Cloud Downtime Due to Memory Errors

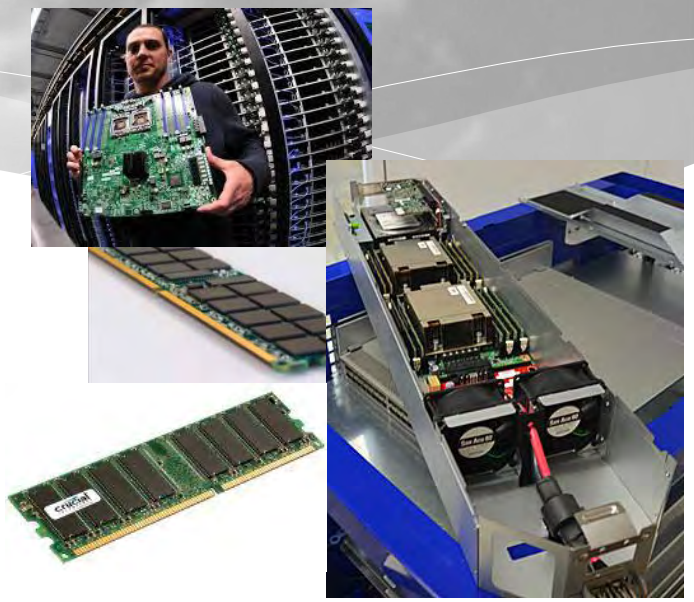
A bigger problem than you might think!

Barbara P Aichinger
Vice President New Business Development
FuturePlus Systems Corporation

Santa Clara, CA
November 2012

Cloud Hardware

- Servers in the Data Center
- Loaded with memory
  - Several DDR3 and soon DDR4 channels per processor
  - Several DIMMs per channel
  - DIMMs becoming denser
- The industry has NO standardized compliance tesing
- Price pressures affecting validation budgets

