Education

Storage Networking Standards: Recent Developments

David L. Black, EMC Corporation
Member, SNIA Technical Council

SNIA Legal Notice

- The material contained in this tutorial is copyrighted by the SNIA.
- Member companies and individuals may use this material in presentations and literature under the following conditions:
  - Any slide or slides used must be reproduced without modification
  - The SNIA must be acknowledged as source of any material used in the body of any document containing material from these presentations.
- This presentation is a project of the SNIA Education Committee and Technical Council

Abstract

Storage Networking Standards: Recent Developments

Interoperability standards play a vital role in customer and advancement of storage networking technologies and systems. Storage networking is based on a broad spectrum of standards (developed by multiple standards organizations) in areas such as Fibre Channel (INCITS T11), SCSI (INCITS T10), iSCSI (IETF), and storage management (SNIA, IETF). The current state and future direction of standards development can provide useful insights into technology developments. This tutorial covers storage networking standards and the role that the resulting standardized interfaces and functionality play in networked storage infrastructure. The tutorial presenter is a member of the SNIA Technical Council who is actively involved in development of many storage networking standards.

About the Author

David L. Black

Senior Technologist, EMC Corporation

Member, SNIA Technical Council

Chair, T11 FC-SP-2 (Fibre Channel Security-2) Working Group

Chair, IETF IP Storage (IPS) Working Group

Chair, IETF Internet and Management Support for Storage (IMSS) Working Group

Co-Chair, SNIA Fixed Content Aware Storage (FCAS) Technical Working Group

Overview

- Interoperability standards for storage
  - Connect products from different vendors
- Standards can provide technology insight
  - Emerging technology
  - Evolution of existing technology
- This talk: Standards developments and directions
  - Implications for technology vendors and users

Standards: End User Benefits

- Protect technology investment
- Ensure a base level of interoperability
- Provide choice among products
- Ensure continuing innovation
- Commonality leads to less training, simpler deployment

Storage Networking Standards

Data Protocols

- SCSI (T10)
- Fibre Channel (T11)
- ATA and SATA (T13)
- NAS (IETF, Microsoft)
- IP Storage (IETF)

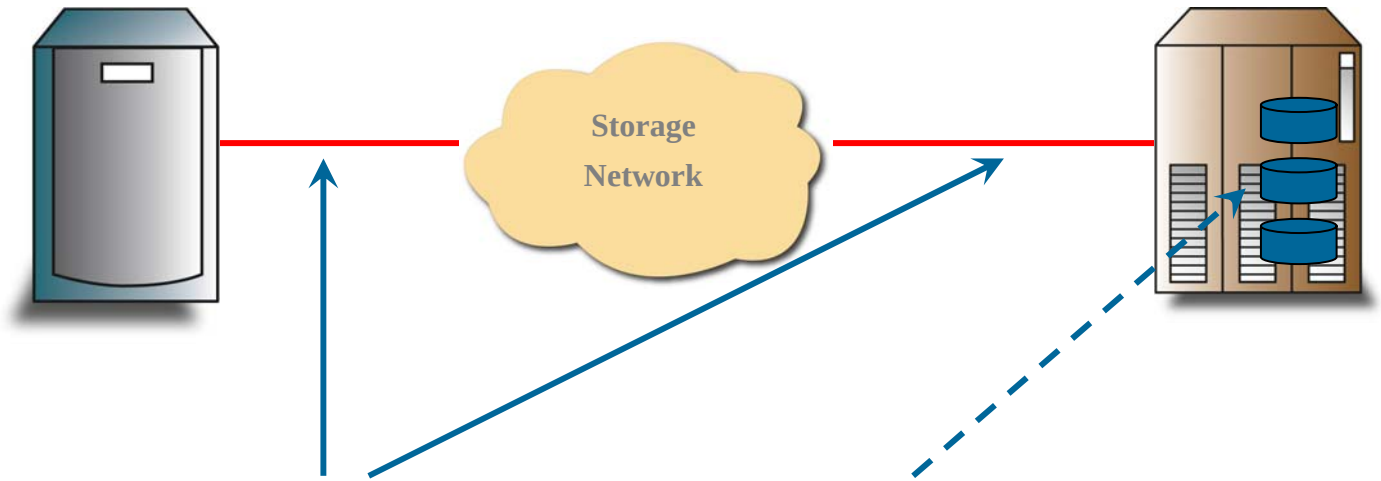
Management

- SMI-S (SNIA)
  - Uses CIM (DMTF)
- SNMP (IETF)
- Web (IETF, W3C)

Stored Data

- RAID Layout (SNIA)
- Encryption (IEEE)
- Fixed Content (SNIA)

Storage Protocol Classes



Network Communication

- Fibre Channel Fabric
- IP Storage (iSCSI, FCIP, iFCP)
- Network Attached Storage (NAS)

