

Internet stability amid change

Andre Broido
broido@caida.org, CAIDA/SDSC/UCSD

kc claffy
kc@caida.org, CAIDA/SDSC/UCSD
10 dec 01

Overview

- Data description and caveats
- Taxonomy of prefixes (definitions)
- Topology changes
- Prefix set changes
- More specifics
- Multihoming
- Routing updates

Data analysis: BGP peer and prefix selection

- Select only full-size tables from RouteViews peers
 - full: > 104K prefixes
 - filtered: 89-97K prefixes

Prefix selection:

- globally routed prefixes (those common to all tables)
- semiglobal (found in the majority of backbone tables)
 - prefixes dropped by one (few) peers still observed
 - prefixes seen only locally won't get into table
 - more robust than either full or global prefix set
- We use only semiglobal prefixes

Caveats:

- many metrics are sensitive to prefix/peer choice
- important when trends are small and/or undecided
- instant and daily influx or loss of prefixes
- Table sizes differ noticeably (few %) by day

Data sources

UOregon ANTC RouteViews BGP tables

- Oct. 31, 1999 (14 tables)
- Nov. 1, 2000 (18 tables)
- Nov. 1, 2001 (26 tables)

unless otherwise specified

Sampling for prefix analysis:

- from 6 months (1997-2001)
- to 2 hours (data from the end of 2001.)
- updates in continuous mode, all of Nov 2001

Prefix taxonomy

Prefix types:

- **standalone** -- no subsets/supersets in the table
- **root** -- a least specific route, has subsets in the table
- **more specific** -- a subset of some other prefix

In addition: **top** prefixes = standalone + roots

Caveats:

- depends on the prefix set
- not an inherent property of a prefix

prefixes seen by given number of peers

Nov.01, 2001, 26 tables

1	3925
2	1880
3	490
4	259
5	595
6	217
7	131
8	86
9	65
10	29
11	220
12	57
13	14
14	98
15	20
16	12
17	105
18	8
19	13
20	15
21	7
22	7
23	55
24	215
25	996
26	102000

Bulk measures: Dec. 2001 BGP tables

Global prefixes (26 peers): 102394

Semiglobal (13 or more peers): 103828

Multiorigin prefixes (union of 26 tables): 1170

Smallest blocks (/24s): 57% of global

More specifics: 50% of global

ASes: 12399

Transit ASes: 2001 (16%)

AS paths: 15-16K/table

links in BGP graph: 25630

Atoms: 25698

Average prefixes per AS: >8

Maximum #prefixes announced by single AS: 2103

ASes with one prefix: 40%

Prefixes announced by such ASes: 5%

ASes with over 100 prefixes: 1% <--- we define as "large ASes"

Prefixes announced by such ASes: 32%

Growth in prefixes, ASes, and degree

	1999	2000	2001
Semiglobal prefixes	64769	88714 (+37%)	102394 (+15%)
ASes	6107	9116 (+50%)	12399 (+36%)
AS links/nodes (avg in/outdegree)	1.97	2.00	2.07

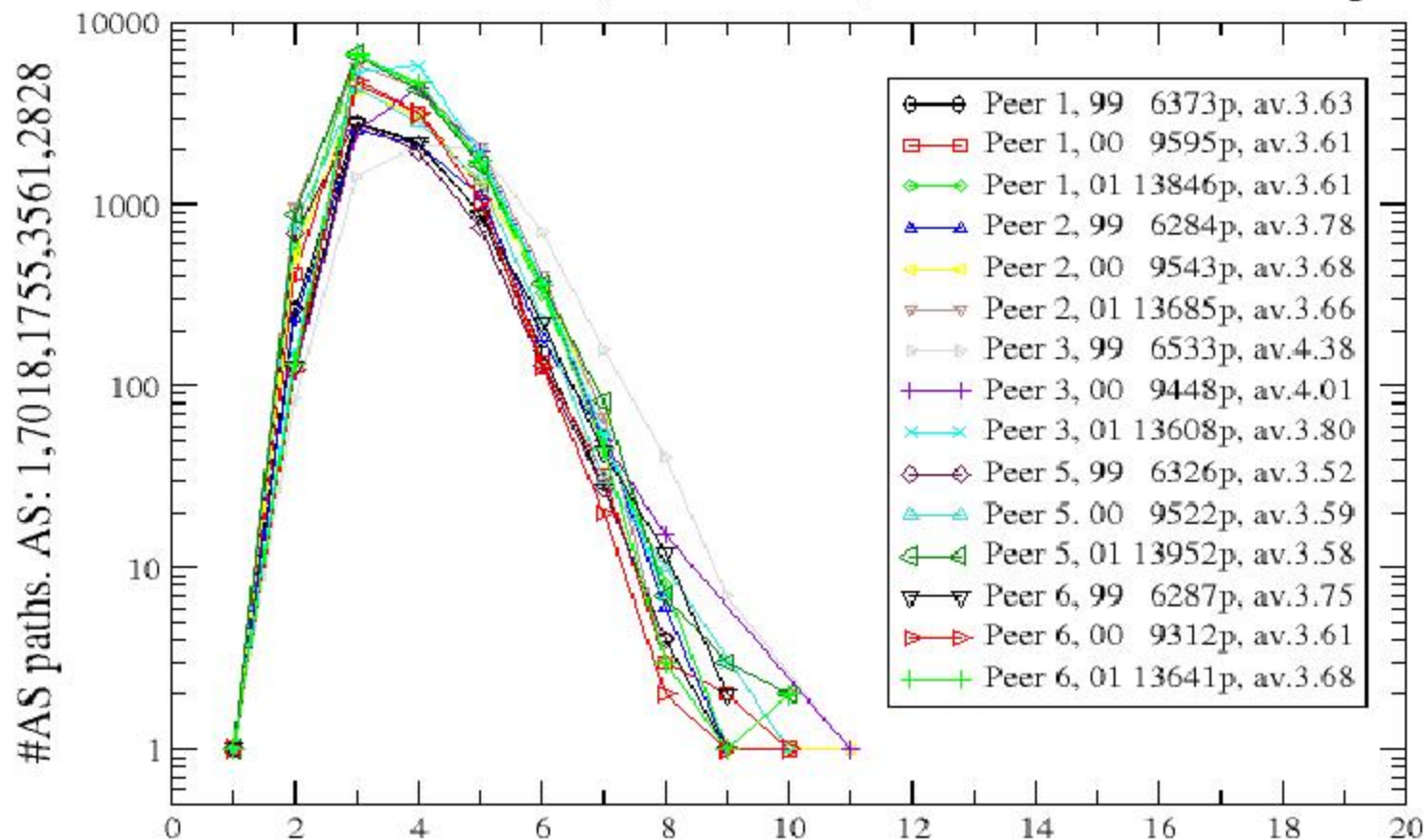
- prefix and AS growth slowed in 2001
- prefix growth is close to 2/3 of AS growth
- link/node ratio grows slowly, almost invariant

peering richness, and navigation complexity of BGP AS graph did not significantly change between Nov 2000 and May 2001, although there were a lot of changes for individual ASes

Is average AS path length decreasing?

