Thermal Covert Channels on Multicore Systems

Ramya Jayaram Masti*, Devendra Rai+, Aanjhan Ranganathan*,
Christian Müller+, Lothar Thiele+, Srdjan Čapkun*

* Institute of Information Security
+ Computer Engineering and Networks Laboratory

ETH zürich

Virtual machine used to steal crypto keys from other VM on same server

New technique could pierce a key defense found in cloud environments.

Scientist-devised crypto attack could one day steal secret Bitcoin keys

Technique exposes weaknesses not only in Bitcoin but also in OpenSSL.

Researchers can slip an undetectable trojan into Intel's Ivy Bridge CPUs

New technique bakes super stealthy hardware trojans into chip silicon.

Covert channel tool hides data in IPv6

But VoodooNet not new magic

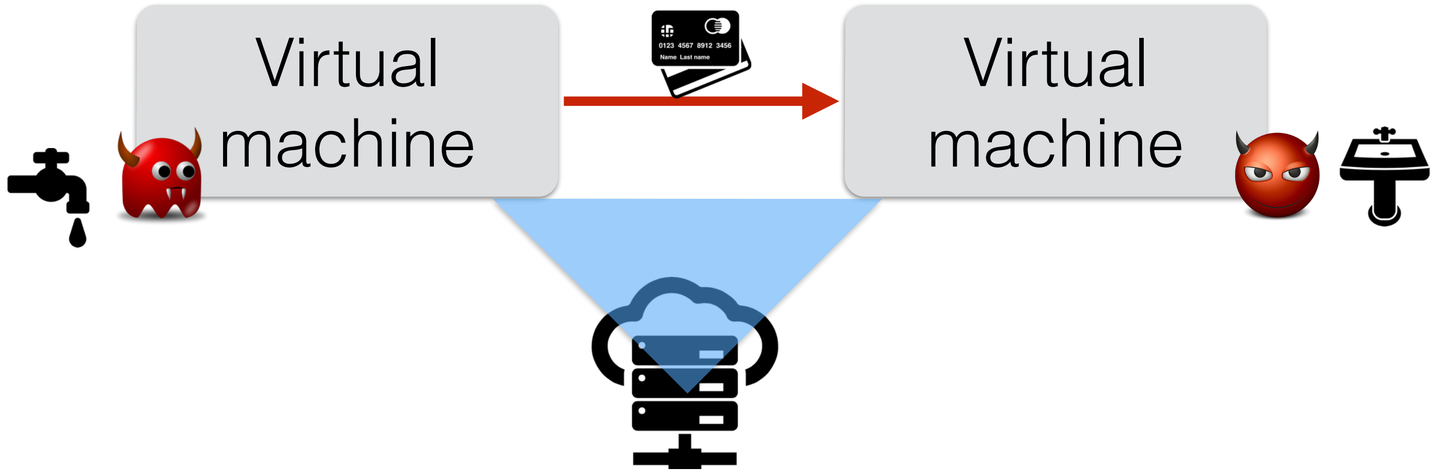
NSA uses covert radio transmissions to monitor thousands of bugged computers

And a Federal District Court judge says Surveillance Court should not be changed.

Scientist-developed malware prototype covertly jumps air gaps using inaudible sound

Malware communicates at a distance of 65 feet using built-in mics and speakers.

Overt and Covert Communication Channels



Overt channels

- Through intended data containers (e.g, IPC, files)
- Typically detected using information flow tracking and tainting

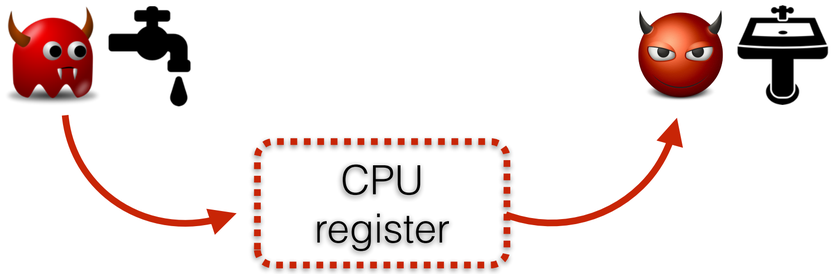
Covert channels

- Through a channel not designed for any kind of information transfer (e.g., system registers, cache access time)
- Hard to detect.

Covert Communication Channels

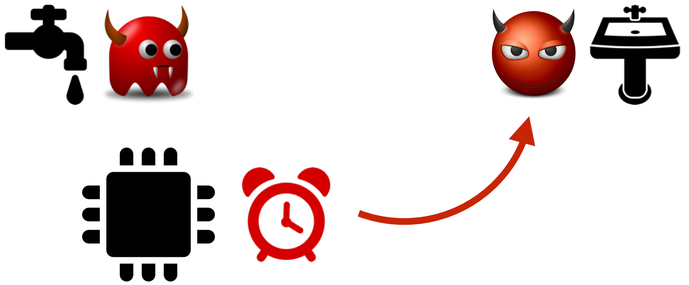
Storage Channels

- Use shared system storage areas (e.g., registers) to communicate data.
- Raising exceptions, file attributes etc.



Timing Channels

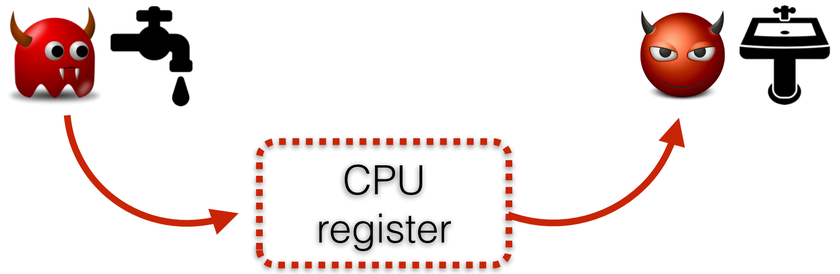
- Modifying the time taken to accomplish a task.
- Cache access times, CPU time etc.



Covert Communication Channels

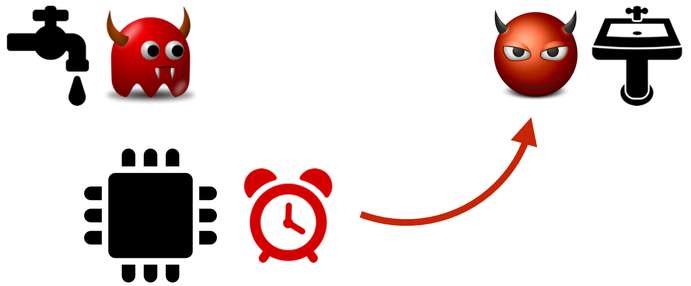
Storage Channels

- Use shared system storage areas (e.g., registers) to communicate data.
- Raising exceptions, file attributes etc.



Timing Channels

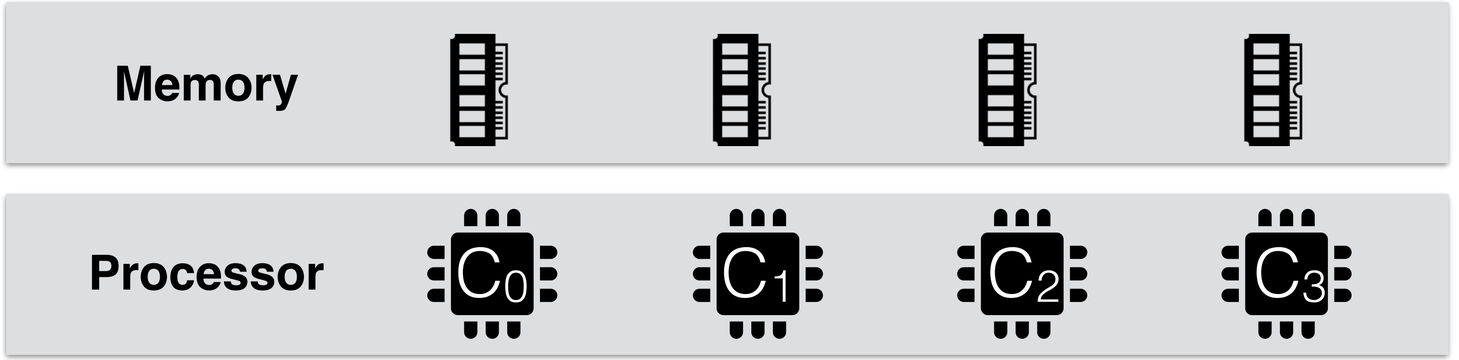
- Modifying the time taken to accomplish a task.
- Cache access times, CPU time etc.



Its all about sharing!

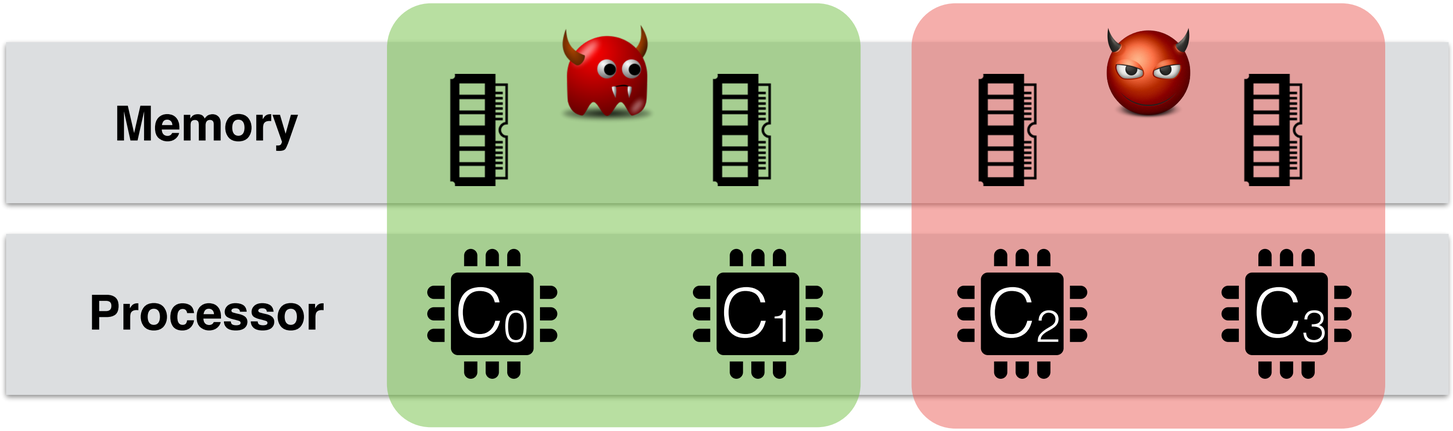
Mitigating Covert Channels: Resource Partitioning

Spatial Partitioning



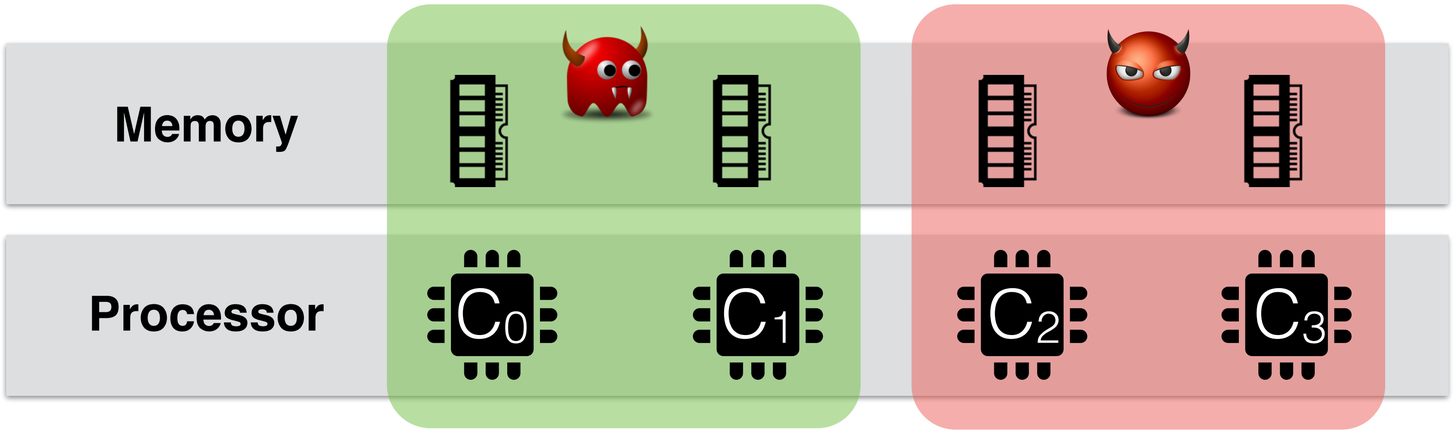
Mitigating Covert Channels: Resource Partitioning

