BGP Route Stability

Alexander Asimov
<aa@highloadlab.com>
Highload Lab

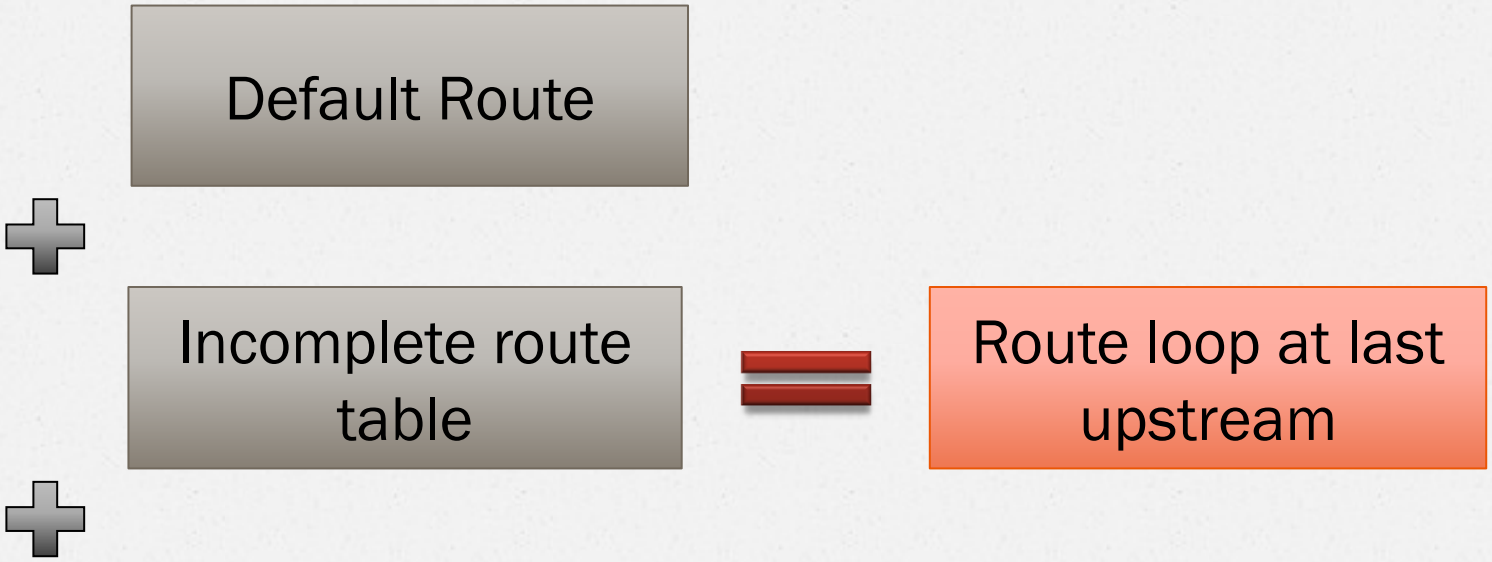
Research history

- 2009-2010 – study of BGP convergence process:
  - Imitation testbed (BGPC + PRIME);
  - BGP convergence time equations;
  - Experiments with BGP timers;
- Since 2010 – BGP policy recovery model;
- Since 2011 – BGP route monitor.

Route instability

1. Router misconfiguration;
2. BGP route loops;
3. Link failure.

Default route



Default route : Stat

- Prefixes affected 12514
- Vulnerability: could be used for DDoS amplification on AS (~ TTL)
  - Link exhaustion
  - Billable bandwidth

DDoS amplifier

- Prefixes affected 688
- Vulnerability: exploit makes AS work like DDoS amplifier.
  - DDoS at multiple ASes at the same time with N-time increase of attack magnitude

BGP Route Loops

Built-in defense for static loops

Dynamic loops!

BGP dynamic loops

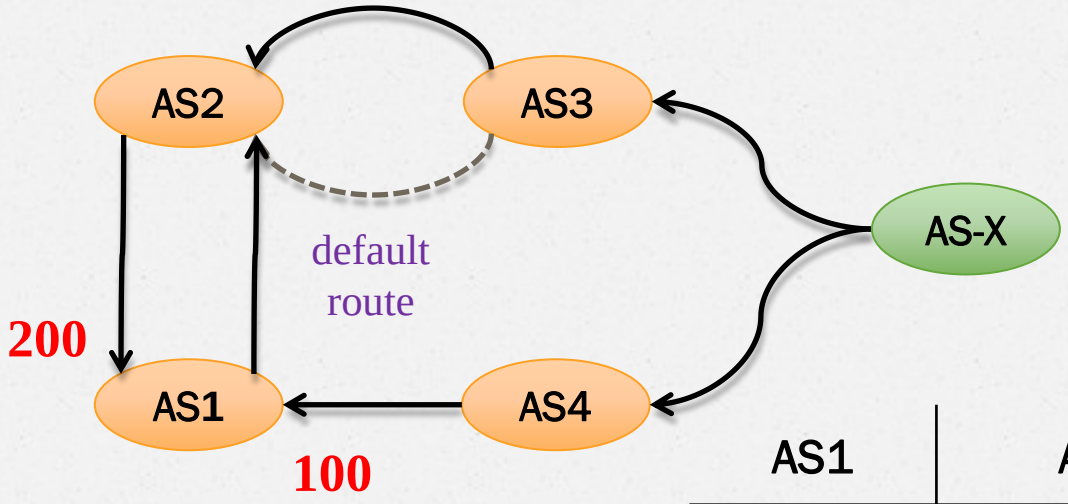
- Route withdraw;
- Changing prepend policy;
- Changing LOCAL_PREF.

BGP convergence process

- $ET_{up} = \theta(d \times Et_{wait})$, where d – diameter of AS graph;
- $ET_{down} = \theta(D \times Et_{wait})$, where D – Hamiltonian path of AS graph;

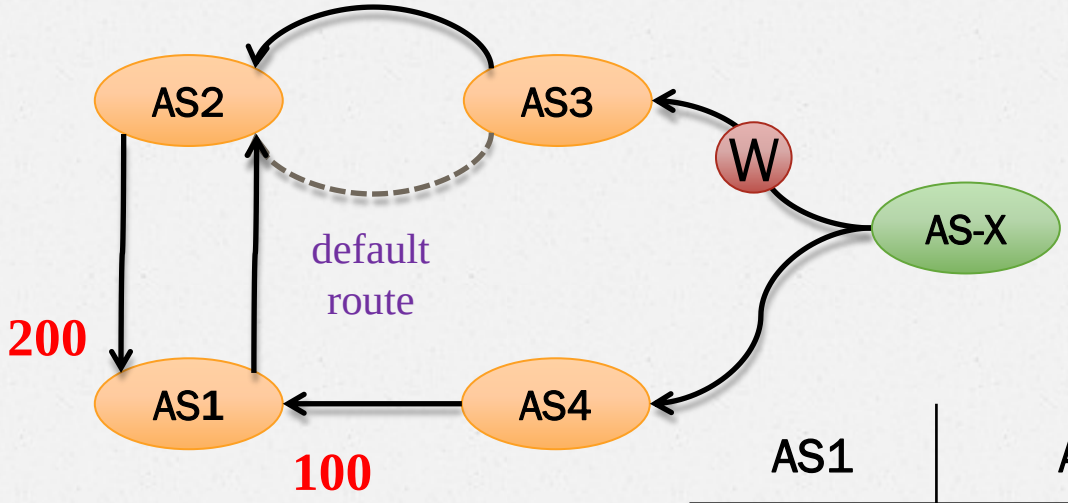
Alternatives paths!

