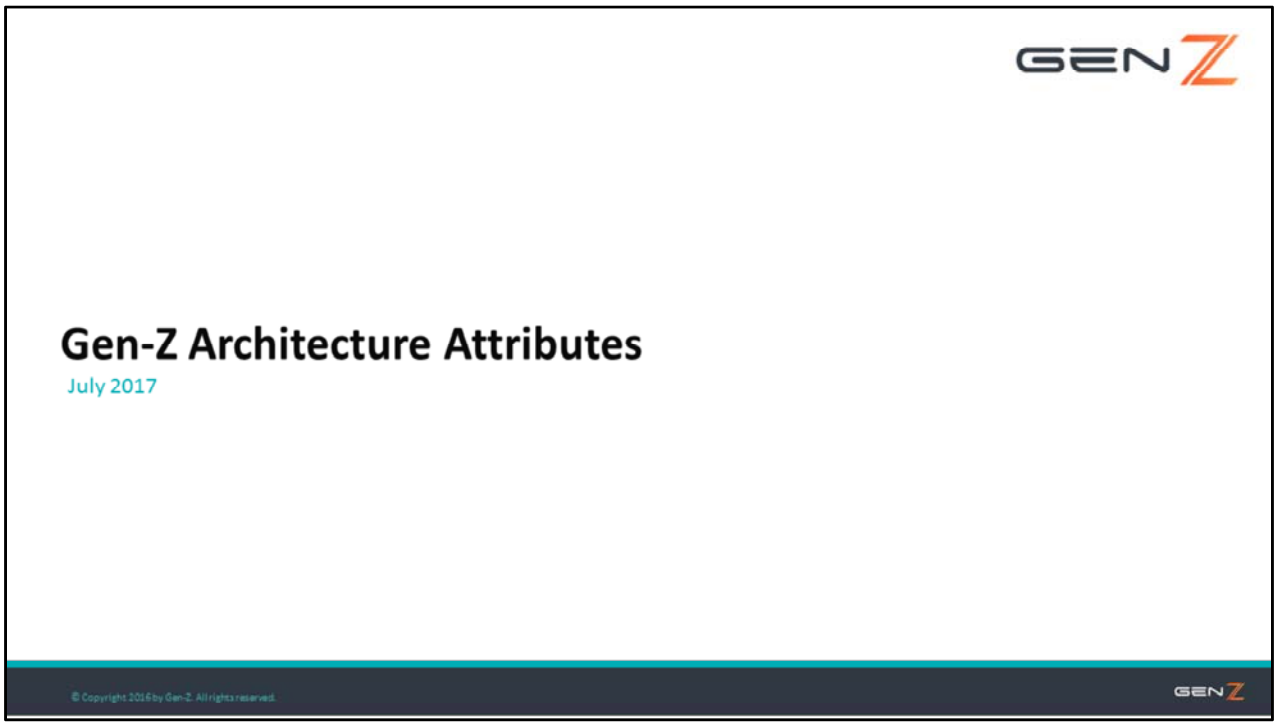


This presentation covers key architectural attributes of Gen-Z.

Disclaimer

This document is provided 'as is' with no warranties whatsoever, including any warranty of merchantability, noninfringement, fitness for any particular purpose, or any warranty otherwise arising out of any proposal, specification, or sample. Gen-Z Consortium disclaims all liability for infringement of proprietary rights, relating to use of information in this document. No license, express or implied, by estoppel or otherwise, to any intellectual property rights is granted herein.

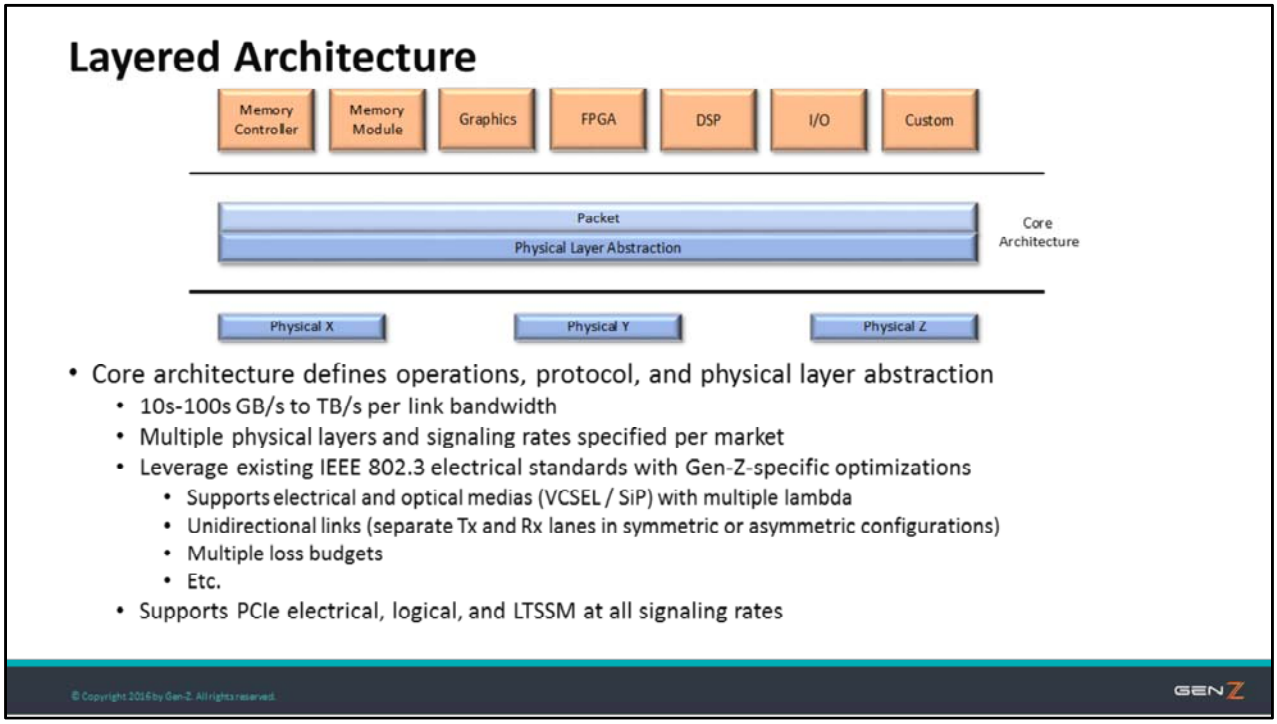
Gen-Z is a trademark or registered trademark of the Gen-Z Consortium.

