雲端經驗: 從Google到台達

Cloud Experience: from Google to Delta

台達電子 翟本喬
Ben Jai, Delta Electronics, Inc.

COSCUP 2010　　2010/8/15

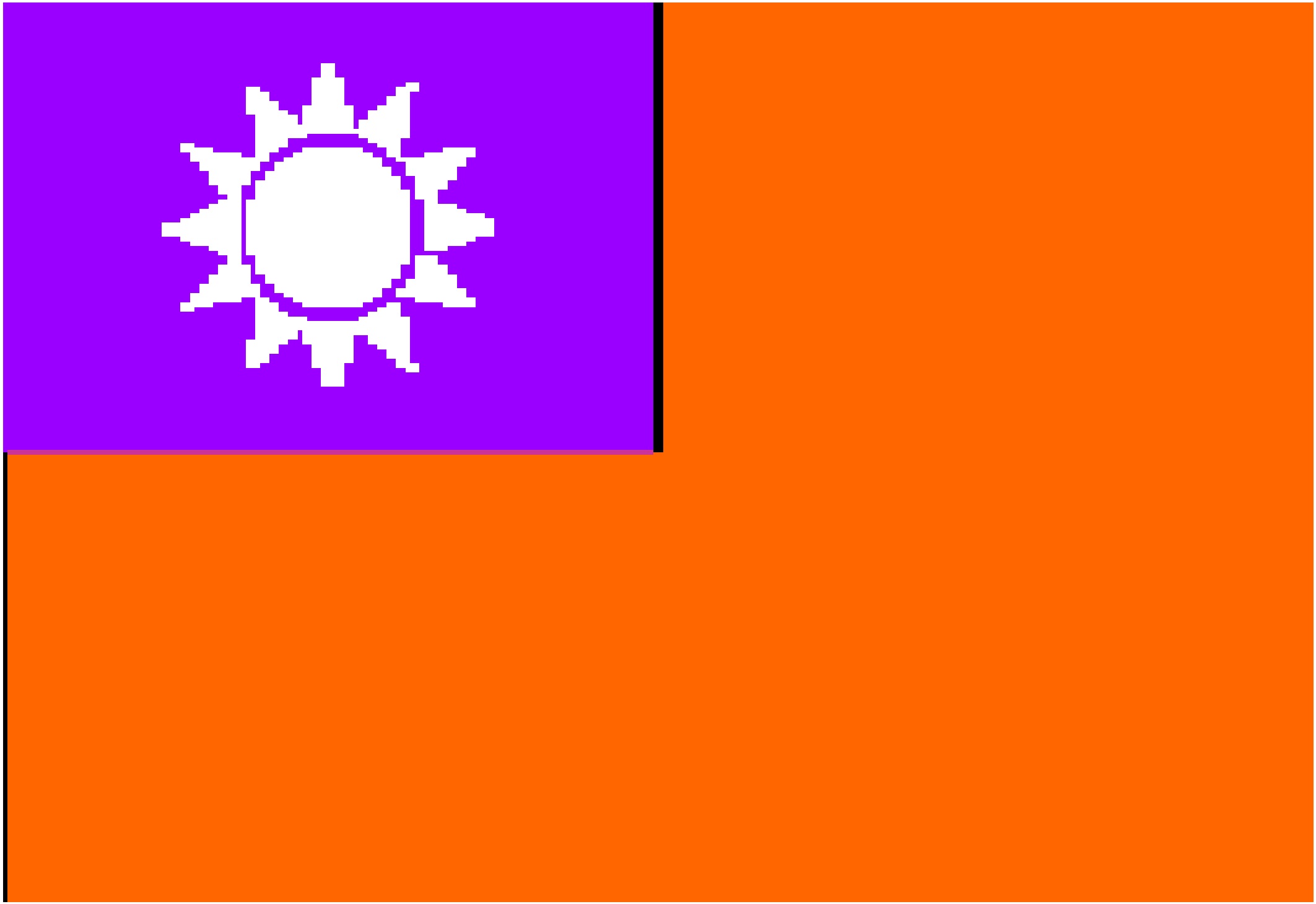
Personal Introduction

- 2003-2010: Google
  - Data Center Research, Manager, Hardware Architect
  - Single-V PSU, server with built-in battery, container data center
- 1999-2003: Bell Labs Research
  - Network management, middleware
- 1999: Ph.D. Computer Science, NYU
- 1984-1999: Computer consultant
  - Traffic control system, Taipei (and Tainan and other cities)
  - Taipower 3rd Nuclear Power Plant Safety Control System
- 1980: 創立建中電研社
- Contribution to Open Source
  - 1990: NYUMinix: Minix on DOS
  - 1982: …

```
 10   HGR2
 20   HCOLOR= 5
 30   FOR I = 0 TO 191
 40   HPLOT 0,I TO 279,I
 50   NEXT
 60   HCOLOR= 2
 70   FOR I = 0 TO 95
 80   HPLOT 0,I TO 139,I
 90   NEXT
 100   HCOLOR= 3
 110 M = 6.283 / 12
 120   FOR I = 0 TO 6.283 STEP .03

 130   HPLOT 70,48 TO 70 +  COS (I
       ) * 17,48 +  SIN (I) * 17
 140   HPLOT 70 +  COS (I) * 21,48
       +  SIN (I) * 21 TO 70 +  COS
      ((  INT (I / M + 0.5)) * M) *
      35,48 +  SIN ((  INT (I / M +
      0.5)) * M) * 35
 150   NEXT

]
```



Cloud Computing Vendor Lists

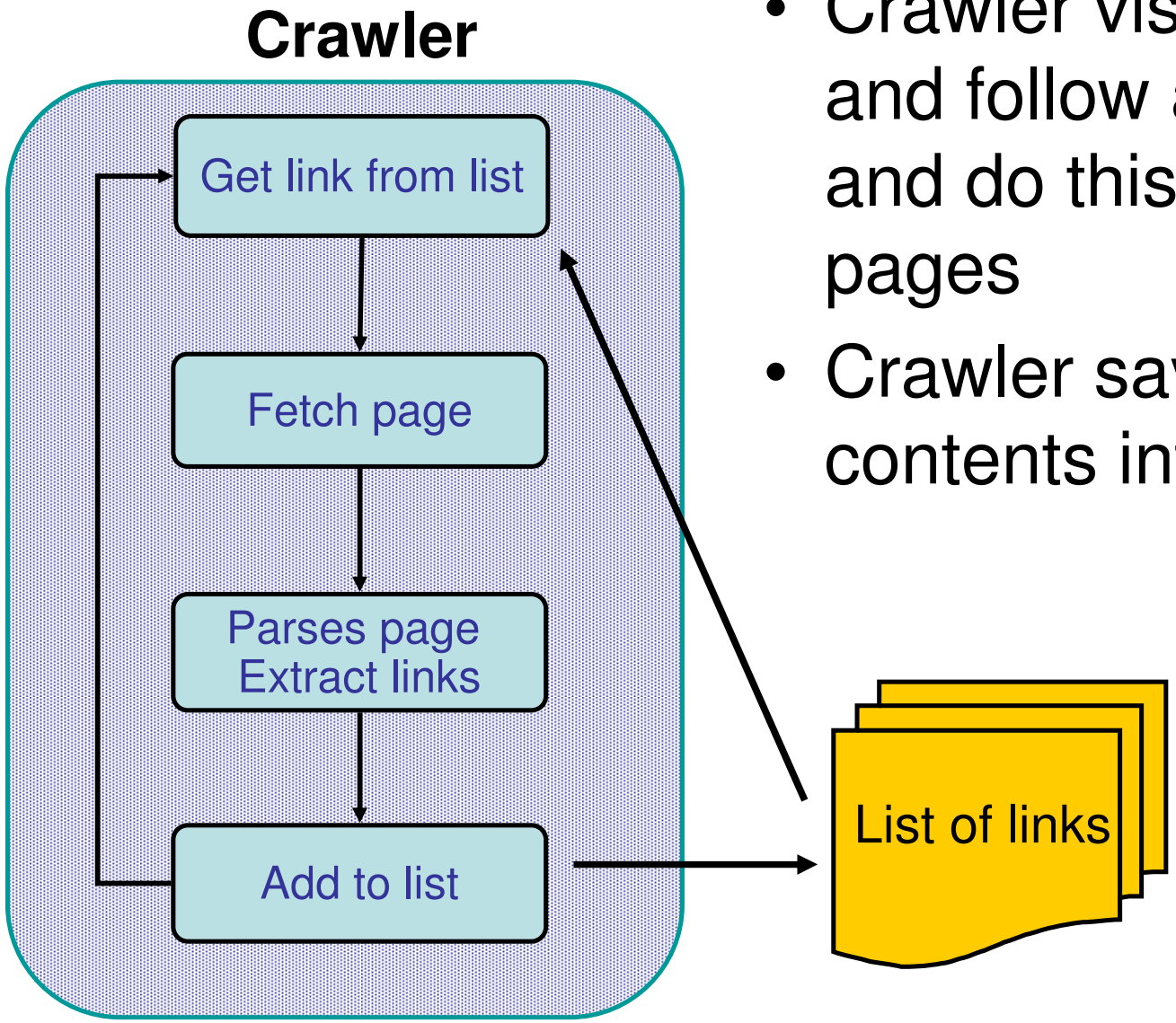
- Google, Salesforce.com
- Google, Microsoft, VMWare
- Google, Amazon
- Cisco, HP, IBM, Dell

