Technology Discussion: Lightning Fast I/O with PacketDirect

Gabriel Silva
Program Manager
Microsoft – Datacenter Network Platform Team

Microsoft

Agenda

- Networking State of the Union
- What is PacketDirect (PD)?
- Inner workings and architecture of PD
- Target scenarios
- Future possibilities
- Demo

State of the Union

Inflection point

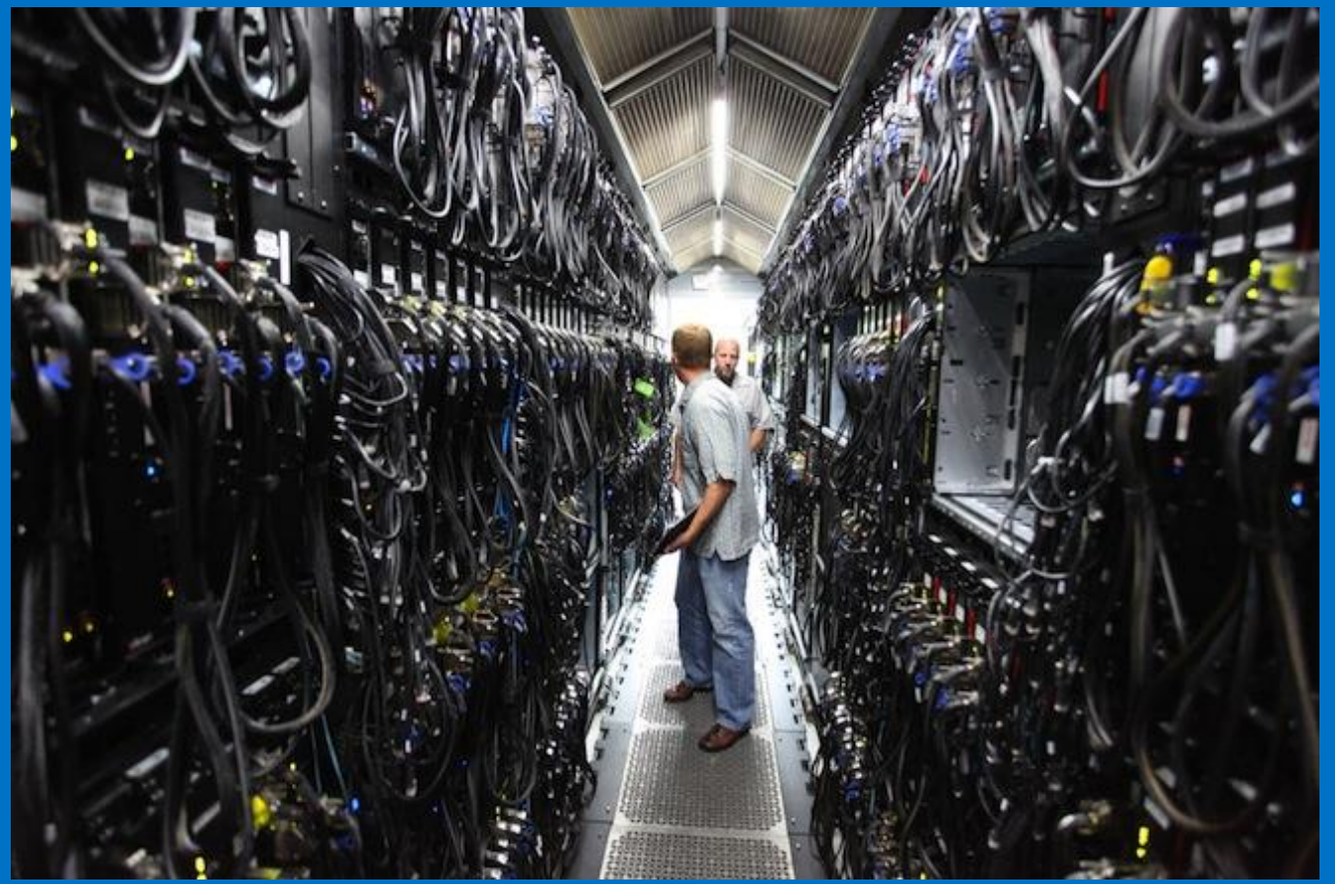
- Disaggregation of the networking stack
  - Multiple offload technologies – allow higher bandwidths, better CPU utilization and greater workload VM densities
- Disaggregation of hardware
  - Networking moving from custom hardware to Virtualized Network Functions
- Silicon grows up
  - Merchant silicon providing the foundations for Software Defined Networks, with an open control interface

