Improving Network Agility with Seamless BGP Reconfigurations

S. Vissicchio, L. Vanbever, C. Pelsser, L. Cittadini, P. Francois, and O. Bonaventure

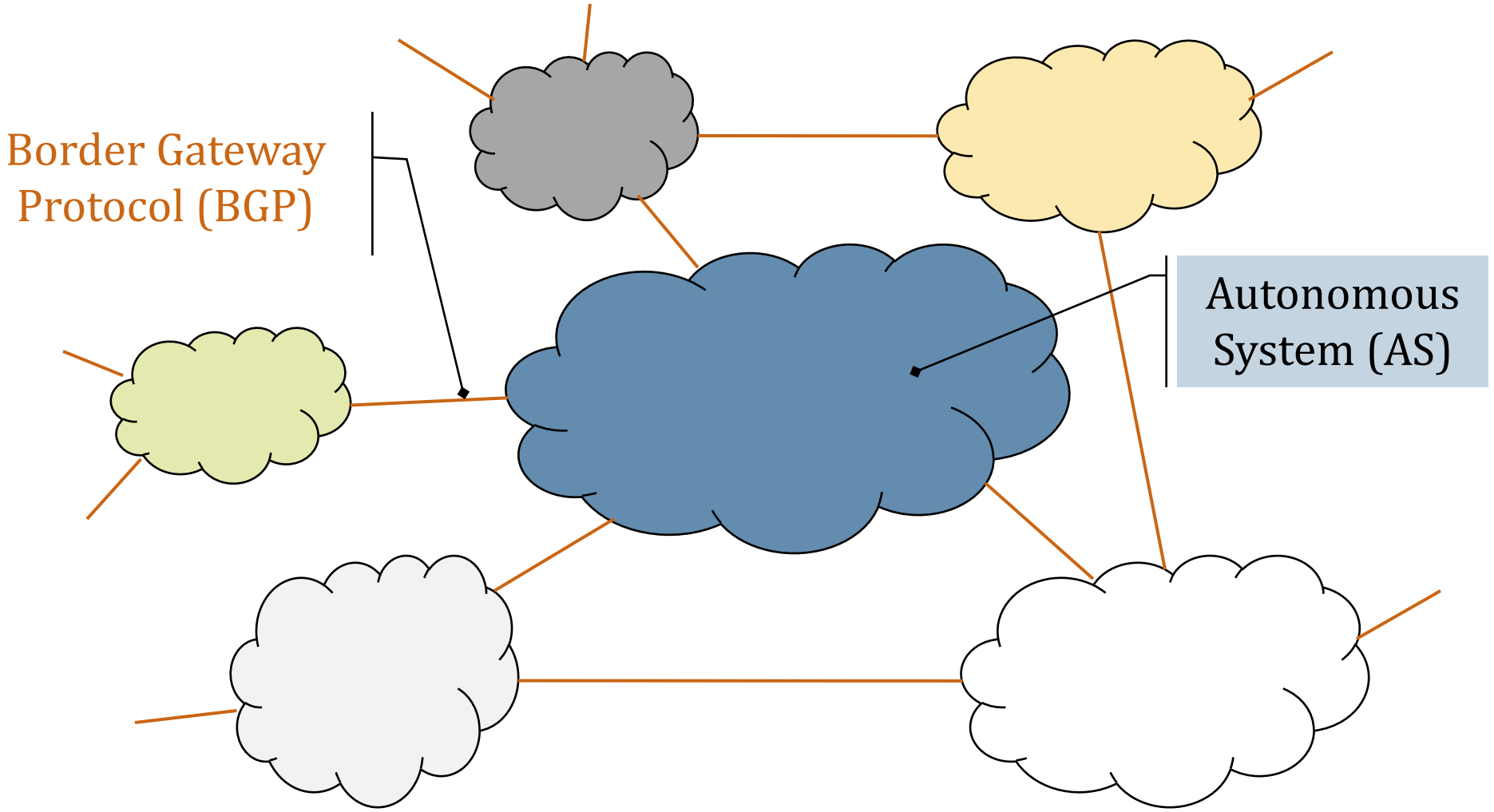
Université de Strasbourg -19 October 2012

Overview

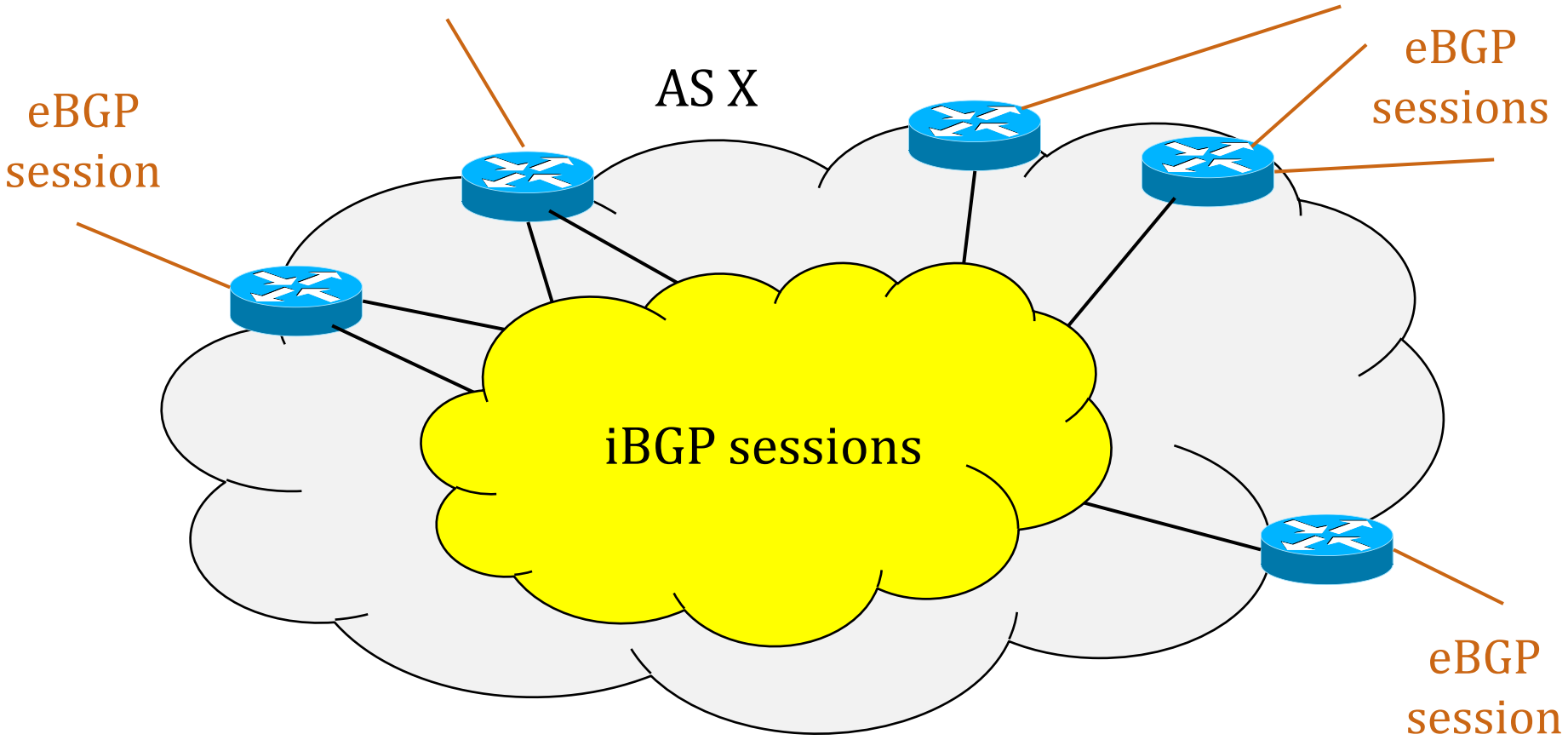
- BGP reconfigurations are needed
  - BGP is critical for ISPs
  - configuration needs to be evolved
  - running services should not be affected
- BGP reconfigurations are difficult
  - current best practices do not work
  - no algorithmic approach is viable
- We propose a reconfiguration framework
  - enabling anomaly-free reconfigurations
  - prototype implementation
  - evaluation and case study