Drive Interface

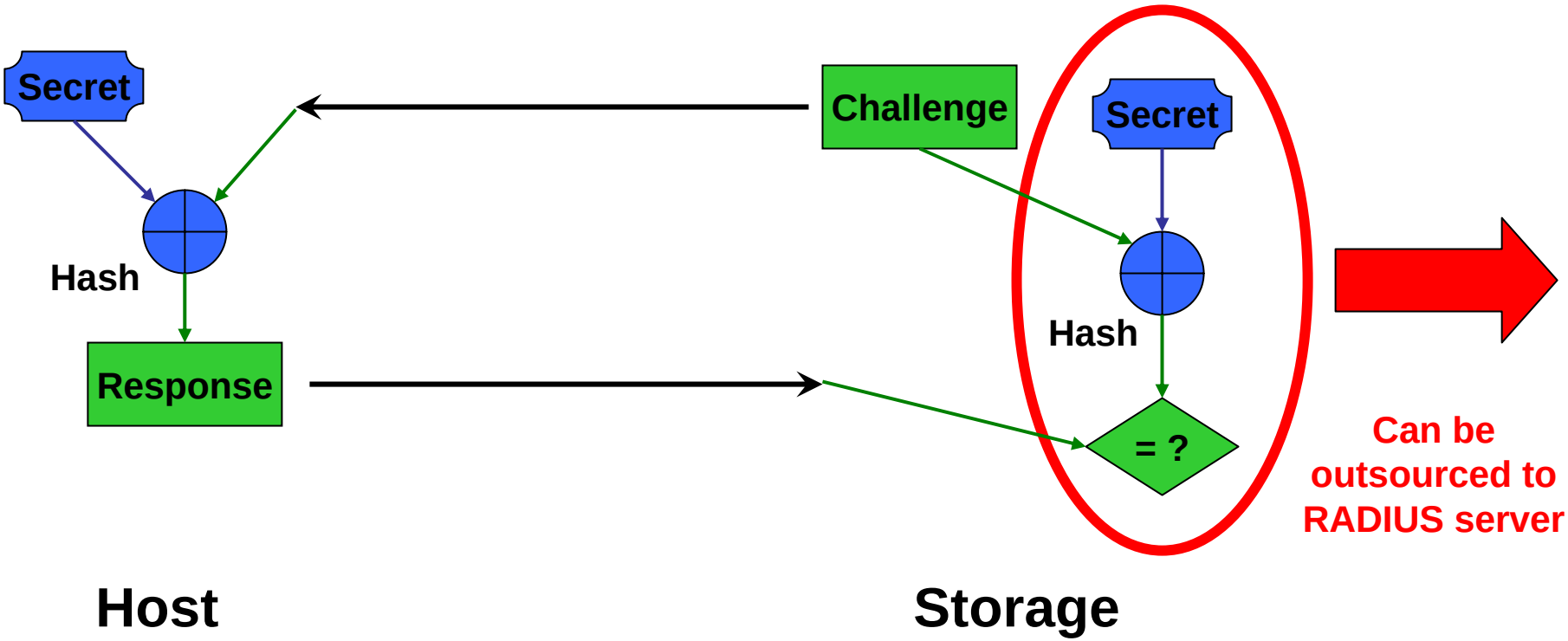
- FC-AL (Arbitrated Loop)
- Parallel SCSI and Serial Attach SCSI (SAS)
- ATA and Serial ATA (SATA)

Fibre Channel 1: Security (T11 FC-SP)

- Fibre Channel (FC) Fabric access and config controls
  - Control fabric structure and what can join the fabric
- In-band Authentication
  - Secret (CHAP, DH-CHAP)
    - CHAP = Challenge Handshake Authentication Protocol
    - DH-CHAP = Diffie-Hellmann CHAP (adds a DH exchange)
  - Public Key (FCAP) and Password (FCPAP)
    - FCAP = Fibre Channel Authentication Protocol
    - FCPAP = Fibre Channel Password Authentication Protocol
- Secure communication channels
  - Adaptation of IPsec subset to Fibre Channel

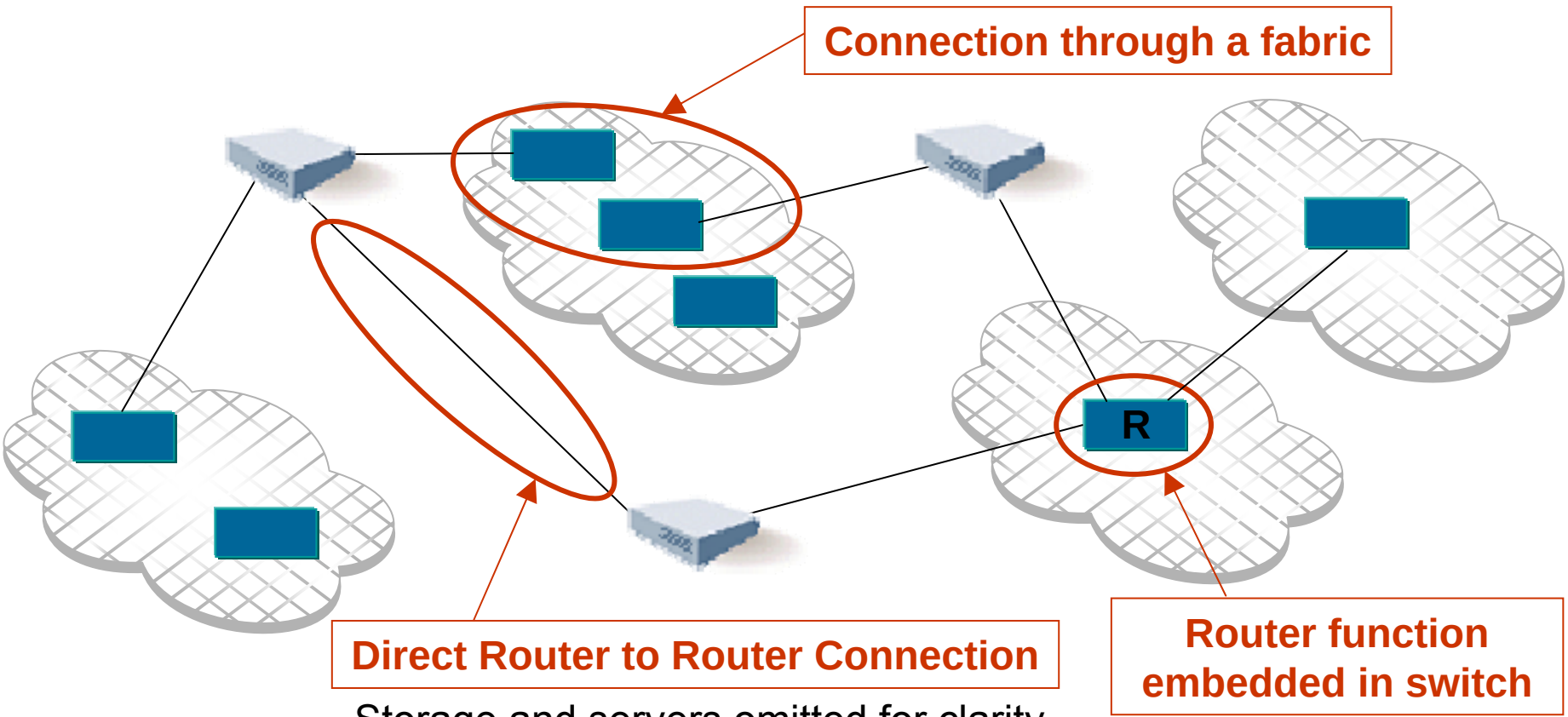
CHAP Authentication Protocol

- Authenticate with a shared secret, random challenge
  - Based on a secure hash, usually MD5



