SIGCOMM 2026
DENVER

Alibaba Cloud
Worldwide Cloud Services Partner

Verifying Non-Deterministic Convergence on a Global Production WAN

Xing Fang[1,2], Fangdan Ye[2], Yifei Yuan[2], Zhongyu Guan[2], Duncheng She[2], Xiaobo Zhu[2], Qiao Xiang[1]

[1] Xiamen University, [2] Alibaba Cloud

Alibaba Cloud

Alibaba Cloud Operates a Global WAN

Alibaba Cloud
Worldwide Cloud Services Partner

32 regions **105 AZs** **3200+ Edge sites**

Operating Alibaba's Global WAN Is Hard

Alibaba Cloud
Worldwide Cloud Services Partner

Config Update

Post-change behavior?

Large Scale

~10^3 routers, ~10^6 prefixes

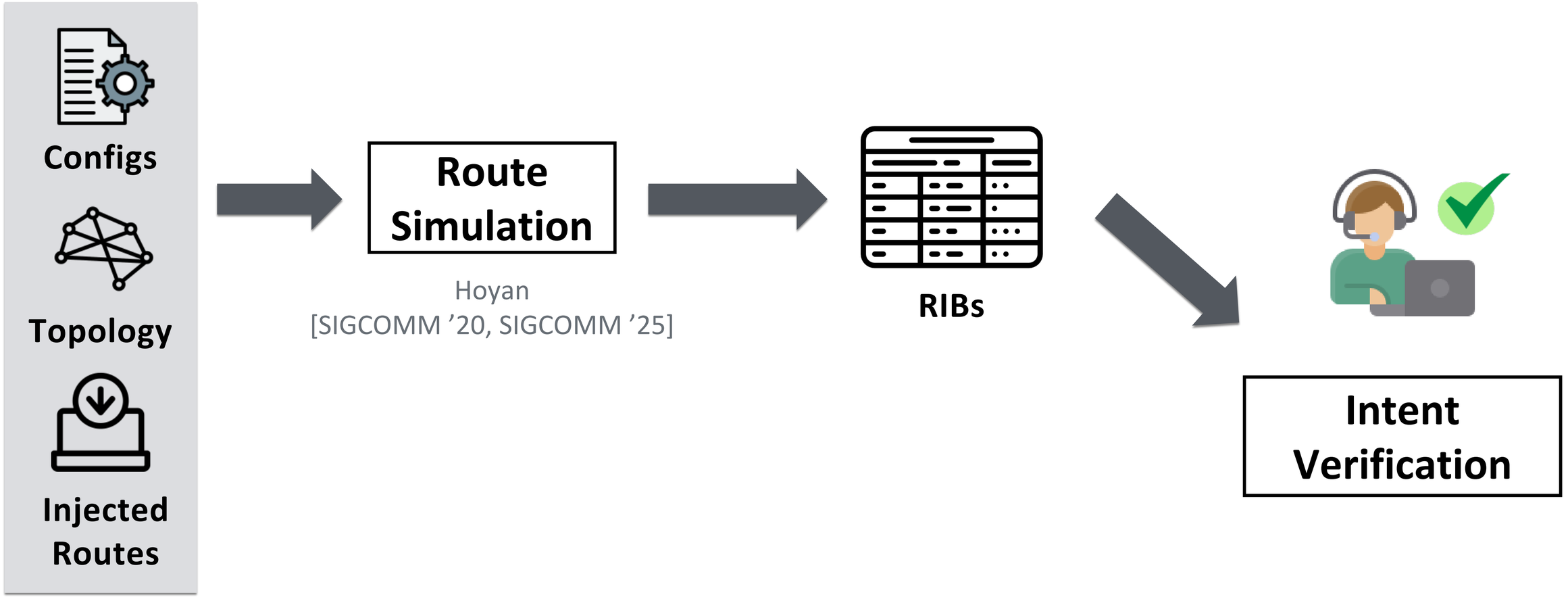
Protocol Complexity

BGP, IS-IS, MPLS/SR, SRv6, …

Vendor Differences

Multiple Vendors & software versions

Verifying Network Changes through Simulation



Verifying Network Changes through Simulation

Route Simulation

Hoyan
[SIGCOMM ’20, SIGCOMM ’25]

RIBs

Traffic Simulation

Jingubang [NSDI ’24]
YU [SIGCOMM ’24]

20%

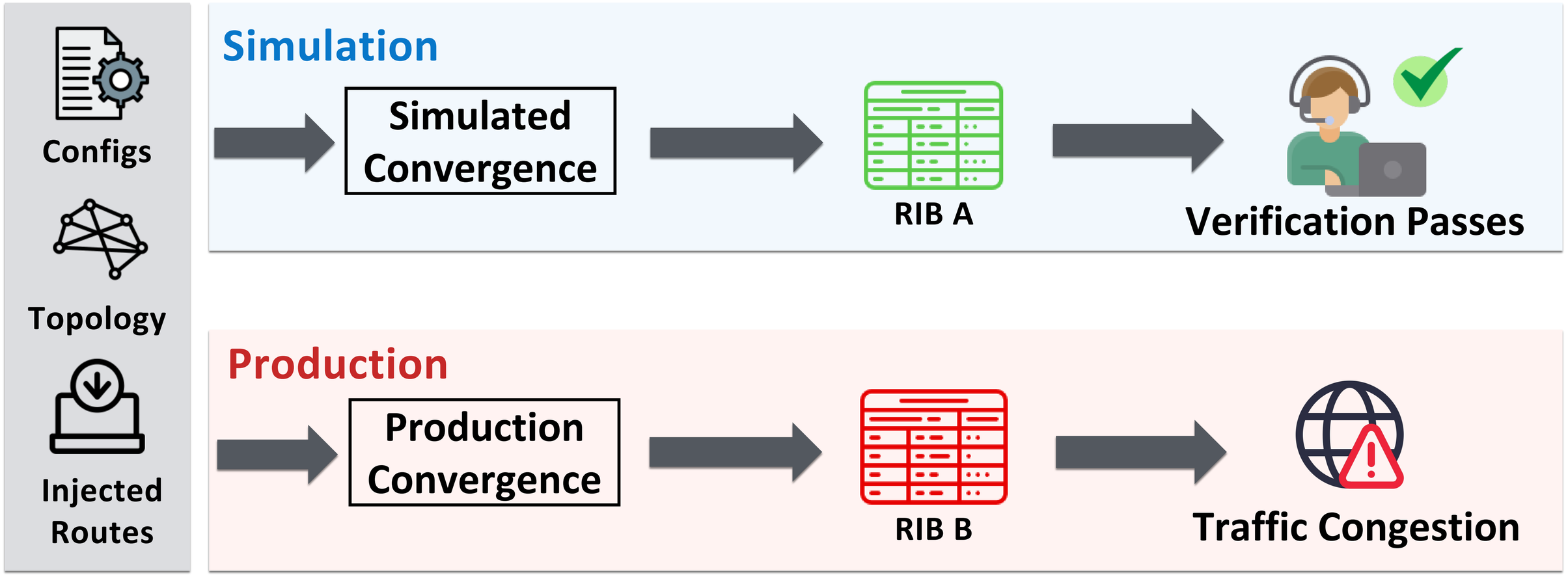
Flow paths / Traffic loads

Intent Verification

In production since 2018

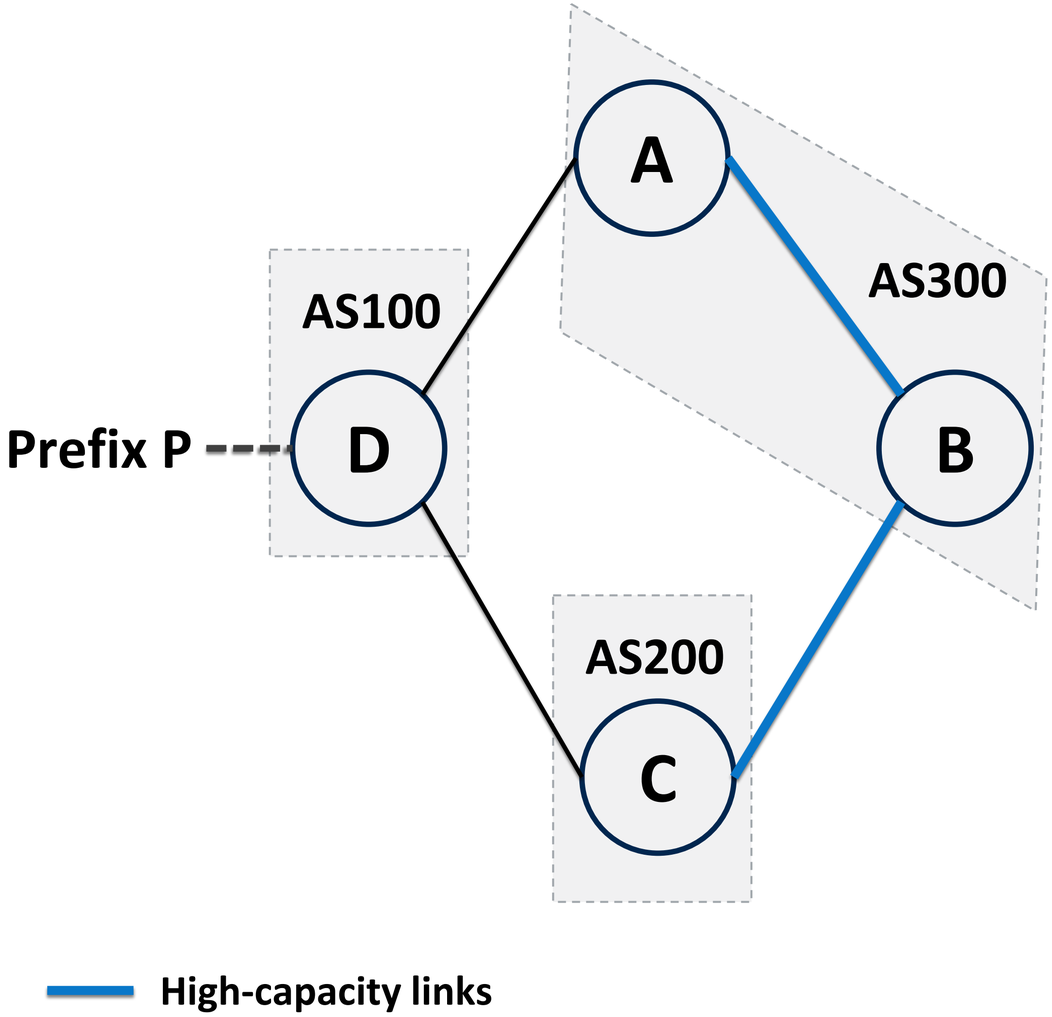
Misconfiguration-caused incidents: 56% → 5%

