Performance Measurements

Improving Latency and Bandwidth of Your DDR4 System

Barbara Aichinger

Vice President New Business Development

FuturePlus Systems Corporation

MEMCON 2014

Outline

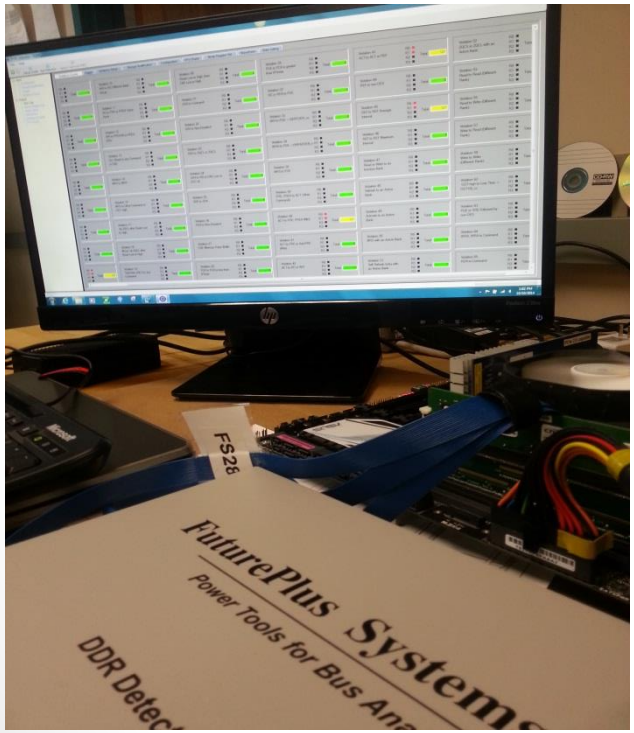
- Performance measurements or Analytics?
  - Events on every cycle @1867MT/s that's is ~ 1 billion events per second
- Work Harder or Smarter?
  - Should we go faster or better use what we have?
- What should we measure to know if we are working hard or working smart?
  - Power Management, Latency, Bandwidth
  - New Metrics

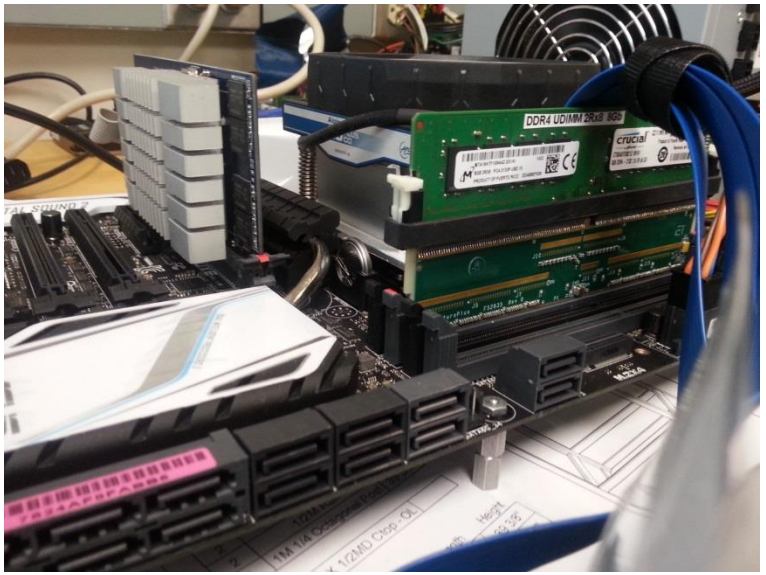
How do I know if I'm working smart?

- Measure it!
- Power Management, Latency, Bandwidth
- But wait…there's more!
  - Page Hit Analysis
  - Multiple Open Banks
  - Bank Group Analysis
  - Bank Utilization
  - Boot Analysis

The Target used

- Asus X99
  - DDR4 1867, Crucial/Micron DIMMs 2Rx8 8Gb
  - Running Google StressApp memory test





Work Smarter not Harder

- For Performance metrics the DDR Detective® uses counters instead the traditional trace memory
  - To capture a second of DDR4 traffic would take 4.5Gbytes of logic analyzer/protocol analyzer trace depth $$$$!
    - 1 hour = 270 Gbytes of trace depth and then time to sift through it and post process!
  - By using large counters and counting events and the time between events we can achieve hours and days worth of metrics with no trace buffer memory and with no time consuming post processing

