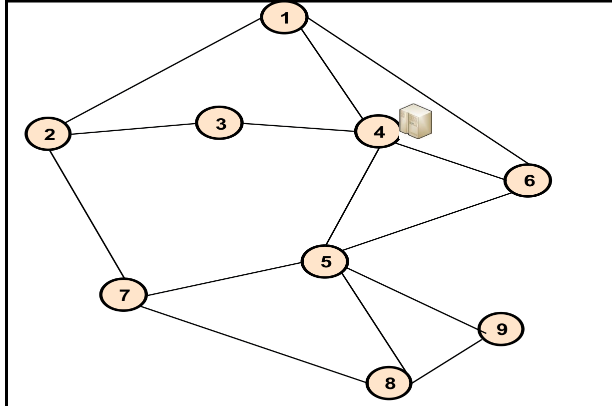
The Impact of Control Path Survivability on Data Plane Survivability in SDN

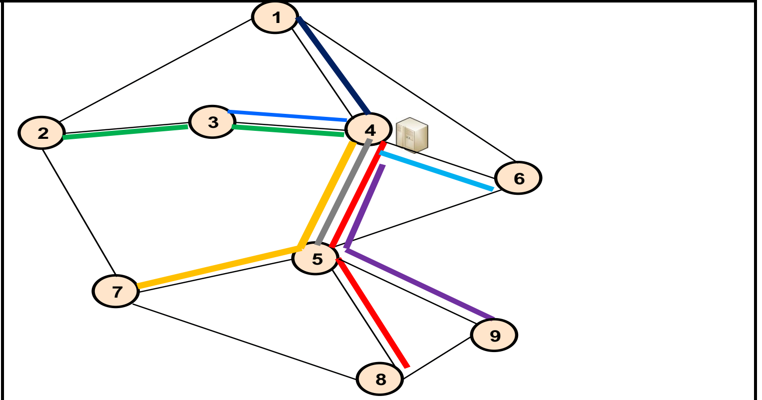
Sedef Savas

Networks Lab, Group Meeting

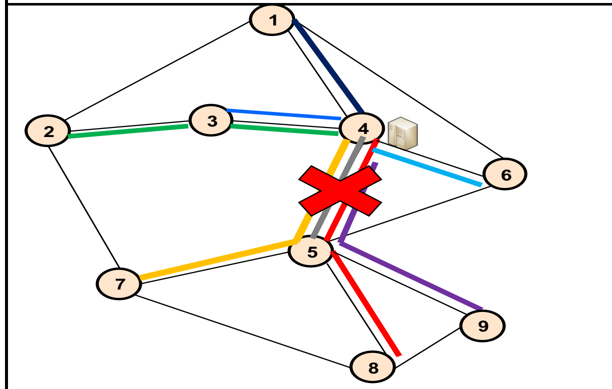
Aug 11, 2017



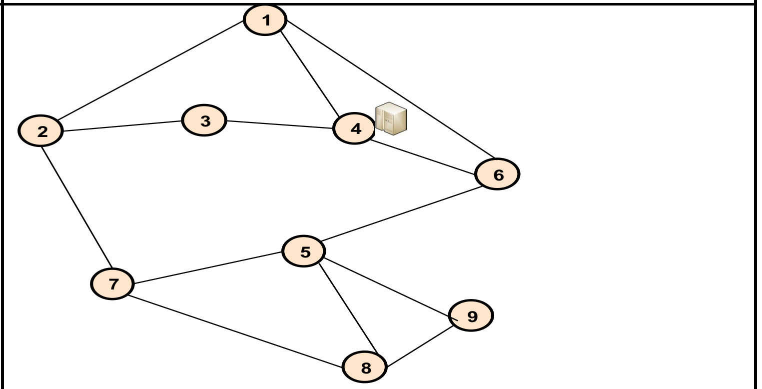
a) 9-node network topology, controller is on node 4.



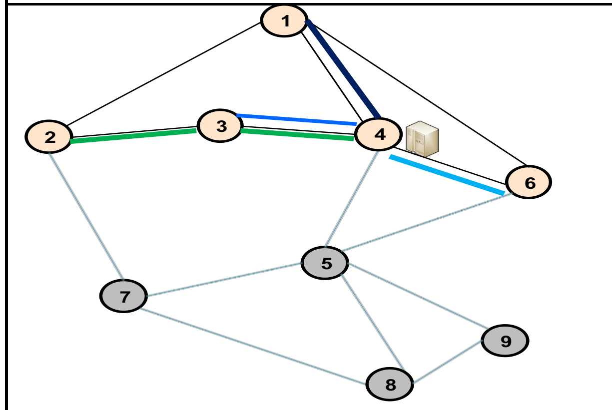
b) Shortest path: controller-switch paths are shown. E.g., switch-controller path for 8 is 8-5-4.



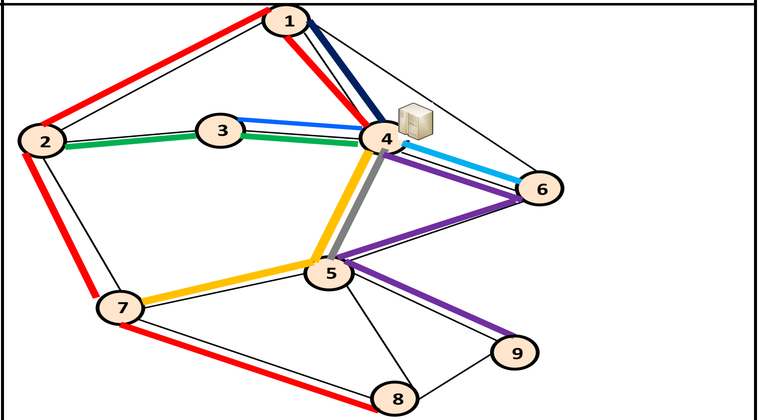
c) Link between node 4 and 5 is down.



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 effected 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

Controller needs to restore communication with node 7 or 5 before being able restore 8 and 9, which requires somewhat sequential recovery in the control plane which is slow.

