MANAGING PERSISTENT MEMORY

Linux Foundation Vault 2016

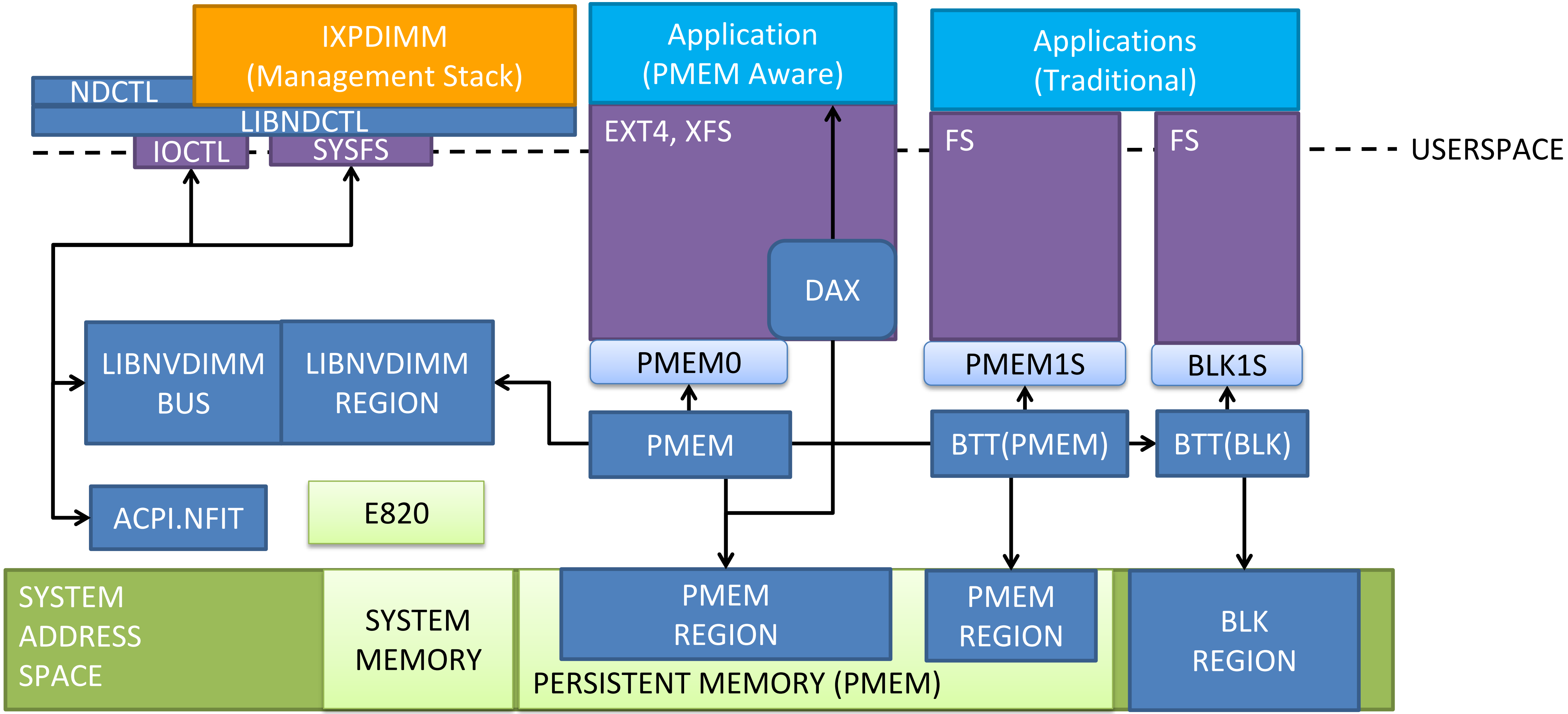
Dan Williams & Tiffany Kasanicky

/dev/pmem0

THANK YOU

/dev/ndblk0.0s

Managing Persistent Memory



Namespaces

```
# ndctl list --namespaces --type=pmem
{
  "dev":"namespace6.0",
  "mode":"raw",
  "size":33554432,
  "uuid":"70a6adce-722e-4ab8-b698-35eaea9750b3",
  "blockdev":"pmem6"
}
```

Namespaces

```
# ndctl list --namespaces --type=pmem
{
  "dev":"namespace6.0",
  "mode":"raw",
  "size":33554432,
  "uuid":"70a6adce-722e-4ab8-b698-35eaea9750b3",
  "blockdev":"pmem6"
}
```

