Optimal Security Response in IT Systems

Defense for the Degree of Doctor of Philosophy

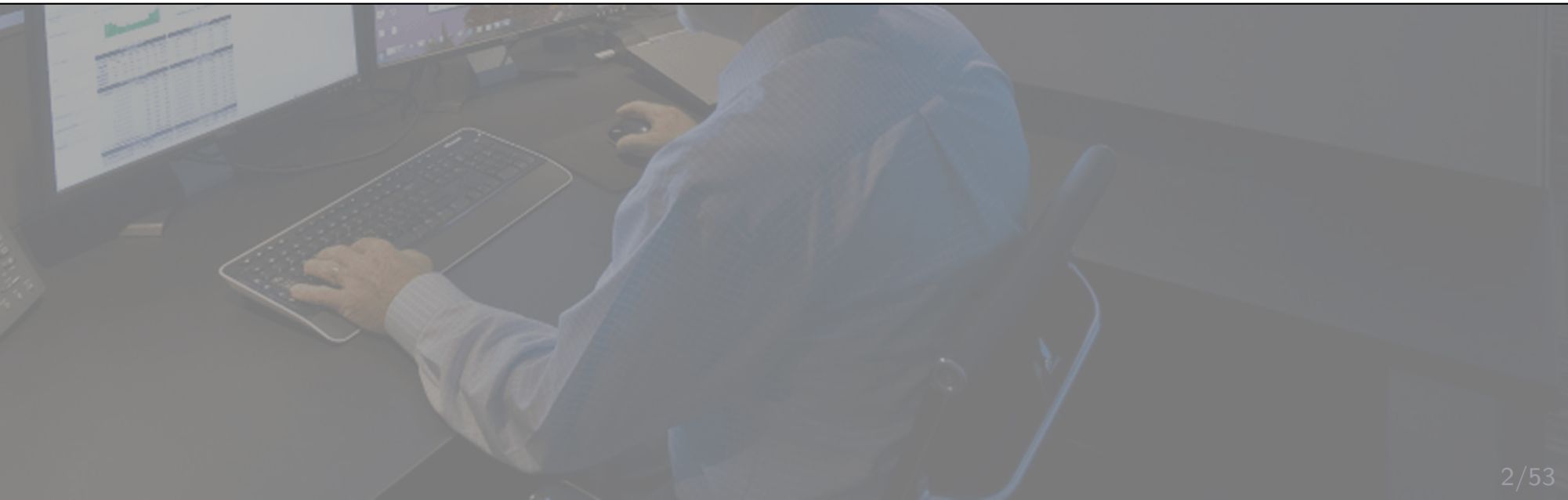
Candidate: Kim Hammar

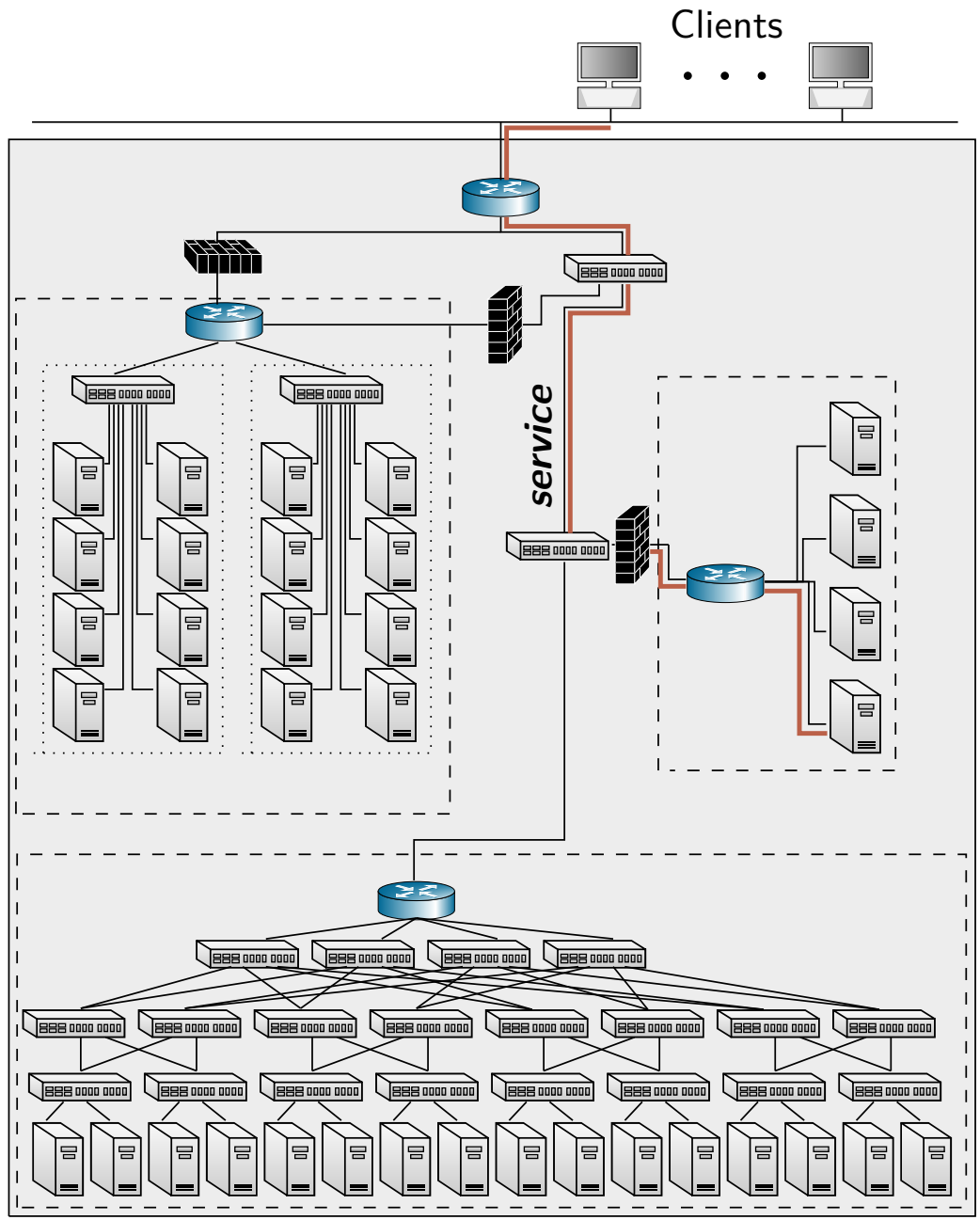
KTH Royal Institute of Technology, F3, Lindstedtsvägen 26.

Supervisor: Prof. Rolf Stadler. Opponent: Prof. Tansu Alpcan.
Chair: Prof. Viktoria Fodor. Reviewer: Prof. Henrik Sandberg.
Committee: Prof. Karl H. Johansson, Prof. Alina Oprea, Prof. Emil Lupu.

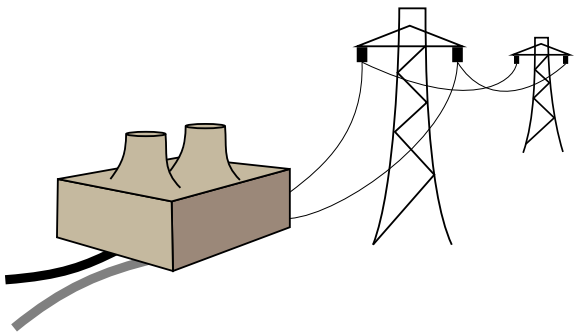
Dec 5, 2024

How to **automate** security response operations in an optimal way?



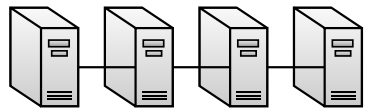

Clients
service

IT systems facilitate



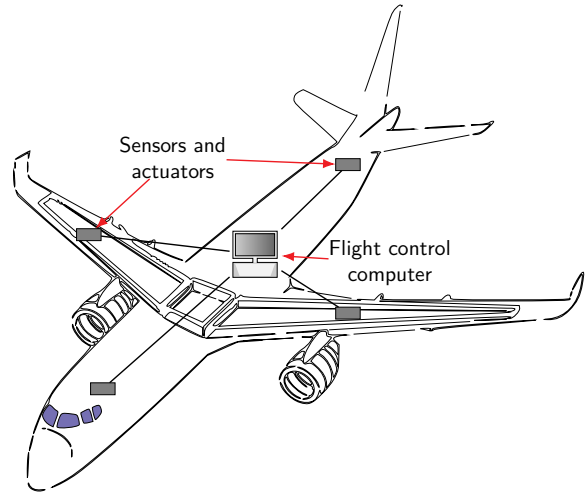
Critical infrastructures

e.g., power and transport infrastructures.



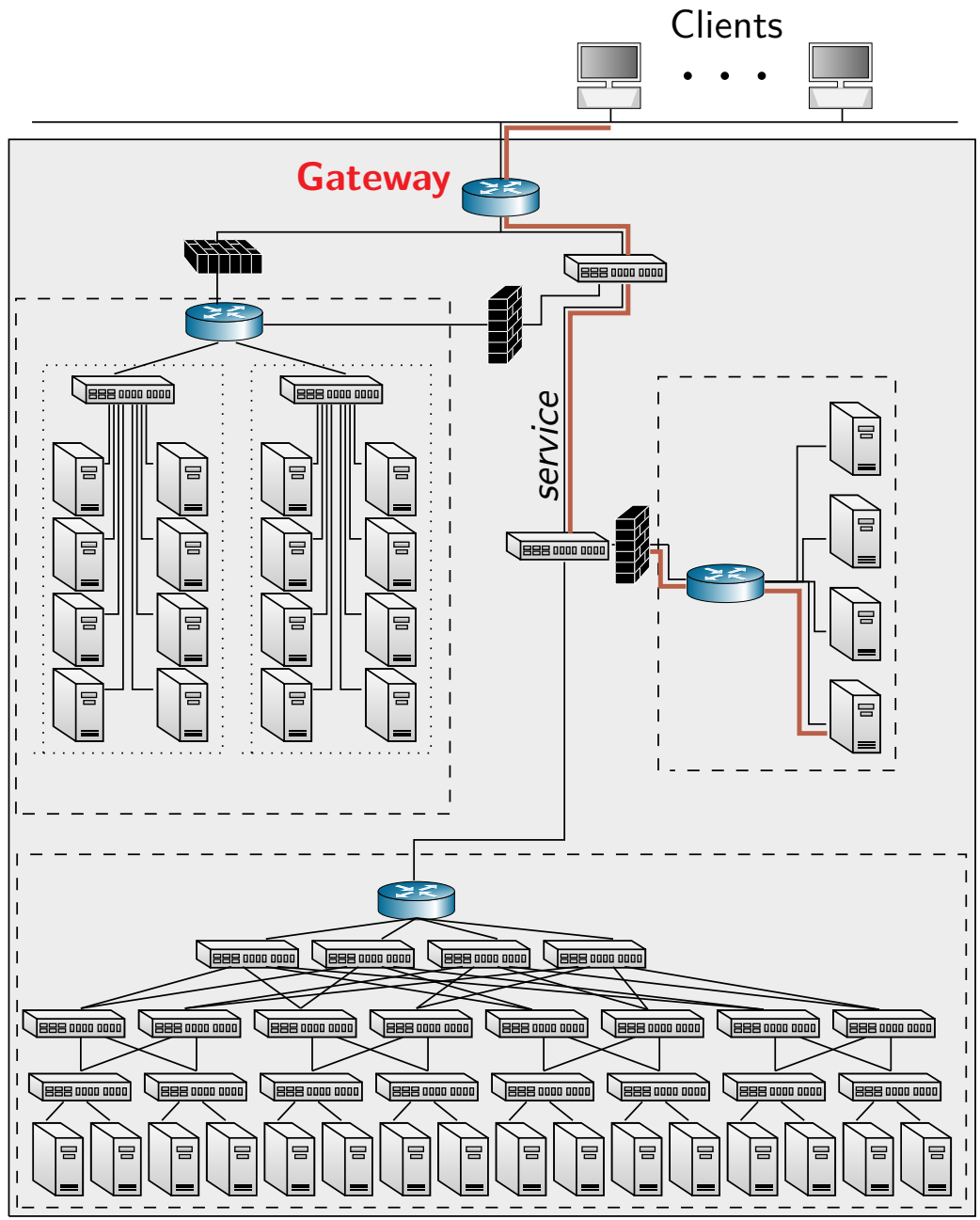
Financial ecosystems

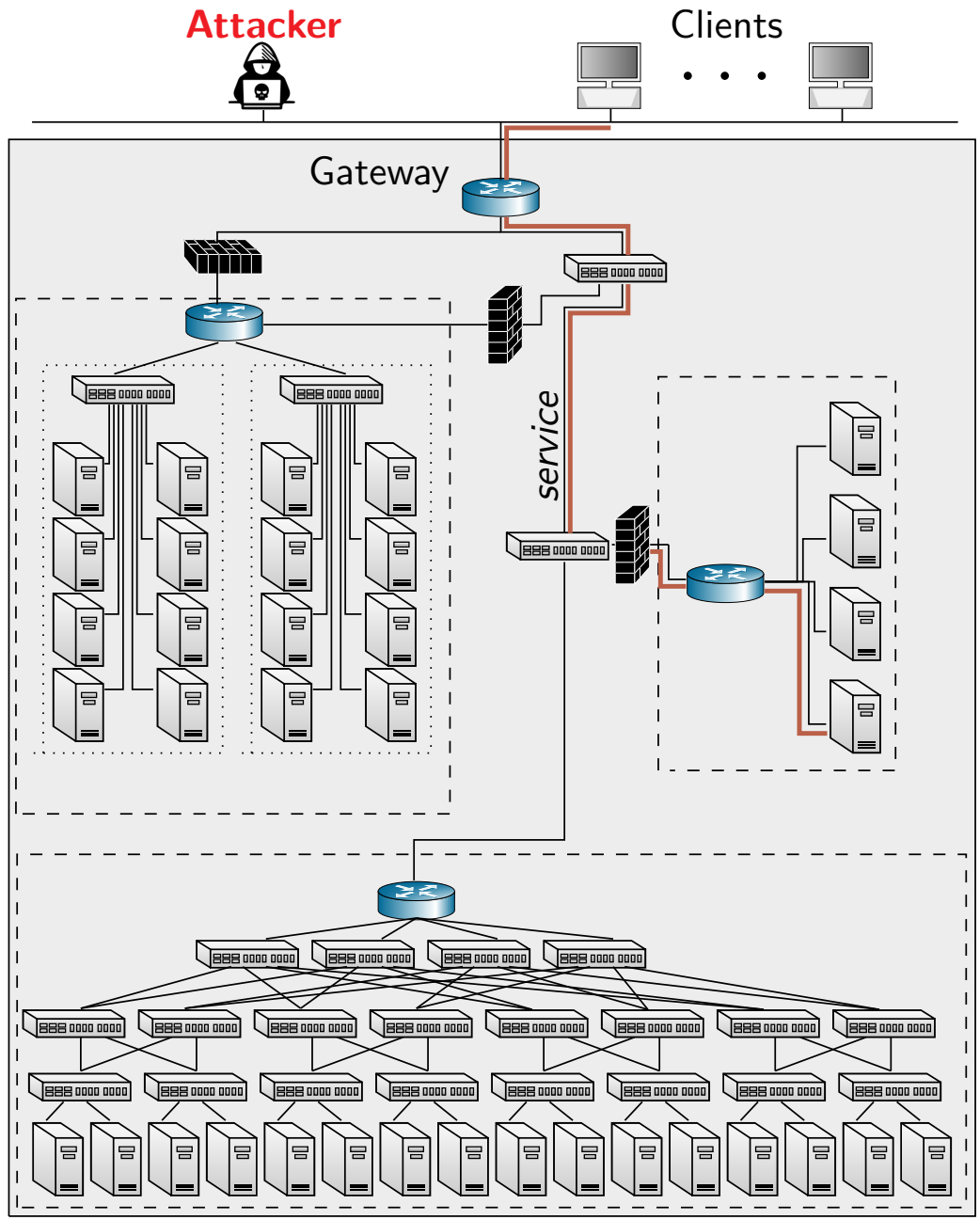
e.g., banking systems, payment processing systems, Swish, etc.

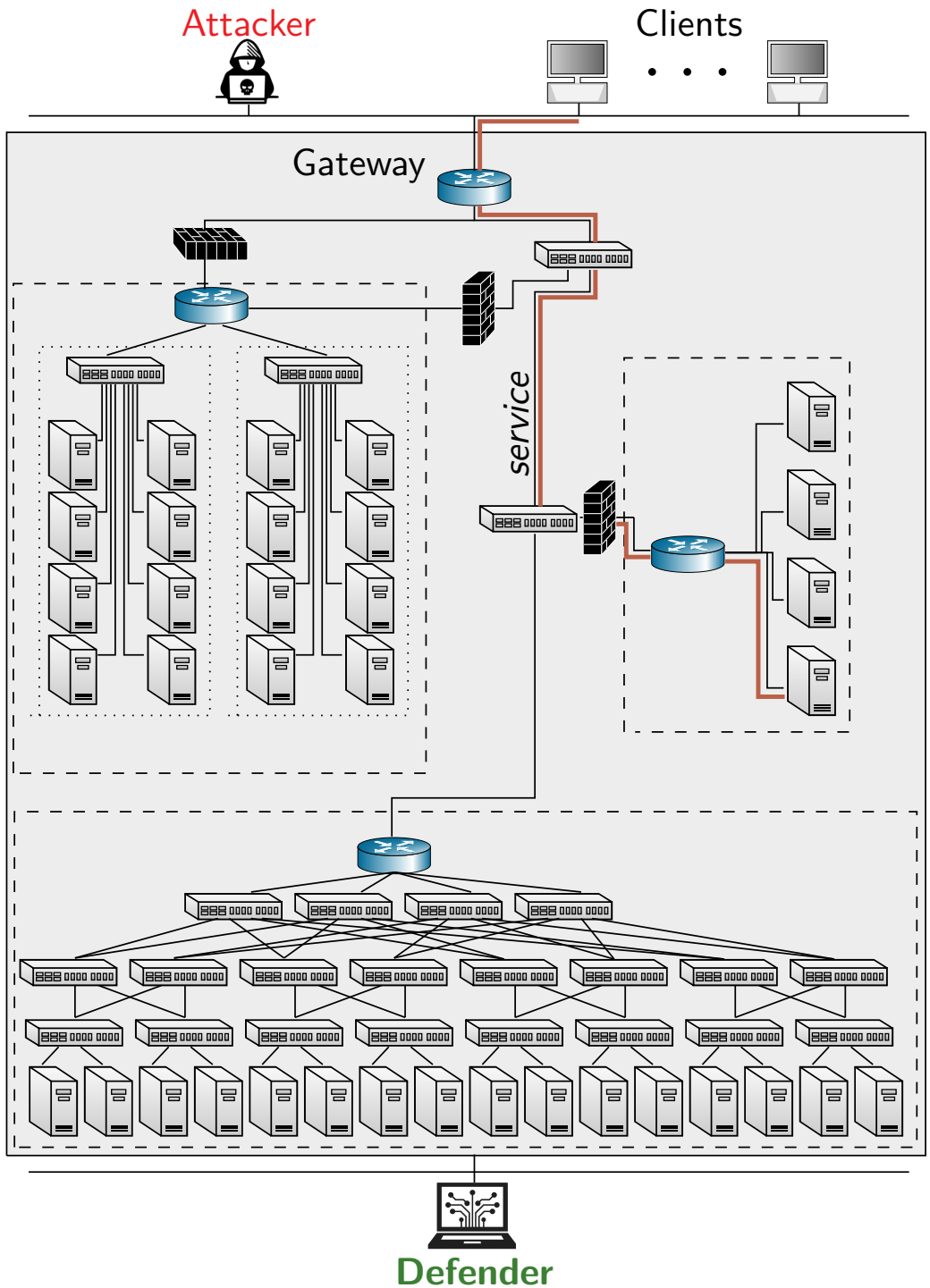


Cyber-physical systems

e.g., flight control systems, train signaling systems, healthcare systems, etc.


Clients
Gateway
service

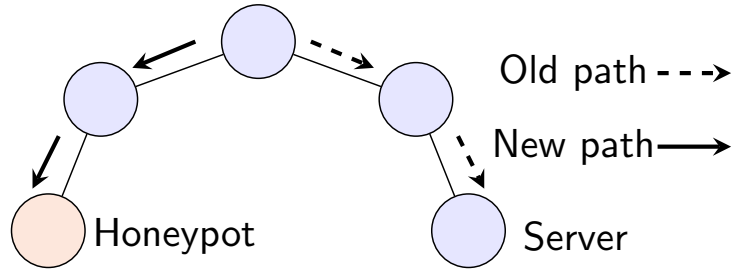

Attacker
Clients
Gateway
service


Attacker
Clients
Gateway
service
Defender

Examples of Response Actions

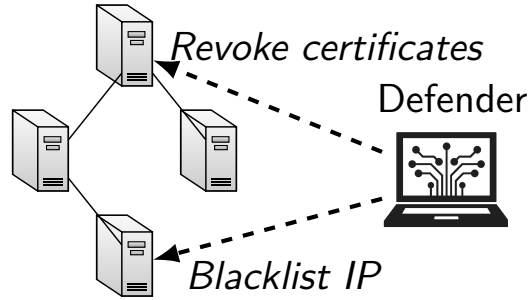
Flow control

By redirecting traffic, the defender can isolate malicious behavior.



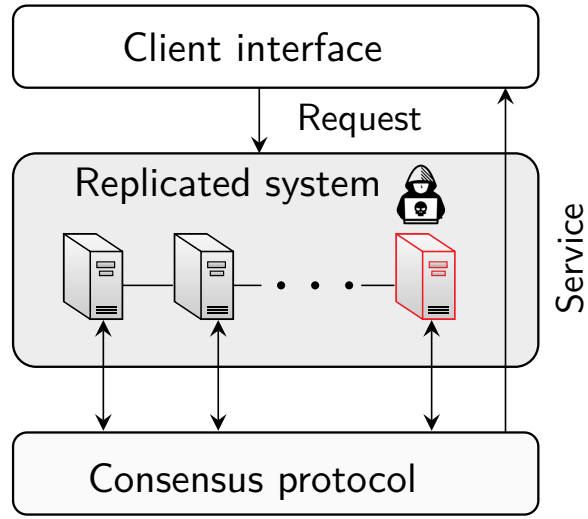
Access control

By adjusting resource permissions, the defender can prevent the attacker from compromising critical assets.



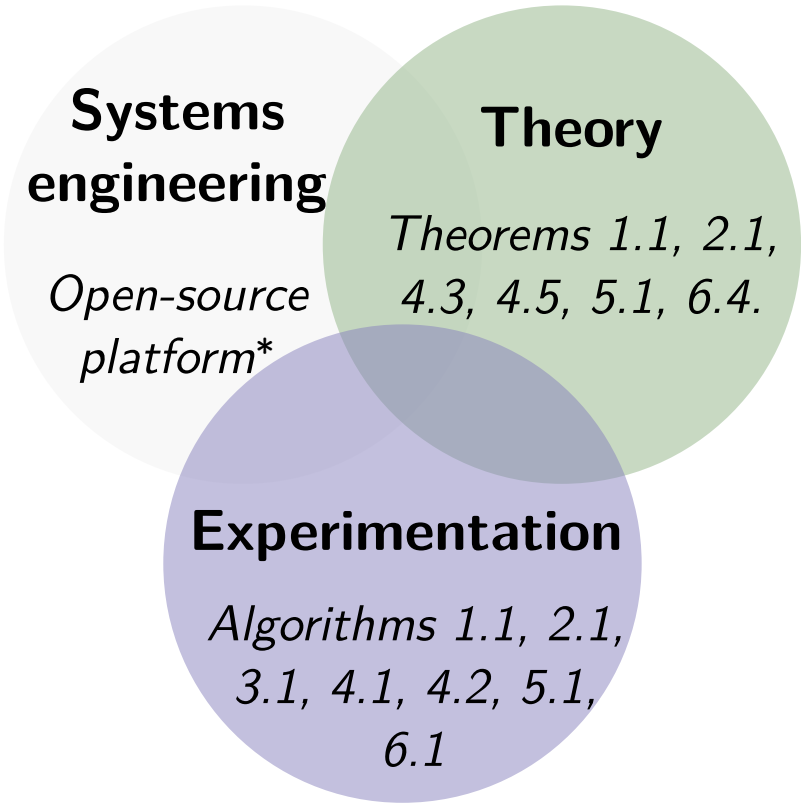
Replication control

Replication can ensure that multiple replicas of services remain available even when some are compromised.



Thesis Contributions - Optimal Security Response

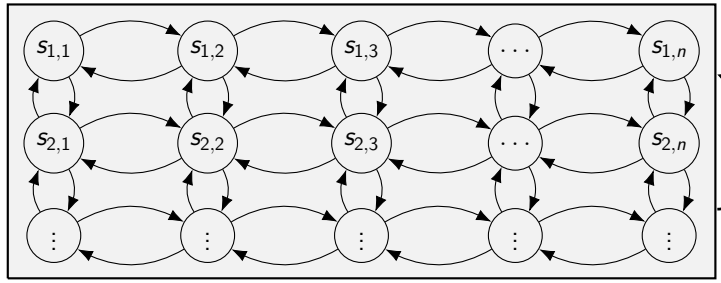
- My thesis advances optimal security response through *theoretical foundations*, *system design*, and *experimental validation.*



*Kim Hammar. *Cyber Security Learning Environment (CSLE)*. Documentation: https://limmen.dev/csle/, traces: https://github.com/Limmen/csle/releases/tag/v0.4.0, source code: https://github.com/Limmen/csle, video demonstration: https://www.youtube.com/watch?v=iE2KPmtIs2A&. 2023. URL: https://limmen.dev/csle/.

Methodology for Optimal Security Response

Simulation System

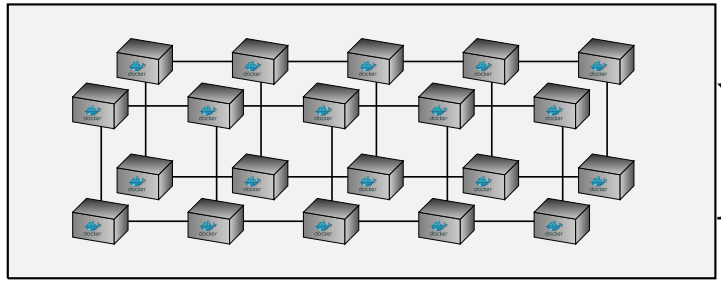


Mathematical Model & Optimization

Strategy Mapping π

System Identification

Emulation System

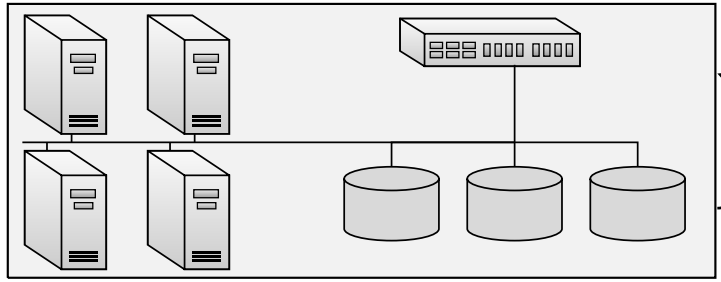


Strategy Evaluation & Model Estimation

Strategy Implementation π

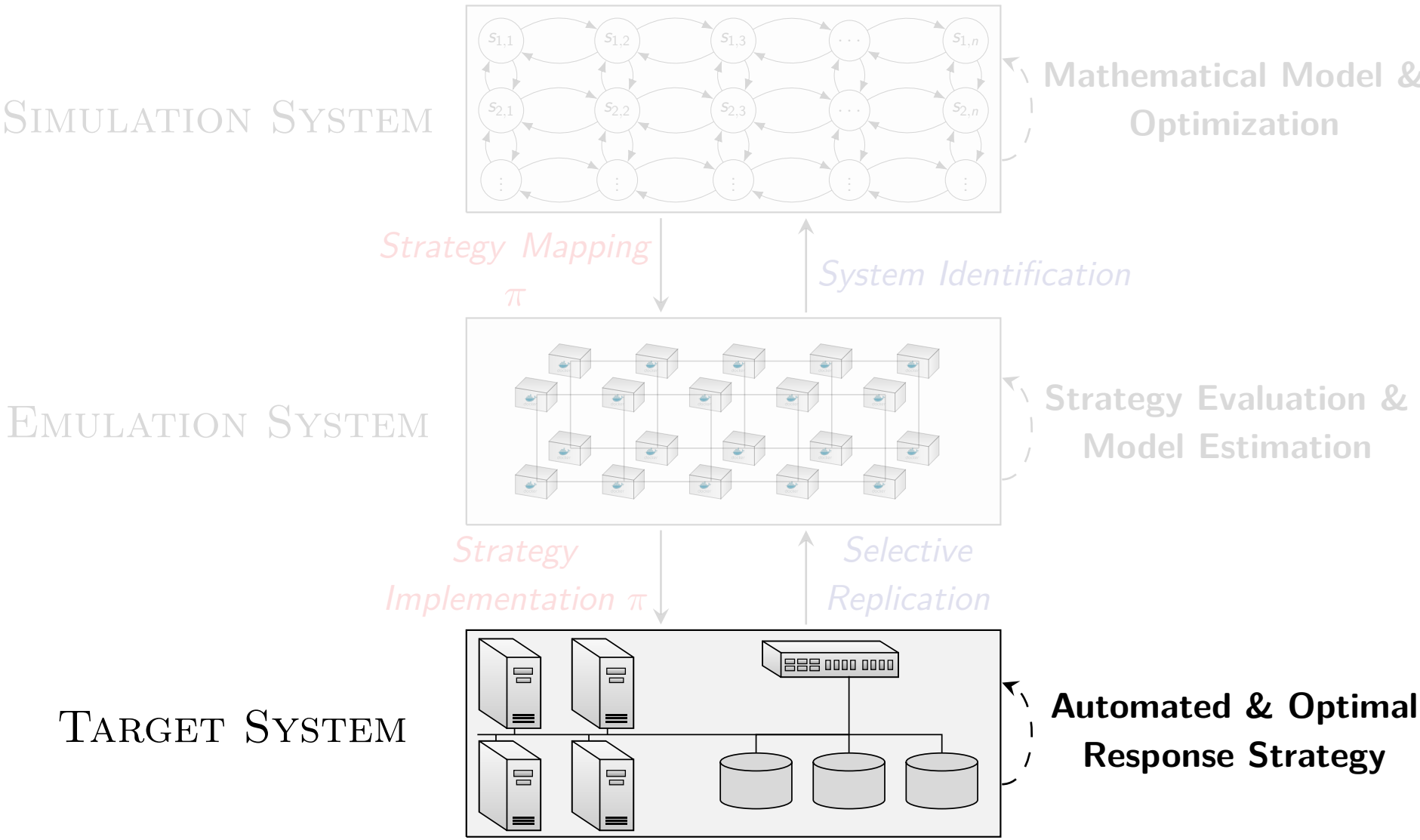
Selective Replication

Target System

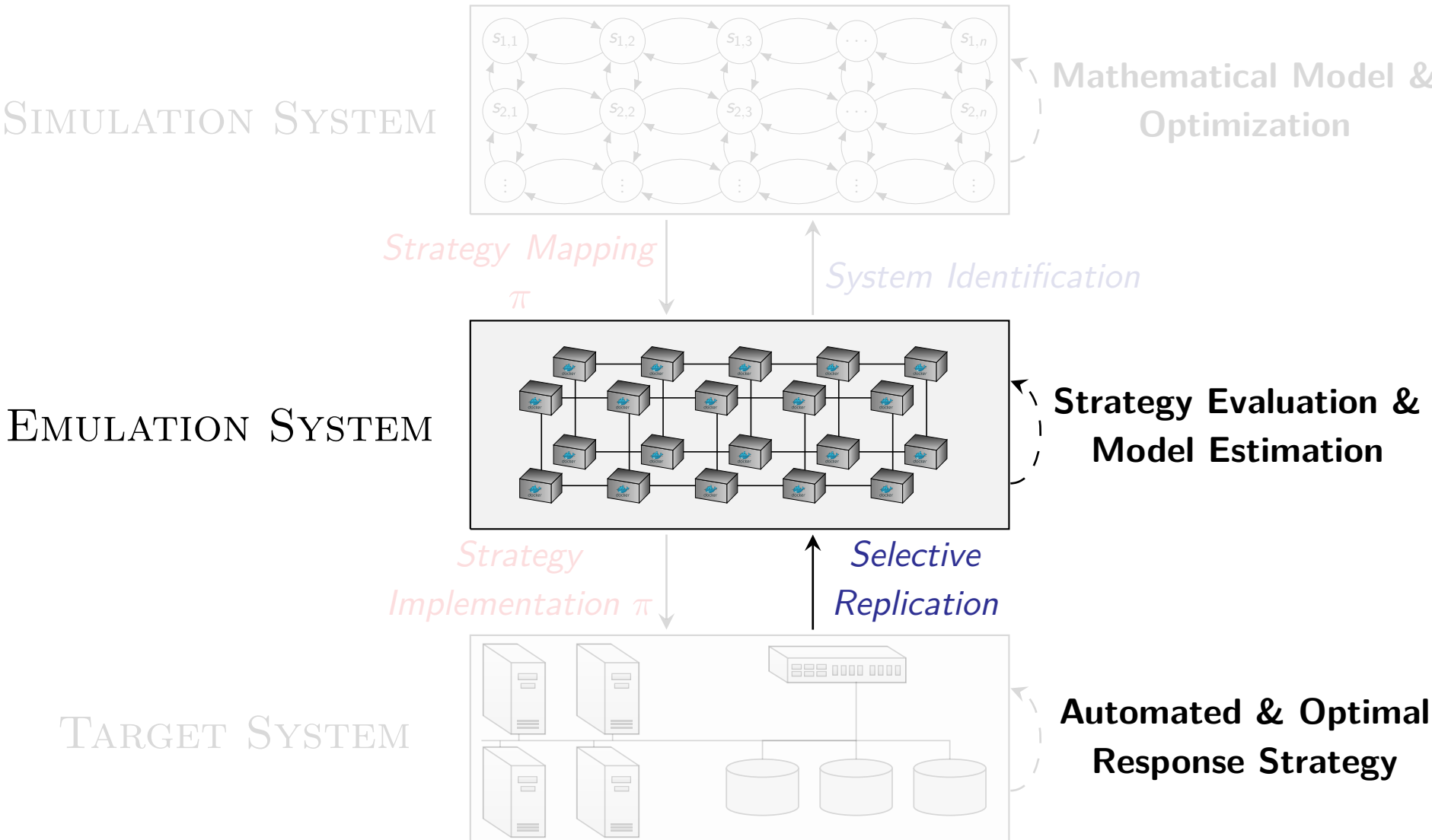


Automated & Optimal Response Strategy

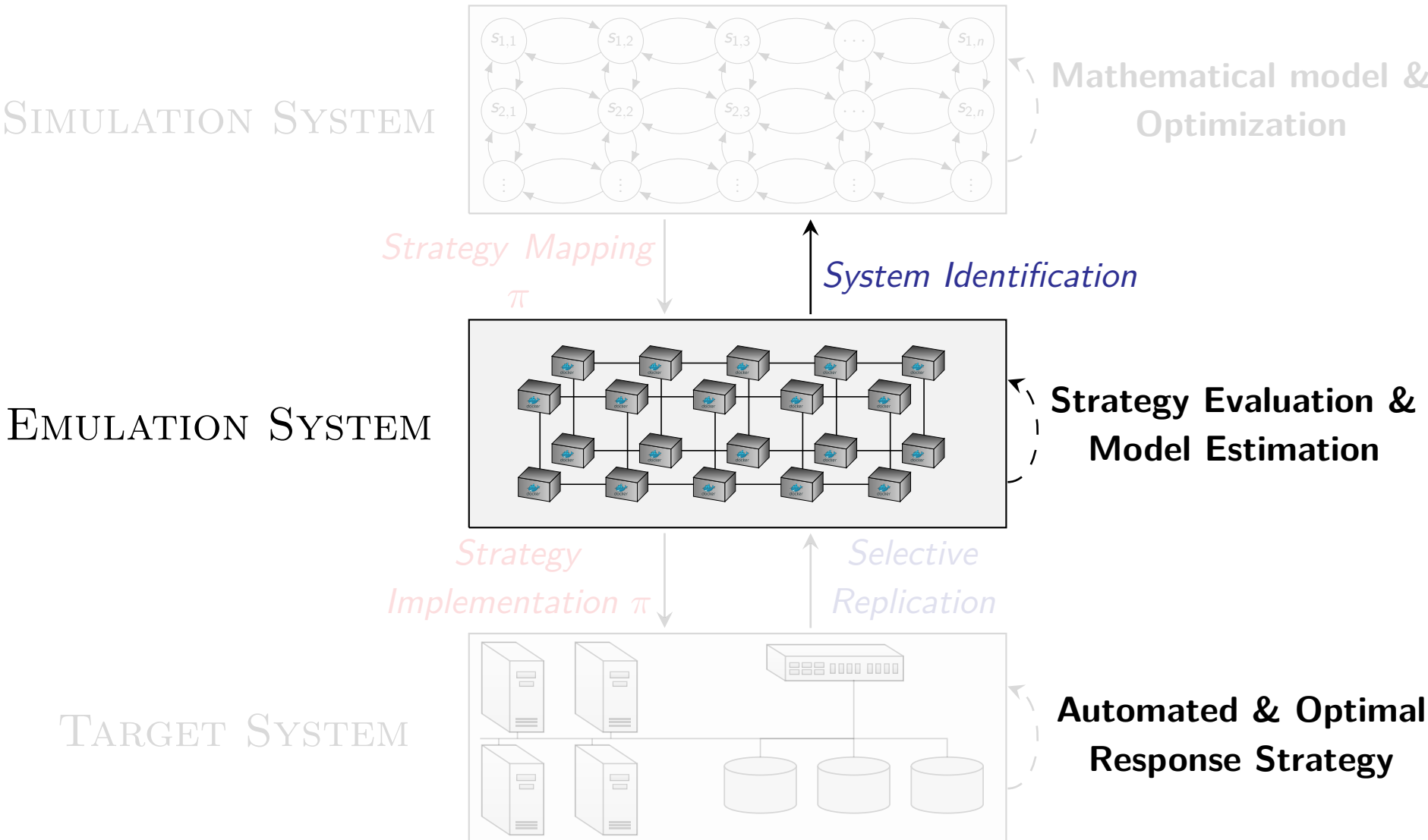
Methodology for Optimal Security Response