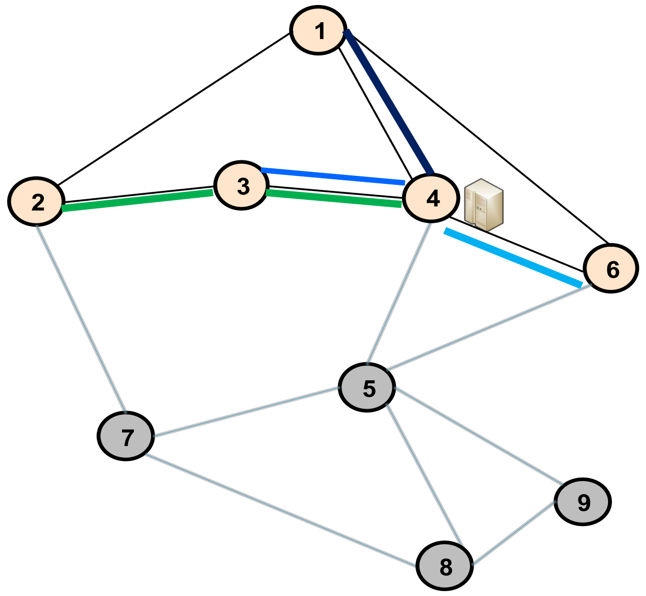
When control paths of nodes 5, 7, 8, 9 is disconnected after a failure (stage 1), we need several stages to fully connect those nodes again.
In stage 1, several data and control paths are disrupted.
Controller recovers whatever it can with remaining network, some unrecovered data paths remain.

In stage 2, node 5 and 7 will find a new control path.
In stage3, 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.

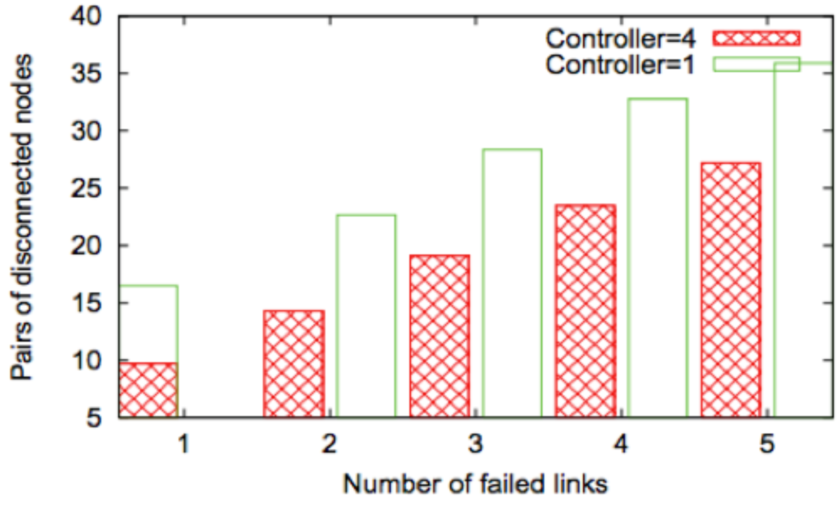
We will show how much traffic is recovered at each stage and the recovery delay.
Also, show how much data path length increase is compared to shortest path on the physical network after the failure.



Related Work

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

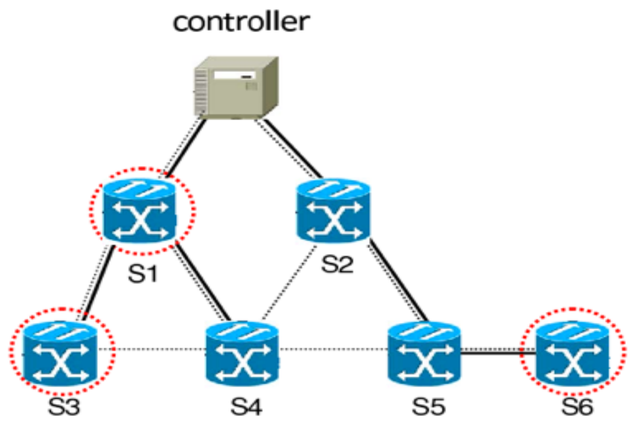
- They propose a resilient controller placement and switch assignment. They define a metric called expected control path loss (switch-controller paths are shortest paths). Considering this metric they optimize the locations. They show how multiple number of link failures affect control-switch paths.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



Related Work (cont.)

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



ig. 1. Protection against link and node failures

Related Work (cont.)

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.

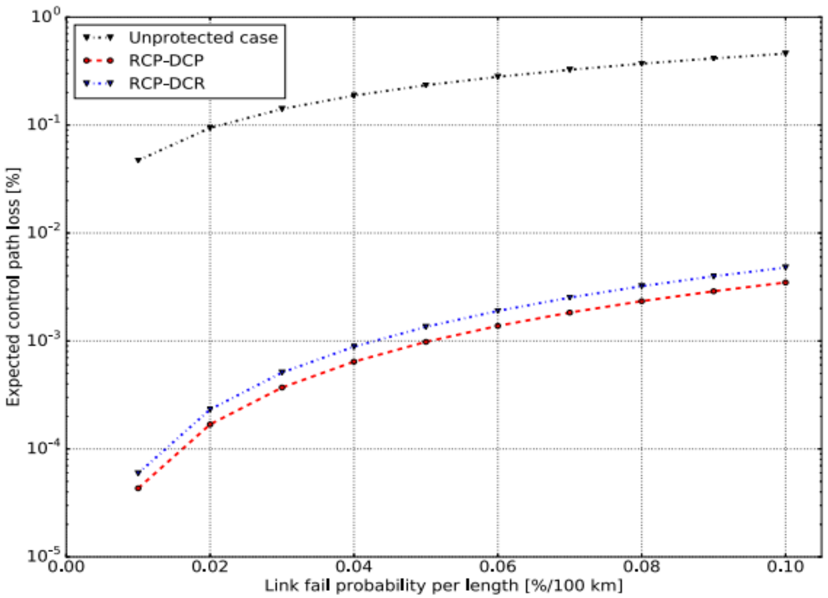


Fig. 5: Expected control path loss when single and double link failures are considered.

