Obligatory rubric

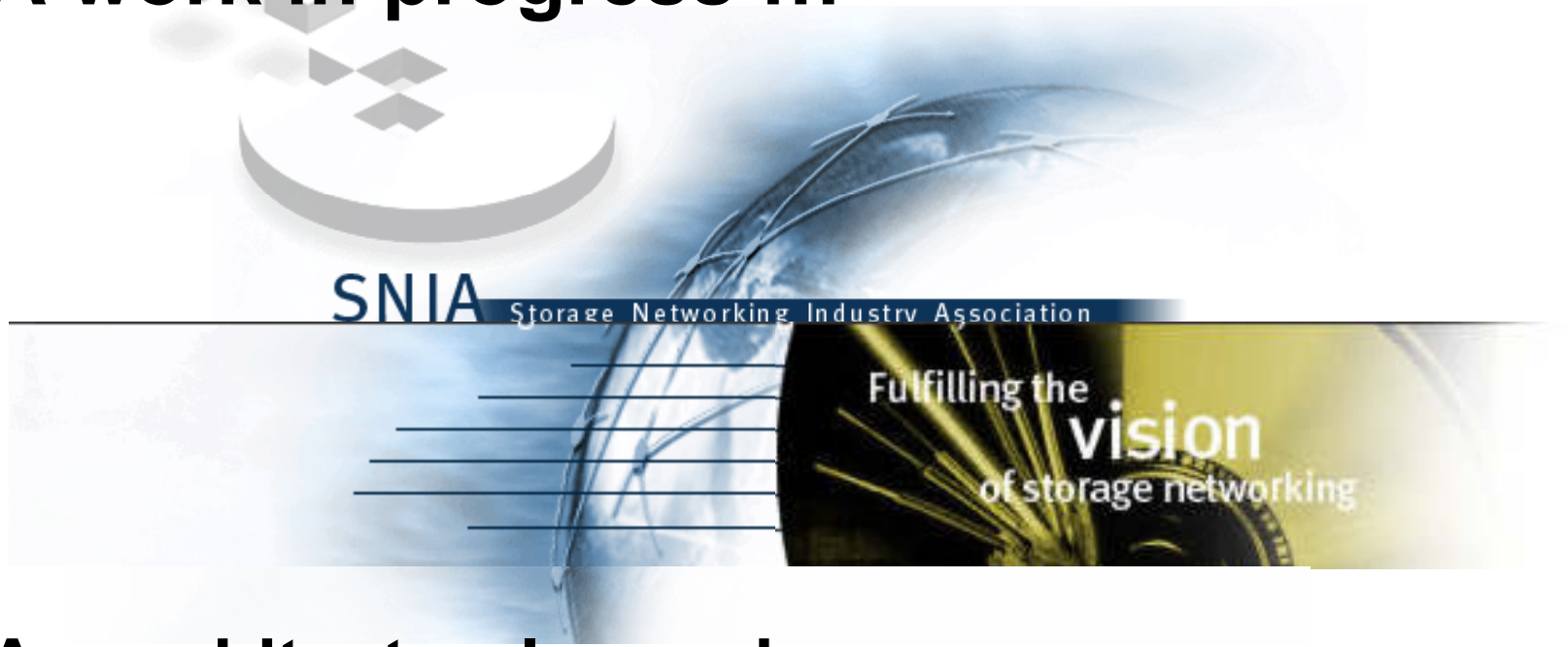
- Copyright © 2000,2003 Storage Networking Industry Association (SNIA).
- SNIA-TC members may use this material freely; all others require permission from the SNIA.

- See the speaker notes for this slide for important usage information.
- See SNIA-SSM-colors.ppt for printing and color/graphics conventions.

SNIA
Storage Networking Industry Association

SNIA shared-storage model

A work in progress …



An architectural overview

This revision:

- 2001-06-05 last content update
- 2003-04-13 last graphics update

Contents

- **Purpose**
- **The SNIA storage model**
  - Layers, functions, and services
  - Networks and interfaces
- **Applying the SNIA storage model**
  - Common storage architectures
- **Conclusions**

SNIA
Storage Networking Industry Association

Purpose

- **Present a simple model for shared storage architectures**
- **Use it to describe common examples graphically**
- **Expose, for each one:**
  - What services are provided, where
  - Where interoperability is required
  - [future] Pros and cons of the architecture

SNIA
Storage Networking Industry Association

Benefits

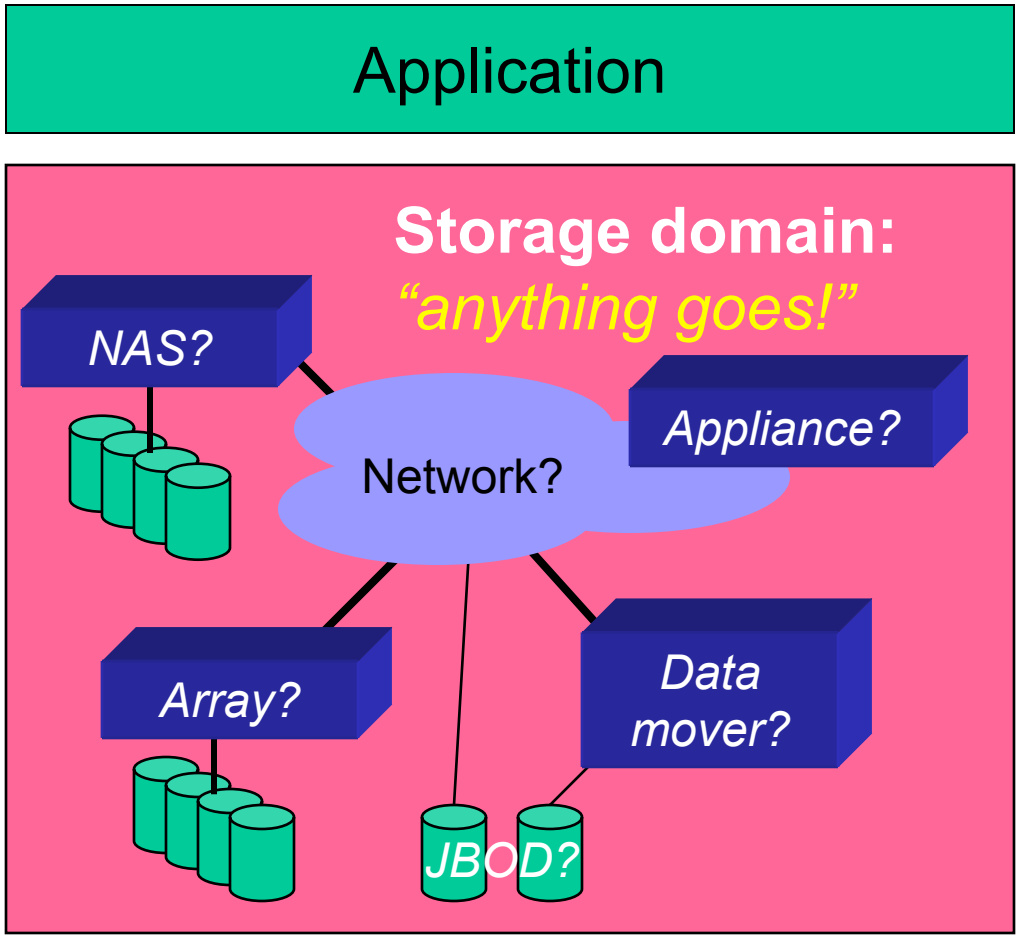
- **A common “architecture vocabulary”**
- **Reference comparisons between common solutions**
- **Help to align the industry**
  - Customers can better structure their choices
  - Vendors can better explain their differences

SNIA
Storage Networking Industry Association

What the model is and is not

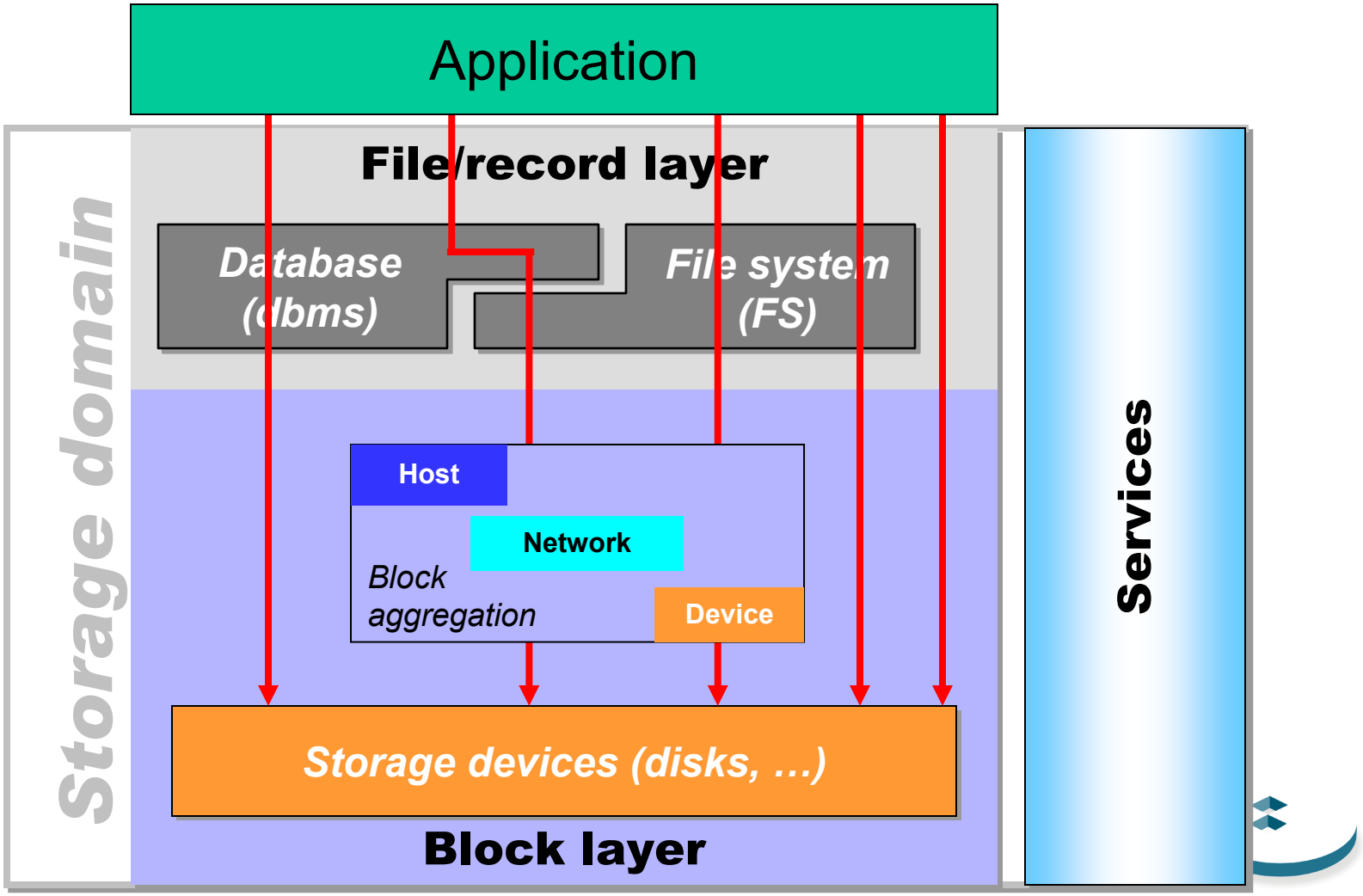
- **It is not:**
  - A specification, an architecture, a design, a product, a recommendation, or an installation
- **It is:**
  - A framework that captures the functional layers and properties of a storage system

Classic storage model



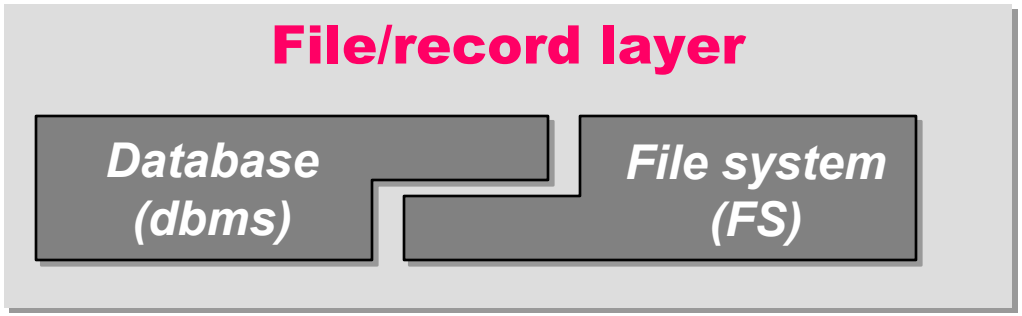
SNIA
Storage Networking Industry Association