All other product names are trademarks, registered trademarks, or servicemarks of their respective owners.

All material is subject to change at any time at the discretion of the Gen-Z Consortium

http://genzconsortium.org/

© Copyright 2016 by Gen-Z. All rights reserved.



Gen-Z specifies a layered architecture:

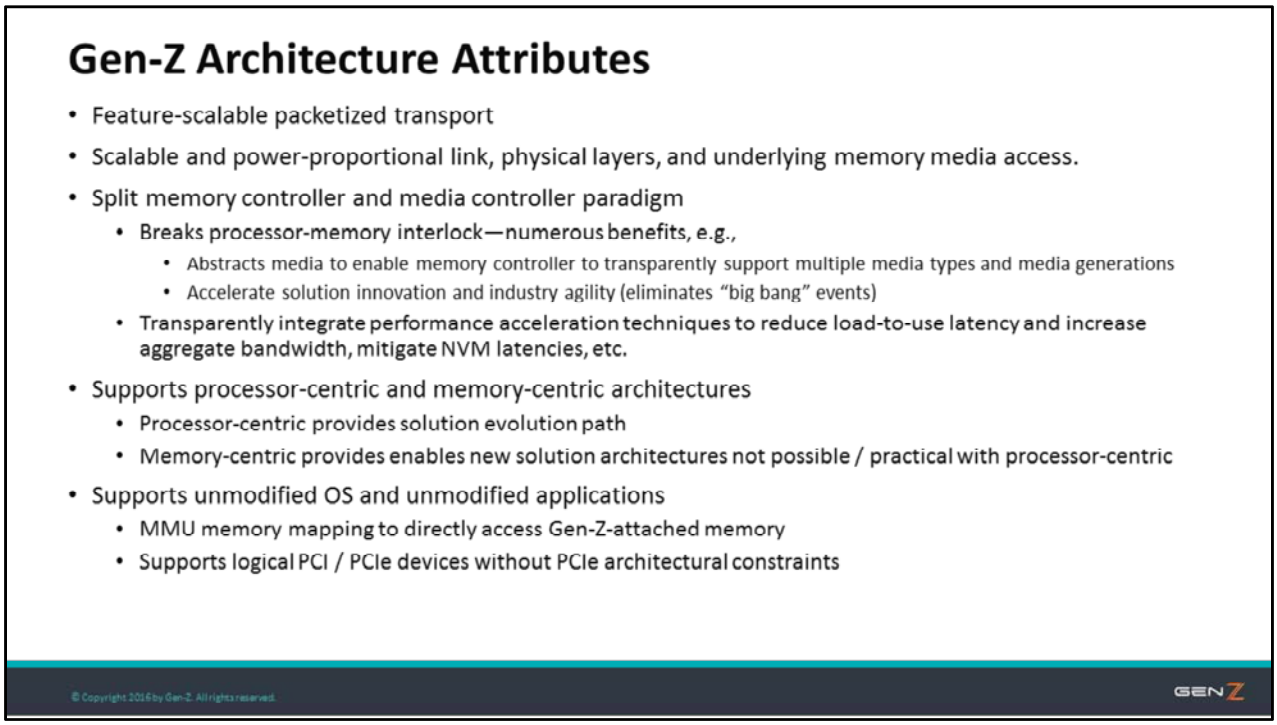
- The packet layer covers link-local and end-to-end protocol packets.
- The physical layer abstraction (PLA) abstracts the physical layer-specifics from the higher layers. This enables a design to operate over multiple physical layers.

