STORAGE DEVELOPER CONFERENCE

SDC 23

BY Developers FOR Developers

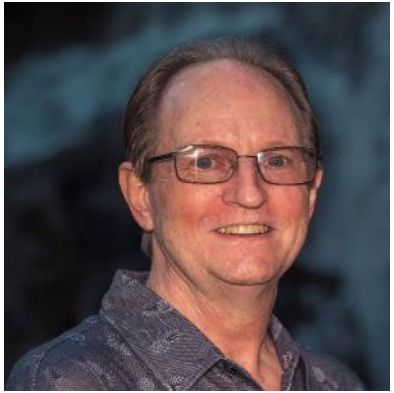
NVM Express® Computational Storage

Presented by Kim Malone and Bill Martin

Speakers

Kim Malone

intel



Bill Martin

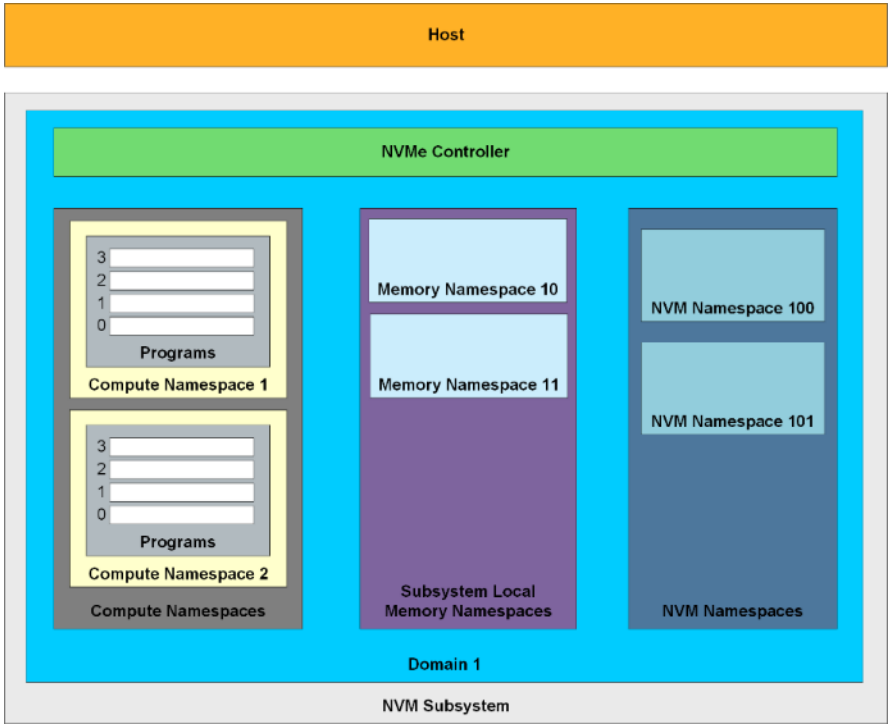
SAMSUNG

SDC 23

Agenda

- Major architectural components
- Memory Namespaces
- Copy
- Compute Namespaces
- Computational programs
- Program Unique Identifier (PUID)
- Memory Range Sets
- Reachability
- Reachability Association example
- Example flow

Major Architectural Components



The NVM Express® (NVMe®) computational storage architecture involves several types of namespaces:

- Compute namespaces (new)
- Memory namespaces (new)
- NVM namespaces
  - NVM, Zoned and Key Value namespaces

This presentation discusses NVMe technology work in progress, which is subject to change without notice.

