

FiDe: Reliable and Fast Crash Failure Detection to Boost Datacenter Coordination

Davide Rovelli^{1,2}, Pavel Chuprikov³, Philipp Berdesinski⁴, Ali Pahlevan², Patrick Jahnke⁴, Patrick Eugster¹

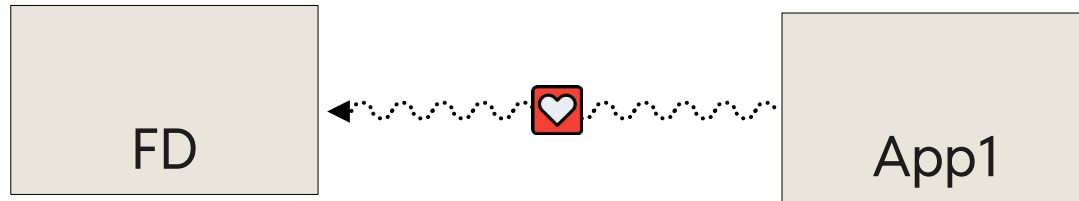
Corresponding author: davide.rovelli@usi.ch



Failure detection in cloud datacenters

- Highly-available core services run 24/7 distributed in cloud datacenters
- Failure detectors (FDs) commonly used as building block

The heartbeat problem



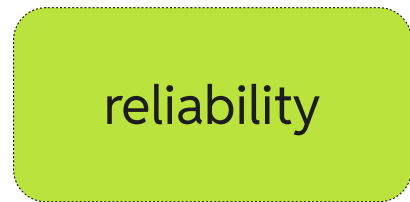
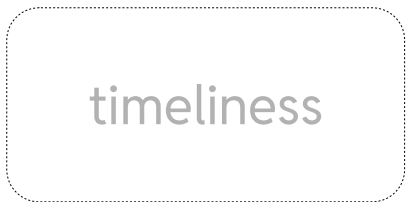
Timeouts must account for jitter in processing and communication

The heartbeat problem

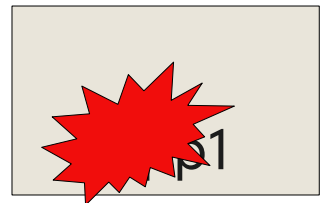


Large timeout

Properties:



The heartbeat problem



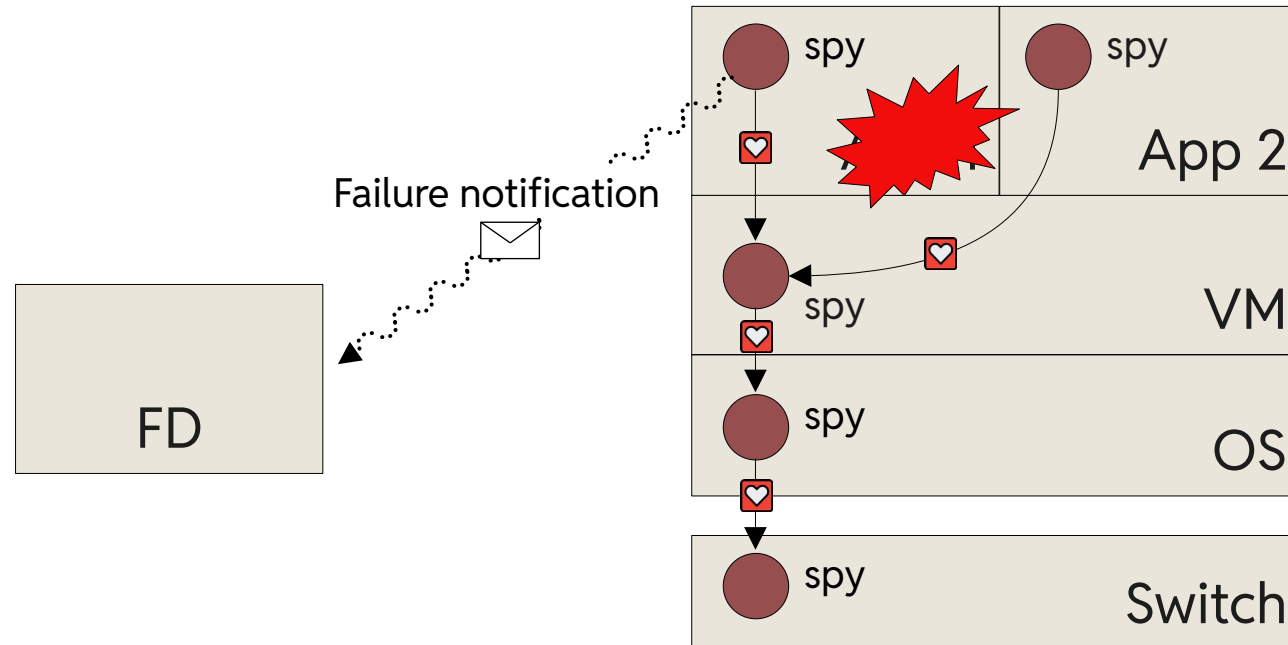
Small timeout

Properties:

timeliness

reliability

Multi-level/gray failure detectors [1][2]