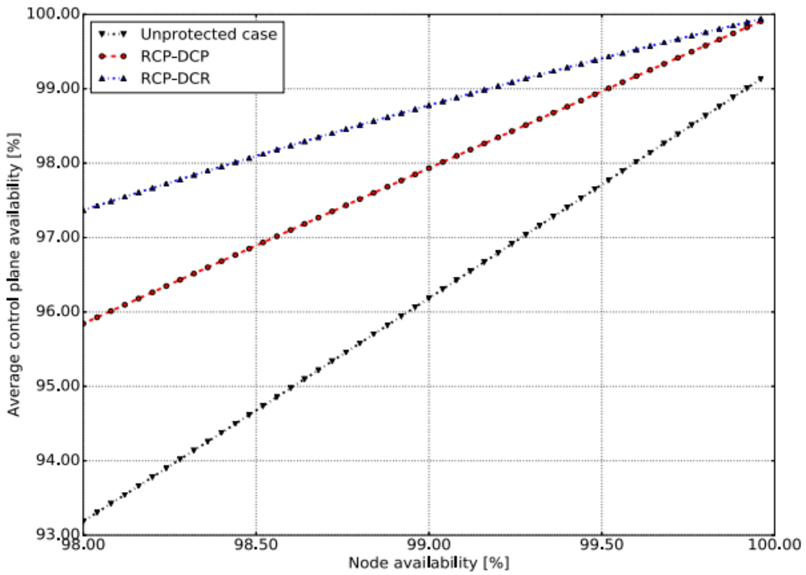


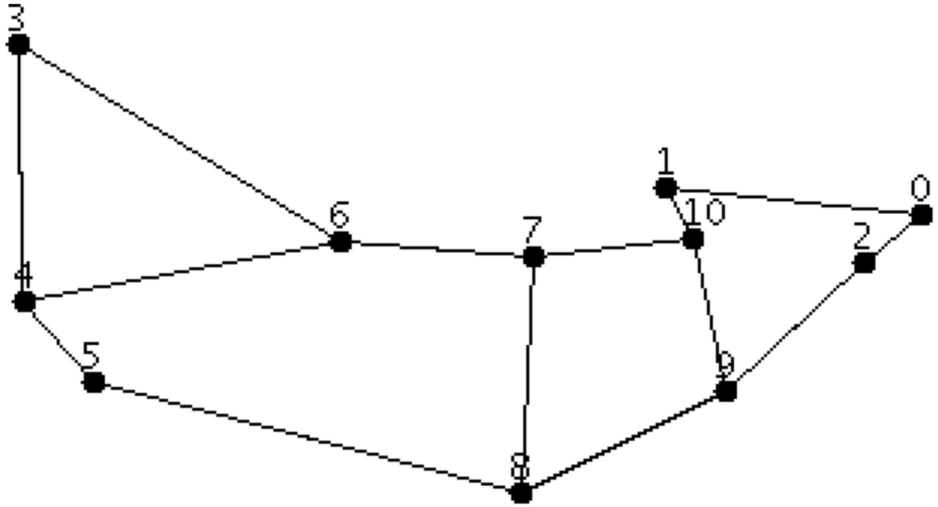
Fig. 6: Average control path availability as a function of node availability. Link failure probability was 0.1%/100 km.

- 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?

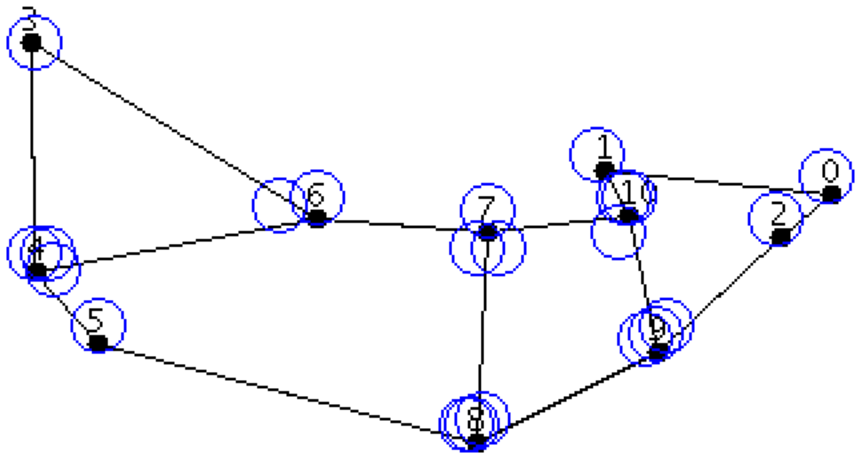
Novelty

- 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 = 100km:

- **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.

	Initially failed data paths	Unrecoverable	Physical Failure	Source/Dest control path failure	Disconnected network
Abilene under 100 km disaster	54	54	20	34	0

Most disruptive link failure in Abilene is link 7-10.

This link failure cause the following nodes' control path loss:

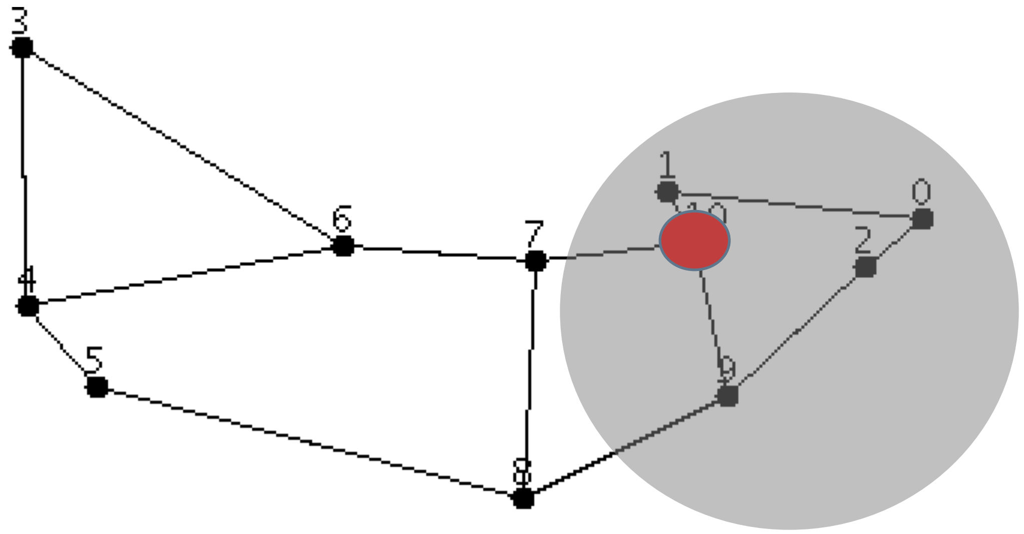
- 0,1,2,9,10

	Initially failed data paths	Unrecoverable	Physical Failure	Source/Dest control path failure	Disconnected network
Abilene under 100 km disaster	54	54	20	34	0
Abilene under link failure	41	41	0	41	0

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

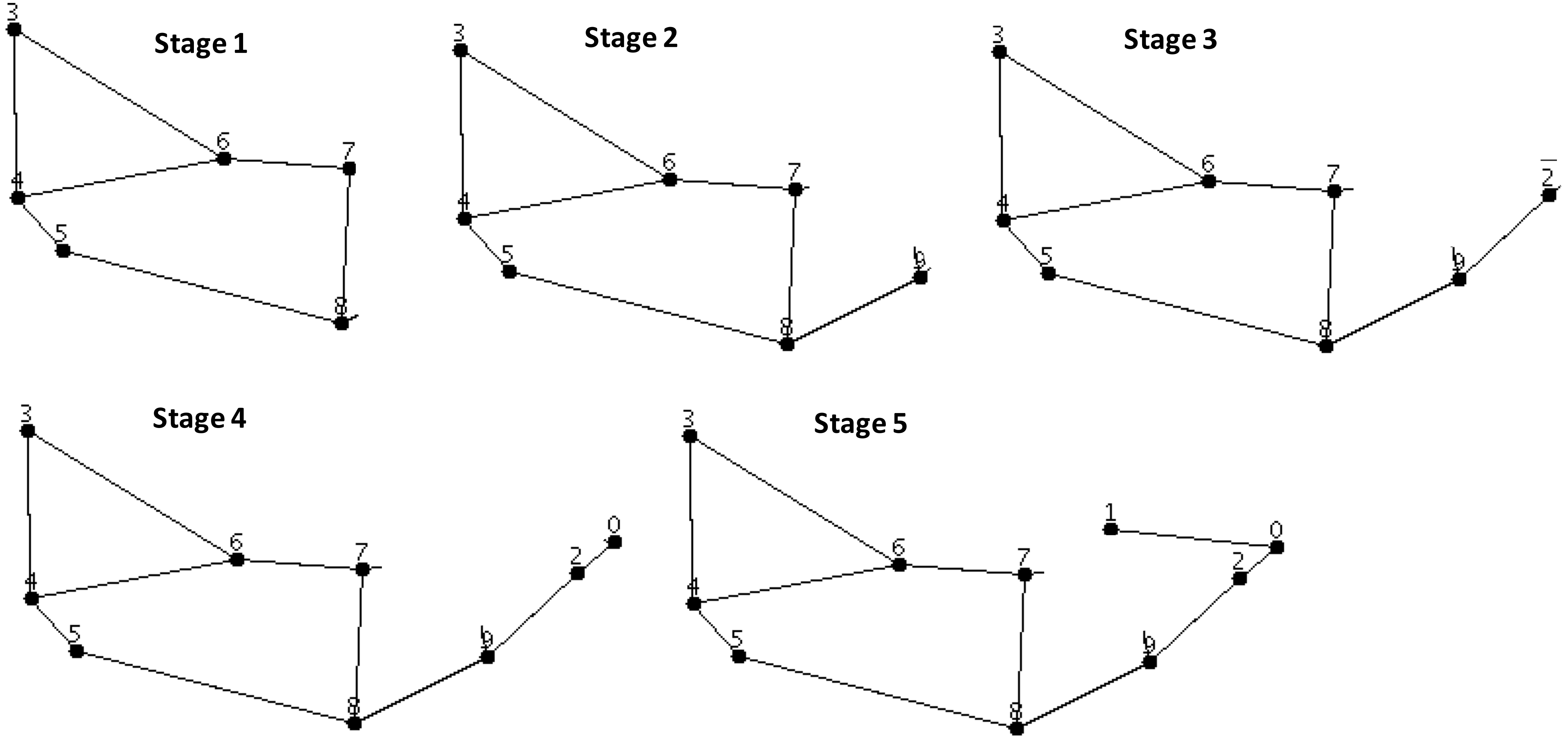


Control Path Routing

0 - 1 - 10 – 7
1 - 10 – 7
2 - 9 - 10 – 7
3 - 6 – 7
4 - 6 – 7
5 - 8 – 7
6 – 7
8 – 7
9 - 10 – 7
10 - 7

Node 10 goes down. 54 data paths are disrupted.

Node 1-0-2-9 lost control paths.



