Operating Systems

Fall 2014

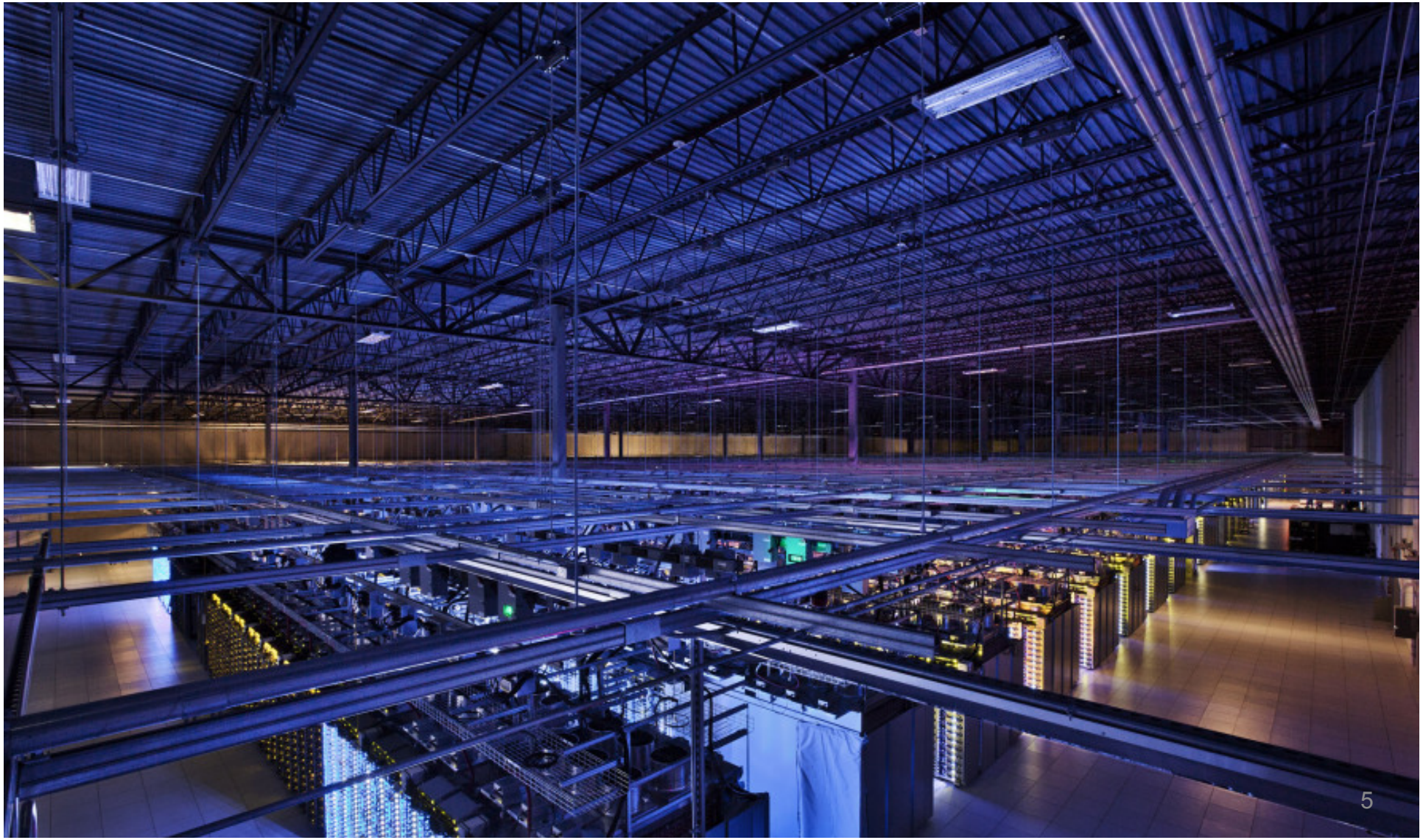
Cloud Computing and Data Centers

Myungjin Lee
myungjin.lee@ed.ac.uk

Google data center locations

A closer look

Inside data center



- A datacenter has 50 - 250 containers
- A container has 1,000 - 2,000 servers
- A server has two processors, 2 disks, tons of memory, battery backup
- Processors are chosen for power efficiency, not performance

Some facts about data centers

- Google has ~0.9 million servers in its all DCs
  - 260M watts of power = 0.01% of global energy
- Facebook processes 750TB of data every day
  - Around 7PB of photo storage from its facility every month
- Amazon serves ~40 PB of videos per month
  - Around 450,000 servers
- Microsoft has 1,000,000 servers
  - It has so far spent around $23 billion

What do these numbers imply?

