OpenFabrics Windows Status

Fab Tillier
SilverStorm Technologies
ftillier@silverstorm.com

Agenda

- Overview
- Current Attractions
- Coming Soon
- Long Term

Goals

- Industry standard IB SW stack
  - Increases adoption rate
  - Decreases development costs
- Broad range of upper level protocols
- Tight integration into the Windows OS
- Windows Compute Cluster Server 2003
- Windows Server 2003
  - x86, x64
- Windows XP
  - x86, x64

Components & Maintainers

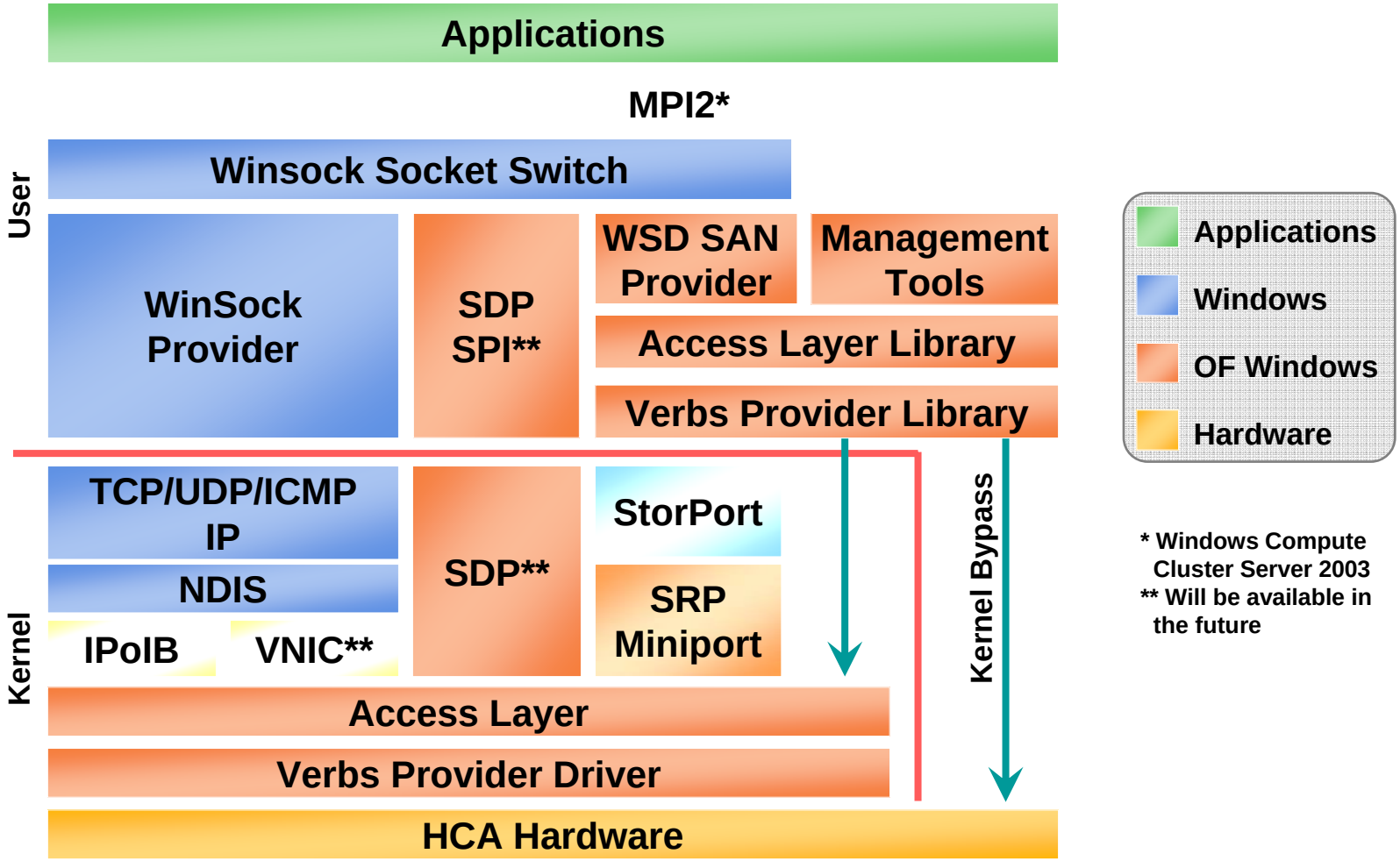
Component	Maintainer
Bus Drivers/Access Layer	Fab Tillier (SilverStorm)
HCA driver	Leonid Keller (Mellanox)
IPoIB	Fab Tillier (SilverStorm)
Winsock Direct Provider*	Fab Tillier (SilverStorm)
SRP Initiator**	Fab Tillier (SilverStorm)
SDP***	Tzachi Dar (Mellanox)
OpenSM	Eitan Zahavi (Mellanox)
uDAPL	Fab Tillier (SilverStorm)
VNIC***	Alex Estrin (SilverStorm)