Gen-Z's architecture is capable of scaling to up to 448 GB/s of raw read-write bandwidth (32 differential pairs operating at 112 GT/s PAM 4 signaling). Components that support multiple links using two 4C connectors could provide nearly 1 TB/s of bandwidth. Future solutions will be able to scale to multiple TB/s.

Gen-Z supports two physical layers that can be instantiated as electrical or optical medias (VCSEL or silicon photonic). The industry has extensive experience with these physical layers, and robust ecosystems have been established.

- Gen-Z leverages the IEEE 802.3 electrical specification. The 802.3 electrical supports any mix of Tx and Rx lanes to optimize communications to meet workload needs, e.g., read-dominate or balanced communications. The 802.3 electrical supports multiple loss budgets to provide more design flexibility, optimize implementation size and power consumption, and reduce complexity and cost.
- Gen-Z support the PCIe electrical, logical and LTSSM at all supported signaling rates (up to 32 GT/s once PCIe 5 is specified)

Gen-Z can also be used in co-packaged solutions, e.g., co-packaged memory. The physical layer is TBD at this stage, though it could leverage the existing co-packaged physical layers.



Gen-Z uses a packet protocol to communicate between components. This enables the protocol to scale from co-packaged, to embedded, to single enclosure, to rack scale.

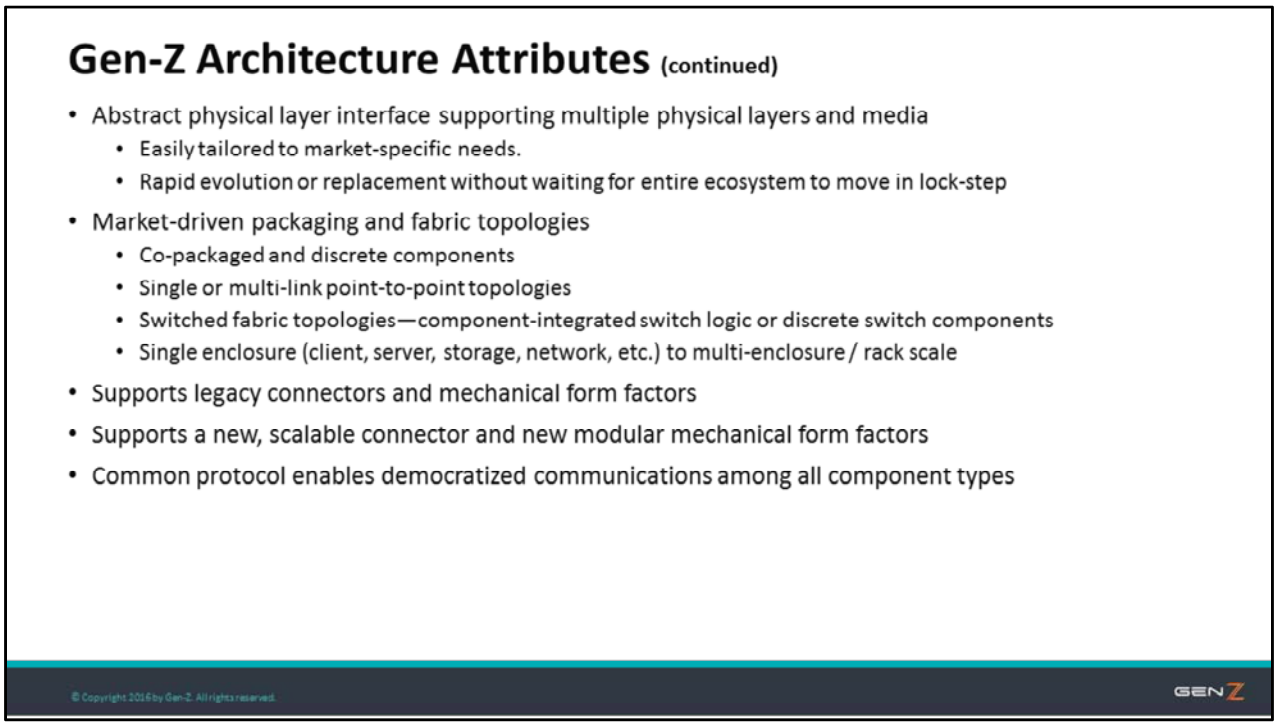
Links can be composed of any number and mix of Tx and Rx lanes and use physical layers that support multiple signaling rates to provide workload-optimized performance.

The architecture supports a split memory controller and media controller functional paradigm. The memory controller takes care of all high-level memory operations, and the media controller takes care of all of the media-specific logic. The two communicate with one another using the Gen-Z protocol. There are numerous technical and business advantages in supporting this split model (see the overview specification for additional details).

The architecture supports today's processor-centric architecture, i.e., where most communication passes through an application processor. This enables existing solutions to transition to Gen-Z with minimal disruption to the overall ecosystem. The architecture also supports a memory-centric architecture. In a memory-centric architecture, components can talk directly to the memory or each other without going through the application processor. A memory-centric architecture reduces the amount of data movement / power consumption / latency / etc. Further, for a number of workloads, 90+% of the data does

not need to be coherently exchanged, hence, flowing through the processor provides no benefit, and can increase solution CAPEX / OPEX costs.