- Fueling the Internet

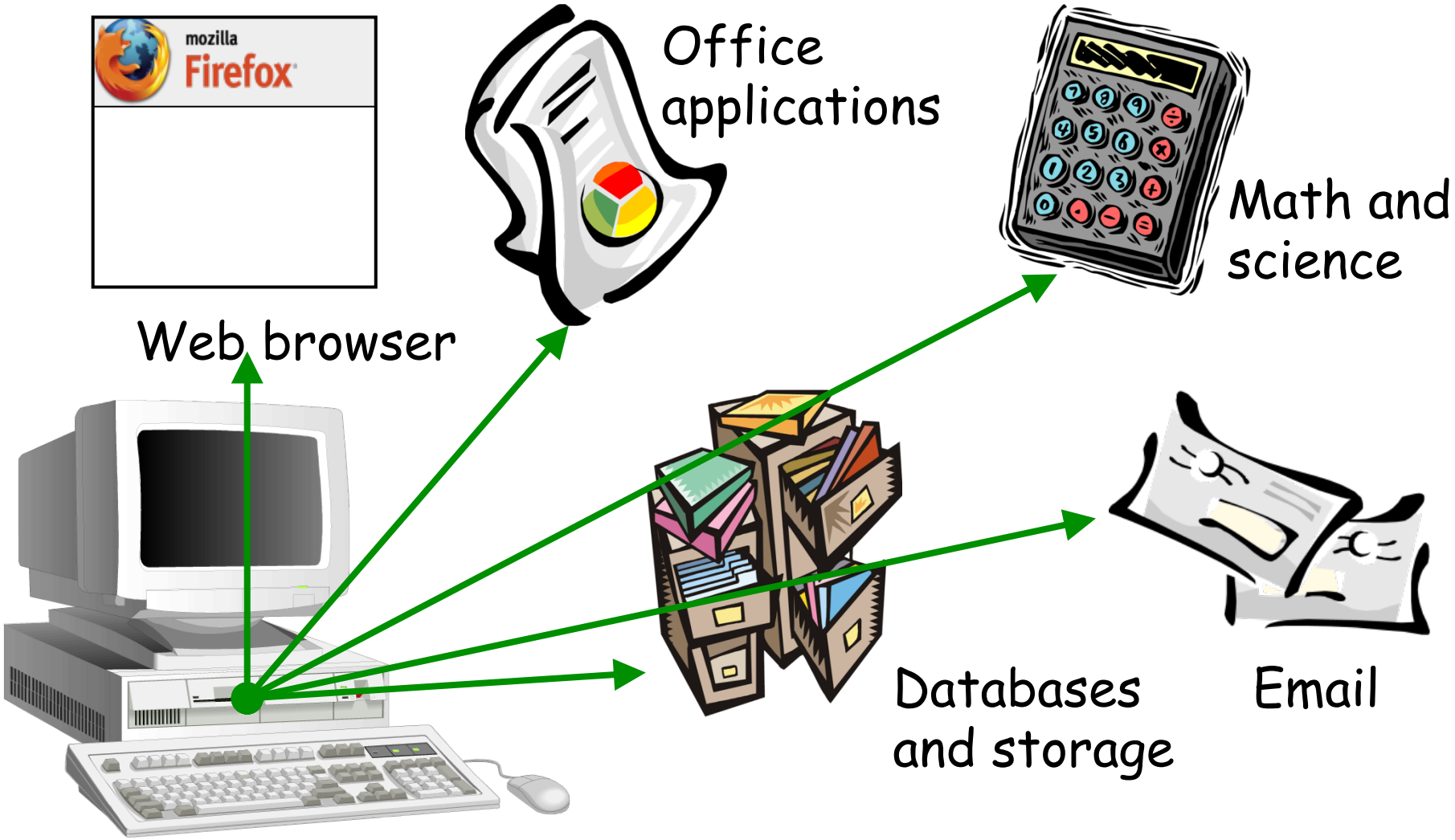
- Too big to fail

Google Outage on 17th Aug 2013

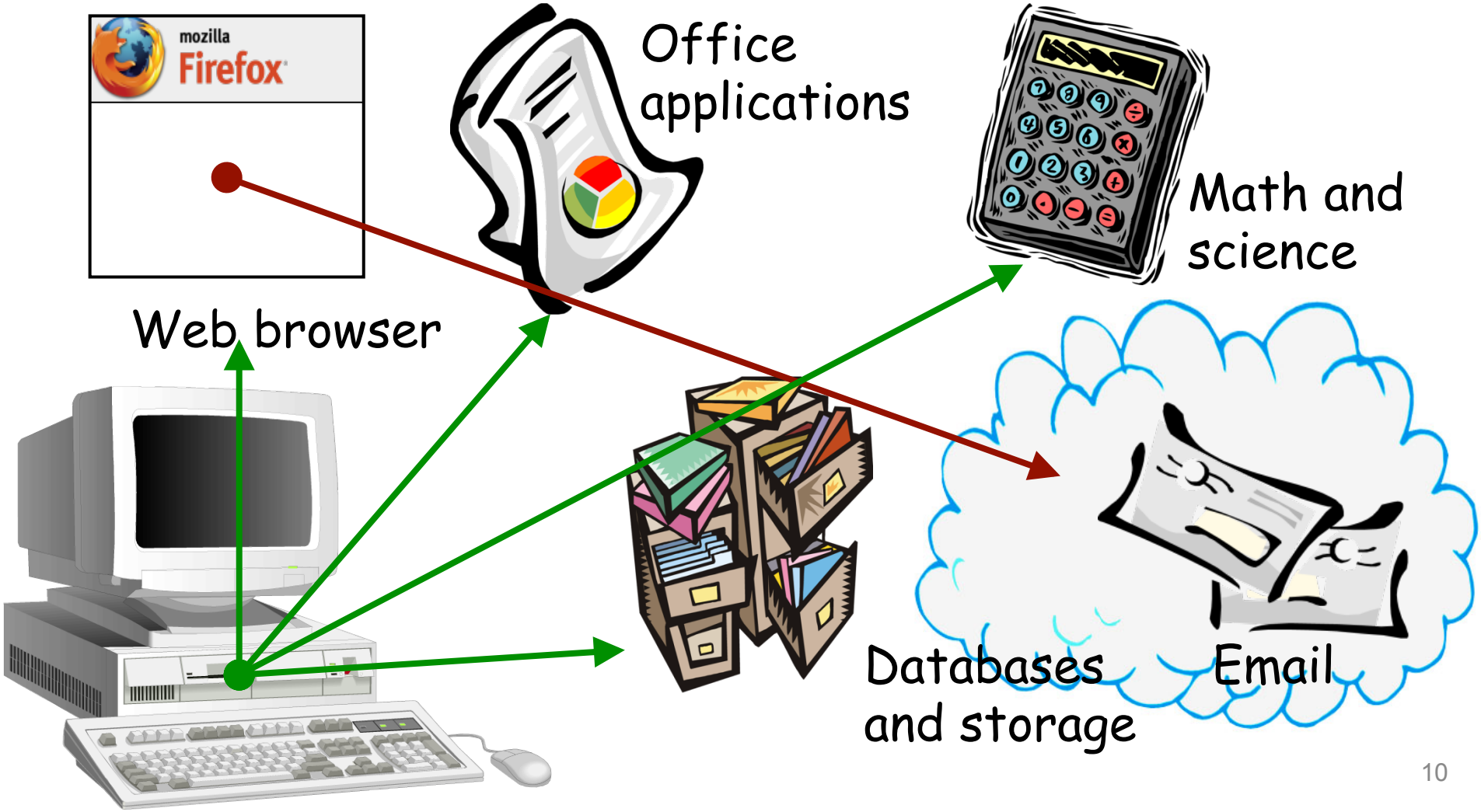
40% drop in Internet traffic

$545,000 revenue loss for 5 min