The SNIA shared storage model



SNIA
Storage Networking Industry Association

The SNIA storage model:
File/record layer



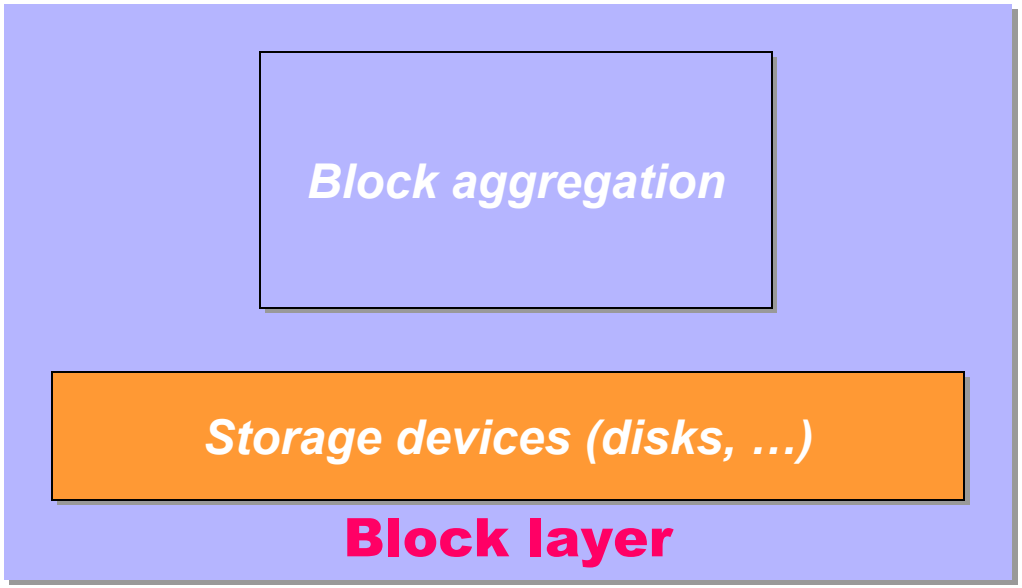
SNIA
Storage Networking Industry Association

The SNIA storage model:
File/record layer — functions

- **Aka "access methods"**
  - File system, database
- **Primary responsibility: packing many smaller things into a few larger ones**
  - Fine-grain naming, space allocation
- **Secondary responsibilities**
  - Caching for performance
  - Coherency in distributed systems

SNIA
Storage Networking Industry Association

The SNIA storage model:
Block layer



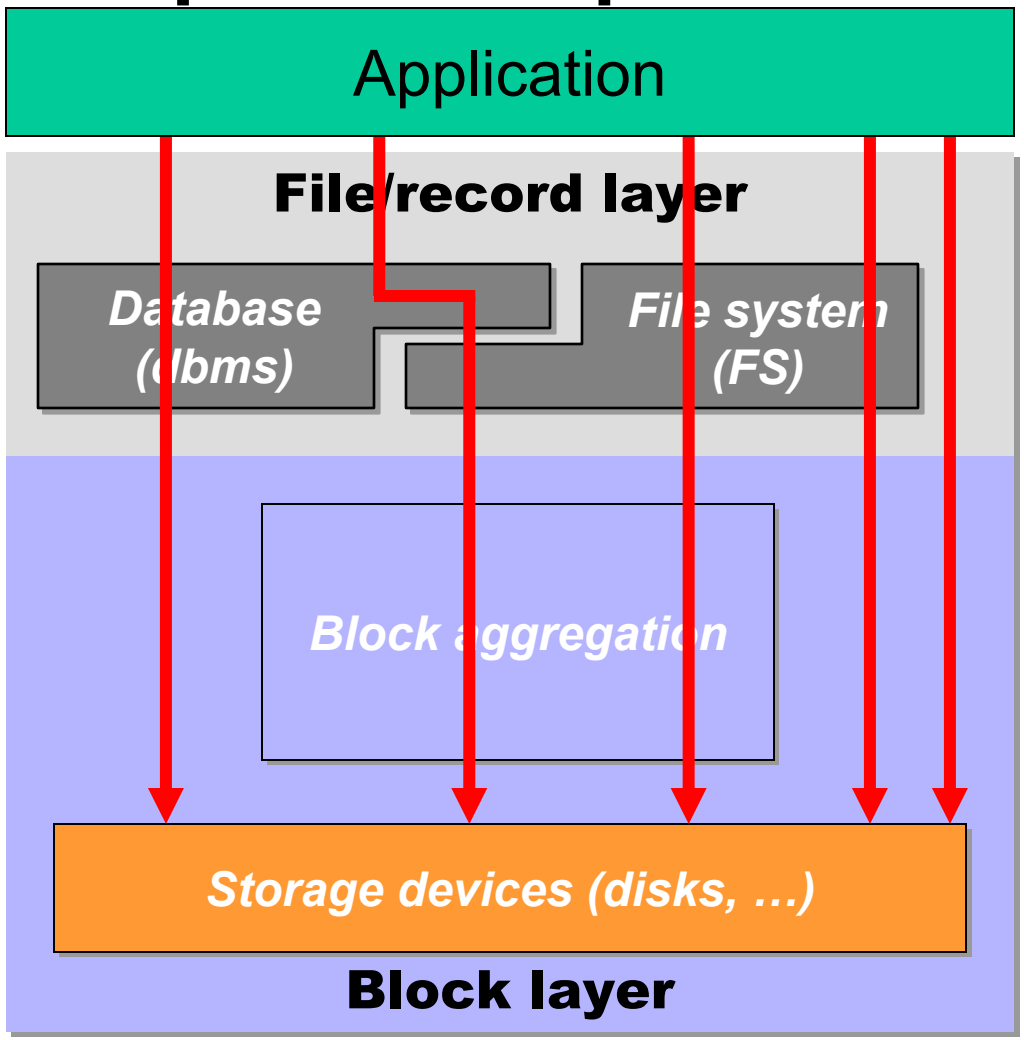
SNIA
Storage Networking Industry Association

The SNIA storage model:
Block layer — functions

- **Storage devices – storing data**
  - disk drives, tape drives, solid-state disks, …
- **Block aggregation – address mapping**
  - in-SN aggregation, or “virtualization”
  - slicing & concatenation, striping
  - local & remote mirroring, RAID-n
- **Examples**
  - volume managers
  - disk array LUs
- **Secondary responsibilities**
  - caching

SNIA
Storage Networking Industry Association

The SNIA storage model:
Access path examples

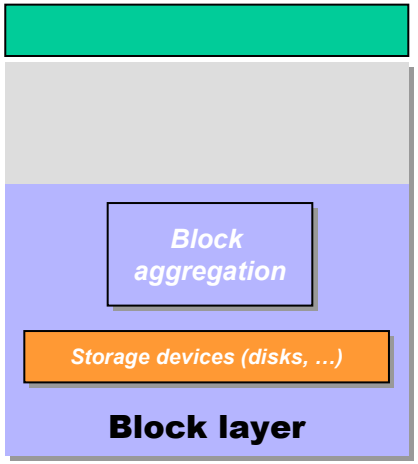


Note: all 8 possible paths can be used!

SNIA
Storage Networking Industry Association

Block layer

- **Block-mapping functions: what can be done**
- **Functional decomposition: where it can be done**
- **Sample architectures**



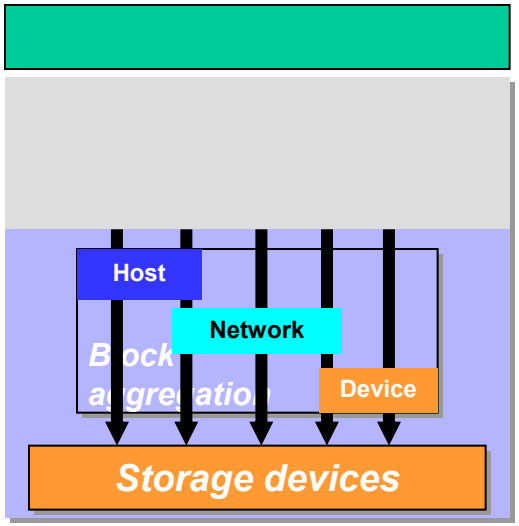
SNIA
Storage Networking Industry Association

Block layer
What can be done

- **Space management**
  - making a large store from many small ones
  - packing many small stores into one large one
- **Striping**
  - for performance (load balancing, throughput)
- **Redundancy**
  - full (local & remote mirroring, RAID-1, -10, …)
  - partial (RAID-3, -4, -5, …)
  - point-in-time copy

Block layer
Where it can be done

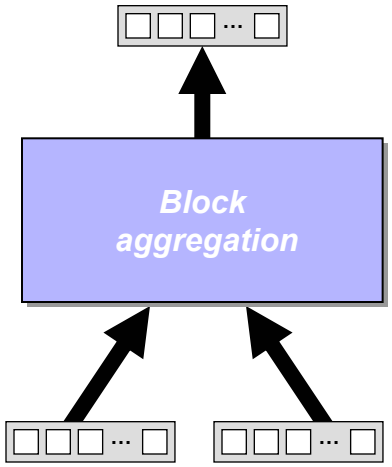
- **Host-side**
  - logical volume managers
  - device drivers, HBAs
- **SN-based**
  - HBAs, specialized SN appliances
- **Device-based**
  - array controllers (e.g., RAID)
  - disk controllers (e.g., sparing)



SNIA
Storage Networking Industry Association

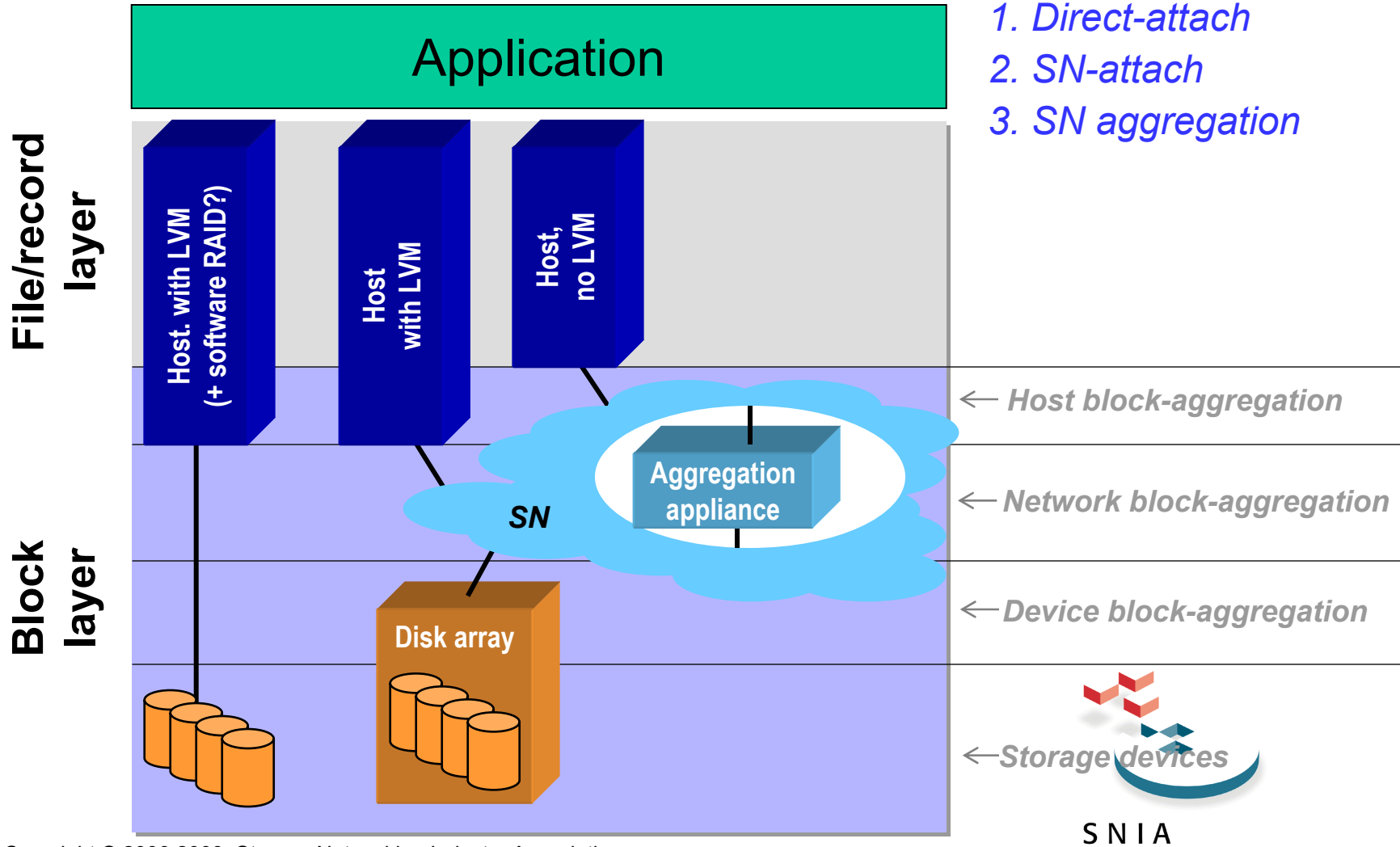
Block layer
How it is done

- **Building blocks**
  - input: vector of blocks
  - output: vector of blocks
- **Result: building blocks can be stacked**
  - enables the 3 layer model for the block layer
  - layers can be nested on one another
  - could be extended to more layers



SNIA
Storage Networking Industry Association

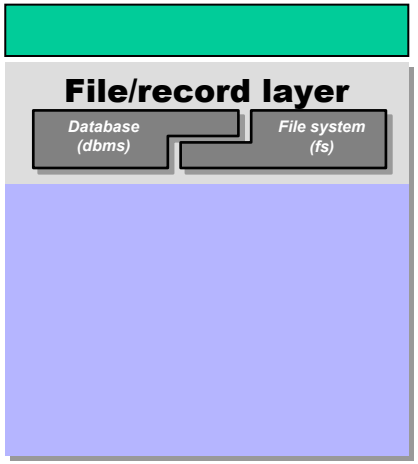
Block layer
Sample architectures



SNIA
Storage Networking Industry Association

File/record layer

- **Byte-mapping functions: what can be done**
- **Functional decomposition: where it can be done**
- **Sample architectures**



