

The Complete PCI Express[†] Reference

Design Insights for Hardware
and Software Developers

Edward Solari
Brad Congdon

INTEL
PRESS

Contents

Acknowledgements xxi

Chapter 1 Architecture Overview 1

The Original PCI Bus	2
PCI Platforms	3
Early Architecture	3
PCI Hub-Based Architecture	5
PCI 64 Data Bits and 66.6 Megahertz Emerge	6
PCI Protocol Basics	7
Limitations of the PCI Bus Architecture	14
PCI Bus Feature Summary	15
PCI-X Platforms	16
Split Transaction Protocol	17
Improved Performance of PCI-X 1.0 over PCI	17
PCI-X 2.0 Increases Performance Further	19
PCI-X 1066	20
PCI-X Bus Feature Summary	21
PCI Express	23
PCI Express Improves Upon PCI and PCI-X	23
Platform Implementations with PCI Express	26
PC and Server Platforms	27
PCI Express Specific Performance	31
Other PCI Express Performance and Features	33

Chapter 2 PCI Express Architecture Overview 41

Generic PCI Express Platform Architecture	41
Hierarchy and Hierarchy Domain	44
Other Interconnects between PCI Express Devices	44
Generic Architecture of PCI Express Devices and Flow	
Control Considerations	48
Root Complex	48
Switches	52
Endpoints	54
Bridges	54
Overview of Transaction Flow and Participants	57
Restrictions of Root Complex as Requester/Completer	57
Restrictions on Switch TLP Porting	58
Endpoints	59
Restrictions on Bridges TLP Porting	61
In-band and Advanced Peer-to-Peer Communications	62
Transactions and Packets	62
Requester/Completer Protocol	63
PCI Express Layers and Packets Overview	65
Transmitting and Receiving Packets: An Example	69
Flow Control Protocol	74
Quality of Service	75
Traffic Classes	76
Virtual Channels	78
Link Flow Control	78
Other PCI Express Compatibilities and Differences	
with PCI and PCI-X	78
Address Spaces	79
Messages	84
Interrupts	84
Lock Function	85
Power Management	85
Hot Plug	86
Mechanical and Electrical Overview	86
Mechanical	86
Electrical	88

Chapter 3 Implementation 93

Generic Root Complex and Endpoints	95
Generic Switch	98

Generic Bridge	100
Other Implementations	101
Other Implementations of the Root Complex	102
Other Implementations of Switches	106
Other Implementations of Bridges	109
Other Bridge Considerations	112
PCI Express/PCI Bridge and PCI Express/PCI-X Bridge Considerations	112

Chapter 4 Addressing and Routing 121

Address Spaces Defined in PCI Express	123
Memory Address Space	123
I/O Address Space	126
Configuration Address Space	128
Message Address Space	140
Participants of Transactions	143
Participants of Memory, I/O and Configuration Transactions	143
Message Transaction Participants	146
Addressing and Routing	150
Addressing of Memory and I/O Transactions	151
Configuration and Completer Transaction ID Routing	152
Routing Message Transaction Packets	154
Associated Addressing Protocols	161
Type 0 and Type 1 Configuration Requester Transactions	161
TAG# and Transaction ID	168

Chapter 5 Transaction Packets 171

Header Field Information	174
Memory Transaction Packets and Execution Protocol	175
Requester ID, Completer ID, and Tag#	176
Memory Read Transaction Packets	178
Memory Read Transaction Execution	180
Treatment of Other Fields in the Memory Read Transaction Packets	188
Memory Write Transaction Packets	191
Memory Write Transaction Execution	192
Treatment of Other Fields in Memory Write Transaction Packets	192
I/O Transaction Packets and Execution Protocol	193
I/O Transaction Packets	194

