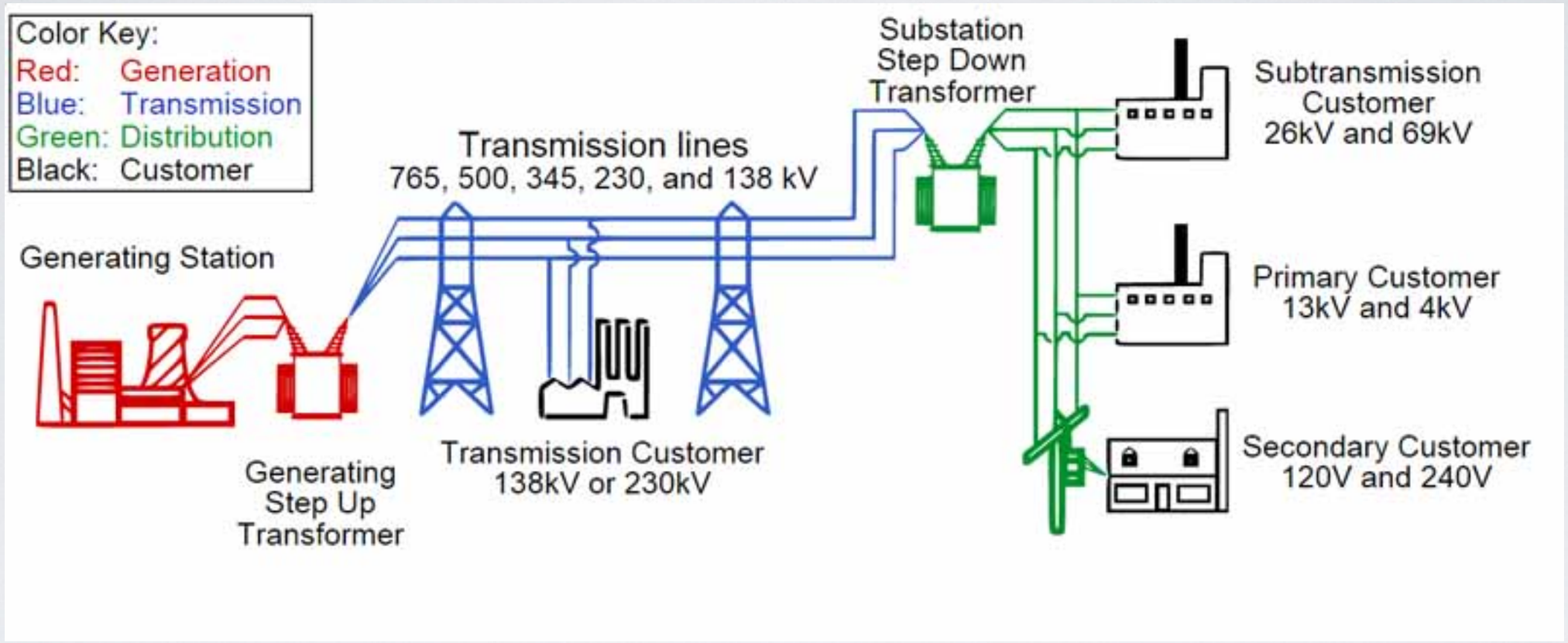
Evaluating BFT Protocols for Spire

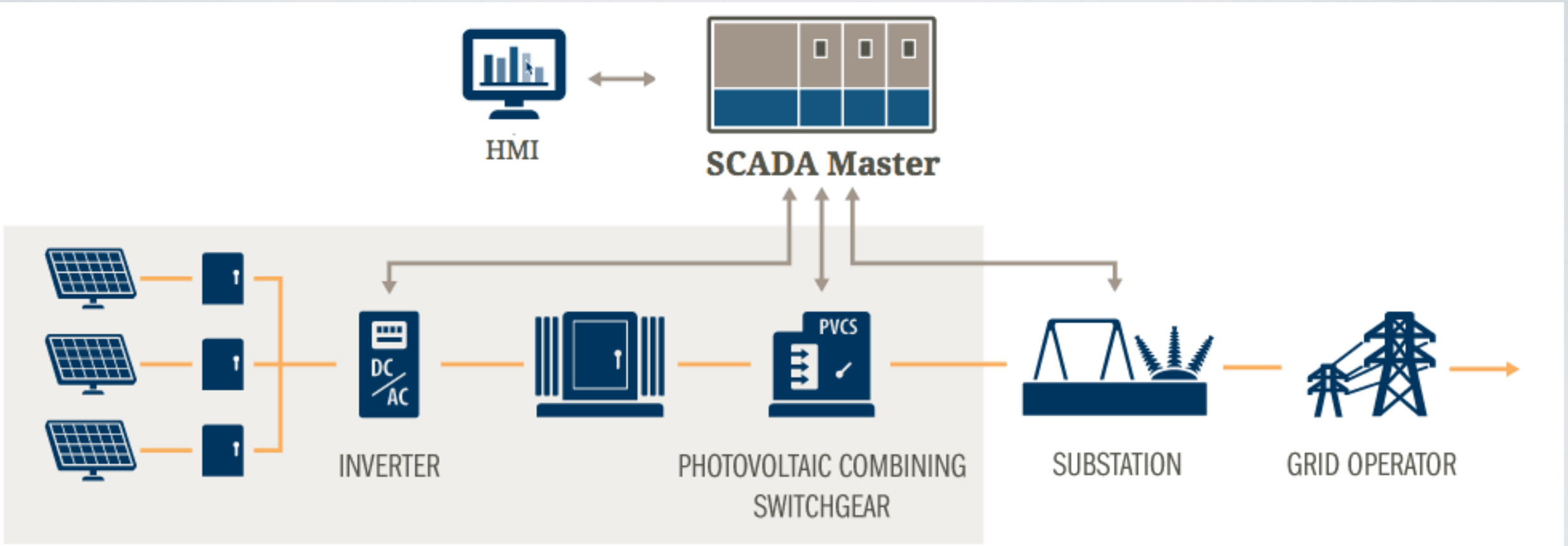
Henry Schuh & Sam Beckley

600.667 Advanced Distributed Systems & Networks

- SCADA & Spire Overview
- High-Performance, Scalable Spire
- Trusted Platform Module
- Known Network Characteristics
- Evaluating BFT-SMART
- Benchmarking Results
- Conclusions



Power Grid Overview



SCADA Overview

SCADA Requirements

- Must have very low latencies (100-200ms)
- Must have very high reliability
- Must be able to run for decades

SCADA Adopting IP & Internet

- In the past SCADA used proprietary protocols on air gapped systems
- Now moving to both IP & the Internet to reduce costs

