3x in 30 Minutes

Application Acceleration Via Improved Network Communication

Joachim Worringen, Senior Software Architect

MySQL Users Conference 2009

Dolphin Interconnect Solutions

Application Acceleration Components

- PCIe cluster switches and host adaptors
- PCIe SSD/storage
- PCIe multi-host IO-Expansion
- Bundled, optimized software stacks

Enterprise DB High Performance Appliances

- Optimized Platforms for MySQL based enterprise and web applications
- Appliances for Data Warehousing and Business Intelligence

Embedded systems Real Time Components

- Reflective Memory
- Simulators
- Medical Imaging
- Mission Computers

connect better

Airtravel Experience

Buying a ticket via the website of a large german airline:

- First experience: system is slooooow
- Logged in and noticed personal information is outdated
  - Updated personal information (email address)
- Reserved flight
  - Ticket was sent to old email address!

Asynchronous scale-out? Inconsistent caching?

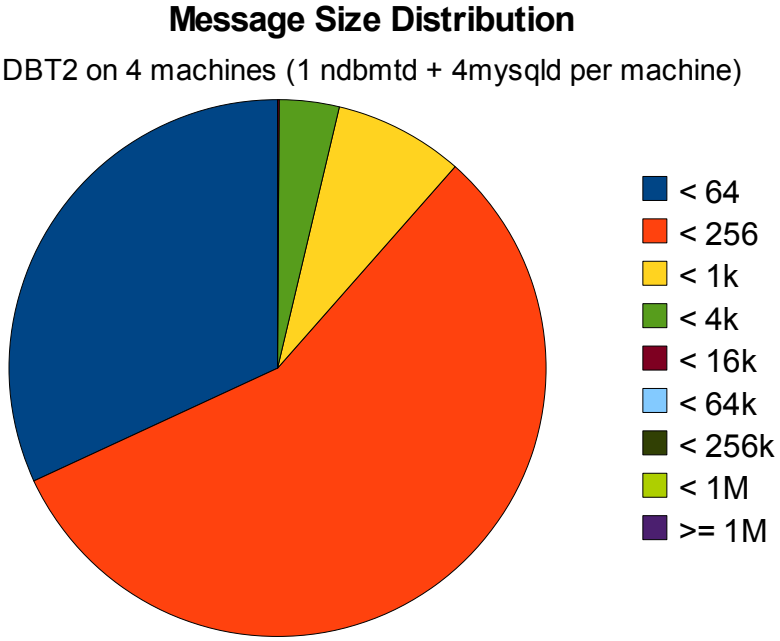
Performance is Key (plus Reliability)

- Well-known performance factors:
  - CPU: # cores, architecture
  - Memory: never can have to much
  - Disk: more spindles, caches, solid state
  - Software: client, query, schema, storage engine, OS, ...
- Network! For certain architectures:
  - MySQL Cluster
  - Synchronous Replication:
    - DRBD
    - mysqld with semi-synchronous replication
    - custom-designed MySQL replication solutions
  - memcached

...whereever response time matters

Latency Drives Database Performance

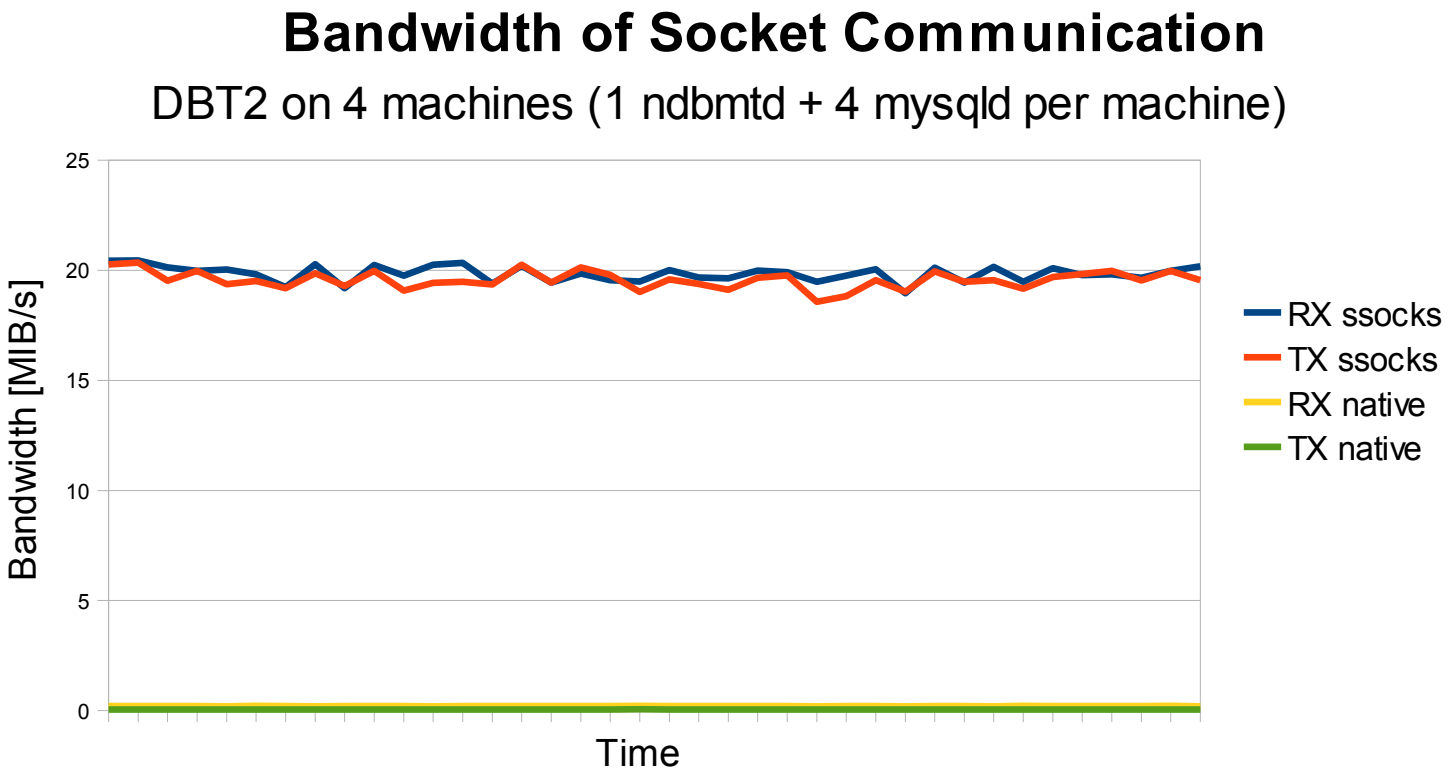
MySQL Cluster 6.4 with DBT2 Benchmark on Sun X2250



