STORAGE DEVELOPER CONFERENCE

SDC22

Fremont, CA
September 12-15, 2022

BY Developers FOR Developers

A SNIA. Event

NVMe® Computational Storage

Standardizing offload of computation

Presented by Kim Malone, Intel

nvm EXPRESS®

Programs as Computational Storage Offloads

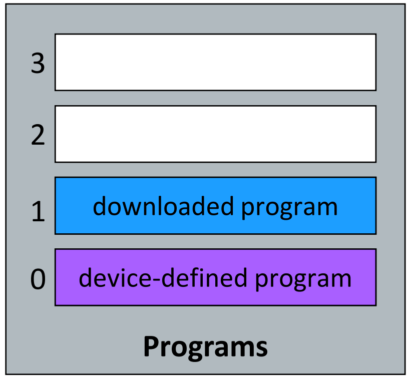
Programs:

- Invoked and used in a standard way
  - Conceptually similar to software functions
  - Called with parameters and run to completion
- Operate only on data in Subsystem Local Memory
- Run on compute resources
- May be in hardware or software
  - Device may offer fixed function programs
  - Device may offer downloadable programs
- A program may only be able to execute on a subset of the compute resources in an NVM subsystem

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

Downloadable and device-defined programs

- Support for both device-defined and downloadable programs
- Device-defined programs
  - “Fixed” programs provided by the manufacturer
  - Functionality implemented by the device that are callable as programs
  - e.g. compression, decryption
- Downloadable programs
  - Programs that are loaded to a Computational Programs namespace by the host

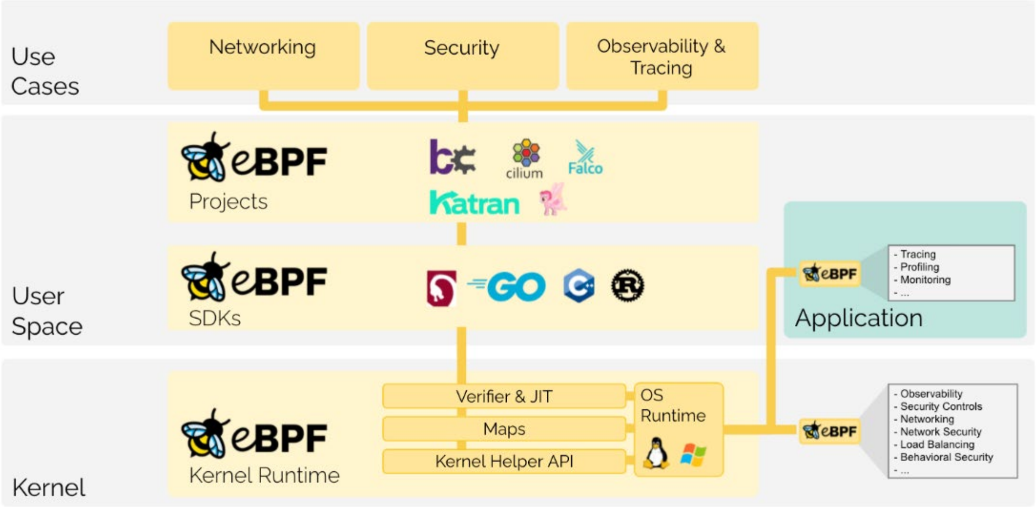


This presentation discusses NVMe® technology work in progress, which is subject to change without notice.