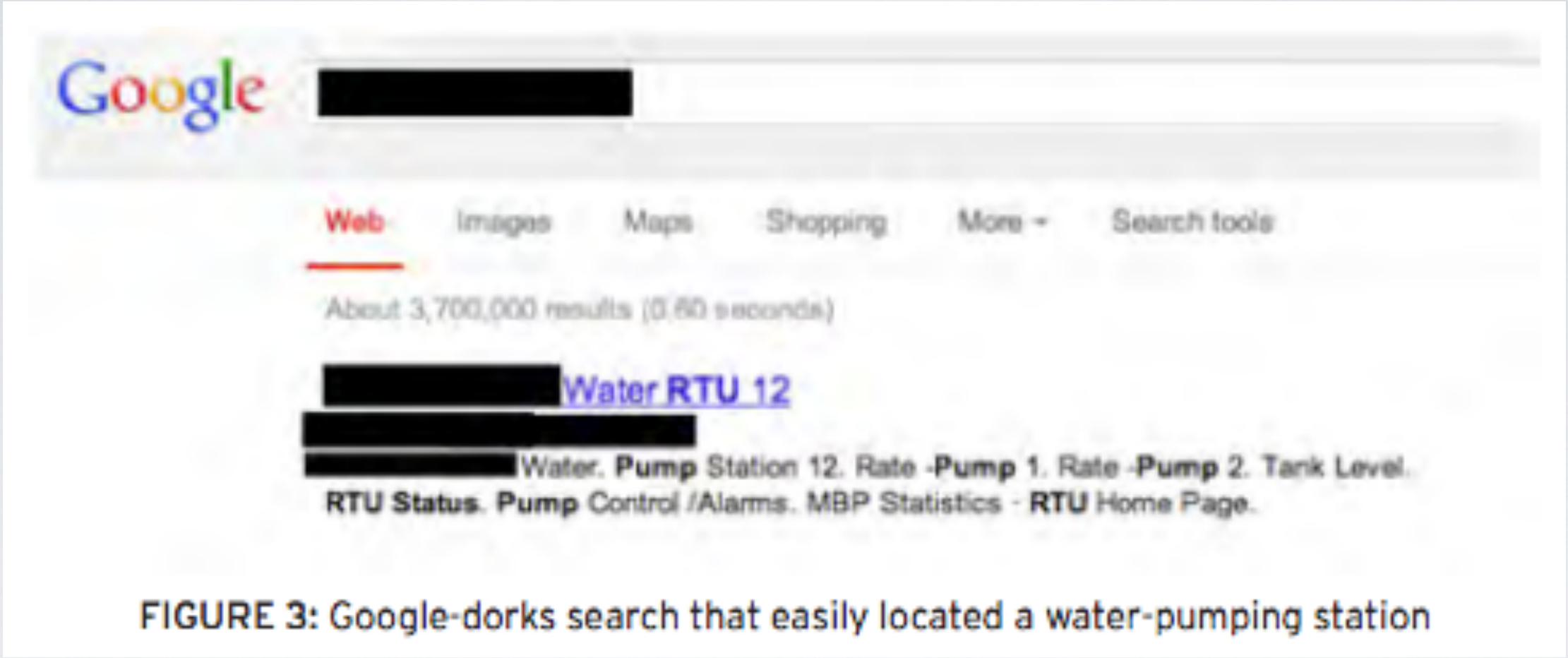


FIGURE 3: Google-dorks search that easily located a water-pumping station

"These devices were not only internet facing, they did not have security mechanisms to prevent unauthorized access"

- Trend Micro Incorporated, *Who's Really Attacking Your ICS Systems*

Attacks on SCADA Systems

28 Days: 39 Attacks
All targeted specifically at SCADA systems

The first attack was within 18 hours of the honeypot going live

Source: Trend Micro Incorporated,
Who's Really Attacking Your ICS Systems

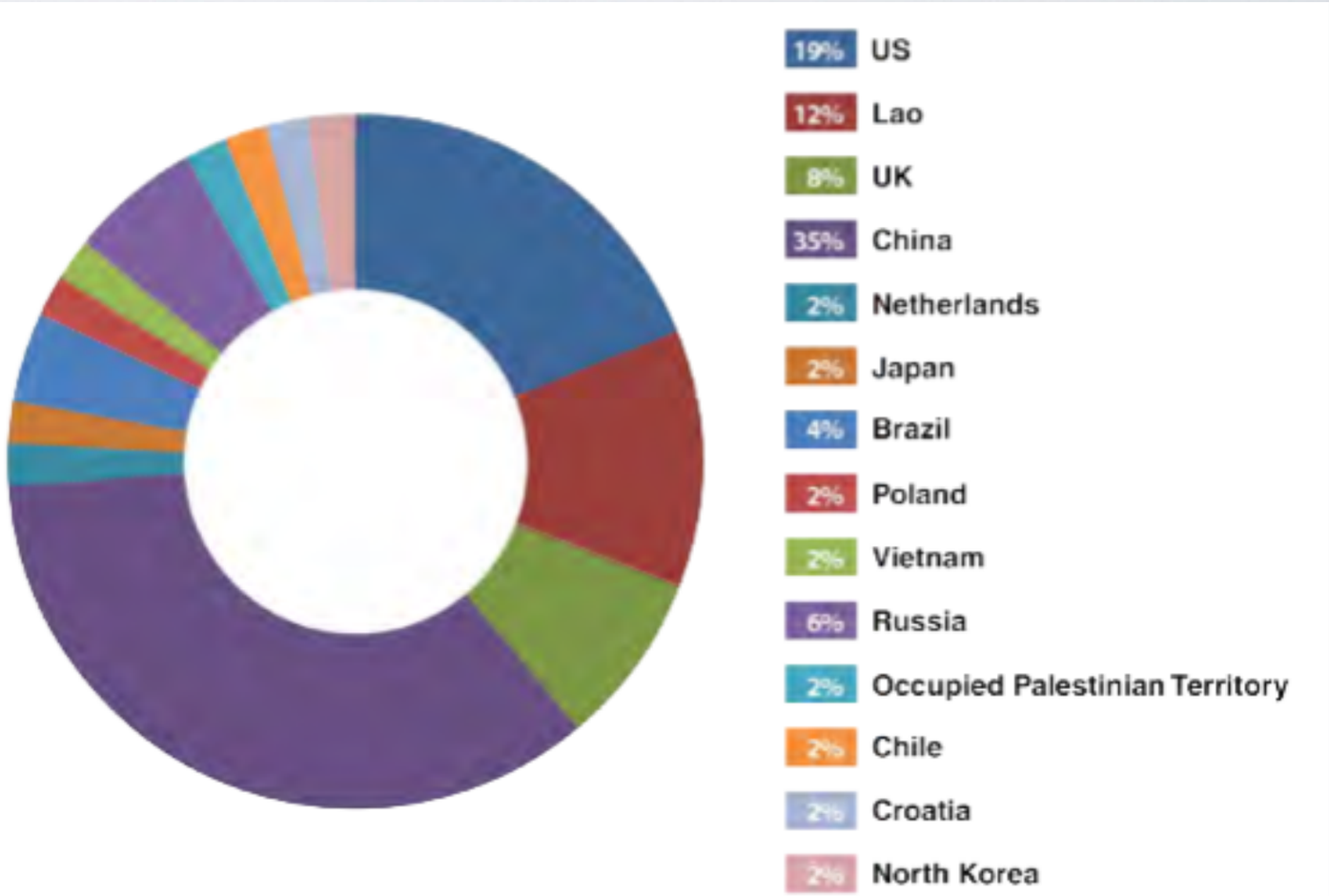


FIGURE 10: Country breakdown indicating the number of attack attempts

Distributed Replication

- Several machines that coordinate their actions such that they appear to be a single unified machine to a client.