Methodology for Optimal Security Response

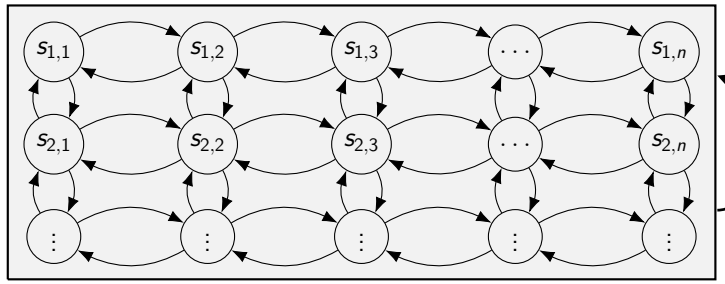


Methodology for Optimal Security Response



Methodology for Optimal Security Response

Simulation System

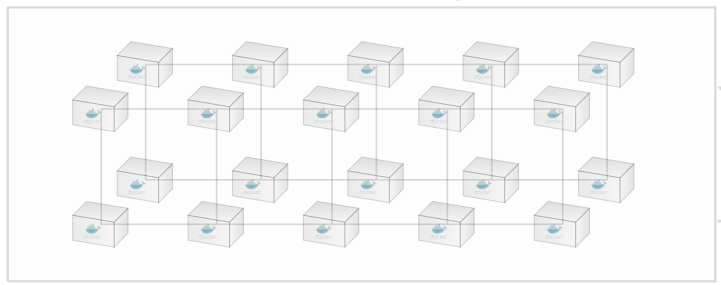


Mathematical Model & Optimization

Strategy Mapping π

System Identification

Emulation System

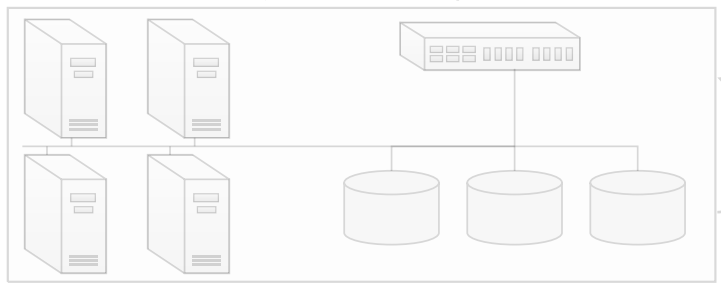


Strategy Evaluation & Model Estimation

Strategy Implementation π

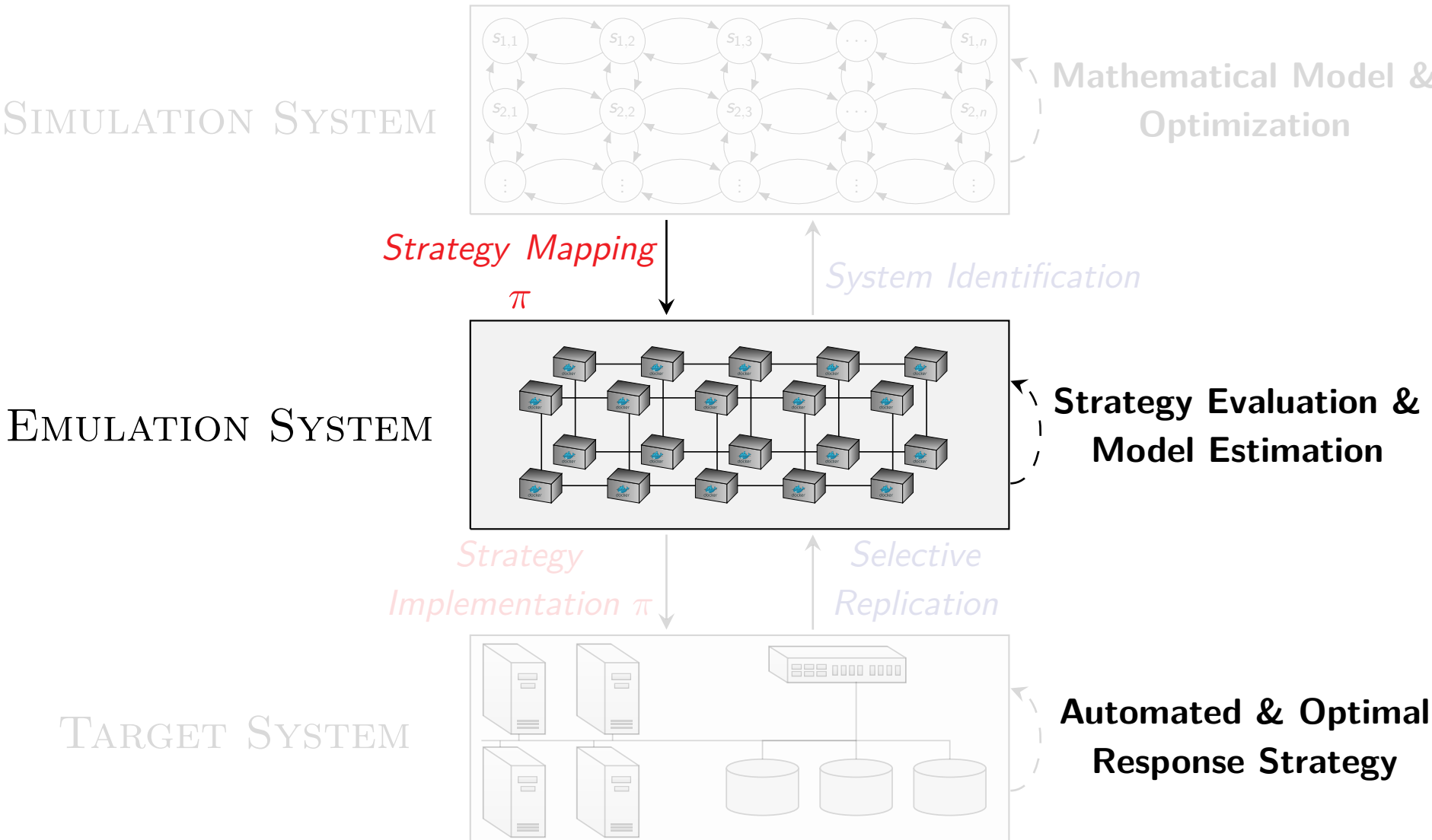
Selective Replication

Target System

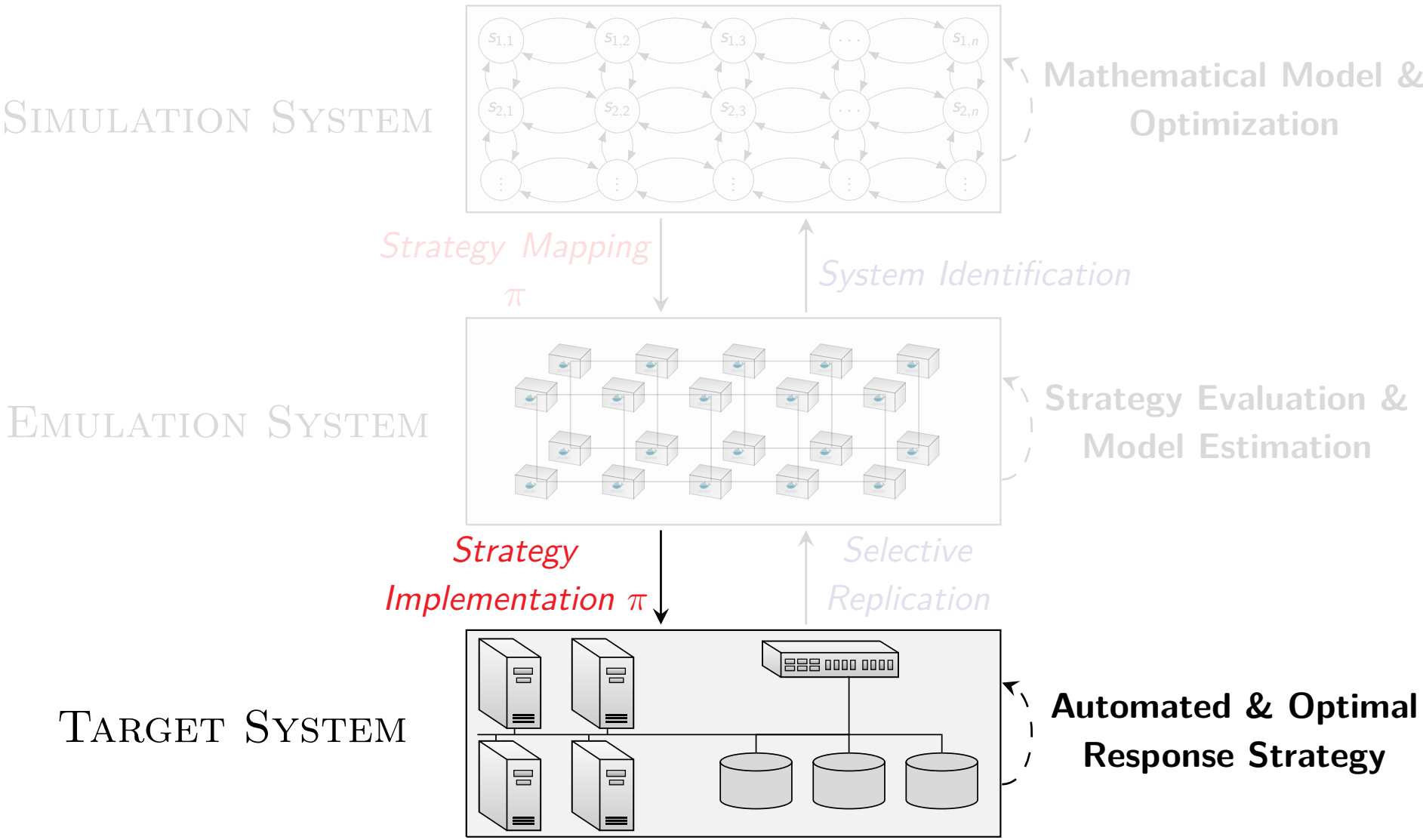


Automated & Optimal Response Strategy

Methodology for Optimal Security Response

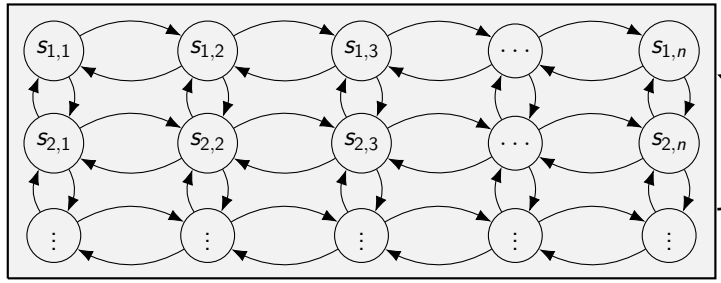


Methodology for Optimal Security Response



Methodology for Optimal Security Response

Simulation System

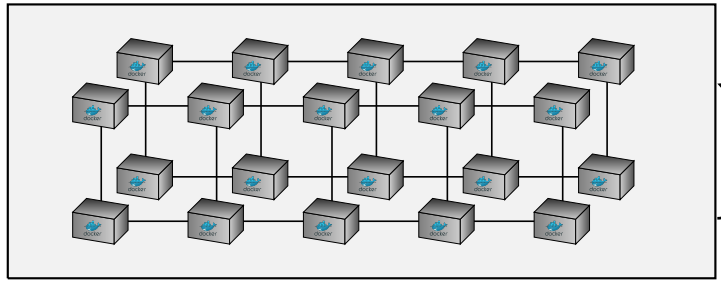


Mathematical Model & Optimization

Strategy Mapping π

System Identification

Emulation System

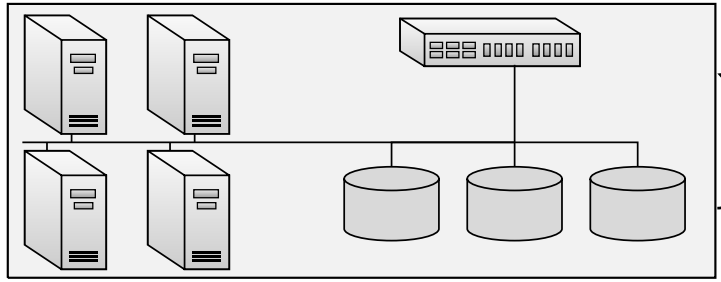


Strategy Evaluation & Model Estimation

Strategy Implementation π

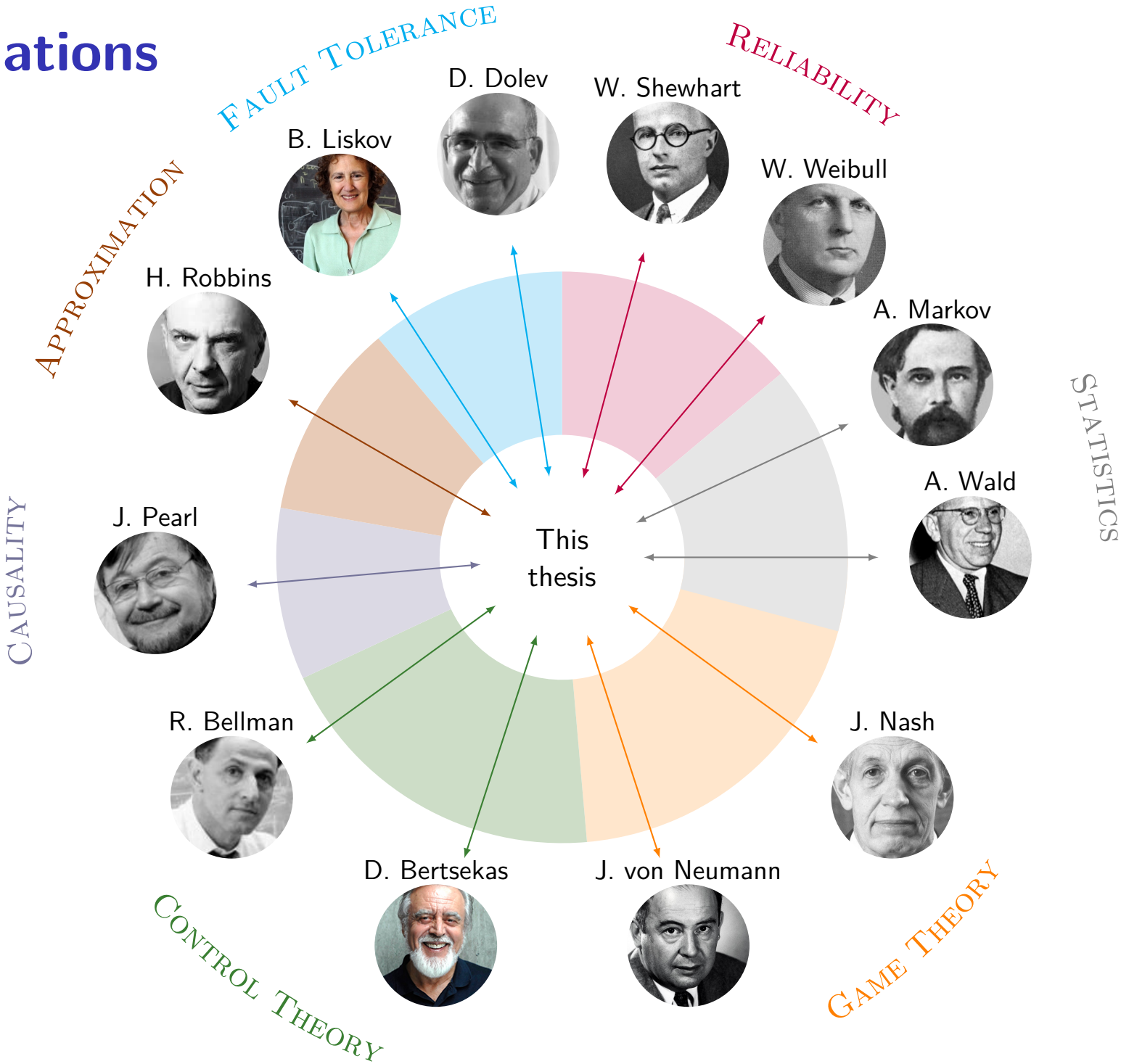
Selective Replication

Target System

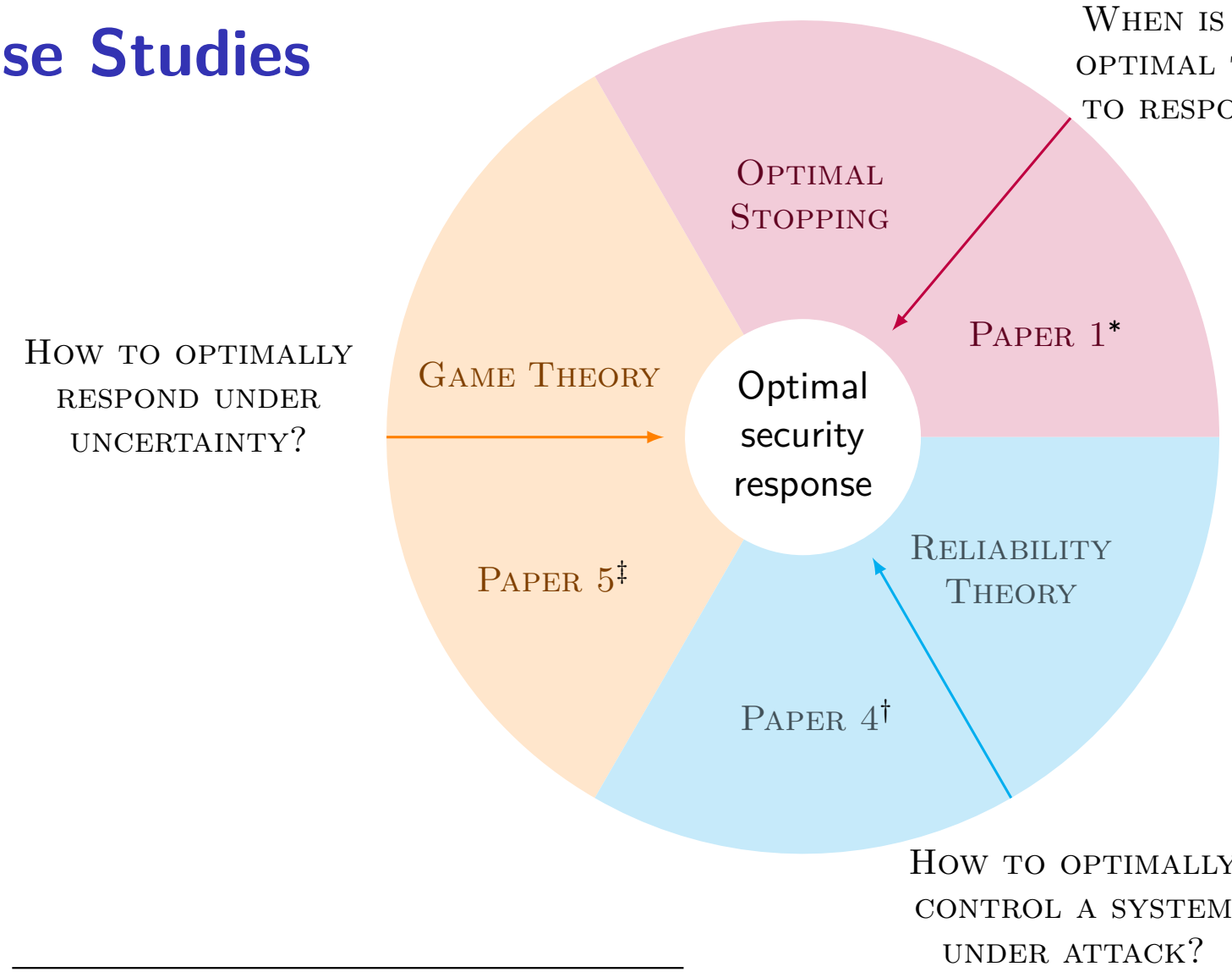


Automated & Optimal Response Strategy

Foundations



Case Studies

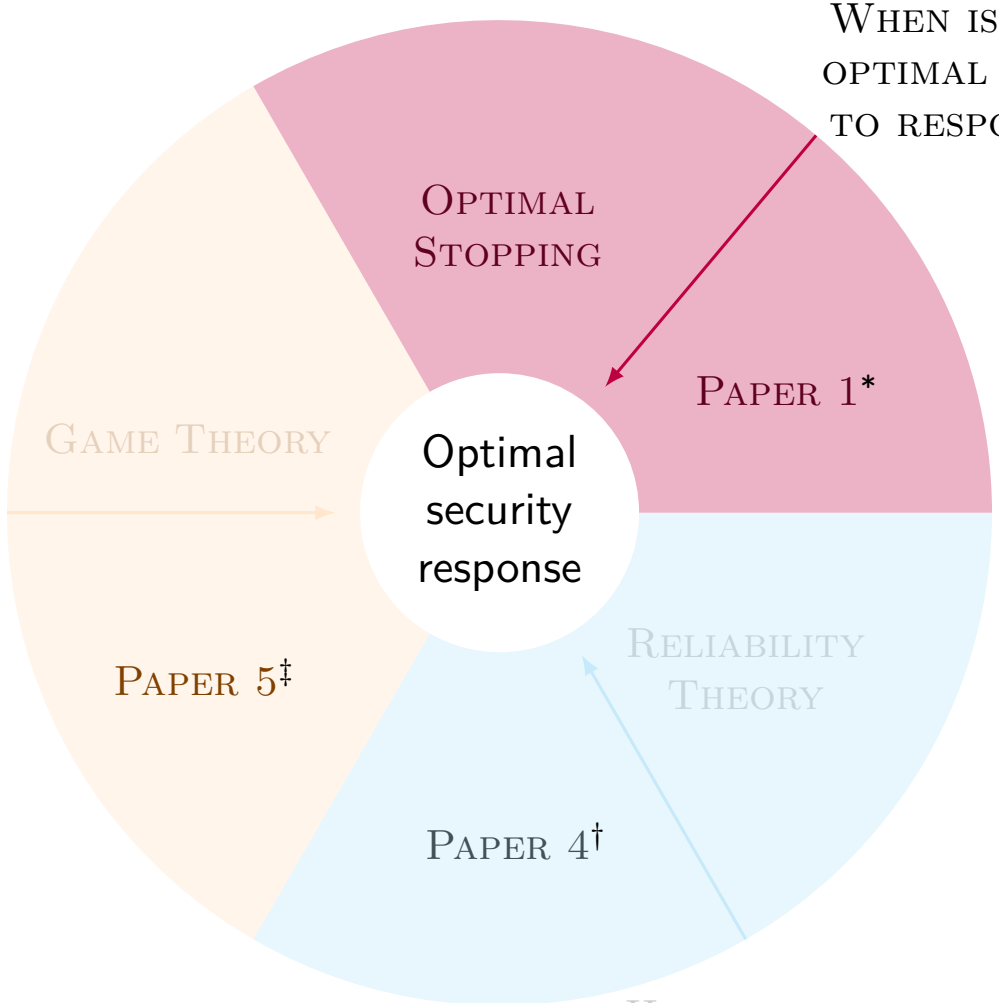


*Kim Hammar and Rolf Stadler. "Intrusion Prevention through Optimal Stopping". In: *IEEE Transactions on Network and Service Management* 19.3 (2022), pp. 2333–2348. DOI: 10.1109/TNSM.2022.3176781.

†Kim Hammar and Rolf Stadler. "Intrusion Tolerance for Networked Systems through Two-Level Feedback Control". In: *2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. 2024, pp. 338–352. DOI: 10.1109/DSN58291.2024.00042.

‡Kim Hammar et al. *Automated Security Response through Online Learning with Adaptive Conjectures*. https://arxiv.org/abs/2402.12499. To appear in IEEE Transactions on Information Forensics and Security

Case Studies

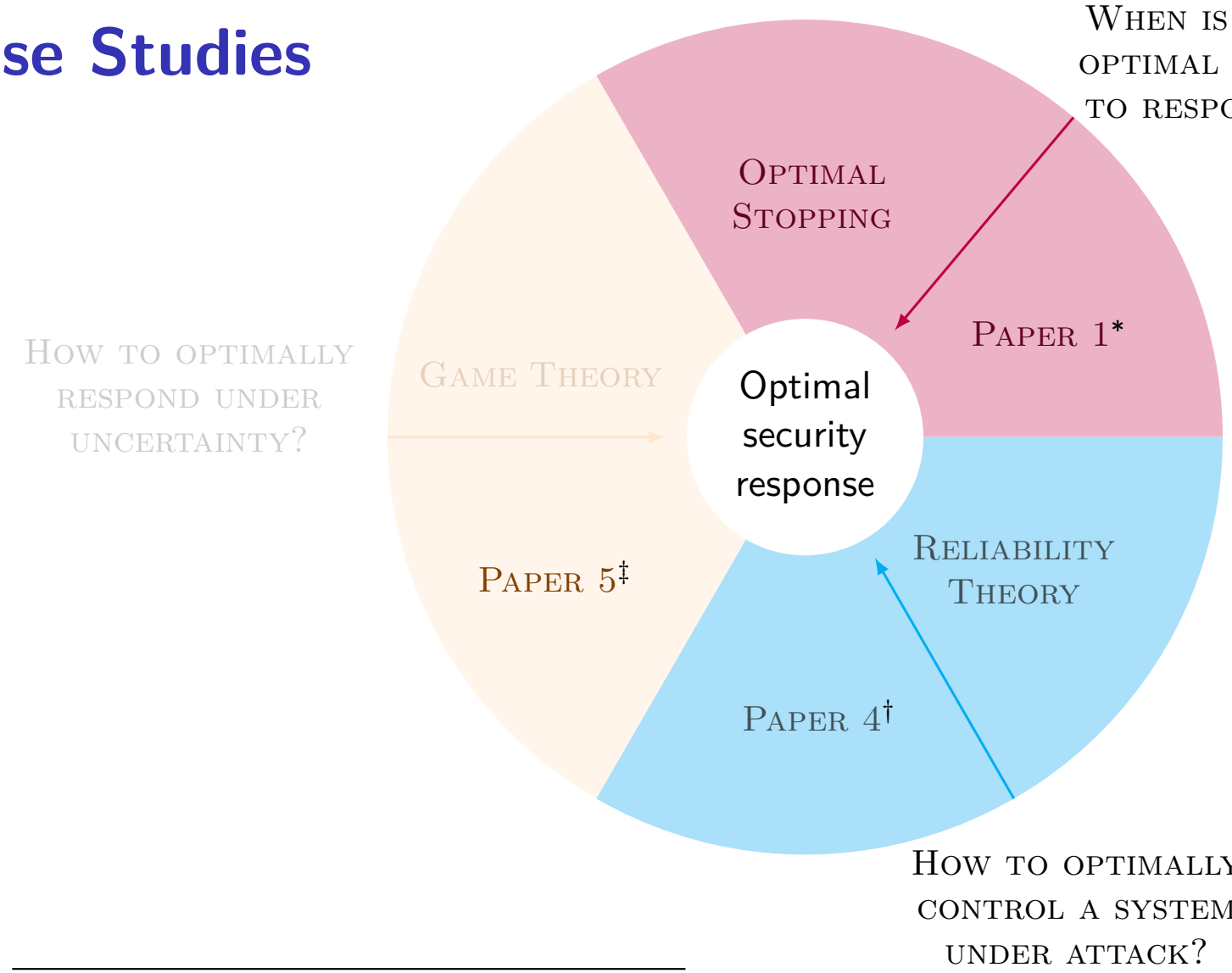


*Kim Hammar and Rolf Stadler. "Intrusion Prevention through Optimal Stopping". In: *IEEE Transactions on Network and Service Management* 19.3 (2022), pp. 2333–2348. DOI: 10.1109/TNSM.2022.3176781.

†Kim Hammar and Rolf Stadler. "Intrusion Tolerance for Networked Systems through Two-Level Feedback Control". In: *2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. 2024, pp. 338–352. DOI: 10.1109/DSN58291.2024.00042.

‡Kim Hammar et al. *Automated Security Response through Online Learning with Adaptive Conjectures*. https://arxiv.org/abs/2402.12499. To appear in IEEE Transactions on Information Forensics and Security

Case Studies

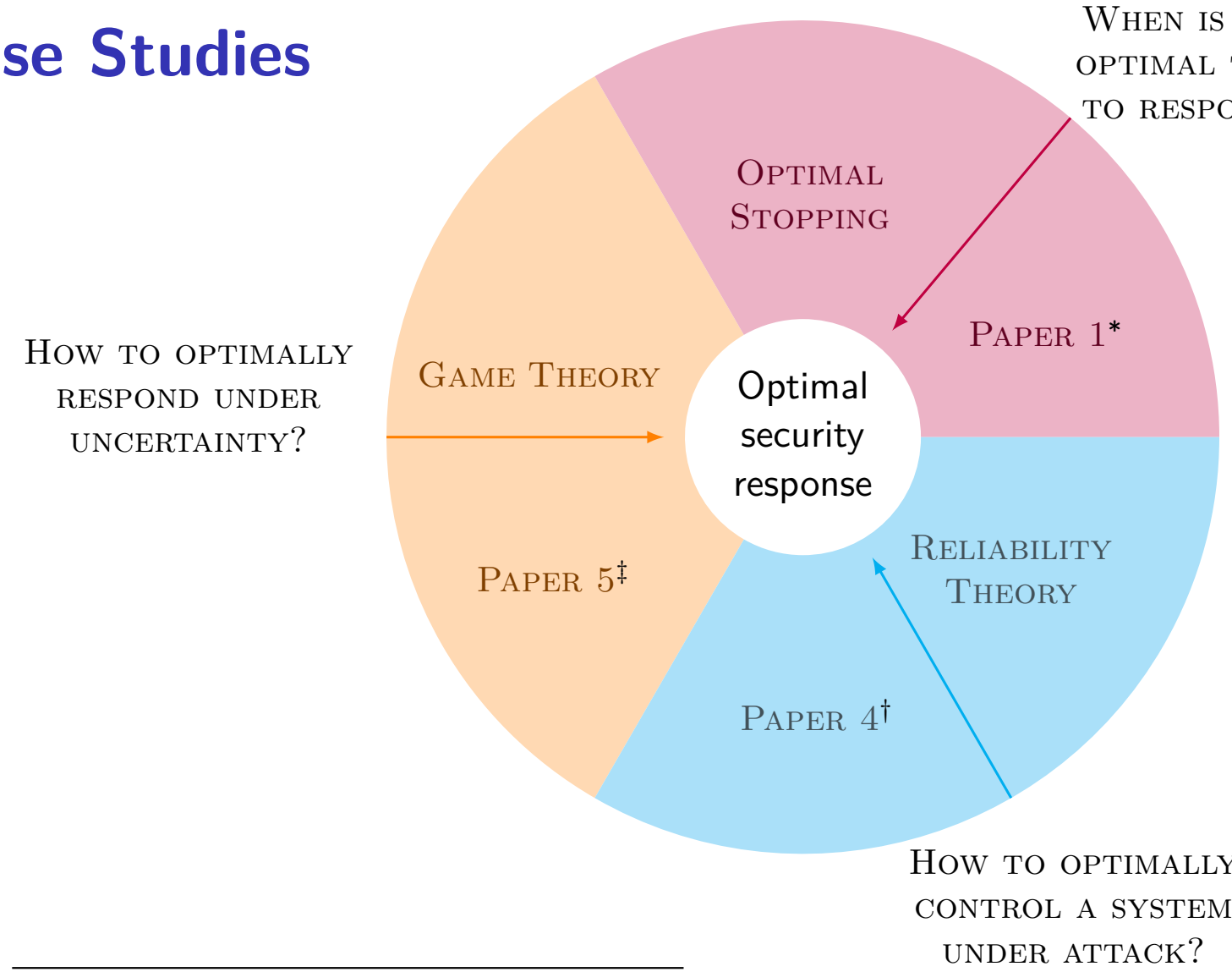


*Kim Hammar and Rolf Stadler. "Intrusion Prevention through Optimal Stopping". In: *IEEE Transactions on Network and Service Management* 19.3 (2022), pp. 2333–2348. DOI: 10.1109/TNSM.2022.3176781.

†Kim Hammar and Rolf Stadler. "Intrusion Tolerance for Networked Systems through Two-Level Feedback Control". In: *2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. 2024, pp. 338–352. DOI: 10.1109/DSN58291.2024.00042.

