NFS INDUSTRY CONFERENCE

September 22-24

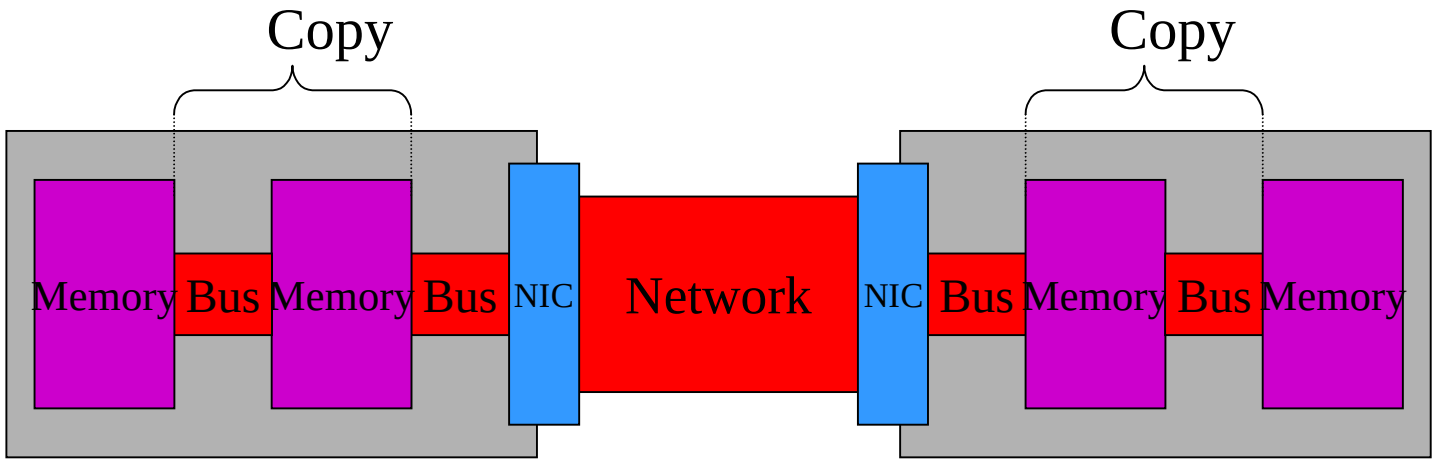
NFS over RDMA

Sweet Spot

Brent Callaghan
Sun Microsystems, Inc.

