SDC
STORAGE DEVELOPER CONFERENCE
SNIA SANTA CLARA, 2016

Design of a WORM Filesystem

Terry Stokes
Dell EMC/Isilon

Who am I?

- 15+ years in network protocols.
- Previously worked at Blue Coat Systems, RadioFrame and Microsoft Anti-trust (MCPP).
- Co-founded a network capture device startup in 2003.
- Currently work in the AIMA group at Isilon.

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

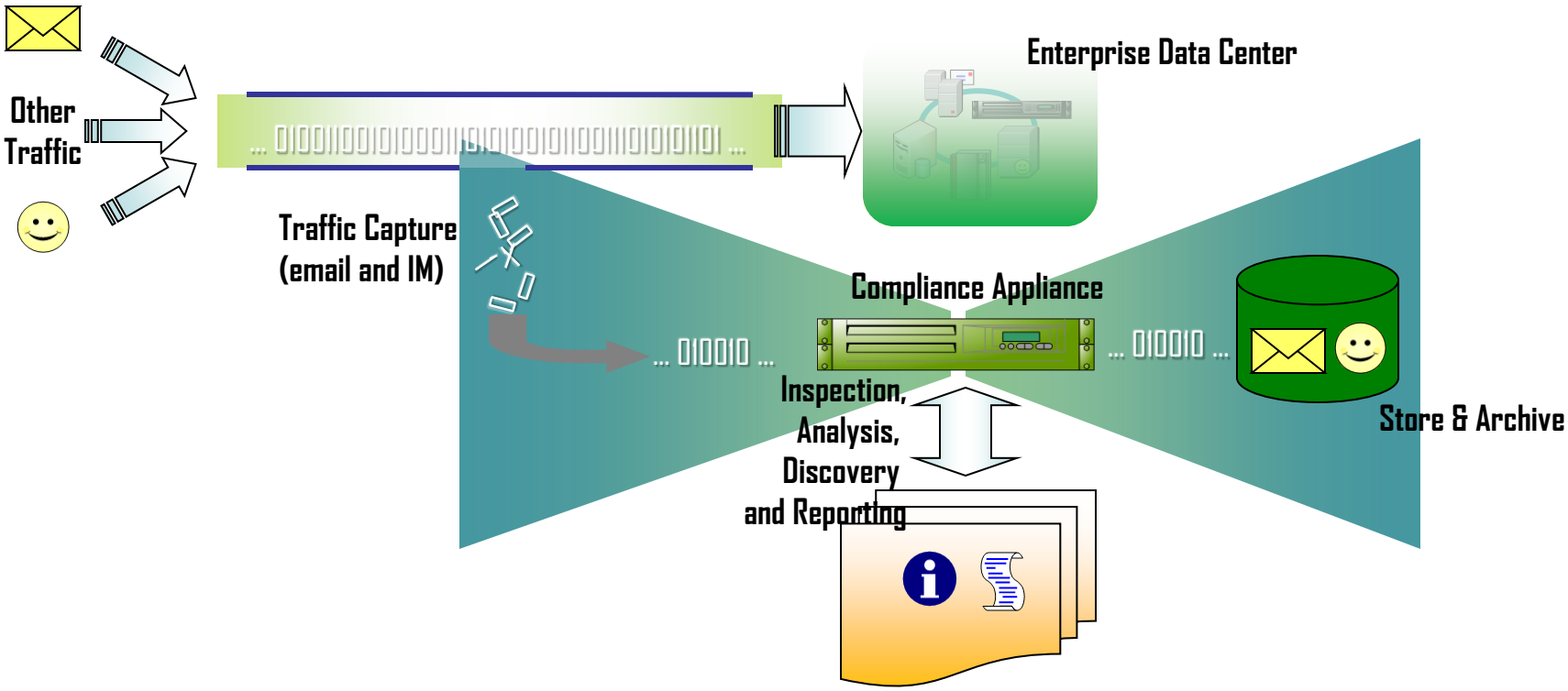
Disclaimers

- Not Isilon Smartlock design.
- Dated design, developed in 2003.
- I’m not a filesystems expert.

