Network-Attack-Resilient Intrusion-Tolerant SCADA for the Power Grid

Amy Babay, Thomas Tantillo, Trevor Aron,
Marco Platania, and Yair Amir

Johns Hopkins University, AT&T Labs, Spread Concepts LLC

JOHNS HOPKINS UNIVERSITY

Distributed Systems and Networks Lab
www.dsn.jhu.edu

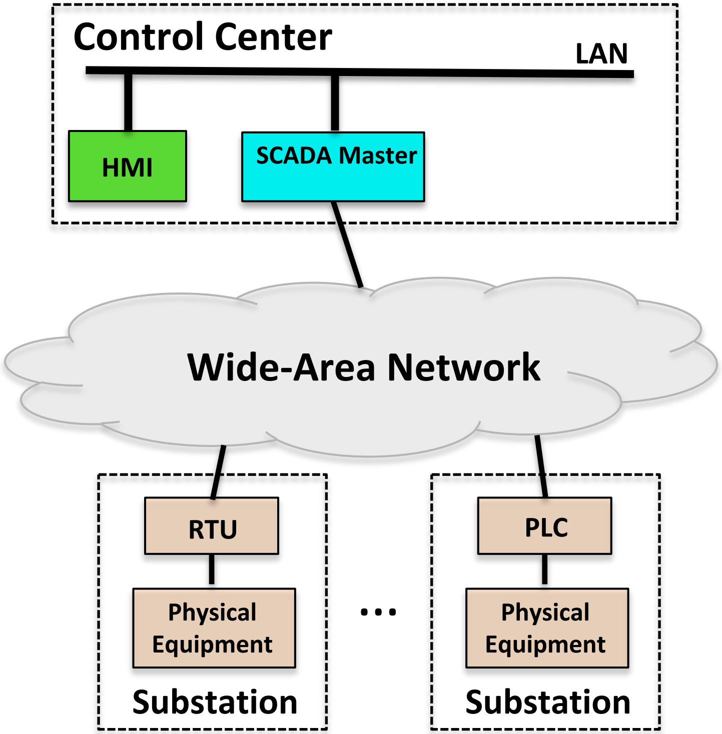
Importance of SCADA Systems

- Supervisory Control and Data Acquisition (SCADA) systems form the backbone of critical infrastructure services
  - **Power grid**, water supply, waste management

- To preserve control and monitoring capabilities, SCADA systems must be constantly available and run at their expected level of performance (able to react within 100-200ms)
- SCADA system failures and downtime can cause catastrophic consequences, such as equipment damage, blackouts, and human casualties

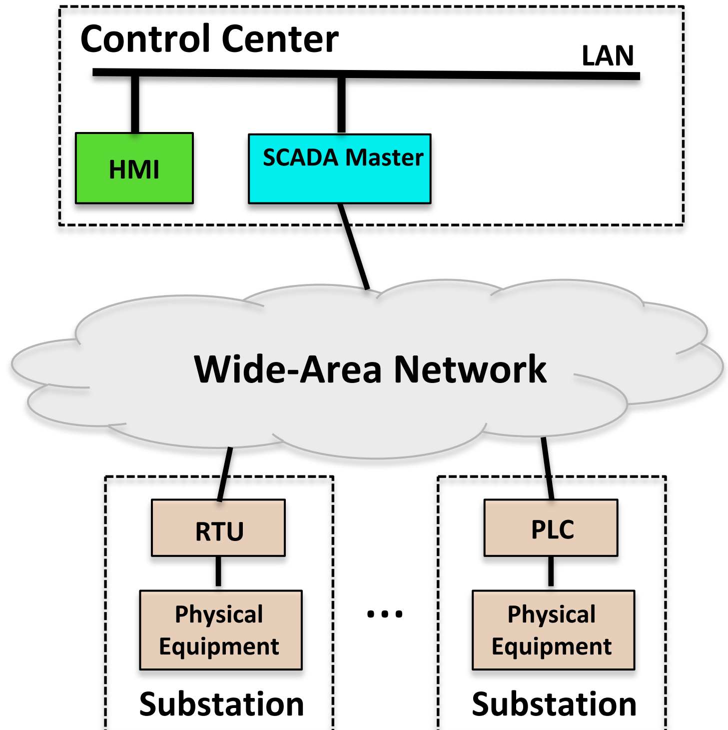
Basic SCADA Architecture

- Remote Terminal Units (RTUs) and Programmable Logic Controllers (PLCs): Control physical equipment in power substations
- SCADA Master: Central control server, maintains current state of the system (based on updates from RTUs/PLCs) and issues supervisory commands
- Human Machine Interface (HMI): provide graphical displays for operators to interact with the system



Basic SCADA Architecture

1



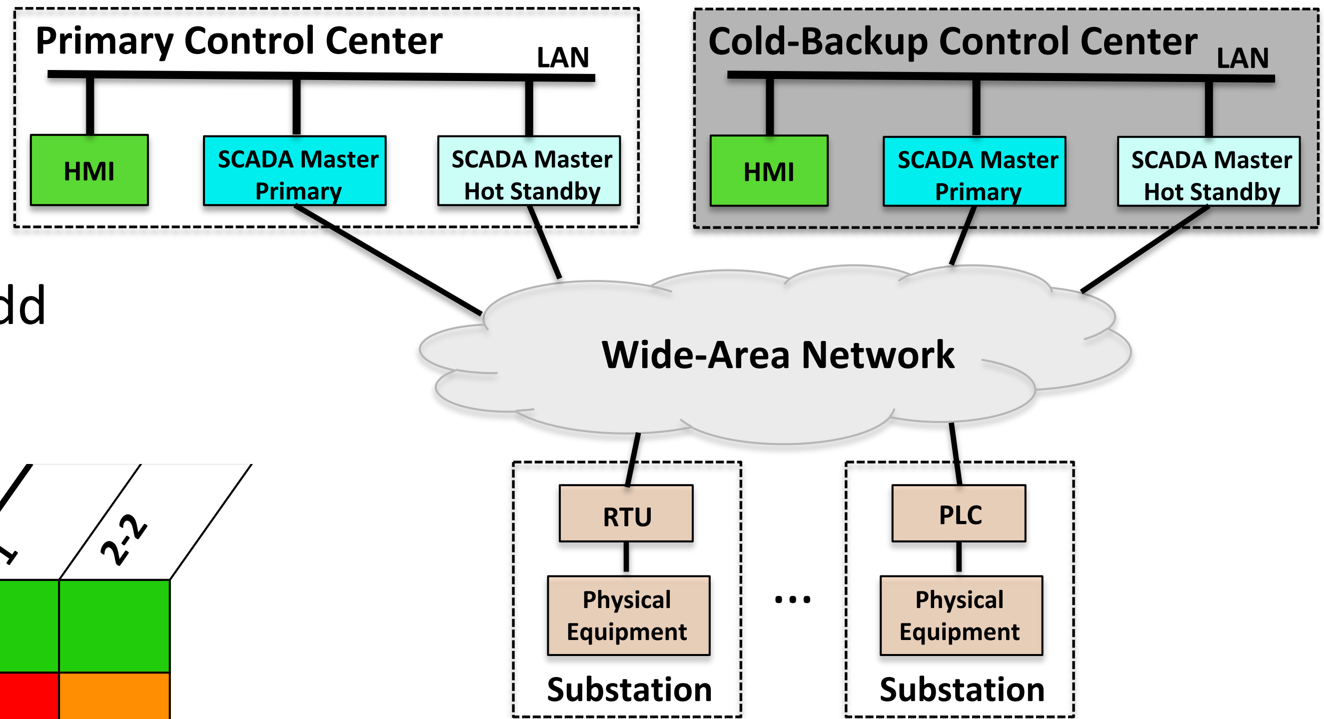
For illustration only – this is a very fragile architecture

	1
All Correct	
Failed Site	
SCADA Master Crash	

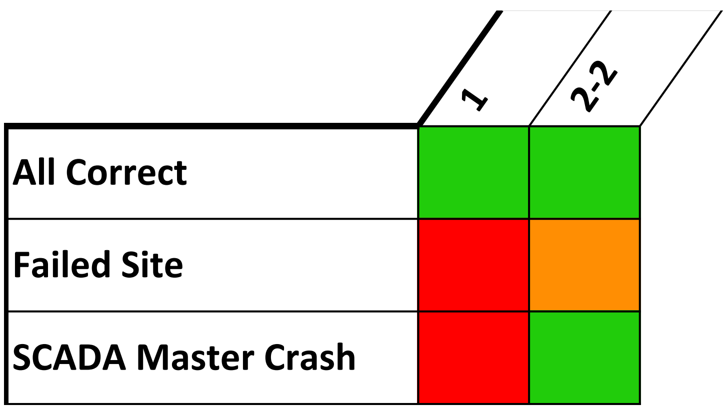
- Bounded Delay
- Bounded Delay, except when rejuvenating any correct replica
- Eventual Progress – Human in the loop. Potentially powering up cold backup control center
- Eventual Progress – No bound. Network has to heal, crash has to be repaired, and/or intrusion needs to be cleansed
- Incorrect System

Modern SCADA Systems

2-2



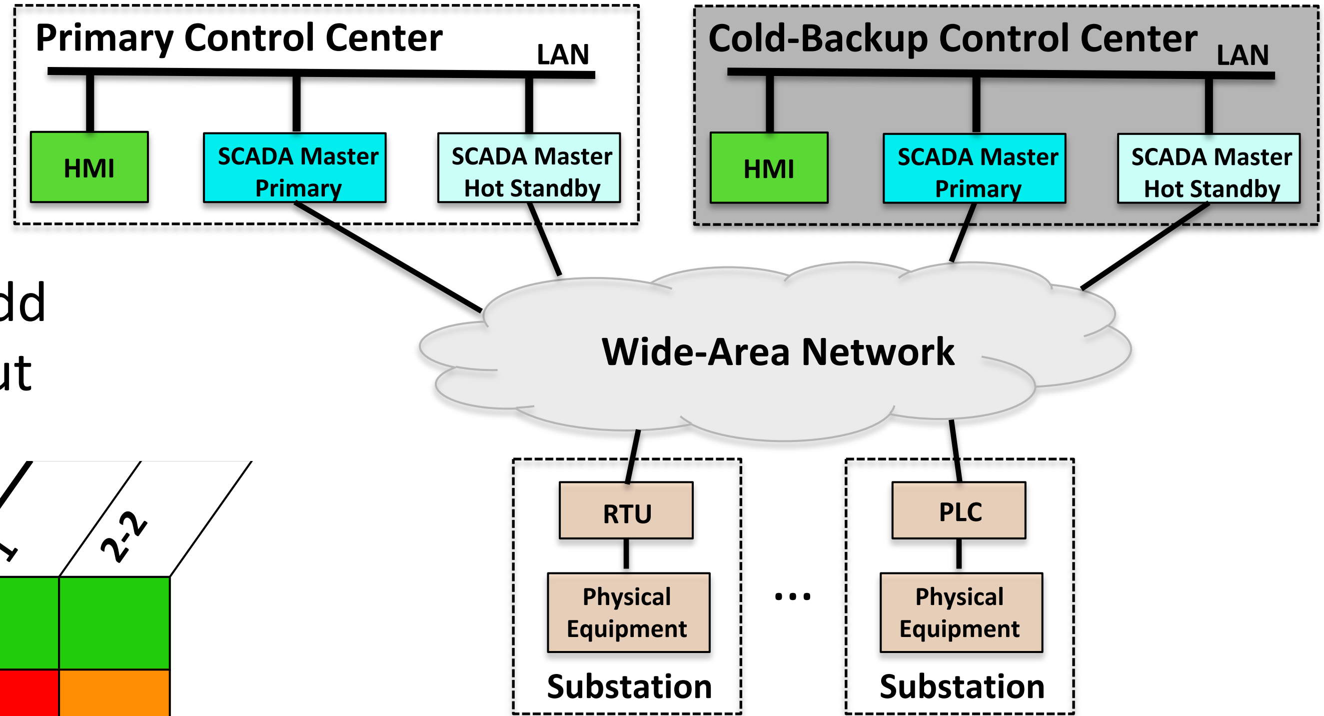
Modern systems add
fault-tolerance



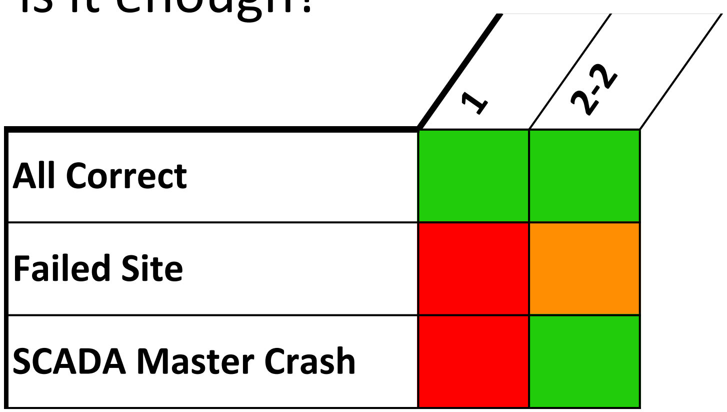
- Bounded Delay
- Bounded Delay, except when rejuvenating any correct replica
- Eventual Progress – Human in the loop. Potentially powering up cold backup control center
- Eventual Progress – No bound. Network has to heal, crash has to be repaired, and/or intrusion needs to be cleansed
- Incorrect System

Modern SCADA Systems

2-2



Modern systems add fault-tolerance…but is it enough?