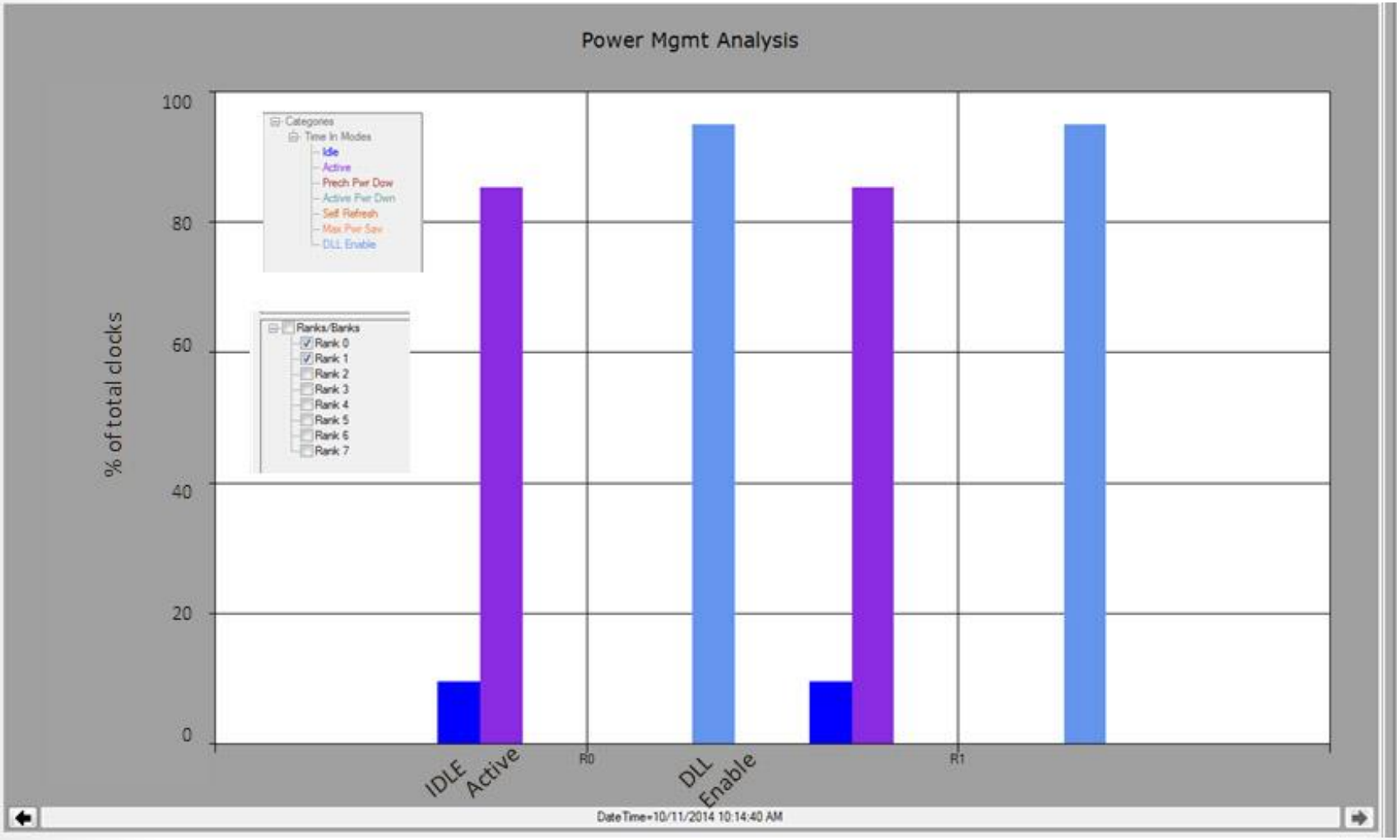
Power Management Metrics

DDR4

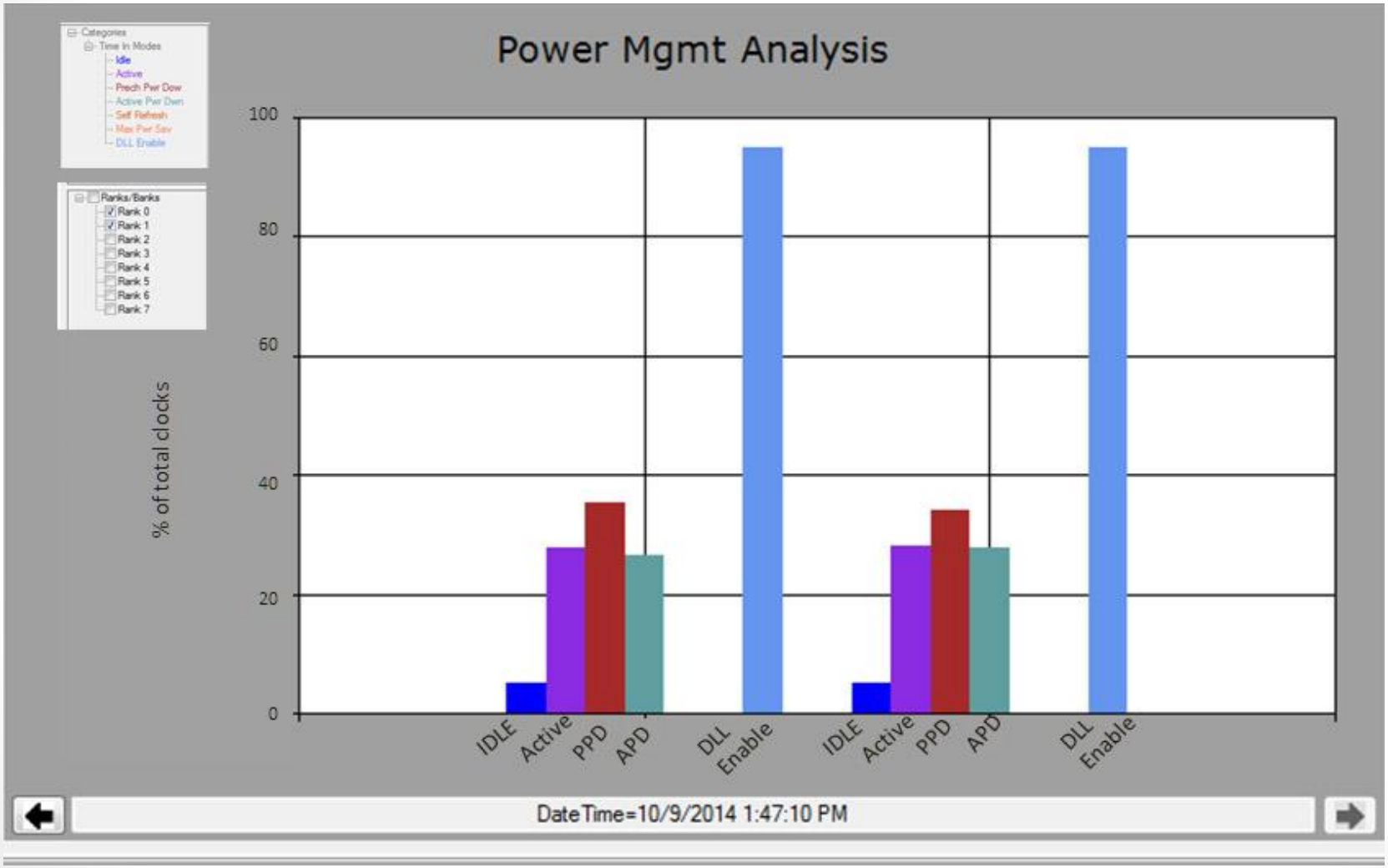
- Idle
- Active
- PreCharge Power Down
- Active Power Down
- Self Refresh
- Max Power Down
- DLL Enable

Power Management

while running StressApp



System Idle


Power Mgmt Analysis
Categories
Time In Modes
Idle
Active
Prech Pwr Dow
Active Pwr Dwn
Self Refresh
Max Pwr Sav
DLL Enable
Ranks/Banks
Rank 0
Rank 1
Rank 2
Rank 3
Rank 4
Rank 5
Rank 6
Rank 7
% of total clocks
100
80
60
40
20
0
IDLE
Active
PPD
APD
DLL Enable
IDLE
Active
PPD
APD
DLL Enable
DateTime=10/9/2014 1:47:10 PM

Power Management

- ~50M servers Servers World Wide
- Each Server averages 16-24 DIMMs
  - 800M to 1.2B DIMMs
- Even a small power savings per DIMM can add up

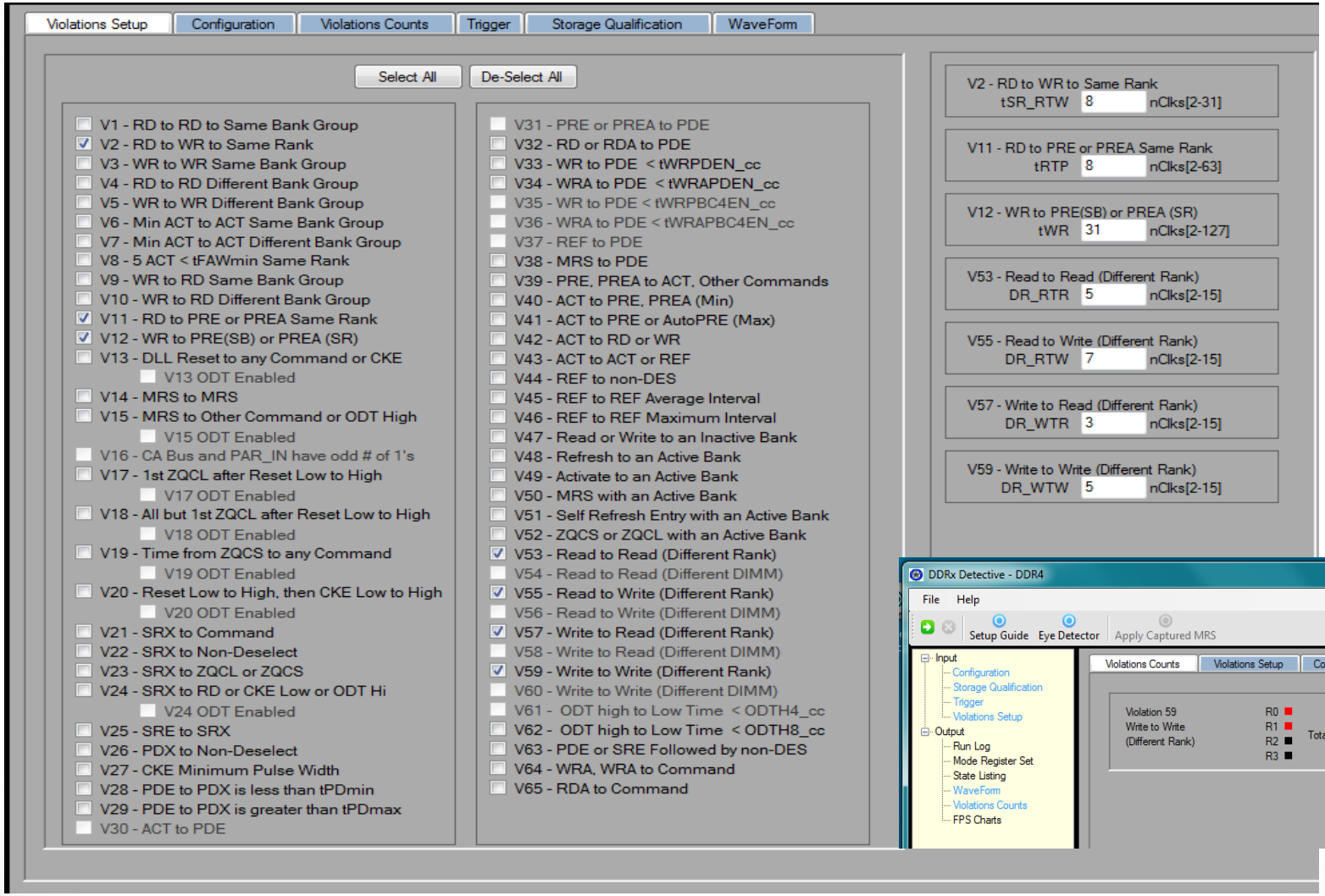
*Every time **Facebook**'s data center engineers figure out a way to **reduce** server consumption **by a single watt**, the improvement, at Facebook's scale, has the potential to add **millions of dollars** to the company's bottom line.*

Yevgeny Sverdlik
Editor in Chief
Data Center Knowledge

Latency

- Several Jedec Paramters apply:
  - RD to WR same rank tSR_RTW
  - RD to PRE/PREA same Rank tRTP
  - WR to PRE(SB) or PREA (SR) tWR
  - Read to Read different Rank tDR_RTR
  - Read to Write different Rank DR_RTW
  - Write to Read different Rank tDR_WTR
  - Write to Write different Rank tDR_WTW

Measure it!



