
MEMORY SYSTEMS,
ETC.

Bruce Jacob

University of
Maryland

SOS-13
March 2009

SLIDE 1

The Memory System and You

A Love/Hate Relationship

Bruce Jacob
Electrical & Computer Engineering
University of Maryland at College Park
blj@umd.edu



Good Quotes from Tues.

Al Geist: Apps get bigger and more complex

Thomas Schulthess: Memory BW primary limiter to solving superconductivity equations

Karl-Heinz Winkler: Memory = 50% power (cf. Zia spec: proc=214W, proc+mem=230W)

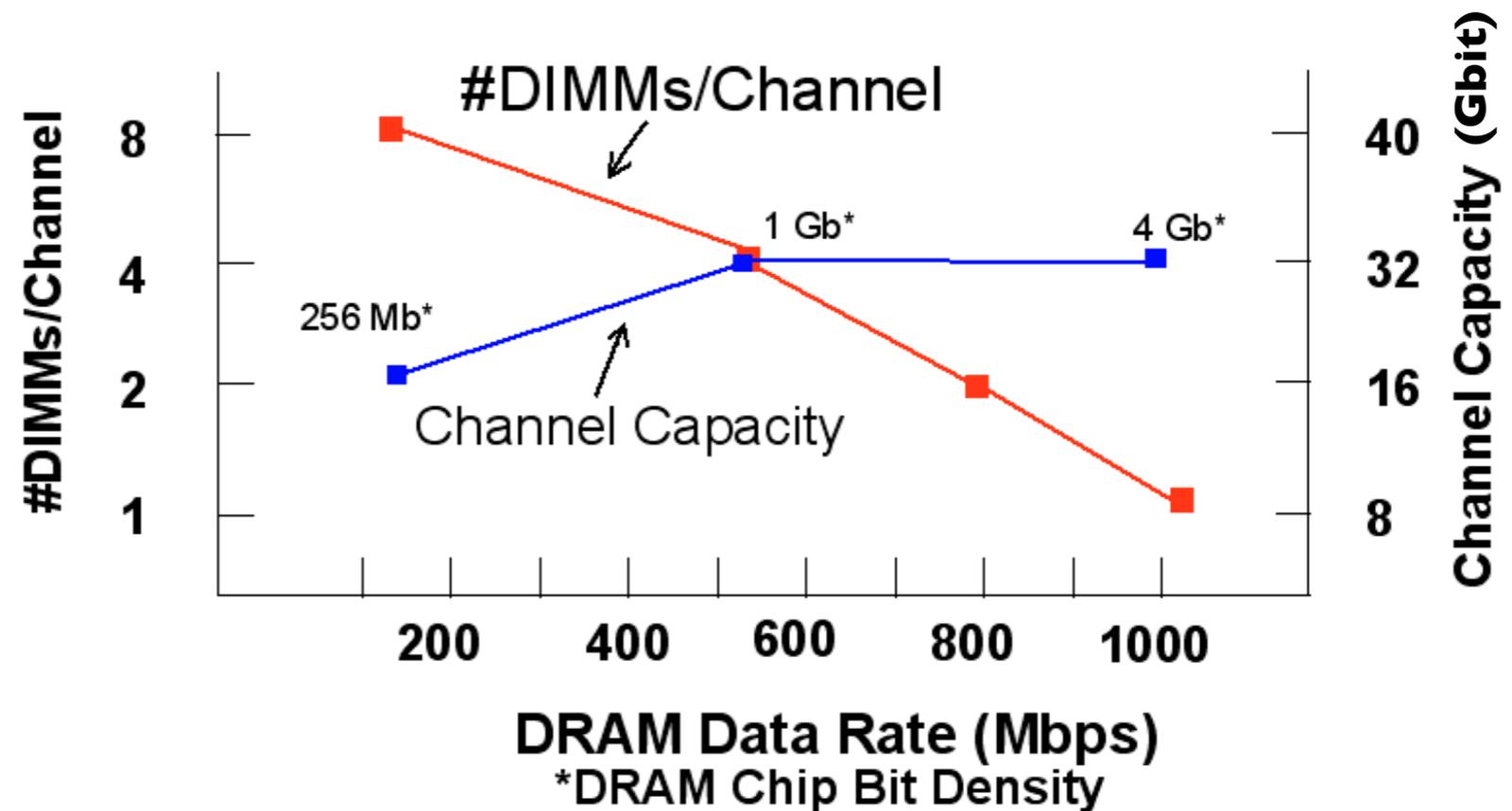
Chuck Moore: image/video as data types => larger working sets, larger data types

Bill Camp:

“DRAM sucks—that’s the real problem”

Some Trends

Storage per CPU socket has been flat:

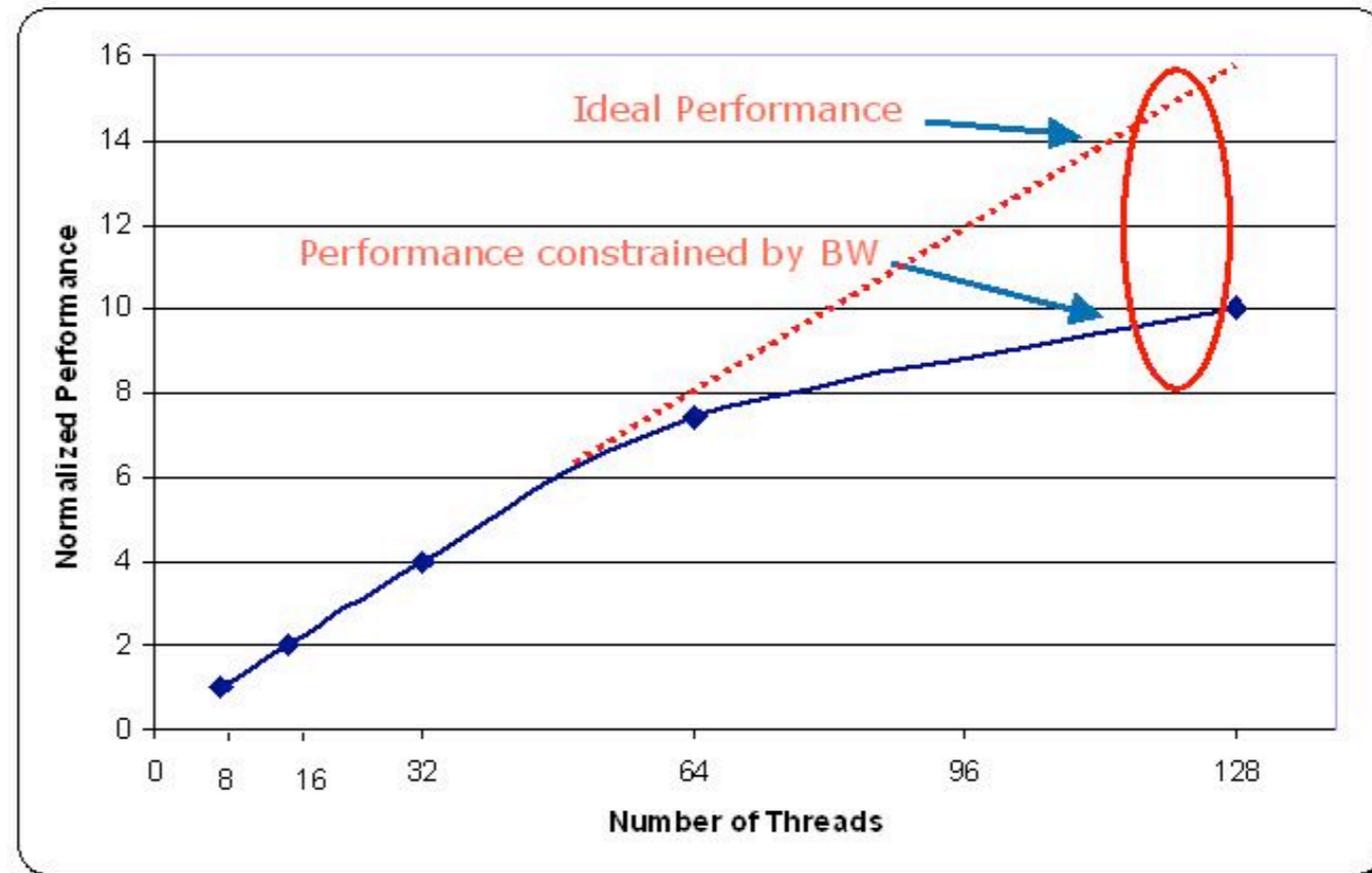


(the capacity problem, Brinda Ganesh's thesis)

Per-core capacity *decr.* as # cores/CPU *incr.*

Some Trends

Required BW per core (~ 1 GB/s):



- Thread-based load (SPECjbb), memory set to 52GB/s sustained
- Saturates around 64 cores/threads (~ 1 GB/s per core)
- cf. 32-core Sun Niagara: saturates at 25.6 GB/s

Some Trends

Commodity Systems:

- Low double-digit GB per CPU socket
- \$10–100 per DIMM

High End:

- Higher (but still not *high*)
double-digit GB per CPU socket
- ~ \$1000 per DIMM

Fully-Buffered DIMM:

- (largely failed) attempt to bridge the gap ...