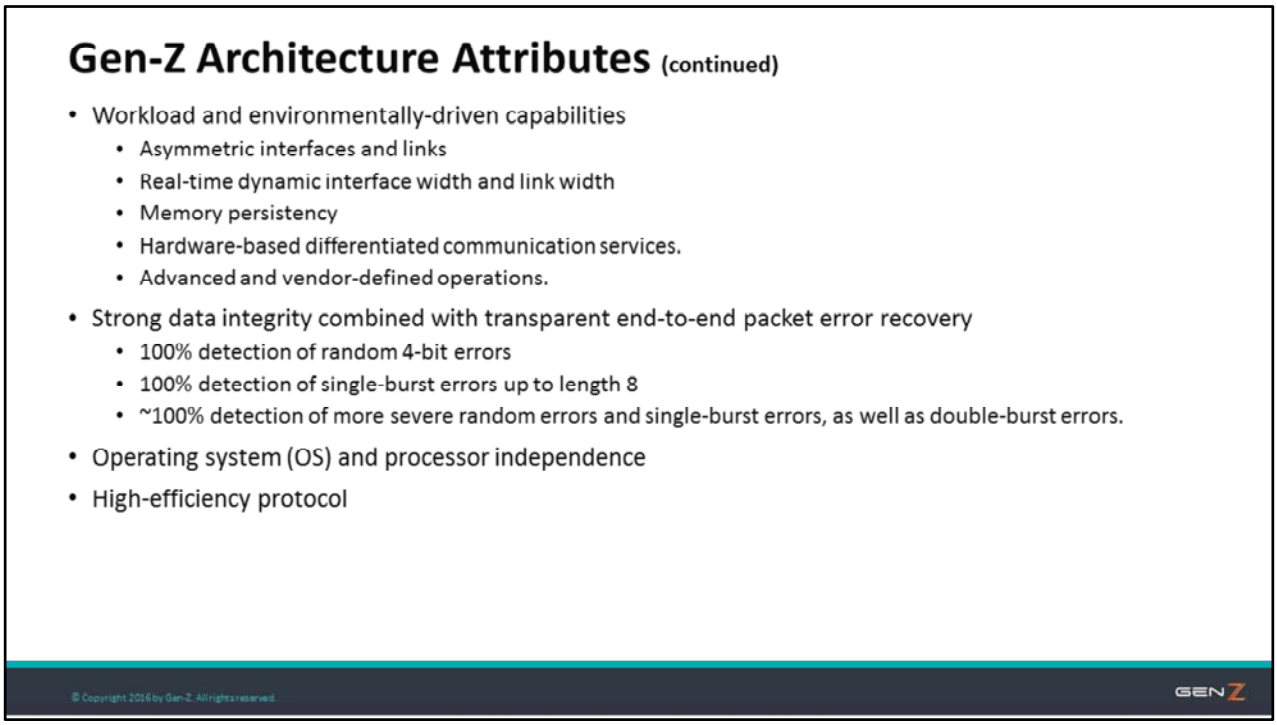
Gen-Z's architecture enables components to be inserted into any solution stack without requiring OS or application modification. For example, a memory component can be mapped just like DDR or HBM memory, and be transparently accessed using load-store semantics. Similarly, Gen-Z specifies logical PCI / PCIe device support that enables I/O components to take advantage of Gen-Z's architectural capabilities to provide customer-visible value, e.g., multipath to provide aggregate bandwidth and resiliency (PCIe supports a single link), full set of atomic operations (PCIe supports just a couple of atomics), multi-node I/O sharing and scale I/O virtualization sharing without PCIe SR-IOV / MR-IOV constraints, etc..



Gen-Z specifies the Physical Layer Abstraction (PLA). The PLA enables designs to operate across multiple physical layers without requiring physical layer-specific changes. This enables designs to be quickly tailored to market-specific needs, e.g., a media controller design can be quickly instantiated in a co-packaged solution, a single enclosure solution using an electrical PHY, and a multi-enclosure solution using photonics.

The architecture supports multiple package options and fabric topologies. This enables Gen-Z to be used across market segments and use cases.

Gen-Z supports a wide-range of mechanical connectors and form factors including: U.2 / U.3, PCIe CEM, OCP mechanical form factors, a new scalable connector developed by the Gen-Z mechanical workgroup that was recently transferred to the SNIA / SFF industry body to maximize industry adoption, and a new modular mechanical form factor intended to reduce mechanical fragmentation.



