DISASTER-RESILIENT CONTROL PLANE DESIGN

Sedef Savas

October 28th, 2016

Netlab Friday Group Meeting

CONTROLLER PLACEMENT STRATEGIES FOR A RESILIENT SDN CONTROL PLANE

Petra Vizarreta, Carmen Mas Machuca, and Wolfgang Kellerer
Technical University of Munich

RNDM September 2016

UCDAVIS

Introduction

Objective: Provide reliable switch-to-controller connection with controller placement for fast failover.

Planning of the backup control paths in advance: By combining the controller placement problem with resilient routing principles to minimize the latency of the control plane, while providing the protection for every control path.

The first approach: switches have to be connected to a controller over two Disjoint Control Paths (RCP-DCP).

The second approach: switches have to be connected to two Different Controller Replicas (RCP-DCR) over two disjoint paths.

They compared wrt the unprotected scenario, in terms of **control path length**, **expected control path loss** in different failure scenarios, and **average control plane availability**.

Both methods improve the resilience, while increasing average control path length.

Motivation

- **High availability** and **low control plane latency** are necessary to guarantee the data plane performance.
- Logically centralized: exchange the information about the network state in a timely manner and maintain the accurate global overview.
- ONIX, Hyperflow solved physically distributed controllers problem??

Reliable controller placement

Controller Placement Problem metrics:

- control plane latency, latency in case of a failure, inter-controller latency, load balance between the controllers

Deterministic reliability metrics: focus on connectivity, measured as number of node or edge disjoint paths between the nodes.

K-critical: minimizes the number of controllers necessary to provide latency guarantees, while also minimizing the number of hops it take to reach a controller. Less hops are more reliable. [1]

Min-cut clustering: A cluster is defined as a set of nodes controlled by the same controller. The algorithm first finds a clustering of the network with the smallest number of edges belonging to different clusters (min-cut) and then it assigns a controller to the node that has shortest average distance to all the other nodes in the same partition. [2]

Another: maximizes number of node disjoint paths between the controllers and their assigned switches.[3]

Probabilistic reliability metrics: includes the probability of the failure of different physical components.

Approach: how many controllers (minimize) the node should connect to in order to achieve ”five nines reliability”?[4]

[1] Y. Jimnez, C. Cervell-Pastor, and A. J. Garca,”On the controller place- ment for designing a distributed SDN control layer,” in *IFIP Networking Conference*, 2014.
[2] Y.Zhang,N.Beheshti,andM.Tatipamula,”Onresilienceofsplit-archi- tecture networks,” in *2011 IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2011, pp. 16.
[3] L.F.Mulleretal.,”Survivor:Anenhancedcontrollerplacementstrategy for improving SDN survivability,” in *2014 IEEE Global Telecommunica- tions Conference (GLOBECOM)*, 2014, pp. 19091915.
[4] F. J. Ros and P. M. Ruiz, ”Five nines of southbound reliability in software defined networks,” in *Proc. 3rd Workshop Hot Topics Software Defined Networks*, 2014, pp. 31-36.

Problem formulation

None considers backup path provisioning. [3] is the closest, do not provision.

There are k controllers to place.

The goal is to minimize the average control path length (a.k.a latency). The length of both working and backup control path is jointly optimized to offer required low latencies.

To provide optimal backup paths, the working path is longer than in the unprotected case.

Disjoint Control Paths (RCP-DCP): Every node must be connected to its assigned controller over two disjoint paths. (Node disjoint.)

Different Controller Replicas (RCP-DCR): Every node must be connected to two different controllers over two disjoint paths.

Assumptions: uniform demand, no link capacities.

UCDAVIS

Baseline (unprotected scenario): Controllers are placed in a such way that the distance between the switches and their assigned controllers is minimized.

All switches being controlled by their nearest controller and all control paths being the shortest paths between the switch and the assigned controller. If a failure happens along the control path the connection with the controller will be lost and it will take time until the new control path is established.

To prevent long restoration times, the backup control paths have to be provided proactively.

Comparison: Average control path length, expected control path loss, average connection availability, and solving time.

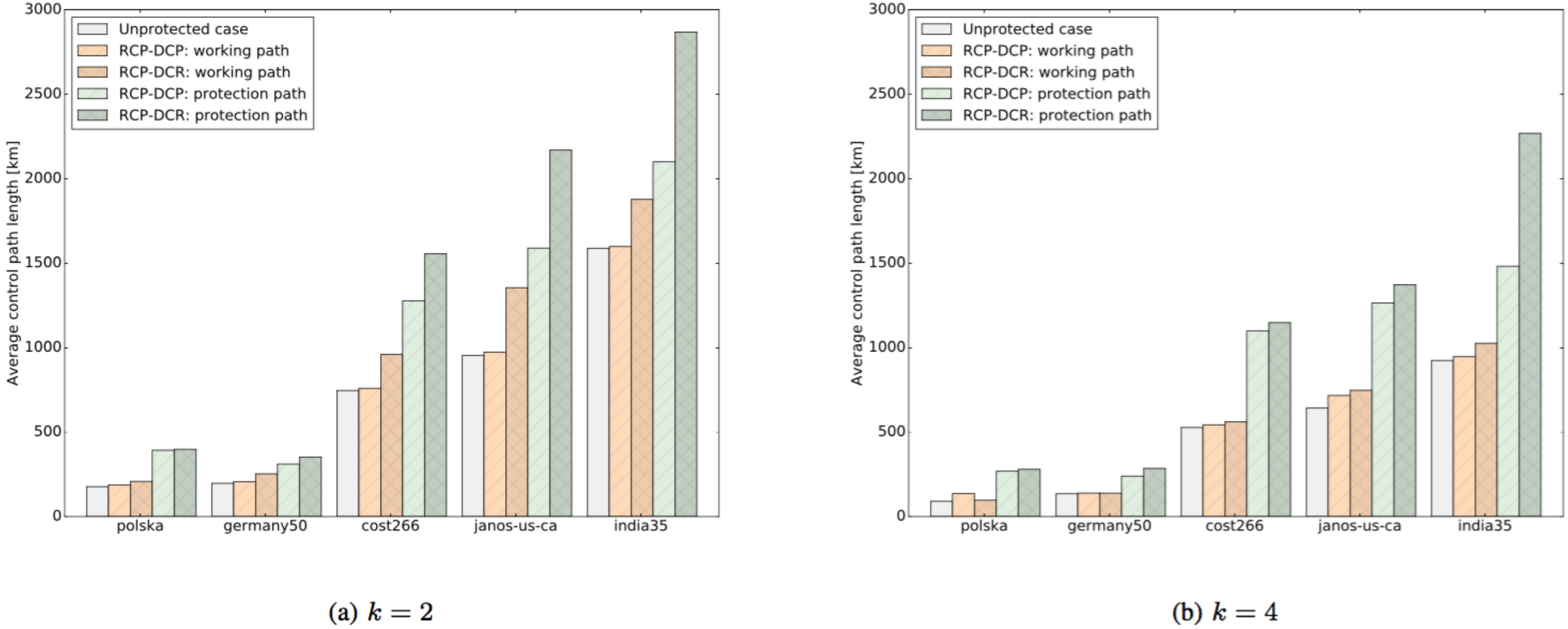


Fig. 1: Average control path length for different number of controllers in selected SND topologies [20].

