Causal Inference-Based Root Cause Analysis for Online Service Systems with Intervention Recognition

Mingjie Li, Zeyan Li, Kanglin Yin, Xiaohui Nie, Wenchi Zhang, Kaixin Sui, Dan Pei

清華大學 Tsinghua University

必示 BIZSEER

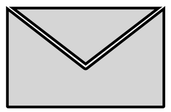
Background

Background
Analysis
Methodology
Evaluation
Conclusion

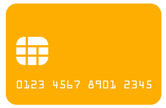
Background

Online Service Systems (OSS)

Online Services



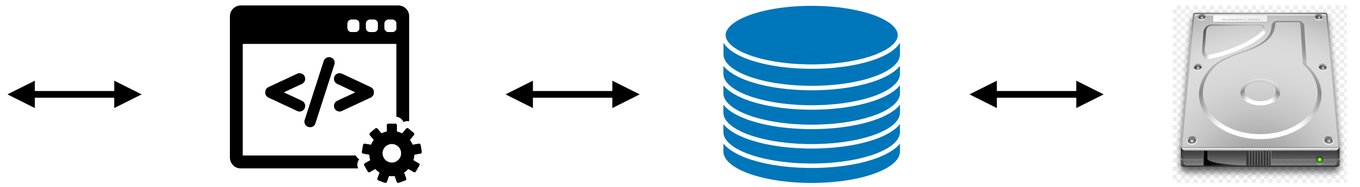
Social Networks



Online Shopping

Search Engine

.....

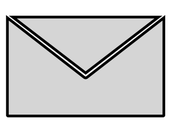


Web Server Database Storage

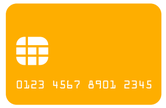
Background

Online Service Systems (OSS)

Online Services



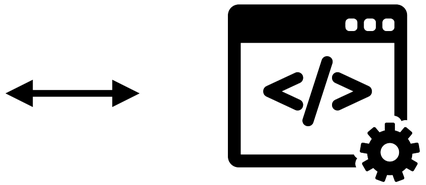
Social Networks



Online Shopping

Search Engine

.....



Web Server

queries per second
response time
success rate
.....

Database

average active session
log file sync
logic read per second
.....

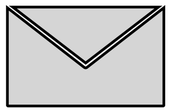
Storage

disk utilization
IO per second
IO wait time
.....

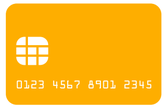
Background

Online Service Systems (OSS)

Online Services



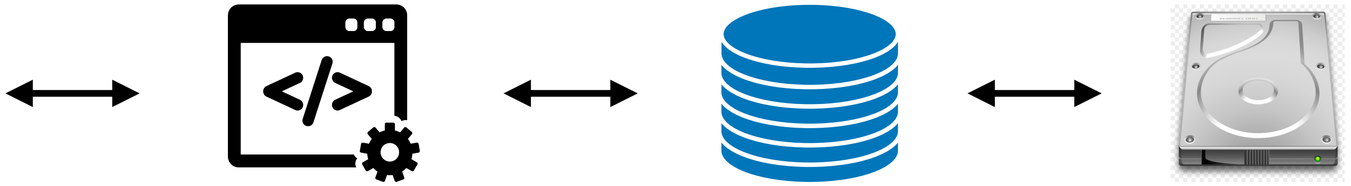
Social Networks



Online Shopping

Search Engine

.....



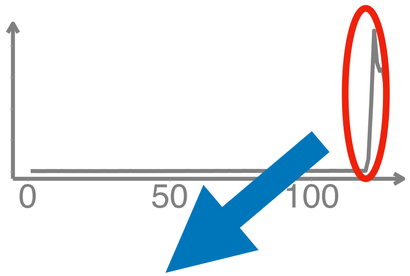
Web Server | Database | Storage

queries per second
response time
success rate
.....

average active session
log file sync
logic read per second
.....

disk utilization
IO per second
IO wait time
.....

Metrics are the most widely available data, often in the form of a time series

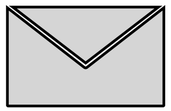


Time	09:23	09:24	09:25	09:26	09:27
Value	302	4095	22142	44936	34745

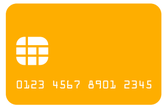
Background

Root Cause Analysis (RCA)

Online Services



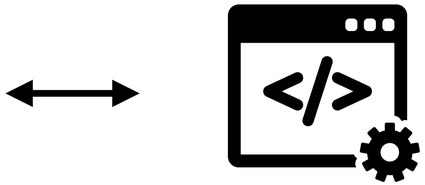
Social Networks



Online Shopping

Search Engine

.....



Web Server

queries per second
response time
success rate
.....

Database

average active session
log file sync
logic read per second
.....

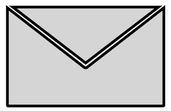
Storage

disk utilization
IO per second
IO wait time
.....

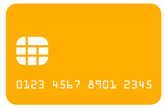
Background

Root Cause Analysis (RCA)

Online Services



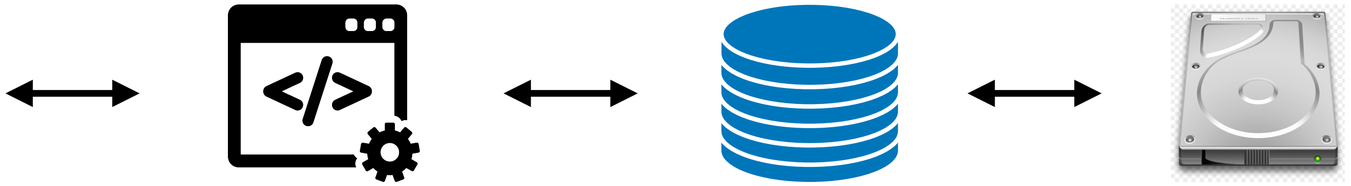
Social Networks



Online Shopping

Search Engine

.....



Web Server | Database | Storage

queries per second
response time
success rate
.....

average active session
log file sync
logic read per second
.....

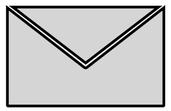
disk utilization
IO per second
IO wait time
.....

A fault may propagate in the system

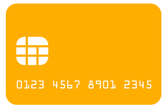
Background

Root Cause Analysis (RCA)

Online Services



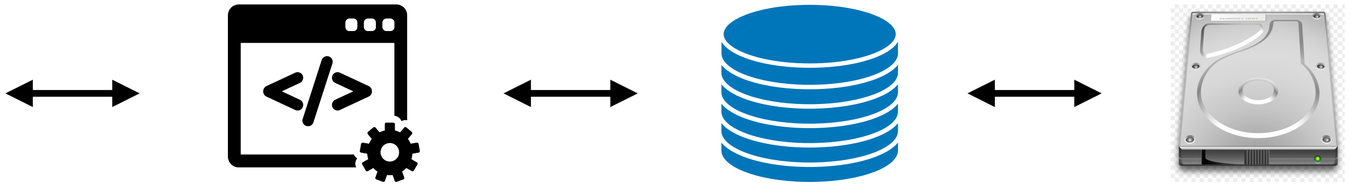
Social Networks



Online Shopping

Search Engine

.....



Web Server	Database	Storage
queries per second	average active session	disk utilization
response time	log file sync	IO per second
success rate	logic read per second	IO wait time
.....	.....	.....

Recognize a small set of root cause indicators to accelerate failure mitigation

Background

Related Works

Background

Related Works

Invariant Network-based
KDD'16, ICDM'17

$$y(t) = 2x(t-1) + x(t-2) + 1$$

x ——— y

Background

Related Works

Invariant Network-based
KDD'16, ICDM'17

$$y(t) = 2x(t-1) + x(t-2) + 1$$

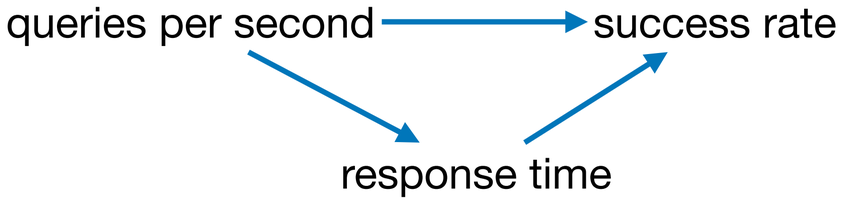
Background

Related Works

Invariant Network-based
KDD'16, ICDM'17

Random Walk-based
CCGRID'18, IWQoS'20, WWW'20, ASE'21

$$y(t) = 2x(t-1) + x(t-2) + 1$$



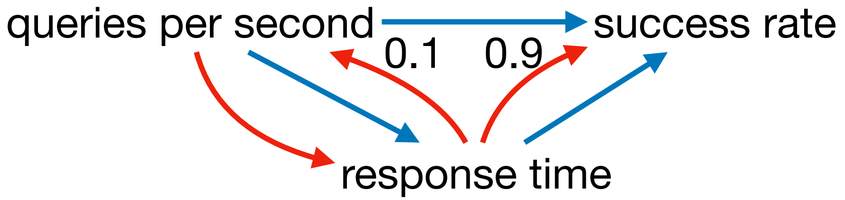
Background

Related Works

Invariant Network-based
KDD'16, ICDM'17

Random Walk-based
CCGRID'18, IWQoS'20, WWW'20, ASE'21

$$y(t) = 2x(t-1) + x(t-2) + 1$$



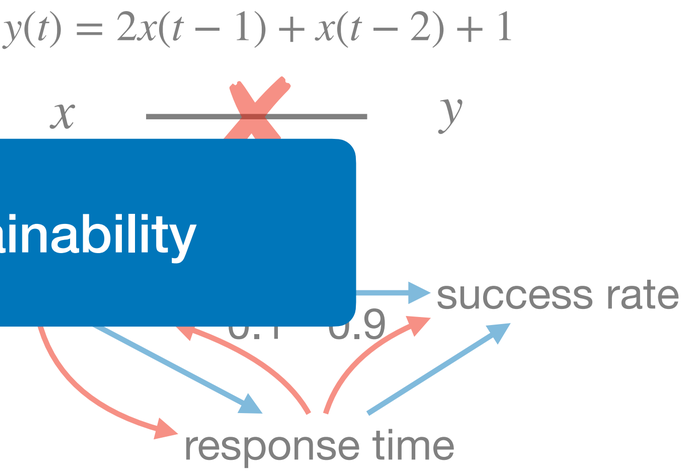
Background

Related Works

Invariant Network-based
KDD'16, ICDM'17

Random Walk-based
CCGRID'18, IWQoS'20, WWW'20, ASE'21

Lack explainability



Background

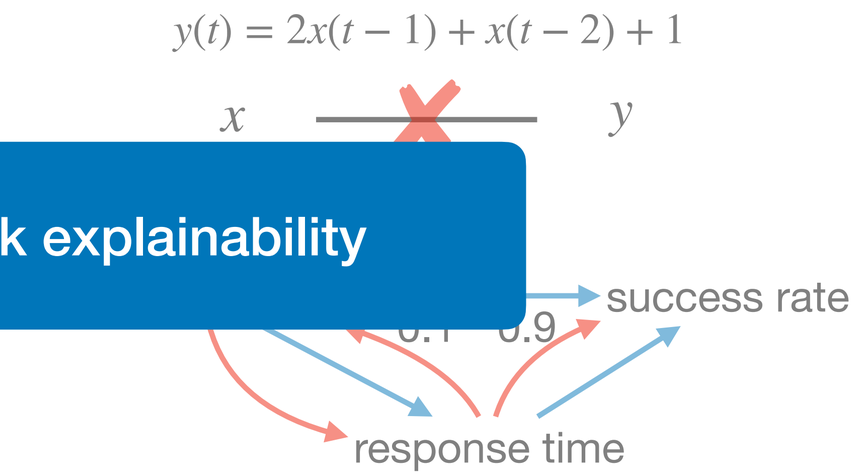
Related Works

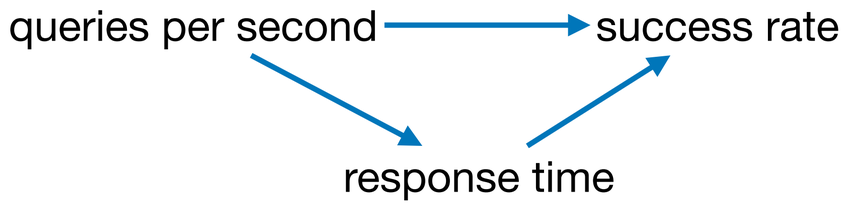
Invariant Network-based
KDD'16, ICDM'17

Random Walk-based
CCGRID'18, IWQoS'20, WWW'20, ASE'21

Search-based
INFOCOM'14, SOC'18, ICSE'21

Lack explainability





Background

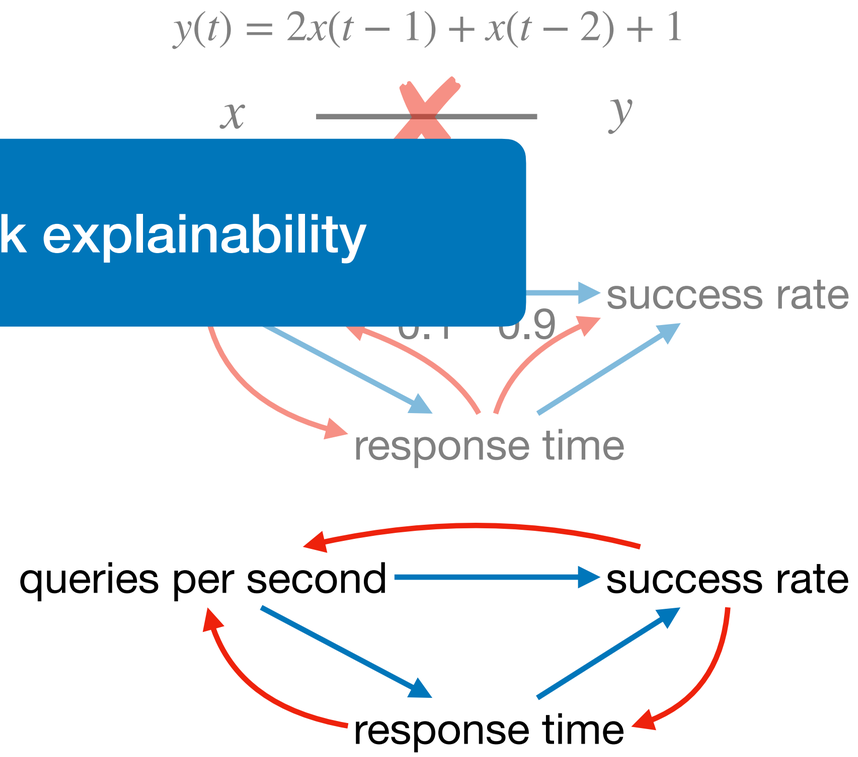
Related Works

Invariant Network-based
KDD'16, ICDM'17

Random Walk-based
CCGRID'18, IWQoS'20, WWW'20, ASE'21

Search-based
INFOCOM'14, SOC'18, ICSE'21

Lack explainability



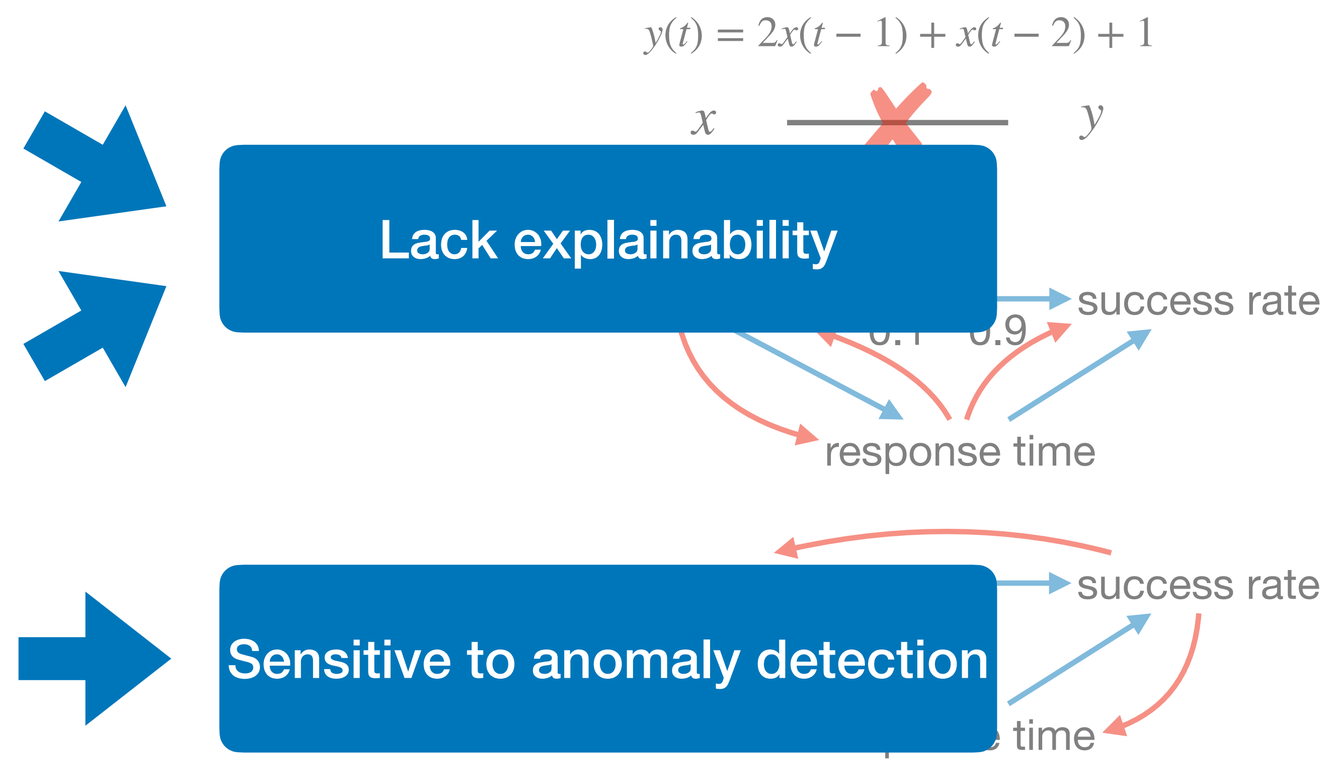
Background

Related Works

Invariant Network-based
KDD'16, ICDM'17

Random Walk-based
CCGRID'18, IWQoS'20, WWW'20, ASE'21

Search-based
INFOCOM'14, SOC'18, ICSE'21



Background

Causal Inference

Sage (ASPLOS'21) based on Counterfactual

The latency is high now.
Would the latency recover if CPU utilization decreased?

Background

Causal Inference

Sage (ASPLOS'21) based on Counterfactual

The latency is high now.
Would the latency recover if CPU utilization decreased?

$\mathscr{L}_3$ Counterfactual

How will the latency distribute
if we make CPU utilization low?

$\mathscr{L}_2$ Interventional

How will the latency distribute when CPU utilization is low?

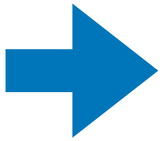
$\mathscr{L}_1$ Associational

Judea Pearl's Ladder of Causation

Background

Causal Inference

Sage (ASPLOS'21) based on Counterfactual



Assume no fault

$\mathscr{L}_3$ Counterfactual

How will the latency distribute
if we make CPU utilization low?

$\mathscr{L}_2$ Interventional

How will the latency distribute
when CPU utilization is low?

