renesys

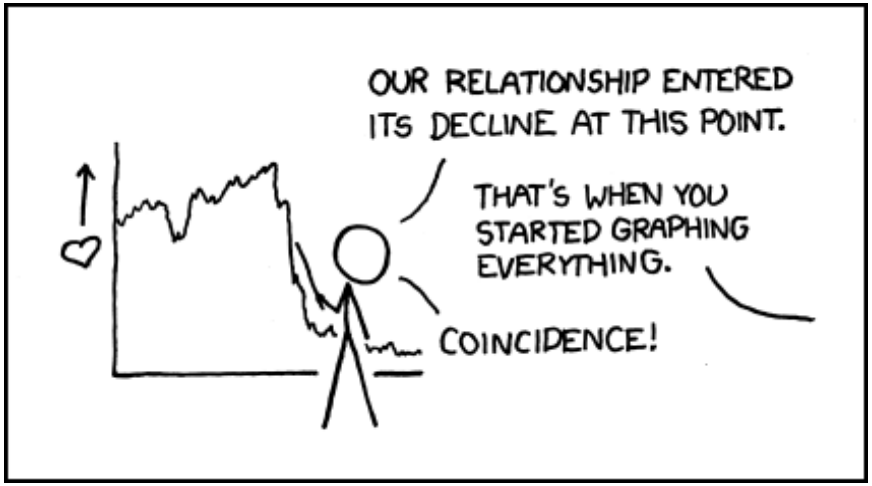
Practical Instability Scoring

Jim Cowie

NANOG 45
January 25-27, 2009

Outline: Practical Instability Scoring

- Practical Reasons to Measure the Stability of your Customer Cone
- Measuring the Problem
  - By transit ASN
  - By geographic region
- Surprising Results
- Fear, Recrimination, Blame
- Absolution



http://xkcd.com/523/

Motivation

- Every time a route changes, there's the potential for packet delay, reordering, or loss
- If route changes are significant, users somewhere are noticing the impact, running traceroutes, blaming transit providers
- If route changes are **really significant**, your routers experience increased load and sessions may reset
- If route changes are **really, really significant,** we all get to discover what a power-law size distribution means for internet outage events

Shift in perception: risk assessment

- Instead of focusing on BGP update rates, let's look for prefixes that have **consistently poor stability** (over long periods of time).
- The most unstable ~1% of the table generates 50%+ of the BGP update traffic each day.
- Stability is a surrogate measurement for those qualities you look for in a customer/peer/partner: good infrastructure, clueful admins, quiet enjoyment of the relationship.

Shoulders of Giants

- Geoff Huston
- Nick Feamster
- Jennifer Rexford
- Lixia Zhang
- Craig Labovitz
- Dan Massey
- Feng Wang
- Lixin Gao
- ...many others

Sources of Route Change

- Links die, boxes die, admins make mistakes
- Sessions go up and down and up and down
- One remote announcement or withdrawal can be perceived as 10-100 “echoes” across local peerset
- “Path exploration” multiplies observed change [Labovitz et al 2000]
- Dampening persistent instability near the origin can isolate instability, but can seriously prolong convergence [Mao et al 2002]

Practical Instability Scoring Thesis

Given enough observations from enough sources over a long enough time, we can identify specific ASNs and geographic regions that contribute significantly to global instability.

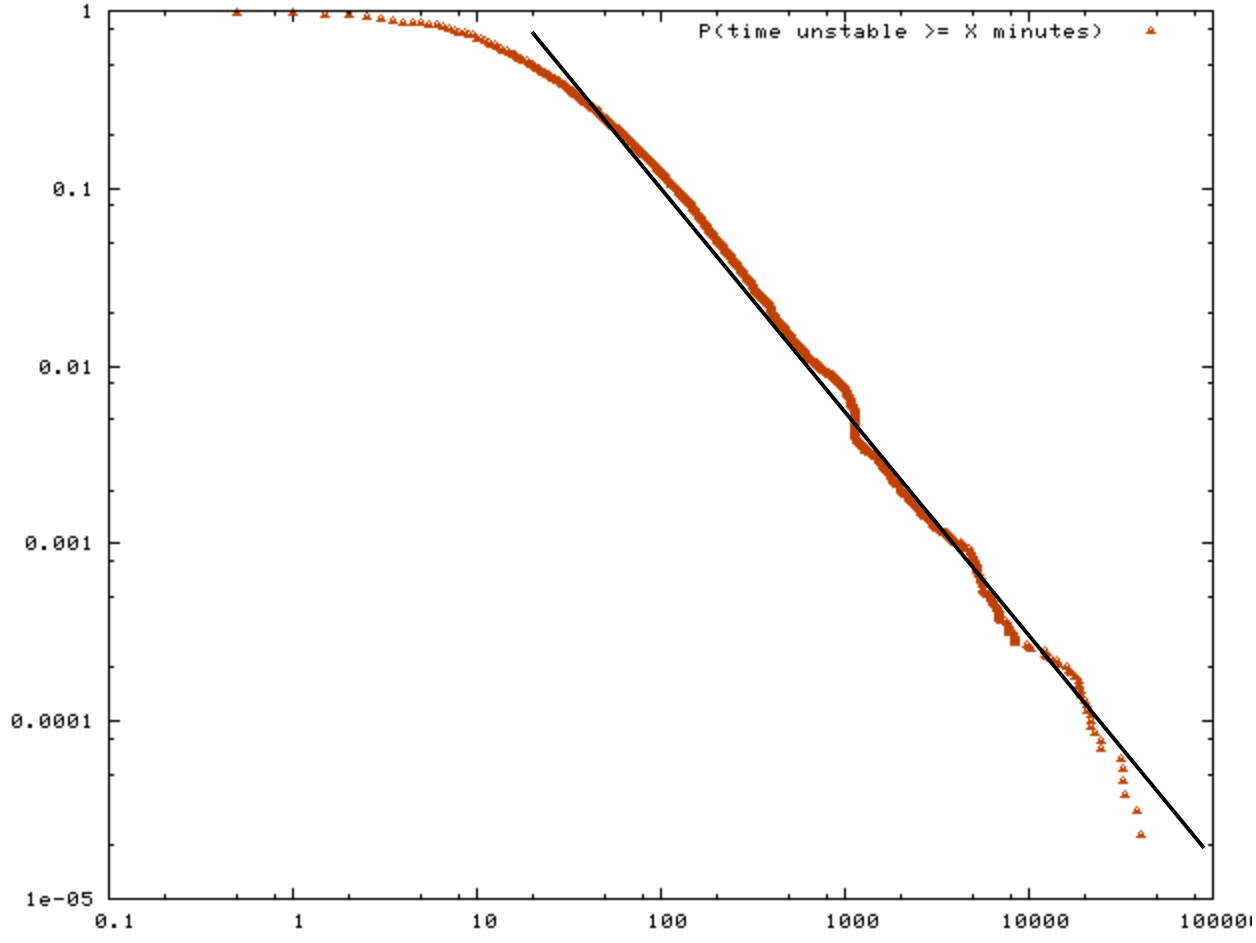
We can assign scores that assess the risk that prefixes transited by a given ASN, or originating in a particular geographic region, will experience “significant” route instability in future periods.

You can use these for bragging rights, to beat up competitors, or as the basis for new kinds of exotic derivative securities. (ok, pls don't do that)

Why Stability Scoring is Difficult

- Instability can originate at the edge, in the core, or anywhere in between
- Different vantage points can report very different experiences during the same event
- Every ASN along the path is potentially at fault
- Everyone has route stability issues sometimes
- Normalization is a nightmare
- The more prefixes you handle, the more instability you are bound to witness and/or contribute to

Many prefixes are unstable very often



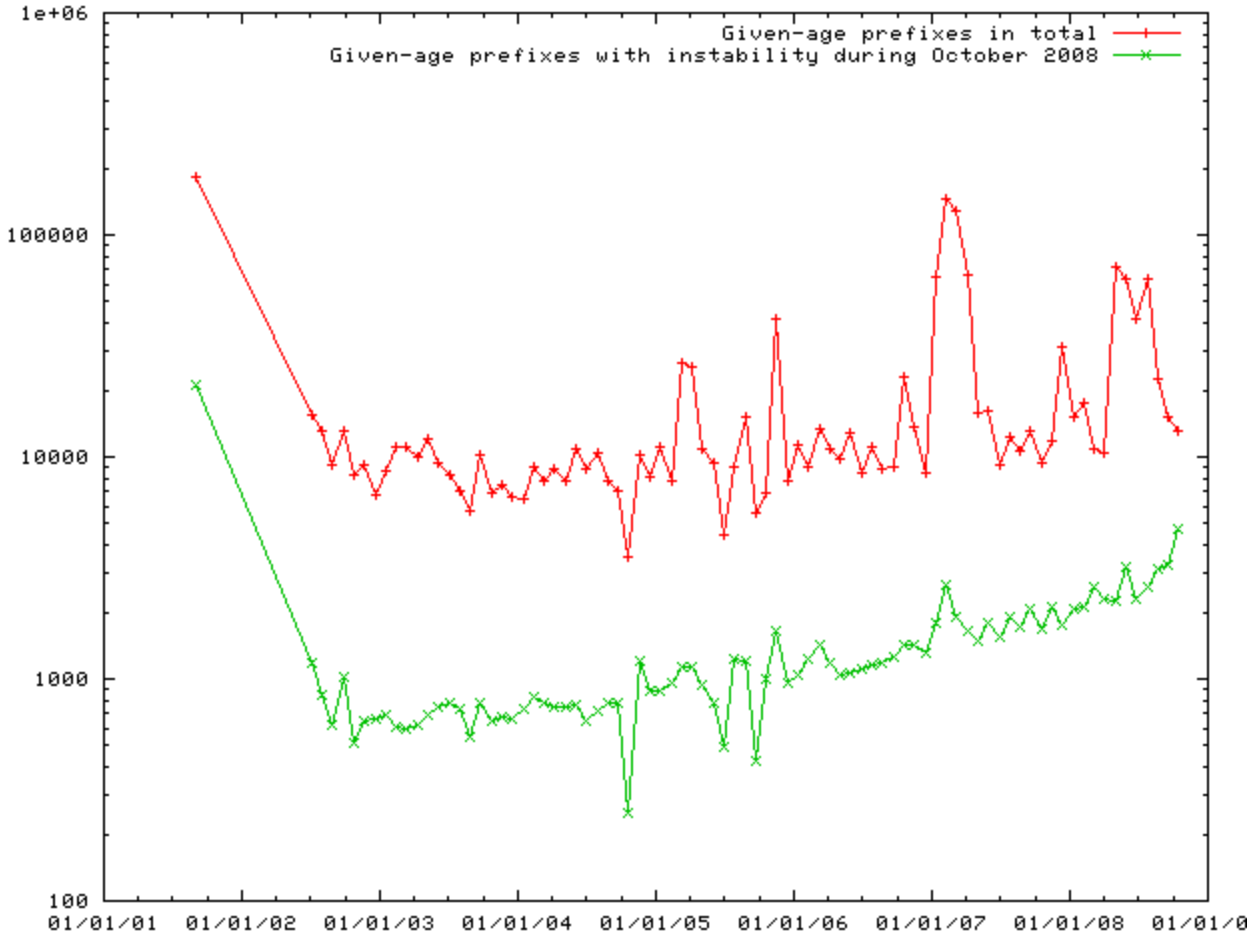
In October 2008, of 129,456 prefixes with at least mild instability:

- **70% >=10m total**
- **21% >= 1h**
- **10% >= 2h**
- **1.6% >= 8h**
- **0.3% >= 24h**
- **0.03% (34) >= 1w**

Who is most consistently bad?

