RAPID: Real-Time Alert Investigation with Context-aware Prioritization for Efficient Threat Discovery

Yushan Liu, Xiaokui Shu, Yixin Sun, **Jiyong Jang**, Prateek Mittal
Princeton University, IBM Research, University of Virginia

This project was sponsored by Air Force Research Laboratory (AFRL) and Defense Advanced Research Agency (DARPA) under the award number FA8650-15-C-7561.The views, opinions, and/or findings contained in this article are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. This project was also supported in part by the Commonwealth Cyber Initiative.

We would like to thank the reviewers and our shepherd for their valuable feedback, and the analysts who participated in our survey.

Advanced Persistent Threat (APT)

- 6,000 severe incidents reported in the past decade
- conducted in **multiple stages**, stealthy

HOME SEARCH The New York Times

A Home Depot in New Orleans. Andrew Burton/Getty Images

Home Depot

September 2014

Home Depot said about 56 million payment cards were probably compromised in an attack that ran from April through September and affected stores in the United States and Canada.

BBC Sign in News Sport Weather Shop Earth Travel

NEWS

Home Video World US & Canada UK Business Tech Science Magazine

Sony cyber-attack: North Korea faces new US sanctions

The US has imposed new sanctions on North Korea in response to a cyber-attack against Sony Pictures Entertainment.

President Barack Obama signed an executive order on Friday allowing sanctions on three North Korean organisations and 10 individuals.

The White House said the move was a response to North Korea's "provocative, destabilising, and repressive actions".

Richard Forrest reports.

03 Jan 2015 | US & Canada

REUTERS World Business Markets Politics TV

#BUSINESS NEWS DECEMBER 18, 2013 / 7:05 PM / 4 YEARS AGO

Target cyber breach hits 40 million payment cards at holiday peak

Jim Finkle, Dhanya Skariachan 8 MIN READ

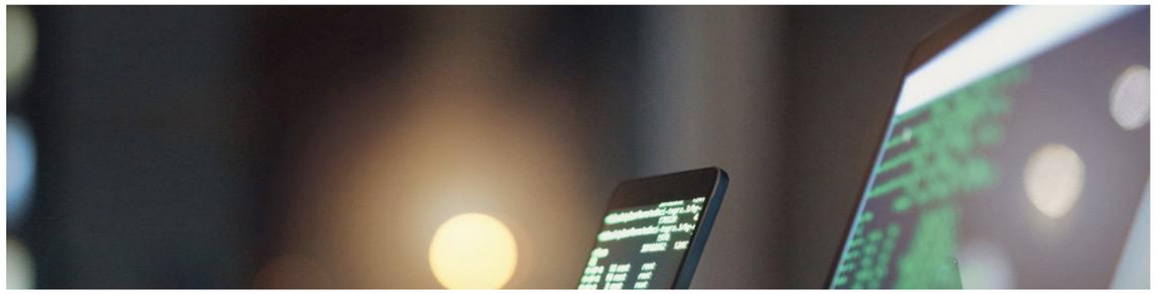
BOSTON (Reuters) - Target Corp said hackers have stolen data from up to 40 million credit and debit cards of shoppers who visited its stores during the first three weeks of the holiday season in the second-largest such breach reported by a U.S. retailer.

Merchandise baskets are lined up outside a Target department store in Palm Coast, Florida, December 9, 2013. REUTERS/Larry Downing

Alert Flooding

- **Security operations center (SOC) is drowning in cybersecurity alerts**
  - more than half see more than 10,000 alerts per day
  - many alerts are false positives or even trivial true positives
  - analysts spend ~30 minutes to investigate each alert

Survey: 27 Percent of IT professionals receive more than 1 million security alerts daily



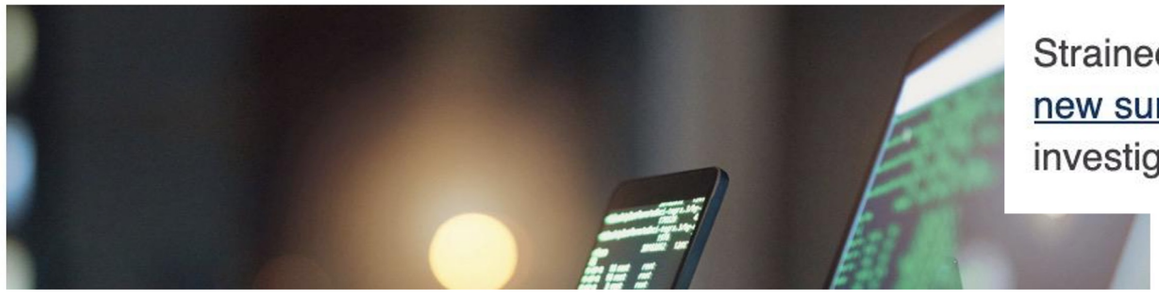
Alert Flooding

- **Security operations center (SOC) is drowning in cybersecurity alerts**
  - fail to respond to true threats within the “golden hour”

Efficient, Automated Alert Triage!

Survey: 27 Percent of IT professionals receive more than 1 million security alerts daily

...



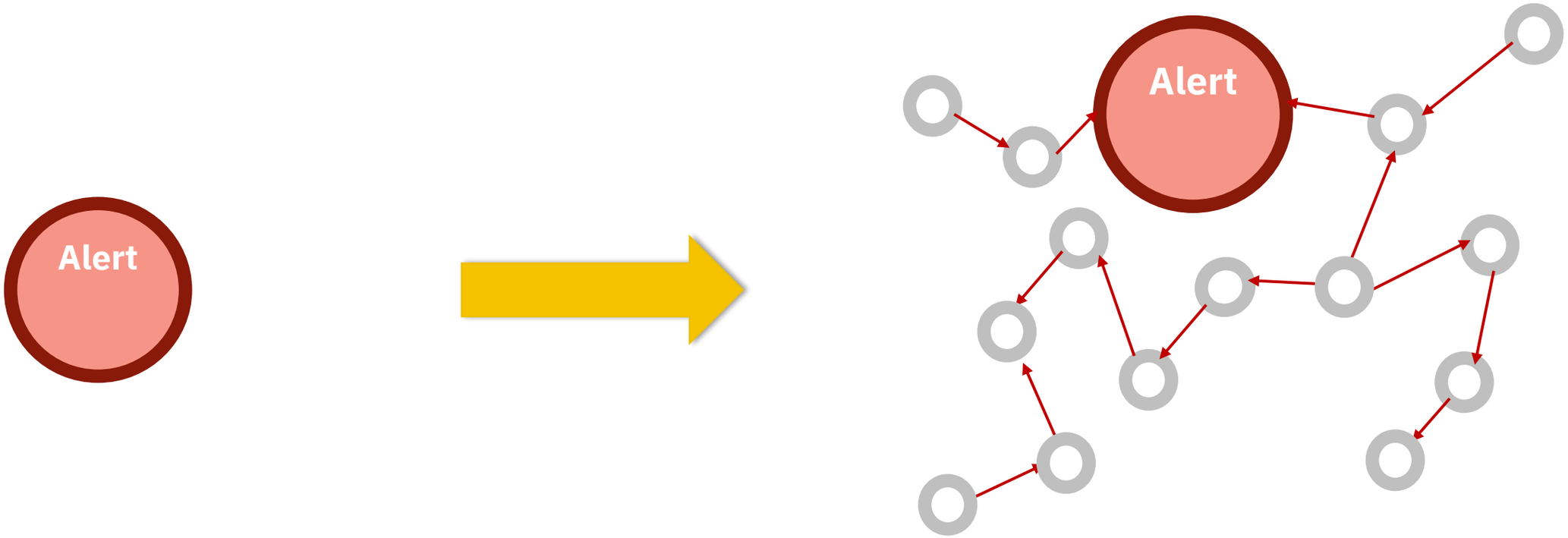
Half of Security Pros Ignore Some Important Alerts

Short-staffed, more than half of organizations admit they ignore alerts that should be investigated because they lack resources to handle the overflow.

Strained by the cybersecurity skills shortage, 54 percent of respondents to a new survey say they are forced to ignore security alerts worthy of further investigation, because they don't have the staff and expertise to handle them