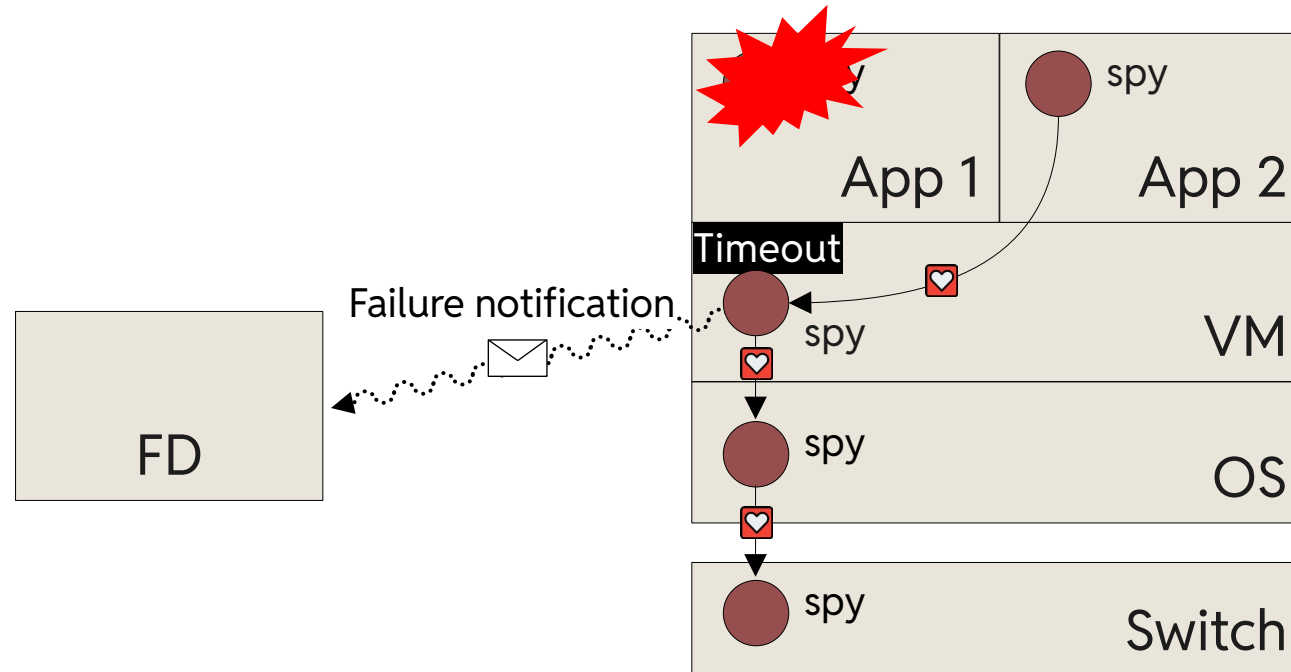


Custom
monitoring spies
at every level

[1] Joshua B. Leners, Hao Wu, Wei-Lun Hung, Marcos K. Aguilera, and Michael Walfish. Detecting failures in distributed systems with the falcon spy network. SOSP11

[2] R. Guerraoui, A. Murat, J. Picorel, A. Xygkis, H. Yan, and P. Zuo. uKharon: A membership service for microsecond applications. ATC22.

