



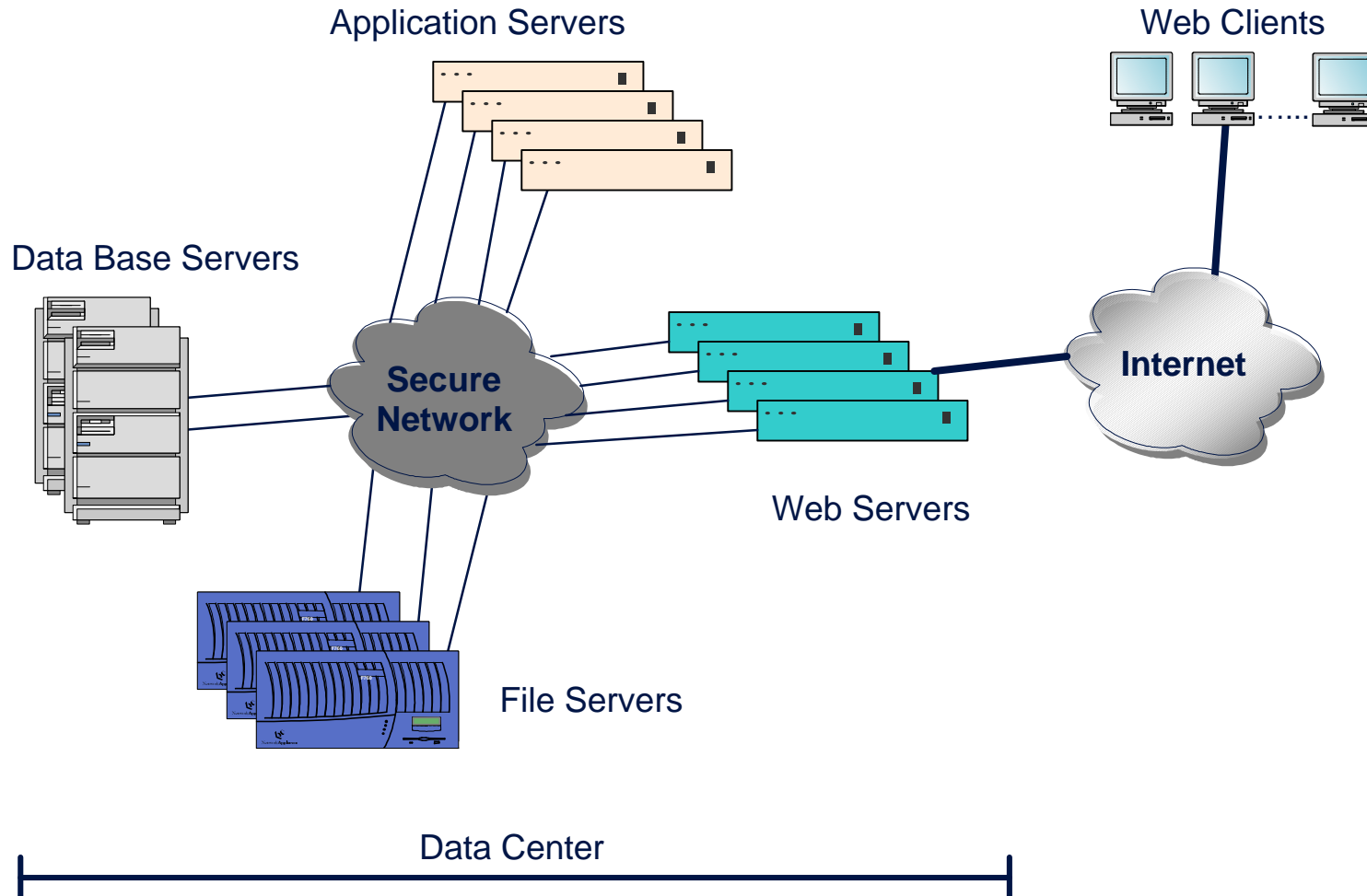
# **An Introduction to the Direct Access File System (DAFS) Protocol**

