Datacenter 101

Dazhao Cheng

What is a Datacenter?

- ***"A data center is a facility used to house computer systems and associated components, such as telecommunications and storage systems. It generally includes redundant or backup power supplies, redundant data communications connections, environmental controls (e.g., air conditioning, fire suppression) and security devices."***

  (source: http://en.wikipedia.org/wiki/Datacenter)

Taking a Step Back: Why Do We Really Want?

- **We want lots of computing power**
- **Why?**
  - Corporate Infrastructure: employees need to read email
  - Academic computing: students need to do assignments
  - Scientific computing: NASA needs to do galaxy simulations
  - Distributed computation: Google needs to index the web

What's the Big Deal? We've Got a Computer Cluster…

- **Why not just use…**
  - Individual PCs for everyone?
  - A computer cluster where people can do assignments?
- **But this is very inefficient and has issues**
  - Rooms to store these computers
  - Requirements for shared storage
  - Staff to maintain machines (TCO)
  - Cabling
  - Remote access?
  - Poor amortization of resources

Ok, Fine: Let's Just Shove Some Computers in a Room

- **Yes: now we have centralized rooms to maintain. TCO is reduced.**
- **But…**
  - What sorts of machines do you use? Not PCs: too much space.
  - PCs get hot. You need to keep them cool.
  - What about amortization of resources?
  - Remote Access strategy?
  - Shared storage?
  - Cabling?

But There's More: What About Availability?

- **Datacenters have strict standards for reliability and availability**
  - Tier 1: 99.671% Availability: 28 hours of downtime/year
  - Tier 2: 99.741% Availability: 22 hours of downtime/year
  - Tier 3: 99.982% Availability: 1.5 hours of downtime/year
  - Tier 4: 99.995% Availability: 26 minutes of downtime/year

We Need a Plan

- **What we need is a new architecture to support our compute and availability requirements.**

So, Finally, What *is* a Datacenter?

- **A datacenter is a collection of computers and storage**
- **A datacenter has many purposes**
  - Provide single location for compute infrastructure for big organizations
  - Provide hosting facilities so companies don’t have to manage their own IT
  - Perform large distributed computations (Google Warehouse computing, scientific simulations)
  - Provide scale-out on-demand IT services (e.g., accommodating temporary usage spikes)

And Why is Datacenter Design Interesting?

- **Analogy to computer architecture**
  - There are many issues that influence the performance/reliability of a desktop
    - Choice of CPU
    - Cache designs and memory hierarchies
    - OS
    - Storage design (e.g., RAID0-RAID5, etc.)
    - Power/Cooling
  - Similar issues are present in a datacenter
    - Storage subsystem
    - Networking infrastructure
    - Power distribution
    - Reliability guarantees (mission-critical?)

So, What are the Rules for Datacenter Design?

- **Standard for Datacenter design: TIA-942**
- **Basic principles**
  - Use racks and rack-mount form factor for high density of compute resources
  - Heat harms components: control it
  - Keep cabling manageable
  - Obey FCC rules for EMI/RFI

