# A Server-to-Server View of the Internet











Bala

Georgios

Arthur

Matthew

KC



Content moved closer to end users to reduce latency.

Connections from end users are terminated at CDN servers close to the end users.



Viewing the Internet's core from the distributed measurement platform of a CDN.



Back-Office Web traffic accounts for a significant fraction of core Internet traffic — *Pujol et al., IMC, Nov. 2014*.

End-user experience is at the mercy of the unreliable Internet and its *middle-mile* bottlenecks — *T. Leighton, CACM, Vol. 52. No. 2, Feb. 2009.* 



A six-month timeline of RTTs between servers in Honk Kong, HK and Tokyo, JP



A six-month timeline of RTTs between servers in Honk Kong, HK and Tokyo, JP



A six-month timeline of RTTs between servers in Honk Kong, HK and Tokyo, JP

Level-shifts in RTTs over both IPv4 and IPv6



To what extent do changes in the AS path affect round-trip times?



A portion of the timeline of RTTs between servers in Honk Kong, HK and Tokyo, JP.

Daily oscillations in RTT between the servers.



How common are periods of daily oscillation in RTT, and where do they occur?



What affects end-to-end RTTs more – routing or congestion?



How does IPv4 and IPv6 compare with respect to routing and performance?

1. To what extent do changes in the AS path affect round-trip times?

2. How common are periods of daily oscillation in RTT, and where do they occur?

# Effect of *routing changes* on endto-end RTTs

# Data Set: Long Term

- $\approx 600$  dual-stacked servers in 70 different countries.
  - US, AU, DE, IN, JP, ...



Traceroutes conducted between servers in both directions over both protocols.



Every **3** hours traceroutes done over the *full-mesh*. All traceroutes in a given **3** hour time frame have the same timestamp.



Traceroutes over the *full-mesh* every **3** hours for **16** months from Jan. 2014 through Apr. 2015.

≈700M IPv4 and ≈600M IPv6 traceroutes

Trace timeline  $S_a \rightarrow S_b$  is different from  $S_b \rightarrow S_a$ 



- Extract two pieces of information from each traceroute
  - **AS path** inferred from interfaces in the traceroute output
  - end-to-end RTT between the two servers



#### A–B trace timeline

(AS-path, end-to-end RTT) tuples spanning the study period



**Popular** AS path observed in A–B trace timeline **AS<sub>1</sub>-AS<sub>2</sub>-AS<sub>3</sub>-AS<sub>4</sub>** with prevalence **60%** 



AS path *prevalence* — Vern Paxson, IEEE/ACM Transactions on Networking 1997



Most paths had one dominant route, with 80% dominant for at least half the period.



*Number of AS-path changes* observed in the A–B trace timeline



80% of the trace timelines experienced 20 or fewer changes over the course of 16-months.

How do the AS-path changes affect the *baseline* RTT of server-to-server paths?



#### Group RTTs by AS paths. Baseline: 10<sup>th</sup>-percentile of each AS-path (bucket).



Optimal Path: path with lowest baseline. Optimal: AS1-AS5-AS9-AS4 Sub-Optimal: AS1-AS2-AS3-AS4



Baseline of sub-optimal path with prevalence of 60% is ~4.5 ms increase in end-to-end RTT.



Typically a routing change causes only a small change in RTT.



But for a minority of cases, the change can be significant. 10% of trace timelines over IPv4 the (sub-optimal) AS paths that led to at least a 20 ms increase in RTTs had a prevalence of at least 30%

# Effect of *periods of daily oscillation* on end-to-end RTTs

## Data Set: Short Term

≈3,500 server clusters in 1,000 locations in 100 different countries.



*ping* measurements every *15 minutes* for one week from Feb. 22, 2015 through Feb. 28, 2015.

 $\approx$ **2.9M** IPv4 and  $\approx$ **1M** IPv6 server pairs

Based on Time Sequence Latency Probes by Luckie et al., IMC 2014





# Identify first *segment* with high-correlation with end-to-end RTT?

# Highlights

3155 links were congested in our study of IPv4 traceroutes.1768 internal & 1121 interconnection links.