Spatial Partitioning

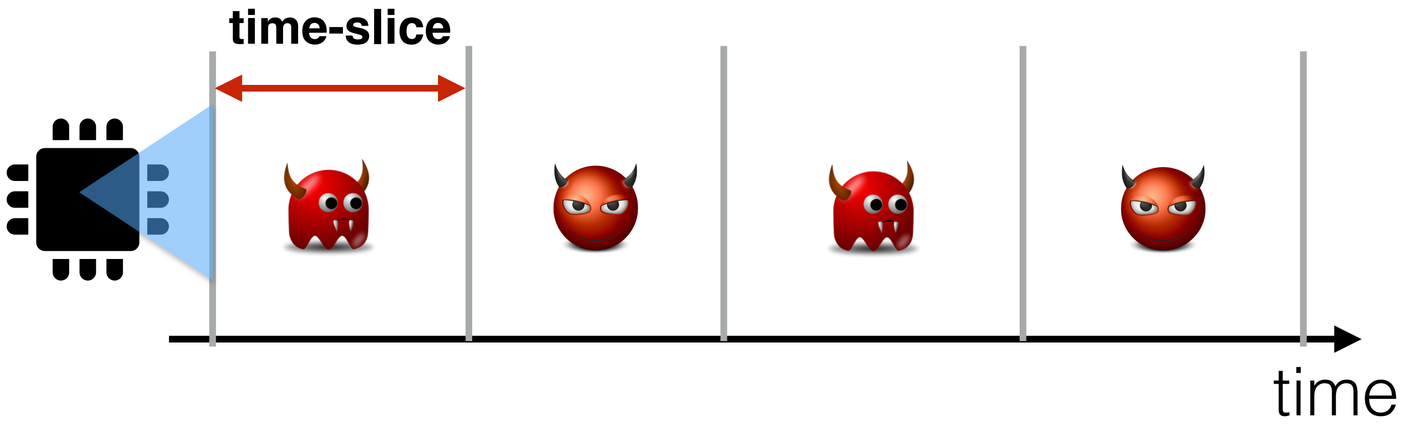


Mitigating Covert Channels: Resource Partitioning

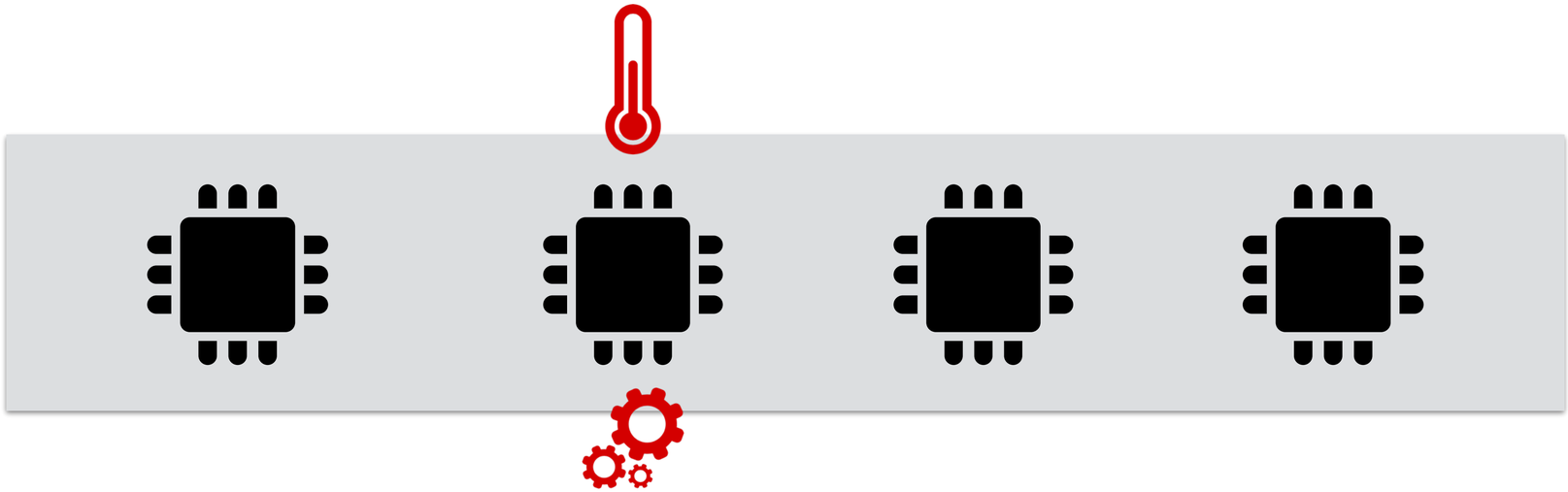
Spatial Partitioning



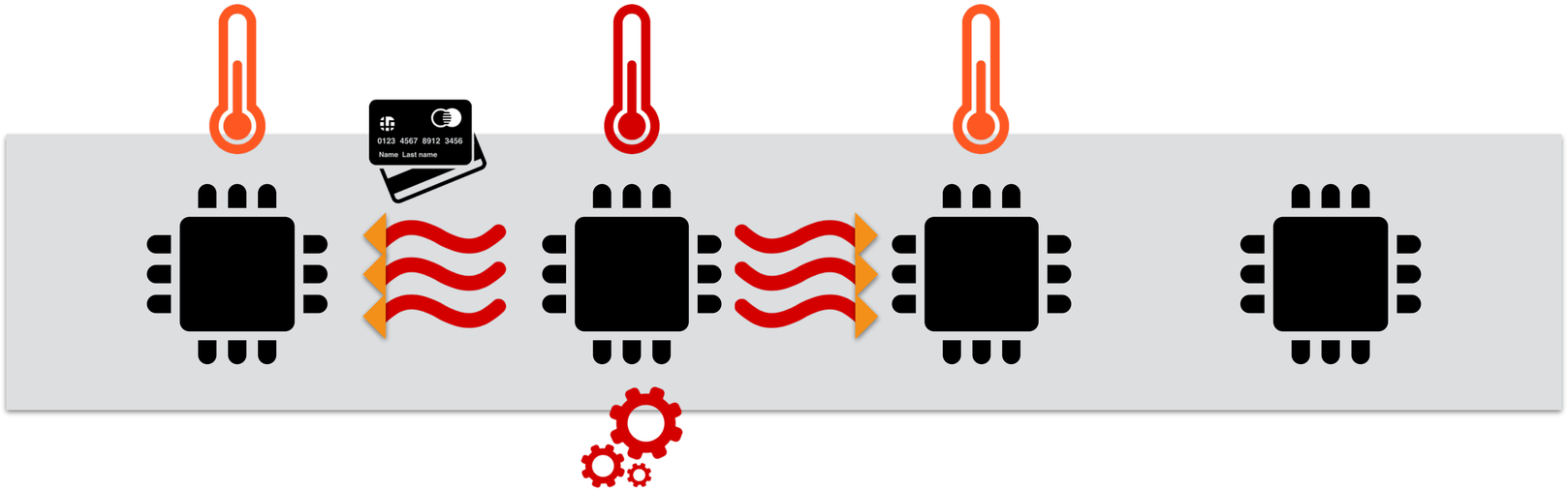
Temporal Partitioning



Thermal Covert Channels on Multicore Systems

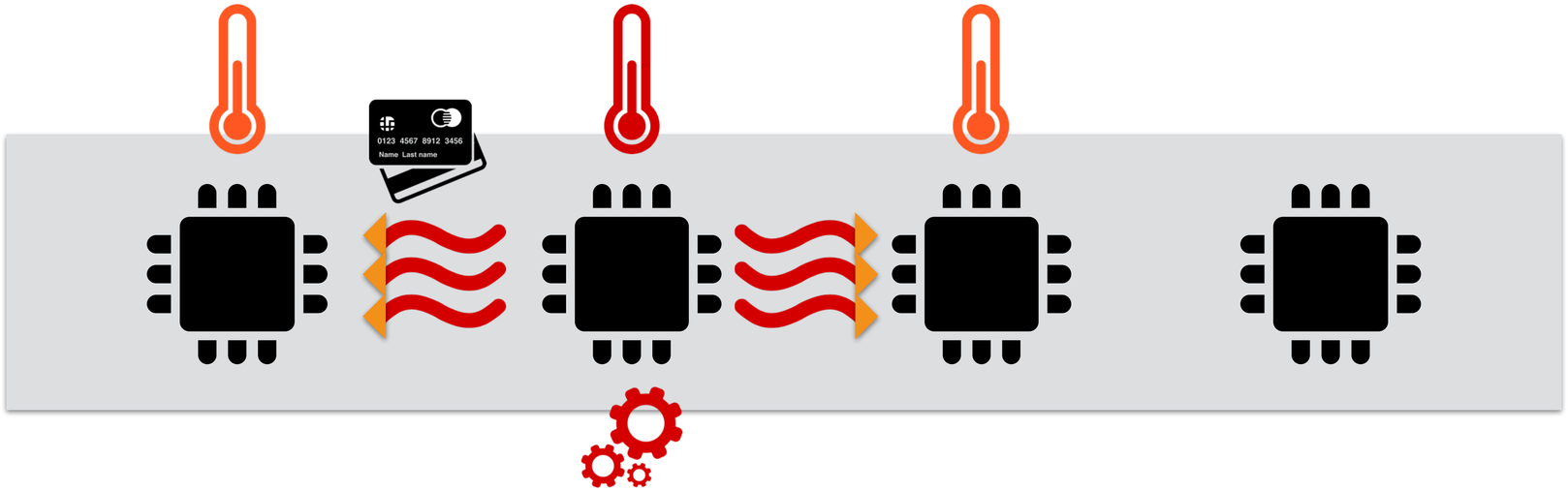


Thermal Covert Channels on Multicore Systems



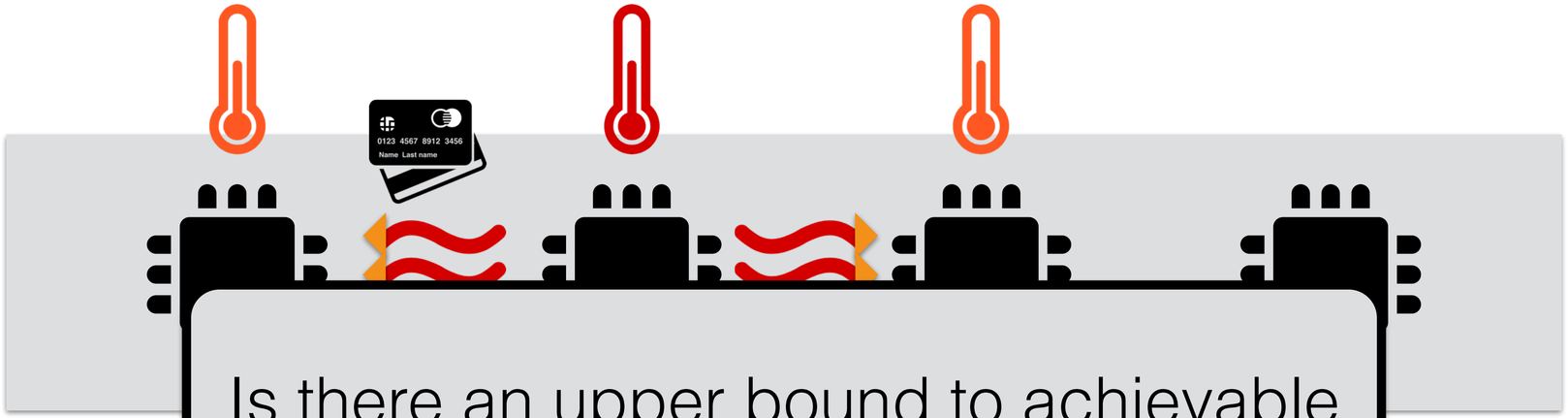
- We demonstrate the feasibility of thermal covert channels in multicore systems despite temporal and spatial resource partitioning and explore the factors that affect its throughput.

Thermal Covert Channels on Multicore Systems



- We demonstrate the feasibility of thermal covert channels in multicore systems despite temporal and spatial resource partitioning and explore the factors that affect its throughput.
- In addition, we show the existence of limited thermal side channel leakage about processes running on neighbouring cores.

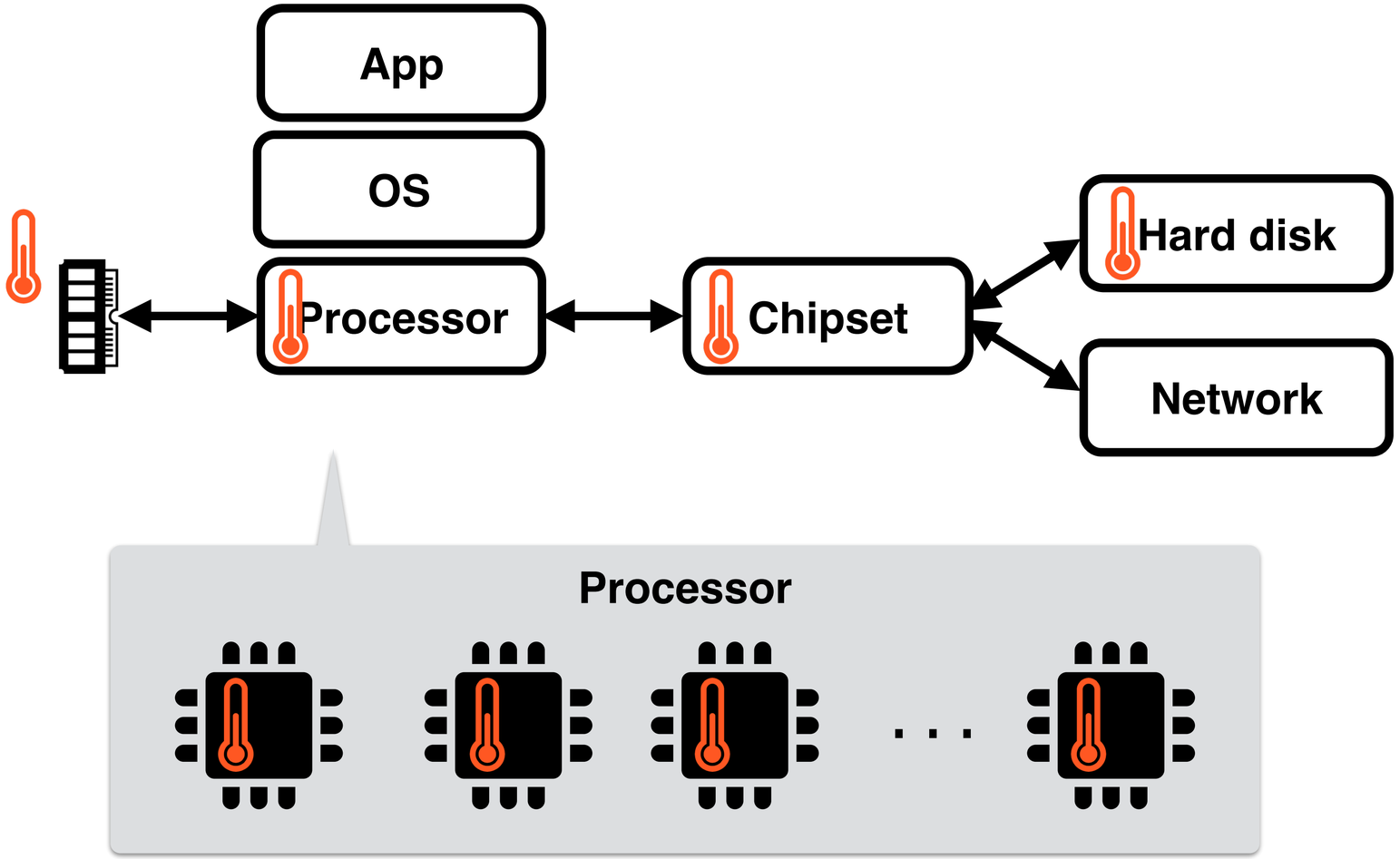
Thermal Covert Channels on Multicore Systems



Is there an upper bound to achievable isolation on multicore platforms?

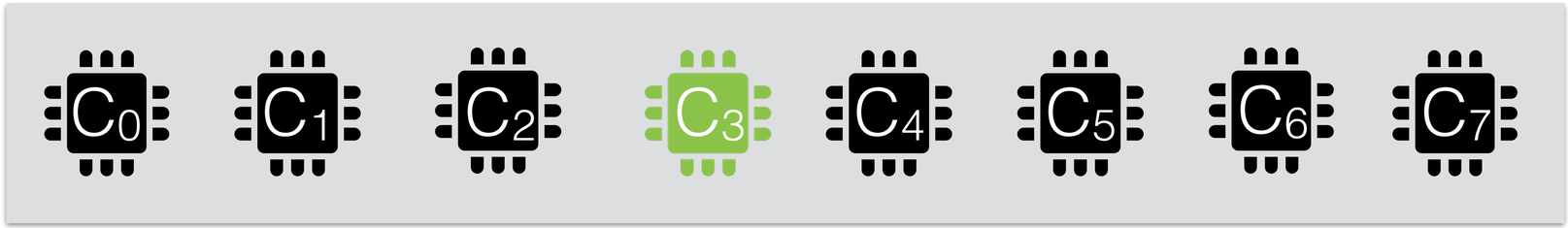
- We demonstrate the feasibility of thermal covert channels in multicore systems despite temporal and spatial resource partitioning and explore the factors that affect its throughput.
- In addition, we show the existence of limited thermal side channel leakage about processes running on neighbouring cores.

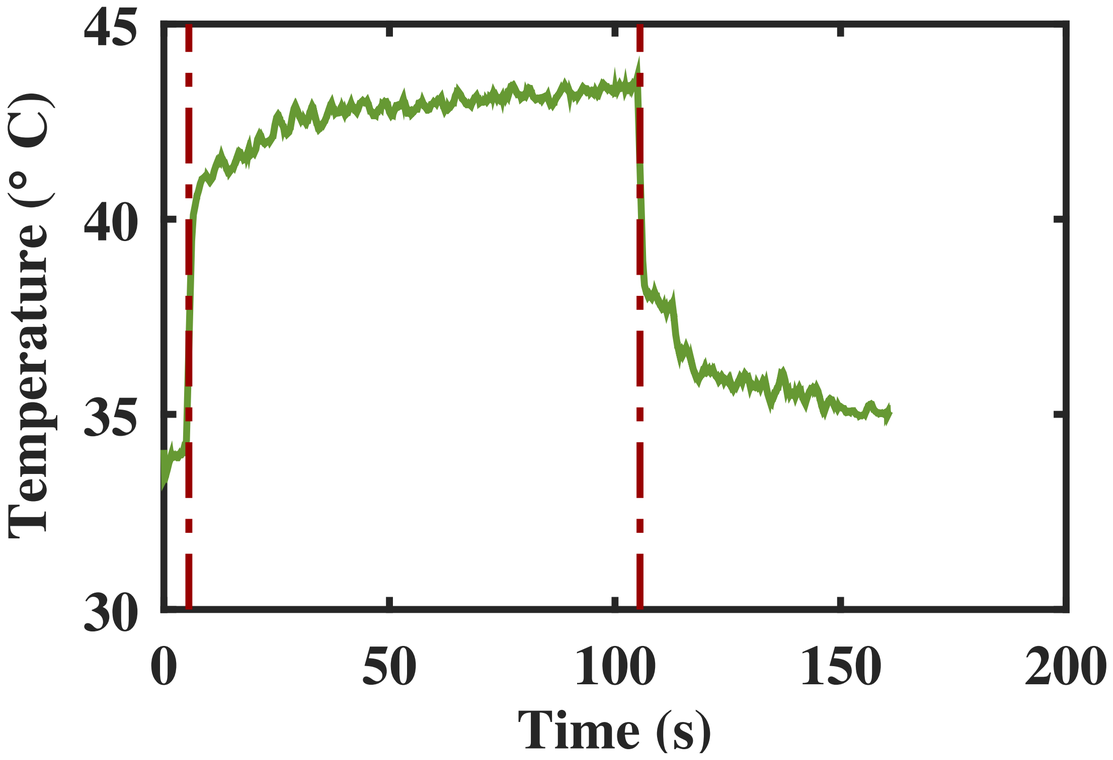
Thermal Management in x86 Systems



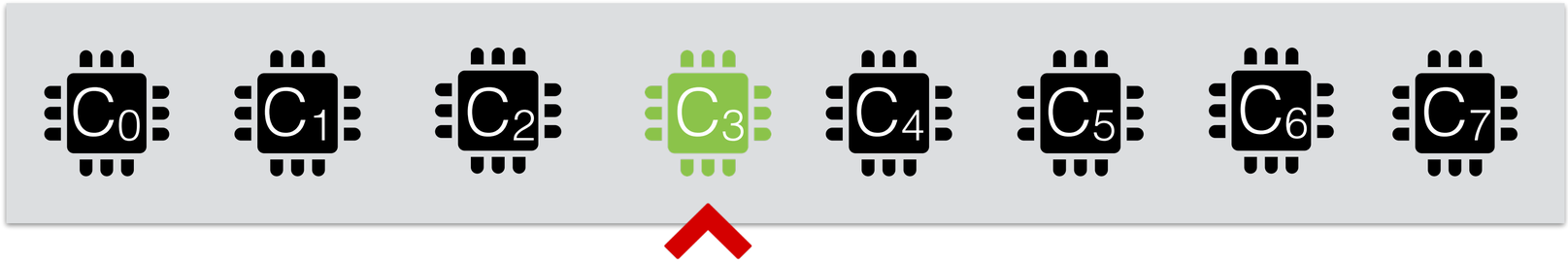
Recent trend towards user-centric thermal management!