$\mathscr{L}_1$ Associational

Judea Pearl's Ladder of Causation

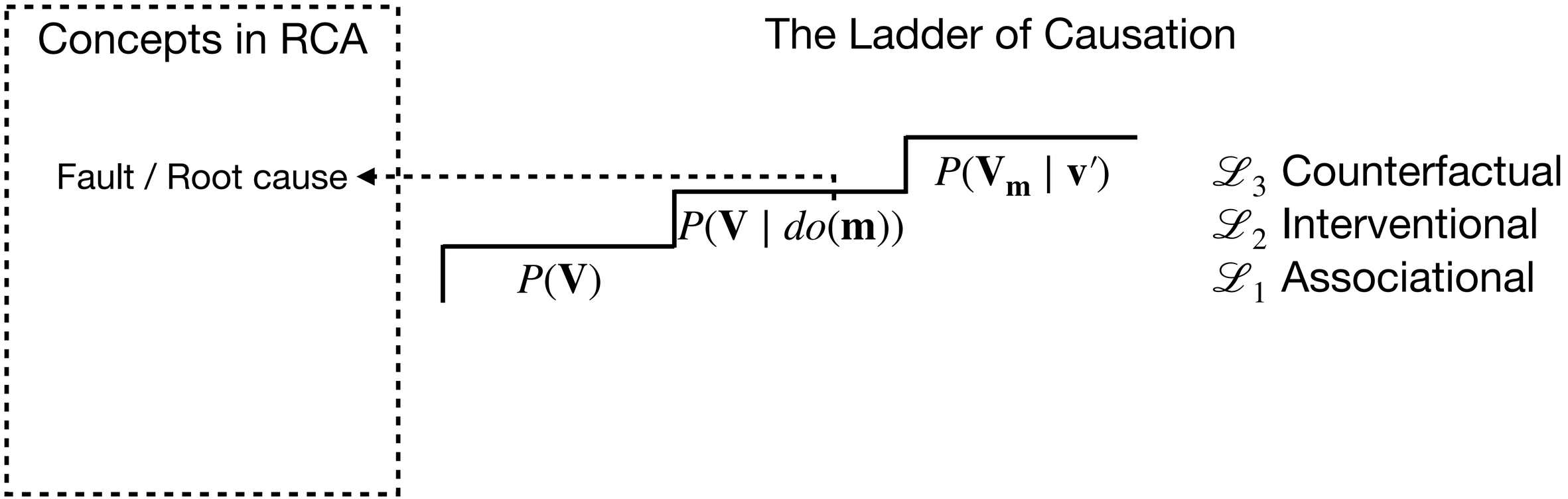
Problem Formulation

Formulate RCA as a causal inference task

$\mathscr{L}_3$ Counterfactual
$\mathscr{L}_2$ Interventional
$\mathscr{L}_1$ Associational

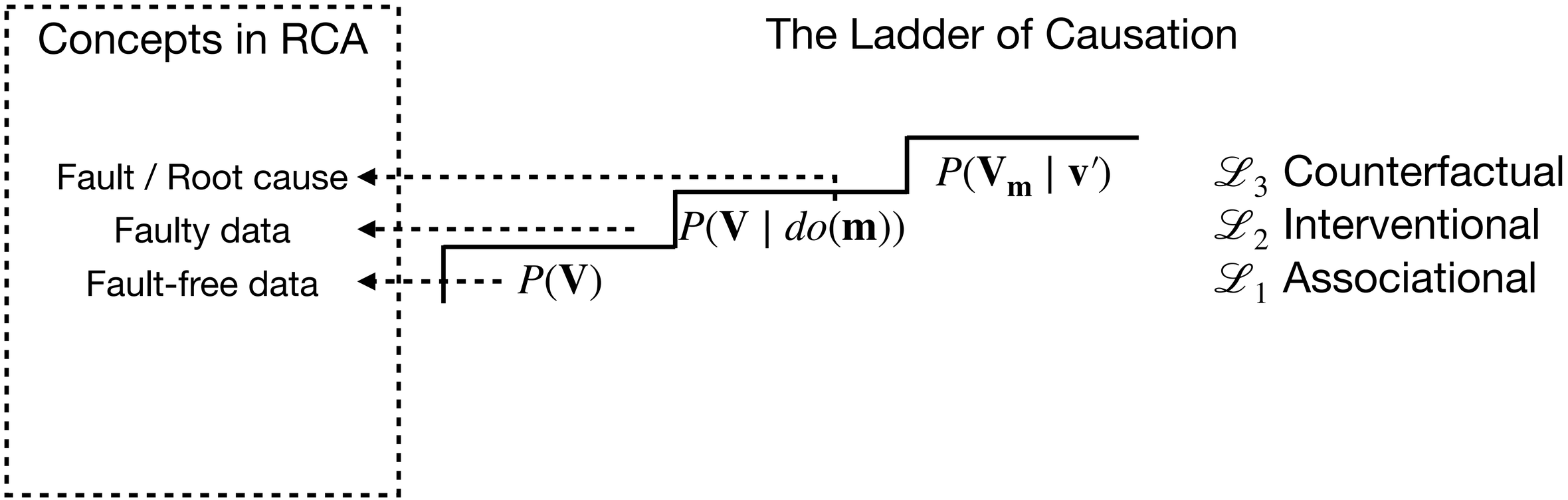
Problem Formulation

Formulate RCA as a causal inference task



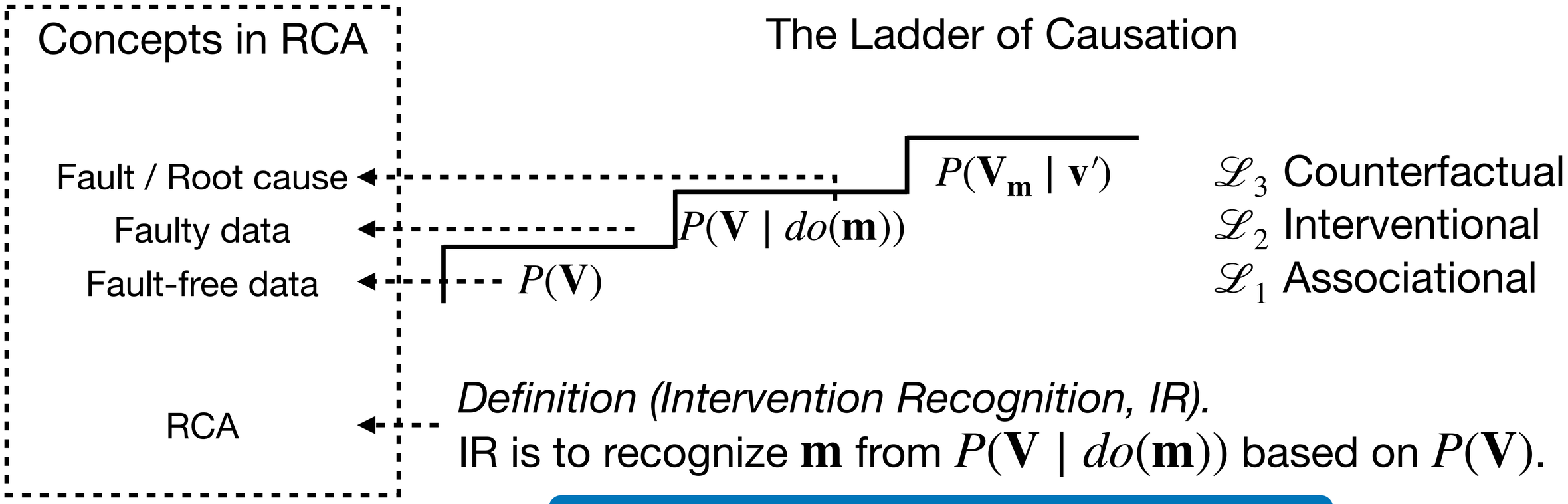
Problem Formulation

Formulate RCA as a causal inference task



Problem Formulation

Formulate RCA as a causal inference task

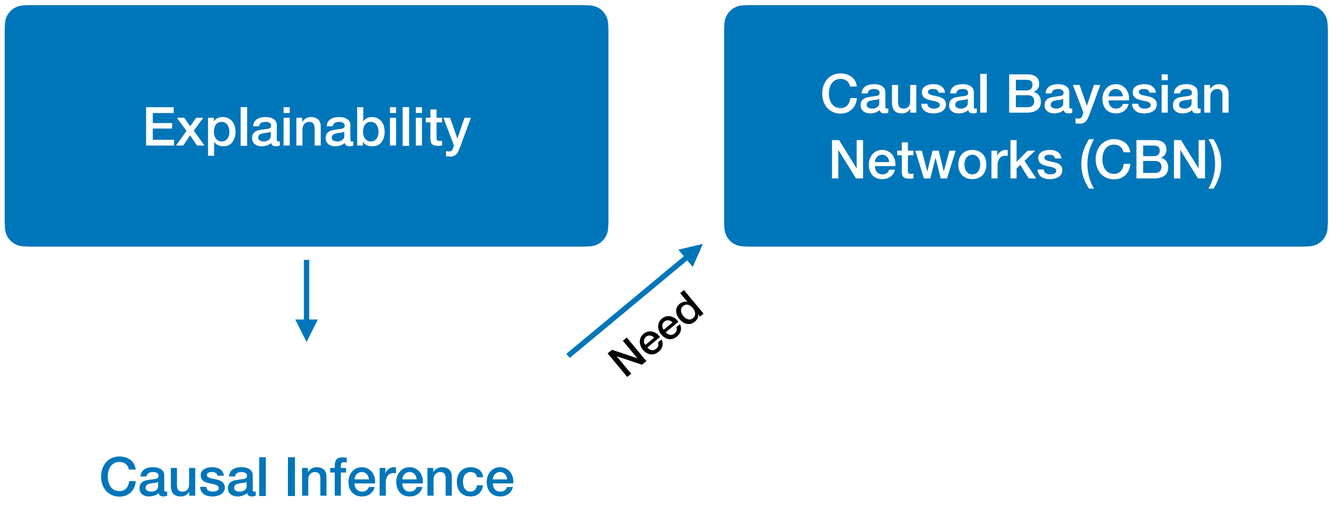


Definition (Intervention Recognition, IR).
IR is to recognize $\mathbf{m}$ from $P(\mathbf{V} \mid do(\mathbf{m}))$ based on $P(\mathbf{V})$.

$\mathbf{M}$ (***M***anipulation): the intervened variables
$\mathbf{m}$: the value of $\mathbf{M}$

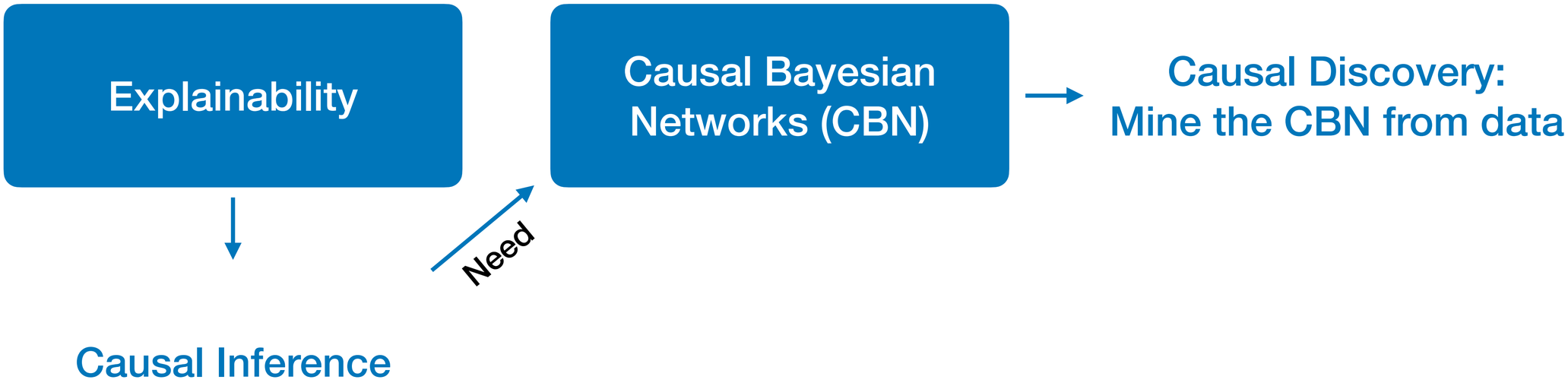
Challenges

Causal inference requires the Causal Bayesian Networks



Challenges

Causal inference requires the Causal Bayesian Networks



Challenges

Causal inference requires the Causal Bayesian Networks

Challenges

Causal inference requires the Causal Bayesian Networks

Explainability

Causal Bayesian Networks (CBN)

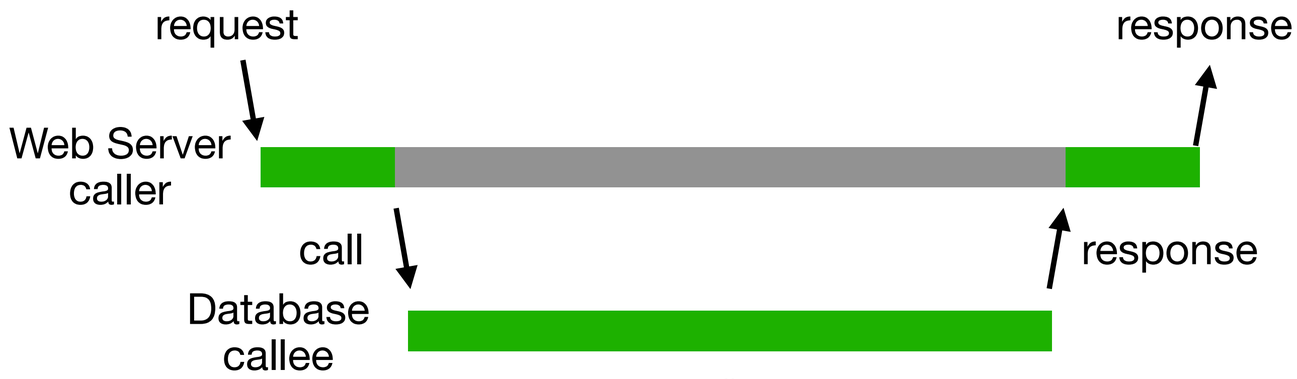
Causal Discovery: Mine the CBN from data

Need

Causal Inference

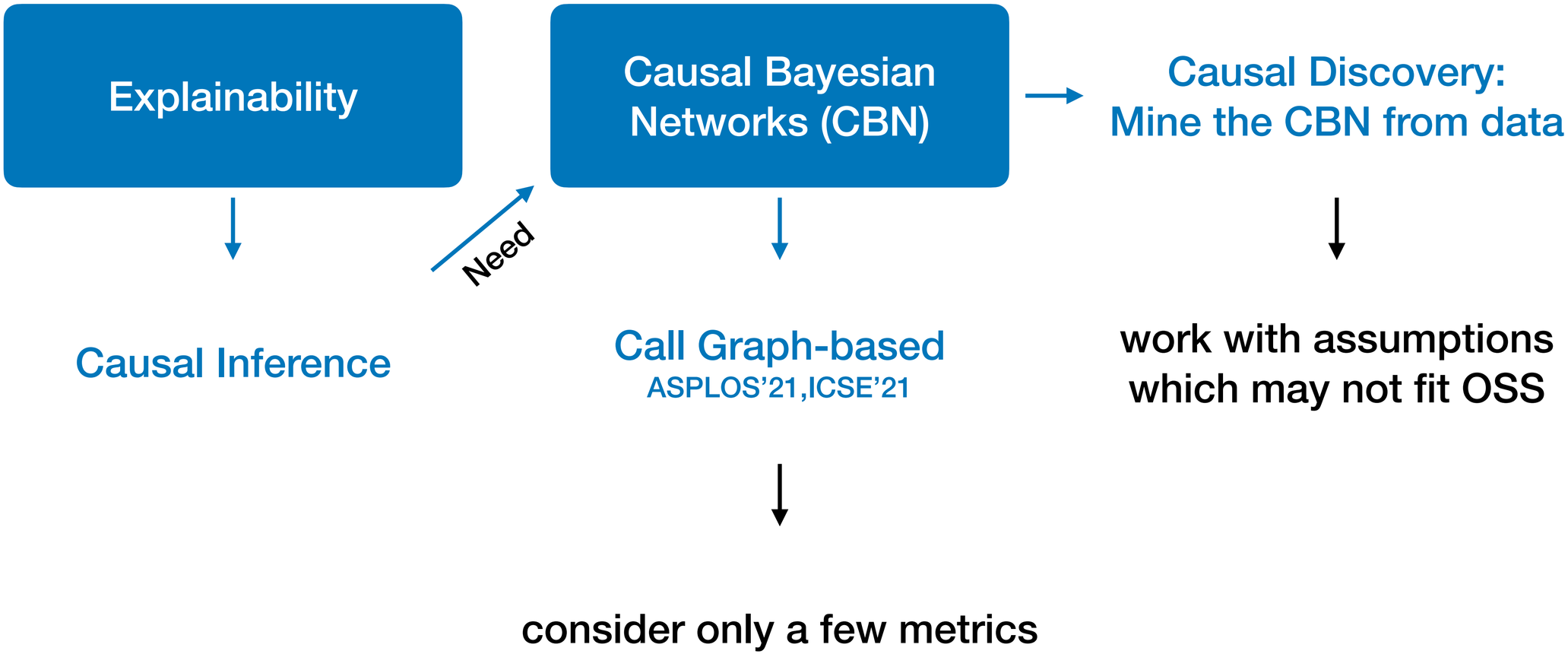
Call Graph-based
ASPLOS'21,ICSE'21

work with assumptions which may not fit OSS



Challenges

Causal inference requires the Causal Bayesian Networks



Challenges

Observational knowledge is incomplete

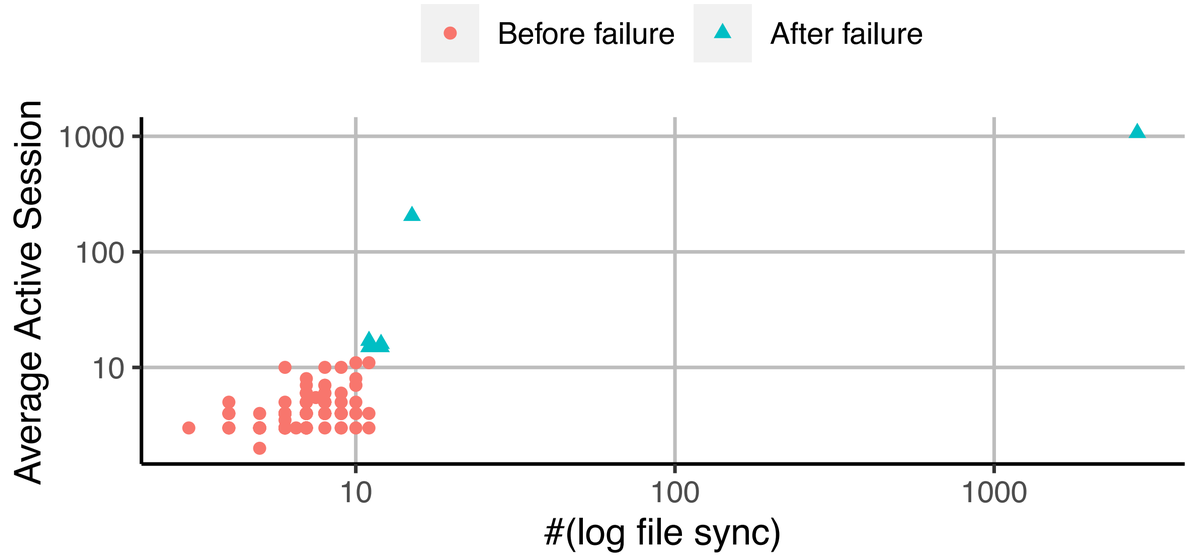
Explainability

Causal Bayesian Networks (CBN)

Incomplete observational knowledge

Need

Causal Inference



Challenges

Observational knowledge is incomplete

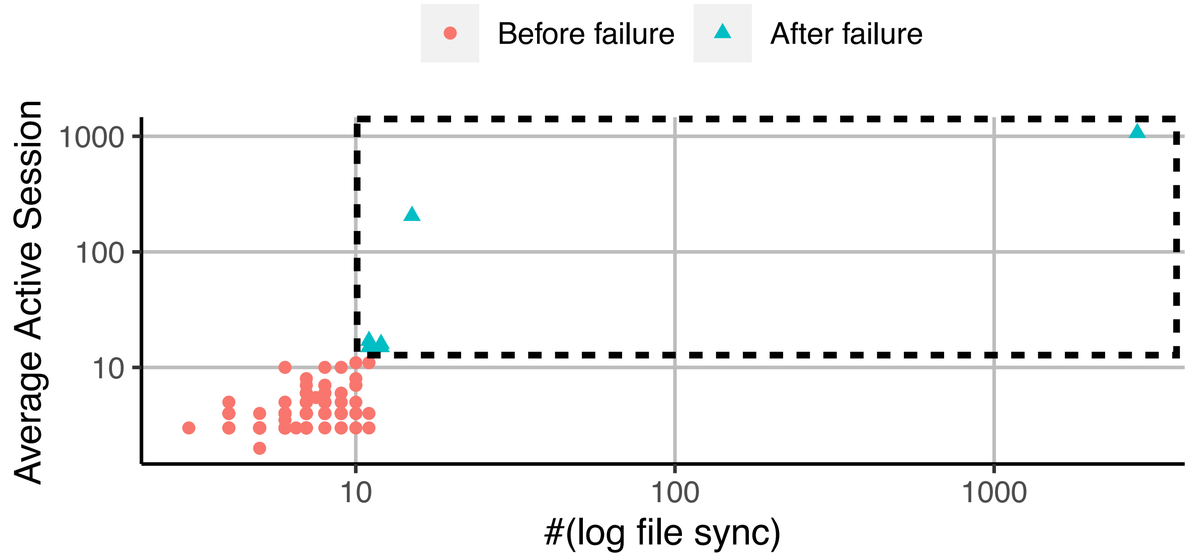
Explainability

Causal Bayesian Networks (CBN)