NFS INDUSTRY CONFERENCE

Network vs Bus Performance

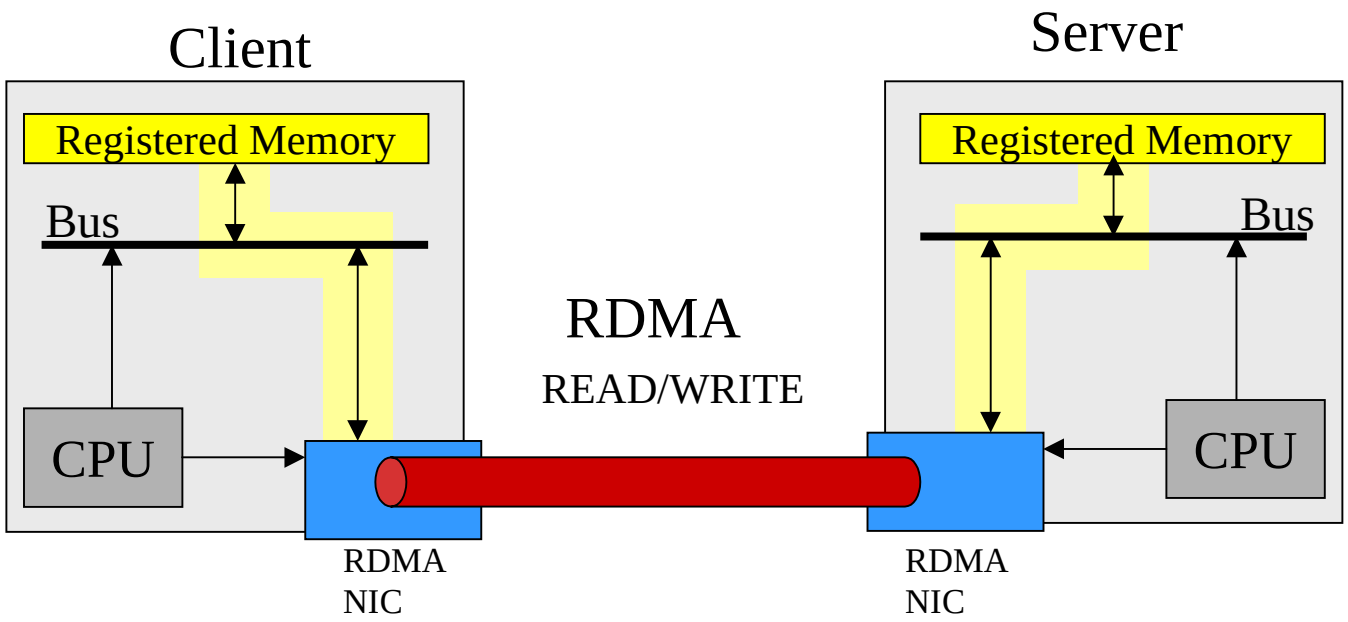


Latency gets worse with each memory copy which further loads the CPU.

NFS INDUSTRY CONFERENCE

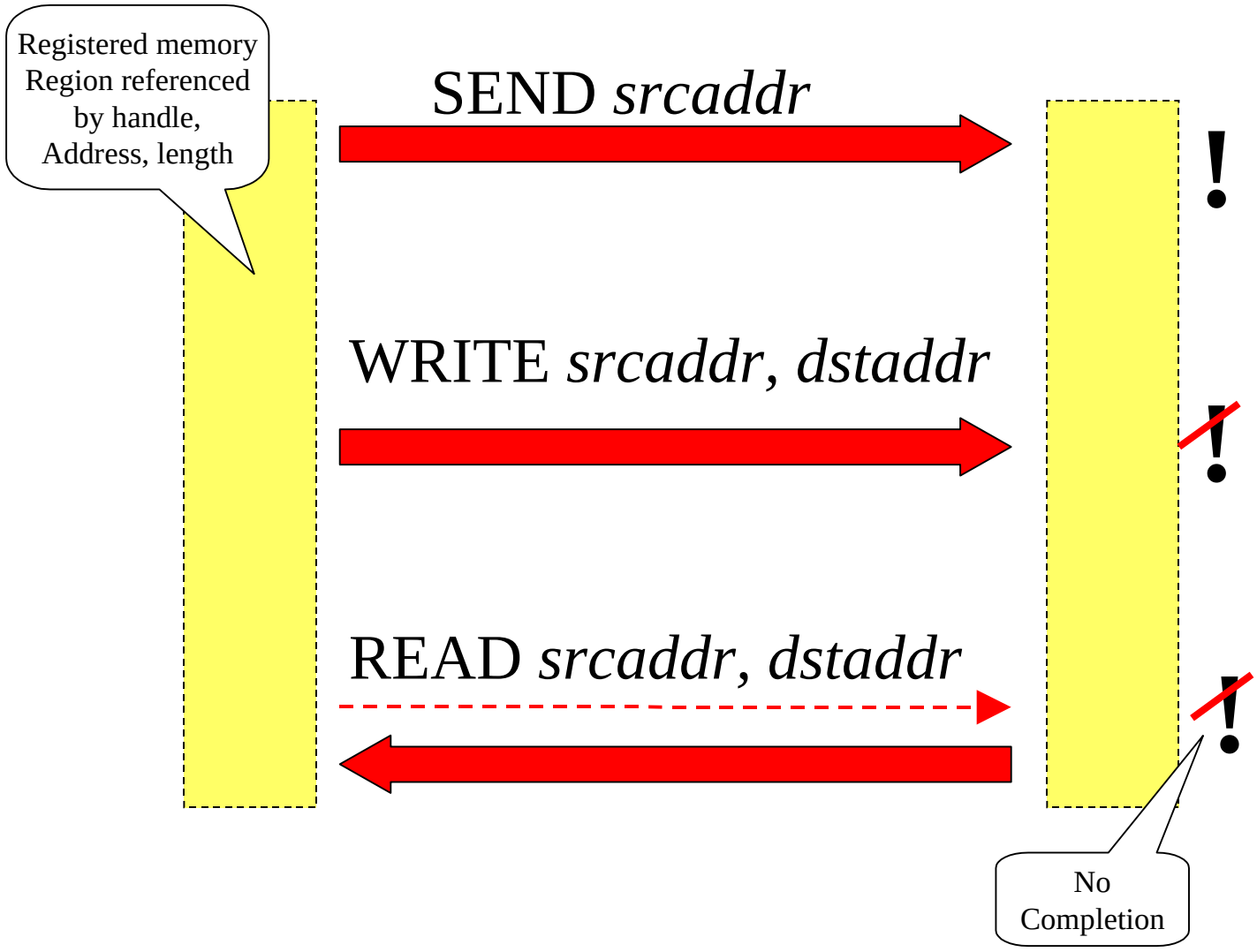
What is RDMA ?

