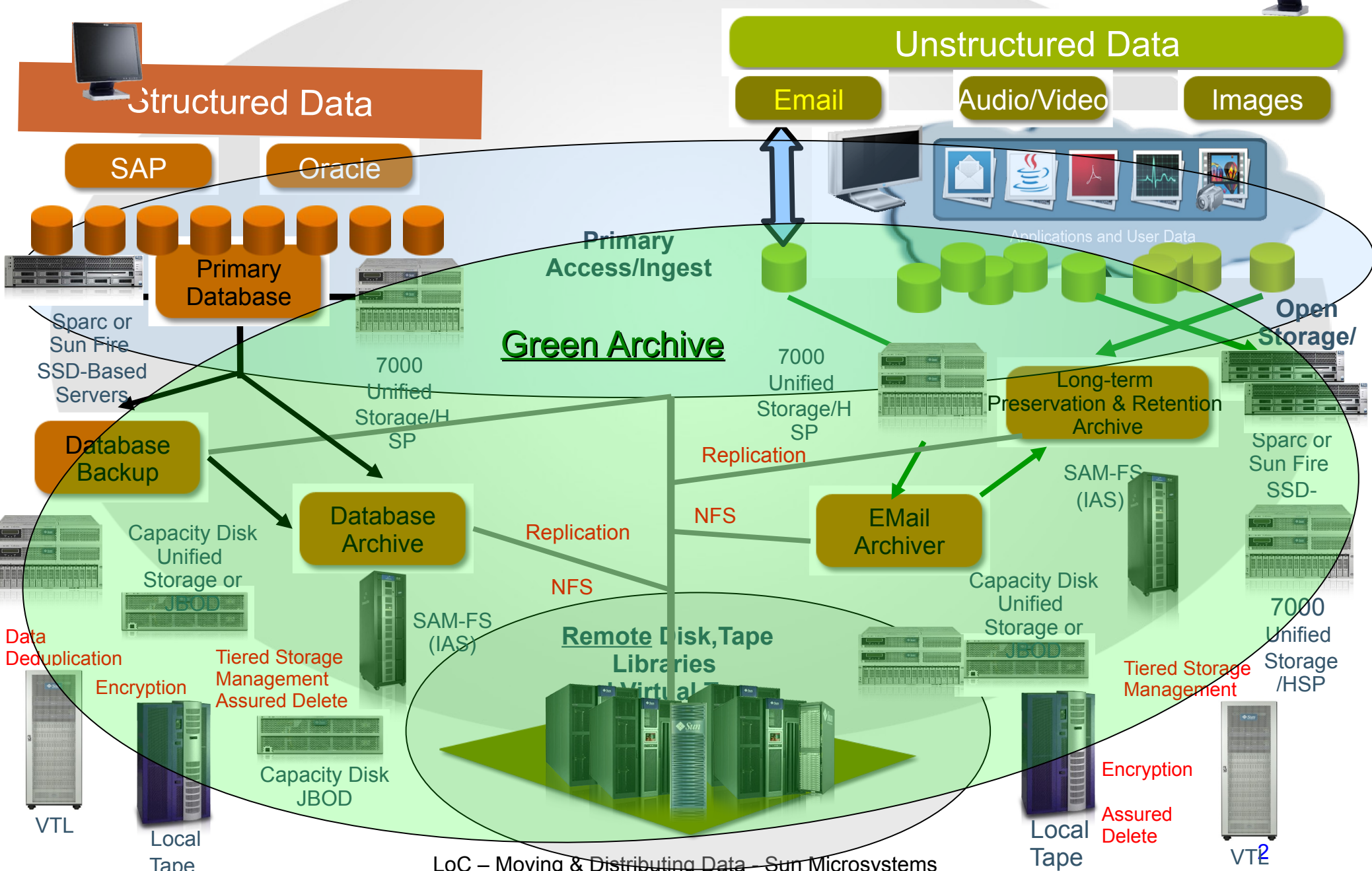# Moving Data and Distributing Data

Raymond A. Clarke

Sr. Enterprise Storage Solutions Specialist, Sun Microsystems - Archive & Backup Solutions