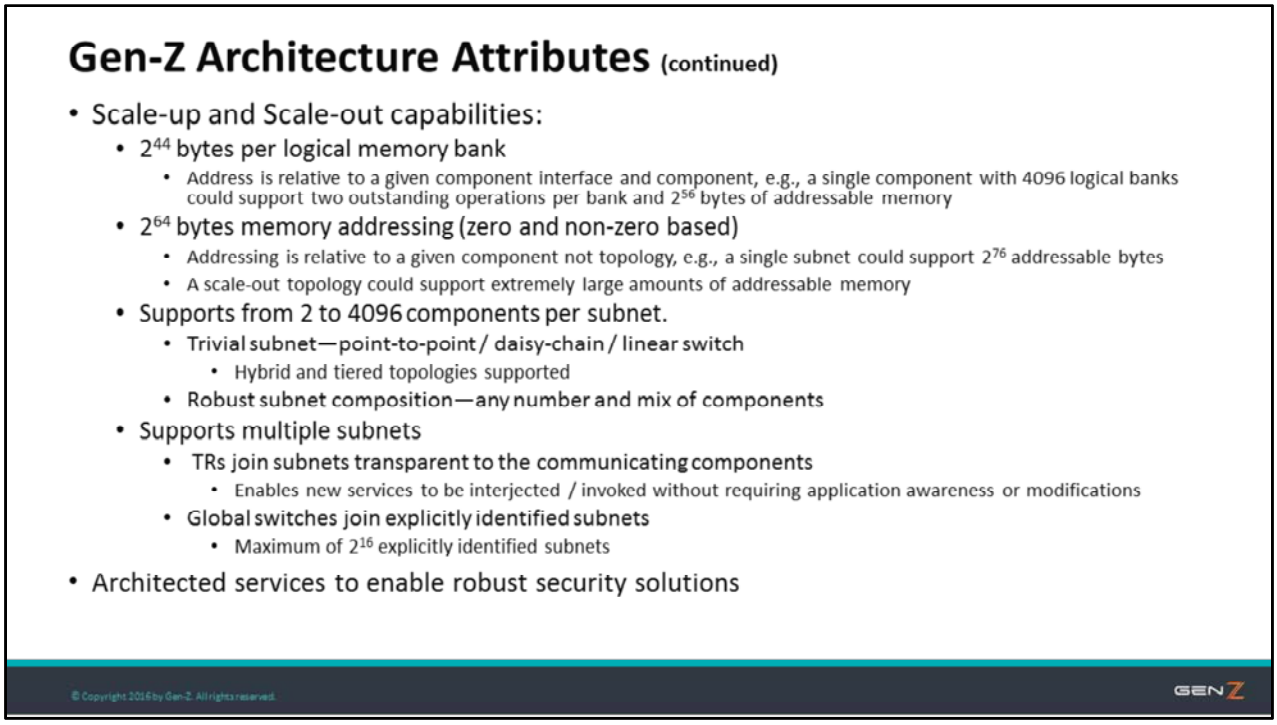
Gen-Z supports workload and application environment capabilities, e.g.,

- Asymmetric interfaces and links to adjust read and write bandwidths to match workload needs without requiring solutions to provision wider links and connectors.
- Real-time dynamic interface and link widths—the architecture enables a link to burst traffic and quickly turn off unneeded lanes. This technique can reduce PHY power consumption by up to 80%.
- Support for memory persistency—per write operation or persistent flush to make all outstanding writes persistent.
- Quality of Service (QoS) services to prioritize traffic as well as adapt to fabric load conditions, e.g., take advantage of multipath options to bypass congestion or failed hardware
- Supports a variety of advanced operations including coherency, collectives, buffers, etc. to accelerate workloads. Also supports up to 8 vendor-defined OpClasses to easily customize communications.

Gen-Z supports robust data integrity to detect a variety of transient and burst errors. It also supports reliable delivery and link-level reliability to reduce application communication overheads.

Gen-Z is OS and processor independent to enable seamless adoption

Gen-Z protocol is very efficient—for example, ~78-83% efficient in point-to-point topologies for 64B moves and 90+% efficient for 256B moves in any topology



