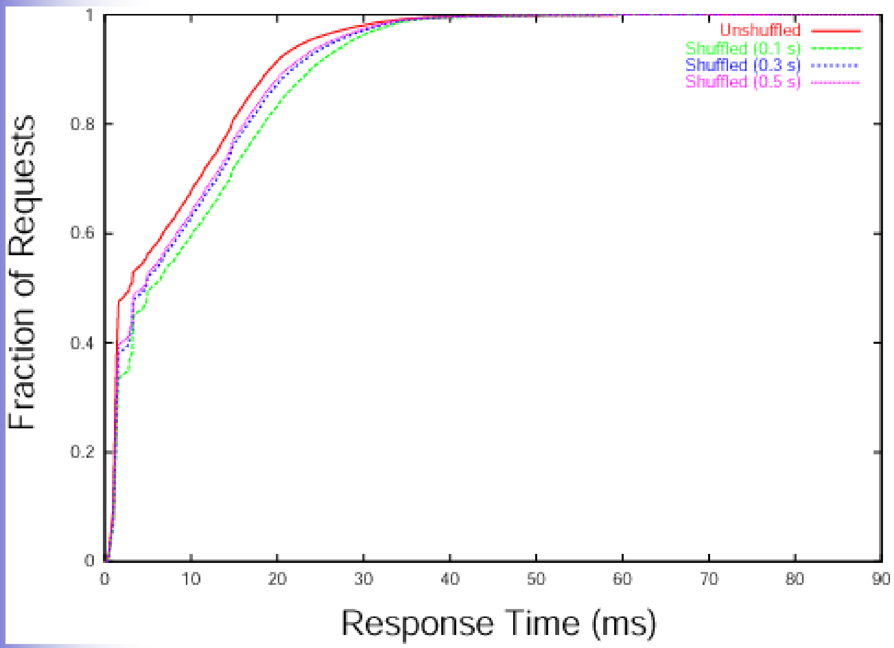


- Shuffling trace intervals removes long range dependence
- The technique was originally proposed in [Grossglauser and Bolot]

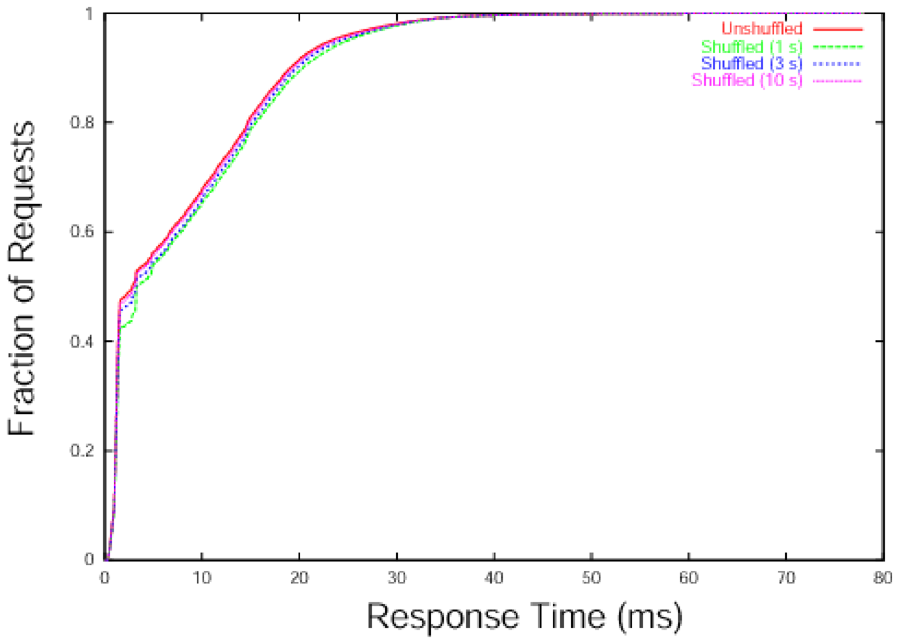
Experimental Methodology

- HPL Pantheon I/O subsystem simulator [Wilkes]
- Metric: root mean squared (RMS) horizontal distance between the CDF of I/O response times [Ruemmler and Wilkes]
- Workloads [Ruemmler and Wilkes]
  - Cello news server traces
    - Random accesses
  - Snake file server traces
    - Sequential accesses