Multi-level/gray failure detectors



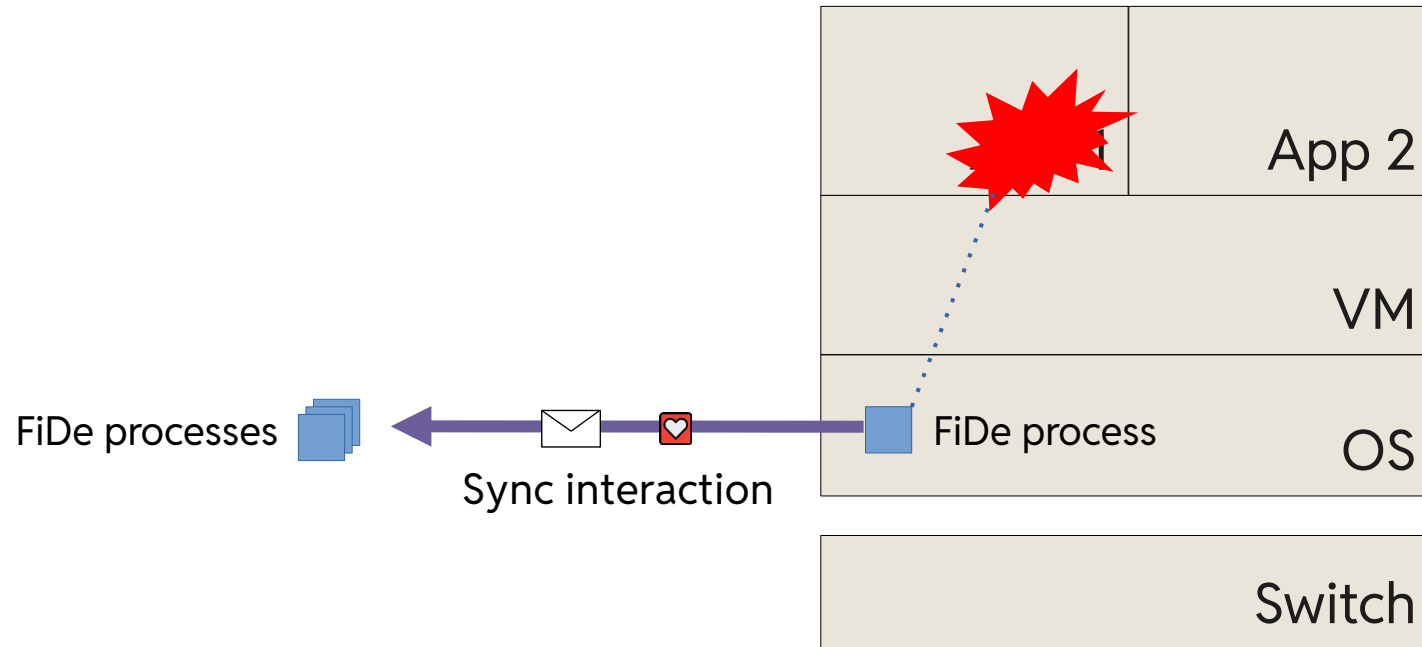
Still large/small
timeout trade-off

Properties:

timeliness

reliability

FiDe



- Focus on crash
- System-driven sync interactions
- OS watchdog

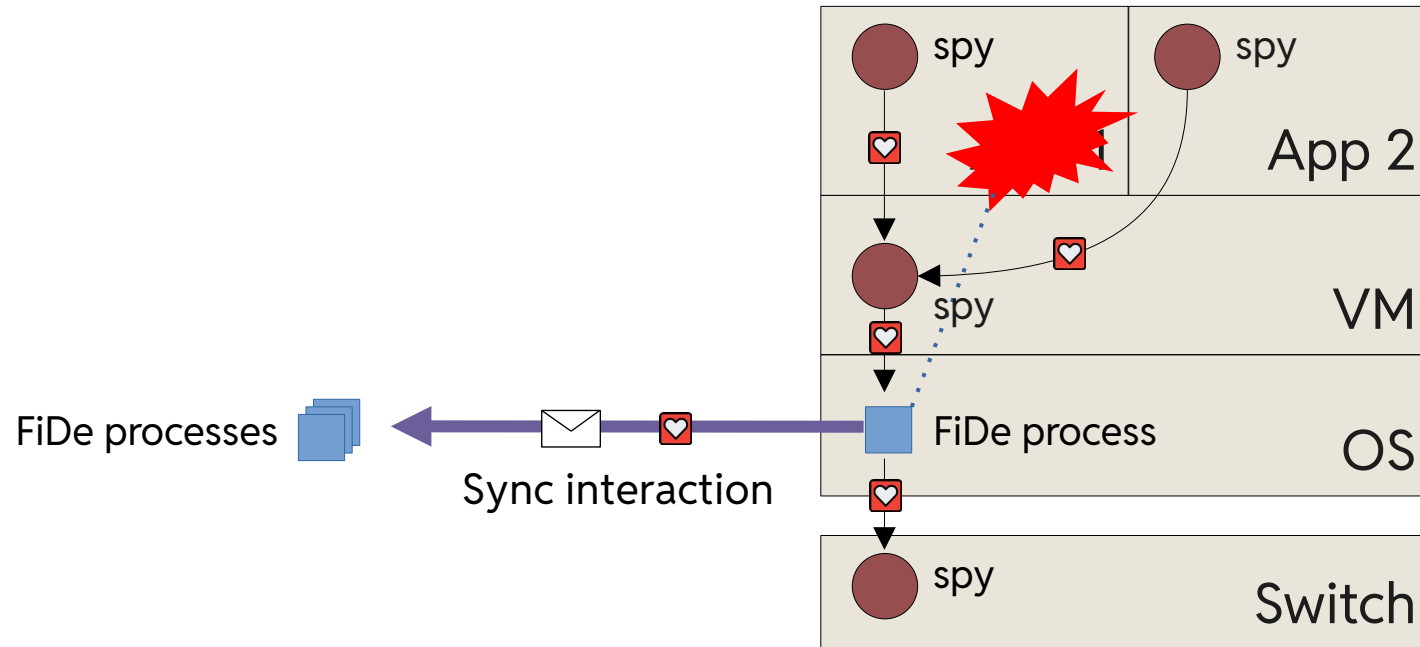
Properties:

timeliness

reliability

non-
intrusiveness

FiDe + gray failure detectors (GFDs)



