

Trends bei Speichertechnologien

3. BITKOM-Anwenderforum IT-Infrastruktur

Frankfurt, 27. November 2008

Christian Bandulet

Principal Engineer

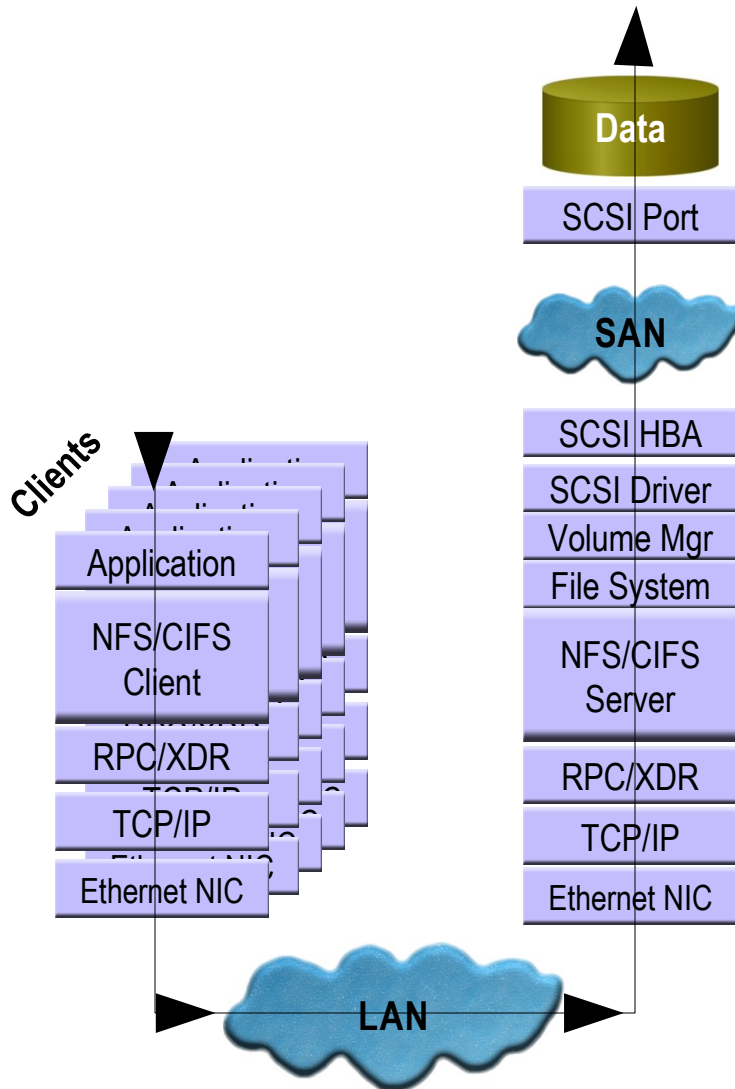
Sun Microsystems



- ▶ **WAFS – Wide Area File Services**
- ▶ **Data De-Duplication**
- ▶ **Storage Virtualization- Thin Provisioning**
- ▶ **Solid State Disk (SSD)**
- ▶ **Storage Grid/Cluster**
- ▶ **Fibre Channel over Ethernet**

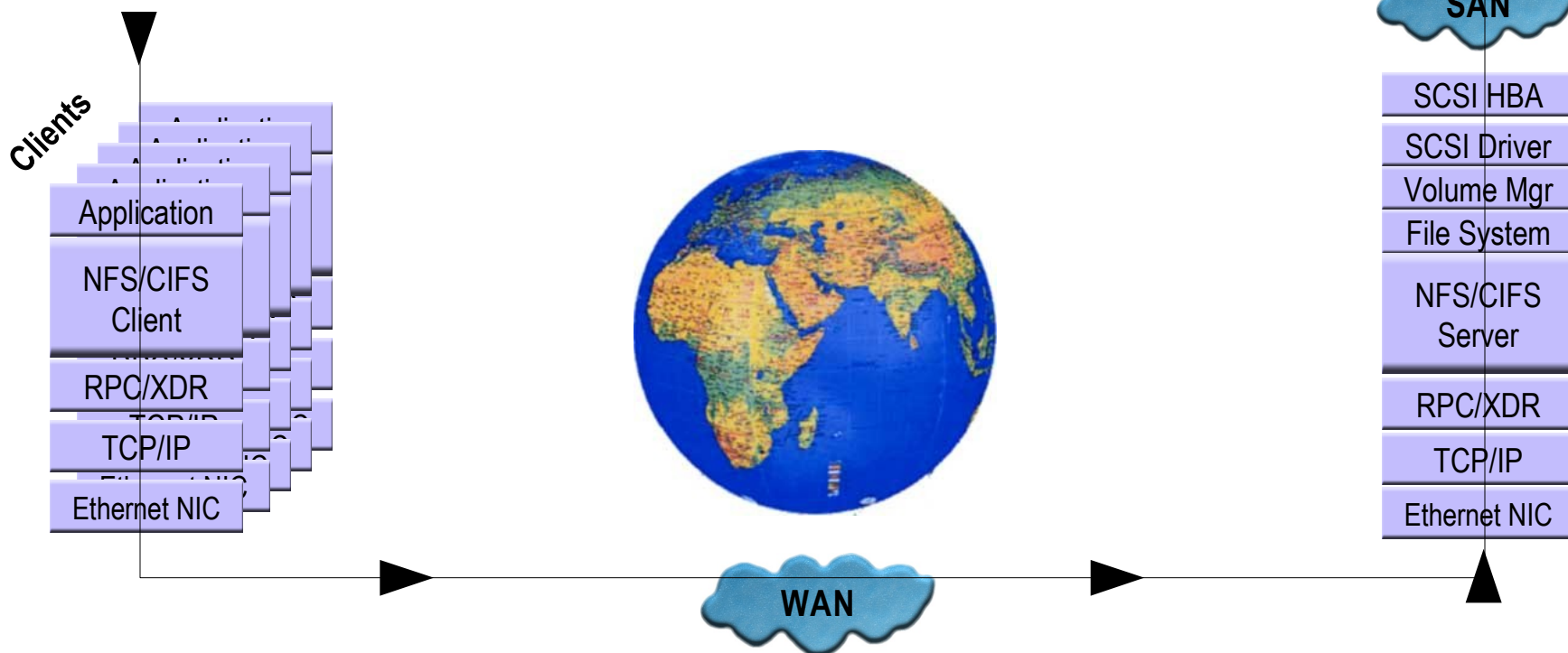
- **WAFS – Wide Area File Services**
- **Data De-Duplication**
- **Storage Virtualization – Thin Provisioning**
- **Solid State Disk (SSD)**
- **Storage Grid/Cluster**
- **Fibre Channel over Ethernet**