Route Withdraw



AS1	AS2 AS3 AS-X
AS2	AS3 AS-X
AS3	AS-X
AS4	AS-X

Route Withdraw

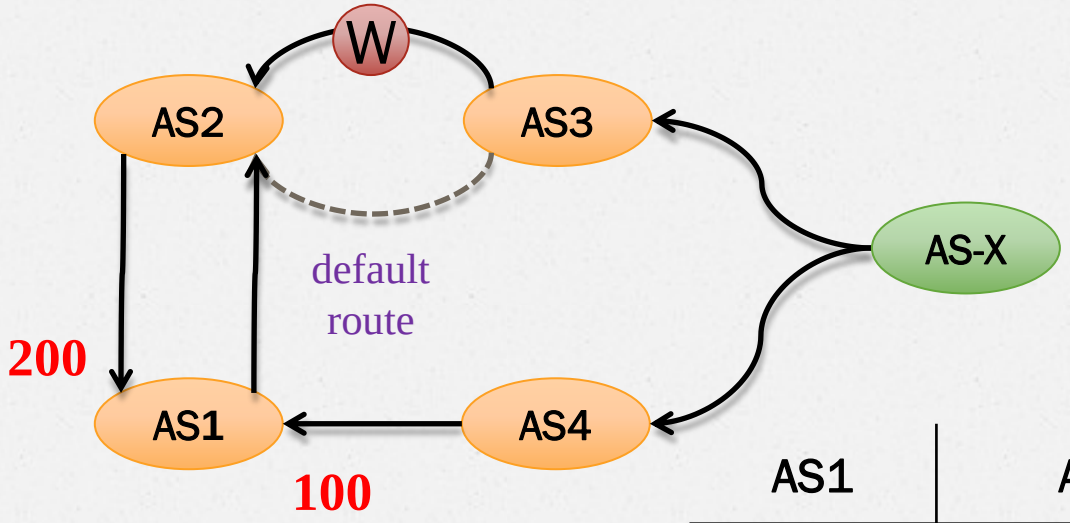


Loop AS2 – AS3

Packet loss AS1, AS2,AS3

AS1	AS2 AS3 AS-X
AS2	AS3 AS-X
AS3	--
AS4	AS-X

Route Withdraw

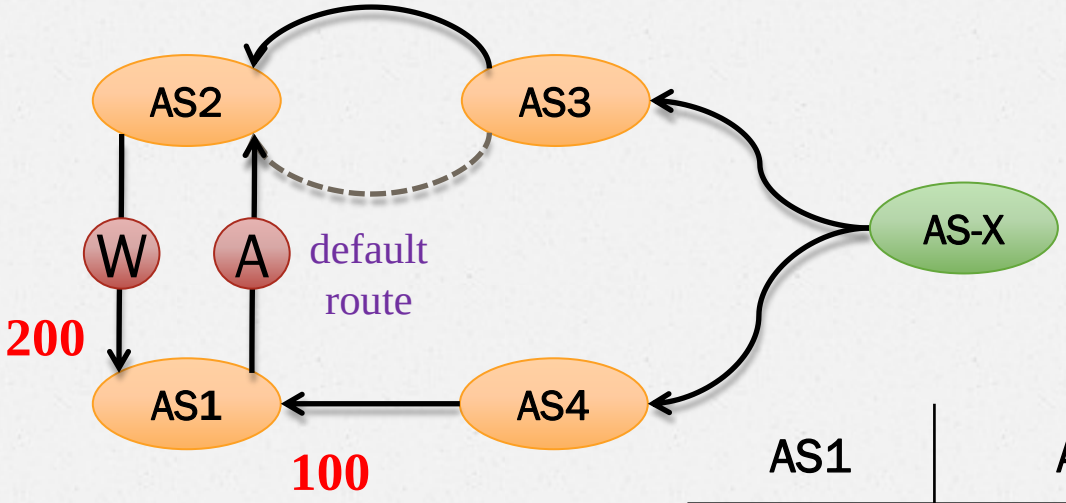


Loop AS2 – AS3

Packet loss AS1, AS2,AS3

AS1	AS2 AS3 AS-X
AS2	--
AS3	--
AS4	AS-X

Route Withdraw



1 minute problem!

$\sim\rho(c2p) * 30s$

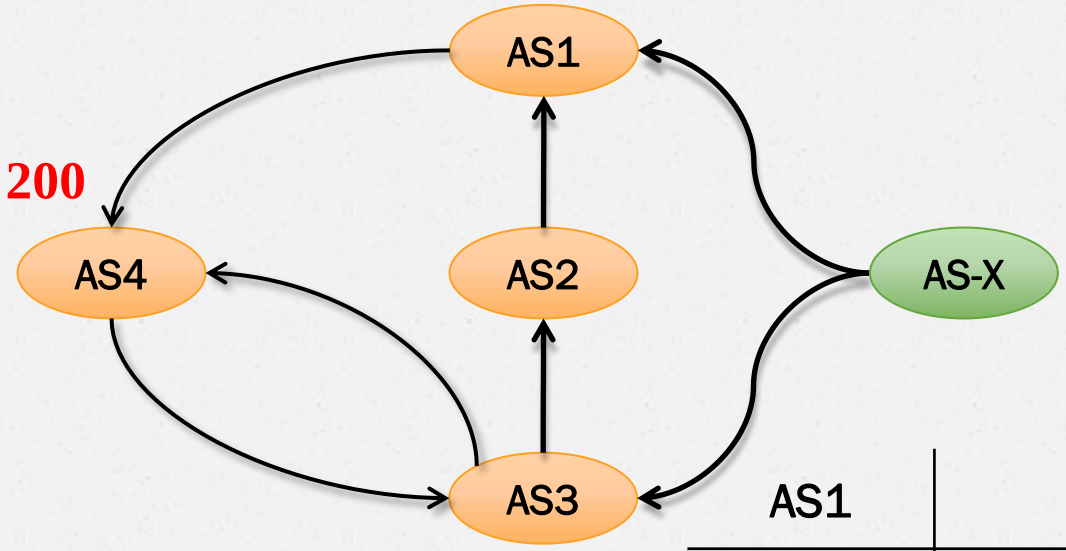
AS1	AS2 AS3 AS-X
AS2	AS1 AS4 AS-X
AS3	--
AS4	AS-X

Route Flaps

One route withdraw could cause unstable dynamic loop.