Downloadable Programs

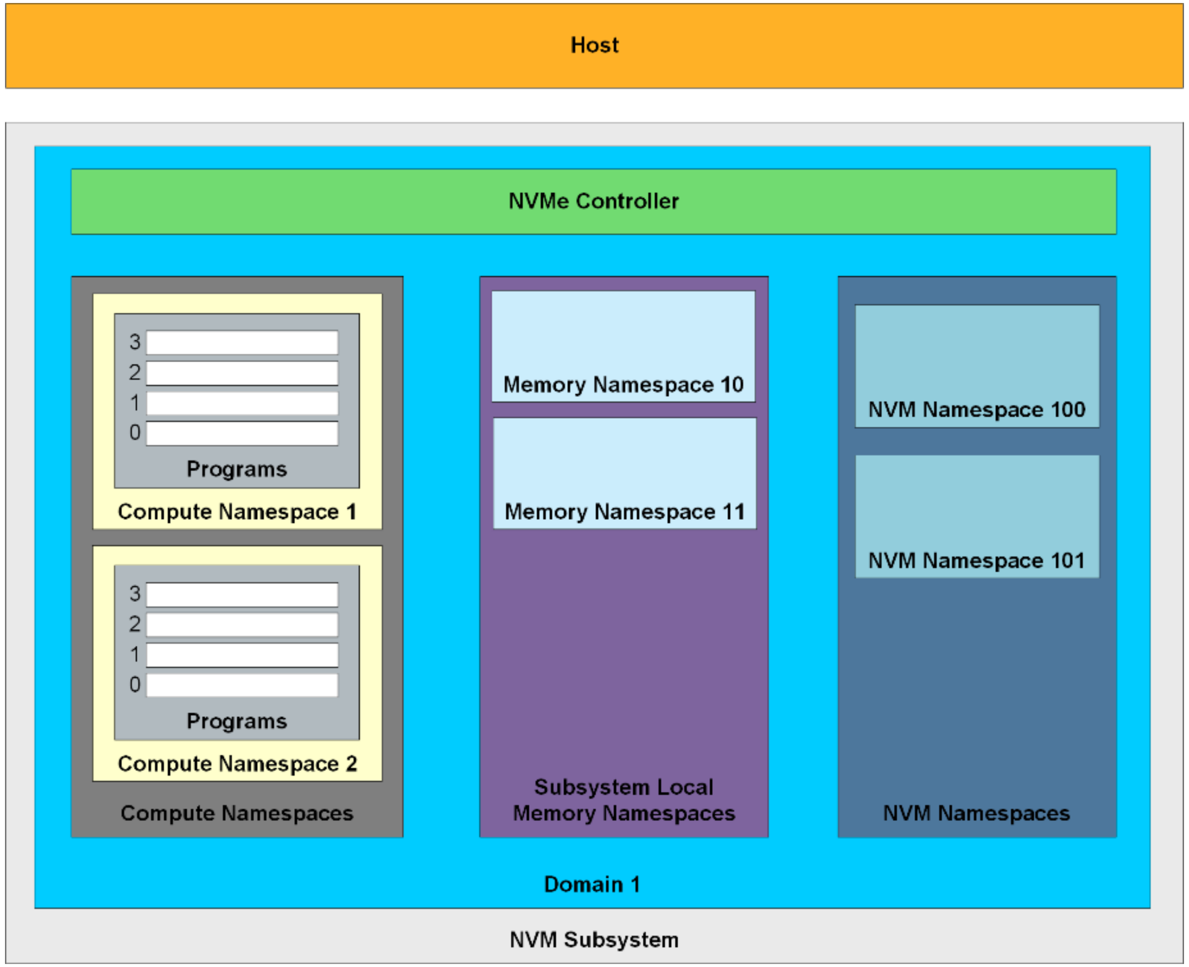
- Why downloadable programs?
  - Flexibility
  - Process complex formats
  - Emerging applications
  - Portability from existing applications
  - Vendor-specific formats or well-known formats (e.g. eBPF)

- Example format: eBPF
  - Vendor agnostic
  - Well understood
  - Existing ecosystems, toolchains
  - LLVM
  - Sits under Linux Foundation



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

Major Architectural Components



The 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

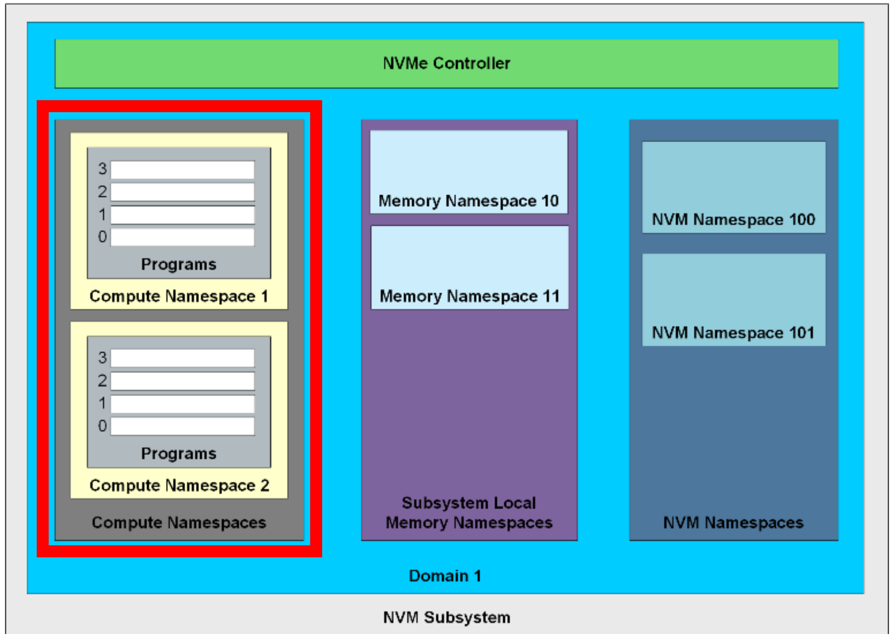
A compute namespace:

- Is a namespace in an NVMe® technology subsystem that is able to execute one or more programs
- May support a subset of all possible program types
- Is a namespace that is associated with the Computational Programs I/O command set
- Programs may access data in one or more memory namespaces

TP4091: Computational Programs

New Computational Programs I/O command set for compute namespaces

- New commands may include:
  - Execute program
  - Load program
  - Activate program
  - Create/Delete Memory Range Set
- Support for Identify Controller, Namespace



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

Memory Namespaces

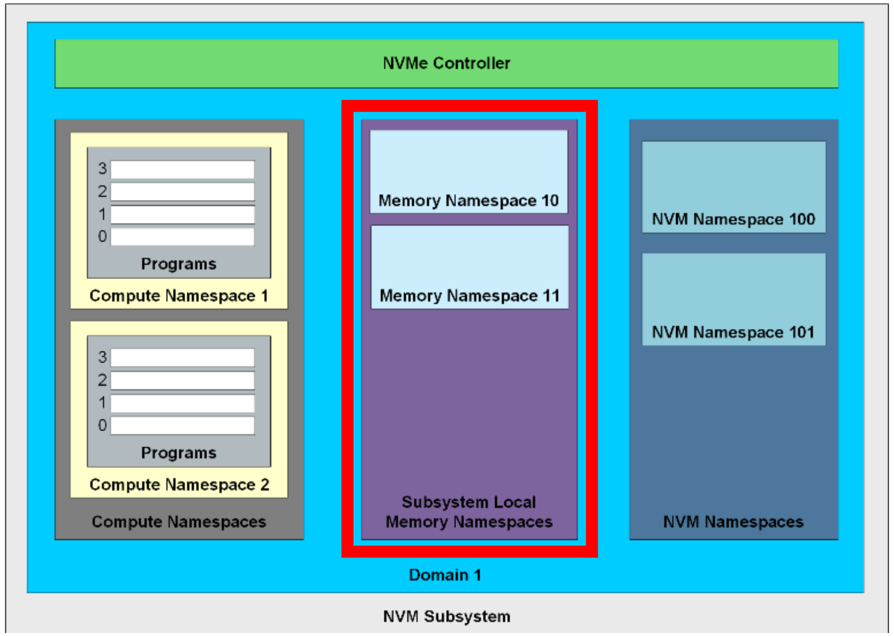
A memory namespace:

- 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)