Background

The Internet Infrastructure



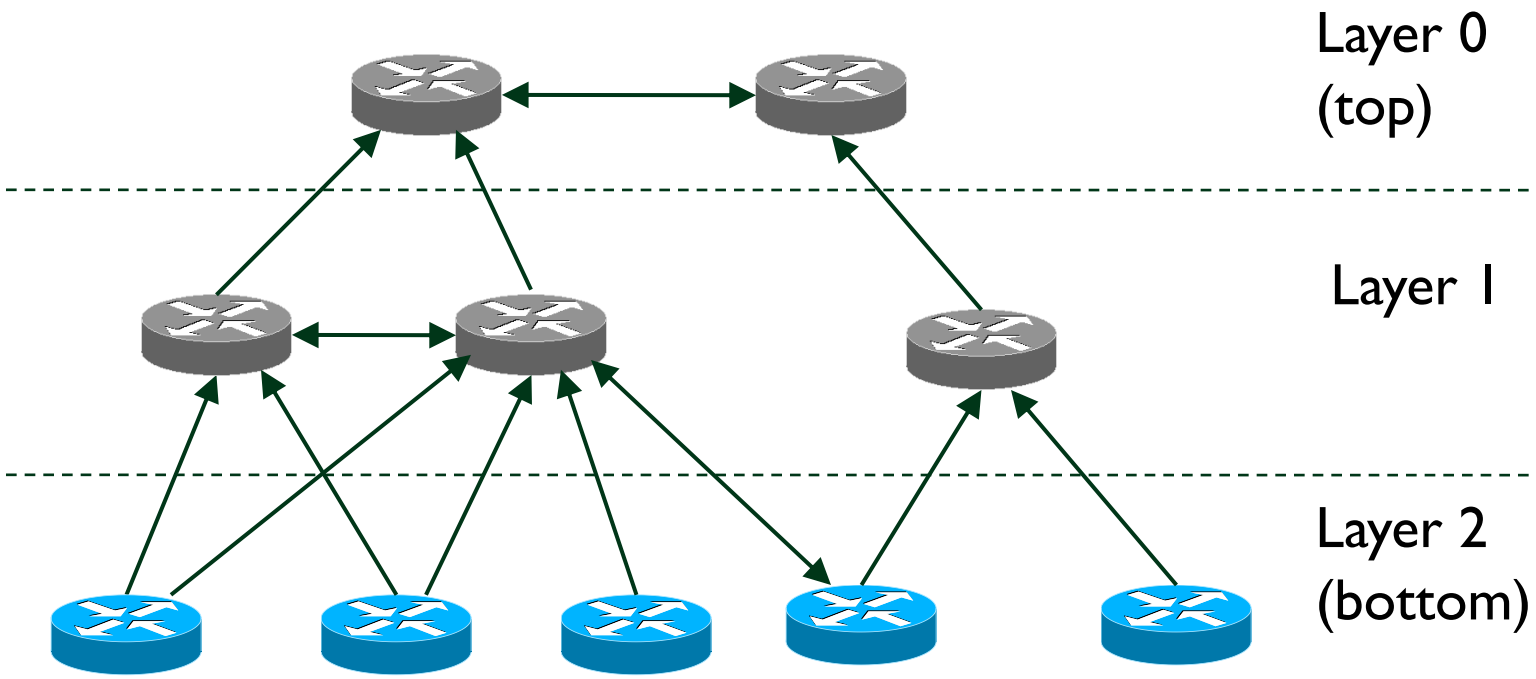
We Take the Perspective of a Single AS



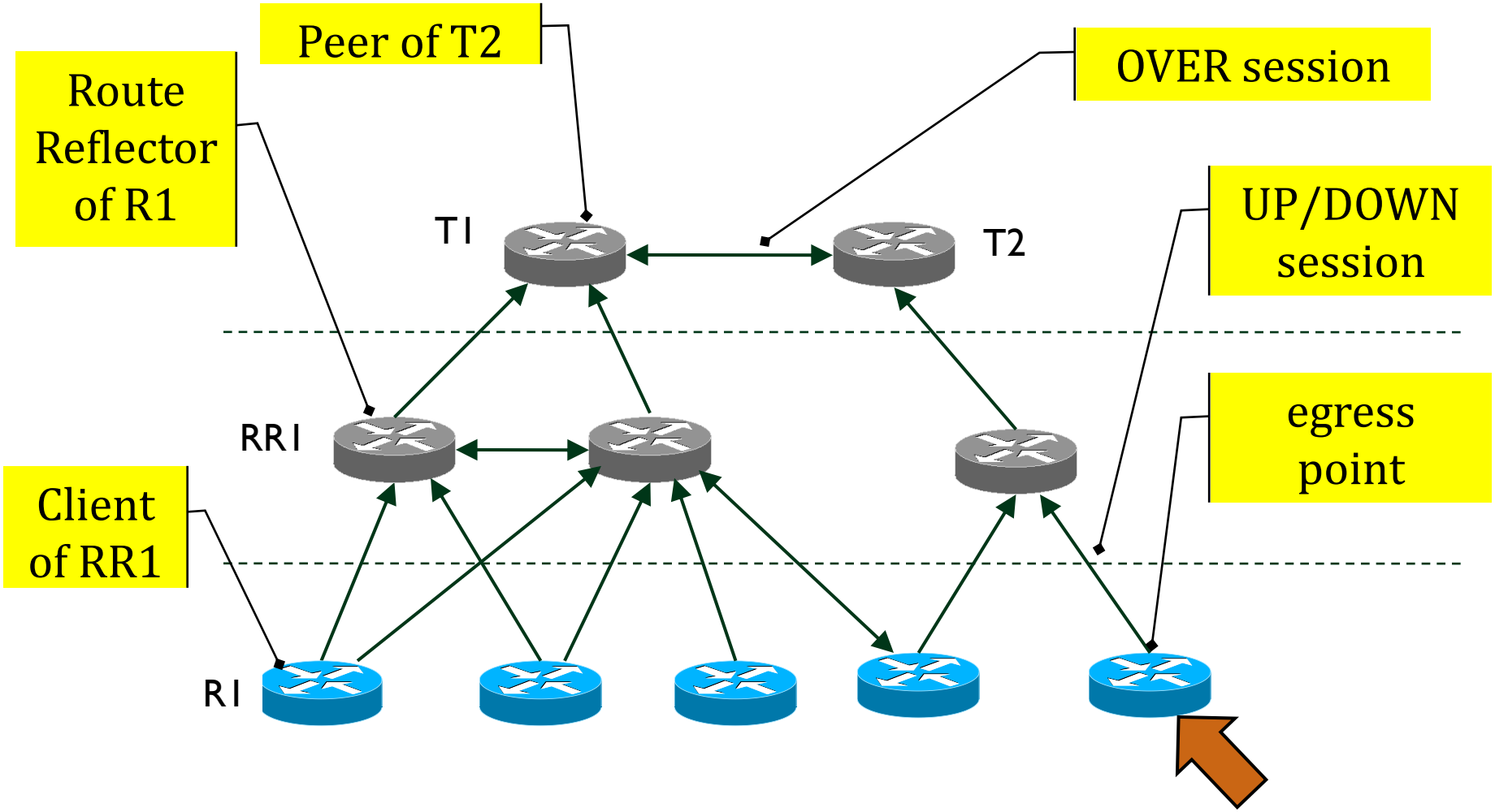
We refer to the organization of iBGP sessions as iBGP topology

iBGP Route Reflection (RR)

- to scale, routers are organized in a hierarchy
  - lower layers rely on upper layer for routing information
  - we assume proper hierarchy and top layer full mesh

