BGP Peering Engineering Automation challenges and enablers

Cloud & Virtualization Group

Camilo Cardona (camcardo@cisco.com), Paolo Lucente (plucente@cisco.com)

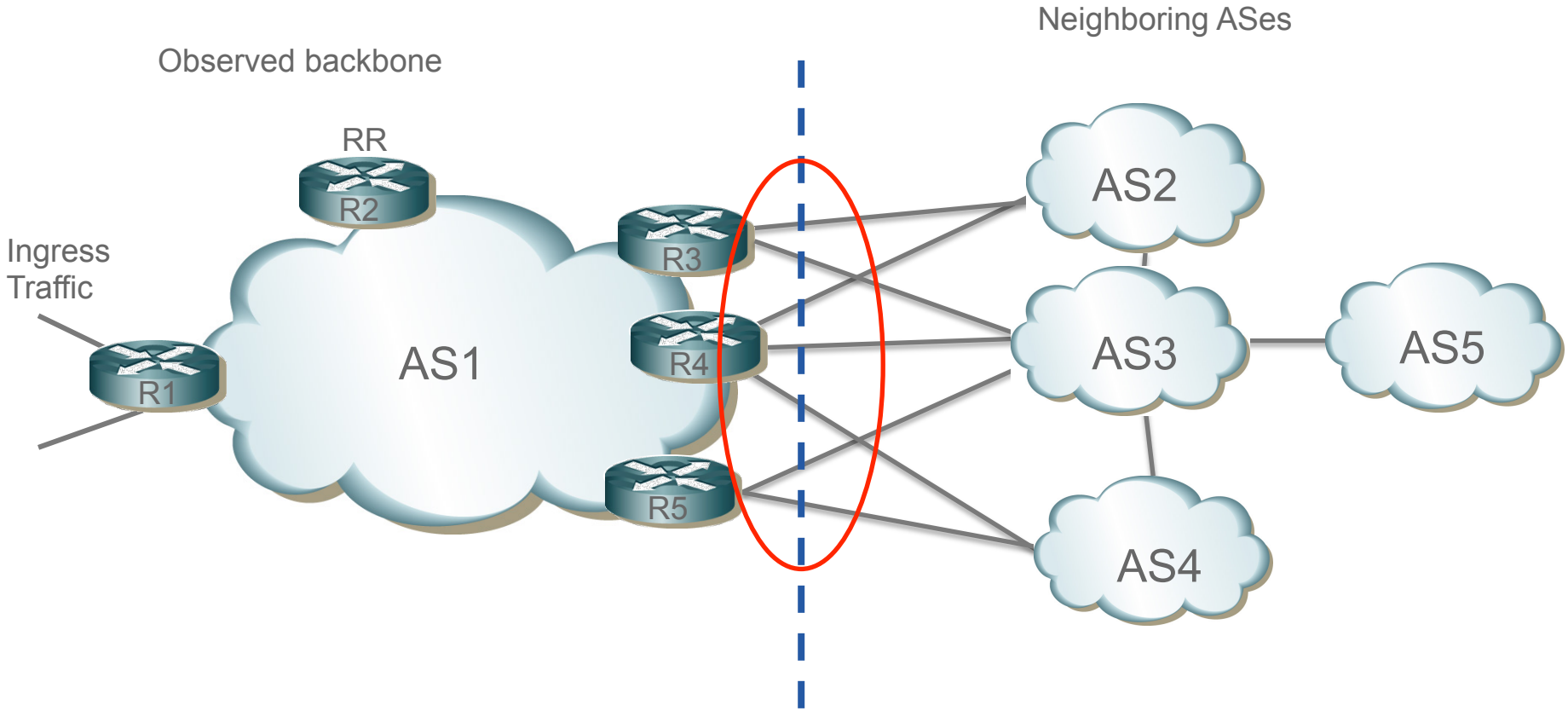
November 2015

v5.1

Introduction

- Inter-domain traffic includes all traffic crossing the boundary of networks
- BGP is used to exchange reachability information among networks
- Inter-domain traffic engineering has been historically hard:
  - Control traffic to meet objectives
  - Estimate the effects of changes
- We describe these problems and discuss potential solutions
- Main objective being (semi-)automation