Results For Snake File Server



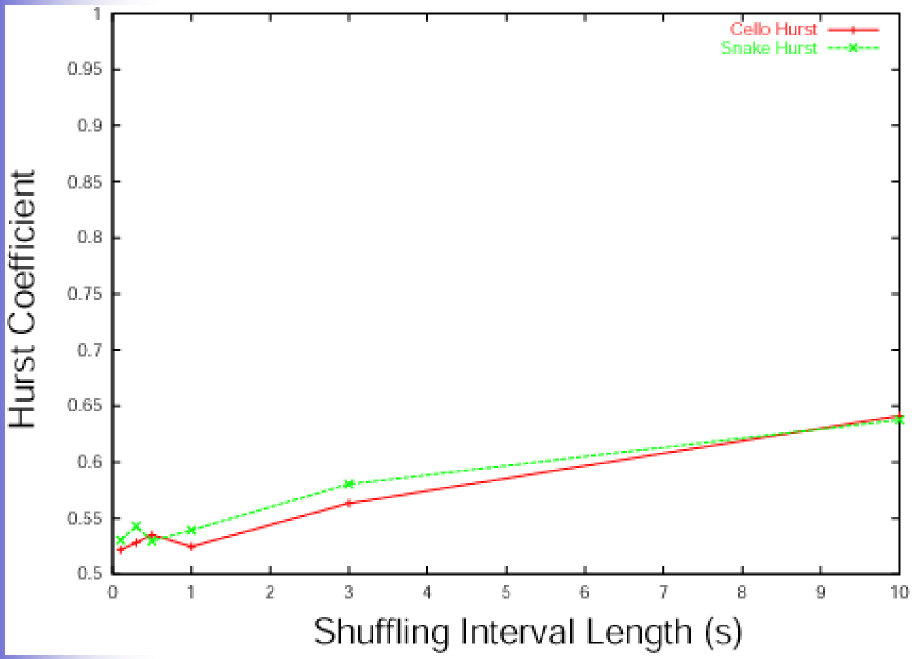
Small shuffling intervals



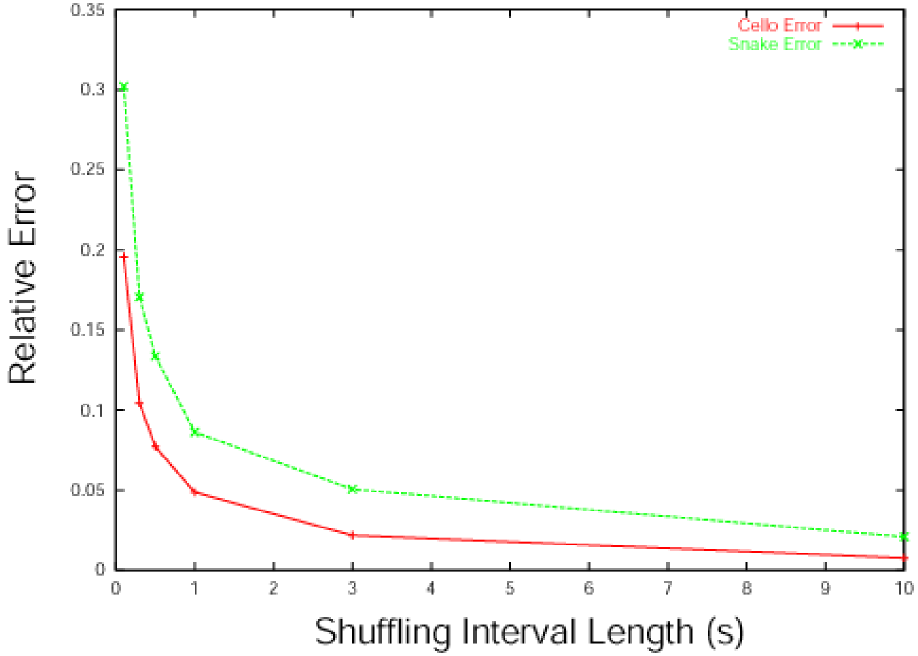
Large shuffling intervals

- Shuffling at very small time scales perturbs disk behavior
- Correlations longer than one second do not affect disk response times appreciably

Experimental Results



Hurst coefficient



Relative error (RMS)

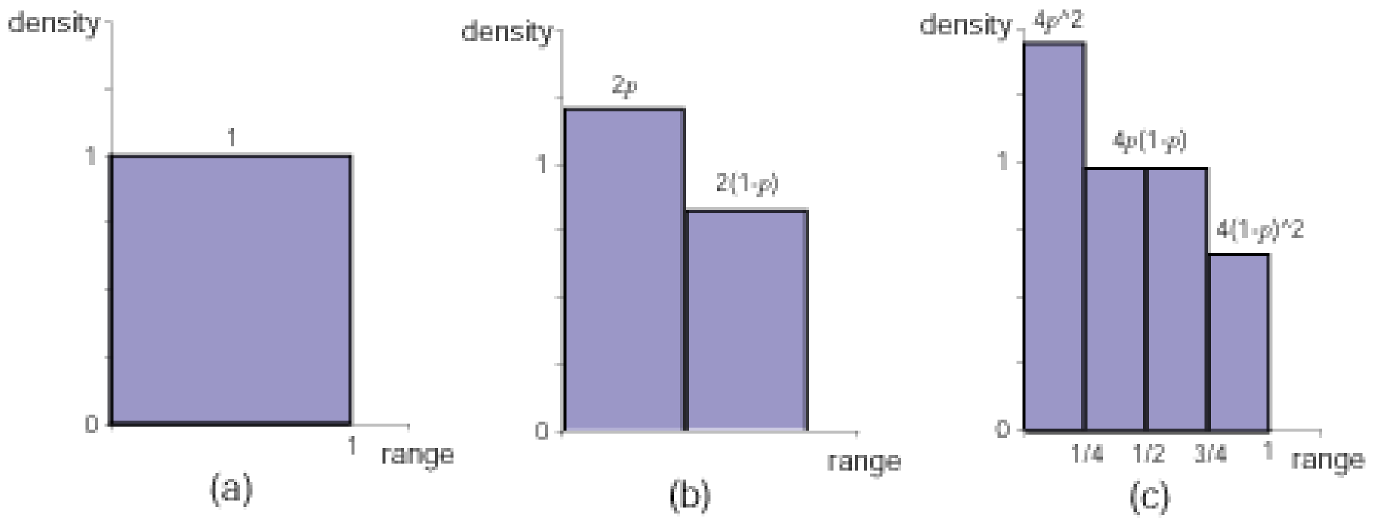
- Long-range dependence is removed by shuffling
- Relative error is still small
- Dependence at large time scales in I/O traffic does not significantly affect disk behaviors with respect to response times

Implications of the Irrelevance of LRD

- Implications to workload modeling and synthesis
  - Short trace intervals can be considered as independent
- Binomial multifractals for local burstiness in I/O access patterns
- Reconstruct quality synthetic workloads from a small set of representative trace intervals, which are selected using cluster analysis based on metrics related to disk behaviors (in another paper)

