

# A Study of BGP Origin AS Changes and Partial Connectivity

Ratul Mahajan<sup>†</sup>

David Wetherall<sup>† ‡</sup>

Tom Anderson<sup>‡ †</sup>

<sup>†</sup> University of Washington

<sup>‡</sup> Asta Networks

# Goals

- Long-Term: What is the extent and impact of configuration errors in BGP?
  - incorrect origin AS, partial connectivity, pvt. ASN, pvt. address space, looping AS-paths, .....
- This talk:
  - Origin AS Changes
  - Partial Connectivity

# Origin AS Changes

- Why does the origin for a prefix change?
  - How many changes are short-lived?
  - How many changes are a result of misconfiguration?
  - How many changes lead to reachability problems?
- Easy ways to detect misconfigured origins?
  - Multiple origins for a prefix
    - misses subset space hijack
    - increasingly common practice
  - IRR's
    - are they accurate?

## IRR: Simple Way to Detect Incorrect Origins?

BGP Table Snapshot: Sep 28, 2001

	<b>Total Prefixes</b>	<b>Registered Origins</b>	<b>Consistent Origin(s)</b>	<b>Inconsistent Origin (s)</b>
<b>Single Origin AS</b>	115228	101952	70458 (69%)	31494 (31%)
<b>Multiple Origin AS's</b>	1720	1523	293 (19%)	1230 (81%)

Verified using RADB, RIPE, APNIC, ARIN

# Origin Changes: Methodology

- Use BGP snapshots archived by Route Views
  - caveat: would miss origin changes that come and go between snapshots
- Identify prefixes that are not announced by the same origin(s) throughout the day
  - includes prefixes not present in all snapshots
- Attribute a cause to every origin change

# Classification of Origin Changes

*Long-lived:* Changes that seem permanent

*Fluctuating:* Short-lived changes with correct origins

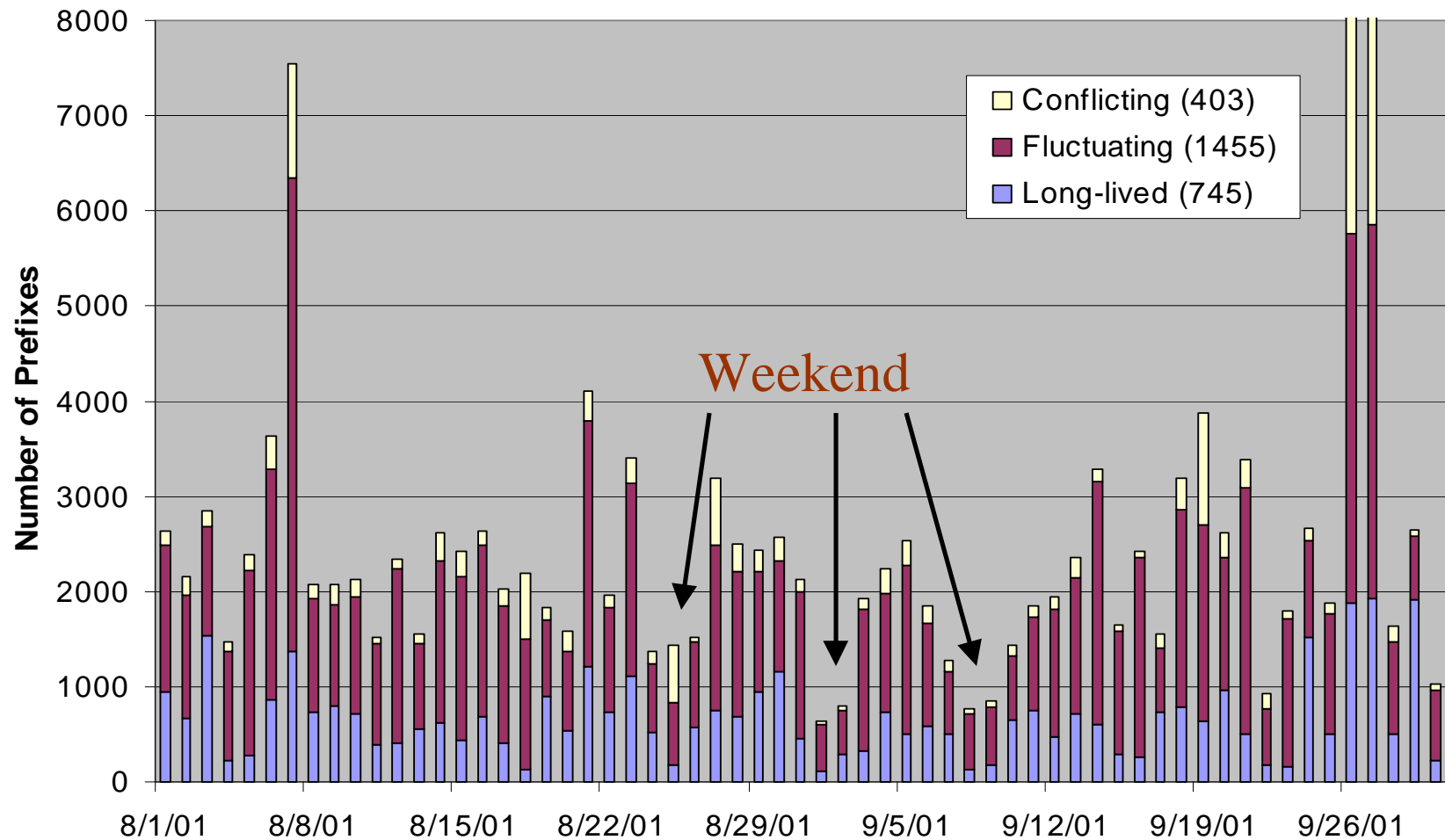
*Conflicting:* Short-lived changes with potentially incorrect origins

