Mitigating Synchronized Hardware Trojan Attacks in Smart Grids

Chenglu Jin, Lingyu Ren, Xubin Liu, Peng Zhang and Marten van Dijk

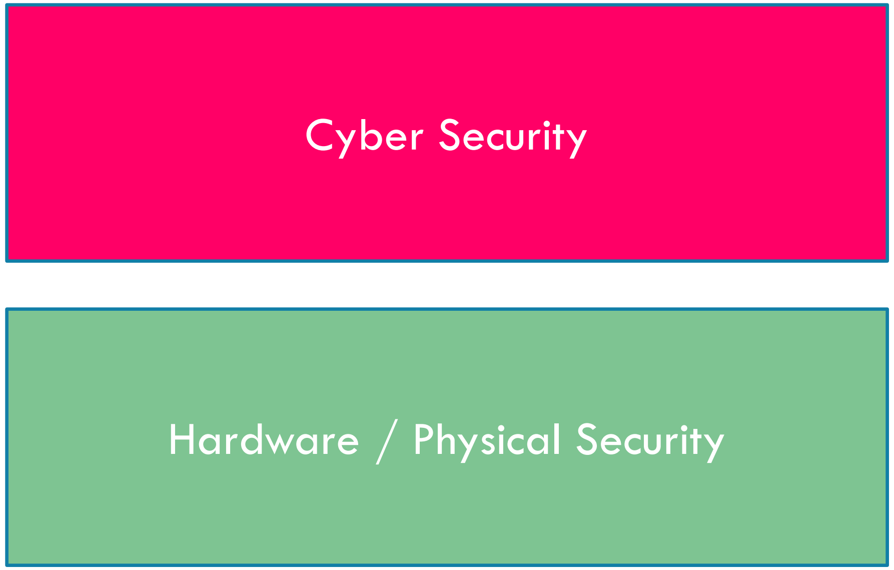
Secure Computation Laboratory
Department of Electrical & Computer Engineering
University of Connecticut

Email: chenglu.jin@uconn.edu

UCONN

Smart Grid Security

- Current researches are more focused on cyber security issues in smart grids.
- This implicitly assumes that the underlying hardware is trusted.
    - i.e. The hardware is doing and only doing what is supposed to do.

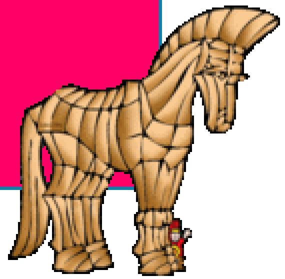


Smart Grid Security

- Current researches are more focused on cyber security issues of smart grids.
- This implicitly assumes that the underlying hardware is trusted.
    - i.e. The hardware is doing and only doing what is supposed to do.
- But this may not the case in the real life.
- Malicious hardware manufacturers can introduce malicious modifications, so called hardware Trojans, into their designs.
- We have to start questioning trustworthiness of the underlying hardware.

Cyber Security

Hardware / Physical Security



Hardware Trojans

Hardware Trojans

- Malicious modification in hardware. Can do anything malicious in theory.

Hardware Trojans

- Malicious modification in hardware. Can do anything malicious in theory.
- Hot research topic in hardware security community for more than one decade.

Hardware Trojans

- Malicious modification in hardware. Can do anything malicious in theory.
- Hot research topic in hardware security community for more than one decade.
- Current state-of-the-art detection methods (HaTCh, FANCI, etc.) require very large computational complexity to detect sophisticated hardware Trojan designs.

Hardware Trojans

- Malicious modification in hardware. Can do anything malicious in theory.
- Hot research topic in hardware security community for more than one decade.
- Current state-of-the-art detection methods (HaTCh, FANCI, etc.) require very large computational complexity to detect sophisticated hardware Trojan designs.
- New Trojan designs are EMERGING.

Hardware Trojans

- Malicious modification in hardware. Can do anything malicious in theory.
- Hot research topic in hardware security community for more than one decade.
- Current state-of-the-art detection methods (HaTCh, FANCI, etc.) require very large computational complexity to detect sophisticated hardware Trojan designs.
- New Trojan designs are EMERGING.
- Hardware Trojans/ backdoors were FOUND in military chips.
  - Reported in “Breakthrough silicon scanning discovers backdoor in military chip”, CHES’12

Hardware Trojans

- Malicious modification in hardware. Can do anything malicious in theory.
- Hot research topic in hardware security community for more than one decade.
- Current state-of-the-art detection methods (HaTCh, FANCI, etc.) require very large computational complexity to detect sophisticated hardware Trojan designs.
- New Trojan designs are EMERGING.
- Hardware Trojans/ backdoors were FOUND in military chips.
  - Reported in “Breakthrough silicon scanning discovers backdoor in military chip”, CHES’12

A hidden 'back door' in a computer chip could allow cyber-criminals a way to override and control computer systems on Boeing 787s

http://www.dailymail.co.uk/sciencetech/article-2152284/Could-vulnerable-chip-allow-hackers-Boeing-787-Back-door-allow-cyber-criminals-way-in.html#ixzz28fcdeOdm

Hardware Trojans

- Malicious modification in hardware. Can do anything malicious in theory.
- Hot research topic in hardware security community for more than one decade.
- Current state-of-the-art detection methods (HaTCh, FANCI, etc.) require very large computational complexity to detect sophisticated hardware Trojan designs.
- New Trojan designs are EMERGING.
- Hardware Trojans/ backdoors were FOUND in military chips.
  - Reported in “Breakthrough silicon scanning discovers backdoor in military chip”, CHES’12

A hidden 'back door' in a computer chip could allow cyber-criminals a way to override and control computer systems on Boeing 787s

- It is still very hard to completely eliminate/ detect hardware Trojans in a large chip.
- Instead, we minimize the damage of a hardware Trojan.

http://www.dailymail.co.uk/sciencetech/article-2152284/Could-vulnerable-chip-allow-hackers-Boeing-787-Back-door-allow-cyber-criminals-way-in.html#ixzz28fcdeOdm

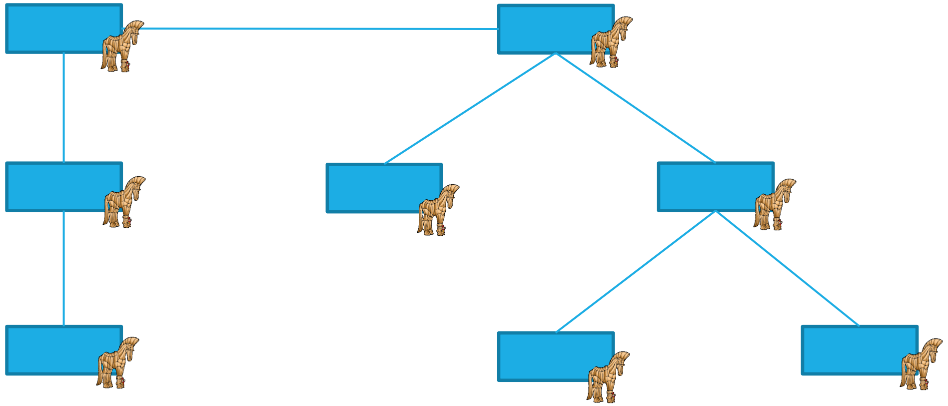
Hardware Trojans in Smart Grids

Synchronized

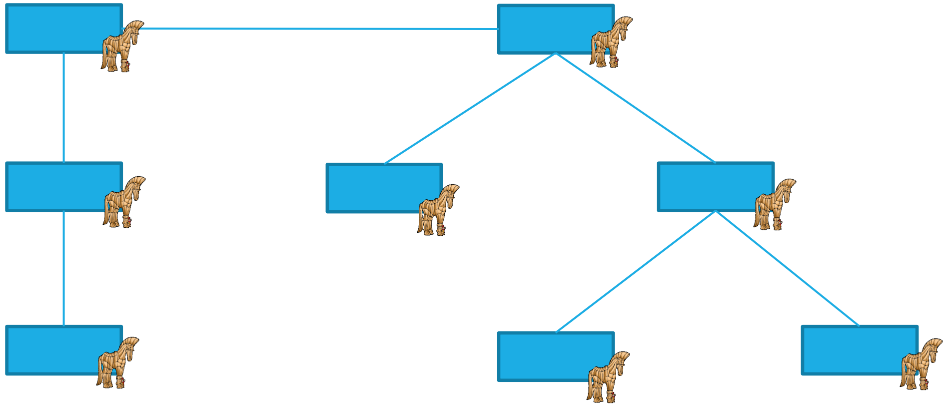
VS

Sporadic

Failure in large portion (or every node) of the smart grid at the same time



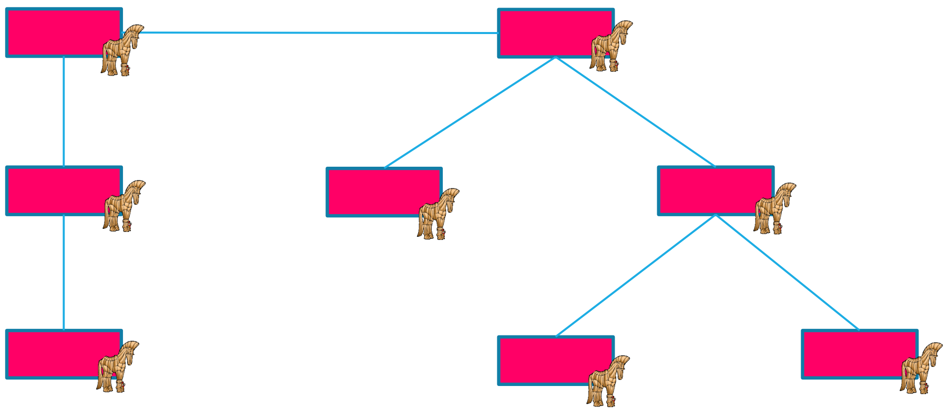
Sporadic single node failures.



Hardware Trojans in Smart Grids

Synchronized

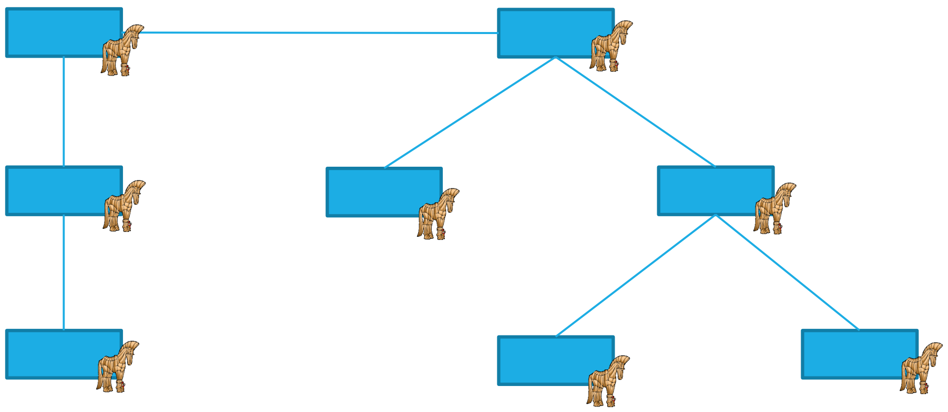
Failure in large portion (or every node) of the smart grid at the same time



VS

Sporadic

Sporadic single node failures.



