

Today

✓ * Quicksilver

Internet scale computing

(Lesson 9)

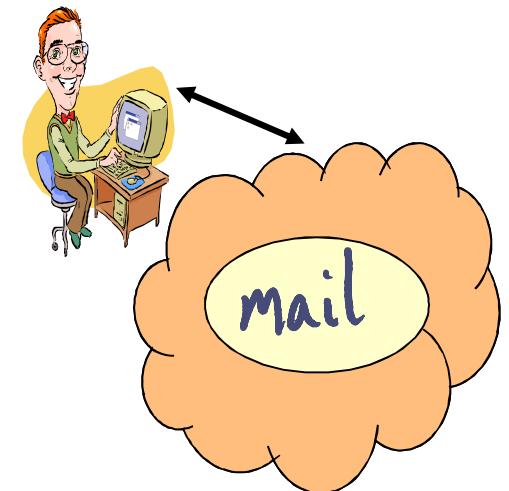
⇒ * giant scale services

* Map-reduce programming model

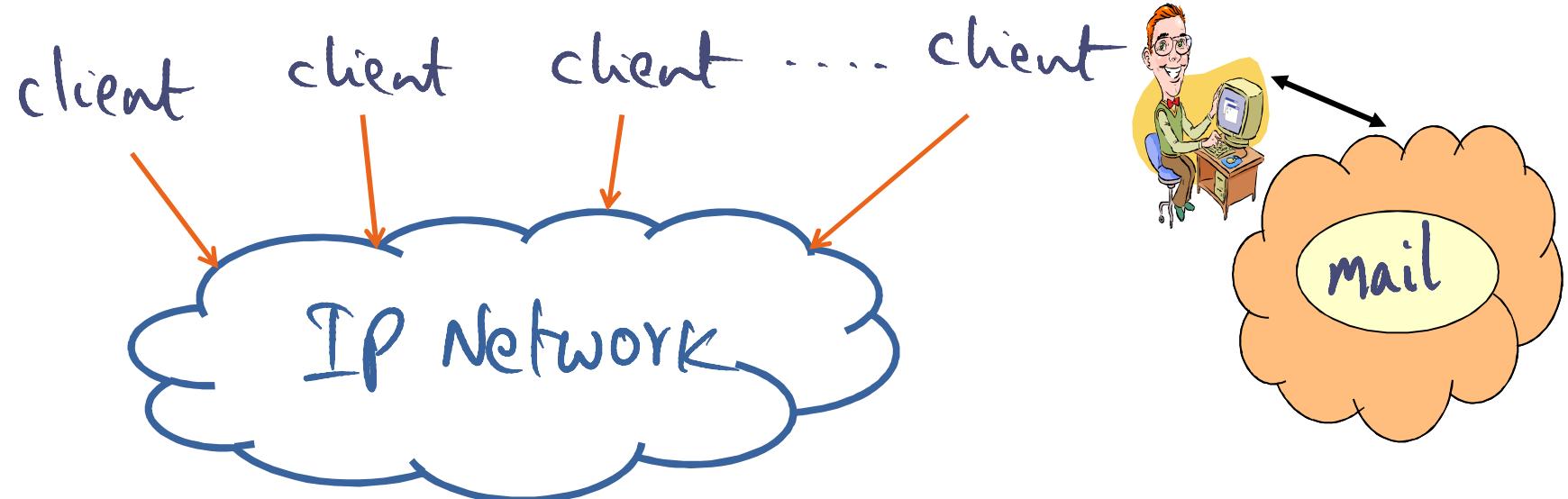
* CDN (Coral)

* Please watch videos before class *

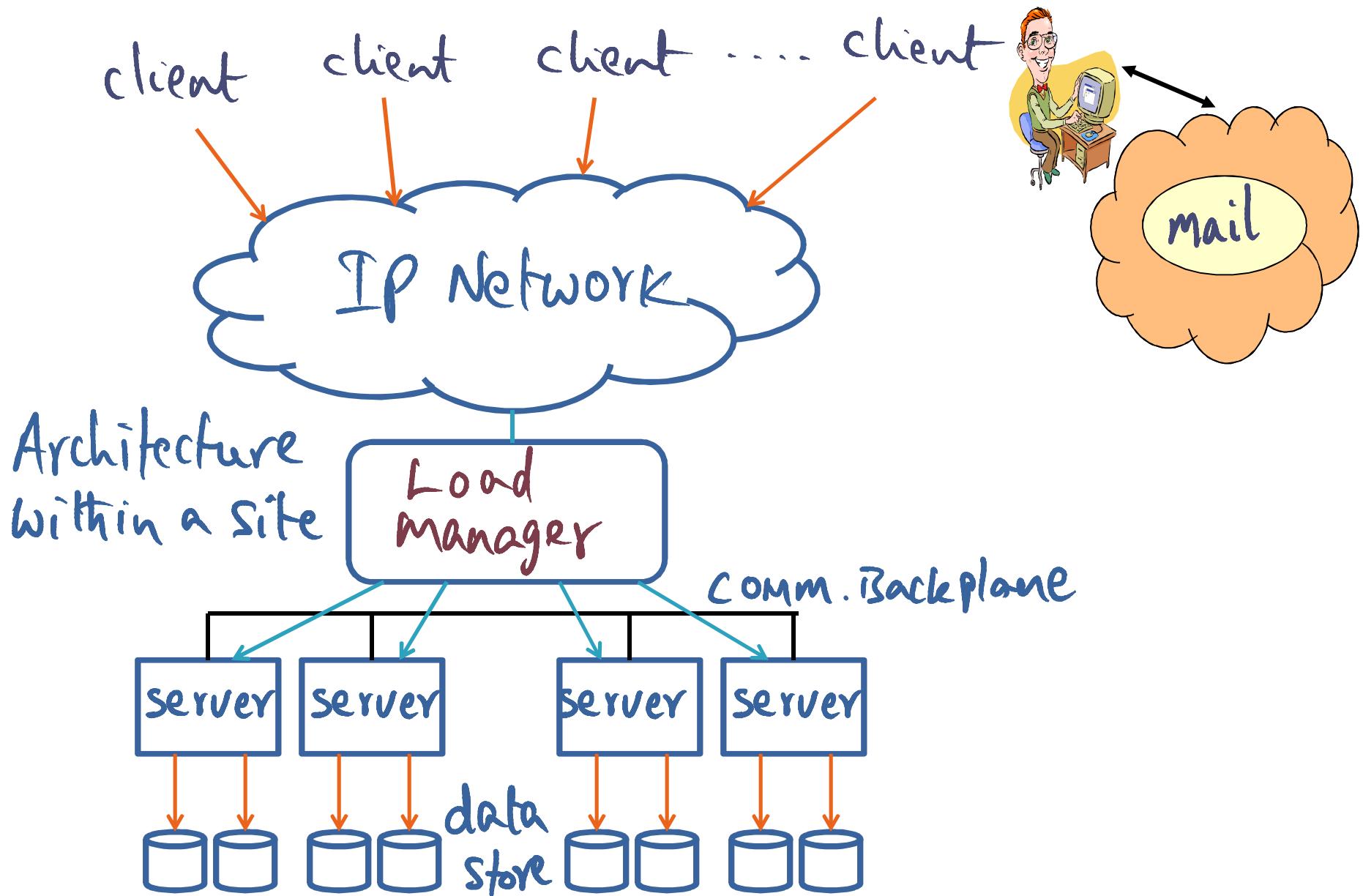
Generic Service model of giant scale Services