Some Trends: FB-DIMM



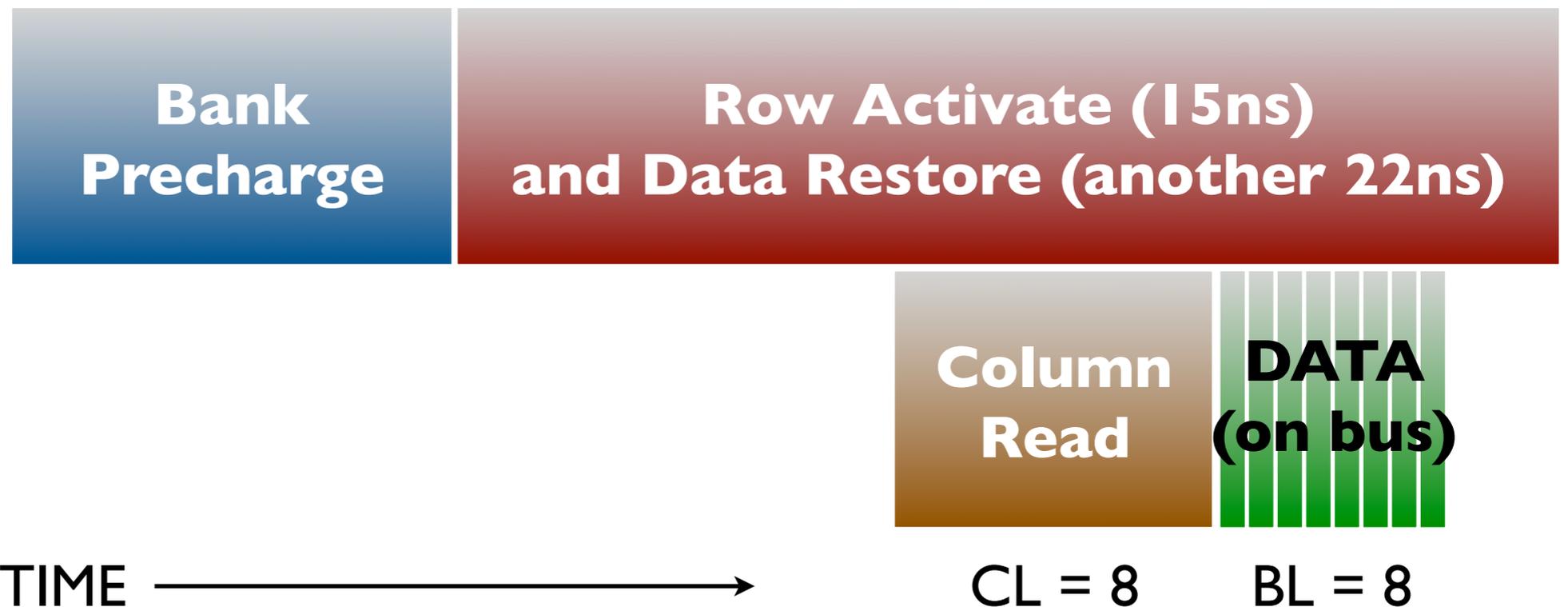
- JEDEC DDRx: ~1W/DIMM, ~10W total
- FB-DIMM: 5–10W/DIMM, ~350W total

Some Perspective

Cost of access is high; requires significant effort to amortize this over the (increasingly short) payoff.

$t_{RP} = 15\text{ns}$

$t_{RCD} = 15\text{ns}$, $t_{RAS} = 37.5\text{ns}$



Bottom Line

DRAM is performance limiter: HPC apps
will always exceed your cache
(if you could run it on a laptop, you would)

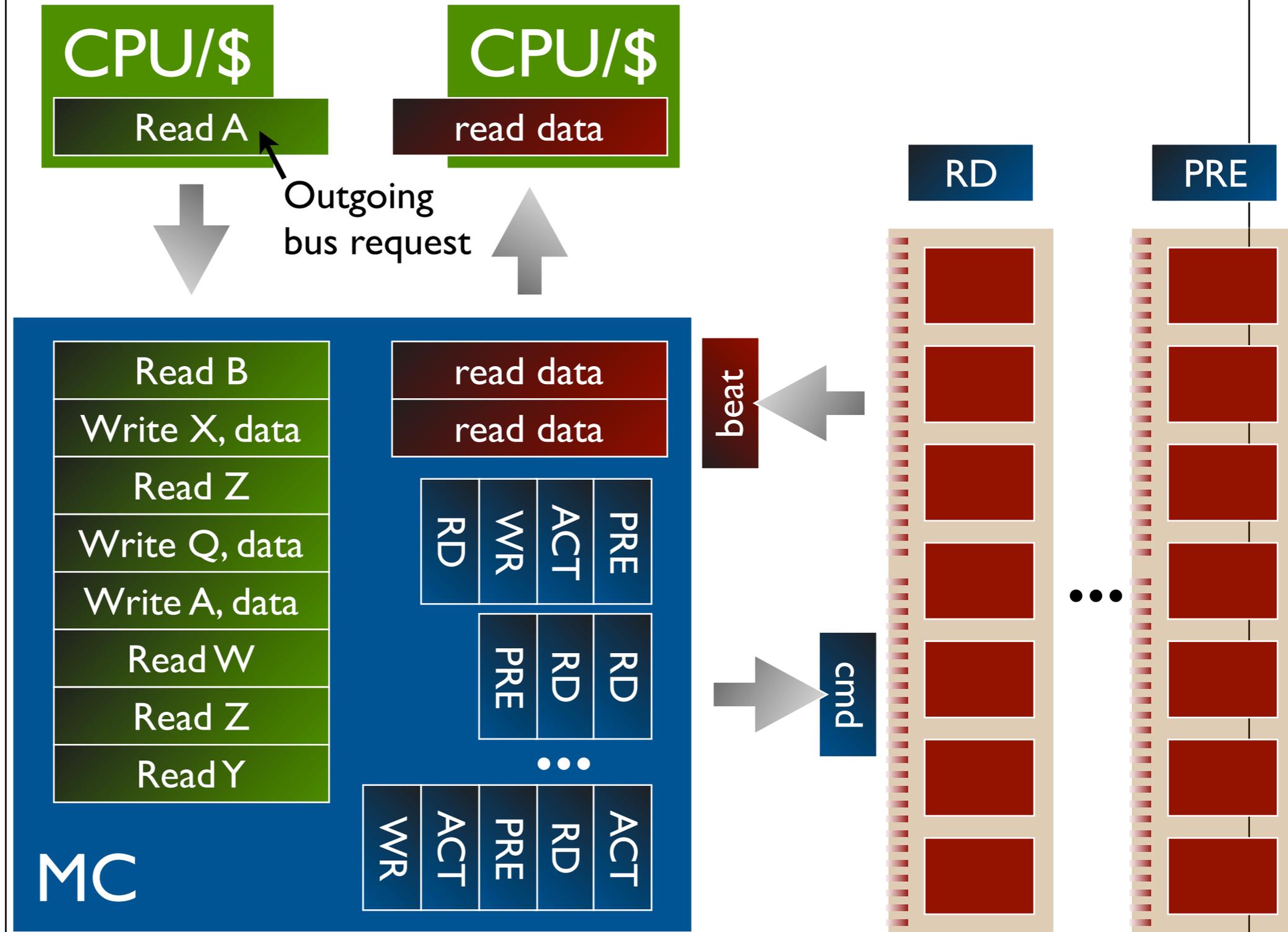
Memory system determines, to large extent:

- Your performance
- Your power dissipation
- Your system cost

PROBLEM:

- Nobody models it accurately

What It Looks Like



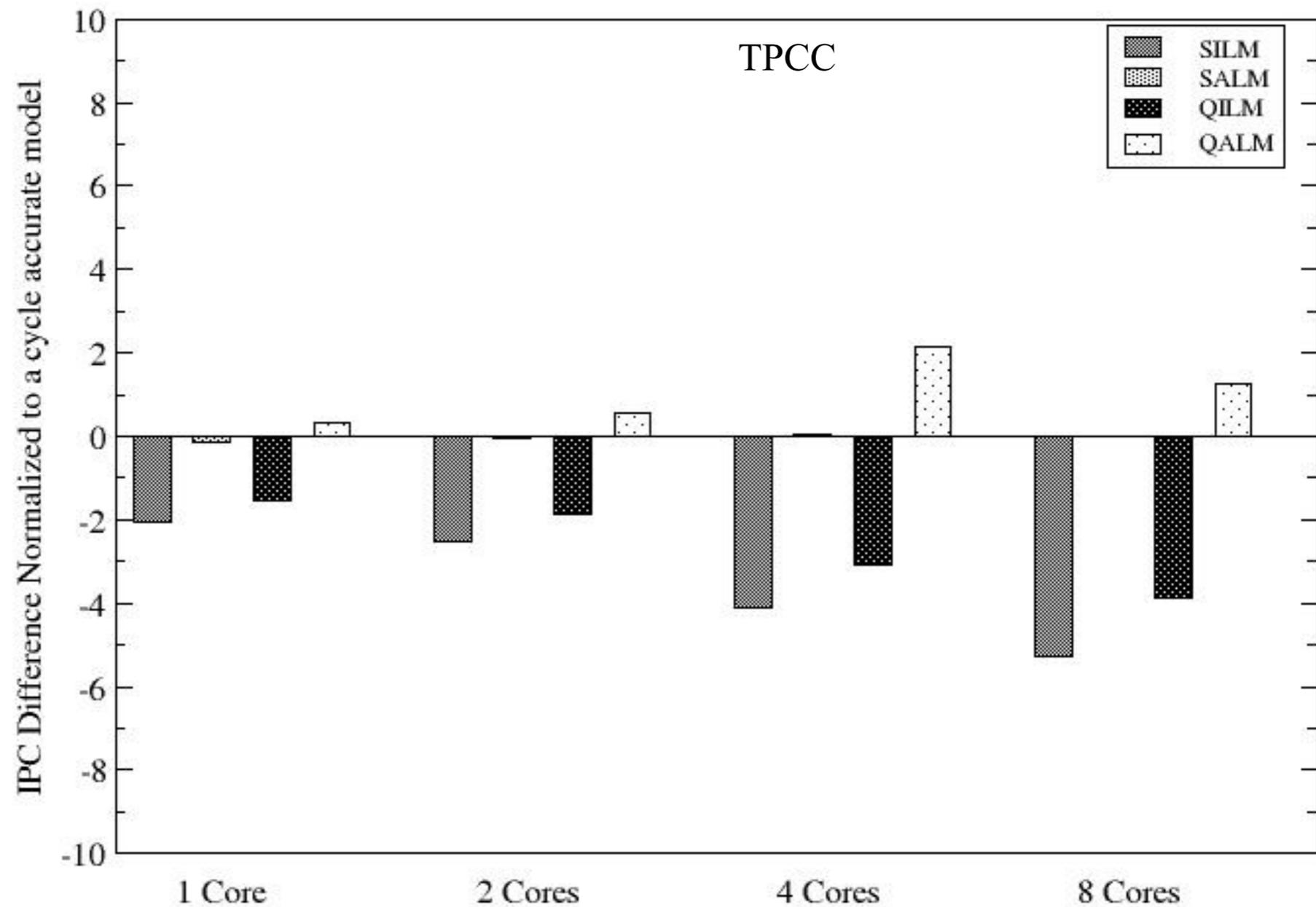
How It Is Represented

```
if (cache_miss(addr)) {  
    cycle_count += DRAM_LATENCY;  
}
```