KEYSIGHT TECHNOLOGIES
Strategic Solutions Partner

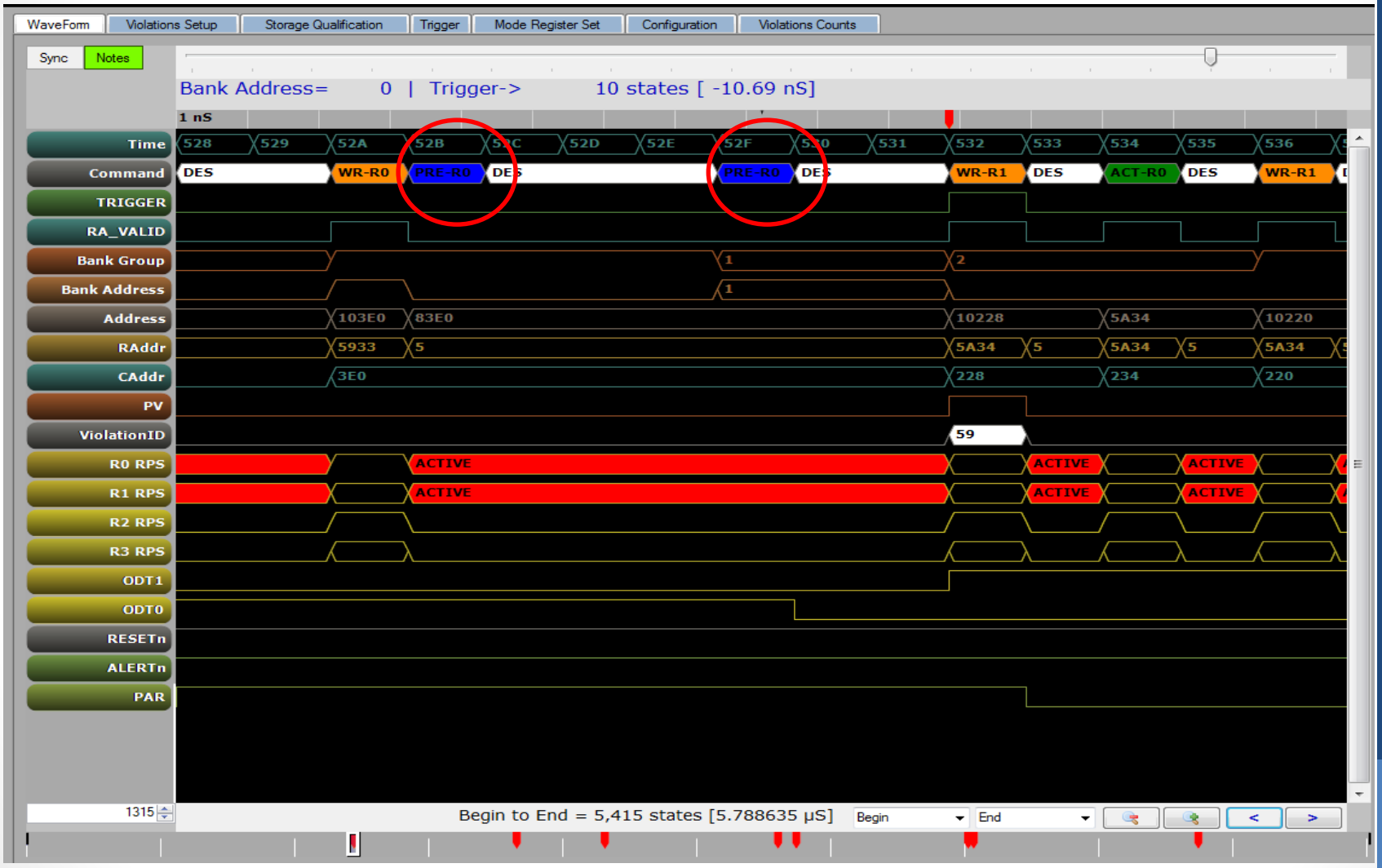
MEMCON 2014

FuturePlus Systems
Power Tools for Memory Bus Analysis

Latency Measurements

V#	Parameter	Description	Spec	Measured
V2	tSR_RTW	RD to WR same Rank	8	10
V11	tRTP	RD to PRE same Rank	8	8
V12	tWR	WR to PRE SB or PREA SR	31	31
V53	tDR_RTR	RD to RD diff Rank	5	6
V57	tDR_WTR	WR to RD diff Rank	3	6
V59	tDR_WTW	WR to WR diff Rank	5	8

Intervening Commands



Latency Measurements

V#	Parameter	Description	Spec	Measured
V1	tCCD_L	RD to RD Same Bank Group	5	6
V3	tCCD_L	WR to WR Same Bank Group	5	6
V4	tCCD_S	RD to RD diff Bank Group	4	4
V5	tCCD_S	WR to WR diff Bank Group	4	4
V6	tRRD_L	ACT to ACT Same Bank Group	5	5
V7	tWTR_L	ACT to ACT diff Bank Group	4	4
V9	tWTR_L	WR to RD Same Bank Group	22	23
V10	tWTR_S	WR to RD Diff Bank Group	17	19

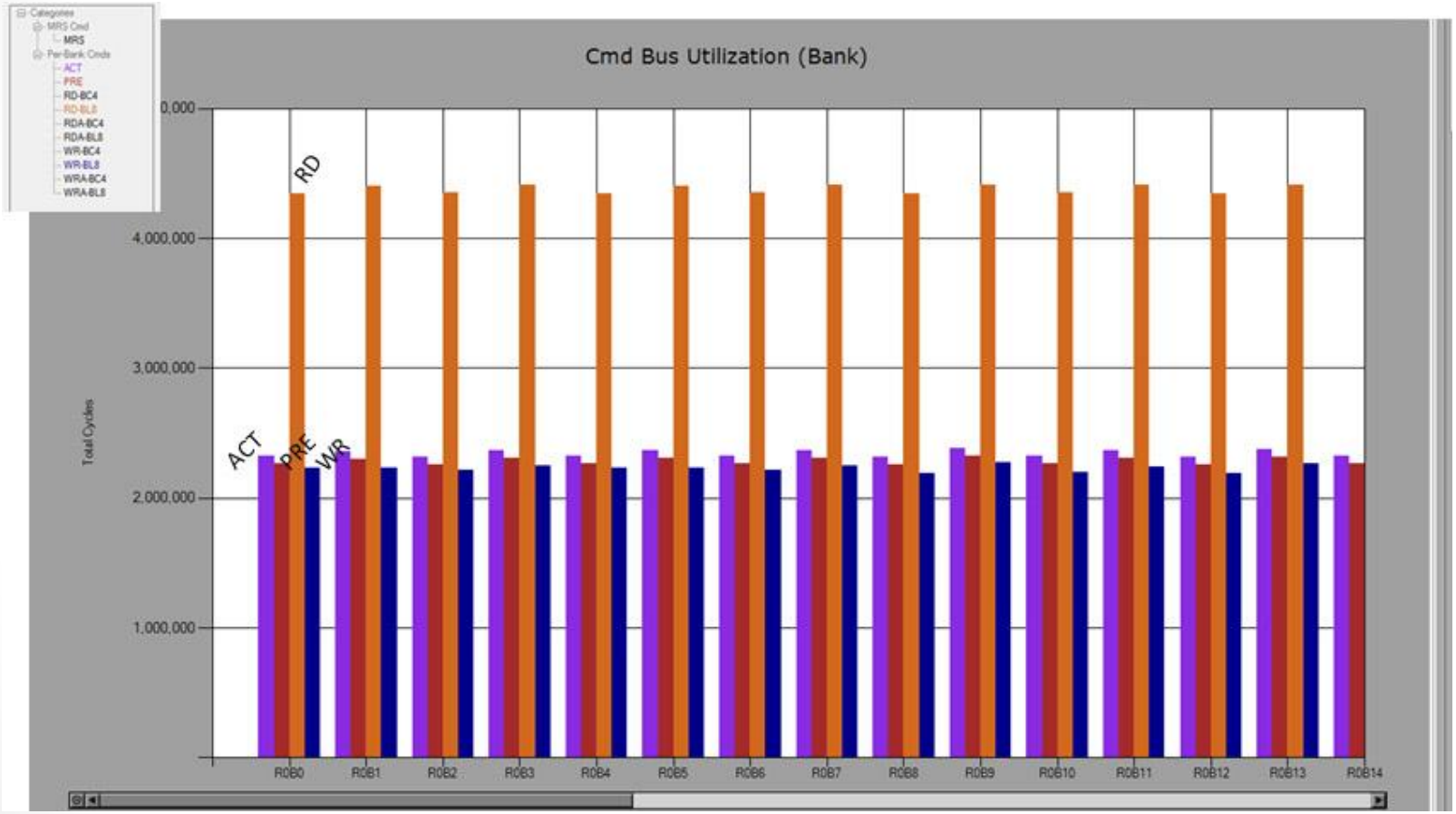
Latency

- Good designs operate on the edge of the spec
- Architectural tradeoffs will occur
- Do I need margin?
  - Design for the worst case and buy quality parts

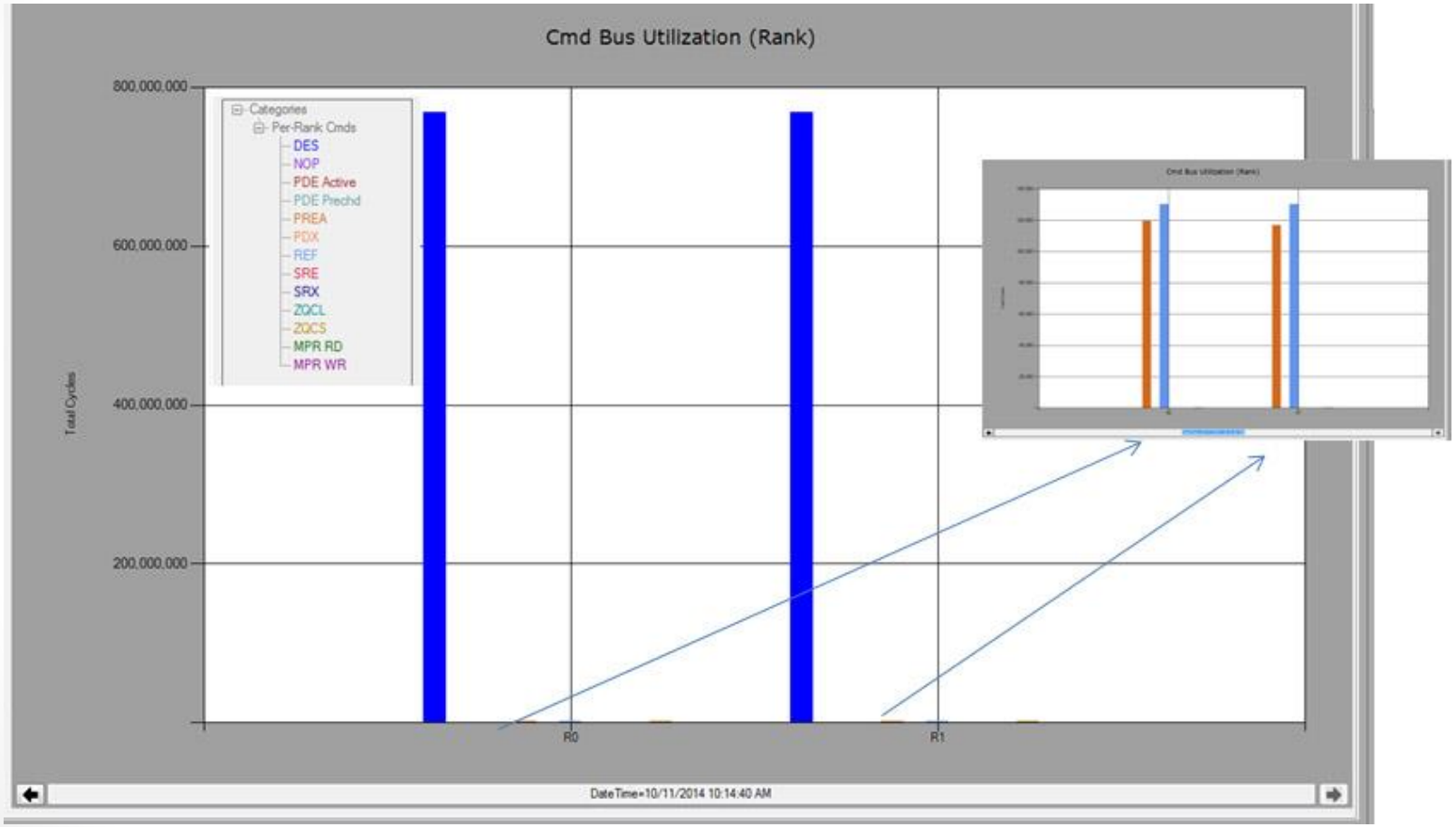
Bandwidth

- Overhead
  - Any use of the bus other than a Read or a Write
  - Command Bus Utilization
- Data Bus
  - Utilization: the % of the time that Read or Write Data is being transferred
  - Bandwidth: the amount of data transferred per second

