Open vSwitch

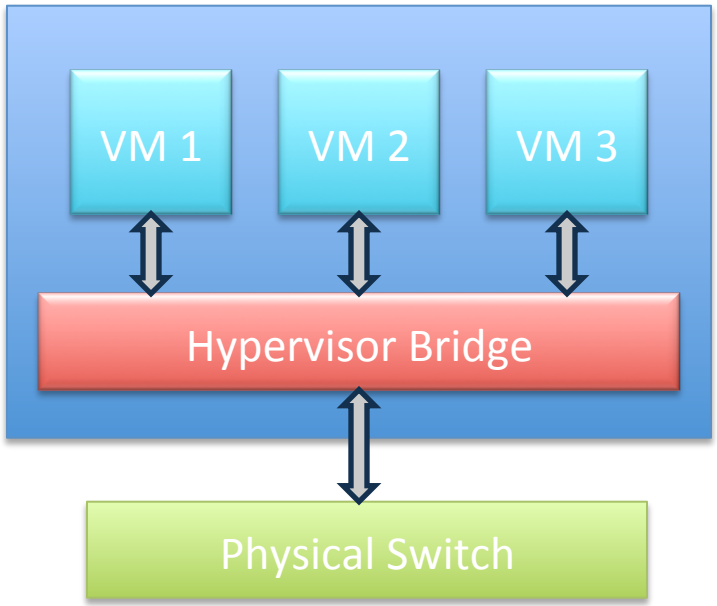
Justin Pettit and Jesse Gross

Linux Collaboration Summit

April 7, 2011

What is Virtualization?

- Multiple virtual machines on the same physical host
- Lowest layer is the hypervisor, which provides the illusion
- Built by OS people
- Historically, simple bridge



What's been missing?

- Lacked visibility that network administrators expected
- Lack of fine-grain control
- Brittle in face of mobility

Virtualized Networking is Different

- Greater context
  - MAC/IP addresses
  - Host identifiers (UUID)
  - Multicast membership
  - Machine start/stop/move events
- Mobility
- Strong isolation between tenants

Advanced Edge Switches

- Tight integration with hypervisor
- Approaching feature-parity with hardware switches
  - Visibility
  - ACLs
  - Quality of Service
- Centralized management
- Hardware off-loading
- Examples: VMware vSwitch, Cisco Nexus 1000V, Open vSwitch

Open vSwitch

- Advanced edge switch
- Works on Linux-based hypervisors: Xen, KVM, VirtualBox
- Open source, commercial-friendly license
- Widely deployed in enterprise, service provider, and Telco production environments