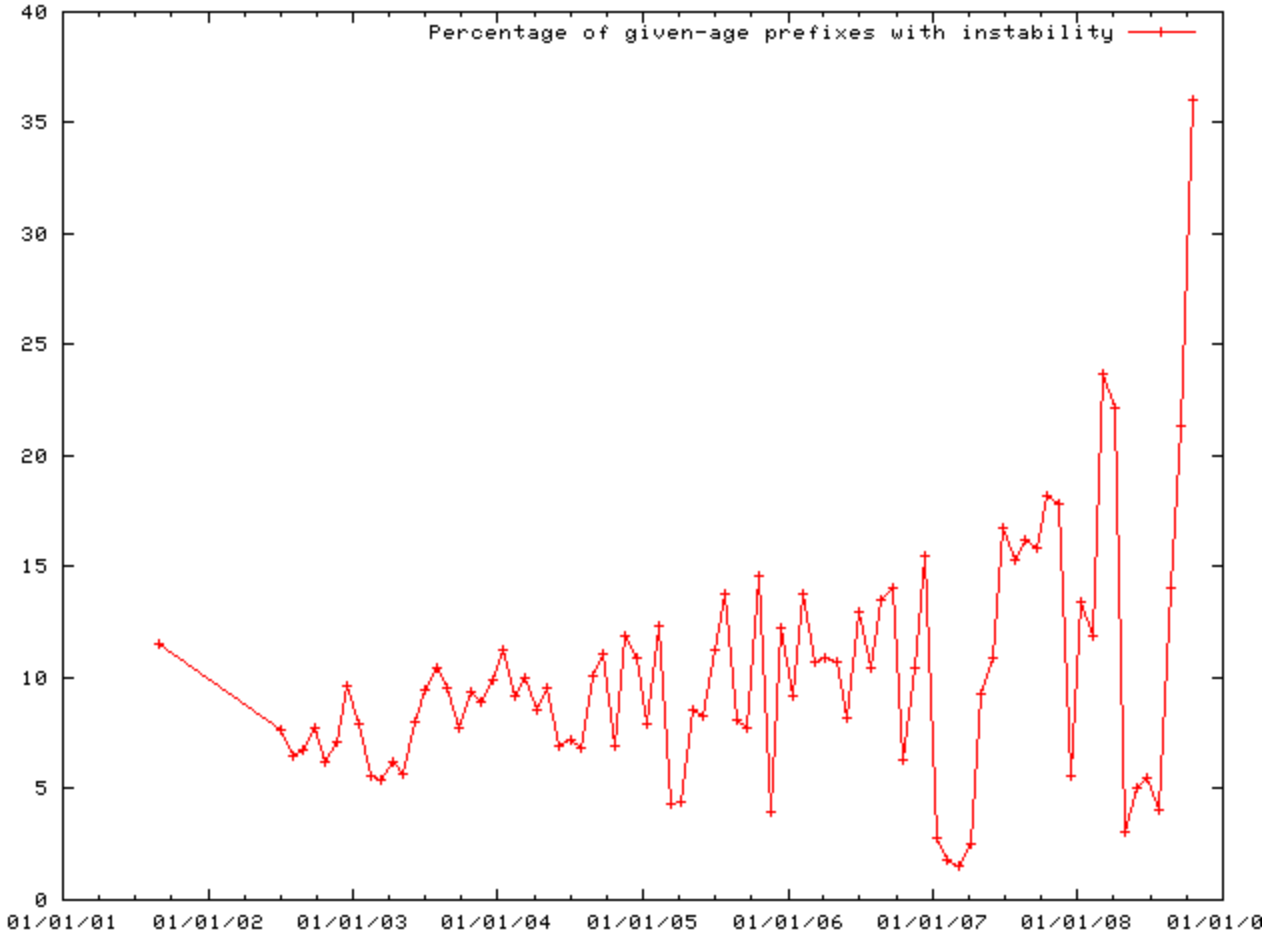
- Examined October 2008
  - 49% of the table was briefly unstable (30s+)
  - 129K of 263K prefixes impacted, very broad
  - First step: identify unstable subpopulations
- Theories included:
  - Age (Does survival favor the stable?)
  - CIDR length (Do really big prefixes flap less?)
  - Geography (Country of origin?)
  - Transit (Blame NANOG?)

Does age imply stability?



No, “elderly” prefixes that have been seen for years suffer mild instability at roughly the same levels as younger prefixes.

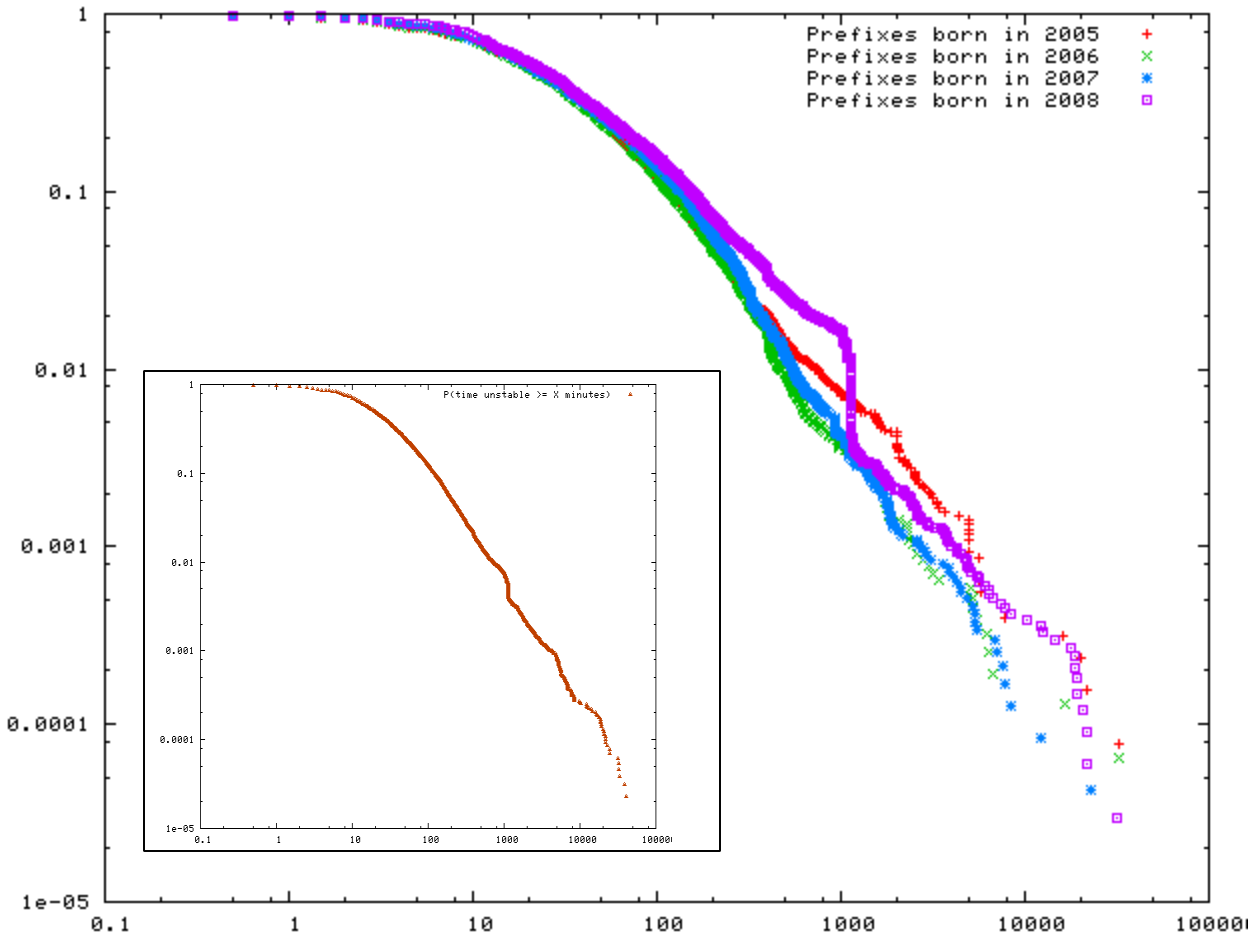
Stability increases very slowly with age.



The exception: the very young and foolish.

Prefixes younger than 30 days have 3x more mild instability (as you might expect).

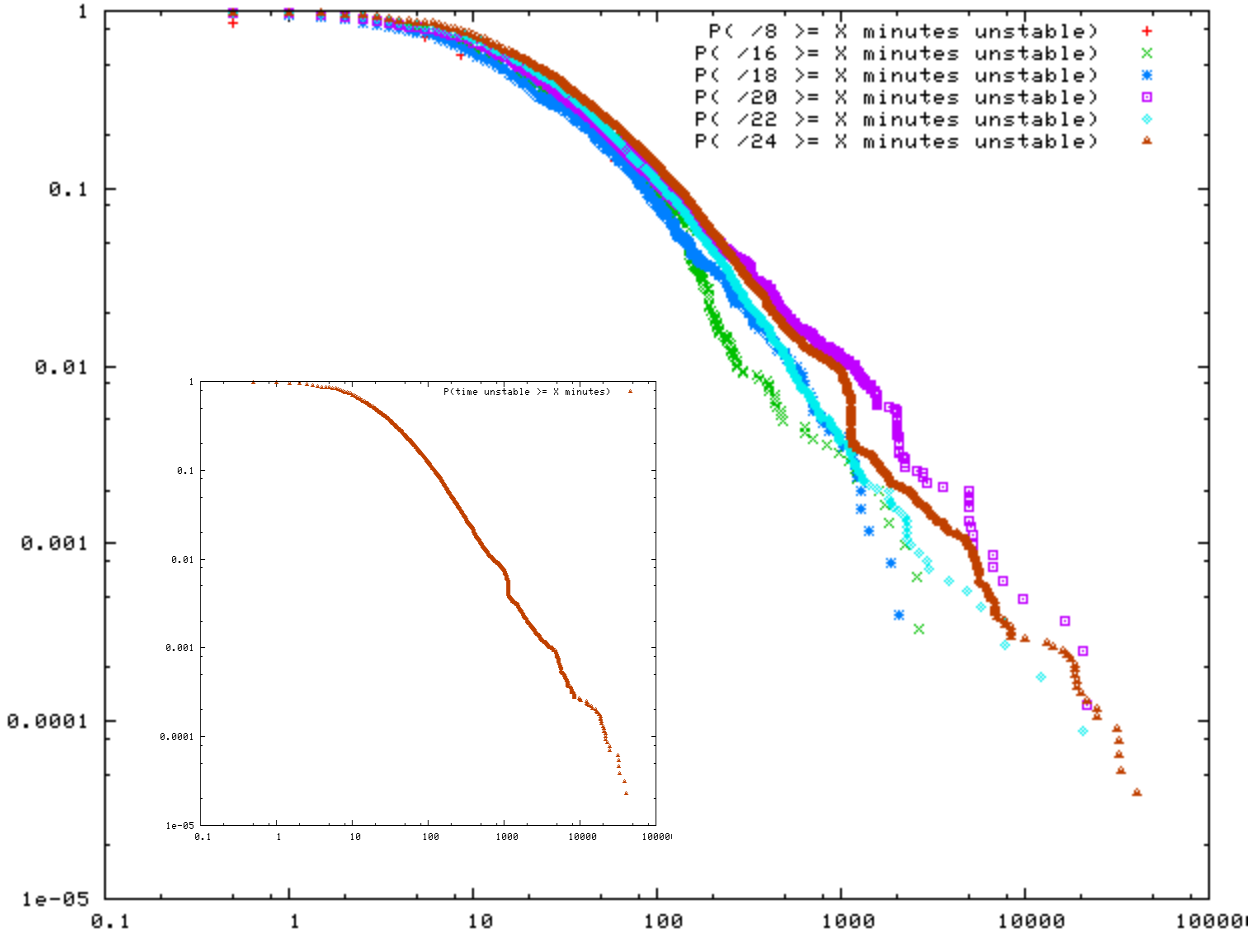
Age and beauty don't correlate.