UCONN

Hardware Trojans in Smart Grids

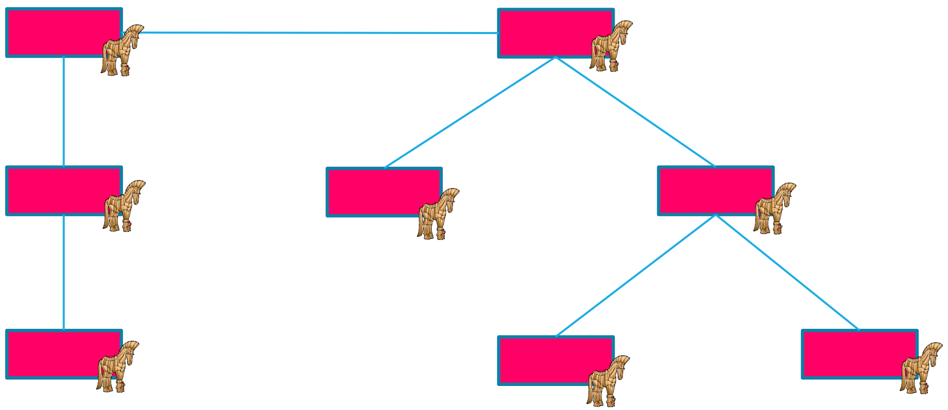
Synchronized

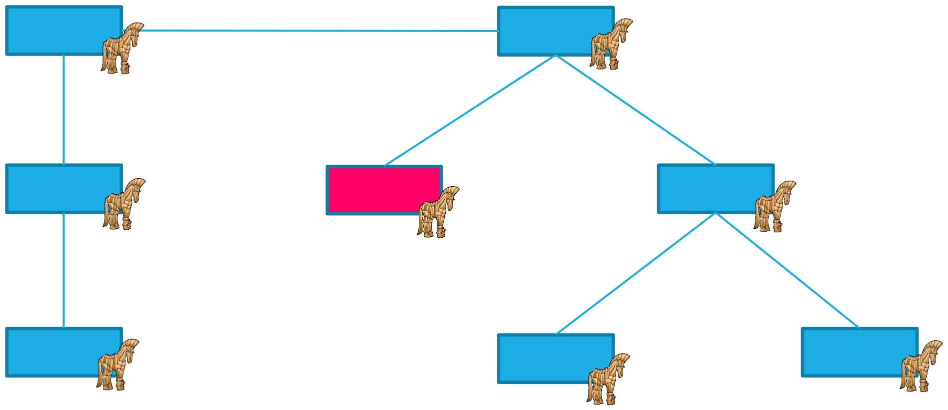
VS

Sporadic

Failure in large portion (or every node) of the smart grid at the same time

Sporadic single node failures.





UCONN

Hardware Trojans in Smart Grids

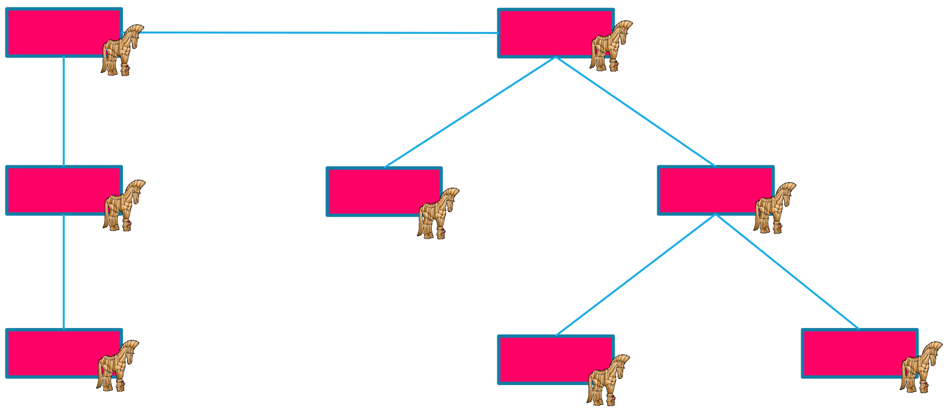
Synchronized

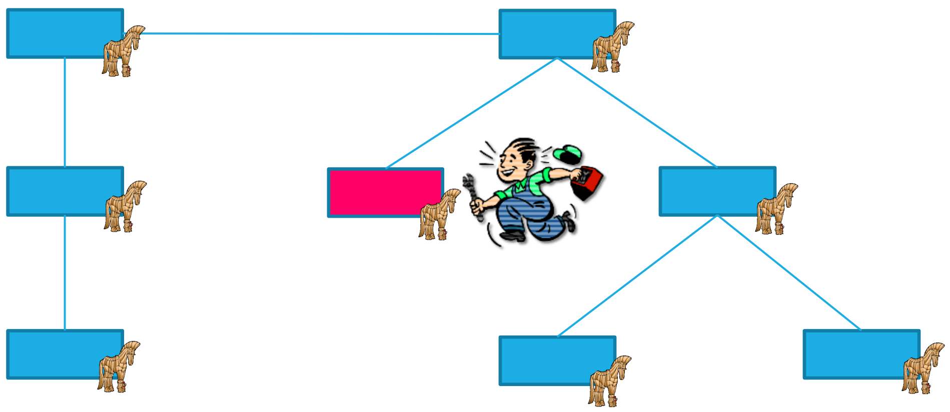
VS

Sporadic

Failure in large portion (or every node) of the smart grid at the same time

Sporadic single node failures.

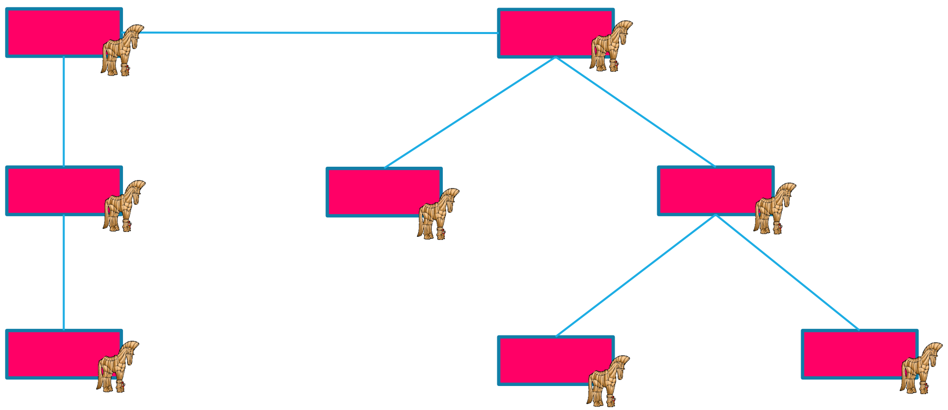




Hardware Trojans in Smart Grids

Synchronized

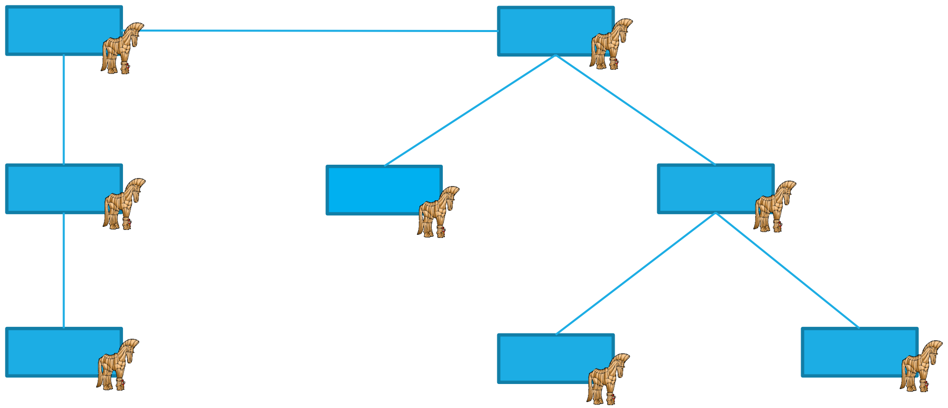
Failure in large portion (or every node) of the smart grid at the same time



VS

Sporadic

Sporadic single node failures.



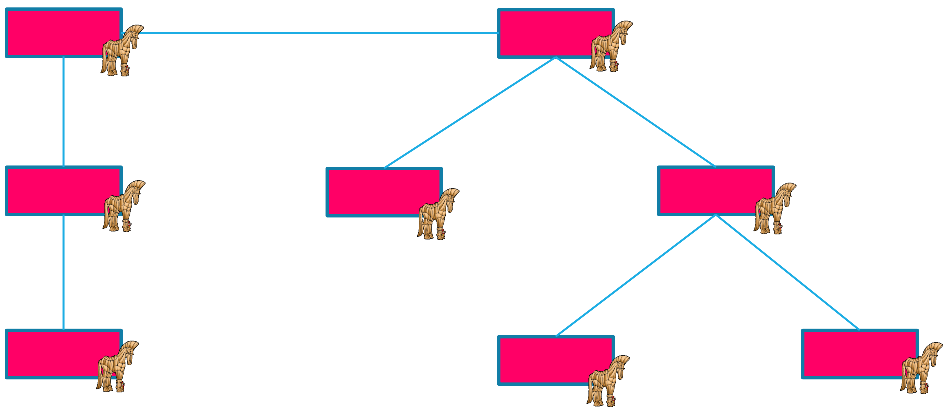
Hardware Trojans in Smart Grids

Synchronized

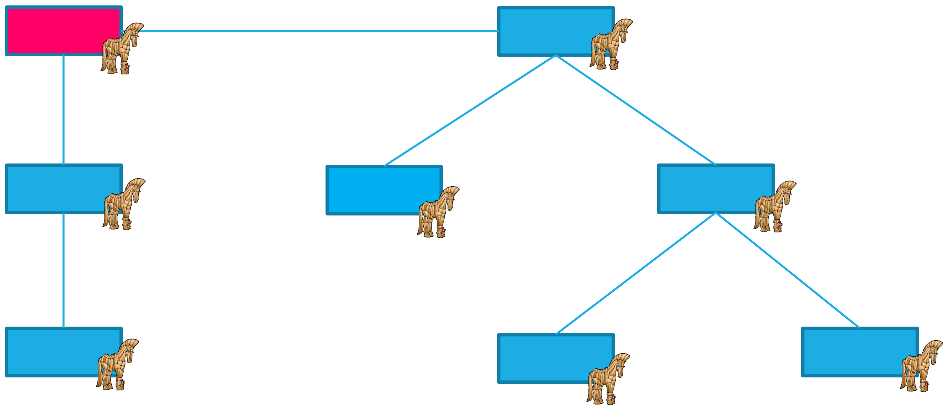
VS

Sporadic

Failure in large portion (or every node) of the smart grid at the same time



Sporadic single node failures.