Fibre Channel 2: Inter-Fabric Routing

- FC Routers inter-connect FC Storage Area Networks (SANs)
  - Particularly useful for isolated SANs (SAN islands)



Storage and servers omitted for clarity

FC Inter-Fabric Routing Properties

- Routing interconnects physical ***and*** virtual fabrics
  - Virtual fabrics can be in the same or different physical SAN
  - Routes can pass through existing fabrics and switches
- The interconnected fabrics ***do not merge***
  - Prevents some disruption propagation (e.g., RSCN)
    - RSCN = Registered State Change Notification
  - Translation of FC addresses required (unlike IP routing)
- Routing is transparent to servers and storage
  - Zoning, name service, etc. continue to work
- Routing function packaging:
  - In a separate router or combined with a fabric switch

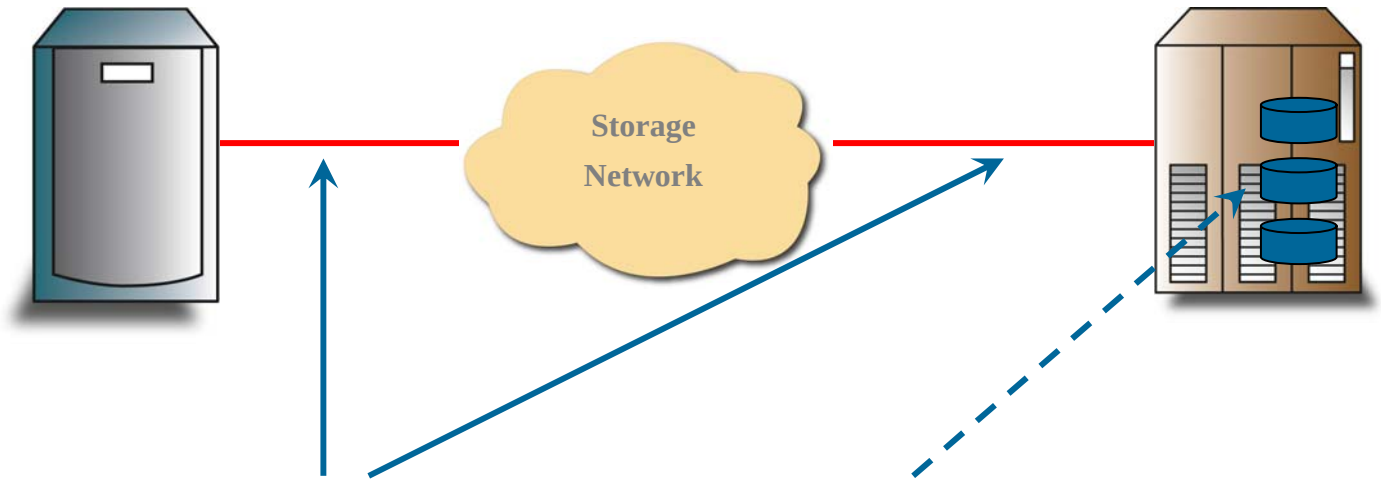
Fibre Channel 3: Communication Media

- 4 Gbit/sec Fibre Channel speed
  - Compatible upgrade, shorter distance limits
  - Next speed upgrade will be to 8 Gbit/sec
    - Limited deployment of 10 Gbit/sec FC
- OM3 multimode optical fiber (50µ)
  - 2 Gbit/sec FC reach: 300m on OM2 fiber
  - 4 Gbit/sec FC reach: 150m on OM2 fiber, 300m+ on OM3 fiber
- FC Base-T: Fibre Channel over twisted pair
  - Twisted Pair: Category (Cat) 5e, 6 and 6a cable
    - Same cables as Gigabit and 10 Gigabit Ethernet
  - 1, 2, 4 Gbit/sec speeds – 100m reach on Category 6a cable
    - 4 Gbit/sec: 40m reach on Cat 6, Cat 5e not recommended

Fibre Channel 4: Protocol Topics

- N_Port (usually server) identity virtualization (NPIV)
  - Logical N_Port per virtual server supports server virtualization
  - Replaces questionable use of multiple FC-AL identifiers
- FCoE – Fibre Channel over Ethernet
  - Motivation: Consolidated server I/O over Ethernet
    - Server racks (e.g., blades) – cost, space, power considerations
  - Encapsulate Fibre Channel frames in Ethernet frames
    - Requires at least baby jumbo (2.5k) Ethernet frames
  - Anticipate Ethernet enhancements to eliminate drops
    - Back-propagation of Ethernet Pause across switches
    - Per-priority Pause (separate priorities for FC vs.IP traffic)
  - FCoE multipathing – will be based on VRRP
    - Virtual Router Redundancy Protocol
    - See IETF RFC 3768

Storage Protocol Classes



Network Communication

- Fibre Channel Fabric
- IP Storage (iSCSI, FCIP, iFCP)
- Network Attached Storage (NAS)