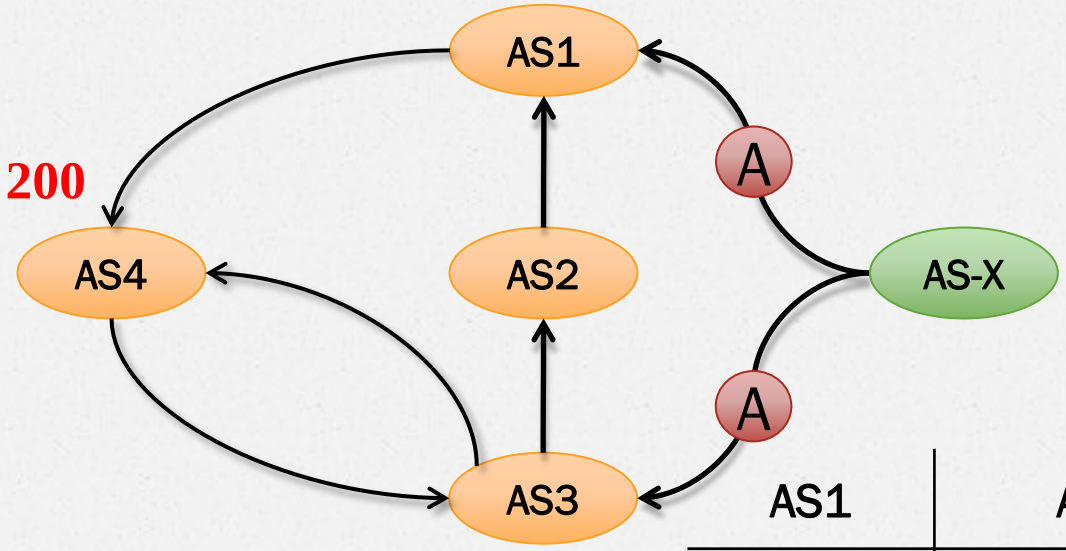
Multiple flaps could make your net unreachable.

