Robust Control Plane for Transport Software Defined Networks

Rafael B. R. Lourenco in January 11, 2018.

UC DAVIS
UNIVERSITY OF CALIFORNIA

Outline

1. Transport SDN Characteristics
2. Disaster Resilient Controller Placement for T-SDN
3. ILP
4. Illustrative Result

UCDAVIS

Transport Software Defined Networking

”Transport SDN (T-SDN) is an SDN-based architecture for the control and management of transport networks, that could involve multi-layer, multi-domain and multi-vendor scenarios. The Optical Internetworking Forum (OIF) defines Transport SDN (T-SDN) as a subset of SDN-architecture functions comprising the transport network relevant components.”

[1] R. Alvizu, G. Maier, Comprehensive Survey on T-SDN: Software-Defined Networking for Transport Networks, IEEE COMMUNICATIONS SURVEYS & TUTORIALS, 2017

T-SDN Characteristics

- High resiliency requirements
- Heterogeneous technologies (OTN, OCh, MPLS, etc.)
- Heterogeneous architectures
- Optical domain impairments
- Multiple layers: L0, L1 and L2

[1] R. Alvizu, G. Maier, Comprehensive Survey on T-SDN: Software-Defined Networking for Transport Networks, IEEE COMMUNICATIONS SURVEYS & TUTORIALS, 2017

T-SDN Characteristics

- SDN in Layers 1 and 0 is not yet fully enabled
  - Some argue that the full capability of SDN will not be deployed in L0
- Historically, network equipment providers have increased their solutions competitive advantage by introducing proprietary technologies and improving their management systems; this led to heterogeneous networks with several interoperability issues

➢ Thus, transport networks are usually composed of several administratively isolated islands (domains) for each of its vendors

[1] R. Alvizu, G. Maier, Comprehensive Survey on T-SDN: Software-Defined Networking for Transport Networks, IEEE COMMUNICATIONS SURVEYS & TUTORIALS, 2017

T-SDN Control Plane Architectures

- Monolithic (SDON): single controller that manages the entire network
- Hierarchical (HT-SDN): Different domain controllers are orchestrated by **master controller** (or Orchestrator) through North/Southbound Interfaces
  - Standardization bodies (ONF and OIF) agreed on hierarchical architecture
  - Better suits multi-domain Transport Networks
- Flat or Mesh (FT-SDN): Direct controller to controller communication through West/Eastbound Interfaces
  - Not the focus of Transport SDN due to difficulty in interoperability among domains

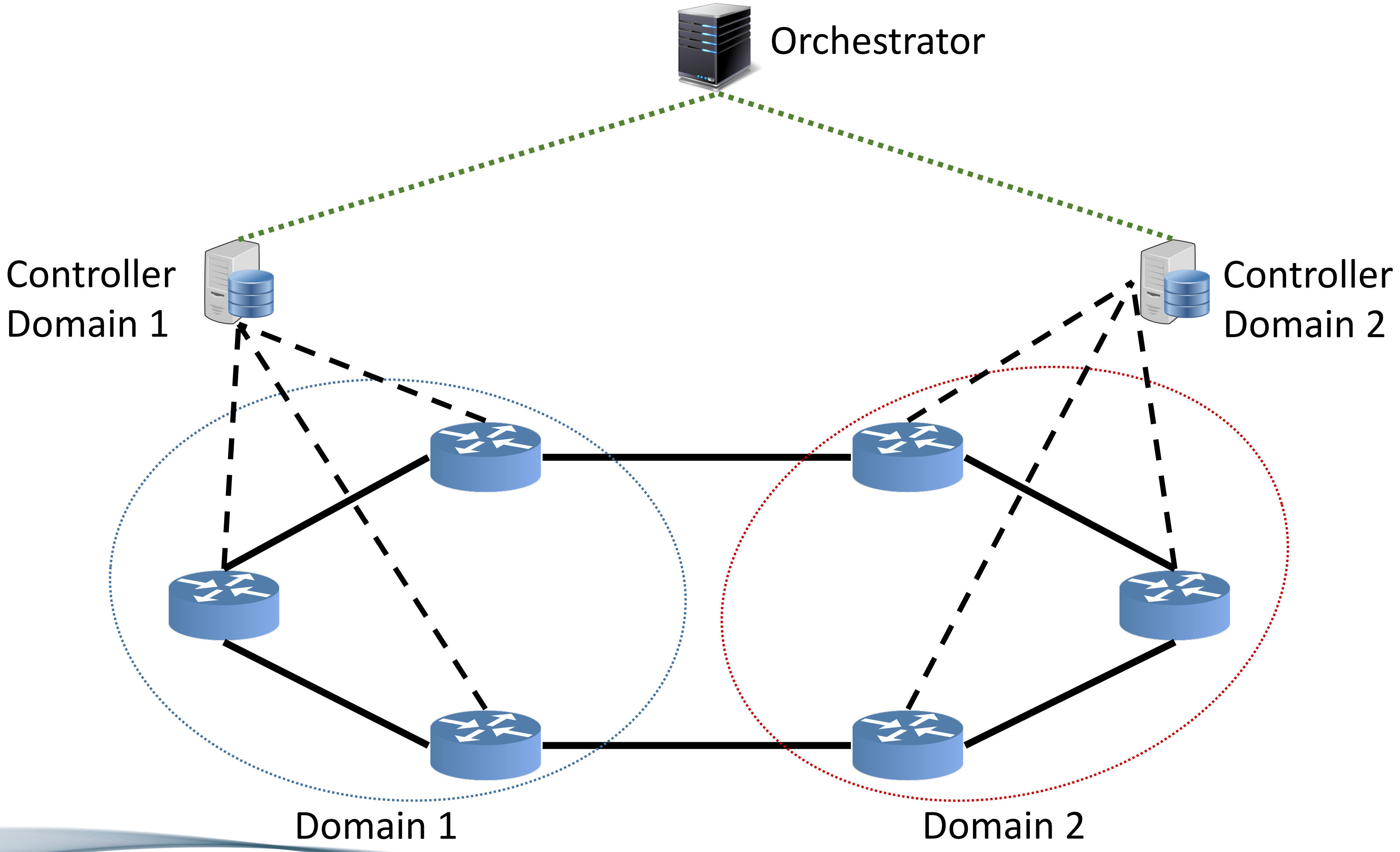
[1] R. Alvizu, G. Maier, Comprehensive Survey on T-SDN: Software-Defined Networking for Transport Networks, IEEE COMMUNICATIONS SURVEYS & TUTORIALS, 2017

Our Proposal

- Study how to place controller in a multi-domain T-SDN, with a centralized orchestrator (and backup) who is in charge of all different domain controllers. The controllers do not interact with one another directly, only through the orchestrator. There are two levels of control traffic in this situation:
  - Controller-to-Orchestrator traffic: this traffic is important for higher level (more than Layer 0 and 1) type of functionalities, as the orchestrator is in charge of managing inter domain traffic, mainly. This results in not so stringent latency requirements as the other type
  - Node-to-controller traffic: in a T-SDN this traffic is more challenging as it relates to lower layers (L0,L1). This requires stringent latencies between node and controller

UCDAVIS

Example of T-SDN Control Plane



UC DAVIS

The Capacitated Controller and Orchestrator Placement Problem (CCOPP)

- Deciding how many controllers and orchestrators to deploy, where to place them in the network, and how to assign switches to controllers and controllers to orchestrators
- **Given**: topology, the controllers' management capacity (and possibly the orchestrator's), the maximum latencies, different nodes domains

CCOPP for T-SDN

As in T-SDN reliability is one of the most important issues, we focus on:

➢ **Disaster-Resilient Placement of Controllers and Orchestrator in a Transport Software Defined Network**

- **Given:** Network topology; network nodes capable of hosting SDN controllers and/or Orchestrators; for each network node, the network area it belongs to; disaster prone regions and their risks; maximum switch-to-controller and controller-to-orchestrator delays; switch management capacity of controllers.
- **Output:** Where to place controllers and orchestrator(s), and how to route switch-to-controller and controller-to-orchestrator connections.
- **Objective:** First: Minimize the risk of control plane disruption (and switch- to-controller communication failure) due to disasters. Second: Once the first is guaranteed, minimize resource utilization.

