Internet Network Management Workshop (INM/WREN) – 27 April 2010

Stefano Vissicchio, Luca Cittadini, Maurizio Pizzonia,
Luca Vergantini, Valerio Mezzapesa, Maria Luisa Papagni
Università degli Studi RomaTre

BEYOND THE BEST: REAL-TIME NON-INVASIVE COLLECTION OF BGP MESSAGES

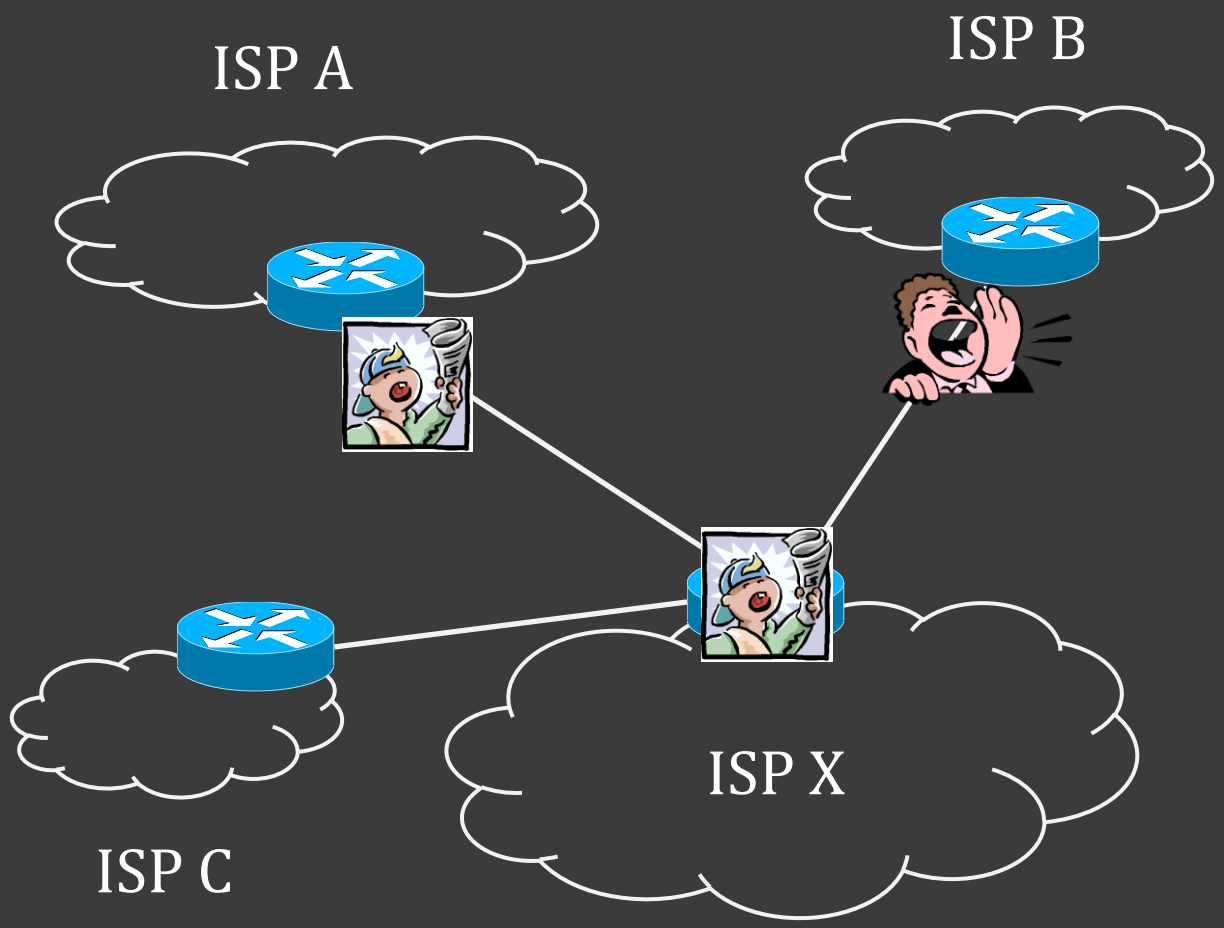
Interdomain Routing = BGP

- BGP is the Internet glue
  - de-facto standard for interdomain routing
- BGP decides traffic forwarding in the Internet
  - BGP has a non-negligible economic impact on the business of the ISPs
- BGP monitoring is crucial for ISPs
  - several applications, from troubleshooting [Roughan04] to traffic engineering [Balon08] and SLA compliance [Feamster04]