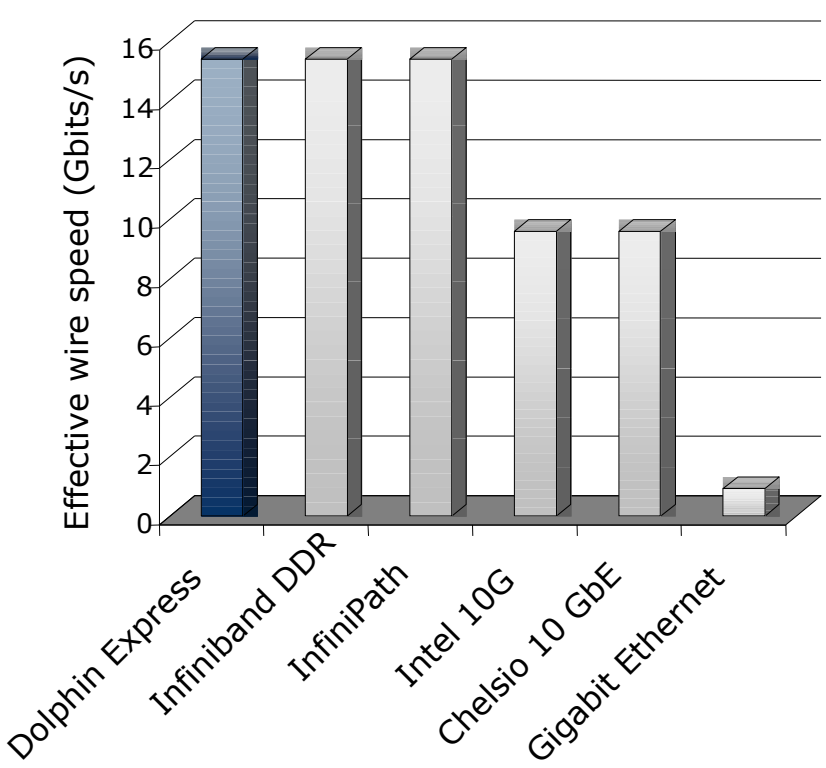
…because database transfers are small!