Generic Service model of giant scale Services



Generic Service model of Giant scale Services

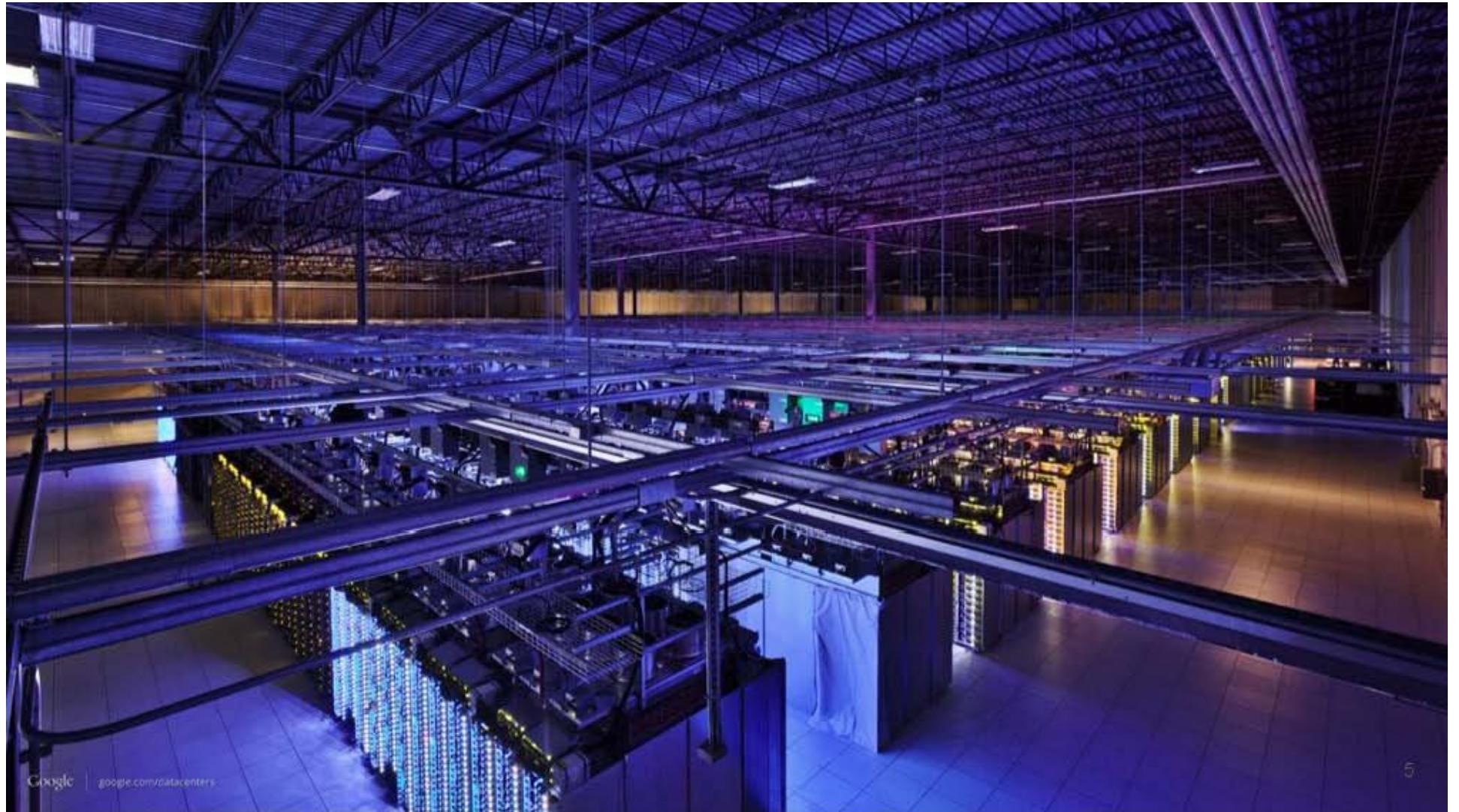


Google Data Center



Google data center: Courtesy Jeff Dean's slide deck from History day at SOSP 2015

