Assert(!Defined(Sequential I/O))

Cheng Li*, Philip Shilane, Fred Douglis, Darren Sawyer, and Hyong Shim

**Rutgers University* *EMC Corporation*

EMC²

RUTGERS

Sequential I/O is Important

- Driven by traditional storage characteristics
  - Optimization based on sequentiality (hard drive, tapes)
  - Applications leverage sequentiality: e.g., cache, prefetch
- Non-rotational devices favor sequential I/O
  - Large unit SSD writes to reduce erasures
- Help classify workload characteristics
  - Benefit system researchers, trace analysis, I/O generation
  - Consistent metric makes results comparable

EMC²

RUTGERS

Sequential I/O is NOT Straightforward

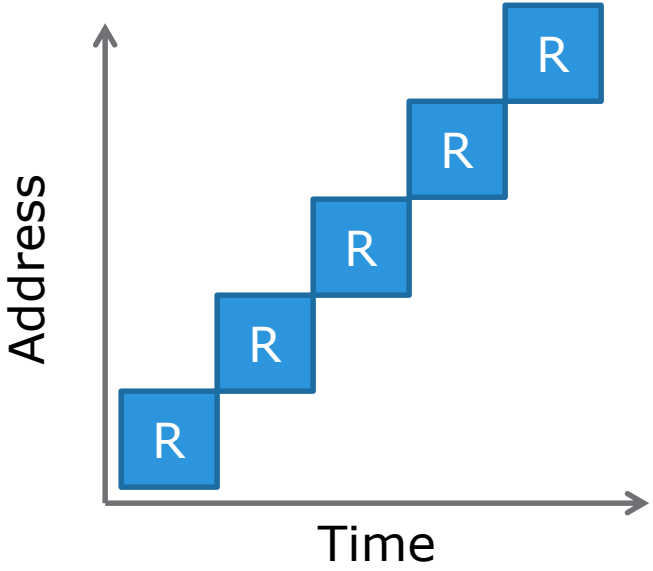
- Definition of sequential I/O matters!
  - Sequentiality changes significantly based on definitions
- Inconsistent definitions of sequential I/O
  - "We consider two read or writes requests to be sequential if they are consecutive in time, and the file offset + request size of the first request equals the file offset of the second request." [Chen'11]
  - "…we consider sequential I/O as the number of bytes transferred before a random seek" [Shim'13]
  - "…this series to be sequential despite the missing 1k between the third and fourth requests." [Ellard'03]
- Data-driven approach to compare sequentiality definitions

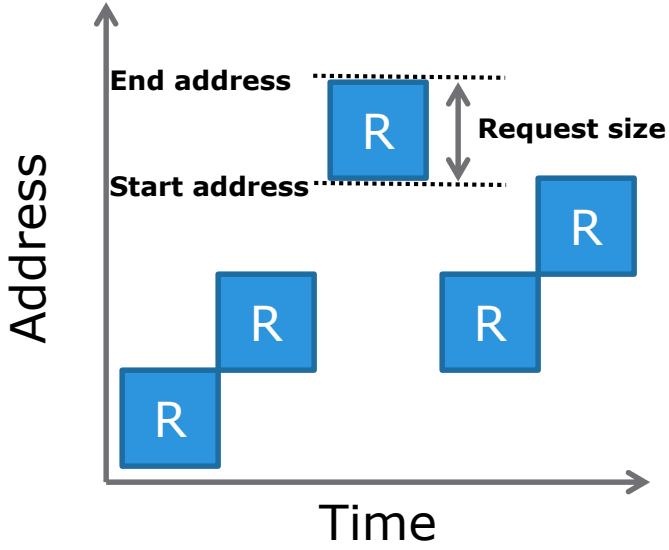
EMC²

RUTGERS

Consecutive Access Ratio (CAR)

- Canonical definition
  - Fraction of consecutive accesses
  - Example: 4/4=100% vs. 2/4=50% (ignore first access)



