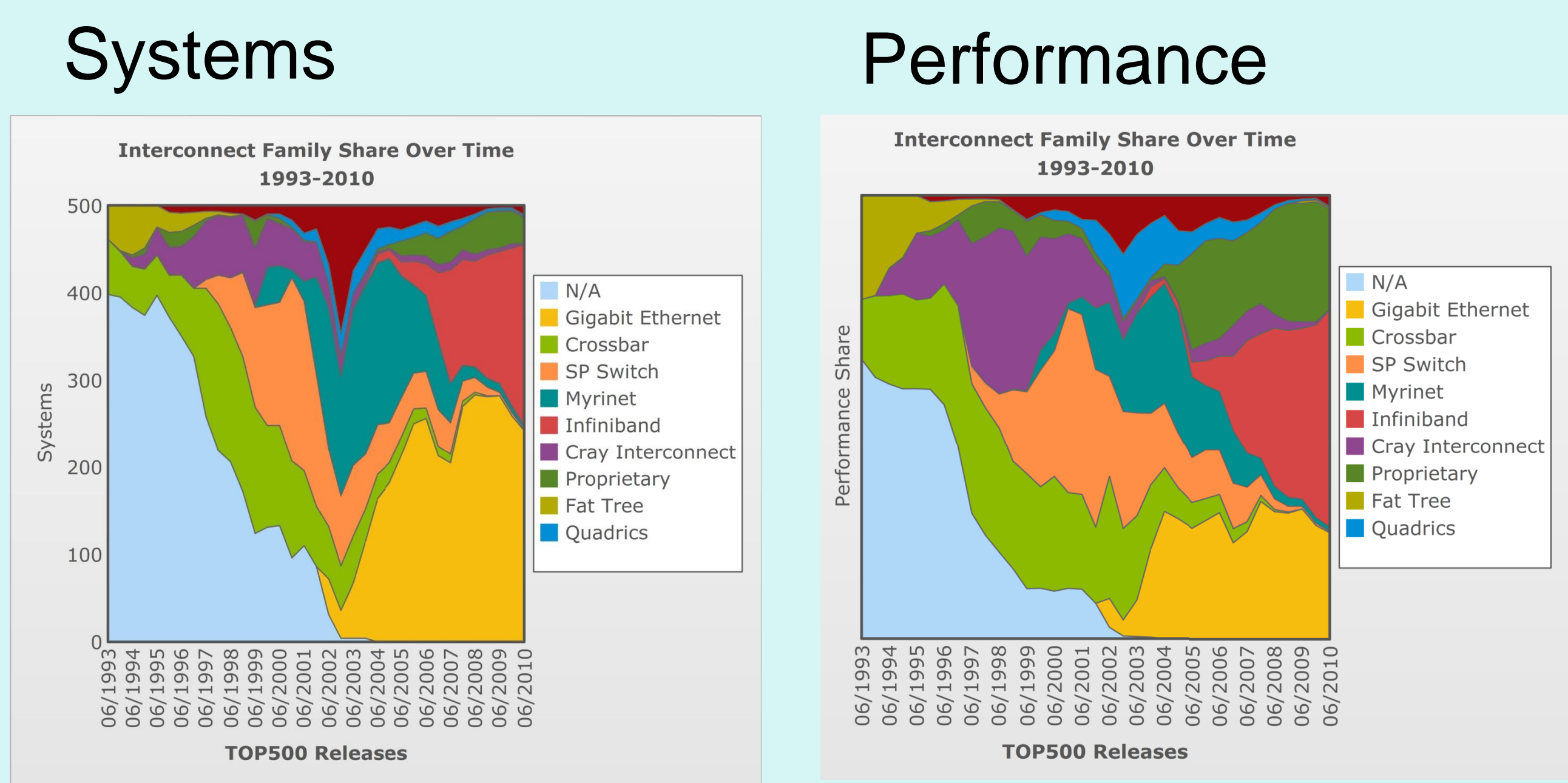