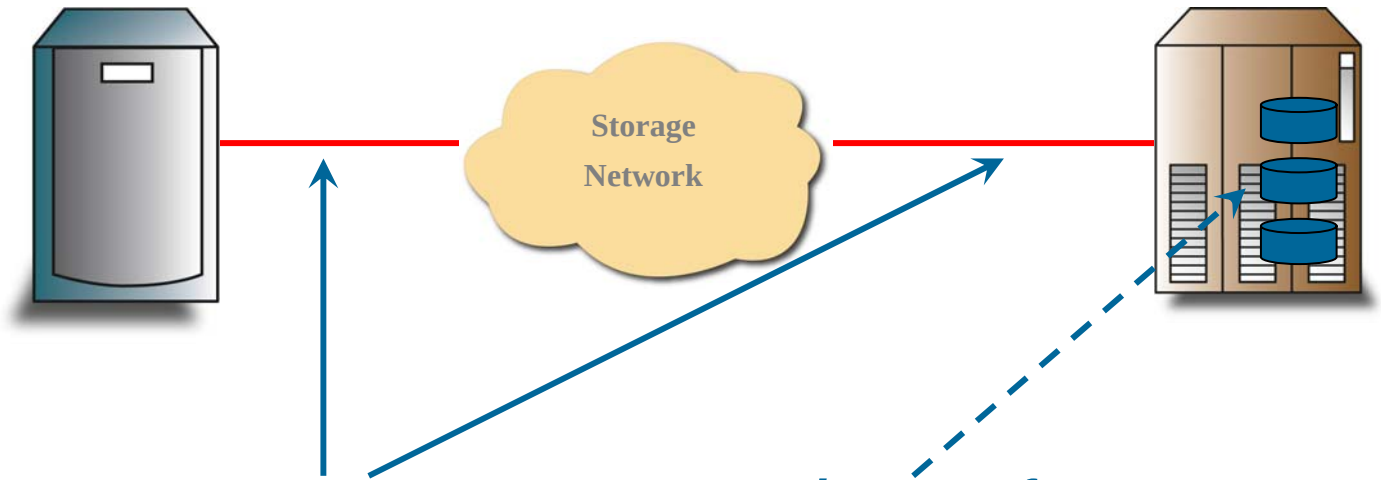
Drive Interface

- FC-AL (Arbitrated Loop)
- Parallel SCSI and Serial Attach SCSI (SAS)
- ATA and Serial ATA (SATA)

IP Storage Developments

- RFC specifications have published
  - iSCSI (RFC 3720), FCIP (RFC 3821), iFCP (RFC 4172)
- iSCSI Corrections and Clarifications draft approved
  - Minor corrections and clarifications
- RDMA for iSCSI (iSER)
  - RDMA = Remote DMA over a TCP/IP network (iWARP)
  - iSER = iSCSI Extensions for RDMA
    - InfiniBand: iSER is an alternative to SRP for storage gateways to FC
    - SRP = SCSI RDMA Protocol
- Fibre Channel Pseudo-Wire over MPLS
  - MPLS: Multi-Protocol Label Switching (carrier infrastructure)
  - Being jointly developed by IETF and T11

Storage Protocol Classes



Network Communication

- Fibre Channel Fabric
- IP Storage (iSCSI, FCIP, iFCP)
- Network Attached Storage (NAS)

Drive Interface

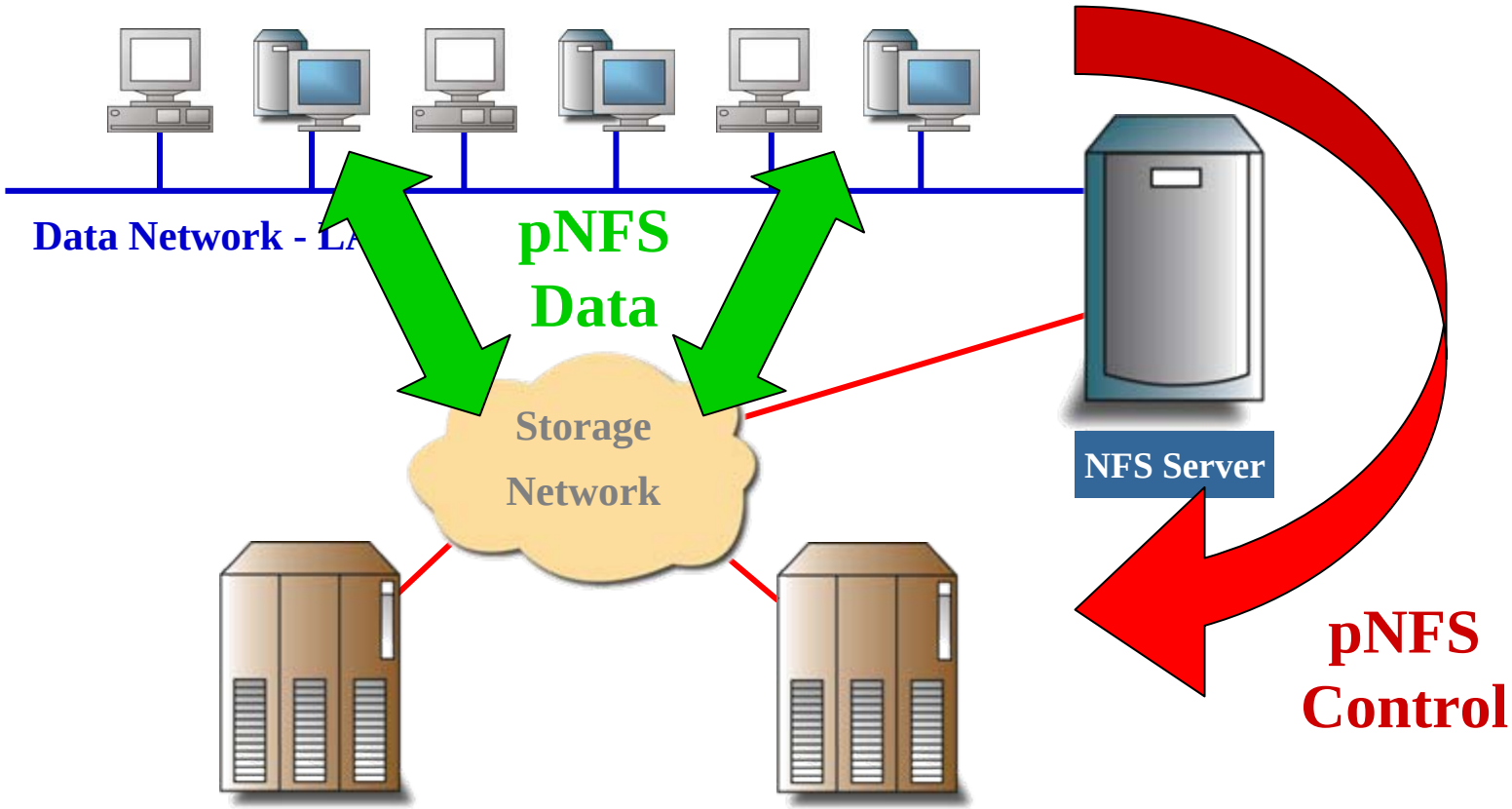
- FC-AL (Arbitrated Loop)
- Parallel SCSI and Serial Attach SCSI (SAS)
- ATA and Serial ATA (SATA)

