



Deadlock Avoidance in PCI Express[®] Based Architectures

Isaac Livny
Applications Engineer
LSI Corporation



Agenda

- Deadlock issues in bus architectures
- Deadlocks in PCI Express® architectures
- Deadlocks at common resource level
 - ✓ Example: ASPM L1 re-entry delay
- Deadlock at RC/chipset level
 - ✓ Example: PME fence mechanism hang-up
- Deadlock at device level – EP or RC
 - ✓ Example 1: Incompliant NAK
 - ✓ Example 2: CLKREQ# readiness
- Deadlock at PCIe® link level
 - ✓ Example: Credit leakage

Agenda

- Deadlock issues in bus architectures
- Deadlocks in PCI Express architectures
- Deadlocks at common resource level
 - ✓ Example: ASPM L1 re-entry delay
- Deadlock at RC/chipset level
 - ✓ Example: PME fence mechanism hang-up
- Deadlock at device level – EP or RC
 - ✓ Example 1: Incompliant NAK
 - ✓ Example 2: CLKREQ# readiness
- Deadlock at PCIe link level
 - ✓ Example: Credit leakage

General

- Deadlock avoidance crucial for every bus architecture
- Occur when multiple agents mutually depend on response from each other or common resource
 - ✓ Access to a resource cannot complete until other agent has completed its access to the same resource
- Conventional PCI back-off schemes
 - ✓ Disconnect
 - ✓ Retry
- PCI Express replaces retry and disconnect termination schemes with split transaction protocols.
 - ✓ Complexity of protocols can create more deadlock pitfalls

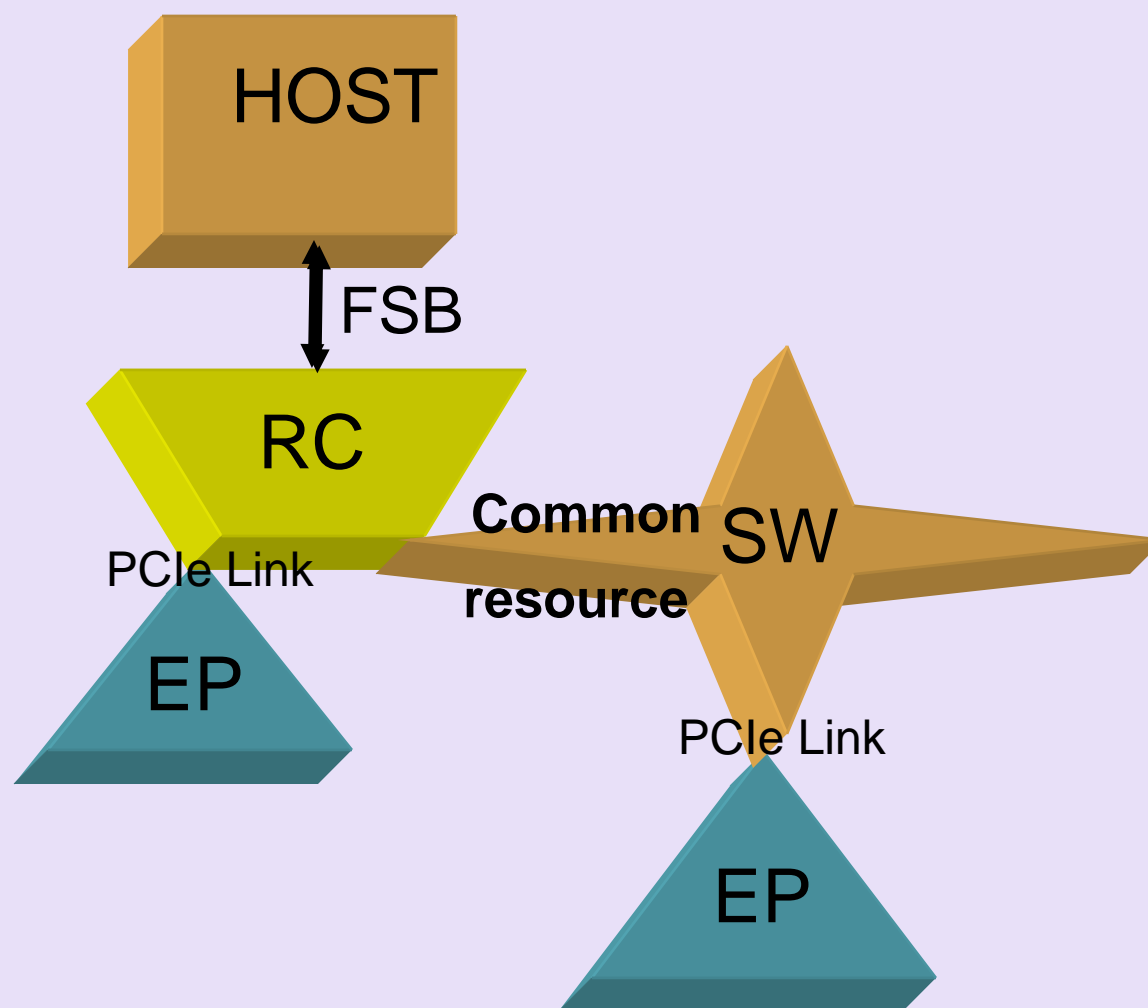
Agenda

- Deadlock issues in bus architectures
- Deadlocks in PCI Express architectures
- Deadlocks at common resource level
 - ✓ Example: ASPM L1 re-entry delay
- Deadlock at RC/chipset level
 - ✓ Example: PME fence mechanism hang-up
- Deadlock at device level – EP or RC
 - ✓ Example 1: Incompliant NAK
 - ✓ Example 2: CLKREQ# readiness
- Deadlock at PCIe link level
 - ✓ Example: Credit leakage

Deadlock scenarios

- 4 types of deadlock scenarios
 - characterized by where they can occur in the PCI Express fabric
 - ✓ Within an endpoint
 - ✓ Associated with a particular PCI Express link
 - ✓ Associated with a resource common to a number of PCI express devices or within a PCI express switch
 - ✓ At the Root Complex level, within the chipset or at the system level