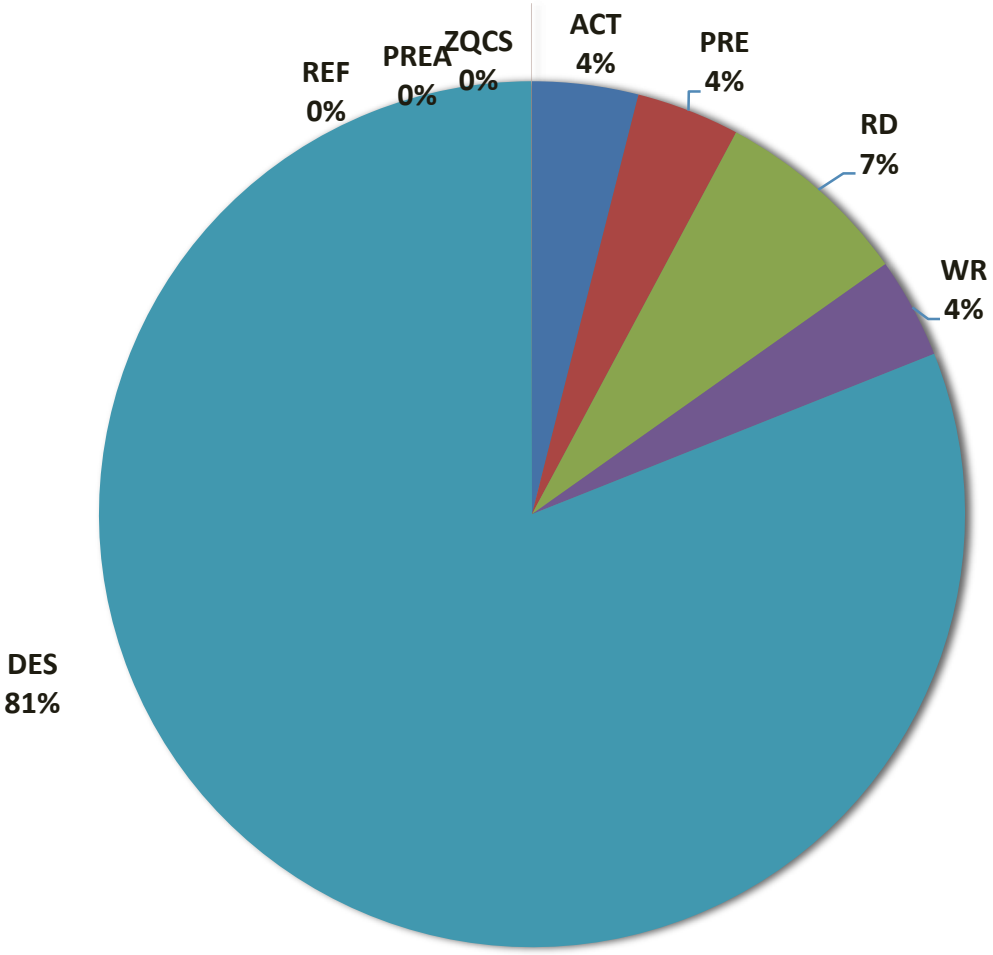
Command Bus Utilization



DES,REF, ZQCL

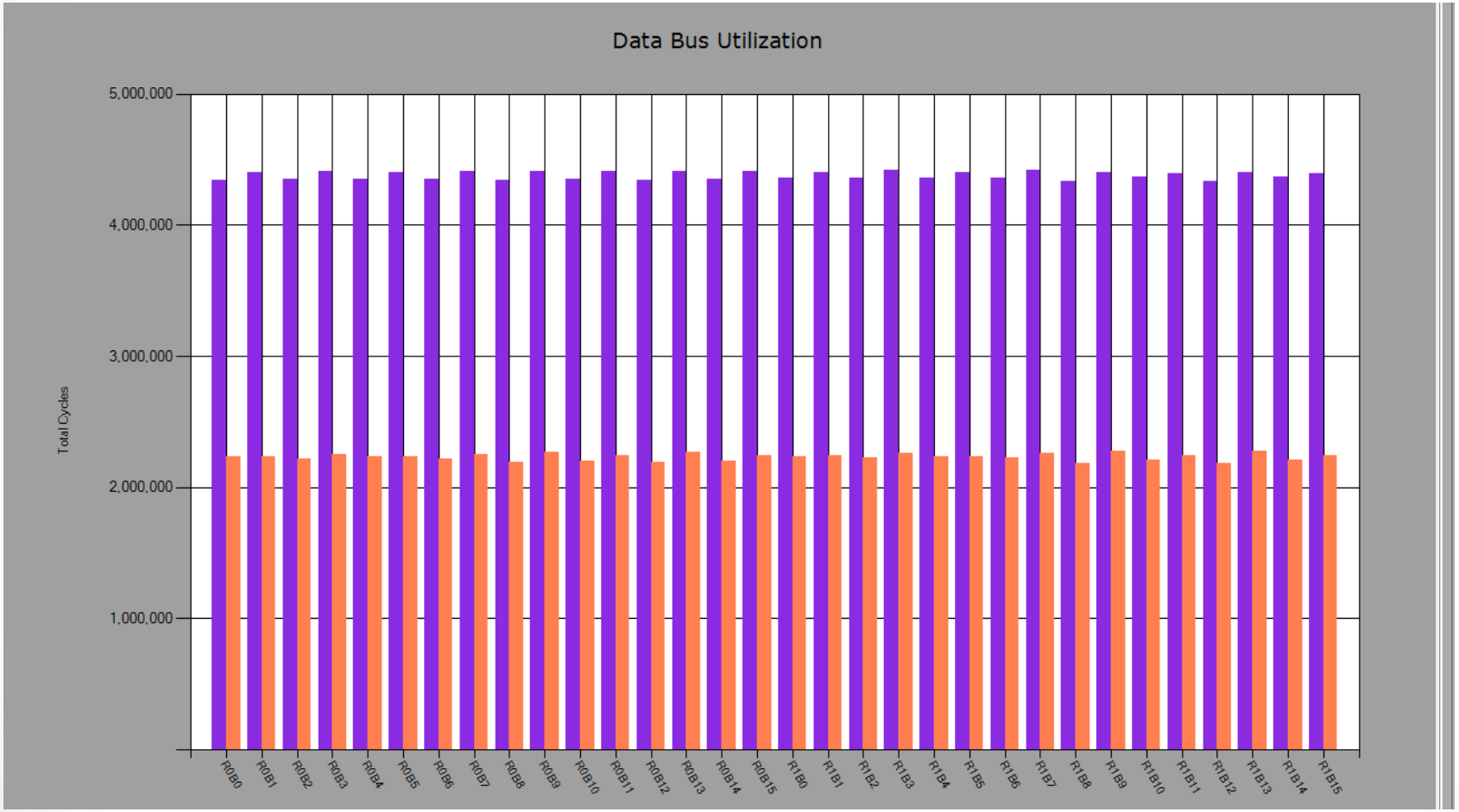

Cmd Bus Utilization (Rank)
800,000,000
600,000,000
400,000,000
200,000,000
Total Cycles
Categories
Per-Rank Cmds
DES
NOP
PDE Active
PDE Precha
PREA
PDX
REF
SRE
SRX
ZQCL
ZQCS
MPR RD
MPR WR
R0
R1
DateTime=10/11/2014 10:14:40 AM