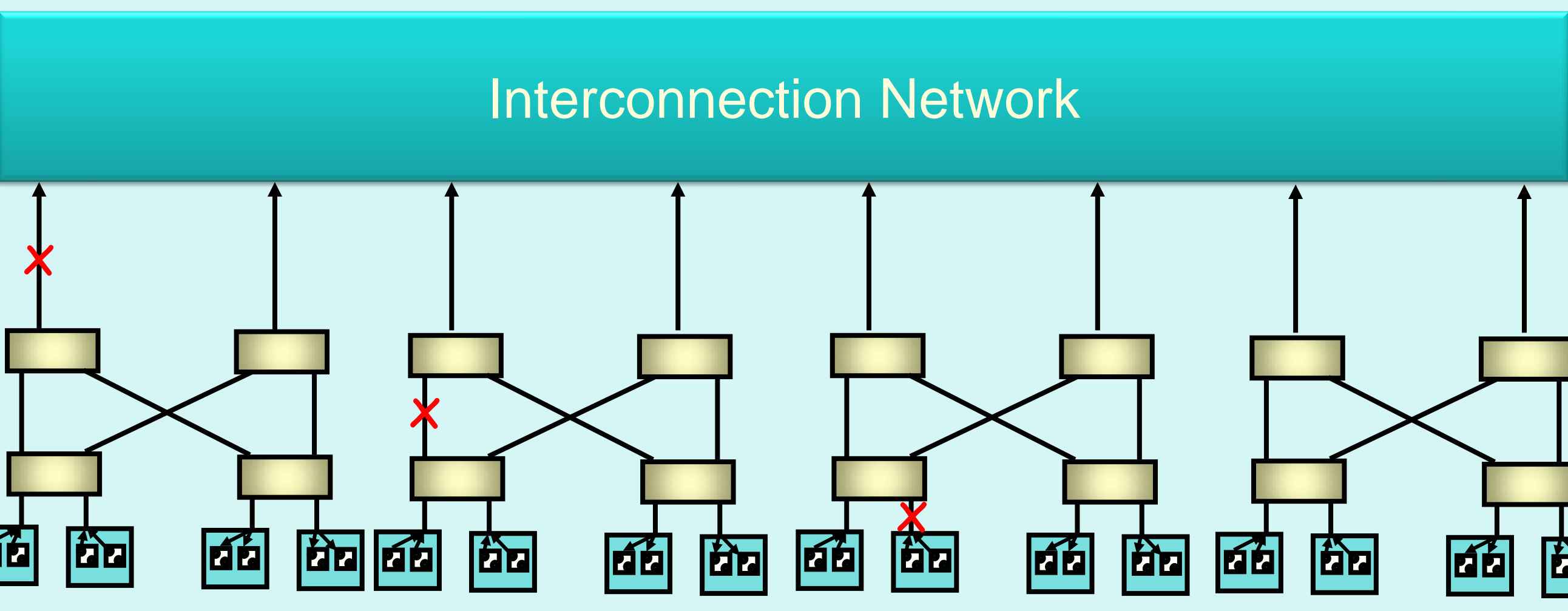


