Technical Track

ARISE Automated Reliable Intelligent Software Engineering

香港中文大學 CUHK

Eadro: An End-to-End Troubleshooting Framework for Microservices on Multi-Source Data

Cheryl Lee[*], Tianyi Yang[*], Zhuangbin Chen[*], Yuxin Su[†], and Michael R. Lyu[*]

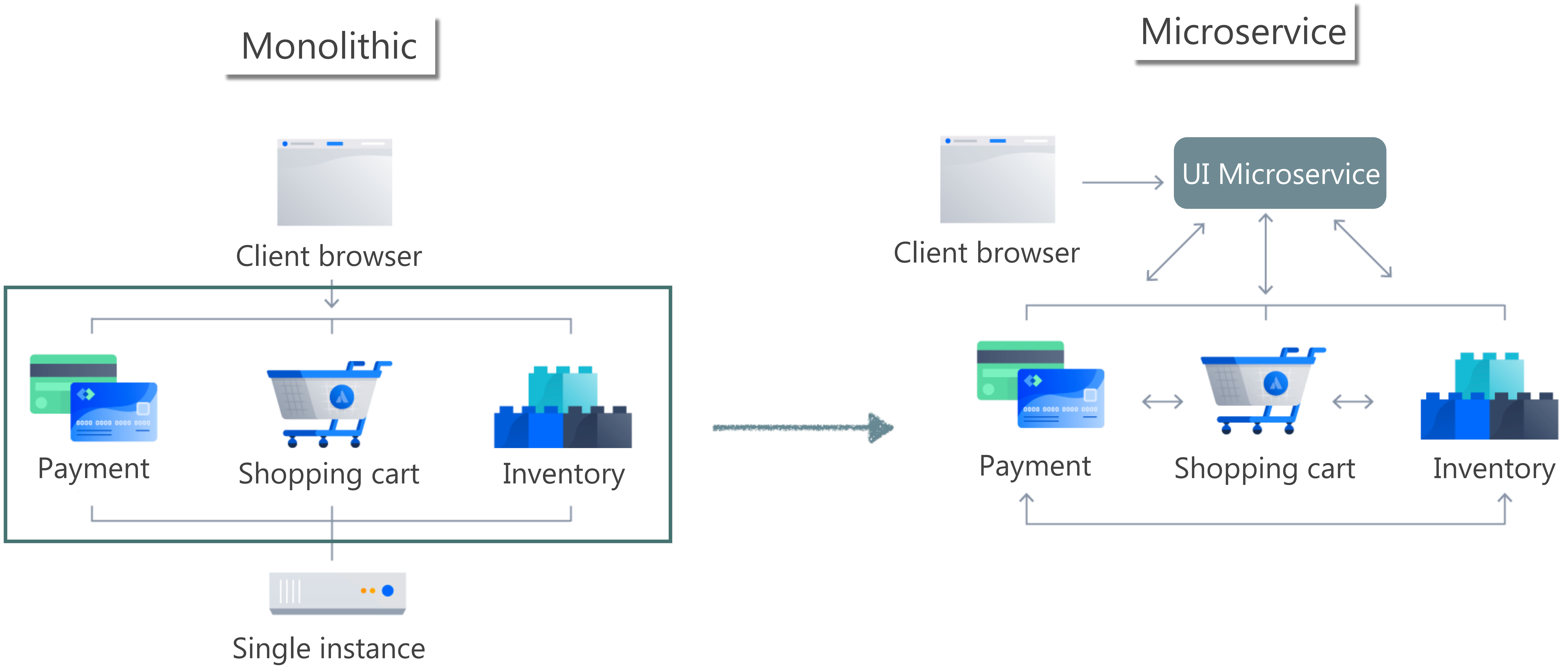
*The Chinese University of Hong Kong
†Sun Yat-sen University

May, 2023

Table of Contents

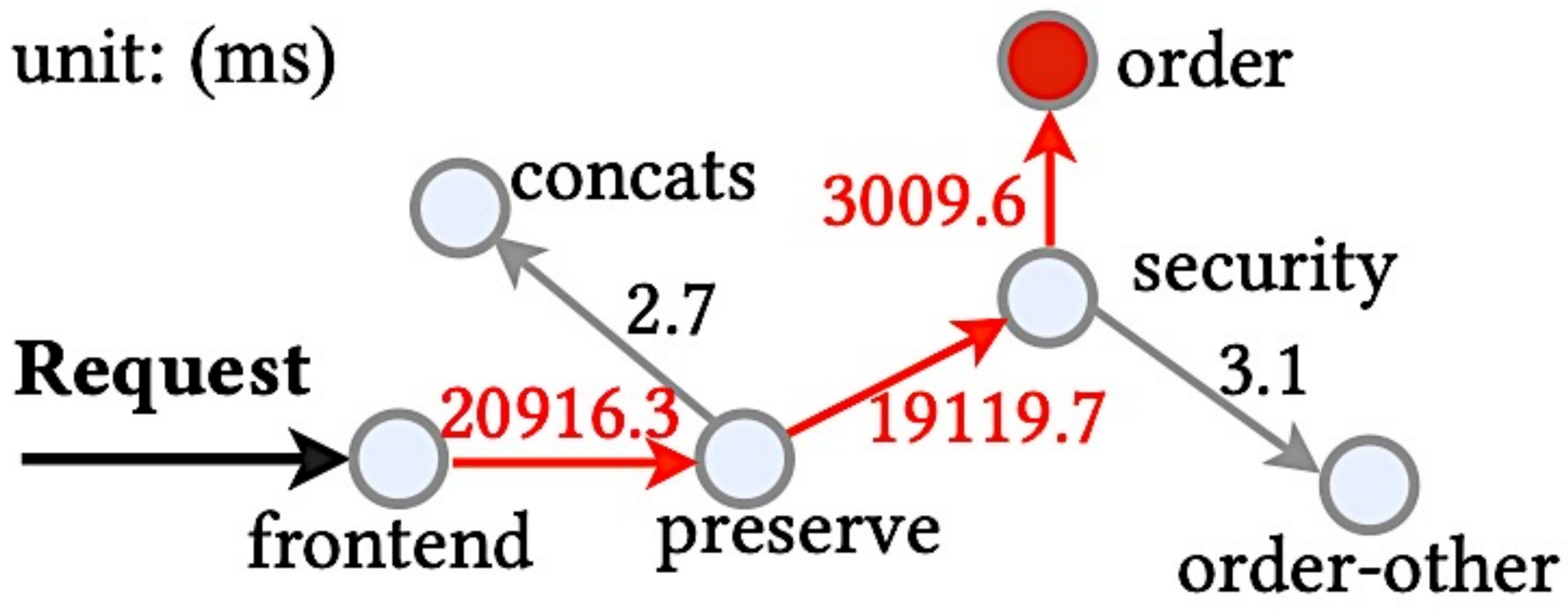

01 INTRODUCTION

Background



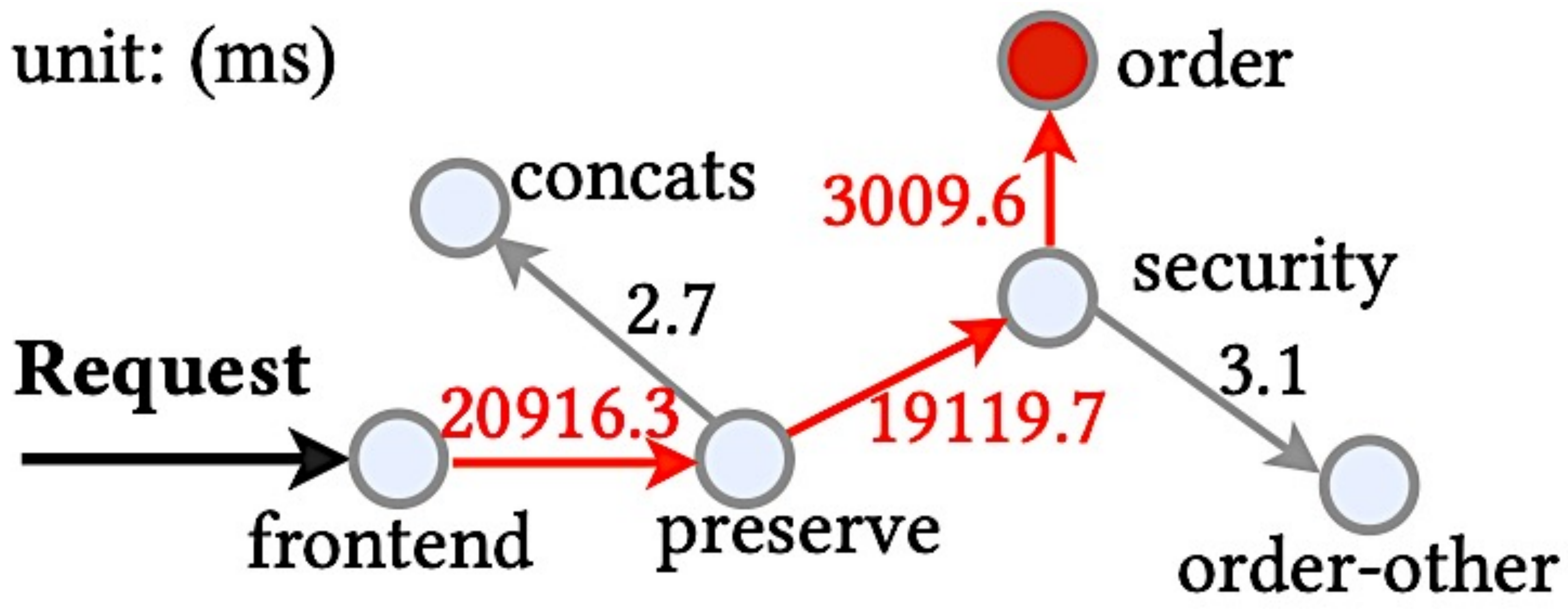
Monolithic
Client browser
Payment
Shopping cart
Inventory
Single instance
Microservice
Client browser
UI Microservice
Payment
Shopping cart
Inventory