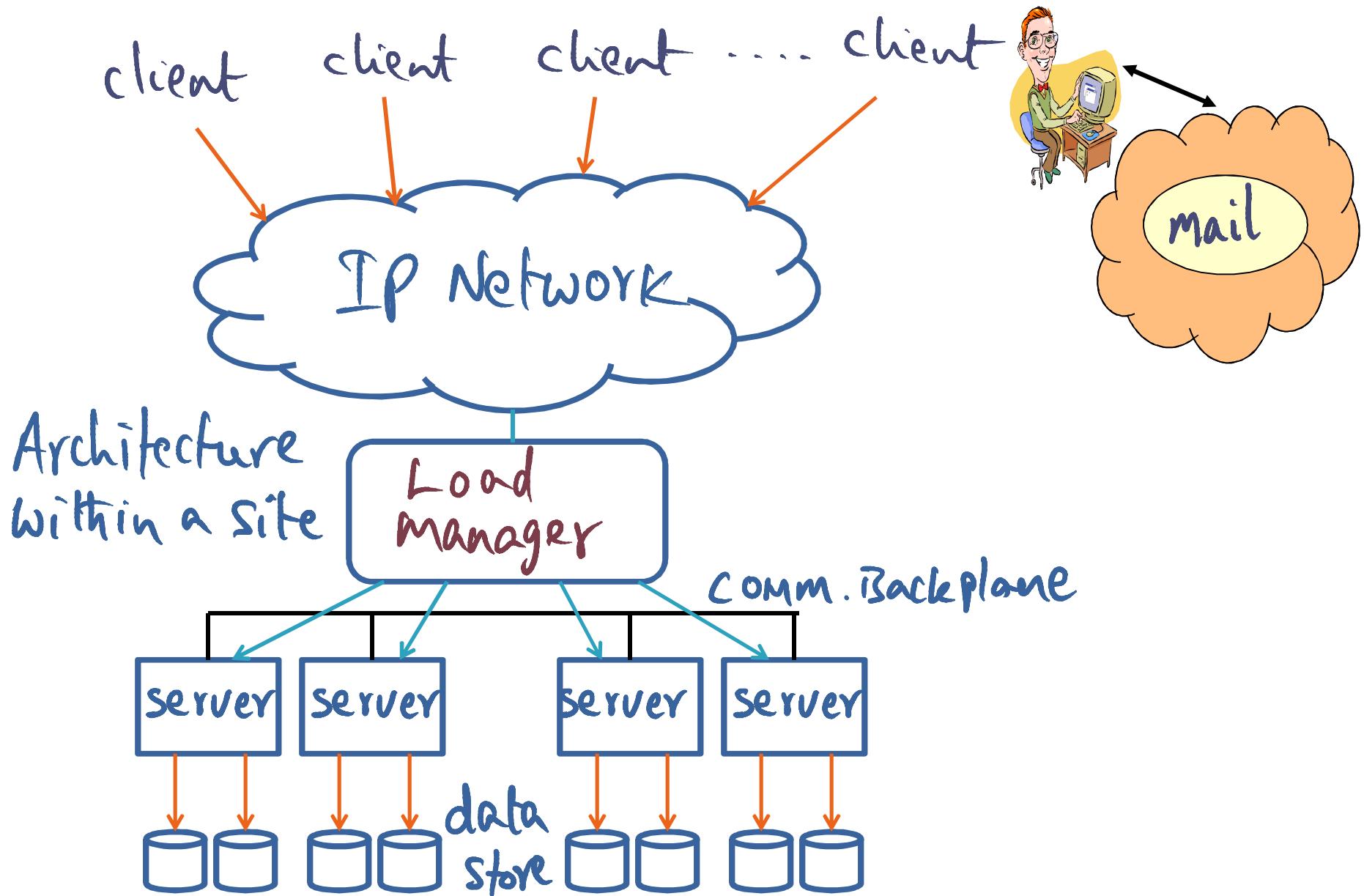
Google Data Center



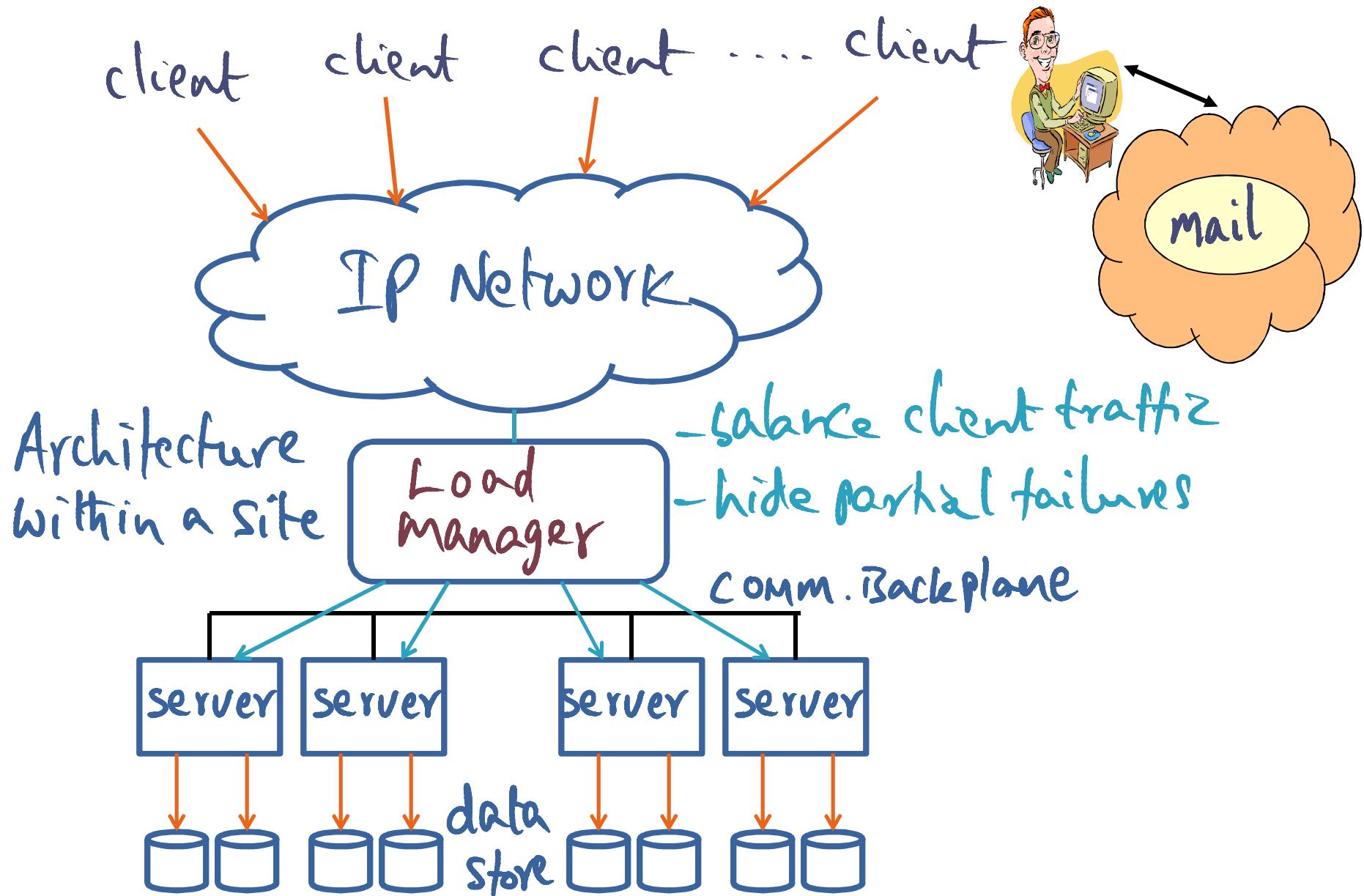
Google | google.com/datacenters

Google data center: Courtesy Jeff Dean's slide deck from History day at SOSP 2015