- Bounded Delay
- Bounded Delay, except when rejuvenating any correct replica
- Eventual Progress – Human in the loop. Potentially powering up cold backup control center
- Eventual Progress – No bound. Network has to heal, crash has to be repaired, and/or intrusion needs to be cleansed
- Incorrect System

Hostile Operating Environments

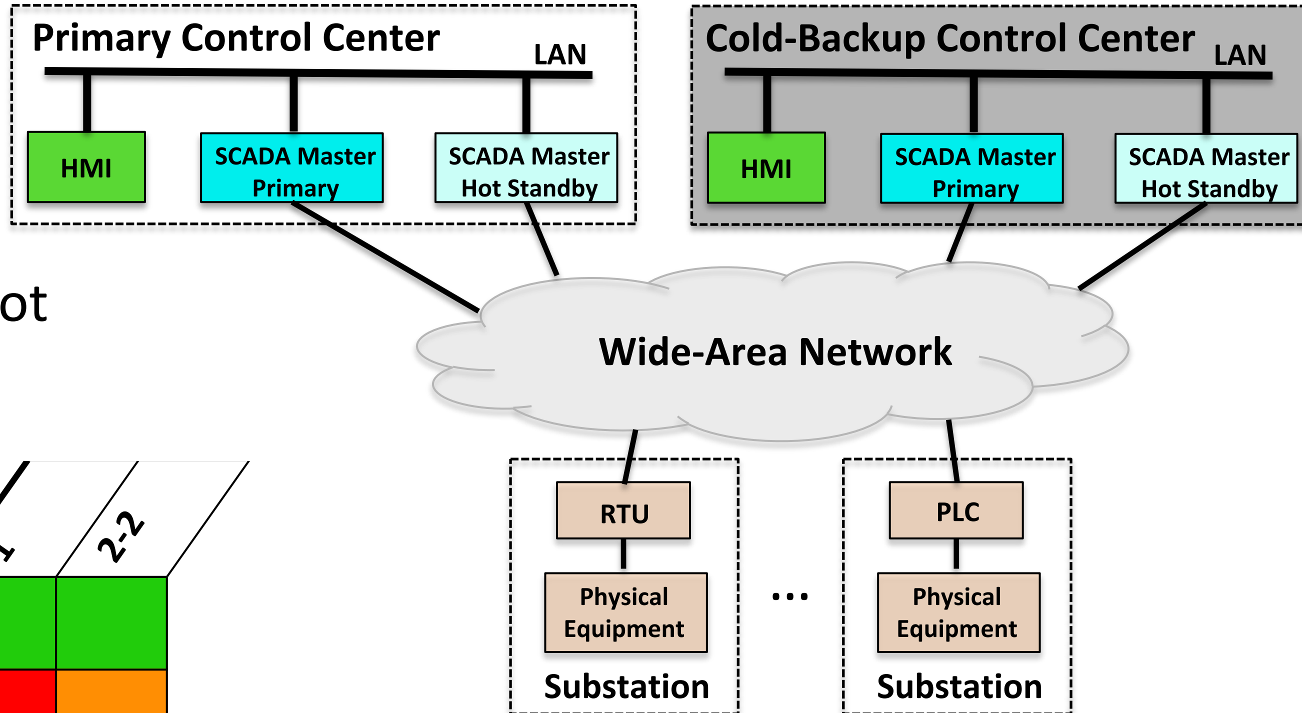
- Traditional SCADA systems ran on proprietary networks
  - Created air gap from outside world and attackers
- Cost benefits and ubiquity of IP networks are driving SCADA to use IP networks
  - Exposes SCADA to hostile environments, removing the air gap
- Raises additional concerns because SCADA systems are:
  - In service for decades
  - Running legacy code with well-known exploits
  - Increasingly becoming a target for attackers
- Attacks: Stuxnet (2010), Ukraine (2015, 2016), and others…

Hostile Environments

- Fault tolerance is no longer sufficient
- System-level attacks
  - Compromised SCADA master can issue malicious control commands and manipulate monitoring information
  - Hot-backup is not effective: compromised primary reports that it is working correctly
- Network-level attacks
  - Sophisticated attacks can take the primary control center offline at the time of the attacker’s choosing
  - Cold-backup inherently incurs downtime (on the order of hours). Not effective when the attacker controls when downtime occurs and can force it to occur repeatedly
  - Hot-backup introduces inconsistency due to “split-brain” problem

Modern SCADA Systems

2-2



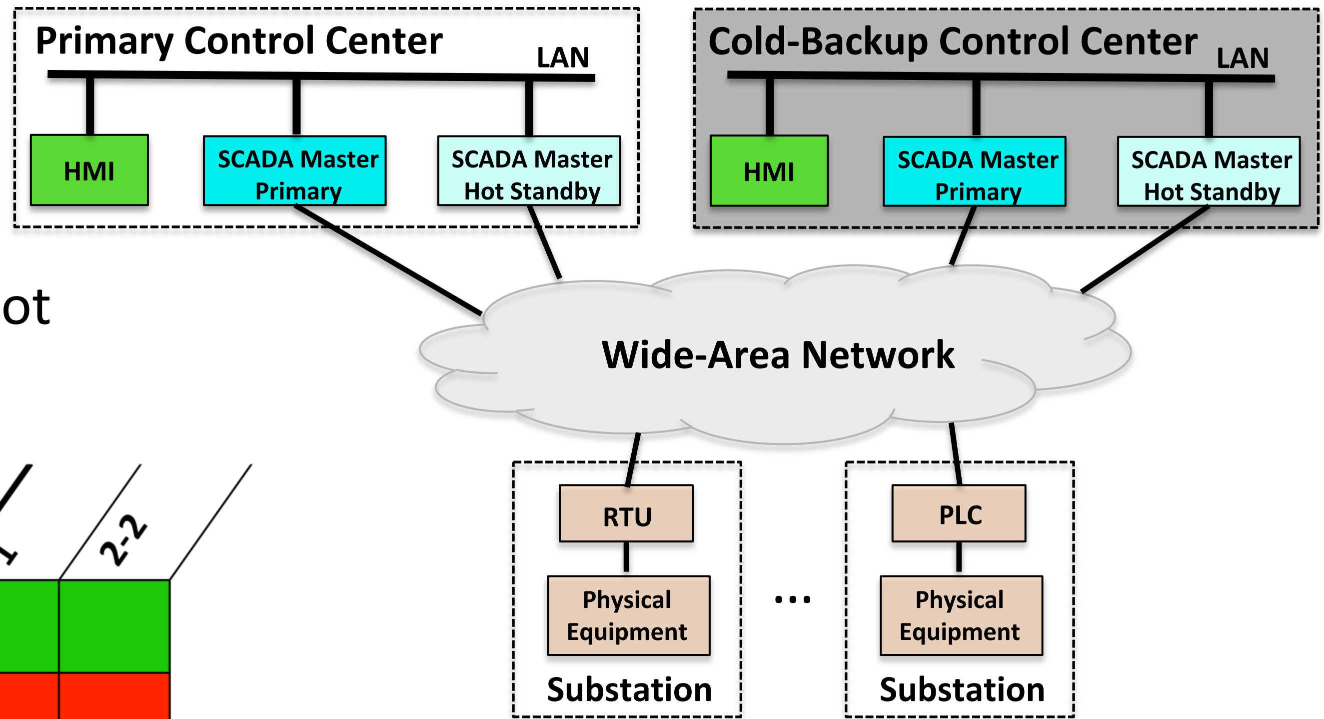
Fault tolerance is not sufficient in hostile environments

	1	2-2
All Correct	Green	Green
✖ Failed Site	Red	Orange
✖ SCADA Master Crash	Red	Green

Green	Bounded Delay
Yellow	Bounded Delay, except when rejuvenating any correct replica
Orange	Eventual Progress – Human in the loop. Potentially powering up cold backup control center
Red	Eventual Progress – No bound. Network has to heal, crash has to be repaired, and/or intrusion needs to be cleansed
Gray	Incorrect System

Modern SCADA Systems

2-2



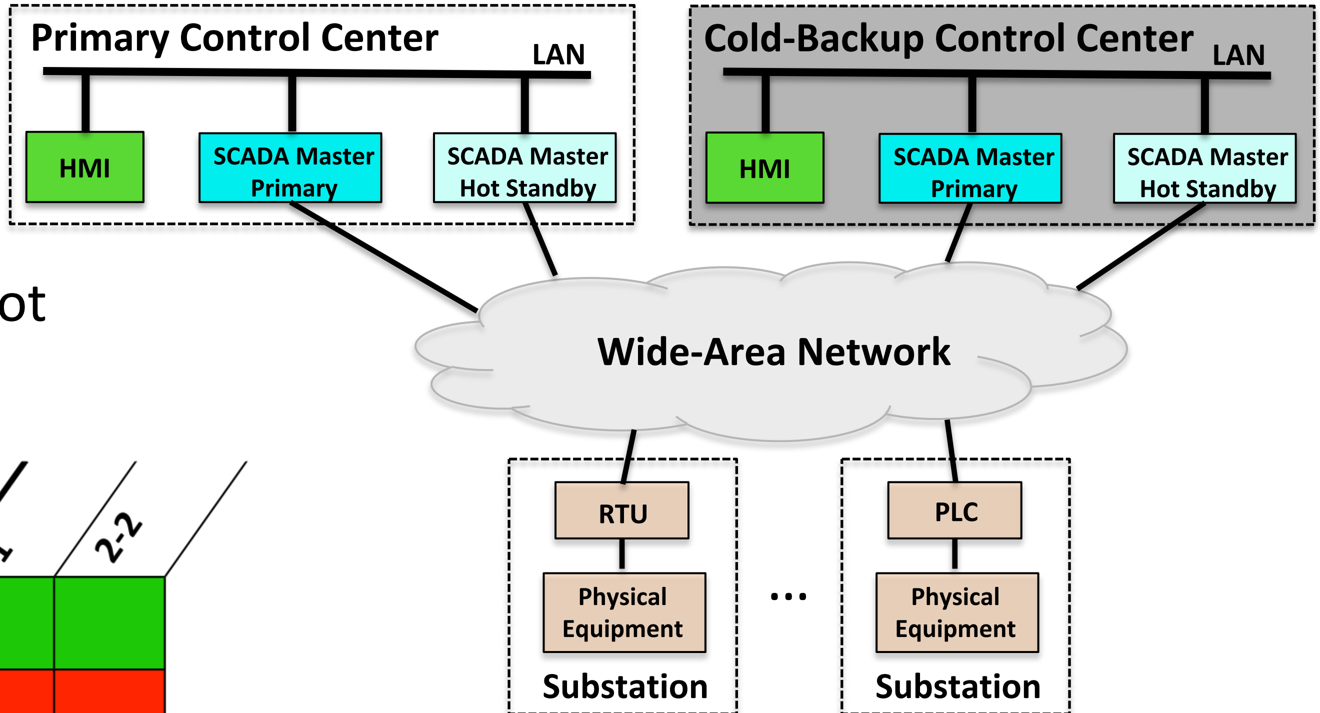
Fault tolerance is not sufficient in hostile environments

	1	2-2
All Correct		
Network Attack		
Intrusion		

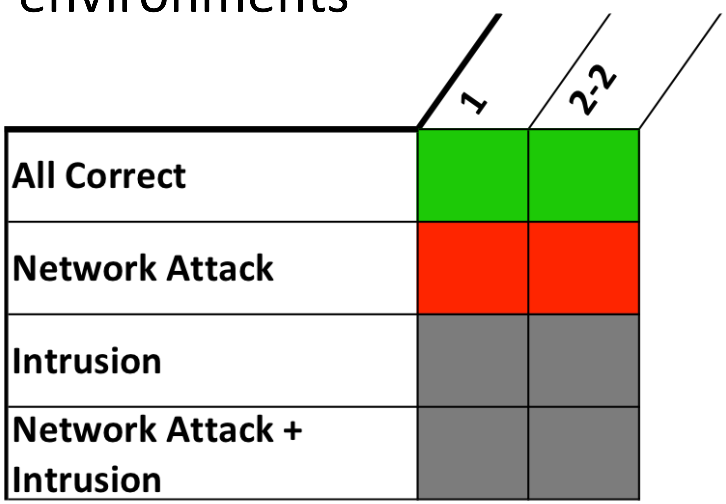
	Bounded Delay
	Bounded Delay, except when rejuvenating any correct replica
	Eventual Progress – Human in the loop. Potentially powering up cold backup control center
	Eventual Progress – No bound. Network has to heal, crash has to be repaired, and/or intrusion needs to be cleansed
	Incorrect System