Mathematical Formulation

1) Input parameters:

- $G(N_t, E)$: network topology, where E is the set of directed links, and N_t is the set of nodes with area information;
- F: set of nodes capable of hosting controllers and/or orchestrators. We consider that each node in this set can host however many orchestrators and/or controllers necessary;
- $T = \{t | t \in \{0, 1, 2, ...\}\}$: set of areas in the network;
- $D_{ij} \in \{0, 1\}$: 1 if node i is within reach of node j, under maximum switch-to-controller latency;
- $L_{ij} \in \{0, 1\}$: 1 if node i is within reach of node j, under maximum controller-to-orchestrator latency;
- P_{ij}: set of possible paths from i to j within latency limits;
- $Y = \{y | y =< E_y, R_y >\}$: set of disasters, E_y links they affect, and the risk R_y of y occurring;
- $U_p^y \in \{0, 1\}$: 1 if path p survives disaster Y;
- $B \in \{1, 2, 3, ...\}$: how many switches each controller can manage;
- $K \in \{1, 2, 3, ...\}$: minimum number of controllers that must be reachable from any switch;
- $N \in \{1, 2, 3, ...\}$: minimum number of paths between a switch and its controller;
- $Q \in \{1, 2, 3, ...\}$: minimum number of paths between a controller and the orchestrators;
- $D \in \{0, 1, 2, ...\}$: number of backup orchestrators that should be deployed;
- $V_i^y \in \{0, 1\}$: 1 if node i survives disaster y. If a disaster completely isolates a node from the network (even though it may remain working), we consider that such node does not survive the disaster.

Mathematical Formulation

2) Variables:

- $c_f^t \in \{0, 1\}$: 1 if controller of domain t is located in node f;
- $o_f^t \in \{0, 1\}$: 1 if orchestrator is located in node f;
- $a_{if}^t \in \{0, 1\}$: 1 if switch i is assigned to controller f;
- $b_{if}^t \in \{0, 1\}$: 1 if all controller in i are assigned to orchestrator f;
- $j_i \in \{0, 1\}$: 1 if any controller is active in node i;
- $h_{if}^{tp} \in \{0, 1\}$: 1 if path p is used for communication between switch i and controller f, both of domain t;
- $w_{if}^p \in \{0, 1\}$: 1 if path p is used for communication between controller i and orchestrator f;
- $z_{if}^y \in \{0, 1\}$: 1 if at least one path from controller i to orchestrator f survives disaster y;
- $s_f^y \in \{0, 1\}$: 1 if all controllers that survive disaster y remain connected to orchestrator f.

UCDAVIS

Mathematical Formulation

3) Objective:

$$minimize\,(\omega + \gamma + \delta + \theta + \lambda)$$

where

$$\omega = \Omega \cdot \sum_{y \in Y} \sum_{i \in N} \sum_{f \in F} \sum_{p \in P_{if}} R_y \cdot (1 - U_p^y) \cdot w_{if}^p$$

is the average number of controller-to-orchestrator paths that fail due to disasters, weighted by the risk of each disaster occurring,

$$\gamma = \Gamma \cdot \sum_{y \in Y} \sum_{i \in F} \sum_{f \in F} \sum_{p \in P_{if}} \sum_{t \in T} R_y \cdot (1 - U_p^y) \cdot h_{if}^{tp}$$

is the average number of switch-to-controller paths that fail due to disasters, weighted by the risk of each disaster occurring,

$$\delta = \Delta \cdot \sum_{i \in F} j_i$$

is the total number of nodes that host controllers (for all areas in the network),

$$\theta = \Theta \cdot \sum_{y \in Y} \sum_{i \in F} R_y \cdot (1 - V_i^y) \cdot j_i$$

is the average number of controllers that fail due to disasters, weighted by the risk of each disaster occurring, and

$$\lambda = \Lambda \cdot \sum_{i \in N} \sum_{j \in N} \sum_{p \in P_{ij}} len(p) \cdot \left(w_{if}^p + \sum_{t \in T} h_{if}^{tp} \right)$$

is the resource utilization measured by the number of links used for switch-to-controller and controller-to-orchestrator communication ($len(p)$ is the number of links in path p).

In the functions above Ω, Γ, Δ, Θ, and Λ are sufficiently large constants whose values may affect the overall solution of the problem. In this paper, we first focus on making the controller to orchestrator communication disaster-resilient (thus, $\Omega \gg \Gamma$); then, the switch-to-controller (thus, $\Gamma \gg \Delta$); then, on minimizing both the number of controllers deployed and the number of controllers that might be damaged by disasters (thus, $\Delta = \Theta$); and, finally, on minimizing resource consumption (thus, $\Theta \gg \Lambda$).

UCDAVIS

Mathematical Formulation

4) Constraints:

a) Reachability:

$$\sum_{f \in F} c_f^t \cdot D_{if} \leq K, \quad \forall i \in N, \forall t \in T$$

enforces that at least K controllers will be reachable from any node.

b) Binarization of c_f^t:

$$c_f^t \leq \sum_{i \in N} a_{if}^t, \quad \forall f \in F, \forall t \in T$$

$$c_f^t \geq \frac{\sum_{i \in N} a_{if}^t}{M}, \quad \forall f \in F, \forall t \in T, M \; large$$

enforces that a controller will be deployed in node f if at least one node i of area t is set to be controlled by such controller.

c) Controller Capacity:

$$\sum_{i \in N} a_{if}^t + 1 \leq B, \quad \forall f \in F, \forall t \in T$$

enforces that at most B nodes will be controlled by any controller.

d) Switch Assignment:

$$\sum_{i \in N} a_{if}^t = 1, \quad \forall f \in F, \forall t \in T$$

enforces that every node must be controlled by exactly one controller.

e) Node area:

$$a_{if}^t = 0 \quad \forall i, f \in N, t \in T, area(i) \neq t$$

$$c_f^t = 0 \quad \forall i, f \in N, t \in T, area(i) \neq t$$

enforces that a node can only be assigned to controller of the same area and that controller can only be deployed in nodes of the area they belong to ($area(i)$ is the area of node i).

f) Switch to Controller Latency:

$$a_{if}^t \leq d_{if}, \quad \forall i \in N, f \in F, \forall t \in T$$

enforces that nodes can only be assigned to controller within their latency limit.

g) Switch to Controller Check:

$$a_{if}^t = c_f^t, \quad \forall i, f \in N, \forall t \in T, i = f$$

$$a_{if}^t + c_i^t \leq 0, \quad \forall i, f \in N, \forall t \in T, i \neq f$$

$$a_{if}^t \leq c_f^t, \quad \forall i, f \in N, \forall t \in T, i \neq f$$

enforces that a node must be assigned to the local controller if it exists (and to a non-local controller if the node does not host a controller locally).

