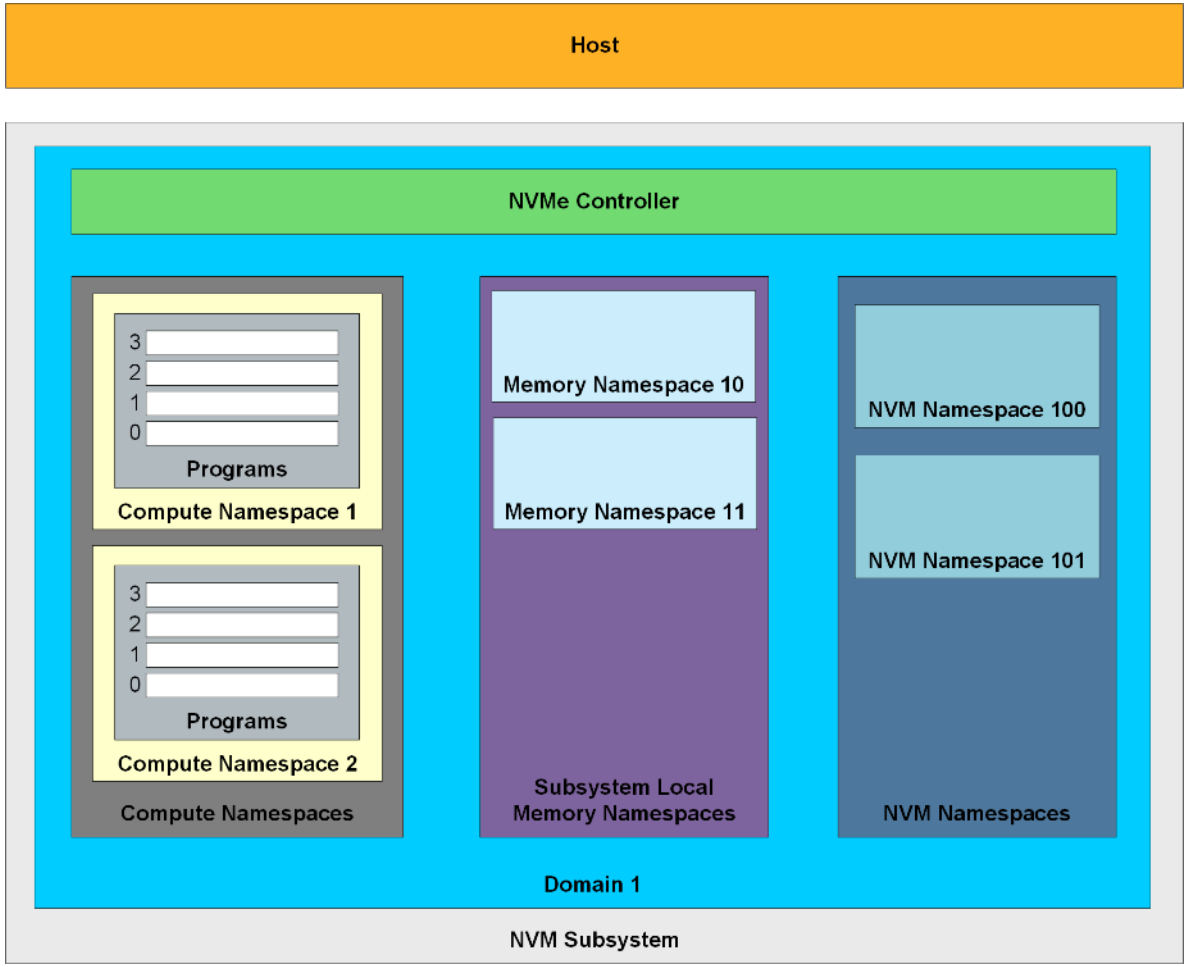
SNIA COMPUTE + MEMORY + STORAGE SUMMIT

Architectures, Solutions, and Community
VIRTUAL EVENT, APRIL 11-12, 2023

NVMe® Computational Storage

Presented by
Kim Malone, Intel
Bill Martin, Samsung

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.

