



*Fabric Computing That Works*

# A Unified Approach to Networks

The OFED stack

Paul Grun

Chief Scientist

System Fabric Works, Inc.

[pgrun@systemfabricworks.com](mailto:pgrun@systemfabricworks.com)



# Abstract

RDMA is a message passing paradigm. Messages are independent of the I/O protocol. The OFED stack presents a standards-based message passing interface for kernel and user applications, independent of the underlying wire protocol.

In this session, we discuss how OFED implements a true unified fabric.

We will also look at the three transport protocols and discuss their usages.



# Message Passing

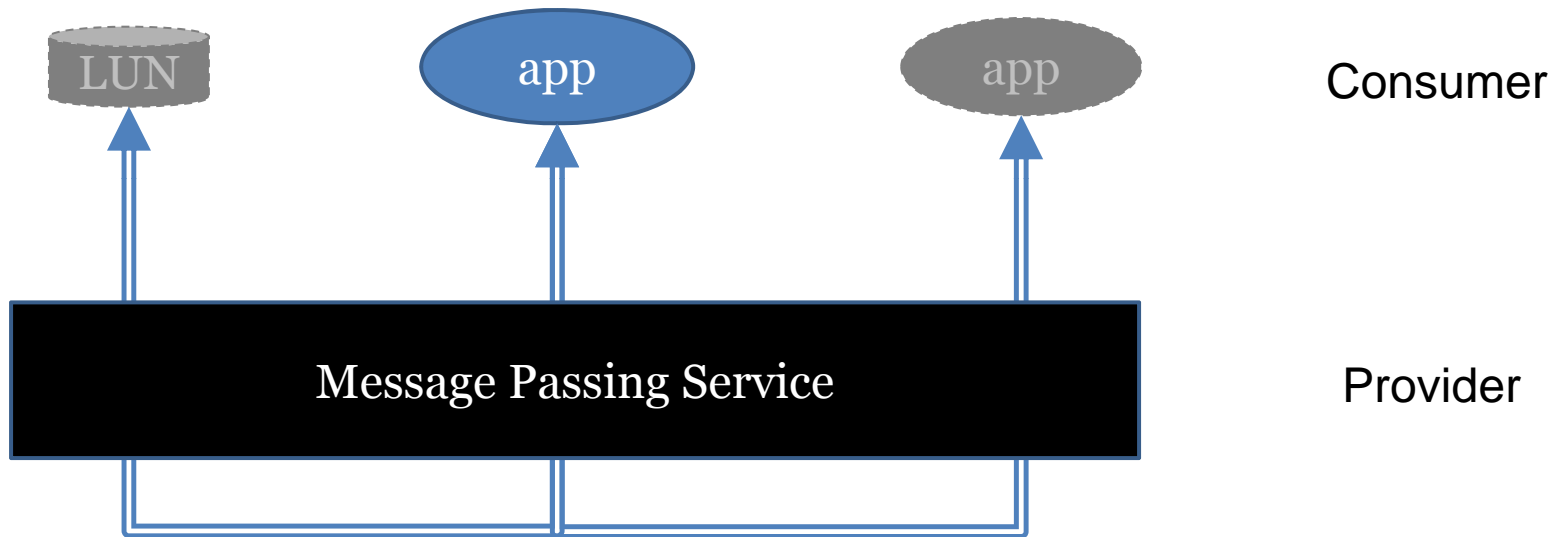


\***Message passing** ...is a form of communication used in [parallel computing](#), [object-oriented programming](#), and [interprocess communication](#). In this model [processes](#) or [objects](#) can send and receive messages ...to other processes. By waiting for messages, processes can also [synchronize](#).