UCDAVIS

Mathematical Formulation

h) Switch to Controller Path:

$$\sum_{p \in P_{if}} h_{if}^{tp} \geq N \cdot a_{if}^{t}, \quad \forall i, f \in N, \forall t \in T, i \neq f$$

enforces that at least N paths must be setup between each node and its controller. Similar to traditional protection against element failures (when $N > 1$).

i) Binarization of o_f:

$$o_f \leq \sum_{i \in N} b_{if}, \quad \forall f \in F$$

$$o_f \geq \frac{\sum_{i \in N} b_{if}}{M}, \quad \forall f \in F, \forall t \in T, M \ large$$

enforces that an orchestrator will be deployed in node f if at least one controller i (of any area) is set to be managed by such orchestrator.

j) Controller Assignment:

$$\sum_{f \in F} b_{if} = (1 + D) \cdot j_i, \quad \forall i \in N$$

enforces that every controller must be managed by exactly $1 + D$ (primary and backups) orchestrator(s).

k) Binarization of j_i:

$$j_i \leq \sum_{t \in T} c_i^t, \quad \forall i \in N$$

$$j_i \geq \frac{\sum_{t \in T} c_i^t}{M}, \quad \forall f \in F, \forall t \in T, M \ large$$

enforces that j_i is one if at least one controller of any area t is placed on node i.

l) Controller to Orchestrator Latency:

$$b_{if} \leq L_{if}, \quad \forall i, f \in N$$

enforces that all orchestrators must be within the latency limits of every controller.

m) Controller to Orchestrator Check:

$$b_{if} \leq o_f, \quad \forall i, f \in F, \forall t \in T$$

$$b_{if} \leq j_i, \quad \forall i, f \in F, \forall t \in T$$

enforces that controllers must be managed by nodes that host orchestrators (and vice-versa). Since all controller are simultaneously assigned to all orchestrators, it is not necessary to check whether the controller local to a node is connected to the orchestrator local to the same node (i.e., if such node contains both an orchestrator and a controller).

UC DAVIS

Mathematical Formulation

n) Controller to Orchestrator Path:

$$\sum_{p \in P_{if}} w_{if}^{p} \geq Q \cdot b_{if}, \quad \forall i, f \in F$$

enforces that at least Q paths must be setup between each controller and each orchestrator. Similar to traditional protection against element failures (when $M > 1$).

o) Binarization of z_{if}^{y}:

$$z_{if}^{y} \leq V_i^y \cdot o_f, \quad \forall i, f \in N, \forall y \in Y$$

$$z_{if}^{y} \leq V_i^y \cdot j_i, \quad \forall i, f \in N, \forall y \in Y$$

$$z_{if}^{y} \leq \sum_{p \in P_{if}} w_{if}^{p} \cdot U_p^y, \quad \forall i, f \in F, \forall y \in Y$$

$$z_{if}^{y} \geq \frac{\sum_{p \in P_{if}} w_{if}^{p} \cdot U_p^y}{M}, \quad \forall i, f \in F, \forall y \in Y, M\ large$$

enforces that z_{if}^{y} is one if at least one path controller to orchestrator survives disaster y (i.e., if both controller and orchestrator also survive such disaster).

o) Binarization of s_f^y

$$\sum_{i \in N} z_{if}^{y} \geq \sum_{i \in N} V_i^y \cdot j_i - M\left(1 - s_f^y\right), \quad \forall f \in N, \forall y \in Y$$

$$\sum_{i \in N} V_i^y \cdot j_i \geq \sum_{i \in N} z_{if}^{y} - M \cdot s_f^y + 1, \quad \forall f \in N, \forall y \in Y$$

note that $\sum_{i \in N} z_{if}^{y}$ is the total amount of controller-hosting nodes that survive disaster y and remain connected to orchestrator f; and $\sum_{i \in N} V_i^y \cdot j_i$ is the total amount of controller-hosting nodes that survive disaster y. Thus, the second is an upper bound of the first. This constraint enforces that s_f^y is one only if all controllers that survive disaster y remain connected to orchestrator f.

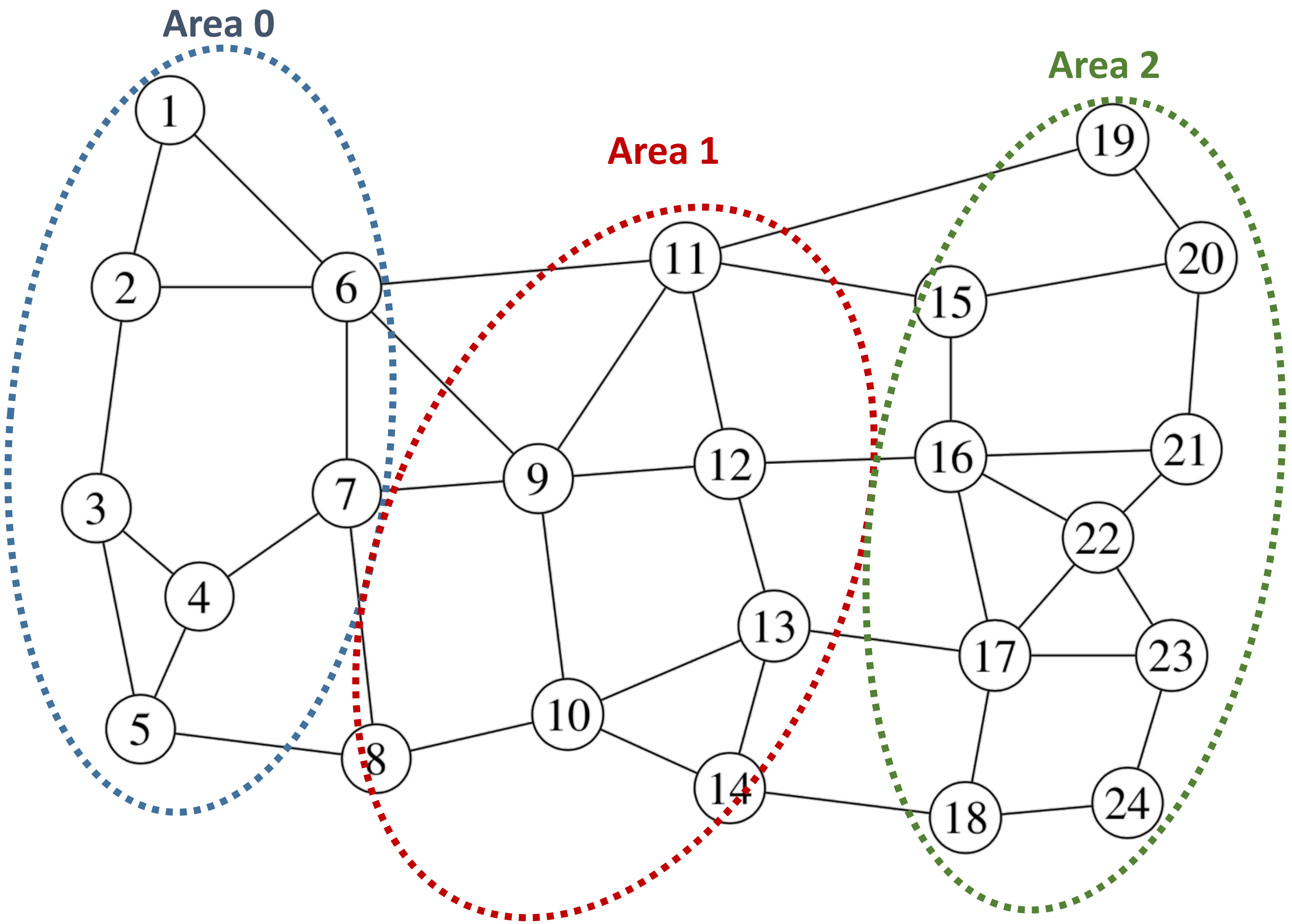
p) Disaster-survivability of controller-to-orchestrator relationship:

$$\sum_{f \in F} s_f^y \geq 1, \quad \forall y \in Y$$

enforces that, after any disaster y, all surviving controllers remain connected to at least one surviving orchestrator. If more than one orchestrator survives, at least one of them will remain connected to every surviving controller.