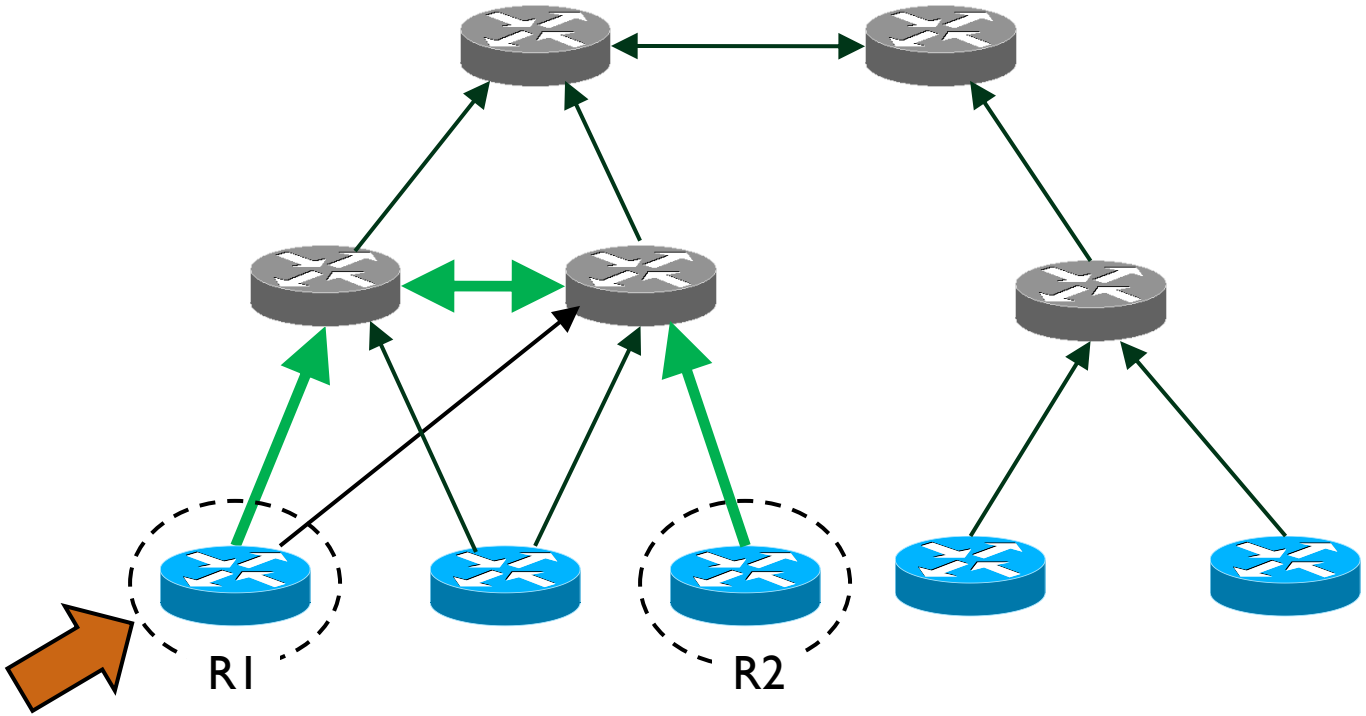


Terminology



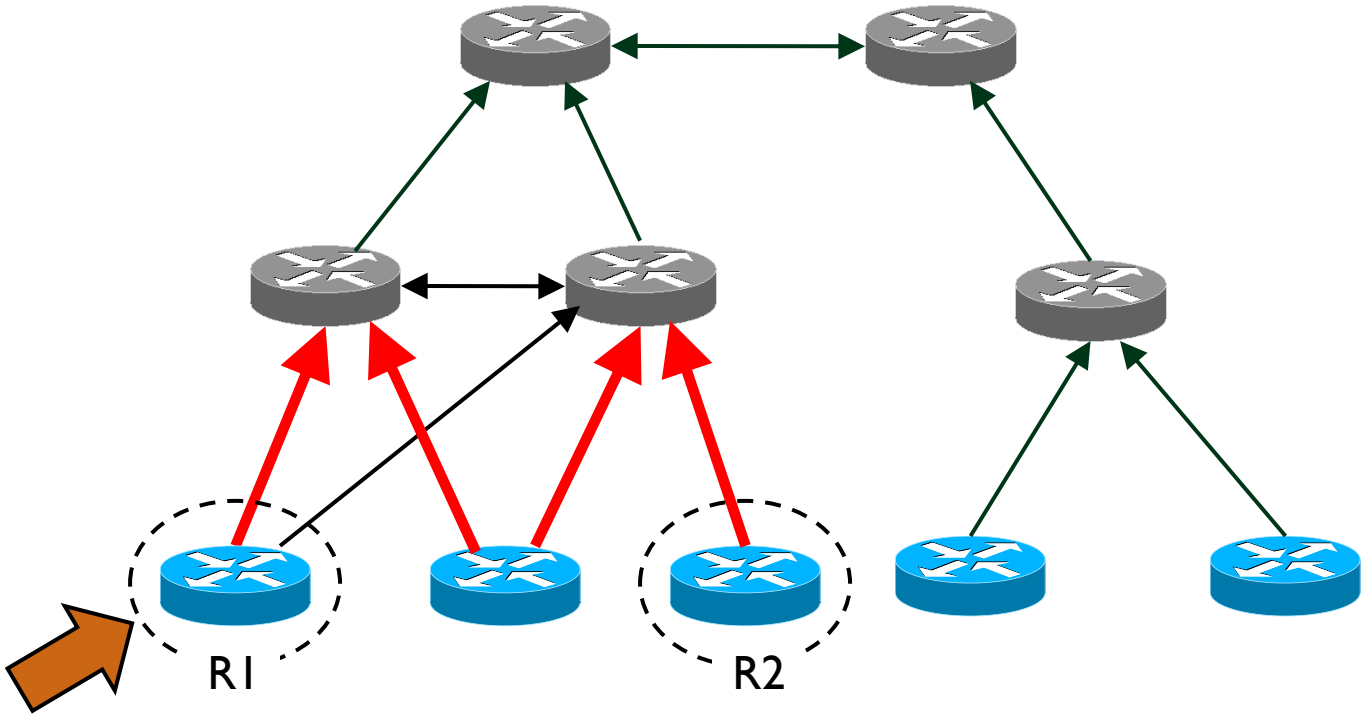
Route Propagation

- Only best routes are propagated on valid signaling paths
  - Routes from non-clients are not propagated to non-clients
  - Valid signaling paths: UP* OVER? DOWN*



Route Propagation

- Only best routes are propagated on valid signaling paths
  - Routes from non-clients are not propagated to non-clients
  - Valid signaling paths: UP* OVER? DOWN*



Configuration Correctness

- Route propagation rules → limited route visibility

- In turn, limited route visibility + IGP/iBGP coupling → routing and forwarding anomalies
  - signaling anomalies: iBGP never converges to a stable state [Griffin00]
  - dissemination anomalies: not all the routers receive a route to a given destination in the stable state [Vissicchio12]
  - forwarding anomalies: data packets incurs in deflections and forwarding loops in the stable state [Griffin00]

Problem Statement

Seamless BGP Reconfigurations

- Change BGP configurations …
  - need for evolution and network agility
- … with no impact on services
  - intermediate configurations should be anomaly-free
  - unintended traffic shifts should not occur

- in this talk, we focus on iBGP topology changes
  - similar considerations hold to eBGP policy modification

The Need for Evolution

- BGP configuration has to be changed regularly
  - e.g., PoP addition / inflation or commercial relationship change

- we performed a study on Tier-1 historical data
  - data set: 20% router configs, April 2010 – July 2011
  - 90 BGP config changes / month, on average
  - BGP session additions and removals are frequent
    - session additions: 4,800 in iBGP, 1,000 in eBGP
    - session removals: 1,400 in iBGP, 230 in eBGP
  - eBGP policy changes are less frequent (but still present)
    - policy name changes: 120 in total

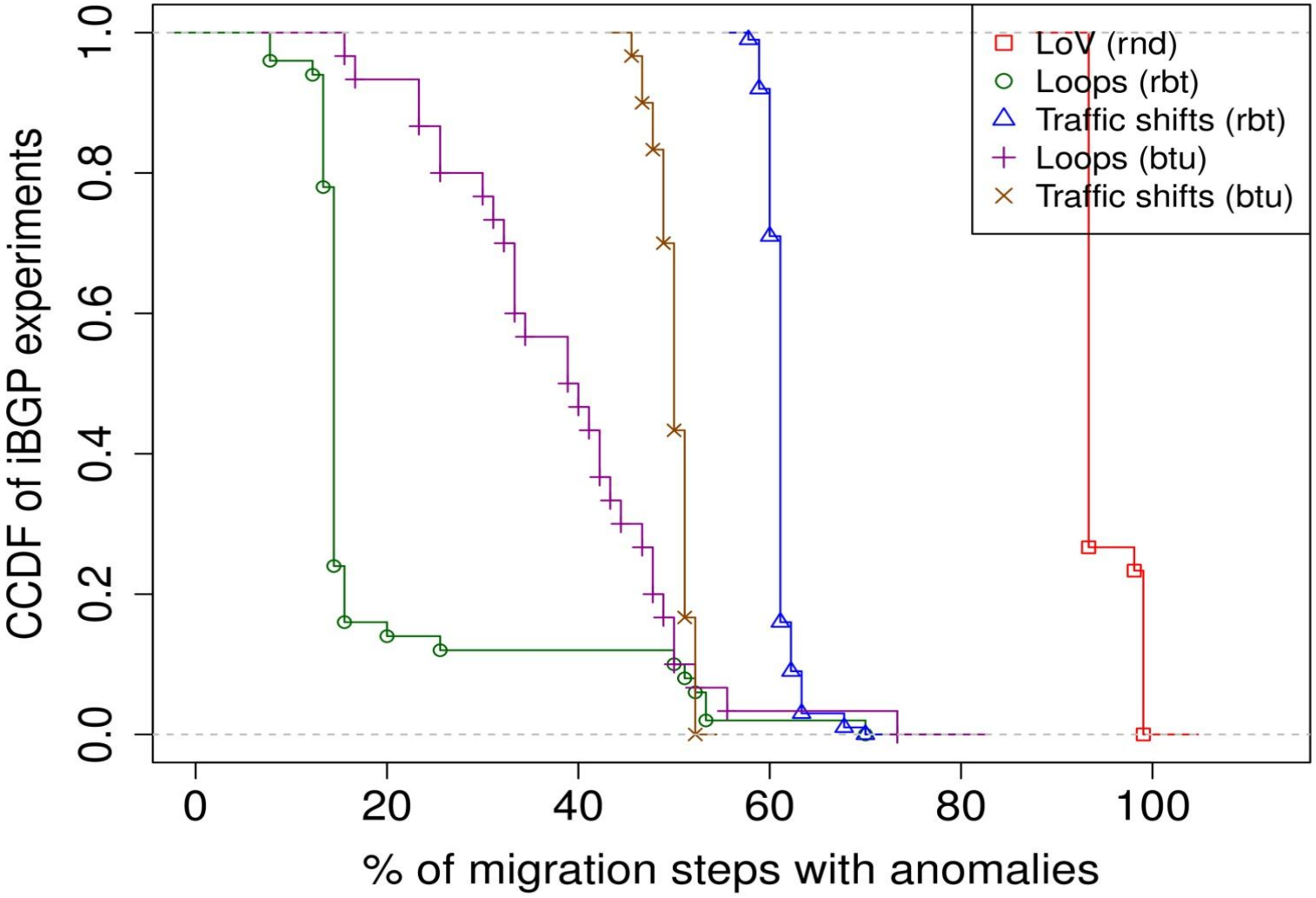
Network Agility

- possibilities of competitive advantage for ISPs
  - adapt routing to traffic flows
    - mirroring daily / nightly trends
    - supporting green networking
  - flexible inter-domain routing policies
    - reduce transit costs
    - flexible upstream connectivity

