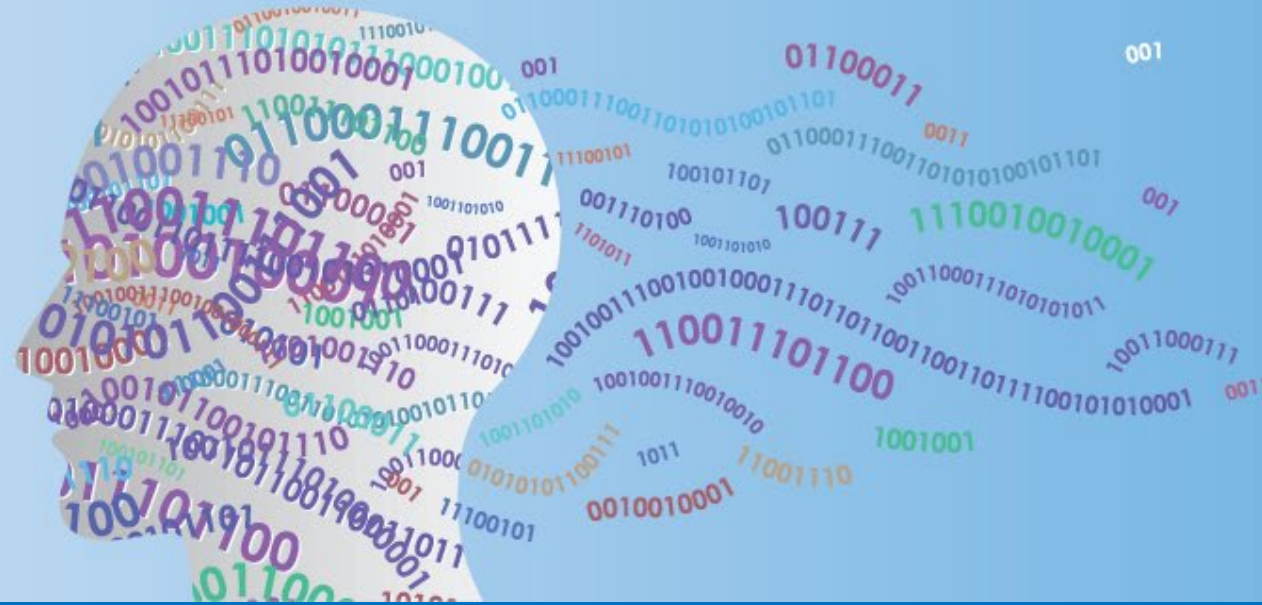




SNIA

# PERSISTENT MEMORY + SUMMIT 2021 COMPUTATIONAL STORAGE

FROM DATACENTER TO EDGE : VIRTUAL EVENT  
APRIL 21-22, 2021



## A New Path to Better Data Movement within System Memory, Computational Memory with SDXI

Shyam Iyer, *Chair, SNIA SDXI Technical Work Group;*  
*Distinguished Engineer, Dell*

(With inputs from SDXI TWG Members)

# SDXI: Smart Data Accelerator Interface

SDXI(Smart Data Accelerator Interface) is a proposed SNIA standard for a memory to memory Data Mover interface.



# Agenda

The problem and the need for a solution

Introducing SDXI(Smart Data Acceleration Interface)



# System Design Challenges

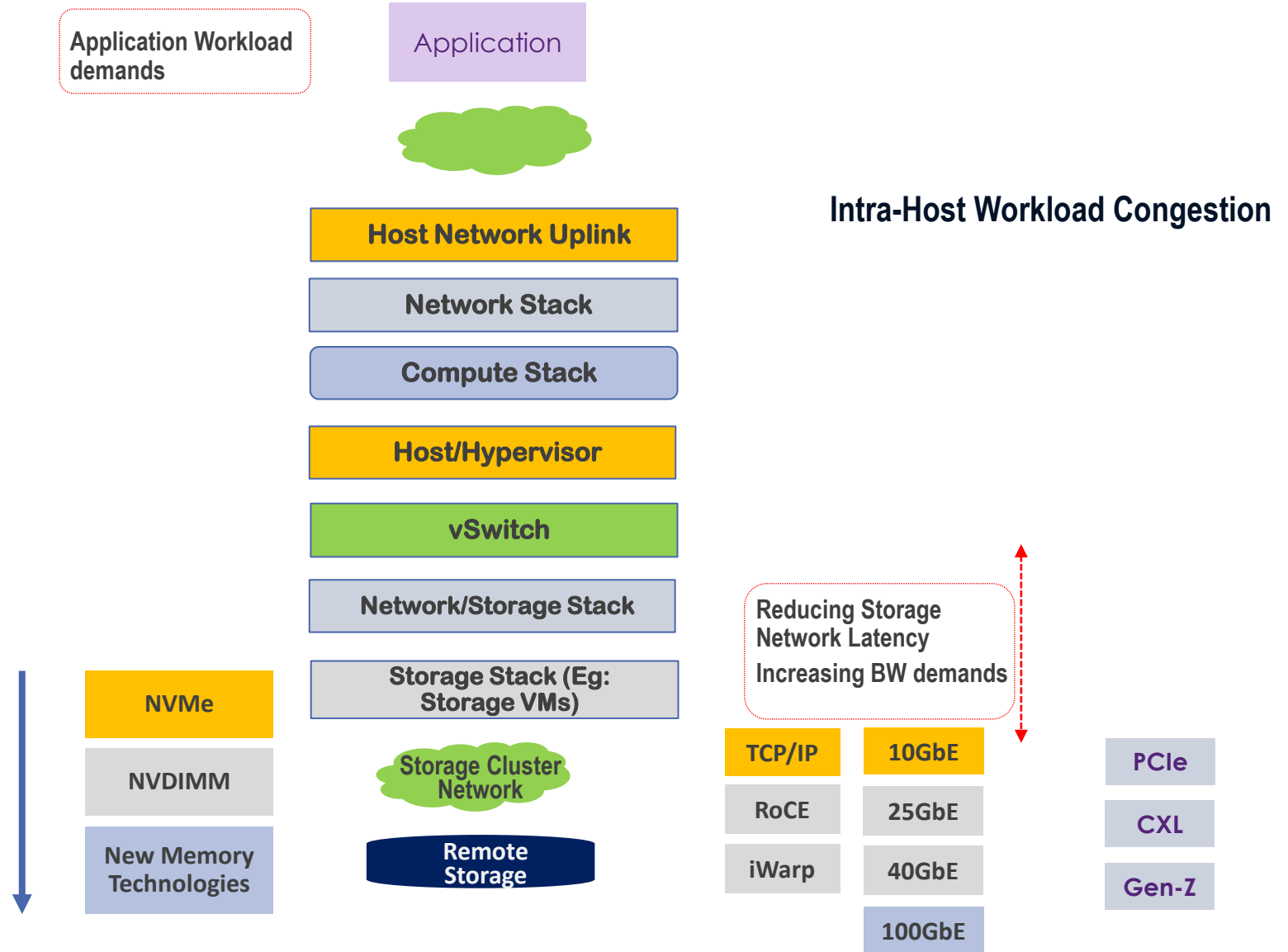
- Core counts increasing to enable Compute scaling
- Compute density is on the rise
- Converged and Hyperconverged Storage appliances are enabling new workloads on server class systems
- **Data locality is important**
- Single threaded performance is under pressure.
- I/O intensive workloads can take away compute CPU cycles available.
- Network and Storage workloads can take compute cycles
- Data Movement, Encryption, Decryption, Compression