Pros: High Availability and Performance

Cons: Cost of Synchronization

Intrusion Tolerant Replication

Somewhat Formally:

The ability to make progress in the presence of some number of malicious replicas with guaranteed correctness. Some protocols also guarantee a level of performance under attack.

Informally:

If some of the replicas get hacked the system still works.

Defense Across Space & Time

Defense Across Time:
Have to periodically regain control of a compromised machine to stop the attacker from eventually gaining control of the entire network.

Defense Across Space:
Every replica must present a unique attack surface so that one attack cannot be used to compromise every replica.

- SCADA & Spire Overview
- High-Performance, Scalable Spire
- Trusted Platform Module
- Known Network Characteristics
- Evaluating BFT-SMART
- Benchmarking Results
- Conclusions

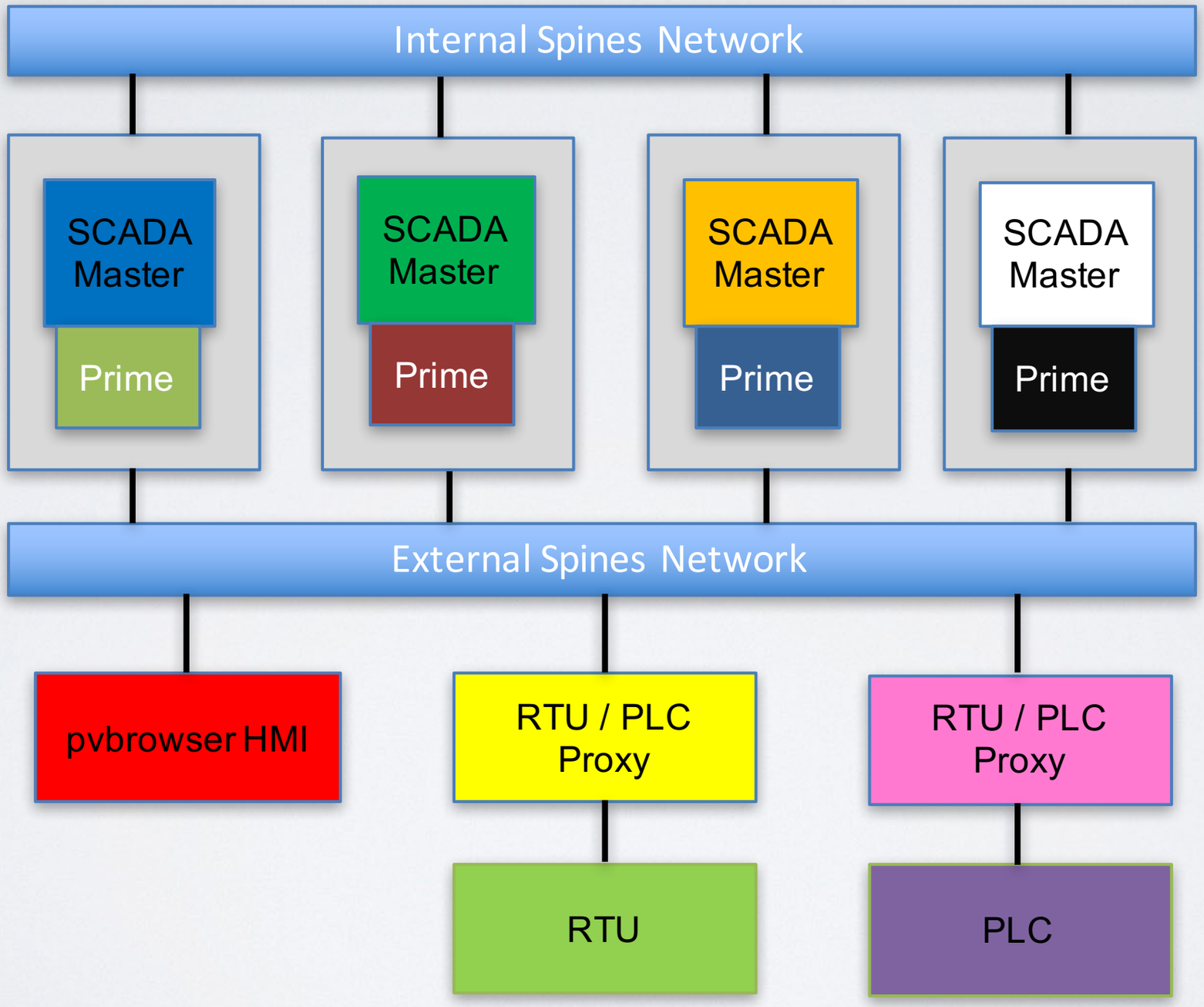
Spire

Open Source SCADA system that provides both standard crypto defense mechanisms as well as an intrusion tolerant SCADA Master.

Spire uses several different technologies

- Prime
- Spines
- PVBrowser

Spire


Internal Spines Network
SCADA Master
Prime
SCADA Master
Prime
SCADA Master
Prime
SCADA Master
Prime
External Spines Network
pvbrowser HMI
RTU / PLC Proxy
RTU / PLC Proxy
RTU
PLC

Scaling Spire

In order to tolerate more intrusions we need more replicas

The more replicas, the higher the latency becomes

We rely on having very low latency

Our Mission

Find a way to make Spire more scalable, to allow for more replicas, and thus more intrusions