(Partial) List of Contributors

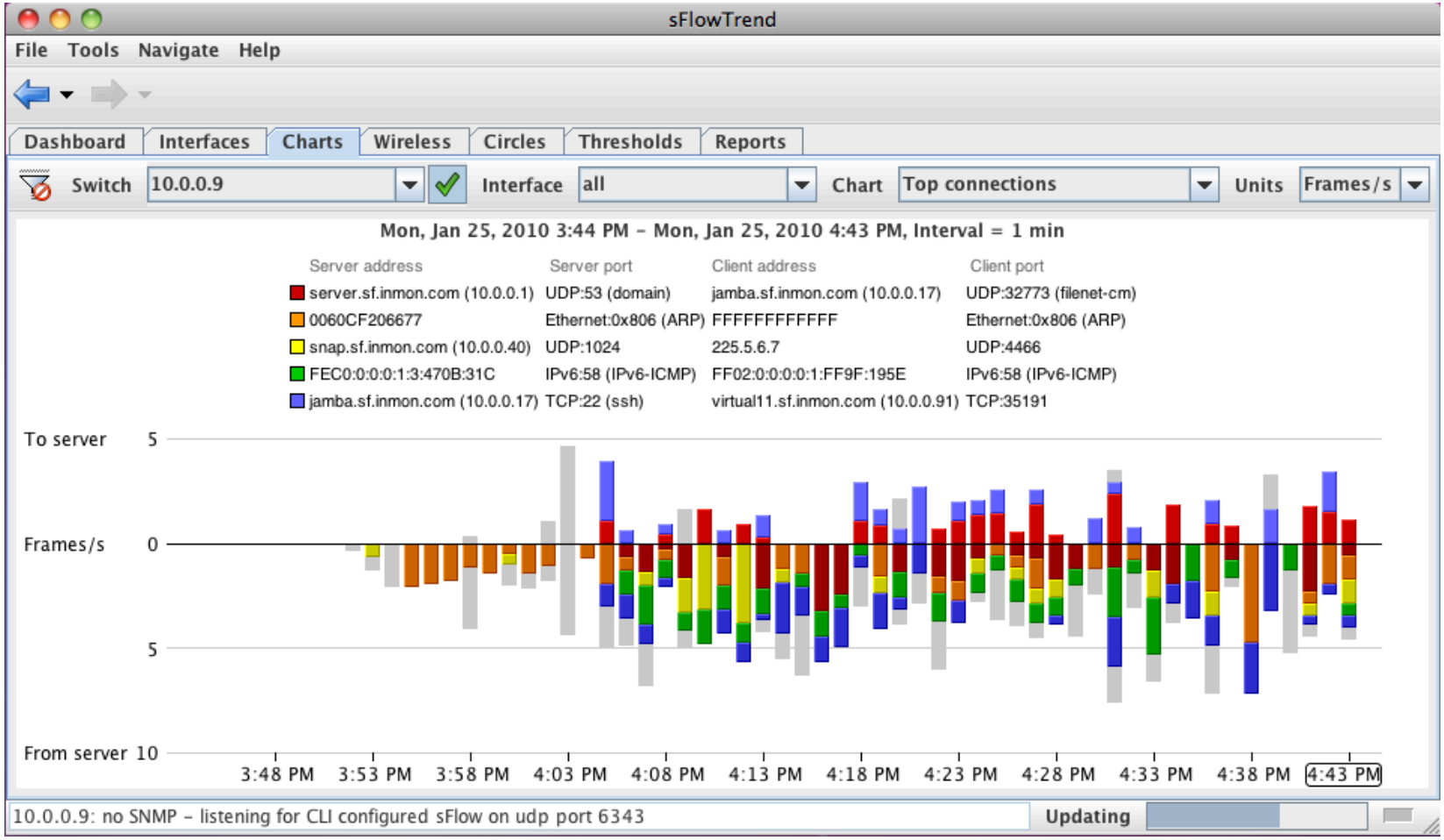
Packaging

- Default networking stack for Xen Cloud Platform (XCP)
- Ships with Citrix XenServer and basis for their Distributed Virtual Switch (DVS)
- Distribution packaging
  - Debian
  - Ubuntu
  - SUSE
  - Red Hat
- Goal is to upstream kernel module

Visibility and Control

- Visibility
  - NetFlow
  - sFlow
  - Mirroring (SPAN/RSPAN/ERSPAN)
- Fine-grained ACL and QoS policies
  - L2-L4 matching
  - Actions to forward, drop, modify, and queue
  - HTB and HFSC queuing disciplines