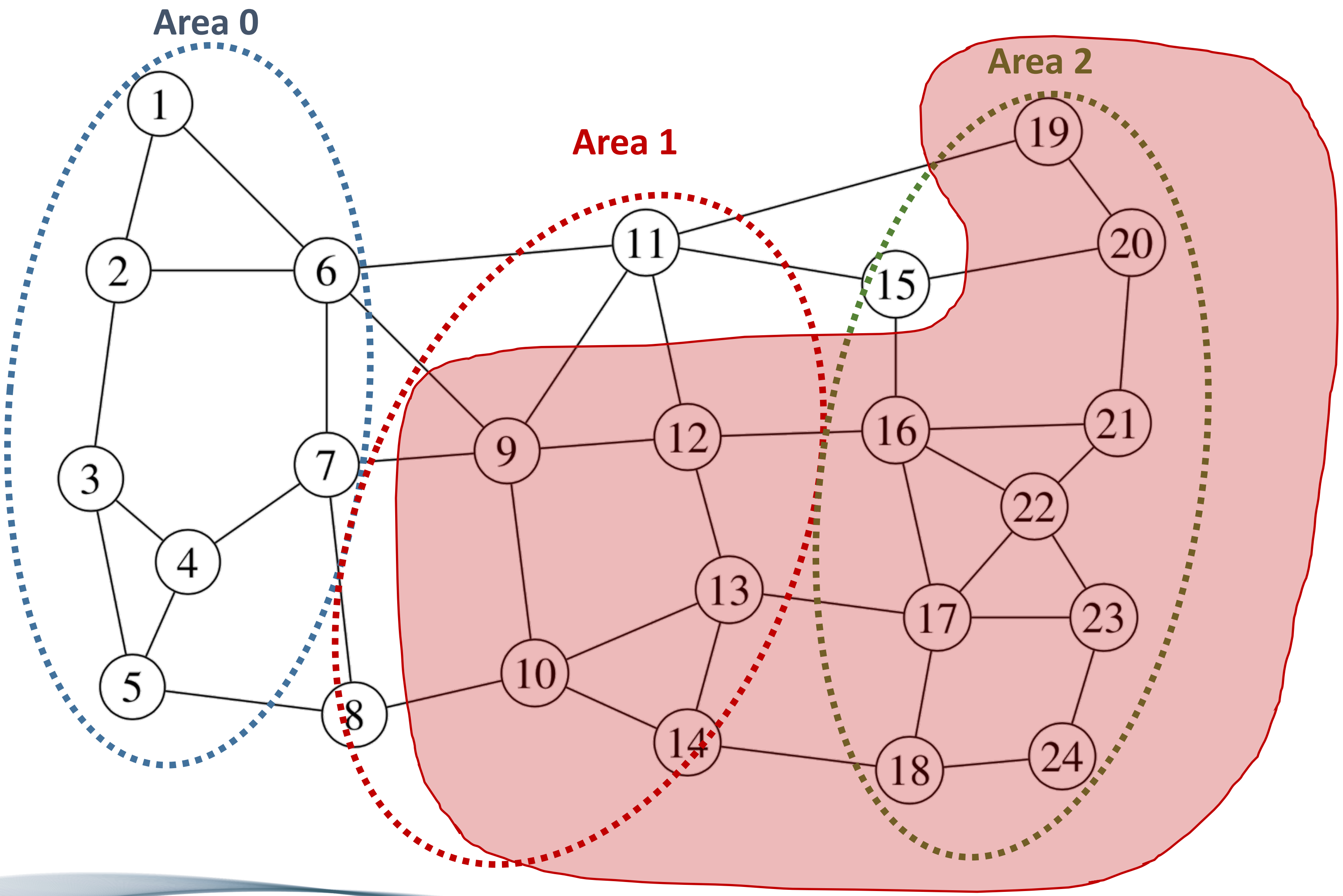
UCDAVIS

Illustrative Example – Topology and Risky Regions



UCDAVIS

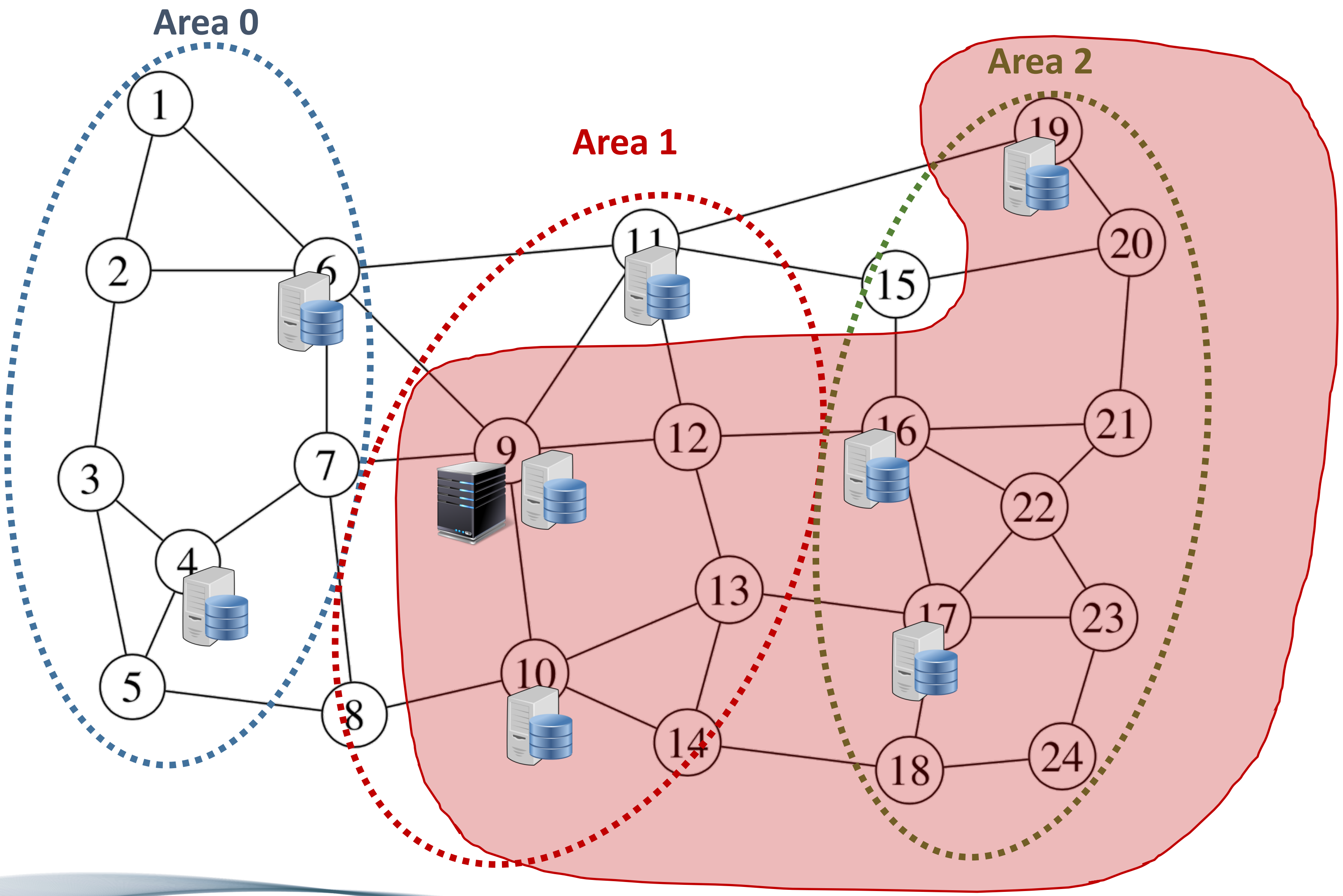
Illustrative Example – Topology and Risky Regions