	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Restored Control Paths		Node 9	Node 2	Node 0	Node 1
Recovered Path Number	0	6	6	8	14

Data Path Recovery Delay

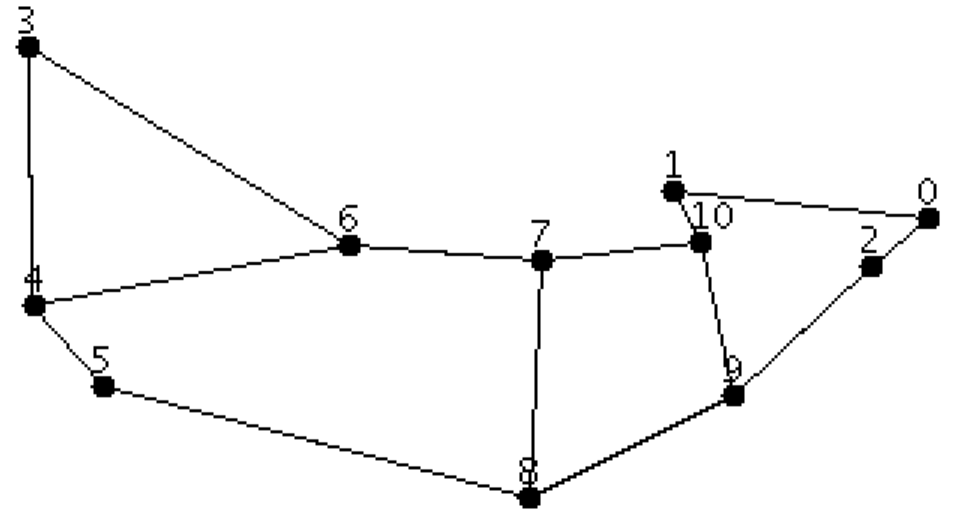
- # of data paths that can be recovered at each stage.
- Say control paths are 3 hop on average: if the hops are 1000km: propagation delay is 15ms.

- 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 * 4 = 140 ms for control-path-recovery. On top of this data-path-recovery delay will be added.

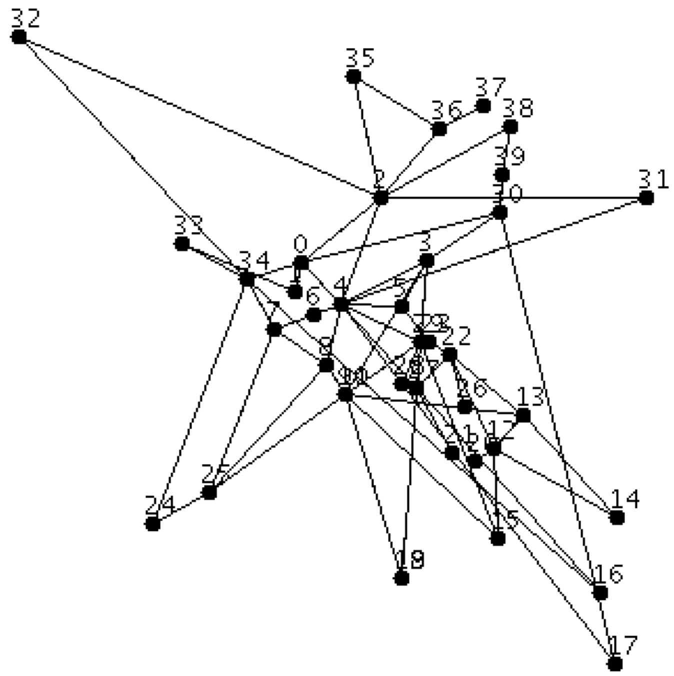
All Possible Single Link Failures: Recovered data paths at each stage

	Stage 1	Stage 2	Stage 3	Stage 4	Effected Paths	Uncontrolled Nodes	Uncontrolled Links
Link:0-1	2	12			14	1	2
Link:0-2	10				10	0	0
Link:1-10	0	10	16		26	2	6
Link:2-9	6	16			22	1	2
Link:3-4	6				6	0	0
Link:3-6	0	14			14	1	2
Link:4-5	14				14	0	0
Link:4-6	2	10			12	1	4
Link:5-8	8	14			22	1	2
Link:6-7	0	8	26		34	3	14
Link:7-8	0	9			9	2	8
Link:7-10	0	6	16	17	39	5	22
Link:8-9	21				21	0	0
Link:9-10	5	18			23	2	8



GEANT (40 node): Total data path number = 1560

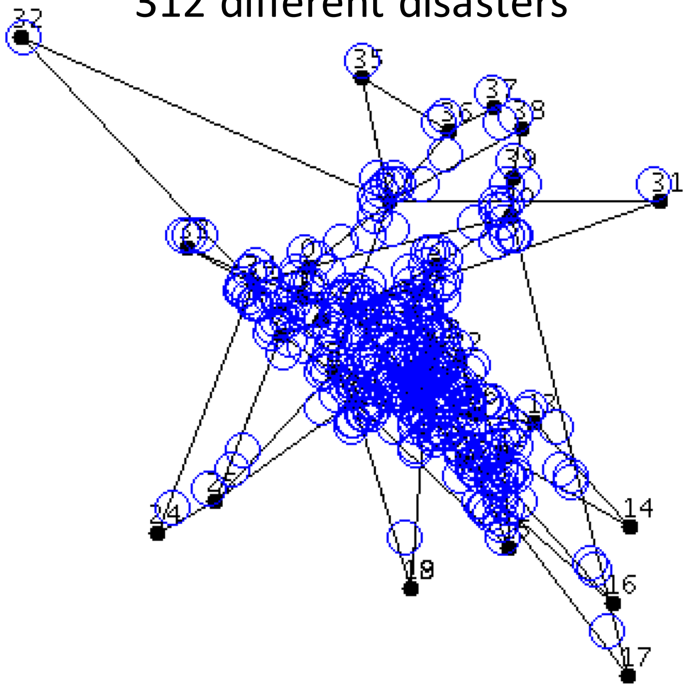
Controller is on node 7



Disaster radius = 100km

All possible disasters that affect more than 1 link:

312 different disasters



<u>Most disruptive disaster among these 312 disasters where r = 100km:</u>

- **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

<u>This disaster cause the following nodes' control path loss:</u>

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

Most disruptive disaster's effect on data paths

	Initially failed data paths	Unrecoverable	Physical Failure	Source/Dest control path failure	Disconnected network
GEANT under 100 km disaster	816	678	78	556	64

Only stage 1 is considered.

Most disruptive link failure in GEANT is link 2-4.

<u>This link failure cause the following nodes' control path loss:</u>

- 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.