Namespaces

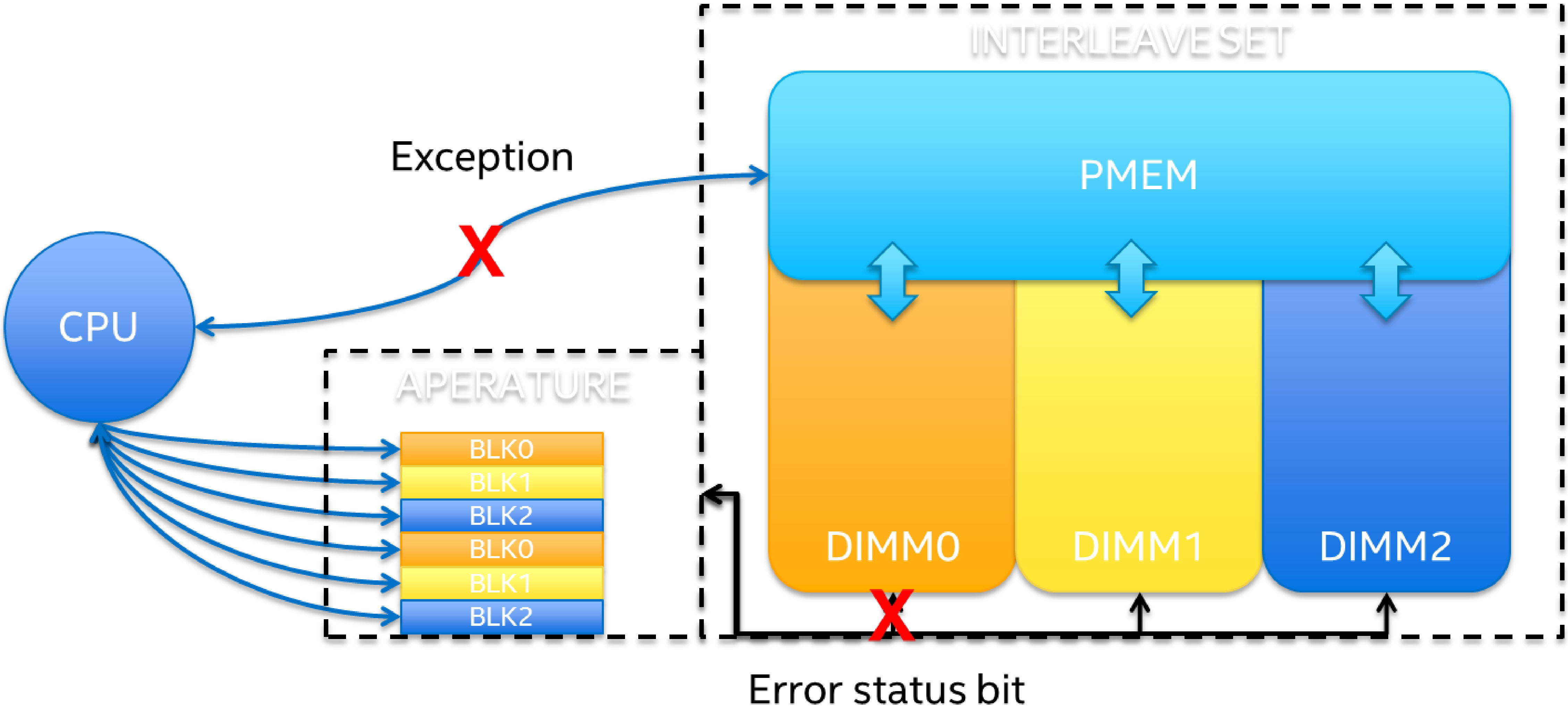
```
# ndctl list --namespaces --type=pmem
{
  "dev":"namespace6.0",
  "mode":"raw",
  "size":33554432,
  "uuid":"70a6adce-722e-4ab8-b698-35eaea9750b3",
  "blockdev":"pmem6"
}
```

“Namespace”: Persistent memory capacity accessed through a PMEM or BLK disk device

Namespaces

```
# ndctl list --namespaces --type=blk
{
  "dev":"namespace0.0",
  "mode":"sector",
  "uuid":"5ce6c34a-88b0-469a-86f5-ea8f462a68ca",
  "sector_size":4096,
  "blockdev":"ndblk0.0s"
}
```

Why BLK?


INTERLEAVE SET
Exception
PMEM
CPU
APERATURE
BLK0
BLK1
BLK2
BLK0
BLK1
BLK2
DIMM0
DIMM1
DIMM2
Error status bit

CONFIGURATION FOR APPLICATIONS

Namespace Modes

RAW

SECTOR

MEMORY

Namespace Modes

RAW

- Byte-addressable
- Limited DAX

SECTOR

MEMORY

Namespace Modes

RAW

- Byte-addressable
- Limited DAX

SECTOR

- Software atomic sector update semantics
- Configurable sector size.
- Applicable to PMEM and BLK namespaces