- What exactly did Google do in Cloud Computing?

Search Engine

- Crawler: collects web pages and images
  - Billions of pages
  - 10-100 KB/page → exabytes (10^18)
  - and emails, blogs, tweets, documents, databases, books, etc.
- Indexer: transforms the data into index
  - Ready for lookup
- Search server: responds to user queries
  - Lookup index and present results
  - Need to figure out what users like to see
- Do all these in large quantities in a reliable way

Crawling the Web



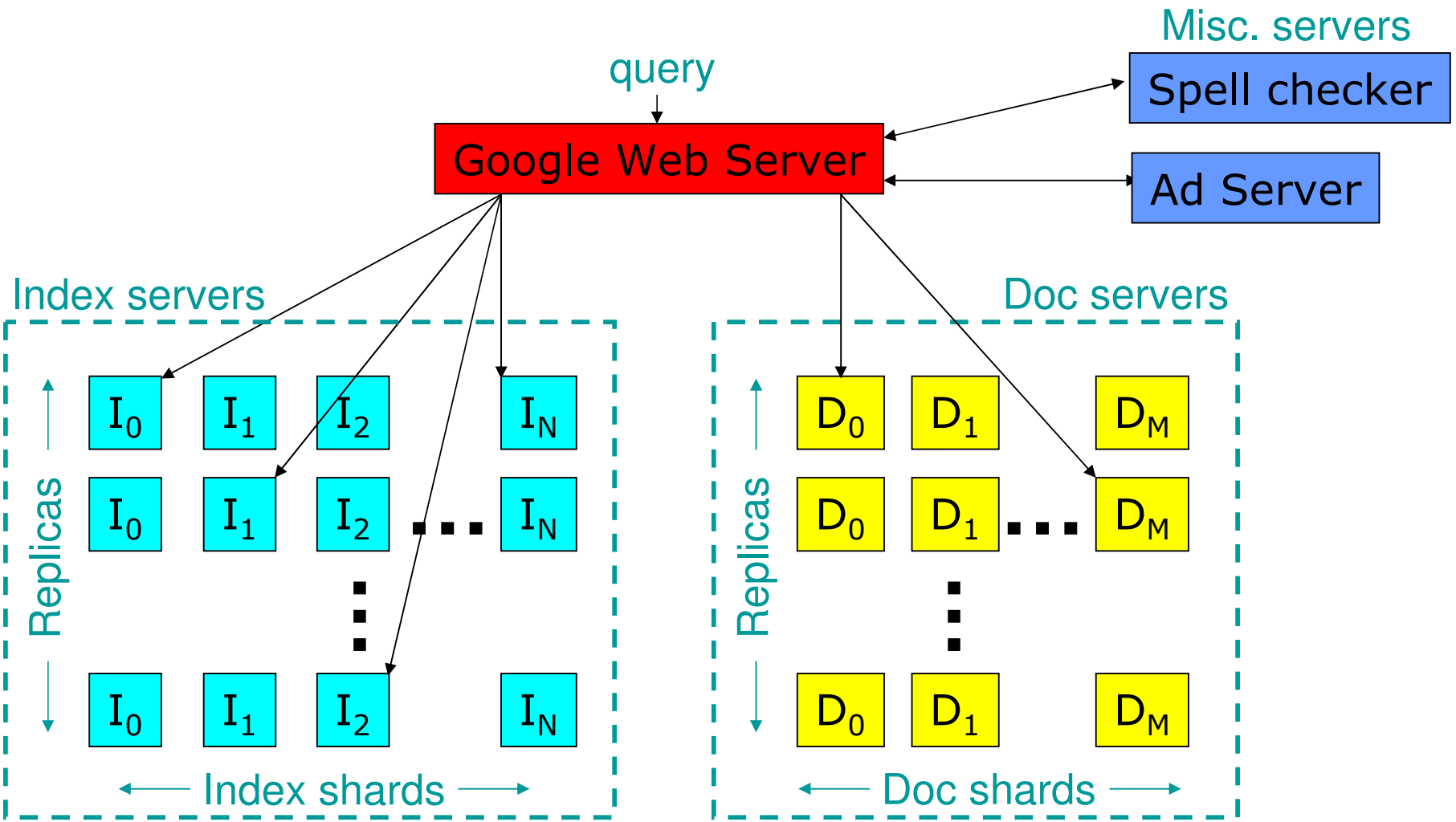
- Crawler visits a web page and follow all the links, and do this to billions of pages
- Crawler saves the page contents into files

Building the Index

- Just like index in the back of books
  - Documents are assigned ID numbers
  - Words are mapped to positions
  - Result is a 3D list

I	doc 1:	word 3, 5, 17, 26;
	doc 3:	word 103, 192;
	doc 4:	word 3, 5, 9, 17;
	doc 55112:	word 6, 335;
you	doc 7:	word 1;
	doc 1293:	word 19, 106, 116, 137;
	doc 55112:	word 8;
love	doc 55112:	word 7, 113;