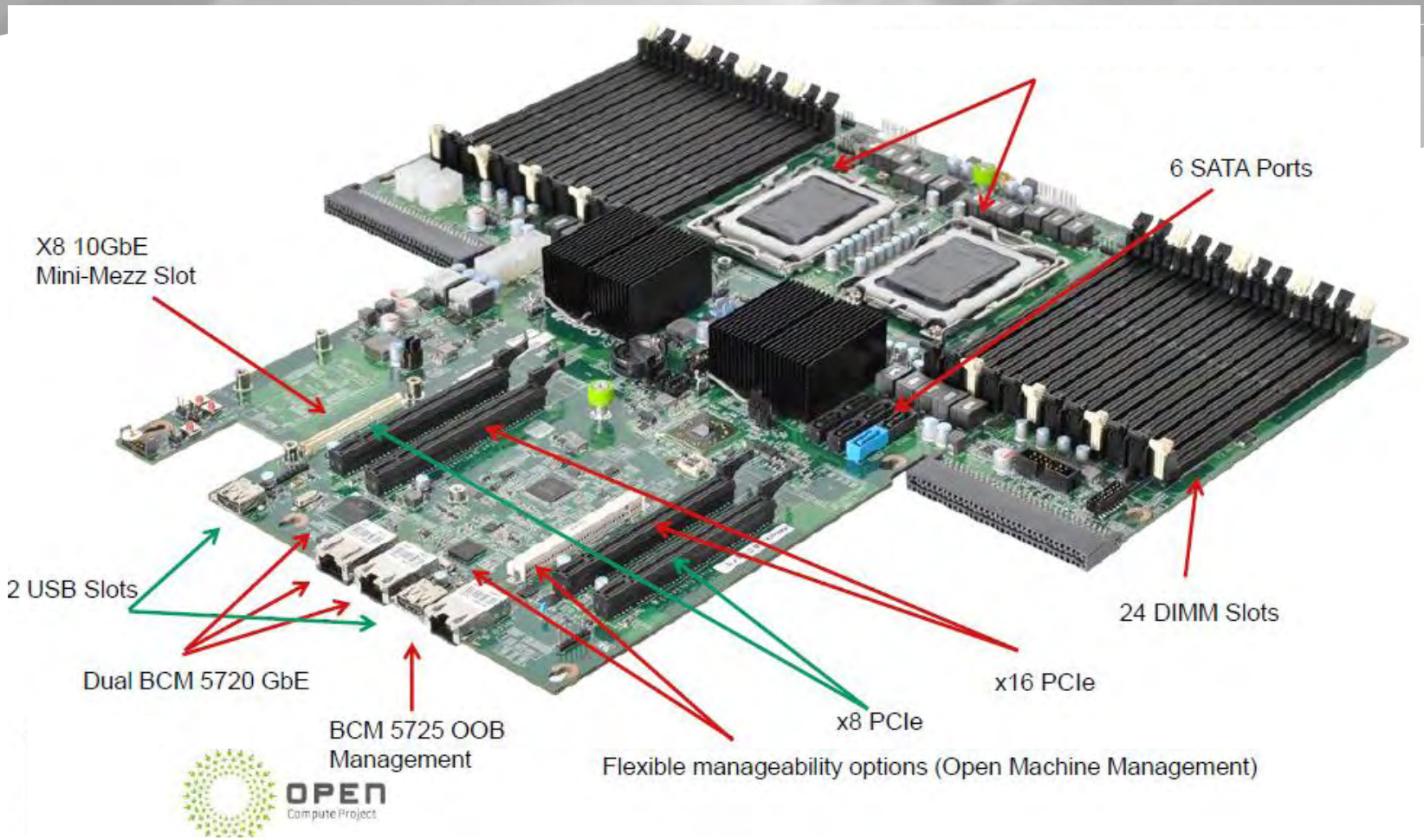


DDR3

JEDEC DDR4

Cloud Hardware

24 total DIMM slots



Studies Show: More Errors than Expected

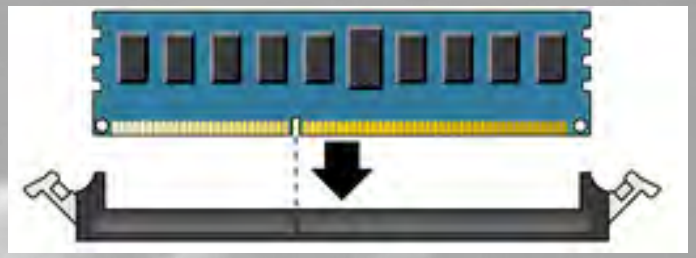
- DRAM Errors in the Wild: A Large-Scale Field Study
- Cosmic Rays Don't Strike Twice: Understanding the Nature of DRAM Errors and the Implications for System Design
- Hard Data on Soft Errors: A Large-Scale Assessment of Real-World Error Rates in GPGPU
- An Empirical Study of Memory Hardware Errors in a Server Farm
- A Realistic Evaluation of Memory Hardware Errors and Software System Susceptibility

The Result?

- Band-Aids and work arounds
  - Chip Kill
  - ECC
  - Page Retirement
  - Mirroring
  - Duplication

You are just rearranging the chairs on the deck of the Titanic….the ship is still going down

Memory Standards



- JEDEC www.JEDEC.com is the international standards body that governs DRAM Chip Specifications
- JEDEC also specifies DIMM mechanical form factors and electrical properties of the DIMM etch
- Standardizes connector pinout
- Standardizes speeds and properties that promote interoperability
- However there is NO standardized Compliance Testing

How big of a problem is this?

- Studies show: 2-4% error rate
- **If** Google has a million servers**
- A 4% error rate per year is 40,000 failures a year
- 3,334 failures per month
- 110 failures per day
- 4.6 failures per hour