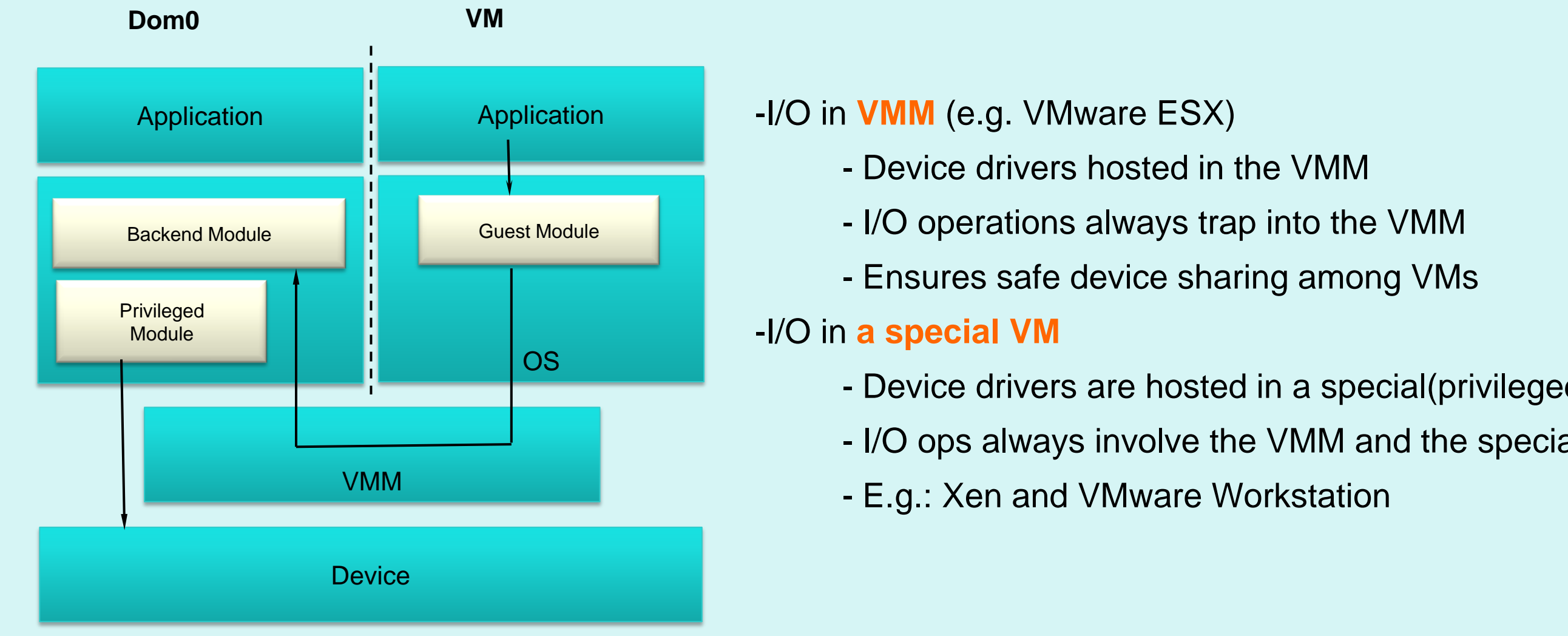
Motivation



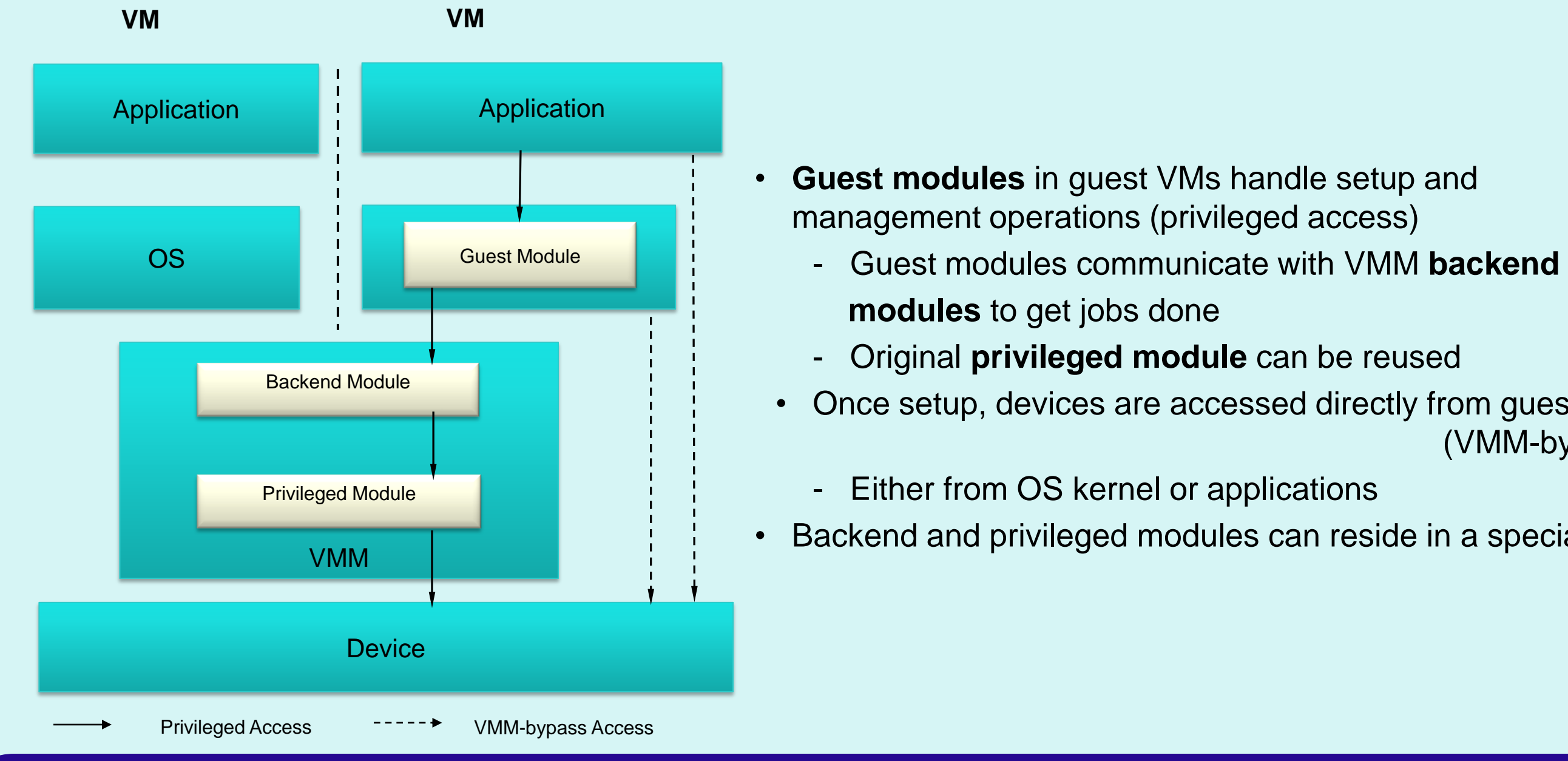
Network Faults



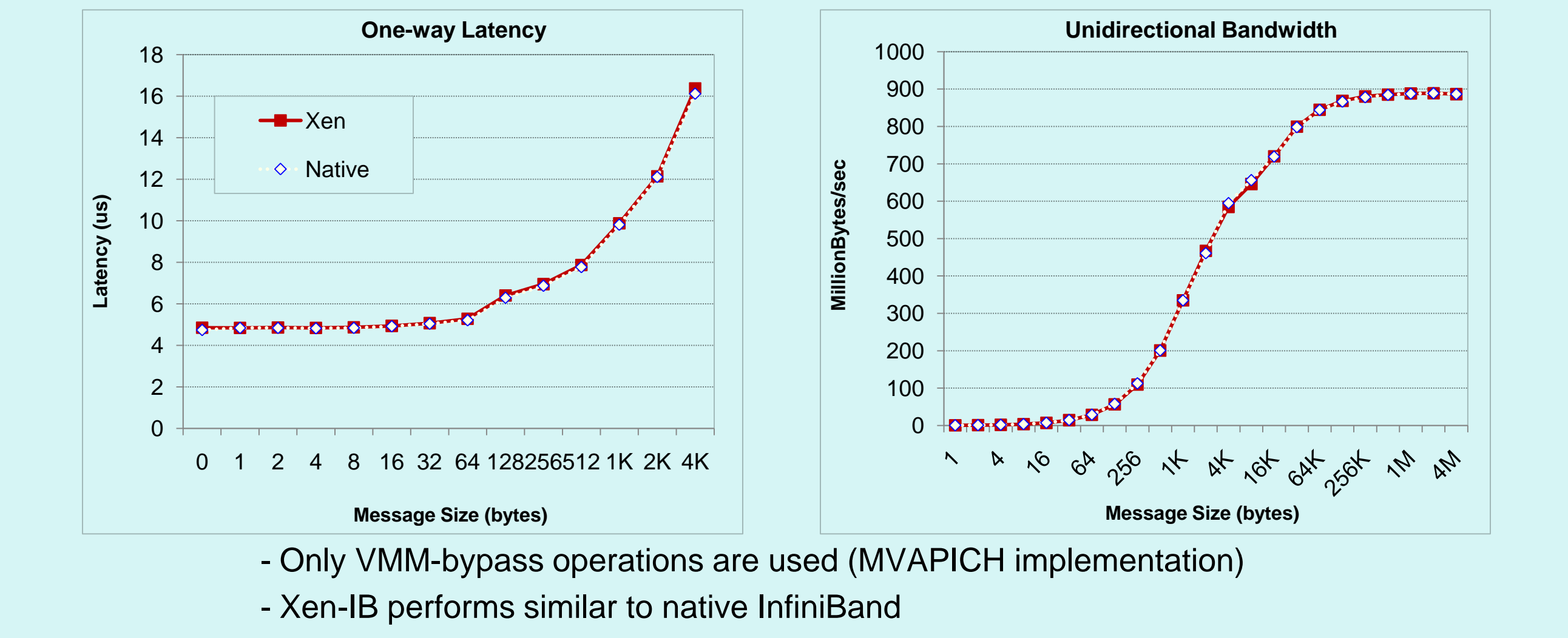
I/O Virtualization