Background

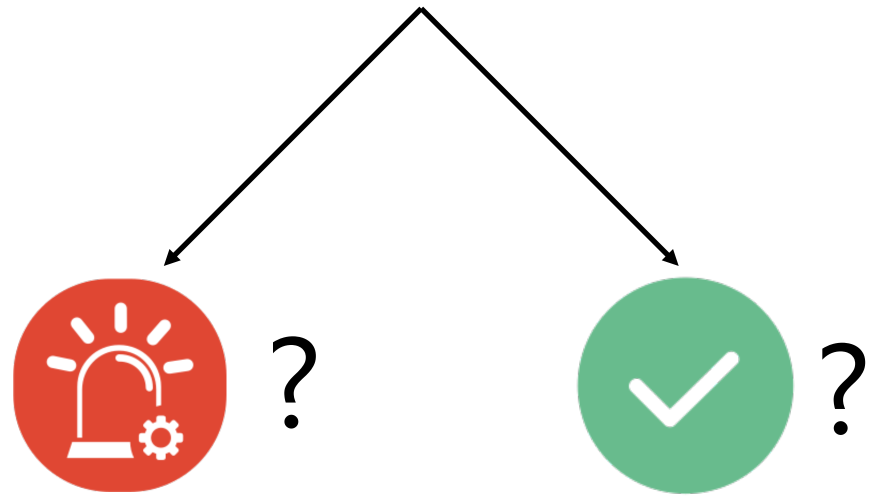


unit: (ms)
order
concats
3009.6
security
2.7
Request
3.1
20916.3
19119.7
frontend
preserve
order-other

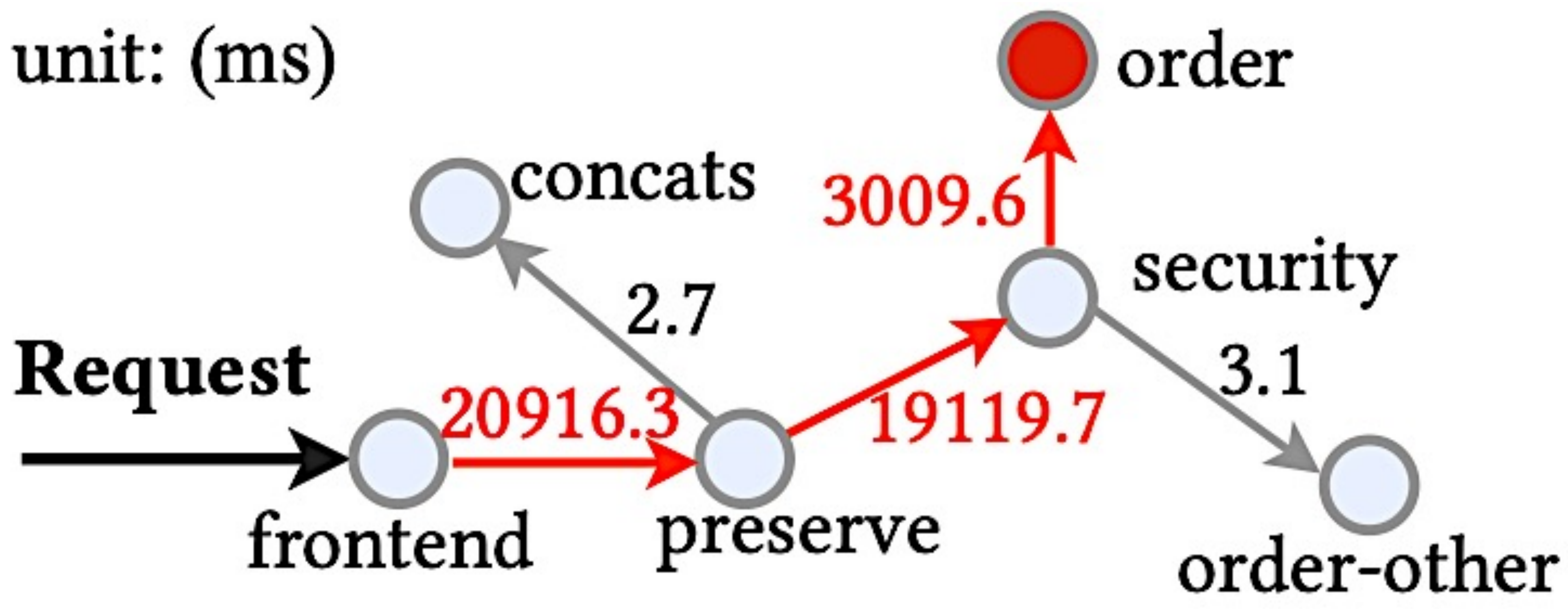
Background



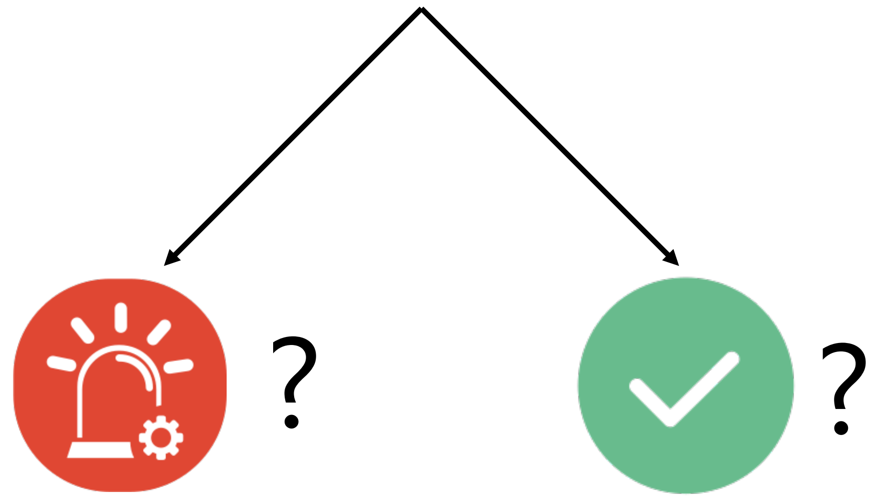
Anomaly detection (AD) identifies the existence of an anomaly.



Background

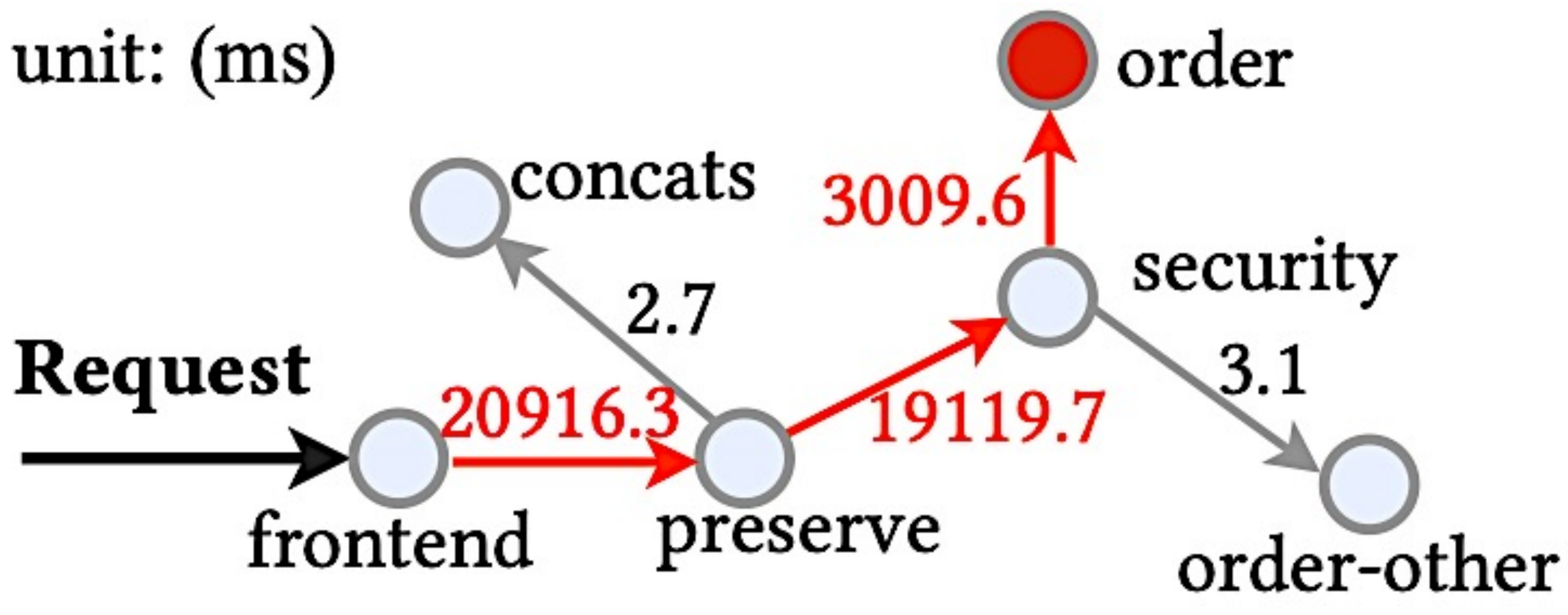


Anomaly detection (AD) identifies the existence of an anomaly.

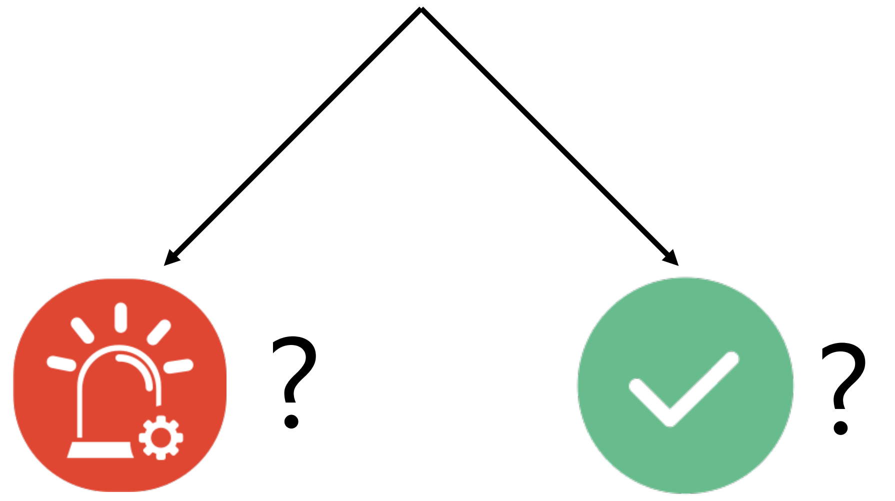


But we need finer-grained information…

Background

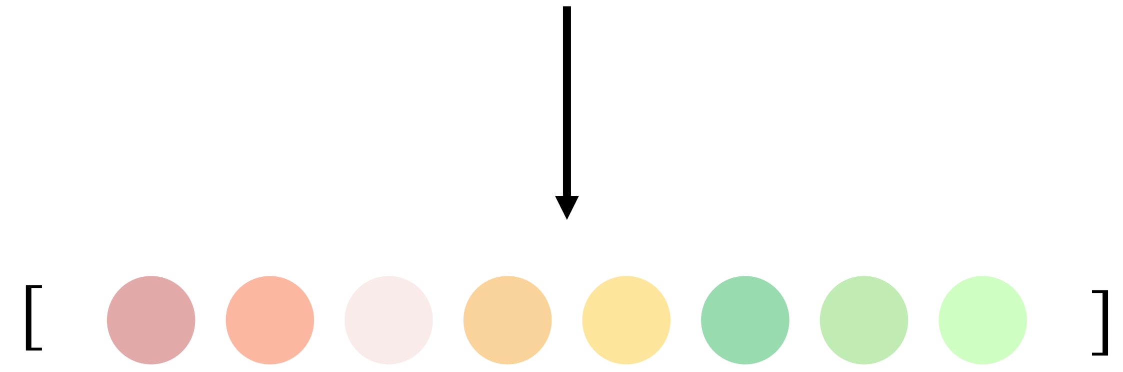


Anomaly detection (AD) identifies the existence of an anomaly.