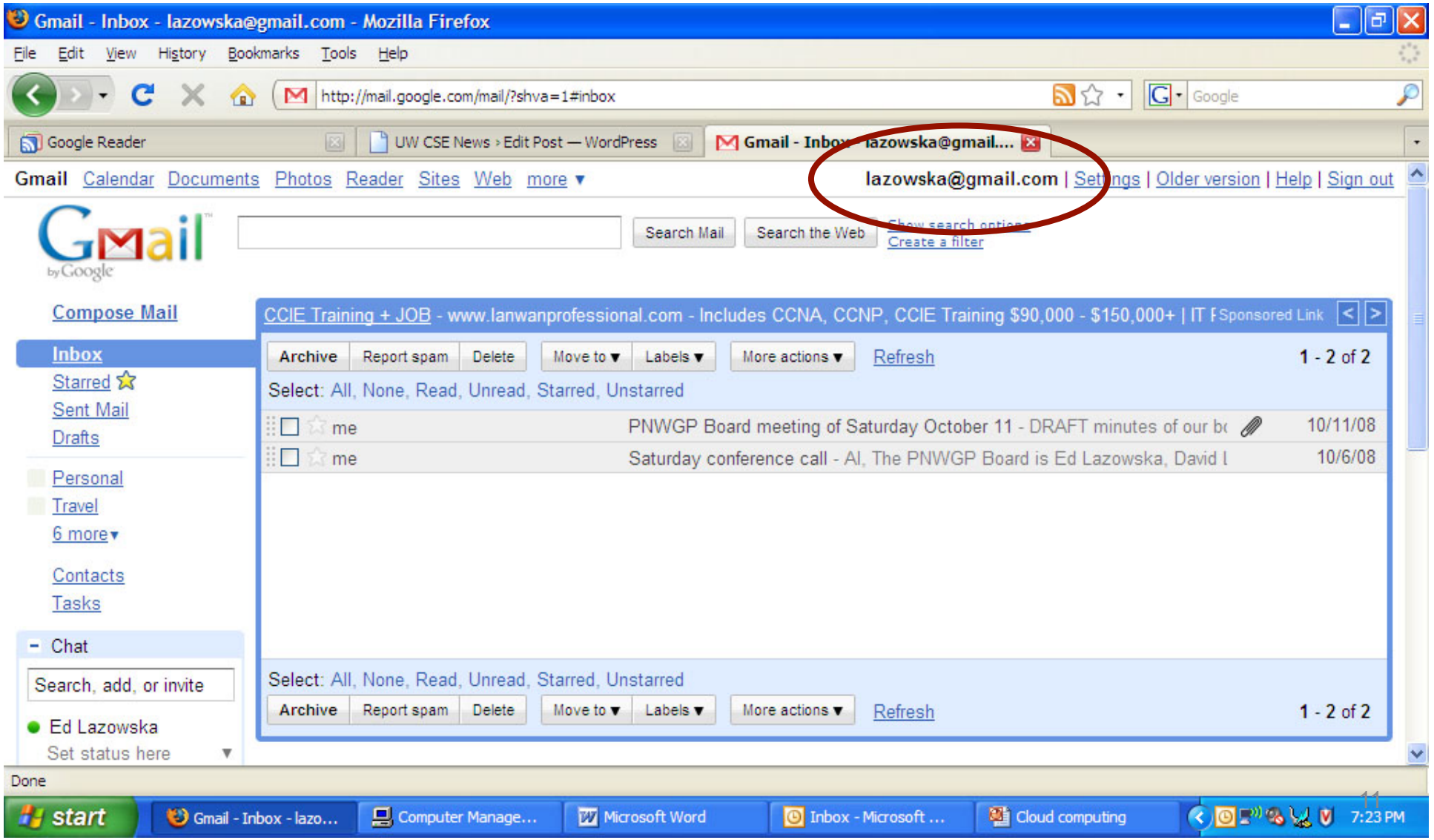
Personal computing



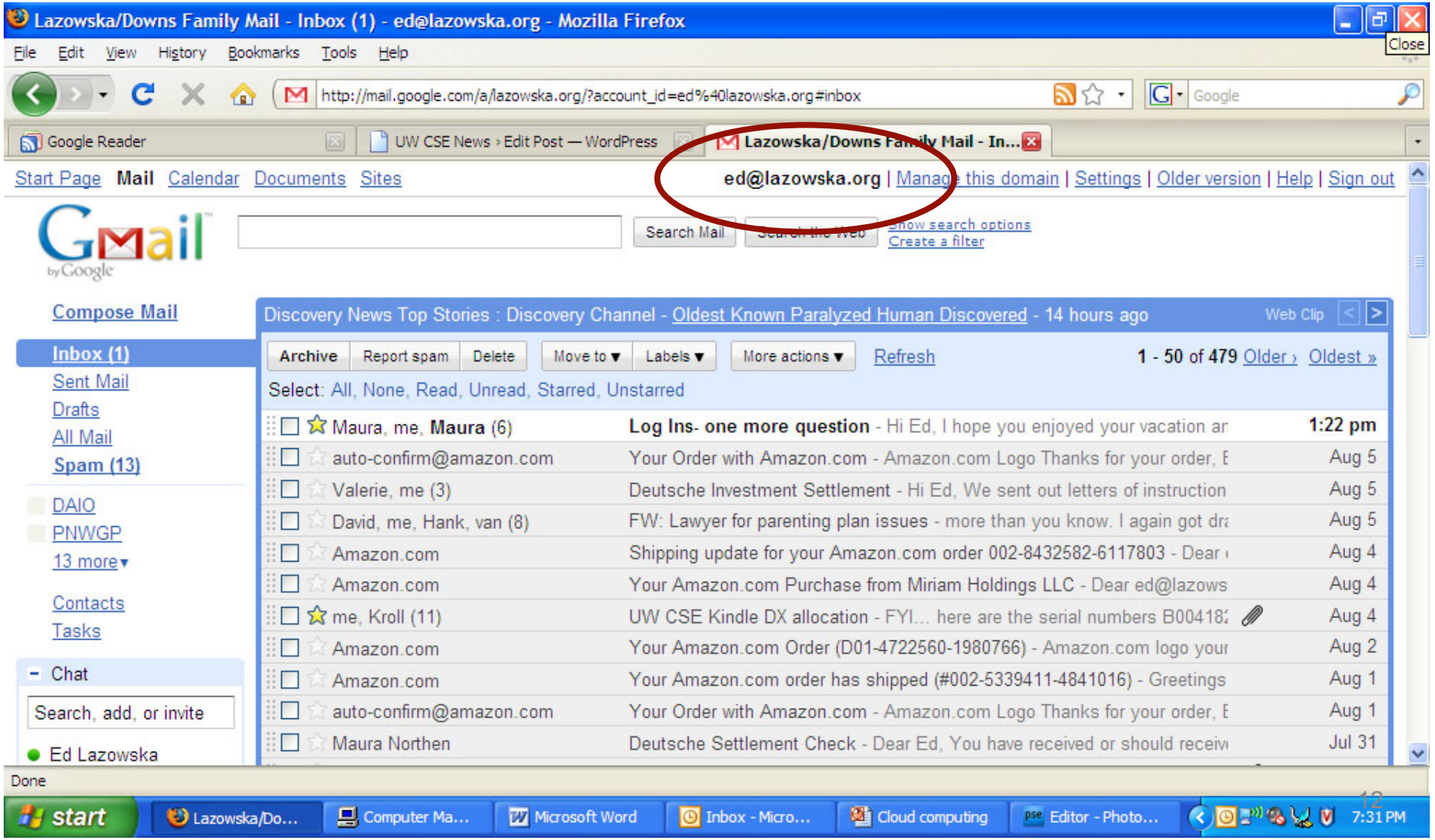
Cloud email accessed through the browser



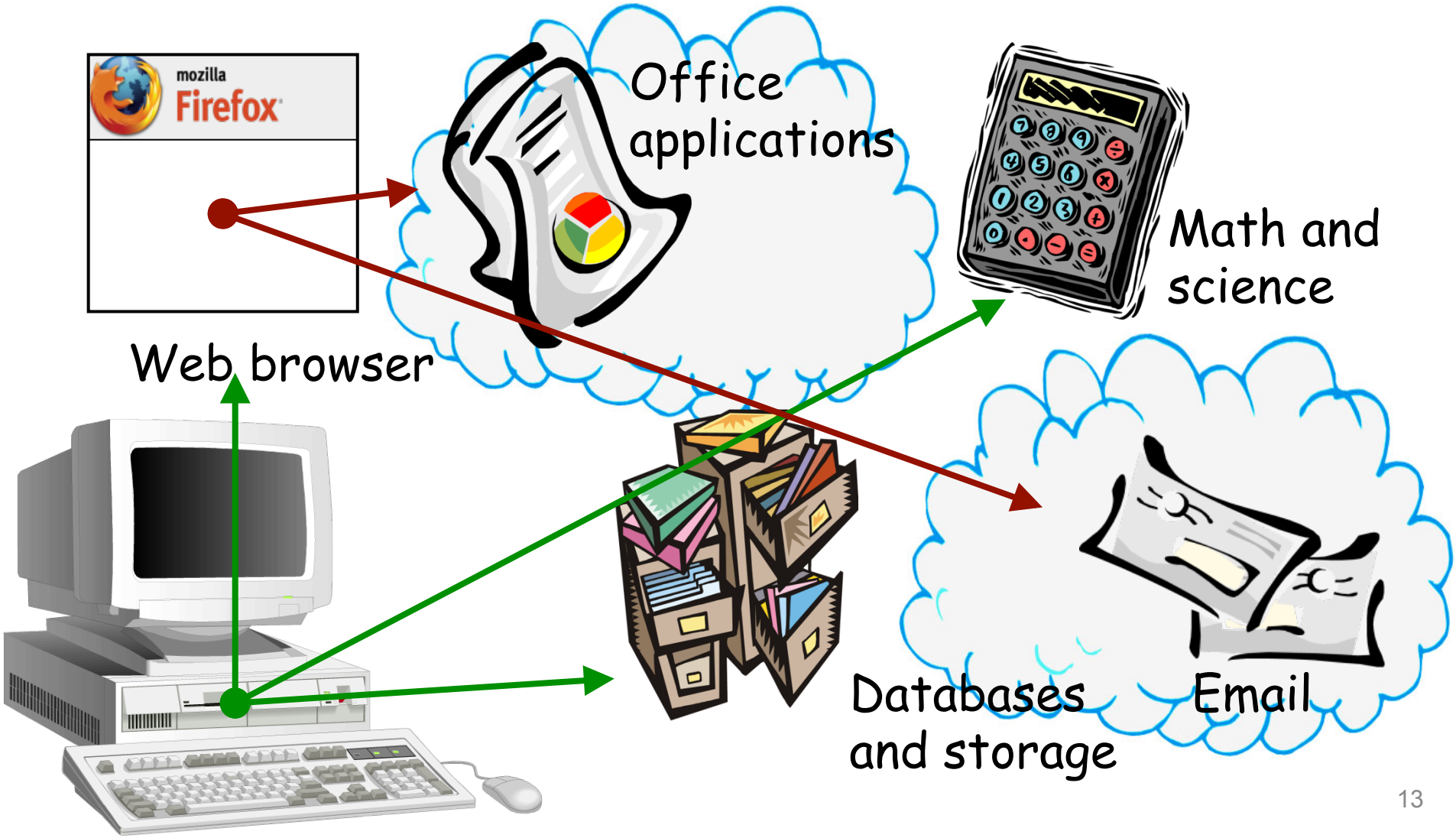
… with the cloud provider's domain name …