SNIA Data Management Forum, Board of Directors

Sun's Enterprise Archive Alternatives

LIBRARY OF CONGRESS

Sun microsystems

Unstructured Data

Email   Audio/Video   Images

Structured Data

SAP   Oracle

Applications and User Data

Primary Access/Ingest

Primary Database

Sparc or Sun Fire SSD-Based Servers

Green Archive

7000 Unified Storage/HSP

7000 Unified Storage/HSP

Long-term Preservation & Retention Archive

Open Storage/

Database Backup

Replication

Database Archive

Replication

NFS

EMail Archiver

SAM-FS (IAS)

Sparc or Sun Fire SSD-

Capacity Disk Unified Storage or JBOD

SAM-FS (IAS)

NFS

Capacity Disk Unified Storage or JBOD

7000 Unified Storage /HSP

Data Deduplication

Encryption

Tiered Storage Management Assured Delete

Remote Disk,Tape Libraries and Virtual Tape

Tiered Storage Management

VTL

Local Tape

Capacity Disk JBOD

Encryption

Assured Delete

Local Tape

VTL

2

Challenge: Manage Data for 75++ Years

- HW typically only backward compatible N-1
- Yearly capacity increases
- Every 2-5 years HW becomes obsolete:
  - > Need to migrate current data to newer HW components
  - > Replace compute parcels
  - > Replace FC Parcel for performance and capacity
  - > Replace tape drives and media to current technology
  - > Replace SATA parcels for capacity / footprint
- Minimize vulnerability
- HW migration is inevitable; PLAN for it

What's SNIA Doing About All This?

- Educates, Defines and Taking Action to Address Industry Challenges
- Specific Activities
  - XAM – eXtensible Access Method
  - Self-Contained Information Retention Format(SIRF)
    - Rationale & Objectives
    - Requirements & Use Cases
  - Bridging Terminology
  - Green Storage Initiative
  - Cloud Storage Initiative(CSI)

What is XAM?

- ## XAM is a SNIA Architecture
  - > The XAM Architecture spec defines the normative semantics of the API for use by applications and implementation by storage systems and standardizes metadata and services across XAM compliant systems

- ## XAM is an Application Programming Interface (API)
  - > The XAM Java API spec defines the binding of the XAM Architecture to the Java Language
  - > The XAM C API spec defines the binding of the XAM Architecture to the C Language

- ## XAM is SNIA Software
  - > The XAM SDK provides a common library and reference implementation to promote widespread adoption of the standard

The need for MetaData Standards

- Which can contains corn?
- Open the cans.
- How much does it cost?
- Ask the clerk.
- How many calories does it have?
- Ask the vendor.
- How does the store automatically manage inventory?
- They can't.

MetaData Standards

Standardized labeling allows multiple vendors to consistently represent information to consumers

Extended labeling for LOB uses

What's Sun Doing About All This?

- Open Source/Standards Community Engagement
- Specific Activities
  - Open Solaris – ZFS(Hybrid Storage Pool)
  - Open Storage
  - Cloud Computing