Single disaster-prone region
Risk = 0.5
No controller capacity limit
Paths switch-to-control. = 1
Paths control-to-orches. = 1
No backup orchestrator
One controller reachable per switch
No latency limits

UCDAVIS

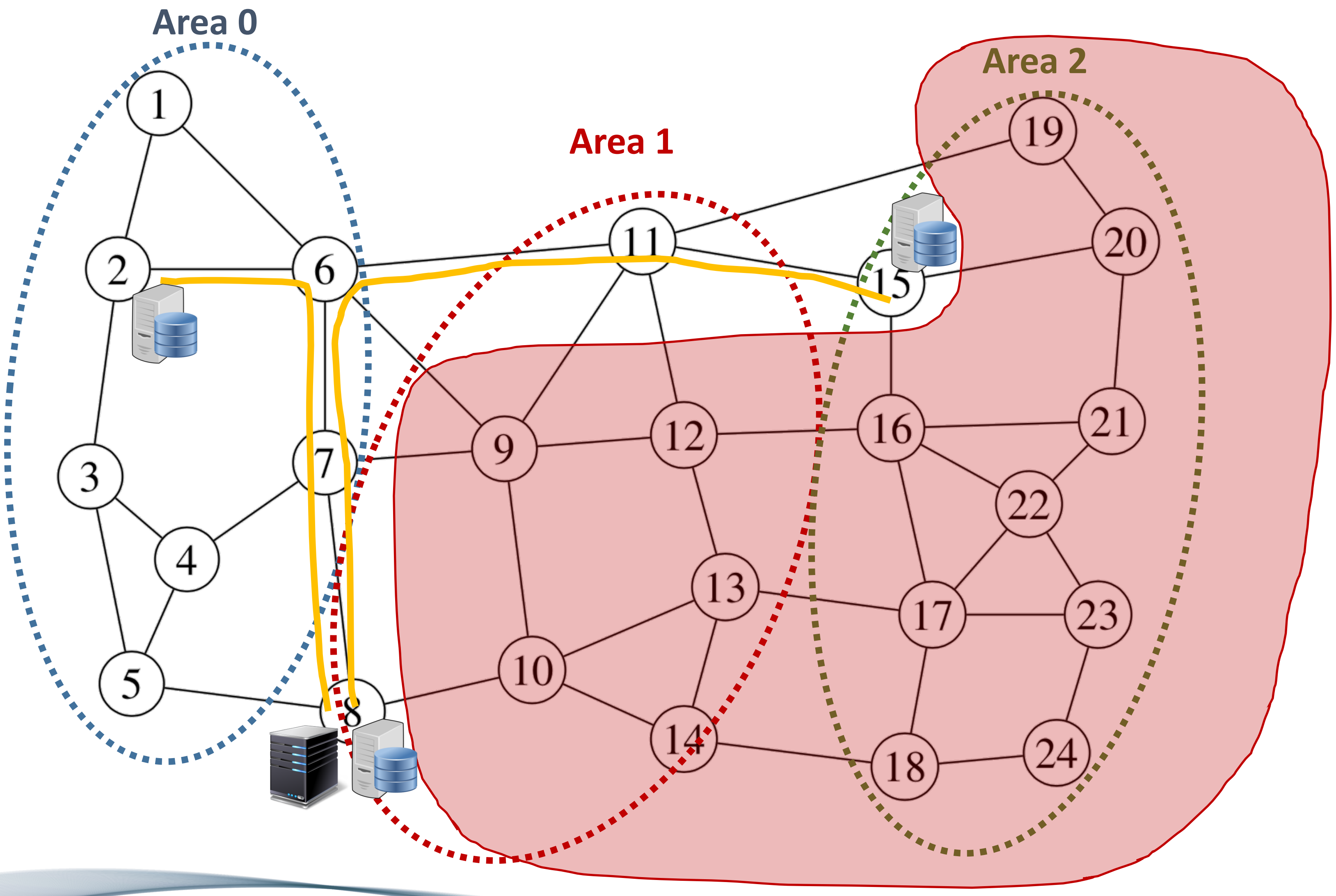
Illustrative Example – Minimize Resource



Without orchestrator survivability constraints.

No latency constraint.