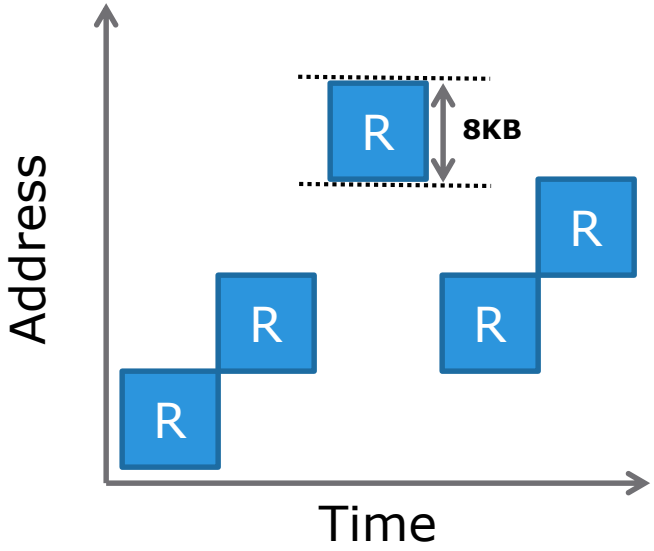


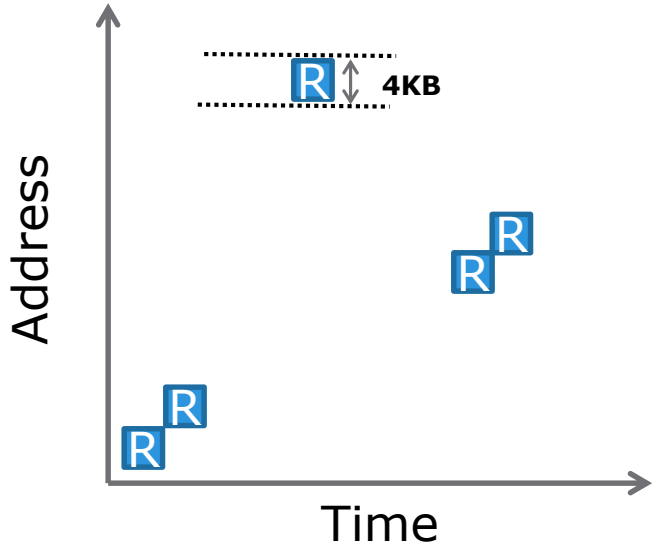
EMC²

RUTGERS

Consecutive Bytes Accessed (CBA)

- I/O size should be incorporated
  - Consecutive bytes accessed between seeks
  - Example: 5*8KB/2 = 20KB vs. 5*4KB/2 = 10KB



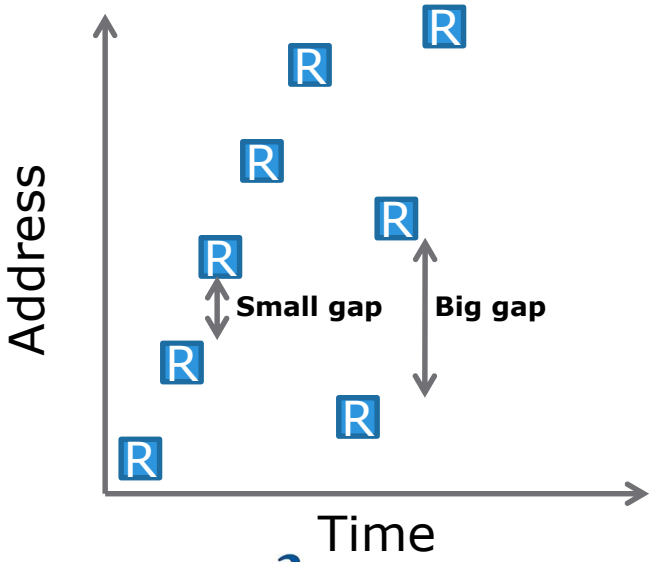


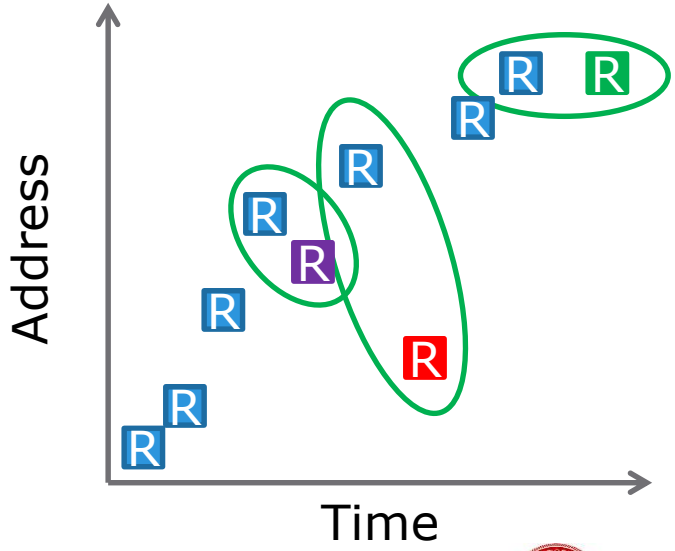
EMC²

RUTGERS

Strided Range (SR)

- Non-consecutive accesses can be sequential
  - The range property allows gaps, small backward seeks, and re-access of an address to still be considered sequential.