OS-bypass to VMM-bypass



Near-native Performance with Xen

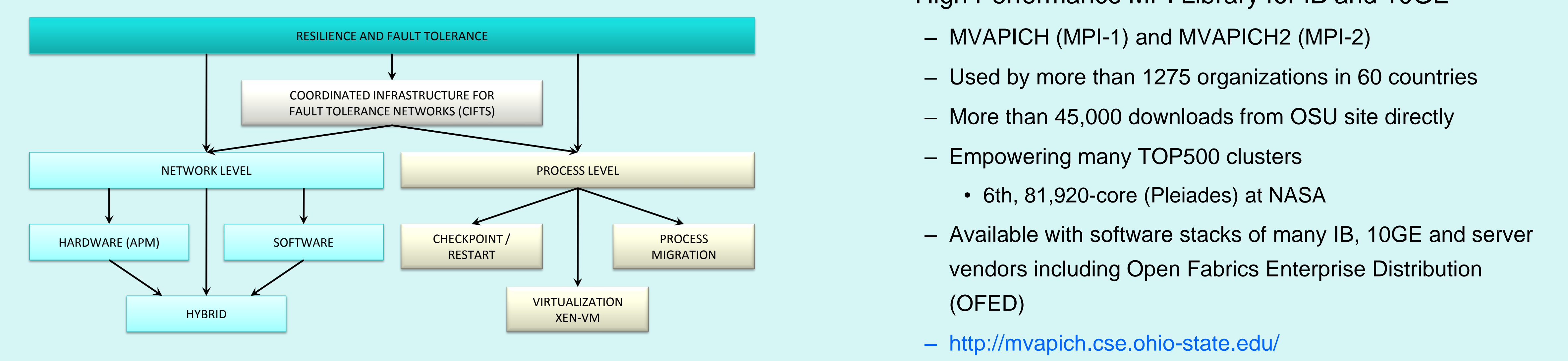


- J. Liu, W. Huang, B. Abali, D. K. Panda. "High Performance VMM-Bypass I/O in Virtual Machines", USENIX '06
- W. Huang, J. Liu, B. Abali, D. K. Panda. "A Case for High Performance Computing with Virtual Machines", ICS '06