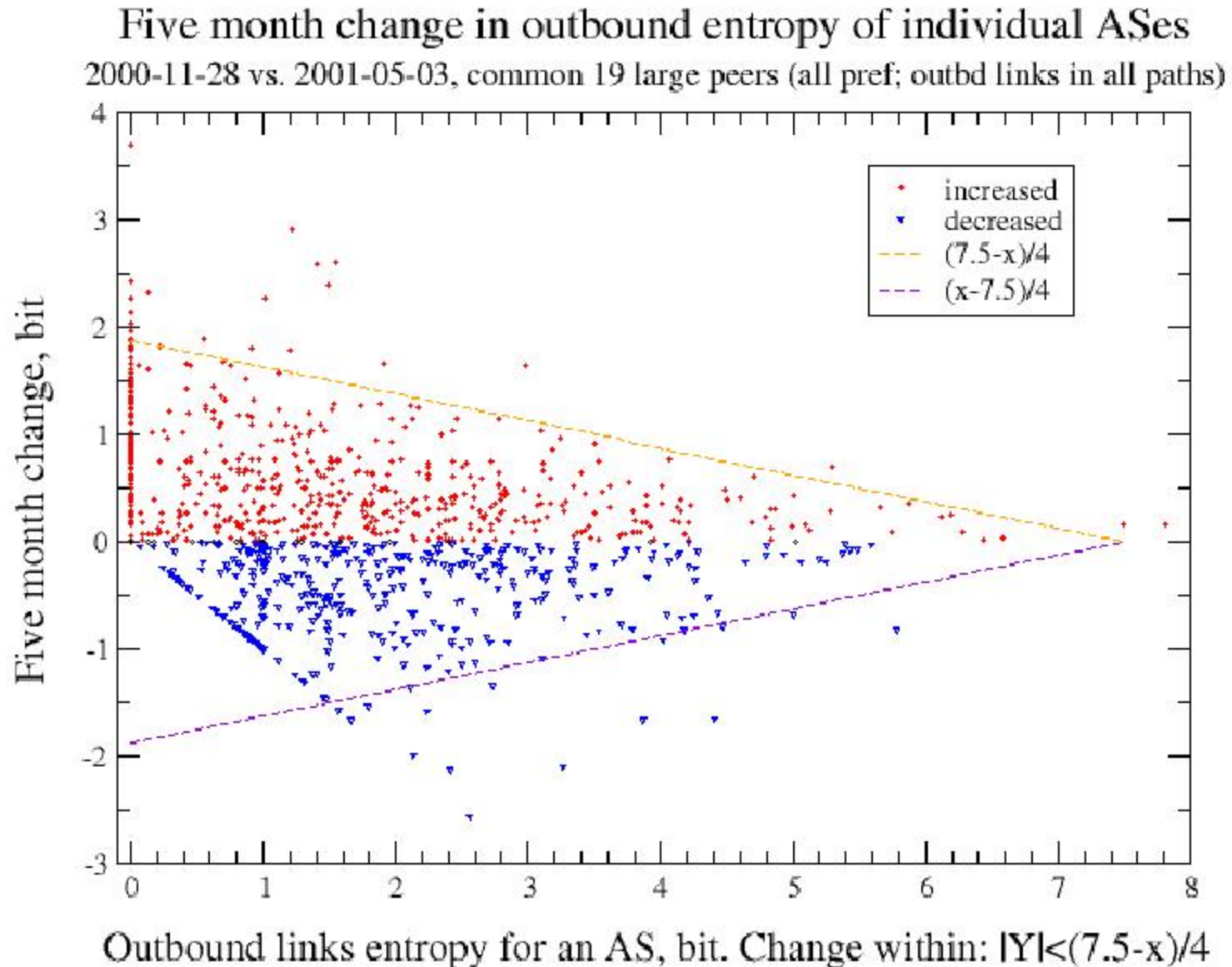
- Since 1999, many AS paths changed either way
- Change in the average AS path length is insignificant

AS path length distributions for 1999, 2000, 2001
Route View 1999-05-01, 2000-05-02, 2001-05-03. 6 common peers



