Graph-based Trace Analysis for Microservice Architecture Understanding and Problem Diagnosis

ESEC/FSE 2020 Industry

Agenda

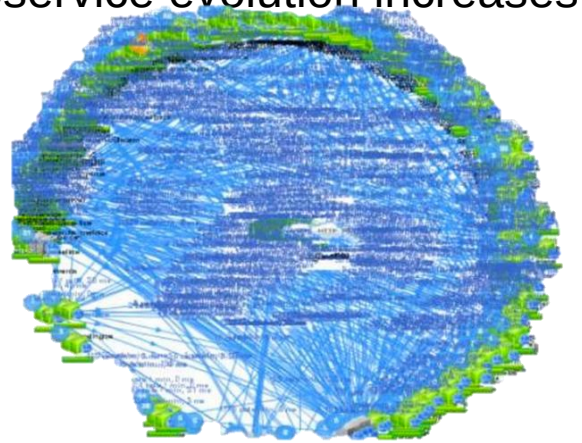
- Background
- System
- Case study
- Evaluation

Background

Infrastructure scale keeps increasing

source: Cisco

Microservice evolution increases complexity

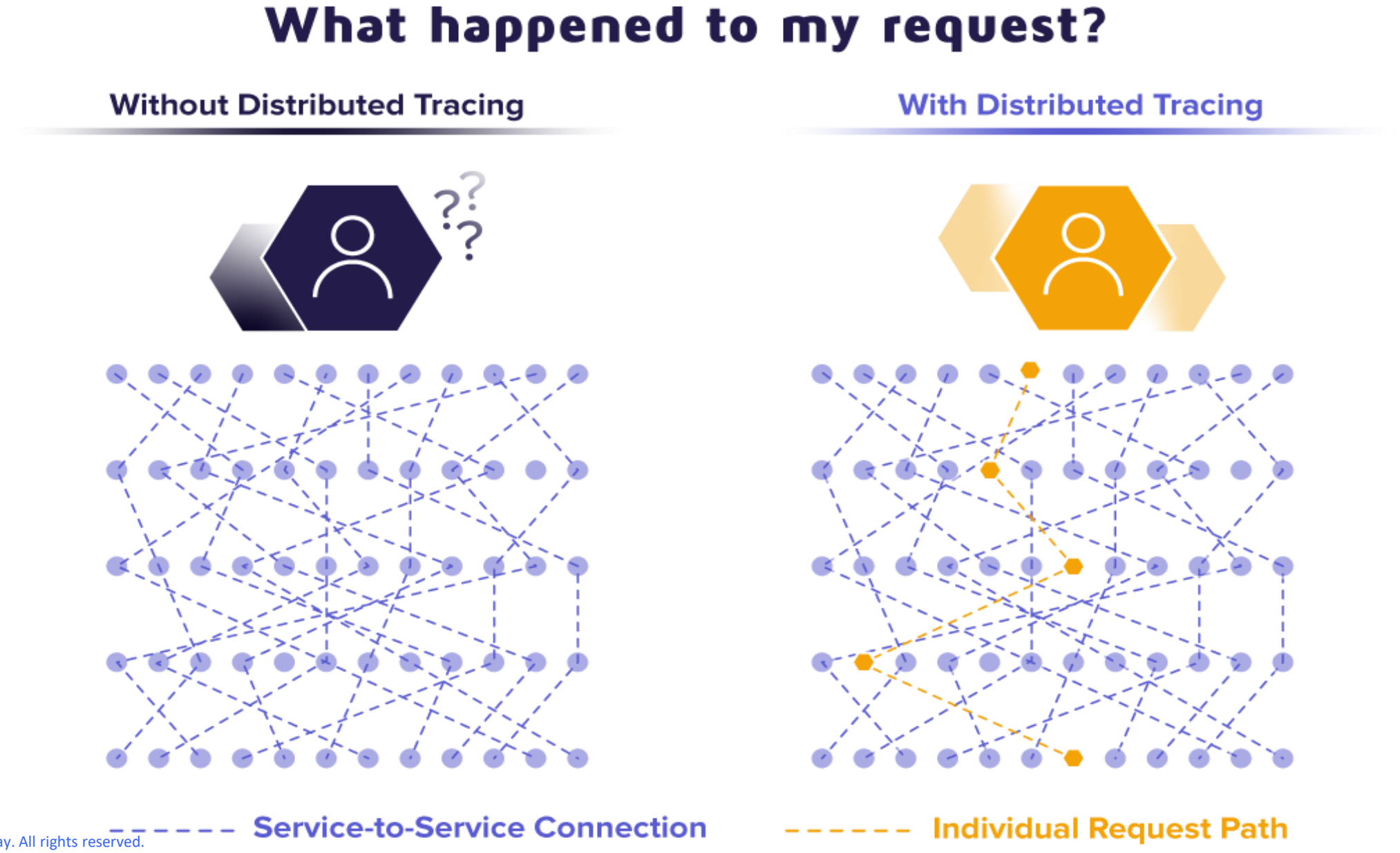


source: Netflix

Harder to operate

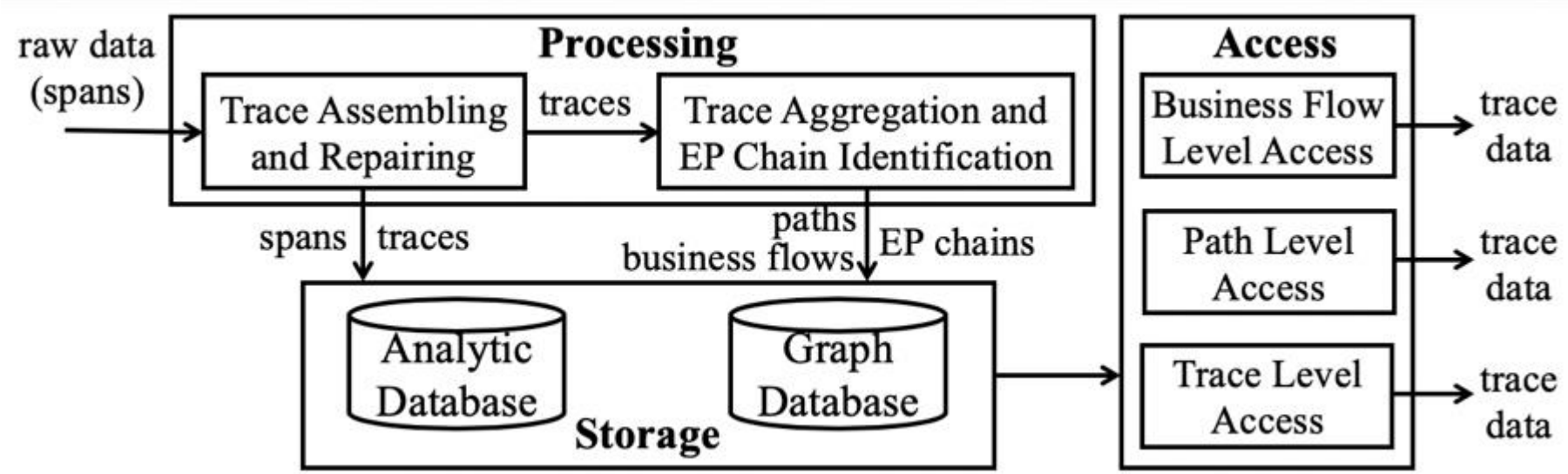
Too many alerts

Background - Distributed Tracing

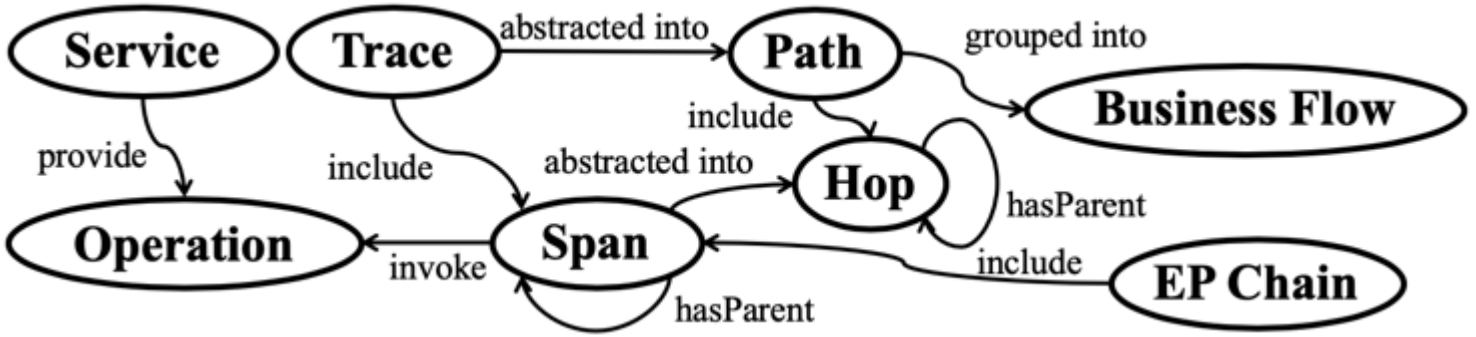


GMTA

System Overview

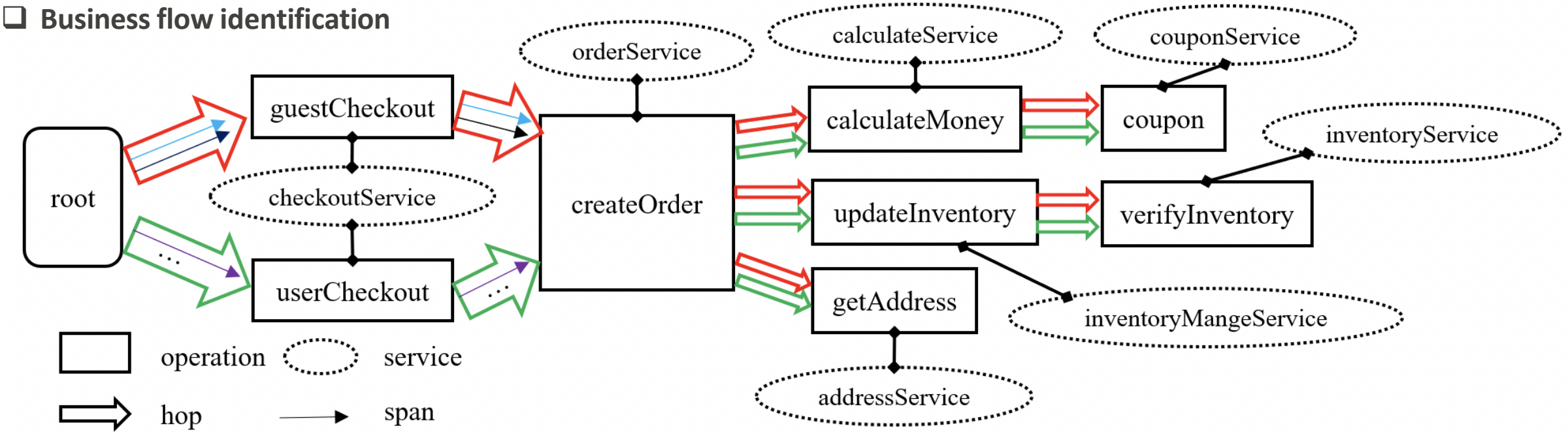


Conceptual Model of Graph based Trace Data Representation



Data Processing

- **Trace Assembling and Repairing**
  - broken trace & bad instrument
  - trace assembling and error propagation chain assembling
- **Path identification**
- **Business flow identification**



Data Processing

Definition

Path is a rooted directed graph:

$$G = (V, E)$$

edge-property(PathID):