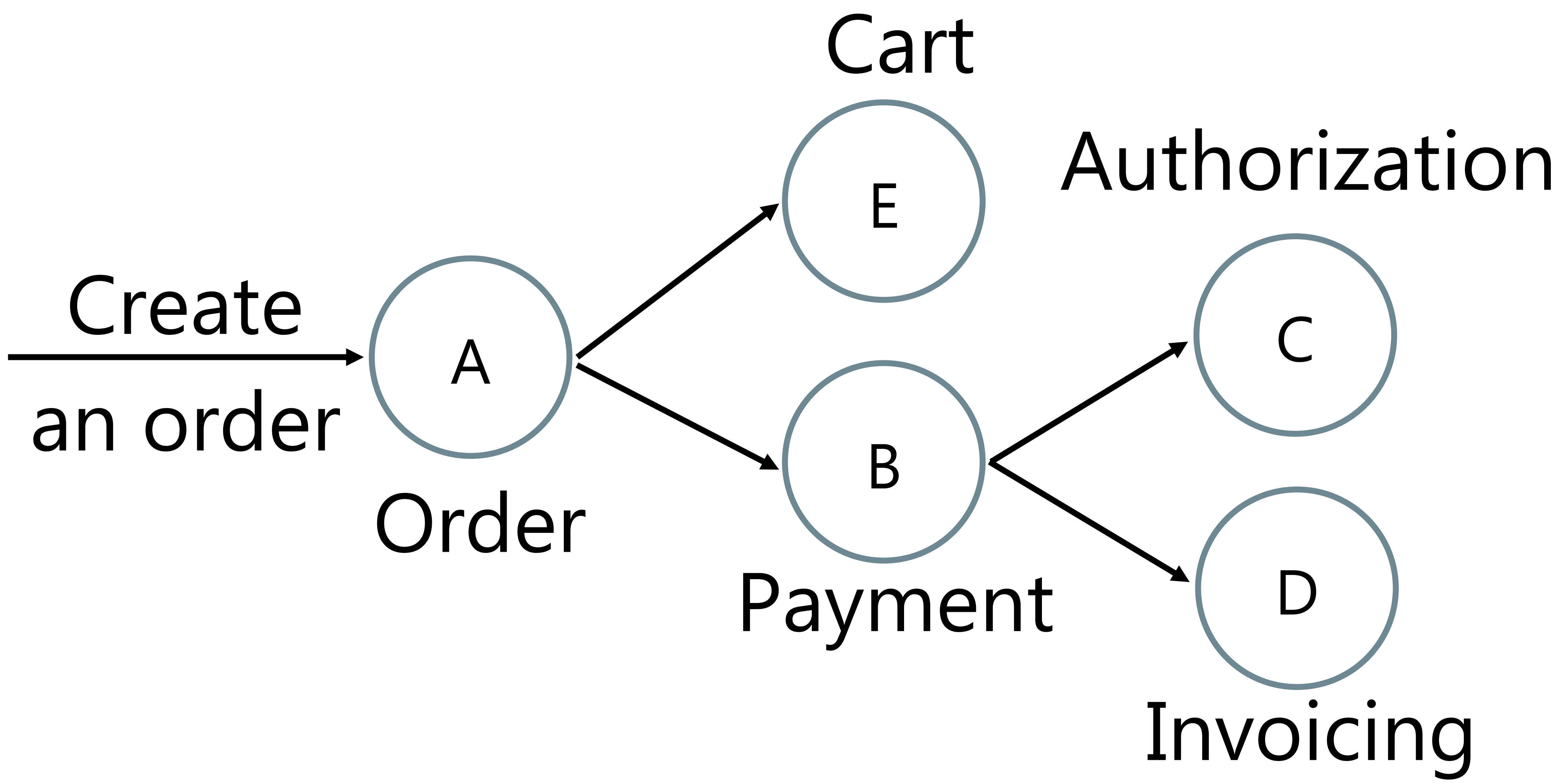


But we need finer-grained information…

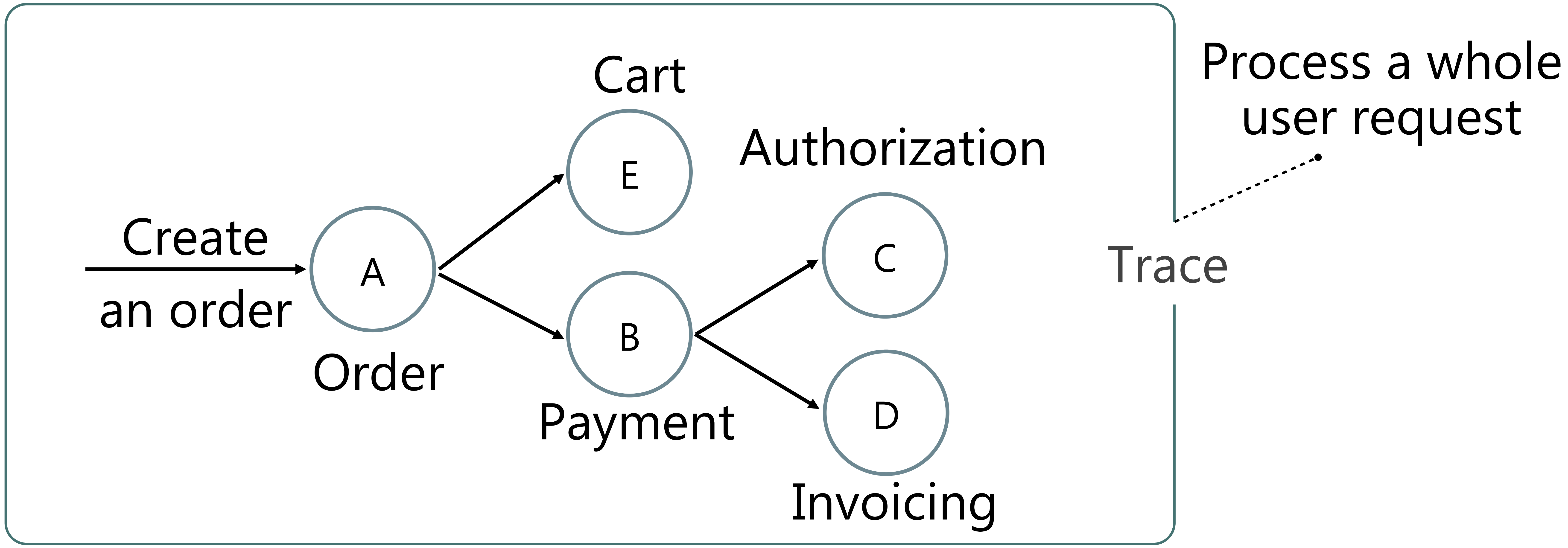
Root cause localization (RCL) answers the probability of each microservice being the culprit.