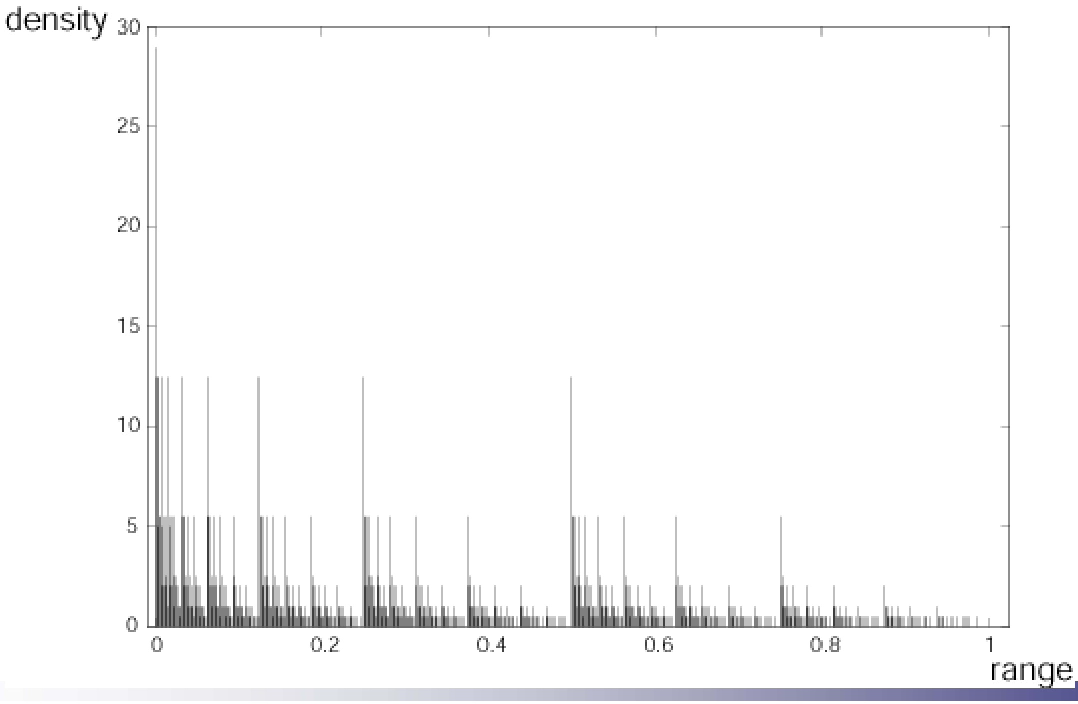
Multifractal-Based I/O Trace Synthesis

- Binomial multifractals
  - 80/20 "laws"
  - Bias ***p*** (> 0.5)
  - Previously used in I/O traffic modeling by [Wang and Madhyastha et al.]



Multifractal-Based I/O Trace Synthesis

- Binomial multifractals
  - 80/20 “laws”
  - Bias ***p*** (> 0.5)
  - Previously used in I/O traffic modeling by [Wang and Madhyastha et al.]



