SDC 19
SNIA EMEA
JANUARY 30, 2019
TEL AVIV, ISRAEL
STORAGE DEVELOPER CONFERENCE

Gen-Z and Storage Class Memory (SCM) integration and path forward

Parmeshwr_Prasad@dell.com
DellEMC

Agenda

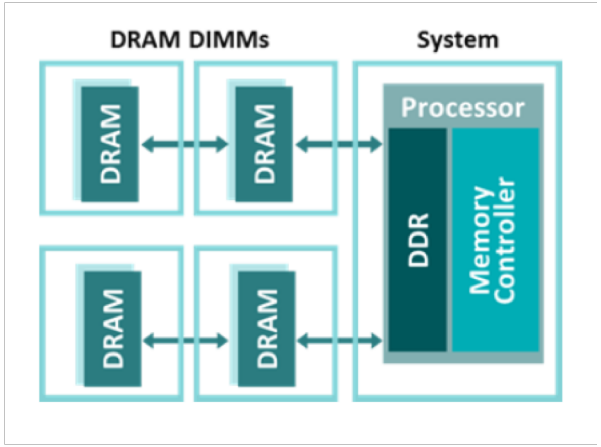
- Gen-z design principal and advantages
- OpClass
- ZMMU
- Memory read
- Memory write

Gen-Z Design Principle

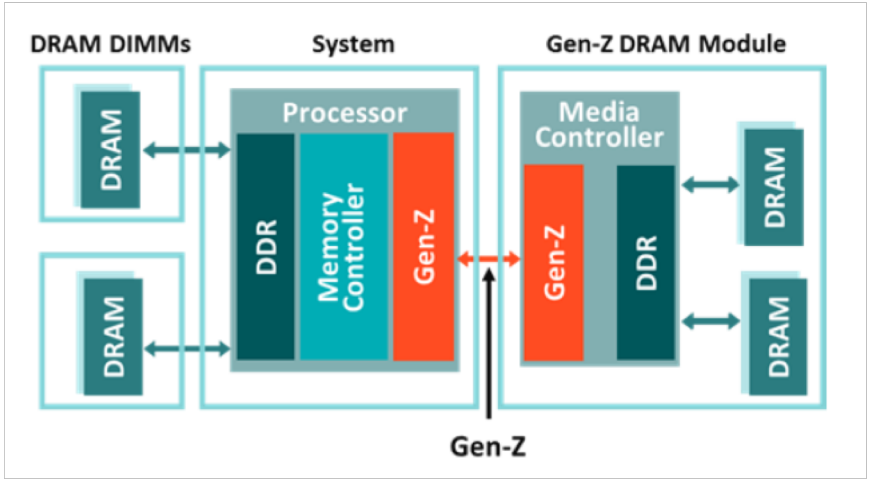
Gen-Z is a scalable, universal system interconnect intended to unify communications, simplify designs and solution architectures, and improve interoperability.

- Simplifying Data Path
- Architectural flexibility
- Resiliency
- Security built into the architecture
- Scalability