# Outline



- **Enabling e-Business Data Centers**
- **Local File-sharing Environment**
- **VI and VI-like Interconnects**
- **Direct Access File System (DAFS) Protocol**
- **DAFS Implementation Examples**
- **The DAFS Collaborative**
- **Summary**

# e-Business Data Centers



# Data Center Requirements

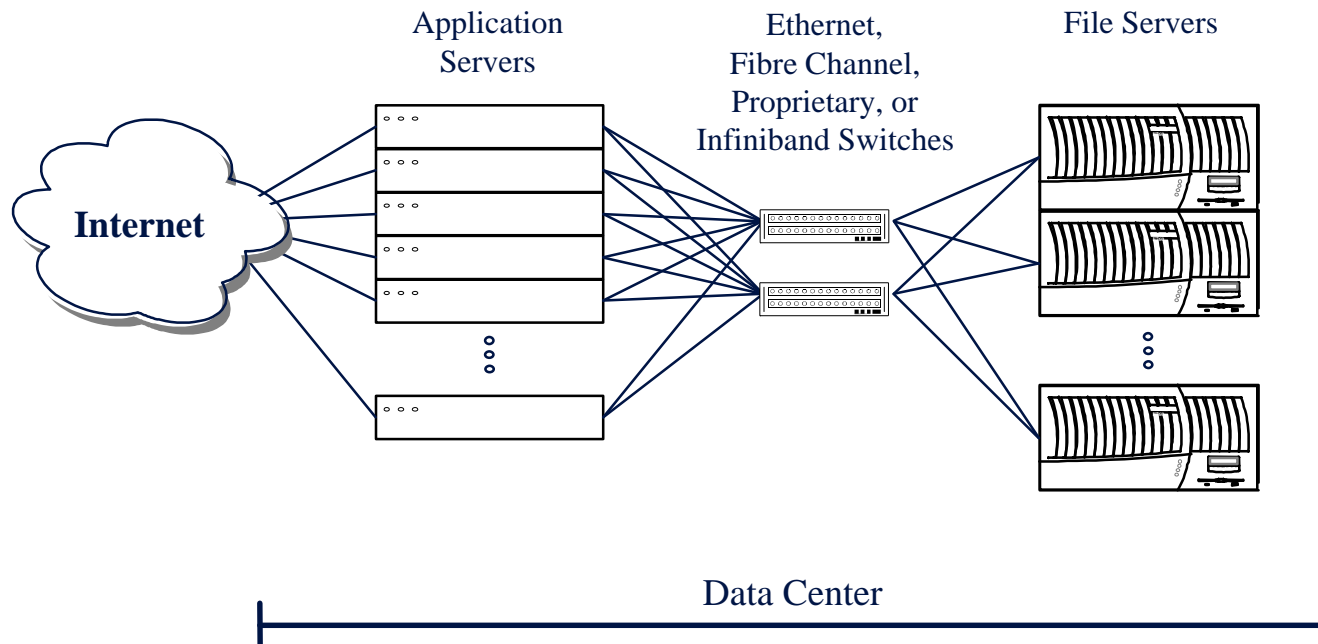


- **Scalability**
  - Independently scalable compute and storage
- **Fast Data Transfer**
  - Low latency; high throughput
- **Resiliency**
  - Application server and file server fail-over
- **Cost**
  - Heterogeneous, standards-based hardware
- **Ease of Management**

# “Local File-Sharing” Defined



- **Geographically constrained**
  - 1 or 2 machine rooms
- **Mostly homogeneous clients**
  - Can be large or small
  - 1 -100 machines
- **Single administrative control**
  - “Trusted environment”
- **High performance applications**
  - Web services; email, etc.
  - Database, GIS