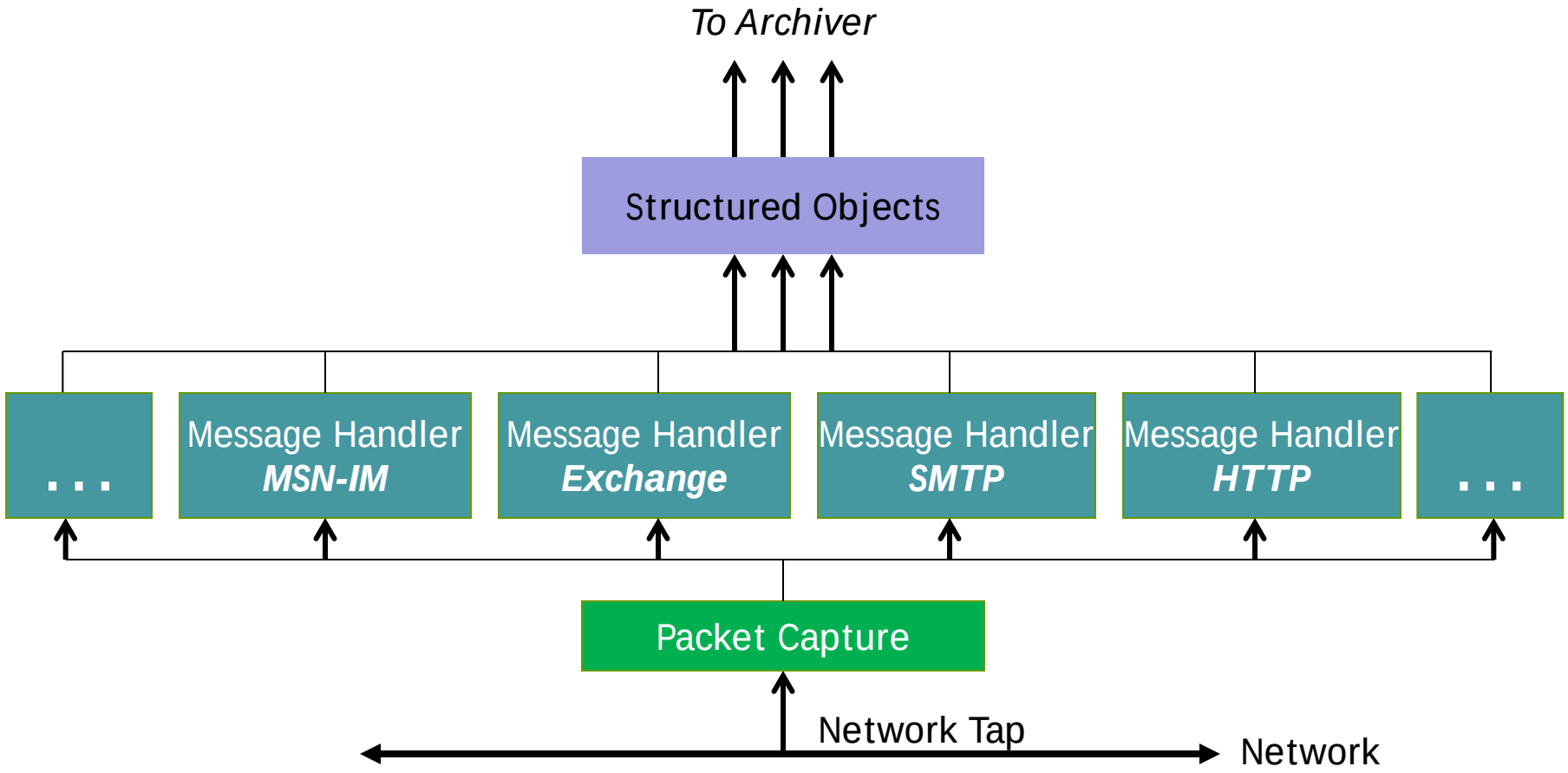
2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

- Founded in 2003, folded in 2005.
- Product called the “Compliance Appliance”.
- Capture and archive electronic communications.
- Incorporated a WORM filesystem.
- 10+ “sweat equity” employees.
- Two WORM filesystem patents.
- Failed to get funding.

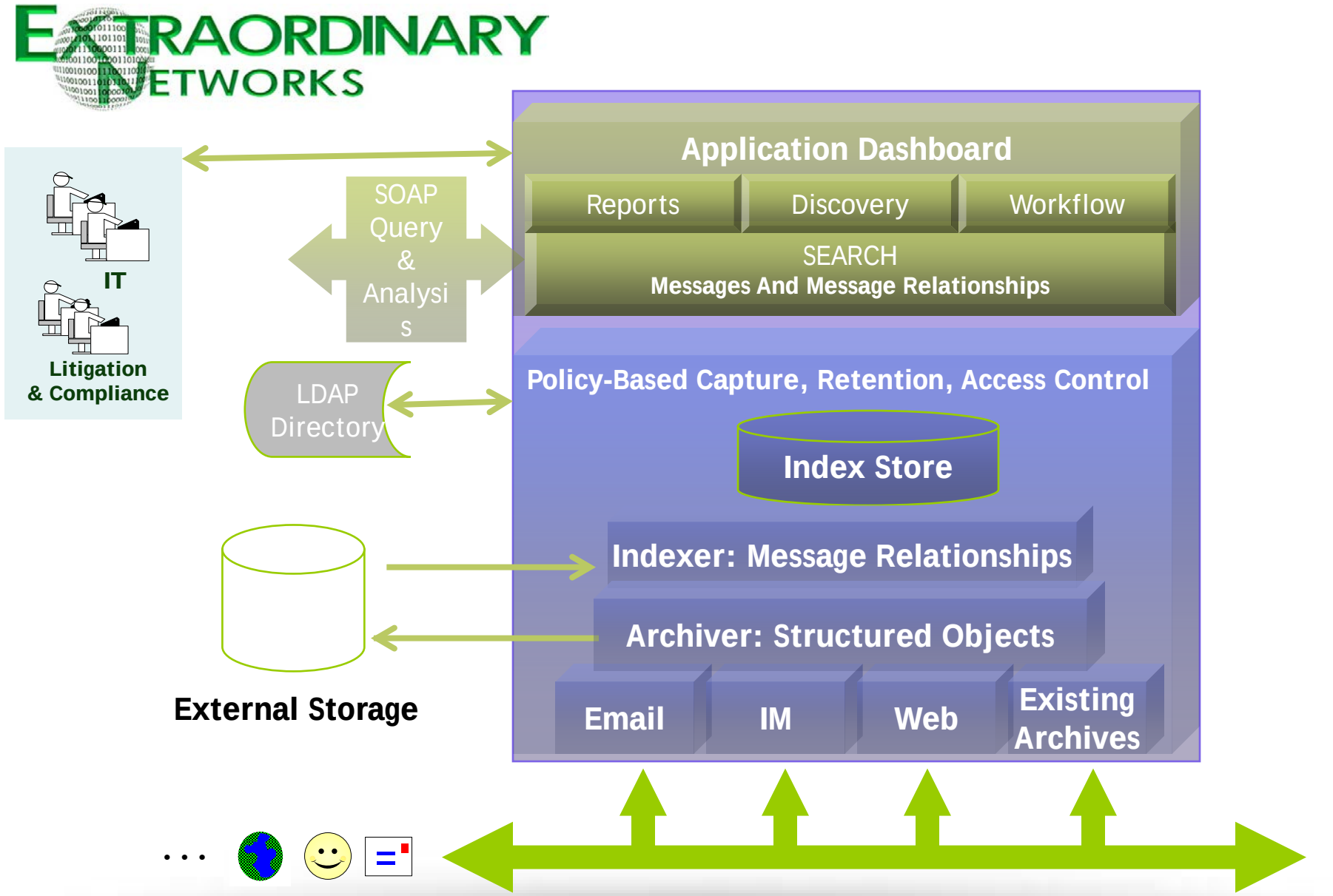
2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.