Unlocking trapped resources

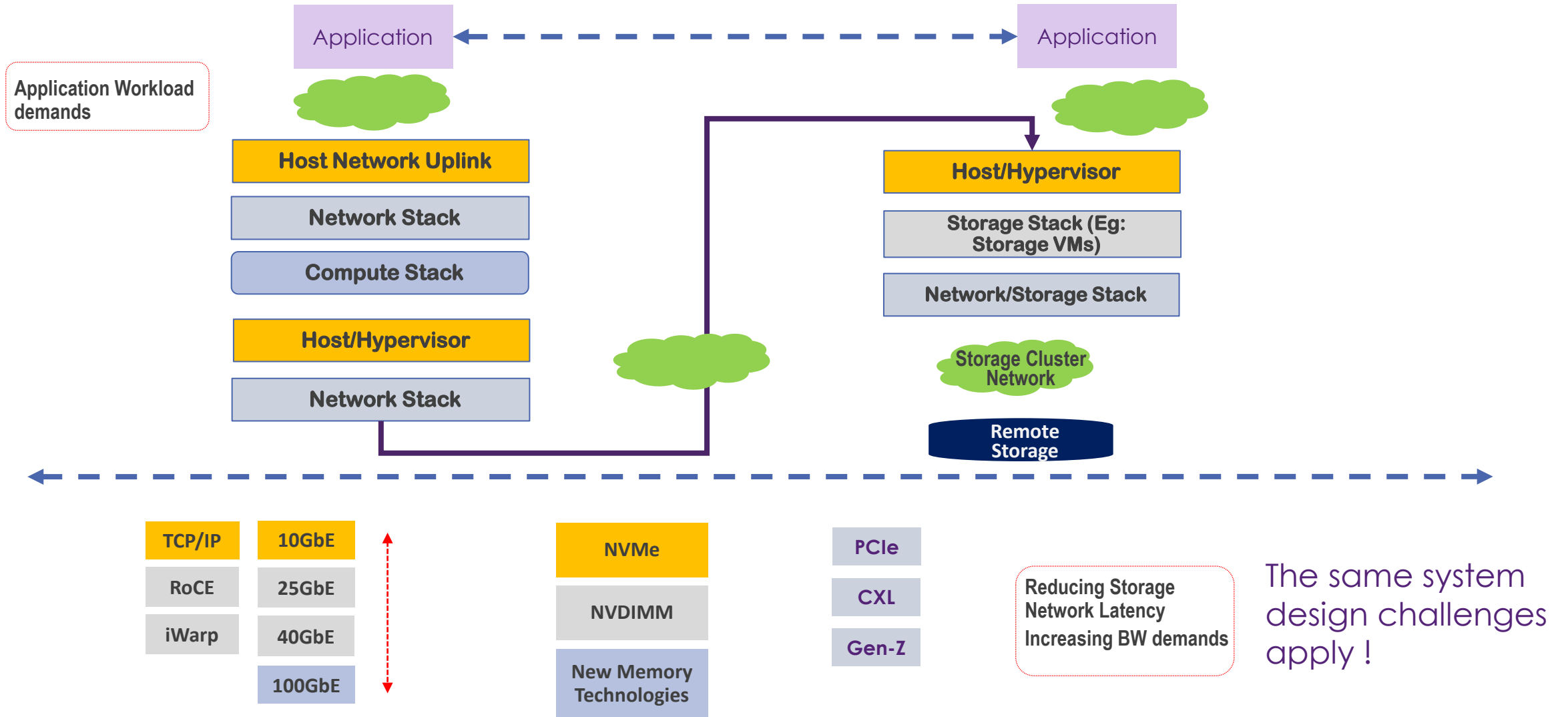


# Host Stack

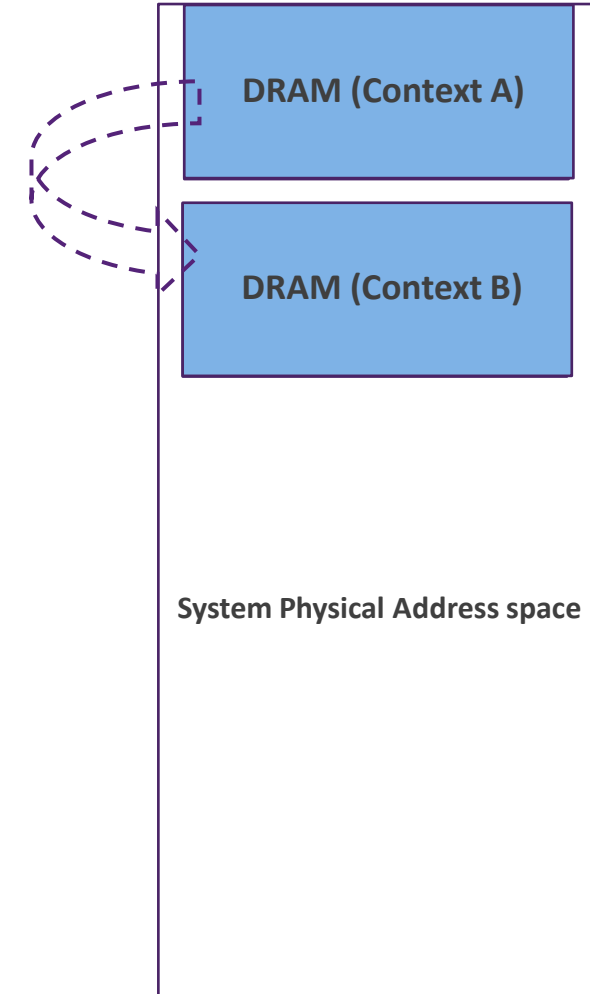
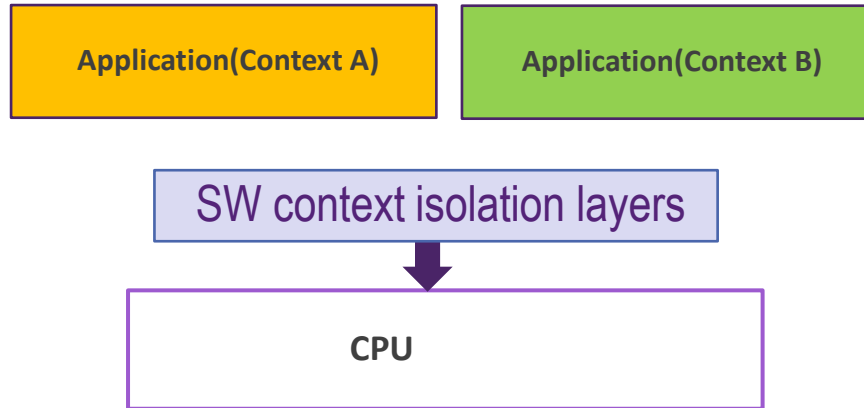
- Each intra-host exchange can comprise multiple memory buffer copies (or transformations)
  - Generally implemented with layers of software stacks:
  - Kernel-to-I/O can leverage I/O-specific hardware memory copy
  - But SW-to-SW usually relies on per-core synchronous software (CPU-only) memory copies