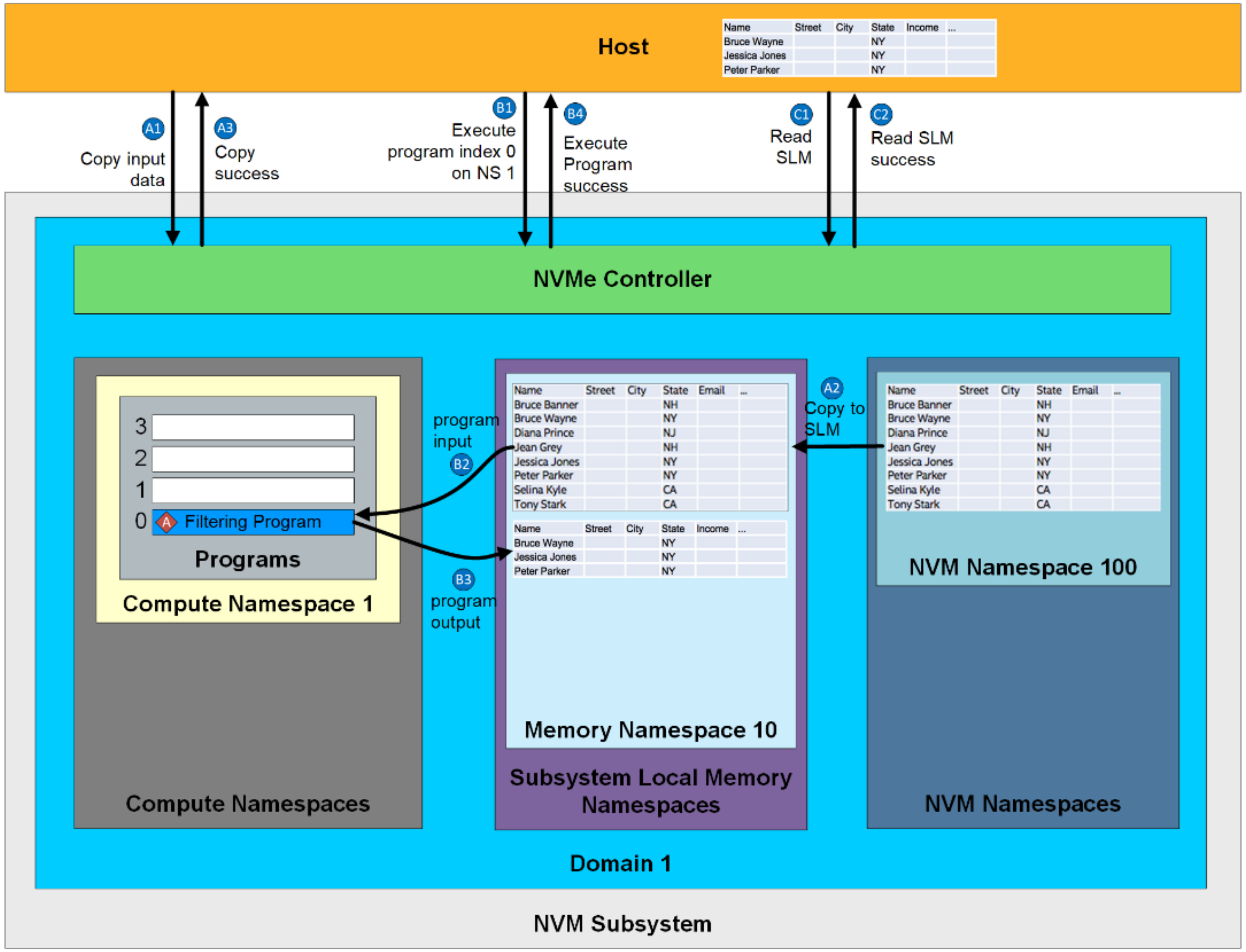
New Subsystem Local Memory I/O command set for memory namespaces

- New commands may include:
  - Commands for reading from a memory namespace into host memory and writing from host memory to a memory namespace
  - Command to allow copying data between NVM and memory namespaces
- Support for Identify Controller, Namespace



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

Flow: Execute Program – Simple Data Filter



Host:

Name	Street	City	State	Income	...
Bruce Wayne			NY		
Jessica Jones			NY		
Peter Parker			NY		

Memory Namespace 10:

Name	Street	City	State	Email	...
Bruce Banner			NH		
Bruce Wayne			NY		
Diana Prince			NJ		
Jean Grey			NH		
Jessica Jones			NY		
Peter Parker			NY		
Selina Kyle			CA		
Tony Stark			CA		

Name	Street	City	State	Income	...
Bruce Wayne			NY		
Jessica Jones			NY		
Peter Parker			NY		

NVM Namespace 100:

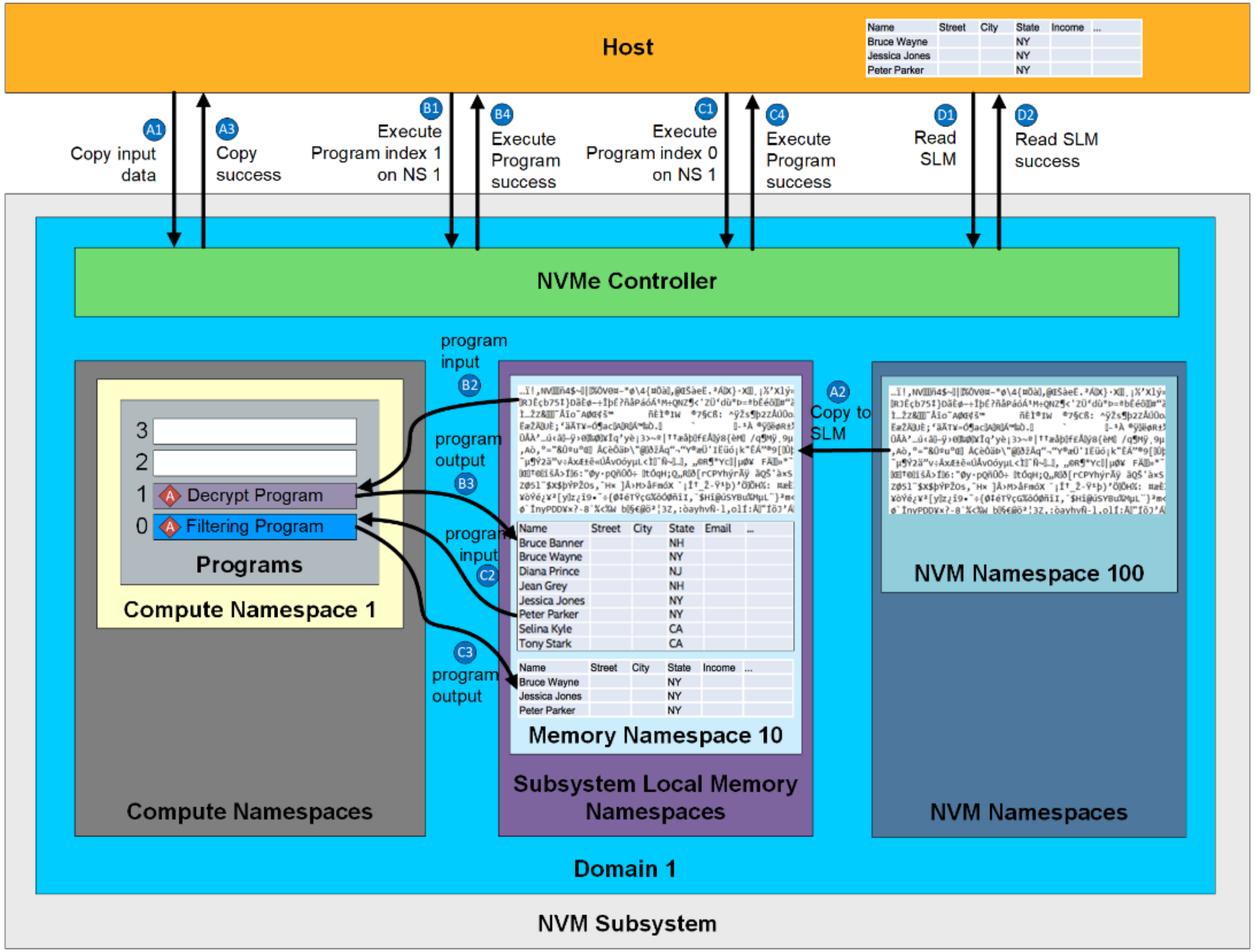
Name	Street	City	State	Email	...
Bruce Banner			NH		
Bruce Wayne			NY		
Diana Prince			NJ		
Jean Grey			NH		
Jessica Jones			NY		
Peter Parker			NY		
Selina Kyle			CA		
Tony Stark			CA		

Flow steps

- A Copy stored data into subsystem memory
- B Execute Program with index 0 on NS 1
- C Read filtered data from subsystem memory to host

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

Flow: Execute Program – Filter Encrypted Data



Name	Street	City	State	Income	...
Bruce Wayne			NY		
Jessica Jones			NY		
Peter Parker			NY		

Name	Street	City	State	Email	...
Bruce Banner			NH		
Bruce Wayne			NY		
Diana Prince			NJ		
Jean Grey			NH		
Jessica Jones			NY		
Peter Parker			NY		
Selina Kyle			CA		
Tony Stark			CA		

Name	Street	City	State	Income	...
Bruce Wayne			NY		
Jessica Jones			NY		
Peter Parker			NY		

Flow steps

- A Copy encrypted data into subsystem memory
- B Execute Program 1 on NS 1
- C Execute Program 0 on NS 1
- D Read filtered data from subsystem memory 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)
  - Stephen Bates (Eideticom)
  - Bill Martin (Samsung)
- Task Group Goals
  - Define the architecture of TP4091
  - Take TP4091 through to ratification
  - Other CS Technical Proposals

JOIN US!

- Membership
  - 228 members from 49 companies
- 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

QUESTIONS?

Please take a moment to rate this session.

Your feedback is important to us.

nvm
EXPRESS®
Architected for Performance