Modern SCADA Systems

2-2



Fault tolerance is not sufficient in hostile environments



- Bounded Delay
- Bounded Delay, except when rejuvenating any correct replica
- Eventual Progress – Human in the loop. Potentially powering up cold backup control center
- Eventual Progress – No bound. Network has to heal, crash has to be repaired, and/or intrusion needs to be cleansed
- Incorrect System

Our Contribution

- The first SCADA architecture that simultaneously addresses system (SCADA master) compromises and network attacks
  - Key idea: distributes SCADA master replicas across three or more active geographic sites to ensure continuous availability
- An extension of the architecture that avoids constructing additional power company control centers by leveraging commodity data centers
  - Makes the architecture feasible for deployment
- An open-source implementation and evaluation of the architecture
  - Spire: www.dsn.jhu.edu/spire

Addressing System Compromises

- Byzantine Fault Tolerant Replication (BFT)
  - Correctly maintains state in the presence of compromises
  - 3f+1 replicas needed to tolerate up to f intrusions
  - 2f+1 connected correct replicas required to make progress
  - Prime protocol – latency guarantees under attack [ACKL11]

Addressing System Compromises

- Byzantine Fault Tolerant Replication (BFT)
  - Correctly maintains state in the presence of compromises
  - 3f+1 replicas needed to tolerate up to f intrusions
  - 2f+1 connected correct replicas required to make progress
  - Prime protocol – latency guarantees under attack [ACKL11]
- What prevents an attacker from reusing the same exploit to compromise more than f replicas?

Addressing System Compromises

- Byzantine Fault Tolerant Replication (BFT)
  - Correctly maintains state in the presence of compromises
  - 3f+1 replicas needed to tolerate up to f intrusions
  - 2f+1 connected correct replicas required to make progress
  - Prime protocol – latency guarantees under attack [ACKL11]
- Diversity
  - Present a different attack surface so that an adversary cannot exploit a single vulnerability to compromise all replicas
  - Multicompiler from UC Irvine [HNLBF13]

Addressing System Compromises

- Byzantine Fault Tolerant Replication (BFT)
  - Correctly maintains state in the presence of compromises
  - 3f+1 replicas needed to tolerate up to f intrusions
  - 2f+1 connected correct replicas required to make progress
  - Prime protocol – latency guarantees under attack [ACKL11]
- Diversity
  - Present a different attack surface so that an adversary cannot exploit a single vulnerability to compromise all replicas
  - Multicompiler from UC Irvine [HNLBF13]
- What prevents an attacker from compromising more than f replicas over time?

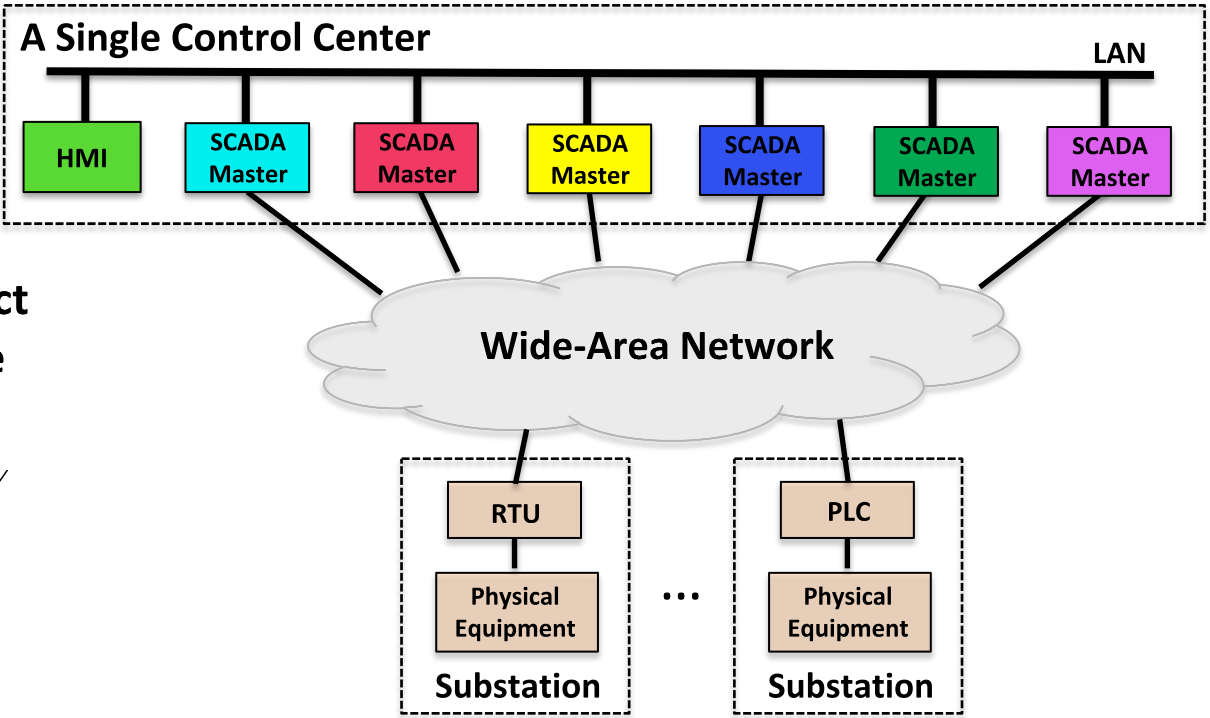
Addressing System Compromises

- Byzantine Fault Tolerant Replication (BFT)
  - Correctly maintains state in the presence of compromises
  - 3f+1 replicas needed to tolerate up to f intrusions
  - 2f+1 connected correct replicas required to make progress
  - Prime protocol – latency guarantees under attack [ACKL11]
- Diversity
  - Present a different attack surface so that an adversary cannot exploit a single vulnerability to compromise all replicas
  - Multicompiler from UC Irvine [HNLBF13]
- Proactive Recovery
  - Periodically rejuvenate replicas to a known good state to cleanse any potentially undetected intrusions
  - 3f+2k+1 replicas needed to simultaneously tolerate up to f intrusions and k recovering replicas [SBCNV10]
  - 2f+k+1 connected correct replicas required to make progress

Intrusion Tolerance State-of-the-Art in Research

6 (progress: 4)

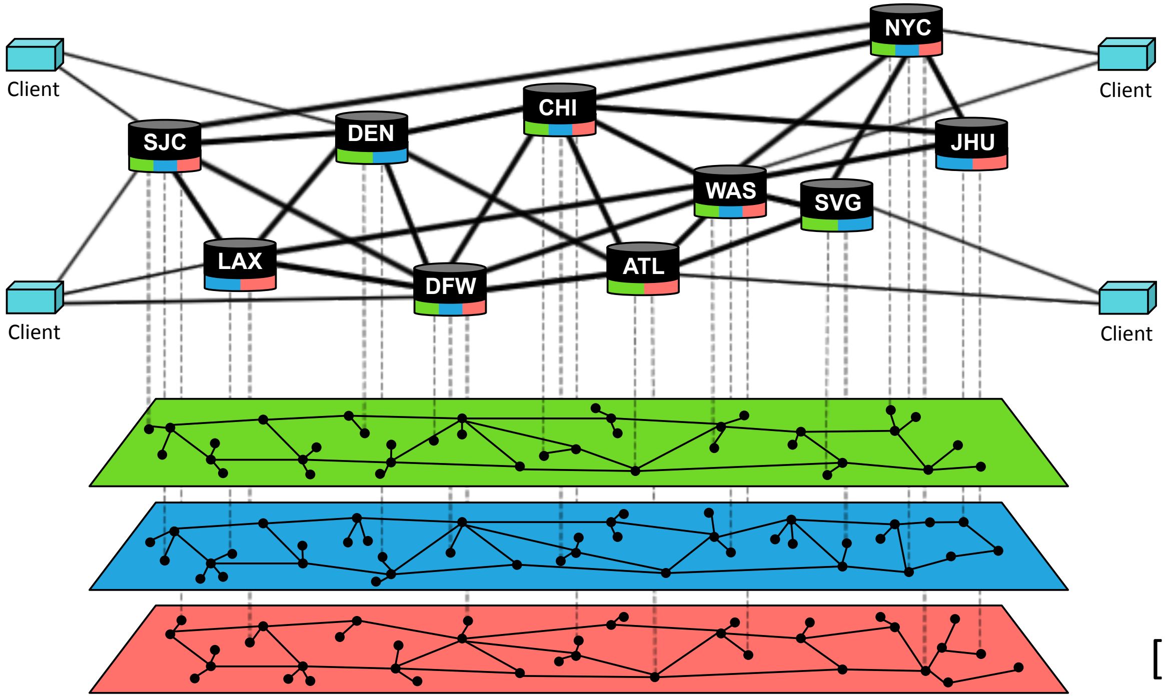
- **3f+2k+1 = 6 total replicas**
- **2f+k+1 = 4 connected correct replicas required to provide bounded delay**



	1	2-2	6
All Correct	Green	Green	Green
Proactive Recovery (PR)	Yellow	Green	Green
Network Attack	Red	Red	Red
Network Attack + PR	Red	Red	Red
Intrusion	Gray	Gray	Green
Intrusion + PR	Gray	Gray	Green
Network Attack + Intrusion	Gray	Gray	Red
Network Attack + Intrusion + PR	Gray	Gray	Red

Green	Bounded Delay
Yellow	Bounded Delay, except when rejuvenating any correct replica
Orange	Eventual Progress – Human in the loop. Potentially powering up cold backup control center
Red	Eventual Progress – No bound. Network has to heal, crash has to be repaired, and/or intrusion needs to be cleansed
Gray	Incorrect System

Addressing Network Attacks (Part 1): Spines Intrusion-Tolerant Network



[IEEE ICDCS 2016]

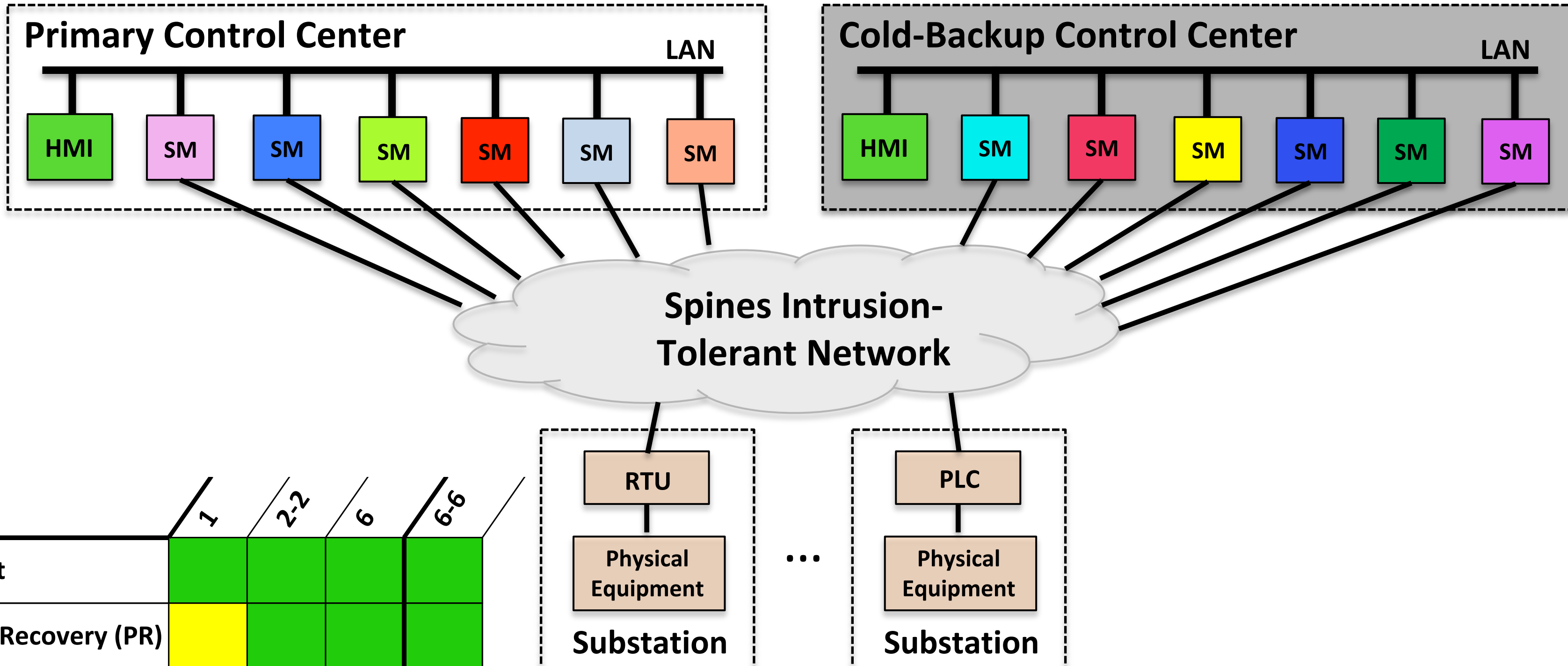
- Resilient Overlay Network Architecture: overcomes compromises in the underlying network infrastructure
- Intrusion-Tolerant Overlay Protocols: overcome compromises of overlay nodes