- DMA: Direct Memory Access
- RDMA: *Remote* Direct Memory Access
- Supports Direct Placement
- Networking offload for CPU



NFS INDUSTRY CONFERENCE

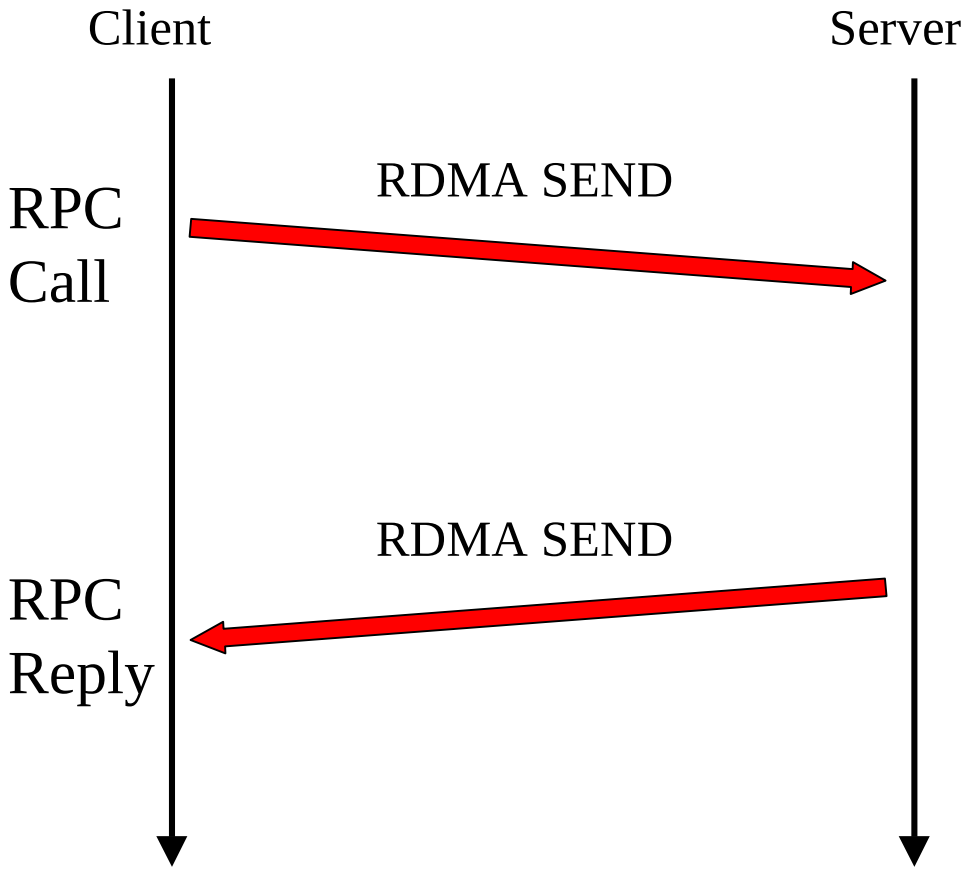