MEMORY

Namespace Modes

RAW

- Byte-addressable
- Limited DAX

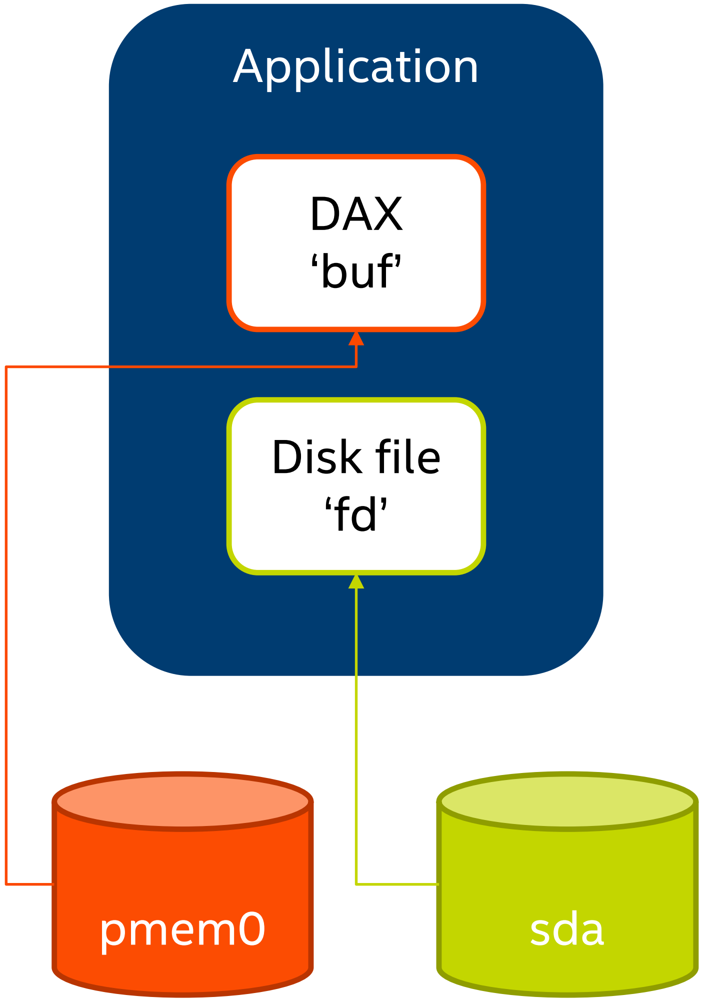
SECTOR

- Software atomic sector update semantics
- Configurable sector size.
- Applicable to PMEM and BLK namespaces