Overview

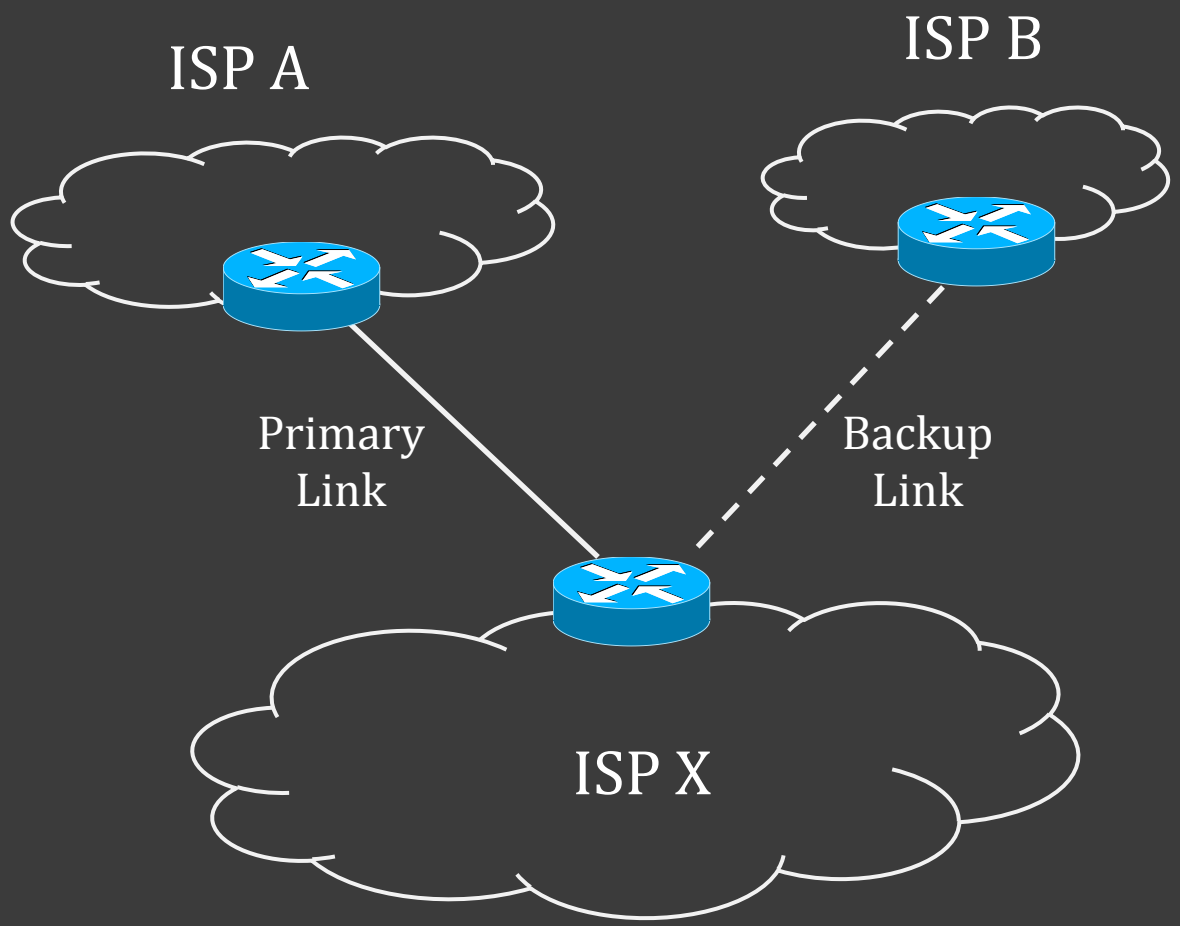
- We identify the basic requirements for an ideal monitoring system
  - cost-effective system for the collection of **all** BGP messages as sent by neighboring ISPs
- We proposed a monitoring infrastructure
  - routers are mandated to copy TCP segments and an ad-hoc software collect and store them
    - exploit an already available feature
  - easily extendable to other protocols
- We experimentally evaluate our solution