‡Kim Hammar et al. *Automated Security Response through Online Learning with Adaptive Conjectures*. https://arxiv.org/abs/2402.12499. To appear in IEEE Transactions on Information Forensics and Security

Case Studies



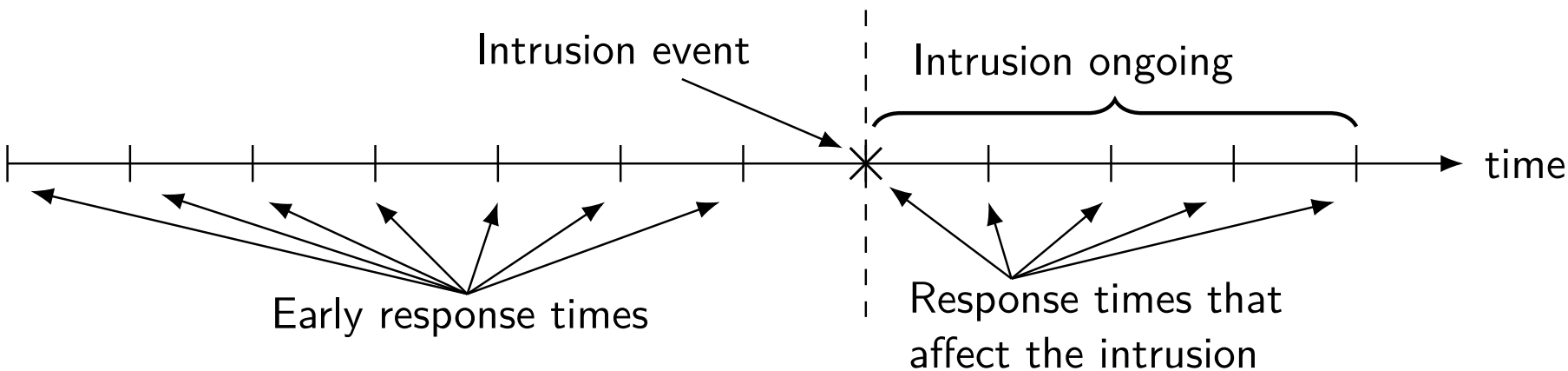
*Kim Hammar and Rolf Stadler. "Intrusion Prevention through Optimal Stopping". In: *IEEE Transactions on Network and Service Management* 19.3 (2022), pp. 2333–2348. DOI: 10.1109/TNSM.2022.3176781.

†Kim Hammar and Rolf Stadler. "Intrusion Tolerance for Networked Systems through Two-Level Feedback Control". In: *2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. 2024, pp. 338–352. DOI: 10.1109/DSN58291.2024.00042.

‡Kim Hammar et al. *Automated Security Response through Online Learning with Adaptive Conjectures*. https://arxiv.org/abs/2402.12499. To appear in IEEE Transactions on Information Forensics and Security

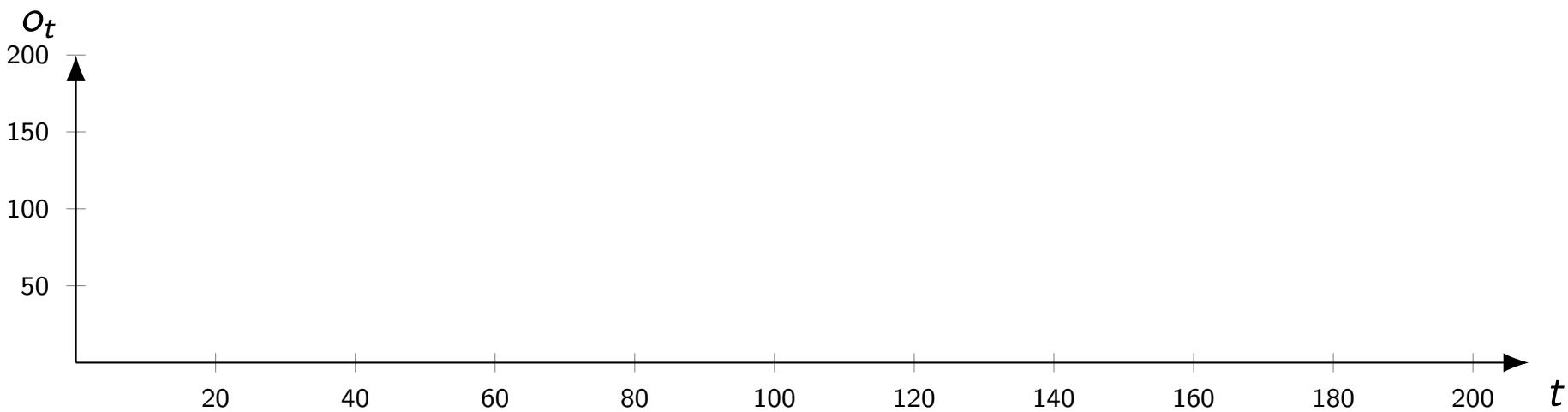
When is the Optimal time to Respond?

- The attacker seeks to intrude on the infrastructure.
- One response action, e.g., block the gateway.



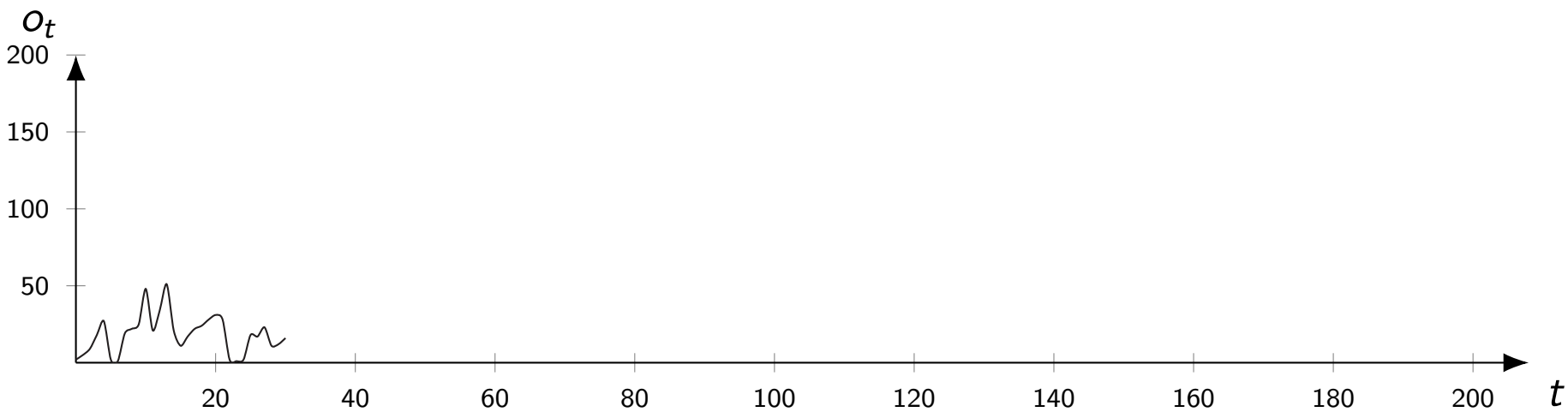
Optimal Stopping

- Observe the system through the stochastic process $(o_t)_{t=1}^{T}$.
- o_t is the number of security alerts at time t.



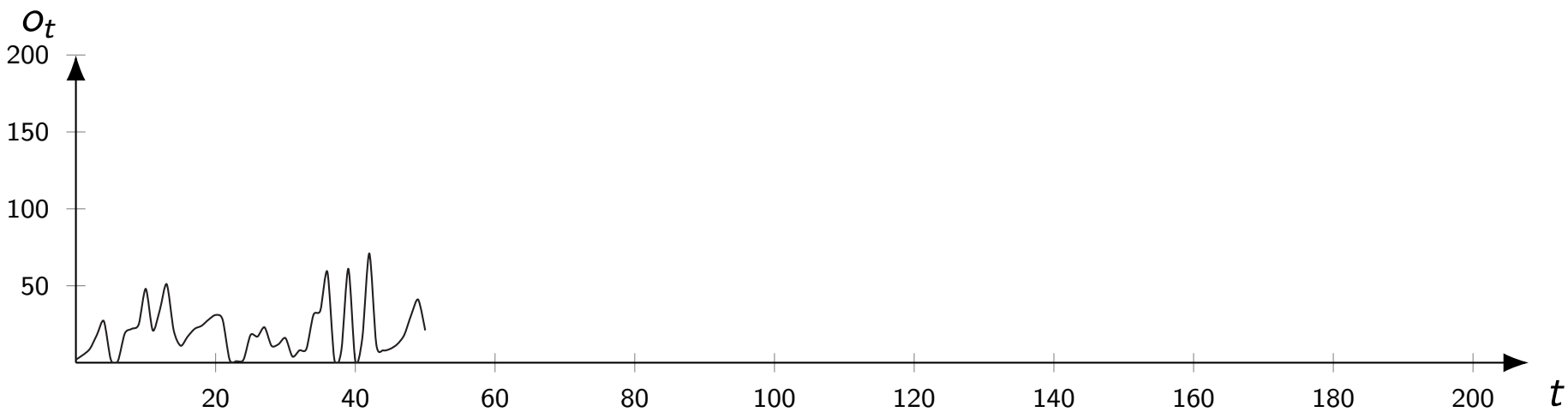
Optimal Stopping

- Observe the system through the stochastic process $(o_t)_{t=1}^{T}$.
- o_t is the number of security alerts at time t.



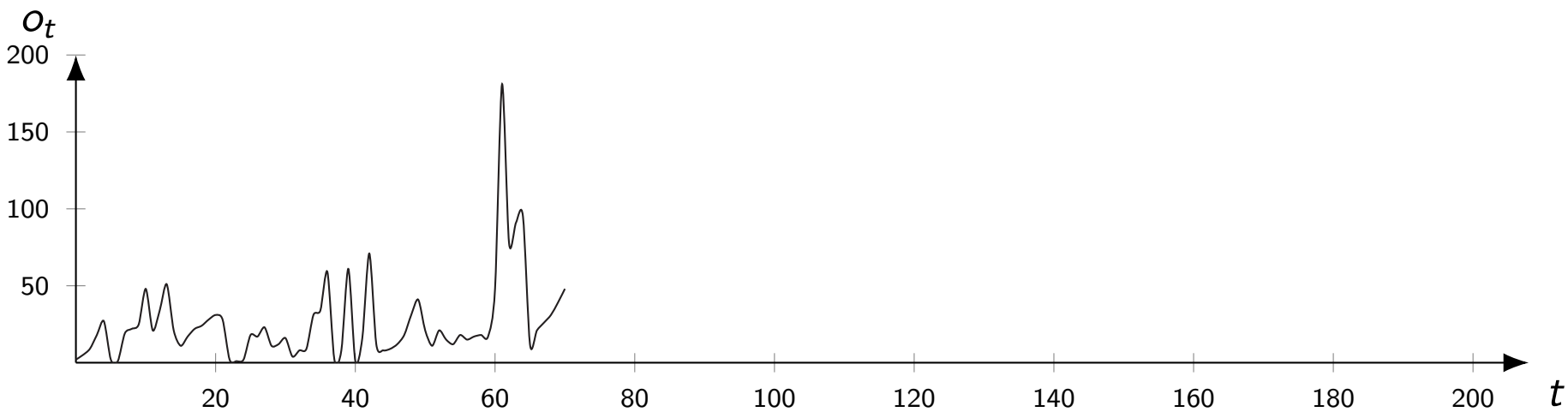
Optimal Stopping

- Observe the system through the stochastic process $(o_t)_{t=1}^{T}$.
- o_t is the number of security alerts at time t.



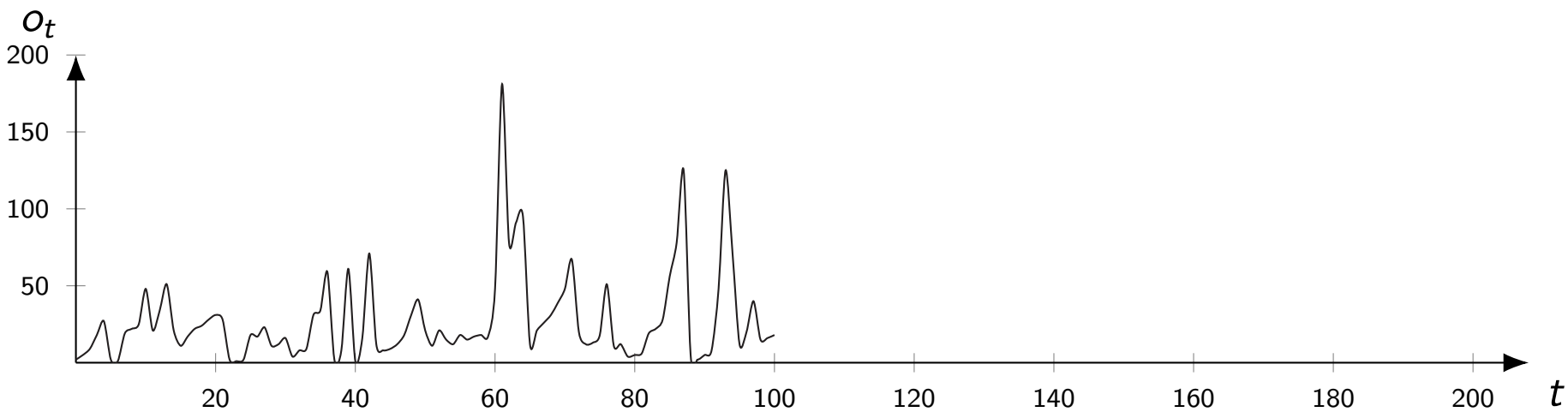
Optimal Stopping

- Observe the system through the stochastic process $(o_t)_{t=1}^{T}$.
- o_t is the number of security alerts at time t.



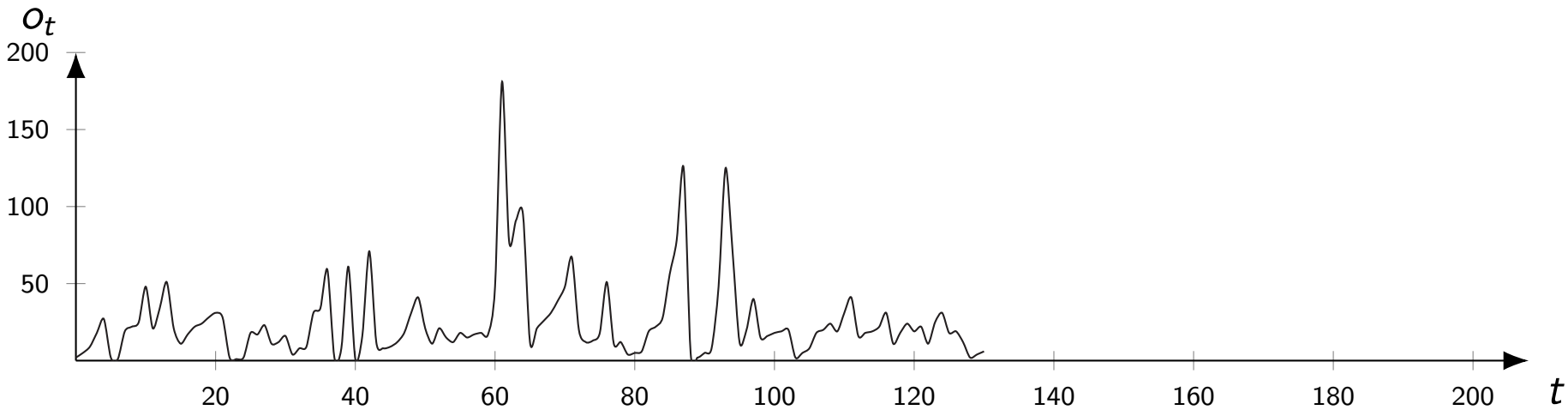
Optimal Stopping

- Observe the system through the stochastic process $(o_t)_{t=1}^{T}$.
- o_t is the number of security alerts at time t.



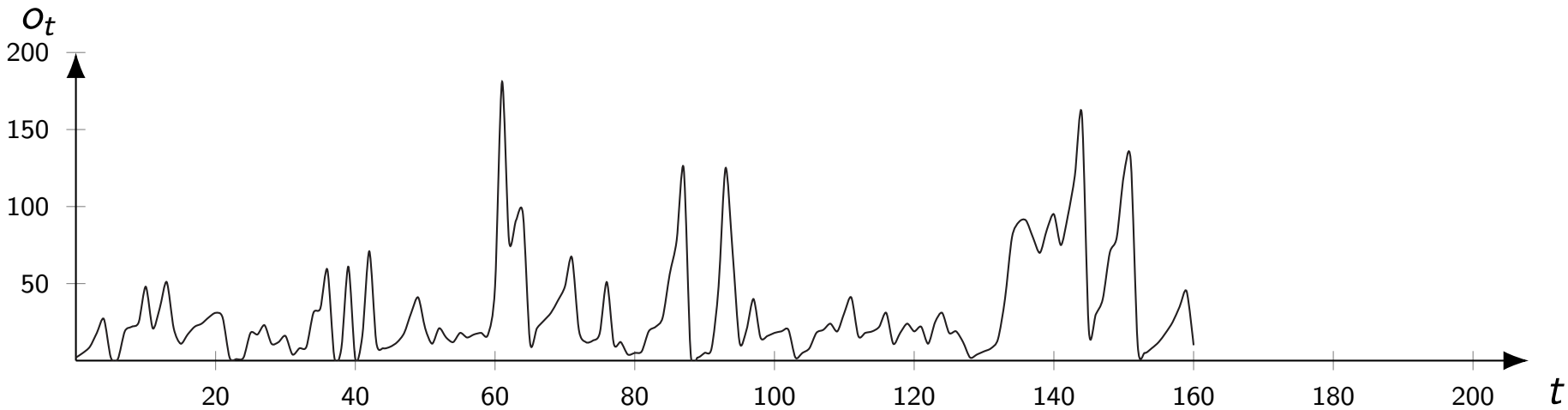
Optimal Stopping

- Observe the system through the stochastic process $(o_t)_{t=1}^{T}$.
- o_t is the number of security alerts at time t.



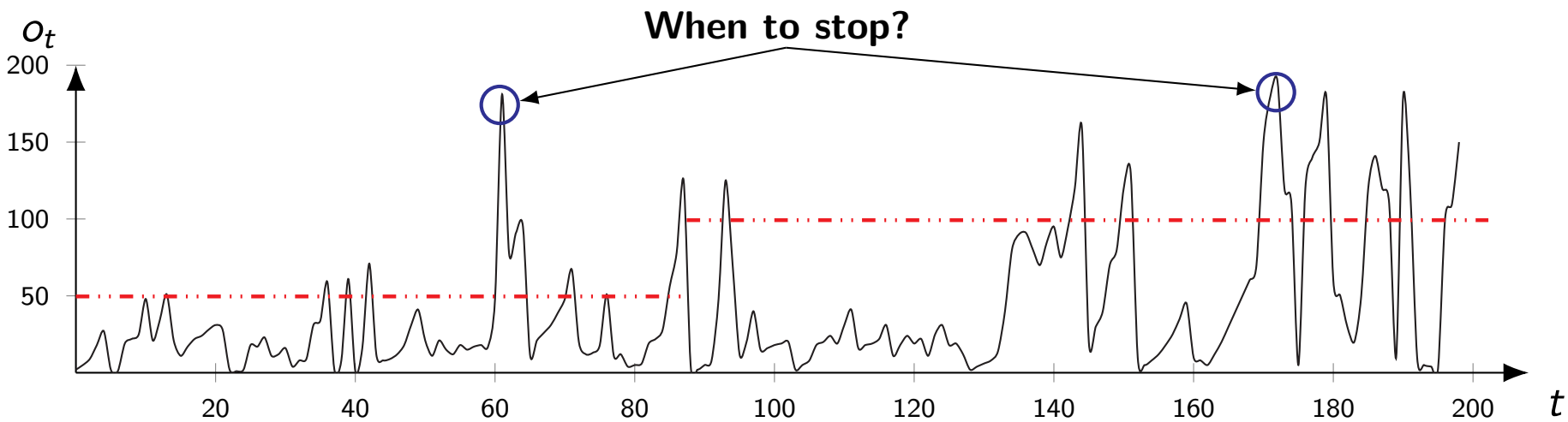
Optimal Stopping

- Observe the system through the stochastic process $(o_t)_{t=1}^{T}$.
- o_t is the number of security alerts at time t.



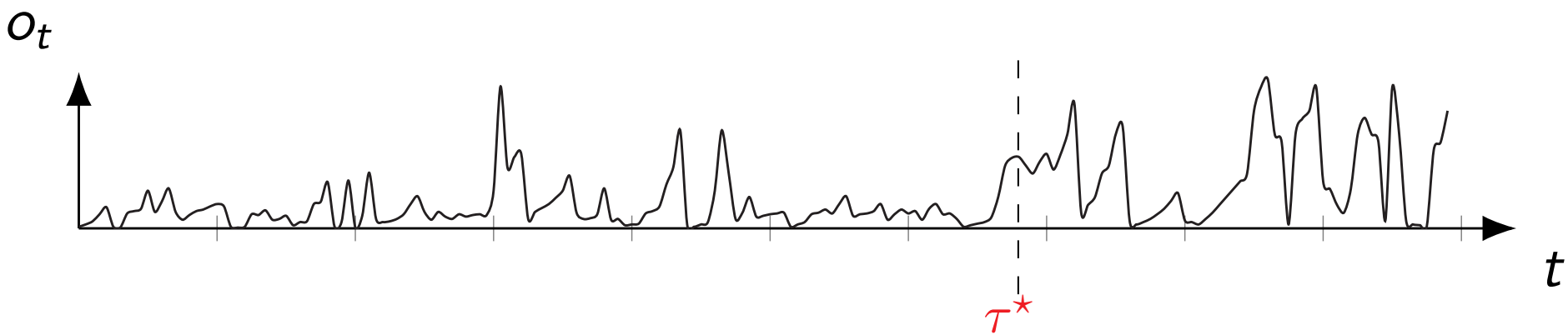
Optimal Stopping

- Observe the system through the stochastic process $(o_t)_{t=1}^{T}$.
- o_t is the number of security alerts at time t.

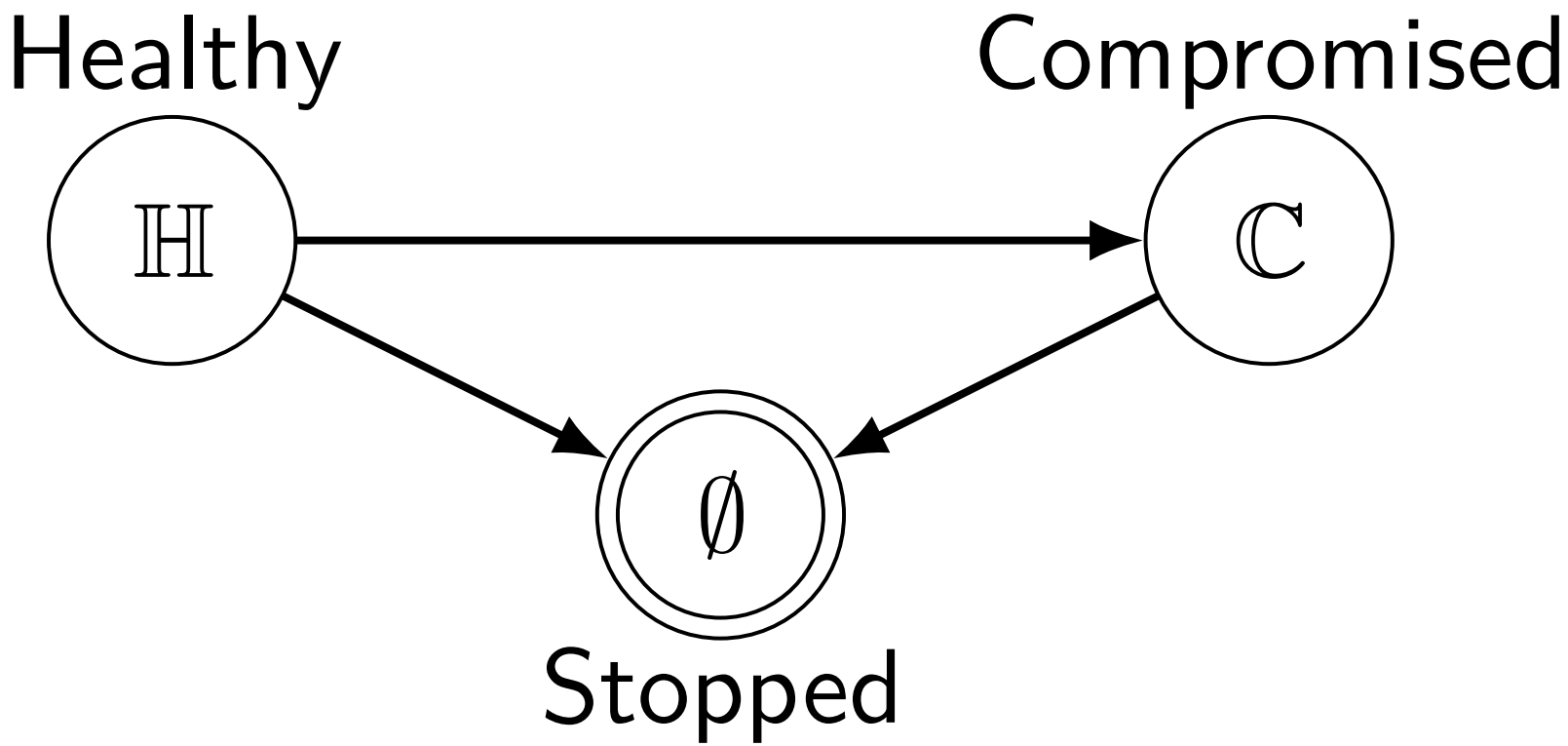


Optimal Stopping Time

- Find the ***optimal stopping time*** $\tau^\star \in \arg\max_\tau \mathbb{E}[J(\tau)]$.



Dynamical System



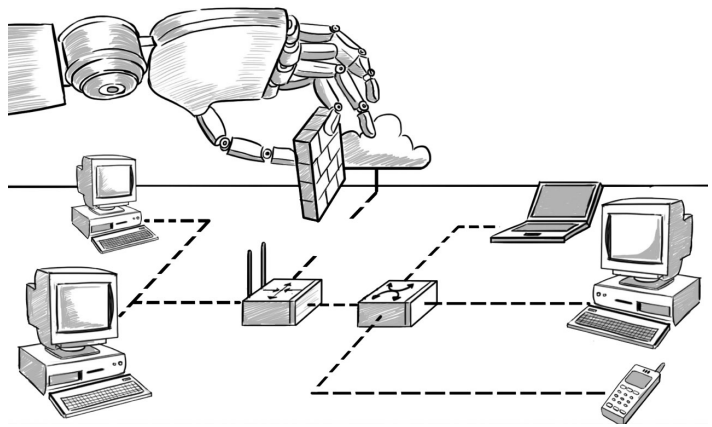
Challenge: System identification.

How to model the observation distribution $o_t \sim z(\cdot \mid s_t)$?

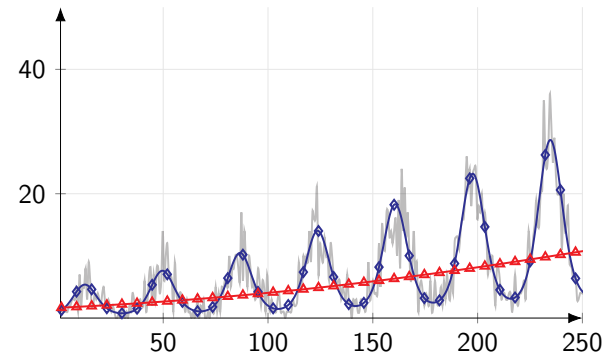
Inputs

Attacks
& responses

IT Infrastructure

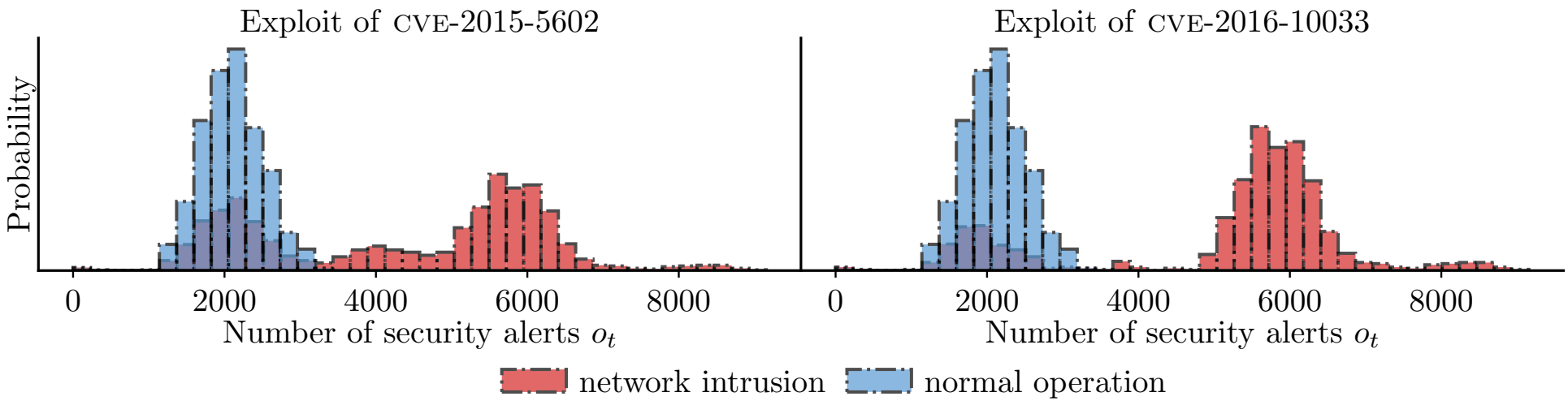


Observations



System Identification

- Measurement data from the digital twin:

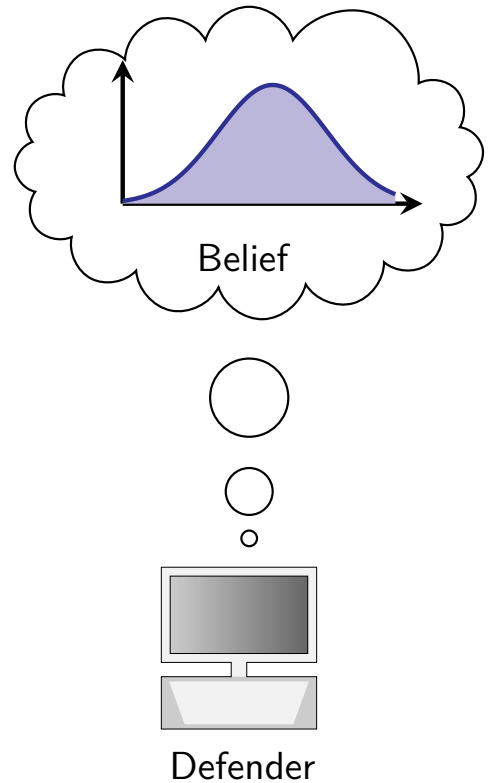


- **Estimate** $o_t \sim z(\cdot \mid s_t)$ with the empirical distribution $\widehat{z}$.
- $\widehat{z} \to^{\text{a.s}} z$ (Glivenko-Cantelli theorem).

Belief State

- The defender can compute the **belief**

$$b_t \triangleq \mathbb{P}[S_t = \mathbb{C} \mid b_1, o_1, o_2, \ldots o_t].$$

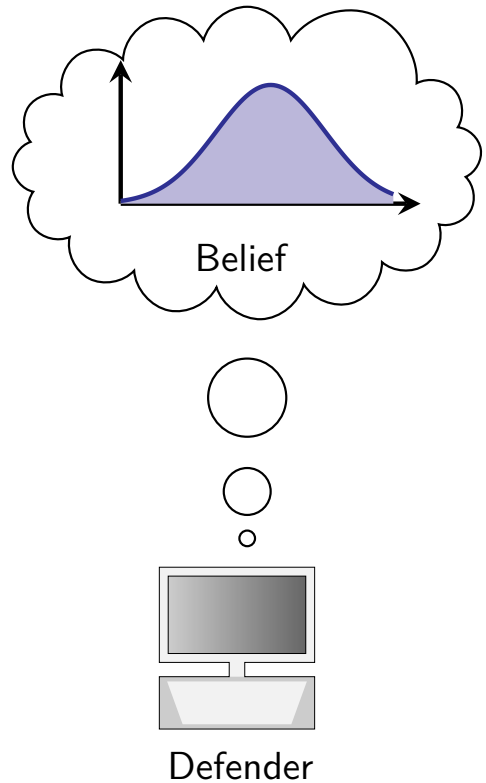


Belief State

- The defender can compute the **belief**

$$b_t \triangleq \mathbb{P}[S_t = \mathbb{C} \mid b_1, o_1, o_2, \ldots o_t].$$

- **Stopping strategy**:
  $\pi(b) : [0, 1] \rightarrow \{\text{Stop}, \text{Continue}\}$.



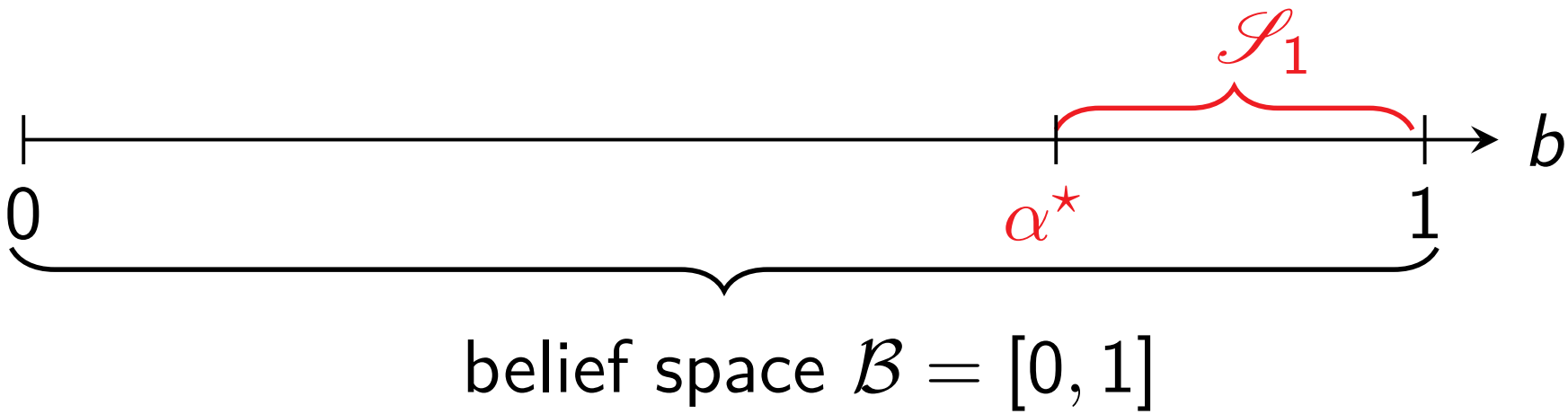
Threshold Structure of an Optimal strategy

Theorem

*There exists an **optimal defender strategy** of the form:*

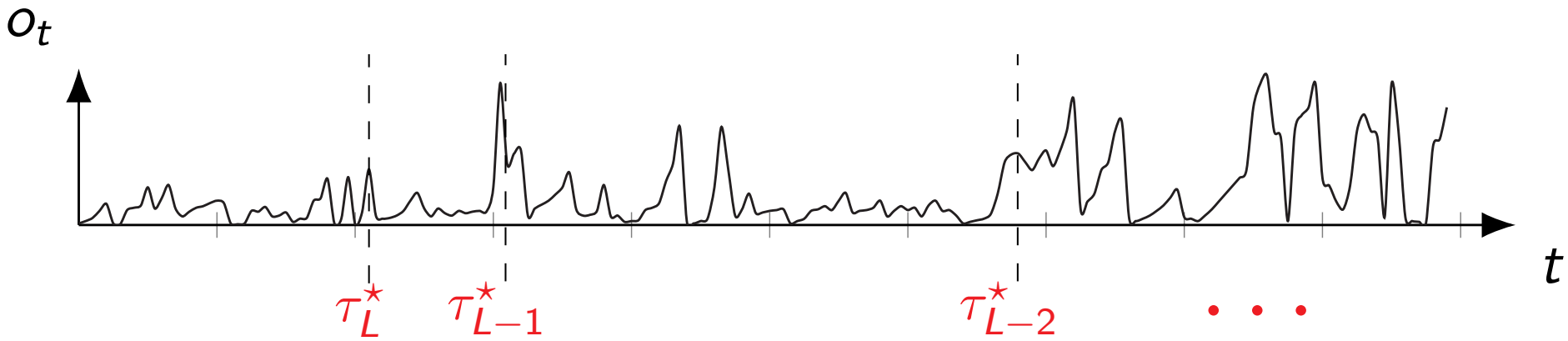
$$\pi^\star(b) = \textit{Stop} \iff b \geq \alpha^\star, \qquad \textit{where } \alpha^\star \in [0,1].$$

The stopping set is $\mathscr{S} = [\alpha^\star, 1]$.