<b>Long-lived</b>	<b>Fluctuating</b>	<b>Conflicting</b>
More Specific Added	Self Deaggregation	AS-Path Stripping
More Specific Deleted	Failures (unreachable)	Strip Deaggregation
Origin Added	Backups	Extra Last Hop
Origin Deleted		Foreign Deaggregation
Origin Changed		Other
New Address Space		
Address Space Deleted		

# Glossary for Short-Lived Changes

	Stable Announcements		Short-lived Announcements	
<b>Self Deaggregation</b>	a.b.0.0/16	X-Y-Z	a.b.c1.0/24 a.b.c2.0/24	X'-Y'-Z X'-Y'-Z
<b>AS-Path Stripping</b>	a.b.c.d/s	X-Y-Z	a.b.c.d/s	X'-Y
<b>Strip</b>	a.b.0.0/16	X-Y-Z	a.b.c1.0/24	X'-Y
<b>Deaggregation</b>			a.b.c2.0/24	X'-Y
<b>Extra Last Hop</b>	a.b.0.0/16	X-Y-Z	a.b.c1.0/24 a.b.c2.0/24	X'-Y'-Z-O X'-Y'-Z-O
<b>Foreign Deaggregation</b>	a.b.0.0/16	X-Y-Z	a.b.c1.0/24 a.b.c2.0/24	X'-Y'-O X'-Y'-O