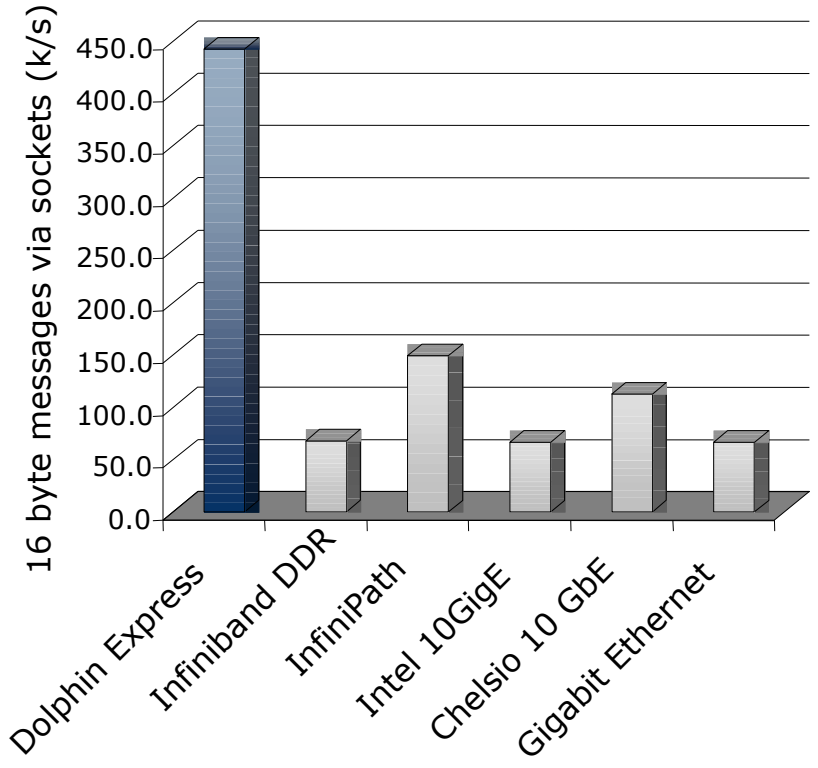
Bandwidth Rarely Matters

- DBT2 benchmark on MySQL Cluster
  - Bandwidth per node does not exceed 20MB/s



The Wire Speed Story

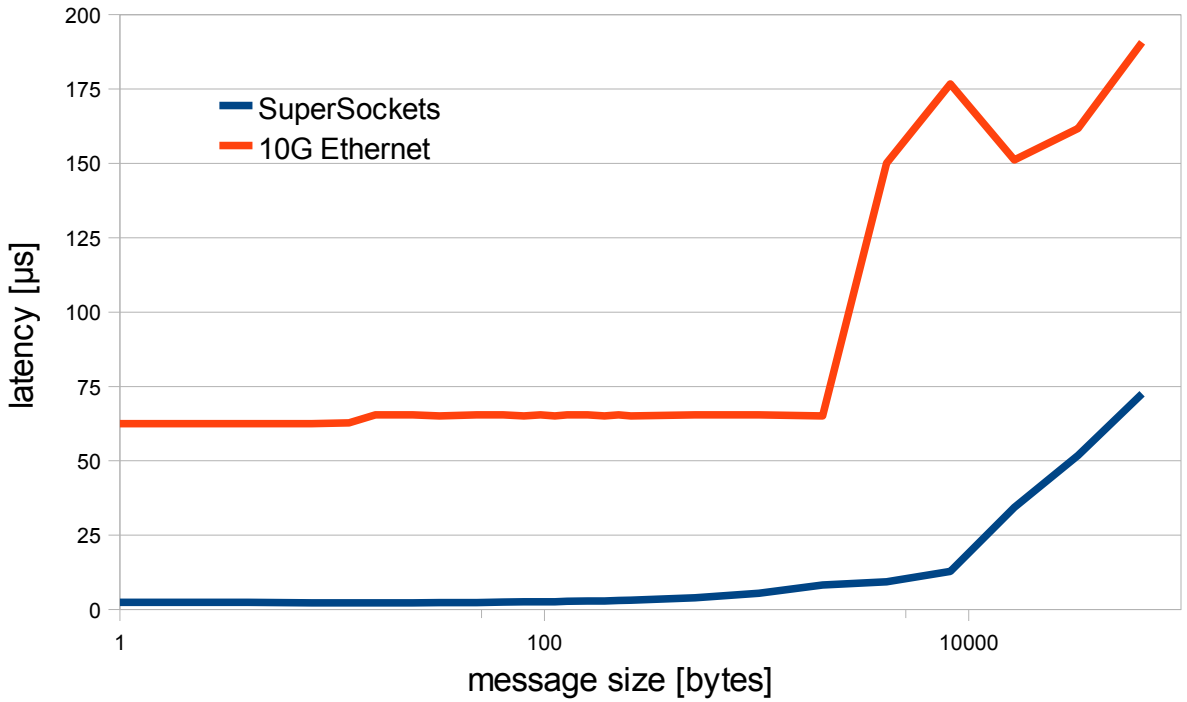




Charts based on publicly available information

Dolphin SuperSockets vs 10GigE

- Latency of 10GigE about the same as 1GigE



data size	ratio
512	16.5
1k	12.0
2k	7.9
4k	16.1
8k	13.8

Measured on Intel E5405 (Xeon 2.0GHz quadcore), SLES 10
10 GigE: Intel 82598EB CX4, driver 1.3.47 (default settings)
Dolphin Express: DXH510 CX4, driver 3.0.1 (default settings)

connect better

16-Bytes TCP Packet Latency

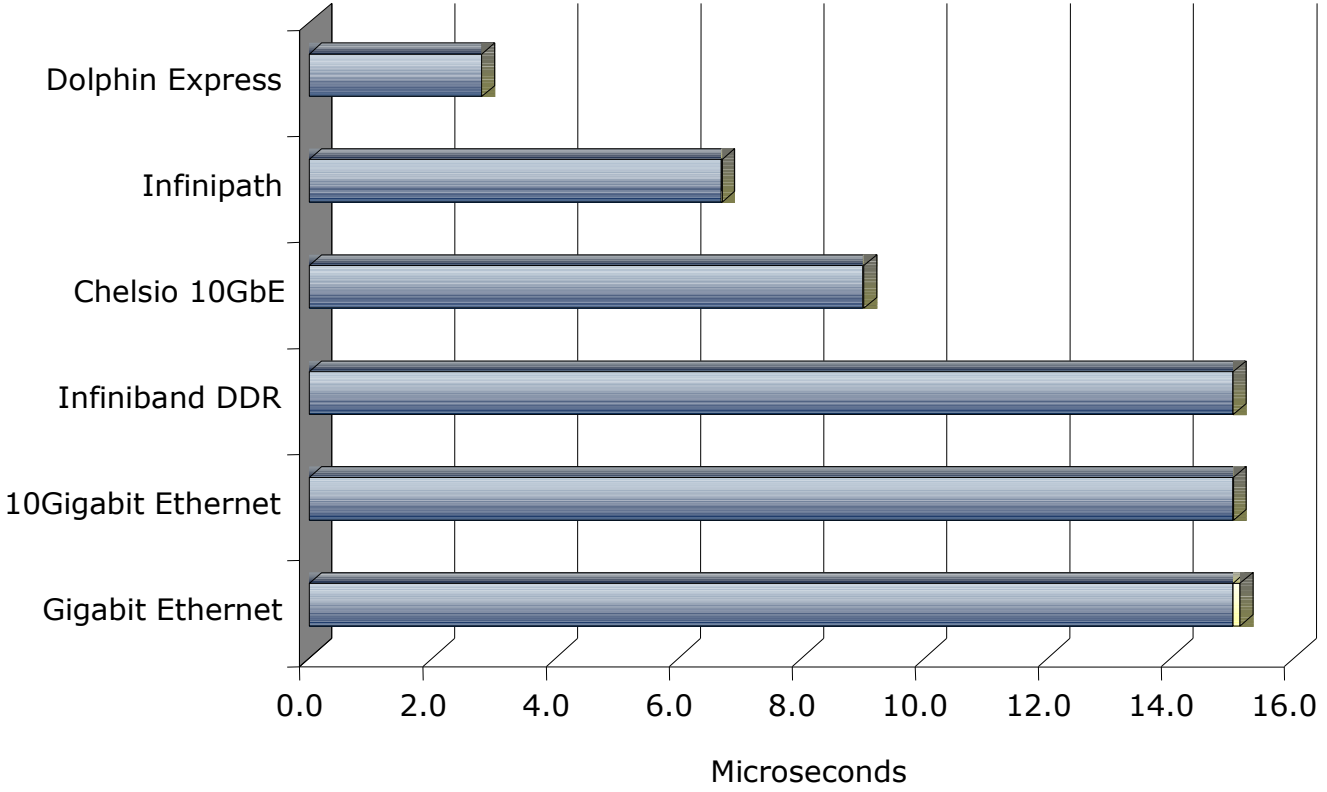
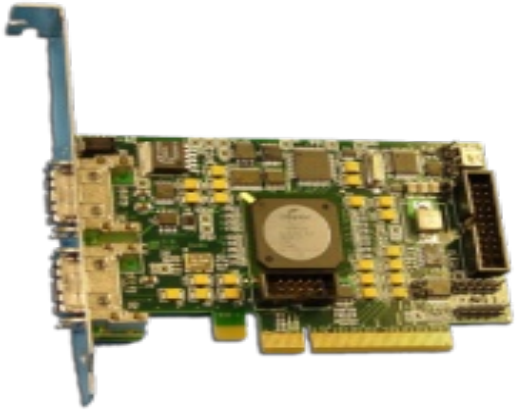