The Evolution of Data Storage:
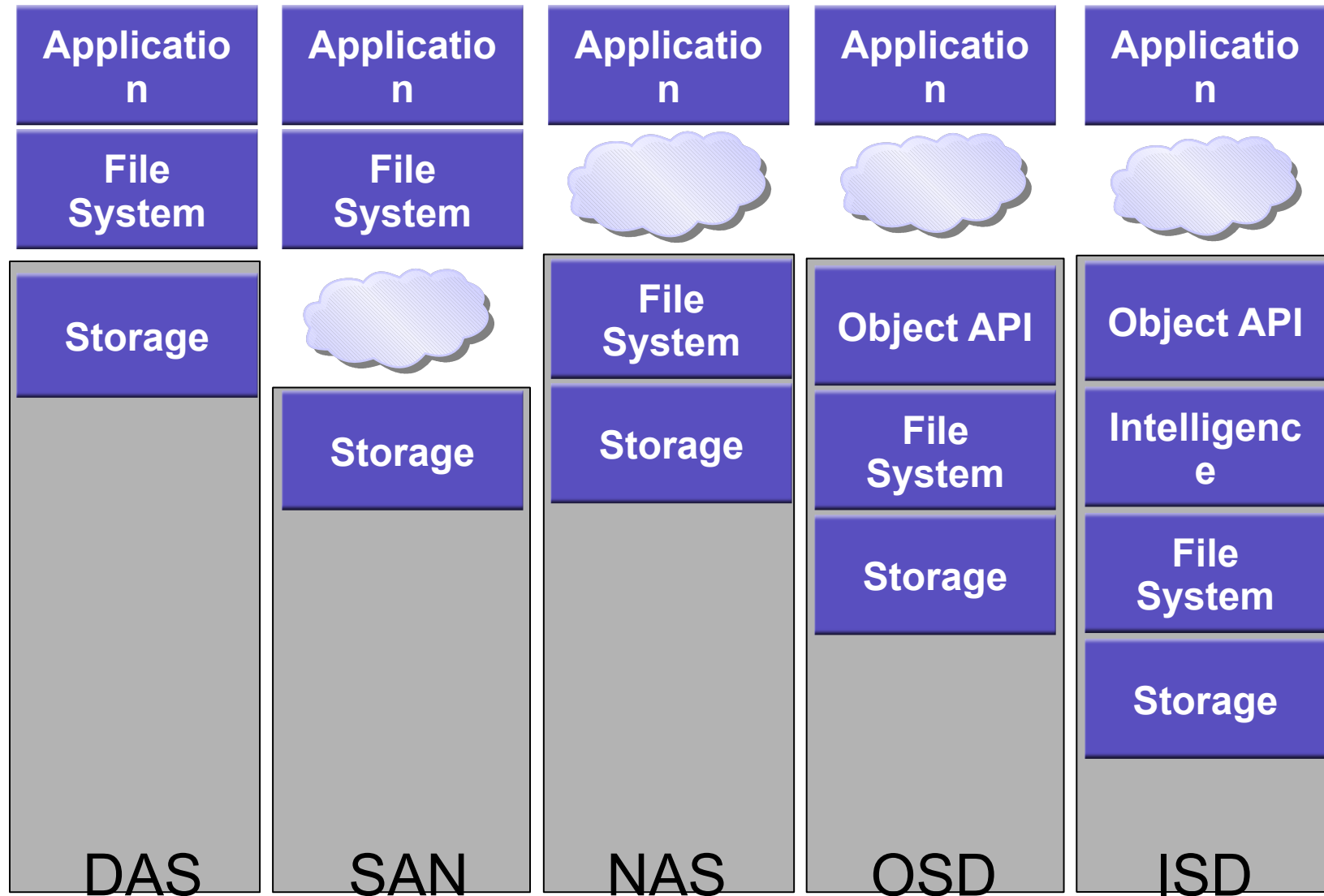
Application	Application	Application	Application	Application
File System	File System	*(cloud)*	*(cloud)*	*(cloud)*
Storage	*(cloud)*	File System	Object API	Object API
	Storage	Storage	File System	Intelligence
			Storage	File System
				Storage
DAS	**SAN**	**NAS**	**OSD**	**ISD**