Motivation: Non-determinism Undermines Verification



Same inputs → Different convergence outcomes → Unreliable verification

A Real-World Case Study



Config Update:

A assigns higher LocalPref to routes learned from B

Goal: utilize B's high-capacity links

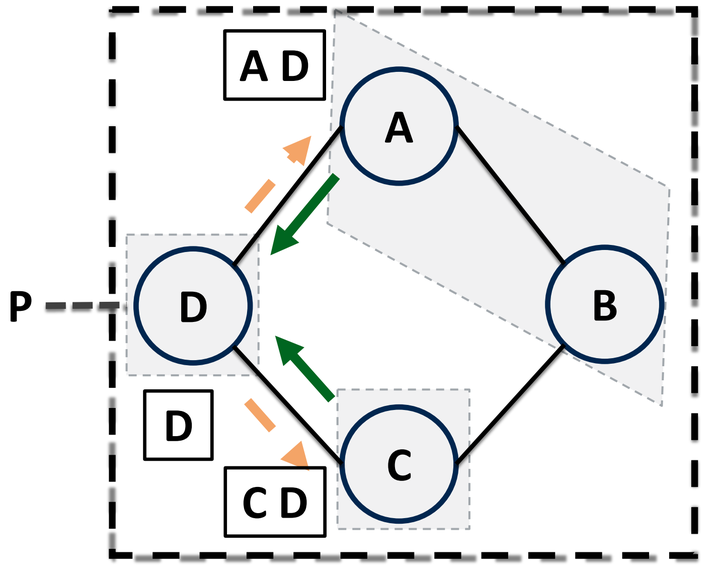
Intents:

Reachability: Every node can reach P

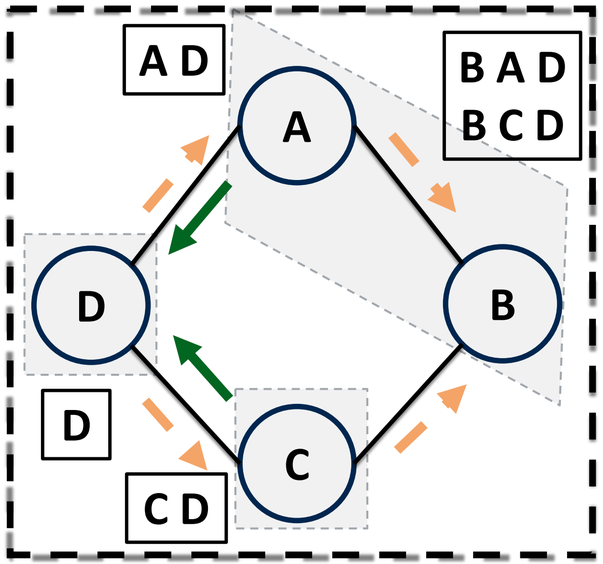
Capacity safety: No link is overloaded

Simulation Converges to a Safe Outcome

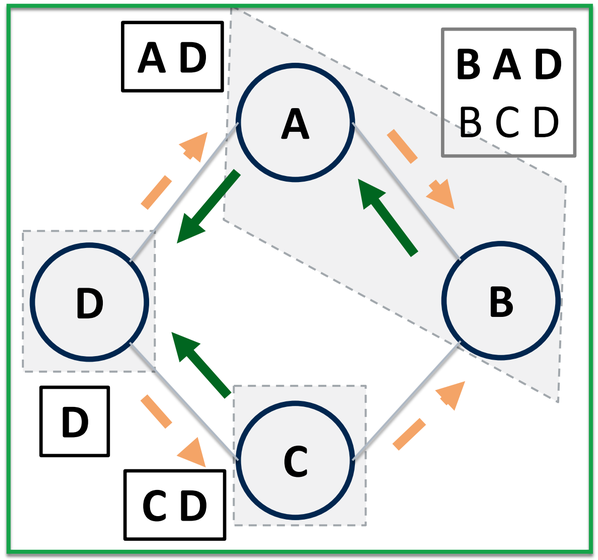
Step 1:
D advertises P



Step 2:
A and C advertise P



Step 3: Converged
B selects A
(shorter AS-PATH)



Verification Passes

Route propagation

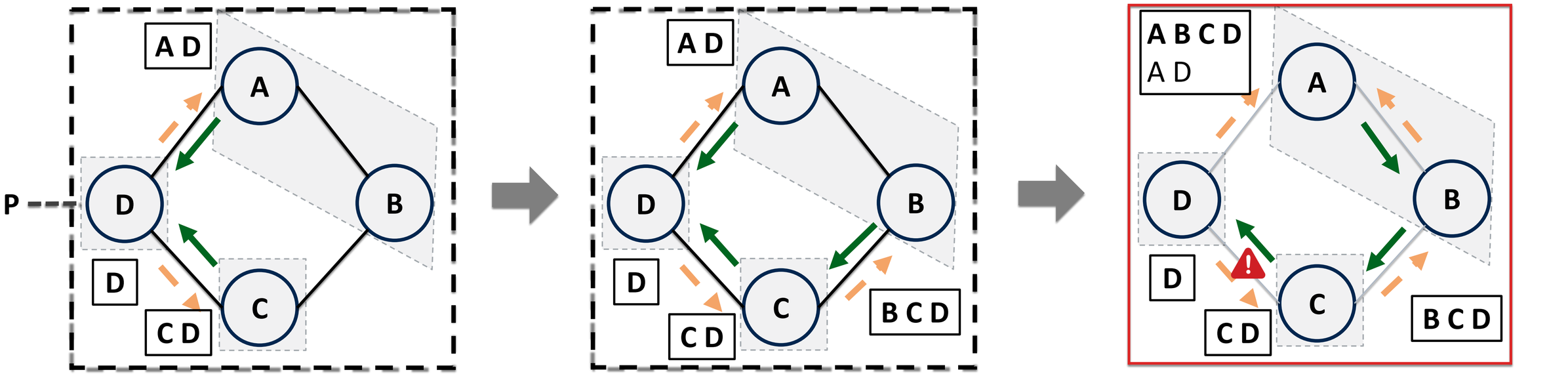
Traffic forwarding

Production Can Converge to an Unsafe Outcome

Step 1:
D advertises P

Step 2:
~~A and~~ C advertise P

Step 3: Converged
B advertises P
A selects B
(higher LocalPref)



C–D Link Overloaded

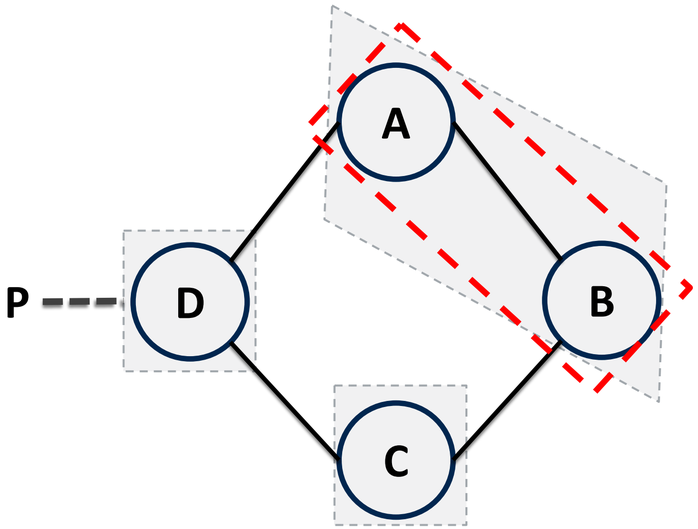
Route propagation

Traffic forwarding

Advertisement Order Determines the Converged State

Root Cause: Mutual Route Preferences

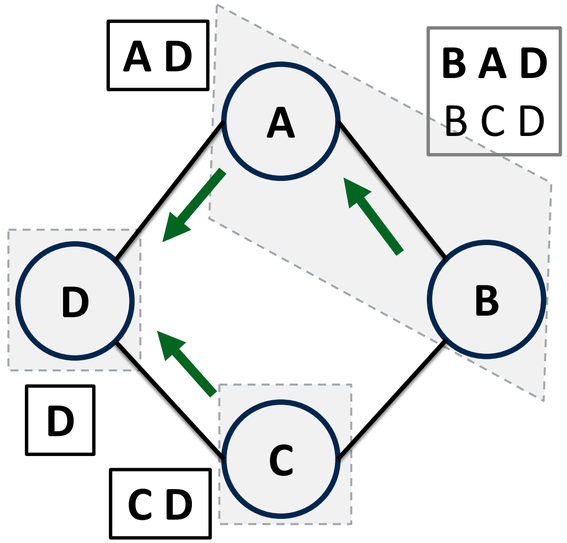
A prefers B; B prefers A



A advertises P first

B selects A

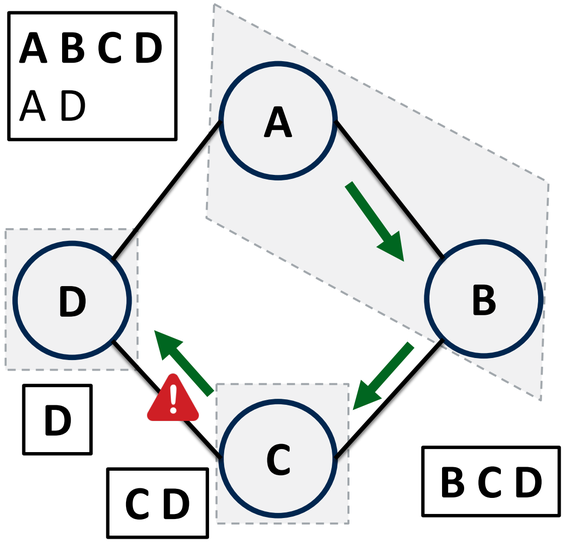
(shorter AS-PATH)



B advertises P first

A selects B

(higher LocalPref)