... even in simulators with “cycle accurate”
memory systems—no lie

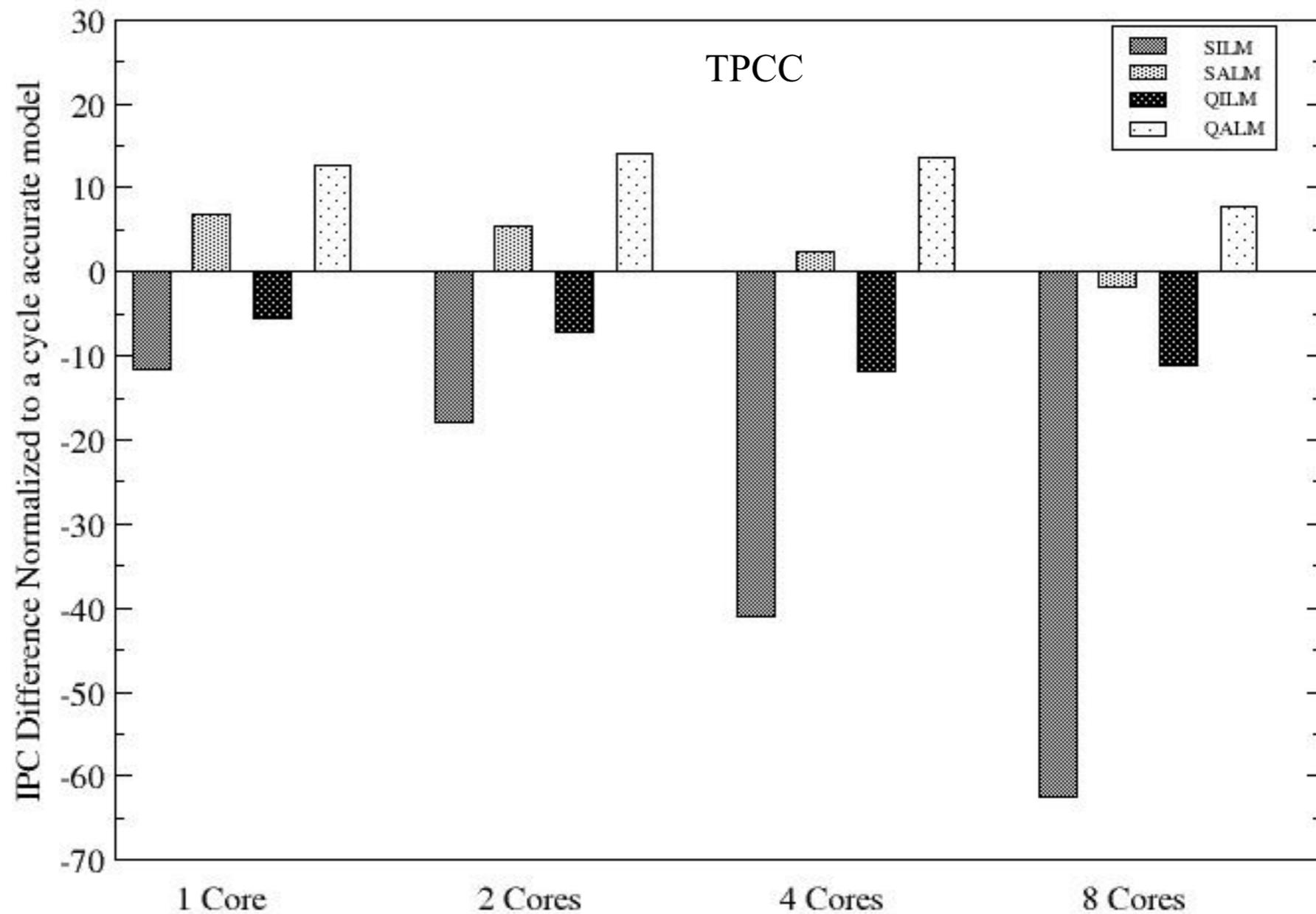
Some Cases In Point

Prefetching, Multicore, etc.



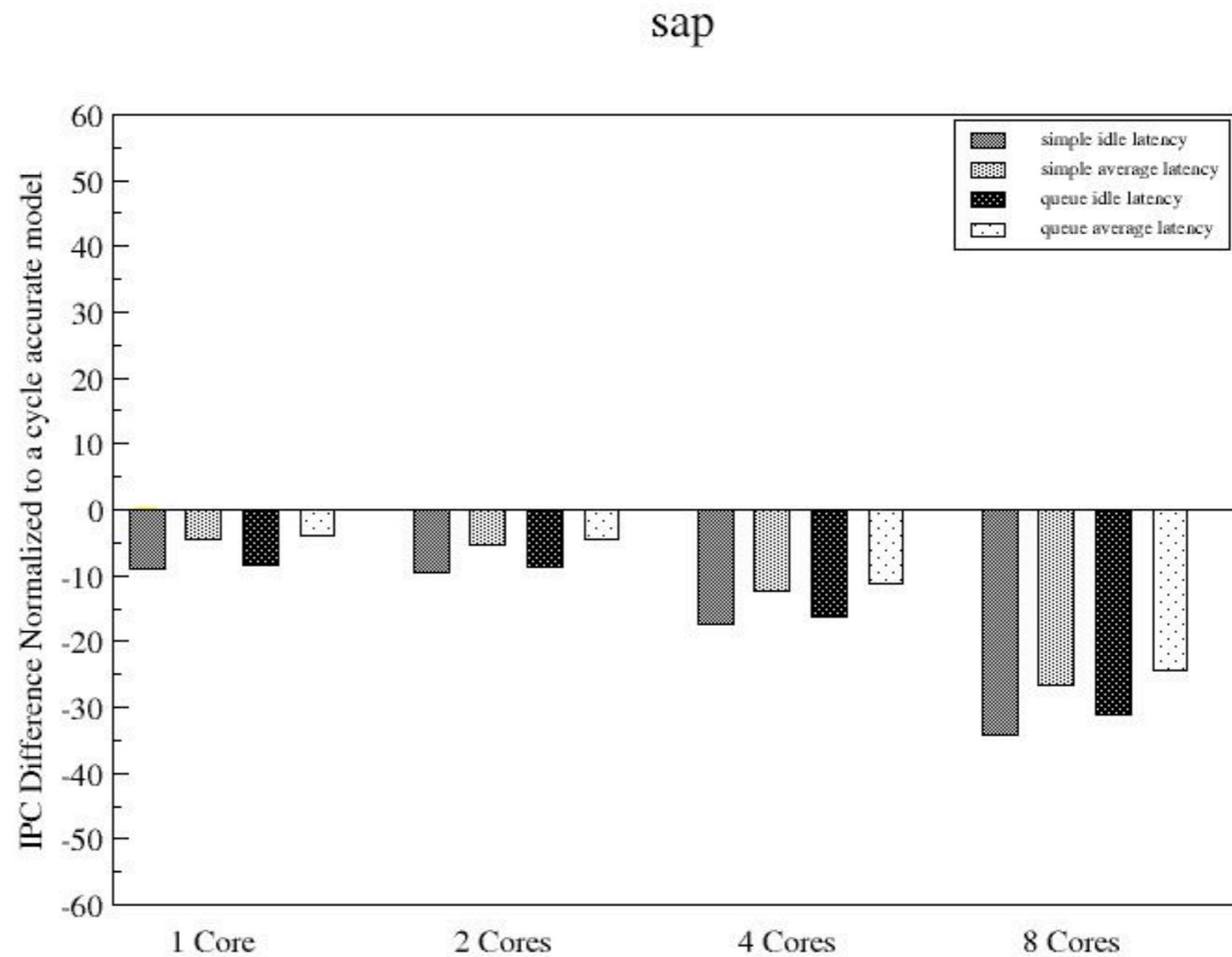
Some Cases In Point

Prefetching, Multicore, etc.



Some Cases In Point

Prefetching, Multicore, etc.



Your Choice

Today, DRAM is primary limiter. If you really want to predict system behavior

- performance
- power
- cost

you have to model the memory system,
accurately.

... so what do you get if you do?

MEMORY SYSTEMS,
ETC.

