Concepts Introduced in Chapter 6

- introduction to warehouse-scale computing
- programming models
- infrastructure and costs
- cloud computing

Warehouse-Scale Computers

- A *cluster* is a collection of desktop computers or servers connected together by a local area network to act as a single larger computer.
- A *warehouse-scale computer* (WSC) is a cluster comprised of tens of thousands of servers.
- The cost may be on the order of $150M for the building, electrical and cooling infrastructure, the servers, and the networking equipment that houses 50,000 to 100,000 servers.
- A WSC can be used to provide internet services.
  - search - Google
  - social networking - Facebook
  - video sharing - YouTube
  - online sales - Amazon
  - cloud computing services - Rackspace
  - and many more applications

Important Design Factors for WSCs

- WSC goals and requirements in common with servers.
  - cost-performance - work done per dollar
  - energy efficiency - work done per joule
  - dependability via redundancy
  - network I/O
  - interactive and batch processing workloads
- WSC aspects that are distinct from servers.
  - Ample parallelism is always available in a WSC.
  - Operational costs represent a greater fraction of the cost of a WSC.
  - Customization is easier for the scale of an WSC.

Programming Models for WSCs

- MapReduce (or the open source Hadoop) is the most popular framework for batch processing in a WSC.
  - *Map* applies a programmer-supplied function to each logical input record to produce a set of key-value pairs.
  - *Reduce* collapses these values using another programmer-supplied function.
- Both tasks are highly parallel.

Programming Models for WSCs (cont.)

- There is often a high variability in performance between the different WSC servers due to a variety of reasons.
  - varying load on servers
  - file may or may not be in a file cache
  - distance over network can vary
  - hardware anamolies
- A WSC will start backup executions on other nodes when tasks have not yet completed and take the result that finishes first.
- Rely on data (file) replication to help with read performance and availability.
- A WSC also has to cope with variability in load.
  - servers
  - entire WSC
- Often WSC services are performed with in-house software to reduce costs and optimize for performance.

Storage for a WSC

- A WSC uses local disks inside the servers as opposed to network attached storage (NAS).
- The Google file system (GFS) uses local disks and maintains at least three replicas to improve dependability by covering not only disk failures, but also power failures to a rack or a cluster of racks by placing the replicas on different clusters.
- A read is serviced by one of the three replicas, but a write has to go to all three replicas.
- Google uses a relaxed consistency model in that all three replicas have to eventually match, but not all at the same time.

WSC Networking

- A WSC uses a hierarchy of networks for interconnection.
- The standard rack holds 48 servers connected by a 48-port Ethernet switch.
- A rack switch has 2 to 8 uplinks to a higher switch. So the bandwidth leaving the rack is 6 (48/8) to 24 (48/2) times less than the bandwidth within a rack.
- There are array switches that are more expensive to allow higher connectivity.
- There may also be Layer 3 routers to connect the arrays together and to the Internet.
- The goal of the software is to maximize locality of communication relative to the rack.

WSC Location

- proximity to Internet backbone optical fibers
- proximity to users of service to reduce Internet access latency
- electricity availability and cost
- property tax rate
- low risk from environmental disasters
- stability of country
- low temperature to decrease cooling cost

WSC Power and Cooling

- power usage of just the WSC IT equipment
  - 33% for processors
  - 30% for DRAM
  - 10% for disks
  - 5% for networking
  - 22% for other components within the servers
- Air conditioning is used to cool server room, requiring 10%-20% of IT equipment power due mostly to fans.
- Chilled water is often used to cool the air, requiring 30% to 50% of IT equipment power. Outside cooling towers can leverage lower outside temperature.

Measuring WSC Efficiency

- power
  - power utilization effectiveness (PUE) is a widely used simple metric.
  - Median PUE reported in a 2006 study was 1.69.

$$PUE = \frac{total\_facility\_power}{IT\_equipment\_power}$$

- performance
  - Bandwidth is an important metric as there may be many simultaneous user requests or metadata generation batch jobs.
  - Latency is also an important metric as it is seen by users when they make requests. Users will use a search engine less as the response time increases. Also users are more productive in responding to interactive information when the system response time is faster as they are less distracted.

Google WSC Innovations to Improve Energy Efficiency

- Modified server containers.
  - Separated hot and cold chambers to reduce variation in air temperative, which allows air to be delivered at higher temperatures due to less severe worst-case hot spots.
  - Operating servers at higher temperatures allowed use of cooling towers instead of the more inefficient traditional chillers.
  - Shrunk distince of air circulation loop to reduce energy required to move air.
- Located WSCs in more temperate climates to allow more use of evaporative cooling.
- Deployed extensive monitoring to measure actual PUE.
- Designed motherboards that only need a single 12-volt supply so that a UPS could be provided using standard batteries with each server.
- Google PUE was 1.23 in 2007 and was 1.12 in 2011.

Cost of a WSC

- capital expenditures (CAPEX)
  - CAPEX is the cost to build a WSC, which includes the building, power and cooling infrastructure, and initial IT equipment (servers and networking equipment).
- operational expenditures (OPEX)
  - OPEX is the cost to operate a WSC, which includes buying replacement equipment, electricity, and salaries.

Advent of Cloud Computing

- Cloud computing can be thought of as providing computing as a utility, where a customer pays for only what they use, just as we do for electricity. Cloud computing relies on increasingly larger WSCs which provide several benefits if properly set up and operated.
  - improvements in operational techniques
  - economies of scale
  - reduces customer risks of over-provisioning or under-provisioning

Improvements in Operational Techniques

- WSCs have led to innovations in system software to provide high reliability.
  - failover - Automatically restarting an application that fails without requiring administrative intervention.
  - firewall - Examines each network packet to determine whether or not it should be forwarded to its destination.
  - virtual machine - A software layer that executes applications like a physical machine.
  - Protection against denial-of-service attacks.

Economies of Scale

- WSCs offer economies of scale that cannot be achieved with a data center.
  - 5.7 times reduction in storage costs
  - 7.1 times reduction in administrative costs
  - 7.3 times reduction in networking costs
  - volume discount price reductions
  - PUE of perhaps 1.2 versus PUE of 2.0 for a data center
  - better utilization of WSC by being available to the public

Reducing Customer Risks

- WSCs reduce risks of over-provisioning or under-provisioning, particularly for start-up companies.
  - Providing too much equipment means overspending.
  - Providing too little equipment means demand may not be able to be met, which can give a bad impression to potential new customers.

Amazon Web Services

- Amazon offered Amazon Simple Storage Service (Amazon S3) and Amazon Elastic Computer Cloud (Amazon EC2) in 2006.
- Relied on virtual machines.
  - Provides better protection for users.
  - Simplified software distribution within a WSC.
  - The ability to reliably kill a virtual machine made it easier to control resource usage.
  - Being able to limit use of resources simplified providing multiple price points for customers.
  - Improved flexibility in server configuration.
- Relied on open source software.
- Provided service at very low cost.
- No contract required.

Fallacies and Pitfalls

- Fallacy: Capital costs of a WSC facility are higher than the servers that it houses.
- Pitfall: Trying to save power with inactive low power modes versus active low power modes.
- Pitfall: Using too wimpy a processor when trying to improve WSC cost-performance.
- Fallacy: Replacing all disks with Flash memory will improve cost-performance of a WSC.