Chart based on publicly available information

connect better

Dolphin Express - Hardware

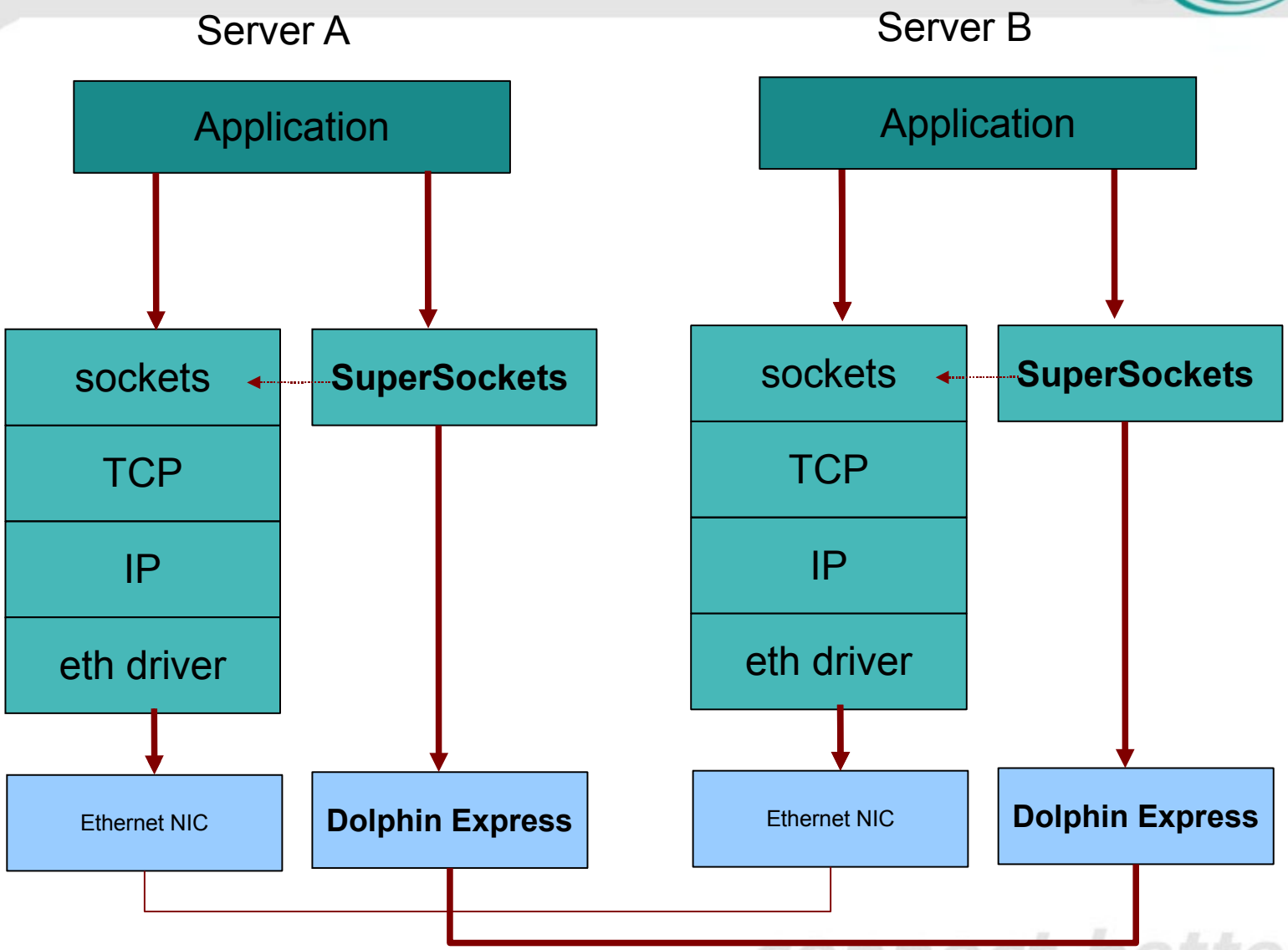
- High Speed Interconnect:
  D and DX Series NICs and Switch
  - Utilizes PCI Express and Dolphin Express interconnect
  - Low latency communication – direct remote memory access
  - Up to 20Gbps transfer rate
  - Low cost switching solution



Dolphin Express - Software

- Software Compatibility: **SuperSockets**
  - Berkeley Sockets API on Dolphin Interconnect
    - New socket transport family AF_SSOCKS
  - TCP/UDP/RDS compatible
  - Data transfer through remote shared memory
    - PIO for low latency
    - DMA for low CPU utilization
  - Kernel Sockets:
    - Full Socket Semantics for Cluster Applications

How do SuperSockets work?