Current Best Practices

- rules of thumb for simple cases
  - e.g., full-mesh to 2-layer route reflection
- per-router reconfigurations
  - final configs are pushed to one route-reflector at the time
  - no operational ordering is specified
- however, operational ordering requires extreme care
  - anomalies can occur in intermediate configs
  - both theoretically and practically

Current Best Practices Do Not Work



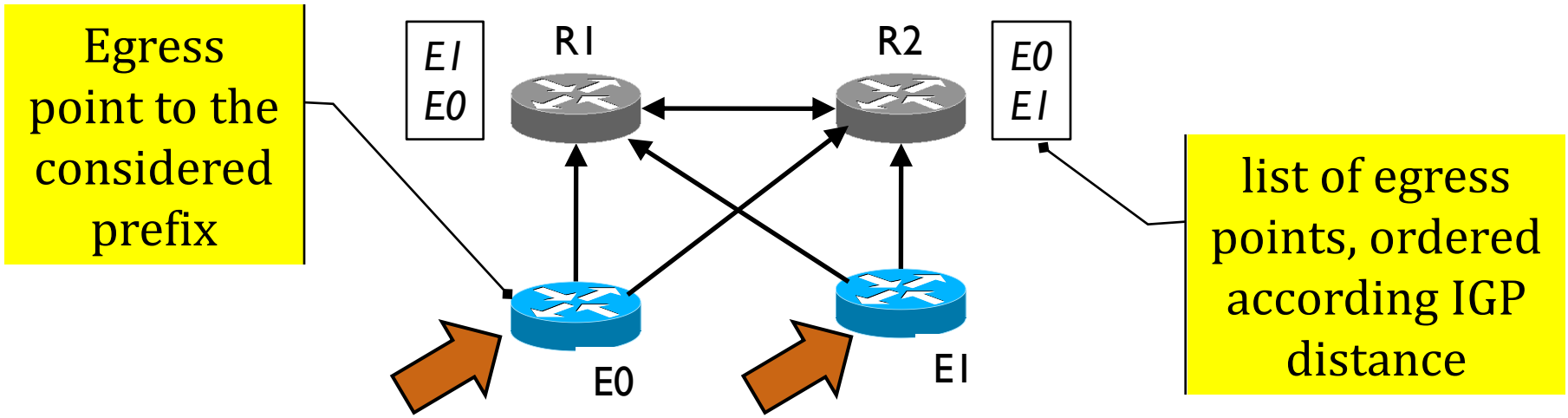
An Algorithmic Perspective

The Algorithmic Problem

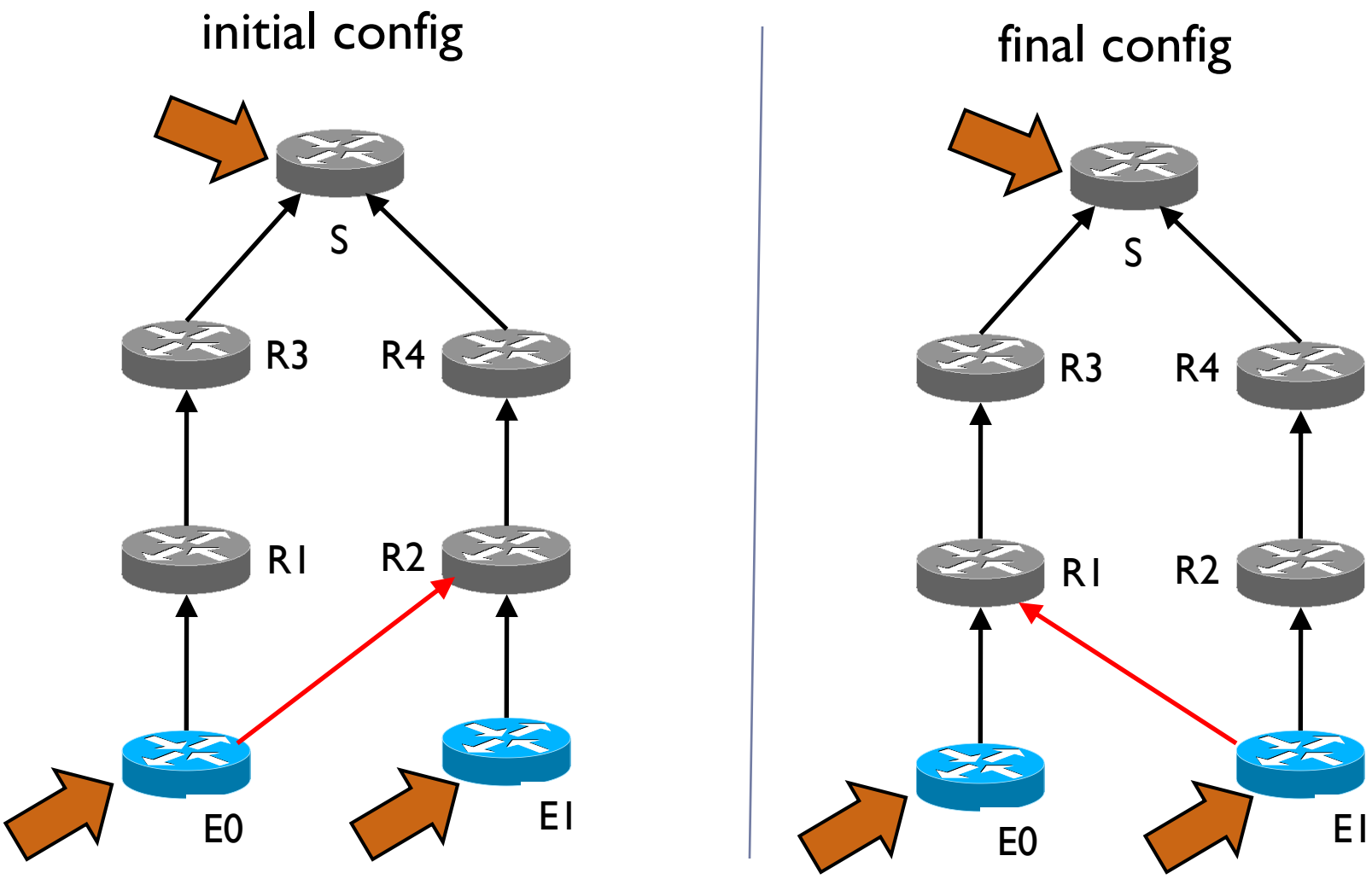
- *Input*: initial and final BGP configs
  - anomaly-free, by hypothesis
- *Problem*: compute an ordering in which to add / remove BGP sessions
  - the ordering must ensure a seamless reconfiguration

