U LISBOA | UNIVERSIDADE DE LISBOA

Experiences with BFT-SMaRt as a consensus substrate of Permissioned Ledgers

Alysson Bessani

This session:

EXPERIENCES WITH DEPLOYED BLOCKCHAINS

NEXUS MEDIA NEWS

POPSCI.COM/BLOGS

These companies claim blockchain could help fight climate change

Is it a breakthrough, or just a buzzword?

By Jeremy Deaton Nexus Media January 18, 2018

Madison McVeigh/CityLab/David Ryder/Reuters

The Tech That's Changing How Cities Help the Homeless

SARAH HOLDER / LINDA POON MAY 31, 2018

From mapping apps to the blockchain, new tools are intended to give cities the information they need to address this growing challenge.

DISRUPTOR DAILY

Home Technologies Industries News We're Hiring!

Blockchain | Consumer Health | FinTech + Cryptocurrency | Health | Healthcare | Industries | Legal | News | Technologies

Sexual Consent, STD Status, and More on the Blockchain

CAS PROFFITT | JANUARY 23, 2018 11:00 AM

Share Tweet LinkedIn

27 APRIL 2018 | ARTICLES

How Blockchain Can Eradicate Poverty in Third-World Countries

BY CHRISTINA COMBEN

TECHINASIA NEWS JOBS COMPANIES EVENTS ABOUT ADVERTISE

CATEGORIES MARKETS WRITERS

FEATURES / PROFILES

Eva Xiao · 22 Nov 2017 · 7 min read

How blockchain in space aims to challenge the dominance of Google and Amazon in internet services

Prime-Core Executive Search
Your Primary Source for Placement

Home | About Us | Clients | Candidates | Prime-Core Blog | News | Contact Us

POSTED BY KEN DROPIEWSKI

Treating Cancer with Blockchain Computing

Computer processing has become an important tool for diagnosing and treating cancer. The idea is that with personalized treatments coming into a clinical setting, rapid analysis of a patient's data becomes more crucial. The same applies to research in finding more cures for cancer.

mobihealthnews | PROVIDER | PAYER | PHARMA | CONSUMER | INVESTOR | Search

OncoPower uses blockchain, cryptocurrency to help users manage cancer

By **Laura Lovett** | March 19, 2018

SHARE Share 101

Correction: An earlier version of this story included an incorrect funding amount received by OncoPower.

A new platform exclusively for cancer patients is now registering new users. Witty Healthy's newly launched platform, called **OncoPower**, uses blockchain technology to help patients keep track of their medical data across providers and offer users incentives.