RDMA Operations



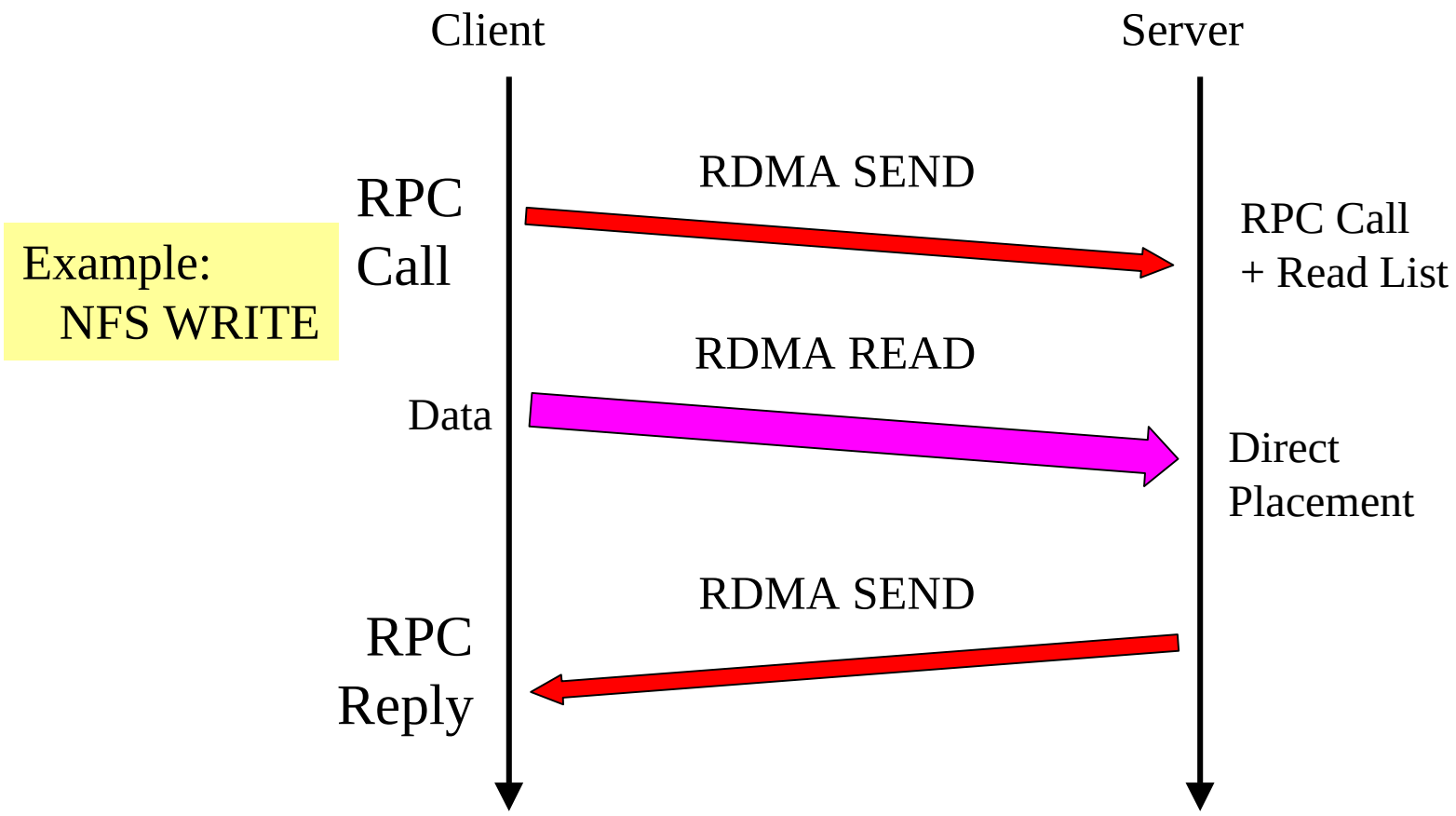
NFS INDUSTRY CONFERENCE

Small RPC Messages

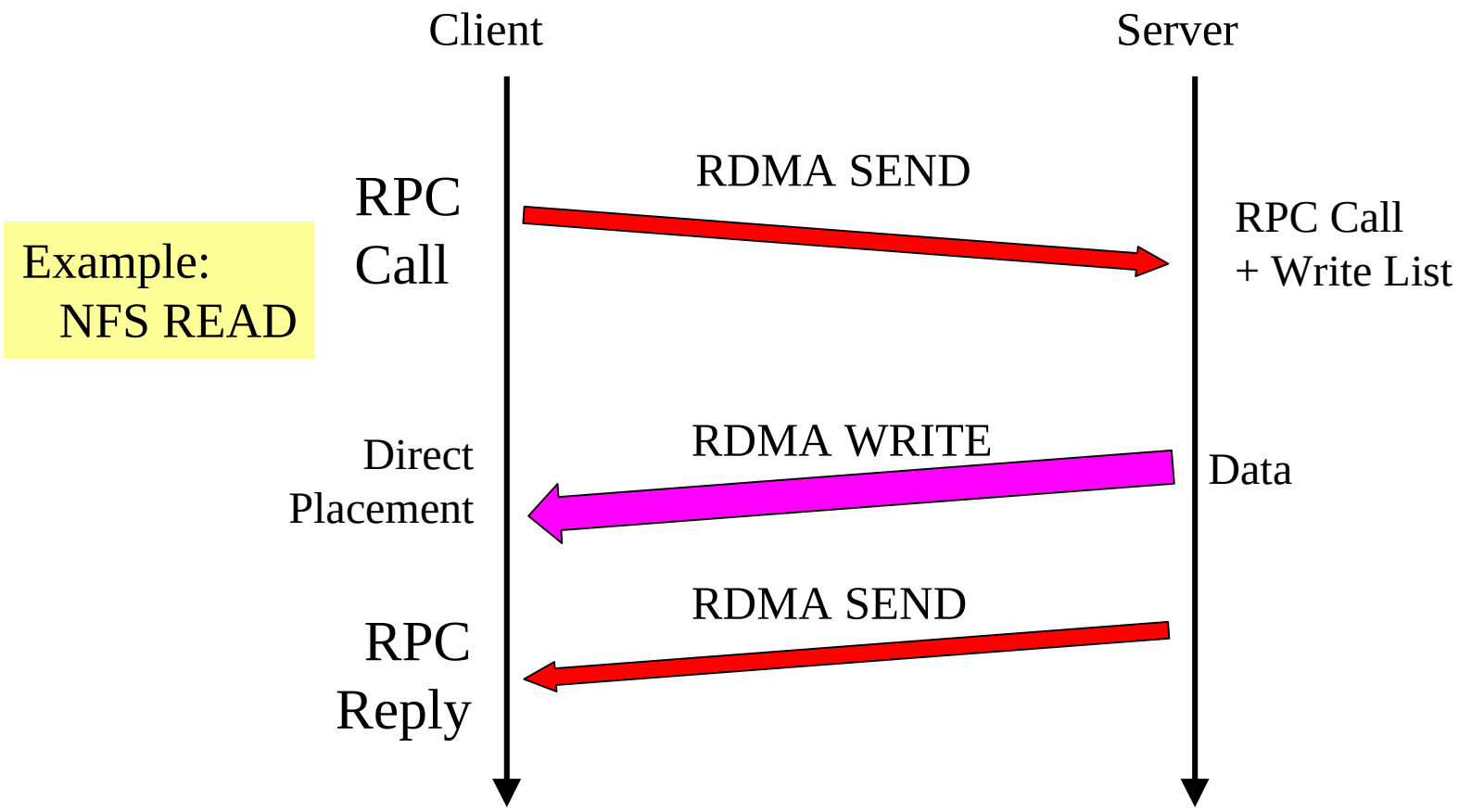
Most RPC Messages are small.
Examples:
- LOOKUP
- GETATTR