Optimal **Multiple** Stopping

- Suppose the defender can take $L \geq 1$ **response actions**.
- Find the *optimal stopping times* $\tau_L^\star, \tau_{L-1}^\star, \ldots, \tau_1^\star$.



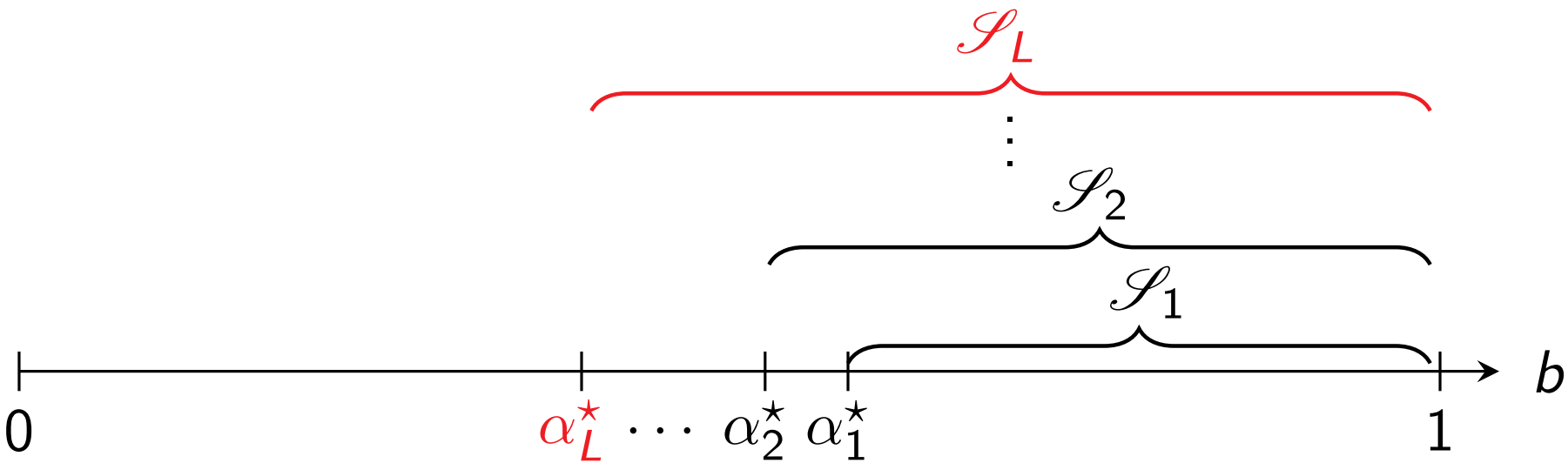
Threshold Structure of an Optimal Strategy

Theorem

- *Stopping sets are nested* $\mathscr{S}_{l-1} \subseteq \mathscr{S}_l$ *for* $l = 2, \dots L$.
- *There exists an* ***optimal defender strategy*** *of the form:*

$$\pi_l^\star(b) = \textit{Stop} \iff b \geq \alpha_l^\star, \qquad l = 1, \dots, L$$

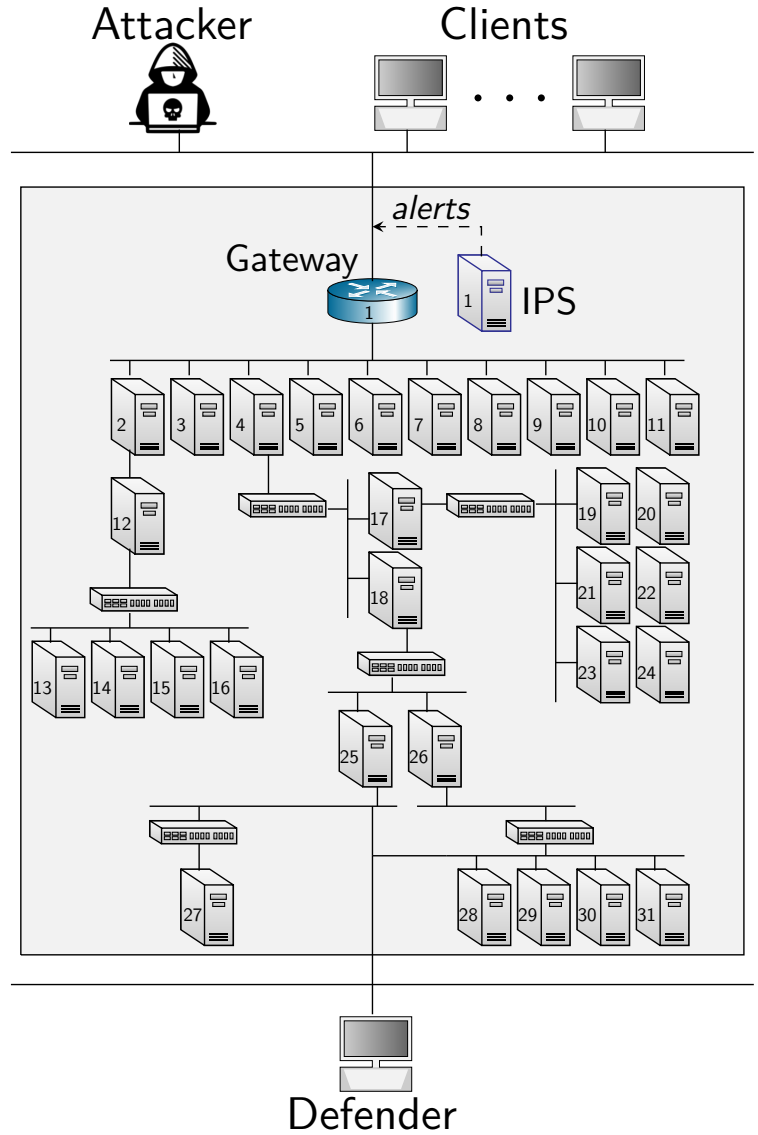
where $\alpha_l^\star \in [0,1]$ *is decreasing in* l.



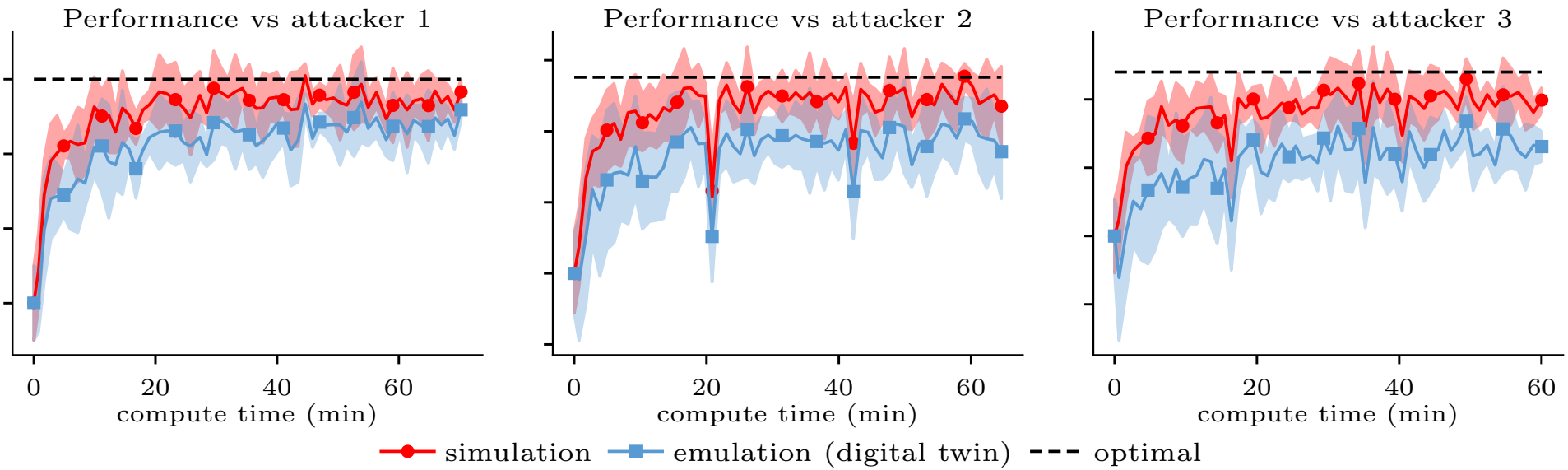
Testbed at KTH

Evaluation Scenario

- **IT Infrastructure**
  - 31 virtual servers (containers).
  - 11 vulnerabilities: CVE-2010-0426, CVE-2015-3306, etc.
- **Attacker**
  - 3 types of attackers.
  - Reconnaissance and exploits.
- **Defender**
  - Response action: block the gateway.
  - Threshold response strategy.

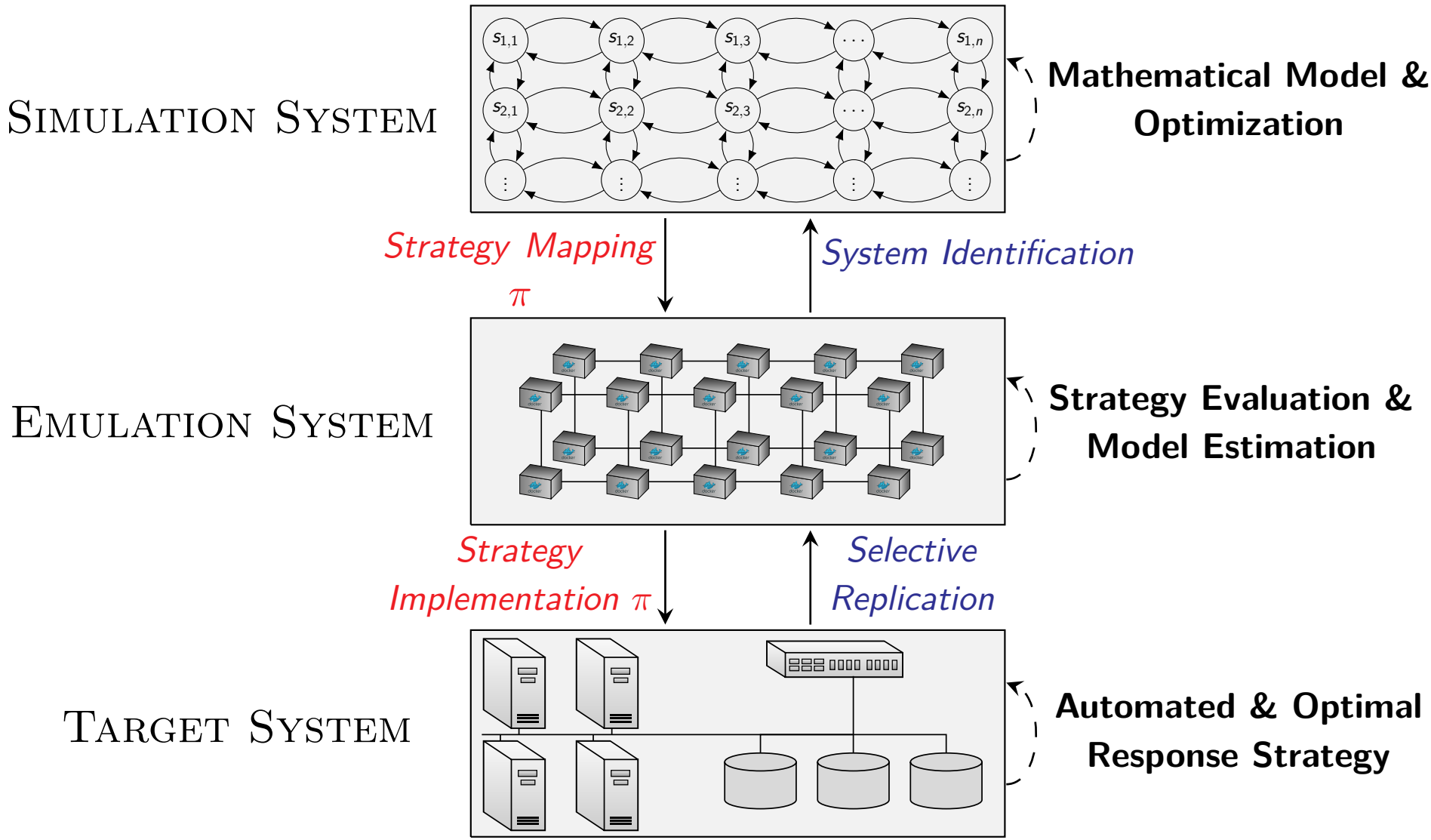


Evaluation Results

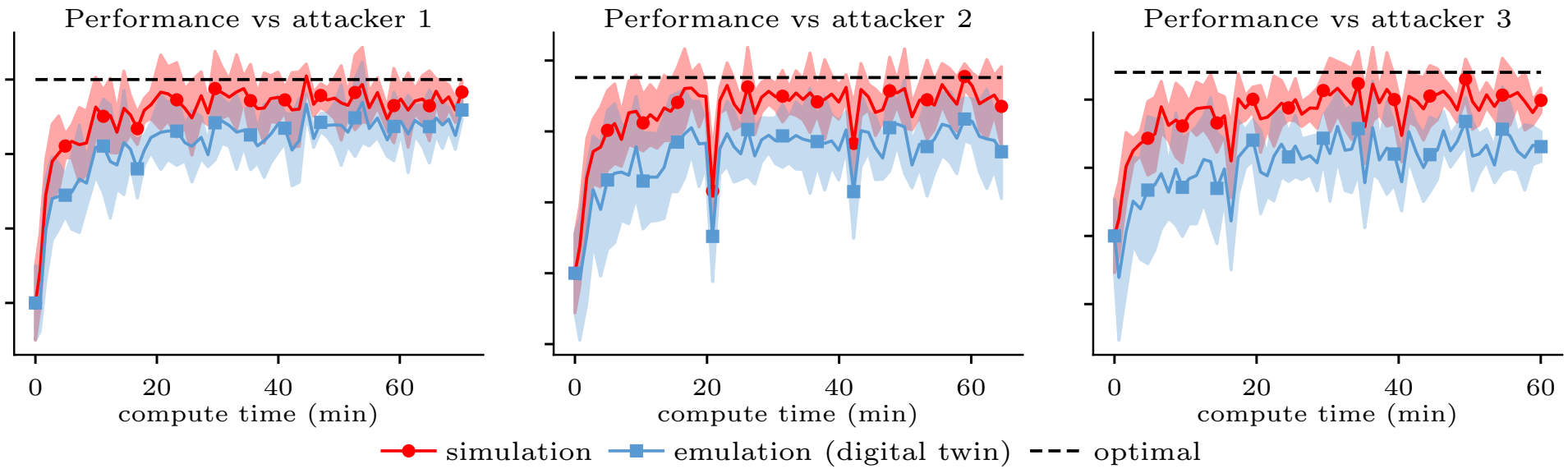


Performance on the simulator transfers to the digital twin.

Narrowing the **Gap between Theory and Practice**



Narrowing the **Gap between Theory and Practice**

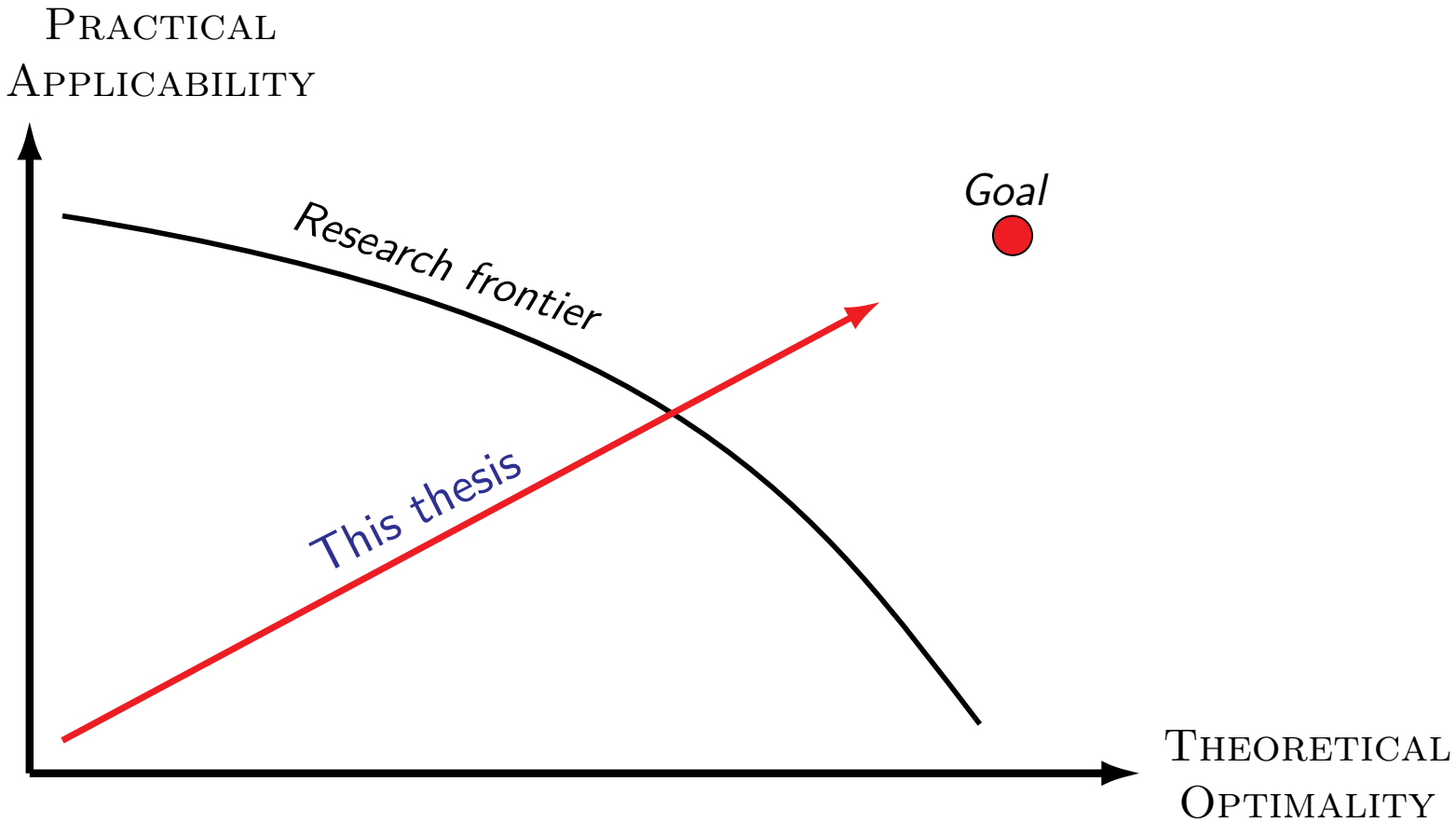


Performance on the simulator transfers to the digital twin.

What's new here?

First demonstration of optimal security response on an infrastructure with a practical configuration (31 servers).

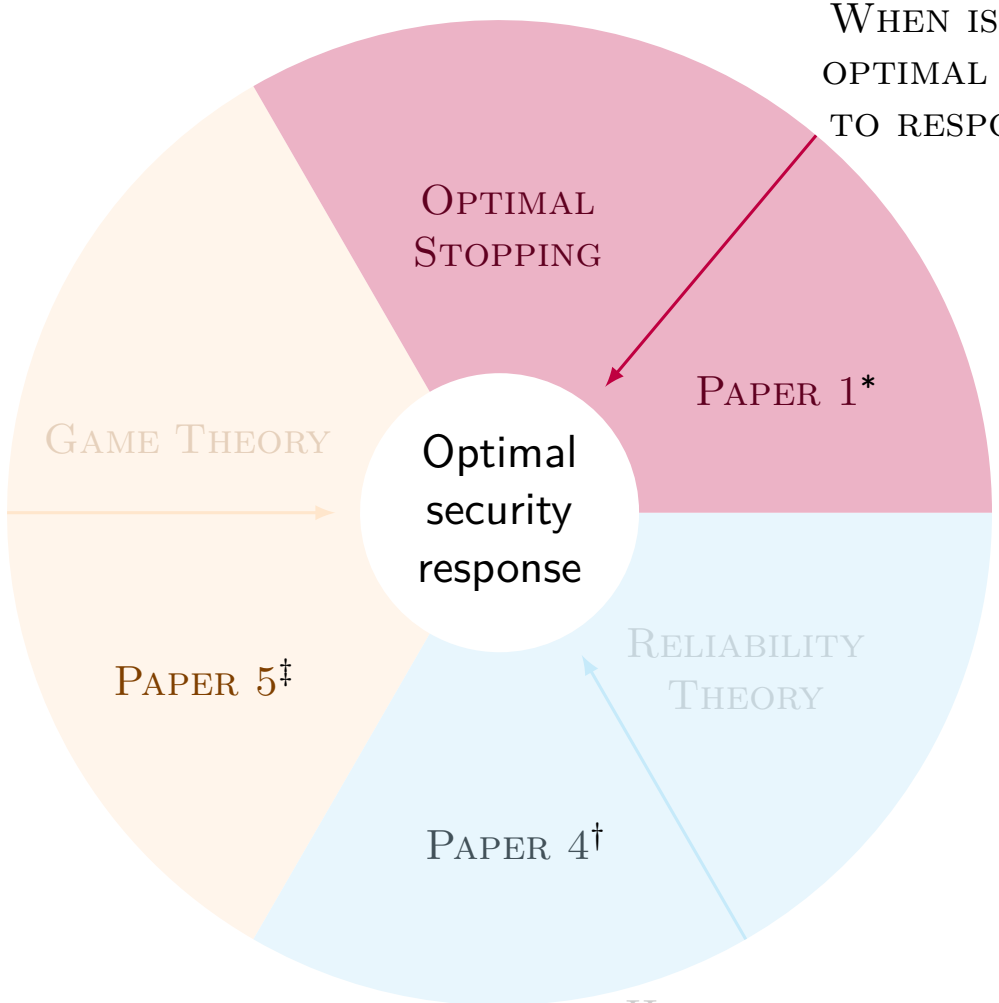
Comparison with Prior Work



Limitation of prior work.

Current response systems are based on **heuristics**. Optimal solutions have only been validated in simulation.

Case Studies

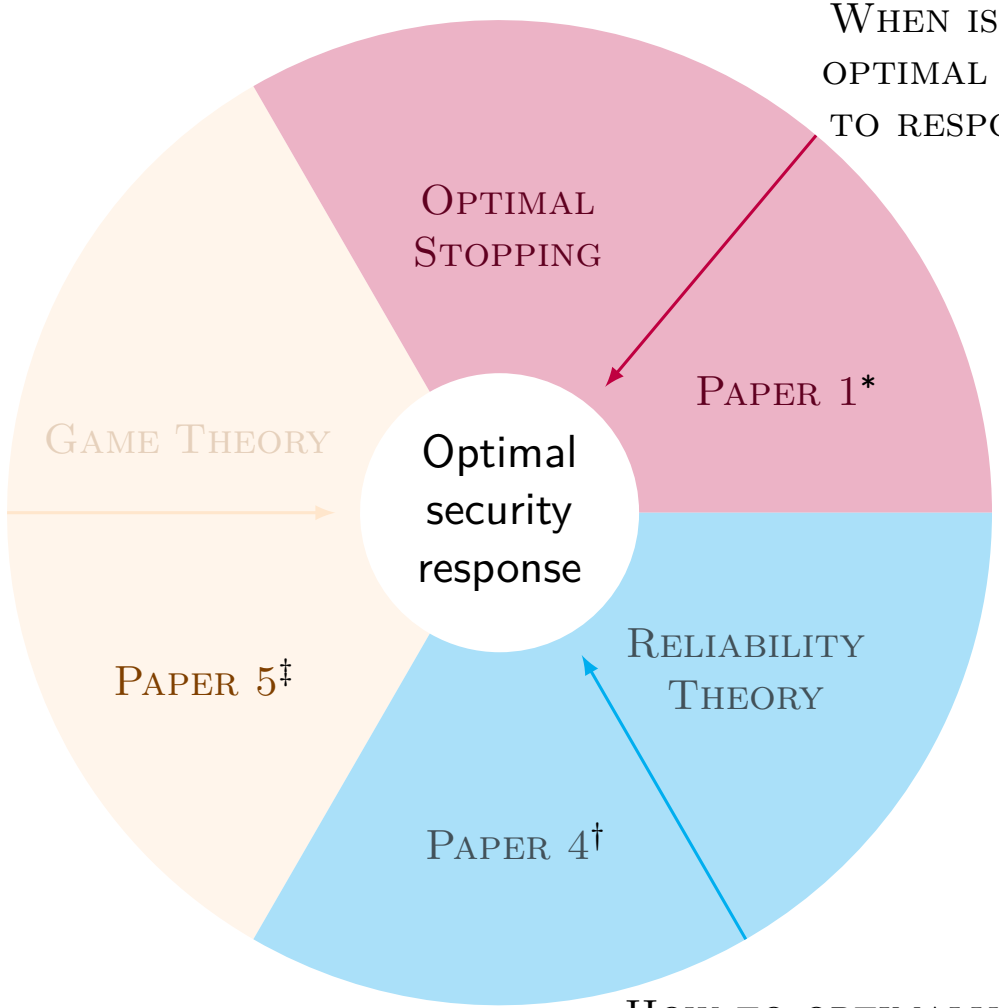


*Kim Hammar and Rolf Stadler. "Intrusion Prevention through Optimal Stopping". In: *IEEE Transactions on Network and Service Management* 19.3 (2022), pp. 2333–2348. DOI: 10.1109/TNSM.2022.3176781.

†Kim Hammar and Rolf Stadler. "Intrusion Tolerance for Networked Systems through Two-Level Feedback Control". In: *2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. 2024, pp. 338–352. DOI: 10.1109/DSN58291.2024.00042.

‡Kim Hammar et al. *Automated Security Response through Online Learning with Adaptive Conjectures*. https://arxiv.org/abs/2402.12499. To appear in IEEE Transactions on Information Forensics and Security

Case Studies



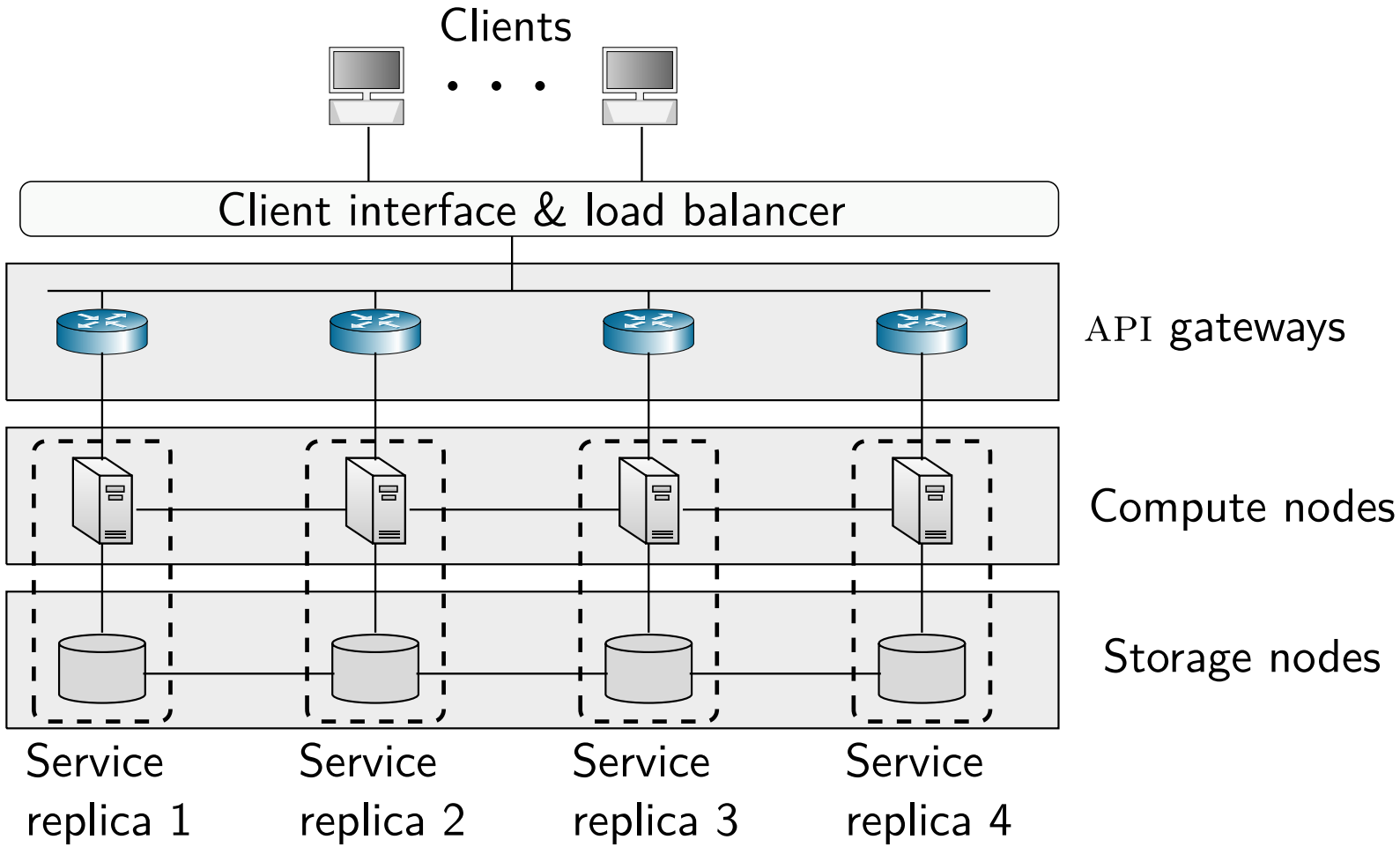
*Kim Hammar and Rolf Stadler. "Intrusion Prevention through Optimal Stopping". In: *IEEE Transactions on Network and Service Management* 19.3 (2022), pp. 2333–2348. DOI: 10.1109/TNSM.2022.3176781.

†Kim Hammar and Rolf Stadler. "Intrusion Tolerance for Networked Systems through Two-Level Feedback Control". In: *2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. 2024, pp. 338–352. DOI: 10.1109/DSN58291.2024.00042.

‡Kim Hammar et al. *Automated Security Response through Online Learning with Adaptive Conjectures*. https://arxiv.org/abs/2402.12499. To appear in IEEE Transactions on Information Forensics and Security

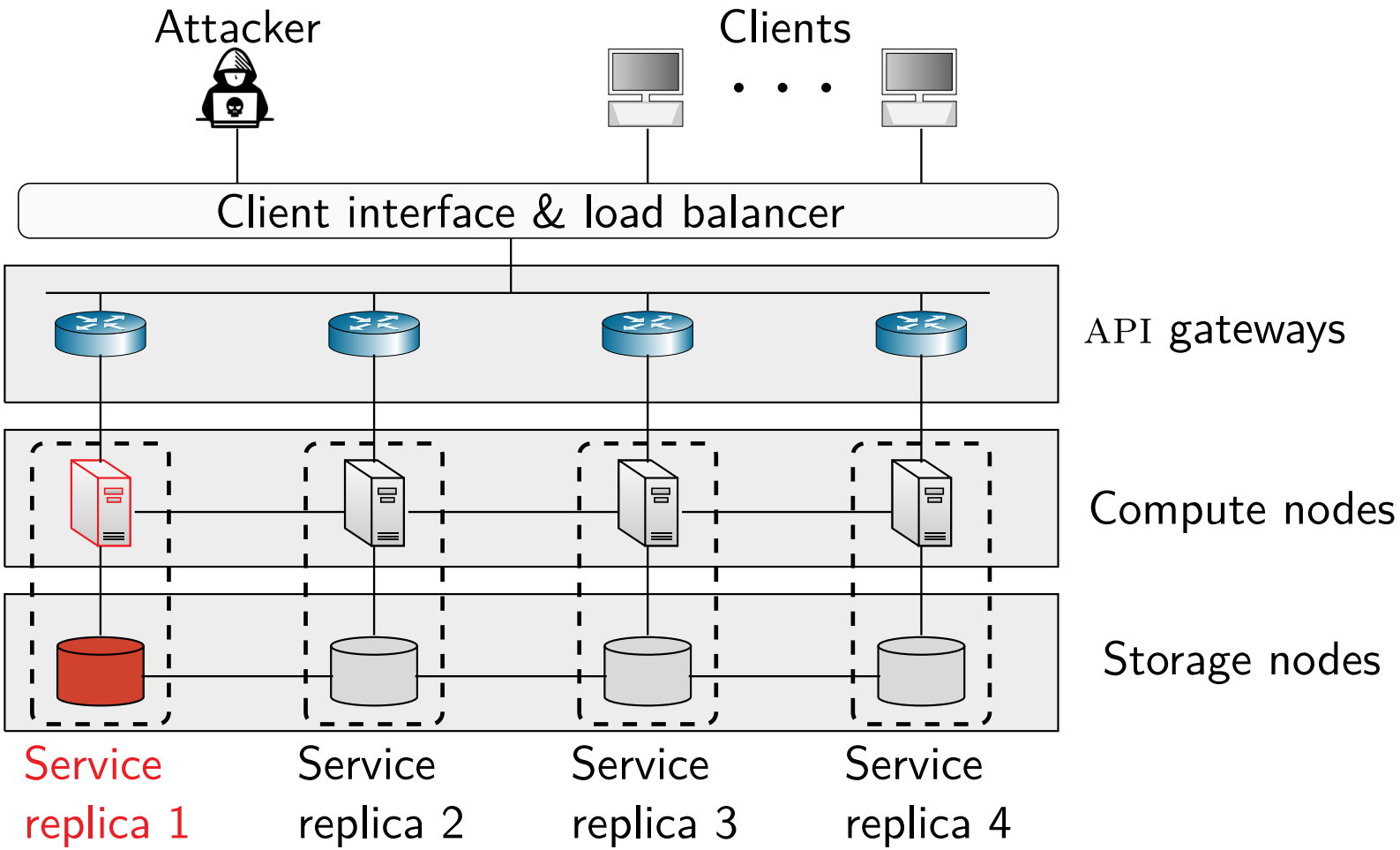
Intrusion Tolerance

Consider a **distributed system** that provides a replicated service.