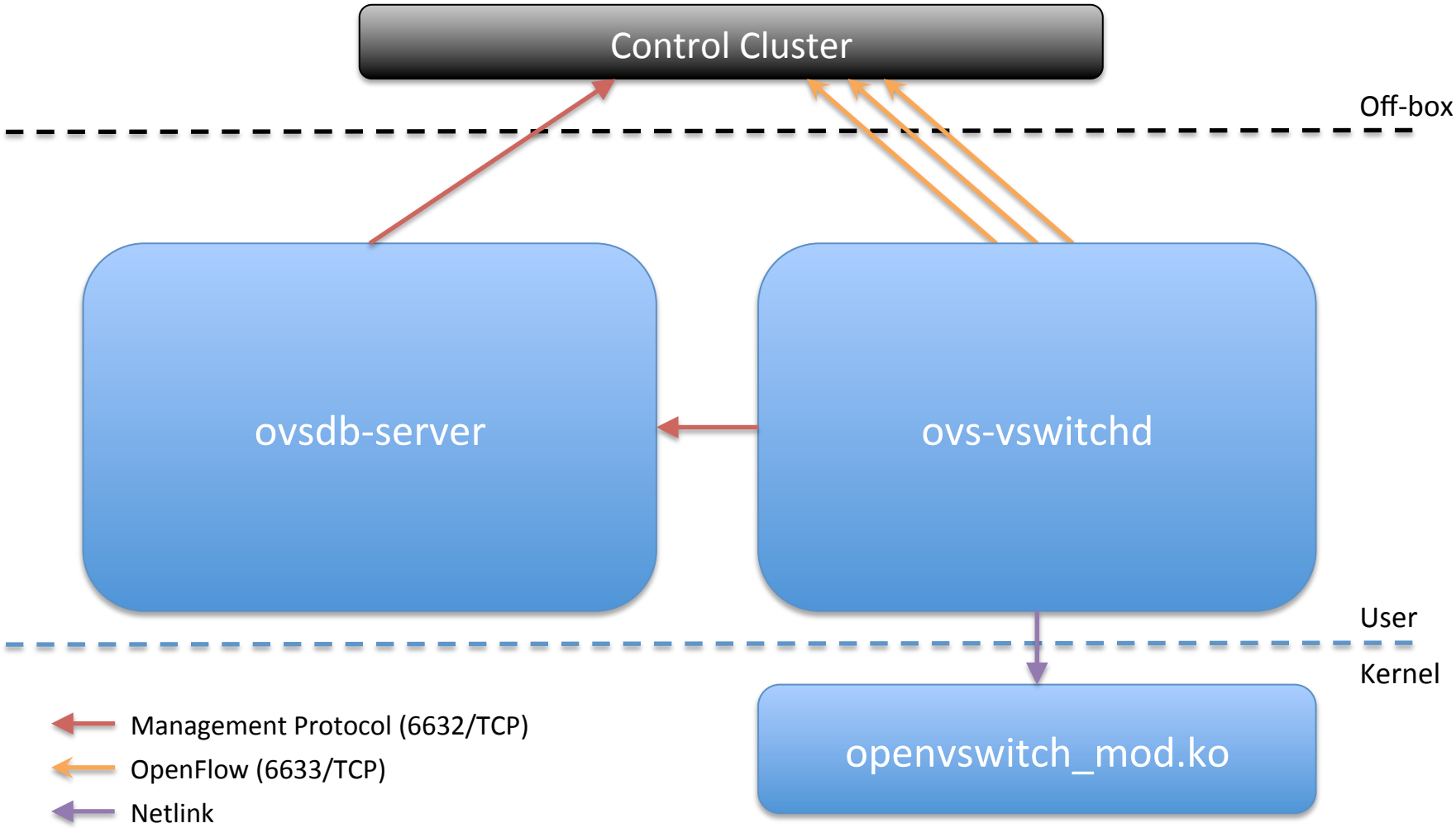
sFlow with Open vSwitch



Forwarding

- LACP
- Port bonding
  - Source-MAC load-balancing
  - TCP load-balancing
  - Active/backup
- 802.1ag CFM (Connectivity Fault Mgmt)
- Fast Ethernet-over-GRE tunneling