Incomplete observational knowledge

Need

Causal Inference



Challenges

Observational knowledge is incomplete

Explainability

Causal Bayesian Networks (CBN)

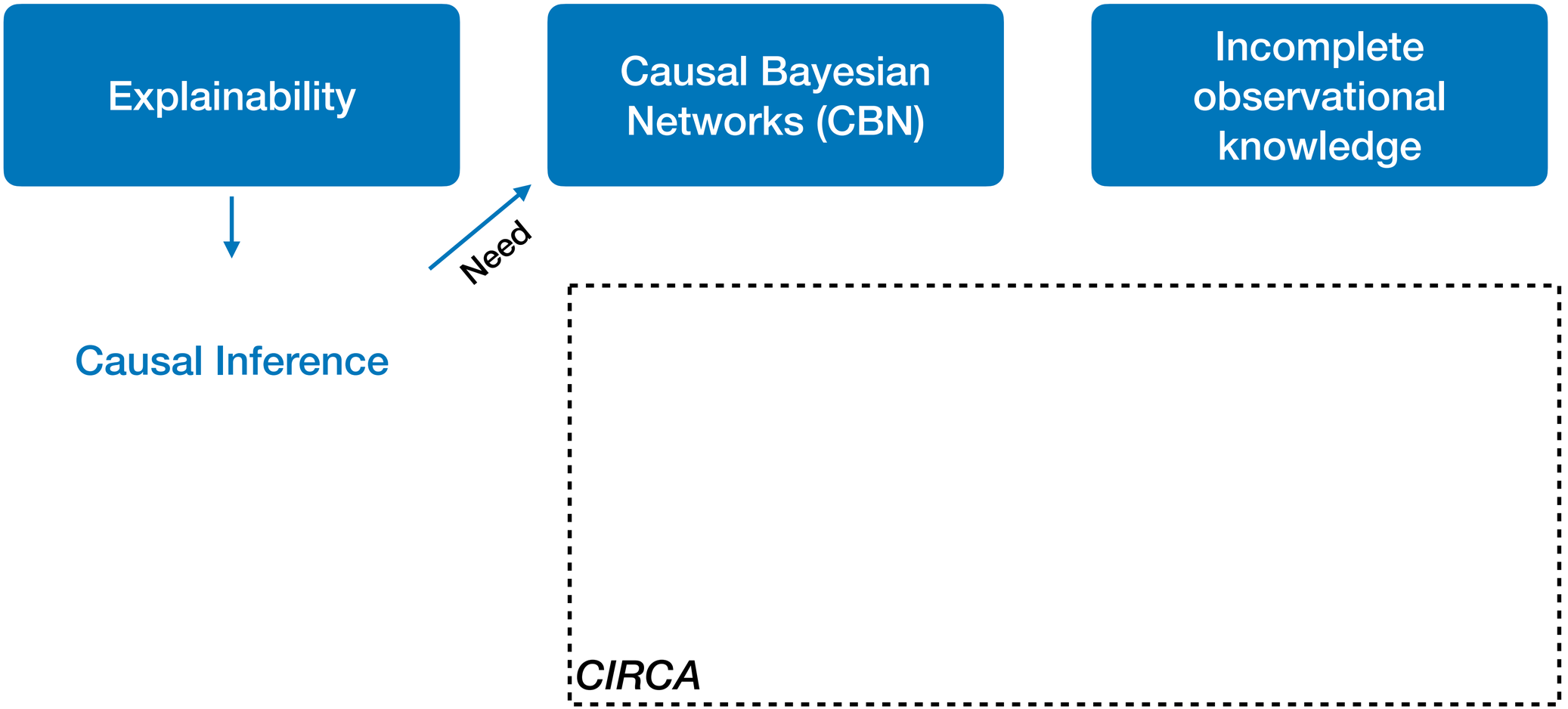
Incomplete observational knowledge

Need

Causal Inference

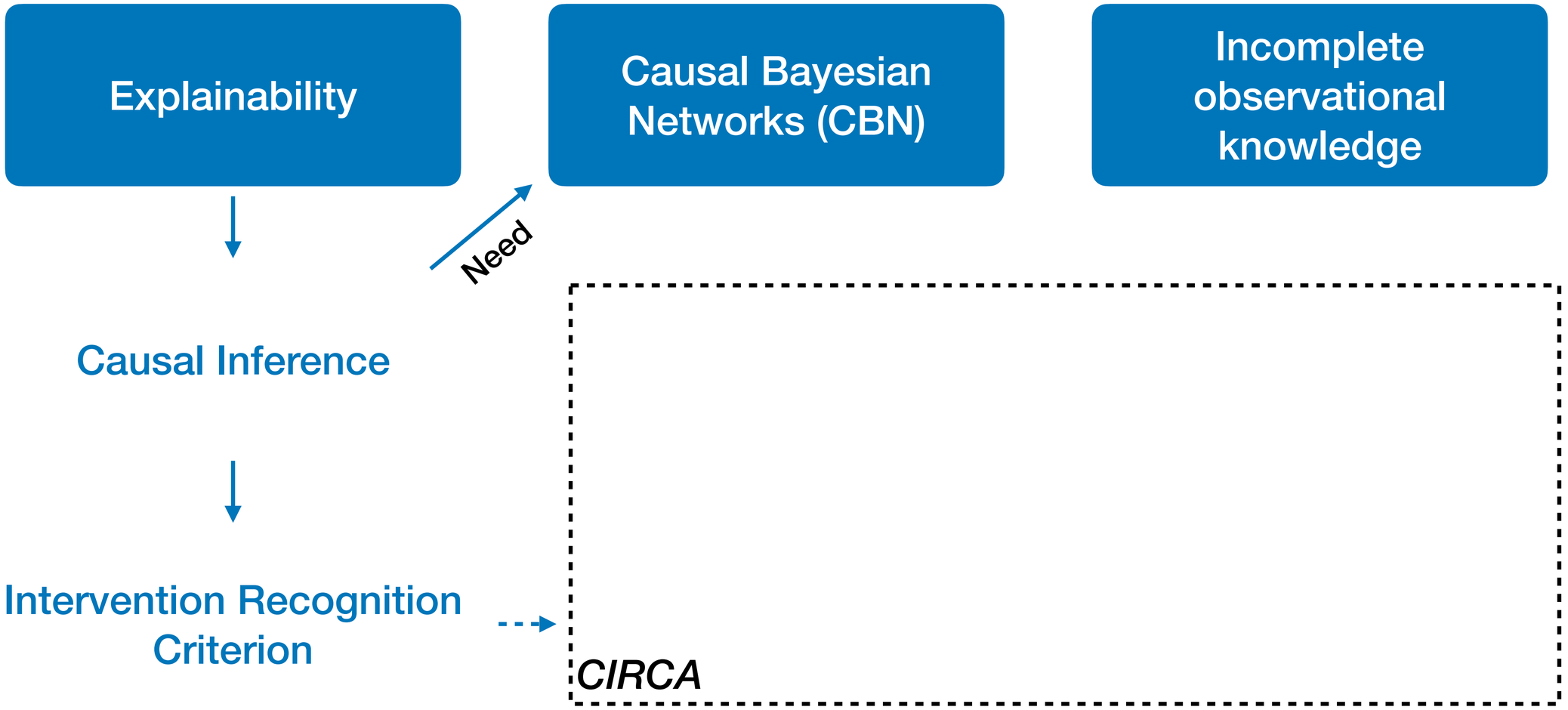
Overview

Causal Inference-based Root Cause Analysis (CIRCA) for OSS



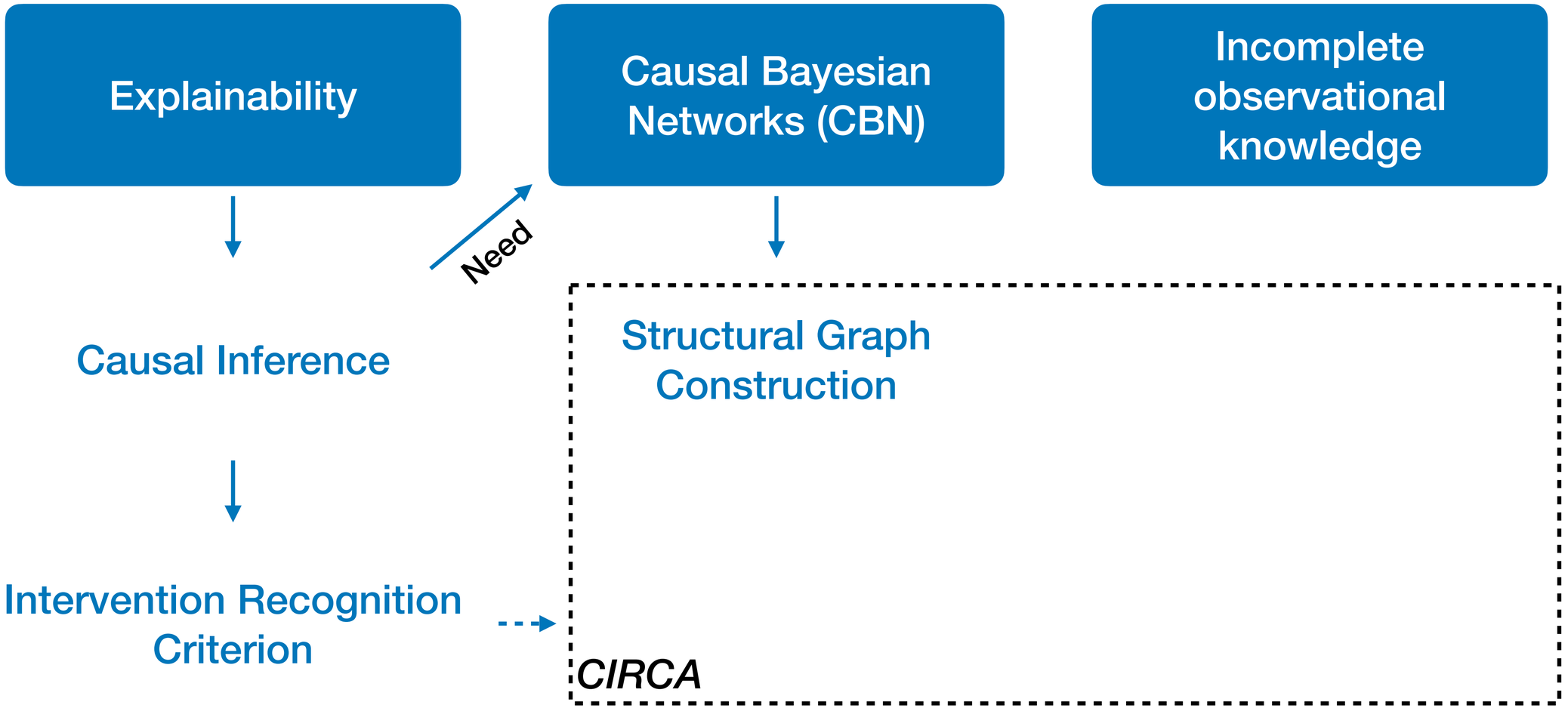
Overview

Causal Inference-based Root Cause Analysis (CIRCA) for OSS



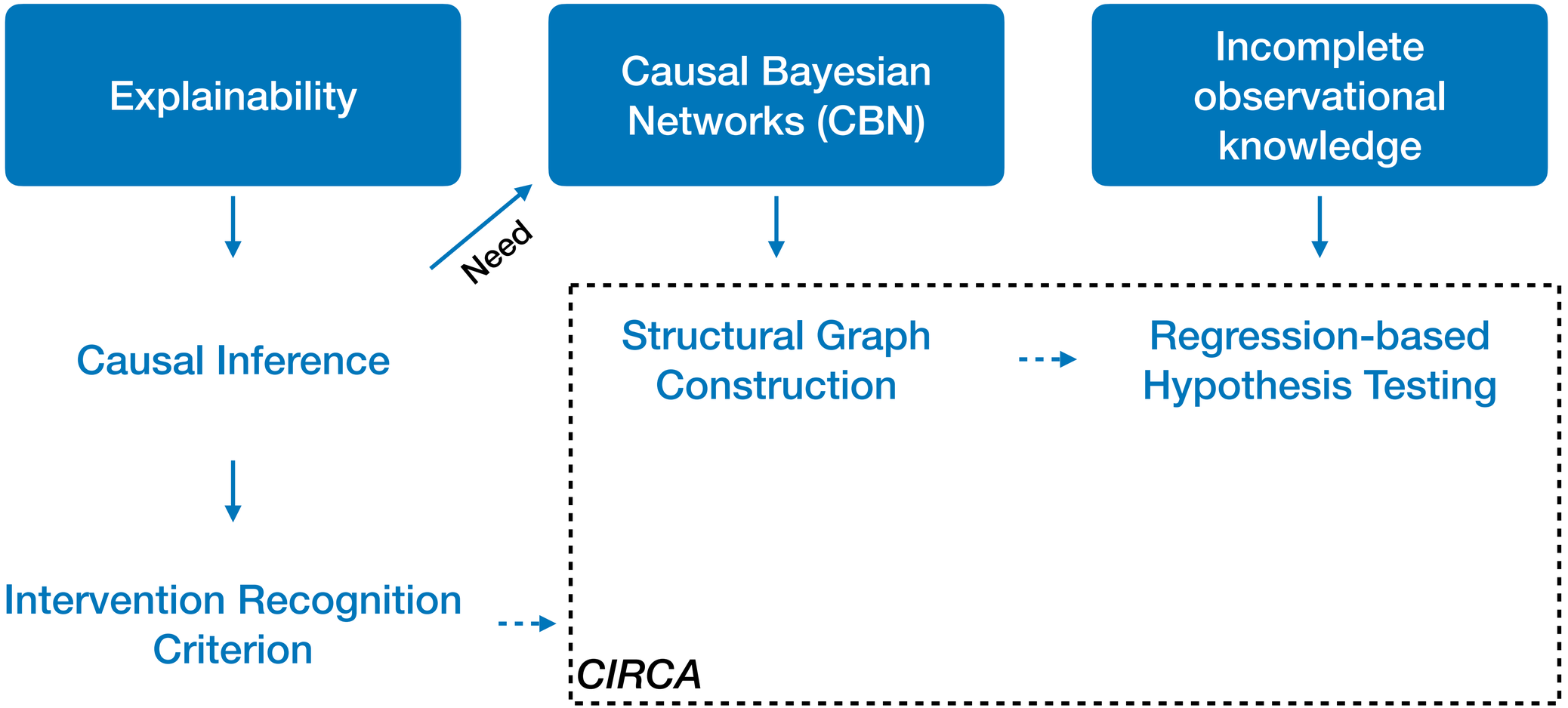
Overview

Causal Inference-based Root Cause Analysis (CIRCA) for OSS



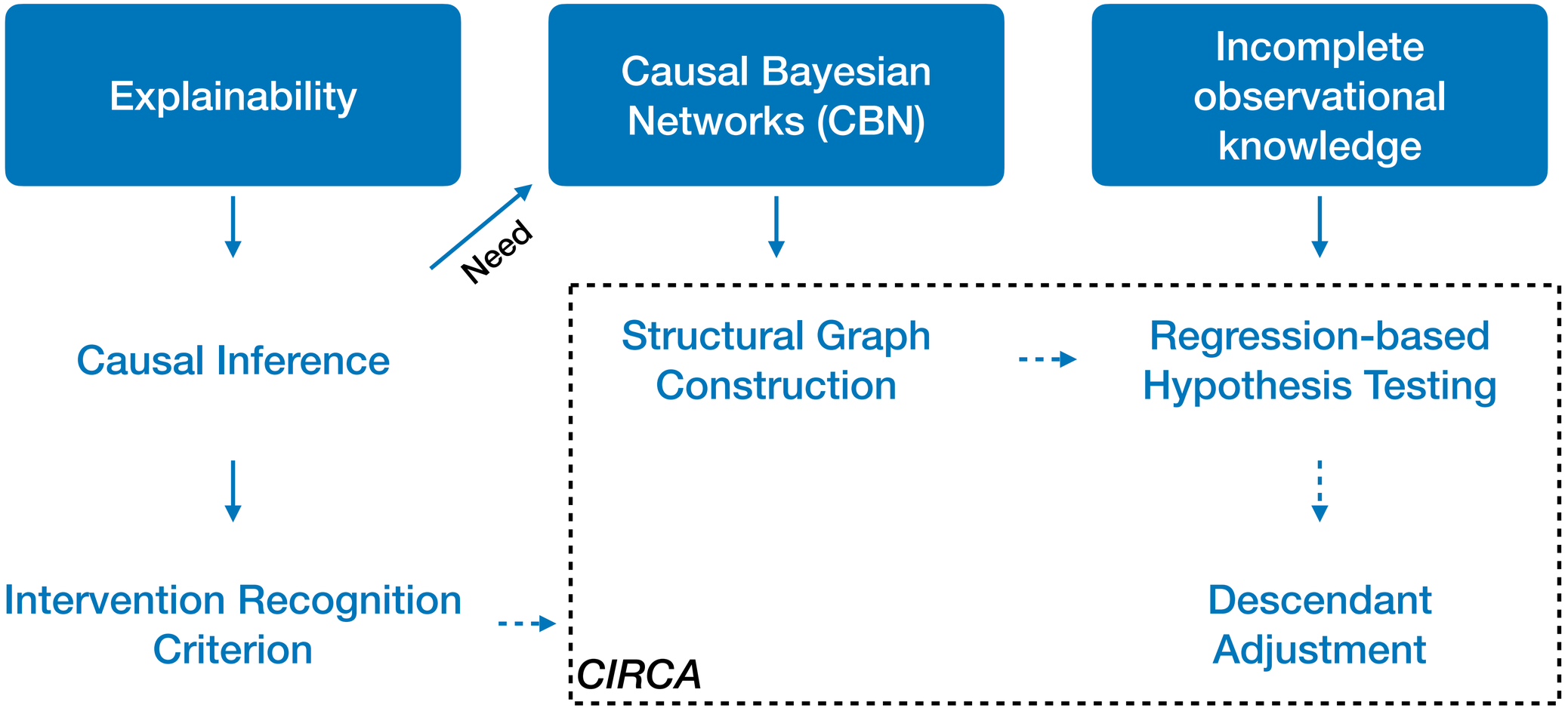
Overview

Causal Inference-based Root Cause Analysis (CIRCA) for OSS



Overview

Causal Inference-based Root Cause Analysis (CIRCA) for OSS



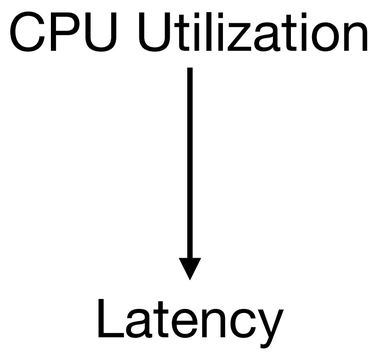
Analysis

Background
Analysis
Methodology
Evaluation
Conclusion

Analysis

Mapping from interventional distributions to interventions

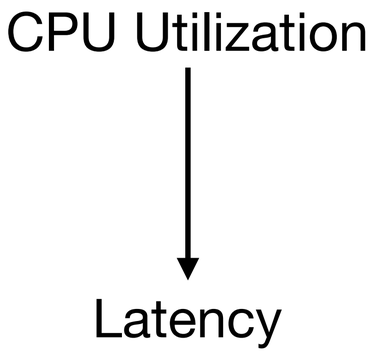
Latency	10s
CPU Utilization	98%



Analysis

Mapping from interventional distributions to interventions

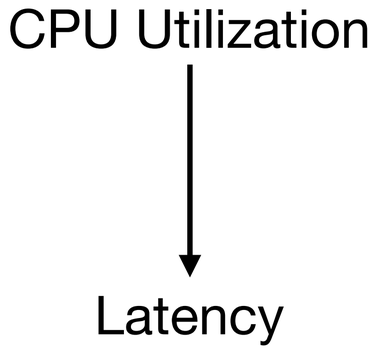
Latency	10s
CPU Utilization	98%



Analysis

Mapping from interventional distributions to interventions

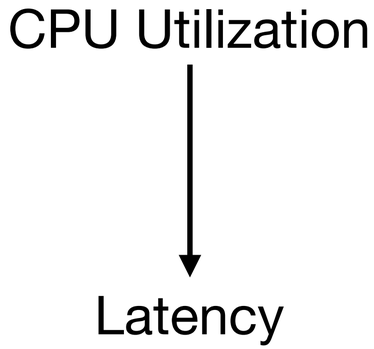
$$P(\text{Latency} = 10s \mid do(\text{CPU Utilization} = 98\%))$$



Analysis

Mapping from interventional distributions to interventions

$$P(\text{Latency} = 10s \mid do(\text{CPU Utilization} = 98\%))$$



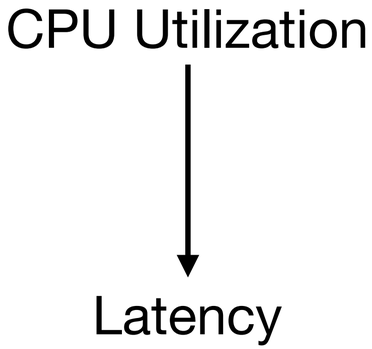
IR $\Rightarrow \mathscr{L}_2$

Analysis

Mapping from interventional distributions to interventions

Latency	10s
CPU Utilization	98%

$$P(\text{Latency} = 10s \mid do(\text{CPU Utilization} = 98\%))$$



IR $\Rightarrow \mathscr{L}_2$

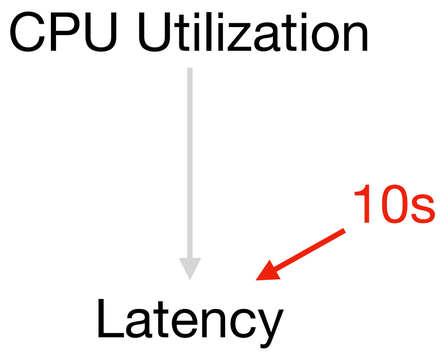
$\mathscr{L}_2 \Rightarrow$ IR?

