

On the structure and application of BGP 'Policy Atoms'

Yehuda Afek

Omer Ben-Shalom

Anat Bremler-Barr

Tel-Aviv University



What are BGP policy atoms

- BGP atoms are a possible mid level aggregate of IP space suggested by Andre Broido and kc claffy from CAIDA
 - Subnets/prefixes
 - BGP Published subnet (CIDR/prefixes)
 - **Atoms**
 - ASs (Autonomous systems)



Why BGP policy atoms ?

- A higher level aggregate over the prefix level can reduce complexity
 - Measurements
 - Processing/Memory
- Created by policy
 - Can help us understand Internet Policy



Lecture Outline

- Atom definition
- Atom calculation methods
- Are atoms stable ?
- Atom Correlation to BGP updates
- Where are atoms created ?
- Using Atoms




BGP 'reminder'

- BGP is the 'Internet' routing protocol
 - Routes traffic between Autonomous system (AS)
- BGP Is a distance vector protocol
 - Uses AS hop count as the vector
 - Keeps the whole path to avoid loops (AS_path attribute)



BGP ‘reminder’

- AS is the ‘base’ unit
- All traffic to the same destination AS should follow the same AS path
- Enforcing policy
 - Per prefix attributes (local pref etc)
 - Selective blocking of advertisement



multiple AS path to prefixes in the same AS

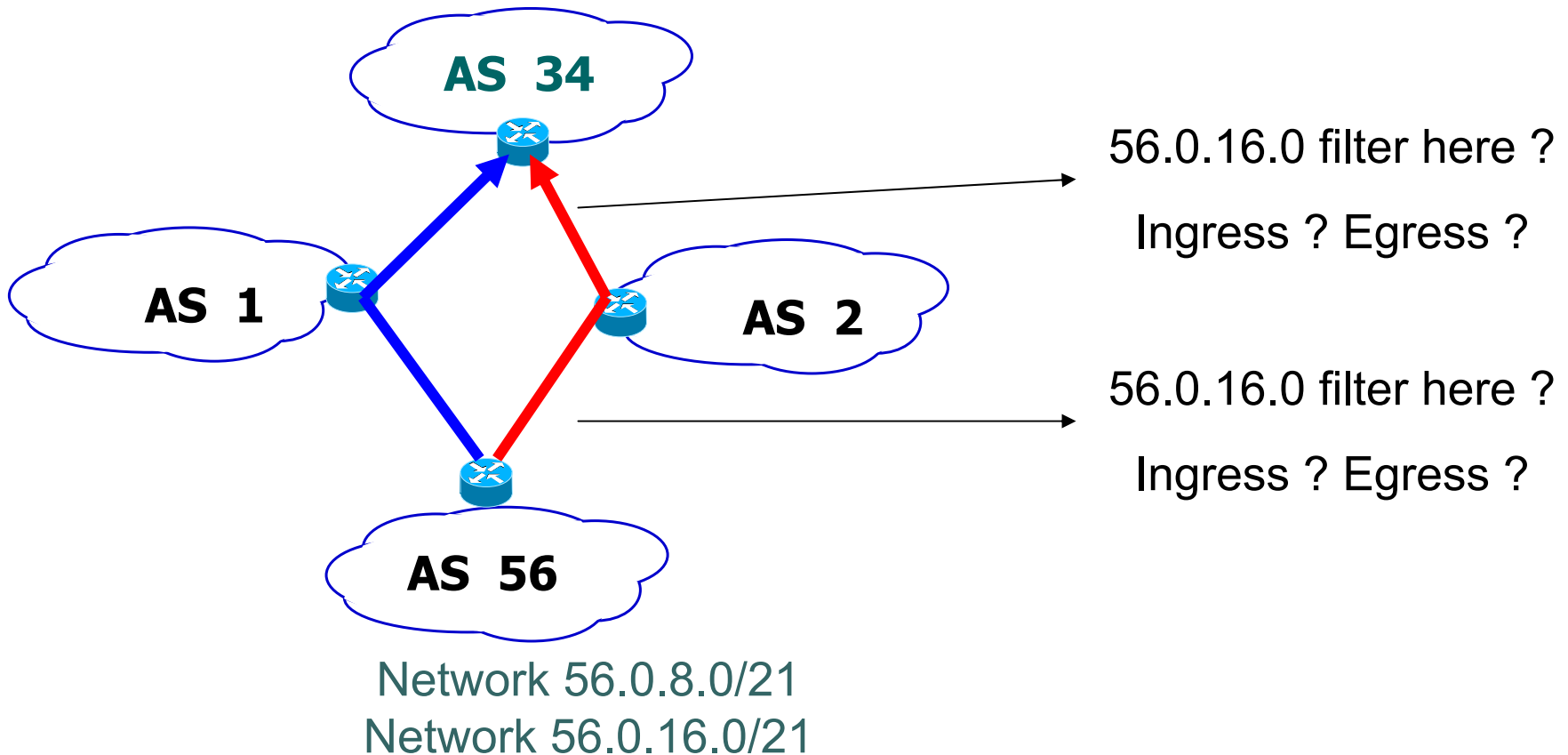
- BGP tables contain different AS path to prefixes in the same destination AS

Prefix	AS path
123.45.68.0/21	12 34 56
123.45.76.0/21	12 34 56
123.45.84.0/21	14 45 56

Finding out where policy was set is not easy !