BGP Routes Propagation



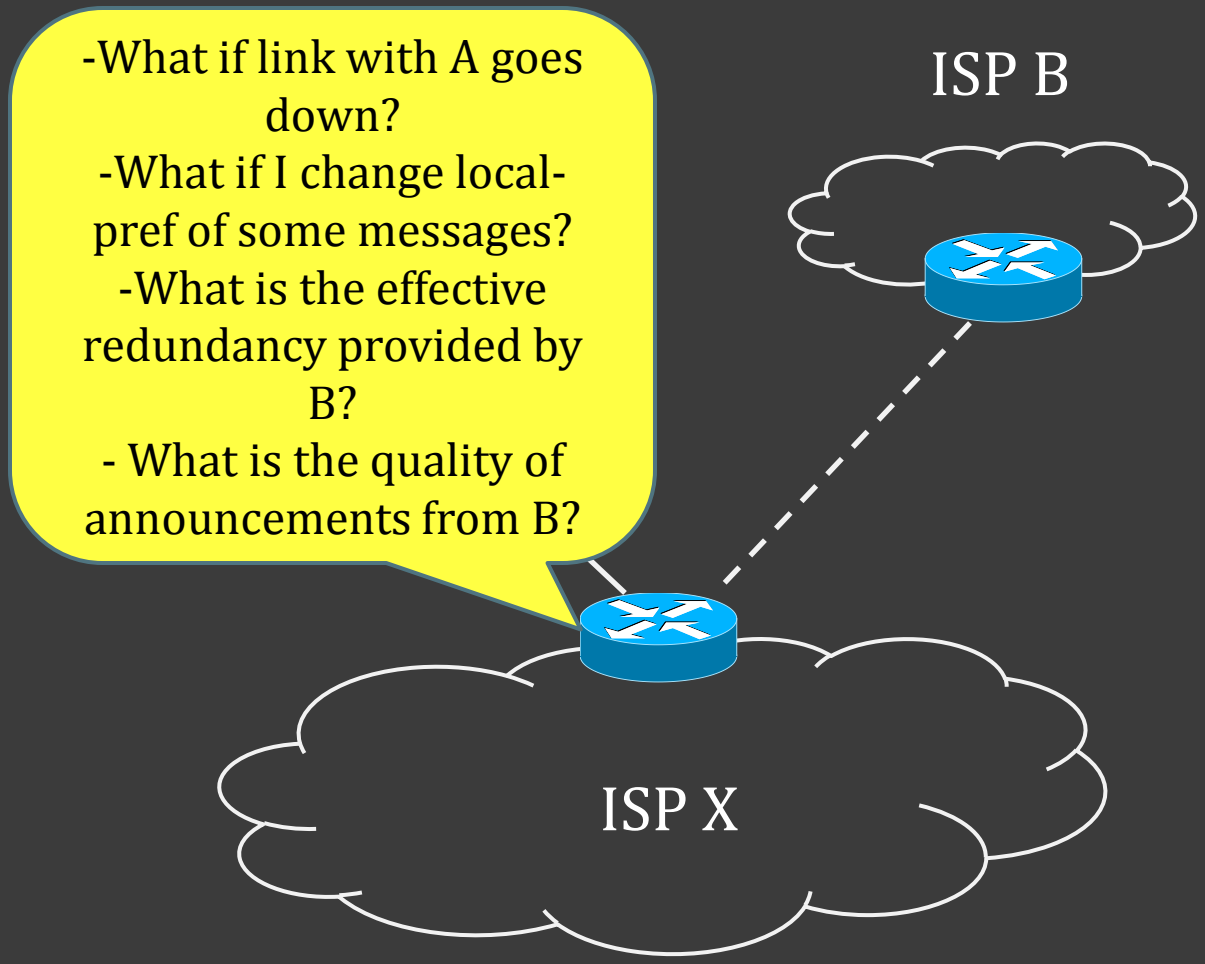
- for each destination, BGP routers receive a set of announcements
- each BGP router autonomously selects the best route among them
  - best routes control traffic flow
- … and propagates it to its neighbors