Prefixes born in 2005, 2006, 2007, 2008 all have basically the same kind of distribution.

2008 vintage is still young, a little more unstable than most.

Size isn't a strong predictor either.

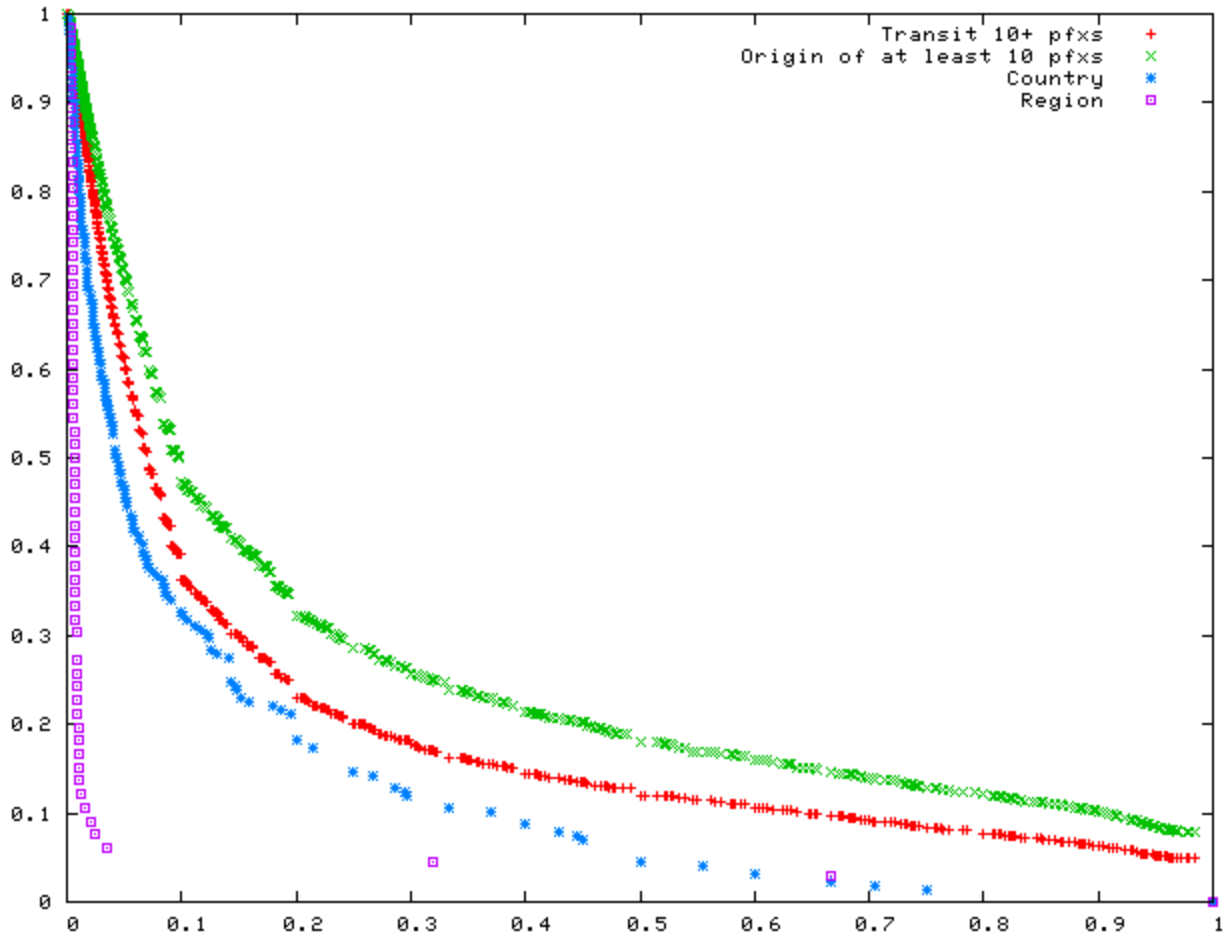


Among unstable prefixes:

- **19% of /16s**
- **15% of /18s**
- **18% of /20s**
- **19% of /22s**
- **23% of /24s**

were unstable for an hour or more during the month.

Nontrivial Transit, Origin, Country, Region



50%+ impact in median hour:

- **3%** of countries
- **no** US states
- **12%** of **nontrivial** transit ASNs
- **18%** of **nontrivial** origin ASNs

How many do YOU have in YOUR customer cone?

Simple scoring metric: percentage impact

- Assign a score to each ASN in a given hour:
  - percentage of on-net prefixes that are significantly impacted by route changes, where
  - "Impacted/Significant" means "withdrawn in any 30 second window within the hour," *or*
  - "at least 3 flaps in at least 5% of the 30 second windows in the hour, seen by majority of peers"
- Simple enough. Let's play the feud!

Grading On A Steep Curve

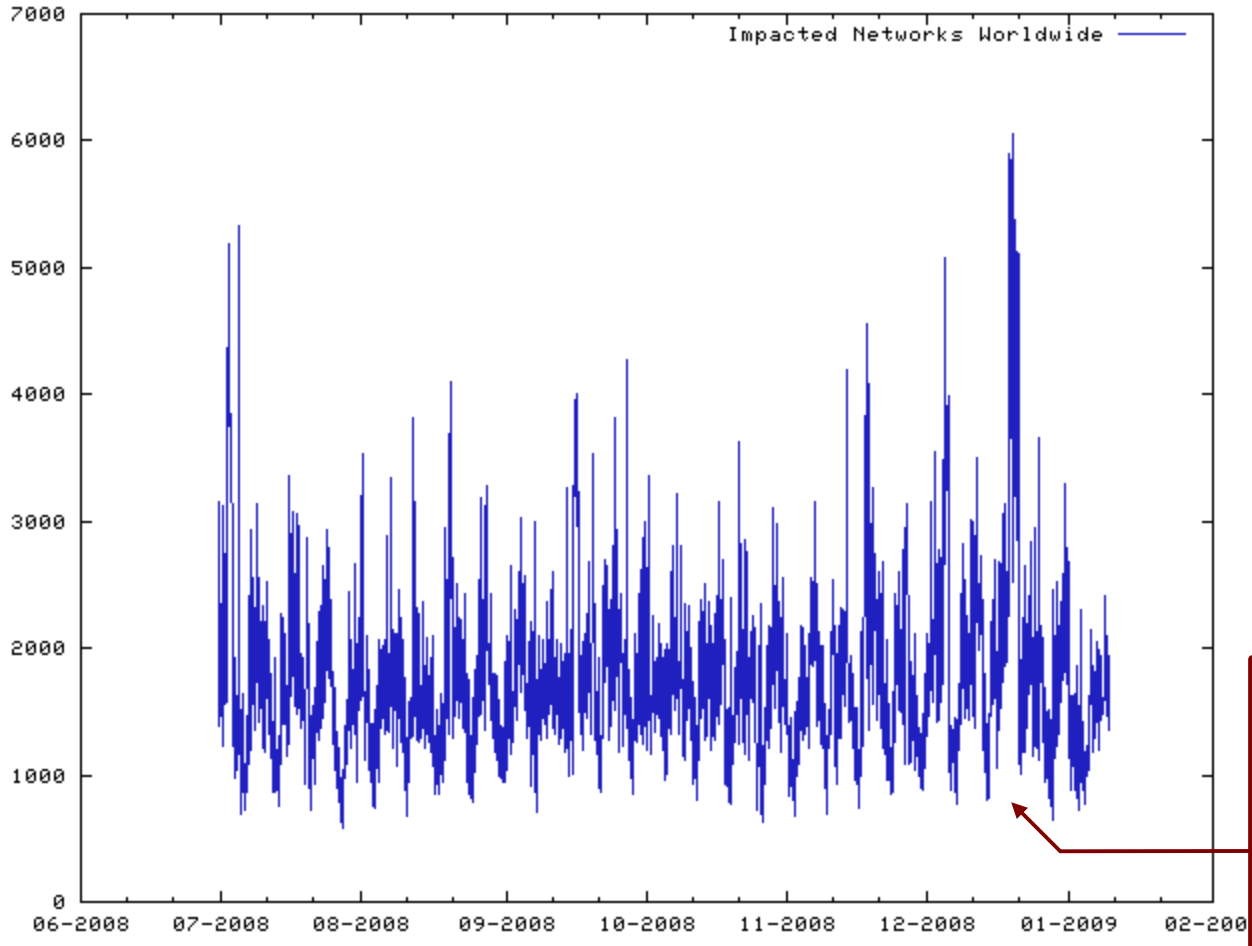
Apply 95-5 rules to 6 months of hourly stability (34Q08).