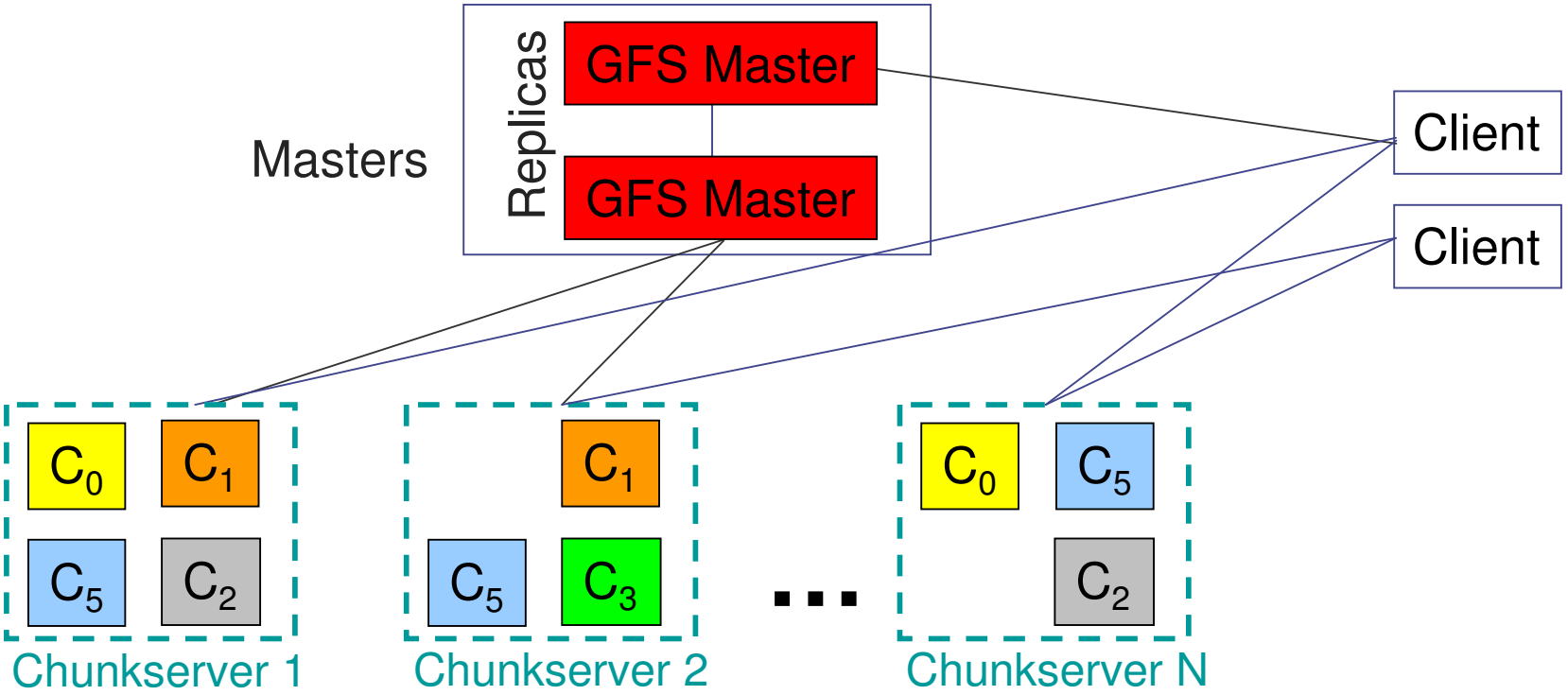
Serving Queries



Reliable Building Blocks

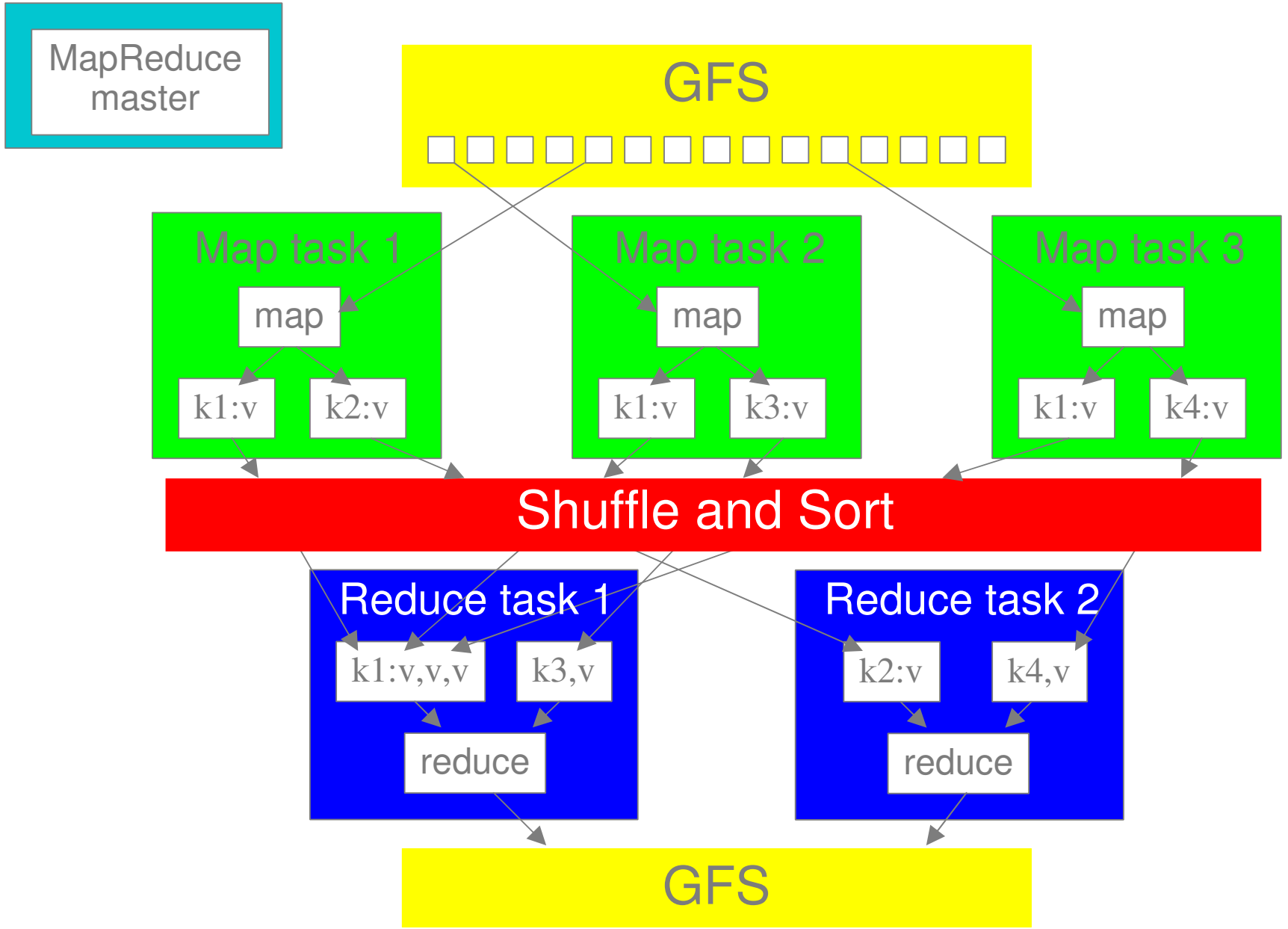
- Need to store data reliably
- Need to run jobs on cluster of machines
- Need to make it easy to apply lots of computational resources to problems
- Need to organize data is easy-to-use form
- Google solutions:
  - Google File System: distributed fault-tolerant storage
  - Global Work Queue: distributed job scheduling
  - MapReduce: simplified large-scale data processing
  - BigTable: distributed structured data management

Google File System (GFS)



