

PCI



SIG[®]



Performance of PCI Express® devices – a case study

Roland Scherzinger
Technical Marketing Engineer
Agilent Technologies



Agenda

- Overview
- Definition of performance parameters
- Test methods and setup
- Measurement results
- Summary

Overview

- Using PCI Express does not automatically optimize performance.
- PCI Express parameters such as TLP size, availability of flow control credits and latencies have a strong influence on the overall performance.
- Due to the number of input factors it's hard to predict the actual performance of a new device under real-life conditions.

Goal of this presentation

- Define performance parameters like Throughput, Efficiency and Utilization.
- Show how TLP size, availability of flow control credits and latencies affect the overall performance.
- Analyze the PCI Express performance of various systems.
- Present potential improvement paths for increasing the overall performance.

A first estimate

- What is the maximum Throughput you can get on a x1 link for read completions ?
 - ✓ The requester is able to accept completion packets at line rate.
 - ✓ The completer is able to send completion packets at line rate.
 - ✓ The completer splits the completion at each 64 byte Read Completion Boundary
 1. 240 MB/s ?
 2. 210 MB/s ?
 3. 190 MB/s ?
 4. 170 MB/s ?

Agenda

- Overview
- **Definition of performance parameters**
- Test methods and setup
- Measurement results
- Summary

Throughput, Utilization and Efficiency

- Throughput
 - ✓ Transferred data in bytes / second
- Utilization
 - ✓ Percentage of link active bytes / total bytes
- Efficiency
 - ✓ Percentage of payload bytes / link active bytes
- Actual Throughput =
Maximum Throughput * Utilization * Efficiency

Latencies

- Request to Completion

- ✓ Time between read request and first completion
- ✓ Time between read request and last completion

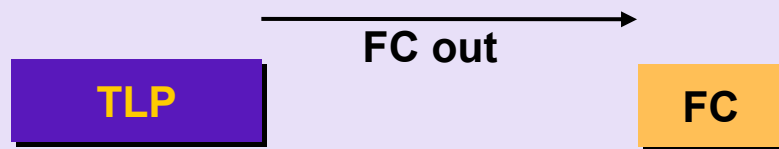


- Latencies that may decrease utilization

- ✓ TLP to FC Update
- ✓ FC Update to TLP
- ✓ TLP to NAK / ACK
- ✓ NAK to Replay / ACK to replay buffer free

Latencies that may decrease utilization

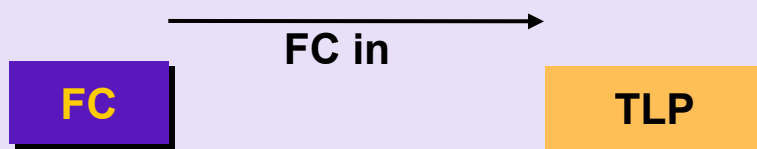
- TLP to FC Update (FC out)
 - ✓ Distance between end of TLP and the FC update that gives back the credits that were used by the TLP



- FC Update to TLP
- TLP to NAK / ACK
- NAK to Replay / ACK to replay buffer free

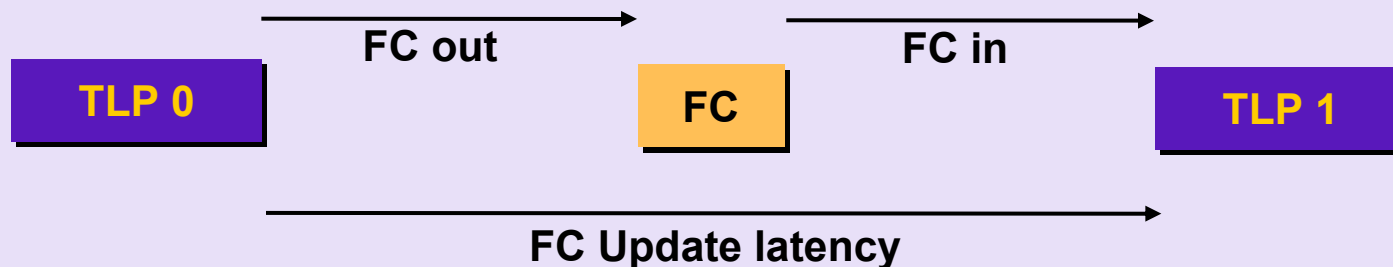
Latencies that may decrease utilization

- TLP to FC Update
- FC Update to TLP (FC in)
 - ✓ Time between reception of a FC Update until the next TLP (that was waiting for credits) is sent