Focus



Some high-level use-cases

- Design/architectural/business:
  - Simulate optimal placement of interconnects
  - Simulate impact of interconnect failures
- Operational:
  - Simulate impact of (BGP policy) changes without disrupting live traffic

Egress / Ingress differences

- Egress and ingress traffic are differently managed
- Egress traffic:
  - Operators have control on the paths they would like to use for their traffic
- Ingress traffic:
  - Depends on how the other networks decide
  - Operators can try to influence the decision of others

Egress traffic

- Routers decide the best path using the BGP decision process
- Operators change attributes of the paths to reflect their policy
- The typical egress TE process consists in tweaking attributes to steer traffic as desired
- Old problem:
  - Feamster et al. “Guidelines for interdomain traffic engineering”. CCR. 2003.
  - Nanog presentations (e.g. Wepman, 2004; Roisman, 2009; etc.)
- Proposed workflow: collection, simulation, optimization, deployment

EPE Challenges: Collection

- Requires the collection of different sources of data:
  - BGP paths, traffic, policy
- Over time mostly non-standard or mature APIs:
  - BGP paths (BGP Add-Path, BMP)
  - Traffic (Netflow/IPFIX, sFlow)
  - Policy (Tail-f, Openconfig project)

Collection | Simulation | Optimization | Deployment

EPE Challenges: Simulation

- What-if scenarios hard to simulate
- BGP decision process is complex:
  - RRs add a level of complexity
  - Not always deterministic
  - Proper network design
- Focus on the important prefixes
  - Identify importance, ie. by service or by IP prefixes/paths making more traffic
  - Prioritize or exclude

Collection | Simulation | Optimization | Deployment

EPE Challenges: Optimization

- Not easy to move traffic in a granular way:
  - Changing LP, MED changes everything
  - Need to include IGP into account
  - iBGP policies are typically not desirable
- Complex metrics:
  - For example: latency, bit-miles calculations
  - No standard way to include in the optimization process