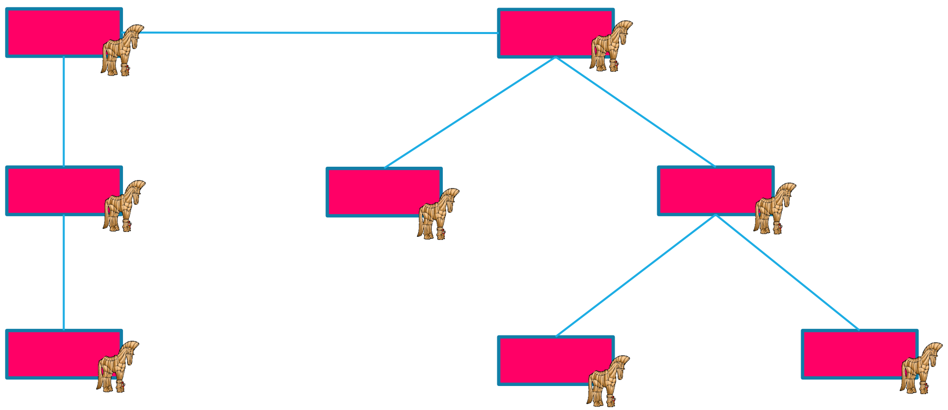
Hardware Trojans in Smart Grids

Synchronized

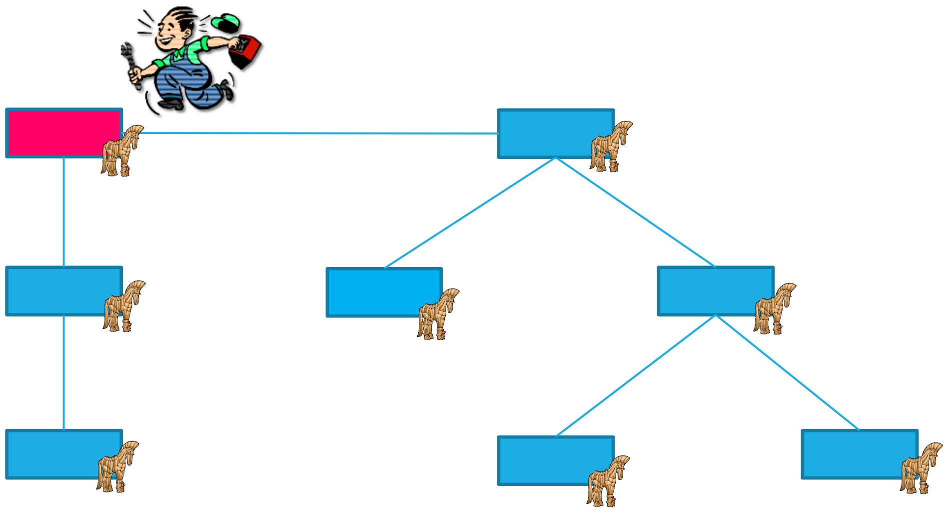
VS

Sporadic

Failure in large portion (or every node) of the smart grid at the same time



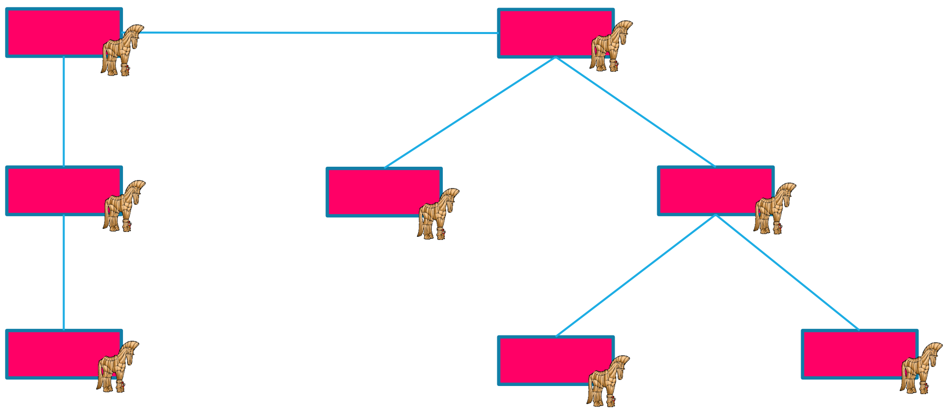
Sporadic single node failures.



Hardware Trojans in Smart Grids

Synchronized

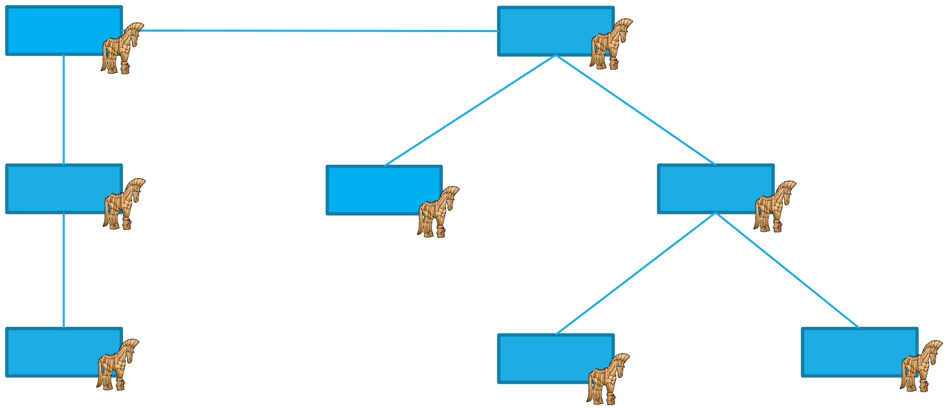
Failure in large portion (or every node) of the smart grid at the same time



VS

Sporadic

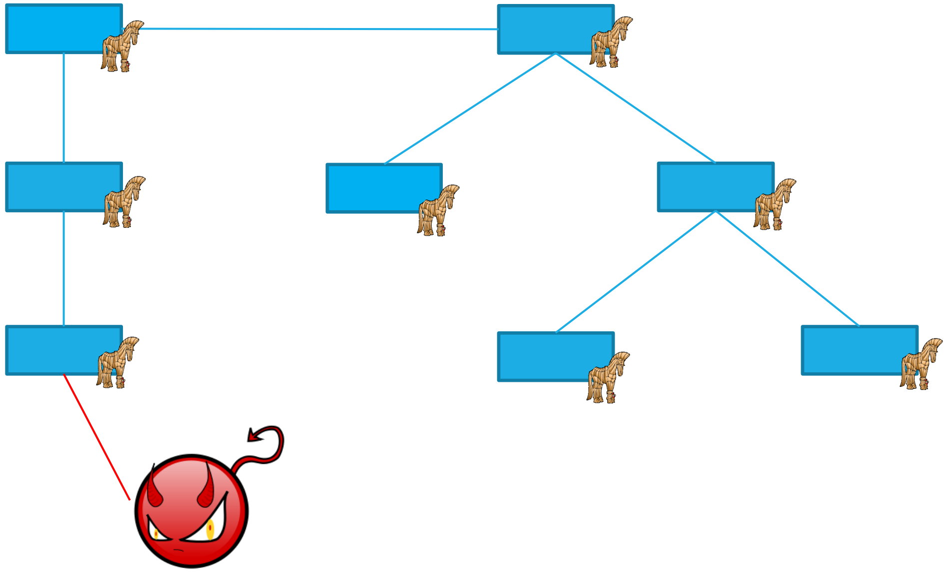
Sporadic single node failures.



Our mitigation strategy is to converte a synchronized hardware Trojan attack into sporadic single node failures.

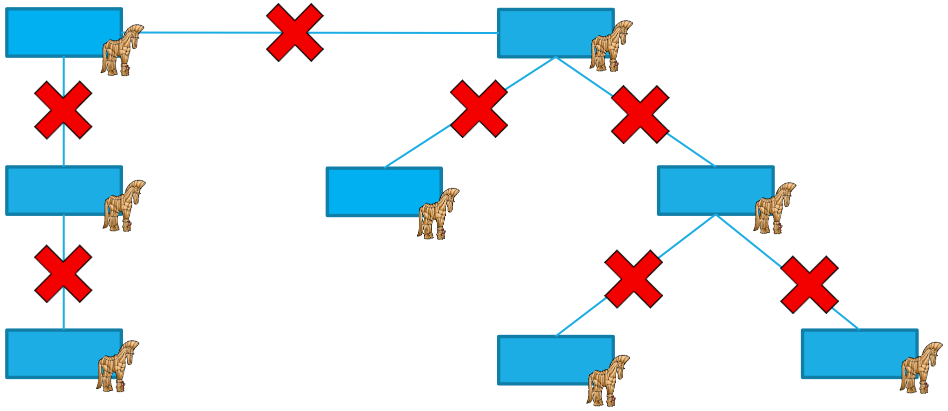
Online vs Offline Trojans

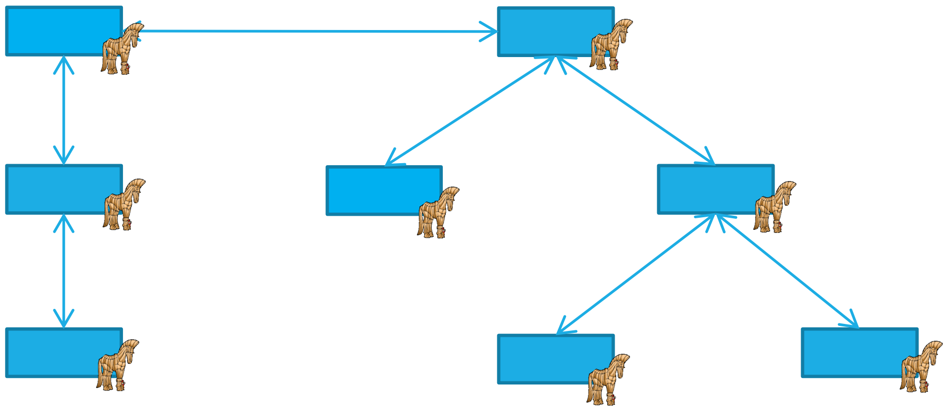
- Online Hardware Trojans:
- The attackers have connection and controllability of the chips (Trojans) after they are deployed.
- It also requires the attackers to first penetrate the network of smart grids to communicate with the Trojans and trigger the payloads.
- Needs to exploit software/ network vulnerabilities.
- Can be solved by software solutions.
- Open problem.



Offline Synchronized Hardware Trojans

- Type A: No inter-Trojan communications.
  - UTC provided by GPS module is a perfect way to synchronize each Trojan with one another.
- Type B: Allow inter-Trojan communications.
  - Trojans can communicate with one another via network or powerline to synchronize with each other.
  - Open problem, some interesting thoughts.





Outline

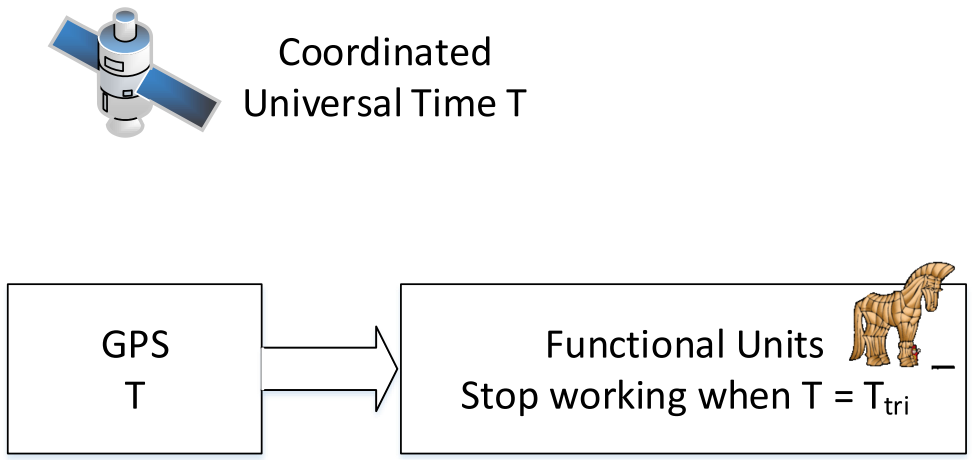
- Type A: No inter-Trojan communications.
  - Attack
  - Mitigation
- Type B: Allow inter-Trojan communications.
  - Attack
  - Possible Mitigation
- Risk Study

Type A Synchronized Attack

- Implemented in a simple killer switch.

Type A Synchronized Attack

- Implemented in a simple killer switch.
- In each critical node of a smart grid, the functional unit (e.g. PMU, RTU) which has a Trojan embedded can check whether the current time information provided by the GPS module is equal to a preset trigger time or not.