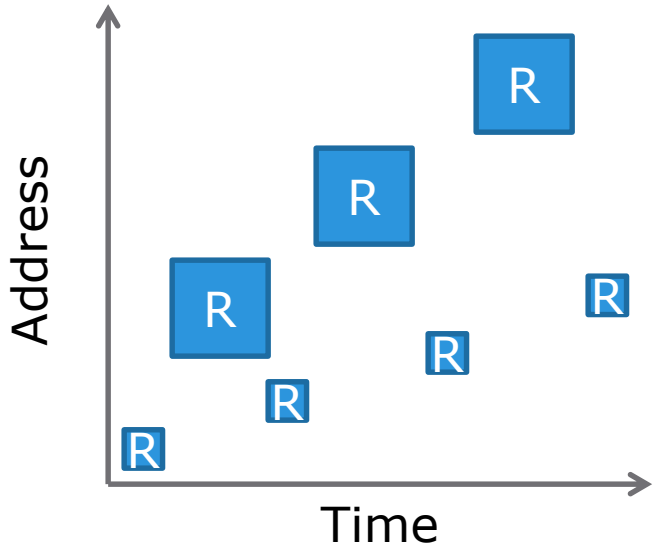


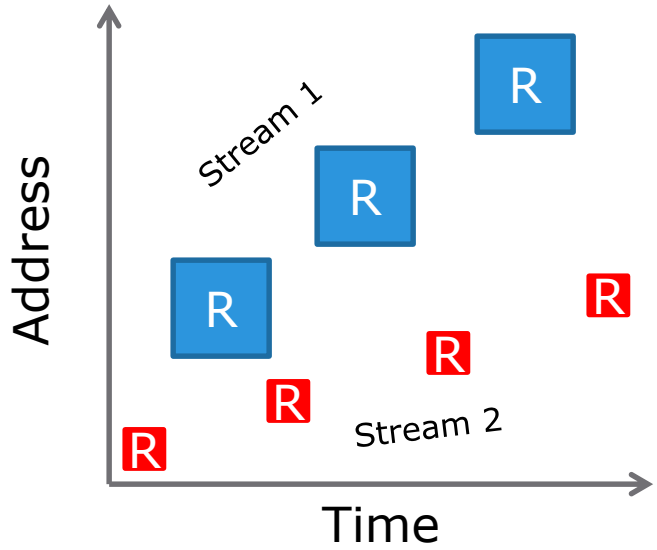
EMC²

RUTGERS

Multi Streams (MS)

- Single stream vs. multiple streams
  - Multiple streams are interleaved
  - Multi-threading, virtual machines