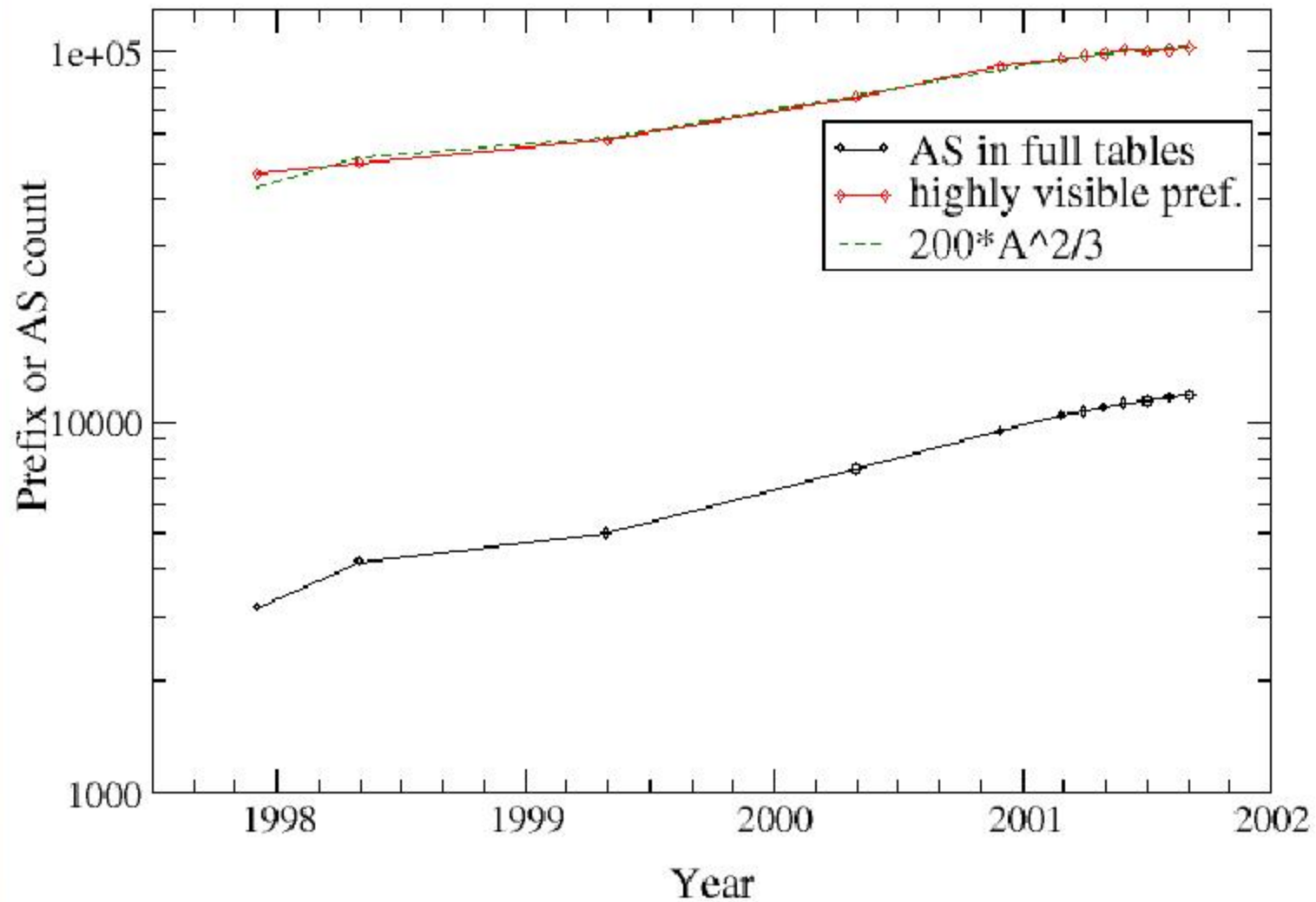
AS path length, ASes. Chg.by Nov.01: same,down,up,down,up

Peering richness: changes per AS, not on average



AS refinement formula: $P = 200 A^{2/3}$

Prefix vs. AS growth: $P = 200 * A^{2/3}$
Route Views 1998-2001. Prefixes present in $> 1/2$ of all full tables



Super-refinement

Refinement formula worked for 3 years

- Current prediction: 107140 (formula overestimates)
- Prefix growth stopped (last few months)
- AS refinement faster than predicted by formula
- Will eventually have less than one prefix per AS
- Already have ASes with 0 announced prefixes
- (85 transit-only ASes for Nov 01, 2001, up from 55 a few months ago)

Prefix distribution: findings

- Changes in global routing not captured by bulk measures such as number of prefixes in a table
- Net change is a sum total of commensurable contributions of opposite sign, so...
- Total variation is not the difference, but rather the sum of their magnitudes.
- Most measures of routing system complexity demonstrate slow growth, dynamic equilibrium and/or decrease for May-Dec. 2001.

Current state of the Internet routing system shows large variation but NOT rapid growth!

Prefix distribution

■ /24 prefixes

- the smallest globally routable address blocks
- 58% of the table, fraction stable since 1997

■ Q: Is proportion of /24s in backbone tables growing?

A: No

- Since 1998, the share of /24s has stayed between 57% and 58.5%
- Has decreased since then by 0.6%-0.9% in all non-filtered tables

■ About 1/2 of all prefixes are more specifics (subsets).

- 1999 50%
- 2000 55%
- 2001 52%

Prefix distribution: refinement of prefix trees

Prefix type breakdown

(semiglobal prefixes in 26 tables, Nov.01,2001)

standalone	44264	42.75%
root	5601	5.41%
more specifics	53686	51.84%
total prefixes	103551	100.00%

Percentage of root prefixes growing steadily, if slowly

- 1999 4.64
- 2000 4.86
- 2001 5.41

Another refinement phenomenon:

- growing share of the table is roots
- constant or decreasing share of more specifics
- average number of subblocks of a root is decreasing

Prefix set churn

Churn (emergence and disappearance of prefixes in the table) was much higher than growth rate.

- highest for more specifics
- lowest for standalone blocks
- net change in table size sometimes < 0

semiglobal prefixes in following data uses all available RouteViews backbone tables, including filtered tables, for 3 sample points: May 3, August 1, and November 1, 2001 (33, 42, and 39 tables, respectively)

Changes in prefix status from May->Aug 2001

- transitions are from columns to rows
 - from May to August 2001
 - equal number of prefixes moved from standalones to roots/m.s.'s
 - more prefixes changed from m.s. to standalone than v.v.
 - "Vacuum" counts prefixes which appeared or disappeared

	St.al	Root	M.sp	Vac	Sum	Out
St.al	38676	544	576	2450	42246	8.45
Root	377	4374	40	307	5098	14.20
M.sp	917	41	42188	8746	51892	18.70
Vac	4232	420	9206	0	13858	100.00
In	13.08	19.71	18.93	83.01	99236r	
Sum	44202	5379	52010	11503	101591c	

Changes in prefix status from Aug->Nov 2001

- transitions are from columns to rows
 - from August to November 2001
 - Root prefixes coming mostly from standalone (607 vs. 433 new)
 - more prefixes changed from standalone to m.s. than v.v. (reversal of May-Aug trend)

	standal.	root	more spec	vac	sum	out
St.al	39530	607	1460	2605	44202	10.57
Root	396	4532	146	305	5379	15.75
M.sp	604	33	41835	9538	52010	19.56
Vac	3738	433	10178	0	14349	100.00
In	10.72	19.95	22.66	86.75	101591r	
Sum	44268	5605	53619	12448	103492c	

Prefix set churn, cont.

May-Aug 2001 (33 (May) and 42 (Aug) RouteViews peers):

	% standalone	% of root	% more spec.
Out	8.45	14.20	18.70
In	13.08	19.71	18.93
Net	4.63	5.51	0.23

No growth for more specifics in May-Aug 2001

Aug-Nov 2001 (42 (Aug) and 39 (Nov) RouteViews peers):

	% standalone	% of root	% more spec.
Out	10.57	15.75	19.56
In	10.72	19.95	22.66
Net	0.15	4.20	3.09

No growth for standalones in Aug-Nov 2001

("In" and "Out" include transitions when existing prefix changes its type, e.g., from standalone to more specific)

Growth of non-transit ASes

- Transit and non-transit ASes contribute about equal numbers of prefixes despite the fact that there are 5 times as many non-transit ASes as transit ASes.
- Transit ASes announce over 1/2 prefixes (53%)
- For 1999-2000-2001:
 - Fraction of nontransit ASes: 82%, 83%, 84%
 - Announcements from nontransit ASes (percentage of semiglobal): 42%, 43%, 46%

Growth of non-transit ASes, cont.