Coordinated Universal Time T

GPS

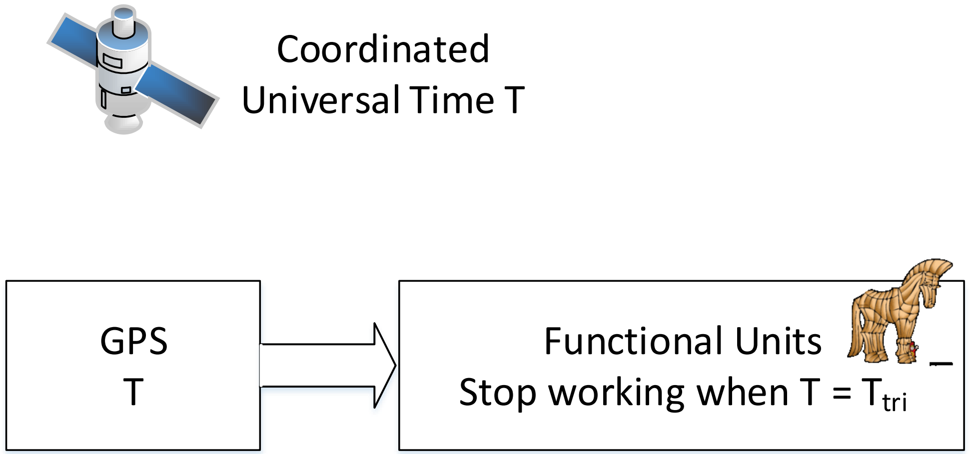
T

Functional Units

Stop working when T = T_{tri}

Type A Synchronized Attack

- Implemented in a simple killer switch.
- In each critical node of a smart grid, the functional unit (e.g. PMU, RTU) which has a Trojan embedded can check whether the current time information provided by the GPS module is equal to a preset trigger time or not.
- If all the Trojans have the same trigger time, then the entire power grid will shut down at the same time.



Coordinated Universal Time T

GPS

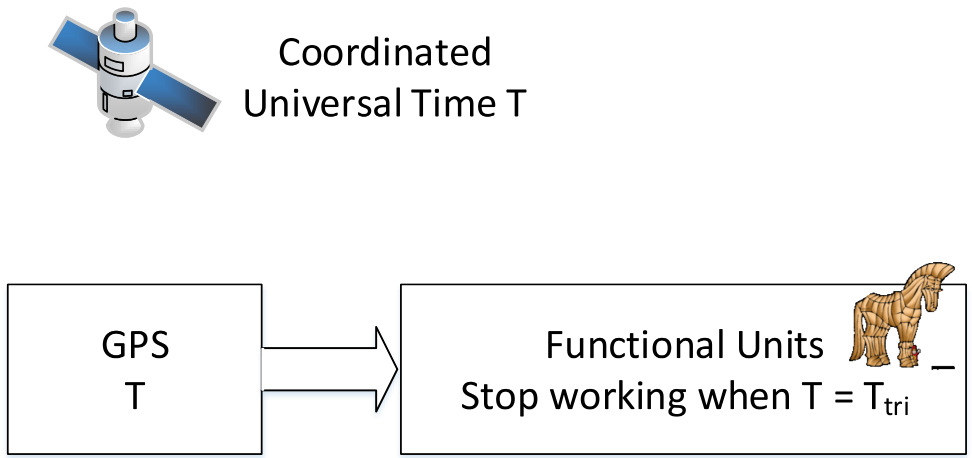
T

Functional Units

Stop working when $T = T_{tri}$

Type A Synchronized Attack

- Implemented in a simple killer switch.
- In each critical node of a smart grid, the functional unit (e.g. PMU, RTU) which has a Trojan embedded can check whether the current time information provided by the GPS module is equal to a preset trigger time or not.
- If all the Trojans have the same trigger time, then the entire power grid will shut down at the same time.
- Assumptions of Type A Trojans:



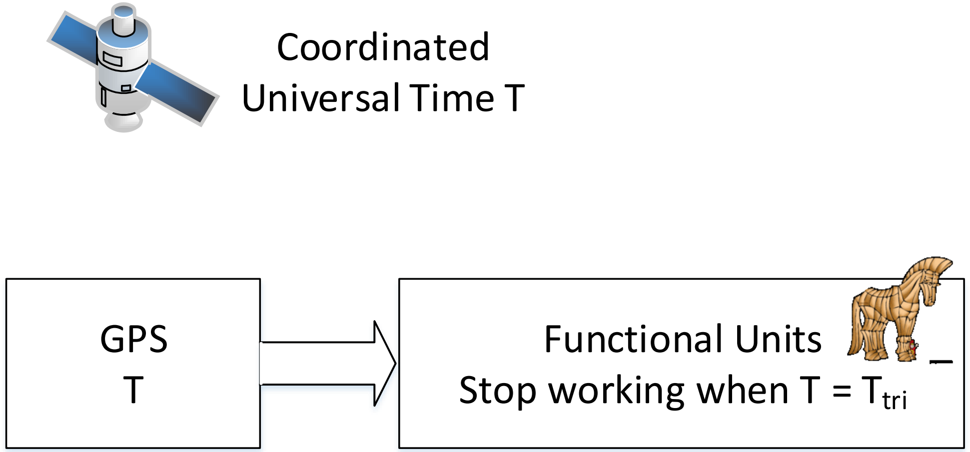
Coordinated Universal Time T

GPS
T

Functional Units
Stop working when T = T_{tri}

Type A Synchronized Attack

- Implemented in a simple killer switch.
- In each critical node of a smart grid, the functional unit (e.g. PMU, RTU) which has a Trojan embedded can check whether the current time information provided by the GPS module is equal to a preset trigger time or not.
- If all the Trojans have the same trigger time, then the entire power grid will shut down at the same time.
- Assumptions of Type A Trojans:
  - No GPS module in Trojans



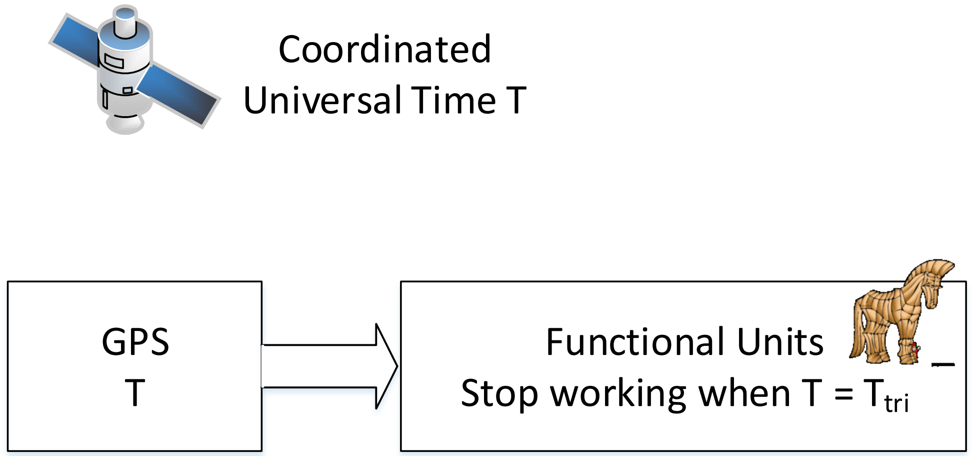
Coordinated Universal Time T

GPS
T

Functional Units
Stop working when $T = T_{tri}$

Type A Synchronized Attack

- Implemented in a simple killer switch.
- In each critical node of a smart grid, the functional unit (e.g. PMU, RTU) which has a Trojan embedded can check whether the current time information provided by the GPS module is equal to a preset trigger time or not.
- If all the Trojans have the same trigger time, then the entire power grid will shut down at the same time.
- Assumptions of Type A Trojans:
  - No GPS module in Trojans
  - Trojans do not access SW clock.



Coordinated
Universal Time T

GPS
T

Functional Units
Stop working when T = T_{tri}

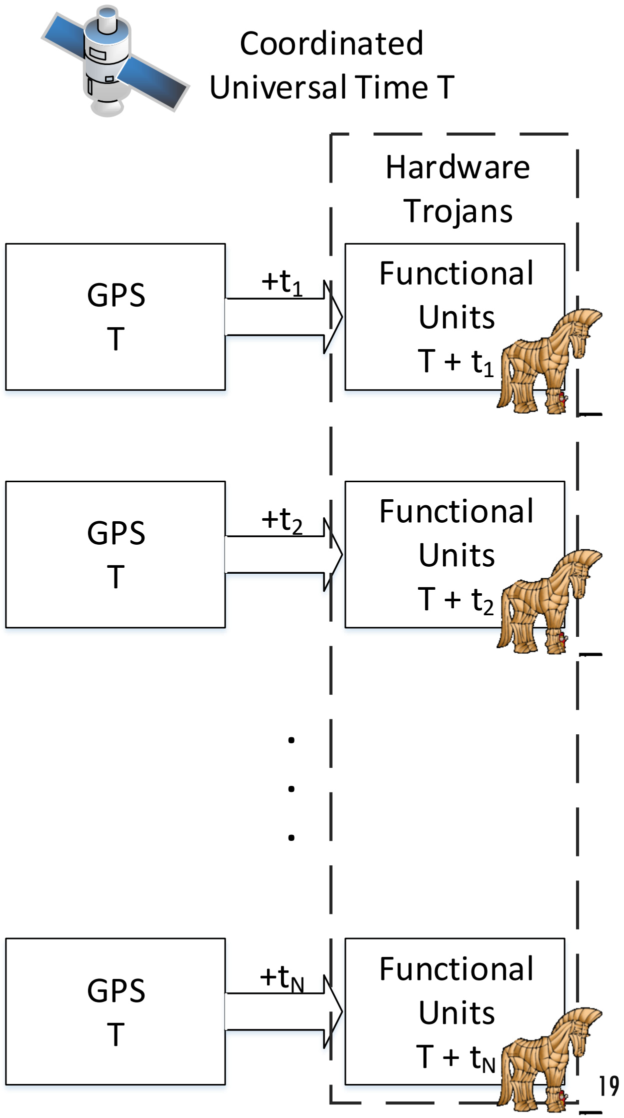
Mitigation for Type A Attack

Mitigation for Type A Attack

- Main idea: prevent Hardware Trojans from accessing to the correct time information.