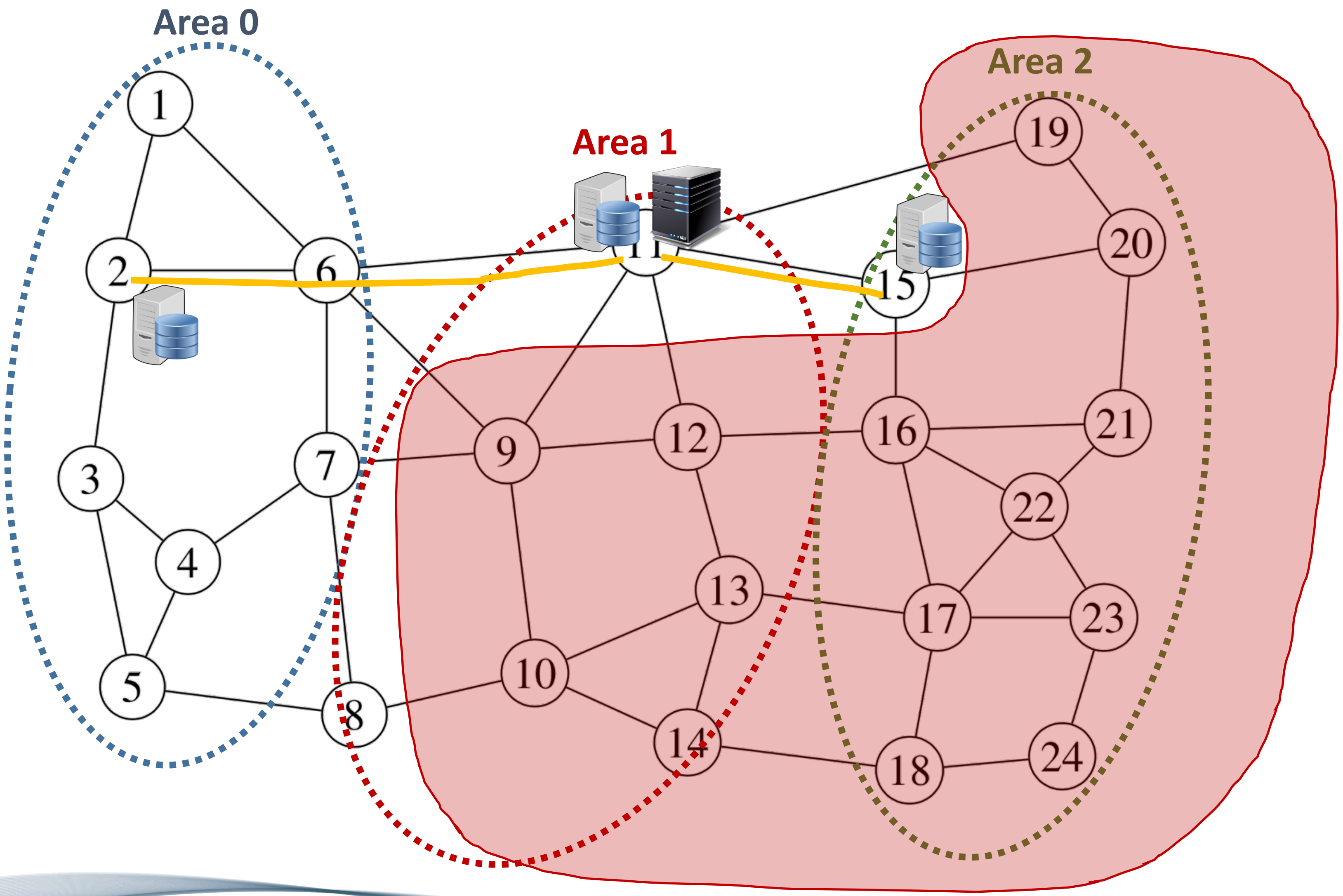
Illustrative Example 1



Single disaster-prone region
Risk = 0.5
No controller capacity limit
Paths switch-to-control. = 1
Paths control-to-orches. = 1
No backup orchestrator
One controller reachable per switch
No latency limits

UCDAVIS

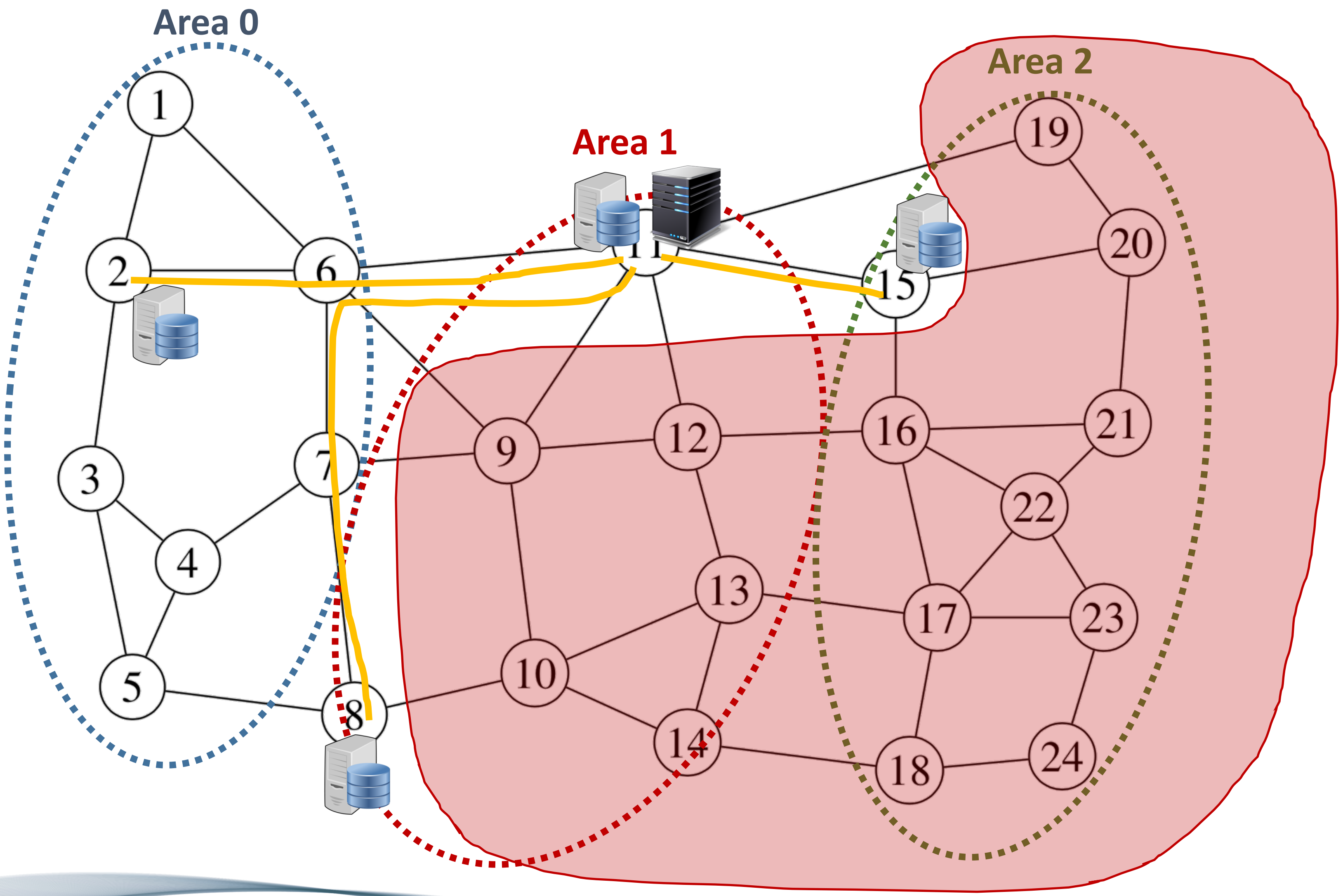
Illustrative Example 2



Single disaster-prone region
Risk = 0.5
No controller capacity limit
Paths switch-to-control. = 1
Paths control-to-orches. = 1
No backup orchestrator
One controller reachable per switch
Switch-to-control. < 4k miles
Control-to-orches. < 4k miles