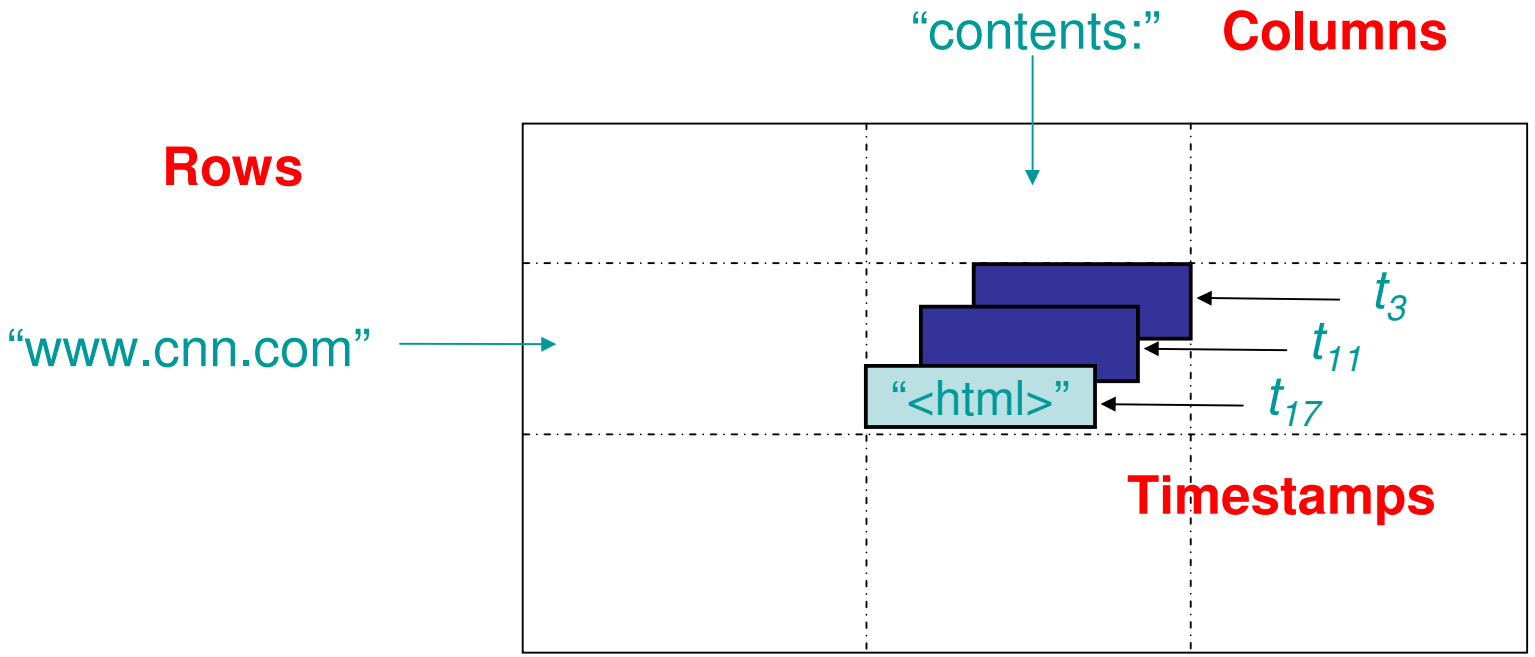
- Files broken into chunks (typically 64 MB)
- Chunks triplicated across three machines for safety
- Master manages metadata
- Data transfers happen directly between clients/chunkservers

MapReduce: Massive Batch Processing



BigTable Database

- Distributed multi-dimensional sparse map
  - *(row, column, timestamp) → cell contents*
- In-memory serving, backed up by GFS



Cost-performance Ratio

High-end SMP server

- HP Integrity Superdome
  - 64 1.5GHz Itanium2s
  - 128GB DDR DRAM
- Performance per CPU
  - 13K tpmC
- Price/CPU: ~$42,000

PC-class DP server

- HP ProLiant DL-360
  - 2 3.0GHz Pentium4 Xeons
  - 4GB DDR DRAM
- Performance per CPU
  - 26K tpmC
- Price/CPU: ~$2,500

- 33x advantage for the PC-class server!
- Get a ~1000 CPU PC-class cluster for about the same price of one 64-CPU high-end server… and possibly 30x the performance!

Reliability Factors

- 3-year MTBF → 1 in 1000 fails each day
- Large applications need 1000's of machines

→Many machines fail every day

- Have to deal with failures in software

- Replication and redundancy necessary for capacity anyway

- Reliable hardware makes software engineers lazy
- Fault-tolerant software makes cheap hardware useful

Power Cost vs. Hardware Cost

- What's in the PUE? (Total power / IT power)
  - Computer power: 400W * 1000 = 400KW
  - UPS: 10% loss → 440KW
  - Delivery: 2% loss → 450KW
  - Air conditioning: + 450KW
  - Lighting, office, camera, etc.: 60KW
  - Total = 960KW
  - PUE = 960KW / 400KW = **2.4**
  - Could go up to 5.0!
- 5-year power cost
  - 960KW * 8760 hr * 5 yr * $0.15/kwh = $6.3M
  - 1000 servers = $2.5M ~ $6M

Power Usage Problems

- CPU companies emphasized performance, not power efficiency
- Equipment makers don’t pay for electricity

- Data center standards (e.g. TIA-942) were set by telecom industry, not IT industry
- IT equipment evolves too quickly

- Lack of end-to-end understanding (IT, power, cooling, facility)

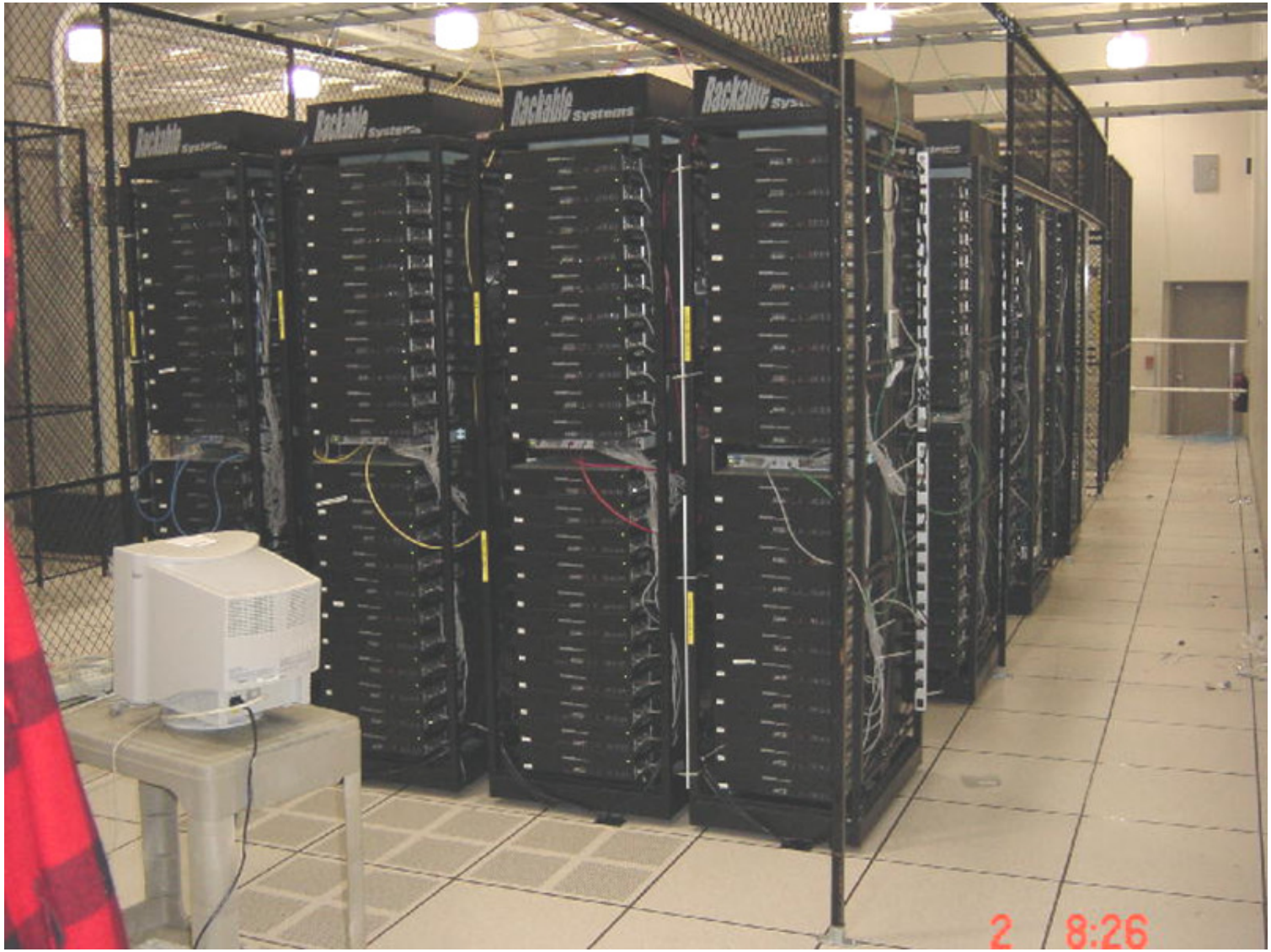
google.stanford.edu (1997)

google.com (1999)

Google Data Center (2000)