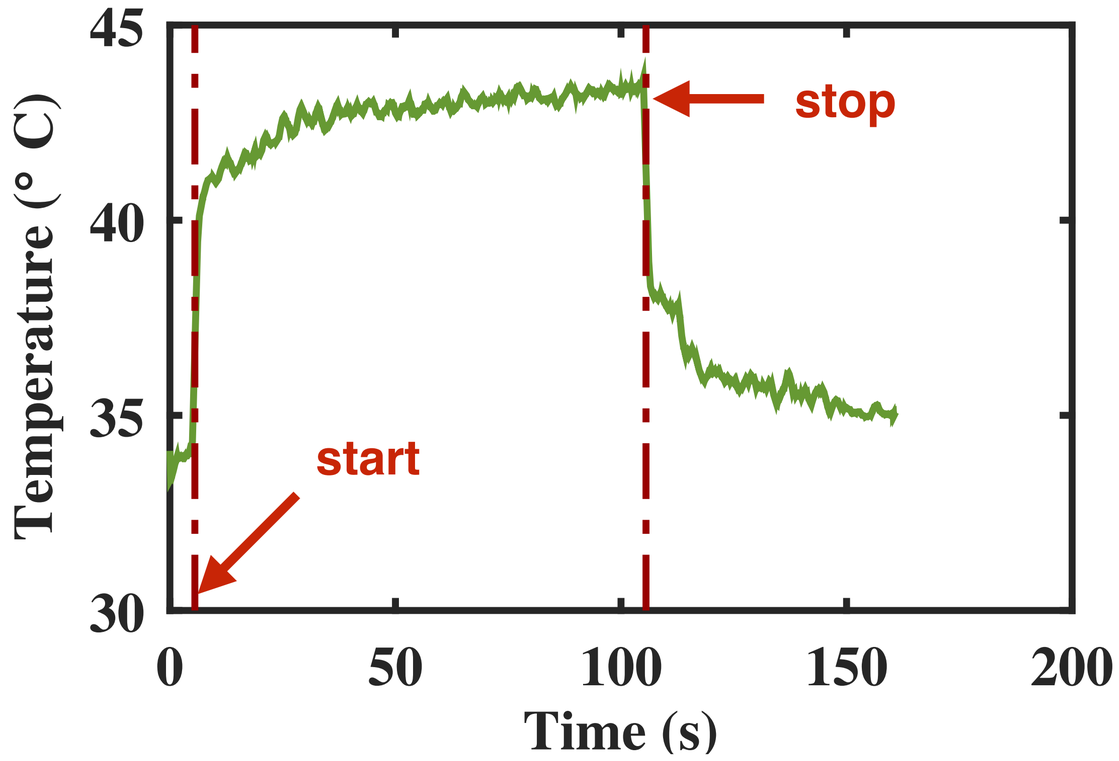
Thermal Response of a CPU Intensive Application





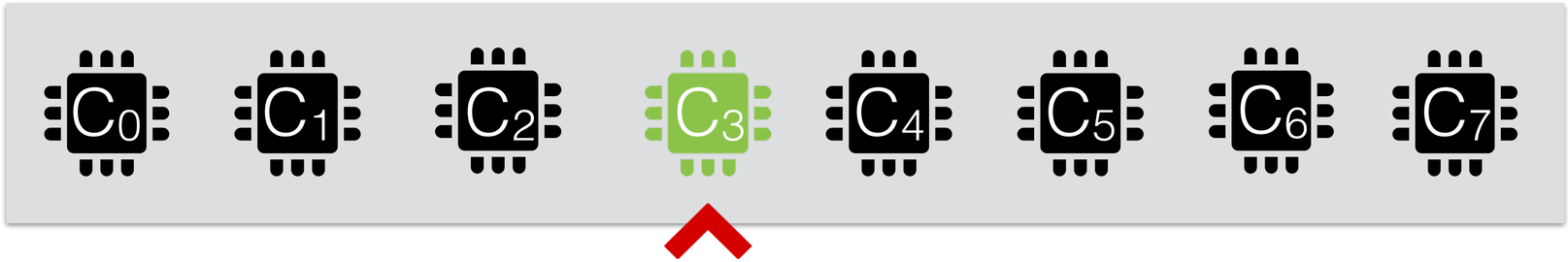
Thermal Response of a CPU Intensive Application

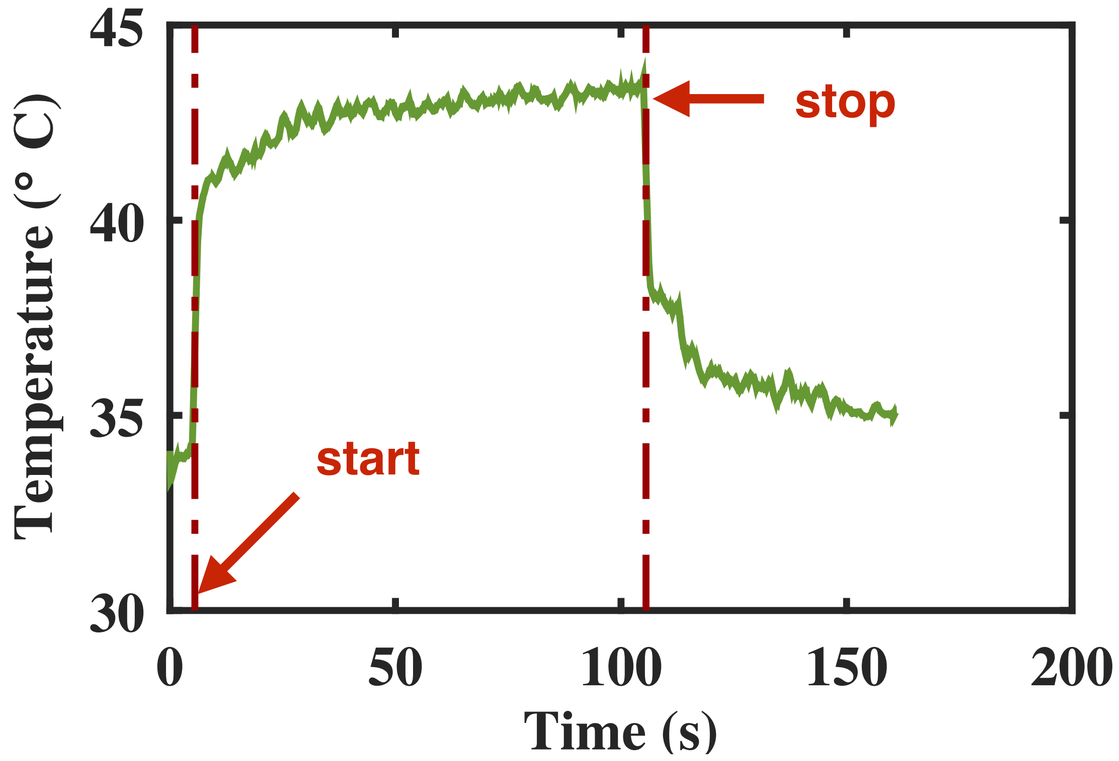




Application run for 100 s

Thermal Response of a CPU Intensive Application

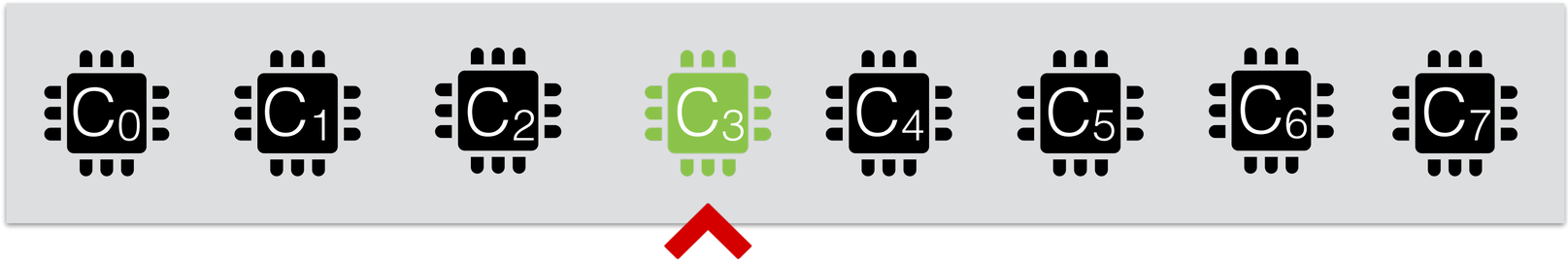


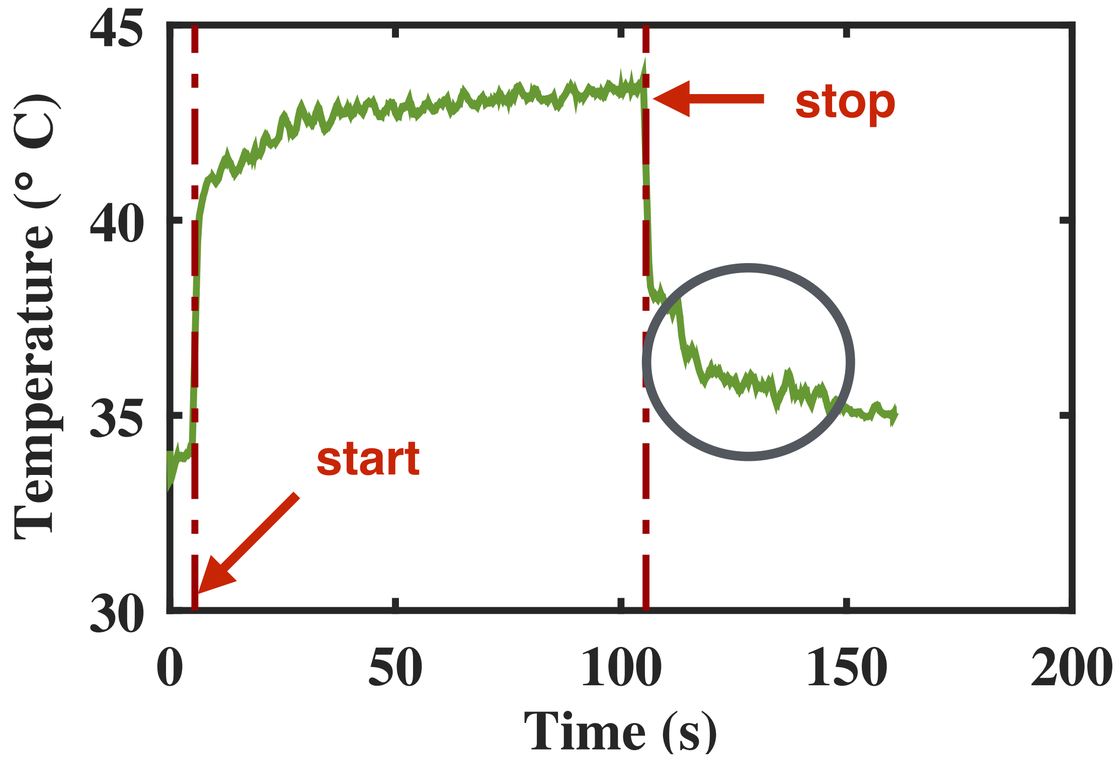


Application run for 100 s

Exponential rise and fall of temperature.

Thermal Response of a CPU Intensive Application



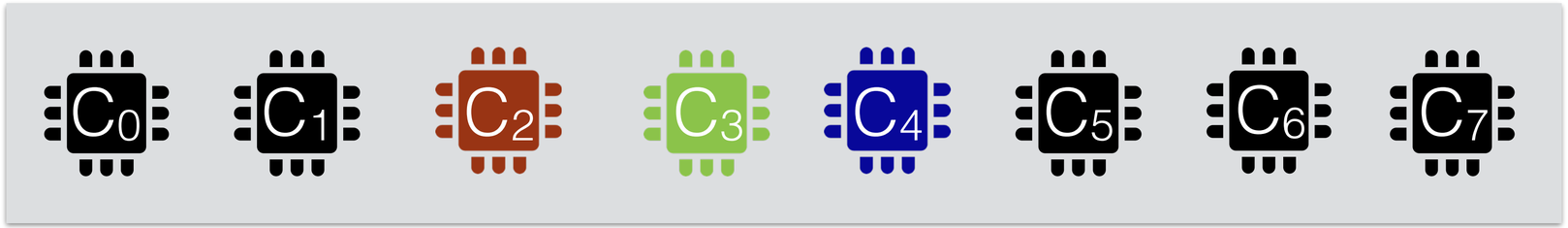


Application run for 100 s

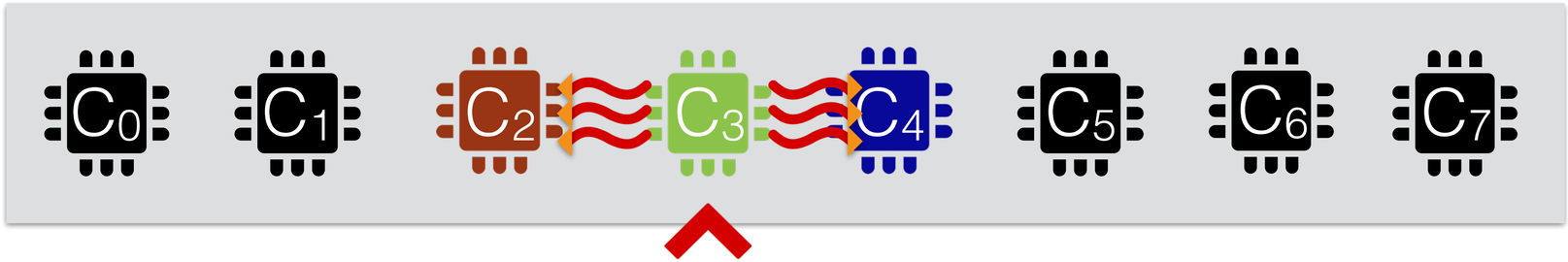
Exponential rise and fall of temperature.