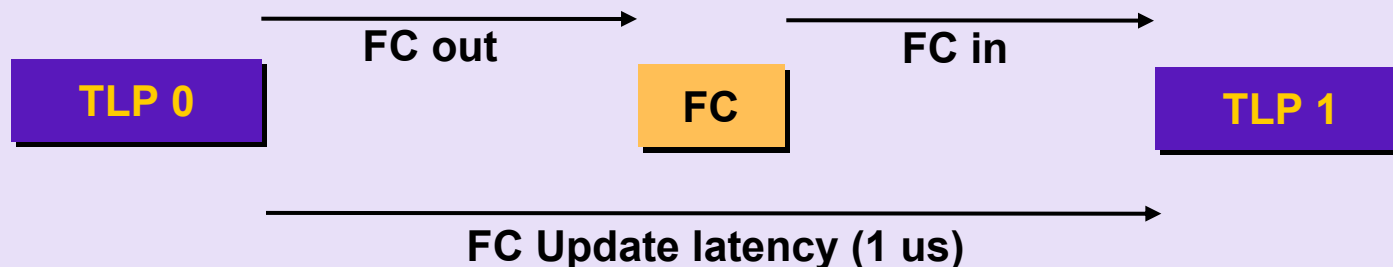
- TLP to NAK / ACK
- NAK to Replay / ACK to replay buffer free

FC Update latency



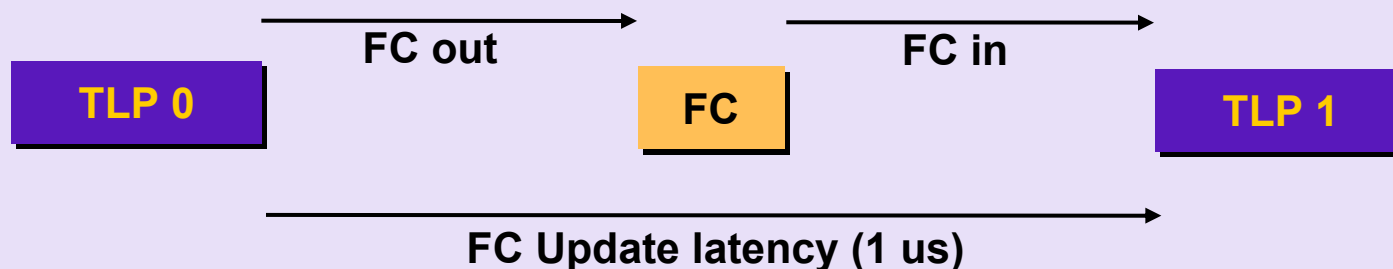
- A device may be forced to wait for credits
- Depending on link speed and number of credits

FC Update latency (x1)



- Need credits for sending 250 bytes (250 MB/s / 1 us, header and payload) in order to achieve full line rate
- Less credits will result in lower Utilization

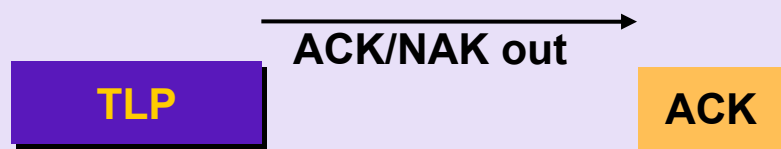
FC Update latency (X8)