Analysis

Same distribution but different interventions

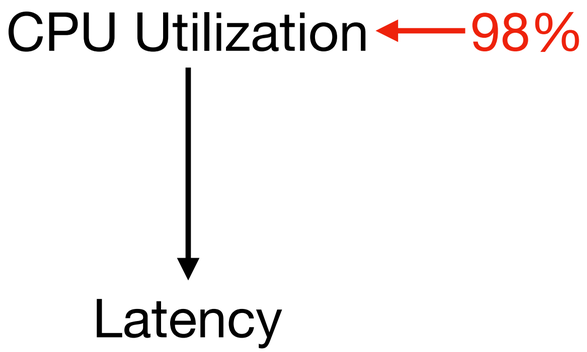
* red means the intervention

Latency	10s
CPU Utilization	98%



World 1

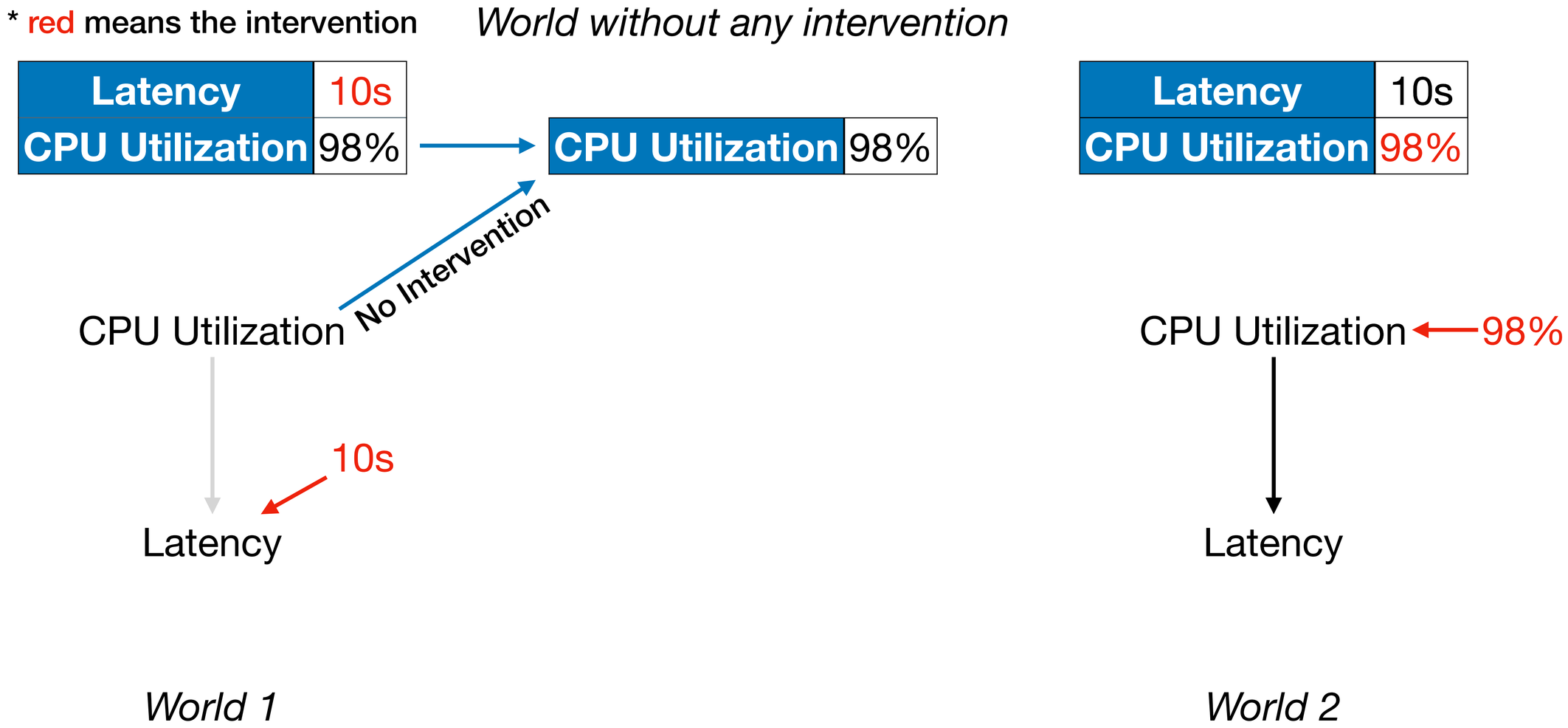
Latency	10s
CPU Utilization	98%



World 2

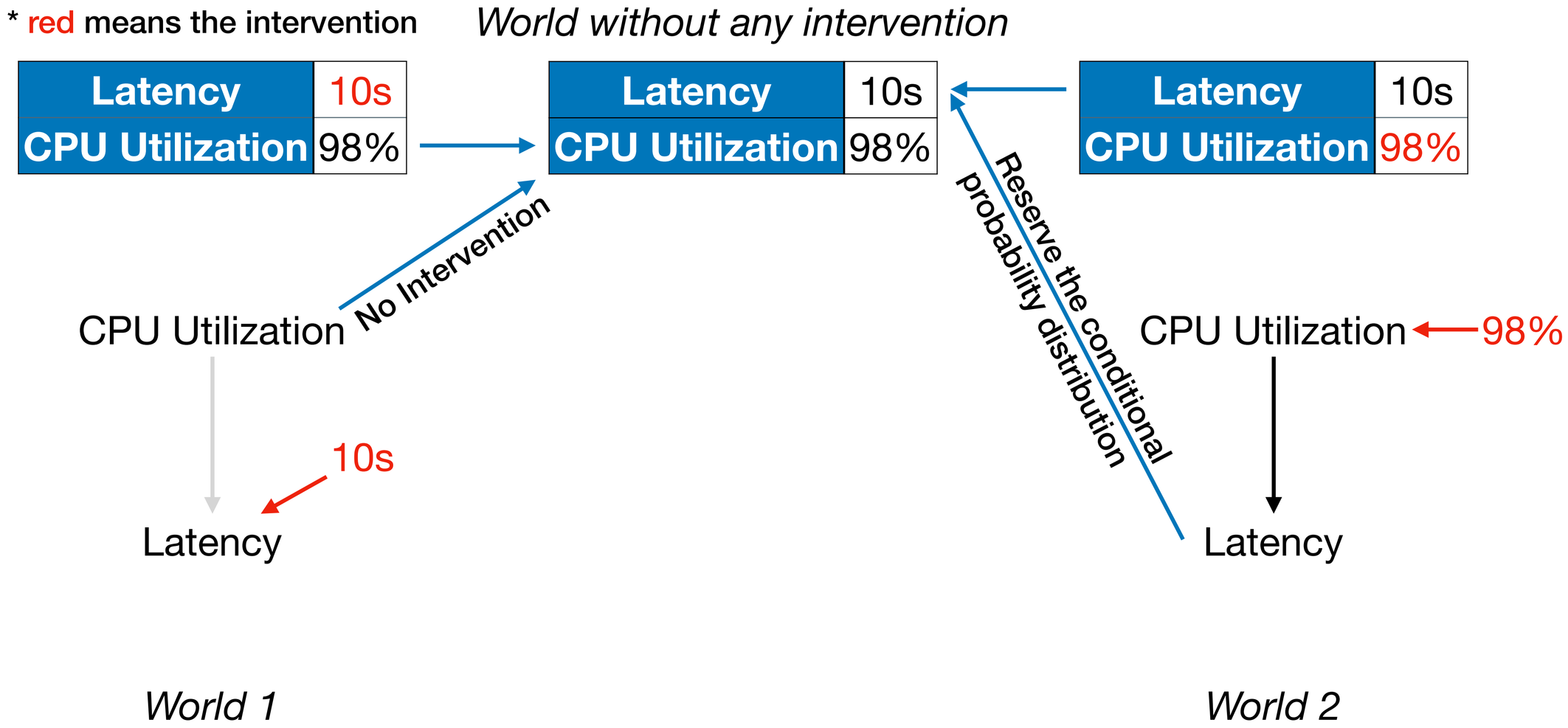
Analysis

Same distribution but different interventions



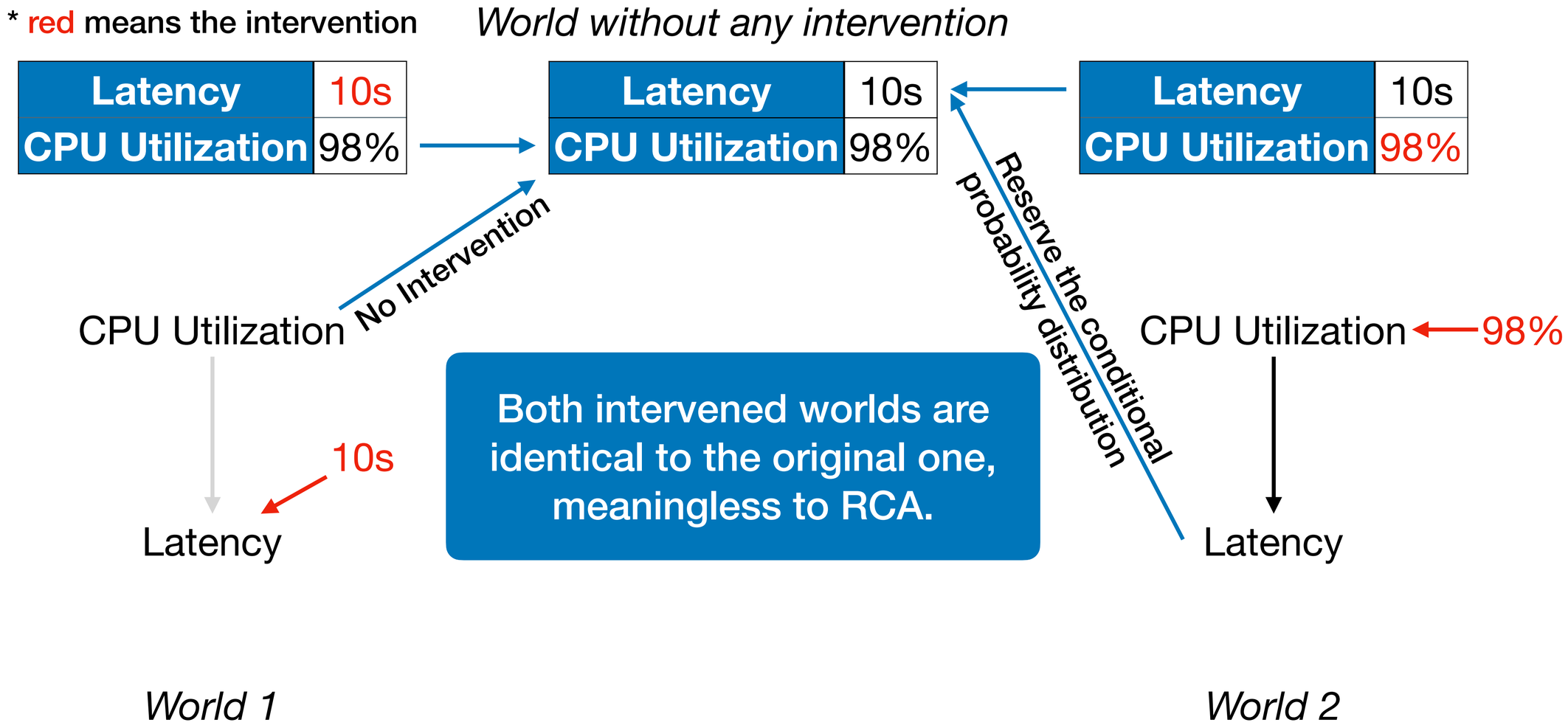
Analysis

Same distribution but different interventions



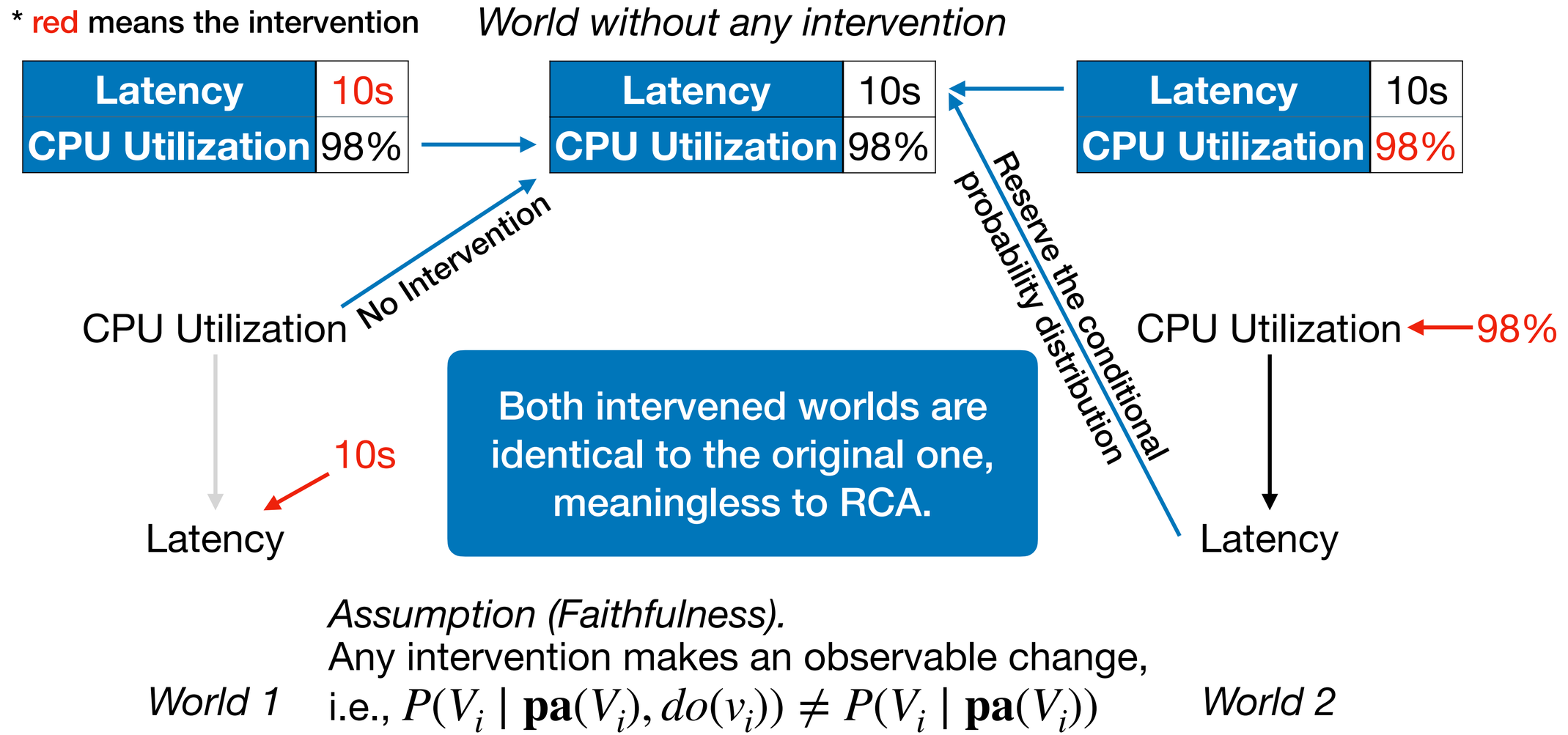
Analysis

Same distribution but different interventions



Analysis

Same distribution but different interventions



Analysis

IR is at the second layer of the causal ladder

IR $\equiv \mathscr{L}_2$

Assumption (Faithfulness).
Any intervention makes an observable change,
i.e., $P(V_i \mid \mathbf{pa}(V_i), do(v_i)) \neq P(V_i \mid \mathbf{pa}(V_i))$

Analysis

IR is at the second layer of the causal ladder

IR $\equiv \mathscr{L}_2$

Causal Hierarchy Theorem [1]

If we want to answer the question at Layer i, we need knowledge at Layer i or higher.

