Network File System (NFS) in High Performance Networks

Wittawat Tantisiriroj, Garth Gibson

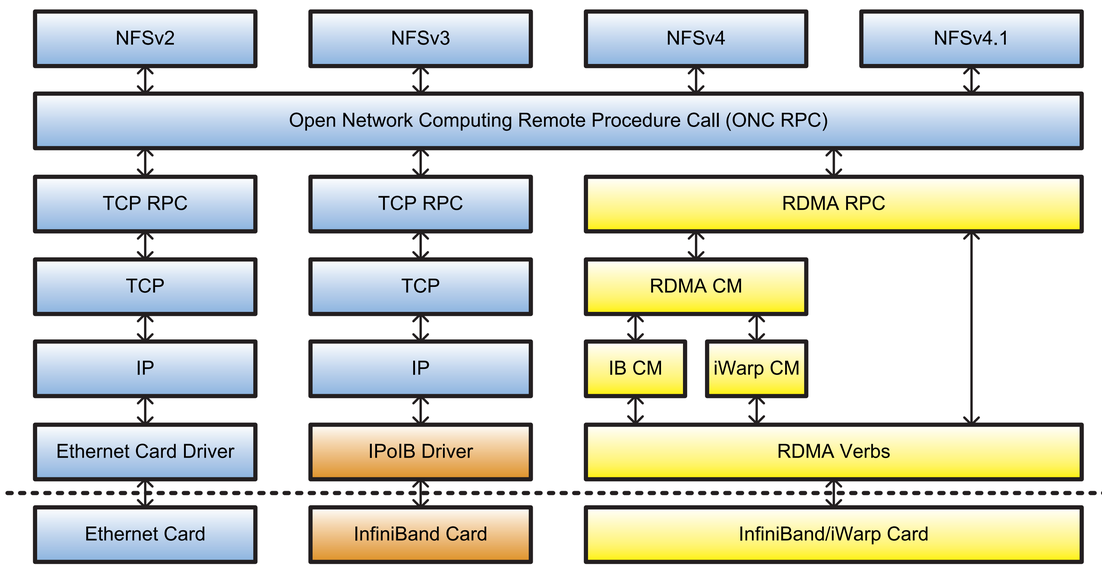
Overview

- **NFS over RDMA was recently released in February 2008**
- **What is the value of RDMA to storage users?**
- **Competing networks**
  - **General purpose network (e.g. Ethernet)**
  - **High-performance network with RDMA (e.g. InfiniBand)**

Type	Bandwidth (Gbps)	~Latency (µs)	~Price per NIC+Port ($)
Gigabit Ethernet	**1**	**40**	**40**
10 Gigabit Ethernet	**10**	**40**	**1,350**
Infiniband 4X SDR	**8**	**4**	**600**
Infiniband 4X SDR	**16**	**4**	**720**
Infiniband 4X SDR	**32**	**4**	**1,200**

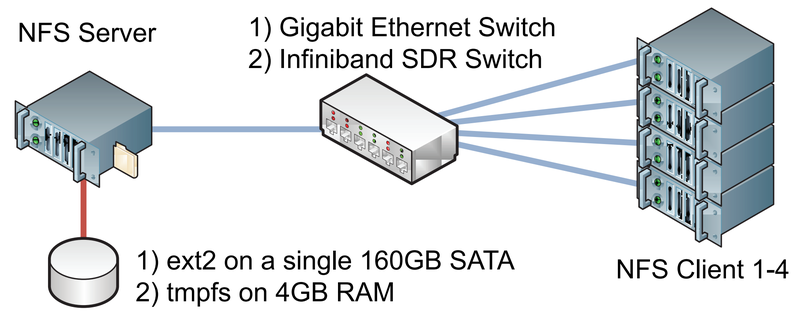
Source: High-Performance Systems Integration group, Los Alamos National Laboratory (HPC-5, LANL)

