

Powered by YouTube:  
**Scaling YouTube**

Cuong Do  
7/10/2008

# Agenda

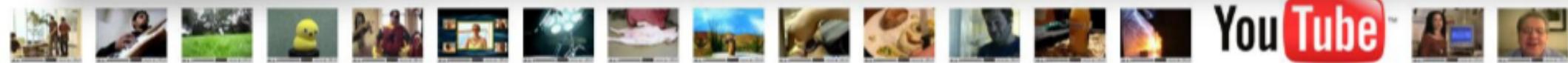
- Historical context
- Scalability highlights
- Challenges
- Lessons learned
- Q & A



# What is this graph?

1. The price of a gallon of gas
2. My weight after joining Google
3. YouTube's daily video views

YouTube Confidential and Proprietary



# Site Growth Timeline

- 2/2005: founded
- 10/2005: first round of VC funding; first office!
- 03/2006: 30 million video playbacks / day
- 07/2006: 100 million video playbacks / day
- 11/2006: acquired by Google: FREE FOOD!  
(already a top 10 site by multiple measures)
- Q3 2007: #4 on Alexa
- Today: 10 hours of video uploaded per minute  
“hundreds of millions” of video playbacks / day

YouTube Confidential and Proprietary



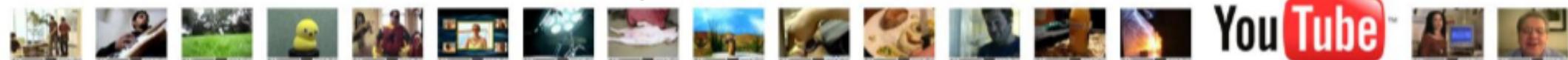
# Pre-Acquisition Team

- 2 system administrators
- 2 scalability software architects
- 2 feature developers
- 2 network engineers
- 1 database administrator

- 0  
chefs!

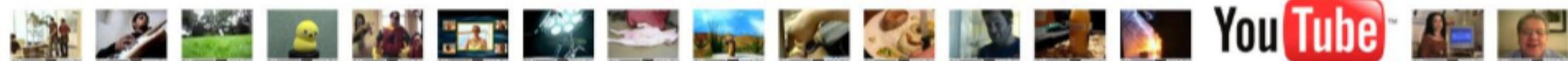


YouTube Confidential and Proprietary



# Handling Rapid Growth

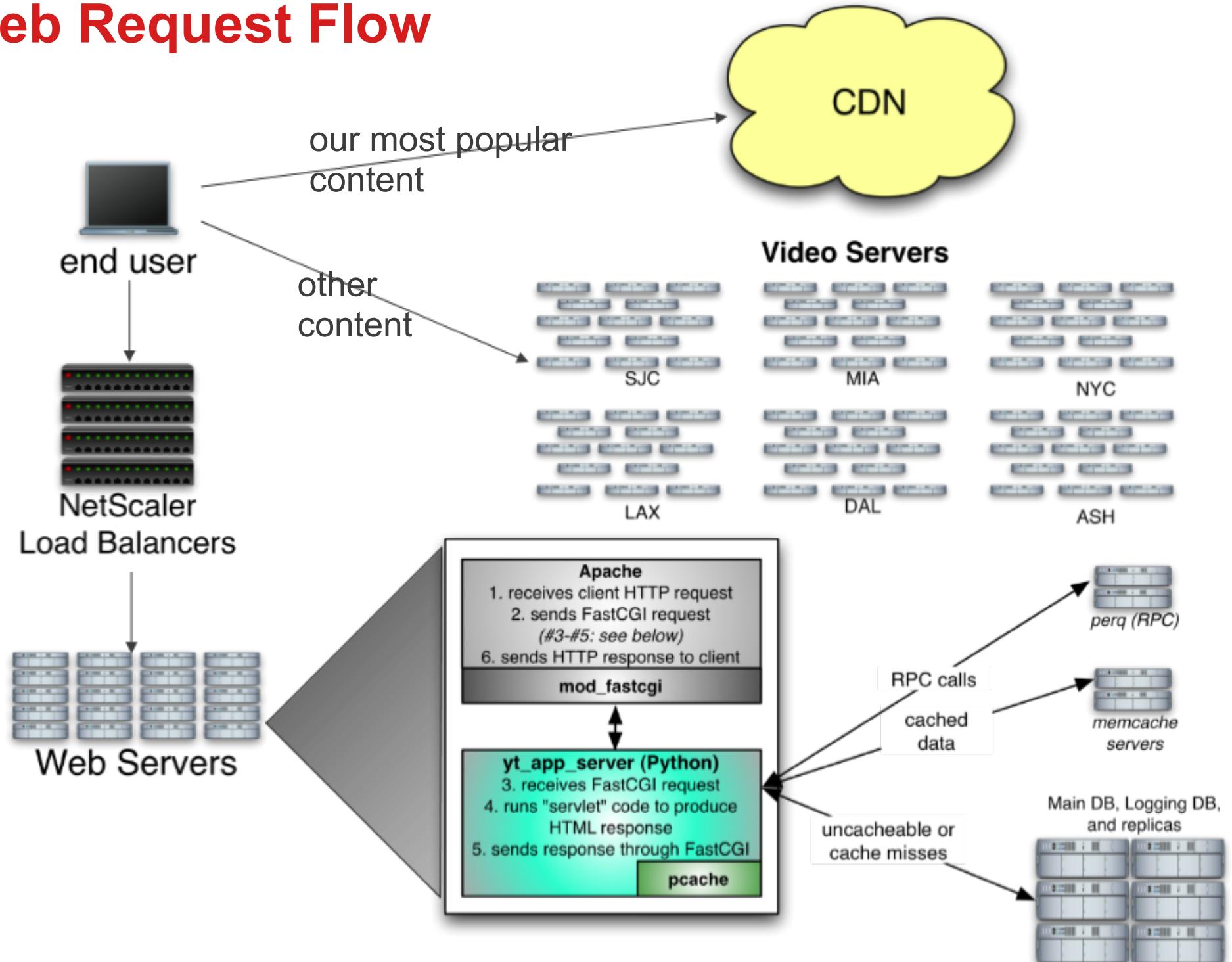
```
while(true)
{
    identify_and_fix_bottlenecks();
    drink();
    sleep(TOO_LITTLE);
    notice_new_bottleneck();
}
```



