



Networking in the Cloud

Crossbow's Revolutionary Impact in a Virtualized World

Ben Rockwood
Director of Systems Engineering
Joyent, Inc.

Problem #1: Bandwidth Accounting

- Accounting for bandwidth is more complicated than it should be
- PCAP based solutions are most common, but often elaborate (*pmacct*, etc)
- Log Trolling is also very common
- IPQoS is horrific; complicated and fragile

Problem #2: Bandwidth Enforcement

- Enforcing bandwidth is even more difficult
- Most hosting environments simply offer “unlimited bandwidth” and terminate abusive customers
- Would love to sell bandwidth plans to customers, ie: “unmetered 10Mb/s”
- Network choke points such as firewalls can do some of the work, but complex to manage and introduces a complex failure point
- Would love to manage bandwidth on the local box

Problem #3: Network Control

- Virtual interfaces (e1000g0:1) have limited capabilities passed on to the customer
- Customers want to snoop, run firewalls, etc.
- Customers themselves want to create virtual interfaces
- Special drivers are a poor solution (tun/tap, etc)
- Managing multiple default gateway networks without VLANs is problematic; a virtual interface routing nightmare
- Customers get bogus network statistics, aggregated for all virtual interfaces

Crossbow & Bandwidth Accounting

- IPQoS no longer required, w00t!
- Extended Accounting has its faults, but works well and can be integrated nicely into an accounting architecture
- Can get not just per interface granularity but per flow granularity too!
- *dladm* and *flowadm* “show-usage” subcommands make it easy for any administrator to query without special tools
- Easy to administer, any Solaris admin can learn it in 5 minutes.
- Allows business to focus on how to use the data, not how to generate it!

Crossbow & Bandwidth Enforcement

- Ability to enforce bandwidth both per interface & per flow!
- Provides a whole new way to sell resources and guarantee them to customers
- Easy to use, just a simple property; not special staff required
- Done locally, no special network firewalls or appliances
- One solution that works for any virtualization technology: xVM, Zones, VirtualBox, etc.

Crossbow & Network Control

- IP Instances make Zones significantly more compelling, but VNICs make them a viable solution.
- No more special IP Instance systems with quad-cards
- Customers get “real” interfaces that they can work with
- Network statistics are real because they aren’t on a shared NIC
- Additional network interfaces are easier to add than ever
- Etherstubs provide excellent way to create private internal networks

Crossbow & Research

- Testing new distributed technology is complicated
- Who actually has dedicated racks of machines for testing?
- Etherstub networks allow testing of distributed technologies within a single system
- Crossbow + Quagga = Network Admins Playbox! :)

The logo features a stylized orange cross symbol on the left, followed by the word "Joyent" in a large, white, sans-serif font.

Joyent

Crossbow can fix that....

Integration Example: *linkstat*

- *linkstat* is a monitoring tool which displays data-link statistics in a similar manner to *iostat -xn*
- **Why?** *netstat -i* and *dladm showlink -s* will show you the data you want, but only for a single interface at a time, you can not view all interfaces simultaneously to answer “who’s sucking up the bandwidth??” questions.
- You could use *dladm show-usage*, but real-time data is helpful in many troubleshooting instances.
- Implemented currently in PERL using ‘link’ class object.

Initial Concerns (From snv_105 Analysis Notes)

- Joyent Internal Documentation started: 27 Sep 2007
- Quoted from “Joyent Edge Cases and Concerns”:
 - * Link Accounting (Per-Day Bandwidth Usage)
 - * Link Constriction; Rate-Limiting
 - * Snoop (Test snooping other Zones on the same interface)
 - * Using an unassigned IP address (IPF)
 - * IP Filtering Inside Zone
 - * IP Filtering from Global Zone
 - * Need to verify kstats; global and internal

How to make Crossbow even better!

- For Zones architecture to optionally create a VNIC if not already present, to facilitate improved Zone migration portability.
- Ability to L2 filter a VNIC from the globalzone; for, but not limited to...
- Ability to restrict the potentially assigned IP Address(es) to a VNIC
 - ie: How do you keep a customer/user from assigning themselves IP addresses they don't have permission to use?
- Kstat lockdown; *kstat -p link* within a non-global-zone will display information for all links, assigned or not.
- New tool, or improvements to *dladm/netstat*, to see real-time stats on unplumbed data-links.