## Open Fabrics Conference in Monterey, California

#### OFED use in the financial industry

Christoph Lameter, Ph.D.

### Overview

- Various use cases in Finance
- Banks (Investment, Retail)
- Stock Exchanges
- Investment Funds
- High Frequency / Algorithmic Trading



## Investment Banks / Exchanges

- Fairness issue for Customers/Investors.
- Reliability is a must.
- Real Time kernels frequently used.
- Fast yes but fair first.
- Hesitant to adopt Infiniband Technology. Mostly use of 10G Ethernet technology
- Use of 40G Infiniband mainly due to bandwidth concerns.
- Simulations.
- HPC style usage patterns



# Algorithmic Trading

- Speed is the all important issue
- Use of Infiniband due to lower latency.
- Kernel bypass frequently used. The IB verbs API is a nice offload protocol for this purpose.

Institution: investors

Funds

Program trading

Black has

Block

Trading

trading

Core execution

- Small latencies introduced by the OS and the environment are a problem.
- Infiniband competes with special 10G NICs (Solarflare, Chelsio).

## Multicast Support

- Special concern for financial industry
- Market data is distributed via Multicast
- Must be effective for short reaction times.
- Large number of listeners that could not be handled with unicast distribution.



## Messaging Middleware "Making Multicast reliable"

### • Products

- 29West
- Tibco FTL
- IBM Websphere LLM
- MAMA/OpenMAMA
- Redhat: AMQP
- Approach
  - Numbering of packets (Sequence ID)
  - Receiver asks for retransmission of missing packets.
  - Packet Storm problems
  - Standardized via RFC 3208 (PGM).

## Multicast Challenges in Infiniband Fabrics

- Backpressure can slow everyone down
- Unique stress on fabric. "Nonblocking" fabrics begin to block.
- Packet delivery is reliable within the fabric
- But packet loss occurs
  - Before entry: Due to the fabric slowing down the sender
  - On exit: Because the NICs will silently drop packets for which no Infiniband buffer is available.

## Comments on OFED

- OFED should not exist separately. Functionality should be standard in the kernel without an extra build step.
- If OFED continues to exist then there needs to be an easy way to build the stuff. OFED builds are based on a custom kernel source that is not upstream. Difficult to patch and contriibute to.
- OFED in its current source form is a headache that needs to go away.
- Progress has been made in merging most of OFED into upstream kernel sources (thanks, Roland) but some pieces are still missing.

## **Offload API Standardization?**

- There is no standard for a off load technology
- OFA software could set the standard for such an endeavor
- API needs to be generalized to work with arbitrary devices
- Less cryptic
- Documentation
- Can the POSIX standard be expanded?
- Can it become even faster?



## Combination with OS low latency

- Problem of OS noise that is relatively long in latency compared to the transfer speeds of Infiniband Host to Host.
- The dynticks patchset could address some of that.
- Does Infiniband work well under realtime OS approaches?





## Conclusion

- Recommend one or several strategies
- Give a summary of the expected results
- Name the next steps to be taken
- Delegate the various tasks