Prepend Policy



AS1	AS-X
AS2	AS3 AS-X
AS3	AS-X
AS4	AS1 AS-X

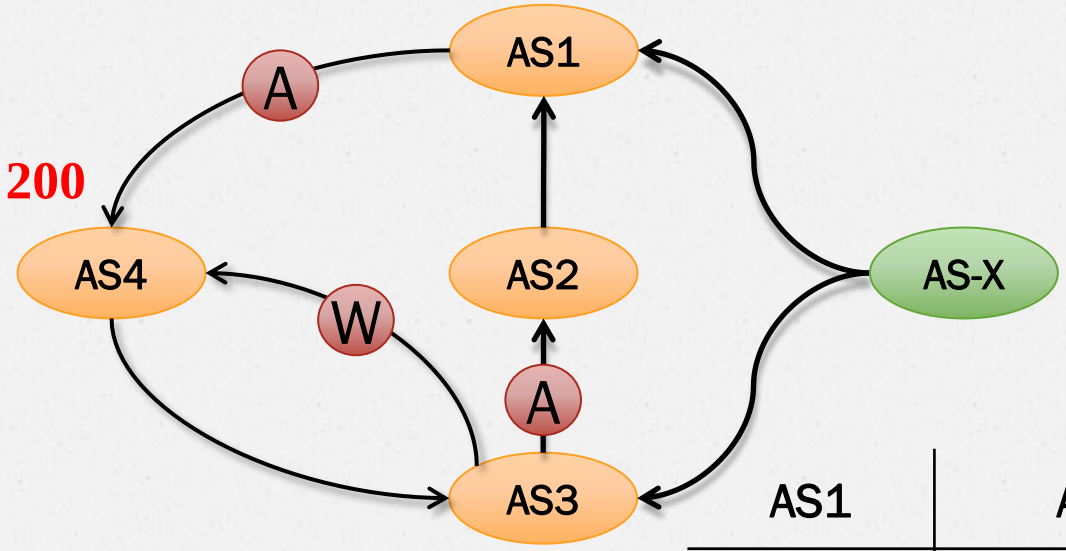
Prepend Policy



AS1	AS2 AS3 AS-X
AS2	AS3 AS-X
AS3	AS4 AS1 AS-X
AS4	AS1 AS-X

as-path prepend 5
Loop AS1-AS2-AS3-AS4

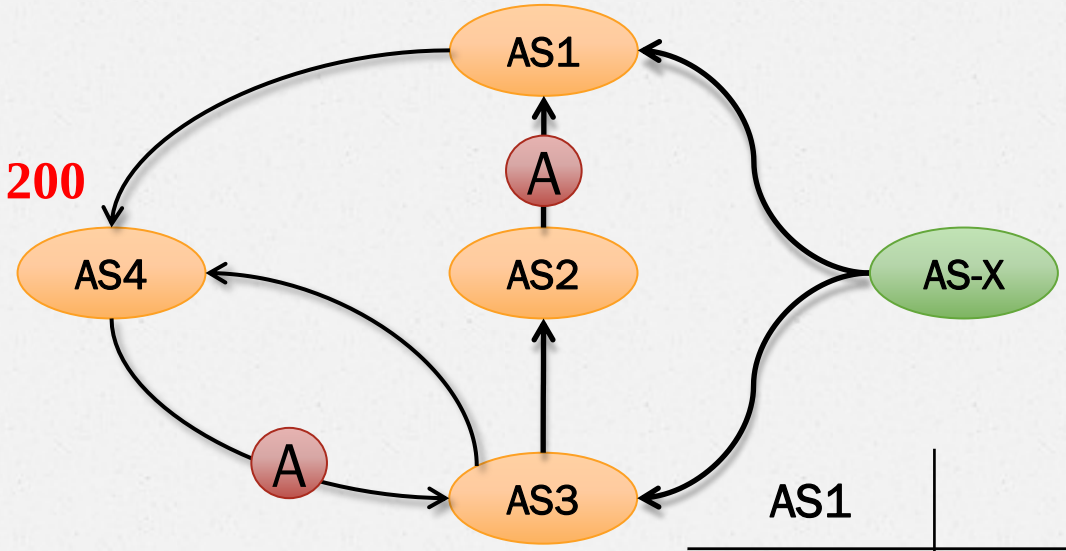
Prepend Policy



Loop AS1-AS2-AS3-AS4

AS1	AS2 AS3 AS-X
AS2	AS3 AS4 AS1 AS-X
AS3	AS4 AS1 AS-X
AS4	AS1 AS2 AS3 AS-X

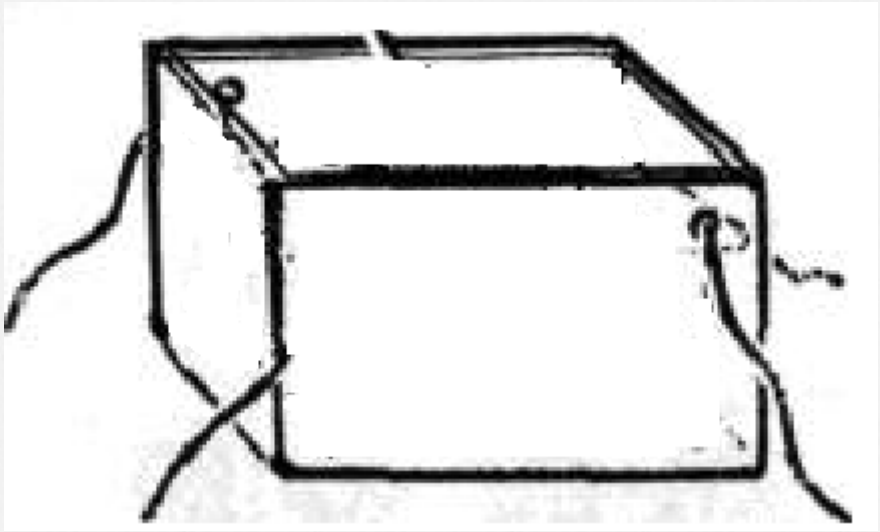
Prepend Policy



AS1	AS-X(5)
AS2	AS3 AS4 AS1 AS-X
AS3	AS-X(5)
AS4	AS1 AS2 AS3 AS-X

More then 1 minute problem

Core of the problem



Traffic flow engineering looks like…

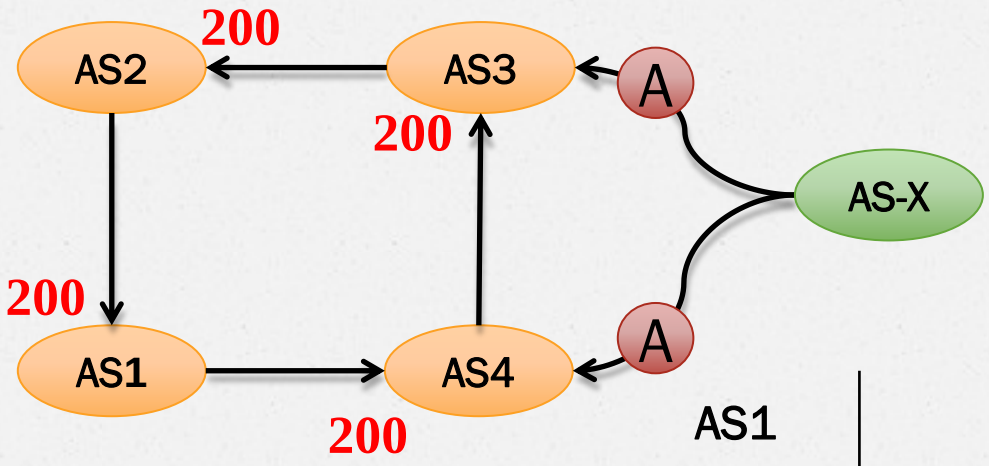
Automated traffic flow engineering

- No opportunity for traffic flow prediction;
- No use of BGP convergence process;
- Could make networks partially unavailable.

Route flaps : Stat

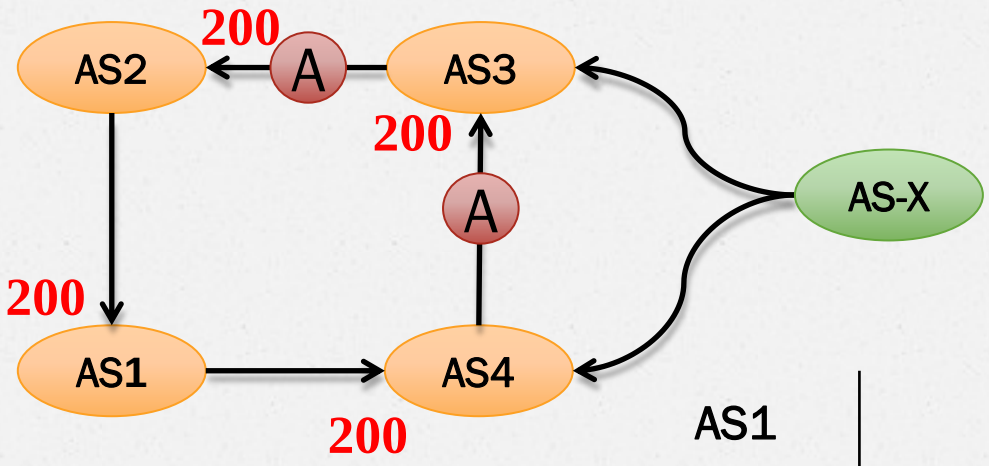
- Prefixes affected 1720
- Vulnerability:
  - Packet loss;
  - BGP message noise.

BGP Route Policy Loops



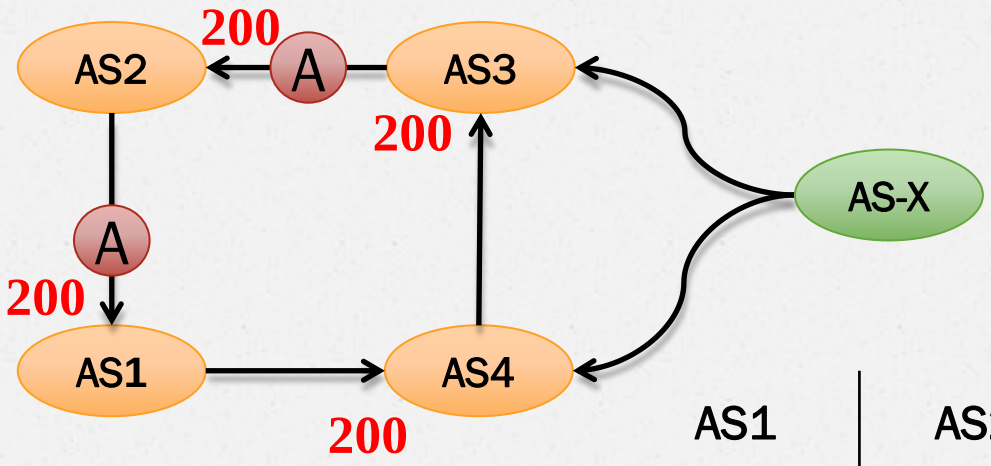
AS1	--
AS2	--
AS3	AS-X
AS4	AS-X

BGP Route Policy Loops



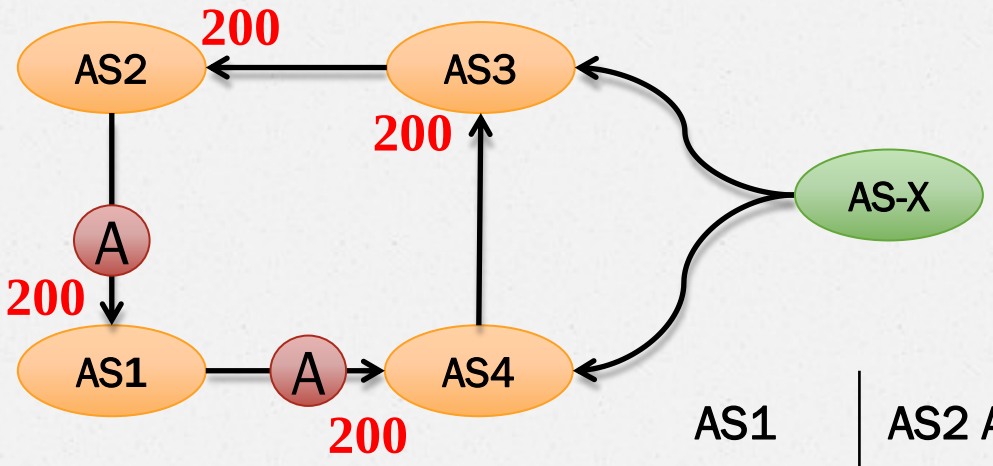
AS1	--
AS2	AS3 AS-X
AS3	AS4 AS-X
AS4	AS-X