Bruce Jacob

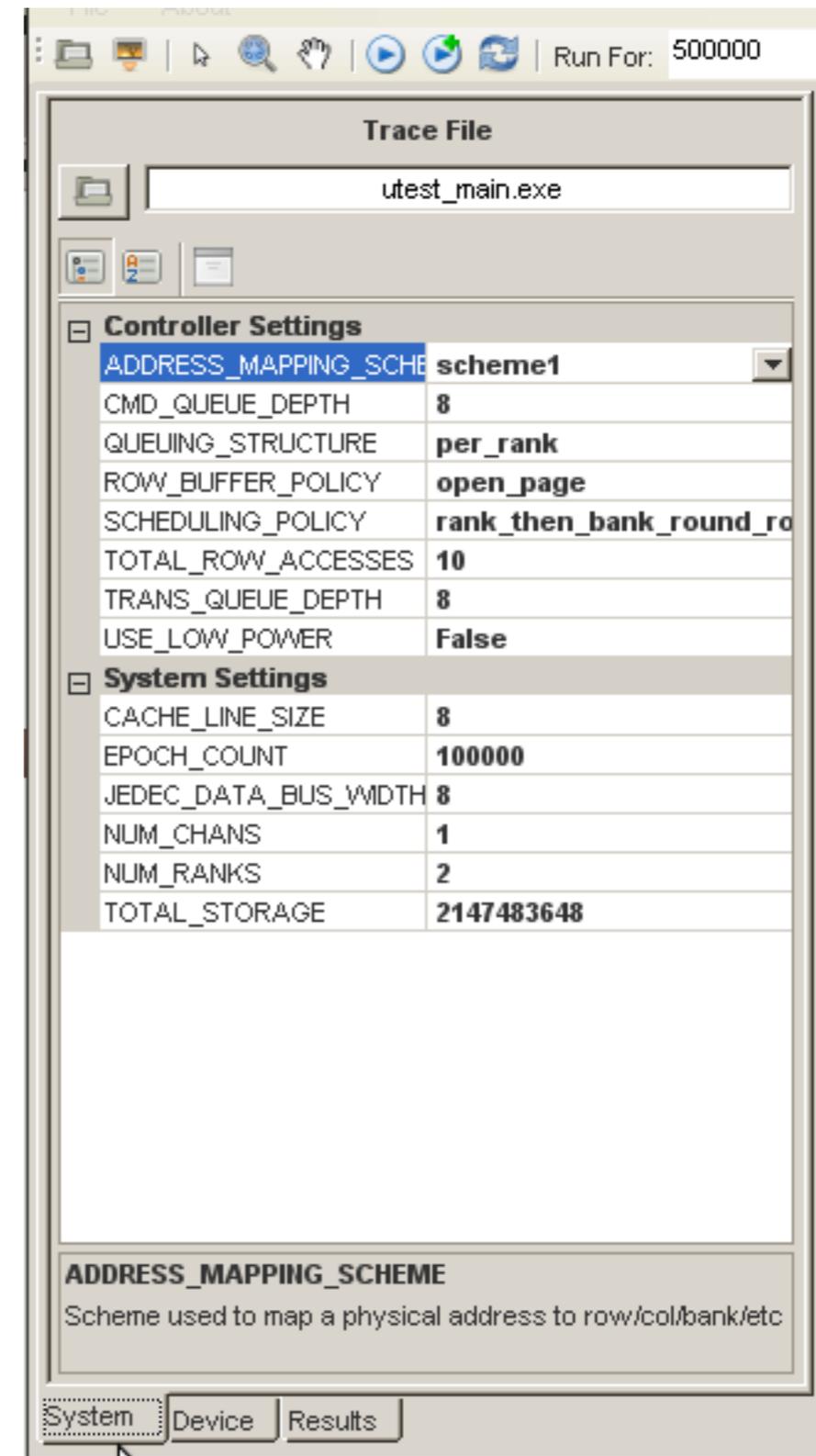
DRAMsim

Latest version
via Sandia & NSA



DRAMsim

System Parameters



The screenshot shows the DRAMsim software interface. At the top, the window title is "Trace File" and the file name is "utest_main.exe". The "Run For:" field is set to "500000". The interface is divided into two main sections: "Controller Settings" and "System Settings".

Controller Settings	
ADDRESS_MAPPING_SCHE	scheme1
CMD_QUEUE_DEPTH	8
QUEUING_STRUCTURE	per_rank
ROW_BUFFER_POLICY	open_page
SCHEDULING_POLICY	rank_then_bank_round_ro
TOTAL_ROW_ACCESSES	10
TRANS_QUEUE_DEPTH	8
USE_LOW_POWER	False

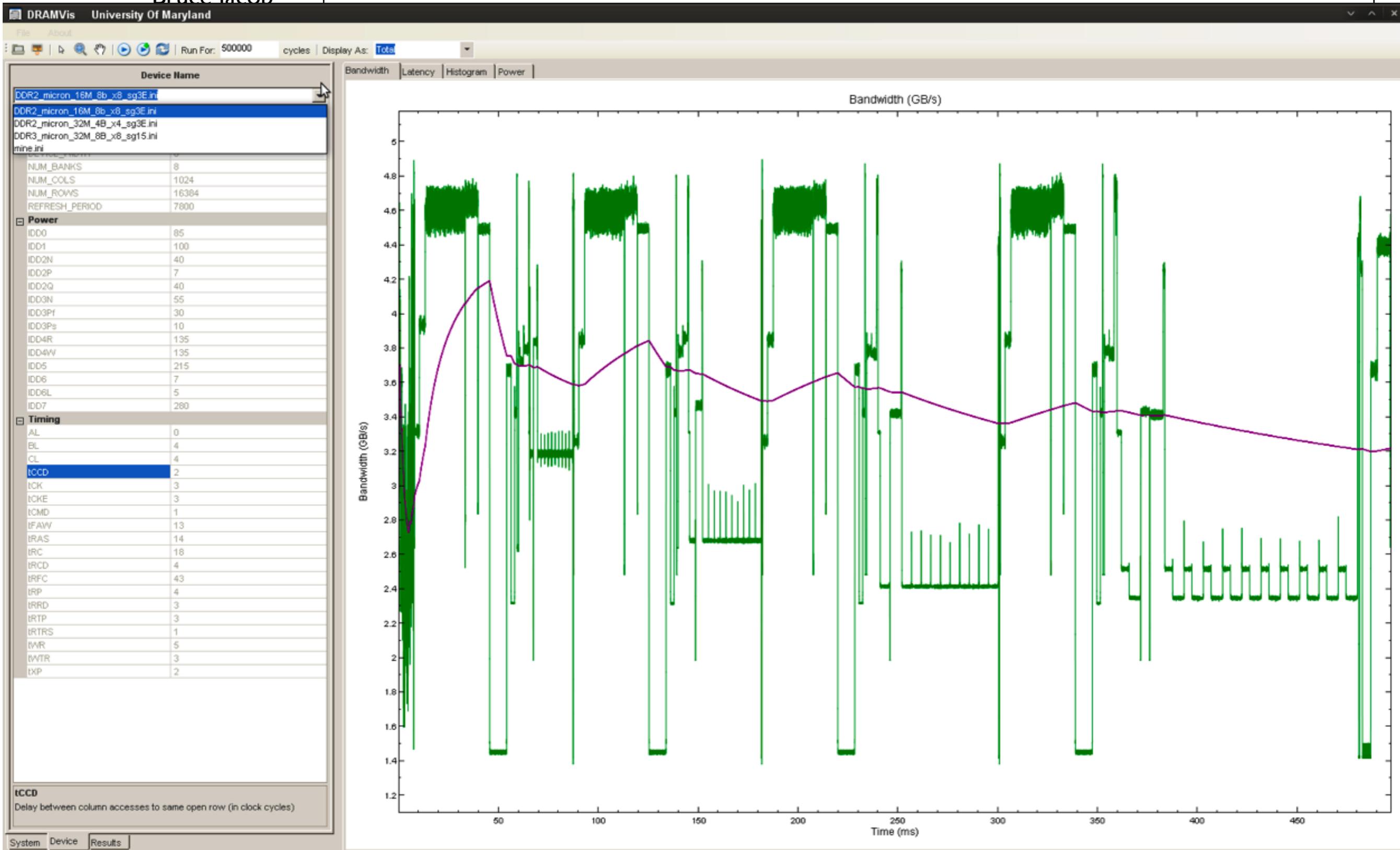
System Settings	
CACHE_LINE_SIZE	8
EPOCH_COUNT	100000
JEDEC_DATA_BUS_WIDTH	8
NUM_CHANS	1
NUM_RANKS	2
TOTAL_STORAGE	2147483648

At the bottom of the interface, there are three tabs: "System", "Device", and "Results". The "System" tab is currently selected. Below the settings, there is a description for the "ADDRESS_MAPPING_SCHEME": "Scheme used to map a physical address to row/col/bank/etc".

MEMORY SYSTEMS,
ETC.