Model

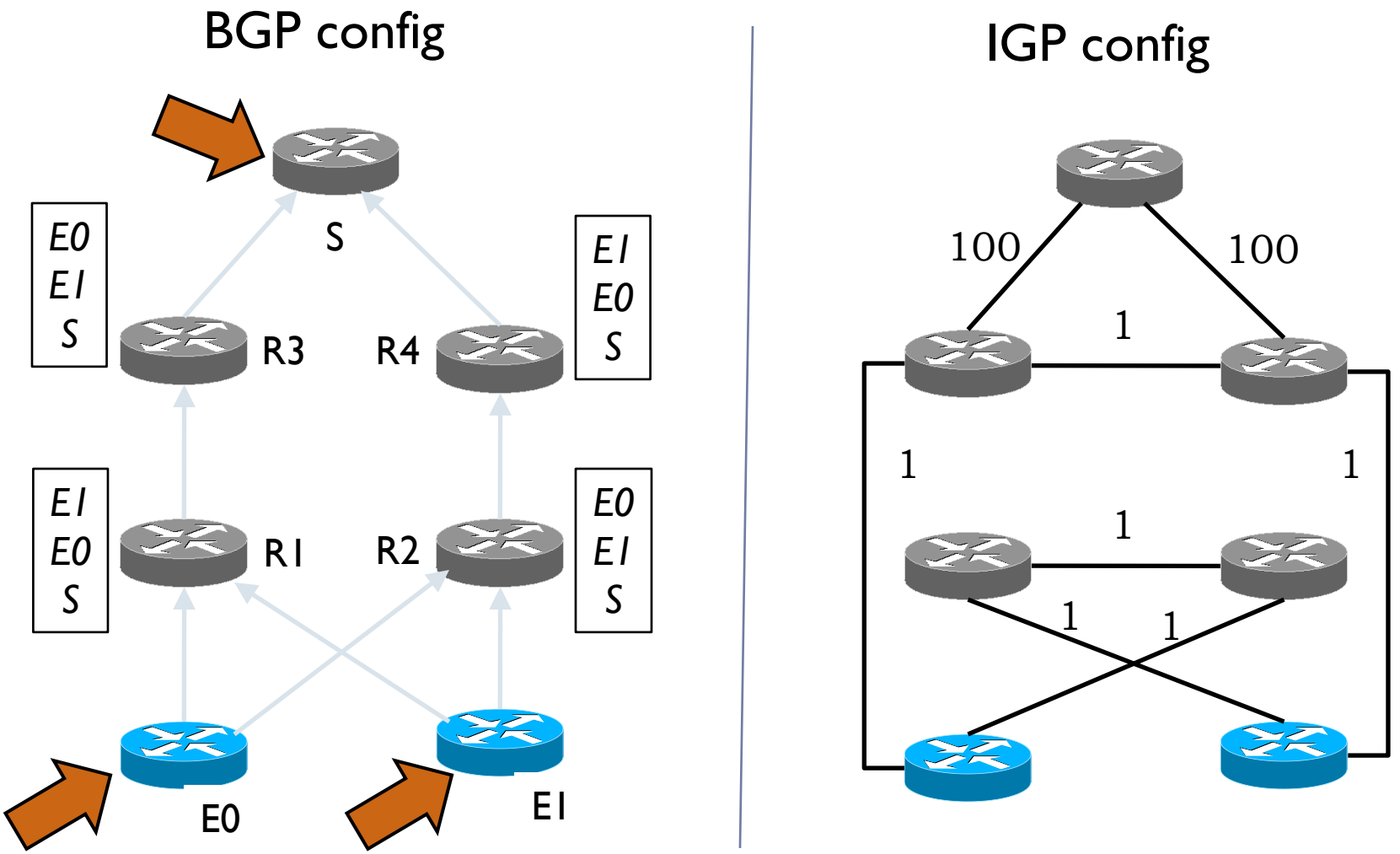
- iBGP configs are modeled as enriched graphs
  - nodes are routers, edges are BGP peerings
  - an ordered list of permitted paths per node
    - only valid signaling paths are permitted
  - path preferences reflect the iBGP decision process