$$\mathcal{P}(x, y) = \{f(root)\} \quad x, y \in V \;\&\&\; (x, y) \in E$$

$$f(v) = hash(v) + distance(v, root) \times \sum_{u \in V} e(v, u) \times f(u)$$

Merge paths in a graph

$$Graph = G1 \oplus G2 = (V1 \cup V2, E1 \cup E2)$$

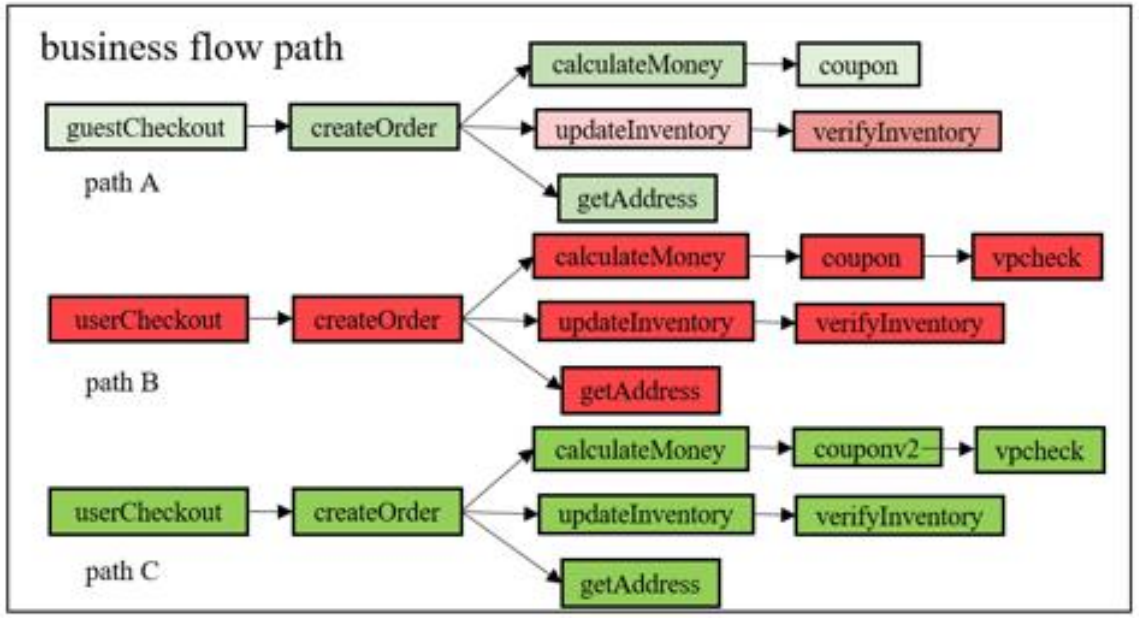
$$\mathcal{P}_{G1 \oplus G2}(x, y) = \mathcal{P}_{G1}(x, y) \cup \mathcal{P}_{G2}(x, y)$$