Network Attached Storage (NAS)

- Two primary file serving protocols:

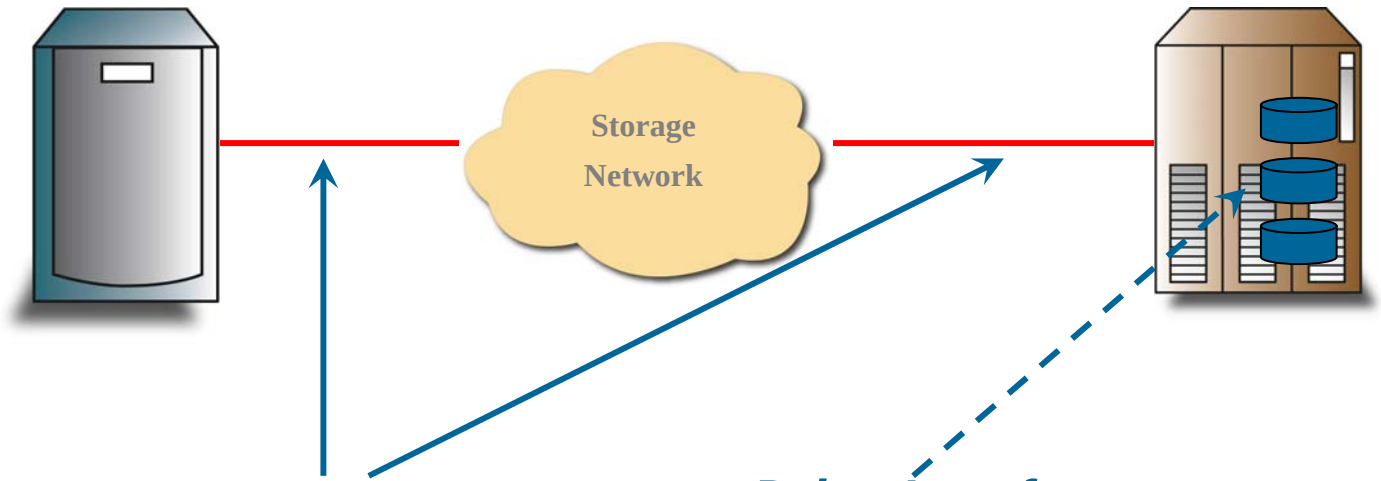
1. NFS, primarily for Unix clients (IETF)
   - Transition to NFSv4 is underway
   - Parallel NFS (pNFS): SAN filesystem support
     - Parallel (e.g. striped) access across NFS servers
   - RDMA (iWARP) for NFS
   - Draft of NFS 4.1 specification:
     - draft-ietf-nfsv4-minorversion1-12.txt
     - Sessions, directory delegations and pNFS
2. CIFS, primarily for Windows clients (Microsoft)
   - Please ask Microsoft

Parallel NFS - pNFS



- NFS file naming, management, and administration
- Parallel high bandwidth file access (via Storage Network)

Storage Protocol Classes



Network Communication

- Fibre Channel Fabric
- IP Storage (iSCSI, FCIP, iFCP)
- Network Attached Storage (NAS)

Drive Interface

- FC-AL (Arbitrated Loop)
- Parallel SCSI and Serial Attach SCSI (SAS)
- ATA and Serial ATA (SATA)

Drive Interface Developments

- ATA drives: Serial ATA (SATA)
  - Replacing Parallel ATA
- SCSI drives: Serial Attached SCSI (SAS)
  - Replacing Parallel SCSI
  - Connection-based protocol (not packet switched)
  - SAS Can carry SATA traffic, attach to SATA drives
  - SAS-2 work in progress, includes:
    - SAS Zoning will allow shared SAS infrastructure and storage
    - 6 Gbit/sec drive interface speed
- Fibre Channel drives: FC-AL
  - 4 Gbit/sec drives available
  - FATA/FC-LC: ATA-class disk with FC-AL interface
  - FC-SATA standard will enable FC to attach to SATA drives

Storage Networking Standards

Data Protocols

- SCSI (T10)
- Fibre Channel (T11)
- ATA and SATA (T13)
- NAS (IETF, Microsoft)
- IP Storage (IETF)