Remnant heat even after computation

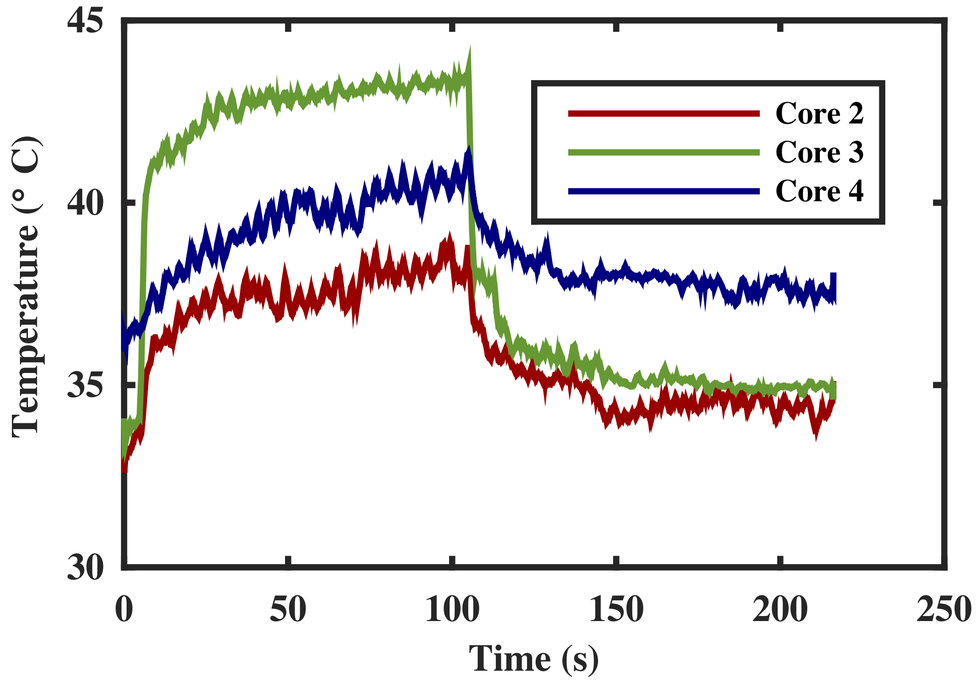
Effect of a Neighbour's Computation



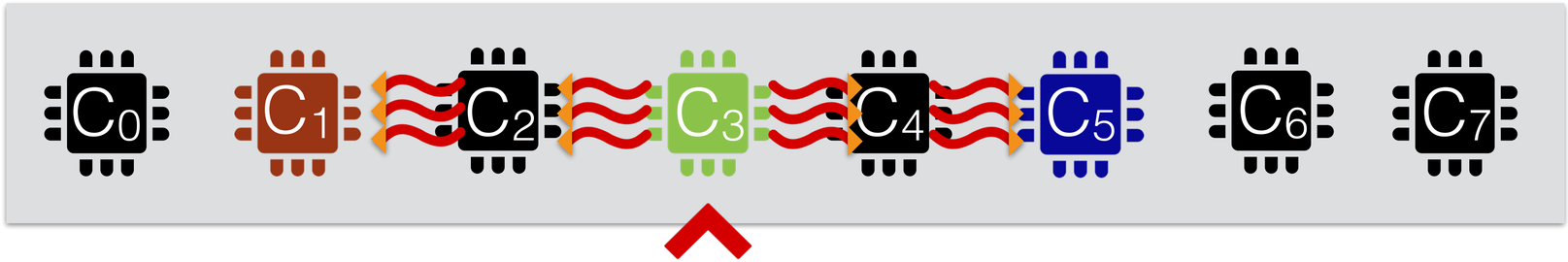
Effect of a Neighbour's Computation



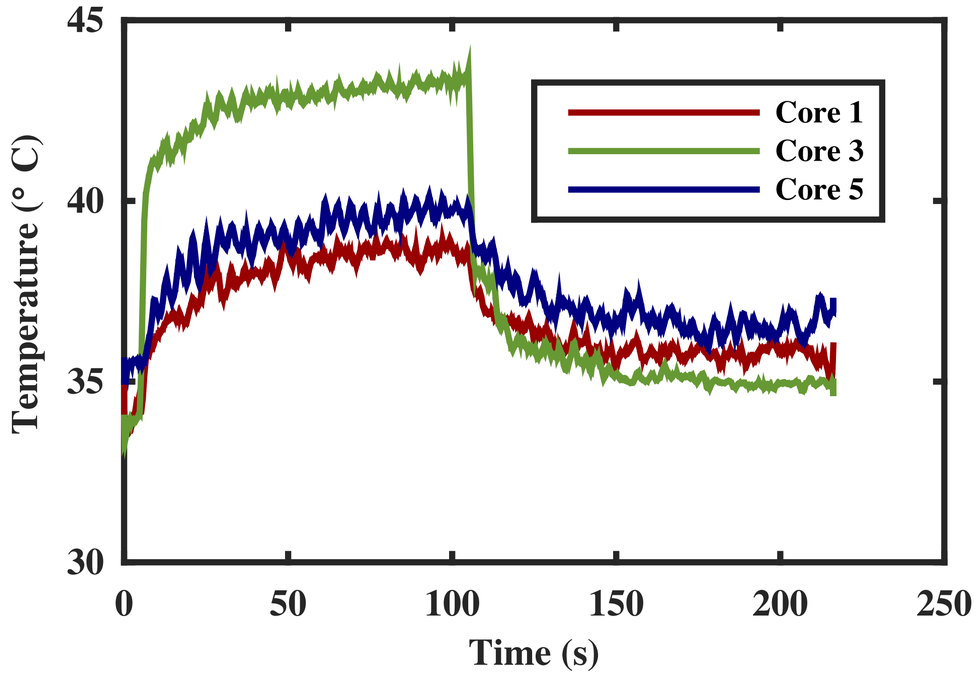
1-hop neighbours



Effect of a Neighbour's Computation



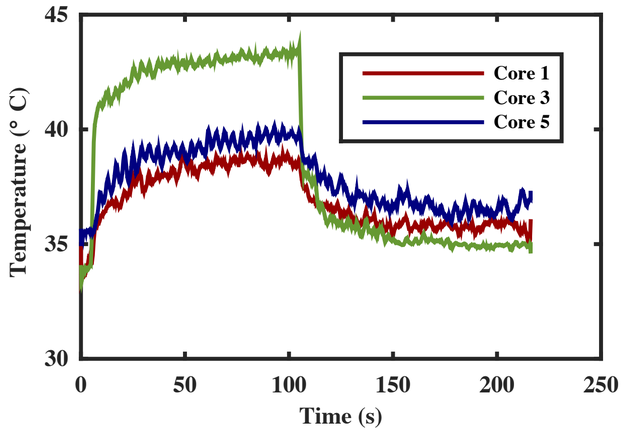
2-hop neighbours



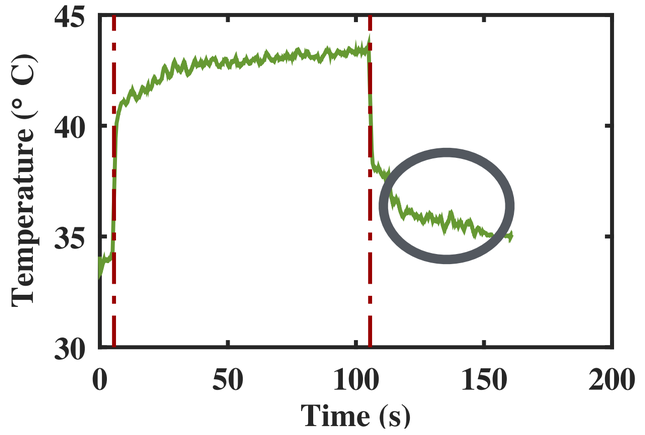
Exploiting Thermal Behaviour

Key Observations:

Heat propagates and affects neighbouring cores.



There is remnant heat even after computation is completed.

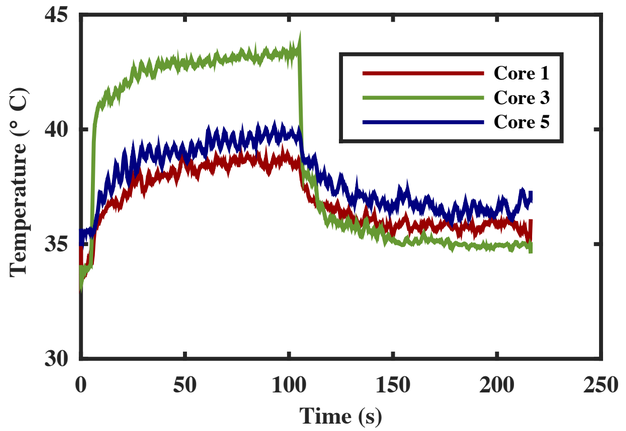


Exploiting Thermal Behaviour

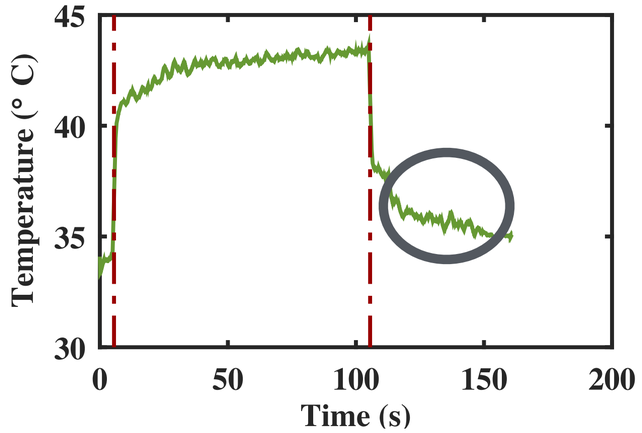
Key Observations:

Heat propagates and affects neighbouring cores.

Implications for spatial partitioning?



There is remnant heat even after computation is completed.

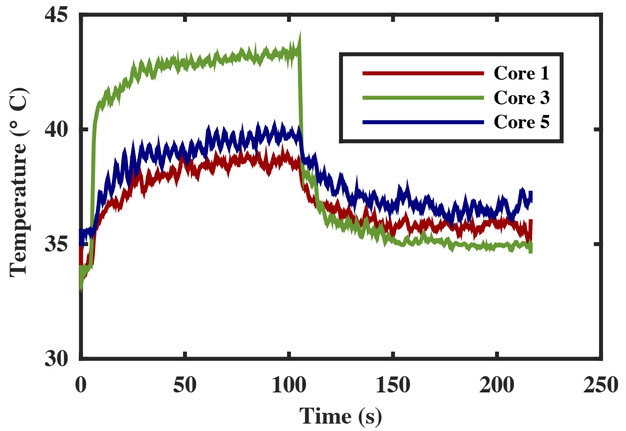


Exploiting Thermal Behaviour

Key Observations:

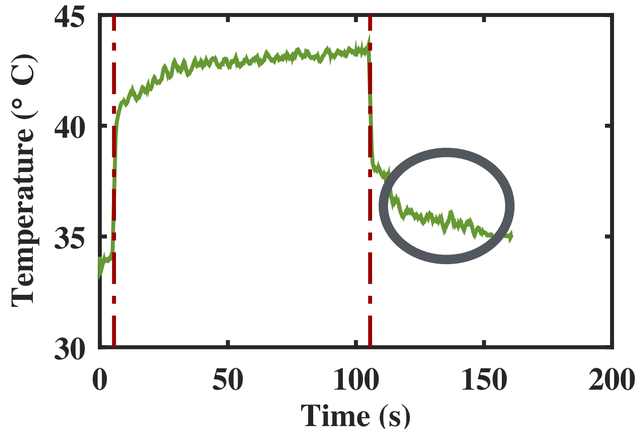
Heat propagates and affects neighbouring cores.

Implications for spatial partitioning?



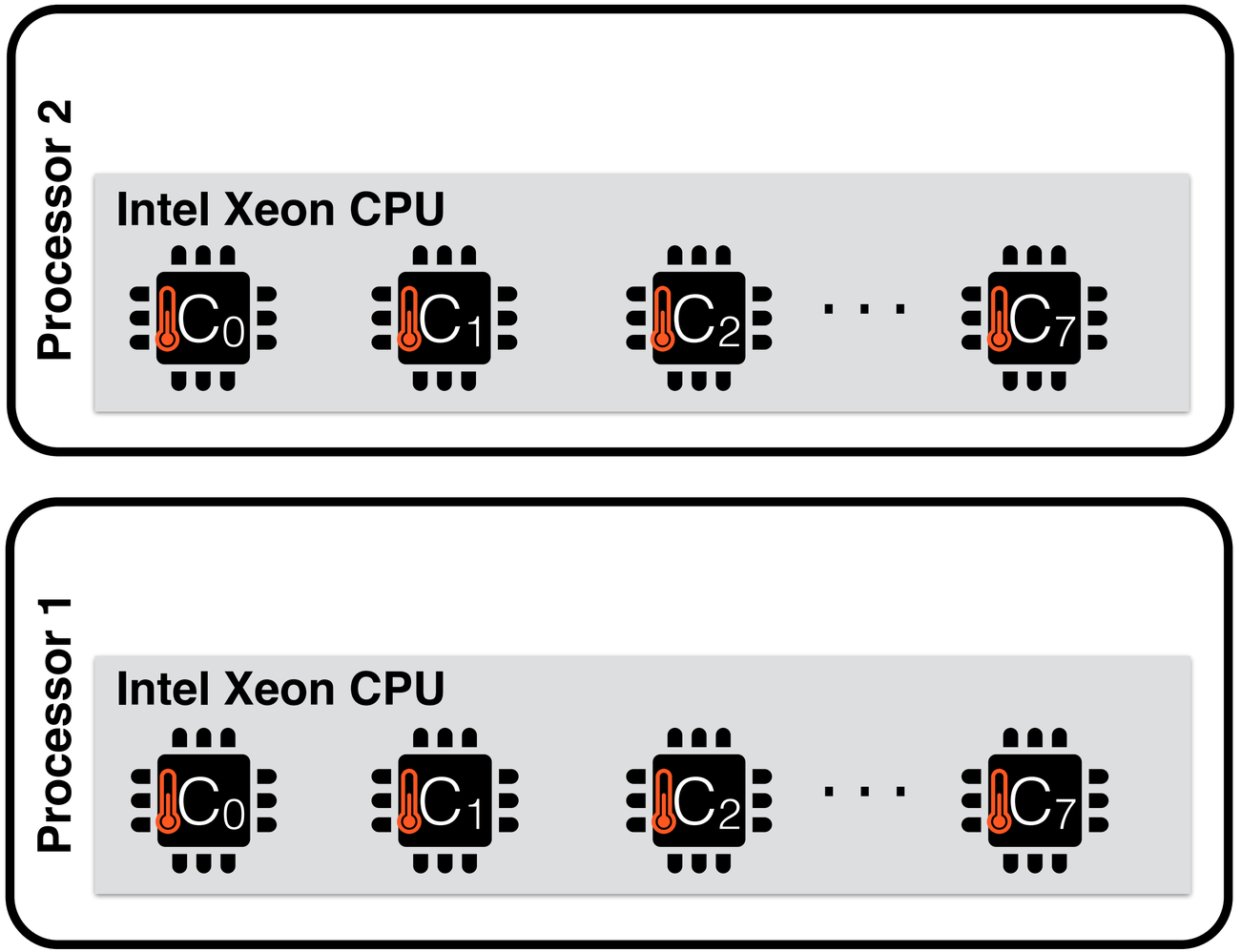
There is remnant heat even after computation is completed.

Implications for temporal partitioning?



