Understanding and Mitigating Packet Corruption in Data Center Networks

Danyang Zhuo, Monia Ghobadi, Ratul Mahajan

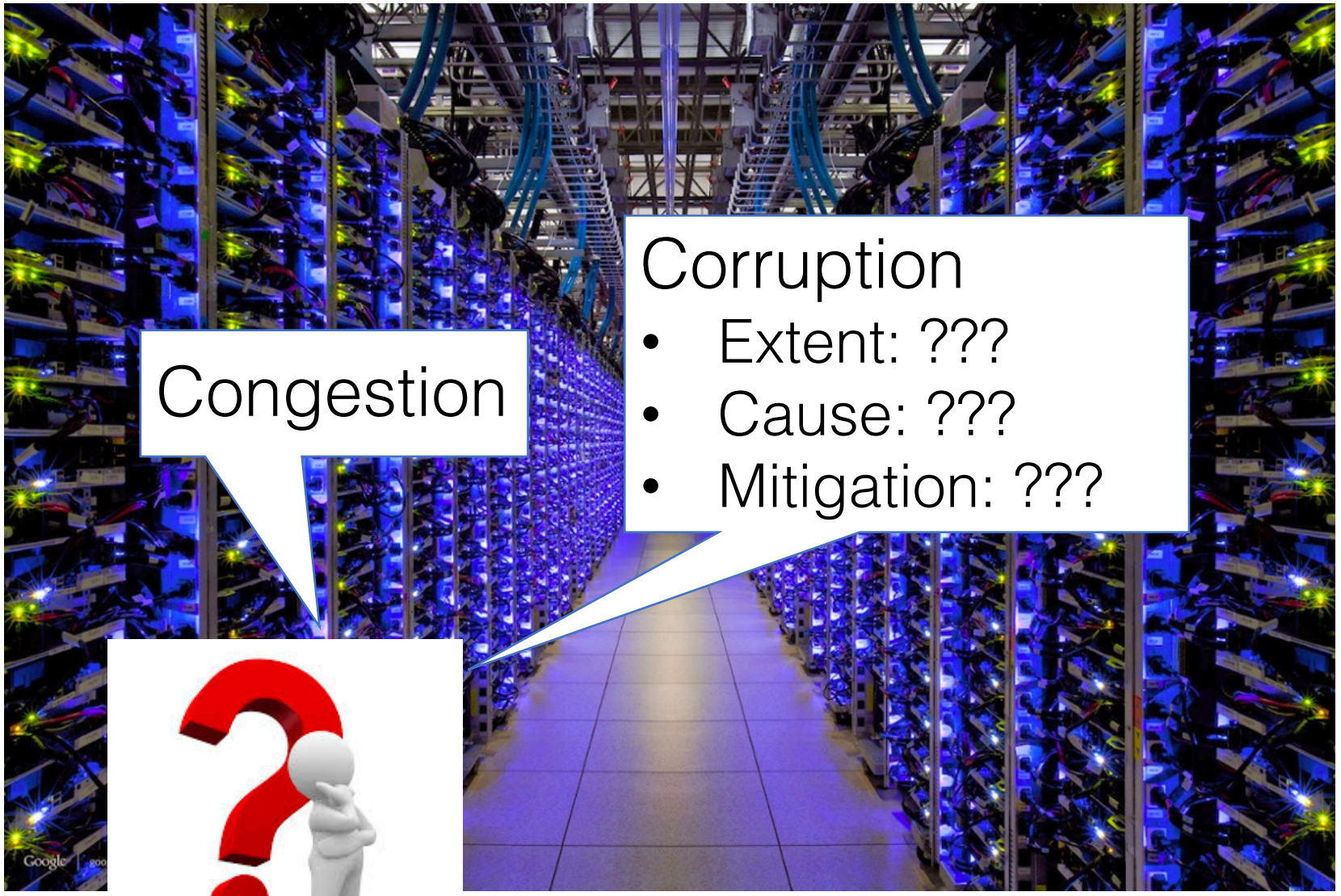
Klaus-Tycho Förster, Arvind Krishnamurthy, Thomas Anderson

UNIVERSITY of WASHINGTON

Microsoft Research

AD NYE VEJE AALBORG UNIVERSITET

INTENTIONET



Packet loss

↓

Hurt application

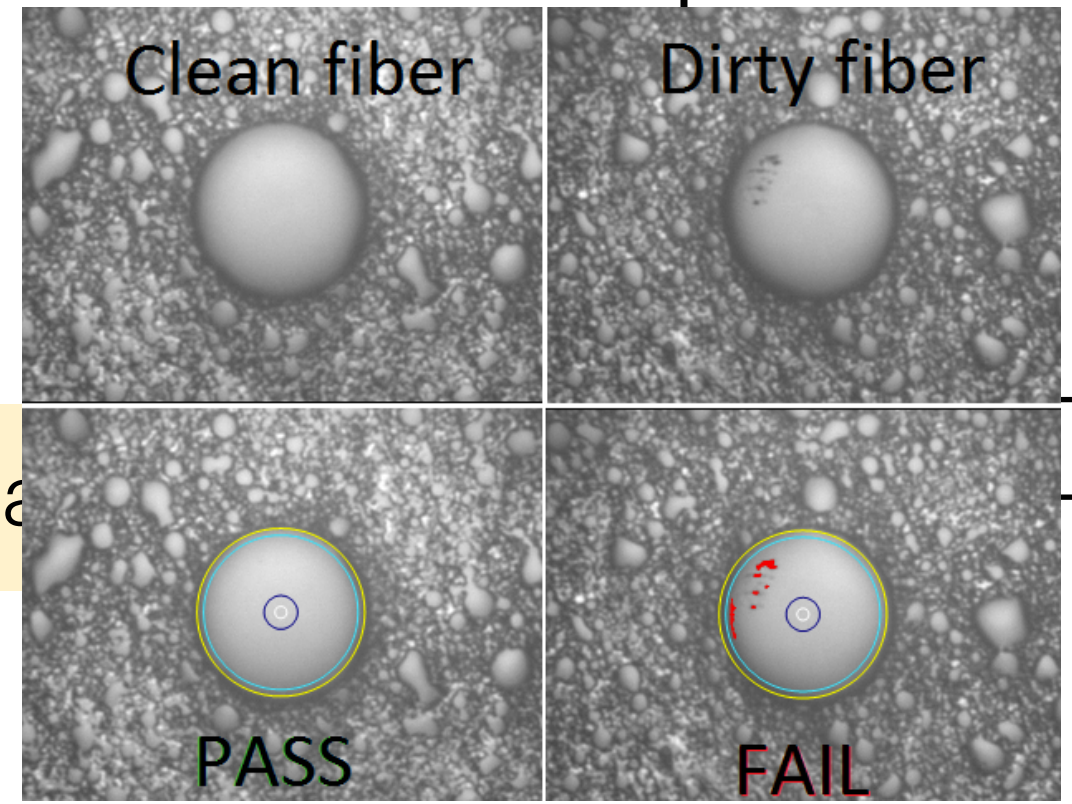
↓

Revenue loss

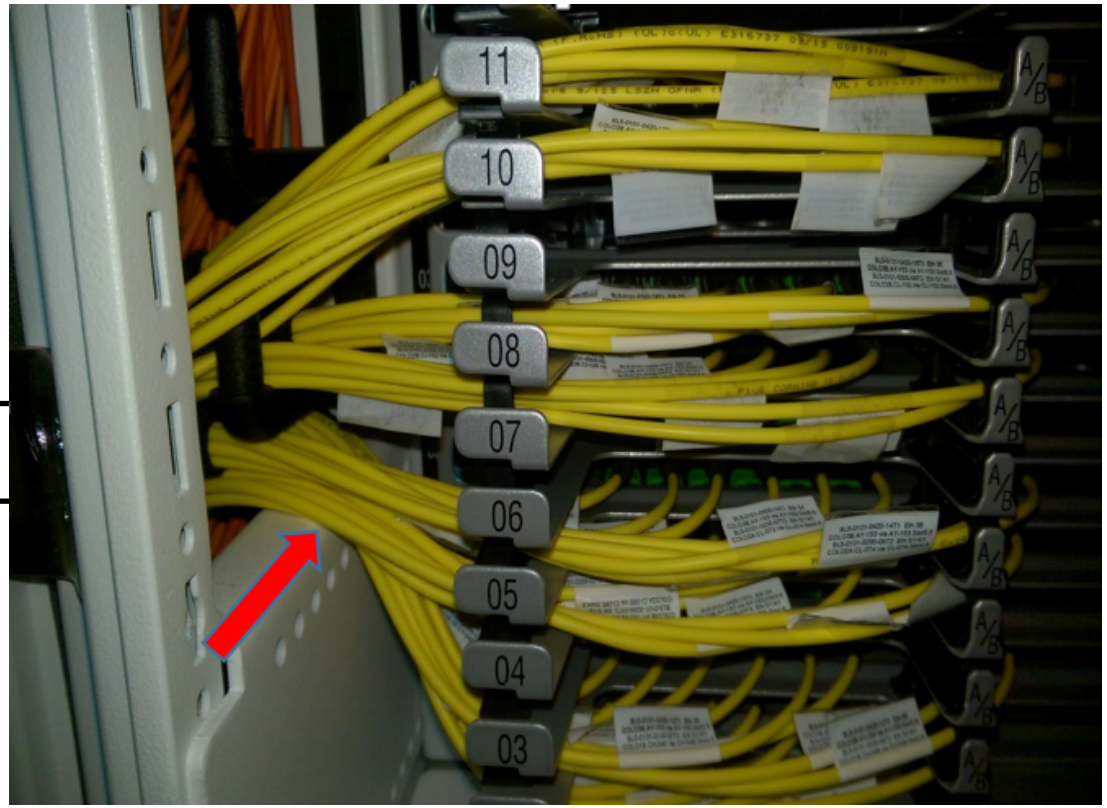
In this talk…

- Packet corruption is a significant source of packet loss
- Packet corruption has distinctive symptoms and root causes
- CorrOpt reduces corruption by 3 orders of magnitude
  - Parts of the system are deployed in Microsoft DCs