- Need credits for sending 2KB (2 GB/s / 1 us, header and payload) in order to achieve full line rate
- Less credits will result in lower Utilization

Latencies that may decrease utilization

- TLP to FC Update
- FC Update to TLP
- TLP to NAK / ACK (ACK/NAK out)
 - ✓ Time between end of TLP and the ACK DLLP that acknowledges the TLP



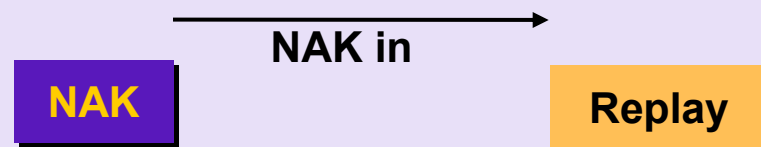
- NAK to Replay / ACK to replay buffer free

Latencies that may decrease utilization

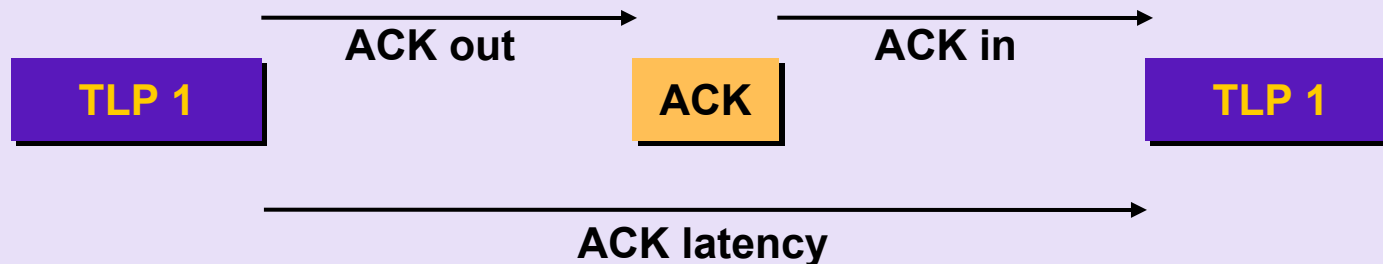
- TLP to FC Update
- FC Update to TLP
- TLP to NAK / ACK
- NAK to Replay / ACK to replay buffer free
 - ✓ Time from reception of the ACK DLLP until the receive buffer is freed



- ✓ Time from reception of the NAK DLLP until the TLP is replayed



ACK Latency



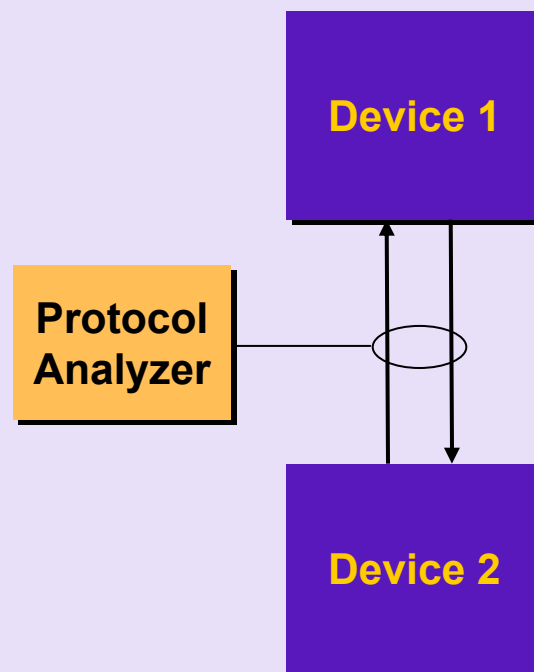
- A device may be forced to wait for replay buffer space depending on link speed and replay buffer size

Agenda

- Overview
- Definition of Terms
- **Test methods and setup**
- Measurement results
- Summary