Beyond Performance: Availability

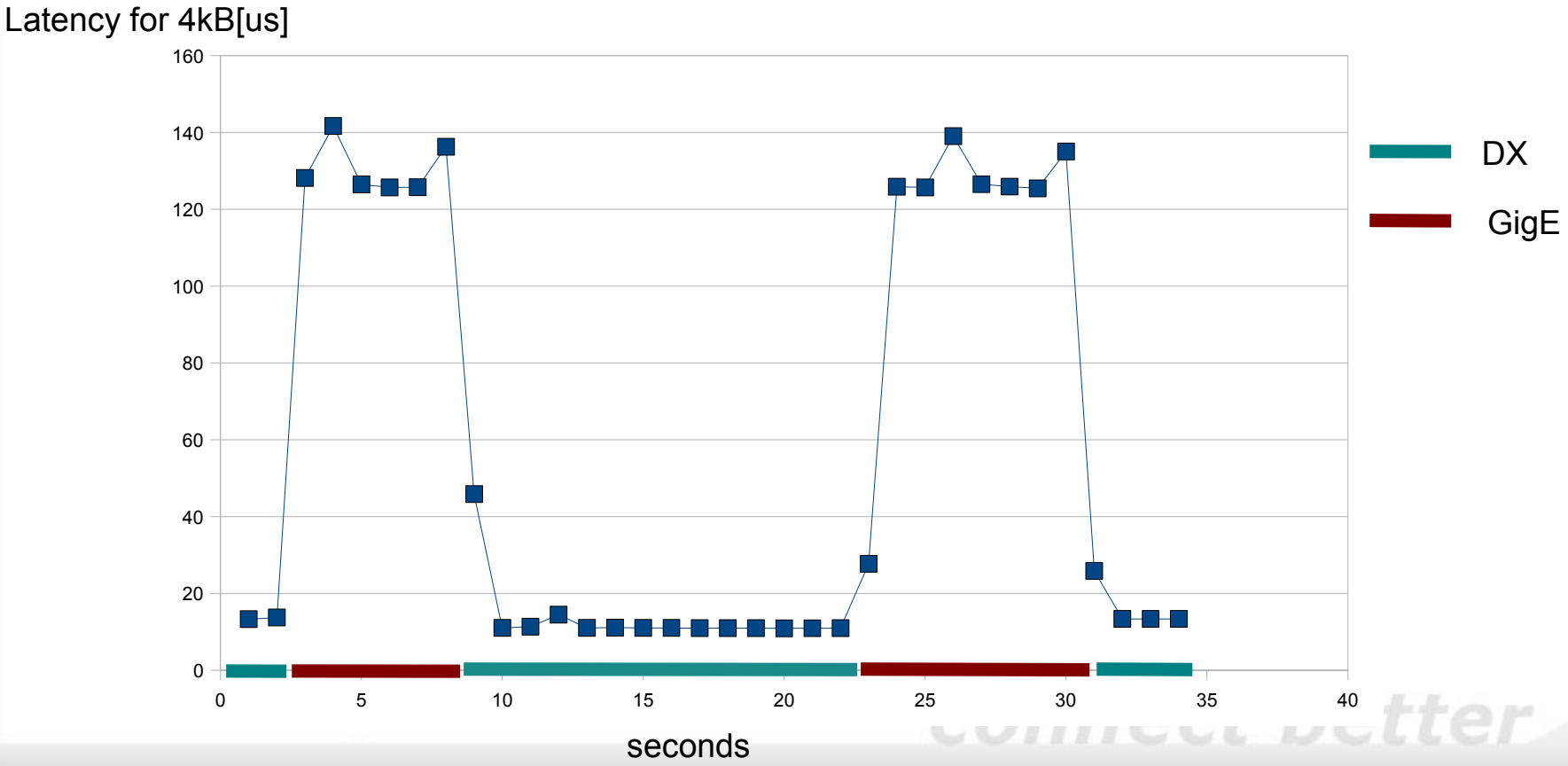
- Integrated High Availability
  - Two independent 10Gbps ports (DX)
    - Hardware-bondable for transparent, true 20Gbps
  - Transparent, automatic failover and failback to Ethernet
  - Integrated link watchdog and heartbeat
    - Detection of unresponsive link endpoints

No packets lost due to node, card, switch or cable failure

No single point of failure

Fail-over cycle: DX - Ethernet - DX

- roundtrip/2 latency for fail-over
  - Pulling and re-inserting DX cable



Beyond Performance: Deployment

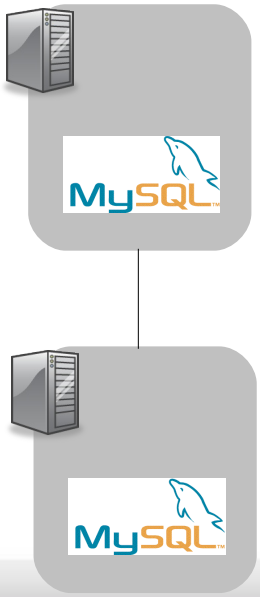
- Universal fit
  - Supports all applications (TCP, UDP, RDS)
  - No changes to applications
  - No specialized database version required
- Operating System Support
  - Linux 2.6 **(all kernels)**
  - Solaris (OpenSolaris release 5/2009 and Solaris 11)
  - Windows

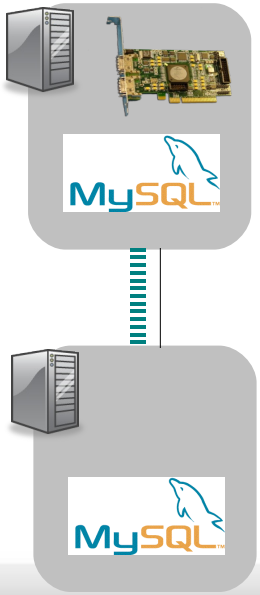
Installation – simple yet flexible

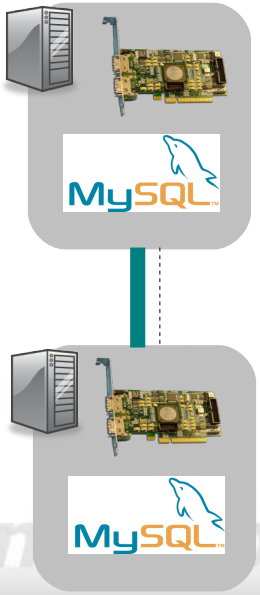
- Install Dolphin Express hardware
- Install SuperSockets software stack
  - Distributed as „self installing archive" (SIA)
    - Executable shell archive
  - Invoke SIA on **one** node to install on **all** nodes
  - Compiles on the fly to match the running kernel
    - Works with **any** 2.6 kernel you might be using
  - Installs itself as RPMs or DEB packages

Deploying SuperSockets

- SuperSockets increase socket performance
  - for new or existing setups
  - without being forced to use a specific kernel
  - without changing application configuration
- With redundant applications, rolling upgrade to SuperSockets is possible:

1. 

2. 

3. 

Example: Accelerating DRBD

Live-upgrade of DRBD-based HA-Cluster:

- Shut down Secondary, install Dolphin Express Hardware and SuperSockets
- Update DRBD if necessary
  - 8.2.7 and above with SuperSockets support
- Restart Secondary, wait for resync to complete
  - This operation will run over Ethernet (SuperSockets fail-over mode)
- Fail over, promoting Secondary to Primary
- Repeat on peer

Now, DRBD runs over SuperSockets!