Packet Corruption

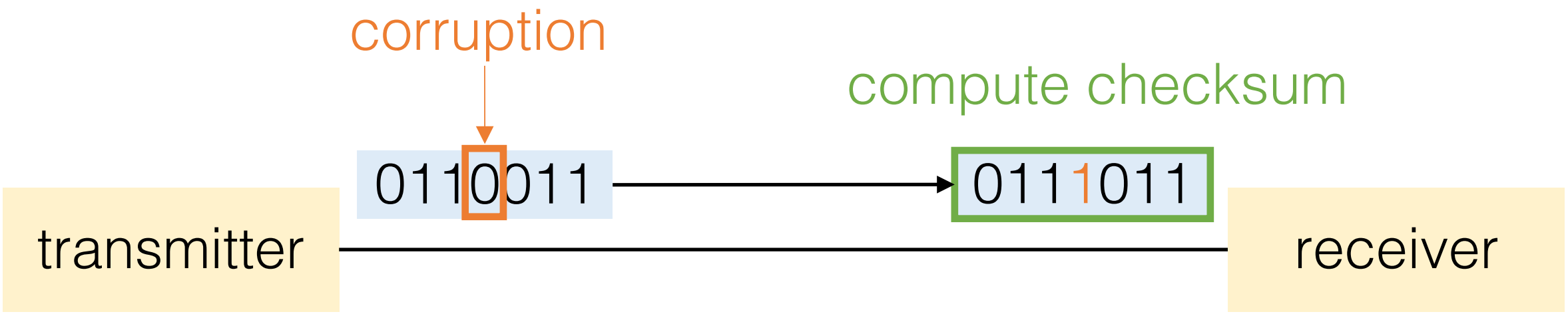


dirty optical connector

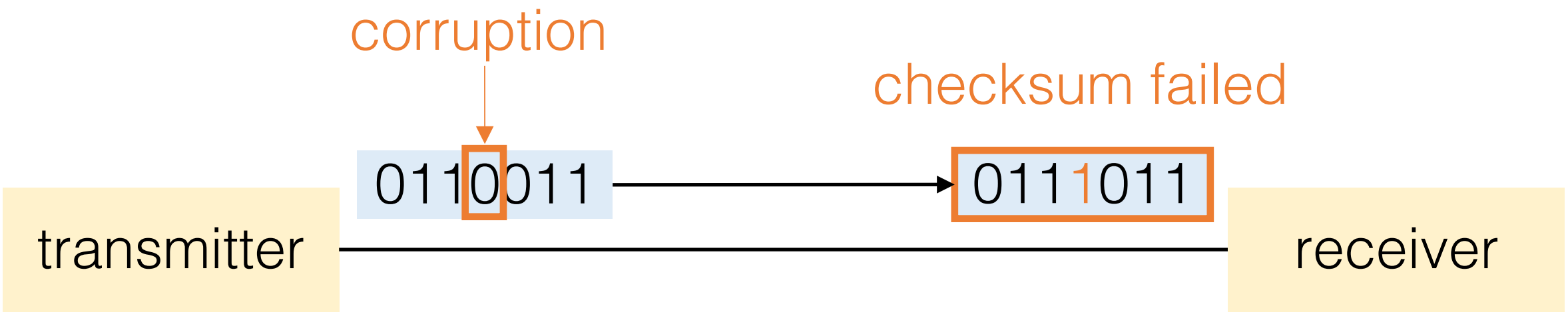


damaged fiber

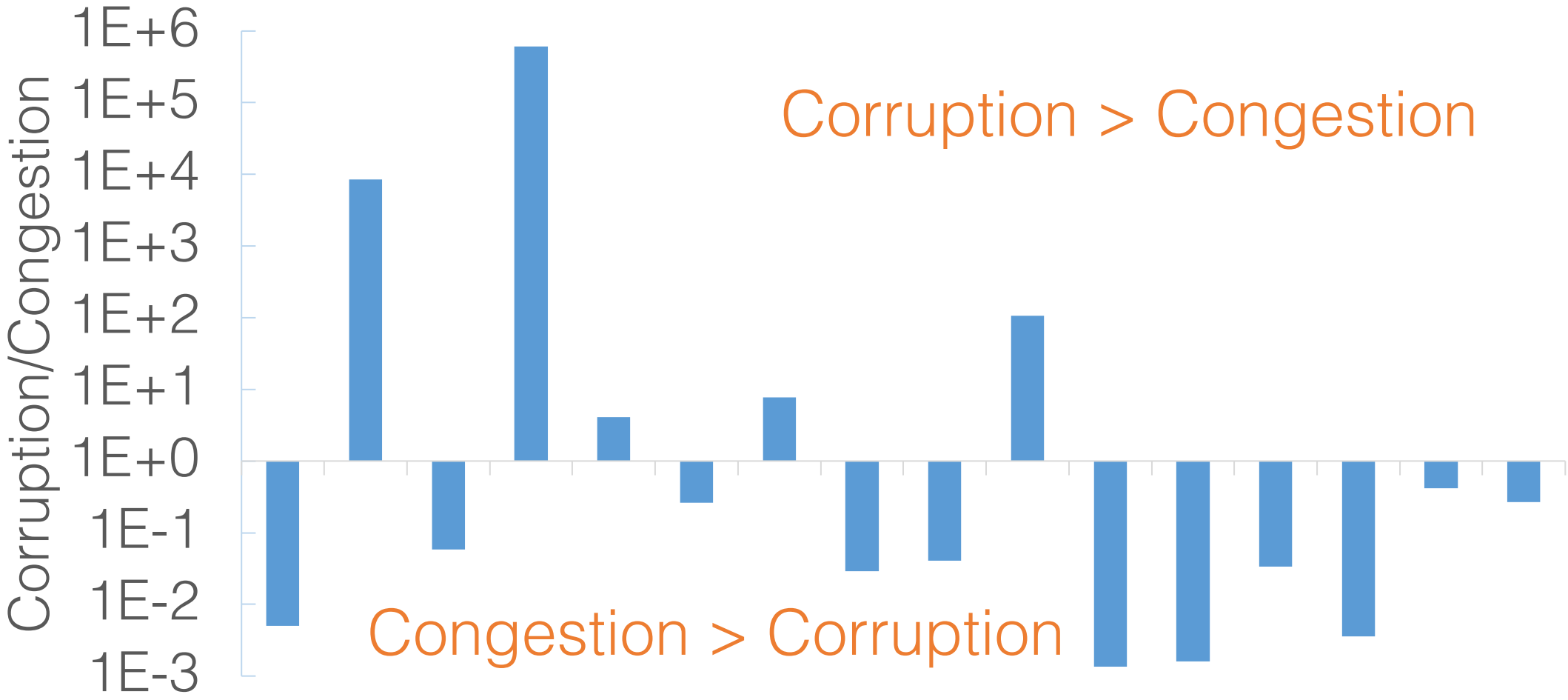
Packet Corruption



Packet Corruption



Packet Corruption is Significant



350K switch-to-switch links, 15 data centers

In this talk…

- Packet corruption is a significant source of packet loss
- Packet corruption has distinctive symptoms and root causes
- CorrOpt reduces corruption by 3 orders of magnitude
  - Parts of the system are deployed in Microsoft DCs

Questions

- How widespread is corruption?
- Is corruption loss rate stable over time?
- Where are corrupting links located?
- What causes packet corruption?

Number of Corrupting Links is Small

- Number of links with corruption is 2-4% of links with congestion

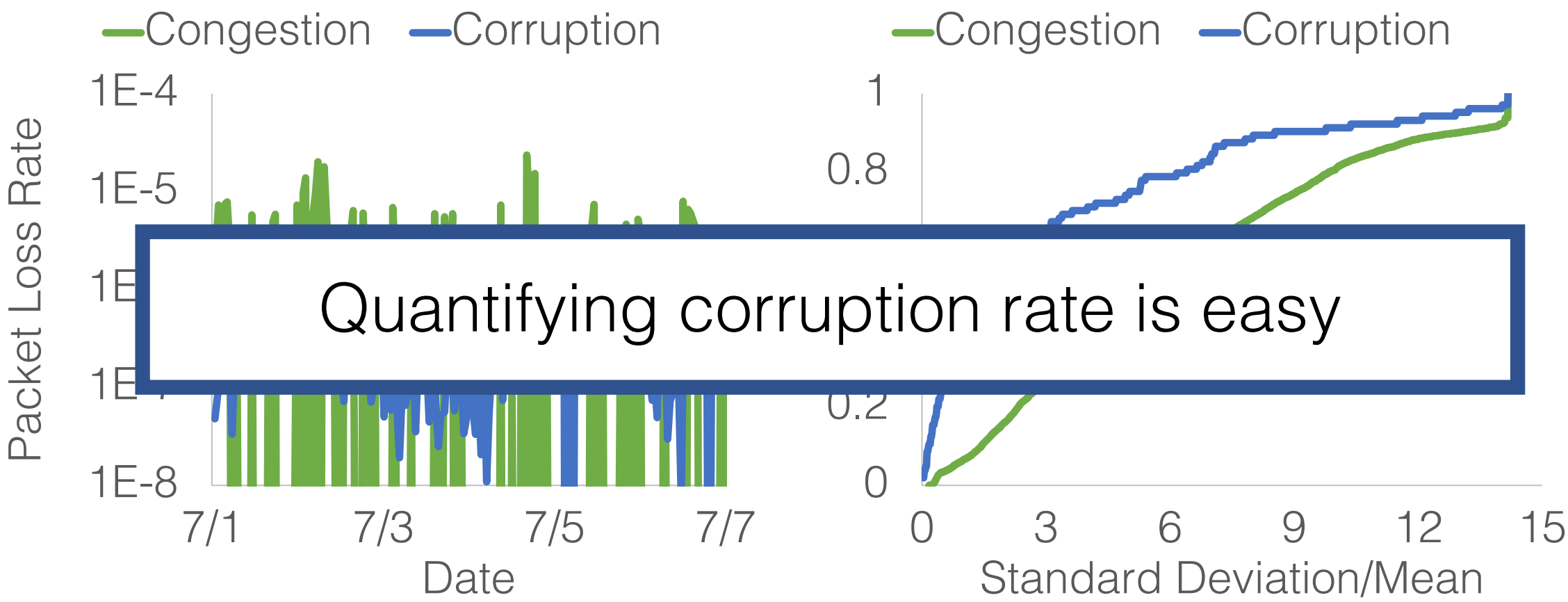
Why is the number of corruption losses still comparable to congestion?

Many Links have High Corruption Rate

Loss Rate Bucket	Links with Congestion	Links with Corruption
$[10^{-8}, 10^{-5})$	92.44%	47.23%
[		
[		
Greater than 10^{-3}	0.22%	12.67%
Total	100%	100%

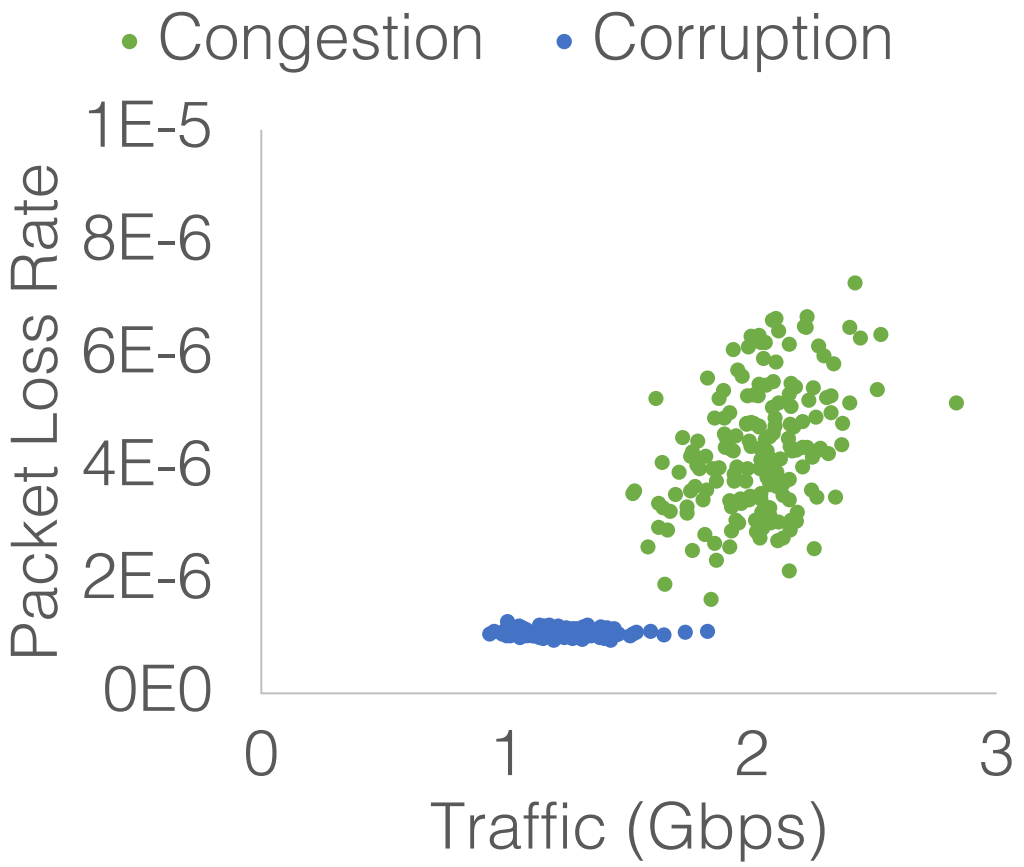
Need a system to mitigate corruption

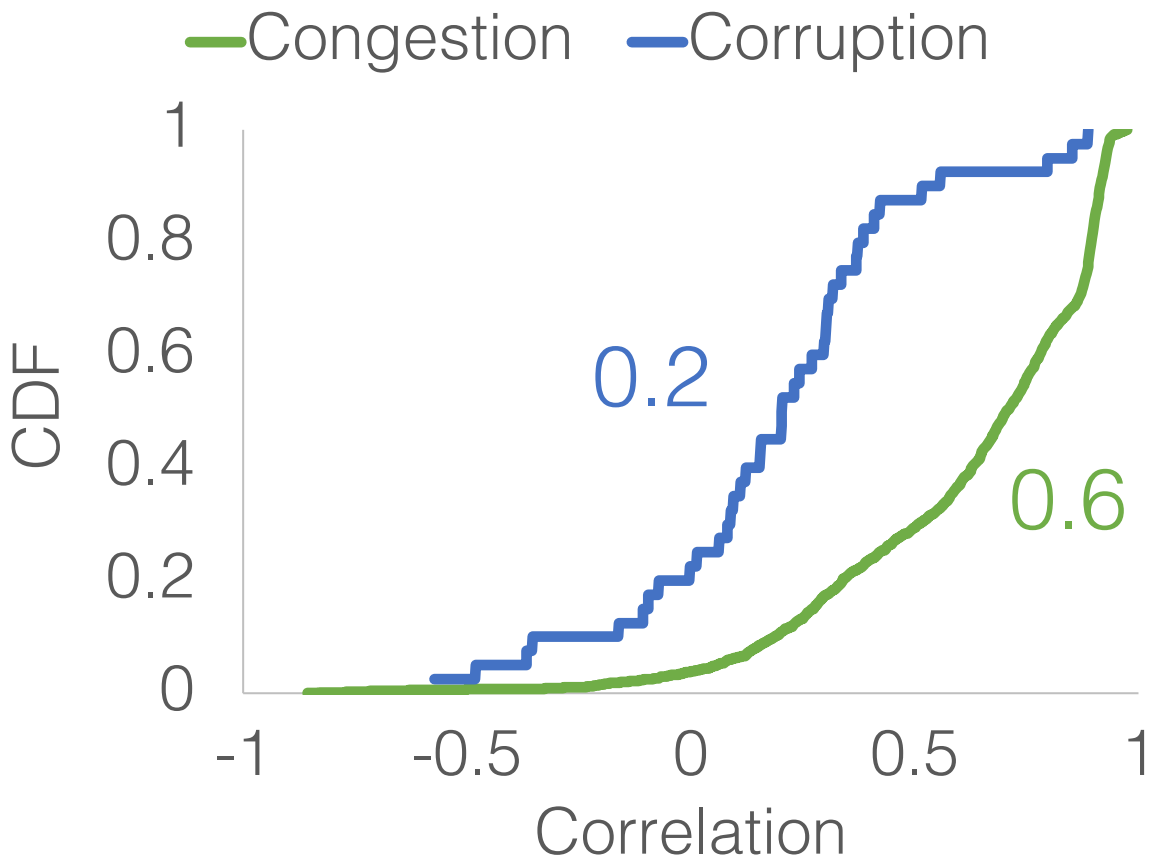
Corruption Rate is Stable



Quantifying corruption rate is easy

Why is Corruption Rate Stable?