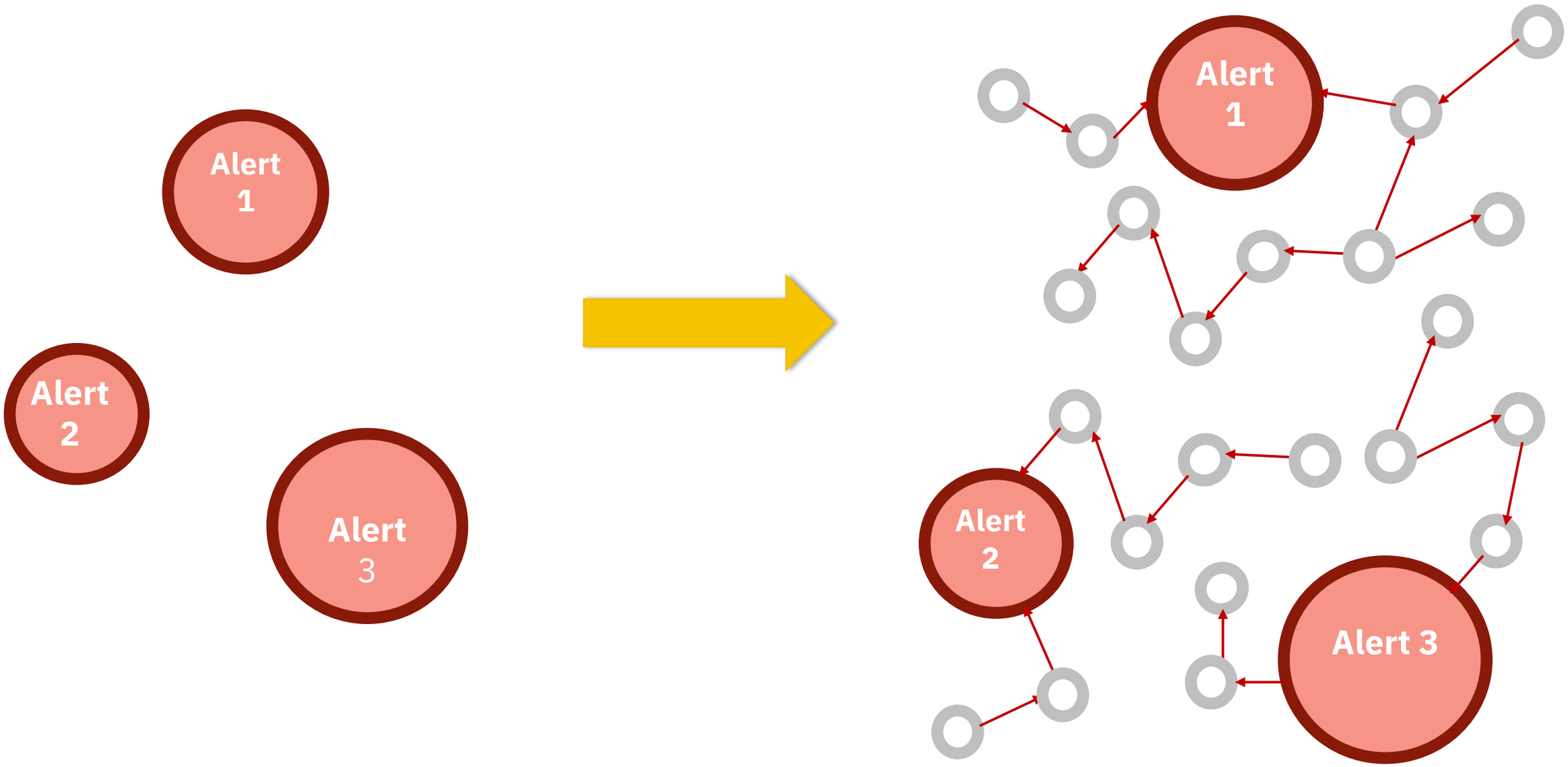
Automated Alert Triage

- **Attack causality analysis**
  - dependency explosion: a large graph around an alert



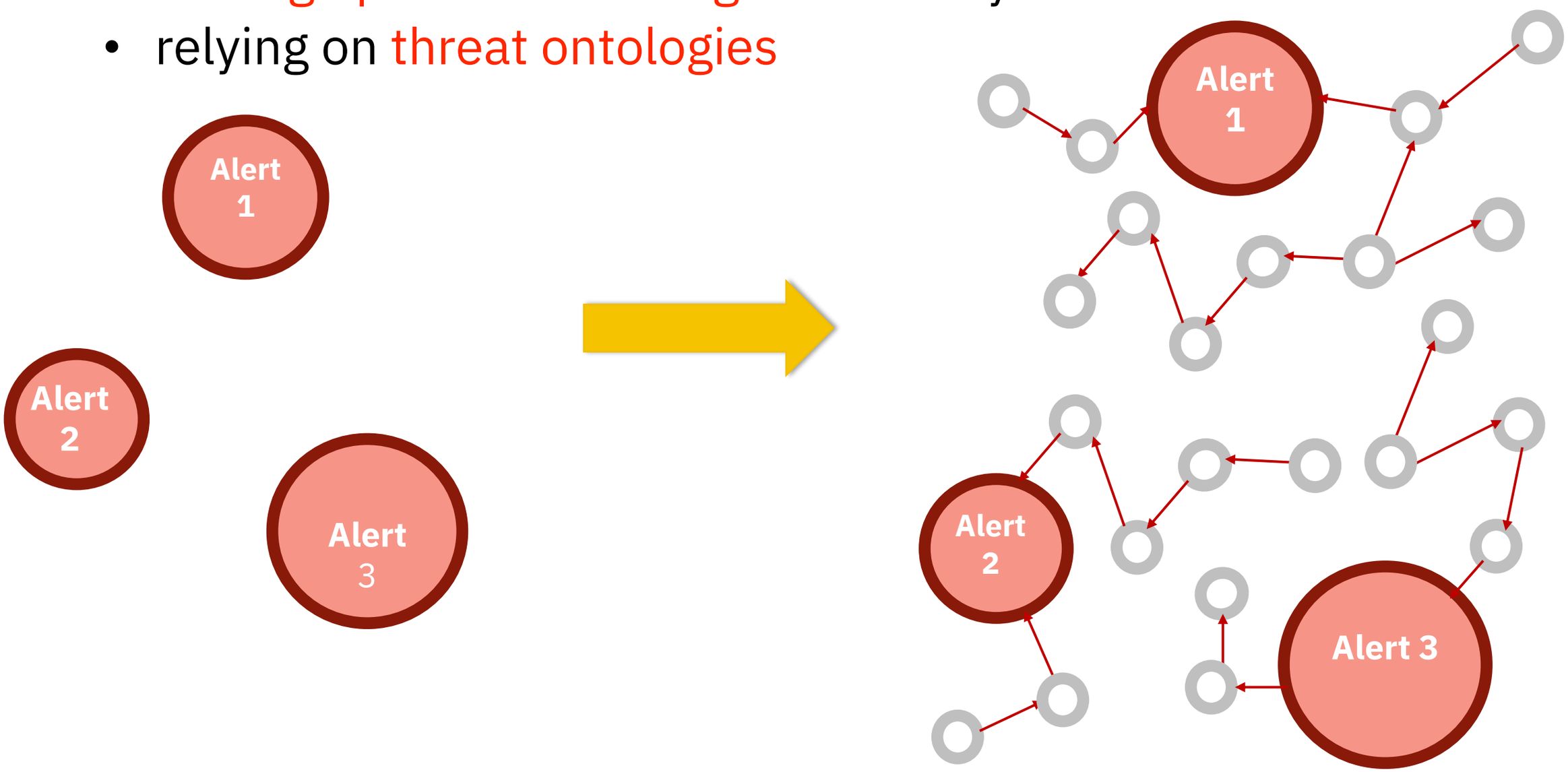
Automated Alert Triage

- **Attack causality analysis**
  - dependency explosion: a large graph around an alert
  - alert flooding: many alerts



Limitations of Existing Approaches

- **Attack causality analysis**
- **Computational resources are limited**
- **Two critical assumptions**
  - scaling up each alert triage is the only solution
  - relying on threat ontologies



Limitations of Existing Approaches

- **Attack causality analysis**
- **Computational resources are limited**
- **Two critical assumptions**
  - scaling up each alert triage is the only solution
  - relying on threat ontologies

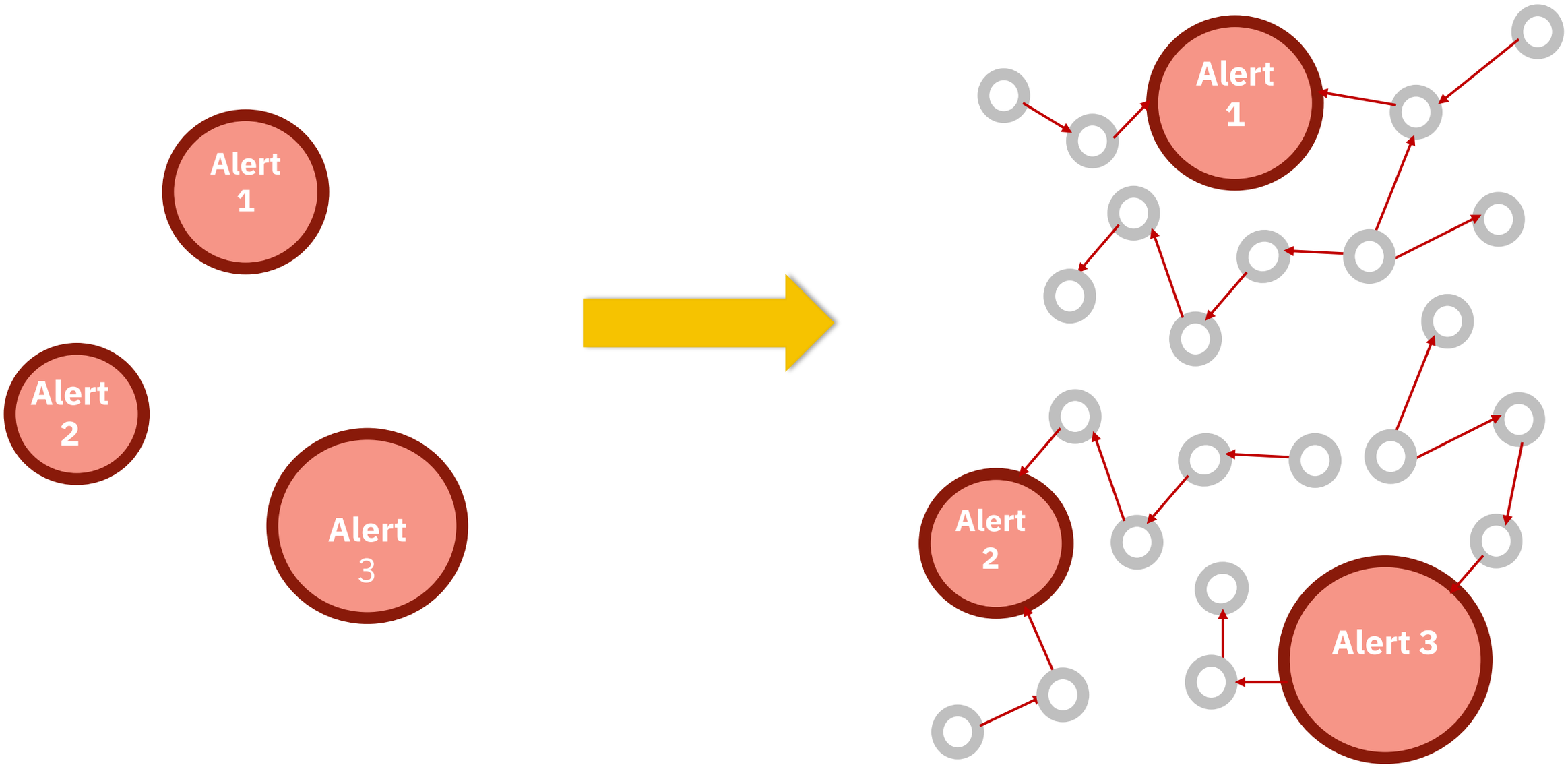
- **Our observations**
  - scaling out with collaborative alert triage is important
  - less dependent on presumed attack patterns

RAPID: scales out the alert processing capability in enterprises without relying on a threat ontology