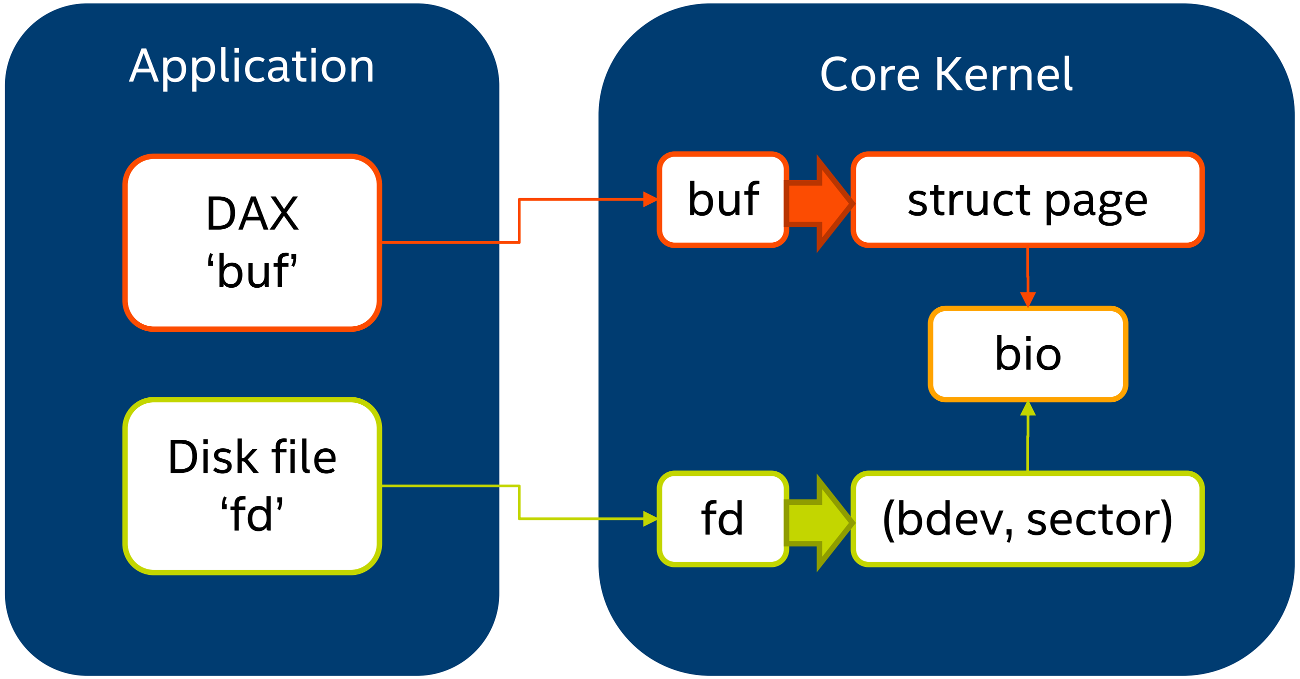
MEMORY

- Enables full DAX (DMA/RDMA/Direct-I/O)
- Only applicable to PMEM namespaces

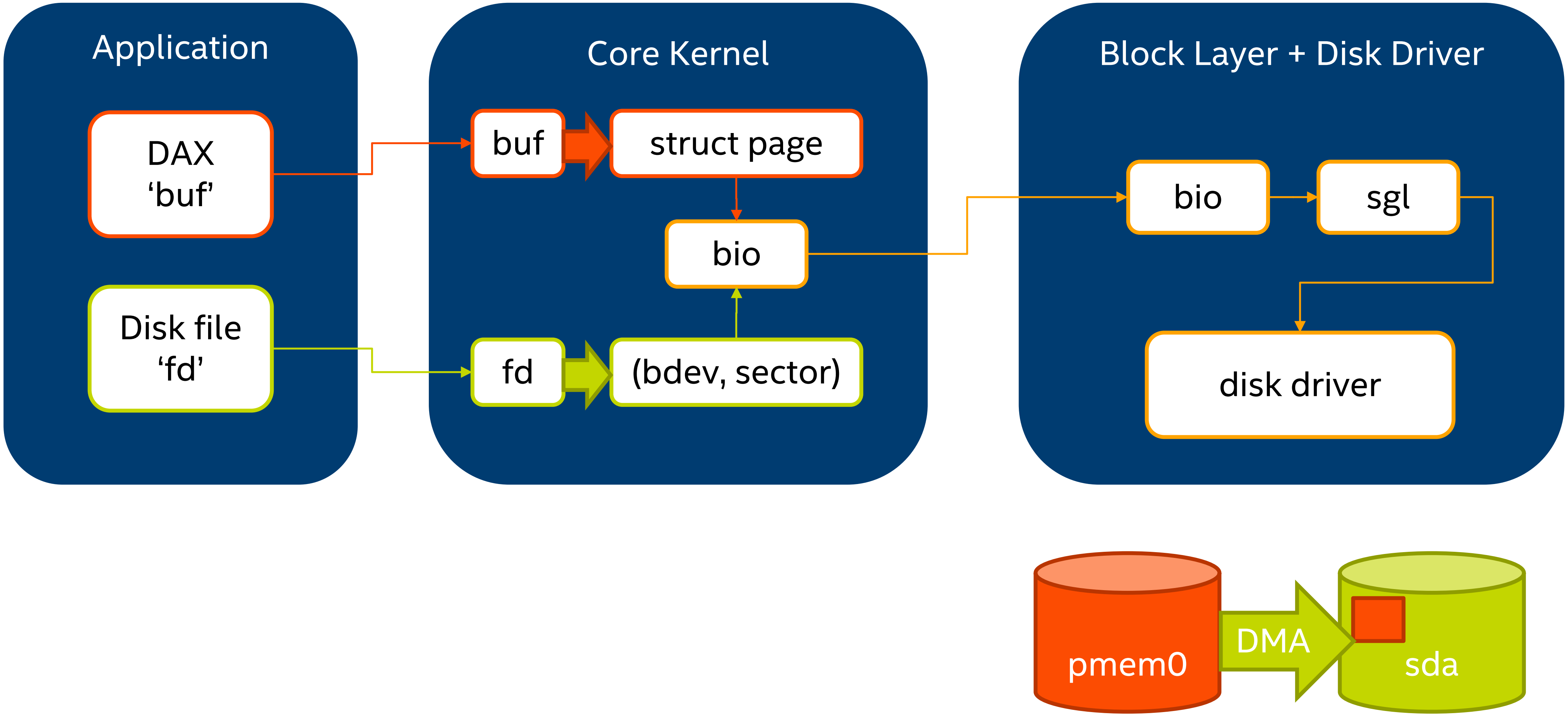
“Memory” Mode DAX: Direct I/O



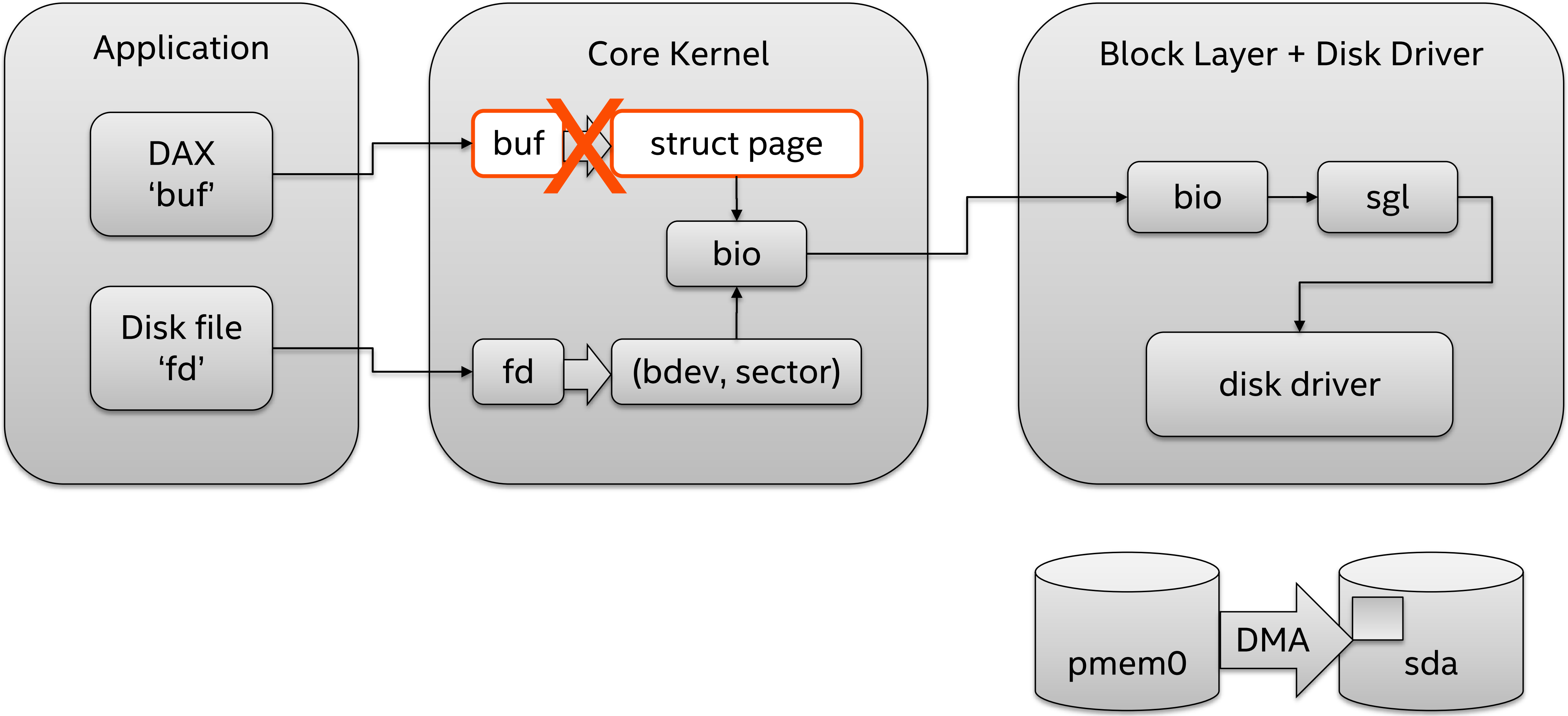
“Memory” Mode DAX: Direct I/O



“Memory” Mode DAX: Direct I/O



“Memory” Mode DAX: Direct I/O



“Memory” Mode DAX: Considerations

- struct page array is 64-bytes per 4K page (16GB per 1TB)

```
# ndctl create-namespace --reconfig=namespace9.0 --mode=memory --map=dev --force
```