Other Characteristics of Corruption

- Corruption tends to be scattered across all stages of data center network
- Corruption tends to affect only a single direction of transmission

Root Causes of Packet Corruption

Root Cause	Symptoms	Contribution
Dirty connector	High transmit and low receive optical power	17-57%
Damaged fiber/cable	High transmit and low receive optical power (bi-directional)	14-48%
Bad or loose transceiver	Good optical power, affect a single link	6-45%
Shared component failure	Co-located failures	10-26%
Decaying transmitter laser	Low transmit optical power	< 1%

In this talk…

- Packet corruption is a significant source of packet loss
- Packet corruption has distinctive symptoms and root causes
- CorrOpt reduces corruption by 3 orders of magnitude
  - Parts of the system are deployed in Microsoft DCs

System Framework

- Pinpoint corrupting links

Corruption rate is stable.

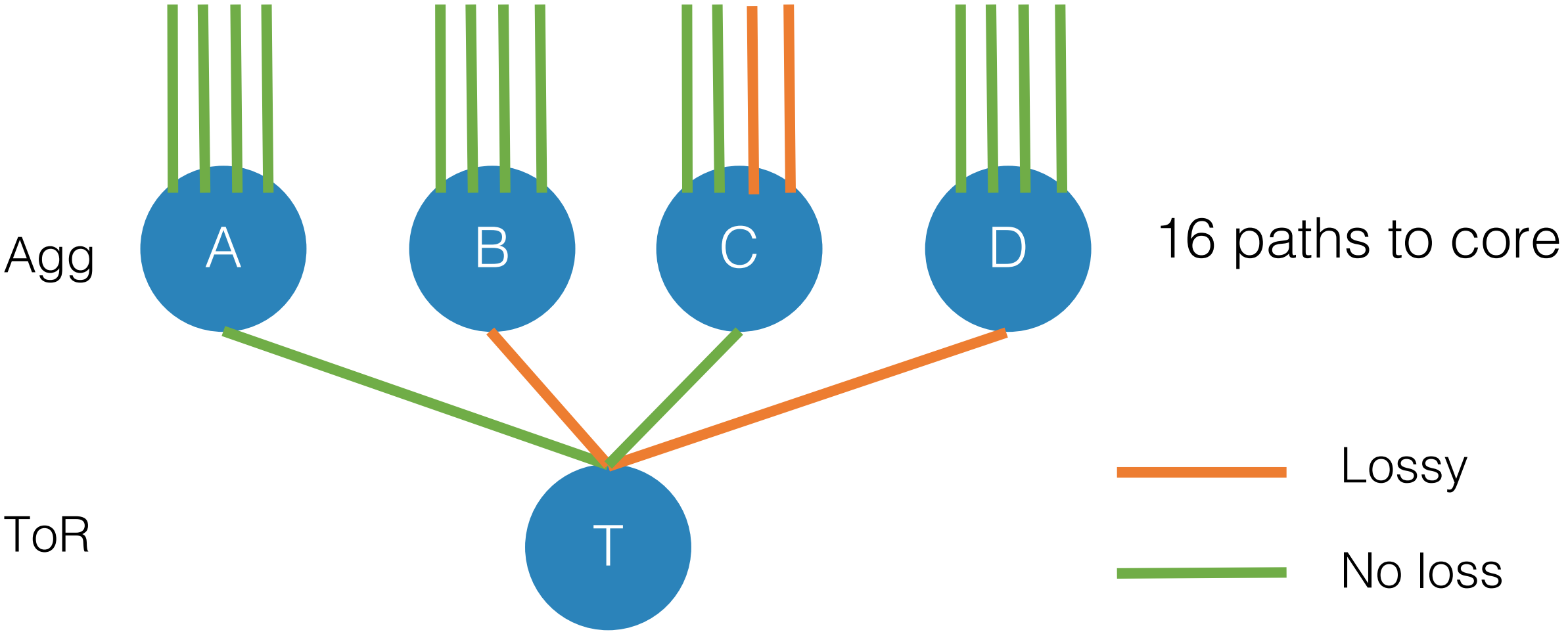
- Disable corrupting links while meeting capacity constraint
  - Every rack has a minimum fraction of paths to reach core

A small subset of links are corrupting.

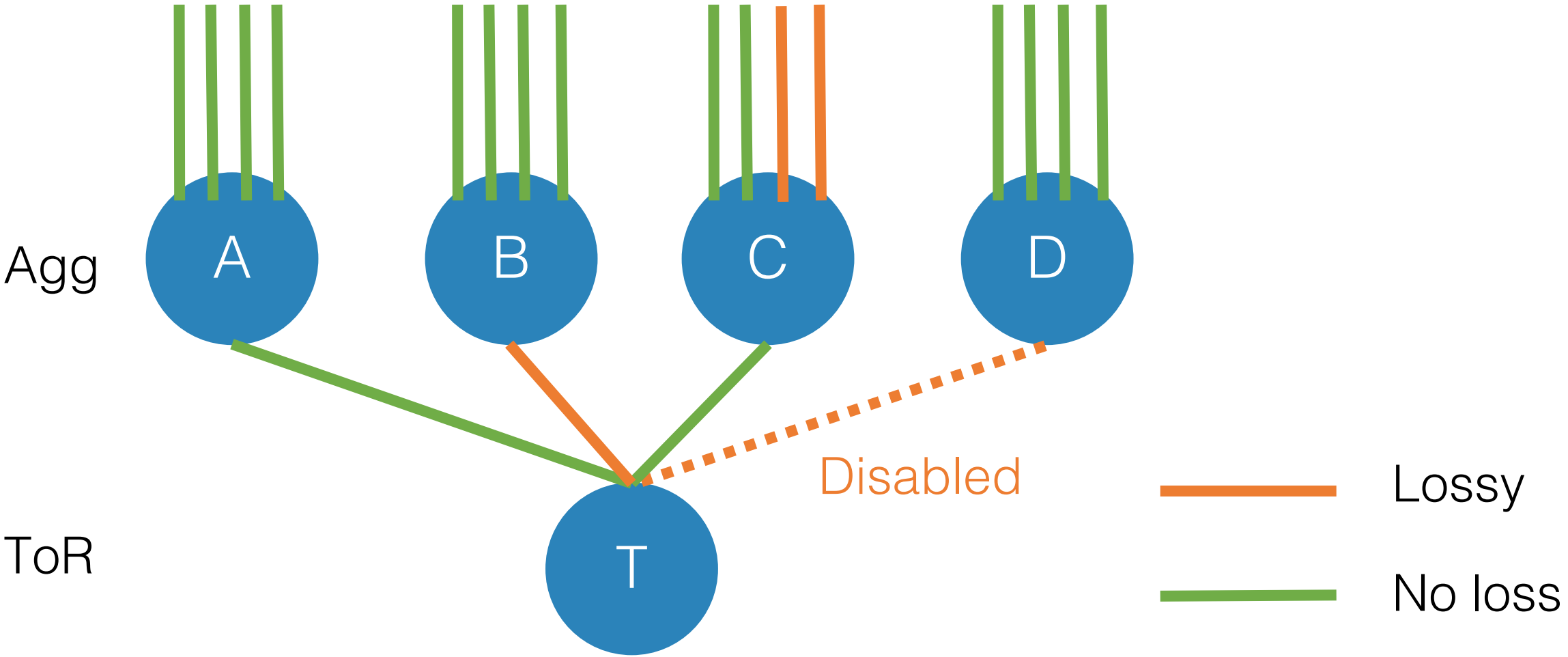
- Diagnose root cause and repair
  - Clean optical connector, replace fiber, replace transceiver, etc.

Multiple root causes. Repair depends on the root cause.