AS 34 BGP table:

Prefix	AS path	Preference
56.0.8.0/21	2 56	200
	1 56	100
56.0.16.0/21	1 56	0

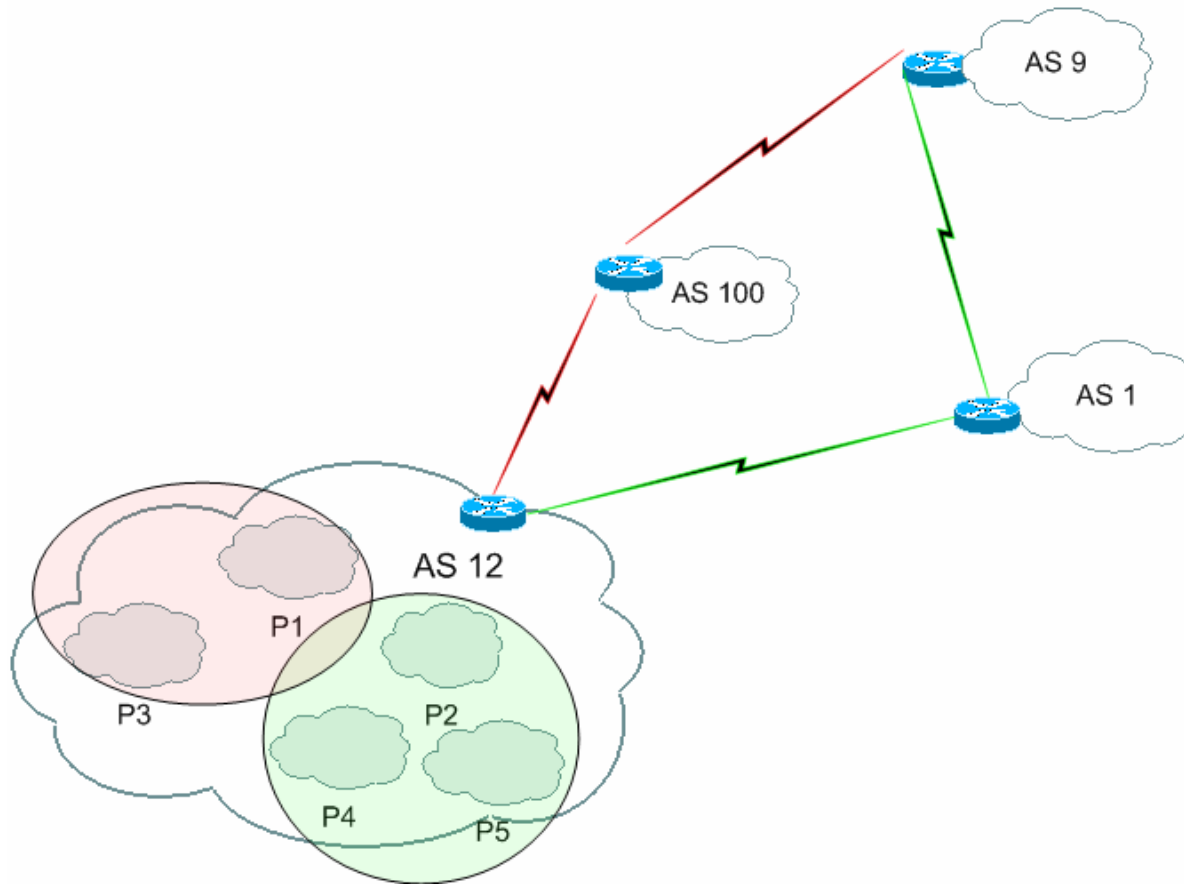




Prefix grouping by AS_path

- group prefixes with same BGP AS path on a single router

View of a single router in AS 9 – two atoms



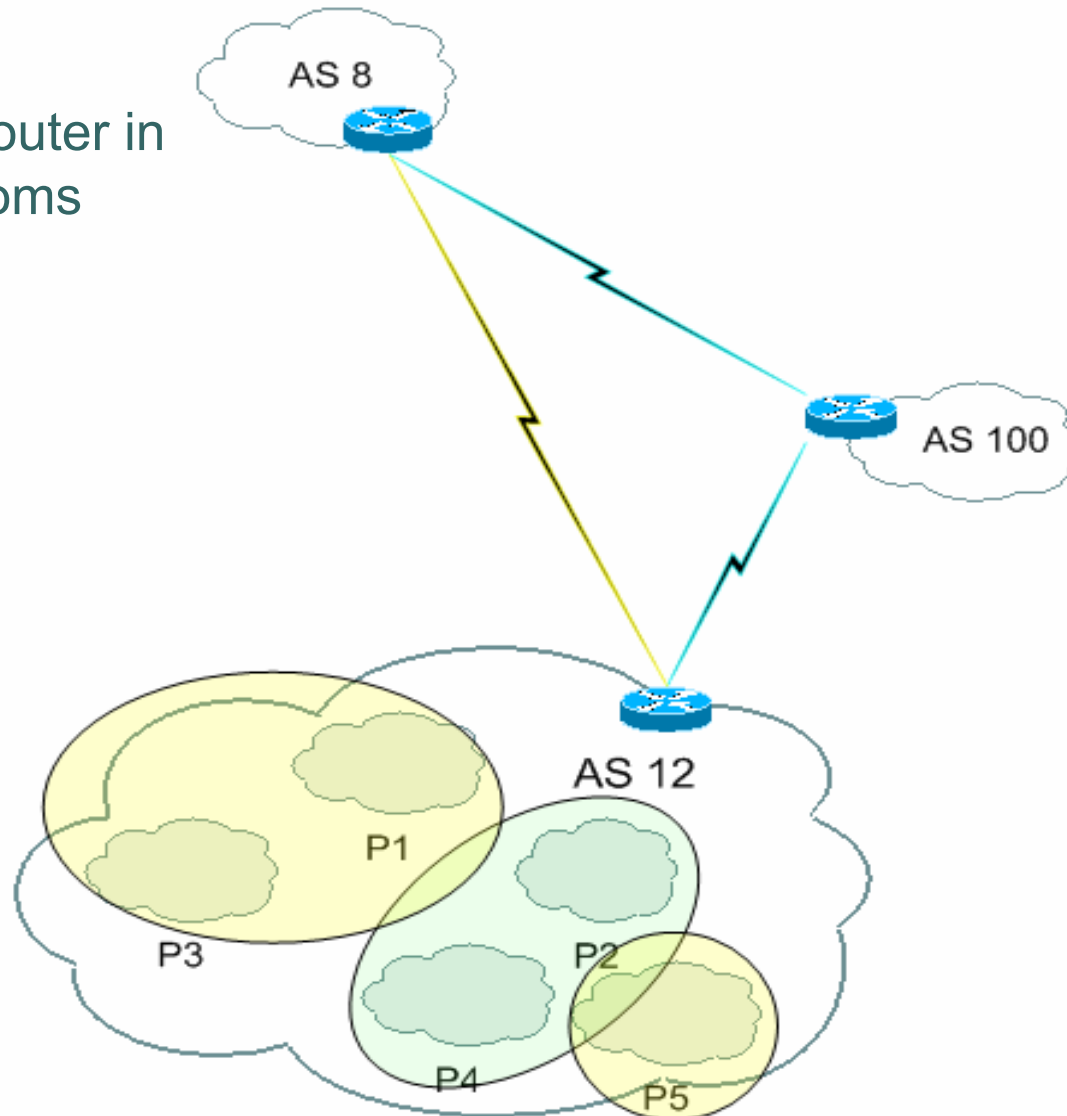
AS 9 BGP table

P1	100 12