2-5 % increase in working paths.

Impact of number of controllers:

Fig. 2: Cost266 network [20].

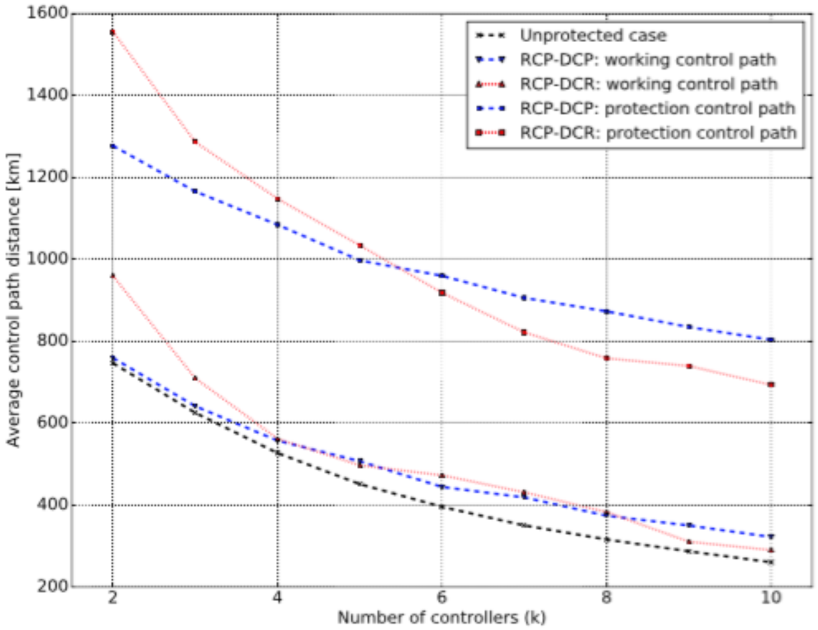


Fig. 3: Impact of the number of controllers k on the average control path lengths in Cost266 network.

UCDAVIS

Expected control path loss (ECPL)

Assumption: Only single and double link failures. Link failure probability is proportional to its length.

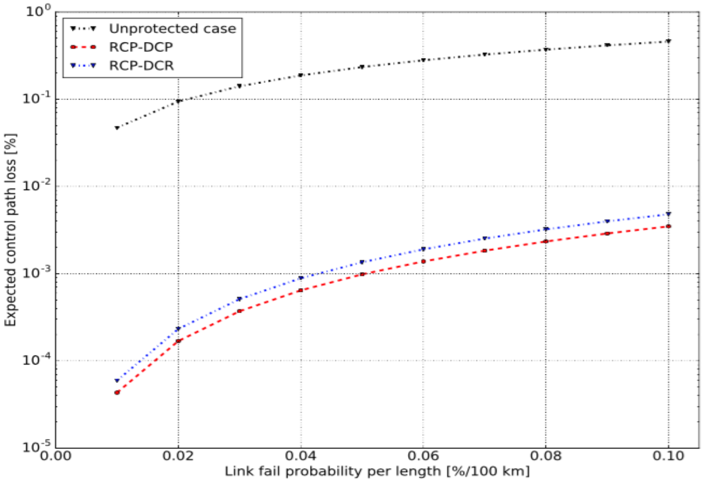


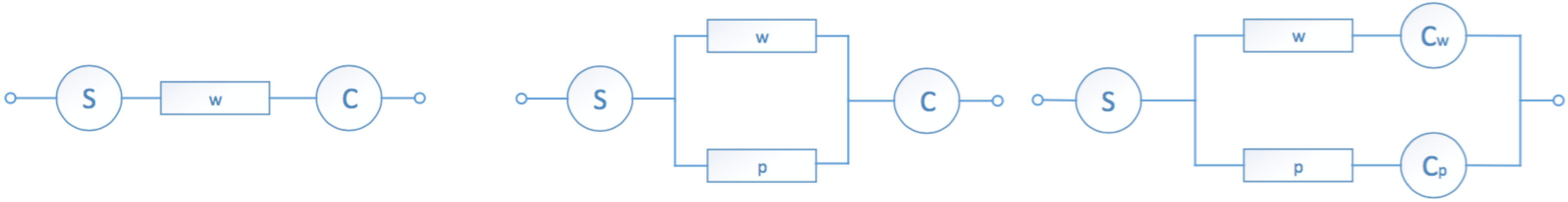
Fig. 5: Expected control path loss when single and double link failures are considered.

Average control plane availability

Control path availability is a function of availability of all nodes and links on the control path.

Failure rate per length of the link = 0.1 %/100 km.

The node availability is varied between 98% and 100%.



(a) Unprotected control path
$A = a_s a_w a_c$

(b) Disjoint Control Paths (RCP-DCP)
$A = a_s(1-(1-a_w)(1-a_p))a_c$

(c) Different Controller Replicas (RCP-DCR)
$A = a_s(1-(1-a_w a_{c_w})(1-a_p a_{c_p}))$

Fig. 4: Reliability block diagram and control path availability expressions for unprotected and RCP-DCP and RCP-DCR models. S stands for switch; w for working and p for protection control path; C for controller (C_w: working and C_p: backup controller).

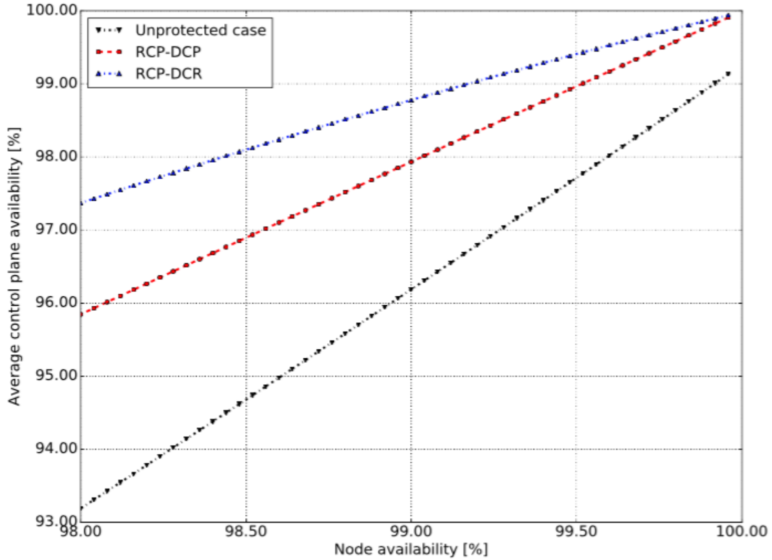


Fig. 6: Average control plane availability as a function of node availability. Link failure probability was 0.1%/100 km.