GFD integration for coverage

Properties:

timeliness

reliability

non-intrusiveness

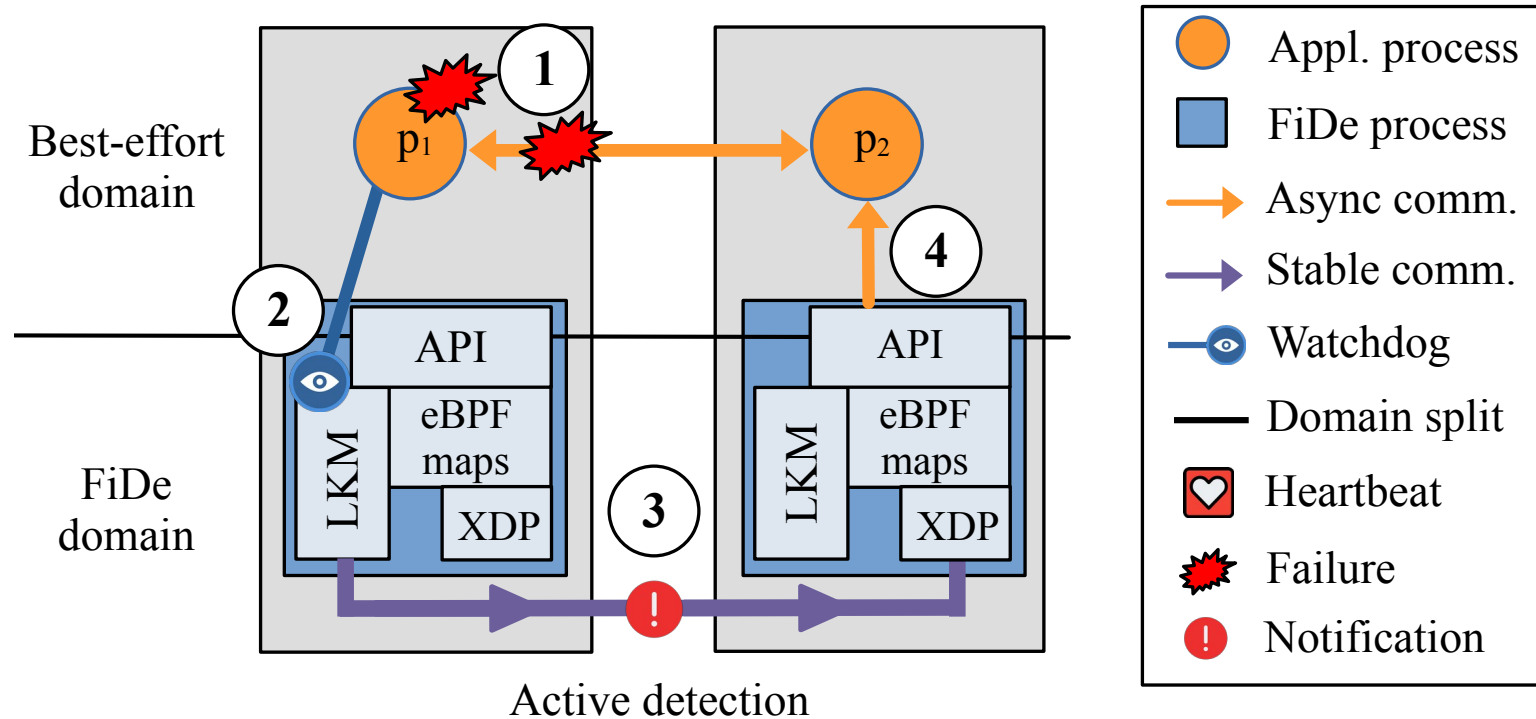
FiDe's contributions

1. Reliable, fast crash detection through synchronous interactions
2. Building block for simple and efficient coordination in *failure-free* executions