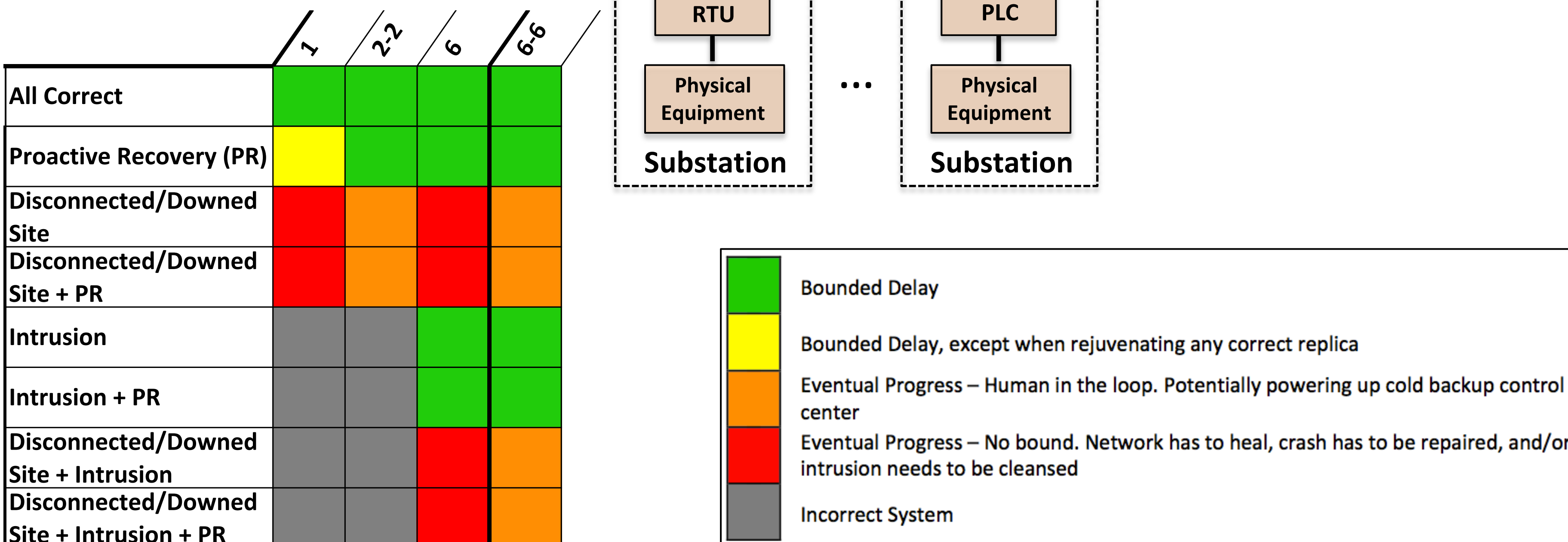
Addressing Network Attacks (Part 1):
Spines Intrusion-Tolerant Network

- Makes widespread network disruption nearly infeasible (requires simultaneous meltdown of multiple ISP backbones)
- A dedicated (e.g. nation-state) attacker can still invest the resources to disconnect a single targeted site
- Reduces the problem of addressing arbitrary network attacks to handling a single downed or disconnected site

New Natural Extensions (1/2): Primary-Backup Sites with Intrusion-Tolerant Replication

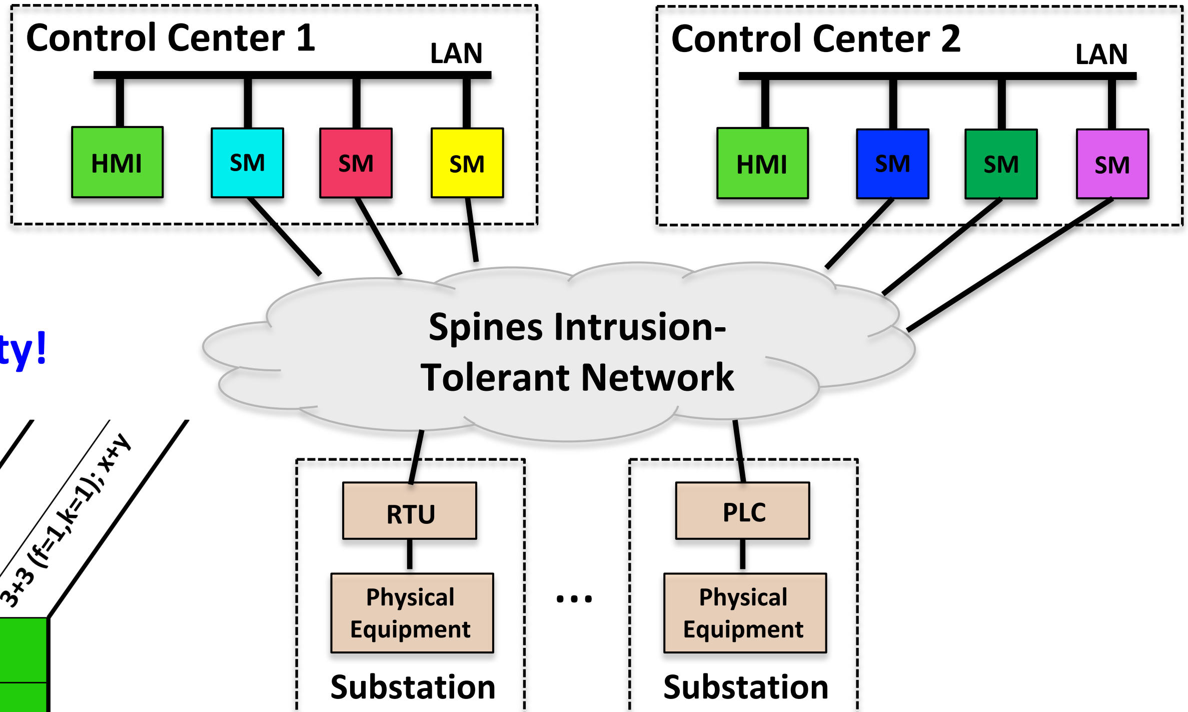
6-6





New Natural Extensions (2/2):
Active Intrusion-Tolerant Replication across Two Sites

3+3 (progress: 4)



Need more than two sites to provide continuous availability!

	1	2-2	6	6-6	3+3 (f=1,k=1); x+y
All Correct	green	green	green	green	green
Proactive Recovery (PR)	yellow	green	green	green	green
Disconnected/Downed Site	red	orange	red	orange	red
Disconnected/Downed Site + PR	red	orange	red	orange	red
Intrusion	gray	gray	green	green	green
Intrusion + PR	gray	gray	green	green	green
Disconnected/Downed Site + Intrusion	gray	gray	red	orange	red
Disconnected/Downed Site + Intrusion + PR	gray	gray	red	orange	red

- (green) Bounded Delay
- (yellow) Bounded Delay, except when rejuvenating any correct replica
- (orange) Eventual Progress – Human in the loop. Potentially powering up cold backup control center
- (red) Eventual Progress – No bound. Network has to heal, crash has to be repaired, and/or intrusion needs to be cleansed
- (gray) Incorrect System

Addressing Network Attacks (Part 2):
Active Replication Across Three or More Sites

- Two sites (even if both active) cannot provide intrusion tolerance and the necessary resilience to network attacks
  - True for any X + Y configuration
  - Replication protocol requires more than half the replicas to work
  - Site disconnection can make at least half of the system unavailable
  - Therefore, a solution requires active replication across three or more sites
- Control centers are expensive!
  - Setup to control, monitor, and communicate with RTUs in the field
  - Therefore, to be feasible, solutions should fit the two-control center model used by power companies
- **New idea**: devise an architecture where additional sites beyond the two control centers do not need to control RTUs or PLCs
  - Commodity data centers provide cost-effective alternative
  - Commodity data centers are becoming prevalent

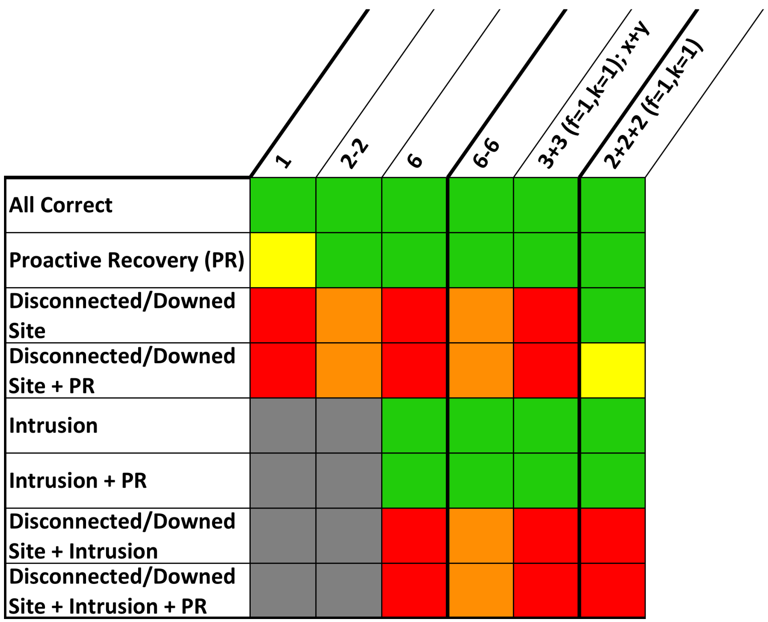
Novel Resilient Configurations (1/6)

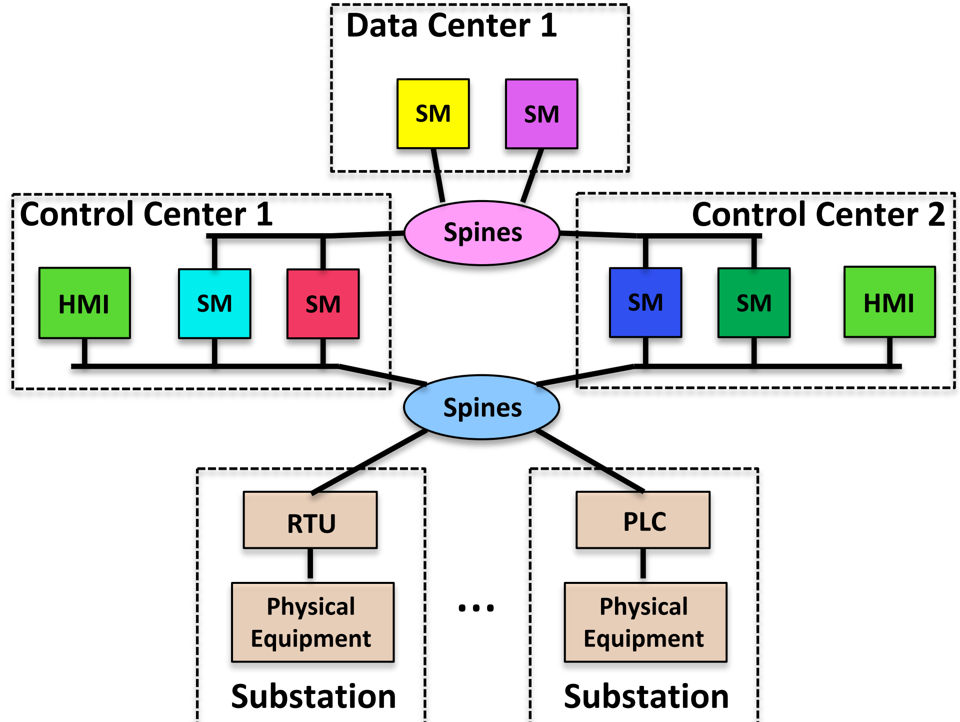
2+2+2 (progress: 4)

Two separate Spines networks:

- **One to communicate with RTUs in the field**
- **One for SCADA Master coordination**

Need to increase the number of replicas to cover disconnected sites due to network attacks!