[1] Elias Bareinboim, Juan D. Correa, Duligur Ibeling, Thomas Icard. On Pearl's Hierarchy and the Foundations of Causal Inference. Last Revision: Mar, 2021

Analysis

IR is at the second layer of the causal ladder

IR $\equiv \mathscr{L}_2$

Causal Hierarchy Theorem [1]

IR needs the knowledge of $\mathscr{L}_2$ (like the CBN [1])

Counterfactual knowledge of $\mathscr{L}_3$ is unnecessary

[1] Elias Bareinboim, Juan D. Correa, Duligur Ibeling, Thomas Icard. On Pearl's Hierarchy and the Foundations of Causal Inference. Last Revision: Mar, 2021

Analysis

Intervention Recognition Criterion

A given variable is a root cause indicator $V_i \in \mathbf{M}$

Faithfulness →

Change in the distribution conditioned on the parents in the CBN

$$P(V_i \mid \mathbf{pa}(V_i), do(\mathbf{m})) \neq P(V_i \mid \mathbf{pa}(V_i))$$

Analysis

Intervention Recognition Criterion

A given variable is a root cause indicator $V_i \in \mathbf{M}$

Faithfulness

Parents do/see [1] a property of the CBN

Change in the distribution conditioned on the parents in the CBN

$$P(V_i \mid \mathbf{pa}(V_i), do(\mathbf{m})) \neq P(V_i \mid \mathbf{pa}(V_i))$$

[1] Elias Bareinboim, Juan D. Correa, Duligur Ibeling, Thomas Icard. On Pearl's Hierarchy and the Foundations of Causal Inference. Last Revision: Mar, 2021

Analysis

Intervention Recognition Criterion

A given variable is a root cause indicator $V_i \in \mathbf{M}$

Faithfulness →

← Parents do/see [1] a property of the CBN

Change in the distribution conditioned on the parents in the CBN $P(V_i \mid \mathbf{pa}(V_i), do(\mathbf{m})) \neq P(V_i \mid \mathbf{pa}(V_i))$

$$V_i \in \mathbf{M} \Leftrightarrow P(V_i \mid \mathbf{pa}(V_i), do(\mathbf{m})) \neq P(V_i \mid \mathbf{pa}(V_i))$$

[1] Elias Bareinboim, Juan D. Correa, Duligur Ibeling, Thomas Icard. On Pearl's Hierarchy and the Foundations of Causal Inference. Last Revision: Mar, 2021

Methodology

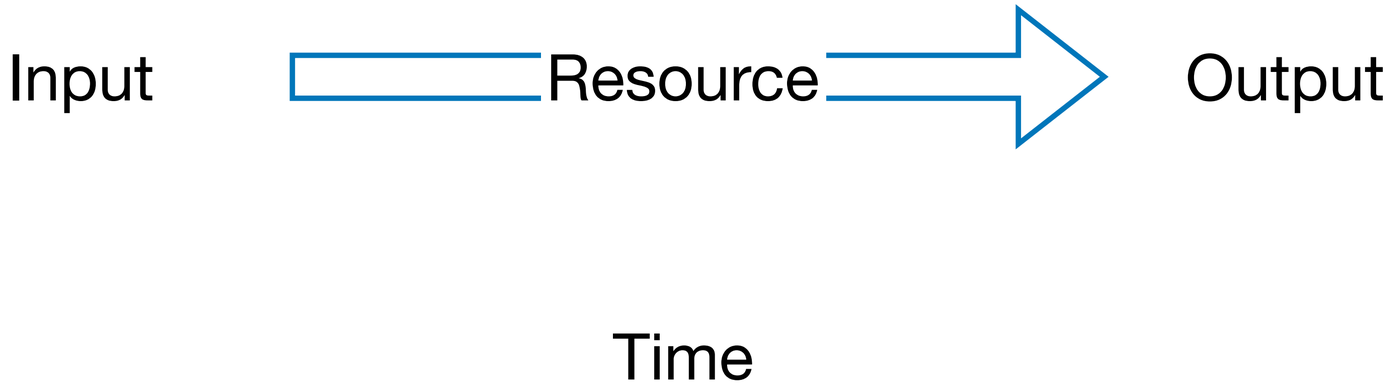
Background
Analysis
Methodology
Evaluation
Conclusion

Structural Graph Construction

Meta Metrics

Structural Graph Construction

Meta Metrics



Structural Graph Construction

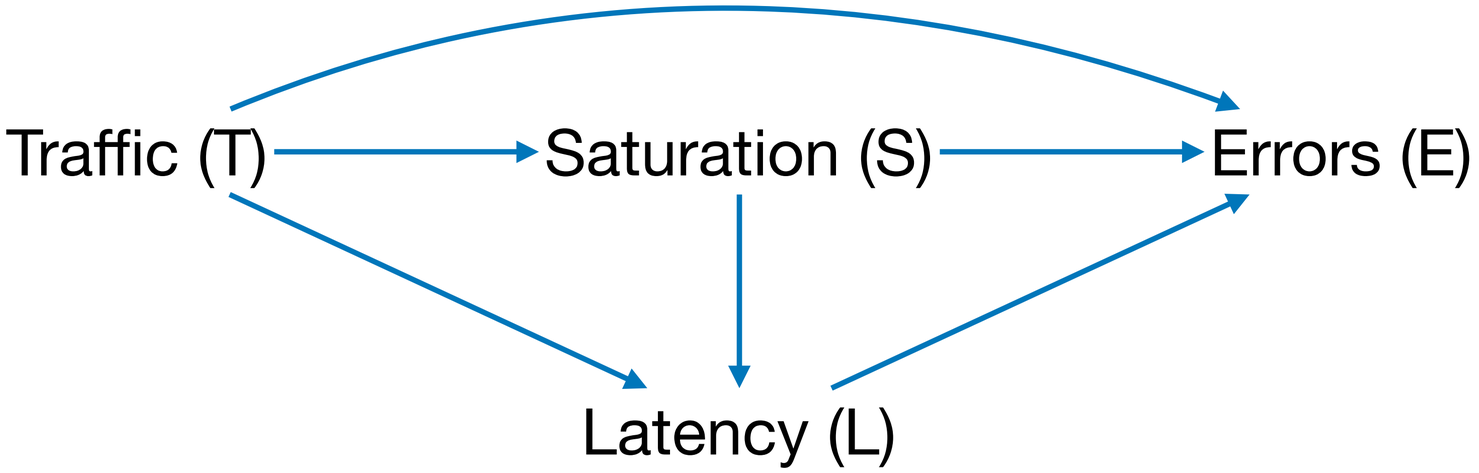
Meta Metrics

Latency (L)

[2] Betsy Beyer, Chris Jones, Jennifer Petoff, and Niall Richard Murphy. Site Reliability Engineering (first ed.). O'Reilly Media, Inc. 2016.

Structural Graph Construction

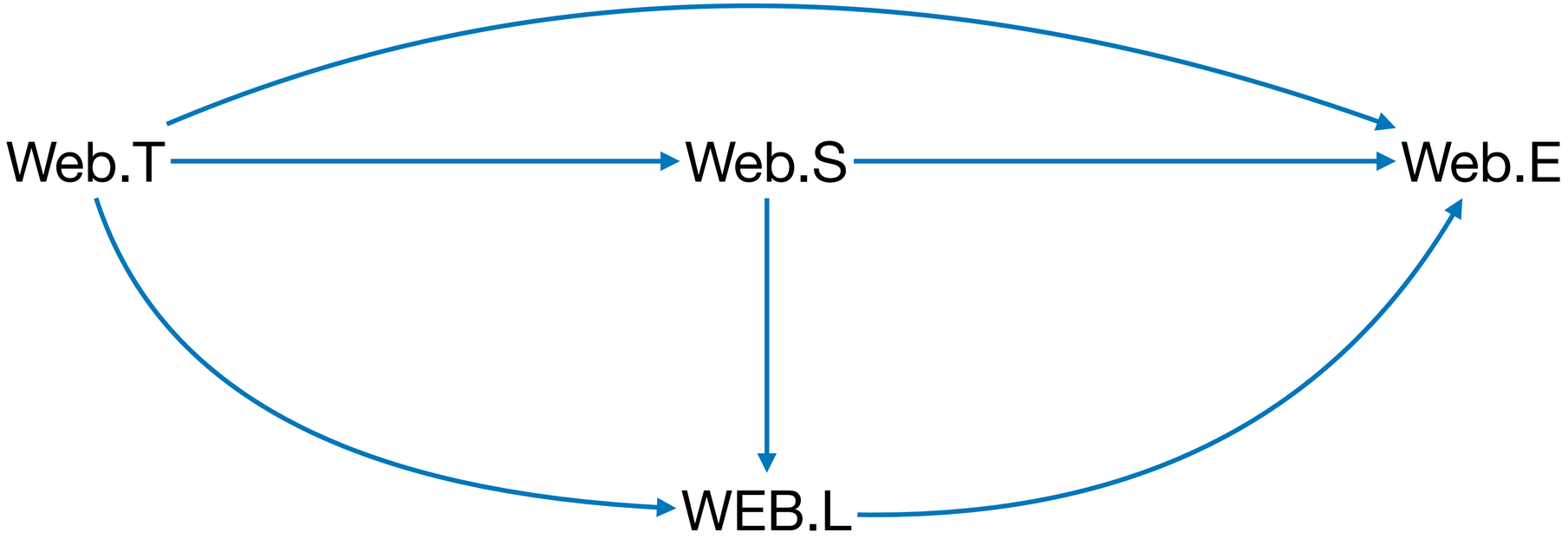
Meta Metrics



Assign directions among meta metrics as causal assumptions

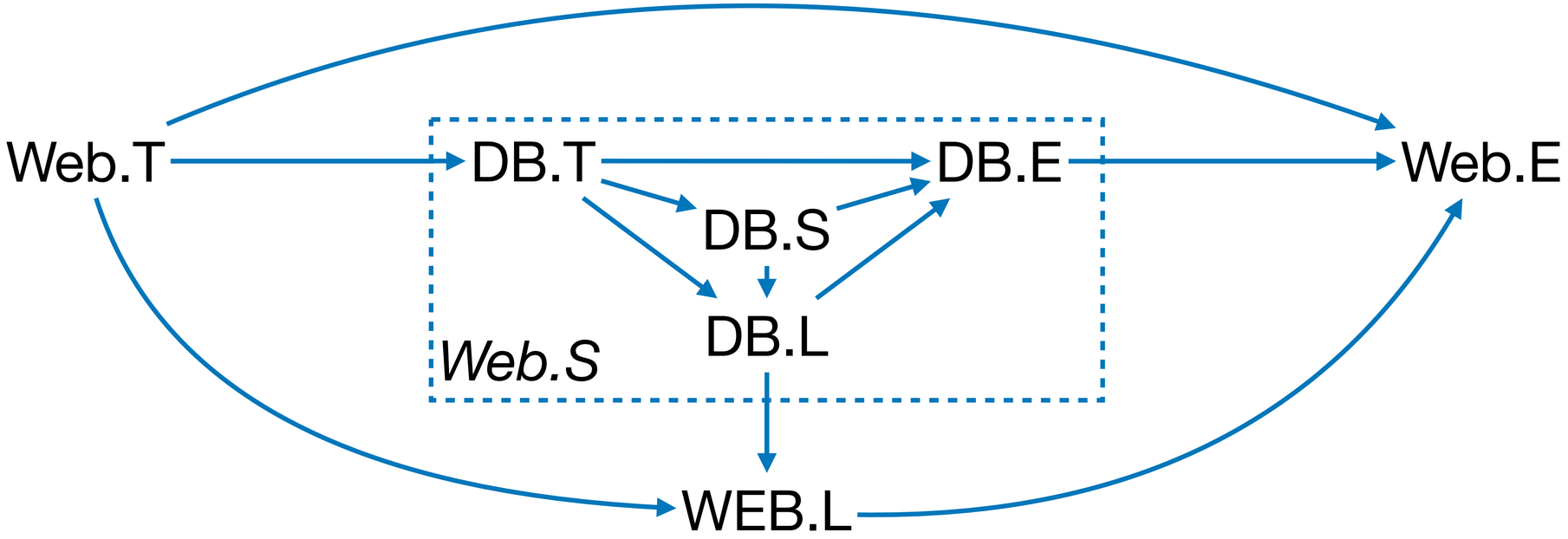
Structural Graph Construction

Skeleton with Architecture Extension



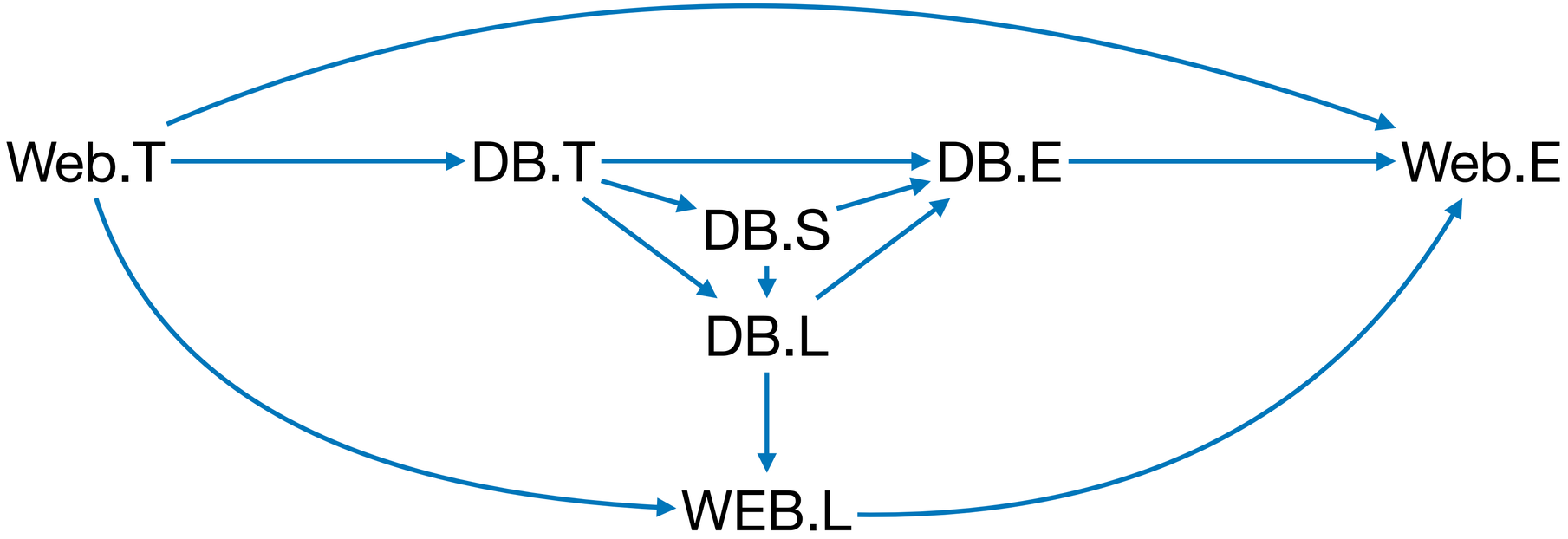
Structural Graph Construction

Skeleton with Architecture Extension



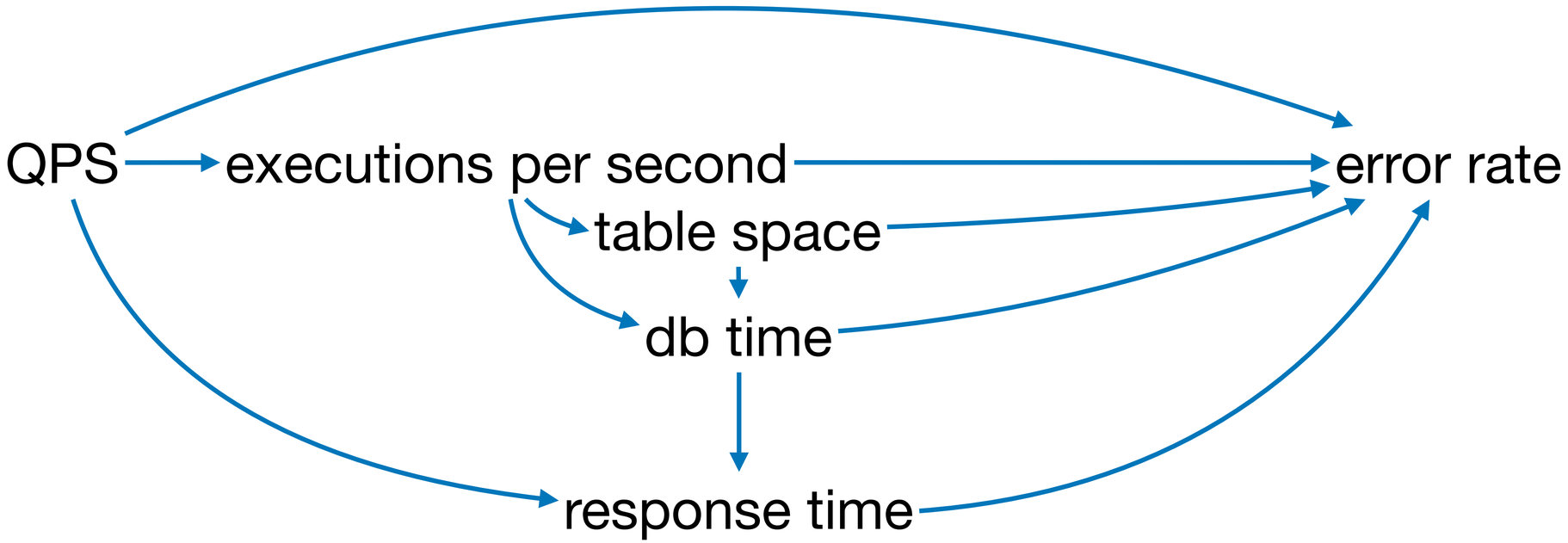
Structural Graph Construction

Monitoring Metric Plugging-in



Structural Graph Construction

Monitoring Metric Plugging-in



Regression-based Hypothesis Testing (RHT)

Handle incomplete observational knowledge

Core Idea: $V_i \in \mathbf{M} \Leftrightarrow P(V_i \mid \mathbf{pa}(V_i), do(\mathbf{m})) \neq P(V_i \mid \mathbf{pa}(V_i))$

Regression-based Hypothesis Testing (RHT)

Handle incomplete observational knowledge

A few faulty data are available

Core Idea: $V_i \in \mathbf{M} \Leftrightarrow P(V_i \mid \mathbf{pa}(V_i), do(\mathbf{m})) \neq P(V_i \mid \mathbf{pa}(V_i))$

Regression-based Hypothesis Testing (RHT)

Handle incomplete observational knowledge

A few faulty data are available

Core Idea: $V_i \in \mathbf{M} \Leftrightarrow P(V_i \mid \mathbf{pa}(V_i), do(\mathbf{m})) \neq P(V_i \mid \mathbf{pa}(V_i))$

Hypothesis Testing

$\mathbf{H}_0(V_i \notin \mathbf{M})$: $V_i^{(t)} \sim P(V_i^{(t)} \mid \mathbf{pa}^{(t)}(V_i))$

The need for faulty data is reduced

Regression-based Hypothesis Testing (RHT)

Handle incomplete observational knowledge

A few faulty data are available

Fault-free data may not cover the given condition

Core Idea: $V_i \in \mathbf{M} \Leftrightarrow P(V_i \mid \mathbf{pa}(V_i), do(\mathbf{m})) \neq P(V_i \mid \mathbf{pa}(V_i))$

Hypothesis Testing

$\mathbf{H}_0(V_i \notin \mathbf{M})$: $V_i^{(t)} \sim P(V_i^{(t)} \mid \mathbf{pa}^{(t)}(V_i))$

The need for faulty data is reduced

Regression-based Hypothesis Testing (RHT)

Handle incomplete observational knowledge

A few faulty data are available

Fault-free data may not cover the given condition

Core Idea: $V_i \in \mathbf{M} \Leftrightarrow P(V_i \mid \mathbf{pa}(V_i), do(\mathbf{m})) \neq P(V_i \mid \mathbf{pa}(V_i))$

Hypothesis Testing

Regression

$\mathbf{H}_0(V_i \notin \mathbf{M})$: $V_i^{(t)} \sim P(V_i^{(t)} \mid \mathbf{pa}^{(t)}(V_i))$