P2	1 12
P3	100 12
P4	1 12
P5	1 12

AS 8 BGP table

P1	12
P2	100 12
P3	12
P4	100 12
P5	12

View of a single router in
AS 8 – two atoms



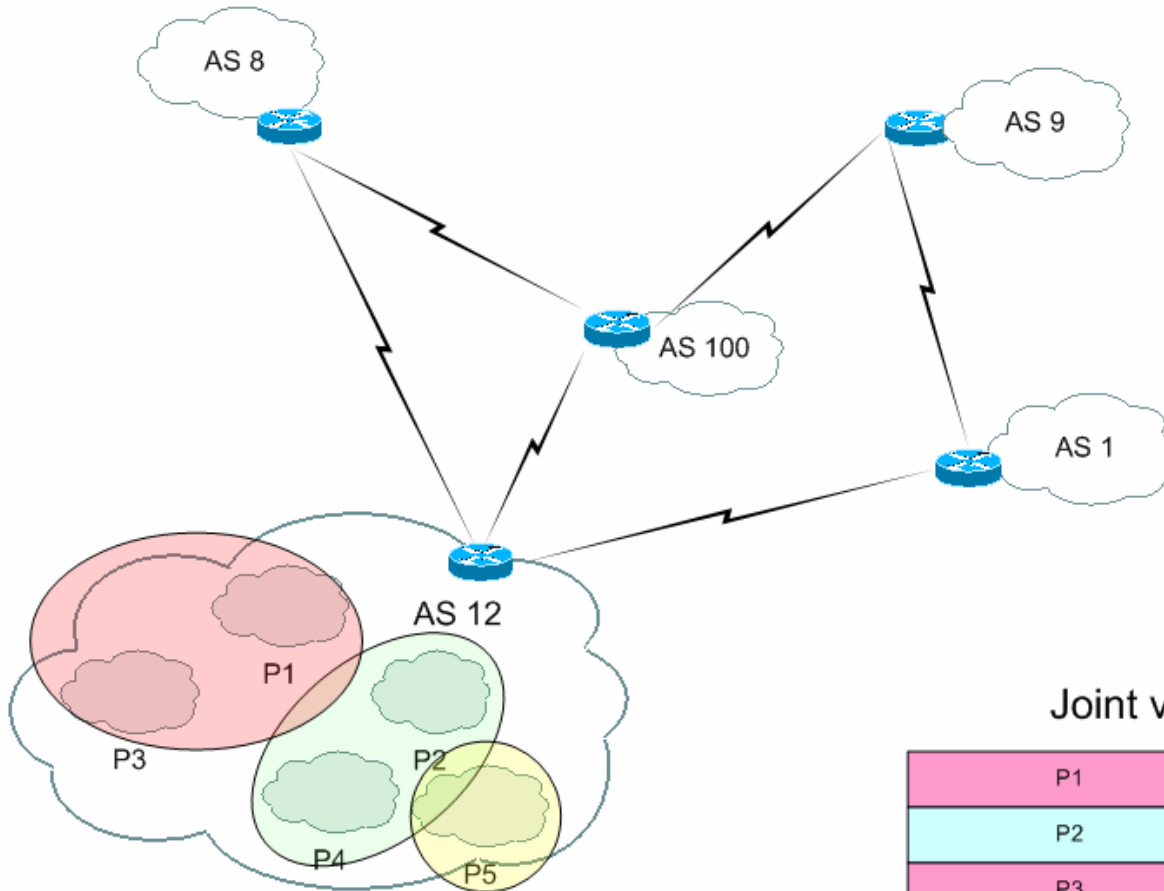
AS 8 BGP table

P1	12
P2	100 12
P3	12
P4	100 12
P5	12

Joint view of AS 8 + 9 Three atoms

AS 9 BGP table

P1	100 12
P2	1 12
P3	100 12
P4	1 12
P5	1 12



Joint view

P1	AS8: 12 AS9: 100 12
P2	AS8: 100 12 AS9: 1 12
P3	AS8: 12 AS9: 100 12
P4	AS8: 100 12 AS9: 1 12
P5	AS8: 12 AS9: 1 12



Atom definition

- Atom definition
 - Prefix group which appear in the same local group on any BGP router / shows no contradiction
- An Atom is assumed the result of policy
- fault should affect full atoms



