The Illumos Home Data Center (revisited)

Daniel L. McDonald - OmniOS Engineering

Background

- I gave an earlier version of this talk almost two years ago at Illumos Day, 2012. You can see the talk on YouTube.

- I should be able to give a proper demonstration now.

- My own HDC hardware is now vastly improved.

- And I have one Big Idea at the end.

Illumos

- You can't kill open-source. Illumos is what was once known as OpenSolaris.

- Illumos itself is the OS and Networking (OSNet) parts.

- Several distributions:

  - Openindiana - desktop and closest to old OpenSolaris

  - SmartOS - Joyent's kernel-as-hypervisor cloud platform

  - OmniOS - OmniTI's server-focussed distribution.

OmniOS

- Uses the Image Packaging System (IPS).
- Optimized for servers, and a traditional deployment.
- General-purpose: file serving, compute, and even VMs.
- Solves our problems --> it can solve yours too!

Let's Recap Illumos

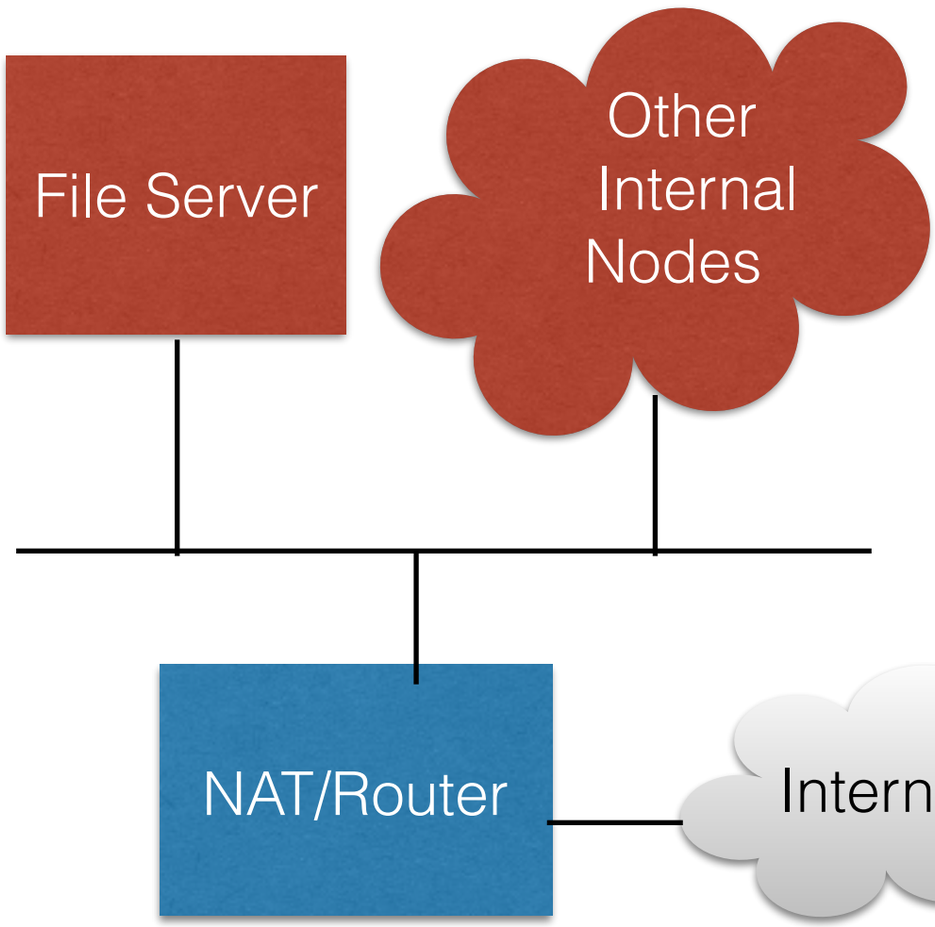
- What comes to mind when you think of Illumos?
    - ZFS
    - DTrace
    - Crossbow
    - Zones
        - Including and especially exclusive-TCP/IP-stack zones.
        - Loopback mounts of global-zone filesystems.
- And people are using these features to Solve Problems.
- But there are more problems we can solve.