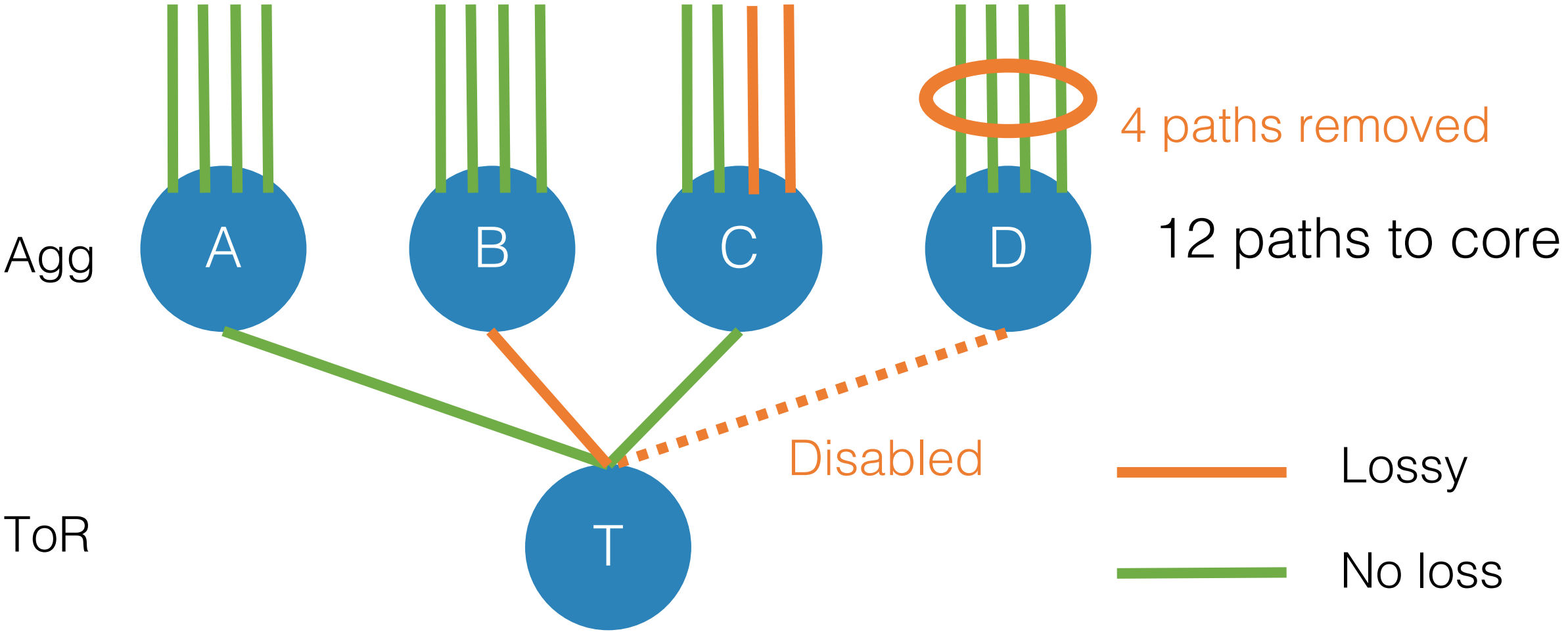
Challenge #1: Meet Capacity Constraint



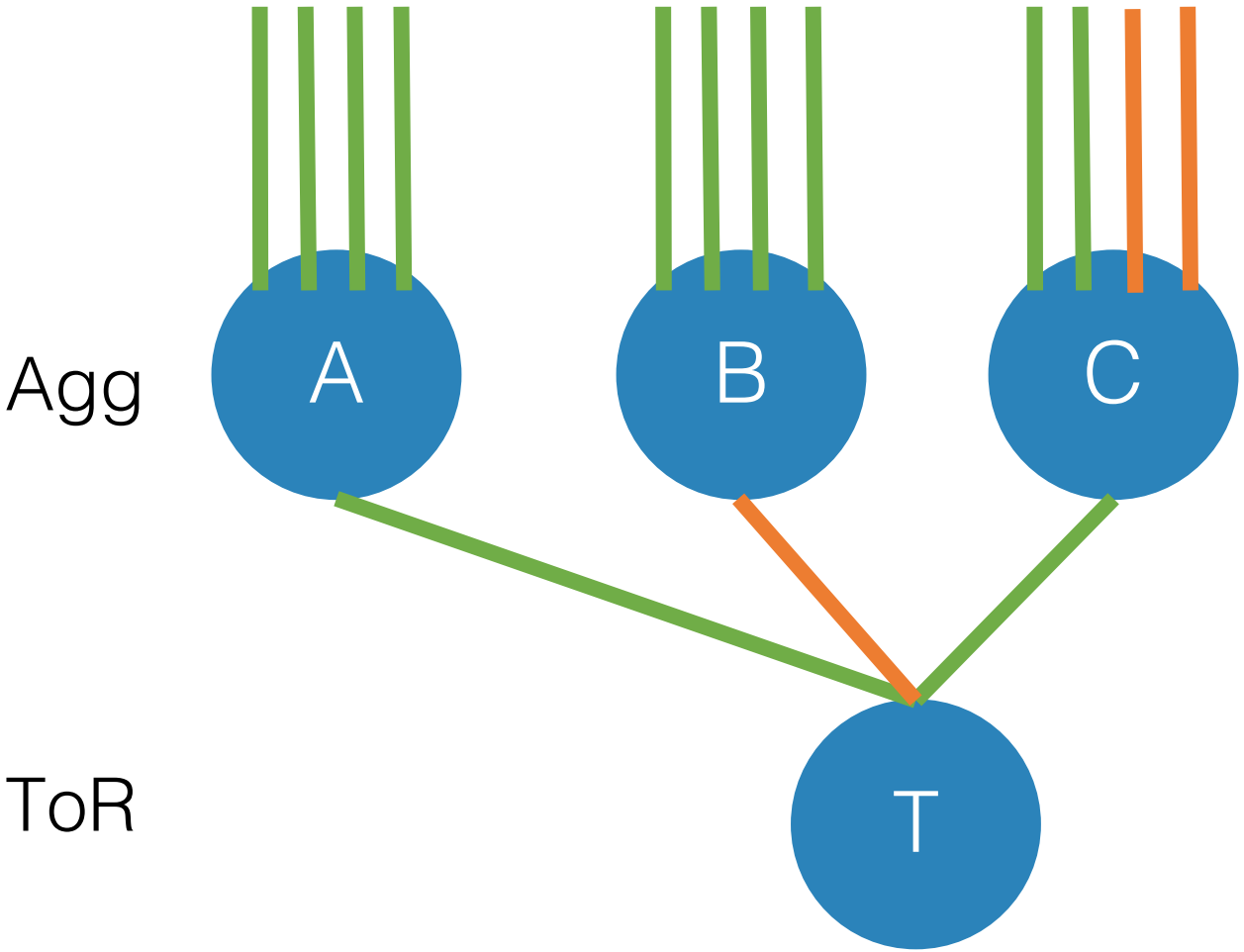
Challenge #1: Meet Capacity Constraint



Challenge #1: Meet Capacity Constraint



Challenge #1: Meet Capacity Constraint

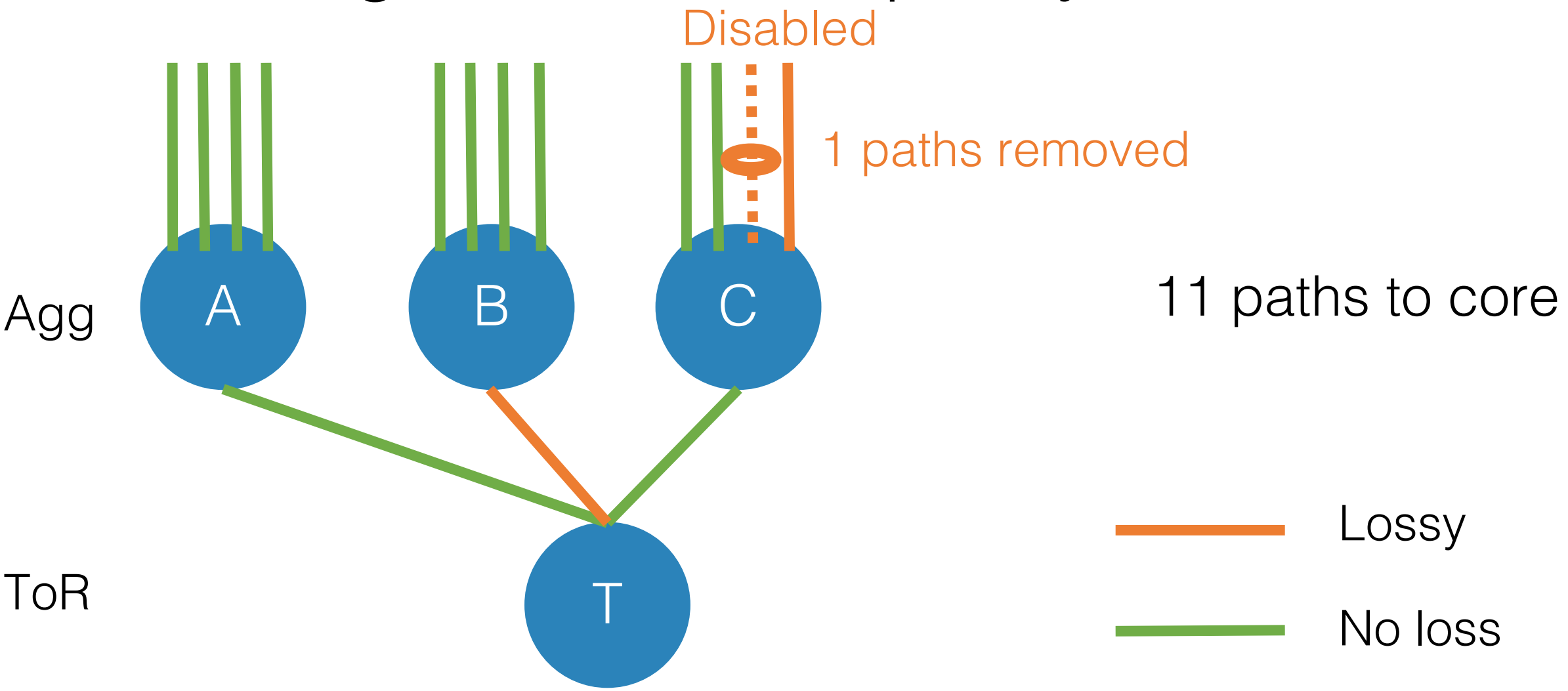


12 paths to core

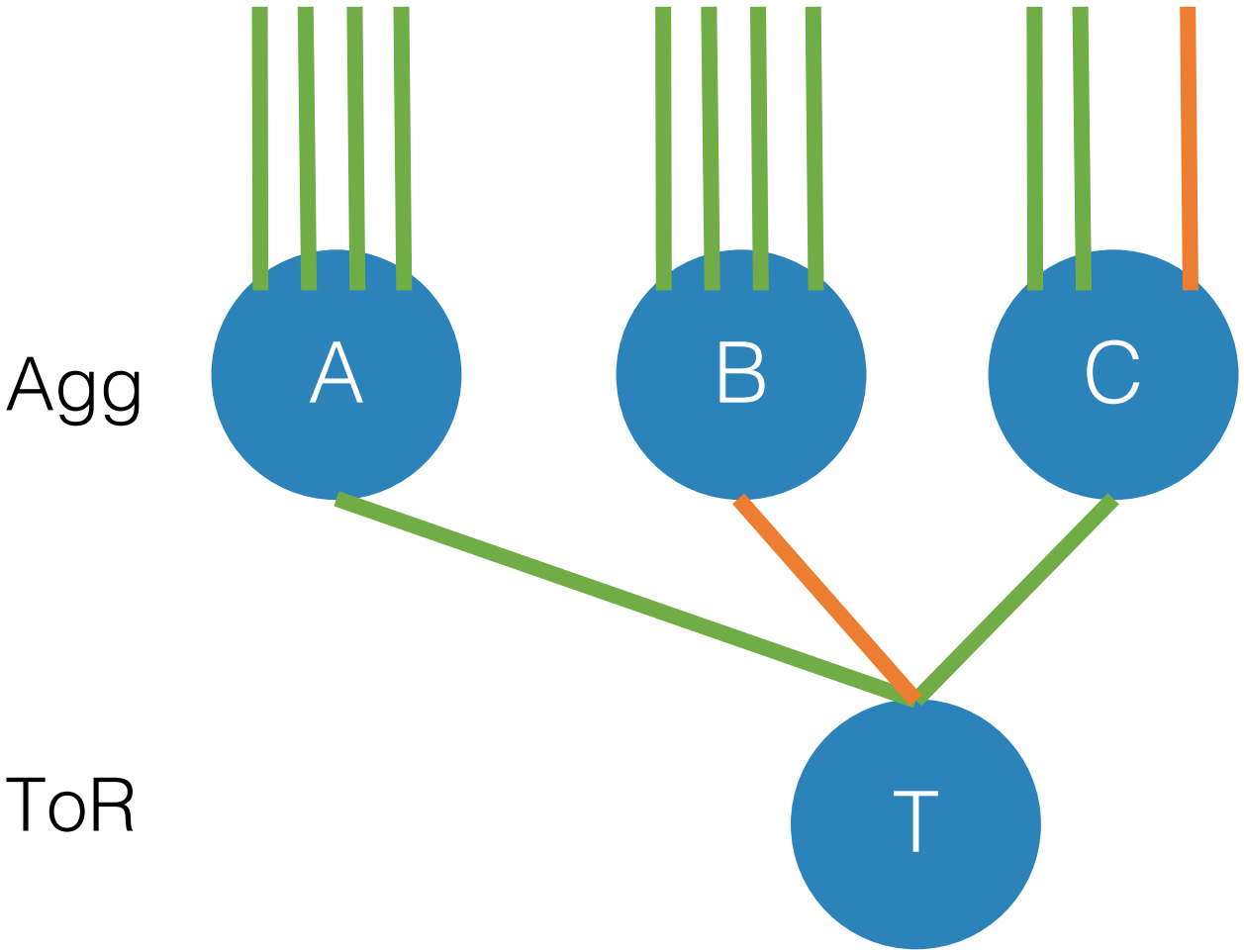
Lossy

No loss

Challenge #1: Meet Capacity Constraint



Challenge #1: Meet Capacity Constraint

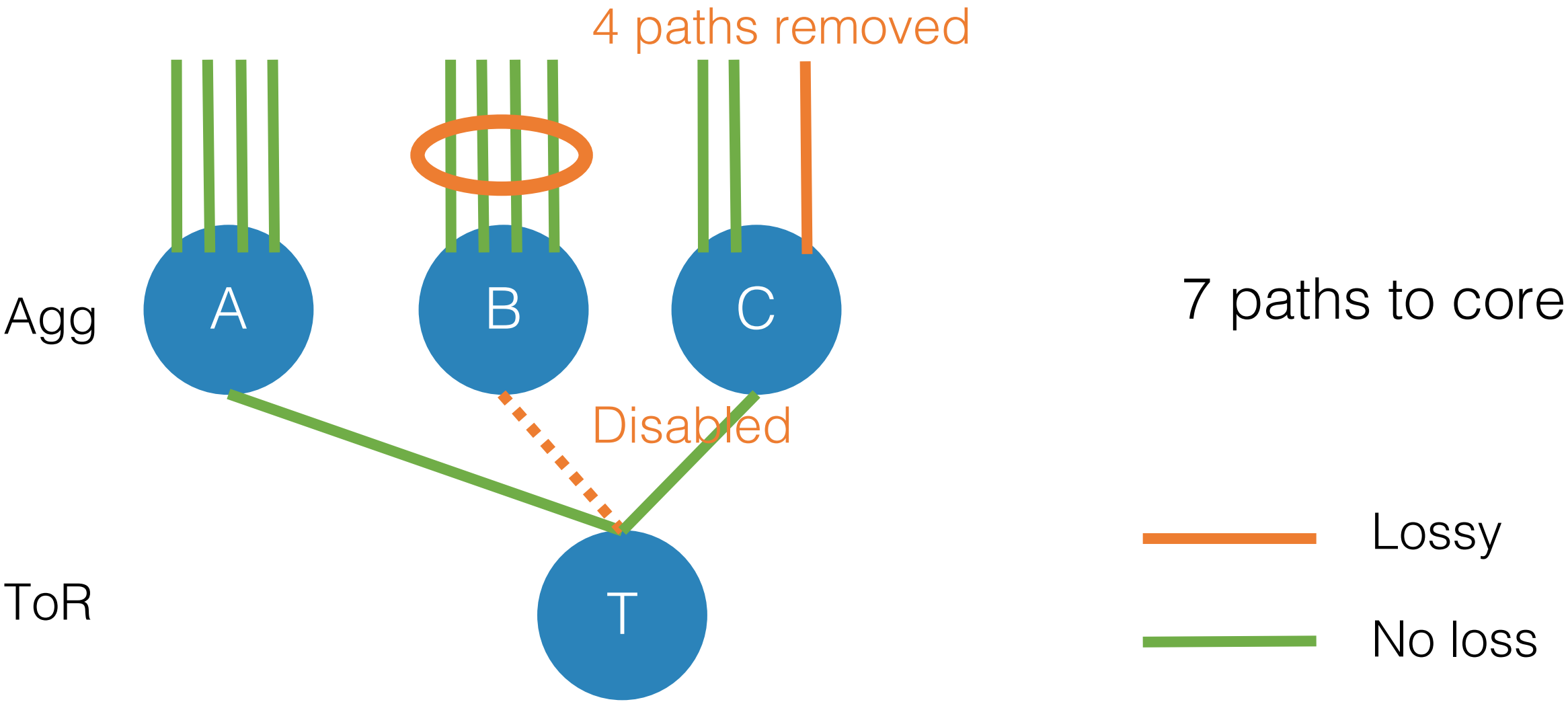


11 paths to core

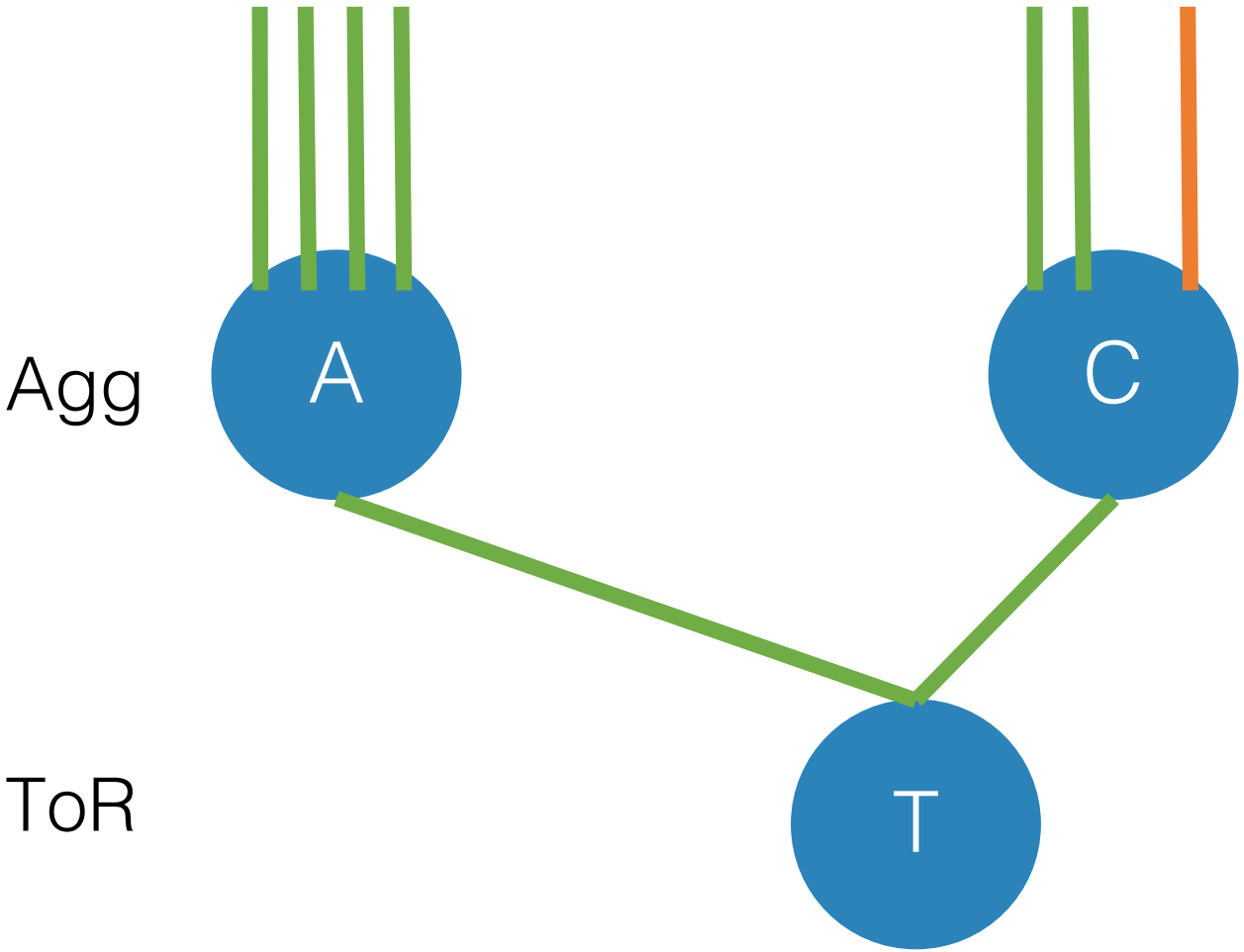
Lossy

No loss

Challenge #1: Meet Capacity Constraint



Challenge #1: Meet Capacity Constraint

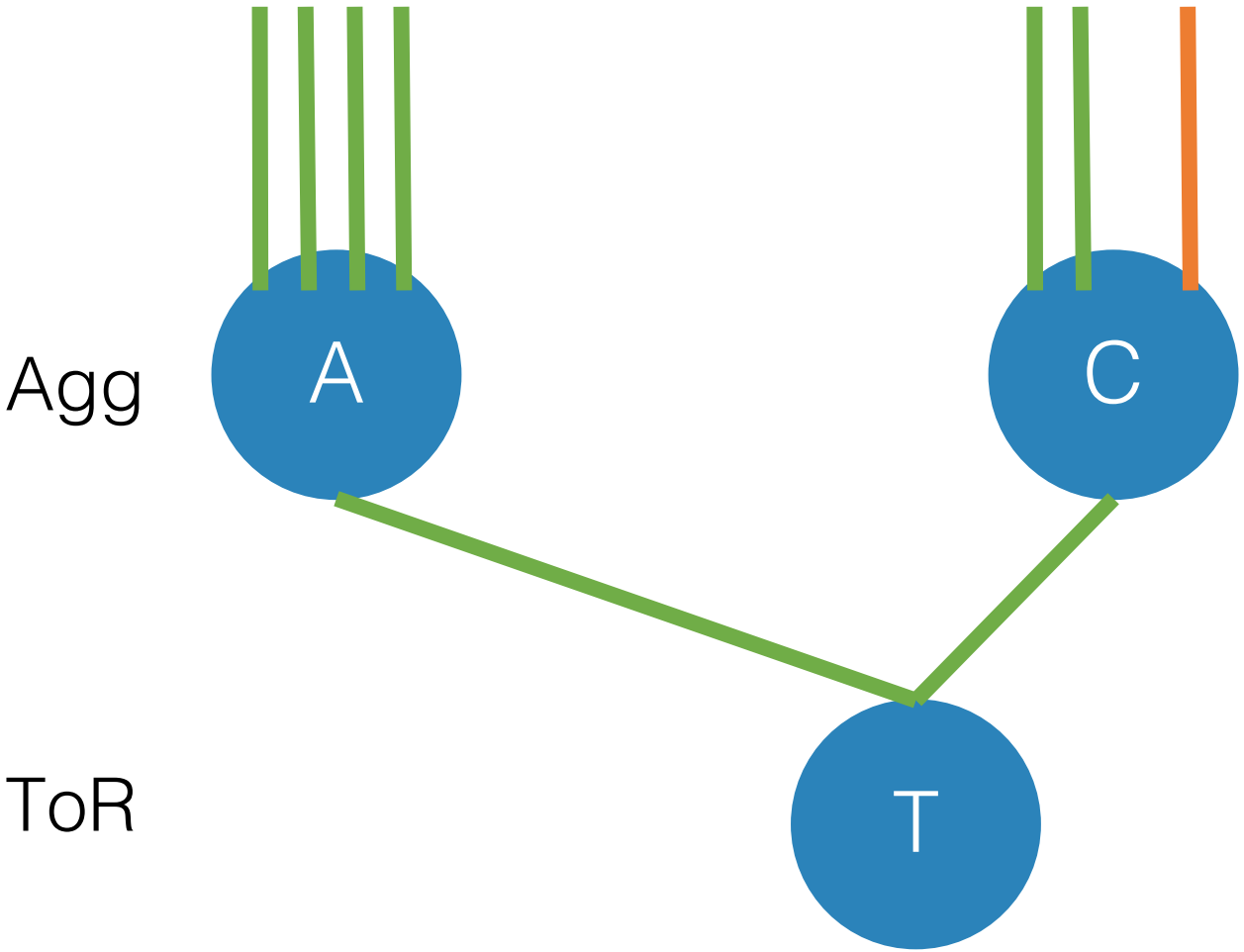