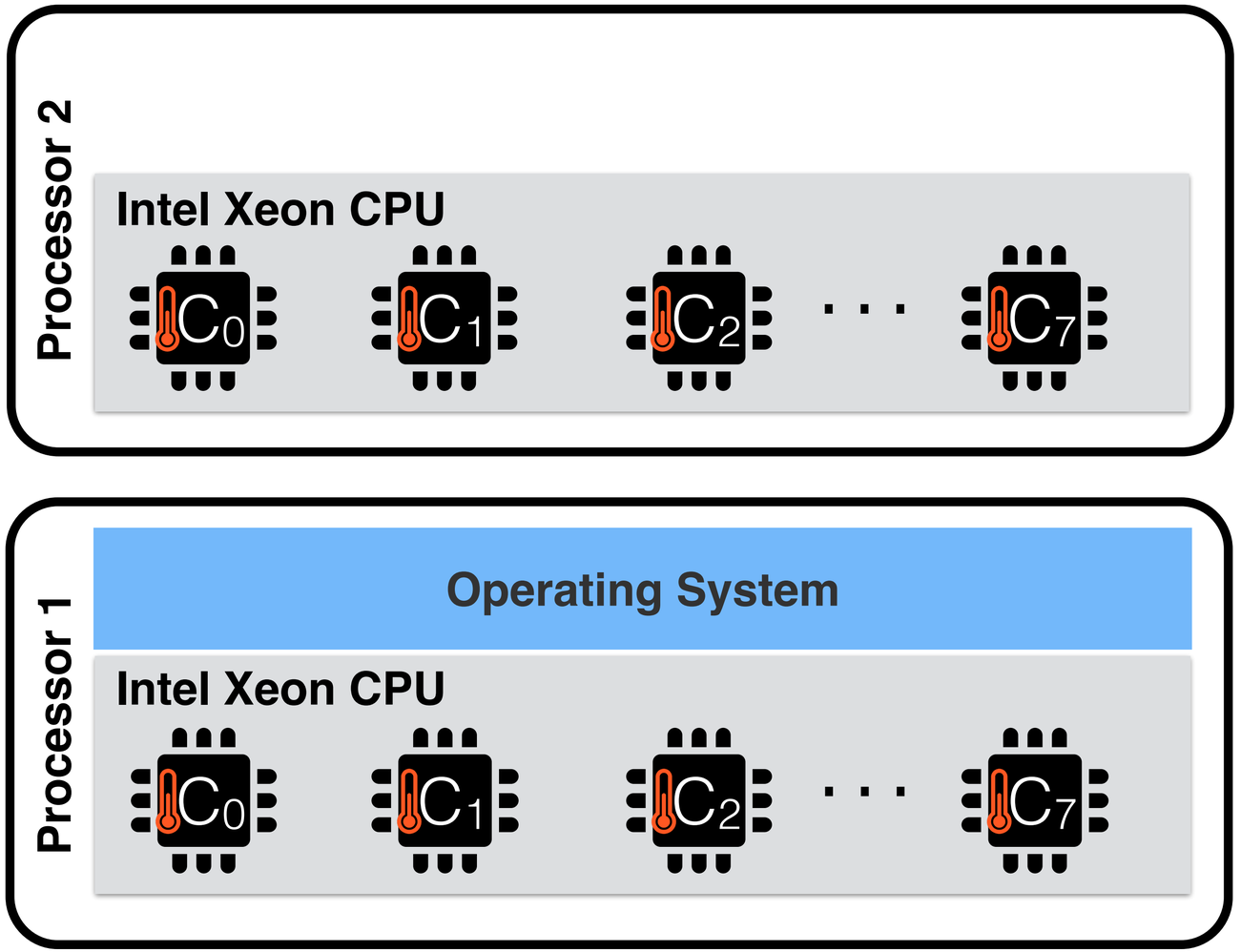
Experimental Setup

- Intel Xeon based server consisting of two 8-core CPUs running Linux
- ‘cpusets’ for implementing spatial and temporal partitioning.
- Allows runtime configuration of CPU frequency, application-core mapping etc.



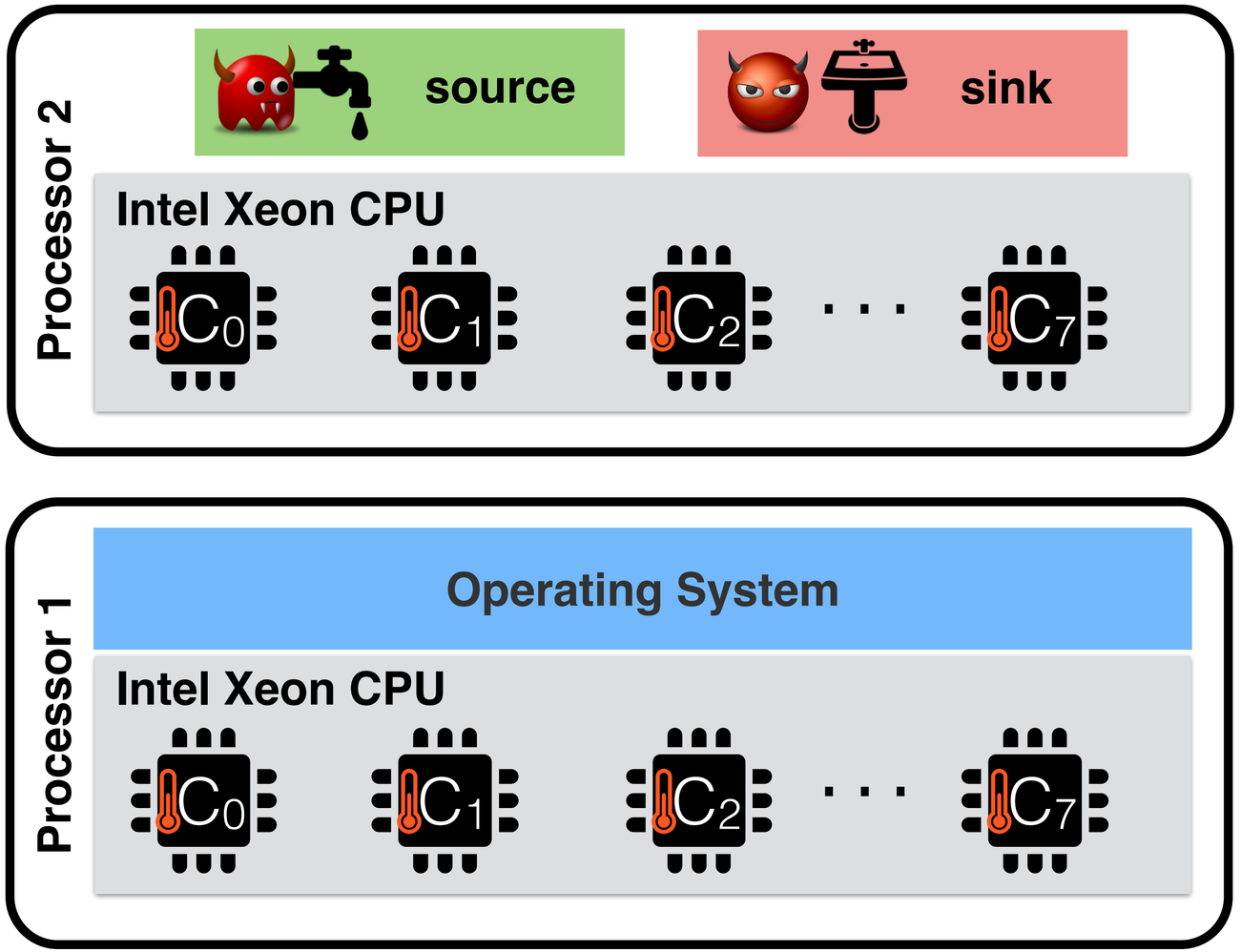
Experimental Setup

- Intel Xeon based server consisting of two 8-core CPUs running Linux
- ‘cpusets’ for implementing spatial and temporal partitioning.
- Allows runtime configuration of CPU frequency, application-core mapping etc.

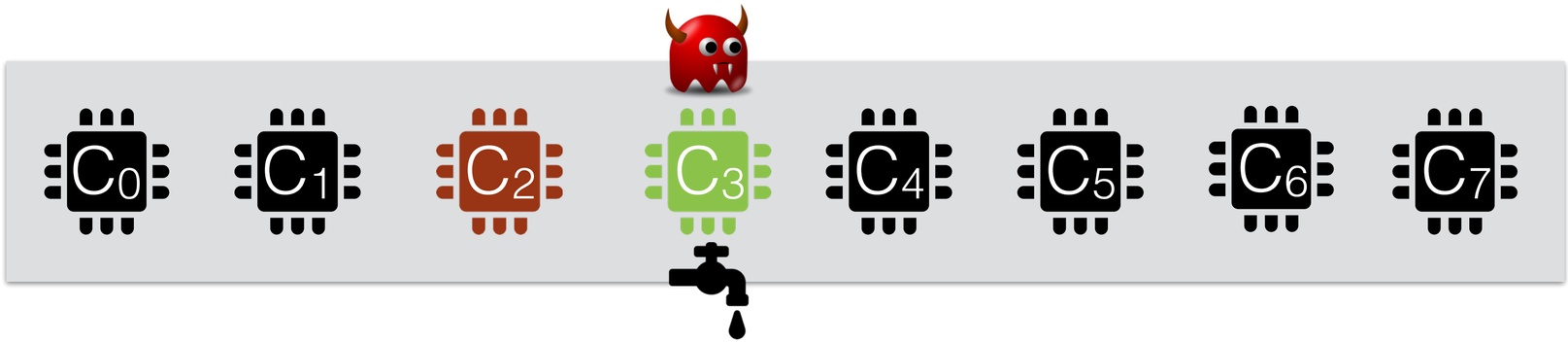


Experimental Setup

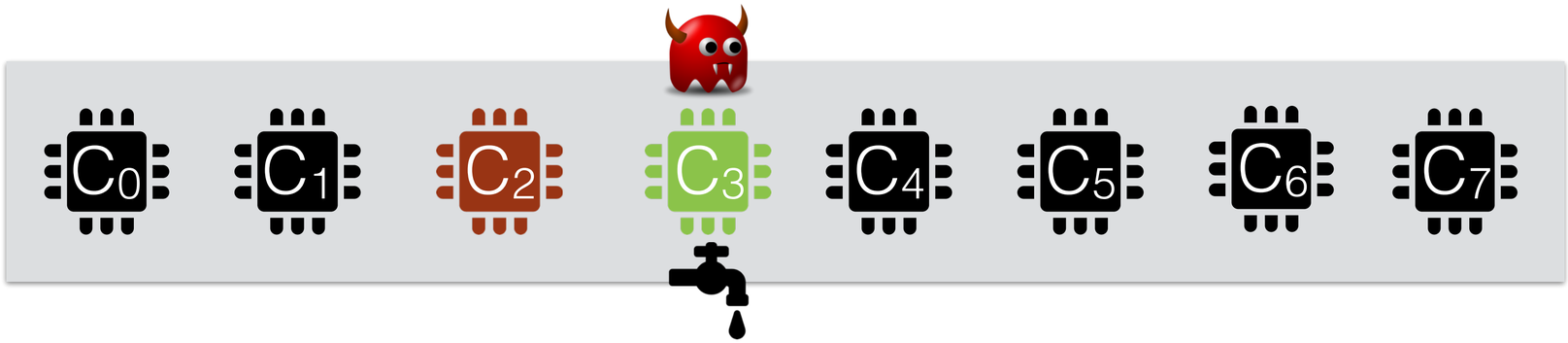
- Intel Xeon based server consisting of two 8-core CPUs running Linux
- ‘cpusets’ for implementing spatial and temporal partitioning.
- Allows runtime configuration of CPU frequency, application-core mapping etc.