3 Angles of Attack

Trusted Hardware - using a TPM

Taking Advantage of Known Network Characteristics

Hierarchy of Protocols

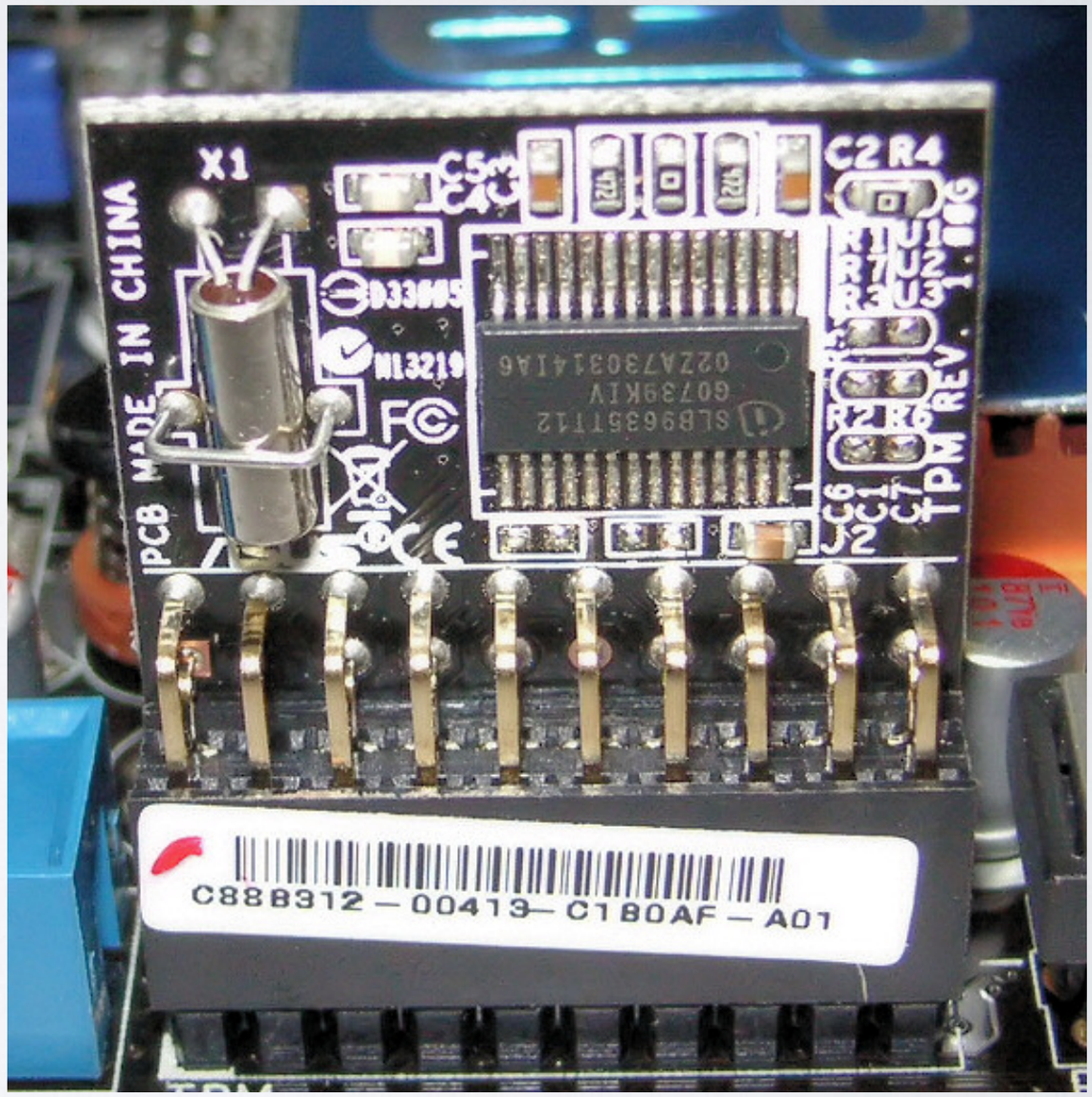
- SCADA & Spire Overview
- High-Performance, Scalable Spire
- Trusted Platform Module
- Known Network Characteristics
- Evaluating BFT-SMART
- Benchmarking Results
- Conclusions

Trusted Platform Module

Specialized chip that holds a secret key and can perform cryptographic functions for the rest of the machine

The key never leaves the TPM

Too slow :’(



- SCADA & Spire Overview
- High-Performance, Scalable Spire
- Trusted Platform Module
- Known Network Characteristics
- Evaluating BFT-SMART
- Benchmarking Results
- Conclusions

Leverage Network Characteristics

SCADA deployments are static and predictable

Most importantly, we know:

- Geographically close - low latency communication
- Consistent number of clients and messaging pattern

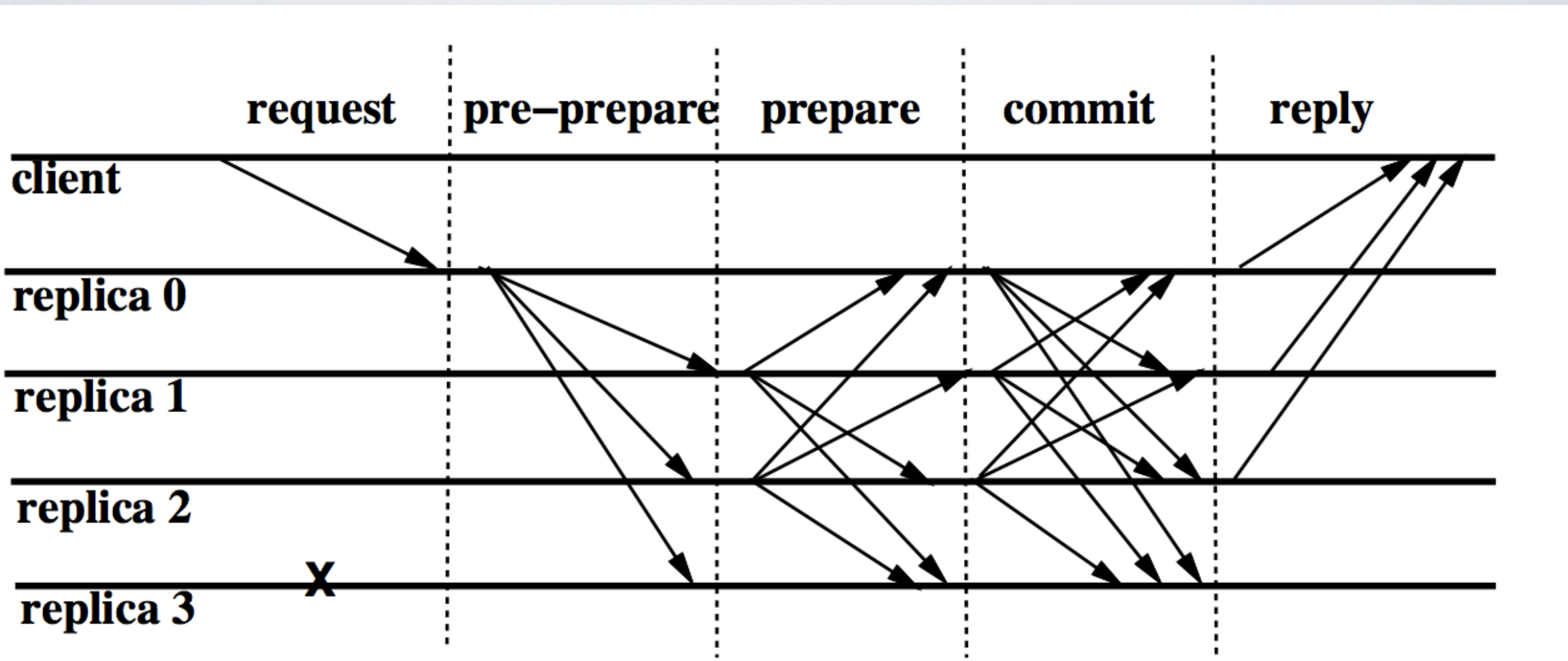
The Three BFT Protocol Families

PBFT

Spinning

Prime

PBFT


request
pre–prepare
prepare
commit
reply
client
replica 0
replica 1
replica 2
replica 3
X

PBFT

When the leader fails we must perform a “view change”
This is by far the most expensive operation in PBFT

“[The view change] is the Achilles Heel”
-Yair Amir

Spinning

Every ordering is done by a different leader

A bad leader can delay exactly one ordering before it is evicted from the protocol

Prime

Designed to remove load from the leader to allow for many clients without performance degradation