Reduced Performance Overhead and Live Migration using Virtualization for HPC

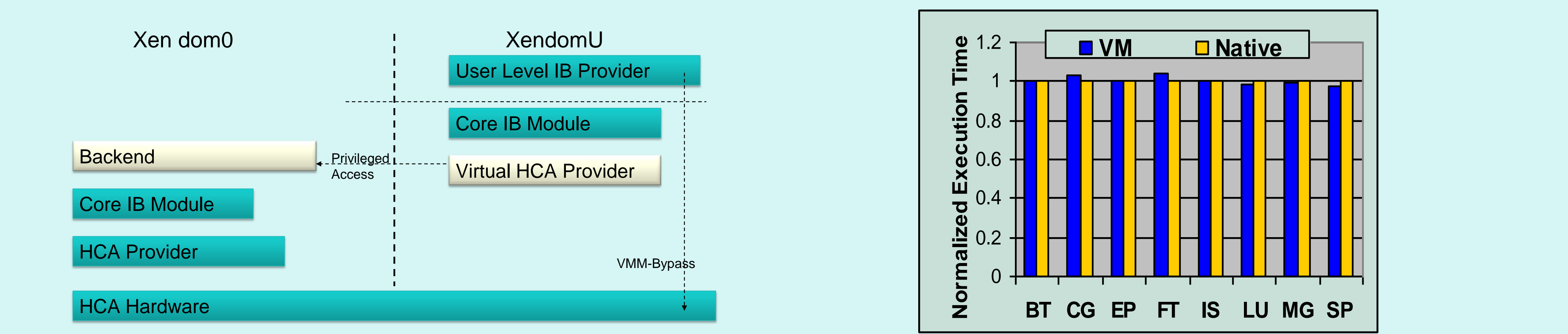
D. K. Panda
The Ohio State University

Overview

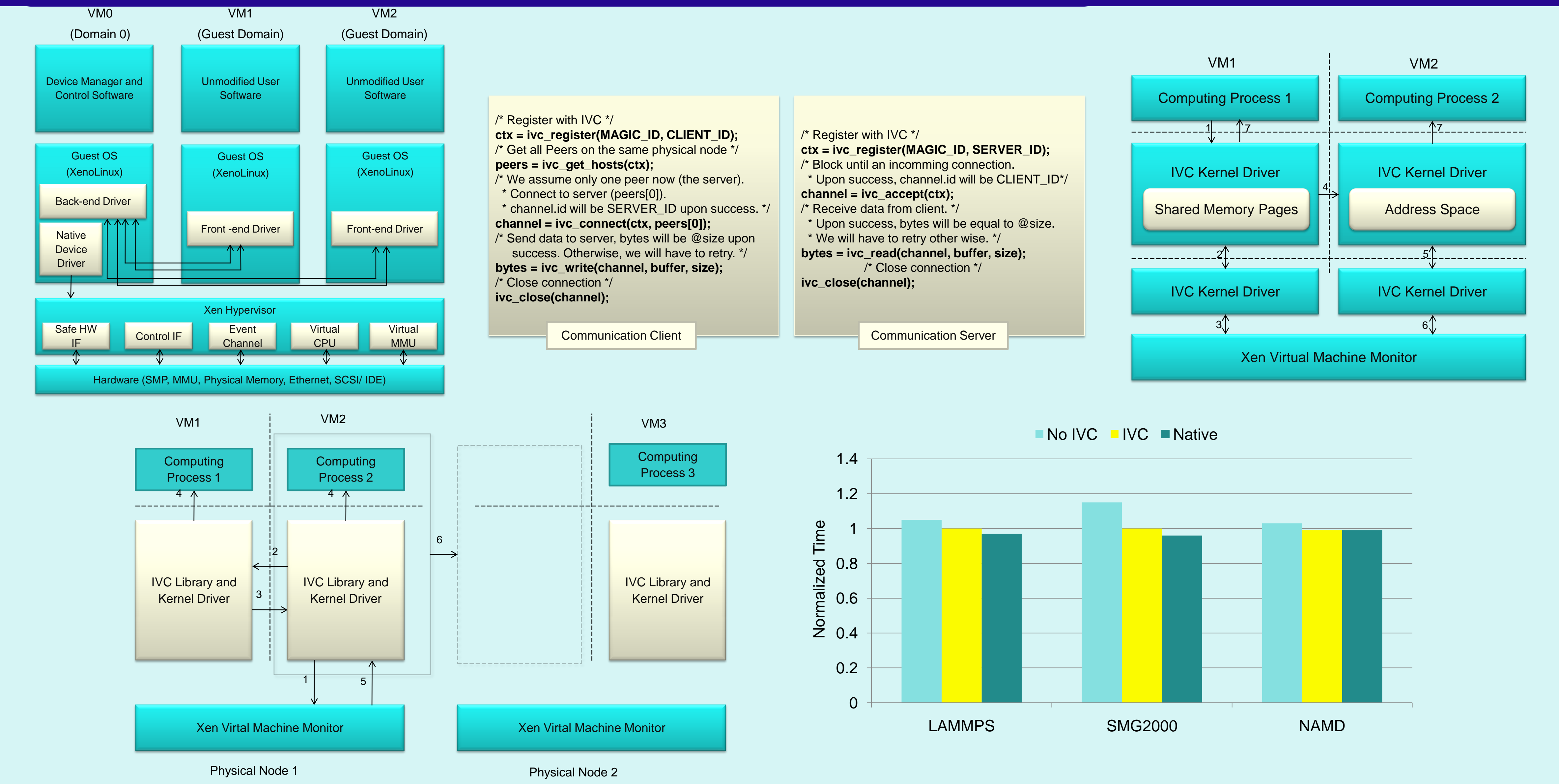


- High Performance MPI Library for IB and 10GE
 - MVAPICH (MPI-1) and MVAPICH2 (MPI-2)
 - Used by more than 1275 organizations in 60 countries
 - More than 45,000 downloads from OSU site directly
 - Empowering many TOP500 clusters
 - 6th, 81,920-core (Pleiades) at NASA
 - Available with software stacks of many IB, 10GE and server vendors including Open Fabrics Enterprise Distribution (OFED)
 - <http://mvapich.cse.ohio-state.edu/>