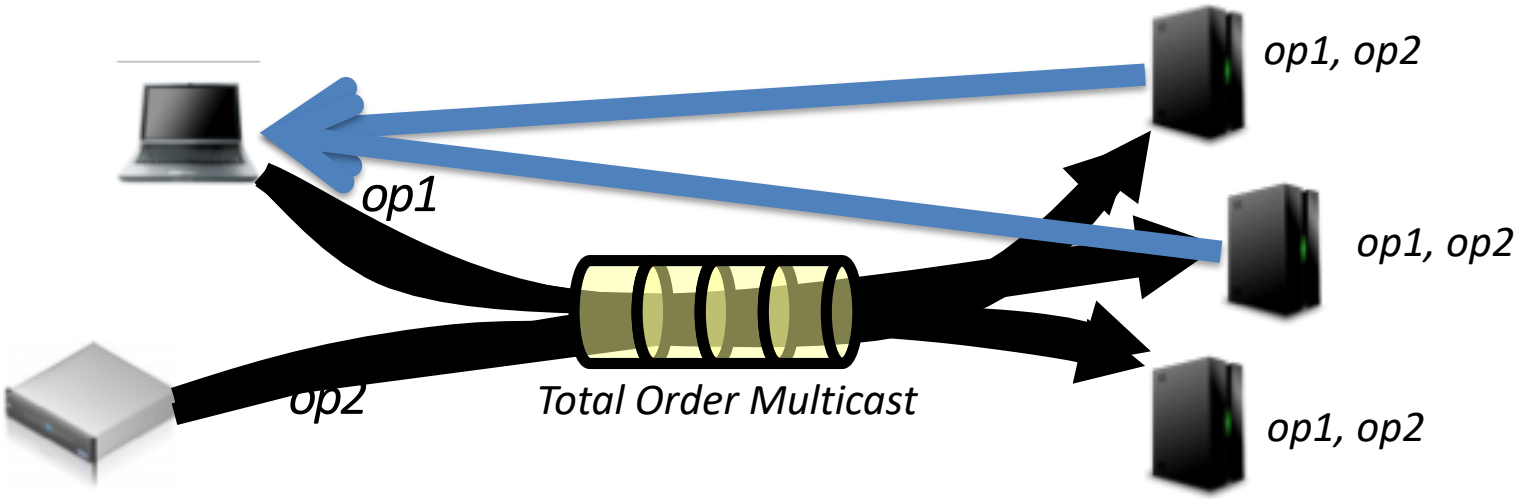
Outline

- BFT-SMaRt
  - Overview
  - Performance
- BFT-SMaRt in Permissioned Ledgers
  - Symbiont
  - Hyperledger Fabric
  - R3 Corda
- Beyond BFT-SMaRt

Part 1

BFT-SMART

State Machine Replication



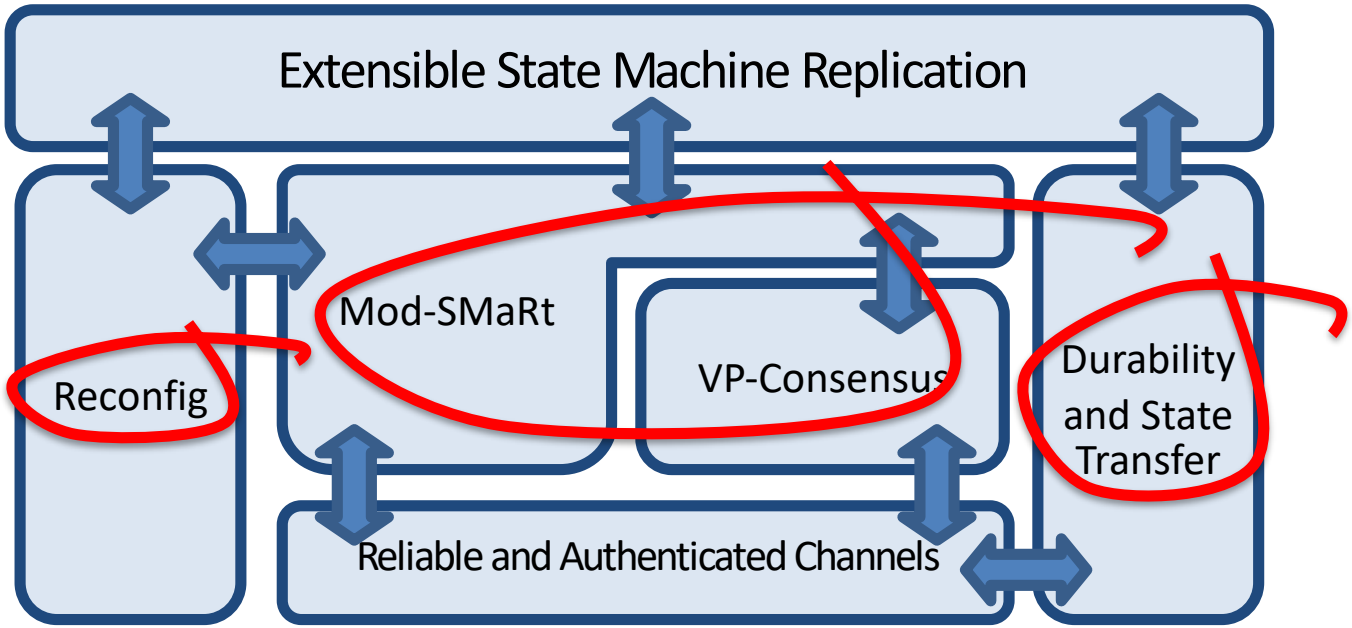
Safety: all replicas execute the same sequence of commands

