Morpheus: Enabling Flexible Interdomain Routing Policies

Yi Wang
Ioannis Avramopoulos
Jennifer Rexford
Princeton University

NANOG42
02/20/2008
http://www.cs.princeton.edu/research/techreps/TR-802-07

Large ISPs Have Rich Path Diversity

- Top 2% ASes have 10 or more AS paths for certain destinations [SIGCOMM'06]
- 5-10 router-level paths per prefix is common in large ISPs [survey on NANOG mailinglist, April 2007]
- 20 router-level paths per prefix on average in a tier-1 ISP [USENIX'2007]

Paths May Differ Significantly

- Security
  - Prefix / sub-prefix hijacking is a real threat
  - Avoiding an undesirable AS along the path
  - Large ASes are likely to have at least one valid / desirable route for most prefixes
- Performance
  - Alternative BGP paths often have better performance than the default path [PAM'07]
- Path diversity gives large ISPs plenty of choices

Convert Path Diversity into Revenue

- Different customers may want different paths
  - Financial companies: secure paths
  - VoIP / online-gaming providers: low latency paths
  - Content providers: high BW paths
  - Many others: any paths with low cost

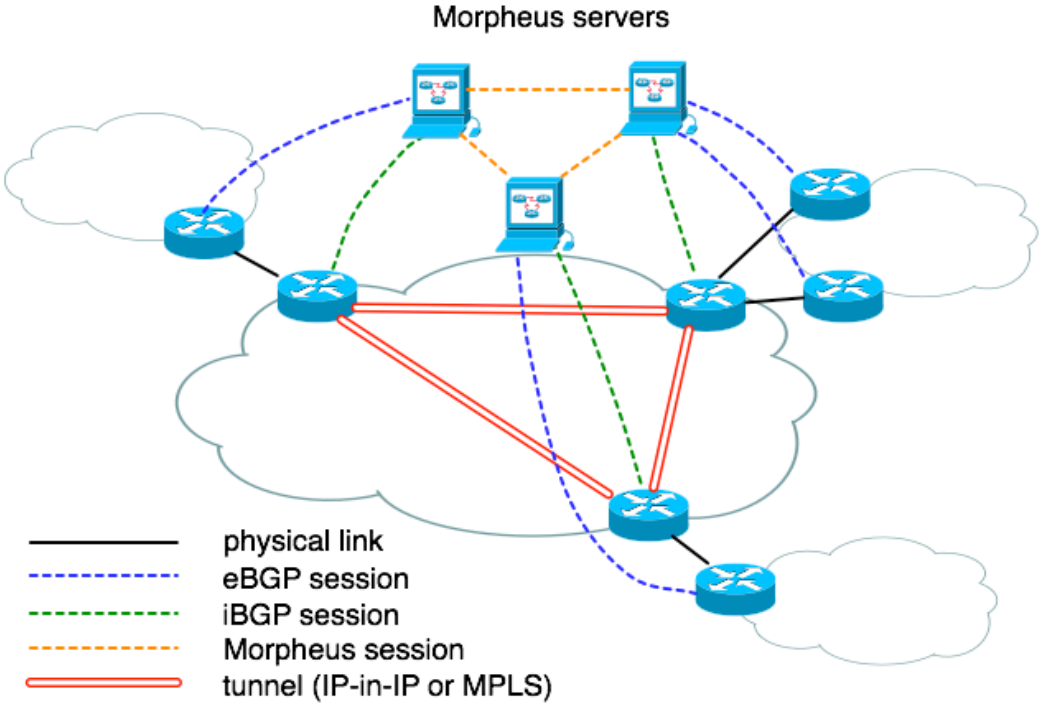
- Unfortunately, large ISPs cannot capitalize their path diversity today
  - One “best” BGP route for all

Morpheus: Enable Flexible Path Selection

- A routing control platform that enables a single ISP to flexibly pick paths for customers

- Two components
  - Supports from intra-AS routing architecture
  - Morpheus servers with flexible path selection processes