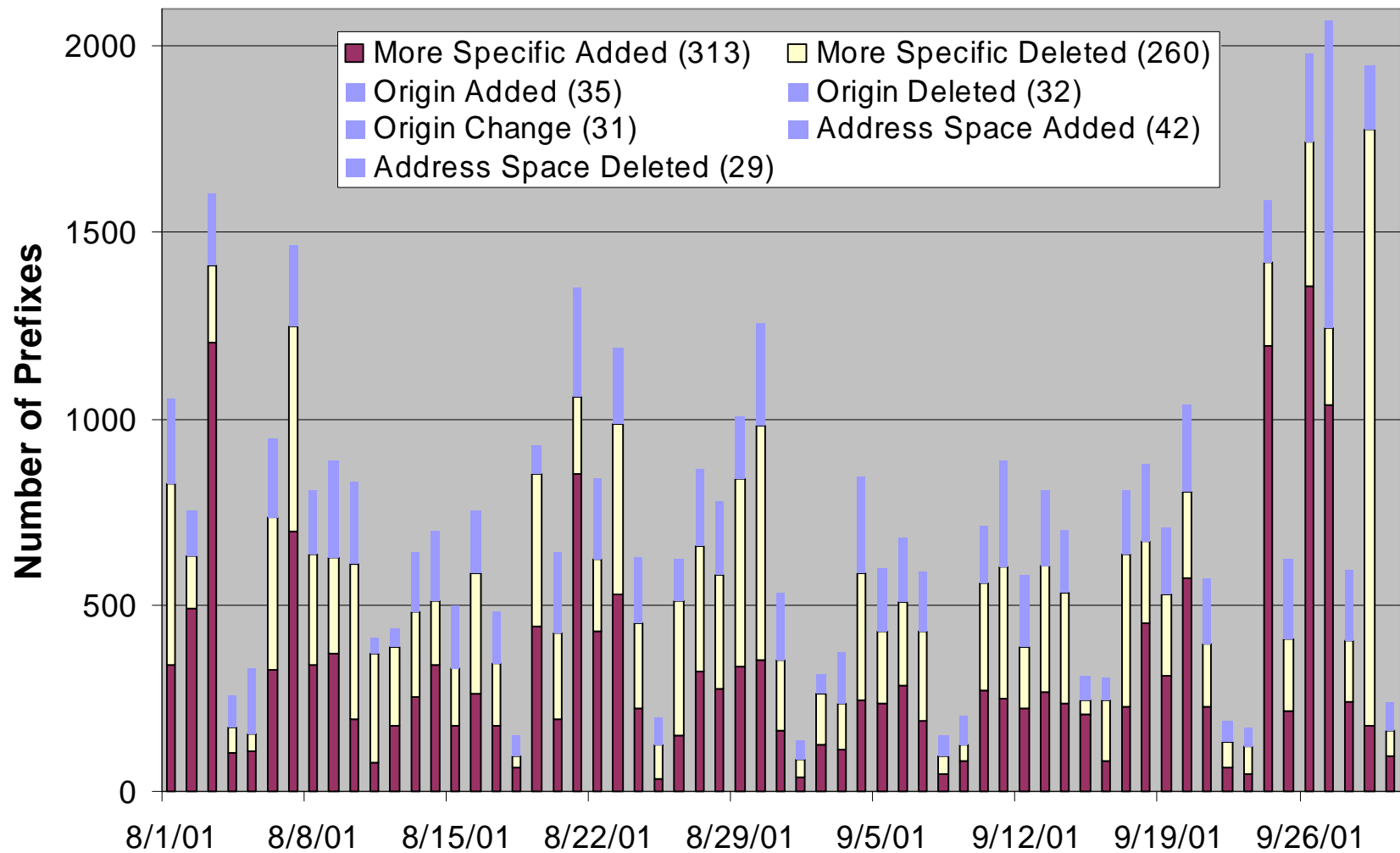
## Distribution of Origin Changes



1. More than 2% of the prefixes experience a change
2. Less than a third of changes are long-lived
3. Weekly pattern in the number of changes seen