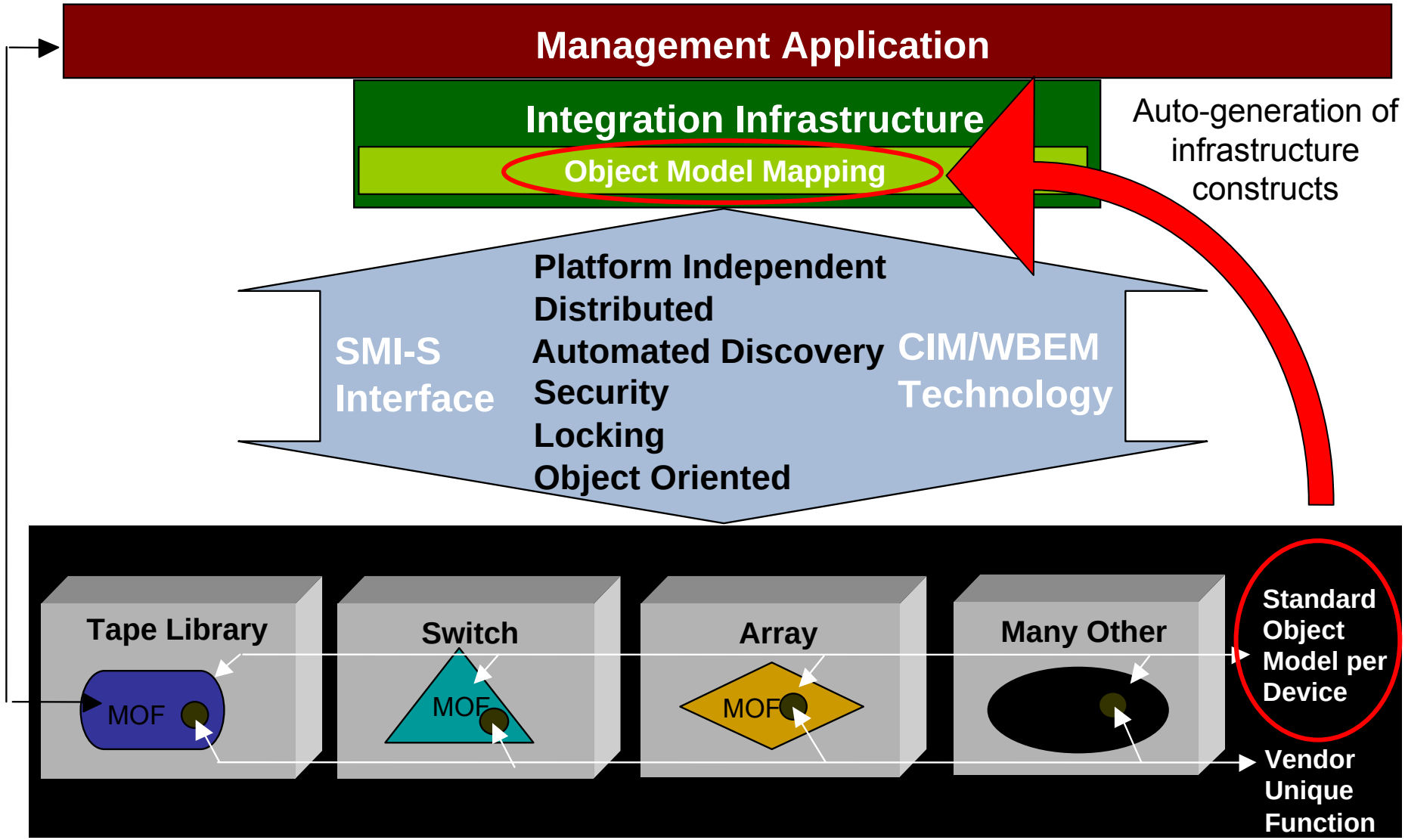
Management

- SMI-S (SNIA)
  - Uses CIM (DMTF)
- SNMP (IETF)
- Web (IETF, W3C)

Stored Data

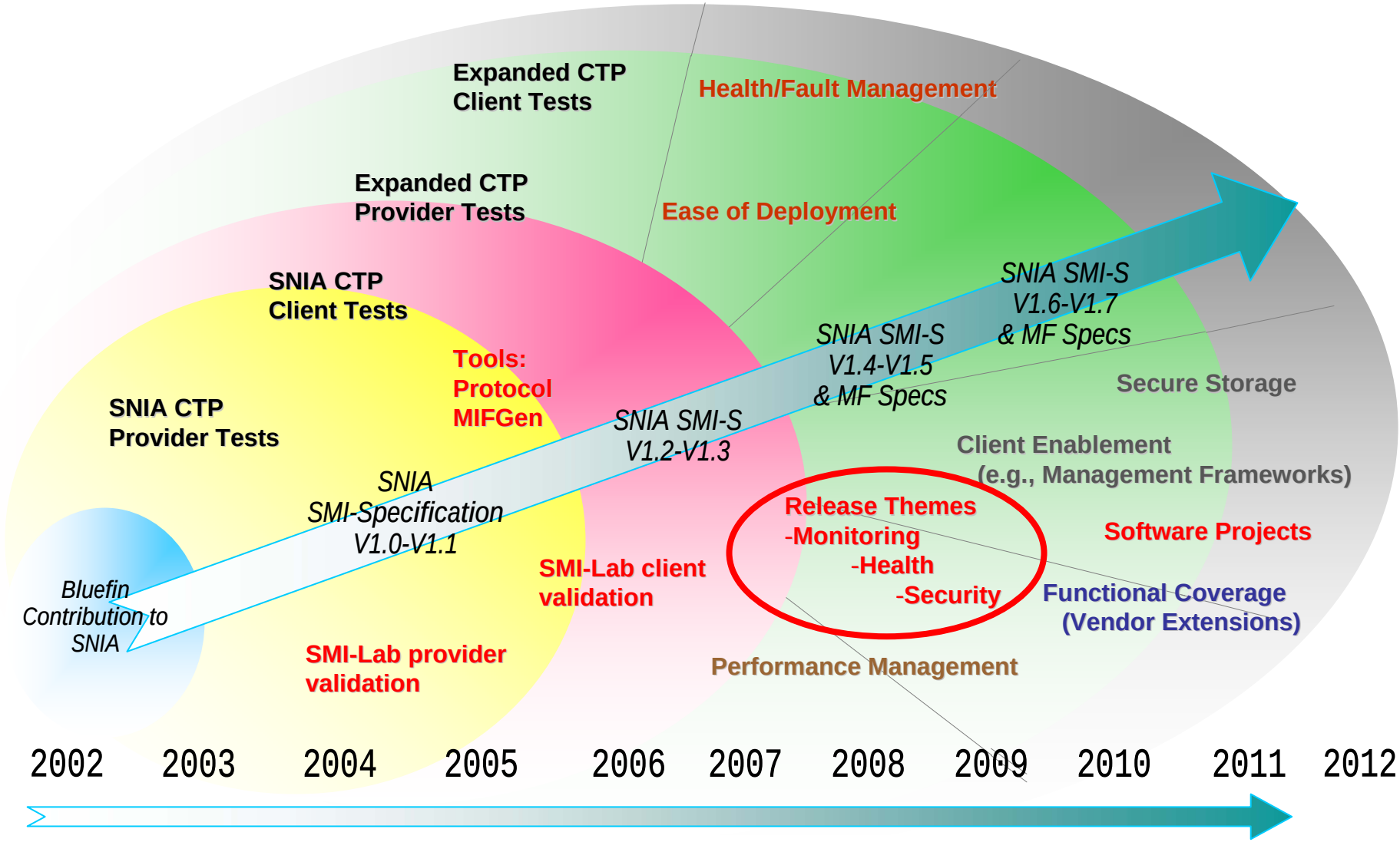
- RAID Layout (SNIA)
- Encryption (IEEE)
- Fixed Content (SNIA)

