

OPEN Compute Engineering Workshop March 9, 2015 San Jose

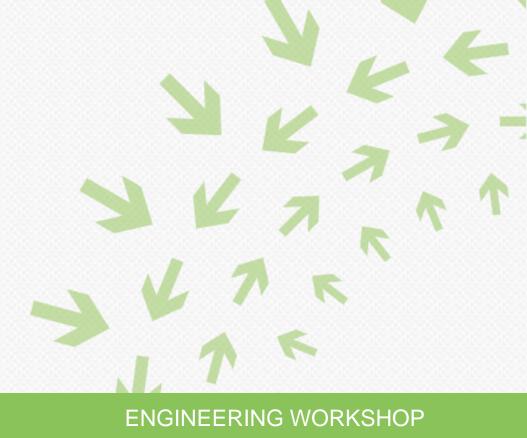


Server Memory Performance

Characterizing Workloads

March 9th 2015

Barbara Aichinger FuturePlus Systems Vice President New Business Development



Agenda

- What is DDR4 Memory?
- Traditional Performance Metrics
- New Performance Metrics
- How can we monitor the DDR Memory?
- Summary



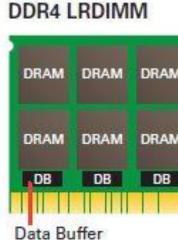


Power Tools for Bus Analysis

DDR4: The Next Generation

- -FASTER: 1600MT/s, 1866MT/s, 2133MT/s, 2400MT/s, 2666MT/s, 3200MT/s (25.6 GB/s)
- Lower Voltages
- More power saving features
- Higher Density
 - 3DS: 3D stacking
- LRDIMM: Load Reduced DIMM
- More robust

Alert signal for ECC errors, Command/Address Parity



M	DRAM	DRAM	RCD	DRAM	DRAM	DRAM	DRAM	
м	DRAM	DRAM		DRAM	DRAM	DRAM		
	DB	DB		DB	DB	DB		

Very short frontside trace Backside same as RDIMM trace length

FuturePlus Systems

Power Tools for Bus Analysis

Traditional Measurements

- Bandwidth
 - Command Bus Utilization
 - Data Bus Utilization
- Power Management
- Latency