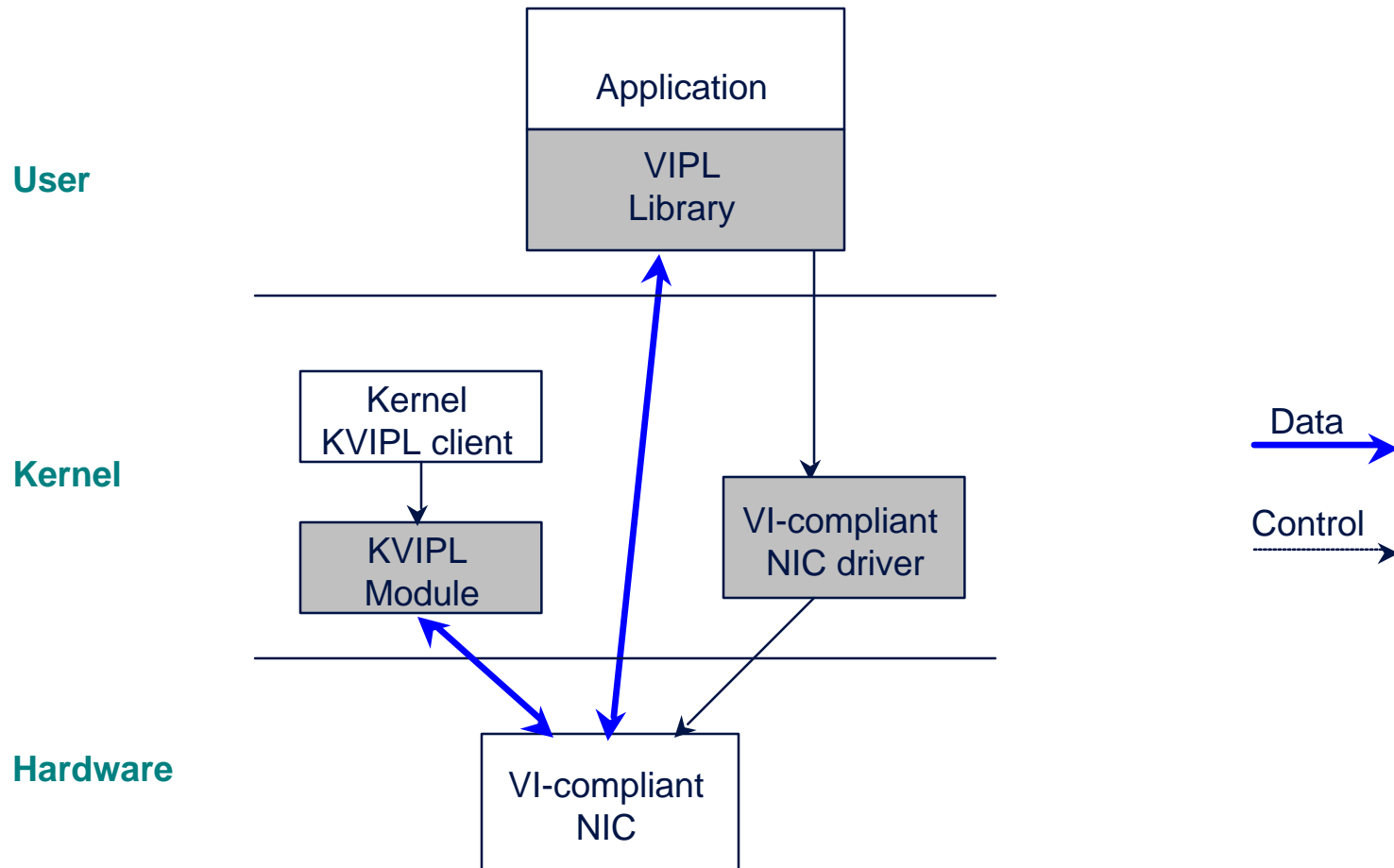


# Virtual Interface Technology



- **Virtual Interface (VI) Architecture**
  - Cluster interconnect architecture defined by Intel, Compaq and Microsoft
  - VI Developer's Forum
- **Standard capabilities**
  - Send/receive message, remote DMA read/write
  - Multiple channels with send/completion queues
  - Data transfer bypasses kernel
- **Transport Independence**
  - VI over TCP/IP (IETF)
  - VI over Fibre channel (ANSI FC-VI draft std)
  - InfiniBand (VIPL 2.0)