*Fabric Computing That Works*

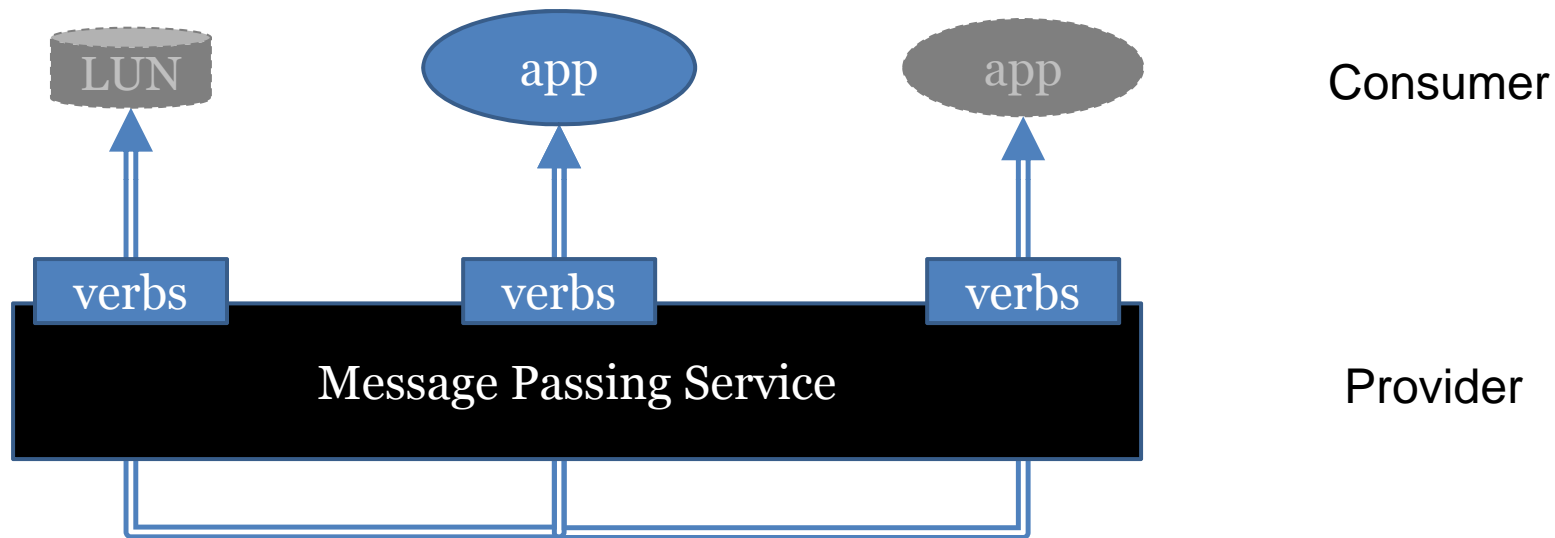
# A message passing service





Fabric Computing That Works

# OFED verbs API

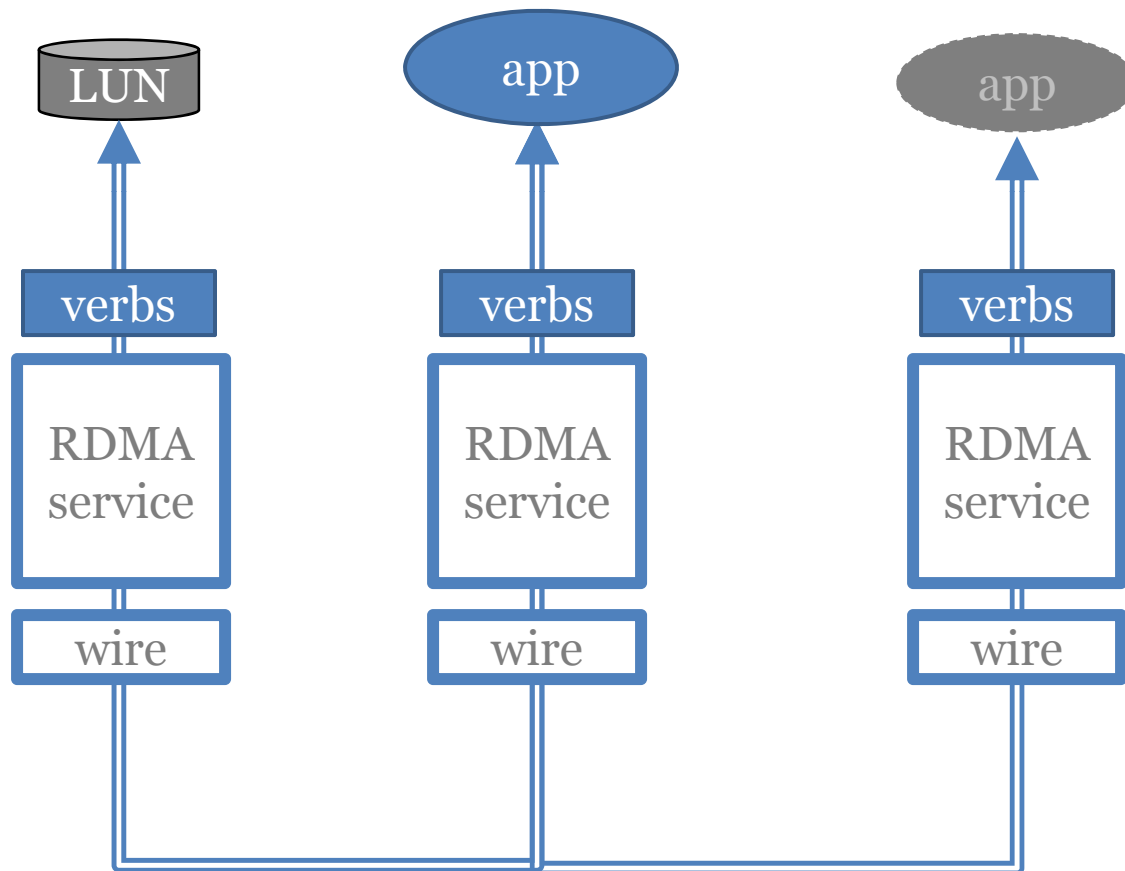


The OFED Verbs API is how an application accesses the message service



Fabric Computing That Works

# The RDMA Service

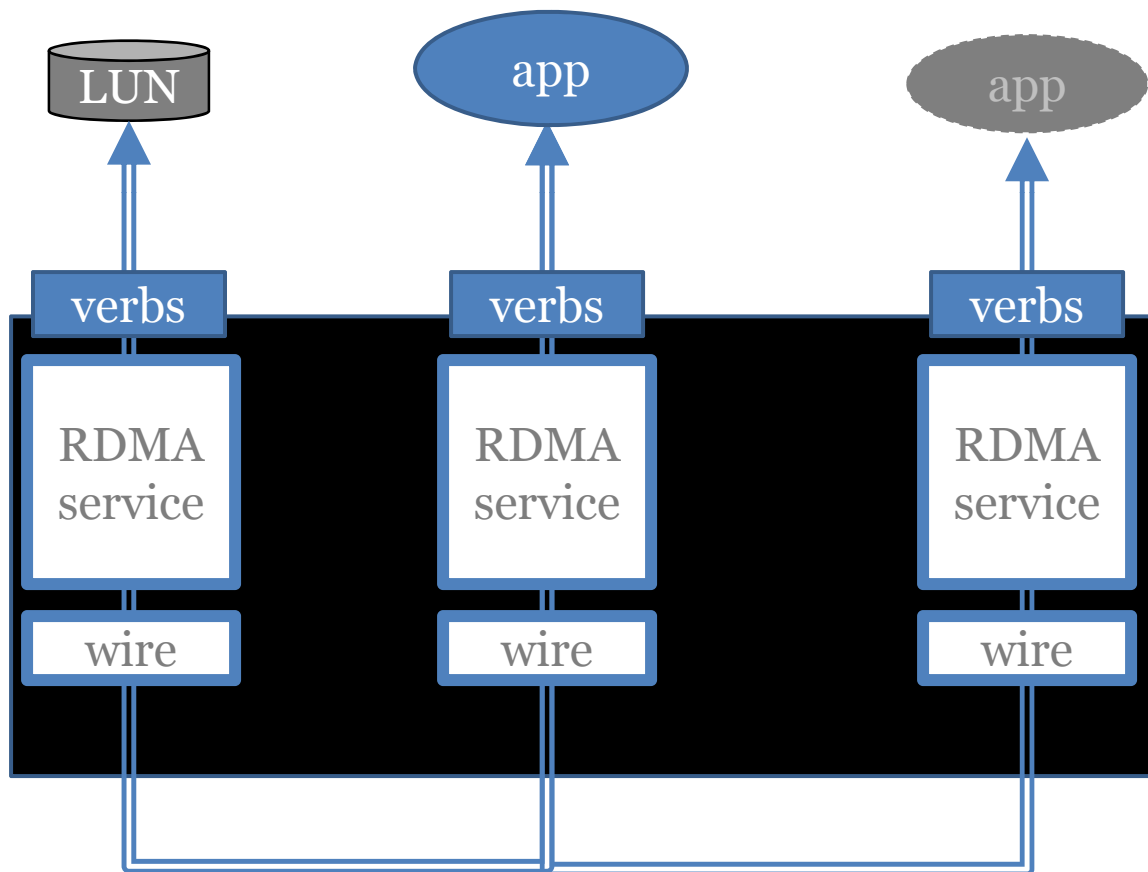