Traditional - Vs - Gen-Z integration



- Memory Capacity
- Channel & sockets
- Lifecycle
- Memory controller

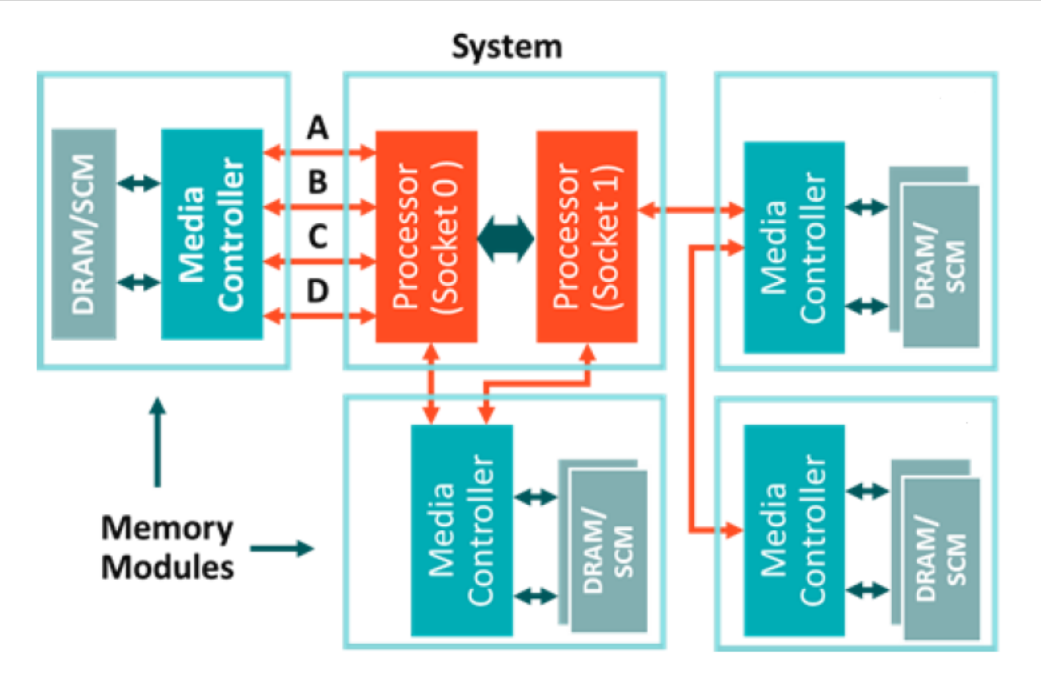


- Media Disaggregation
- Universal memory interface
- High bandwidth, flexible
- Multipath possibility

Memory multipath in Gen-Z and fabric attached memory

- Higher bandwidth, resiliency, robust and flexible topologies
- Multipath, daisy chain
- Reducing load from IPL

- Processors, FPGAs, GPUs can directedly connect to memory.
- Memory module leasing
- Region leasing
- Workload balancing



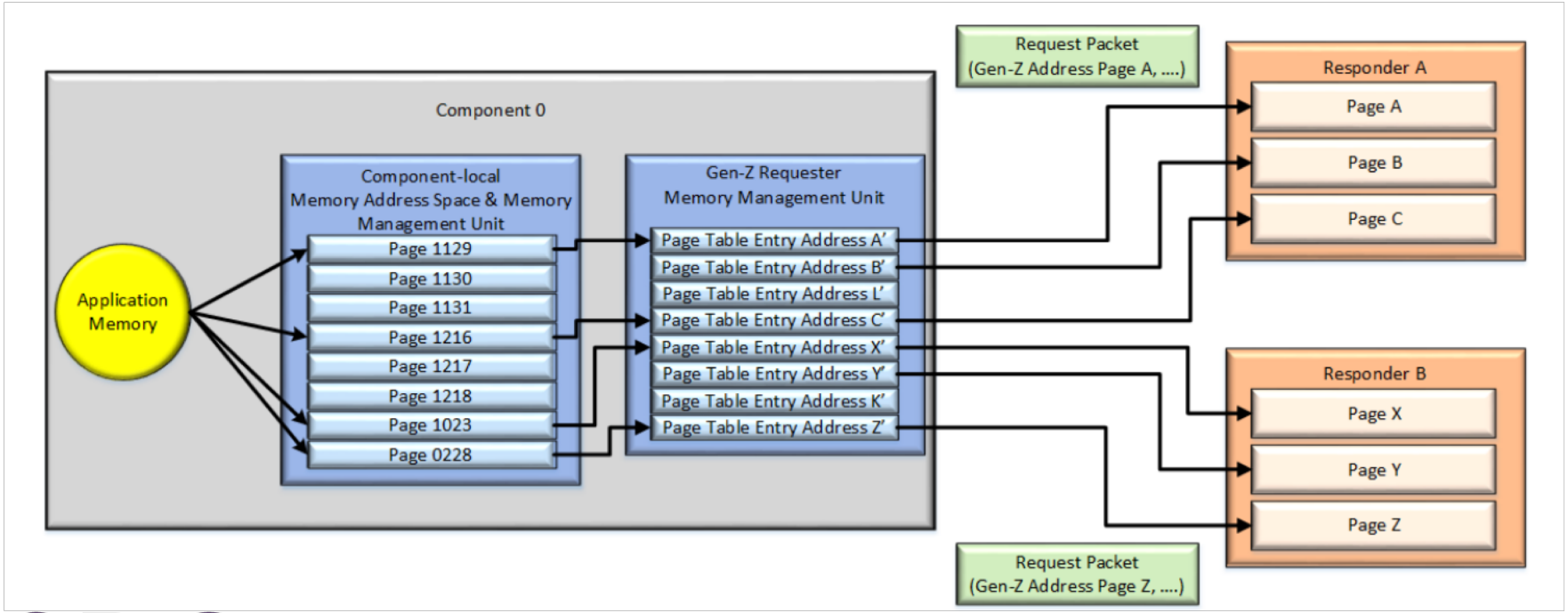
Gen-Z OpClass

Operation: Process of exchanging package(s) between requester and responder to provide a service.