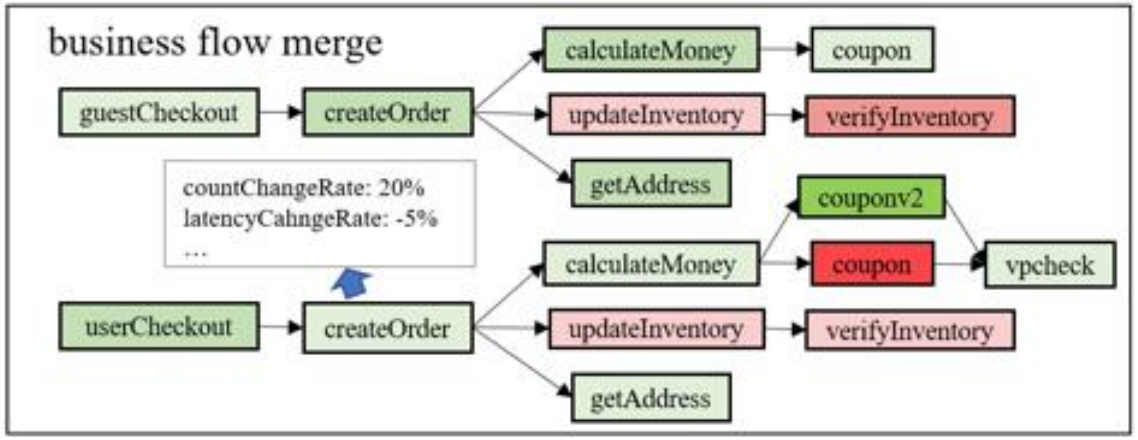
Business flow identification

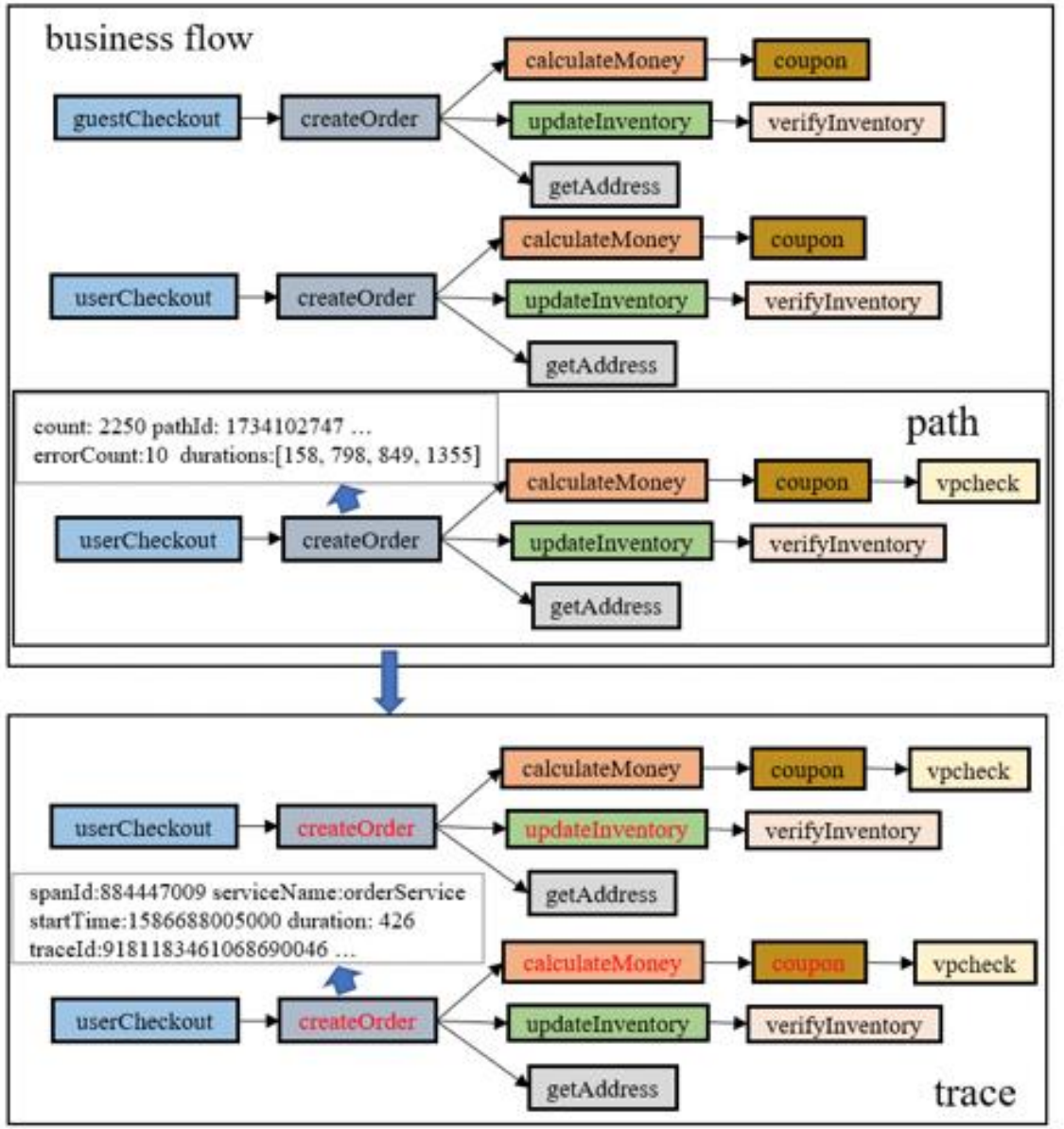
The set of paths belong to business flow going through vertex x, y , could be defined as

$$(\cup_{u \in V} \mathcal{P}_G(x, u) \cup \cup_{u \in V} \mathcal{P}_G(u, x)) \cap (\cup_{v \in V} \mathcal{P}_G(y, v) \cup \cup_{v \in V} \mathcal{P}_G(v, y))$$

