Datacenters

Mendel Rosenblum
Evolution of datacenters

- 1960's, 1970's: a few very large time-shared computers
- Today and into the future:
  - Datacenter contains large numbers of nearly identical machines
  - Individual applications use thousands of machines simultaneously
- Companies consider datacenter technology a trade-secret
  - No public discussion of the state of the art from industry leaders
Typical specs for a datacenter today

- 15-40 megawatts power (Limiting factor)
- 50,000-200,000 servers
- $1B construction cost
- Onsite staff (security, administration): 15
Rack

- Typically is 19 or 23 inches wide
- Typically 42 U
  - U is a Rack Unit - 1.75 inches

- Slots:
Rock Slots

- Slots hold power distribution, servers, storage, networking equipment

- Typical server: 2U
  - 8-128 cores
  - DRAM: 32-512 GB

- Typical storage: 2U
  - 30 drives

- Typical Network: 1U
  - 72 10GB
Row/Cluster

- 30+ racks
Networking - Switch locations

- **Top-of-rack switch**
  - Effectively a cross-bar connecting machines in rack
  - Multiple links going to end-of-row routers

- **End-of-row router**
  - Aggregate row of machines
  - Multiple links going to core routers

- **Core router**
  - Multiple core routers
Multipath routing

Many paths between servers

equal performance paths
Ideal: "full bisection bandwidth"

- Would like network like cross-bar
  - Everyone has a private channel to everyone else

- In practice today: some oversubscription (can be as high as 100x)
  - Assumes applications have locality to rack or row) but this is hard to achieve in practice.
  - Some problem fundamental: Two machines transferring to the same machine

- Consider where to place:
  - Web Servers
  - Memcache server
  - Database servers - Near storage slots

- Current approach: Spread things out
Power Usage Effectiveness (PUE)

- Early datacenters built with off-the-shelf components
  - Standard servers
  - HVAC unit designs from malls
- Inefficient: Early datacenters had PUE of 1.7-2.0

PUE ratio = \[ \frac{\text{Total Facility Power}}{\text{Server/Network Power}} \]

- Best-published number (Facebook): 1.07 (no air-conditioning!)
- Power is about 25% of monthly operating cost
Energy Efficient Data Centers

- Better power distribution - Fewer transformers
- Better cooling - use environment (air/water) rather than air conditioning
  - Bring in outside air
  - Evaporate some water
- Hot/Cold Aisles:
- IT Equipment range
  - OK to +115°F
  - Need containment
Backup Power

- Massive amount of batteries to tolerate short glitches in power
  - Just need long enough for backup generators to startup
- Massive collections of backup generators
- Huge fuel tanks to provide fuel for the generators
- Fuel replenishment transportation network (e.g. fuel trucks)
Fault Tolerance

- At the scale of new datacenters, things are breaking constantly
- Every aspect of the datacenter must be able to tolerate failures
- **Solution: Redundancy**
  - Multiple independent copies of all data
  - Multiple independent network connections
  - Multiple copies of every service
Failures in first year for a new datacenter (Jeff Dean)

~ thousands of hard drive failures
~ 1000 individual machine failures
~ dozens of minor 30-second blips for DNS
~ 3 router failures (have to immediately pull traffic for an hour)
~ 12 router reloads (takes out DNS and external VIPs for a couple minutes)
~ 8 network maintenances (4 might cause ~30-minute random connectivity losses)
~ 5 racks go wonky (40-80 machines see 50% packet loss)
~ 20 rack failures (40-80 machines instantly disappear, 1-6 hours to get back)
~ 1 network rewiring (rolling ~5% of machines down over 2-day span)
~ 1 rack-move (plenty of warning, ~500-1000 machines powered down, ~6 hours)
~ 1 PDU failure (~500-1000 machines suddenly disappear, ~6 hours to come back)
~ 0.5 overheating (power down most machines in <5 mins, ~1-2 days to recover)
Choose datacenter location drivers

- Plentiful, inexpensive electricity
  - Examples - Oregon: Hydroelectric; Iowa: Wind
- Good network connections
  - Access to the Internet backbone
- Inexpensive land
- Geographically near users
  - Speed of light latency
  - Country laws (e.g. Our citizen's data must be kept in our county.)
- Available labor pool
Google Data Centers

**Americas**
- Berkeley County, South Carolina
- Council Bluffs, Iowa
- Douglas County, Georgia
- Quilicura, Chile
- Jackson County, Alabama
- Mayes County, Oklahoma
- Lenoir, North Carolina
- The Dalles, Oregon

**Asia**
- Changhua County, Taiwan
- Singapore

**Europe**
- Hamina, Finland
- St Ghislain, Belgium
- Dublin, Ireland
- Eemshaven, Netherlands
Google Data Center - Council Bluffs, Iowa, USA
Google datacenter pictures: Council Bluffs