I/O Read and Write Execution	196
Treatment of Other Fields in the I/O Transaction Packets	196
Configuration Transaction Packets Protocol	197
Configuration Transaction Packets	198
Configuration Read and Write Execution	200
Treatment of Other Fields in the Configuration Transaction Packets	200
Message Transaction Packets Protocol	202
Message Baseline Transaction Packets	203
Message Baseline Transaction Execution	204
Treatment of Other Fields in Message Baseline Transaction Packets	208
Message Vendor-Defined Transaction Packets	209
Message Vendor-Defined Transaction Execution Packets	211
Treatment of Other Fields in Message Vendor-Defined Transaction Packets	212
Message Advanced Switching Requester Transaction Packets	213

Chapter 6 Transaction Layer 215

Transaction Layer	217
The Control Element	218
The Address Element	218
The Flow Control Protocol Element	219
The Repeater/Completer Protocol Element	219
The Data Element	220
The Integrity Element	220
The Message Element	221
Transaction Layer Implementation	221
Header Field Formats	222
Header Field Formats for Memory Read Transactions	224
MEMORY READ REQUESTER TRANSACTION	227
MEMORY READ COMPLETER TRANSACTION	232
Header Field Formats for Memory Write Transactions	238
MEMORY WRITE REQUESTER TRANSACTION	240
Header Field Formats for I/O Read Transactions	245
I/O READ REQUESTER TRANSACTION	247
I/O READ COMPLETER TRANSACTION	250
Header Field Formats for I/O Write Transactions	255
I/O WRITE REQUESTER TRANSACTION	256

I/O WRITE COMPLETER TRANSACTION	260
Header Field Formats Configuration Read Transactions	263
CONFIGURATION READ REQUESTER TRANSACTION	265
CONFIGURATION READ COMPLETER TRANSACTION	269
Header Field Formats for Configuration Write Transactions	274
CONFIGURATION WRITE REQUESTER TRANSACTION	275
CONFIGURATION WRITE COMPLETER TRANSACTION	279
Header Field Formats for Message Baseline Transactions	282
MESSAGE BASELINE REQUESTER TRANSACTION	284
Header Field Formats for Message Vendor-defined Transactions	293
MESSAGE VENDOR-DEFINED REQUESTER TRANSACTION	295
Header Field Formats for Message Advanced Switching Transactions	300
MESSAGE ADVANCED SWITCHING REQUESTER TRANSACTION	302
Data Field and Digest Field Formats	305
Data Field	305
Digest Field	312

Chapter 7 Data Link Layer and Packets 315

Overview of Data Link Layer Packets	317
DLLP	317
LLTP	317
Other Considerations	318
Link Activity States	318
Link_UP and Link_DOWN	319
Link Activity States Application	319
Details of Link Activity States	322
Flow Control Initialization	324
Set of Buffers	325
Flow Control Initialization Protocol	329
Flow Control Protocol Update of Available Buffer Space	340
DLLP Formats	342
Reserved	343
Flow Control DLLPs	343
Acknowledgement DLLPs	348
Power Management (PM) DLLPs	350
Vendor-Specific DLLPs	353

- LLTP Format 355
- Reserved 355
- LLTPs 355

Chapter 8 Physical Layer and Packets 359

- Parallel LLTPs and DLLPs versus Serial Physical Packets 362
 - Step 1: Conversion of LLTP and DLLP Portion of Physical Packet 365
 - Step 2: Addition of FRAMING to Complete the Physical Packet 370
 - Step 3: Encoding and Decoding Protocols and Reference Clock 372
- Physical Packet and Other Symbol Transmission 381
- Lane Parsing 383
 - Parsing Example with Four Lanes 383
 - K and D Symbols' Application and Lane Parsing 385
 - Application of K and D Symbols Relative to Physical Packets and Actual Lane Parsing 387
 - Receiver Error 395
- Application of Ordered Sets 395
 - Skip OS 396
 - Fast Training Sequence OS 397
 - Electrical Idle OS 398
 - Training Sequence 1 and 2 OSs (also Known as TS1 and TS2 OSs) 401
 - Inclusion of Ordered Sets into Lane Parsing 403
 - Receiver Error 404
- Special Transmitting and Receiving Considerations 405
 - Scrambling 405
 - Lane Polarity Inversion 407
 - Disparity 407
- Link Training and Link Configuration 408
 - Link Configuration Protocol Iterative Process Discussion 409

