WHITE PAPER

August 2011

Accelerating Memcached Solutions with Mellanox VMA

What is Memcached?

As more and more data moves online and is accessed frequently, companies with substantial online infrastructure need datacenters where frequently accessed data is available without any down time. Accessing such data from traditional database infrastructure will create delays and bottlenecks. Web 2.0 based companies have moved such data into another tier of appliances called Memcached servers to host frequently accessed data. This is a way to accelerate web server access to the back-end database tier. Instead of doing a slow DB transaction, the web server tries first to load the object from a ram based cache/storage. If it cannot find the data, it will go to the database. Many web developers currently use this method in small and large sites to increase the number of web transactions/sec while offloading the database as it helps them decrease the response time to user queries. This is part of a family of solutions sometimes referred to as No SQL.

What is VMA?

The Mellanox Message Accelerator (VMA) library is a network-traffic offload, dynamically-linked, user-space Linux library which serves to transparently enhance the performance of socket-based networking-heavy applications over an InfiniBand or Ethernet network. VMA focuses on latency-sensitive and throughput-demanding, unicast and multicast applications and is used to accelerate producer applications and consumer applications, and enhances application performance by orders of magnitude without requiring any modification to the application code.

The VMA library utilizes the direct hardware access and advanced polling techniques of RDMA-CM for connection management in Mellanox HCAs. Utilization of direct hardware access enables the VMA kernel bypass, which causes the VMA library to bypass the kernel's network stack for all IP network traffic transmit and receive socket API calls.

Memcached solutions using the VMA library gain many benefits, including reduced context switches and interrupts, which result in:

- Lower latencies
- Higher throughput

- Improved CPU utilization
- Minimal buffer copies between user data and hardware – VMA needs only a single copy to transfer a unicast or multicast offloaded packet between hardware and the application's data buffers.

Why VMA for Memcached Solutions?

Memcached solutions are fast transaction-based network applications, which require a high rate of request-response type operations over TCP or UDP unicast. These solutions also involve any send or receive to and from an external network entity, such as smartphones accessing Web 2.0 applications, mobile browsers, laptops, etc. They also do caching/data distribution applications, which utilize quick network transactions for cache creation/state maintenance. All these characteristics of Memcached solutions are satisfied by VMA to provide the following:

- Higher Packets per Second (PPS) rates than with kernel
- Lower data distribution latency
- Lower CPU utilization by the multicast consuming/producing application in order to support further application scalability

Typical Memcached server performance is 50-100K transactions/sec. Few Web 2.0 companies have an optimized version which can support 200-300K transactions/sec. Since the client to server ratio is anywhere between 1:8 to 1:100, server is the bottleneck in this setup. With VMA, the transport is accelerated to support >500K transactions/sec. This is achieved by taking the following steps with VMA and no additional changes are required in Memcached solution.

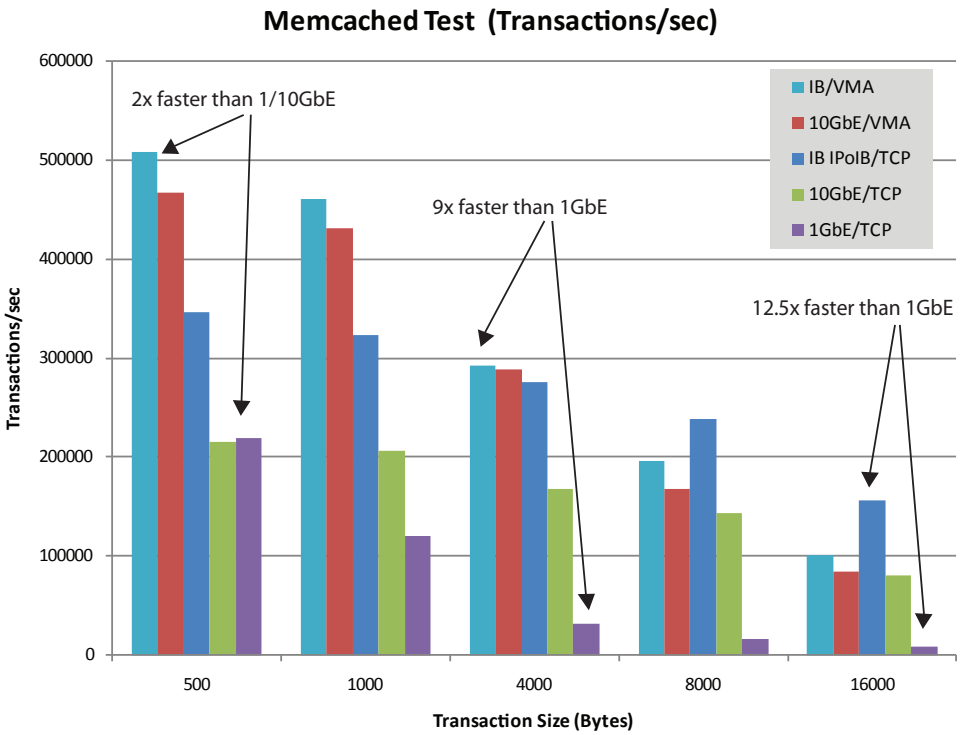
- Reduced number of context switches
  - o Allows the server to process a significantly higher packet rate than would otherwise be possible
- Minimal buffer copies
  - o Data is transferred from the hardware (NIC) straight to the application buffer in user space, with only a single intermediate user space buffer and zero kernel I/O buffers
- Fewer hardware interrupts for received/transmitted packets
- Fewer queue congestion problems witnessed in standard TCP/IP applications
- Supports legacy socket applications – no need for application code rewrite
- Maximizes Messages per second (MPS) rates
- Minimizes message latency
- Reduces latency spikes (outliers)
- Lowers the CPU usage required to handle traffic

VMA for Memcached solution runs on Linux currently and is interoperable with other OSes such as Windows/Linux/VMware/Solaris for non-VMA receive/send side of connectivity.