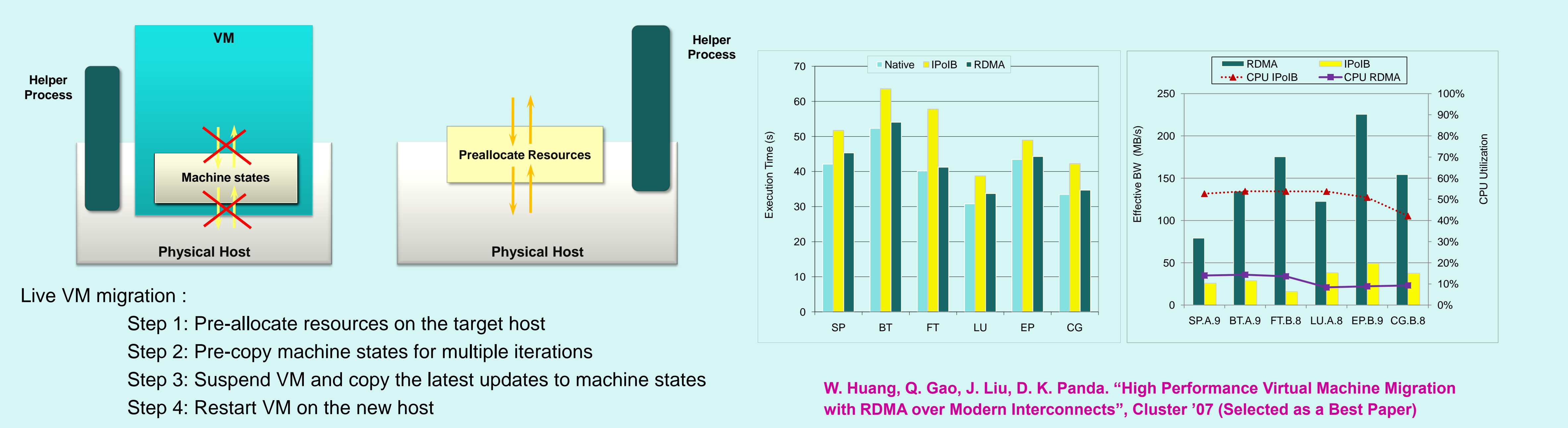
Virtualization-Xen-IB



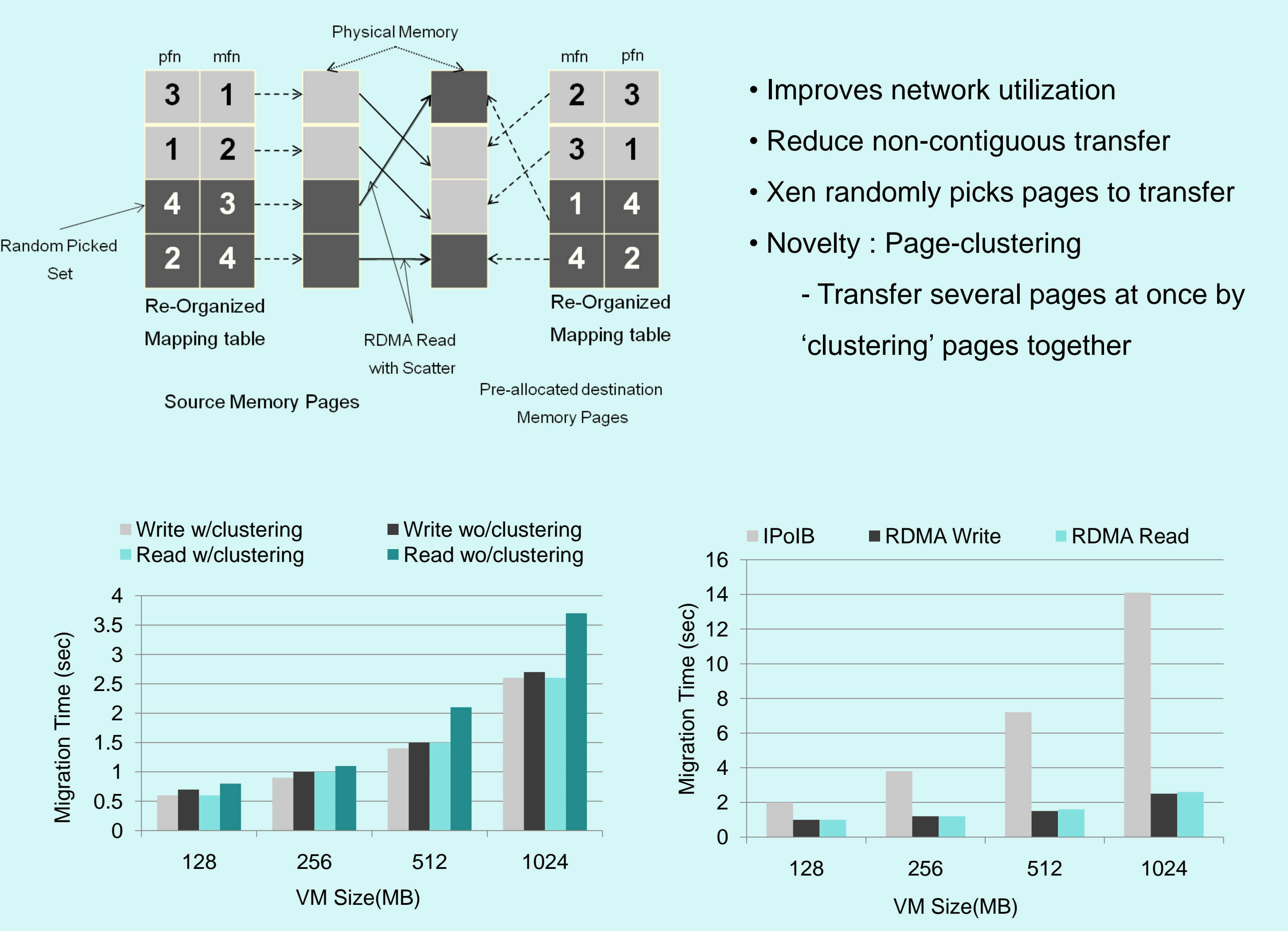
VM aware Communication Libraries for HPC