The RDMA service can pass messages across a network

RDMA Protocols  
Reliable Transport  
Physical Wire



# The RDMA Service

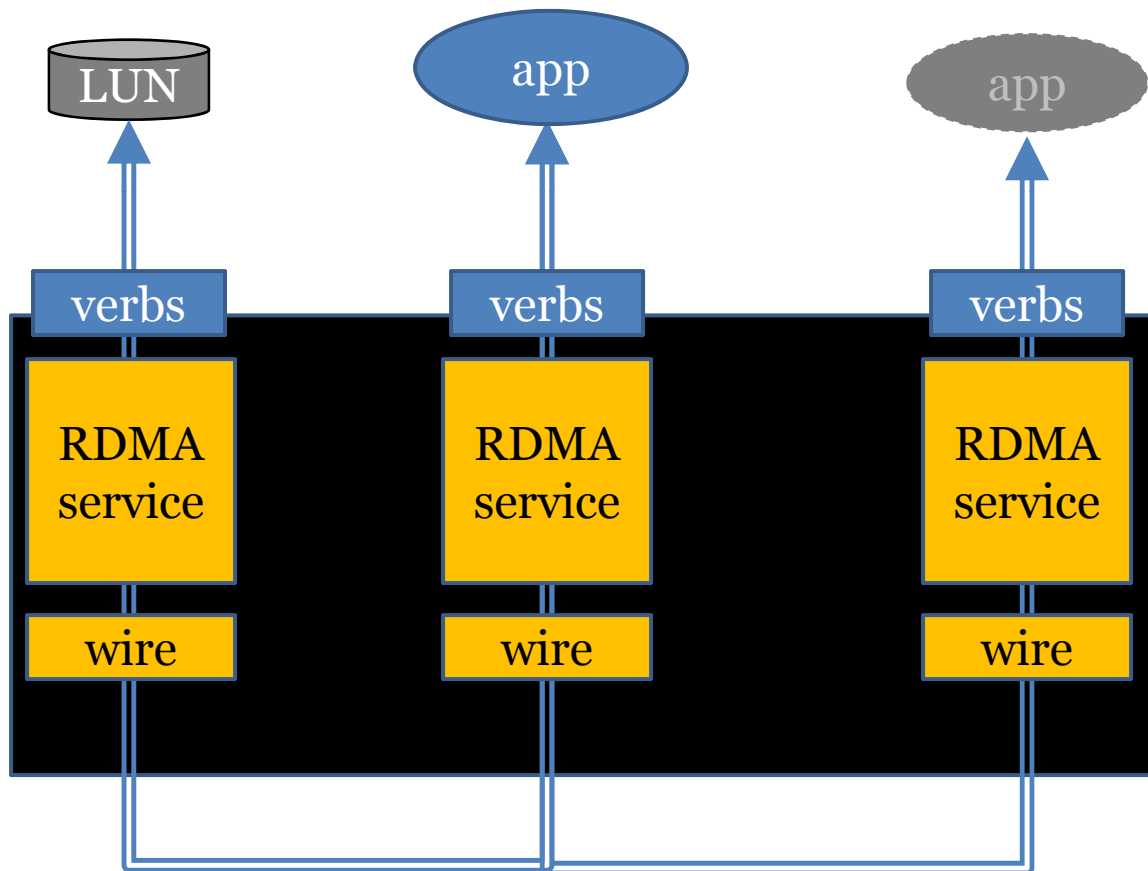


The provider appears to the consumer as a black box.

- a well known set of inputs and outputs (e.g. verbs),
- a well defined set of services



# The RDMA Service



This means we can easily change what is inside the black box...