PCI Express Fabric



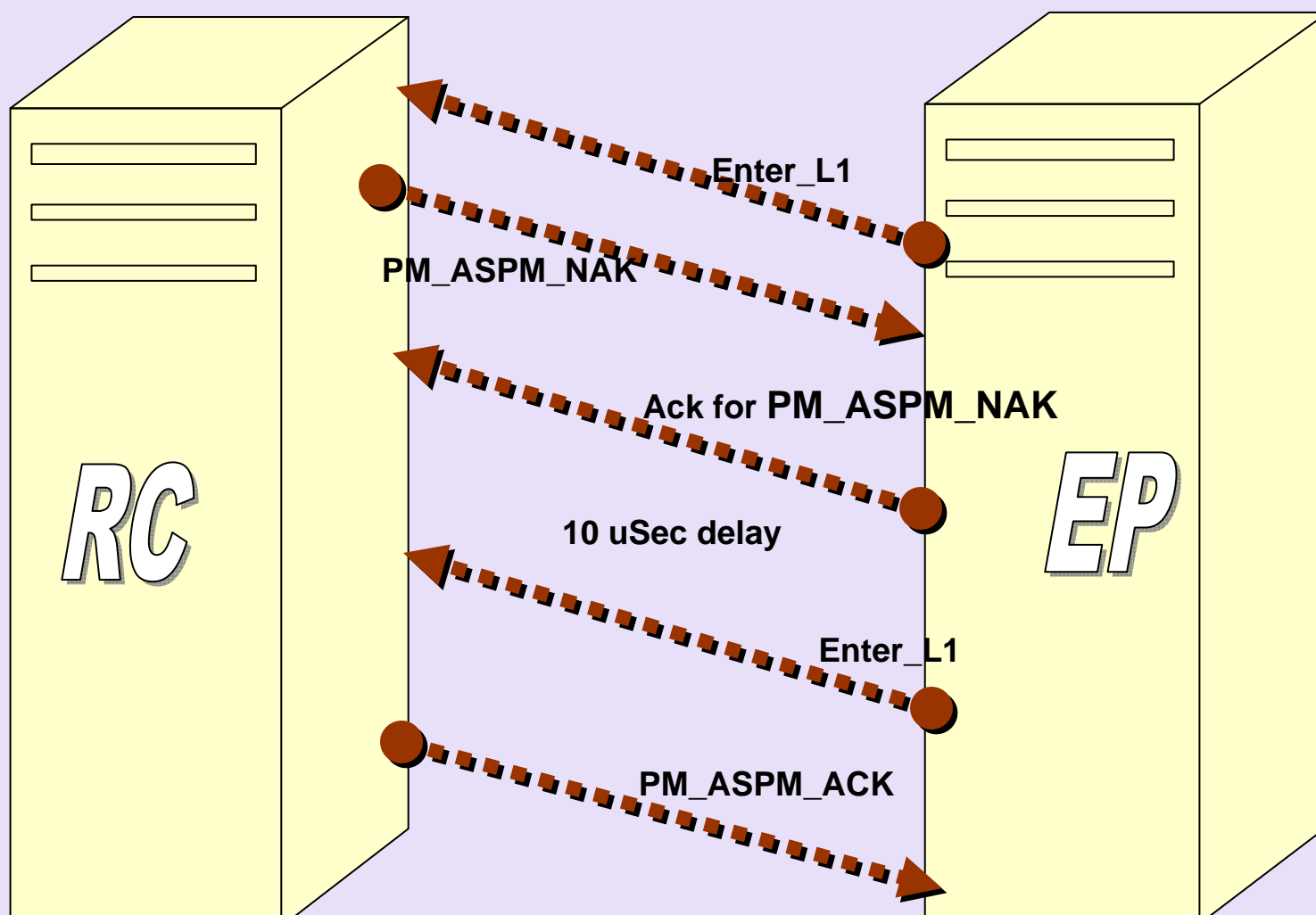
Deadlocks and avoidance

- PCIe protocols affected
 - ✓ ASPM L1 entry and exit protocols
 - ✓ PME fence mechanism
 - ✓ Error reporting schemes
 - ✓ Credit handling for special cases
 - ✓ Clock power management readiness
- Deadlock avoidance schemes
 - ✓ At the device level
 - ✓ Root Complex level
 - ✓ System level
 - ✓ Possible specification enhancements

Agenda

- Deadlock issues in bus architectures
- Deadlocks in PCI Express architectures
- Deadlocks at common resource level
 - ✓ Example: ASPM L1 re-entry delay
- Deadlock at RC/chipset level
 - ✓ Example: PME fence mechanism hang-up
- Deadlock at device level – EP or RC
 - ✓ Example 1: Incompliant NAK
 - ✓ Example 2: CLKREQ# readiness
- Deadlock at PCIe link level
 - ✓ Example: Credit leakage

ASPM L1 Re-entry delay



Deadlock example at common resource level

ASPM L1 re-entry delay deadlock

Packet	R→	2.5	TLP	Cpl	CplID	RequesterID	Tag	CompleterID	Status	BCM	Byte Cnt	Lwr Addr	Register Data
140713	R→	x1	1138		10:01010	000:00:0	1	004:00:0	SC	0	4	0x00	0x10110042
Packet	R→	2.5	DLLP	ACK	AdkNak Seq Num	CRC 16	Time Delta						
140714	R→	x1			1138	0x14AD	536.000 ns						
Packet	R→	2.5	DLLP	UpdateFC-NP	VC ID	HdrFC	DataFC	CRC 16	Time Delta				
140716	R→	x1			0	105		16-bit CRC	8.568 μs				
Packet	R→	2.5	DLLP	UpdateFC-P	VC ID	HdrFC	DataFC	CRC 16	Idle				
140719	R→	x1			0	68	81	0x6285	0.000 ns				
Packet	R→	2.5	DLLP	UpdateFC-NP	VC ID	HdrFC	DataFC	CRC 16	Time Delta				
140720	R→	x1			0	105	146	0x726C	2.704 μs				
Packet	R→	2.5	DLLP	UpdateFC-NP	VC ID	HdrFC	DataFC	CRC 16	Idle				
140723	R→	x1			0	68	0	0x2D8B	96.000 ns				
Packet	R→	2.5	DLLP	UpdateFC-P	VC ID	HdrFC	DataFC	CRC 16	Time Delta				
140724	R→	x1			0	75	136	0xFFFF1	13.768 μs				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140729	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140730	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140731	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140732	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140733	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140734	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140735	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140740	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140741	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140740	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140741	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140740	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140741	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140740	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140741	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140740	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140741	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140740	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140741	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140740	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140741	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140740	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140741	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140740	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140741	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140740	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140741	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140740	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140741	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140740	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140741	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta							
140736	R→	x1			0xEB05	24.000 ns							
Packet	R→	2.5	TLP	Msg	Msg	Msg Routing	RequesterID	Tag	Message Code				
140737	R→	x1	1191		01:10100	Local	000:00:0	0	PME_Active_State_Nak				
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140738	R→	x1			0xEB05	0.000 ns							
Packet	R→	2.5	DLLP	PM_Active_State_Request_L1	CRC 16	Idle							
140739	R→	x1			0xEB05	0.000 ns							
Packet	R→												

L1 Re-entry delay pitfall

- ASPM L1 enabled at endpoint, disabled at root port
 - ✓ PCIe 1.1 Requires correct software sequencing
 - ✓ Used to force PM_Active_State_NACK
- Traffic recess – EP requests L1, Root port NAKs
- EP Data Link Layer Acks the NACK message
- EP required ASPM L1 re-entry delay > 10 usec
- If re-entry is violated – RC interprets as old request
- RC resources congestion blocks additional traffic
- Deadlock situation arises.
- If RC low on resources deadlock can also happen when re-entry delay is within spec
- RC should flush L1 requests not to consume resources

Agenda

- Deadlock issues in bus architectures
- Deadlocks in PCI Express architectures
- Deadlocks at common resource level
 - ✓ Example: ASPM L1 re-entry delay
- Deadlock at RC/chipset level
 - ✓ Example: PME fence mechanism hang-up
- Deadlock at device level – EP or RC
 - ✓ Example 1: Incompliant NAK
 - ✓ Example 2: CLKREQ# readiness
- Deadlock at PCIe link level
 - ✓ Example: Credit leakage

D3 Cold Transition Protocol

- System software puts device to D3
- Endpoint initiates L1 transition
- L1 exit, Link transitions back to L0
- System software/RC initiate PME fence mechanism
- Endpoint Acks PME Shutdown message
- Transition to L2/L3 ready
- L23 entry message ack'ed
- Link goes to electrical idle

D3 Cold Transition Protocol

Packet	R→	2.5	TS2	COM	Link	Lane	N_FTS	Training Control	Data Rate	TS2 Symbols	Idle				
1484949		x1		K28.5	0	0	24	0 0 0 0	2.5 GT/s	D05.2 ...	0.000 ns				
Packet	R→	2.5	TS2	COM	Link	Lane	N_FTS	Training Control	Data Rate	TS2 Symbols	Idle				
1484950		x1		K28.5	0	0	24	0 0 0 0	2.5 GT/s	D05.2 ...	0.000 ns				
Packet	R→	2.5	TS2	COM	Link	Lane	N_FTS	Training Control	Data Rate	TS2 Symbols	Idle				
1484951		x1		K28.5	0	0	24	0 0 0 0	2.5 GT/s	D05.2 ...	0.000 ns				
Packet	R→	2.5	TS2	COM	Link	Lane	N_FTS	Training Control	Data Rate	TS2 Symbols	Idle				
1484952		x1		K28.5	0	0	24	0 0 0 0	2.5 GT/s	D05.2 ...	180.000 ns				
Packet	R→	2.5	TLP	Cfg	CfgRd0	RequesterID	Tag	DeviceID	Register	1st BE	LCRC	Time Delta			
1484953		x1	2815	00:00100	000:00:0	1	002:00:0	0x044	0011	0xB1E45B7D	292.000 ns				
Packet	R←	2.5	TLP	Cpl	CplID	RequesterID	Tag	CompleterID	Status	BCM	Byte Cnt	Lwr Addr	Register Data	LCRC	
1484955		x1	1806	10:01010	000:00:0	1	002:00:0	SC	0	4	0x00	0x00000003	0x341B74A8		
Packet	R→	2.5	TLP	Cfg	CfgRd0	RequesterID	Tag	DeviceID	Register	1st BE	LCRC	Time Delta			
1484957		x1	2816	00:00100	000:00:0	0	002:00:0	0x004	0011	0x541B3A60	324.000 ns				
Packet	R←	2.5	TLP	Cpl	CplID	RequesterID	Tag	CompleterID	Status	BCM	Byte Cnt	Lwr Addr	Status	Command	LCRC
1484959		x1	1807	10:01010	000:00:0	0	002:00:0	SC	0	4	0x00	0x0010	0x0106	0x868B509	
Packet	R→	2.5	TLP	Cfg	CfgWr0	RequesterID	Tag	DeviceID	Register	1st BE	Status	Command	LCRC	Time Delta	
1484961		x1	2817	10:00100	000:00:0	1	002:00:0	0x004	0011	0x0000	0x0100	0xE9889B5A	344.000 ns		
Packet	R←	2.5	TLP	Cpl	Cpl	RequesterID	Tag	CompleterID	Status	BCM	Byte Cnt	Lwr Addr	LCRC	Time Delta	
1484963		x1	1808	00:01010	000:00:0	1	002:00:0	SC	0	4	0x00	0x400DF844	5.389 sec		
Packet	R→	2.5	TLP	Meg	Meg	Meg Routing	RequesterID	Tag	Message Code	LCRC	Time Delta				
1484965		x1	2818	01:10011	Broadcast	000:00:0	0	PME_Turn_Off	0xDF294396	248.000 ns					
Packet	R←	2.5	TLP	Meg	Meg	Meg Routing	RequesterID	Tag	Message Code	LCRC	Idle				
1484967		x1	1809	01:10101	Gathered	002:00:0	0	PME_TO_Ack	0xD14D2D5F	0.000 ns					
Packet	R←	2.5	TLP	Meg	Meg	Meg Routing	RequesterID	Tag	Message Code	LCRC	Idle				
1484968		x1	1810	01:10101	Gathered	002:00:0	0	PME_TO_Ack	0x5516B70C	456.000 ns					
Packet	R←	2.5	DLLP	PM_Enter_L23	CRC 16	Idle									
1484970		x1		0x1055	0.000 ns										
Packet	R←	2.5	DLLP	PM_Enter_L23	CRC 16	Idle									
1484971		x1		0x1055	0.000 ns										