UCDAVIS

Limitations

Do not decide on number of controllers.

No specific controller locations.

Unlimited link capacity.

Comparison

No latency limit while selecting controllers and paths. Primary path length do not differ much, but what about backup path? (Disjoint)

OPTIMAL PLACEMENT OF CONTROLLERS IN A RESILIENT SDN ARCHITECTURE

Nancy Perrot, Thomas Reynaudt
Orange Labs, France

DRCN May 2016

Introduction

Focus: The optimal number of controllers and their location while embedding QoS and load balancing constraints for several level of backup controllers.

Minimizes the number of active controllers needed in a WAN while considering several levels of back up controllers and maintaining tight latency, capacity, and balancing constraints.

Questions to deploy SDN in WAN:

How many controllers are required to manage the whole network?

What are the clusters of nodes depending on each controller ?

What are the right nodes to place them in the cluster ? and so on.

The capacity of a controller in terms of processing, memory and in/out bandwidth limits the number of nodes (e.g. switch, routers) that could be managed by a single controller.

Different criteria explored for CPP:

minimum number of controllers, minimization of the worst case latency between nodes and controllers, minimization of the inter controller latency, optimization of the balancing of clusters of routers for each controller, taking into account or not the controllers capacities, and considering failure of both controllers and routers.

CPP reduces to a Facility Location Problem and is proved to be **NP Hard.**

Routers are automatically assigned to their nearest available controller and a maximal latency constraint is introduced to ensure good performance.

In a pre-processing procedure, all the candidate nodes which a router could be reached within the maximal required latency is computed.

Aim is to find the minimal number of active controllers such that each router node is assigned to one of its closest controller, and such that all the controllers have an equivalent number of nodes to manage (the so-called load balancing constraints).

Mathematical formulation

- Only specific nodes can be controller locations.
- Max latency between router-controller(30% of the graph diameter) and controller-controller (70% of the graph diameter) is set.
- Max allowed difference between controller loads is set (2).
- Each router must be covered by at least one controller within the latency bound.
- Each router to exactly one controller.
- Routers are assigned to their nearest available controller.

For consistency purpose, the delays are as ratio of the graph diameter.

Relaxing the constraint on the load balancing of the controllers clusters doesn't allow to save controller (still 4 optimal controllers) while it yields to very unbalanced clusters, 3 to 18 nodes.

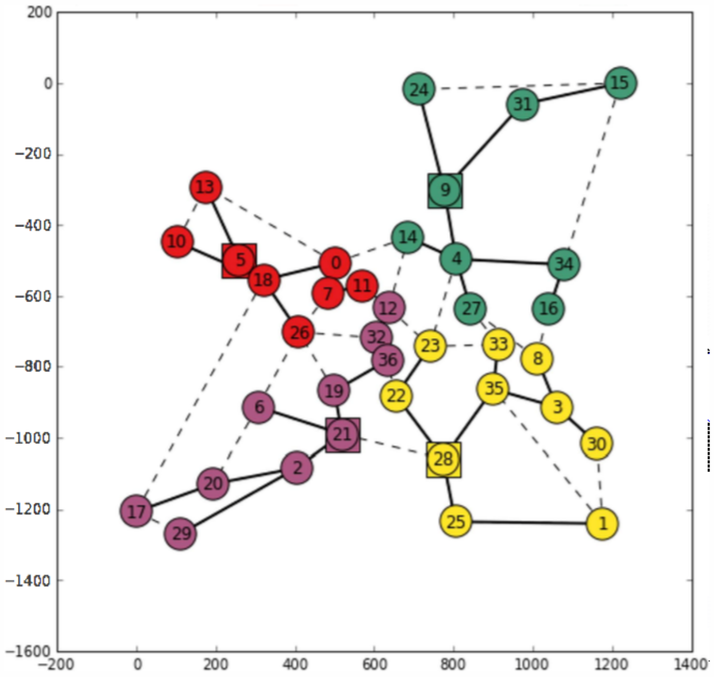


Fig. 1. Optimal solution with $l_{max} = 30\%$, $l_{cc\text{-}max} = 70\%$ and $\delta = 2$

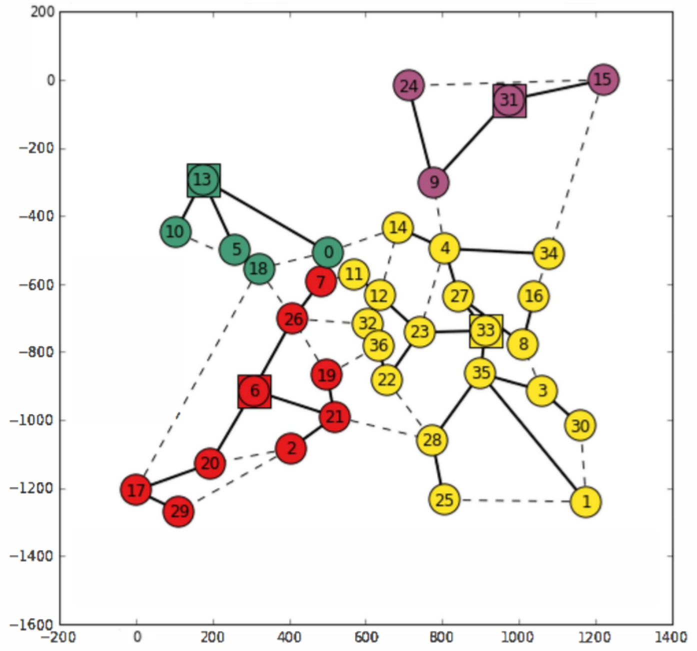
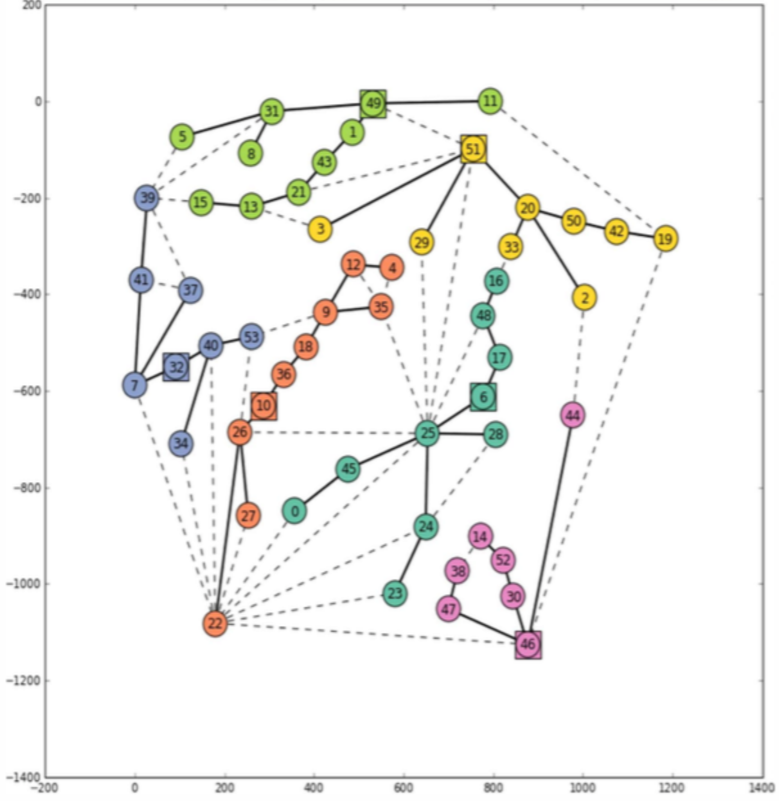