Estimating Bias *p*

- Entropy values at different aggregation levels

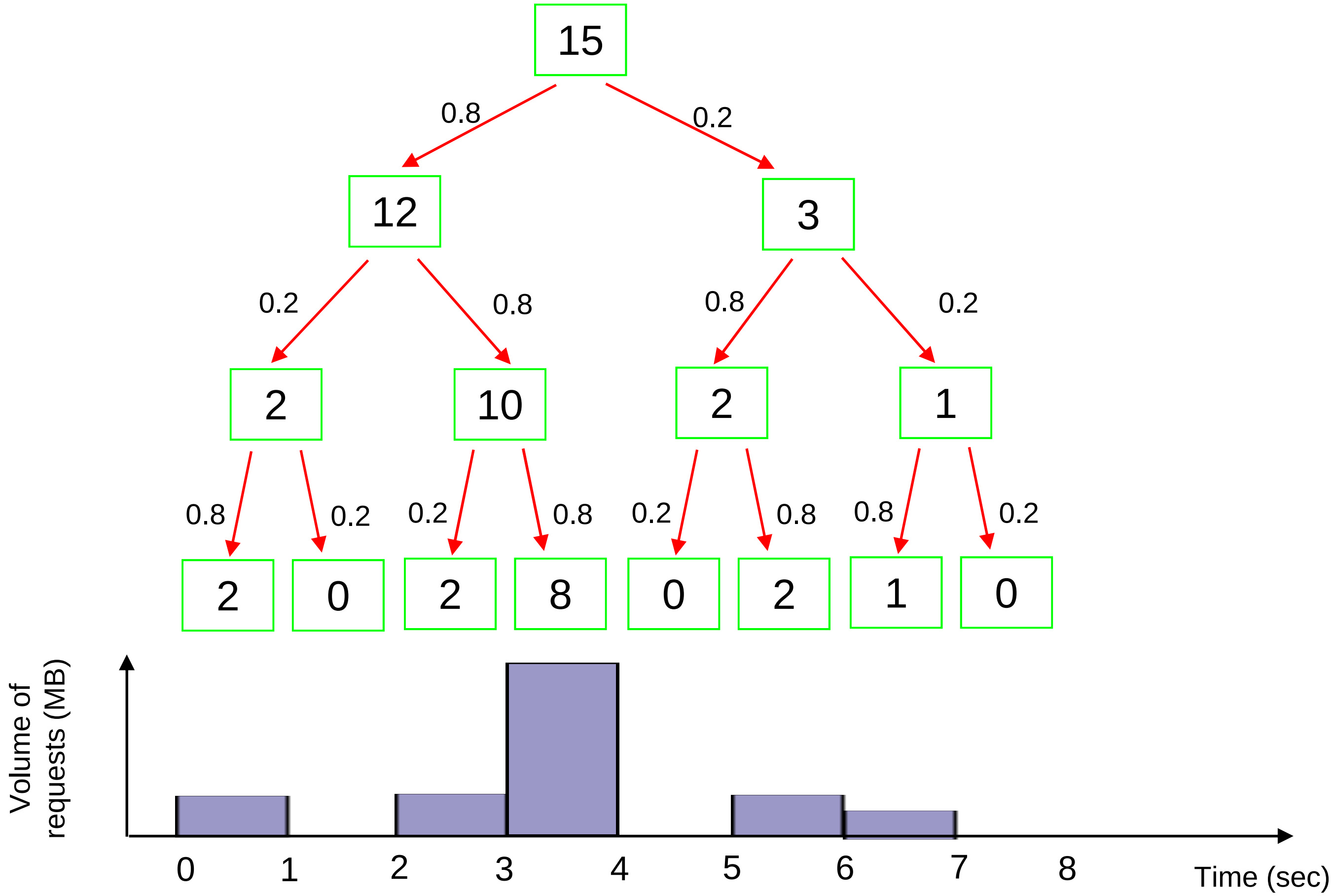
$$Y_t^{(n)}(k) = \int_{k2^{-n}}^{(k+1)2^{-n}} Y_t dt$$ **, where *k* = 0, 1, …,** $2^n - 1$

$$E_p^{(n)} = -\sum_{k=0}^{2^n-1} \frac{Y_t^{(n)}(k)}{\int_0^1 Y_t dt} \log_2 \frac{Y_t^{(n)}(k)}{\int_0^1 Y_t dt}$$

$$E_p^{(1)} = -p\log_2 p - (1-p)\log_2(1-p)$$

Binomial Multifractal Generation

Bias ***p***: 0.8, time interval length ***l***: 8 sec, volume of requests ***m***: 15 MB