■ L1 exit, Link back to L0

■ PME mechanism

■ L2/L3 ready trans

■ L1 exit, Link back to L0

■ PME Fence mechanism

■ L2/L3 ready transition

D3 Cold Transition Protocol

- L23 entry message ack'ed
- Link goes to electrical idle

Packet 1484985	R←	2.5 x1	DLLP	PM_Enter_L23	CRC 16 0x1055	Time Delta 0.000 ns
Packet 1484986	R→	2.5 x1	DLLP	PM_Request_Ack	CRC 16 0x930C	Time Delta 32.000 ns
Packet 1484987	R←	2.5 x1	DLLP	PM_Enter_L23	CRC 16 0x1055	Idle 32.000 ns
Packet 1484988	R←	2.5 x1	EIOS	COM K28.5	EIOS Symbols K28.3 K28.3 K28.3	Time Delta 0.000 ns
Packet 1484989	R→	2.5 x1	DLLP	PM_Request_Ack	CRC 16 0x930C	Idle 0.000 ns
Packet 1484991	R→	2.5 x1	DLLP	PM_Request_Ack	CRC 16 0x930C	Idle 88.000 ns
Packet 1484992	R→	2.5 x1	DLLP	PM_Request_Ack	CRC 16 0x930C	Idle 0.000 ns
Packet 1484993	R→	2.5 x1	DLLP	PM_Request_Ack	CRC 16 0x930C	Idle 616.000 ns
Packet 1484996	R→	2.5 x1	EIOS	COM K28.5	EIOS Symbols K28.3 K28.3 K28.3	

Power Sequencing Deadlock Scenario

- Power on
- Wait for the BIOS to kick in
- Toggle main supply/RESET
 - ✓ Typical for mobile systems
- Windows starts
- Disable the device in device manager, or go to Hibernate
- Machine hangs

Observed Sequence

Packet	R→	2.5	TLP	Cfg	CfgWr0	RequesterID	Tag	DeviceID	Register	1st BE	Register Data
4320		x1	2773		10:00100	000:00:0	1	003:00:0	0x044	1111	0x00008003

Packet	R←	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16	Idle
4321		x1			2773	0x3F37	80.000 ns

Packet	R←	2.5	TLP	Cpl	Cpl	RequesterID	Tag	CompleterID	Status	BCM	Byte Cnt	Lv
4322		x1	3107		00:01010	000:00:0	1	003:00:0	SC	0	4	0

Packet	R→	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16	Time Delta
4323		x1			3107	0x8519	216.000 ns

Packet	R←	2.5	DLLP	PM_Enter_L23		CRC 16	Idle
4324		x1				0x1055	0.000 ns

Packet	R←	2.5	DLLP	PM_Enter_L23		CRC 16	Idle
4325		x1				0x1055	0.000 ns

Packet	R←	2.5	DLLP	PM_Enter_L23		CRC 16	Idle
4326		x1				0x1055	0.000 ns

Packet	R←	2.5	DLLP	PM_Enter_L23		CRC 16	Idle
4327		x1				0x1055	0.000 ns

Packet	R←	2.5	DLLP	PM_Enter_L23		CRC 16	Time Delta
4328		x1				0x1055	8.000 ns

Packet	R→	2.5	DLLP	PM_Request_Ack		CRC 16	Time Delta
4329		x1				0x930C	24.000 ns

Observed Sequence

Packet 4338	R→	2.5 x1	DLLP	PM_Enter_L23	CRC 16 0x1055	Time Delta 8.000 ns
Packet 4339	R→	2.5 x1	DLLP	PM_Request_Ack	CRC 16 0x930C	Time Delta 24.000 ns
Packet 4340	R→	2.5 x1	DLLP	PM_Enter_L23	CRC 16 0x1055	Time Delta 8.000 ns
Packet 4341	R→	2.5 x1	DLLP	PM_Request_Ack	CRC 16 0x930C	Idle 0.000 ns
Packet 4342	R→	2.5 x1	DLLP	PM_Request_Ack	CRC 16 0x930C	Time Delta 24.000 ns
Packet 4343	R→	2.5 x1	EIOS	COM K28.5	EIOS Symbols K28.3 K28.3 K28.3	Time Delta 8.000 ns
Packet 4344	R→	2.5 x1	DLLP	PM_Request_Ack	CRC 16 0x930C	Time Delta 24.000 ns
Packet 4345	R→	2.5 x1	Link Event Link Down	Time Delta 8.000 ns		
Packet 4346	R→	2.5 x1	DLLP	PM_Request_Ack	CRC 16 0x930C	Idle 0.000 ns
Packet 4347	R→	2.5 x1	DLLP	PM_Request_Ack	CRC 16 0x930C	Idle 16.000 ns
Packet 4348	R→	2.5 x1	EIOS	COM K28.5	EIOS Symbols K28.3 K28.3 K28.3	Time Delta 824.000 ns

PME Fence Mechanism Deadlock Summary

Packet 4320	R→	2.5	TLP	Cfg	CfgWr0	RequesterID	Tag	DeviceID	Register	1st BE	Register Data
		x1			10:00100	000:00:0	1	003:00:0	0x044	1111	0x00008003
Packet 4321	R→	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16	Idle				
		x1			2773	0x3F37	80.000 ns				
Packet 4322	R→	2.5	TLP	Cpl	Cpl	RequesterID	Tag	CompleterID	Status	BCM	Byte Cnt
		x1			00:01010	000:00:0	1	003:00:0	SC	0	4
Packet 4323	R→	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16	Time Delta				
		x1			3107	0x6519	216.000 ns				
Packet 4324	R→	2.5	DLLP	PM_Enter_L23		CRC 16	Idle				
		x1				0x1055	0.000 ns				
Packet 4325	R→	2.5	DLLP	PM_Enter_L23		CRC 16	Idle				
		x1				0x1055	0.000 ns				
Packet 4326	R→	2.5	DLLP	PM_Enter_L23		CRC 16	Idle				
		x1				0x1055	0.000 ns				
Packet 4327	R→	2.5	DLLP	PM_Enter_L23		CRC 16	Idle				
		x1				0x1055	0.000 ns				
Packet 4328	R→	2.5	DLLP	PM_Enter_L23		CRC 16	Time Delta				
		x1				0x1055	8.000 ns				
Packet 4329	R→	2.5	DLLP	PM_Request_Ack		CRC 16	Time Delta				
		x1				0x930C	24.000 ns				