Compute Namespaces

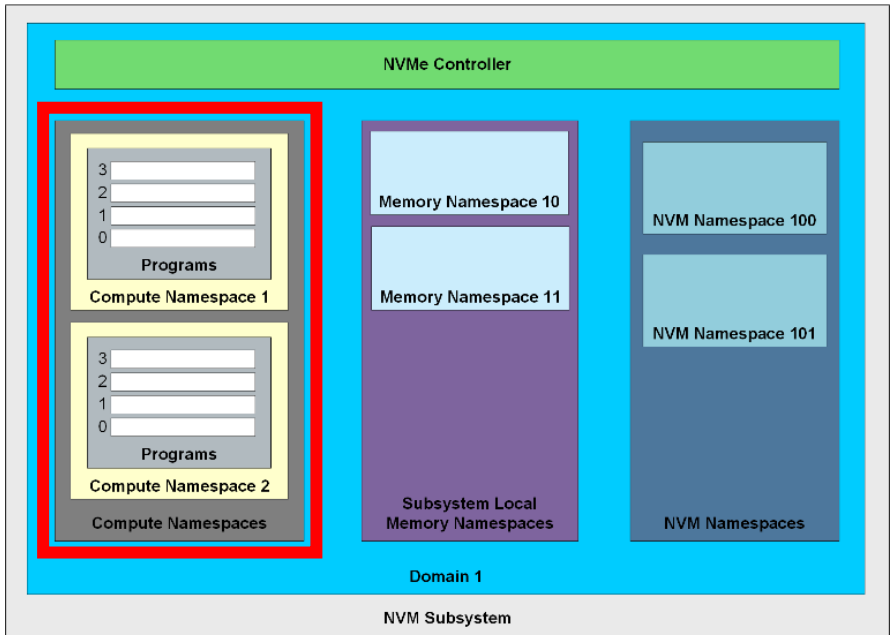
A compute namespace:

- Is a namespace in an NVMe technology subsystem that is able to execute one or more programs
- Is a namespace that is associated with the Computational Programs I/O command set
- Contains compute resources

TP4091: Computational Programs

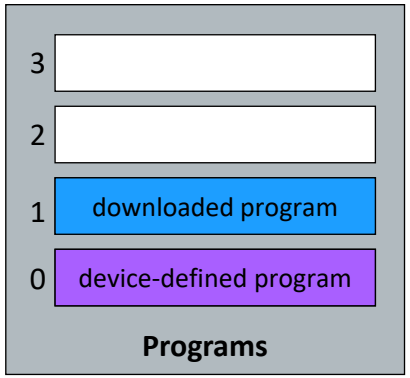
New Computational Programs I/O command set for compute namespaces

- New commands include:
  - Execute program
  - Load program
  - Activate program
  - Create/Delete Memory Range Set
- 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
- May be device-defined or downloadable
  - Device-defined programs
    - Programs provided at time of manufacture e.g., compression, encryption
  - Downloadable programs
    - Programs that are loaded to a Computational Programs namespace by the host
- A program may only be able to execute on a subset of the compute resources in an NVM subsystem
  - A program may be implemented in an ASIC
  - A program may be executed on a CPU core



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

Memory Namespaces

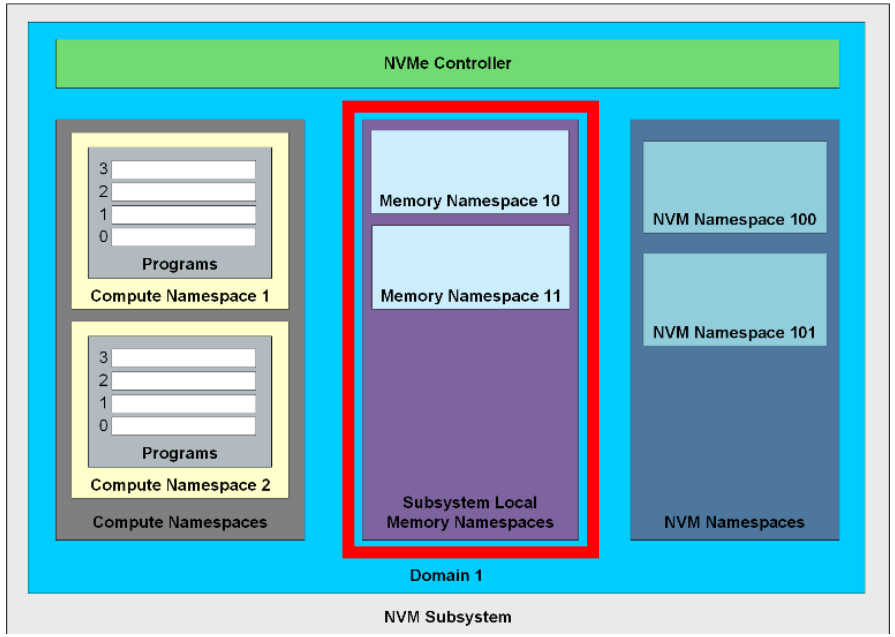
A memory namespace:

- Is a namespace in an NVMe technology subsystem that provides host command access to memory in the NVMe technology subsystem
- Is a namespace that 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 between NVM and memory namespaces



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

Reachability

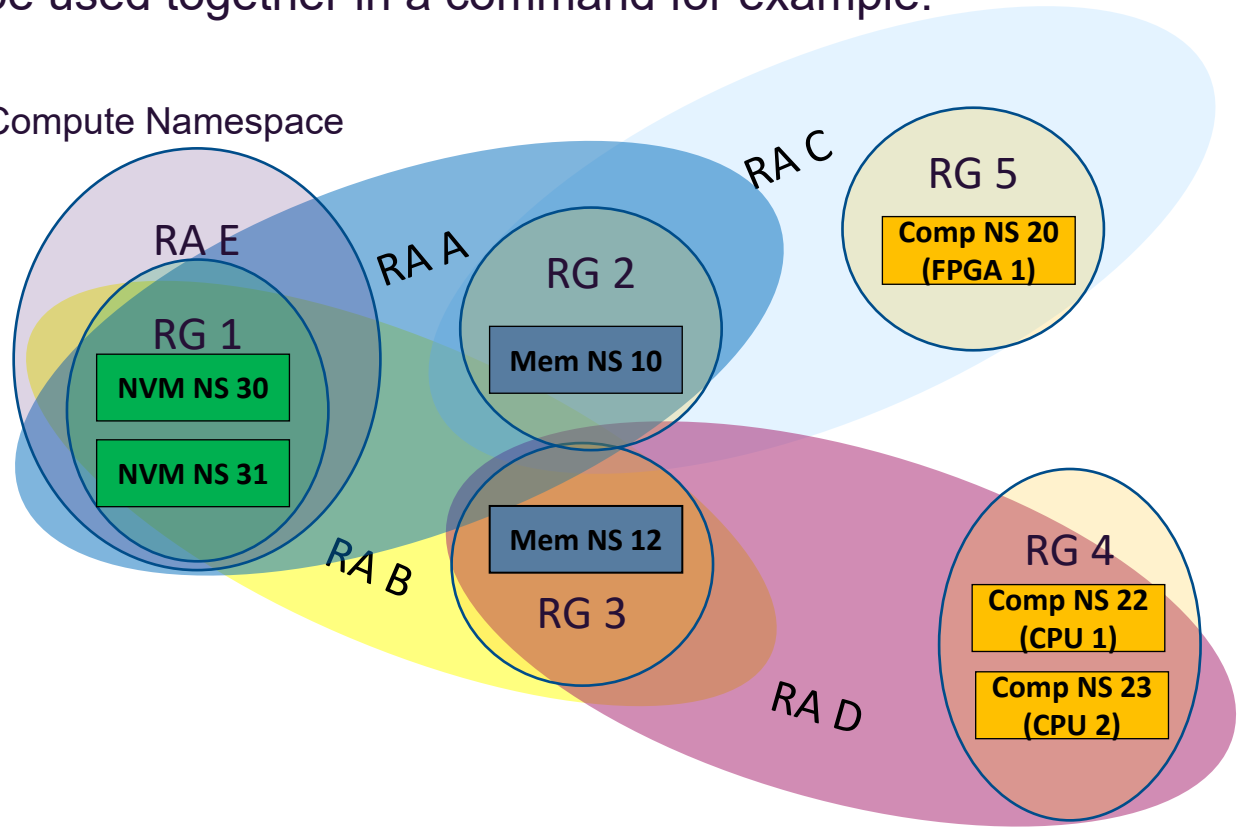
Reachability:

- A new feature in the NVM Express Base Specification
- 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