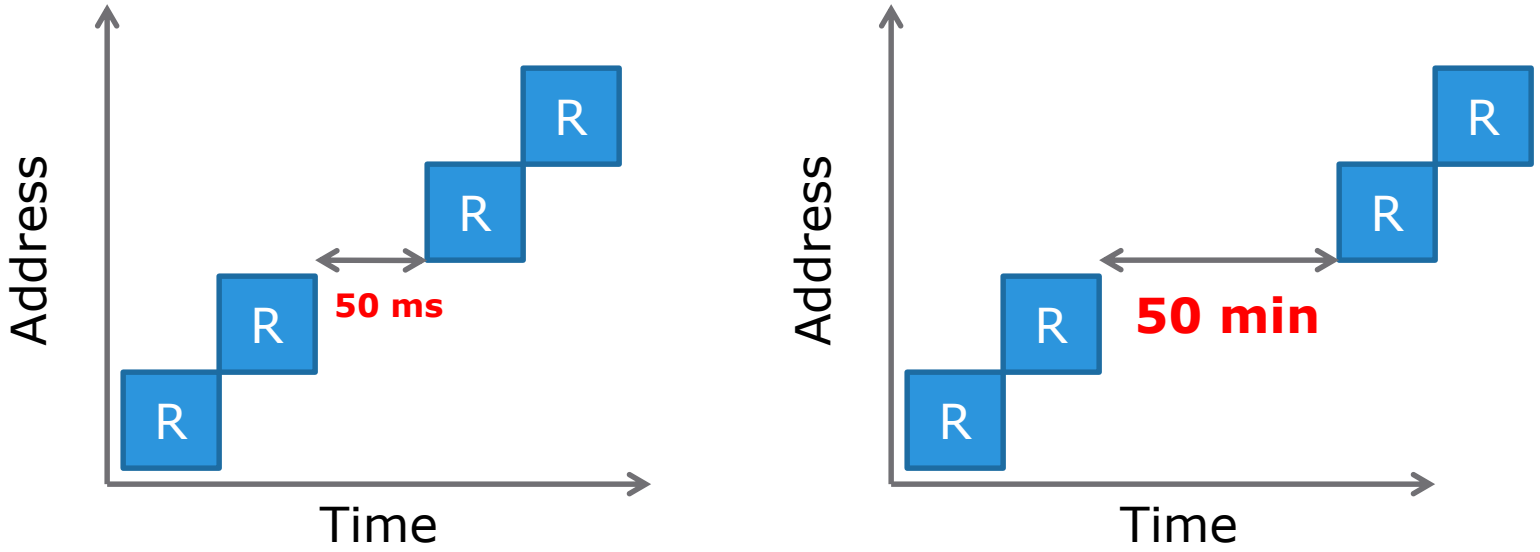


EMC²

RUTGERS

Inter-arrival Time (IT)

- Think time between two I/Os
- Background tasks cause disk movement



Problem Recap

- Canonical consecutive access ratio (CAR)
  - Can be too restrictive
  - Insufficient to distinguish different access patterns
- Need to explore other metric alternatives
  - Captures important properties
  - Calibrates locality accurately
- Study different sequentiality metrics
  - Pick a metric that best aligns with the use case
  - Use standard statistics tools to study correlation

EMC²

RUTGERS

Sequentiality Metrics

- Metric table explores all property combinations

Family	Metric	SR	MS	IT	Metric	Family
F1 CAR-based (ratio)	M1				M9	F2 CBA-based (bytes)
	M2	X			M10	
	M3		X		M11	
	M4			X	M12	
	M5	X	X		M13	
	M6	X		X	M14	
	M7		X	X	M15	
	M8	X	X	X	M16	

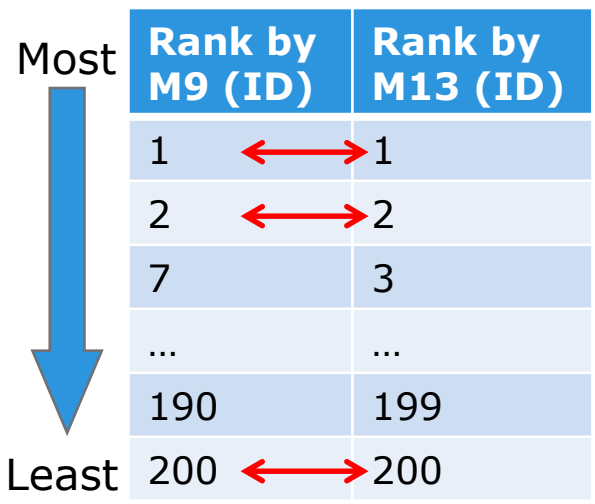
SR=Stride Range
MS=Multi-stream
IT=Inter-arrival Time