Our Goal

Prevent Non-Deterministic Convergence from Harming Verification Accuracy

Identify — Affected prefixes

Pinpoint — Their root causes

Improve — Verification accuracy

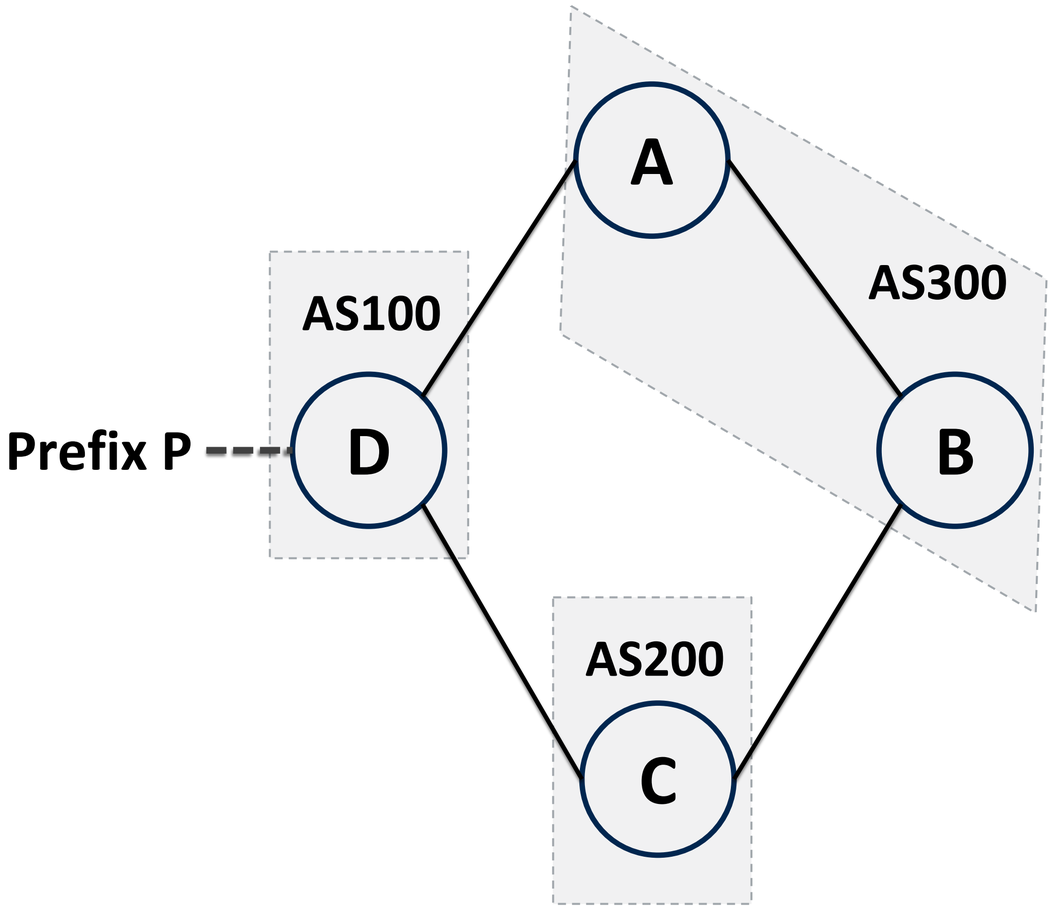
Why Can’t Existing Tools Help?

Alibaba Cloud
Worldwide Cloud Services Partner

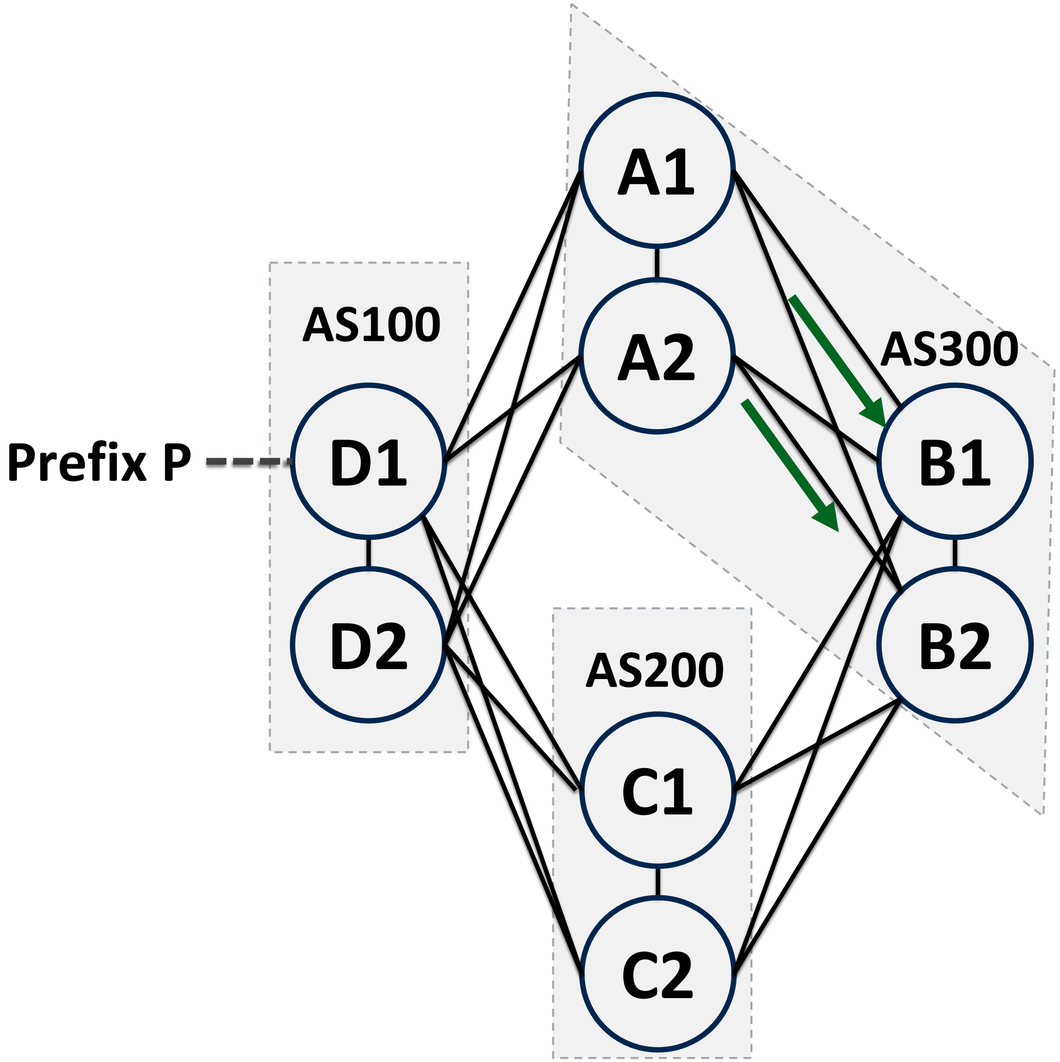
SMT-based	Encode network constraints → Solve formulas Solver bottleneck ⚠️	Minesweeper [SIGCOMM’17] BiNode [INFOCOM’20, TON’22] NetSMT [INFOCOM’24]
SPP-based	Enumerate routing paths → Check conditions Path explosion ⚠️	Greedy+ [TNSM’11] Hoyan [SIGCOMM’20] UPEA [ICNP’25]
Model-checking-based	Explore state transitions → Find converged states State explosion ⚠️	Plankton [NSDI’20]

Struggle to Scale to WANs with $O(10^3)$ Routers

Key Insight: Group-Level Routing Similarity



Key Insight: Group-Level Routing Similarity



Device Groups:

Routers are grouped by site and role for redundancy.

Configuration Similarity:

Group members share configuration templates.

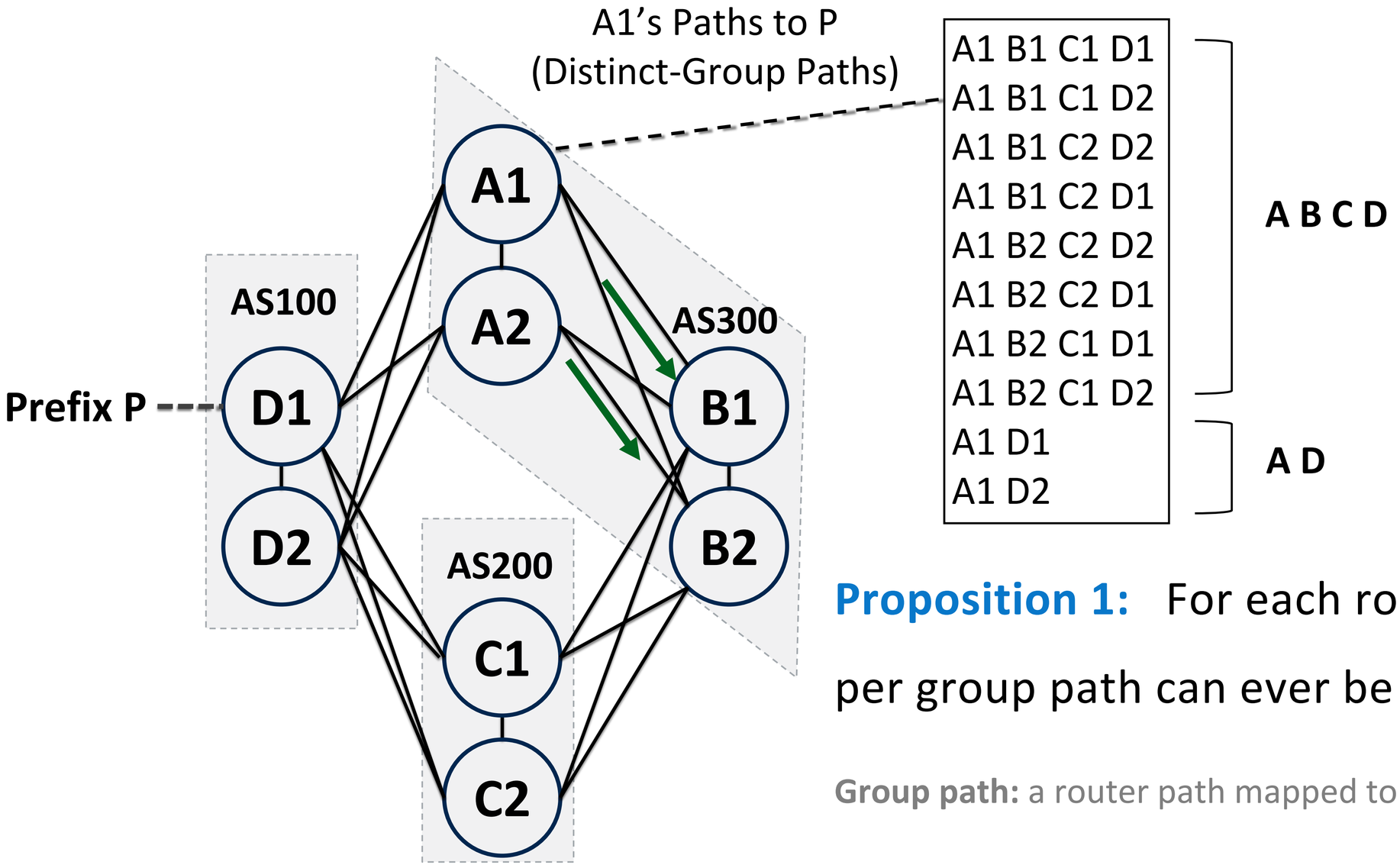
designed to achieve

Routing Similarity:

Group members select next hops from the same device group.