- Multiple OpCode in an OpClass
  - Enables applications to operate on multiple topologies.
  - 32 OpCodes, 35 OpClass.
  - A single component can support 1K+ operation types.
- Part of Gen-Z package.
  - Implicit OpClass, P2P-Core, P2P-Coherency and P2P-Vendor defined
- Core 64 is for switch-base solutions like I/O and memory expansion, 64-bit.
- Atomic OpClass for atomic operation and management different OpClass.
- For bigger data size different OpClass, others for pattern operations, lightweight notification, precision time, unicast packet encapsulation
- Each OpCode represent unique operation
- OpClasses are managed by OpCode set structure, supported and enabled field.

ZMMU

To enable application-transparent access to a Responder's addressable resources, a Requester maps each Responder's addressable resources into the Requester's memory address space.

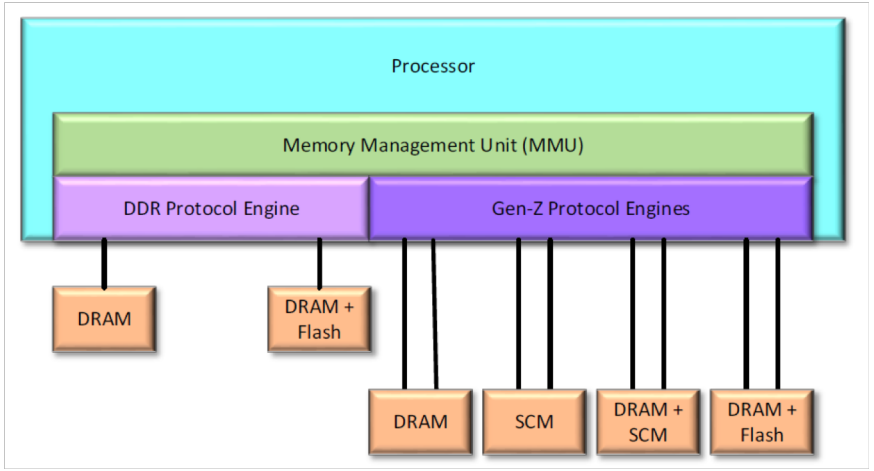


ZMMU …

- Requester ZMMU
  - Map responder address space
  - Application transparent access
- Responder ZMMU
  - Translate package address to media address
  - Enforce access permission (RO R-Key, RW R-Key)

SCM Operation

- Processor has PCIe address range, DDR memory controller address range and Gen-Z address range.
- Any load-store operation maps to any of the address range.
- Later respective protocol will decide outgoing channel/interface.
- Gen-Z has three address range:-
  - P2P-Core address for memory
  - P2P-Coherency address range for I/O, accelerator.
  - Explicit Core 64 FAM (switch, P2P) memory I/O and storage.



Processor and memory attachment

Different package requests

- P2P-Core address range
  - The physical address is mapped to one of the provisioned Requester Bank structures that correspond to a Responder memory module that has been directly mapped by software to a processor physical address range.
  - Using the Requester Bank structure, the physical address is mapped to a specific logical bank, and the Tag that will be used in the Gen-Z request packet is encoded to target the bank and Responder component.
  - An egress interface is selected, and the request is passed to the Gen-Z P2P-Core protocol engine, and a Gen-Z request packet is generated and transmitted.