Outline

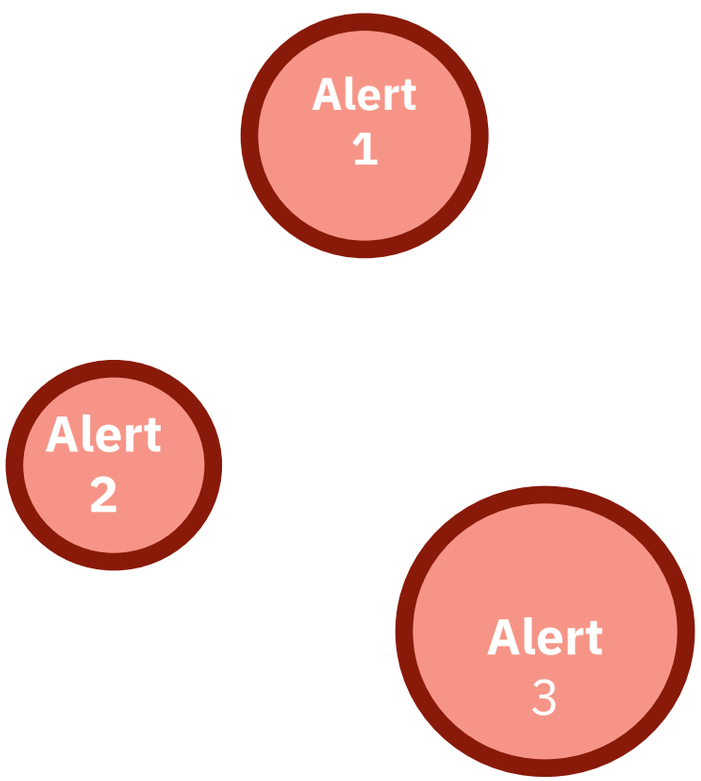
- Introduction
- **System Overview**
- Interruptible Priority-based Tracking
- Evaluation
- Conclusion

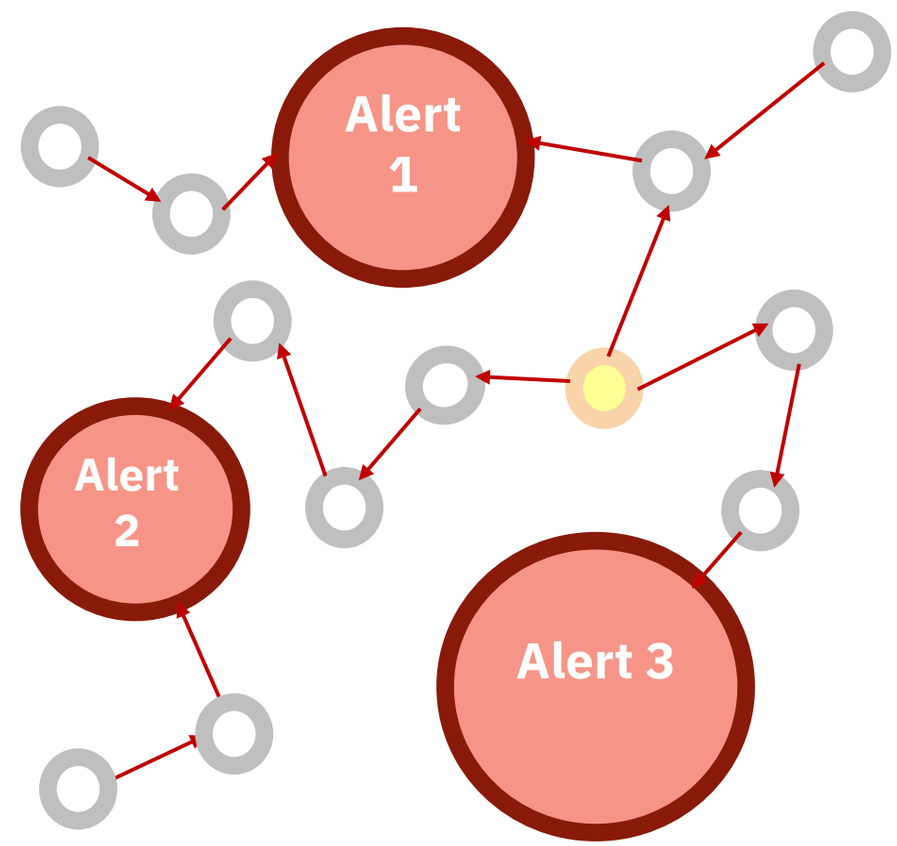
Non-collaborative Alert Triage



RAPID: Collaborative Alert Triage

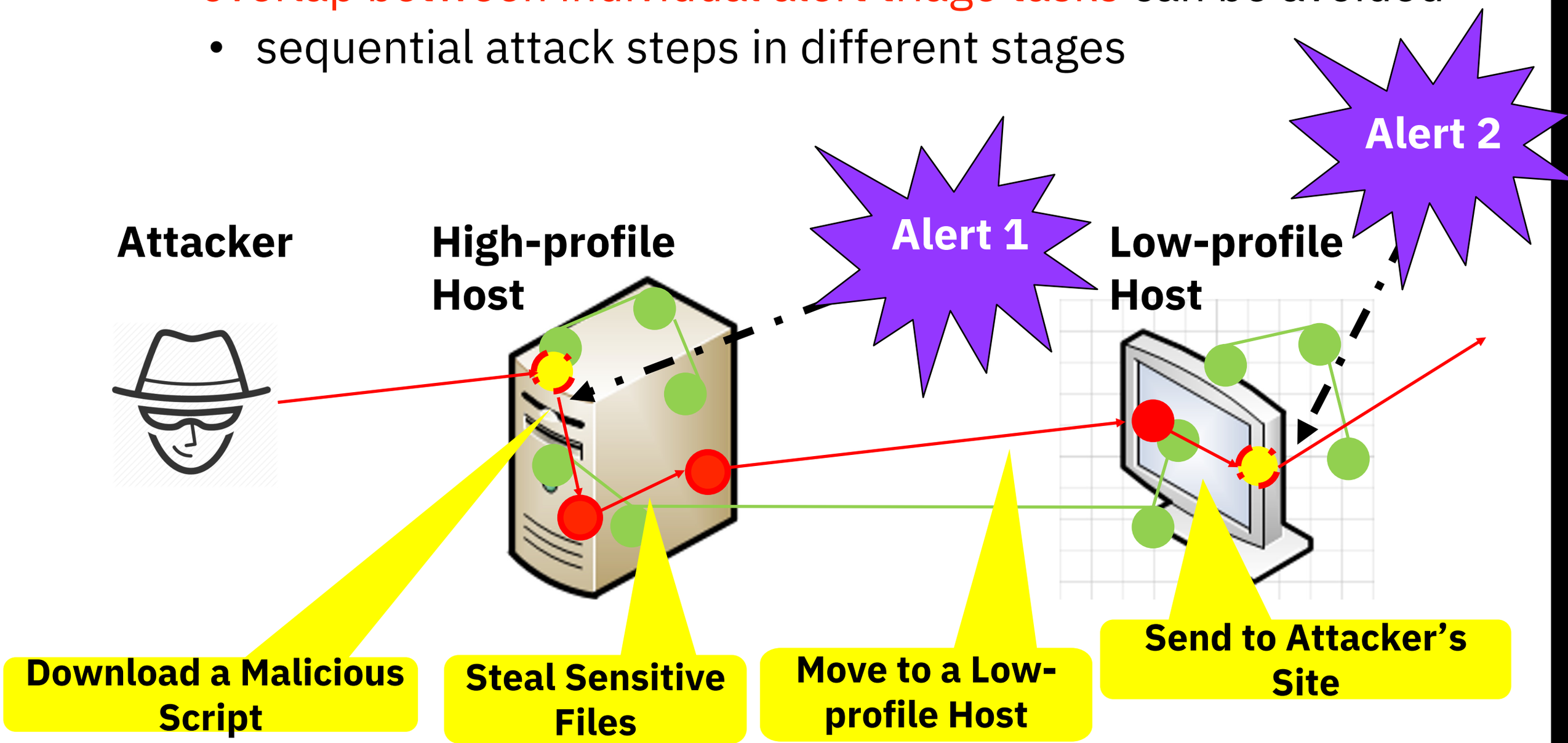
- **Idea: cluster the causally connected alerts to improve efficiency**





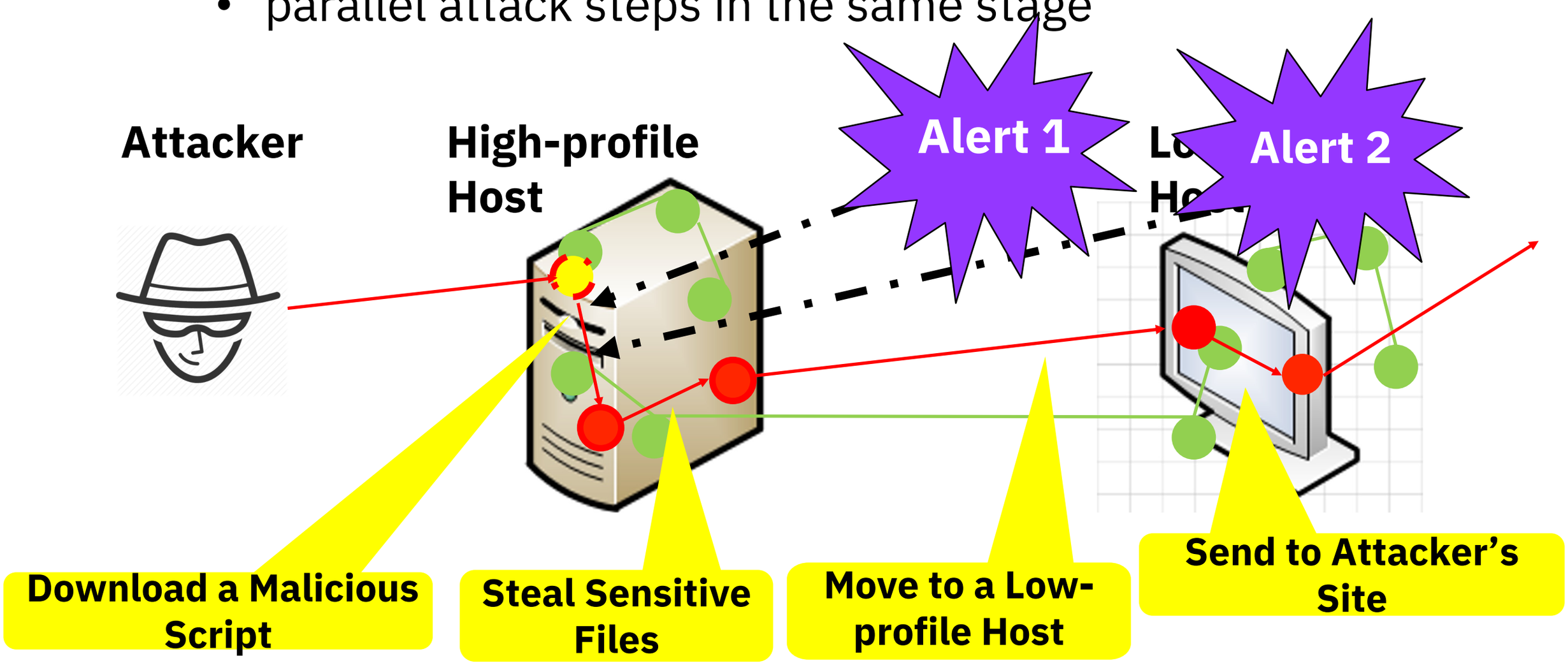
Why Collaborative Alert Triage?

