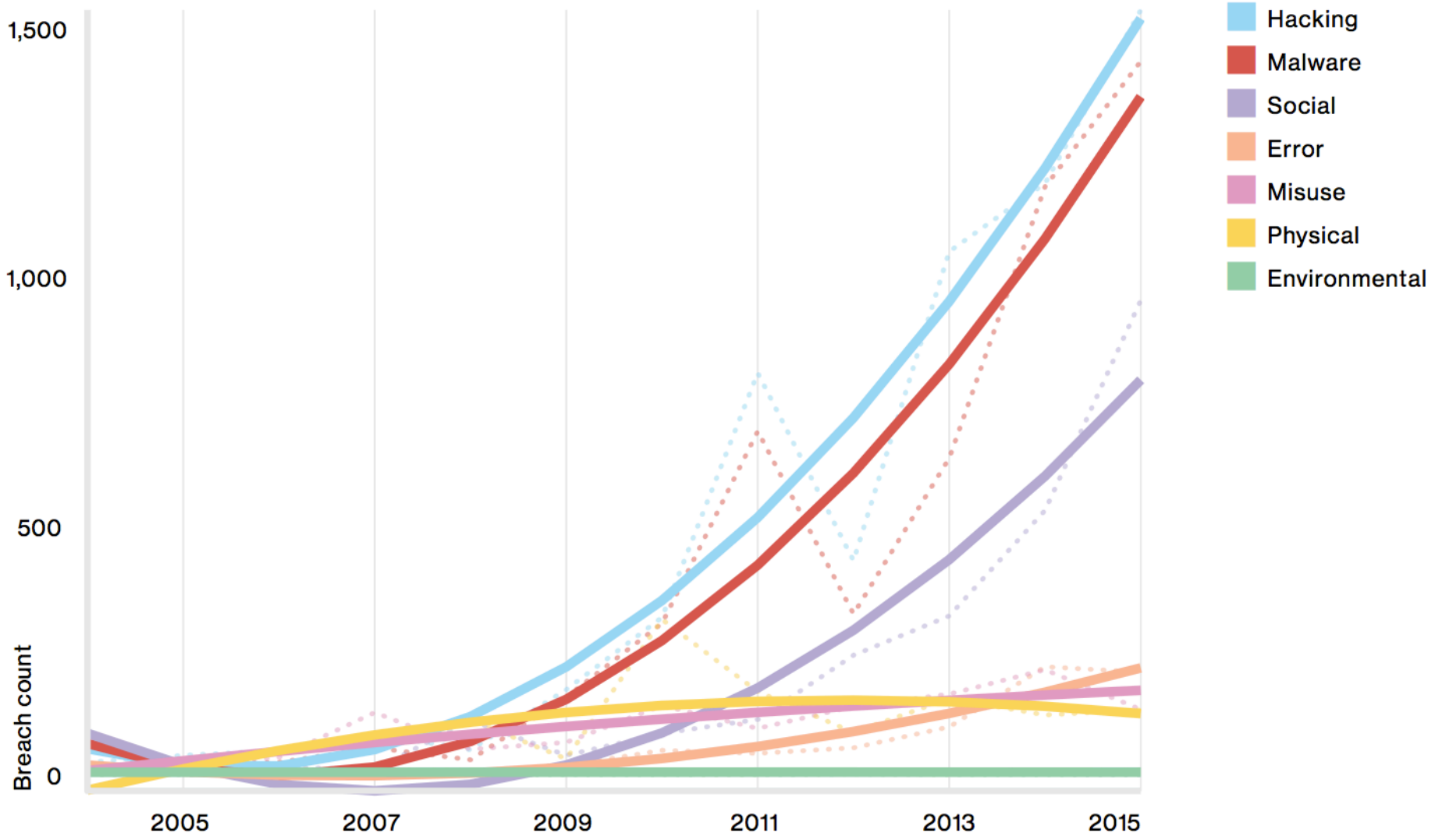
Behavioral Analysis for Intrusion Resilience