Fig. 2. Optimal solution with $l_{max} = 30\%$, $l_{cc\text{-}max} = 70\%$ and $\delta = 37$

TABLE I. SENSITIVITY ANALYSIS OF THE CONSTRAINTS ON THE OPTIMAL SOLUTION

l_{max}	$l_{cc\text{-}max}$	δ	Placement
0.4	0.7	3	4 21
0.3	0.7	3	9 18 21 28
0.27	0.7	3	2 4 11 26 28 34
0.25	0.7	3	2 9 18 22 23 24 26 30
0.35	0.4	3	$\emptyset$
0.35	0.7	3	10 19 22 29
0.35	0.9	3	14 16 19 32
0.35	1	3	2 5 8 9
0.35	0.7	1	2 3 5 9
0.35	0.7	7	6 9 28
0.35	0.7	15	6 31 35



ig. 6. Optimal solution with $l_{max} = 30\%$, $l_{cc\text{-}max} = 70\%$ and $\delta = 3$

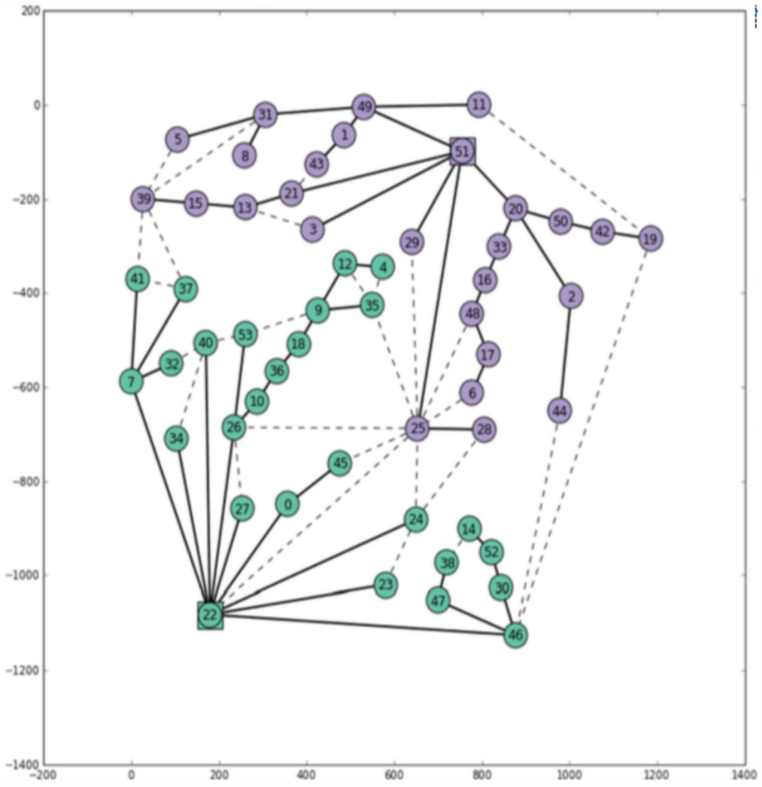


Fig. 7. Optimal solution with $l_{max} = 50\%$, $l_{cc\text{-}max} = 70\%$ and $\delta = 3$

Resilient controller placement problem

If a controller fails, its routers must be assigned to another one, leading then to an increase of the latency between routers and controller and potentially to unbalanced domains, especially if the secondary controller takes the management of all the routers of the failed controller.

Consider a failure probability for each controller assuming that these failure probabilities are inherent to the controller and independent om each other.

No link failures.

Main objective is still to minimize the number of active controllers.

A router is assigned to its k^{th} controller if and only if all the primary controllers fail. When there is no controller at level k for router, it induces a penalty cost.

Constraints

- A router j is either assigned to a controller or receive a penalty at each level k.
- Each router has a minimum of x level of backup controllers.
- Prevent a router to be assigned to the same controller at two different levels.
- The load balancing constraints.
- Latency constraints.

This is a bi-objective problem that is solved in two stages :

- First, given latency const., find number of controllers.
- Second, locate the controllers considering resiliency.

Numerical examples

Failure prob. = 0.95

PARAMETERS USED FOR SIMULATIONS

Parameter	Value
Number of nodes	[10,80]
l_{max}	{3000,5000,7000}
$l_{cc\text{-}max}$	7000
δ	3
γ	2
η	1