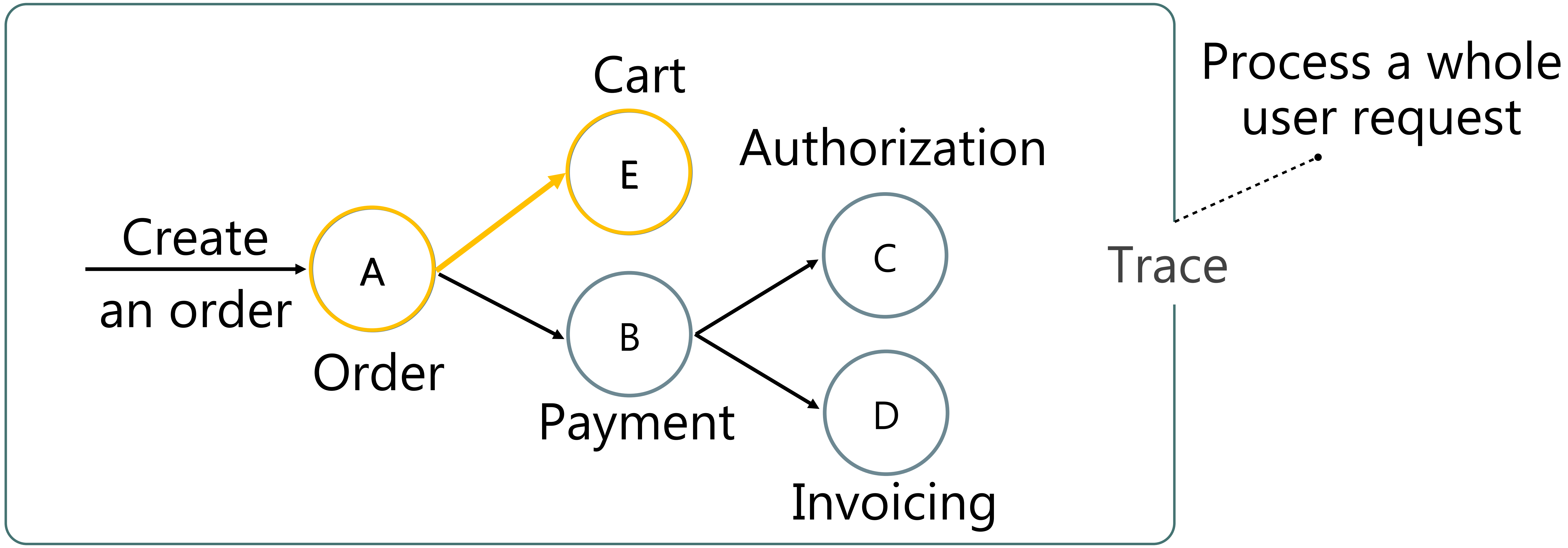
Terminologies



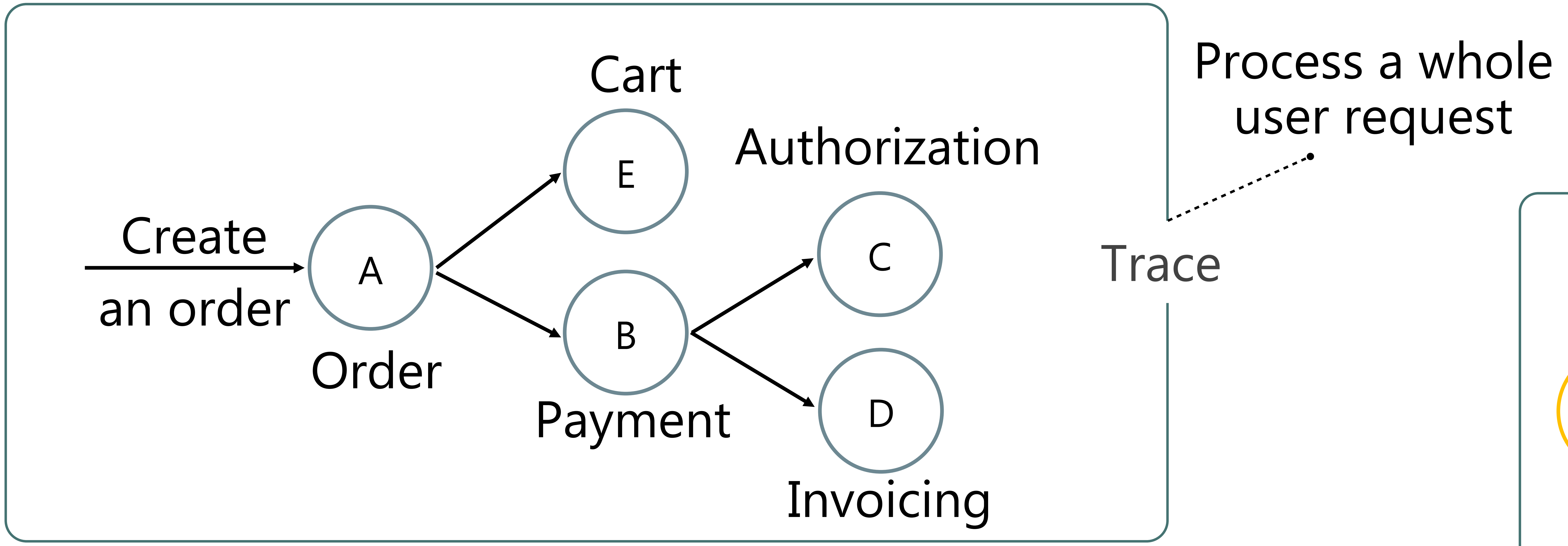
Terminologies

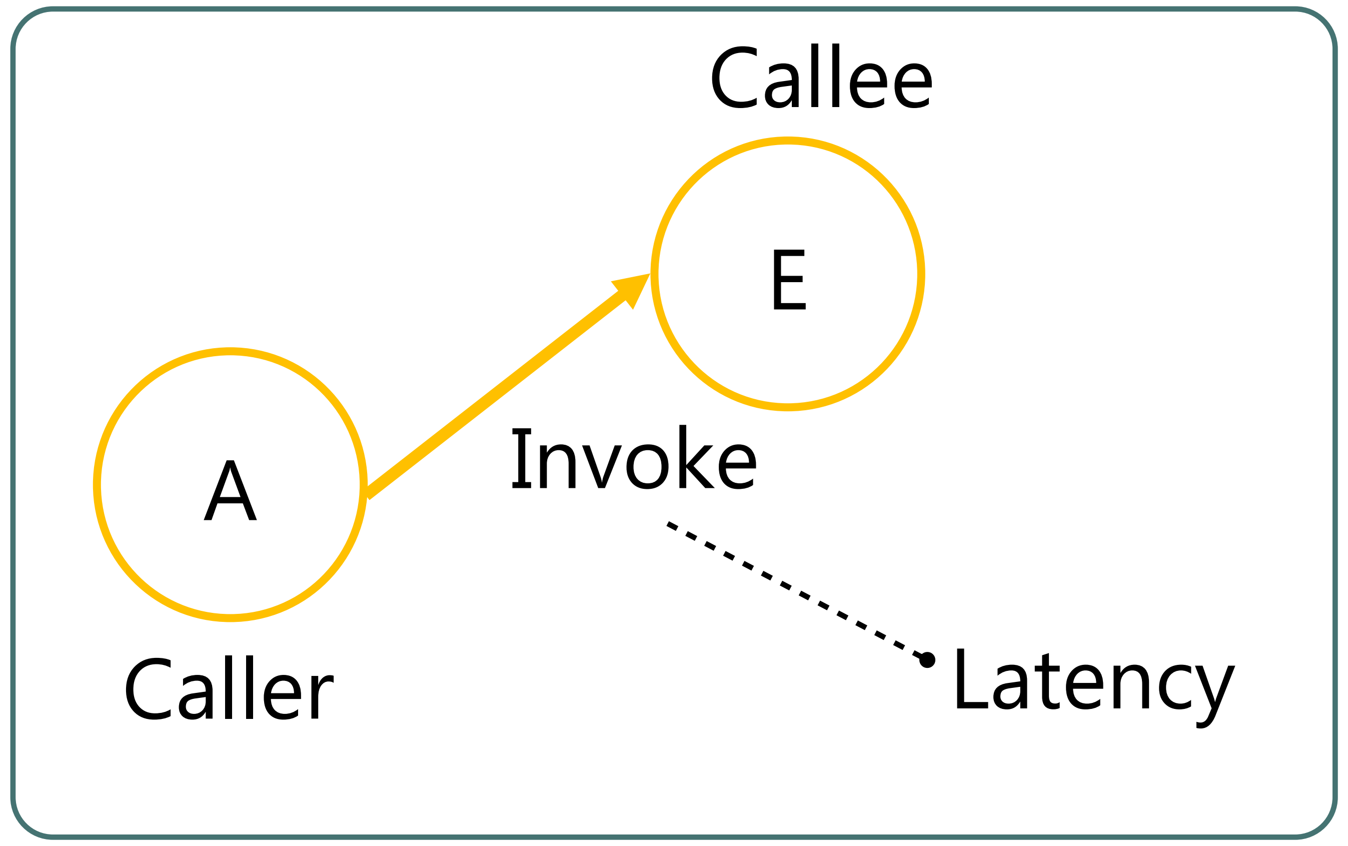


Terminologies

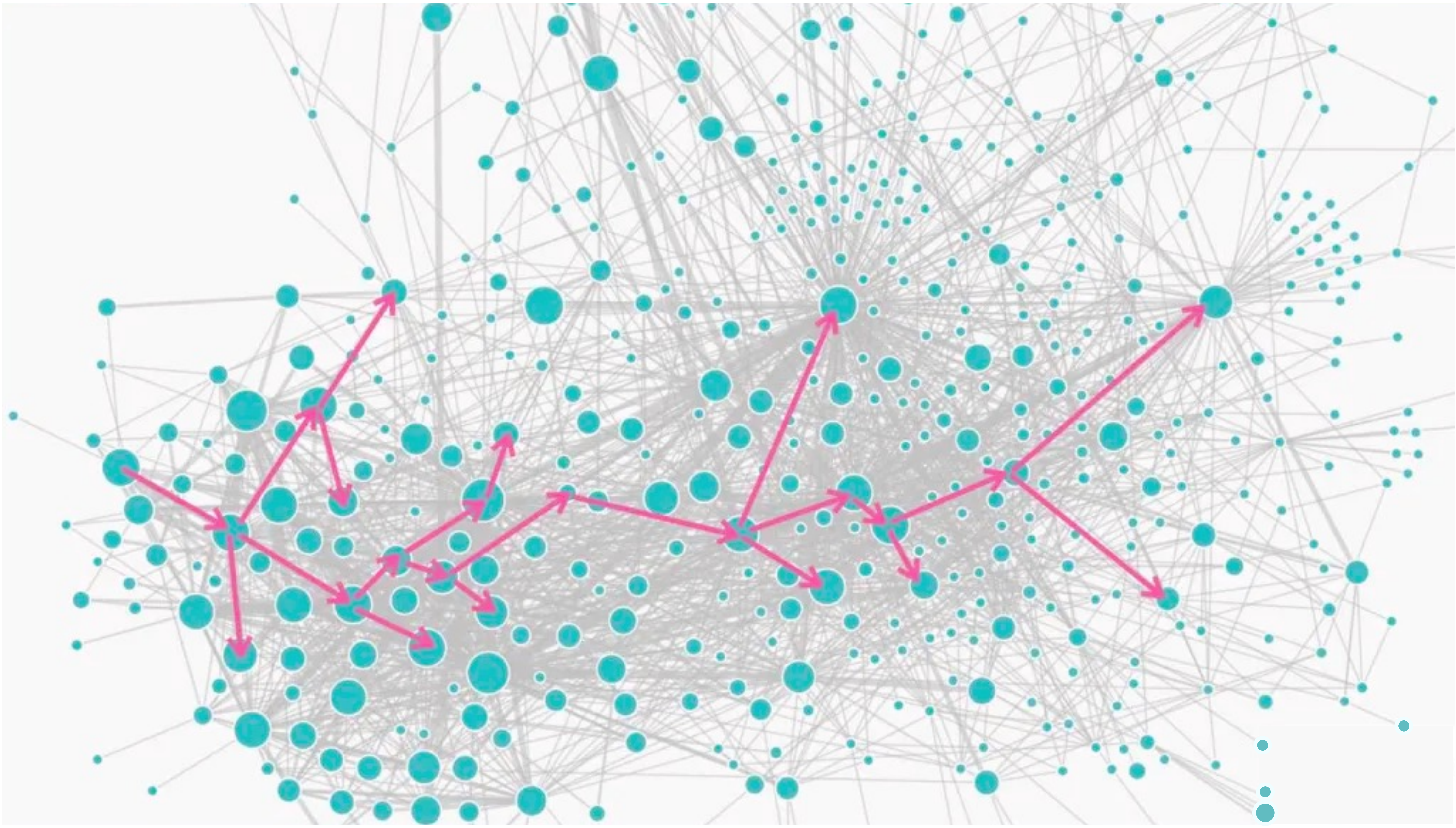


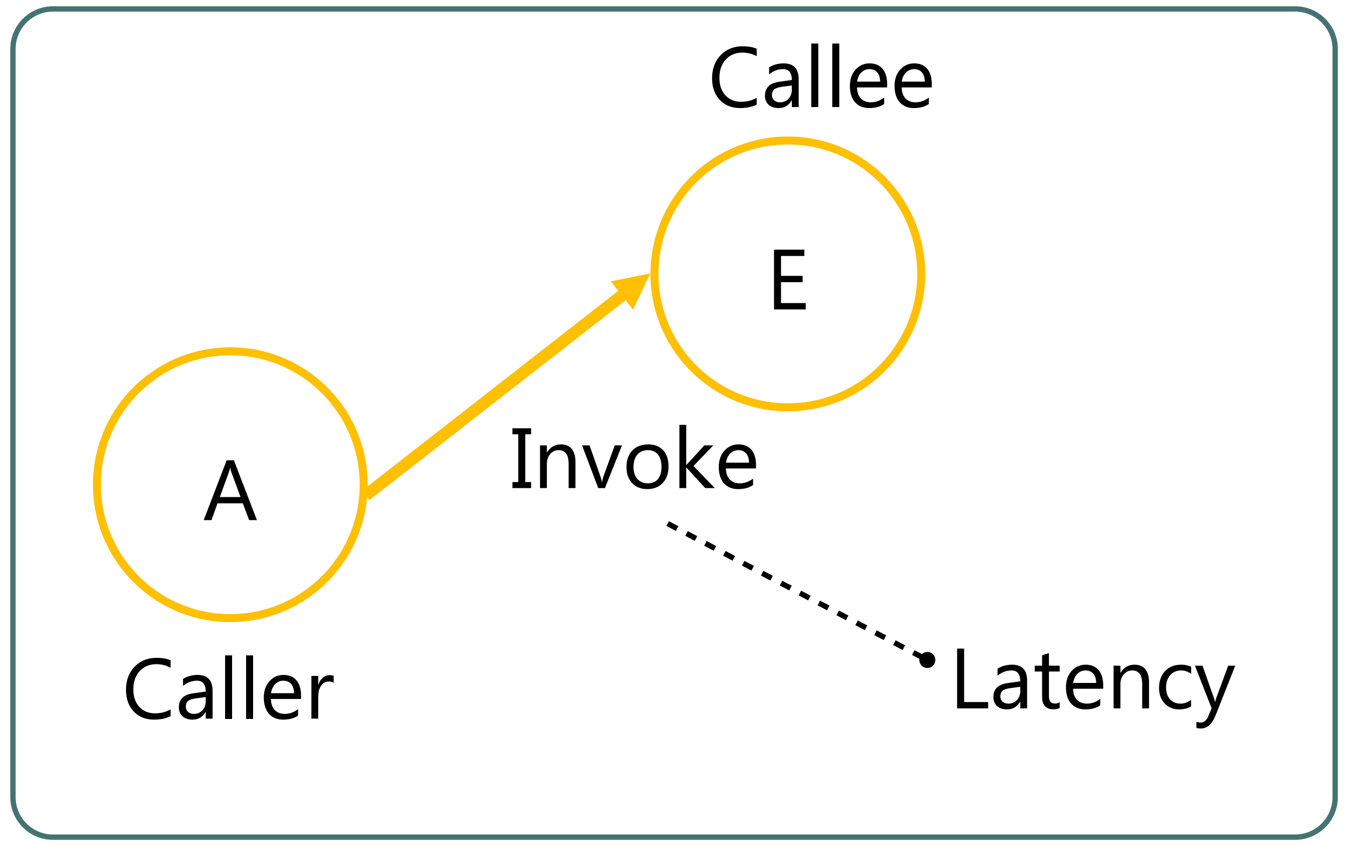
Terminologies





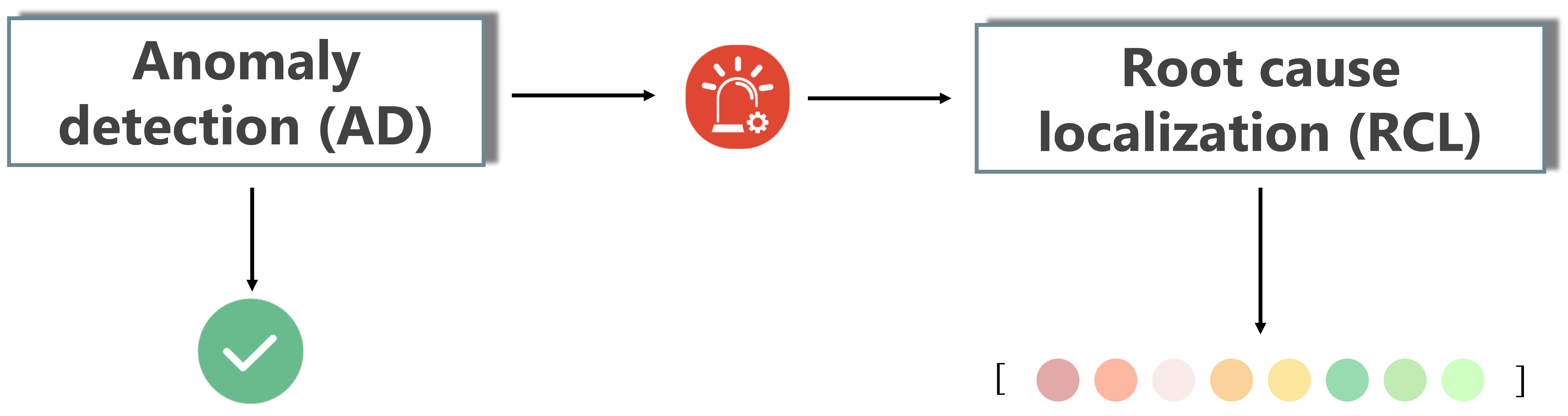
Terminologies





