Morpheus: Enabling Flexible Interdomain Routing Policies

Yi Wang
Ioannis Avramopoulos
Jennifer Rexford
Princeton University

NANOG42
02/17/2008

http://www.cs.princeton.edu/research/techreps/TR-802-07
Large ISPs Have Rich Path Diversity

• Top 2% ASes have 10 or more ASes 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 industry: secure paths
  – VoIP / online gaming providers: low latency 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

• Two components
  – Supports from intra-AS routing architecture
  – Morpheus server 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

• Support for multipath already available
  – “Virtual routing and forwarding (VFR)” (Cisco)
  – “Virtual router” (Juniper)

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

R3’s FIB entries
Limitations of Current BGP Implementations

**Limitation 1:** Overloading of BGP attributes

- Policy objectives are forced to “share” BGP attributes

- 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**
  - Measurement data
  - Business relationships database
  - Registry of prefix ownership

  **Internal State Information**
  - History of (prefix, origin) pairs
  - Statistics of route instability

- 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: 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 traffic engineering
  
  "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

<table>
<thead>
<tr>
<th>ASN</th>
<th>701</th>
<th>7018</th>
<th>172</th>
<th>1239</th>
<th>3356</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customers</td>
<td>2634</td>
<td>2053</td>
<td>1667</td>
<td>1651</td>
<td>1425</td>
</tr>
<tr>
<td>Stub (%)</td>
<td>84.4%</td>
<td>86.1%</td>
<td>66.9%</td>
<td>78.9%</td>
<td>60.0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ASN</th>
<th>209</th>
<th>3549</th>
<th>2914</th>
<th>3561</th>
<th>5511</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customers</td>
<td>1233</td>
<td>924</td>
<td>460</td>
<td>449</td>
<td>131</td>
</tr>
<tr>
<td>Stub (%)</td>
<td>86.7%</td>
<td>57.8%</td>
<td>48.9%</td>
<td>72.8%</td>
<td>40.5%</td>
</tr>
</tbody>
</table>
Summary

• Morpheus: enable flexible path selection to capitalize the path diversity in large ISPs
  – Significantly more flexible
  – No impact on stability
  – Efficient and scalable enough for large ISPs
  – Backwards compatible
Questions for Operators

• How much flexibility is desired? (More flexibility could mean more knobs to tweak…)

• Potential applications of Morpheus? (Concrete examples)

• Practical concerns?

Very interested in feedback and collaboration

yiwang@cs.princeton.edu

More information:

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