- **Triage task deduplication**
  - overlap between individual alert triage tasks can be avoided
    - sequential attack steps in different stages



Why Collaborative Alert Triage?

- **Triage task deduplication**
  - overlap between individual alert triage tasks can be avoided
    - sequential attack steps in different stages
    - parallel attack steps in the same stage



Why Collaborative Alert Triage?

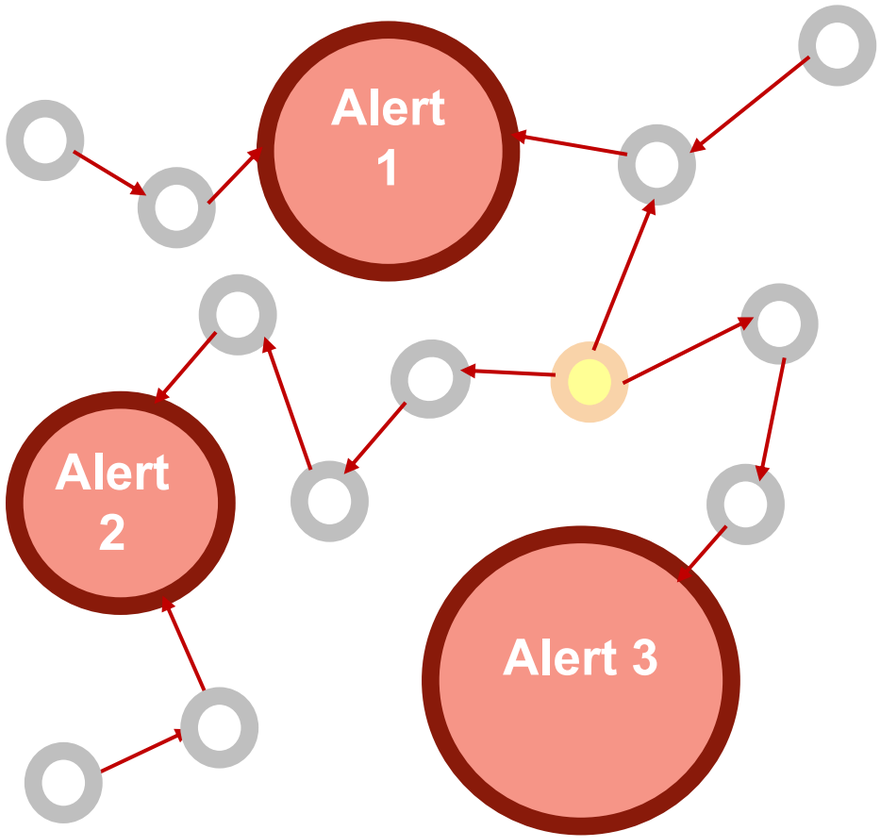
- **Triage task deduplication**
- **Cross-task prioritization: non-collaborative**
  - later tasks starve if computational resources are occupied
  - dropped tasks are not resumed even if other tasks are later diagnosed as less important

Why Collaborative Alert Triage?

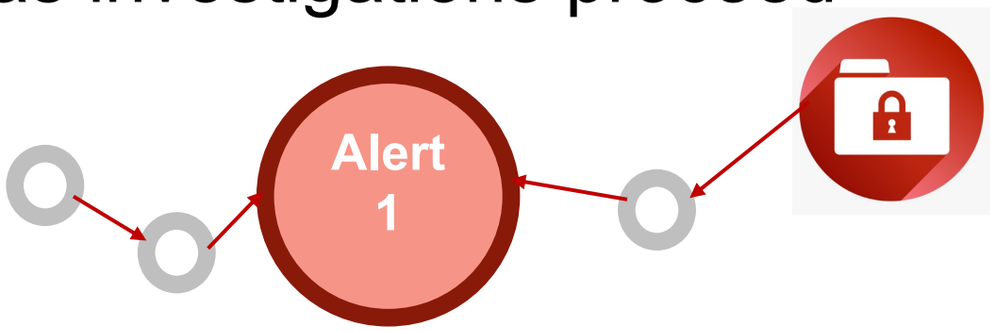
- **Triage task deduplication**
- **Cross-task prioritization**
  - re-allocate computational resources to most critical alerts
  - adapt alert prioritization to newly learned context
    - true alerts can be initially assigned with low severity

Why Collaborative Alert Triage?

- **Triage task deduplication**
- **Cross-task prioritization**
  - re-allocate computational resources to most critical alerts
  - adapt alert prioritization to newly learned context
    - re-assess alert priorities as investigations proceed