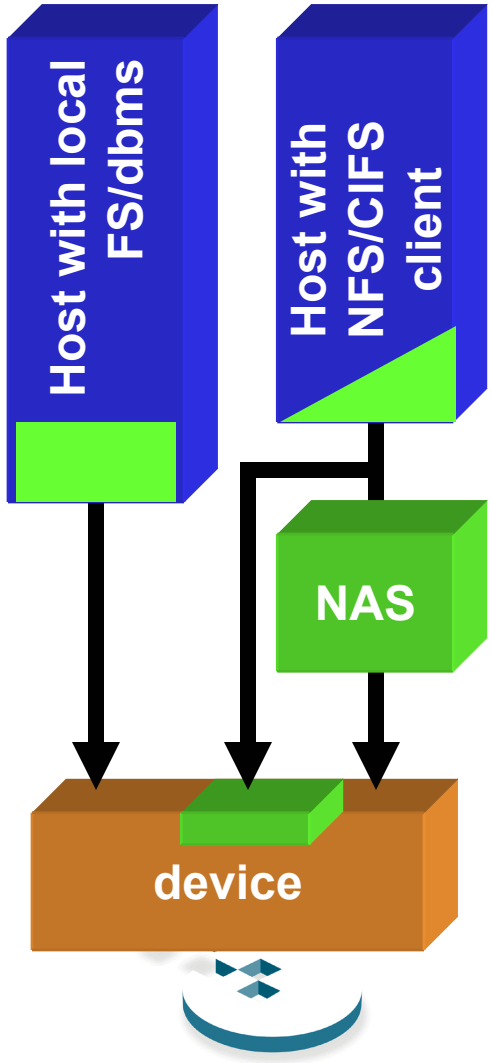
SNIA
Storage Networking Industry Association

File/record layer
What can be done

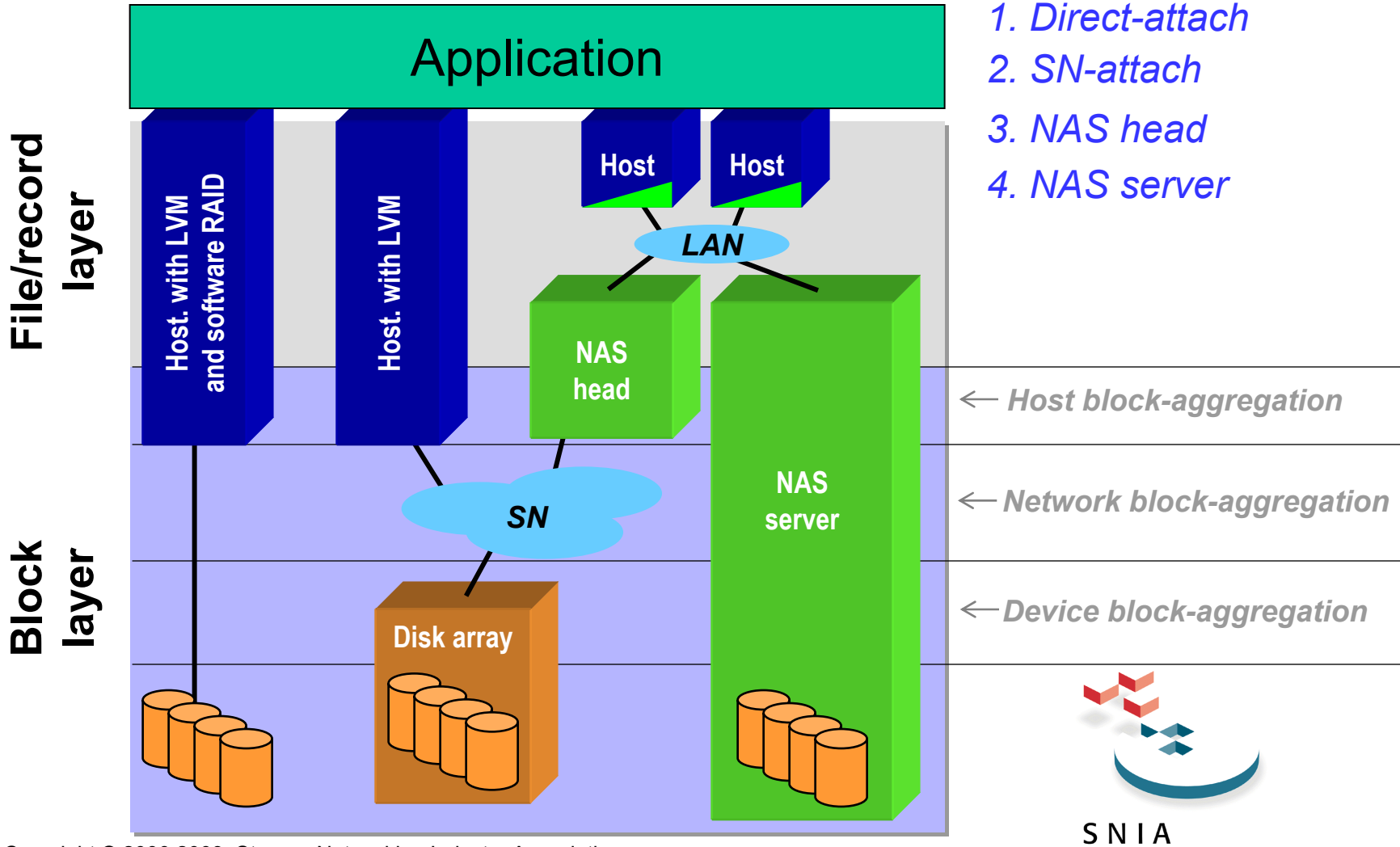
- **Database management systems**
  - tuples → tables
  - tables → table-spaces
  - table-spaces → volume
- **File systems**
  - files → volume
- **New types**
  - http caches: a kind of distributed file system?

File/record layer
Where it can be done

- **Host-side**
  - file systems and databases
  - NFS, CIFS, etc. are client-server splits inside the file system
- **SN-based**
  - NAS head
- **Device-based**
  - NAS functions in array box



File/record layer



SNIA
Storage Networking Industry Association

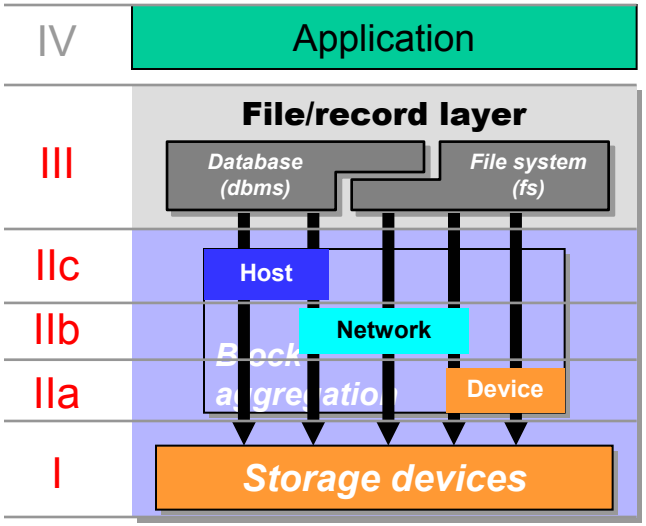
The SNIA storage model
A layered view

IV. Application

III. File/record layer
- IIIa. Database
- IIIb. File system

II. Block aggregation
- IIa. Host
- IIb. Network
- IIc. Device

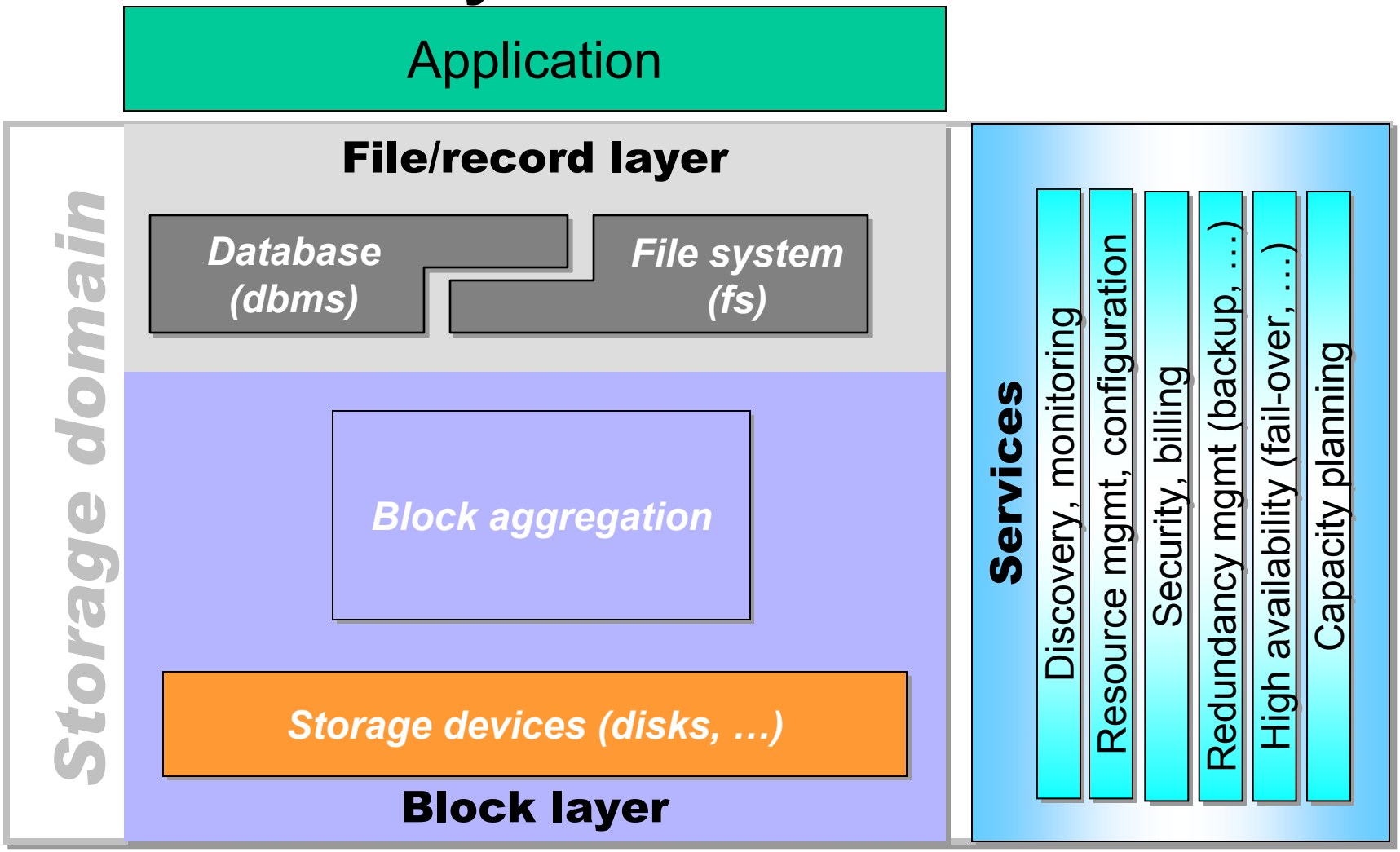
I. Storage devices



SNIA
Storage Networking Industry Association

The SNIA storage model

Services subsystem



Services

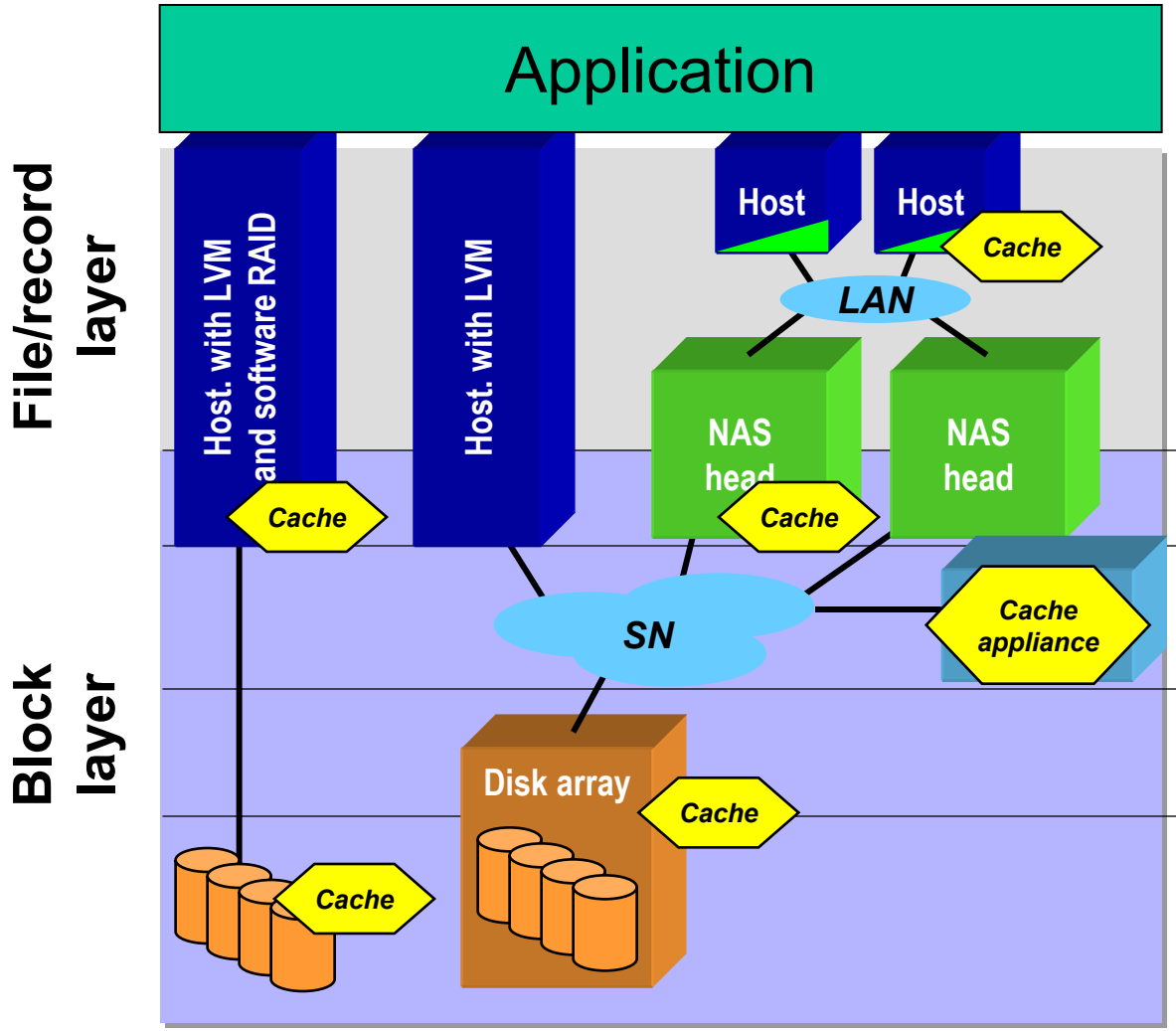
- **Operations off the critical path**
  - naming, discovery, monitoring, configuration, security, billing, redundancy management (backup, …), high availability management (fail-over, …), capacity planning, …
  - strong ties into system-wide management services