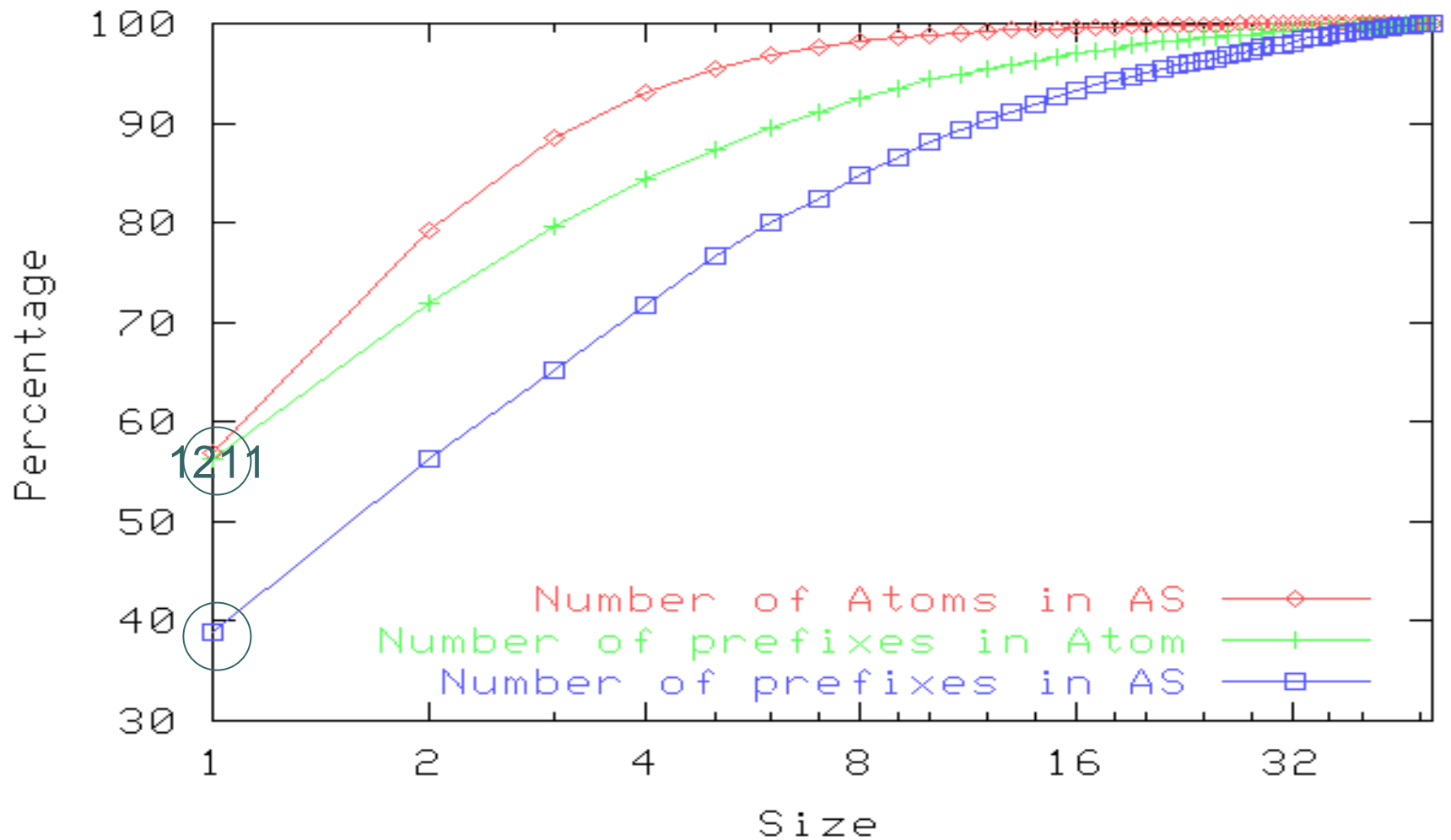
Scale comparison

Entity	Count
Announced Prefixes	> 110K (now ~115K)
Atoms	> 20K
AS	> 12K (now ~14K)

Atoms are much closer to AS in scope !!!

AS, Atom and Prefix size

Cummulative size distribution





Practical Atom definition

- Defined in theory by all BGP tables
- Can be calculated well by sample (8) of BGP views taken in a 'snapshot'
- Analog to viewing a very complex '3D' graph from 8 spatial locations
 - Beware of 'same angle'