Performs one ordering every X milliseconds

Prime

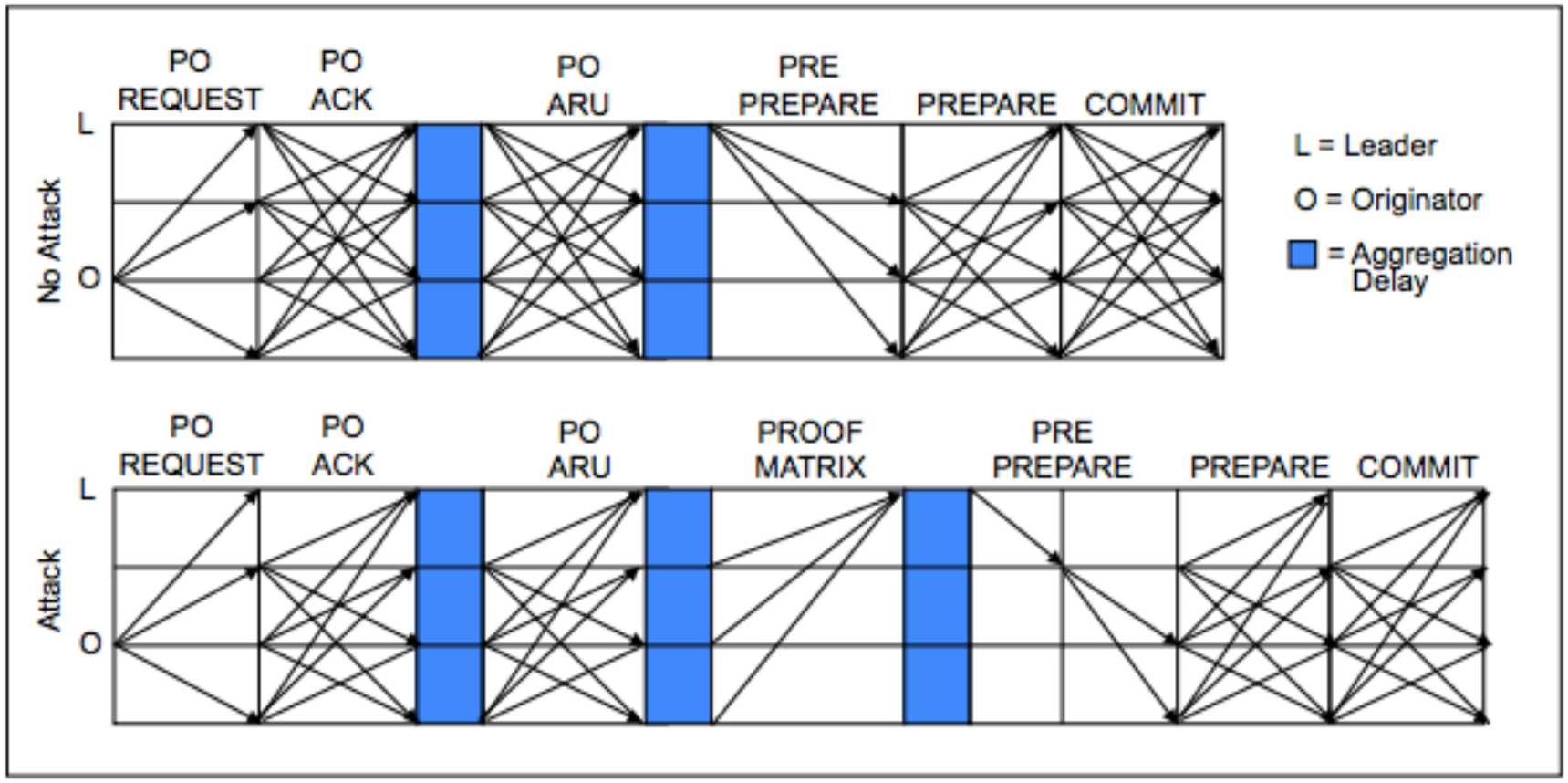


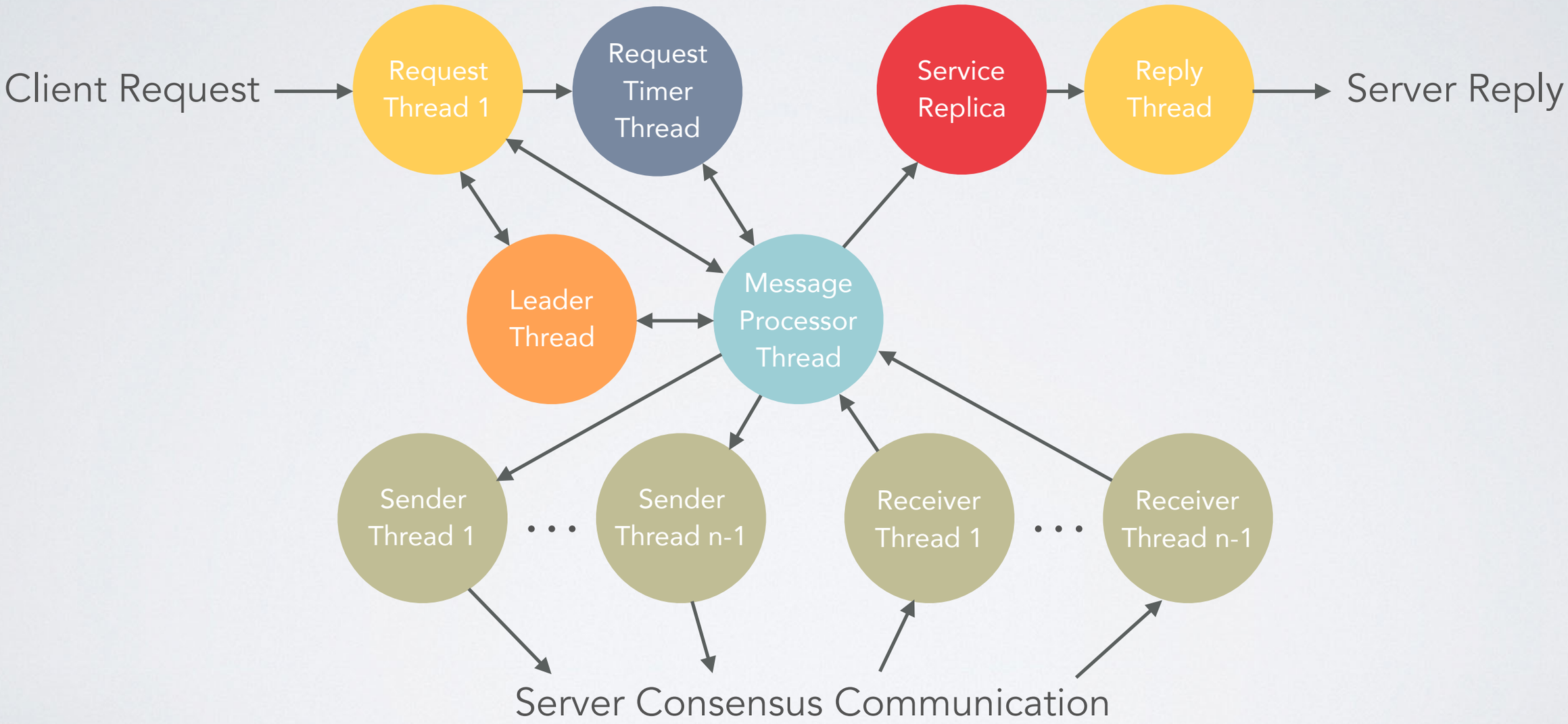
Fig. 1. The pre-ordering phase in Prime helps to detect a delay attack and replace a malicious leader.