Q: Does the Internet grow mostly at the periphery?

A: Most likely, although BGP graph may not have enough coverage at network edge to answer

- may show transit ASes as nontransit
- may undercount indegree (multi-homing)
- captures only downstream, not lateral, connectivity
 - 18%, 17%, 16% transit ASes in 1999, 2000, 2001
 - absolute number grows but proportion shrinks
 - bidirectional part of the graph (core): 3.13%, 2.24%, 2.09%
- compare to skitter (traceroute-based) AS graph:
 - 32% transit ASes, twice as many as BGP graph
 - 28% ASes in combinatorial core (BGP: 2%)

Multihomed ASes

Definitions: AS with indegree > 1 in BGP AS graph is called **multihomed**
AS with outdegree > 0 in BGP AS graph is called **transit**

BGP gives only a lower bound on multi-homing,
although a consistent one

- (26 vs. 39 tables differ by 0.3% in multi-homed ASes)

For Nov 01, 2001 RouteViews data, multihomed networks originated

- 76% of all prefixes
- 74% of all more specific (subset) prefixes
- includes transit & non-transit
- **multihomed nets originate a proper fraction of more specifics**
- Non-transit multihomed ASes contribute **less** prefixes than their

fair share:

Year	%AS	%prefixes
1999	43.65	27.46
2000	45.84	29.37
2001	48.75	29.74

- Fraction of non-transit multihomed ASes grew in 1999-2001
- Fraction of prefixes grew in 1999-2000, was constant in 2000-2001

Transit AS prefixes by type (Nov.01 2001, 26 tables)

Prefixes from transit multihomed ASes:

standalone	23801	49.28%	
root	3723	7.71%	
more specific	20777	43.02%	
total	48301	100.00%	--> 46.64% of all prefixes

Prefixes from transit singlehomed ASes:

standalone	2608	37.74%	
root	344	4.98%	
more specific	3958	57.28%	
total	6910	100.00%	--> 6.67% of all prefixes

Non-transit AS prefixes by type

Prefixes from non-transit multihomed ASes:

standalone	10715	34.80%
root	999	3.24%
more specifics	19080	61.96%
total	30794	100.00% --> 29.74% of all prefixes

Prefixes from non-transit singlehomed ASes:

standalone	6507	39.51%
root	452	2.74%
more specifics	9509	57.74%
total	16468	100.00% --> 15.90% of all prefixes

New top prefixes

- New address space announcements and deaggregation of existing prefixes were two major sources of new root and standalone prefixes between Nov 2000-May 2001.

Multihomed ASes and traffic engineering

Q: Are extra routes introduced for traffic engineering?

A: Yes, but only a minority

■ Multihomed ASes originating on one link

▶ Nov.01, 2001, 26 tables; multiorigin prefixes are excluded

- 22.2% of all prefixes
- 15.8% of all top prefixes
- 28.6% of all more specifics
- more specifics are announced on one link more frequently
- 6-12% more specifics may be traffic engineered splinter blocks

■ Majority of semiglobal prefixes are announced on all available connections

■ Largest groups:

- On 1 connection for singly homed AS, 22.6%
- On 2 connections for doubly homed AS, 15.8%
- On 3 connections for triply homed AS, 31.1%
- On 2 connections for triply homed AS, 7.3%

▶ (triply homed = indegree ≥ 3)

More specifics and traffic engineering

- traffic engineering would imply the more specific announced on only one connection
- some more specifics are involved in traffic engineering, but most of them aren't

Routing changes: Oct -Nov 2001

Semiglobal prefixes, 26 full size backbone tables

104555 prefixes on Oct 01

103815 prefixes on Nov 30

Union: 137374

There are lots of

- flips (withdrawals and reannouncements)
- transients (initial announcements and final withdrawals)

Flips (reannouncements after withdrawals)

	number	% of total	% of Oct30
Long-lived:	93127	67.79%	89.07%
Apparently emerged:	32819	23.89%	31.39%
Apparently disapp.:	33559	24.43%	32.10%
Transients (came&went):	22131	16.11%	21.17%
Entries w.a flip:	31007	22.57%	
Long-lived entries w.a flip:	20951	15.25%	
Long-lived entries' flips:	35591	60.06% (of flips)	
Transient entries w.a flip:	3743	2.72%	
Transient entries' flips:	9285	15.67% (of flips)	

long-lived + emerged + disappeared - transients = total observed = 137K

Flips (/24's flip more frequently; /21-24's most transients)

Pf.len	prefxs	flp.pf	flips	trans	%all	%flp.pfs	%flips	%trans
1	1	0	0	1	0.00	0.00	0.00	0.00
2	1	0	0	1	0.00	0.00	0.00	0.00
8	28	8	264	11	0.02	0.03	0.45	0.05
9	5	0	0	0	0.00	0.00	0.00	0.00
10	9	0	0	1	0.01	0.00	0.00	0.00
11	12	1	2	0	0.01	0.00	0.00	0.00
12	39	8	12	1	0.03	0.03	0.02	0.00
13	95	13	21	0	0.07	0.04	0.04	0.00
14	236	27	43	5	0.17	0.09	0.07	0.02
15	410	55	81	5	0.30	0.18	0.14	0.02
16	7387	1256	1966	115	5.38	4.05	3.32	0.52
17	1452	203	302	62	1.06	0.65	0.51	0.28
18	2605	401	581	115	1.90	1.29	0.98	0.52
19	7450	1213	1792	203	5.42	3.91	3.02	0.92
20	6852	1737	3981	356	4.99	5.60	6.72	1.61
21	7176	1193	1866	2009	5.22	3.85	3.15	9.08
22	8803	1792	3367	1169	6.41	5.78	5.68	5.28
23	11636	2600	4503	2028	8.47	8.39	7.60	9.16
24	83177	20500	40480	16049	60.55	66.11	68.31	72.52
Total	137374	31007	59261	22131	100.00	100.00	100.00	100.00

- The fraction of flips in /24 range is close to fair share

Flips in /24s per class (A,B,C)

Statistics of /24s

Class	/24s	flp.pfs	flips	%/24s	%fpg/24s	%allfp	P{fp cl}
A	11266	1890	3338	13.54	9.22	8.25	16.78
B	6666	2034	4587	8.01	9.92	11.33	30.51
C	65245	16576	32555	78.44	80.86	80.42	25.41
Total	83177	20500	40480	100.00	100.00	100.00	24.65