- **Vital for successful operation**
  - and a major opportunity for SNIA …
  - … but not discussed further in this presentation

SNIA
Storage Networking Industry Association

Caching

… can be added to almost any layer

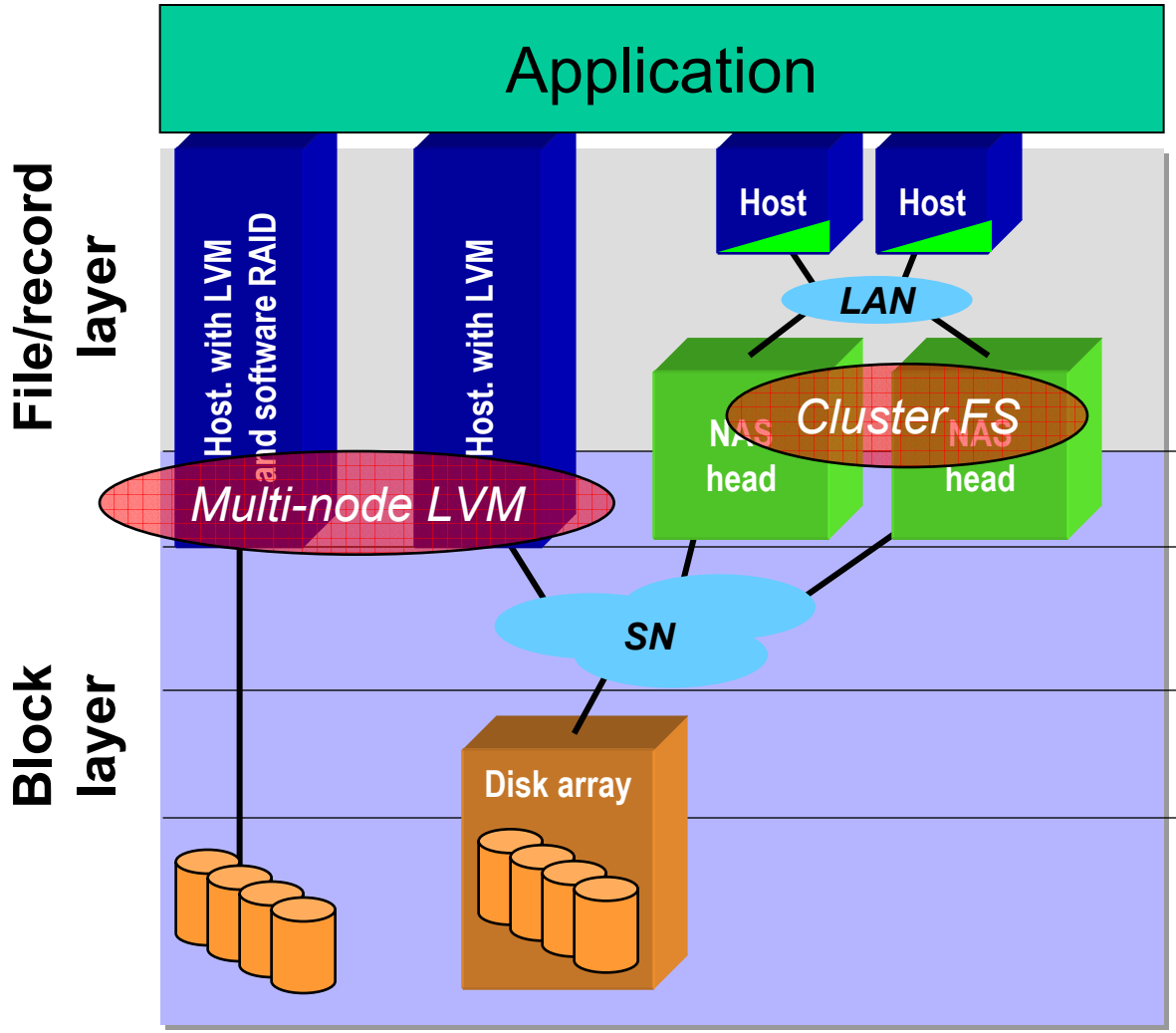


Ideally, caching only affects performance.

But: coherency implications do affect management (protocols, interfaces)

SNIA
Storage Networking Industry Association

Clustering
Inter-box aggregation



Purposes:

- ***load spreading across peers (scalability)***
- ***alternate paths (high availability, scalability)***

SNIA
Storage Networking Industry Association

Q: Data versus storage?
A: Putting information into containers

- user: data *(“learning my preferences”)*
  application: container*(“user keystroke history”)*
- application: data *(“user keystroke history file”)*
  file system: container *(“byte vector”)*
- file system: data *(“a named file”)*
  volume system: container *(“blocks in volume”)*
- volume system: data *(“replicated, striped layout”)*
  disk array: container *(“blocks in LU”)*

Sharing
Content sharing or resource sharing?

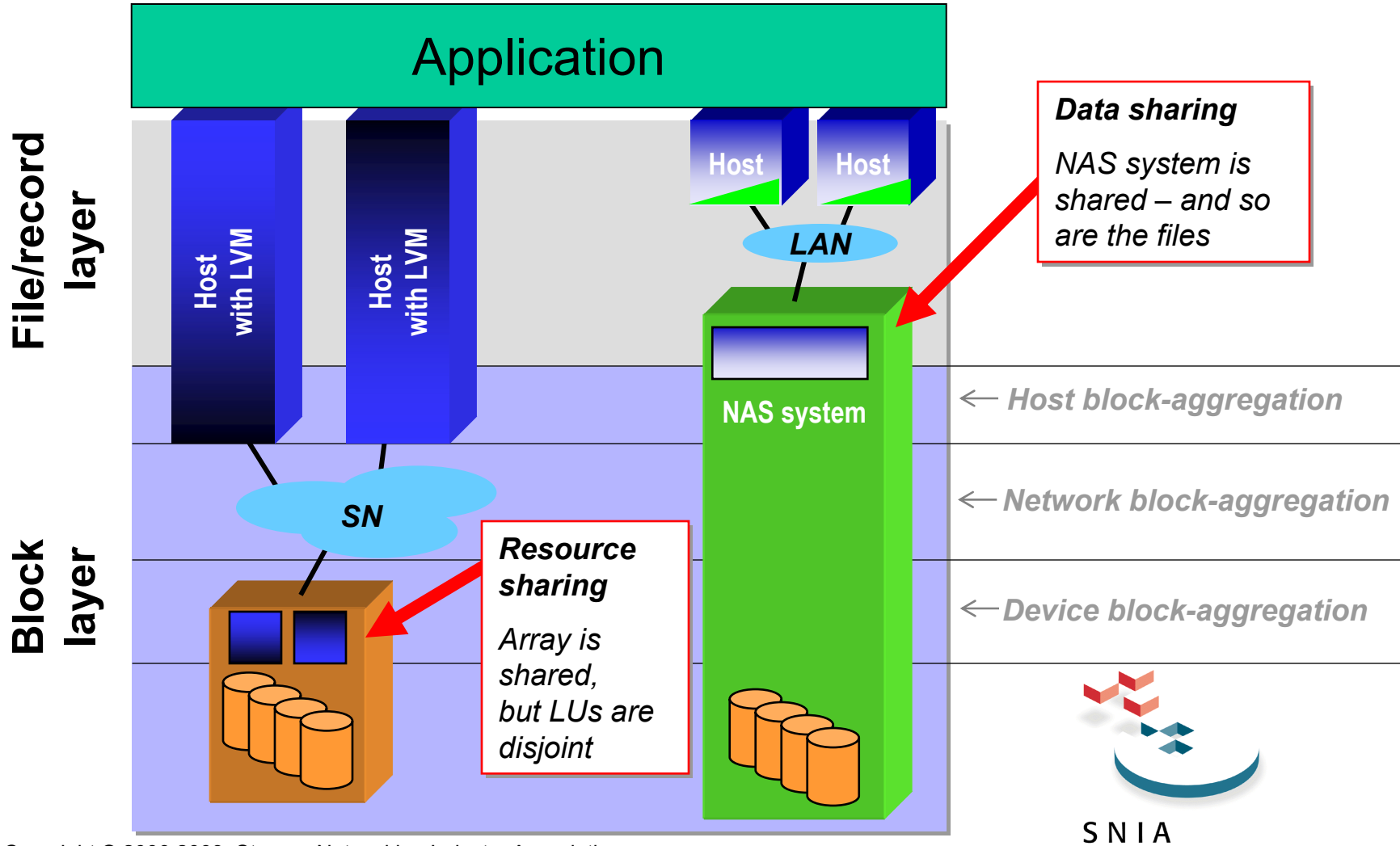
- **Content sharing ("logical", "data")**
  - contents accessed and understood by multiple clients
    - **e.g., file system, Oracle Parallel Server dbms**
  - some of the hard issues:
    - **coherency**
    - **heterogeneous data formats**

- **Resource sharing ("container", "physical")**
  - e.g., disk array where hosts access disjoint LUs

SNIA
Storage Networking Industry Association

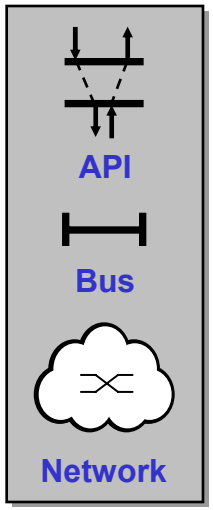
Sharing

Content sharing and resource sharing



SNIA
Storage Networking Industry Association

Networks and interfaces
are pervasive in the model



Application

Operating System

File layer

Block layer

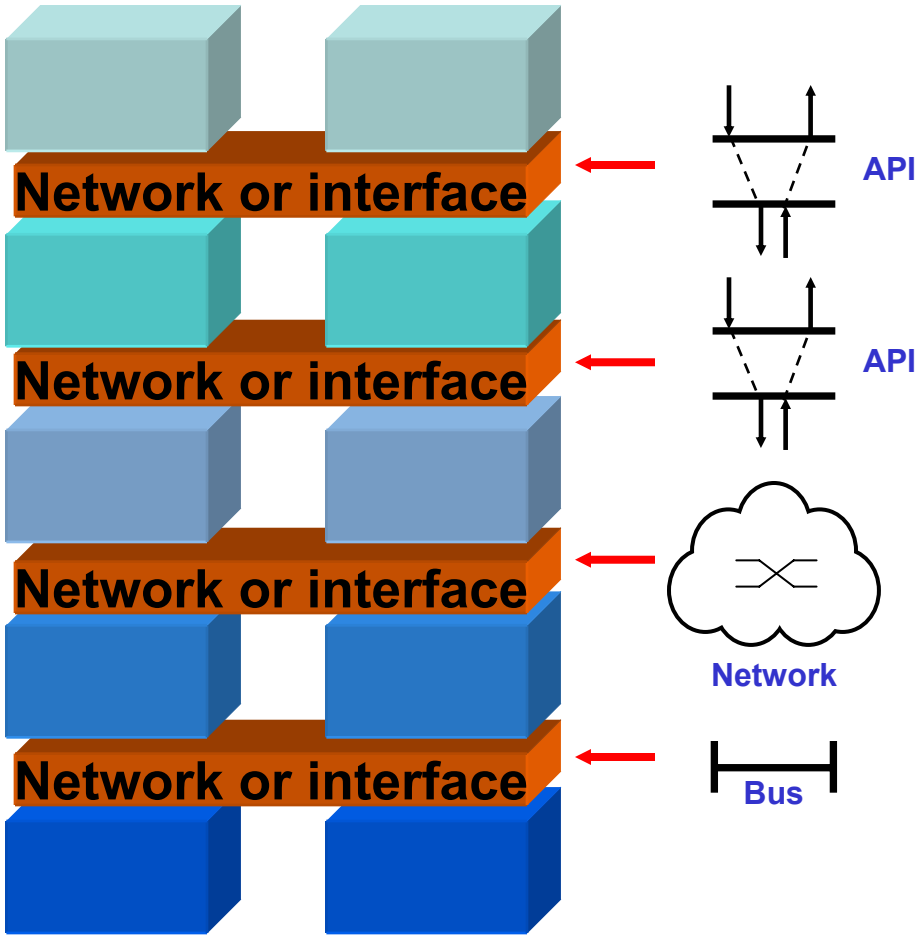
Storage device

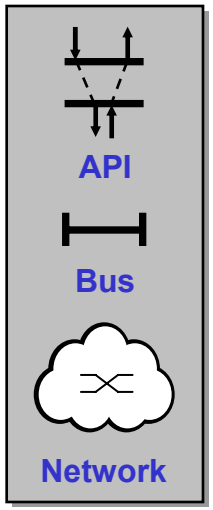
Full benefits come only from ***common, open*** interfaces