- SCADA & Spire Overview
- High-Performance, Scalable Spire
- Trusted Platform Module
- Known Network Characteristics
- Evaluating BFT-SMART
- Benchmarking Results
- Conclusions

BFT-SMART

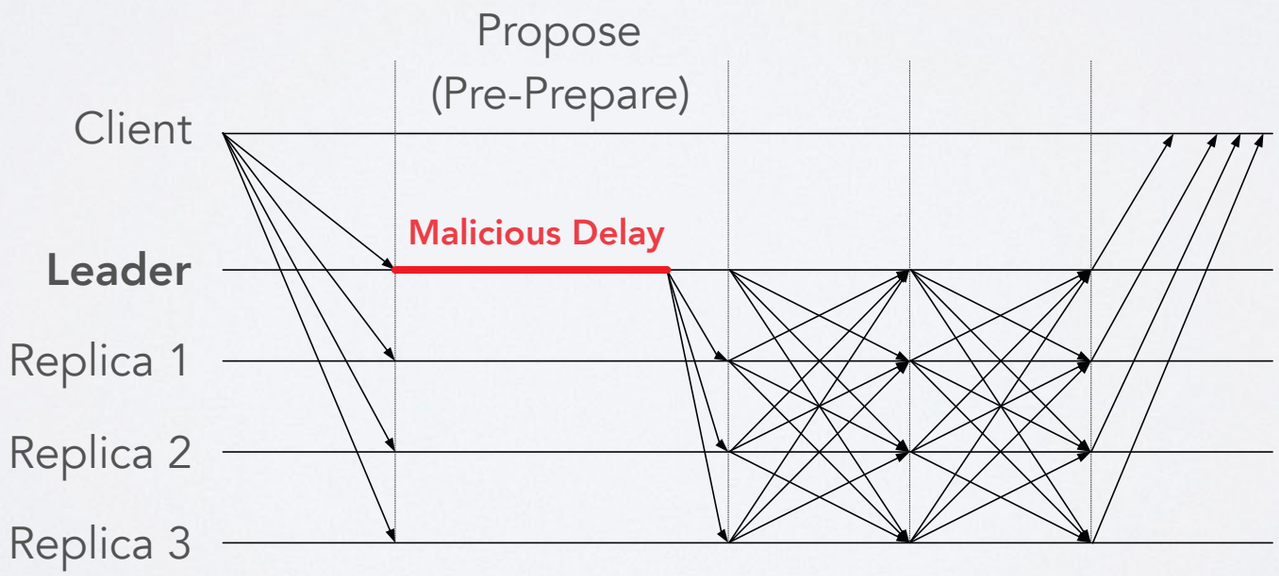
- Implements *"Yet Another Visit to Paxos"* protocol (IBM Zurich) in Java
- Modular, multi-threaded server replicas
- Standard BFT message pattern
- Modern protocol with ongoing development

Multithreaded Design



BFT-SMART and Performance Attacks

- Consensus relies on leader to order messages
- A malicious leader could delay progress
- Timeouts limit the leader's worst-cast performance



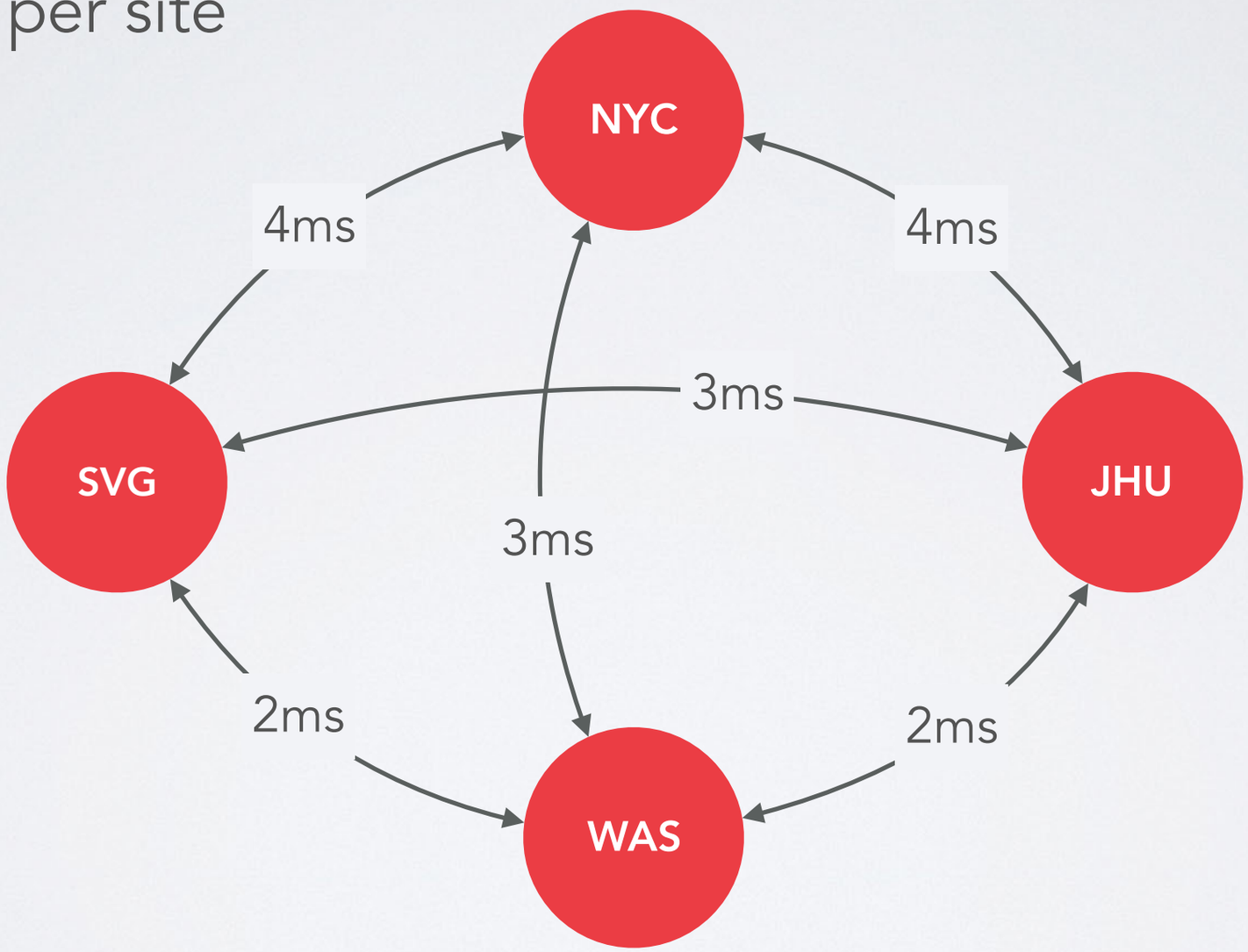
- SCADA & Spire Overview
- High-Performance, Scalable Spire
- Trusted Platform Module
- Known Network Characteristics
- Evaluating BFT-SMART
- Benchmarking Results
- Conclusions