Flips per lifetime of a prefix

#) AS	Av.#pf	Flips	Fp/Kupt	Fp/Klft	Pc.fp	Acc.pc	NewAnn	TmpW/drw	Reann	FinW/hdr
1) 4323	421.72	1466	4.89	4.59	2.47	2.47	34	52	51	16
2) 752	3.60	1444	564.72	335.58	2.44	4.91	1	659	658	2
3) 701	2196.94	1379	0.88	0.82	2.33	7.24	101	351	350	91
4) 702	956.48	924	1.36	1.32	1.56	8.80	68	373	346	45
5) 3908	652.29	921	1.99	1.79	1.55	10.35	63	105	109	43
6) 705	397.04	661	2.34	2.18	1.12	11.47	62	162	165	46
7) 1580	180.75	613	4.77	4.36	1.03	12.50	5	147	144	7
8) 3475	28.91	540	26.27	21.79	0.91	13.41	23	237	238	22
9) 4755	308.19	510	2.33	2.20	0.86	14.27	31	195	180	31
10) 8708	161.29	497	4.33	4.14	0.84	15.11	21	52	59	19
18) 8112	48.03	378	11.07	8.71	0.64	20.58	7	7	7	4
27) 724	244.60	278	1.60	1.56	0.47	25.14	12	76	81	11
39) 1239	692.89	190	0.39	0.38	0.32	30.21	55	94	90	33
56) 13609	109.98	150	1.92	1.90	0.25	35.03	14	40	40	4
80) 18618	64.45	114	2.49	2.45	0.19	40.17	1	10	9	2
110) 7629	21.04	83	5.55	5.40	0.14	45.10	7	22	23	6
150) 11623	6.90	66	13.46	13.26	0.11	50.01	0	12	12	0
201) 9800	104.11	52	0.70	0.70	0.09	55.07	6	18	17	2
266) 11554	18.78	39	2.92	2.90	0.07	60.01	0	3	3	1
356) 17625	6.84	29	5.97	4.97	0.05	65.05	4	7	9	3
399) 10223	42.14	25	0.83	0.83	0.04	67.01	1	10	10	2
478) 600	89.94	20	0.31	0.31	0.03	70.01	1	17	17	1
654) 6539	239.63	15	0.09	0.09	0.03	75.02	6	10	10	3
902) 224	14.94	10	0.94	0.94	0.02	80.00	0	10	10	1
1262) 2508	1.16	7	8.47	4.92	0.01	85.01	0	7	7	0

Total origins of flips: 5259 AS. Sum of temp. w/draw AS events: 22657. Reannouncement: 22816

All AS initiating temp. withdrawals: 5259. Reannouncing AS: 5259

Note: when a prefix changes origin, an interval of AS lifetime sum is lost

Several origin ASes over time

#)	#or.AS	# Prefixes
1)	1	130786
2)	2	4901
3)	3	1500
4)	4	96
5)	5	50
6)	6	16
7)	7	17
8)	8	5
9)	9	1
10)	11	1
11)	12	1
	Total	137374

- 6588 prefixes changed hands (origin or some of origins)
- 4220 prefixes were multiorigin (at least once)
- 3861 (91.49% of above) multiorigin prefixes also changed hands

multi-origin prefixes are highly dynamic, some may be artifacts of convergence to a new origin at snapshot time

BGP table change contributors (summary)

- ASes with new and/or (temporarily) lost prefixes
 - sum total of all changes in prefix status
- 50% of (all) changes from 127 most "dynamic" ASes
- Sorted by total lost+found prefix occurrences
- for Oct-Nov 2001 (61 days, 2 hour int, 711 tables)

BGP table change: top 10 contributors

type	ASN	Av.#pf	ap+ds	2-hour intervals with changes				new app	tmp disp	reapp	final disp
				a+d /Kupt	a+d /Klft	%a+d	cum %a+d				
1 Bac	701	2196.94	6790	4.35	4.04	3.67	3.67	101	351	350	91
2 Bac	3908	652.29	5915	12.75	11.49	3.20	6.87	63	105	109	43
3 AOL	4323	421.72	3236	10.79	10.14	1.75	8.62	34	52	51	16
4 Bac	1239	692.89	3199	6.49	6.47	1.73	10.35	55	94	90	33
5 Res	752	3.60	2893	1131.4	672.32	1.56	11.92	1	659	658	2
6 Bac	702	956.48	2883	4.24	4.12	1.56	13.48	68	373	346	45
7 Tech	71	46.22	2589	78.79	77.60	1.40	14.88	3	10	10	6
8 Bac	705	397.04	2409	8.53	7.94	1.30	16.18	62	162	165	46
9 Bac	4999	21.64	1493	97.02	28.93	0.81	16.99	6	2	2	6
10 Mil	1580	180.75	1448	11.27	10.30	0.78	17.77	5	147	144	7

(ap = appeared, ds = disappeared)

- Both background noise and storm-like events are present

BGP table change contributors: prefix flips (reannouncements after withdrawals)

- 50% flips from 150 ASes (3% of 5259 flipping ASes)
- 75% flips from 12.4% ASes
- #events for AOL much larger than for other ASes
 - not just background noise

type	ASN	Av.#pf	Flips	Fp/Kupt	Fp/Klft	Cum		newap	tmpdisp	re-app	finaldisp
						%fps	%fps				
1 AOL	4323	421.72	1466	4.89	4.59	2.47	2.47	34	52	51	16
2 Res	752	3.60	1444	564.72	335.58	2.44	4.91	1	659	658	2
3 Bac	701	2196.94	1379	0.88	0.82	2.33	7.24	101	351	350	91
4 Bac	702	956.48	924	1.36	1.32	1.56	8.80	68	373	346	45
5 Bac	3908	652.29	921	1.99	1.79	1.55	10.35	63	105	109	43
6 Bac	705	397.04	661	2.34	2.18	1.12	11.47	62	162	165	46
7 Mil	1580	180.75	613	4.77	4.36	1.03	12.50	5	147	144	7
8 Mil	3475	28.91	540	26.27	21.79	0.91	13.41	23	237	238	22
9 Dev	4755	308.19	510	2.33	2.20	0.86	14.27	31	195	180	31
10 Dev	8708	161.29	497	4.33	4.14	0.84	15.11	21	52	59	19

