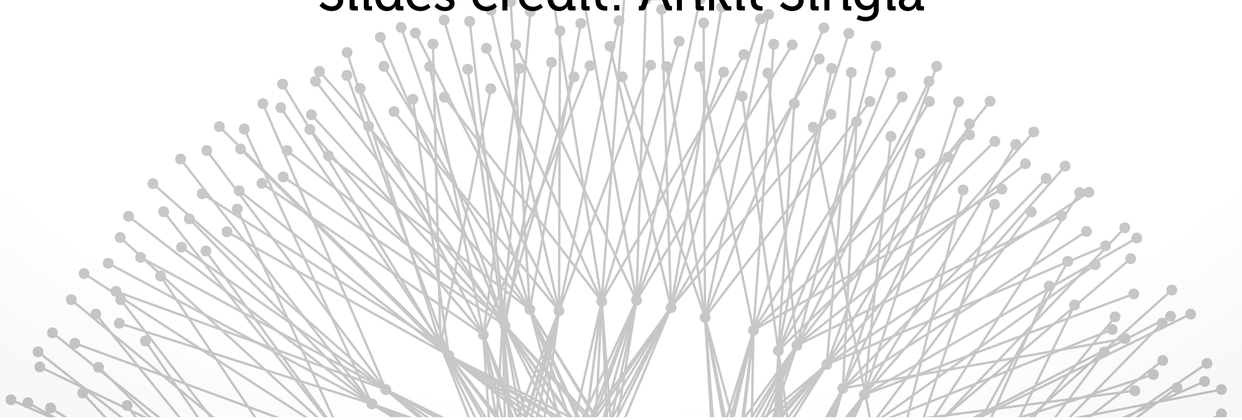
# Cloud Services

Brighten Godfrey
CS 538 April 10 2017

Slides credit: Ankit Singla



# Applications and network traffic




Dropbox

coursera



amazon

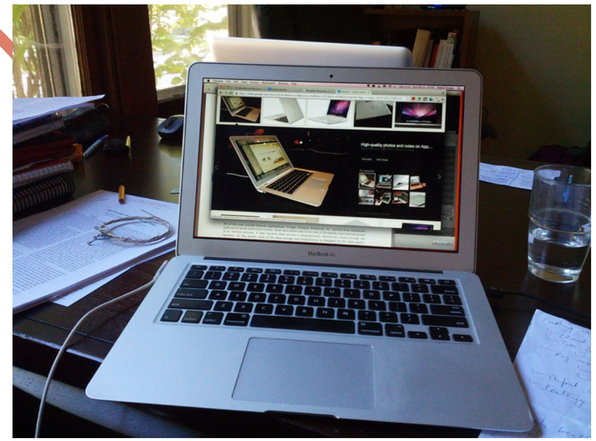
Google

bing

NETFLIX

# How a Web search works

# How a Web search works



