The Impact of Control Path Survivability on Data Plane Survivability in SDN

Sedef Savas
Networks Lab, Group Meeting
Aug 11, 2017
a) 9-node network topology, controller is on node 4.

b) Shortest path: controller-switch paths are shown. E.g., switch-controller path for 8 is 8-5-4.

c) Link between node 4 and 5 is down.

d) Physical effect on the topology is just 1 link. All data traffic passes link 4-5 needs restoration.

e) Controller lost its communication to switches 5-7-8-9, hence cannot use that portion of network to route affected data traffic. Controller needs

f) Min-max algorithm for switch-controller path: Minimize the maximum number of control paths passing a link. When link 4-5 down, only switches
When control paths of nodes 5, 7, 8, 9 is disconnected after a failure (stage 1), we need several stages to fully connect those nodes again.

In stage 1, several data and control paths are disrupted. Controller recovers whatever it can with remaining network, some unrecovered data paths remain.

In stage 2, node 5 and 7 will find a new control path. In stage 3, node 8 and 9 can be connected.

As controller cannot reach 7 and 8 on stage 2, it has no knowledge of when they will be up again, hence it will try to recover remaining disrupted data paths and at each stage some of the disconnected data paths will be tried to recovered.

This will cause increase in data path lengths.

We will show how much traffic is recovered at each stage and the recovery delay. Also, show how much data path length increase is compared to shortest path on the physical network after the failure.
Related Work

On Resilience of Split-Architecture Networks, Ying Zhang, Neda Beheshti, Mallik Tatipamula from Ericsson, Globecom 11.

• They propose a resilient controller placement and switch assignment. They define a metric called expected control path loss (switch-controller paths are shortest paths). Considering this metric they optimize the locations. They show how multiple number of link failures affect control-switch paths. They focus on placement instead of routing.
Related Work (cont.)

Fast Failover for Control Traffic in Software-defined Networks, Ying Zhang, Neda Beheshti from Ericsson, Globecom 12.

- They focus on control path routing as we do. A binary tree routing is used for control paths. In figure below, if link between s4 and s1 goes down, s4 locally change the port it forward the control traffic and use s4-s2.

- But as far as I know current OpenFlow switches cannot do that, if the control path goes down, controller reaches out to the switch.

![Diagram of network with controller and switches]

*Fig. 1. Protection against link and node failures*
Controller Placement Strategies for a Resilient SDN Control Plane, Petra Vizarreta, Carmen Mas Machuca and Wolfgang Kellerer TUM, RNDM

• They focus on both placement and control path routing as we do. This one is the only one that uses backup paths.

• Their first approach considers that switches have to be connected to a controller over two Disjoint Paths.

• Their second approach considers that switches have to be connected to two Different Controller Replicas over two disjoint paths. Both approaches are finding working and protection control paths of minimum length to enable fast and efficient failover.

• They perform resilient controller placement also: The goal is to find controller placement that provides working and backup control paths of the shortest length.

• The metrics used to evaluate and compare the controller placement strategies are average control path length, expected control path loss, average connection availability. They compared unprotected control path with having 2 disjoint path to a) same controller and b) different controllers.

• They don’t mention resource consumption of control paths. They consider single link/node failures and double link failures. The below results are the only results they show other than showing average path length of the compared schemes under different # of controllers.
• Our backup path approach does not propose anything novel on top of this.

• But, for unprotected ones, we consider all control paths, try to minimize the number of control paths that are disrupted per failure by doing load balancing. So our argument is that even failure detection takes time to switch to a backup path, so routing primary control path good is important?
Novelty

• The effect of control path routing on recovery cost of the data plane (both in terms of bw and latency).
• We do optimization considering all control path routings to minimize less number of control path disruptions.
• Both data paths and other control paths are considered while routing a control path.
Abilene network (11 node) Total data path number = 110
Controller is on node 7

Disaster radius = 100km
All possible disasters that affect more than 1 link:
22 different disasters

Most disruptive disaster among these 22 disasters where $r = 100$km:

- **Failed Links** are 1-10 10-1 7-10 10-7 9-10 10-9
- **Failed Nodes** are 10
- **This disaster cause the following nodes’ control path loss:**
  - 0,1,2,9,10
Instant Recovery Blockers

- **Unrecoverable**: Right after the disaster, with the remaining physical network that is still connected to the controller, which disrupted data paths cannot find a route?

