PATHLET ROUTING

P. Brighten Godfrey
pgb@illinois.edu

Igor Ganichev, Scott Shenker, and Ion Stoica
{igor,shenker,istoica}@cs.berkeley.edu

SIGCOMM 2009
Internet routing challenges

- Multipath
  - reliability
  - path quality
- Scalability
- Policy
Internet routing challenges

- Multipath
- reliability
- path quality

- Scalability

- Policy

[ F. Wang, Z. M. Mao, J. Wang, L. Gao, R. Bush ’06]
Internet routing challenges

- Multipath
- Reliability
- Path quality
- Scalability
- Policy

- Lowest latency path
- Highest bandwidth path
- Path the network picked for you
Internet routing challenges

Multipath reliability path quality

Scalability

Policy

Internet forwarding table size [Huston ’09]
Internet routing challenges

- Multipath
- reliability
- path quality

- Scalability

- Policy
Internet routing challenges

Multipath
  reliability
  path quality

Scalability

Policy
Internet routing challenges

Multipath
- reliability
- path quality

Scalability

Policy
**Pathlet routing**

- **vnode**: virtual node
- **pathlet**: fragment of a path: a sequence of vnodes

**Source routing** over pathlets.
Pathlet routing

vnode  virtual node

pathlet  fragment of a path:
          a sequence of vnodes

Source routing over pathlets.
**Pathlet routing**

**vnode** virtual node

**pathlet** fragment of a path: a sequence of vnodes

Source routing over pathlets.
Pathlet routing

**vnode**  virtual node

**pathlet** fragment of a path: a sequence of vnodes

Source routing over pathlets.
Pathlet routing

**vnode** virtual node

**pathlet** fragment of a path: a sequence of vnodes

Source routing over pathlets.

virtual graph: flexible way to define policy constraints
vnode virtual node

pathlet fragment of a path: a sequence of vnodes

Source routing over pathlets.

virtual graph: flexible way to define policy constraints

provides many path choices for senders
Flexibility
• can emulate BGP, source routing, MIRO, LISP, NIRA
Flexibility

• can emulate BGP, source routing, MIRO, LISP, NIRA

• local transit policies provide multipath and small forwarding tables
Flexibility

- can emulate BGP, source routing, MIRO, LISP, NIRA

- local transit policies provide multipath and small forwarding tables

- coexistence of different styles of routing policy
Flexibility

- can emulate BGP, source routing, MIRO, LISP, NIRA
- local transit policies provide multipath and small forwarding tables
- coexistence of different styles of routing policy
Flexibility

• can emulate BGP, source routing, MIRO, LISP, NIRA

• local transit policies provide multipath and small forwarding tables

• coexistence of different styles of routing policy
“Design for variation in outcome, so that the outcome can be different in different places, and the tussle takes place within the design, not by distorting or violating it.”

— Clark, Wroclawski, Sollins & Braden, 2002
“Tussle in Cyberspace”
• The protocol
• Uses
• Experimental results
• Comparing routing protocols
Pathlet routing

vnode virtual node

pathlet fragment of a path: a sequence of vnodes

Source routing over pathlets.
**vnode**: virtual node within an AS
 vnode: virtual node
within an AS
**vnodes**

**vnode**: virtual node within an AS

- Walla Walla
- New York
- Crumstown
- San Diego
- Roosterville
**vnode**: virtual node within an AS
**vnode**: virtual node within an AS
vnode: virtual node within an AS
vnode: virtual node within an AS
vnode: virtual node within an AS

designated ingress vnode for each neighbor
**vnodes**

**vnode:** virtual node within an AS

designated **ingress vnode** for each neighbor
**vnodes**

**vnode**: virtual node within an AS

designated **ingress vnode** for each neighbor
vnode: virtual node within an AS

designated ingress vnode for each neighbor

Internally: a forwarding table at one or more routers
Pathlets

Packet route field  Forwarding table
Pathlets

Packet route field  Forwarding table
Pathlets

Packet route field  Forwarding table

A

B 7

C 2

D
Pathlets

Packet route field

Forwarding table

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>7</td>
<td>fwd to C</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>fwd to D</td>
</tr>
</tbody>
</table>
Pathlets

Packet route field

Forwarding table

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>fwd to C</td>
</tr>
<tr>
<td>2</td>
<td>fwd to D</td>
</tr>
</tbody>
</table>
Pathlets

Packet route field

Forwarding table

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7,2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>fwd to C</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>fwd to D</td>
<td></td>
</tr>
</tbody>
</table>
Pathlets

Packet route field

Forwarding table

A

B 7

C 2

D

7,2

2

... ...