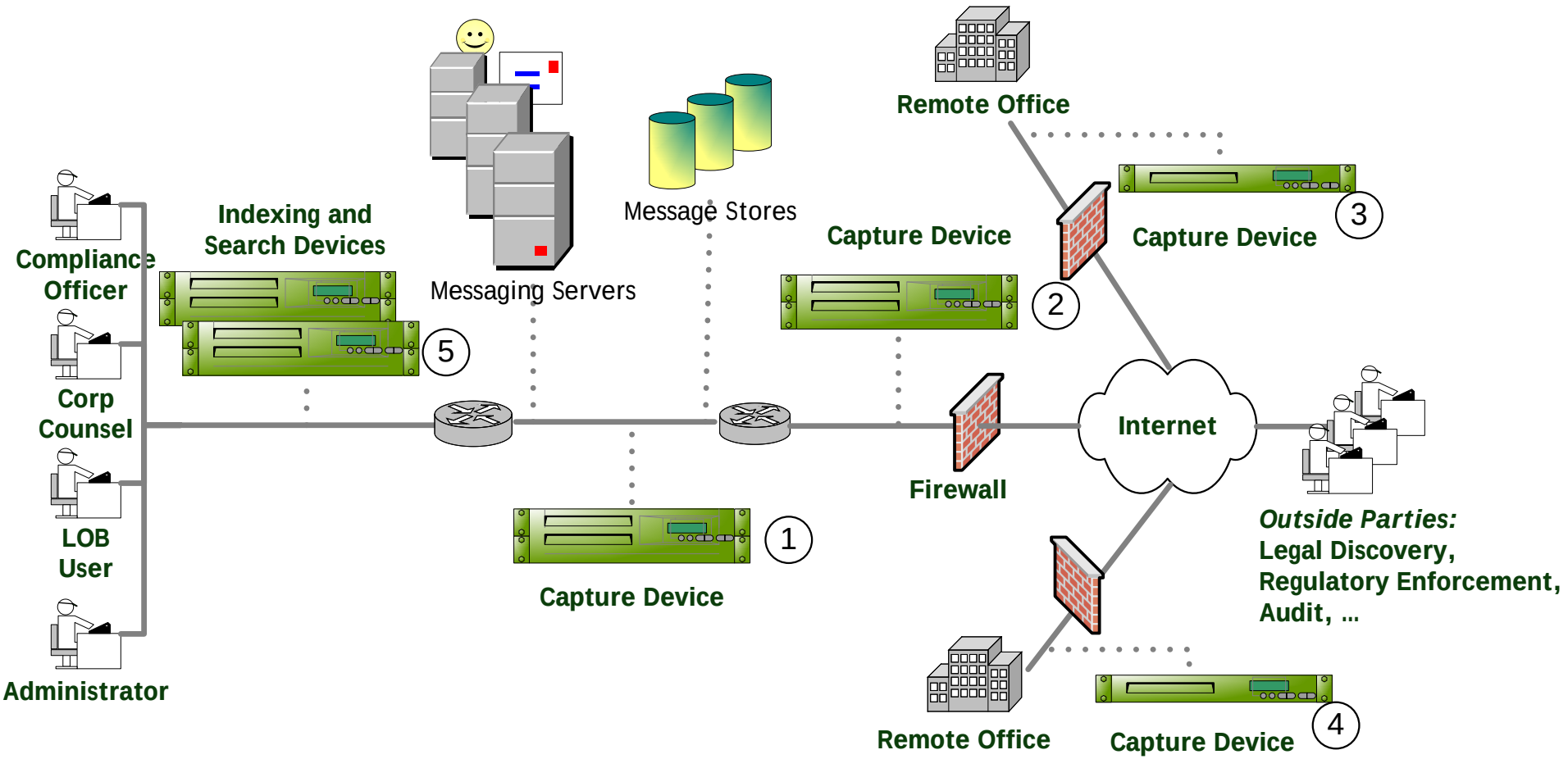
2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.



2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.



SDC 16
2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.



2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

WORM Storage 2003

- Optical storage (CD-R, optical tape, etc).
- EMC Centera CE
  - Content-Addressable Storage
  - Content hashed for storage location.
- NetApp Snaplock
  - Plugin for Data OnTap
  - Software prevents file deletion.

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Customer Research

- Talked to 10-15 potential customers.
- Banks, Brokers, Healthcare, Manufacturers.
- 10-15 investment firms.
- Several had home grown software solutions.
- A few simply backed up optical storage.
- Issue was greater than simply providing a better WORM storage solution.

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Regulations

SOX	Section 404, 802
HIPAA	Protected Health info
Securities Regulations	SEC 17a-4 NASD 3010/3110 NYSE 440/472
Gramm Leach Bliley Act	Fin Privacy Rule Safeguards Rule
Basel-II	Supervisory Control
CA SB 1386	Privacy and Protection
US Patriot Act	Cust Identification Pgm
ISO 17799	Retention & Safeguarding
DOD 5015.2-STD	ERecords Management

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Multiple Protocols

- SEC rules are ambiguous:
  - All broker-dealer communications must be held for 3 years on non-rewriteable and non-erasable storage.
  - Originally SEC only enforced email storage.
  - Added instant messaging requirement.
  - What about broker web pages? VoIP?
- Each protocol requires a separate solution.

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Slow Message Retrieval

- Message discovery requests (from legal and SEC) have a time limit.
- Many SEC requests involved several brokers, with each broker search taking 2-3 days.
- By the time search was complete, only a day or two left to review before handover.
- Didn’t know what was being handed over.
- Legal Discovery industry born from this.