- Initial boot-power-sequence leaves device at D3-cold ready state

- Device transitions to L23 ready instead of L1

Packet 4338	R→	2.5	DLLP	PM_Enter_L23		CRC 16	Time Delta				
		x1				0x1055	8.000 ns				
Packet 4339	R→	2.5	DLLP	PM_Request_Ack		CRC 16	Time Delta				
		x1				0x930C	24.000 ns				
Packet 4340	R→	2.5	DLLP	PM_Enter_L23		CRC 16	Time Delta				
		x1				0x1055	8.000 ns				
Packet 4341	R→	2.5	DLLP	PM_Request_Ack		CRC 16	Idle				
		x1				0x930C	0.000 ns				
Packet 4342	R→	2.5	DLLP	PM_Request_Ack		CRC 16	Time Delta				
		x1				0x930C	24.000 ns				
Packet 4343	R→	2.5	EIOS	COM	EIOS Symbols	Time Delta					
		x1			K28.5 K28.3 K28.3 K28.3	8.000 ns					
Packet 4344	R→	2.5	DLLP	PM_Request_Ack		CRC 16	Time Delta				
		x1				0x930C	24.000 ns				
Packet 4345	R→	2.5	Link Event	Link Down		Time Delta					
		x1				8.000 ns					
Packet 4346	R→	2.5	DLLP	PM_Request_Ack		CRC 16	Idle				
		x1				0x930C	0.000 ns				
Packet 4347	R→	2.5	DLLP	PM_Request_Ack		CRC 16	Idle				
		x1				0x930C	16.000 ns				
Packet 4348	R→	2.5	EIOS	COM	EIOS Symbols	Time Delta					
		x1			K28.5 K28.3 K28.3 K28.3	824.000 ns					

Power Sequencing Hangup Analysis

- Host tries to communicate while device in D3 cold
- RC blocks transactions since it has transitioned the link to L23
- Machine hangs on no completions coming back from the device
- Power down state machine should not transition to L23 before receiving PME_Turn_Off message and fully executing the PME fence mechanism

Agenda

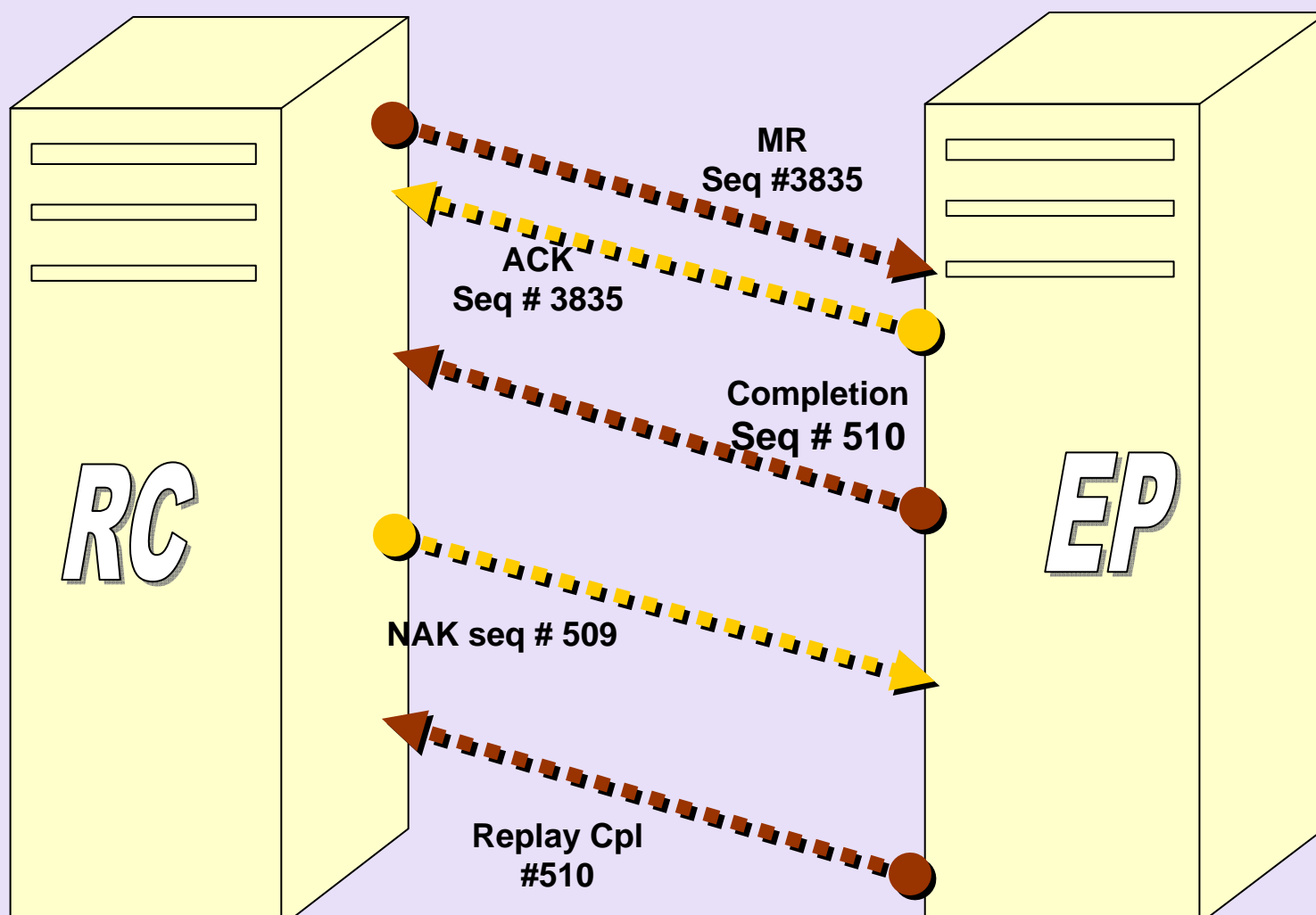
- Deadlock issues in bus architectures
- Deadlocks in PCI Express architectures
- Deadlocks at common resource level
 - ✓ Example: ASPM L1 re-entry delay
- Deadlock at RC/chipset level
 - ✓ Example: PME fence mechanism hang-up
- Deadlock at device level – EP or RC
 - ✓ Example 1: Incompliant NAK
 - ✓ Example 2: CLKREQ# readiness
- Deadlock at PCIe link level
 - ✓ Example: Credit leakage

Deadlock at the Device Level: Handling of Illegitimate NAK

- Compliant NAK
 - ✓ TLP # 510 Nak'd by NAK 509
 - ✓ TLP #510 is replayed

Packet	R→	2.5	TLP	Mem	MRd(32)	TC	TD	EP	Attributes	Length	RequesterID	Tag	Address	1st BE	Last BE	LCRC	Time Delta	
93203		x1	3835		00:00000	0	0	0	01	1	000:00:0	1	FEAFF180	1111	0000	0xBD332CCA	272.000 ns	
Packet	R→	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16		Idle										
93204		x1			3835	0x1139		56.000 ns										
Packet	R→	2.5	TLP	Cpl	CplID	TC	TD	EP	Attributes	Length	RequesterID	Tag	CompleterID	Status	BCM	Byte Cnt	Lwr Addr	Data
93205		x1	510		10:01010	0	0	0	01	1	000:00:0	1	002:00:0	SC	0	4	0x00	11040000
Packet	R→	2.5	TLP	Cpl	CplID	TC	TD	EP	Attributes	Length	RequesterID	Tag	CompleterID	Status	BCM	Byte Cnt	Lwr Addr	Data
93207		x1	510		10:01010	0	0	0	01	1	000:00:0	1	002:00:0	SC	0	4	0x00	11040000
Packet	R→	2.5	TLP	Cpl	CplID	TC	TD	EP	Attributes	Length	RequesterID	Tag	CompleterID	Status	BCM	Byte Cnt	Lwr Addr	Data
93210		x1	510		10:01010	0	0	0	01	1	000:00:0	1	002:00:0	SC	0	4	0x00	11040000
Packet	R→	2.5	DLLP	NAK	AckNak_Seq_Num	CRC 16		Time Delta										
93211		x1			509	0xCAC4		280.000 ns										
Packet	R→	2.5	TLP	Cpl	CplID	TC	TD	EP	Attributes	Length	RequesterID	Tag	CompleterID	Status	BCM	Byte Cnt	Lwr Addr	Data
93214		x1	510		10:01010	0	0	0	01	1	000:00:0	1	002:00:0	SC	0	4	0x00	11040000
Packet	R→	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16		Idle										
93215		x1			510	0xC28F		480.000 ns										
Packet	R→	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16		Idle										
93216		x1			510	0xC28F		1.312 us										

Compliant NAK



Compliant NAK

- ✓ NAK of TLP #866 Nak's TLPs # 867, #868
- ✓ TLP #867, 868 are replayed
 - Note: TLP# 869 sent before NAK processed