**** Aug 2011 DataKnowledgeCenter.com reported 900,000**

How is it handled?
DIMM swaps and system replacement

Source: Google

What are the causes of these memory errors

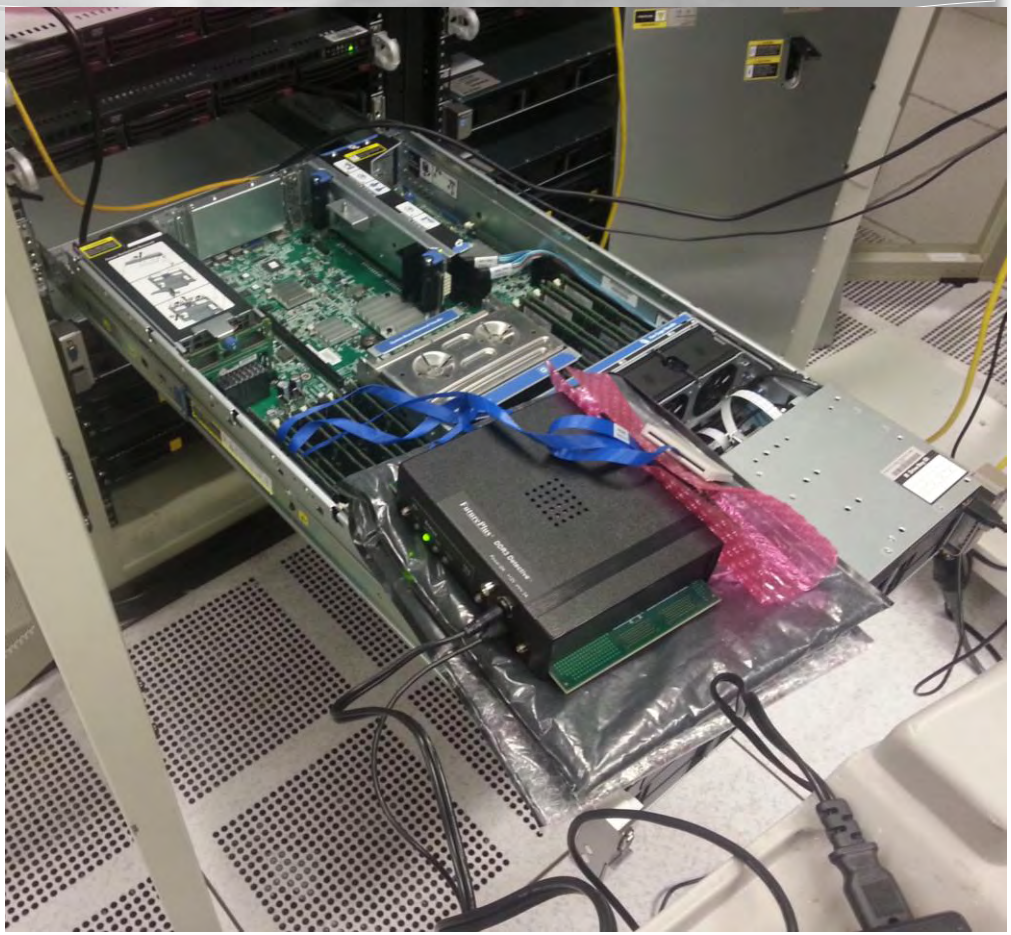
Other than the Cosmic Rays, Aging and Environmental Issues

- Motherboard Signal Integrity
- Poor quality Connectors
- Incompatible BIOS Settings
- Protocol Violations by the Memory Controller
- Poor Quality DIMMs