Ahmed Fawaz

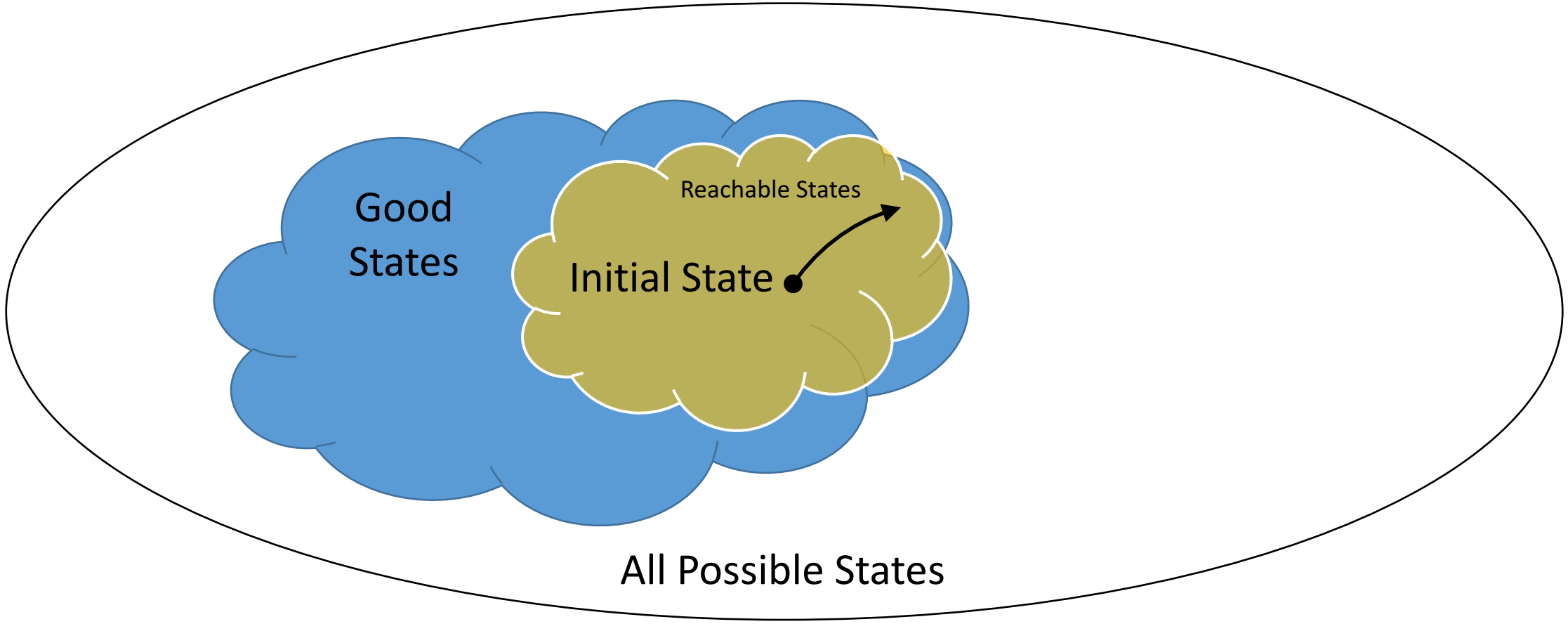
Dec 6, 2016

Recent Cyber Attacks on Private and Public Entities

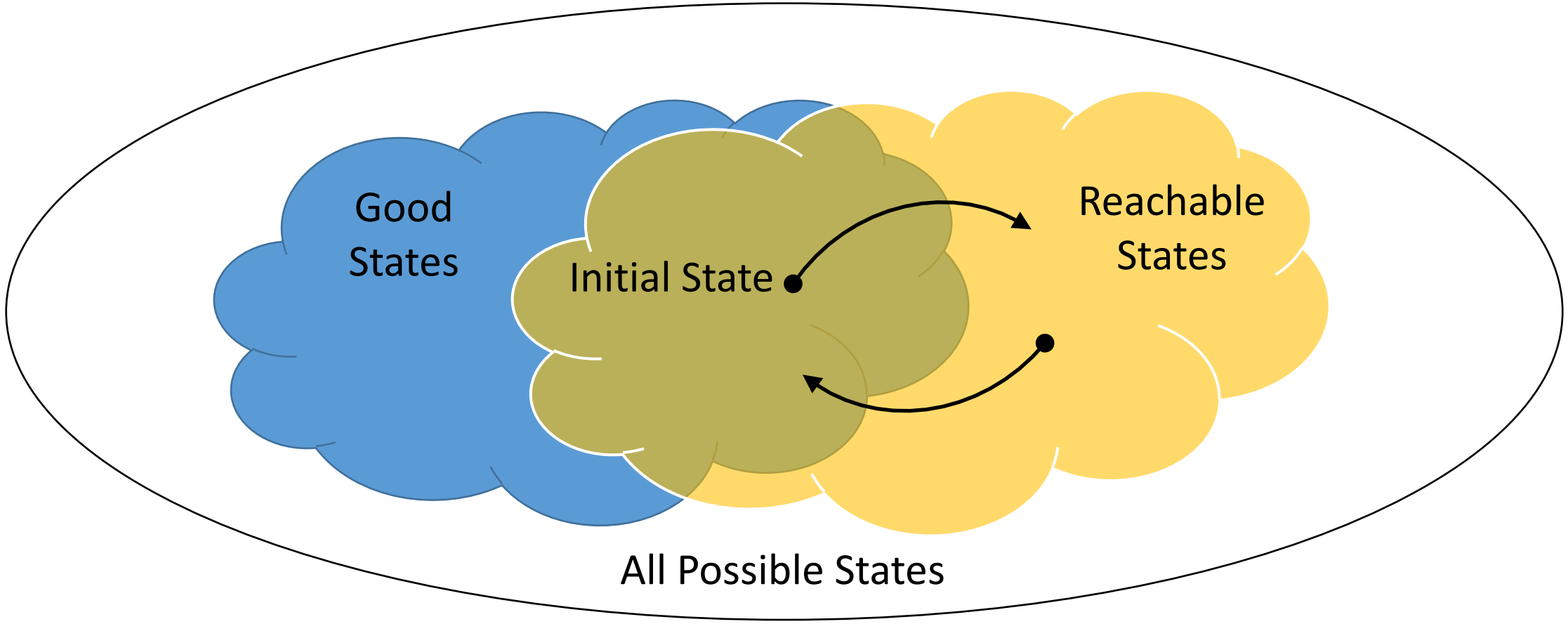


Source: Verizon 2016 Data Breach Investigation Report

Traditional Security



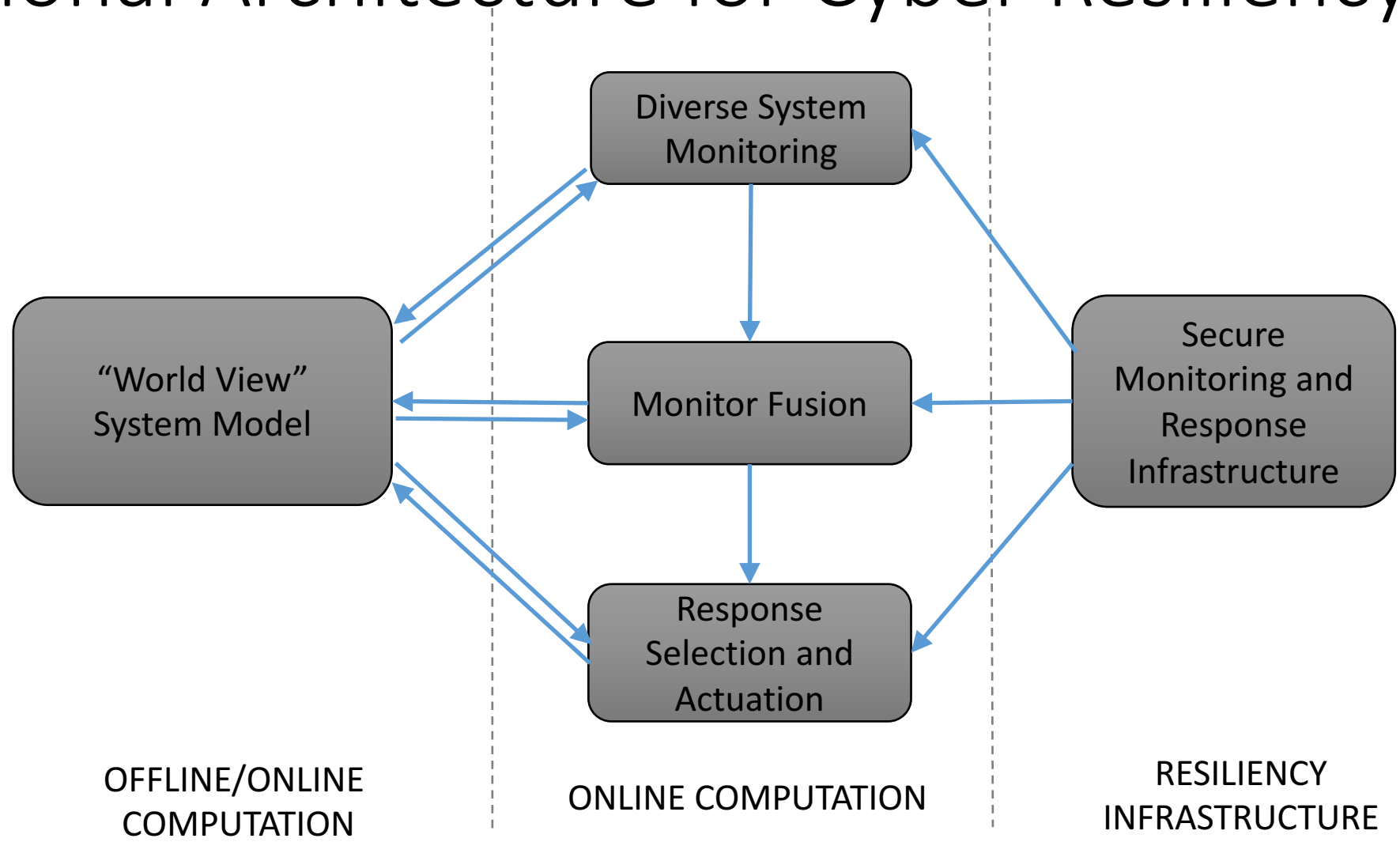
Resiliency Approach



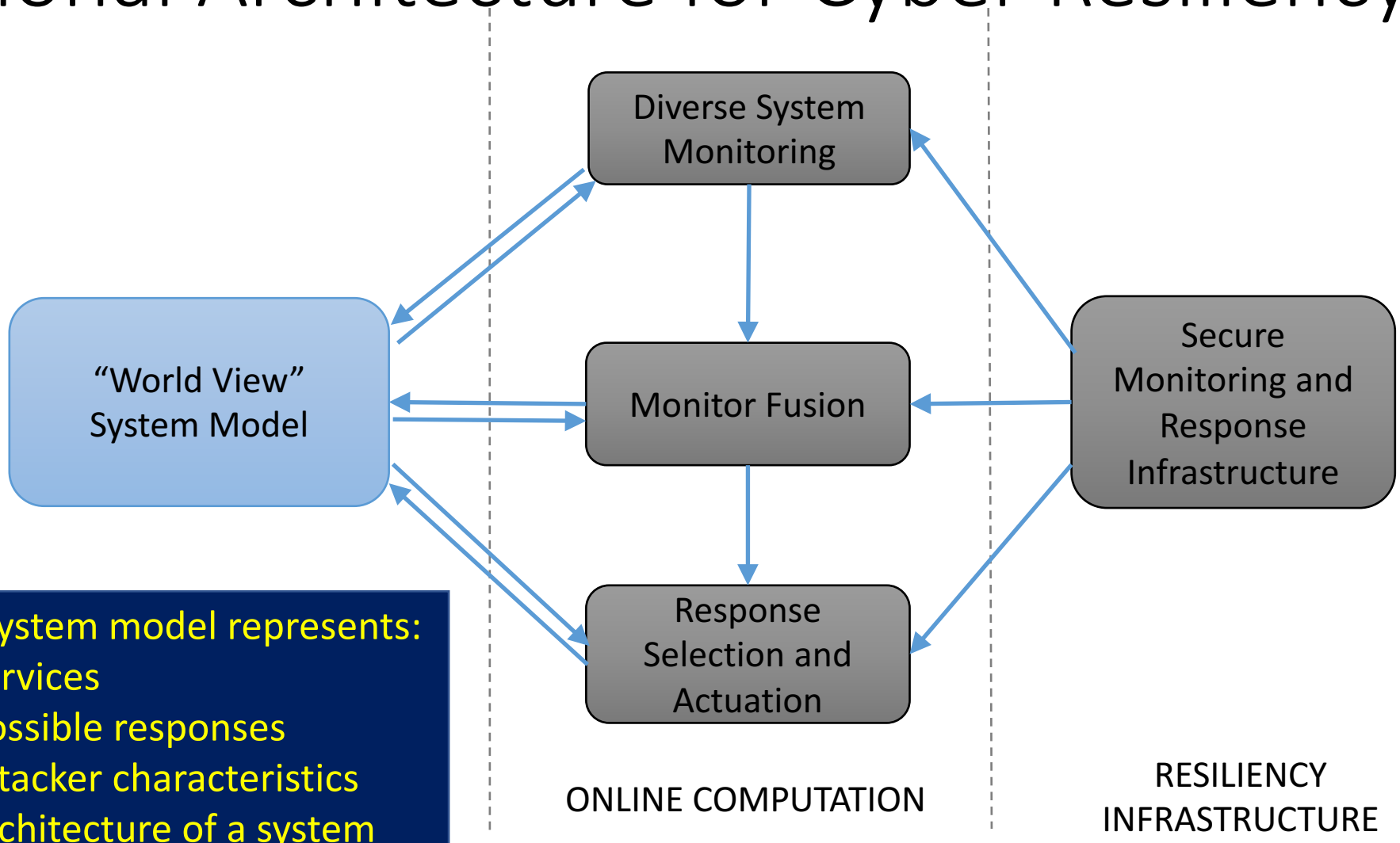
Cyber resilience is the ability to *identify*, *prevent*, *detect* and *respond* to malicious and random process or technology failures and recover— while maintaining an acceptable level of service.

Adapted from:
Presidential Policy Directive 21 (PPD-21), "Critical Infrastructure Security and Resilience," February 12, 2013.
Accenture Consulting, Making your Enterprise Cyber Resilient, 2015

Notional Architecture for Cyber Resiliency



Notional Architecture for Cyber Resiliency

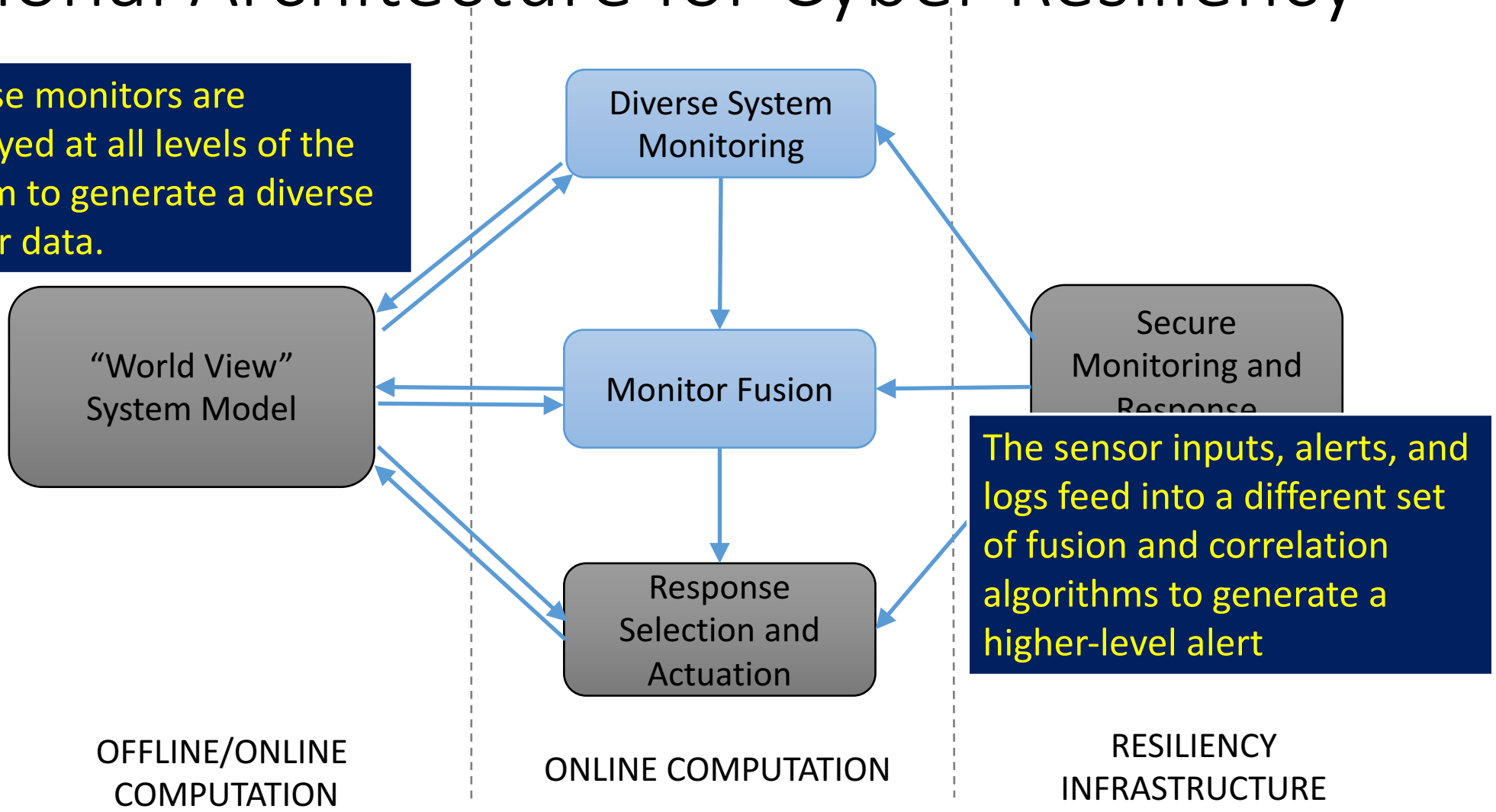


The system model represents:

- services
- possible responses
- attacker characteristics
- architecture of a system

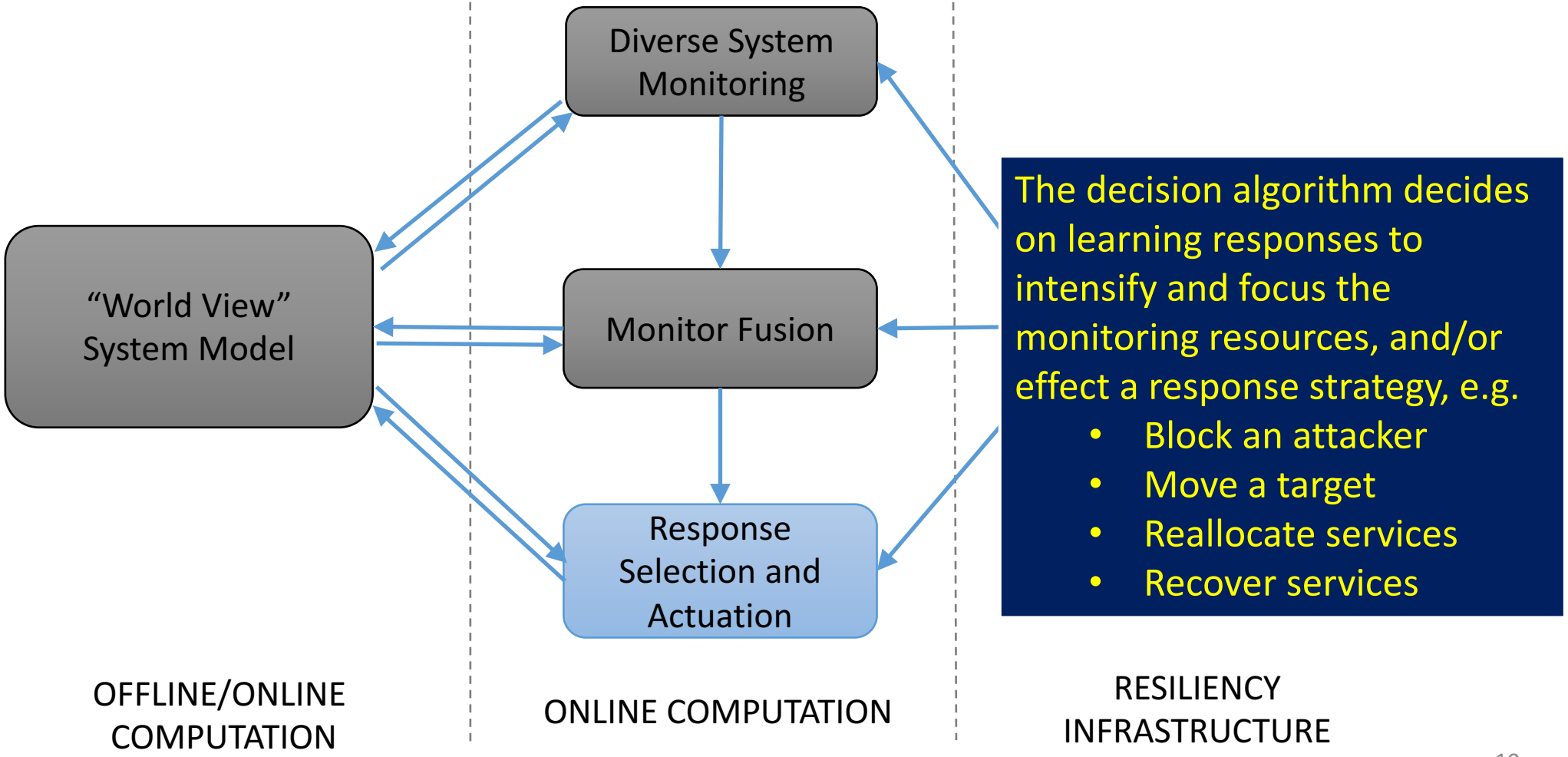
Notional Architecture for Cyber Resiliency

Diverse monitors are deployed at all levels of the system to generate a diverse sensor data.