02 MOTIVATION

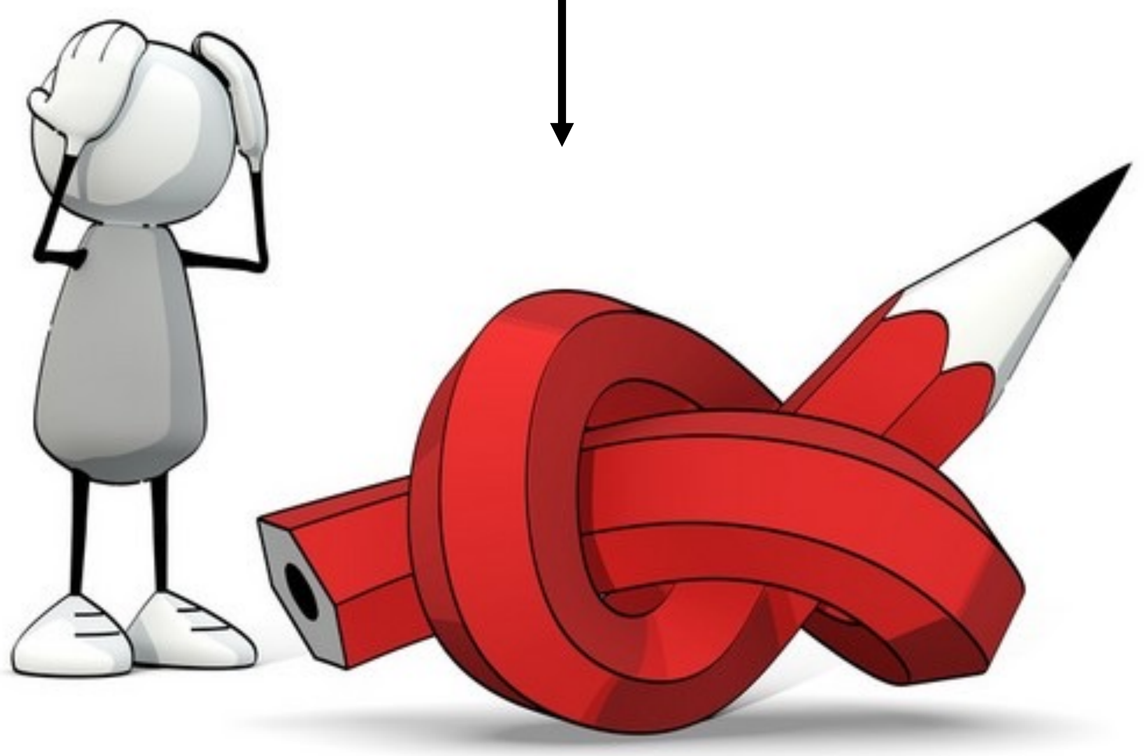
Inaccurate AD Results Limits RCL's Accuracy



Inaccurate AD Results Limits RCL’s Accuracy

Anomaly detection (AD)

Root cause localization (RCL)



Inaccurate AD Results Limits RCL's Accuracy

Current detectors attached with localizers cannot deliver satisfying accuracy.