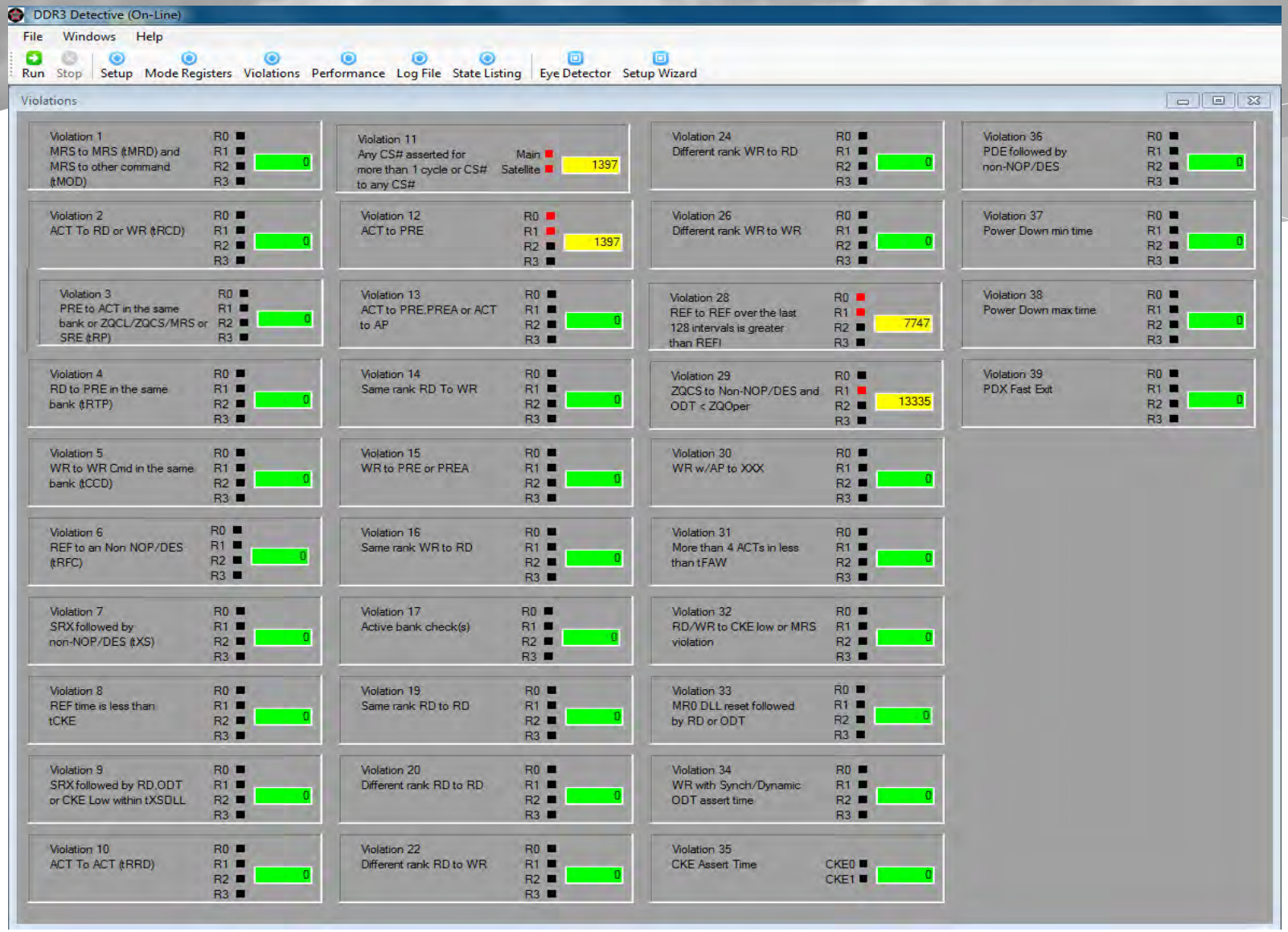
Design Flaws!

Case Study

- Server in the Data Center
- Equipment
  - DDR3 Detective®
  - Google StressAPP test