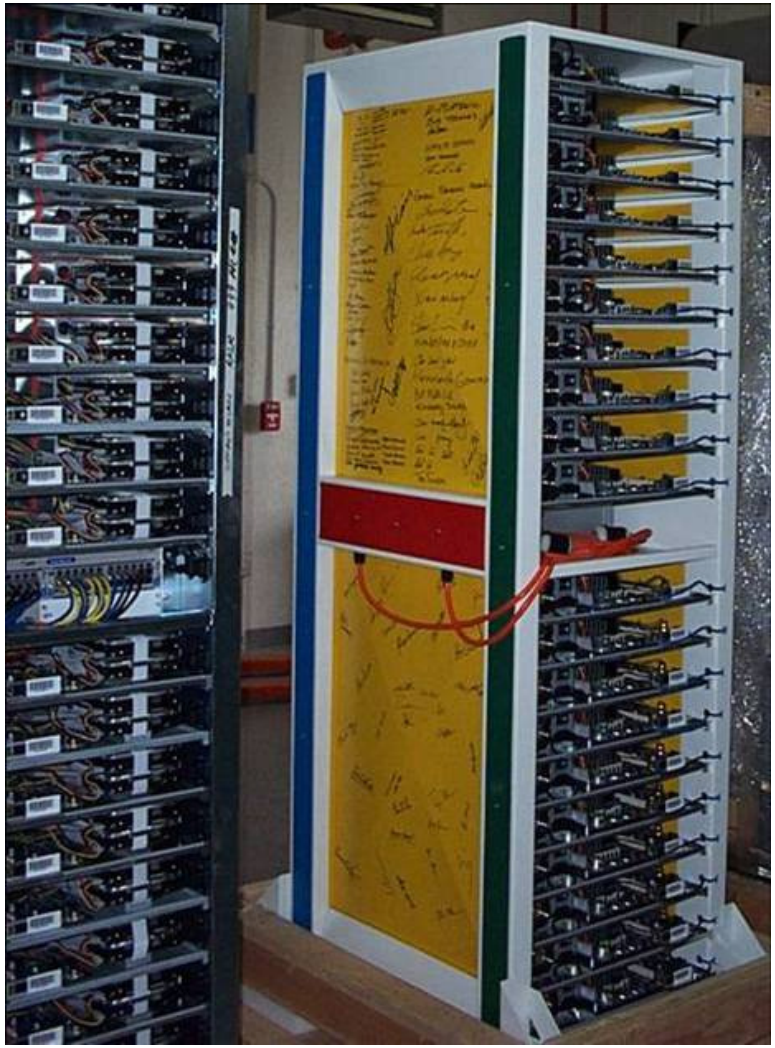
Empty Google Data Center (2001)