Different package requests Cont …

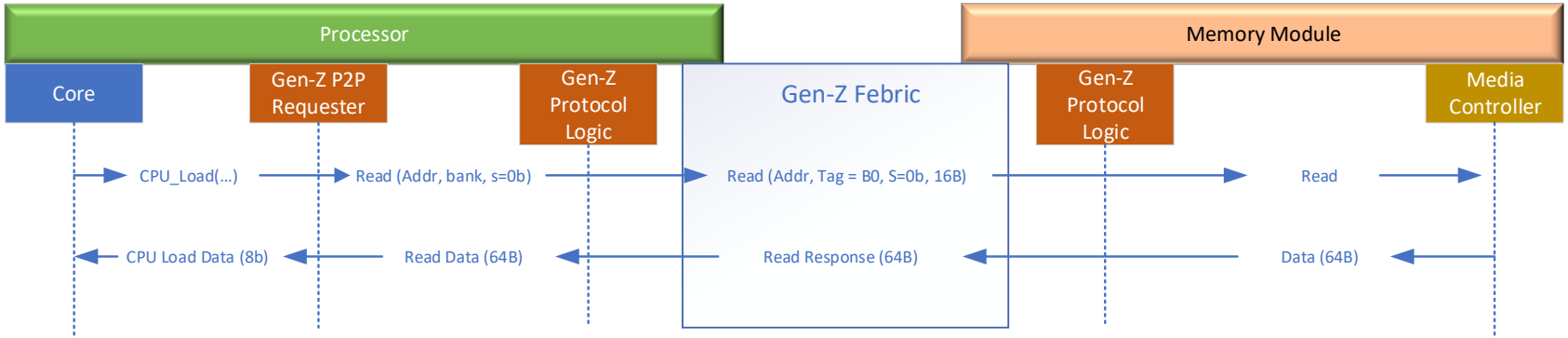
- P2P-Coherency address range
  - The physical address is mapped to one of the provisioned Requester Bank structures that correspond to a Responder (bridge, I/O, accelerator).
  - This has been directly mapped by software (e.g., system firmware) to a processor physical address range.
  - An egress interface is selected, and the request is passed to the Gen-Z P2P-Coherency protocol engine, and a Gen-Z request packet is generated and transmitted.

Different package requests Cont.

- Explicit OpClass
  - The physical address acts as an input into the provisioned Requester ZMMU to identify a Requester PTE (Page Table Entry). Each Requester PTE has been preconfigured by software with all of the information necessary to target a Responder and to translate the processor physical address to a Responder address.
  - Using information derived from the Requester PTE, the request is passed to the Gen-Z explicit OpClass protocol engine, an egress interface is selected, and the request packet is generated and transmitted.

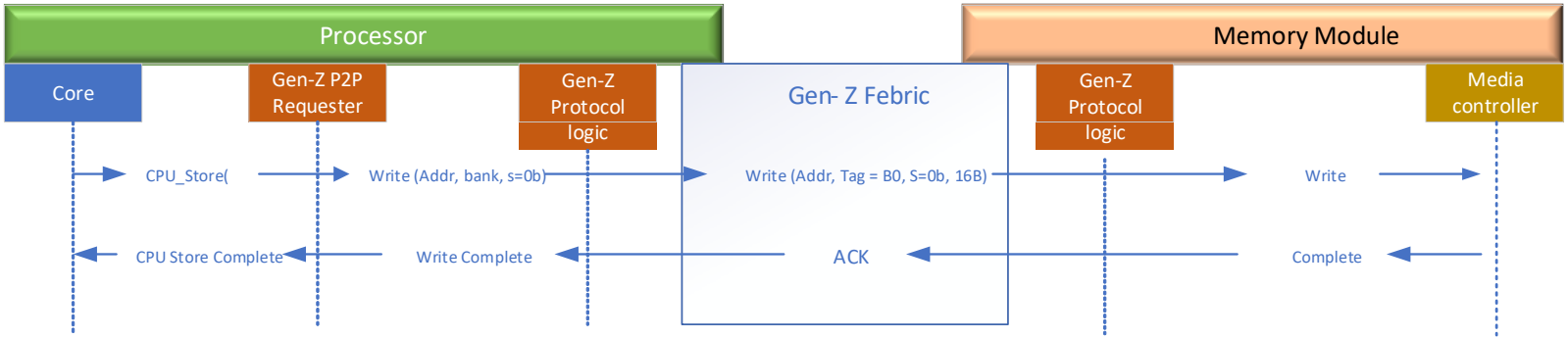
P2P-Core Memory Read

- Memory controller receives the read request and forward the request to Gen-Z logic, which will determine the target address.
- Gen-Z logic determines the egress interface, logical bank, row and generate the P2P-Core read package.
- P2P-Core reads 64 read request type, Gen-Z logic transmit the package.