Intra-AS Routing Architecture

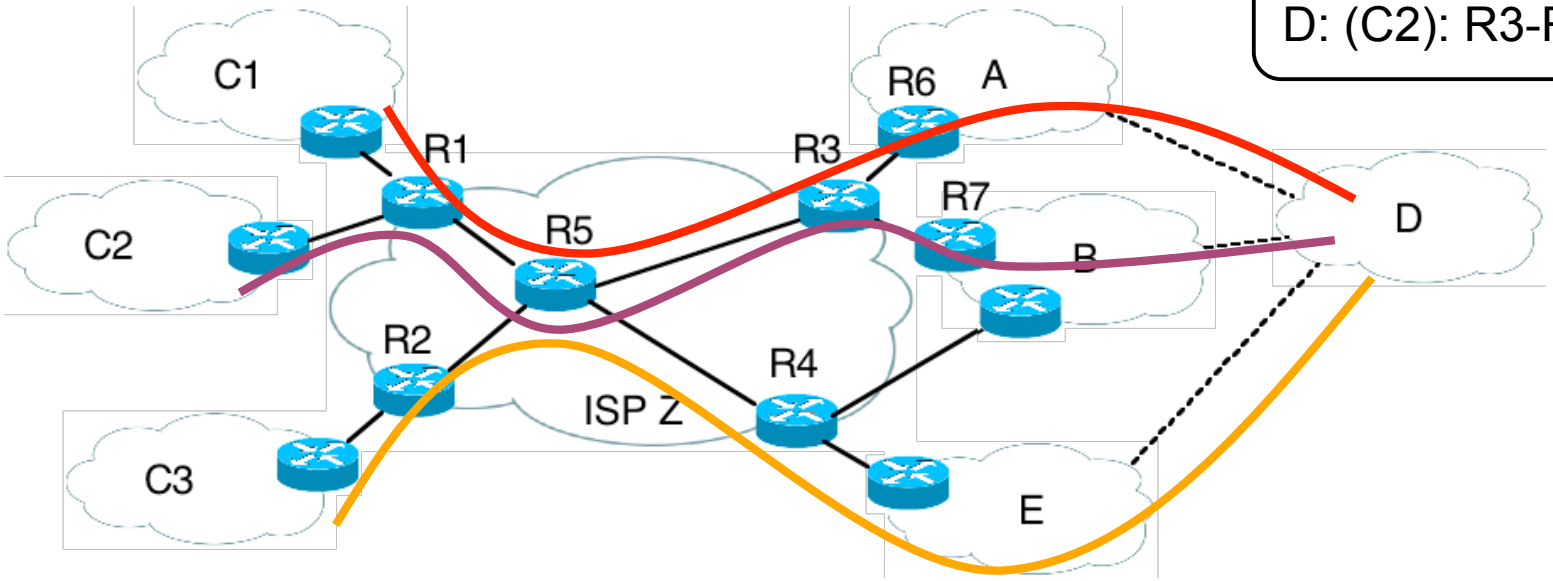


- Backward compatible
  - No changes in neighboring domains
  - No changes to the routers

Intra-AS Routing Architecture

R3's FIB entries

D: (C1): R3-R6
D: (C2): R3-R7



- Support for multipath already available
  - "Virtual routing and forwarding (VFR)" (Cisco)
  - "Virtual router" (Juniper)