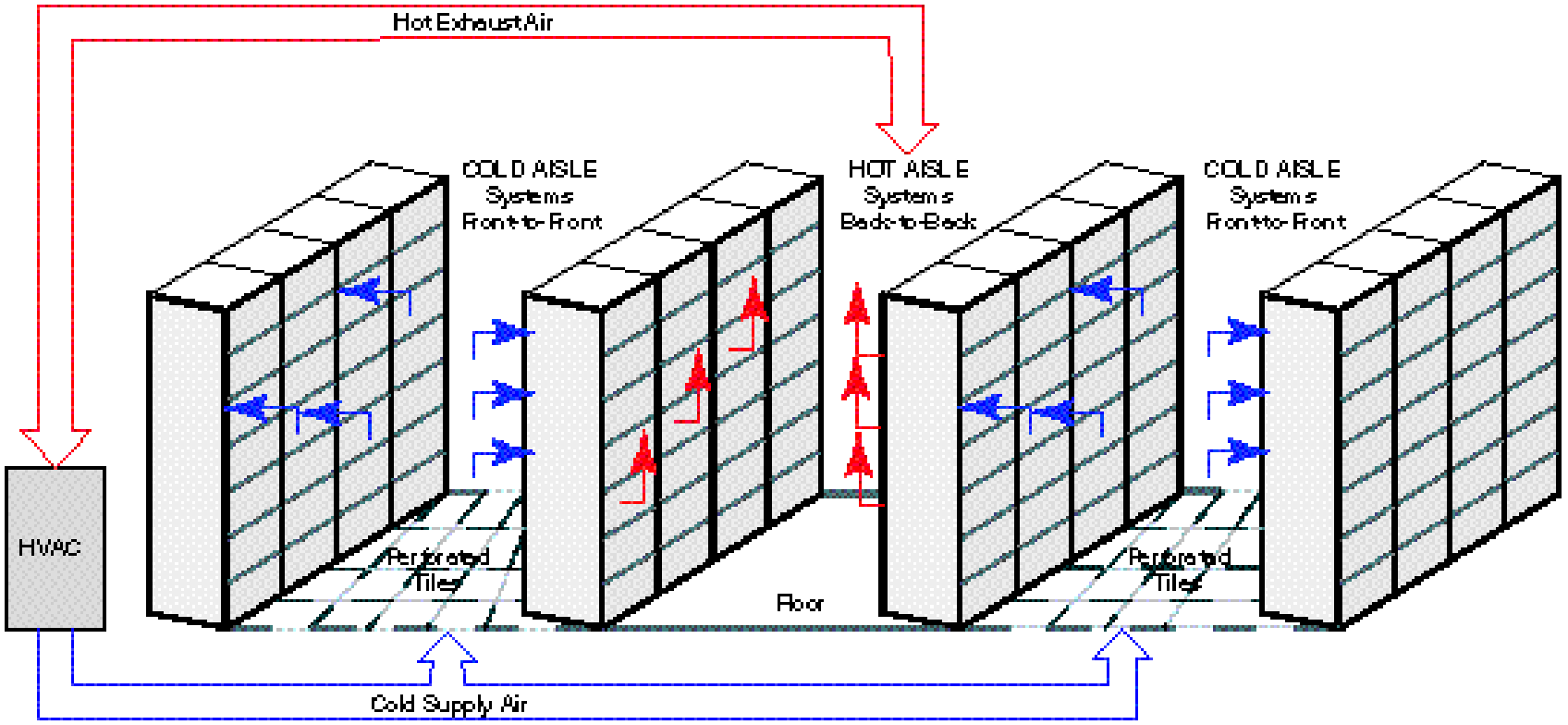
Rack Internals

- **What goes inside a rack?**
  - Servers
  - Storage
  - Cables
  - Power distribution units (PDUs)
  - Fans
  - Switches
  - KVM

Cooling in a Datacenter

- Inflow: Cool Air from HVAC.
- Outflow: Warmed Air. Eventually cooled by HVAC.
- Never let warmed air re-circulate to inflow or mix with cool air
- Same idea as in a standard PC chassis
- Hot Aisle/Cold Aisle

Hot Aisle/Cold Aisle

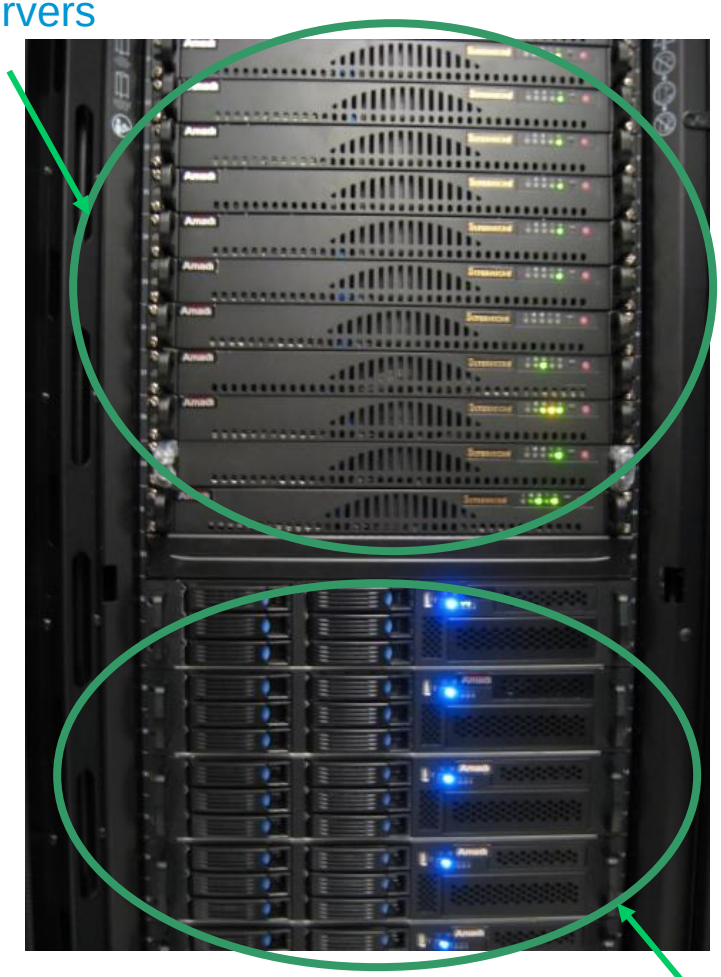


Scenes from a Typical Datacenter (1)