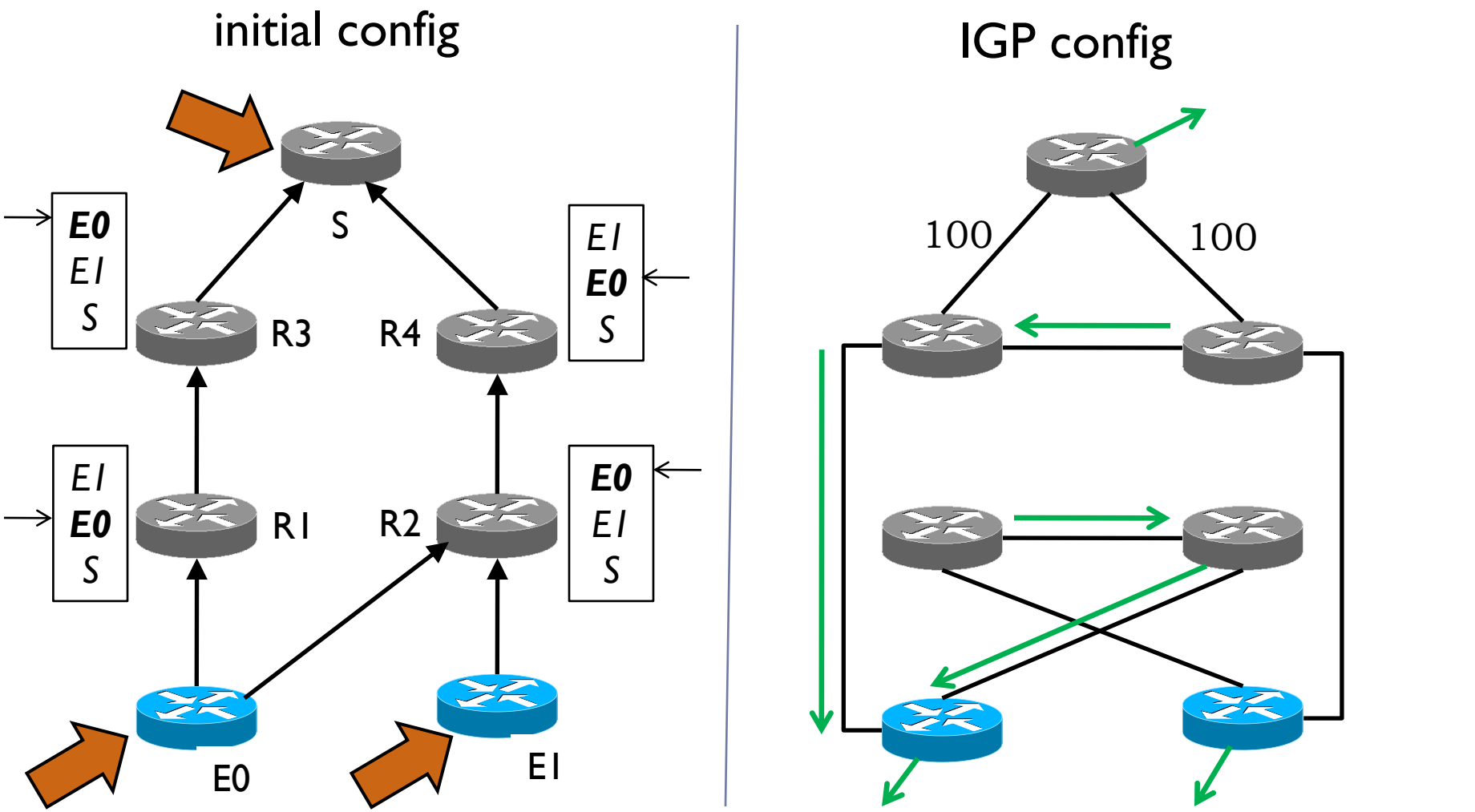
Non-Solvable Example



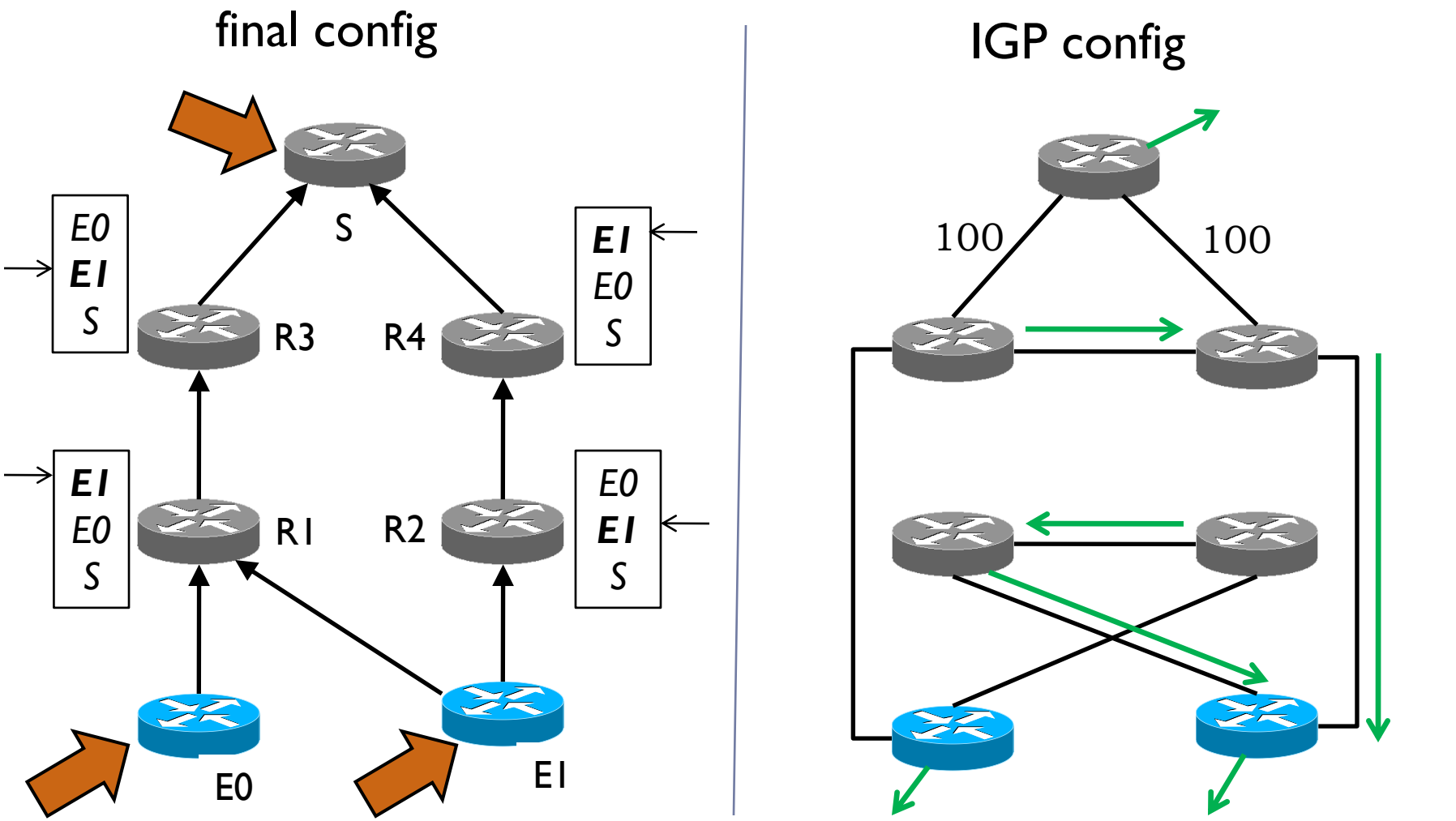
BGP Route Preferences



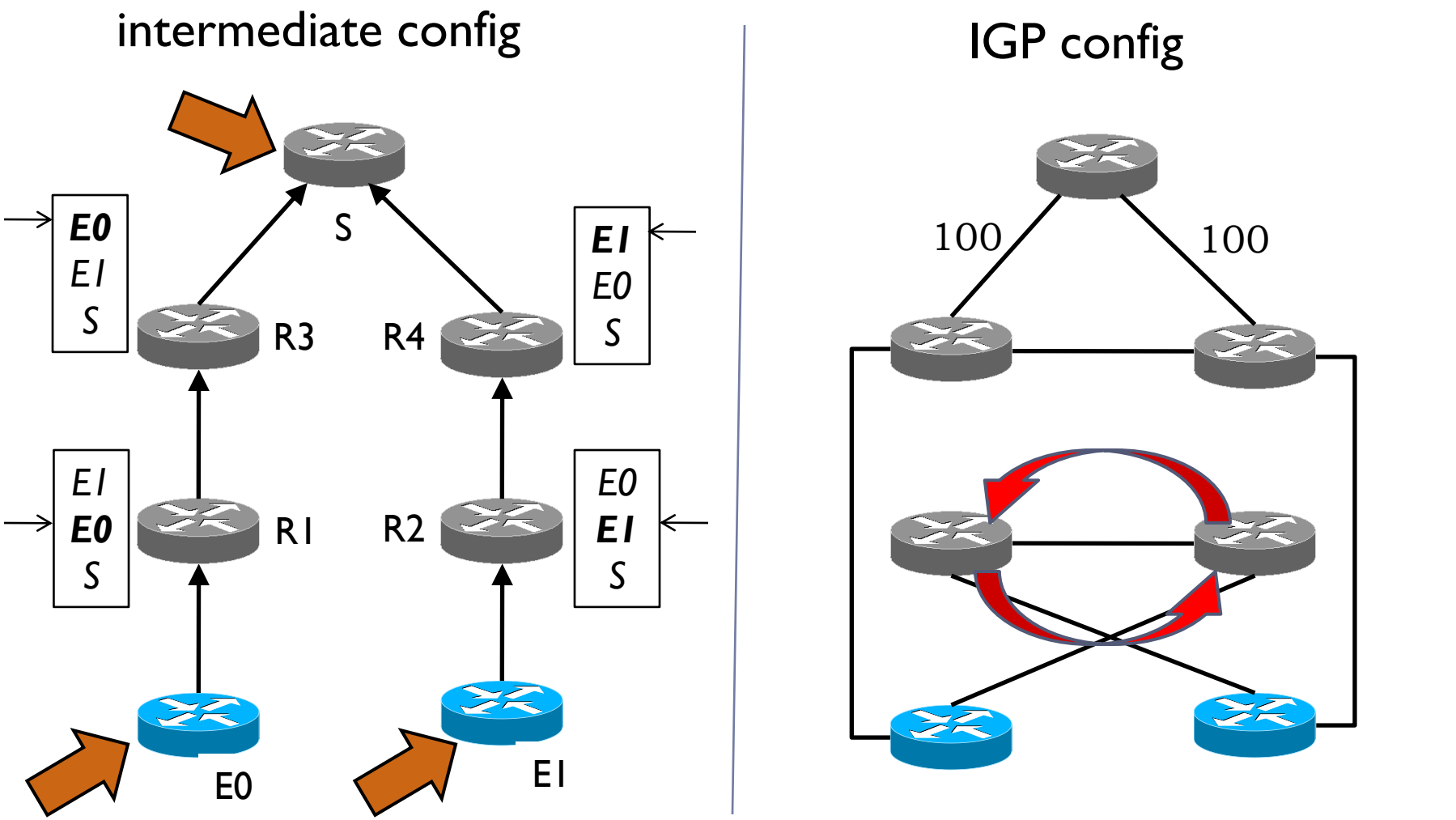
Initial Config is Anomaly-Free



Final Config is Anomaly-Free

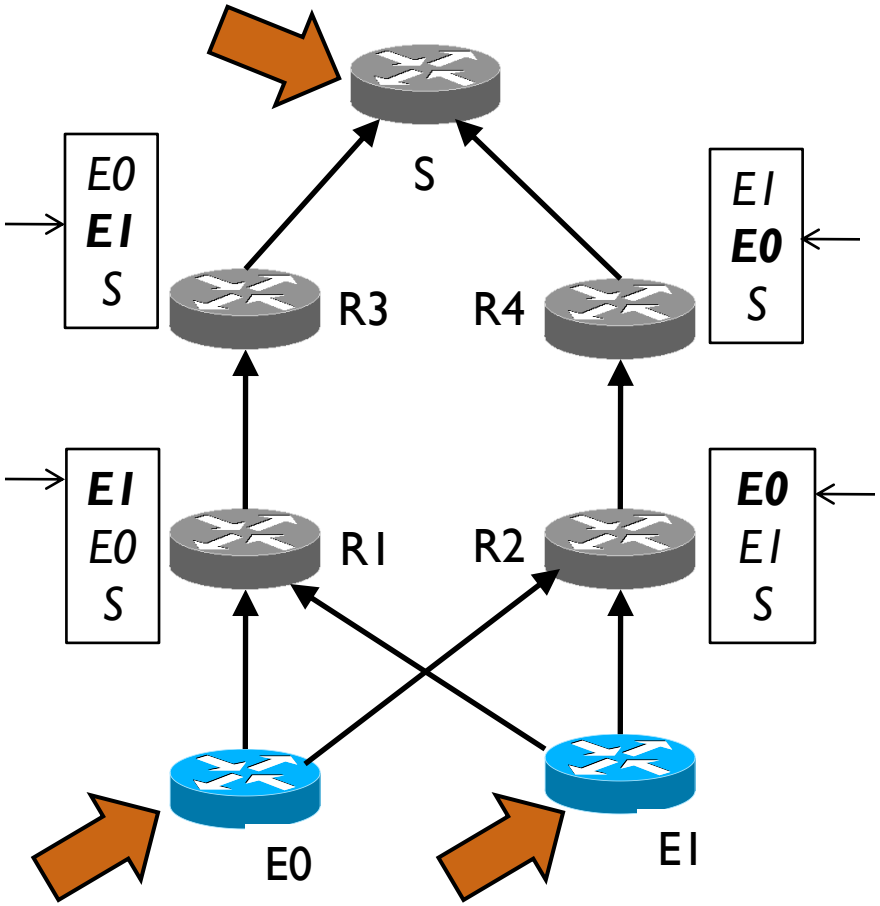


Case 1: Remove Before Adding

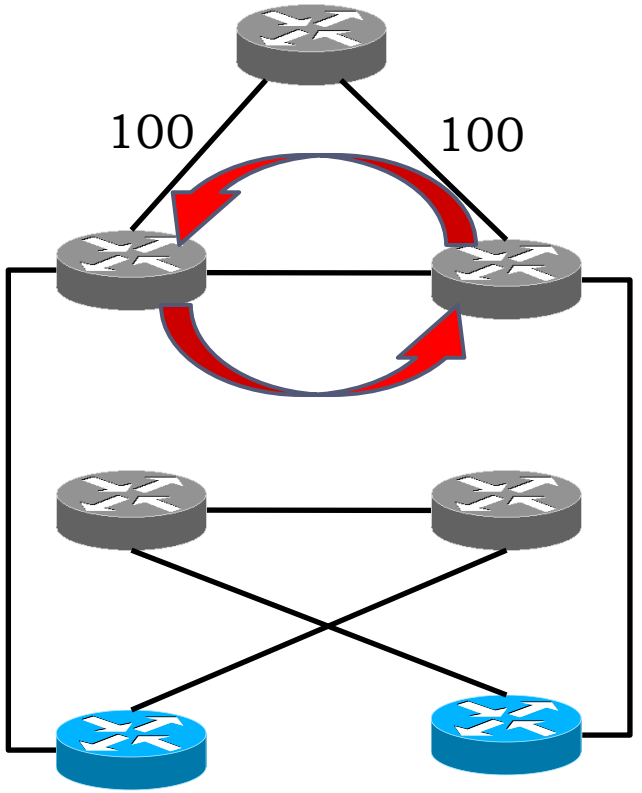


Case 2: Add Before Removing

intermediate config



IGP config



More in General

- Examples exist in which every intermediate config is not free from
  - signaling anomalies
  - dissemination anomalies
  - forwarding anomalies
  - unintended traffic shifts
  - a combination of the above anomalies
- Those examples exist for both iBGP and eBGP reconfigurations