Main Components



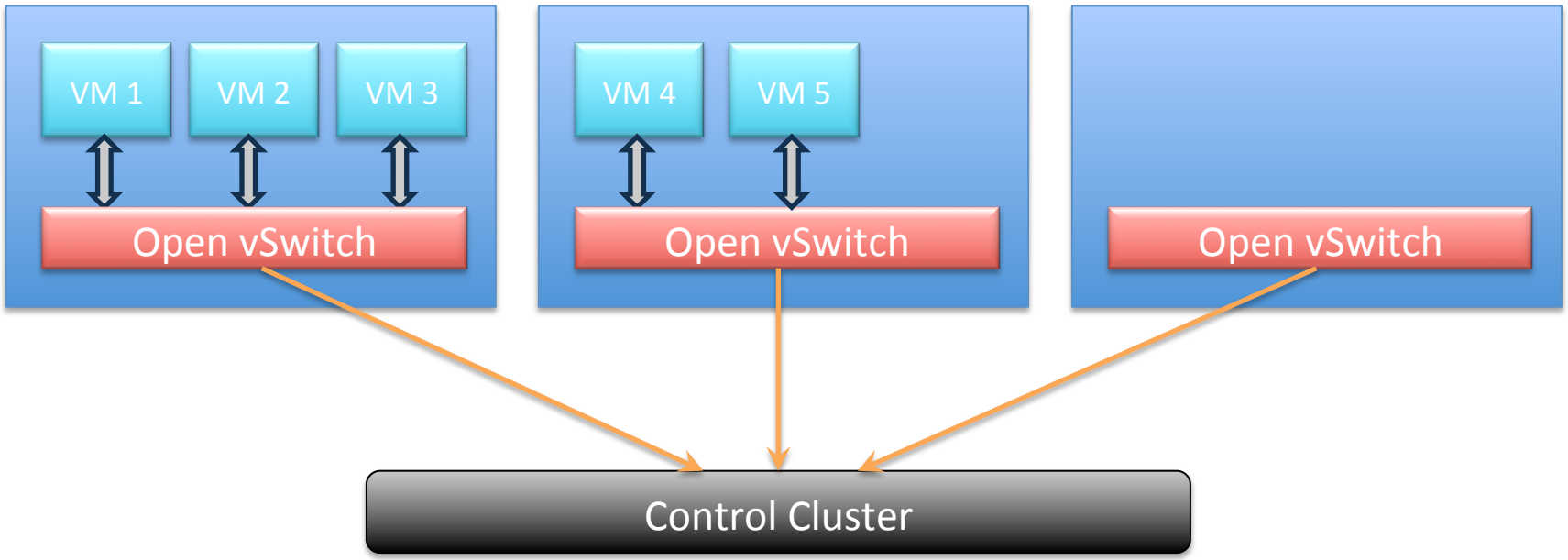
Forwarding Components

- ovs-vswitchd (Slow Path)
  - Forwarding logic (learning, mirroring, VLANs, and bonding)
  - Remote configuration and visibility
- openvswitch_mod.ko (Fast Path)
  - Packet lookup, modification, and forwarding
  - Tunnel encapsulation/decapsulation