# Disaggregated Host Stacks



# Current memory to memory data movement standard:



- Stable CPU ISA for SW based memory copies
  - Takes away from application performance
  - Software overhead to provide context isolation
  - Synchronous SW copies stall applications
  - Less portable to different ISAs(Instruction Set Architectures)
  - Finely tuned CPU data movement algorithms can break with new microarchitectures

# Offload DMA engines: A new concept ?

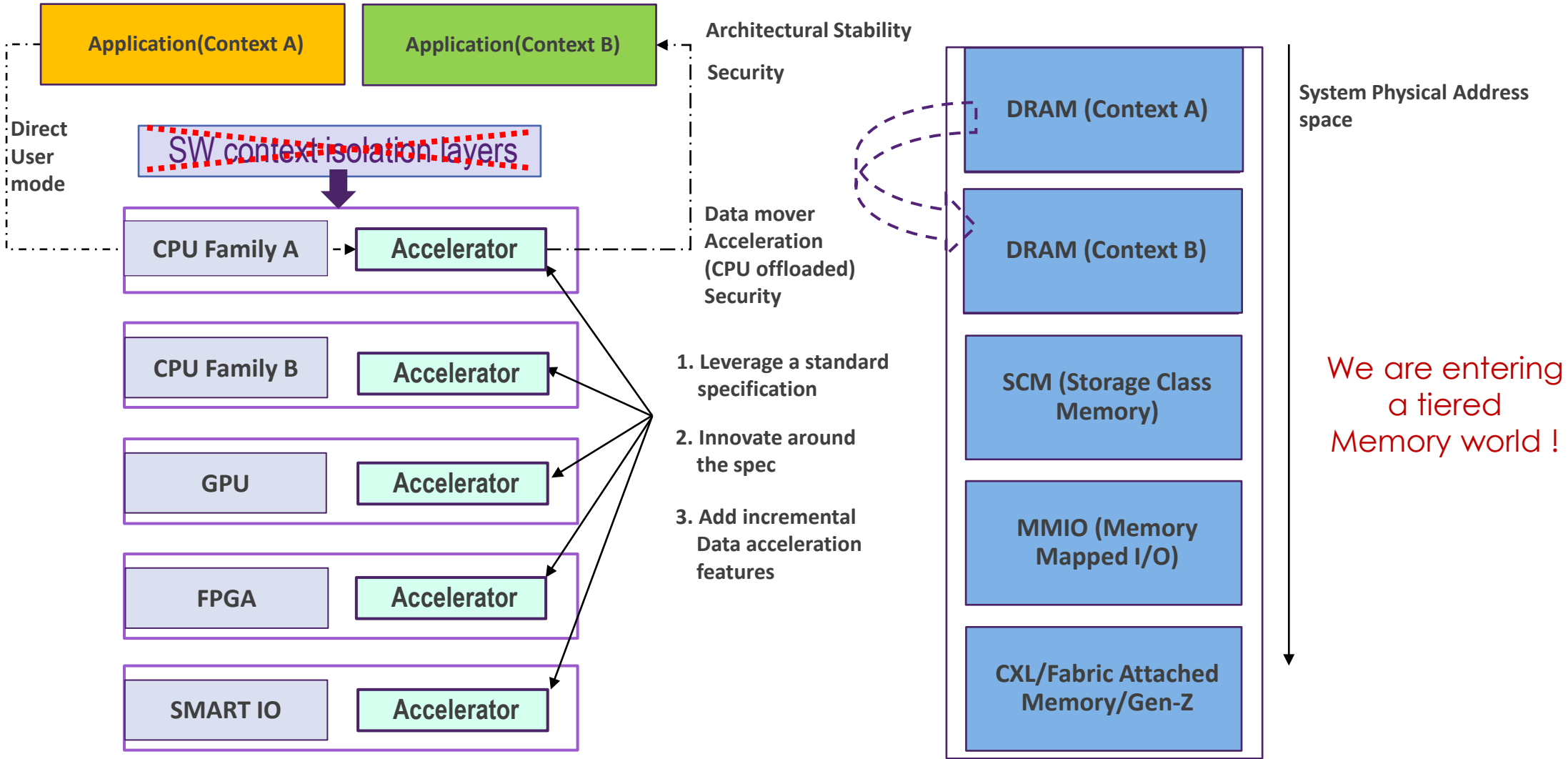
- Fast DMA offload engines are -
  - Vendor-specific HW
  - Vendor specific drivers, APIs
  - Vendor specific work submission/completion models
  - Direct access by user level software is difficult
  - Limited Usage Models
  - Vendor specific DMA states – Makes it harder to abstract/virtualize and migrate the work to other hosts



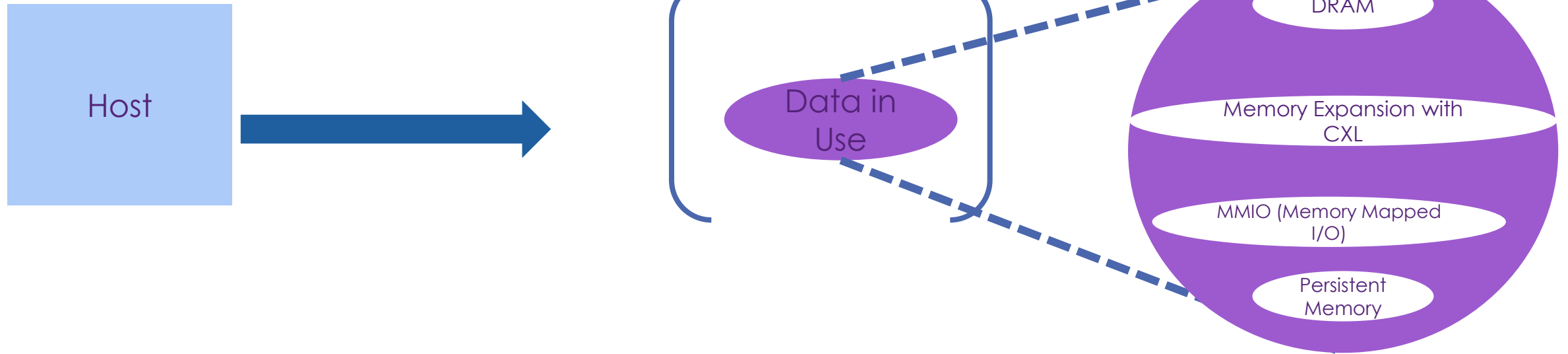
# Solution Requirements

1. Need to offload I/O from Compute CPU cycles
2. Need Architectural Stability
3. Enable Application/VM acceleration but,
  - Help migration from existing SW Stacks
4. Create abstractions in Control Path for scale and management
5. Enable performance in data path with offloads