Bruce Jacob

DRAMsim



DRAMsim

Device Parameters

Run For: 500000

Device Name: DDR2 micron 16M 8b x8 sq3E.ini

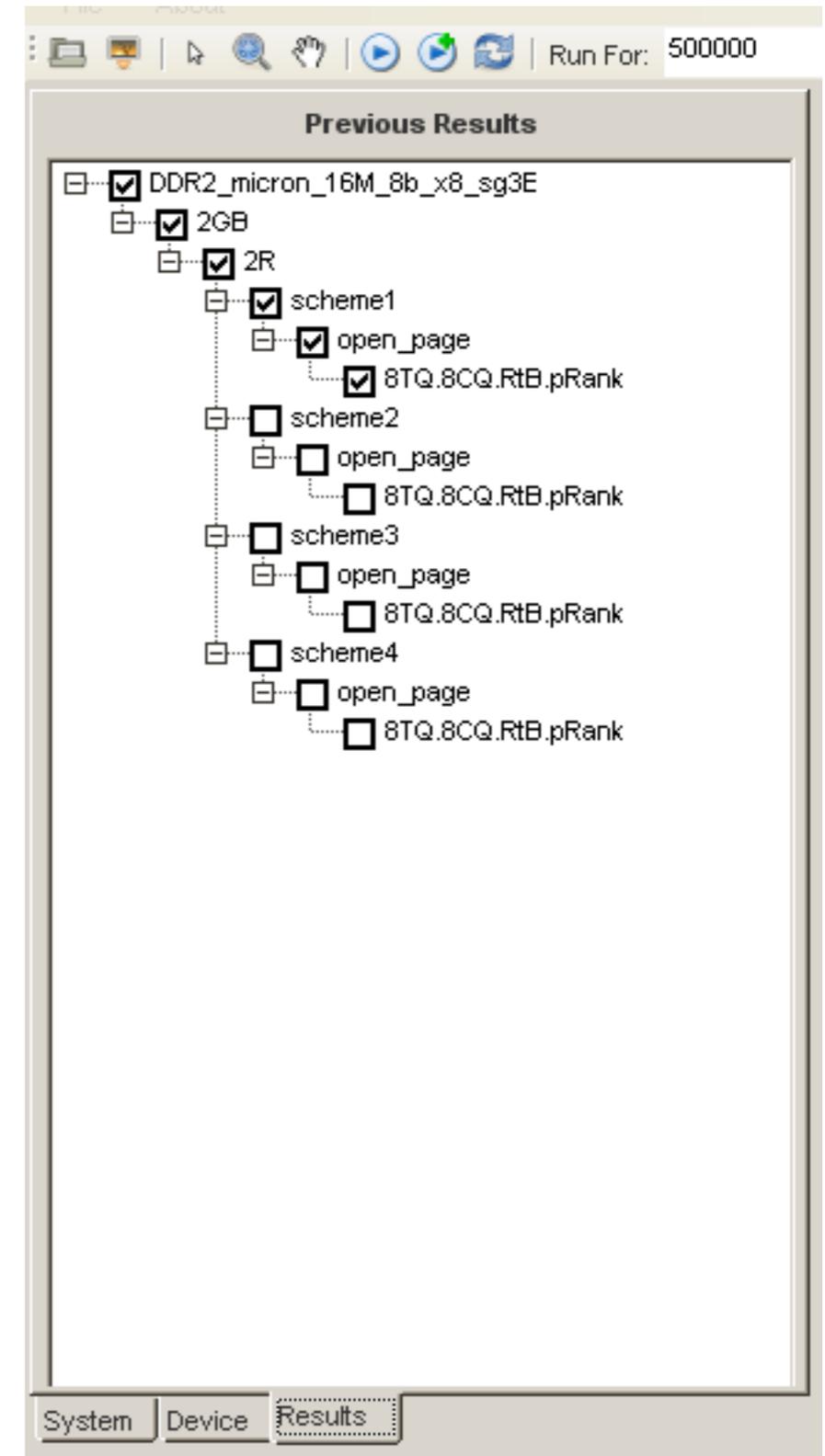
Category	Parameter	Value
Physical	DEVICE_WIDTH	8
	NUM_BANKS	8
	NUM_COLS	1024
	NUM_ROWS	16384
	REFRESH_PERIOD	7800
Power	IDD0	85
	IDD1	100
	IDD2N	40
	IDD2P	7
	IDD2Q	40
	IDD3N	55
	IDD3Pf	30
	IDD3Ps	10
	IDD4R	135
	IDD4W	135
IDD5	215	
IDD6	7	
IDD6L	5	
IDD7	280	
Timing	AL	0
	BL	4

IDD3Pf
All-banks open in power-down mode - fast (mA)