- **A** 99%+ stable
- **B** 98% stable
- **C** 95% stable
- **D** 90% stable
- **F** <90% stable

Transit Customers with grades A-F, Nov 2008

ASN	Org	A	B	C	D	F
209	Qwest	1290	13	18	10	**15**
3549	GLBX	967	59	86	38	**33**
1239	Sprint	1416	50	63	18	**28**
6453	Teleglobe	295	30	49	28	**33**
2516	KDDI	162	8	6	1	**0**
7018	AT&T	2005	28	38	20	**22**
701	Verizon	2174	43	53	25	**43**
3561	Savvis	398	30	18	14	**5**
3257	Tiscali	312	31	39	24	**12**
3356	Level(3)	1730	92	118	48	**25**
2914	Verio	430	27	46	16	**18**
1299	Telia	333	44	70	24	**15**
174	Cogent	1792	32	68	30	**40**
4134	China Telecom	24	8	7	4	**2**

Worldwide Impact, July-December 2008

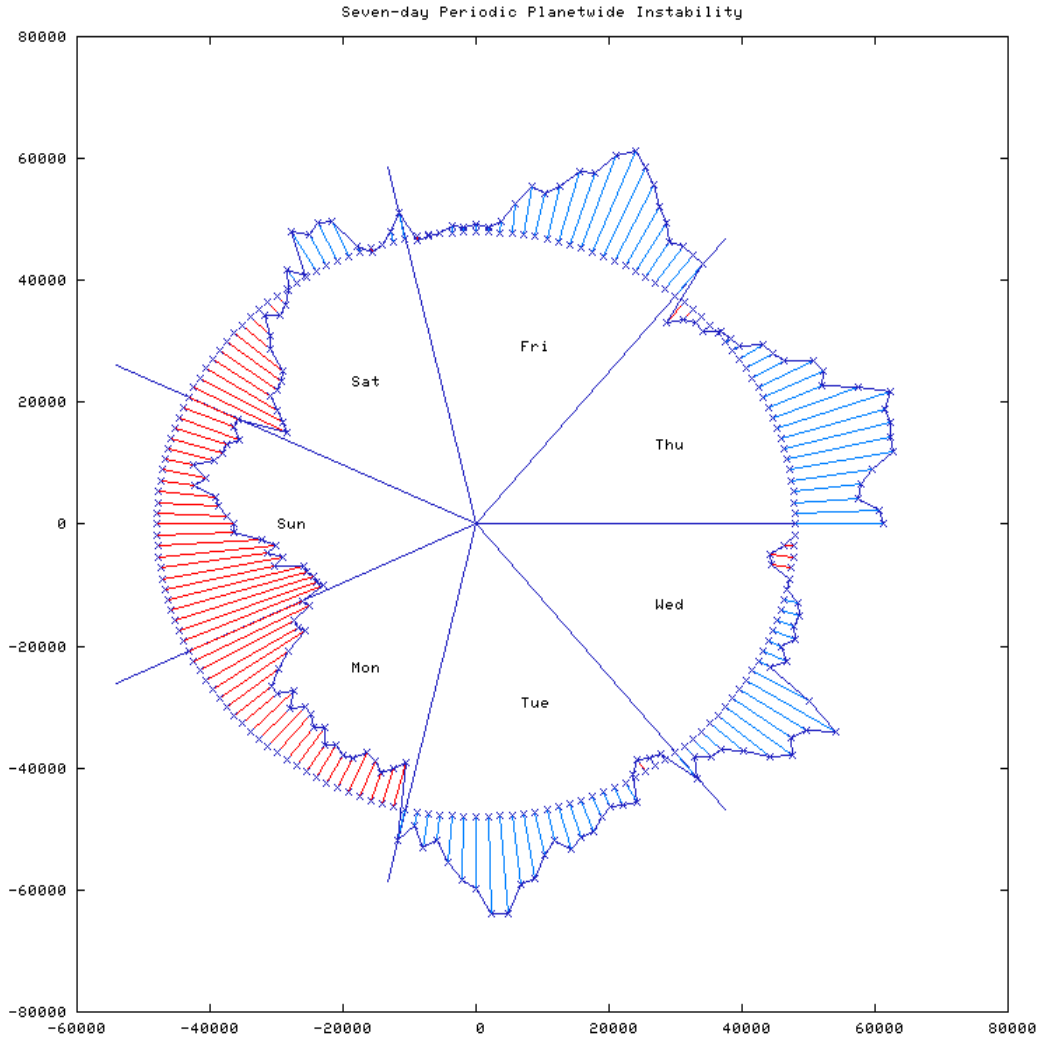


Roughly 1700 networks (0.62%) impacted in average hour.

Seven-day cycle clearly visible.

(December 19th: SMW3-SMW4-Flag cable cut) – 2.24% impacted

Worldwide Impact, July-December 2008



Weekly cycle, clockwise with Sunday at 9 o'clock

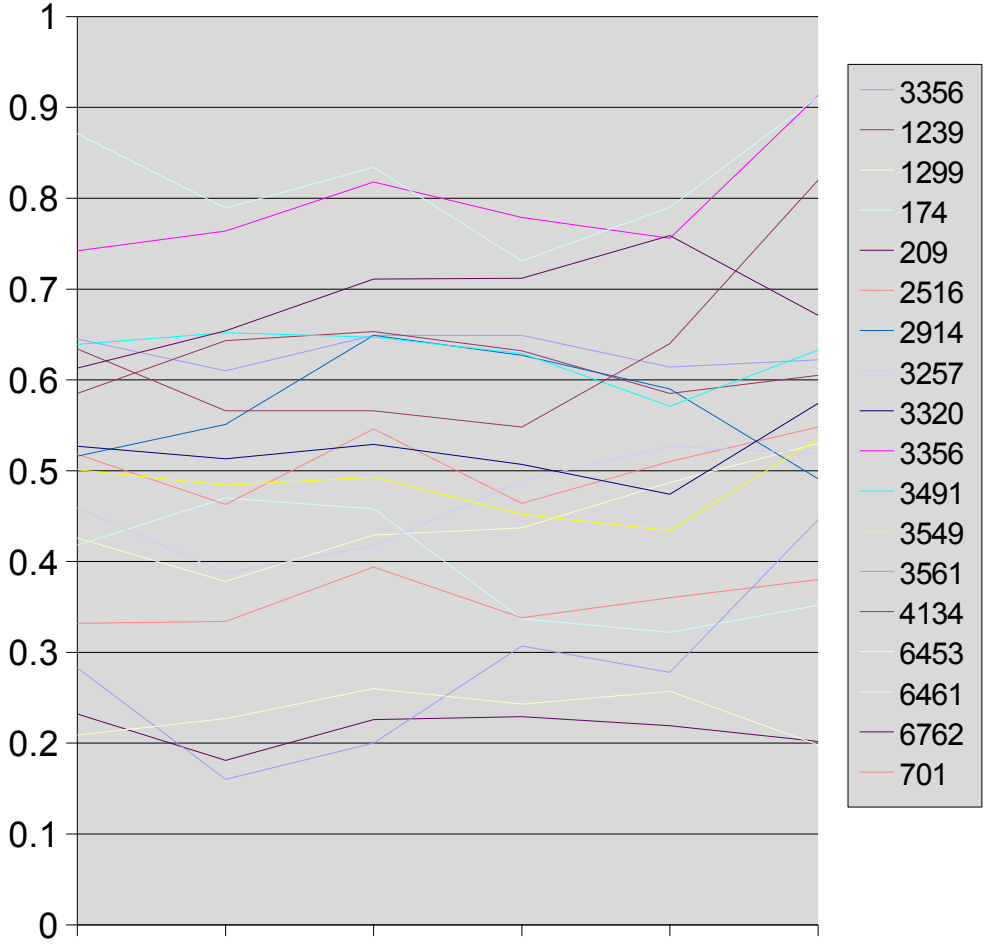
Tuesday through Fridays exhibit diurnal instability (workdays for engineers)

Ironically, Mondays are quieter than average.

Daily low ~23h00 UTC

Whose clock is this?

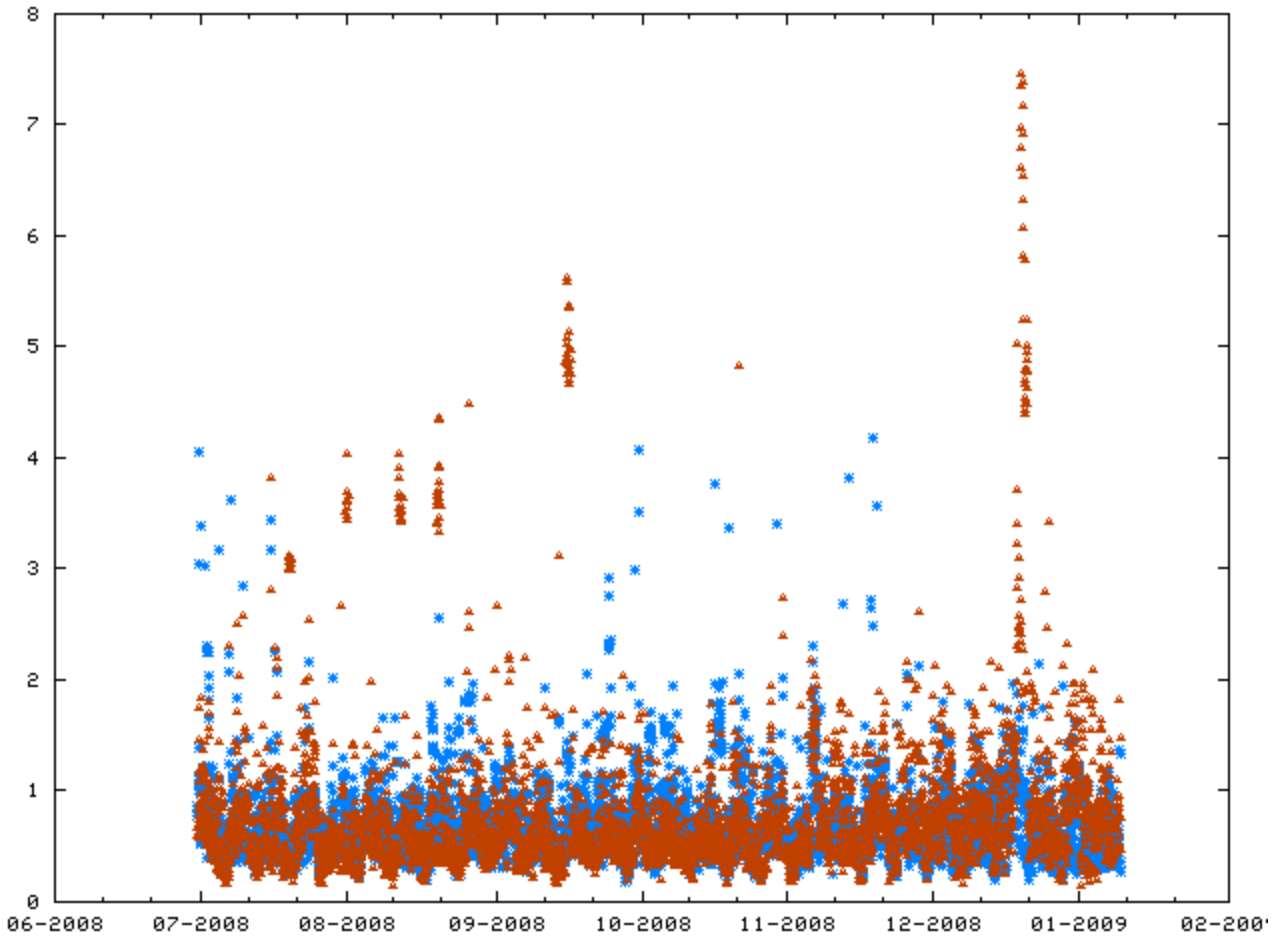
Top 20 ASNs Worldwide



Percentage impact (less than 1%) for median hour in each month July-December 2008. Lower is better.

Note that each of these very large ASNS maintains a pretty characteristic level across many months, based on the makeup of its customer cone.