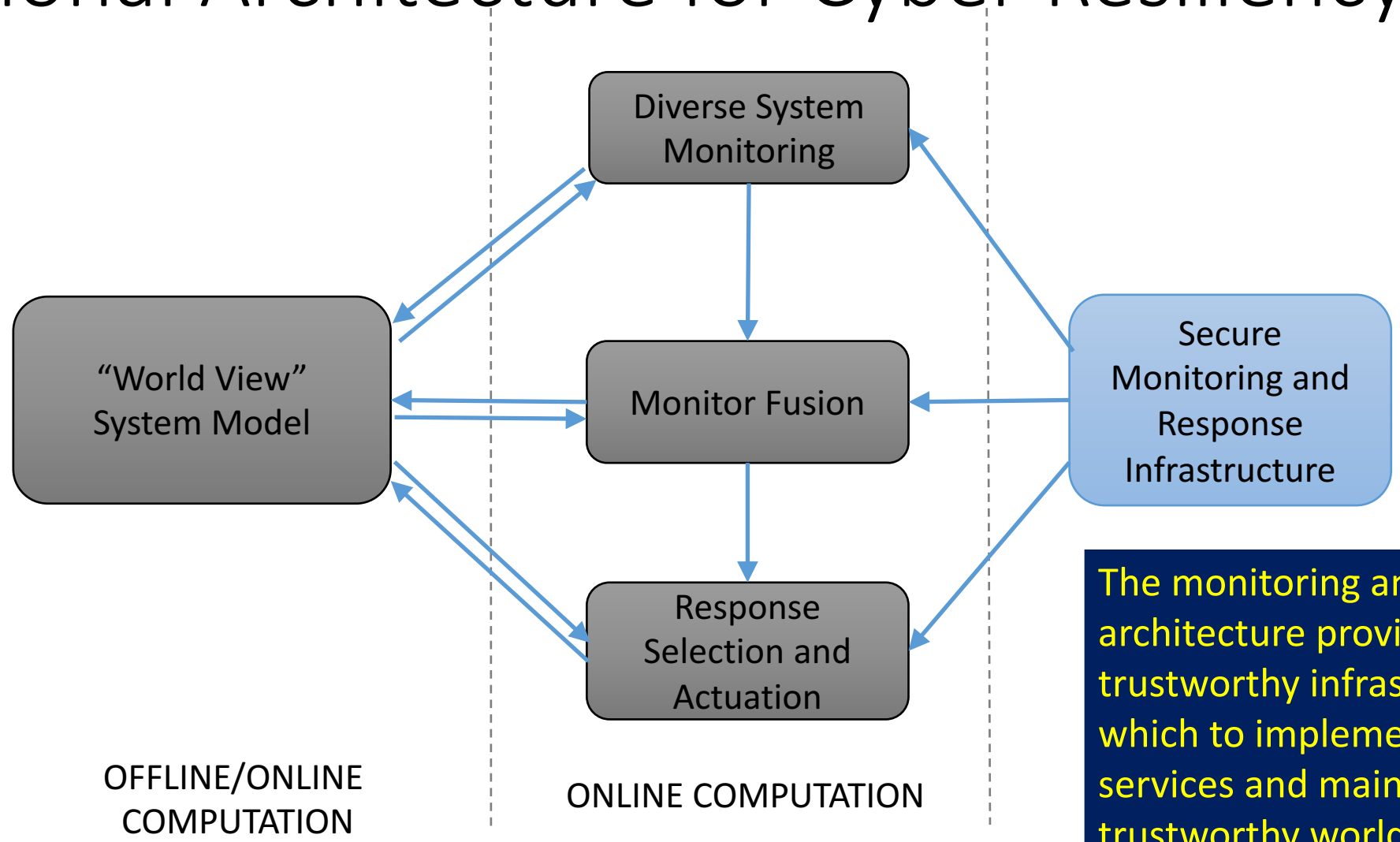


The sensor inputs, alerts, and logs feed into a different set of fusion and correlation algorithms to generate a higher-level alert

Notional Architecture for Cyber Resiliency

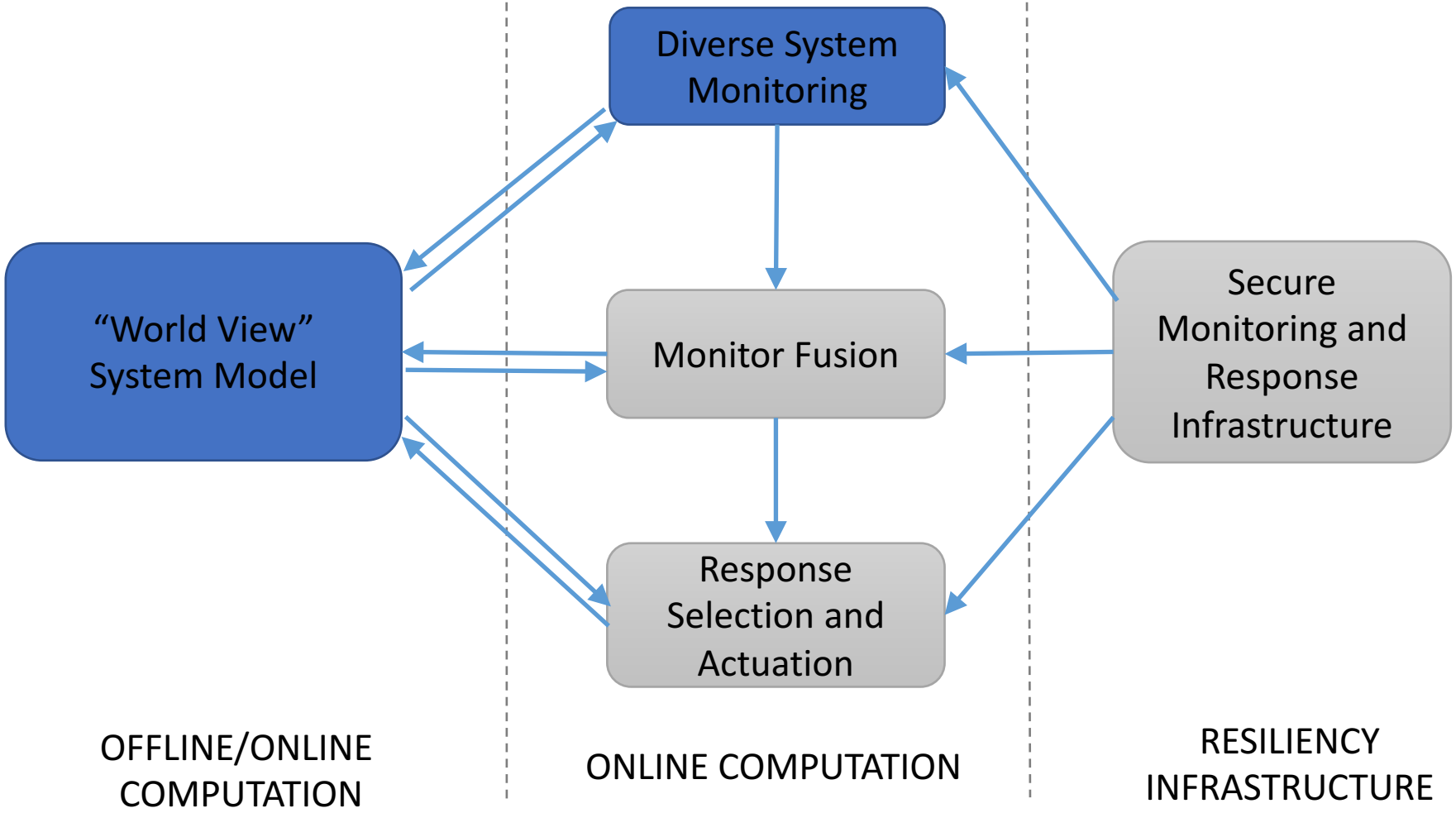


Notional Architecture for Cyber Resiliency



The monitoring and response architecture provides a trustworthy infrastructure on which to implement resiliency services and maintain a trustworthy world view.

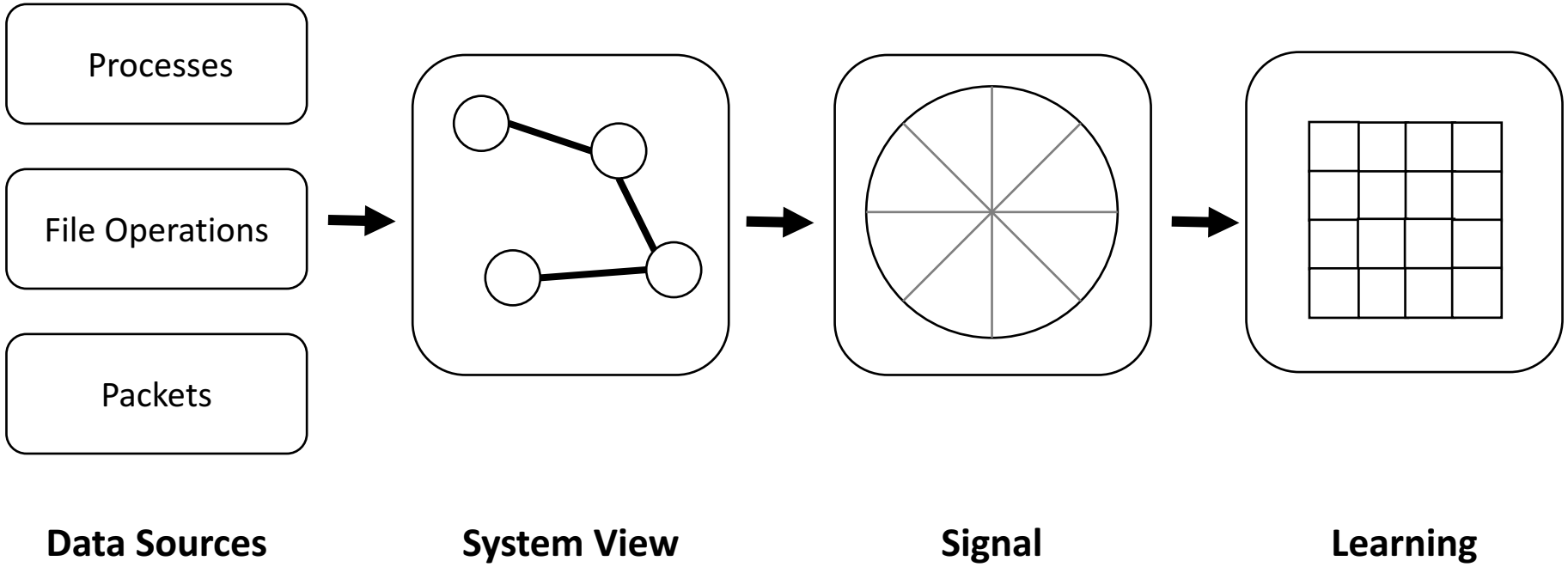
Kobra: A Kernel Monitoring Engine



Problem Description

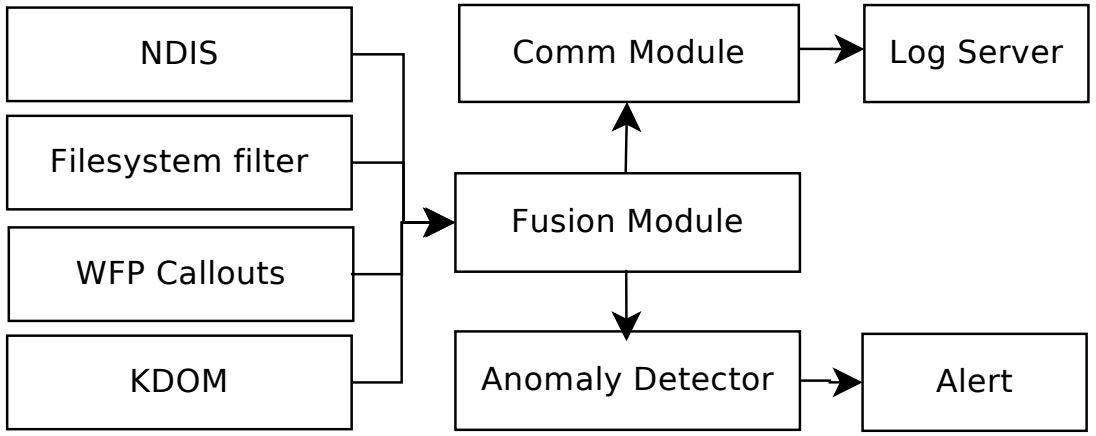
How to use diverse data types to model application behavior for anomaly detection?