	Initially failed data paths	Unrecoverable	Physical Failure	Source/Dest control path failure	Disconnected network
GEANT under 100 km disaster	816	678	78	556	64
GEANT under link failure	287	249	0	198	51

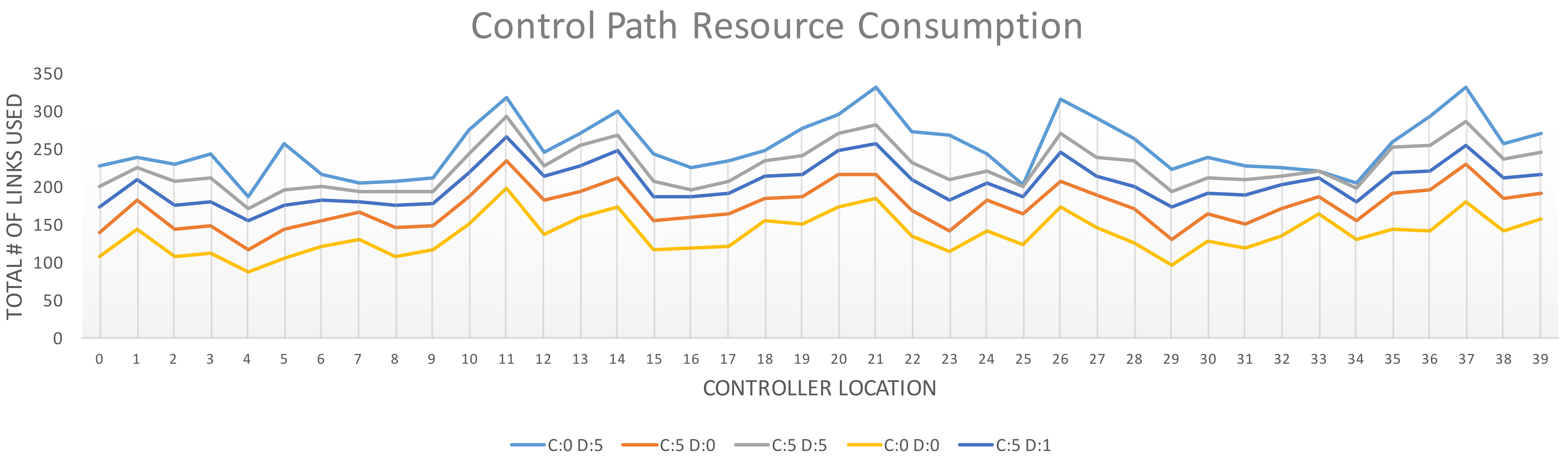
Most disruptive node failure is same with the most disruptive disaster, node 4.

0 1 2 4 30 34
1 0 33
2 0 4 31 32 35 36 38
3 4 5 10 19 30
4 0 2 3 5 6 8 16 17 29 31
5 3 4 23
6 4 7
7 6 8 25 34
8 4 7 9 25
9 8 15 18 25 29
10 3
11 13
12 13 14 15 20 22
13 11 12 14 22
14 12 13
15 9 12 29
16 4 34
17 4 30
18 9
19 3
20 12
21 27
22 12 13 23 26 27
23 5 22 29
24 25 34
25 7 8 9 24
26 22
27 21 22 28
28 27 29
29 4 9 15 23 28
30 0 3 17 39
31 2 4
32 2 34
33 1 34
34 0 7 16 24 32 33
35 2 36
36 2 35 37
37 36
38 2 39
39 30 38

Control-Path-Routing Methods

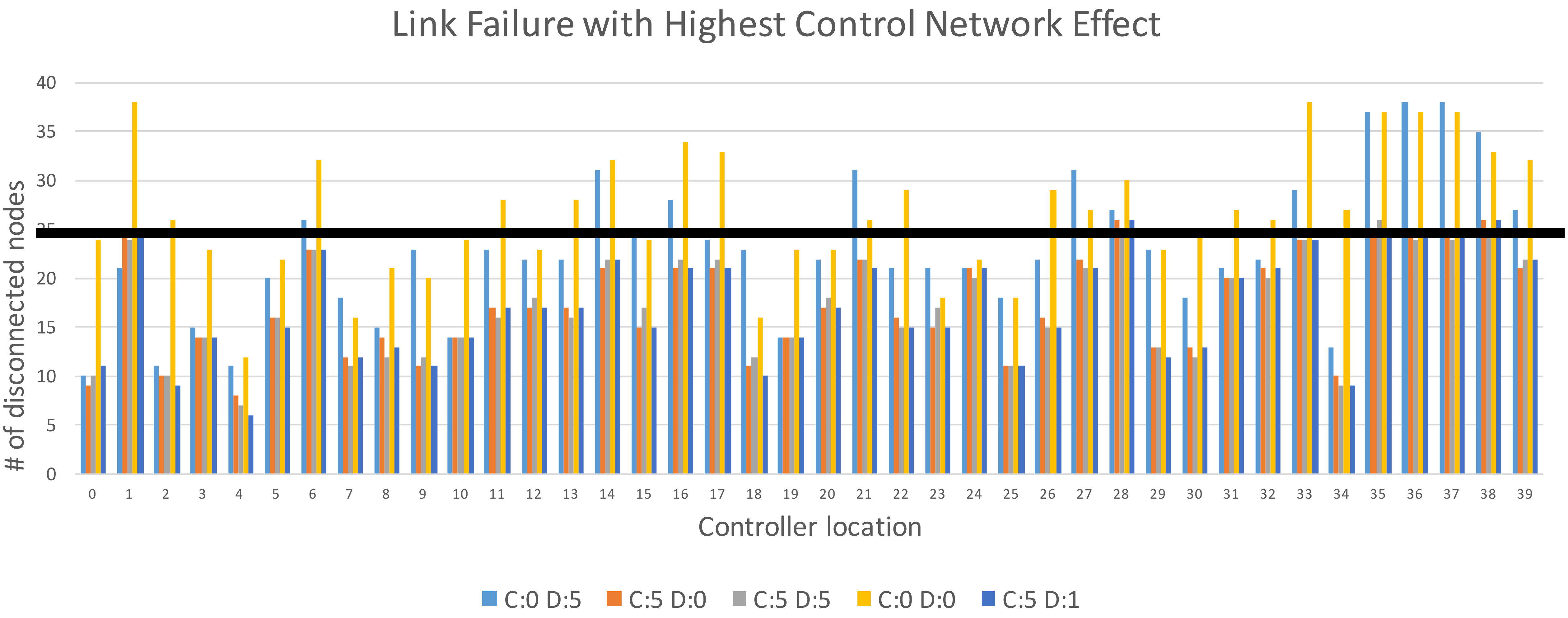
- Shortest path
- Minimize maximum control paths passing through each link (good for single link failures)
- Minimize maximum number of control paths passing through each node (good for single node failures, as when a node is down, less control paths will be affected.)
- Data-path aware routing
- Link/node/disaster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