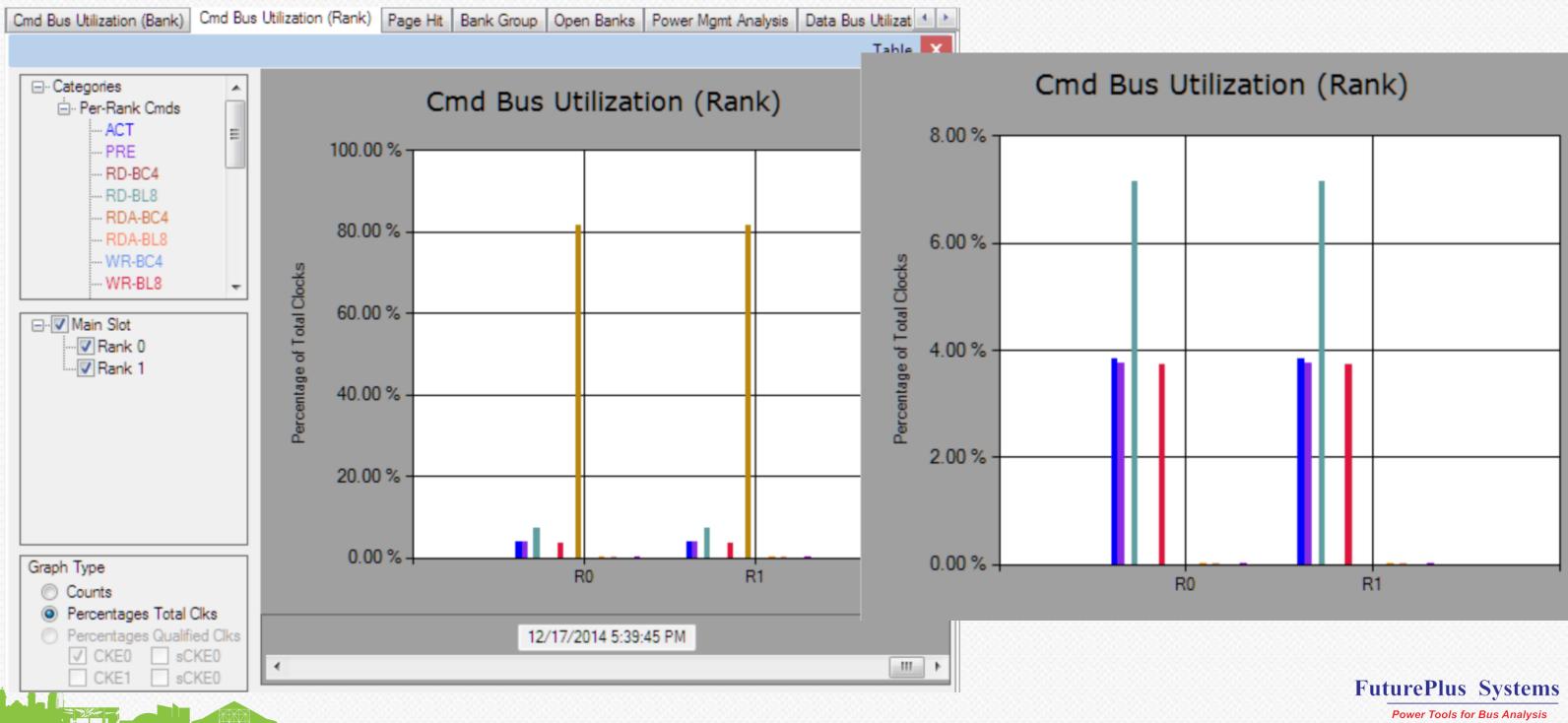
Labo



Power Tools for Bus Analysis

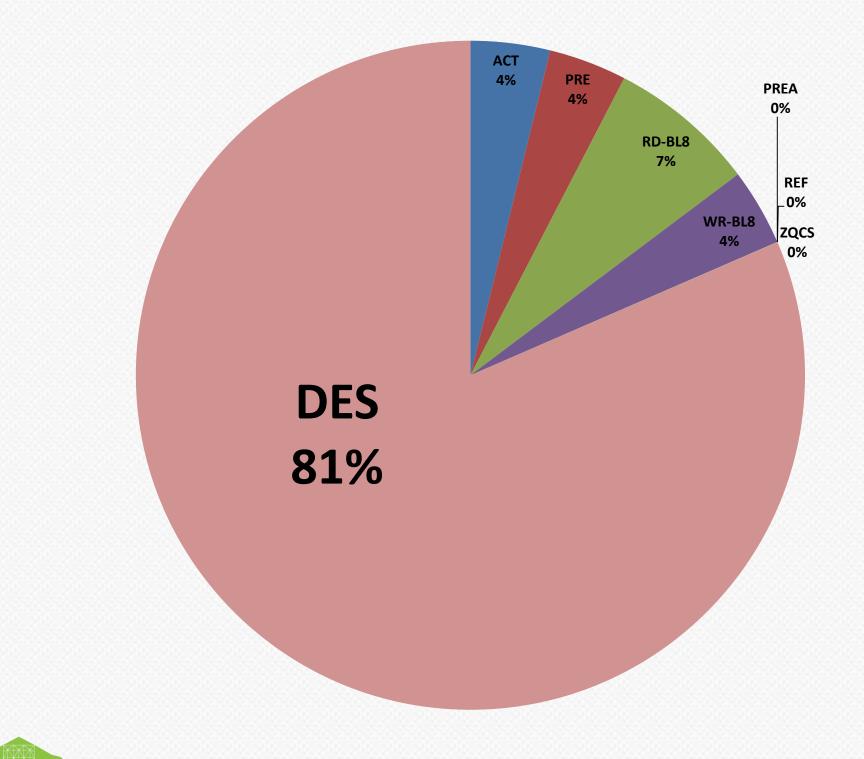
Command Bus Utilization

2400MT/s



Command Bus Utilization

Label





Power Tools for Bus Analysis

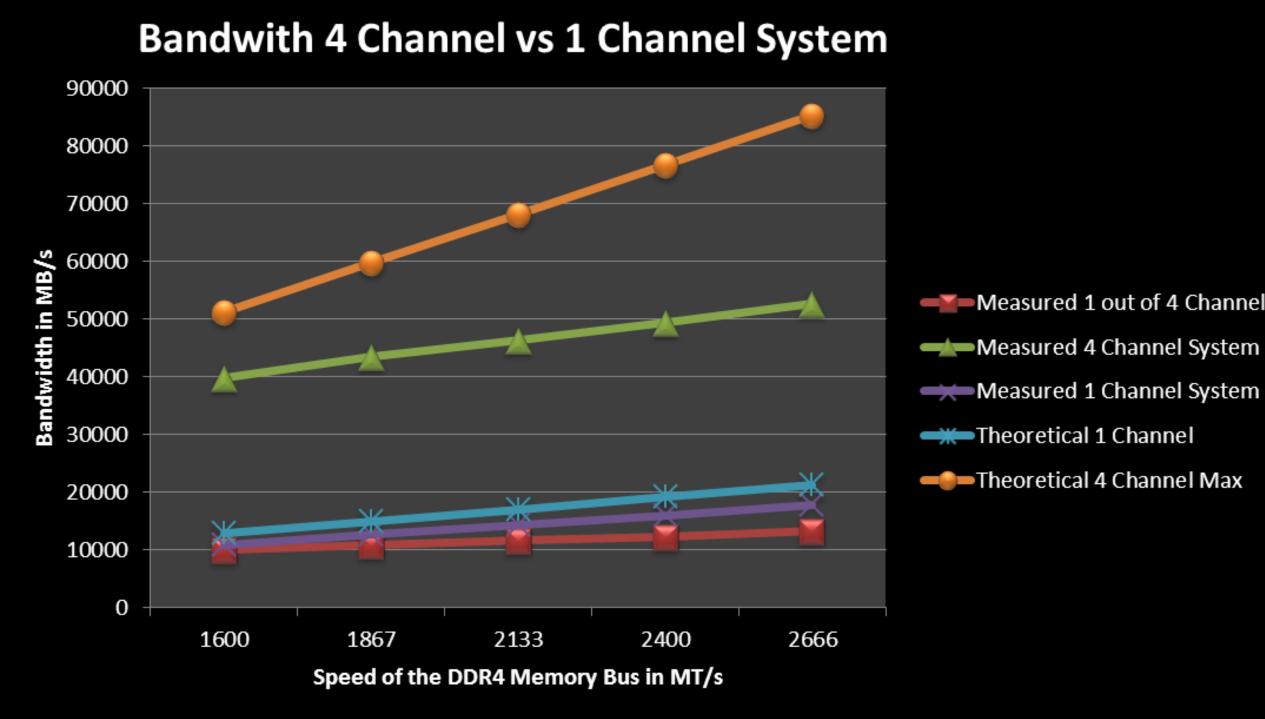
Data Bus Utilization

2400MT/s





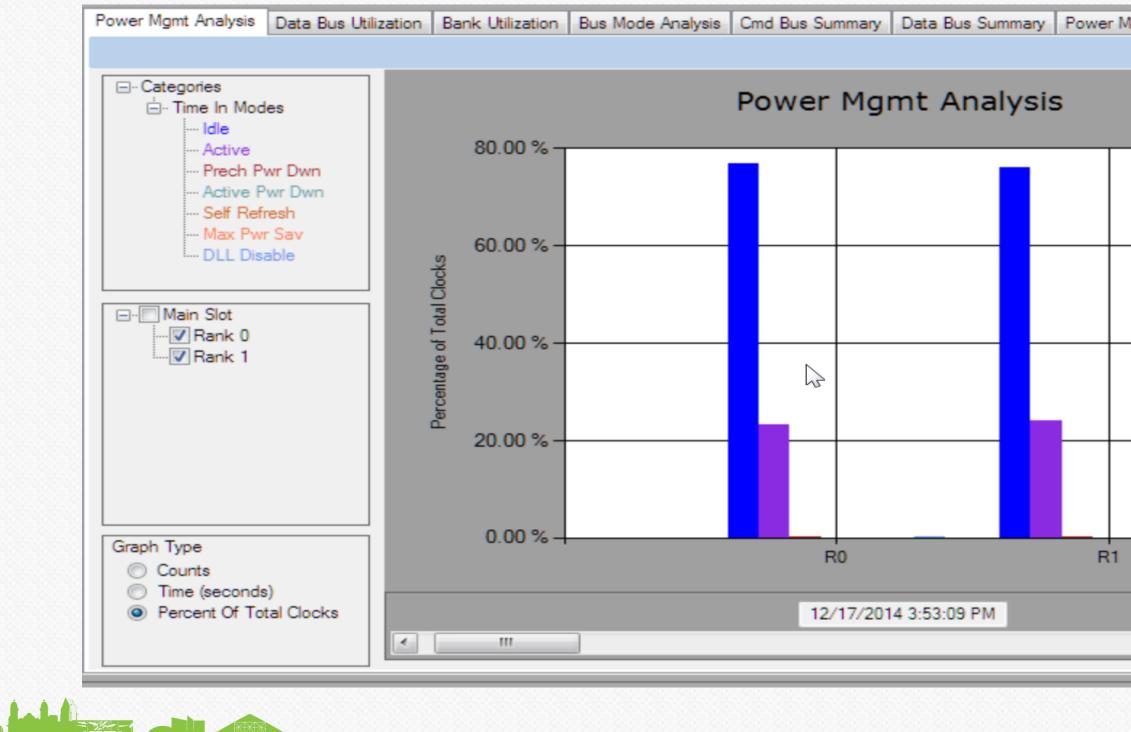
DDR4 Bandwidth



Measured 1 out of 4 Channels



Power Management Analysis



Embedded Video

gmt Summary Bank Group Sur	4 >	
Table	×	
	4	
	FuturePlus	Systems

Power Tools for Bus Analysis

Latency Several JEDEC Parameters 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

1.h



Power Tools for Bus Analysis

Latency Measurements

measurement made at 1867

Labor

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

FuturePlus Systems

Power Tools for Bus Analysis

Intervening Commands Performance versus Power Management Tradeoffs