2001 + 3 days

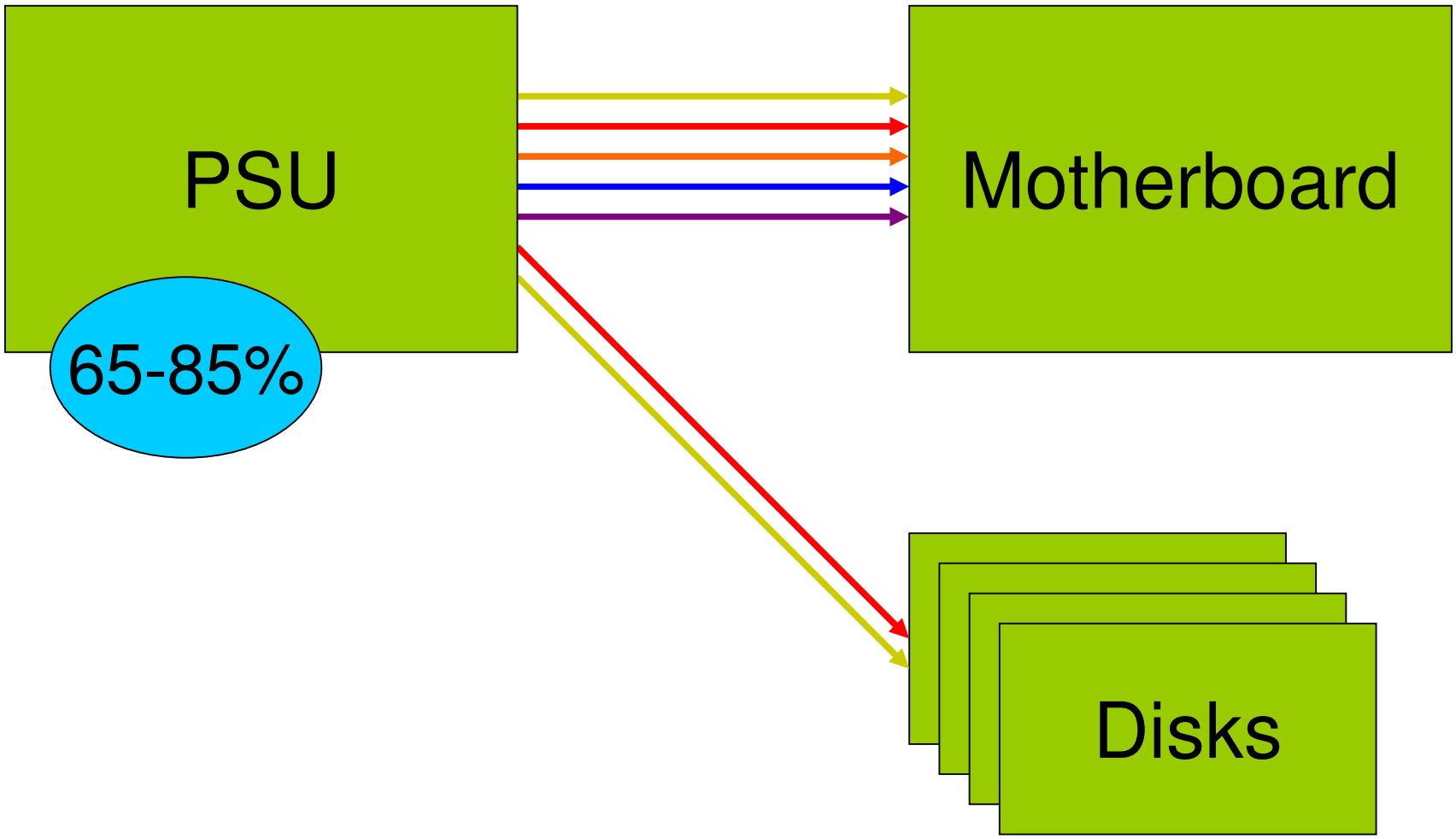


Google 2003

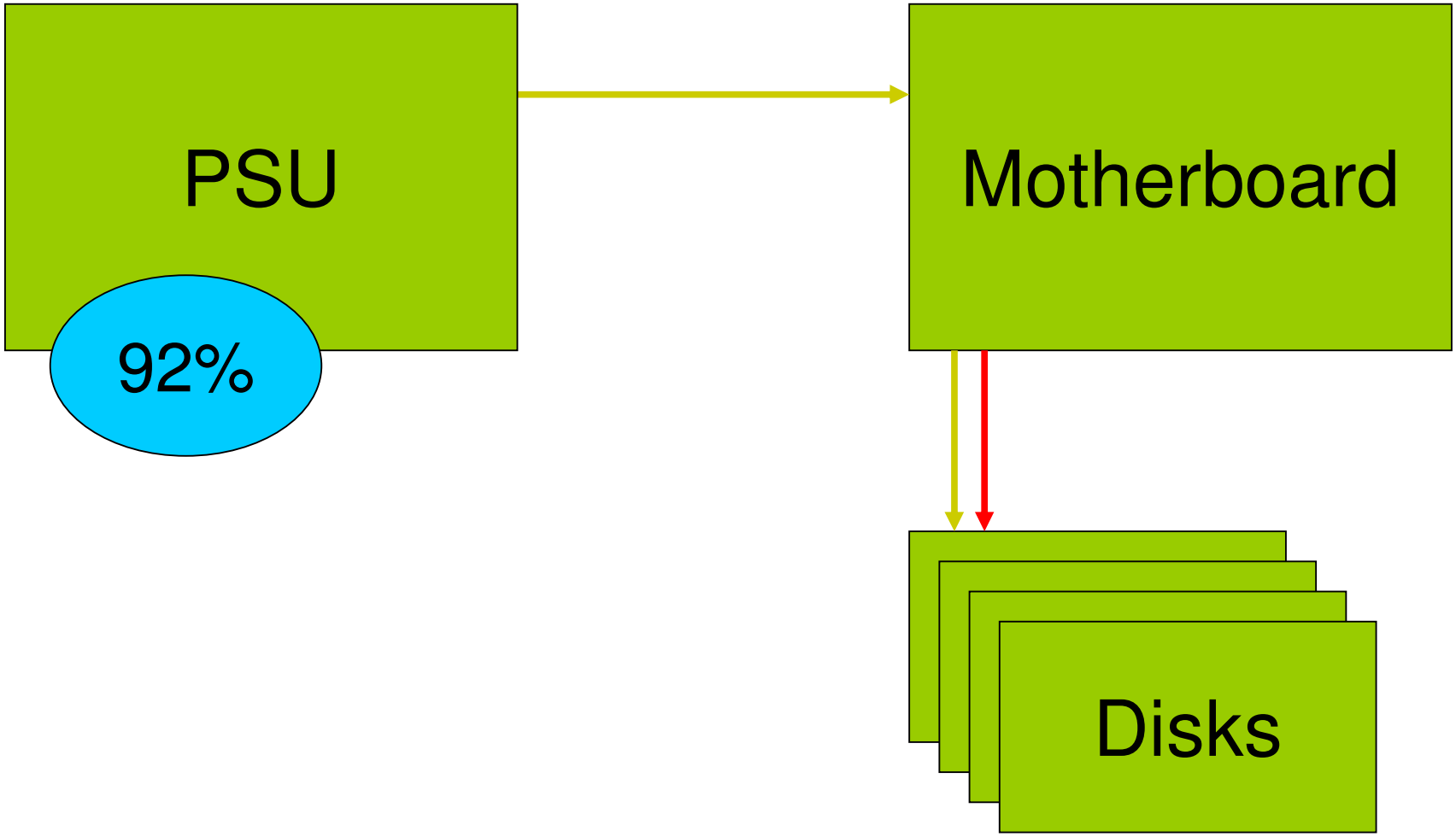
- Custom mechanical, thermal, electrical and electronics design
- Highly customized PC-class motherboards
- Low-end storage and networking gear
- Running Linux
- In-house management & application software