* Not available on Windows XP

** Not available on Windows XP 32 bits

*** Will be available in the future

Architecture



* Windows Compute Cluster Server 2003
** Will be available in the future

Application Development

- WSD preferable for TCP applications
  - Shields users from verb interface changes
  - Not for Windows to Linux communication
- SDP for Windows to Linux
  - Not yet available through OFW
- IPoIB for UDP, Multicast
- IBAL interface for highest performance
  - Dynamic interface
  - No mechanisms for backward compatibility yet
  - Not always compatible between releases

Agenda

- Overview
- Current Attractions
- Coming Soon
- Long Term

Latest Build

- Available from OpenFabrics website:

  http://windows.openib.org/downloads/binaries

  - 1.0 release stream
  - SVN revision number included in release number
  - Latest as of June 16, 2006: 1.0.0.384

1.0.0.384 New Features

MTHCA Driver

- Replaces TVPD driver
- Supports all Mellanox HCA devices
  - Mem and memfree, SDR, DDR, etc.
- Windows Software Tracing (WMI)
- Tracking Linux development
  - Lower latency, higher bandwidth
  - Fewer LOC
  - Leverage Linux development

MTHCA Device Support

Device	Dev ID	FW	TVPD	MTHCA
InfiniHost (Tavor) 	23108	3.4.0	☑	☑
InfiniHost III-Ex (Arbel) InfiniHost Mode 	25208	4.7.600	☑	☑
InfiniHost III-Ex (Arbel) InfiniHost III Mode 	25218	5.1.400	☒	☑
InfiniHost III-Lx (Sinai) 	25204	1.0.800	☒	☑

1.0.0.384 New Features

IBAL

- Windows Software Tracing (WMI)
- Mellanox Fast Memory Registration support
  - Vendor specific extension

1.0.0.384 New Features

IPoIB

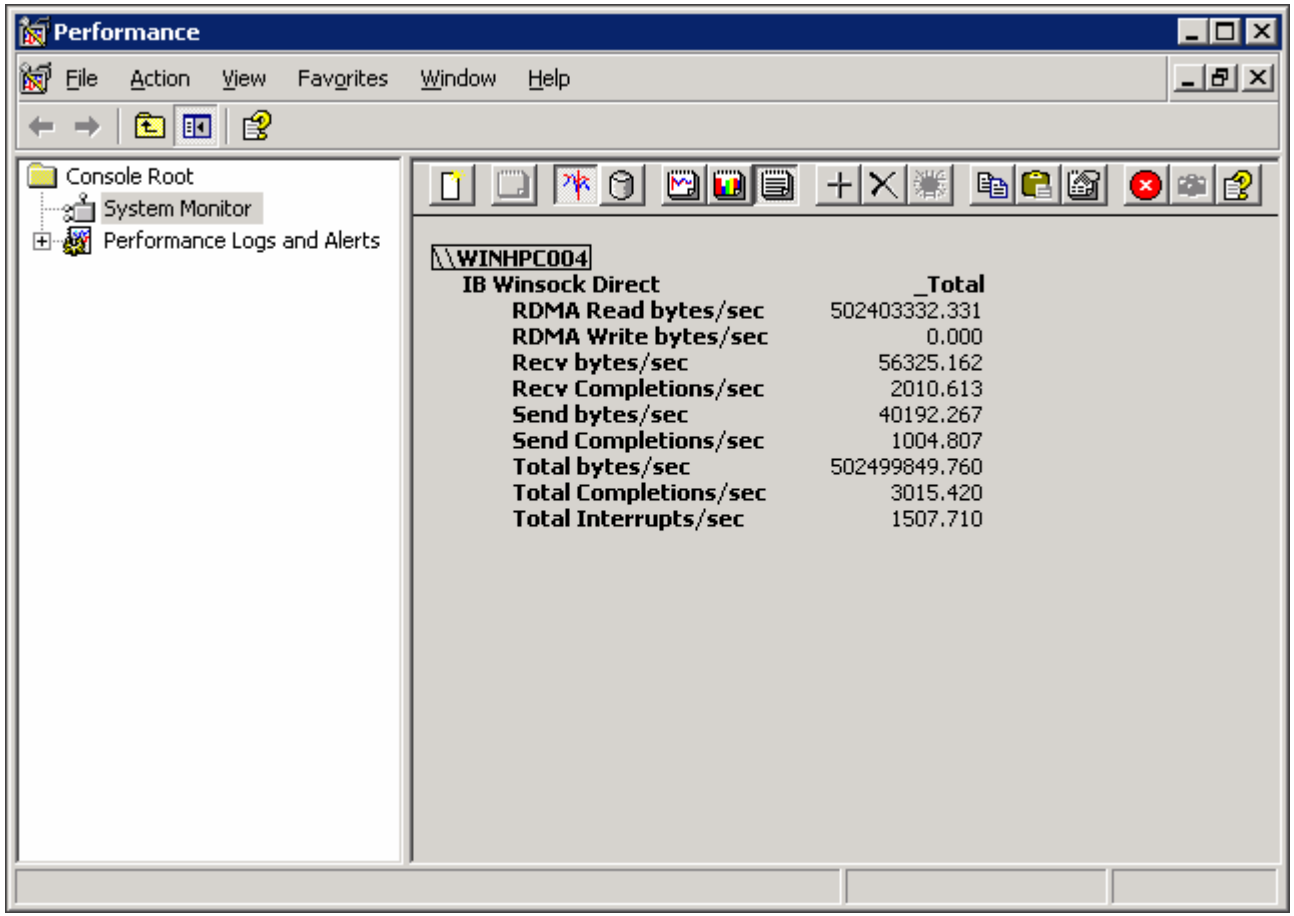
- Event logging to system log
- Windows Software Tracing (WMI)
- Increased robustness and stability
  - Receive queue starvation
  - SM reregistration handling
- WHQL testing