# VI Architecture Example



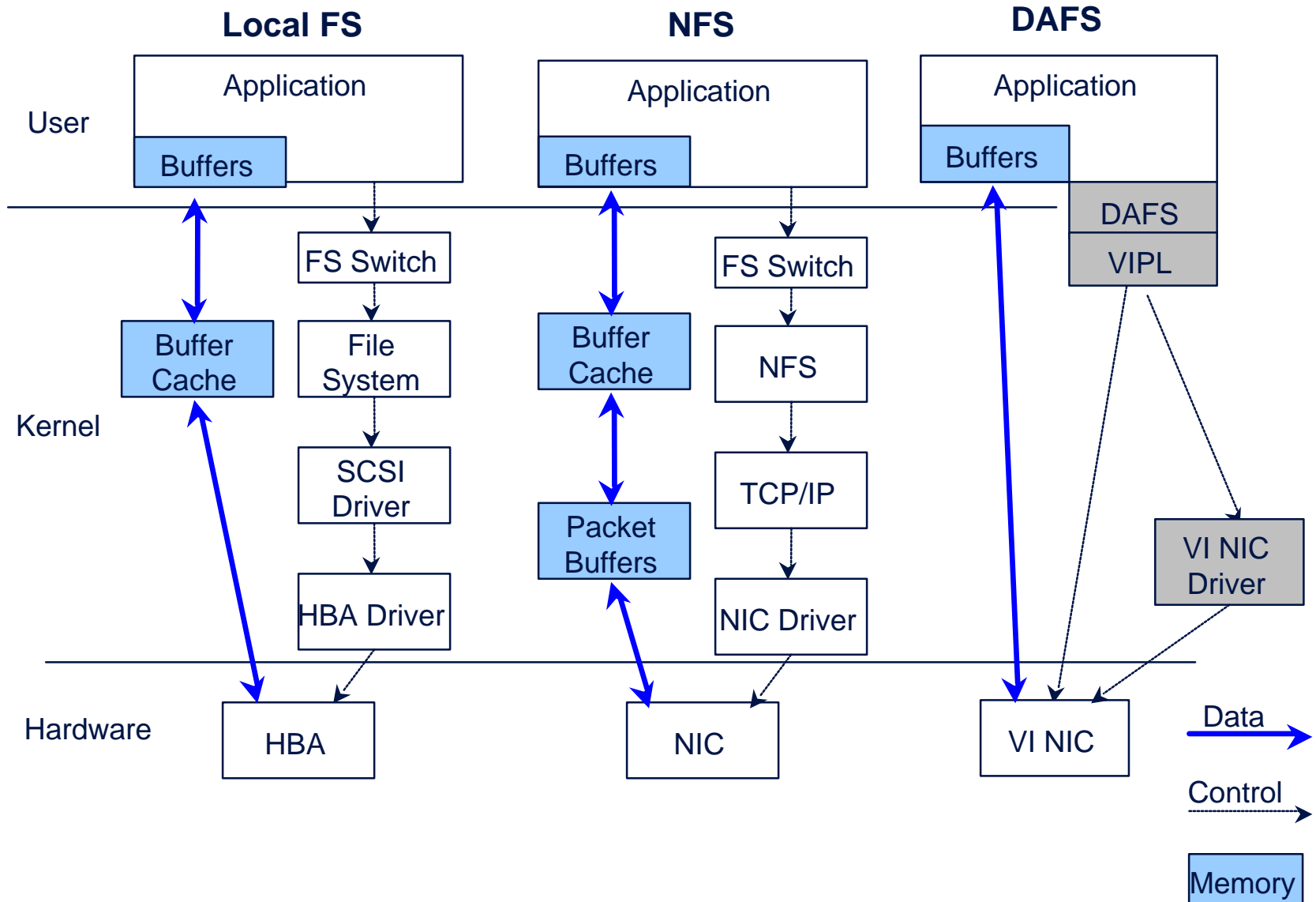
\* VI Provider Layer specification maintained by the VI Developers Forum