WaveForm	Violation	ns S e tup	Storage G	ualification	Trigger	Mode Register Se	t Configuratio	on Via	lations Counts				
Sync N	lotes	-											
		Bank	Address=			ger->							
		1 nS											
	Time	528	×529	X52A	528	X52C X52D) X52 E	(52F	<u> </u>	X531	532	X533	X534
Co	ommand	DES		WR-R0	PRE-RO	DES		PRE-R	DES		WR-R1	DES	ACT-R0
	RIGGER												
R/	VALID												
Ban	k Group			у				\1			χ2		
Bank	Address				\			<u></u> 1			λ		
	Address			X103E0	X83E0						10228		×5A34
	RAddr			X5933	χ5						X5A34	χ5	X5A34
	CAddr			JEO							228		234
	PV												
Vio	lationID										59		
	R0 RPS	_		y	ACTIVE						X	ACTIVE	X
	R1 RPS			X	ACTIVE						X	ACTIVE	X
	R2 RPS				<u>`</u>						_/	<u>`</u>	/
	R3 RPS				λ							λ	λ
	ODT1												
	ODT0 RESETn												
	ALERTN	——											
	PAR	_											
	1315 💠				B	egin to End = !	5,415 states	[5.7886	35 µS1	Begin	▼ End	-	
				-									