Our Approach



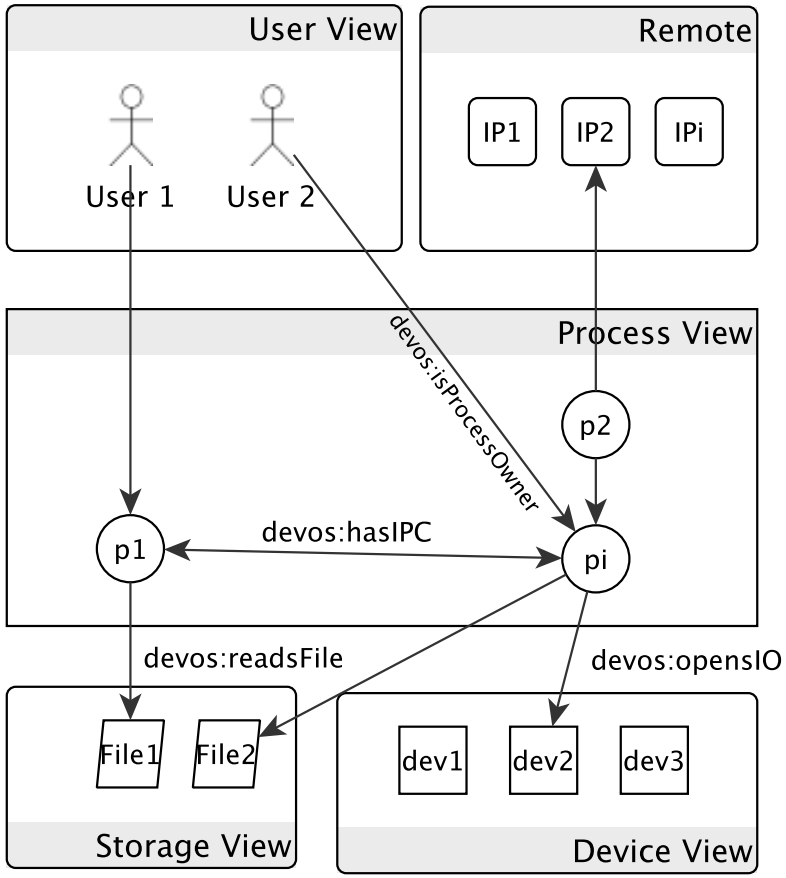
Kobra's Architecture

- Kernel-level monitor for Windows kernel
- Cooperative drivers that captures:
  - Network activity
  - Process communications
  - Process creation/termination
  - Objects access
  - File system activity
- Low-overhead



What is the System View?

- The intent of the system view is to provide high-level information about host state.
- Reflects the methods by which users and user processes access different resources.



File and Network behaviors

```
insert-edge{“VID.mp4” :?”
  2044} {devos:read} {512}
```

```
insert-edge{“VID.mp4” :?”
  2044} {devos:read} {512}
```

```
insert-edge{“VID.mp4” :?”
  2044} {devos:read} {4096}
```

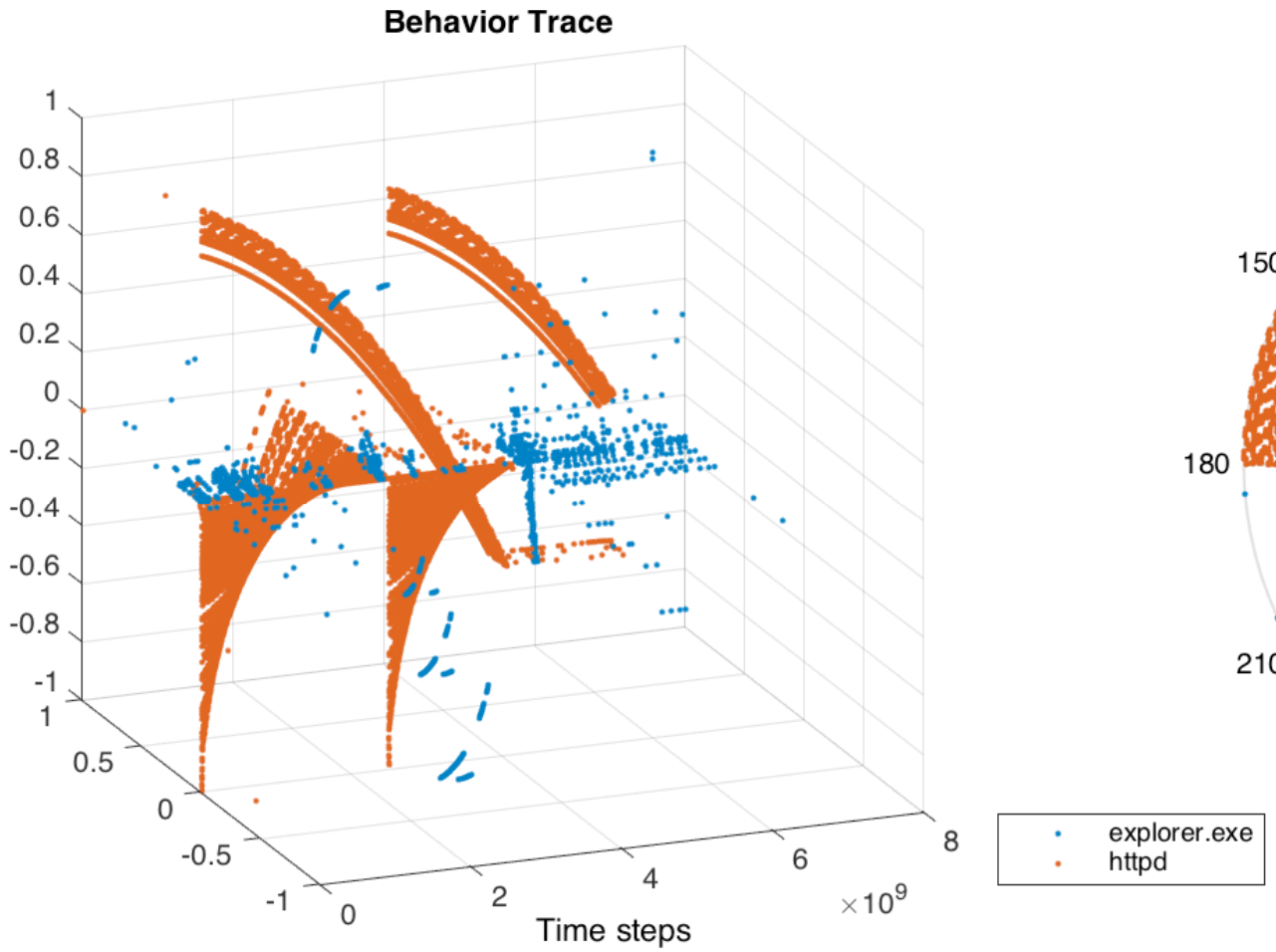
- Chromium
- VLC

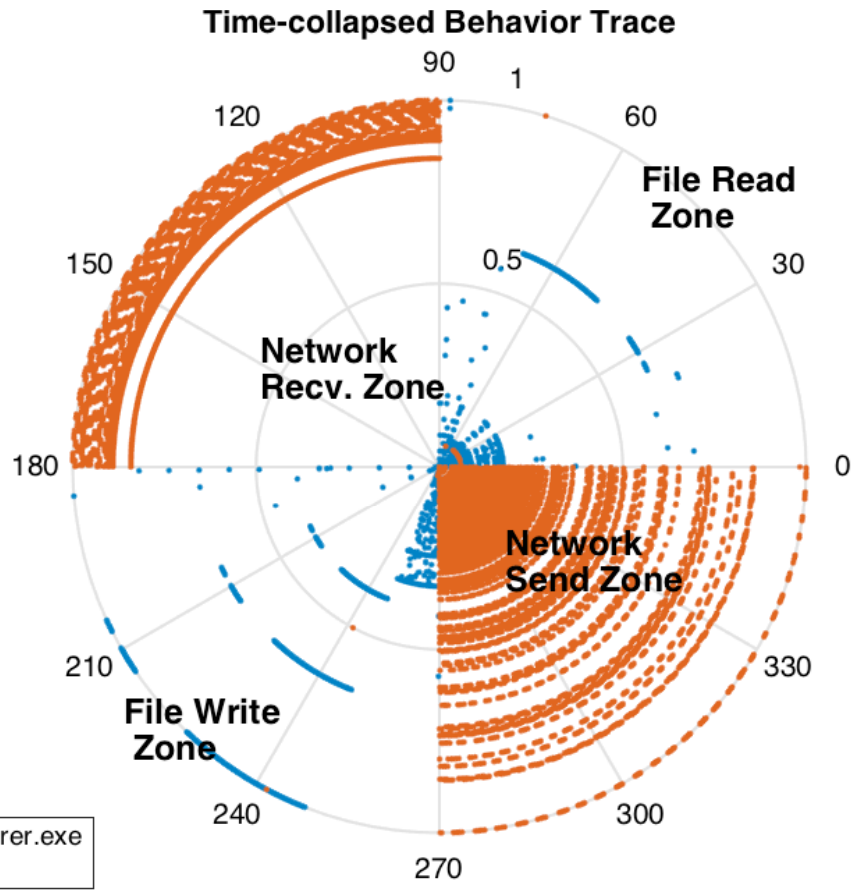
- Filter by Process and Application
- Data is converted to a discrete time signal

Map Discrete Events to a Polar Space

- Mapping inspired by digital modulation methods
- Partition space by quadrants according to type of events
- Map each event to a part of the quadrant
- The magnitude is a function of the “size” of event

Exampled


Behavior Trace
Time steps
×10^9
explorer.exe
httpd

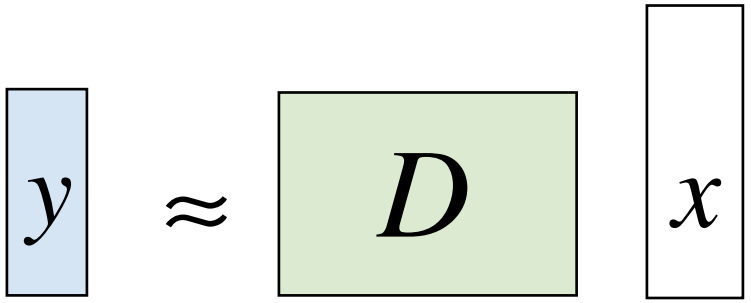

Time-collapsed Behavior Trace
File Read Zone
Network Recv. Zone
Network Send Zone
File Write Zone

Application Behavior Model

- Learn local patterns in the signal (sliding window)
- Learn the co-occurrence relationships between the patterns
- Model: **<Local Patterns, Co-occurrence>**

Learning Local Patterns

- Learn sparse representation dictionary on the time signals
- Dictionary atoms correspond to the local patterns