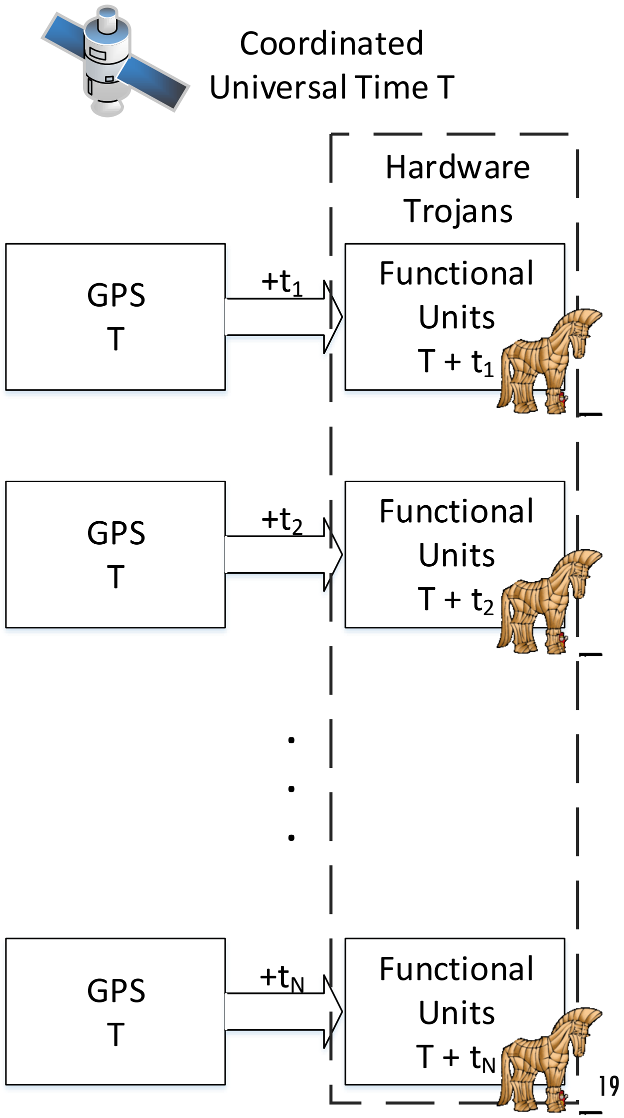
Mitigation for Type A Attack

- Main idea: prevent Hardware Trojans from accessing to the correct time information.
- We propose to enforce **each power grid node** to work in **an unique time domain** which has an unique time offset to the Universal Coordinated Time (UTC).



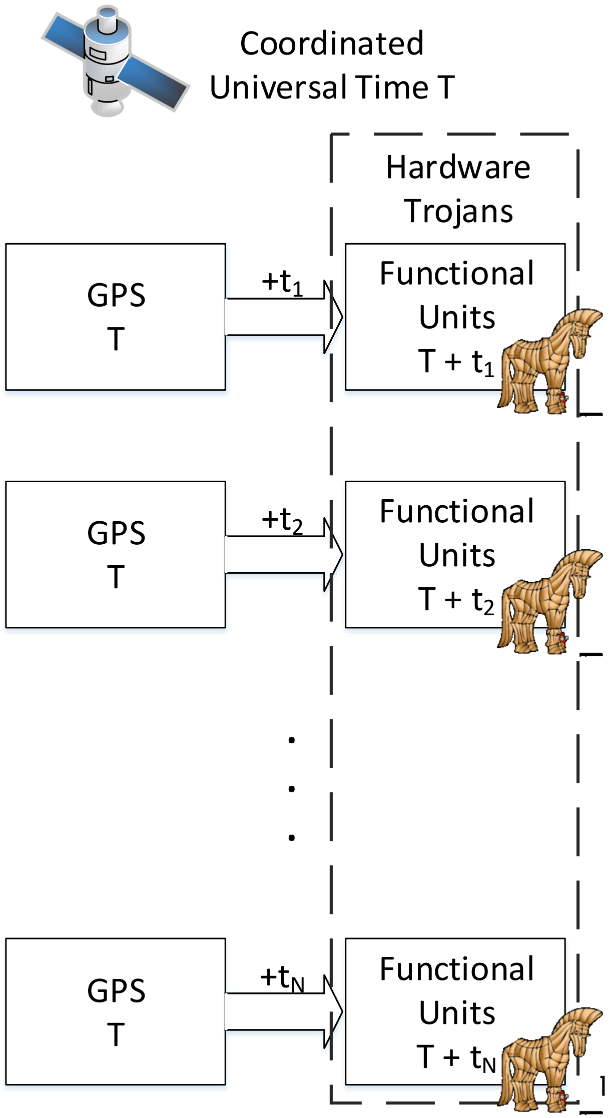
Mitigation for Type A Attack

- Main idea: prevent Hardware Trojans from accessing to the correct time information.
- We propose to enforce **each power grid node** to work in **an unique time domain** which has an unique time offset to the Universal Coordinated Time (UTC).
  - Time offsets are randomly generated, and fixed after initialization.



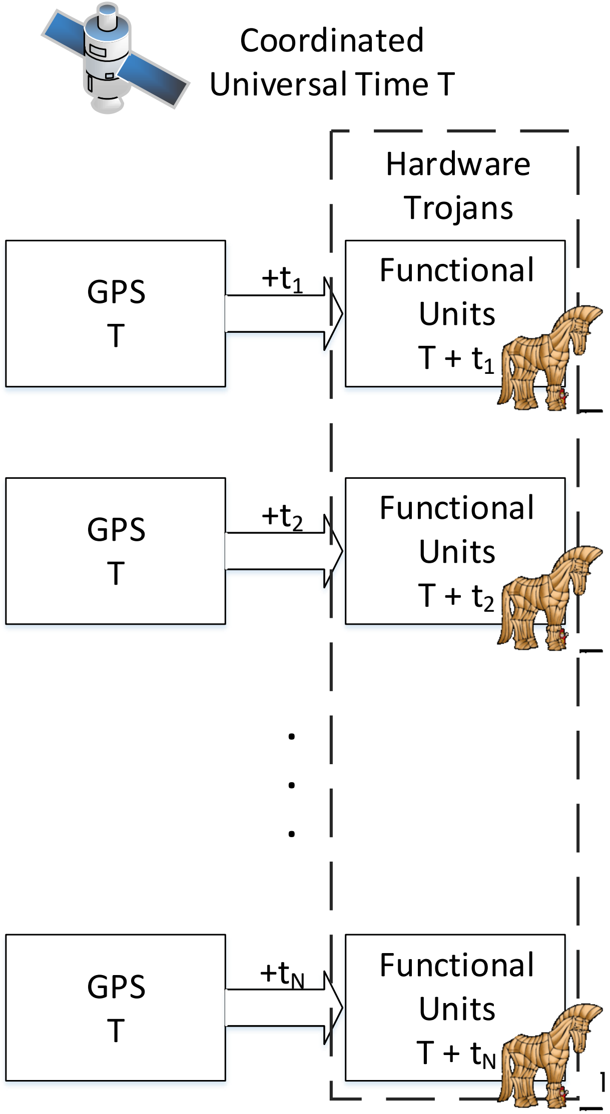
Mitigation for Type A Attack

- Main idea: prevent Hardware Trojans from accessing to the correct time information.
- We propose to enforce **each power grid node** to work in **an unique time domain** which has an unique time offset to the Universal Coordinated Time (UTC).
  - Time offsets are randomly generated, and fixed after initialization.
  - Time offsets do not need to be secret, because they are generated after the fabrication of Trojans



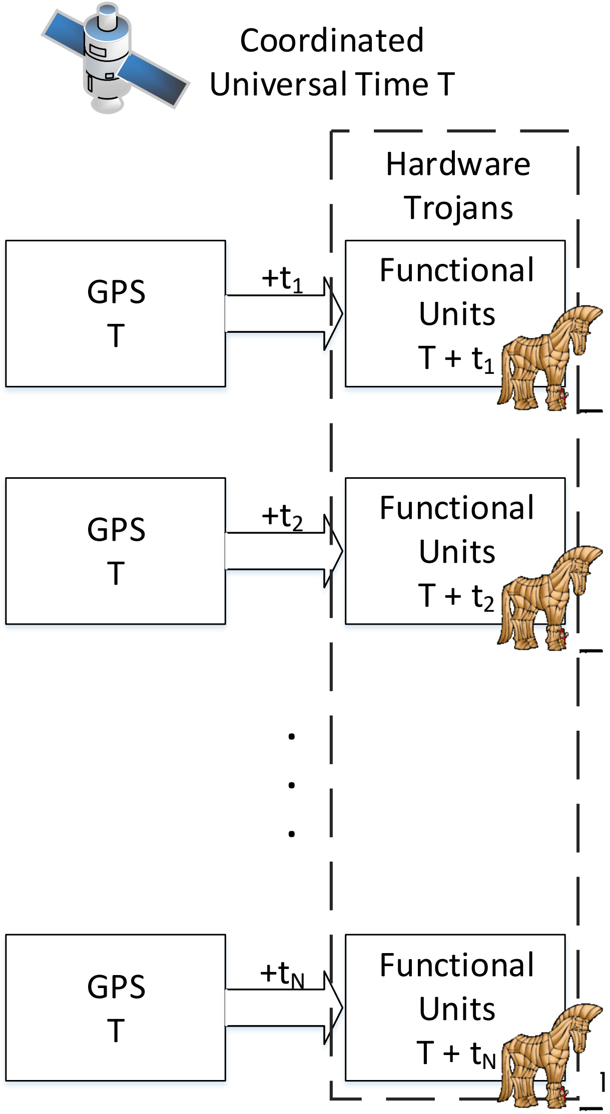
Mitigation for Type A Attack

- Main idea: prevent Hardware Trojans from accessing to the correct time information.
- We propose to enforce **each power grid node** to work in **an unique time domain** which has an unique time offset to the Universal Coordinated Time (UTC).
  - Time offsets are randomly generated, and fixed after initialization.
  - Time offsets do not need to be secret, because they are generated after the fabrication of Trojans
- A synchronized failure of all the nodes is converted to sporadic single node failures.



Mitigation for Type A Attack

- Main idea: prevent Hardware Trojans from accessing to the correct time information.
- We propose to enforce **each power grid node** to work in **an unique time domain** which has an unique time offset to the Universal Coordinated Time (UTC).
  - Time offsets are randomly generated, and fixed after initialization.
  - Time offsets do not need to be secret, because they are generated after the fabrication of Trojans
- A synchronized failure of all the nodes is converted to sporadic single node failures.
- Adding an additional interface between the GPS modules and the other functional units.
- We **reduce** the Trusted Computing Base (TCB) from all the modules in one node to **a trusted GPS module and a trusted additional interface.**