Limitations of Current BGP Implementations

Limitation 1: Overloading of BGP attributes

- Policy objectives are forced to "share" BGP attributes

Business Relationships **Local-preference** **Traffic Engineering**

- Difficult to add new policy objectives

Limitations of Current BGP Implementations

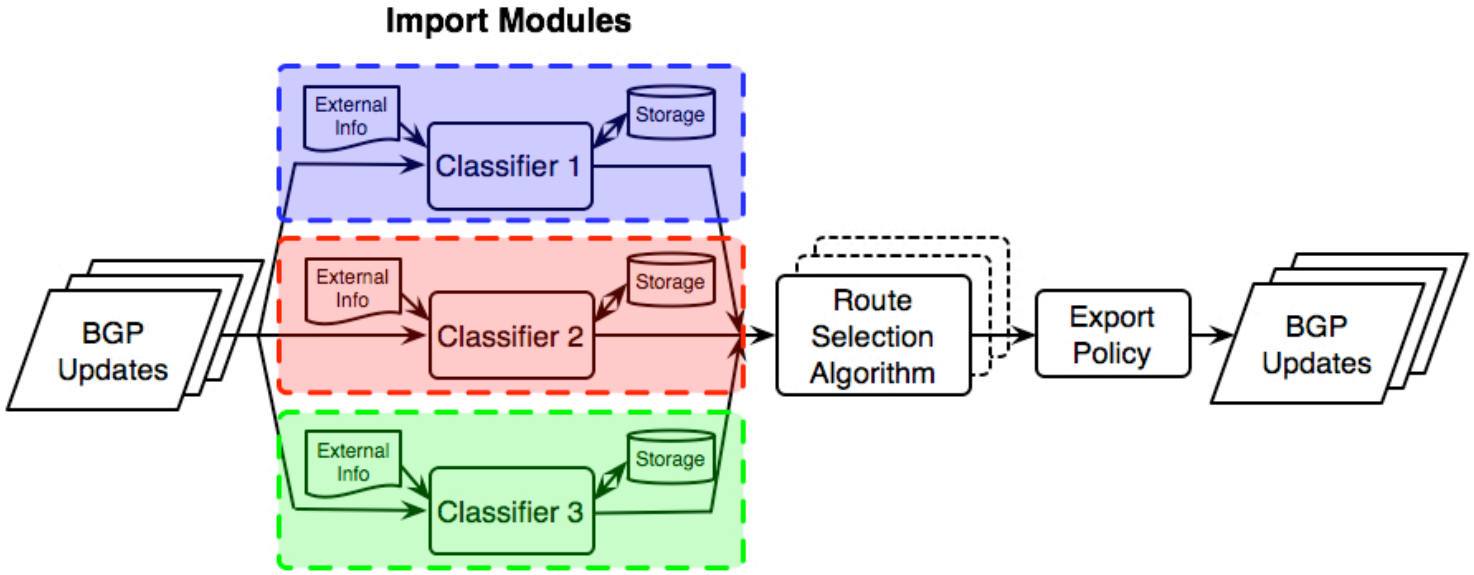
Limitation 2: Difficulty in incorporating "side information"

- Many policy objectives require "side information"

External Information	Internal State Information
Measurement data	History of (prefix, origin) pairs
Business relationships database	Statistics of route instability
Registry of prefix ownership	

- Side information is very difficult to incorporate today

Inside Morpheus Server: Policy Objectives As Independent Modules



- Each module tags routes in separate spaces (solves limitation 1)
- Easy to add side information (solves limitation 2)
- Different modules can be implemented independently (e.g., by third-parties) – evolvability

Limitations of Current BGP Implementations

Limitation 3: Strictly rank one attribute over another (not possible to make trade-offs between policy objectives)

- E.g., a policy with trade-off between business relationships and stability

  "If all paths are somewhat unstable,
  pick the most stable path (of any length)
  Otherwise,
  pick the shortest path through a customer"