- Under best conditions, a prefix flips once in 3 months
- Aside: global (all-peer) prefixes have twice as much flux as semiglobal (flux estimates depend on the prefix set)

analysis of BGP updates

- now we look at granularity of **individual BGP updates**
- RouteViews started archiving October 2001
- we analyze BGP updates over one month (Nov 2001)
 - 15 min updates per file, 22 peers contribute
 - Contribution sizes span an order of magnitude
 - use 50K cutoff to avoid events like full-table transmission
 - leaves about 90% of all time intervals (2434 out of 2711)
 - and about 2/3 of all packets

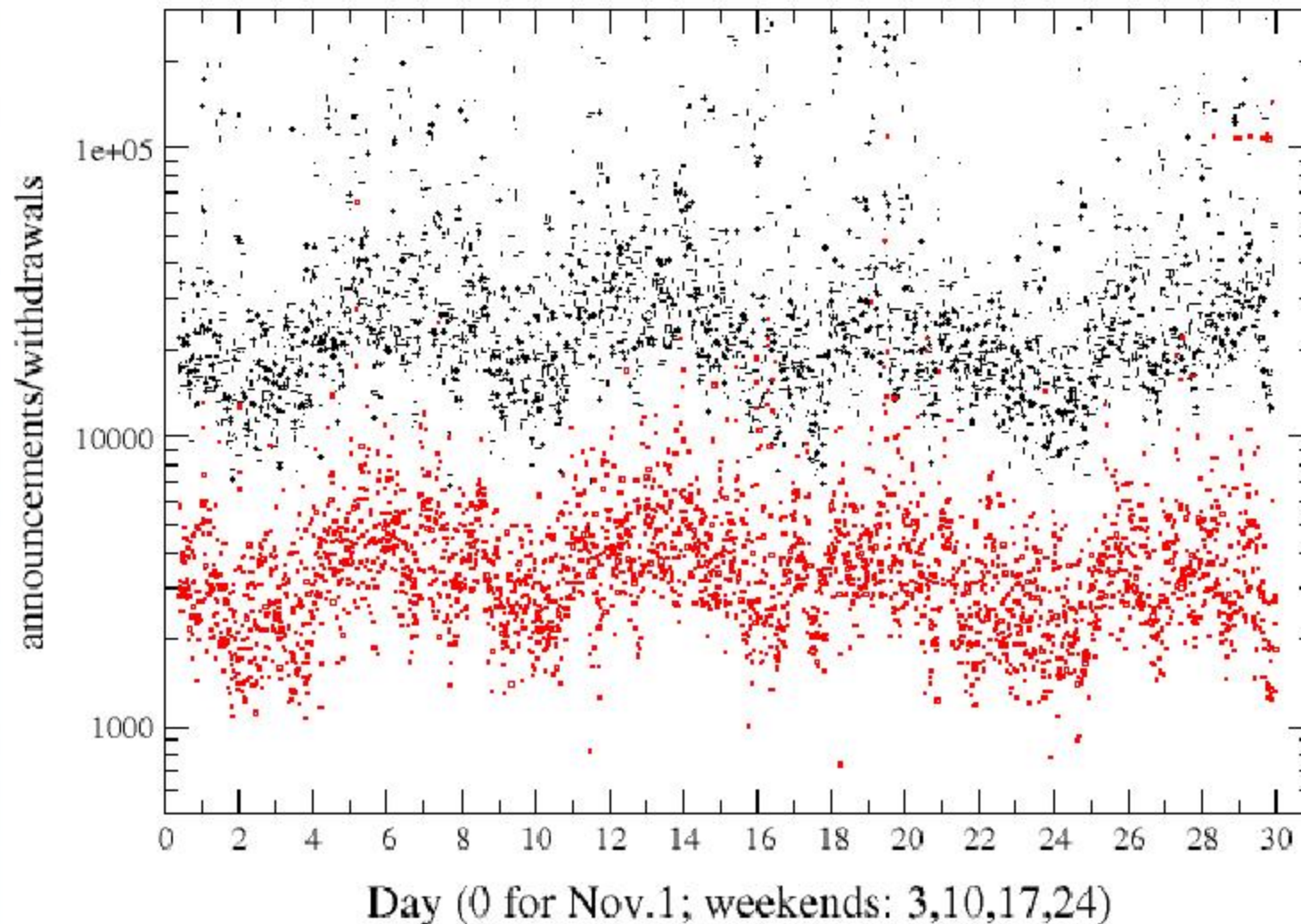
Results:

- Announcements more frequent than withdrawals
- Many whole table transmissions, $O(100K)$ prefixes
- Weekly and diurnal variations clearly visible
- Match traffic pattern on the 'Net
- Traffic load translates into path fluctuations
 - Is it human intervention or automatic tools output?

BGP updates over one month

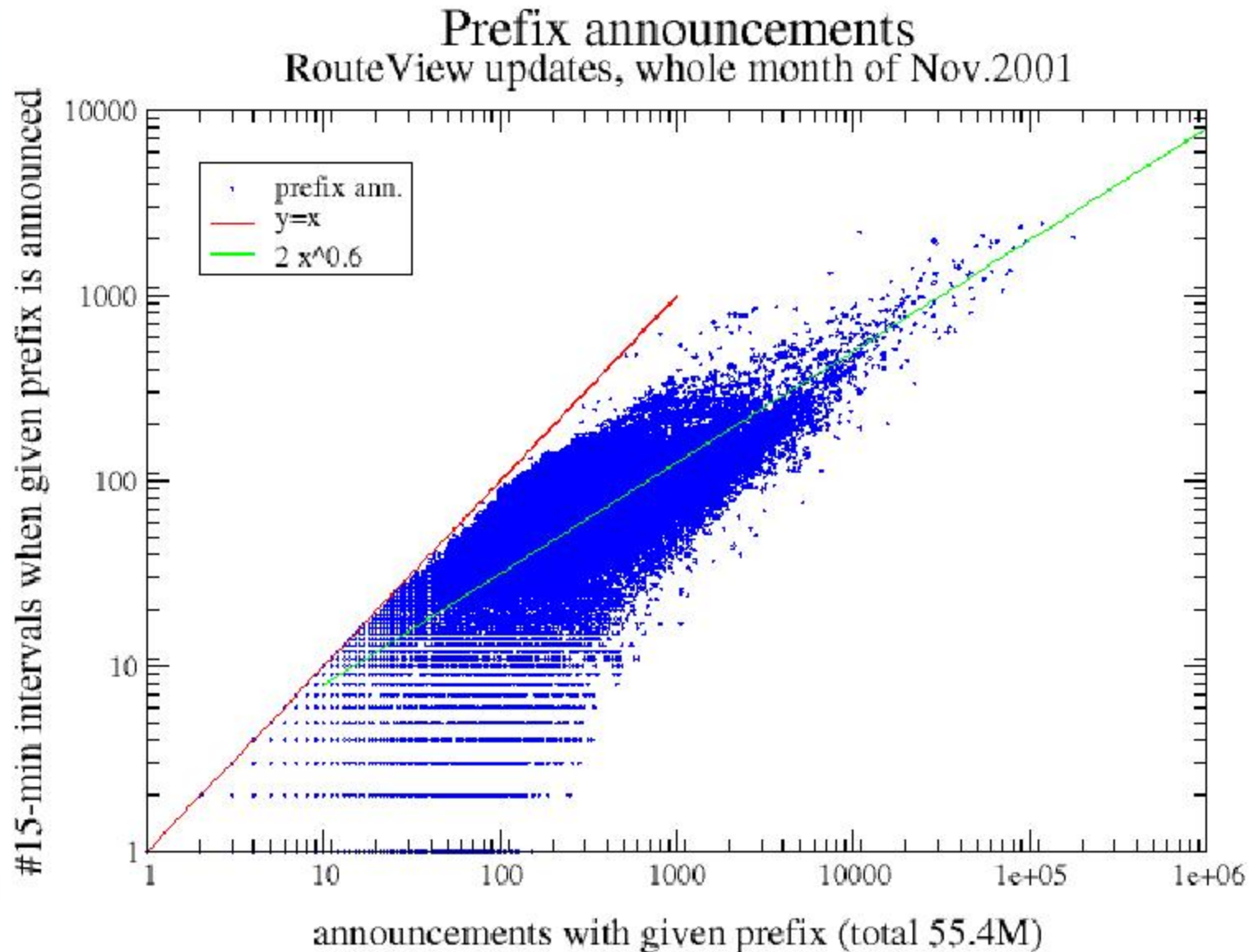
- withdrawal vs announcements (and some full table updates)