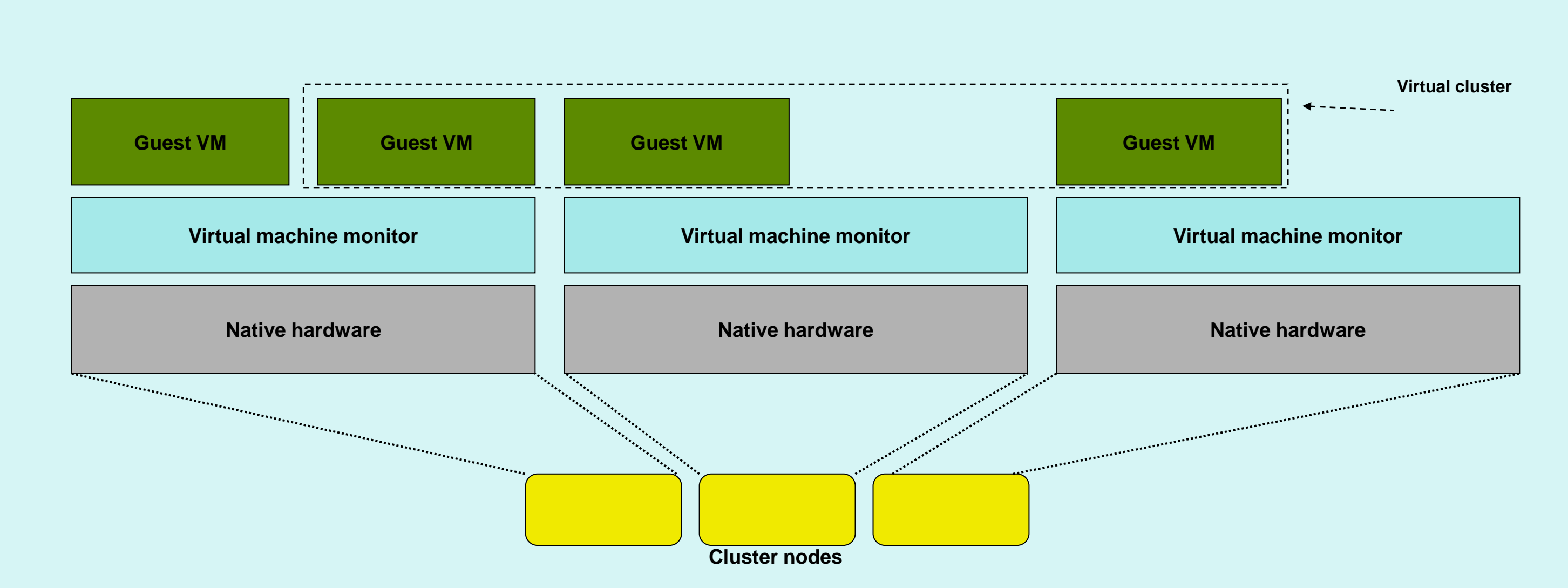
VM Migration through RDMA



VM Migration with Page-Clustering



Issues with In-band Management on VM



- Root access on a single VM would allow a user to control the subnet manager and subnet agents on all IB components
 - Virtualization does not help, since each "virtual network adapter" is connected to the entire network fabric
- Subnet agents do not require any authentication from the subnet manager
- Other networks such as Ethernet only support out-of-band management
 - More secure since who gets access to the management network can be restricted
 - User-defined authentication possible on out-of-band interfaces

Publications

- W. Huang, M. Koop and D.K. Panda, Efficient One-Copy MPI Shared Memory Communication in Virtual Machines, IEEE Cluster 2008, September 2008
- W. Huang, M. Koop, Q. Gao and D. K. Panda, Virtual Machine Aware Communication Libraries for High Performance Computing, Supercomputing (SC), 2007
- W. Huang, Q. Gao, J. Liu and D. K. Panda, High Performance Virtual Machine Migration with RDMA over Modern Interconnects. IEEE Conference on Cluster Computing (Cluster'07), September 2007 (**Best Paper Award**)
- W. Huang, J. Liu, M. Koop B. Abali and D. K. Panda, Nomad: Migrating OS-Bypass Networks in Virtual Machines, The 3rd ACM/USENIX Conference on Virtual Execution Environments (VEE), 2007
- J. Liu, W. Huang, B. Abali, D. K. Panda, High Performance VMM-Bypass I/O in Virtual Machines, USENIX Annual Technical Conference (USENIX'06), May, 2006
- W. Huang, J. Liu, B. Abali, D. K. Panda, A Case for High Performance Computing with Virtual Machines, ACM International Conference on Supercomputing (ICS '06), June, 2006

Acknowledgements & Collaborations



The IB support package for XEN Hypervisor can be downloaded from:
http://www.mellanox.com/content/pages.php?pg=xen_ib_drivers
<http://www.openfabrics.org/>