Packet	R→	2.5	TLP	Mem	MWr(32)	RequesterID	Tag	Address	1st BE	Last BE	Data	LCRC	Time Delta
3213	R→	x1	866	Mem	10:00000	002:00:0	0	06037880	1111	1111	32 dwords	0xC872881E	952.000 ns
3214	R→	x1	DLLP	ACK	AdtNak_Seq_Num	CRC 16	Time Delta						
					866	0x36A9	376.000 ns						
3215	R→	x1	867	Mem	10:00000	002:00:0	0	06037900	1111	1111	32 dwords	0x59E9112C	720.000 ns
3216	R→	x1	868	Mem	10:00000	002:00:0	0	06037980	1111	1111	32 dwords	0x6F85E849	1.128 µs
3217	R→	x1	DLLP	NAK	AdtNak_Seq_Num	CRC 16	Time Delta						
					866	0xDDCE	168.000 ns						
3218	R→	x1	869	Mem	10:00000	002:00:0	0	06037A00	1111	1111	32 dwords	0x46BE0908	0.000 ns
3219	R→	x1	867	Mem	10:00000	002:00:0	0	06037900	1111	1111	32 dwords	0x59E9112C	0.000 ns
3220	R→	x1	868	Mem	10:00000	002:00:0	0	06037980	1111	1111	32 dwords	0x6F85E849	24.000 ns

Incompliant NAK

Packet	R→	2.5	TLP	Mem	MRd(32)	RequesterID	Tag	Address	1st BE	Last BE	LCRC	Time Delta
75124		x1	576		00:00000	000:00:0	1	FEAFF180	1111	0000	0x0BD503A0	272.000 ns

Packet	R→	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16	Idle
75125		x1			576	0xC655	56.000 ns

Packet	R→	2.5	TLP	Cpl	CplID	RequesterID	Tag	CompleterID	Status	BCM	Byte Cnt	Lwr Addr	Data	LCRC	Time Delta
75126		x1	3852		10:01010	000:00:0	1	002:00:0	SC	0	4	0x00	1 dword	0x99A353D5	472.000 ns

Packet	R→	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16	Idle
75127		x1			3852	0xC912	224.000 ns

Packet	R→	2.5	TLP	Mem	MWr(32)	RequesterID	Tag	Address	1st BE	Last BE	Data	LCRC	Time Delta
75128		x1	577		10:00000	000:00:0	0	FEAFF180	1111	0000	1 dword	0x1AAAC0DA	288.000 ns

Packet	R→	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16	Time Delta
75129		x1			577	0x674E	2.337 ms

Packet	R→	2.5	DLLP	NAK	AckNak_Seq_Num	CRC 16	Time Delta
75130		x1			3852	0x2275	2.764 ms

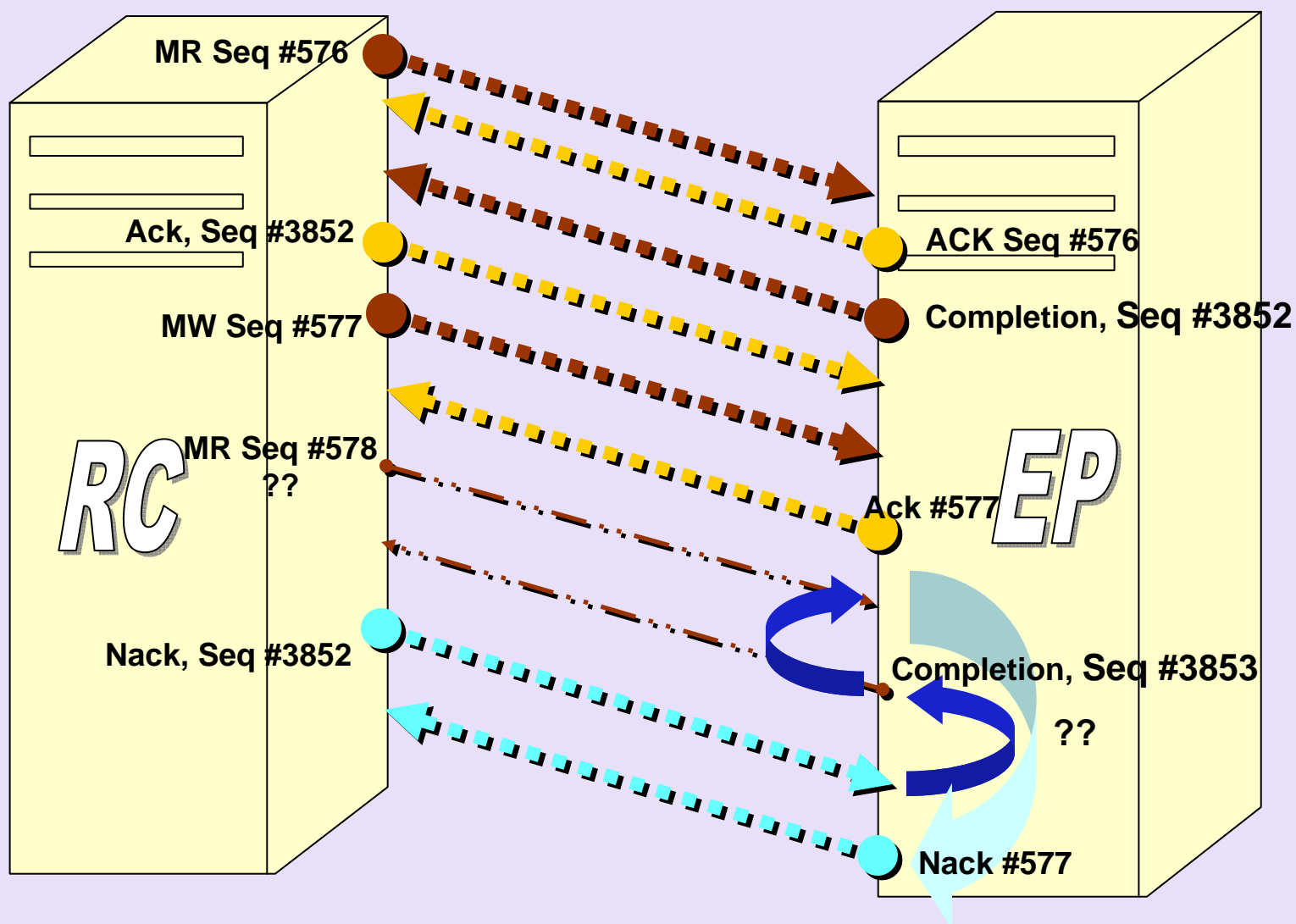
Packet	R→	2.5	DLLP	NAK	AckNak_Seq_Num	CRC 16	Time Delta
75131		x1			577	0x8C29	345.602 ms

Packet	R→	2.5	TLP	Mem	MRd(32)	RequesterID	Tag	Address	1st BE	Last BE	LCRC	Time Delta
75132		x1	578		00:00000	000:00:0	0	FEAFF180	1111	0000	0xE5B3030B	256.000 ns

Incompliant NAK

- Undelivered TLP #3853 Nak'd by NAK #3852
 - ✓ NAK applies to completion for TLP never sent?
- EP interprets NAK #3853 as for bad completion it never sent
- EP alerts RC by NAK #577 that if it expects completion #3853 it should replay Request #578
- RC hangs for 345 msec processing NAK #577
- RC finally plays request #578
- Repeated incompliant NAKs lead to deadlock

Incompliant NAK deadlock



Agenda

- Deadlock issues in bus architectures
- Deadlocks in PCI Express architectures
- Deadlocks at common resource level
 - ✓ Example: ASPM L1 re-entry delay
- Deadlock at RC/chipset level
 - ✓ Example: PME fence mechanism hang-up
- Deadlock at device level – EP or RC
 - ✓ Example 1: Incompliant NAK
 - ✓ Example 2: CLKREQ# readiness
- Deadlock at PCIe link level
 - ✓ Example: Credit leakage



ASPM L1 and Clock Power Management



- L1 exit latency does not account for CLKREQ# enabled, used in miniExpress cards to shut down REFCLK# and further power savings
- Upon deassertion of CLKREQ# EP shuts down PLL in preparation for the loss of REFCLK#
- Device does not account for its own increased PLL recovery time
- System timeout due to long L1 recovery
- PCIe 2.0 requires components to account for additional latency due to loss of REFCLK#
- Both EP and RC need to account for additional latency due to clock power management
- Use backup slow clock to replace bit clock when CLKREQ# de-asserted

Agenda

- Deadlock issues in bus architectures
- Deadlocks in PCI Express architectures
- Deadlocks at common resource level
 - ✓ Example: ASPM L1 re-entry delay
- Deadlock at RC/chipset level
 - ✓ Example: PME fence mechanism hang-up
- Deadlock at device level – EP or RC
 - ✓ Example 1: Incompliant NAK
 - ✓ Example 2: CLKREQ# readiness
- Deadlock at PCIe link level
 - ✓ Example: Credit leakage