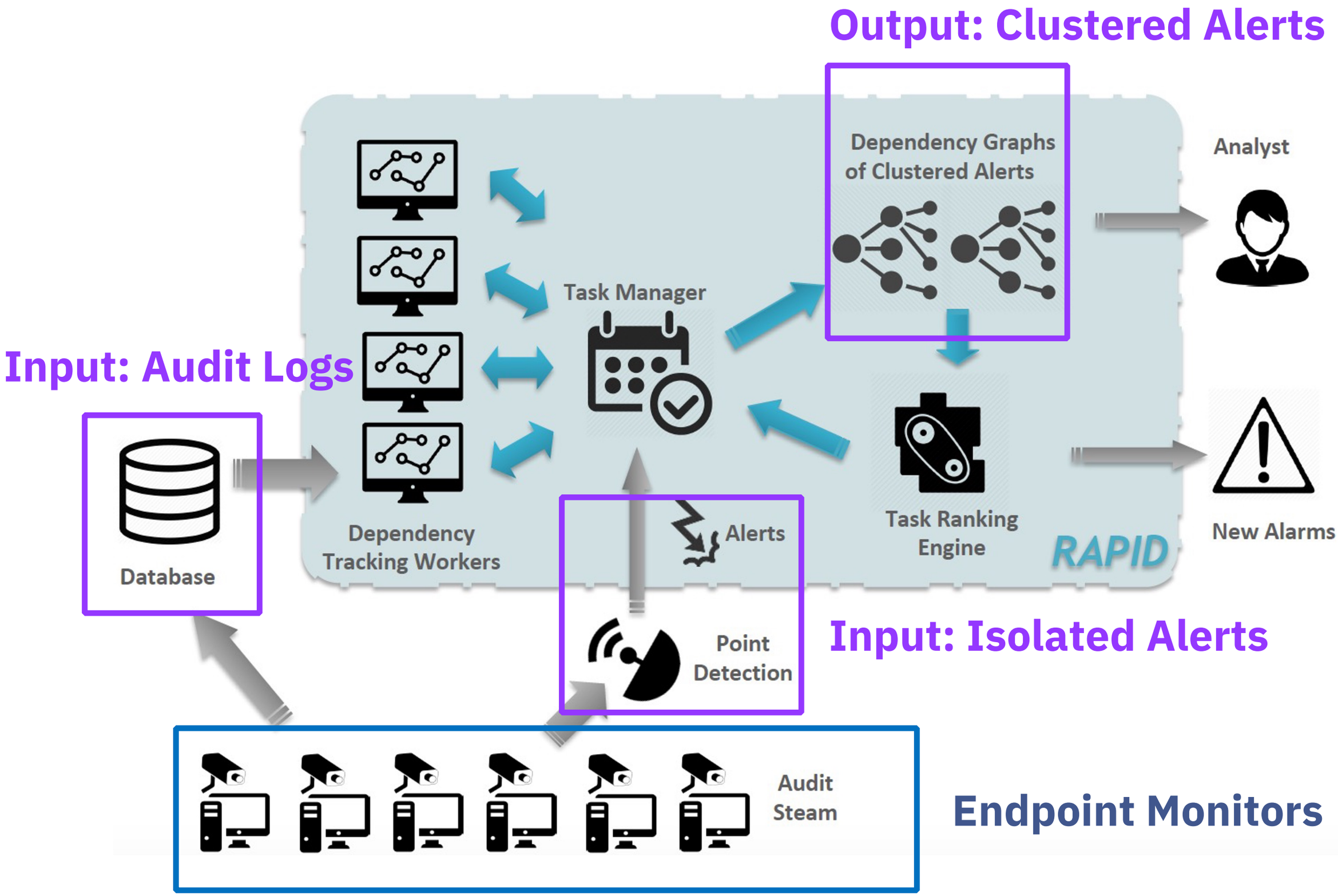
Densely connected alerts



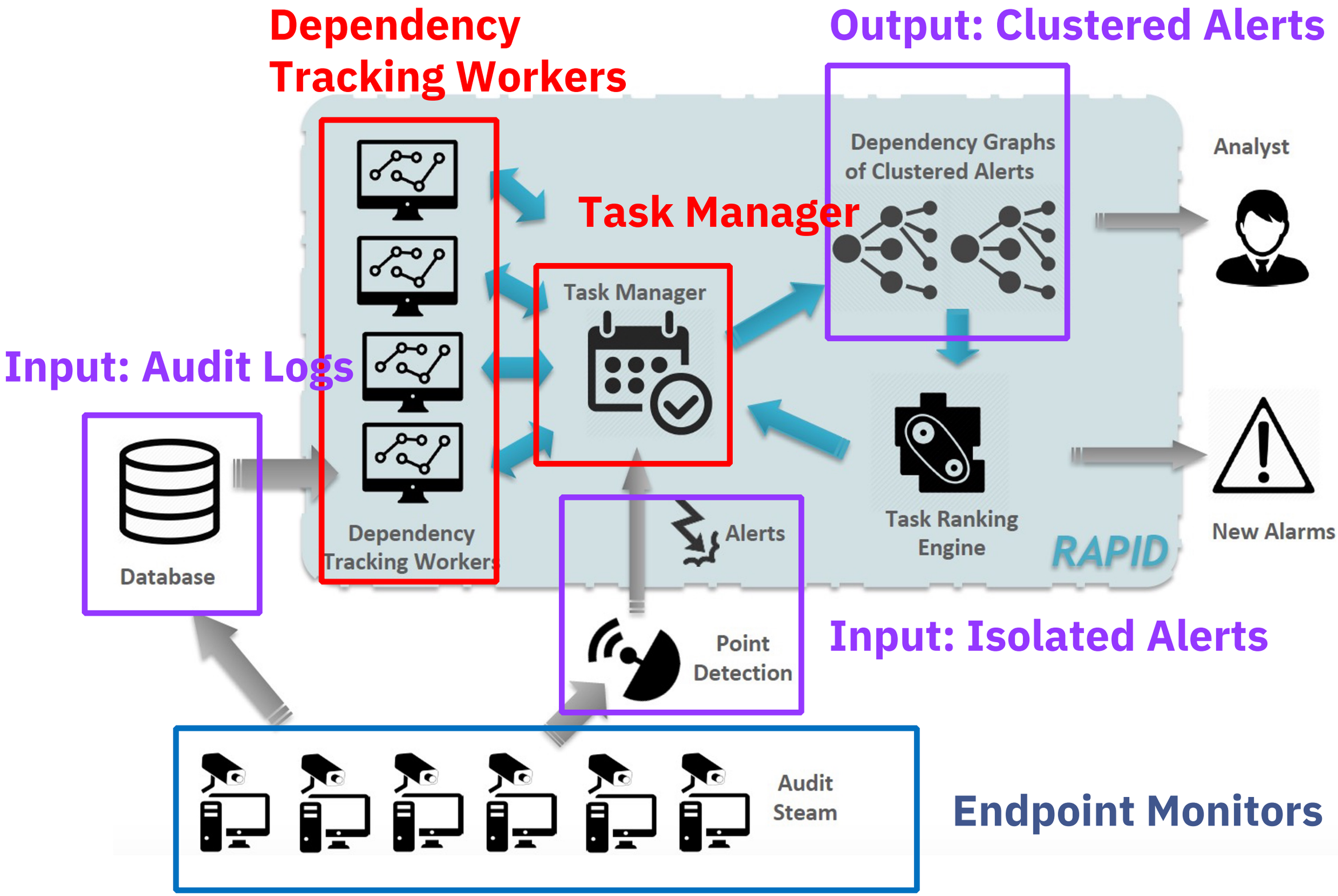
Touches sensitive files

Freshness

System Overview



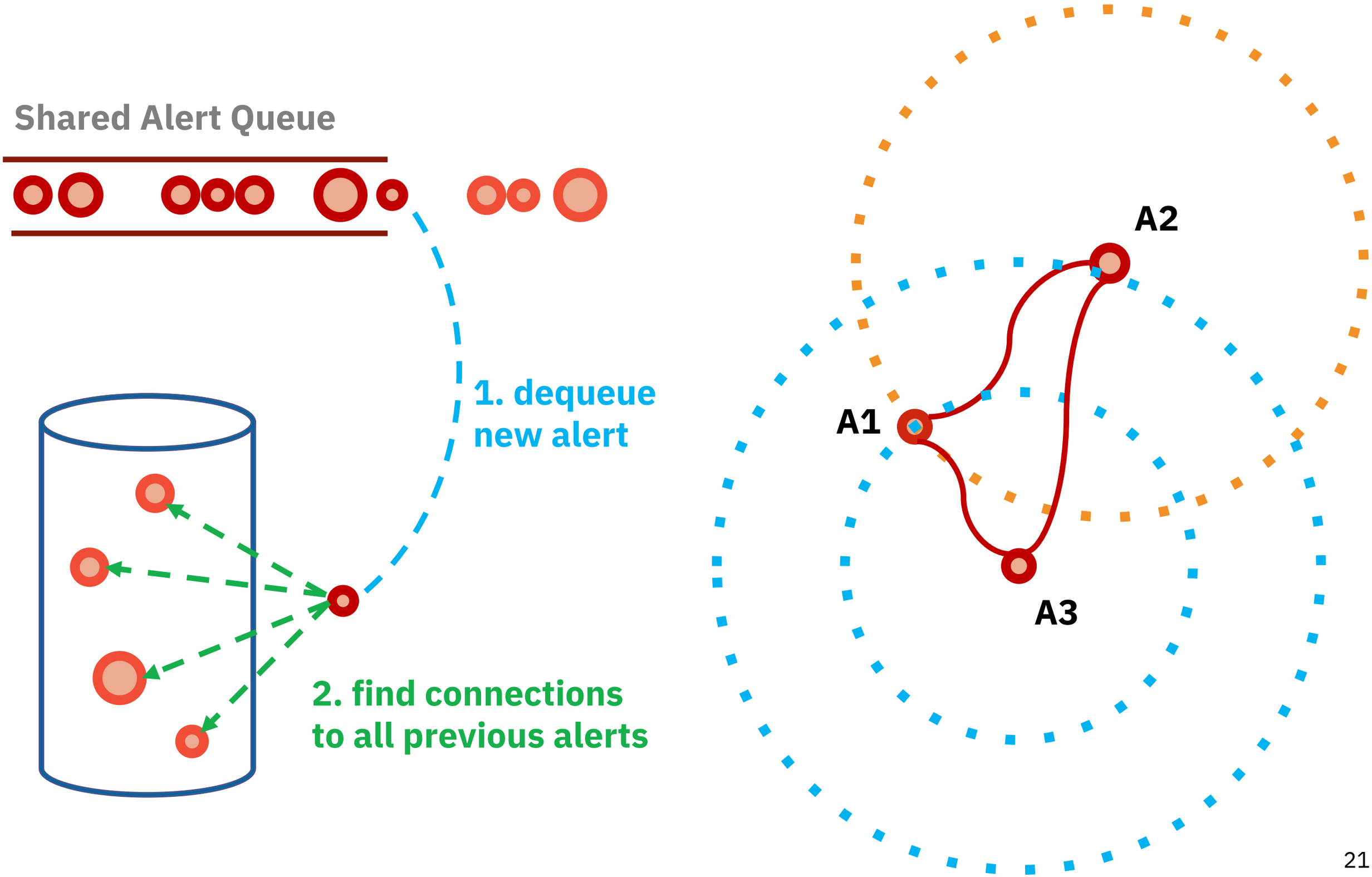
System Overview



Outline

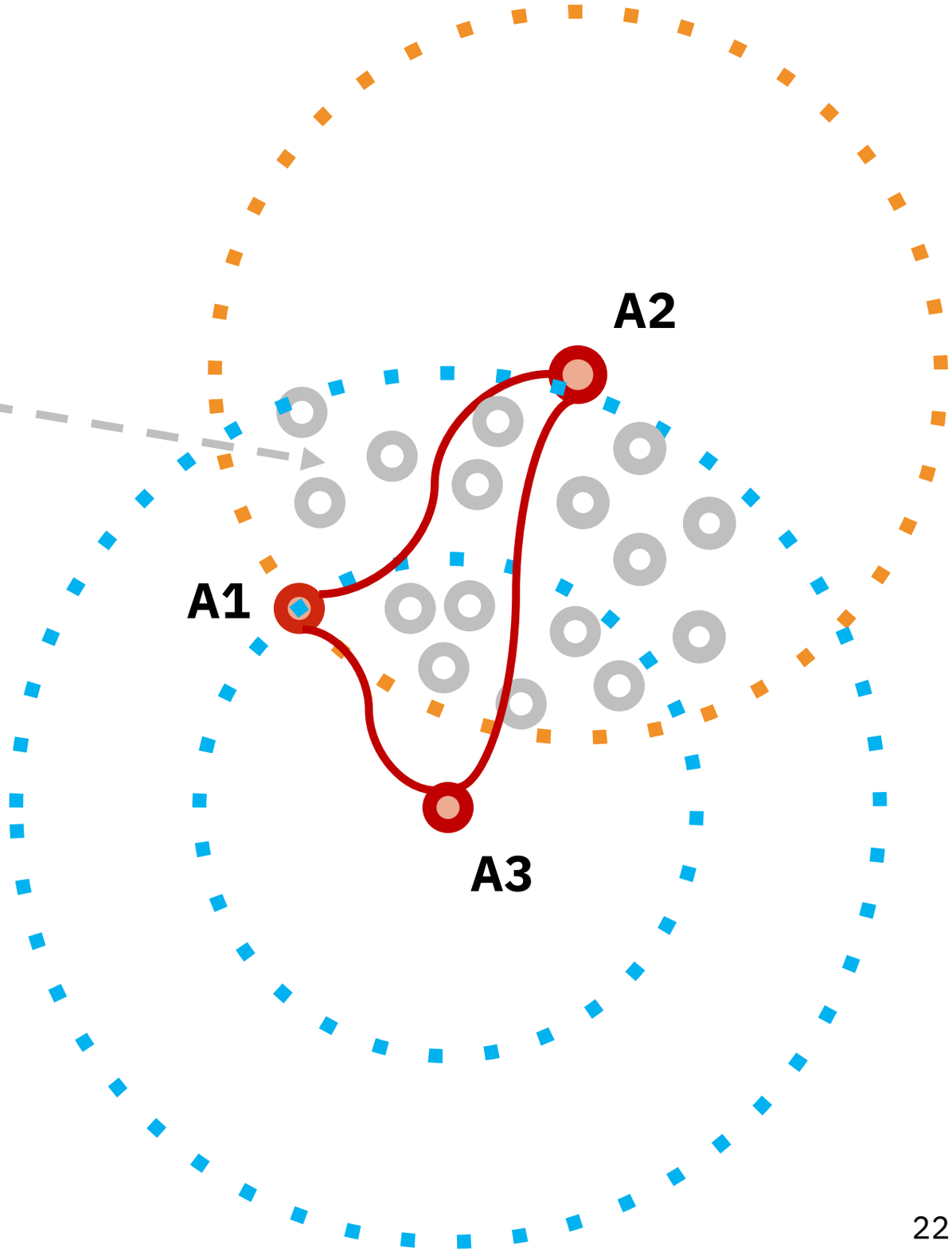
- Introduction
- System Overview
- **Interruptible Priority-based Tracking**
- Evaluation
- Conclusion