Chapter 9 Errors 421

- Relationship to PCI 422
- Error Protocol Specific to PCI Express 423
- Error Reporting 424
- Basic Transaction Layer Errors 425
 - Errors Related to TLPs and Checked at Requester and Completer Destinations 427

Special Error Considerations Specific to Message Requester Transaction Packets	438
Special TLP Error Considerations for Switches and Bridges (Excluding Error Poisoning)	439
Errors Related to Transaction Layer: Error Forwarding (also Known as Data Poisoning)	442
Errors Related to Transaction Layer: Completion Timeout Error	444
Errors Related to Transaction Layer: Receiver Overflow Error	445
Errors Related to Transaction Layer:	
Flow Control Protocol Error (FCPE)	446
Transaction Layer: TLP Fields Checked for Errors	447
MEMORY and I/O REQUESTER TRANSACTIONs contain TYPE, ATTRIBUTE, REQUESTER, and TAG HDWs:	448
CONFIGURATION REQUESTER TRANSACTIONs contain TYPE, ATTRIBUTE, REQUESTER, and TAG HDWs:	452
COMPLETER TRANSACTIONs contain TYPE, ATTRIBUTE, COMPLETER, STATUS, REQUESTER, and TAG HDWs:	455
MESSAGE BASELINE REQUESTER TRANSACTIONs contain TYPE, ATTRIBUTE, REQUESTER, and TAG HDWs	460
MESSAGE VENDOR-DEFINED REQUESTER TRANSACTIONs contain TYPE, ATTRIBUTE, REQUESTER, and TAG HDWs	464
MESSAGE ADVANCED SWITCHING REQUESTER TRANSACTIONs contain TYPE, ATTRIBUTE, REQUESTER, and TAG HDWs	467
Data Link Layer Errors	469
Special Error Considerations for Switches	471
Errors Related to CRC, LCRC, and Nullified TLP Packets	471
Errors Related to Retry_Timer Timeout and Retry_Num# Counter Rollover	473
Data Link Layer Protocol Errors (DLLPEs)	473
Physical Layer Errors	474
Special Error Considerations for Switches	475
Receiver Errors	475
Training Errors	477
CRC, LCRC, ECRC, and Digest Field	478
CRC and LCRC	478
ECRC	479
Protocols for Calculating and Checking CRC, LCRC and ECRC	480

Error Reporting and Logging	484
General Error Considerations	485
Message Error Requester Transaction Packets	485
Minimum Error Reporting Protocol	488
Advance Error Reporting Protocol	492

Chapter 10 Transaction Ordering 501

Transactions Begin and End	501
Simplest Transaction Execution	501
More Complex Transaction Execution	502
Buffers in the PCI Express Fabric	503
Buffers versus FIFOs	504
Solution to Livelock and Deadlock Situations	505
PCI Express Transaction Ordering Protocol	505
Inclusion of Endpoints	505
Attribute	506
Traffic Class	506
Transaction Ordering Protocol Implementation	507
Specifics	512
Special Considerations	516

Chapter 11 Flow Control Protocol: Part 1 519

Introduction to the Flow Control Protocol, Part 1	519
Flow Control Protocol Elements Overview	520
Elements of the Flow Control Protocol	526
Flow Control Protocol: Virtual Channels and Traffic Classes	528
Flow of Transaction Packets and Definition of <i>Packet</i> in This Chapter	528
Virtual Channels	531
Additional Considerations	539
Traffic Classes	541
Flow Control Protocol: Buffer Management	553
FCC and Set of Buffers	556
FCC	557
Flow Control Protocol Execution	559

Chapter 12 Flow Control Protocol: Part 2 581

Introduction to the Flow Control Protocol: Part 2	581
Receiving Ports versus Transmitting Ports	582
VC0 versus VCx	582