What is ZFS?
A new way to manage data

End-to End Data Integrity

With check-summing and copy-on-write transactions

Easier Administration

A pooled storage model – no volume manager

Immense Data Capacity

The world's first 128-bit file system

Data Services

Snapshots
Clones
Replication

FS/Volume Model vs. ZFS

Traditional Volumes

Abstraction: virtual disk

Partition/volume for each FS

Grow/shrink by hand

Each FS has limited bandwidth

Storage is fragmented, stranded
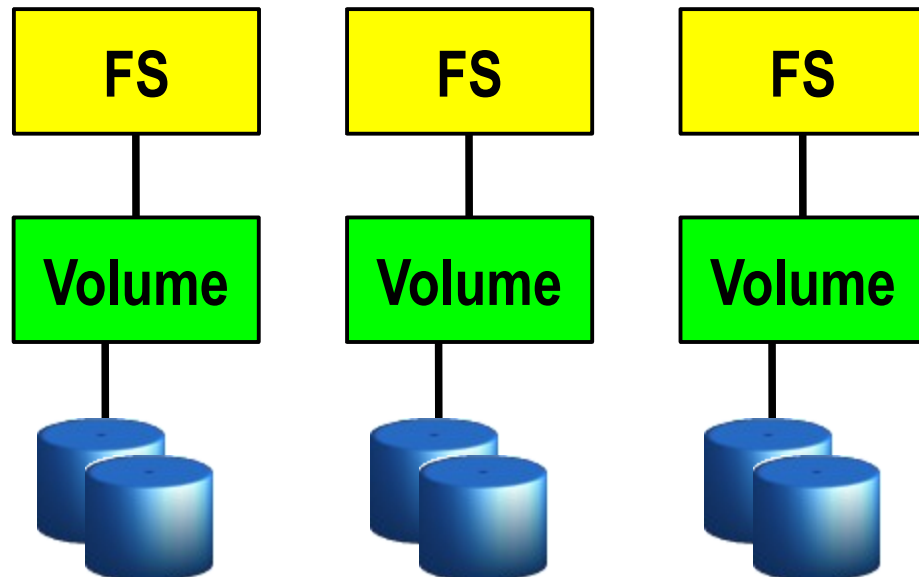
ZFS Pooled Storage

Abstraction: malloc/free

No partitions to manage

Grow/shrink automatically

All bandwidth always available

All storage in the pool is shared
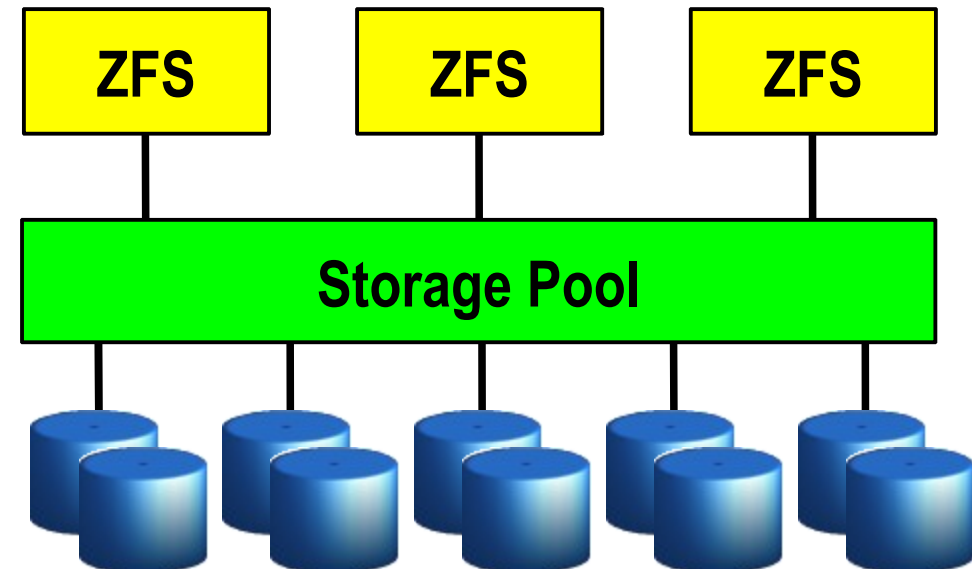
What is ZFS?
A new way to manage data