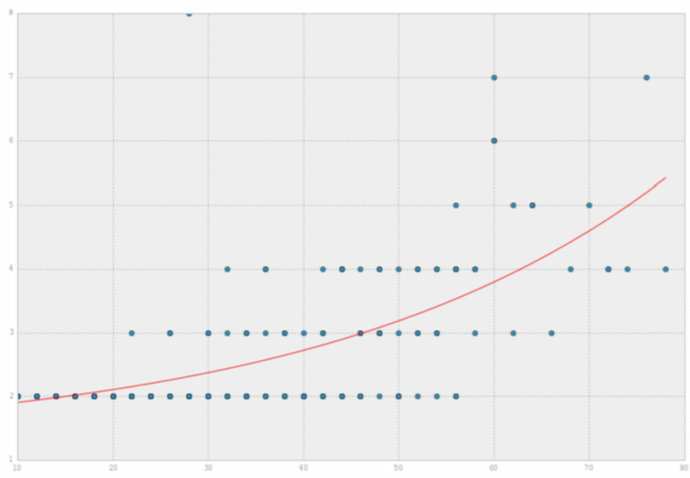


Fig. 8. Number of controllers on the number of nodes

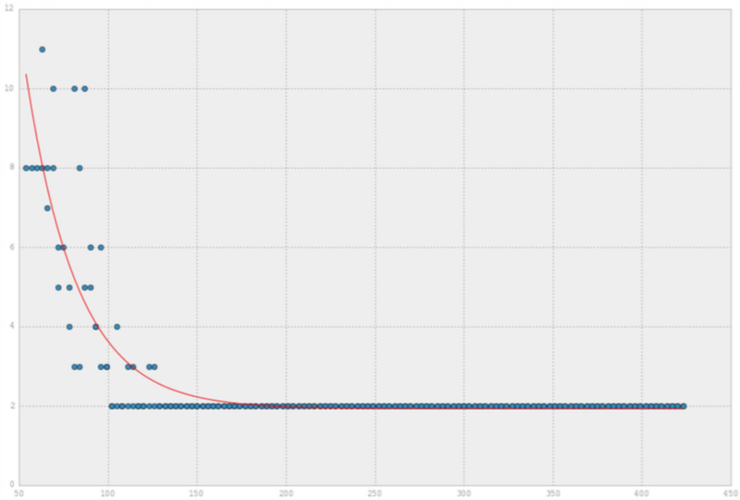


Fig. 9. Number of controllers on the number of arcs

35-nodes graph and the minimum number of required controllers is 2. Decrease is very fast as from a degree of 10 each controller is connected, in average, to 30% of the routers which is relatively high.

The diameter of graph increases with the number of nodes resulting in the need for more controllers.

When graphs are more dense (i.e., more arcs), more paths between routers leads fewer required controllers.

UCDAVIS

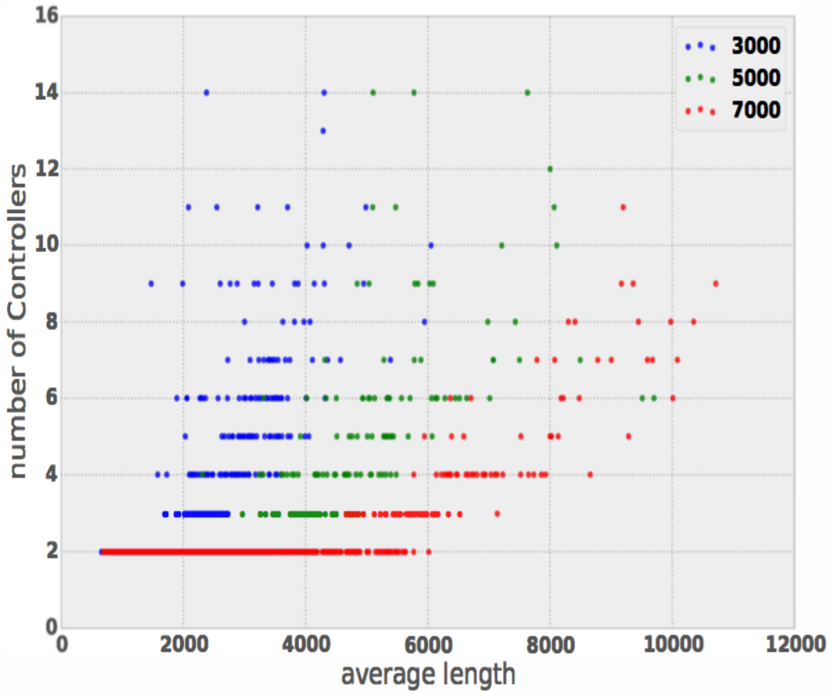


Fig. 10. Number of controllers on the average length of the shortest path in the graph and on l_{max}

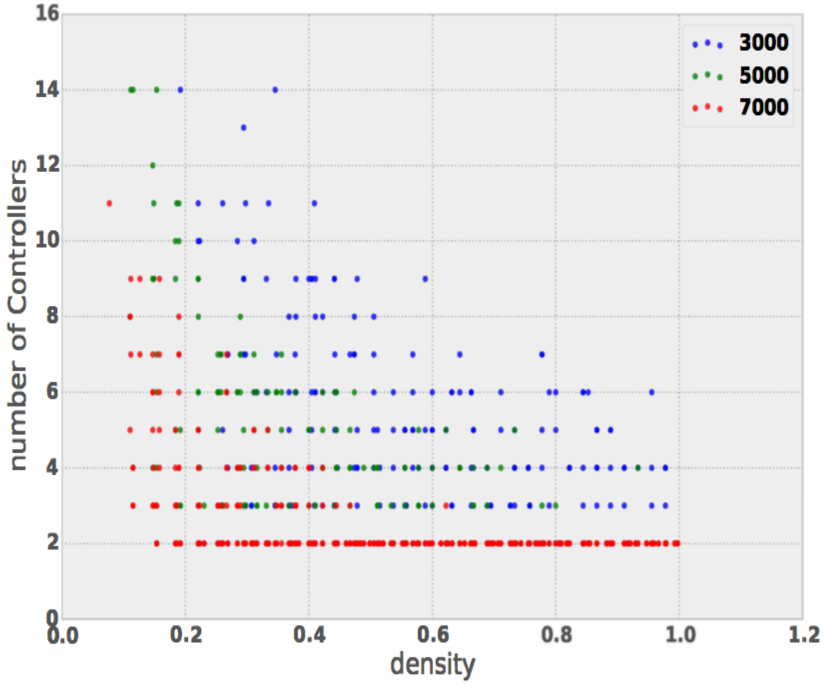


Fig. 11. Number of controllers on the graph density and on l_{max}

UCDAVIS

DISASTER-RESILIENT CONTROL PLANE DESIGN AND MAPPING IN SOFTWARE-DEFINED NETWORKS

S. Sedef Savas[1], Massimo Tornatore[1,2], M. Farhan Habib[1], Pulak Chowdhury[1], and Biswanath Mukherjee[1]

[1]*University of California, Davis, USA*

[2]*Politecnico di Milano, Italy*

HPSR July 2015

Motivation

Survivable control plane problem

- push intelligence into the switch
- distributed SDN control plane
  - how many controllers?
  - where to place them?
  - delay, survivability, capacity requirements, synchronization overhead, etc.

Distributed design works for small scale failures but doesn't give any survivability/connectivity guarantee in case of large-scale disasters.

Reprovisioning control comm. channels before recovering disrupted connections may cause huge delay in the event of disasters.

Problem Statement

We propose to design distributed control-plane as an overlay network mapped over physical network, *Control Network Mapping* problem.

ensure control-plane connectivity against both single point of failures and large-scale disaster failures in SDN.

The ultimate aim is to make the control plane resilient to controller failure, inter-controller communication and controller-switch communication failure.

Control Plane Mapping

A distributed control plane can be designed as an overlay (i.e., virtual/logical) network mapped over a physical (i.e., backbone) network.