# The need for an industry standard



# Why focus on standardizing a memory to memory interface ?



- Heterogenous Compute resource usage on the rise
- Blurring boundaries between a Server host and Storage host
- Disaggregation unlocks trapped resources

## Memory to Memory Interface

Computational Memory Interface

Interface to accelerate Storage & Memory Convergence

Interface to accelerate resource disaggregation with memory links/fabrics

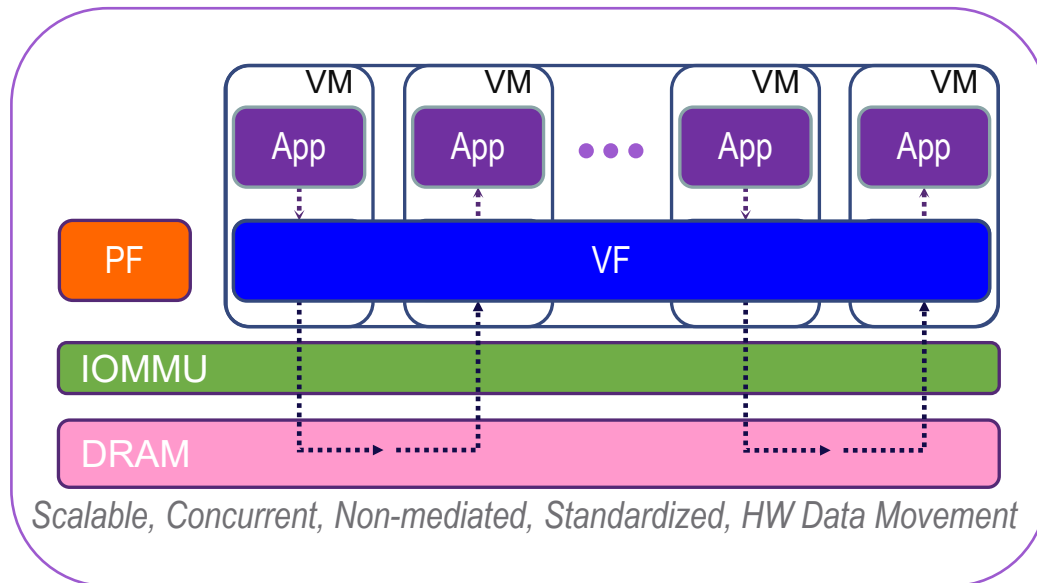
# Agenda

The problem and the need for a solution

Introducing SDXI(Smart Data Acceleration Interface)

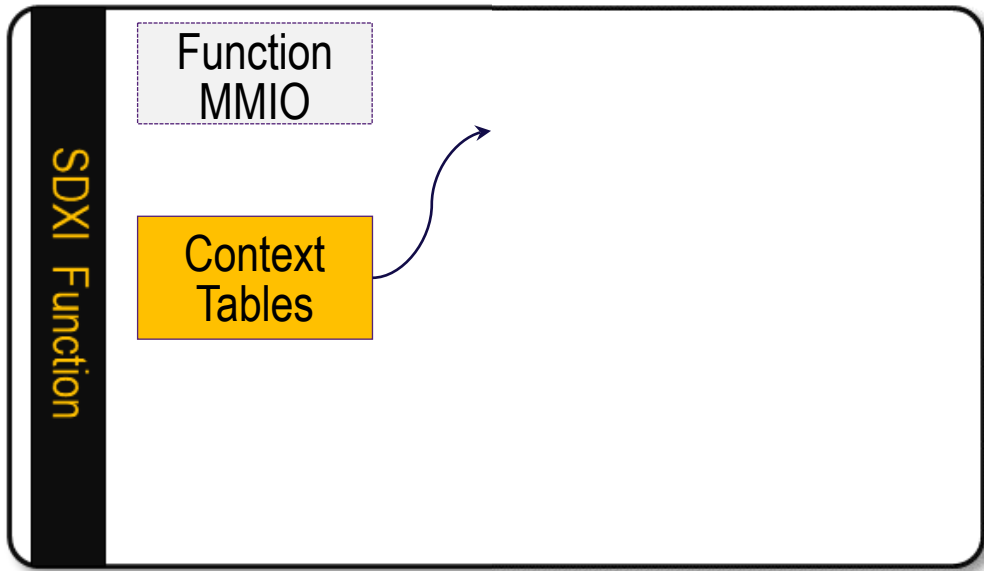


# Accelerated, Virtualized, Standardized HW Data Movement



1. A standardized data mover interface independent of actual implementations & underlying I/O.
2. Data movement between different address spaces both within and across VMs.
3. Data movement without mediation by privileged software(Hypervisor).
4. An interface and architecture that can be abstracted or virtualized by privileged software.
5. Concurrent DMA model.
6. Allow “live” workload or virtual machine migration between servers.
7. Forwards and backwards compatibility.
  - Allow Hardware, software Interoperability
8. Incorporate additional offloads in the future.