# Scatter/gather traffic pattern

Extremely short response deadlines for each server — 10ms

*"Up to 150 stages, degree of 40, path lengths of 10 or more"*

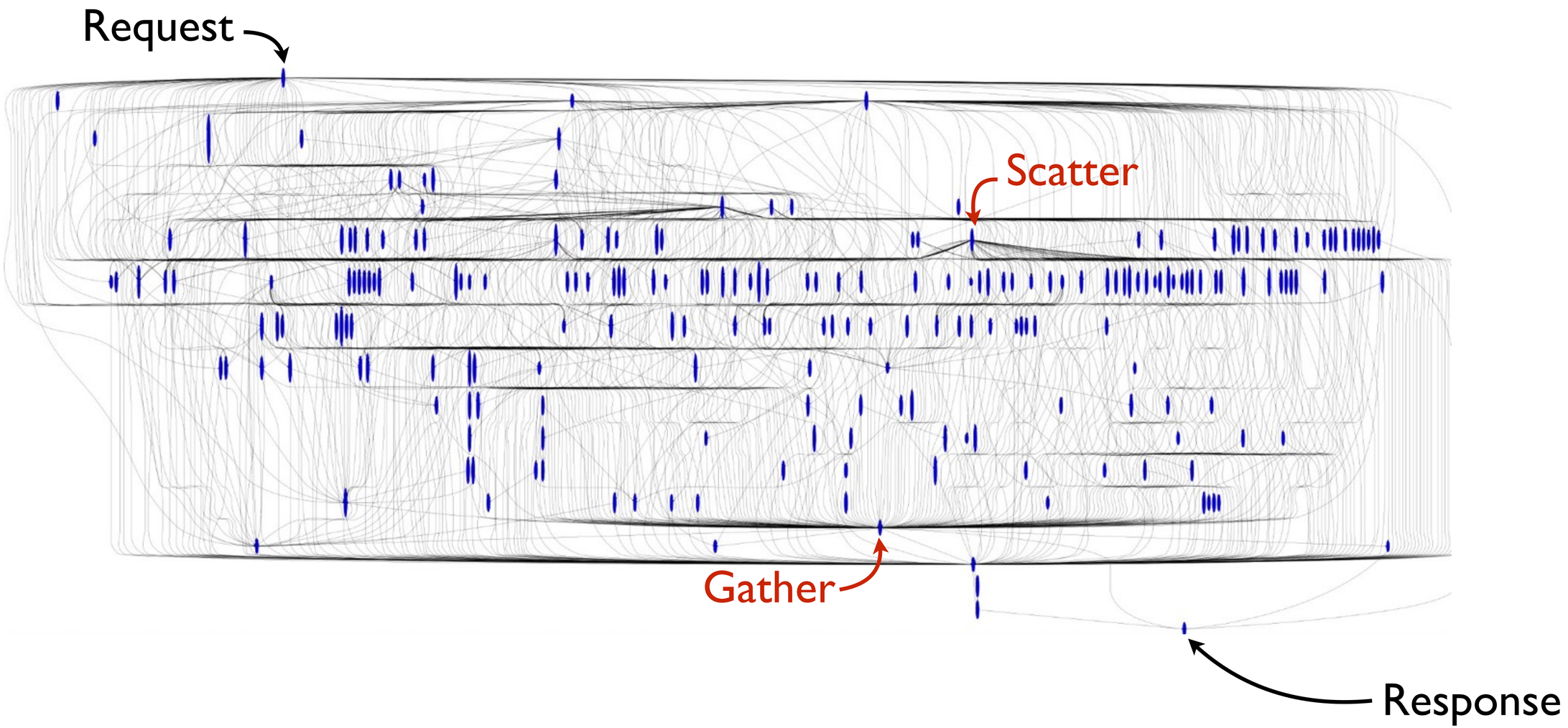



Image source: Talk on "Speeding up Distributed Request-Response Workflows"
by Virajith Jalaparti at ACM SIGCOMM'13