Architecture Understanding

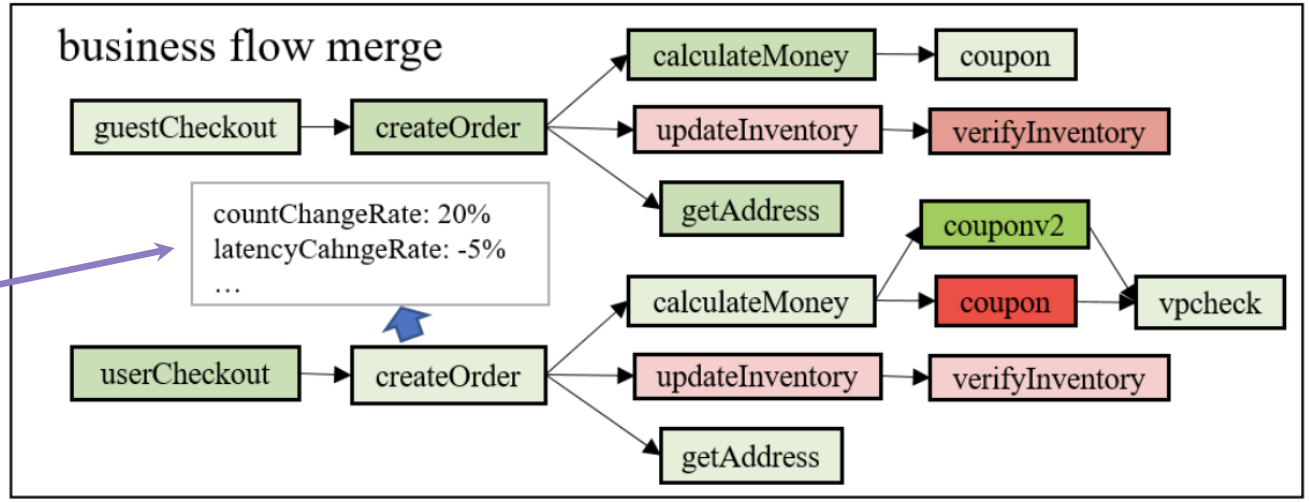




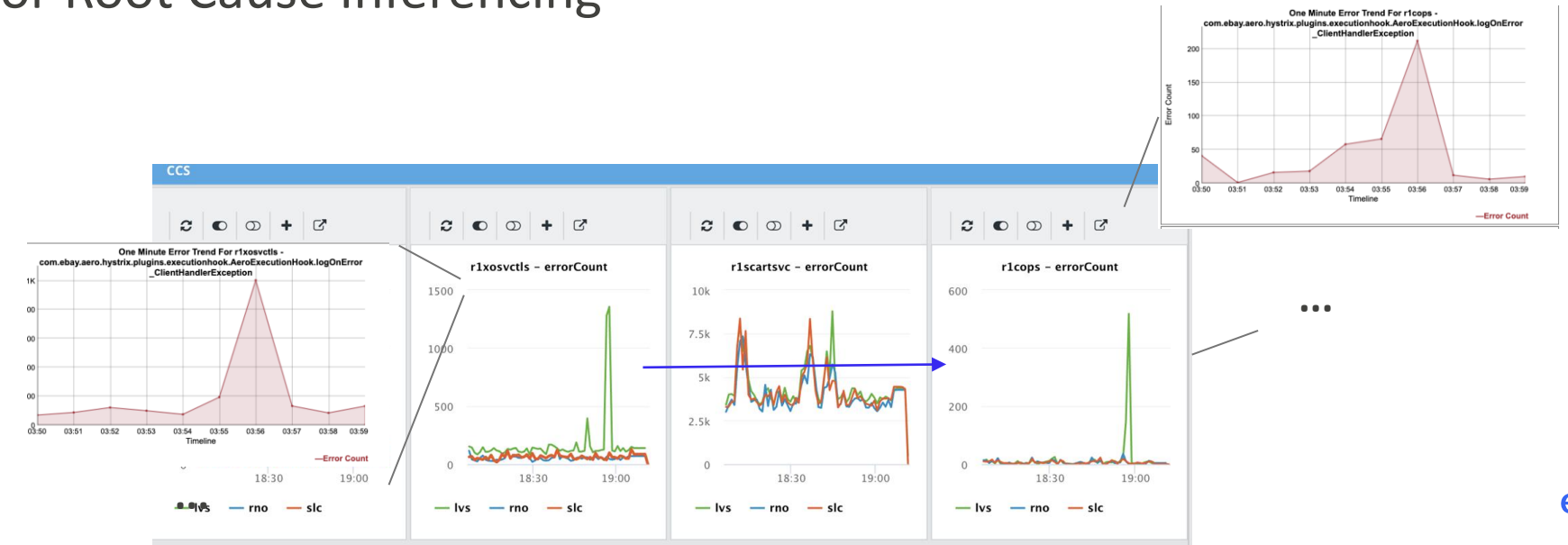


Problem diagnostics

Fault localization

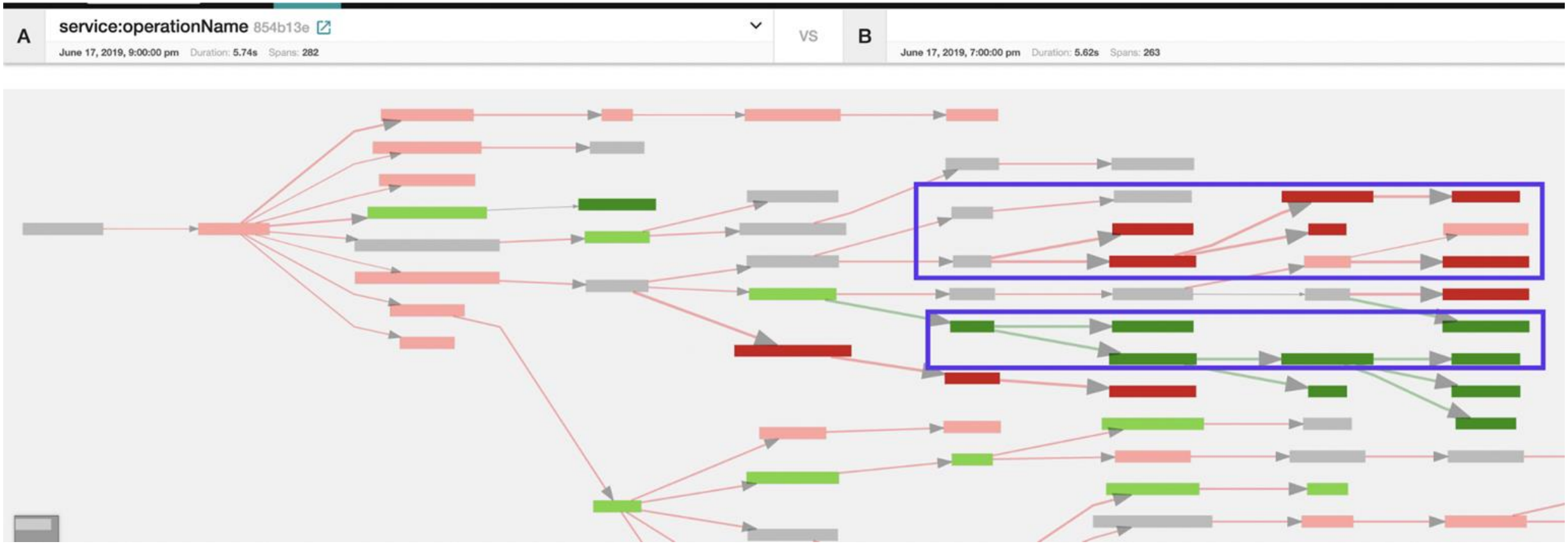


Error Monitor for Root Cause Inferencing