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Deletion of Expired Messages

- Emails can hurt you!
  - Harmful emails have been used in several notable trials, Microsoft for one.
  - One unexpired email can indirectly hold hundreds of expired emails on optical storage media.
- It is important for a company to delete emails as soon as regulations allow.

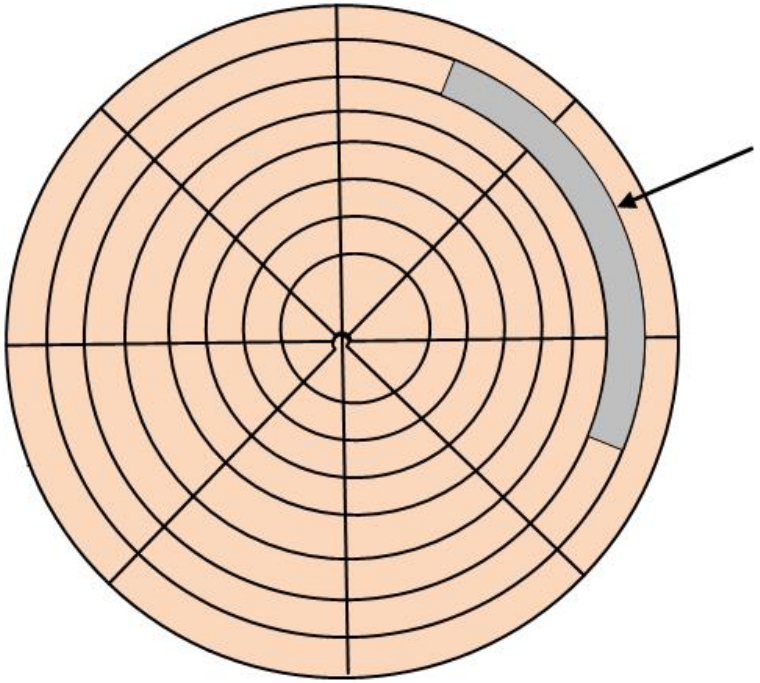
2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Survey of Filesystems

- Reviewed top ten filesystems for ideas.
- ReiserFS:
  - Uses B+ trees for file objects (inodes, directories and file data).
  - Employs tail packing.
- Sun XFS:
  - Use B+ trees for tracking free extents and free inodes.
  - Uses extents instead of blocks.

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

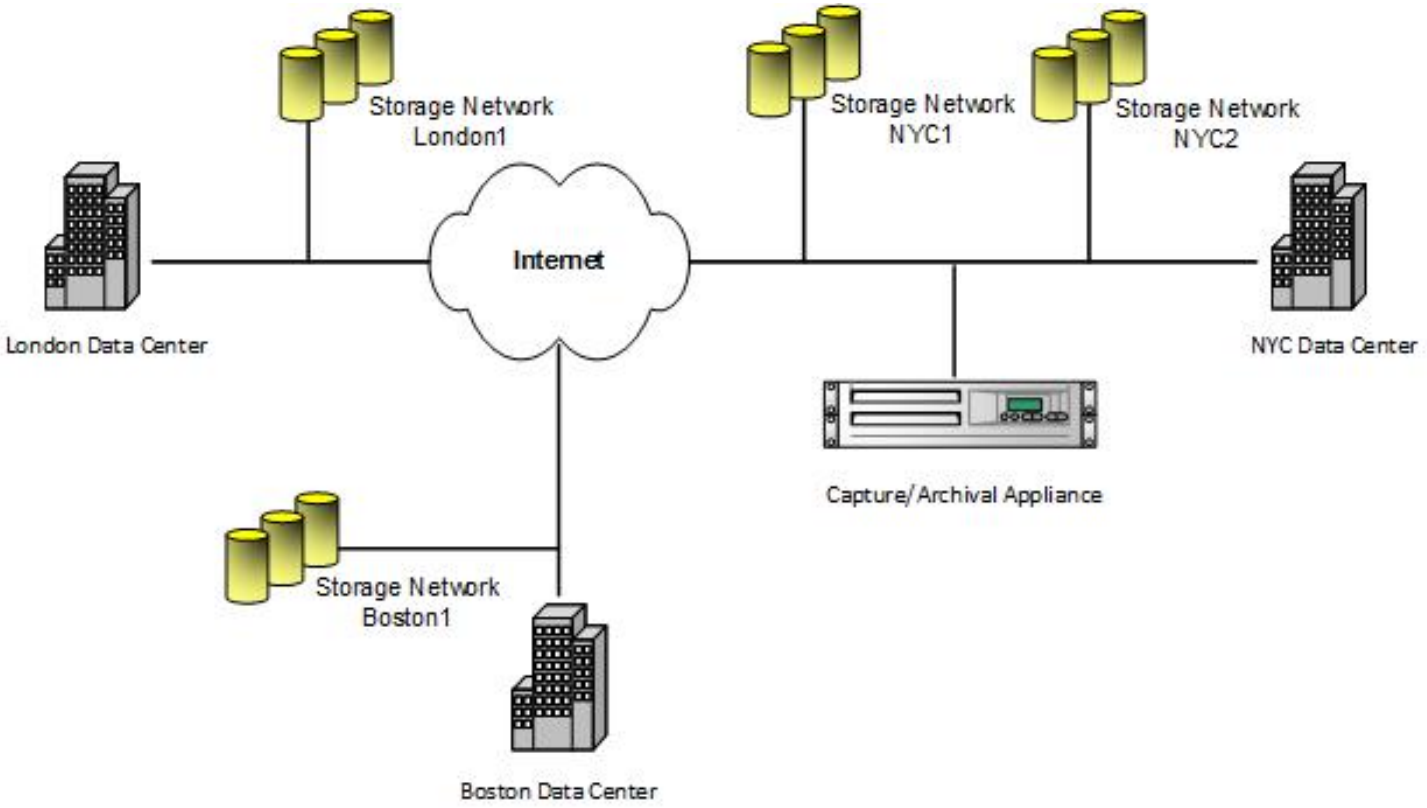
Extents



SDC 16

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Distributed Storage



SDC 16

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Distributed Storage

Start Date	ID Start	ID Stop	Location	Storage Partition	State	Free MB	Access ms
2/3/06	0000	1234	London1	\groupL1a	ready	52369	132
10/23/05	1235	2254	NYC1	\groupN1a	ready	43221	23
10/23/05	2255	3378	NYC2	\groupN2a	ready	96676	34
10/23/05	3378	4865	NYC2	\groupN2b	ready	45312	35
9/18/06	4866	7697	Boston1	\groupB1a	ready	12314	80
6/16/06	7698	8745	NYC2	\groupN2c	ready	23890	34
9/18/06	8746	9999	Boston1	\groupB1b	ready	67114	85
3/27/04	2687	3956	NYC1	\groupN1a	read only	43221	23
7/14/03	5586	6132	NYC1	\groupN1b	read only	324	23