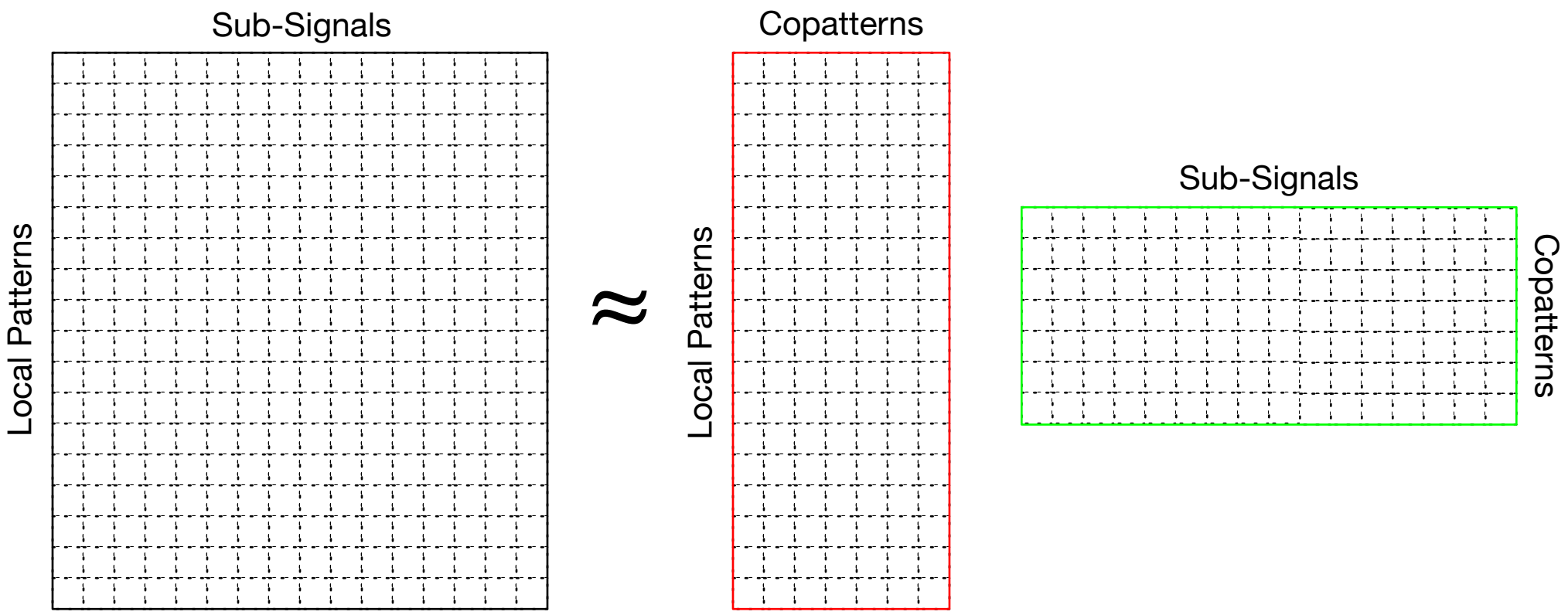
$y \in \mathbb{R}^n$ Input Signal

$D \in \mathbb{R}^{n \times p}$ Dictionary

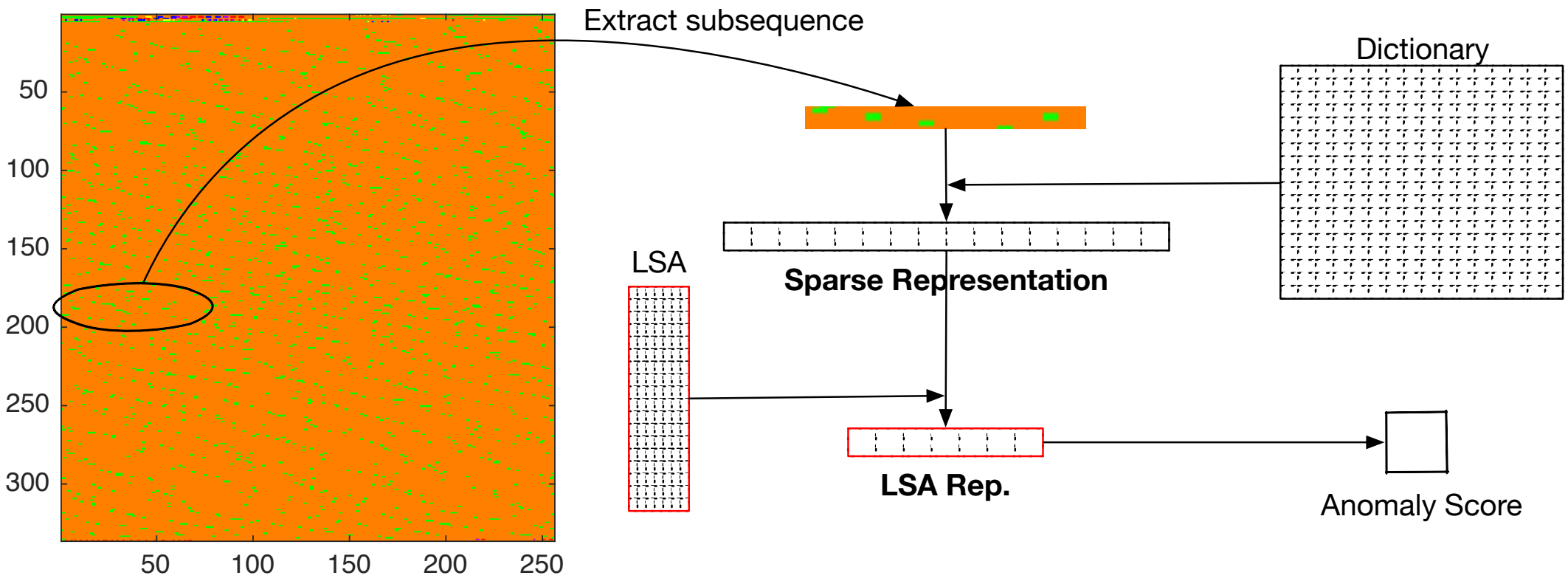
$x \in \mathbb{R}^p$ Sparse Approximation

$$\mathbf{D} = \arg\min_{D} \sum_{i=1}^{n} \min\{\|Dx_i - y_i\|^2 + \lambda \|x_i\|_1\}$$

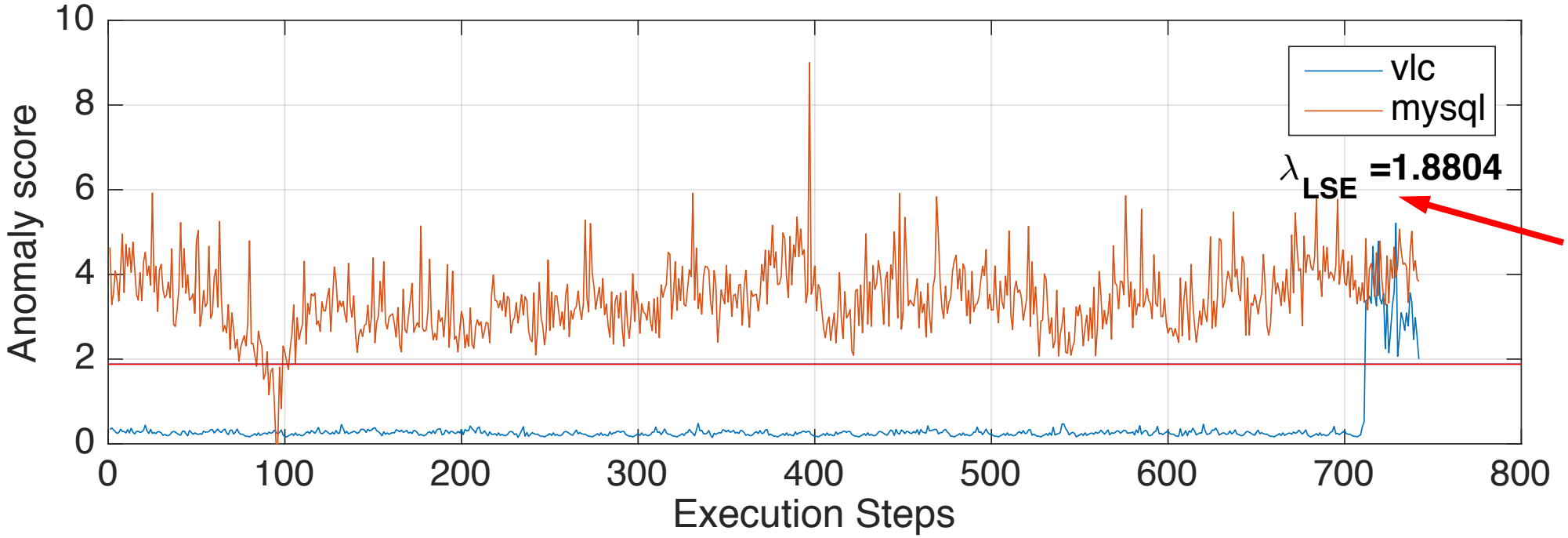
Learning Co-occurrence (LSA)



Anomaly Detection using Model



Reconstruction of MySQL using VLC Model



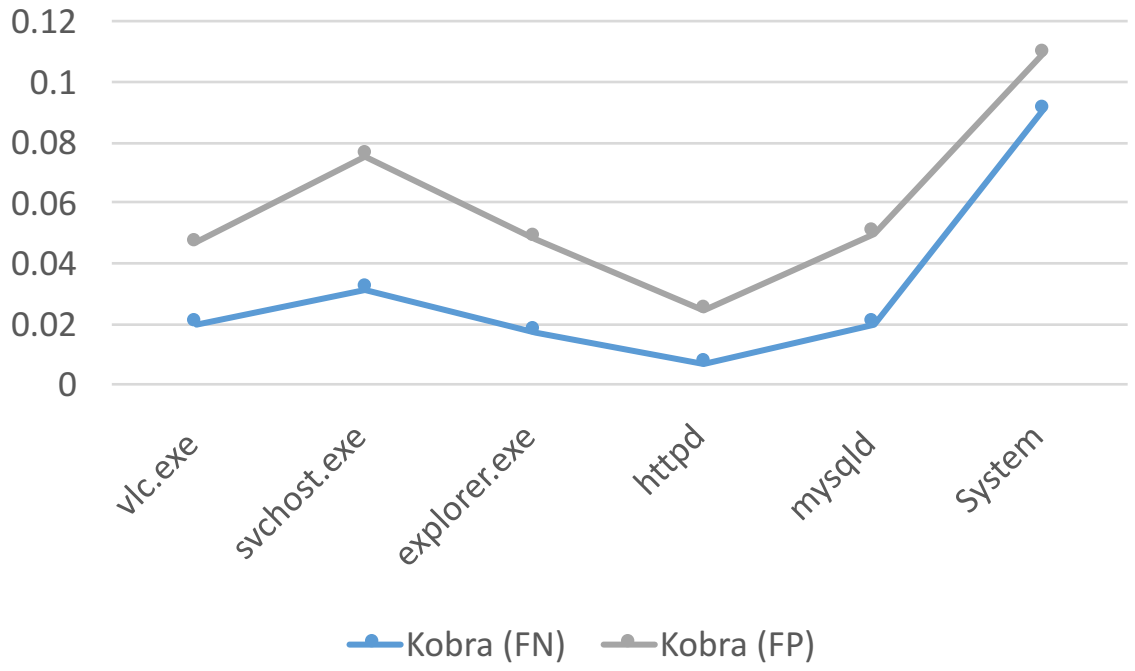
95th percentile of reconstruction error

Evaluation Methodology

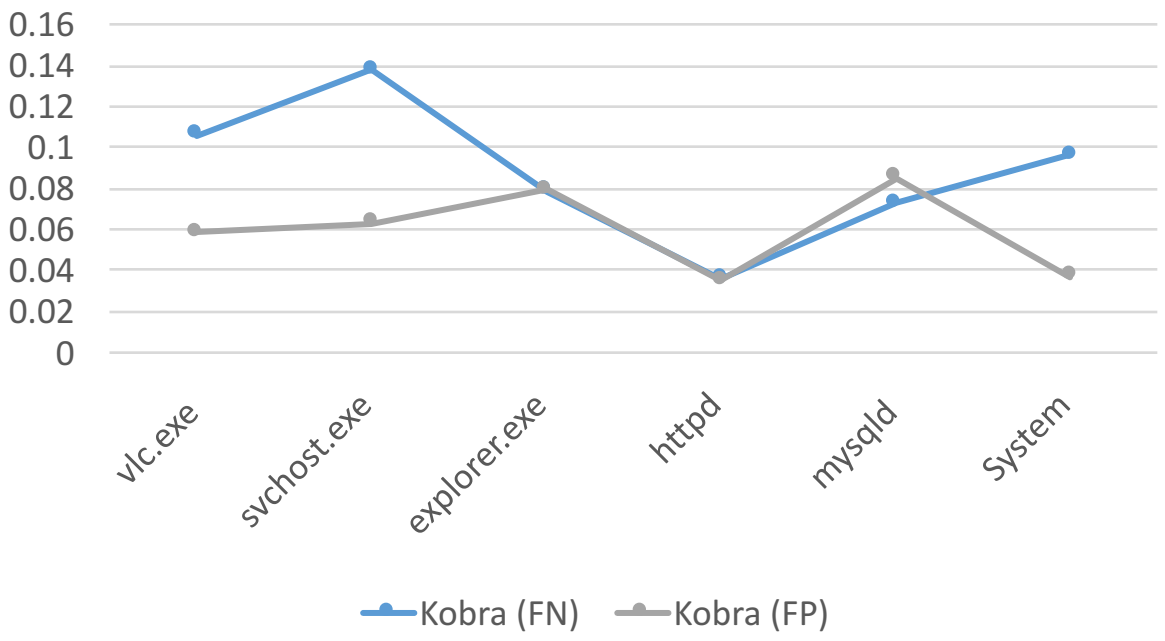
1. Generate traces of normal behavior of application
   - VLC playing local files
   - Apache + Mysql running wordpress
   - Windows services