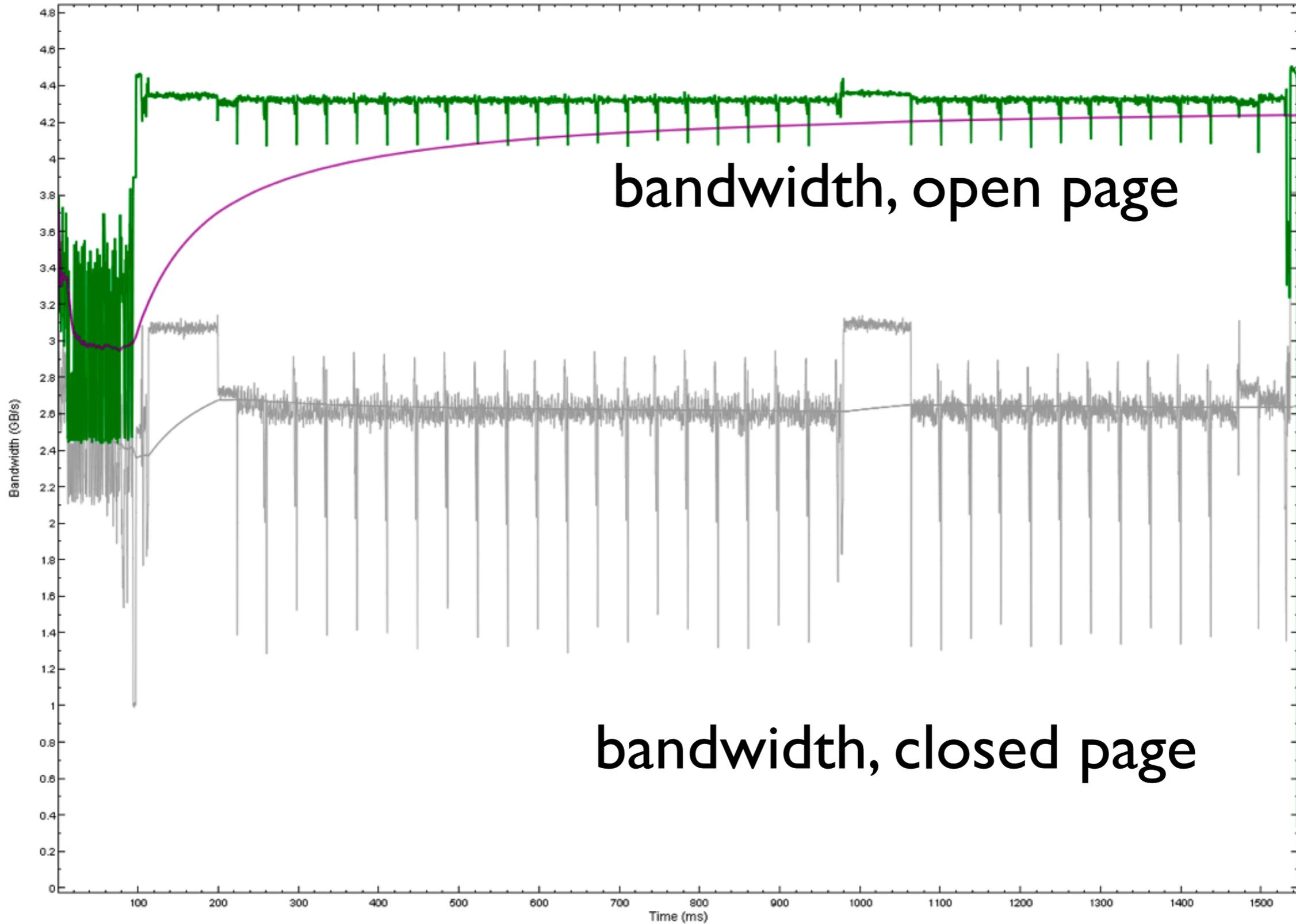
System Device Results

DRAMsim

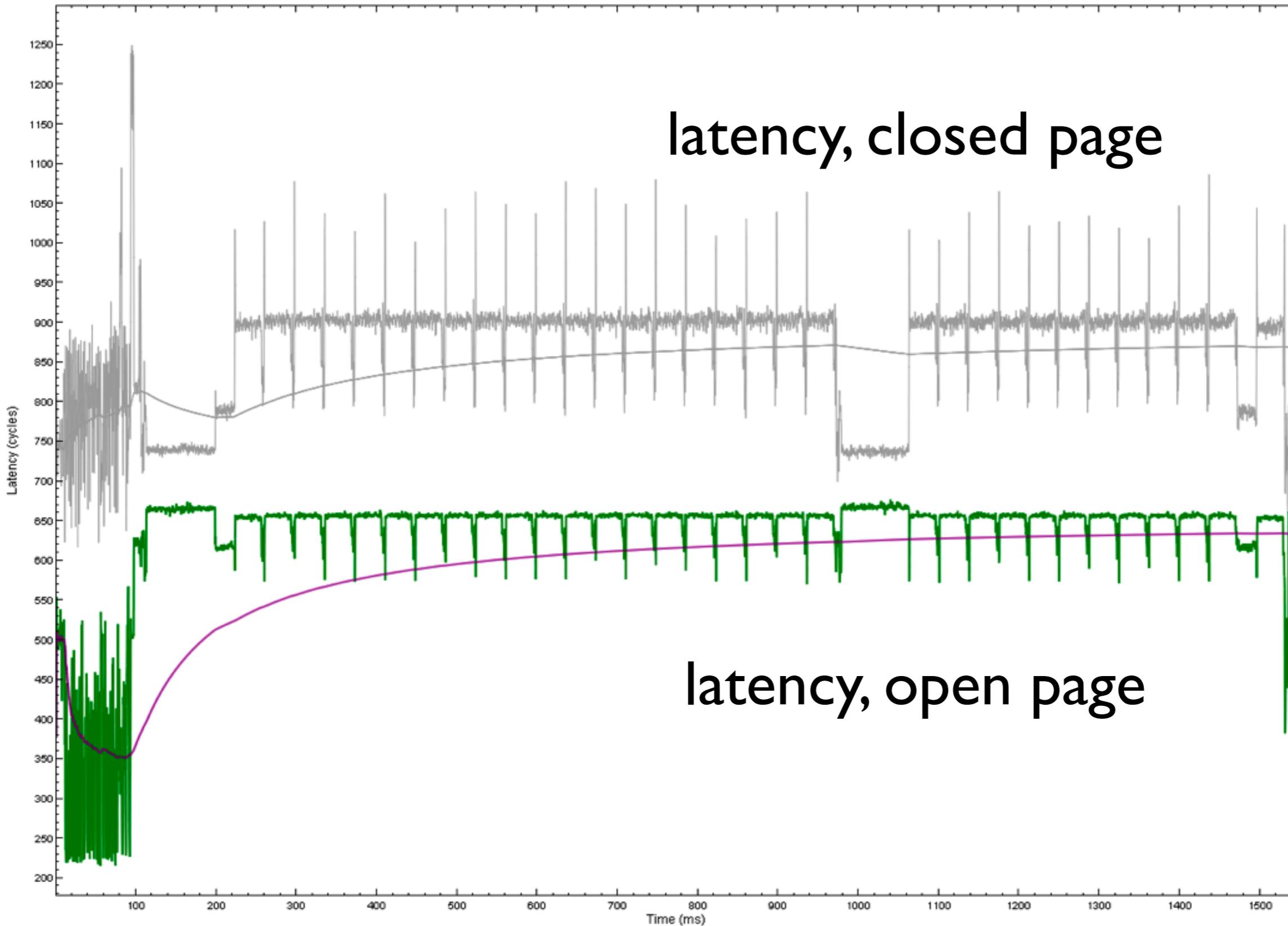
Overlay Graphs from Multiple Runs



DRAMsim Results: Epetra



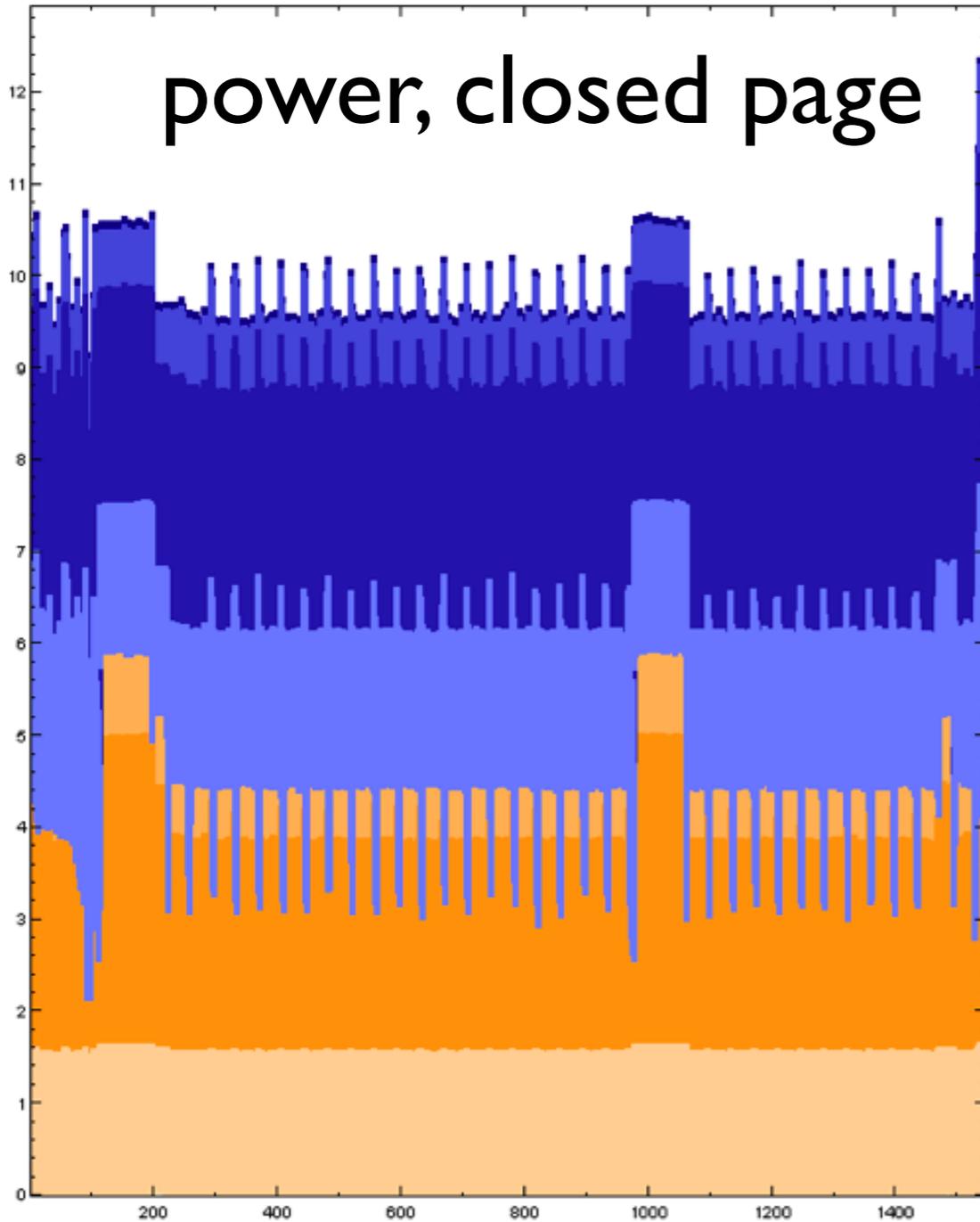
DRAMsim Results: Epetra



DRAMsim Results: Epetra

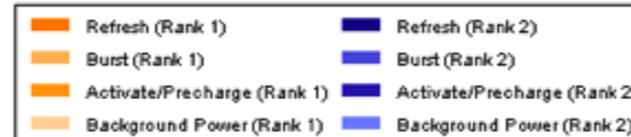
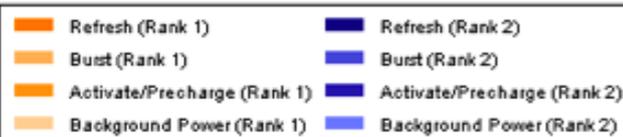
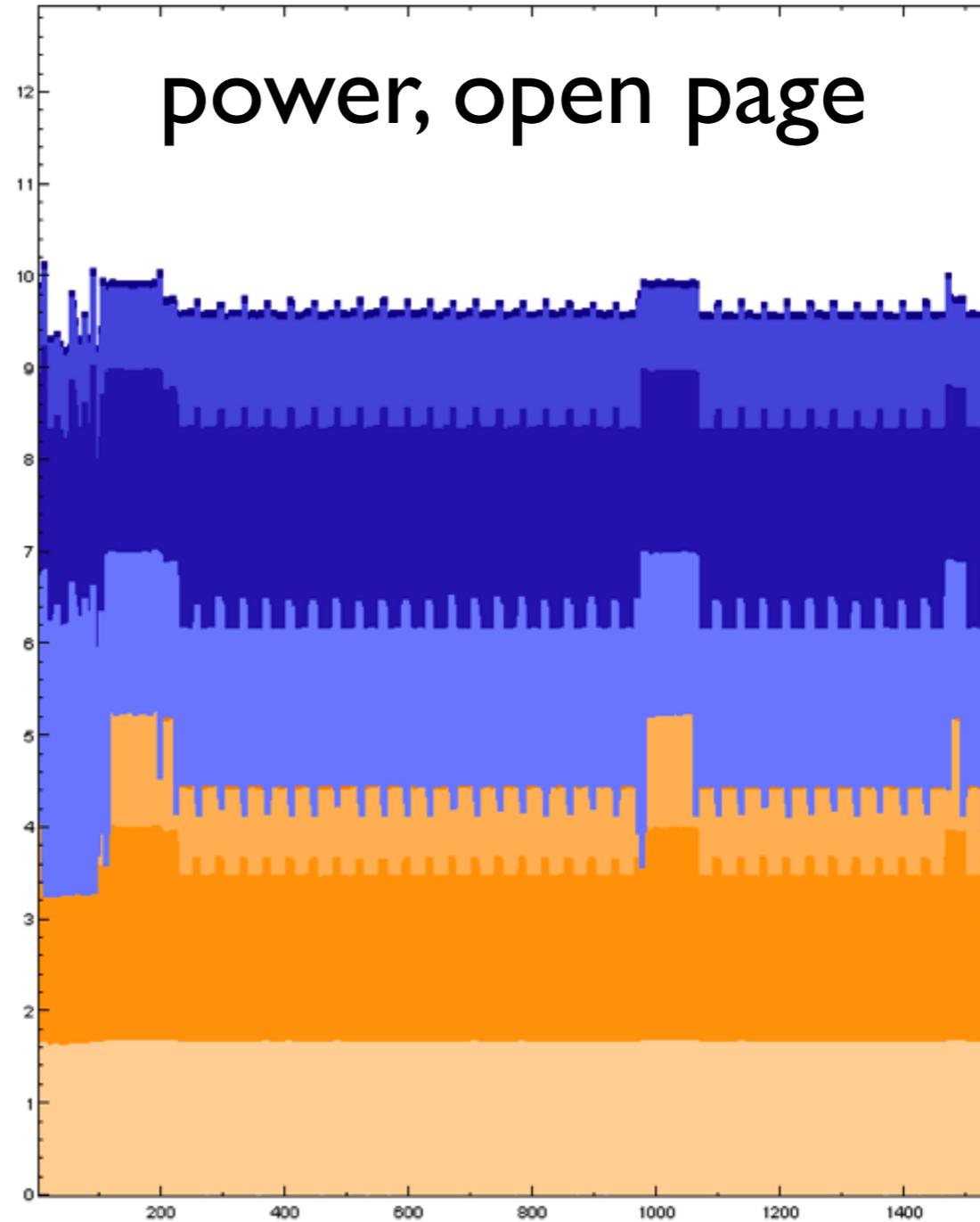
DDR2_micron_16M_8b_x8_sg3E.ini
2GB.1Ch.2R.scheme2.close_page.8TQ.8CQ.RtB.pRank.vis