ndctl/libndctl < IXPDIMM

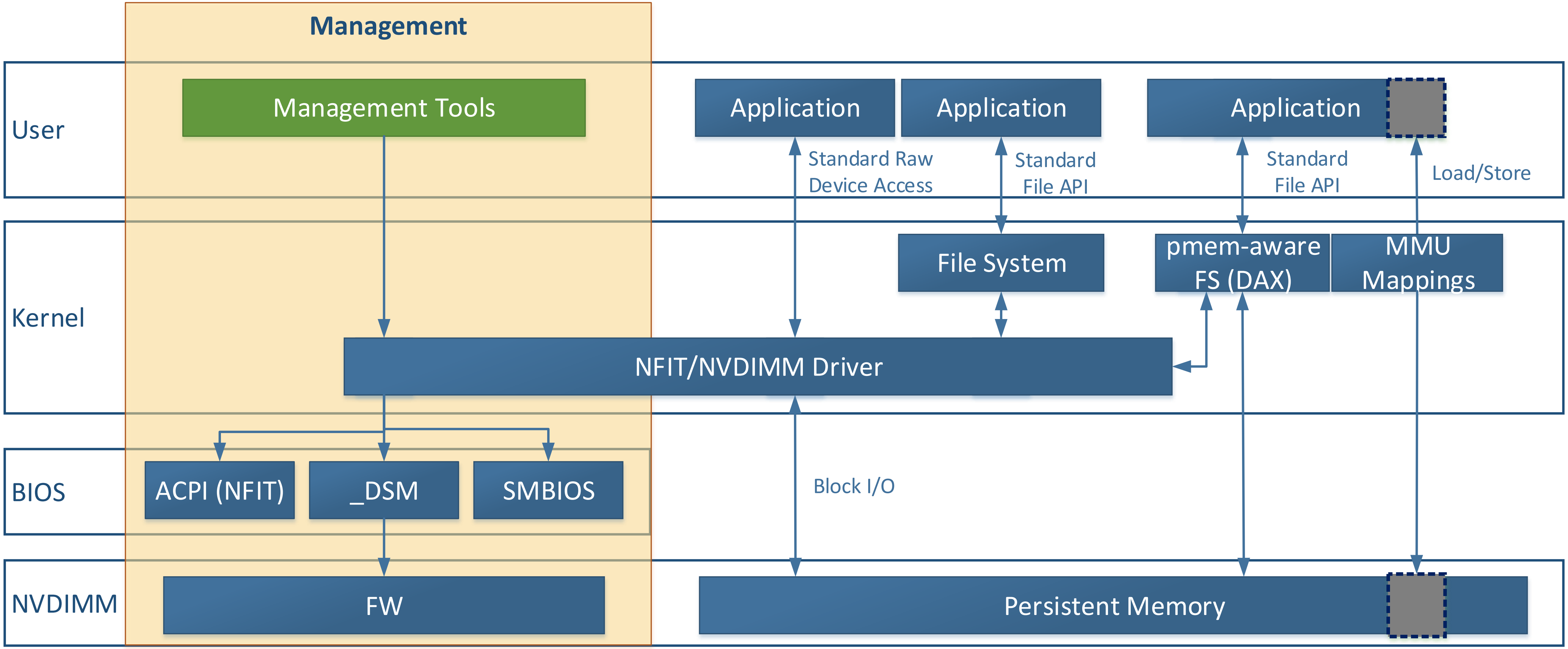
ndctl/libndctl: low level generic primitives

IXPDIMM: Coherent / comprehensive management stack

INTEL IXPDIMM SOFTWARE

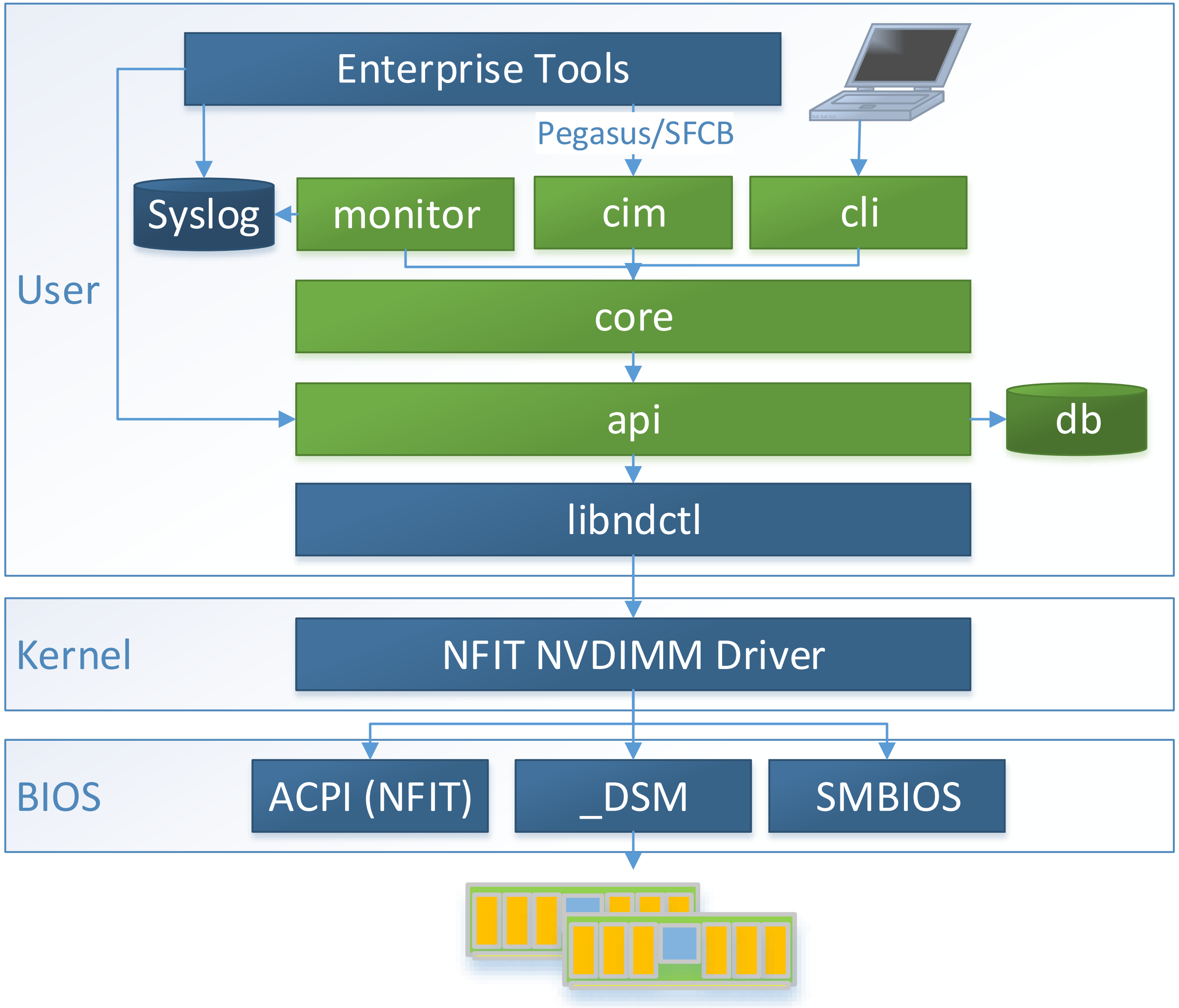
Tiffany Kasanicky
tiffany.j.kasanicky@intel.com