Complexity of the Problem

- Deciding if a ordering free from signaling anomalies exists is (at least) NP-hard
  - P-time reduction from 3-SAT

- The same reduction applies to
  - dissemination- and forwarding-free orderings
  - cases in which only egress points are added
  - iBGP and eBGP config changes

Proposed Solution

Basic Idea

- *main issue*: local changes affect remote routing decisions
- *key idea:* BGP Ships-In-The-Night (SITN)
  - two BGP processes running at the same time
  - reconfigure by switching from one process to the other
- *related ideas*
  - IGP reconfiguration techniques [Herrero10, Vanbever11]
    - but applied to BGP
  - Shadow Configuration [Alimi08]
    - but more lightweight, and working today

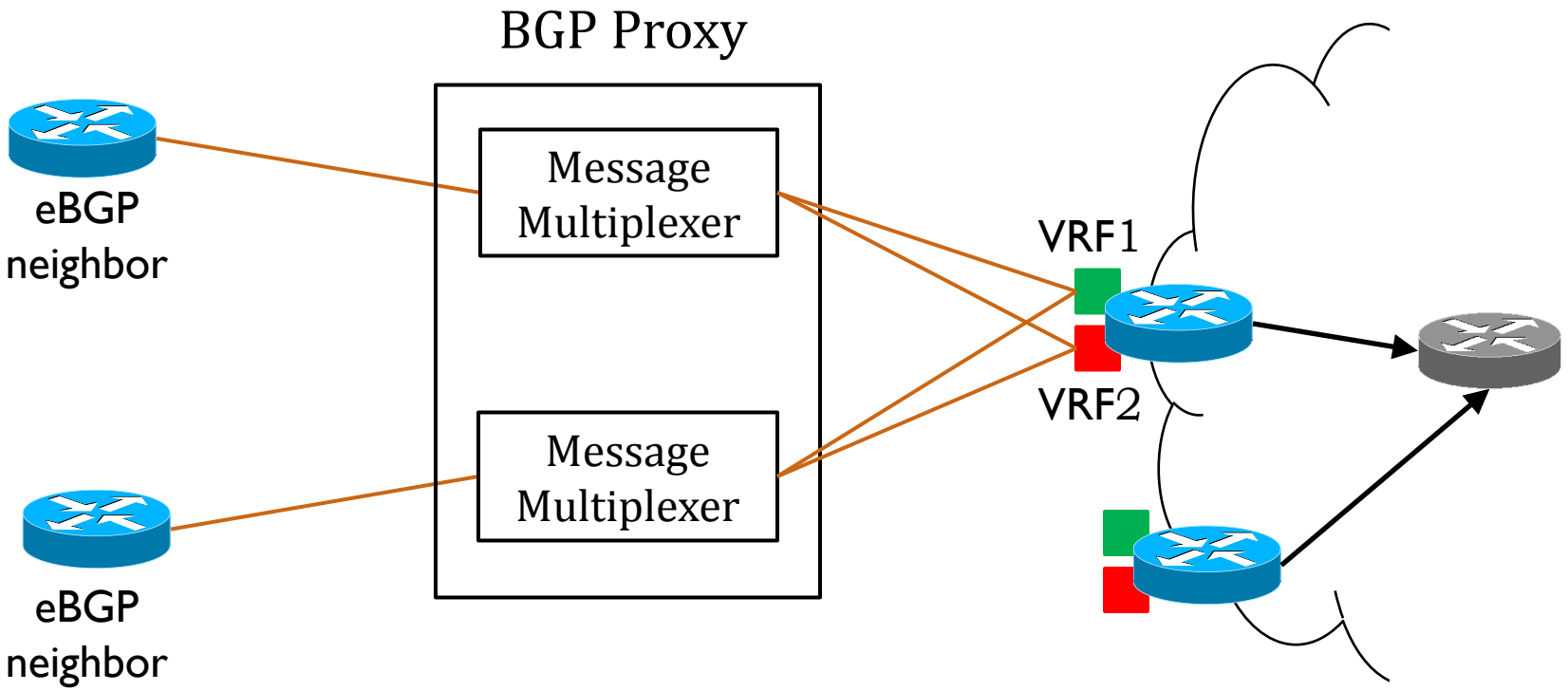
Requirements for SITN

- multiple isolated routing processes
  - inside each BGP router
- independent propagation of routes
  - to all the processes within the same router
- network-wide consistent forwarding
  - each packet must be forwarded according to only one process

Matching the Requirements

- multiple isolated routing processes
  - VRFs provide isolated routing namespaces
- independent propagation of routes
  - only best routes can be imported / exported between VRFs
- network-wide consistent forwarding
  - tagging mechanisms are available
    - packet encapsulation in GRE or MPLS
    - implicit tagging, e.g., VRF-lite

The Solution in a Nutshell

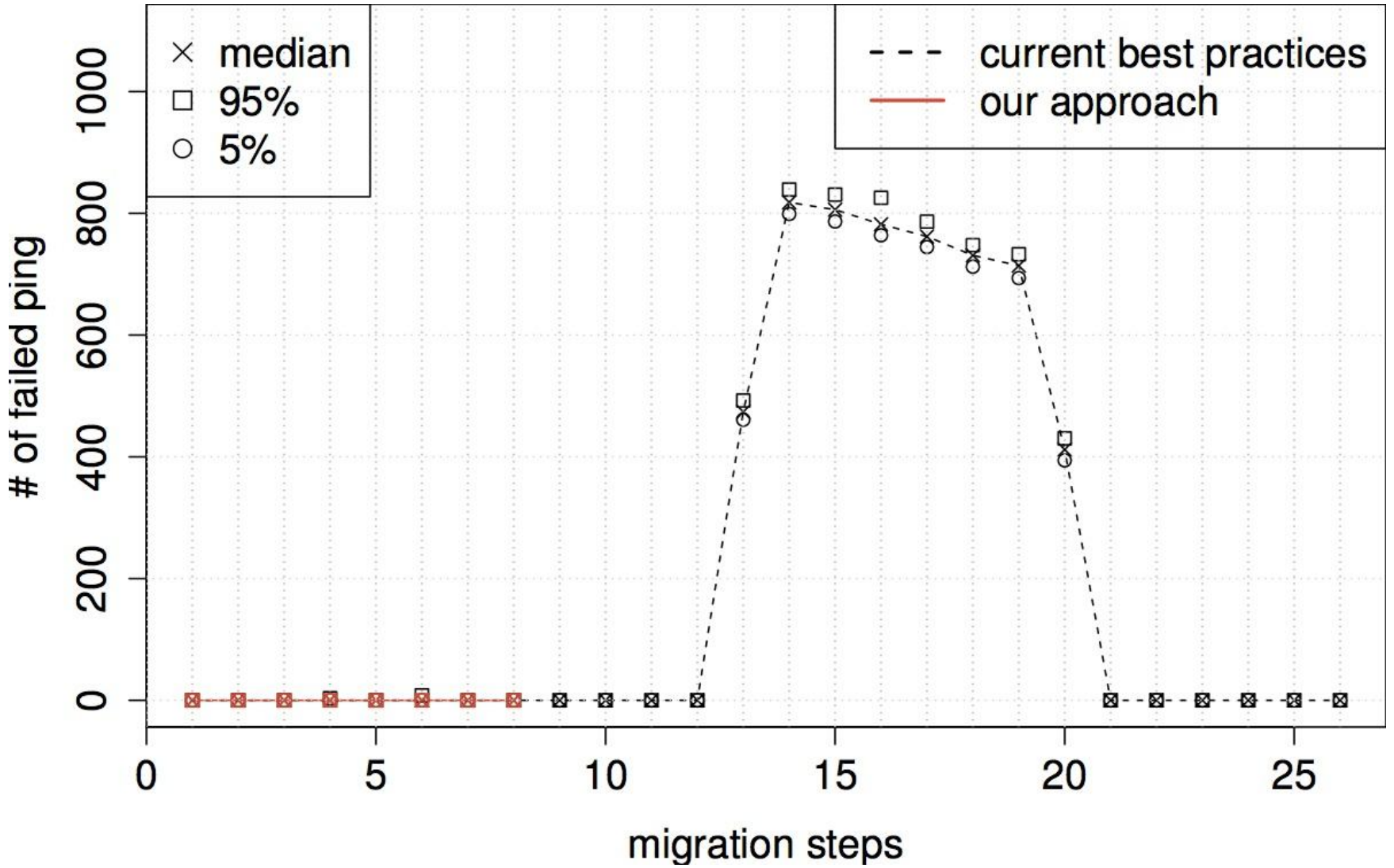


We implemented a prototype BGP proxy in Perl (400 lines) and a provisioning system as an extension of NCGuard [Vanbever08]