tot 1





New Performance Metrics

Page Hit Analysis

- Read Hit: Page was Open
- Read Miss : Page was not Open, Transaction was preceded by an ACT
- Write Hit: Page was Open
- Write Miss: Page was not Open, Transaction was preceded by an ACT
- Unused: Page was opened and closed and never accessed

Multiple Open Banks

- Open Banks make for faster access IF your going to that bank on the next access...performance hit if your not
- Power hit when banks are open

Bank Group Analysis

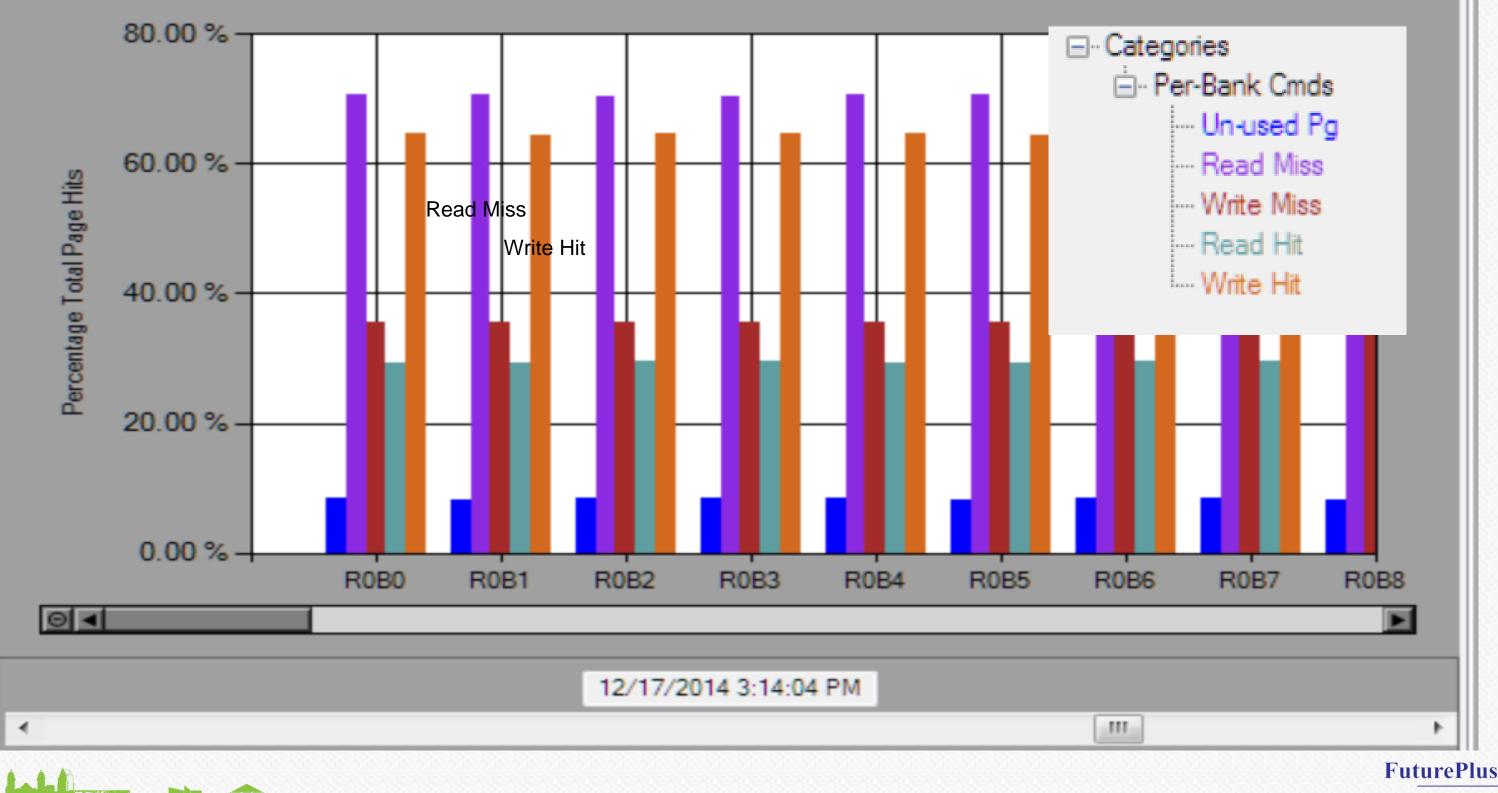
- New for DDR4: Back to back access to same bank is a performance hit
- Faster to have back to back accesses to different bank groups