- Racks for servers

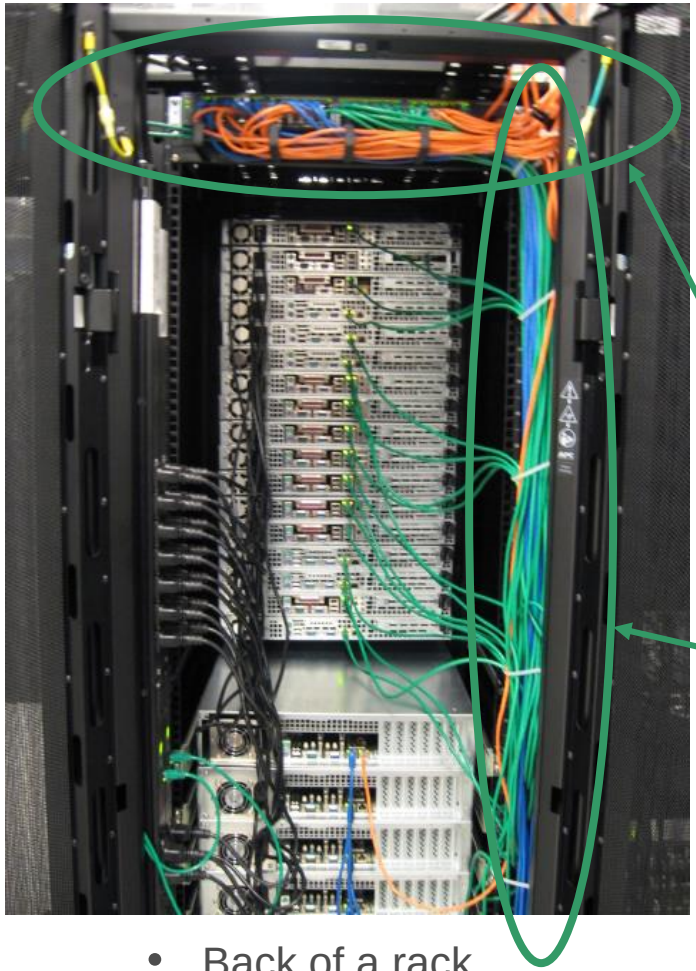
Scenes from a Typical Datacenter (2)

Servers



Storage

- Front of a rack



Network or Storage Cables

- Back of a rack

Scenes from a Typical Datacenter (3)

- HVAC (Heating, Ventilation, and Cooling) for servers

Scenes from a Typical Datacenter (4)

- Power cables above racks

Datacenter Internals

- **Let's delve into some details of some important datacenter pieces.**
  - Servers
  - Storage
  - Switches
  - Power distribution
  - Management software

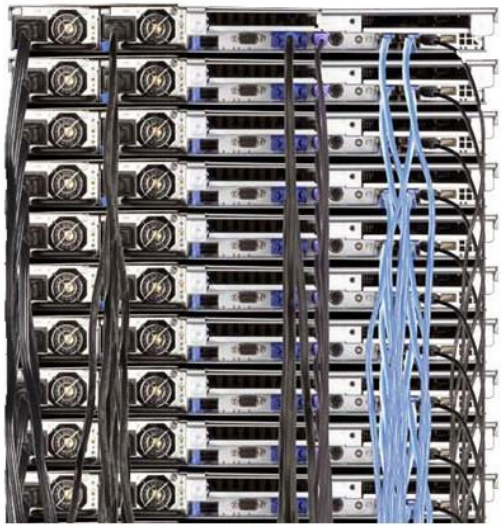
Datacenter Internals (1): Server

- **Server**
  - Similar to a simple desktop
  - Different form factor (rack unit, blade)
  - Different number of CPUs from desktop
  - Different applications (less graphics)
  - Headless, with KVM or console access
  - Different peripherals (networking, shared storage)

Rackmounts vs. Blades

- Rackmount server (2U)

- 1U servers in a rack



Ten 1U dual processor servers and the associated cabling.