Telia versus Teleglobe



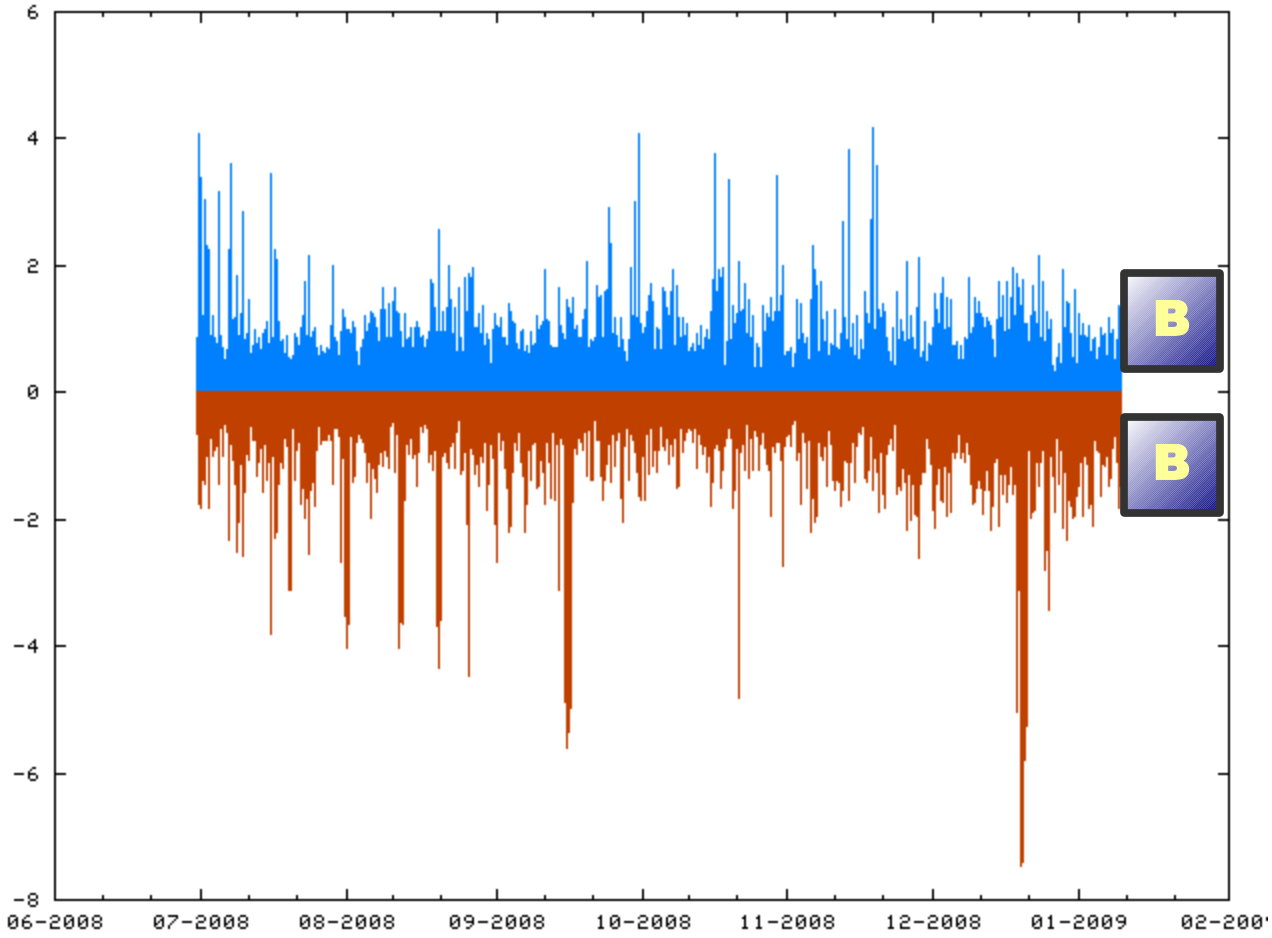
Telia vs Teleglobe

"Percent stable in hour" for each hour July-December '08.

Lower is better.

Evenly matched, both in terms of mean and in terms of sporadic "bursts of failure" (4-8%)

Winner: TELIA (but only by a nose)



Telia: 98.61%

Teleglobe: 98.21%

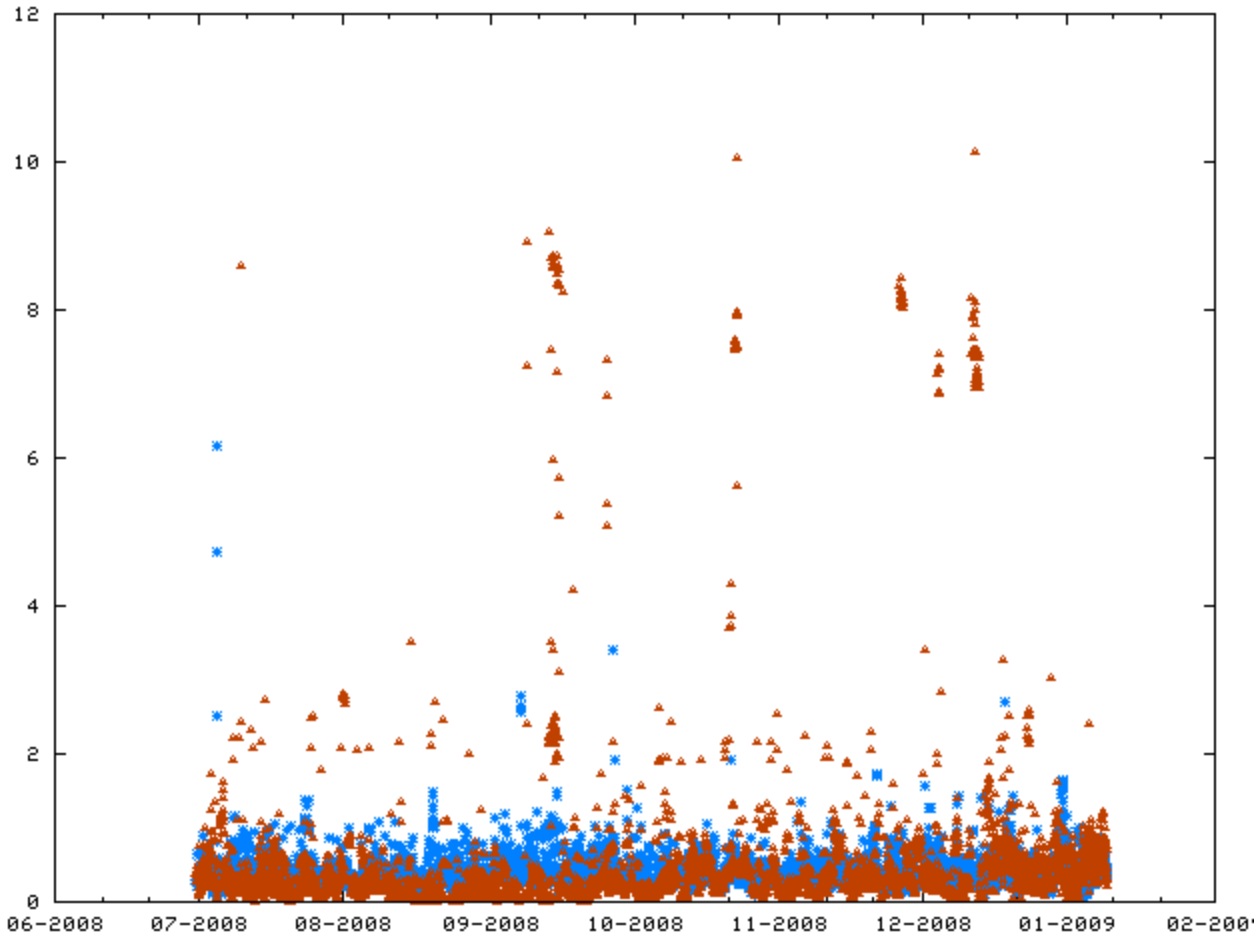
"Percent unstable" for each hour Jul 2008 – Jan 2009.

Smaller is better.

Intuition preserved: evenly matched in average hour:

50.6% :: 49.4%

AT&T versus China Telecom

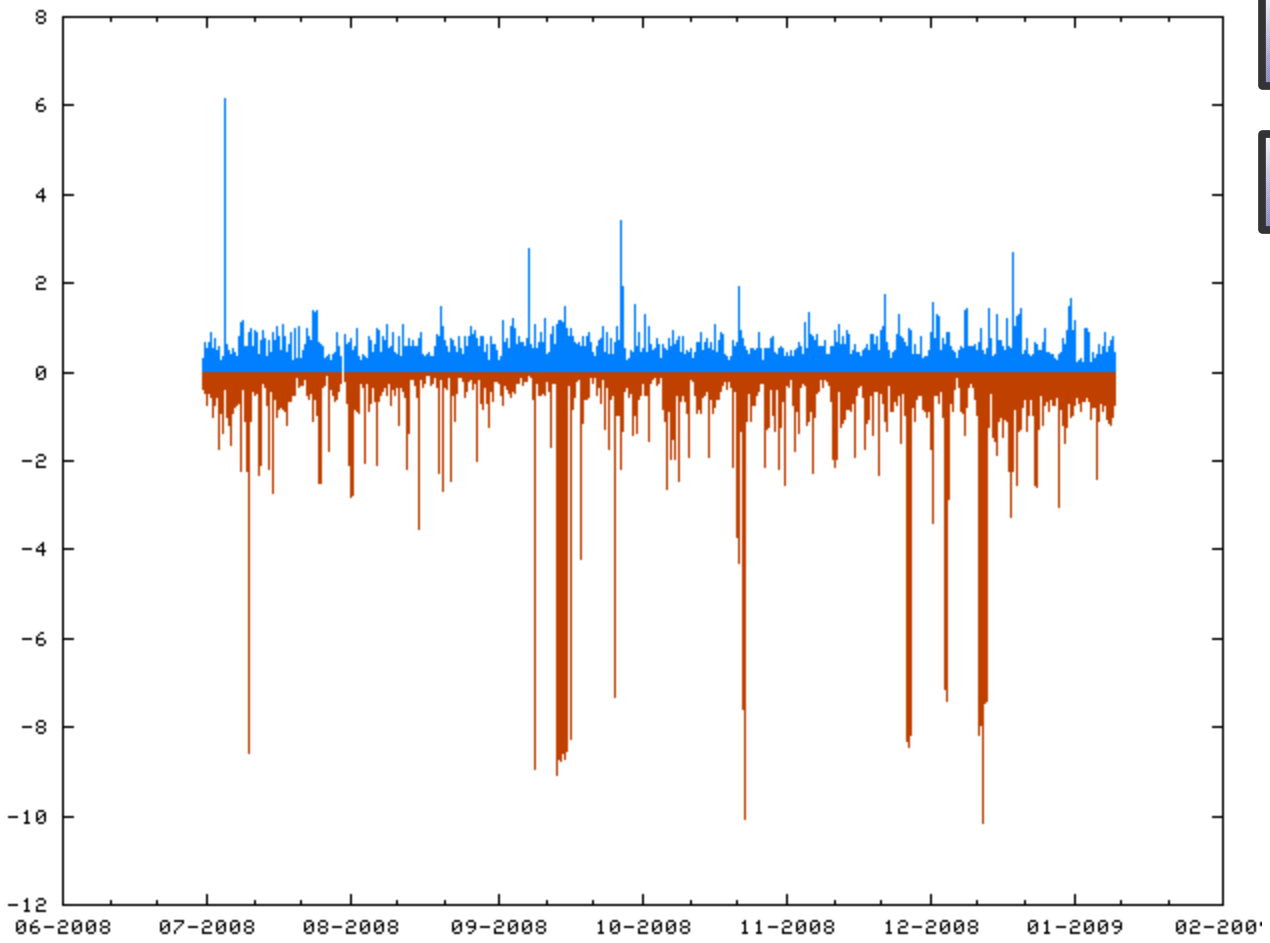


7018 vs 4134

"Percent stable" each hour July-December '08.