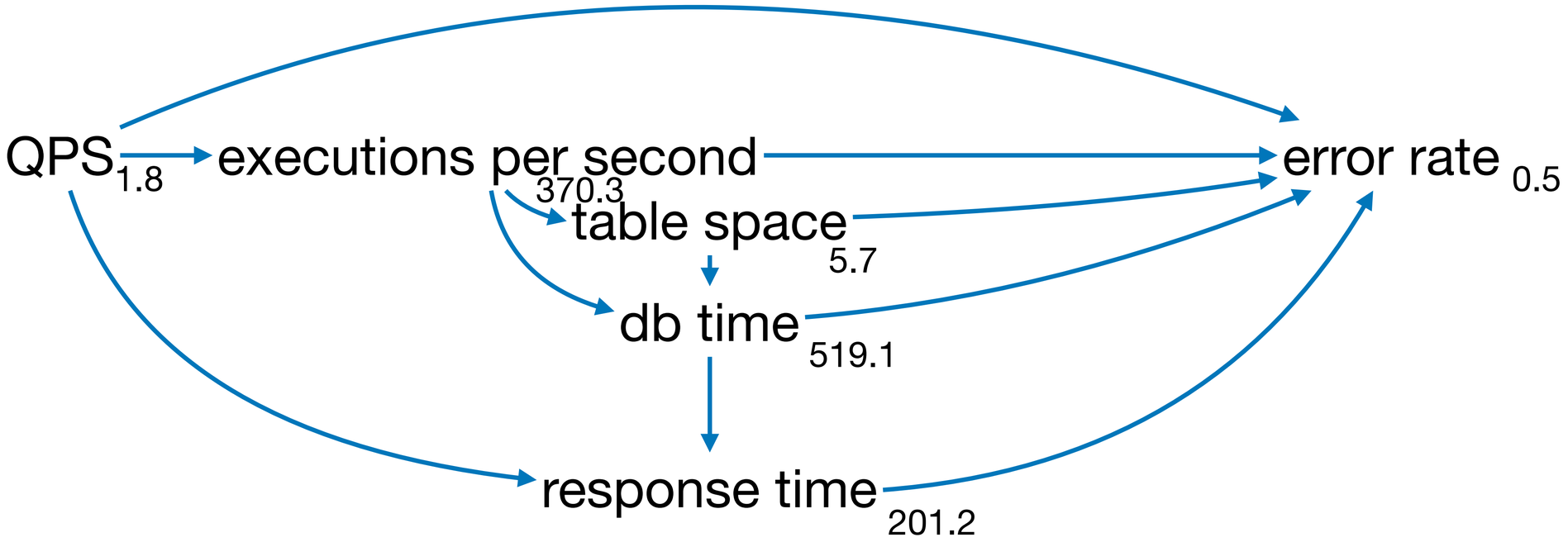
The need for faulty data is reduced

$\approx$

$\mathcal{N}(\text{expectation}, \sigma_{residuals})$

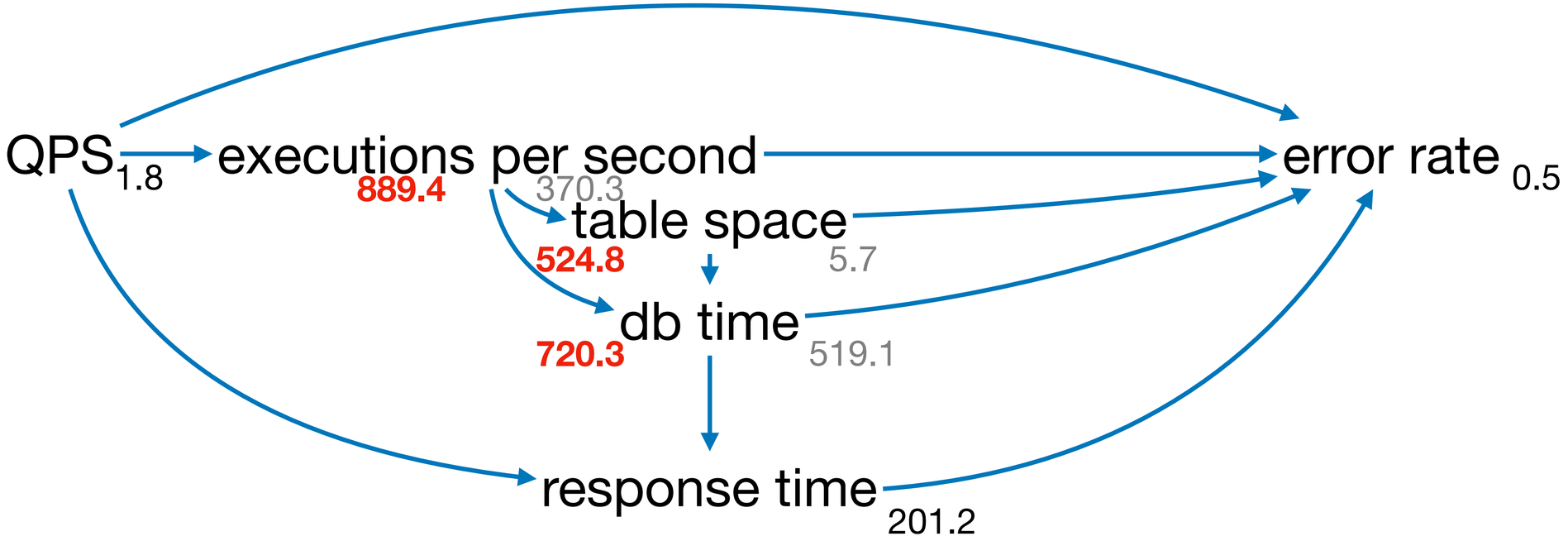
Descendant Adjustment

Alleviate the bias introduced in hypothesis testing



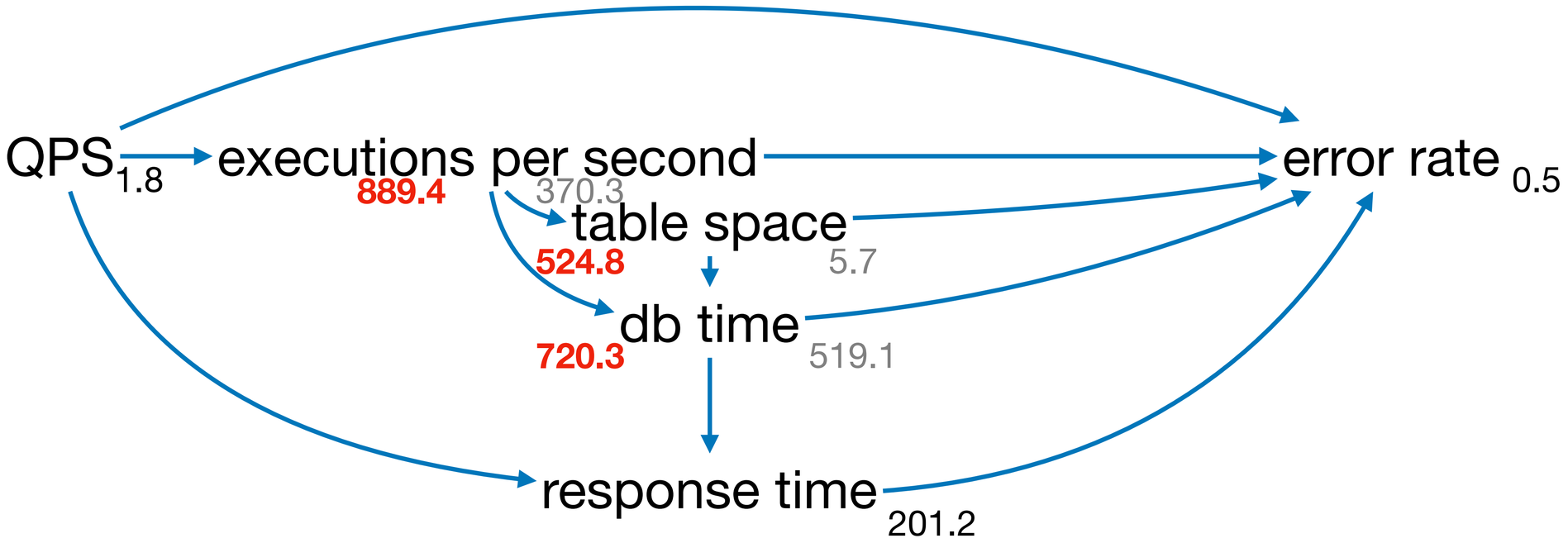
Descendant Adjustment

Alleviate the bias introduced in hypothesis testing



Descendant Adjustment

Alleviate the bias introduced in hypothesis testing



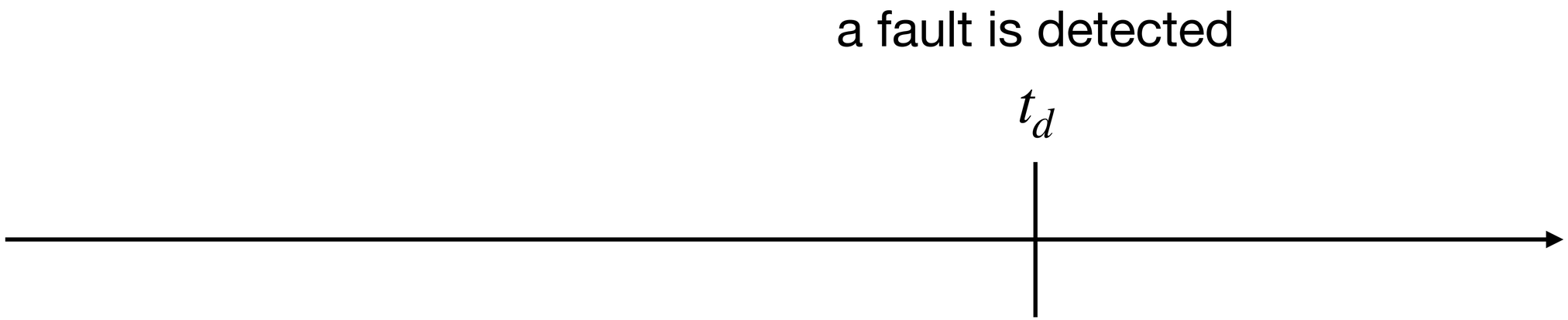
Intuition: A variable may point to an actionable mitigation method more likely than its descendants

Evaluation

Background
Analysis
Methodology
Evaluation
Conclusion

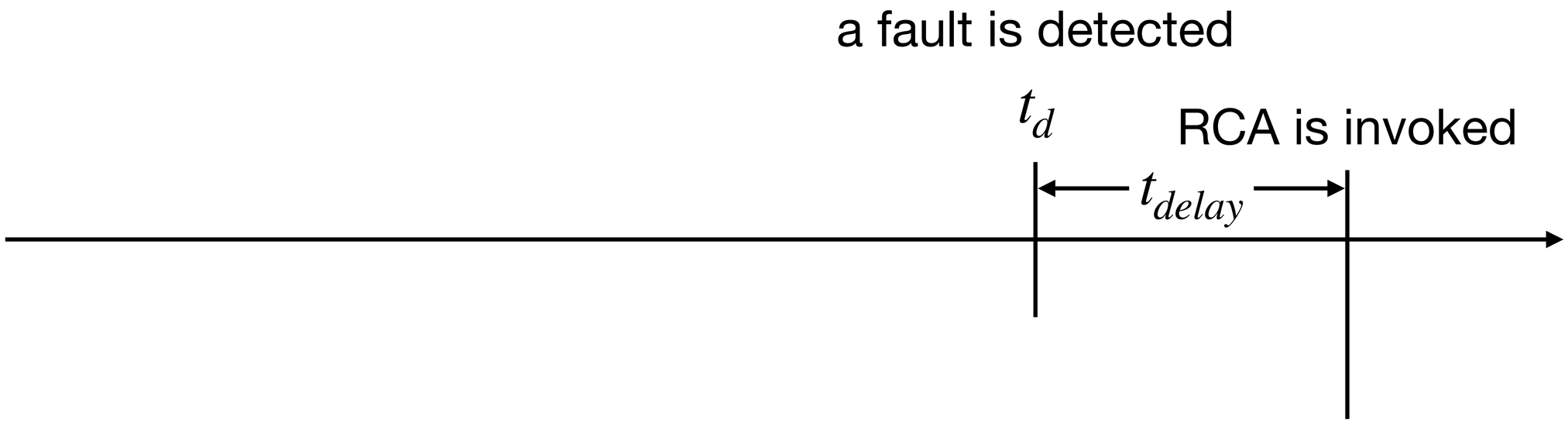
Experimental Setup

Hyperparameters



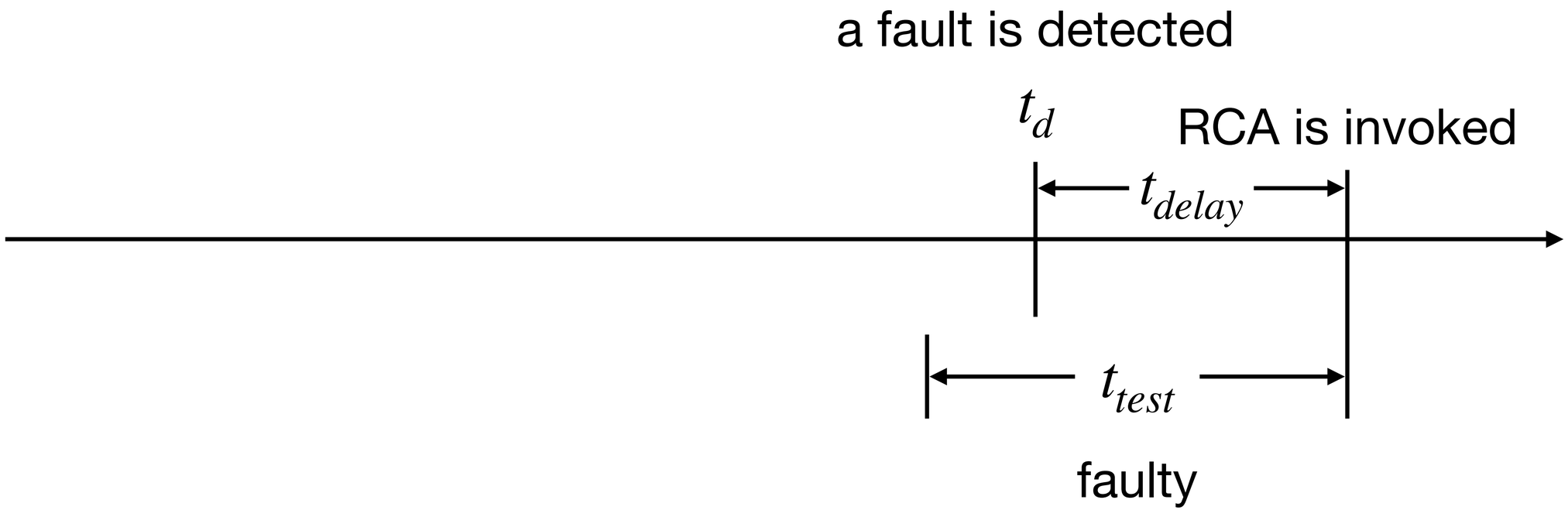
Experimental Setup

Hyperparameters



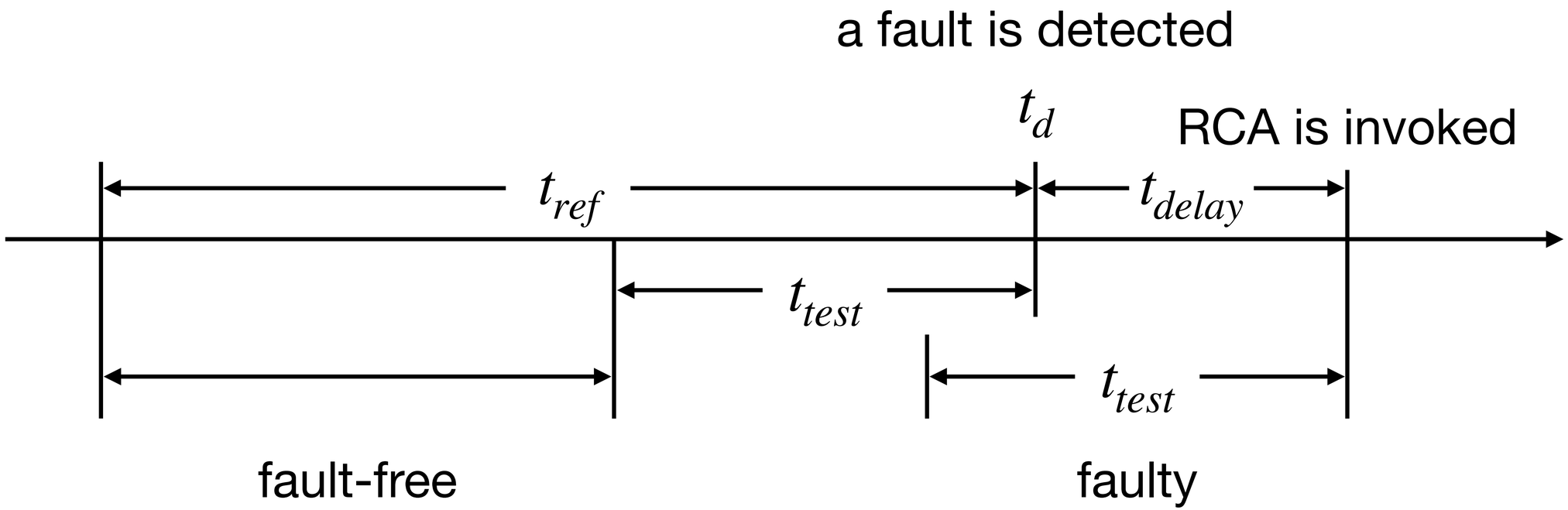
Experimental Setup

Hyperparameters



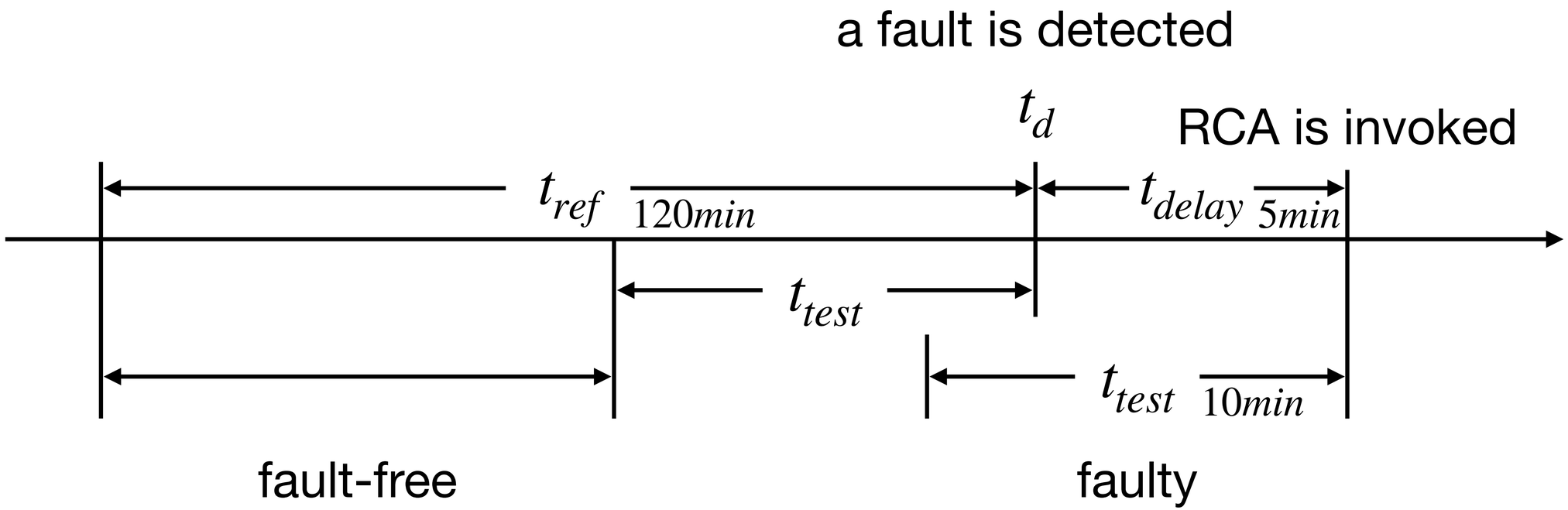
Experimental Setup

Hyperparameters



Experimental Setup

Hyperparameters



Experimental Setup

Evaluation Metrics

- Recall with the top-k results
  - $$AC@k = \frac{1}{|\mathcal{F}|} \sum_{\mathbf{M} \in \mathcal{F}} \frac{|\mathbf{M} \cap \{R_i(\mathbf{M}) \mid i = 1,2,\cdots,k\}|}{|\mathbf{M}|}$$
  - $k \leq K = 5$

Simulation Study

Data Generation

- $\mathbf{x}^{(t)} = \mathbf{A}\mathbf{x}^{(t)} + \beta\mathbf{x}^{(t-1)} + \epsilon^{(t)}$
  - $\mathbf{A}$ encodes the CBN, enforced to be a connected DAG with only the first node (service level indicator) having no children.

	#Node	#Edge	#Graph	#Case/Graph
$\mathscr{D}_{Sim}^{50}$	50	100	10	100
$\mathscr{D}_{Sim}^{100}$	100	500		
$\mathscr{D}_{Sim}^{500}$	500	5,000		

Simulation Study

Performance Evaluation

- RHT-PG: RHT with the **p**erfect **g**raph
  - $\mathbf{Pa}(X_i^{(t)}) = \mathbf{Pa}^{(t)}(X_i) \cup \left\{ X_i^{(t-1)} \right\}$

Scoring Method	$\mathscr{D}_{Sim}^{50}$		
	AC@1	AC@5	T (s)
NSigma	0.432	0.733	0.306
SPOT	0.508	0.761	6.601
DFS	0.541	0.682	0.308
DFS-MS	0.515	0.682	0.502
DFS-MH	0.178	0.217	0.501
RW-Par	0.188	0.433	0.714
RW-2	0.188	0.433	0.437
ENMF	0.116	0.278	0.624
CRD	0.074	0.223	4.844
RHT	*0.598*	*0.880*	0.338
RHT-PG	**0.615**	**0.952**	0.346
Ideal	0.617	0.999	

Simulation Study

Performance Evaluation

- RHT-PG: RHT with the **p**erfect **g**raph
  - $\mathbf{Pa}(X_i^{(t)}) = \mathbf{Pa}^{(t)}(X_i) \cup \left\{ X_i^{(t-1)} \right\}$
- Takeaways
  - RHT has theoretical reliability.

Scoring Method	$\mathscr{D}_{Sim}^{50}$		
	AC@1	AC@5	T (s)
NSigma	0.432	0.733	0.306
SPOT	0.508	0.761	6.601
DFS	0.541	0.682	0.308
DFS-MS	0.515	0.682	0.502
DFS-MH	0.178	0.217	0.501
RW-Par	0.188	0.433	0.714
RW-2	0.188	0.433	0.437
ENMF	0.116	0.278	0.624
CRD	0.074	0.223	4.844
RHT	*0.598*	*0.880*	0.338
RHT-PG	**0.615**	**0.952**	0.346
Ideal	0.617	0.999	

Simulation Study

Performance Evaluation

- RHT-PG: RHT with the **p**erfect **g**raph
  - $\mathbf{Pa}(X_i^{(t)}) = \mathbf{Pa}^{(t)}(X_i) \cup \left\{ X_i^{(t-1)} \right\}$
- Takeaways
  - RHT has theoretical reliability.
  - A broken CBN cannot guarantee a correct answer to RCA.

