



Datacenter Fabric Workshop



InfiniBand Midlayer

Roland Dreier, Sean Hefty, Hal
Rosenstock



Agenda



- Communication Manager (CM)
- RMPP
 - _ Reliable Multi-packet Protocol (large MADs)
- Userspace verbs



MAD and CM



- CM in kernel 2.6.13
 - _ both kernel and userspace supported
 - _ IBA 1.1/1.2 compliant
- RMPP in kernel 2.6.13
- Address Translation (AT) in OpenIB subversion tree
 - _ not upstream



Userspace verbs



- Userspace verbs support in kernel 2.6.13
 - _ 32-bit userspace on 64-bit kernel supported
 - _ Shared Receive Queues (SRQ) in 2.6.14
- Hardware support:
 - _ All Mellanox HCAs (incl. mem-free)
 - _ PathScale HTX and IBM eHCA pending



Userspace libraries



- Userspace libraries in OpenIB svn
 - _ libibverbs: generic verbs API
 - _ libmthca: device-specific plug-in for Mellanox HCAs
 - _ libraries pending for PathScale, IBM HCAs
- Applications link with libibverbs; device-specific libraries loaded by libibverbs
 - _ Apps are device independent.



Design Notes



- Thread-safe
- Cleans up after app crashes
- Completion events and async events delivered via `read()` on file desc.
 - _ Apps can use threads, `poll()`, `SIGIO`, ... to multiplex events



User Verbs Future Work



- `fork()` support
 - _ `system()` probably works fine now
- Multiple CQ event handlers with MSI-X
- Hotplug support
 - _ gracefully fail when device is removed
- iWARP convergence



Questions?



???