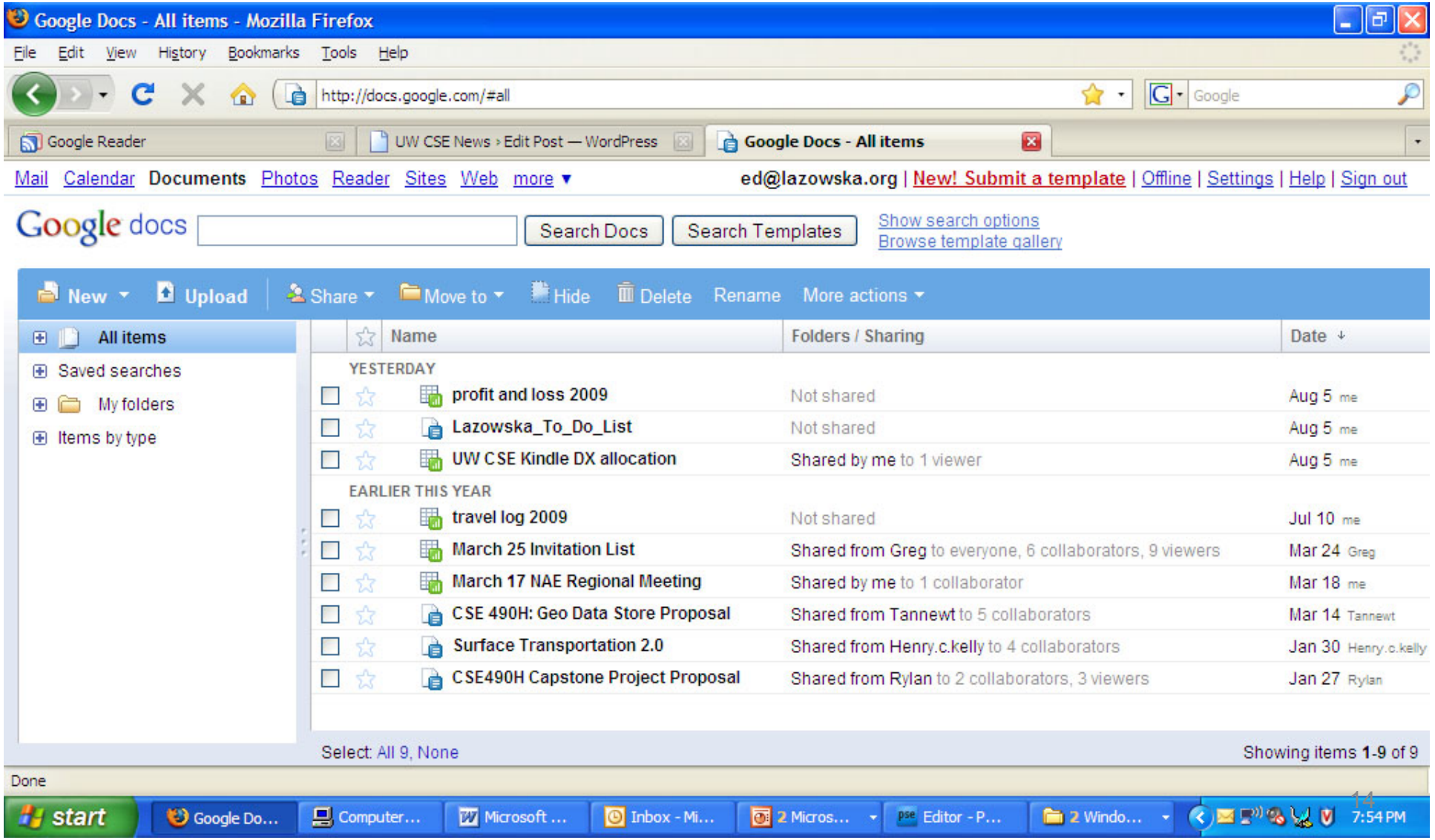


… or with your own

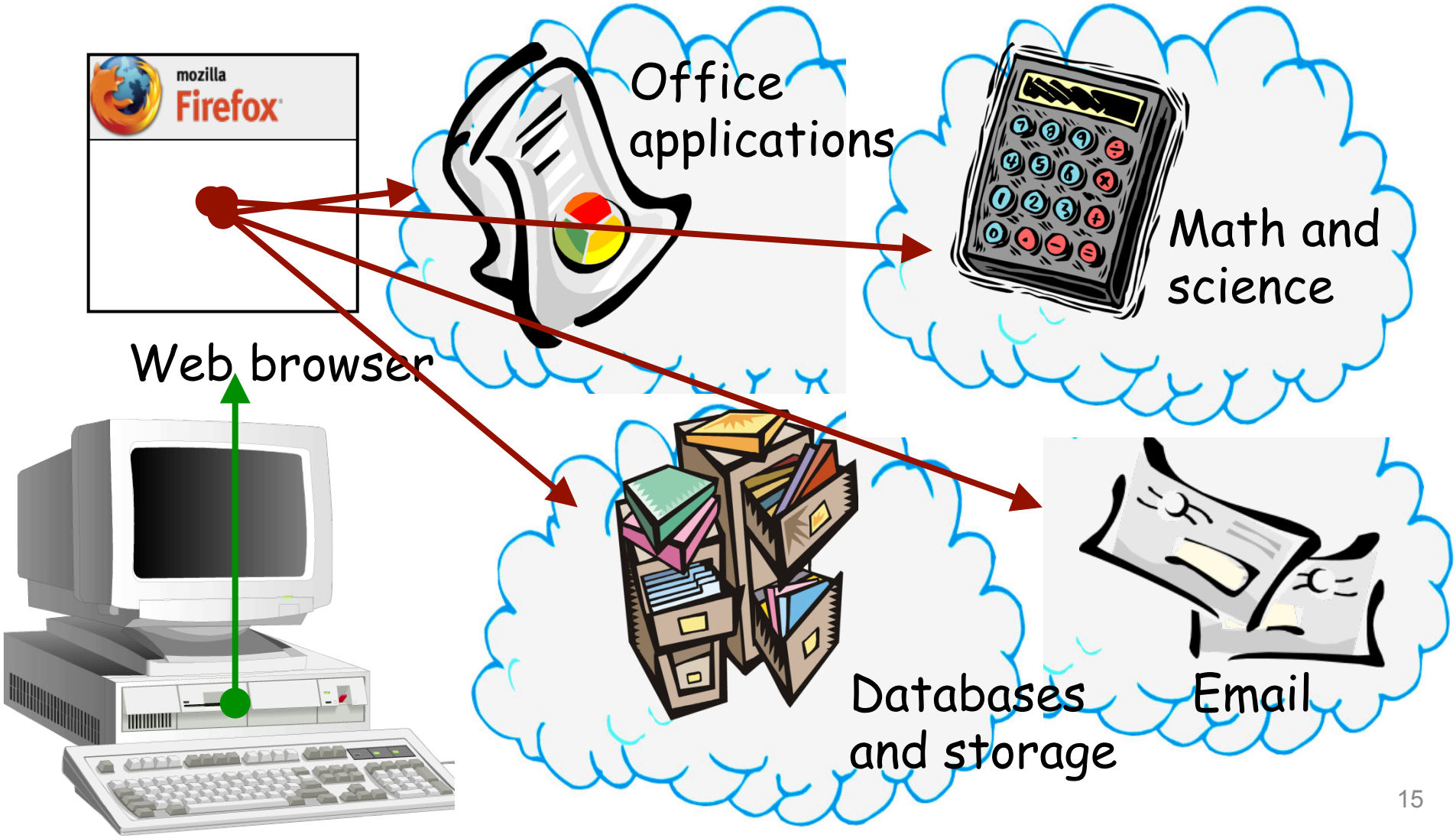


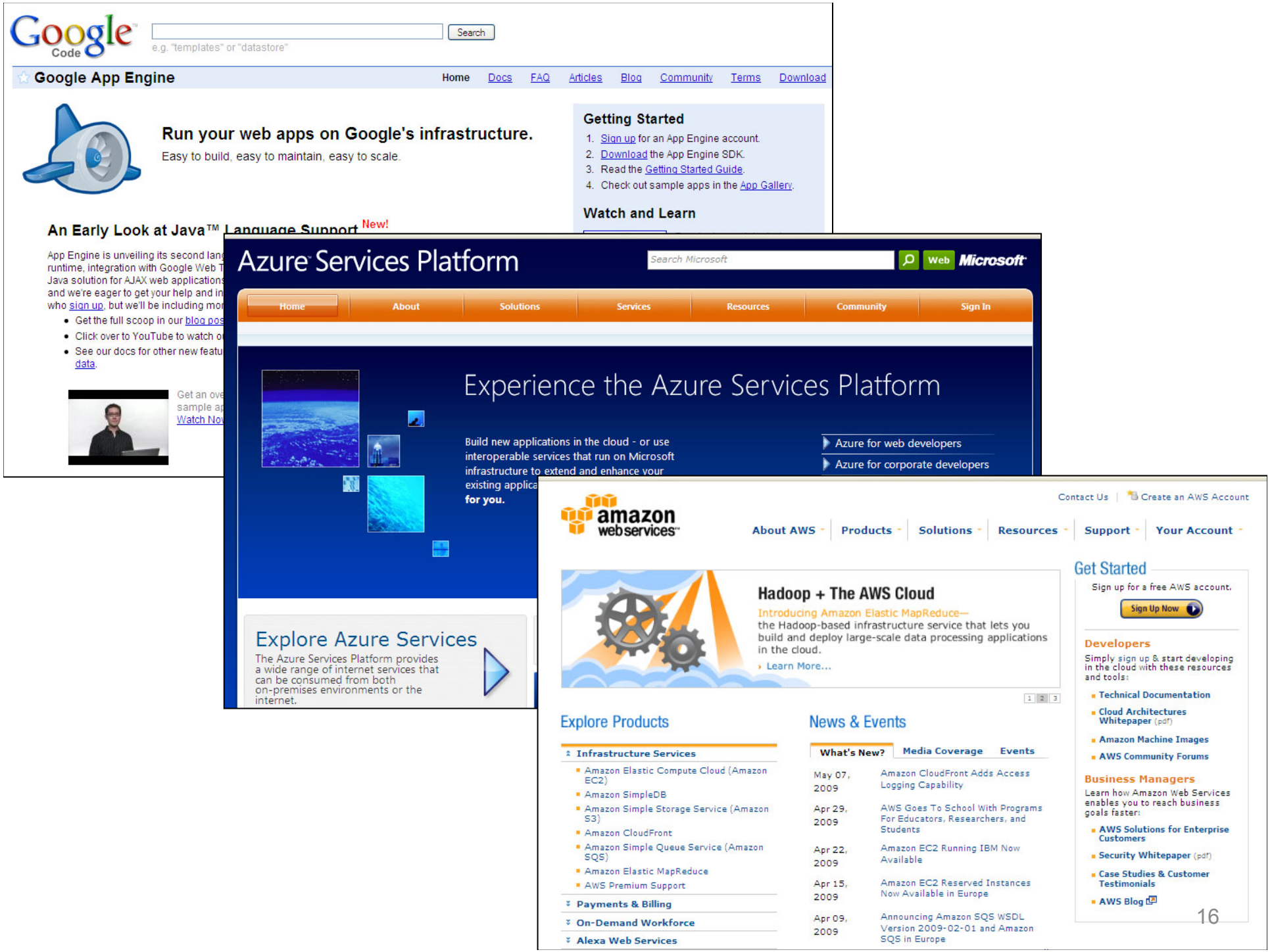
Why not office applications too?




Google Docs - All items - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://docs.google.com/#all
Google
Google Reader
UW CSE News › Edit Post — WordPress
Google Docs - All items
Mail Calendar Documents Photos Reader Sites Web more ▾
ed@lazowska.org | New! Submit a template | Offline | Settings | Help | Sign out
Google docs
Search Docs
Search Templates
Show search options
Browse template gallery
New
Upload
Share
Move to
Hide
Delete
Rename
More actions
All items
Saved searches
My folders
Items by type
Name
Folders / Sharing
Date
YESTERDAY
profit and loss 2009 Not shared Aug 5 me
Lazowska_To_Do_List Not shared Aug 5 me
UW CSE Kindle DX allocation Shared by me to 1 viewer Aug 5 me
EARLIER THIS YEAR
travel log 2009 Not shared Jul 10 me
March 25 Invitation List Shared from Greg to everyone, 6 collaborators, 9 viewers Mar 24 Greg
March 17 NAE Regional Meeting Shared by me to 1 collaborator Mar 18 me
CSE 490H: Geo Data Store Proposal Shared from Tannewt to 5 collaborators Mar 14 Tannewt
Surface Transportation 2.0 Shared from Henry.c.kelly to 4 collaborators Jan 30 Henry.c.kelly
CSE490H Capstone Project Proposal Shared from Rylan to 2 collaborators, 3 viewers Jan 27 Rylan
Select: All 9, None
Showing items 1-9 of 9
Done
start
Google Do...
Computer...
Microsoft ...
Inbox - Mi...
2 Micros...
Editor - P...
2 Windo...
7:54 PM

Why not everything else?




Google Code
Search
e.g. "templates" or "datastore"
Google App Engine
Home Docs FAQ Articles Blog Community Terms Download
Run your web apps on Google's infrastructure.
Easy to build, easy to maintain, easy to scale.
Getting Started
1. Sign up for an App Engine account.
2. Download the App Engine SDK.
3. Read the Getting Started Guide.
4. Check out sample apps in the App Gallery.
Watch and Learn
An Early Look at Java™ Language Support New!
Azure Services Platform
Search Microsoft
Web
Microsoft
Home About Solutions Services Resources Community Sign In
Experience the Azure Services Platform
Build new applications in the cloud - or use interoperable services that run on Microsoft infrastructure to extend and enhance your existing applica
for you.
Azure for web developers
Azure for corporate developers
Explore Azure Services