# Big data analytics

Hadoop

Spark

Dryad

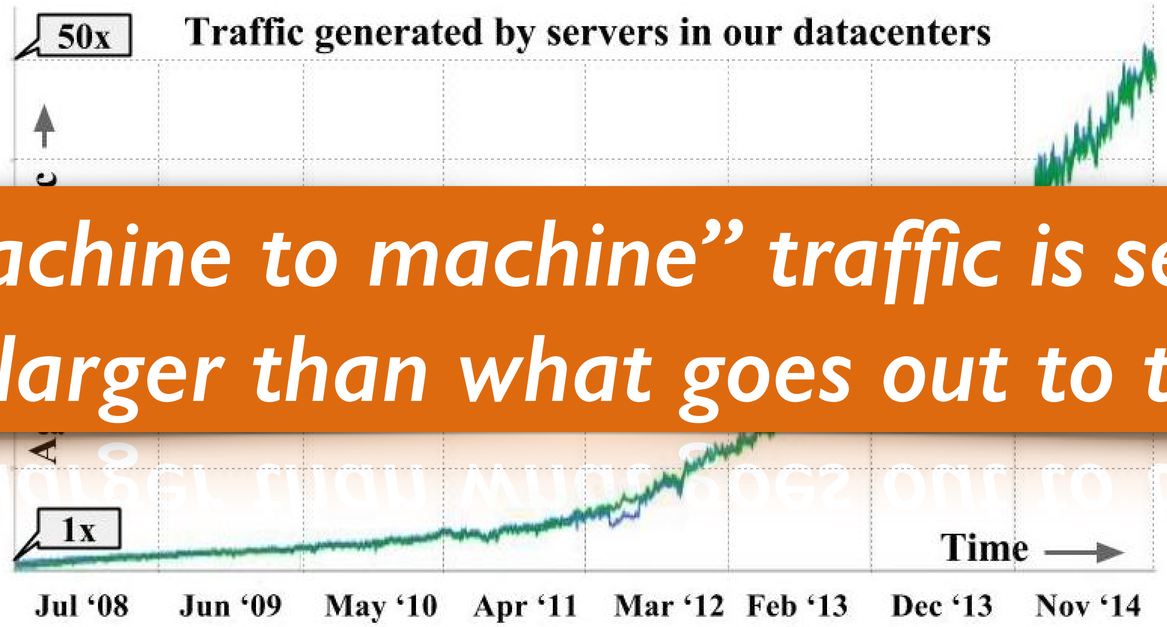
Database *joins*

⋮

# What does data center traffic look like?

It depends … on applications, scale, network design, …

# Traffic characteristics: growing volume




**Facebook: *"machine to machine" traffic is several orders of magnitude larger than what goes out to the Internet***

"Jupiter Rising: A Decade of Clos Topologies and Centralized Control in Google's Datacenter Network", Arjun Singh et al. @ **Google**, ACM SIGCOMM'15

# Traffic characteristics: rack locality

## Facebook

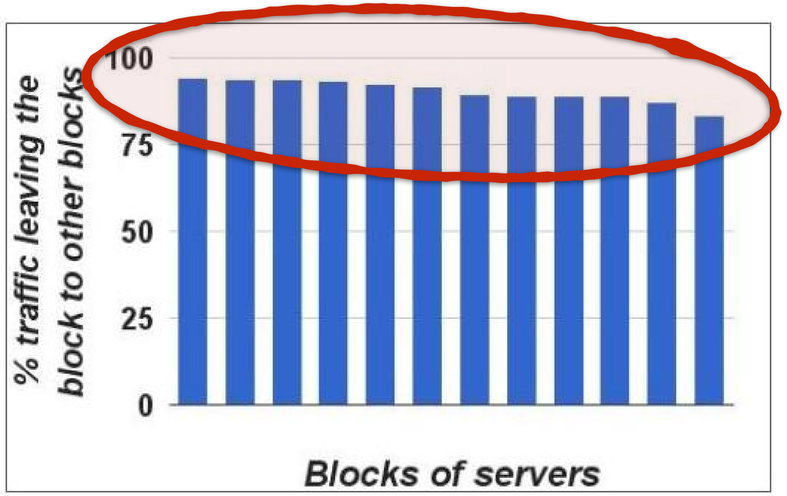
"Inside the Social Network's (Datacenter) Network"
Arjun Roy et al., ACM SIGCOMM'15

| Locality | All | Hadoop | FE | Svc. | Cache | DB |
|---|---|---|---|---|---|---|
| Rack | 12.9 | 13.3 | 2.7 | 12.1 | 0.2 | 0 |
| Cluster | 57.5 | 80.9 | 81.3 | 56.3 | 13.0 | 30.7 |
| DC | 11.9 | 3.3 | 7.3 | 15.7 | 40.7 | 34.5 |
| Inter-DC | 17.7 | 2.5 | 8.6 | 15.9 | 16.1 | 34.8 |
| Percentage | | 23.7 | 21.5 | 18.0 | 10.2 | 5.2 |