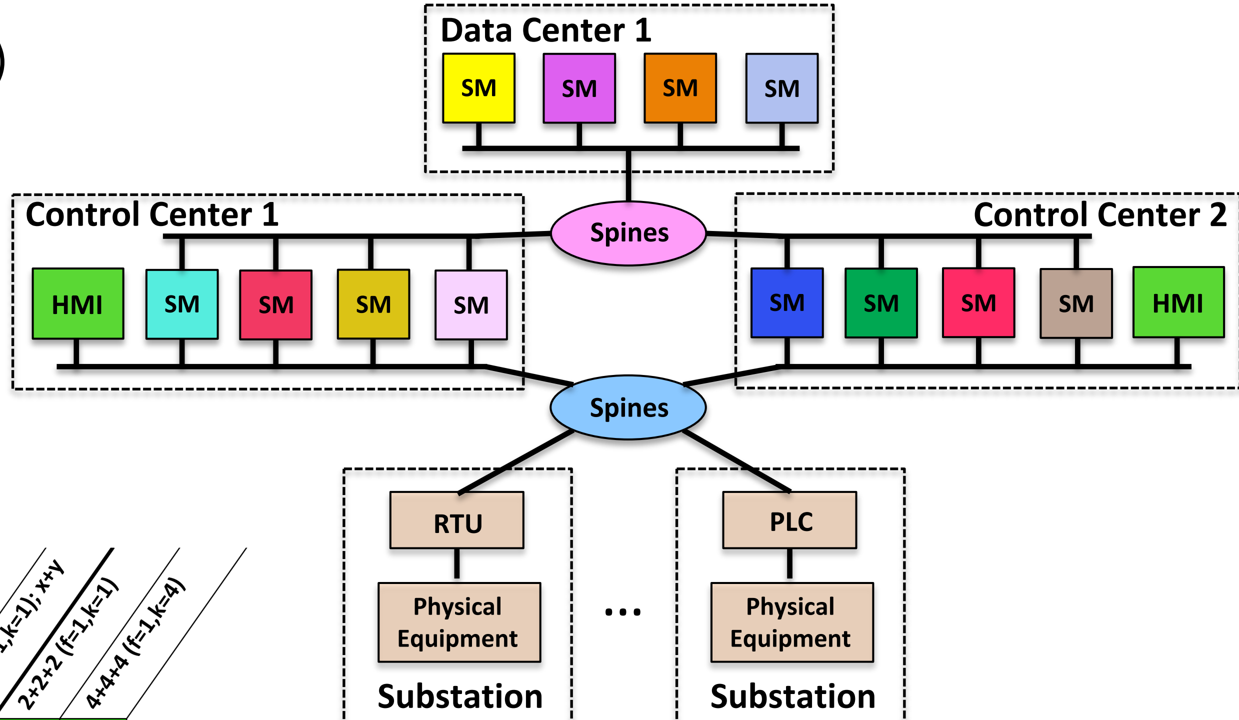
- Green: Bounded Delay
- Yellow: Bounded Delay, except when rejuvenating any correct replica
- Orange: Eventual Progress – Human in the loop. Potentially powering up cold backup control center
- Red: Eventual Progress – No bound. Network has to heal, crash has to be repaired, and/or intrusion needs to be cleansed
- Grey: Incorrect System

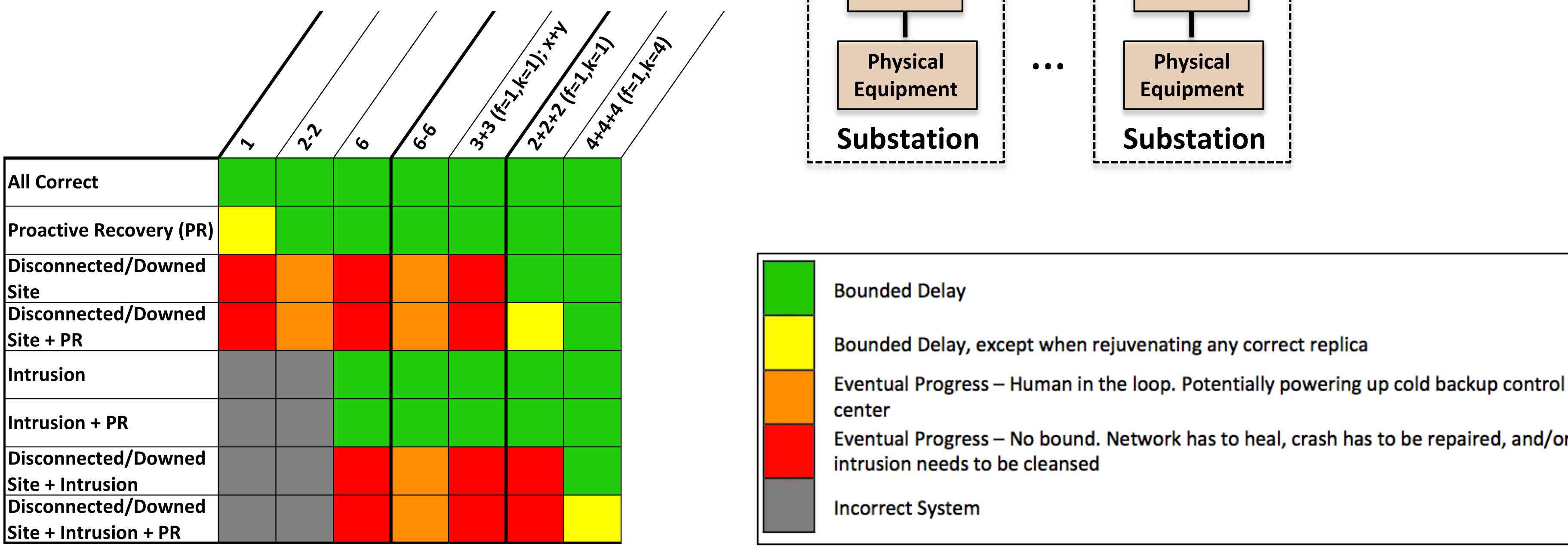
Novel Resilient Configurations (2/6)

4+4+4 (progress: 7)

- **Increase *k*** **to include the number of replicas in the largest site. In this case, k = 4.**

  3f+2k+1 = 3(1)+2(4)+1 = 12

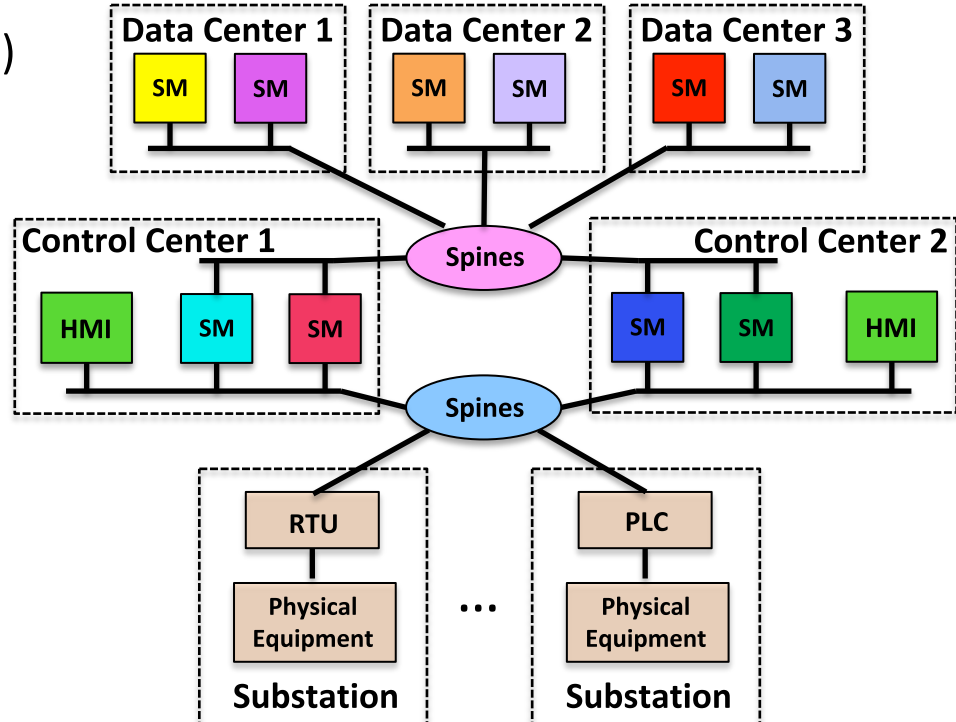




Novel Resilient Configurations (3/6)

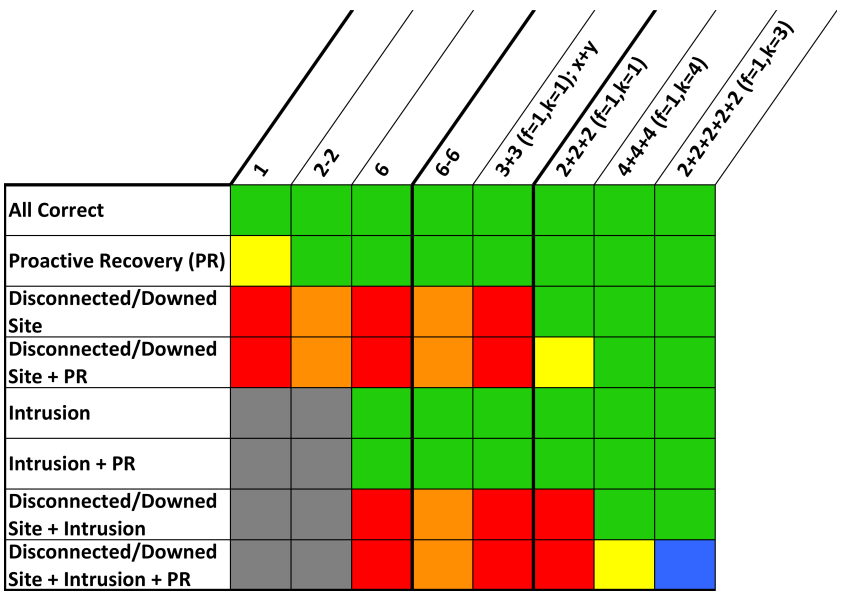
2+2+2+2+2
(2 control centers)

(progress: 6)



- **Increase *k* to include the size of largest site plus rejuvenating replica. In this case, k = 3.**