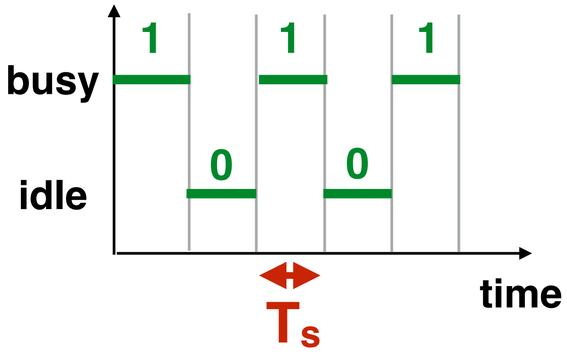
Overcoming Spatial Partitioning



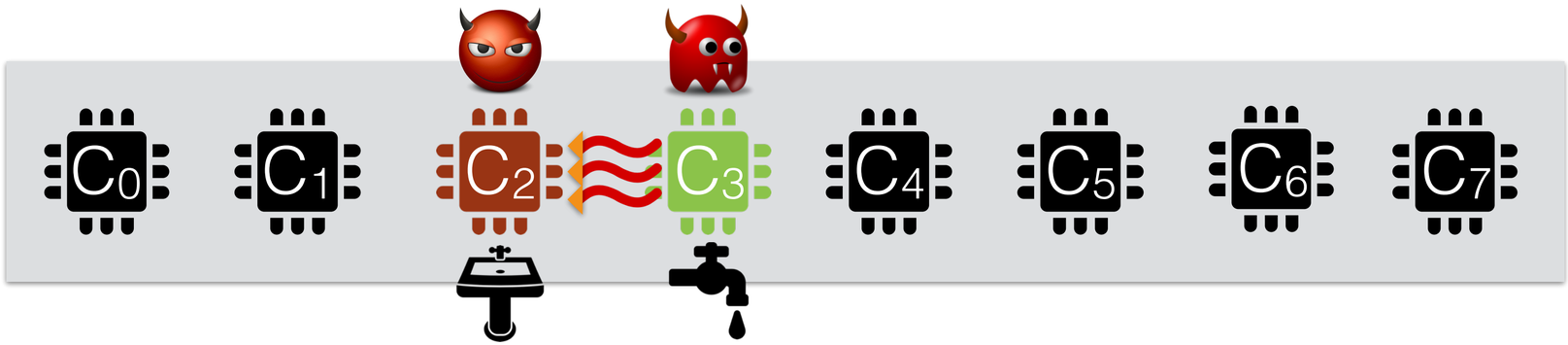
Overcoming Spatial Partitioning



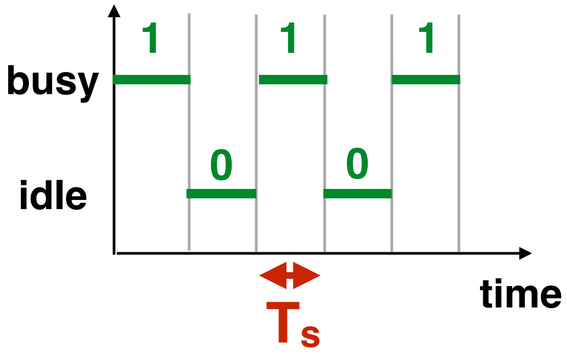
CPU activity at Core 3



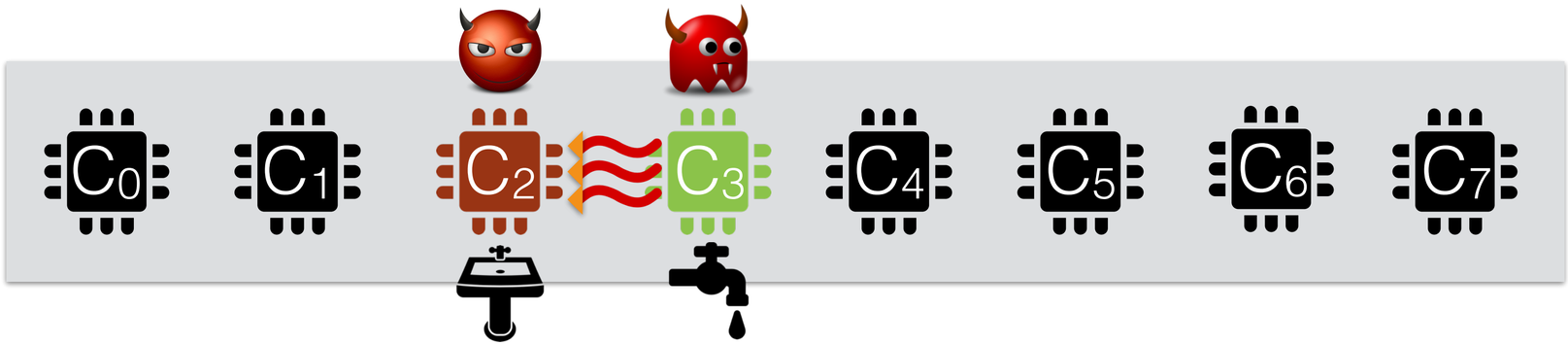
Overcoming Spatial Partitioning



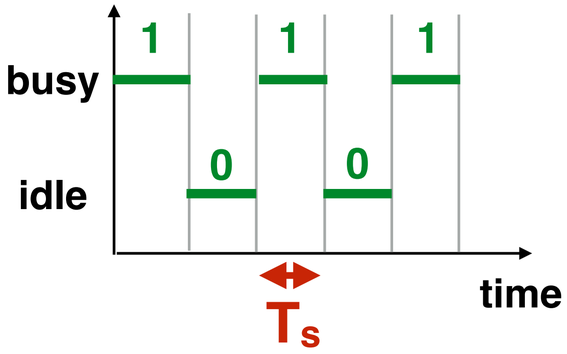
CPU activity at Core 3



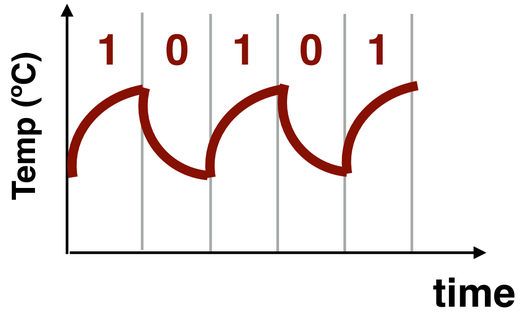
Overcoming Spatial Partitioning



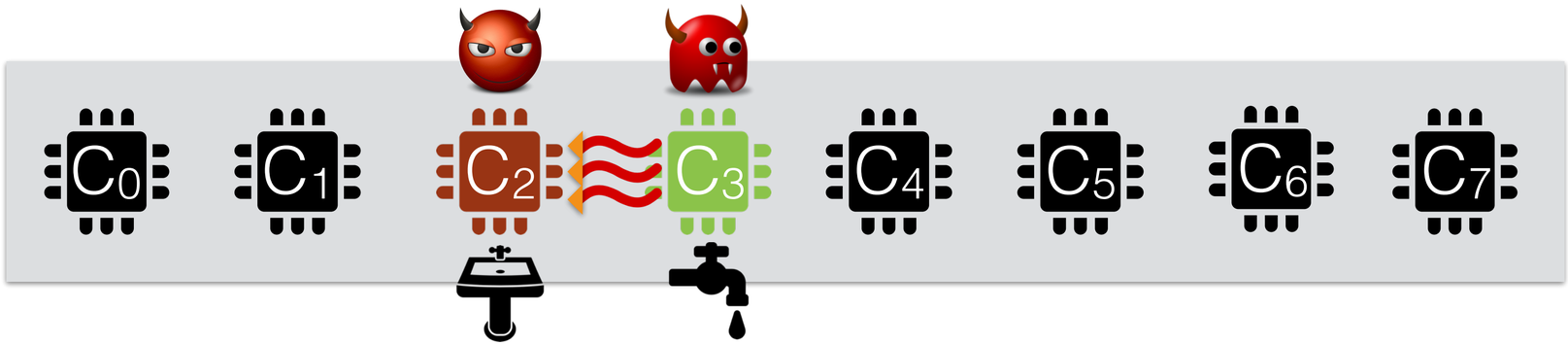
CPU activity at Core 3



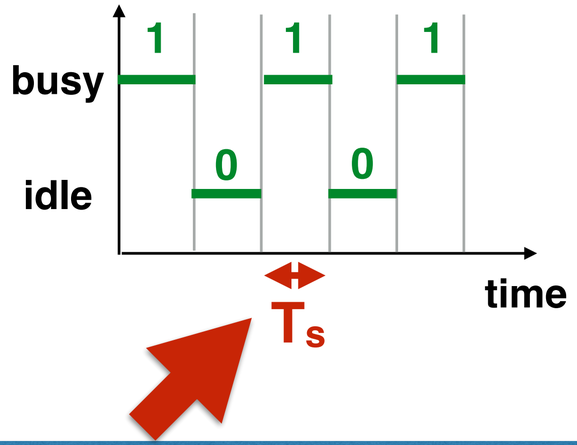
Temperature at Core 2



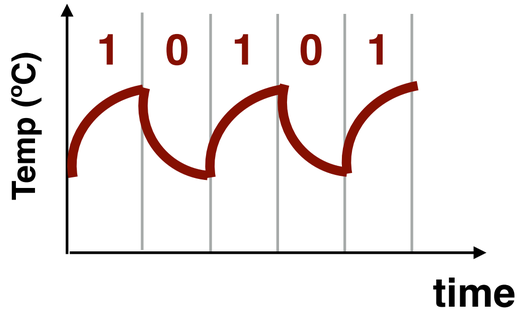
Overcoming Spatial Partitioning



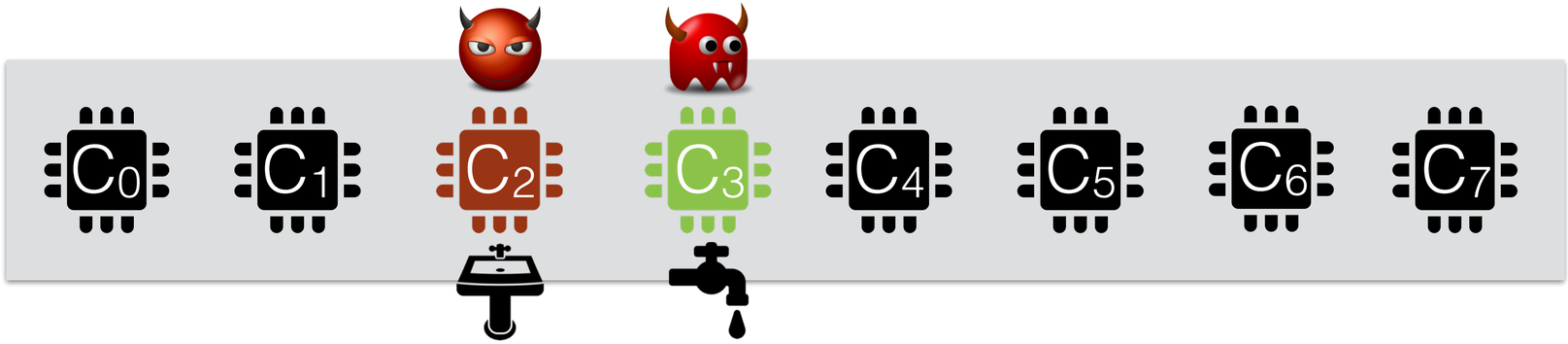
CPU activity at Core 3

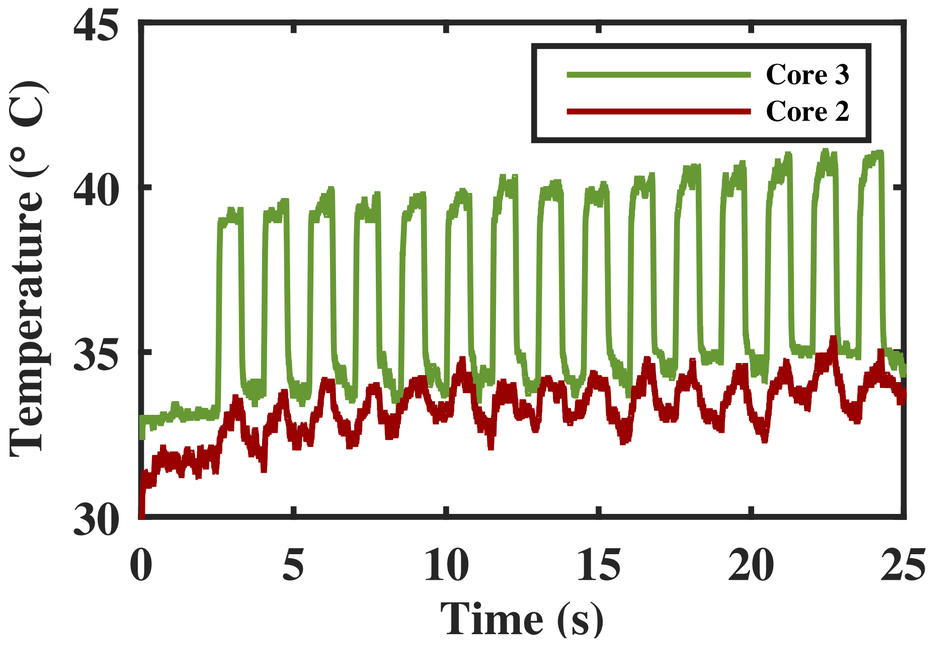


Temperature at Core 2

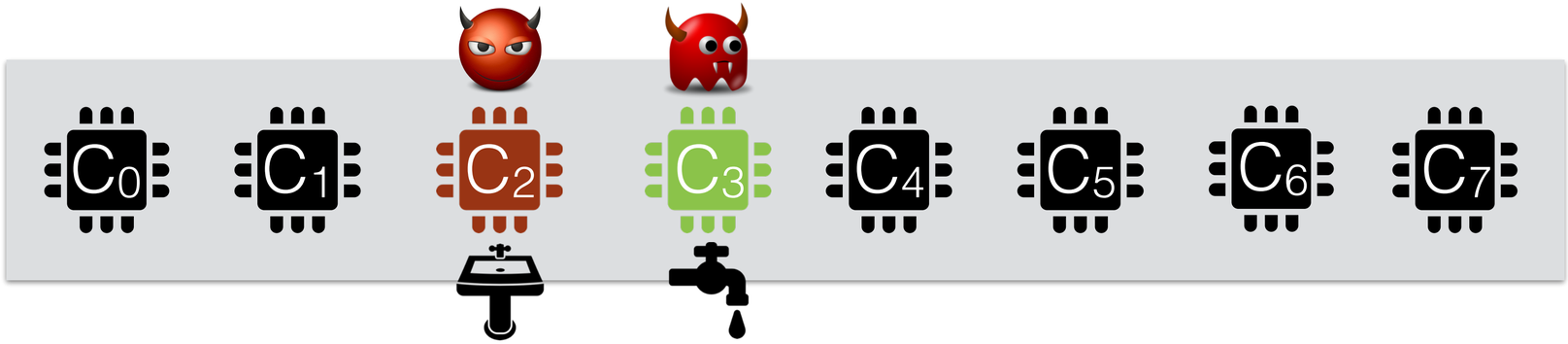


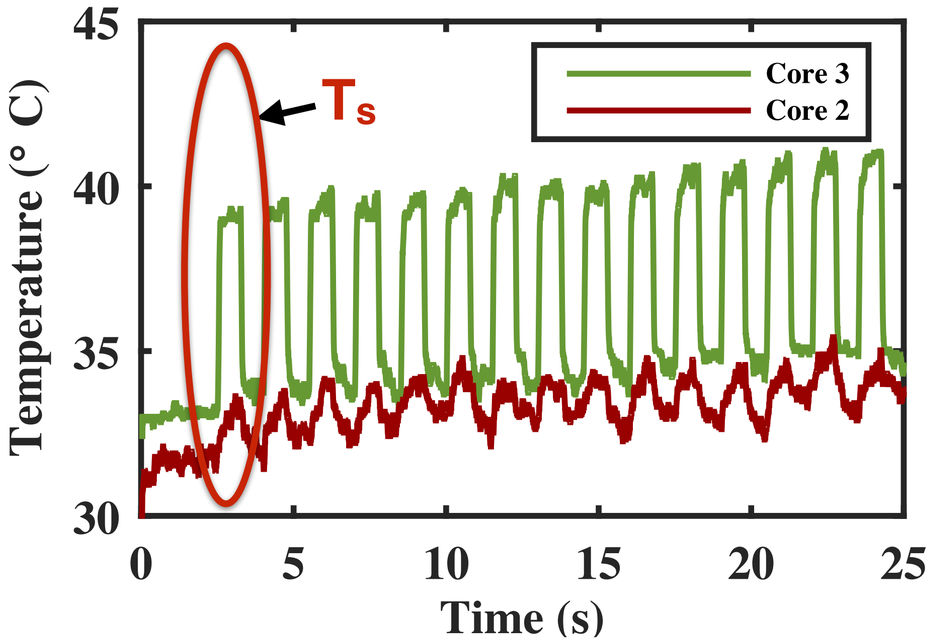
Spatial Partitioning: Error and Throughput Results





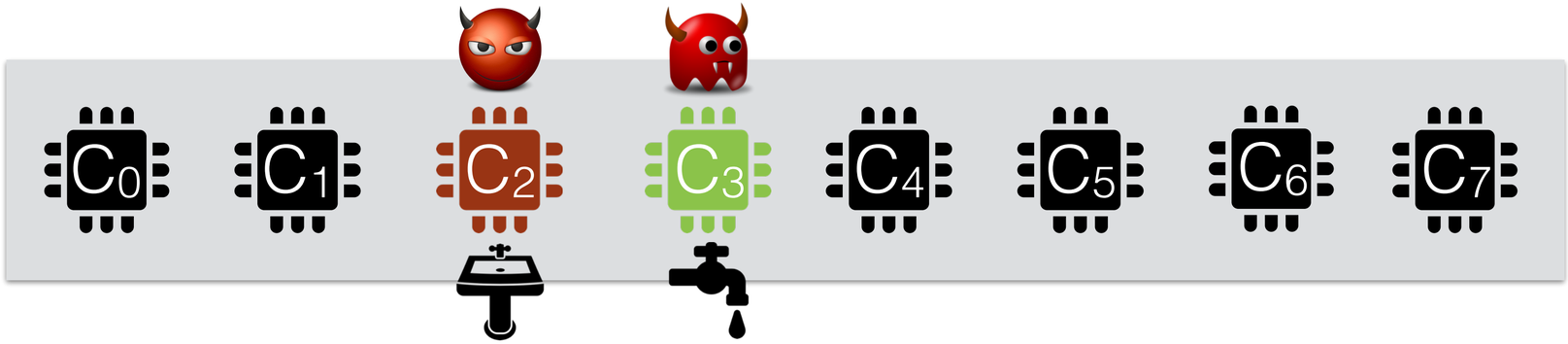
Spatial Partitioning: Error and Throughput Results

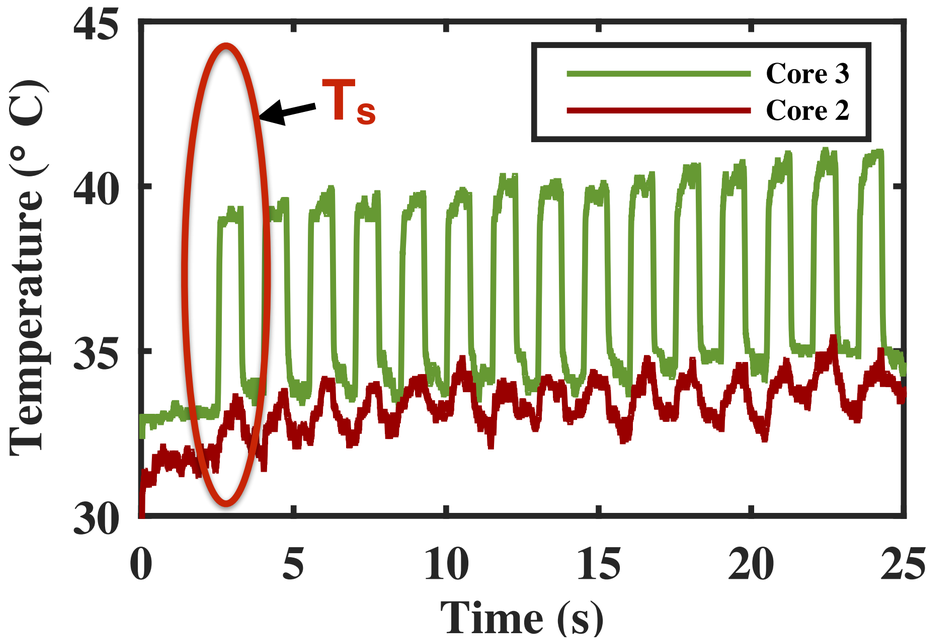




- **T_s = 750 ms**
- 11% errors over 1000 bits
- Achievable throughput with error correction: **0.33 bps**

