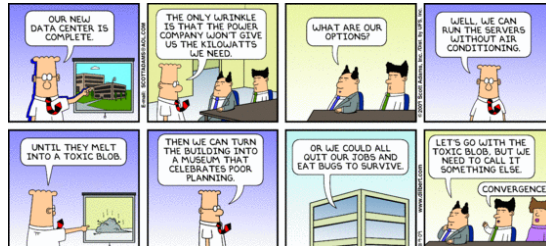


Advanced Computer Architecture Datacenters

Fall
2016



Pejman Lotfi-Kamran

Adapted from slides originally developed by Profs. Hill, Hoe, Falsafi and Wenisch of CMU, EPFL, Michigan, Wisconsin

Fall 2016

Lec 25 - Slide 1

Where Are We?

Fr	Sa	Su	Mo	Tu
	27-Shahrivar		29-Shahrivar	
	3-Mehr		5-Mehr	
	10-Mehr		12-Mehr	
	17-Mehr		19-Mehr	
	24-Mehr		26-Mehr	
	1-Aban		3-Aban	
	8-Aban		10-Aban	
	15-Aban		17-Aban	
	22-Aban		24-Aban	
	29-Aban		1-Azar	
	6-Azar		8-Azar	
	13-Azar		15-Azar	
	20-Azar		22-Azar	
	27-Azar		29-Azar	
	4-Dey		6-Dey	

Fall 2016

Lec 25 - Slide 2

◆ This Lecture
● Data Center (1)

◆ Next Lecture:
● Data Center (2)

The following slides are from
Adam Wierman of CalTech

Fall 2016

Lec 25 - Slide 3

What is a datacenter?

◆ Oxford dictionary defines “data center” as:

“A **large group of networked computer servers** typically used by organizations for the **remote storage, processing, or distribution of large amounts of data.**”

◆ Term originated in the 1990s with the advent of client-server architecture.

◆ Dot-com bubble → internet data centers.

Fall 2016

Lec 25 - Slide 4

Growth of large internet services...



Warehouse-scale computers

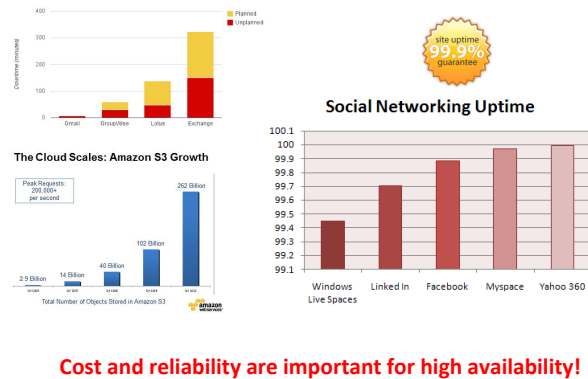
- ◆ **Program**
 - ◆ Internet service (e.g. web search, email, video streaming, maps, etc.).
- ◆ **Computer**
 - ◆ Thousands of individual computing nodes.
 - ◆ Networking and storage subsystems.
 - ◆ Power distribution and cooling system.



Fall 2016

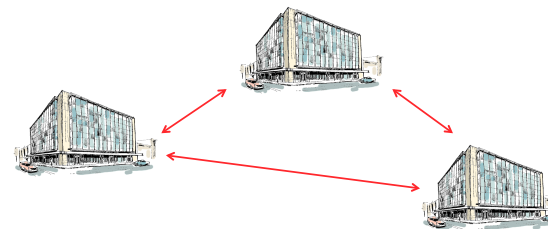
Lec 25 - Slide 6

Datacenter availability requirements



Multi-datacenter scenarios

- ◆ User queries may involve computation across multiple datacenters.
- ◆ Inter-datacenter communications are of much poorer quality than intra-datacenter communications.



Fall 2016

Lec 25 - Slide 8

Architectural overview



Fall 2016

Lec 25 - Slide 9

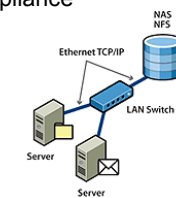
Storage management

Distributed File System

- ◆ Disk drives are directly attached to server nodes
- ◆ Replication across different machines
- ◆ Poorer write performance
- ◆ Higher read performance
- ◆ Can exploit data locality
- ◆ Google File System(GFS)

Network Attached Storage

- ◆ Disk drives are connected to cluster-level switch
- ◆ Replication within each appliance



Fall 2016

Lec 25 - Slide 10

Network fabric

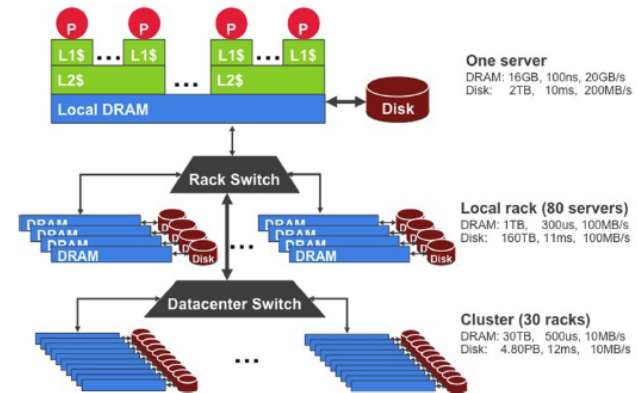


Software should exploit rack-level locality!