E.g., A1 selects B1, A2 selects B2

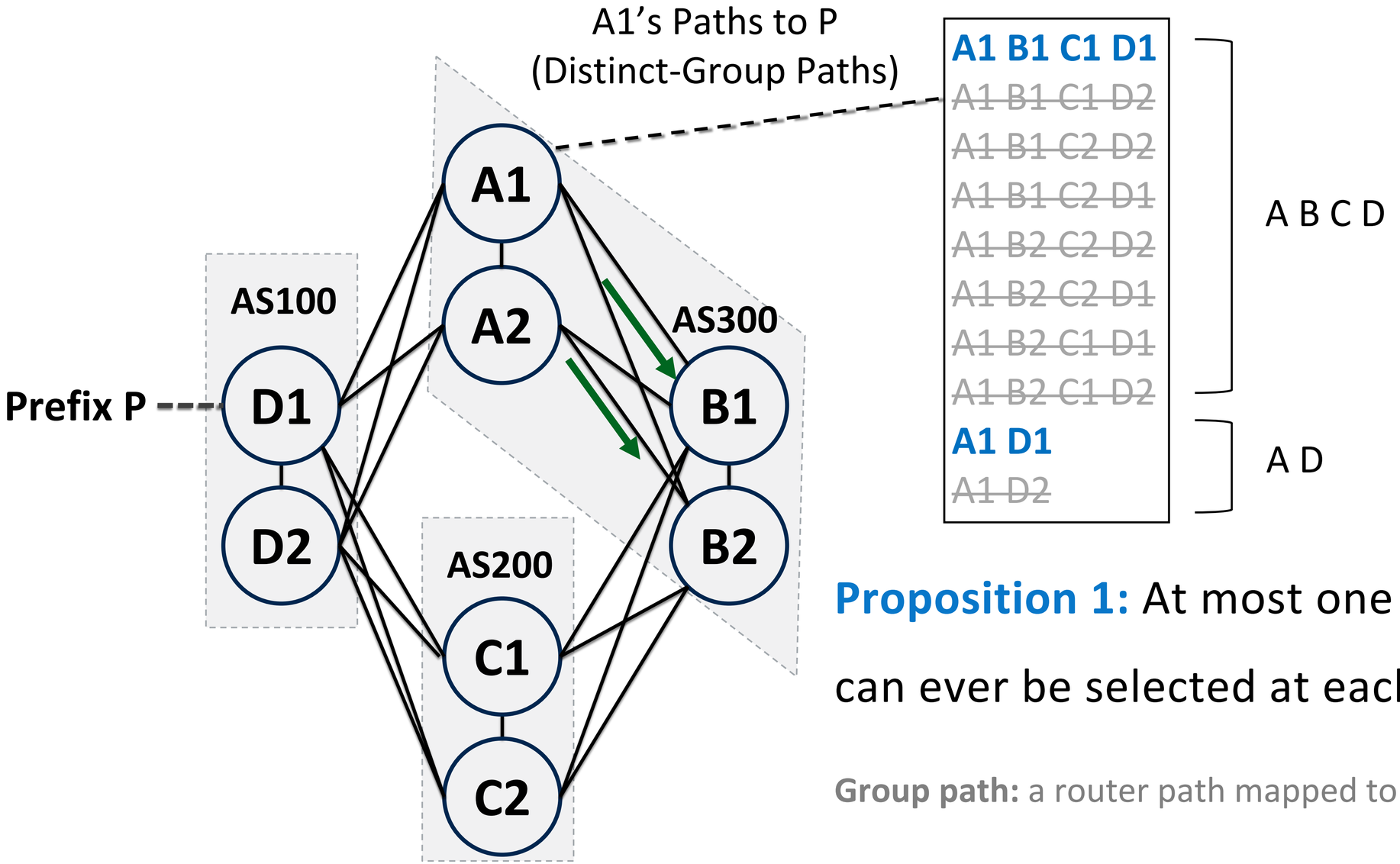
Routing Similarity Prunes Redundant Search



Proposition 1: For each router, at most one path per group path can ever be selected.

Group path: a router path mapped to its sequence of device groups

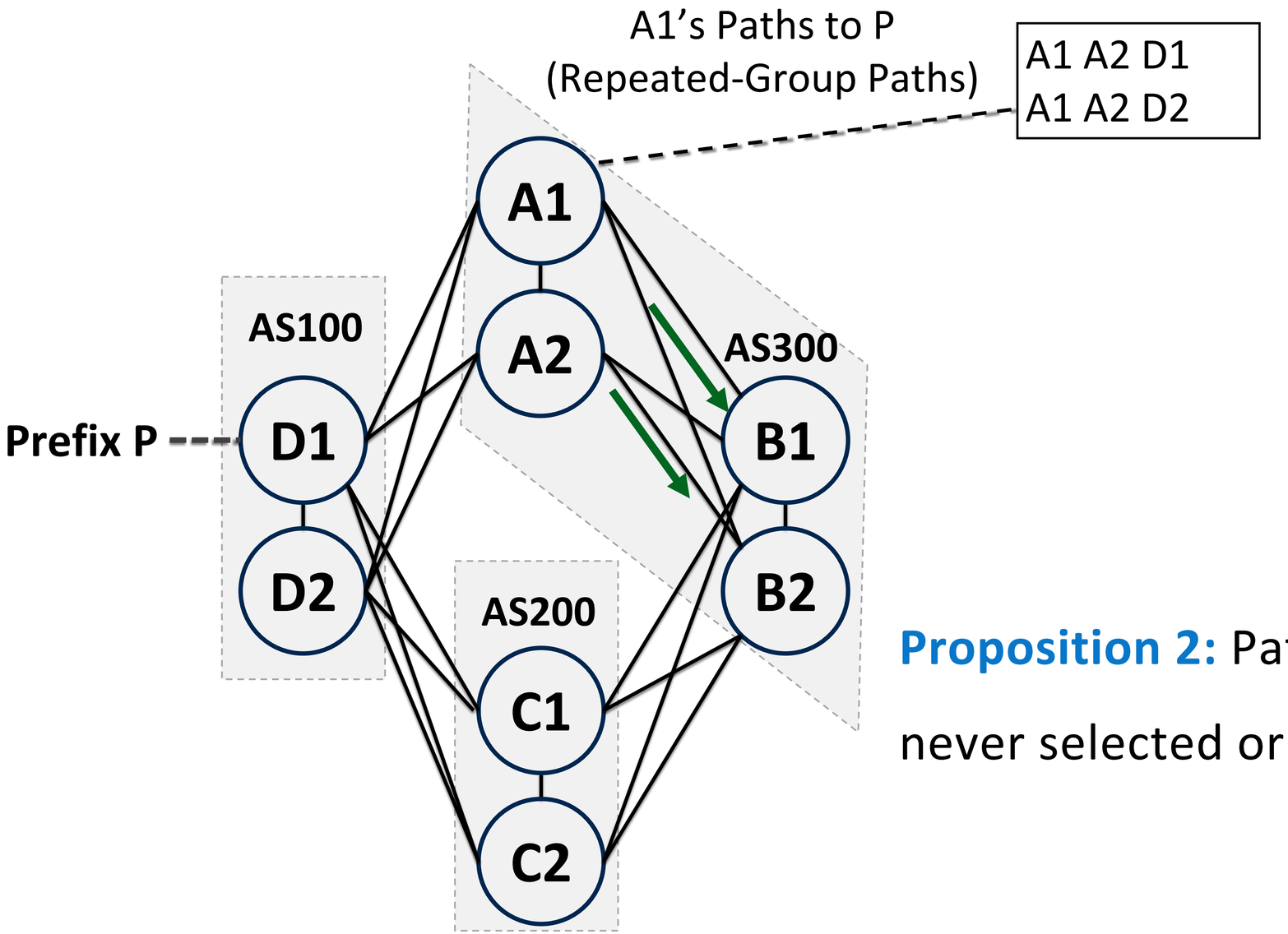
Routing Similarity Prunes Redundant Search



Proposition 1: At most one path per group path can ever be selected at each router.

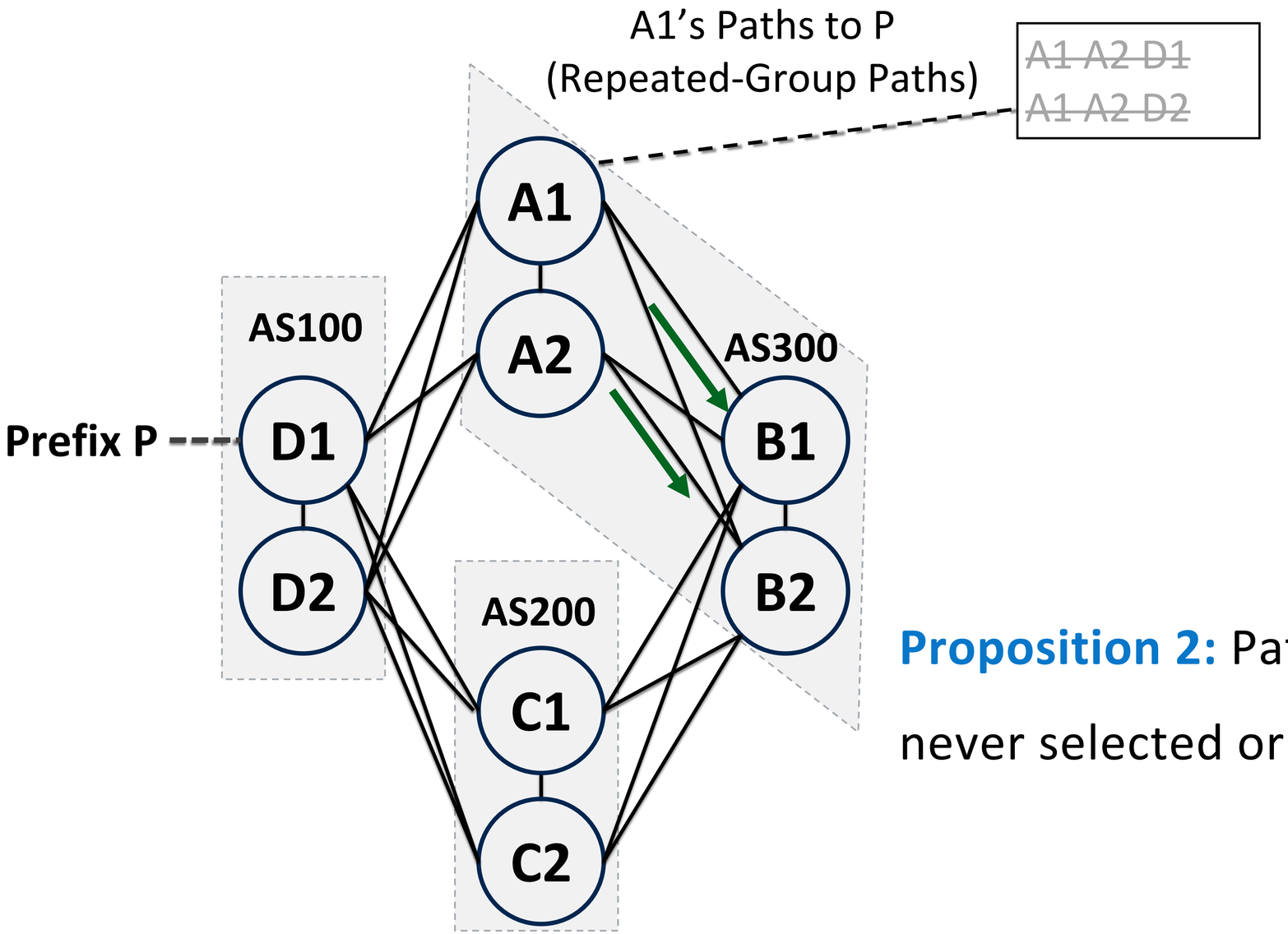
Group path: a router path mapped to its sequence of device groups

Routing Similarity Prunes Redundant Search



Proposition 2: Paths with repeated groups are never selected or advertised.