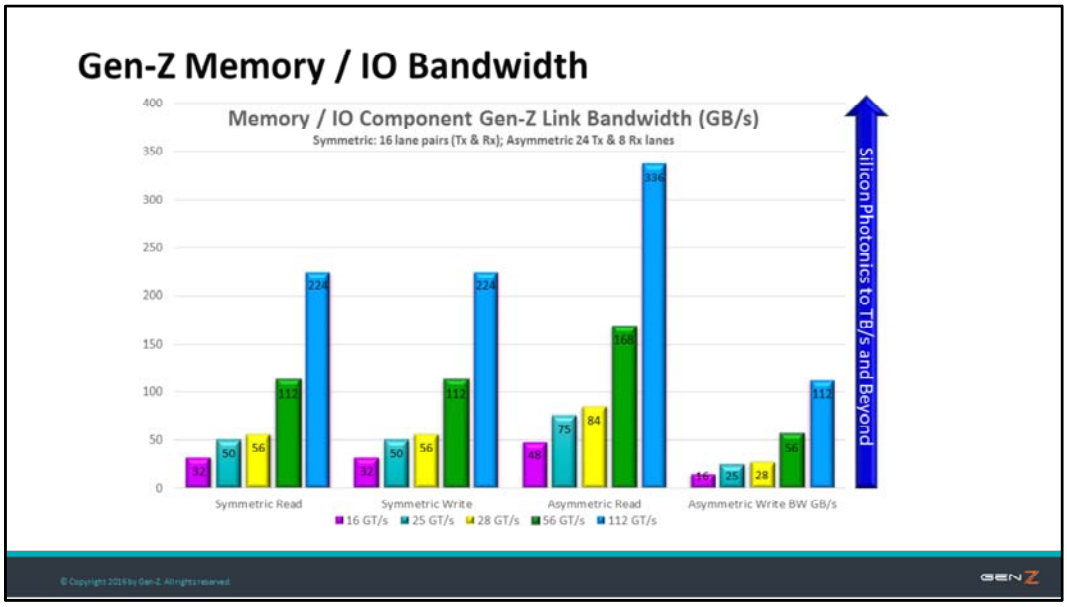
Gen-Z is extremely scalable. For example, a point-to-point optimized memory component can scale to 2e56 bytes of addressable memory and up to 4096 logical banks. Logical banks enable components to optimize memory access without being constrained by the limited number of physical banks. This enables greater parallelism and higher bandwidth to be achieved. Logical bank designs can incorporate a variety of performance optimizations to reduce latency, reduce power, improve resiliency, simplify wear-leveling, etc. A switch-capable memory component can support up to 2e64 bytes of addressable memory. Switch-based memory components do not explicitly support logical banks since they can be accessed by multiple Requesters, e.g., multiple processors. Instead, they advertise a large addressable space and the media controller transparently takes care of the logical bank operation, and it can apply the same techniques as used in the point-to-point memory component to increase parallelism and performance.

Gen-Z supports simple topologies with as few as two components. It can scale to support larger topologies, e.g., a single subnet can support up to 4096 components, and multi-subnet solutions can support up to 2e16 subnets, or up to 2e28 components (over 256 million components) and up to 2e92 bytes of addressable memory.

Gen-Z also supports transparent routers (TRs). A TR can be used to transparently join subnets, e.g., a TR appears as a Responder with a large amount of addressable resources

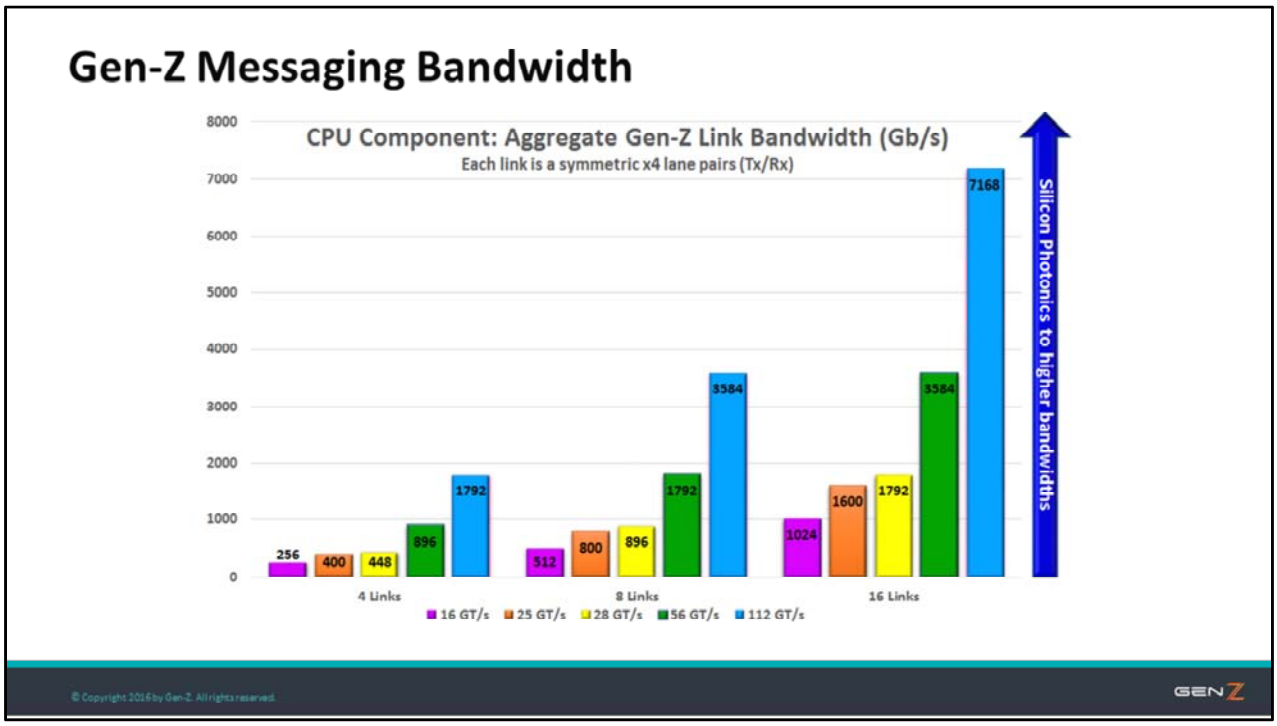
which are actually distributed across multiple memory or storage components. Their transparency enables a variety of capabilities without requiring application or middleware modifications.