NFS over IPoIB / NFS over RDMA

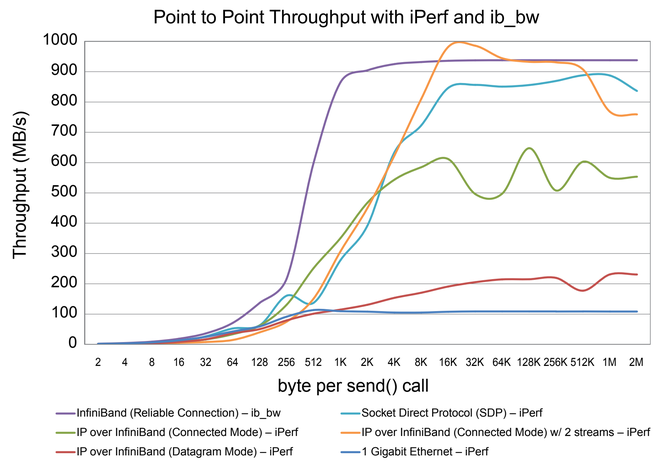


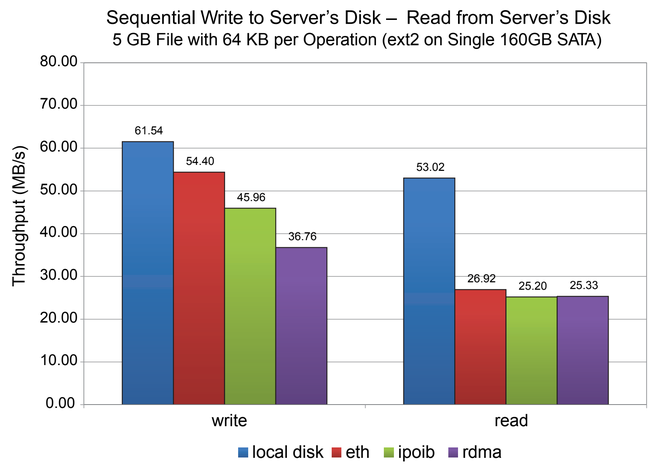
- **IPoIB: Implemented as a standard network driver**
- **RDMA: Implemented as a new RPC type**

Experiment Setup

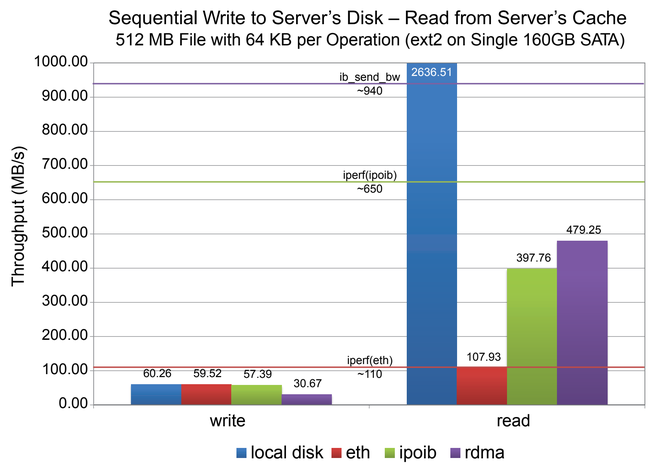


Experimental Results

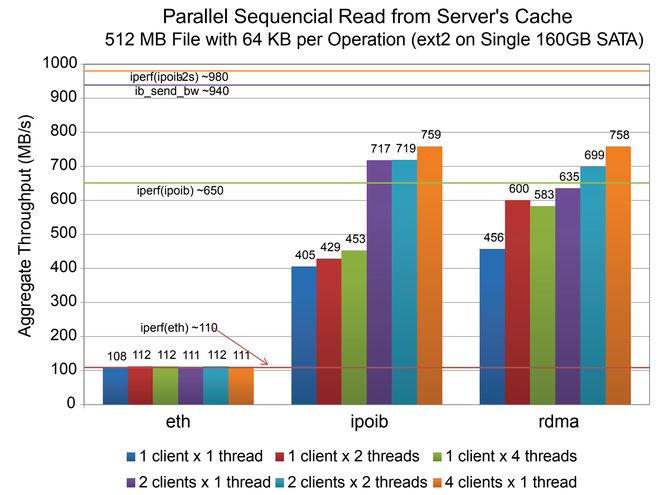




- **For point-to-point throughput, IP over InfiniBand (Connected Mode) is comparable to a native InfiniBand**



- **When a disk is a bottleneck, NFS can benefit from neither IPoIB nor RMDA**



- **When a disk is not a bottleneck, NFS benefits significantly from both IPoIB and RDMA**
- **RDMA is better than IPoIB by ~20%**

- **As the number of concurrent read operations increases, aggregate throughputs achieved for both IPoIB and RDMA significantly improve with no disadvantage for IPoIB**

Carnegie Mellon

Parallel Data Laboratory

PDLSVD08-02