Collection | Simulation | Optimization | Deployment

EPE Challenges: Deployment

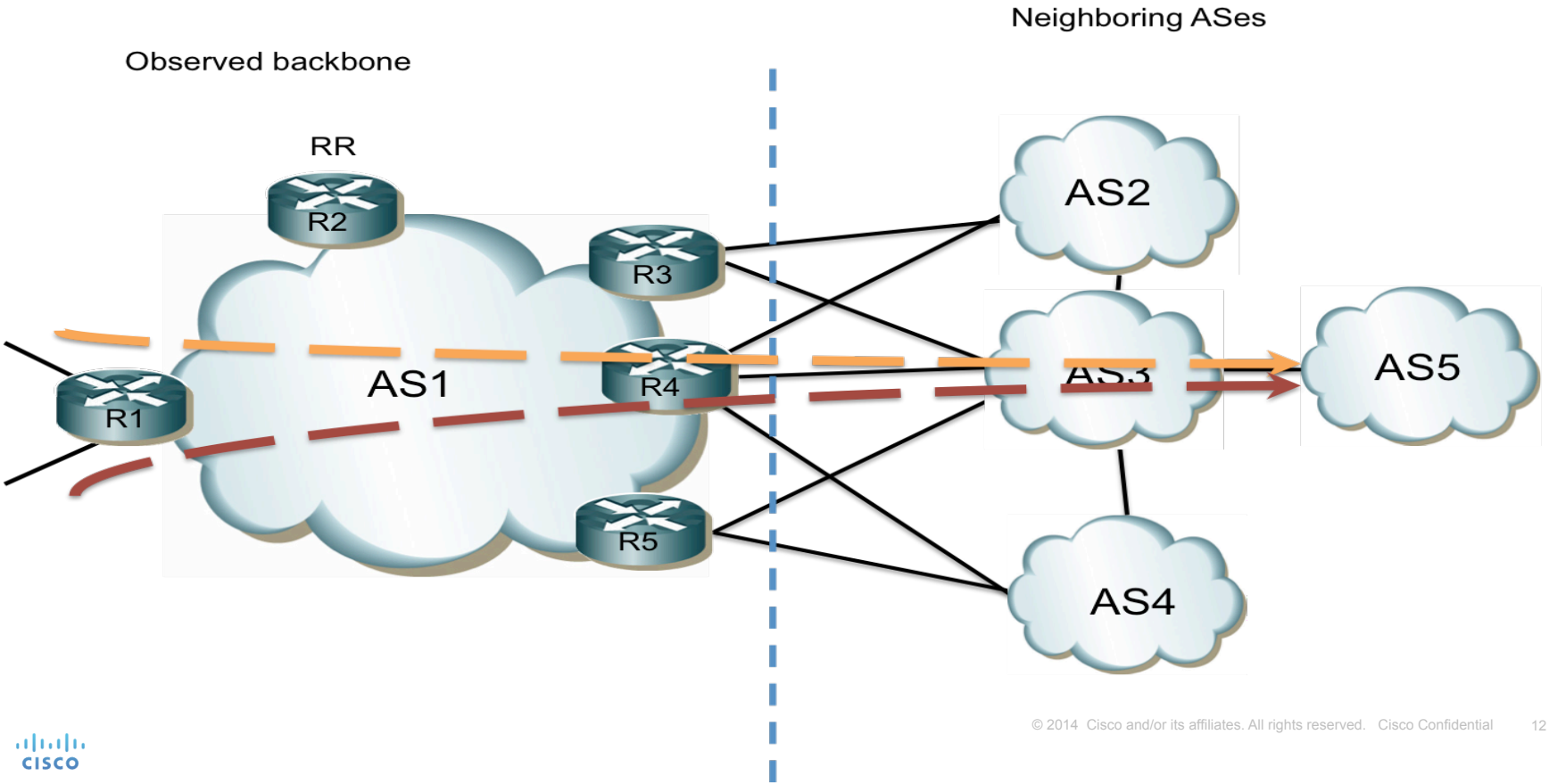
- Alternatives to deploy:
  - **Operate changes (ie. policies) at network edges**
  - Injection of best paths via a BGP controller
- Using controllers to operate changes:
  - Collection would be similar
  - Southbound interface might vary (BGP itself, Openflow, etc.)
  - Segment Routing as a solution