# Direct Access File System Protocol



- **Lightweight file access protocol with implicit native data sharing**
- **High performance**
  - Designed for low latency, low packet loss networks
- **Leverages VI for transport independence**
- **Direct application access to transport resources**
  - File data transfers directly to application buffers
  - Bypasses Operating System
- **Leverages NFSv4 Semantics**

# File Access Methods



# DAFS and NFS



- **NFS**
  - Maps file access semantics onto standard IP networking capabilities
  - Wide area and local file sharing
- **DAFS**
  - Maps file access semantics onto VI capabilities
  - Local file sharing focus

# DAFS Performance vs. NFS over VI



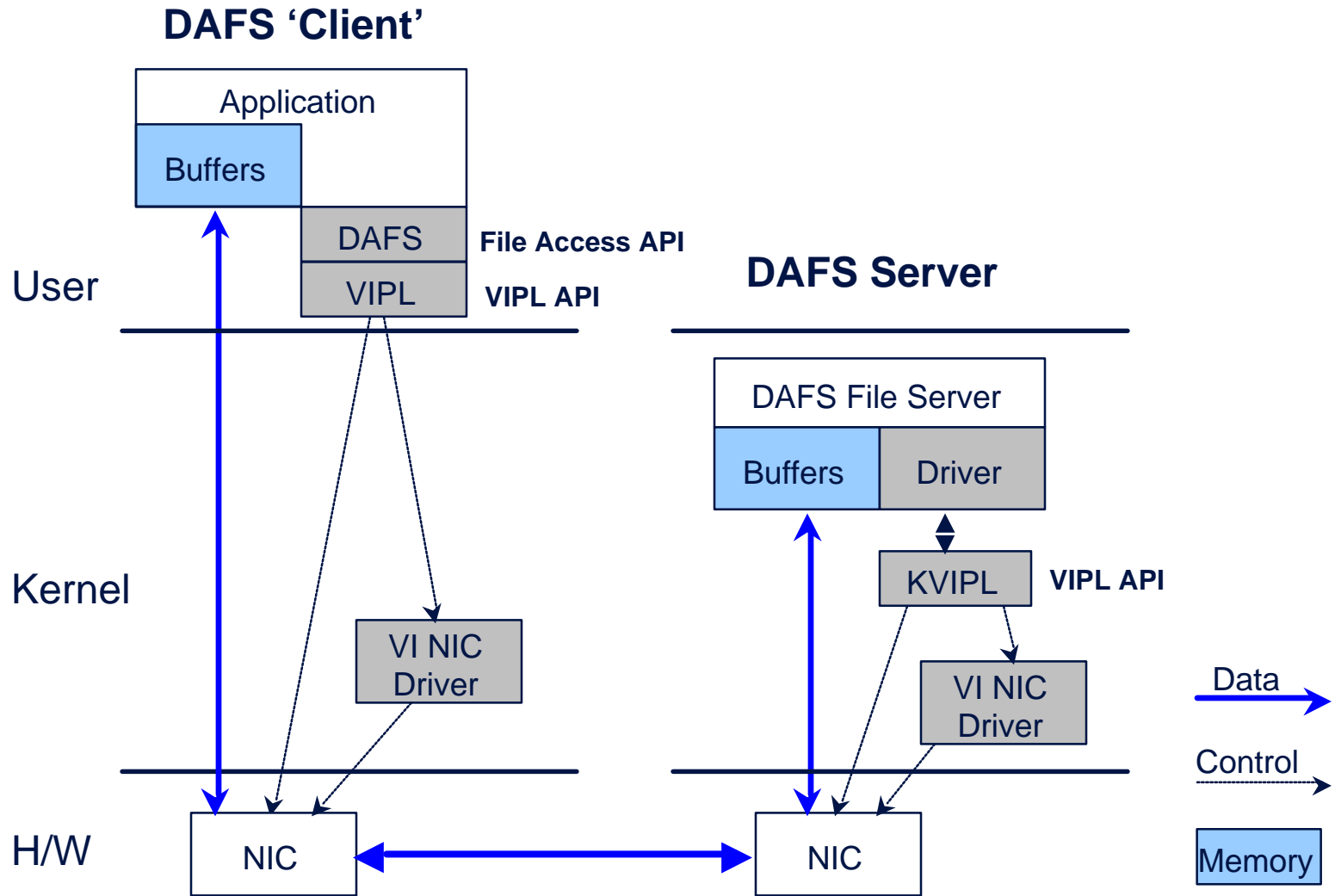
Implementation	Client CPU msec/op
User NFS over VI/sockets prototype	<b>90</b>
Raw access to direct-attached storage	<b>40</b>
User DAFS prototype	<b>25</b>