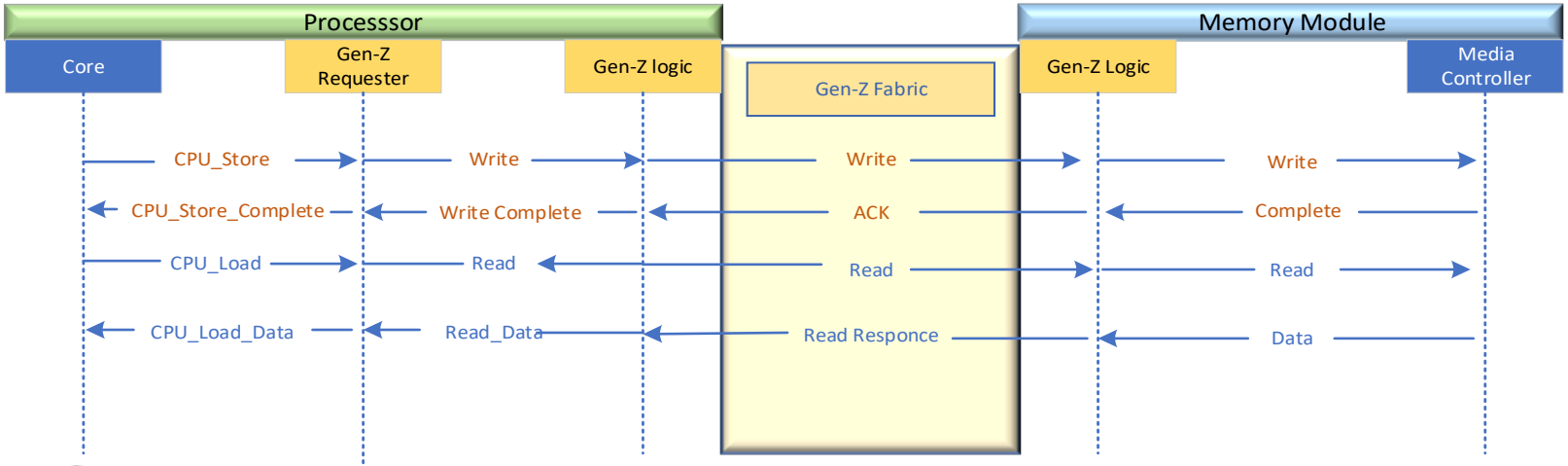
P2P-Core Memory Write

- Memory controller receives the store request, forward the request to Gen-Z logic.
- Gen-Z logic determines the egress interface, logical bank, logical row.
- Memory controller will control the uniqueness of the command.
- Media controller receives and validate the package, transmit the ACK if configured.



Core 64 Write and Read

- ZMMU to identify the component
- Byte addressable



Credits/Acknowledgements

- Gen-Z Consortium Website
  - https://genzconsortium.org
- YouTube Videos
  - Gen-Z Consortium Tutorial Tracks

