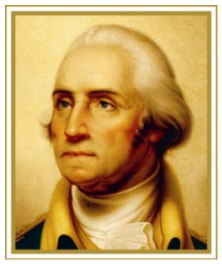


THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

Reconfigurable Active Drive
An FPGA Accelerated Storage Architecture for Data-Intensive Applications

Teng Li, Miaoqing Huang, Tarek El-Ghazawi, and **Howie Huang**

July 30, 2009

SAAHPC'09, Urbana-Champaign, IL

Overview

- Motivations
- Related Work
- RAD Architecture/Model
- Evaluation
- Conclusion

Motivations

- Data set continues to grow rapidly
- Hard disk performance improves slowly
- Focus on end-to-end throughput to memory
  - I/O bottleneck remains

Related Work

- Active disk: moving data processing units close to data
  - Acharya et al. proposed a stream-based programming model to allow "disklet" to execute efficiently and safely
  - Huston et al. proposed a similar "Diamond" architecture to use "searchlet" to filter data
- Franklin et al. proposed an architecture including chip multiprocessors (CMPs) and FPGAs
- Netezza, http://www.netezza.com

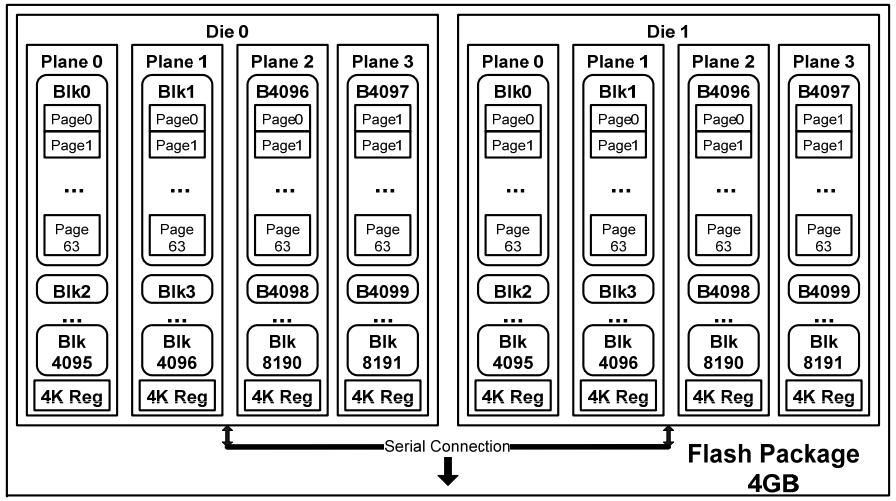
Reconfigurable Active Disk (RAD)

- Explore parallelism in both FPGAs and solid-state drives (SSDs)
- Maximize SSD data transfer rate to match FPGA processing throughput
- Killer applications
  - Data mining
  - Crypto analysis
  - Image processing
  - Many more

Solid-State Drives

- Solid State Drives (SSDs)
  - Use non-volatile flash memory
  - NAND and NOR
  - SLC and MLC

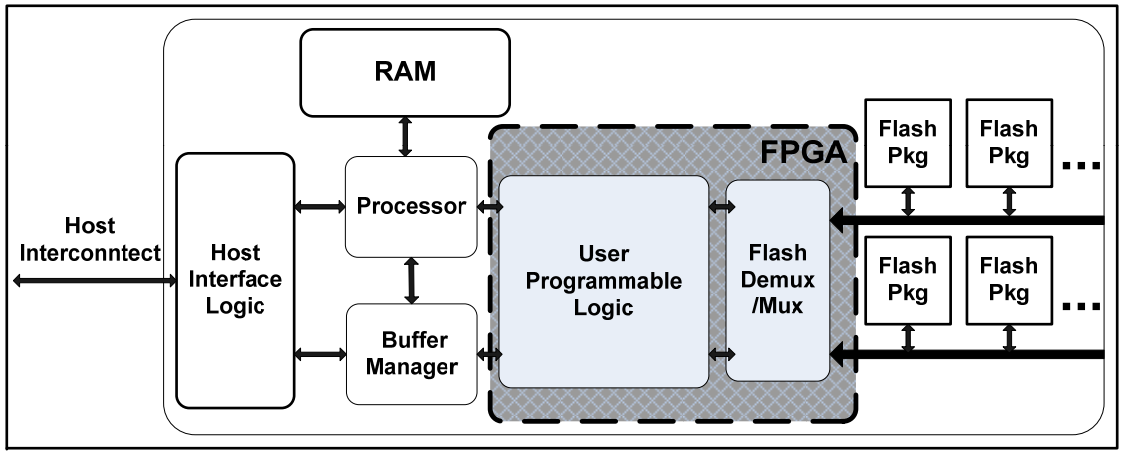
- Samsung K9XXG08UXM
  - Take 100 µs to transfer one page (4 KB) of data from internal register to I/O
  - 40 MB/s bandwidth for one chip
  - Page access time 25 µs
  - Page write time 200 µs
  - Block erase time 1.5 ms
  - Page sizes of 4 Kbytes
  - Block sizes of 64 pages
  - Non-uniform read/write performance
  - Limited erase cycles