Network File System Stack



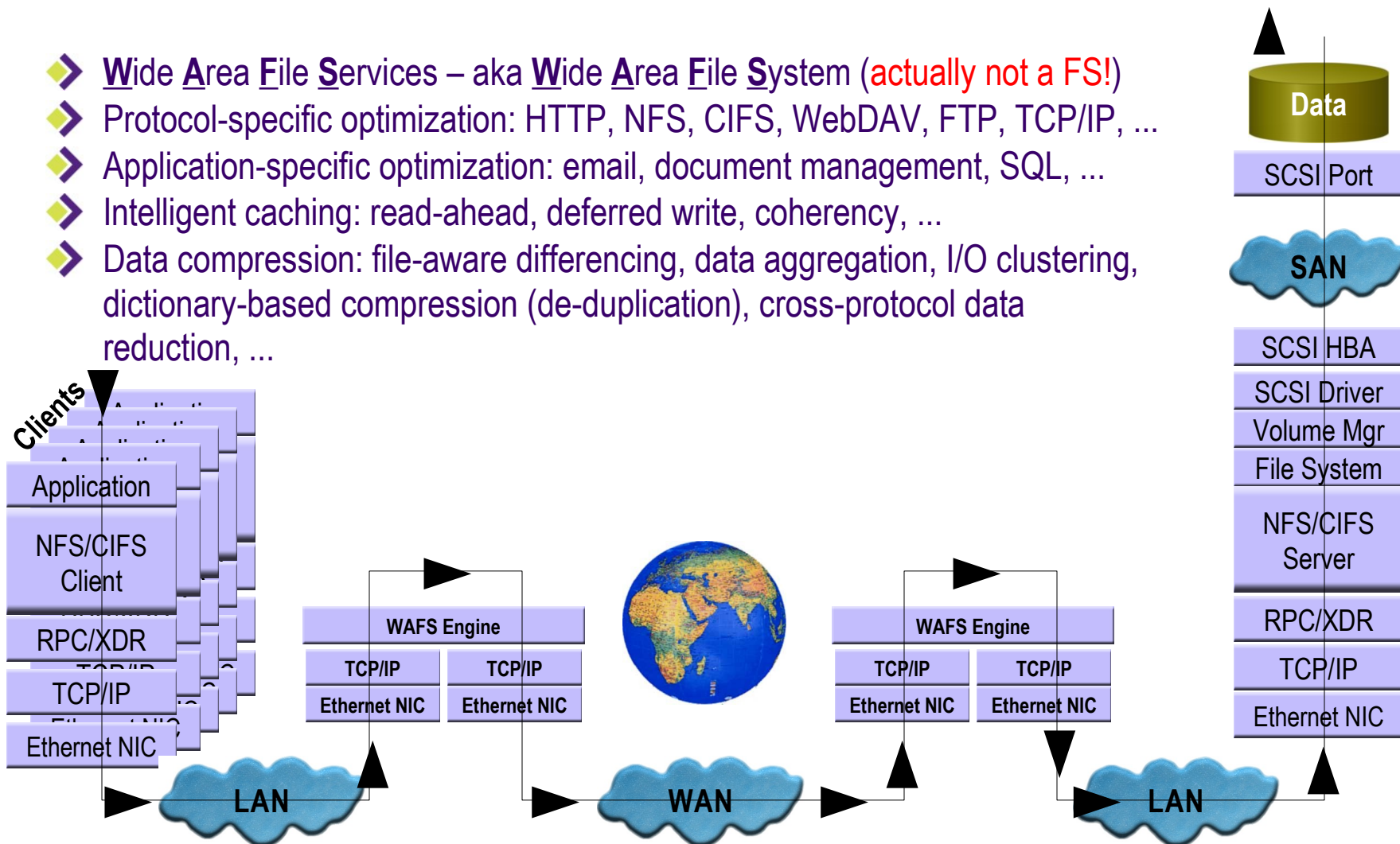
Network FS in a Distributed World

- Consolidating file and storage resources into the data center eases management, administration, cost, and compliance
- Global file sharing and collaboration
- Remote office consolidation and optimization
- **Most application an file access protocols perform poorly over the WAN**



Wide Area File Services

- **Wide Area File Services** – aka **Wide Area File System** (actually not a FS!)
- Protocol-specific optimization: HTTP, NFS, CIFS, WebDAV, FTP, TCP/IP, ...
- Application-specific optimization: email, document management, SQL, ...
- Intelligent caching: read-ahead, deferred write, coherency, ...
- Data compression: file-aware differencing, data aggregation, I/O clustering, dictionary-based compression (de-duplication), cross-protocol data reduction, ...



- ◆ **WAFS – Wide Area File Services**
- ◆ **Data De-Duplication**
- ◆ **Storage Virtualization- Thin Provisioning**
- ◆ **Solid State Disk (SSD)**
- ◆ **Storage Grid/Cluster**
- ◆ **Fibre Channel over Ethernet**