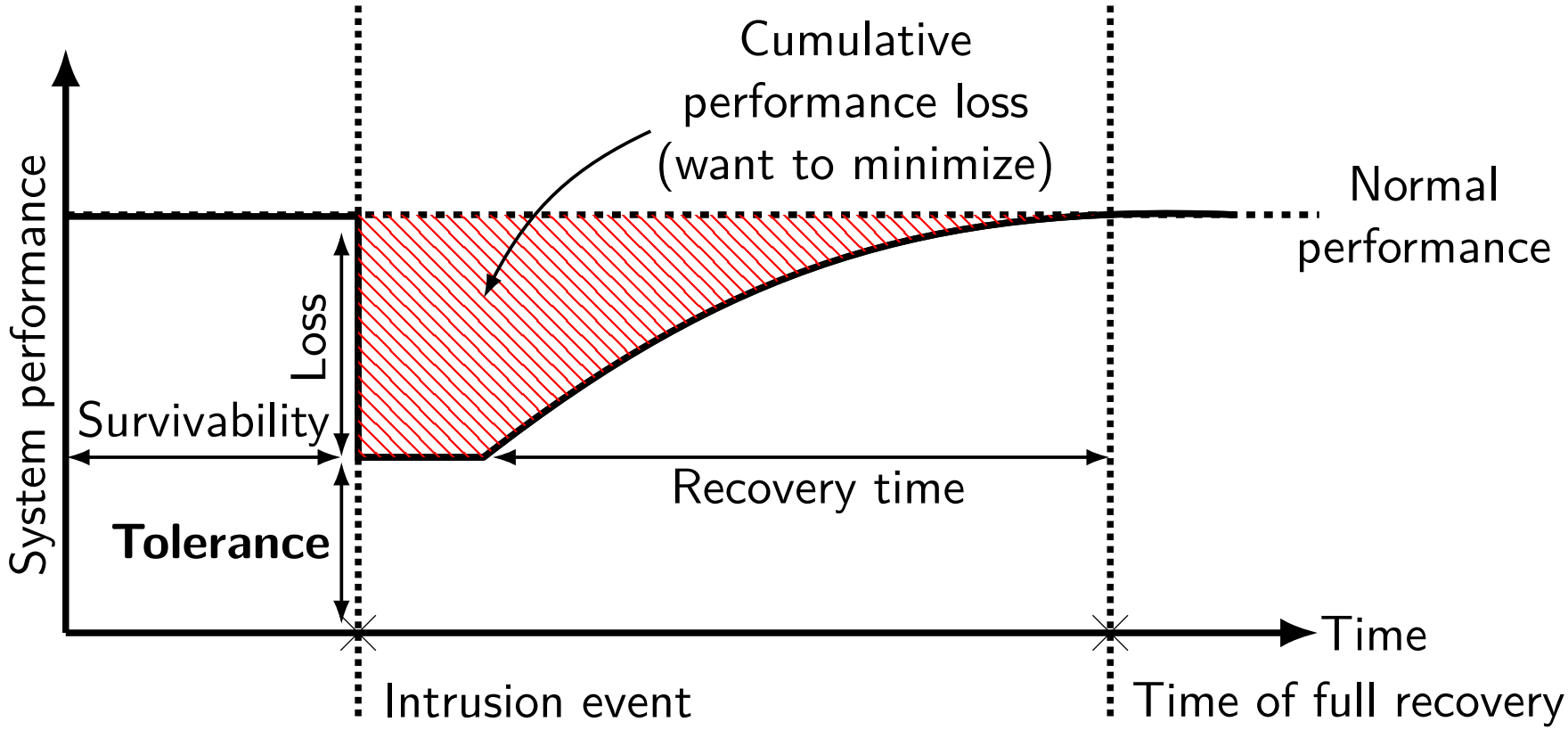


Intrusion Tolerance

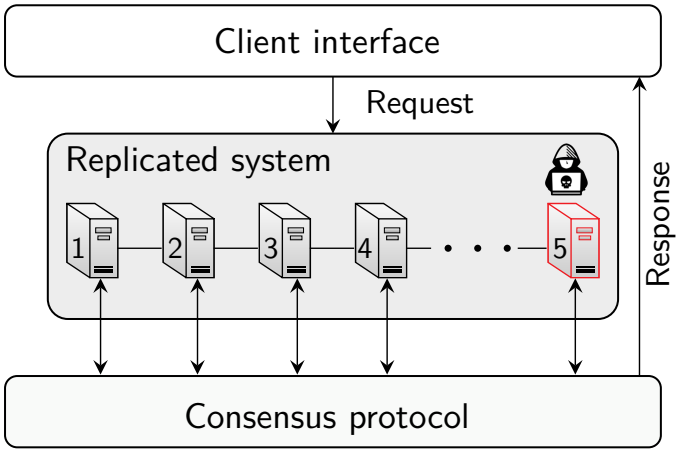
Consider a **distributed system** that provides a replicated service.
The system should tolerate intrusions.

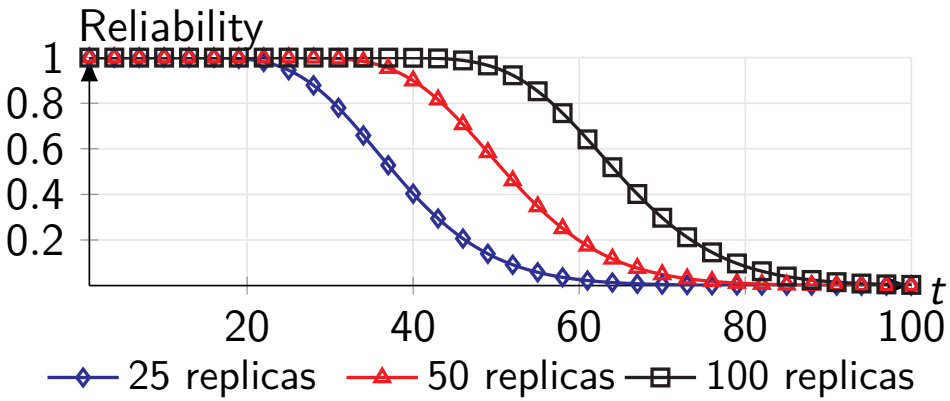


Intrusion Tolerance



Building Blocks of An Intrusion-Tolerant System



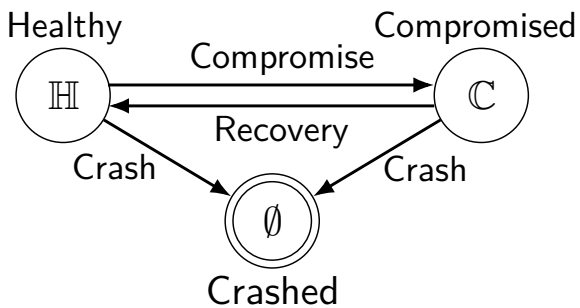


1. Intrusion-tolerant consensus protocol

A quorum needs to reach agreement to tolerate f compromised replicas.

2. Replication strategy

Cost-reliability trade-off.



3. Recovery strategy

Compromises will occur as $t \to \infty$.

Prior Work on Intrusion-Tolerant Systems

The Rampart Toolkit for Building High-Integrity Services

Michael K. Reiter

AT&T Bell Laboratories, Holmdel, New Jersey, USA
reiter@research.att.com

Abstract. Rampart is a toolkit of protocols to facilitate
ment of *high-integrity* services, i.e., distributed se
availability and correctness despite the maliciou
component servers by an attacker. At the core of
tocols that solve several basic problems in dist
cluding asynchronous group membership, reliabl
agreement), and atomic multicast. Using these p
ports the development of high-integrity services v
machine replication, and also extends this technique with a new approach
to server output voting. In this paper we give a brief overview of Rampart, focusing primarily on its protocol architecture. We also sketch its performance in our prototype implementation and ongoing work.

Published 1995

- Fixed number of replicas
- No recoveries

Prior Work on Intrusion-Tolerant Systems

The Rampart Toolkit for Building High-Integrity Services

Michael K. Reiter

AT&T Bell Laboratories, Holmdel, New Jersey, USA
reiter@research.att.com

Abstract. Rampart is a toolkit of protocols to facilitate the development of *high-integrity* services, i.e., distributed services that retain their availability and correctness despite the malicious penetration of some component servers by an attacker. At the core of Rampart are new protocols that solve several basic problems in distributed computing, including asynchronous group membership, reliable multicast (Byzantine agreement), and atomic multicast. Using these protocols, Rampart supports the development of high-integrity services via the technique of *state machine replication*, and also extends this technique with a new approach to server output voting. In this paper we give a brief overview of Rampart, focusing primarily on its protocol architecture. We also sketch its performance in our prototype implementation and ongoing work.

The SecureRing Protocols for Securing Group Communication*

Kim Potter Kihlstrom, L. E. Moser, P. M. Melliar-Smith
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
kimk@alpha.ece.ucsb.edu, moser@ece.ucsb.edu, pmms@ece.ucsb.edu

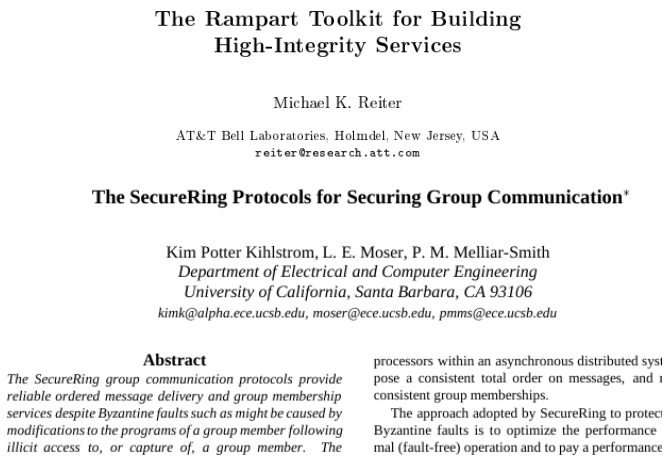
Abstract

The SecureRing group communication protocols prov
reliable ordered message delivery and group members
services despite Byzantine faults such as might be caused
modifications to the programs of a group member follow
illicit access to, or capture of, a group member.

Published 1998

- Fixed number of replicas
- No recoveries

Prior Work on Intrusion-Tolerant Systems



The Rampart Toolkit for Building
High-Integrity Services

Michael K. Reiter

AT&T Bell Laboratories, Holmdel, New Jersey, USA
reiter@research.att.com

The SecureRing Protocols for Securing Group Communication*

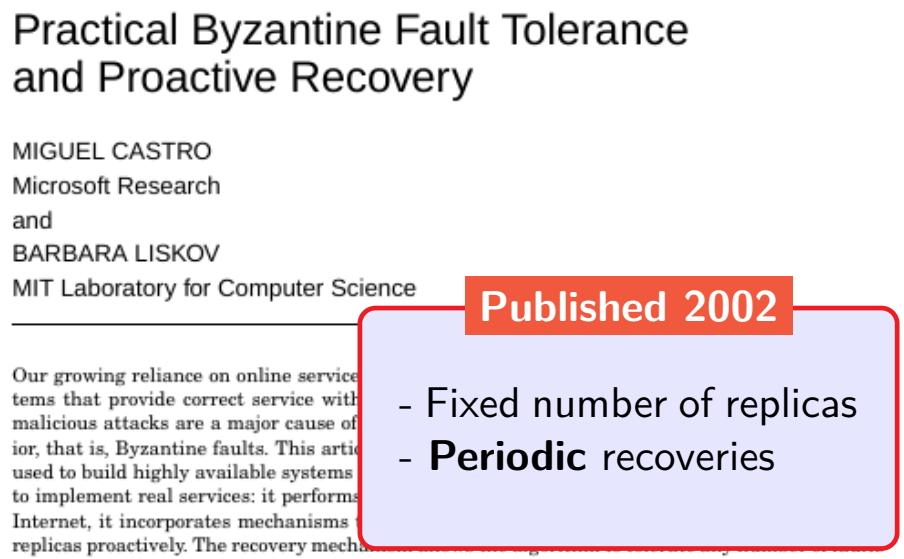
Kim Potter Kihlstrom, L. E. Moser, P. M. Melliar-Smith
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
kimk@alpha.ece.ucsb.edu, moser@ece.ucsb.edu, pmms@ece.ucsb.edu

Abstract

The SecureRing group communication protocols provide reliable ordered message delivery and group membership services despite Byzantine faults such as might be caused by modifications to the programs of a group member following illicit access to, or capture of, a group member. The

processors within an asynchronous distributed syst
pose a consistent total order on messages, and
consistent group memberships.

The approach adopted by SecureRing to protec
Byzantine faults is to optimize the performance
mal (fault-free) operation and to pay a performance



Practical Byzantine Fault Tolerance
and Proactive Recovery

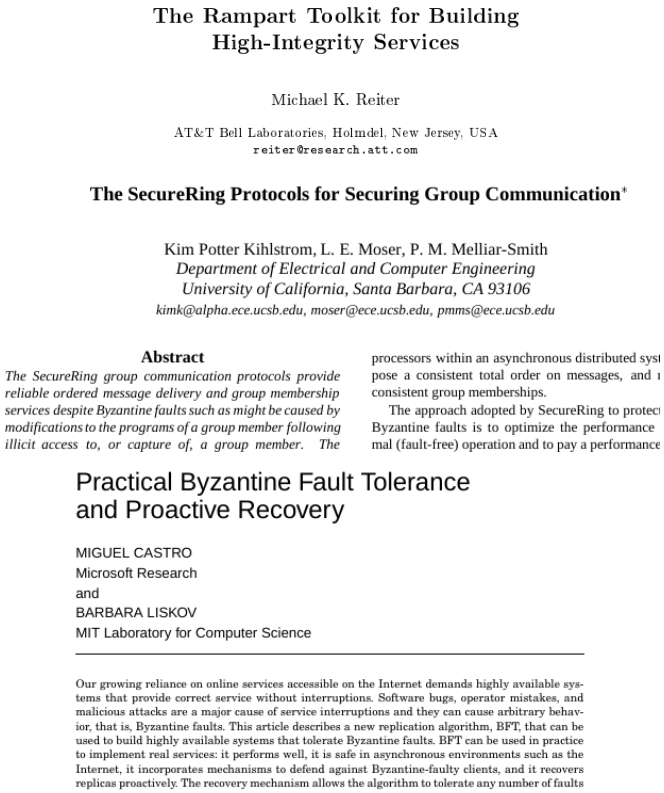
MIGUEL CASTRO
Microsoft Research
and
BARBARA LISKOV
MIT Laboratory for Computer Science

Our growing reliance on online service
tems that provide correct service with
malicious attacks are a major cause of
ior, that is, Byzantine faults. This arti
used to build highly available systems
to implement real services: it performs
Internet, it incorporates mechanisms
replicas proactively. The recovery mech

Published 2002

- Fixed number of replicas
- **Periodic** recoveries

Prior Work on Intrusion-Tolerant Systems



The Rampart Toolkit for Building
High-Integrity Services

Michael K. Reiter

AT&T Bell Laboratories, Holmdel, New Jersey, USA
reiter@research.att.com

The SecureRing Protocols for Securing Group Communication*

Kim Potter Kihlstrom, L. E. Moser, P. M. Melliar-Smith
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
kimk@alpha.ece.ucsb.edu, moser@ece.ucsb.edu, pmms@ece.ucsb.edu

Abstract

The SecureRing group communication protocols provide reliable ordered message delivery and group membership services despite Byzantine faults such as might be caused by modifications to the programs of a group member following illicit access to, or capture of, a group member. The

processors within an asynchronous distributed syst
pose a consistent total order on messages, and
consistent group memberships.

The approach adopted by SecureRing to protec
Byzantine faults is to optimize the performance
mal (fault-free) operation and to pay a performance

Practical Byzantine Fault Tolerance
and Proactive Recovery

MIGUEL CASTRO
Microsoft Research
and
BARBARA LISKOV
MIT Laboratory for Computer Science

Our growing reliance on online services accessible on the Internet demands highly available systems that provide correct service without interruptions. Software bugs, operator mistakes, and malicious attacks are a major cause of service interruptions and they can cause arbitrary behavior, that is, Byzantine faults. This article describes a new replication algorithm, BFT, that can be used to build highly available systems that tolerate Byzantine faults. BFT can be used in practice to implement real services: it performs well, it is safe in asynchronous environments such as the Internet, it incorporates mechanisms to defend against Byzantine-faulty clients, and it recovers replicas proactively. The recovery mechanism allows the algorithm to tolerate any number of faults

A Qualitative Analysis of the Intrusion-Tolerance Capabilities of the MAFTIA Architecture

Robert Stroud, Ian Welch[1], John Warne, Peter Ryan,
School of Computing Science, University of Newcastle upon Tyne, UK
{R.J.Stroud, J.P.Warne, Peter.Ryan}@ncl.ac.uk
Ian.Welch@mcs.vu

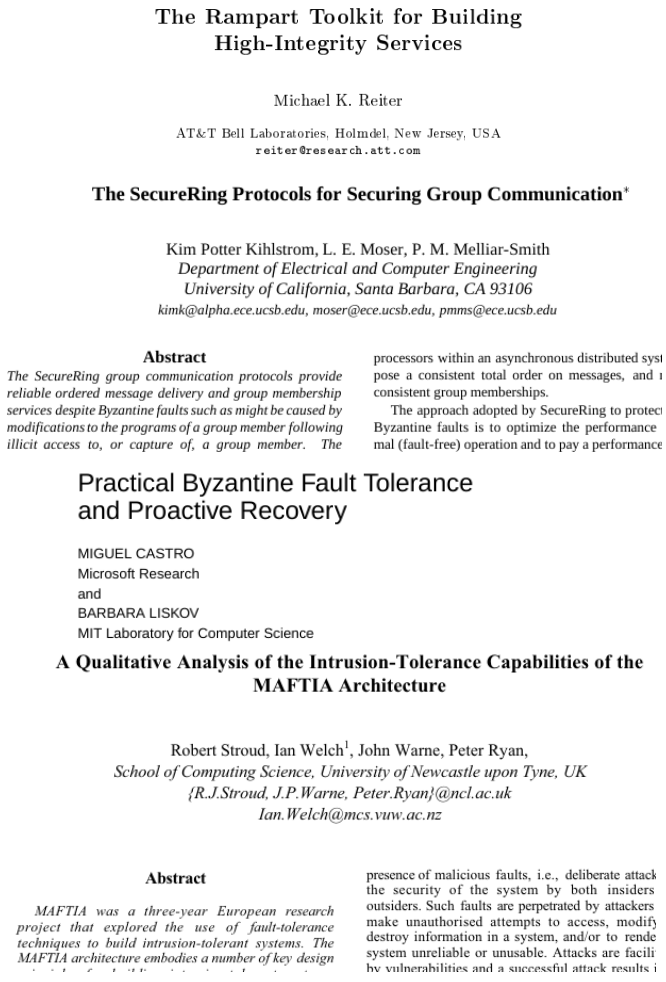
Abstract

MAFTIA was a three-year European resea
project that explored the use of fault-toleran
techniques to build intrusion-tolerant systems. T
MAFTIA architecture embodies a number of key des

Published 2004

- Fixed number of replicas
- Periodic recoveries

Prior Work on Intrusion-Tolerant Systems



The Rampart Toolkit for Building
High-Integrity Services

Michael K. Reiter

AT&T Bell Laboratories, Holmdel, New Jersey, USA
reiter@research.att.com

The SecureRing Protocols for Securing Group Communication*

Kim Potter Kihlstrom, L. E. Moser, P. M. Melliar-Smith
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
kimk@alpha.ece.ucsb.edu, moser@ece.ucsb.edu, pmms@ece.ucsb.edu

Abstract

The SecureRing group communication protocols provide reliable ordered message delivery and group membership services despite Byzantine faults such as might be caused by modifications to the programs of a group member following illicit access to, or capture of, a group member. The

processors within an asynchronous distributed syst
pose a consistent total order on messages, and
consistent group memberships.

The approach adopted by SecureRing to protect
Byzantine faults is to optimize the performance
mal (fault-free) operation and to pay a performance

Practical Byzantine Fault Tolerance
and Proactive Recovery

MIGUEL CASTRO
Microsoft Research
and
BARBARA LISKOV
MIT Laboratory for Computer Science

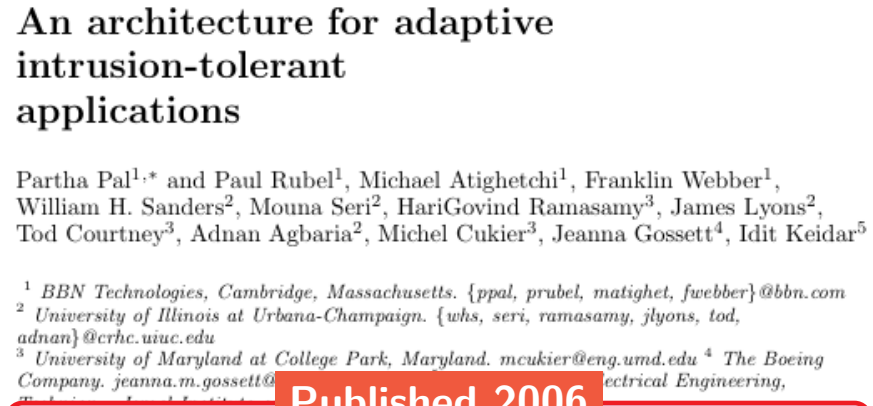
A Qualitative Analysis of the Intrusion-Tolerance Capabilities of the MAFTIA Architecture

Robert Stroud, Ian Welch[1], John Warne, Peter Ryan,
School of Computing Science, University of Newcastle upon Tyne, UK
{R.J.Stroud, J.P.Warne, Peter.Ryan}@ncl.ac.uk
Ian.Welch@mcs.vuw.ac.nz

Abstract

MAFTIA was a three-year European research project that explored the use of fault-tolerance techniques to build intrusion-tolerant systems. The MAFTIA architecture embodies a number of key design

presence of malicious faults, i.e., deliberate attack
the security of the system by both insiders
outsiders. Such faults are perpetrated by attackers
make unauthorised attempts to access, modify
destroy information in a system, and/or to rende
system unreliable or unusable. Attacks are facili
by vulnerabilities and a successful attack results i



An architecture for adaptive intrusion-tolerant applications

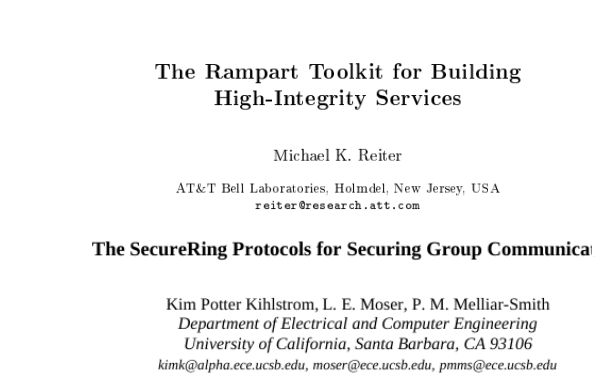
Partha Pal[1,*] and Paul Rubel[1], Michael Atighetchi[1], Franklin Webber[1], William H. Sanders[2], Mouna Seri[2], HariGovind Ramasamy[3], James Lyons[2], Tod Courtney[3], Adnan Agbaria[2], Michel Cukier[3], Jeanna Gossett[4], Idit Keidar[5]

[1] *BBN Technologies, Cambridge, Massachusetts. {ppal, prubel, matighet, fwebber}@bbn.com*
[2] *University of Illinois at Urbana-Champaign. {whs, seri, ramasamy, jlyons, tod, adnan}@crhc.uiuc.edu*
[3] *University of Maryland at College Park, Maryland. mcukier@eng.umd.edu* [4] *The Boeing Company. jeanna.m.gossett@* ectrical Engineering,

Published 2006

- **Adaptive replication** based on heuristics
- Periodic recoveries

Prior Work on Intrusion-Tolerant Systems



The Rampart Toolkit for Building High-Integrity Services

Michael K. Reiter

AT&T Bell Laboratories, Holmdel, New Jersey, USA
reiter@research.att.com

The SecureRing Protocols for Securing Group Communica

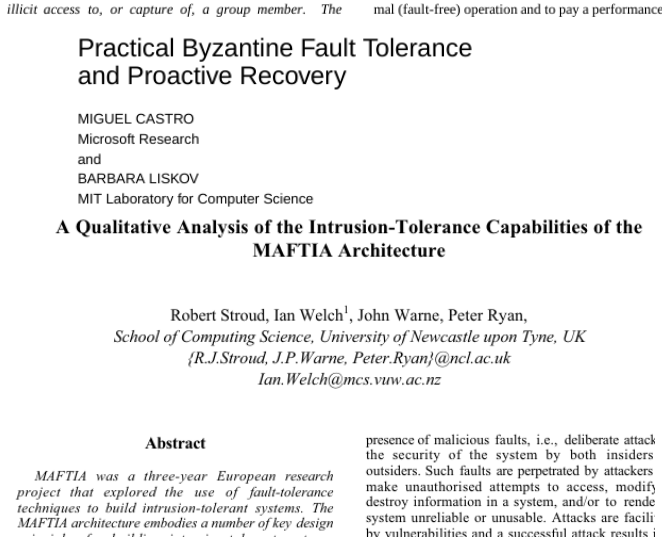
Kim Potter Kihlstrom, L. E. Moser, P. M. Melliar-Smith
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
kimk@alpha.ece.ucsb.edu, moser@ece.ucsb.edu, pmms@ece.ucsb.edu

Abstract

The SecureRing group communication protocols provide reliable ordered message delivery and group membership services despite Byzantine faults such as might be caused by modifications to the programs of a group member following illicit access to, or capture of, a group member. The

processors within an asynchronous distributed syst
pose a consistent total order on messages, and
consistent group memberships.

The approach adopted by SecureRing to protec
Byzantine faults is to optimize the performance
mal (fault-free) operation and to pay a performance



Practical Byzantine Fault Tolerance and Proactive Recovery

MIGUEL CASTRO
Microsoft Research
and
BARBARA LISKOV
MIT Laboratory for Computer Science

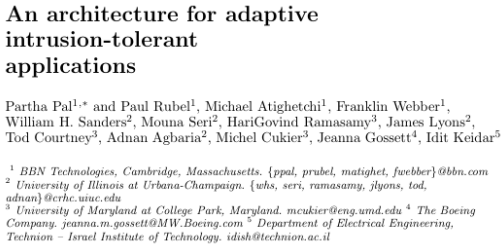
A Qualitative Analysis of the Intrusion-Tolerance Capabilities of the MAFTIA Architecture

Robert Stroud, Ian Welch[1], John Warne, Peter Ryan,
School of Computing Science, University of Newcastle upon Tyne, UK
{R.J.Stroud, J.P.Warne, Peter.Ryan}@ncl.ac.uk
Ian.Welch@mcs.vuw.ac.nz

Abstract

MAFTIA was a three-year European research project that explored the use of fault-tolerance techniques to build intrusion-tolerant systems. The MAFTIA architecture embodies a number of key design

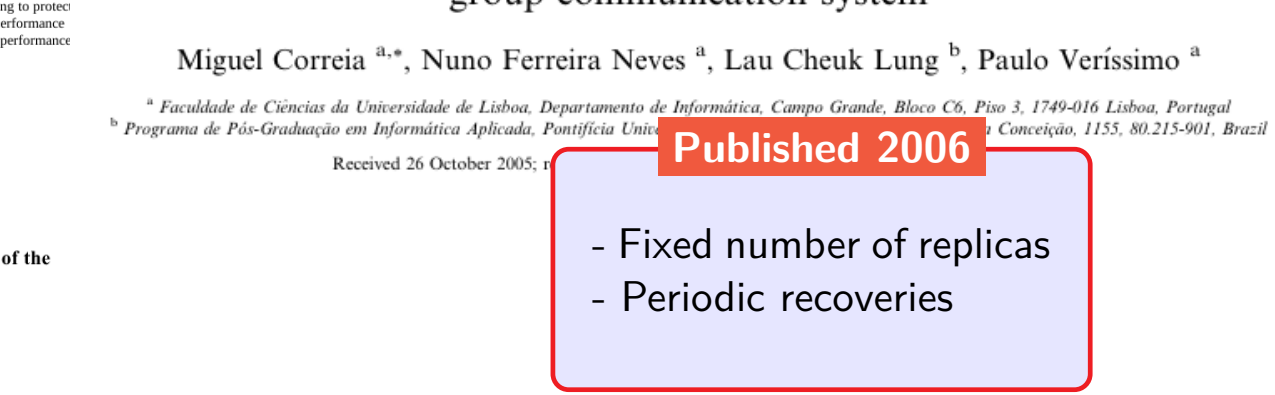
presence of malicious faults, i.e., deliberate attack
the security of the system by both insiders
outsiders. Such faults are perpetrated by attackers
make unauthorised attempts to access, modify
destroy information in a system, and/or to rende
system unreliable or unusable. Attacks are facili
by vulnerabilities and a successful attack results i



An architecture for adaptive intrusion-tolerant applications

Partha Pal[1,*] and Paul Rubel[1], Michael Atighetchi[1], Franklin Webber[1], William H. Sanders[2], Mouna Seri[2], HariGovind Ramasamy[3], James Lyons[2], Tod Courtney[3], Adnan Agbaria[2], Michel Cukier[3], Jeanna Gossett[4], Idit Keidar[5]

[1] *BBN Technologies, Cambridge, Massachusetts. {ppal, prubel, matighet, fwebber}@bbn.com*
[2] *University of Illinois at Urbana-Champaign. {whs, seri, ramasamy, jlyons, tod, adnan}@crhc.uiuc.edu*
[3] *University of Maryland at College Park, Maryland. mcukier@eng.umd.edu* [4] *The Boeing Company. jeanna.m.gossett@MW.Boeing.com* [5] *Department of Electrical Engineering, Technion – Israel Institute of Technology. idish@technion.ac.il*



Worm-IT – A wormhole-based intrusion-tolerant group communication system

Miguel Correia [a,*], Nuno Ferreira Neves [a], Lau Cheuk Lung [b], Paulo Veríssimo [a]

[a] *Faculdade de Ciências da Universidade de Lisboa, Departamento de Informática, Campo Grande, Bloco C6, Piso 3, 1749-016 Lisboa, Portugal*
[b] *Programa de Pós-Graduação em Informática Aplicada, Pontifícia Univ… Conceição, 1155, 80.215-901, Brazil*

Received 26 October 2005; r…

Published 2006

- Fixed number of replicas
- Periodic recoveries

Prior Work on Intrusion-Tolerant Systems

The Rampart Toolkit for Building High-Integrity Services

Michael K. Reiter

AT&T Bell Laboratories, Holmdel, New Jersey, USA
reiter@research.att.com

The SecureRing Protocols for Securing Group Communica

Kim Potter Kihlstrom, L. E. Moser, P. M. Melliar-Smith
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
kimk@alpha.ece.ucsb.edu, moser@ece.ucsb.edu, pmms@ece.ucsb.edu

Abstract

The SecureRing group communication protocols provide reliable ordered message delivery and group membership services despite Byzantine faults such as might be caused by modifications to the programs of a group member following illicit access to, or capture of, a group member. The

processors within an asynchronous distributed syst
pose a consistent total order on messages, and
consistent group memberships.

The approach adopted by SecureRing to protect
Byzantine faults is to optimize the performance
mal (fault-free) operation and to pay a performance

An architecture for adaptive intrusion-tolerant applications

Partha Pal[1,*] and Paul Rubel[1], Michael Atighetchi[1], Franklin Webber[1], William H. Sanders[2], Mouna Seri[2], HariGovind Ramasamy[3], James Lyons[2], Tod Courtney[3], Adnan Agbaria[2], Michel Cukier[3], Jeanna Gossett[4], Idit Keidar[5]

[1] *BBN Technologies, Cambridge, Massachusetts. {ppal, prubel, matighet, fwebber}@bbn.com*
[2] *University of Illinois at Urbana-Champaign. {whs, seri, ramasamy, jlyons, tod, adnan}@crhc.uiuc.edu*
[3] *University of Maryland at College Park, Maryland. mcukier@eng.umd.edu* [4] *The Boeing Company. jeanna.m.gossett@MW.Boeing.com* [5] *Department of Electrical Engineering, Technion – Israel Institute of Technology. idish@technion.ac.il*

Practical Byzantine Fault Tolerance and Proactive Recovery

MIGUEL CASTRO
Microsoft Research
and
BARBARA LISKOV
MIT Laboratory for Computer Science

A Qualitative Analysis of the Intrusion-Tolerance Capabilities of the MAFTIA Architecture

Robert Stroud, Ian Welch[1], John Warne, Peter Ryan,
School of Computing Science, University of Newcastle upon Tyne, UK
{R.J.Stroud, J.P.Warne, Peter.Ryan}@ncl.ac.uk
Ian.Welch@mcs.vuw.ac.nz

Worm-IT – A wormhole-based intrusion-tolerant group communication system

Miguel Correia [a,*], Nuno Ferreira Neves [a], Lau Cheuk Lung [b], Paulo Veríssimo [a]

[a] *Faculdade de Ciências da Universidade de Lisboa, Departamento de Informática, Campo Grande, Bloco C6, Piso 3, 1749-016 Lisboa, Portugal*
[b] *Programa de Pós-Graduação em Informática Aplicada, Pontifícia Universidade Católica do Paraná, Rua Imaculada Conceição, 1155, 80.215-901, Brazil*

Received 26 October 2005; received in revised form 28 March 2006; accepted 30 March 2006

Resilient Intrusion Tolerance through Proactive and Reactive Recovery*

Paulo Sousa Alysson Neves Bessani Miguel Correia
Nuno Ferreira Neves Paulo Verissimo
LASIGE, Faculdade de Ciências da Universidade de Lisboa – Portugal
{pjsousa, bessani, mpc, nuno, pjv}@di.fc.ul.pt

Published 2007

- Fixed number of replicas
- **Supports both periodic and reactive recoveries**
- Does not provide reactive recovery strategies

Prior Work on Intrusion-Tolerant Systems

The Rampart Toolkit for Building High-Integrity Services

Michael K. Reiter

AT&T Bell Laboratories, Holmdel, New Jersey, USA
reiter@research.att.com

An architecture for adaptive intrusion-tolerant applications

Partha Pal[1,*] and Paul Rubel[1], Michael Atighetchi[1], Franklin Webber[1], William H. Sanders[2], Mouna Seri[2], HariGovind Ramasamy[3], James Lyons[2], Tod Courtney[3], Adnan Agbaria[2], Michel Cukier[3], Jeanna Gossett[4], Idit Keidar[5]

[1] *BBN Technologies, Cambridge, Massachusetts. {ppal, prubel, matighet, fwebber}@bbn.com*
[2] *University of Illinois at Urbana-Champaign. {whs, seri, ramasamy, jlyons, tod, adnan}@crhc.uiuc.edu*
[3] *University of Maryland at College Park, Maryland. mcukier@eng.umd.edu* [4] *The Boeing Company. jeanna.m.gossett@MW.Boeing.com* [5] *Department of Electrical Engineering, Technion – Israel Institute of Technology. idish@technion.ac.il*

The SecureRing Protocols for Securing Group Communica

Kim Potter Kihlstrom, L. E. Moser, P. M. Melliar-Smi
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
kimk@alpha.ece.ucsb.edu, moser@ece.ucsb.edu, pmms@ece.ucsb.