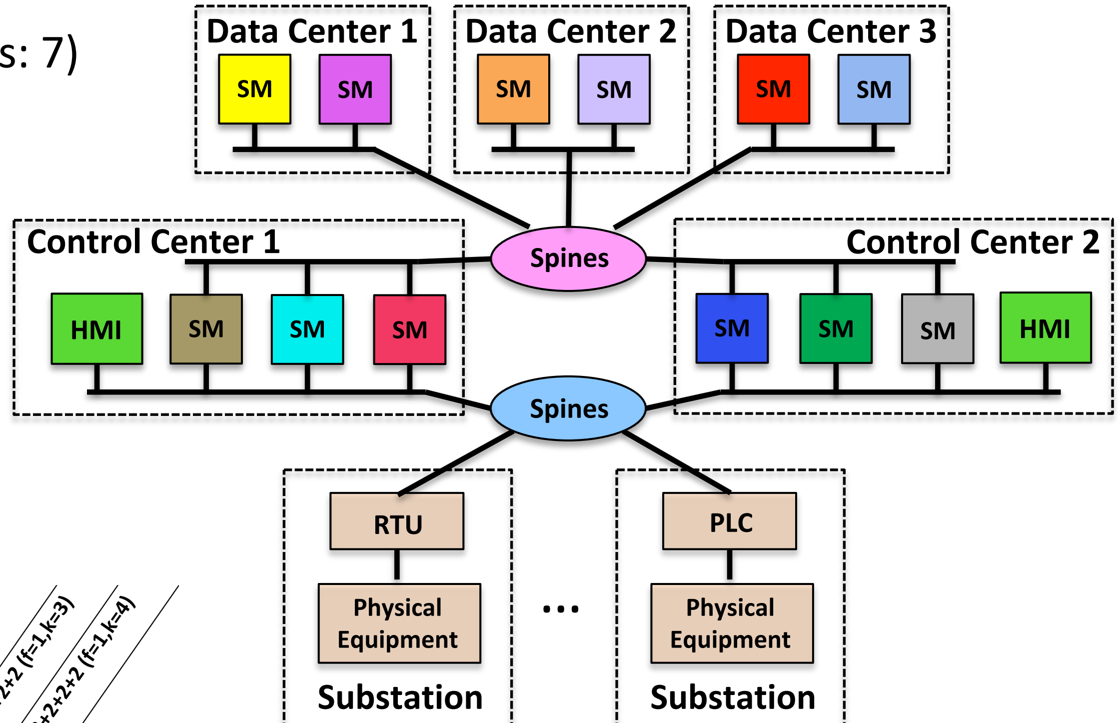
3f+2k+1 = 3(1)+2(3)+1 = 10



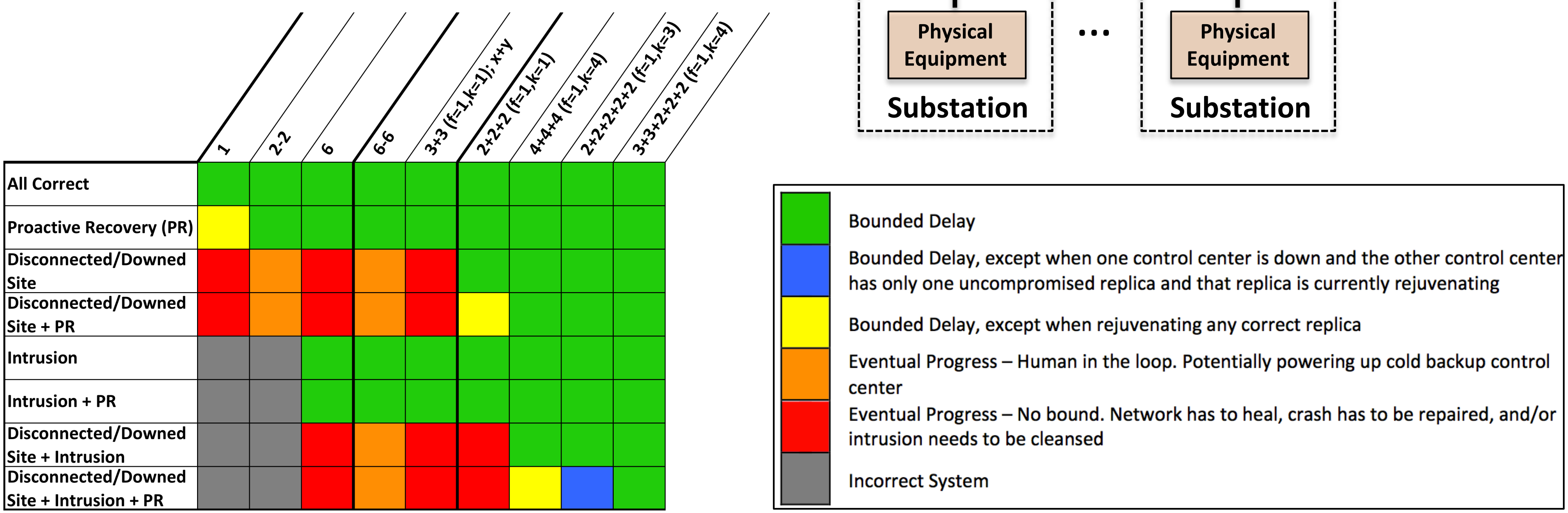
- Bounded Delay
- Bounded Delay, except when one control center is down and the other control center has only one uncompromised replica and that replica is currently rejuvenating
- Bounded Delay, except when rejuvenating any correct replica
- Eventual Progress – Human in the loop. Potentially powering up cold backup control center
- Eventual Progress – No bound. Network has to heal, crash has to be repaired, and/or intrusion needs to be cleansed
- Incorrect System

Novel Resilient Configurations (4/6)

3+3+2+2+2 (progress: 7)



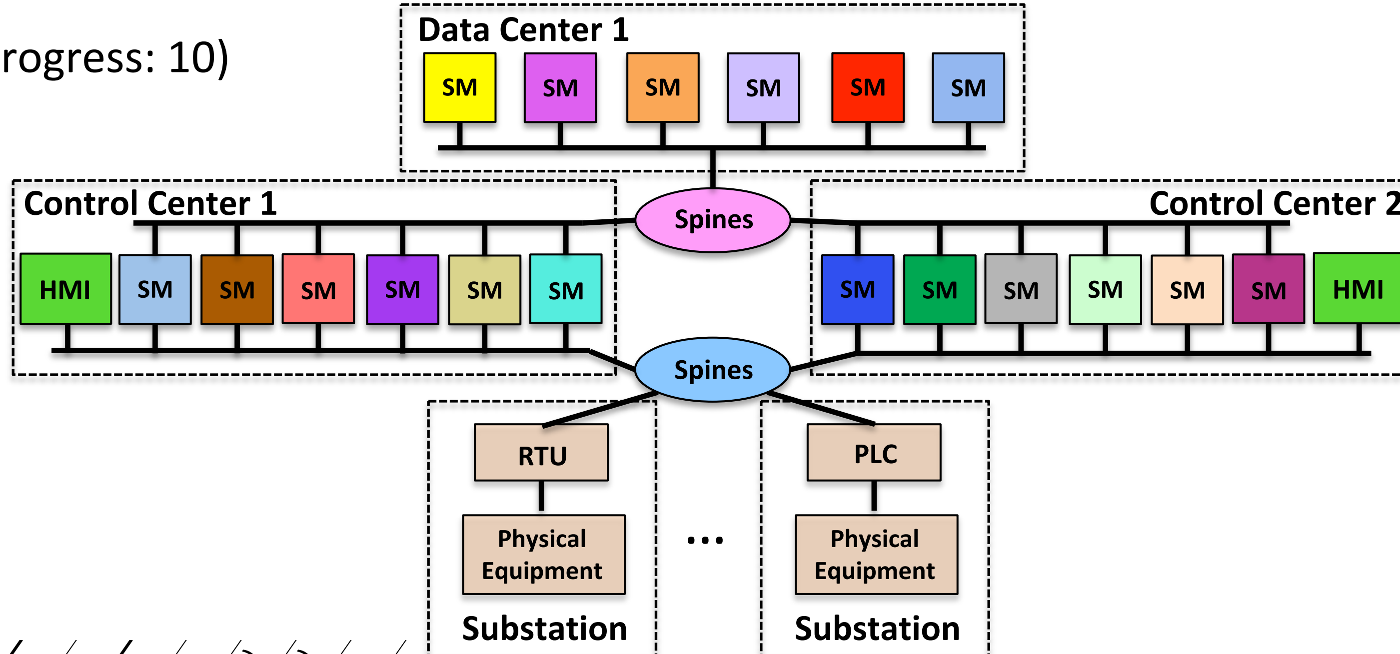
- **At least *f+2* replicas in each control center ensures one correct replica that can control RTUs even with intrusion and ongoing rejuvenation in connected control site**
- **First complete solution**

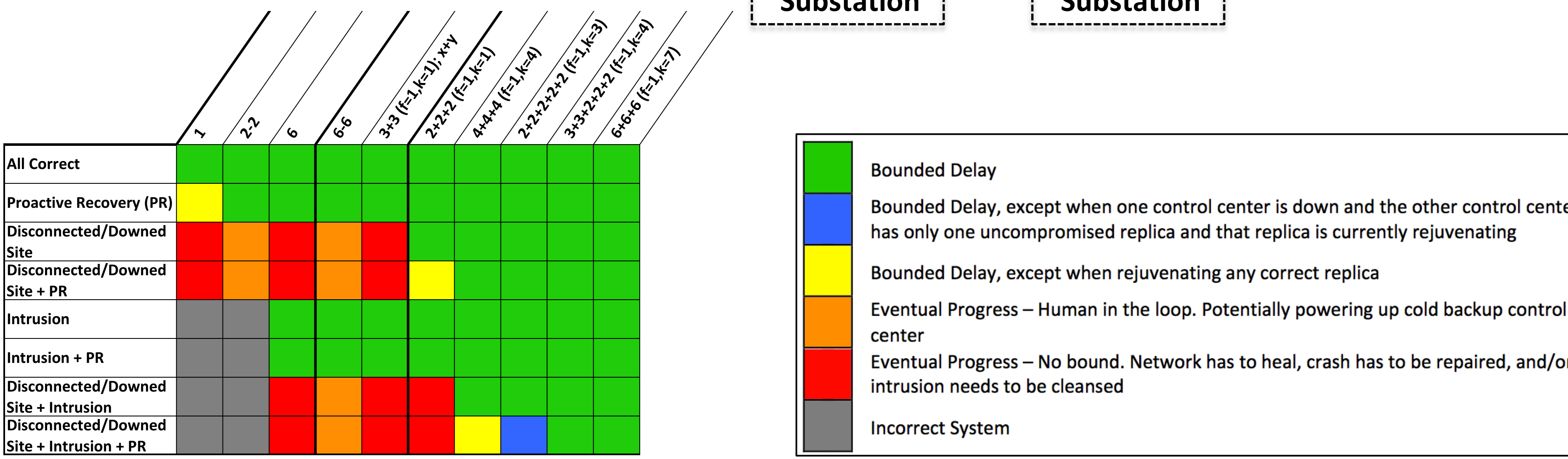


Novel Resilient Configurations (5/6)

6+6+6 (progress: 10)

- **Complete solution for 3 total sites:** **(2 control centers, 1 data center)**

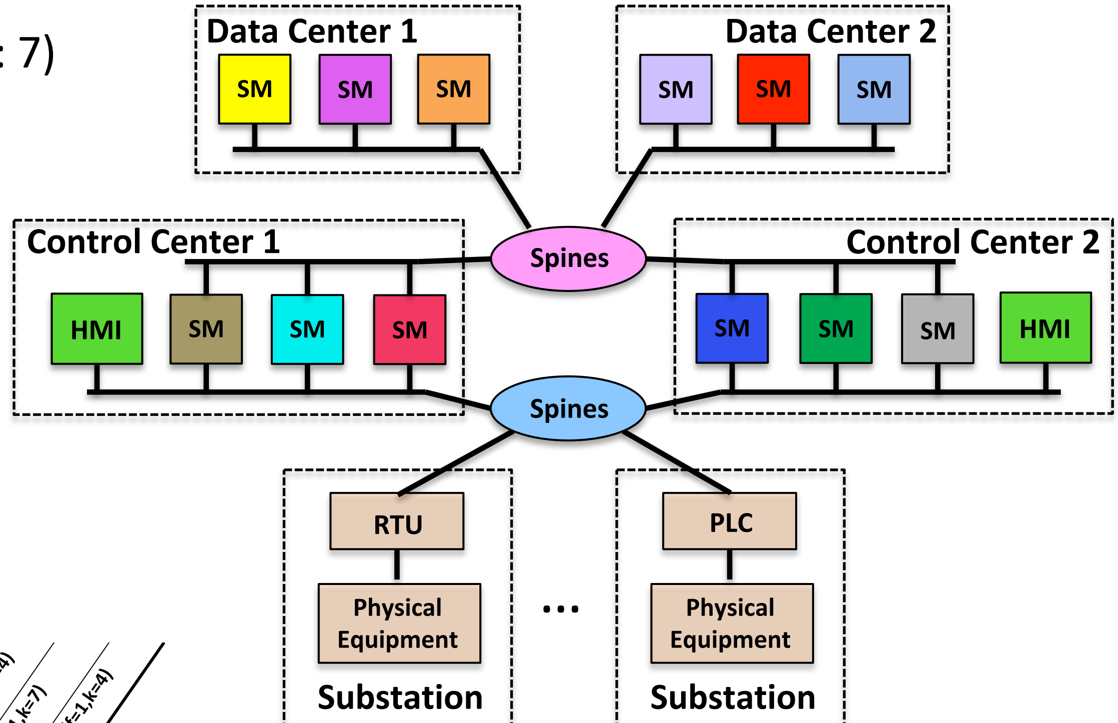


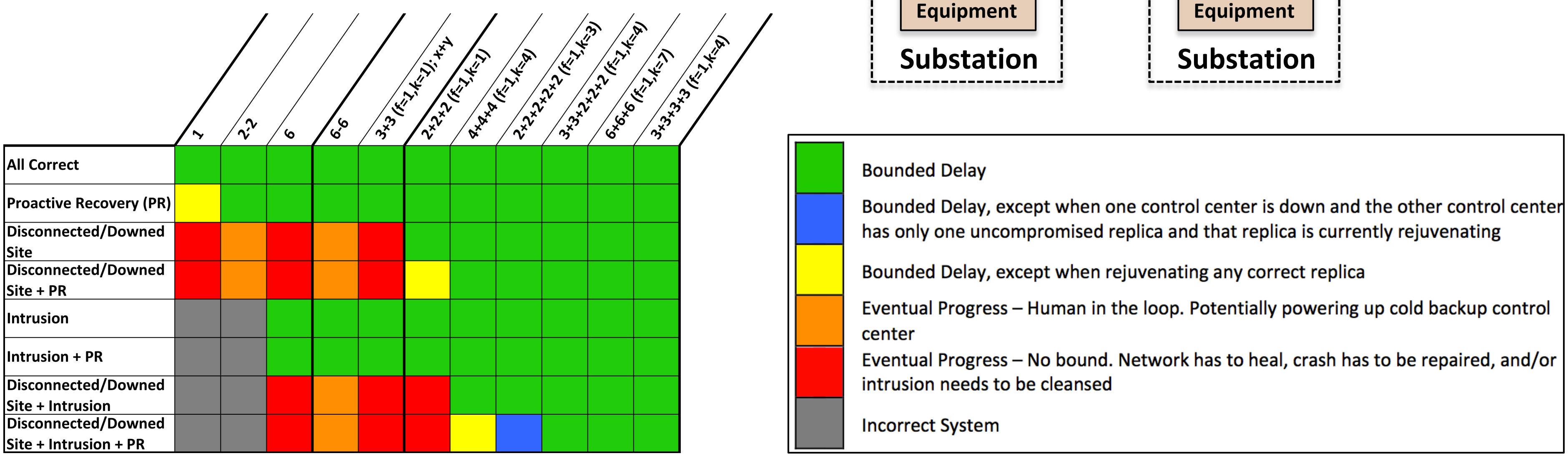


Novel Resilient Configurations (6/6)