Generic Service model of Giant scale Services



Generic Service model of Giant scale Services



Clusters as Workhorses

Circa 2000

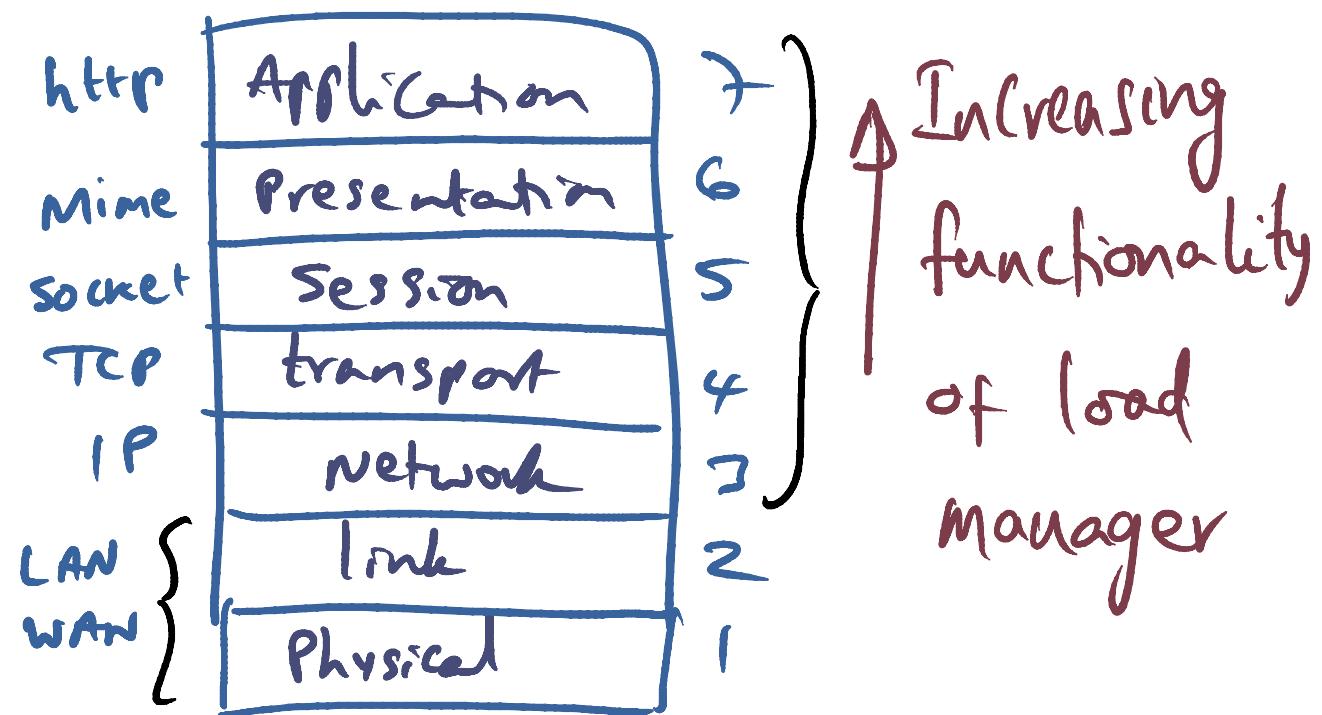
Table A. Example clusters for giant-scale services.