Networking is evolving

- Software Defined Networking (SDN) provides the under pinning for the Software Defined Datacenter
  - Network Controller – self-optimizing and self-healing networks
- Network Function Virtualization
  - Rapid scaling out/in of network functions
- Innovation in performance in scale on merchant silicon
  - 40Gbps appliances running inside VMs

Cloud scale requirements

- Networking for the datacenter
  - 10G to 40G to 100G
  - Very low latency, 1-2ms
  - 1 core, 12 million pps



Networking for Windows

- NDIS for Windows
  - General purpose platform
  - Support for client and datacenter alike
- Is NDIS in its current form enough?
  - No
- We can do better
  - General purpose IO
  - Memory
  - Application is not in full control of its packet management
- DPDK Technology for Intel NICs
- PacketDirect

What is PD?

What is PD?

- New IO model, extension of NDIS
- Does not replace the current I/O model
- Physical and virtual scenarios

Focus of PD

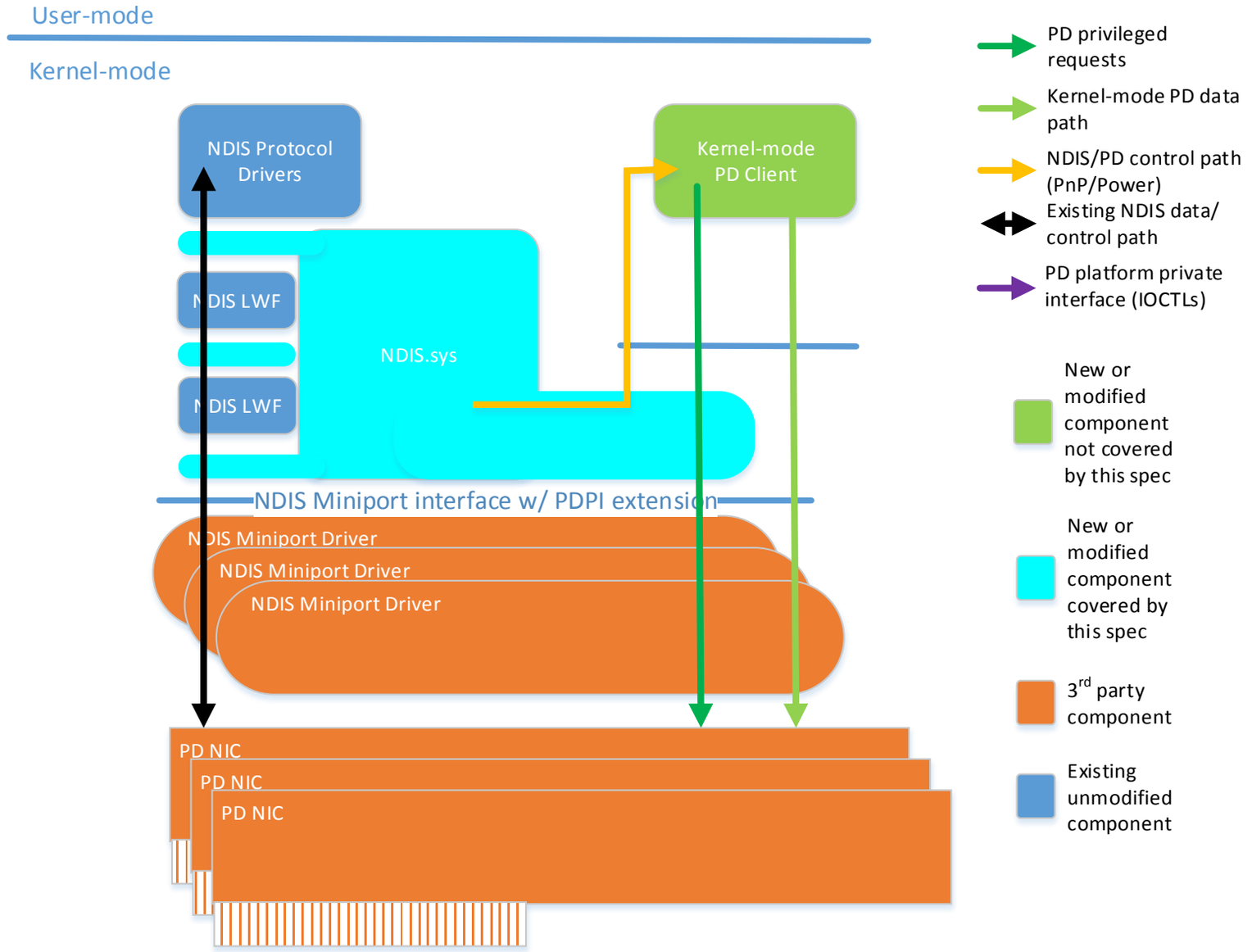
- Physical is interesting…
- Virtualization is our focus
  - Take Network Appliances inside a VM
    - DDoS
    - SLB
    - Gateway
    - Firewall