- N. Agrawal, V. Prabhakaran, T. Wobber, J. D. Davis, M. Manasse and R. Panigrahy, "Design Tradeoffs for SSD Performance," in *Usenix Annual Technical Conference* (USENIX '08), June 2008, pp. 57--70.
- *256M x 8 Bit / 128M x 16 Bit NAND Flash Memory*, Samsung, Inc., Aug. 2003.

RAD Architecture

- FPGA resides between the SSD controller and flash package array



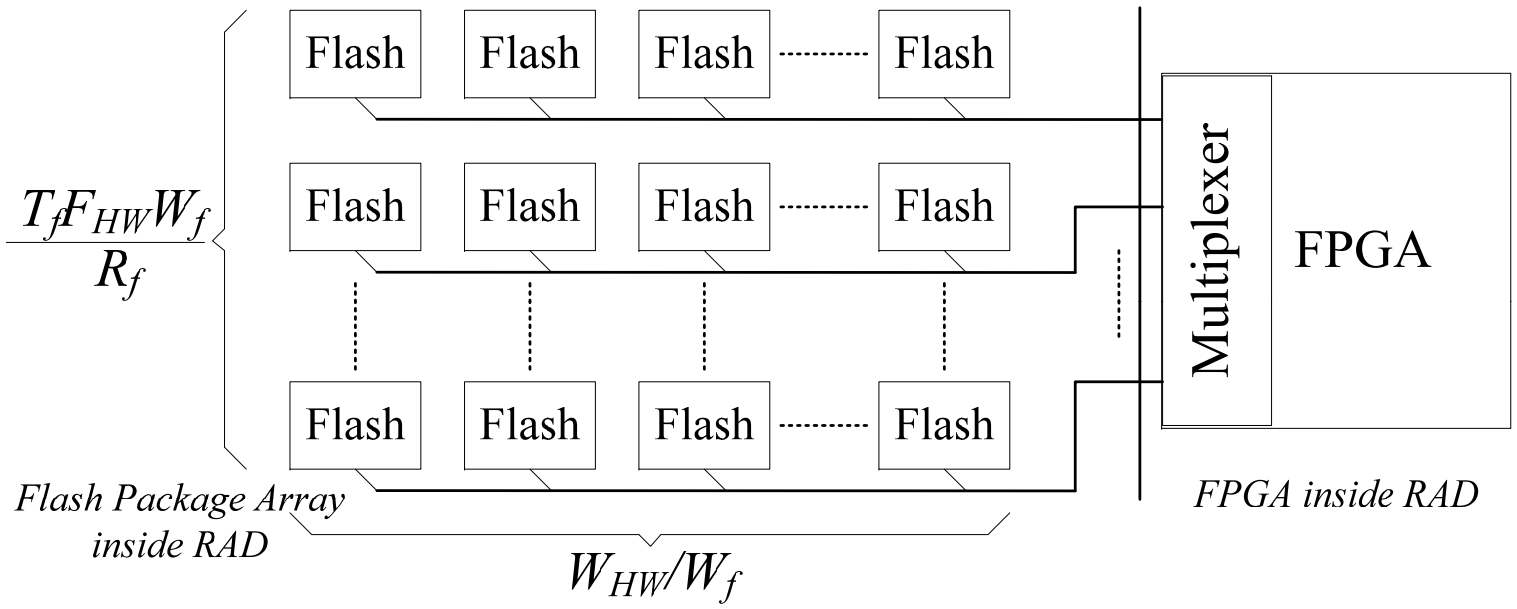
- N. Agrawal, V. Prabhakaran, T. Wobber, J. D. Davis, M. Manasse and R. Panigrahy, “Design Tradeoffs for SSD Performance,” in *Usenix Annual Technical Conference* (USENIX '08), June 2008, pp. 57--70.