➤ **Compression**

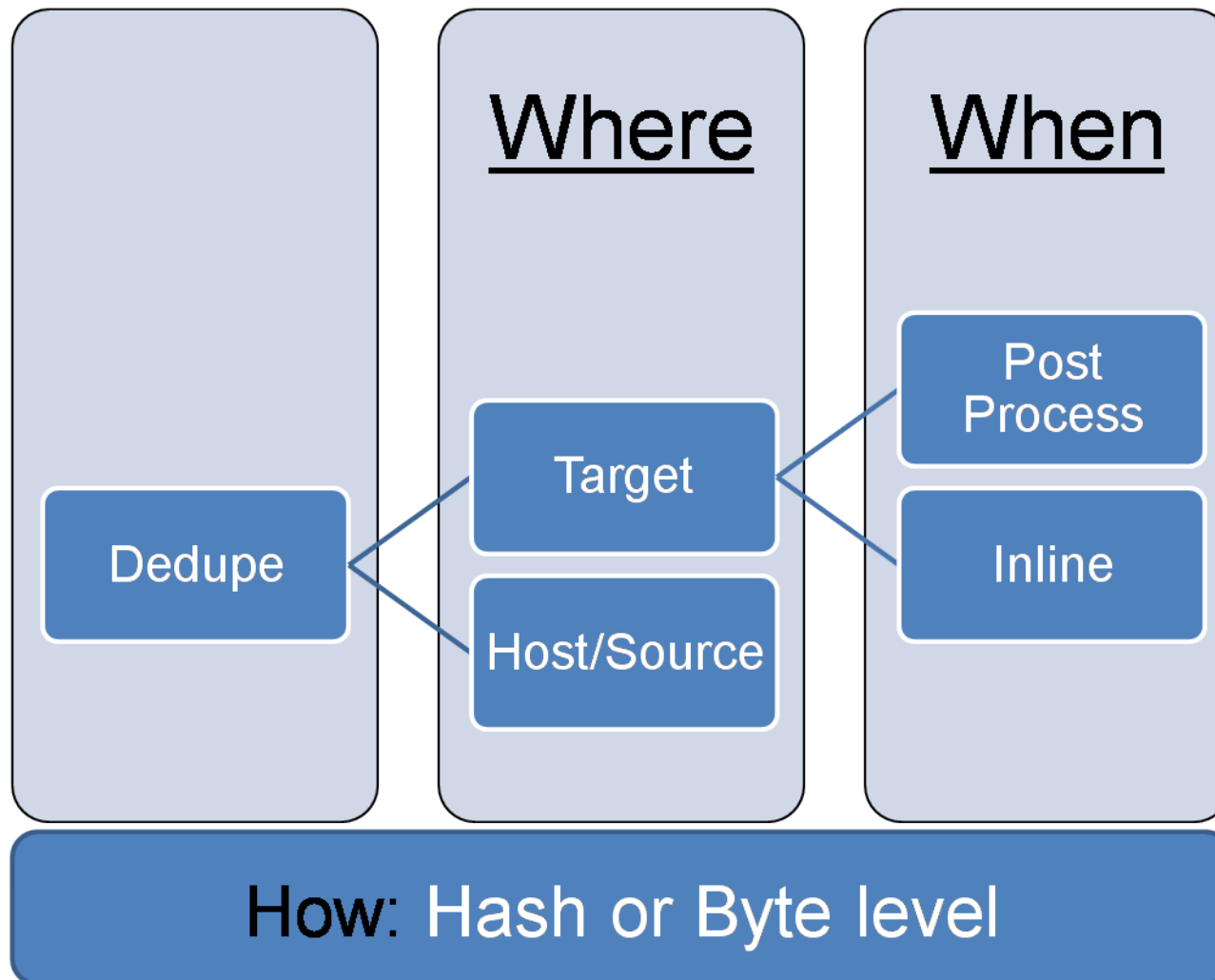
➤ **Single Instancing**

➤ **De-duplication**

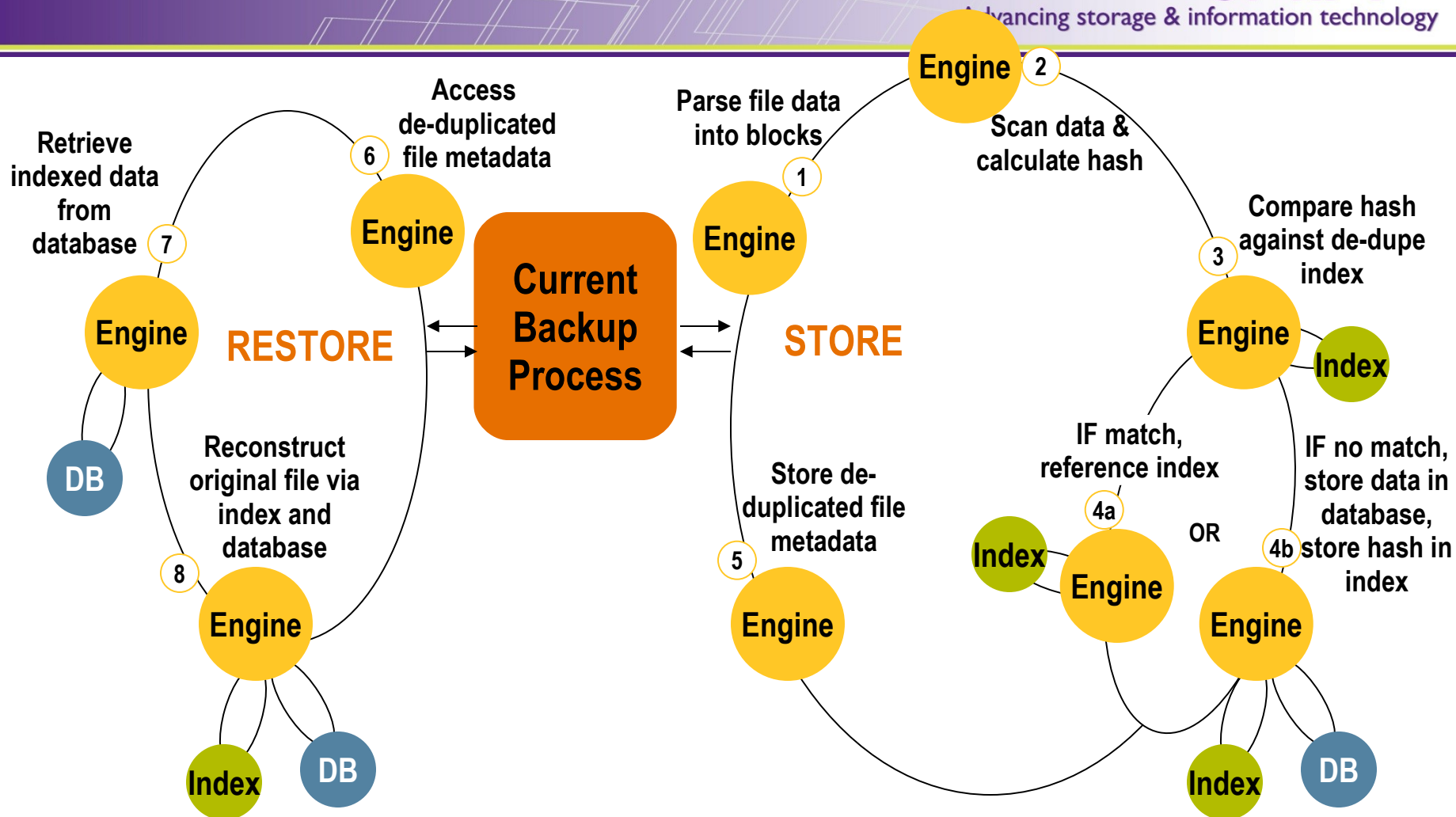
➤ **Attention:**

- ◆ **Capacity optimization has to occur before data encryption**
- ◆ **Single copy for multiple occurrences**
- ◆ **Significant computational overhead**
- ◆ **Claim 1:50 compression depending on data**