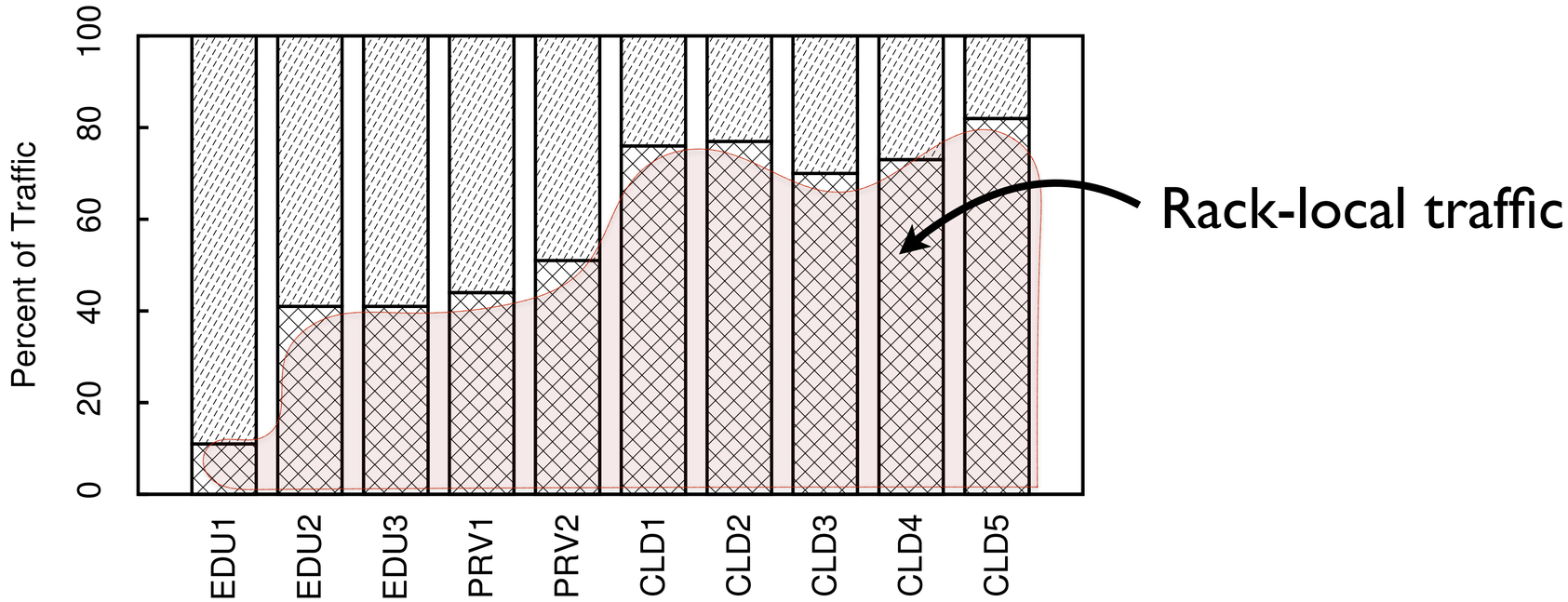
## Google

"Jupiter Rising: A Decade of Clos Topologies and Centralized Control in Google's Datacenter Network"
Arjun Singh et al., ACM SIGCOMM'15

| Job Category | B/w (%) |
|---|---|
| Storage | 49.3 |
| Search Serving | 26.2 |
| Mail | 7.4 |
| Ad Stats | 3.8 |
| Rest of traffic | 13.3 |



# Traffic characteristics: rack locality




"Network Traffic Characteristics of Data Centers in the Wild"
Theophilus Benson et al., ACM IMC'10

# Traffic characteristics: concurrent flows

## Facebook

"Inside the Social Network's (Datacenter) Network"
Arjun Roy et al., ACM SIGCOMM'15

*"Web servers and cache hosts have **100s to 1000s** of concurrent connections"*

*"Hadoop nodes have approximately **25** concurrent connections on average."*

## 1500 server cluster @ ??

"The Nature of Datacenter Traffic: Measurements & Analysis"
Srikanth Kandula et al. (Microsoft Research), ACM IMC'09

*"median numbers of correspondents for a server are two (other) servers within its rack and four servers outside the rack"*