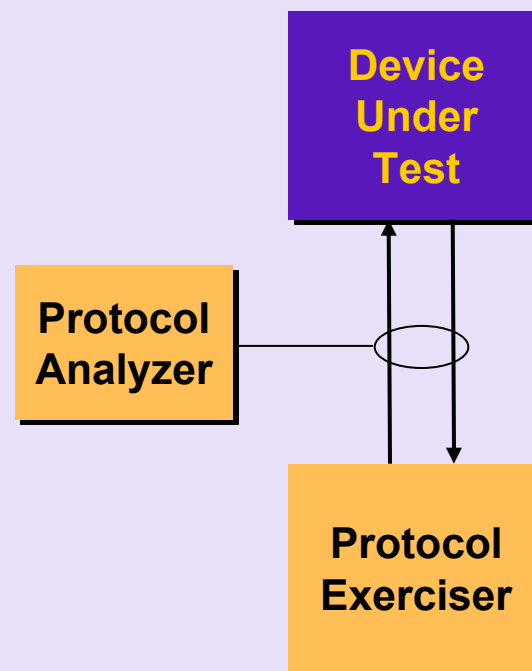
Throughput, Efficiency and Utilization

- Measure real-time actual performance numbers with a protocol analyzer
 - ✓ Does not evaluate the maximum capabilities of Device 1 or Device 2



Measure maximum performance values

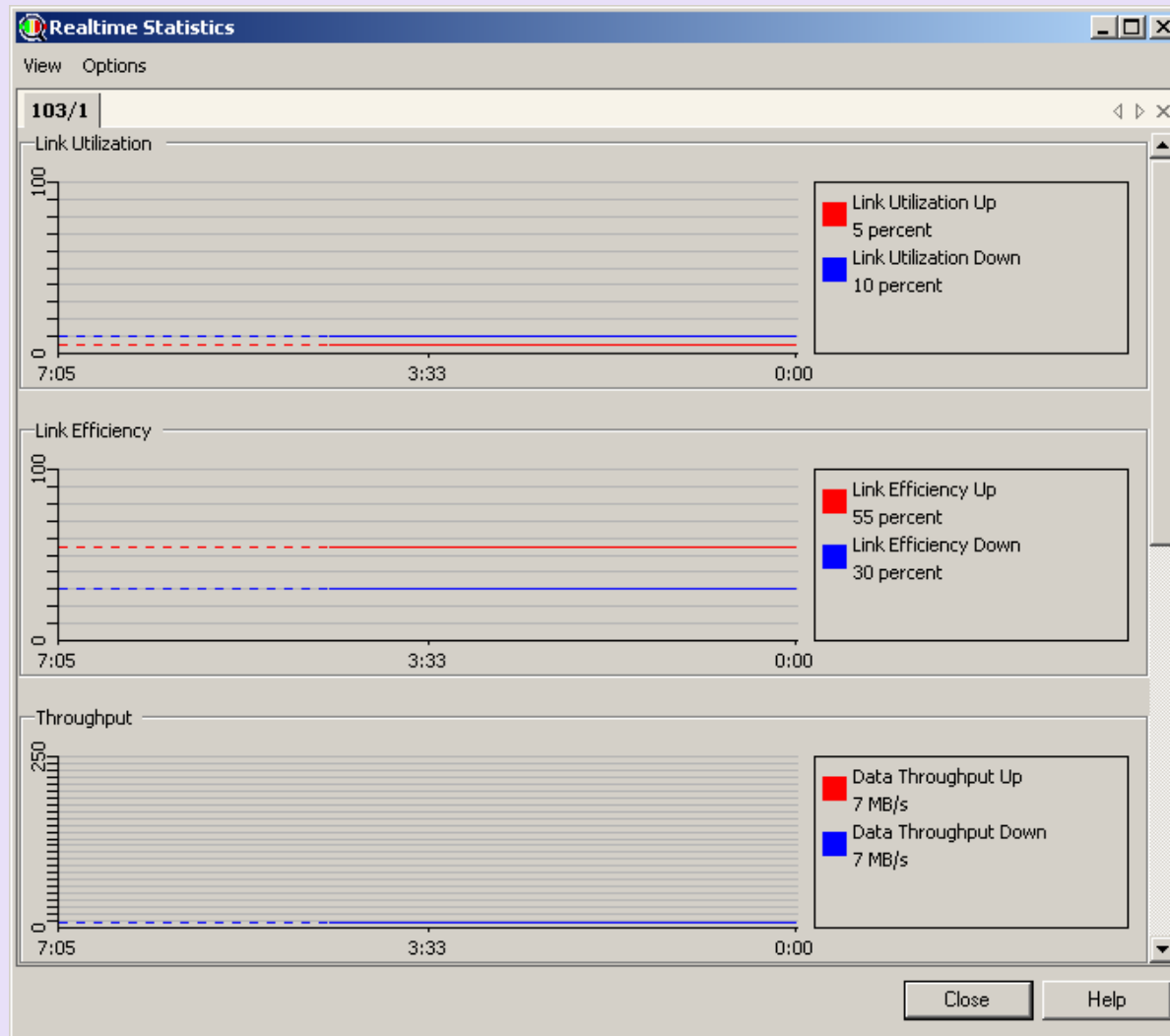
- Use the protocol exerciser as ideal stimulus
- Use the protocol analyzer to measure the actual performance numbers