De-dupe Technology Approaches



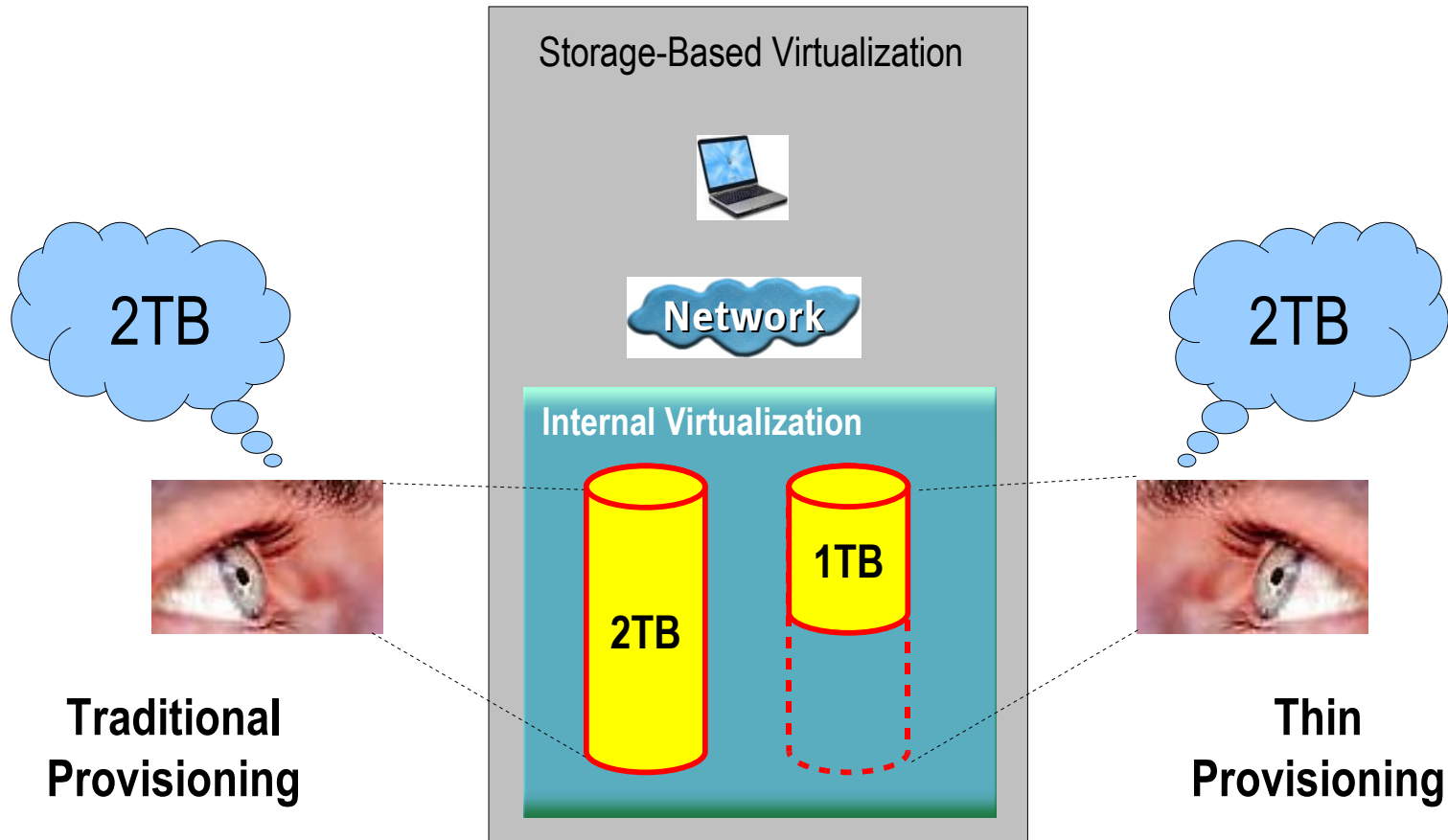
Data De-Duplication – Flow Chart



When implementing de-duplication, consider the effect of adding complexity to backup and restore processes

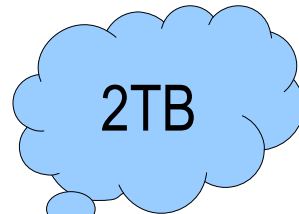
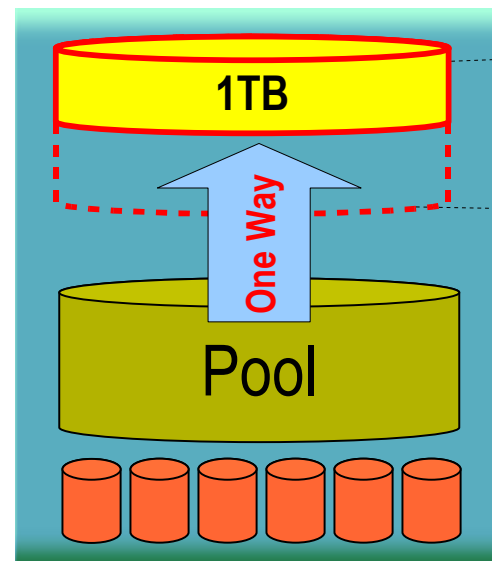
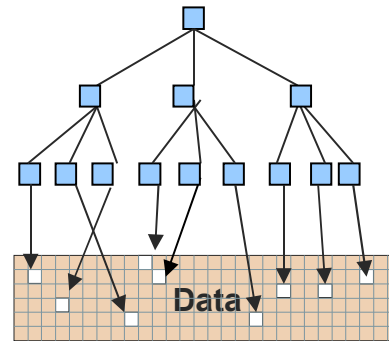
- **WAFS – Wide Area File Services**
- **Data De-Duplication**
- **Storage Virtualization - Thin Provisioning**
- **Solid State Disk (SSD)**
- **Storage Grid/Cluster**
- **Fibre Channel over Ethernet**

Storage Virtualization - Traditional Provisioning vs. Thin Provisioning



Thin Provisioning & Applications (e.g. file systems)

- Filesystem/Application initialization **pre-allocates** physical storage
- Scattered **pre-allocation** requires lots of physical space because it is **allocated-on-write** in fixed size units
- Once written allocated units (i.e. dirty blocks) will leave the pool forever, even if application makes data obsolete (e.g. deletion of a file)
- There is no communication between storage consumer and provider

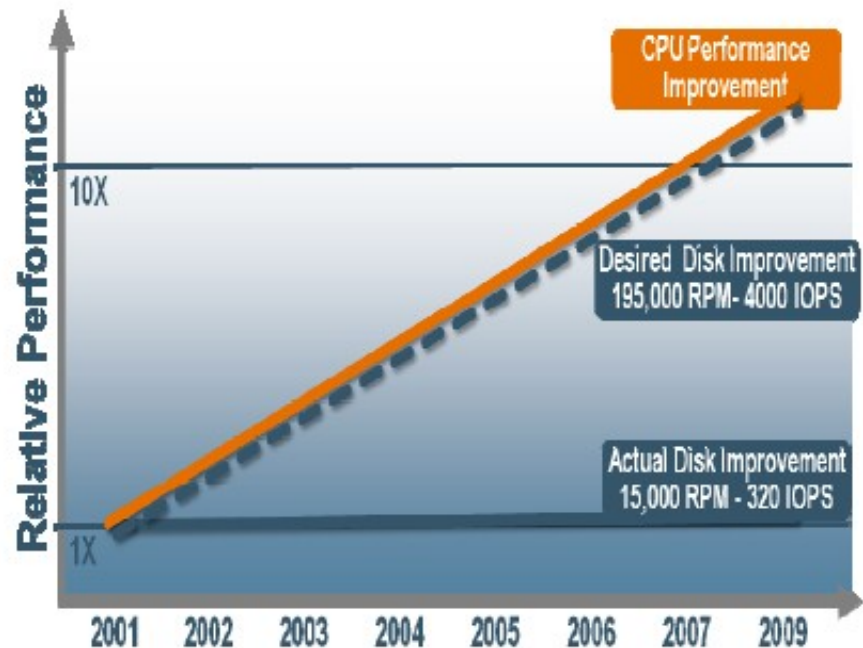


**Thin
Provisioning**

- ◆ **WAFS – Wide Area File Services**
- ◆ **Data De-Duplication**
- ◆ **Storage Virtualization- Thin Provisioning**
- ◆ **Solid State Disk (SSD)**
- ◆ **Storage Grid/Cluster**
- ◆ **Fibre Channel over Ethernet**

Disk to CPU Discontinuity

- ▶ Moore's Law is **outstripping disk drive performance** (rotational speed)
- ▶ As a result, servers and storage systems are **hopelessly unbalanced** between CPU/controller capability and storage pool performance
- ▶ The objective of modern systems design is to **rebalance the CPU-storage ecosystem** while optimizing both low \$/GB and \$/IOPS



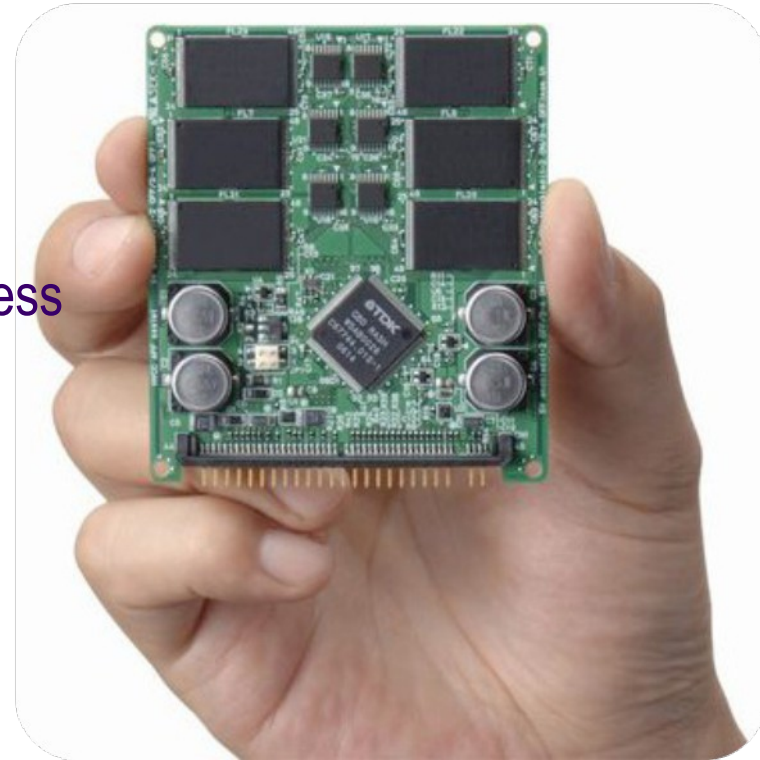
Source: MASS002 Intel Developer Forum 2008

“A solid state drive (SSD) has no moving parts.
It is a semiconductor-based
block storage device that behaves as a
virtual HDD and appears to the host device as a
disk drive.”