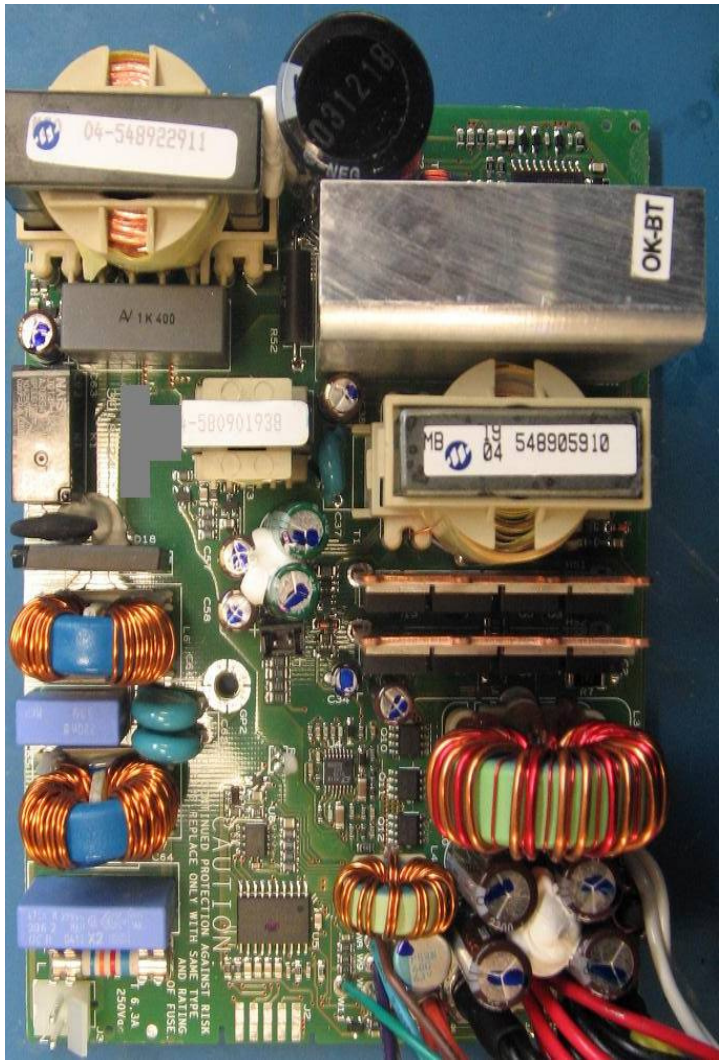
Typical Servers

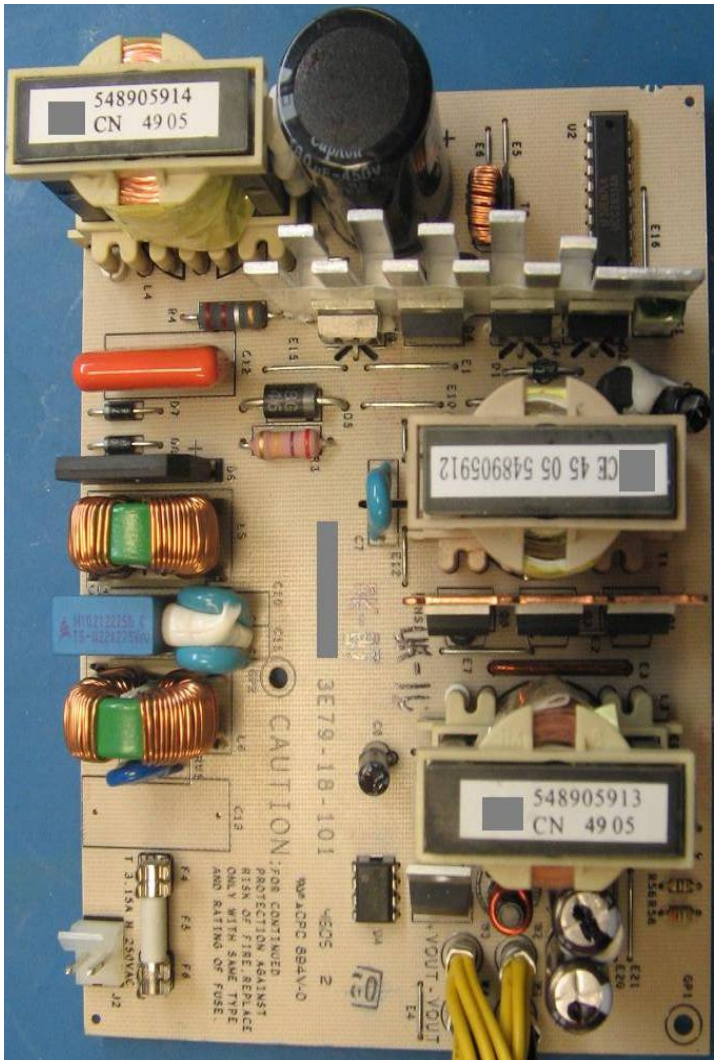


Google Servers

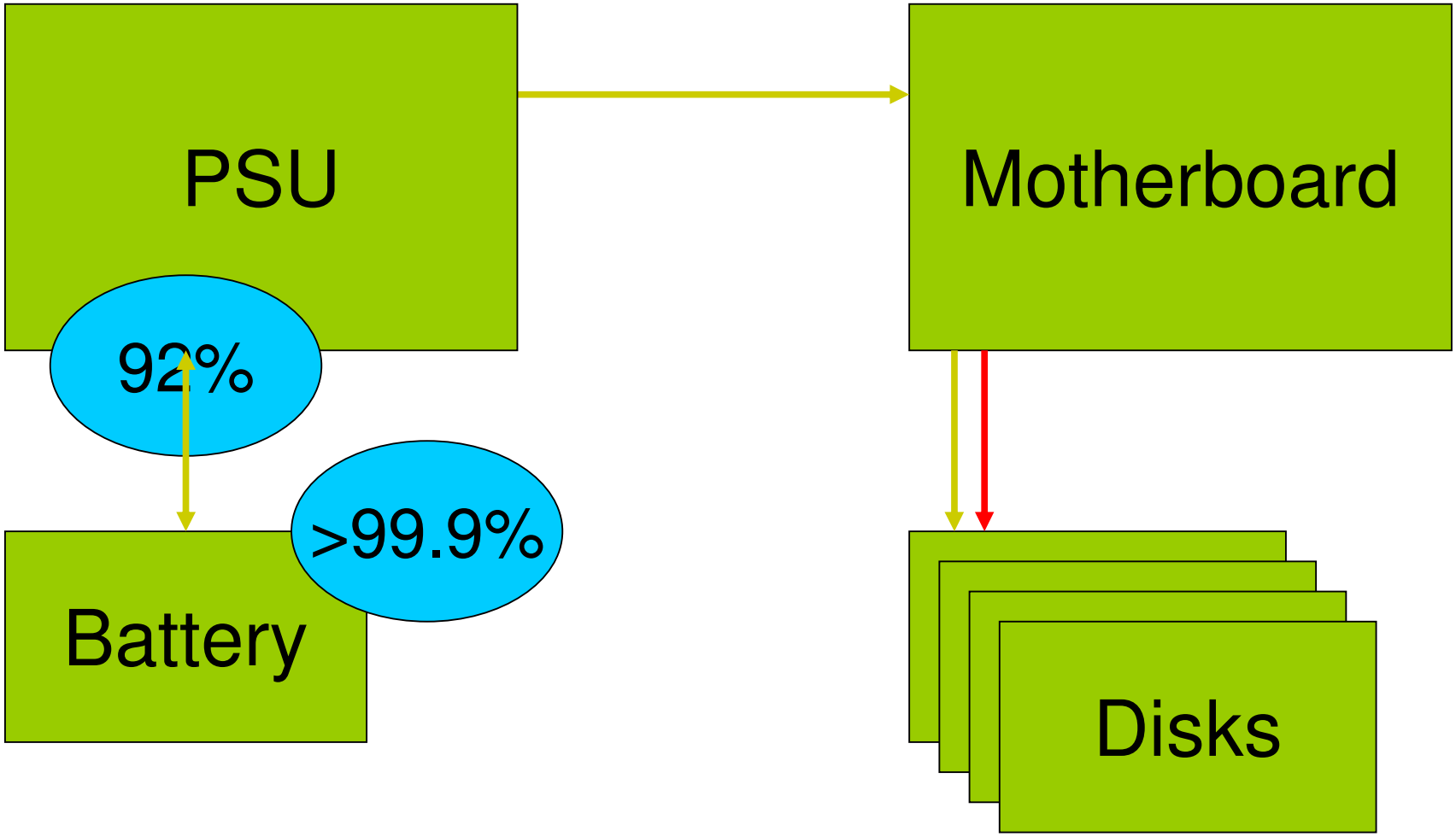


Single-Voltage PSU

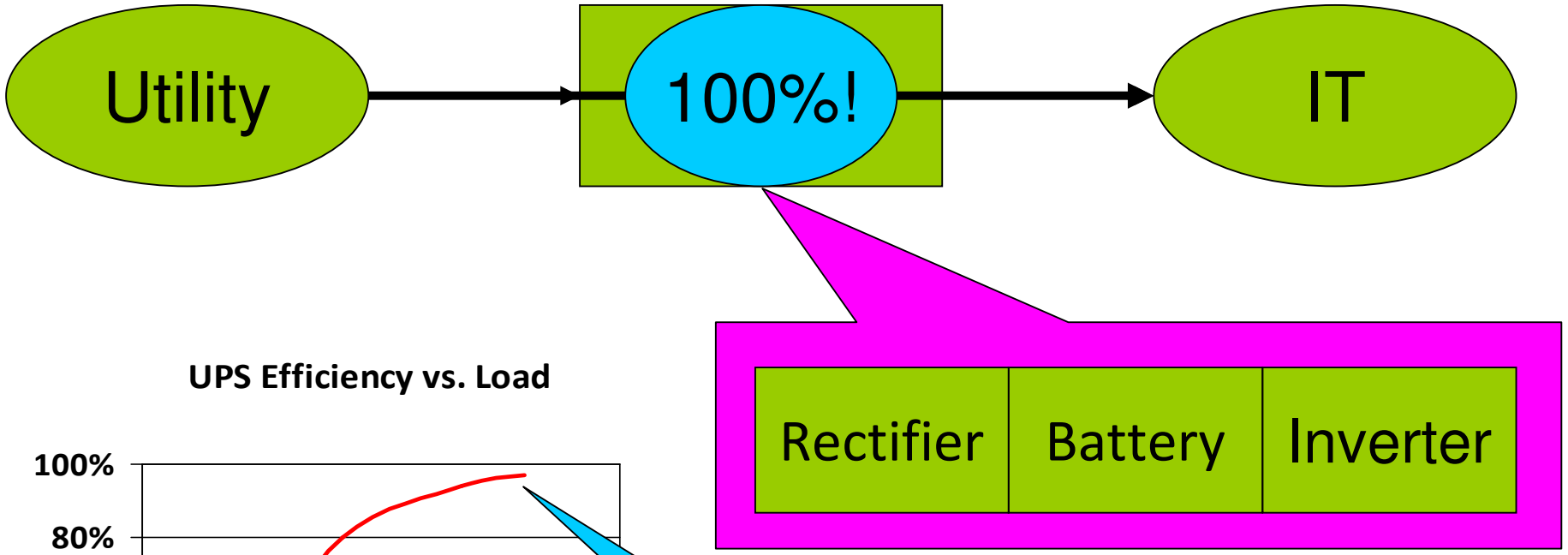


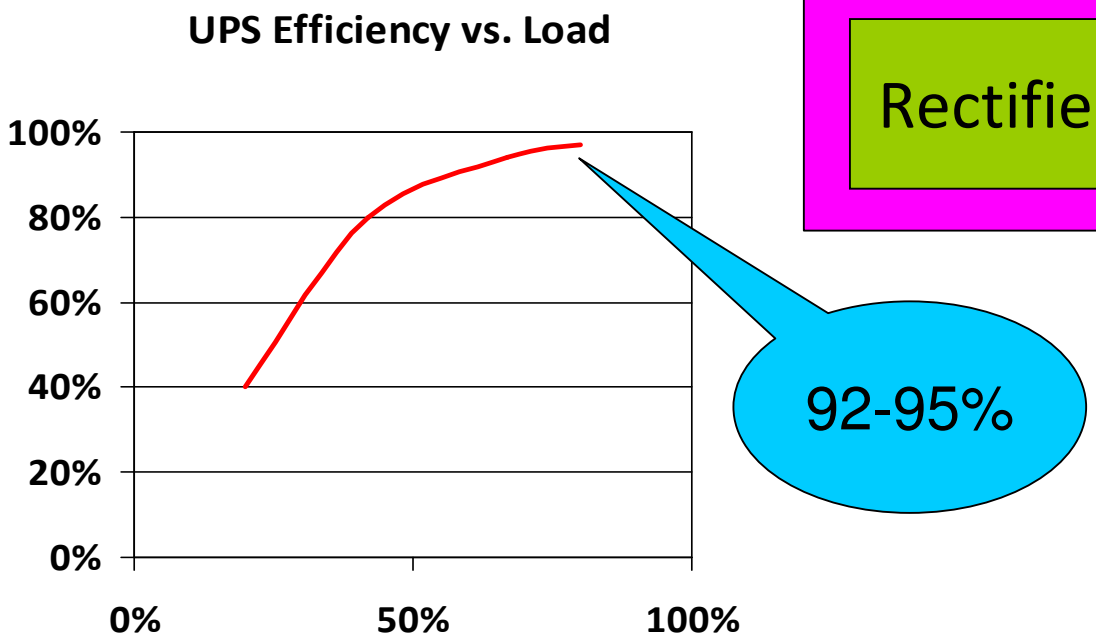


Google Servers

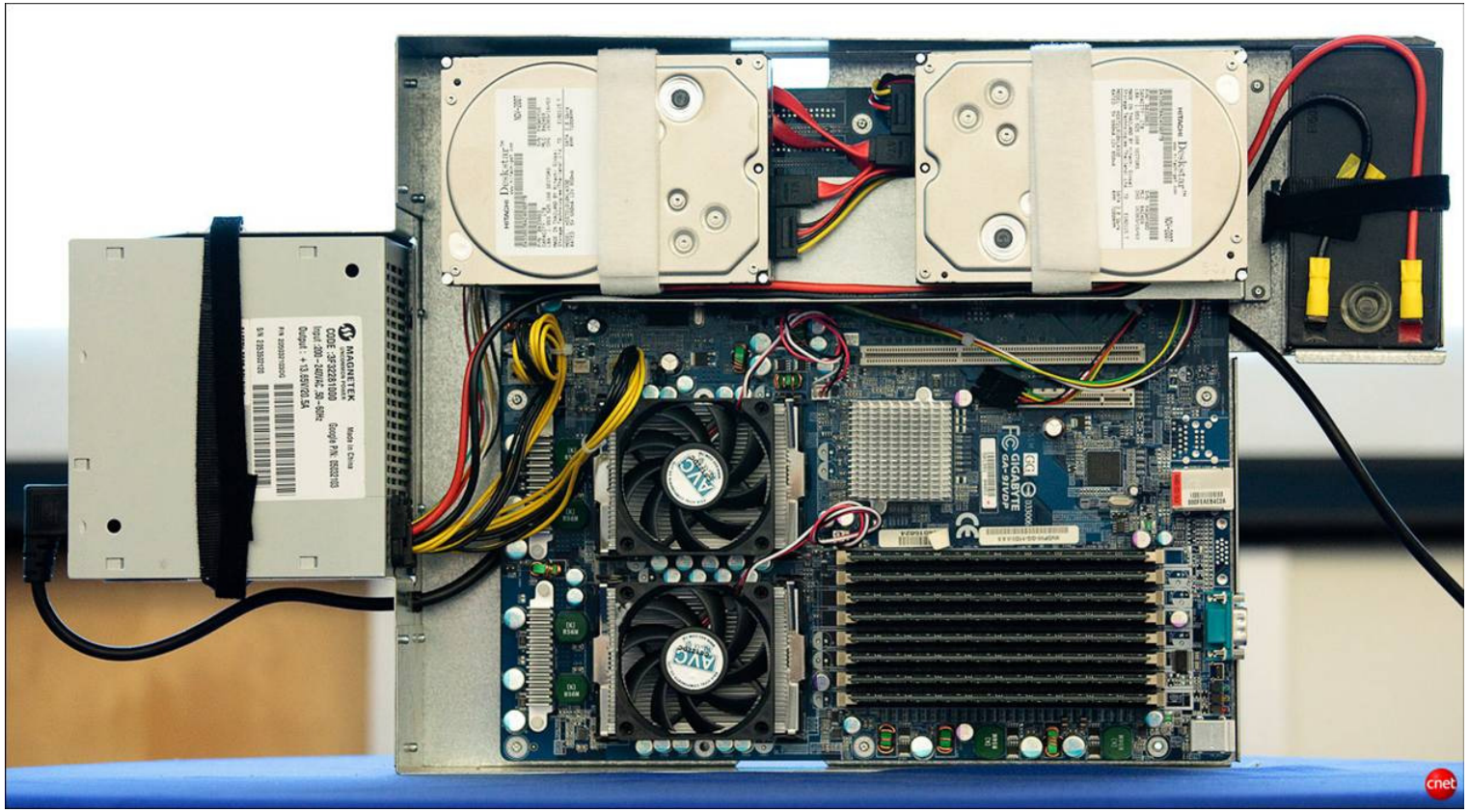


Data Center Power Flow





Google 2005



- Single-voltage power supply with built-in battery
- Shipping container design
- Universal design for compute and storage servers

So, what exactly did Google do?

- Large data sets
- Parallel and distributed computing
- Fault tolerance
- Cheap hardware
- Containerized data center
- Power efficiency

- Are those what cloud computing is about?

Is This a Cloud?

- Shipping container filled with servers
- Redundant networking connections
- Linux
- Hadoop
- Hypervisor
- Server consolidation through virtualization
- VPN access
- Remote desktop

NIST Definition of Cloud Computing

Cloud computing is a model for enabling **convenient, on-demand network access** to a shared pool of **configurable computing resources** (e.g., networks, servers, storage, applications, and services) that can be rapidly **provisioned and released** with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models.

http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc

Cloud Computing *Characteristics*

- On-demand self-service
  - Users can unilaterally provision computing capabilities
- Broad network access
  - Resources are available over the network and accessed through standard mechanisms
- Resource pooling
  - Resources dynamically assigned and reassigned, independent of location
- Rapid elasticity
  - Resources can be rapidly and elastically provisioned to quickly scale out and rapidly released to quickly scale in
- Measured Service
  - Control / optimize resource use by metering

Cloud Computing *Service Models*

- Cloud Software as a Service (SaaS)
  - Ability to use the provider's applications running on a cloud infrastructure.
- Cloud Platform as a Service (PaaS)
  - Ability to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider.
- Cloud Infrastructure as a Service (IaaS)
