ISCAS 中国科学院软件研究所学术年会'2023 暨计算机科学国家重点实验室开放周
学术论文

覆盖导向的分布式系统故障注入测试

高钰，窦文生，王栋，冯文翰，魏峻，钟华，黄涛

Coverage Guided Fault Injection for Cloud Systems

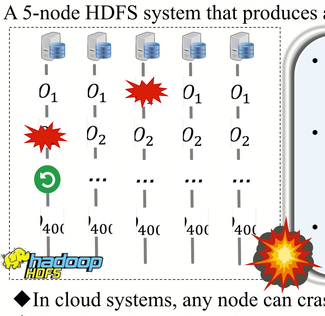
The 45th IEEE/ACM International Conference on Software Engineering (ICSE'23)

联系方式：高钰，13552557320，gaoyu15@otcaix.iscas.ac.cn

Systematic Recovery Testing for Cloud Systems is Challenging

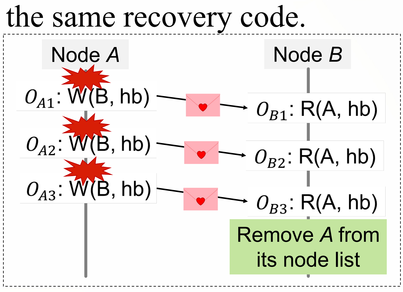
- Cloud systems face huge crash scenario space.

A 5-node HDFS system that produces around 400 I/O operations for each node



- $400 \times 5 = 2{,}000$ crash scenarios when injecting only **one node crash** on these I/O points
- About $400^2 \times C_5^2 = 1{,}600{,}000$ crash scenarios when injecting **two crashes** on the I/O points from two different nodes
- When considering **more node crashes** and even **node reboots** …

- Some crash scenarios may trigger the same recovery code.

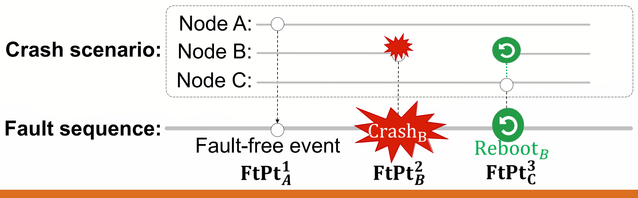


◆In cloud systems, any node can crash or reboot at any time. Node crashes/reboots can trigger crash recovery procedures.

◆Specific node crashes/reboots can trigger crash recovery bugs hidden in incorrect crash recovery mechanisms and implementations.

Fault Model

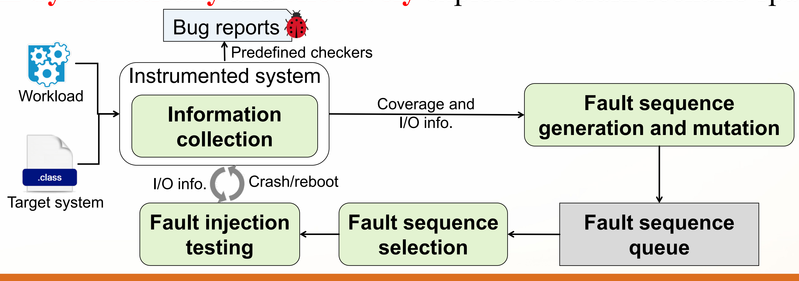
- We use fault sequences to represent various crash scenarios
- Fault sequence: all the I/O points executed in a run and their corresponding events.



Fault point info.
Node ID
Call stack
Timestamp
Event type
…

CrashFuzz

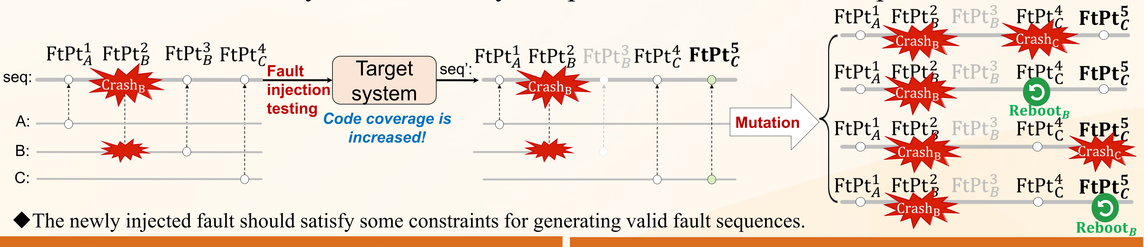
- **Systematically** and **effectively** explore the crash scenario space of cloud systems.



◆Guide a cloud system to cover new crash recovery code and increase the chance of triggering crash recovery bugs.
- Take a fault sequence as a special system input.
- Adjust fault sequences according to system feedbacks .

Fault Sequence Generation and Mutation

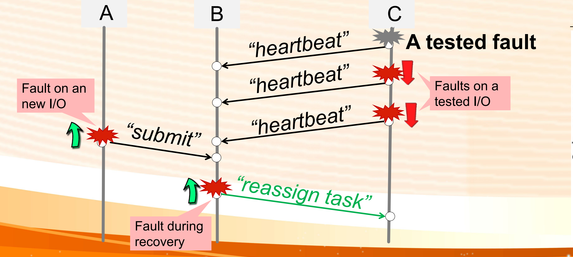
- Keep the fault sequence that increases the code coverage.
- Create a new fault sequence based on the executed I/O points and injected faults.
- Crash or reboot only one non-faulty I/O point in the new fault sequence.



◆The newly injected fault should satisfy some constraints for generating valid fault sequences.

Fault Sequence Selection

- Prioritize the sequences that inject faults on new I/O points.
- Prioritize the sequences that inject faults on I/O points occurring during recovery.



Evaluation

Bug ID	Failure Symptoms
HBASE-26883	Data loss
ZOOKEEPER-4503	Data staleness
HBASE-26897	Cluster out of service
HBASE-26370	Misleading error message
HDFS-16508	Operation failure