VMA for Memcached Solutions Performance Results

Transactions per second and latency was measured for set configuration where client accesses Memcached server by several network solutions such VMA using InfiniBand, VMA using 10GbE, traditional 10GbE and GbE.

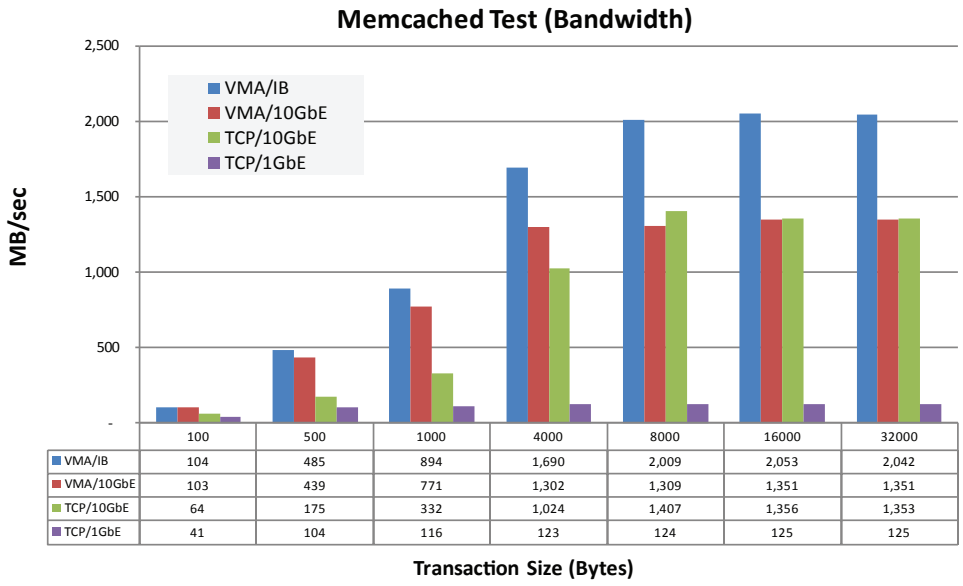
Transactions per second:



It can be seen that VMA with InfiniBand provides the highest transaction per seconds compared to other solutions. Mellanox 10GbE with VMA outperforms traditional 10GbE infrastructures by 2X on smaller packets.

Bandwidth:

TP/s using 2 processes per memcache server.

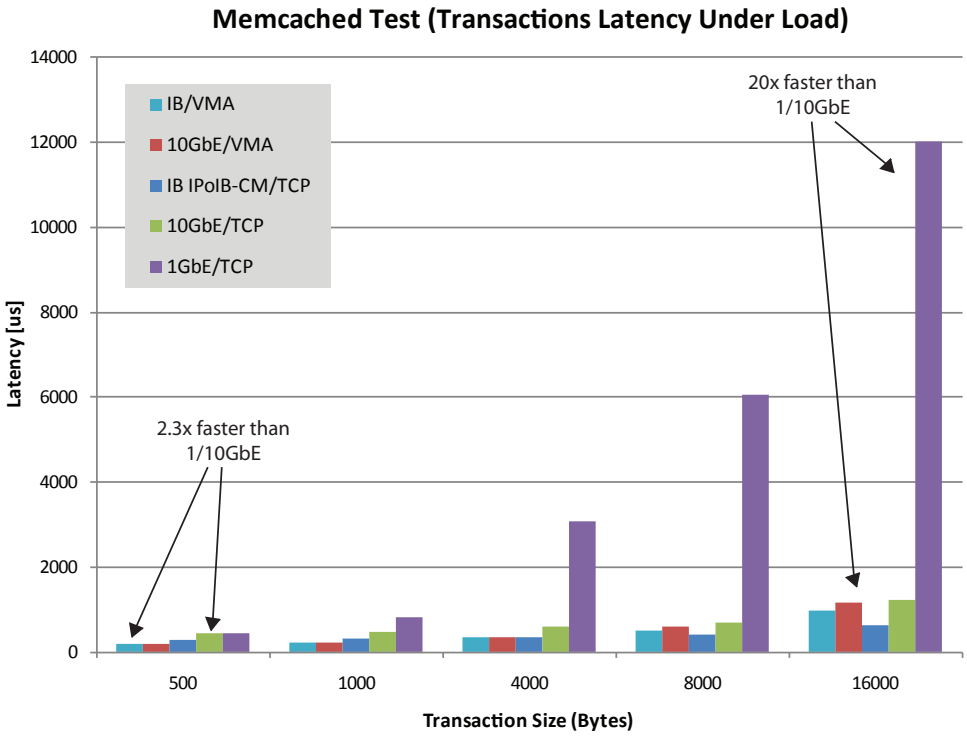


	100	500	1000	4000	8000	16000	32000
VMA/IB	104	485	894	1,690	2,009	2,053	2,042
VMA/10GbE	103	439	771	1,302	1,309	1,351	1,351
TCP/10GbE	64	175	332	1,024	1,407	1,356	1,353
TCP/1GbE	41	104	116	123	124	125	125

Highest bandwidth was achieved with InfiniBand infrastructure and VMA compared to 10GbE and TCP.

Latency:

Latency for Memcached test showed that VMA with IB is 2.3X faster than traditional 10GbE. Memcached solution users can get a response time 20 times faster to their users by using VMA without any modification to their existing Memcached solution.



Conclusion

Mellanox's VMA and InfiniBand or Ethernet solution can accelerate response time, bandwidths and transaction per second for existing and future Memcached solutions without any modifications.

350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com