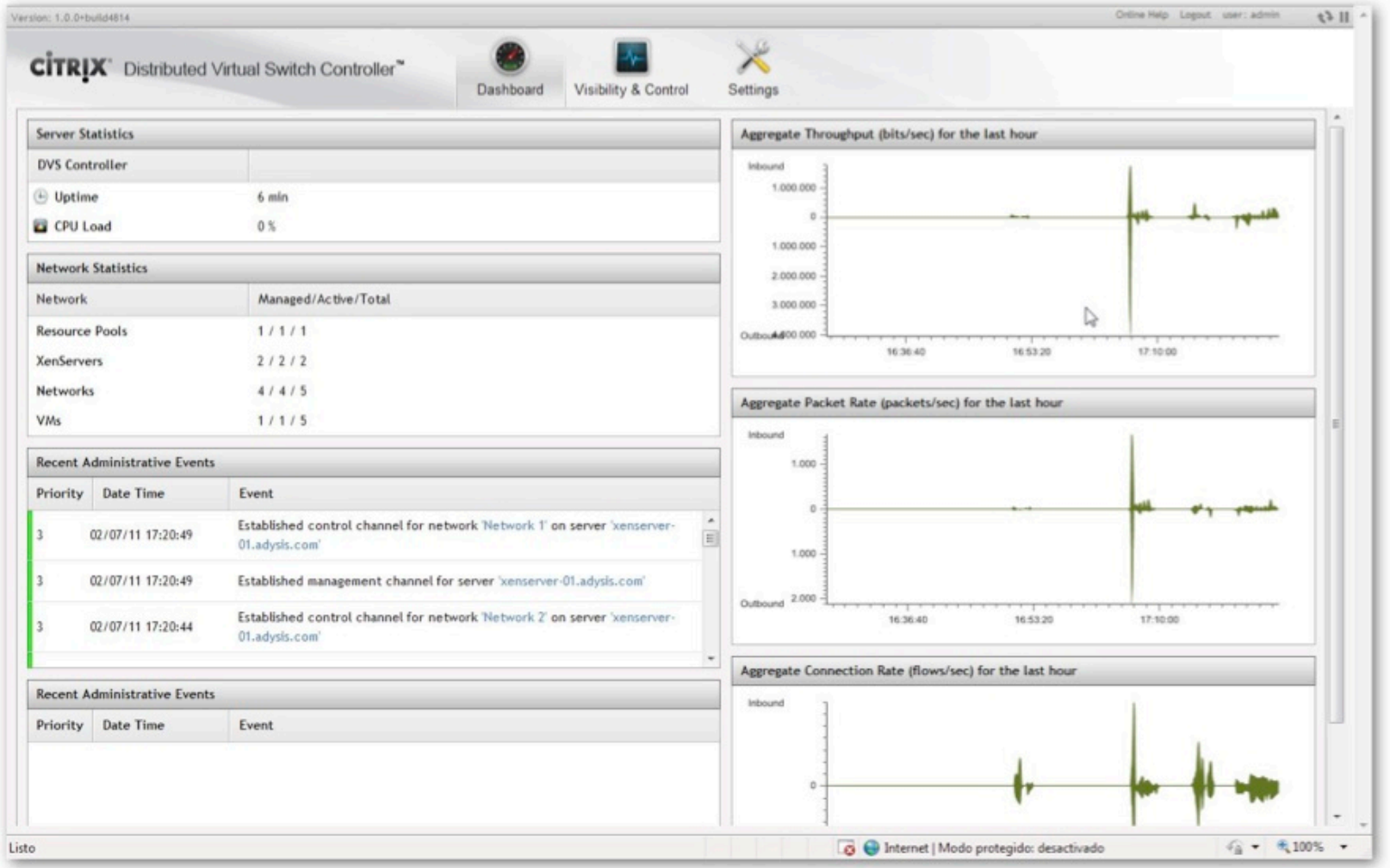
Centralized Control

- One OpenFlow connection per datapath
  - Exports idealized view of switch's datapath
    - Lookup based on L2-L4
    - Full wildcarding and priorities
    - Actions: forward, drop, modify, and queue
    - Missed flows go to central controller
- One management channel per system
  - Switch-level configuration
  - Resources
  - Counters

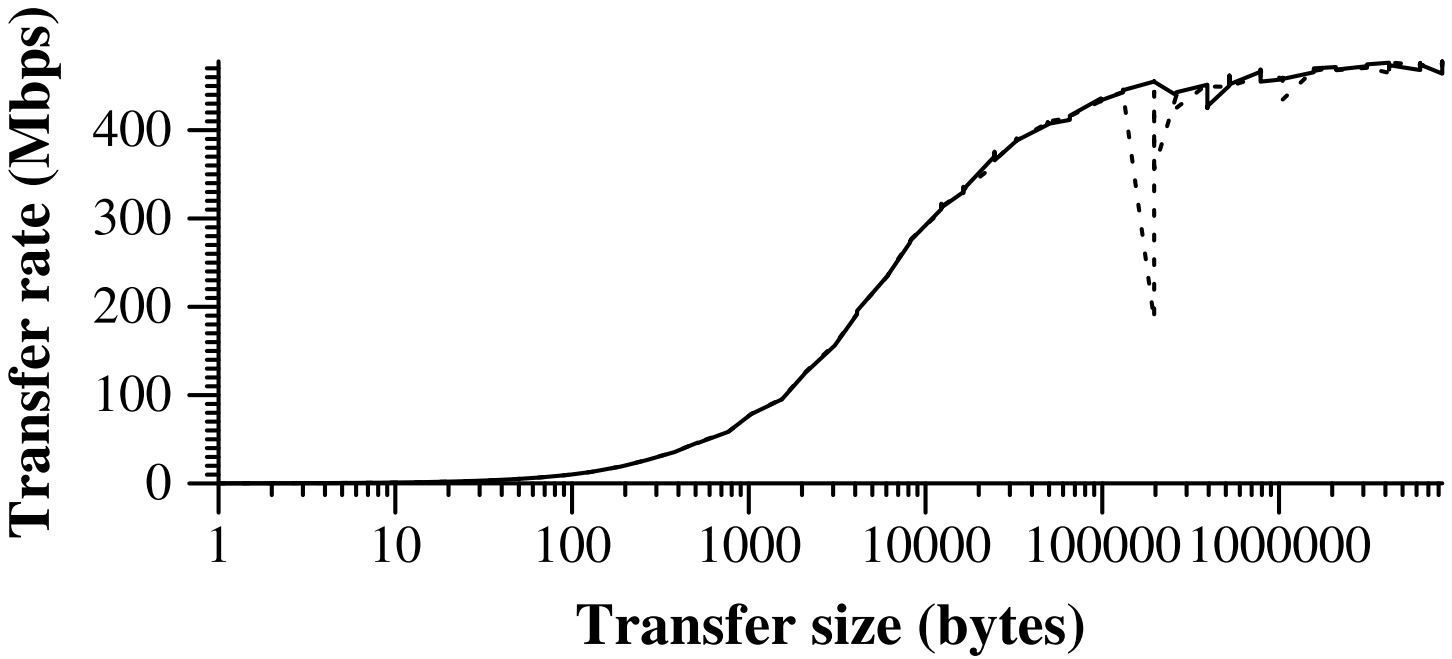
Distributed Virtual Switch



Citrix DVS Controller



Performance



Throughput versus flow size for Xen virtual machines with Linux bridge (dashed) and Open vSwitch (solid).

Project

- http://openvswitch.org
- Mailing Lists
  - Announcements: announce@openvswitch.org
  - User-level discussion: discuss@openvswitch.org
  - Dev (code review, etc): dev@openvswitch.org
  - Archives available in site sidebar
- Source repository:

  git clone git://openvswitch.org/openvswitch