  - Ability to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications.

Cloud Computing *Deployment Models*

- Private cloud
  - The cloud infrastructure is operated solely for one organization
- Community cloud
  - The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, compliance)
- Public cloud
  - The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services
- Hybrid cloud
  - Two or more clouds (private, community, or public) bound together and enable data and application portability

Is This a Cloud?

- Shipping container filled with servers
- Redundant networking connections
- Linux
- Hadoop
- Hypervisor
- Server consolidation through virtualization
- VPN access
- Remote desktop

So, what else did Google do?

- GMail, Calendar (WebApp/SaaS)
- Docs, Sites (WebApp/SaaS, IaaS)
- App Engine (PaaS)
- PicasaWeb online album (WebApp/SaaS, IaaS)
- Android (Client)
- Chrome, Chrome OS (Client)
- GoogleFreeWiFi (Connect)
- FTTH (Connect)

Delta's Mission Statement

- To provide innovative, clean and efficient energy solutions for a better tomorrow.
- Google Translate: 提供創新，廉潔和高效的能源解決方案的一個更好的明天。
- Our own translation: 環保，節能，愛地球

Delta Products

- Power management (converter, inverter, charger)
- Renewable energy (solar, wind, fuel cell)
- Fan and thermal management
- ePaper, projector (3D), digital signage
- LED lights
- Industrial automation
- Automotive (electric and hybrid car parts)
- Networking
- Voice recognition
- Data center and cloud computing

Delta and Cloud Computing

- Applications
- Client
- Connectivity
- Cloud
  - SaaS
  - PaaS
  - IaaS
    - Data center
    - IT equipment
    - Systems software
    - Cloud infrastructure software

- Data center
  - Building, cooling
  - Power, backup power
  - Management system
- IT equipment
  - Server
  - Storage
  - Networking
  - Management system
- Systems software
  - Operating systems
  - Device drivers
- Cloud infrastructure software
  - Reliable storage
  - Efficient, scalable database
  - Cluster management
  - Resource and task management

Q&A

Copyright Notice

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/3.0/ or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.