- virtual nodes where controllers are located and virtual links connects them.

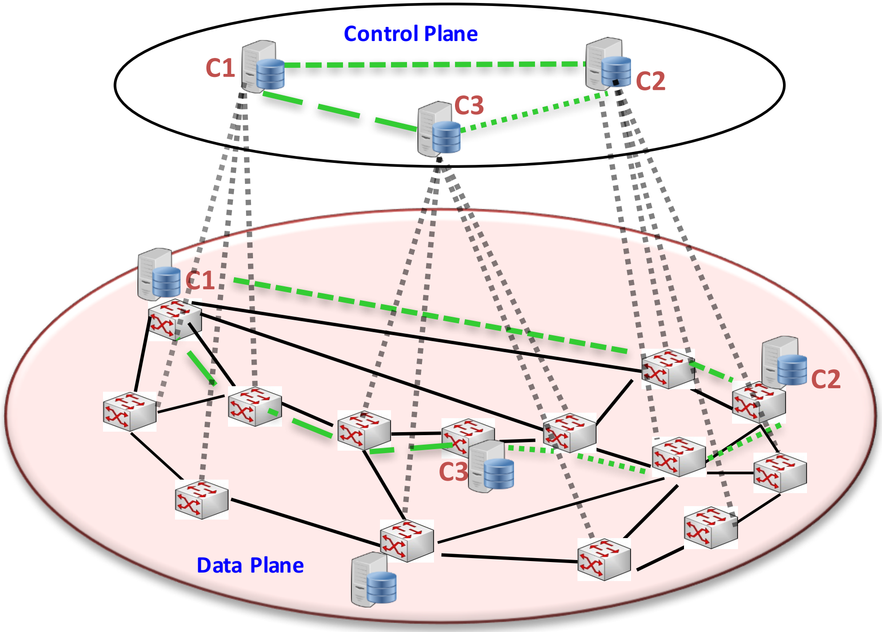
We propose a survivable control plane mapping scheme to ensure control-plane connectivity against both single point of failures and large-scale disaster failures in SDN.

Novelty of the Problem

Joint optimization of

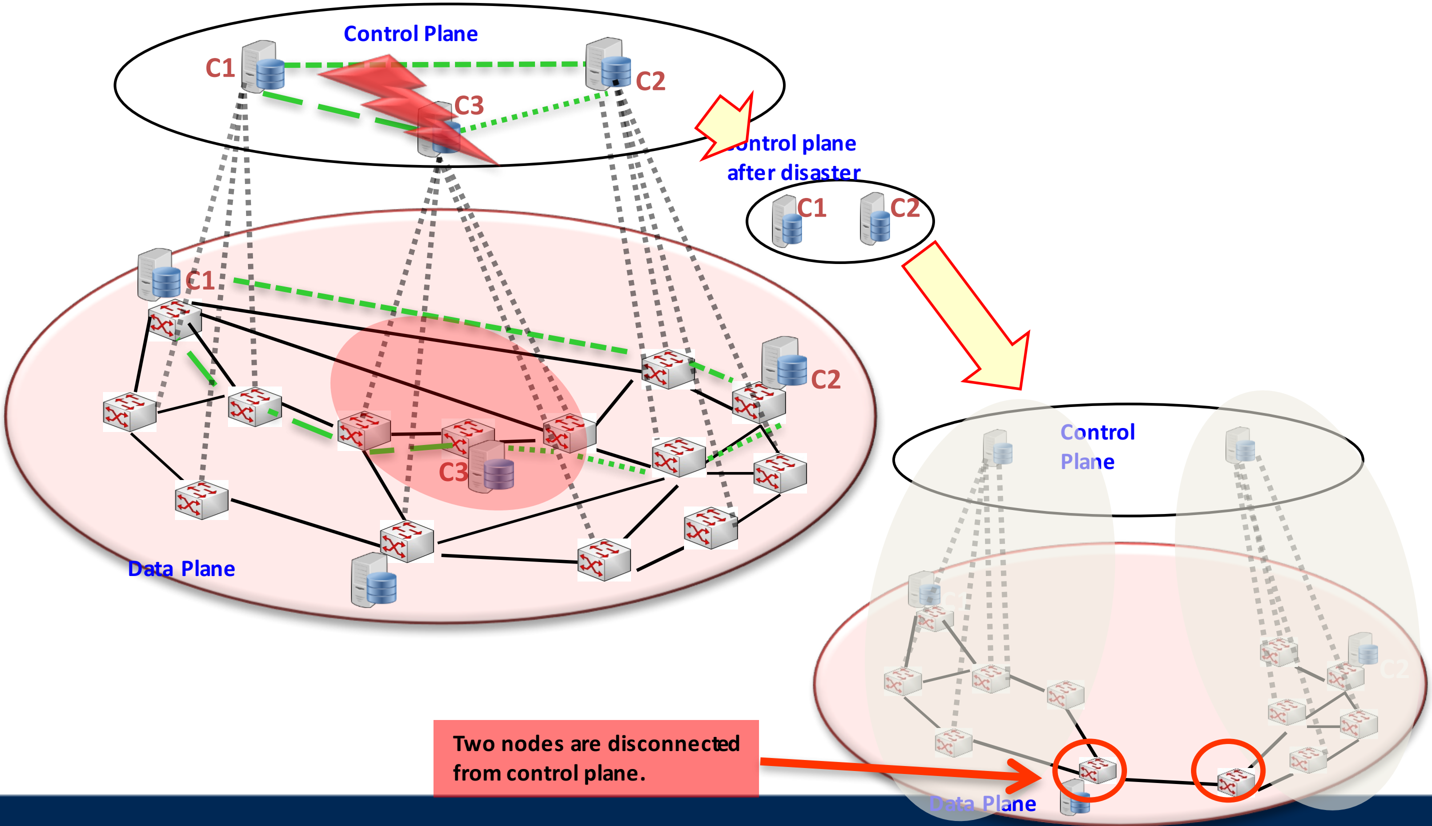
- # of controllers
- Virtual network topology
- Virtual network mapping (virtual node and link mapping which exist in the literature)

