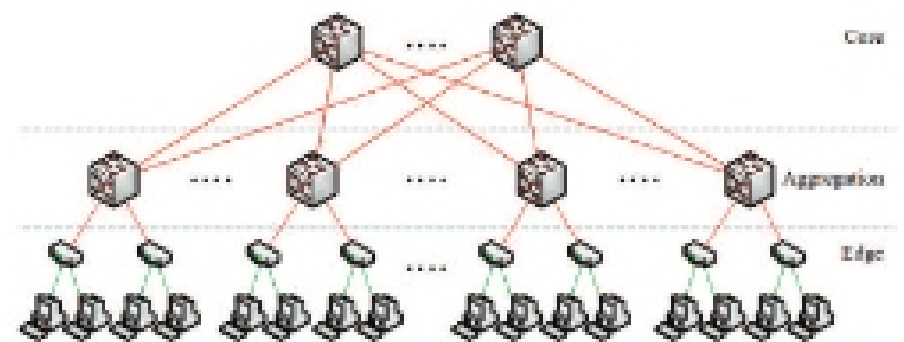


Data Center Networking

Stanford CS144 Lecture 17
Philip Lewis, 11/30/11

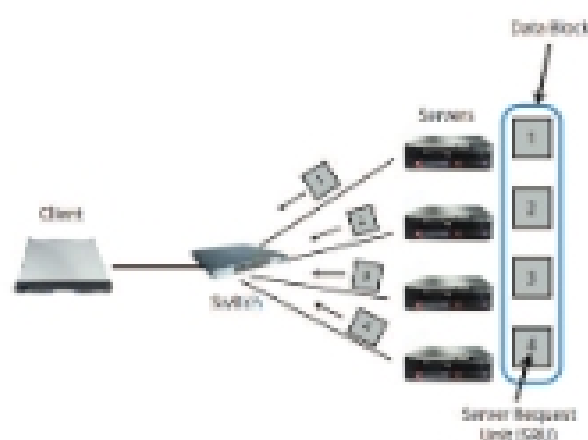


Topology



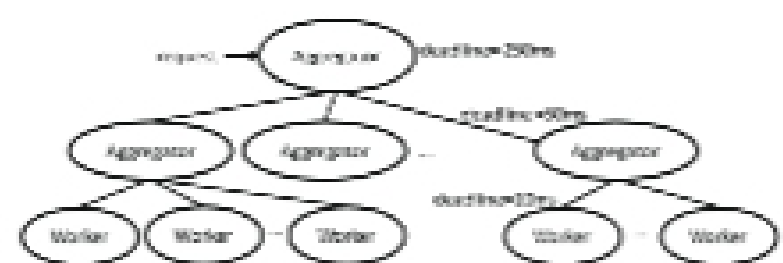
(picture courtesy of Al-Fares et al, "A Scalable, Commodity Data Center Network Architecture")

Storage Workload



(picture courtesy of Phanishayee et al, "Measurement and Analysis of TCP Throughput Collapse in Cluster-based Storage Systems")

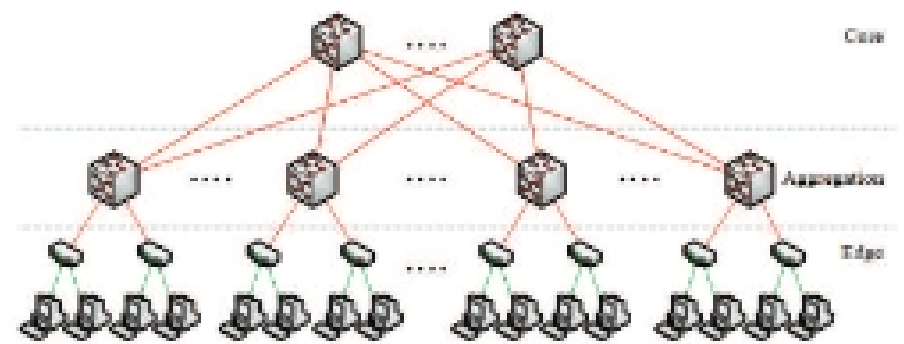
Query Workload



(picture courtesy of Alizadeh et al, "Data Center TCP (DCTCP)")

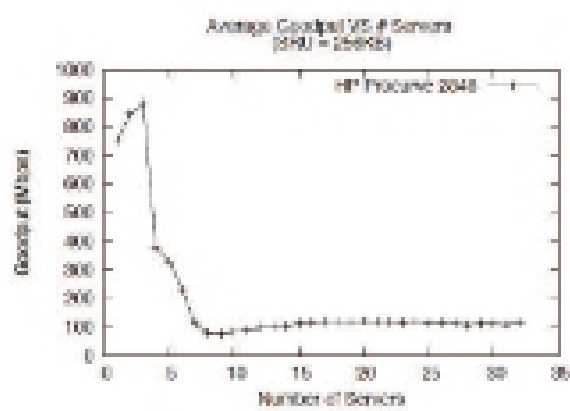
Problems

Per-Pair Bandwidth



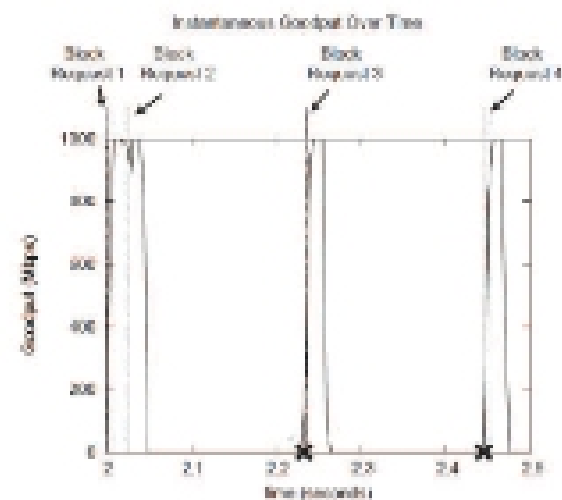
(picture courtesy of Al-Fares et al, "A Scalable, Commodity Data Center Network Architecture")