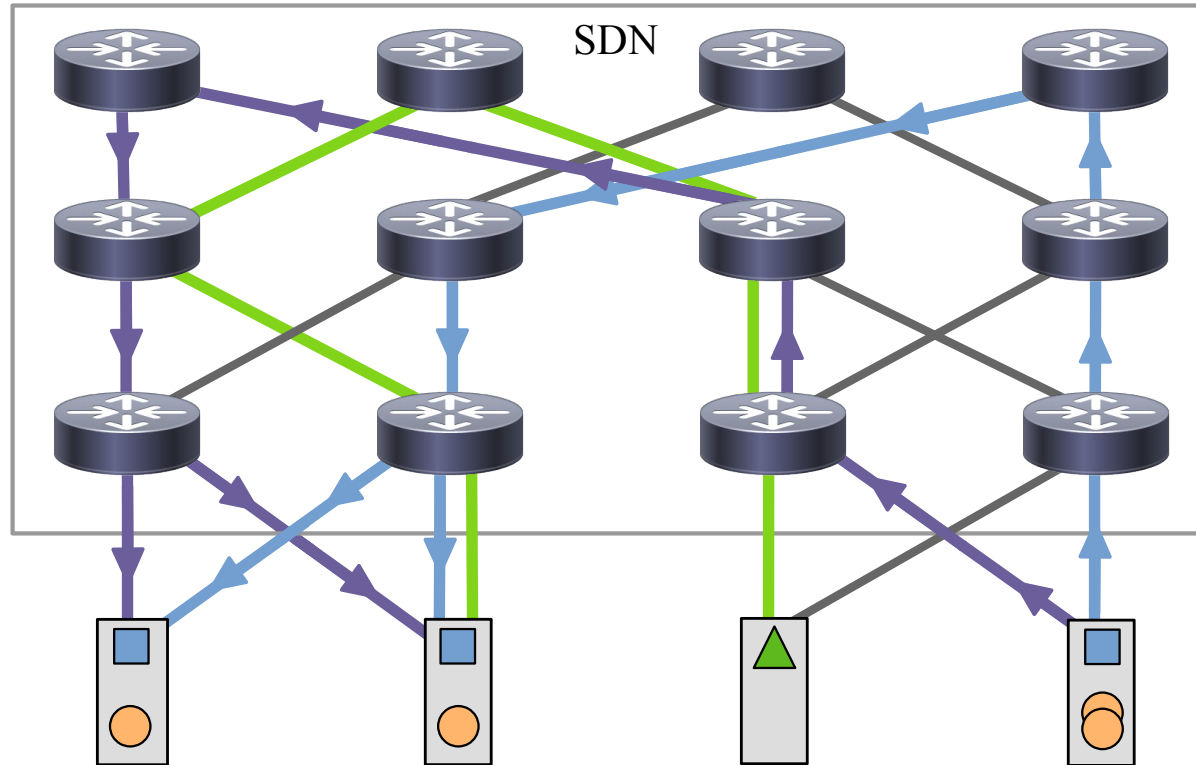
FiDe's contributions

1. Reliable, fast crash detection through synchronous interactions
2. Building block for simple and efficient coordination in *failure-free* executions

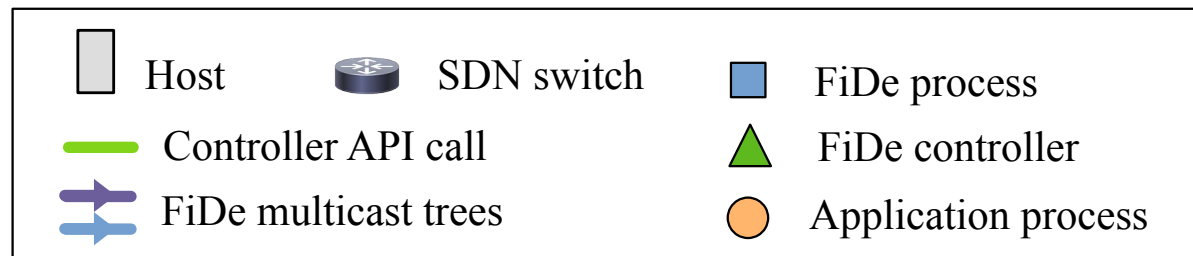
1. Reliable, timely detection & processing