7 paths to core

Lossy

No loss

Challenge #1: Meet Capacity Constraint



Capacity constraint : 50%

7/16 < 50%

Lossy

No loss

Challenge #2: Find Root Cause

- Different repair is needed for different root cause
  - Dirty connector → Clean connector
  - Damaged fiber → Replace fiber
  - Dying transceiver laser → Replace transceiver
- Corruption persists if repair attempt is incorrect
  - Each attempt takes 2 days

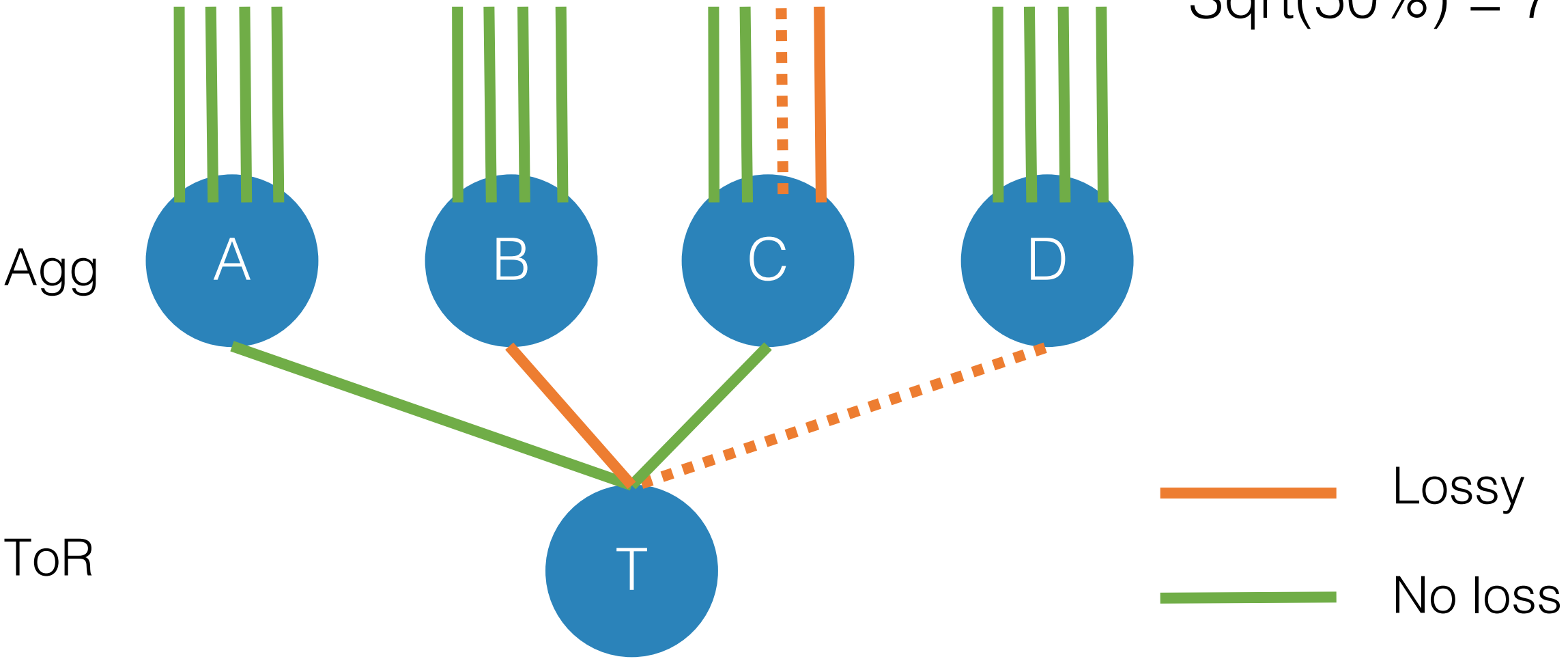
Framework	State-of-art	CorrOpt
Disable corrupting links and maintain capacity constraint	**Switch-local approach**	Conservative
Diagnose & repair	**Symptom-agnostic**	Takes too long to repair.