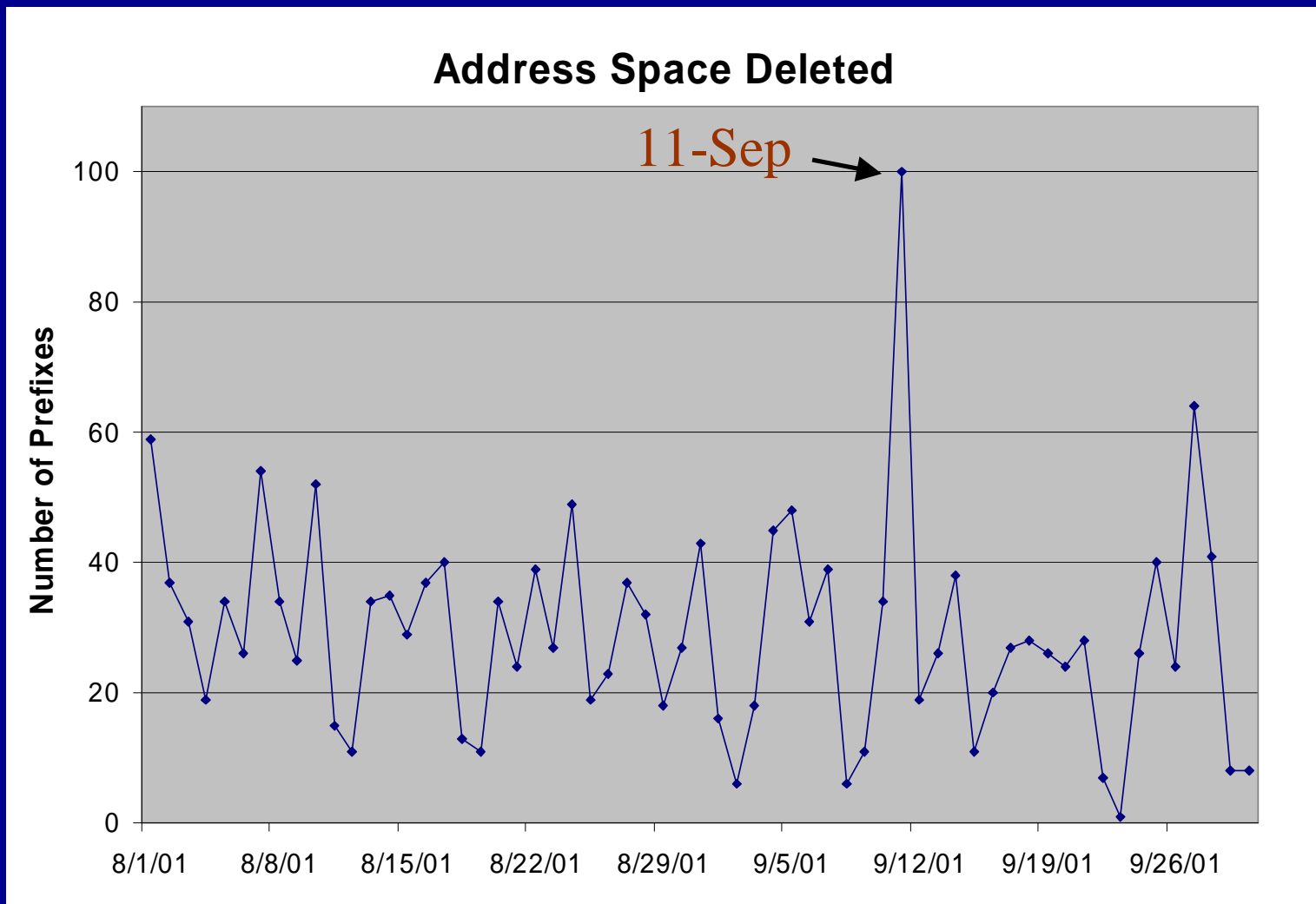


## Breakdown of Long-Lived Changes

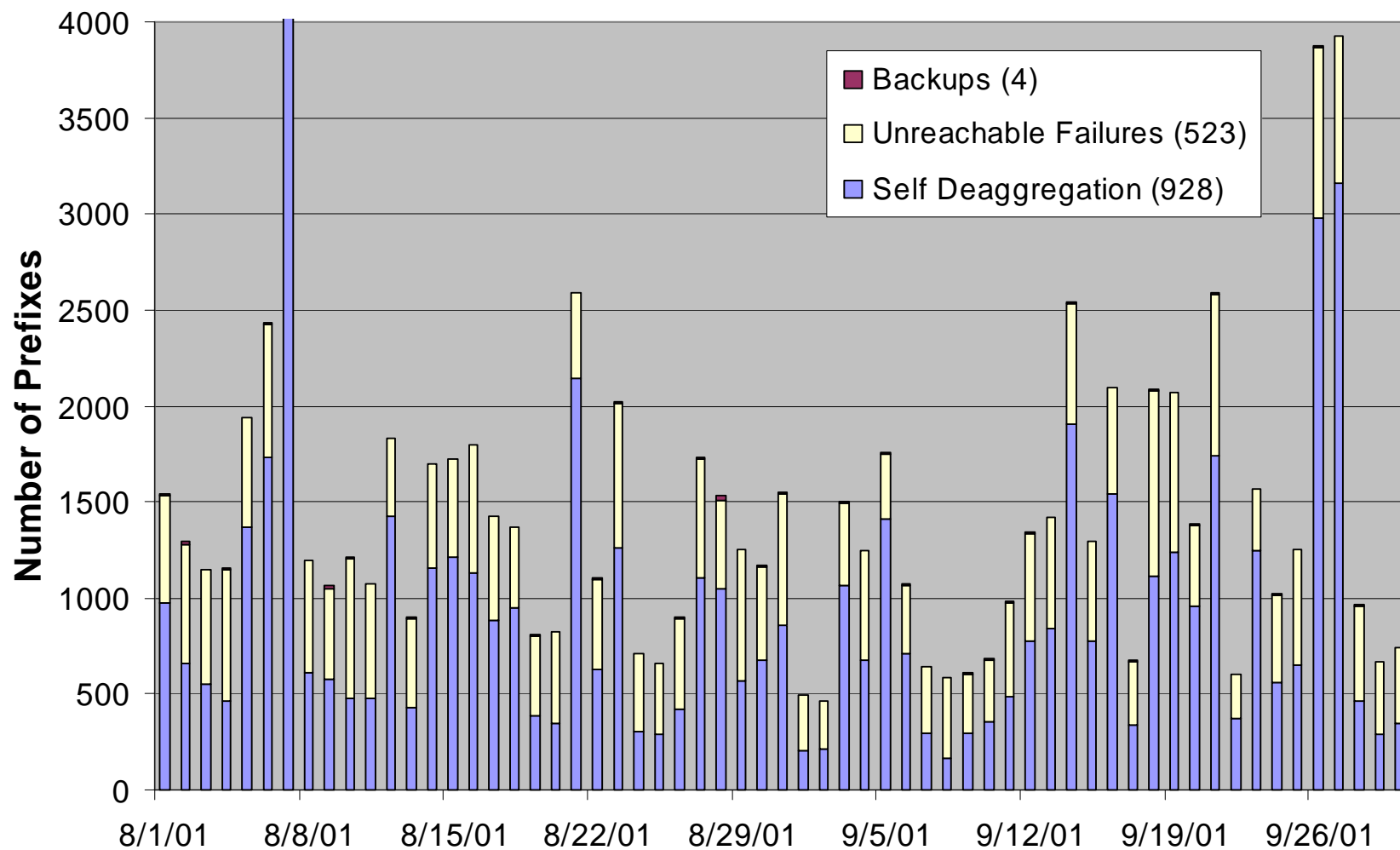


1. 85% of long-lived changes persist beyond a week
2. Most action is in more-specifics (added,deleted)