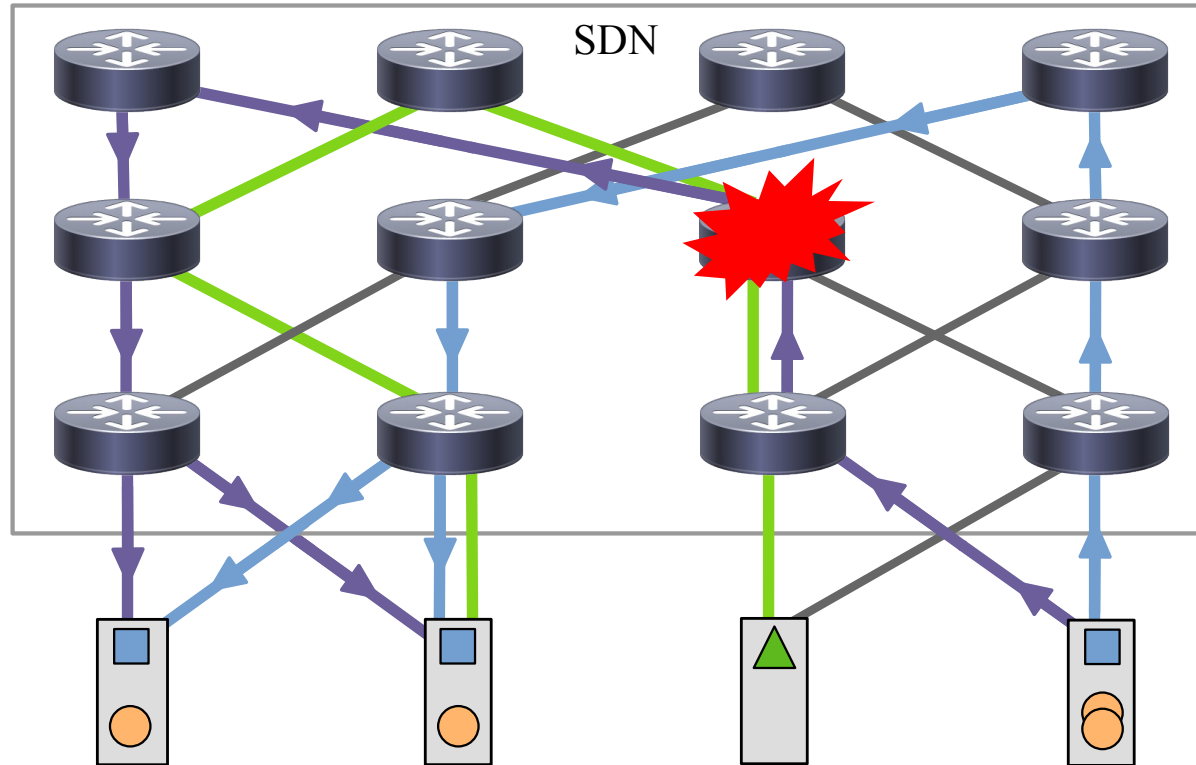
1. Reliable, timely communication



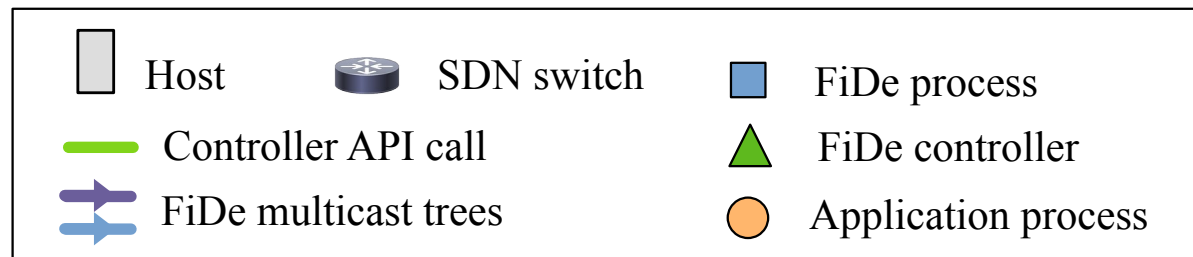
Traffic engineering +
multicast for bounded
latency