LLTP Tracking	582
LLTP Tracking at Transmitting and Receiving Ports	583
PCI Express Arbitration	605
Other Arbitration Considerations	607
Transaction Layer Packet (TLP) Arbitration	607
Switches with Upstream TLP Flow	608
Port Arbitration	611
VC Arbitration	613
Switches with Downstream Flowing and Peer-to-Peer TLP	616
Non-Switch Port Arbitration, VC Arbitration, and Buffer Considerations	617
Priority Schemes	620
Quality of Service	624
Latency	625
Effective Bandwidth	625
Bandwidth Consistency	625
Backpressure	625
Total Latency	626
Isochronous Transfers	627

Chapter 13 PME Overview, Wake Events, and Reset 637

Introduction	637
Other Background Information	638
Power Management Protocol Overview	644
Transition from Sleep (L2) or Powered Off (L3) to Powered On (L0)	644
Lowered Power	646
Placement into Sleep or Powered off Link State	648
Power Cycling	650
Main Power and Vaux	651
Powered On to Powered Off or Sleep	651
Powered Off to Powered On	653
Sleep to Powered On	654
Wake Events: Wakeup Protocol and Wakeup Event	655
The Other Wake Event	656
PCI Legacy PME# Signal Line	656
Wakeup Protocol	657
Resets	662
PERST# Signal Line Considerations	663
Fundamental Reset Protocol	664
Hot Reset Protocol	666

Chapter 14 Link States 671

Introduction to Link States	671
Other Considerations	672
Normal Operation Link State	677
L0 Link State	677
Power Management Category of Link States	682
L0s Link State	682
L1 Link State	686
L2/3 Ready Link State	688
L2 Link State	689
L3 Link State	692
Link Training Link States	692
Detect Link State	693
Polling Link State	696
Configuration Link State	700
Other Link States	714
Recovery Link State	714
Disabled Link State	720
Hot Reset	723
Testing and Compliance	724
Loopback Link State	724
Polling.Compliance Link Sub-state	729

Chapter 15 Power Management 731

Introduction to Power Management	731
Support Requirements	732
Powered-on, Powered-off, and Sleep	733
L0 and Power Management Link States	734
L0 Link State	734
L0s Link State	735
L1 Link State	736
L2/3 Ready Link State	738
L2 Link State	739
L3 Link State	740
PCI-Power Management Software Protocol	740
“D” States for PCI-PM Software Protocol	741
TLPs for PCI-PM Software Protocol	744
DLLPs for PCI-PM Software Protocol	753
PCI-PM Software Protocol Application of “D” States, TLPs, and DLLPs	753

Active-State Power Management Protocol	758
TLP for Active-State PM Protocol	759
DLLPs for Active-State PM Protocol	759
Active-State PM Protocol Application of the TLP and DLLPs	760

Chapter 16 Hot Plug Protocol 775

Introduction to the Hot Plug Protocol	775
Hardware Components for the HP Protocol	776
Power Indicator	777
MRL and MRL Sensor	777
Attention Indicator	778
Attention Button	779
Power Controller (Fault)	780
Add-in Card Present	780
Non-Hardware Components for the HP Protocol	781
Message Baseline Requester Transactions	782
Configuration Address Space Registers for Hot Plug Protocol	786
Hot Plug Protocol	791
HP Protocol with Normal Power	792
HP Protocol with Sleep or Lower Power State	794
Other Considerations for HP Protocol	795

Chapter 17 Slot Power Protocol 797

Introduction to Slot Power	797
Slot Power (SP) Protocol	798
Power Limits	798
Power Associated Message Transactions and Configuration Registers	799
Slot Power Limit Protocol Implementation	803
Message Baseline Requester Transaction Packets: Slot Power Messages	804
Associated Slot Power Items	805

Chapter 18 Advanced Switching 807

Introduction to Advanced Switching	807
Advanced Peer-to-Peer Links	810
Advanced Peer-to-Peer Communications	811