Practical Atom calculation

- Get 'snapshot' of many BGP tables

p1	1 2 3 4
p2	1 9 5 4
p3	1 2 3 4
p4	1 9 5 4
p5	1 9 5 4

p1	8 5 6 4
p2	5 3 1 4
p3	8 5 6 4
p4	5 3 1 4
p5	8 5 6 4

- Create an AS_path set for each prefix

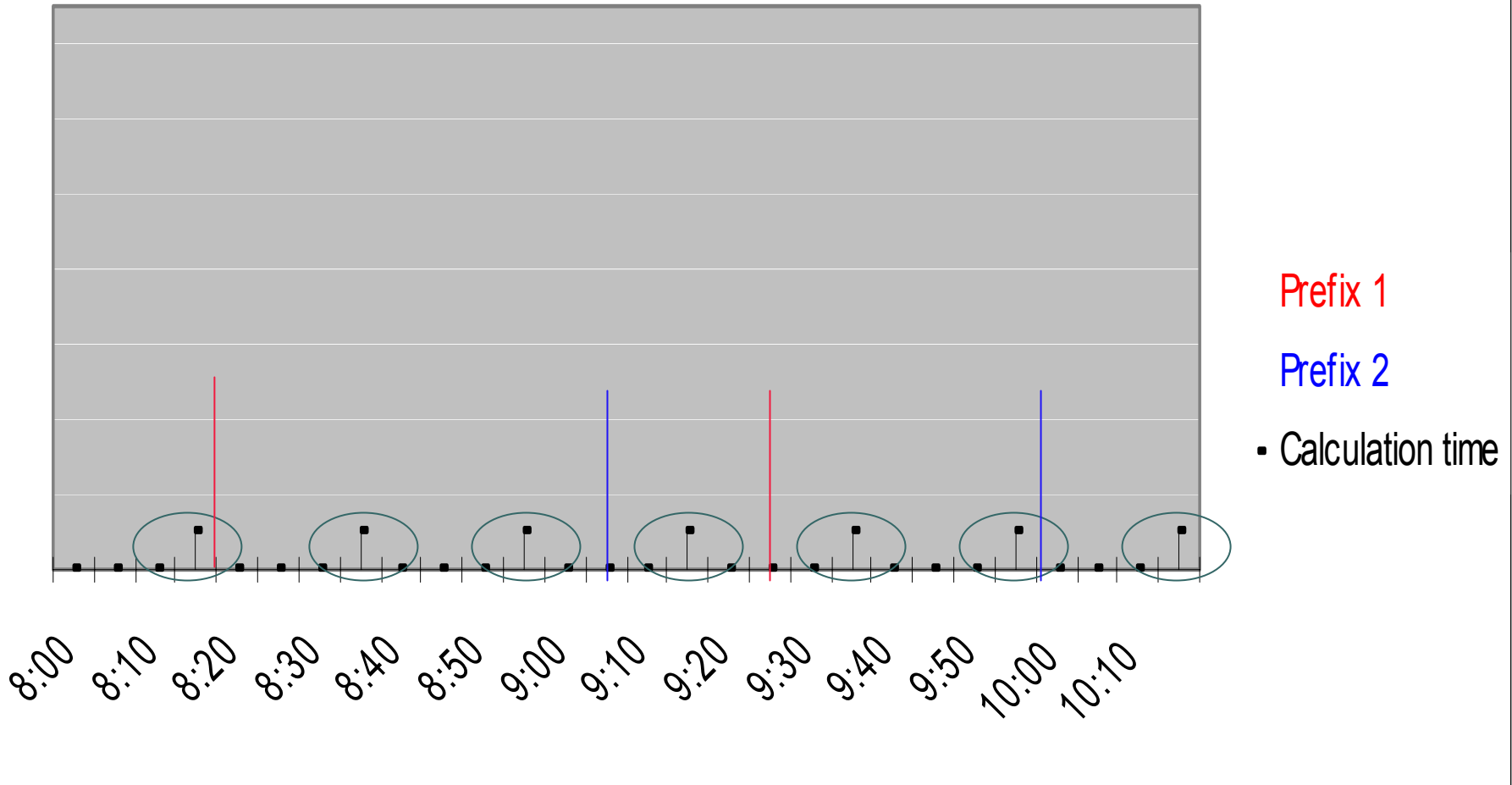
p1	1 2 3 4 8 5 6 4
p2	1 9 5 4 5 3 1 4
p3	1 2 3 4 8 5 6 4
p4	1 9 5 4 5 3 1 4
p5	1 9 5 4 8 5 6 4

- Each AS_path set is an atom



Alternate Calculation

- Is the snapshot method valid ?
 - Uses distributed snapshot
 - No guarantee of synchronization
- Filter out prefixes not assured converged
- Repeat a number of times and group prefixes with no contradiction to clique



Prefix 1 cannot be included at the 08:15 calculation

Prefix 2 cannot be included at the 10:00 calculation



Alternate calculation

- Results close to 'snapshot' method (2-3%)
 - Snapshot is valid ?
 - Different noise ?



Atom Stability

- Atoms calculated stable to 3% of prefixes in 8 hour period
- Much more stable then the prefix AS path attribute.
- Still Less stable then expected
 - Policy set Manually or by script but by design



