Stable Internet Route Selection

Brighten Godfrey
Matthew Caesar
Ian Haken
Scott Shenker
Ion Stoica
UC Berkeley

pb@cs.berkeley.edu

NANOG 40
June 6, 2007
BGP instability: trouble

control plane

CPU cycles
update processing uses majority of cycles on some core routers

data plane

degraded path quality
BGP causes majority of packet loss bursts

Stable Route Selection:
simple technique to significantly improve stability
What about Route Flap Damping?

- Introduces pathologies
- Impacts availability

“...the application of flap damping in ISP networks is NOT recommended.”

--RIPE Route Working Group, May 2006

- Only helps for very unstable routes
Stable Route Selection

Given a choice between routes, select routes that are less likely to fail.

RFD philosophy: shut off bad routes
SRS philosophy: always pick a route if possible, but prefer more stable routes
Challenges

• **Inferring stability** of paths, locally

• **Dependence:** does one ISP’s benefit require others’ participation?

• **Flexibility required**
Outline

• Design

• Evaluation
  Improvement in stability
  Dependence
  Flexibility

• Conclusion
Design

BGP decision process

1. Highest local pref
2. Shortest path length
3. Lowest origin type
4. Lowest MED
5. eBGP- over iBGP-learned
6. Lowest IGP cost
7. Lowest router ID

SRS heuristic

2. Current route
3. Shortest path length
4. Longest uptime
Simplified processes

- Simulator has one router per AS, at most one link between AS’s
- So...

<table>
<thead>
<tr>
<th>Standard BGP</th>
<th>SRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Highest local pref</td>
<td>1. Highest local pref</td>
</tr>
<tr>
<td>2. Shortest path length</td>
<td>2. Current route</td>
</tr>
<tr>
<td>3. Lowest router ID</td>
<td>3. Shortest path length</td>
</tr>
<tr>
<td></td>
<td>4. Longest uptime</td>
</tr>
</tbody>
</table>
Outline

- Design
- **Evaluation**
  - Improvement in stability
  - Dependence
  - Flexibility
- Conclusion
Evaluation methodology

- Event-based BGP simulator

<table>
<thead>
<tr>
<th>Topology</th>
<th>Internet AS-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local prefs</td>
<td>cust./prov./peer</td>
</tr>
<tr>
<td>AS-adjacency failures</td>
<td>inferred from RouteViews</td>
</tr>
</tbody>
</table>

- Measuring **interruptions**: route changes/withdrawals
The bottom line

Mean interruptions per month per src-dst pair

<table>
<thead>
<tr>
<th>Method</th>
<th>Interruptions</th>
<th>Availability loss relative to Std BGP</th>
</tr>
</thead>
<tbody>
<tr>
<td>RFD &amp; SRS</td>
<td>2.8</td>
<td>4%</td>
</tr>
<tr>
<td>SRS</td>
<td>5.3</td>
<td>0%</td>
</tr>
<tr>
<td>RFD</td>
<td>8.9</td>
<td>5%</td>
</tr>
<tr>
<td>Std. BGP</td>
<td>26</td>
<td>0%</td>
</tr>
</tbody>
</table>
Dependence between ISPs

- 44% reduction for first adopters
- 80% reduction at full deployment

Std. BGP
Dependence between ISPs

Tier 1s benefit more than other ISPs when unilaterally running SRS
Route flexibility

Flexibility also needed for:
- load balancing
- business relationships
- path length

What is the tradeoff with these other objectives?
How much flexibility?

- Flexibility for SRS
  - Business LP’s: 5.3
  - Realistic traffic eng., etc.
  - Std BGP: 26

Flexibility for other objectives

Interruptions per month per src/dst pair
Tradeoffs: path length

SRS only
4% longer!

(Hypothetical “Longest paths” strategy: 32% longer)
Outline

• Design

• Evaluation
  Improvement in stability
  Dependence
  Flexibility

• Conclusion
Summary

• **Stable Route Selection**: use flexibility in path selection to optimize for stability
  - Significantly more stable
  - No impact on availability
  - Very low stretch

• Ongoing work: implementation & deployment
Questions for operators

1. How useful is stability?

2. What are the barriers?
   (nondeterminism, traffic engineering...)

3. How much flexibility would be available to SRS?

Very interested in feedback and collaborations

pbg@cs.berkeley.edu
Backup slides
AS adjacency mean session time distribution
Attribution of improvement

\[
\text{instability} = \text{interruptions per event} \times \# \text{ events}
\]

- speed convergence
  - \(\sim 8\% \text{ better}\)
- avoid failures
  - Majority of the improvement
SRS vs. flap damping

SRS: better mean improvement

RFD: better for worst $\sim0.5\%$ of src-dst pairs
SRS vs. flap damping

SRS is more conservative
SRS is more aggressive

always pick a route if one is advertised
use any flexibility available for stability
SRS with less flexibility

But how much flexibility in practice?

Substantial benefit with just 1.2 choices

1.85 possible paths with business class Local Pref

Interruptions per month

Mean number of route options
SRS convergence

- Convergence depends on decision process
- If heuristic is passive
  - Any stable state for Std BGP is still stable
  - Gao-Rexford constraints still sufficient to guarantee convergence to stable state
- (Simulations: *slightly* faster convergence)

SRS Heuristic
2. Current route
3. Lowest path length
4. Longest uptime
SRS can converge where standard BGP doesn’t
Acknowledgements

• This material is based upon work supported by a National Science Foundation Graduate Research Fellowship

• Traffic sign image from Manual of Traffic Signs, by Richard C. Moeur (http://www.trafficsign.us)