Simulating a SCADA Network

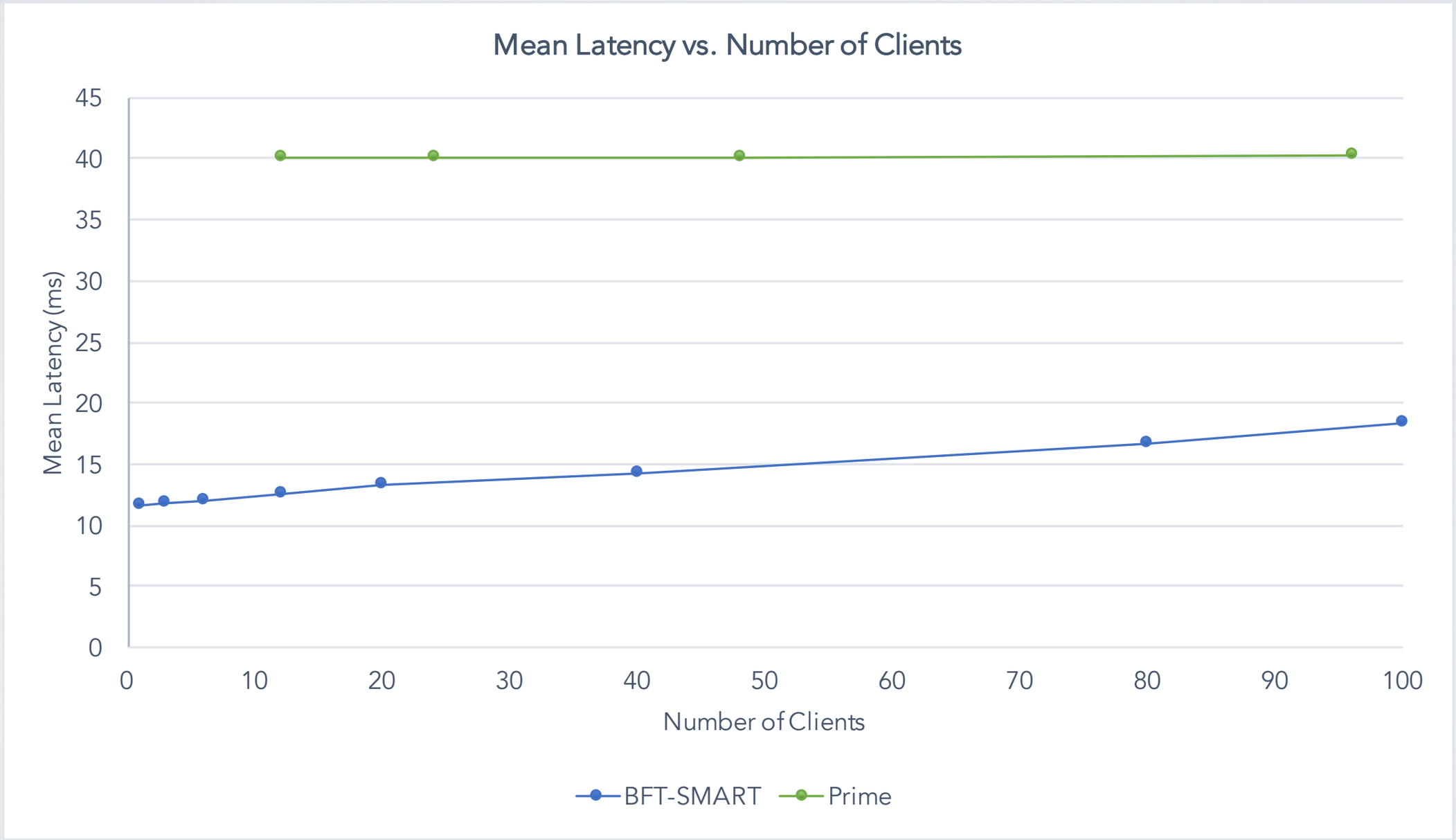
3 replicas per site

n = 12

f = 3



Normal-Case Latency



Normal-Case Latency

- Significantly lower with BFT-SMART, but increasing with number of clients
  - Matches expectations given fewer consensus rounds
- Constant with Prime, due to batch ordering on a preset interval of 20ms

Performance Attack Latency

- Tested 4 timeouts, chosen based on normal performance

    1. **8ms** (aggressive)
    2. **10ms** (conservative)

Performance Attack Latency

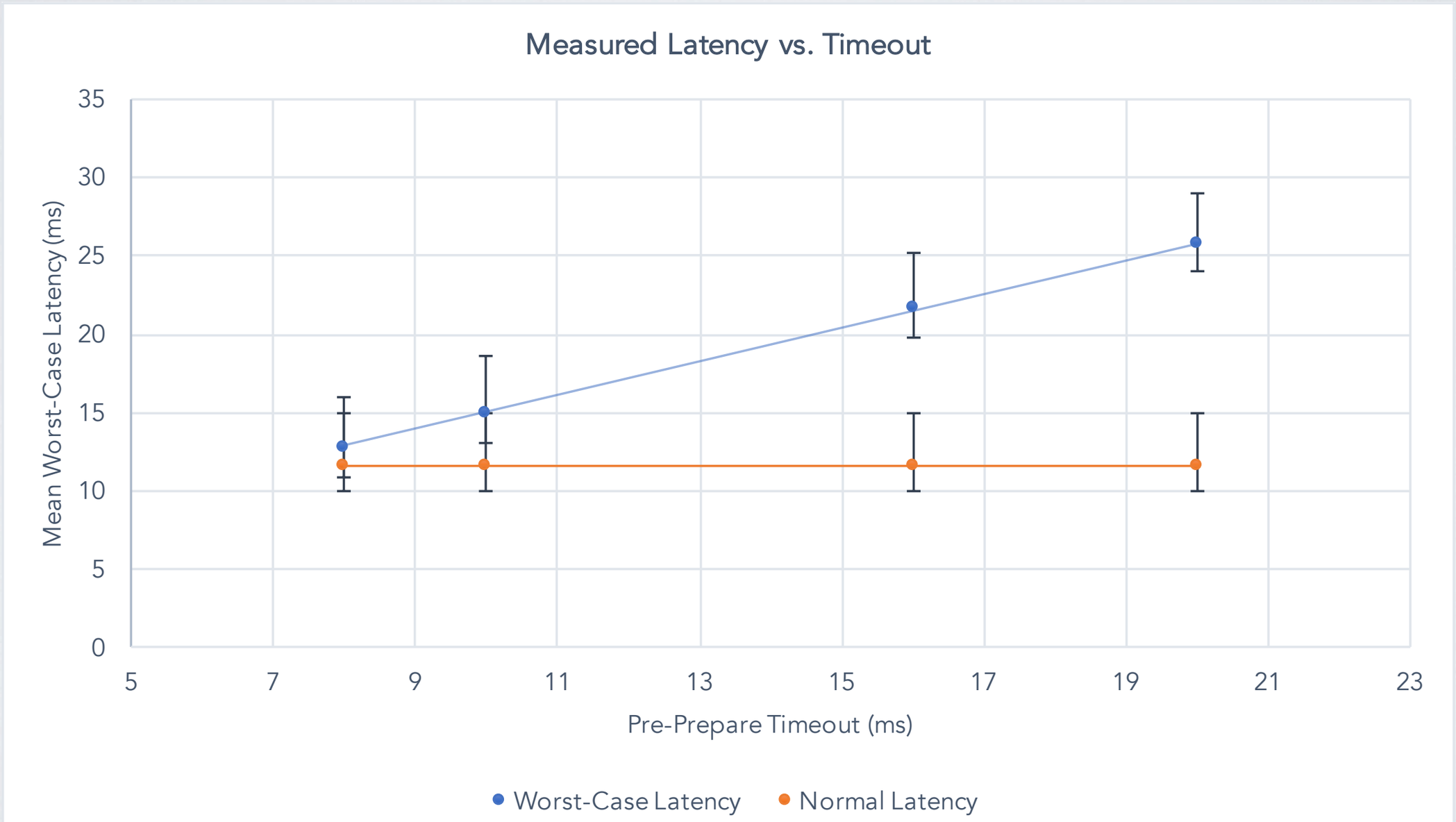
- Tested 4 timeouts, chosen based on normal performance