Routing Similarity Prunes Redundant Search



Proposition 2: Paths with repeated groups are never selected or advertised.

Routing Similarity Prunes Redundant Search

Alibaba Cloud
Worldwide Cloud Services Partner

Proposition 1: ≤ 1 path per group path.

Proposition 2: No repeated-group paths

Worst-Case Path Space

O(N!)

Router-Level Combinations

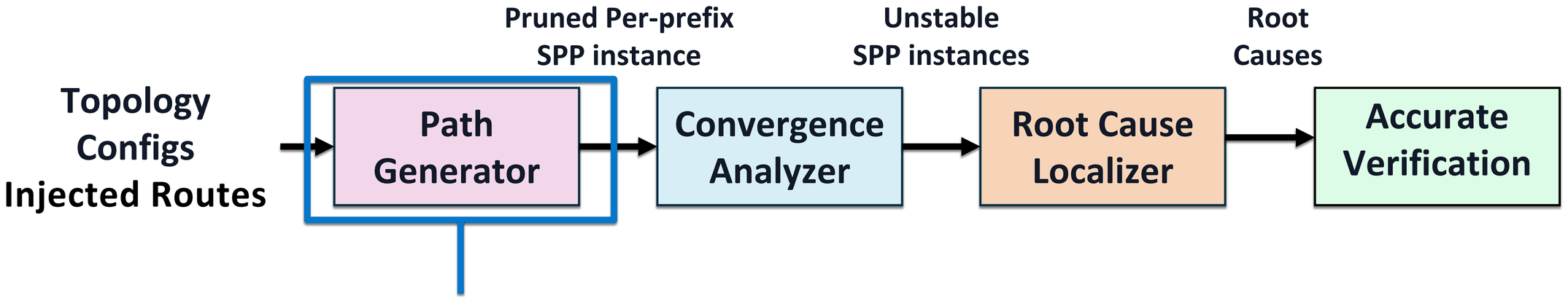
N: routers, **O(10^3)**

O(G!)

Group-Level Combinations

G: device groups, **O(10^2)**

TianYan: Scalable Verification via Path Pruning



Prune paths using routing similarity

Step 1: Generate Per-Prefix SPP Instances

Send the best path

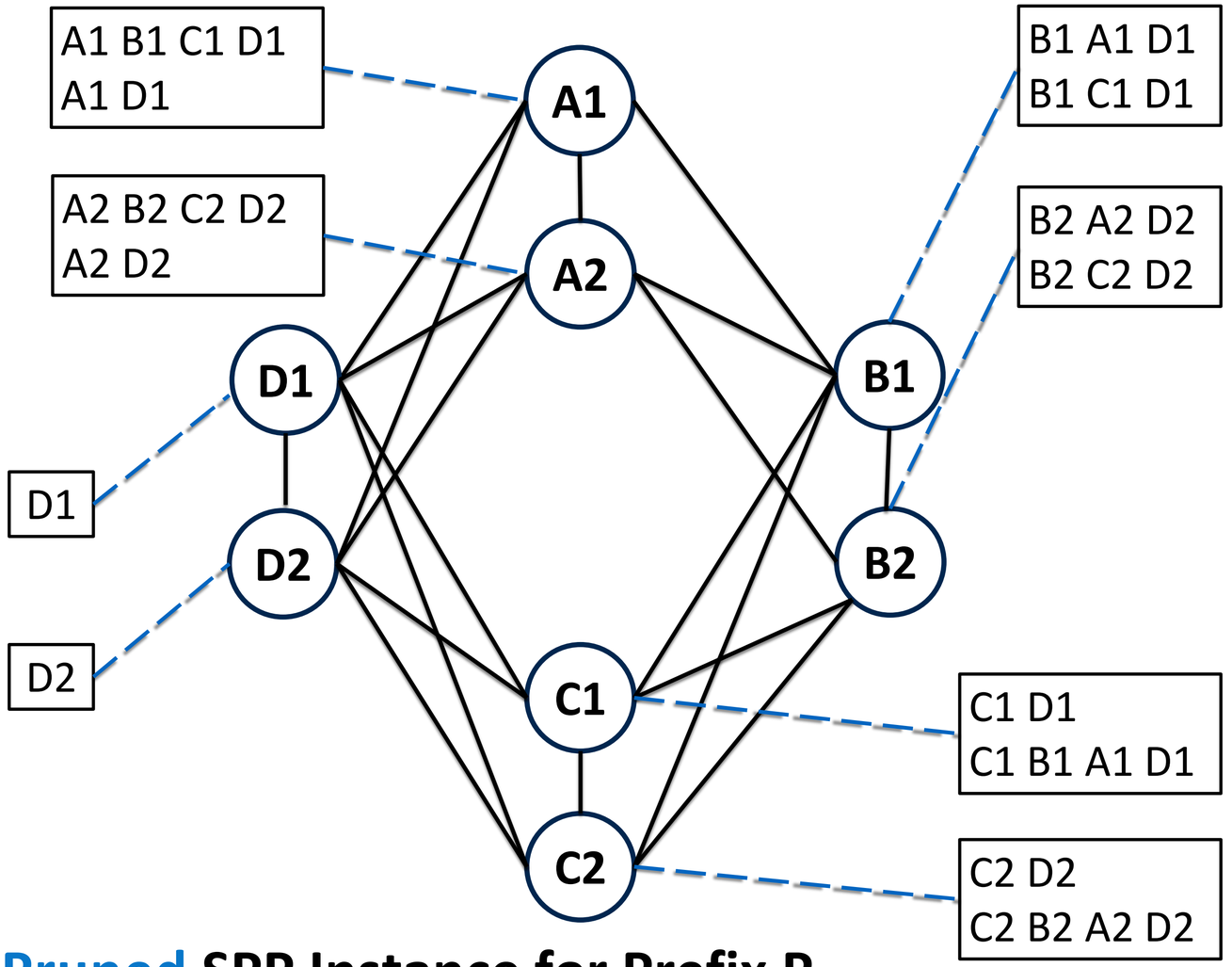
↓

Send all received paths

Modified
Route Simulator

Prune with Propositions 1 & 2

116 paths → 14 paths



Pruned SPP Instance for Prefix P

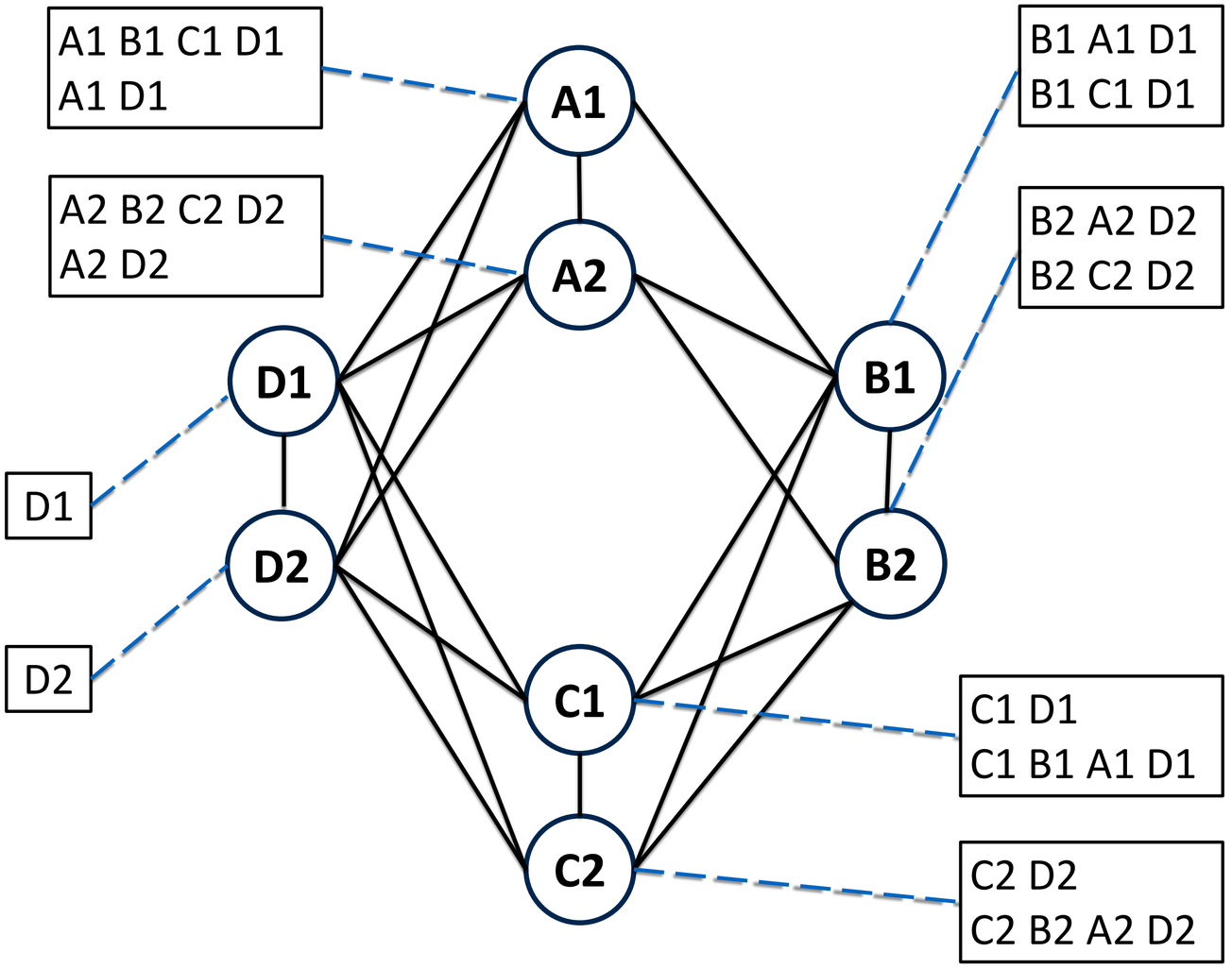
SPP instance = topology + permitted paths + rankings

Step 2.1: Greedy⁺ Quickly Resolves Most Prefixes

Greedy⁺: Find the stable node set

All nodes are stable → deterministic

Polynomial-time and sound



Stable node: its route choice is fixed, independent of advertisement order.