Synthetic Trace Generation

Input: original trace file ***f***, interval length ***s***

Output: a time-stamped request sequence

$$((t_1, m_1), (t_2, m_2), \ldots, (t_n, m_n))$$

For each non-empty interval ***w*** in ***f***

Calculate bias ***p***

mass = aggregated request size in interval ***w***

Binomial-Multifractal-Generation (***p***, ***s***, ***mass***)

map local time-stamps to real time-stamps

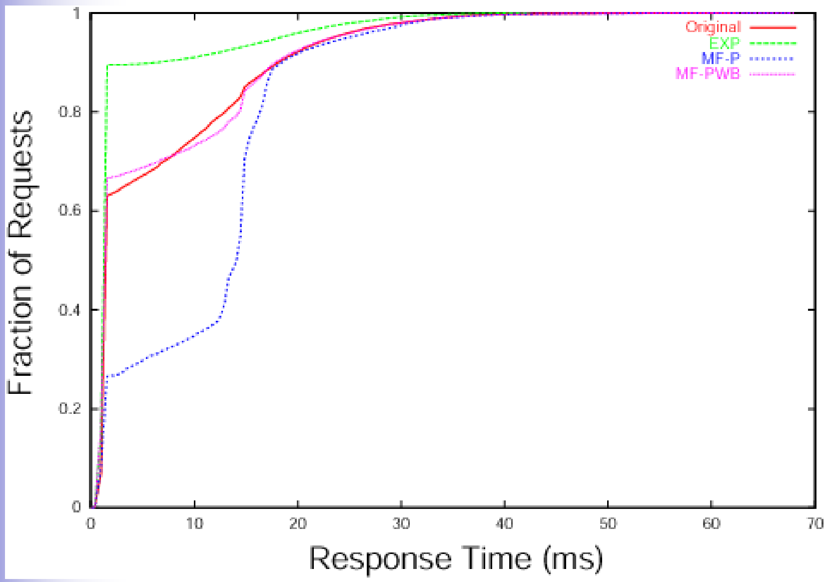
End For

Synthesis Method Description

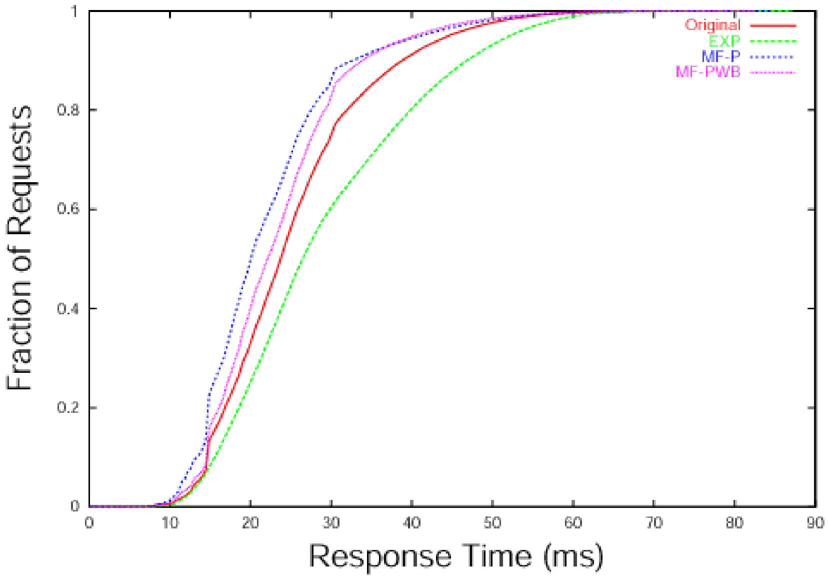
	EXP	MF-P (CMU)	MF-PWB
Interarrival pattern	Exponential	Binomial Multifractals	Binomial Multifractals
Timescale	N/A	Large (1 day)	Small (seconds)
Synthesizes	Interarrival time	Interarrival time and request size	Interarrival time and request size
Does not synthesize*	Request size, sector ID, r / w	Sector ID, r / w	Sector ID, r / w