- **Physical Failure**: among the unrecoverable paths, which ones are unrecoverable because of the physical failure (either lost source or destination.)

  - **Source/Dest control path failure**: among the unrecoverable paths, which ones are unrecoverable because of the control paths loss in their source or destination.

- **Disconnected network**: some data paths are unrecoverable because there is no controlled physical path exists, which means disconnected network.

<table>
<thead>
<tr>
<th>Initially failed data paths</th>
<th>Unrecoverable</th>
<th>Physical Failure</th>
<th>Source/Dest control path failure</th>
<th>Disconnected network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abilene under 100 km disaster</td>
<td>54</td>
<td>54</td>
<td>20</td>
<td>34</td>
</tr>
</tbody>
</table>
Most disruptive link failure in Abilene is link 7-10. This link failure cause the following nodes’ control path loss:
• 0,1,2,9,10

<table>
<thead>
<tr>
<th></th>
<th>Initially failed data paths</th>
<th>Unrecoverable</th>
<th>Physical Failure</th>
<th>Source/Dest control path failure</th>
<th>Disconnected network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abilene under 100 km disaster</td>
<td>54</td>
<td>54</td>
<td>20</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>Abilene under link failure</td>
<td>41</td>
<td>41</td>
<td>0</td>
<td>41</td>
<td>0</td>
</tr>
</tbody>
</table>

Most disruptive node failure is same with the most disruptive disaster, node 10.

Even if connectivity is low as in Abilene networks, as long as S/D control path is up, new middle nodes can be found.
Abilene network (11 node) Total data path number = 110
Controller is on node 7

Node 10 goes down. 54 data paths are disrupted.

Node 1-0-2-9 lost control paths.
Data Path Recovery Delay

• # of data paths that can be recovered at each stage.
• Say control paths are 3 hop on average: if the hops are 1000km: propagation delay is 15ms.

• Delay to set up a control path = (15 ms propagation delay) + (10 ms flow installation delay) + (10 ms processing for route calculation in the controller) = 35ms per stage.

• For instance, to recover the 14 paths in stage 4, we need to wait \(35 \times 4 = 140\) ms for control-path-recovery. On top of this data-path-recovery delay will be added.
All Possible Single Link Failures: Recovered data paths at each stage

<table>
<thead>
<tr>
<th>Link:0-1</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Stage 4</th>
<th>Effected Paths</th>
<th>Uncontrolled Nodes</th>
<th>Uncontrolled Links</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>12</td>
<td></td>
<td></td>
<td>14</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Link:0-2</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Link:1-10</td>
<td>0</td>
<td>10</td>
<td>16</td>
<td></td>
<td>26</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Link:2-9</td>
<td>6</td>
<td>16</td>
<td></td>
<td></td>
<td>22</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Link:3-4</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Link:3-6</td>
<td>0</td>
<td>14</td>
<td></td>
<td></td>
<td>14</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Link:4-5</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Link:4-6</td>
<td>2</td>
<td>10</td>
<td></td>
<td></td>
<td>12</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Link:5-8</td>
<td>8</td>
<td>14</td>
<td></td>
<td></td>
<td>22</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Link:6-7</td>
<td>0</td>
<td>8</td>
<td>26</td>
<td></td>
<td>34</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>Link:7-8</td>
<td>0</td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Link:7-10</td>
<td>0</td>
<td>6</td>
<td>16</td>
<td>17</td>
<td>39</td>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>Link:8-9</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td>21</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Link:9-10</td>
<td>5</td>
<td>18</td>
<td></td>
<td></td>
<td>23</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>
Disaster radius = 100km
All possible disasters that affect more than 1 link: 312 different disasters

Most disruptive disaster among these 312 disasters where r = 100km:

- **Failed Nodes** are 4
- **Failed Links** are 0-4 4-0 2-4 4-2 3-4 4-3 4-5 5-4 4-6 6-4 4-8 8-4 4-16 16-4 4-17 17-4 4-29 29-4 4-31 31-4