Connect the Alerts: Naïve Approach



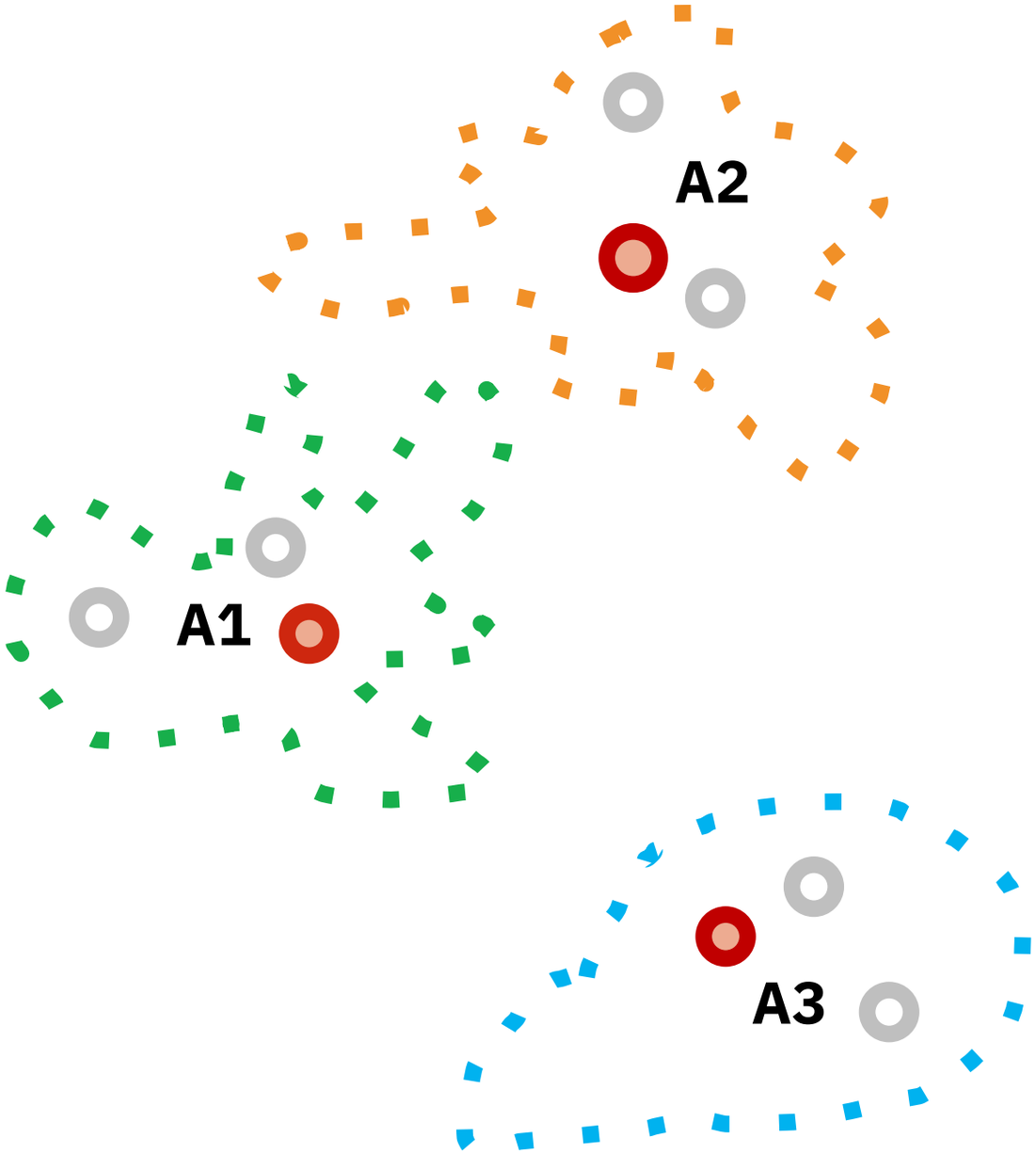
Problem with Complete Computation

- **Inefficiency**
  - reprocessing
    - grey nodes twice in connect(A2, A1), connect(A3, A2)
  - batch processing
    - connect(A2, A1), connect(A3, A2) are bundled in a session



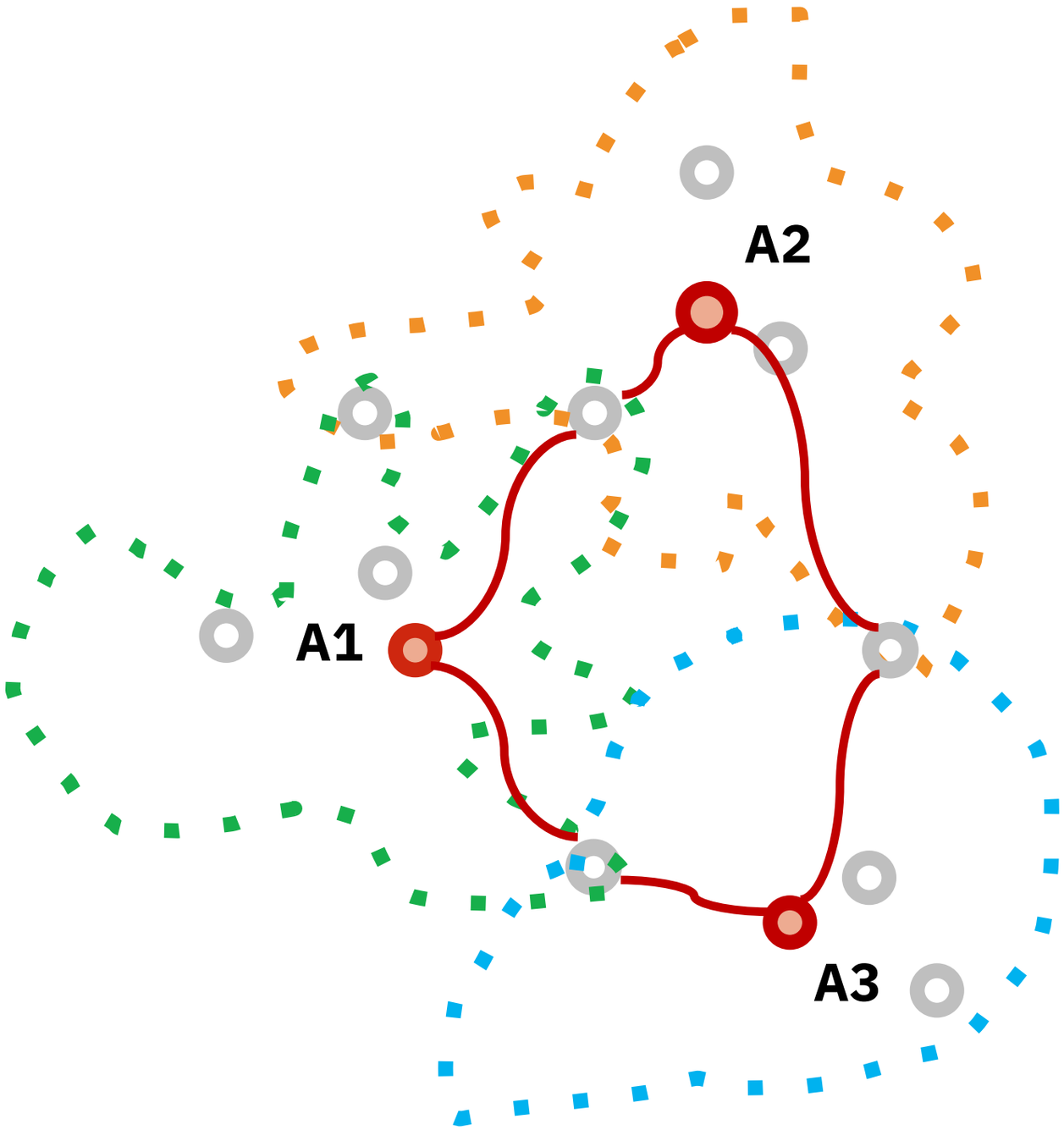
Our solution: Partial Computation

- **Sub-problems**
  - each alert tracks causal dependencies backwards
  - store intermediate entities as waypoints
- **Waypoints**
  - last event timestamp
  - shortest paths from entities to alerts