# 9/11, As Seen by Origin Changes

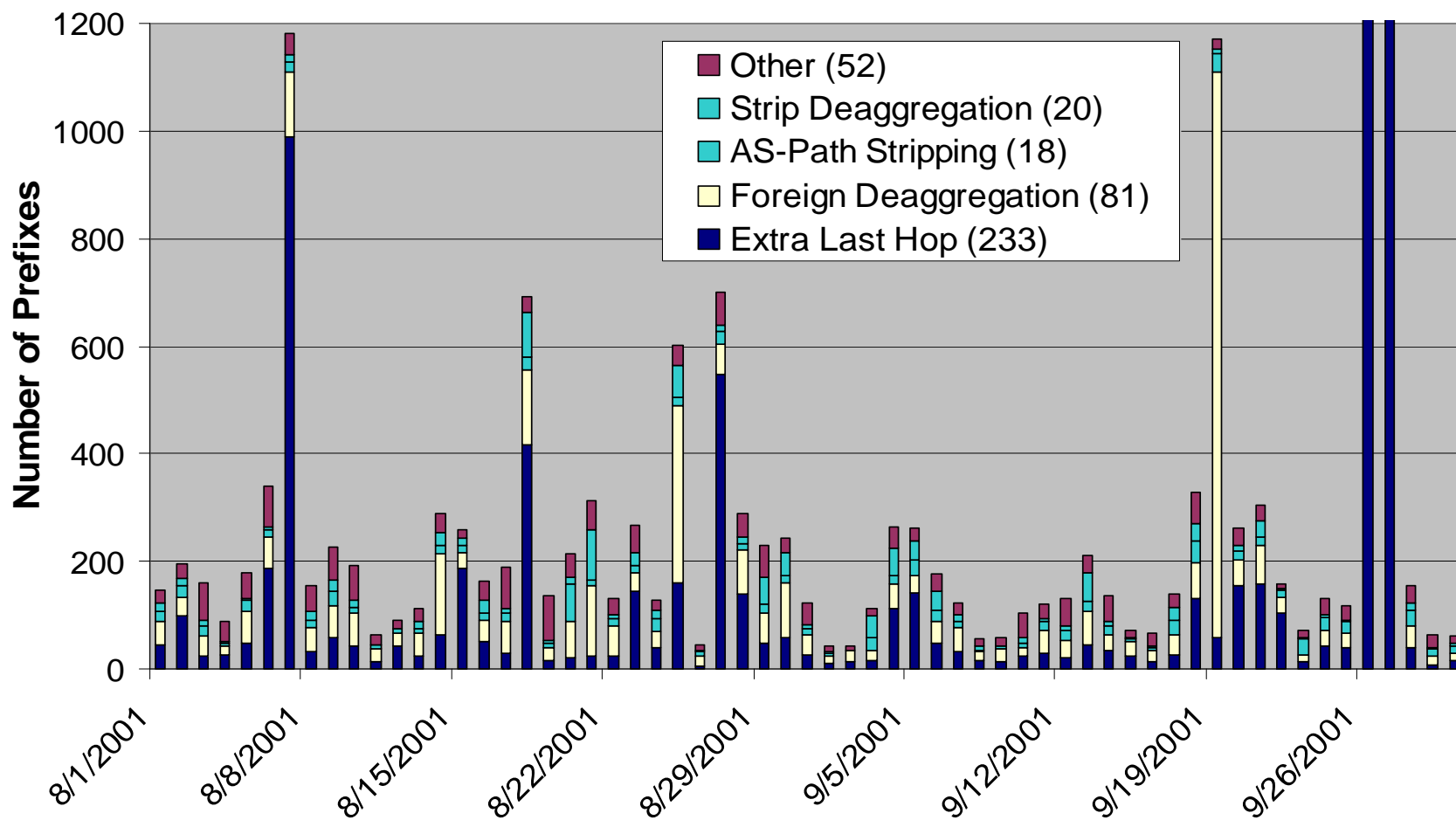


## Breakdown of Fluctuating