TP4156: Reachability

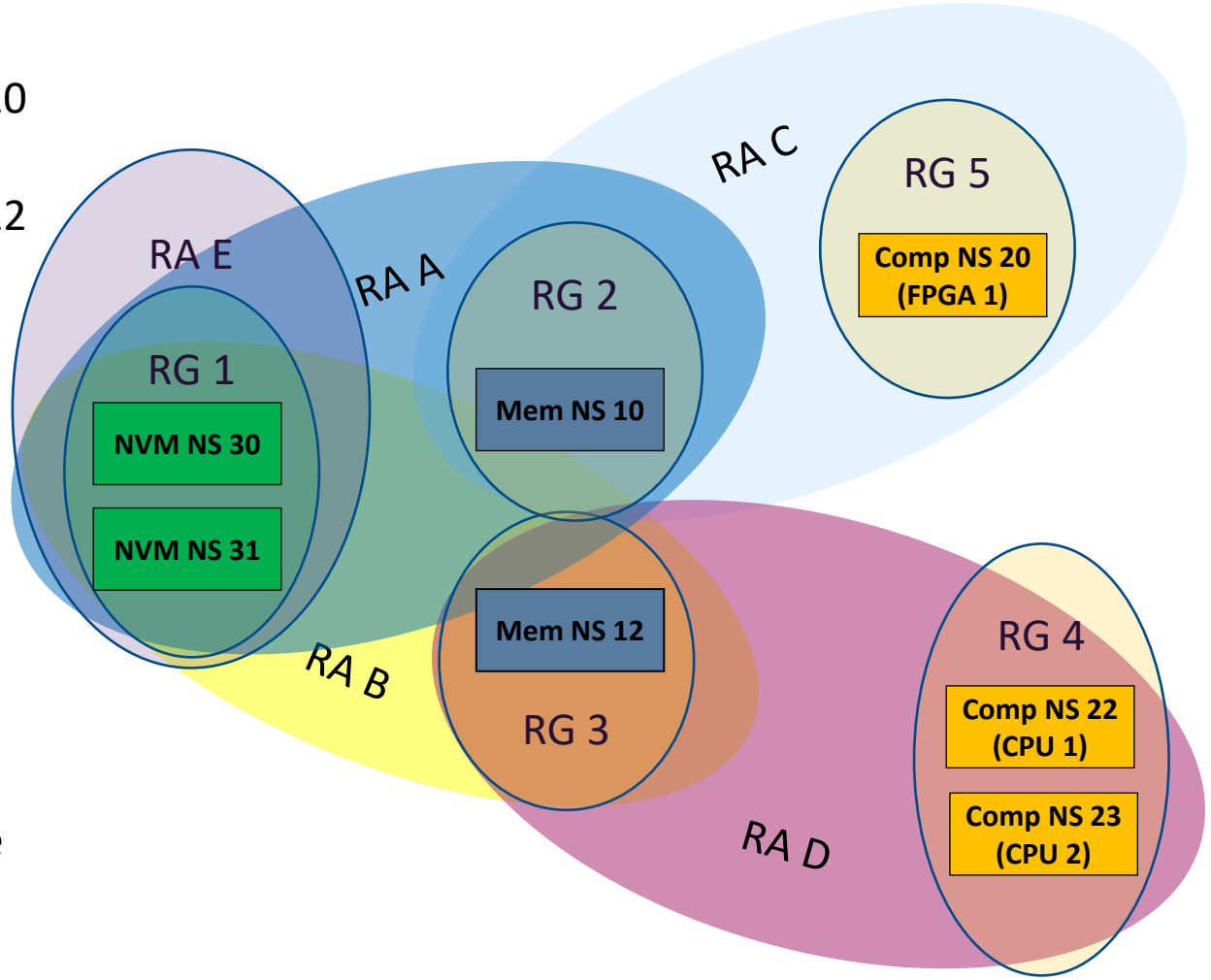
New feature in the NVM Express Base specification

- Defines a Reachability Groups log page (RG)
- Defines a Reachability Association log page (RA)
- The Reachability Association log page defines characteristics, if any, of each Reachability Association