LINBIT
YOUR WAY TO HIGH AVAILABILITY

connect better

Customer Story: 3P Learning

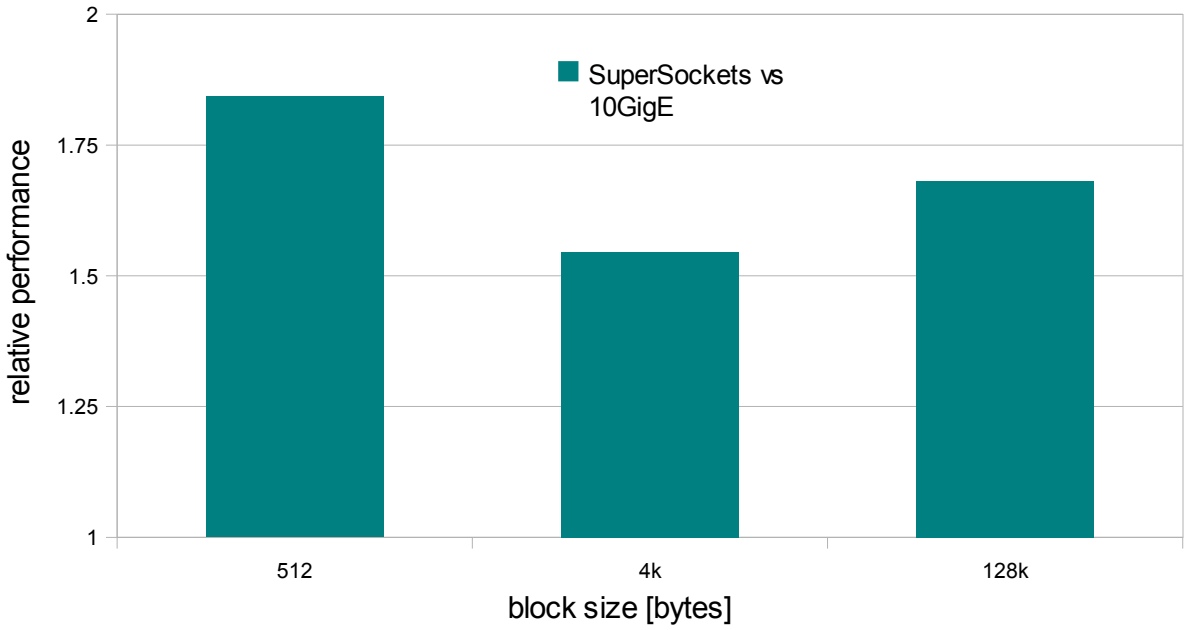
- 3P Learning hosts www.worldmathday.com for a world-wide math contest for schools
  - Very high load during the contest days – several hundred million of „answers" are collected
- 2008 system
  - 6 Windows-based web servers
  - 1 MySQL server: RHEL, Quad-core, 64GB RAM
- 2009 system
  - 40 machines for NDB API clients, 1 for NDB manager
  - 8 machines for NDB data servers
  - All connected with Dolphin Express D / SuperSockets
- Collected 430.000.000 „correct answers" within a few days

Customer Story: Blue Dog Training

- Blue Dog Training (Australia) offers web-based training-courses for the construction industry
- DRBD via DX
  - Two IBM x3650 connected
  - Storage synchronized via DRBD uses controller with non-volatile cache
  - MySQL INSERT operations reported **6 times faster**
- Other DRBD results:
  - Generic block-write latency reduced by more than a factor of 2 (vs 10GigE)
  - File-system bandwidth merely limited by storage system: > 400MB sustained for Dolphin StorExpress

DRBD with Dolphin StorExpress

- DRBD with DX vs 10GigE on solid state storage
- fio benchmark: concurrent file system write



RHEL 5.2 on dual quad-core Intel Xeon 3GHz / Intel 10GigE 82598EB CX4, driver ixgbe 1.1.8 / Dolphin DX x4 / DRBD 8.3.0 / benchmark: fio --numjobs=4 --sync=0 --rw=write --size=100m --bs=512 --ioengine=libaio --iodepth=1 --loops=5 --direct=1