This disaster cause the following nodes’ control path loss:

- 2, 35, 36, 38, 3, 10, 19, 5, 23, 26, 21, 22, 28, 31, 4, 17, 29 35

GEANT (40 node): Total data path number = 1560
Controller is on node 7
Most disruptive disaster’s effect on data paths

<table>
<thead>
<tr>
<th></th>
<th>Initially failed data paths</th>
<th>Unrecoverable</th>
<th>Physical Failure</th>
<th>Source/Dest control path failure</th>
<th>Disconnected network</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEANT under 100 km disaster</td>
<td>816</td>
<td>678</td>
<td>78</td>
<td>556</td>
<td>64</td>
</tr>
</tbody>
</table>
Most disruptive link failure in GEANT is link 2-4. This link failure cause the following nodes’ control path loss:
• 2, 35, 36, 38
• Here, node 37 is disconnected from the rest of the network, although it did not lose its control path, because it was only connected to node 36.

<table>
<thead>
<tr>
<th></th>
<th>Initially failed data paths</th>
<th>Unrecoverable Failure</th>
<th>Physical Failure</th>
<th>Source/Dest control path failure</th>
<th>Disconnected network</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEANT under 100 km disaster</td>
<td>816</td>
<td>678</td>
<td>78</td>
<td>556</td>
<td>64</td>
</tr>
<tr>
<td>GEANT under link failure</td>
<td>287</td>
<td>249</td>
<td>0</td>
<td>198</td>
<td>51</td>
</tr>
</tbody>
</table>

Most disruptive node failure is same with the most disruptive disaster, node 4.
Neighbors of node 0 1 2 4 30 34
Neighbors of node 1 0 33
Neighbors of node 2 0 4 31 32 35 36 38
Neighbors of node 3 4 5 10 19 30
Neighbors of node 4 0 2 3 5 6 8 16 17 29 31
Neighbors of node 5 3 4 23
Neighbors of node 6 4 7
Neighbors of node 7 6 8 25 34
Neighbors of node 8 4 7 9 25
Neighbors of node 9 8 15 18 25 29
Neighbors of node 10 3
Neighbors of node 11 13
Neighbors of node 12 13 14 15 20 22
Neighbors of node 13 11 12 14 22
Neighbors of node 14 12 13
Neighbors of node 15 9 12 29
Neighbors of node 16 4 34
Neighbors of node 17 4 30
Neighbors of node 18 9
Neighbors of node 19 3
Neighbors of node 20 12
Neighbors of node 21 27
Neighbors of node 22 12 13 23 26 27
Neighbors of node 23 5 22 29
Neighbors of node 24 25 34
Neighbors of node 25 7 8 9 24
Neighbors of node 26 22
Neighbors of node 27 21 22 28
Neighbors of node 28 27 29
Neighbors of node 29 4 9 15 23 28
Neighbors of node 30 0 3 17 39
Neighbors of node 31 2 4
Neighbors of node 32 2 34
Neighbors of node 33 1 34
Neighbors of node 34 0 7 16 24 32 33
Neighbors of node 35 2 36
Neighbors of node 36 2 35 37
Neighbors of node 37 36
Neighbors of node 38 2 39
Neighbors of node 39 30 38
Control-Path-Routing Methods

- Shortest path
- Minimize maximum control paths passing through each link (good for single link failures)
- Minimize maximum number of control paths passing through each node (good for single node failures, as when a node is down, less control paths will be affected.)
- Data-path aware routing
- Link/node/disaster radius disjoint backup path
• Controller Disjointness
• Data Path Disjointness
Control Path Resource Consumption

C: 0 D: 0 is shortest path algorithm.
C shows controller disjointness importance
D shows data path disjointness importance
Single link failure, mostly cause 0 or 1 control path loss, recoveries are completed at first 2 stages.

The worst case control path loss is extremely high. Those failures tend to be the ones within first two hops away of the controller.

Shortest path method uses less resources, but in the worst case cause more control loss.
Controller is on node 4 (has highest node degree)

Averaged over all link failures.
Controller is on node 4 (has highest node degree)

Controller is on node 0 (has average node degree)

Average Recovered Paths Per Stage

Averaged over all link failures.