China Telecom wins in the average hour, but has more sporadic failure, bursting to 10%.

Winner: AT&T



A 7018: 99.16%

B 4134: 98.11%

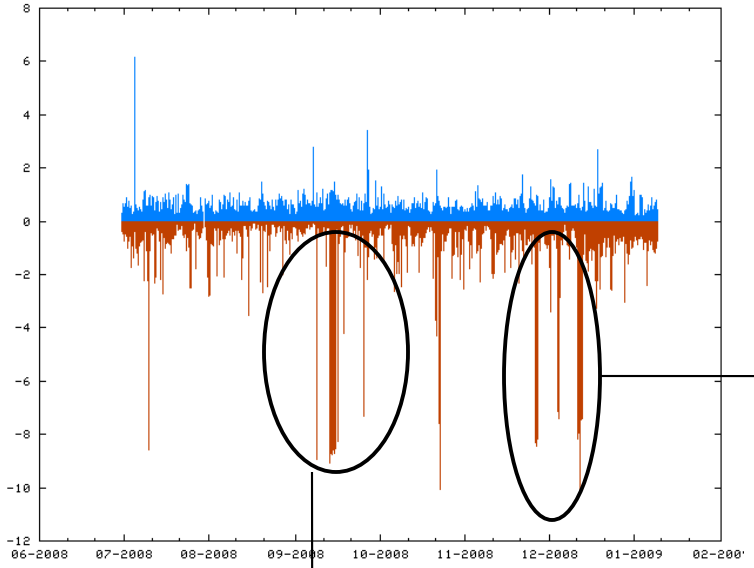
China Telecom is more stable in the average hour…

37.7% :: 62.3%

…but has more high-instability (>1%) hours:

118 :: 388

China Telecom Observations



As well as:

- **AS7552, Vietel (VN)**
- **AS4837, CNCGROUP China169**

C

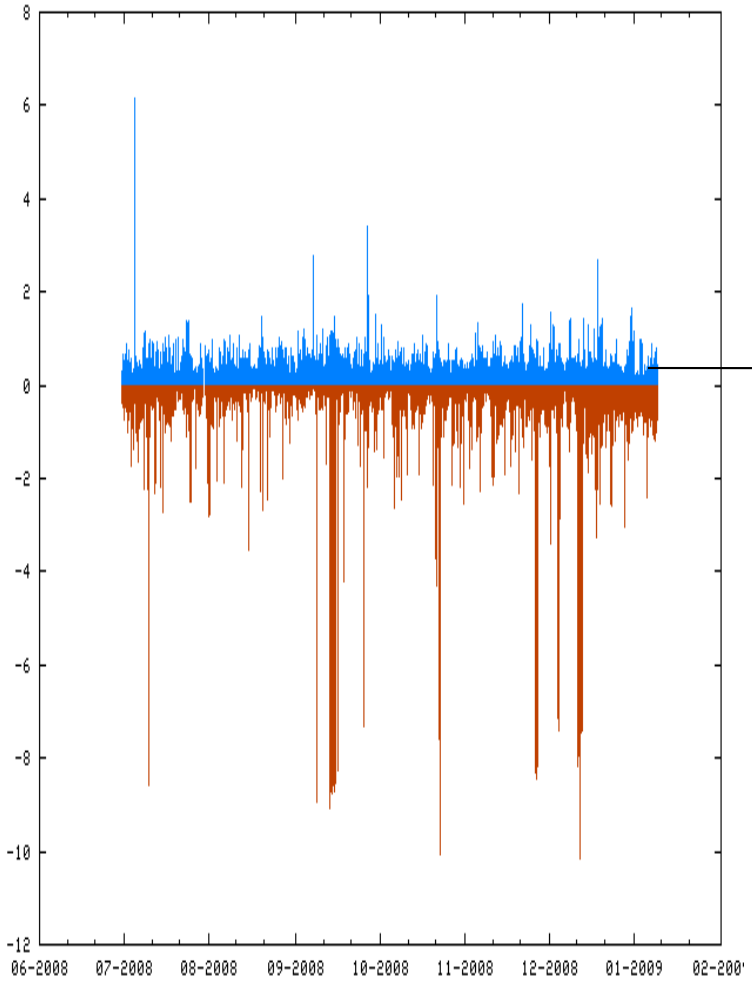
China Telecom transits hundreds of unstable Vietnamese and Cambodian prefixes on behalf of AS7643 (VN)…

D

…and hundreds more on behalf of AS4538 (China Education and Research Network Center)

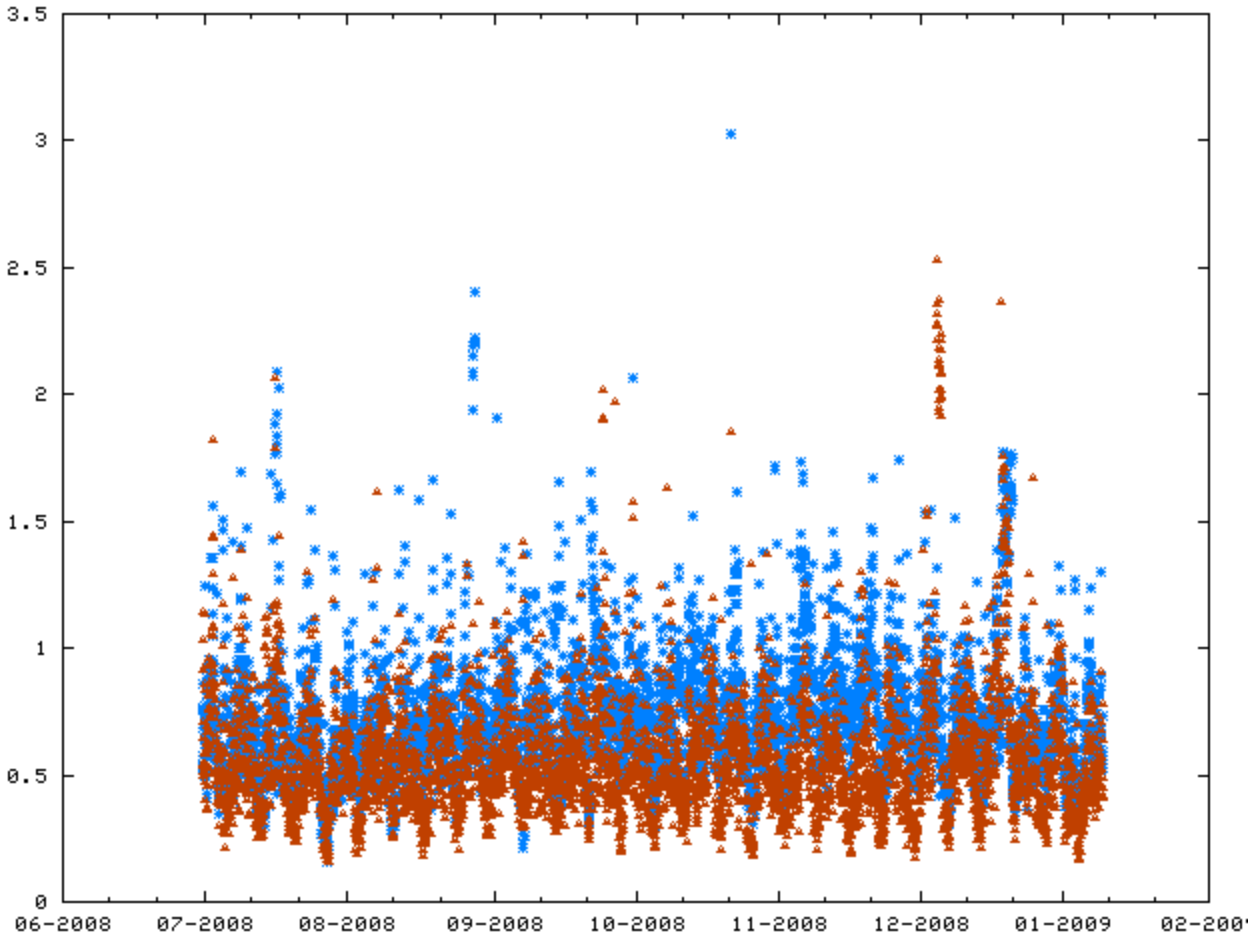
D

AT&T 's Unstable Downstreams



- D •AS8151, Uninet (MX)
- F •AS39386, Saudi Telecom
- D •AS4788, TMNet (MY)
- B •AS3786, LG Dacom (KR)
- C •AS9929, China Netcom (CN)
- C •AS28513, Uninet (MX)
- B •AS9498, Bharti Airtel (IN)
- C •AS4837, CNCGroup China169

Verizon (UUNet) Versus Level(3)



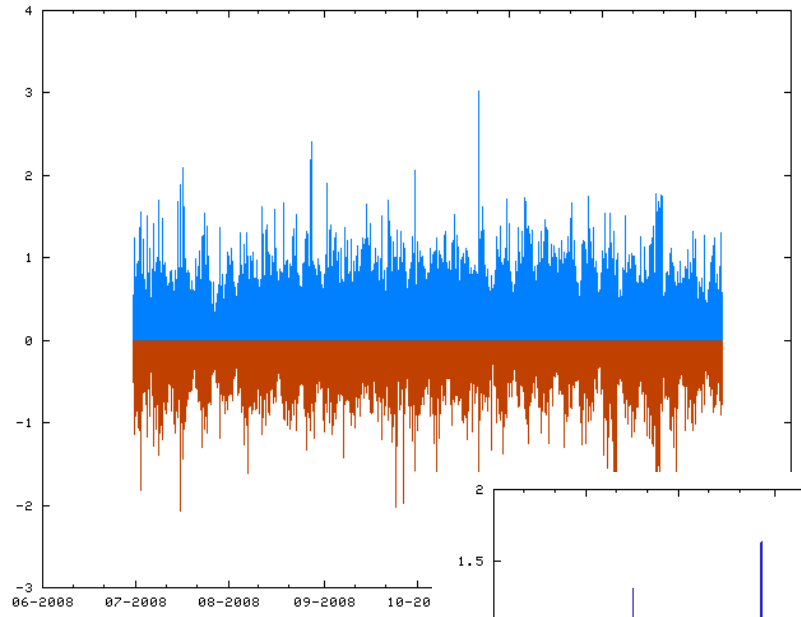
701 vs 3356

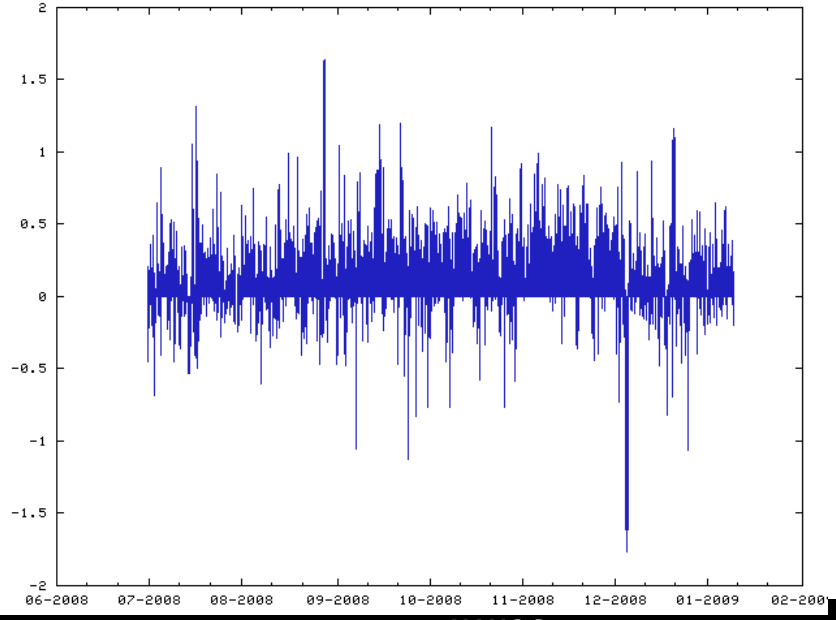
Again, hourly pct% unstable on-net, from July 2008 to January 2009.