- Infeasible today

Use Weighted Sum Instead of Strict Ranking

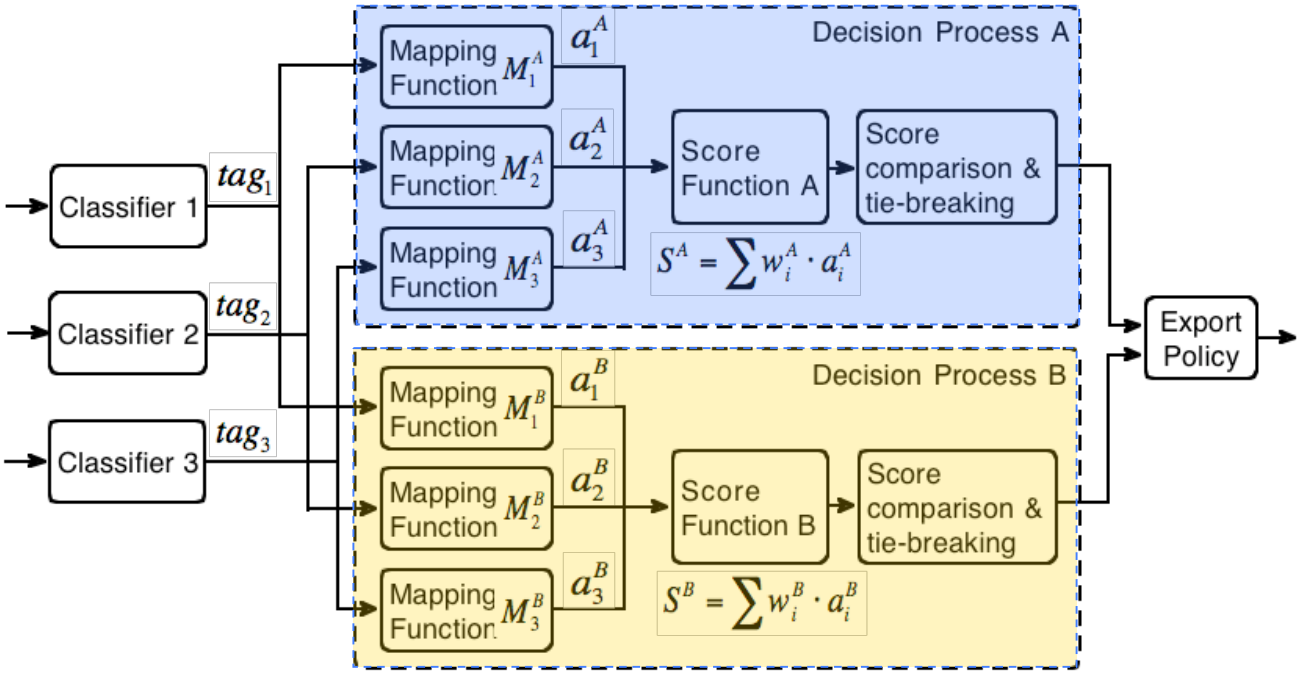
- Every route r gets a value $a_i(r)$ of each criterion (policy objective) c_i (assigned by classifiers)
- Each criterion c_i is assigned a weight w_i
- Every route r has a final score $S(r)$:

$$S(r) = \sum_{c_i \in C} w_i \cdot a_i(r)$$

- The route with highest $S(r)$ is selected as best:

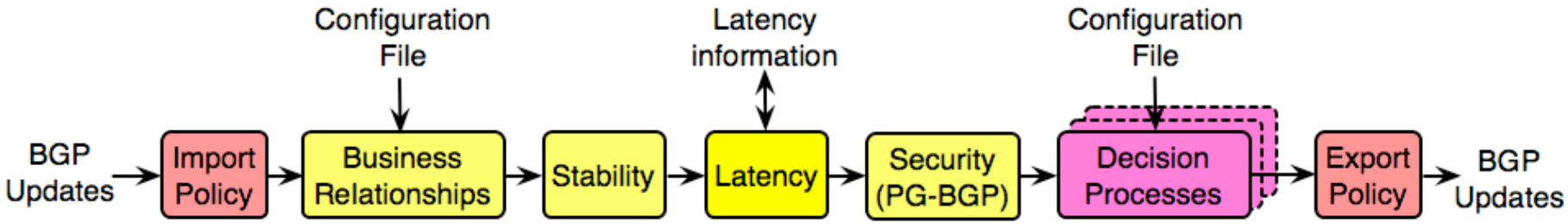
$$r^* = \arg\max_{r \in R} \left(\sum_{c_i \in C} w_{c_i} \cdot a_{c_i} \right)$$

Multiple Decision Processes



- Multiple decision processes running in parallel
- Each with a different set of weights, selecting potentially different best routes