2. Learn model of each application
3. Inject traces of shellcode behavior into testing traces
4. Compute anomaly scores

Evaluation Results

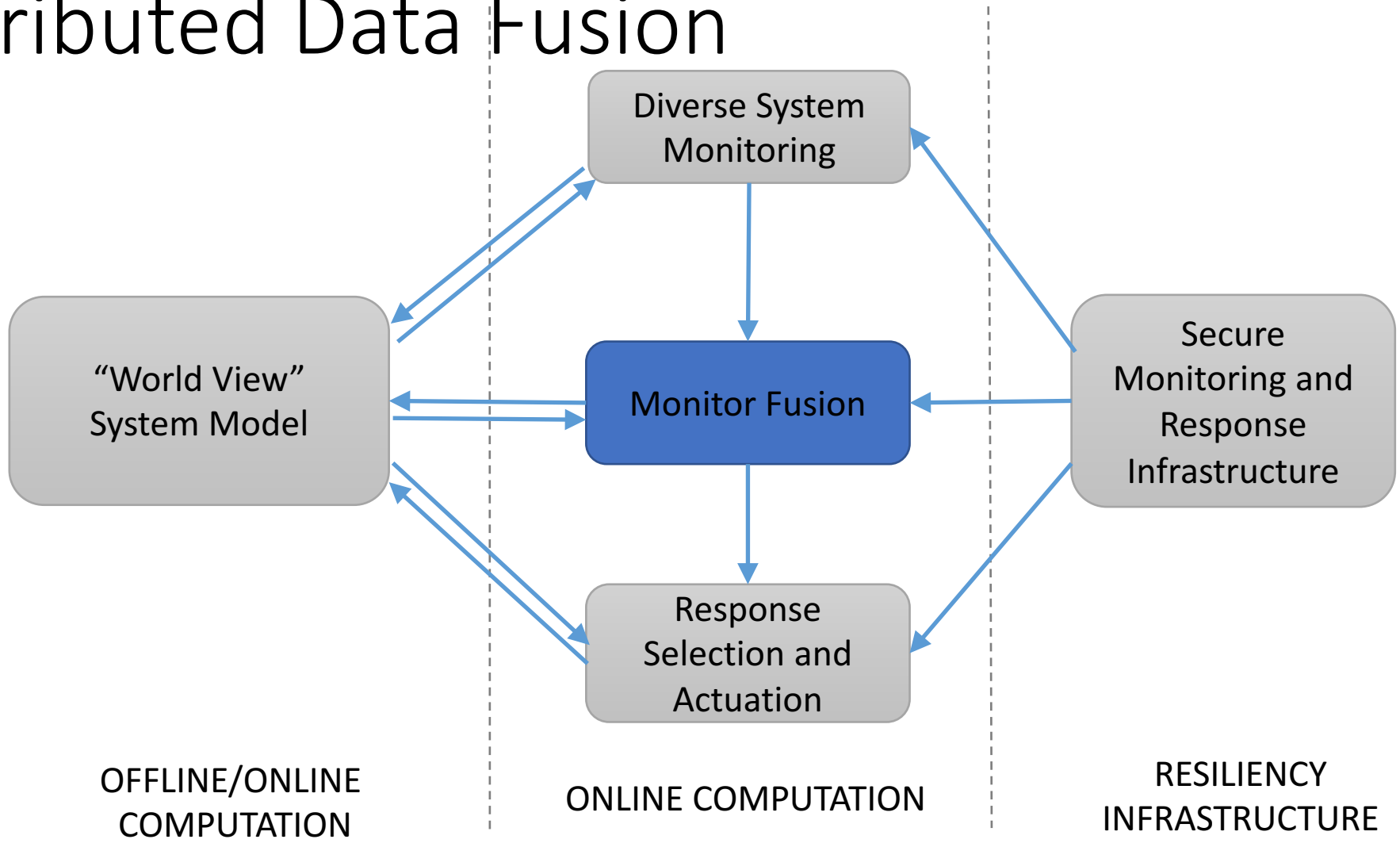


False Positive/Negative Rates for Reverse Shell



False Positive/Negative Rates for Drive-by-Download

Lateral Movement Detection Using Distributed Data Fusion

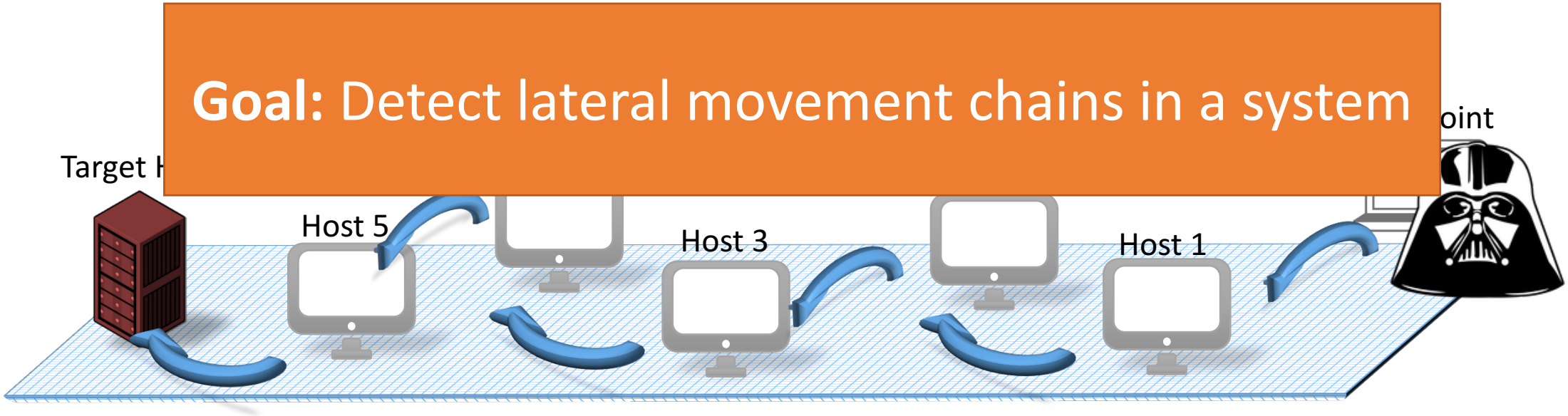


Problem Description

How do we fuse diverse data sources using a distributed agent-based system to detect lateral movement in a network while maintaining scalability?

Lateral Movement Explained

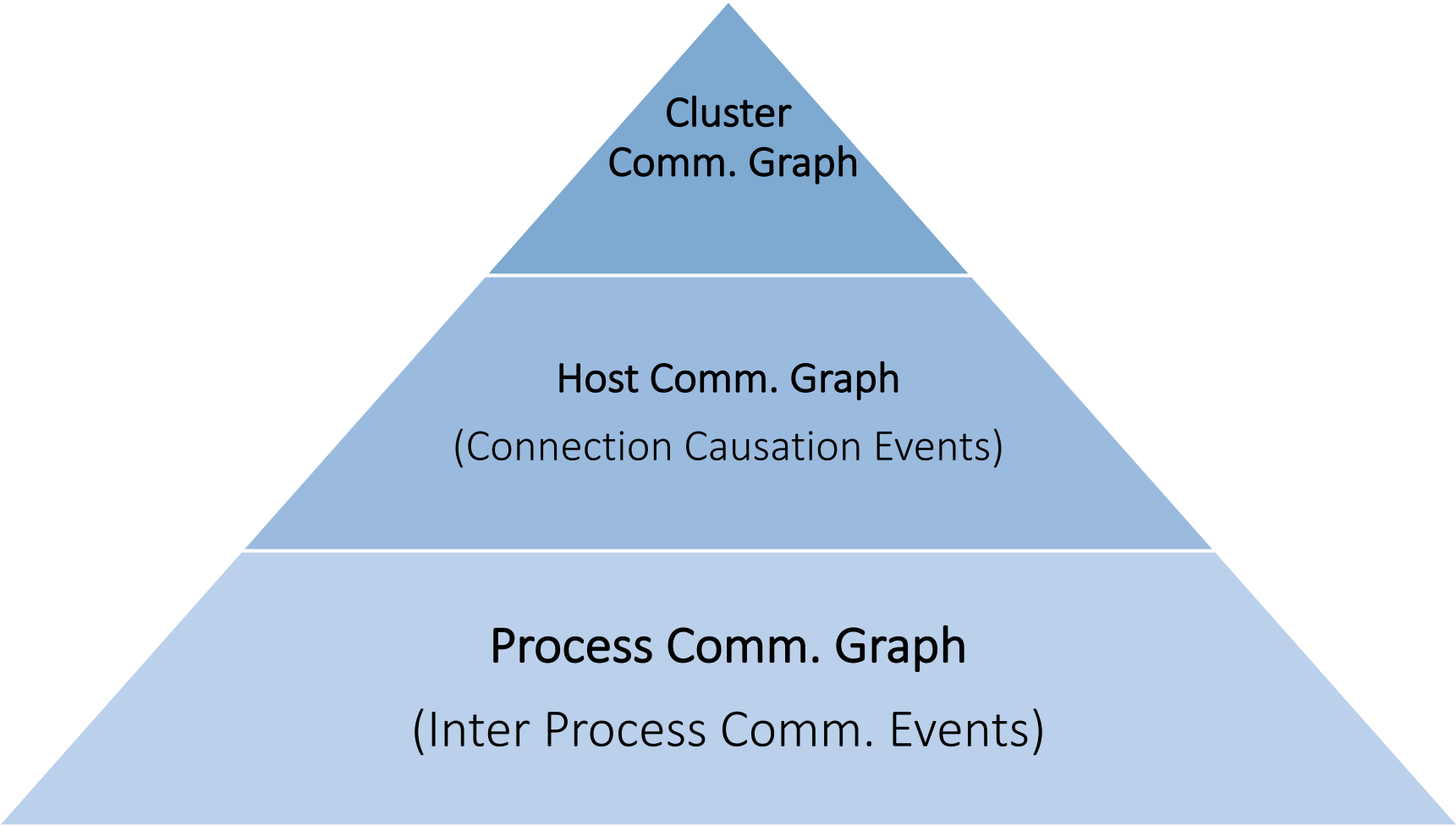
- Starting from the entry point attacker moves to target host
- Uses system services or custom tools



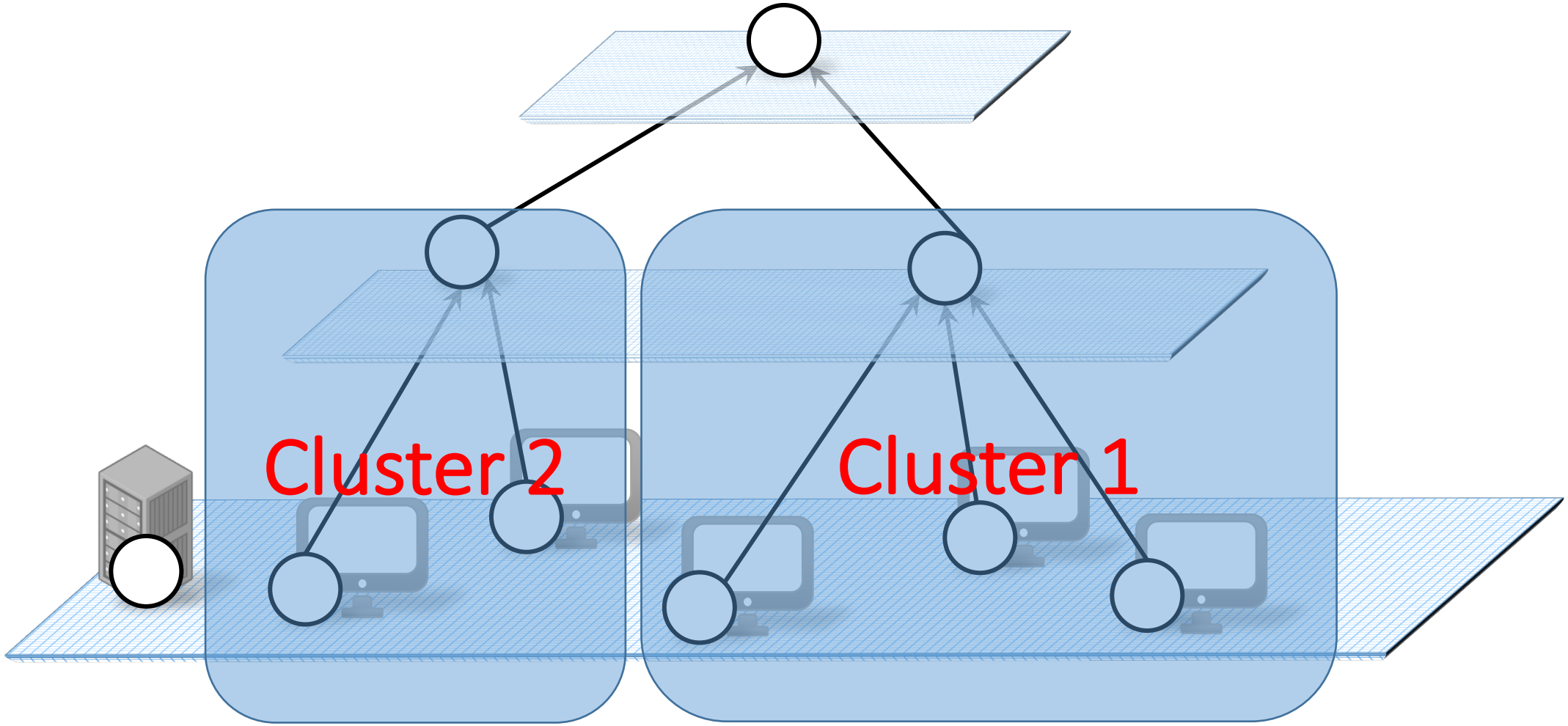
State-of-the-art

- Centrally correlate NetFlows to detect lateral movement
- NetFlow correlation method is not accurate
- Amount of information is too large to be handled centrally