Lower is better.

Level(3) is more stable overall, and less bursty to boot.

Winner: LEVEL(3)





Difference tells the story…

B **701: 98.8%**

A **3356: 99.01%★**

Level(3) is more stable in more hours:

19.0% :: 81.0%

…and has many fewer high-instability (1%+) hours:

502 :: 218

Level(3) and Uunet Observations

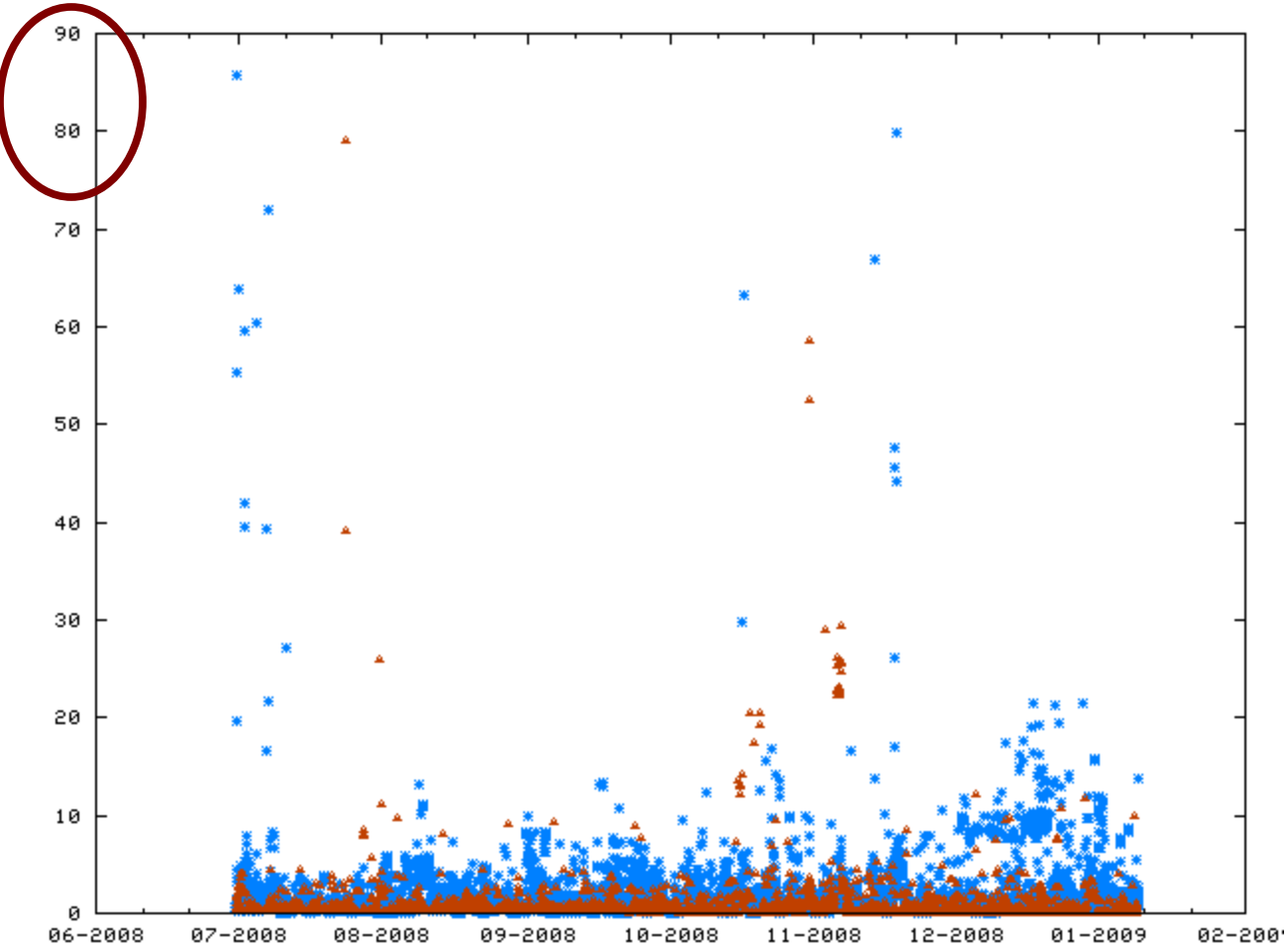
Level(3) selected unstable

- C •AS1273, C&W
- C • … AS20485, TransTelecom (RU)
- C •AS7473, SingTel (SG)
- C •AS8342, RTComm.RU
- C •AS30890, Evolva (Romania)
- C •AS8359, COMSTAR (RU)
- C •AS9498, Bharti Airtel (IN)
- D •AS9121, TTNet (TR)

UUNet selected unstable

- C •AS28513, Uninet (MX)
- F •AS38040, TOT (TH)
- D •AS4788, TM Net (MY)
- C •AS20485, TransTelecom (RU)
- •AS702 (themselves)
- F • …AS9070, ITD (BG)
- • …AS8866, BGTel — C
- C • …AS17557, PK Telecom
- • …AS38193, Transworld (PK) — D
- C • …AS12883, Vega Telecom (UA)
- • …..

Turkish Telecom Vs Bulgarian Telecom

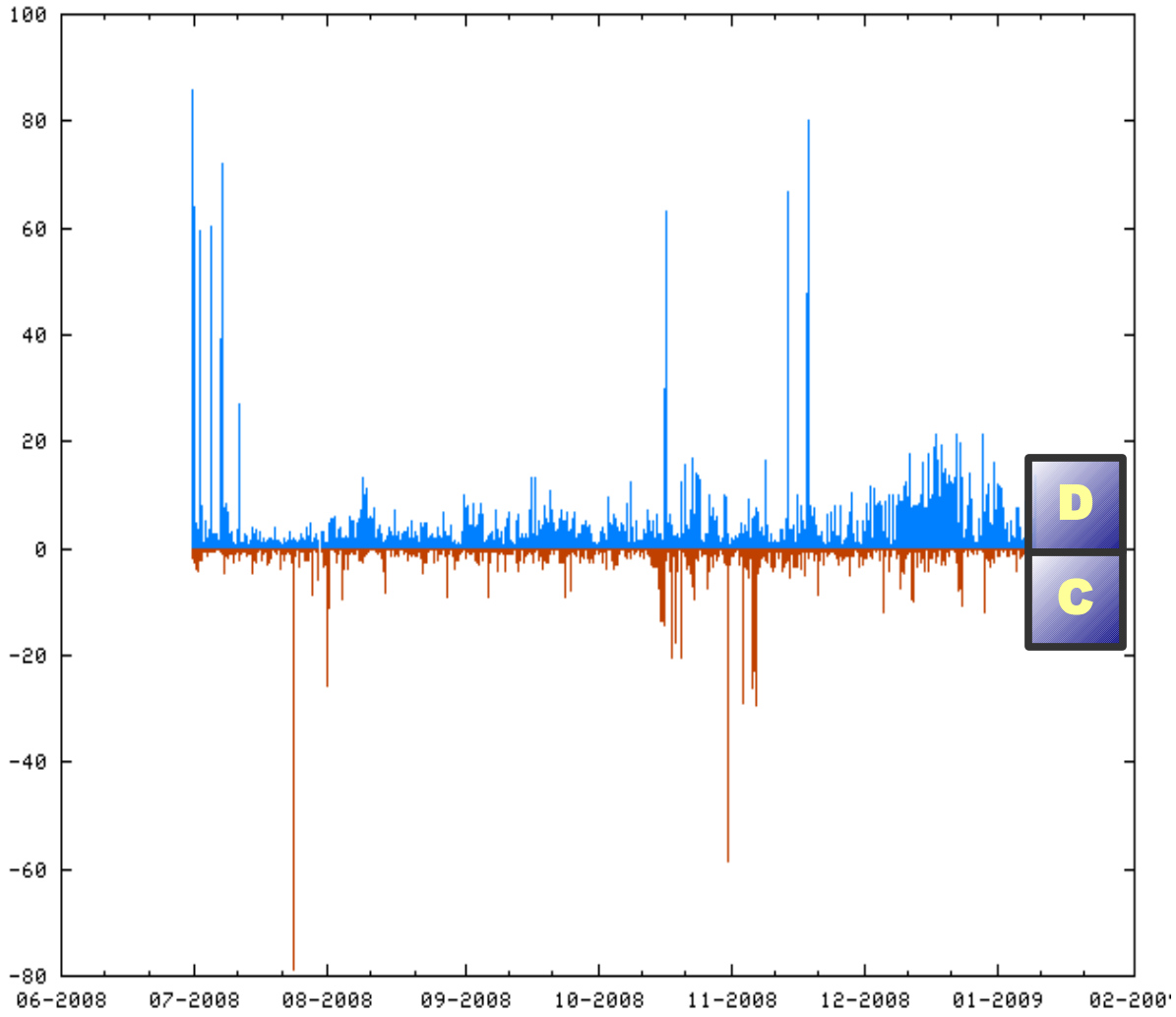


Turks (9121) vs Bulgarians (8866)

Smaller carriers, more unstable networks. Note change of scale!

This one is hard to call. Let's look at the head-to-head numbers....