Resilient Intrusion Tolerance through Proactive and Reactive Recovery*

Paulo Sousa Alysson Neves Bessani Miguel Correia
Nuno Ferreira Neves Paulo Verissimo
LASIGE, Faculdade de Ciências da Universidade de Lisboa – Portugal
{pjsousa, bessani, mpc, nuno, pjv}@di.fc.ul.pt

Abstract

The SecureRing group communication protocols provide reliable ordered message delivery and group membership services despite Byzantine faults such as might be caused by modifications to the programs of a group member following illicit access to, or capture of, a group member. The

processors within an asynchronous distributed syst
pose a consistent total order on messages, and
consistent group memberships.

The approach adopted by SecureRing to prote
Byzantine faults is to optimize the performance
mal (fault-free) operation and to pay a performanc

Practical Byzantine Fault Tolerance and Proactive Recovery

MIGUEL CASTRO
Microsoft Research
and
BARBARA LISKOV
MIT Laboratory for Computer Science

A Qualitative Analysis of the Intrusion-Tolerance Capabilities of the MAFTIA Architecture

Robert Stroud, Ian Welch[1], John Warne, Peter Ryan,
School of Computing Science, University of Newcastle upon Tyne, UK
{R.J.Stroud, J.P.Warne, Peter.Ryan}@ncl.ac.uk
Ian.Welch@mcs.vuw.ac.nz

Worm-IT – A wormhole-based intrusion-tolerant group communication system

Miguel Correia [a,*], Nuno Ferreira Neves [a], Lau Cheuk Lung [b], Paulo Veríssimo [a]

[a] *Faculdade de Ciências da Universidade de Lisboa, Departamento de Informática, Campo Grande, Bloco C6, Piso 3, 1749-016 Lisboa, Portugal*
[b] *Programa de Pós-Graduação em Informática Aplicada, Pontifícia Universidade Católica do Paraná, Rua Imaculada Conceição, 1155, 80.215-901, Brazil*

Received 26 October 2005; received in revised form 28 March 2006; accepted 30 March 2006

State Transfer for Hypervisor-Based Proactive Recovery of Heterogeneous Replicated Services

Tobias Distler Rüdiger Kapitza
Friedrich-Alexander University
Erlangen-Nuremberg, Germany
{distler,rrkapitz}@cs.fau.de

Hans P. Reiser
LASIGE
Universidade de Lisboa, Portu
hans@di.fc.ul.pt

Published 2011

- Fixed number of replicas
- Periodic recoveries

Prior Work on Intrusion-Tolerant Systems

The Rampart Toolkit for Building High-Integrity Services

Michael K. Reiter

AT&T Bell Laboratories, Holmdel, New Jersey, USA
reiter@research.att.com

The SecureRing Protocols for Securing Group Communica

Kim Potter Kihlstrom, L. E. Moser, P. M. Melliar-Smi
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
kimk@alpha.ece.ucsb.edu, moser@ece.ucsb.edu, pmms@ece.ucsb.

Abstract

The SecureRing group communication protocols provide reliable ordered message delivery and group membership services despite Byzantine faults such as might be caused by modifications to the programs of a group member following illicit access to, or capture of, a group member. The

processors within an asynchronous distributed syst
pose a consistent total order on messages, and
consistent group memberships.

The approach adopted by SecureRing to protect
Byzantine faults is to optimize the performance
mal (fault-free) operation and to pay a performance

An architecture for adaptive intrusion-tolerant applications

Partha Pal[1,*] and Paul Rubel[1], Michael Atighetchi[1], William H. Sanders[2], Mouna Seri[2], HariGovind Rama Tod Courtney[3], Adnan Agbaria[2], Michel Cukier[3], Jea

[1] *BBN Technologies, Cambridge, Massachusetts. {ppal, prubel, adnan}@crhc.uiuc.edu*
[2] *University of Illinois at Urbana-Champaign. {whs, seri, rama*
[3] *University of Maryland at College Park, Maryland. mcukier@*
Company. jeanna.m.gossett@MW.Boeing.com [5] *Department of Electrical Engineering, Technion – Israel Institute of Technology. idish@technion.ac.il*

State Transfer for Hypervisor-Based Proactive Recovery of Heterogeneous Replicated Services

Tobias Distler Rüdiger Kapitza
Friedrich-Alexander University
Erlangen-Nuremberg, Germany
{distler,rrkapitz}@cs.fau.de

Hans P. Reiser
LASIGE
Universidade de Lisboa, Portugal
hans@di.fc.ul.pt

Resilient Intrusion Tolerance through Proactive and Reactive Recovery*

Paulo Sousa Alysson Neves Bessani Miguel Correia
Nuno Ferreira Neves Paulo Verissimo
LASIGE, Faculdade de Ciências da Universidade de Lisboa – Portugal
{pjsousa, bessani, mpc, nuno, pjv}@di.fc.ul.pt

Practical Byzantine Fault Tolerance and Proactive Recovery

MIGUEL CASTRO
Microsoft Research
and
BARBARA LISKOV
MIT Laboratory for Computer Science

A Qualitative Analysis of the Intrusion-Tolerance Capabilities of the MAFTIA Architecture

Robert Stroud, Ian Welch[1], John Warne, Peter Ryan,
School of Computing Science, University of Newcastle upon Tyne, UK
{R.J.Stroud, J.P.Warne, Peter.Ryan}@ncl.ac.uk
Ian.Welch@mcs.vuw.ac.nz

Worm-IT – A wormhole-based intrusion-tolerant group communication system

Miguel Correia [a,*], Nuno Ferreira Neves [a], Lau Cheuk Lung [b], Paulo Veríssimo [a]

[a] *Faculdade de Ciências da Universidade de Lisboa, Departamento de Informática, Campo Grande, Bloco C6, Piso 3, 1749-016 Lisboa, Portugal*
[b] *Programa de Pós-Graduação em Informática Aplicada, Pontifícia Universidade Católica do Paraná, Rua Imaculada Conceição, 1155, 80.215-901, Brazil*

Received 26 October 2005; received in revised form 28 March 2006; accepted 30 March 2006

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

Amy Babay*, Thomas Tantillo*, Trevor Aron, Marco Platania, and Yair Amir
Johns Hopkins University — {babay, tantillo, taron1, yairamir}@cs.jhu.edu
AT&T Labs — {platania}@research.att.com
Spread Concepts LLC — {yairamir}@spreadconcepts.com

Published 2018

- Fixed number of replicas
- Periodic recoveries

Prior Work on Intrusion-Tolerant Systems

The Rampart Toolkit for Building High-Integrity Services

Michael K. Reiter

AT&T Bell Laboratories, Holmdel, New Jersey, USA
reiter@research.att.com

The SecureRing Protocols for Securing Group Communica

Kim Potter Kihlstrom, L. E. Moser, P. M. Melliar-Smi
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
kimk@alpha.ece.ucsb.edu, moser@ece.ucsb.edu, pmms@ece.ucsb.

Abstract

The SecureRing group communication protocols provide reliable ordered message delivery and group membership services despite Byzantine faults such as might be caused by modifications to the programs of a group member following illicit access to, or capture of, a group member. The

processors within an asynchronous distributed syst
pose a consistent total order on messages, and
consistent group memberships.

The approach adopted by SecureRing to protect
Byzantine faults is to optimize the performance
mal (fault-free) operation a

An architecture for adaptive intrusion-tolerant applications

Partha Pal[1,*] and Paul Rubel[1], Michael Atighetchi[1], William H. Sanders[2], Mouna Seri[2], HariGovind Rama Tod Courtney[3], Adnan Agbaria[2], Michel Cukier[3], Jea

[1] *BBN Technologies, Cambridge, Massachusetts. {ppal, prubel,*
[2] *University of Illinois at Urbana-Champaign. {whs, seri, rama adnan}@crhc.uiuc.edu*
[3] *University of Maryland at College Park, Maryland. mcukier@*
Company. jeanna.m.gossett@MW.Boeing.com [5] *Department of Electr*
Technion – Israel Institute of Technology. idish@technion.ac.il

Resilient Intrusion Tolerance through Proactive a

Paulo Sousa Alysson Neves Bessani Migu
Nuno Ferreira Neves Paulo Verissin
LASIGE, Faculdade de Ciências da Universidade de Lisboa – Portugal
{pjsousa, bessani, mpc, nuno, pjv}@di.fc.ul.pt

State Transfer for Hypervisor-Based Proactive Recovery of Heterogeneous Replicated Services

Tobias Distler Rüdiger Kapitza
Friedrich-Alexander University
Erlangen-Nuremberg, Germany
{distler,rrkapitz}@cs.fau.de

Hans P. Reiser
LASIGE
Universidade de Lisboa, Portugal
hans@di.fc.ul.pt

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

Amy Babay*, Thomas Tantillo*, Trevor Aron, Marco Platania, and Yair Amir
Johns Hopkins University — {babay, tantillo, taron1, yairamir}@cs.jhu.edu
AT&T Labs — {platania}@research.att.com
Spread Concepts LLC — {yairamir}@spreadconcepts.com

Practical Byzantine Fault Tolerance and Proactive Recovery

MIGUEL CASTRO
Microsoft Research
and
BARBARA LISKOV
MIT Laboratory for Computer Science

A Qualitative Analysis of the Intrusion-Tolerance Capabilities of the MAFTIA Architecture

Robert Stroud, Ian Welch[1], John Warne, Peter Ryan,
School of Computing Science, University of Newcastle upon Tyne, UK
{R.J.Stroud, J.P.Warne, Peter.Ryan}@ncl.ac.uk
Ian.Welch@mcs.vuw.ac.nz

Worm-IT – A wormhole-based intrusion-tolerant group communication system

Miguel Correia [a,*], Nuno Ferreira Neves [a], Lau Cheuk Lung [b], Paulo Veríssimo [a]

[a] *Faculdade de Ciências da Universidade de Lisboa, Departamento de Informática, Campo Grande, Bloco C6, Piso 3, 1749-016 Lisboa, Portugal*
[b] *Programa de Pós-Graduação em Informática Aplicada, Pontifícia Universidade Católica do Paraná, Rua Imaculada Conceição, 1155, 80.215-901, Brazil*

Received 26 October 2005; received in revised form 28 March 2006; accepted 30 March 2006

Skynet: a Cyber-Aware Intrusion Tolerant Overseer

Tadeu Freitas, João Soares, Manuel E. Correia, Rolando Martins
Department of Computer Science, Faculty of Science, University of Porto
Email:{tadeufreitas, joao.soares, mdcorrei, rmartins}@fc.up.pt

Published 2023

- Fixed number of replicas
- Periodic recoveries

Prior Work on Intrusion-Tolerant Systems

The Rampart Toolkit for Building High-Integrity Services

Michael K. Reiter

AT&T Bell Laboratories, Holmdel, New Jersey, USA
reiter@research.att.com

The SecureRing Protocols for Securing Group Communica

Kim Potter Kihlstrom, L. E. Moser, P. M. Melliar-Smi
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
kimk@alpha.ece.ucsb.edu, moser@ece.ucsb.edu, pmms@ece.ucsb.

Abstract

The SecureRing group communication protocols provide reliable ordered message delivery and group membership services despite Byzantine faults such as might be caused by

processors within an asynchronous distributed syst
pose a consistent total order on messages, and
consistent group memberships.
The approach adopted by SecureRing to protec

An architecture for adaptive intrusion-tolerant applications

Partha Pal[1,*] and Paul Rubel[1], Michael Atighetchi[1], William H. Sanders[2], Mouna Seri[2], HariGovind Rama Tod Courtney[3], Adnan Agbaria[2], Michel Cukier[3], Jea

[1] *BBN Technologies, Cambridge, Massachusetts. {ppal, prubel,*
[2] *University of Illinois at Urbana-Champaign. {whs, seri, rama adnan}@crhc.uiuc.edu*
[3] *University of Maryland at College Park, Maryland. mcukier@*
Company. jeanna.m.gossett@MW.Boeing.com [5] *Department of Electr*
Technion – Israel Institute of Technology. idish@technion.ac.il

Resilient Intrusion Tolerance through Proactive a

Paulo Sousa Alysson Neves Bessani Migu
Nuno Ferreira Neves Paulo Verissin
LASIGE, Faculdade de Ciências da Universidade de Lisboa – Portugal
{pjsousa, bessani, mpc, nuno, pjv}@di.fc.ul.pt

State Transfer for Hypervisor-Based Proactive Recovery of Heterogeneous Replicated Services

Tobias Distler Rüdiger Kapitza
Friedrich-Alexander University
Erlangen-Nuremberg, Germany
{distler,rrkapitz}@cs.fau.de

Hans P. Reiser
LASIGE
Universidade de Lisboa, Portugal
hans@di.fc.ul.pt

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

Amy Babay*, Thomas Tantillo*, Trevor Aron, Marco Platania, and Yair Amir
Johns Hopkins University — {babay, tantillo, taron1, yairamir}@cs.jhu.edu
AT&T Labs — {platania}@research.att.com
Spread Concepts LLC — {yairamir}@spreadconcepts.com

Can we do better by applying **our methodology** for optimal security response?

A Qualitative Analysis of the Intrusion-Tolerance Capabilities of the MAFTIA Architecture

Robert Stroud, Ian Welch[1], John Warne, Peter Ryan,
School of Computing Science, University of Newcastle upon Tyne, UK
{R.J.Stroud, J.P.Warne, Peter.Ryan}@ncl.ac.uk
Ian.Welch@mcs.vuw.ac.nz

Worm-IT – A wormhole-based intrusion-tolerant group communication system

Miguel Correia [a,*], Nuno Ferreira Neves [a], Lau Cheuk Lung [b], Paulo Veríssimo [a]

[a] *Faculdade de Ciências da Universidade de Lisboa, Departamento de Informática, Campo Grande, Bloco C6, Piso 3, 1749-016 Lisboa, Portugal*
[b] *Programa de Pós-Graduação em Informática Aplicada, Pontifícia Universidade Católica do Paraná, Rua Imaculada Conceição, 1155, 80.215-901, Brazil*

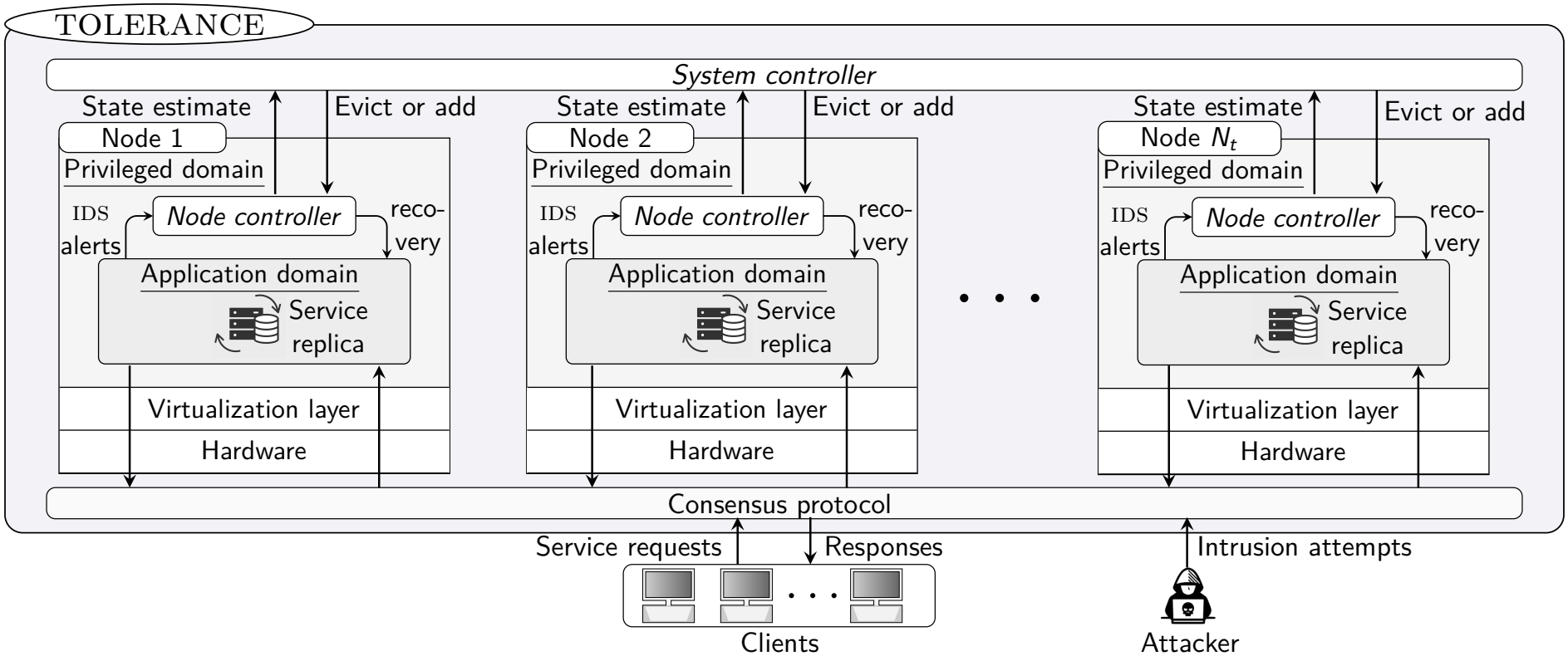
Received 26 October 2005; received in revised form 28 March 2006; accepted 30 March 2006

Published 2023

- Fixed number of replicas
- Periodic recoveries

The TOLERANCE Architecture

Tw**o**-**le**vel **r**ecovery **an**d replication **c**ontrol with f**e**edback.



Definition (**Correct service**)

*The system provides **correct service** if the healthy replicas satisfy the following properties:*

Each request is eventually executed. (Liveness)

Each executed request was sent by a client. (Validity)

Each replica executes the same request sequence. (Safety)

Proposition (**Correctness of TOLERANCE**)

A system that implements the TOLERANCE *architecture* ***provides correct service*** *if*

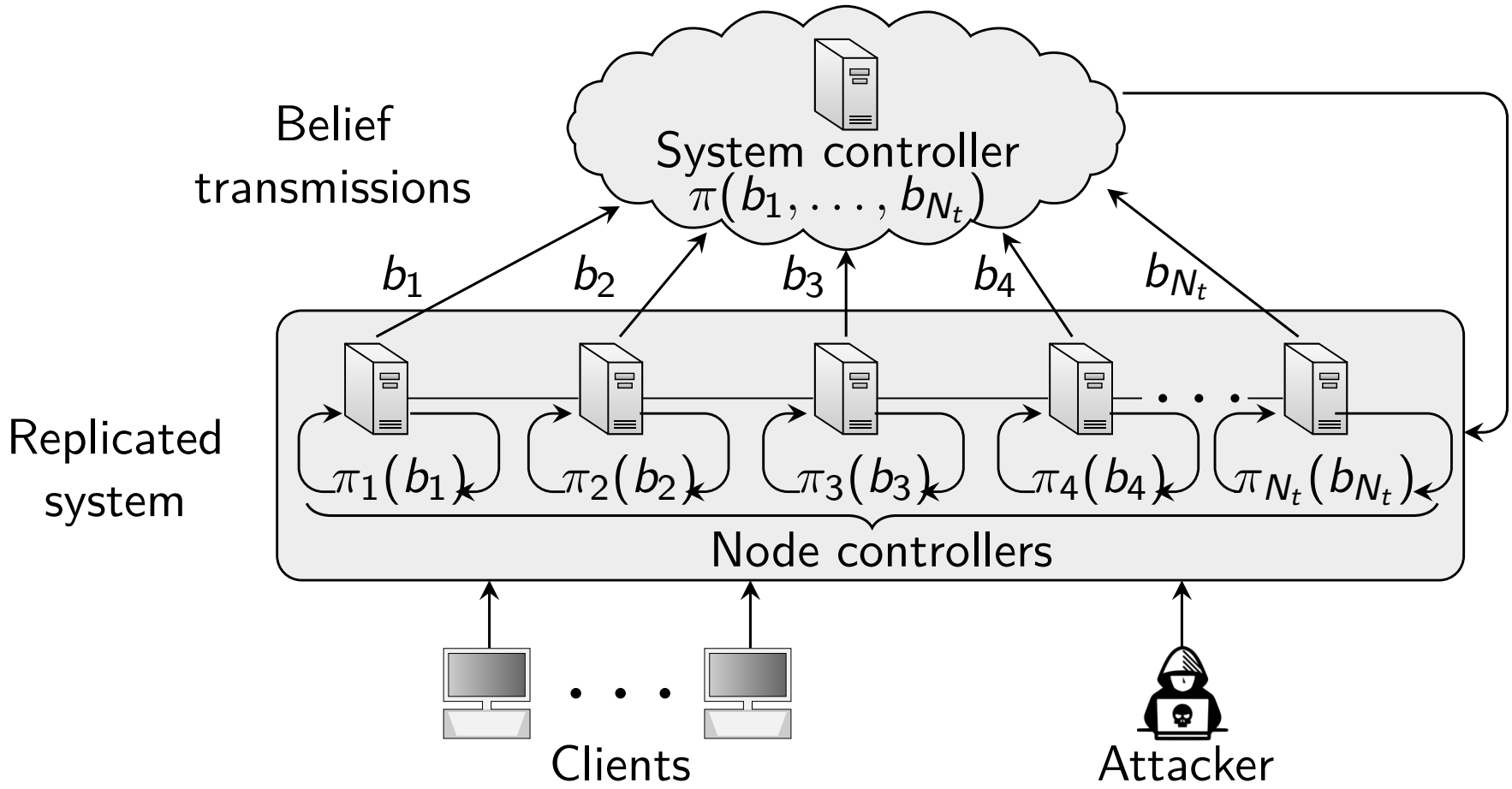
Network links are authenticated.

At most f nodes are compromised or crashed simultaneously.

$N_t \geq 2f + 1$.

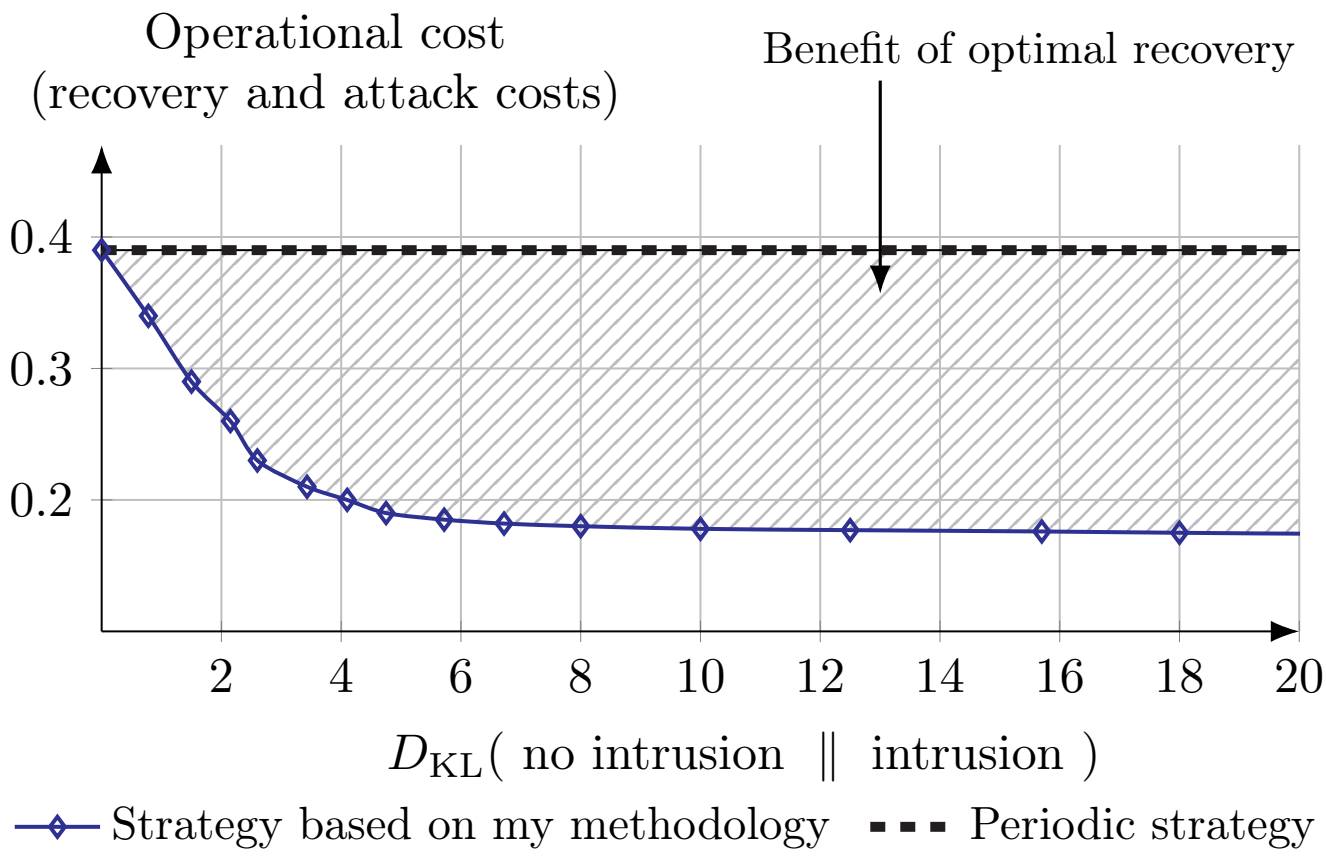
The system is partially synchronous.

Intrusion Tolerance as a **Two-Level Game**



- We formulate intrusion tolerance as a two-level game.
- The local game models intrusion recovery.
- The global game models replication control.

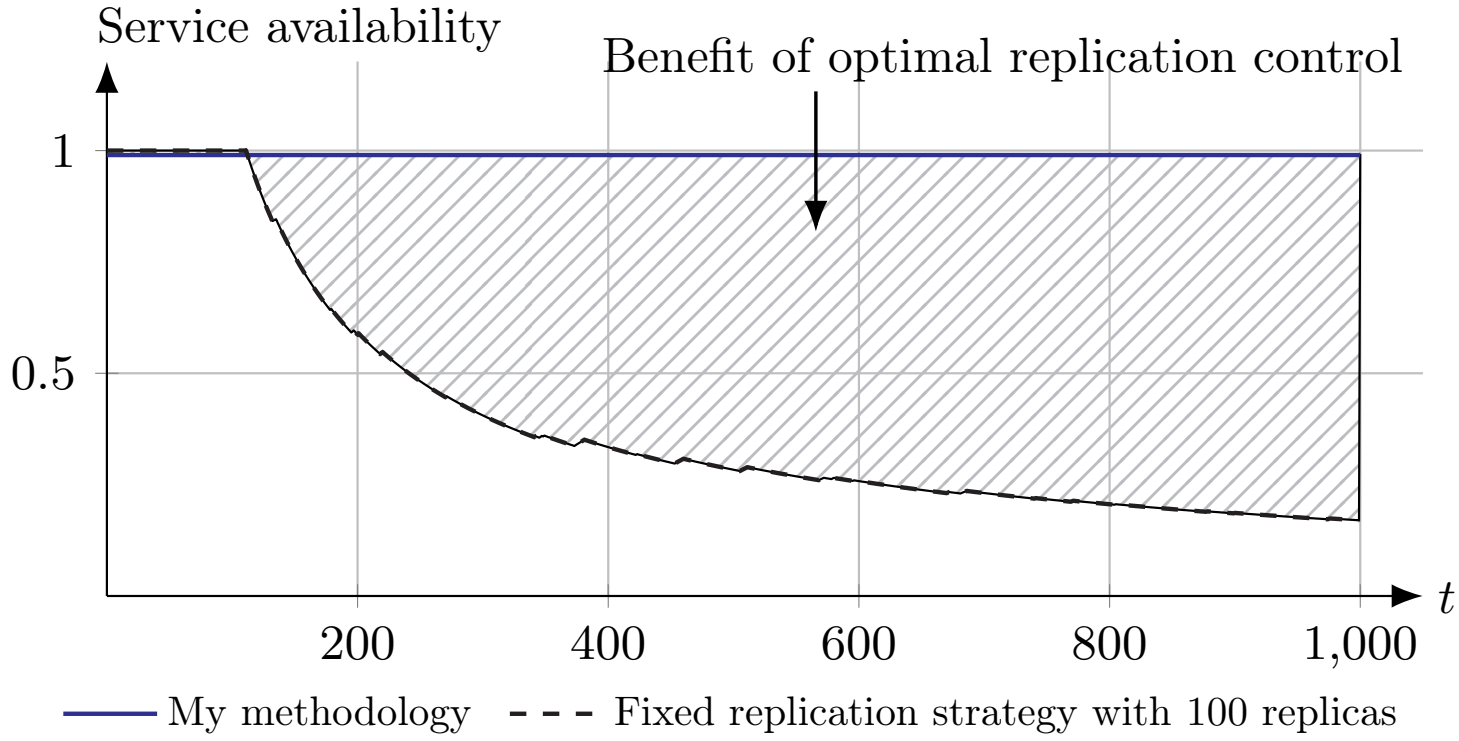
The Benefit of Optimal Recovery



Key insight

Optimal recovery can significantly reduce operational cost given that an **intrusion detection model is available**.

The Benefit of Optimal Replication



Key insight

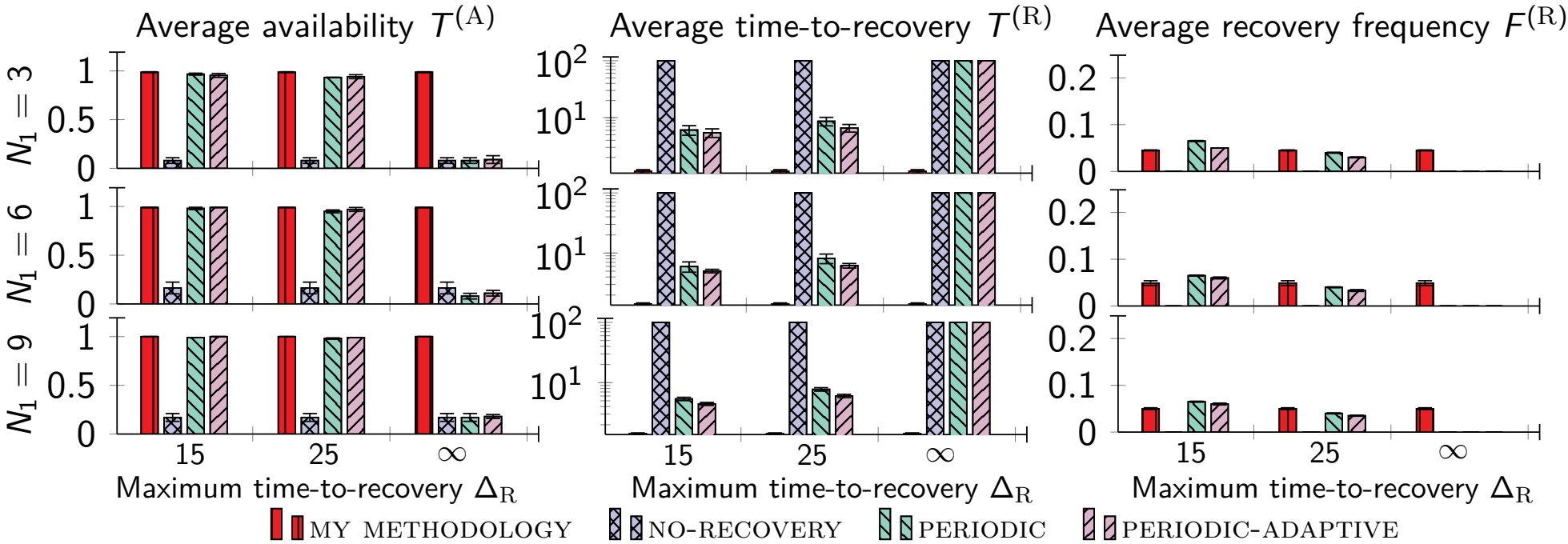
Optimal replication can **guarantee a high service availability in expectation**. The benefit of optimal replication is mainly prominent for long-running systems.

Experimental Evaluation

Experiment Setup - Emulated Intrusions

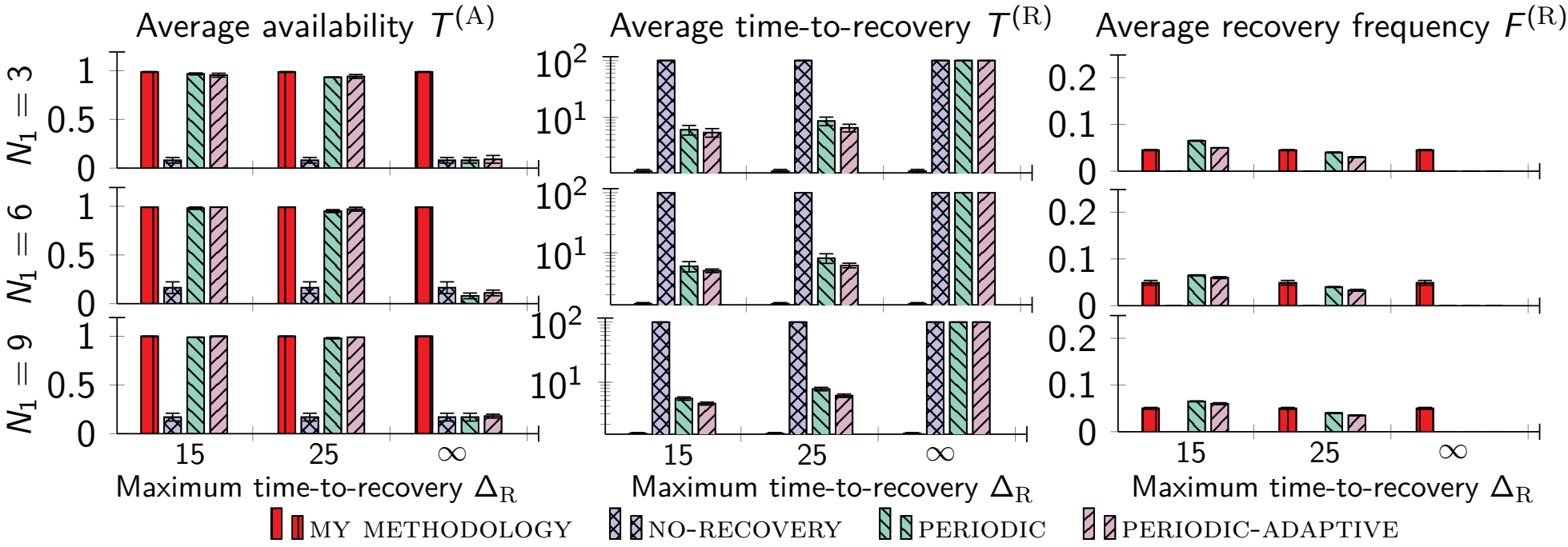
Replica ID	*Intrusion steps*
1	TCP SYN scan, FTP brute force
2	TCP SYN scan, SSH brute force
3	TCP SYN scan, TELNET brute force
4	ICMP scan, exploit of CVE-2017-7494
5	ICMP scan, exploit of CVE-2014-6271
6	ICMP scan, exploit of CWE-89 on DVWA
7	ICMP scan, exploit of CVE-2015-3306
8	ICMP scan, exploit of CVE-2016-10033
9	ICMP scan, SSH brute force, exploit of CVE-2010-0426
10	ICMP scan, SSH brute force, exploit of CVE-2015-5602

Comparison with State-of-the-art Systems



Comparison between the control strategies produced by my methodology and the baselines; Δ_R is the maximum allowed time-to-recovery; N_1 is the number of initial nodes.