EMC²

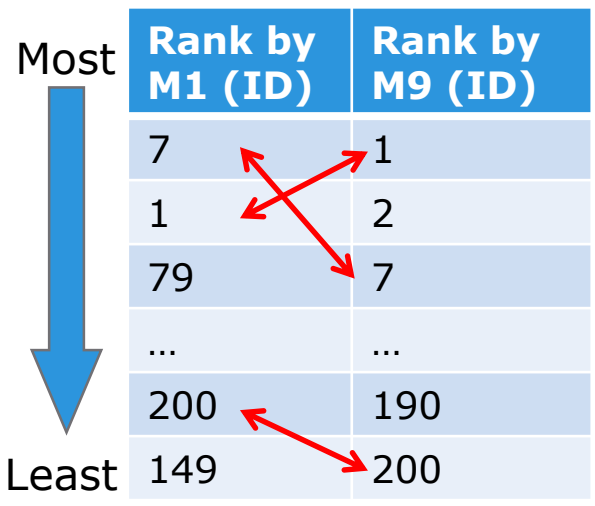
RUTGERS

Experimental Methodology

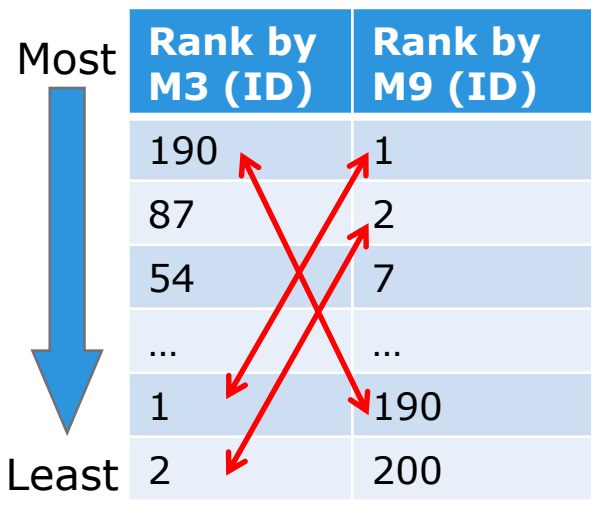
- Correlation analysis
  - Metrics have different range (ratio vs. size)
  - Compare rank order of traces to study correlation



High correlation



Weak correlation



Negative correlation

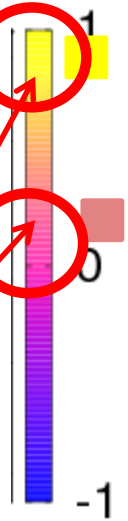
Correlation Matrix

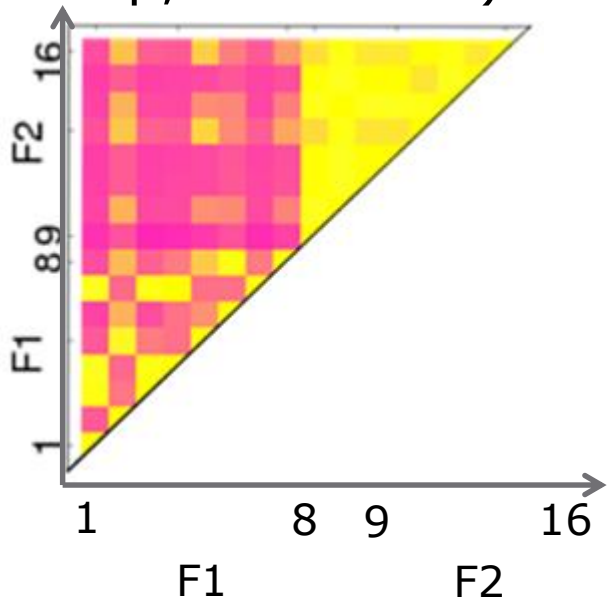
- Compute for each metric-pair (total 16 x 16 = 256)
  - Use standard statistics tools (Spearman's ρ, Kendall's τ)