Winner: BULGARIAN TELECOM



TTNet 9121: 92.33%

BTC 8866: 97.42%

The Bulgarians are big winners in the average hour:

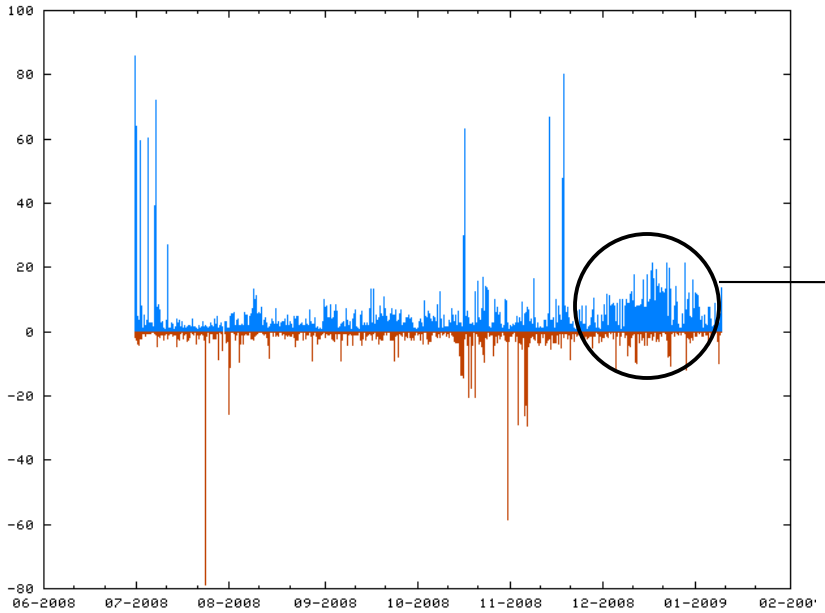
19.8% :: 80.2%

...and have fewer (10%+) hours:

F

111 :: 48

Turkish Telecom Observations



D **AS9121 transits:**

F •**Hundreds of Georgian and Armenian prefixes via United Telecom (AS35805)**

F •**Hundreds of Iranian prefixes via Data Communications of Iran (AS12880)**

D •**Over a thousand Egyptian prefixes via TEDATA (AS8452)**

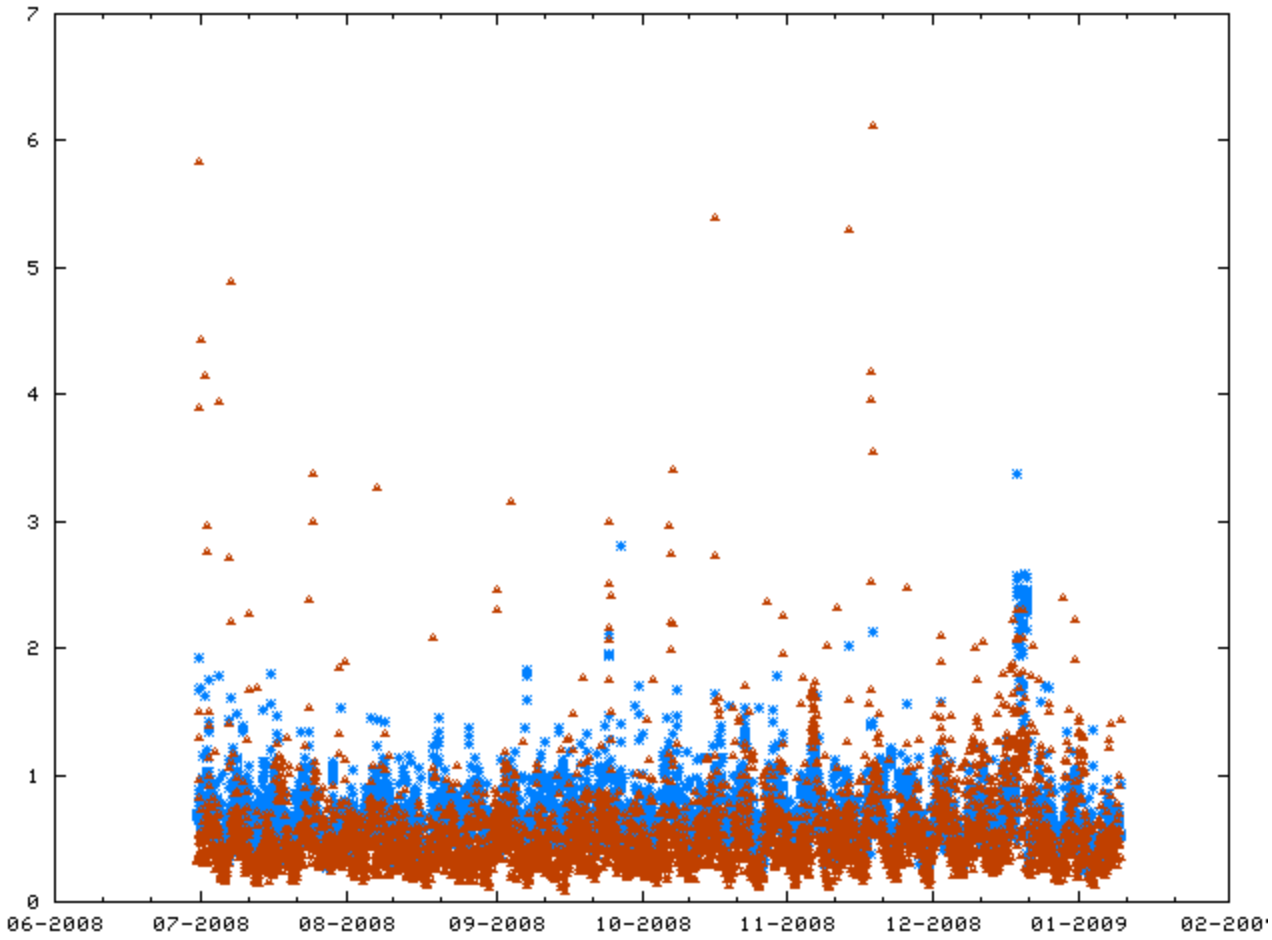
SPRINT versus COGENT

1239 vs 174

Who's going to be more unstable?

The well-respected #1 ranked global transit provider, or the “cheapest to deliver” solution who built their business on ROCK BOTTOM PRICING!?

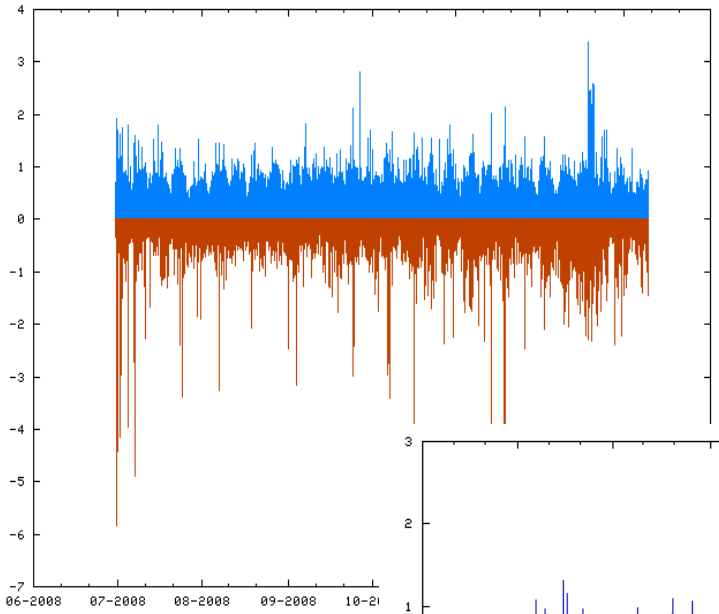
SPRINT versus COGENT

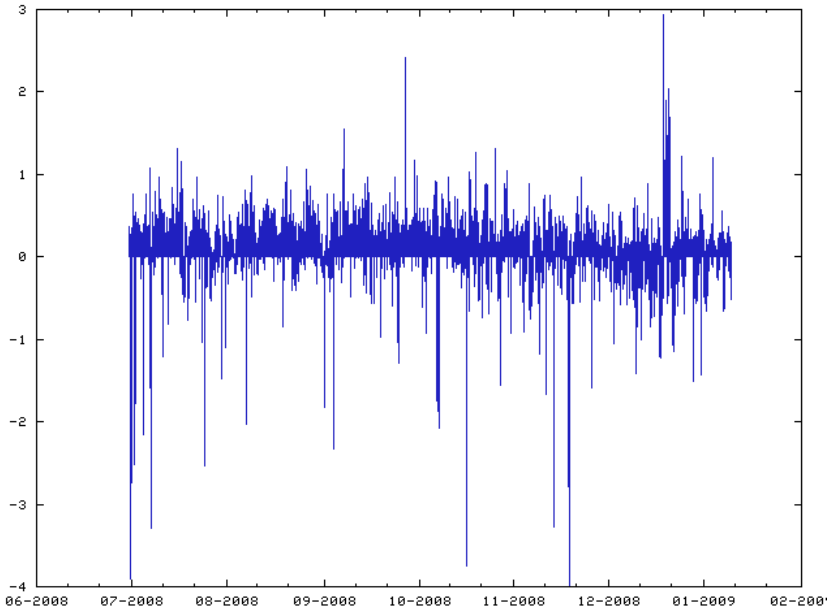


1239 vs 174

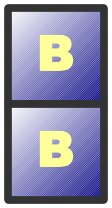
Wait for it ...

Winner: COGENT by a nose





Again, compare the difference.



1239: 98.87% (tie)

174: 98.82% (tie)

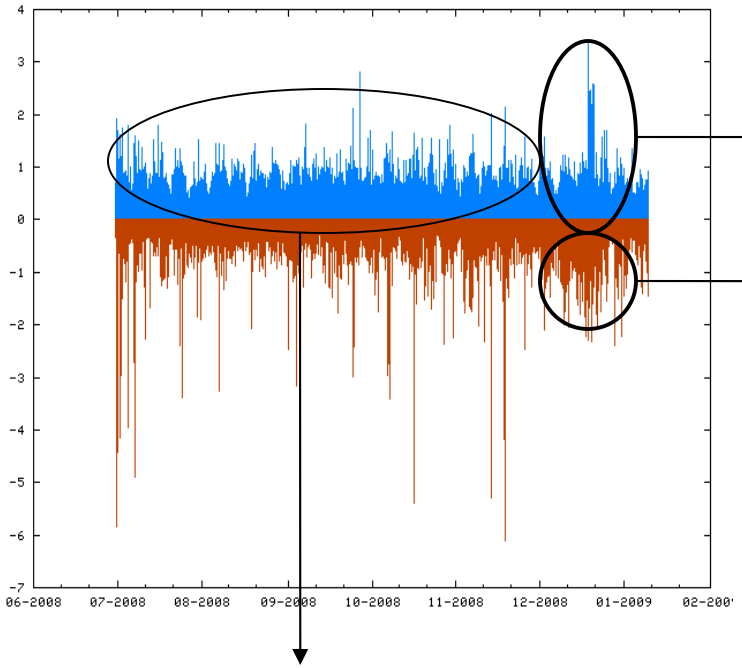
But Cogent wins head-to-head in more hours July-January:

23.9% :: 76.1%

And has fewer hours bursting above 1%+:

389 :: 380

Sprint-Cogent Observations



Cogent transits Turkish Telecom (9121) in the Middle East. D

Sprint provides transit to Telecom Italia (6762) for similar customers. B

Sprint's unstable on-net customers are diverse:

- **AS11830 (Costa Rica)** B
- **AS5588 (GTS Central Europe) / Antel Germany** C
- **AS4837 (CNCGROUP China169)** C
- **AS39386 (Saudi Telecom Company)** F
- **....**

Conclusion: Why You Should Care

- Some prefixes are **significantly** more unstable than others, over long periods of time; they cluster by the ASNs whose cones they're in
- ASN customers who contribute to route instability are potentially more expensive to support
- **The relative stability of your customer cone can be a significant differentiator in the eyes of an enlightened customer**

renesys

Thank You!

Jim Cowie
cowie@renesys.com