**Sometimes we  
combine a few steps...**

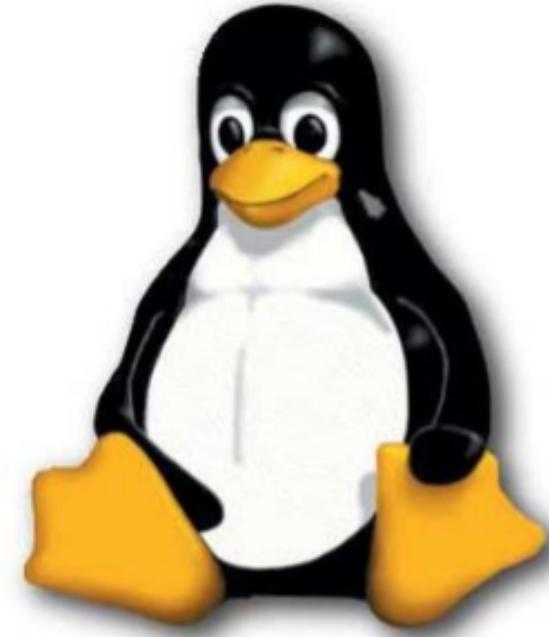


# Web Request Flow

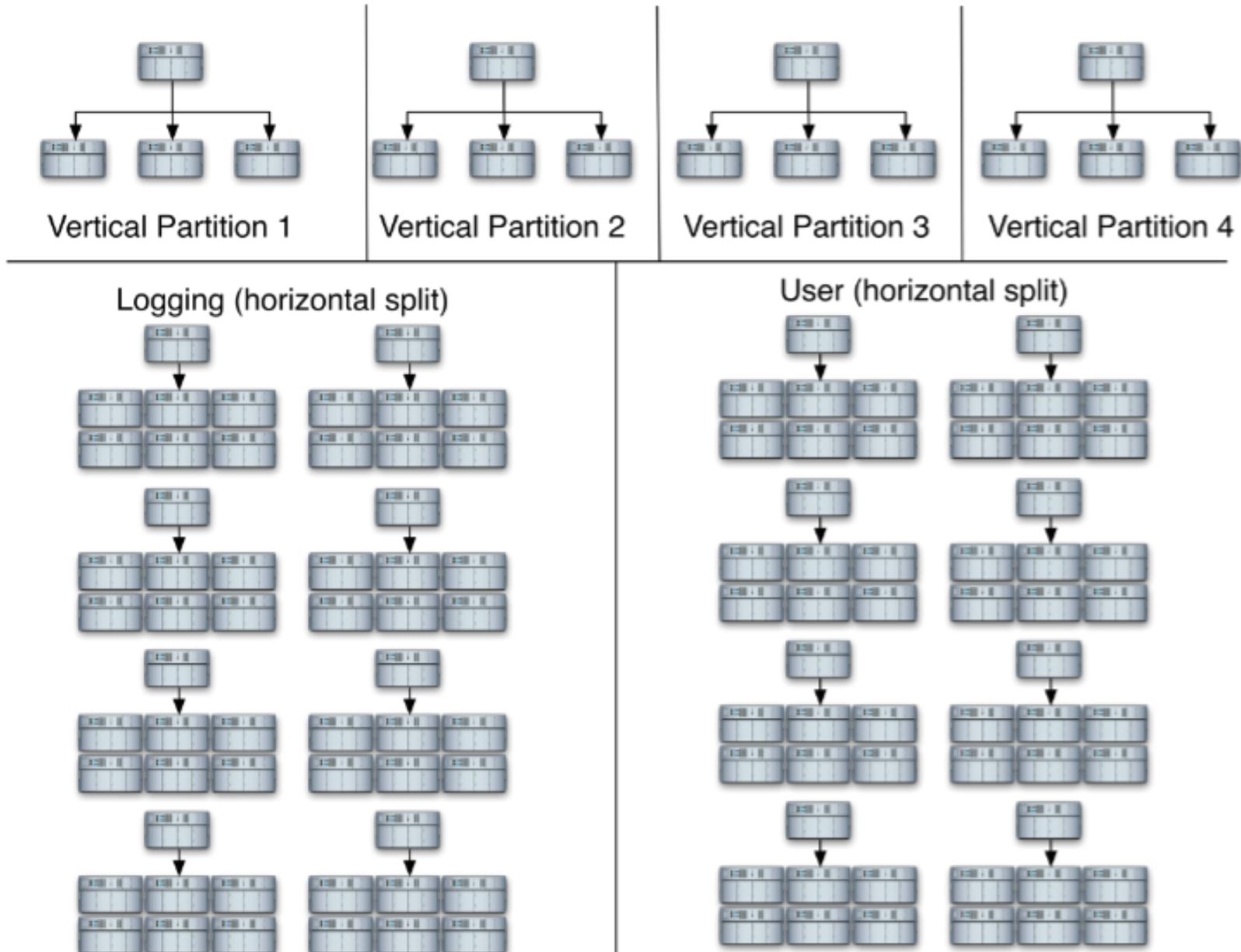


# Key technologies

- Linux (SuSE 10.x)
- Apache (2.2.x) / lighttpd 1.4.x
- MySQL 5.0.x
- Python for 99.999% of application code
  - C extensions for performance-critical code
  - Web machines are cheaper than talented engineers
  - Cheetah template system
- + Google technologies (Search, Bigtable, GFS, etc.)



# Databases

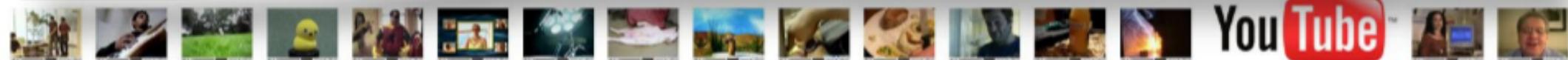


YouTube Confidential and Proprietary



# Scalability Challenges

- Rapid, unpredictable growth
- Passionate users
- New features
- Pushing hardware and software boundaries
- Performance bottlenecks getting more obscure



# No More Videos?

**From:** [foo@youtube.com](mailto:foo@youtube.com)

**Subject:** oh @#!%.

**Date:** October 22, 2005 2:24:33 AM PDT

**To:** all of youtube

we can't accept any more videos.  
too many videos.