power, closed page



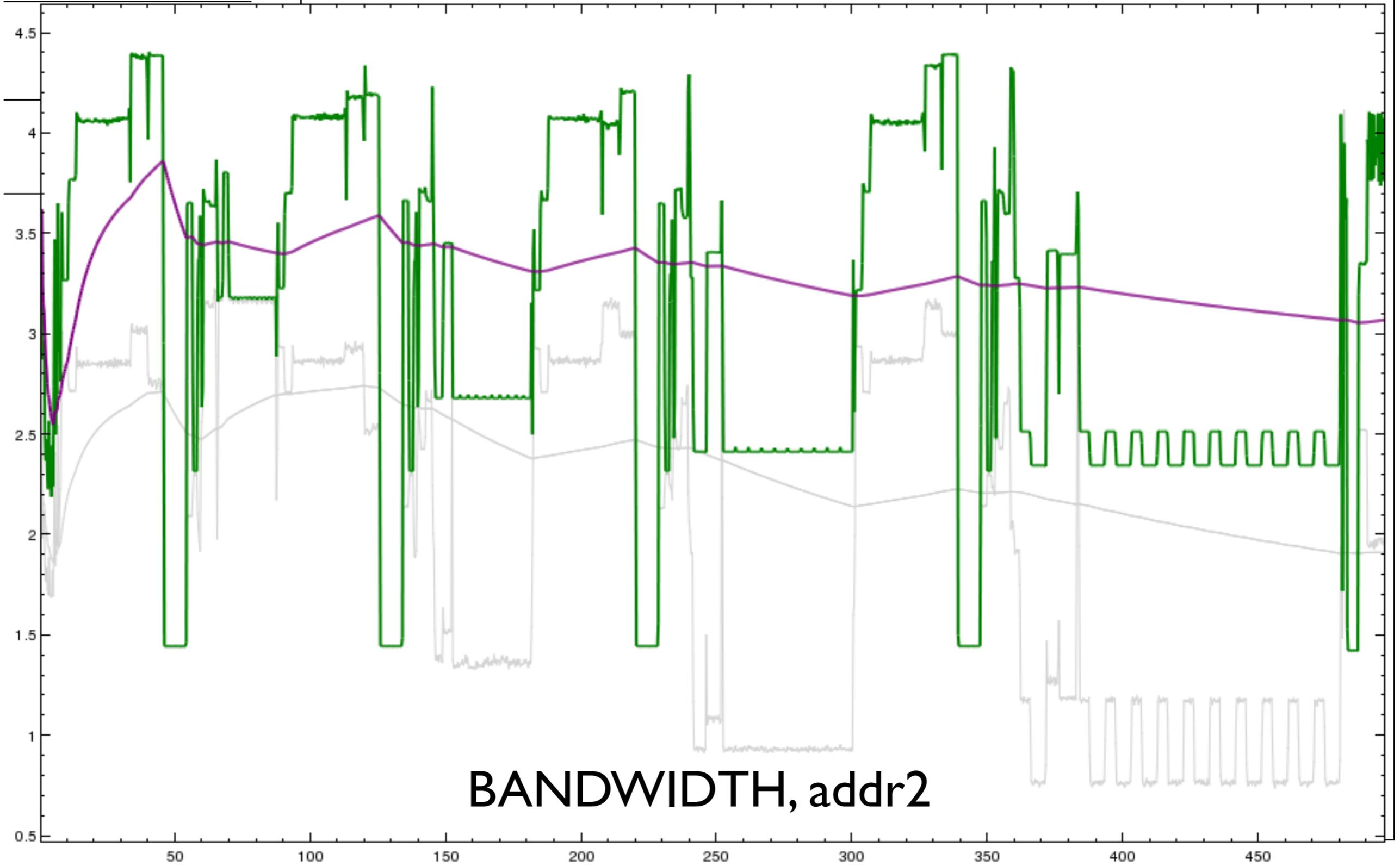
DDR2_micron_16M_8b_x8_sg3E.ini
2GB.1Ch.2R.scheme2.open_page.8TQ.8CQ.RtB.pRank.vis

power, open page



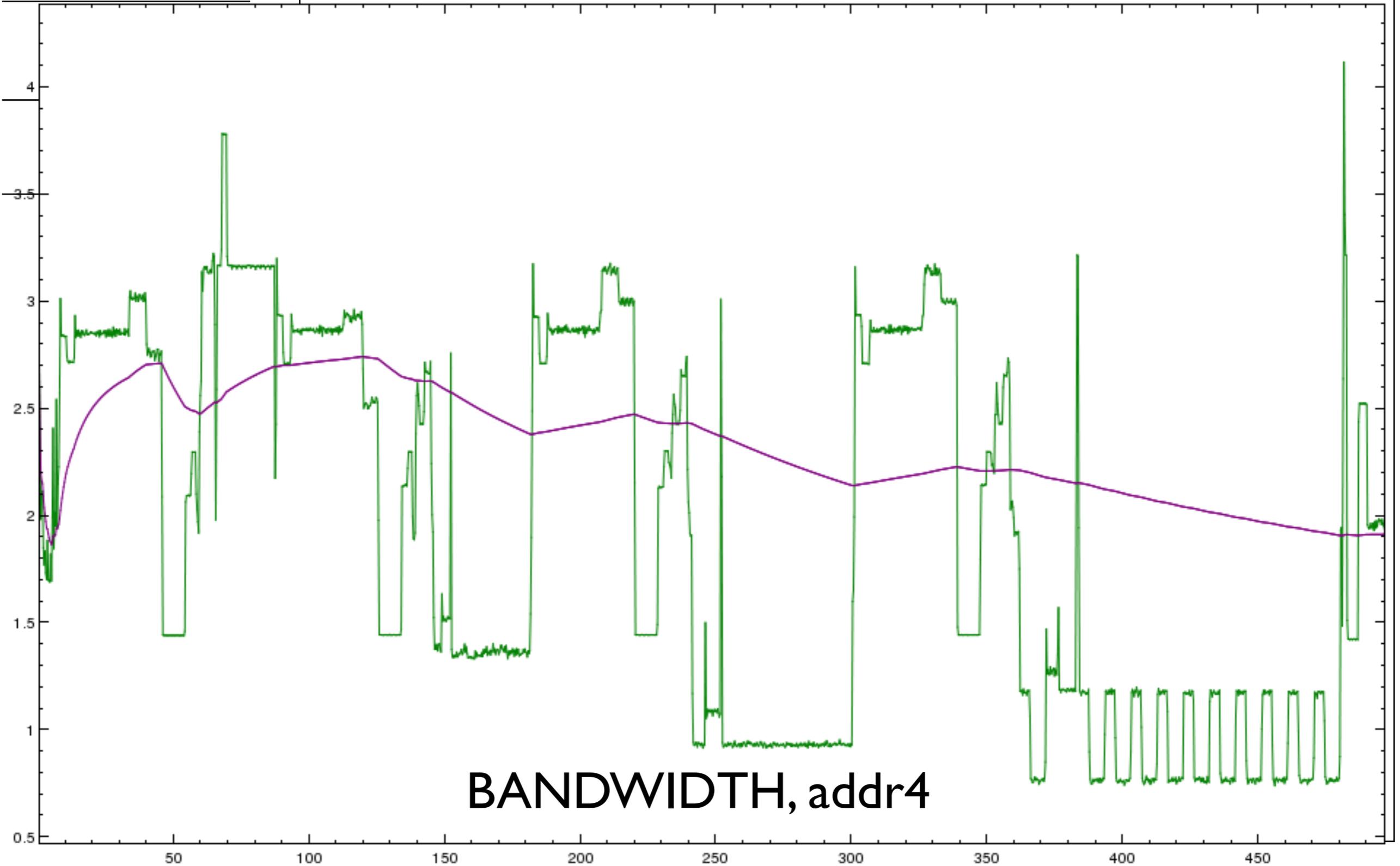
DRAMsim Results: Epetra

Bandwidth (GB/s)



DRAMsim Results: Epetra

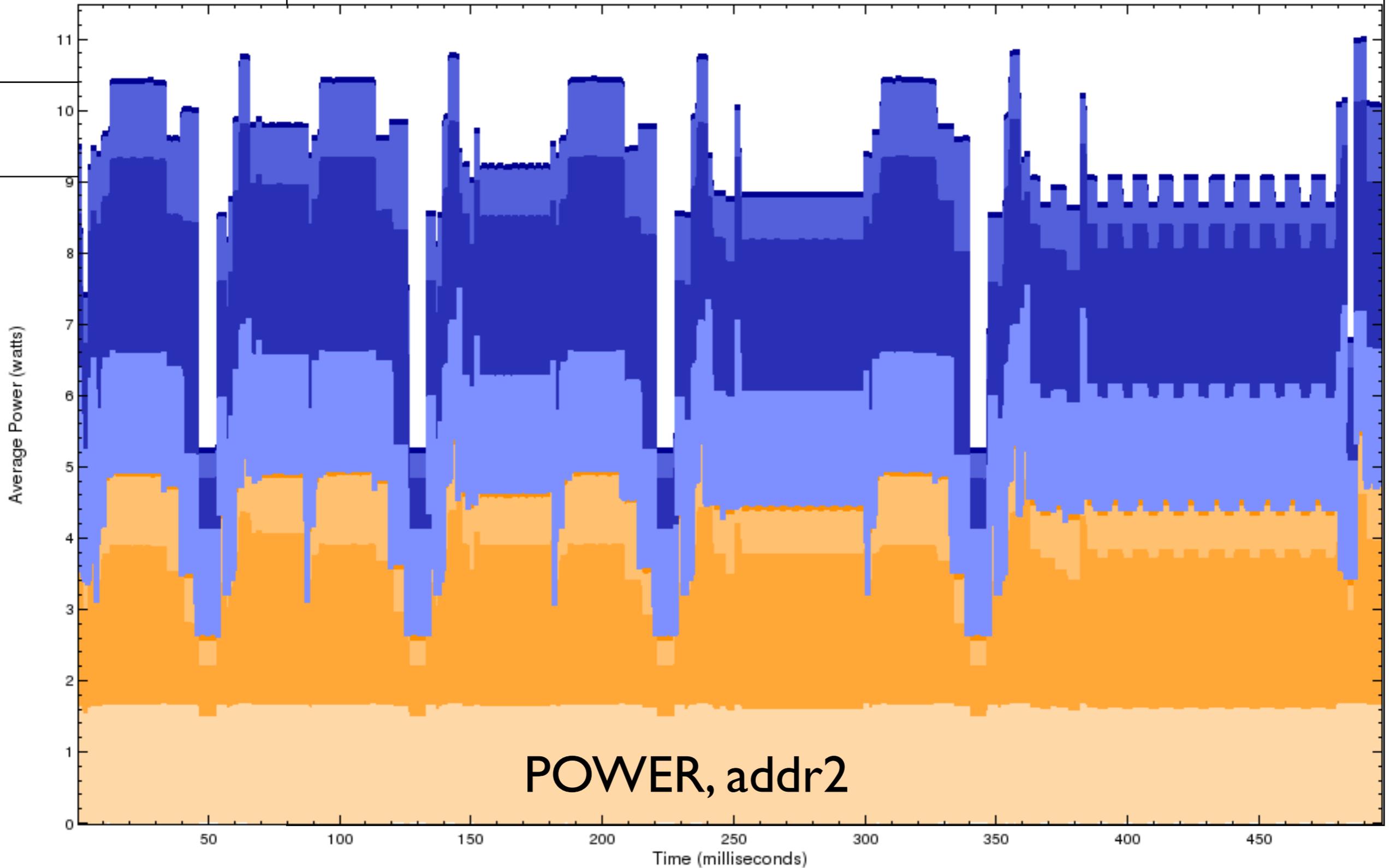
Bandwidth (GB/s)