Collection | Simulation | Optimization | Deployment

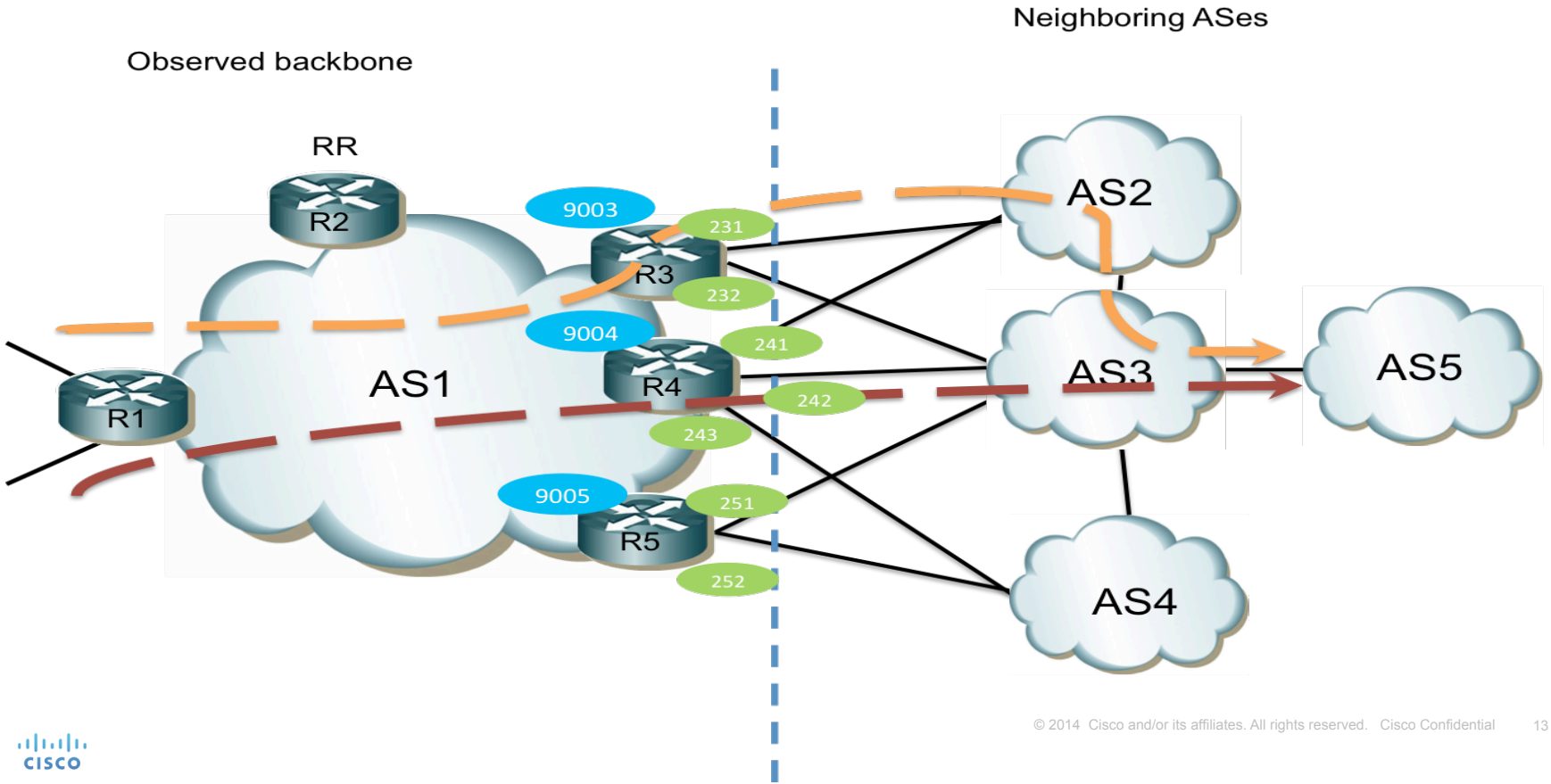
Segment Routing: very quick intro (1/3)

- Segment Routing uses Segment IDs to identify links (or services)
- MPLS or IPv6 to define the labels
- Ingress routers and can use SR to control internal path and external path
- Filsfils et al. “The Segment Routing Architecture”. Globecomm. 2015.
- For inter-domain traffic, SR allows for granularly steering traffic without being impacted by IGP or other metrics

Segment Routing: very quick intro (2/3)



Segment Routing: very quick intro (3/3)



Ingress traffic

- An operator attempts to attract traffic over specific links
- Complex by design: the Internet is formed by networks with different policies:
  - Conflicts could be unsolvable
  - Some Content providers don’t even use BGP to select a source and path
- Proposed workflow: deployment, collection, assessment, negotiation

Ingress: Deploy

- Different tools can be used to influence other ASNs:
  - AS-Path prepending
  - MEDs, Communities
  - Prefix de-aggregation or hiding
- Other routers paradigms such as LISP provide more direct ways of influencing the inbound path:
  - But policies would still prevail

Deployment | Collection | Assessment | Negotiation

Ingress: Collect, Assess, Negotiate

- Collect:
  - Similar to the egress case
  - External data might be useful
- Assess:
  - Keep track of the implemented mechanism. Ensure that it works
- Negotiate:
  - Talk to your neighbors. They might be willing to cooperate

Deployment | Collection | Assessment | Negotiation

Conclusions and closing words

- Take control of your data
- Security was not discussed, but it is an important issue:
  - New data to monitor
  - more decisions to make
  - (Discard a route with problems or lower its preference?)

TOMORROW starts here.