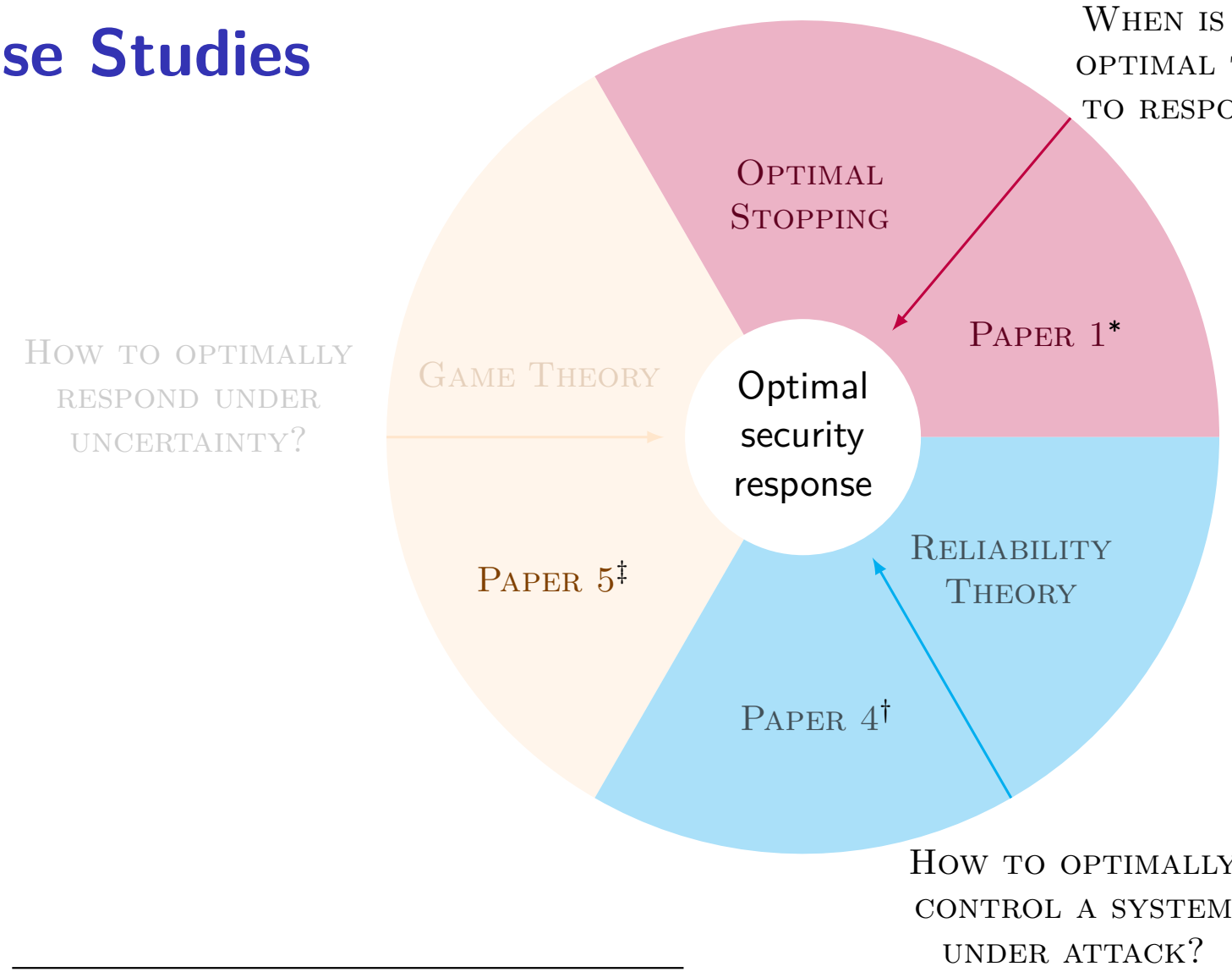
Comparison with State-of-the-art Systems



What's new here?

First intrusion-tolerant system that ensures a chosen level of service availability while minimizing operational cost.

Case Studies

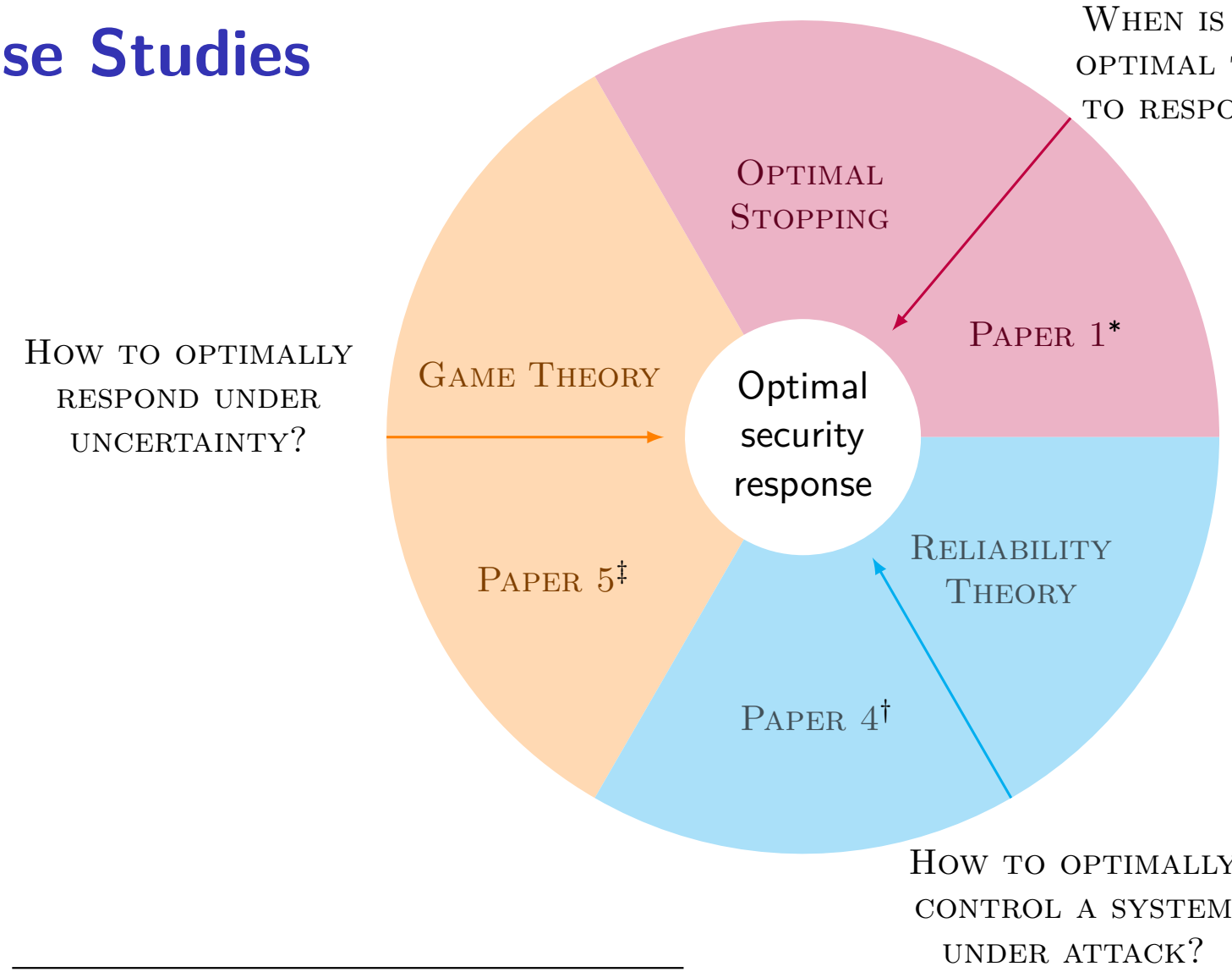


*Kim Hammar and Rolf Stadler. "Intrusion Prevention through Optimal Stopping". In: *IEEE Transactions on Network and Service Management* 19.3 (2022), pp. 2333–2348. DOI: 10.1109/TNSM.2022.3176781.

†Kim Hammar and Rolf Stadler. "Intrusion Tolerance for Networked Systems through Two-Level Feedback Control". In: *2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. 2024, pp. 338–352. DOI: 10.1109/DSN58291.2024.00042.

‡Kim Hammar et al. *Automated Security Response through Online Learning with Adaptive Conjectures*. https://arxiv.org/abs/2402.12499. To appear in IEEE Transactions on Information Forensics and Security

Case Studies

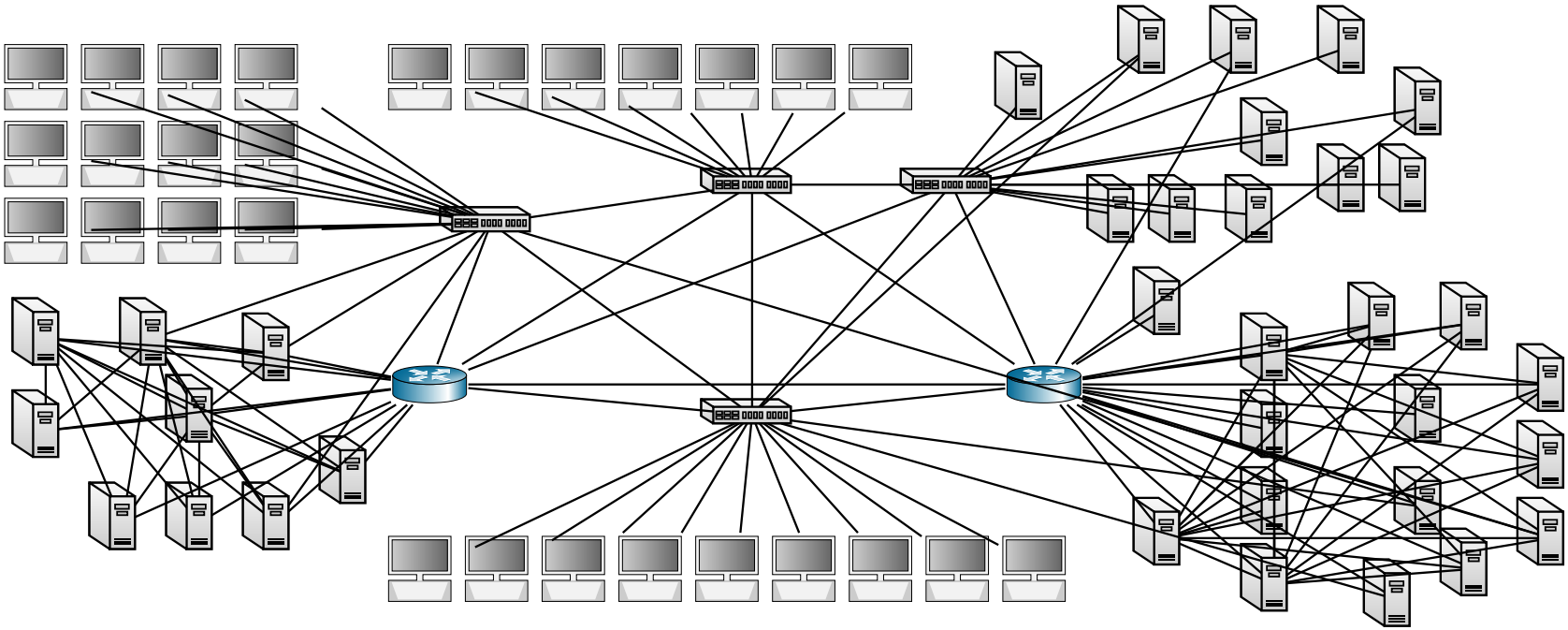


*Kim Hammar and Rolf Stadler. "Intrusion Prevention through Optimal Stopping". In: *IEEE Transactions on Network and Service Management* 19.3 (2022), pp. 2333–2348. DOI: 10.1109/TNSM.2022.3176781.

†Kim Hammar and Rolf Stadler. "Intrusion Tolerance for Networked Systems through Two-Level Feedback Control". In: *2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. 2024, pp. 338–352. DOI: 10.1109/DSN58291.2024.00042.

‡Kim Hammar et al. *Automated Security Response through Online Learning with Adaptive Conjectures*. https://arxiv.org/abs/2402.12499. To appear in IEEE Transactions on Information Forensics and Security

Challenge: IT systems are complex.



- It is not realistic that any model will capture all the details.
  - $\Longrightarrow$ We have to work with **approximate models**.
  - $\Longrightarrow$ model misspecification.
- How does misspecification affect optimality and convergence?

Prior Work

- Assumes a stationary model with no misspecification.
  - Limitation: fails to capture many real-world systems.
- Focuses on offline computation of defender strategies.
  - Limitation: computationally intractable for realistic models.

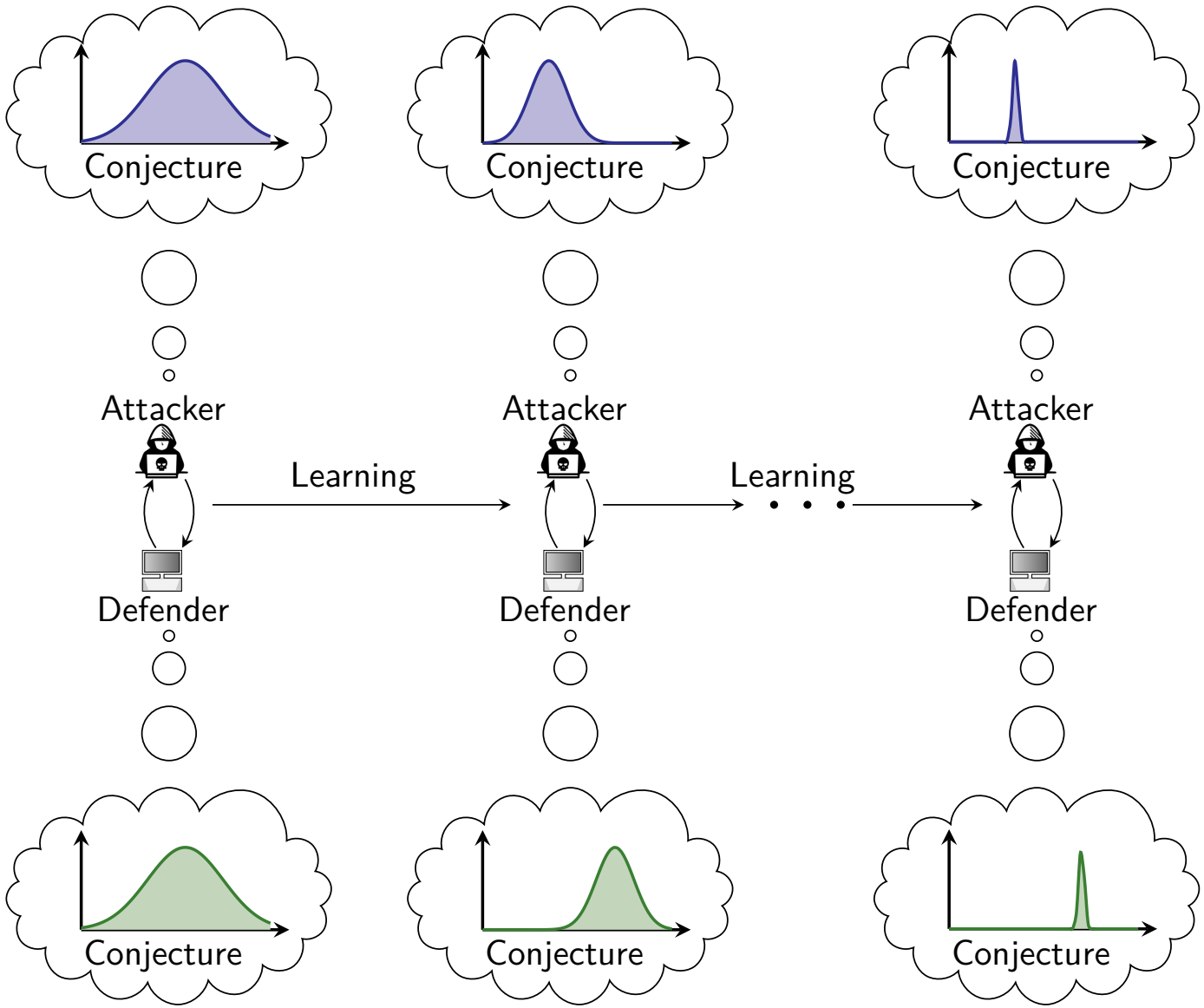
Prior Work

- Assumes a stationary model with no misspecification.
  - Limitation: fails to capture many real-world systems.
- Focuses on offline computation of defender strategies.
  - Limitation: computationally intractable for realistic models.

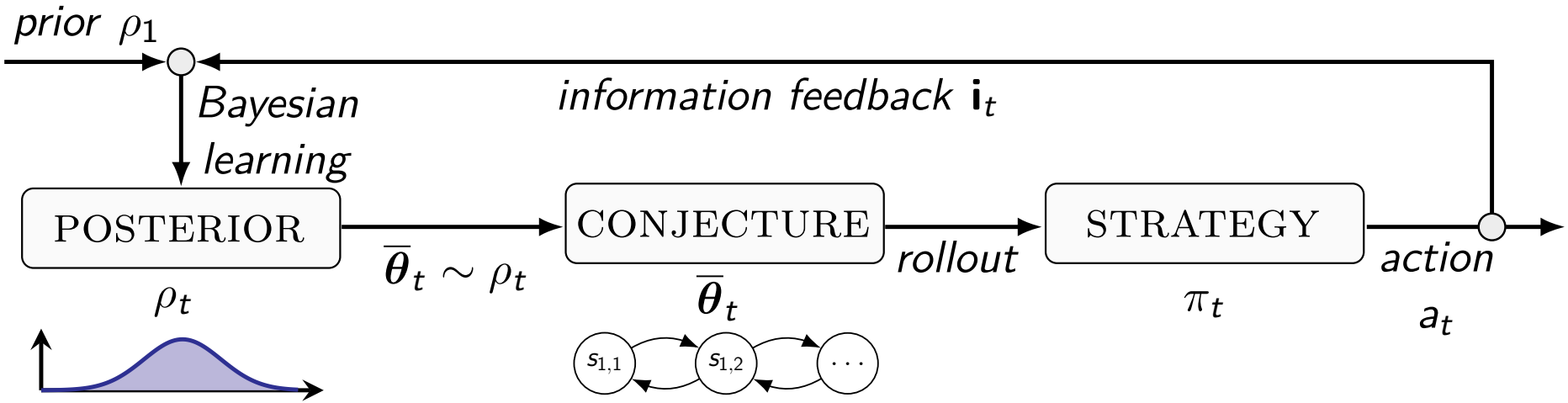
Problem

What if the model is misspecified and non-stationary?

Method: Conjectural Online Learning

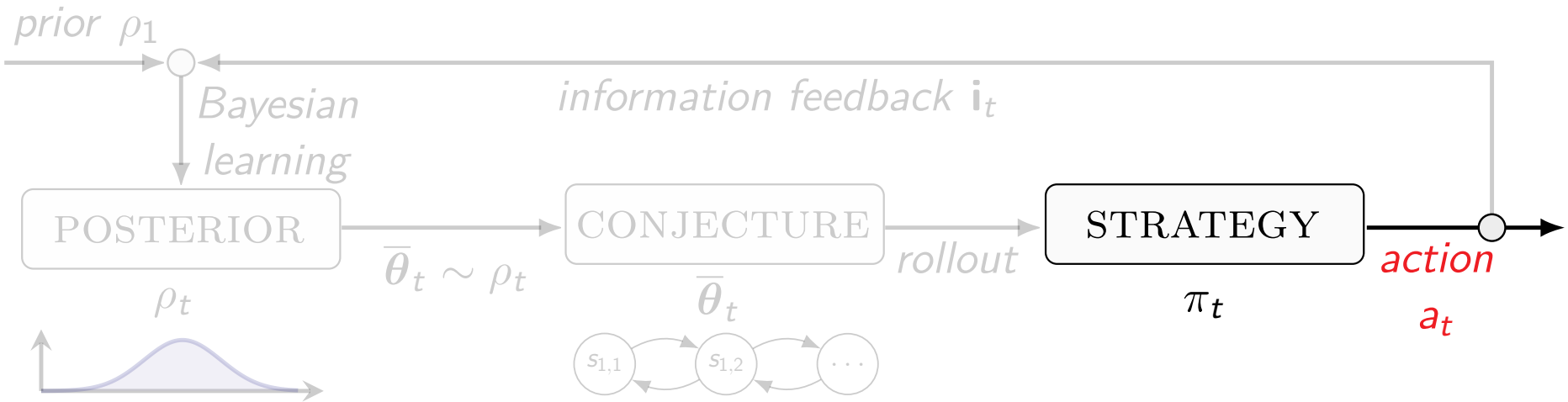


Method: Conjectural Online Learning



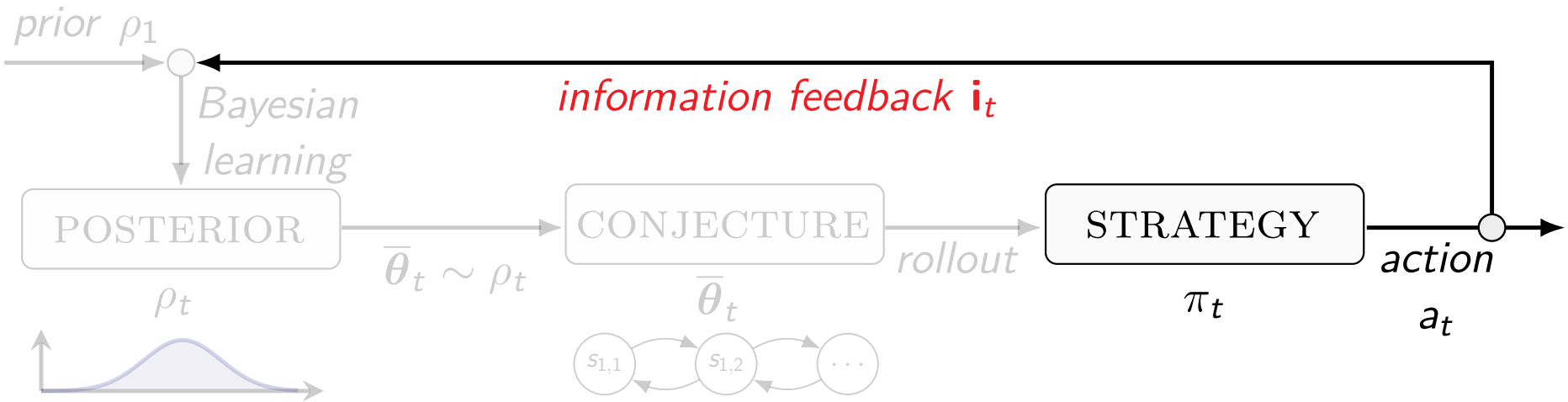
- The **model parameters** are denoted by $\boldsymbol{\theta}$.
- The defender has a conjecture $\overline{\boldsymbol{\theta}} \sim \rho_t \in \Delta(\Theta)$.
- The defender is misspecified if $\boldsymbol{\theta} \notin \Theta$.

Method: Conjectural Online Learning



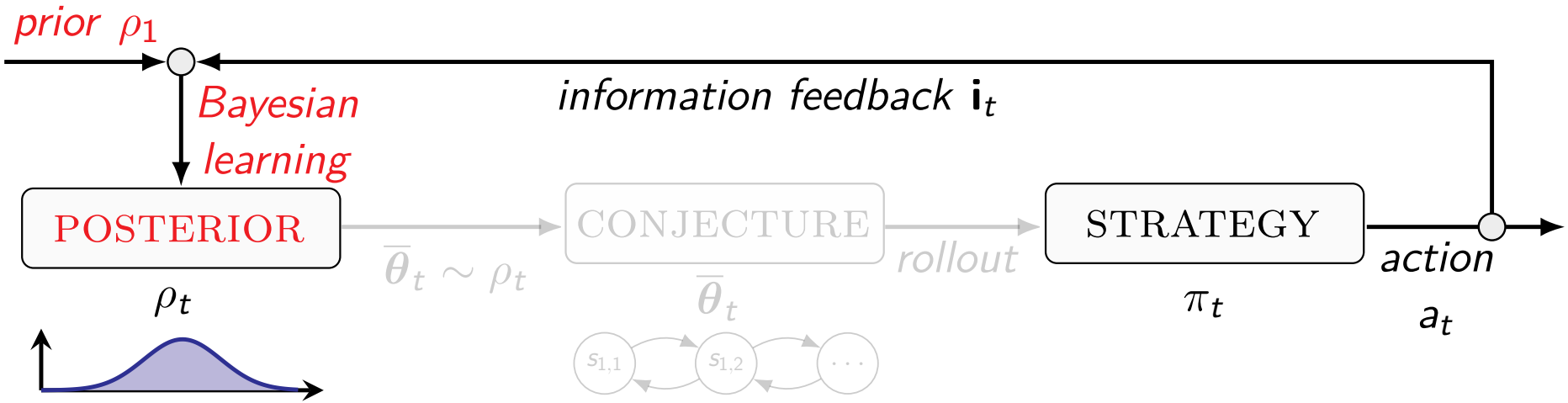
- The **model parameters** are denoted by $\boldsymbol{\theta}$.
- The defender has a conjecture $\overline{\boldsymbol{\theta}} \sim \rho_t \in \Delta(\Theta)$.
- The defender is misspecified if $\boldsymbol{\theta} \notin \Theta$.

Method: Conjectural Online Learning



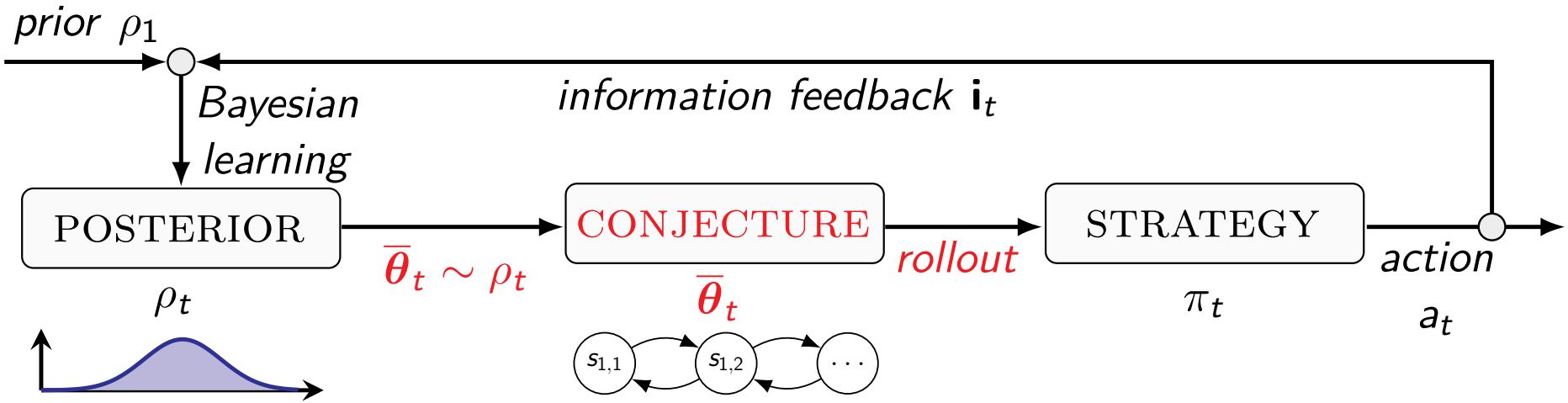
- The **model parameters** are denoted by $\boldsymbol{\theta}$.
- The defender has a conjecture $\overline{\boldsymbol{\theta}} \sim \rho_t \in \Delta(\Theta)$.
- The defender is misspecified if $\boldsymbol{\theta} \notin \Theta$.

Method: Conjectural Online Learning



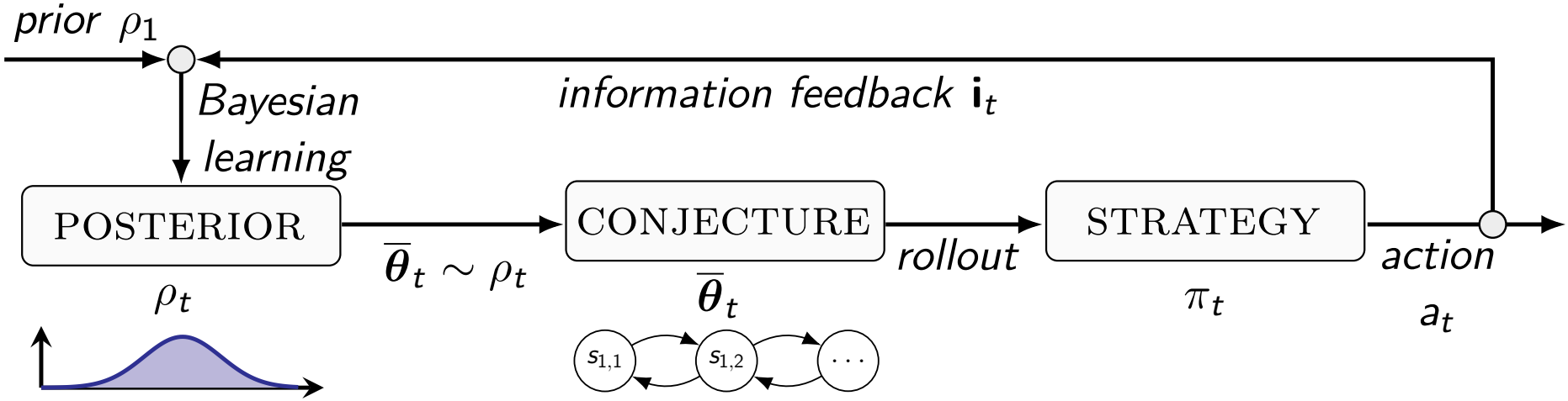
- The **model parameters** are denoted by $\boldsymbol{\theta}$.
- The defender has a conjecture $\overline{\boldsymbol{\theta}} \sim \rho_t \in \Delta(\Theta)$.
- The defender is misspecified if $\boldsymbol{\theta} \notin \Theta$.

Method: Conjectural Online Learning



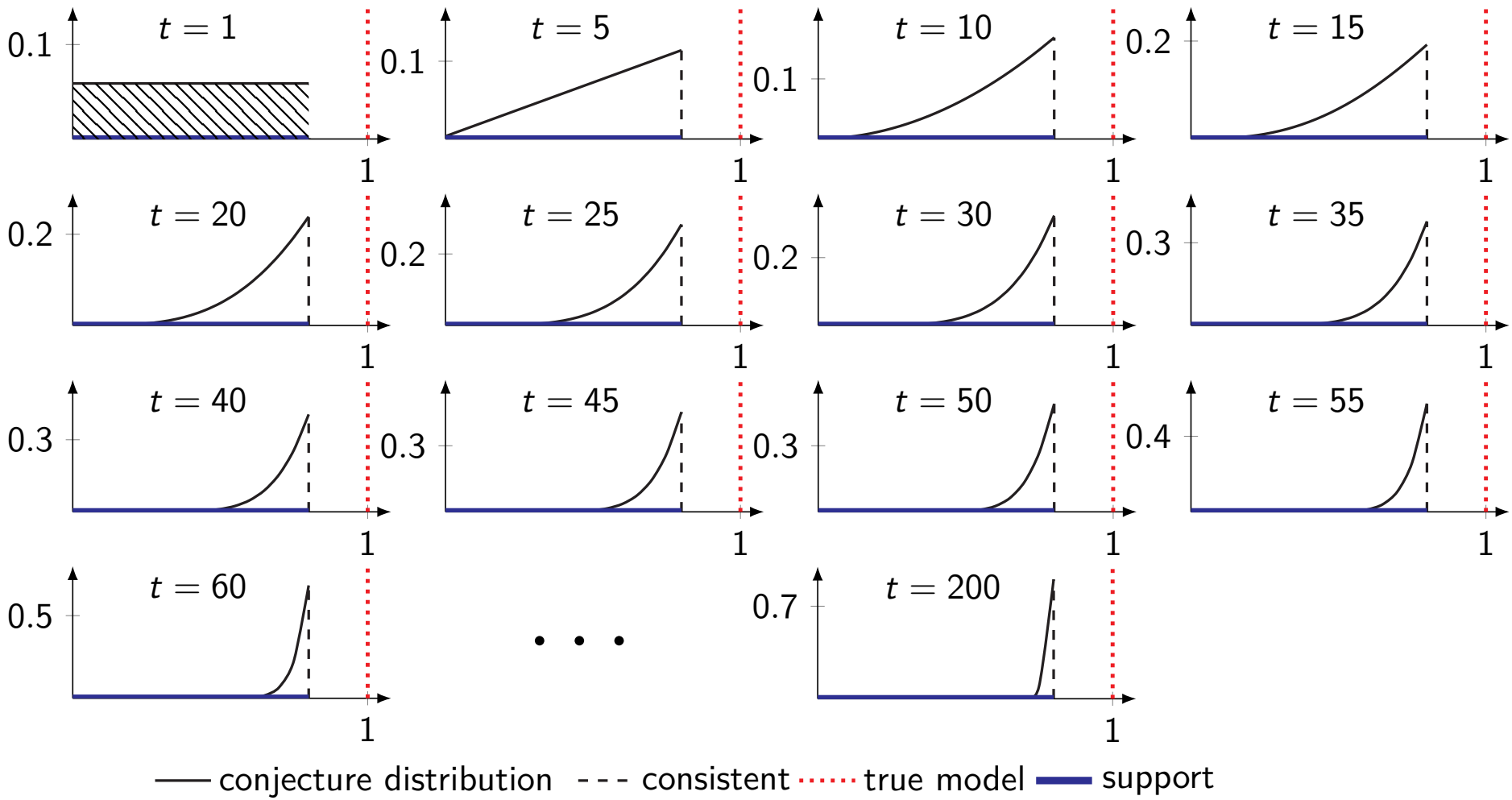
- The **model parameters** are denoted by $\boldsymbol{\theta}$.
- The defender has a conjecture $\overline{\boldsymbol{\theta}} \sim \rho_t \in \Delta(\Theta)$.
- The defender is misspecified if $\boldsymbol{\theta} \notin \Theta$.