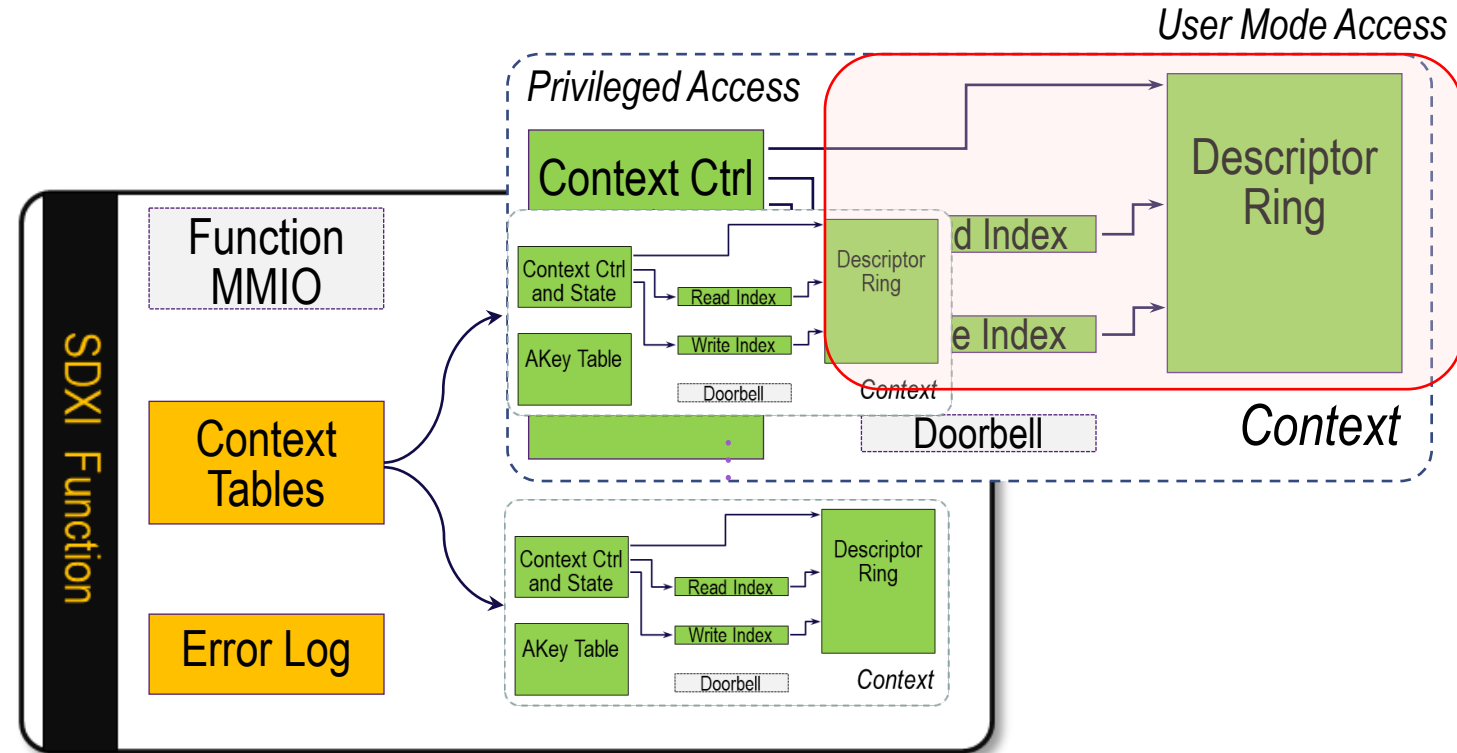
# SDXI Function Architecture



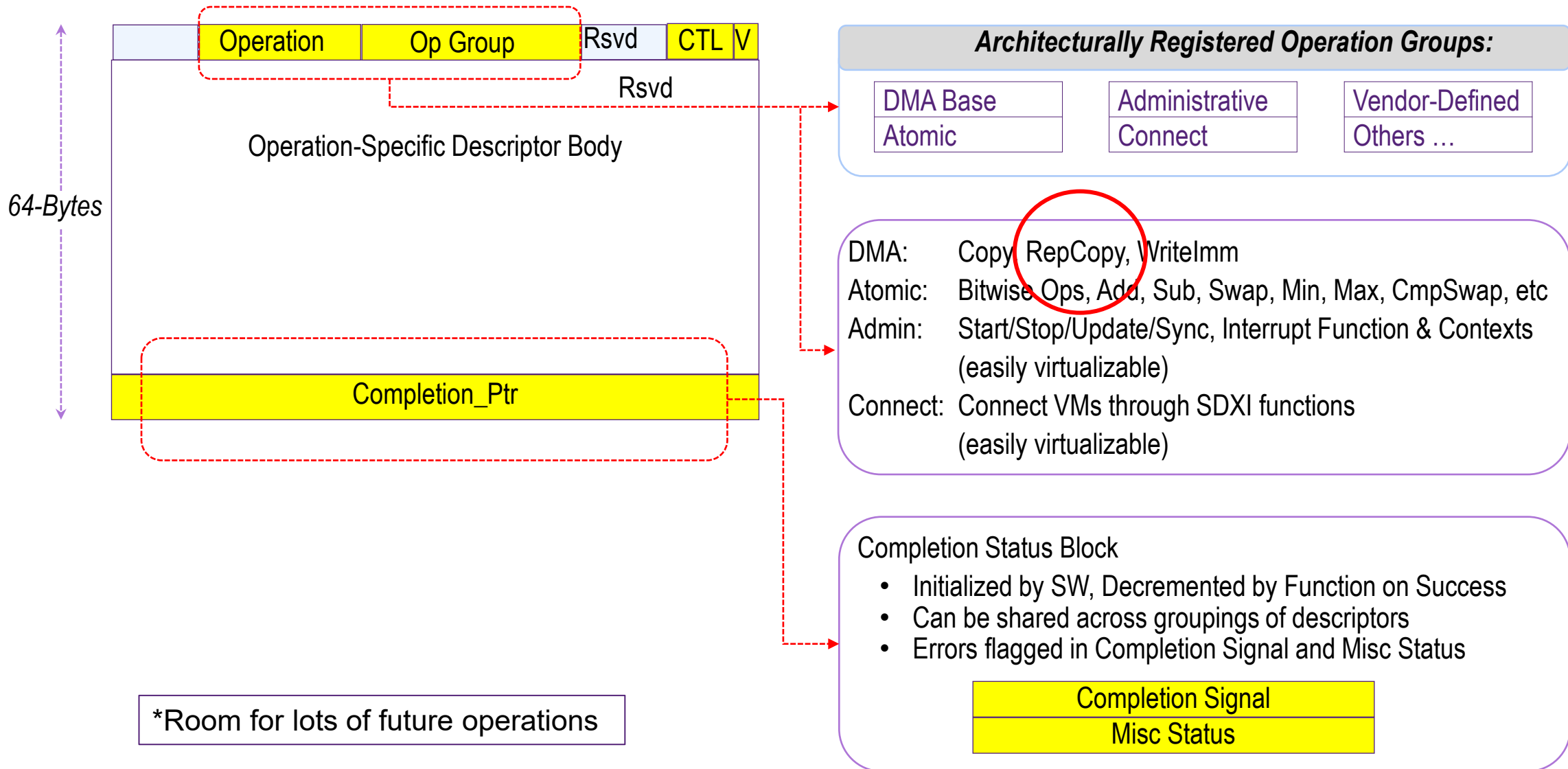
- Function setup & control accelerator-independent.
- One standard descriptor format.
- All SDXI context state resides in memory.
  - No special mechanisms to serialize state.
- Very easy to virtualize.

# SDXI Function Architecture

- Ring state directly managed by user mode.
- Multiple contexts per function.
  - Allows easier multi-producer support within one VM.
- One way to setup, control, and quiesce contexts.
- One flexible way to log errors.