Weighting links by the number of server-to-server paths that cross them ...

interconnection links are more popular!

Large majority of the interconnection links with congestion were *private interconnects*.



Typical overhead due to congestion is *20-30 ms*.



Values between *20-30 ms* — *US:* accounts for *90%* of density. *Europe & Asia:* accounts for *30%* of density.



Transcontinental links in *Europe* & *Asia*.

Routing changes typically do not affect end-to-end RTTs. Congestion is not the norm.

# What about *non-typical* cases?

### Routing

For 10% of server pairs the (sub-optimal) AS paths that led to 20 ms increase in RTTs pertained for at least 30% of the study period for IPv4 & 50% for IPv6.

### Congestion

Only 2% of the server pairs over IPv4, and just 0.6% over IPv6, experience a strong diurnal pattern with an increase in RTT of least 10 ms.

## Routing

10% of trace timelines the (sub-optimal) AS paths that led to at least
20 ms increase in RTTs pertained for at least
30% of the study period for IPv4 & 50% for IPv6.

### Congestion

Only pairs over IPv4, and just 0.6% experience a strong diurnal pattern with an increase in RTT of least 10 ms



- Focus on bandwidth
- No packet loss measurements; platform limitations
- Explore IPv4 & IPv6 infrastructure sharing







Use measurements over paths between CDN servers to understand the state of the Internet core.



*Number of unique AS paths* observed in the A–B trace timeline

 $AS_1 - AS_2 - AS_3 - AS_4$  and  $AS_1 - AS_5 - AS_9 - AS_4$ 



**80%** of trace timelines have **5** or fewer AS paths in IPv4, and **6** or fewer in IPv6.



**80%** of trace timelines have <u>5</u> or fewer AS paths in IPv4, and 6 or fewer in IPv6.



# **80%** of trace timelines have **5** or fewer AS paths in IPv4, and <u>**6** or fewer in IPv6</u>.



Combine AS paths observed in the forward direction with



Combine AS paths observed in the forward direction with those in the reverse direction.



Pairing AS paths in the forward & reverse directions still reveals 80% of server pairs to have 8 or fewer path pairs in IPv4, and 9 or fewer in IPv6.



Pairing AS paths in the forward & reverse directions still reveals *80%* of server pairs to have <u>8 Or fewer</u> path pairs in IPv4, and 9 or fewer in IPv6.



Pairing AS paths in the forward & reverse directions still reveals *80%* of server pairs to have *8* or fewer path pairs in IPv4, and *9* or fewer in IPv6.



[18.0h, 3.9D] [7.6, 15.7M) [6.0, 18.0h) [3.9, 15.9D) 1.0M) [2.0, 4.0M) [4.0, 7.6M) [3.0, 6.0h) [1.0,2.0M) [15.9D' 0.2' [48.3ms, 2.55), 2.80010 6. 0.81 0.76 0.70 0.59 0.31  $\hat{\mathbf{x}}$ [26.1, 48.3ms) 2.40010 5. 33 2.72 0.88 0.98 0.69 0.49 from best AS-path [17.3,26.1ms) 2.000% 1.09 0.95 0.98 0.86 3. 0.91 0.69 .2ª [12.3, 17.3ms] 1,60010 1.63 25 1.07 0.97 1.02 1.02 0.85 <u>\_</u>01 [8.8, 12.3ms) 1.48 0 1.04 1.10 1.08 **Difference in RTT** . 0 , 0<sup>0</sup> 'ن 1.20010 [6.1,8.8ms) 203 0.96 0.98 20  $\sqrt{0}^{0}$ [4.0, 6.1ms) 0.0000 00 0.89 0.92 0.94 ્રે [2.2, 4.0ms) 0.40010 0.94 0.19 0.00 000  $\tilde{\mathcal{Y}}$ 30 [0.8,2.2ms) 0.000% 0.70 18 .22 0<sup>,9</sup> N, ્રે [0.0, 0.8ms) **AS-path lifetime** 

Comparing magnitudes of increase in (baseline) 10<sup>th</sup> percentile of RTTs of AS paths (each relative to the best AS path of the corresponding trace timeline) with the lifetime of AS paths ...

