

## The DNA of Next-Generation Datacenters Schooner Information Technology

Dr. John R. Busch, Co-Founder and CEO August 12, 2009

Thursday, August 6, 2009

#### Too Much Rack, Power, Pipe, Complexity



SCHOONER

Thursday, August 6, 2009



## Four transformational technologies:

- Multi-core processors
- State-of-the-art, enterprise-class flash memory
- Low-latency interconnects
- Optimized data access and caching applications

#### **Typical Web 2.0 and Cloud Deployment**

SCHOONER



© 2009



### **Application and administrator managed scale-out:**

- Add more commodity boxes, GbE, storage to support growing workload
- Applications must provide partitioning, parallelism, concurrency control, replication, and recovery
  - Complex development, integration, re-partitioning
- Exploiting multi-core, flash and low latency interconnect is a customer by customer problem
  - Effectively a research project
  - Ineffective results to date: single-digit utilization

© 2009

### Integrated, optimized, multi-node appliances:

- Effectively leverage multi-core processors, high-speed interconnect, flash memory, middleware software
- Incorporate highly optimized, balanced hardware platform, operating environment, networked middleware data access applications
- Potential: cut power consumption and datacenter floor space by up to 90%
- Compatible with existing web/app tier client applications and monitoring tools
- Provide higher level building blocks that eliminate complex integration projects

## **Tightly Coupled Hardware Architecture**



SCHOONER

## **Tightly Coupled Software Architecture**



#### **Networked Service Application**

1/10 G E-Net Management ■ Application Protocol Handling

#### **Data Fabric**

Object Attributes ■ Thread and Core Management Synchronization/Concurrency Management ■ DRAM Cache Management Container Management ■ Object Metadata Management ■ Replication Management

#### Flash Management Subsystem

Space Allocation and Shard Management Object Replacement (cache mode) Persistency Management Tiered Storage Management

#### Flash Access

Asynchronous IO Handling, Data Striping, Interrupt Batching Admin Configure

> Monitor Control Optimize

© 2009

- Optimizes transactions/sec/watt, transactions/sec/core, transactions/sec/\$
- Supports application level protocols
- Optimizes multi-core
- Manages highly parallel flash memory devices
- Optimizes memory hierarchy: L1/L2/L3/DRAM/Flash/HD
- Performs efficient multi-node replication with load balancing and transparent failure recovery





#### **Schooner Data Access Appliances**

- Purpose-built for Web 2.0 and cloud computing datacenters
- Distributed caching and database appliances
- 8x performance improvements
- 1/8th the power and space requirements
- 60% lower TCO
- 100% compatible with existing client applications and management tools

#### **IBM Partnership**

•IBM manufactures, sells, and provides global, 24/7/365, single-point-ofcontact service and support for every Schooner appliance

© 2009



# Tightly Integrated Multi-Core + Flash + Interconnect + Distributed Caching

- 10 x throughput and network utilization
- Cache mode or persistent key/value store mode
- 90% reduction in power
- Instantaneous and transparent data persistence, replication, and recovery

#### **Example: Memcached distributed caching tier appliance**







© 2009

#### Tightly Integrated Multi-Core + Flash + Interconnect + Database

- Tightly integrated, highly optimized
- High read AND write flash IOPS (>200k/sec) + multi-core scaling
  - -Read-
    - W
    - rite intensive workloads (OLTP), as well as OLAP (read mostly workloads)
  - Fast recovery and warm-up after restart or failover
- High Speed Interconnect enables fast failover and replication/recovery
- Example : MySGlatz langa Bae (т. 2.3) Database Appliance







#### Data Center Impact: Consolidation, Power, TCO



SCHOONE

	Legacy Memcache	Legacy MySQL	Schooner Appliance	Schooner Savings
Nodes	2,150	3,800	595	
CapEx	\$13,480,500	\$50,236,000	\$26,995,150	58%
OpEx	\$25,206,600	\$44,551,200	\$16,257,780	77%
тсо	\$38,687,100	\$94,787,200	\$43,252,930	68%

© 2009





X-Architecture technology



SC

-----

SCHOONER



## Thank you.

X-Architecture technology





## **Backup Slides**

X-Architecture technology

Thursday, August 6, 2009

## Flash Drive Comparison



	Read B/W	Write B/W	Erase Lat.	Read Lat.	Cost per GB
HDD	100 mb/s	150.00 mb/s		5,000.00 us	\$0.10
NAND MLC	250 mb/s	70.00 mb/s	3.5 ms	85.00 us	\$3.50
NAND SLC	250 mb/s	170.00 mb/s	1.5 ms	75.00 us	\$11.00
NOR SLC	58 mb/s	0.13 mb/s	5,000.00 ms	0.27 us	\$70.00
DRAM	2,000 mb/s	2,000.00 mb/ s		0.08 us	\$75.00

© 2009

### **Flash Drive Comparison**



	Fusion-io io-drive	X25-E 8x array		
Read latency uS 4KB	75	75	uSEC	
Write Latency uS 4KB	250	85	uSEC	Write latency of hardware
Read CPU uS	50	15	uSEC	
Write CPU uS	50-100	15	uSEC	Increases with garbage collection
Read MBPS 16KB	650	1400	MBPS	
Write MBPS 16KB	150	260	MBPS	Sustained with garbage collection
Read IOPS 4K	116,000	230,000	IOPS	
Write IOPS 4K	20,000	50,000	IOPS	Sustained with garbage collection
Capacity	160	512	GB	
Cost	\$7,200	\$5,800	USD	
Cost/Gbyte	\$56.25	\$11.33	USD	

\* Write throughput assuming continuous random writes spread across the drive, with 20% reserve capacity configured

© 2009