Changes



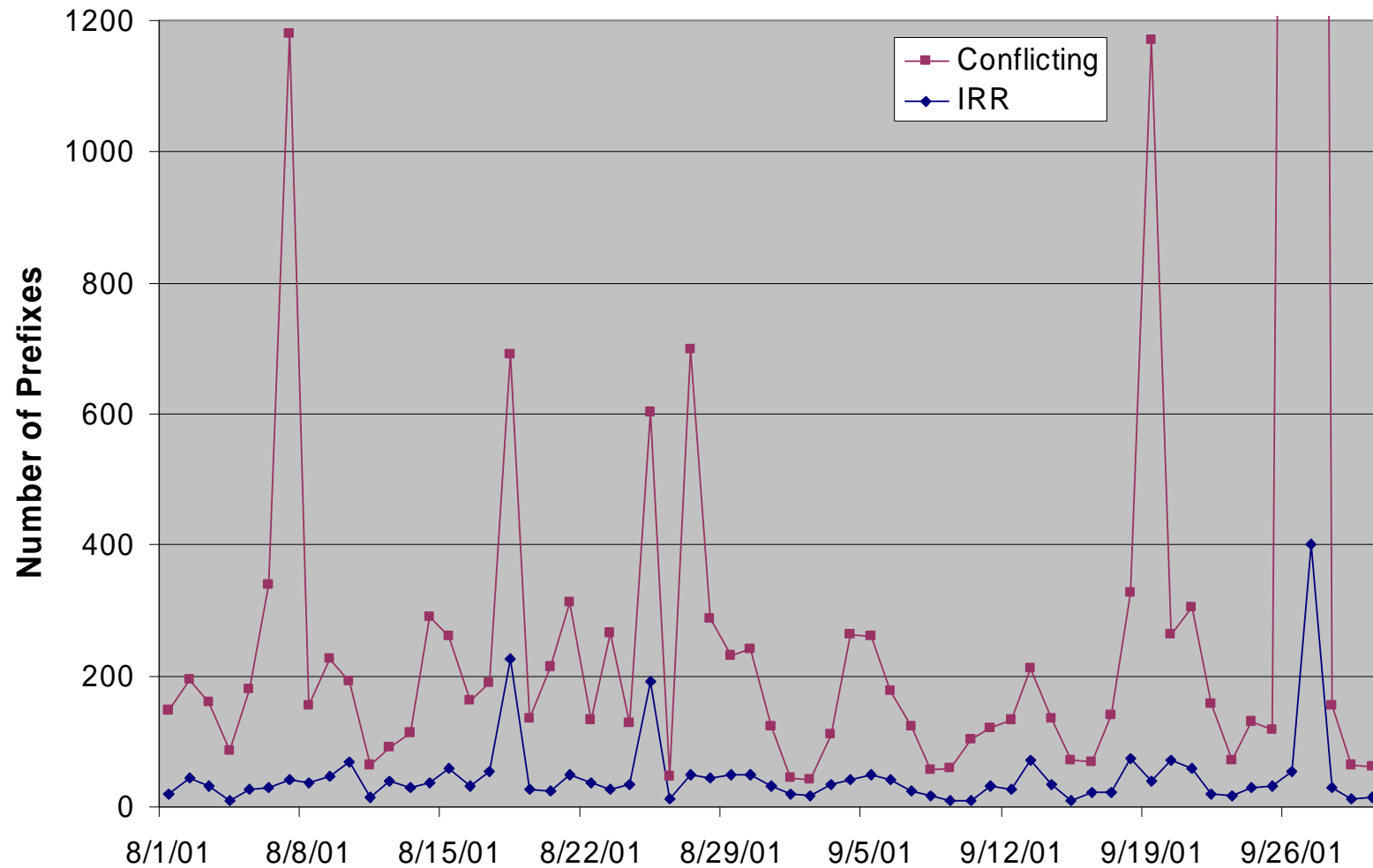
Self Deaggregation causes more origin changes than failures