Big RPC Call

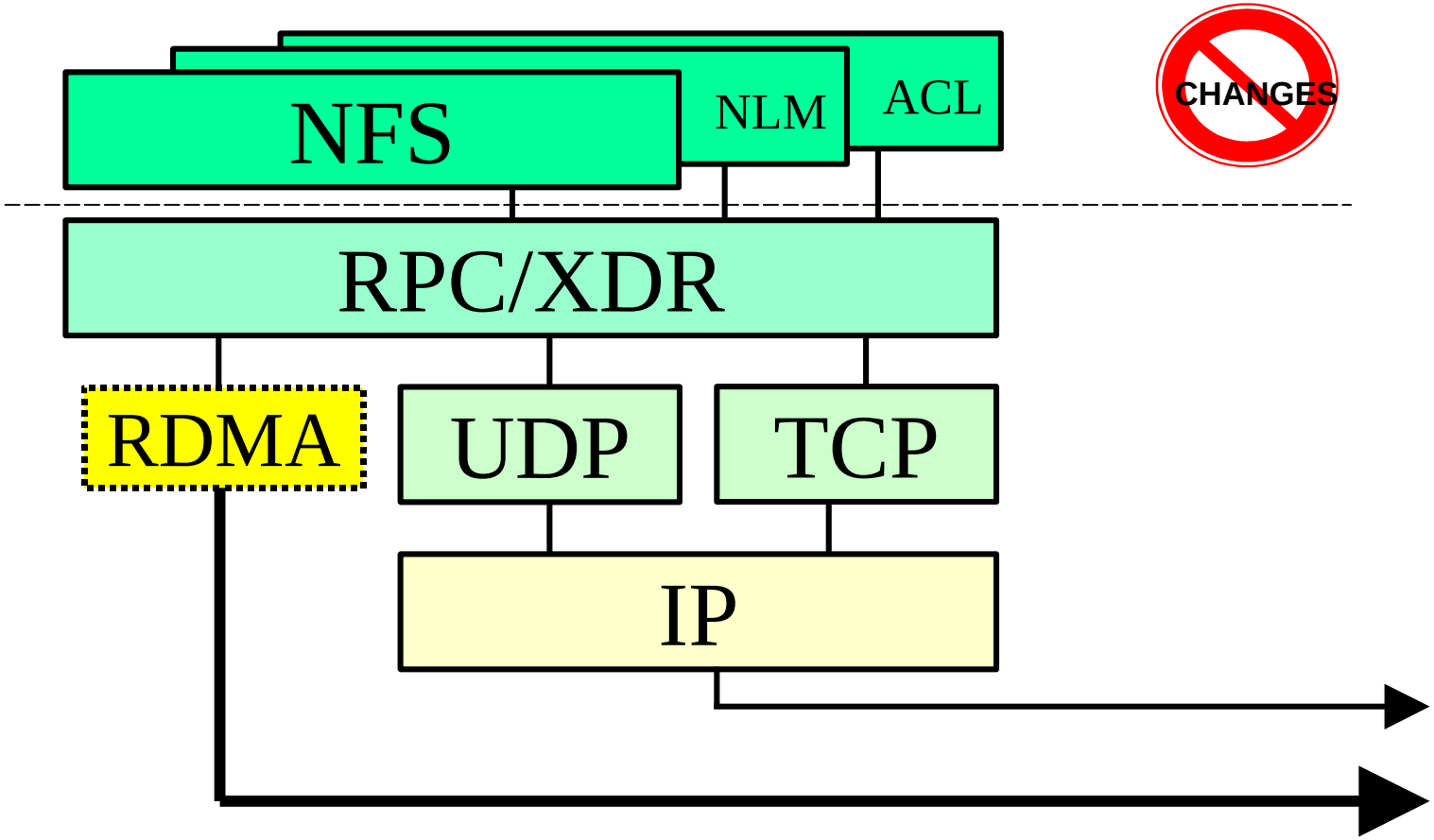

Client
Server
RPC Call
RDMA SEND
RPC Call + Read List
Example: NFS WRITE
RDMA READ
Data
Direct Placement
RDMA SEND
RPC Reply