# DAFS Improvements over NFSv4



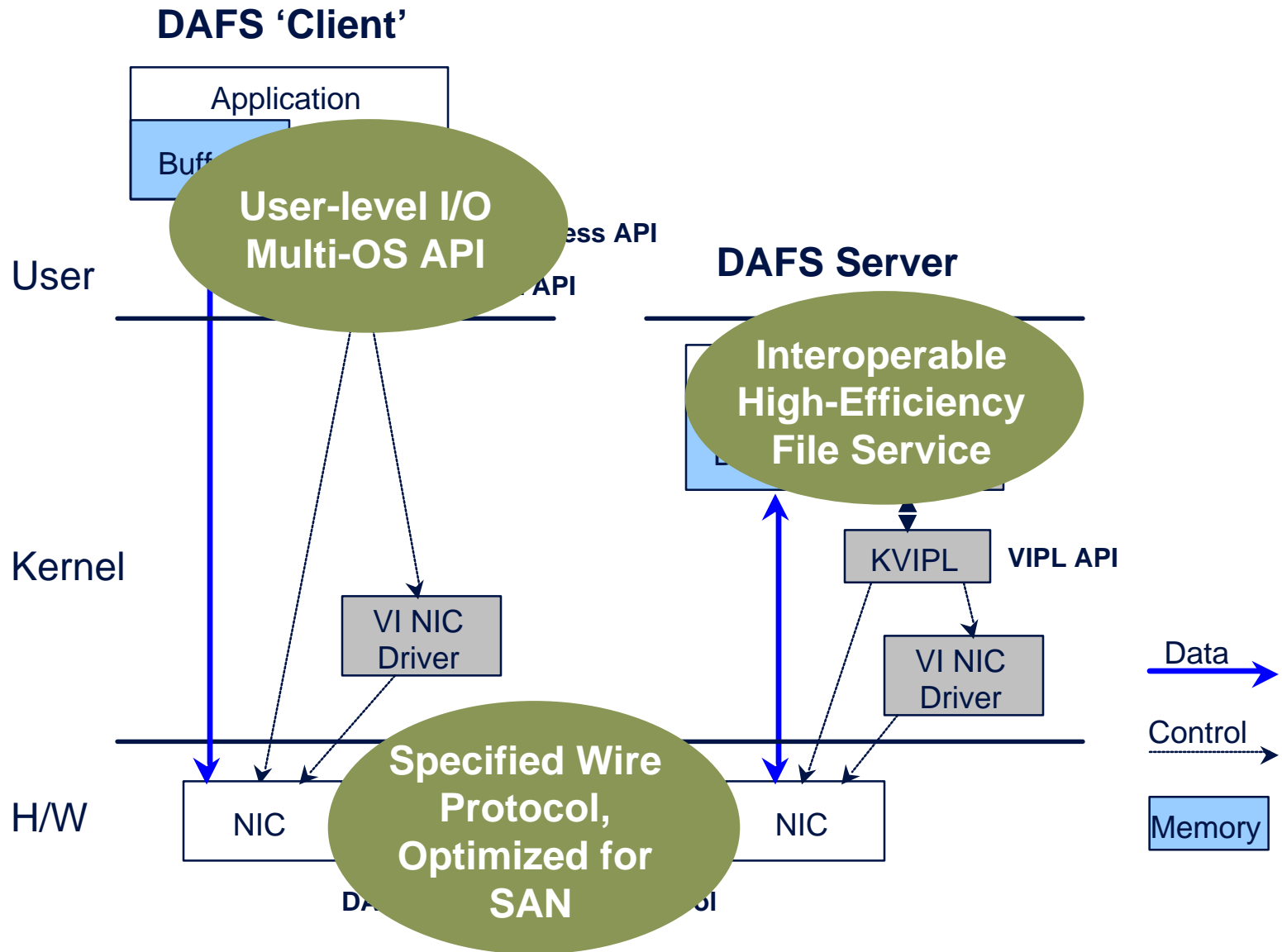
- **Designed to run over VI-capable interconnects**
- **Optimized for high performance**
  - Remote DMA
  - Chaining
  - Application-tunable read-ahead, batch write, cache control
- **Semantics improvements for local file sharing environment**
  - Consistent high-speed locking
  - Graceful client and server fail-over and data migration
  - Fencing