Network or interface

SNIA
Storage Networking Industry Association

Networks and interfaces
Composition and scaling

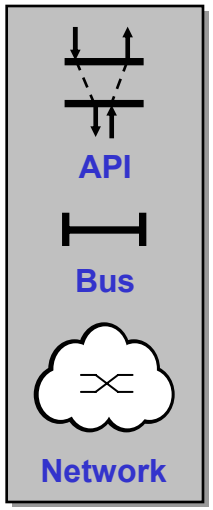




Open interfaces allow:

1. vertical composition
2. horizontal scaling
3. supplier independence

Networks and interfaces
Open interfaces require …



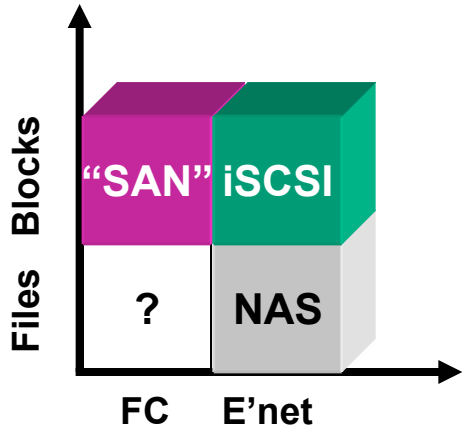
- **Well defined:**
  - *functions* (what they do)
  - *interface protocols* (data formats)
  - *access protocols*
    (system call, RPC, flow control, …)
- **That are:**
  - published
  - supported by multiple products

  => *standards* (which is where SNIA comes in)

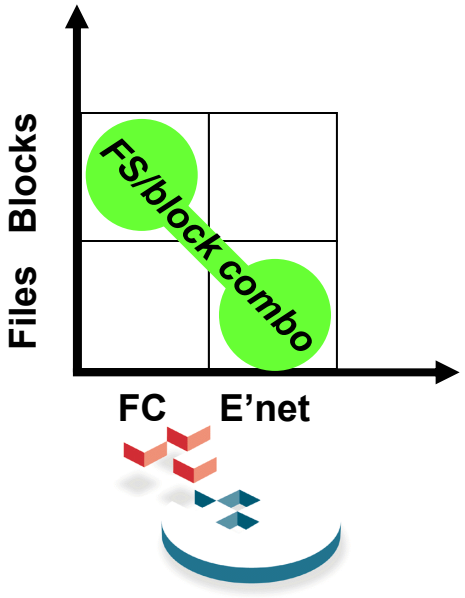
SNIA
Storage Networking Industry Association

Q: “SAN” versus “NAS”?

A: a poorly-formed question



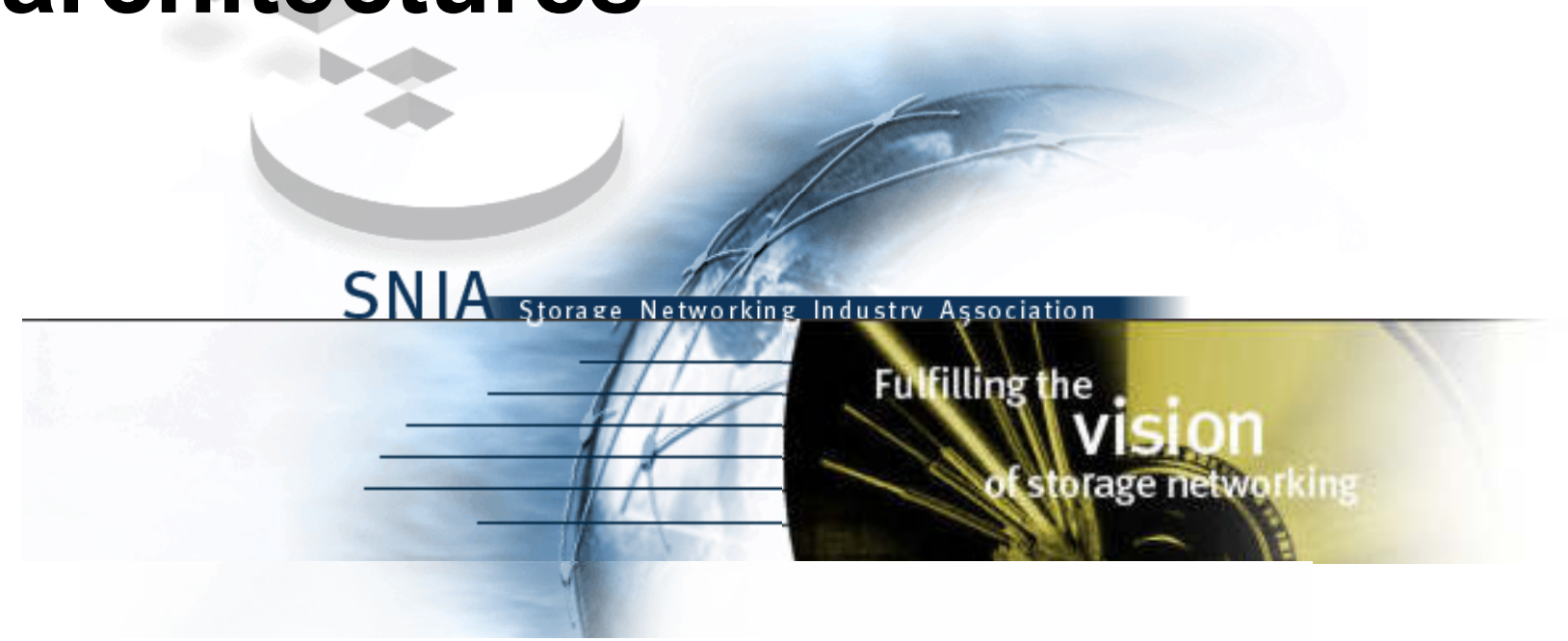
- Q: hardware: FibreChannel vs Ethernet vs InfiniBand?
- Q: API: blocks vs files (aka “NAS”) vs objects (OSD)?
- Q: protocol: FCP vs TCP/IP vs … ?
- A: (to all the above) *it depends* …



- **Storage network (SN):**
  - any (mostly) dedicated network, installed (mostly) for storage traffic
  - *whatever* the hardware, API, or protocol

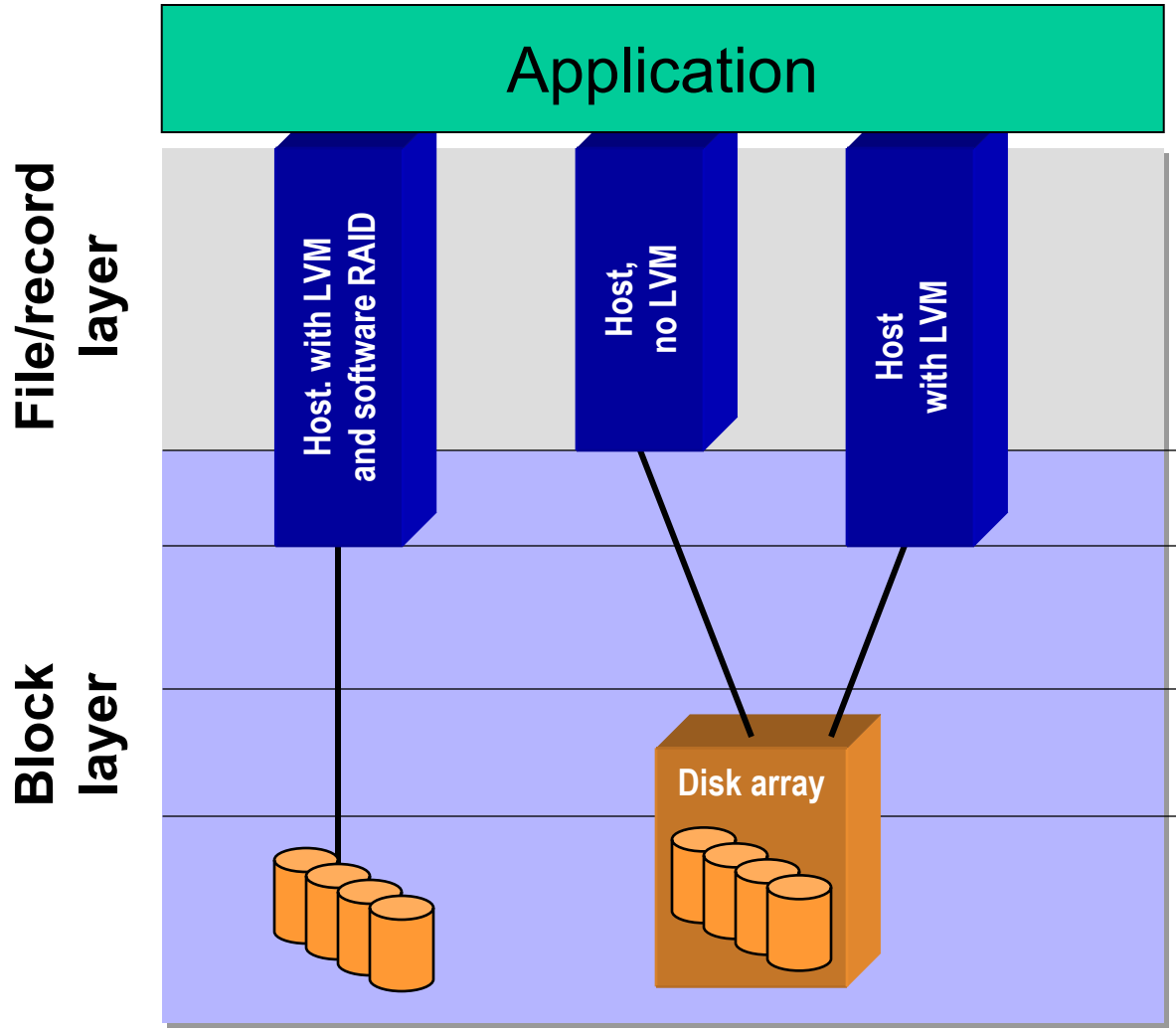
SNIA
Storage Networking Industry Association

