Gen-Z Invigorating Management Architecture Inducing a Change in the Ecosystem

Parmeshwr Prasad
DellEMC
Gen-Z

- Broad industry support
- Supports fast adoption
- PCIe and IEEE 802.3 PHY
- Spec: Core, Mechanical, Physical, Connectors and Management.
Benefits of Gen-Z

- Solution scalability
- Secure architecture
- Resiliency support
- Architectural flexibility
- Unifying data path
Gen-Z Operation

- **Requester**
  - Generate package, Ordering, identify egress channel, reliability, validate and execute response package.

- **Responder**
  - Validate and execute request package, generate response, identify egress channel
Management agenda

Management stack
Gen-Z Managers
Type of management
Discovery/Enumerations/Security/Power
Example
Management software

- No one “owns” any resource
- HW as resource
- Management as per “grand plan”
- Local management service on each resource
Gen-Z management overview

_stack_

Composability manager

Resource Managers, Services, & Monitors

Fabric (or local) Manager

HW init

Device setup

Low-level Fabric setup / Authentication & Routes

Addressing / Firewalls / Partitioning

HW Infrastructure Management Services

Composable Infrastructure Management Services

Data Infrastructure Management Services

Orchestration Tools

Internal Use - Confidential
## Gen-Z management overview Cont...

<table>
<thead>
<tr>
<th>Device Setup</th>
<th>Low level component setup operations such as asserting power, clocks, and resets. Establish component hardware defaults, additional initial state, and link states as required by the platform to prepare components for being fully configured by an in-band Gen-Z manager.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low-level Fabric Setup / Authentication and Routes</strong></td>
<td>Establishing physical connections among components.</td>
</tr>
<tr>
<td><strong>Addressing / Firewalls / Partitioning</strong></td>
<td>Enabling specific endpoints to access specific logical resources using the available physical connections.</td>
</tr>
<tr>
<td><strong>Data Infrastructure Management Services</strong></td>
<td>The general name of the collection of services that manage the persistent data and the logical (data) address space of the Gen-Z fabric.</td>
</tr>
<tr>
<td><strong>Hardware (HW) Infrastructure Management Services</strong></td>
<td>The general name for the collection of services that manage physical resources and their lifecycles.</td>
</tr>
<tr>
<td><strong>Composable Infrastructure Management Services</strong></td>
<td>A central management toolset that abstracts the layers below it into pools of resources and establishes logical servers/platforms and binds them to data structures, other servers, accelerators, networks, boot images, storage and persistent main memory.</td>
</tr>
</tbody>
</table>
Management assumption and tenets

- Primary and secondary manager, both in sync.
- Should support unreliable control write MSG opcode
- Should support manipulation of package contents at all stage of processing
- Should support completion handler
- Should support setting manager type
Control Structure

- Access enforcement
- In-band/OOB accessibility
- Pointer access mechanism
- Core structure at address zero.
Types of Gen-Z Management
In-Band Management

- Access control space
- CTL-read and CTL-write
- P2P OpClass
- Configuration progress measurement.
- Event management
Out-Of-Band management

- I2C or something else
- Managers are impacted
- Platform and device specific
- Certain new changes are required
Manager types
Different managers

- HW Init Manager
- Resource Manager
- Primary Manager
- Fabric Manager
- Composability Manager
Managers role

- **Primary manager:** Manages same physical assembly (such as rack/enclosure)
- **Fabric Manager:** Manages multiple enclosures
- **Resource Manager:** logical abstraction based management such as SCM, FAM, Accelerator.
- **Manager/Director Domain:** Kind of namespace
Composability Manager

- Initial ownership of all resources
- Help to bind a resource to resource manager
- Enforce security policies at bigger level
- Coordinate creation and allocation of R-keys
Zones/Directors

- Manager Zones
  - logical collection of fabric manager domains
  - Aim to improve efficiency
- Fabric Directors
  - Creates hierarchical management zones
  - Separate logical manager service
Component Ownership

- PMCID or PFMCID can be filled by capturing SCID.
- Existing manager can set type of manager of newly added manager.
- Control field can be set to valid value after power reset.
- Existing manager can help other manager to discover any component.
Discovery
Discovery/Enumeration

- In-band, OOB and side band discovery
  - Connected and power-up state (L-UP State)
  - OOB by “presence pin” connection
- Enumeration collects base class, supported capability and current operating state.
  - Redfish and MCTP capability for Gen-Z components.
Configuration Management

- Hand over the capability to upper stack.
- Primary manager/Fabric manager should bring component to usable state.
- Rich in-band management support
- Change only control space
Power Management

- Physical layer power management
- Component/Resource power management
- Software-directed power management
  - OpClass c-state power control package
  - Policies for power management architecture
Linux changes

- ZMMU, IOMMU
- Sysfs changes
- Interrupts events
- Gen-Z bus and device tree
- Managed node sysfs, fabric manager node sysfs
Standalone system Management

Local Fabric
- C2, C3 media controller, byte addressable, core64 opClass
- C4, C5: media controller, block addressable, core64 opClass
- S0: Gen-Z Switch, enclosure class
- B1/SB1: Gen-Z bridge and integral switch (same CID), core64 opClass
- SoC: Any compatible CPU/SoC
- C1: media controller, byte addressable, P2P opClass.

Management
- Primary Manager
Namespace and topology creation

- Primary manager discovers and enumerates the control space, in control of Gen-Z CID namespace, controls all Gen-z component control and data space.
- Receives all errors, signals, events, and signals related to any component.
- Primary manager creates JSON/redfish inventory, supports control function from power on.
- Exploration, enumeration, and configuration by the primary manager.
Discovery and enumeration

- Primary manager establishes control for all components.
- Avoid race condition (multipath, other manager)
- Bridge enumeration (B1)
- Establish OOB management
- Enables appropriate messaging, error and event signaling from bridge control structures.
Memory

- HW prep enables in-band management, sets default attributes (C1, C2, C3, C4 and C5).
- Discover media controller, write bridge’s ZMMU WRT media controller, RO and RW keys.
- Discover media controller of persistent memory.
Component and subsystem integration

- Discover components in single manager domain, get total data.
- Create Gen-z control plane
  - Creates graph form (vertex, edge, root)
- Create Gen-z data space
- Inventory handoff
Security

- No hardware or firmware is allowed onto the fabric without a physical presence of an authorized person who is responsible for validating the certificates, authorizations, versions, physical identities.
- R-Keys are used to support fine grained.
- Resource Managers map the logical resources requested into physical resources.
- Fabric Manager is responsible for actually installing the requested Access Keys, and R-Keys / R-Key Domains throughout the fabric and the responders and requestors.
References

• https://genz.causewaynow.com
• https://genzconsortium.org/
• Gen-Z software
  • https://github.com/linux-genz
Thank You 😊