Chapter 19 Interrupts 813

- Introduction to Interrupts 813
- Legacy Interrupt Protocol Implementation 814
 - Message Interrupt Requester Transaction Packets 815
 - Message Interrupt Requester Transaction Packet Flow 816
 - Legacy Interrupt Protocol Implementation by Legacy Endpoints, Bridges, and Switches 817
 - Legacy Interrupt Protocol Implementation by the Root Complex 821
 - Boundary Conditions for Disabling INTx 821
- Unique Considerations for the Legacy Interrupt Protocol for Re-Mapping through Each Virtual PCI/PCI Bridge 823
- Message Signaled Interrupt Protocol 832

Chapter 20 Lock Protocol 833

- Introduction to the Lock Protocol 833
 - EHA Participants 834
- Lock Function (LF) Protocol 835
 - Lock Transaction Packets 835
 - LF Protocol to Establish EHA 835
 - LF Protocol after EHA Established 836
 - Restriction of Accesses to Locked Device 837
 - LF Protocol to Terminate LF between Root Complex and Locked Device 838
- MW_r Issue 839

Chapter 21 Mechanical and Electrical Overview 841

- Introduction 841
- Mechanical 841
 - Non-Mobile Add-in Cards 842
 - Mobile Add-in Cards: PCI Express Mini Cards 850
 - Cable Modules 853
 - Signal Line Definitions 854
- Electrical 857
 - PHY 858
 - Waveform 863

Chapter 22 Configuration Overview 865

- Introduction 865
- Features of Configuration Space 866
- The Target of a Configuration Address 867
- Configuration Space Size 867
- Configuration Space Hierarchy 868
- Creating a Configuration Address 872
- Legacy PCI Express Configuration Mechanism 872
 - Config_Address Register 872
 - Config_Data Register 873
 - Config_Address and Config_Data Example 873
 - Limitation of Config_Address and Config_Data 875
- Native PCI Express Configuration Mechanism 875
- How Configuration Addresses Reach Their Destination 876
- Type 0 and Type 1 Configuration Accesses 876
 - Configuration Access Example 1 877
 - Configuration Access Example 2 878
 - Configuration Access Example 3 879
- Access Rules for Configuration Space 880
 - Access Rules for Configuration Space Reads 880
 - Access Rules for Configuration Space Writes 880
- How to Detect PCI Express Devices 881
- PCI Express Bus Enumeration 882
- PCI Express Bridge Architecture 884
 - Two Rules for Connecting PCI Express Bridges 885
 - Other Rules for PCI Express Bridges 891

Chapter 23 Configuration Registers 893

- Introduction to Configuration Registers 893
- Basic Function of Configuration Space Registers 894
- Configuration Space Architecture 895
- Configuration Space Layout 896
- Root Complex Register Block 898
- Configuration Header Register Block 899
- Common Header Region 899
- Type 0 Header Region 908
- Type 1 Header Region 919

Chapter 24 Configuration Capabilities 937

- Introduction to Configuration Capabilities 937
- Defined Capabilities 938

Required Capabilities	938
Extended Capabilities	938
Configuration Capability Architecture	939
Device-independent Region	940
Header Type Region	940
Device-Dependent Region	940
Extended Configuration Region	940
Shared Registers	941
Capabilities List Bit	941
How to Locate the Capabilities List	941
How to Scan the Capabilities List	943
Capability Structure Layout	943
Capability Linked List	945
Required Capabilities Overview	946
PCI Express Capability Structure	946
PCI Power Management Capability Structure	947
Message Signaled Interrupts Capability Structure	948
PCI Express Capability Structure	948

Chapter 25 Real-Time Analysis 977

Configurations	977
X1 and X4 Configuration Test Setup	977
X8 Configuration Test Setup	978
Mid-Bus Probe:	979
PCI Express Transaction Flow Overview	979
Transaction Layer	981
TLP Fields	981
TLP Routing	982
TLP Request/Completion	983
Data Link Layer	986
Physical Layer	987
Configuration Space	992
Traffic Class and Virtual Channels	999

Appendix A Capability IDs 1001

Glossary 1005

Index 1019