1. **8ms** (aggressive)
2. **10ms** (conservative)
3. **16ms** (aggressive, forwarding request at 8ms)
4. **20ms** (conservative, forwarding request at 10ms)

Performance Attack Latency

- Developed a malicious replica to delay sending *pre-prepare* messages as leader
- Experimentally maximized delay up to each view change timeout
- Measured worst-case latency seen by client under this condition

Performance Attack Latency



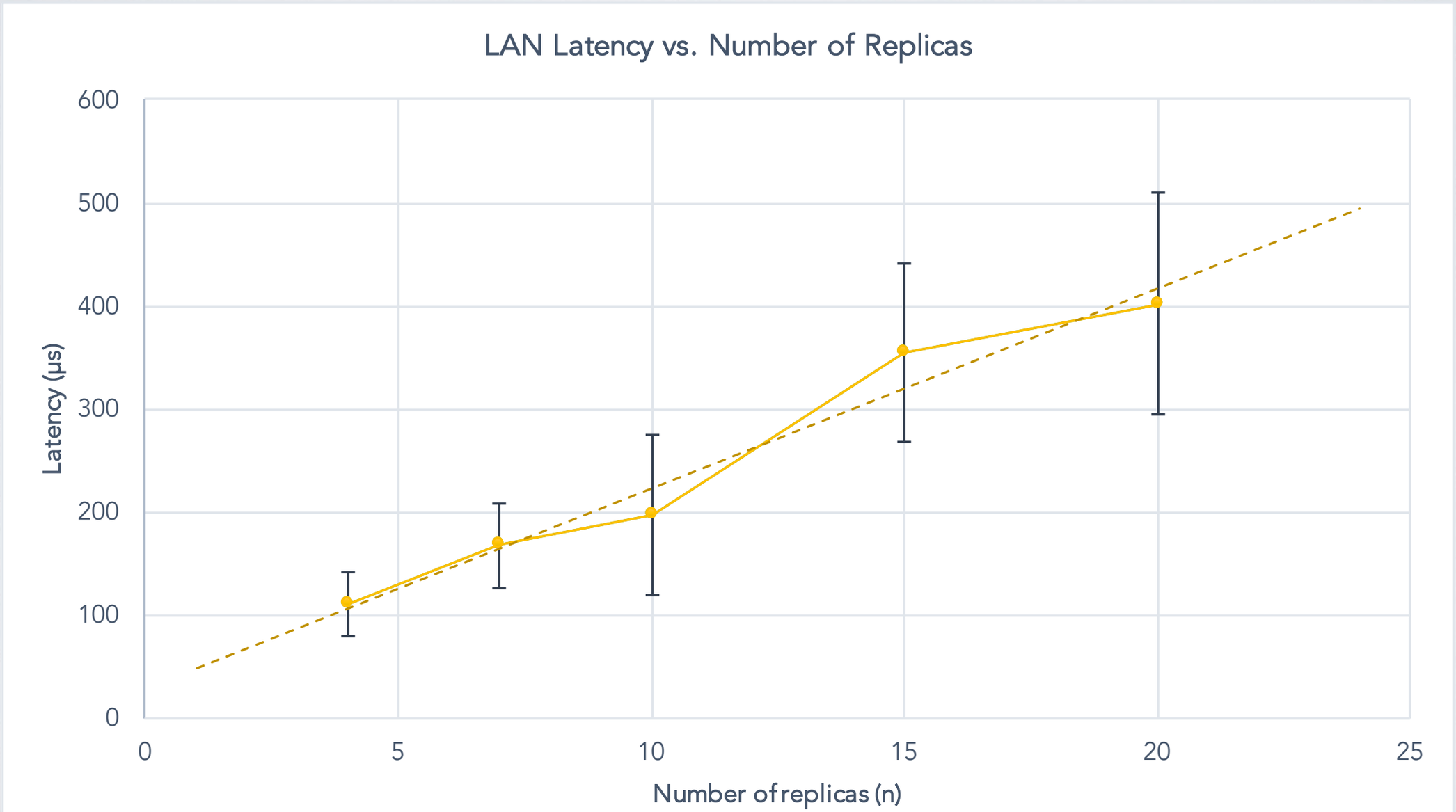
Performance Attack Latency

- With a tight timeout, performance degradation is minimal
- With a conservative timeout, performance degradation approaches 50% (26ms latency)
- In either case, lower than normal-case Prime and exceeds the required performance
- This performance attack would not pose a risk to the SCADA system

View Change

- **50-70ms** depending on number of pending requests
- Slow due to unoptimized serialization, data structures, taking up to **40ms**
- Sequential view changes are an issue with multiple faulty replicas
    - With $f \geq 3$, view change must be improved to meet the 200ms requirement
- Prime view changes are on the order of **60-90ms**

Scalability Overhead



Scalability Overhead

- Shows the computational overhead of increasing *n*
- Latency appears linear with *n*, and grows at a reasonable rate
- Actual latency determined by location of added replicas
    - Another geographic site vs. more replicas per site

- SCADA & Spire Overview
- High-Performance, Scalable Spire
- Trusted Platform Module
- Known Network Characteristics
- Evaluating BFT-SMART
- Benchmarking Results
- Conclusions

BFT-SMART: Pros & Cons

PROS

- Lightweight protocol & implementation
- Possible to apply aggressive timeout
- Low normal-case latency
- Support for dynamic state transfer, reconfiguration/recovery

CONS

- Latency increases with number of clients, concurrent requests
- High view change cost
- Java implementation

Prime: Pros & Cons

PROS

- Leader is not burdened by client requests
- Bounded performance guarantee under attack
- Latency remains constant as number of clients increases
- Measurements performed so replicas can adapt to network conditions

CONS

- 2 more consensus rounds per ordering
- High view change cost
- Significantly higher normal-case latency

Conclusions

- Strict limit on performance attacks possible with a lightweight protocol and bounded network latencies
- View change still a high cost, but could be optimized
- A viable path to scaling Spire
- However, BFT-SMART introduces some new issues

Conclusions: BFT-SMART

- BFT-SMART is a good implementation, but not exactly what we need
- Very good proof of concept that something with weaker guarantees than Prime could outperform Prime in this specific context using known network characteristics

Conclusions: Prime

- We want some of the features Prime has, specifically, network measurements and batching.
- We can live without Prime's expensive offloading of the leader - we can assume that the computers can do the intended job fast enough (need to measure how long it takes for a full update compared with how long it takes for immediate response in Prime).

Next Steps

- Consider diversity and client-server communication

- Interface with the Spines and SCADA hardware

- *Or, apply this approach to something new?*

Thank You

- To Yair, Tom, Amy and Trevor
- To the class
- To Alysson Bessani and the BFT-SMART group