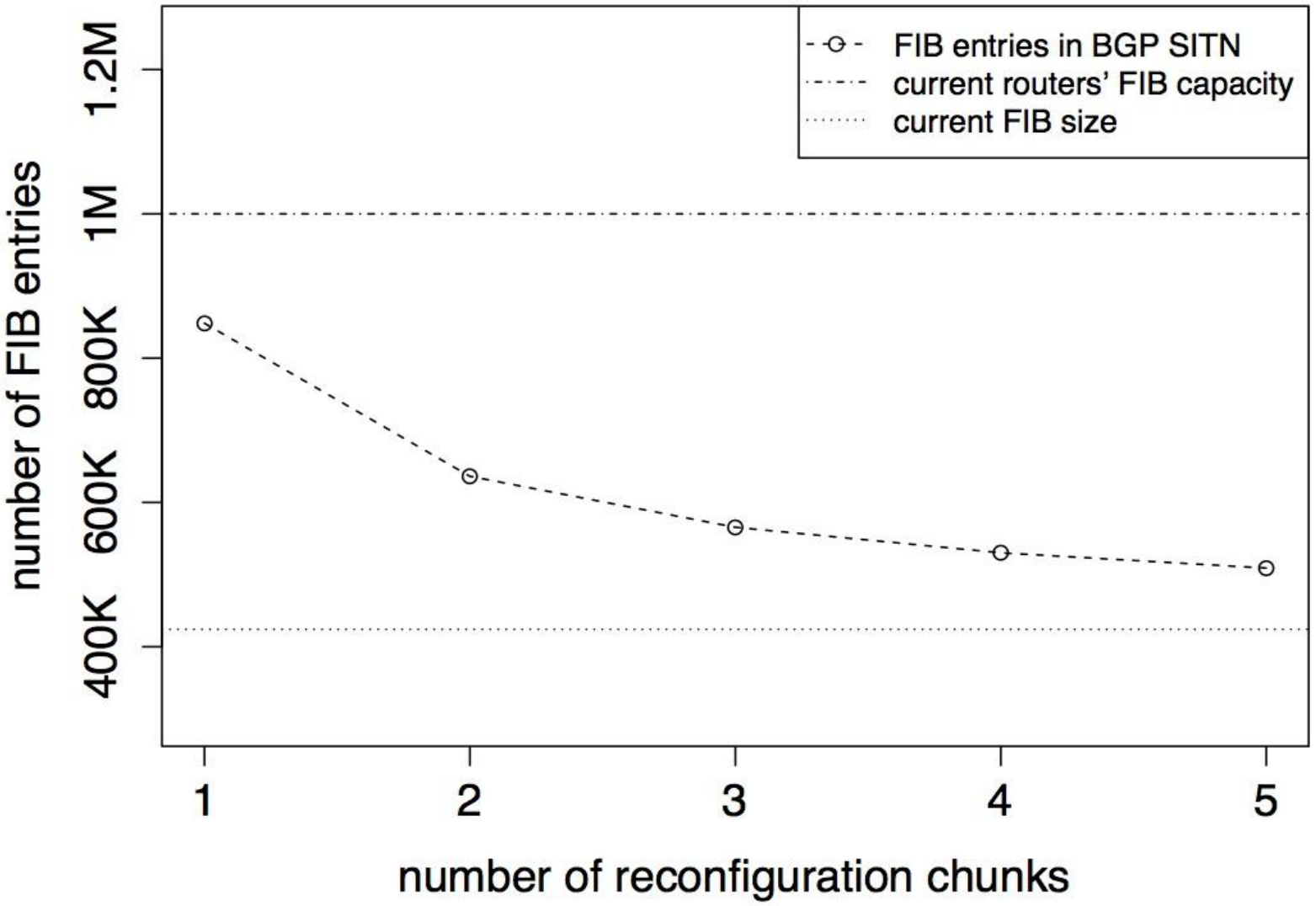
Case Study

- Reconfiguration of the Geant topology
  - from BGP full-mesh to a three-layer RR hierarchy
  - simulation in a virtual environment
- Methodology
  - IGP topology is consistent with L1
  - geographically-based RR topology
- Experiments
  - comparison between current best practices and our approach

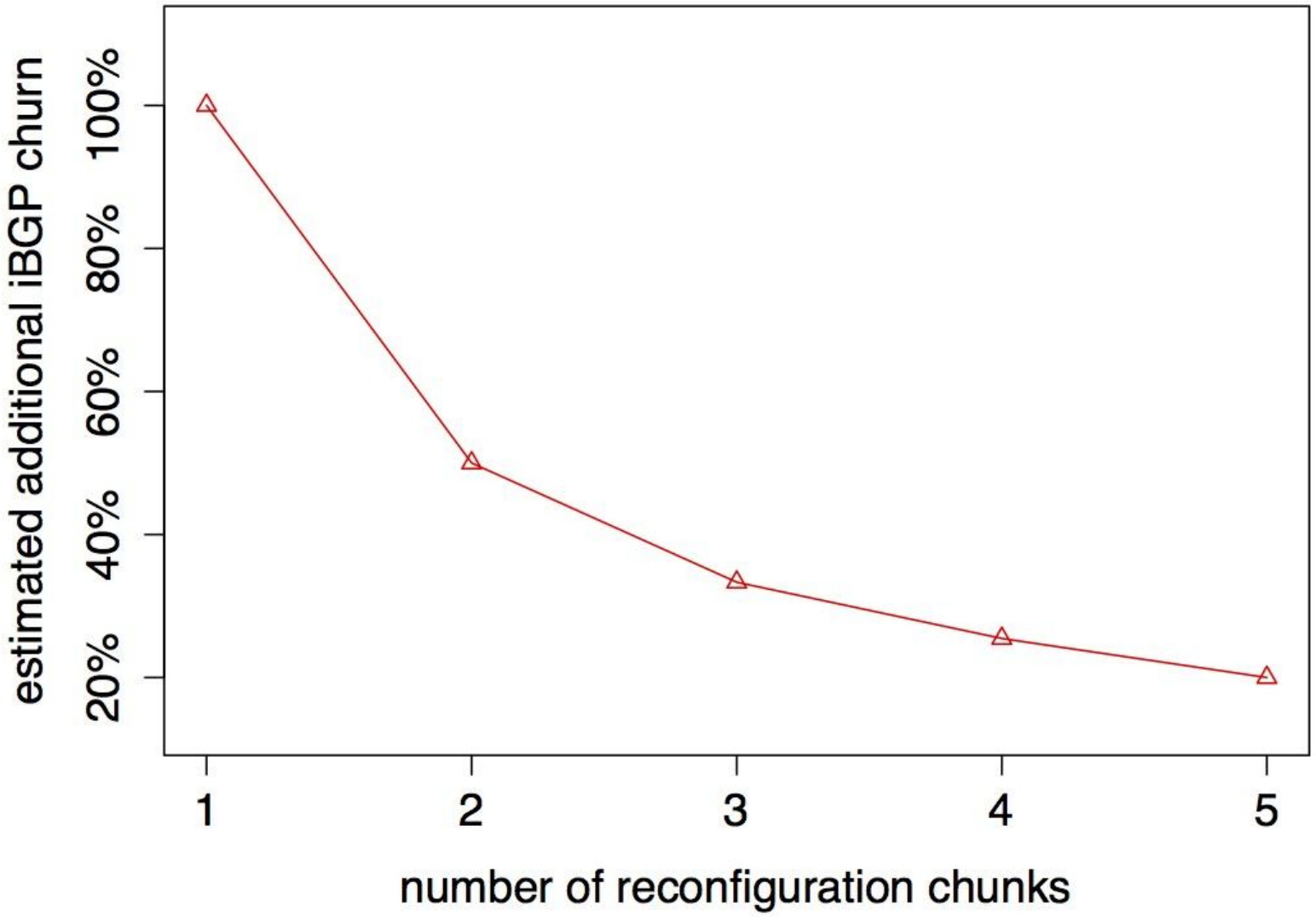
Experimental Results



Scalability (FIB)



Scalability (Churn)



Conclusions

- First study of the BGP reconfiguration problem
  - BGP reconfiguration is a need for operators
  - ensuring service continuity is challenging
- A suitable operational ordering does not always exist
  - CBPs do not work in the general case
  - deciding if such an ordering exists is computationally hard
- *Proposal*: leverage two BGP control-planes
  - VRF-based framework + working prototype
  - consistent forwarding by packet tagging

Open Directions and Future Work

- Algorithmic approach to simpler cases
  - simple reconfigs (e.g., full-mesh to RR)
  - stronger assumptions on the initial and final configs
- Deployment in real networks
  - real-world case studies
- Extension to different scenarios
  - AS renumbering
  - combined IGP and BGP reconfigurations

Thank You !!

- Contacts:

ste.vissicchio@gmail.com