- Blade servers

Datacenter Internals (2): Storage

- **Storage**
  - Local storage (e.g., SCSI)
    - Only one host can access at a time
    - Block-oriented
  - Shared storage
    - Multi-host access
    - Many protocols
      - SAN: block-oriented
      - NAS: file-oriented
      - iSCSI: block-oriented

21

Shared-Storage Protocols

- **SAN (storage-attached network). Example: Fibre Channel**
  - Remote storage looks like a local SCSI block device
  - Use special fiber-optic cables for high-speed
- **NAS: network-attached storage (e.g., NFS, AFS)**
  - Server attached to disks
  - Remote hosts “mount” files on server: file-based access
- **iSCSI**
  - Remote storage looks like a local SCSI block device
  - Uses TCP/IP as the transport mechanism.
  - Cheaper than SAN/FC (cabling, HBA cost)

Datacenter Internals (3): Switches

- **Switches**
  - Network switches
    - Connect servers to each other
    - Connects servers to NAS devices.
    - Connect switches to other switches (hierarchy)
    - Used for communication within a rack and between racks
  - Storage switches
    - Connects servers to their storage
      - Host to SAN device.
      - Host to iSCSI device.
    - Connects storage switches to other switches.
    - Usually use fiber-optic cabling.

Datacenter Internals (4): Cabling and Infrastructure

- **Cabling**
  - Fiber vs. Copper
  - Distance vs. Cost
- **Network types**
  - Ethernet: 10/100 and GigE; T1, T3, etc.
  - Scientific computing: Myrinet, Quadrics, Infiniband (different bandwidth/latency characteristics)
- **Industry: 10GigE mostly going forward.**

Datacenter Internals (5): Power Distribution

- **Components/Issues**
  - PDU: Power distribution unit
  - Power inside a rack
  - Power across racks
  - Redundancy and surge protection (UPS)

Datacenter Internals (6): Management software

- **Requirements**
  - Alerts
  - Usage statistics
  - Remote access
  - Logging for *post mortem* analysis
  - Ease of deployment for software or upgrades

Other Issues

- AC vs. DC
- Position of server within rack (hot air rises!)
- Switch topology in the LAN/SAN (impacts latency)
- Knowing when something breaks (alerts)
- Cost of cooling (power & money)

Challenges

- **Large-scale system management**
  - How do you administer 1000s of servers at once?
  - How do you design for component failure?
- **Power consumption**
  - 2005: Datacenters ~1.2% of electricity in US
  - Overall cost expected to increase tenfold from 2005 to 2012 ($18.5B to $250B).
- **Space**
  - $19B in construction projects as of June 2008

Future directions

- Virtualization of Datacenters and within the datacenter
- Cloud computing
- On-premise and off-premise computing with seamless migration between the two
- Green Grid/Green Computing

Modularity, Modularity, Modularity

- Containers

- Many containers

References (1)

- **Basic Definitions**
  - http://www.webopedia.com/TERM/D/data_center_tiers.html
  - http://en.wikipedia.org/wiki/Datacenter
  - http://www.webopedia.com/quick_ref/blade_servers.asp
- **Pictures of a Datacenter**
  - http://steadfast.net/forum/showthread.php?t=574
- **Datacenter and Rack Standards Documents**
  - TIA-942 (Datacenter Standard): http://*www.tiaonline.org/**standards***
  - EIA-310-d (Rack Standard): http://electronics.ihs.com/collections/abstracts/eia-310.htm

References (2)

- **Racks and Airflow**
  - http://docs.sun.com/source/816-1613-14/Chapter3.html
  - http://en.wikipedia.org/wiki/Rack-mounted
- **Cabling**
  - http://en.wikipedia.org/wiki/Category_5_cable
  - http://en.wikipedia.org/wiki/Category_6_cable
- **Switches**
  - www.gocomsys.com/product-links.html
  - http://www.myri.com/
  - http://www.infinibandta.org/home
  - http://www.quadrics.com/quadrics/QuadricsHome.nsf/DisplayPages/Homepage

Reference (3)

- **Power usage in datacenters**
  - http://enterprise.amd.com/Downloads/svrpwrusecompletefinal.pdf
  - http://greenlight.greentechmedia.com/2008/06/26/data-center-power-consumption-by-the-numbers-341/
  - http://blogs.computerworld.com/node/844
  - Fan, et al, *Power Provisioning for a Warehouse-sized Computer*, In *Proceedings of the International Symposium on Computer Architecture (ISCA) 2007,* pp. 13-23
- **Future Directions**
  - http://www.vmware.com/resources/techresources/1080
  - http://en.wikipedia.org/wiki/The_Green_Grid