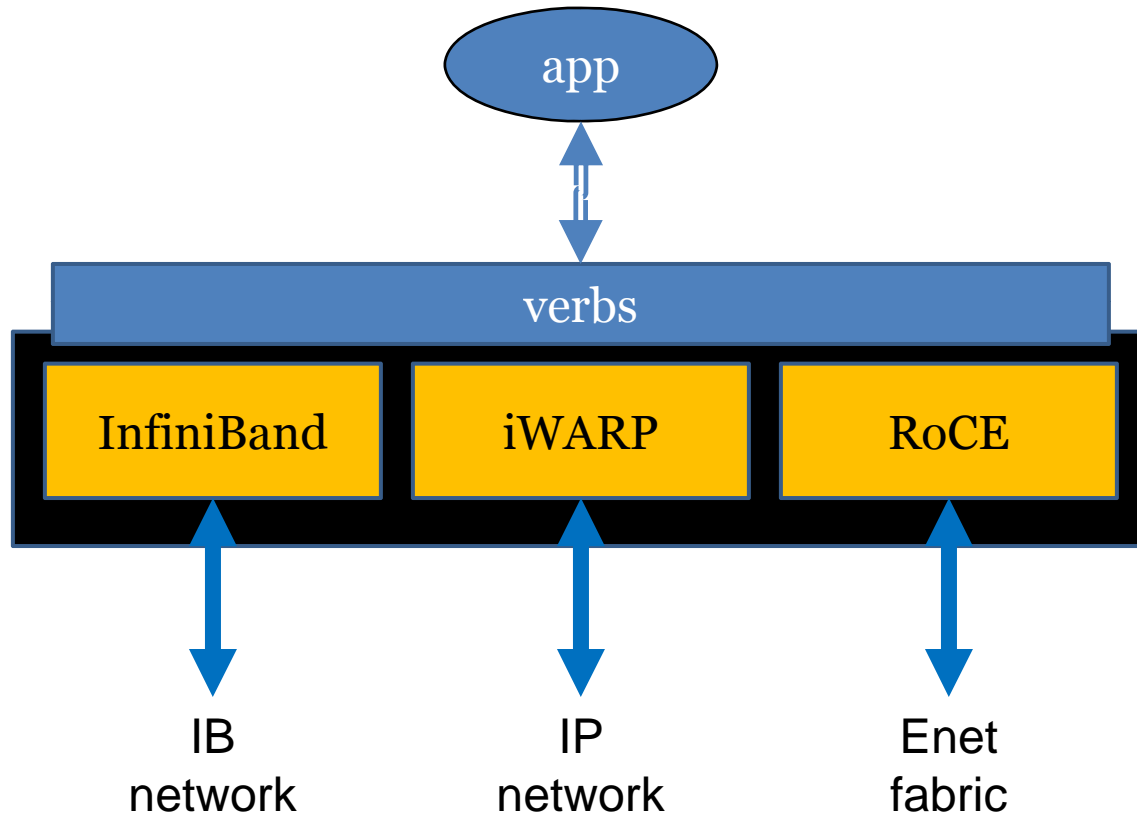
...with no impact to the consumer.