Thank you

Backup OpClass

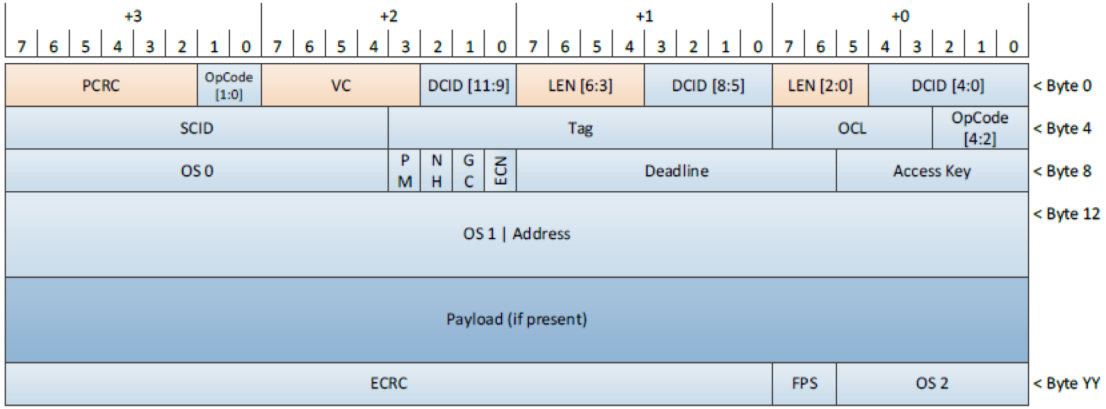


Figure 5-5: Explicit OpClass Request Generic Format

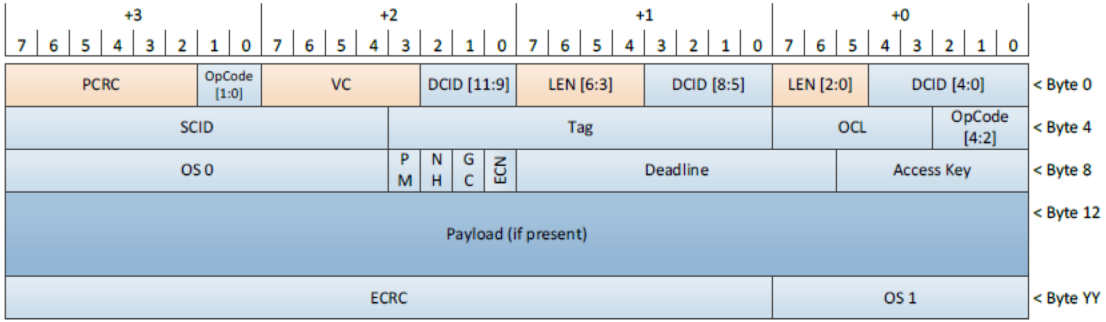
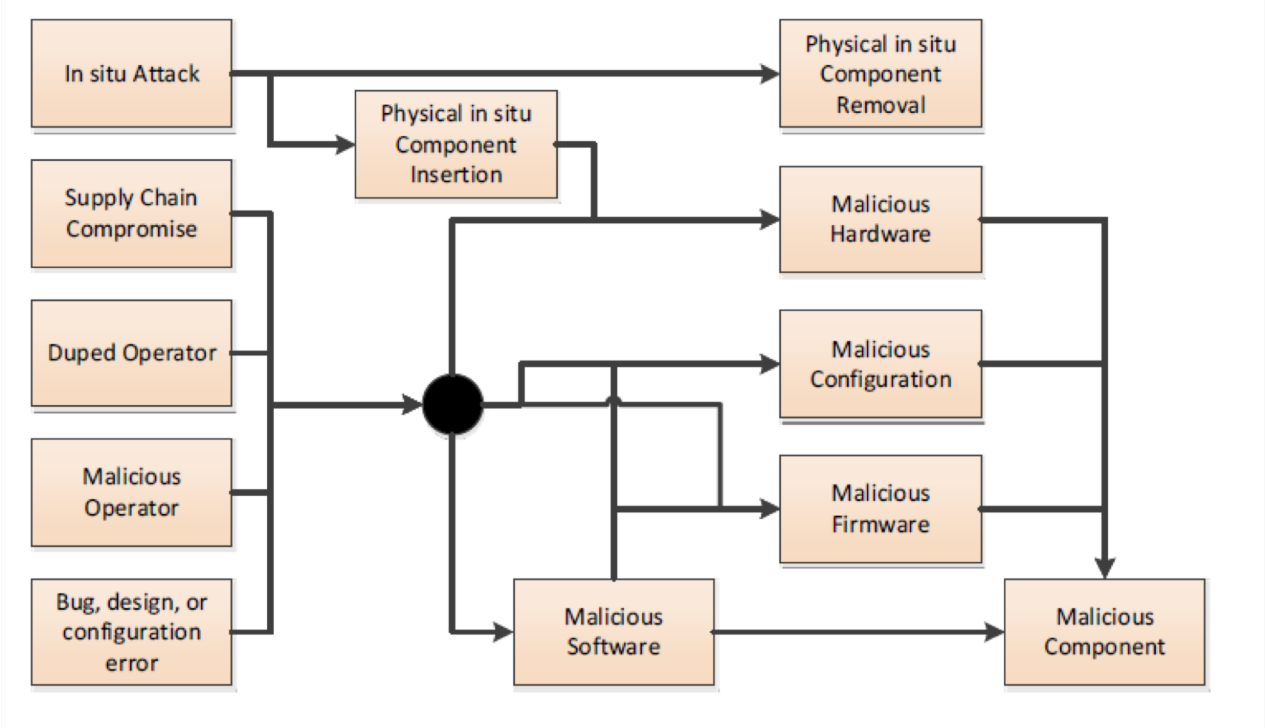