Service	Nodes	Queries	Nodes
AOL Web cache	>1,000	10B/day	4-CPU DEC 4100s
Inktomi search engine	>1,000	>80M/day	2-CPU Sun Workstations
Geocities	>300	>25M/day	PC Based
Anonymous Web-based e-mail	>5,000	>1B/day	FreeBSD PCs

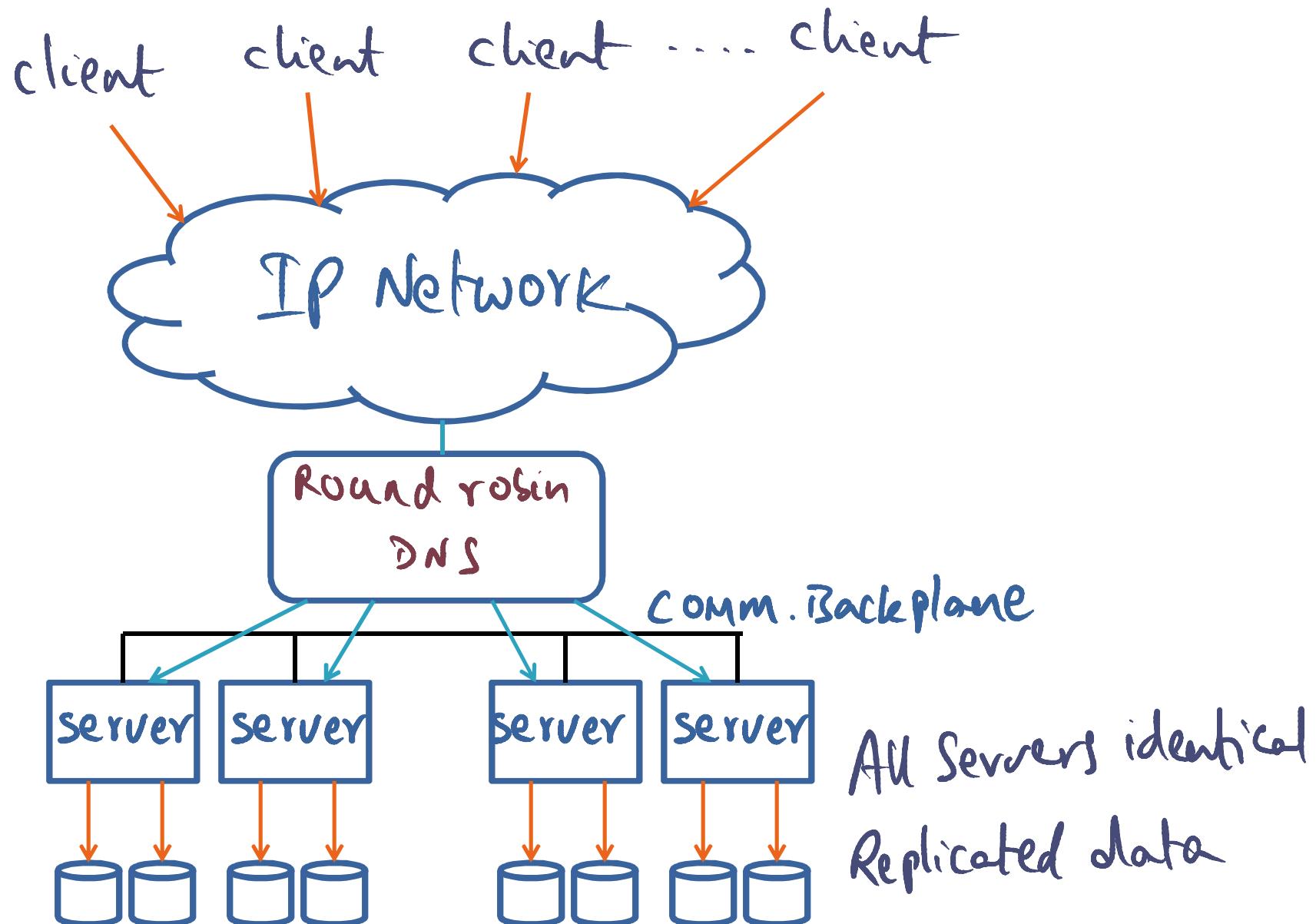
10x to 100x
today compared
to circa 2000