RAD Model

- Maximum data access rate from a single flash chip is $\mathbf{R_f / T_f}$ Bytes/Second
- FPGA processing throughput $\mathbf{W_{HW} \times F_{HW}}$
- Required number of flash chips is $\mathbf{(T_f \times W_{HW} \times F_{HW}) / R_f}$
- The number of flash package the FPGA could access at one time is determined by
  - The number of row is $\mathbf{W_{HW}/W_f}$
  - The number of column is $\mathbf{(T_f \times W_f \times F_{HW}) / R_f}$

	Parameters			
	Description	**Symbol**	**Unit**	**Setup Values**
Flash Package inside RAD Model	Data Register Size	R_f	Byte (B)	4096 B
	Serial Data Access Time to Data Register	T_f	Second (S)	0.0001 S
	I/O Width	W_f	Byte (B)	1 B
FPGA inside RAD Model	Frequency	F_{HW}	Hertz (Hz)	200 MHz
	I/O Width	W_{HW}	Byte (B)	16 B

RAD Architecture



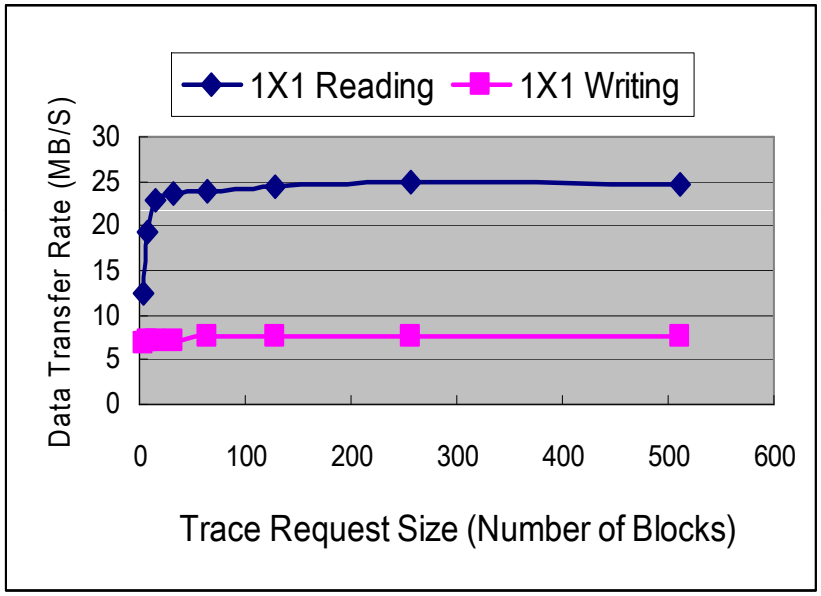
- In most current SSDs, the performance bottleneck is the device interface