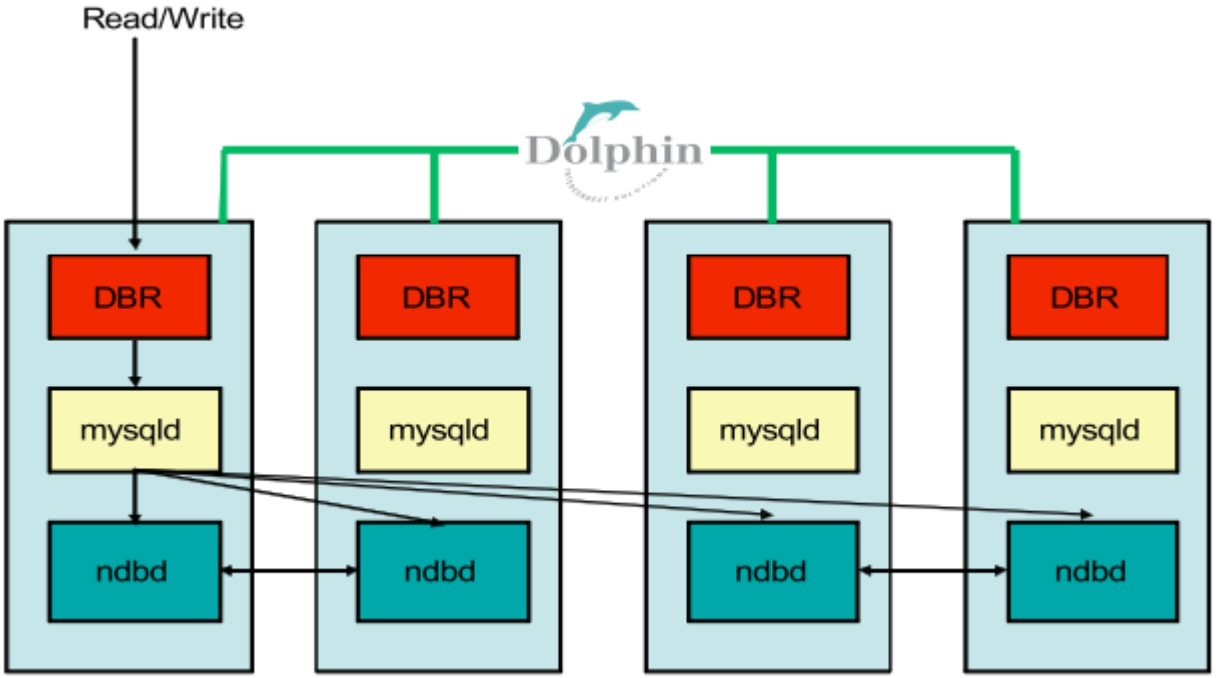
Customer Story: Teligent

- Teligent develops value-added services for Telco carriers
  - Based on „Teligent Application Server" platform
- Use MySQL databases (InnoDB and Cluster)
  - Soft realtime constraints
  - Synchronous replication is required
  - Additional „DBR" tier between client and mysqld
    - Load Balancing
    - Locking
    - Replication

connect better

Teligent Results: MySQL Cluster

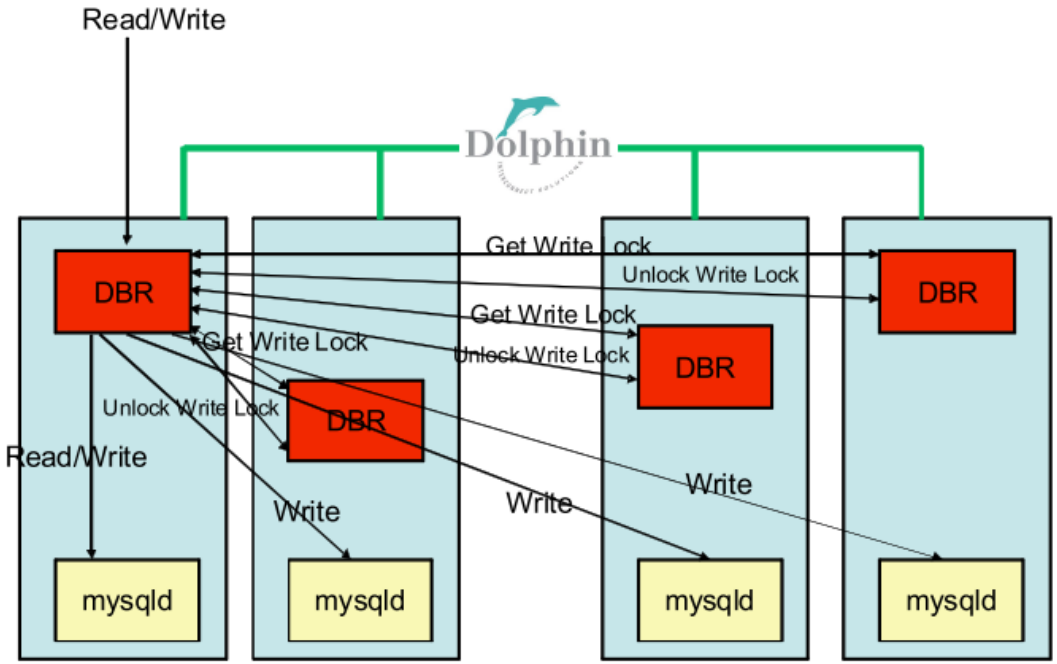
- MySQL Cluster with Load Balancing



  - Increase in TPM: **Factor 2.2**

Teligent Results: InnoDB

- InnoDB with custom synchronous replication
  - Requires locking between mysqld's for consistency