Liveness: commands issued by correct clients are answered

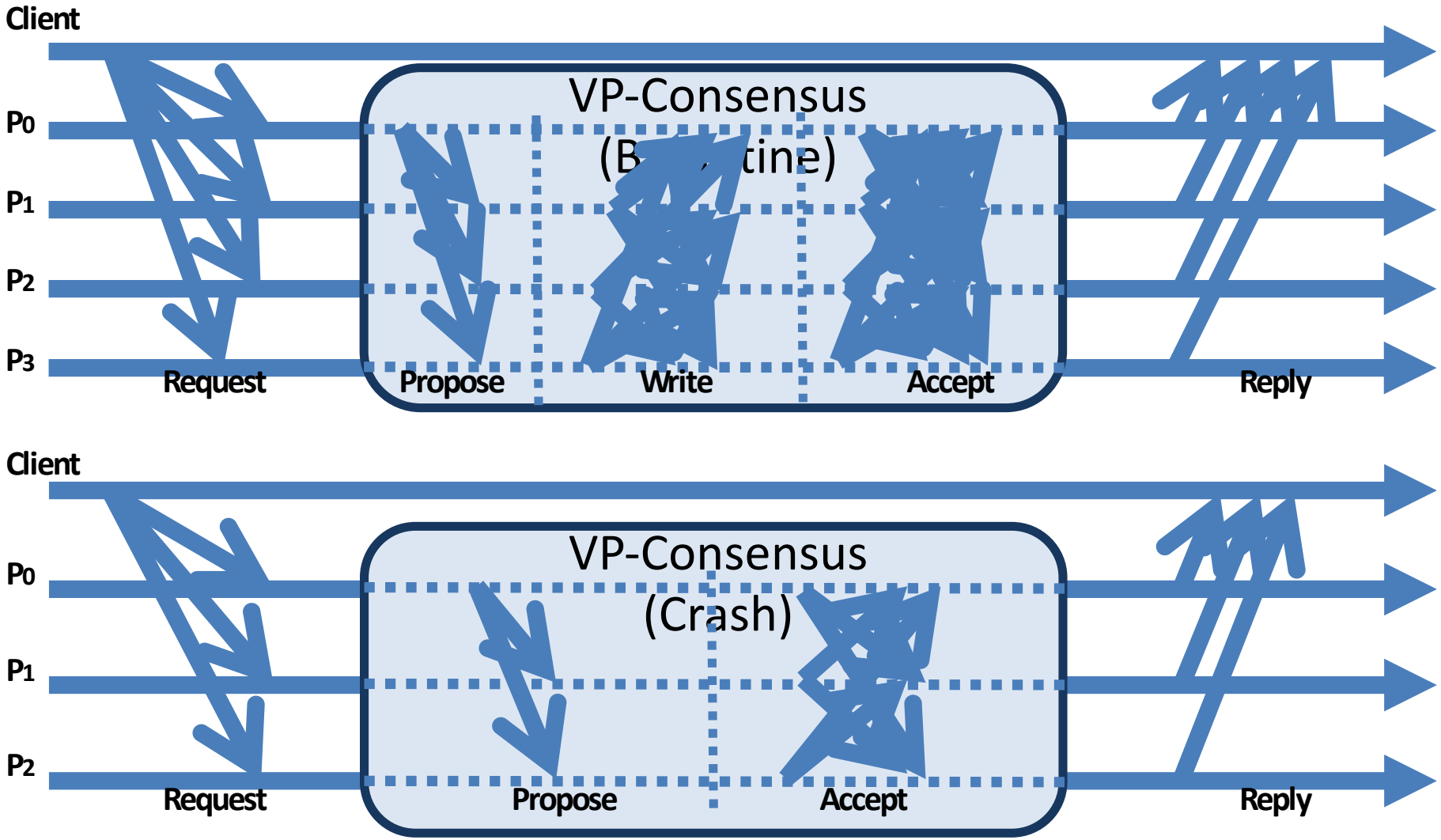
BFT-SMaRt

http://bft-smart.github.io/library/ [Bessani et al. DSN'14]