Failures


DDR3 Detective (On-Line)
File
Windows
Help
Run
Stop
Setup
Mode Registers
Violations
Performance
Log File
State Listing
Eye Detector
Setup Wizard
Violations
Violation 1
MRS to MRS (tMRD) and MRS to other command (tMOD)
R0 R1 R2 R3
0
Violation 11
Any CS# asserted for more than 1 cycle or CS# to any CS#
Main
Satellite
1397
Violation 24
Different rank WR to RD
R0 R1 R2 R3
0
Violation 36
PDE followed by non-NOP/DES
R0 R1 R2 R3
0
Violation 2
ACT To RD or WR (tRCD)
R0 R1 R2 R3
0
Violation 12
ACT to PRE
R0 R1 R2 R3
1397
Violation 26
Different rank WR to WR
R0 R1 R2 R3
0
Violation 37
Power Down min time
R0 R1 R2 R3
0
Violation 3
PRE to ACT in the same bank or ZQCL/ZQCS/MRS or SRE (tRP)
R0 R1 R2 R3
0
Violation 13
ACT to PRE,PREA or ACT to AP
R0 R1 R2 R3
0
Violation 28
REF to REF over the last 128 intervals is greater than REFI
R0 R1 R2 R3
7747
Violation 38
Power Down max time
R0 R1 R2 R3
0
Violation 4
RD to PRE in the same bank (tRTP)
R0 R1 R2 R3
0
Violation 14
Same rank RD To WR
R0 R1 R2 R3
0
Violation 29
ZQCS to Non-NOP/DES and ODT < ZQOper
R0 R1 R2 R3
13335
Violation 39
PDX Fast Exit
R0 R1 R2 R3
0
Violation 5
WR to WR Cmd in the same bank (tCCD)
R0 R1 R2 R3
0
Violation 15
WR to PRE or PREA