Prefix announcements from Route View peers
November 2001, 15-min intervals



BGP prefix announcements

- ann/prefix related to counts of 15-min.int, esp. at the cusp



BGP prefix announcements/withdrawals

- Nov 2001 RouteViews
- 2711 files (no cutoff for #announcements per file)

total packets:	30M
packets with announcements	28.2M
packets with withdrawals	1,955,188
packets with both	127,429
total announcements	83.5M
total withdrawals	11.7M
average announcements/packet	2.7736
average withdrawals/packet	0.3874
average withdrawals/wd.packet	5.9615

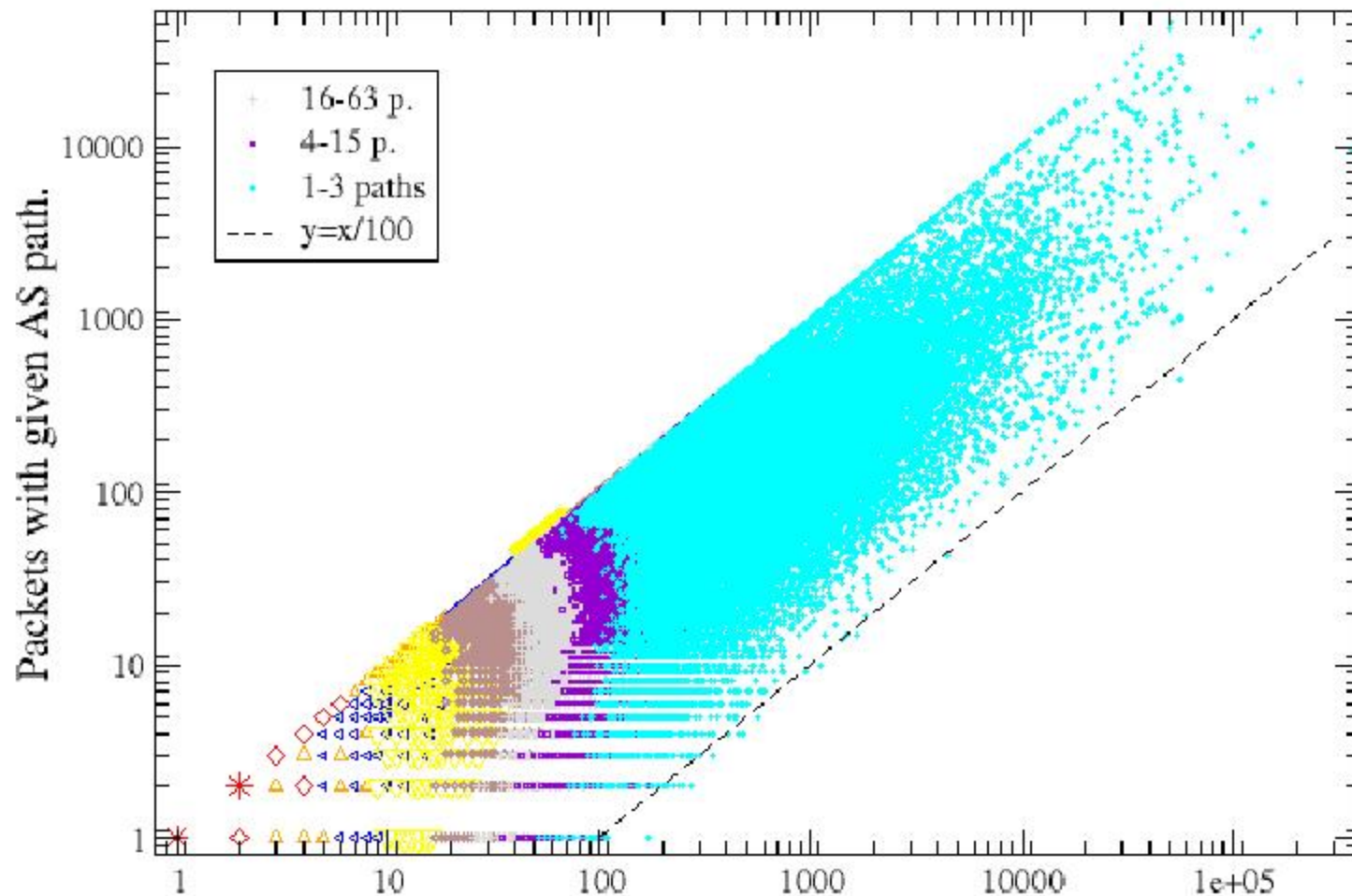
- 15X as many announcement as withdraws
- 15X as many withdraws as ann.&withdraw
- withdrawals packed at prefixes/AS path ratio of ~6.5