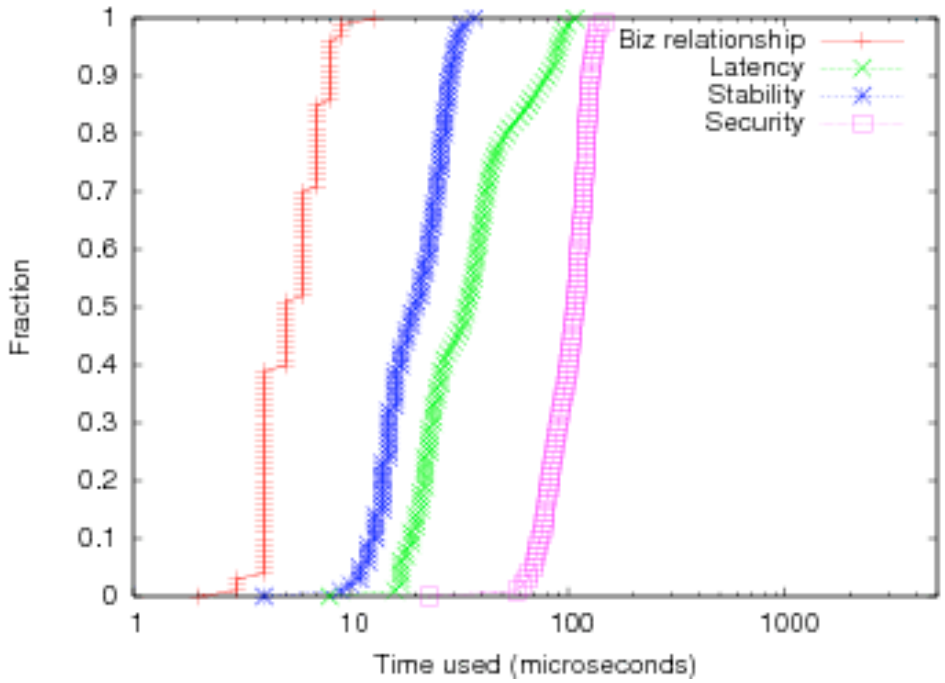
Prototype Implementation



- Implemented as an extension to XORP
- A pipeline of classifier modules

Evaluation - Classification Time

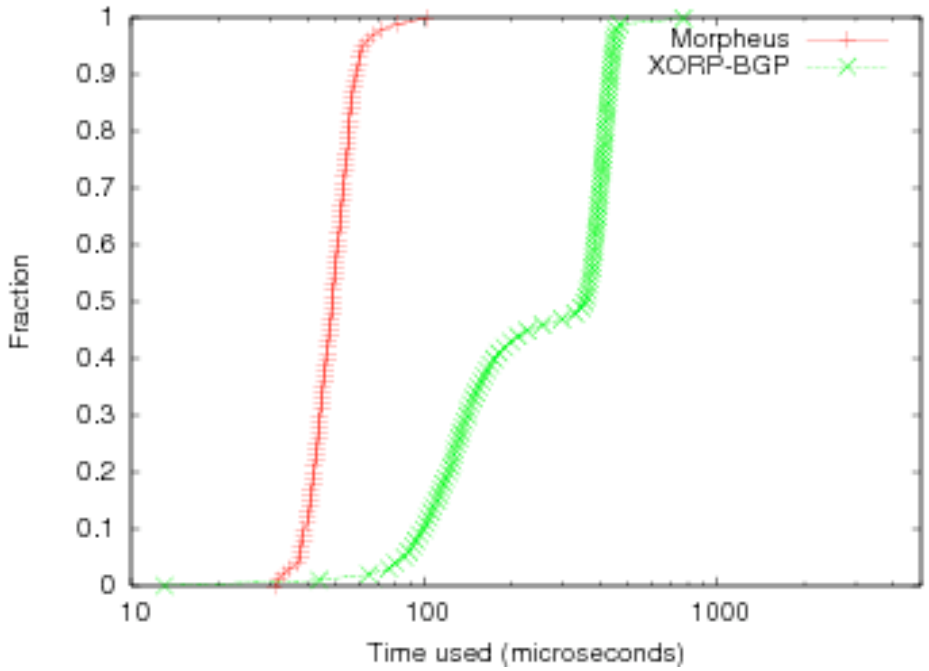
- Classifiers work very efficiently



Average classification time:
- Biz relationship: 5 us
- Stability: 20 us
- Latency: 33 us
- Security: 103 us