- Controller Disjointness
- Data Path Disjointness



C: 0 D: 0 is shortest path algorithm.
C shows controller disjointness importance
D shows data path disjointness importance

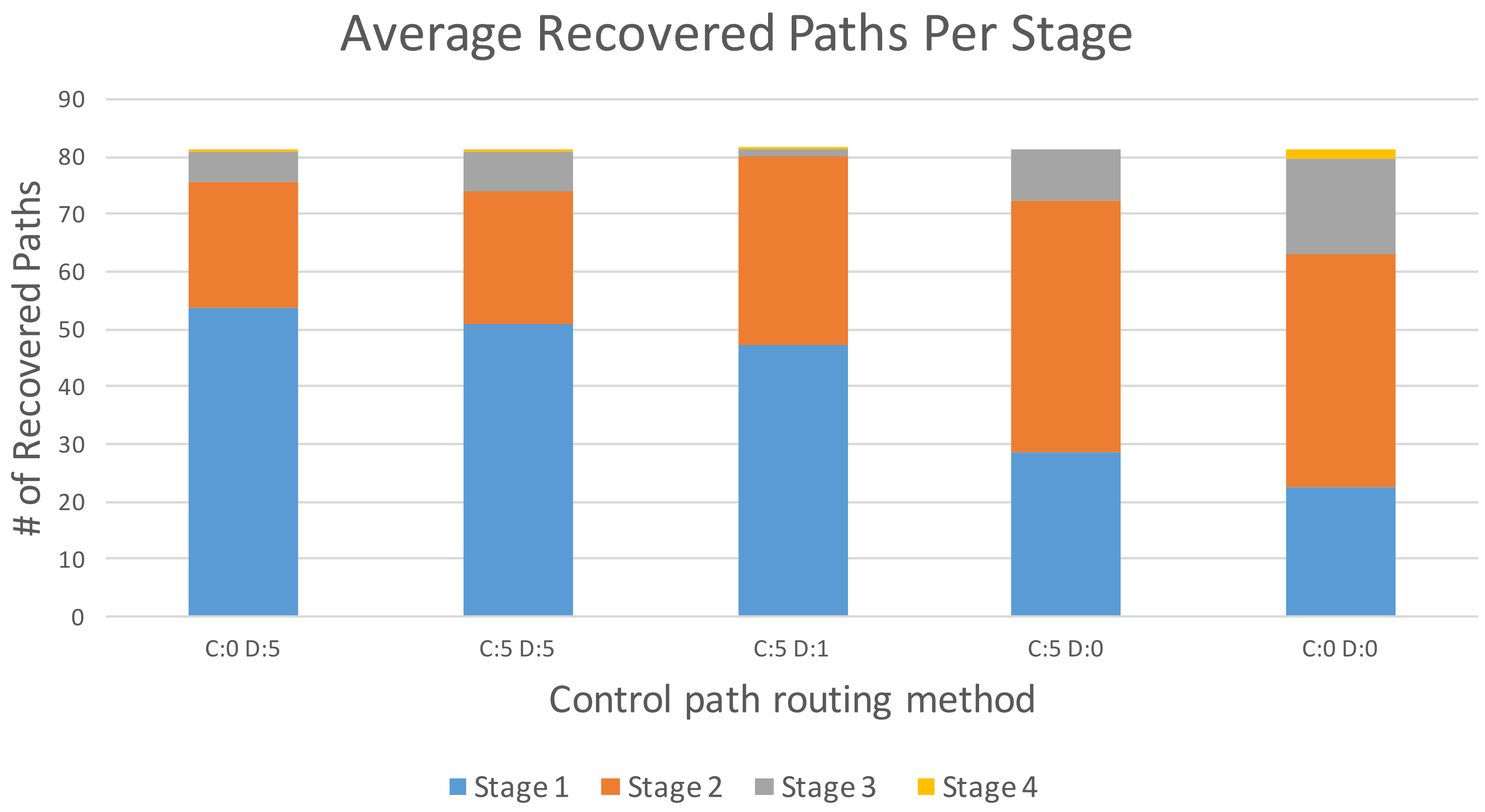
Single link failure, mostly cause 0 or 1 control path loss, recoveries are completed at first 2 stages.