End-to End Data Integrity

With check-summing and copy-on-write transactions

Easier Administration

A pooled storage model – no volume manager

Immense Data Capacity

The world's first 128-bit file system
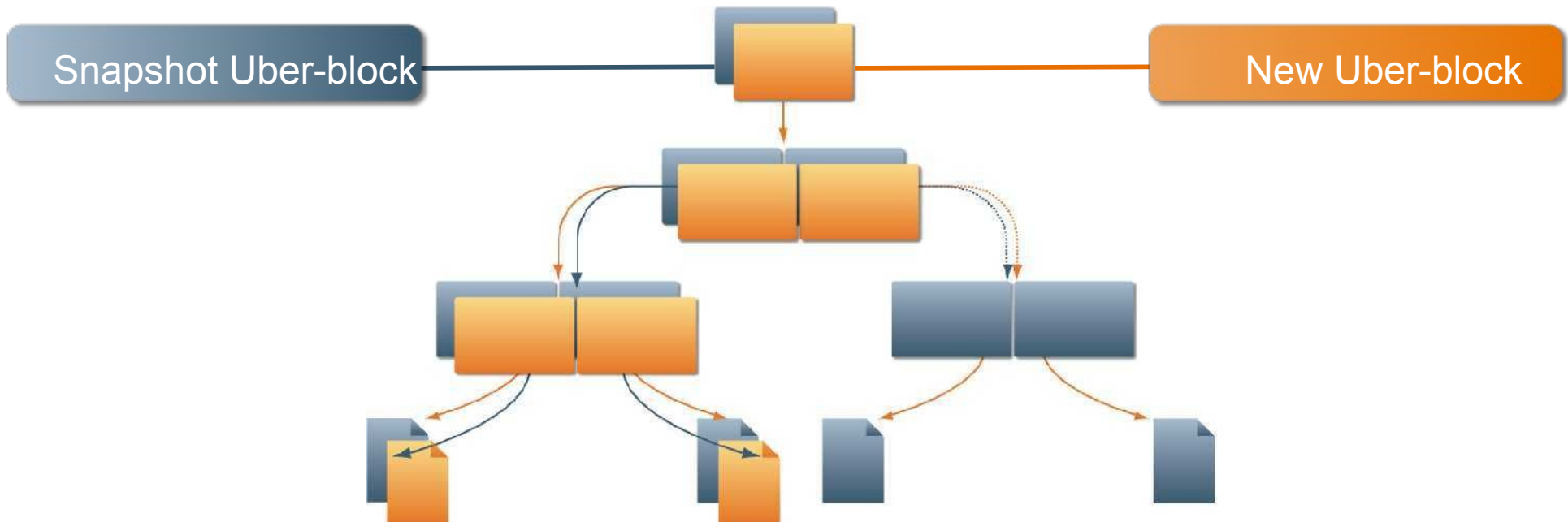
Integrated Data Services