Monitoring BGP Best Routes



- monitor BGP messages
  - quality
  - SLA
  - history
- check egress traffic flow
- … but only on the primary link

Monitoring *All* BGP Routes



- monitor BGP messages on both links
  - quality
  - SLA
  - history
- X is enabled to analyze what-if scenarios, check SLA compliance for A and B, perform other value-added activities

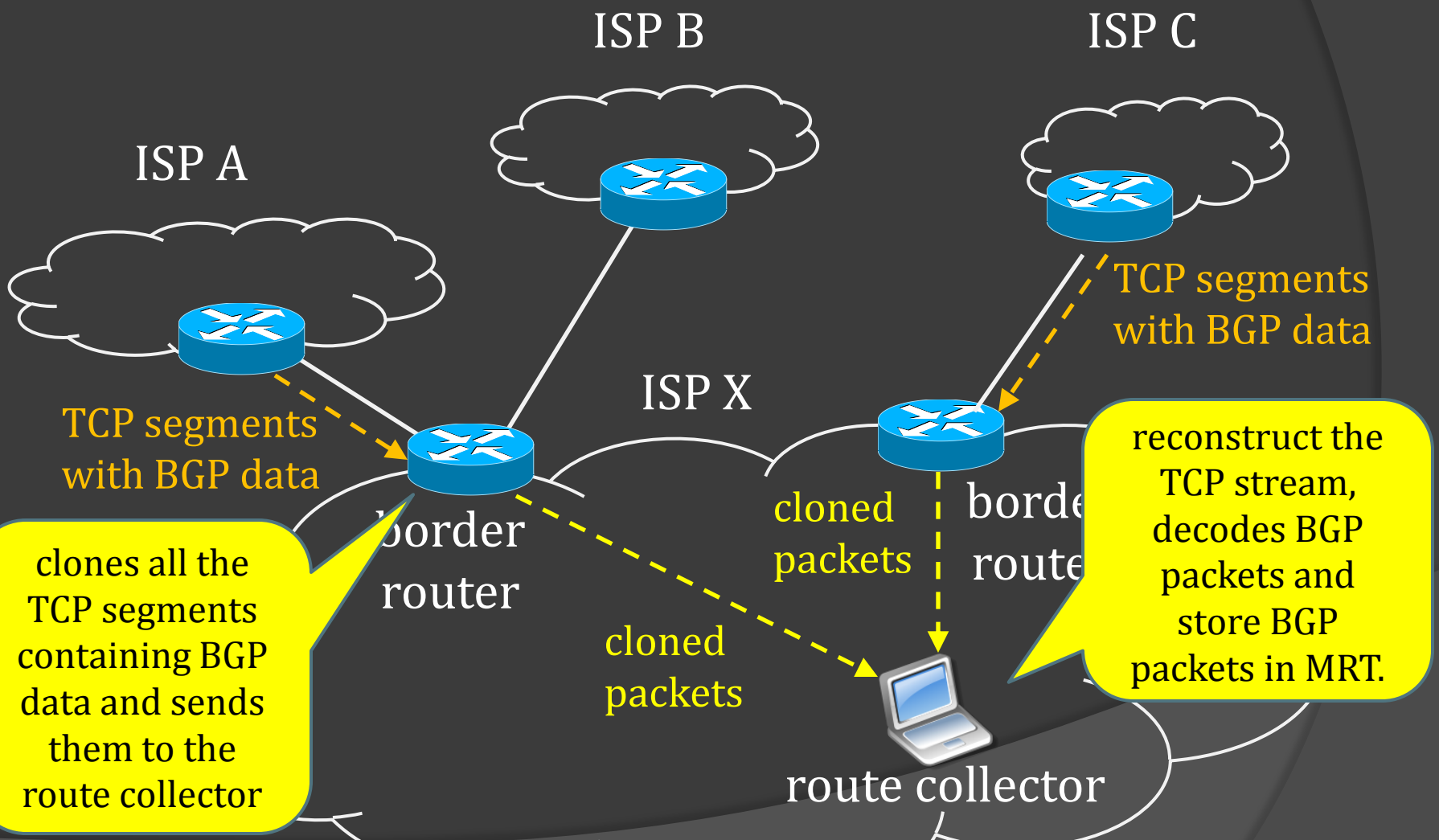
An Ideal Monitoring System

- Collection of all the BGP routes
- Policy independent data
- Real-time collection
- Low impact on router resources
- Cost-efficient deployment

Existing Monitoring Systems

- a collector maintains iBGP peerings with routers that push data to it
  - open source daemons (Quagga, Pyrt, …)
  - not possible to collect all the messages and policy independent data
- a separate management protocol can be used to pull information from routers
  - SNMP, screen scraping