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Unsupported Operations

Supported Operations

- Filesystem Mount/Unmount
- Filesystem Statistics
- Volume Statistics
- File Creation
- File Reading
- File Deletion (if past retention period)
- File Annotation (auditing functions)

Unsupported Operations

- File Deletion (prior to retention period)
- File Content Modification or Appending
- File Attribute Modification
- Directory Support (creation, listing, deletion, etc)
- Symbolic Links

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Specialized Storage

- Formatted for electronic communications.
  - Email
  - Instant Messaging
  - HTTP – Web Pages
  - VoiceOverIP
- File content and meta data encrypted.

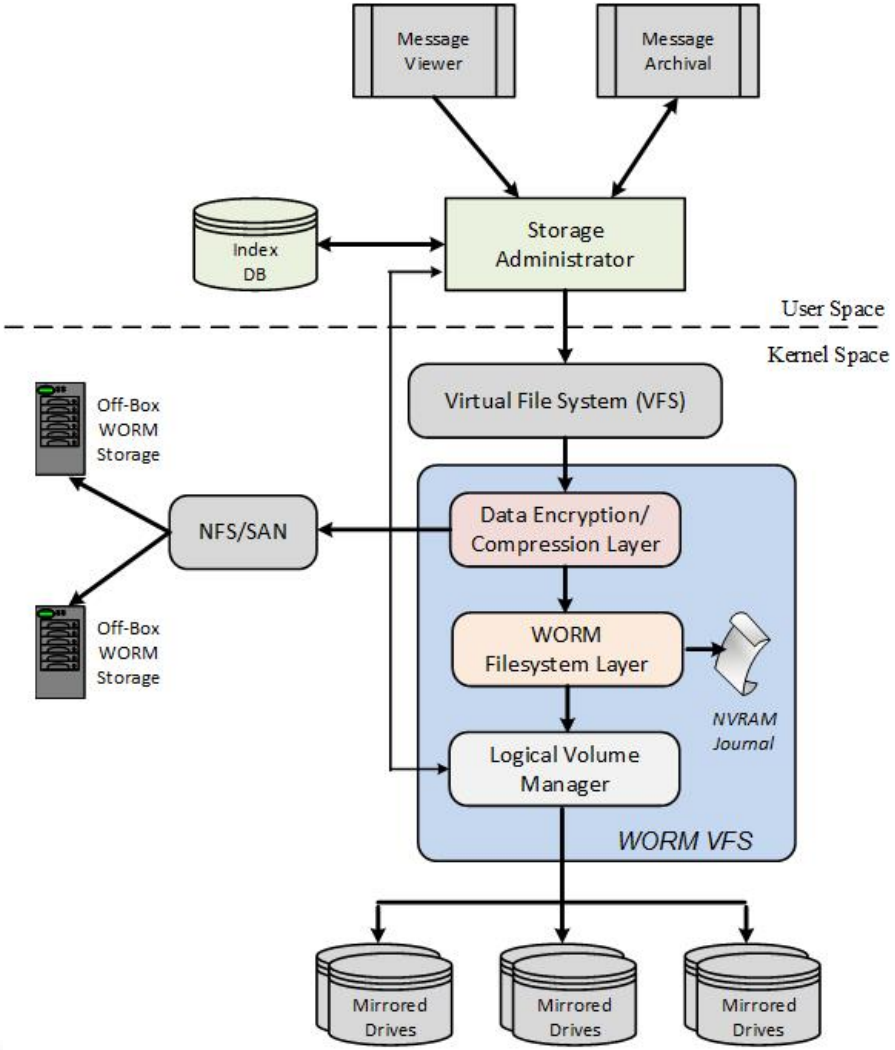
2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Specialized Storage

- Restricted File Writes
  - Write out entire file or not at all.
  - No appends or modifications.
  - NVRAM backed journal for crashes.
- Restricted File Reads.
  - Only two types:
    - Read inode (for searches)
    - Read entire file (for viewing/gathering)