Switch-local Approach

- Capacity constraint: Every rack has x% of available paths to the core
  - For 3-level Clos network, each switch sets local uplink threshold to be sqrt(x%)
- Switches react to corrupting link independently

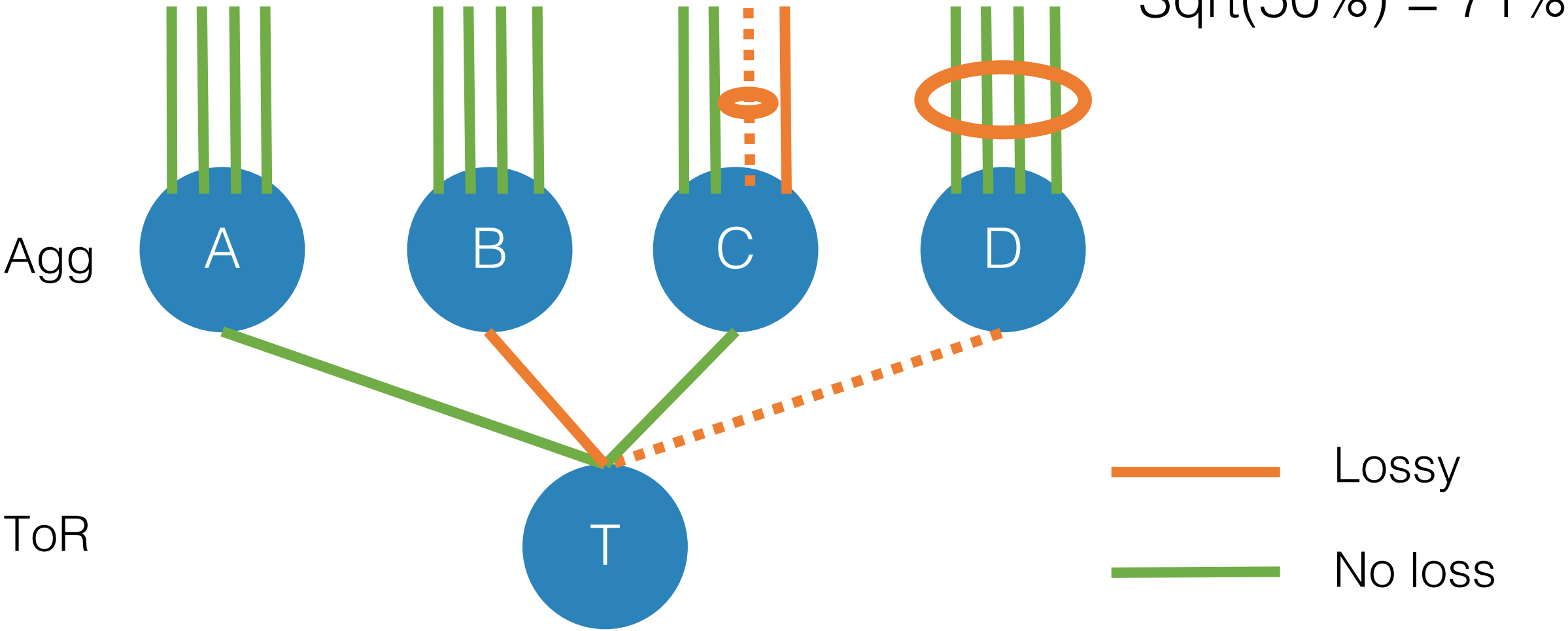
Switch-local is Conservative

Sqrt(50%) = 71%

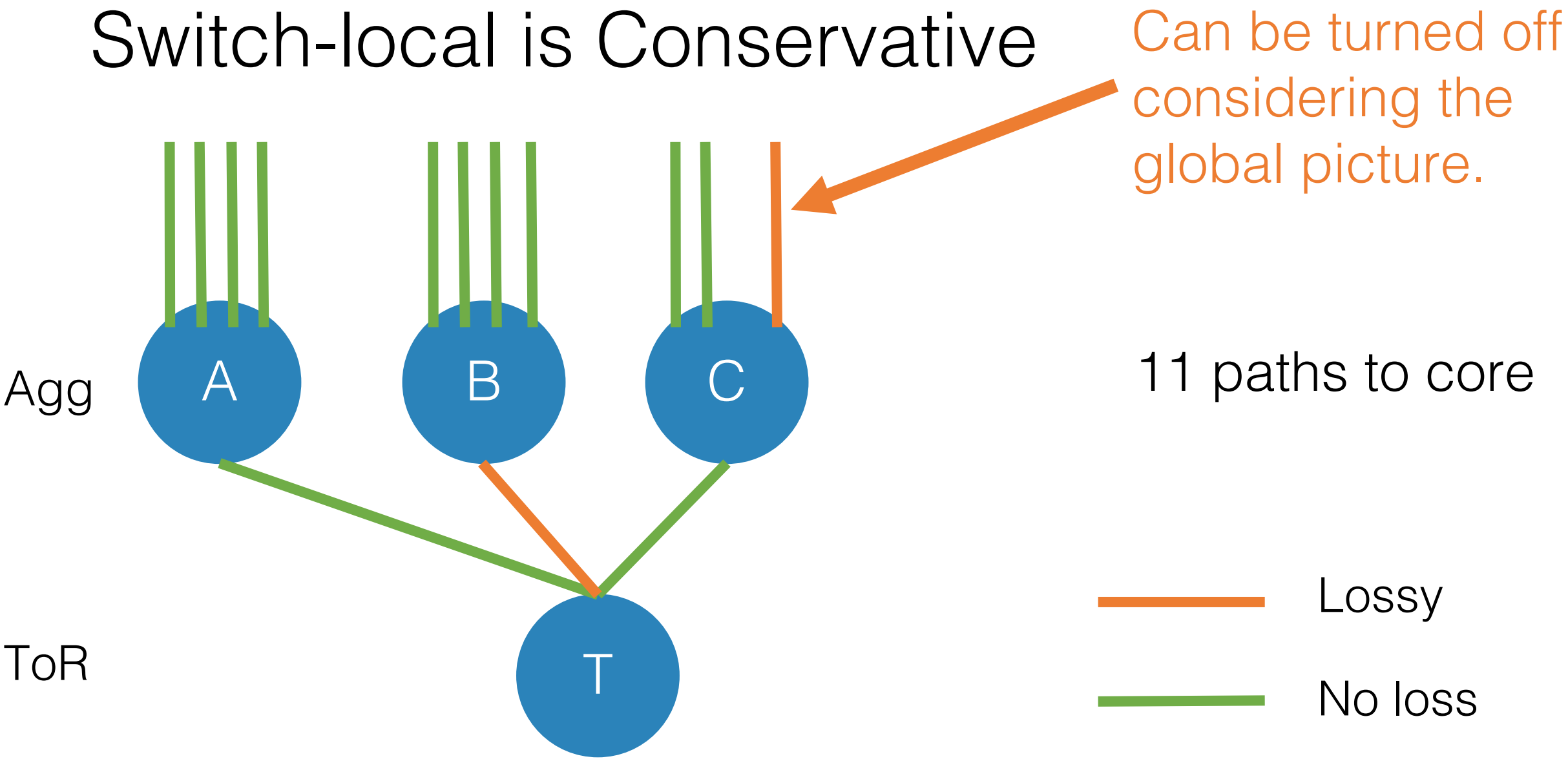


Switch-local is Conservative

Sqrt(50%) = 71%



Switch-local is Conservative

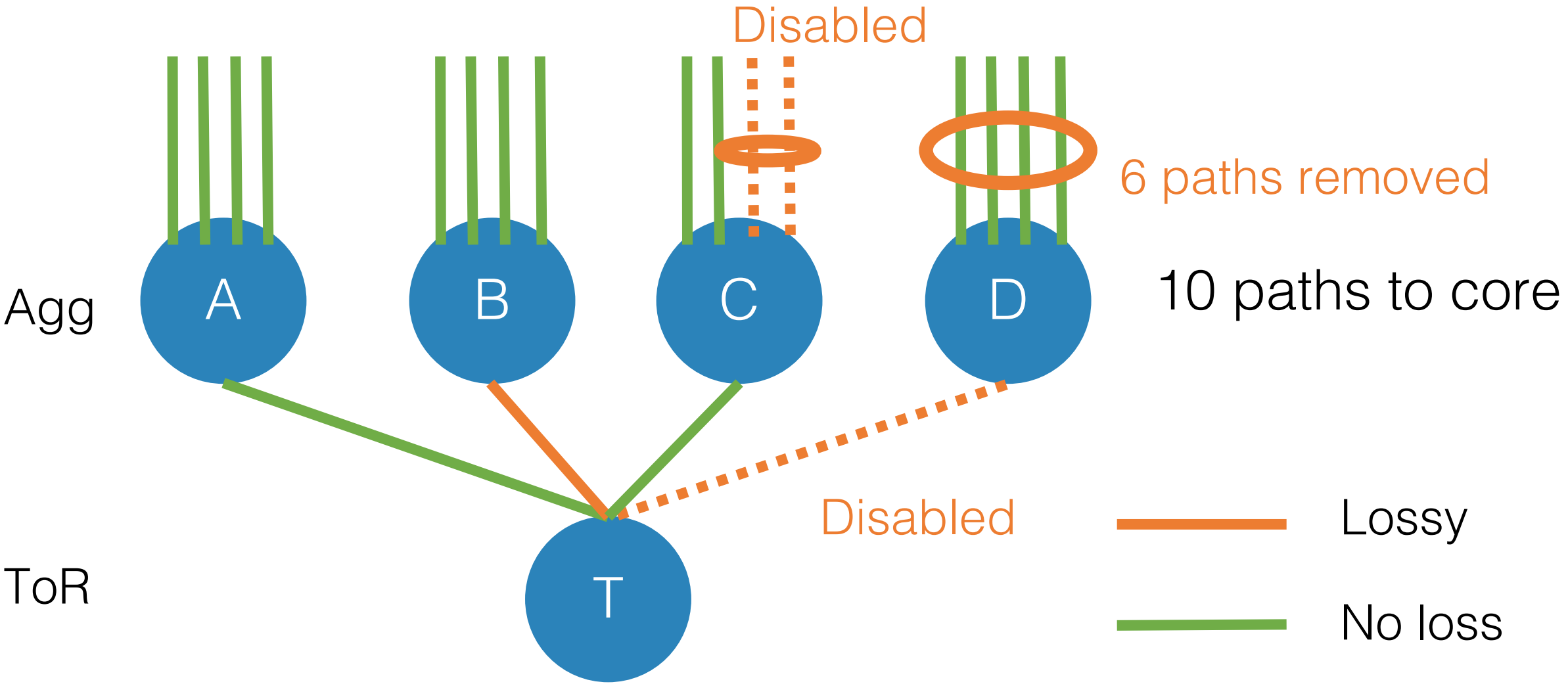


Framework	State-of-art	CorrOpt
Disable corrupting links and maintain capacity constraint	Switch-local approach	**Global approach**
Diagnose & repair	Symptom-agnostic	

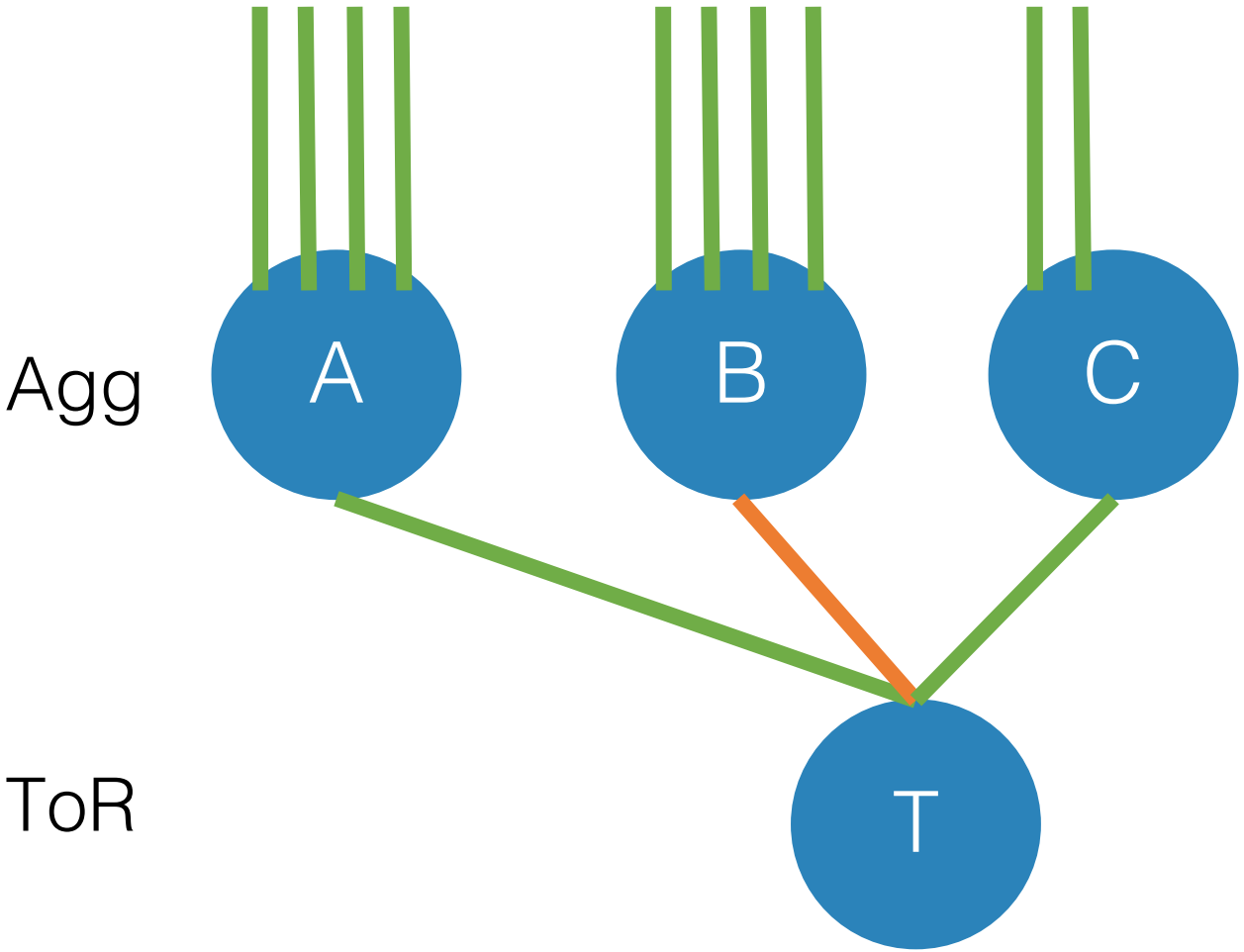
CorrOpt: Global Approach

- Given the set of links with corrupting rates
- Find a subset of corrupting links to disable such that
  - Meet capacity constraint
  - Minimize ∑ link loss rate