The Azure Services Platform provides a wide range of internet services that can be consumed from both on-premises environments or the internet.
amazon web services
Contact Us | Create an AWS Account
About AWS Products Solutions Resources Support Your Account
Hadoop + The AWS Cloud
Introducing Amazon Elastic MapReduce—
the Hadoop-based infrastructure service that lets you build and deploy large-scale data processing applications in the cloud.
Learn More...
Get Started
Sign up for a free AWS account.
Sign Up Now
Developers
Simply sign up & start developing in the cloud with these resources and tools:
Technical Documentation
Cloud Architectures Whitepaper (pdf)
Amazon Machine Images
AWS Community Forums
Business Managers
Learn how Amazon Web Services enables you to reach business goals faster:
AWS Solutions for Enterprise Customers
Security Whitepaper (pdf)
Case Studies & Customer Testimonials
AWS Blog
Explore Products
Infrastructure Services
Amazon Elastic Compute Cloud (Amazon EC2)
Amazon SimpleDB
Amazon Simple Storage Service (Amazon S3)
Amazon CloudFront
Amazon Simple Queue Service (Amazon SQS)
Amazon Elastic MapReduce
AWS Premium Support
Payments & Billing
On-Demand Workforce
Alexa Web Services
News & Events
What's New? Media Coverage Events
May 07, 2009 Amazon CloudFront Adds Access Logging Capability
Apr 29, 2009 AWS Goes To School With Programs For Educators, Researchers, and Students
Apr 22, 2009 Amazon EC2 Running IBM Now Available
Apr 15, 2009 Amazon EC2 Reserved Instances Now Available in Europe
Apr 09, 2009 Announcing Amazon SQS WSDL Version 2009-02-01 and Amazon SQS in Europe

Consider …

- Sharing is easy
- Someone else does backup
- Someone else handles software updates
- There's 7x24x365 operations support, auxiliary power, redundant network connections, geographical diversity
- Scalability – both up and down – is instantaneous
- Many fewer demands on the local operating system and machine

Amazon Elastic Compute Cloud (EC2)

- $0.68 per hour for
  - 4 cores of 2.5 GHz 64-bit 2007 Xeon or Opteron
  - 15 GB memory
  - 1.69 TB scratch storage
- Need it 24x7 for a year?
  - $3900

- $0.085 per hour for
  - 1 core of 1.2 GHz 32-bit Intel or AMD
  - 1.7 GB memory
  - 160 GB scratch storage
- Need it 24x7 for a year?
  - $490

- This includes
  - Purchase + replacement
  - Housing
  - Power
  - Operation
  - Reliability
  - Security
  - Instantaneous expansion and contraction
- 1000 processors for 1 day costs the same as 1 processor for 1000 days!