Some common storage architectures



Mapping the SNIA model onto some current implementations

Direct-attach block storage



- **Direct-attach**
- **Multi-attach box**

E.g., SCSI

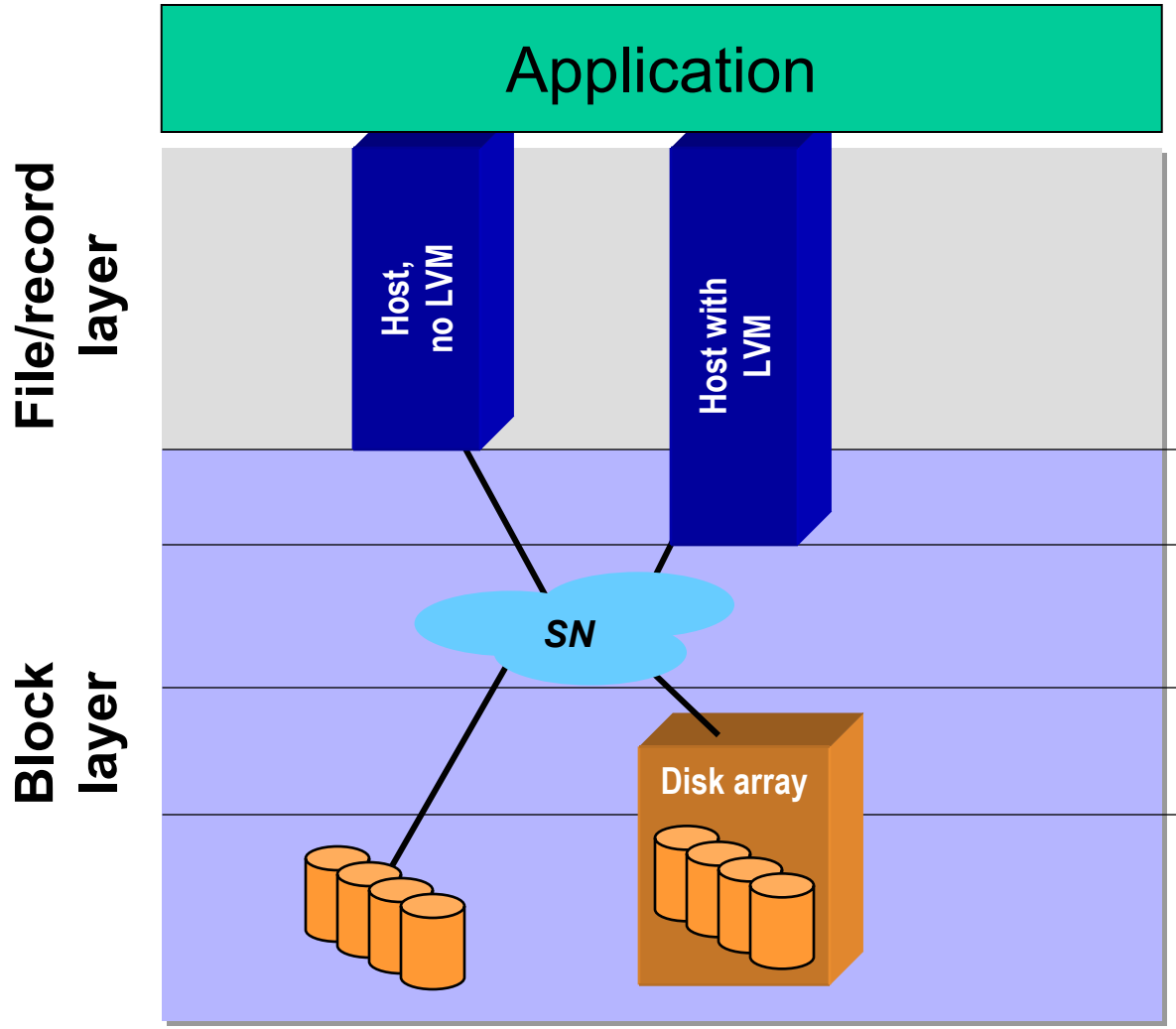
← *Host block-aggregation*

← *Network block-aggregation*

← *Device block-aggregation*

SNIA
Storage Networking Industry Association

SN-attached block storage



"SN" = any network used for storage access.
E.g., Fibre Channel, Ethernet, …

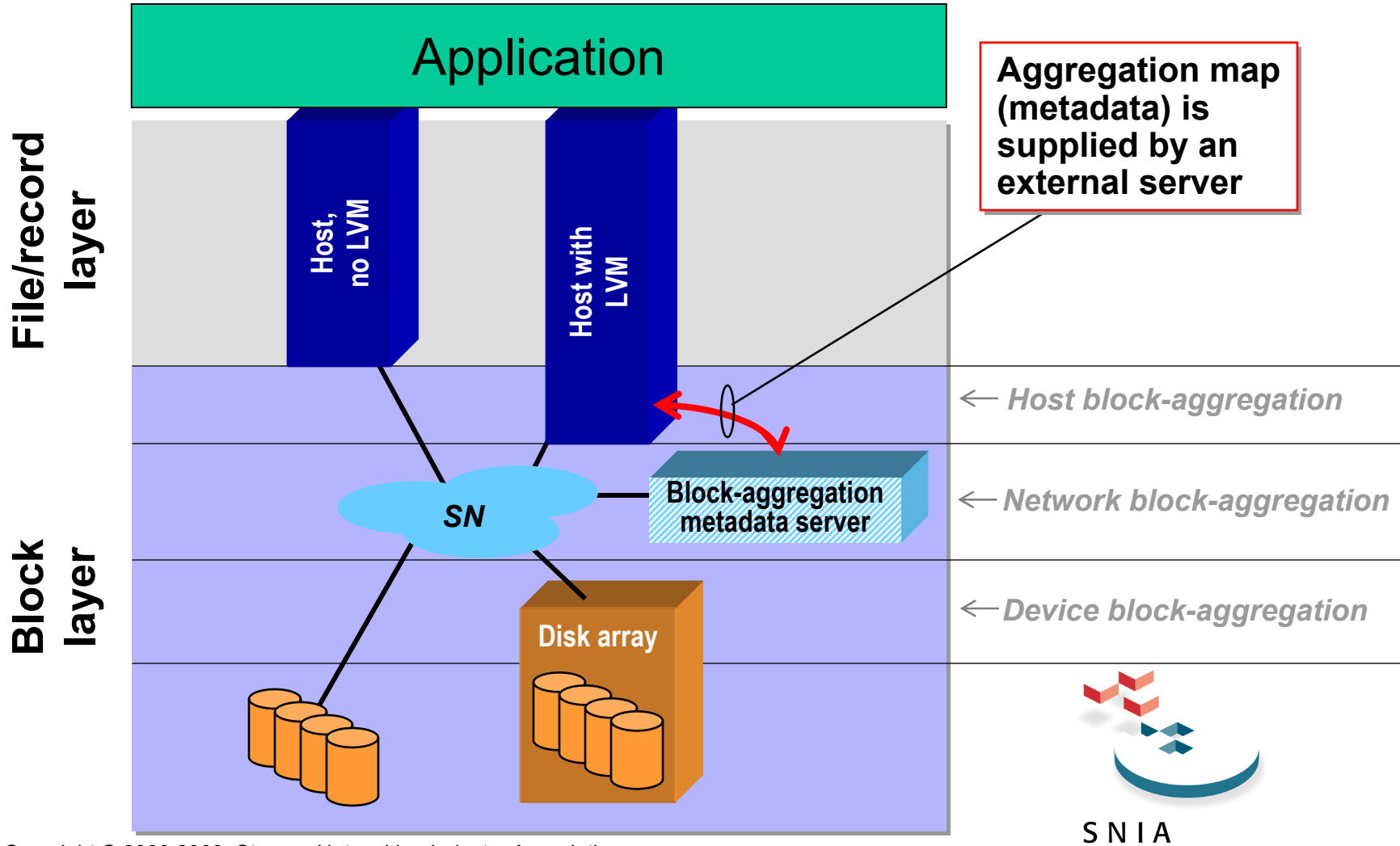
← *Host block-aggregation*

← *Network block-aggregation*

← *Device block-aggregation*

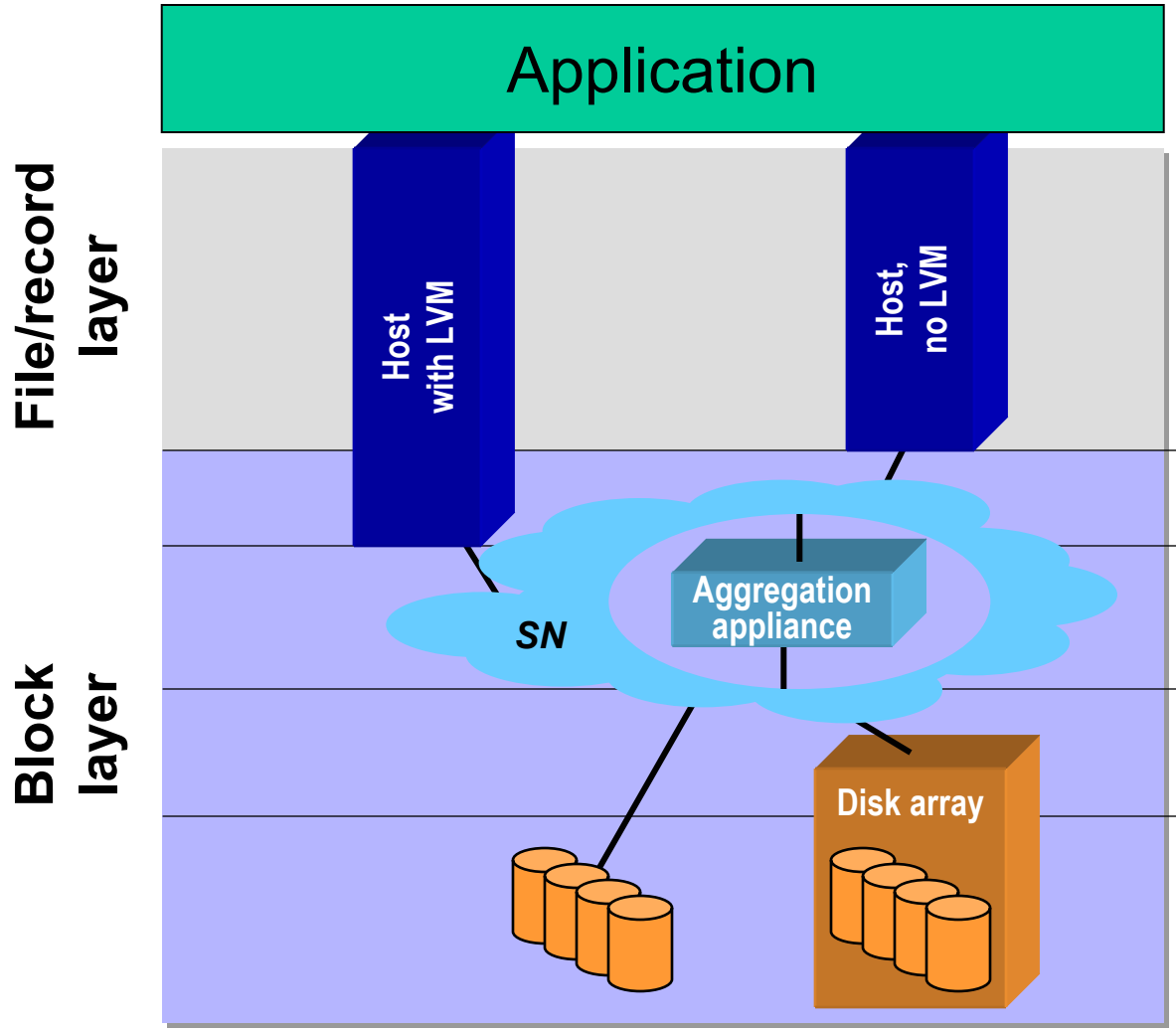
SNIA
Storage Networking Industry Association

SN-attached block storage
with metadata server



SNIA
Storage Networking Industry Association

Block storage aggregation
in a storage network appliance



aka "SN virtualization"