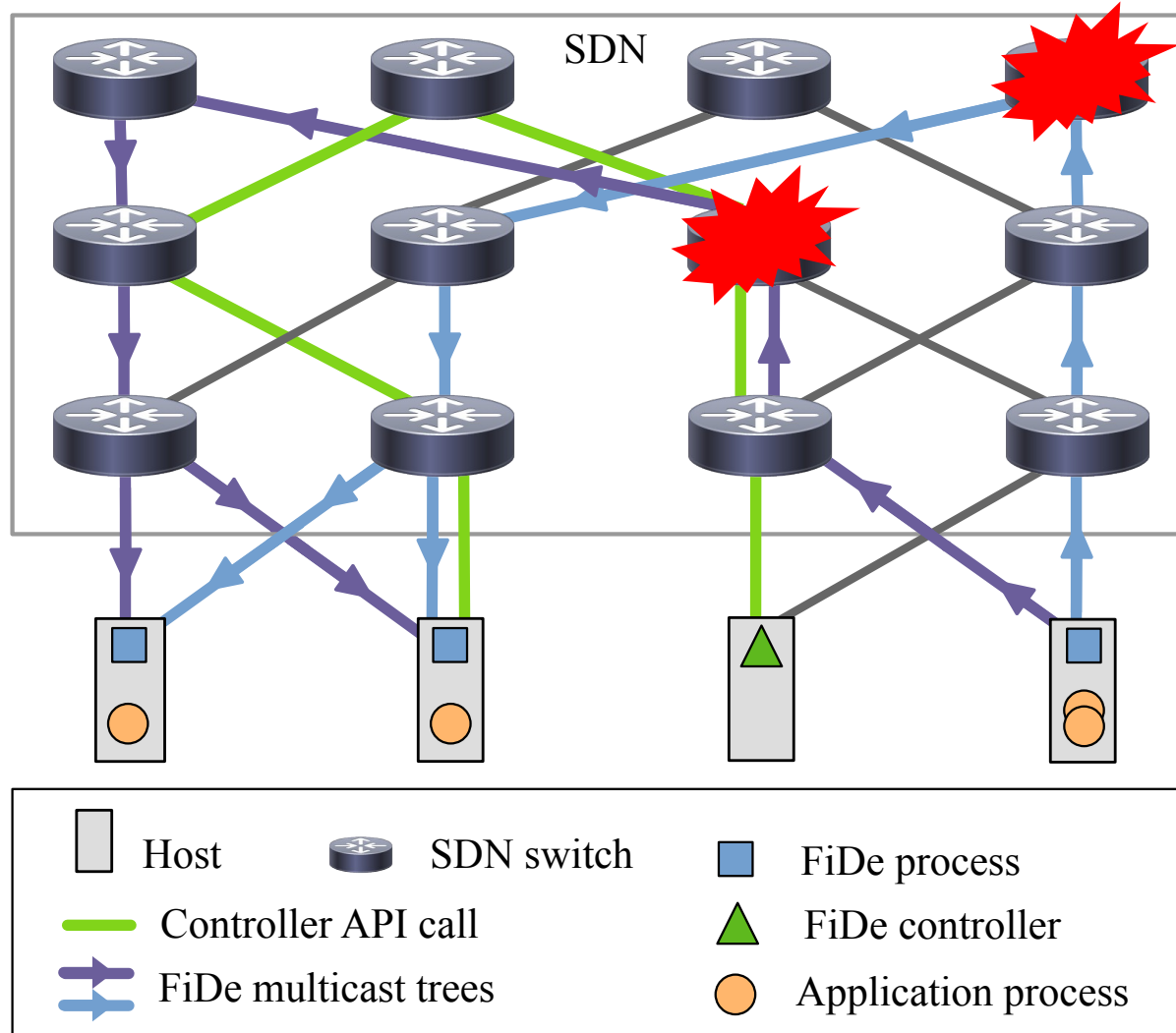
1. Reliable, timely communication



Redundancy and network recovery to overcome network failures



1. Reliable, timely communication



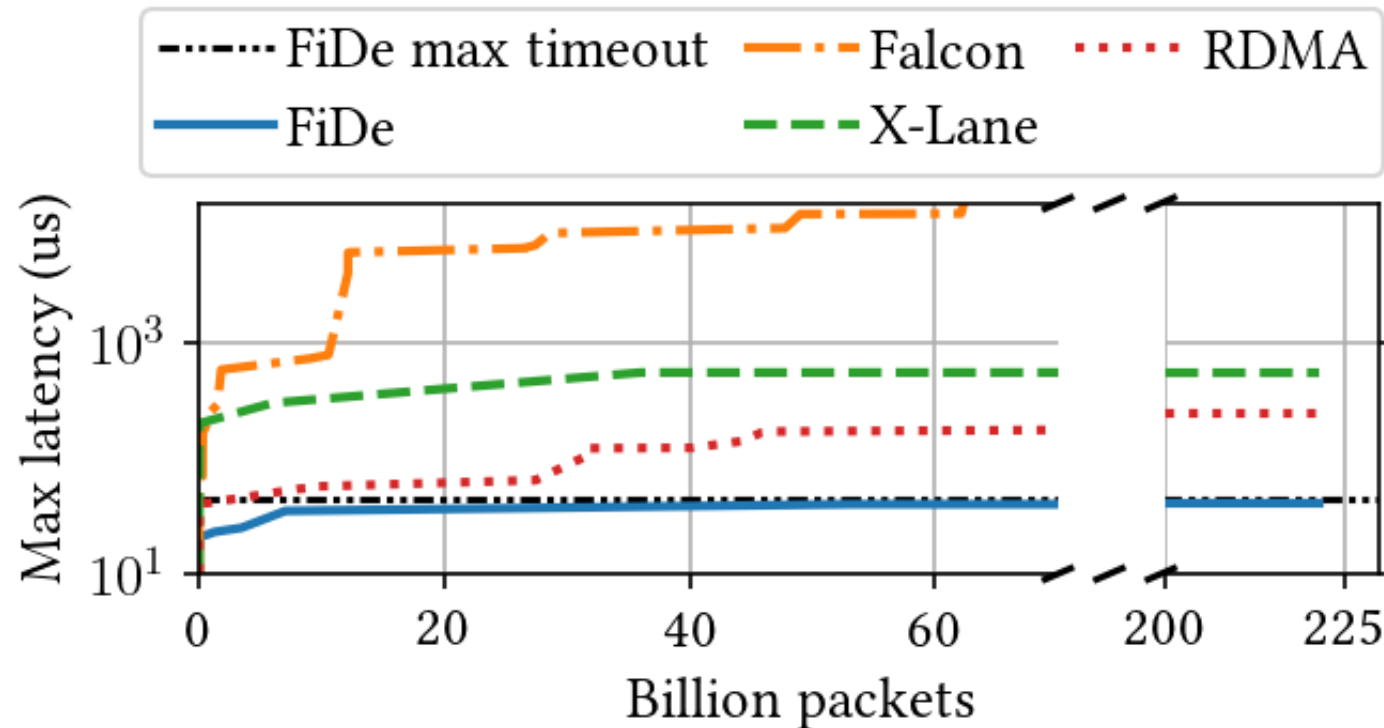
FiDe does not tolerate critical compound failure (i.e. quick network failure succession)

Probability: $1/22.7 \text{ years}^*$

*assuming 500Mbps traffic with failure statistics from Phillipa Gill et al. Understanding network failures in data centers: measurement, analysis, and implications. SIGCOMM 2011.