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

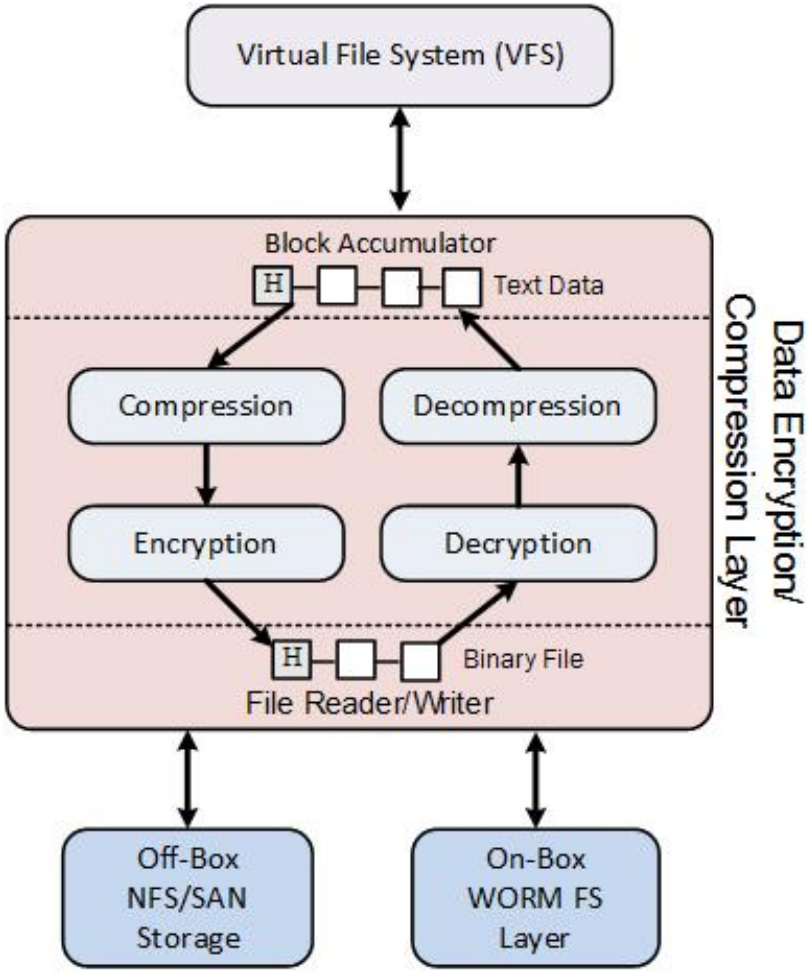
Layered Filesystem



SDC 16

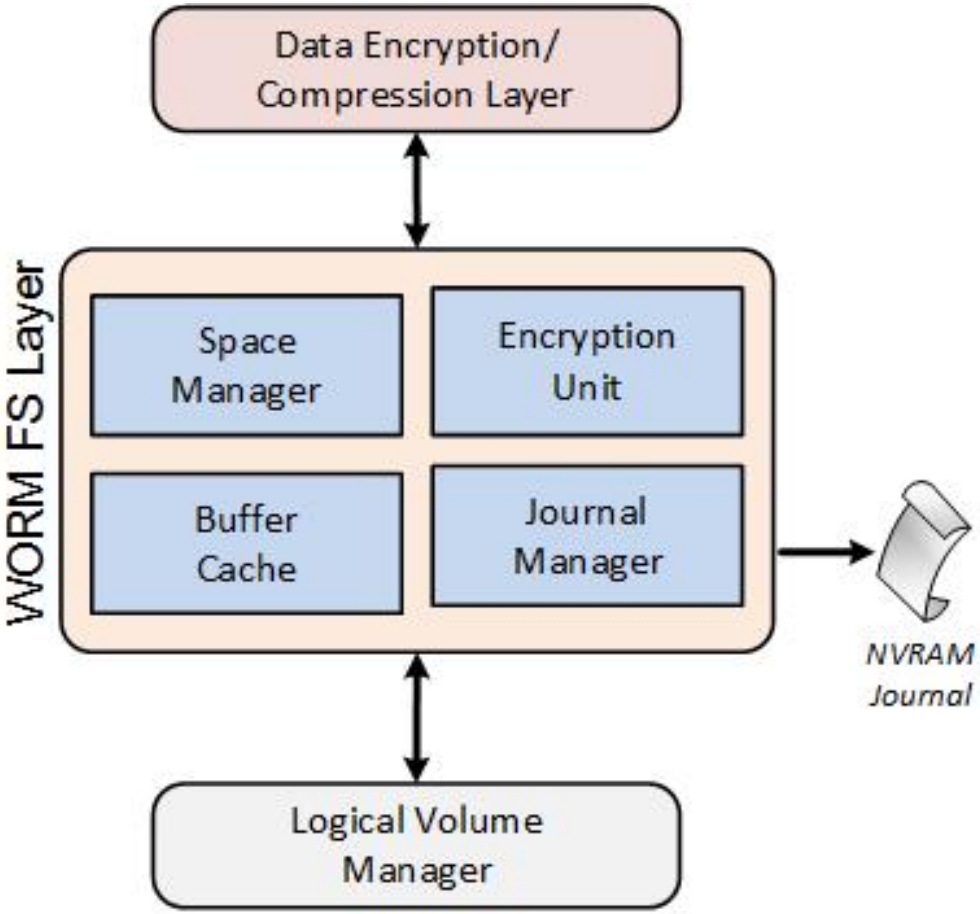
2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Data Encryption/Compression Layer



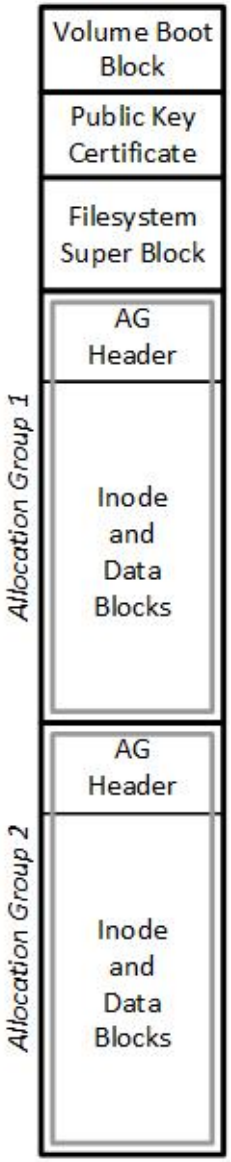
2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