BGP Route Policy Loops



AS1	AS2 AS3 AS5 AS-X
AS2	AS3 AS4 AS5 AS-X
AS3	AS4 AS5 AS-X
AS4	AS5 AS-X

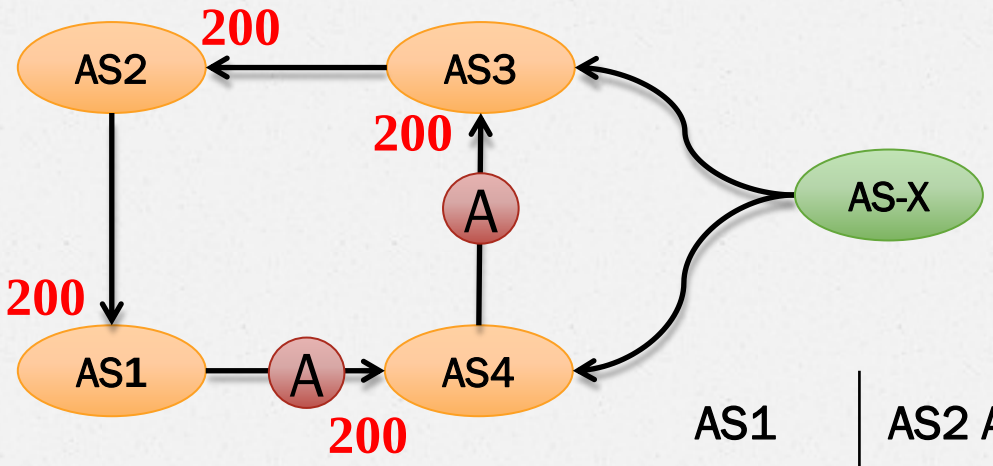
BGP Route Policy Loops



Loop AS1-AS2-AS3-AS4

AS1	AS2 AS3 AS4 AS5 AS-X
AS2	AS3 AS4 AS5 AS-X
AS3	AS4 AS5 AS-X
AS4	AS1 AS2 AS3 AS5 AS-X

BGP Route Policy Loops



Again!

AS1	AS2 AS3 AS4 AS5 AS-X
AS2	AS3 AS4 AS5 AS-X
AS3	AS5 AS-X
AS4	AS5 AS-X

BGP Route Policy Loops : Stat

- Prefixes affected 149
- Vulnerability:
  - Packet loss
  - BGP message noise

Breaking down

1. Route flap
2. Change prepend policy

   AS-X AS-X
3. AS_PATH poisoning

   AS-X AS2 AS-X

   Makes AS2 ignore your route

Breaking down

1. Route flap
2. Change prepend policy

   AS-X AS-X
3. AS_PATH poisoning

   AS-X AS2 AS-X

   Makes AS2 ignore your route

How not to fall into another loop?

One “killer” feature

Autonomous Systems Reverse Map:

1. Traffic flow engineering;
2. AS architecture design;
3. BGP loop prediction.

AS Mapping projects

- Physical links discovery

No use of route policy;

- Macro map

Shows links that are used by somebody, not by your AS;

- Route policy data

Outdated, incomplete.

- Model abstraction too far from reality

BGP Policy recovery project: Current state

- Mathematical model for BGP route policy recovery;
- Mathematical model for AS graph;
- New active verification methods;
- Working prototype;
- Negotiations with RIPE ATLAS.

They had problems…

AS174	AS3356	AS7018	AS6939	AS701
AS3549	AS209	AS4323	AS1239	AS12389
AS2848	AS3257	AS6461	AS2914	AS8468
AS23148	AS8447	AS20485	AS6830	AS8220
AS8928	AS3303	AS4589	AS42708	AS6453
AS6730	AS31130	AS3491	AS3320	AS8218
AS286	AS702	AS3561	AS20764	AS31323
AS20632	AS4766	AS680	AS29686	AS5089
AS10026	AS12350	AS2516	AS3786	AS12741
AS7575	AS1916	AS2273	AS9498	AS1785

And more then 4k other ASes!