Power Tools for Bus Analysis

Running Google StressApp @2133

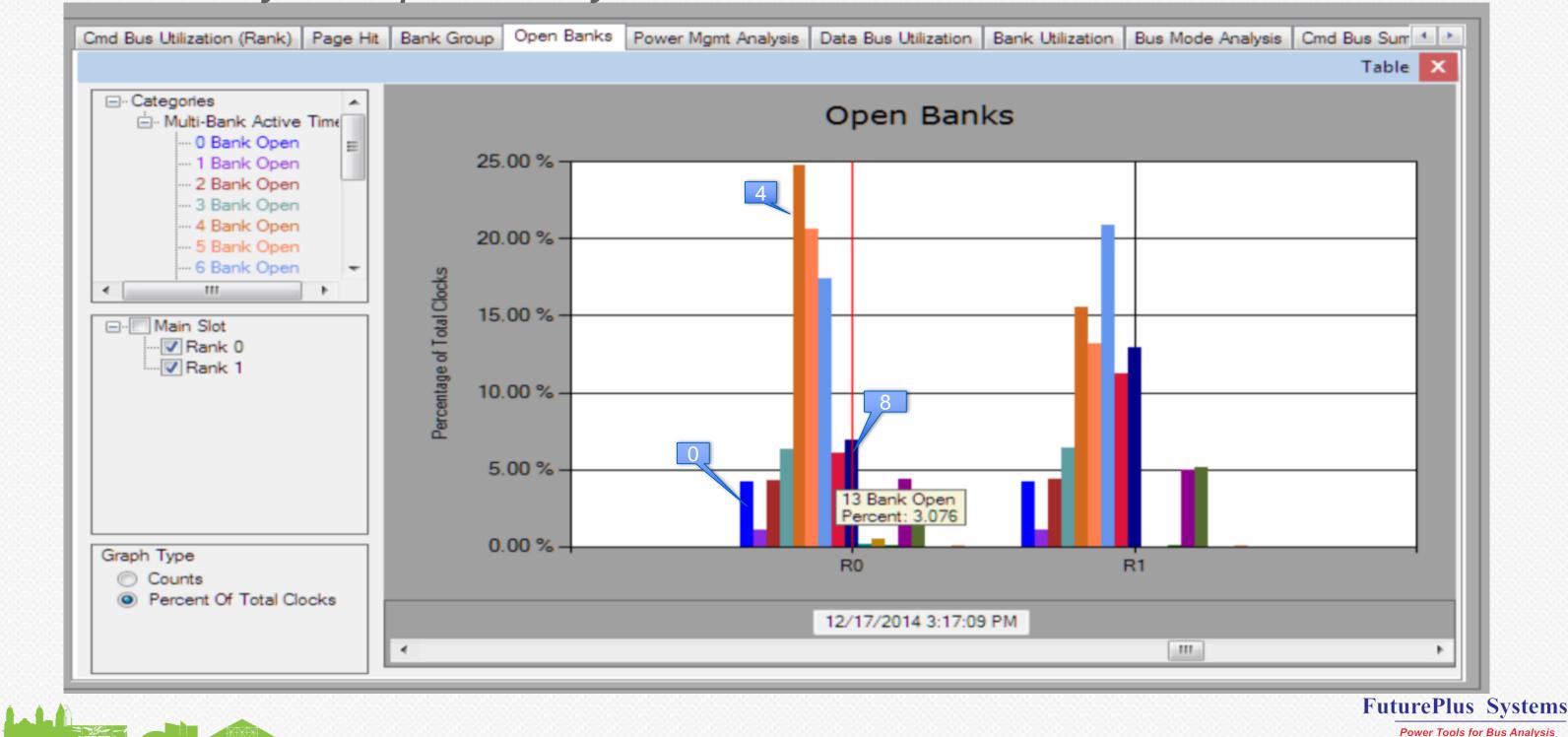
Page Hit



FuturePlus Systems

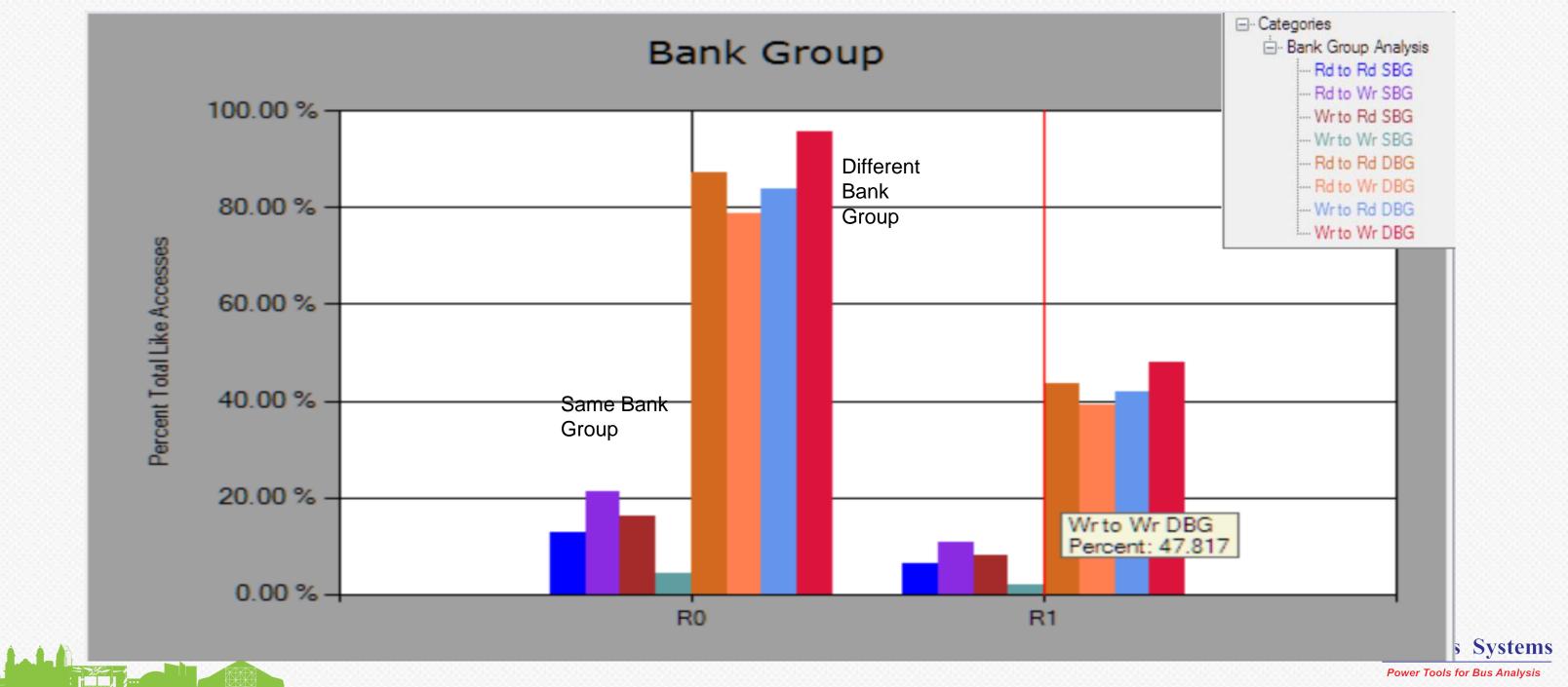
Power Tools for Bus Analysis