Approach Overview

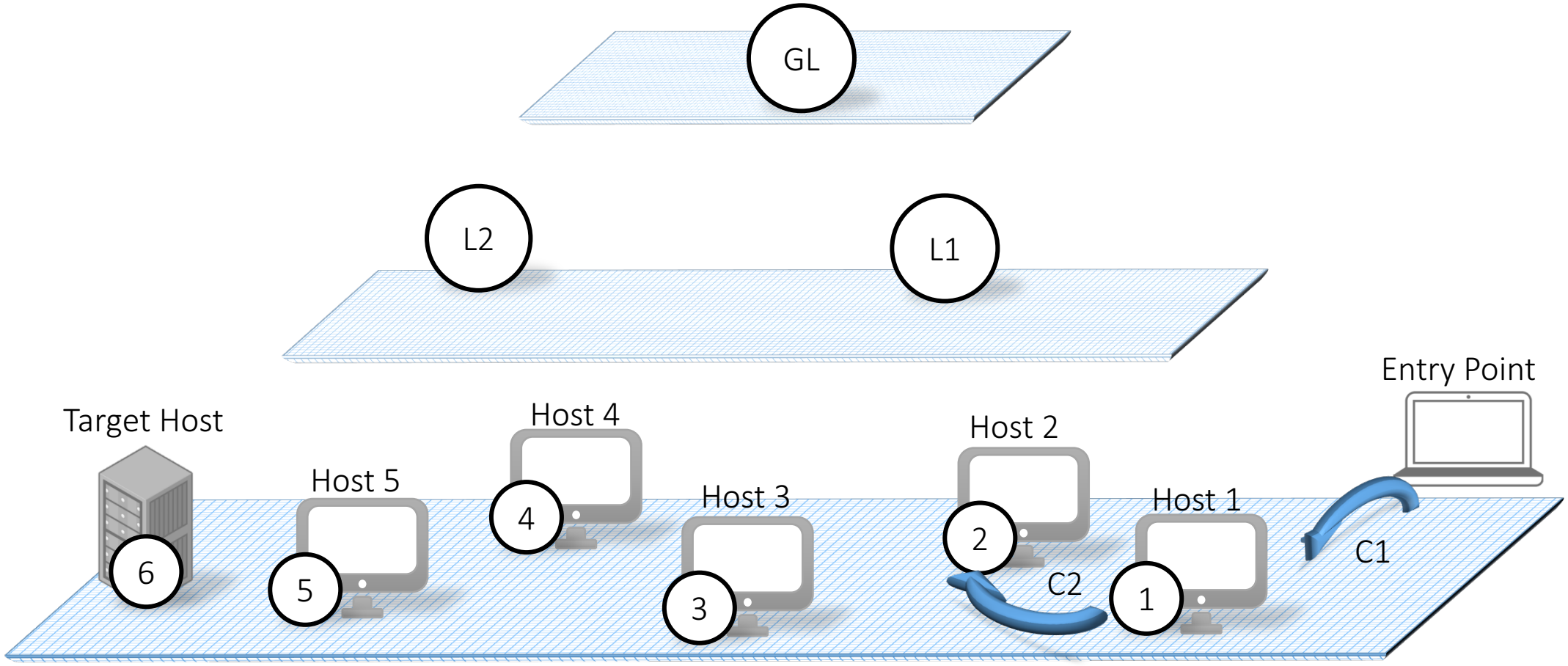


System Model



Lateral Movement

A critical step during APT to move from the entry point to target host



Inside Host 1

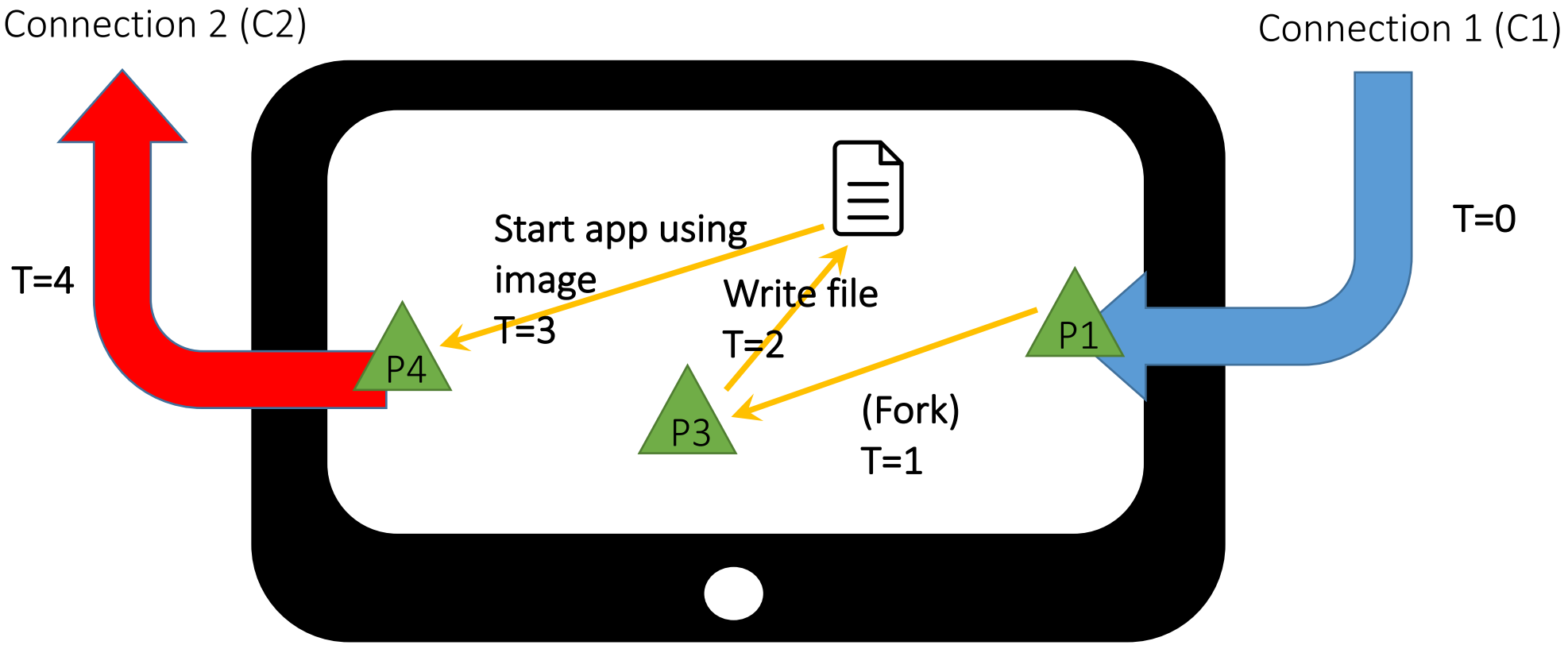
Process Communication View created by Kobra using timestamped events:

- Processes running
- Process communication (pipes, messages,…)
- Network connections (with a unique ID across system)
- File access

Connection causation event is generated when the agent find a path between incoming and outgoing connections

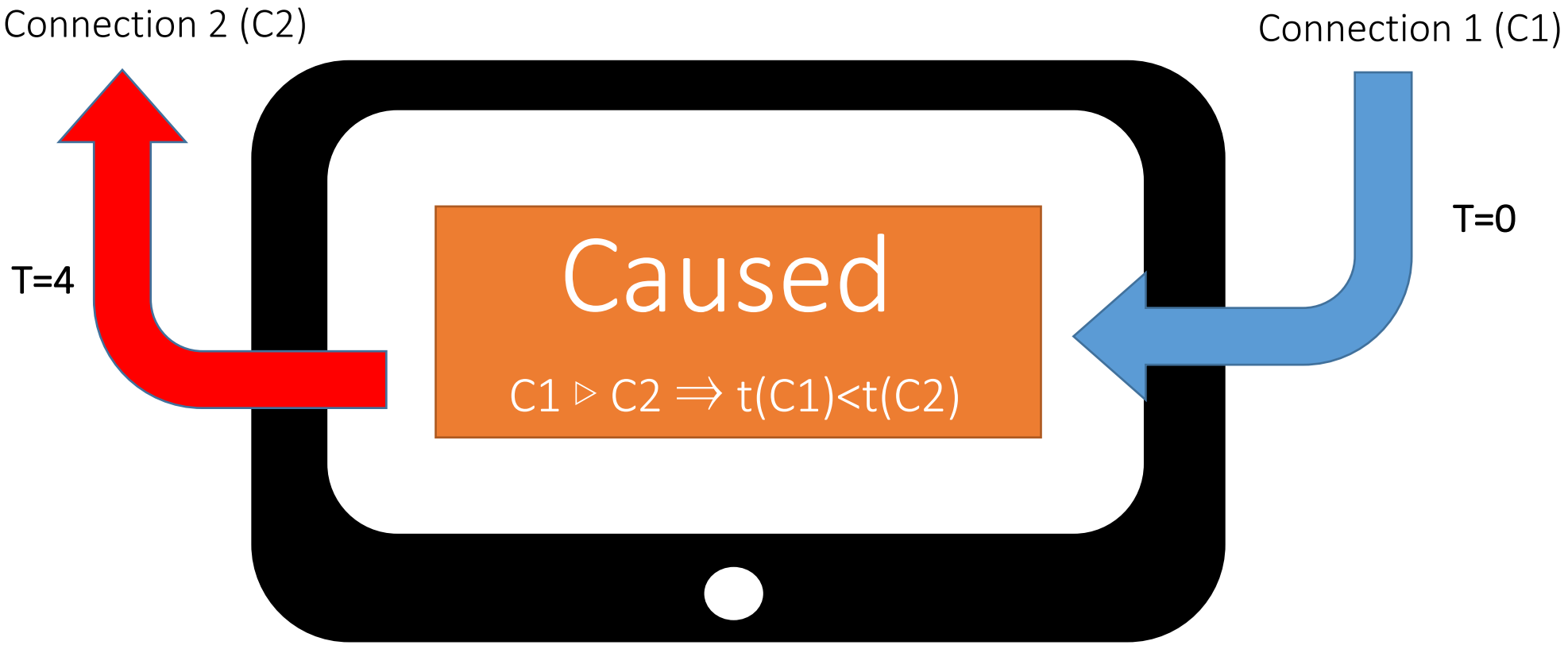
Inside Host 1

Local agent infers connection causation using the **Process Communication Graph**