Gen-Z was designed with built-in security services to address ever growing cyber-threats.



Gen-Z supports multiple signaling rates that when combined with a variable number of transmit and receive lanes, can scale to meet demanding application needs. The bandwidths shown can be achieved using electrical or optical physical layers. Long-term, as silicon photonics evolve and start to displace electrical, Gen-Z's architecture will enable solutions to scale bandwidth to multi-Terabyte/s ranges.

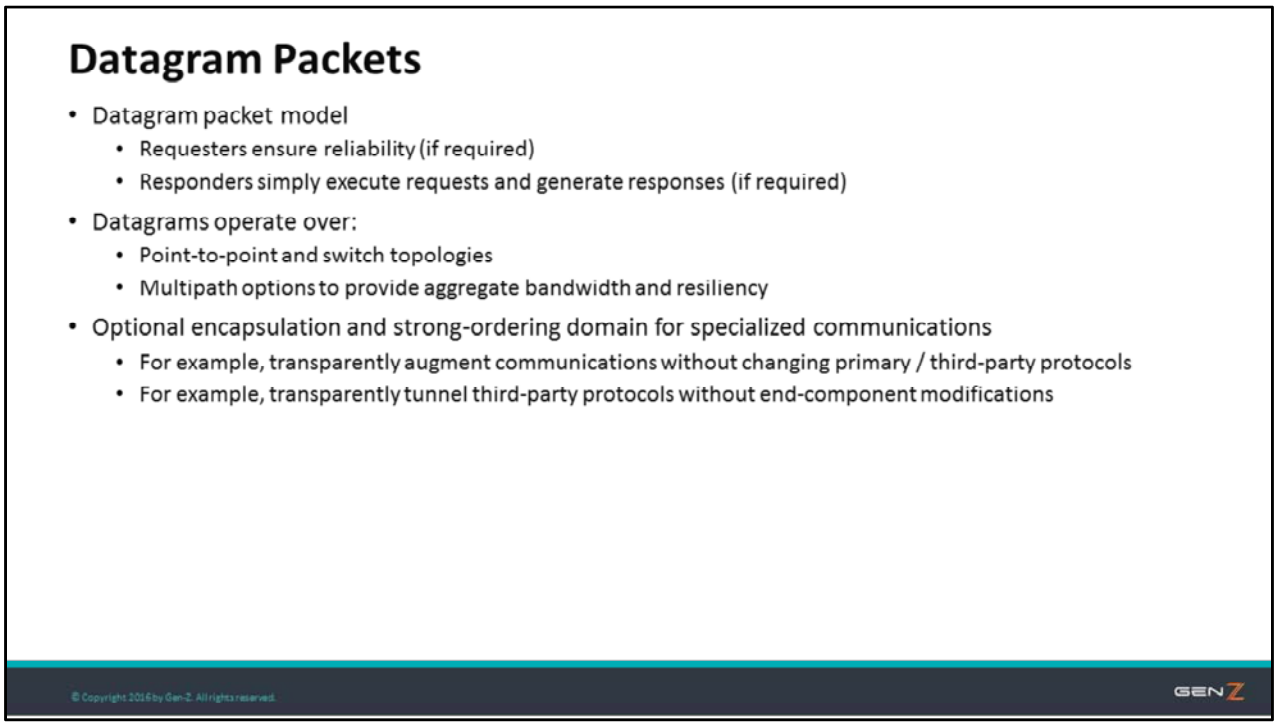
Gen-Z supports symmetric and asymmetric links. A symmetric link contains the same number of transmit and receive lanes. An asymmetric link contains a different number of transmit and receive lanes. For example, many workloads are read-dominant, and often have a 3:1 or higher read-write ratio. An implementation that uses asymmetric links can deliver 50% more read bandwidth than a symmetric link using the same total number of lanes (i.e., no mechanical or physical changes required). As a result, asymmetric links are more flexible, more cost-effective, and more power efficient than symmetric links that require additional lanes to deliver equivalent bandwidth.



The same physical layers and signaling rates are equally applicable to messaging. Though links can support both symmetric and asymmetric configurations to adapt to workload needs, most networks use symmetric links. This could change in the future.

Traditional NICs typically support a single PCIe link to attach to a host and emit 1-2 Ethernet links to attach to the network fabric. A host that integrates Gen-Z or coherently-attaches a discrete Gen-Z bridge can support significantly more links, e.g., 8 or 12. Further, since Gen-Z enables light-weight switching designs, a switch can be integrated into the host / bridge. This provides multiple benefits:

- Simple, efficient peer-to-peer communication between links without requiring host involvement or consuming coherency fabric resources.
- Flattens switch topologies—hosts can be meshed together without requiring discrete switches
- Enables support of advanced routing topologies, e.g., hyper-x, that reduces the number of switch hops between components, thus enabling greater scale with fewer switches and lower latency
- Enables transparent workload segregation and load-balancing to improve QoS and reduce the probability of congestion events
- And more.