Functions: LVM, caching

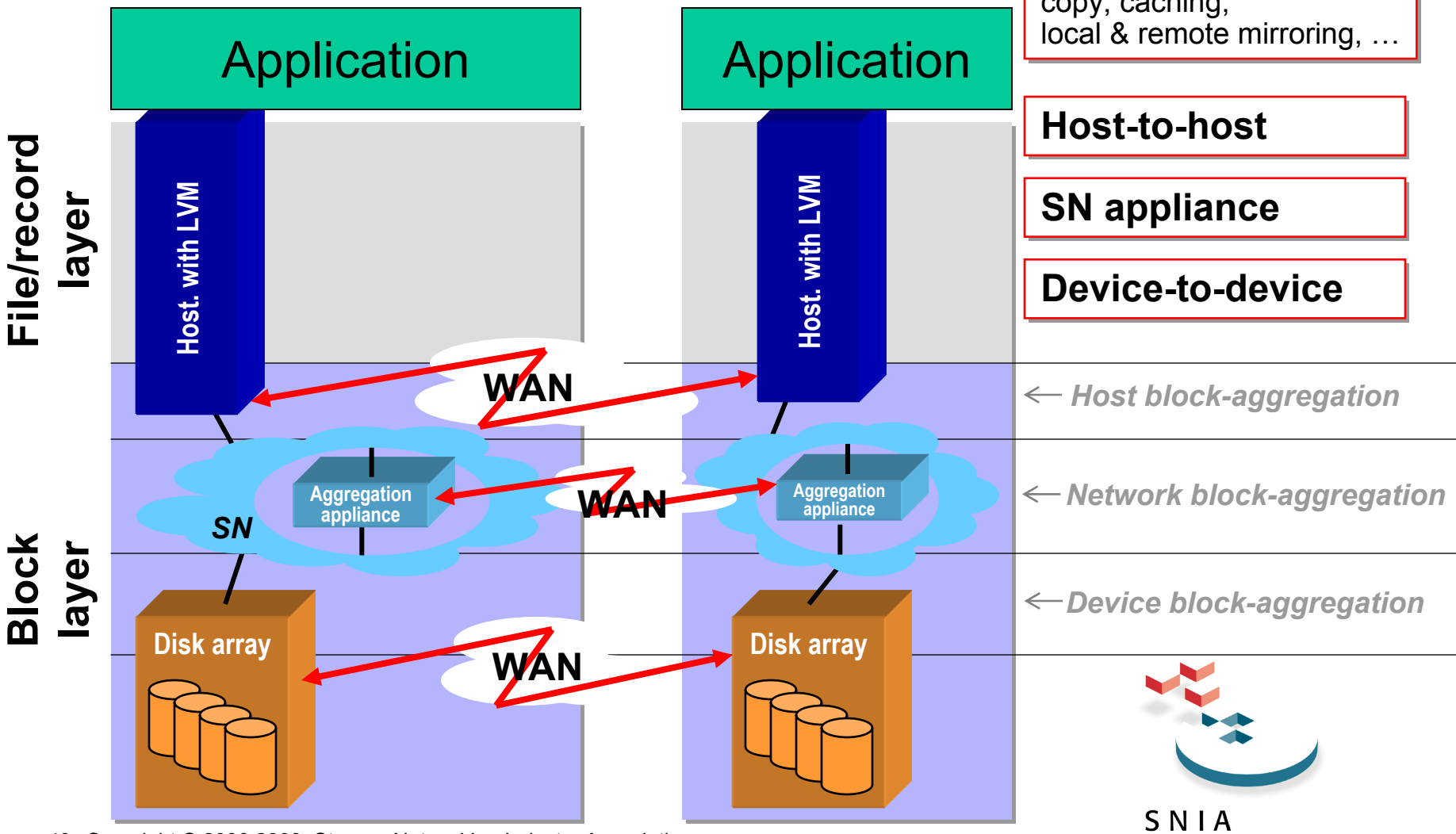
← *Host block-aggregation*

← *Network block-aggregation*

← *Device block-aggregation*

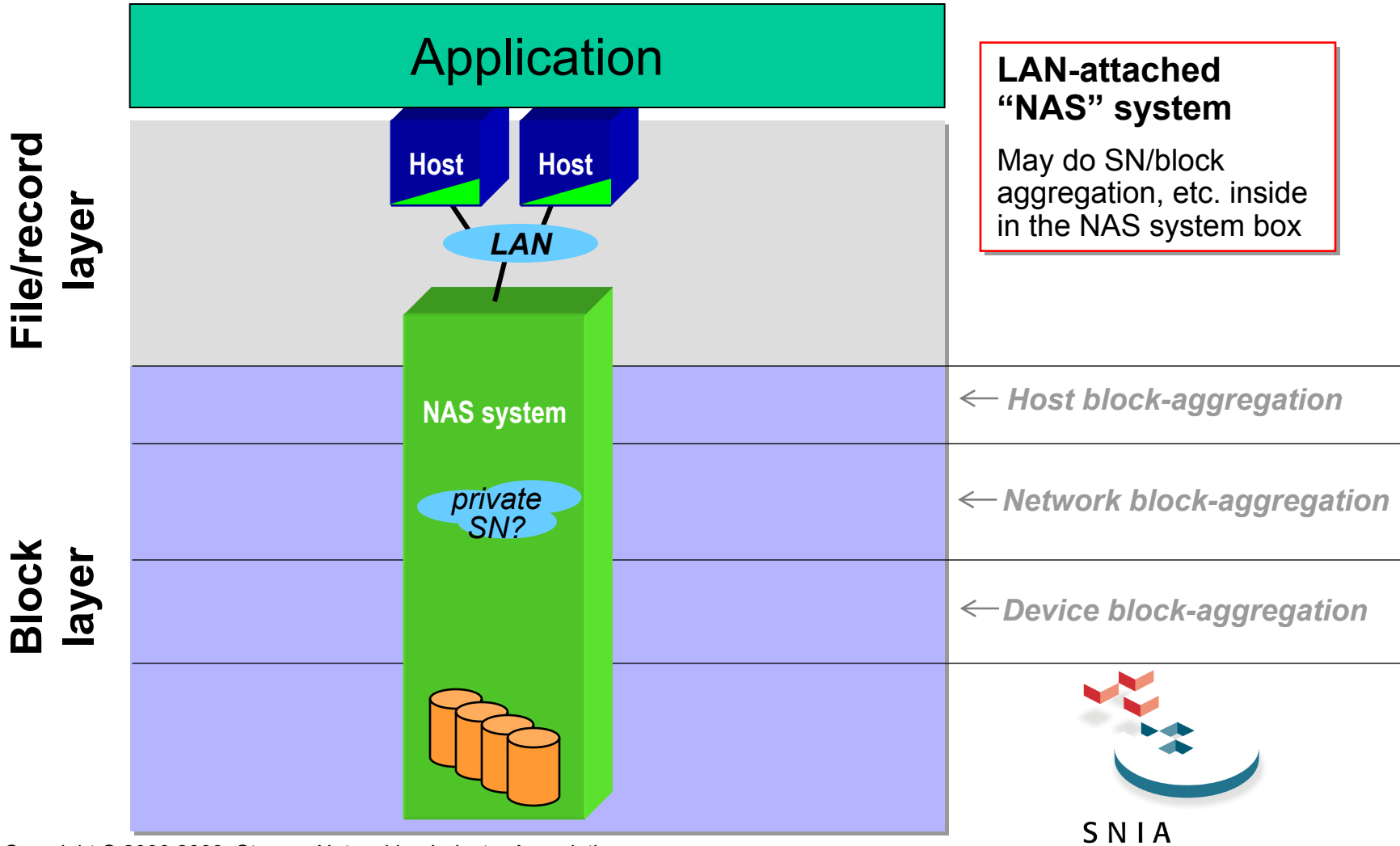
SNIA
Storage Networking Industry Association

Multi-site block storage



SNIA
Storage Networking Industry Association

File server

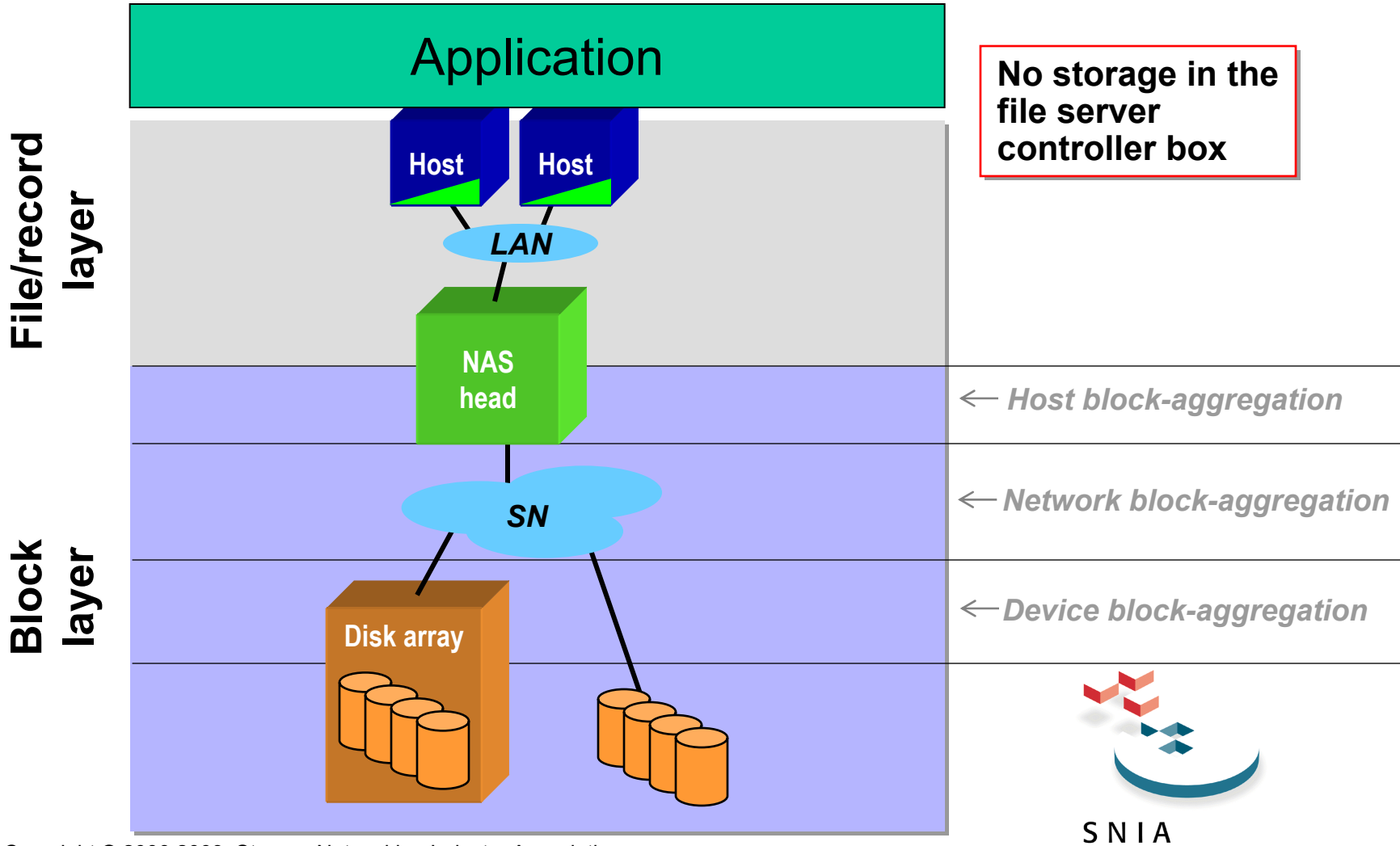


LAN-attached "NAS" system

May do SN/block aggregation, etc. inside in the NAS system box

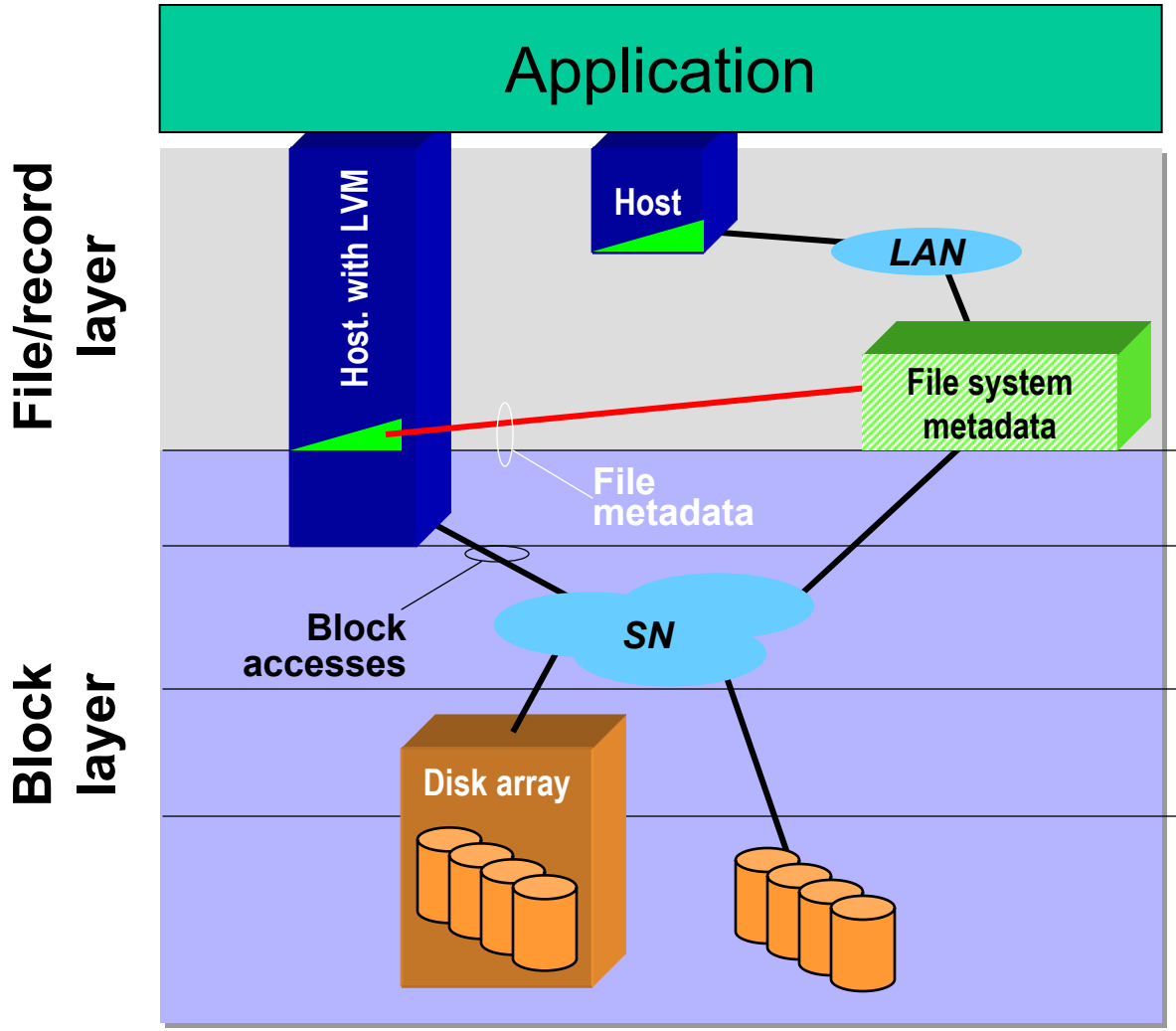
SNIA
Storage Networking Industry Association

File server controller ("NAS head")



SNIA
Storage Networking Industry Association

NAS/file server metadata manager ("asymmetric")