Scoring Method	$\mathscr{D}_{Sim}^{50}$		
	AC@1	AC@5	T (s)
NSigma	0.432	0.733	0.306
SPOT	0.508	0.761	6.601
DFS	0.541	0.682	0.308
DFS-MS	0.515	0.682	0.502
DFS-MH	0.178	0.217	0.501
RW-Par	0.188	0.433	0.714
RW-2	0.188	0.433	0.437
ENMF	0.116	0.278	0.624
CRD	0.074	0.223	4.844
RHT	*0.598*	*0.880*	0.338
RHT-PG	**0.615**	**0.952**	0.346
Ideal	0.617	0.999	

Simulation Study

Performance Evaluation

- RHT-PG: RHT with the **p**erfect **g**raph
  - $\mathbf{Pa}(X_i^{(t)}) = \mathbf{Pa}^{(t)}(X_i) \cup \left\{ X_i^{(t-1)} \right\}$
- Takeaways
  - RHT has theoretical reliability.
  - A broken CBN cannot guarantee a correct answer to RCA.
  - There may be statistical errors due to limited faulty data.

Scoring Method	$\mathscr{D}_{Sim}^{50}$		
	AC@1	AC@5	T (s)
NSigma	0.432	0.733	0.306
SPOT	0.508	0.761	6.601
DFS	0.541	0.682	0.308
DFS-MS	0.515	0.682	0.502
DFS-MH	0.178	0.217	0.501
RW-Par	0.188	0.433	0.714
RW-2	0.188	0.433	0.437
ENMF	0.116	0.278	0.624
CRD	0.074	0.223	4.844
RHT	*0.598*	*0.880*	0.338
RHT-PG	**0.615**	**0.952**	0.346
Ideal	0.617	0.999	

Simulation Study

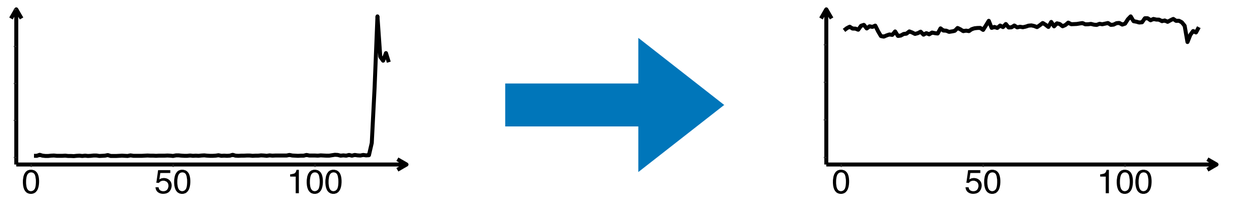
Robustness Evaluation

- Classify faults based on the change of the root cause metrics when the service level indicator is abnormal
  - **Weak**: dramatically
  - **Strong**: slight
  - **Mixed**: both

Simulation Study

Robustness Evaluation

- Classify faults based on the change of the root cause metrics when the service level indicator is abnormal
  - **Weak**: dramatically
  - **Strong**: slight
  - **Mixed**: both

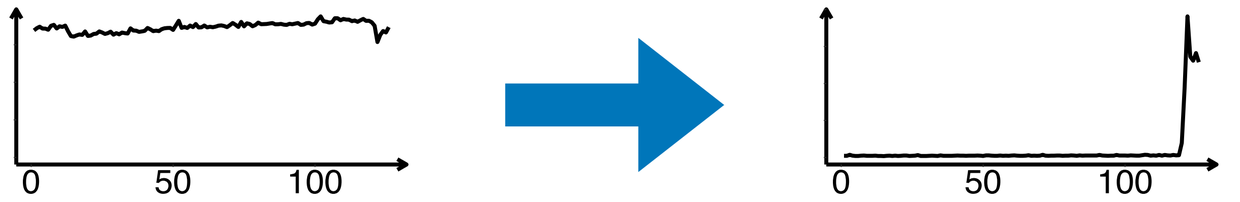


Weak fault **Service Level Indicator**

Simulation Study

Robustness Evaluation

- Classify faults based on the change of the root cause metrics when the service level indicator is abnormal
  - **Weak**: dramatically
  - **Strong**: slight
  - **Mixed**: both



Strong fault **Service Level Indicator**

Simulation Study

Robustness Evaluation

- Classify faults based on the change of the root cause metrics when the service level indicator is abnormal
  - **Weak**: dramatically
  - **Strong**: slight
  - **Mixed**: both
- Takeaway
  - RHT is more robust.

Scoring Method	Weak (916)		Mixed (64)		Strong (20)	
	AC@1	AC@5	AC@1	AC@5	AC@1	AC@5
NSigma	0.454	0.753	0.249	0.498	0.000	0.550
SPOT	0.534	0.783	0.293	0.503	0.000	0.550
DFS	0.558	0.707	0.282	0.368	0.550	0.550
DFS-MS	0.531	0.707	0.277	0.368	0.550	0.550
DFS-MH	0.184	0.223	0.069	0.123	0.250	0.250
RW-Par	0.194	0.445	0.142	0.300	0.050	0.300
RW-2	0.194	0.445	0.142	0.300	0.050	0.300
ENMF	0.111	0.269	0.124	0.321	0.300	0.550
CRD	0.071	0.207	0.088	0.353	0.150	0.550
RHT	*0.613*	*0.888*	*0.325*	*0.730*	*0.800*	**1.000**
RHT-PG	**0.624**	**0.954**	**0.358**	**0.914**	**1.000**	**1.000**
Ideal	0.627	1.000	0.358	0.995	1.000	1.000

Empirical Study on Oracle Database Data

Setup

- Dataset
  - 99 faults with high Average Active Sessions (AAS) from Oracle databases in a large banking system
- Implementation
  - Our structural graph contains 197 monitoring metrics with 2,641 edges
    - denoted as *Structural*
  - Equip RHT with descendant adjustment
    - denoted as *CIRCA*

Empirical Study on Oracle Database Data

Performance Evaluation

- Takeaways
  - CIRCA outperforms baselines.

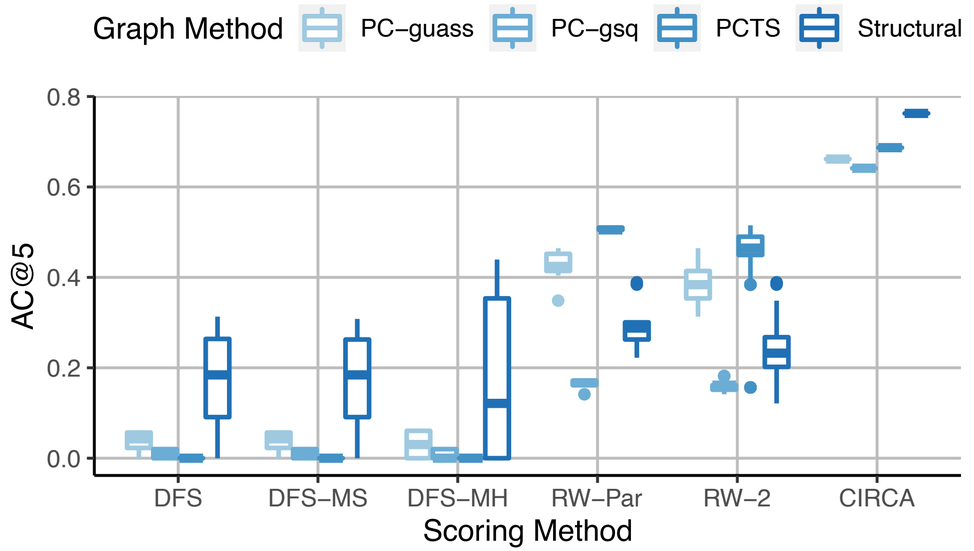
Scoring Method	Graph Method	AC@1	AC@5	T (s)
NSigma	Empty	0.323	0.662	0.472
SPOT	Empty	0.152	0.419	5.027
DFS	Structural	0.187	0.313	0.483
DFS-MS	Structural	0.207	0.308	0.839
DFS-MH	Structural	0.268	0.439	0.844
RW-Par	PCTS	0.086	0.449	24.695
RW-2	PCTS	0.086	0.449	24.559
ENMF	Empty	0.111	0.374	0.771
CRD	Empty	0.035	0.313	4.787
CIRCA	Structural	**0.404**	**0.763**	0.578
Ideal		0.929	1.000	

Empirical Study on Oracle Database Data

Ablation Study

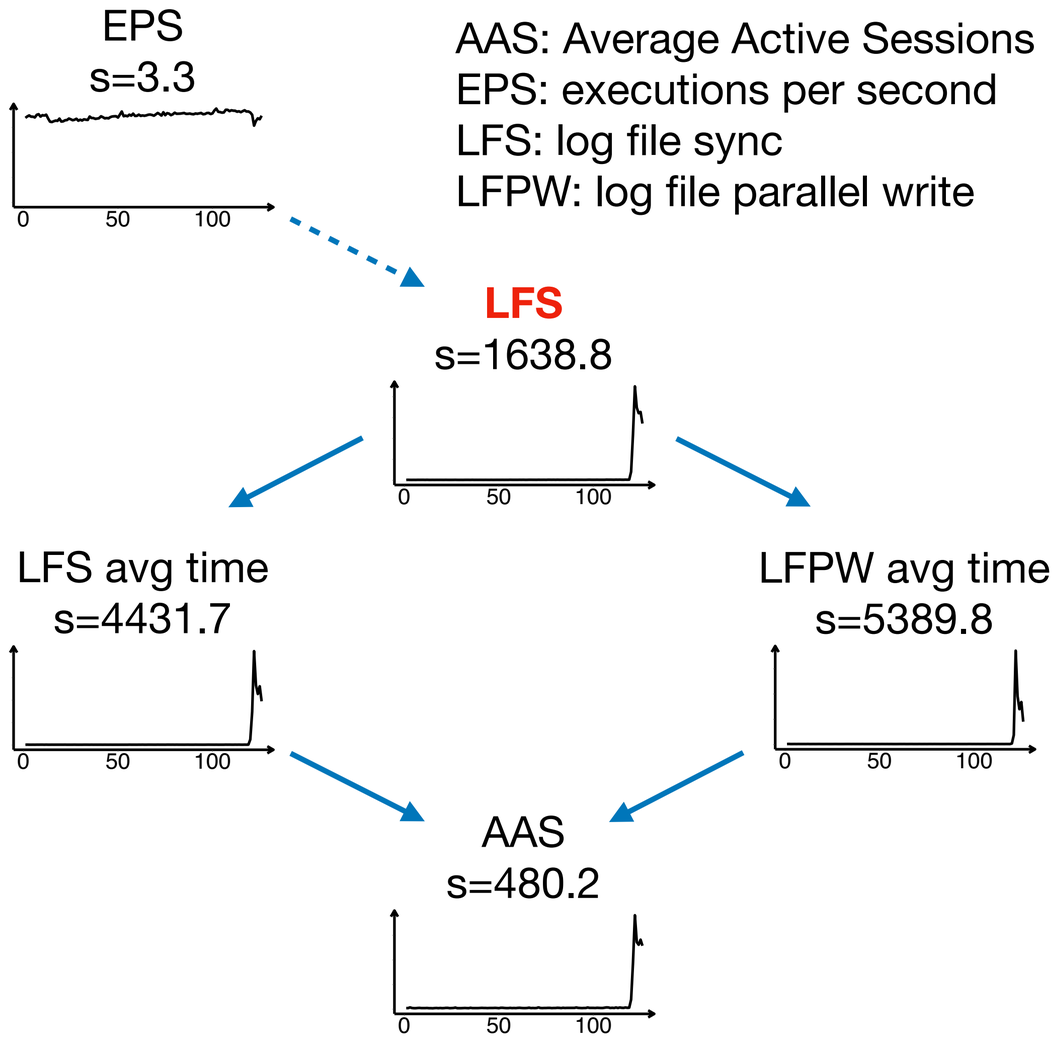
- Takeaways
  - CIRCA outperforms baselines.
  - Each of the 3 proposed techniques has a positive effect.
  - Search-based methods also benefit from the proposed structural graph.

Scoring Method	Graph Method	AC@1	AC@5	T (s)
NSigma	Empty	0.323	0.662	0.472
RHT	Structural	0.328	0.677	0.576
CIRCA	Structural	**0.404**	**0.763**	0.578
Ideal		0.929	1.000	