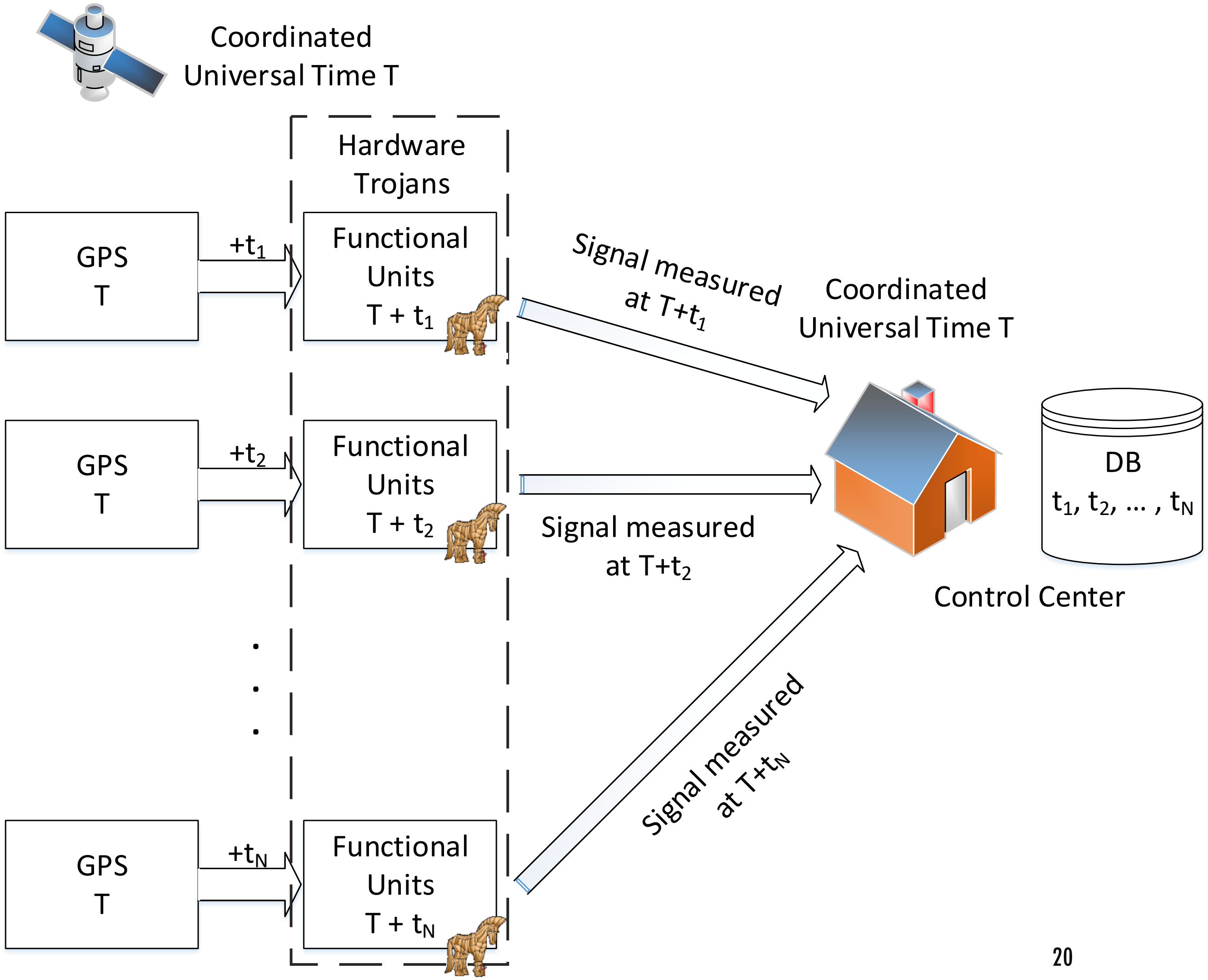
UCONN

Mitigation for Type A Attack

- Time information is critical for the normal functionalities of smart grids.
  - E.g. PMUs in different nodes need to do measurement using the same time reference.

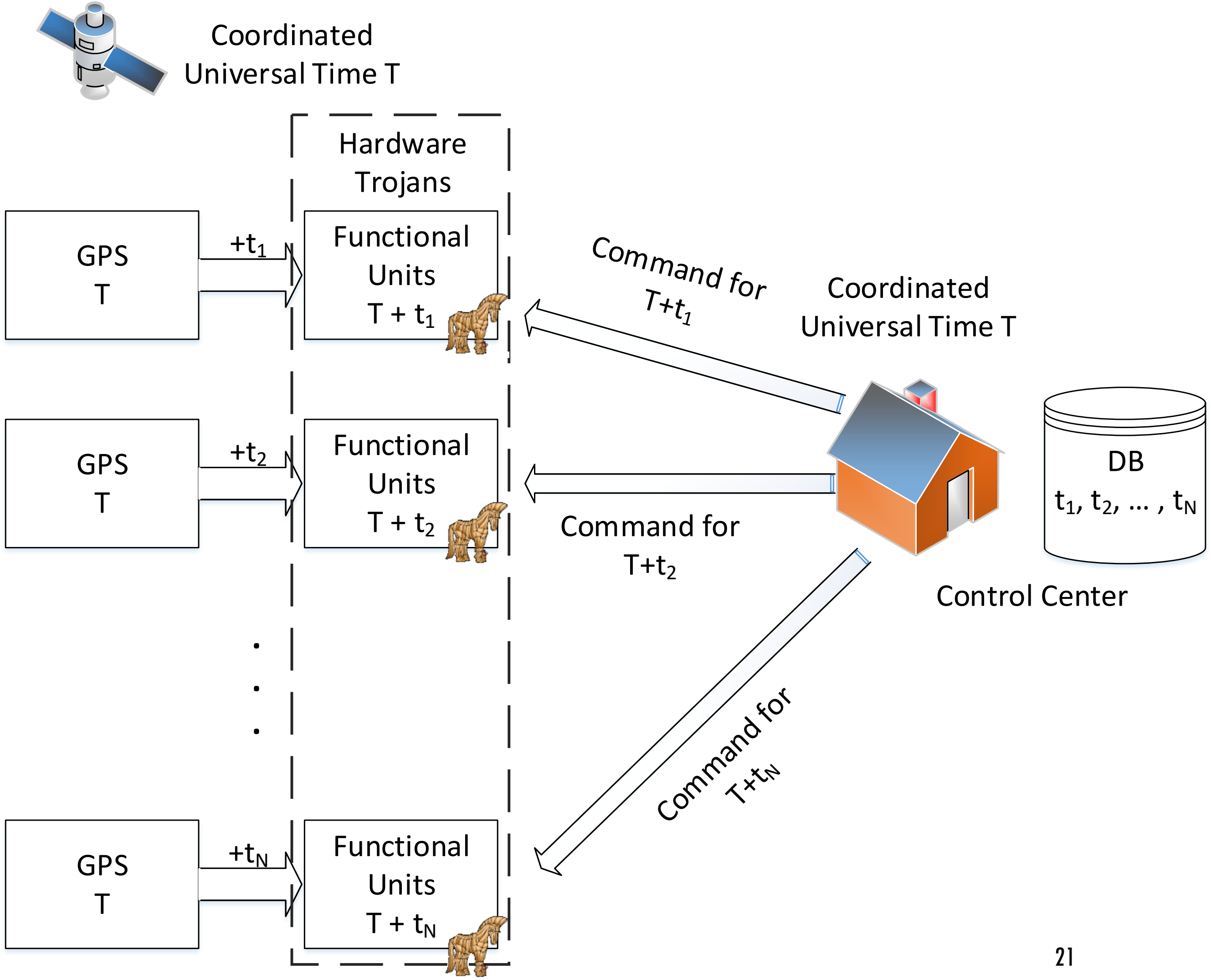
Mitigation for Type A Attack

- Time information is critical for the normal functionalities of smart grids.
  - E.g. PMUs in different nodes need to do measurement using the same time reference.
- When the interfaces are deployed, these time offsets are initialized randomly and sent to the database of the control center.
- The control center can adjust the timestamps of received messages and sent commands accordingly.



Mitigation for Type A Attack

- Time information is critical for the normal functionalities of smart grids.
  - E.g. PMUs in different nodes need to do measurement using the same time reference.
- When the interfaces are deployed, these time offsets are initialized randomly and sent to the database of the control center.
- The control center can adjust the timestamps of received messages and sent commands accordingly.



Outline

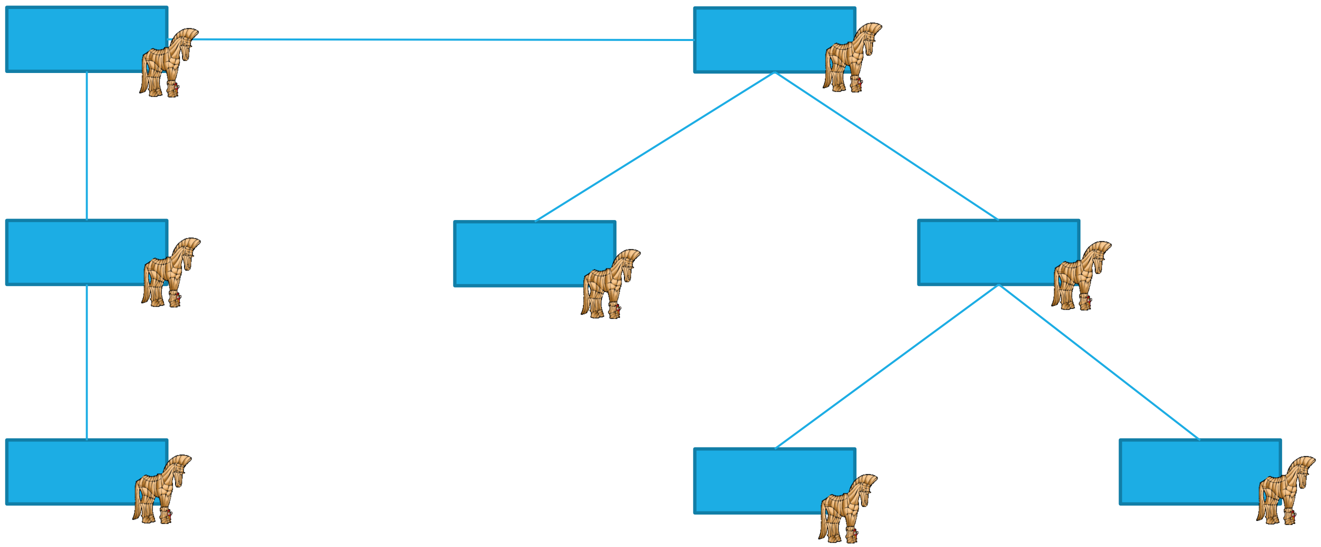
- Type A: No inter-Trojan communications.
  - Attack
  - Mitigation
- Type B: Allow inter-Trojan communications.
  - Attack
  - Possible Mitigation
- Risk Study

Type B Synchronized Attack

- If inter-Trojan communication is allowed in the smart grids, then these hardware Trojans can synchronize with one another just by sending out a synchronization signal.

Type B Synchronized Attack

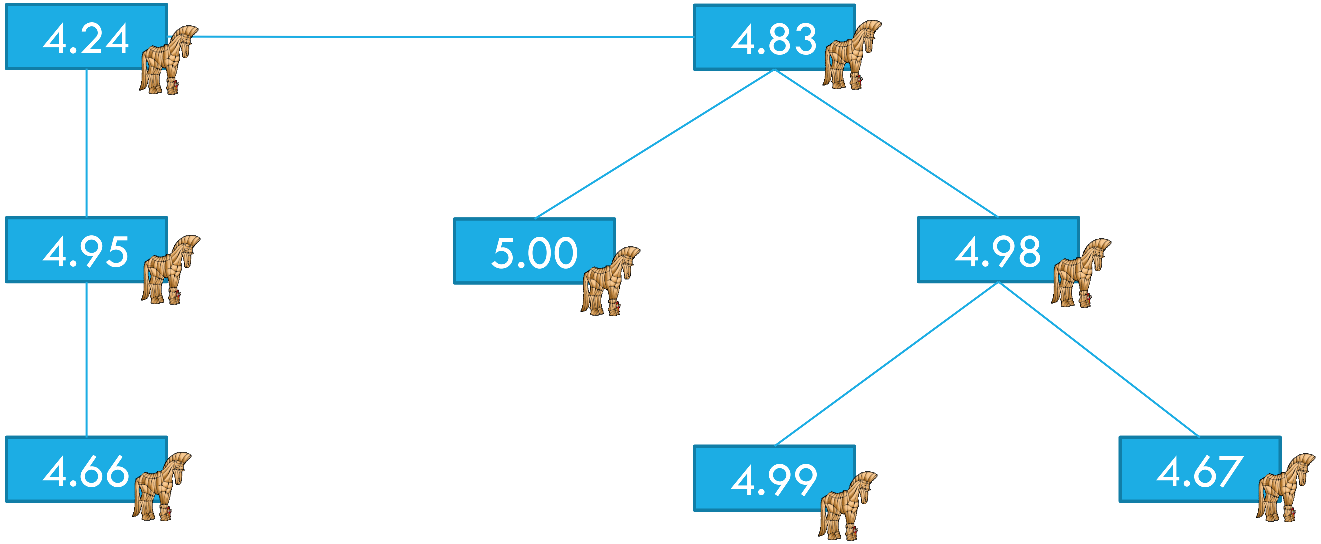
- If inter-Trojan communication is allowed in the smart grids, then these hardware Trojans can synchronize with one another just by sending out a synchronization signal.