UCDAVIS

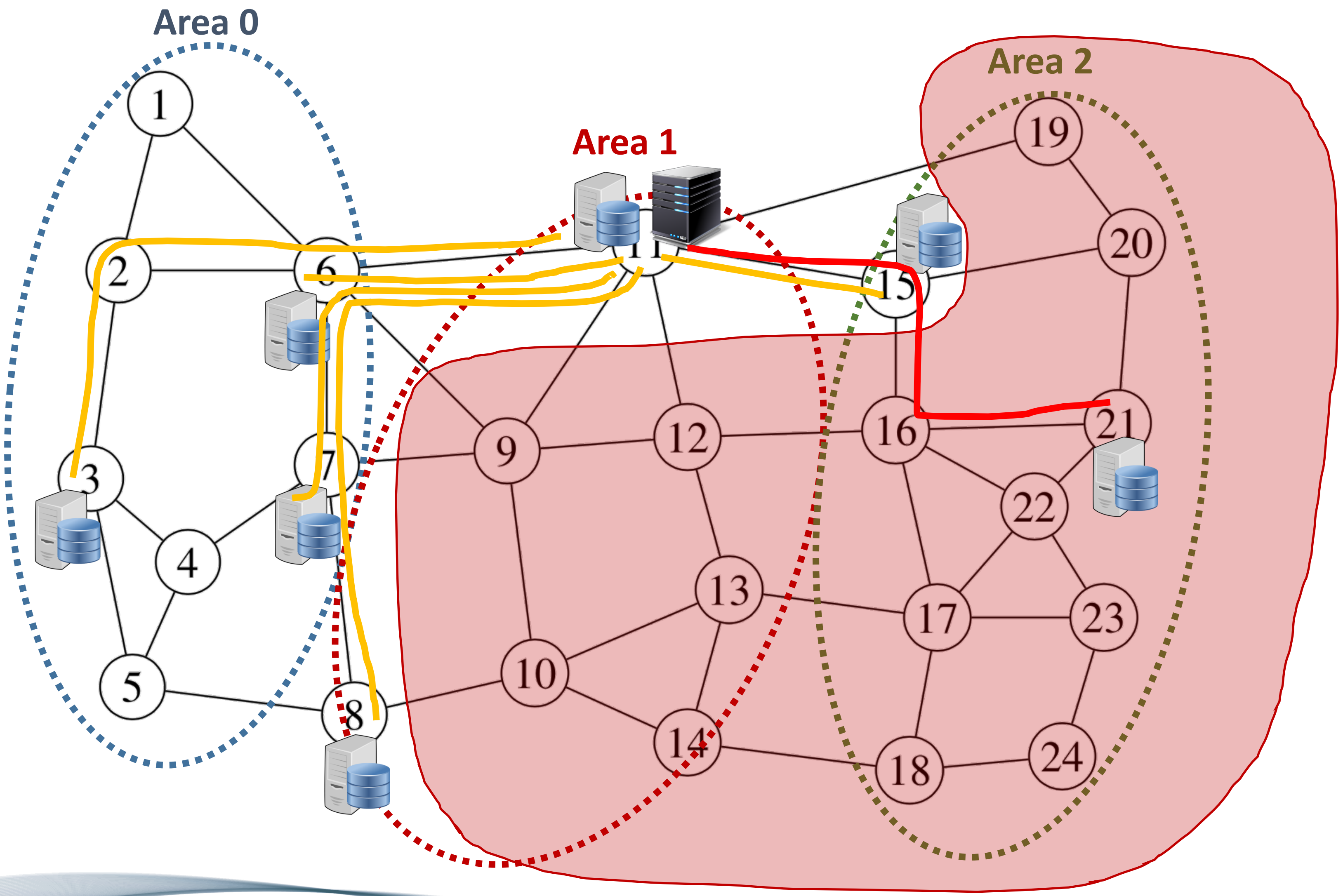
Illustrative Example 3



Single disaster-prone region
Risk = 0.5
No controller capacity limit
Paths switch-to-control. = 1
Paths control-to-orches. = 1
No backup orchestrator
One controller reachable per switch
Switch-to-control. < 3k miles
Control-to-orches. < 4k miles