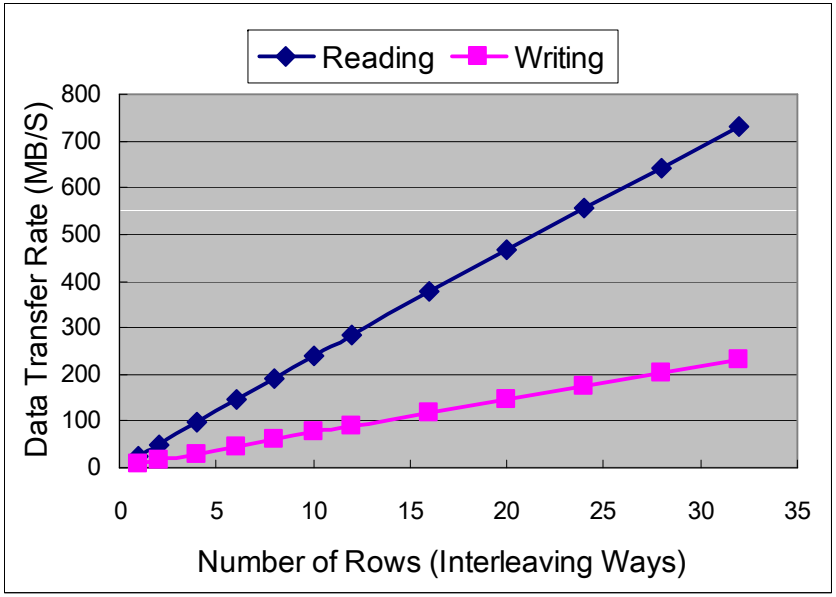
SSD Simulation

- Simulate 1x1 flash package
  - DiskSim (CMU) and SSD extension (MSR)
  - Trace Files collected from OS level using BlkTrace
  - Modify the trace files slightly to increase SSD bandwidth

- Use N x 1 flash package as the baseline
  - N rows and 1 column, i.e., N-way interleaving
  - For 16-byte I/O width, 16 Nx1 flash packages are needed

Baseline Performance





- Assuming that bus bandwidth would scale with the bus width

Applications

- Real-life applications
  - DES (16 bytes I/O on FPGA) running on SGI RC100
  - Sobel Edge Detection (8 bytes I/O on FPGA) on Cray XD1

	DES	**SED**
FPGA Processing Rate	1.14 GB/s (collected on SGI)	1.04 GB/s (collected on Cray)
Required Data Rate for Nx1 SSD model	72.96 MB/s	133.12 MB/s
Required Flash Matrix	10 x 16	18 x 8
RAD Reading Bandwidth	3.72 GB/s	3.33 GB/s
RAD Writing Bandwidth	1.18 GB/s	1.04 GB/s

Speedup

- Compare the end-to-end throughput with
  - Reconfigurable computing (RC)
  - Software implementation of the same applications
- Maximum speedup achieved is 540x

	DES		SED	
RAD end-to-end throughput	1.14 GB/s		0.79 GB/s	
RC end-to-end throughput	23.08 MB/s (collected on SGI)		14.13 MB/s (collected on Cray)	
SW end-to-end throughput	7.34 MB/s (collected on SGI)		1.97 MB/s (collected on Cray)	
Speedup (RC | SW)	50.6	159	75.4	540.6

Conclusion

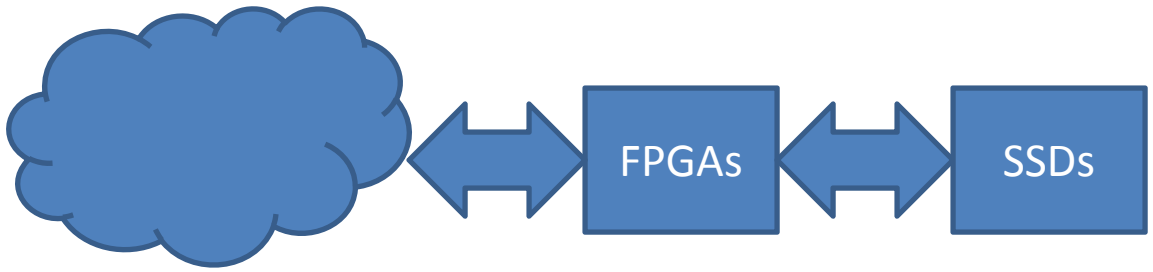
- Reconfigurable Active Drive
  - An FPGA accelerated SSD architecture
  - Fully pipeline the three steps of data retrieval, FPGA processing, and data storage
  - Explore the theoretical model

- Performance advantage
  - RAD vs. reconfigurable computing approach
    - Up to 75 X speedup
  - RAD vs. software approach
    - Up to 540 X speedup

Future Work

- Build prototypes

- Research challenges
  - Interfaces between FPGAs, SSDs, and HDDs
  - Data management
  - Performance
  - Scalability



Thank You