Summarize Command Bus Utilization

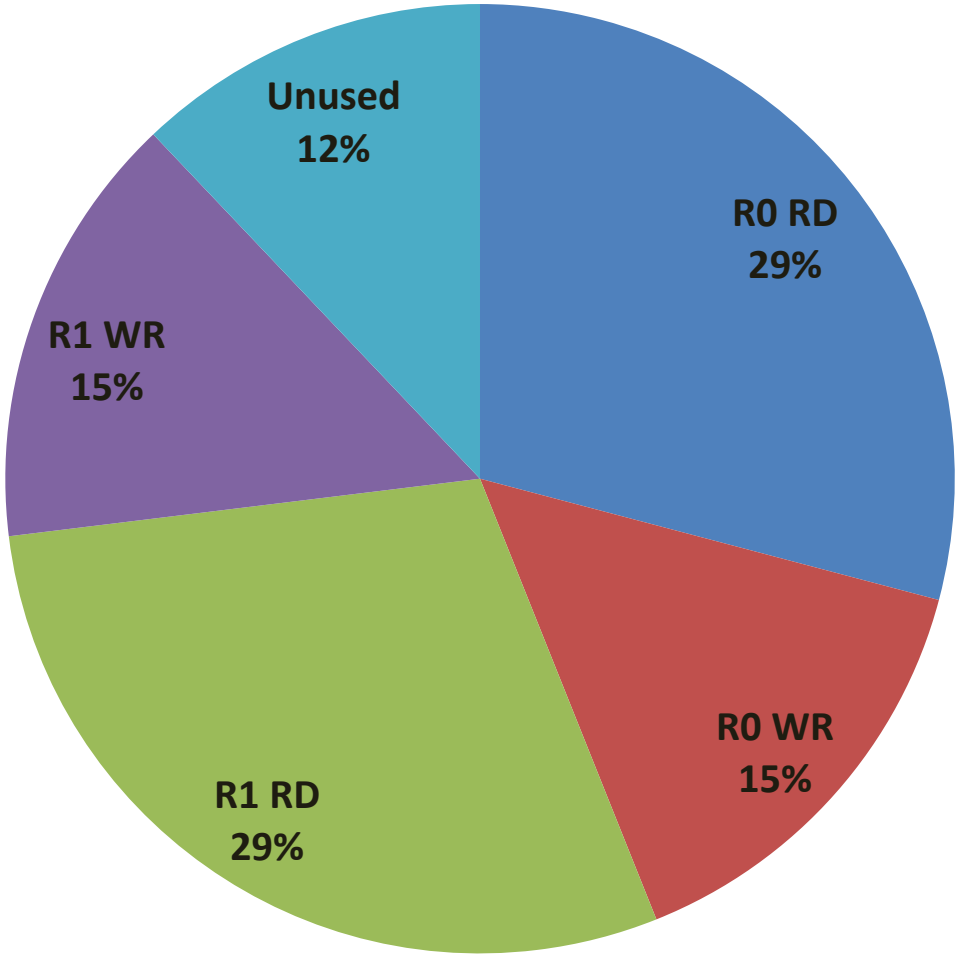


Command Bus Utilization

Data Bus Utilization

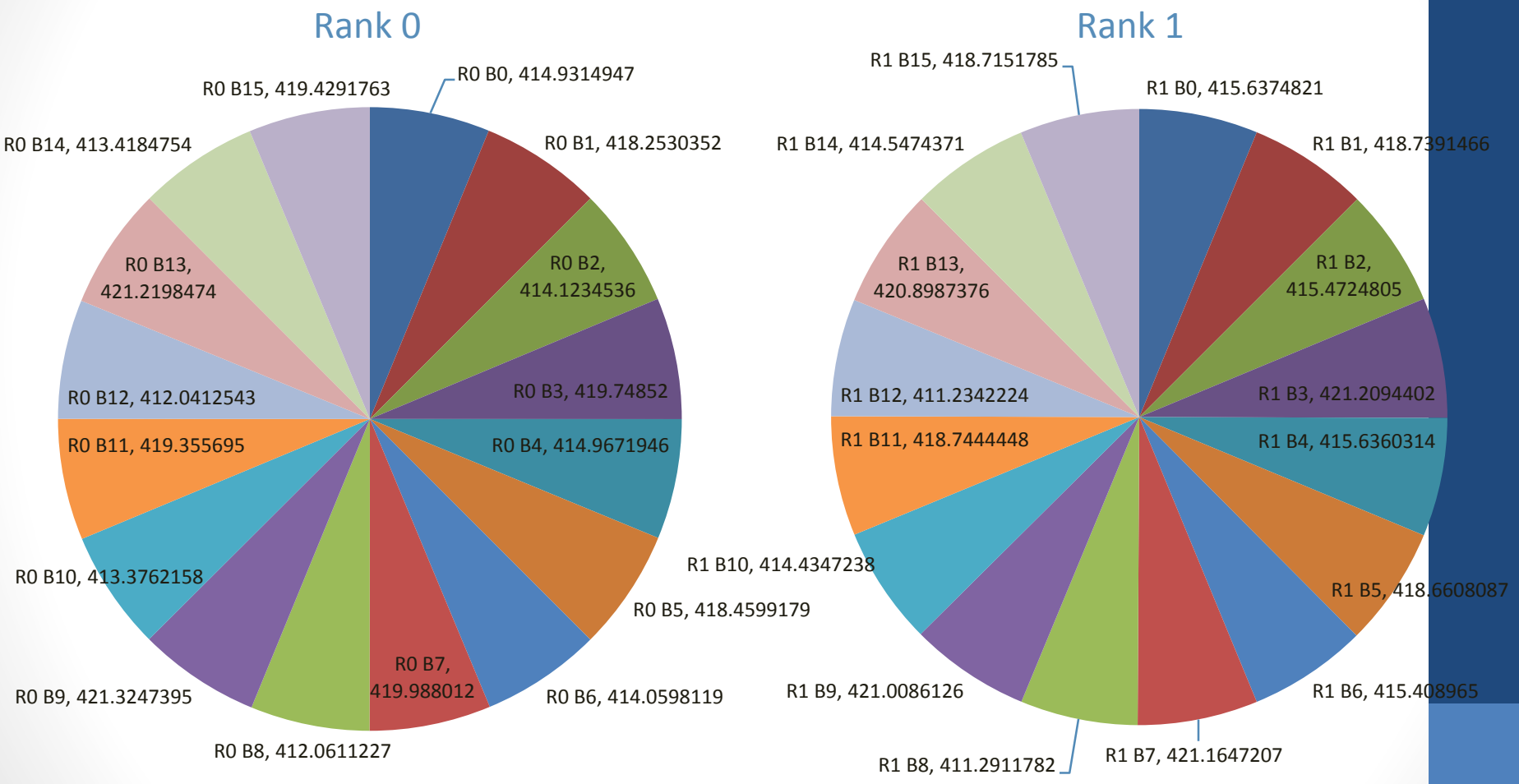


Data Bus Utilization Summary

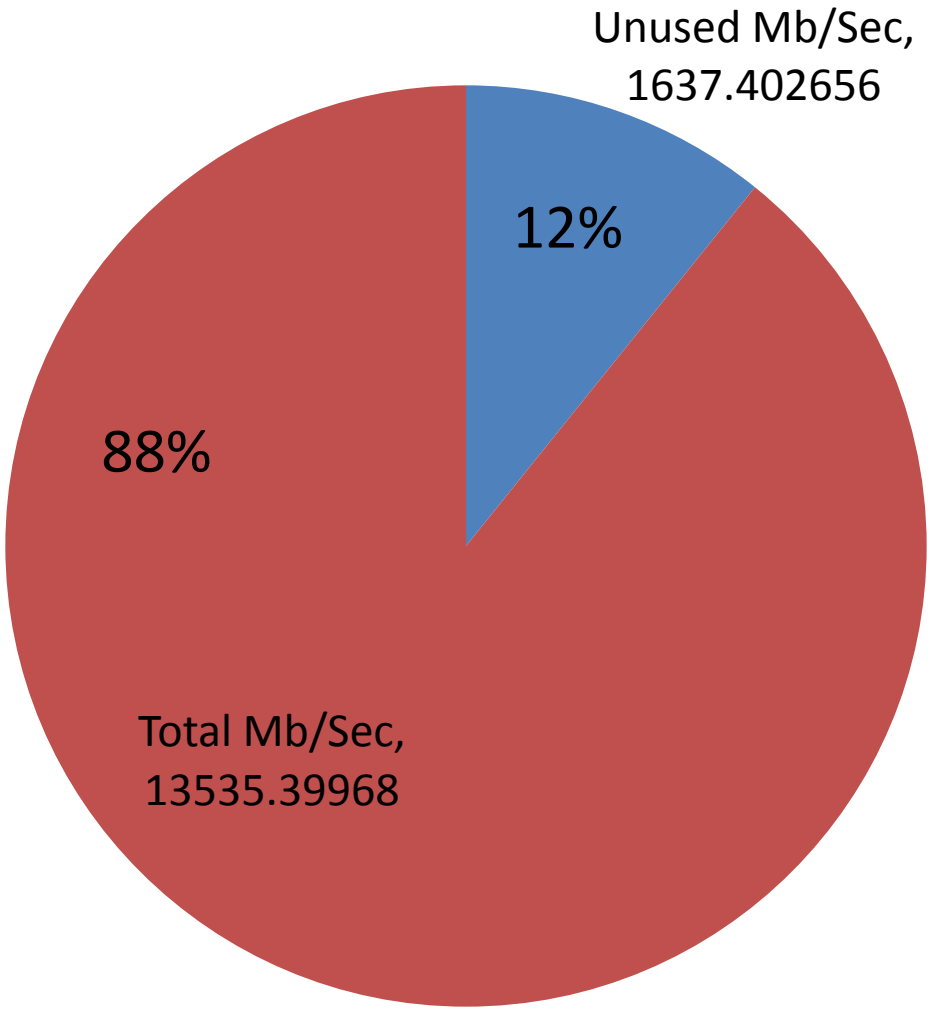


Data Bus Bandwidth

Mbytes transferred in 1 second



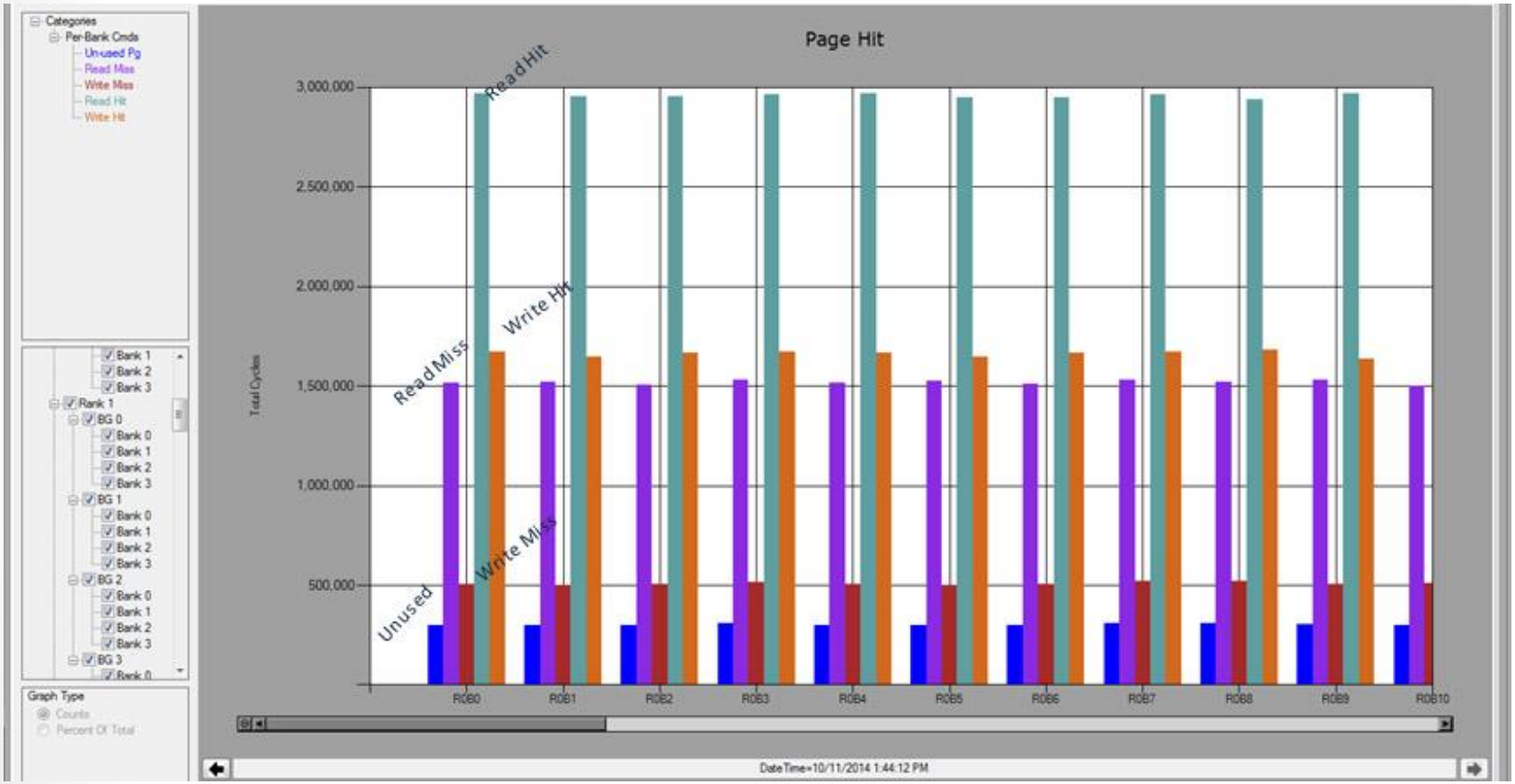
Total Bandwidth



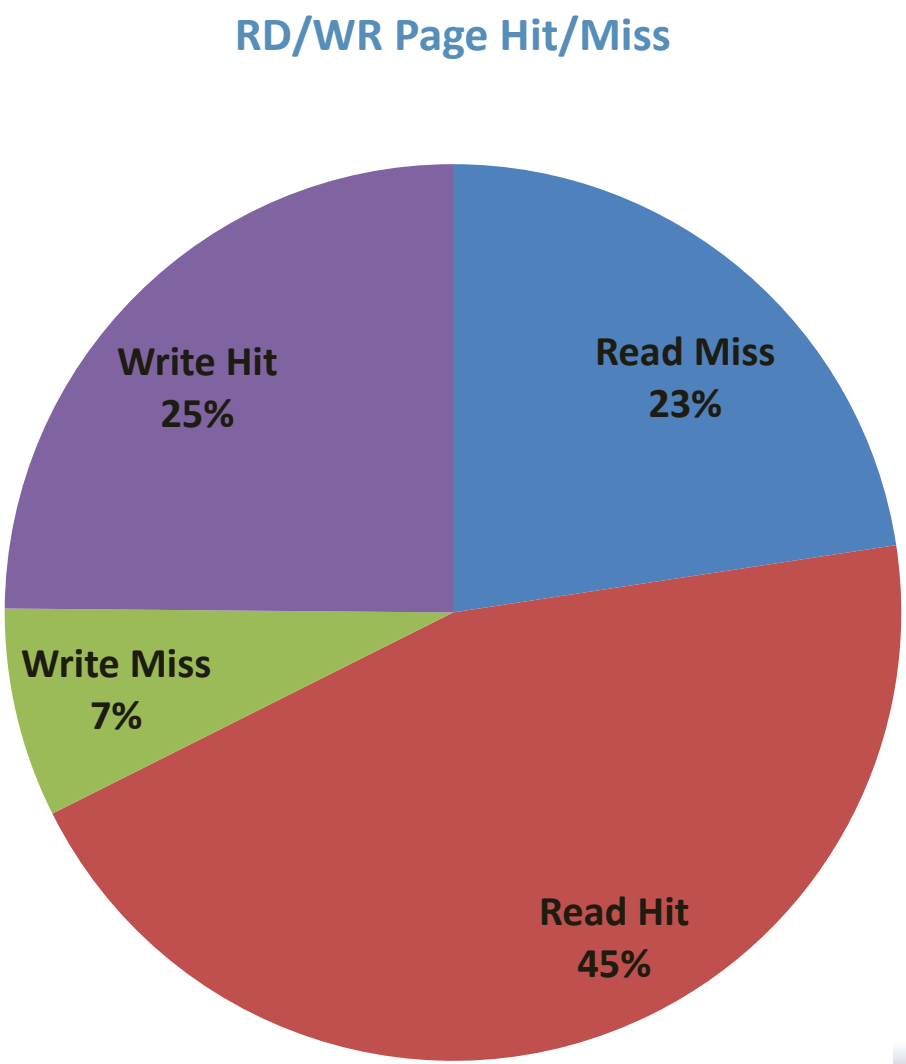
Insight beyond the basics

- Page Hit Analysis
  - Read Miss
  - Write Miss
  - Unused
- Multiple Open Banks
  - Open Banks make for faster access if your going there…performance hit if your not
  - Power hit when banks are open
- Bank Group Analysis
  - New for DDR4 back to back access to same bank performance hit
  - Faster to go back to back to different bank groups