Memory Namespaces

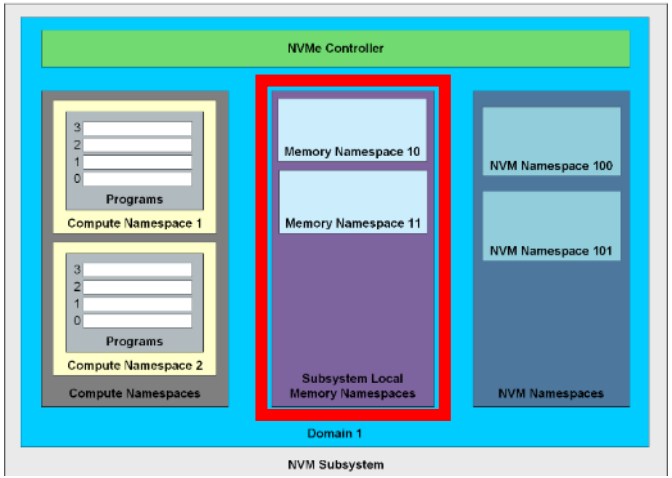
A memory namespace:

- Provides host command access to memory in the NVM subsystem
- Is associated with the Subsystem Local Memory I/O command set
- Is used by the Computational Programs command set to provide access to SLM for program execution

TP4131: Subsystem Local Memory (SLM)

New Subsystem Local Memory I/O command set for memory namespaces

- New commands include:
  - Memory read and memory write
    - Commands for transferring data between host memory and a memory namespace
  - Memory copy
    - Command for copying data from NVM and memory namespaces to a memory namespace
- Host accesses are dword addressable & dword granular
- Compute Namespace accesses are byte addressable & byte granular



This presentation discusses NVMe® technology work in progress, which is subject to change without notice.

Copy

Memory Copy command:

- Defined in SLM command set
- Copies from NVM namespaces or from memory namespaces to a memory namespace
- Copies from NVM namespaces to memory namespace
    - Total length has to be on granularity of LBA size
    - Does conversions from blocks to bytes
- From memory namespaces to memory namespace
    - Total length has to be dword granular

NVM Copy command:

- Defined in the NVM command set
- Copies from memory namespaces or NVM namespaces to an NVM namespace
- From memory namespaces to NVM namespace
    - Does conversion from bytes to blocks
- Total length has to be on granularity of LBA size

Compute Namespaces

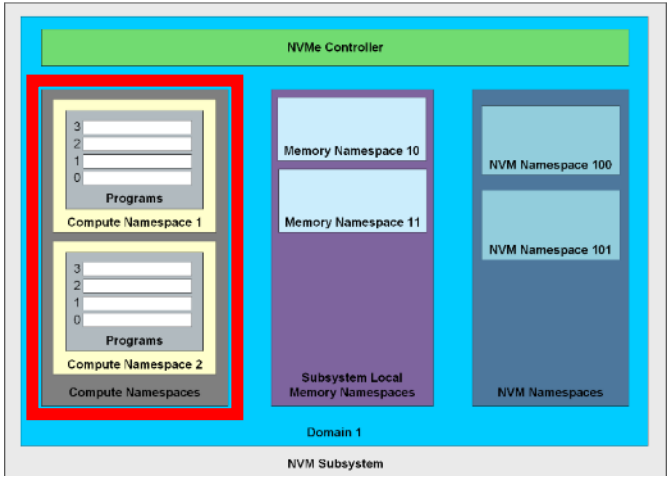
A compute namespace:

- Is able to execute one or more programs
- Is associated with the Computational Programs I/O command set
- Contains compute resources

TP4091: Computational Programs

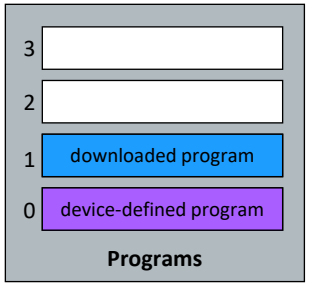
New Computational Programs I/O command set for compute namespaces

- New commands include:
  - Load program
  - Activate program
  - Create/Delete Memory Range Set
  - Execute program
- Provides log pages for program discovery