Culprit:  
inodes / directory

YouTube Confidential and Proprietary



# Pushing our database to its limits

The site, which allows viewers to tune into uploaded videos ranging from seniors making their homespun video debut to polished user-submitted film clips, went down at 7:30 a.m. (PST), leaving viewers with a tongue-in-cheek graphic of the company's database woes.

- CNET, 8/15/2006

## Culprit:



YouTube Confidential and Proprietary



**“What? You didn’t want those servers there?”**

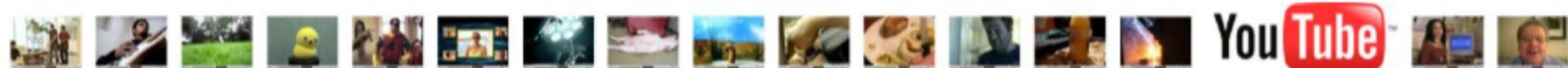


YouTube Confidential and Proprietary



# Ingredients for Scalability

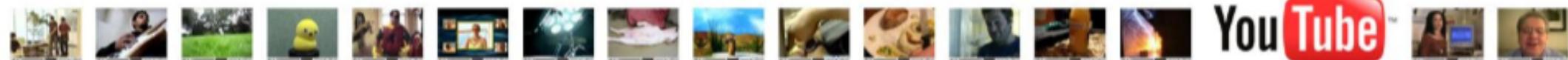
- Keep it simple!
- Iterate, iterate, iterate
  - Software: DB, caching, web server tweaks
  - OS tweaks
  - Hardware: memory, RAID
- Be persistent
- And...



# Last Ingredient: A Great Engineering Team



YouTube Confidential and Proprietary



# Q & A

YouTube Confidential and Proprietary