Multiple Open Banks How many are open at any one time

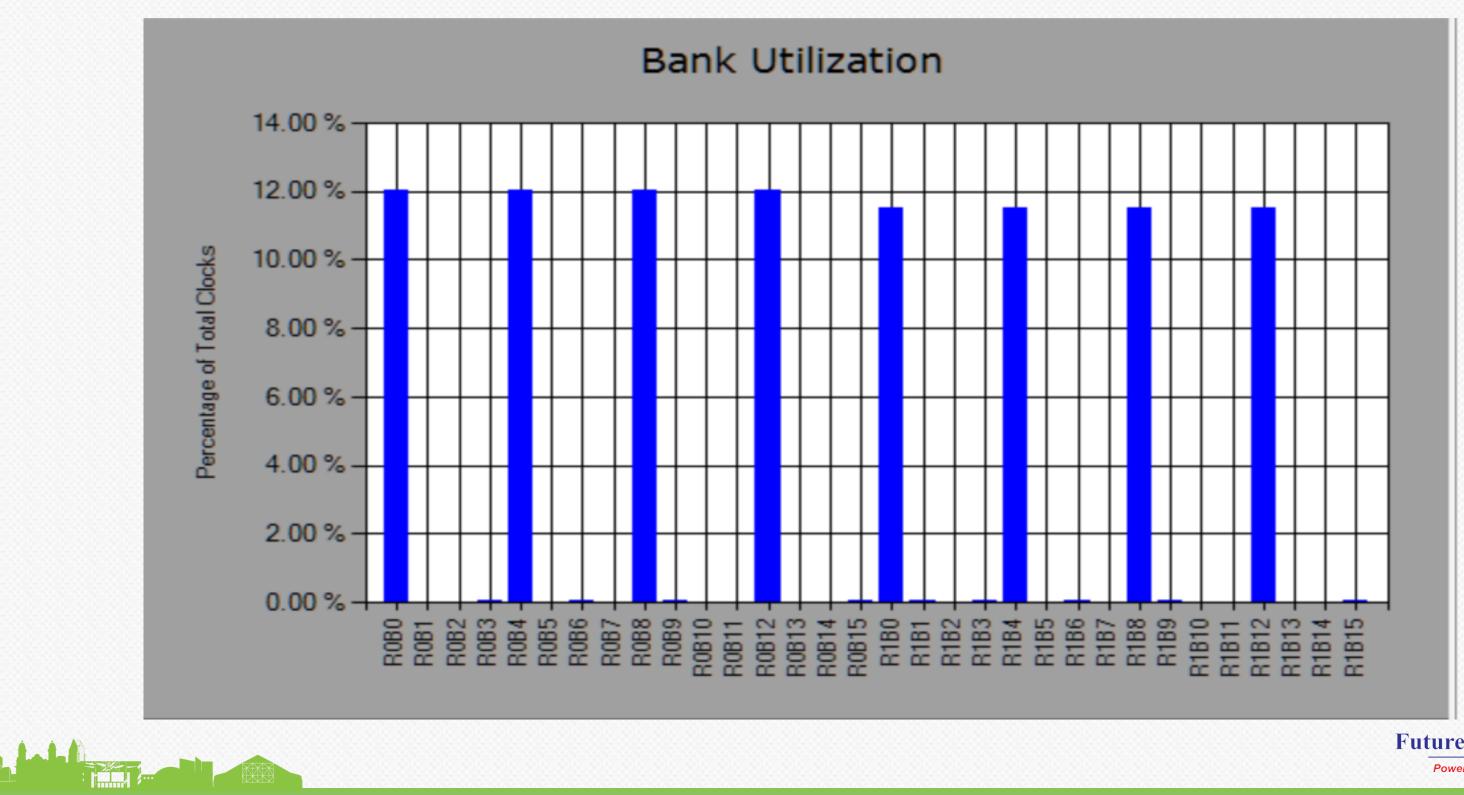


Bank Group Access Analysis

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



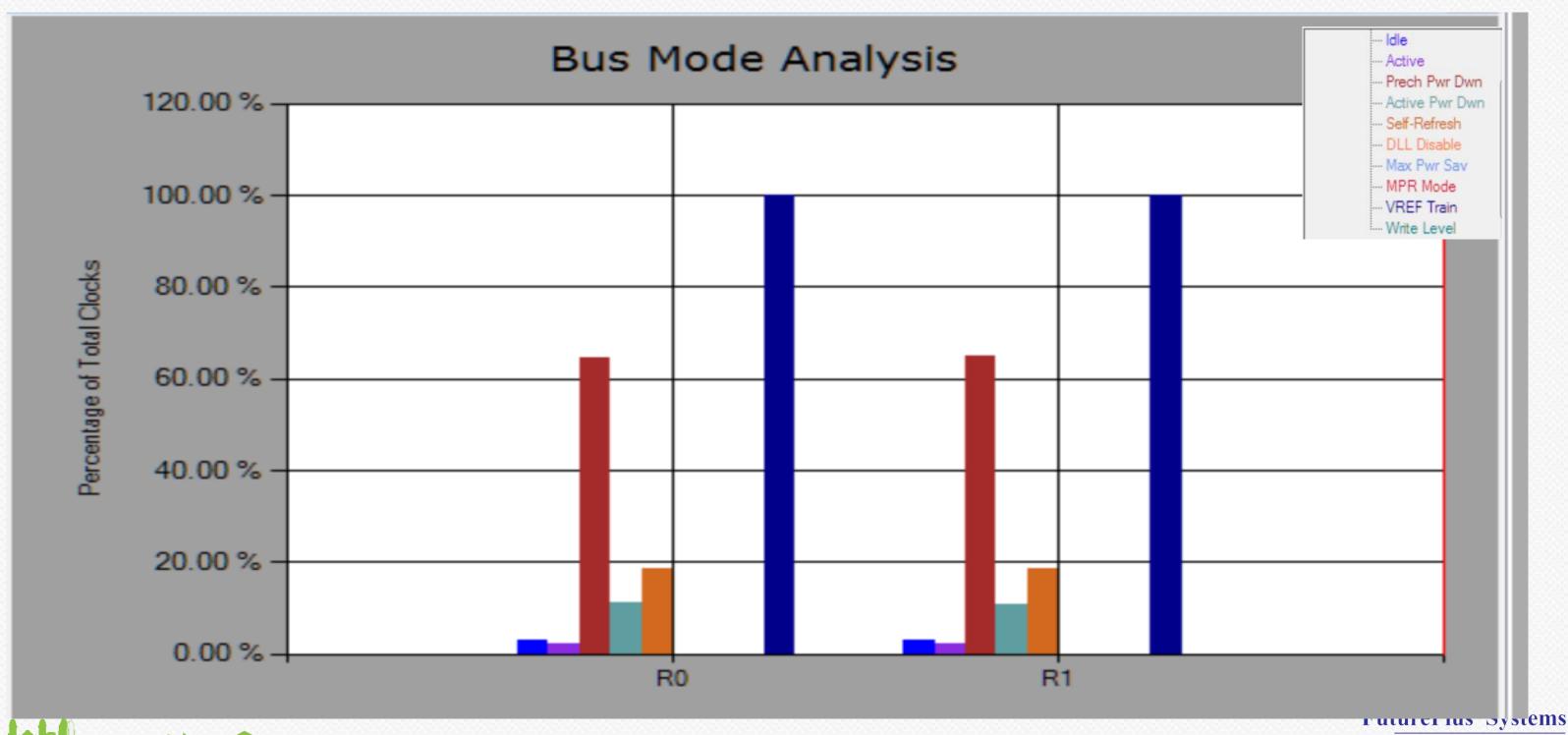
Bank Utilization



FuturePlus Systems

Power Tools for Bus Analysis

Boot Analysis



Power Tools for Bus Analysis

DDR Memory dominates the Data Center Memory power and cooling consumes 16%¹ of the Data