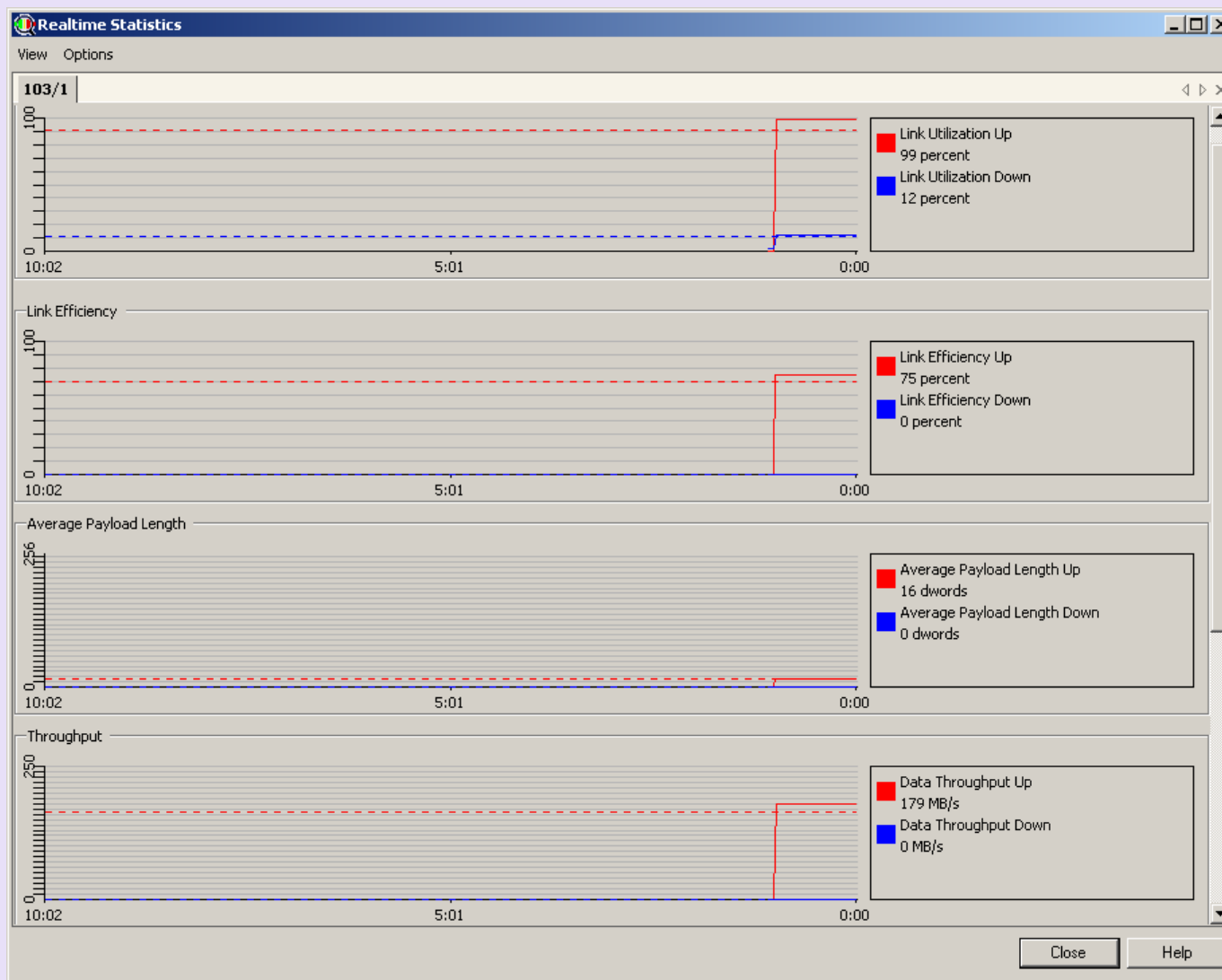
Agenda

- Overview
- Definition of Terms
- Test methods and setup
- **Measurement results**
- Summary

Actual Performance on a x1 link



Maximum Completion Throughput (x1 link)



FC out latency (posted)

624ns – 240ns: 384 ns

Protocol Analyzer - System Protocol Tester

File Edit View Capture Help

100%

103 to 102 = 624.000 ns

Port Overview

D	R	Port	Speed	Name	Link	Records	Trigger	Activity
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	103/1	2.5 Gb/s	Up Down	OK OK	19890	AGT_SEQUENCE...	

103/1

	Channel	Record...	Timestamp	Size...	Seque...	Type	Address	Tag	Data
+	Down	4974	86.469.348.772	4		Skip Ordered Set			
+	Up	4975	86.469.352.144	4		Skip Ordered Set			
+	Up	4976	86.469.352.528	8		UpdateFC-P			
+	Up	4977	86.469.352.608	8		UpdateFC-NP			
+	Up	4978	86.469.352.640	8		UpdateFC-Cpl			
+	Down	4979	86.469.353.596	4		Skip Ordered Set			
M1	+	Up	4980	86.469.354.712	60	8 B6	Memory Write	00 10 00 00	DE 71 FB 53 ED 6D 77 78 DE 7F FE EA F...