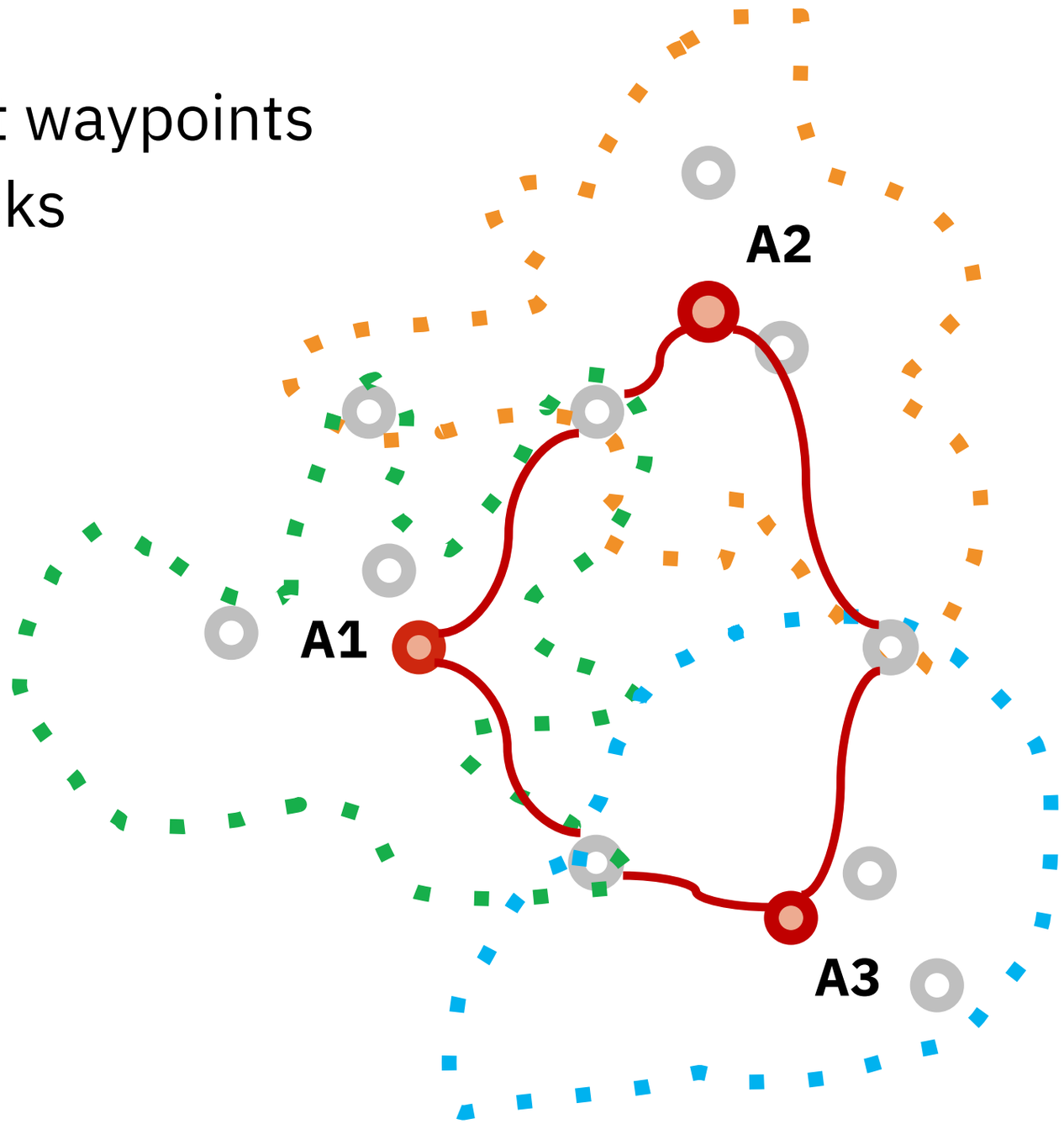
Our solution: Partial Computation

- **Sub-problems**
  - each alert tracks causal dependencies backwards
  - store intermediate entities as waypoints
- **Waypoints**
  - last event timestamp
  - shortest paths from entities to alerts
- **Paths**
  - hit overlapped waypoints



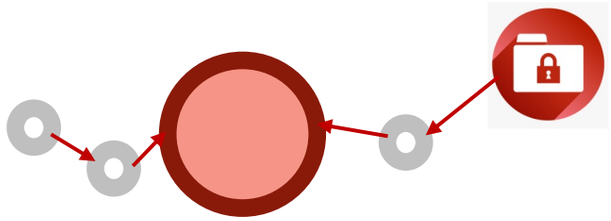
Each Alert: Priority-based Tracking

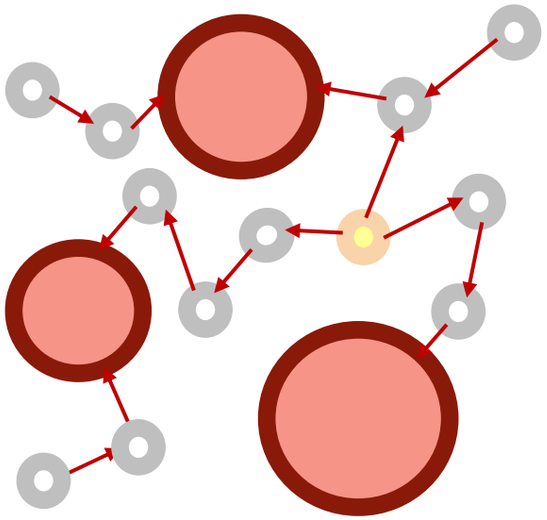
- **Deduplication:** equips a new alert with information from triage of previous alerts
  - connect to the closest waypoints
  - merge most triage tasks



Each Alert: Priority-based Tracking

- **Deduplication:** equips a new alert with information from triage of previous alerts
- **Prioritization:** incorporates contextual information into alert priorities
  - increases an aggregate priority
    - suspicious neighbors
    - connected severe alerts
    - new





Alerts: Interruptible Priority-based Tracking

- Re-allocate resources if a more critical alert is found?
- **Traditionally, no control over the task progress:** kills an old triage task to switch to another alert
  - no partial results from the killed task
  - computation resources wasted when no alerts

Alerts: Interruptible Priority-based Tracking

- Re-allocate resources if a more critical alert is found?
- **Our design:** task manager dynamically prioritizes alert triage tasks
  - task state <unprocessed entities, priority, time>
- **Interruptible alert triage tasks**
  - pause a running old task
  - save the task state
  - resume the task later
- **Results harvested continuously**
  - partial tracking results for analytics

Outline

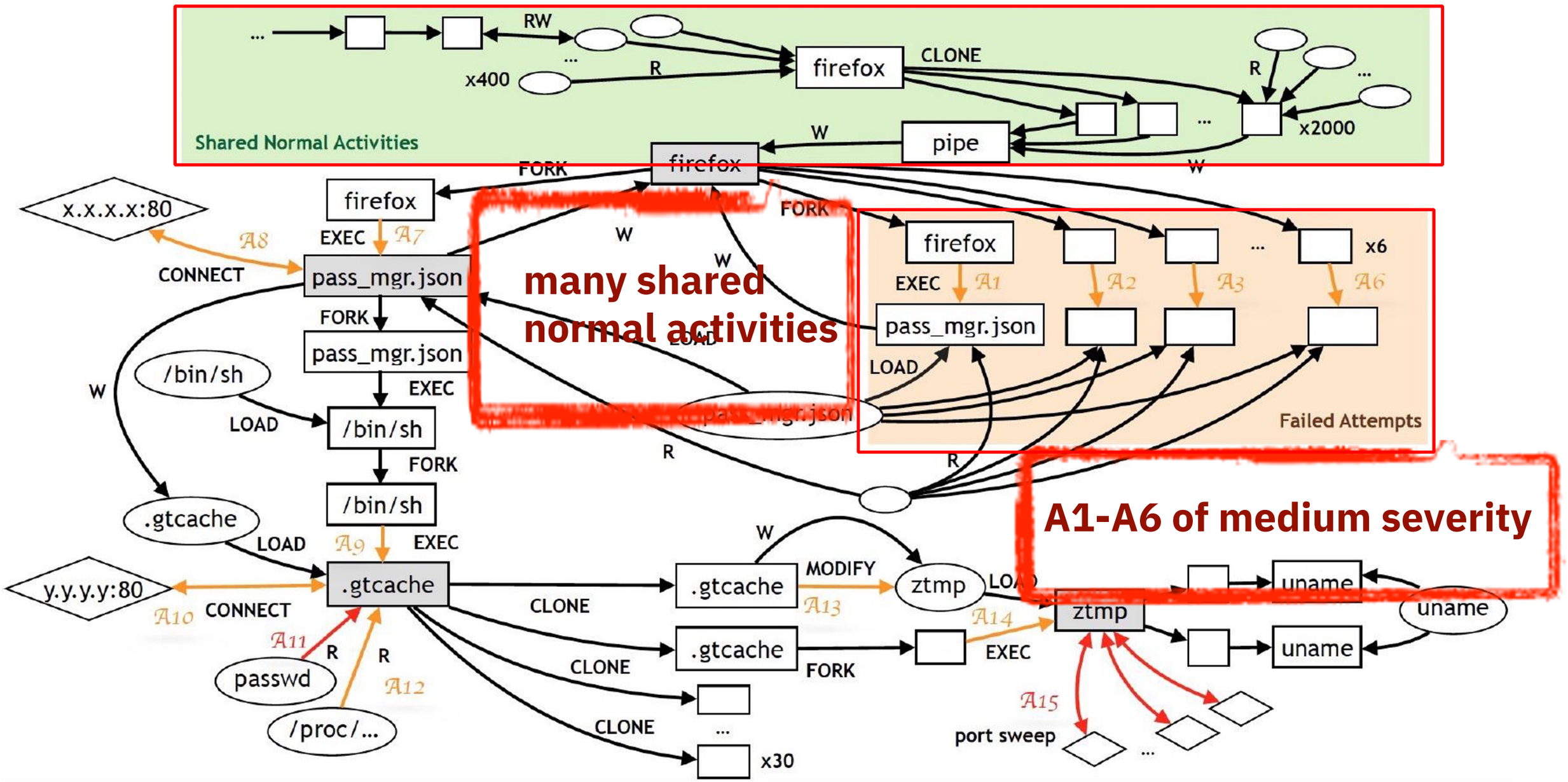
- Introduction
- System Overview
- Interruptible Priority-based Tracking
- **Evaluation**
- Conclusion

Experiment Setup

- **Adversarial engagement of DARPA Transparent Computing (TC) program**
  - 1TB data including 411m events on 7 hosts over 2 weeks
    - FreeBSD, Linux, and Windows
  - 300k alerts from real detectors