Page Hit Analysis

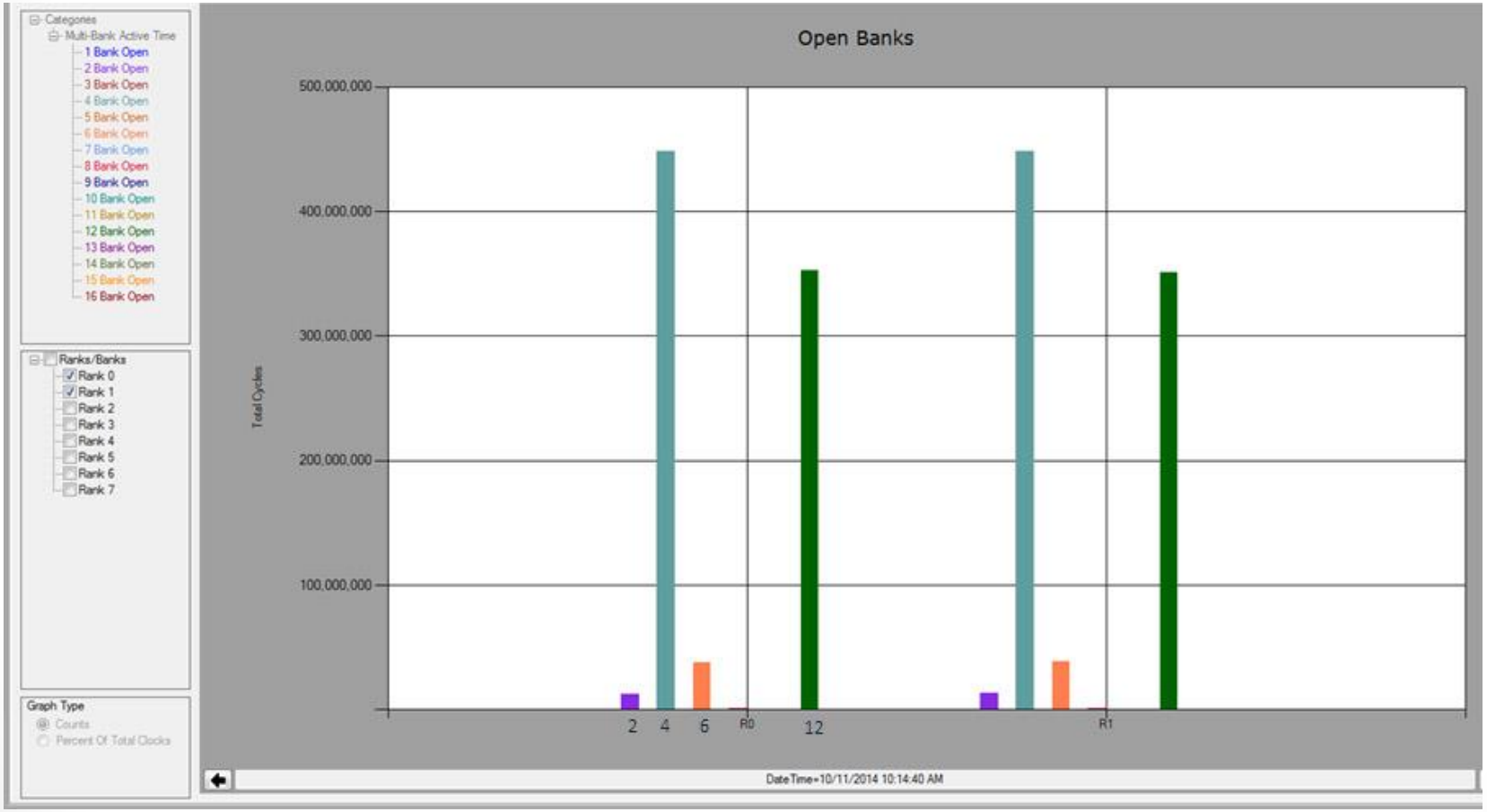


Page Hit by percentages



Multiple Open Banks

How many are open at any one time

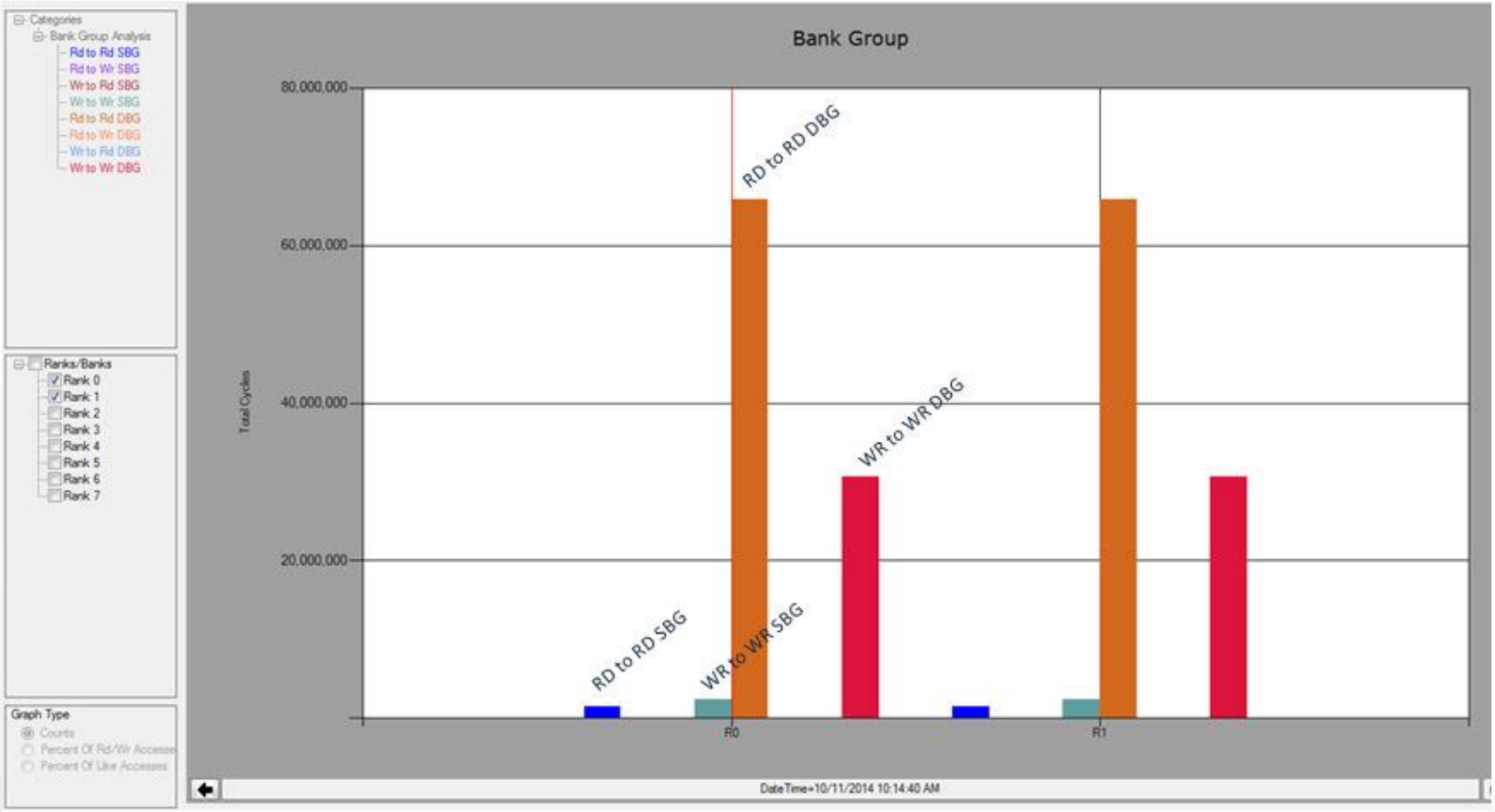


Bank Group Access Analysis

- tCCD_L
  - Takes longer for back to back RD/WR accesses to the same bank group
- tCCD_S
  - Can reduce latency by going to different bank groups

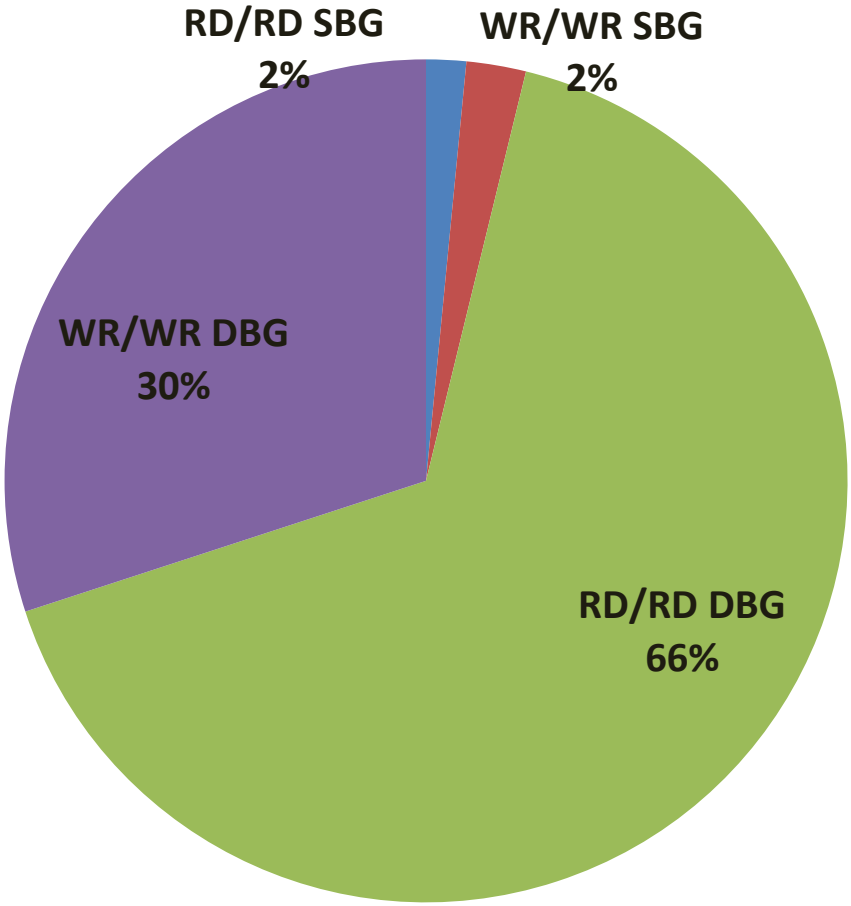
Bank Group Access Analysis

Relative to the previous transaction how many times did the following transaction go to the same/different bank group



Bank Group Access Analysis

by Percentage

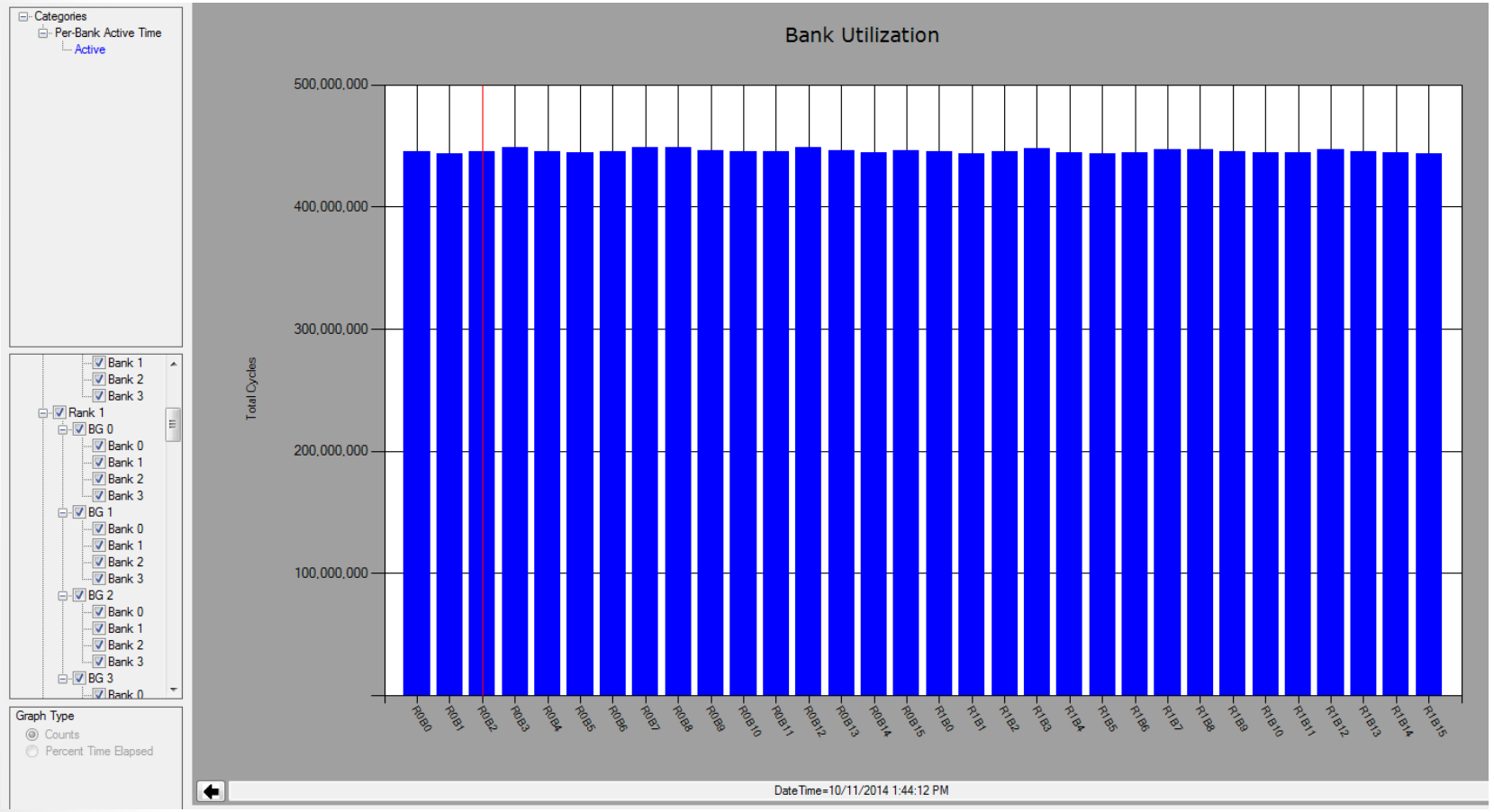


But Wait there's more!