* The information that is not synthesized is retained from real traces

Synthetic Traces



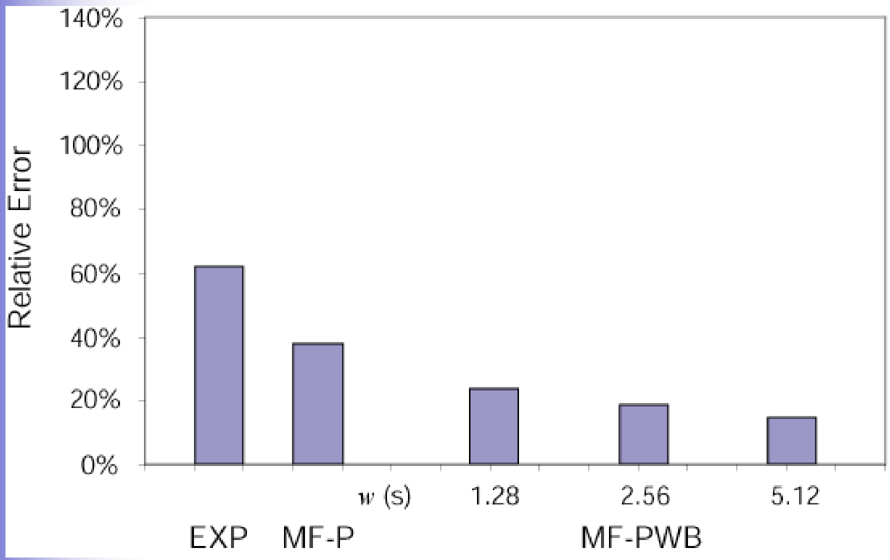
Snake file server



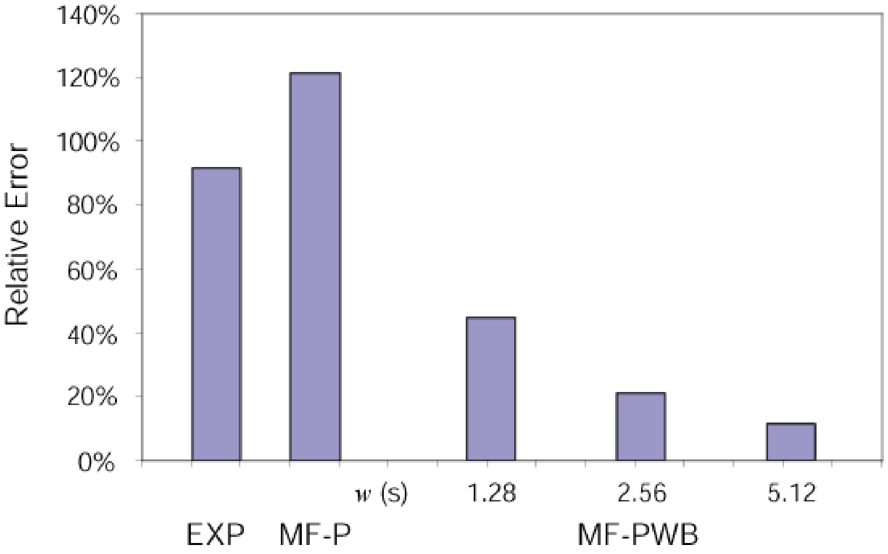
Cello news server

- Binomial multifractals models the interarrival pattern well
- Exponential does not
- Generating at small timescales improves synthesis quality