Type B Synchronized Attack

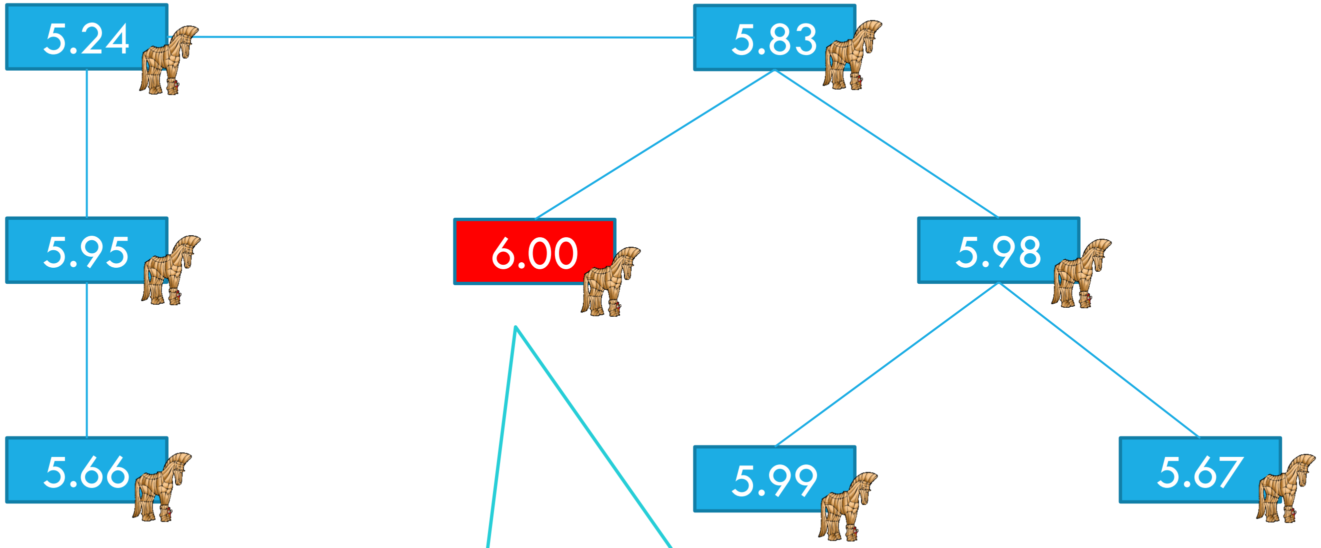
- If inter-Trojan communication is allowed in the smart grids, then these hardware Trojans can synchronized with one another just by sending out a synchronization signal.

Each Trojan has a counter embedded in it to count how long it has been deployed.



Type B Synchronized Attack

- If inter-Trojan communication is allowed in the smart grids, then these hardware Trojans can synchronized with one another just by sending out a synchronization signal.

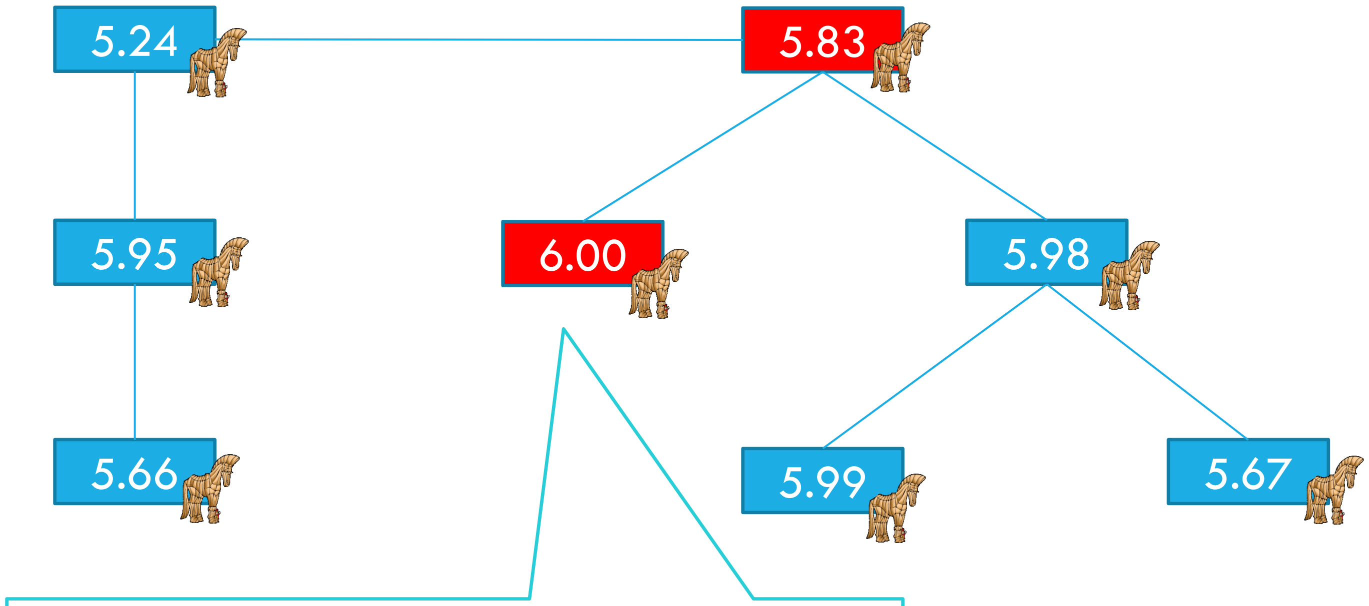


Each Trojan has a counter embedded in it to count how long it has been deployed.

E.g. If one of the counters reaches 6 years, then this Trojan becomes master and broadcasts one special message to trigger all the other Trojans in the network.

Type B Synchronized Attack

- If inter-Trojan communication is allowed in the smart grids, then these hardware Trojans can synchronized with one another just by sending out a synchronization signal.

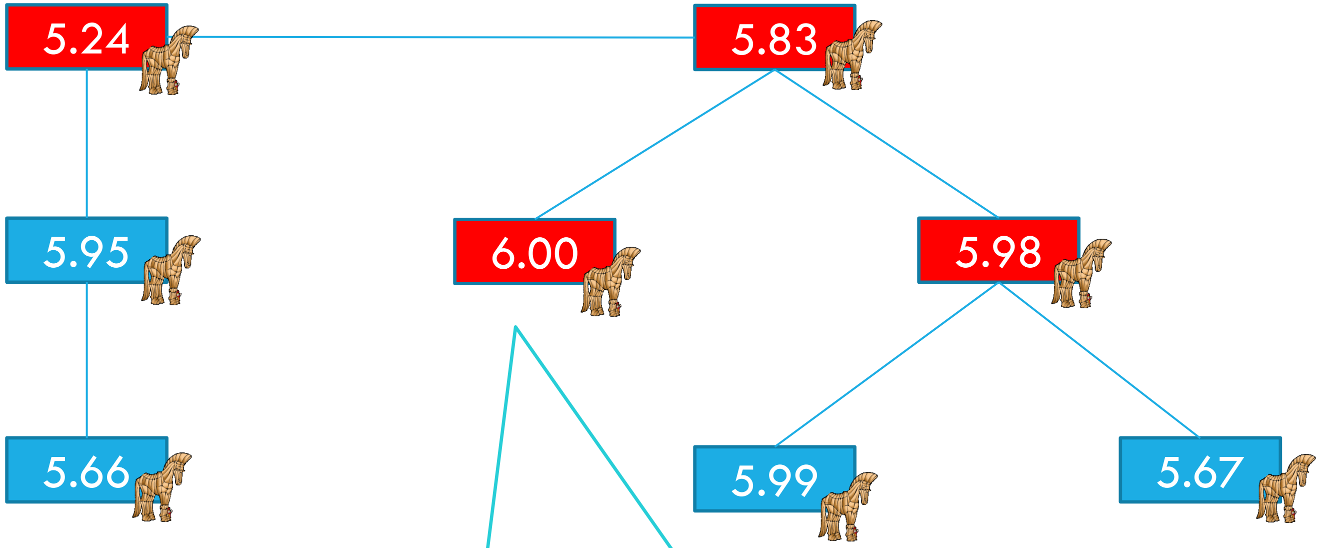


Each Trojan has a counter embedded in it to count how long it has been deployed.

E.g. If one of the counters reaches 6 years, then this Trojan becomes master and broadcasts one special message to trigger all the other Trojans in the network.

Type B Synchronized Attack

- If inter-Trojan communication is allowed in the smart grids, then these hardware Trojans can synchronized with one another just by sending out a synchronization signal.



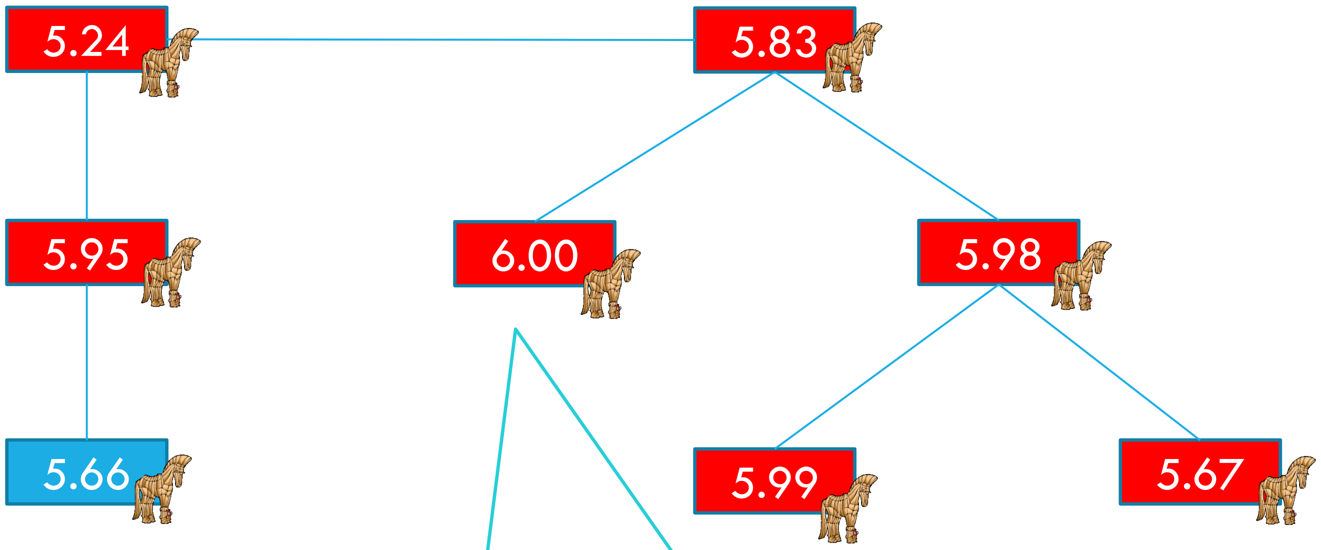
Each Trojan has a counter embedded in it to count how long it has been deployed.

E.g. If one of the counters reaches 6 years, then this Trojan becomes master and broadcasts one special message to trigger all the other Trojans in the network.

Type B Synchronized Attack

- If inter-Trojan communication is allowed in the smart grids, then these hardware Trojans can synchronized with one another just by sending out a synchronization signal.

Each Trojan has a counter embedded in it to count how long it has been deployed.