Reachability Associations Example

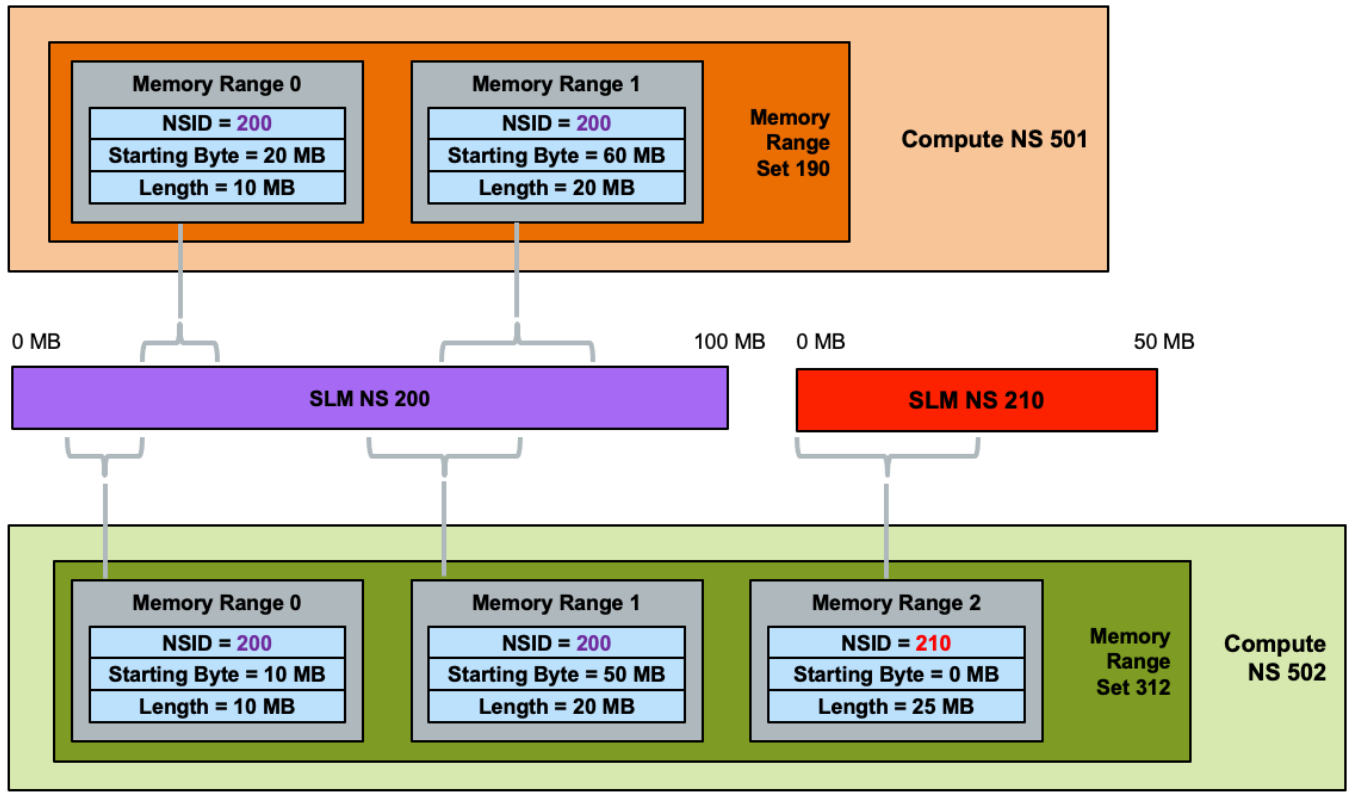
- RA A
  - Copy is possible between NS 30 or NS 31 and NS 10
- RA B
  - Copy is possible between NS 30 or NS 31 and NS 12
- RA C
  - Compute NS 20 may use memory in NS 10
- RA D
  - Compute NS 22 and NS 23 may use memory in NS 12
- RA E
  - Copy is possible between NS 30 and NS 31
- Memory NS 10 and NS 11 CANNOT be used in a copy command to each other
- Compute NS 22 and NS 23 CANNOT communicate with each other