SMI-S: Model-Based Management



Storage Management Initiative

Ecosystem Road Map – The journey continues



SMI-S Protocol Developments

- Web Services (OASIS): increasing interest
  - Two management stacks (WSDM, WS-Man)
    - Convergence whitepaper has been published
  - WS-Man support for SMI-S: version 1.4 or later
- IPv6 deployment
  - US government IPv6 requirements profiles:
    - DISA = Defense Information Systems Agency
    - NIST = National Institute of Standards and Technology
  - IPv4 to IPv6 transition: Mostly transparent to SMI-S
- CIM over SCSI: SCSI commands that carry SMI-S
  - Enables inband SMI-S management of storage arrays
    - Management shares port(s) used for storage access
    - SMI-S over FC, iSCSI, SAS, etc.

Storage Networking Standards

Data Protocols

- SCSI (T10)
- Fibre Channel (T11)
- ATA and SATA (T13)
- NAS (IETF, Microsoft)
- IP Storage (IETF)

Management

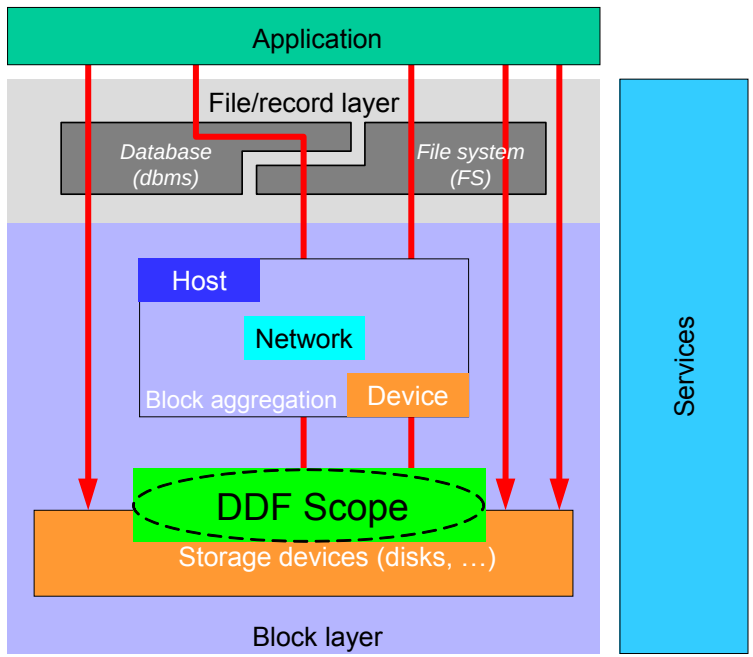
- SMI-S (SNIA)
  - Uses CIM (DMTF)
- SNMP (IETF)
- Web (IETF, W3C)

Stored Data

- RAID Layout (SNIA)
- Encryption (IEEE)
- Fixed Content (SNIA)

SNIA Disk Data Format (DDF)

- DDF - Data structures describing how data is distributed across the drives in a RAID implementation.
- Primary intended scope: RAID controllers for internal and direct attach storage.
- Does not standardize operating system/RAID controller interface or create a single driver



SNIA Shared Storage Model

IEEE P1619 : Encrypt Stored Data at Rest

- Threat: Move encrypted blocks
  - Attacker may know location of data, even if it's encrypted
- Disks: No visible room for additional integrity checks
- "Tweaked" Encryption modes prevent block swapping
  - Two keys used: encryption key and tweak key
  - If encrypted blocks moved: Decryption produces gibberish
- P1619 disk encryption: XEX tweak
  - XTS-AES mode: XTS = XEX Tweak + ciphertext Stealing
  - 128-bit and 256-bit key sizes (2 keys needed)

IEEE P1619.1: Tape Encryption

- Tapes have room for additional integrity checks
- Combined modes: 256-bit AES key
  - Combined = encryption + cryptographic integrity
  - AES-GCM (Galois Counter Mode) – hardware friendly
  - AES-CCM (Counter with CBC-MAC) – simpler
- Other modes: 256-bit AES key + integrity key(s)
  - AES-CBC with HMAC-SHA-(1, 256, or 512)
  - XTS-AES with HMAC-SHA-512

New IEEE P1619 efforts

- P1619.2: Wide Block encryption for disks
  - Write a disk block, make a small change, rewrite the block
  - Attacker compares old and new ciphertext
    - Change may be as small as 16 bytes (for 128-bit AES)
    - Wide block AES modes: Change is at least 512 bytes
- P1619.3: Key Management for protecting stored data
  - Very important: Loss of encryption key **IS** Loss of data
  - Initial goal: keys for use with P1619-defined encryption
  - Scope: All stored data, not just use of P1619 encryption
    - Includes storage, management and distribution of keys
- P1619.2 and P1619.3 – In early stages of work efforts

XAM API for Fixed Content

- Purpose: Fixed Content Storage Access
  - Store content (data) that does not change
  - Independent of location of storage system or data
    - Motivated by migration to new systems and technology
- Vendor independent API + file system interface (FSI)
  - FSI: Present fixed content storage system as file system(s)
  - Language Independent functionality (+ mappings to Java and C)
- Combine content & metadata into single “record” (XSet)
  - Multiple application and system metadata formats
  - Support for unstructured metadata (e.g., thumbnails)
  - Flat namespace for system scaling
- Basic Query, Management and Security functionality
  - Rest of functionality left to applications and/or other interfaces

SNIA: XAM API status

- FCAS Technical WG (TWG) developing specs
  - FCAS = Fixed Content-Aware Storage
- Initial XAM API specs are functionally complete
  - Architecture, language bindings to C and Java
    - C binding also covers C++, but does not use C++ objects
  - 2007: Review and revision to produce final v1 spec
- Reference implementation under development
  - XAM SDK (Software Development Kit) TWG

Storage Networking Standards

Data Protocols

- SCSI (T10)
- Fibre Channel (T11)
- ATA and SATA (T13)
- NAS (IETF, Microsoft)
- IP Storage (IETF)