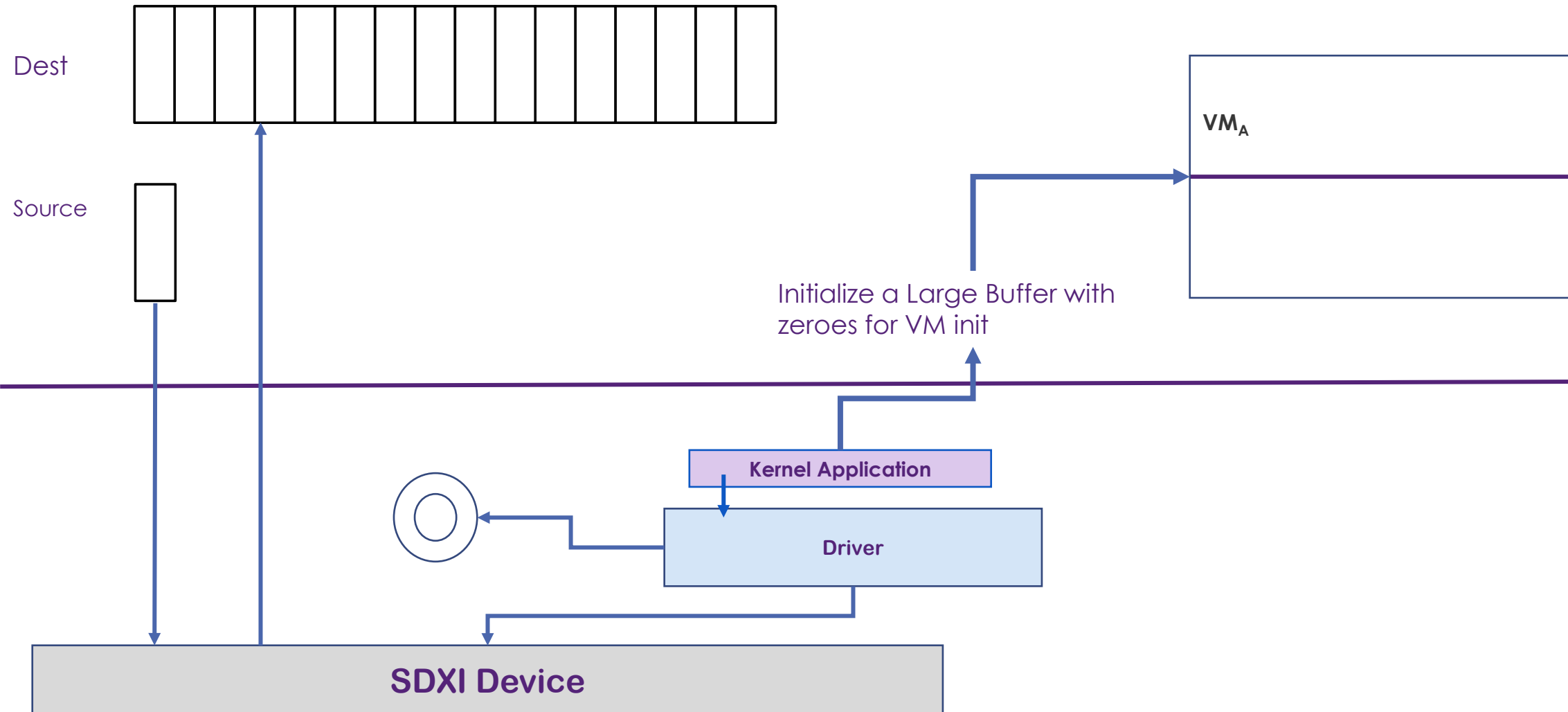


# A Standard Descriptor Format (1)

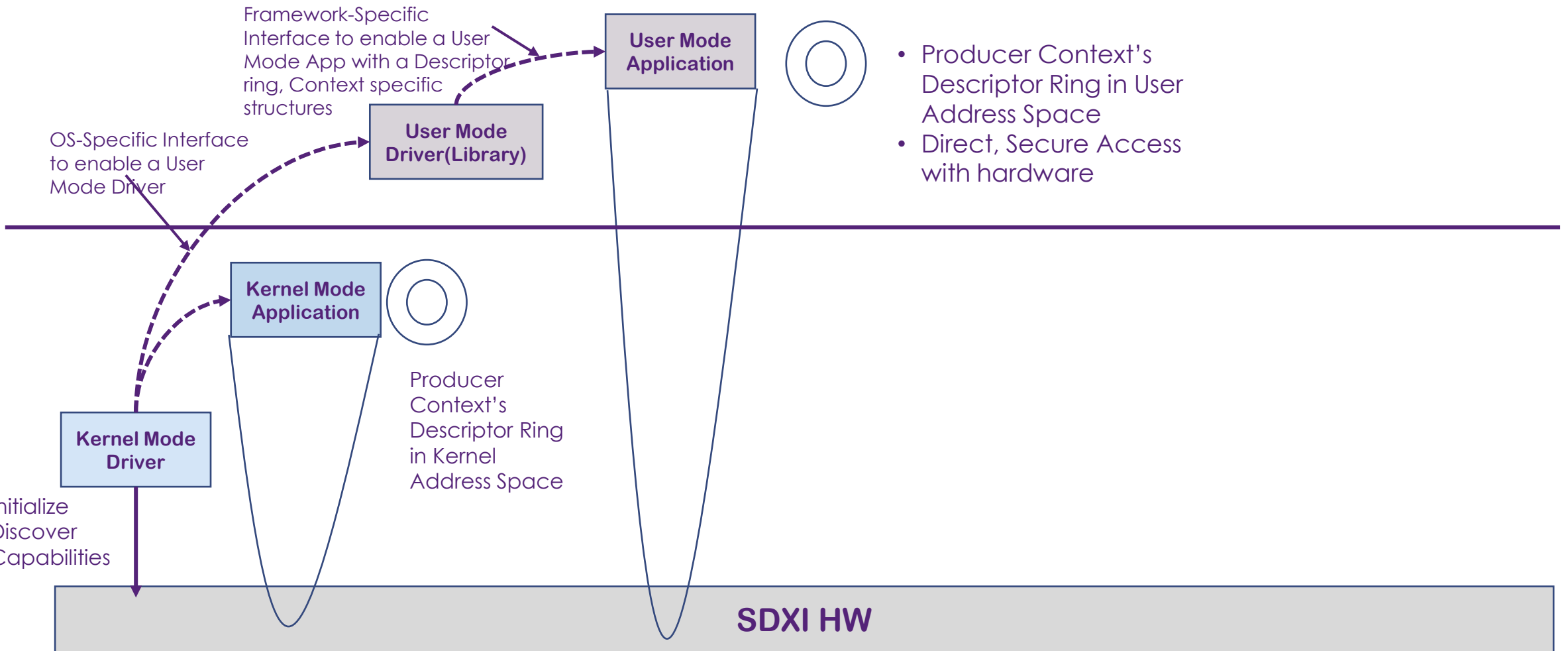




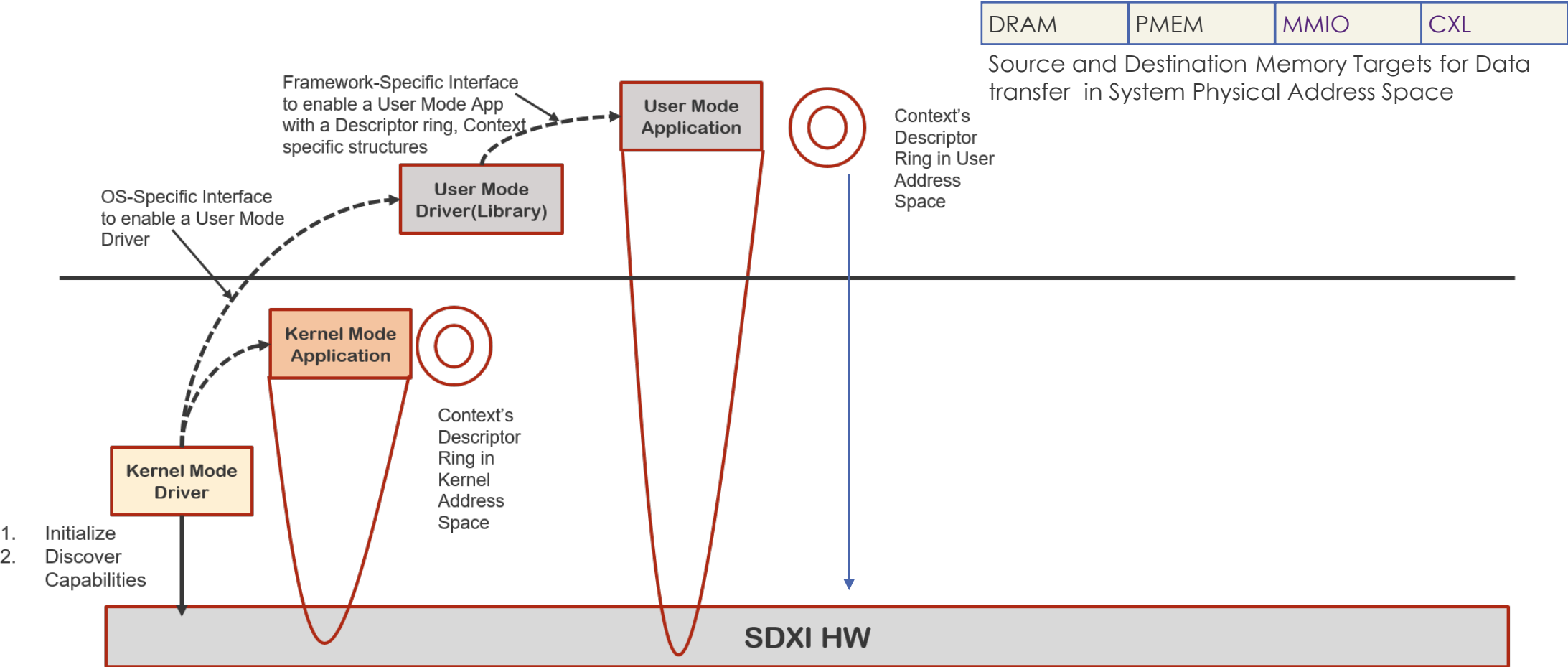
# RepCopy Example



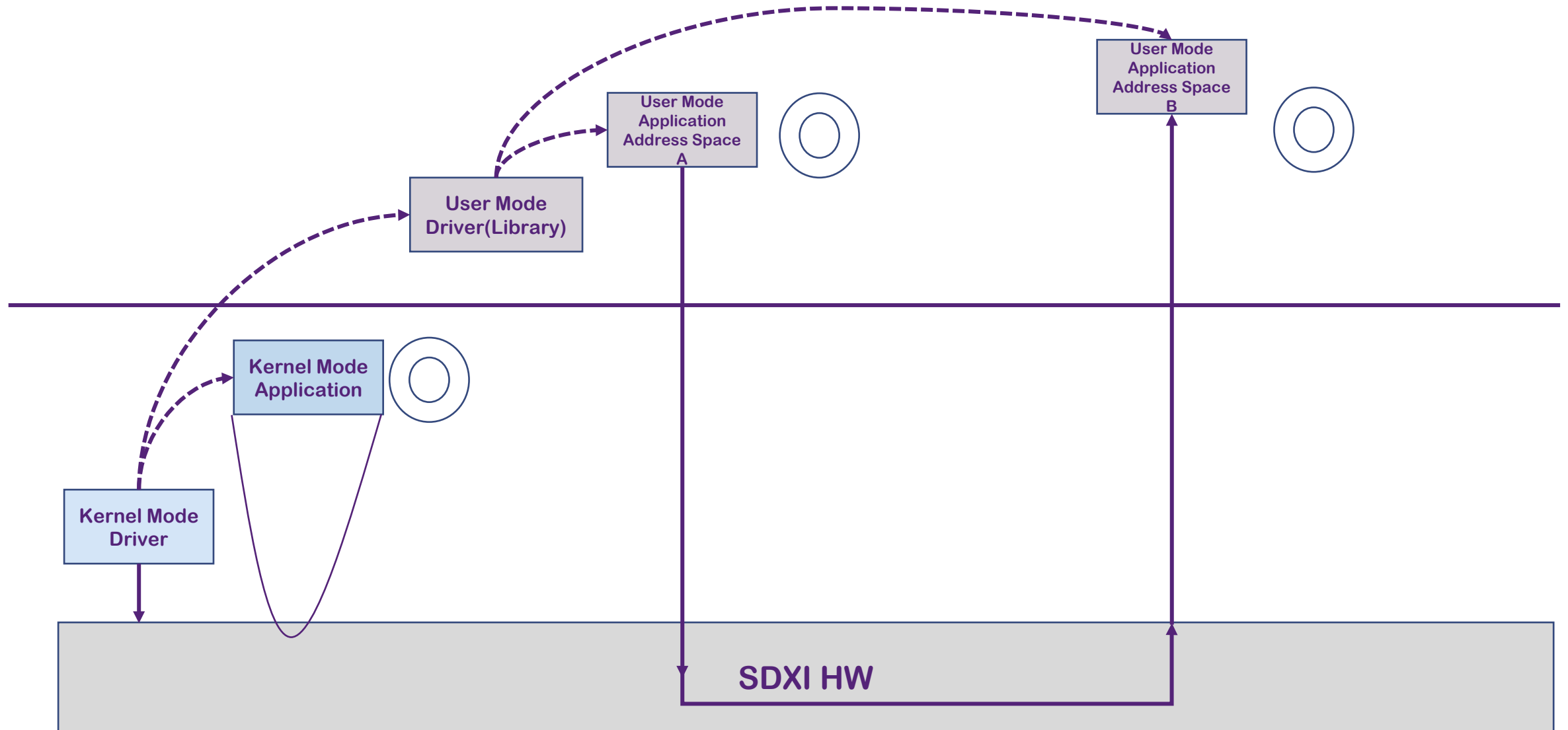
# Baremetal Stack View



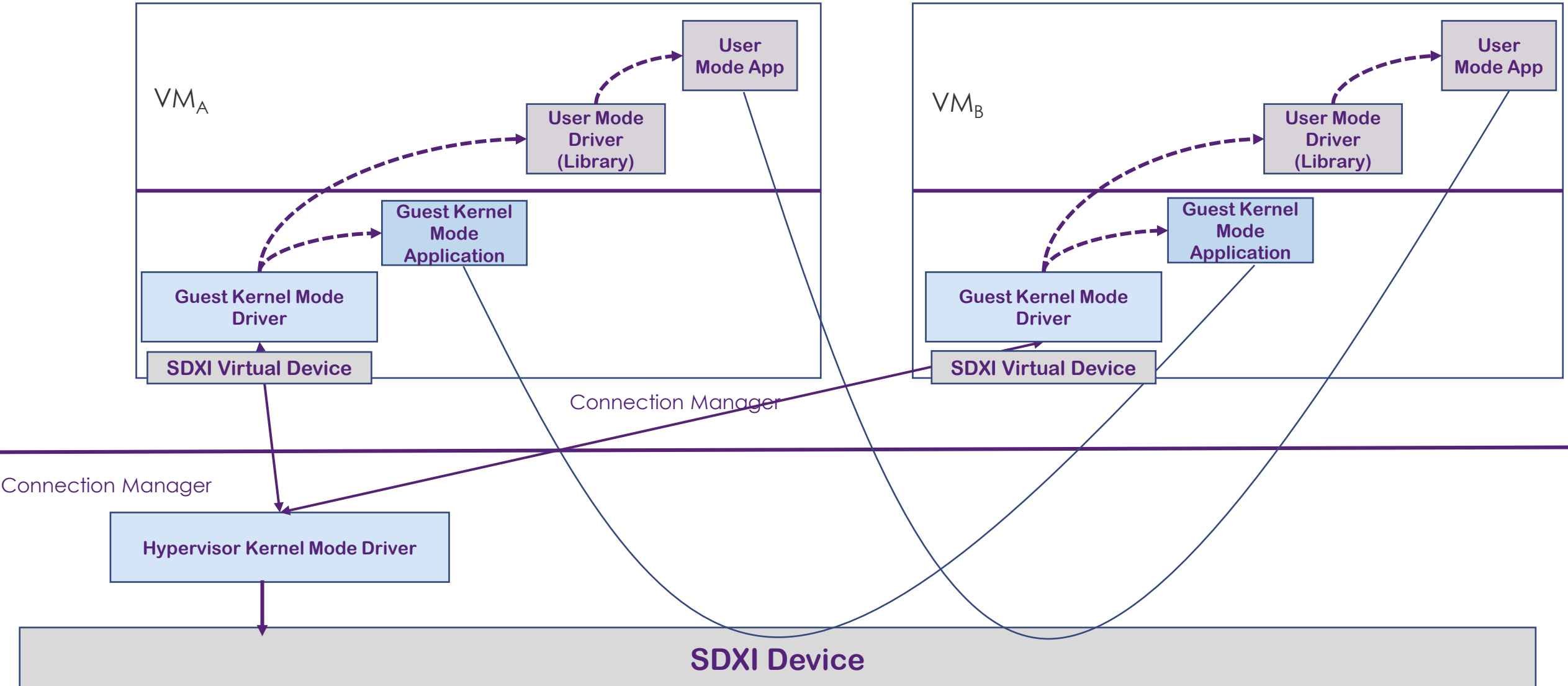
# Direct HW access, Tier across Memory Tiers



# Scale Baremetal Apps – Multi-Address Space



# Scale with Compute Virtualization – Multi-VM address space



# SDXI TWG 2020 Accomplishments

- SDXI(Smart Data Accelerator Interface) is a proposed SNIA standard for a memory to memory Data Mover interface.
- Charter and Program of Work
  - <https://members.snia.org/document/dl/31306>
- Recent completed work - summary
  - TWG approved by SNIA Technical Council and Board in July 2020
  - TWG had its first meeting on July 30, 2020
  - AMD, Dell, VMware provided contributed v0.7 draft specification as a starting point for this TWG(September)
  - 22 Companies and over 50 individual members and growing..

# SDXI TWG 2021 Work Items

- Initial RFC review process – Completed