Hosts get file meta-data from FS/NAS controller, then access the data directly

FS controller can also be NAS server

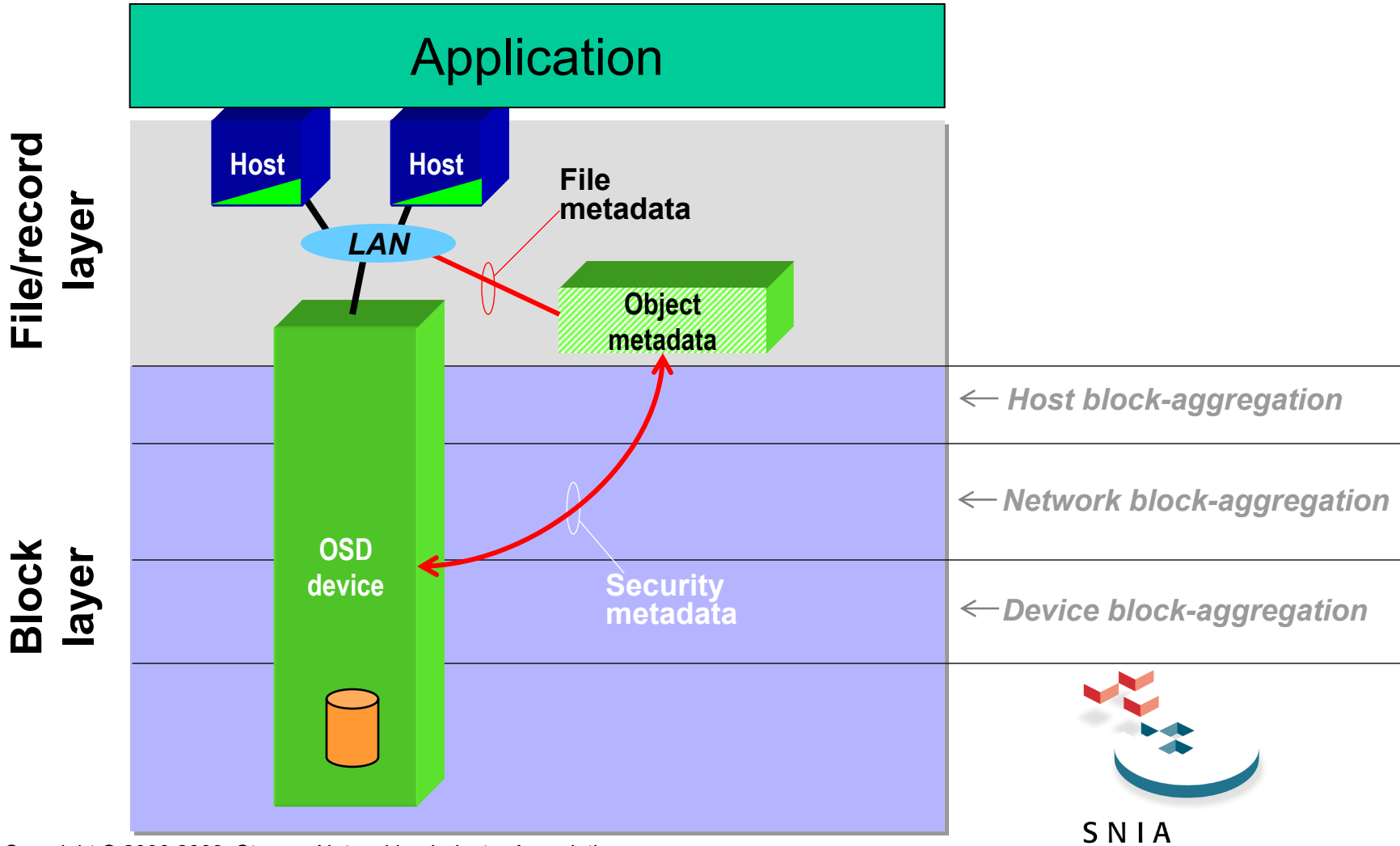
← *Host block-aggregation*

← *Network block-aggregation*

← *Device block-aggregation*

SNIA
Storage Networking Industry Association

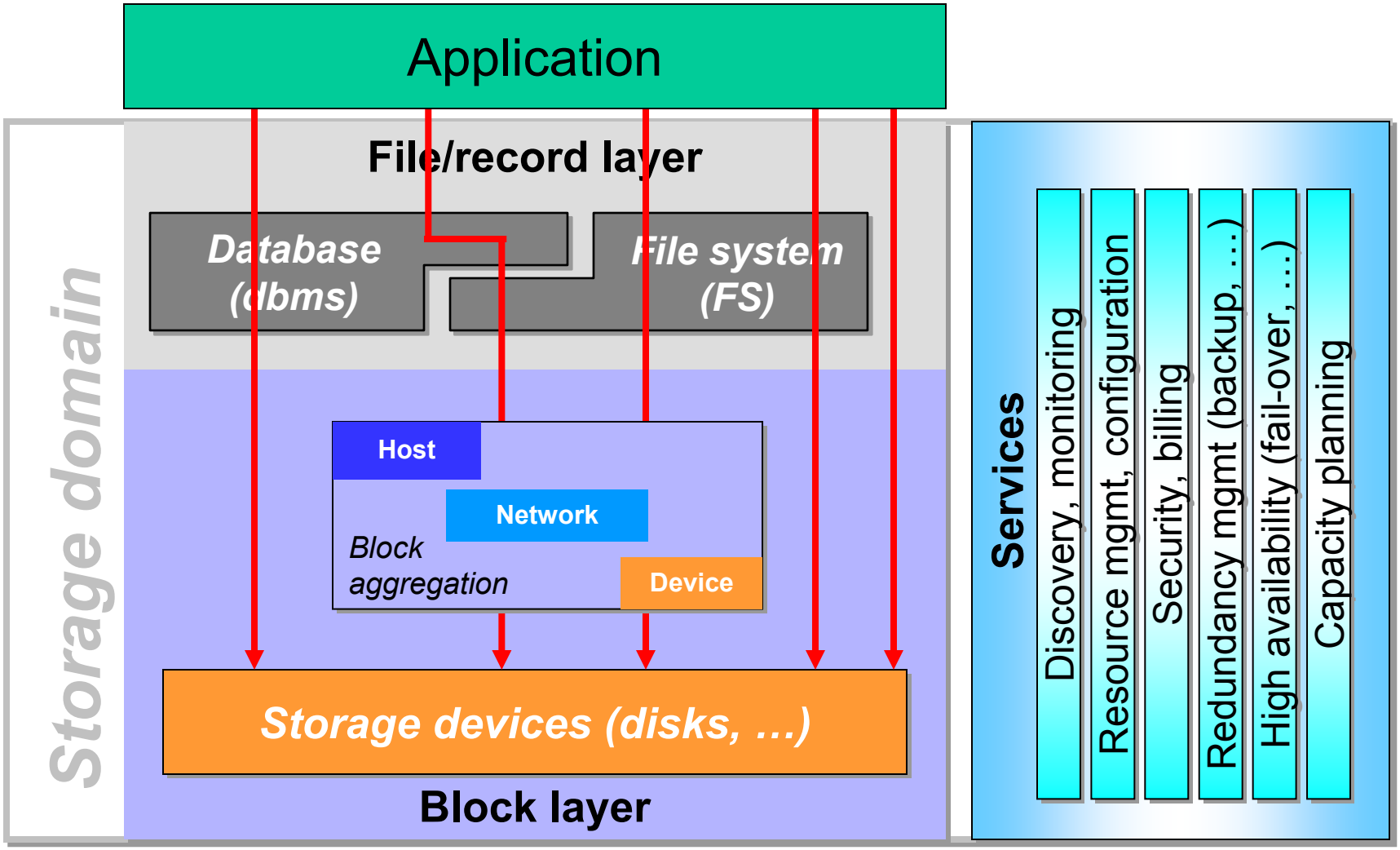
Object-based Storage Device (OSD), CMU NASD



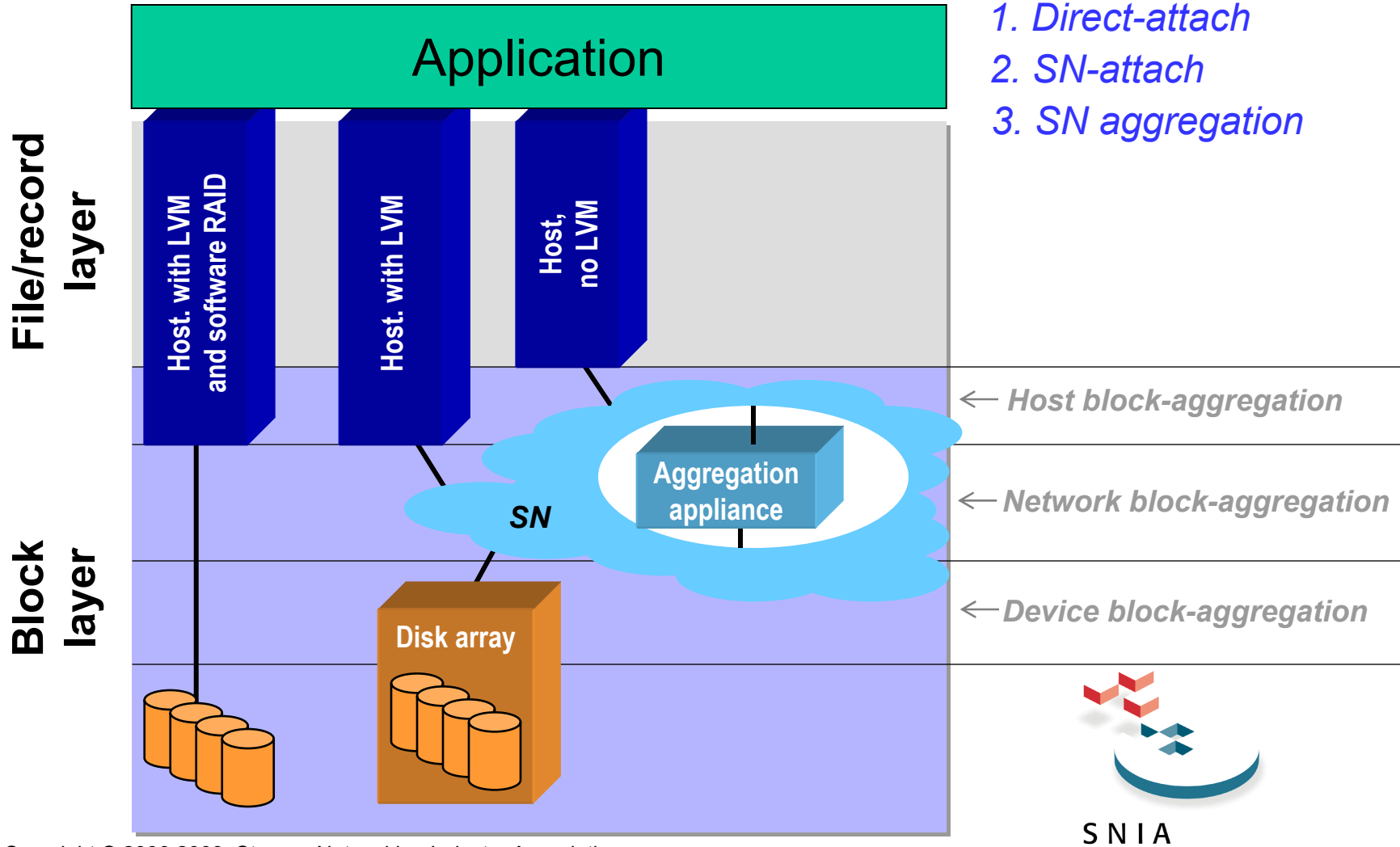
SNIA
Storage Networking Industry Association

Summary & conclusions

The SNIA shared storage model

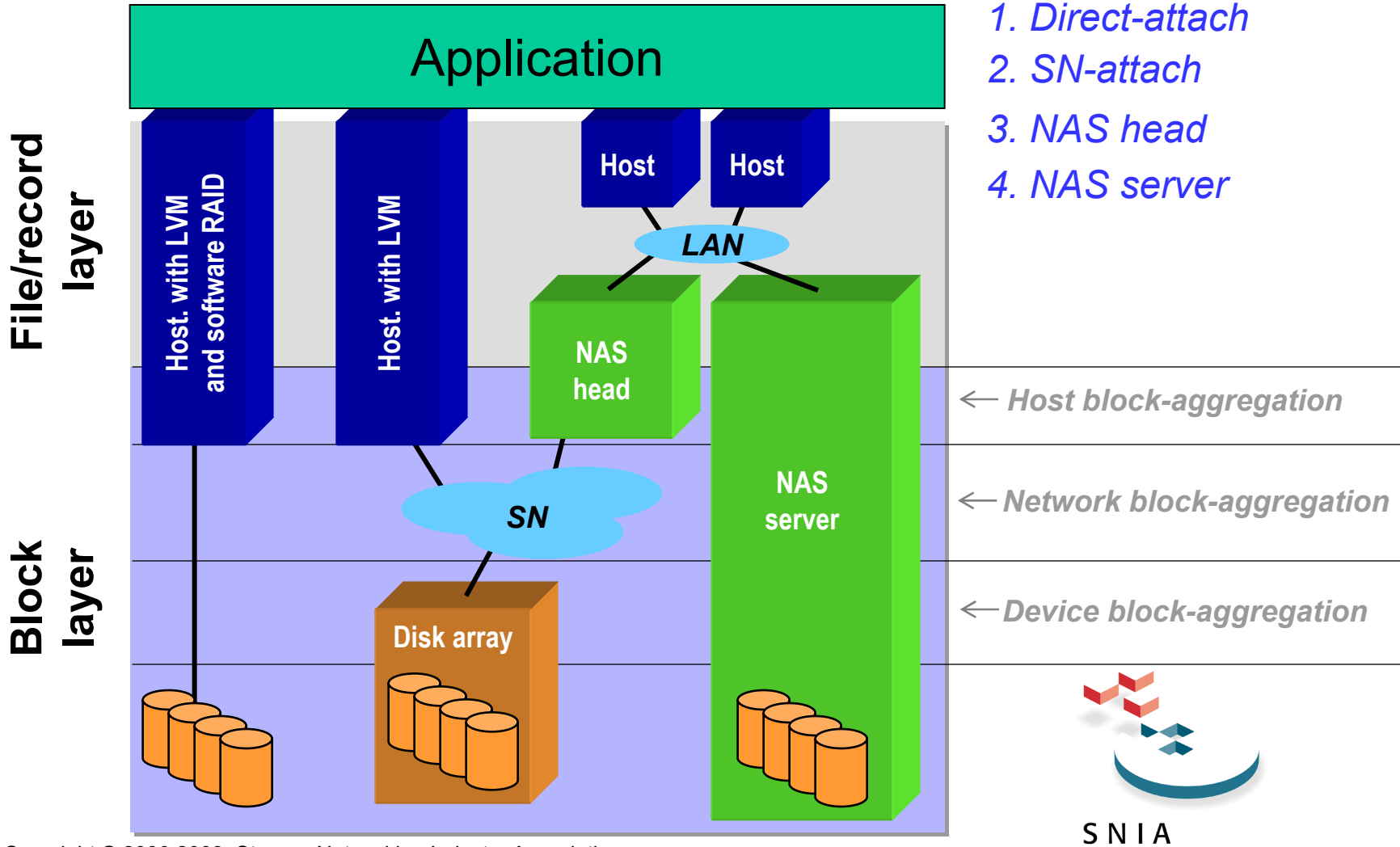


Block layer
Sample architectures



SNIA
Storage Networking Industry Association

File/record layer
Sample architectures



SNIA
Storage Networking Industry Association

Uses for the model

- **Vendors**
  - place products in the space of architectures
  - clarify product differences
- **Customers**
  - understand vendor offerings better
- **The industry**
  - basis for common definitions, communication, understanding, interoperability

SNIA
Storage Networking Industry Association

Conclusions

- **The SNIA shared storage model is both simple and useful**
  - to highlight similarities and differences
  - as a basis for comparisons
- **Still a work in progress**
  - data movers, tape drives, …
  - better comparisons …
  - suggestions?
- **The SNIA-TC welcomes input:**
  - <snia-tc@snia.org>

SNIA
Storage Networking Industry Association

Authors

- **John Wilkes, Hewlett-Packard (project lead)**
- **Harald Skardal, NetApps**
- **David Black, EMC**
- **Wayne Rickard (SNIA TC chairperson), Gadzoox**
- **Co-conspirators: the rest of the SNIA Technical Council**
  - Dave Anderson, Seagate Technology
  - Jim Carlson, IBM
  - Garth Gibson, CMU/Panasas
  - Kevin Phaup, HighGround Systems
  - David Thiel, Compaq

SNIA
Storage Networking Industry Association