Management

- SMI-S (SNIA)
  - Uses CIM (DMTF)
- SNMP (IETF)
- Web (IETF, W3C)

Stored Data

- RAID Layout (SNIA)
- Encryption (IEEE)
- Fixed Content (SNIA)

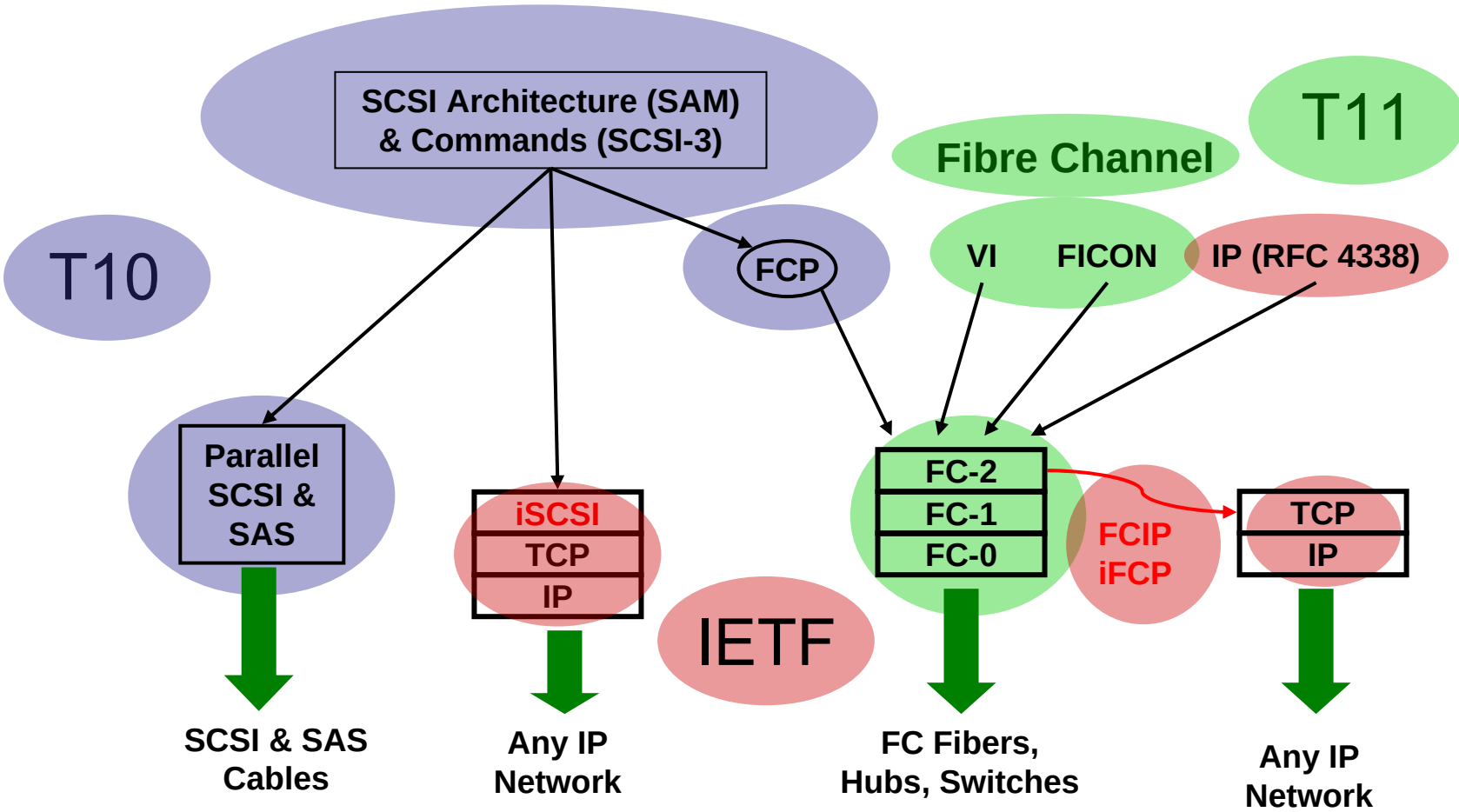
Getting Involved in Standards

- End users are always welcome
  - Remind participants why the standard matters
  - Help make the end result usable and useful
- Lots of opportunities to participate
  - Voting or contributing member
  - Observer: Still very important
    - No Substitute for hallway conversations

National and International Standards

- Standards Progression Path:
  1. Development in standards body (e.g., T11 for Fibre Channel),
  2. Becomes a national standard (e.g., ANSI)
  3. Becomes an international standard (ISO)
- INCITS: Umbrella Standards Organization
  - Umbrella for T10 (SCSI), T11(FC), T13 (ATA) and SNIA
  - Shepherds completed standards to ANSI and ISO
  - Usual path: Completed standard to INCITS to ANSI to ISO
- Not all standards follow this path
  - IETF RFCs are internationally recognized without ISO approval
  - Industry consortia standards (e.g., PCI)
  - Vendor de-facto standards (e.g., CIFS)

Block Storage Communication Protocols



Storage-Related Standards Organizations

- Storage Networking Industry Association (www.snia.org)
  - Storage Management (SMI-S) and other topics
- Distributed Management Task Force (www.dmtf.org)
  - Systems Management
- INCITS Technical Committees
  - SCSI and SAS: T10 (www.t10.org)
  - Fibre Channel: T11 (www.t11.org)
  - ATA and SATA: T13 (www.t13.org)
- IETF: Internet Engineering Task Force (www.ietf.org)
  - IP and Internet-related protocols, including IP Storage and NFS
- IEEE: Institute of Electrical and Electronics Engineers
  - Encrypted Media: P1619 (ieee-p1619.wetpaint.com)

Q&A / Feedback

- Please send any questions or comments on this presentation to SNIA: trackvirtualization@snia.org and the SNIA Technical Council: snia-tc@snia.org

Many thanks to the following individuals for their contributions to this tutorial.

SNIA Education Committee

David L. Black
Mark Carlson
Zoran Cakeljic
Eric Hibbard
Steven Wilson
Bill Dawkins
SW Worth
Walter Dey