Deadlock Associated with Particular PCIe Link

- Credit starvation deadlocks
 - ✓ Malformed TLP going upstream
 - ✓ Downstream device accounts for credit
 - ✓ Upstream device does not account for credit
 - ✓ Additional malformed TLPs lead to credit starvation

Malformed TLP Deadlock

- Downstream FC Init for posted transactions is 16, or $16 \times 16 = 256$ bytes

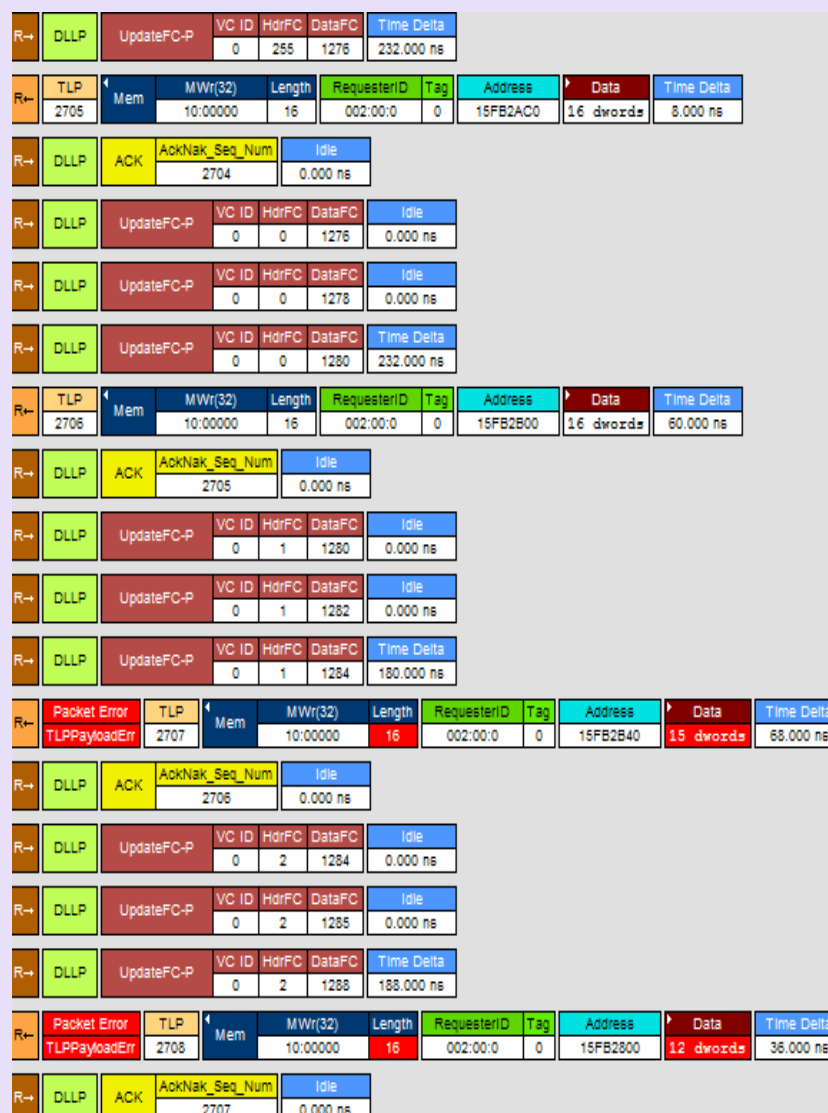
R→	2.5 x1	TS2	COM	Link	Lane	N_FTS	Training Control	Data Rate	TS2 Symbols	Time Delta
			K28.5	1	0	31	0 0 0 0	2.5 GT/s	D05.2 ...	16.000 ns
R→	2.5 x1	TS2	COM	Link	Lane	N_FTS	Training Control	Data Rate	TS2 Symbols	Time Delta
			K28.5	1	0	15	0 0 0 0	2.5 GT/s	D05.2 ...	48.000 ns
R→	2.5 x1	TS2	COM	Link	Lane	N_FTS	Training Control	Data Rate	TS2 Symbols	Time Delta
			K28.5	1	0	31	0 0 0 0	2.5 GT/s	D05.2 ...	16.000 ns
R→	2.5 x1	TS2	COM	Link	Lane	N_FTS	Training Control	Data Rate	TS2 Symbols	Time Delta
			K28.5	1	0	15	0 0 0 0	2.5 GT/s	D05.2 ...	48.000 ns
R→	2.5 x1	TS2	COM	Link	Lane	N_FTS	Training Control	Data Rate	TS2 Symbols	Idle
			K28.5	1	0	31	0 0 0 0	2.5 GT/s	D05.2 ...	96.000 ns
R→	2.5 x1	SKIP	COM	SKIP Symbols			Idle			
			K28.5	K28.0 K28.0 K28.0			0.000 ns			
R→	2.5 x1	DLLP	InitFC1-P	VC ID	HdrFC	DataFC	CRC 16	Idle		
				0	4	16	0xFBB9	0.000 ns		
R→	2.5 x1	DLLP	InitFC1-NP	VC ID	HdrFC	DataFC	CRC 16	Idle		
				0	8	4	0x9809	0.000 ns		
R→	2.5 x1	DLLP	InitFC1-Cpl	VC ID	HdrFC	DataFC	CRC 16	Idle		
				0	0	0	0xD892	0.000 ns		
R→	2.5 x1	DLLP	InitFC1-P	VC ID	HdrFC	DataFC	CRC 16	Time Delta		
				0	4	16	0xFBB9	32.000 ns		
R→	2.5 x1	DLLP	InitFC1-P	VC ID	HdrFC	DataFC	CRC 16	Time Delta		
				0	5	18	0x55E0	0.000 ns		
R→	2.5 x1	DLLP	InitFC1-NP	VC ID	HdrFC	DataFC	CRC 16	Time Delta		
				0	8	4	0x9809	32.000 ns		
R→	2.5 x1	DLLP	InitFC1-NP	VC ID	HdrFC	DataFC	CRC 16	Time Delta		
				0	5	18	0xBE87	0.000 ns		
R→	2.5 x1	DLLP	InitFC1-Cpl	VC ID	HdrFC	DataFC	CRC 16	Time Delta		
				0	0	0	0xD892	32.000 ns		
R→	2.5 x1	DLLP	InitFC1-Cpl	VC ID	HdrFC	DataFC	CRC 16	Time Delta		
				0	0	0	0xD892	0.000 ns		