BANDWIDTH, addr4

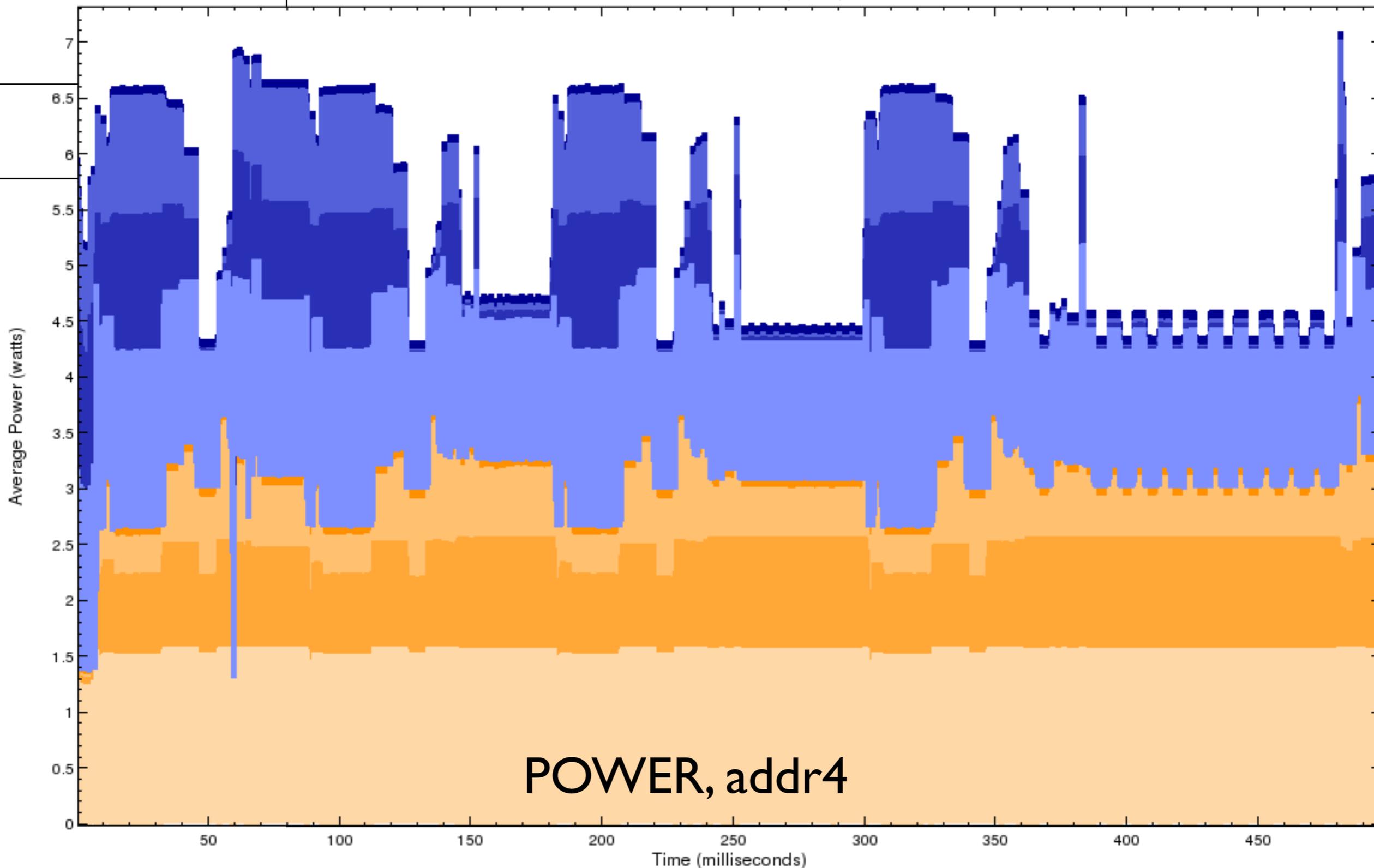
DRAMsim Results: Epetra

Power Dissipation (watts)

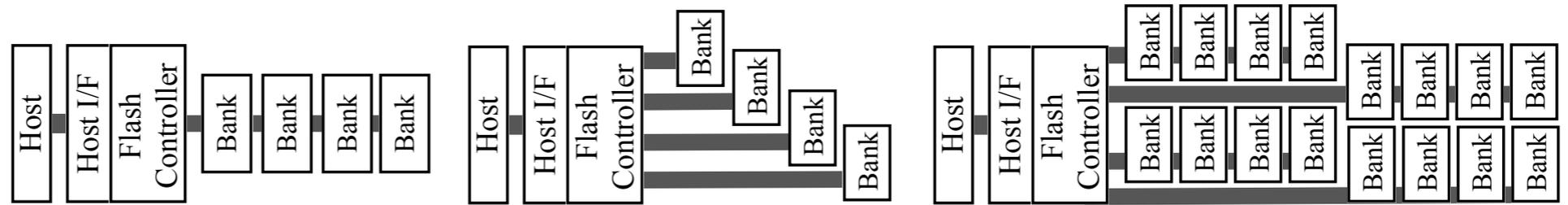


DRAMsim Results: Epetra

Power Dissipation (watts)



DRAMsim Results: Flash

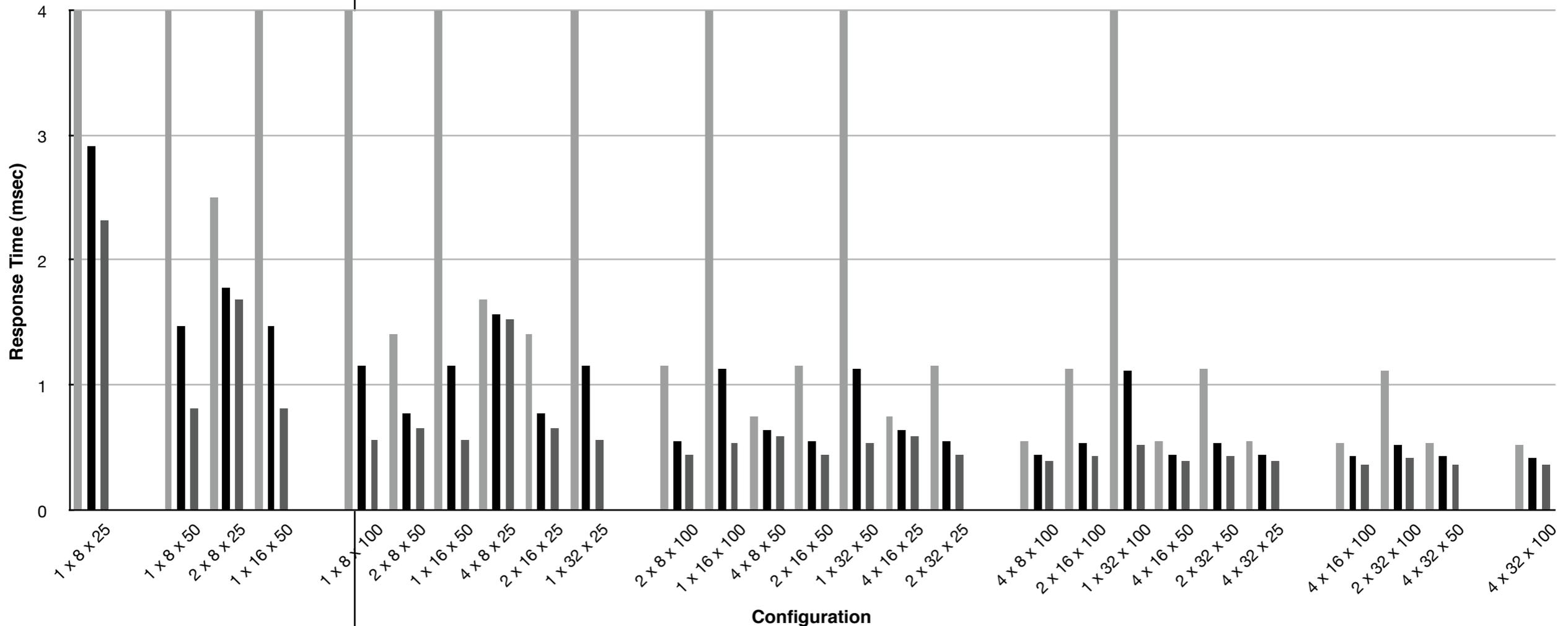


(a) Single channel

(b) Dedicated channel for each bank

(c) Multiple shared channels

Read Response Time vs. Organization



... and tons more

Address-mapping policies (2–10x)

Scheduling policies (2–10x)

(e.g., improved Cray MC by 20% sustained BW)

Paging policies (2x)

System organization (10+x)

Blocking, cache interactions, OS interactions
(e.g. virtual memory), file system interactions,
queueing mechanisms, device architectures, ...

If you want to reduce power, this is where to look

If you want to increase performance, this is where to look

MEMORY SYSTEMS,
ETC.

Bruce Jacob

University of
Maryland

SOS-13
March 2009

SLIDE 30

Shameless Plug