# Traffic characteristics: flow arrival rate

**Facebook**

"Inside the Social Network's (Datacenter) Network"
Arjun Roy et al., ACM SIGCOMM'15

*"median inter-arrival times of approximately 2ms" at a server*

**1500 server cluster @ ??**

"The Nature of Datacenter Traffic: Measurements & Analysis"
Srikanth Kandula et al. (Microsoft Research), ACM IMC'09

< 0.1x Facebook's rate

# Traffic characteristics: flow sizes

## Facebook

"Inside the Social Network's (Datacenter) Network"
Arjun Roy et al., ACM SIGCOMM'15

Hadoop: median flow <1KB
<5% exceed 1MB or 100sec

Caching: most flows are long-lived
… but bursty internally

Heavy-hitters ≈ median flow, not persistent

## 1500 server cluster @ ??

"The Nature of Datacenter Traffic: Measurements & Analysis"
Srikanth Kandula et al. (Microsoft Research), ACM IMC'09

> 80% of the flows last <10sec
> 50% bytes are in flows lasting less <25sec

# What does data center traffic look like?

It depends … on applications, scale, network design, …

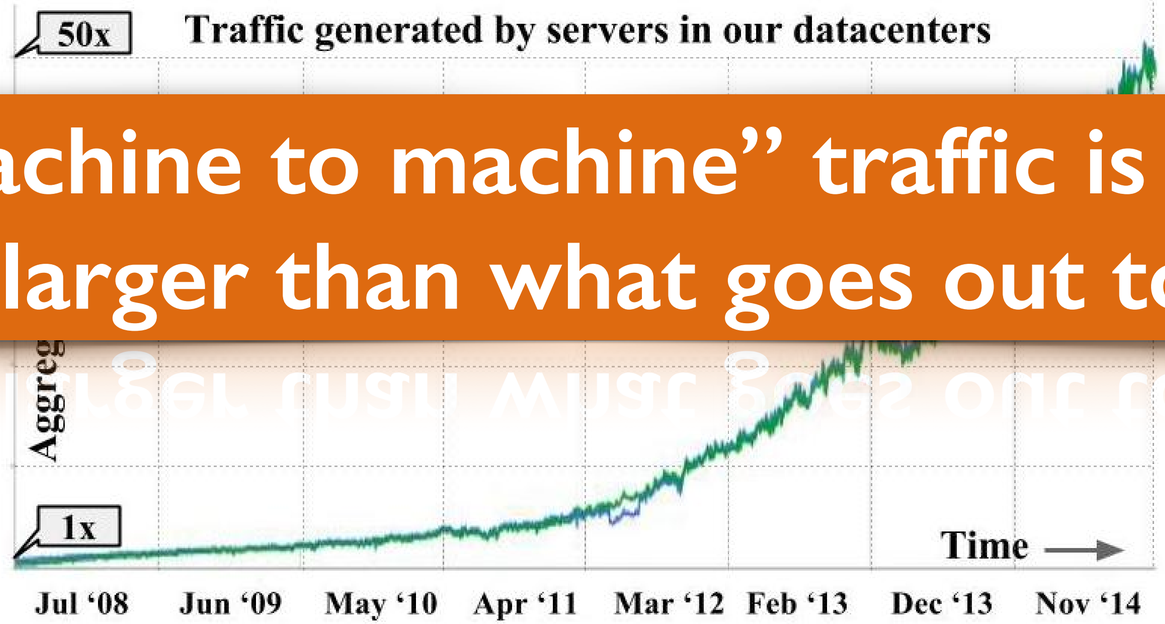
**… and right now, not a whole lot of data is available.**

# Implications for networking

1. Data center internal traffic is BIG
2. Tight deadlines for network I/O
3. Congestion and TCP incast
4. Need for isolation across applications
5. Centralized control at the flow level may be difficult