Examples

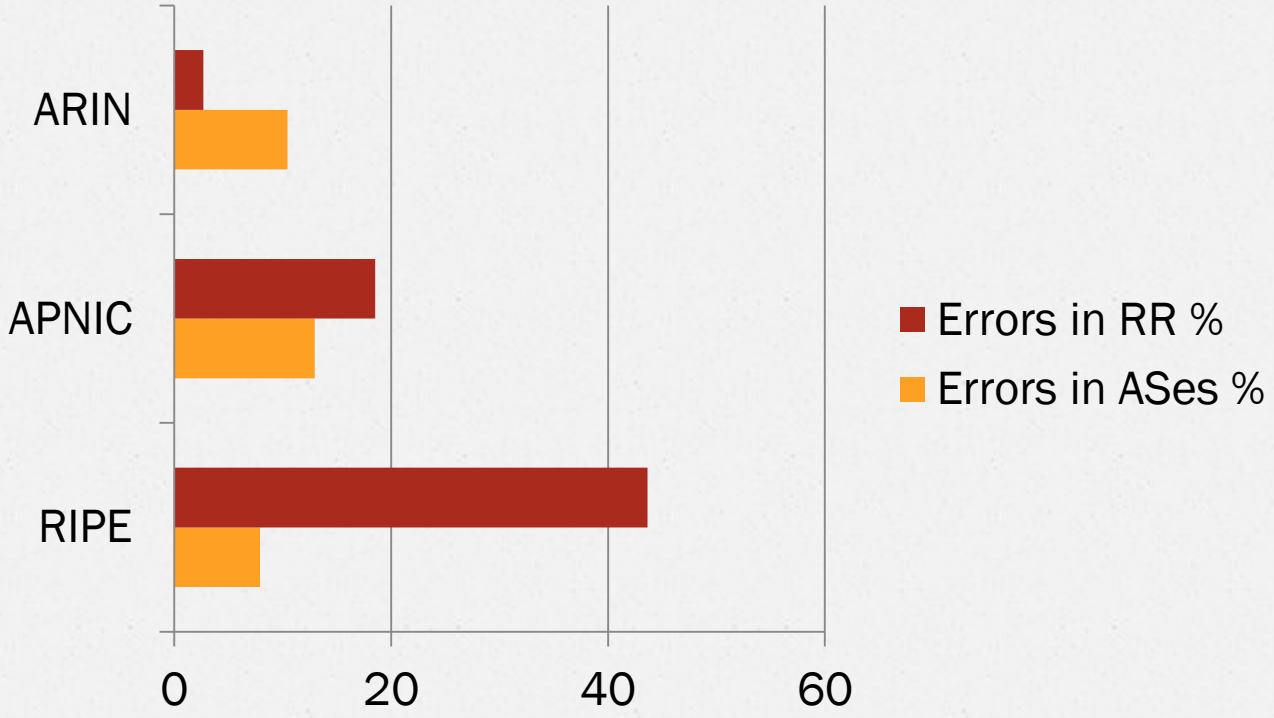
AS174

- increases DDoS 17 times
- Default route: 25 prefixes affected

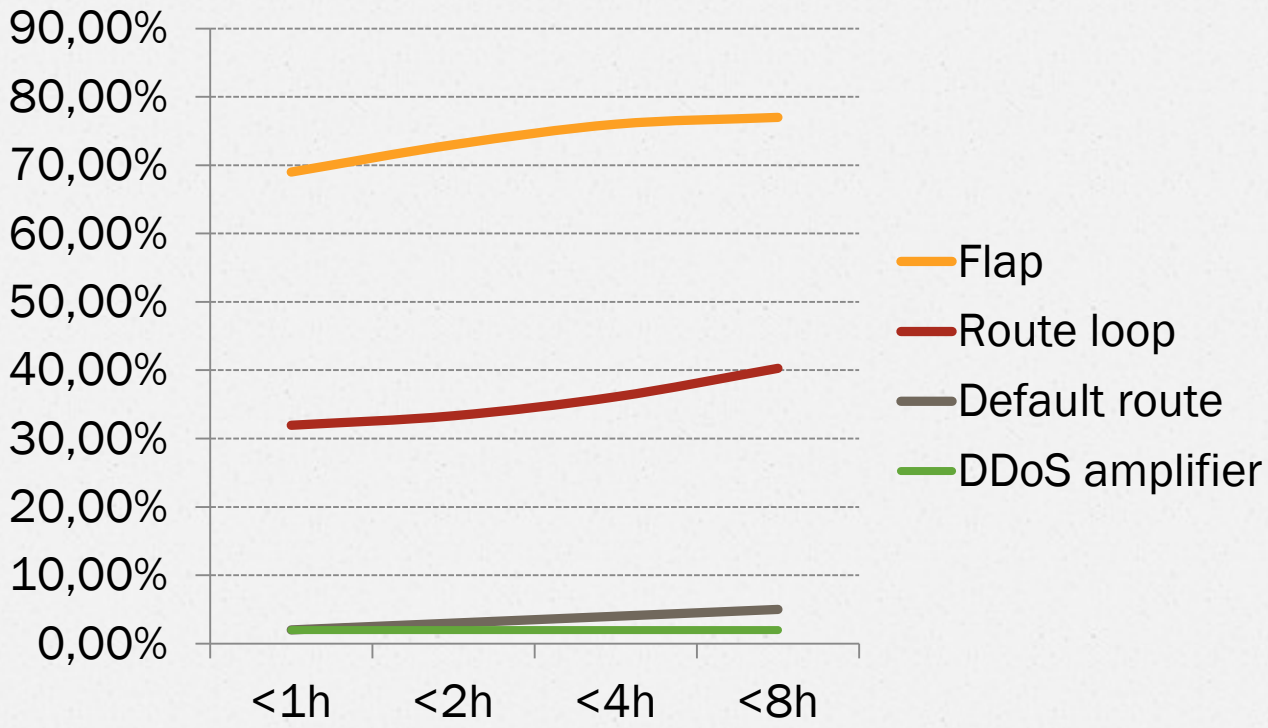
AS3356-AS3549

- increases DDoS 8 times
- Route flap: 12 prefixes affected
- Default route: 86 prefixes affected

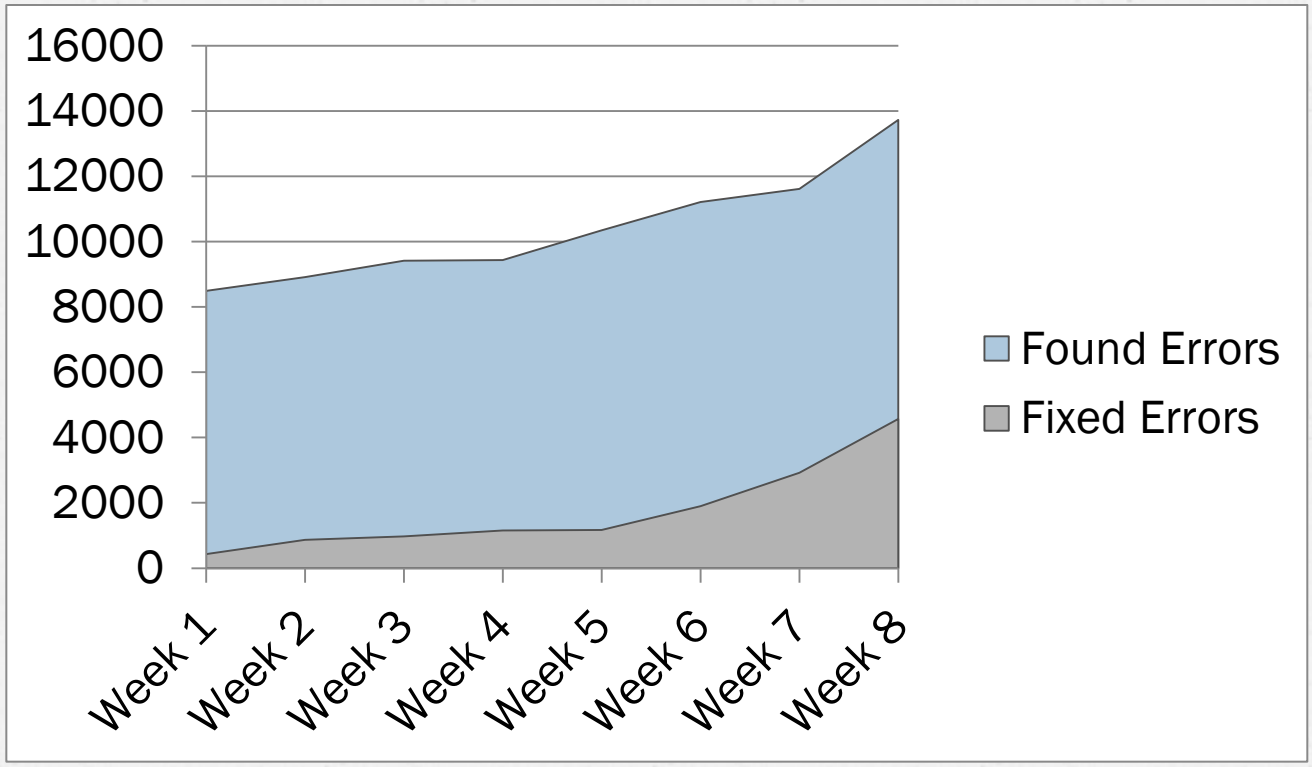
Errors grouped by RR



Duration


90,00%
80,00%
70,00%
60,00%
50,00%
40,00%
30,00%
20,00%
10,00%
0,00%
<1h
<2h
<4h
<8h
Flap
Route loop
Default route
DDoS amplifier

Trends


16000
14000
12000
10000
8000
6000
4000
2000
0
Week 1
Week 2
Week 3
Week 4
Week 5
Week 6
Week 7
Week 8
Found Errors
Fixed Errors

Robin Hood

Hello.

During our research project we have detected IP address “*.*.*.*" in AS***, which multiplies ICMP packets. 11 times is a mean value.

It is dangerous vulnerability because this IP could be used by attackers to N-times increase DoS attack.