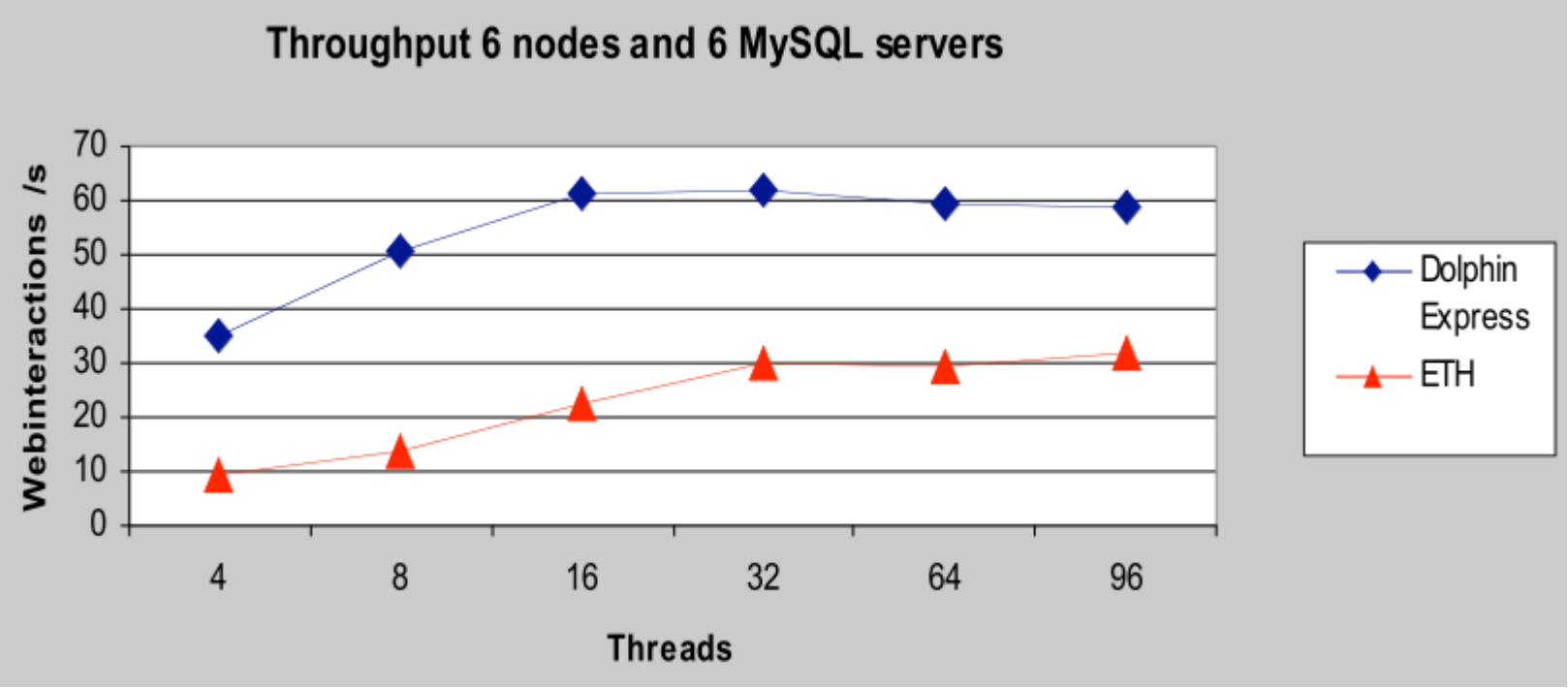
  - Increase in TPM: **Factor 2.3**

Telecom Voting System

- British customer developed telephone voting system based on MySQL
- Custom layer for synchronous replication
- Performance increase with Dolphin SuperSockets (Series D hardware)
  - From 1400 TPM with GigE to 6000 TPM with SuperSockets: **Factor 4.3**

Web Application Test

- TPC-W benchmark with MySQL cluster
  - Apache and Java/Tomcat application server



Throughput increases by factor 1.9 ... 5.25

Response time decreases up to factor 7.

More Advantages

- memcached
  - replace disk latency with network latency
  - memcached fetch() latency reduced by a **factor of 3**

- Loopback sockets:
  - SuperSockets also accelarate loopback communication (TCP sockets within a machine)
  - Becomes more common the more cores per machine
  - Get the same latency between and inside a machine

Try Dolphin Express

- We know: each application is different.
- You've seen: Dolphin Express is simple to deploy.
- We offer: **Turbo Test**
  - Buy & evaluate Dolphin Express SuperSockets
  - If not satisfied: return within 30 days for full refund

- Visit www.dolphinics.com
  - Register for Turbo Test
  - Download
    White Papers & Software
  - Online Shop

Dolphin is more...

- Dolphin StorExpress

  - High capacity (up to 2TB), scalable, ultra high performance solid state storage system
  - 270K IOPS (read/write, 512 to 4k bytes)
  - 2,800MB/s sustained bandwidth

- Dolphin and ScaleDB

  - Newly announced partnership between ScaleDB and Dolphin ICS
  - Delivery of MySQL-based database appliance with revolutionary scaleability – stay tuned.

Meet us at our booth!

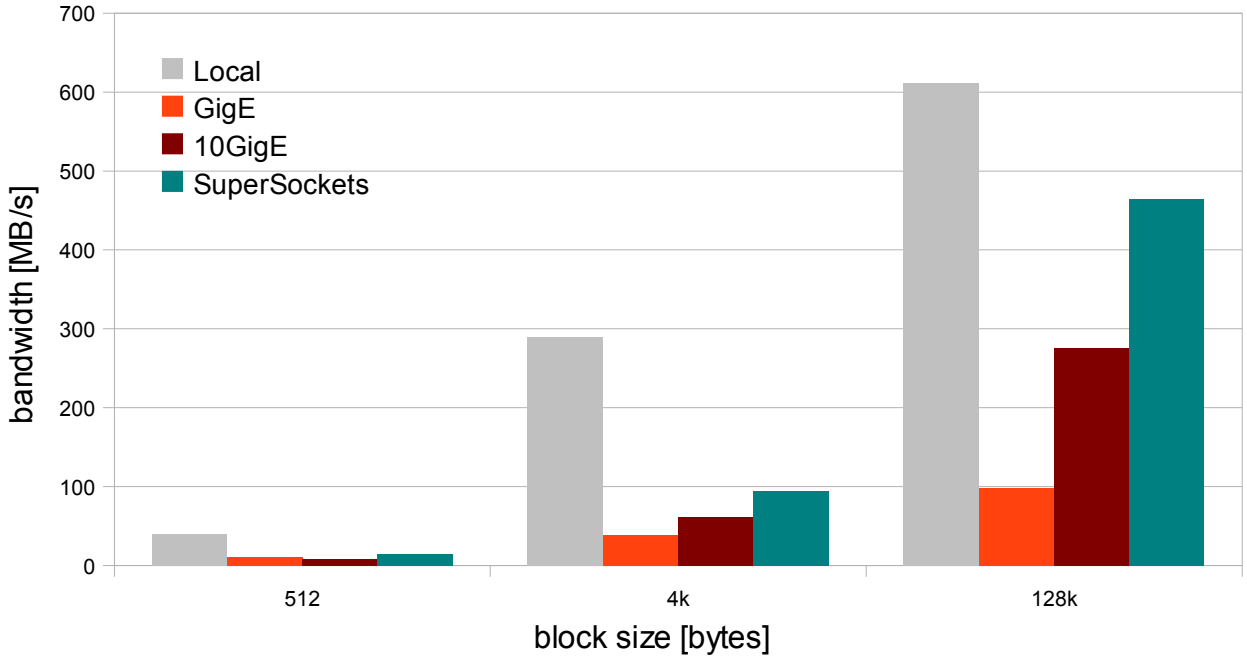
connect better

Thank you!

Q & A

DRBD with Dolphin StorExpress

- DRBD with DX vs 10GigE on solid state storage
- fio benchmark: concurrent file system write



RHEL 5.2 on dual quad-core Intel Xeon 3GHz / Intel 10GigE 82598EB CX4, driver ixgbe 1.1.8 / Dolphin DX x4 / DRBD 8.3.0 / benchmark: fio --numjobs=4 --sync=0 --rw=write --size=100m --bs=512 --ioengine=libaio --iodepth=1 --loops=5 --direct=1