Big RPC Reply

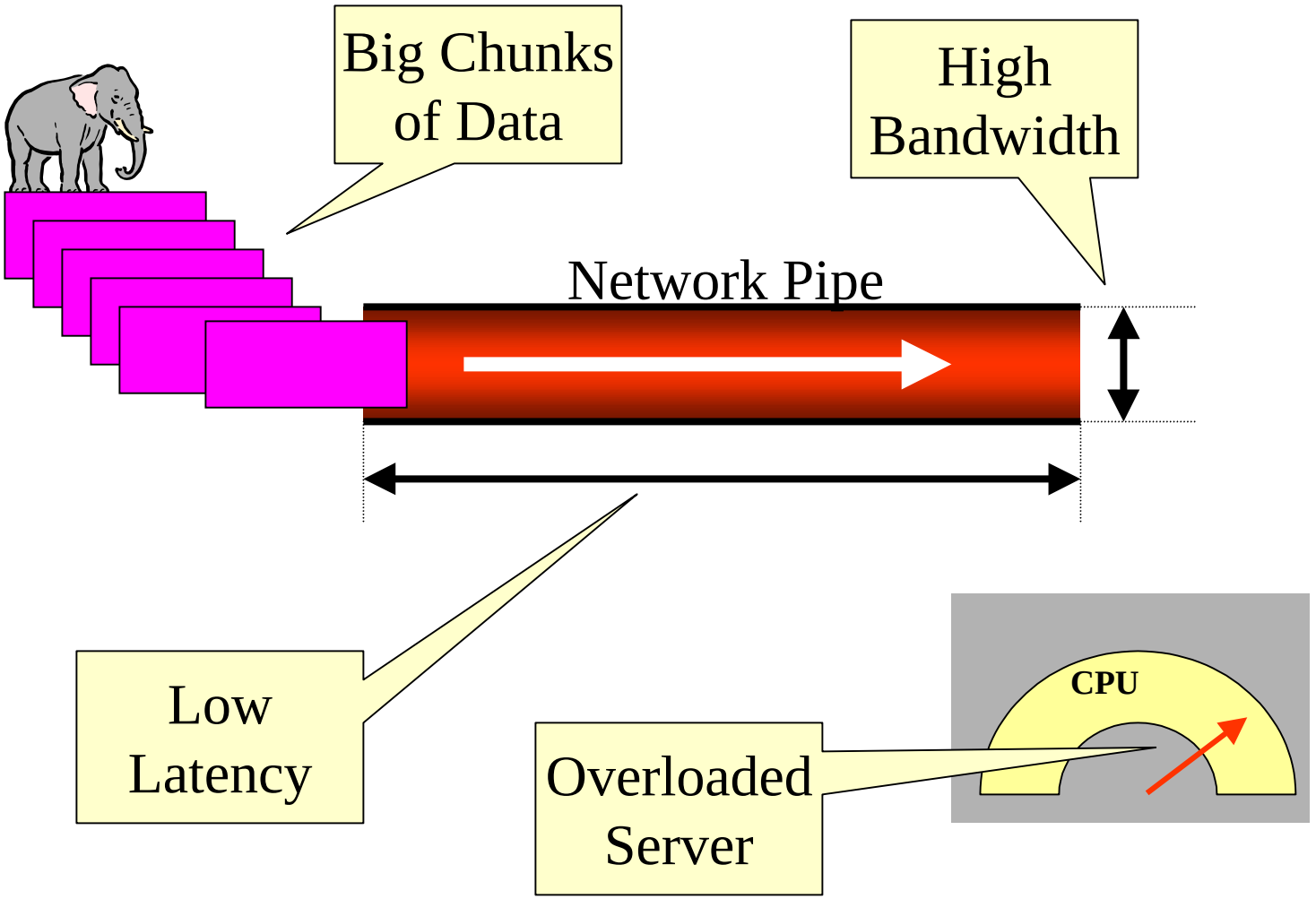


NFS INDUSTRY CONFERENCE

Adapting NFS to RDMA

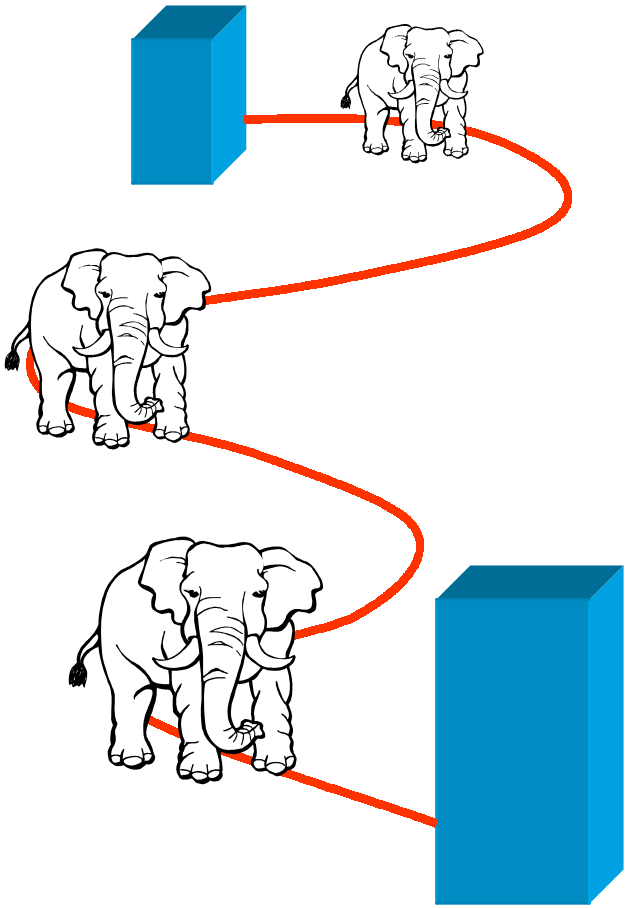


RDMA Sweet Spot



Sweet Spot

Big Chunks of Data