Three main kinds of **RCL-oriented** anomaly detectors:

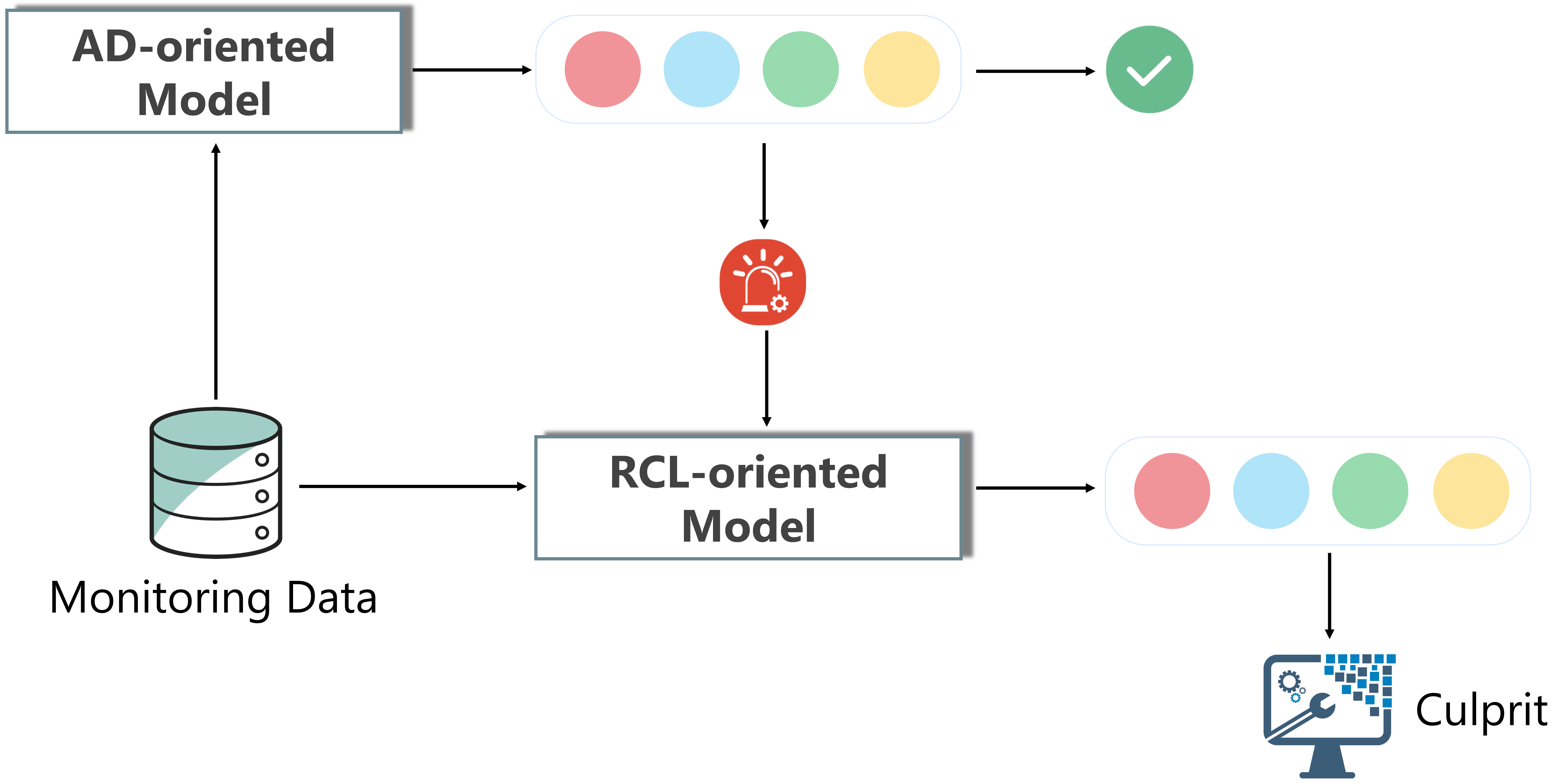
- Statistical tools (e.g., N-sigma)
- Feature engineering + Machine Learning (e.g., OC-SVM)
- SPOT (based on Extreme Value Theory)

COMPARISON OF COMMON ANOMALY DETECTORS

	N-sigma	FE+ML	SPOT
FOR	0.632	0.830	0.638
FDR	0.418	0.095	0
#Infer/ms	0.207	1.361	549.169

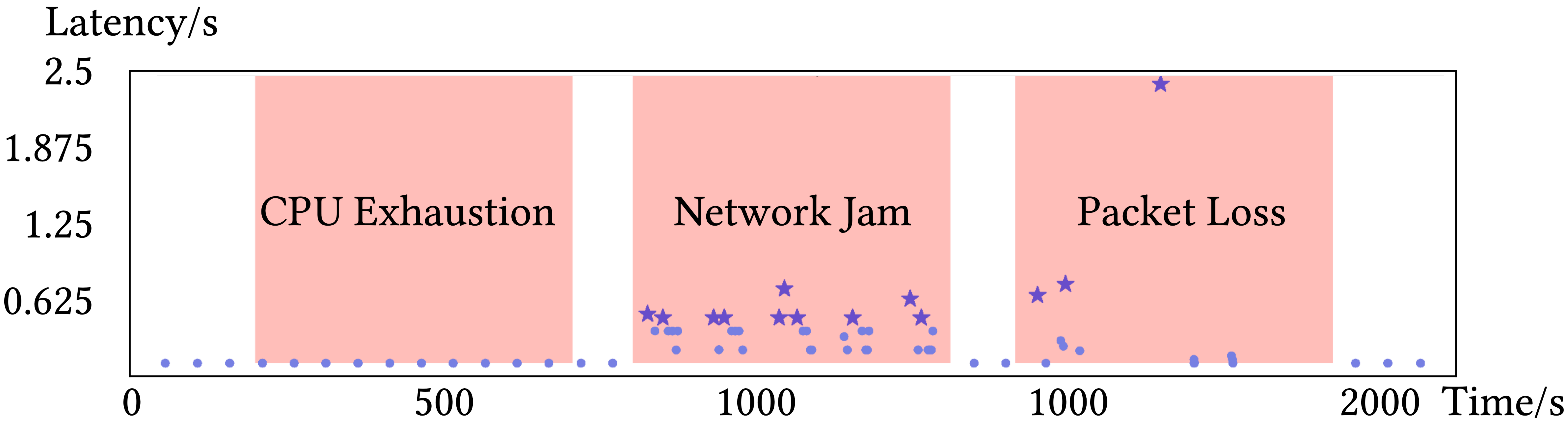
$$FOR = \frac{FN}{FN+TN}, FDR = \frac{FP}{FP+TN}$$

Disconnection in two closely related tasks



Consider data besides traces

Traces are insufficient to reveal all potential faults despite their wide usage.



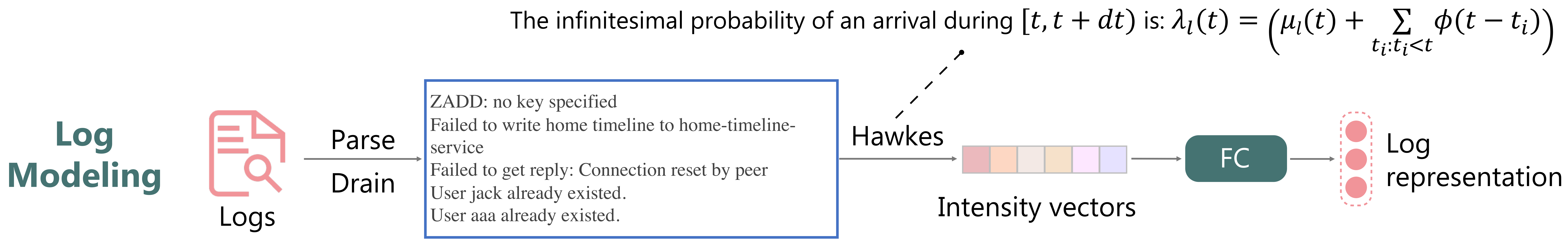
For example, network-related faults incur obvious anomalies in latency of "travel", but the CPU exhaustion fault does not.

03 MOTHODOLOGY

Overview

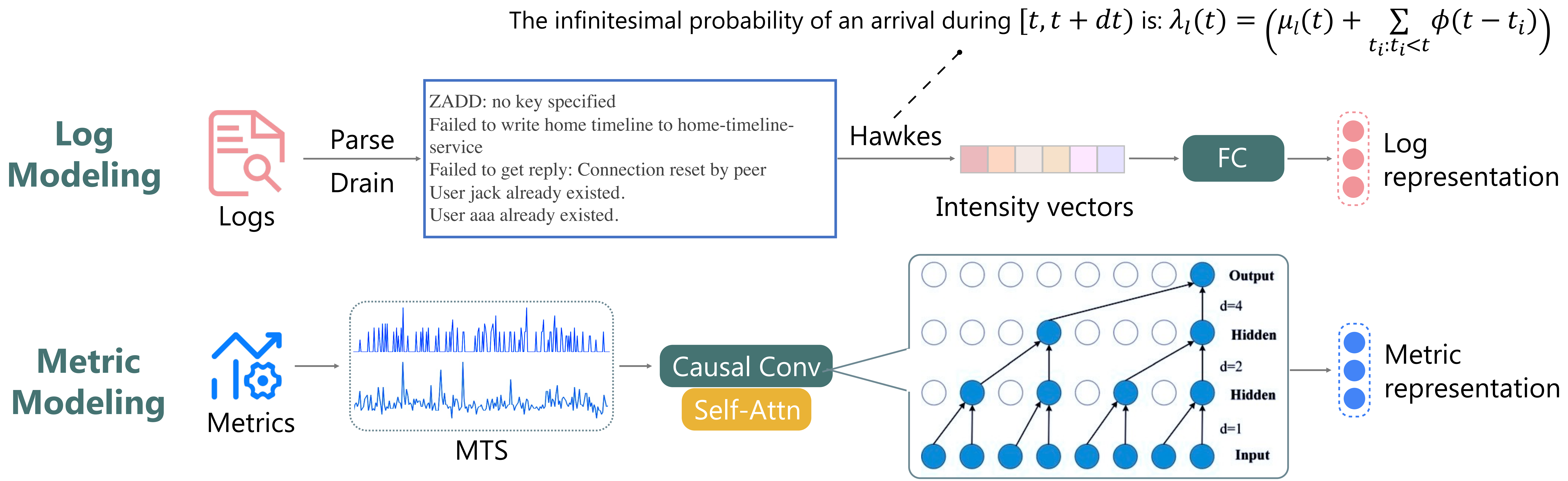

1 Modal-wise Learning
2 Dependency-aware Status Learning
3 Detection & Localization
Logs
Parse
Hawkes
Intensity vectors
FC
Metrics
MTS
Causal Conv
Traces
Latency
Causal Conv
Gated Fusion
GAT
Dependency Graph
Status Representation
Detector
Normal?
Yes
No
Localizer
$P_{[1:M]}$
Culprit
Joint Learning

1 Modal-wise Learning



Parsing → Estimating → Embedding

1 Modal-wise Learning



1 Modal-wise Learning

The infinitesimal probability of an arrival during $[t, t+dt)$ is: $\lambda_l(t) = \left(\mu_l(t) + \sum_{t_i:t_i<t} \phi(t - t_i)\right)$

Log Modeling

Logs

Parse

Drain

ZADD: no key specified
Failed to write home timeline to home-timeline-service
Failed to get reply: Connection reset by peer
User jack already existed.
User aaa already existed.