WORM Layer



2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

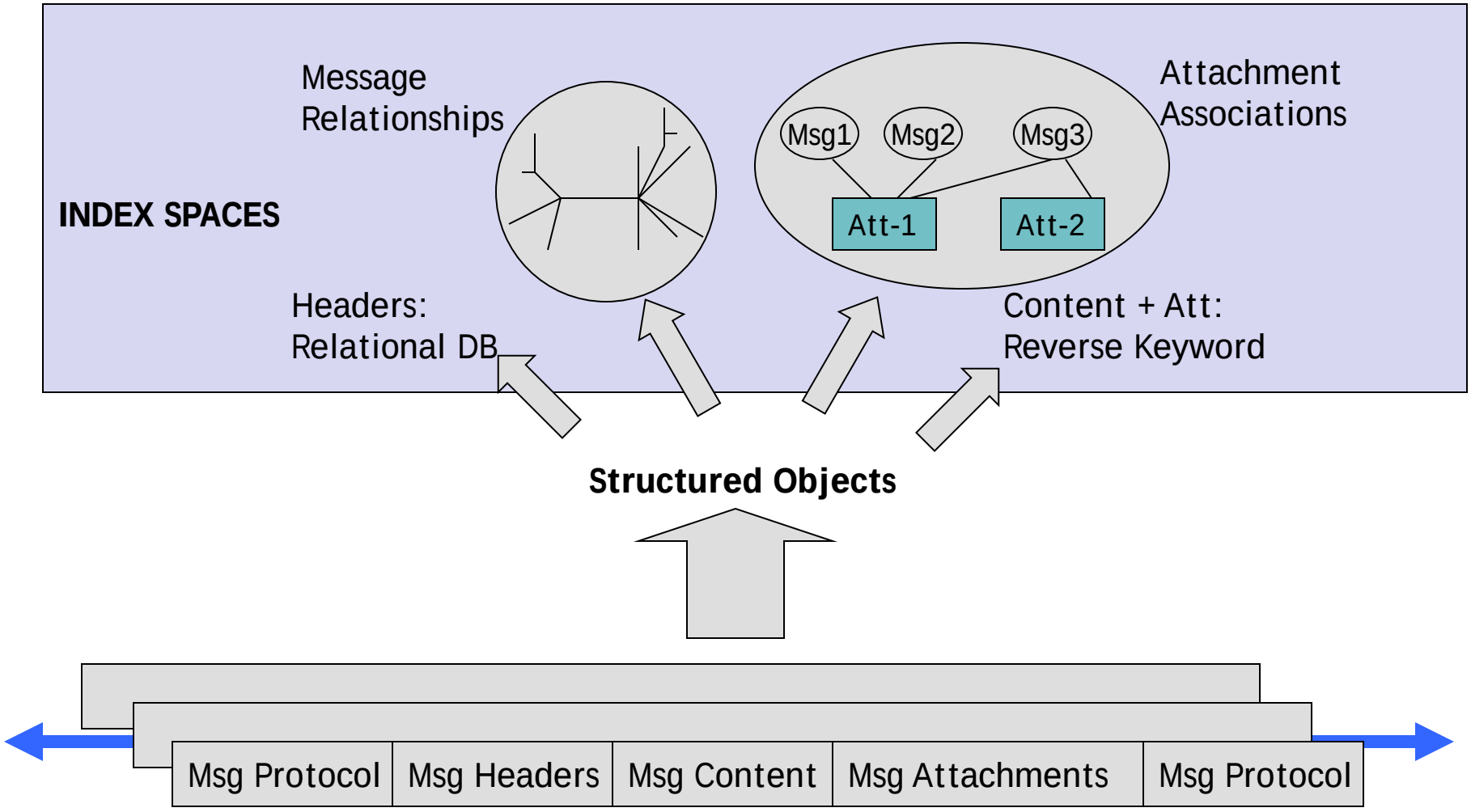
Volume Layout



SDC 16

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Structure Data



SDC 16

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

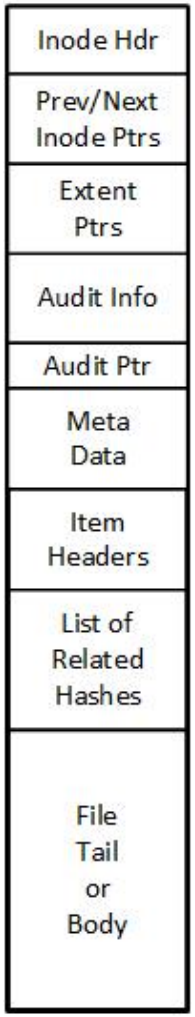
Superblock Format

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

AG Header Format

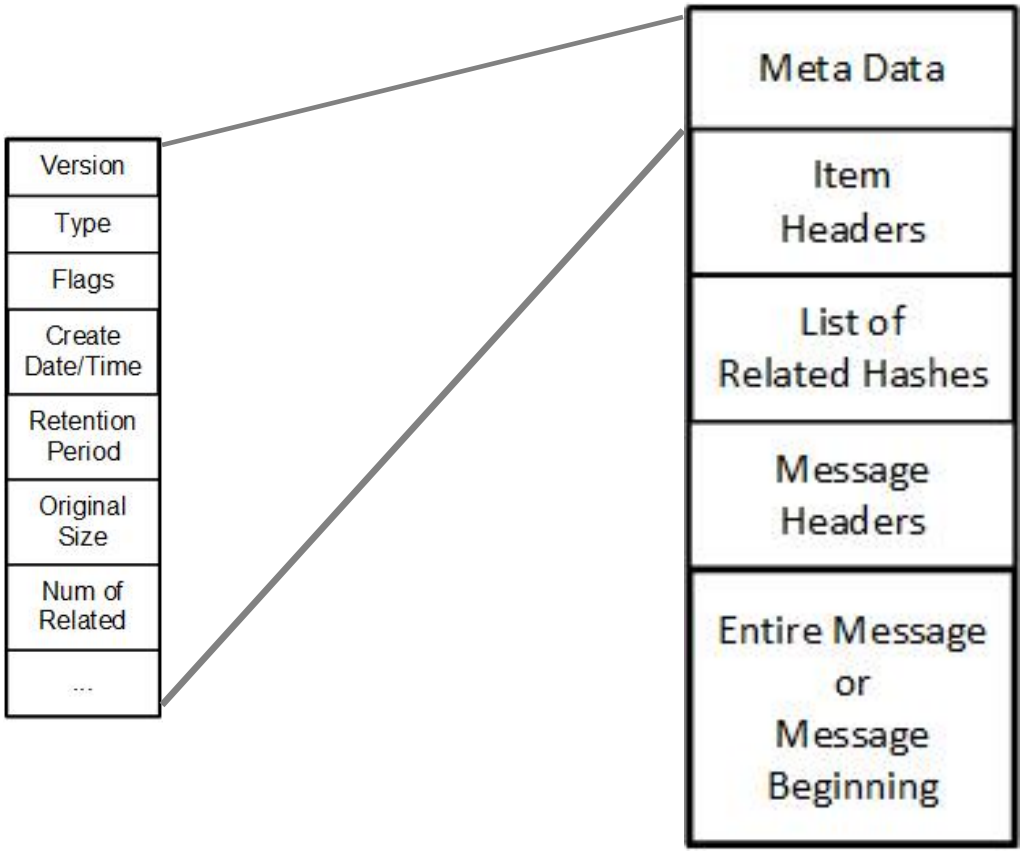
2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Inode Format