	+	Up	4981	86.469.354.952	60	8 B7	Memory Write	00 10 00 00	DE 71 FB 53 ED 6D 77 78 DE 7F FE EA F...
	+	Up	4982	86.469.355.192	60	8 B8	Memory Write	00 10 00 00	DE 71 FB 53 ED 6D 77 78 DE 7F FE EA F...
	+	Down	4983	86.469.355.304	8	8 B6	Ack		
M2	+	Down	4984	86.469.355.336	8		UpdateFC-P		
	+	Up	4985	86.469.355.432	60	8 B9	Memory Write	00 10 00 00	DE 71 FB 53 ED 6D 77 78 DE 7F FE EA F...
	+	Down	4986	86.469.355.544	8	8 B7	Ack		
	+	Down	4987	86.469.355.576	8		UpdateFC-P		
	+	Down	4988	86.469.355.784	8	8 B8	Ack		
	+	Down	4989	86.469.355.816	8		UpdateFC-P		
	+	Down	4990	86.469.356.024	8	8 B9	Ack		
	+	Down	4991	86.469.356.056	8		UpdateFC-P		
	+	Up	4992	86.469.356.648	8		UpdateFC-P		
	+	Up	4993	86.469.356.728	8		UpdateFC-NP		
	+	Up	4994	86.469.356.760	8		UpdateFC-Cpl		
	+	Up	4995	86.469.356.976	4		Skip Ordered Set		

Record Decode

fb 08 b6 40
00 00 0a 03
00 00 ff 00
10 00 00 de
71 fb 53 ed
6d 77 78 de
7f fe ea f...