# DAFS Session Diagram

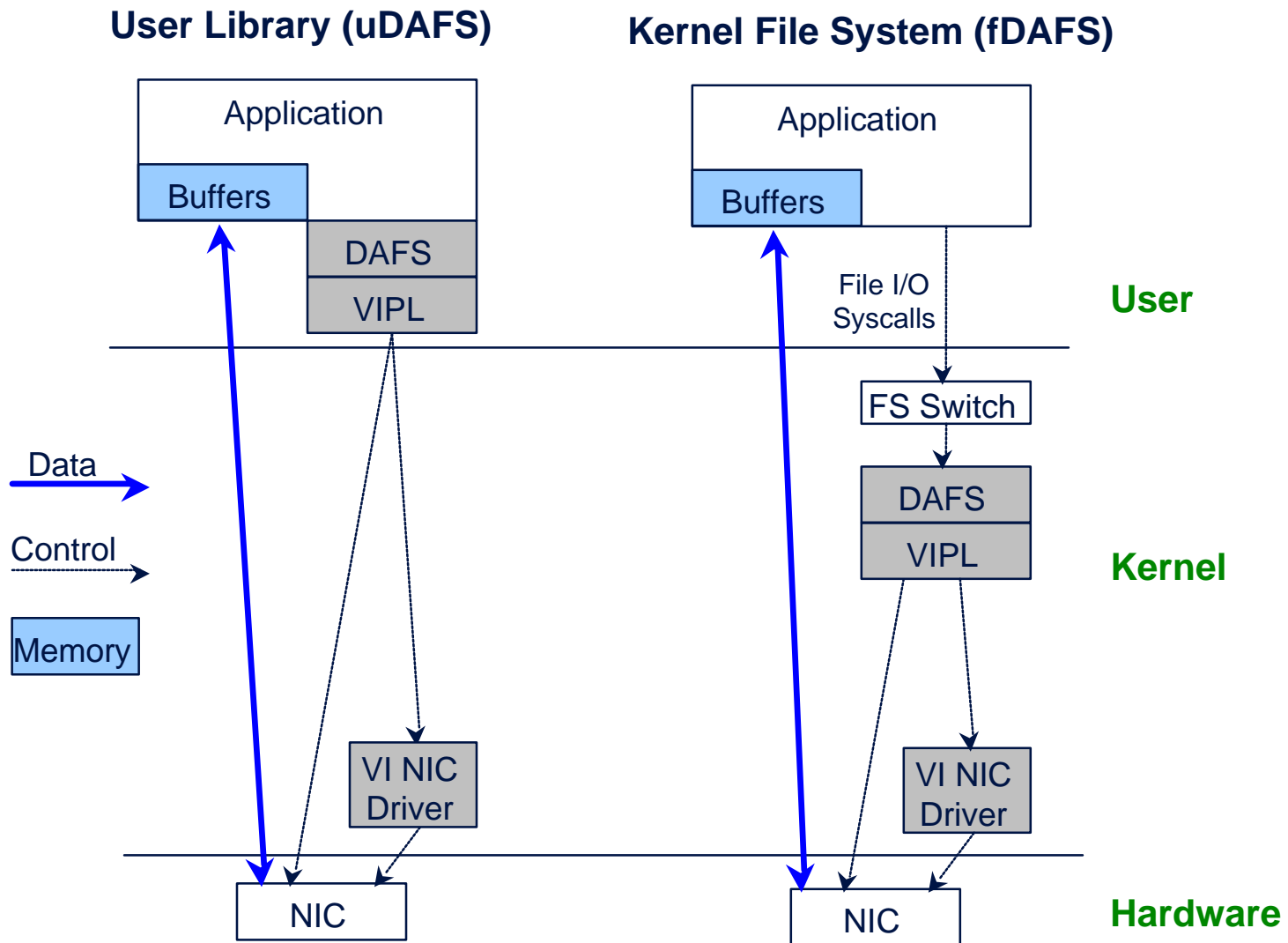


DAFS Wire Format & Protocol

# DAFS Session Diagram



# Implementation Approaches



# uDAFS Client



- **Application invokes uDAFS library for file access**
- **Benefits**
  - Revolutionary performance
  - Gets the OS out of the way
  - True asynchronous file I/O on all OS platforms
  - Direct application access to DAFS Local File Sharing enhancements (e.g., fencing, locking)
- **Cost**
  - Application must be modified to take advantage of portable OS-independent APIs

# fDAFS Client



- **DAFS implemented as an IFS/VFS**
  - Peer to NTFS,CIFS,UFS,NFS
- **Benefits**
  - No application changes (file I/O interface)
  - Performance comparable to direct-attached storage
- **Cost**
  - No opportunity to eliminate kernel transition overhead
  - Asynchronous I/O capability limited by OS architecture
  - No access to DAFS “Local File Sharing” enhancements (e.g., fencing, locking) that are not architecturally supported by the OS

# DAFS Collaborative



- **Hosted by Intel and Network Appliance**
- **Charter:**
  - Drive DAFS specification
  - Submit to standards body Q3CY01
  - Encourage broad industry acceptance
  - Ensure open-source implementations
- **Open to all**
  - New participants welcome
- ***[www.DAFScollaborative.org](http://www.DAFScollaborative.org)***