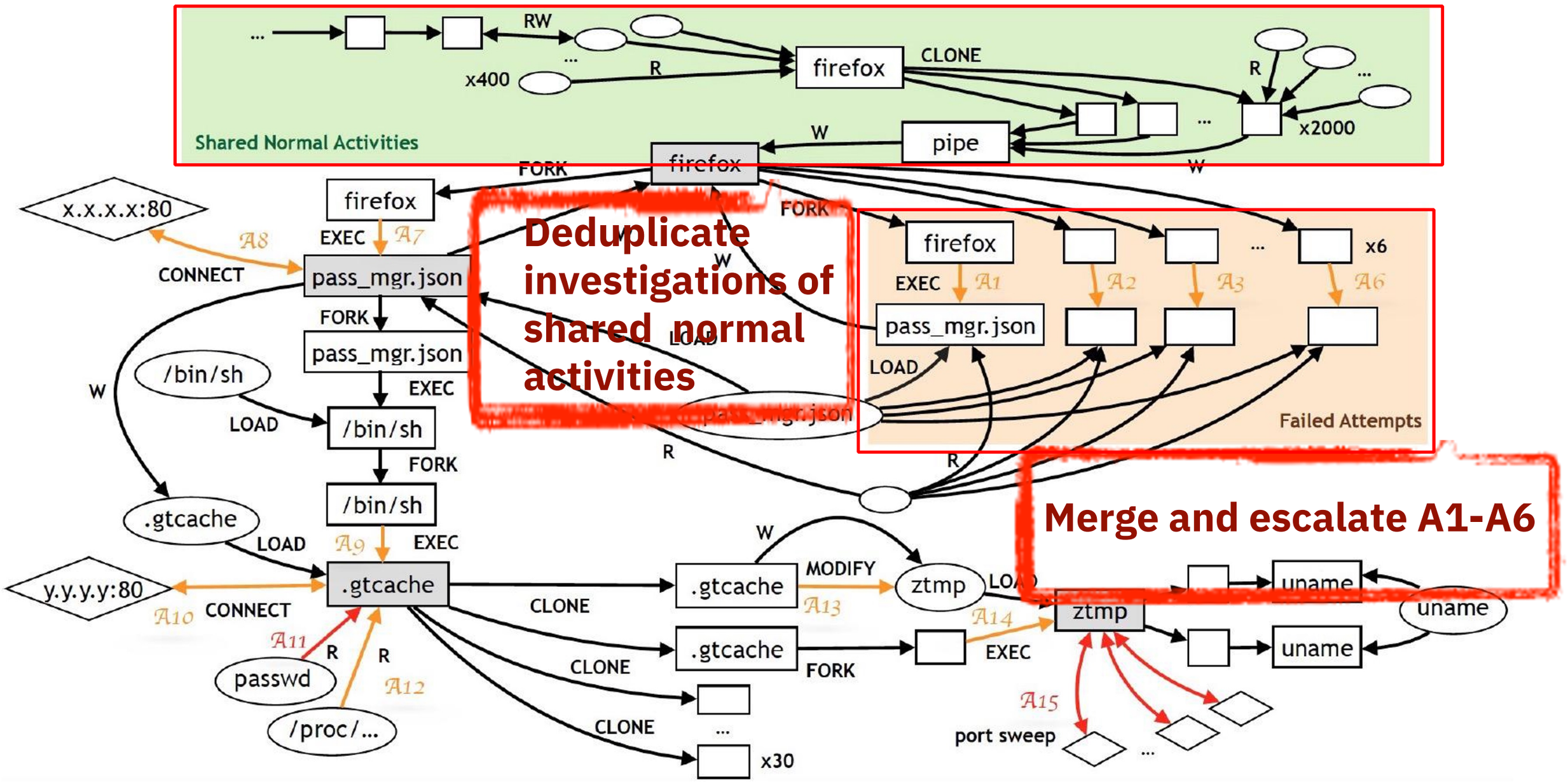
- **Three representative attack cases**
  - attacks from red teams
  - ground truth collected as critical events (CE)

Case Study: Firefox Extension Attack



A compromised Firefox password manager extension drops a malicious script, exfiltrates data and performs port scanning

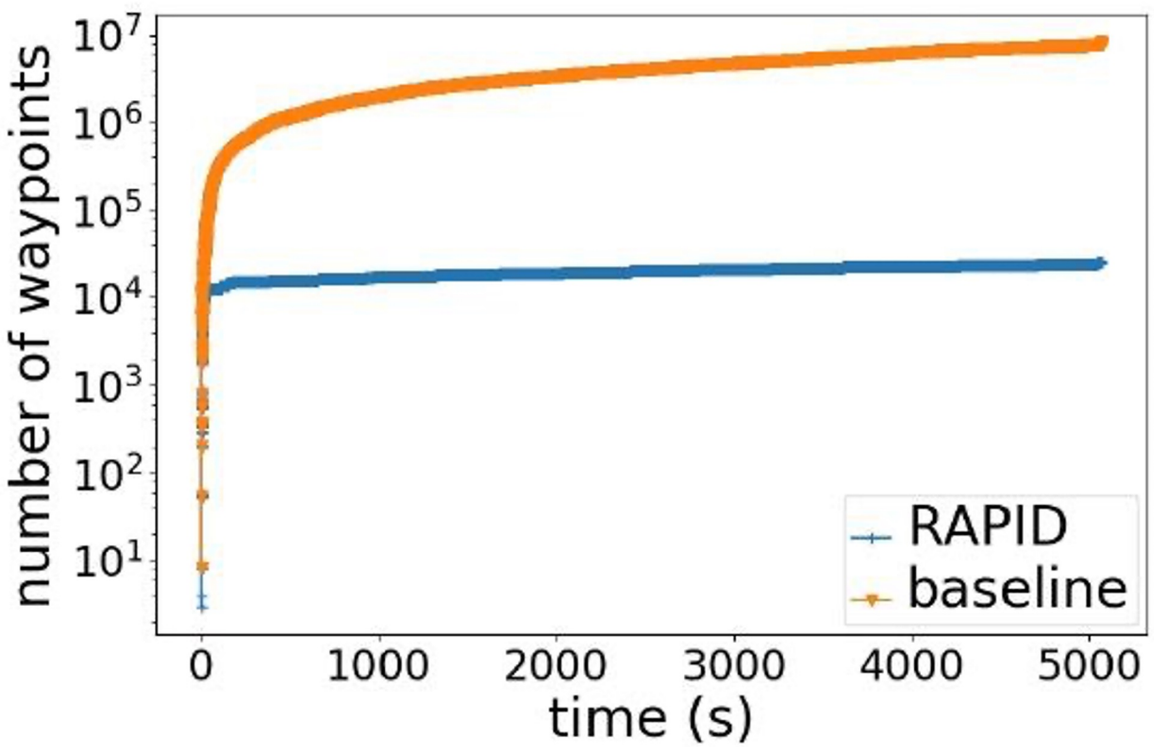
Case Study: Firefox Extension Attack

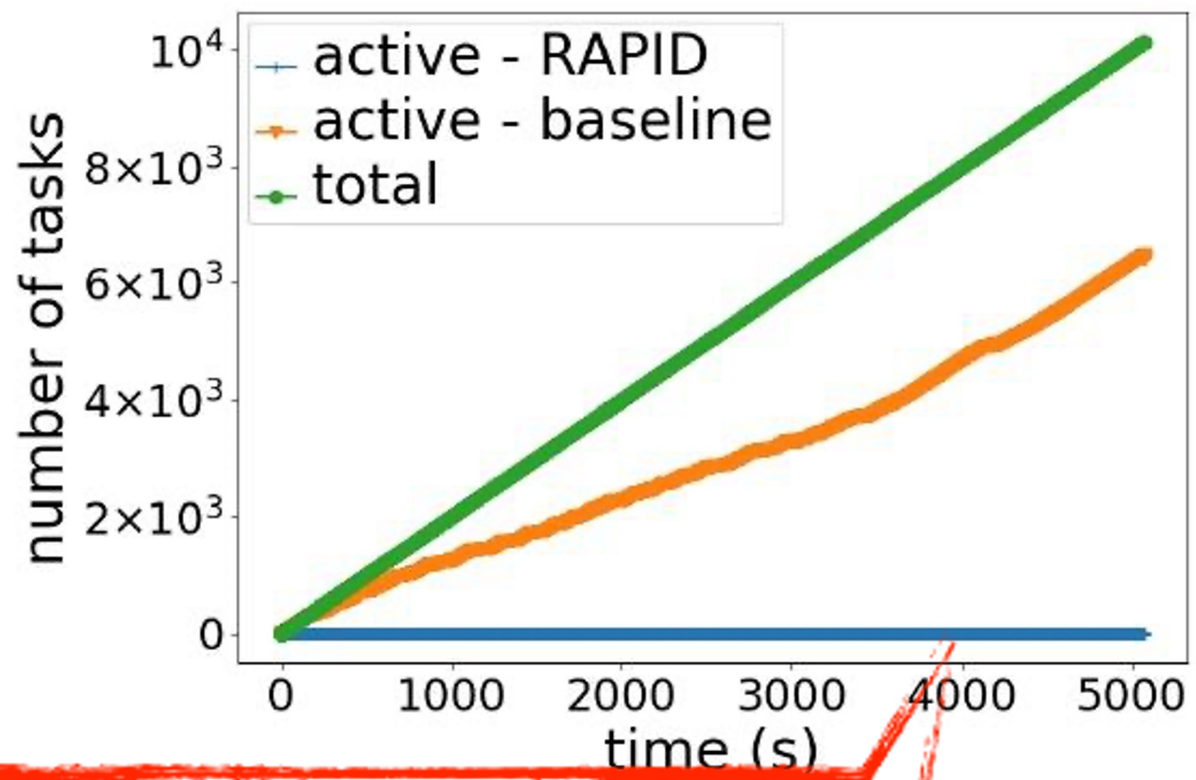


- We avoid reprocessing alerts and prioritize densely connected alerts over normal activities

Results: Efficiency

- **Deduplication improves space efficiency by up to three orders of magnitude**





Thousands of alert triage tasks quickly merge into tens of tasks via the waypoint mechanism

Results: Accuracy

- **Prioritization discovers more attack traces in causality analysis within the time limit**

Attack Case	Critical Events	Baseline		RAPID	
		CE	Missing	CE	Missing
Firefox Extension - 1	66	59	10.61%	66	0%
Firefox Extension - 2	66	58	12.12%	66	0%
Vulnerable Nginx- 1	77	51	20.78%	77	0%
Vulnerable Nginx - 2	77	51	20.78%	77	0%
Phishing E-mail - 1	217	109	49.77%	217	0%
Phishing E-mail - 2	217	132	39.17%	217	0%

Baseline:

- **~10% missing** for Firefox Extension
- **~20% missing** for Vulnerable Nginx
- **~40% missing** for Phishing Email

Rapid:

- **0% missing** for all

Results: Time Effectiveness

- **Prioritization accelerates the discovery of attack traces by up to two orders of magnitude**

Attack Case	Runtime (100%CEs)	
	Baseline	RAPID
Firefox Extension Attack - 1	> 7,600s	94.10s
Firefox Extension Attack - 2	> 12,346s	243.75s
Vulnerable Nginx Attack - 1	> 3,622s	22.61s
Vulnerable Nginx Attack - 2	> 10,023s	4430.74s
Phishing E-mail Attack - 1	> 8,060 s	4,052.76s
Phishing E-mail Attack - 2	> 32,695s	6,075.68s

Baseline:
- **> one hour** for both

Rapid:
- **94s** for Firefox Extension
- **22s** for Nginix

Conclusion

- **Proposed context-aware alert investigation platform**
  - with causality analysis capabilities
  - without relying on threat ontology
- **Evaluated on a 1TB dataset from DARPA TC program**

Thank you!