UCDAVIS

Illustrative Example 4



Single disaster-prone region

Risk = 0.5

No controller capacity limit

Paths switch-to-control. = 1

Paths control-to-orches. = 1

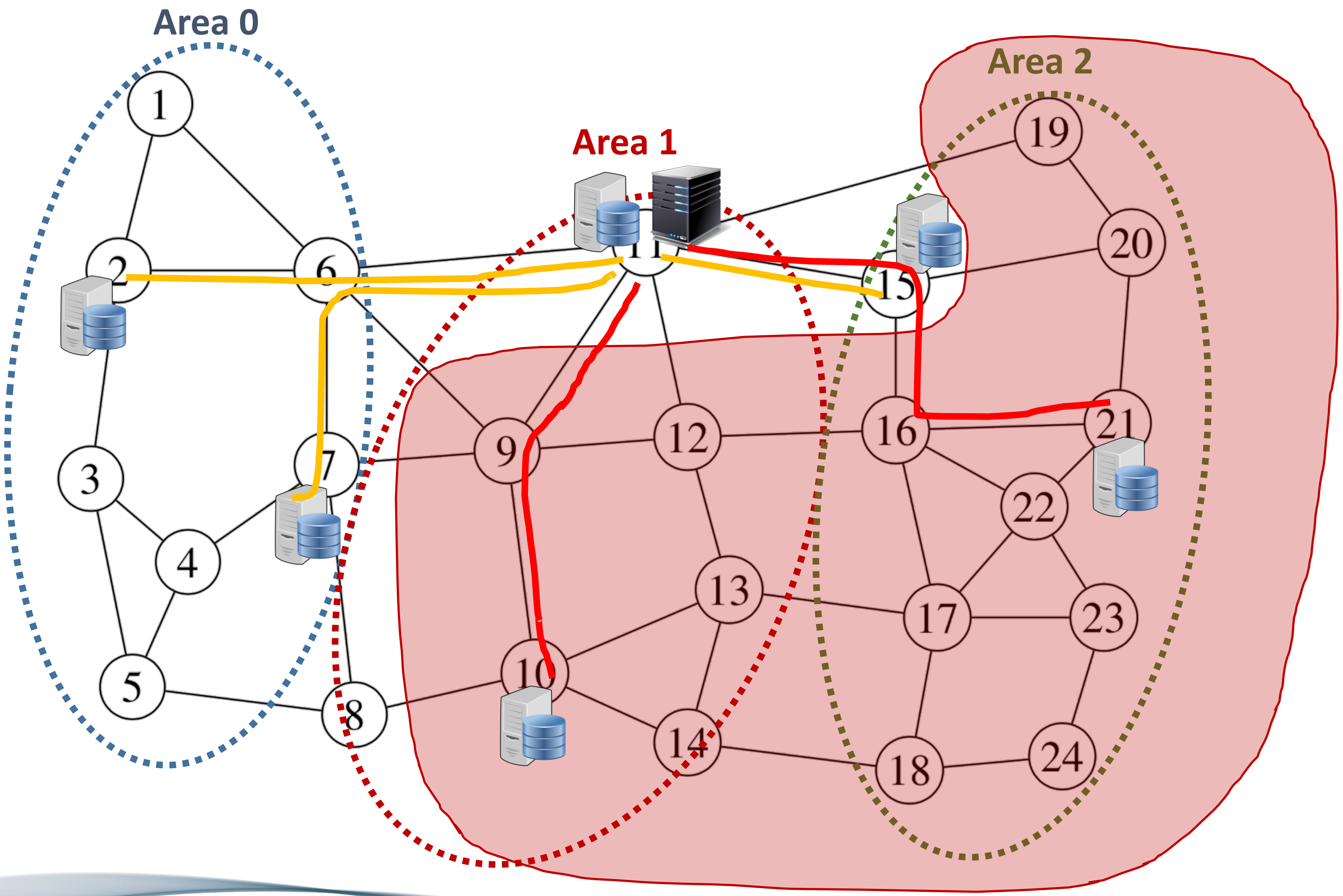
No backup orchestrator

One controller reachable per switch

Switch-to-control. < 2k miles

Control-to-orches. < 4k miles

Illustrative Example 5



Single disaster-prone region

Risk = 0.5

No controller capacity limit

Paths switch-to-control. = 1

Paths control-to-orches. = 1

No backup orchestrator

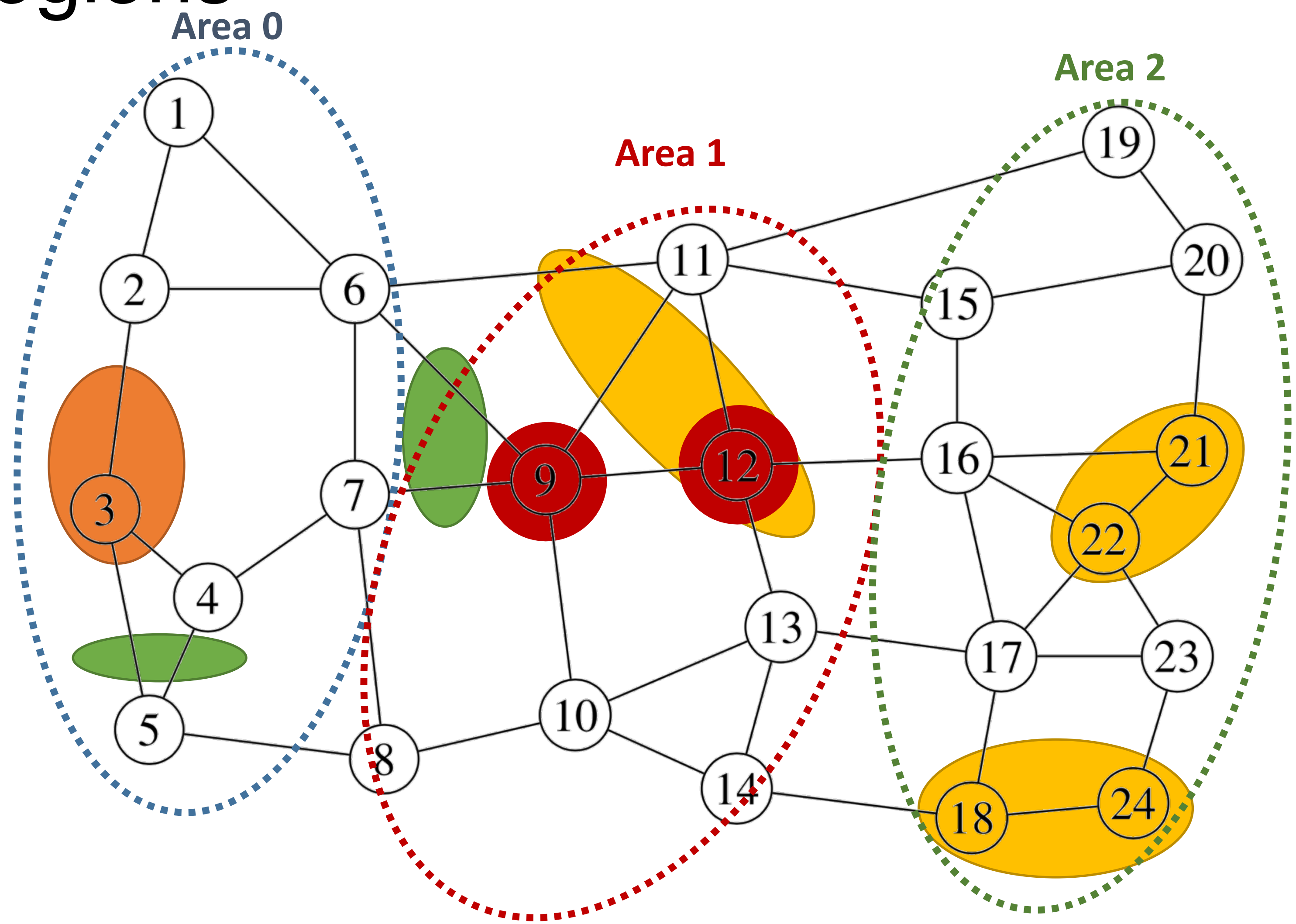
One controller reachable per switch

Switch-to-control. < 2k miles

Control-to-orches. < 3k miles

UCDAVIS

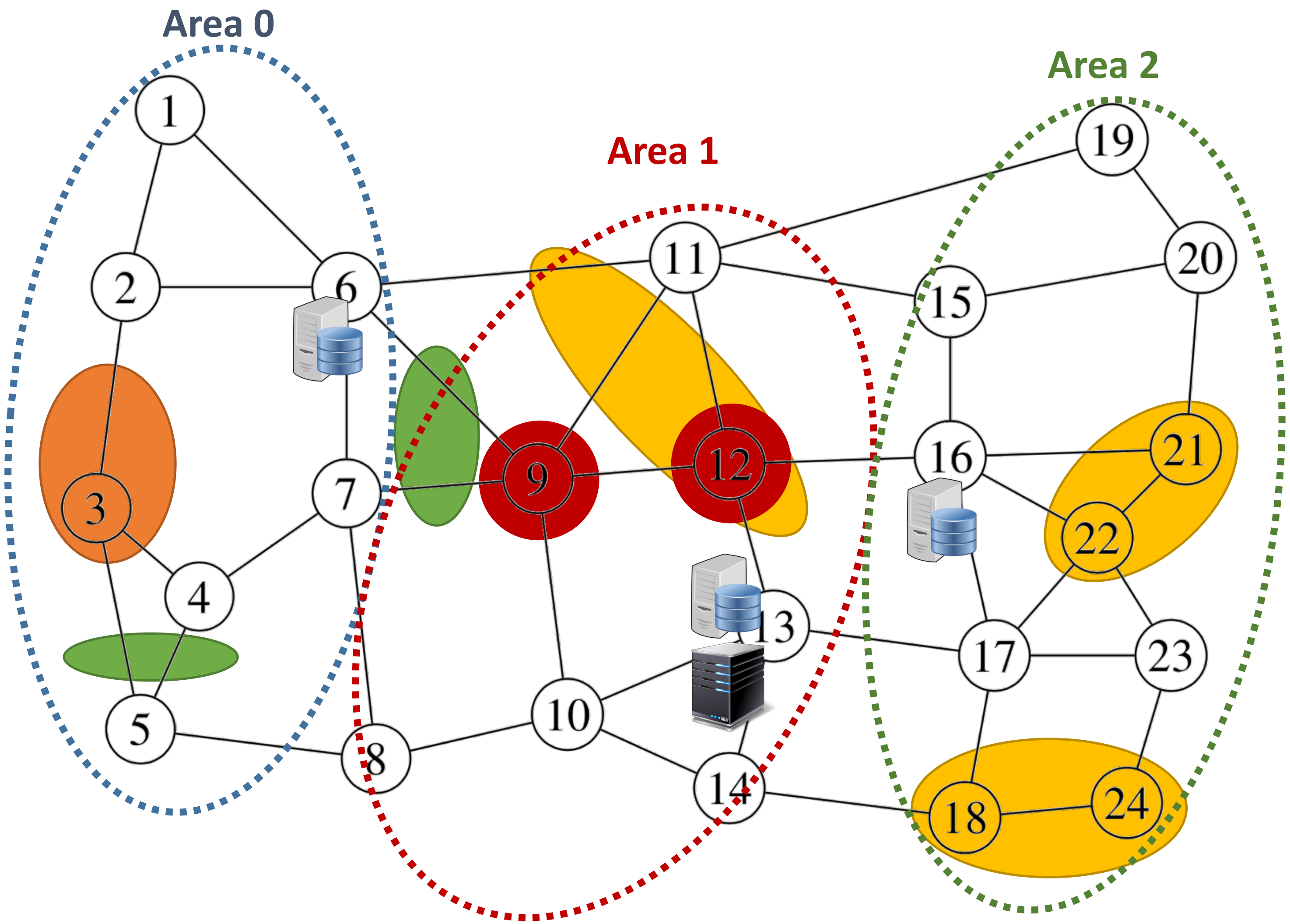
Illustrative Example 6 – Topology and Risky Regions



Color	Probability
Green	0.02
Yellow	0.06
Orange	0.10
Red	0.20

UCDAVIS

Illustrative Example 6 – ILP Result



No controller capacity limit
Paths switch-to-control. = 1
Paths control-to-orches. = 1
No backup orchestrator
One controller reachable per switch
No latency limits

Results

- Average path length?
- Total number of disconnected switches for all disaster scenarios?
- Resource utilization (avg path length)?
- Sensitivity analysis: for large controller capacity and large latency tolerance, what factors influence number of controllers?
- Different topologies (possibly larger)?
- Discrete event simulation comparing placement versus other?