Inner workings & Architecture

Main Components of PD

- PD Client
- PD Platform
- PD Provider Interface (PDPI)

PD Architecture



PD Concepts

- PD execution context (EC)
  - Thread for PD client processing
- PD buffer
  - New packet descriptor format
- PD provider filters
  - Determine where to process which packets/flows, i.e., in control of how to partition the workload as far as CPU/memory/cache goes
- Post-and-Drain
  - Brings all the above concepts together

PD Provider

- Implemented by NIC based on PD Provider Interface (Public)
  - Worked closely w Intel for feedback on APIs
  - Intel Niantic / Fortville among first to implement
- Direct HW queue access for PD clients
  - Rx Queues
  - Tx Queues

PD Platform

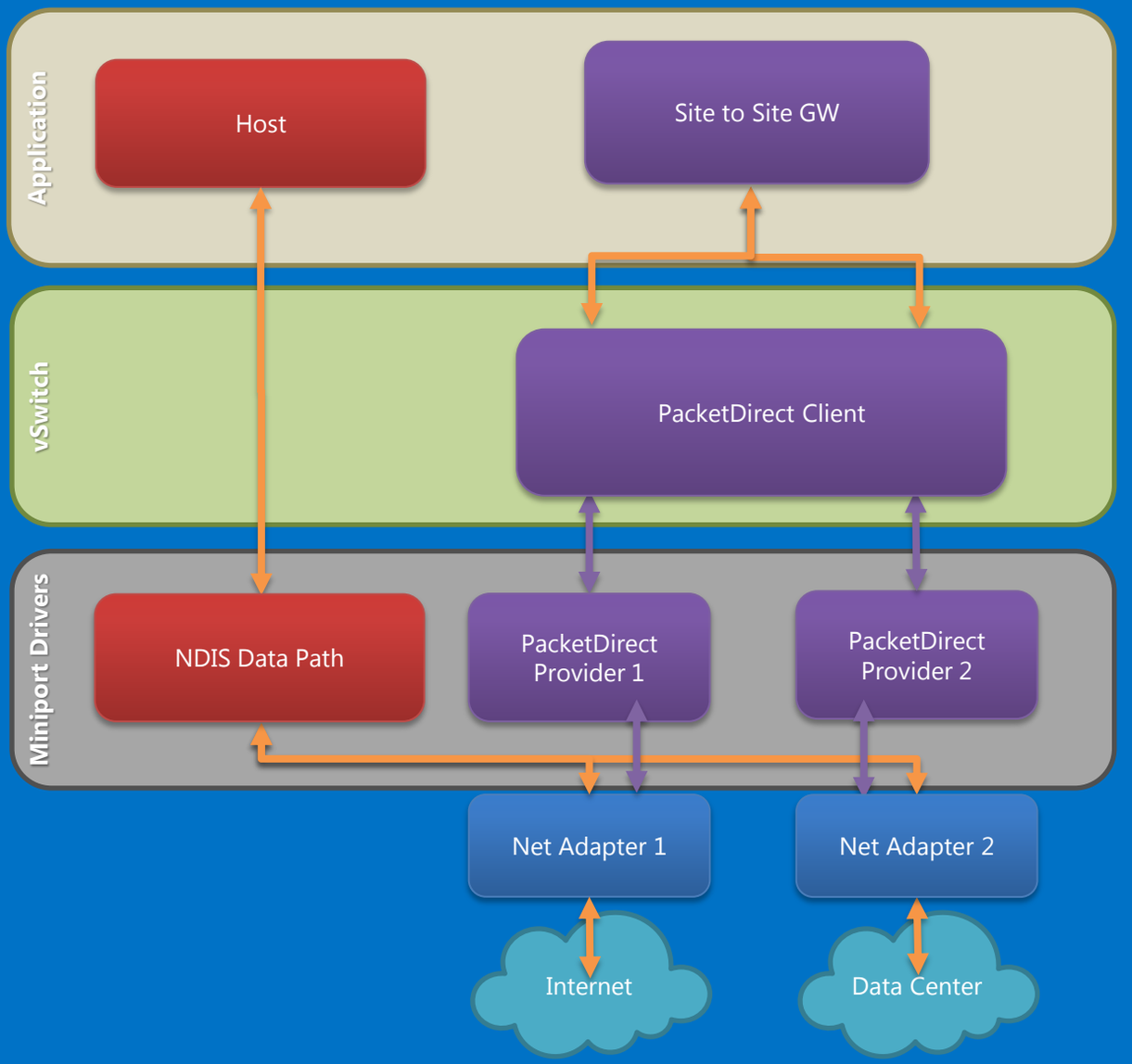
- Buffer management
- Queue abstractions
- Manageability / diagnostics