What are Solid State Drive's (SSDs)?

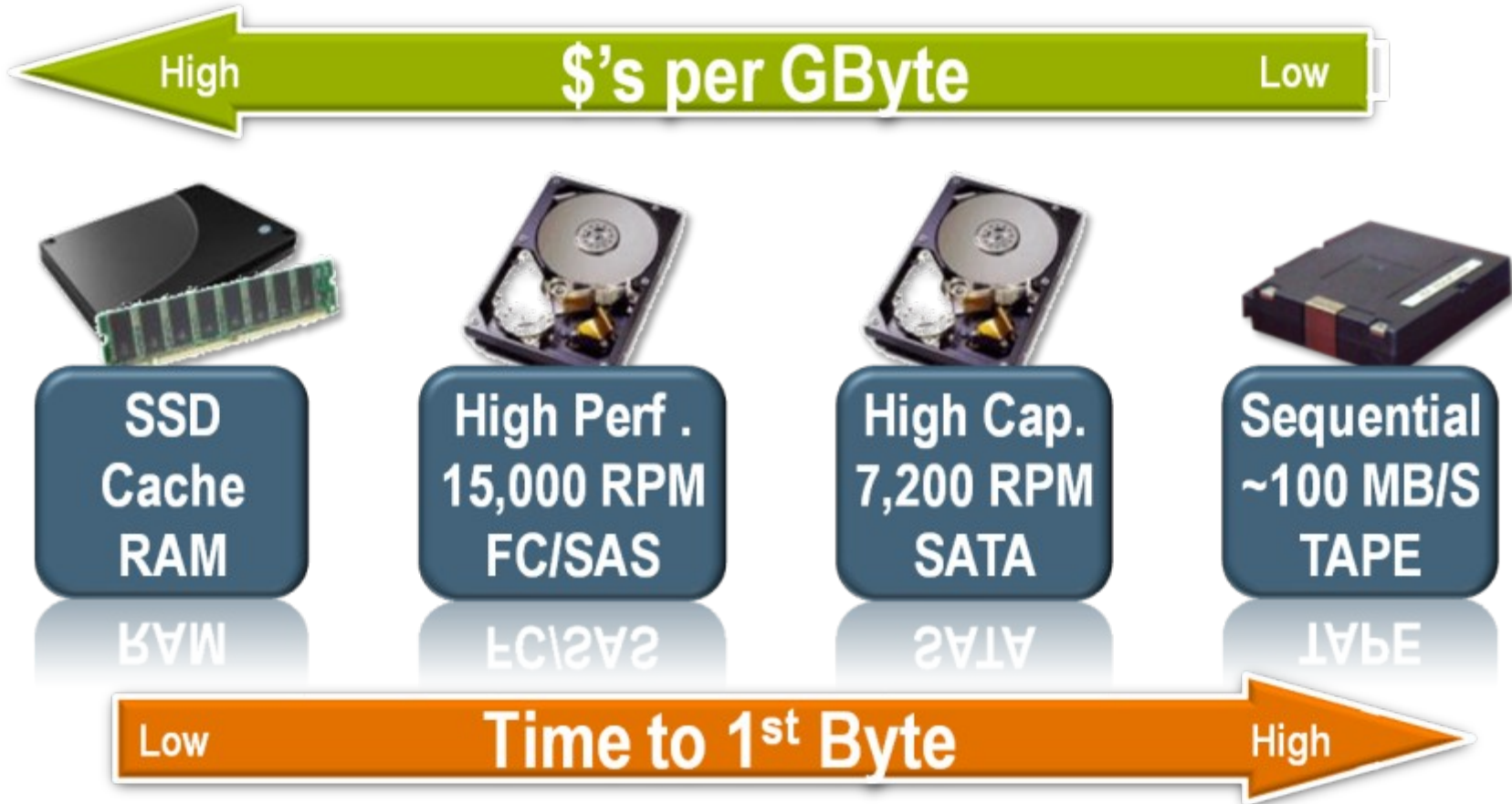
SSD has three major parts:

- A) Controller
- B) DRAM
- C) FLASH bank
- Individual FLASH chips are **pooled** and address space **virtualized** by the controller
- Controller also performs
 - Wear leveling
 - CRC
 - Bad block mapping
- Controller provides the host interface such as SATA, PATA, SAS or FC