[7.6, 15.7M] [6.0, 18.0h] 1.0M [3.9, 15.90 3.90) [4.0, 7.6M] [2.0, 4.0M) [3.0, 6.0h] [1.0, 2.0M) [18.0h, [15.90, 20010 1.49 0.45 0.2' [48.3ms, 2.55], 2.80010 ~°? 0.81 0.16 0.59 0.70 0.37 <u>,</u>6 <u>`</u>?` [26.1, 48.3ms) 2.400% 2.2 0.88 0.88 0.69 0.49 رب<sup>5</sup> 0.84 ્ઝે from best AS-path [17.3,26.1ms) 2.000% 0.95 1.09 0.98 0.86 0.69 80 0.91 , <sup>39</sup> .20 [12.3, 17.3ms) 1,60010 6. 25 2.72 1.07 0.97 1.02 1.02 0.85 <u>,</u>01 [8.8, 12.3ms) 1.48 1.10 1.08 Difference in RTT  $\sqrt{0^{1}}$ 1.04 . 0A , 0A 'ن 1.20010 [6.1,8.8ms) 22 ~03 0.96 1.20 0.98 ,06 [4.0, 6.1ms) • 0.90<sup>0/0</sup> 0. 0.89 0.92 0.94 3. [2.2, 4.0ms) 0.40010 0.94 0.19 0.00 08. رکی ્ઝુ [0.8,2.2ms) 0.000% 0.70 18 .22 0<sup>90</sup> <u>`</u>?` [0.0, 0.8ms) **AS-path lifetime** 

*X-axis*: deciles of the distribution of AS-path lifetimes. half-open intervals [0.0, 3.0h) has no data points Same value for 0<sup>th</sup>% and 10<sup>th</sup>% of the AS-path lifetime distribution

[18.0h, 3.9D] [15.9D, 1.0M) [7.6, 15.7M) [6.0, 18.0h) [3.9, 15.9D) [2.0, 4.0M) [4.0, 7.6M) [3.0, 6.0h) [1.0,2.0M) 50010 1.49 0.45 0.2' [48.3ms, 2.55] 2.80010 1.31 103 0.76 0.81 0.70 0.59 0.31 <u>,</u>6 [26.1, 48.3ms, 2.40010 5. 2.2 0.88 0.98 0.69 0.49 0.84 ્ઝે from best AS-path [17.3,26.1ms] 2.000% 1.09 0.95 3. 0.98 0.86 0.69 0.91 .20 [12.3, 17.3ms] 1,60010 co. 25 1.07 0.97 1.02 0.85 1.07 1.02 [8.8, 12.3ms] 1.04 1.10 **Difference in RTT** 1.08  $\sqrt{0^{1}}$ 1.04 . 0 `ڧ 1.20010 [6.1,8.8ms) 03 0.96 0.98 20  $\sqrt{0}^{6}$ [4.0, 6.1ms] • 0.90<sup>0/0</sup> 0.89 0.92 0.94 3. [2.2, 4.0ms] 0.40010 0.19 0.00 0.81  $\tilde{\mathcal{Y}}$ ્રેંગ [0.8,2.2ms] 0.000% 0.70 18 .22 42 0<sup>,9</sup> ્રે [0.0, 0.8ms] **AS-path lifetime** 

*Y-axis*: deciles of the distribution of magnitudes of increase in 10th percentile of RTTs of AS paths (each relative to the best AS path of the corresponding trace timeline).