Method: Conjectural Online Learning

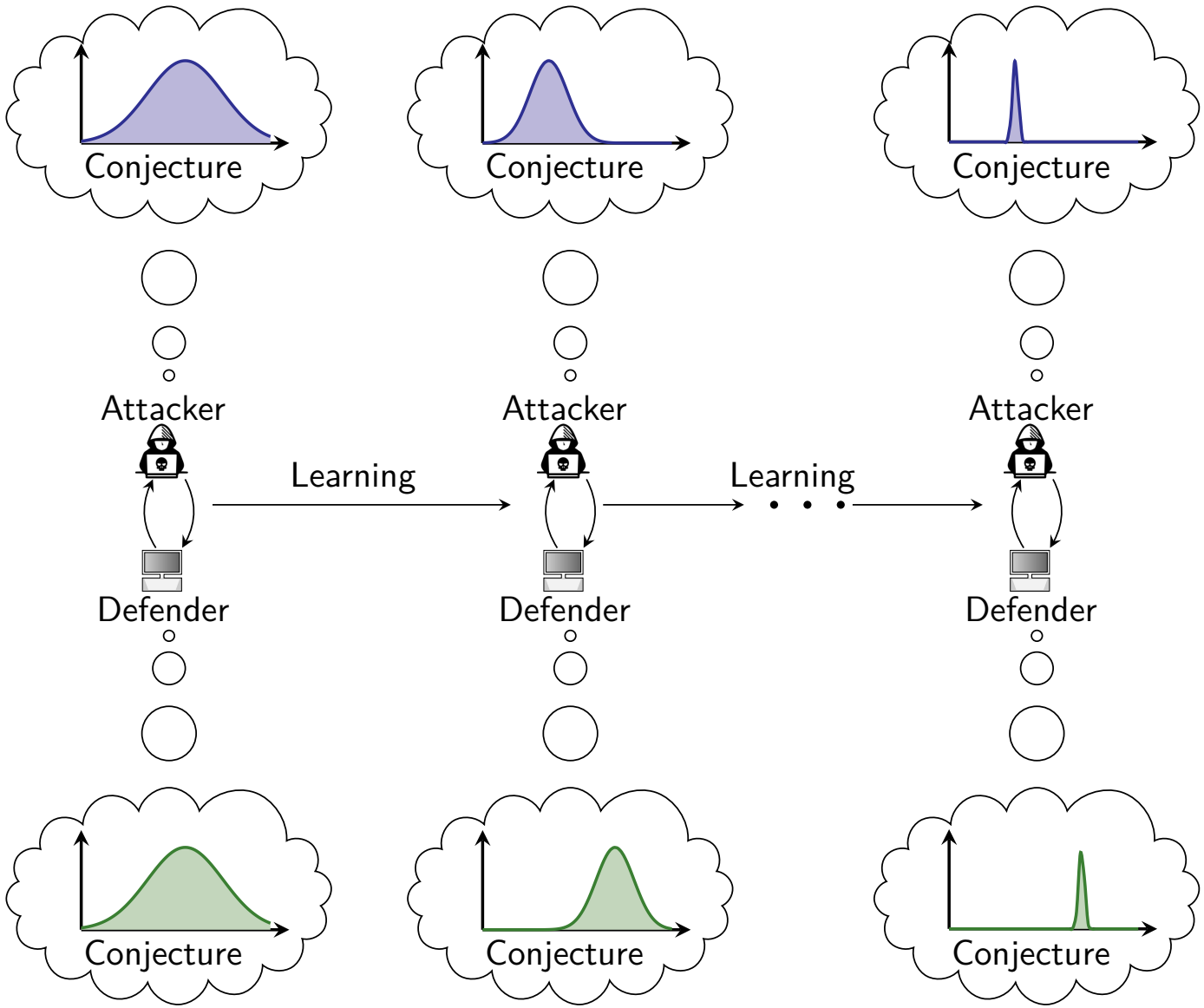


- Theorem 5.3: performance *improvement bound* of COL.
- Theorem 5.4: *asymptotic consistency* of COL.

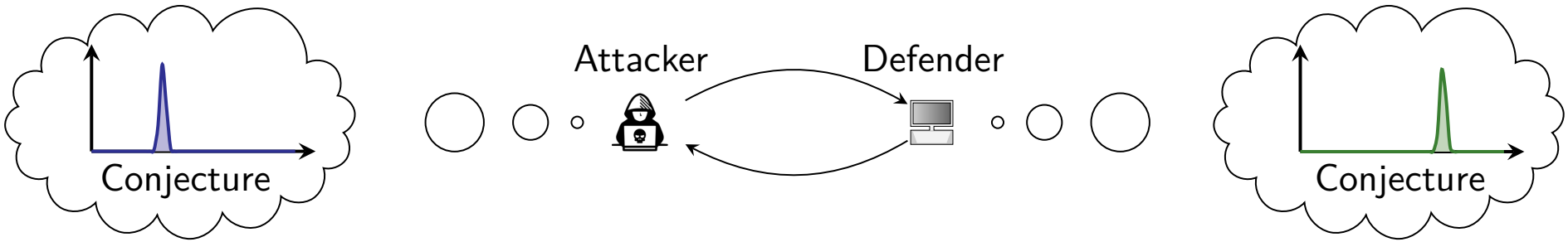
COL Converges to Consistent Conjectures



COL Converges to Consistent Conjectures



The Berk-Nash Equilibrium



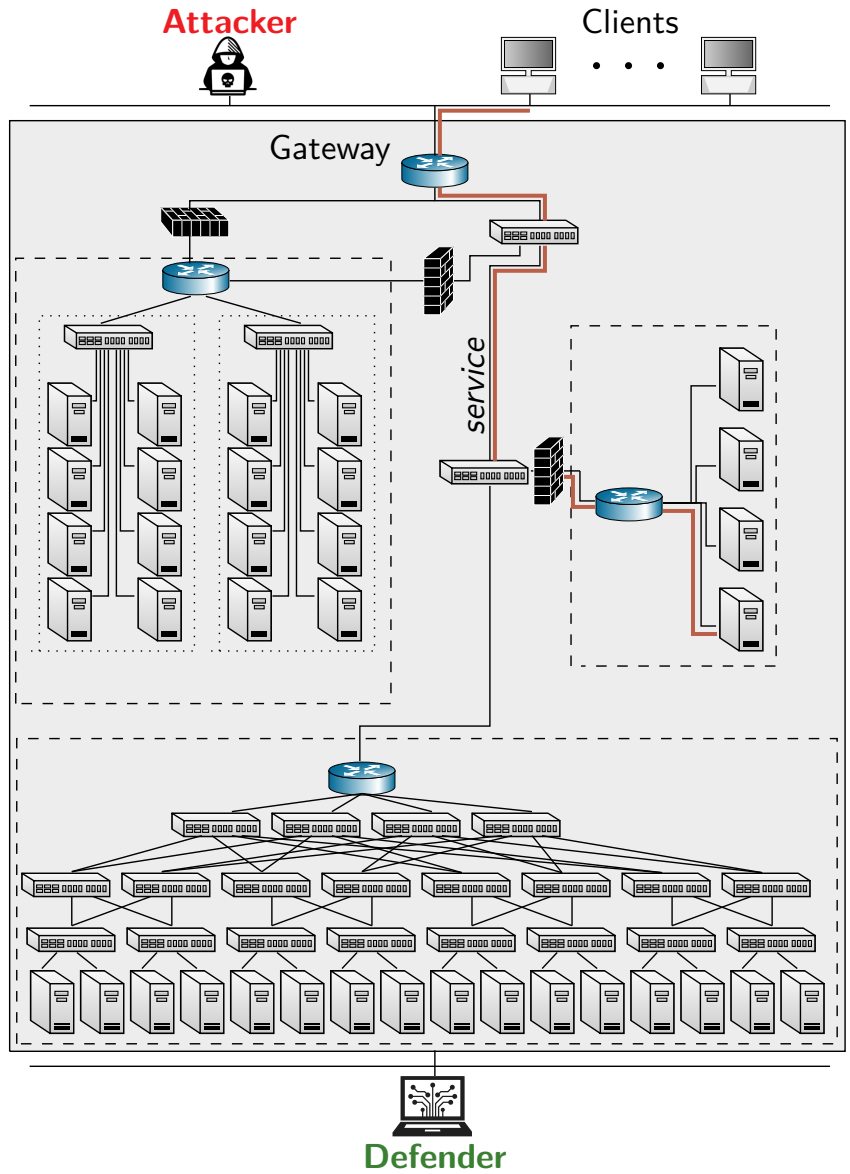
Definition (Berk-Nash Equilibrium (Informal))

A strategy profile $\boldsymbol{\pi}$ and an occupancy measure $\nu \in \Delta(\mathcal{B})$ is a Berk-Nash equilibrium iff

1. NASH. π_{k} *is a best response against* $\pi_{-\mathrm{k}}$.
2. BERK. *Each player's conjecture is consistent.*
3. STATIONARITY. ν *is a limit point of* COL.

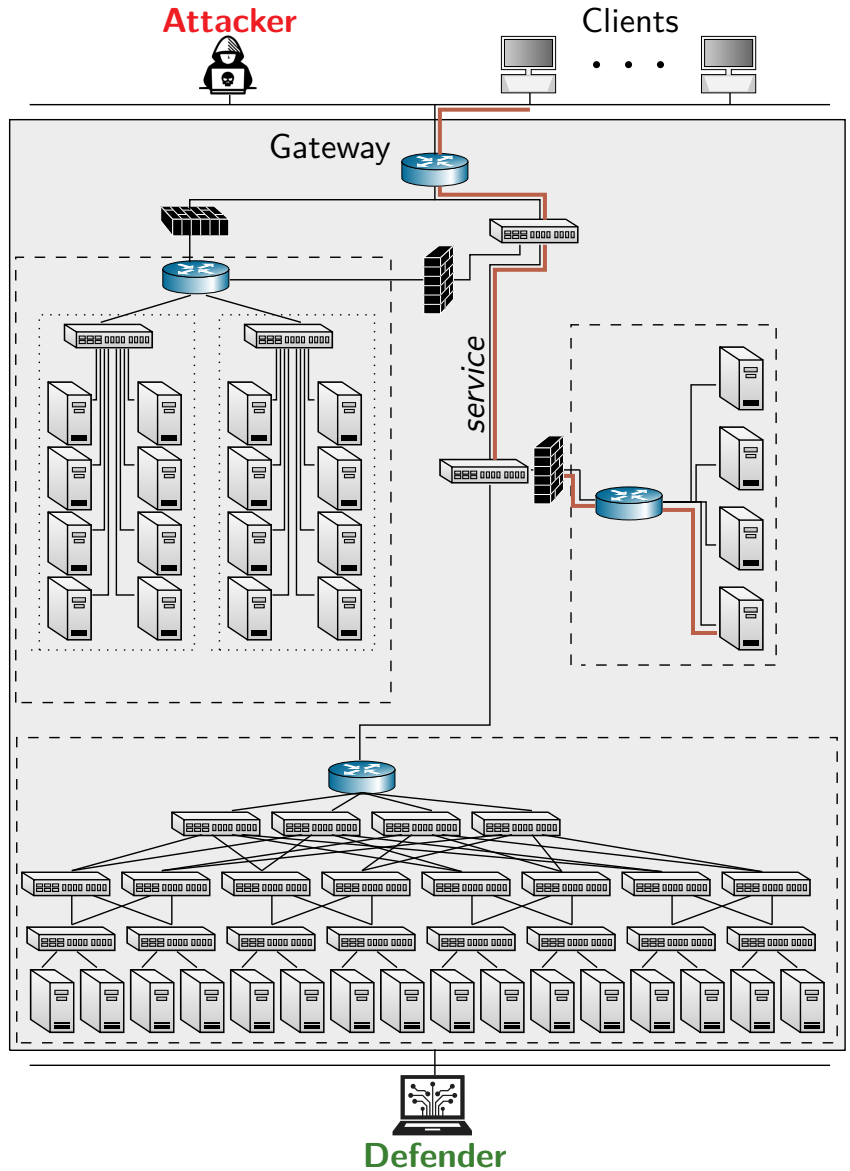
Experimental Evaluation

- Model $\boldsymbol{\theta}$: distribution of security alerts.
- Defender: controls the blocking threshold.

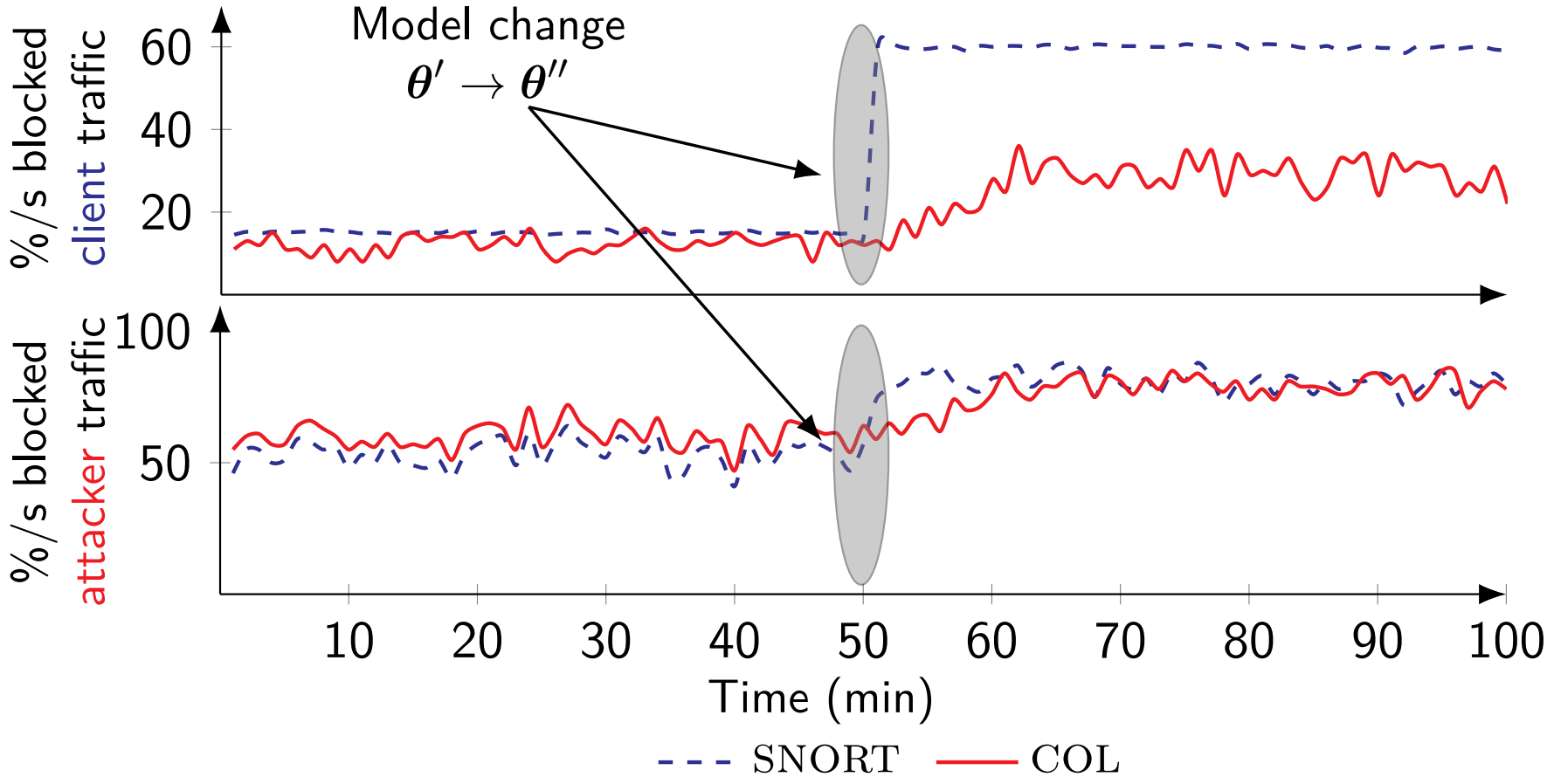


Experimental Evaluation

- Model $\boldsymbol{\theta}$: distribution of security alerts.
- Defender: controls the blocking threshold.
- **Baseline**: SNORT
  - A rule-based intrusion prevention system.

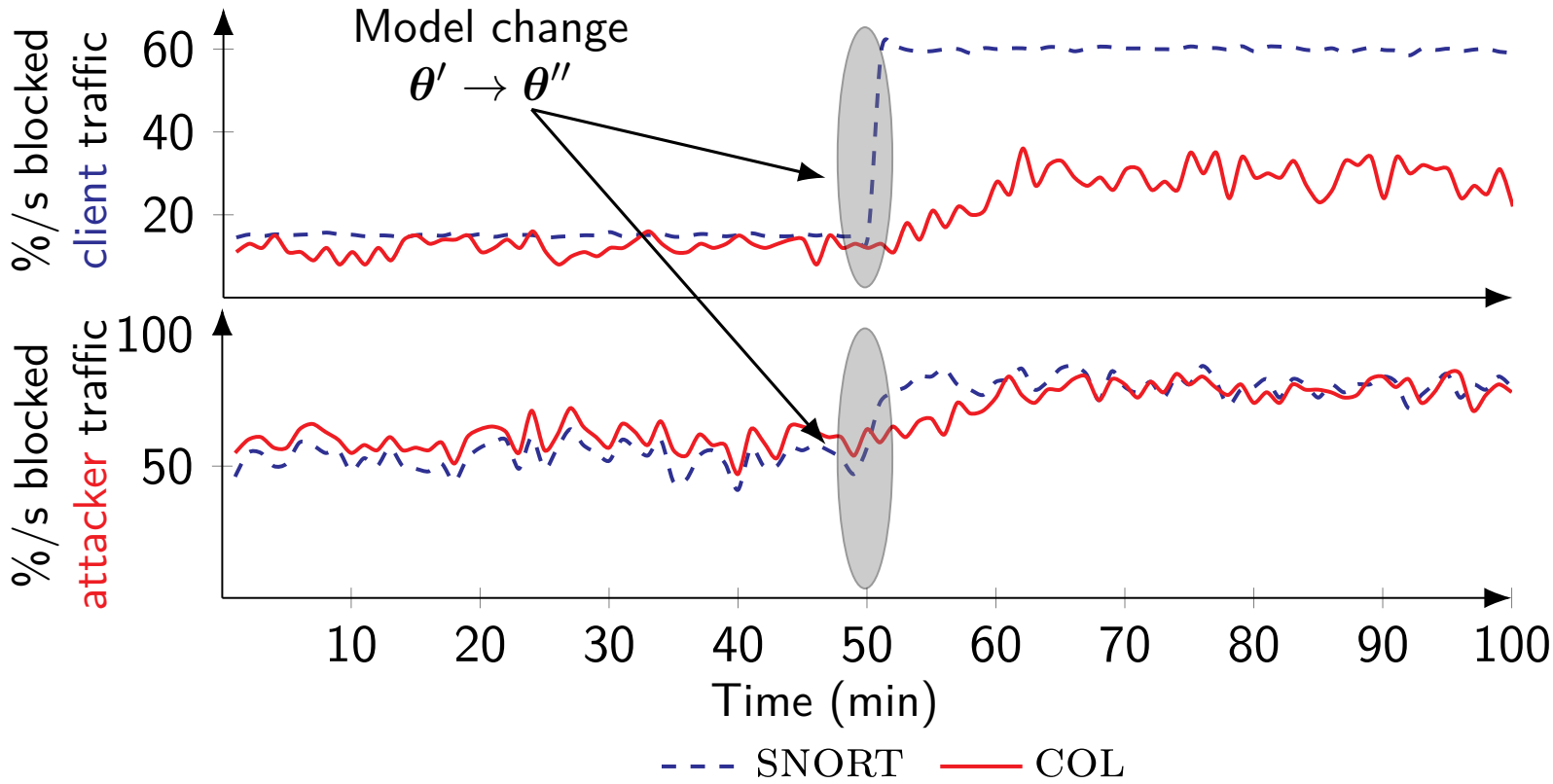


Comparison with an Industry Standard



- θ represents distributions of intrusion detection alerts.

Comparison with an Industry Standard



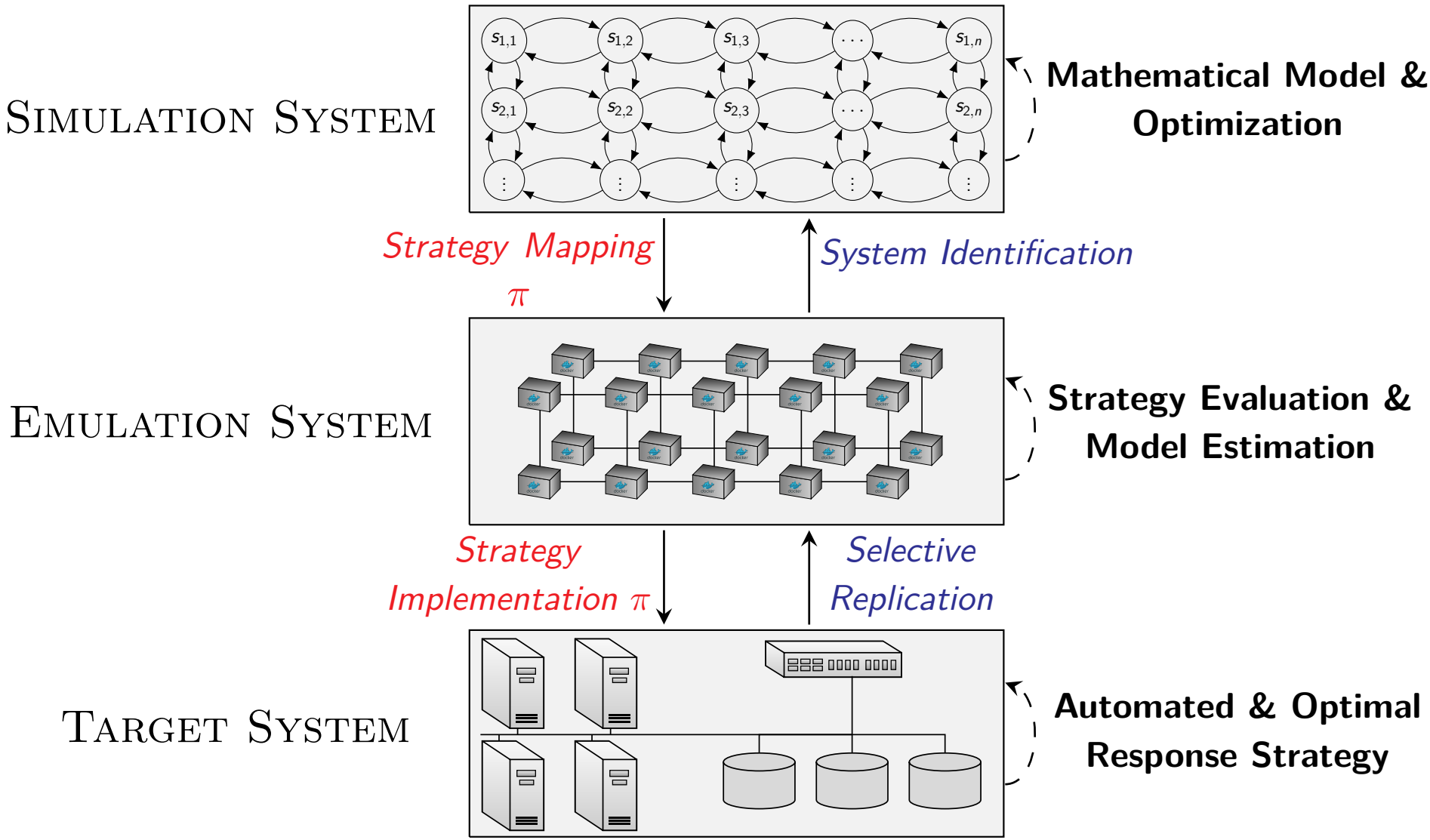
What's new here?

COL provides a higher level of automation than SNORT.

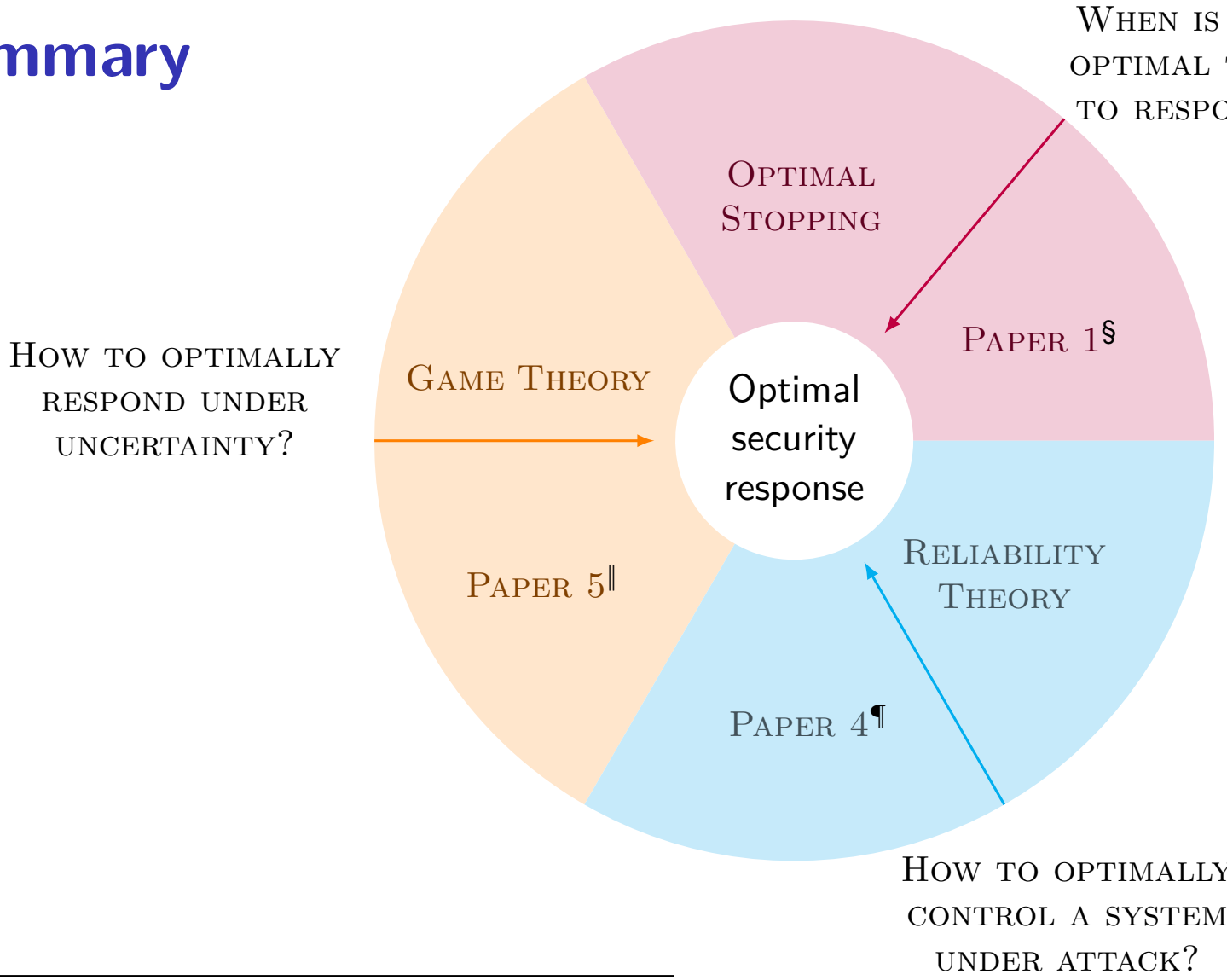
Summary

How to achieve **automated** and optimal security response?

Summary


Simulation System
$s_{1,1}$ $s_{1,2}$ $s_{1,3}$ $\cdots$ $s_{1,n}$
$s_{2,1}$ $s_{2,2}$ $s_{2,3}$ $\cdots$ $s_{2,n}$
Mathematical Model & Optimization
Strategy Mapping π
System Identification
Emulation System
Strategy Evaluation & Model Estimation
Strategy Implementation π
Selective Replication
Target System
Automated & Optimal Response Strategy

Summary

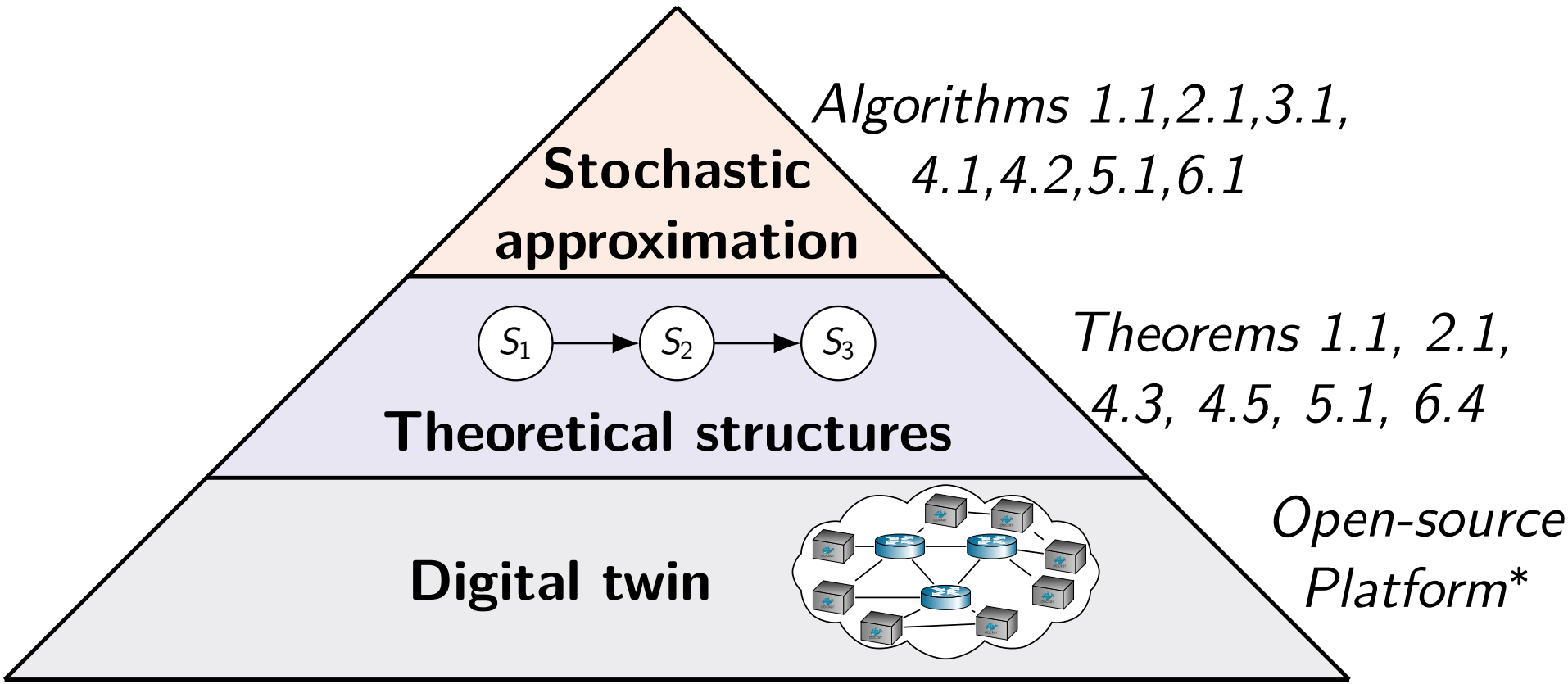


§Kim Hammar and Rolf Stadler. "Intrusion Prevention through Optimal Stopping". In: *IEEE Transactions on Network and Service Management* 19.3 (2022), pp. 2333–2348. DOI: 10.1109/TNSM.2022.3176781.

¶Kim Hammar and Rolf Stadler. "Intrusion Tolerance for Networked Systems through Two-Level Feedback Control". In: *2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. 2024, pp. 338–352. DOI: 10.1109/DSN58291.2024.00042.

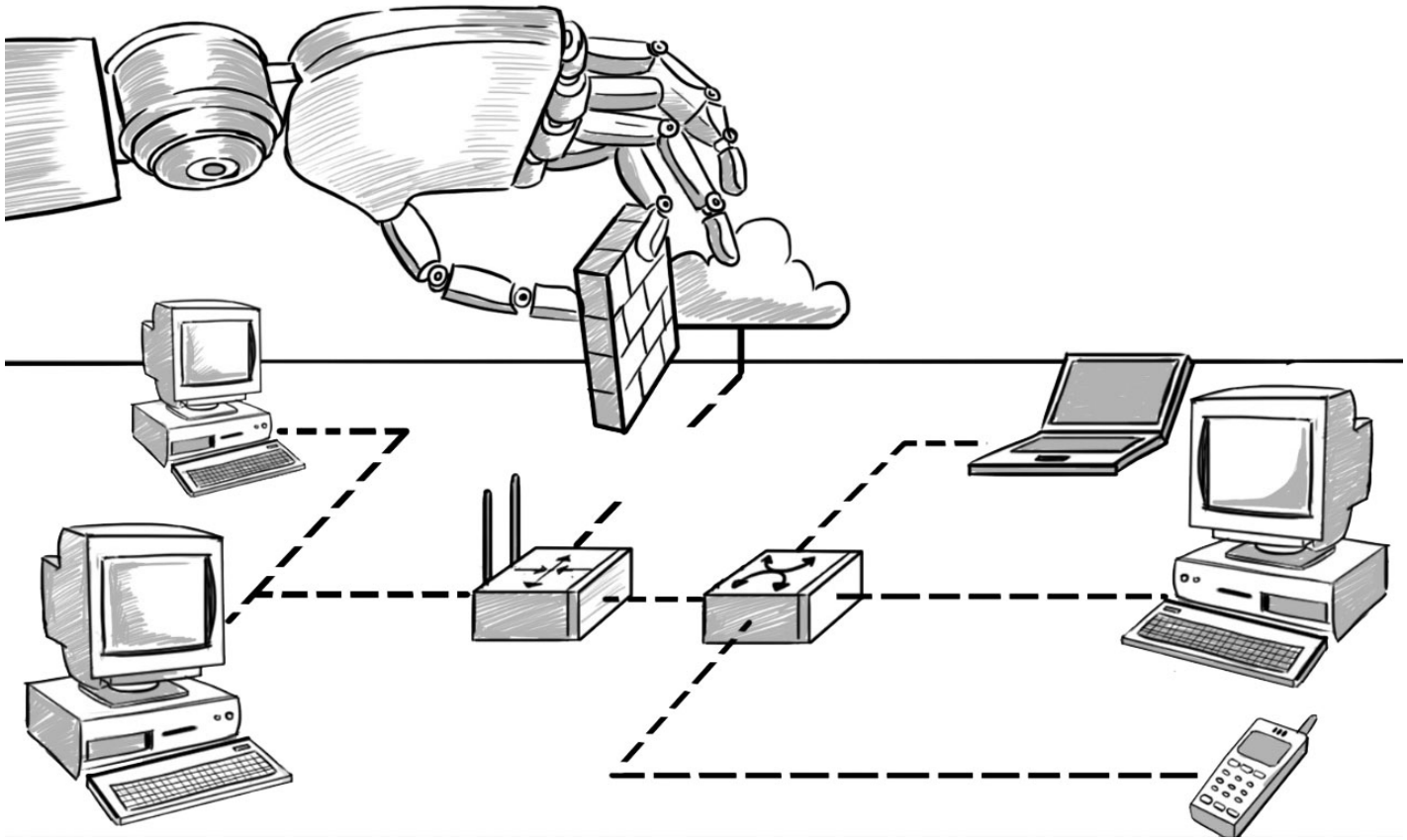
‖Kim Hammar et al. *Automated Security Response through Online Learning with Adaptive Conjectures*. https://arxiv.org/abs/2402.12499. To appear in IEEE Transactions on Information Forensics and Security

Key Elements for Optimal Security Response



*Kim Hammar. *Cyber Security Learning Environment (CSLE)*. Documentation: https://limmen.dev/csle/, traces: https://github.com/Limmen/csle/releases/tag/v0.4.0, source code: https://github.com/Limmen/csle, video demonstration: https://www.youtube.com/watch?v=iE2KPmtIs2A&. 2023. URL: https://limmen.dev/csle/.

Conclusion



- **Optimal and automated security response** is feasible using a methodology based on
  - engineering principles for self-adaptive systems.
  - mathematical models for numerical computations.