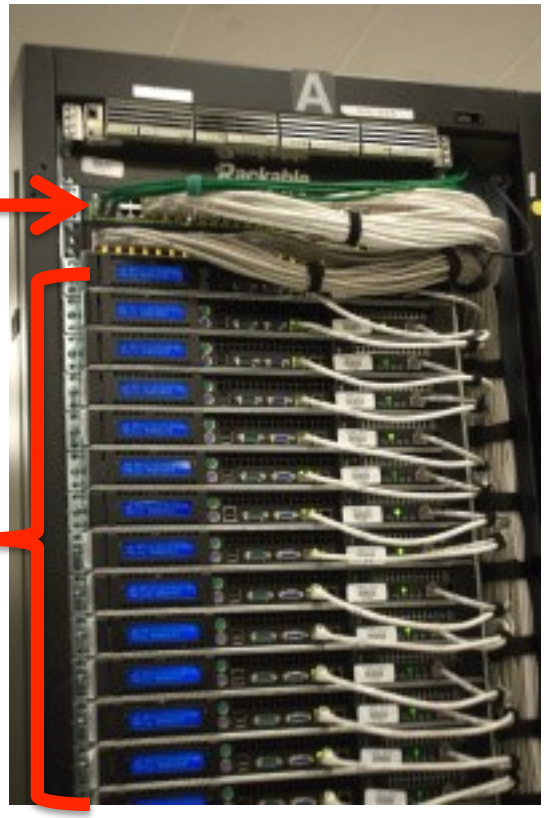
The nuts and bolts of data center

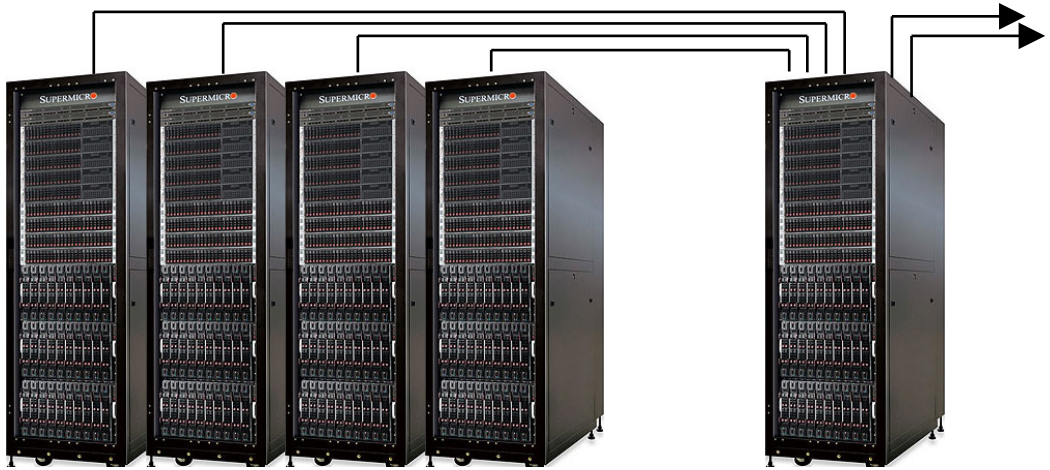
- Networks
- Servers
- Storages
- Software
- Power systems
- Cooling systems
- …

How should we design a data center network?

Interconnecting 10,000s of machines

- Top-of-Rack architecture
  - Rack of commodity servers
  - Top-of-Rack Switch
- Aggregation of ToRs

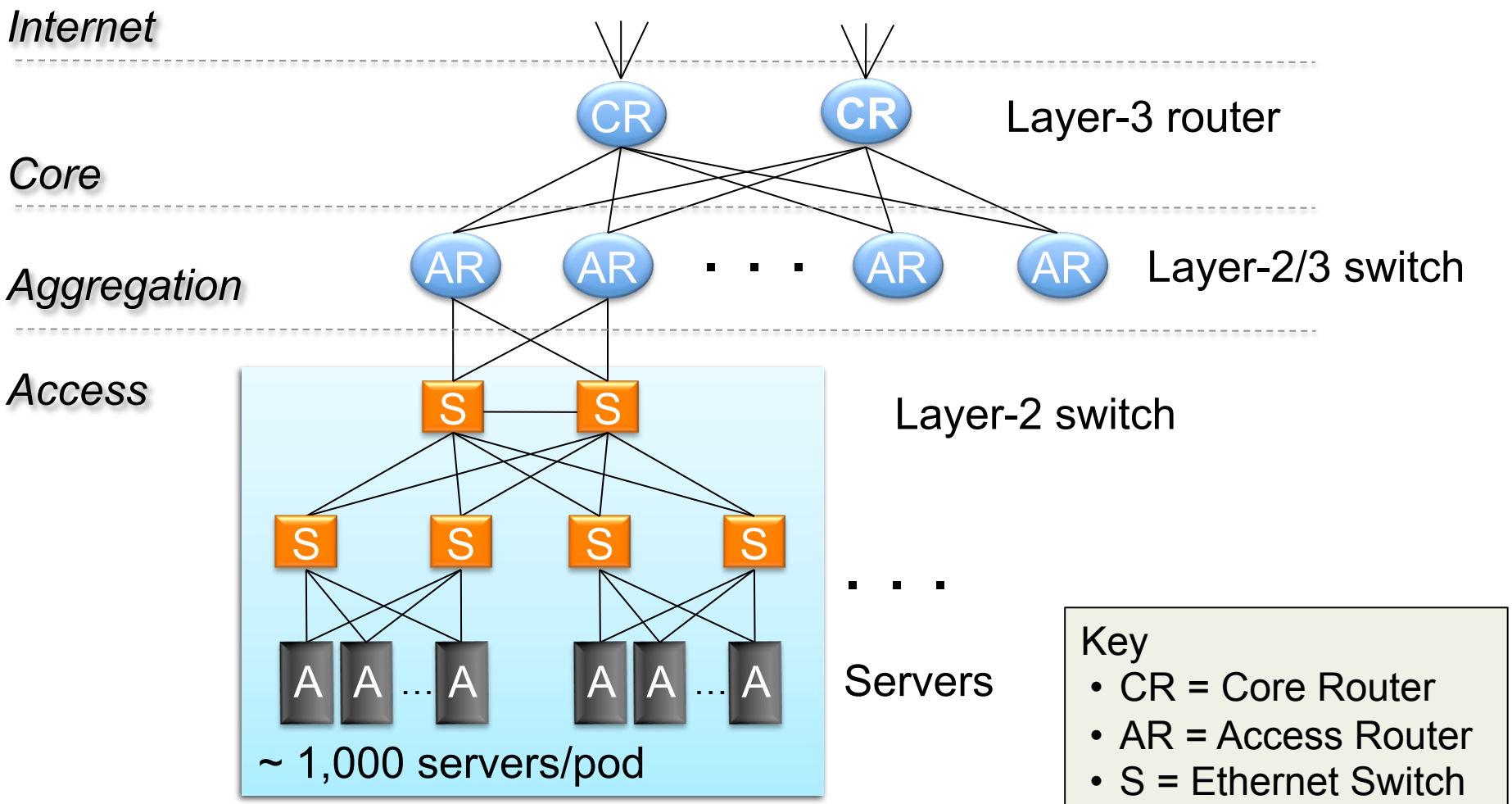




Format borrowed from Jen Rexford's COS 561 slides

Overall picture

Common data center network topology



Source: Jen Rexford's COS 561 slides

Characteristics of data center (networks)

