RDMA stacks, migrating from older versions to newer

Christoph Lameter, Ph.D.
Gentwo
March 19, 2015, Monterey, CA
Introduction

- At some point one needs to migrating to newer RDMA software
- OFED 1.5.4 -> Redhat 7
- Not complete.
- Testing Lab experienced various issues
- Partially in production environment.
- RDMA software in house
RDMA Software Stacks in the game

- Mellanox OFED 1.X and 2.X
- OFED
  - 1.X
  - 3.X
- Redhat
  - 6.*
  - 7.0
- Ubuntu
  - 10.04
  - 12.04
  - 14.04
- Upstream Linux
  - 2.6.32
  - 3.10, 3.14
  - 3.19, 4.0
Migration issues

- Code does not compile due to API changes
- Features missing in some stacks
- Failures during link
- Versioning issues
- Binary compatibility issues
- Third party binary kernel modules
OFED, MOFED source issues

• Building from source should allow combination of features that one needs while mitigating the risk and the deviation from upstream.
• OFED 1.X has patches (but strange base)
• OFED 3.X changed approach. Has modified tree in there but git archives are publicly available.
• MOFED 1.X also has patches based on OFED 1.X
• MOFED 2.X has modified tree but no patches or publicly available git archive.
Library issues

• Subtle dependencies. Frequent breakage because of library / kernel issues.
• ABI and API screwed up with vectors of functions etc. Weird indirect calls.
• Multiple approaches to “solving” the issue.
• Established ways of ensuring binary compatibility are worked around.
• The higher you go up the stack the less likely breakage becomes because the linker convention and ABI conventions are followed.
Ideal world

- One source tree that is continually moving forward
- Less breakage
- Everyone follow upstream
- New features as git branches on kernel.org that are reviewable
- No strange OFED archives with strange patches please
- Experimental release contain all patches and refer to a definite point in the upstream tree. Patches are applicable against that tree.
OFA software devel issues

• Minimal interaction with the Linux ecosystem. Staying within academic communities.
• Not communicating with the kernel community.
• Reinventing the wheel, non standard solutions
• Conventions are not followed
• Misunderstanding how the community works.
• Result is a very fragile ecosystem
• For us this requires an extremely high support effort.
Questions
Discussions
and
Thank You

OpenFabrics Software
User Group Workshop

#OFSUserGroup