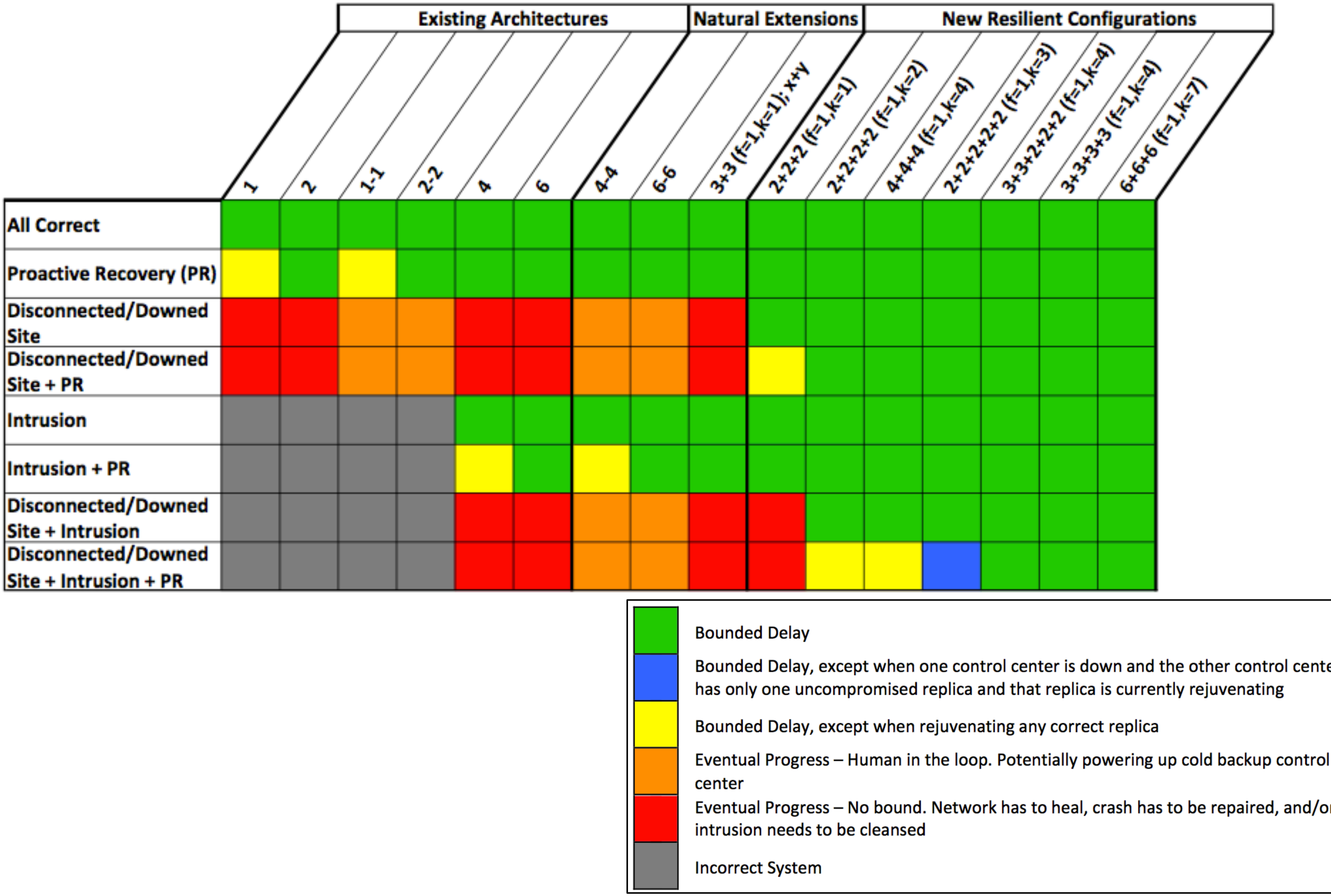
3+3+3+3 (progress: 7)

- **Complete solution for 4 total sites: (2 control centers, 2 data centers)**
- **Sweet-spot balancing the number of data center sites, the number of total replicas, and the communication overhead**





SCADA Architecture Comparison



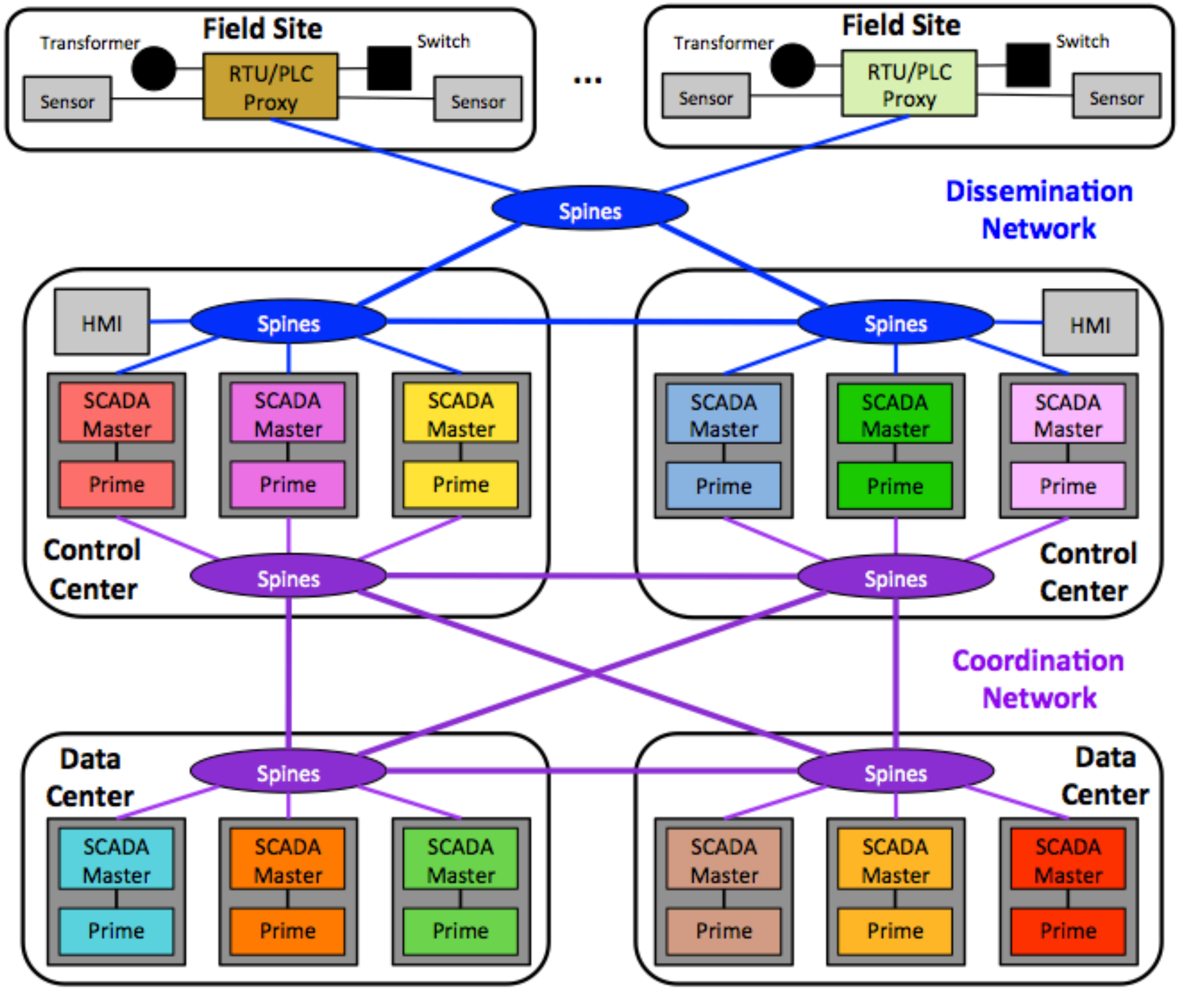
Intrusion-Tolerant SCADA Configuration Framework

- General framework to create SCADA configurations that use *S* total sites ($S > 2$) and tolerate *f* intrusions

	2 control centers + 1 data center	2 control centers + 2 data centers	2 control centers + 3 data centers
$f = 1$	6+6+6	3+3+3+3	3+3+2+2+2
$f = 2$	9+9+9	5+5+5+4	4+4+3+3+3
$f = 3$	12+12+12	6+6+6+6	5+5+4+4+4

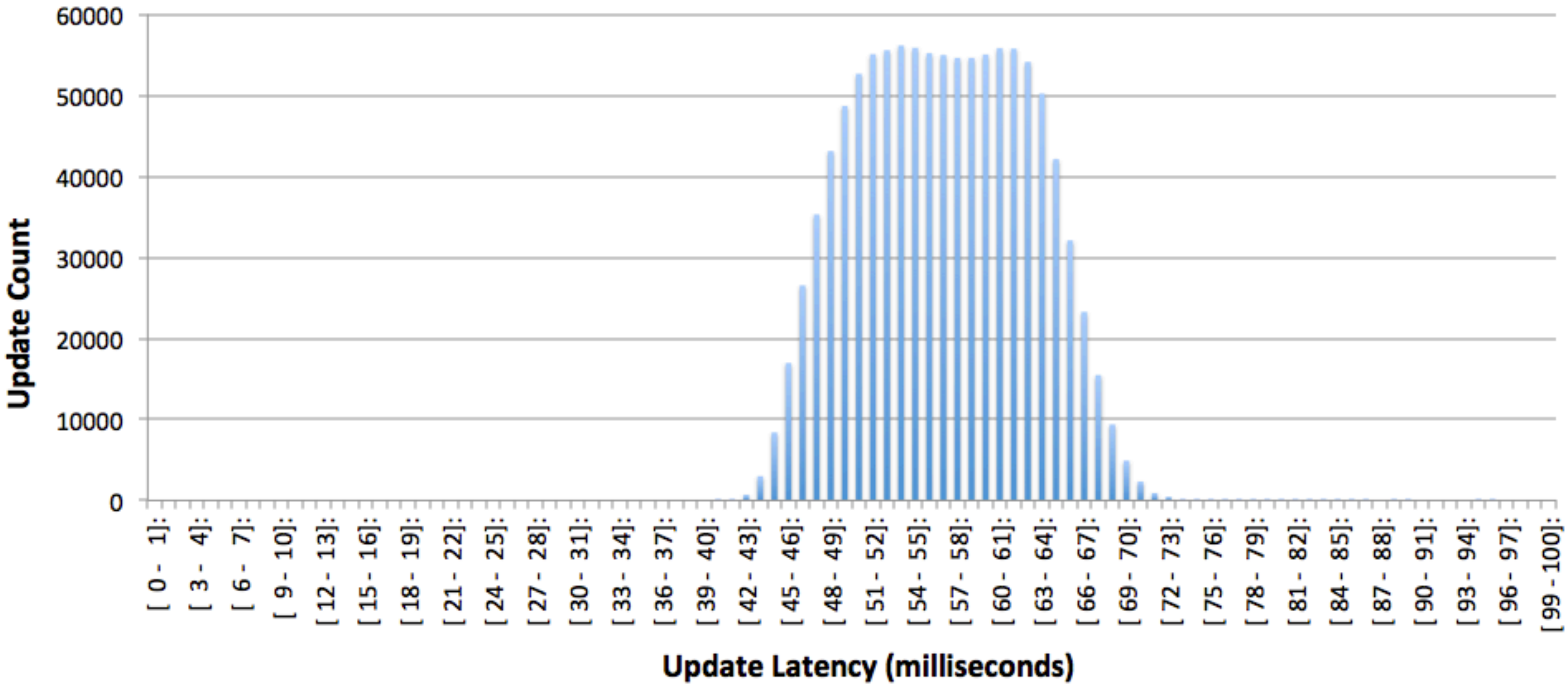
Minimum number of replicas required to overcome *f* intrusions, a single rejuvenating replica, and a single disconnected site, varying *f* and *S* (total number of sites).