Spatial Partitioning: Error and Throughput Results

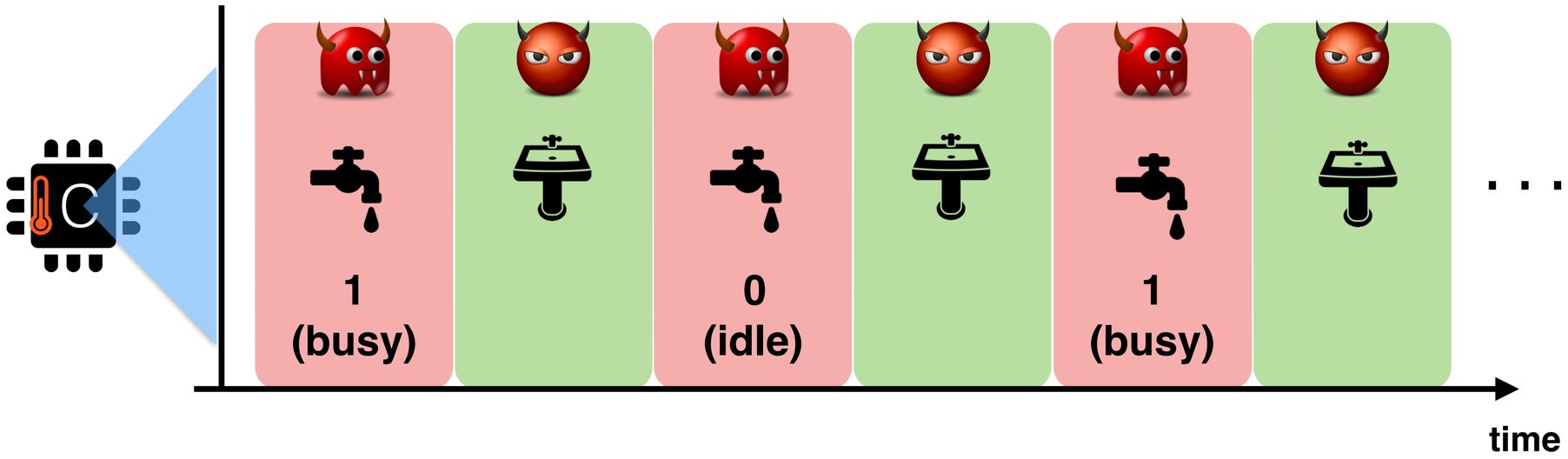


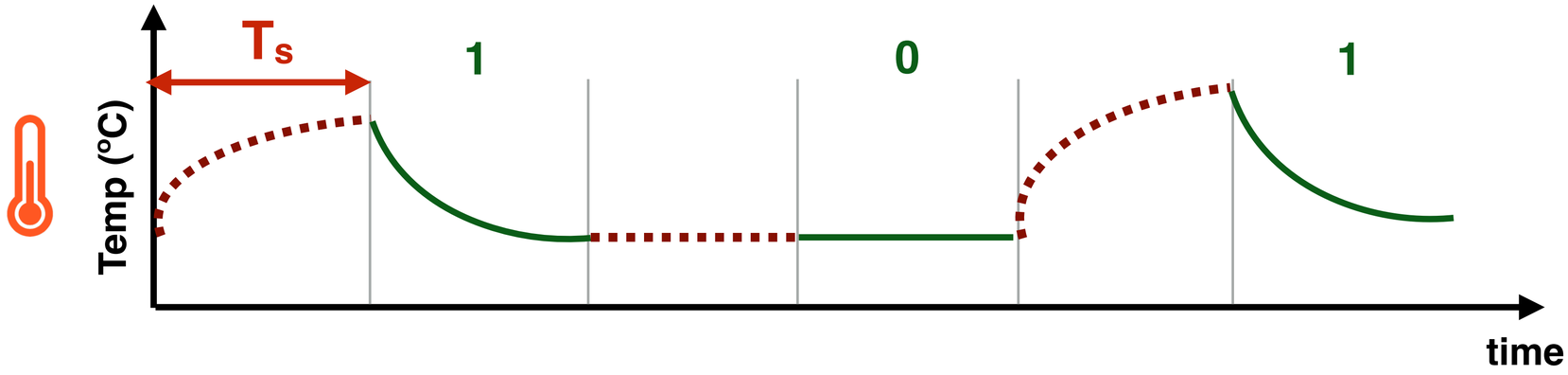


- **T_s = 750 ms**
- 11% errors over 1000 bits
- Achievable throughput with error correction: **0.33 bps**

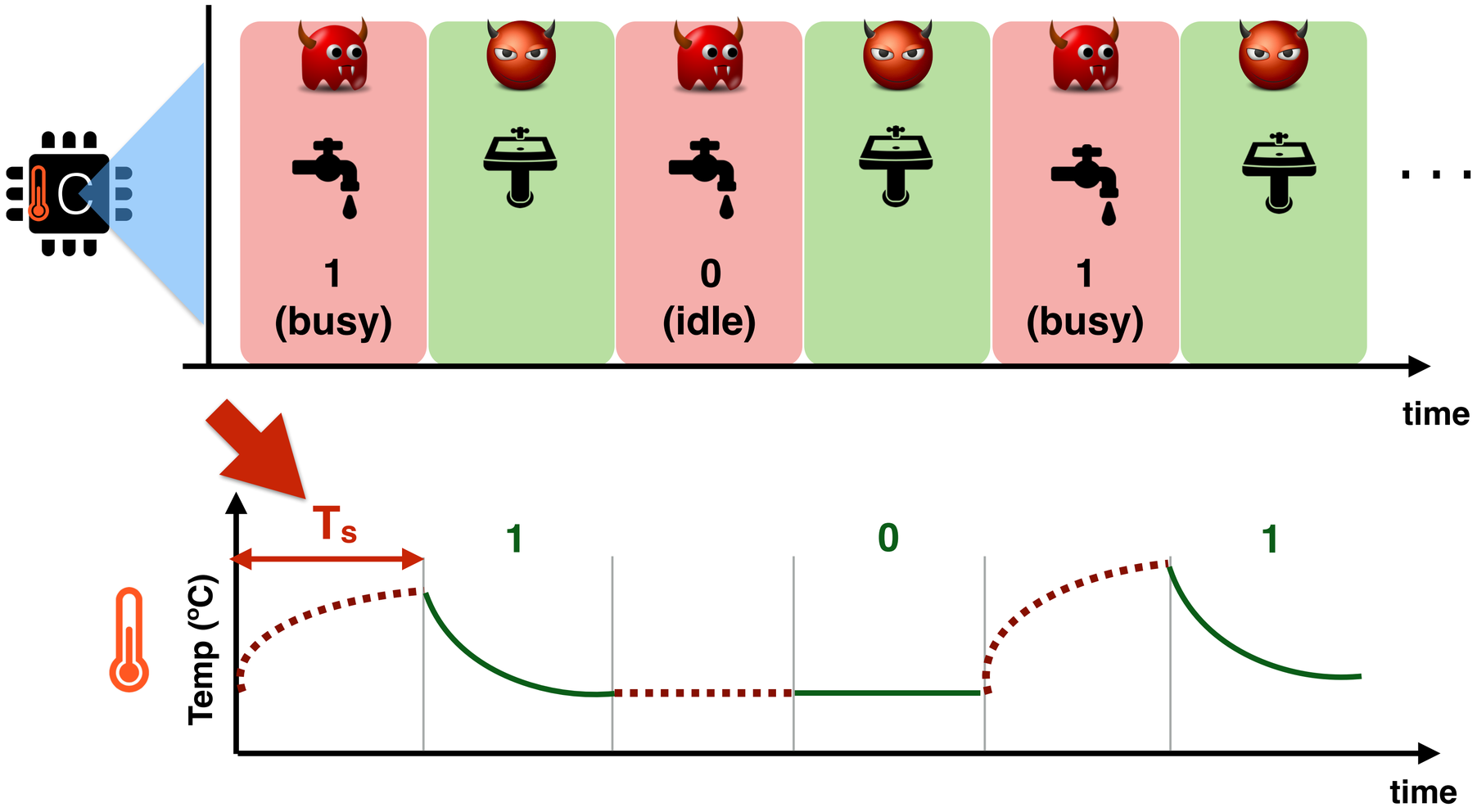
Credit card details can be transmitted in 4 min

Overcoming Temporal Partitioning

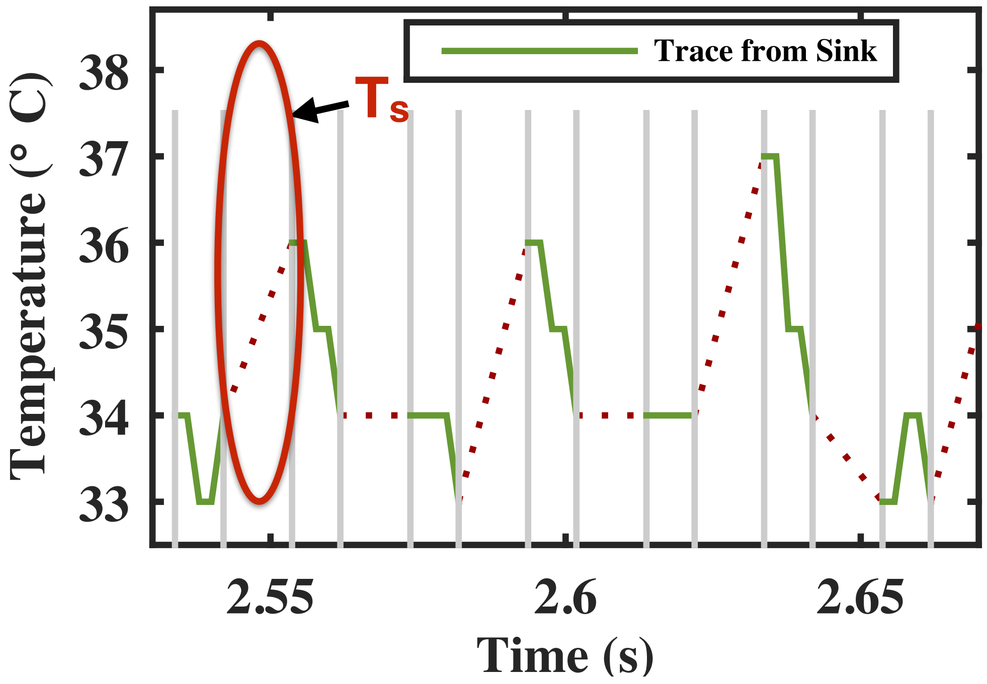




Overcoming Temporal Partitioning

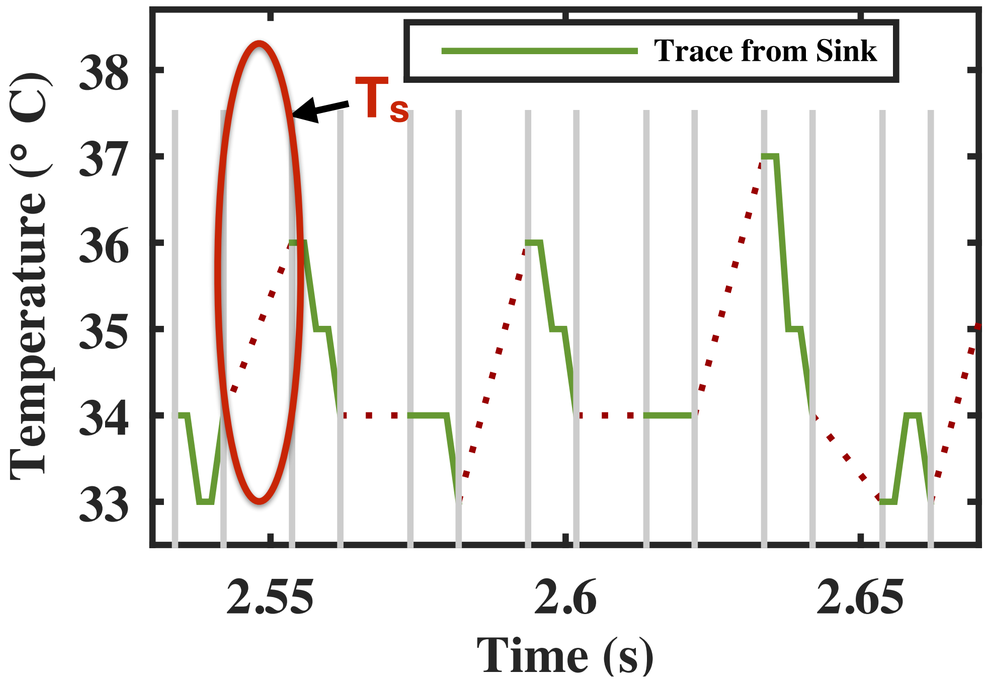


Temporal Partitioning: Errors and Throughput Results



- **T_s = 10 ms**
- 7% errors over 1000 bits
- Achievable throughput with error correction: **12.5 bps**

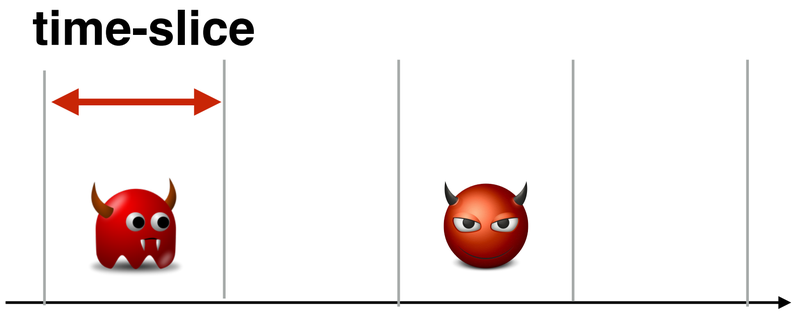
Temporal Partitioning: Errors and Throughput Results



- **T_s = 10 ms**
- 7% errors over 1000 bits
- Achievable throughput with error correction: **12.5 bps**

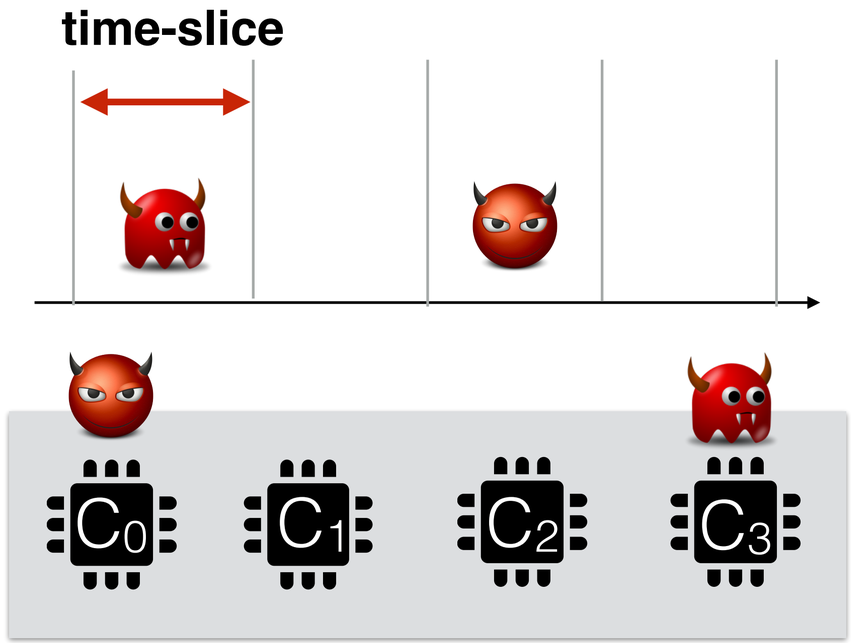
Credit card details can be transmitted in 5 sec