Experimental Results



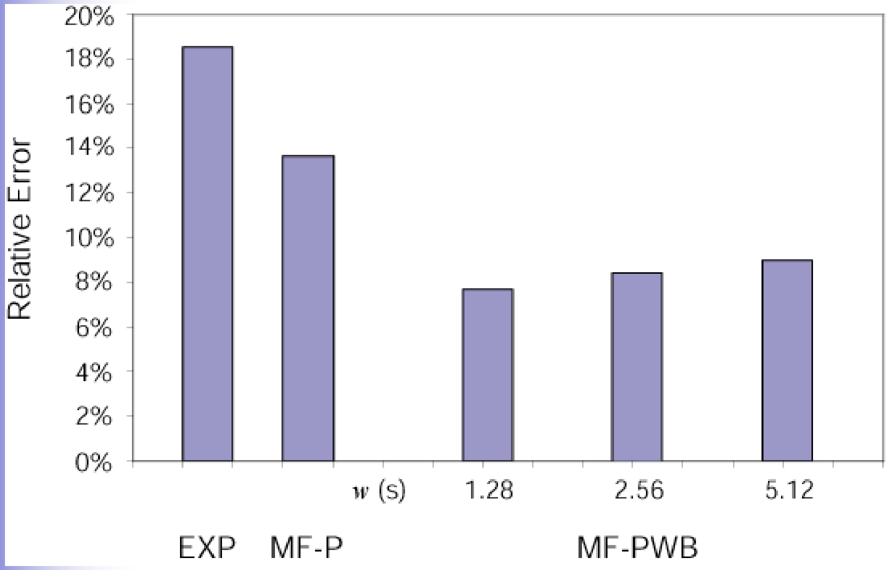
Snake file server (Day 1)



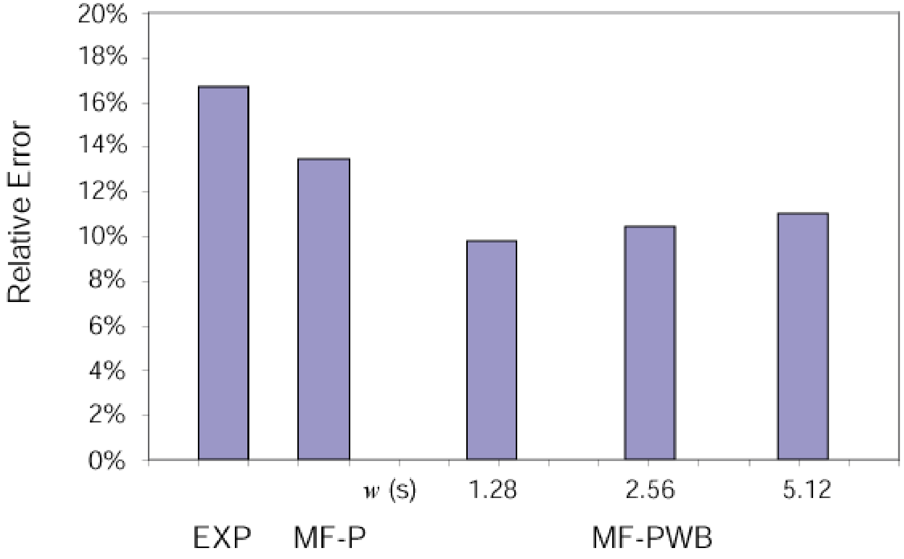
Snake file server (Day 2)

- MF-PWB method can generate the most accurate synthetic traces

Experimental Results



Cello news server (Day 1)



Cello news server (Day 2)

- The generation interval length matters
- The majority of synthesis errors by MF-P and MF-PWB comes from synthetic arrival time (results shown in paper)

Conclusions and Future Work

- Long range dependence is not important for reproducing certain performance metrics
- Binomial multifractals well capture local burstiness in I/O interarrival patterns
  - Achieve a 8 –12% demerit factor for I/O response times on random and sequential workloads

- Automatically determine appropriate interval length
- Need to quantify spatial locality

- PQRS model [Wang et al.]
  - Use a two-dimensional version of binomial multifractals
  - Model both spatial and temporal localities as well as their correlations

Thank You!

- Acknowledgements
  - Our shepherds Julian Satran and Robert Chadduck
  - Mengzhi Wang and Christos Faloutsos from CMU PDL
  - John Wilkes and Kimberly Keeton from HPL SSP
  - Eitan Bachmat from the Ben-Gurion University, Israel
  - UCSC Storage Systems Research Center
- More information:
  - http://www.soe.ucsc.edu/~tara/stargroup
  - http://www.soe.ucsc.edu/~hongbo/publications.html
- Questions?