BGP updates

- BGP groups prefixes in same update if they share same attributes

TIME: 09/01/01 12:23:27

TYPE: BGP4MP/MESSAGE/Update

FROM: 64.211.147.146 AS3549

TO: 193.0.0.1 AS12654

ORIGIN: IGP

ASPATH: 3549 6453 12956 4926

NEXT HOP: 64.211.147.146

COMMUNITY : 3549:2246 3549:9840

ANNOUNCE

200.16.216.0/24

192.67.345.0/24



Correlating Atoms to updates

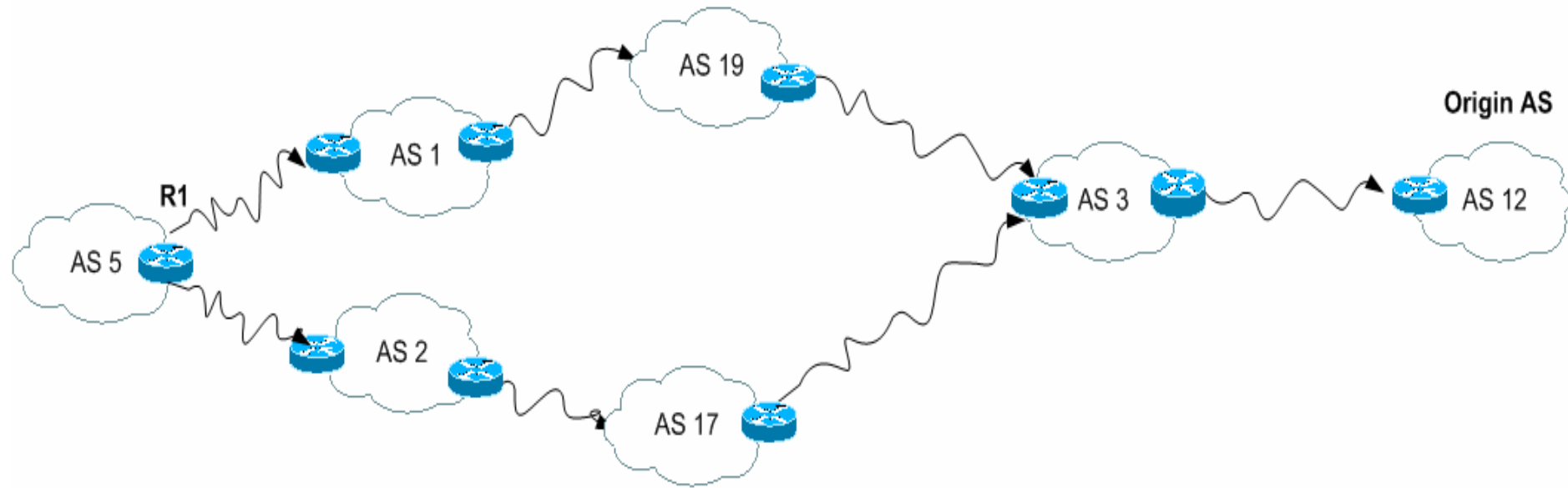
- Fault should affect full atoms and be in one update
 - All attributes are the same
- **75% of updates include single, full atom**
- 86% contain prefixes from one atom only
- **Full AS prefix set appears in just 20% of updates**



Atom split locations

- Where do atoms get created ?
- Policy normally enforced in source AS or neighbor
- How can we calculate the ‘split’ location ?

Atom split example (single router view)



R1 partial BGP table

Prefix group	AS path
11.0.1.0/24	1 19 3 12
11.0.2.0/24	1 19 3 12
11.0.3.0/24	2 17 3 12
11.0.4.0/24	2 17 3 12
11.0.5.0/24	2 17 3 12



Atom split example

Atom	Prefixes	AS path set
#1	11.0.1.0/24 11.0.2.0/24 11.0.3.0/24 11.0.4.0/24 11.0.5.0/24	R1: 12 R2: 12 R3: 12

All prefixes belong to AS 12 (len 1)

Atom split example

Atom	Prefixes	AS path set
#1	11.0.1.0/24 11.0.2.0/24	R1: 3 12 R2: 4 12 R3: 5 12
#2	11.0.3.0/24 11.0.4.0/24 11.0.5.0/24	R1: 3 12 R2: 4 12 R3: 6 12

New atom due to different path at R3 (len 2)

Atom split example

Atom	Prefixes	AS path set
#1	11.0.1.0/24 11.0.2.0/24	R1: 17 3 12 R2: 54 4 12 R3 25 5 12
#2	11.0.3.0/24	R1: 17 3 12 R2: 54 4 12 R3 22 6 12
#3	11.0.4.0/24 11.0.5.0/24	R1: 19 3 12 R2: 54 4 12 R3 22 6 12