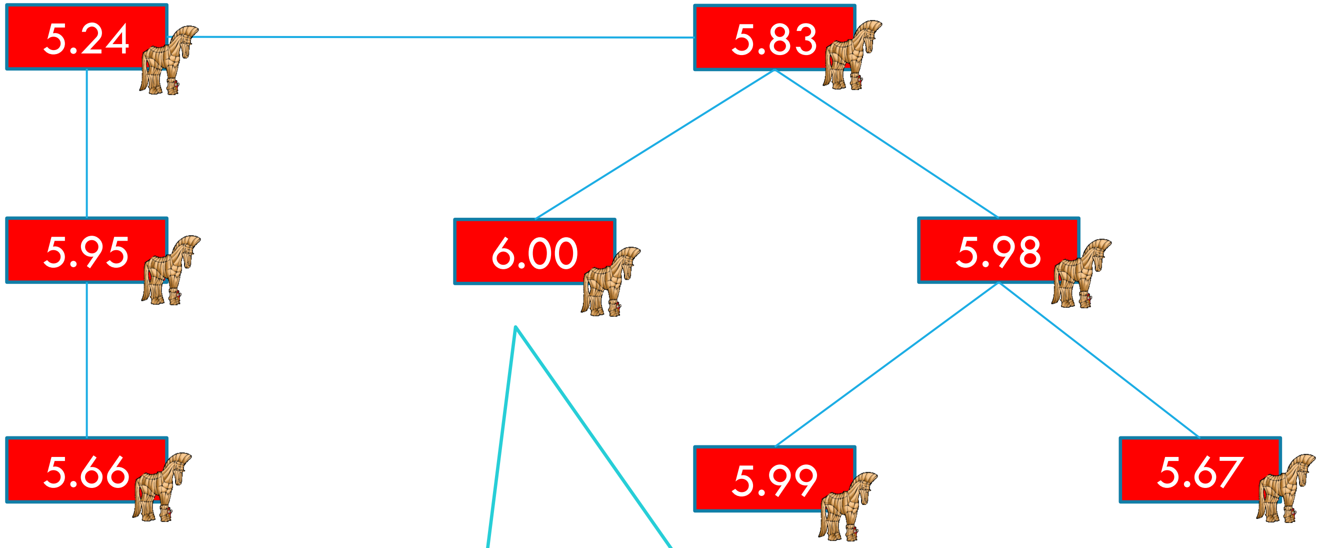


E.g. If one of the counters reaches 6 years, then this Trojan becomes master and broadcasts one special message to trigger all the other Trojans in the network.

Type B Synchronized Attack

- If inter-Trojan communication is allowed in the smart grids, then these hardware Trojans can synchronized with one another just by sending out a synchronization signal.

Each Trojan has a counter embedded in it to count how long it has been deployed.



E.g. If one of the counters reaches 6 years, then this Trojan becomes master and broadcasts one special message to trigger all the other Trojans in the network.

Possible Mitigation for Type B Attack

Possible Mitigation for Type B Attack

- Open problem.

Possible Mitigation for Type B Attack

- Open problem.
- Possible mitigations:

Possible Mitigation for Type B Attack

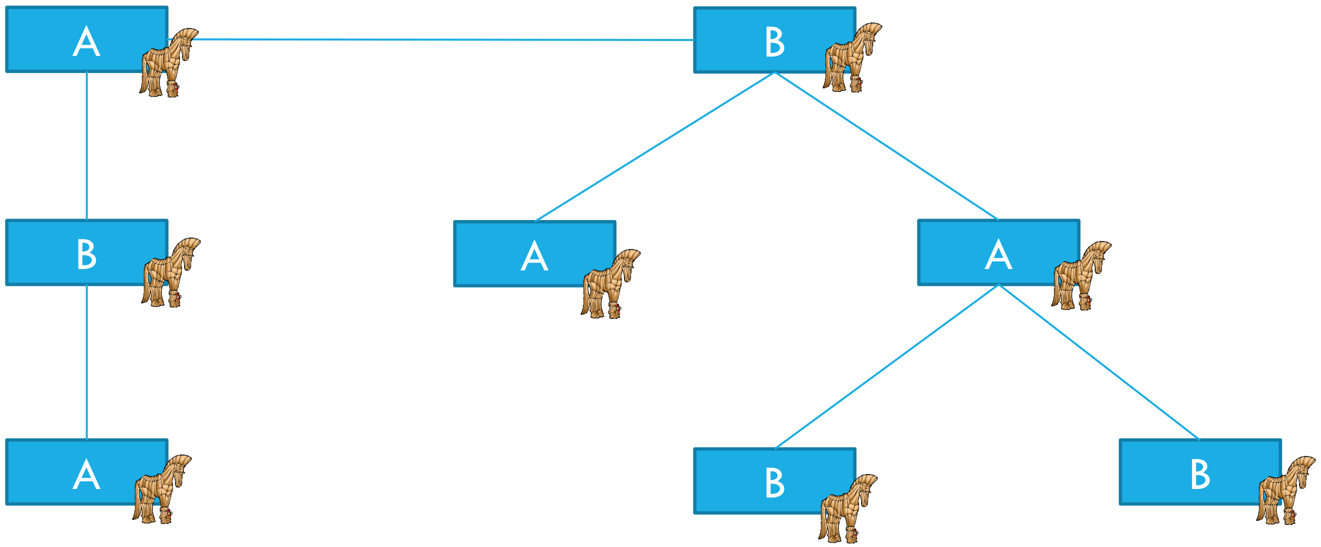
- Open problem.
- Possible mitigations:
  - Formally verified finite state machine in the communication module
    - Filter out all out-of-spec/ invalid messages.
    - But it does not prevent attackers from using a rarely happened valid message as a trigger.

Possible Mitigation for Type B Attack

- Open problem.
- Possible mitigations:
  - Formally verified finite state machine in the communication module
    - Filter out all out-of-spec/ invalid messages.
    - But it does not prevent attackers from using a rarely happened valid message as a trigger.
  - Split manufacturing.
    - Ask two manufacturers to fabricate the communication modules, assuming they do not collude with each other, and they cannot interpret one another's trigger message.
    - Neighboring nodes in the network topology originate from the different manufacturers.

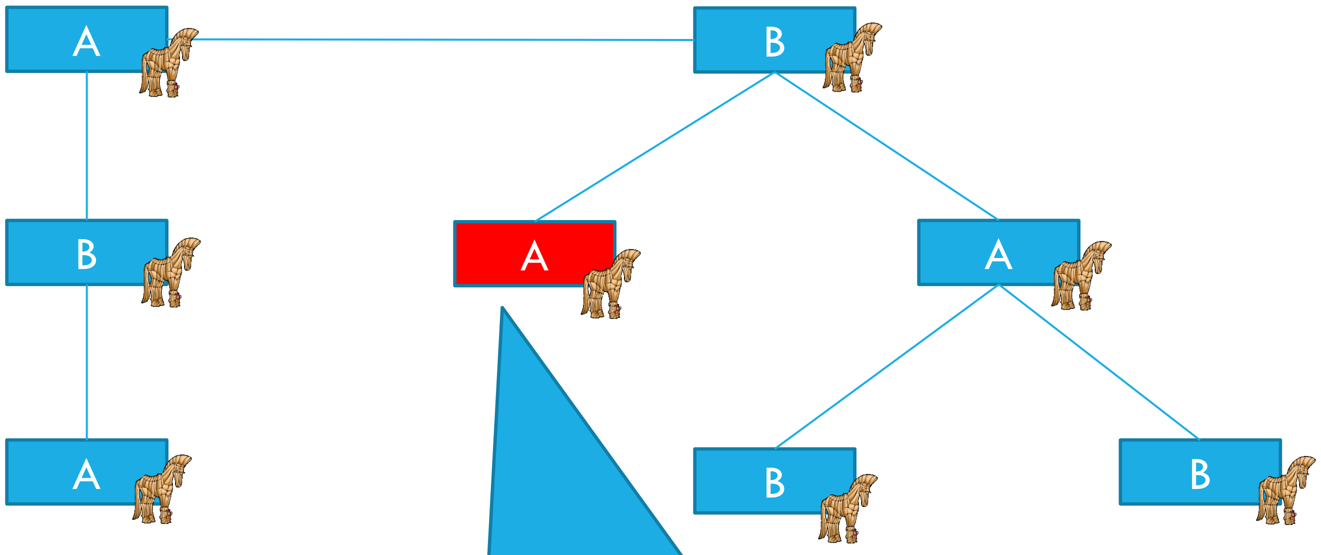
Mitigation for Type B Attack

- Suppose that we have devices from two non-colluding malicious manufacturers A and B.



Mitigation for Type B Attack

- Suppose that we have devices from two non-colluding malicious manufacturers A and B.



One of the Trojans is activated first, but ideally its broadcasting message cannot be interpreted by the neighboring nodes, so the package is dropped.

Outline

- Type A: No inter-Trojan communications.
  - Attack
  - Mitigation
- Type B: Allow inter-Trojan communications.
  - Attack
  - Possible Mitigation
- Risk Study

Risk Study

- Both online and offline hardware Trojan attacks are valid and possible in theory.
- In practice, a software attack is more likely to happen, because a large scale hardware attack is harder to prepare and launch.
- Hardware Trojans can be used to support software attacks, and the malicious behavior is controlled/ triggered by software.

Conclusion

- We studied the feasibility and risk of synchronized hardware Trojan attacks in smart grids. We conclude that hardware Trojan attacks are more difficult to launch a damaging attack in smart grids than software attacks.

Conclusion

- We studied the feasibility and risk of synchronized hardware Trojan attacks in smart grids. We conclude that hardware Trojan attacks are more difficult to launch a damaging attack in smart grids than software attacks.
- For Type A offline attack:
  - We propose to isolate the time domain of each node to prevent type A offline hardware Trojans from being activated at the same time.
  - It converts a failure of the entire power grid to sporadic single node failures.
  - Our solution reduces the TCB to a GPS module with a small additional interface in each node.
  - Applicable to the current power grid infrastructure.

Conclusion

- We studied the feasibility and risk of synchronized hardware Trojan attacks in smart grids. We conclude that hardware Trojan attacks are more difficult to launch a damaging attack in smart grids than software attacks.
- For Type A offline attack:
  - We propose to isolate the time domain of each node to prevent type A offline hardware Trojans from being activated at the same time.
  - It converts a failure of the entire power grid to sporadic single node failures.
  - Our solution reduces the TCB to a GPS module with a small additional interface in each node.
  - Applicable to the current power grid infrastructure.
- For Type B offline attack:

Conclusion

- We studied the feasibility and risk of synchronized hardware Trojan attacks in smart grids. We conclude that hardware Trojan attacks are more difficult to launch a damaging attack in smart grids than software attacks.
- For Type A offline attack:
  - We propose to isolate the time domain of each node to prevent type A offline hardware Trojans from being activated at the same time.
  - It converts a failure of the entire power grid to sporadic single node failures.
  - Our solution reduces the TCB to a GPS module with a small additional interface in each node.
  - Applicable to the current power grid infrastructure.
- For Type B offline attack:
  - Open problem.
  - Possible mitigations: Formally verified communication modules, Split Manufacturing.

Conclusion

- We studied the feasibility and risk of synchronized hardware Trojan attacks in smart grids. We conclude that hardware Trojan attacks are more difficult to launch a damaging attack in smart grids than software attacks.
- For Type A offline attack:
  - We propose to isolate the time domain of each node to prevent type A offline hardware Trojans from being activated at the same time.
  - It converts a failure of the entire power grid to sporadic single node failures.
  - Our solution reduces the TCB to a GPS module with a small additional interface in each node.
  - Applicable to the current power grid infrastructure.
- For Type B offline attack:
  - Open problem.
  - Possible mitigations: Formally verified communication modules, Split Manufacturing.

Thank you!