Hawkes

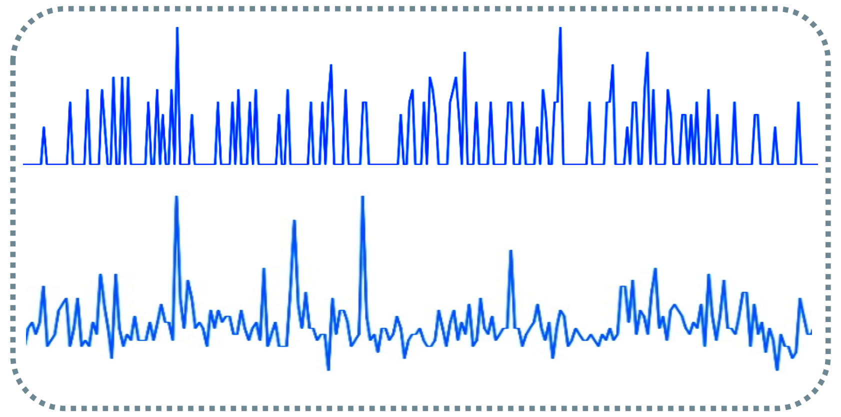
Intensity vectors

FC

Log representation

Metric Modeling

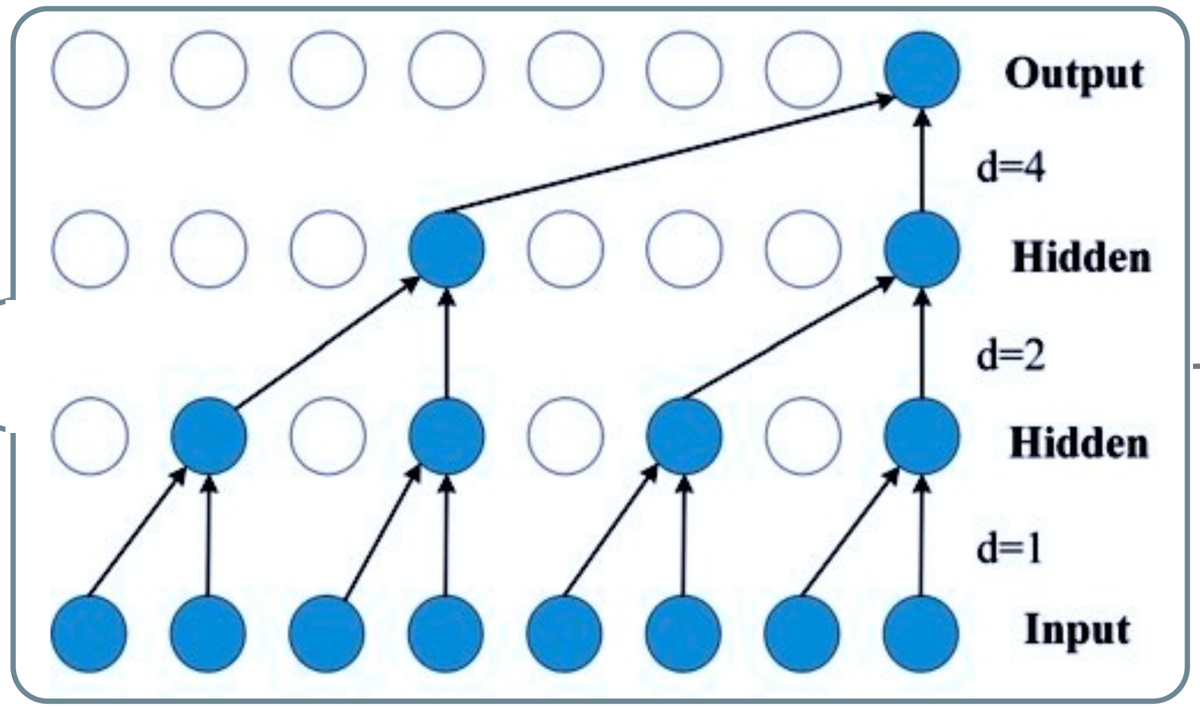
Metrics



MTS

Causal Conv

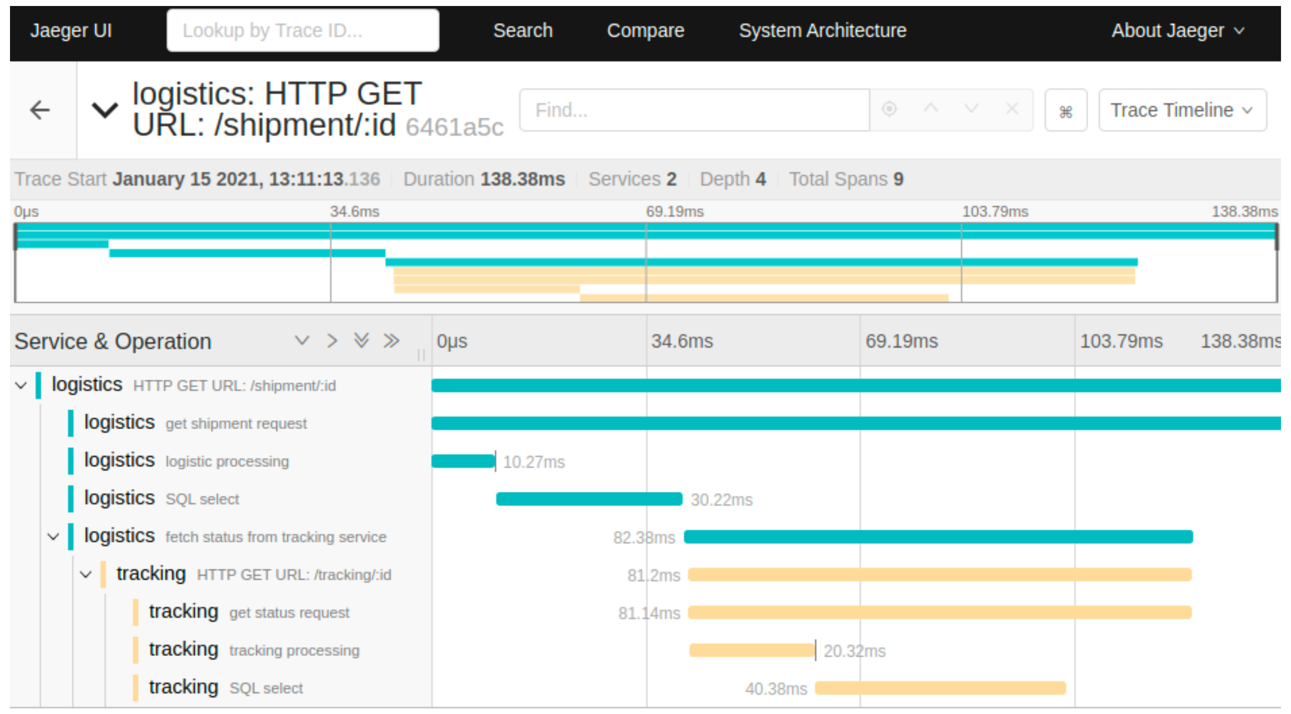
Self-Attn

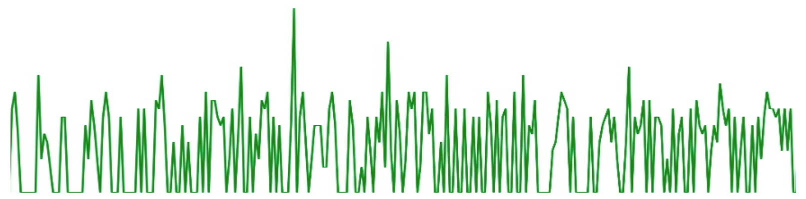


Metric representation

Latency Modeling (of traces)