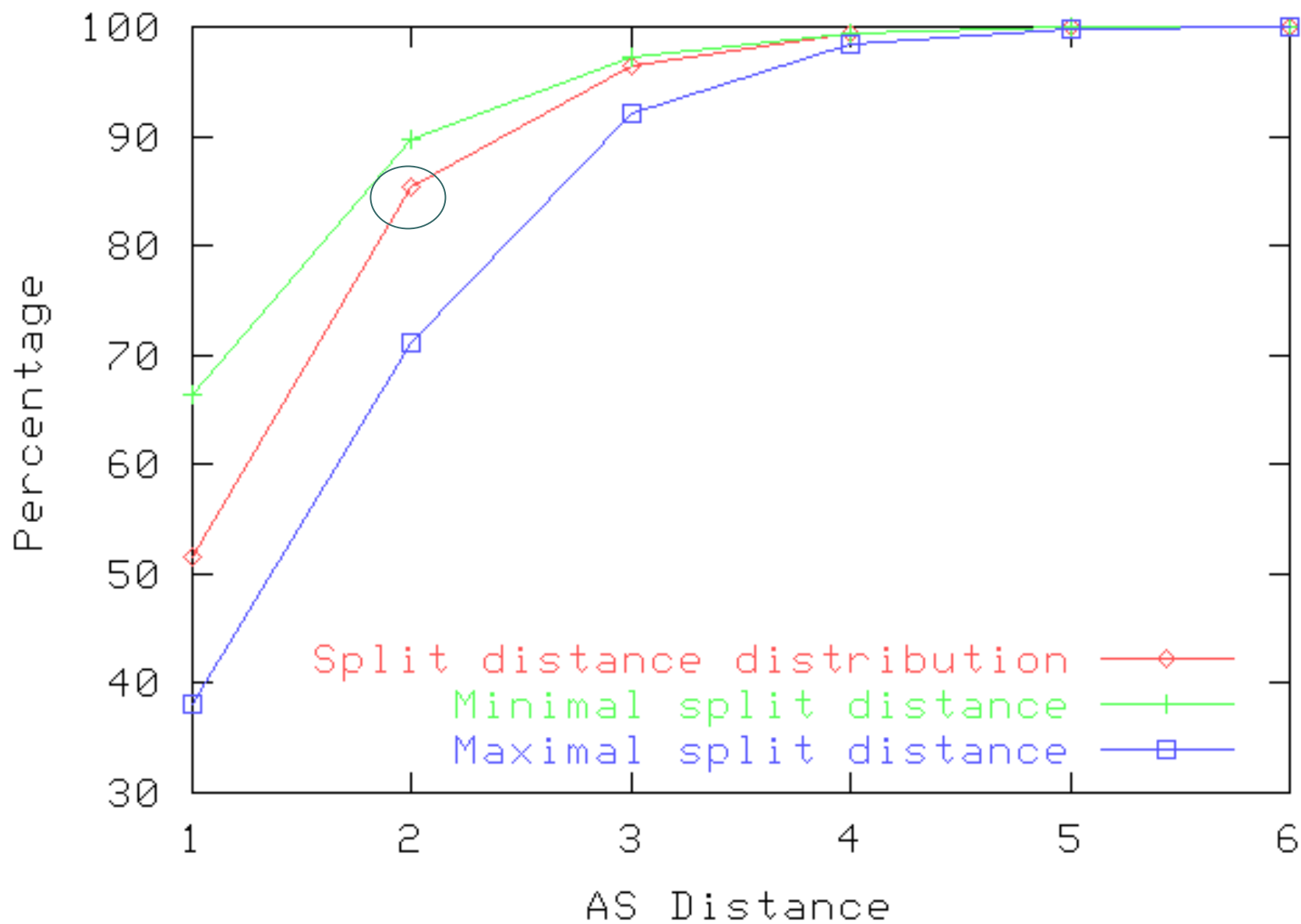
New atom by different path at R1 (len 3)

Atom split example

Atom	Prefixes	AS path set
#1	11.0.1.0/24 11.0.2.0/24	R1: 2 17 3 12 R2: 100 34 54 4 12 R3: 22 11 25 5 12
#2	11.0.3.0/24	R1: 2 17 3 12 R2: 100 34 54 4 12 R3: 99 13 22 6 12
#3	11.0.4.0/24 11.0.5.0/24	R2: 1 19 3 12 R1: 100 34 54 4 12 R3: 99 13 22 6 12

Longer paths will not provide extra resolution

Distribution of atom creation distance





Application of Atoms

1. Research

- Differentiate Fault from Policy
- Track policy

2. Tweak update hold timers:

- Panelize partial Atom updates
- Reward full updates

3. Reduce scope of tags used in MPLS

- When/If MPLS is global in the Internet
- Set on a per prefix basis
- Could be reduced by as much as $1:5^{31}$



Application of Atoms

4. Reduce load of active measurements

- Use Atoms sampling and not prefix sampling
- Results are not very favorable

Type	Count	Mean SD
AS	66	55.7
Atom	226	37.9
Prefix	9146	20.1



Application of Atoms

5. Compress BGP updates
 - Gain limited to 66% of prefix section

6. Shows limit on route table reduction while keeping information
 - Renumbering to CIDR by atoms gain 50% only



Administration and distribution of Atom structure

1. Central body (CAIDA ?)
 - 'Perfect' results
 - How to distribute ?
2. Origin AS tag by extended community
 - Knowledge propagation within BGP
 - How to automate and not introduce more administration overhead
3. Use 'local' versions of Atoms
 - No distribution
 - Reduces potential benefits



Thank you for your time

Questions ?

BGP 'reminder'

- BGP routing between AS analogue to RIP with in AS (AS == Router)