R0 R1 R2 R3
0
Violation 30
WR w/AP to XXX
R0 R1 R2 R3
0
Violation 6
REF to an Non NOP/DES (tRFC)
R0 R1 R2 R3
0
Violation 16
Same rank WR to RD
R0 R1 R2 R3
0
Violation 31
More than 4 ACTs in less than tFAW
R0 R1 R2 R3
0
Violation 7
SRX followed by non-NOP/DES (tXS)
R0 R1 R2 R3
0
Violation 17
Active bank check(s)
R0 R1 R2 R3
0
Violation 32
RD/WR to CKE low or MRS violation
R0 R1 R2 R3
0
Violation 8
REF time is less than tCKE
R0 R1 R2 R3
0
Violation 19
Same rank RD to RD
R0 R1 R2 R3
0
Violation 33
MR0 DLL reset followed by RD or ODT
R0 R1 R2 R3
0
Violation 9
SRX followed by RD,ODT or CKE Low within tXSDLL
R0 R1 R2 R3
0
Violation 20
Different rank RD to RD
R0 R1 R2 R3
0
Violation 34
WR with Synch/Dynamic ODT assert time
R0 R1 R2 R3
0
Violation 10
ACT To ACT (tRRD)
R0 R1 R2 R3
0
Violation 22
Different rank RD to WR
R0 R1 R2 R3
0
Violation 35
CKE Assert Time
CKE0
CKE1
0

Violations Found

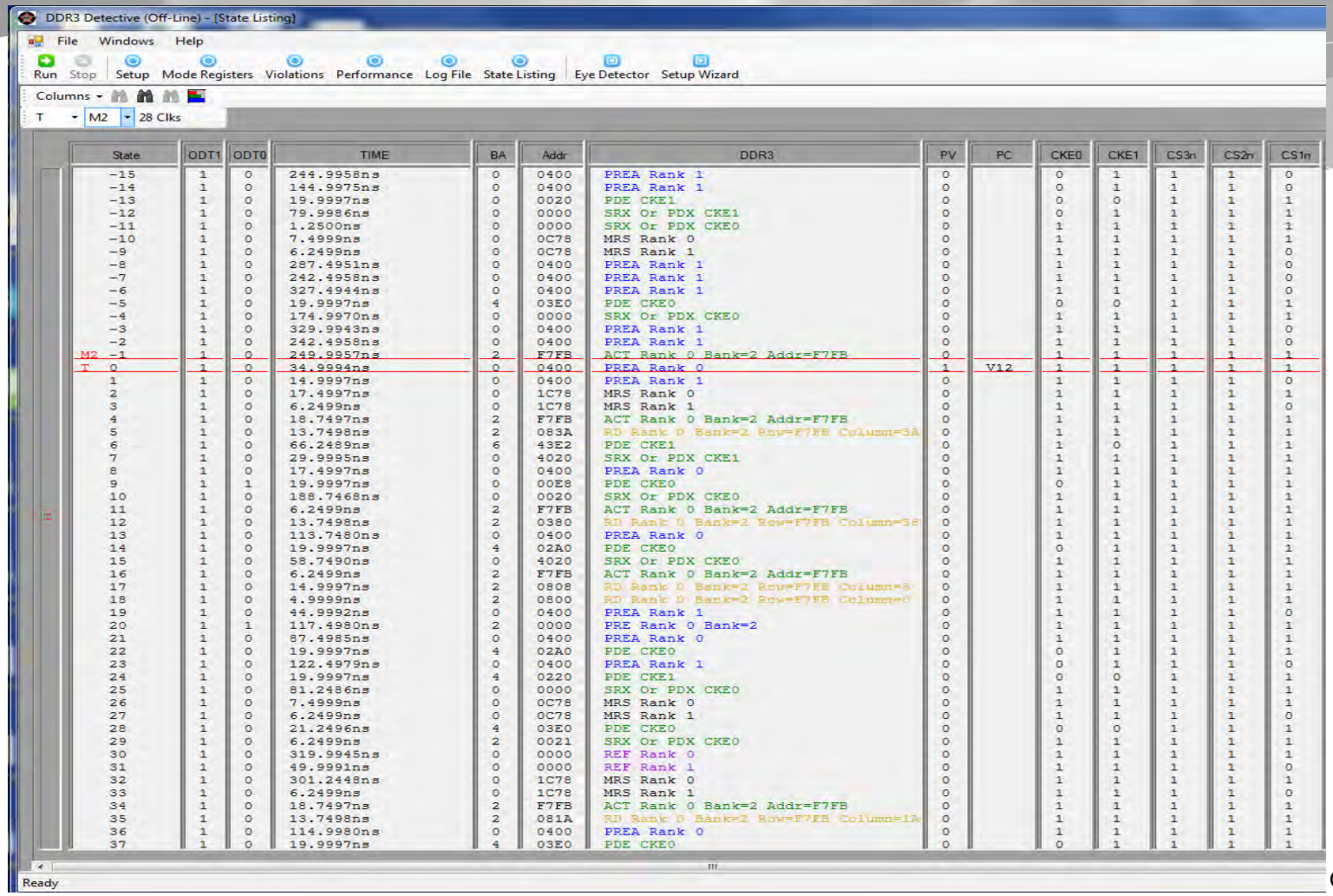
- Protocol Compliance is correct timing between events on the DDR memory bus
- What we found on the server
  - Simultaneous access to two DIMMs
  - Opening and Closing DRAM Banks too close together
  - In adequate Refreshes
  - Incorrect protocol during a Calibrate Command