Computational Programs

- Conceptually similar to software functions
  - Called with parameters and run to completion
- Are addressed via a compute namespace program index
- May be identified by a globally unique program identifier
- Operate only on data in Subsystem Local Memory
- A program may only be able to execute on certain compute namespace(s)
- May be device-defined or downloaded
  - Device-defined programs
    - Programs provided at time of manufacture e.g., compression, encryption
  - Downloaded programs
    - Programs that are loaded to a Computational Programs namespace by the host
- May be implemented in a number of ways. For example,
  - A program may be implemented in an ASIC or an FPGA
  - A program may be executed on a CPU core

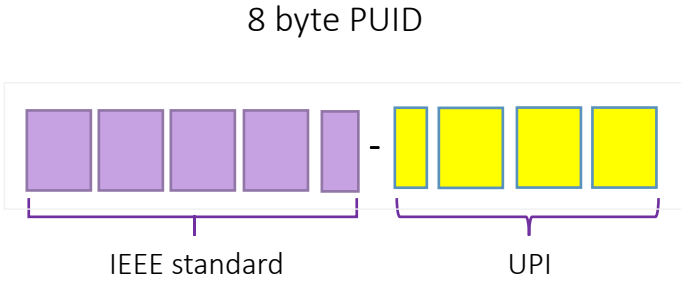


This presentation discusses NVMe® technology work in progress, which is subject to change without notice.

Program Unique Identifier (PUID)

A Program Unique ID:

- Is an identifier associated with specific program functionality
- Is composed of an IEEE assigned value and a Unique Program Identifier (UPI)
  - The UPI is unique to and assigned by the organization identified by the IEEE standard value
- For programs that are NVMe® technology defined
  - The IEEE standard value is set to DA-E6-D7-00-0, which contains the CID assigned to NVMe by IEEE
  - The association of a PUID to functionality will be defined in an NVM Express® maintained registry
- For programs that are vendor defined
  - The IEEE standard value is set using the vendor's IEEE assigned value (OUI-24, CID-24, or OUI-36)



The IEEE standard value is created using one of the following:

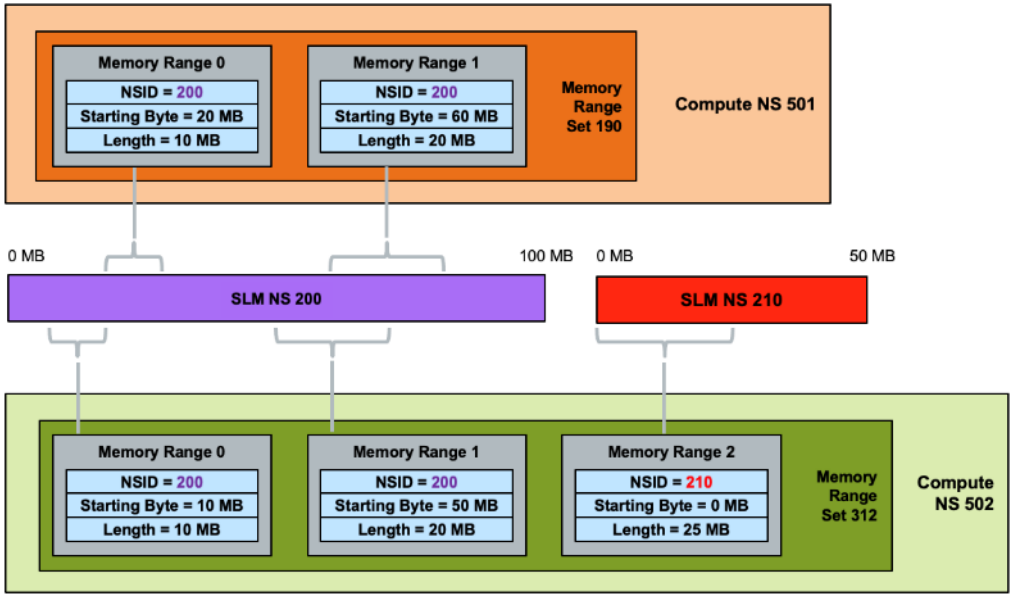
OUI – 24 bit IEEE assigned Organizationally Unique Identifier
CID – 24 bit IEEE assigned Company ID
OUI-36 – 36 bit IEEE assigned Organizationally Unique Identifier

This presentation discusses NVMe technology work in progress, which is subject to change without notice.