Countermeasures



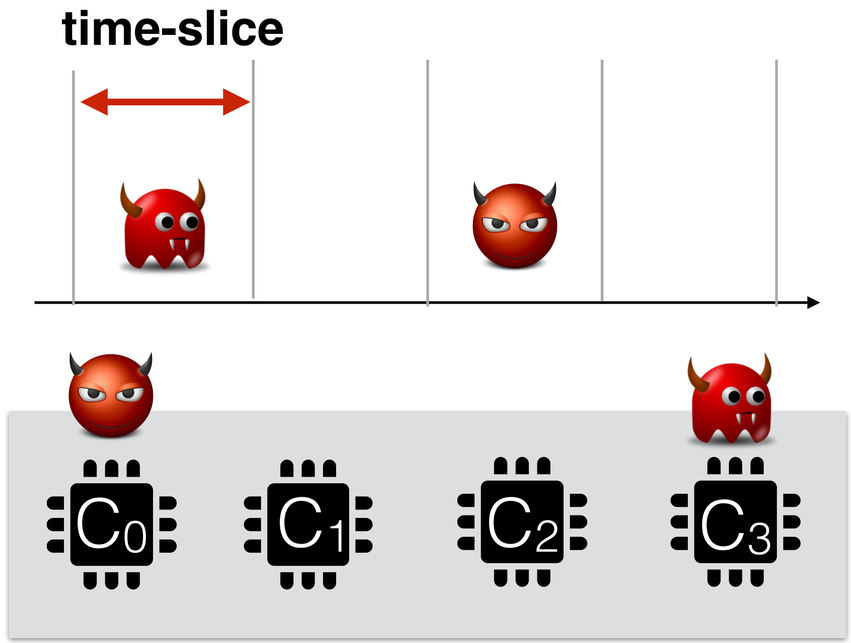
Separate processes temporally and spatially

Countermeasures



Separate processes temporally and spatially

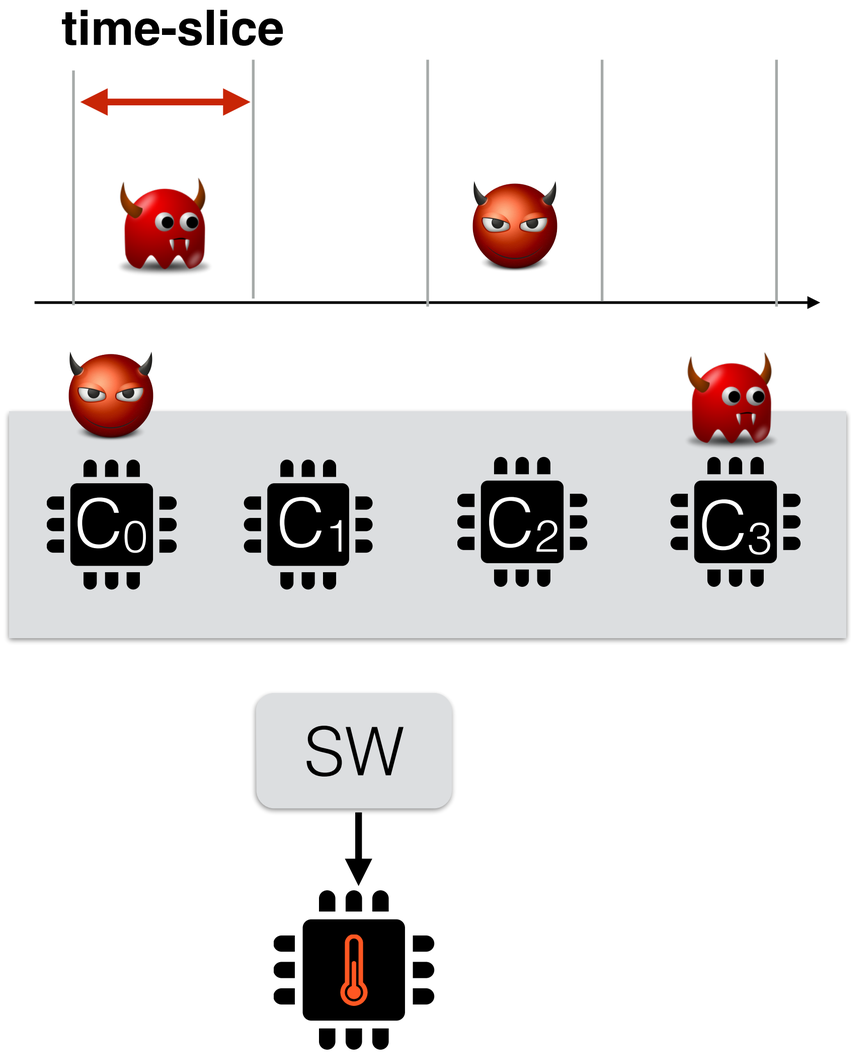
Countermeasures



Separate processes temporally and spatially

Waste of resources

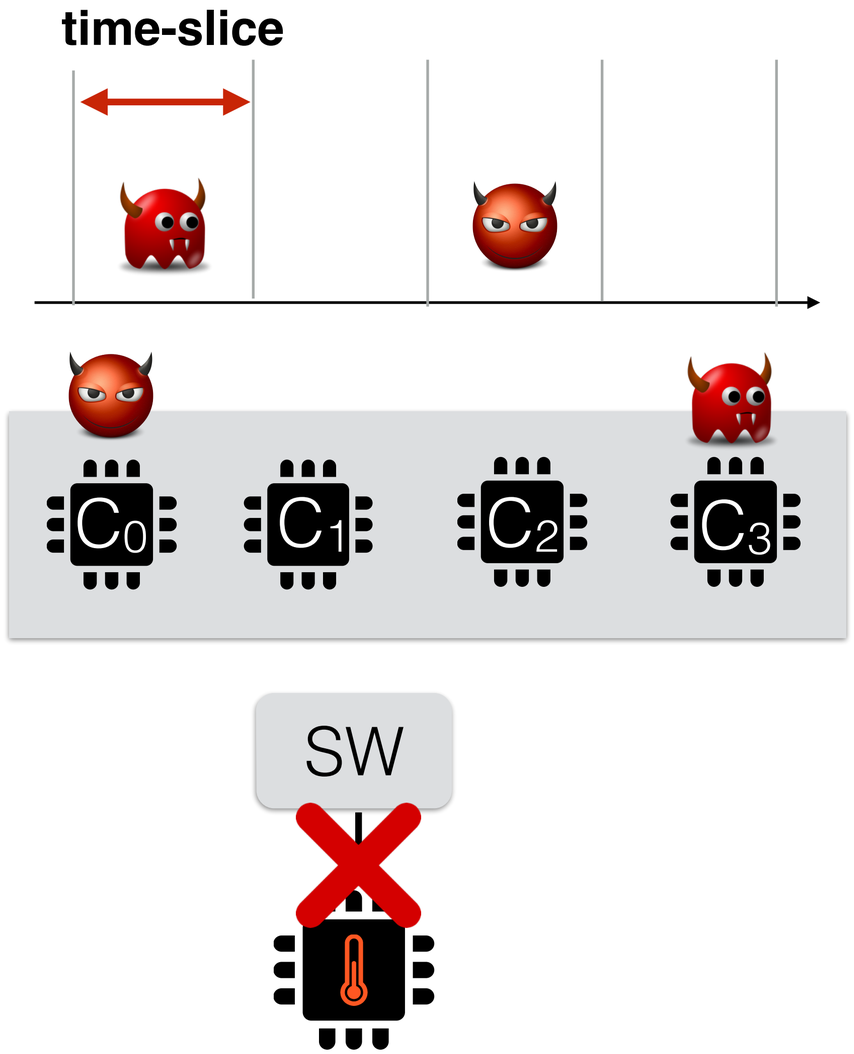
Countermeasures



Separate processes temporally and spatially

Waste of resources

Countermeasures

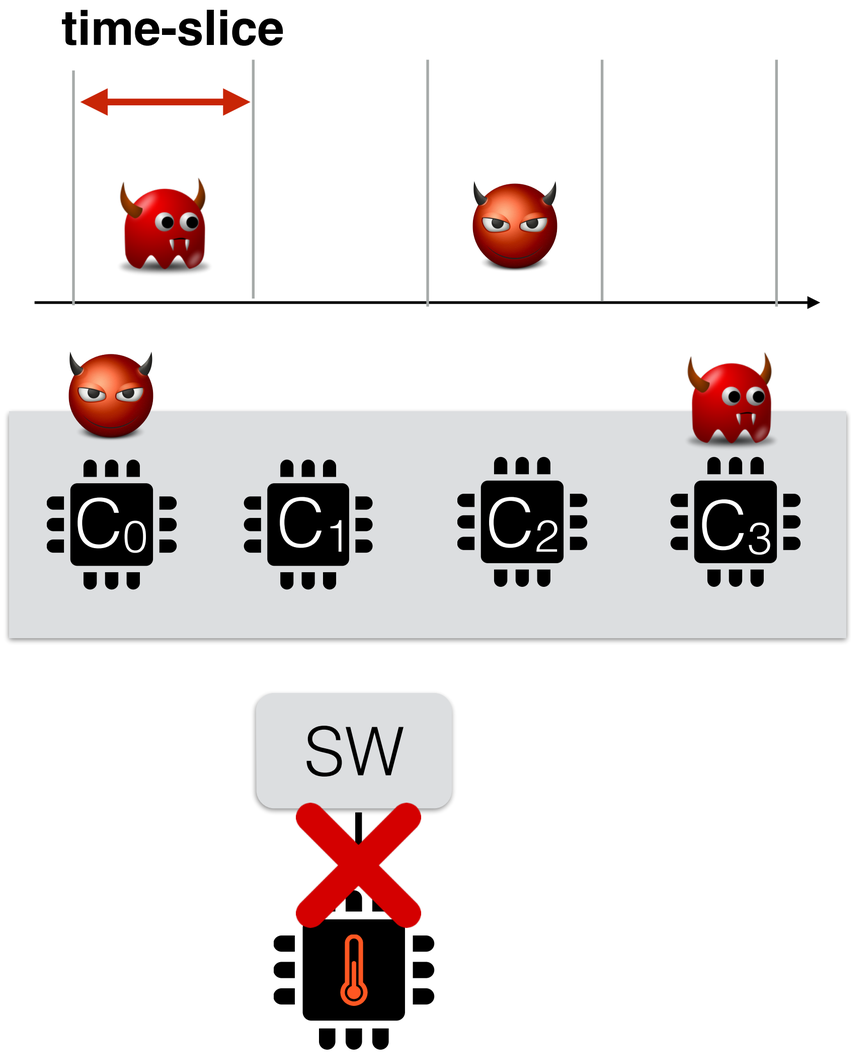


Separate processes temporally and spatially

Waste of resources

Restrict application access to thermal sensors

Countermeasures



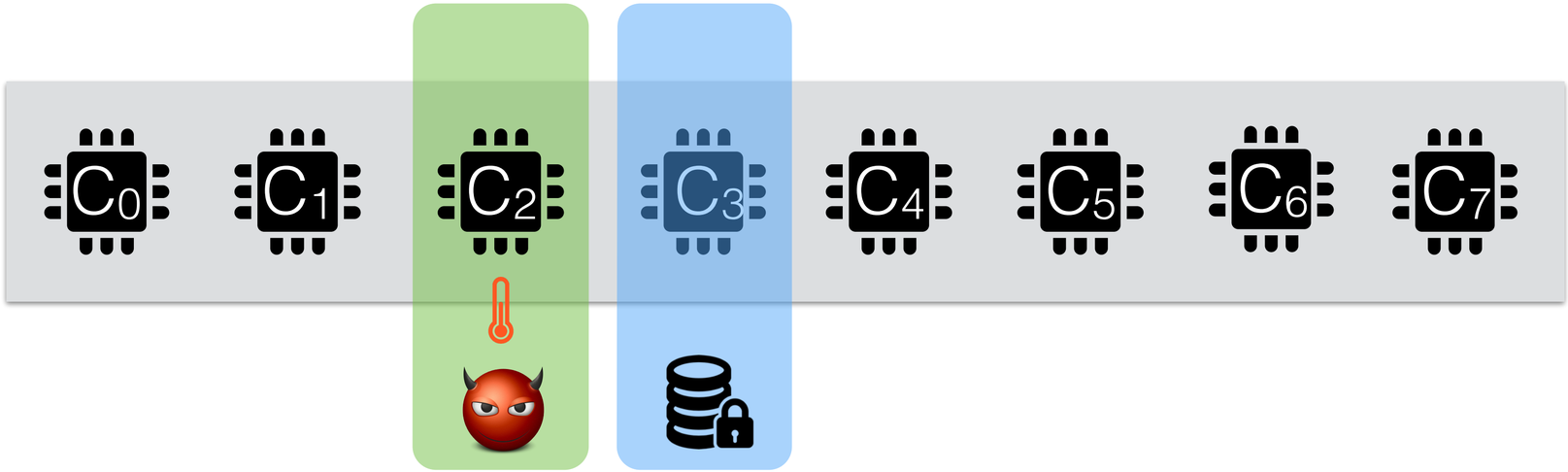
Separate processes temporally and spatially

Waste of resources

Restrict application access to thermal sensors

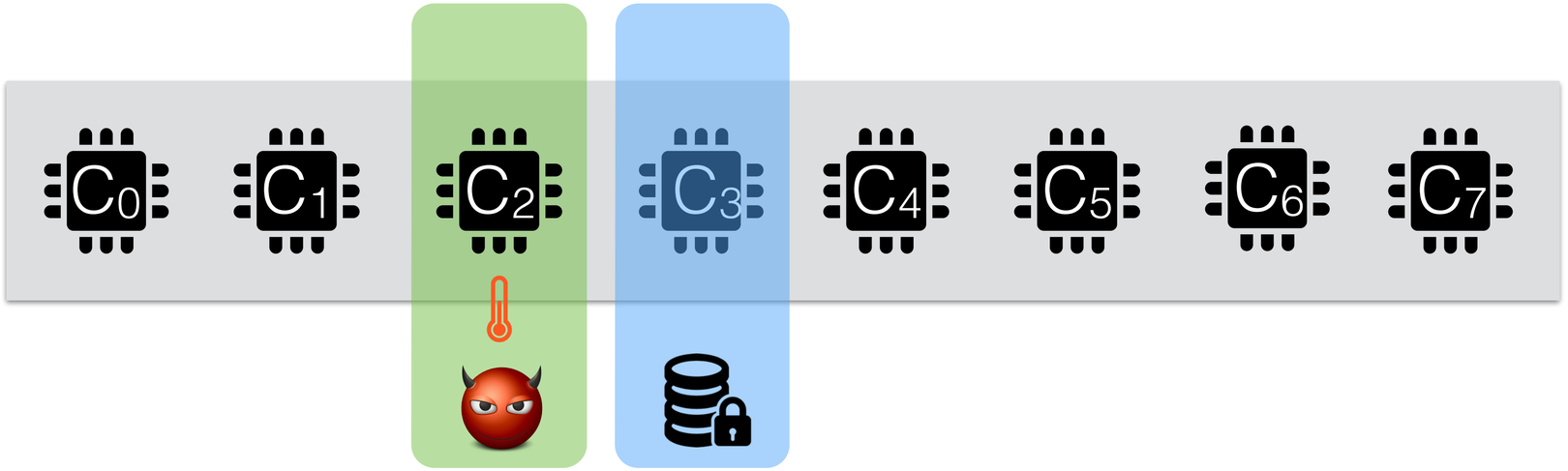
Thermal data is the key to building thermal-aware systems

Application Identification based on Thermal Traces



Is it possible for an attacker to exfiltrate information regarding the application running on core 3 by just observing its own core's thermal trace?

Application Identification based on Thermal Traces



Is it possible for an attacker to exfiltrate information regarding the application running on core 3 by just observing its own core's thermal trace?

Will more advanced metrics such as thermal models, machine learning based classifiers result more fine-grained data extraction? (e.g., secret keys)

Summary

- We demonstrated how thermal channels can be exploited to circumvent strong isolation guarantees provided by spatial and temporal resource partitioning schemes.
- Our work points to the limitation in the isolation guarantees achievable in modern multicore systems.
- Given the trend towards “energy efficient” computing, our work highlights the tension between designing thermally-aware energy efficient systems and their security guarantees.

Summary

- We demonstrated how thermal channels can be exploited to circumvent strong isolation guarantees provided by spatial and temporal resource partitioning schemes.
- Our work points to the limitation in the isolation guarantees achievable in modern multicore systems.
- Given the trend towards “energy efficient” computing, our work highlights the tension between designing thermally-aware energy efficient systems and their security guarantees.

email: aanjhan@inf.ethz.ch