All of these can cause the DRAM to lose state thus resulting in Data Corruption

Response from the System Vendor

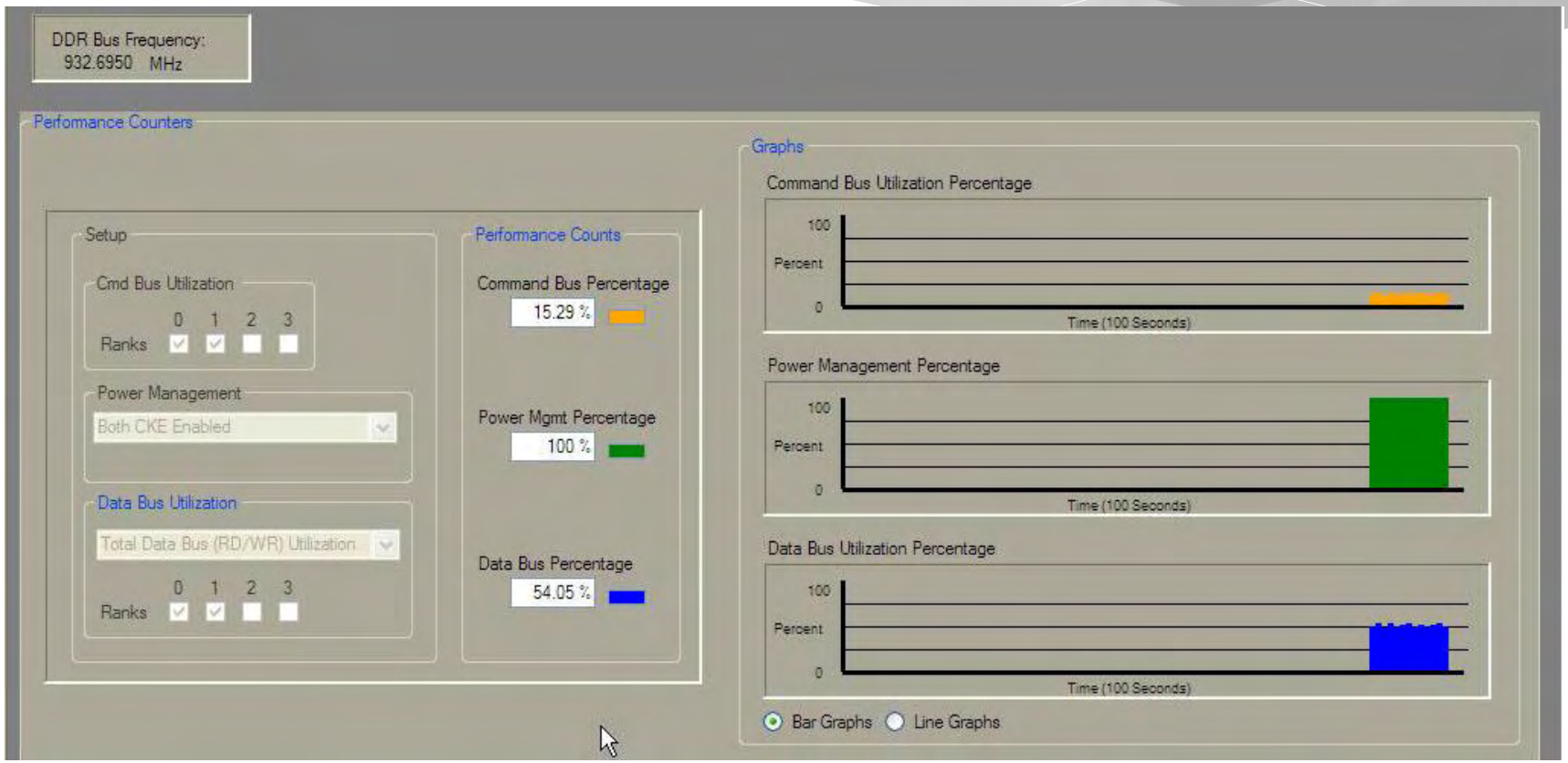
- Thanks for telling us that. It's a BIOS Setting
- Its ok only a few clocks off on the Refreshes
- We tie ODT high and disable that feature in the Mode Registers
- Really?