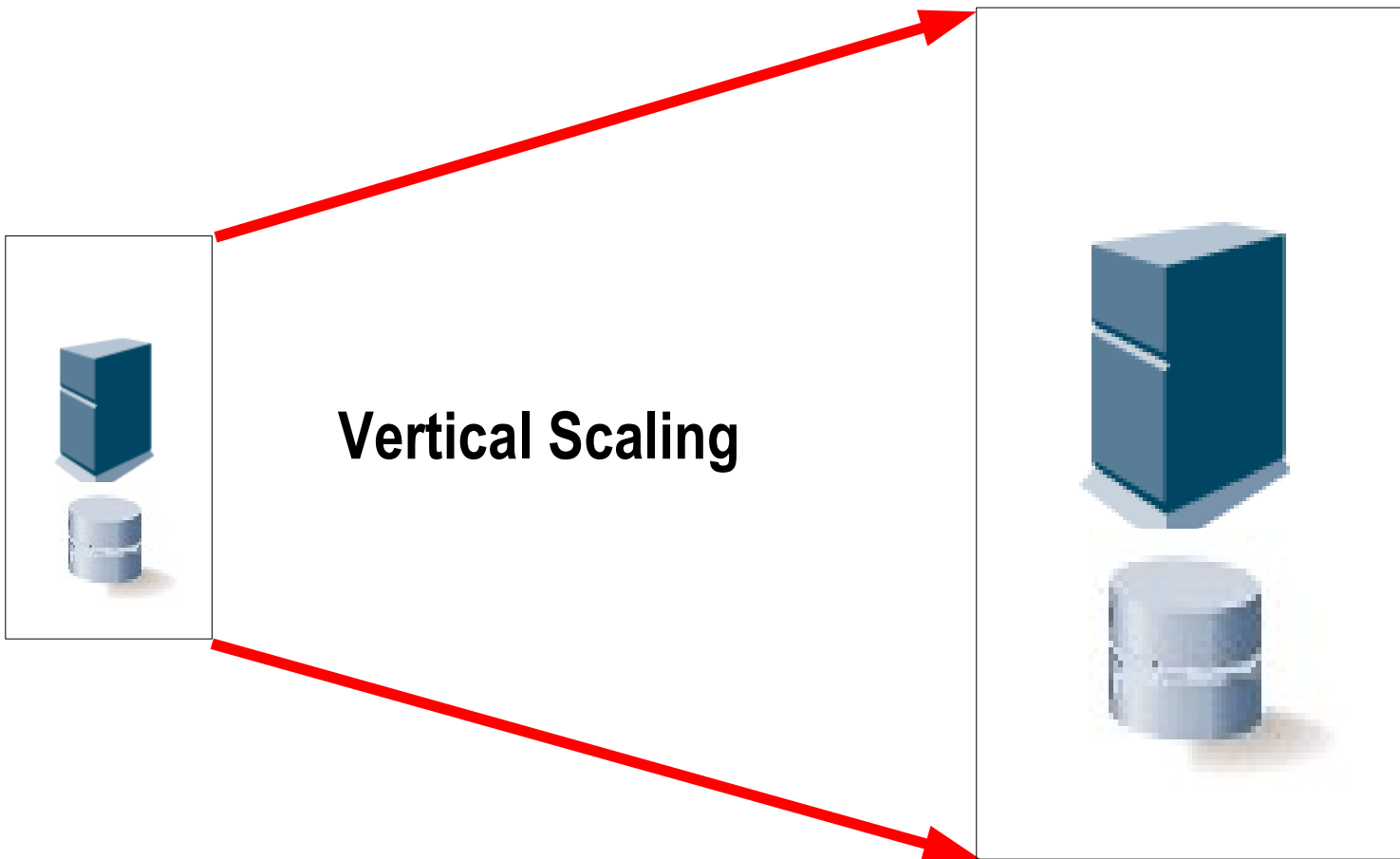


Advantages:
fast random access
50,000 IOPS vs. 250 IOPS
extreme resilient against shock & vibration
low power consumption
lighter
quieter

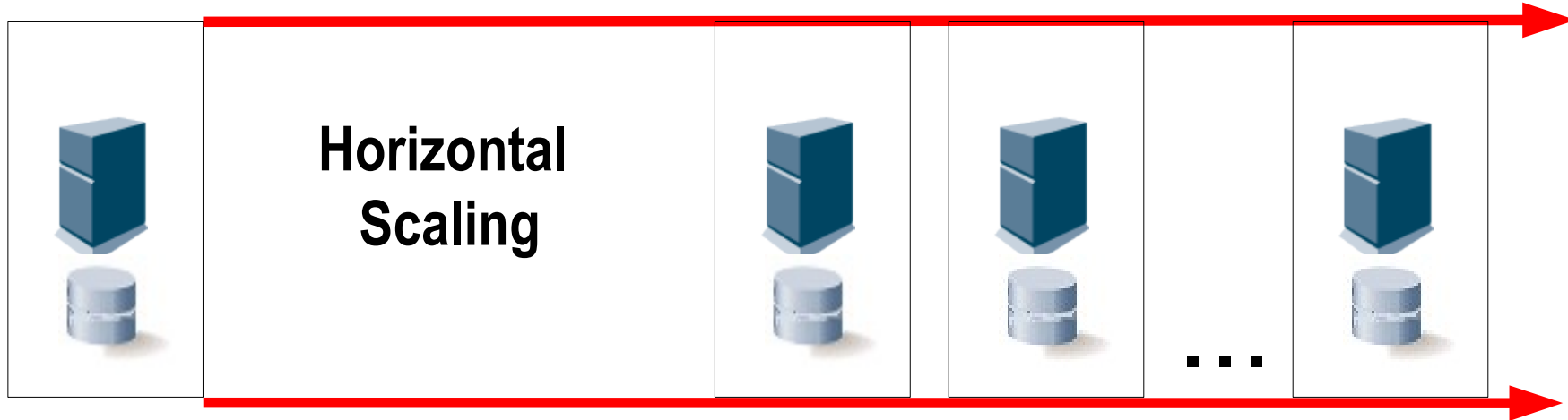
Where to Store Data?



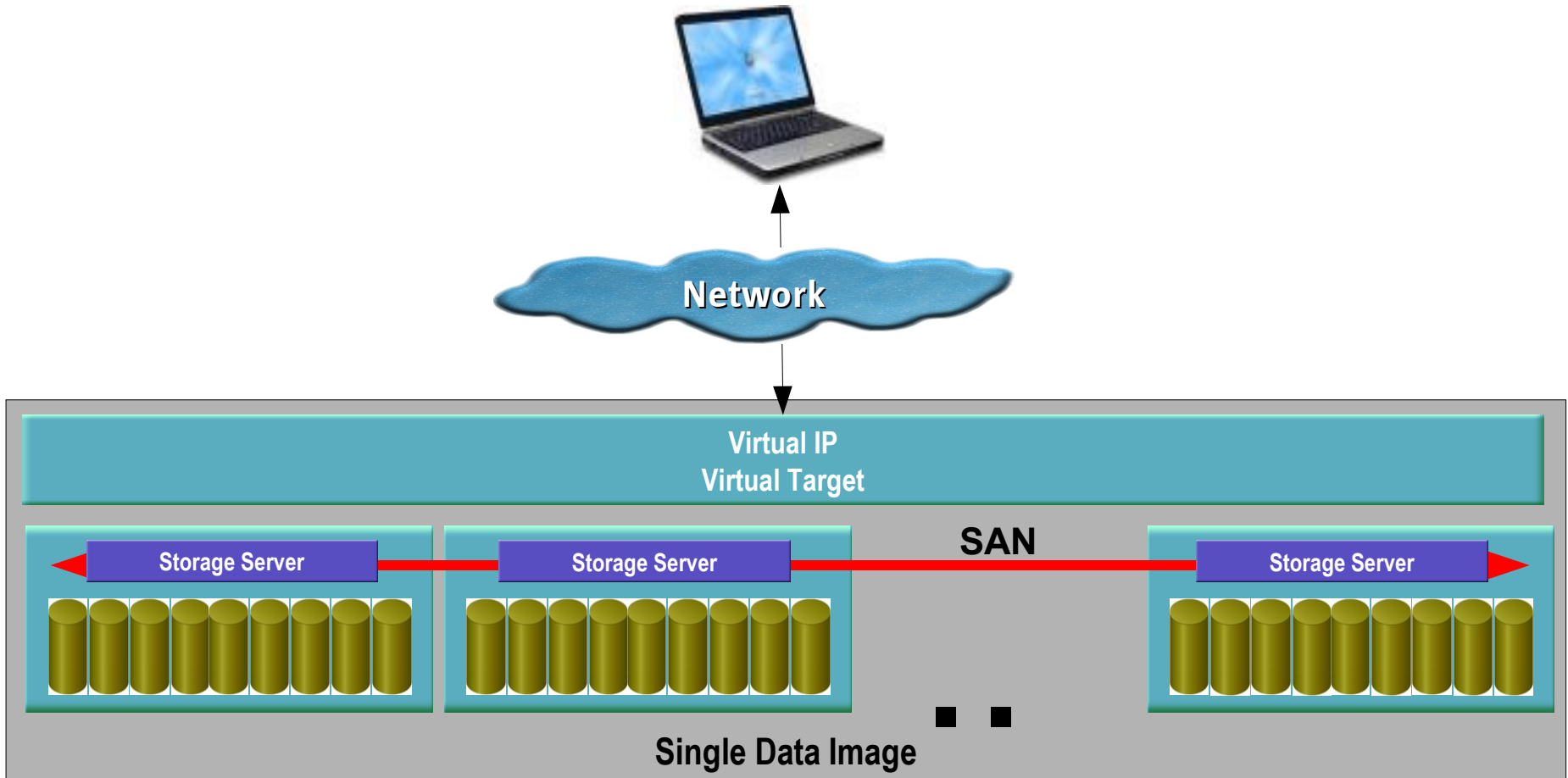
- ▶ **WAFS – Wide Area File Services**
- ▶ **Data De-Duplication**
- ▶ **Storage Virtualization- Thin Provisioning**
- ▶ **Solid State Disk (SSD)**
- ▶ **Storage Grid/Cluster**
- ▶ **Fibre Channel over Ethernet**