PD Client

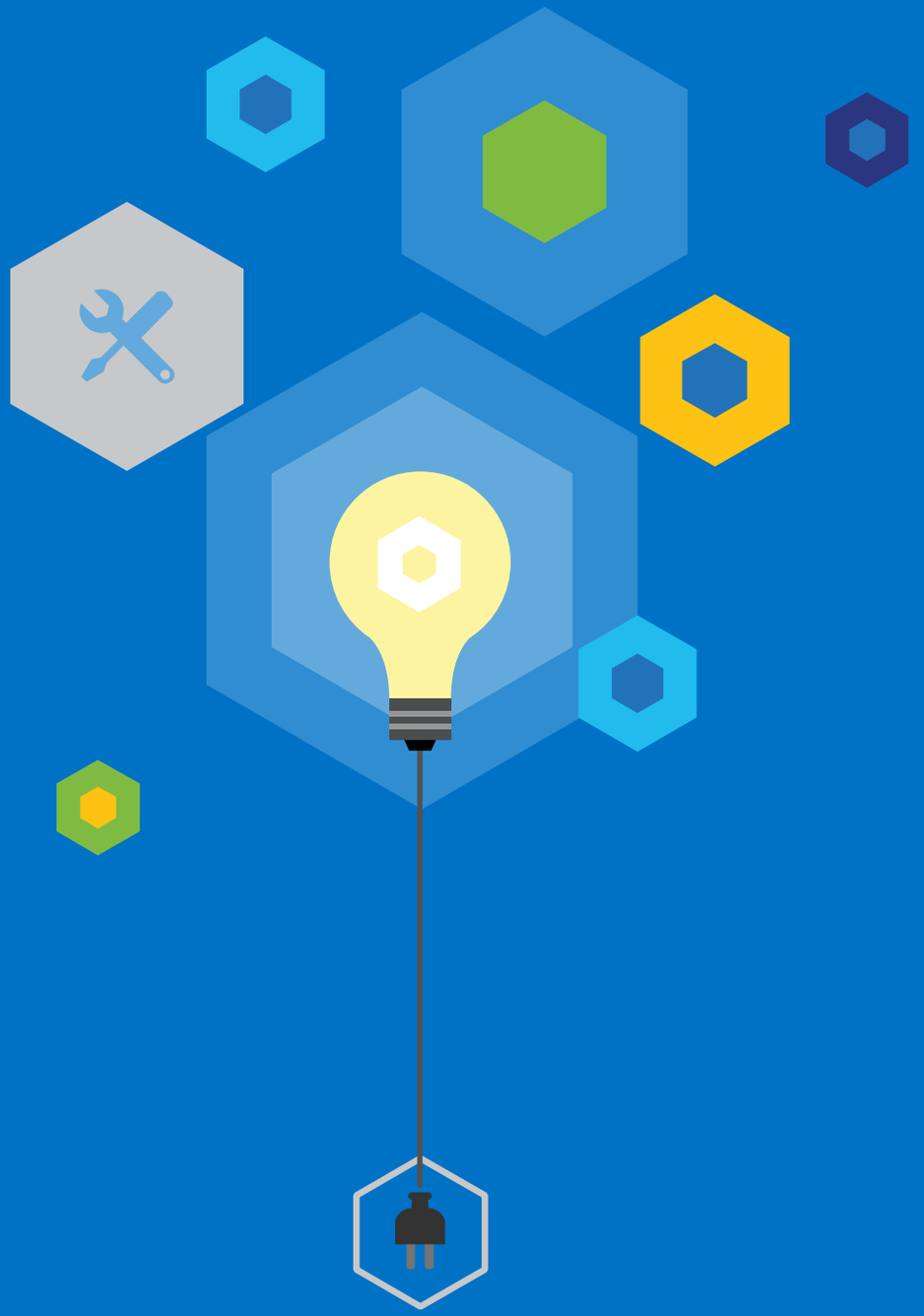
- Network Service / Application
- Take advantage of the new IO model to speed up the data path
- Packet processing context owned and managed by the client
  - Own buffers
  - Create / Maintain ECs
  - Create filters on queues

Scenario

- vSwitch enlightened w PD
- Accelerate path into and out of VM



PD Demo



Futures

Future Possibilities

- PD publically available
- SR-IOV
- PD inside the VM
- User mode PD

Thank you

- Questions?

Microsoft