Control Network Design and Mapping

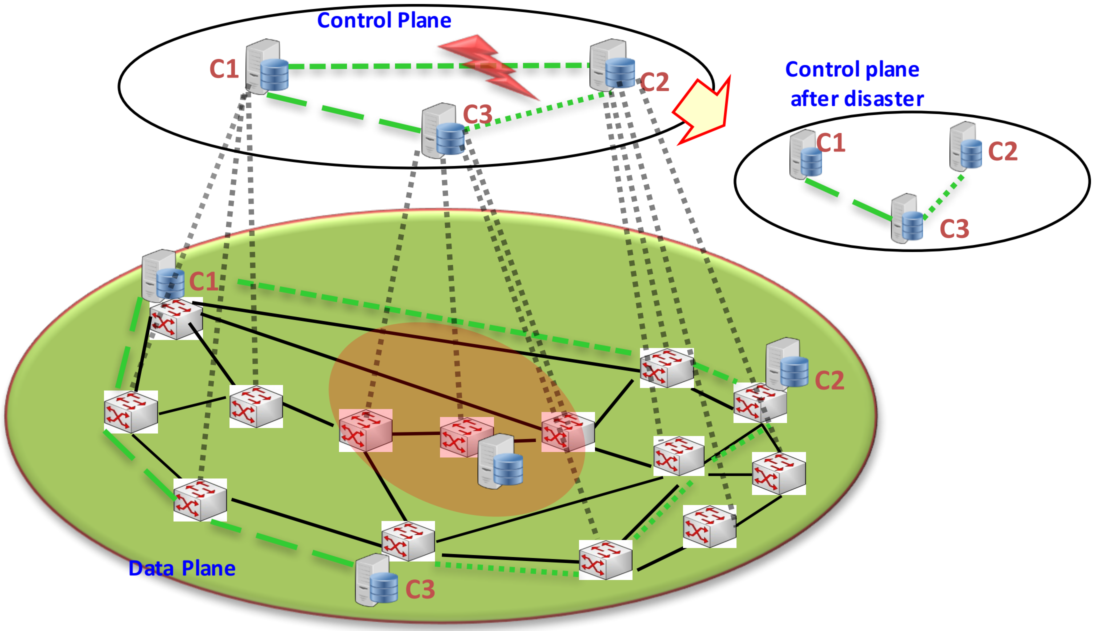


- Determining # of controller and their placements ✓
- Determining logical control plane topology ✓
- Mapping control plane to physical network ✓
- Controller-switch assignment ✓

Disaster-Unaware Control Network Design & Mapping



Disaster-Aware Control Network Design & Mapping



Problem Formulation

Given

$G(N,E)$: Network topology where V is the set of nodes and E is the set of directed links.

F: Set of candidate nodes where a controller can be deployed.

$d_{ij} \in \{0,1\}$: is 1 if node j is located within node i's reachability island.Reachability Island of one node is a circular region where shortest path distance to every node within this region from this node satisfies the latency constraint of this node.

k: Number of controllers that are guaranteed to be located by every switch within a latency limit. For example, in our case we guarantee that there will be at least 2 controllers that will be located within the latency constraint by each switch.

q: At least q controllers are active at any time in the system.

t_i: Traffic passing through node i.

CA: Maximum amount of traffic a controller can maintain (controller capacity).

DIP: Number of disjoint paths between every controller pair in the virtual topology. Used to decide on virtual links.

P_{ij}: is the set of possible paths to use for virtual link mapping between node i and node j.

$U_p^y \in \{0,1\}$: is 1 if path $p$survives from disaster y.

$Y=\{y|\ y=<E_y,\rho_y>\}$: Set of DZs where $E_y$is the set of links that is member of Disaster y and ρ_y is the probability that disaster y causes a failure.

Assumption: No resource limit on the links.

Variables

$C_f \in \{0,1\}$: is 1 if a controller is deployed and active on node f.

$C_{st}^p \in \{0,1\}$: is 1 path p is used for virtual link between controller s and controller t.

$a_{if} \in \{0,1\}$: is 1 if switch $i$is assigned to controller f.

$m_{ij}^{st} \in \{0,1\}$: is 1 if virtual link between controller$i$ and controller j carries flow for (s,t) $\in (F,F)$. We use flow formulation to make sure that there exists some paths which use virtual links between every controllers.

$o_{ifst} \in \{0,1\}$: is 1 if nodes$i,f,s$and$t$are controllers.

$V_{if} \in \{0,1\}$: is 1 if there is a virtual link between controller$i$ *and* controller f.

$A_{st}^p \in \{0,1\}$: is 1 path p is used for virtual link between controller s and controller t.

$K_{if}^y \in \{0,1\}$: is 1 if virtual link V_{if} survives from disaster y.

$X_{if}^{sty} \in \{0,1\}$: is 1 if virtual link between controller i and controller j carries flow for (s,t) $\in (F,F)$ after disaster y

OBJECTIVE

Minimize (# of Controllers + # of virtual links between controllers)

of controllers

$\Sigma_{f \in N} C_f \cdot d_{if} \geq k \cdot 1 - C_i \quad \forall i \in N$ where k is 2 in our case (1)

$\Sigma_{f \in F} C_f \geq Min_{controller} \quad \forall i \in N$ where $Min_{controller}$ is 2 in our case (2)

$\Sigma_{f \in F} C_f \leq Max_{controller} \quad \forall i \in N$ where $Max_{controller}$ is 6 in our case. (3)

Switch-controller assignment

$a_{if} = 0$ if f is not a server node $\quad \forall f \in N$ (3.5)

$C_f \leq \Sigma_{i \in N} a_{if} \quad \forall f \in N$ (4)

$C_f \geq \Sigma_{i \in N} a_{if}\ / M \quad \forall f \in N$ (5)

$\Sigma_{i \in N} a_{if} \cdot t_i \leq CA \quad \forall f \in N$ (6)

$a_{if} \geq C_f \quad \forall i,f \in N$ when $i = f$ (7)

$\Sigma_{f \in F} a_{if} = 1 \forall i \in N$ (7)

$a_{if} \leq d_{if} \quad \forall i \in N, \forall f \in F$ (8)

$a_{if} \leq \neg C_i \wedge C_f \quad \forall i \in N, \forall f \in F$ (8.5)

Virtual link selection between controllers

$o_{ifst} = C_f \wedge C_i \wedge C_s \wedge C_t \quad \forall f,i,s,t \in F$ (11)

$m_{if}^{st} \leq o_{ifst} \quad \forall f \in F \quad \forall i \in F \quad \forall (s,t) \in (F,F)$ (12)

$m_{if}^{st} = 0$ when $i = f$ and $s = t \ \forall f \in F \quad \forall i \in F \quad \forall (s,t) \in (F,F)$

$\Sigma_{f \in F} m_{if}^{st} - \Sigma_{f \in F} m_{fi}^{st} = \begin{Bmatrix} \geq 2, & i = s \\ \leq -2, & i = t \\ 0, & ow \end{Bmatrix} \quad \forall i,s,t \in F$ when i ≠f && s≠t (13)

$V_{if} = \begin{Bmatrix} m_{if}^{st}, & i = s \text{ and } f = t \\ 0, & ow \end{Bmatrix} \quad \forall i,s,t \in F$ (14)

Controller-controller virtual link mapping to physical

$\Sigma_{p \in P_{if}} A_{if}^p = V_{if} \quad \forall i,f \in F$ (15)