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

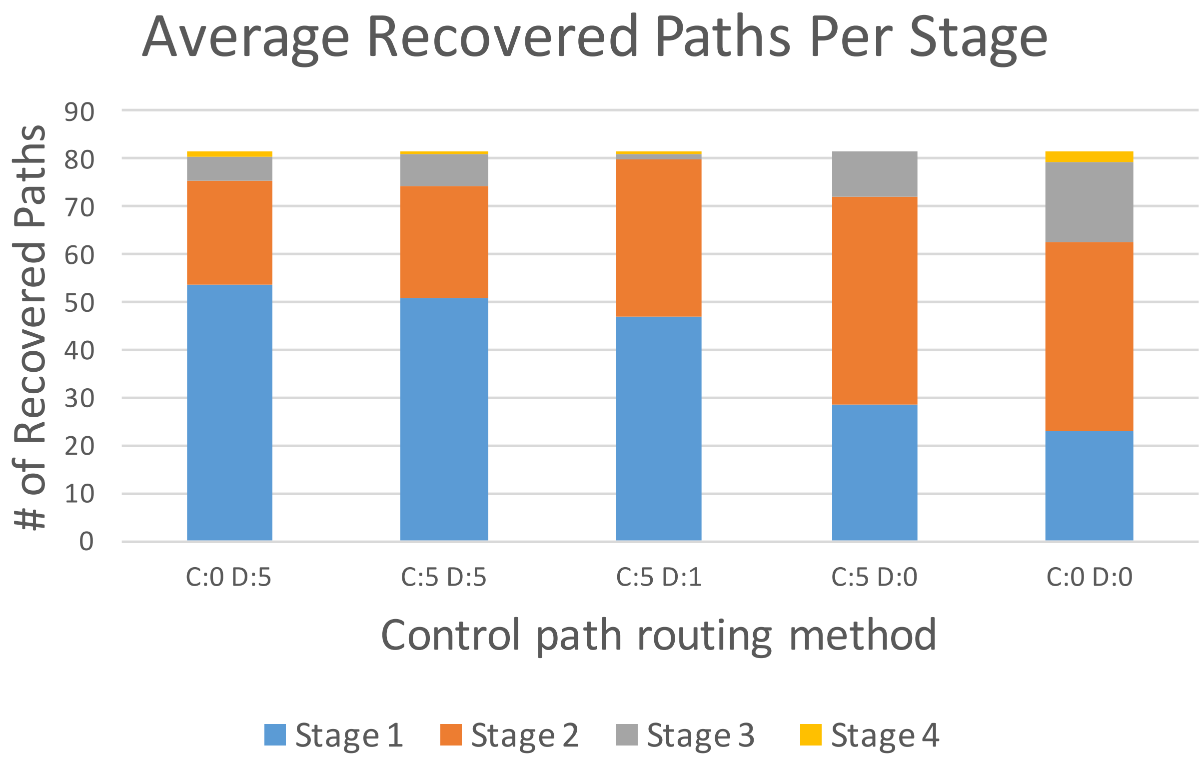
Shortest path method uses less resources, but in the worst case cause more control loss.

Controller is on node 4 (has highest node degree)



Averaged over all link failures.

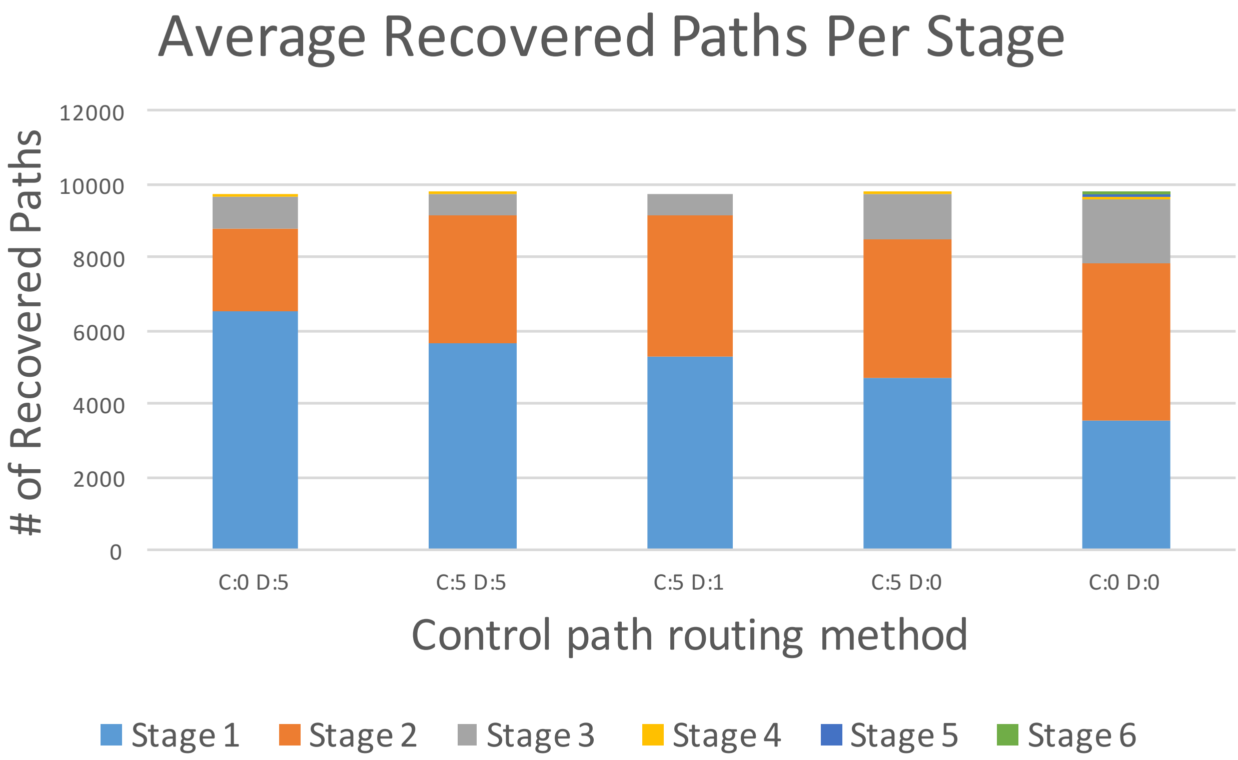
Controller is on node 4 (has highest node degree)



	Stage 1	Stage 2	Stage 3	Stage 4
C:0 D:5	6425	2619	644	72
C:5 D:5	6090	2800	812	58
C:5 D:1	5664	3918	174	4
C:5 D:0	3410	5244	1106	
C:0 D:0	2723	4815	2009	213

Averaged over all link failures.

Controller is on node 0 (has average node degree)



	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6
C:0 D:5	6569	2247	886	58		
C:5 D:5	5640	3546	522	52		
C:5 D:1	5322	3808	630			
C:5 D:0	4680	3791	1263	26		
C:0 D:0	3536	4310	1726	66	74	48