Malformed TLP Deadlock

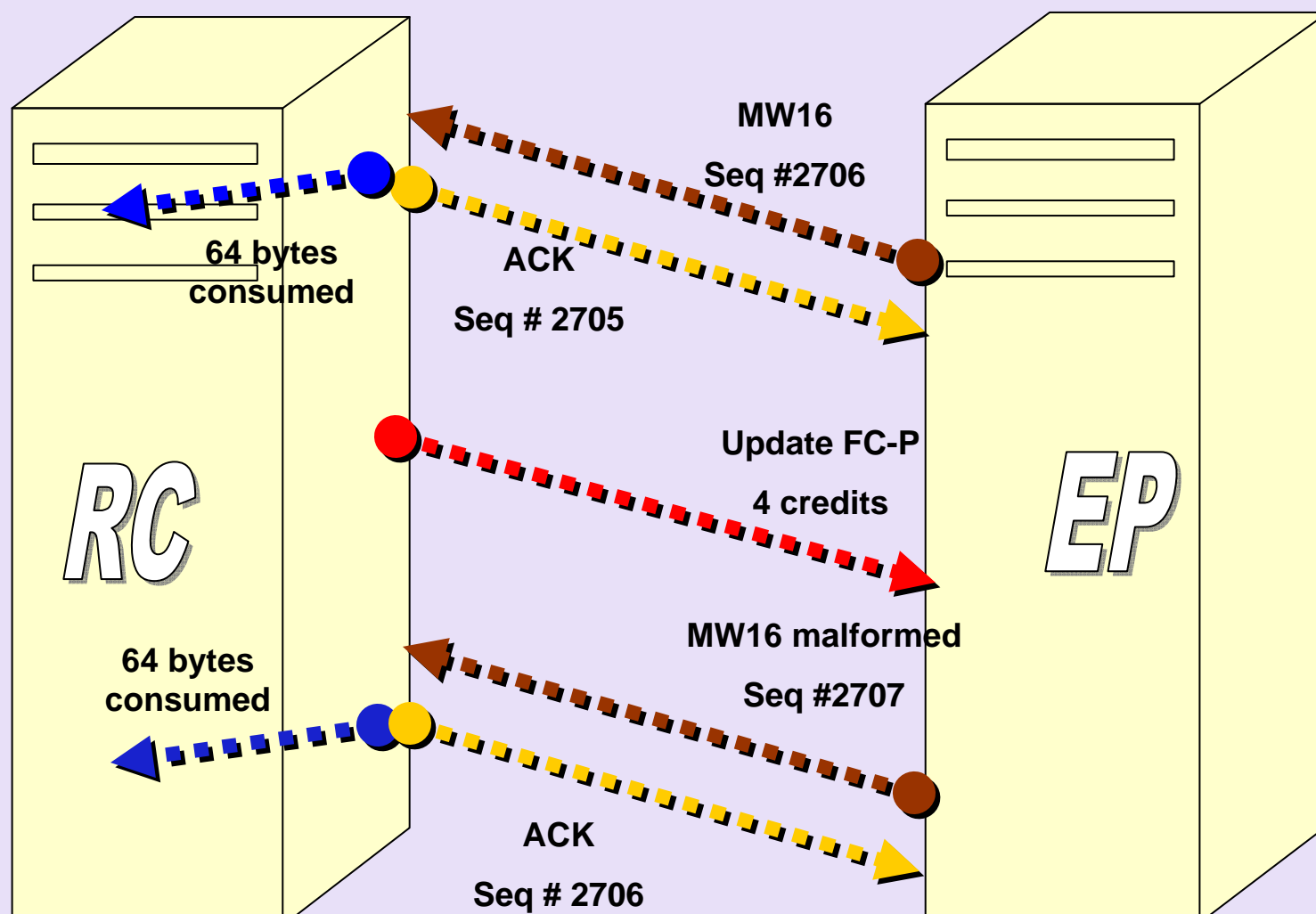
- RC advertises buffer space for a request for $4 \times 16 = 64$ posted data bytes
- EP sends MW request for $16 \times 4 = 64$ bytes
- RC consumes 64 bytes and advertises release of buffer space
- EP sends additional memory write request
- RC receives a malformed TLP
- RC does not consume, does not update
- Result is credit leakage

Malformed TLP Deadlock

- Memory Write TLPs require free resource of $16 \times 4 = 64$ bytes
= 4 credits
- TLP 2705 is sent, but only TLP 2704 is ack'd
- 4 credits are consumed and FC updated
- EP lags 4 credits behind RC
- Each malformed TLP adds additional lag of 4 credits



Malformed TLP Deadlock

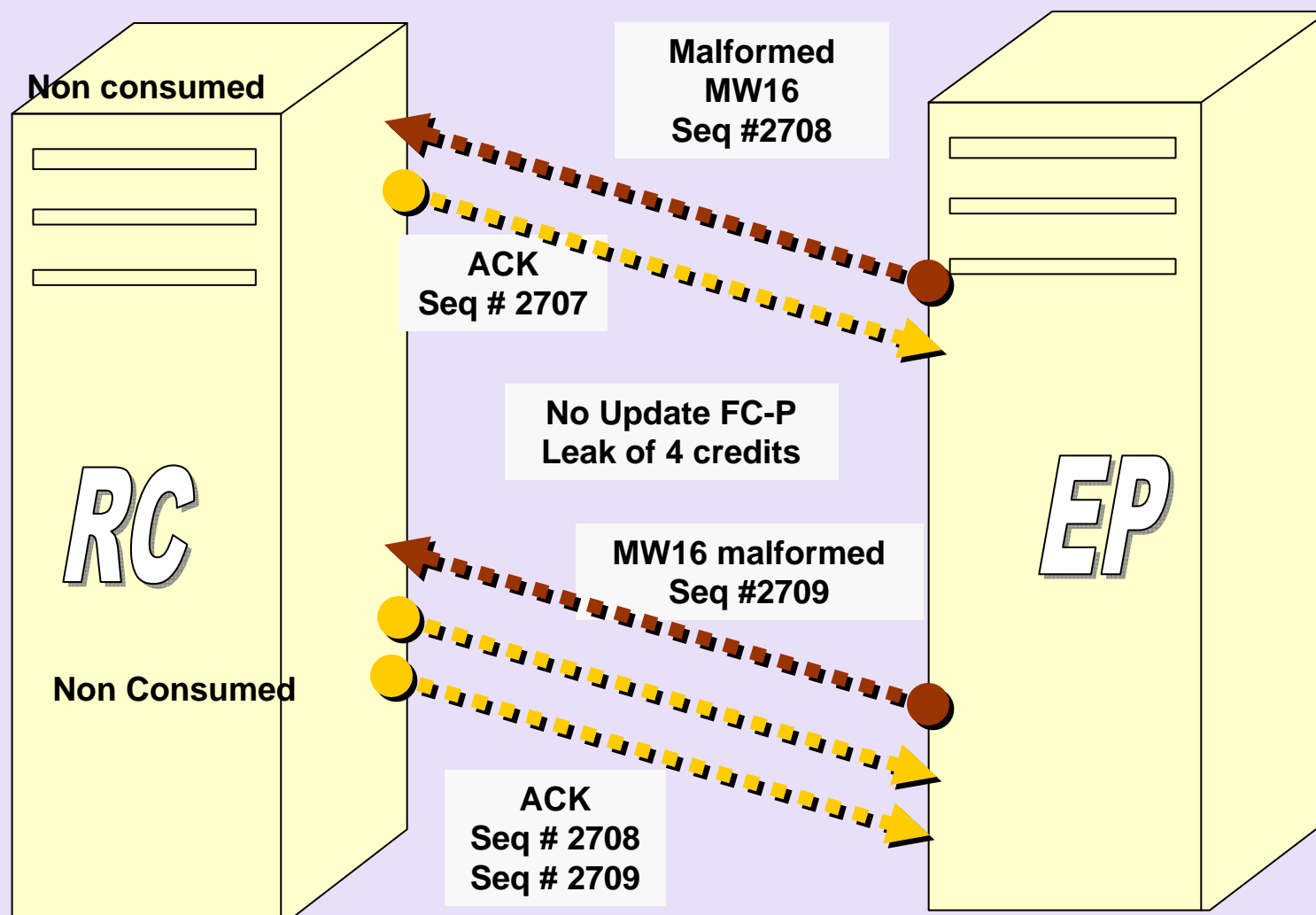


Malformed TLP Deadlock

- 3 malformed TLPs create a lag of 12 credits
- Leads to EP credit starvation and blocks further traffic
- Deadlock scenario resulting from credit starvation

R←	Packet Error TLPPayloadErr	TLP 2708	Mem	MWr(32) 10:00000	Length 16	RequesterID 002:00:0	Tag 0	Address 15FB2800	Data 12 dwords	Time Delta 36.000 ns
R→	DLLP	ACK	AckNak_Seq_Num 2707	Idle 0.000 ns						
R→	DLLP	UpdateFC-NP	VC ID 0	HdrFC 242	DataFC 4	Time Delta 236.000 ns				
R←	Packet Error TLPPayloadErr	TLP 2709	Mem	MWr(32) 10:00000	Length 16	RequesterID 002:00:0	Tag 0	Address 15FB2800	Data 12 dwords	Time Delta 40.000 ns
R→	DLLP	ACK	AckNak_Seq_Num 2708	Idle 232.000 ns						
R→	DLLP	ACK	AckNak_Seq_Num 2709	Time Delta 3.560 μs						
R←	DLLP	UpdateFC-P	VC ID 0	HdrFC 44	DataFC 1849	Idle 0.000 ns				
R←	DLLP	UpdateFC-NP	VC ID 0	HdrFC 176	DataFC 95	Time Delta 23.960 μs				
R→	DLLP	UpdateFC-P	VC ID 0	HdrFC 2	DataFC 1288	Idle 964.000 ns				
R→	DLLP	UpdateFC-NP	VC ID 0	HdrFC 242	DataFC 4	Time Delta 7.788 μs				
R←	DLLP	UpdateFC-P	VC ID 0	HdrFC 44	DataFC 1849	Idle 0.000 ns				
R←	DLLP	UpdateFC-NP	VC ID 0	HdrFC 176	DataFC 95	Time Delta 20.296 μs				
R→	DLLP	UpdateFC-P	VC ID 0	HdrFC 2	DataFC 1288	Idle 972.000 ns				
R→	DLLP	UpdateFC-NP	VC ID 0	HdrFC 242	DataFC 4	Time Delta 11.460 μs				
R←	DLLP	UpdateFC-P	VC ID 0	HdrFC 44	DataFC 1849	Idle 0.000 ns				
R←	DLLP	UpdateFC-NP	VC ID 0	HdrFC 176	DataFC 95	Time Delta 16.660 μs				
R→	DLLP	UpdateFC-P	VC ID 0	HdrFC 2	DataFC 1288	Idle 976.000 ns				
R→	DLLP	UpdateFC-NP	VC ID 0	HdrFC 242	DataFC 4	Time Delta 15.100 μs				

