The Open Fabrics Verbs Working Group

Pavel Shamis and Liran Liss

Introduction

- Verbs is a high-performance mature and robust interface
  - Widely adopted
  - Stable
  - Scalable
  - Standards-based
  - Interoperable

- Open-source development
  - Peer-reviewed code patches
  - Peer tested

- Development rate and novelty constantly increasing
  - New applications and features
  - New systems and architectures

OFV Mission

- Enhance the Verbs interface to meet the needs of future systems and applications
  - HPC, storage, cloud applications and more
  - CPUs, GPUs, and other compute elements
  - RDMA and other interconnect accelerated capabilities

- Guide the development of the Verbs eco-system
  - Innovation!
    - Forum to discuss new ideas and approach for new features
  - Raise community awareness to trends and requirements
  - Partner with collaborators for introducing new features

The Verbs Eco-System

- **Extended Verbs**
  - Enable new features

- **Infrastructure**
  - Common services
  - Common abstractions

- **Middleware and ULPs**
  - Application-facing APIs
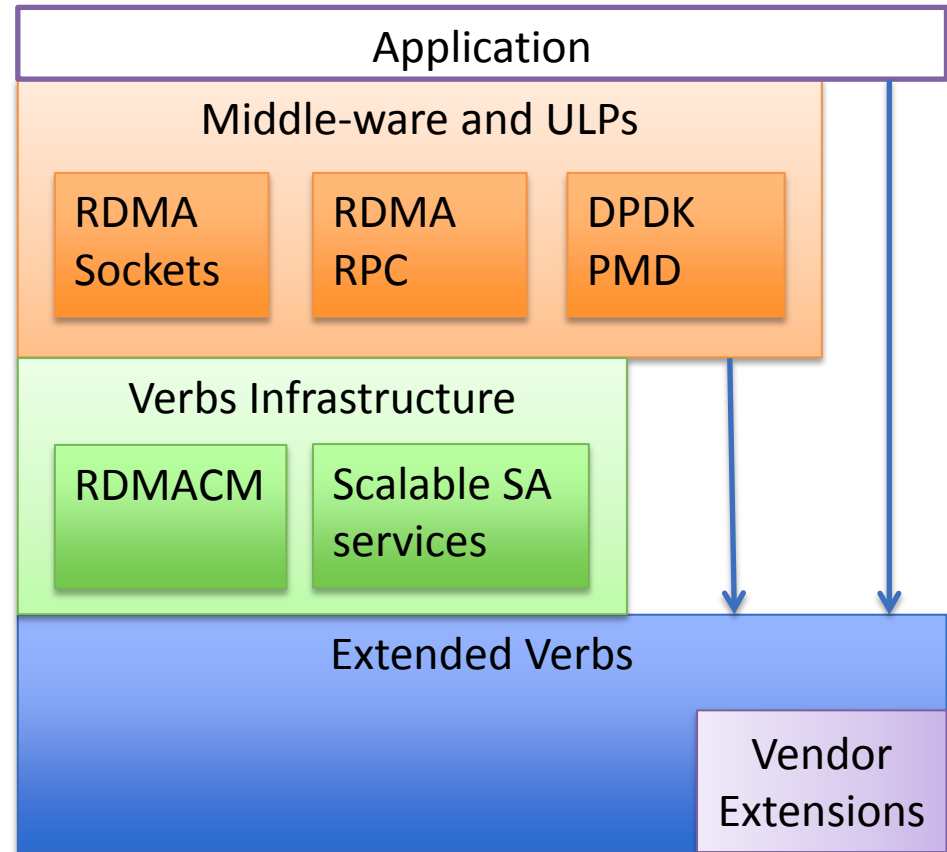
The Verbs Eco-System: Extended Verbs

- **Accelerated Verbs**
  - HPC
  - Packet processing

- **Raw Ethernet support**
  - Rx flow steering
  - Tx flow anti-spoofing
  - VLAN stripping
  - RSS, TSS
  - Tunneling (VXLAN, NVGRE)