Memory Range Set (MRS)

A Memory Range Set:

- Is used for the purpose of limiting program access to a specific subset of SLM.
  - Each execution of a program is restricted from accessing any SLM other than what is specified by the Memory Range Set in the Execute Program command
- Describes one or more ranges of Subsystem Local Memory (SLM).
  - Each range is specified by a memory Namespace ID, an offset, and a length.
  - A range must only specify a memory namespace that is reachable by the compute namespace
- Is created and exists in a specific compute namespace.



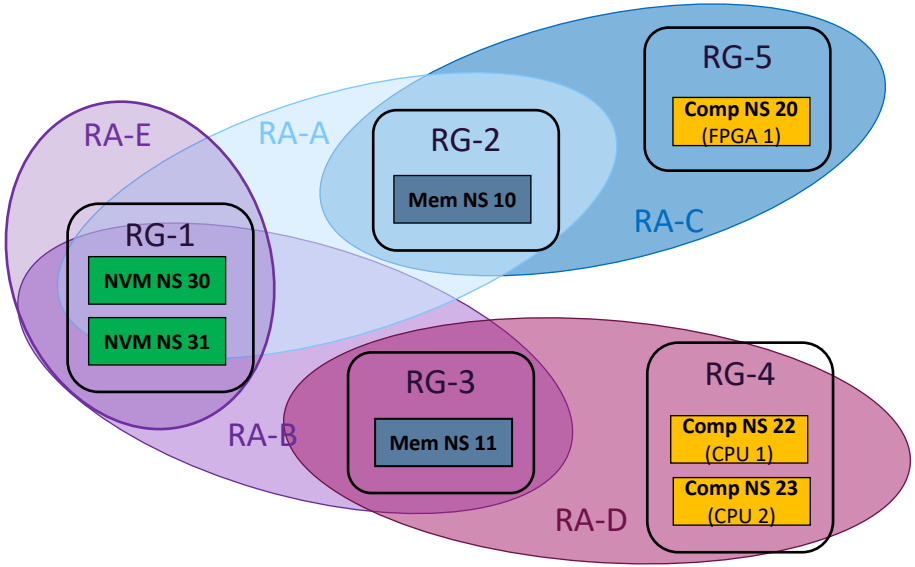
This presentation discusses NVMe® technology work in progress, which is subject to change without notice.

Reachability

Reachability is a new feature in the NVM Express® Base Specification

TP4156: Reachability

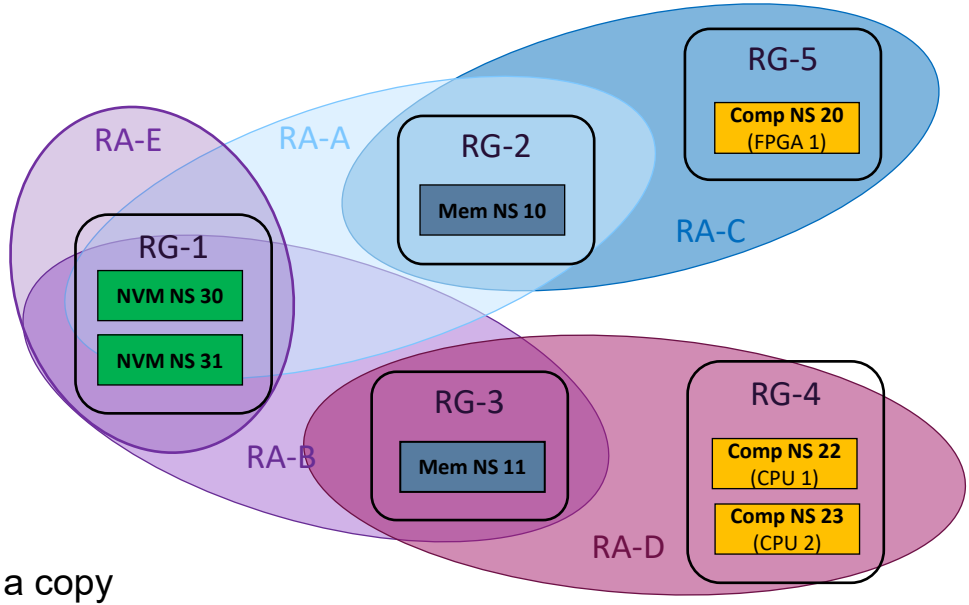
- Defines descriptors that specify what namespaces are able to be used together in a command. For example:
  - Which namespaces may be specified in a copy command
  - Which memory namespace may be specified in a command to a Compute Namespace
- Mechanism
  - Defines a Reachability Groups (RG) log page
  - Defines a Reachability Association (RA) log page
    - This log page defines characteristics, if any, of each Reachability Association