Load management choices

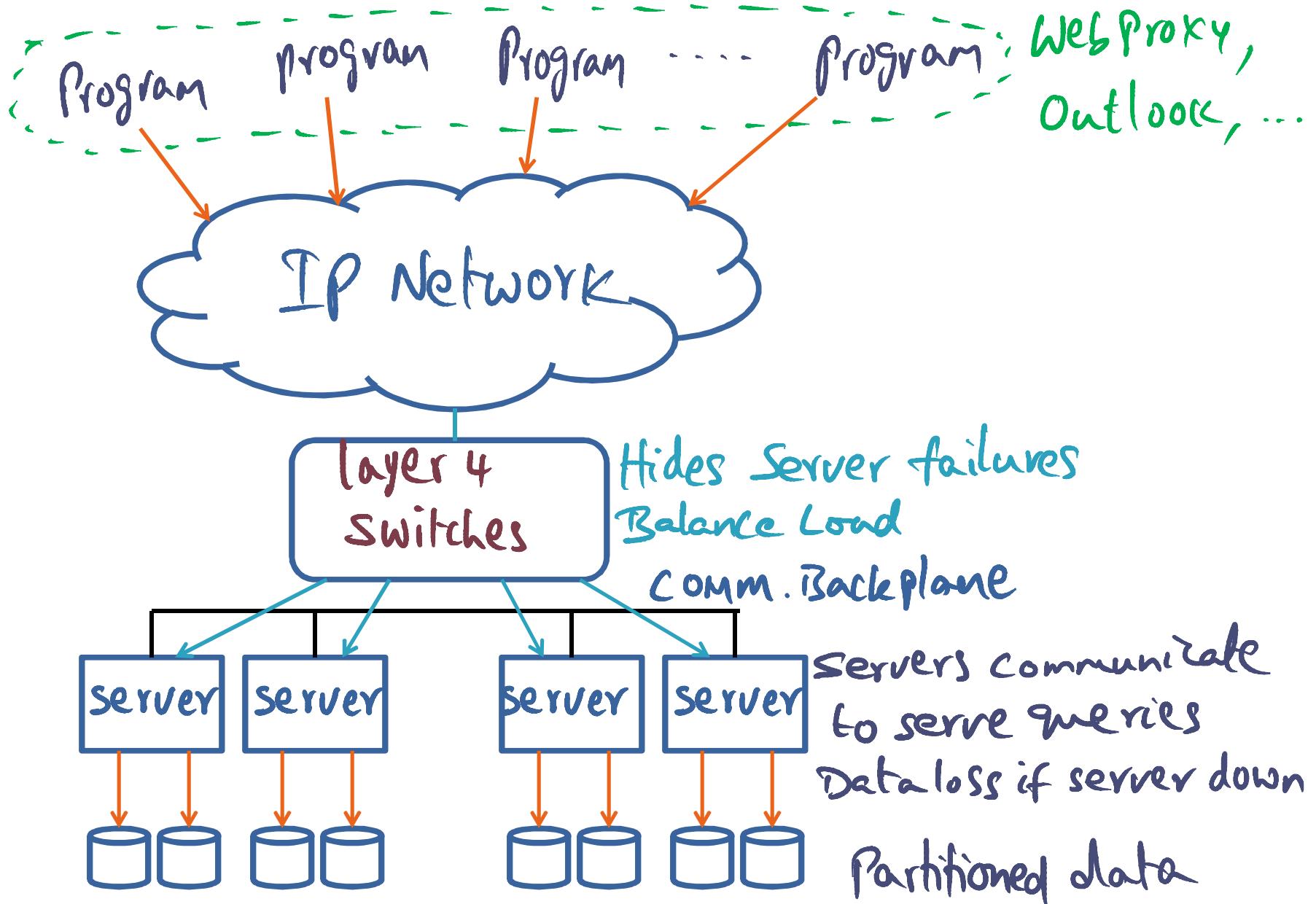
OSI reference model



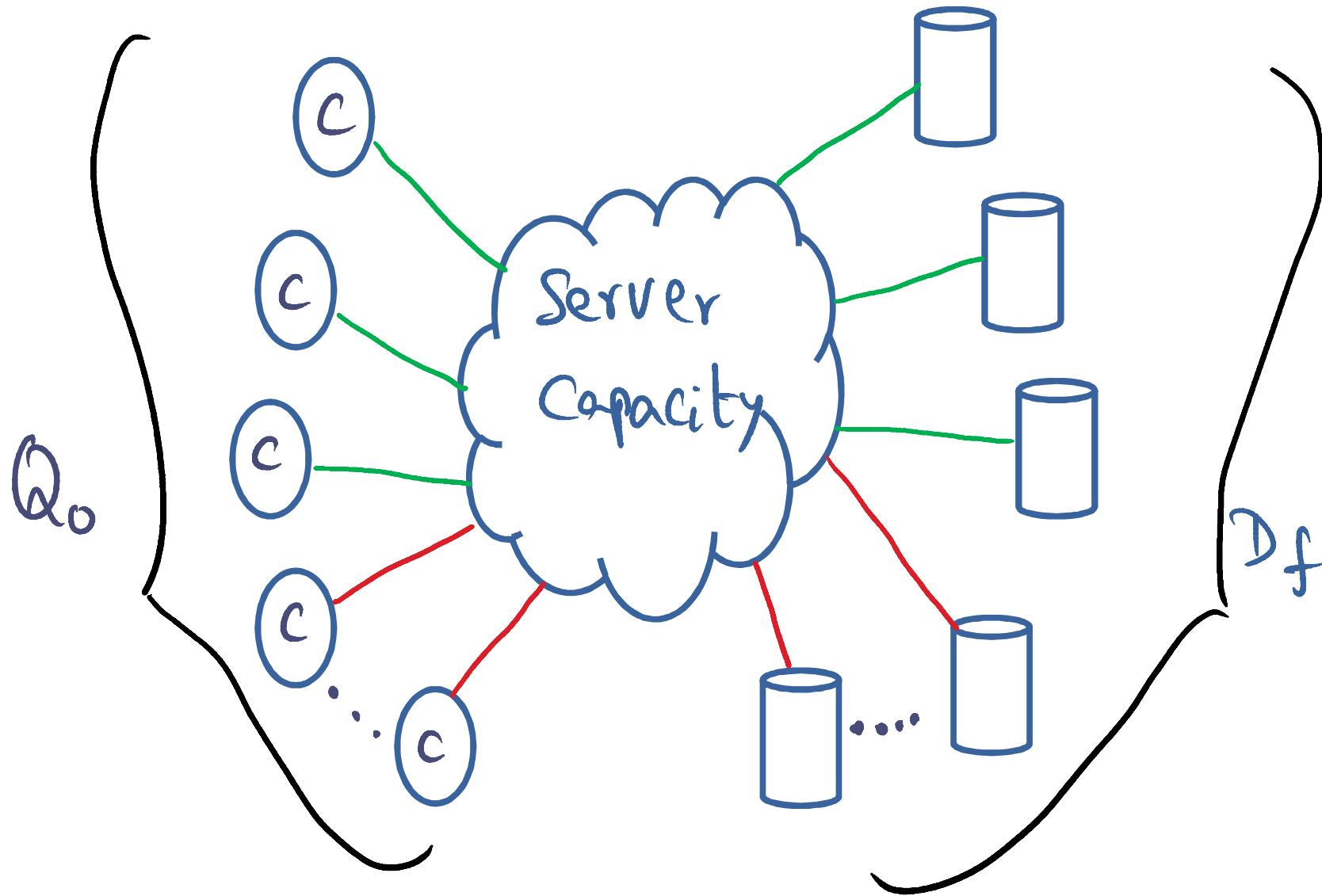
Load management at Network level



Load management at transport level or higher