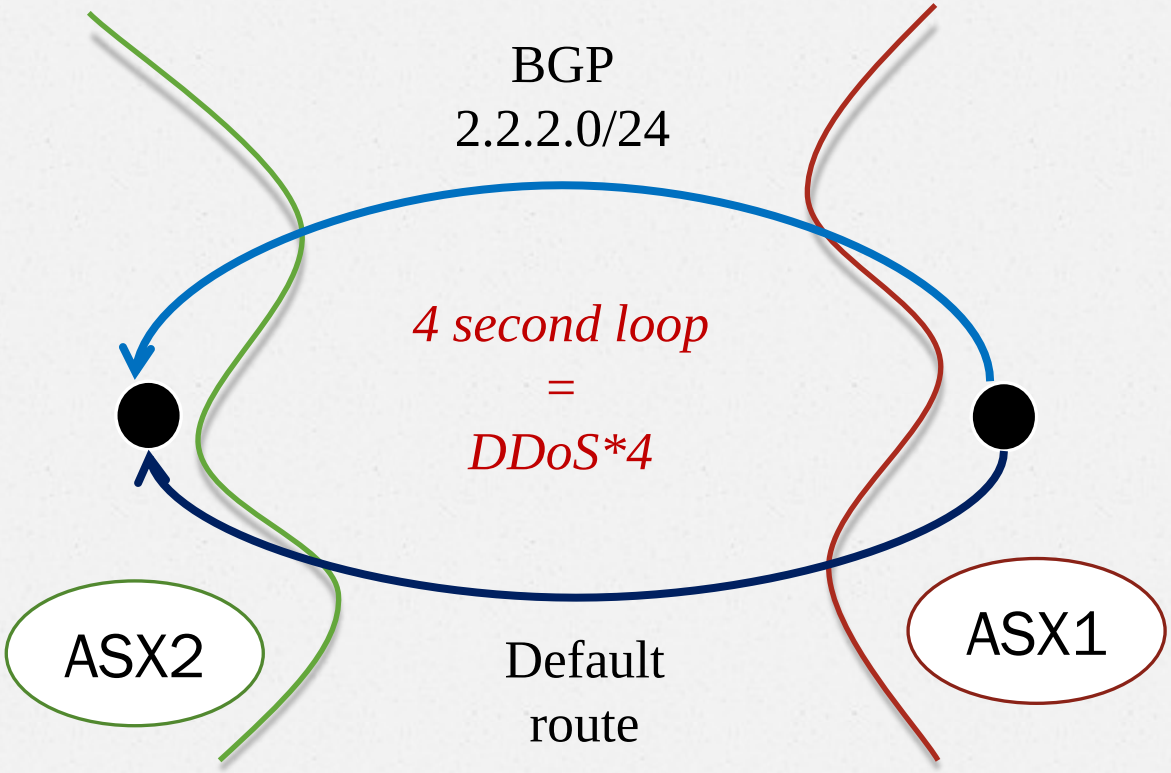
This could be harmful for your own network infrastructure and also break down attacked network.

Could you tell me for what purpose your router was configured by such way?

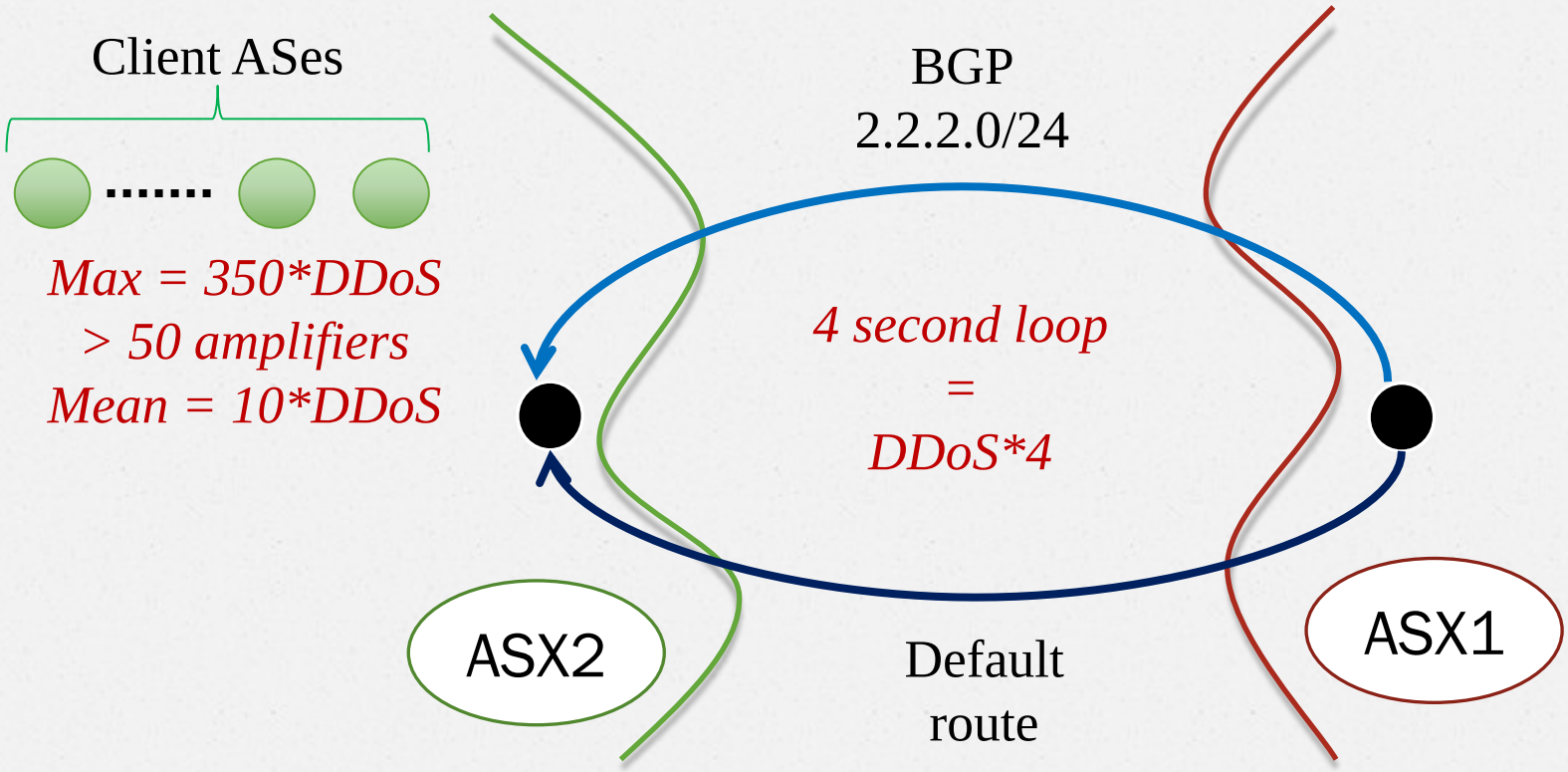
Hello

This is assigned to one of our customers. We cannot see how there equipment is configured. If you believe this is in violation of our Acceptable use policy you can email abuse@***.com to report that.