- **Centers Power Budget**
- Memory is 12%² of the Data Centers TCO over a 3 yr period.
- Memory is up to 50% of Server Capital Cost
- Servers are 25% of a Data Centers TCO

1: Source: Samsung 2: Source: "The Data Center as a Computer", by Luiz Barroso, Jimmy Clidaras, and Urs Hölzle (Morgan and Claypool, 2009)

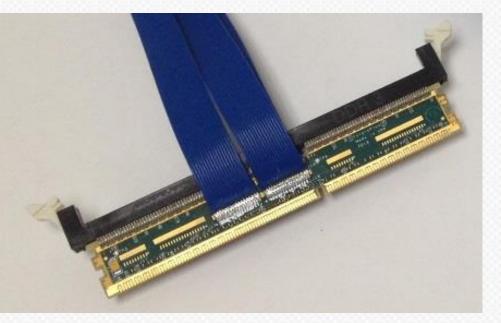
How to Monitor the DDR4 Memory

Use a slot interposer to 'listen' to the traffic between the DIMM and the Memory Controller

- A small amount of current is 'tapped' off the bus
- Only the Address, Command and Control bus needs to be monitored



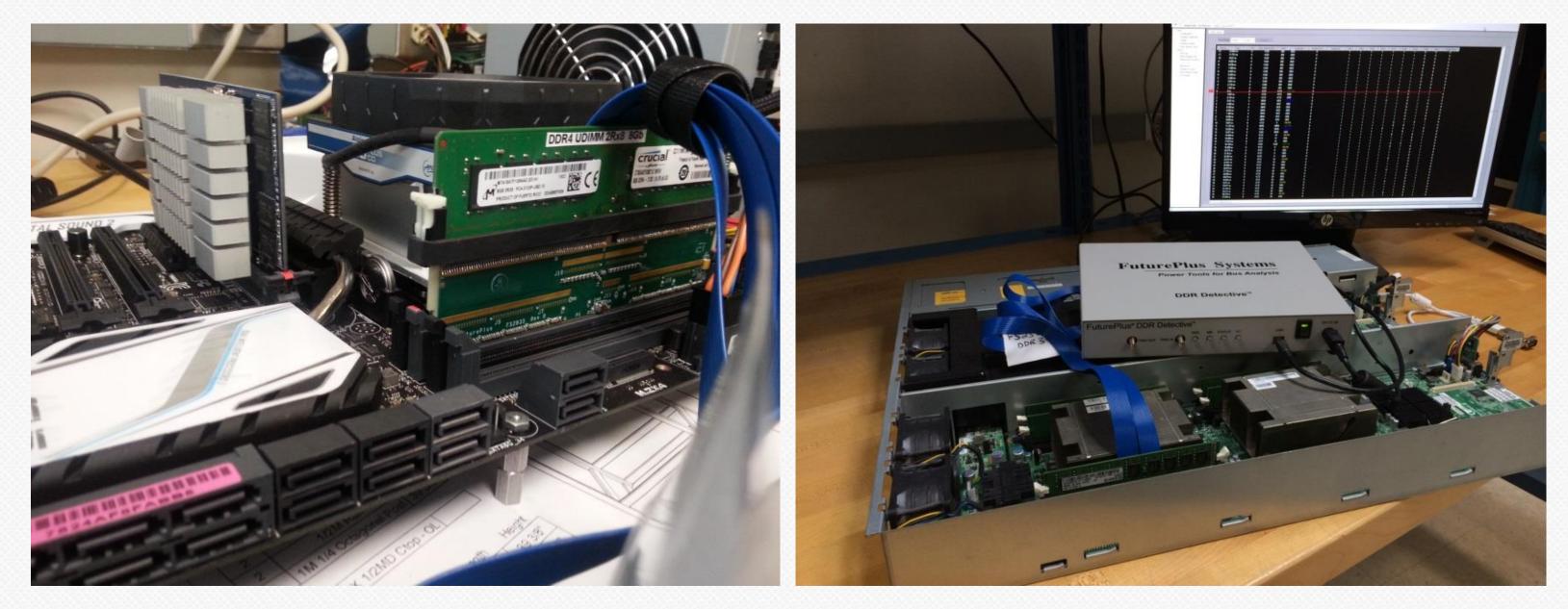






Power Tools for Bus Analysis

The system boots and runs never knowing the equipment is present



FuturePlus Systems

Power Tools for Bus Analysis

Knowledge is King!

Memory Controller/System Architecture

- Can this insight lead to better designs?
- Benchmark Servers Memory Performance

Workload Analysis

- Should the Memory Controller settings be based on criteria set by the workload?
- Can compilers be made better?

Do we all need a DDR5?

Work Smarter not Harder and understand what we have





Power Tools for Bus Analysis

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

FuturePlus Systems

Power Tools for Bus Analysis