Smoking Gun



	State	ODT1	ODT0	TIME	BA	Addr	DDR3	PV	PC	CKE0	CKE1	CS3n	CS2n	CS1n
	-15	1	0	244.9958ns	0	0400	PREA Rank 1	0		0	1	1	1	0
	-14	1	0	144.9975ns	0	0400	PREA Rank 1	0		0	1	1	1	0
	-13	1	0	19.9997ns	0	0020	PDE CKE1	0		0	0	1	1	1
	-12	1	0	79.9986ns	0	0000	SRX Or PDX CKE1	0		0	1	1	1	1
	-11	1	0	1.2500ns	0	0000	SRX Or PDX CKE0	0		1	1	1	1	1
	-10	1	0	7.4999ns	0	0C78	MRS Rank 0	0		1	1	1	1	1
	-9	1	0	6.2499ns	0	0C78	MRS Rank 1	0		1	1	1	1	0
	-8	1	0	287.4951ns	0	0400	PREA Rank 1	0		1	1	1	1	0
	-7	1	0	242.4958ns	0	0400	PREA Rank 1	0		1	1	1	1	0
	-6	1	0	327.4944ns	0	0400	PREA Rank 1	0		1	1	1	1	0
	-5	1	0	19.9997ns	4	03E0	PDE CKE0	0		0	0	1	1	1
	-4	1	0	174.9970ns	0	0000	SRX Or PDX CKE0	0		1	1	1	1	1
	-3	1	0	329.9943ns	0	0400	PREA Rank 1	0		1	1	1	1	0
	-2	1	0	242.4958ns	0	0400	PREA Rank 1	0		1	1	1	1	0
M2	-1	1	0	249.9957ns	2	F7FB	ACT Rank 0 Bank=2 Addr=F7FB	0		1	1	1	1	1
T	0	1	0	34.9994ns	0	0400	PREA Rank 0	1	V12	1	1	1	1	1
	1	1	0	14.9997ns	0	0400	PREA Rank 1	0		1	1	1	1	0
	2	1	0	17.4997ns	0	1C78	MRS Rank 0	0		1	1	1	1	1
	3	1	0	6.2499ns	0	1C78	MRS Rank 1	0		1	1	1	1	0
	4	1	0	18.7497ns	2	F7FB	ACT Rank 0 Bank=2 Addr=F7FB	0		1	1	1	1	1
	5	1	0	13.7498ns	2	083A	RD Rank 0 Bank=2 Row=F7FB Column=3A	0		1	1	1	1	1
	6	1	0	66.2489ns	6	43E2	PDE CKE1	0		1	0	1	1	1
	7	1	0	29.9995ns	0	4020	SRX Or PDX CKE1	0		1	1	1	1	1
	8	1	0	17.4997ns	0	0400	PREA Rank 0	0		1	1	1	1	1
	9	1	1	19.9997ns	0	00E8	PDE CKE0	0		0	1	1	1	1
	10	1	0	188.7468ns	0	0020	SRX Or PDX CKE0	0		1	1	1	1	1
	11	1	0	6.2499ns	2	F7FB	ACT Rank 0 Bank=2 Addr=F7FB	0		1	1	1	1	1
	12	1	0	13.7498ns	2	0380	RD Rank 0 Bank=2 Row=F7FB Column=38	0		1	1	1	1	1
	13	1	0	113.7480ns	0	0400	PREA Rank 0	0		1	1	1	1	1
	14	1	0	19.9997ns	4	02A0	PDE CKE0	0		0	1	1	1	1
	15	1	0	58.7490ns	0	4020	SRX Or PDX CKE0	0		1	1	1	1	1
	16	1	0	6.2499ns	2	F7FB	ACT Rank 0 Bank=2 Addr=F7FB	0		1	1	1	1	1
	17	1	0	14.9997ns	2	0808	RD Rank 0 Bank=2 Row=F7FB Column=8	0		1	1	1	1	1
	18	1	0	4.9999ns	2	0800	RD Rank 0 Bank=2 Row=F7FB Column=0	0		1	1	1	1	1
	19	1	0	44.9992ns	0	0400	PREA Rank 1	0		1	1	1	1	0
	20	1	1	117.4980ns	2	0000	PRE Rank 0 Bank=2	0		1	1	1	1	1
	21	1	0	87.4985ns	0	0400	PREA Rank 0	0		1	1	1	1	1
	22	1	0	19.9997ns	4	02A0	PDE CKE0	0		0	1	1	1	1
	23	1	0	122.4979ns	0	0400	PREA Rank 1	0		0	1	1	1	0
	24	1	0	19.9997ns	4	0220	PDE CKE1	0		0	0	1	1	1
	25	1	0	81.2486ns	0	0000	SRX Or PDX CKE0	0		1	1	1	1	1
	26	1	0	7.4999ns	0	0C78	MRS Rank 0	0		1	1	1	1	1
	27	1	0	6.2499ns	0	0C78	MRS Rank 1	0		1	1	1	1	0
	28	1	0	21.2496ns	4	03E0	PDE CKE0	0		0	0	1	1	1
	29	1	0	6.2499ns	2	0021	SRX Or PDX CKE0	0		1	1	1	1	1
	30	1	0	319.9945ns	0	0000	REF Rank 0	0		1	1	1	1	1
	31	1	0	49.9991ns	0	0000	REF Rank 1	0		1	1	1	1	0
	32	1	0	301.2448ns	0	1C78	MRS Rank 0	0		1	1	1	1	1
	33	1	0	6.2499ns	0	1C78	MRS Rank 1	0		1	1	1	1	0
	34	1	0	18.7497ns	2	F7FB	ACT Rank 0 Bank=2 Addr=F7FB	0		1	1	1	1	1
	35	1	0	13.7498ns	2	081A	RD Rank 0 Bank=2 Row=F7FB Column=1A	0		1	1	1	1	1
	36	1	0	114.9980ns	0	0400	PREA Rank 0	0		1	1	1	1	1
	37	1	0	19.9997ns	4	03E0	PDE CKE0	0		0	1	1	1	1

Performance



Summary

- Its Cloudy with a chance of Errors
- Ask your vendor about their validation strategy
- Don’t fill your data center with problems!
- The DDR3 Detective can help
  - Prequalification before purchase
  - Measure effectiveness of your software
  - Validate that BIOS changes are compatible with your DIMMs

FuturePlus Systems

Computer System Validation experts for over 20 years!

www.FuturePlus.com

Barbara Aichinger

Vice President New Business Development

Barb.Aichinger@FuturePlus.com

Office: 603-472-5905

Cell: 603-548-5037