Persistent Memory



Components

- Basic Management
  - End-user provisioning and management via CLI
- Enabling
  - SFCB/Pegasus CIM provider for remote access and 3rd party integration
  - C library for programmatic access and abstraction
- Monitoring
  - Daemon for health monitoring

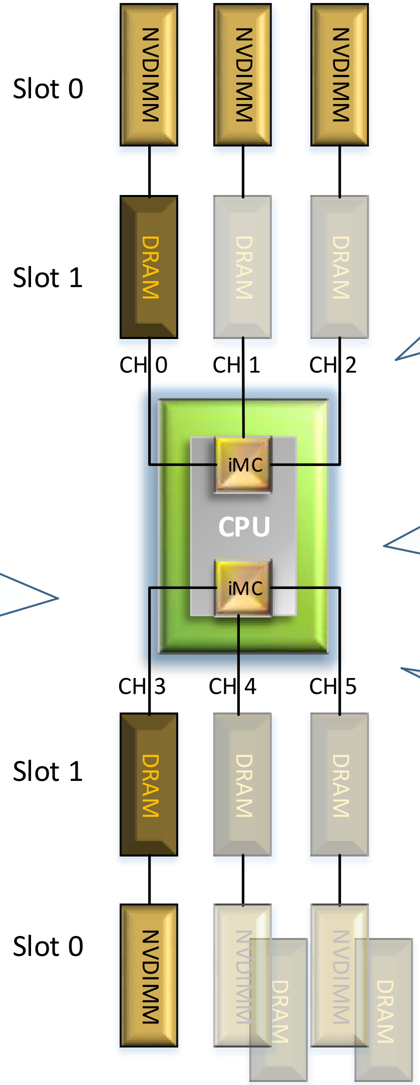


NVDIMM Management

Discovery:
DRAM + NVDIMM topology, identifying information, capabilities

Memory Configuration:
Volatile and persistent partitions, interleave settings

Namespaces:
Create/delete/inventory namespaces



Sensors & Settings:
Thermal, wear, spare, power, errors

Performance:
Bytes read/written, host and block read/writes

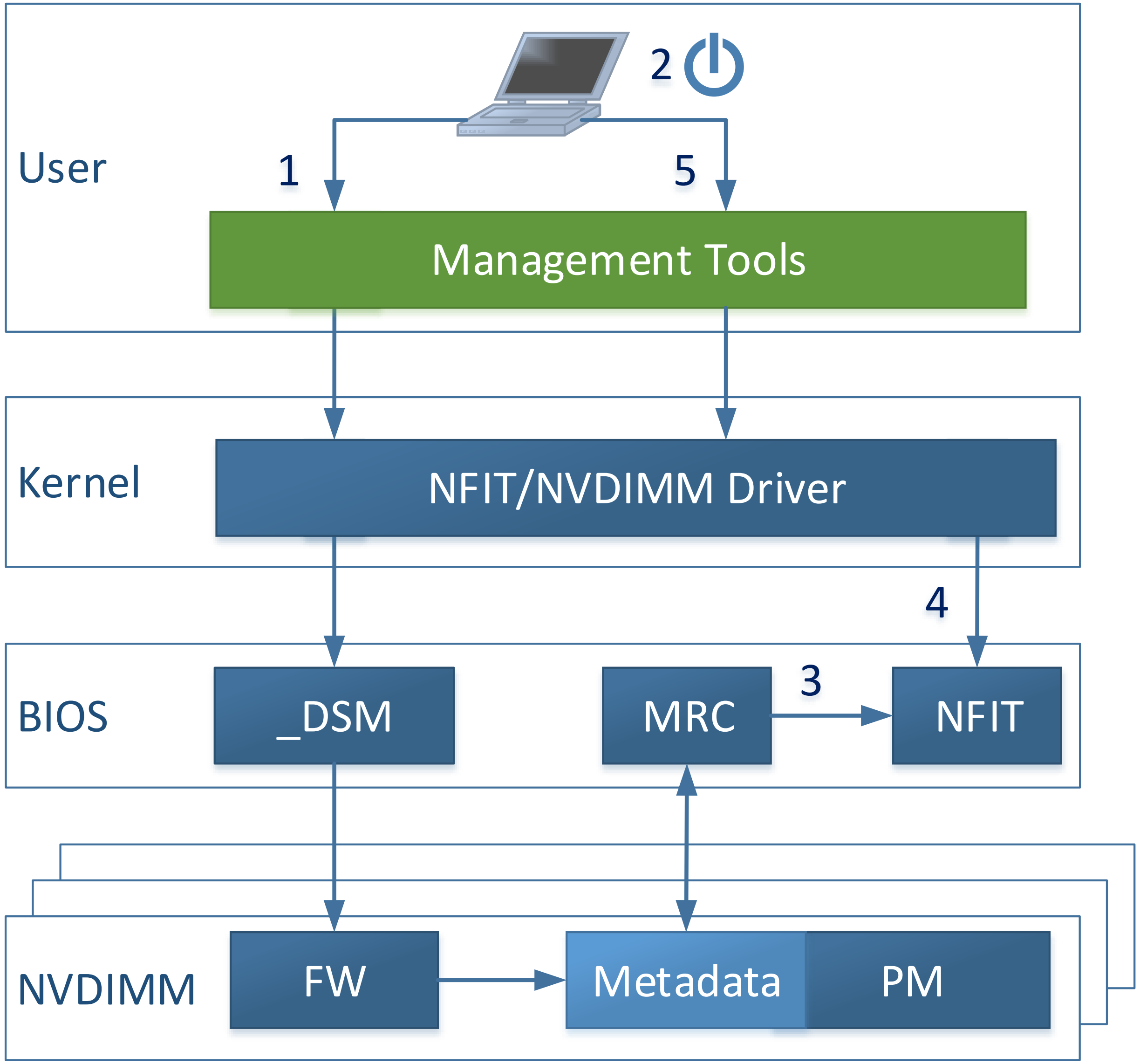
Instrumentation:
FW update, SW versioning, data-at-rest security, FW settings and policies

Discovery

- DRAM/NVDIMM Topology
  - SMBIOS Type 17 (memory device) data
  - NVDIMM socket, memory controller, channel population
- Aggregated Memory Resources
- Capabilities
  - Platform BIOS, NVDIMM, FW, SW
- NVDIMM Information
  - Identifying - serial number, model number, device ID
  - Status – manageability, health, security
  - Provisioning – partitioning, attributes, state

Memory Provisioning

1. Create memory allocation goal
2. Reboot