Evaluation - Decision Time

- Morpheus is faster than the standard BGP decision process, when there are multiple alternative routes for a prefix



20 routes per prefix

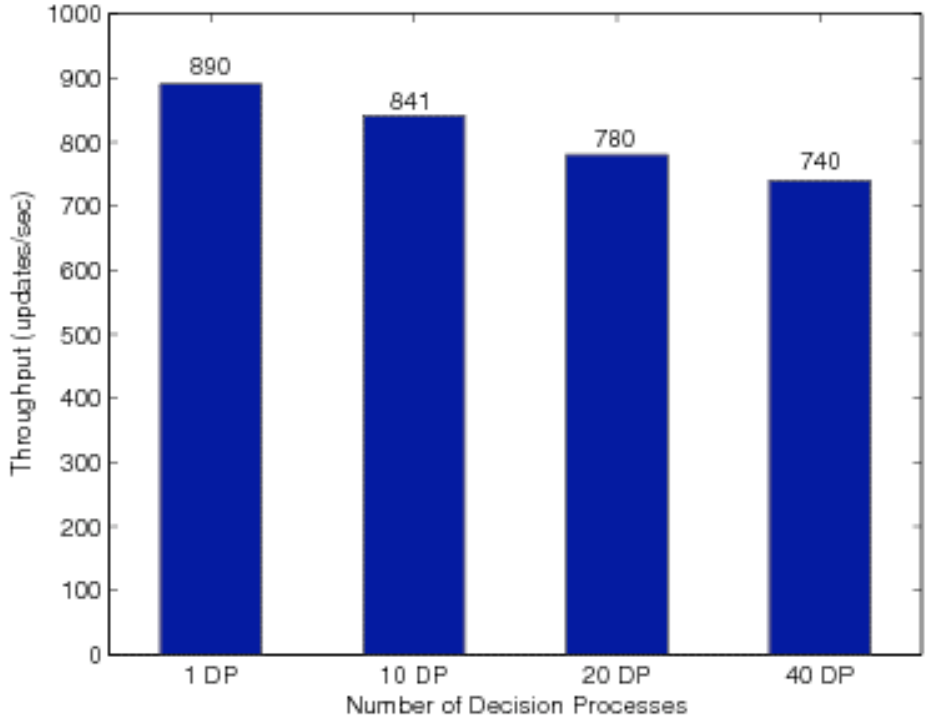
Average decision time:
- Morpheus: 54 us
- XORP-BGP: 279 us

Evaluation - Throughput

- Setup
  - 40 POPs, 1 Morpheus server in each POP
  - Each Morpheus server: 240 eBGP / 15 iBGP sessions, 39 sessions with other servers
  - 20 routes per prefix

Evaluation - Throughput

- Morpheus can efficiently support a large number of decision processes in parallel



No Threat to Stability

- Only announce “non-default” routes to stub customers
- A significant portion of customers are stubs

ASN	701	7018	172	1239	3356
Customers	2634	2053	1667	1651	1425
Stub (%)	84.4%	86.1%	66.9%	78.9%	60.0%

ASN	209	3549	2914	3561	5511
Customers	1233	924	460	449	131
Stub (%)	86.7%	57.8%	48.9%	72.8%	40.5%

Summary

- Morpheus: a simple, practical way for ISPs to capitalize path diversity
- Benefits
  - Significantly more flexible
  - No impact on stability
  - Efficient and scalable enough for large ISPs
  - Backwards compatible

Questions for Operators

- What are your top policy objectives?
- Real examples of customers demanding different routes / more control of the routes they get?
- How much control are you willing to give to your customers?
- Practical concerns?

Very interested in feedback and collaboration

yiwang@cs.princeton.edu

More information:

http://www.cs.princeton.edu/research/techreps/TR-802-07

Backup Slide

How to Setup the Weights?

- Simple configuration interface based on Analytical Hierarchy Process (AHP)
- How does it work?
  - Operators specify preference of each pair of objectives using number 1 (equally prefer) to 9 (extremely prefer one over another)
  - AHP automatically derives the appropriate weights from the preference matrix
  - More information: http://www.cs.princeton.edu/research/techreps/TR-802-07