# Implications for networking

1. Data center internal traffic is BIG

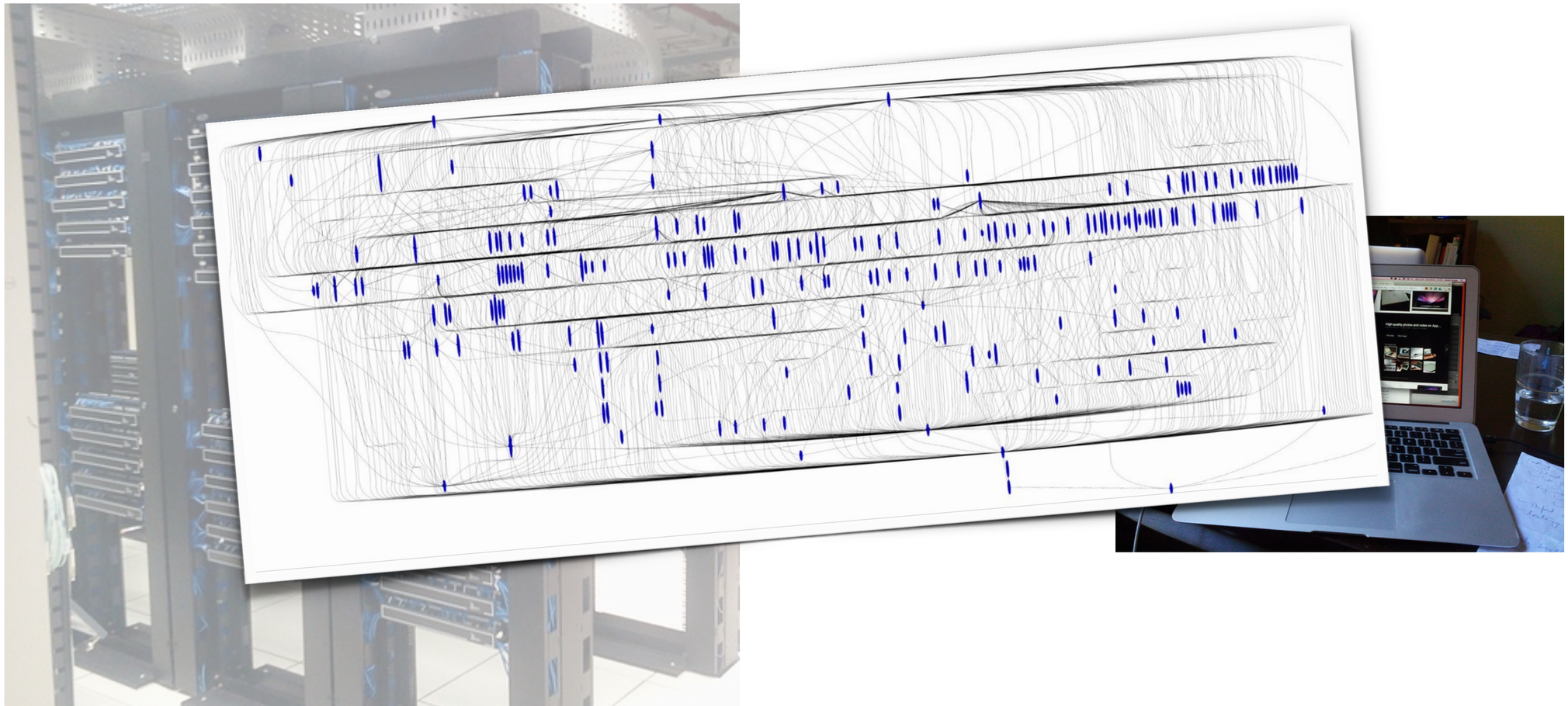



**Facebook: "machine to machine" traffic is several orders of magnitude larger than what goes out to the Internet**

"Jupiter Rising: A Decade of Clos Topologies and Centralized Control in Google's Datacenter Network", Arjun Singh et al. @ **Google**, ACM SIGCOMM'15