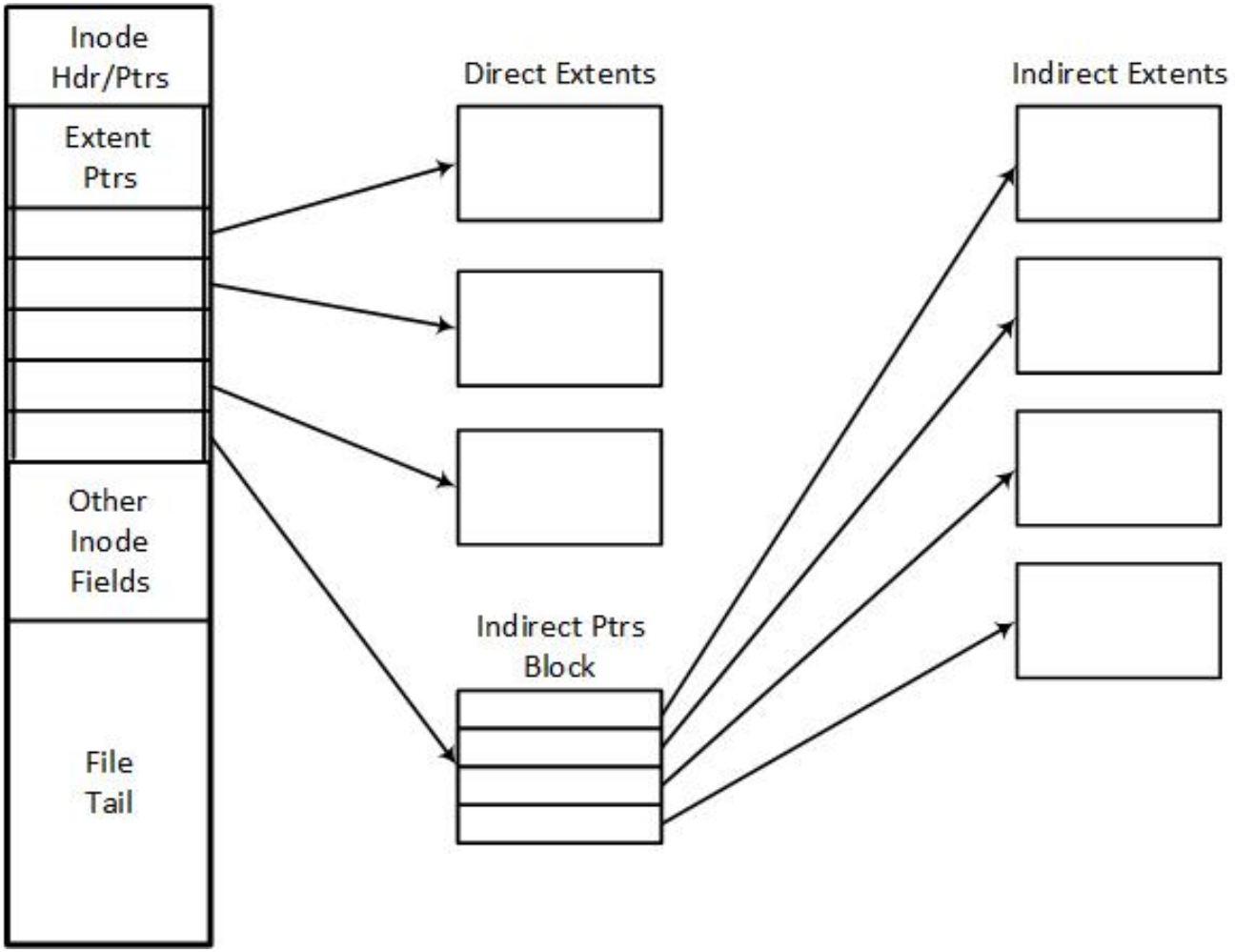
2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

File Format



2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Disk File Format



2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Automatic File Deletion

- Hourly process to delete expired files.
- Deletion holds table:

Rule	Condition
1.	Type = Any and Groups = Execs
2.	Type = Email and Users = John Smith, George Martin, Pete Nobody
3.	Type = Spreadsheets, Email, Reports and Groups = Execs, Finance
4.	Type = VoIP and Groups = Customer Service
5.	Type = Any and Users = All

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

WORM Characteristics

- Disk structures and data encrypted.
- Non-standard disk format.
- Unsupported file and directory operations.
- Access to file contents restricted via software.

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Performance Characteristics

- Use of extents.
- File contents optimized on disk.
- Tail packing in inode.
- Read operations optimized for archival.
- Structured file format.
- B+ trees for free inodes and extents.

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

If I did it again…

- Investigate SED drives.
- Drop support for off-box storage.
- Include disk defragmentation scheme.
- Add rudimentary directory structure.
- Remove distributed storage by location.

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.

Questions?

terry.stokes@emc.com

2016 Storage Developer Conference. © Terry Stokes. All Rights Reserved.