Most

Rank by M9 (ID)	Rank by M13 (ID)
1	1
2	2
7	3
…	…
190	199
200	200

Least





Correlation(M1,M9)=0.12
Correlation(M9,M13)=0.91

EMC²

RUTGERS

Traces

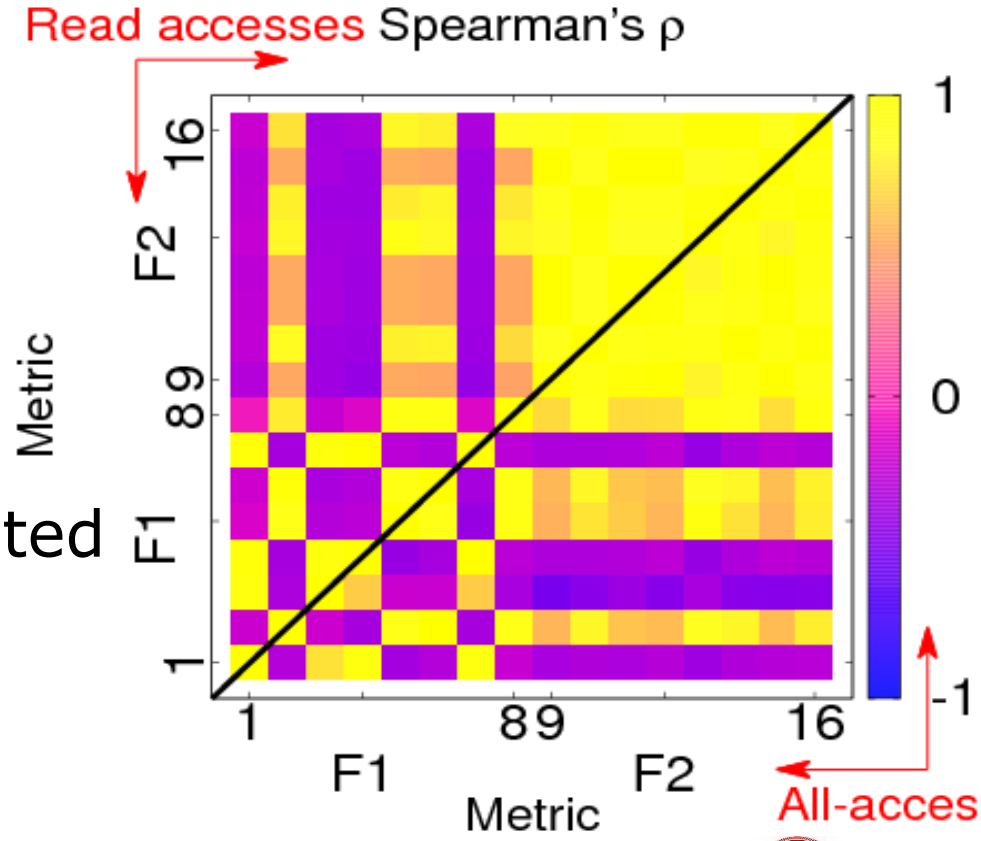
- Private traces
  - 294 EMC VMAX traces
- Public traces (94 in total)
  - 5 IBM-HP
  - 36 Microsoft research
  - 50 Microsoft production
  - 3 Florida International University (FIU)
- Covers both sequential and random traces

EMC²

RUTGERS

Comparing Sequentiality Metrics

- All accesses
  - Cross family
  - Family 1 (CAR)
  - Family 2 (CBA)
- Read accesses
- Negatively correlated