  - heavy impact on routers, can not be real-time
- BMP (comparison in the following)

Proposed Architecture



Border Routers

- border routers have to selectively clone incoming traffic to a destination
  - supported by major vendors on most routers
    - RITE/ERSPAN (Cisco), port mirroring (Juniper)
    - originally designed for supporting IDSes
  - cloned packets can typically be sent to the collector via VLANs or IP tunnels
  - management overhead is limited

Configuring Border Routers

```
access-list 100 permit tcp any any
                                eq bgp
```

define traffic to be cloned

```
ip traffic-export profile <pr-name>
    interface <vlan-interface>
    incoming access-list 100
        mac-address <addr>
```

configure destination interface

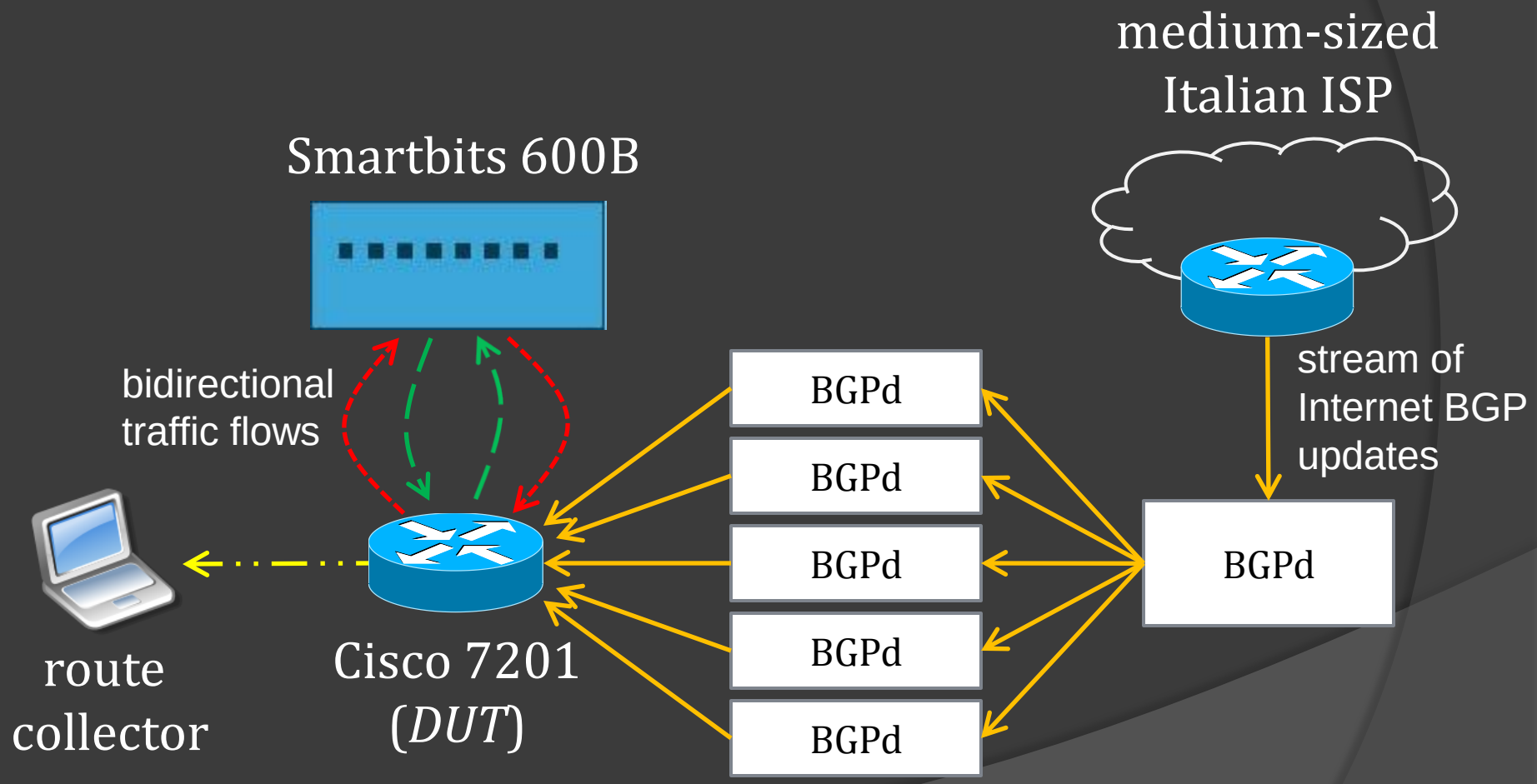
```
interface <src-interface>
    ip traffic-export apply <pr-name>
```