CorrOpt: Global Approach



CorrOpt: Global Approach



10 paths to core

Lossy

No loss

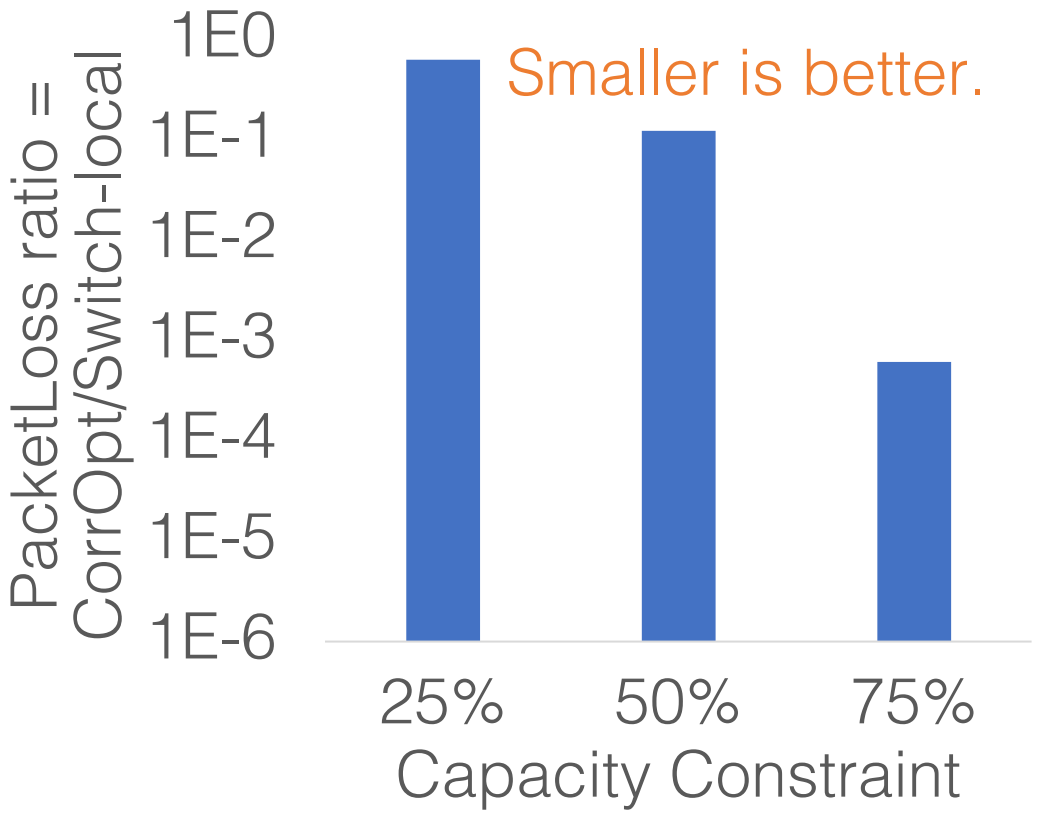
CorrOpt: Online Optimization

- On detecting a new corrupting link
  - Check whether turning off the corrupting link violates capacity constraint
  - O(number of links)
- On link repair
  - Optimize for the best subset of links that minimize corruption
  - NP-complete with a small problem size

Evaluation Methodology

- Replay corruption events from Oct 2016 – Dec 2016
- Assume corrupting link takes 2 or 4 days to repair after taken down
- Count total packet loss over the entire time duration

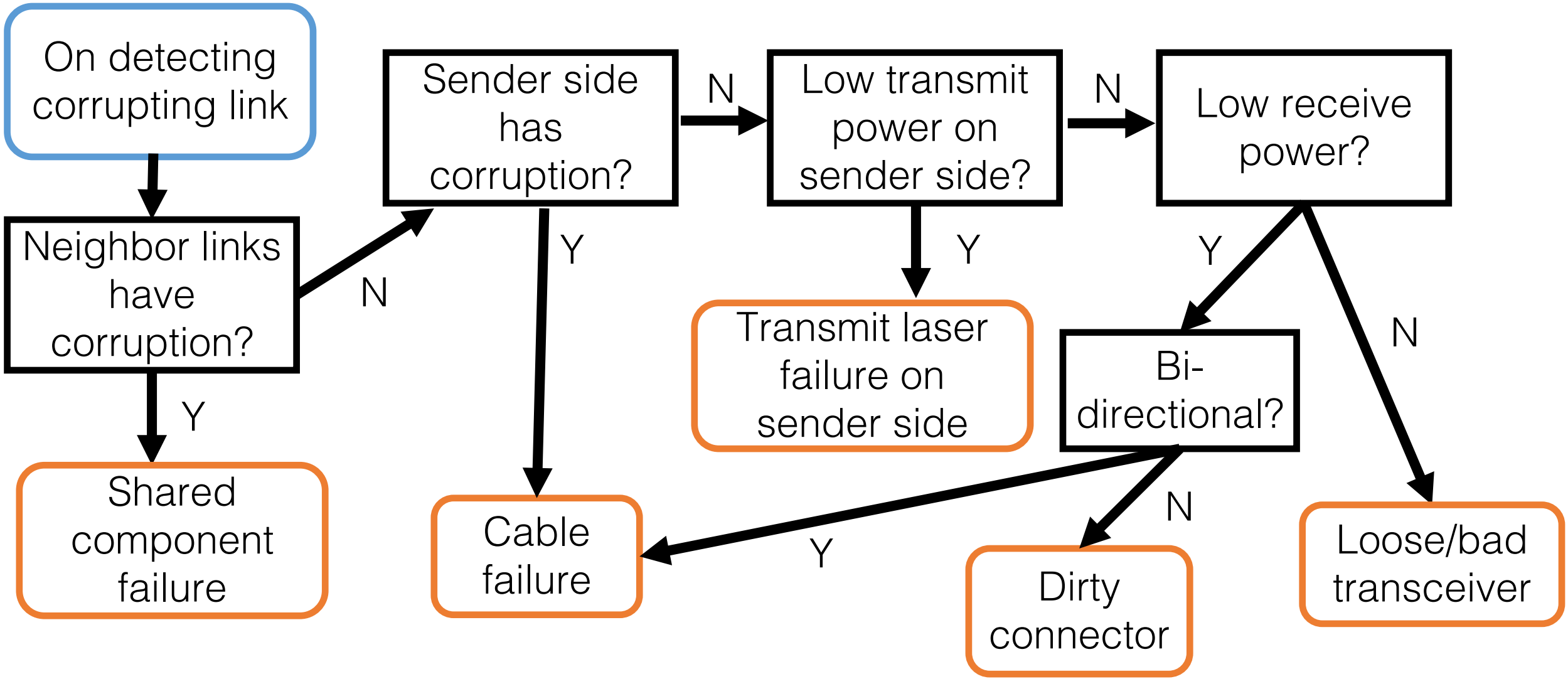
Evaluations of Disabling Corrupting Link



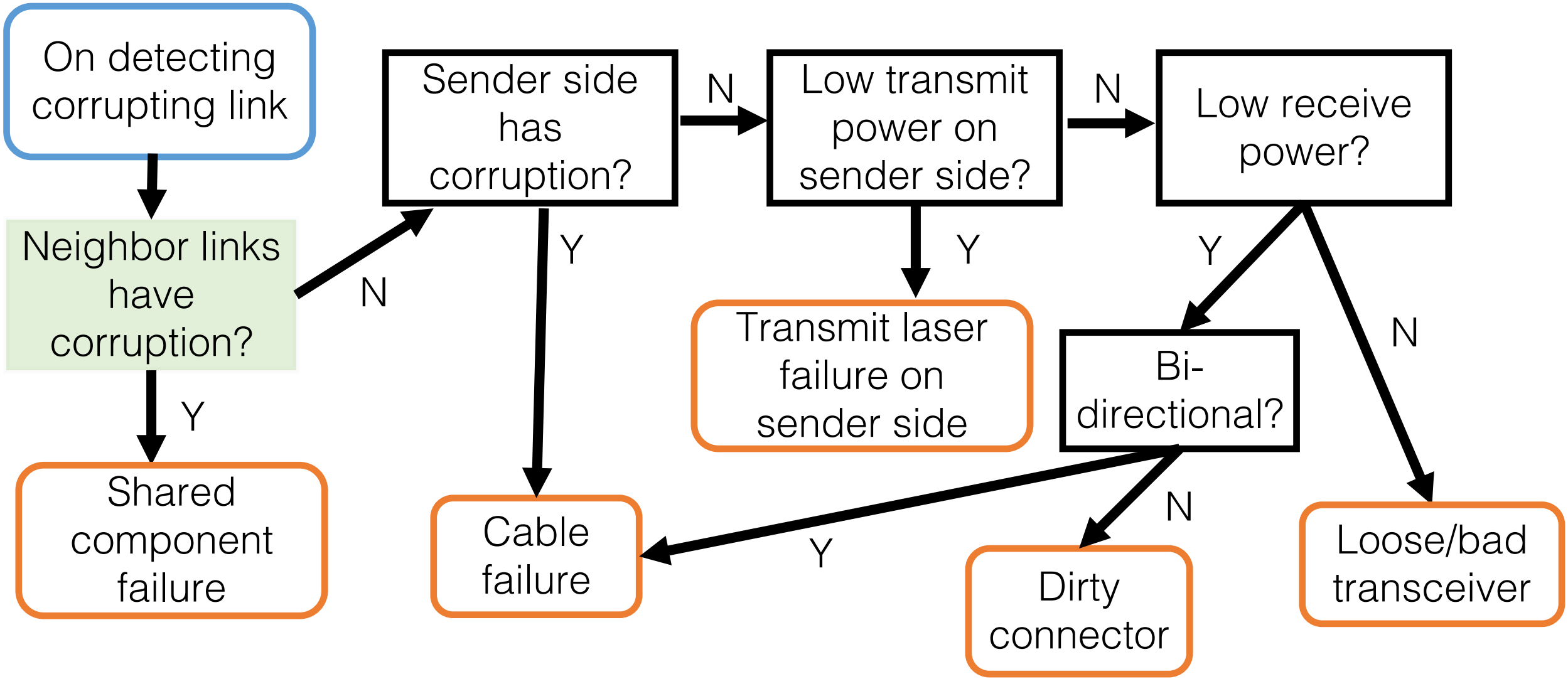
- Reduce corruption
- Fast
  - On detecting new corrupting link: 100-300ms
  - On link repair: <1 minute

Framework	State-of-art	CorrOpt
Disable corrupting links and maintain capacity constraint	Switch-local approach	Global approach
Diagnose & repair	Symptom-agnostic	**Symptom-aware**