This presentation discusses NVMe® technology work in progress, which is subject to change without notice.

Reachability Associations (RA) Example

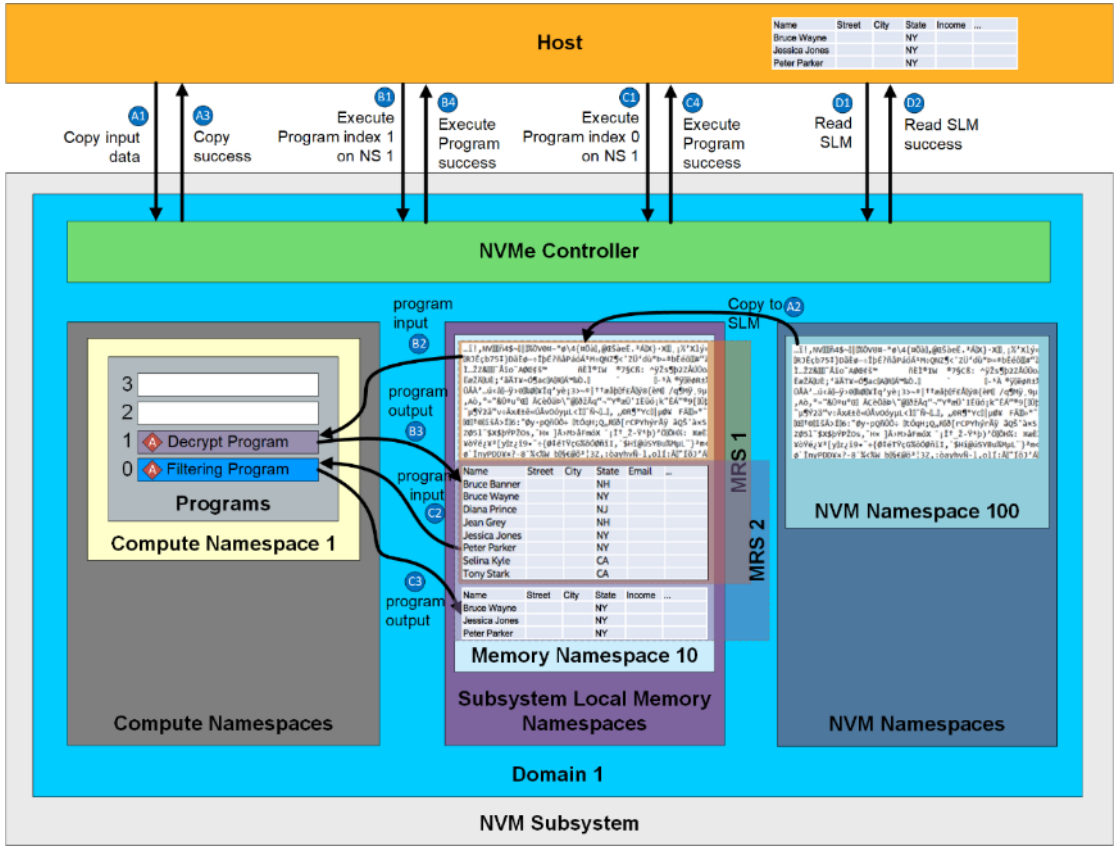
RA	Meaning
A	Copy is possible between NS 30 or NS 31 and NS 10
B	Copy is possible between NS 30 or NS 31 and NS 11
C	Compute NS 20 may use memory in NS 10
D	Compute NS 22 and NS 23 may use memory in NS 11
E	Copy is possible between NS 30 and NS 31