Scale-Out

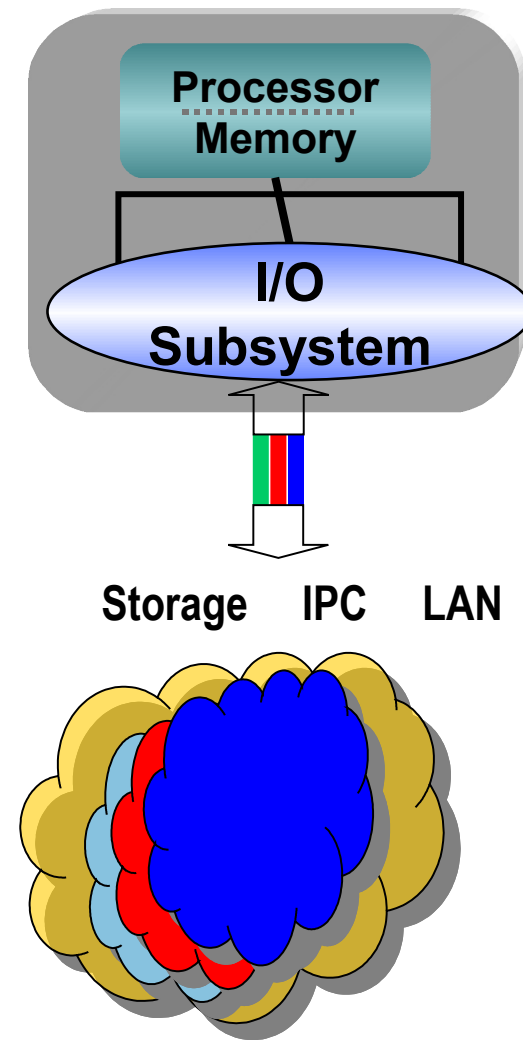
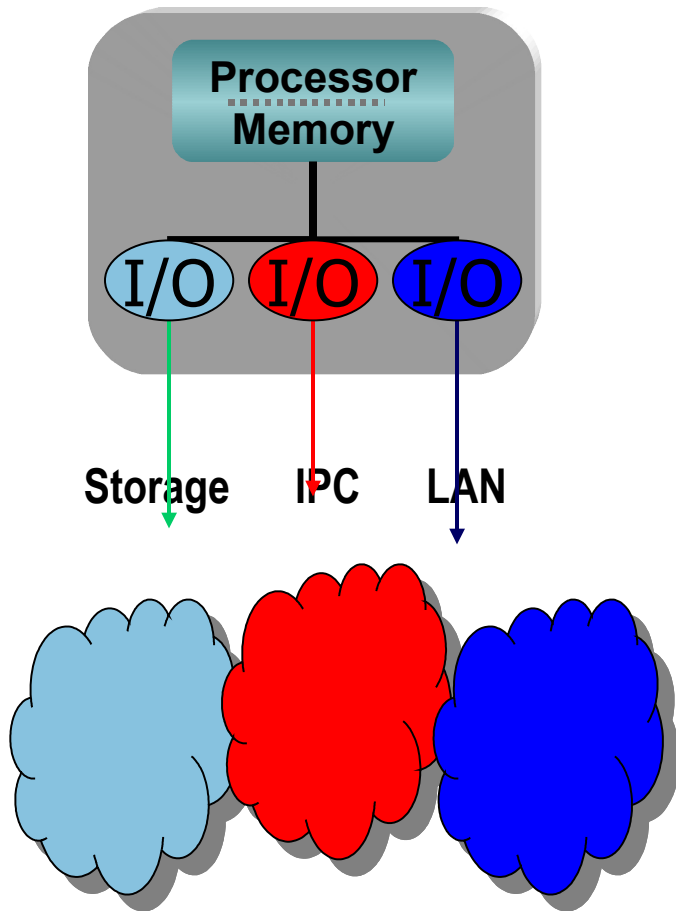


Storage Cluster/Grid Horizontal Scaling / Scale-Out

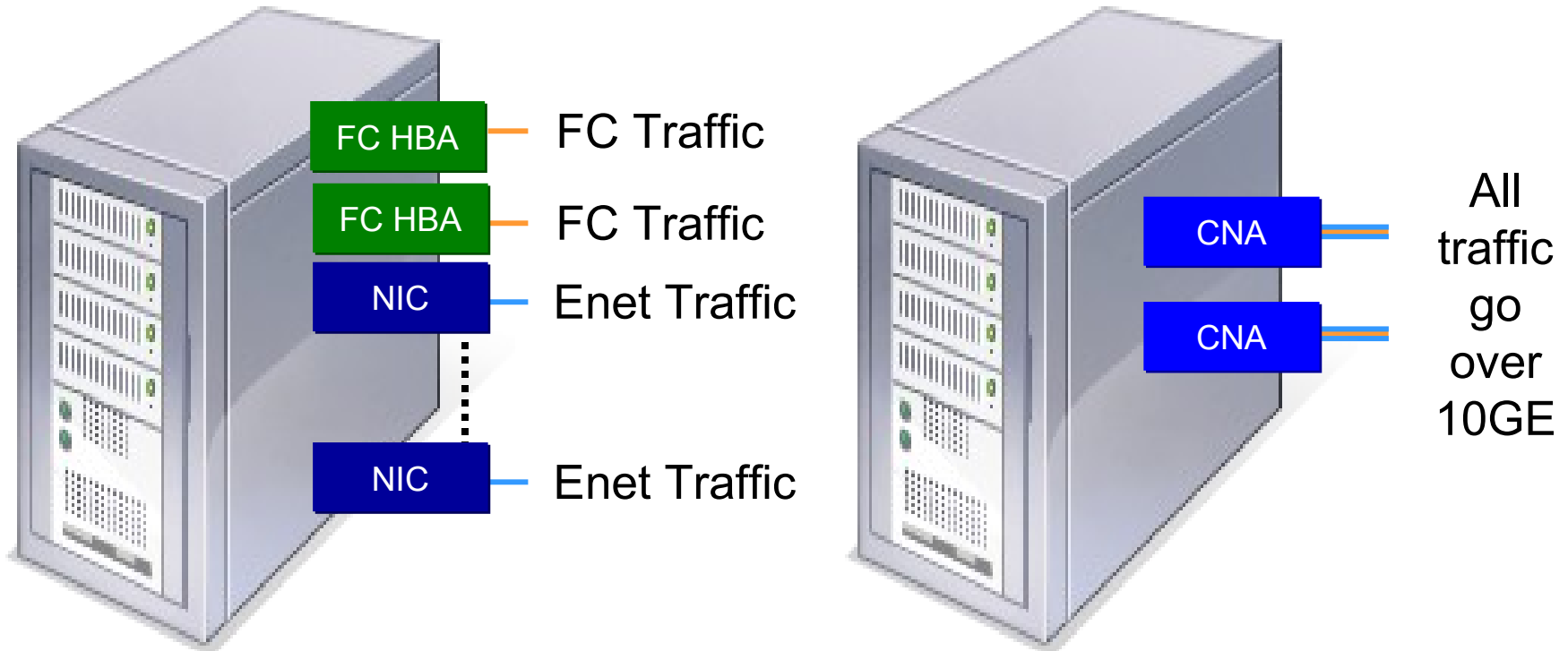


- ▶ **WAFS – Wide Area File Services**
- ▶ **Data De-Duplication**
- ▶ **Storage Virtualization- Thin Provisioning**
- ▶ **Solid State Disk (SSD)**
- ▶ **Storage Grid/Cluster**
- ▶ **Fibre Channel over Ethernet**