# Implications for networking

2. Tight deadlines for network I/O



# Implications for networking

2 Tight deadlines for network I/O

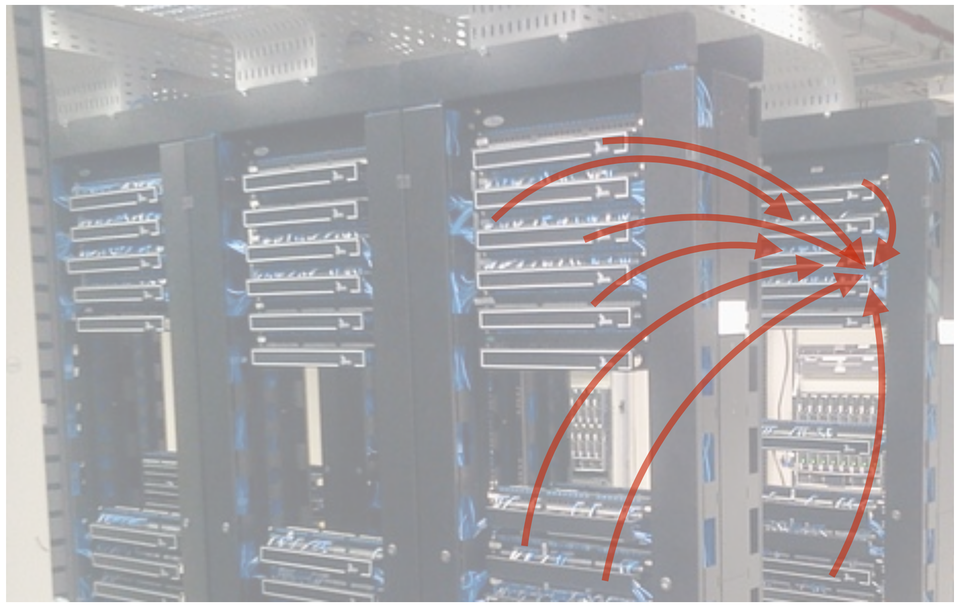
Suppose: server response-time is 10ms for 99% of requests; 1s for 1%