1. Sync interaction & FD timeout

SAP prod. cluster: 6x servers with ConnectX-4 NICs, 1 Arista 7280CR-48 ToR switch



FiDe worst-case latency/timeout < 45us over several weeks

FiDe's contributions

1. Reliable, fast crash detection through synchronous interactions
2. Building block for simple and efficient coordination in *failure-free* executions

2. FiDe-based coordination services

Two novel consensus algorithms

Heartbeat stabilizing uniform consensus (HSUC)

- No majority, use FiDe for reliable leader election

Heartbeat uniform consensus (HUC)

- Piggyback messages to FiDe for consistent + timely delivery

2. FiDe-based coordination services

Two novel consensus algorithms

Heartbeat stabilizing uniform consensus (HSUC)

- No majority, use FiDe for reliable leader election

Heartbeat uniform consensus (HUC)

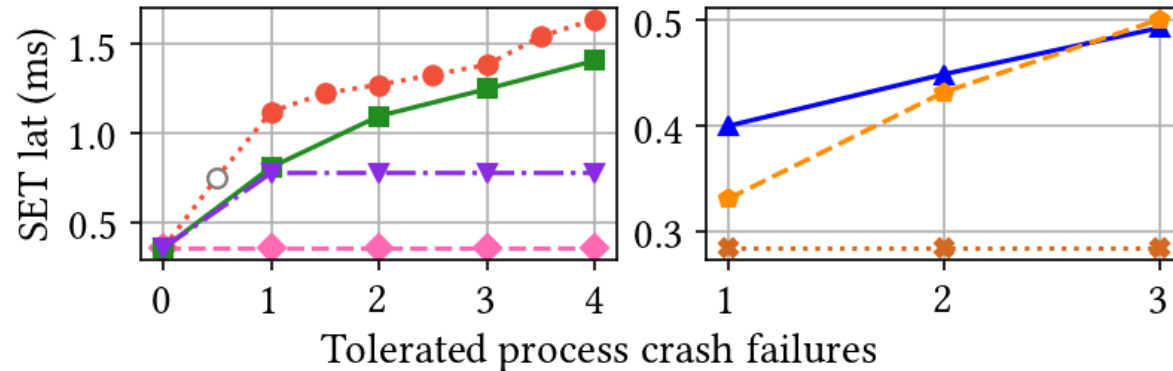
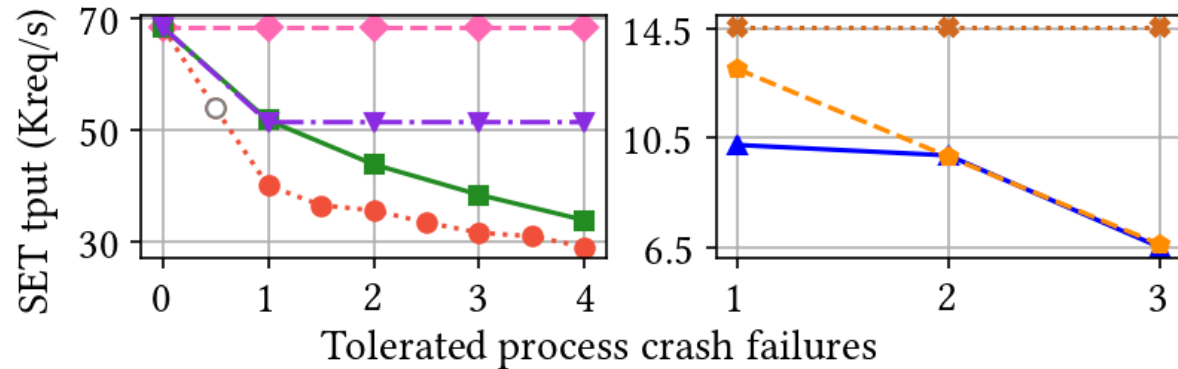
- Piggyback messages to FiDe for consistent + timely delivery

2. FiDe-based coordination services

Takeaways:

- improved protocol simplicity
- improve performance in the *failure-free* case

2. FiDe-based services in real-world apps



HSUC/HUC vs classic SMR services

■ Redis KV store and Zookeeper

■ ~1.5x throughput

■ ~0.5x latency

Conclusions

- Robust timeouts via datacenter “selective” synchrony
 - kernel support for stable processing
 - network reservation and traffic engineering for communication
- FiDe achieves reliable crash μ s-scale detection
- FiDe enables simpler and more efficient protocols in failure-free executions
- Future: hardware-offload suited for the selective-sync model

Thank you!

FiDe: Reliable and Fast Crash Failure Detection to Boost Datacenter Coordination

me → davide.rovelli@usi.ch

group → swsystems.usi.ch



turba