Case Study

Case study: Change impact evaluation

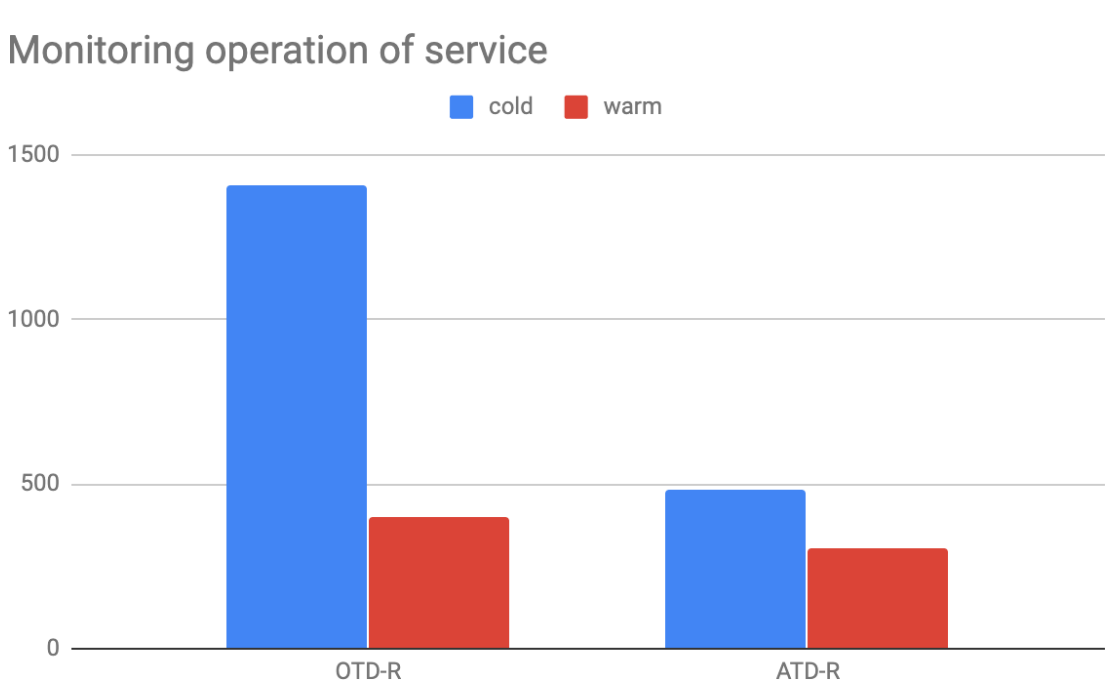


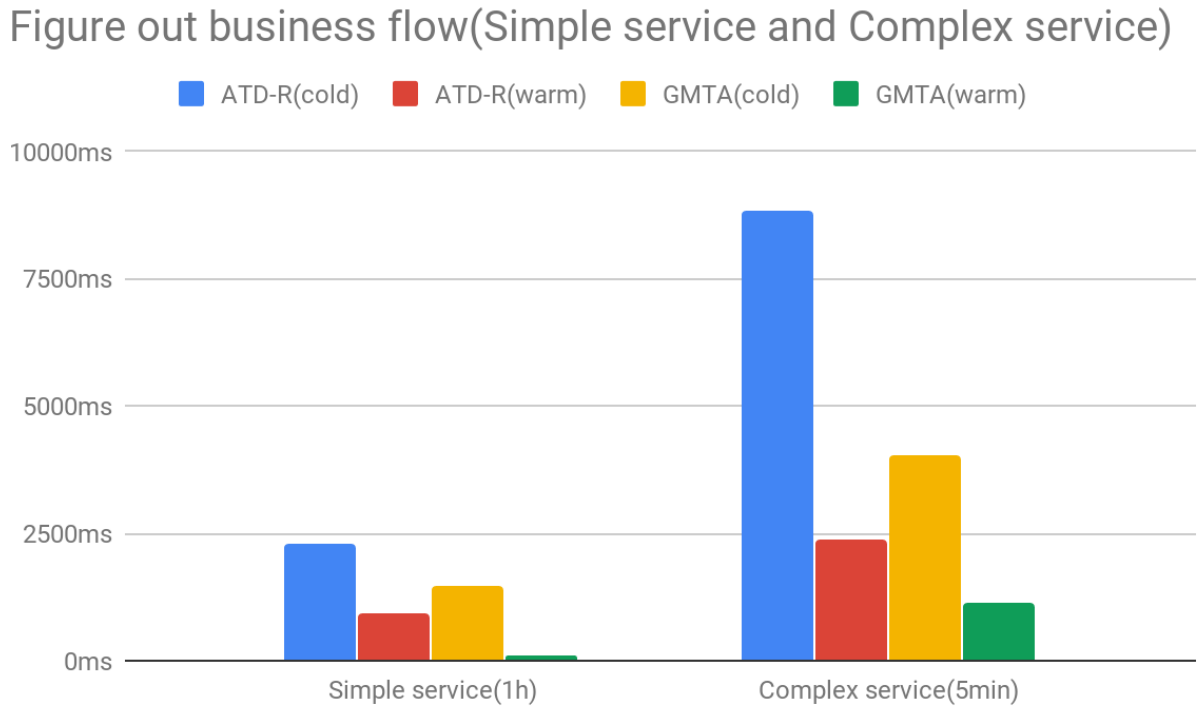
Evaluation:Effectiveness

Approach	Traces	Paths	Business Flow	Graph-based Storage
OTD-R	√	×	×	×
ATD-R	√	√	×	×
GMTA	√	√	√	√

Use case	OTD-R	ATD-R	GMTA
Single Trace Query	√	√	√
Operation Metrics Query	√	√	√
Single Path Query	x	√	√
Error Propagation Chain Query	x	√	√
Business Flow Generation	x	√	√
Service Dependency analysis	x	x	√

Evaluation:Efficiency





Thanks