Figure 5-6: Explicit OpClass Response Generic Format

Backup-Security



Backup - Security

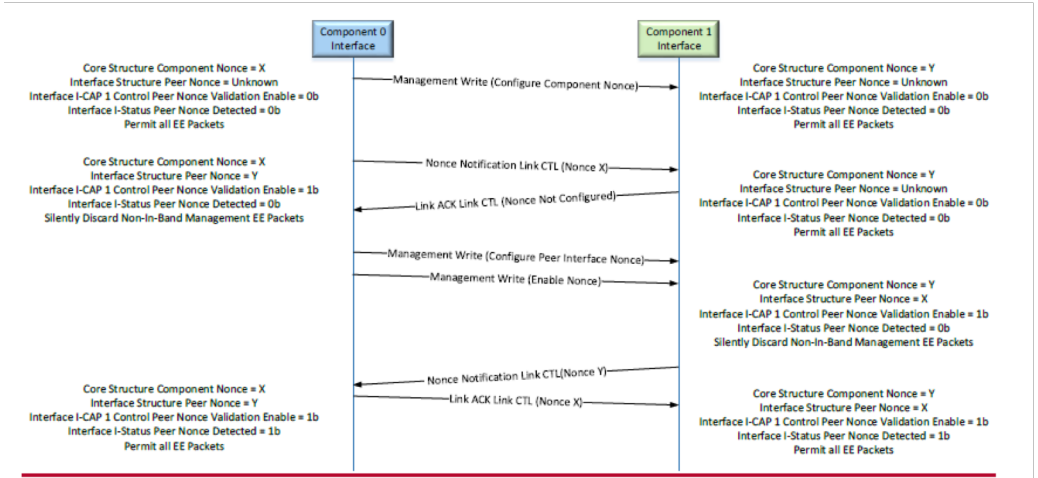


Figure 16-2: Example Nonce Configuration Sequence

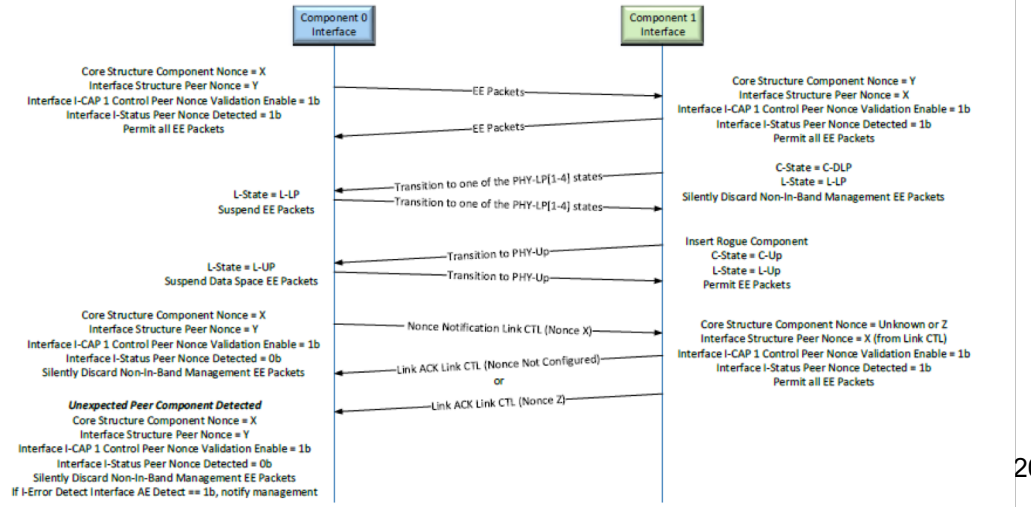


Figure 16-3: Example Unexpected Peer Component Detection Sequence