  - Work on building an OS-independent Software reference - Underway
  - Initiate Public review of a pre 1.0 draft specification
  - Release a v1.0 SNIA architecture document
  - Plan and begin work on Post v1.0 features
  - Work on designing Compliance Testing tools
- SNIA Group collaboration
  - Leverage expertise in other SNIA groups around persistent memory.
  - Coordinate synergies with Computational Storage Work Group
- External group collaboration / Alliance work items
  - PCI SIG, CXL, Gen-Z

# SDXI TWG's Program of Work

- Post v1.0 Focus
  - New data mover operations for smart acceleration
  - Data mover operations involving persistent memory targets
  - Cache coherency models for data movers
  - Security Features involving data movers
  - Connection Management architecture for data movers
- Encourage adopting companies to work towards compliant software implementations and driver models.
- Educate and encourage adoption by OS, Hypervisors, OEMs, Applications and Data Acceleration vendors



# SDXI TWG Membership as of 4-14-2021

- Advanced Micro Devices(AMD)
- ARM
- Broadcom
- Dell Inc.
- Fujitsu America Inc.
- Futurewei Technologies, Inc
- Hewlett Packard Enterprise
- Huawei Technologies Co. Ltd
- IBM(includes Red Hat, Inc)
- Inspur Electronic Information Industry Co Ltd.
- MemVerge
- Micron
- Microsemi a Microchip Company
- Microsoft Corporation
- NetApp
- NGD Systems, Inc
- Samsung Electronics Co., LTD
- Scaleflux
- SK Hynix
- VMware
- Western Digital
- Xilinx, Inc



## 1. How to get more involved ?

- <https://www.snia.org/sdxi>
- Participate in public review of draft specification – **Coming Soon!**

## 2. Need more details ?

- [SDC 2020 Conference presentation on SDXI](#)
- <https://www.youtube.com/watch?v=iv2GUfnxG-A>

## 3. Questions ?

- [sdxitwgchair@snia.org](mailto:sdxitwgchair@snia.org)

## 4. Acknowledgement

- SDXI TWG members – Rich Brunner, Philip NG, Glen Sescila, Don Dutile, Beau Beachamp, Jason Wohglemuth, Murali Ravirala, Frederick Knight, Alexandre Romana, Dwight Riley, Paul Hartke, Paul Von Stamwitz, James Leighton, Srinivas Gowda , Bill Martin & others

# Thank you

Please visit [www.snia.org/pm-summit](http://www.snia.org/pm-summit) for presentations