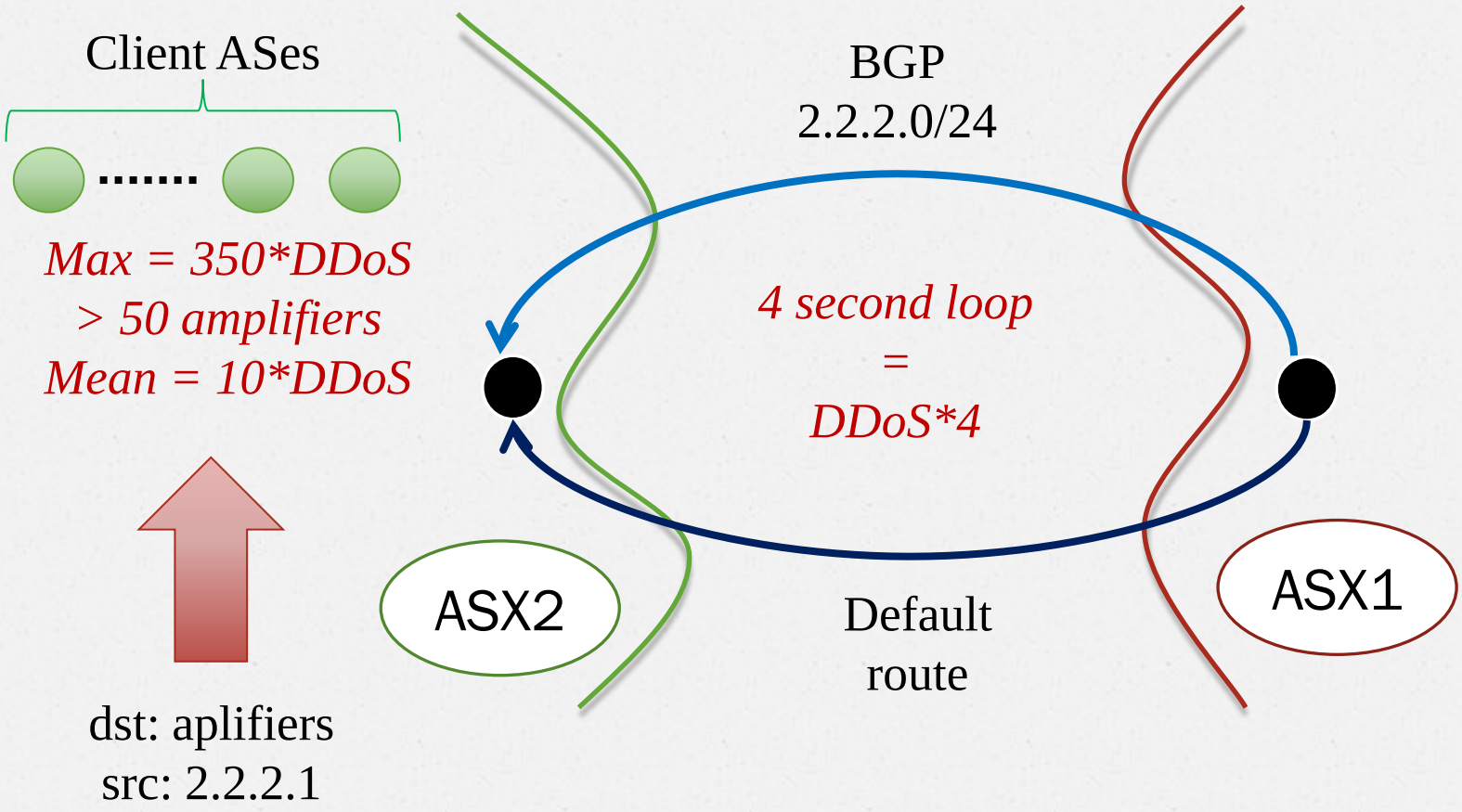
In the core of the Internet



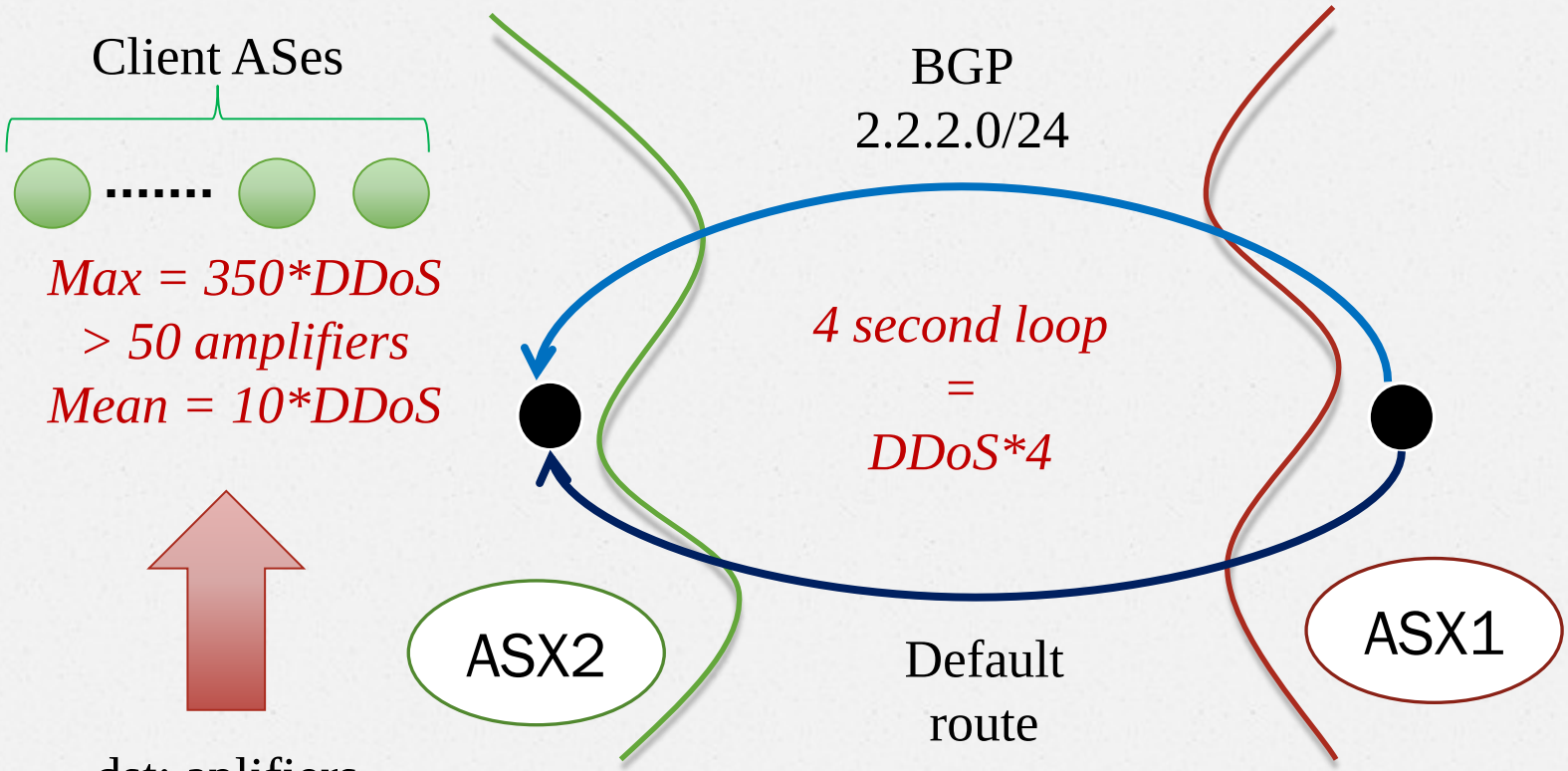
In the core of the Internet



In the core of the Internet



In the core of the Internet



Result = 40*DDoS

Results

- The problem in client network is your problem;
- The problem in your upstream network is your problem too;
- Unstable prefix is invisible from origin AS;
- BGP balancing without prediction could make network unavailable;
- We have monitor. Community needs some trusted mechanism for notification.

Thank you for listening!

Questions?