Decentralized cloud WAN traffic engineering with BlastShield

Umesh Krishnaswamy  Rachee Singh  Nikolaj Bjorner  Himanshu Raj

Microsoft
Software-driven WAN

SIGCOMM ’13
SWAN traffic

Traffic growth

- 2018 to 2022
- Y-axis: 0 to 10x
- X-axis: Period
SWAN outage of global scope

Bug 1 in controller removed all routes

Bug 2 in fallback routing amplified congestion

Bug 3 in controller increased time to recover
Blast radius

Blast radius = customer traffic at risk from a controller failure

<table>
<thead>
<tr>
<th>Traffic Type</th>
<th>Tier</th>
<th>Service level objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer traffic</td>
<td>Tier-0</td>
<td>99.999</td>
</tr>
<tr>
<td>Discretionary traffic</td>
<td>Tier-1</td>
<td>99.9</td>
</tr>
<tr>
<td></td>
<td>Tier-2</td>
<td>99</td>
</tr>
</tbody>
</table>
BlastShield slices
Design assumptions

<table>
<thead>
<tr>
<th>Decentralized</th>
<th>Hierarchical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global view</td>
<td>Local view</td>
</tr>
<tr>
<td>No coordination</td>
<td>With coordination</td>
</tr>
<tr>
<td>Inter-slice traffic &lt; Intra-slice traffic</td>
<td></td>
</tr>
</tbody>
</table>
BlastShield controller

- Topology service
- Demand predictor
- Traffic engineering scheduler
- Route programmer

feeds

global topology

global demands

slice TE routes

TE routes

global
slice-specific
Inter-slice routing

TE path segments programmed by slice controllers

$s \rightleftharpoons a \rightleftharpoons d \rightleftharpoons t (w_1)$
$s \rightleftharpoons c \rightleftharpoons b \rightleftharpoons d \rightleftharpoons t (w_2)$
$s \rightleftharpoons c \rightleftharpoons e \rightleftharpoons t (w_3)$
Blast ripple and routing loops

Traffic uses fallback routes to destination when downstream controller fails.

Enter-leave constraints restrict paths to achieve loop-free routing.
Source routing

TE paths programmed by source slice controller
Traffic engineering scheduler

Path computer
- Max flow
- Penalizing

Constraints

Global topology

Global demands

TE solver
- Priority
- Fairness
- tier-0
  - Max-min + min cost + diverse
- tier-1/2
  - Max-min + min util

Route generator
- slice configuration
- slice TE routes
Symphony or cacophony

TE inefficiency 3%
Blast radius reduction

Blast radius reduction 6%
Blast radius reduction

Safe deployment lowers failure probability

Applications have option to fail out
Summary

BlastShield slices

- Decentralized vs. Hierarchical
- Global view vs. Local view
- No coordination vs. With coordination
- Inter-slice traffic < Intra-slice traffic

Blast radius reduction 6%

TE inefficiency 3%