L. Cittadini et al., "From Theory to Practice: Efficiently Checking BGP Configurations for Guaranteed Convergence," IEEE TNSM, 2011.

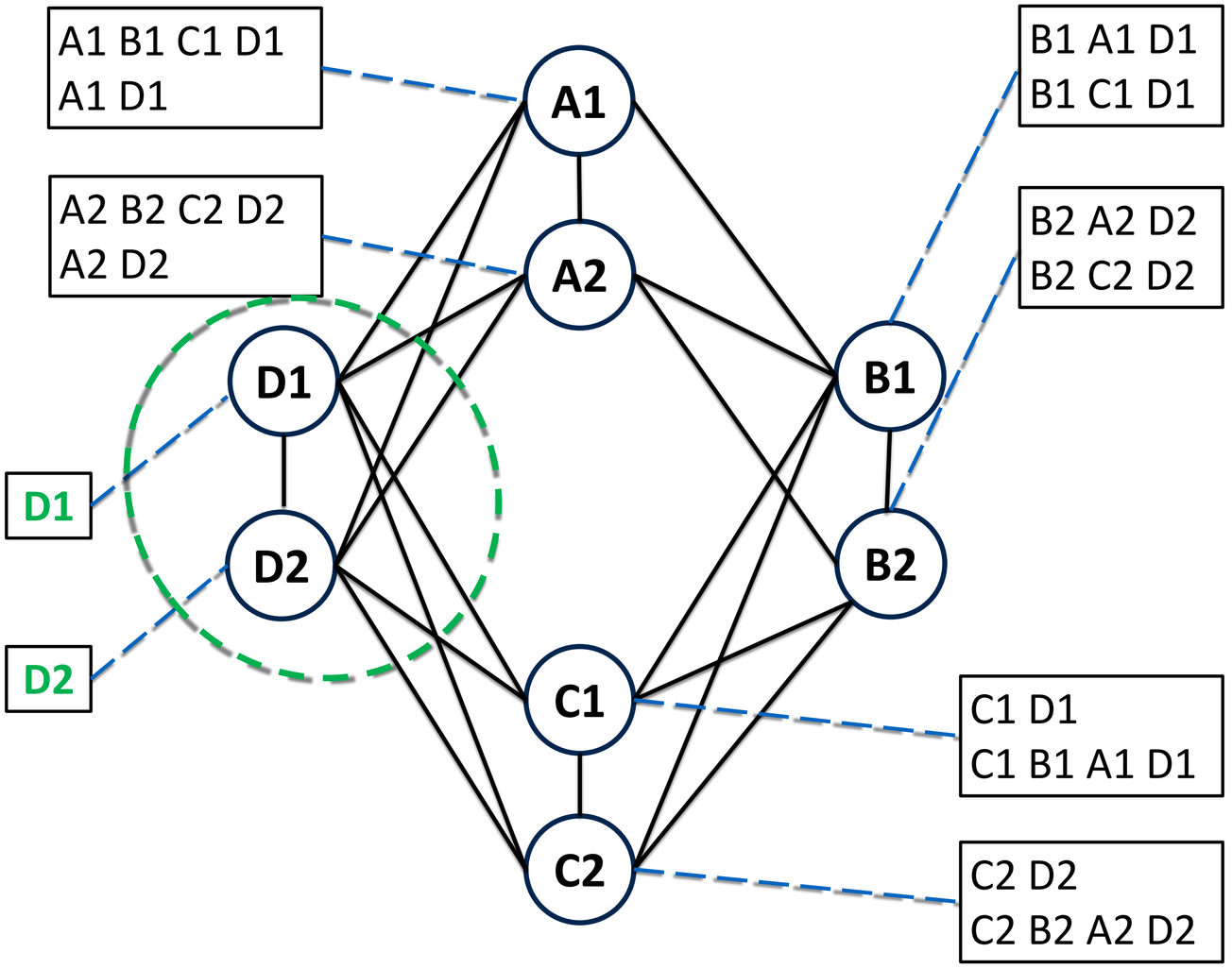
Step 2.1: Greedy+ Quickly Resolves Most Prefixes

Greedy+: Find the stable node set

All nodes are stable → deterministic

Polynomial-time and sound

1. Origin nodes are stable (D1/D2)



Stable node: its route choice is fixed, independent of advertisement order.

L. Cittadini et al., "From Theory to Practice: Efficiently Checking BGP Configurations for Guaranteed Convergence," IEEE TNSM, 2011.

Step 2.1: Greedy⁺ Quickly Resolves Most Prefixes

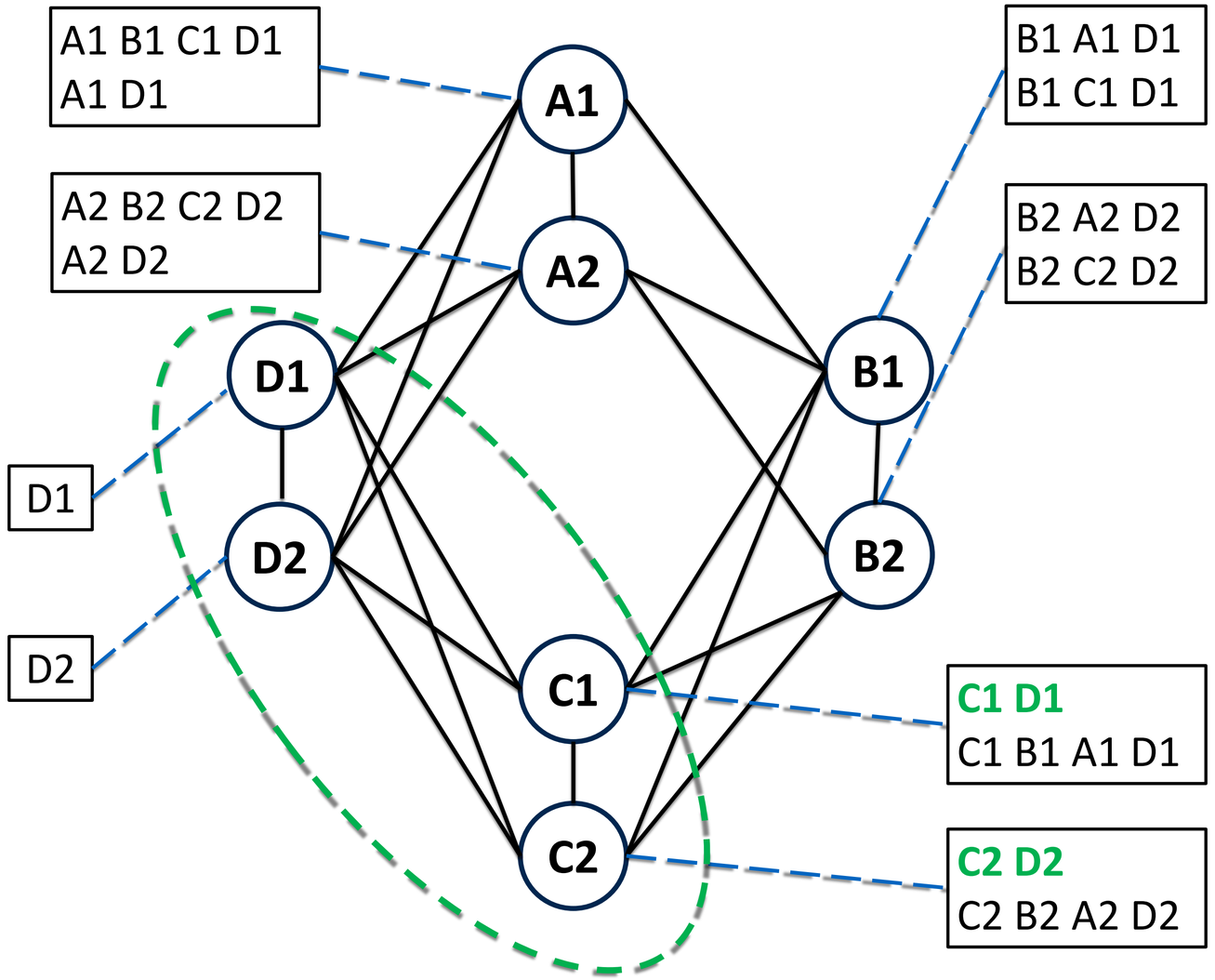
Greedy⁺: Find the stable node set

All nodes are stable → deterministic

Polynomial-time and sound

1. Origin nodes are stable (D1/D2)
2. C1/C2 prefer paths from stable nodes, thus stable
3. Stop ⇒ Potentially non-deterministic

Stable node: its route choice is fixed, independent of advertisement order.



L. Cittadini et al., "From Theory to Practice: Efficiently Checking BGP Configurations for Guaranteed Convergence," IEEE TNSM, 2011.

Step 2.2: SMT Detects Multiple Stable Outcomes

Encode stable-path constraints

1. Path Availability

Origin path, or its suffix is selected by the next hop.