Virtual Network Machines

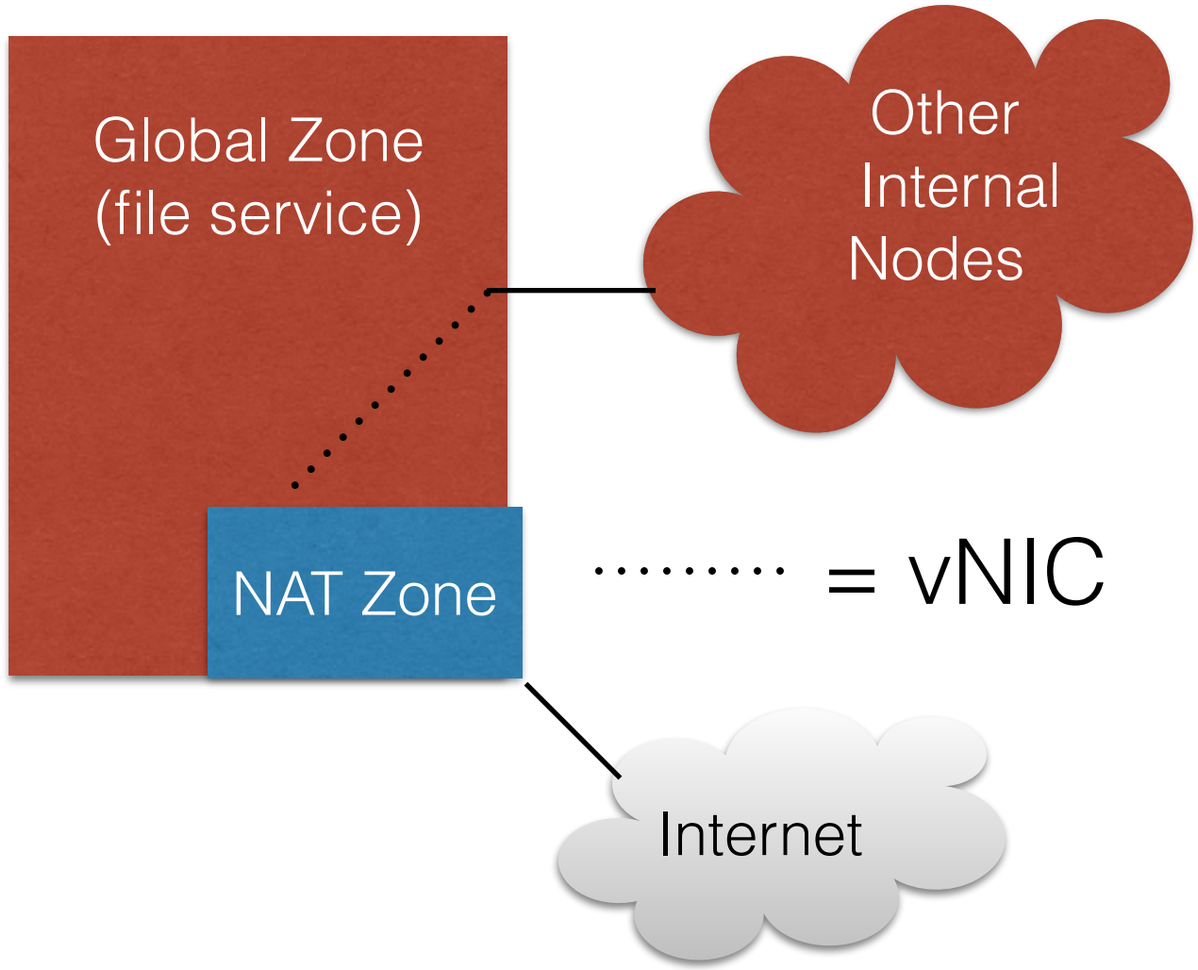
- LISA 2009 Paper by Tripathi, Droux, Belgaid, and Khare
- Combines unique-stack zones, virtual network interfaces (vNICs), and virtual network attachment points (etherstubs) into one.
- A zone can act as a network device:
    - Router
    - NAT
    - Firewall
    - Security Gateway

Example of a Virtual Network Machine

What it looks like:
two machines



What it is:
one machine



Let's Recap Zones

- **Zones are lightweight virtual machines.**
  - Like BSD Jails, or LXC containers.
- They share the same kernel as the global zone.
- Some kernel resources are re-instantiated per zone. (e.g. A TCP/IP stack)
- They are (relatively) cheap.
- Beside getting their own resources, zones can inherit a subset of the global zone's filesystem.
- If one needs a full VM (to run another OS), KVM can run as a zone's boot process.