Fall 2016

Lec 25 - Slide 11

Storage hierarchy



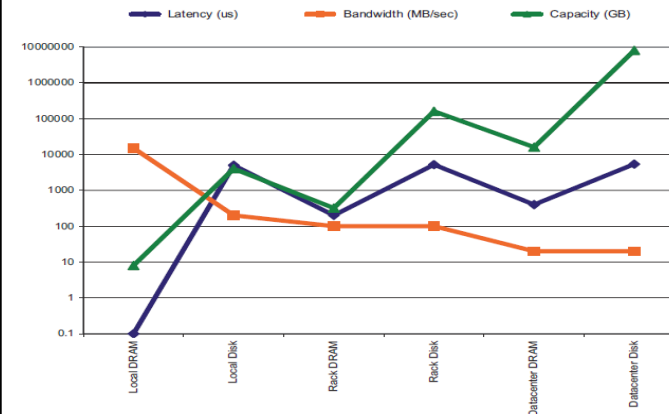
Fourth-Generation Datacenters

- ◆ 17x football field
- ◆ 10^6 cores
- ◆ 10^{17} bytes
- 20 MW

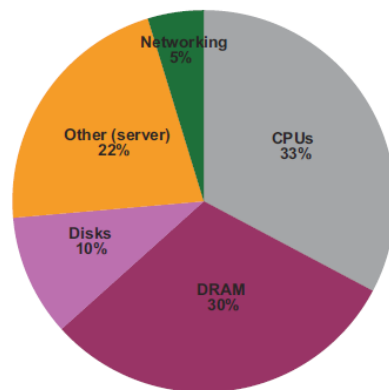


Fall 2016

Performance variations



Power usage

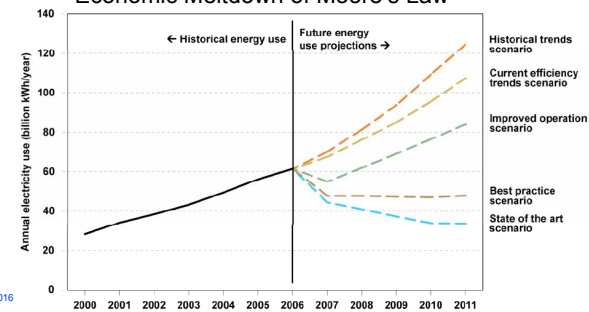


Fall 2016

Lec 25 - Slide 15

Datacenter Power

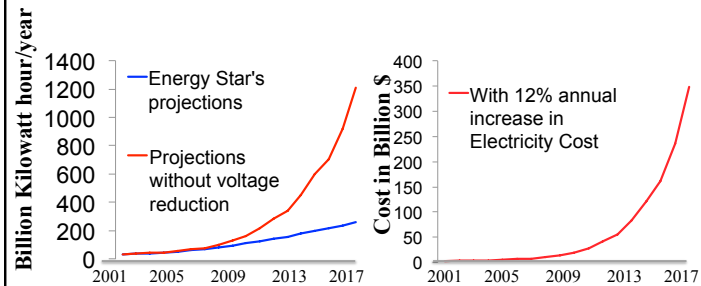
- ◆ 2% of worldwide electricity
 - Growing at 17% annually
- ◆ Referred to by Kenneth Brill as the "Economic Meltdown of Moore's Law"



Fall 2016

Slide 16

Extrapolated from Energy Star (2007)



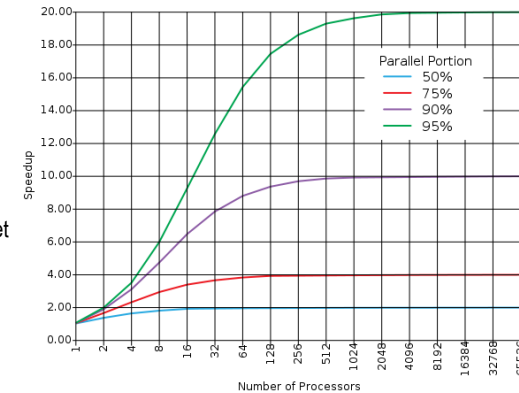
◆ An unsustainable trend!

Fall 2016

Lec 25 - Slide 17

Don't Forget End of Dennard Scaling

- ◆ Limits of Manycores (Amdahl's Law)
- ◆ Lack of Vdd scaling will offset gains from parallelism



Fall 2016

Lec 25 - Slide 18

Software layers

- ◆ **Platform-level software** – common firmware, kernel, operating system distribution, and libraries.
- ◆ **Cluster-level infrastructure** – distributed file systems, schedules, and remote procedure call (RPC) layers.
- ◆ **Application-level software** – specific services.
 - ◆ Online services: web search, email, maps.
 - ◆ Offline services: building of web index.

Fall 2016

Lec 25 - Slide 19

Datacenter software development

- ◆ Applications have inherent data parallelism or request **parallelism**.
- ◆ Each platform generation has significant **homogeneity**.
- ◆ Isolation of users from service implementation makes it much **easier to deploy new software quickly**.
- ◆ Cluster-level software must deal with **expected frequent hardware failure**.

Fall 2016

Lec 25 - Slide 20

Performance and availability toolbox

	Performance	Availability
Replication	Yes	Yes
Sharding (partitioning)	Yes	Yes
Load-balancing	Yes	
Health checking and watchdog timers		Yes
Integrity checks		Yes
Application-specific compression	Yes	
Eventual consistency	Yes	Yes

Fall 2016

Lec 25 - Slide 21

Cluster-level infrastructure software

- ◆ **Resource management**
 - ◆ Maps user tasks to hardware resources.
 - ◆ Enforces priorities and quotas.
 - ◆ Users should be able to specify job requirements at a relatively high level
 - e.g., CPU performance, memory capacity, bandwidth
 - ◆ Increasingly important that cluster schedulers consider **power limitations** and **energy usage optimization**.
- ◆ Hardware abstraction
- ◆ Deployment and maintenance
- ◆ Programming frameworks

Fall 2016

Lec 25 - Slide 22