14th ANNUAL WORKSHOP 2018

RDMA CONTAINERS THINK TANK

Doug Ledford/Jason Gunthorpe

[April 13, 2018]

WHO PARTICIPATED

- **People from the upstream linux community (too many to name individually)**
- **From the consumer side, we had representatives from:**
  - Microsoft/Azure Cloud
  - NASA Ames/Charlie Cloud
  - DoD
  - OSI

WHAT DID WE DISCUSS

- **What requirements do people have in terms of container needs?**
  - Most people need the same basic things
  - Multi-tenancy was a strong need from everyone
  - Access to RDMA resources from user space apps and not just the kernel supplied ULPs was also big
  - Fine grained tracking of resources on a per host basis is needed
  - Tracking of security related issues inside the container is needed
  - Device level enforcement of QoS

WHAT DID WE DISCUSS (CONT)

- **Where are we now?**
  - You can create normal containers that use unconstrained host provided RDMA features (like IPoIB or NFSoRDMA)
  - You can create constrained containers using SRIOV devices that you lock to a namespace (for RoCE/iWARP devices)
  - Nothing else that you might want to do works today

WHAT DID WE DISCUSS (CONT)

- **Things to do**
  - Since the RDMA subsystem extensively uses sysfs files for device lookup and configuration lookup, the sysfs files provided by the RDMA subsystem needs to be made namespace aware
  - Finalize how we will do namespace boundary elements for non Ethernet based RDMA devices (LID/GID/P_Key tuple is the current thinking, but that may change)
  - Disable all RoCE_V1 access from within a container as the link layer GID can't be constrained and will break any containerization
  - Probably just preserve and use the current net device namespace semantics on RoCE_V2 and iWARP devices
  - Possibly look at using JKeys for PSM devices
  - Try to work on getting IPoIB working. Possibly requiring a modification to the on-wire rdmacm protocol :-(
  - There was a proposal to make RoCE devices use the semantics of IB devices instead of their parent net device semantics…this will need more discussion

14th ANNUAL WORKSHOP 2018

THANK YOU

Doug Ledford/Jason Gunthorpe