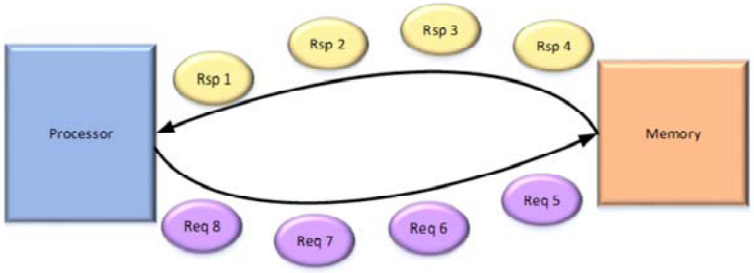


The majority of packets are exchanged as datagrams. Datagram communications provide numerous benefits, e.g.,:

- Support multiple topologies
- Support multipath and real-time adaption to fabric load and operating conditions
- Reduce hardware resource requirements
- Simplify hardware design, lower implementation cost
- Datagrams are used by numerous applications due to ability to easily scale up and out
- Etc.

Gen-Z supports packet encapsulation to enable third-party protocols to be tunneled across Gen-Z, as well as to enable Gen-Z operations to be tunneled in specific situations (this simplifies design and maximizes re-use).

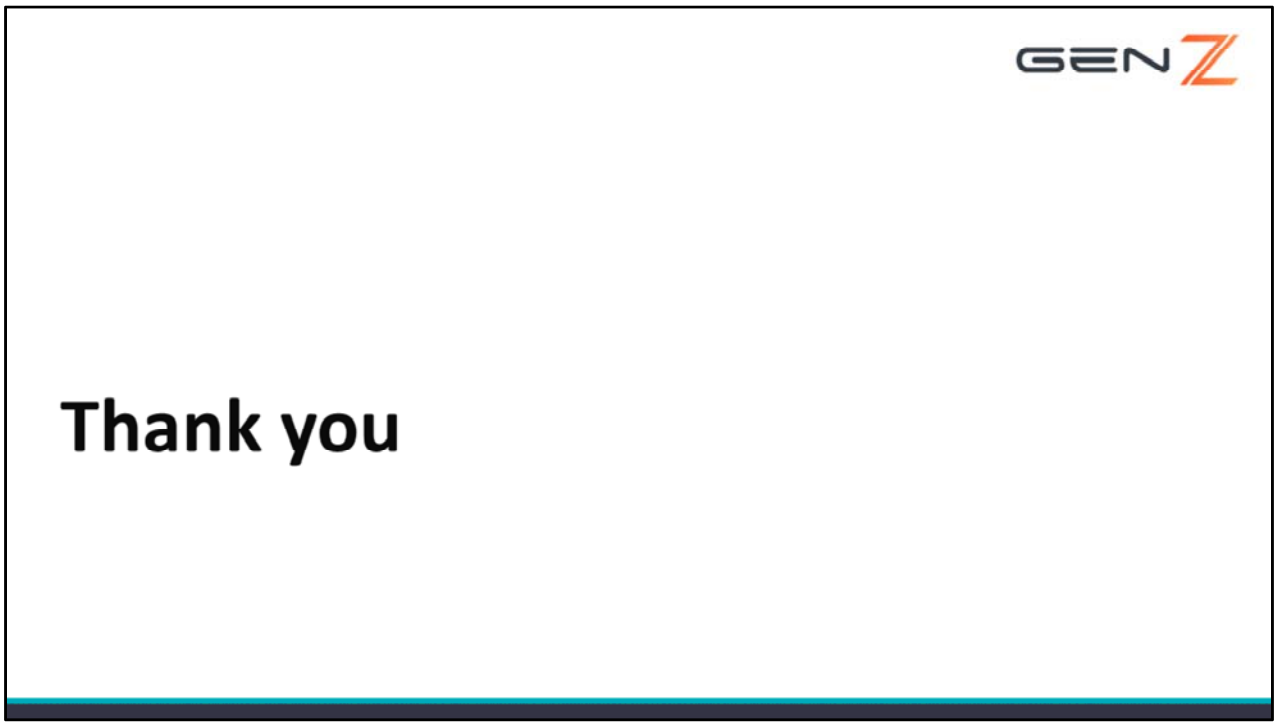
Asynchronous Communications



- Non-blocking
  - Packet pipelining
  - Asynchronous completions
- Low-latency and long-lived ops
- Simplified memory controller
  - Comprehends ordering, multipath, etc.
- Simplified media controller
  - Optimized for internal media architecture
- Naturally tolerates variability
  - Media material latencies (read / write)
  - Internal media contention
  - Diverse topologies
  - Fabric load
- Naturally enhances resiliency
  - Transparent end-to-end reliability
  - Transparent fail-over and recovery
  - Simplified serviceability

Gen-Z Confidential
© Copyright 2016 by Gen-Z, All Rights Reserved

Gen-Z uses asynchronous communications to improve performance, simplify hardware design and implementation, account for variable media and operating conditions, etc. This slide lists just some of the benefits and operational aspects.



This concludes this presentation. Thank you.