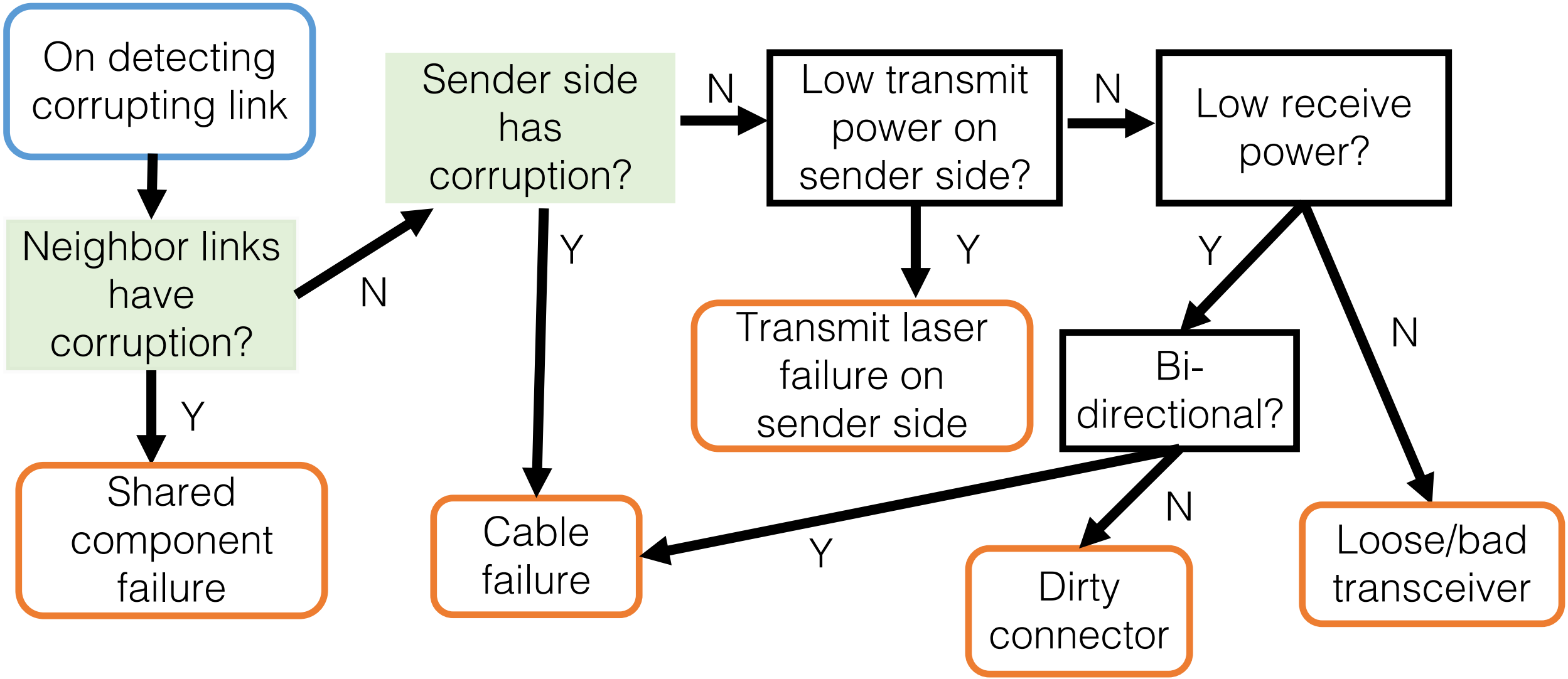
CorrOpt: Symptom-aware Diagnosis



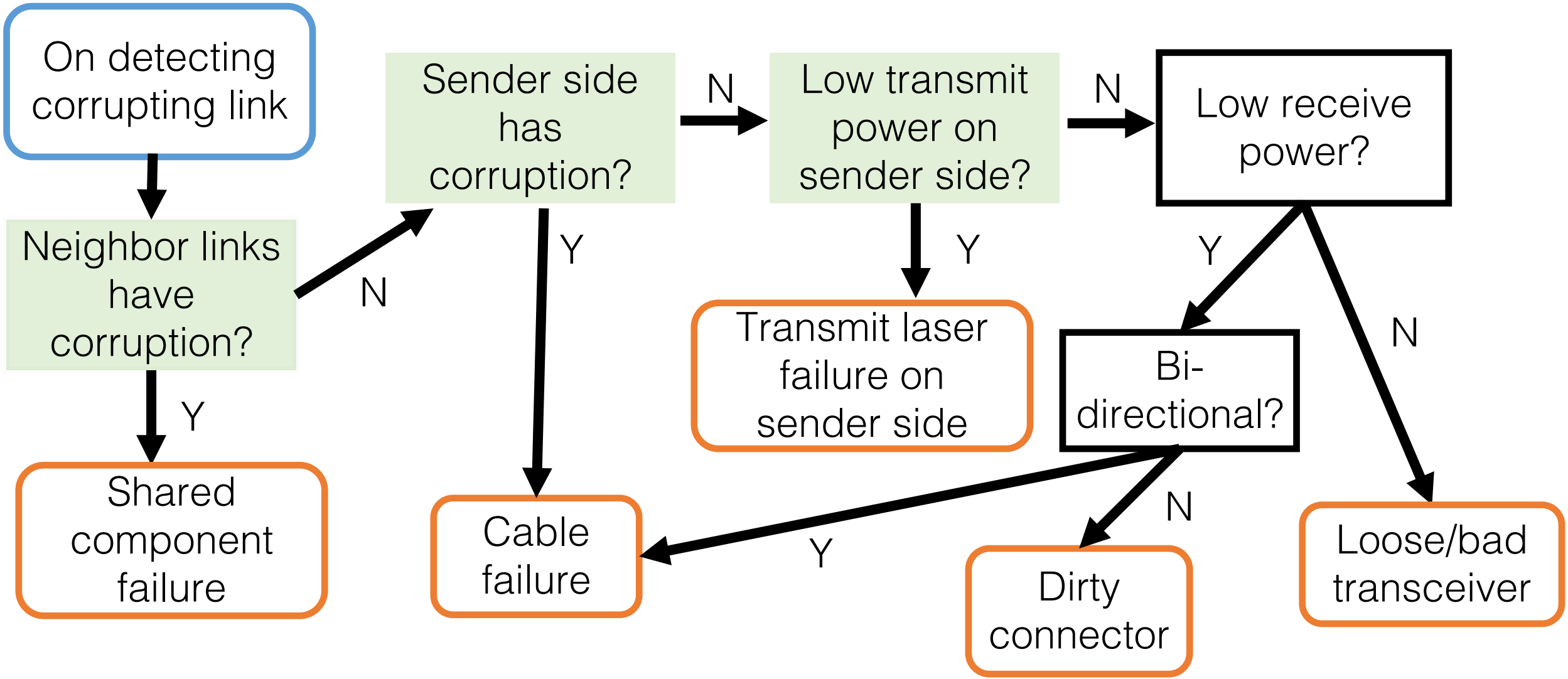
CorrOpt: Symptom-aware Diagnosis



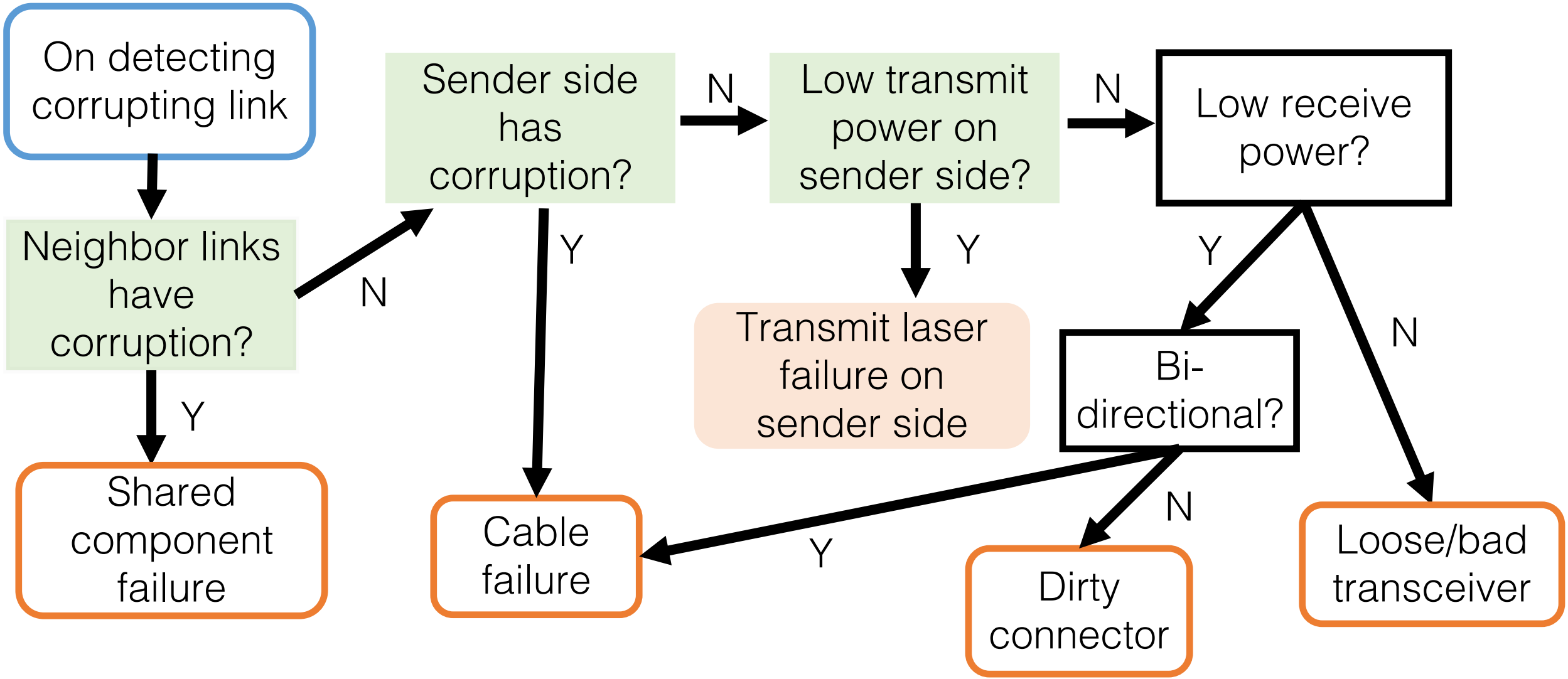
CorrOpt: Symptom-aware Diagnosis



CorrOpt: Symptom-aware Diagnosis



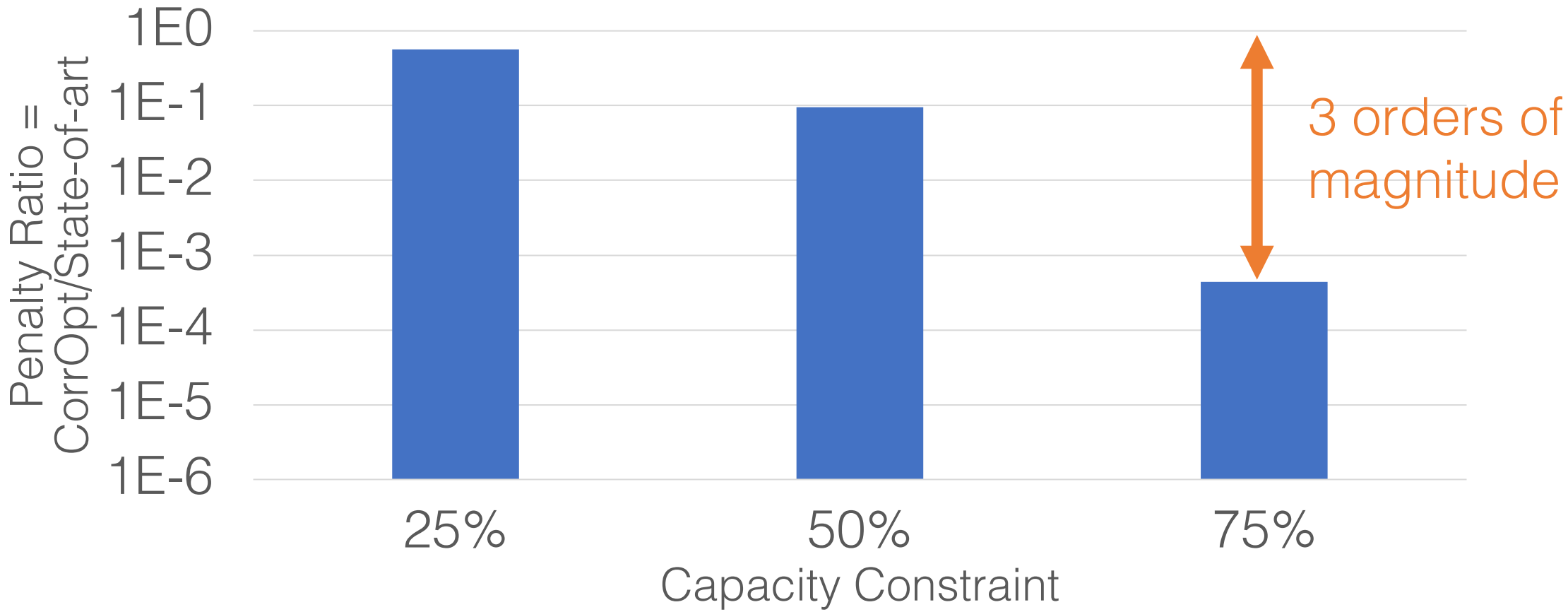
CorrOpt: Symptom-aware Diagnosis



Evaluations of Symptom-aware Diagnosis

- Deployed in Microsoft data centers
- Analyzing 300 repair tickets
  - Accuracy: 50% → 80%
  - Link becomes active sooner
  - Allow other corrupting links to be turned off

Combined Impact



Summary

- Packet corruption is a significant source of packet loss
- Packet corruption has distinctive symptoms and root causes
- CorrOpt reduces corruption by 3 orders of magnitude