Dependence between packets seen & bgp announcements

- some AS paths update 100s of prefixes at once

Packets and announcements with a common AS path

Route Views BGP updates, full month of November 2001

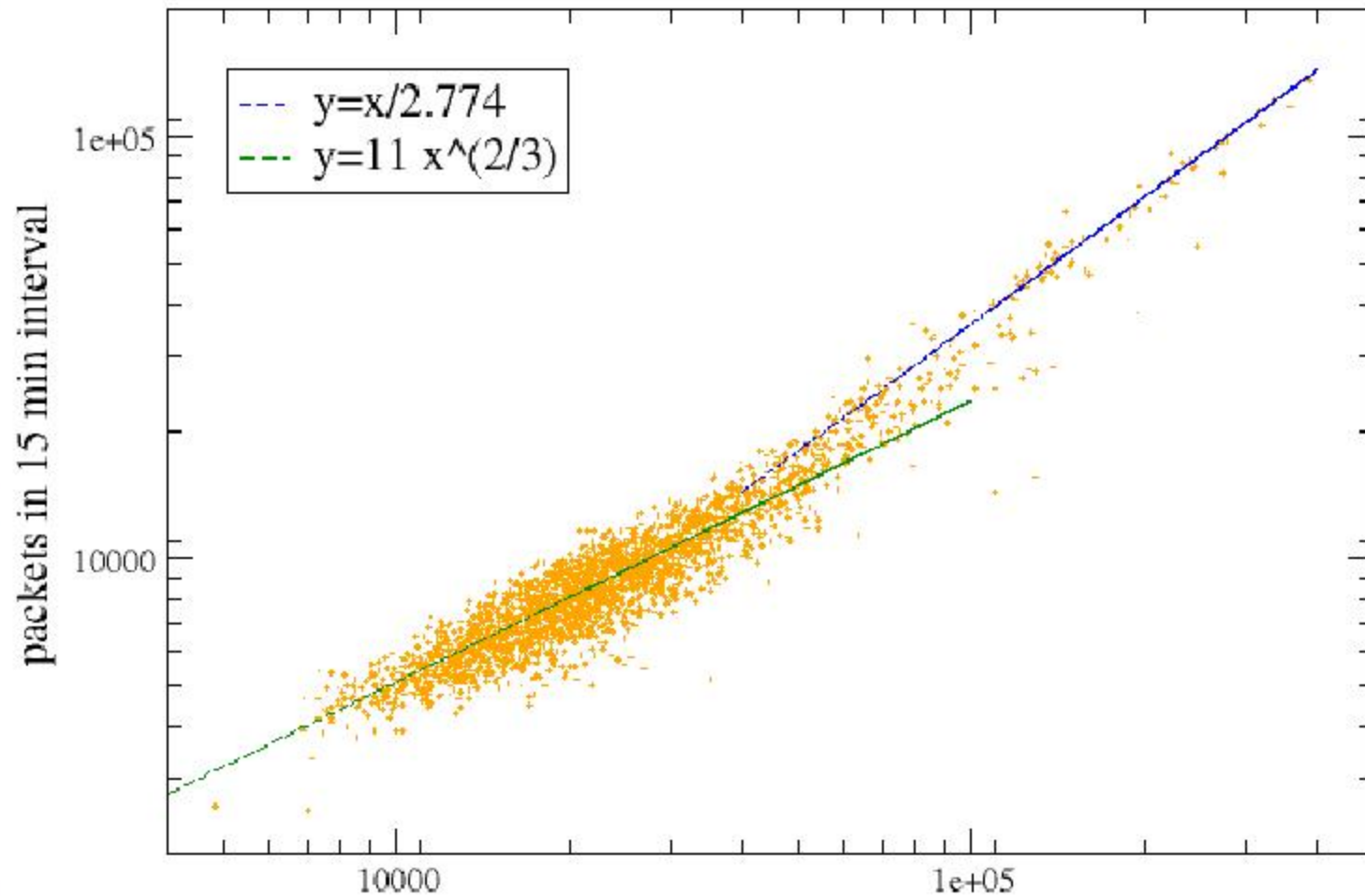


Announcements with given AS path. Total 1,116,845 paths w.prepending

packing of prefixes into announcements

- two regions: full tables (blue) and incremental updates (green)

Prefix announcements and packet counts
RouteView updates, whole month of Nov.2001



announcements in 15 min interval (total 55.4M). 2711 intervals

BGP updates contributors

Q: Are small ASes major contributors of BGP updates?

A: The bulk of BGP announcements comes from ISPs who announce many (tens to hundreds) prefixes:

- backbone providers
- government networks
- ISPs in developing nations
- ISPs in Canada and US

Large batches of announcements often contain deaggregated intervals of IP space (as in "avoiding a small block" scenario)

Top 10 announcers, Nov 2001

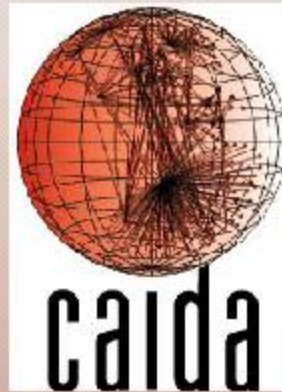
type	ASN	#Anns	%ann	cum%	/15mi
1 ISP Can	852	777540	1.40	1.40	732
2 Dev Eu	12302	578932	1.05	2.45	681
3 Mil	6595	564201	1.02	3.47	1622
4 Bac	577	508675	0.92	4.39	1085
5 Dev Asia	4755	471458	0.85	5.24	1985
6 Bac	702	436282	0.79	6.02	2421
7 Dev SAm	20043	412334	0.74	6.77	749
8 Gov ZA	2018	370335	0.67	7.44	337
9 Gov	2572	365096	0.66	8.10	592
10 Bac	6453	346551	0.63	8.72	2442

- top 2 announce 538 and 102 prefixes, resp. (backbones up to 2K)
- announcements are spread across time, not just storms
- reminder: only 1% of ASes announce >100 prefixes

Conclusions

- Internet undergoes refinement process
- Some bulk measures are stable
- Others change slowly
- Opposite trends cancel each other
- More churn than bulk measures suggest
- Internet churn comes from everywhere
 - mostly from large ISPs, .gov/.mil, and developing nation telecoms
- Small ASes generate little churn and table growth
- Invariant or almost invariant relations do exist
 - e.g. AS refinement formula

still a lot left to analyze...



www.caida.org/outreach/presentations/

Andre Broido
kc claffy
UCSD/SDSC/CAIDA
kc@caida.org
www.caida.org