2. Best-Path

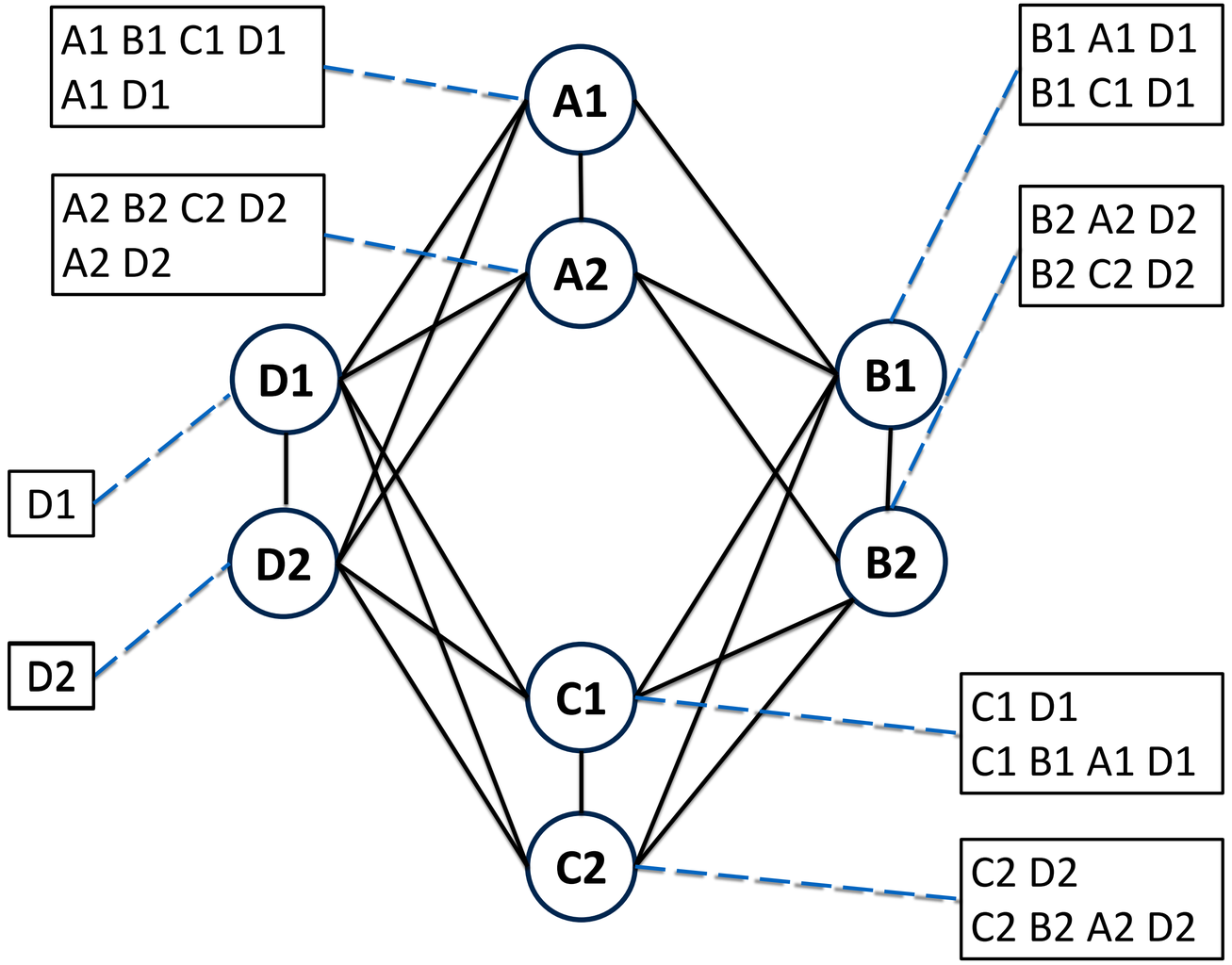
Select the highest-ranked available path.

Stable-path constraints

SMT Solver

Two distinct assignments

Non-deterministic convergence



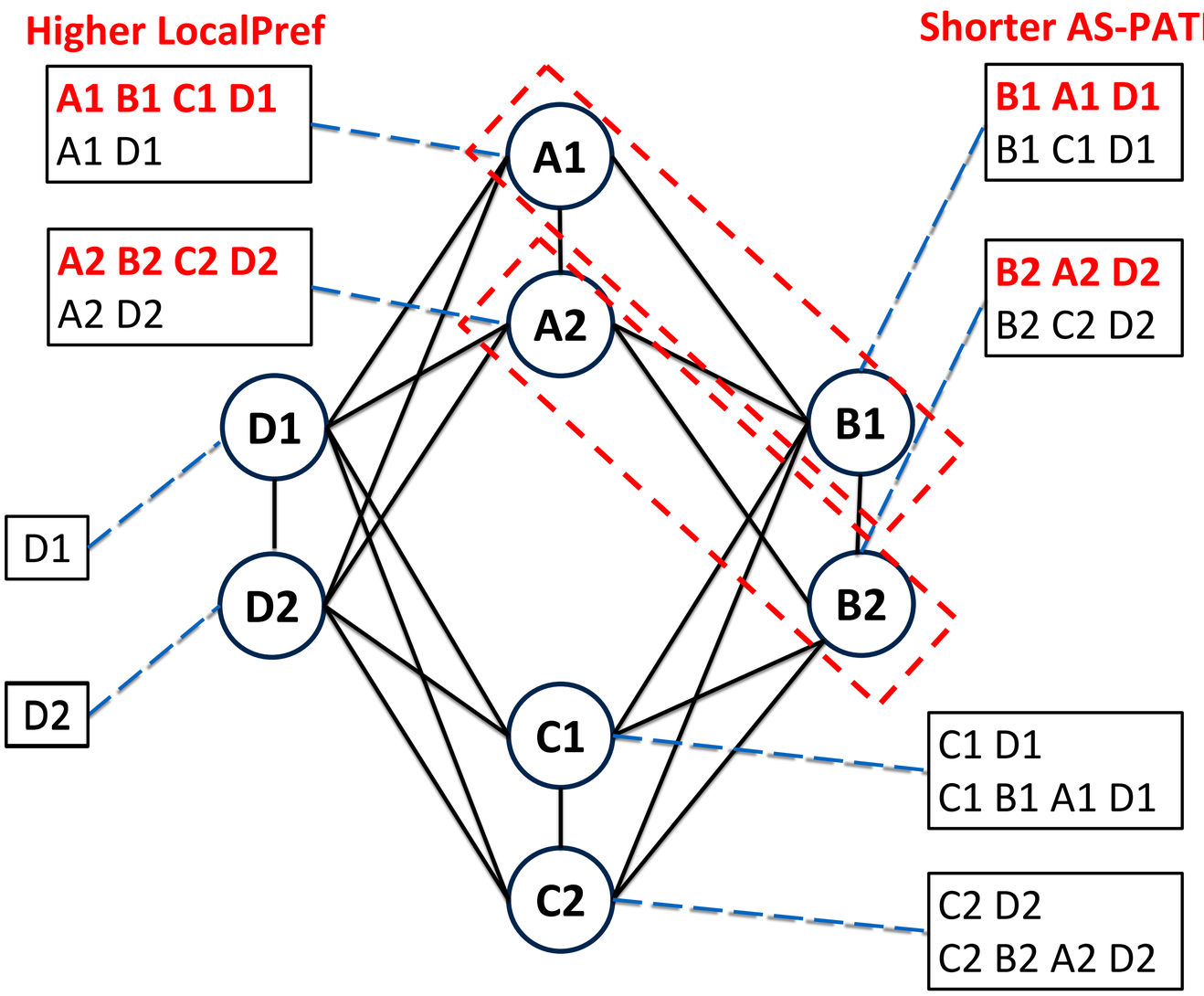
Step 3: Root Cause Localization

1. Find path preference loop

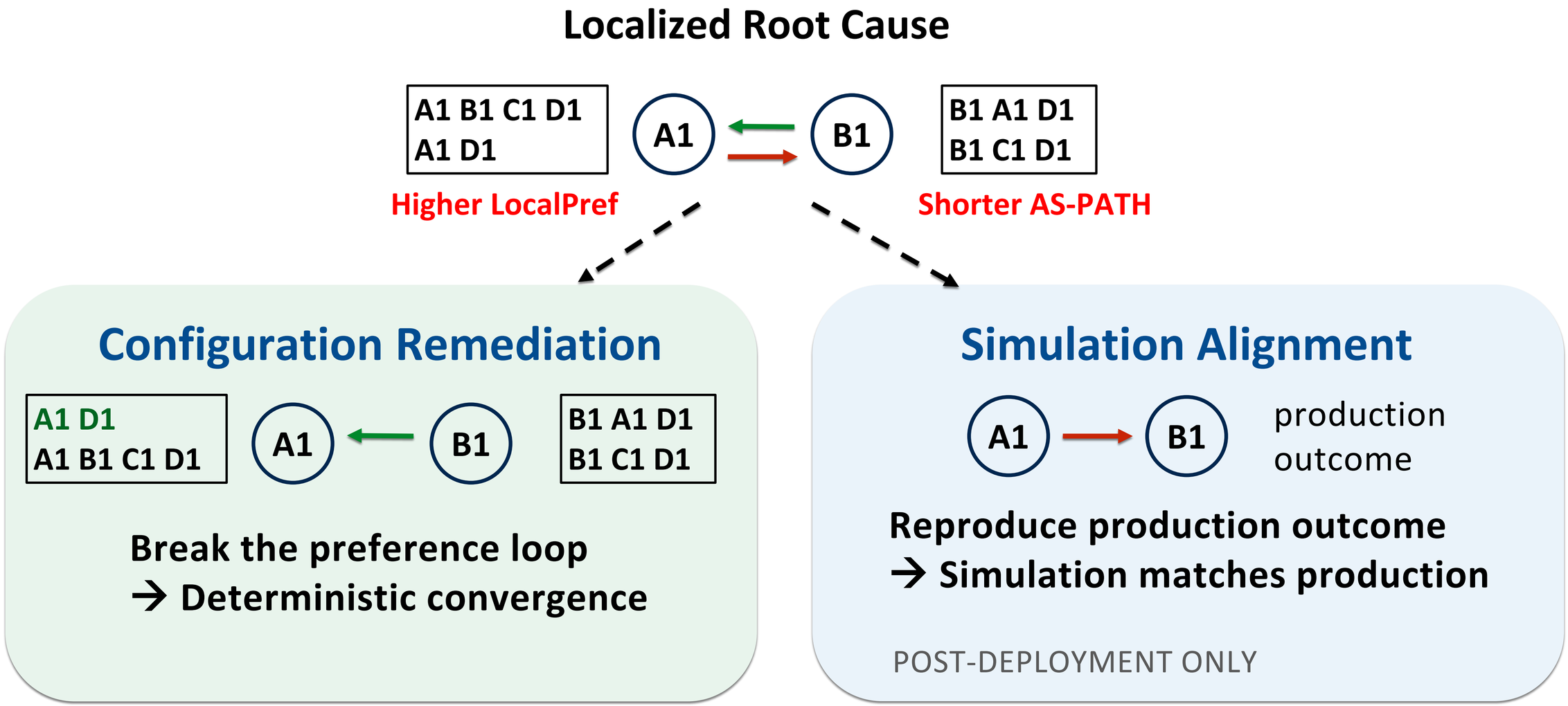
Trace preferred paths among unresolved nodes to identify cyclic dependencies.

2. Identify Tie-Breaking Attributes

Find the BGP attribute that makes each preferred path win.