| #Servers | Requests 1s or slower |
|---|---|
| 1 | 1% |
| 100 | 63% |

**Need to reduce variability and tolerate *some* variation**

# Implications for networking

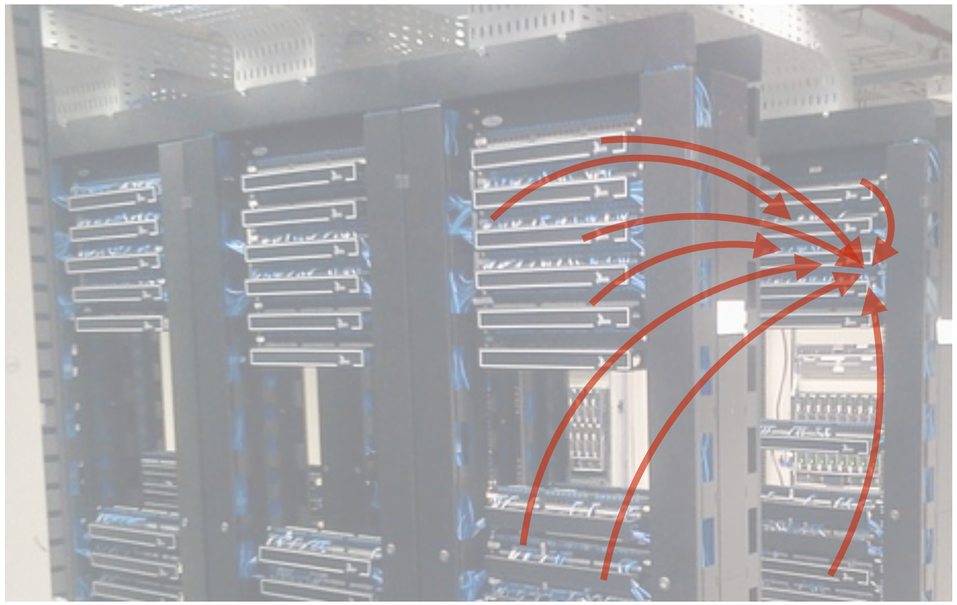
3 Congestion and TCP incast



**TCP does not work very well**

# Implications for networking

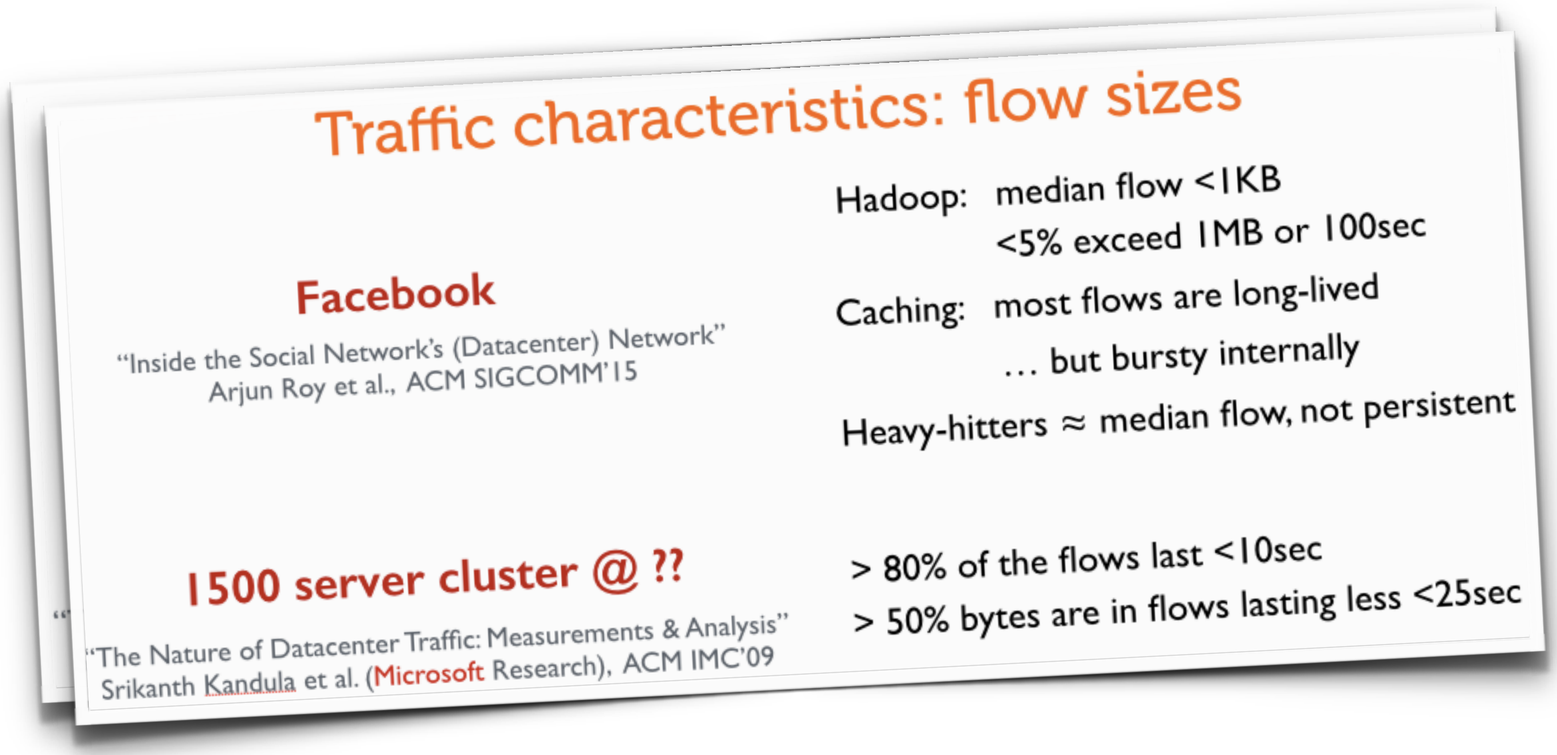
4 Need for isolation across applications



**Applications with different objectives sharing the network**

# Implications for networking

5 Centralized control at the flow level may be difficult




**Distributed control, perhaps with some centralized tinkering**