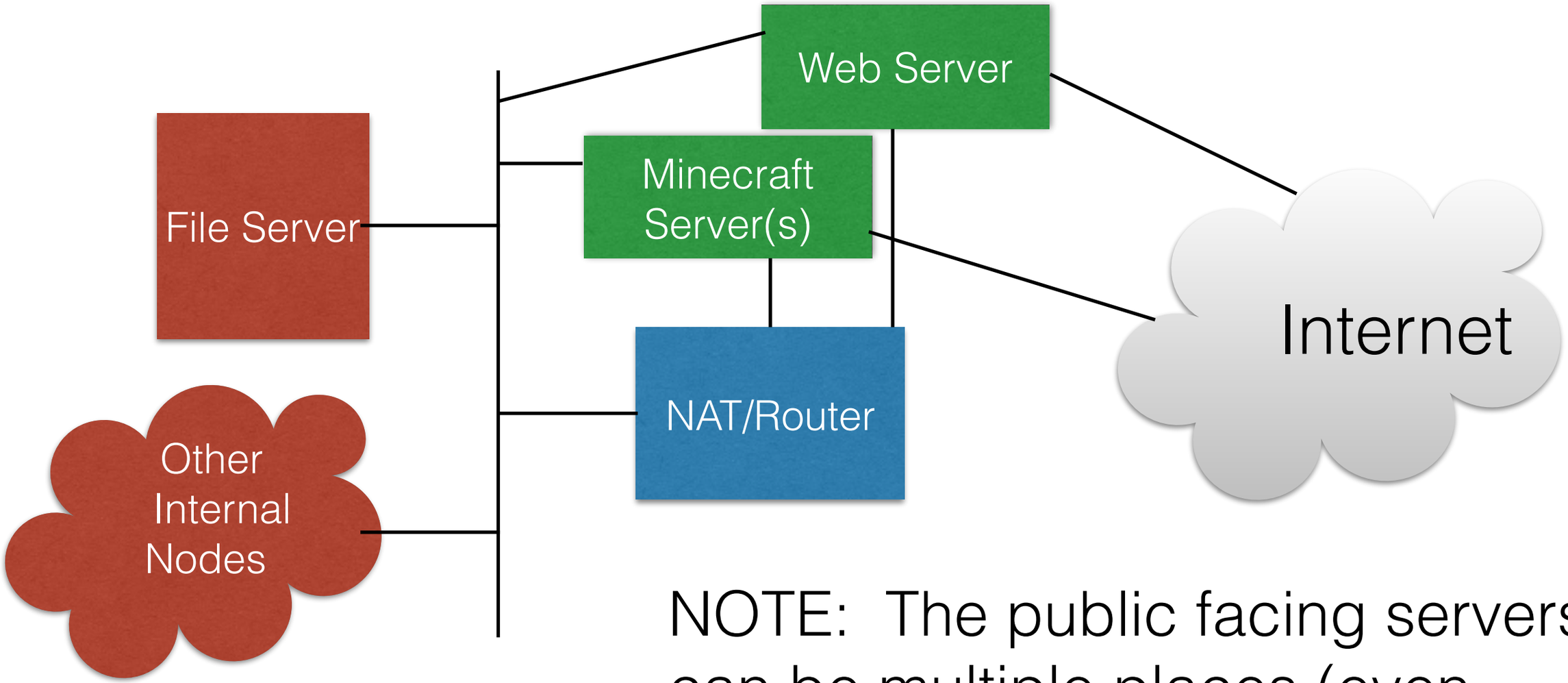
Now Consider the Home (or Even the Small Business)

- Today a home or small business usually has:
  - A NAT box (usually misnamed a "router").
  - A wifi box (often combined with the NAT).
- Sometimes a NAS or fileserver. More common in the small business case.
- A small business that's actually a branch office, or has multiple locations, may connect to other locations with {,d}TLS or IPsec.

Now Consider the Home (or Even the Small Business, contd.)