select source interfaces

Route Collector

- the route collector has to reconstruct the TCP stream and to decode and store BGP messages
  - TCP segments are reordered and duplicated packets are silently ignored
  - prototype based on two Perl scripts
    - the first script reconstruct the TCP stream
    - the second script decodes and stores BGP packets in MRT

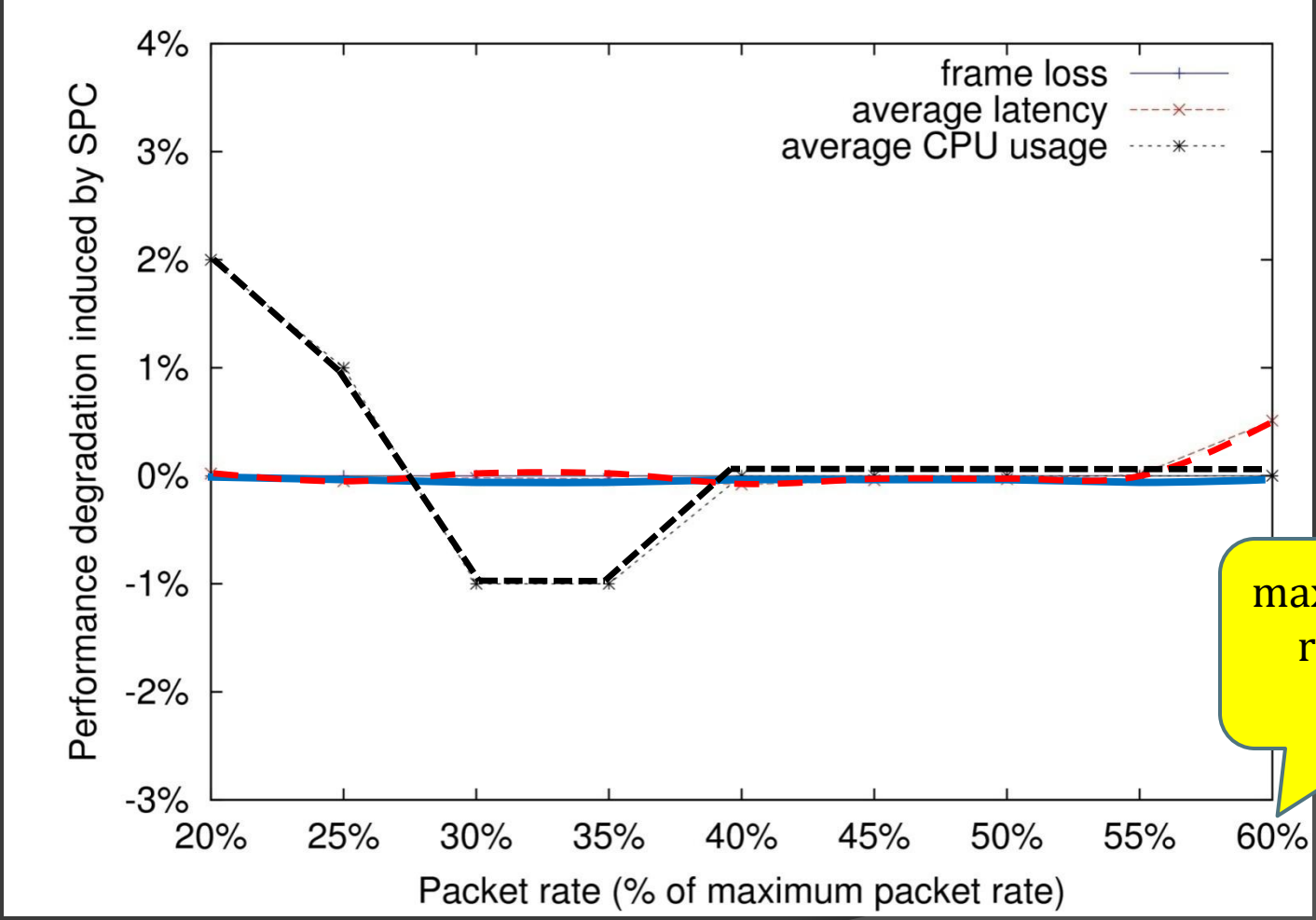
Testbed


medium-sized
Italian ISP
Smartbits 600B
bidirectional
traffic flows
stream of
Internet BGP
updates
BGPd
BGPd
BGPd
BGPd
BGPd
BGPd
route
collector
Cisco 7201
(DUT)

Evaluation of our Solution

- We checked solution for correctness
  - no cloned packet was dropped
  - BGP messages were always correctly reconstructed and stored on disk
- We also evaluate performance of both border routers and route collector
  - throughput
  - CPU usage
  - latency

Evaluation: Border Routers



Evaluation: Route Collector

- Transfer of five full BGP RIBs is replayed using tcpreplay at top speed

	original transfer	tcpreplay	stream reconstruction	BGP decoding and storage
elapsed time	> 2 min	3.38 sec	2.6 sec	1.7 sec

- A single route collector can handle hundreds of border routers
  - processing a single prefix took about 5 µsec
- Performance can be further improved

Comparison with Related Work

	BGP daemons (Quagga, Pyrt)	SNMP screen scraping	Our Approach and BMP
non-best collection			
policy independency			
real-time			
impact on router resources			
cost efficiency			

Detailed Comparison with BMP

- Our solution pushes complexity to the collector side

	BMP	Our Approach
solution deployability	Internet draft, not widely supported yet	readily deployable
reliable delivery to the collector	yes, TCP connection	only check for lost packets
router performance	additional daemon, routers maintain a state	leverage optimized switching mechanisms
extendability to other protocols	extensions require software changes	easily extendable

Conclusions and Future Work

- what is the impact on production networks?
  - we exploit optimized packet copying mechanisms
  - experimental results are promising
  - a couple of companies already contacted us
- we plan to
  - deploy this solution in real networks
  - extend the approach to monitor all the control plane
  - integrate with iBGPlay: www.ibgplay.org

Thank you!!

- Questions?