- Bank Utilization
  - What happens during a chip kill or page retirement scenario?
  - How does the traffic reallocate?
  - What are the performance implications?
- Do I have system hot spots?
  - Row Hammer (excessive Activates)
- Fast Boot
  - Why does the system take so long to boot?

Bank Utilization

Number of cycles the banks are open

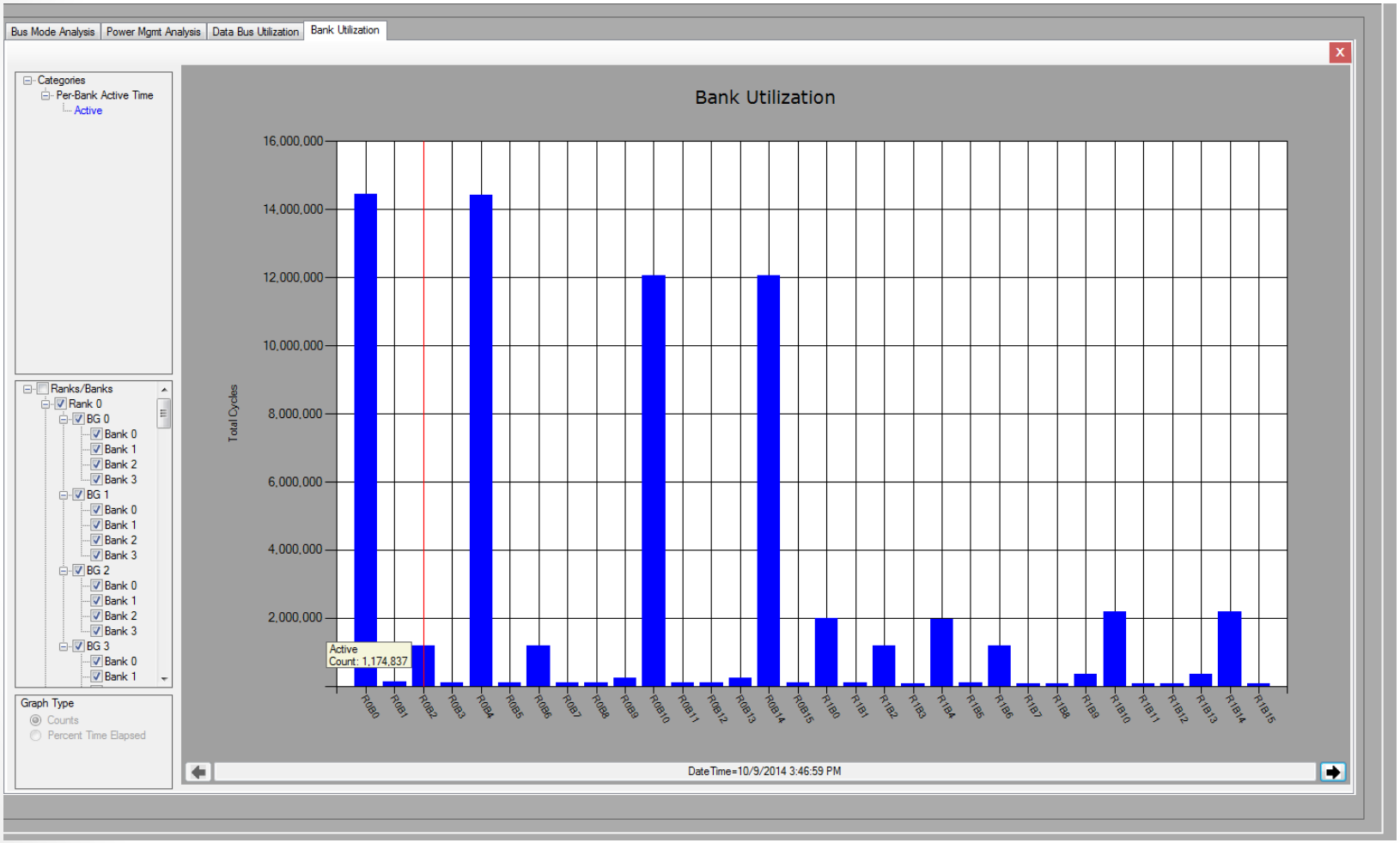


KEYSIGHT TECHNOLOGIES
Strategic Solutions Partner

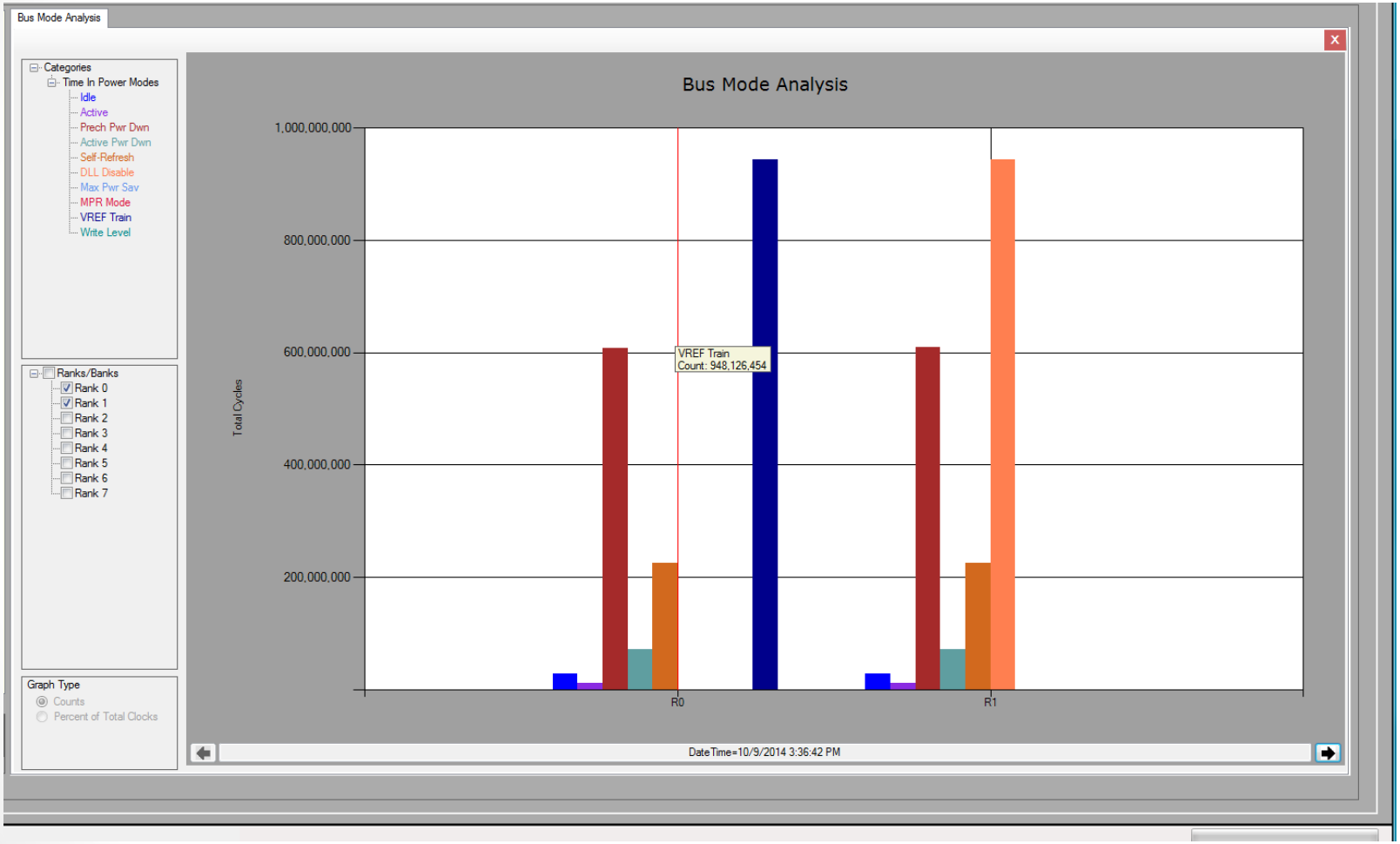
FuturePlus Systems
Power Tools for Memory Bus Analysis

Bank Utilization

When system sitting idle



Boot Analysis



Advancing the State of the Art

- Memory Controller/System Architecture
  - Can this insight lead to better designs?
  - Dynamic architecture based on workload?
- Which software to run that stresses the system the best and shows architectural flaws? (benchmarking)
  - Looking for feedback from the industry…we can test using the DDR Detective®

Summary

- Power Management, Bandwidth, Latency
- NEW Metrics:
  - Page Hit Analysis
  - Multiple Open Banks
  - Bank Group Analysis
  - Bank Utilization
  - Boot Analysis
  - New Measurements give insight into new designs and better architectures

Contact Information

Barbara Aichinger

Vice President New Business Development

FuturePlus Systems

Barb.Aichinger@FuturePlus.com

www.FuturePlus.com

Check out our new website dedicated to DDR Memory! www.DDRDetective.com