3. BIOS writes NFIT
4. Driver reads NFIT
5. Create namespace
6. Mount file system



Diagnostics

- Quick Health Check
- Platform Configuration Check
- Security Check
- FW Consistency and Settings Check
- Persistent Memory Metadata Check
- Address Range Scrub Results

Packages

Component	Package	Repository
cli	ixpdimm-cli	https://github.com/01org/IXPDIMMSW
cim	libixpdimm-cim	https://github.com/01org/IXPDIMMSW
core	libixpdimm-core	https://github.com/01org/IXPDIMMSW
api	libixpdimm-api libixpdimm-api-devel	https://github.com/01org/IXPDIMMSW
monitor	ixpdimm-monitor	https://github.com/01org/IXPDIMMSW
cli framework	libintelnvm-cli libintelnvm-cli-devel	https://github.com/01org/intelnvmclilibrary
i18n framework	libintelnvm-i18n libintelnvm-i18n-devel	https://github.com/01org/intelnvmi18nlibrary
cim framework	libintelnvm-cim libintelnvm-cim-devel	https://github.com/01org/intelnvmcimlibrary

Distribution Plan

- Open source 3-clause BSD license
- Hosted on 01.org/github – Intel maintainers
  - https://01.org/ixpdimm-sw
  - https://01.org/intel-nvm-cim-library
  - https://01.org/intel-nvm-cli-library
  - https://01.org/intel-nvm-i18n-library
- Targeted OS Distributions:
  - RHEL/Fedora
  - SLES/OpenSuSE

tiffany.j.kasanicky@intel.com