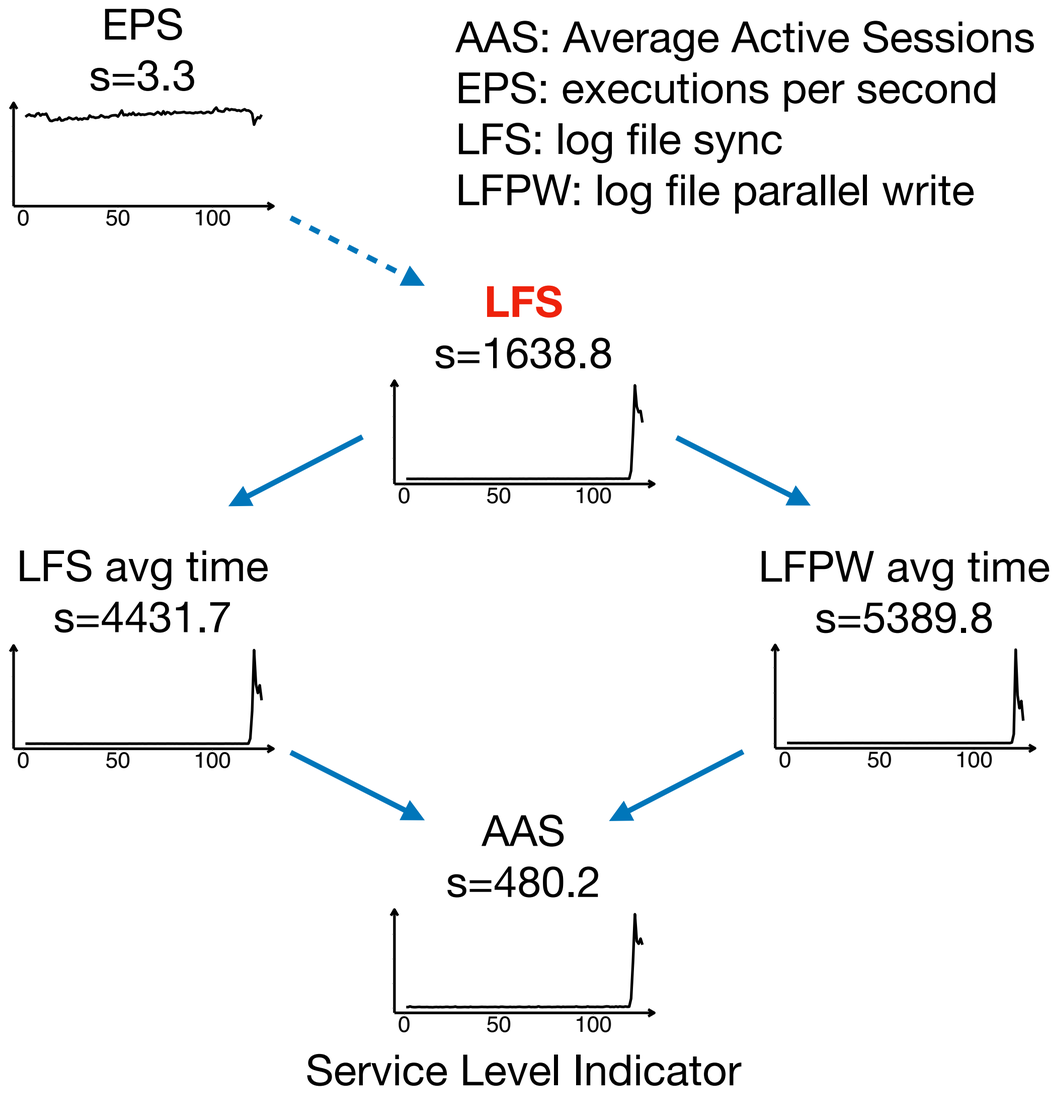
Empirical Study on Oracle Database Data

Case Study



Empirical Study on Oracle Database Data

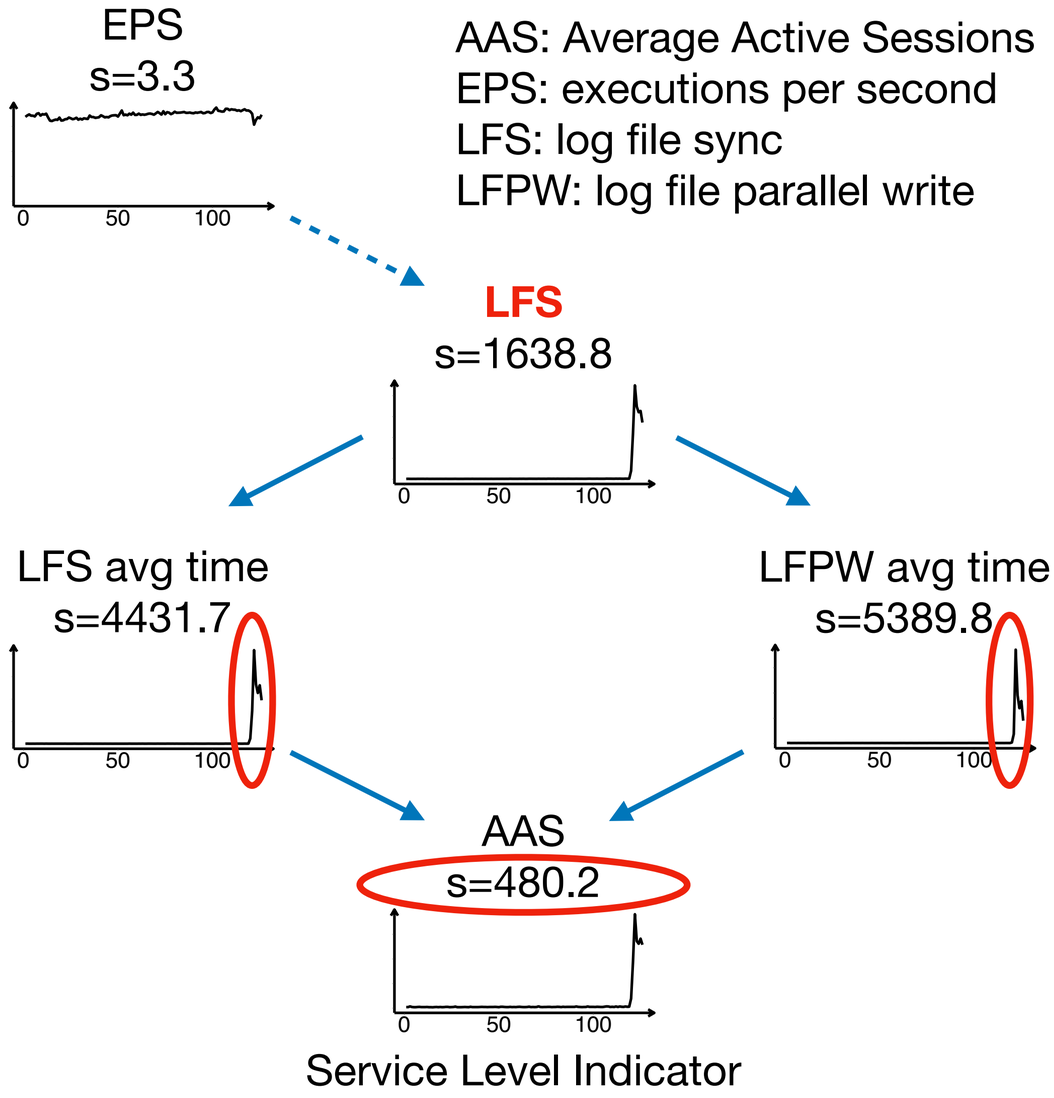
Case Study



Empirical Study on Oracle Database Data

Case Study

RHT confronts incomplete observational knowledge

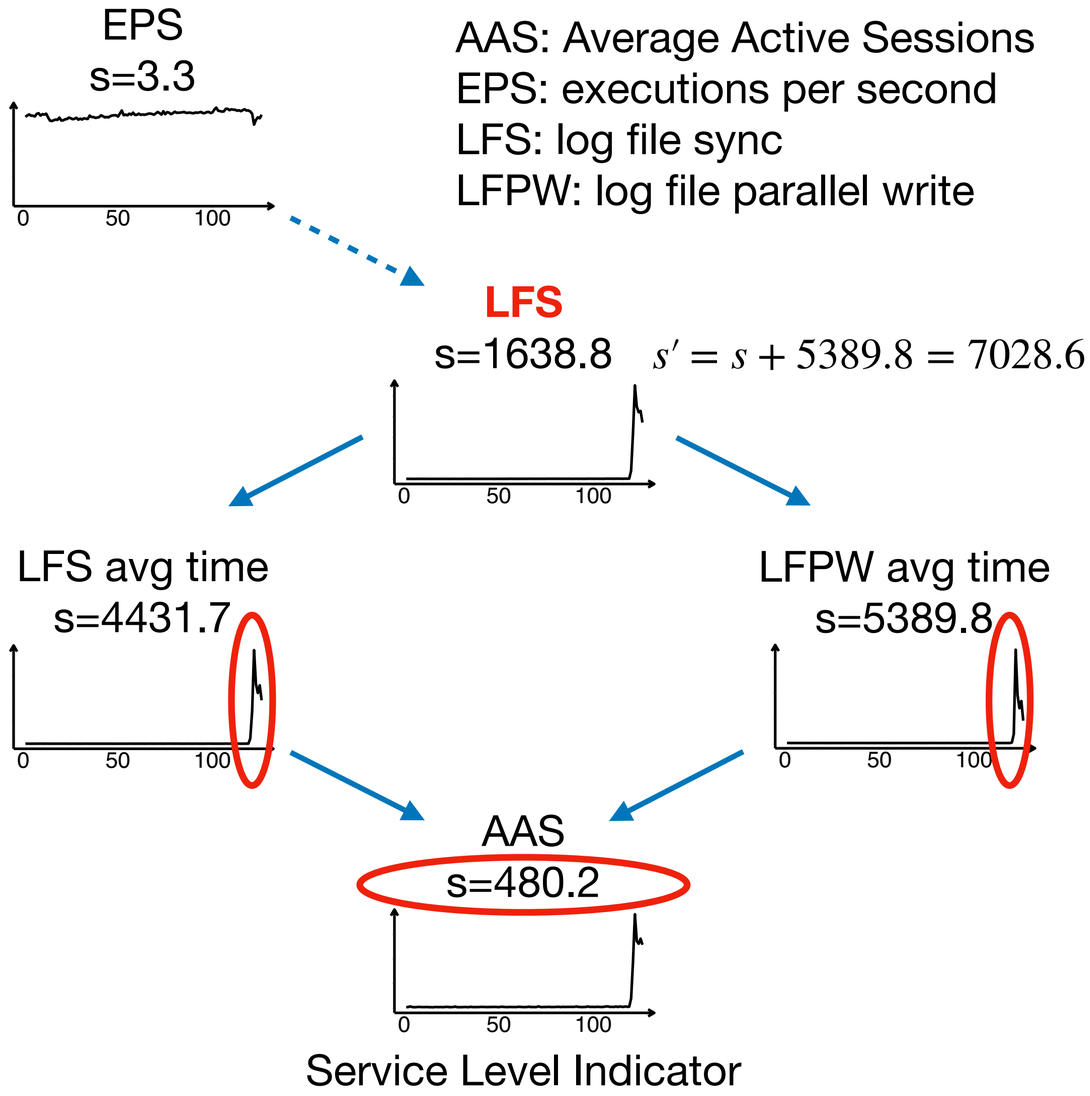


Empirical Study on Oracle Database Data

Case Study

RHT confronts incomplete observational knowledge

Descendant adjustment helps CIRCA rank LFS ahead



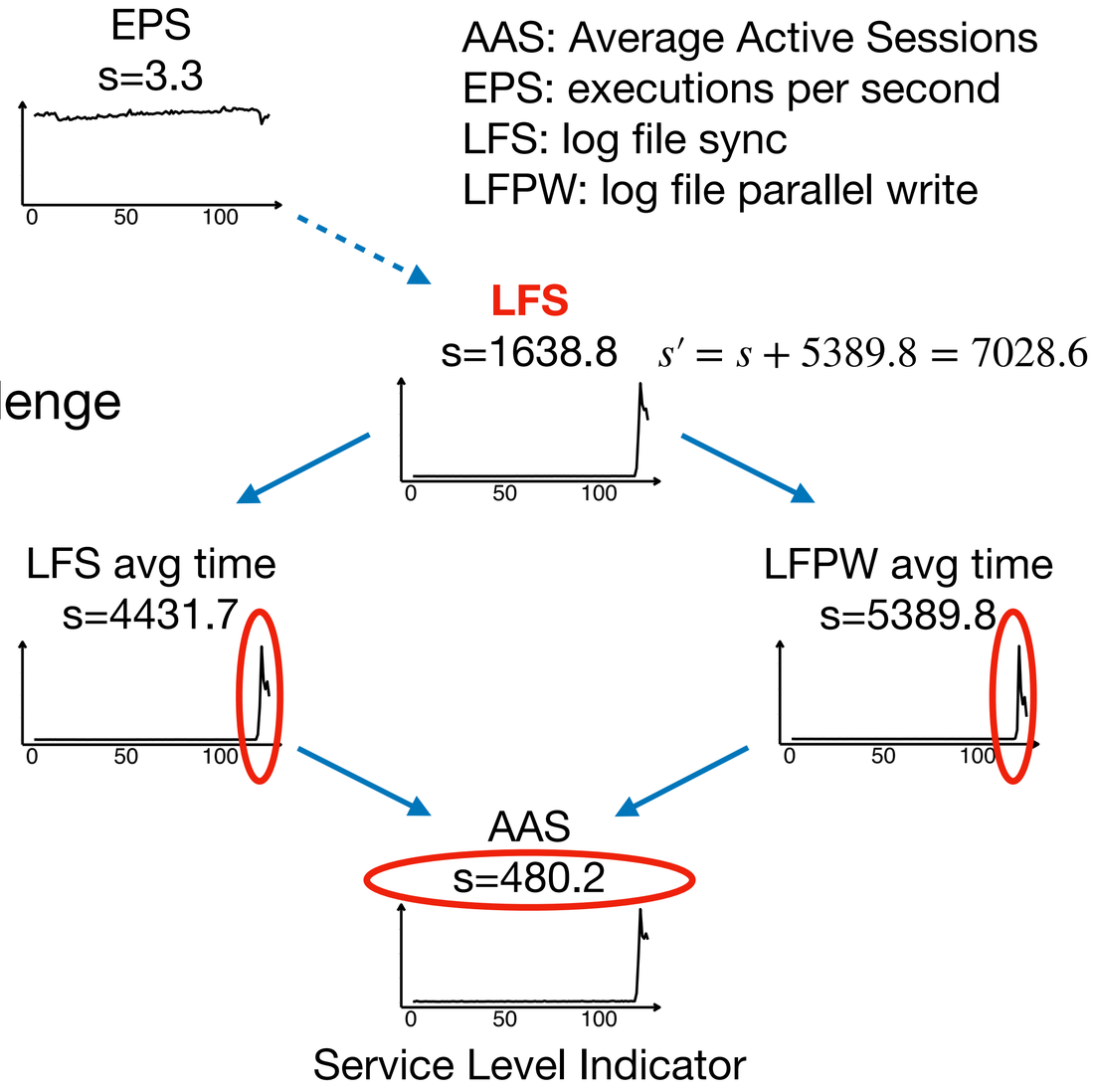
Empirical Study on Oracle Database Data

Case Study

RHT confronts incomplete observational knowledge

- Further advancement should handle this challenge

Descendant adjustment helps CIRCA rank LFS ahead



Empirical Study on Oracle Database Data

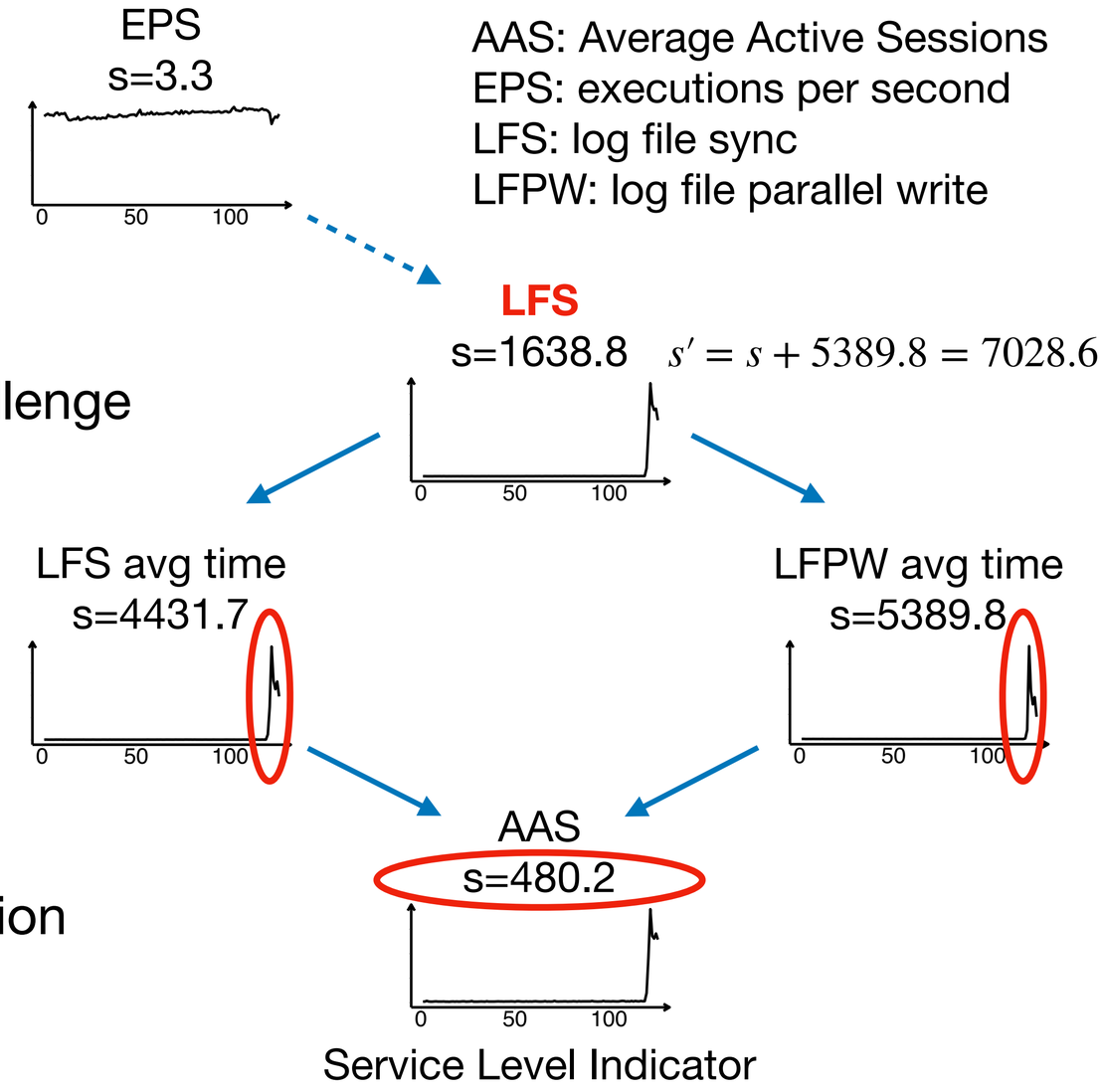
Case Study

RHT confronts incomplete observational knowledge

- Further advancement should handle this challenge

Descendant adjustment helps CIRCA rank LFS ahead

- CIRCA outperforms pure RHT in this case
- Descendant adjustment needs more verification

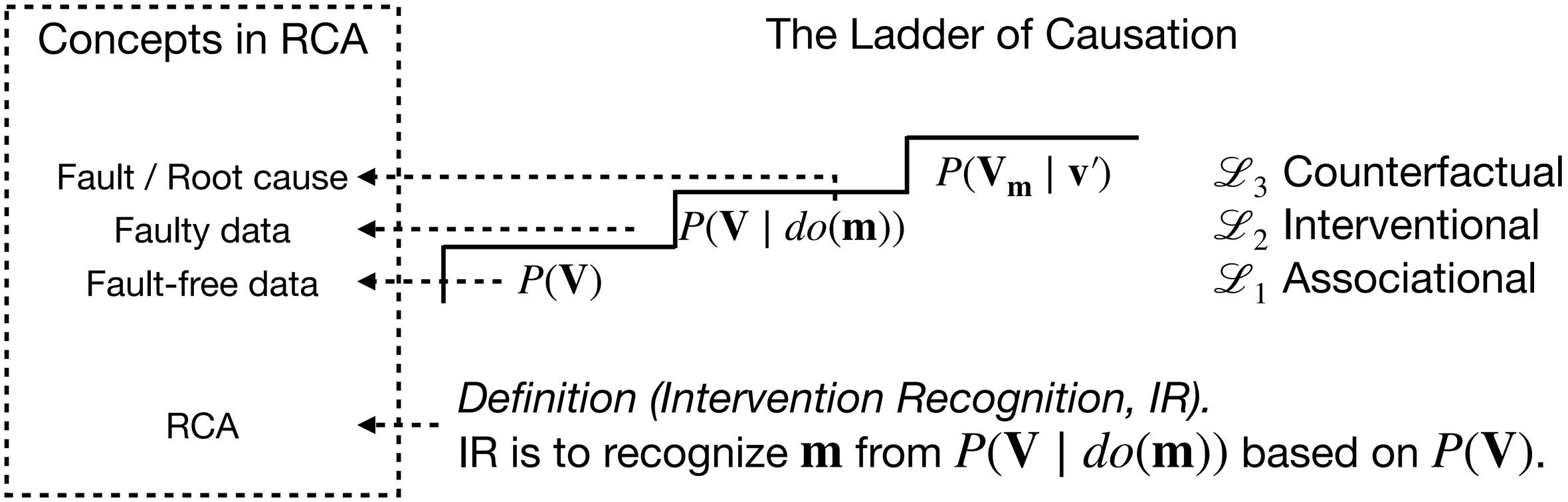


Conclusion

Background
Analysis
Methodology
Evaluation
Conclusion

Contributions

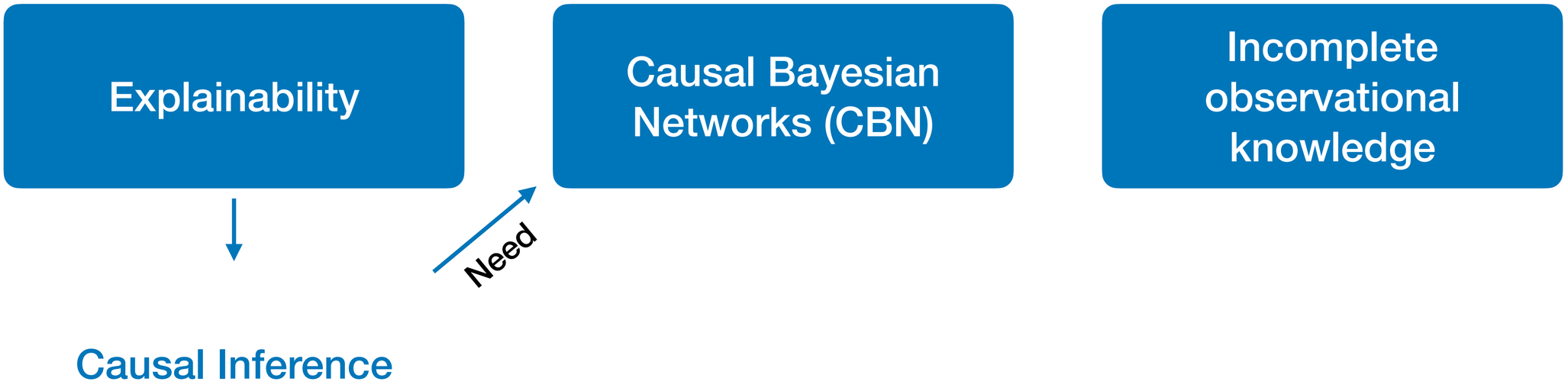
Formulate RCA as a causal inference task



Definition (Intervention Recognition, IR).
IR is to recognize $\mathbf{m}$ from $P(\mathbf{V} \mid do(\mathbf{m}))$ based on $P(\mathbf{V})$.

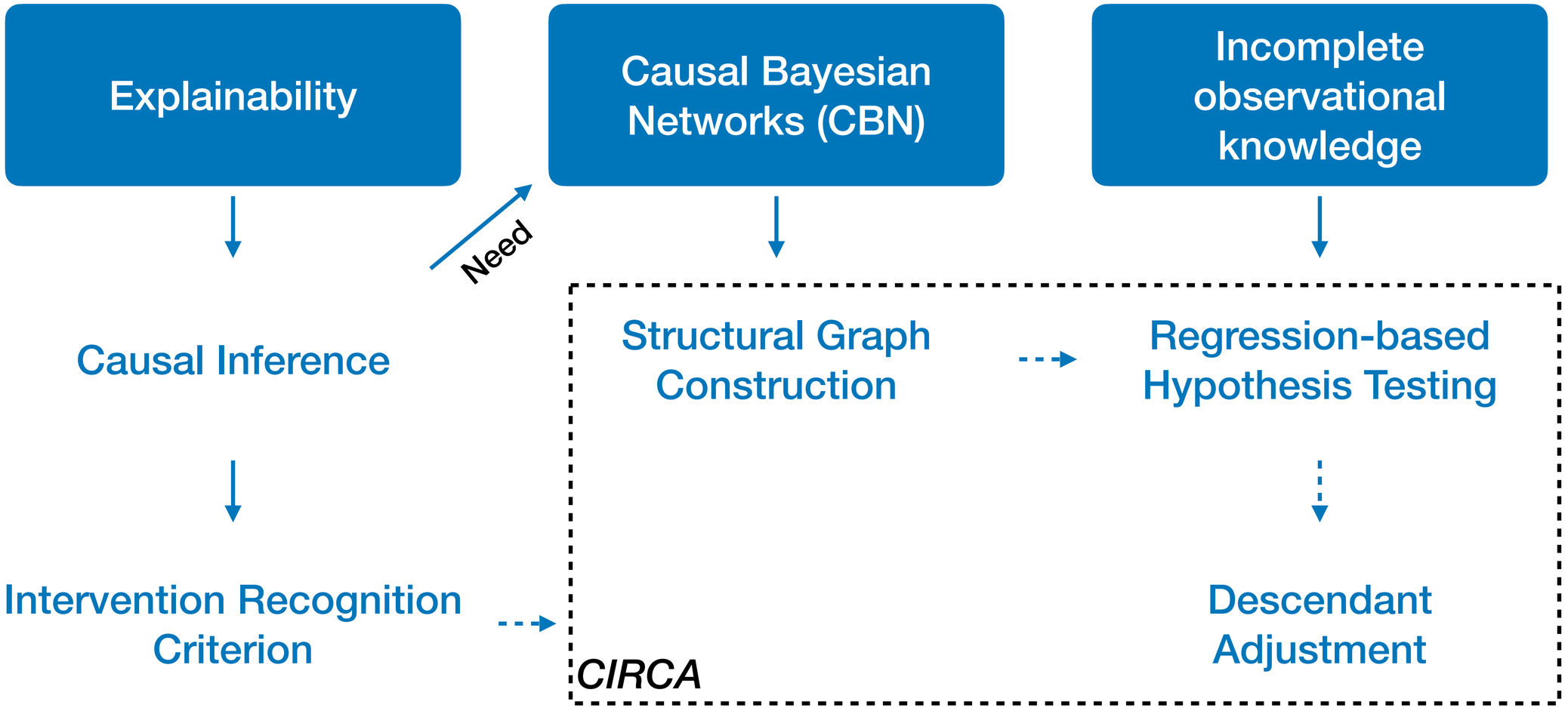
Contributions

Causal Inference-based Root Cause Analysis (CIRCA) for OSS



Contributions

Causal Inference-based Root Cause Analysis (CIRCA) for OSS



Contributions

Evaluation with both simulation and real-world datasets

Simulation Study

Empirical Study

Contributions

Evaluation with both simulation and real-world datasets

Simulation Study → Theoretical Reliability

Empirical Study → Practical Value

Causal Inference-based Root Cause Analysis

Thanks for listening

https://github.com/NetManAIOps/CIRCA