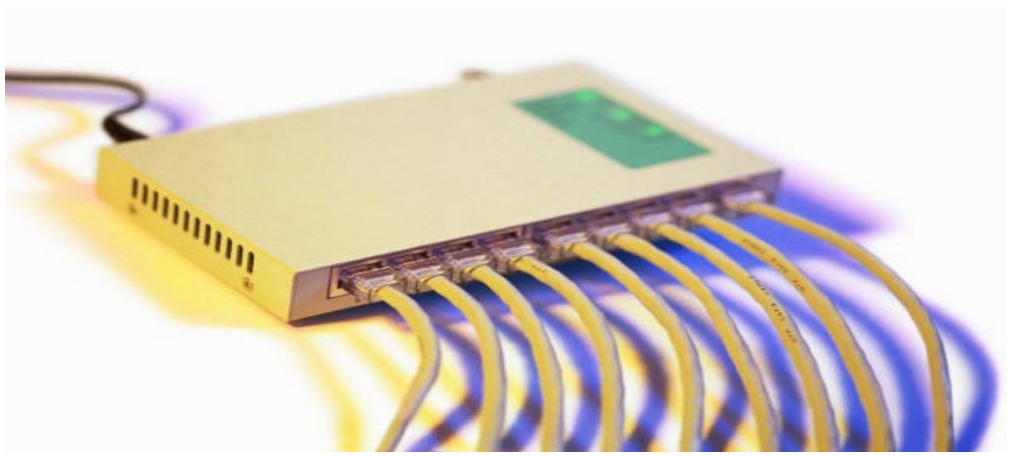


- RDMA Excels at Moving big data
  - NFS
  - Database
  - Backup
  - Replication
- Small data ?
  - May still be some improvement from protocol offload.