- Memory NS 10 and NS 11 CANNOT be used in a copy command to each other
- Compute NS 22 and NS 23 CANNOT communicate with each other

This presentation discusses NVMe® technology work in progress, which is subject to change without notice.

Flow: Execute Program – Filter Encrypted Data



Precondition:

- Memory Range Sets MRS1 and MRS2 have been created

Flow steps

A. Copy encrypted data into SLM

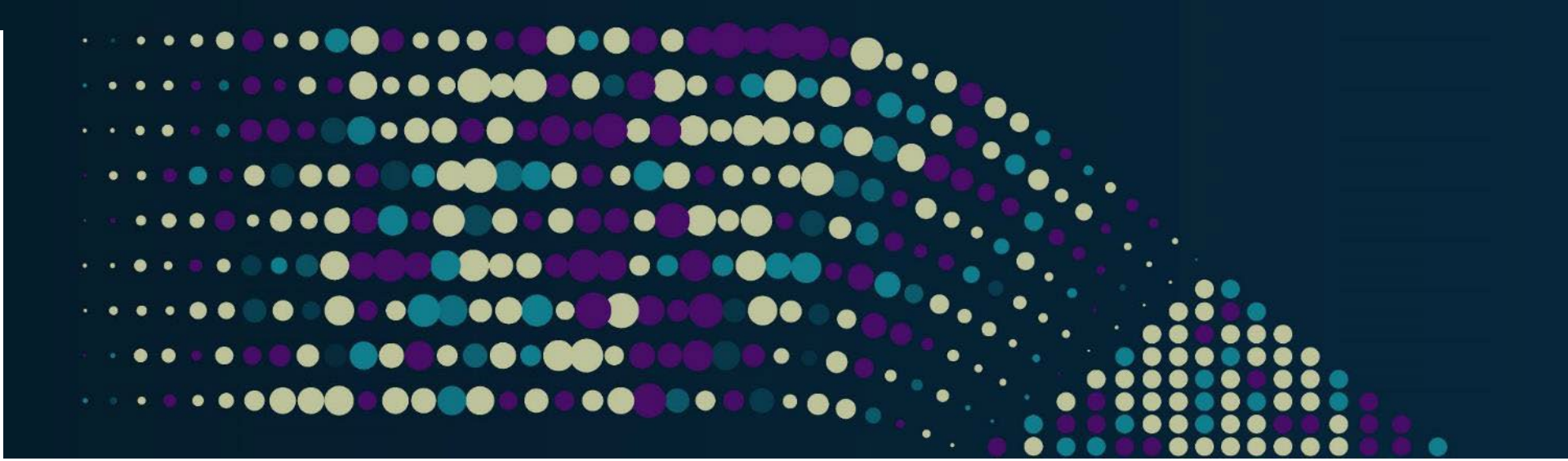
B. Execute Program 1 on compute NS 1 using MRS1

C. Execute Program 0 on compute NS 1 using MRS2

D. Read filtered data from SLM to host

This presentation discusses NVMe® technology work in progress, which is subject to change without notice.

Questions?



Please take a moment to rate this session.

Your feedback is important to us.