- **Memory**
  - Indirect Memory Regions

- **Storage support**
  - Signature offloads
  - RAID offloads

- **Virtualization support**
  - Namespace control
    - QPNs, CQNs, MRs, etc.
  - QP suspend/resume

- **Time stamp operations**
  - Read HCA clock
  - Time stamp completions

The Verbs Eco-System: Verbs Infrastructure

- RDMACM APM support

- On-Demand-Paging

- Scalable address resolution

- SR-IOV

- Container support

- Multi-path RDMA

- SoftRoCE

- Improving connection rate

- SELinux support

- Kernel-managed user-space QPs
  - Connection management
  - fork() support

The Verbs Eco-System: Verbs Middleware

- RDMA accelerators
  - E.g., Hadoop, Ceph

- Packet processing
  - E.g., DPDK Poll-Mode Driver

- RDMA RPC

- RDMA sockets

- User-space TCP/IP

- Message queuing
  - E.g., ZeroMQ RDMA support

- AIO
  - E.g., libevent RDMA support

- Java/Python bindings

Discuss New Approaches

- Solicit feedback from a larger community and users
  - Not all of our contributors track kernel mailing lists
  - Storage, Big-Data, HPC, etc.

- Discuss concepts before writing code or detailed [RFC] patches

- Converge faster by interactive feedback
  - Faster acceptance

- Tackle hard-to-crack concepts by focused discussions
  - Weekly increments

Community Awareness

- Raise the need for important features
  - Introduce new use-cases
  - Describe feature requests
  - Aid in prioritizing the focus of the development community

- Point out pain points and urgent issues
  - E.g., connection rate, "debugability"

Collaboration

- Raise issues that you would like to work on

- See if other community members are currently tackling the same issues

- Establish collaboration to speed up development and increase efficiency

- Cross-community collaboration
  - Discuss challenges with broader community

OFVWG Backlog Planning

- Backlog items
  - Clear title, description, and scope
  - Determine priority
    - High – high impact and/or working on this now
    - Med – medium impact and/or working on this in the near future
    - Low – not a major feature or not in near term plans

- Revisit backlog each meeting
  - Propose new items
  - Extend discussion over existing items
  - Reprioritize and arrange backlog accordingly

- Harden agenda for next 3-4 meetings
  - Get commitment from owners

OFVWG Work has Started

- 2/10 – First OFVWG meeting
- 2/17 – OFVWG procedures and initial backlog
- 2/24 – Accelerated Verbs Framework
- 3/3 – Verbs Extensions Framework
- 3/11 – Verbs Extensions Framework (cont.)

Current Backlog

Meeting order	Item	Effort estimation (weeks)	Priority (high/med/low)	Owner
1	HPC Accelerated Verbs	1-2	High	Mellanox
2	RoCE address management	1	High	Mellanox
3	Scalable SA overview	2-3	High	
4	SoftRoCE	1-2	High	
5	QP packet pacing	1	High	Mellanox
6	Fast connection management	2-3	High	Mellanox
7	Standardize Verbs error codes	1-2	High	
8	APM support for RDMACM	1-2	Med	Mellanox
9	MP-RDMA	3-4	Med	Mellanox
10	Verbs versioning API	2-3	Med	
11	Verbs capability query framework	1-2	Med	
12	Ethernet APIs	3-4	Med	

Summary

- The pace of Verbs development is increasing
- The user base of the Verbs Eco-system is expanding
- The OFV WG will facilitate
  - Agreement on approach for new fronts
    - Faster acceptance and convergence
  - Prioritize and focus OFA development efforts
  - Encourage collaboration
  - Broaden the use of the interface
  - Increase adoption of RDMA technology
- Join us!
  - 11:00AM PST on Tuesdays:
    https://mtlmeet.mellanox.com/shainer/27GVHTN2
  - List: http://lists.openfabrics.org/mailman/listinfo/ofvwg
  - Download: https://www.openfabrics.org/downloads/ofv/

Thank You