StartTag = FB
Reserved = 0000
Sequence Number = 8 B6
Reserved = 0
Fmt = 10
Type = 00000

Stopped Online

FC in latency (completion)

432ns

Protocol Analyzer - System Protocol Tester

File Edit View Capture Help

▶ ◀ T [Icons] 100% [Icons]

M1 to M2 = 432.000 ns

Port Overview

D	R	Port	Speed	Name	Link	Records	Trigger	Activity
✓	✓	103/1	2.5 Gb/s	Up Down	OK OK	19811	AGT_SEQUENCE...	

103/1

	Channel	Record...	Timestamp	Size...	Seque...	Type	Address	Tag	Data
⊕	Down	5095	6.665.343.380	4		Skip Ordered Set			
⊕	Up	5096	6.665.344.432	4		Skip Ordered Set			
⊕	Up	5097	6.665.347.080	8		UpdateFC-P			
⊕	Up	5098	6.665.347.160	8		UpdateFC-NP			
M1	Up	5099	6.665.347.192	8		UpdateFC-Cpl			
M2	Down	5100	6.665.347.624	84	0 02	Completion with...		1F	DE 71 FB 53 ED 6D 77 78 DE 7F FE EA F...
⊕	Down	5101	6.665.348.204	4		Skip Ordered Set			
⊕	Up	5102	6.665.348.872	8	0 02	Ack			
⊕	Up	5103	6.665.349.248	4		Skip Ordered Set			
⊕	Up	5104	6.665.351.192	8		UpdateFC-P			
⊕	Up	5105	6.665.351.272	8		UpdateFC-NP			
⊕	Up	5106	6.665.351.304	8		UpdateFC-Cpl			
⊕	Down	5107	6.665.351.736	84	0 03	Completion with...		1F	DE 71 FB 53 ED 6D 77 78 DE 7F FE EA F...
⊕	Up	5108	6.665.352.984	8	0 03	Ack			
⊕	Down	5109	6.665.353.020	4		Skip Ordered Set			
⊕	Down	5110	6.665.353.428	8		UpdateFC-P			
⊕	Down	5111	6.665.353.460	8		UpdateFC-NP			
⊕	Up	5112	6.665.354.064	4		Skip Ordered Set			
⊕	Up	5113	6.665.355.296	8		UpdateFC-P			
⊕	Up	5114	6.665.355.376	8		UpdateFC-NP			
⊕	Up	5115	6.665.355.408	8		UpdateFC-Cpl			
⊕	Down	5116	6.665.355.840	84	0 04	Completion with...		1F	2C 0D BE 0D C3 0D CD 0D D3 0D 34 09 D...

Record Decode

5c a0 00 80
14 42 7b fd

StartTag = 5C
Type = 10100
Virtual Channel = 0
Reserved = 00
HdrFC = 02
Reserved = 00

Stopped Online

Performance improvement ways

- Efficiency:
 - ✓ Maximize payload size
 - ✓ Minimize overhead: Avoid unnecessary DLLP's
- Utilization:
 - ✓ Minimize FC latencies
 - ✓ Maximize (FC) buffer sizes
 - ✓ Minimize ACK latencies
 - ✓ Maximize retry buffer size

Agenda

- Overview
- Definition of Terms
- Test methods and setup
- Measurement results
- **Summary**

Summary

- PCI Express parameters such as TLP size, availability of flow control credits and latencies have a strong influence on the overall performance.
- Both sides of a link are influencing the overall performance.
- For corner case measurements an ideal stimulus is required.

Thank you for attending the
PCI-SIG Developers Conference 2005.

For more information please go to
www.pcisig.com



Performance of PCI Express devices – a case study

Roland Scherzinger
Technical Marketing Engineer
Agilent Technologies



PCI



SIG[®]