## Breakdown of Conflicting Changes



Deaggregation in general, and extra last hop in particular, causes most changes

## How effective is IRR during conflicts?



Consulting the IRR when you see conflicts does not help

## Some Examples of Misconfigurations

- Small AS's announce /8's (61.0/8, 62.0/8, ...)
- An AS in Romania originated routes for most of Romania but NO reachability problems
- AS's accept their own deaggregated address space adverts, and pass them on

Not all origin misconfigurations cause reachability problems. How to figure out which ones do?

# Reachability Test

1. Download the current BGP table
2. Identify the announcements with new origins
3. Divide the AS's into two sets
  - *loyals*: AS's that believe the old origin
  - *converts*: AS's that believe the new origin
4. Use public looking glass servers to check if one set can reach the prefix while the other cannot

# Reachability Test: Possible Results

- Pass
  - both sets can reach the prefix, or
  - both get blocked at the *same place* in the network
- Inconclusive
  - both sets cannot reach the prefix, and get blocked at different places in the network
- Fail
  - one set can reach the prefix, while the other cannot
  - loyals win or converts win



# Reachability Test: Initial Results

	<b>Total</b>	<b>Pass</b>	<b>Inconclusive</b>	<b>Fail: Loyals Win</b>	<b>Fail: Converts Win</b>
<b>Long-lived</b>	837	755	52	18 (2.1%)	12
<b>Fluctuating</b>	117	103	8	3 (2.5%)	3
<b>Conflicting (except Other)</b>	206	186	15	3 (1.4%)	2
<b>Other</b>	91	69	11	11 (12.0%)	0

# Partial Connectivity

- Advertised address space not reachable from all places in the Internet
- Causes:
  - convergence delays
  - route flap damping and failures
  - policy
    - filtering (prefix length, commercial relationships)
- Failures should not lead to partial connectivity by themselves