Traces



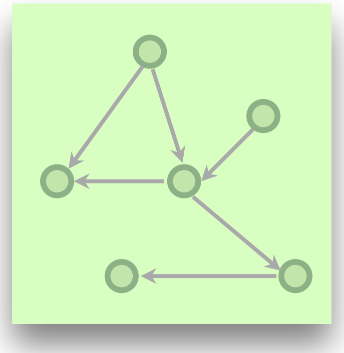


Latency

Causal Conv

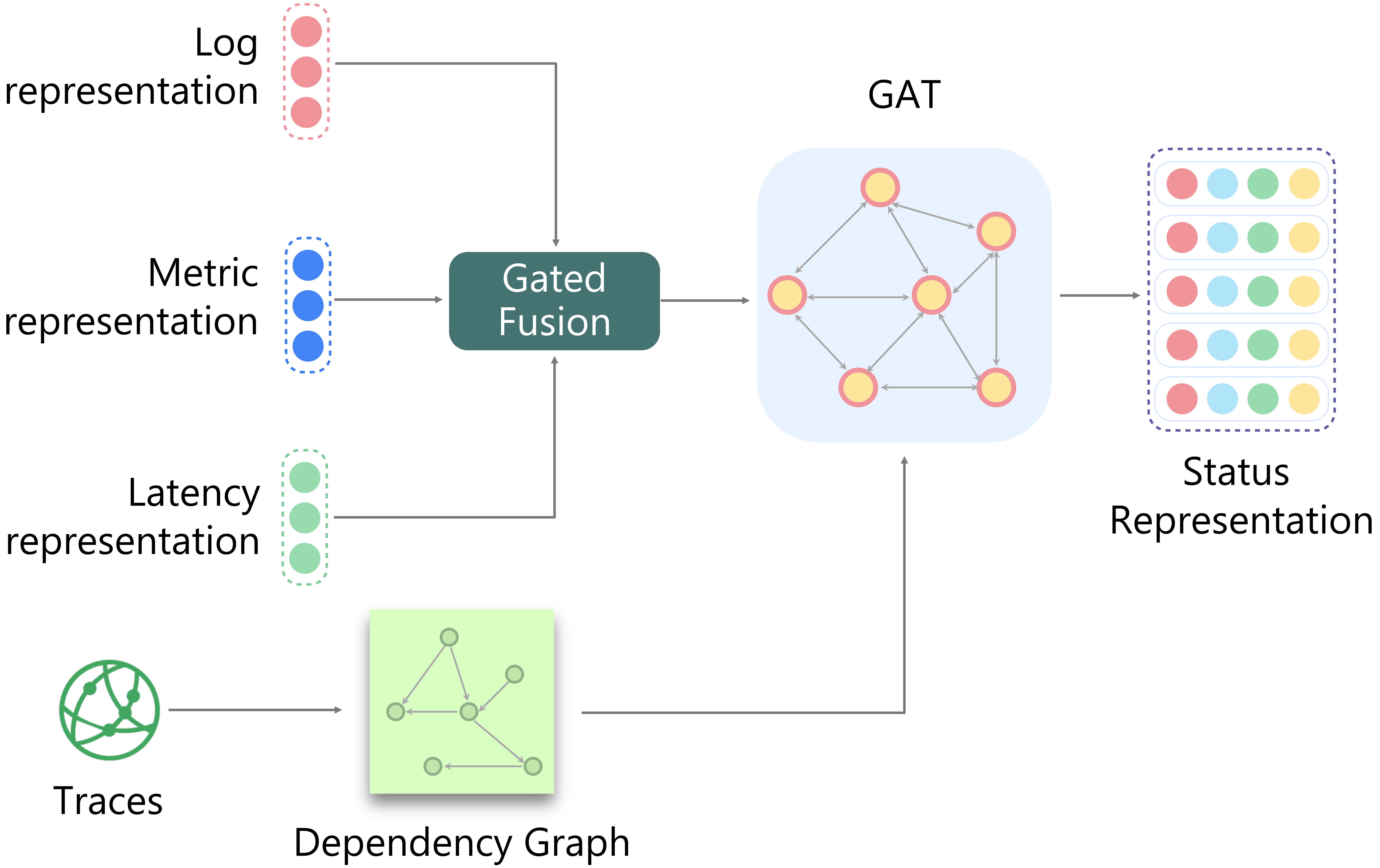
Self-Attn

Latency representation

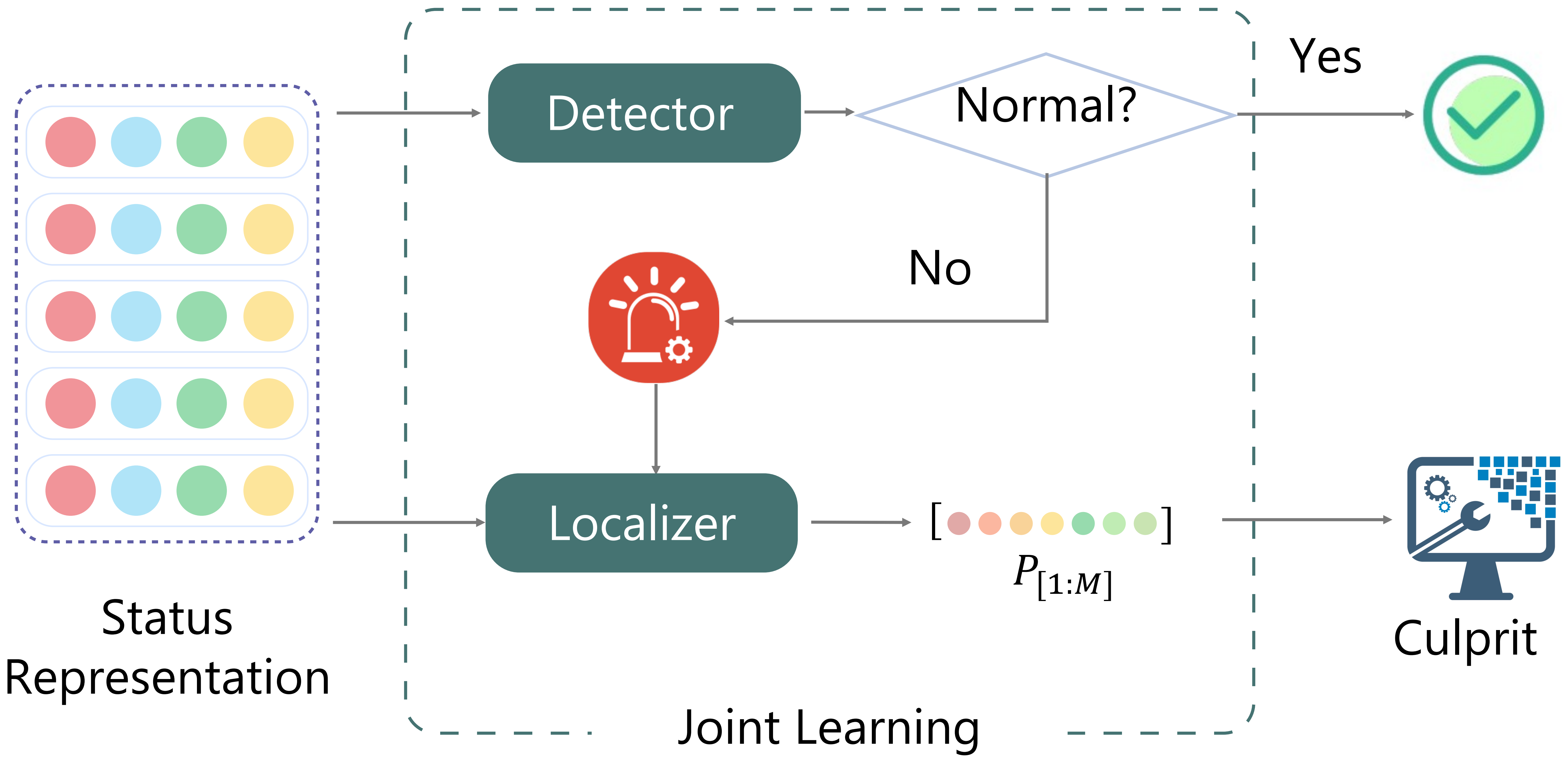


Dependency graph

2 Dependency-aware Status Learning



3 Detection & Localization




Eadro
Overview
Details
Settings
User: Tim
tx/b
rx/b
5000
4000
3000
2000
1000
9:28
9:30
9:32
9:34
9:36
9:38
Frequency: 2T
ts-order-service
Dependency
order
Frontend
Log File Download
cpu_sys
cpu_total
cpu_user
0.01
0.008
0.005
0.003
0
9:28
9:30
9:32
9:34
9:36
9:38
Latency
8.0
6.0
4.0
2.0
0.0
9:28
9:30
9:32
9:34
9:36
9:38
Root Cause List
Service
Probability
order
0.972
preserve
0.087
security
0.011
frontend
0.010
Trace File Download

04 EVALUATION

Evaluation

RQ1: How effective is Eadro in anomaly detection?

RQ2: How effective is Eadro in root cause localization?

RQ3: How much does each data source contribute?

RQ1: Effectiveness in AD

Eadro improves F1-score by 53.82%~92.68% compared to baselines and 3.13%~25.32% compared to derived methods.

PERFORMANCE COMPARISON FOR ANOMALY DETECTION

Approaches	$\mathcal{TT}$			$\mathcal{SN}$		
	F1	*Rec*	*Pre*	*F1*	*Rec*	*Pre*
TraceAnomaly	0.486	0.414	0.589	0.539	0.468	0.636
MultimodalTrace	0.608	0.576	0.644	0.676	0.632	0.726
MS-RF-AD	0.817	0.705	0.971	0.773	0.866	0.700
MS-SVM-AD	0.787	0.678	0.938	0.789	0.770	0.808
MS-LSTM	0.967	0.997	0.940	0.948	0.959	0.937
MS-DCC	0.965	0.993	0.938	0.948	0.962	0.934
Eadro	**0.989**	**0.995**	**0.984**	**0.986**	**0.996**	**0.977**

RQ2: Effectiveness in RCL

Eadro increases Top-1 Hit Rate by 290%~5068% than baselines and 26.93%~66.16% than the derived methods.

PERFORMANCE COMPARISON FOR ROOT CAUSE LOCALIZATION

Approaches	$\mathcal{TT}$					$\mathcal{SN}$				
	HR@1	*HR@3*	*HR@5*	*NDCG@3*	*NDCG@5*	*HR@1*	*HR@3*	*HR@5*	*NDCG@3*	*NDCG@5*
TBAC	0.037	0.111	0.185	0.079	0.109	0.001	0.085	0.181	0.048	0.087
NetMedic	0.094	0.257	0.425	0.195	0.209	0.069	0.187	0.373	0.146	0.218
MonitorRank	0.086	0.199	0.331	0.142	0.196	0.068	0.118	0.221	0.095	0.137
CloudRanger	0.101	0.306	0.509	0.218	0.301	0.122	0.382	0.629	0.269	0.370
DyCause	0.231	0.615	0.808	0.448	0.607	0.273	0.636	0.727	0.301	0.353
MS-RF-RCL	0.637	0.922	<u>0.970</u>	0.807	0.827	0.704	<u>0.908</u>	<u>0.970</u>	0.825	<u>0.851</u>
MS-SVM-RCL	0.541	0.908	0.944	0.814	0.820	0.614	0.838	<u>0.955</u>	0.741	0.790
MS-LSTM	0.756	0.930	<u>0.969</u>	0.859	0.877	0.757	<u>0.884</u>	<u>0.907</u>	0.834	<u>0.844</u>
MS-DCC	0.767	0.938	0.972	0.870	0.882	0.789	0.968	0.985	0.898	0.905
Eadro	**0.990**	**0.992**	**0.993**	**0.994**	**0.994**	**0.974**	**0.988**	**0.991**	**0.982**	**0.983**

RQ3: Usefulness of Each Data Source

All of the involved data sources can all contribute to Eadro, and traces contribute the most.

EXPERIMENTAL RESULTS OF THE ABLATION STUDY

Variants	$\mathcal{TT}$			$\mathcal{SN}$		
	HR@1	*HR@5*	*F1*	*HR@1*	*HR@5*	*F1*
Eadro	**0.990**	**0.993**	**0.989**	**0.974**	**0.991**	**0.986**
Eadro w/o $\mathcal{L}$	0.926	0.993	0.964	0.902	0.954	0.972
Eadro w/o $\mathcal{M}$	0.776	0.962	0.960	0.684	0.947	0.974
Eadro w/o $\mathcal{T}$	0.785	0.930	0.945	0.627	0.930	0.957
Eadro w/o $\mathcal{G}$	0.803	0.982	0.970	0.791	0.960	0.946

THANKS

Presenter: Cheryl LEE

Arise Lab

Full Paper

My Homepage