7  fwd to C

... ...

2  fwd to D
Pathlets

Packet route field

A
B 7
C 2
D

Forwarding table

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7,2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>fwd to C</td>
</tr>
<tr>
<td>2</td>
<td>fwd to D</td>
</tr>
</tbody>
</table>

delivered!
### Packet route field

- **A**
  - 3

- **B**
  - 7

- **C**
  - 2

- **D**

### Forwarding table

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>fwd to C</td>
</tr>
<tr>
<td>3</td>
<td><strong>push 7,2; fwd to B</strong></td>
</tr>
<tr>
<td>2</td>
<td>fwd to D</td>
</tr>
</tbody>
</table>

Delivered!
Pathlets

Packet route field

Forwarding table

A

B

C

D

3

7,2

2

3

push 7,2; fwd to B

7

fwd to C

2

fwd to D

delivered!
• Global gossip fine, except for scalability

• So, let routers choose not to disseminate some pathlets

• Leads to (ironic) use of path vector — only for pathlet dissemination, not route selection
Outline

• The protocol
• Uses
• Experimental results
• Comparing routing protocols
Local transit policies

Each ingress \( \rightarrow \) egress pair is either allowed or disallowed.

Subject to this, any path allowed!

Represented with few pathlets: small FIB
“All valley-free” is local

“customers can route to anyone; anyone can route to customers”
“All valley-free” is local

“customers can route to anyone; anyone can route to customers”
“All valley-free” is local

“customers can route to anyone; anyone can route to customers”
“All valley-free” is local

“customers can route to anyone; anyone can route to customers”
“All valley-free” is local

“customers can route to anyone; anyone can route to customers”
“All valley-free” is local

“customers can route to anyone; anyone can route to customers”
“All valley-free” is local

“customers can route to anyone; anyone can route to customers”
“All valley-free” is local

“customers can route to anyone; anyone can route to customers”
“All valley-free” is local

“customers can route to anyone; anyone can route to customers”

Forwarding table size: 3 + #neighbors
Emulating BGP

128.2.0.0/16
Emulating BGP

128.2.0.0/16
Emulating BGP

128.2.0.0/16
Emulating BGP

128.2.0.0/16
Emulating BGP
Mixed policies

local  BGP-like  local  local  local  local
Outline

• The protocol
• Uses
• Experimental results
• Comparing routing protocols
Improved connectivity

BGP-style

Mixed

LT policies
Tiny forwarding tables

Forwarding table size CDF
Tiny forwarding tables

Forwarding table size CDF

- BGP: 132,158+ entries: one per IP prefix
- pathlet routing, valley-free LT policies
  - 2,264 entries, max
  - 8.48 entries, mean

current Internet (CAIDA/APNIC):
Control overhead

2.23x more messages, 1.61x more memory in LT than PV

This can likely be improved.
Outline

- The protocol
- Uses
- Experimental results
- Comparing routing protocols
Comparing protocols
Comparing protocols

- Feedback-based routing
- Strict source routing
- Loose source routing
- MIRO
- NIRA
- Pathlet routing
- LISP
- Routing deflections, path splicing
- BGP
Comparing protocols

Feedback-based routing

Strict source routing

Loose source routing

Pathlet routing

MIRO

LISP

BGP

NIRA

Routing deflections, path splicing
Conclusion

• Pathlet routing: source routing over a virtual topology formed by pathlets and vnodes

• Highly flexible; supports both “local” policies with small forwarding tables and many paths, and complex BGP policies

• Challenges for source routing: Incentives to provide multiple paths; selecting paths; security; ...