Putting it all Together: Complete SCADA System



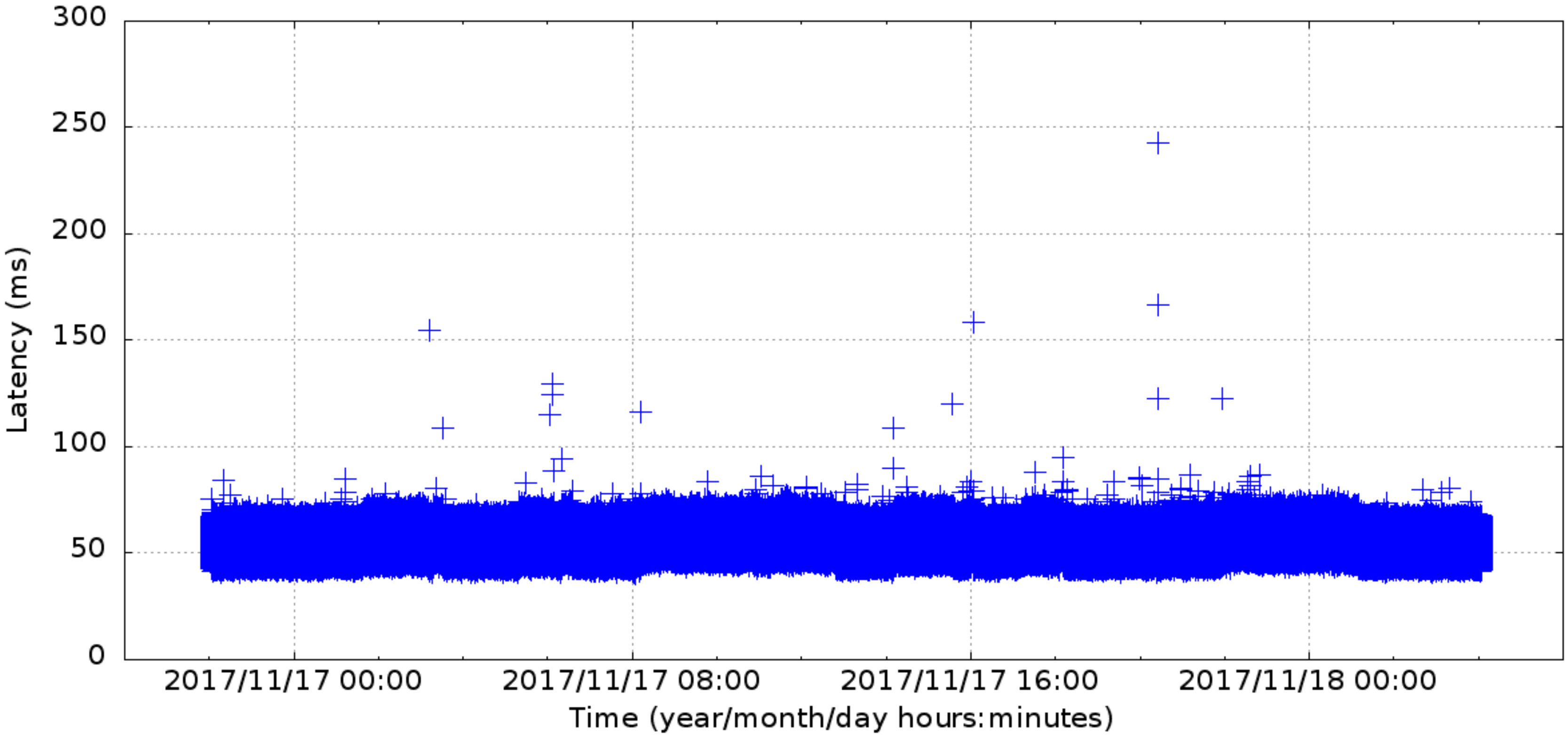
Spire Software Architecture for Configuration 3+3+3+3

Wide Area Update Latency Histogram



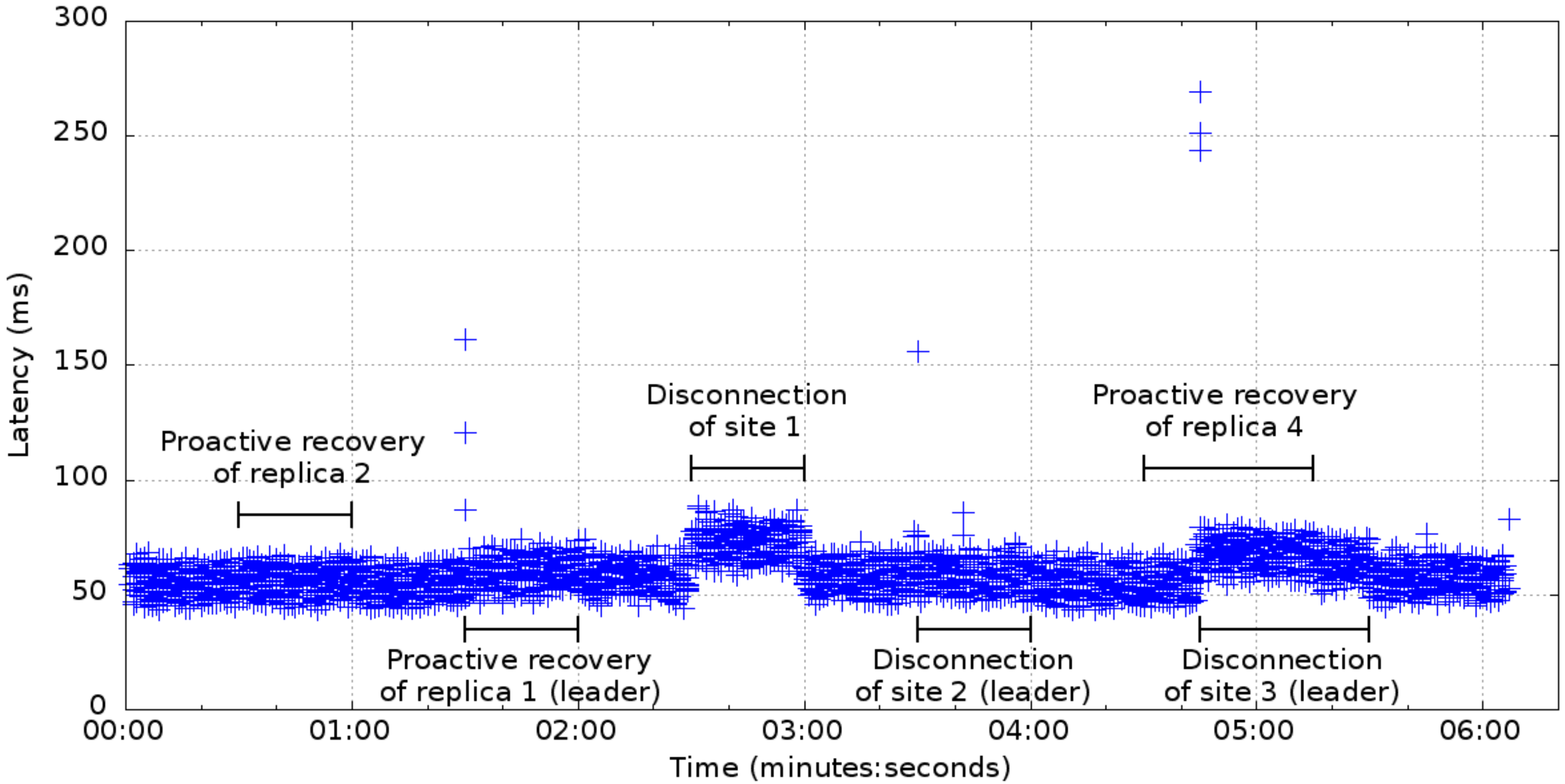
- 30-hour wide-area deployment of configuration 3+3+3+3
  - Control centers at JHU and SVG, data centers at WAS and NYC
  - 10 emulated substations sending periodic updates
  - 1.08 million updates (108K from each substation)
  - Nearly 99.999% of updates delivered within 100ms (56.5ms average)

Wide Area Update Latency Plot



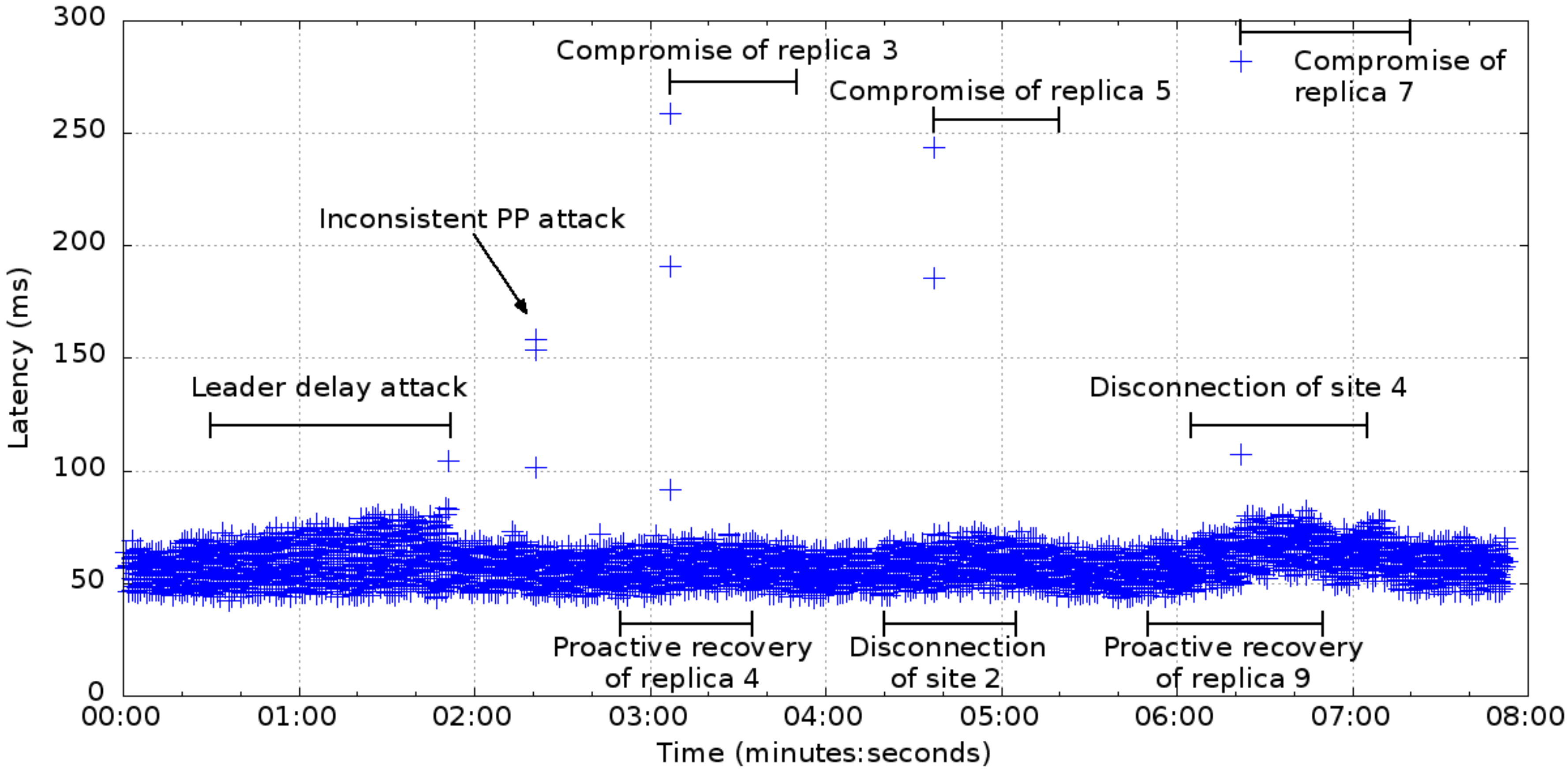
- 30-hour wide-area deployment of configuration 3+3+3+3
  - Only 13 updates out of 1.08 million delivered with latency exceeding 100ms
  - Only 1 update exceeded 200ms

Wide Area: Latency Under Attack



- Targeted attacks designed to disrupt the system
  - All combinations of site disconnection (due to network attack) + proactive recovery

Wide Area: Latency Under Attack



- Targeted attacks designed to disrupt the system
  - All combinations of intrusion + site disconnection (due to network attack) + proactive recovery

The Spire Forum

- Forum focused on open source intrusion-tolerant control systems for the power grid
- Please join the Spire forum if interested
- http://dsn.jhu.edu/spire

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Distributed Systems
and Networks Lab