1.0.0.384 New Features

WSD

- Support MTHCA driver
- Performance Monitoring in Perfmon
- WHQL Testing

WSD Performance Counters



WHQL Overview

- Microsoft defined WHQL Program for IB
  - Requires both IPoIB and WSD
  - OpenFabrics Windows first candidate
- Working with Microsoft to refine process
  - Exemptions needed for IPoIB
  - WSD testing methodology
- Mellanox serving as certifying company
- To be available from OFW downloads

WHQL Status

- Currently top priority in OFW
- 1.0.0.384 is rc0 for WHQL
- IPoIB running all tests to completion
  - Negotiating contingencies and errata
- WSD failing 5-10% of tests
  - Socket duplication
  - Possible test case issues
  - Debugging with Microsoft

Agenda

- Overview
- Current Attractions
- Coming Soon
- Long Term

Tools

- Automated Installation/Upgrade
  - Manual installation not practical for clusters
- User-friendly FW update
- Diagnostics

IPoIB

- Linux Interoperability
  - DHCP packet format and handling
  - Windows DHCP server requires unique MAC conversion for port GUIDs
  - Use IETF-compliant MGIDs for multicast
    - Currently works between Windows systems only

SRP

- Currently Beta, tested with:
  - SilverStorm VFx
  - LSI Logic IB Storage
  - Mellanox IB Storage reference design
- Preliminary error handling
  - IB and FC cable outages
- Stabilization effort to begin shortly
- August release targeted

VNIC

- Implements SilverStorm EVIC Protocol
  - SilverStorm Ethernet Virtual I/O Modules
- New development
- September release targeted

SDP

- Working logistics to add source to SVN
- Plan to have binary version available sooner
- Technology preview
- Currently bCopy only
- zCopy roadmap TBD

Agenda

- Overview
- Current Attractions
- Coming Soon
- Long Term

General Direction

- Better integration into Windows
- Leverage key design elements from Linux OpenFabrics stack where it makes sense
  - CQ Polling WC array vs. linked list
  - Public structures for objects instead of opaque handles to store static attributes
- Partition Support

Kernel Future

- Power Management Support
- Boot Support
- Verbs at DISPATCH_LEVEL
  - IRP-based
  - Client controls sync/async policy
  - IOCTL and Direct Call
- HCA device object resolution
  - Reparse from well-known name to actual HCA device object
  - All verb calls made on actual HCA device object

User-Mode Future

- Win32 Look and Feel
  - Reduce learning curve for users
- Support true single-threaded apps
- Client controls async/sync event policy
  - Expose Win32 I/O completion mechanisms and semantics

Call To Action

- Use OpenFabrics Windows
  - End users
  - Application developers
- Communicate
  - Questions, issues, feature requests

  http://openib.org/mailman/listinfo/openib-windows
- Contribute
  - Send patches and fixes
  - Add driver capabilities
  - Add new upper layer protocols

  http://windows.openib.org/openib/contribute.aspx

Resources

- OpenIB WiKi
  - https://openib.org/tiki/tiki-index.php?page=OpenIB+Windows
- Openib-windows mailing list
  - http://openib.org/mailman/listinfo/openib-windows
- Binaries
  - http://windows.openib.org/downloads/binaries
- Symbol Server
  - http://windows.openib.org/downloads/symbols

Q & A