# DAFS Collaborative Contributors



Adaptec  
Agilent  
ATTO technologies  
Bakbone Software  
Banderacom  
Berg Software Design  
Cisco  
Ciprico  
CMD Technology  
Compaq Computer Corp  
ComVault Systems  
Delphi Engineering Group  
Earthlink Network Inc  
ECCS  
Emphora  
Eurologic Systems  
Fujitsu Ltd

Gadzoox Networks  
Giganet  
Hewlett Packard  
IBM  
Infocruiser  
Intel  
LSI Logic Storage Systems  
NEC Corp  
Network Appliance  
Network Engines  
Network Storage Solutions  
Open Source Asia  
Pirus Networks  
Procom Technology  
Qlogic  
Rhapsody Networks  
SAN Experts Facility

SANgate Systems  
Seagate Technology  
Seek Systems  
Sendmail  
SINTESI  
Solution-Soft  
Spinnaker Networks  
Toigo Productions  
Troika Networks  
Update Systems  
VALinux  
Veltrek  
VERITAS Software  
Viathan  
Vieo Inc  
Voltaire  
Zerowait Corp

# DAFS Protocol Benefits



- **High-performance file access protocol with implicit native data sharing**
- **Optimized for high throughput and low latency**
- **Leverages VI for transport independence**
- **Direct application access**
  - File data transfers directly to application buffers
  - Bypasses Operating System
- **Consistent high speed locking**
- **Graceful recovery/failover of clients and servers**
- **Fencing**

> **Enabling Technology for the Data Center of the Future**