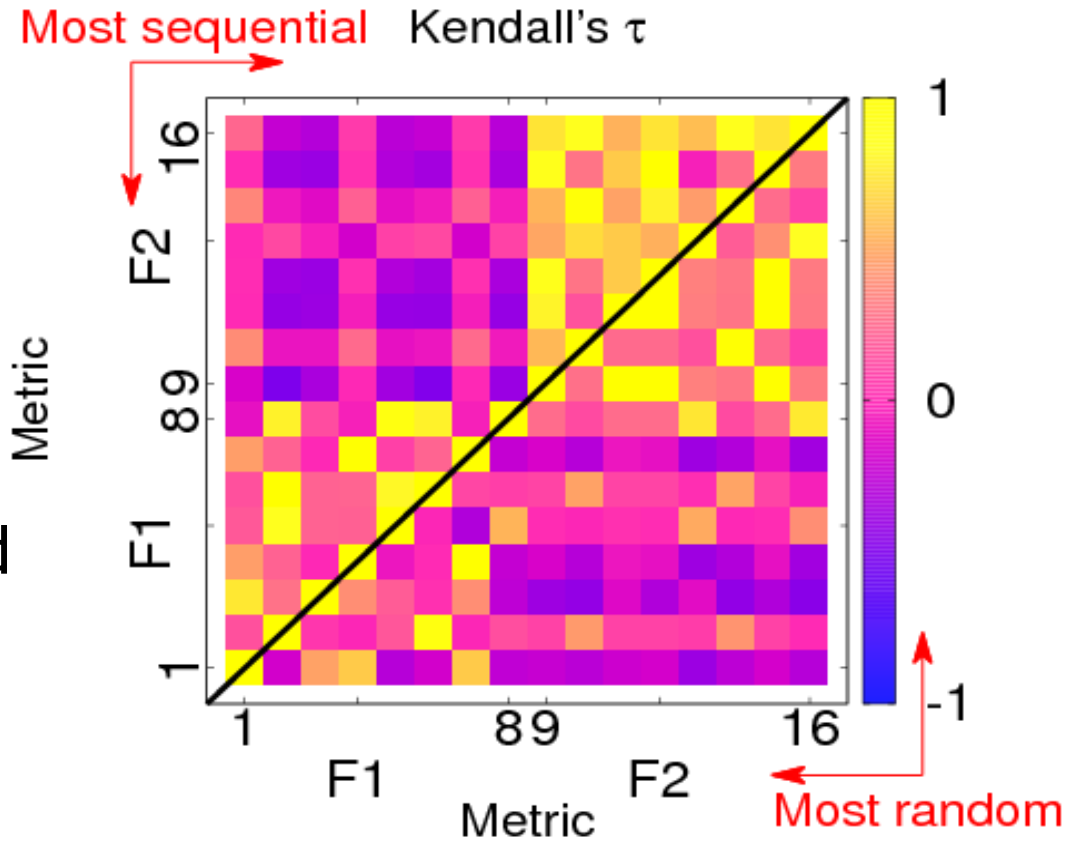


EMC²

RUTGERS

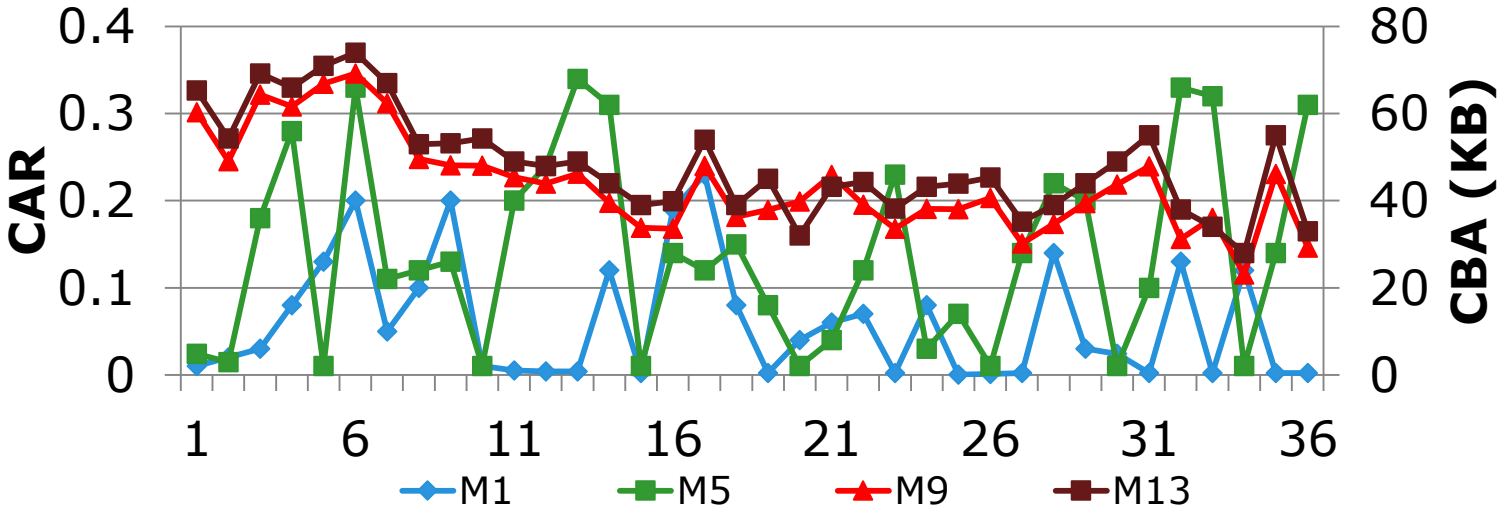
Most/Least Sequential Comparison

- Most random
  - Cross family
  - Family 1 (CAR)
  - Family 2 (CBA)
- Most sequential
- Negative correlated



Sequentiality Over Time

- Calculate metric value every 10min
  - CAR-based M1, M5 (left-axis), CBA-based M9, M13 (right-axis)
- F2 metrics are more consistent than F1 metrics



EMC²

RUTGERS

Discussion

- Controversy of sequential I/O metrics
- I/O size matters
  - CBA based metrics are more consistent
- Strided range matters
  - Non-consecutive accesses can be sequential
- Incorporate domain knowledge
  - Apply domain knowledge and state your metric

EMC²

RUTGERS

Q&A

Assert(!Defined(Sequential I/O))

Cheng Li*, Philip Shilane, Fred Douglis, Darren Sawyer, and Hyong Shim

Cheng Li* *Rutgers University* Philip Shilane, Fred Douglis, Darren Sawyer, and Hyong Shim *EMC corporation*

EMC²

RUTGERS