$\Sigma_{p \in P_{if}} A_{if}^p \cdot U_{pif}^y \leq K_{if}^y \quad \forall y \in Y \quad \forall i,f \in F$ (16)

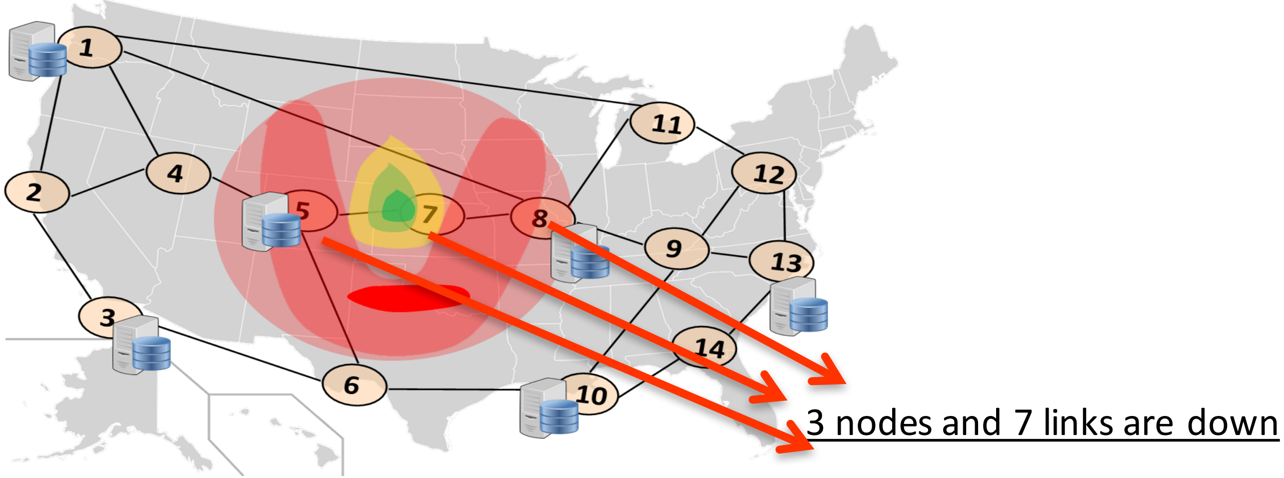
$\Sigma_{p \in P_{if}} A_{if}^p \cdot U_p^y / M \geq K_{if}^y \quad \forall y \in Y \ \forall i,f \in F$ (17)

$\Sigma_{p \in P_{if}} A_{if}^p \cdot U_p^y = K_{if}^y \quad \forall y \in Y \quad \forall i,f \in F$ (16)

$K_{if}^{sty} \leq K_{if}^y \forall i,f,s,t \in F \ \forall y \in Y$ (18)

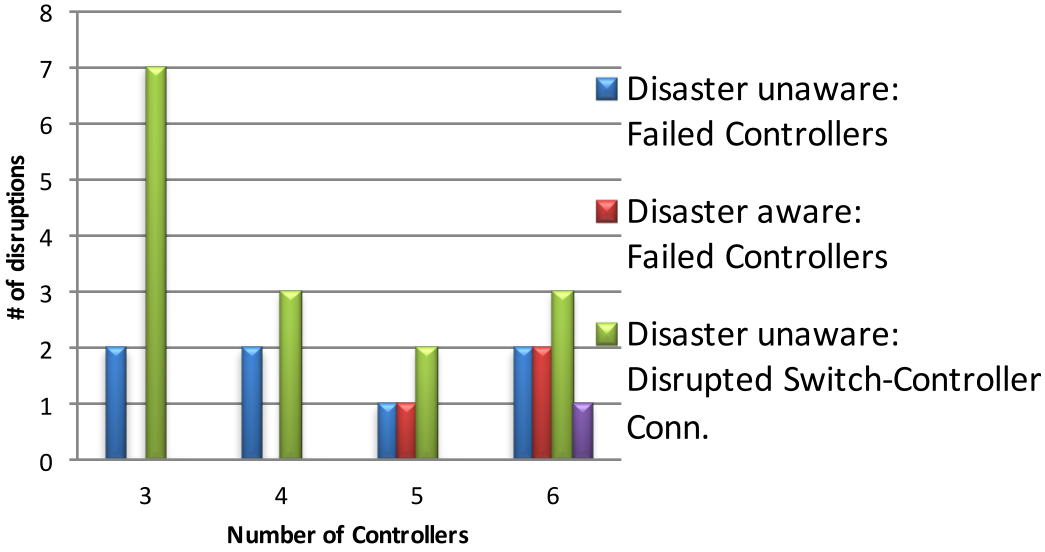
$\Sigma_{f \in F} X_{if}^{sty} - \Sigma_{f \in F} X_{if}^{sty} = \begin{Bmatrix} +1, & i = s \\ -1, & i = t \\ 0, & ow \end{Bmatrix} \quad \forall i,s,t \in F \quad \forall y \in Y$ (19)

ILLUSTRATIVE NUMERICAL EXAMPLES -1



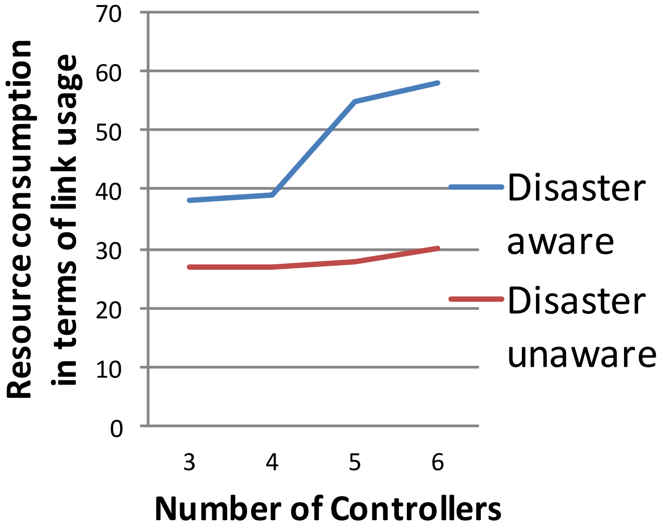
NSFNet topology with a potential EMP attack.
Different colors show EMP fields with different strengths.

ILLUSTRATIVE NUMERICAL EXAMPLES -2



Comparison of disaster-aware (Min-Risk) and disaster-unaware (Min-Resource) schemes in terms of disruptions on the control plane caused by an EMP attack.

Illustrative Numerical Examples -3



Comparison of resource consumption of control plane and switch-to-controller communication channels.

Conclusion

Our survivable control plane design reduces the probability of losing communication between controller-to-controller and switch-to-controller in case of very large-scale disasters with reasonable increase in resource usage.