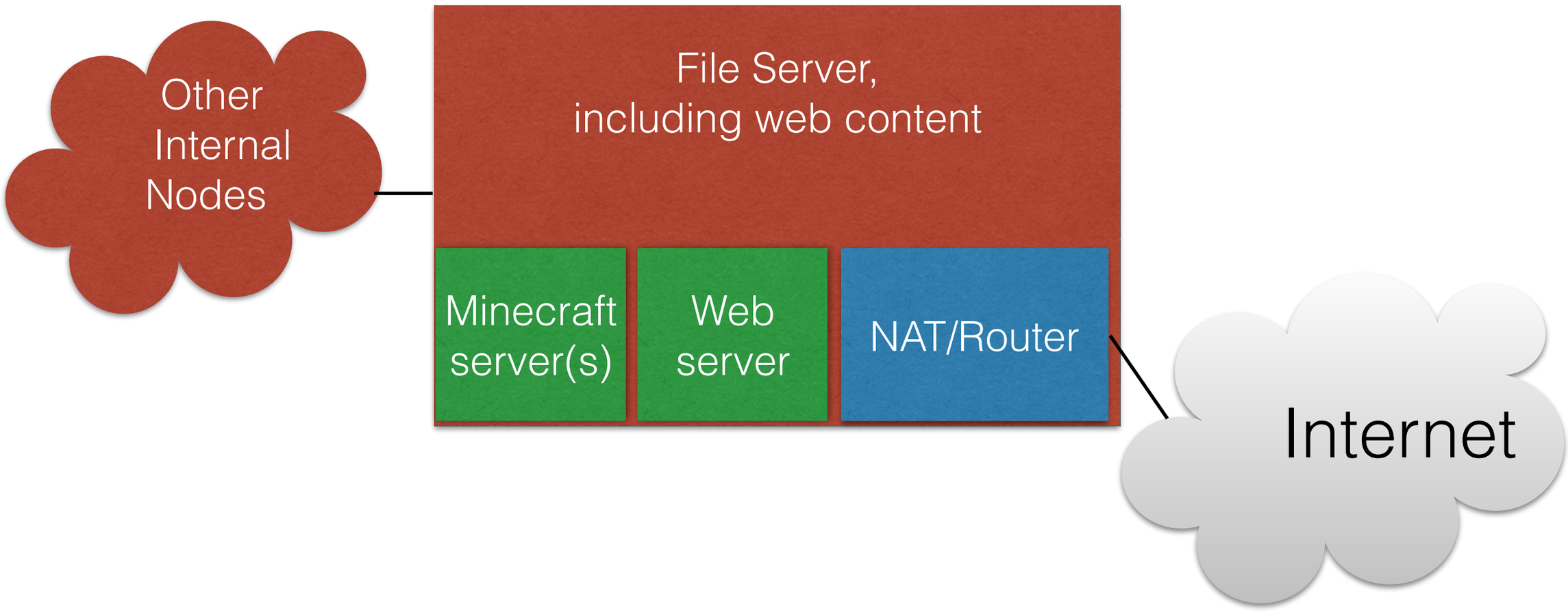
- NAT or wifi boxes can have problems.
    - Tiny, often downrev, Linux or proprietary kernel.
    - Provider-provided ones sometimes have open TCP ports ("For customer service").
    - Bufferbloat.
- Non-trivial NAS or non-trivial VPN can involve real money to real vendors.
- Small business may even have an on-site web server. Homes SHOULD, but let's table that for now.

Let's Take a Look



NOTE: The public facing servers can be multiple places (even multiple machines), depending on your deployment architecture.

Let's Take ANOTHER Look



Zone Recap

- Global zone
    - File services (most are not zone-instantiable yet).
    - Might not need a default route (save when upgrading software).
    - Can provide other local network services if default route is in place.
- Router zone
    - Performs NAT, and possibly firewall services.
        - Port-redirection is important here.
    - Can also VPN to other sites or serve as a remote-access server.
- Webserver zone
    - "As shown on the tin."
    - If Global is not routing off-site, this zone can provide local network services (e.g. DNS).
    - Can inherit (even read-only) filesystems from global, allowing easy drop-off of site content.
- Minecraft zone
    - Can be resource-limited if need be.

In-use Today @kebe.com

- I use loopback mounts (lofs) of global-zone subsets.
  - E.g. /export/home/danmcd/external becomes /export/home/danmcd in the webserver zone.
  - And with a symlink on the global zone: ~danmcd/public_html becomes an easy way to internally edit what is visible on the webserver.
    - E.g. http://kebe.com/~danmcd/
- Global zone has no default route.
  - Also uses a modicum of IPsec "drop" policy to narrow acceptable local IP range.
- NAT directs local TCP port 80 (and soon 443) to/from webserver zone, and Minecraft ports as well.
- Webserver zone serves as DHCP and DNS server for local network.
- Router zone also serves as remote-access server.

So Many Functions, So Little Hardware

- **Downside: Can be a single point of failure.**

- Upside: VERY cheap. Also, lower power consumption if HW is built right.

- Illumos capability can cover most cases today.

- An Illumos capable of running a wifi chip as a base station would complete the picture.

    - Rumor has it someone inside Sun had this working at one point.

Potential Future Work

- **Illumos can benefit from some additional work.**
- Open-source IKE/IKEv2, and/or properly integrated IP-over-{,d}TLS.
- Base-station wifi support.
- For bigger or better software defined networks: Openflow. (May be already done by Pluribus.)
- Atom support?
  - Avoton should be able to do everything we'd want today, assuming it has VT-x with EPT.
- ARM support?
  - With ARMv8, the ZFS 64-bit arithmetic problems go away.

A Big Idea

- Services like Facebook, Instagram, etc. turn you into the product. You don't OWN your data.

- A **properly-productized** HDC could allow people to eschew these services. Or better control them.

- There would be a lot of barriers:

  - Properly productizing this would be HARD (but not untenable).

  - Last-mile ISPs don't want to traffic engineer for bidirectional traffic.

  - Some outsourced infrastructure would be needed (e.g. DNS).

  - Fun with the National Security State? :)

Wrapping Up

- **Exploit illumos for routing/forwarding path uses as well!!!**

- Maximize ALL of your hardware with zones (including NICs).

- A small box can do big jobs with the right software and configuration.

- Demo!

- Any questions?

Daniel L. McDonald
Twitter: @kebesays
Blog: http://kebesays.blogspot.com/
WWW: http://kebe.com/~danmcd/