Malformed TLP Deadlock

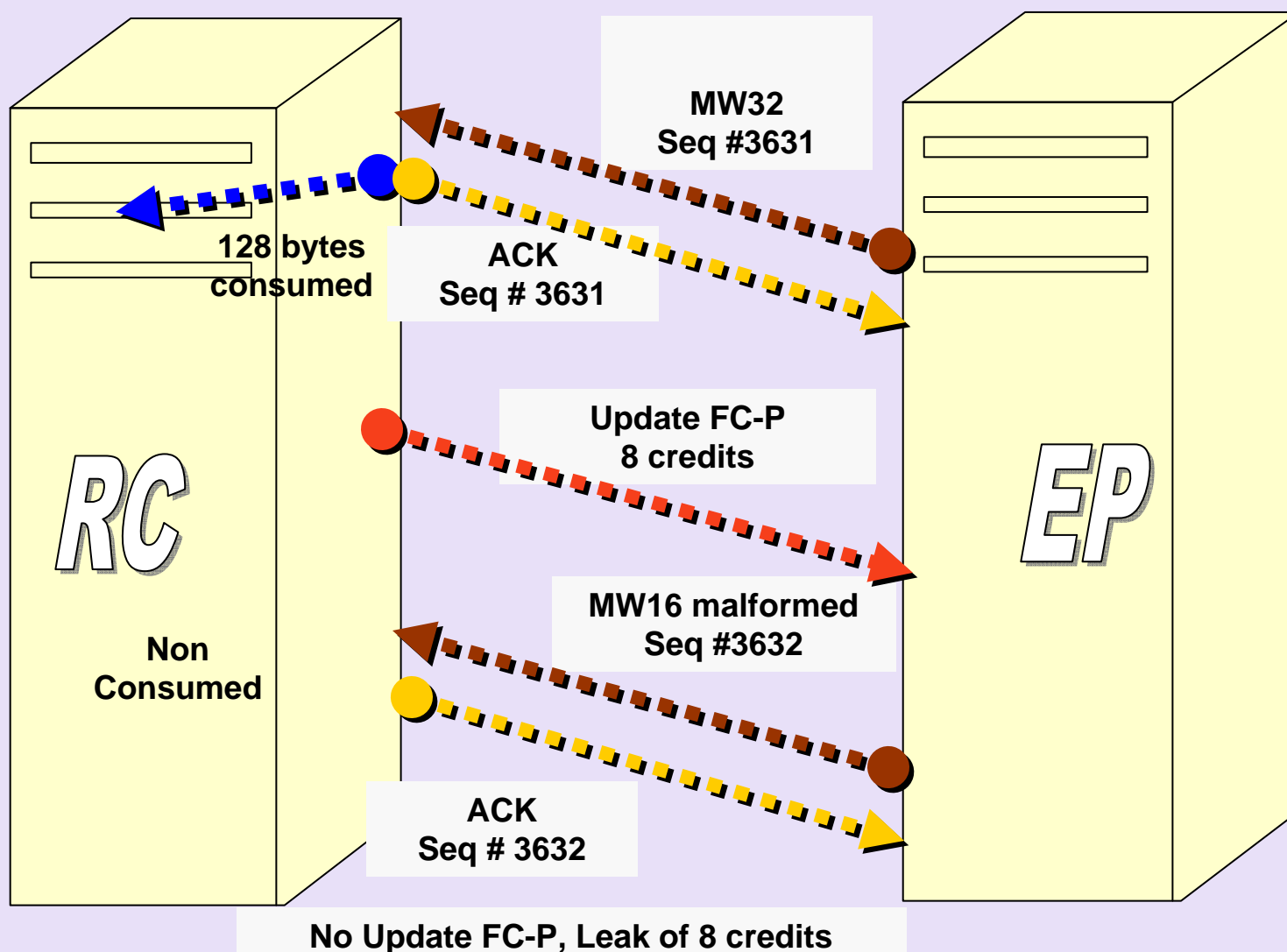


Malformed TLP Deadlock

- TLP of length 32 requires $32 * 4 / 16 = 8$ credits
- No Updates
- Credit starvation will occur with only 2 malformed TLPs

R→	2.5	DLLP	UpdateFC-P	HdrFC	DataFC	Idle	Time Stamp				
	x1			170	7	0.000 ns	0147 . 191 591 984 s				
R→	2.5	DLLP	UpdateFC-P	HdrFC	DataFC	Time Delta	Time Stamp				
	x1			170	9	48.000 ns	0147 . 191 592 016 s				
R→	2.5	DLLP	UpdateFC-P DLLP: Updating Flow Control Credits for Virtual Channel								
	x1			170	11	224.000 ns	0147 . 191 592 064 s				
R→	2.5	TLP	Mem	MWrr(32)	Attributes	Length	RequesterID	Tag	Address	Data	Time Delta
	x1	3631		10:00000	00	32	002:00:0	0	285DC580	32 dwords	672.000 ns
R→	2.5	DLLP	ACK	AckNak_Seq_Num	Idle	Time Stamp					
	x1			3631	0.000 ns	0147 . 191 592 960 s					
R→	2.5	DLLP	UpdateFC-P	HdrFC	DataFC	Idle	Time Stamp				
	x1			171	11	0.000 ns	0147 . 191 592 992 s				
R→	2.5	DLLP	UpdateFC-P	HdrFC	DataFC	Idle	Time Stamp				
	x1			171	13	0.000 ns	0147 . 191 593 024 s				
R→	2.5	DLLP	UpdateFC-P	HdrFC	DataFC	Idle	Time Stamp				
	x1			171	15	68.000 ns	0147 . 191 593 056 s				
R→	2.5	DLLP	UpdateFC-P	HdrFC	DataFC	Idle	Time Stamp				
	x1			171	17	0.000 ns	0147 . 191 593 156 s				
R→	2.5	DLLP	UpdateFC-P	HdrFC	DataFC	Time Delta	Time Stamp				
	x1			171	19	260.000 ns	0147 . 191 593 188 s				
R→	2.5	Packet Error	TLP	Mem	MWrr(32)	Attributes	Length	RequesterID	Tag	Address	Data
	x1	TLPayloadErr	3632		10:00000	00	32	002:00:0	0	285DC600	15 dwords
R→	2.5	DLLP	ACK	AckNak_Seq_Num	Time Delta	Time Stamp					
	x1			3632	9.912 µs	0147 . 191 593 820 s					
R→	2.5	DLLP	UpdateFC-NP	HdrFC	DataFC	Time Delta	Time Stamp				
	x1			116	4	18.088 µs	0147 . 191 603 732 s				
R→	2.5	DLLP	UpdateFC-P	HdrFC	DataFC	Time Delta	Time Stamp				
	x1			171	19	11.052 µs	0147 . 191 621 820 s				
R→	2.5	DLLP	UpdateFC-NP	HdrFC	DataFC	Time Delta	Time Stamp				
	x1			116	4	18.084 µs	0147 . 191 632 872 s				
R→	2.5	DLLP	UpdateFC-P	HdrFC	DataFC	Time Delta	Time Stamp				
	x1			171	19	11.024 µs	0147 . 191 650 956 s				
R→	2.5	DLLP	UpdateFC-NP	HdrFC	DataFC	Time Delta	Time Stamp				
	x1			116	4	18.068 µs	0147 . 191 661 980 s				
R→	2.5	DLLP	UpdateFC-P	HdrFC	DataFC	Time Delta	Time Stamp				
	x1			171	19	11.060 µs	0147 . 191 680 048 s				

Malformed TLP Deadlock



Malformed TLP Deadlock

- Error reporting
 - ✓ Malformed TLP Error reporting message required for downstream TLP, meaningless for upstream TLP
- Specification enhancement recommendation
 - ✓ Downstream FC Init or link retrain should be mandatory following malformed TLP, or other scenarios leading to credit starvation
 - ✓ ECR proposal under study

Thank you for attending the
PCI-SIG Developers Conference
Europe 2007.

For more information please go to
www.pcisig.com



Deadlock Avoidance in PCI Express Based Architectures

Isaac Livny
Applications Engineer
LSI Corporation