Fabric Computing That Works

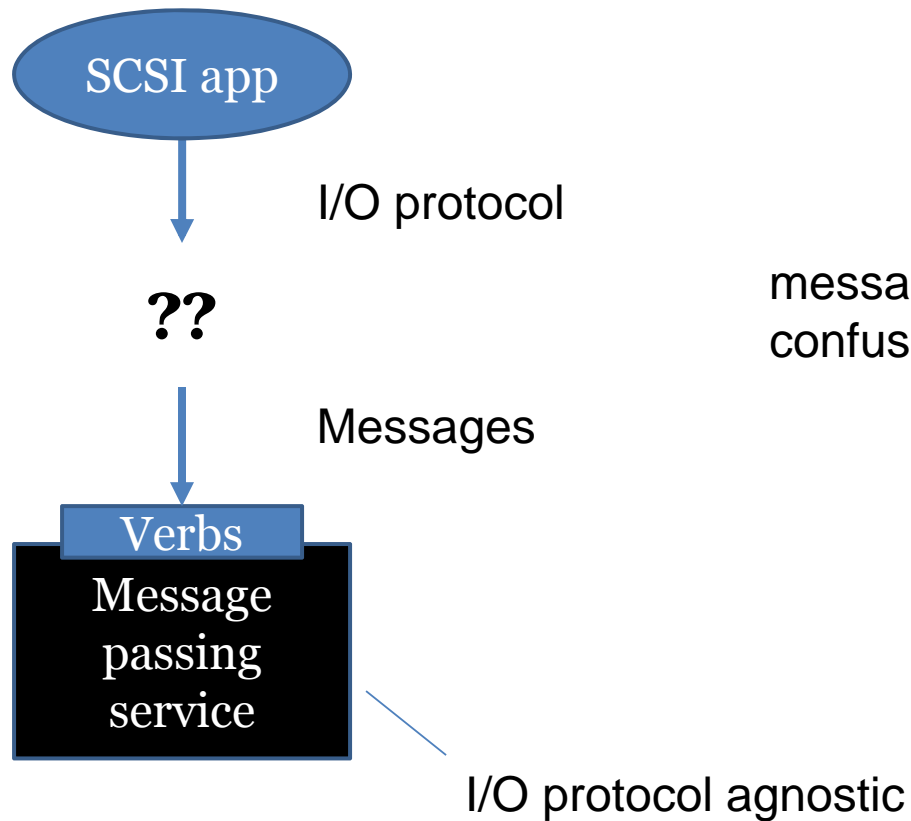
# The RDMA Service



OFED fully supports three wire protocols



# I/O versus message passing?

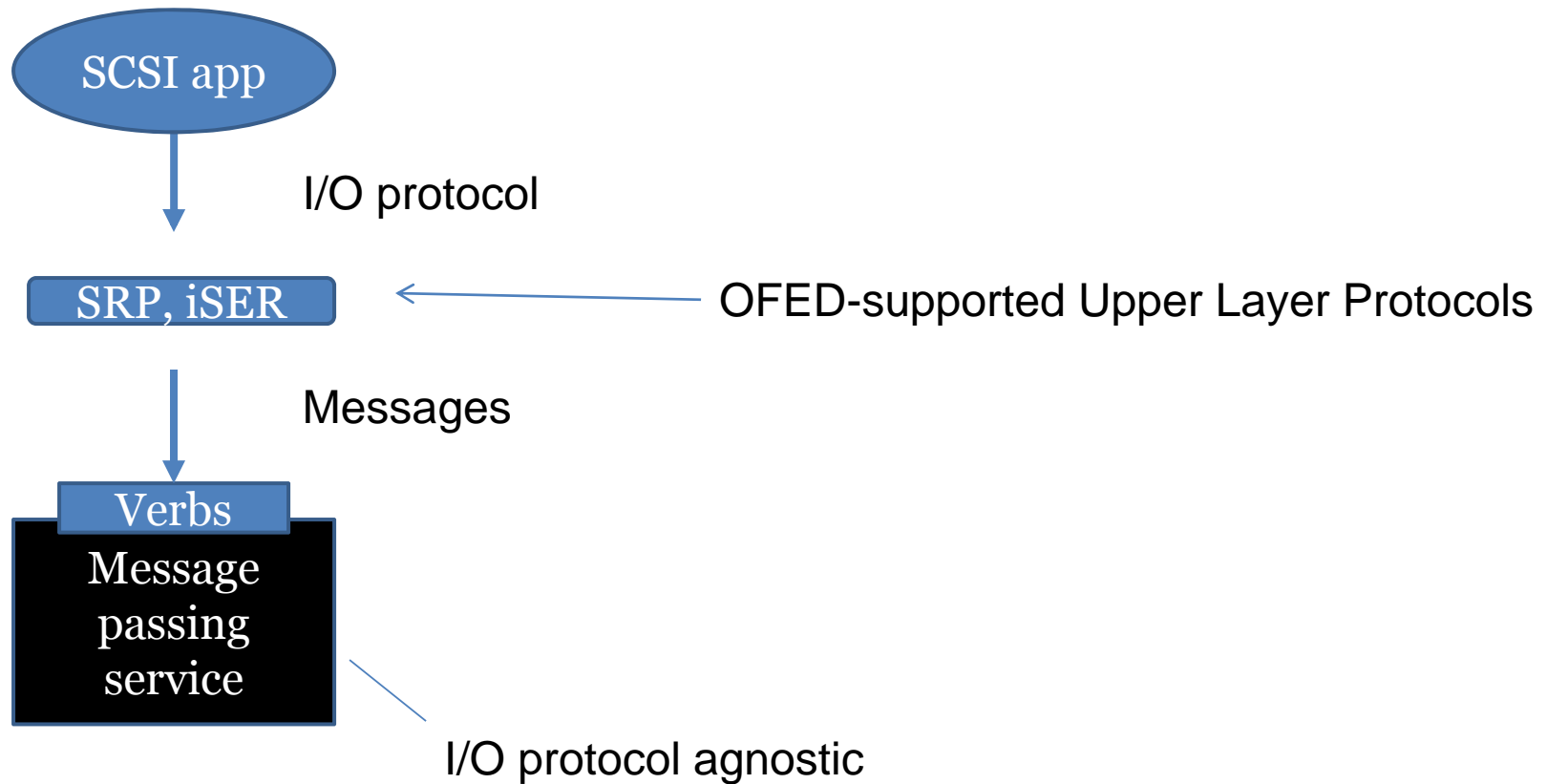


message passing shouldn't be confused with an I/O protocol



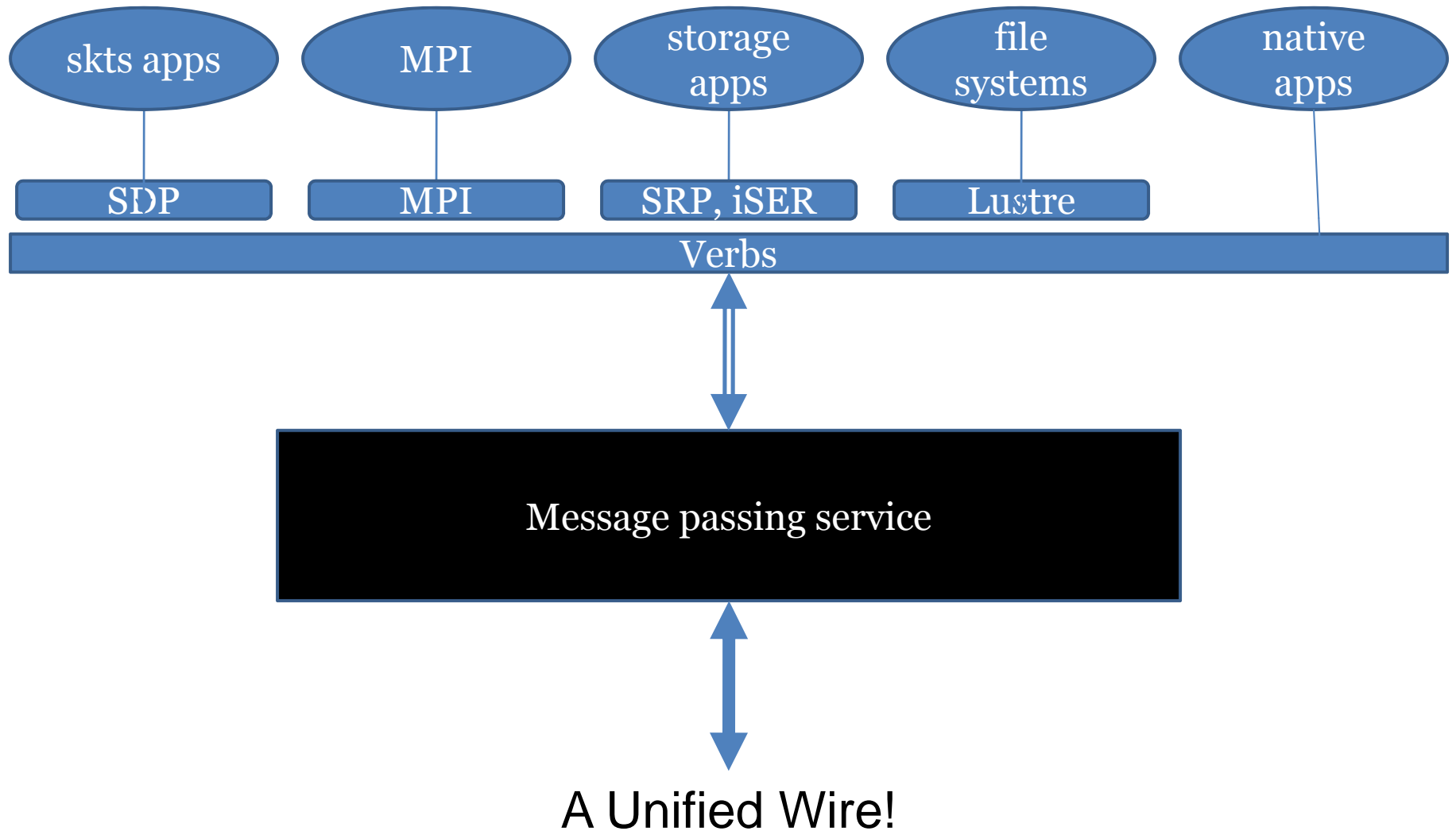
Fabric Computing That Works

# OFED upper layer protocols





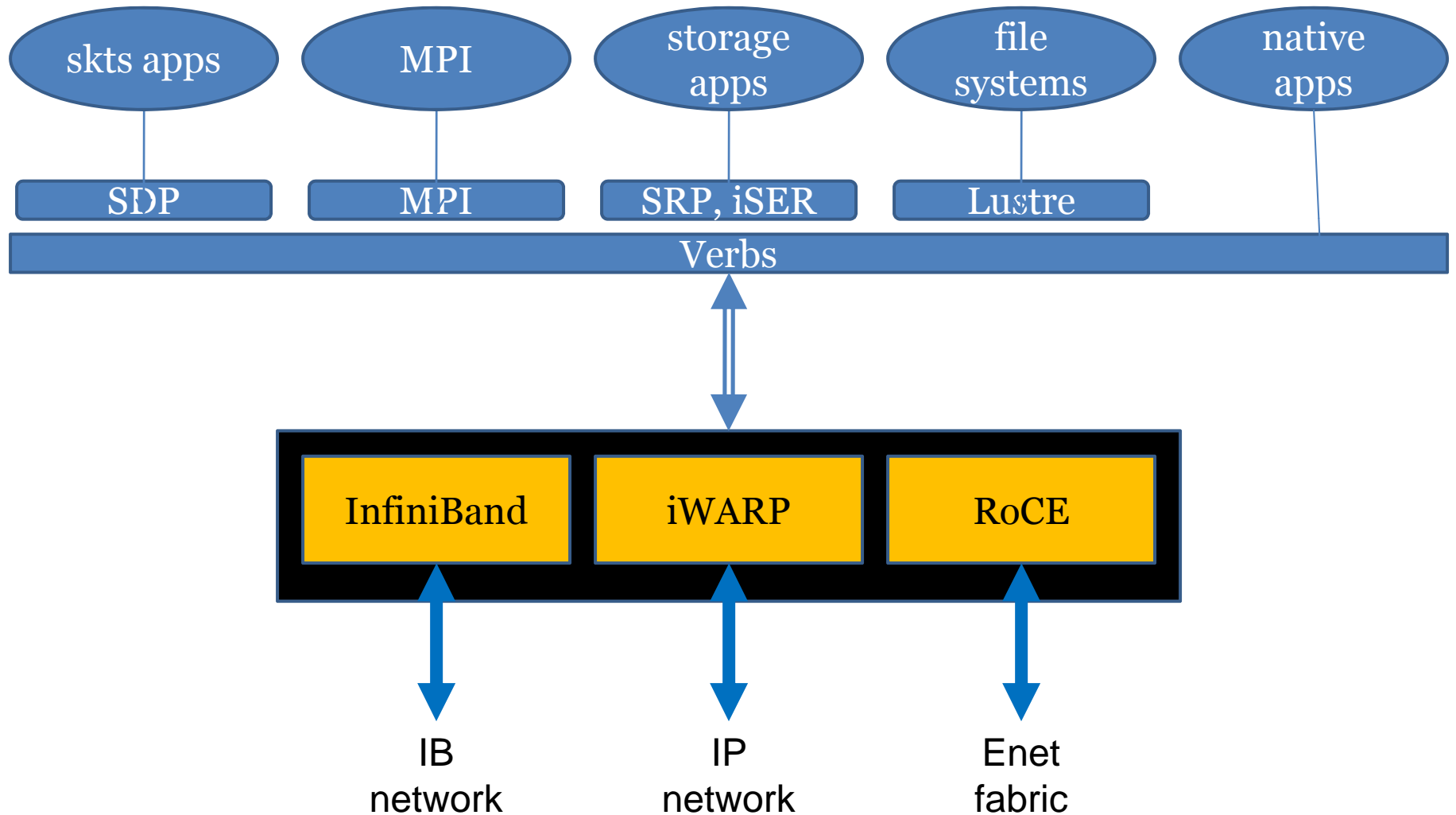
# A unified wire





*Fabric Computing That Works*

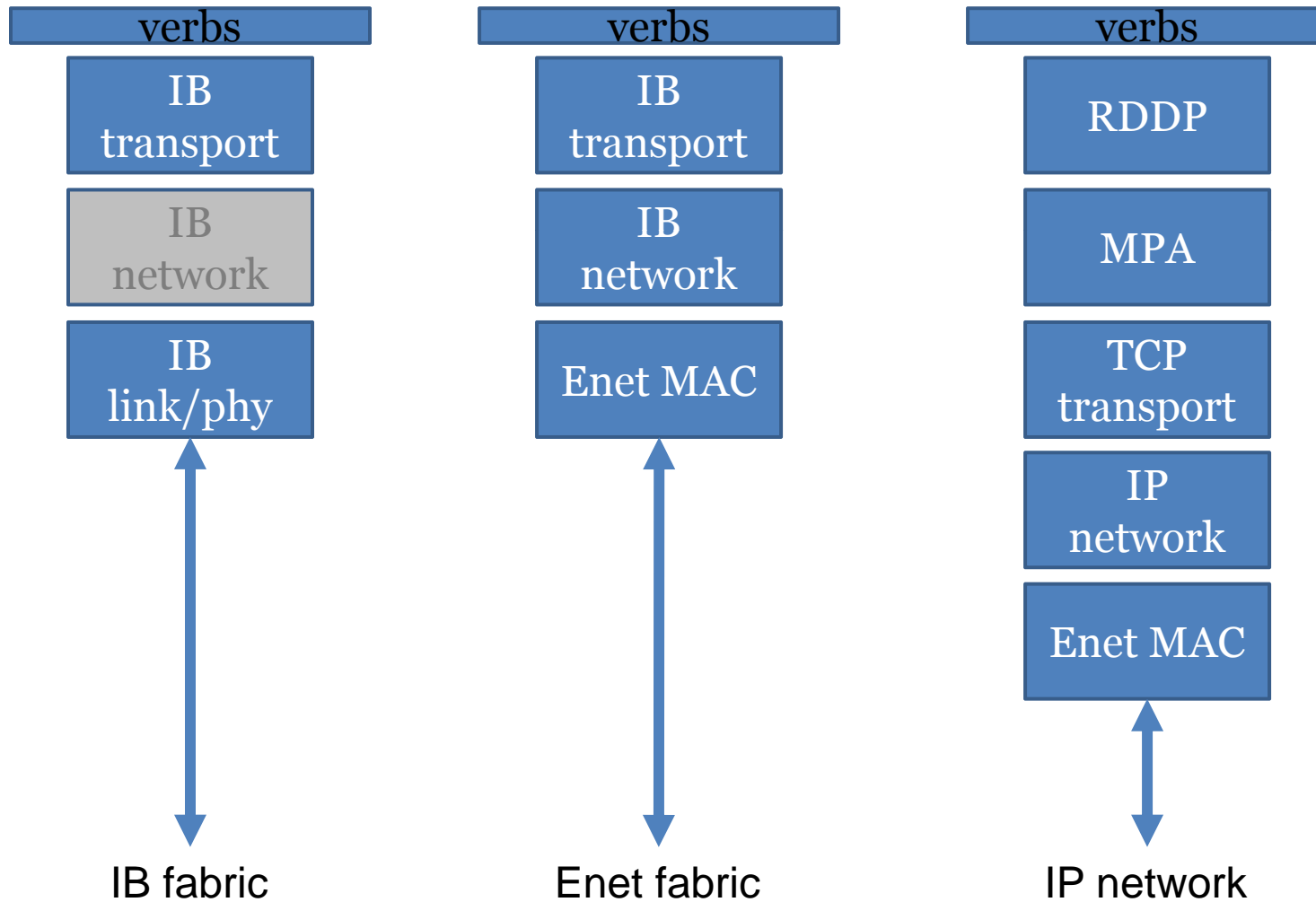
# Actually, a choice of 3 unified wires





*Fabric Computing That Works*

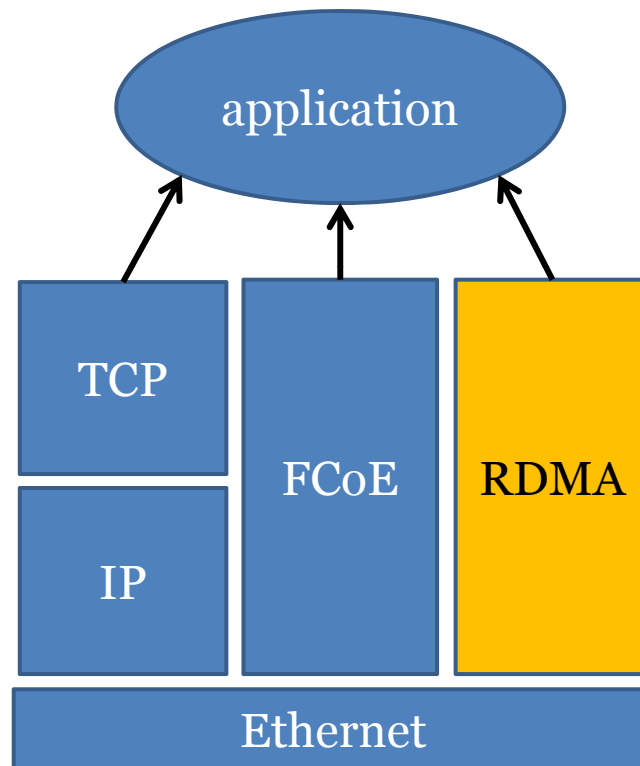
# 3 Transport Protocols





*Fabric Computing That Works*