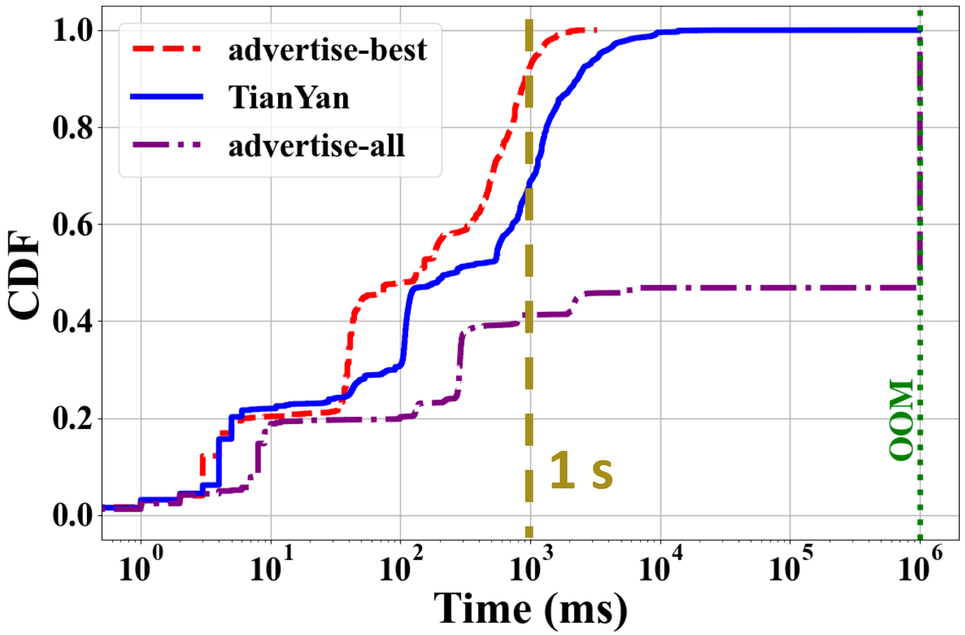


Step 4: Improve Verification Accuracy



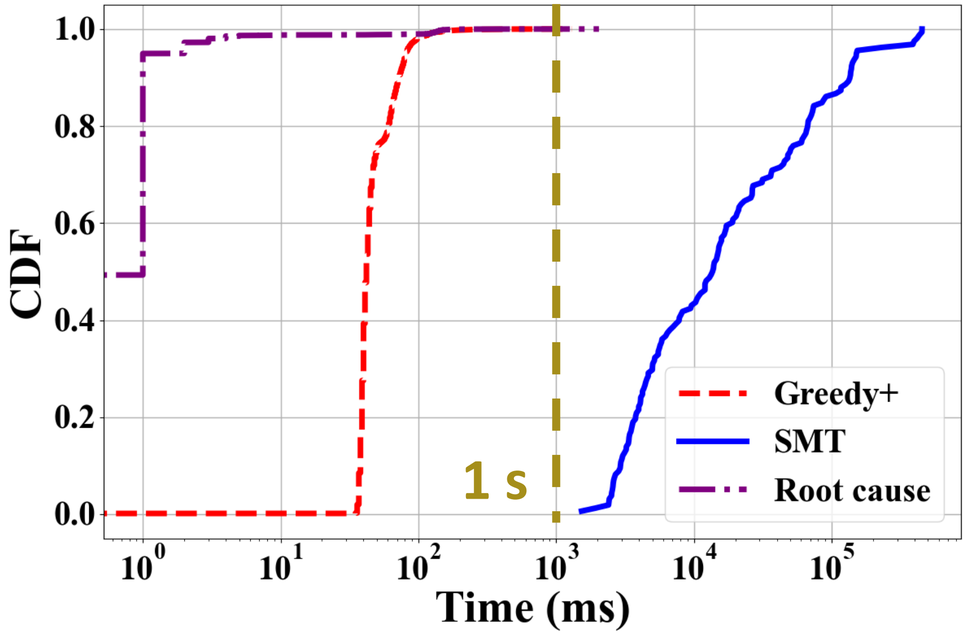
Evaluation on Alibaba's Production WAN

$O(10^3)$ routers · $O(10^6)$ prefixes · Single server



Path generation

Path pruning avoids explosion



Convergence analysis

SMT only for suspicious cases

Most prefixes can be analyzed within 1 second

Evaluation on Alibaba's Production WAN

Alibaba Cloud
Worldwide Cloud Services Partner

TOTAL PREFIXES

2×10^6

production prefixes evaluated

NON-DETERMINISTIC PREFIXES

4×10^4 (~2%)

prefixes with multiple outcomes

ROOT-CAUSE TYPES

1×10^2

root-cause types identified

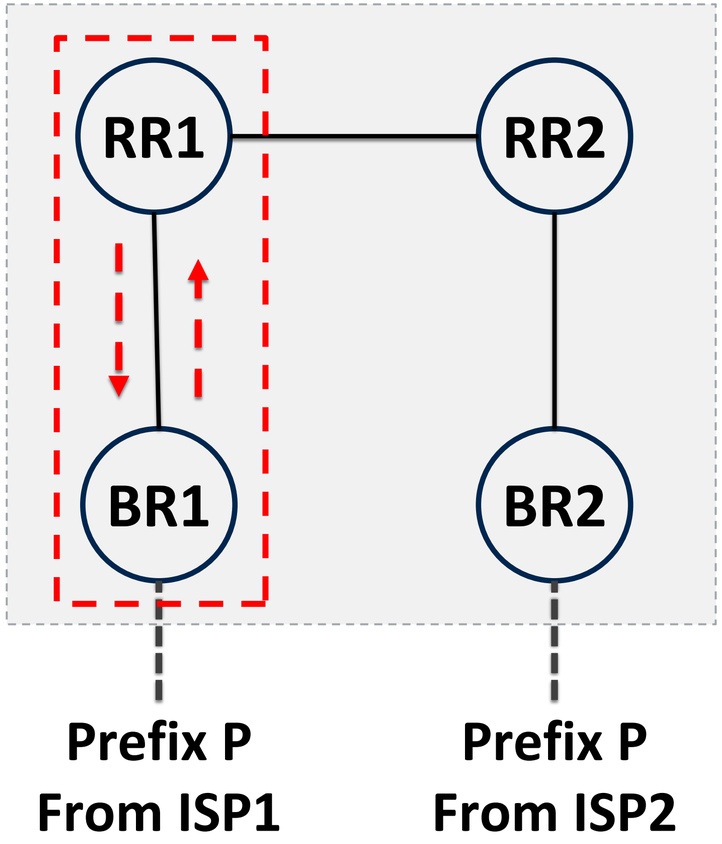
Without alignment, ~30K prefixes differ from production.

Production Case: Non-Deterministic Egress

Alibaba Cloud
Worldwide Cloud Services Partner

RR1 prefers BR1:
Higher Weight

BR1 prefers RR1:
Higher LocalPref



Configuration Flaws

1. RR2 omitted from RR1's weight policy

RR2's routes are unconditionally de-preferred.

2. BR2's regional LocalPref affects BR1's egress

Policy unintentionally affects BR1.

Two Converged Egress Outcomes

BR1 → ISP1 or BR1 → ISP2

ISP2: unintended, but observed in production.

Configuration flaws lead to non-deterministic egress.

BR: Border Router · RR: Router Reflector

Lessons from Production

Alibaba Cloud
Worldwide Cloud Services Partner

01 REALITY

Non-determinism is widespread.

~2% affected
~30K simulations differ from production

02 BLIND SPOTS

Capable verifiers still miss violations.

Missing intents + coarse checks
→ Verify stability directly

03 OPERATIONS

Late remediation is often impractical.

Network-wide fixes are often too risky;
operators prioritize business-critical prefixes.

04 PREVENTION

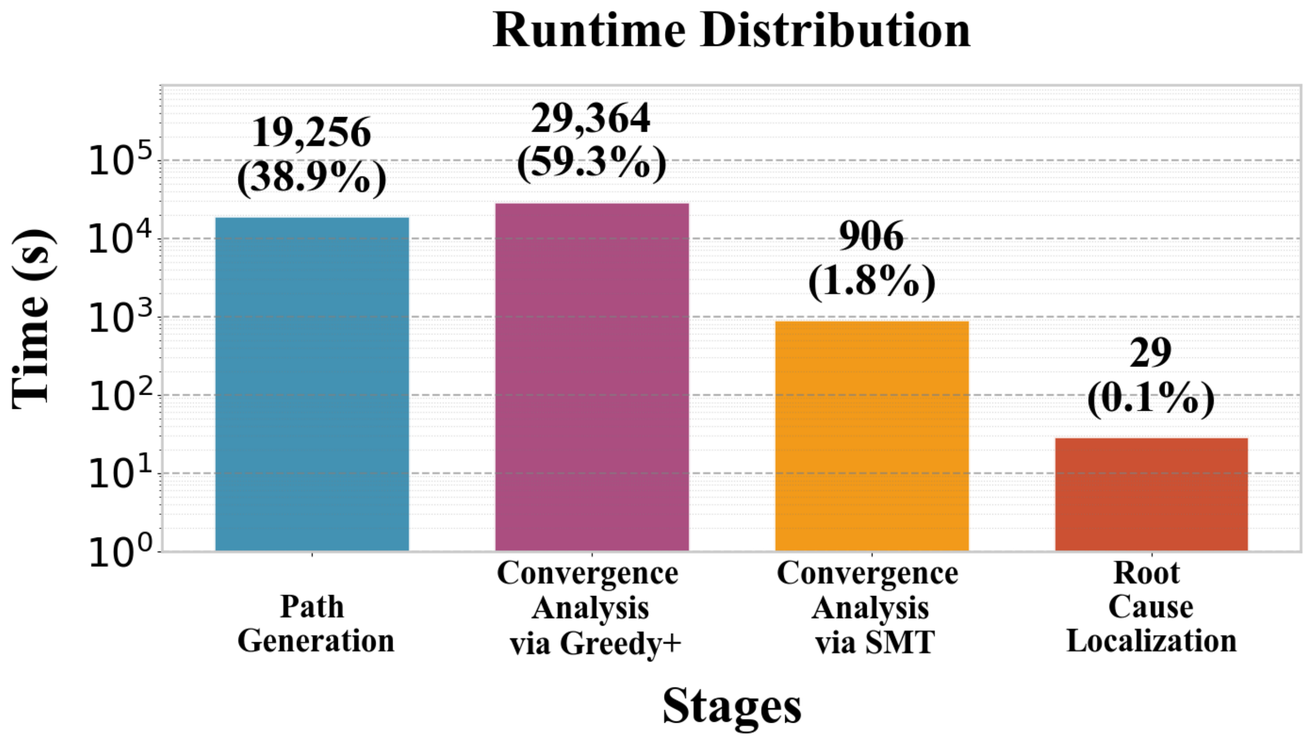
Preserve intent through network evolution.

Fine-grained policies · explicit intent database
Design–implementation consistency

Next: Incremental verification · Broader input coverage
From Converged-State Verification to Transient-State Verification

THANKS

Where Does TianYan Spend Its Runtime?



13.8 h total runtime

98.2% spent on path generation and Greedy+

Prefix-level parallelism enables straightforward distributed scaling

Limitations

- **Reliance on routing similarity**
- **Group-path explosion**
- **Time-dependent and inter-protocol behavior**
- **Route aggregation**
- **Arbitrary link failures**