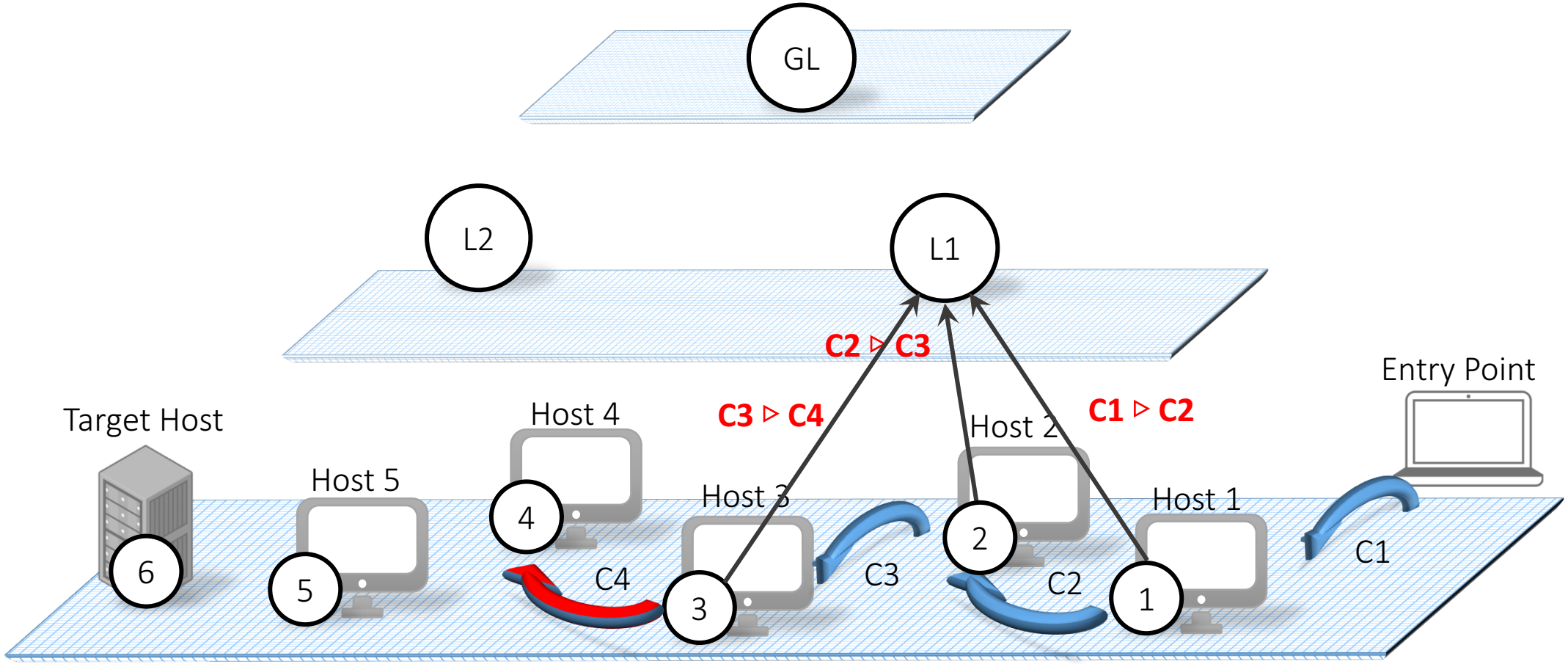
Inside Host 1

Local agent infers connection causation using the **Process Communication** Graph



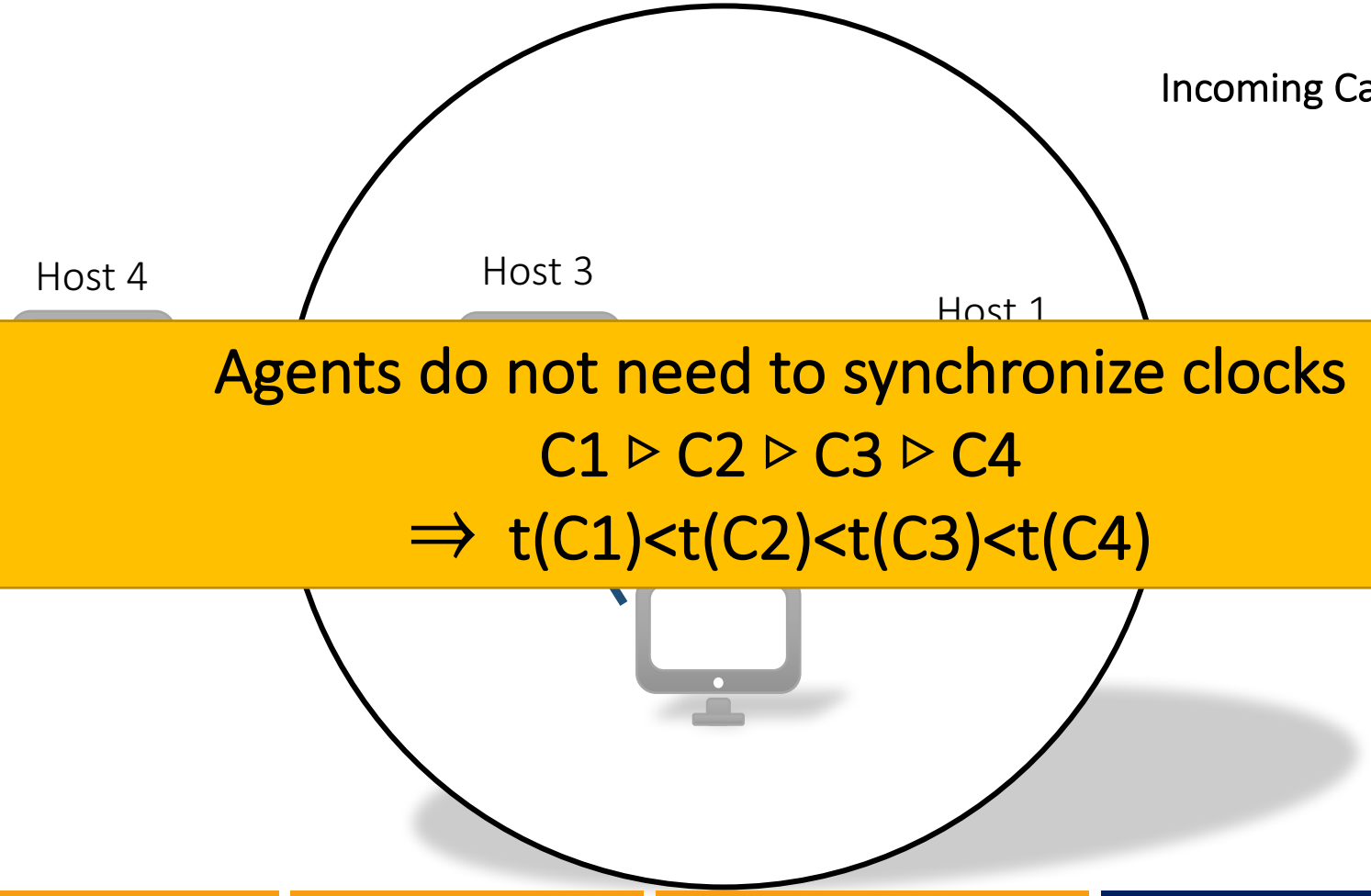
Lateral Movement

A critical step during APT to move from the entry point to target host



Inside Cluster Leader 1

Cluster head maintains **Host Communication** Graph



Incoming Causation Events:

C1 ▹ C2

C2 ▹ C3

C3 ▹ C4

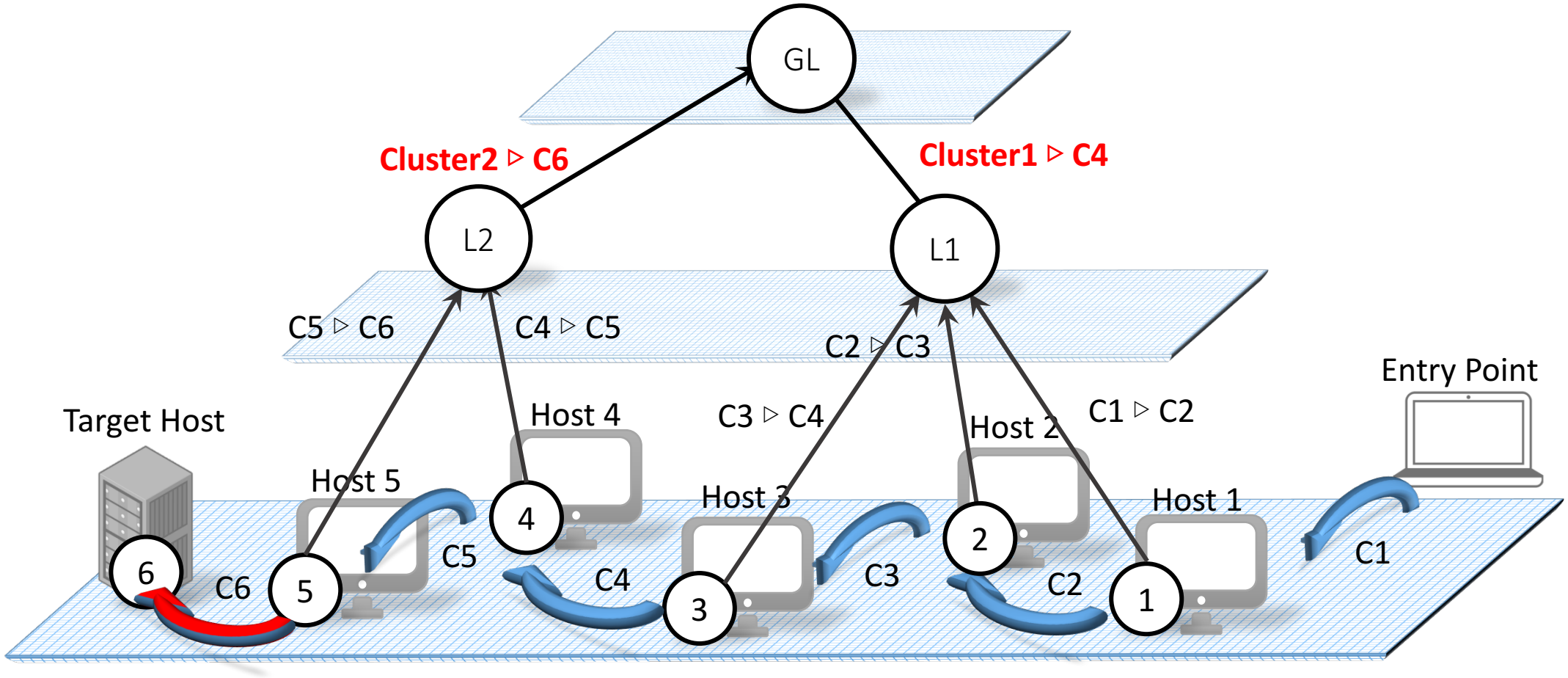
Agents do not need to synchronize clocks

C1 ▹ C2 ▹ C3 ▹ C4

⇒ t(C1)<t(C2)<t(C3)<t(C4)

Lateral Movement

A critical step during APT to move from the entry point to target host



Discussion

- Network level causation inference using host-level calls
- Detection load distributed over all agents via distributed fusion
- Eliminate the need for global clocks by abstracting data using hierarchy

Conclusion

- We designed an “end-to-end solution” that provides cyber resiliency against coordinated threats
  - Kobra generate views of a host and to learn models of applications
  - In a hierarchical manner, we used Kobra’s views to generate a network-wide chain of a coordinated attack

Future Work

- We will formulate a theory for resilient integrity checking when an attacker is attempting evasion
  - PowerAlert
  - Integrity checking of an SDN
  - Rekeying of smart meters
- We plan to develop a response mechanism for lateral movement using adaptive control
  - The attacker model is unknown, to be learned
  - Response actions change network topology and healing rates of machines

Bibliography

- [PRDC'17] A. M. Fawaz and W. H. Sanders, "Learning Process Behavioral Baselines for Anomaly Detection" Proceedings of the 22nd IEEE Pacific Rim International Symposium on Dependable Computing (PRDC 2017), Christchurch, New Zealand, January 22-25, 2017, to appear.
- [SRDS'16] A. Fawaz, A. Bohara, C. Cheh, and W. H. Sanders, "Lateral Movement Detection Using Distributed Data Fusion", Proceedings of the 35th Symposium on Reliable Distributed Systems (SRDS), Budapest, Hungary, Sept. 26-29, 2016, to appear.
- [RAID'16] Fawaz, A., and Sanders, W. H, "Poster: Learning Process Behavioral Baselines for Anomaly Detection", RAID 2016[Poster]
- [GameSec'16] M. A. Noureddine, A. Fawaz, W. H. Sanders, and T. Basar, "A Game-Theoretic Approach to Respond to Attacker Lateral Movement" Proceedings of the 7th Conference on Decision and Game Theory for Security (GameSec 2016), New York, New York, November 2-4, 2016, *Lecture Notes in Computer Science* vol. 9996, Springer, 2016, pp. 294-313.
- [TSG'16] Fawaz, A., Berthier, R., and Sanders, W. H., "A Response Cost Model for Advanced Metering Infrastructures", IEEE Transactions on Smart Grid, vol. 7, no. 2, March 2016, pp. 543-553.
- [JSAC'13] Stephen McLaughlin, Brett Holbert, Ahmed Fawaz, Robin Berthier and Saman Zonouz, "A Multi-Sensor Intrusion and Energy Theft Detection Framework for Advanced Metering Infrastructures", IEEE JSAC Smart Grid Communications Series, vol. 31, no. 7, pp. 1319-1330, July 2013.
- [SmartGridComm'12] Fawaz, A., Berthier, R., and Sanders, W. H., "Cost Modeling of Response Actions for Automated Response and Recovery in AMI," In Proceedings of the Third IEEE International Conference on Smart Grid Communication (SmartGridComm 2012), Tainan City, Taiwan, Nov. 5-8, 2012, pp. 348-353.
- [NISTCPS'12] Fawaz, A., Berthier, R., Sanders, W. H., and Pal., P., "Understanding the Role of Automated Response Actions in Improving AMI Resiliency," In Proceedings of the NIST Cybersecurity for Cyber-Physical Systems Workshop, Gaithersburg, Maryland, Apr. 23-24, 2012.