NFS INDUSTRY CONFERENCE

Sweet Spot

High Bandwidth



- ☐ 10 Mb/sec
- ☐ 100 Mb/sec
- ☑ 1,000 Mb/sec
- ☑ ☑ 10,000 Mb/sec

- 10 & 100Mb Ethernet - easily handled by CPU & bus
- 1 Gb Ethernet drains CPU & bus cycles
- At 10Gb, CPU & bus overwhelmed.

  Must have RDMA.

NFS INDUSTRY CONFERENCE

Sweet Spot

Low Latency

- RDMA Protocols more “chatty”
  - More network messages
  - Per message latency is cumulative
- RDMA performance advantages may fall off with distance

NFS INDUSTRY CONFERENCE

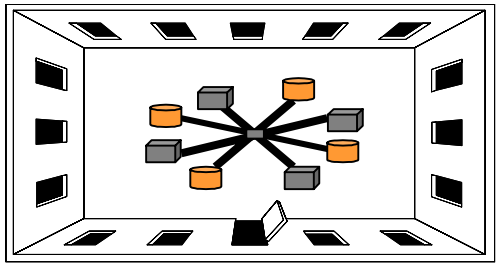
Sweet Spot

Overloaded Servers

- RDMA Hardware offloads protocol processing from busy CPUs.
- RDMA **Direct Data Placement** relieves congested memory and I/O buses.

Sweet Spot

RDMA in the Data Center



- Low latency - nodes are meters apart
- High bandwidth - runs are short, cheap
- Lots of data movement
- Infiniband clusters

NFS INDUSTRY CONFERENCE

NFS Storage Upgrade

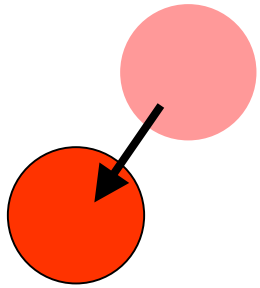
Sweet Spot

- Many customers already using NFS for data center storage.
- RDMA over Ethernet
  - Just a NIC upgrade.
  - Use existing fabric, switches, routers
  - No change to administration

NFS INDUSTRY CONFERENCE

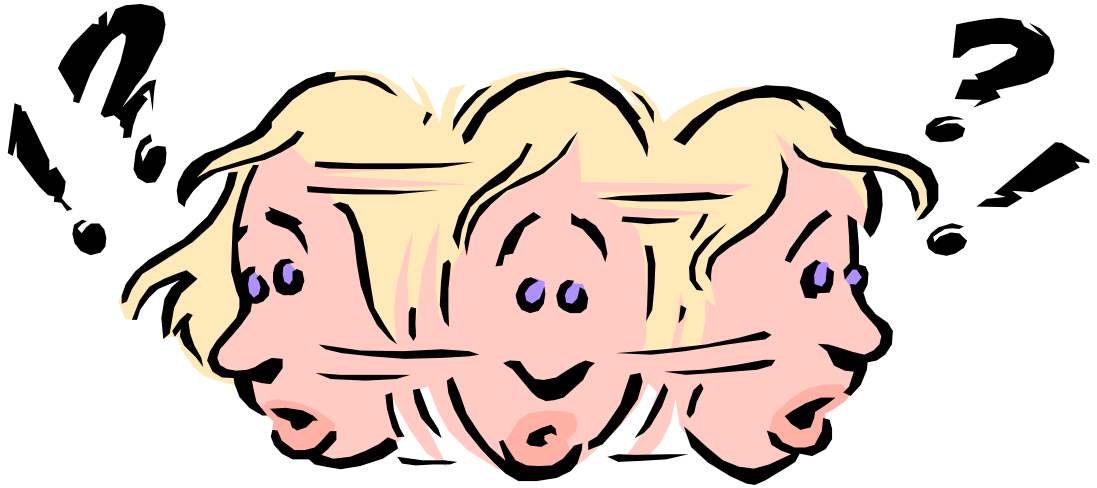
Sweet Spot will shift

- As NFS/RDMA becomes available on Ethernet
- As the price drops
  - RDMA on the motherboard



- As available bandwidth grows
- As big data increases
- As CPU performance improves

NFS INDUSTRY CONFERENCE



Questions & Answers