[18.0h, 3.9D] [15.9D, 1.0M) [7.6, 15.7M) [6.0, 18.0h) [39,15.9D) [2.0, 4.0M) [4.0, 7.6M) [3.0, 6.0h) [1.0, 2.0M) 50010 1.49 0.45 29 0.21 [48.3ms, 2.55], 2.80010 ~°? 0.76 3. 0.81 0.70 0.59 0.31 <u>,</u>6 [26.1, 48.3ms) 2.40010 2.72 0.88 0.88 0.69 0.49 رب<sup>5</sup> 0.84 ્ઝે from best AS-path [17.3,26.1ms) 2.000% 1.09 0.95 0.86 3. 0.98 0.69 0.91 .20 [12.3, 17.3ms) 1,60010 6. 25 2.72 1.07 0.97 1.02 1.02 1.07 0.85 [8.8, 12.3MS) Difference in RTT 1.48 1.08 04 , 0Å <u>,</u>01 <u>^</u>0 , 0<sup>0</sup> Ò 1.20010 [6.1, 8.8ms) 27 ~°? 0.96 0.98 20 ,06 [4.0, 6.1ms) 0.80010 0.89 0.00 0.92 0.94 ્રે [2.2, 4.0ms) 0.40010 0.94 0.19 0.83 08.  $\hat{\gamma}$ ઝે [0.8,2.2ms) 0.000% 0.70 18 0<sup>90</sup>  $\mathcal{N}$ ્રે [0.0, 0.8ms) **AS-path lifetime** 

Baseline RTTs of AS paths with longer lifetimes are close in value to that of the best AS path of corresponding trace timelines.

[18.0h, 3.9D] [15.9D, 1.0M) [7.6, 15.7M) [6.0, 18.0h) [3.9, 15.9D) [4.0, 7.6M) [3.0, 6.0h) [2.0, 4.0M) [1.0,2.0M) 3.20010 0.45 0.58 0.74 0.27 0.1 [48.3ms, 2.55] 2.80010 30. 0.59 0.81 0.76 0.70 0.31 ્ર્હ [26.1, 48.3ms) 2.40010 0.69 0.49 N. 0.88 0.88 0.84 from best AS-path ŝ. Ś [17.3,26.1ms) 2.00010 0.95 0.98 0.86 0.69 1.09 0.91 20 [12.3, 17.3ms) 1,600lo 107 1.02 1.63 25 1.72 0.97 1.07 1.02 0.85 [8.8, 12.3ms) 1.48 1.07 1.10 1.08 1.04 1.04 1.3 1.02 Difference in RTT . 0A 1.20010 [6.1,8.8ms) 0.96 0.98 22 1.03 1.06 1.20 ~3 1.18 Ŷ [4.0, 6.1ms) • 0.00<sup>0</sup>0 0.09 0.89 0.94 0.92 1.18 A <u>`</u>?` [2.2, 4.0ms) 0.40010 0.94 0.19 0.83 0.87 3. 1.23 66 N. [0.8,2.2ms] 0.000% 0.58 0.70 18 22 42 0<sup>90</sup> ્રેટે Or [0.0, 0.8ms) **AS-path lifetime** 

Paths with poor-performance are often those with relatively short lifetimes.

[17.0,27.20] [21.0h, 5.2D) [27.20, 1.7M) [5.3, 14.6M) [6.0,21.0h] [5.2, 17.00) [3.0, 6.0h) [2.9, 5.3M] [1.7,2.9M) 2.400% 0.46 0.86 0.39 0.83 0.59 1.50 [59.0ms, 1.95) 0.65 0.95 0.88 0.13 2.10010 1.10 0.48 20 [31.1,59.0ms) 0.69 1.00 0.96 0.93 0.59 0.18 from best AS-path 1,80010 N. [19.4,31.1ms) 0.80 0.97 0.94 0.80 0.72 5. 1.50010 [13.1, 19.4ms) 1.06 1.16 3.5 1.02 1.00 0.89 0.91 0.81 [9.1, 13.1ms) 1.20010 1.29 1.04 52 1.09 1.02 1.04 1.03 1.00 Difference in RTT 0.96 [6.3,9.1ms) < 0.90<sup>0/0</sup> 1.2 0.98 0.97 3. 1.05 1.06 10 1.16 [4.1, 6.3MS) • 0.60<sup>0/0</sup> 0.93 N. 0.97 0.88 30 1.09 1.04 22 N.X. [2.2, A. 1ms) 0.30010 0.85 0.86 0.97 0.77 1.04 10 1.40 ىكى [0.7,2.2ms) 0.000% 0.76 0.19 0.61 0.67 18 A <u>, 0</u>0 60 [0.0, 0.7ms) AS-path lifetime

#### Similar observations from IPv6 traceroutes.