Incast



(from Phanishayee et al, "Measurement and Analysis of TCP Throughput Collapse in Cluster-based Storage Systems")

Incast Details



(from Phanishayee et al, "Measurement and Analysis of TCP Throughput Collapse in Cluster-based Storage Systems")

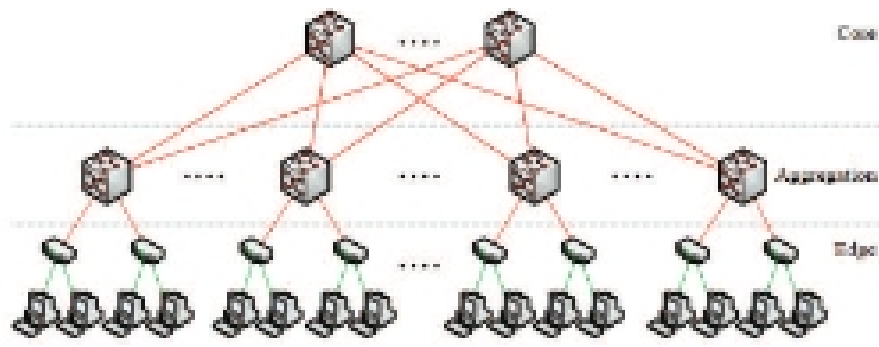
Mixed traffic

- Low latency for short flows
- High burst tolerance (incast)
- High throughput for long flows

Recent Research

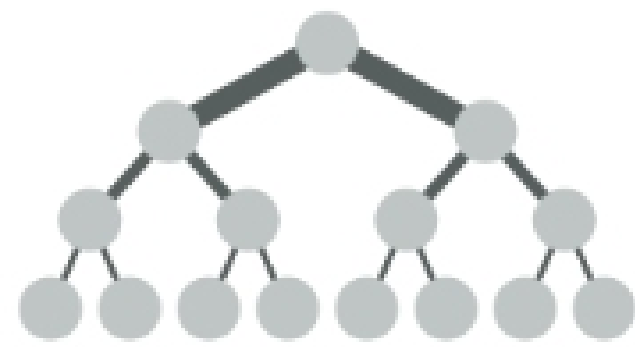
- New switching topology: Al-Fares et al.
- Fix TCP incast: Vasudevan et al.
- Data Center TCP: Alizadeh et al.

Per-Pair Bandwidth

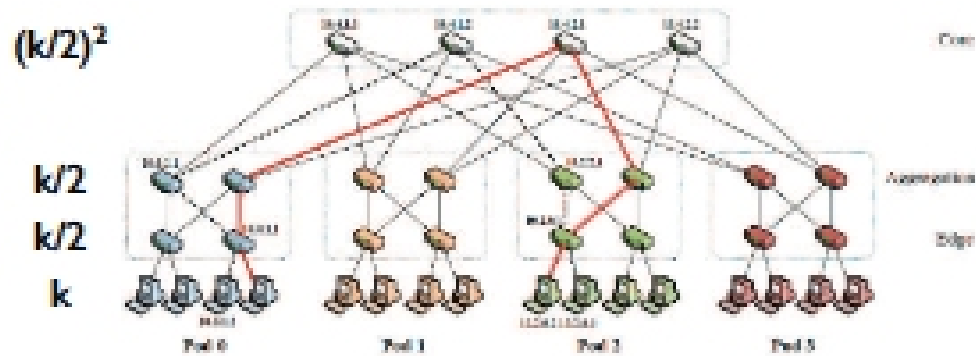


(picture courtesy of Al-Fares et al, "A Scalable, Commodity Data Center Network Architecture")

Fat Tree



Fat Tree



Switching

Prefix	Port
10.2.0.0/24	0
10.2.1.0/24	1
0.0.0.0/0	

Suffix	Port
0.0.0.2/8	2
0.0.0.3/8	3

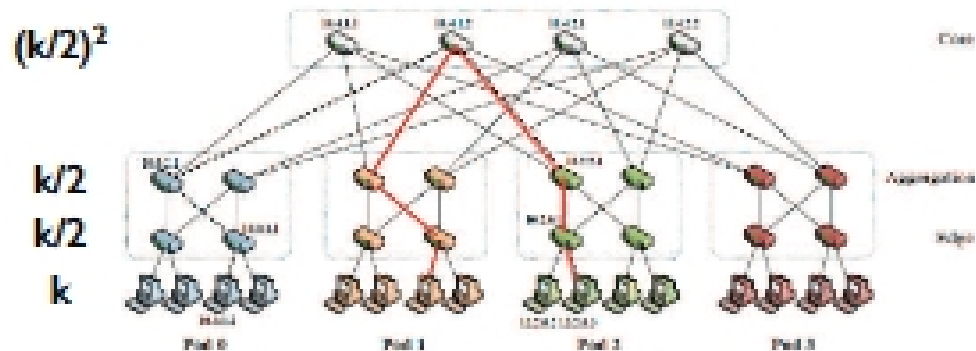
TCAM

10.2.0.X
10.2.1.X
X.X.X.2
X.X.X.3

Encoder

Prefix	Next Hop	Port
00	10.2.0.1	0
01	10.2.1.1	1
10	10.4.1.1	2
11	10.4.1.2	3

Not Perfect



Fat-Tree Status