I/O consolidation

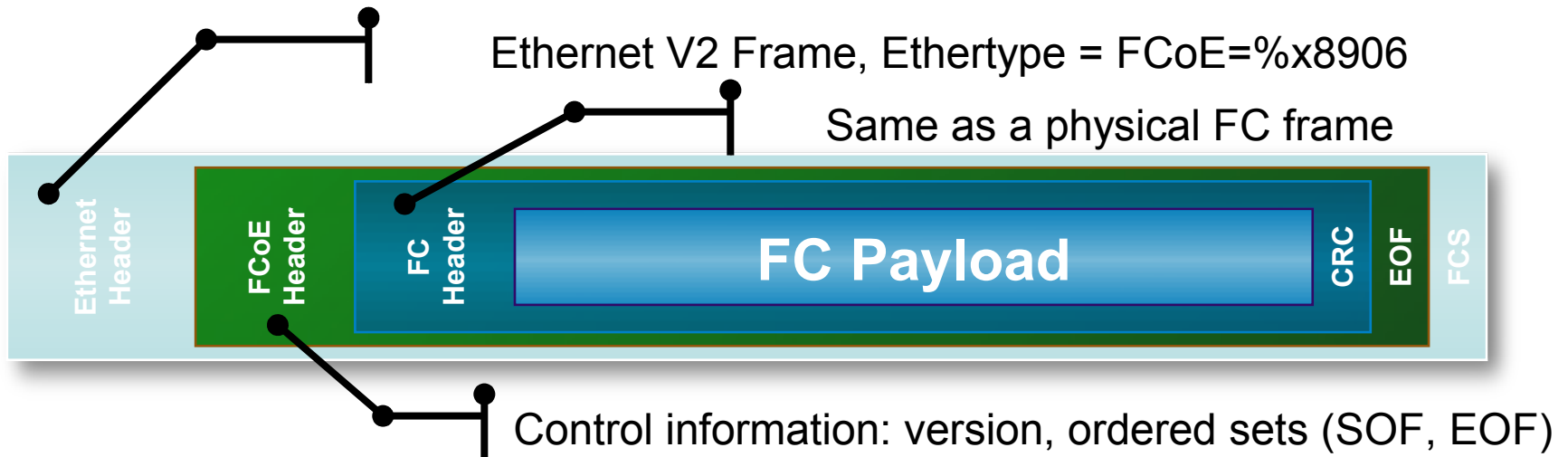


I/O Consolidation Benefits



- ◆ **Adaptor:** NIC for Ethernet/IP, HCA for InfiniBand, Converged Network Adaptor (CNA) for FCoE
- ◆ **Customer Benefit:** Fewer NIC's, HBA's and cables, lower CapEx, OpEx (power, cooling)

- **10 GbE**
- **Data Center Ethernet (Lossless Ethernet)**
- **Ethernet jumbo frames (2180 bytes)**
 - ◆ **Max FC frame payload (2112 bytes)**

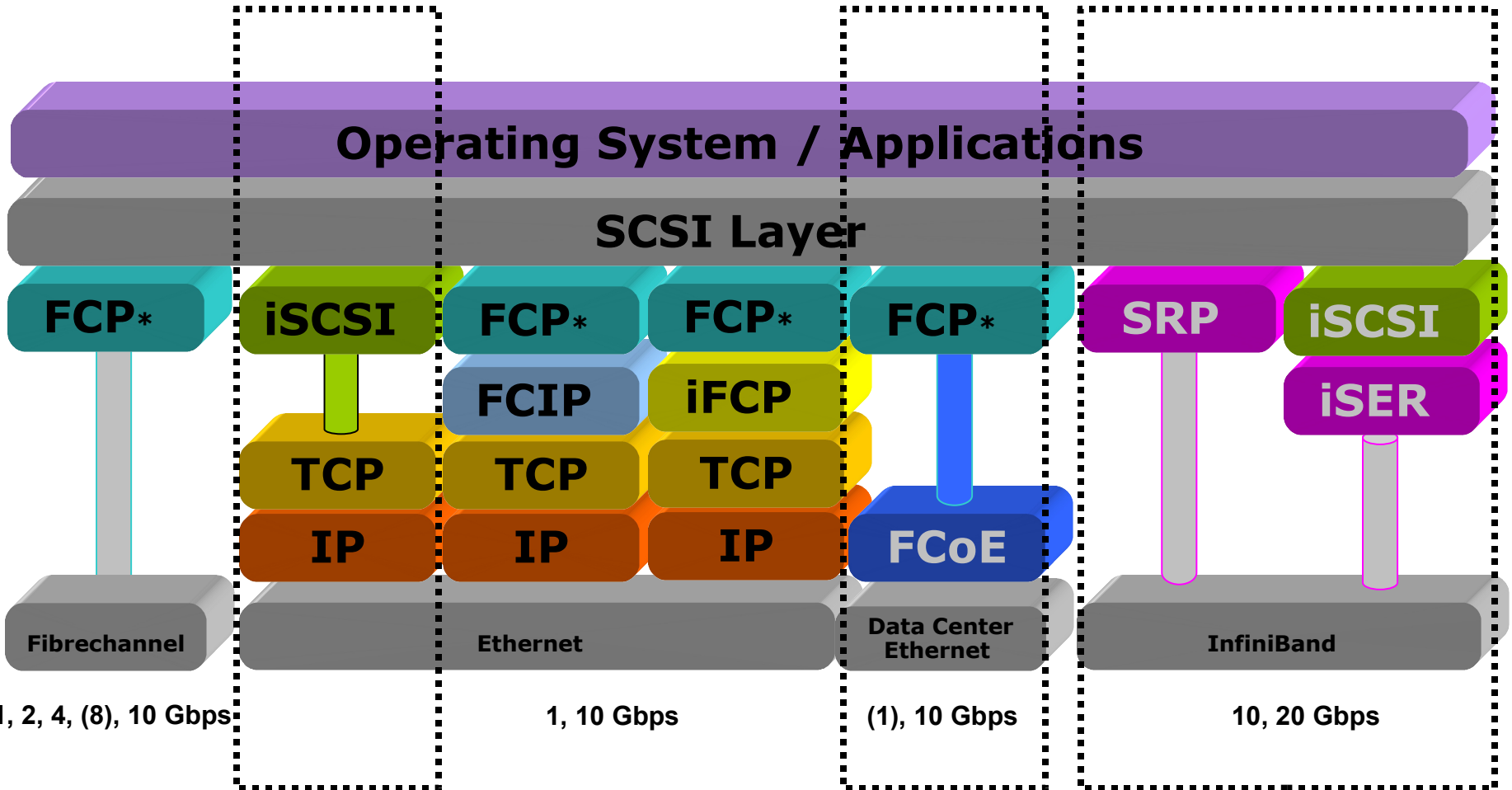


Encapsulation Technologies

iSCSI

FCoE

InfiniBand



* Includes FC Layer

Trends bei Speichertechnologien

3. BITKOM-Anwenderforum IT-Infrastruktur

Frankfurt, 27. November 2008

Christian Bandulet

Principal Engineer

Sun Microsystems