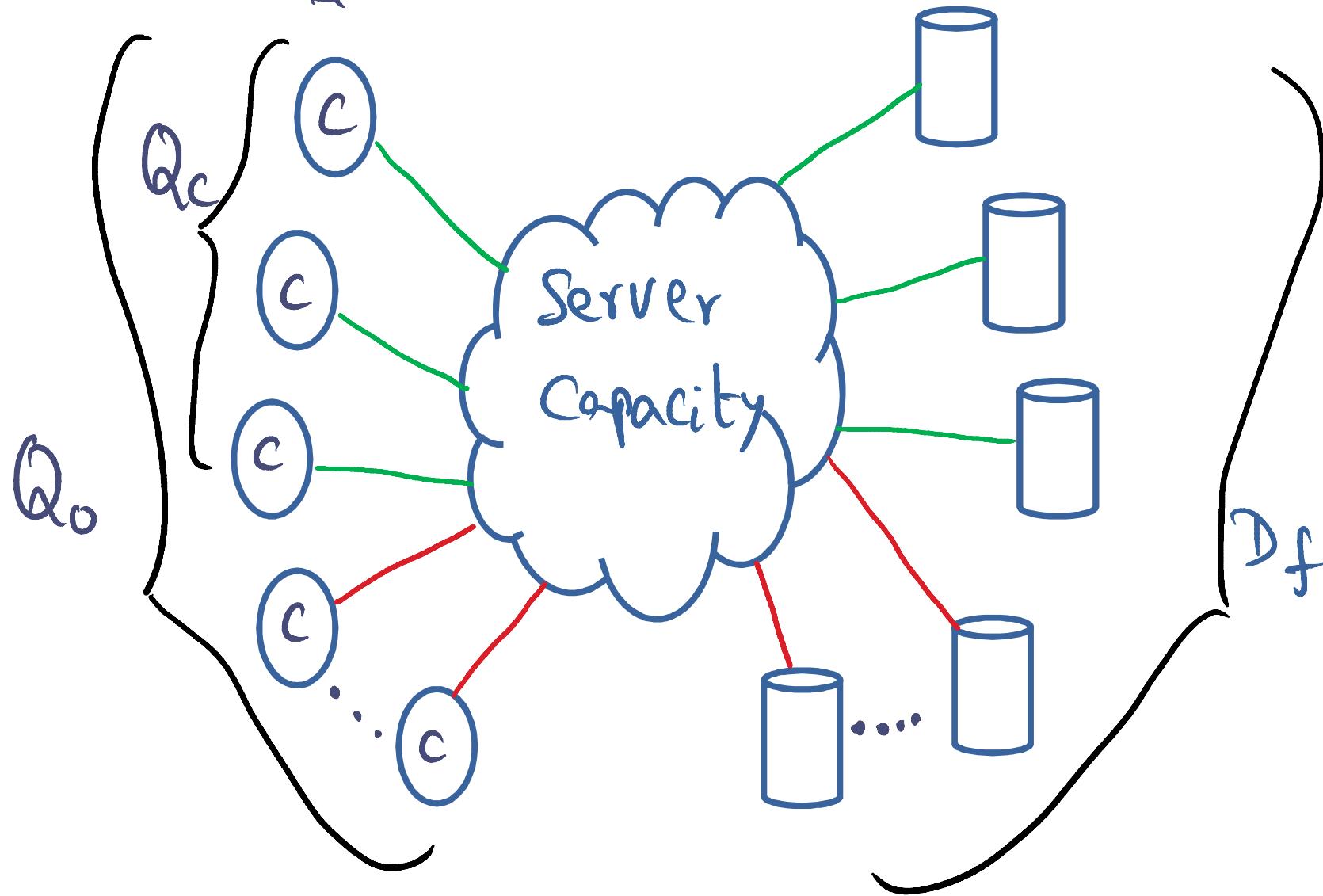


DQ Principle



DQ Principle

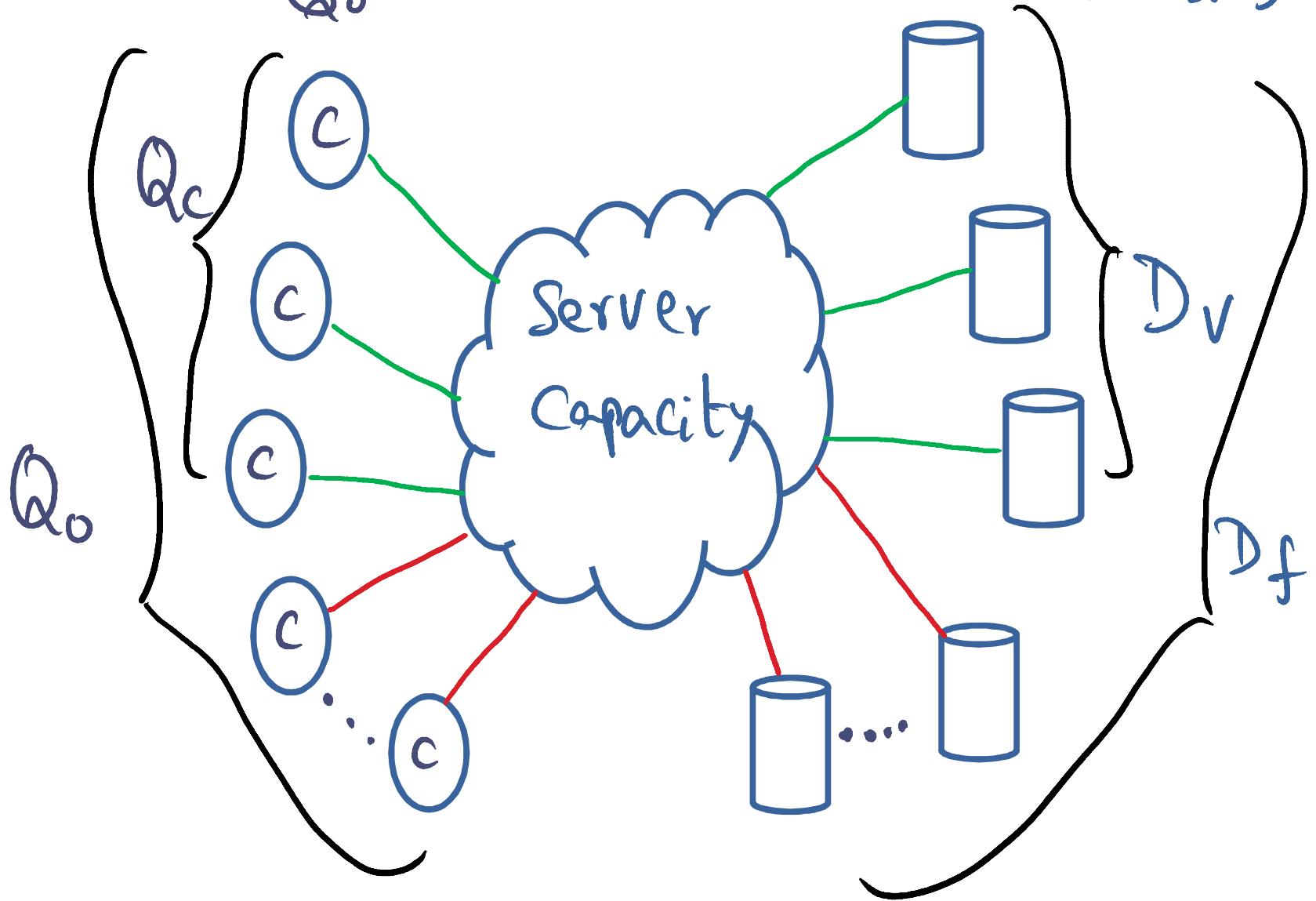
$$\text{Yield: } Q = \frac{Q_c}{Q_0}$$



DQ Principle