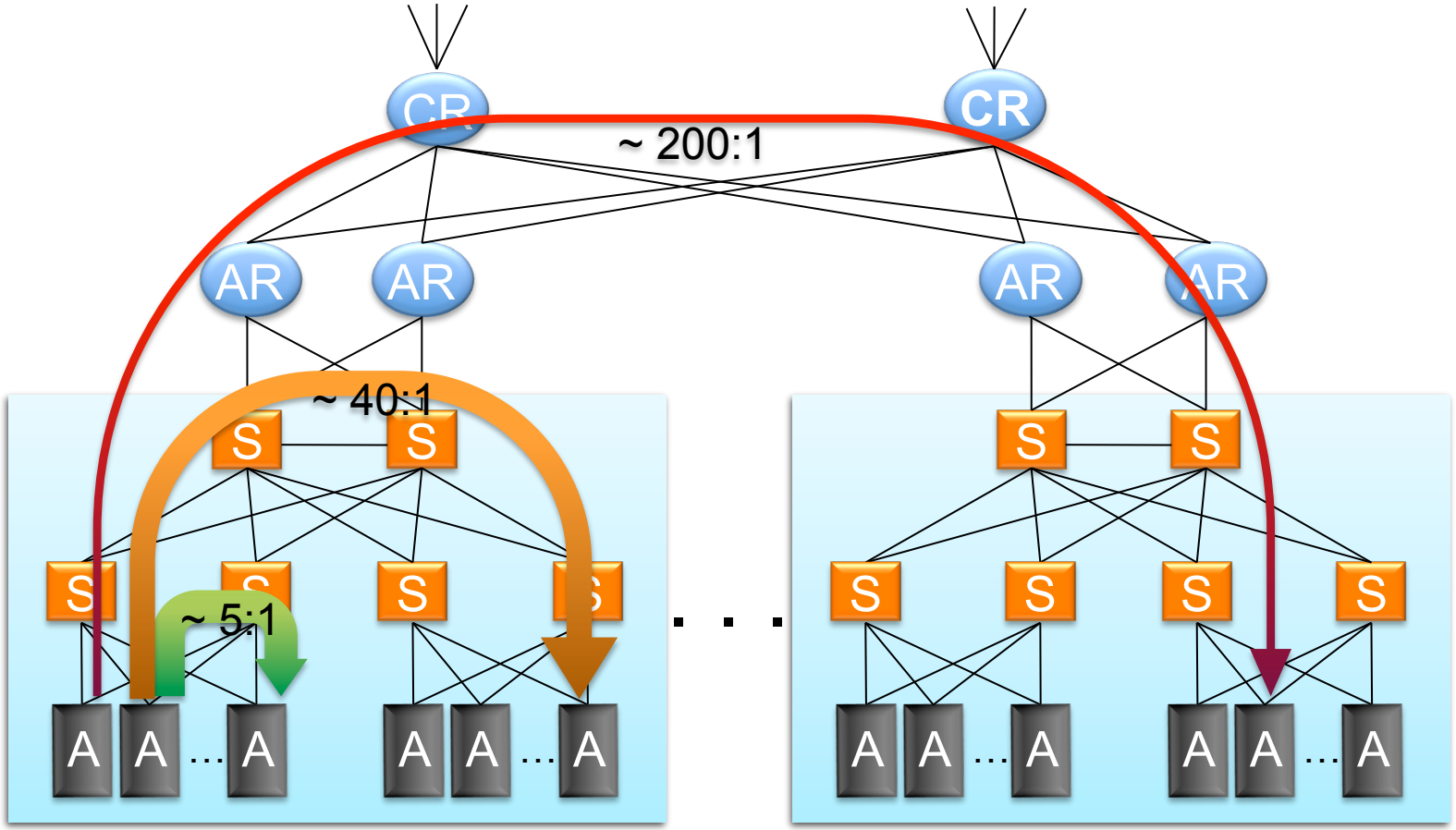
- Single ownership
  - Allows full control over an entire system
  - Less concern about standards and interoperability
- Less heterogeneous environments
  - Similar servers, storage, topology, software stack
- Multiple end-to-end paths
  - E.g. Clos topology, multi-rooted tree topology
- Low end-to-end delays when no congestion
  - Servers in a geographically small region
  - DC:100s of μs vs. Internet: 10s of ms to 100s of ms

Applications in data centers

- Web services
- Web search
  - Google Search, Microsoft Bing
- High performance computing (HPC)
- Big data analytics
  - Hadoop, MapReduce, Twitter Storm, etc.
- Machine learning
- Cloud applications
  - DropBox, Google Drive, etc.