## Why RoCE?



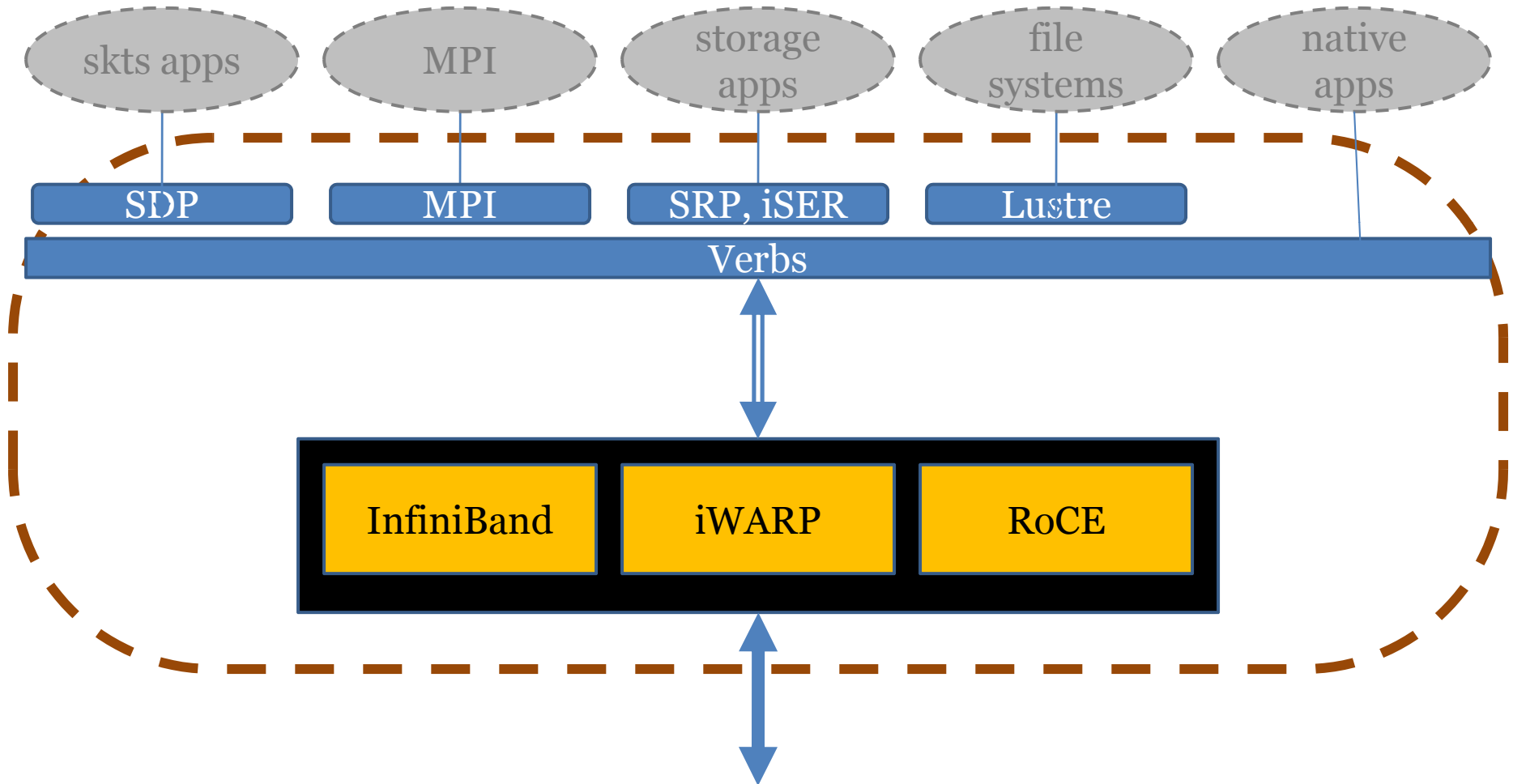
- Networking via TCP/IP
- Storage via FCoE
- Low latency via RoCE

This completes the Ethernet TrilogY – networking, storage, IPC



Fabric Computing That Works

# OFED – a complete solution







*Fabric Computing That Works*

- RDMA has been shown to deliver value propositions that are not available through any other communications paradigm
- These value propositions are compelling, particularly in the HPC space
  - Low latency, improved resource utilization, flexible resource allocation,, scalability, unified fabric...



*Fabric Computing That Works*

- RDMA as an excellent message passing service
- Verbs – a standards-based interface for passing messages
- Upper layer protocols: A true unified fabric
- Seamless support for three different wires



*Fabric Computing That Works*

# A shameless plug

Open Fabrics Alliance announces the January availability of a new course called

## Application Programming for RDMA

Targeted at application programmers who want to learn how to write applications for RDMA

Registration is now open:

<http://www.openfabrics.org/training/index.html>



*Fabric Computing That Works*

# System Fabric Works

System Fabric Works, Inc. delivers engineering, system integration and strategic consulting services to organizations seeking to deploy high productivity computing and storage systems, low latency high performance networks and the optimal software to meet our customer's application requirements. SFW also offers custom integration and deployment of commodity servers and storage systems at levels of performance, scale and cost effectiveness that are not available from other suppliers. SFW personnel are widely recognized experts in the fields of high performance computing, networking and storage systems particularly in OpenFabrics Software, InfiniBand, Ethernet and energy saving, efficient computing technologies such as RDMA.

[www.systemfabricworks.com](http://www.systemfabricworks.com)