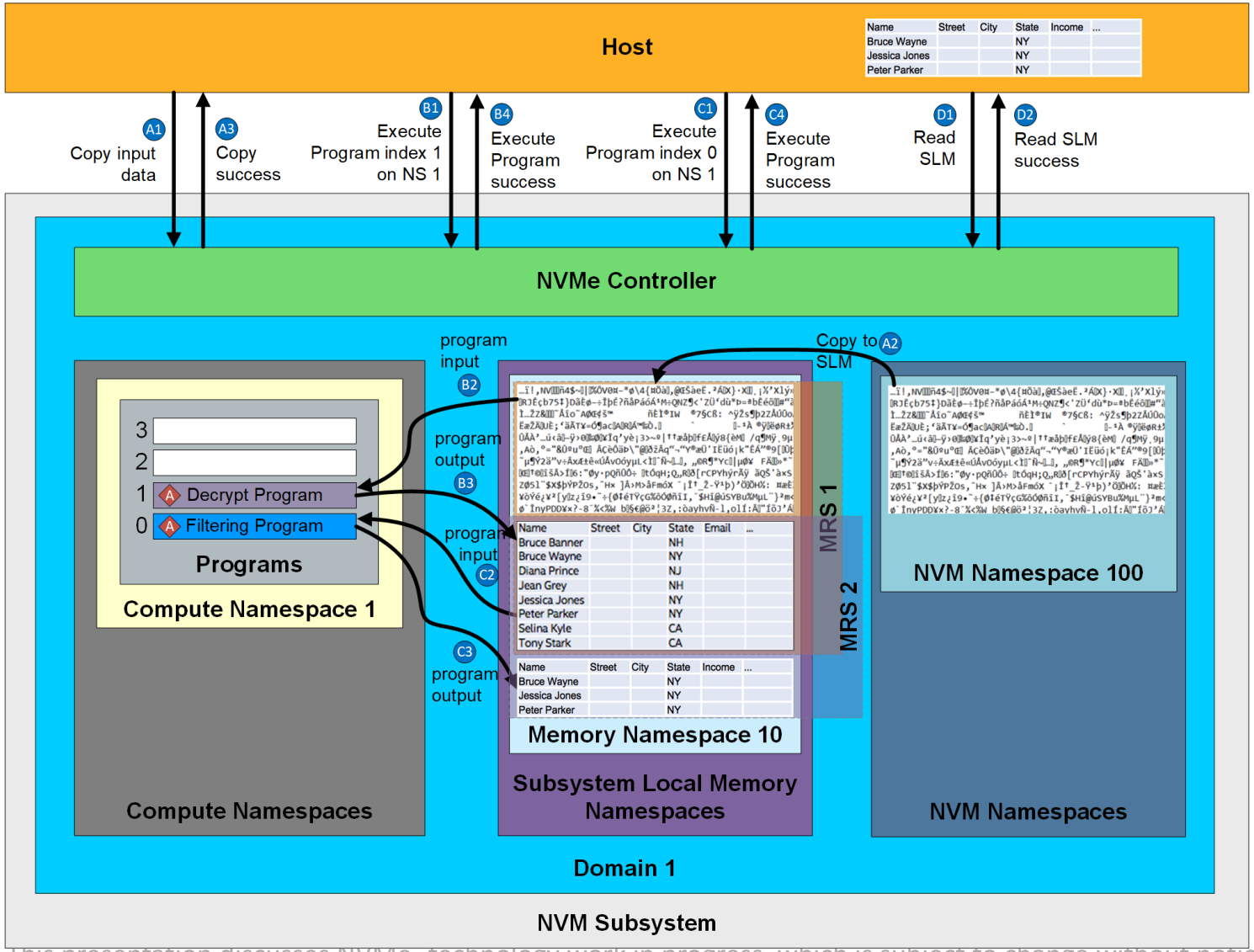
Memory Range Set

- A Memory Range Set describes one or more ranges of Subsystem Local Memory.
- Memory Range Sets are used for the purpose of limiting program access to a specific subset of Subsystem Local Memory.

- A Memory Range Set is created in a specific compute namespace.
- Each execution of a program is restricted from accessing any Subsystem Local Memory other than what is specified by the Memory Range Set described in the Execute Program command
- Each range is specified by a memory NamespaceID, an offset, and a length.
- Each range must only specify a namespace that is reachable by the compute namespace



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.

NVM Express Computational Storage Task Group

- Task Group co-chairs
  - Kim Malone (Intel)
  - Bill Martin (Samsung)
- Join the task group
  - Go to the NVMe workgroup portal
  - Select the CS Task Group
  - Click on the “Join Group” link
- Task group meetings
  - Thursdays 9 – 10 am Pacific time

JOIN US!

SNIA COMPUTE + MEMORY + STORAGE SUMMIT

Architectures, Solutions, and Community
VIRTUAL EVENT, APRIL 11-12, 2023

Please take a moment to rate this session.

Your feedback is important to us.

Post-Summit, visit www.snia.org/cms-summit for additional content.