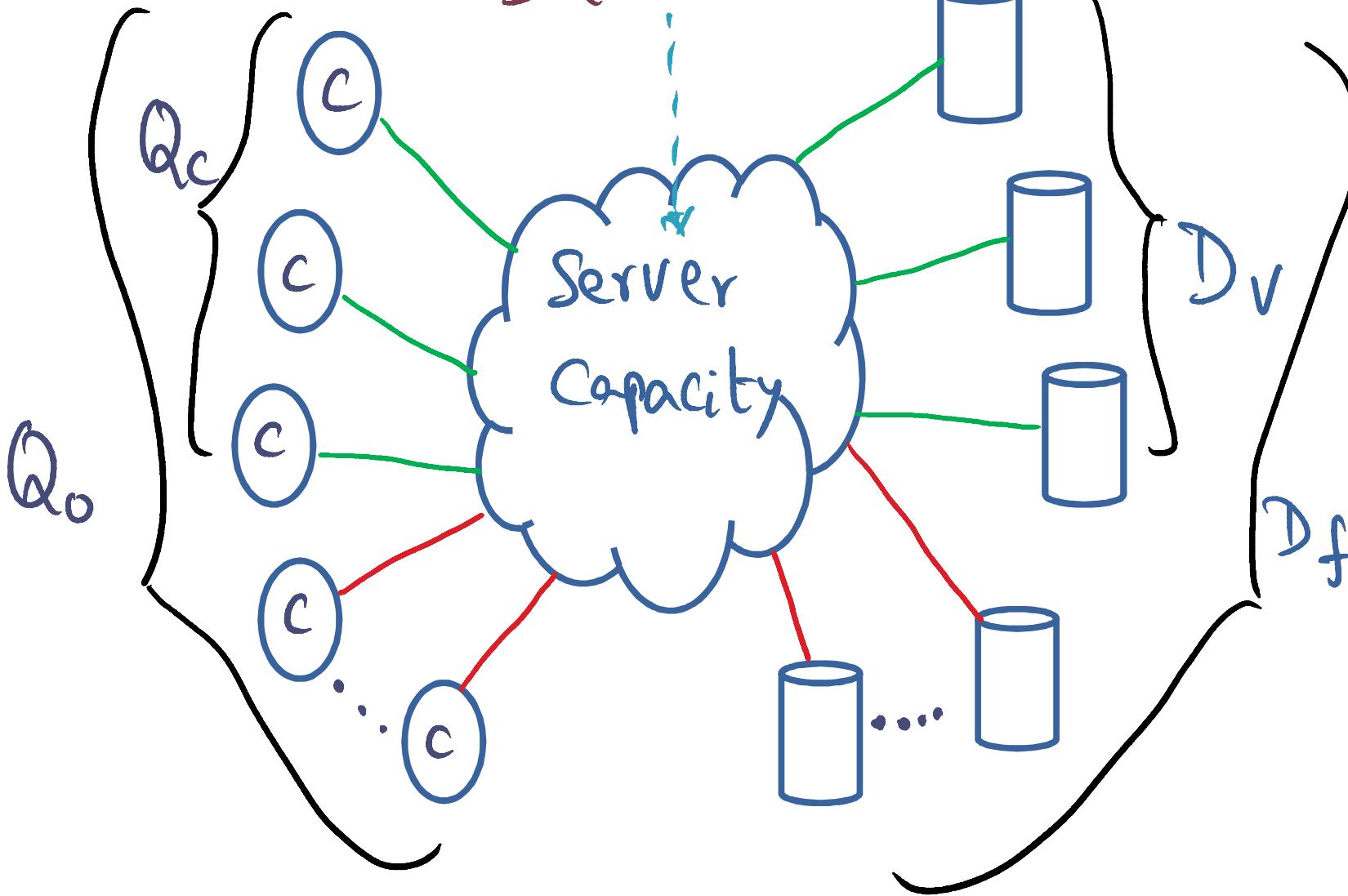
$$\text{Yield: } Q = \frac{Q_c}{Q_0}$$

$$\text{Harvest: } D = \frac{D_v}{D_f}$$



DQ Principle

$$\text{Yield: } Q = \frac{Q_c}{Q_0}$$



$$\text{Harvest: } D = \frac{D_v}{D_f}$$