Snapshots
Clones
Replication

ZFS Snapshots

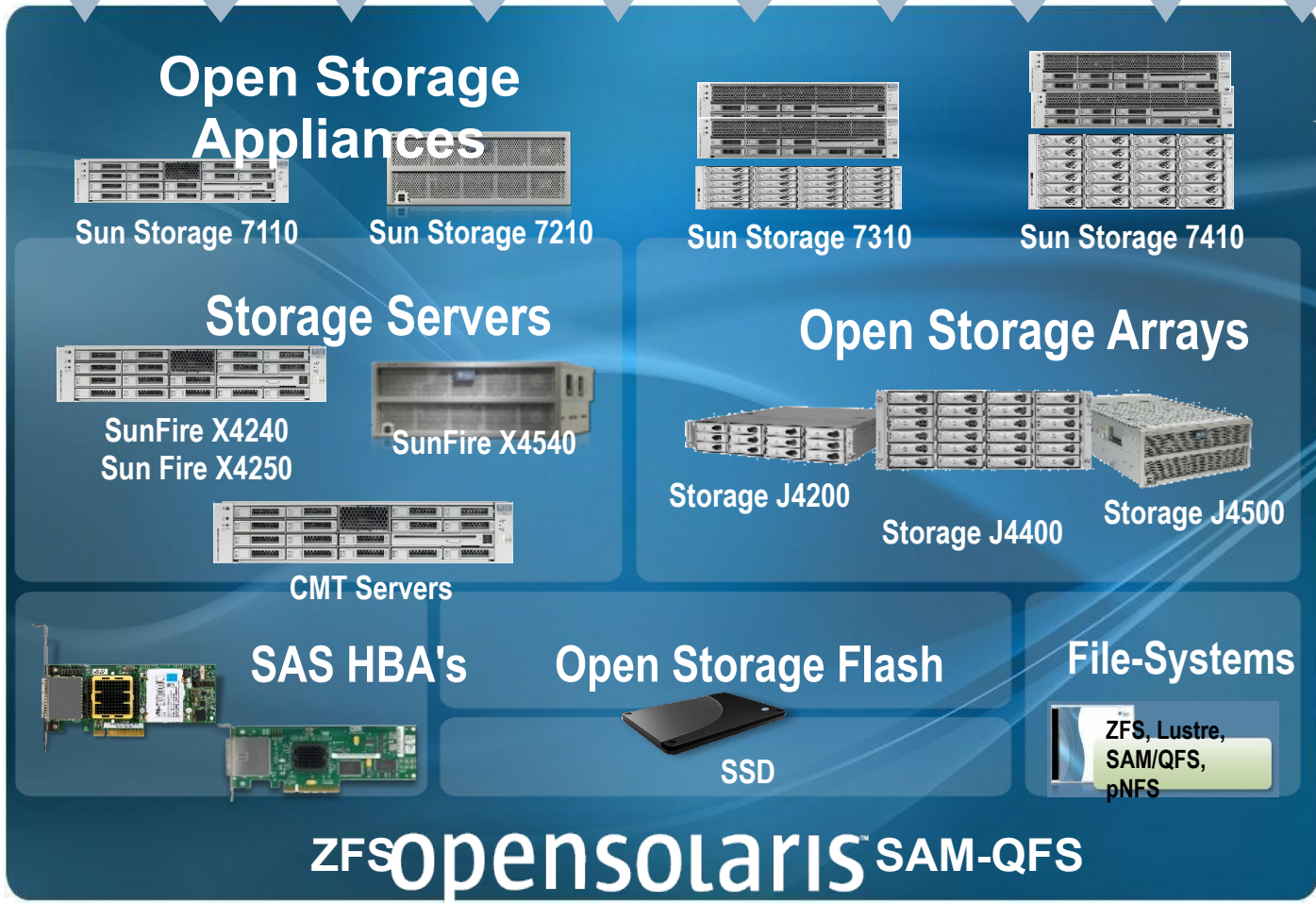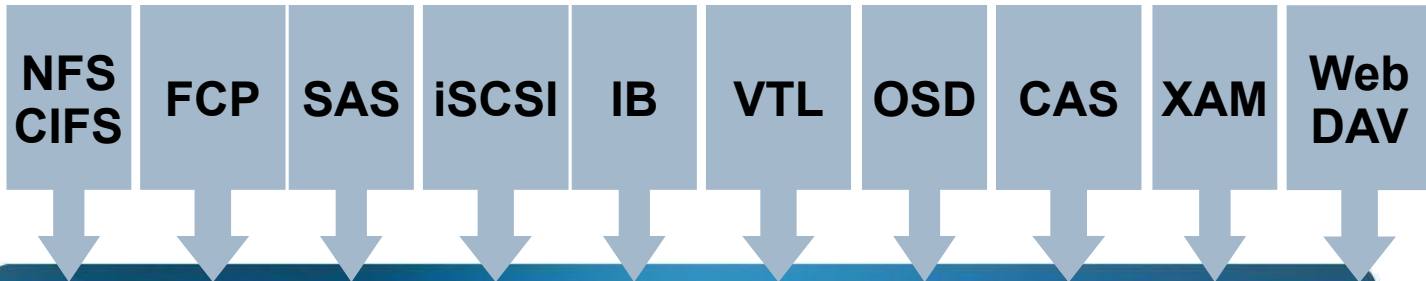Provide a read-only point-in-time copy of file system

Copy-on-write makes them essentially "free"

Very space efficient – only changes are tracked

And instantaneous – just doesn't delete the copy

Snapshot Uber-block

New Uber-block

Thank You
for
Your Time and
Attention

Raymond.Clarke@Sun.com

(212) 558-9321