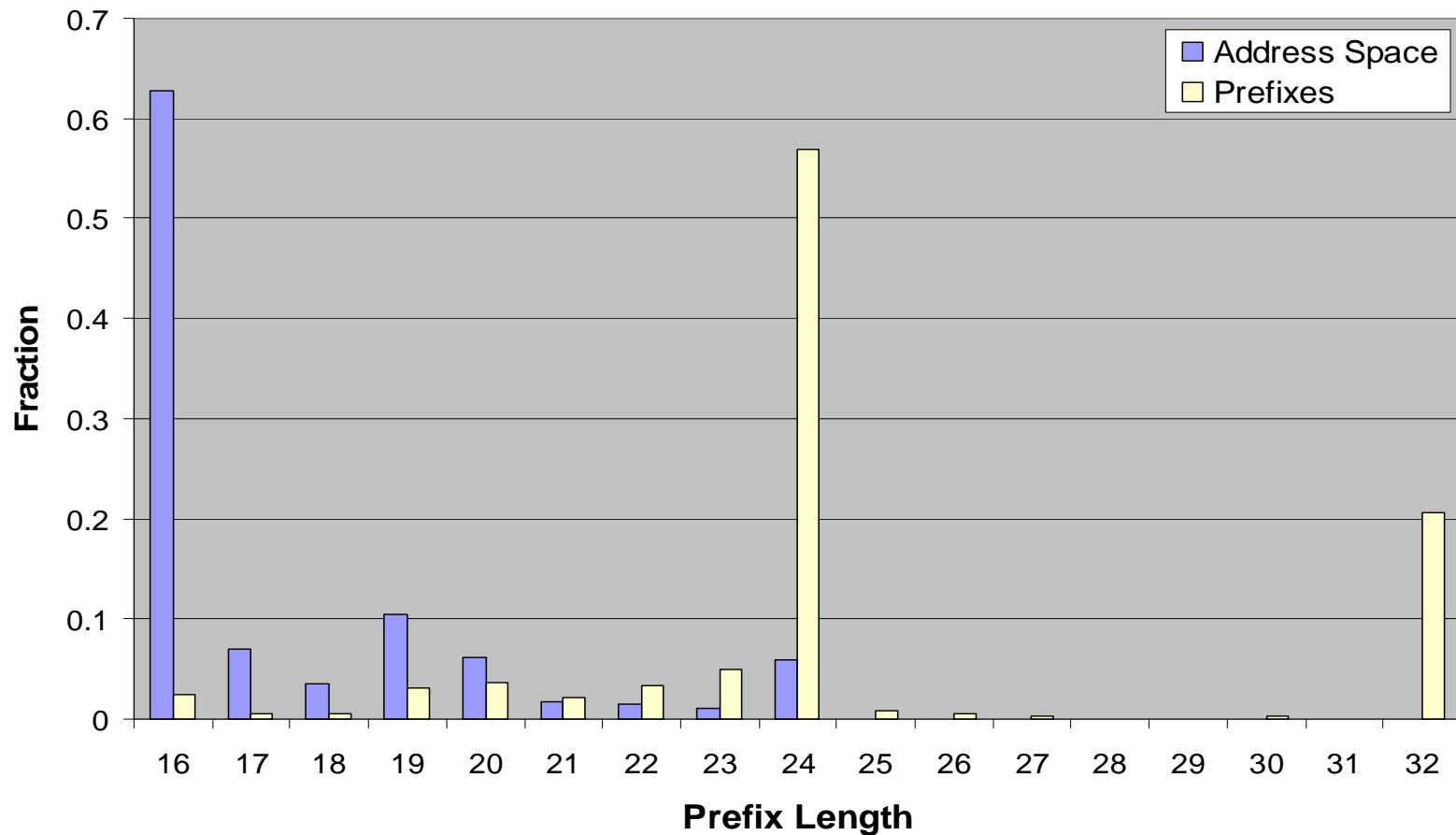
# Partial Connectivity: Methodology

- Identify partially connected address space ( $\neq$  prefix) from the BGP table
- Consult snapshots 15 minutes before and after to identify partial connectivity due to convergence delays
- Correlate across days to identify policy based partial connectivity
- Verify using public looking glasses to guard against restrictive export policies and default pointing

# Partial Connectivity: Results

- Expressed as % of advertised address space:
  - convergence: 0.005-0.02%
  - flap damping and failures: 0.1-0.8%
  - policy: 0.7%

## Prefix Length Distribution of Policy-based Partially Connected Address Space



Most partially connected prefixes are /24's

Most partially connected address space is due to /16's

## Conclusions

- More than 2% of the prefixes experience an origin change during the day
- Less than a third of the changes are long-lived
- Only a small fraction of the changes lead to reachability problems
- 0.7% of address space is partially connected due to filtering policies

Feedback:

<http://www.cs.washington.edu/homes/ratul/bgp/>  
<*ratul@cs.washington.edu*>