Applications compete for data center resources

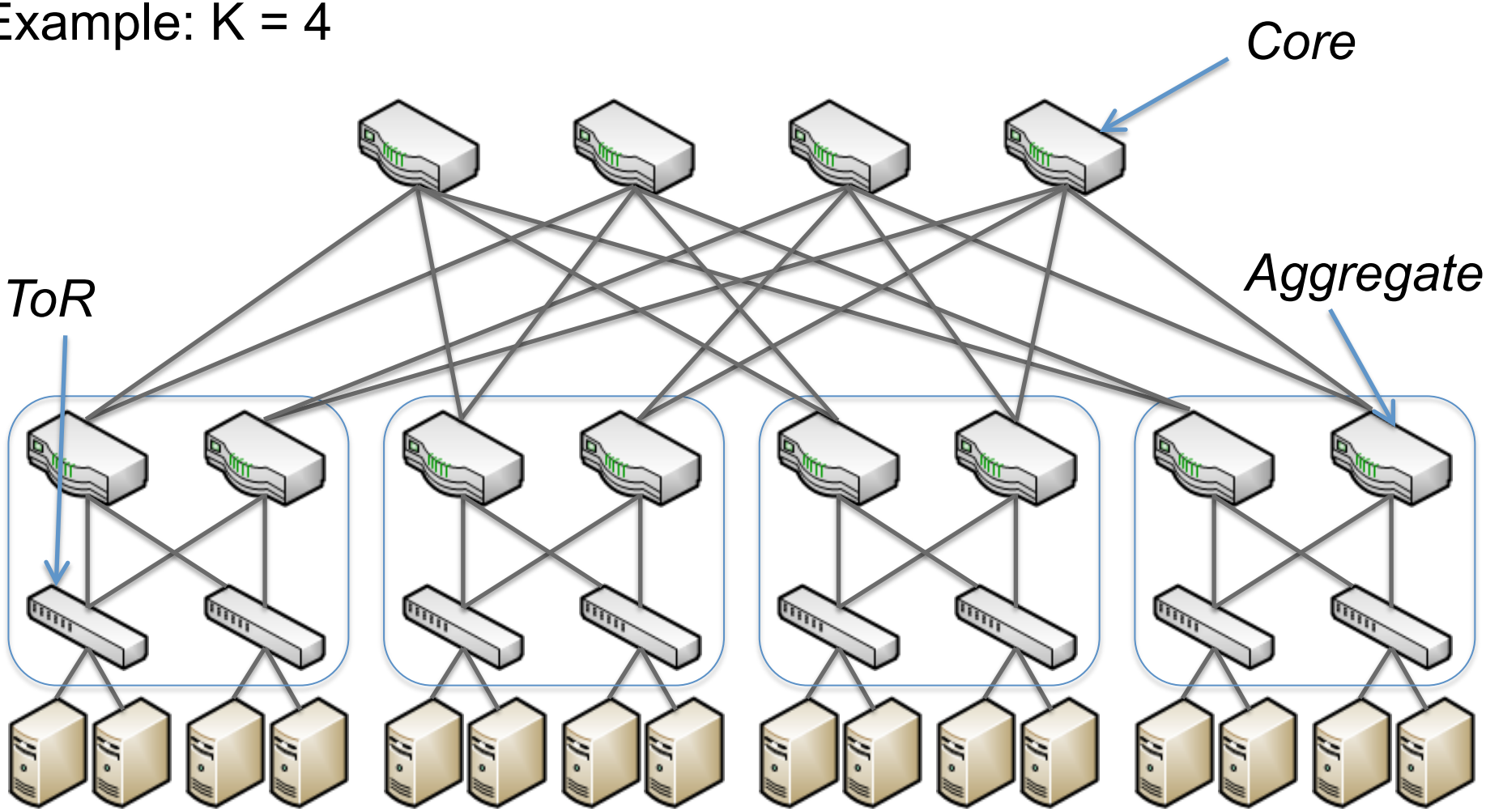
Capacity mismatch



Source: Jen Rexford's COS 561 slides

Example: Fat-tree topology

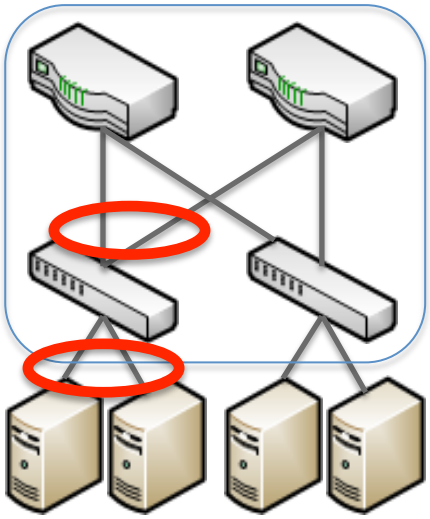
Example: K = 4



Fat-tree topology

A set of K/2 ports used for upper level connectivity, another set for lower level connectivity

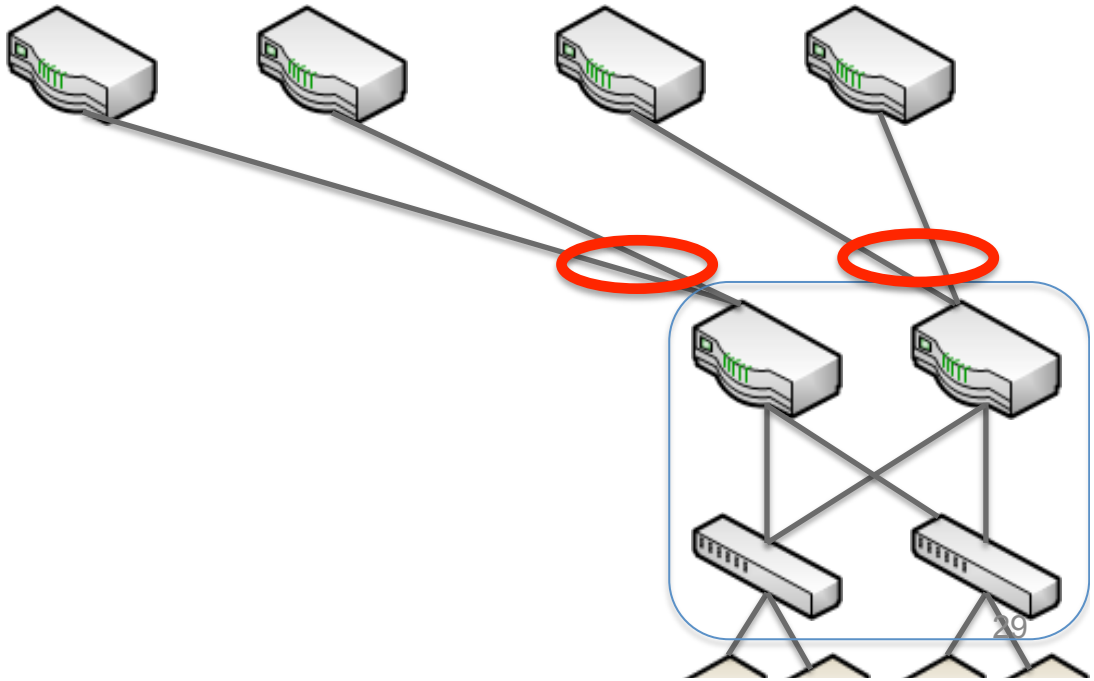
K-port switches/routers



Pod

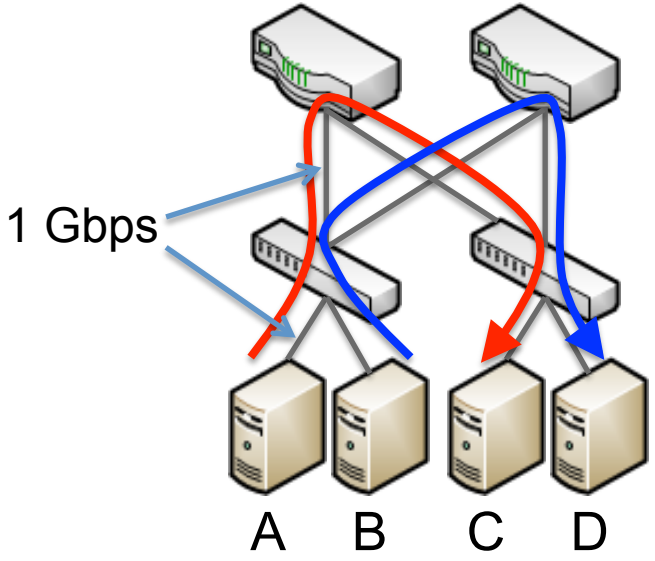
Example: K = 4

Top Level: core routers

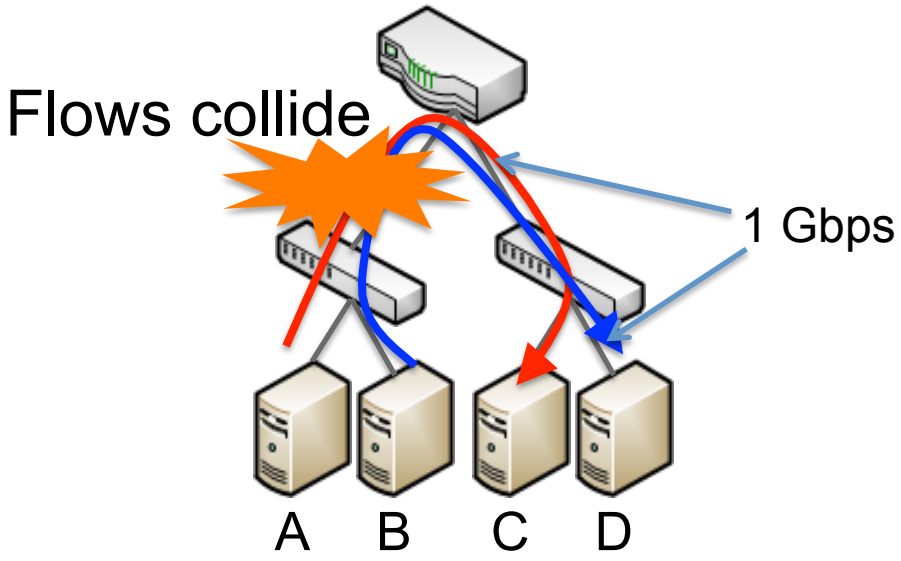


Benefit of multiple equal-cost paths

Each link = 1 Gbps; A talks to C; B talks to D



Throughput = 1 Gbps



Throughput = 500 Mbps

Deciding an end-to-end path of flows is an important scheduling task to fully utilize multiple paths

Exploiting multiple equal-cost paths

- Many approaches
  - ECMP (Equal-Cost Multi-Path) forwarding
  - Monsoon [PRESTO'08]
  - VL2 [SIGCOMM'09]
  - Hedera [NSDI'10]
  - Mahout [Infocom'11]
  - MPTCP (Multpath TCP) [SIGCOMM'11]
  - Packet Spraying [Infocom'13]
  - …

Data centers are cool!

Google data center, Lenoir, North Carolina, US