- Byzantine Fault tolerant state machine replication library written in Java (under development since 2010)
- Tolerates either crash (*2f+1* replicas) or Byzantine faults (*3f+1* replicas), under a partially synchronous system model
- Available under Apache license



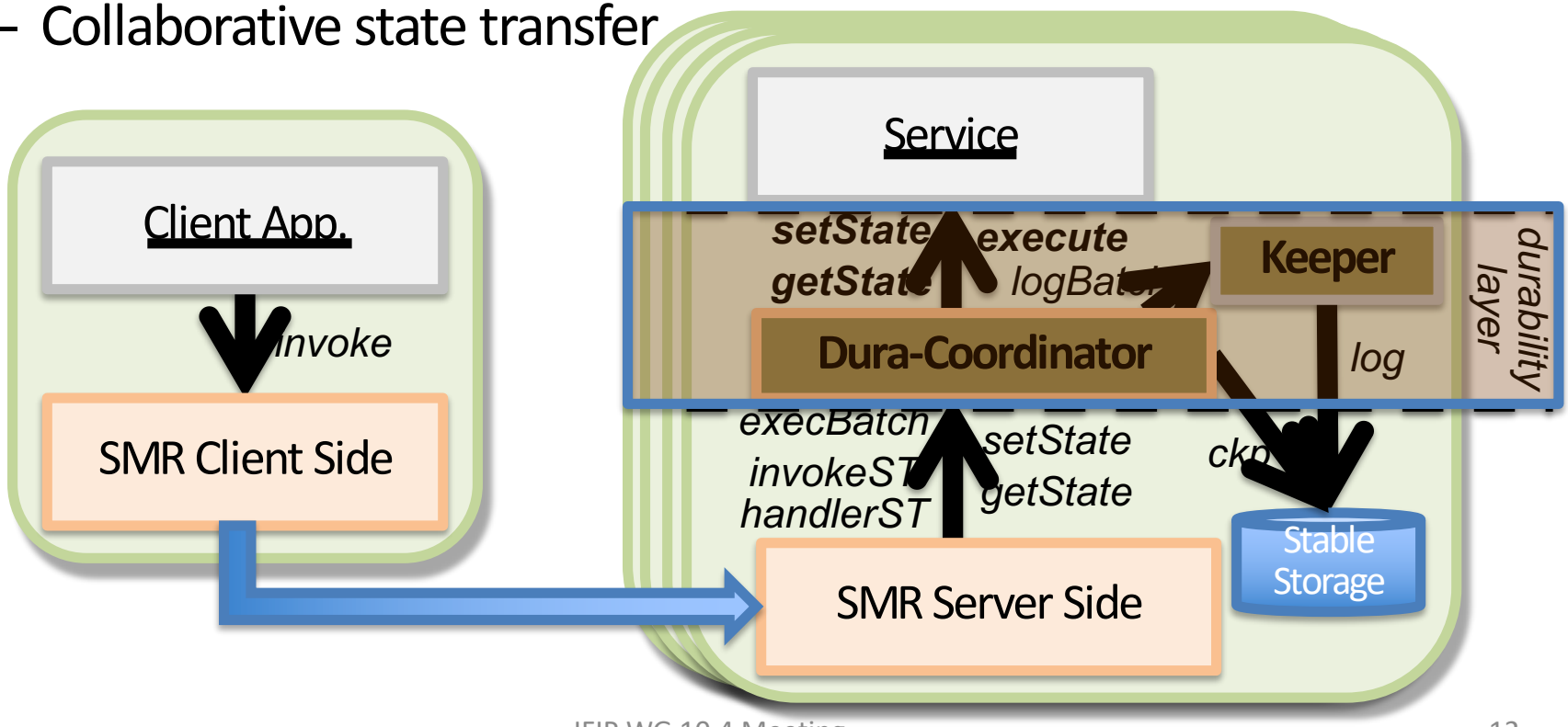
BFT-SMaRt Ordering Protocols



Durability in BFT-SMaRt

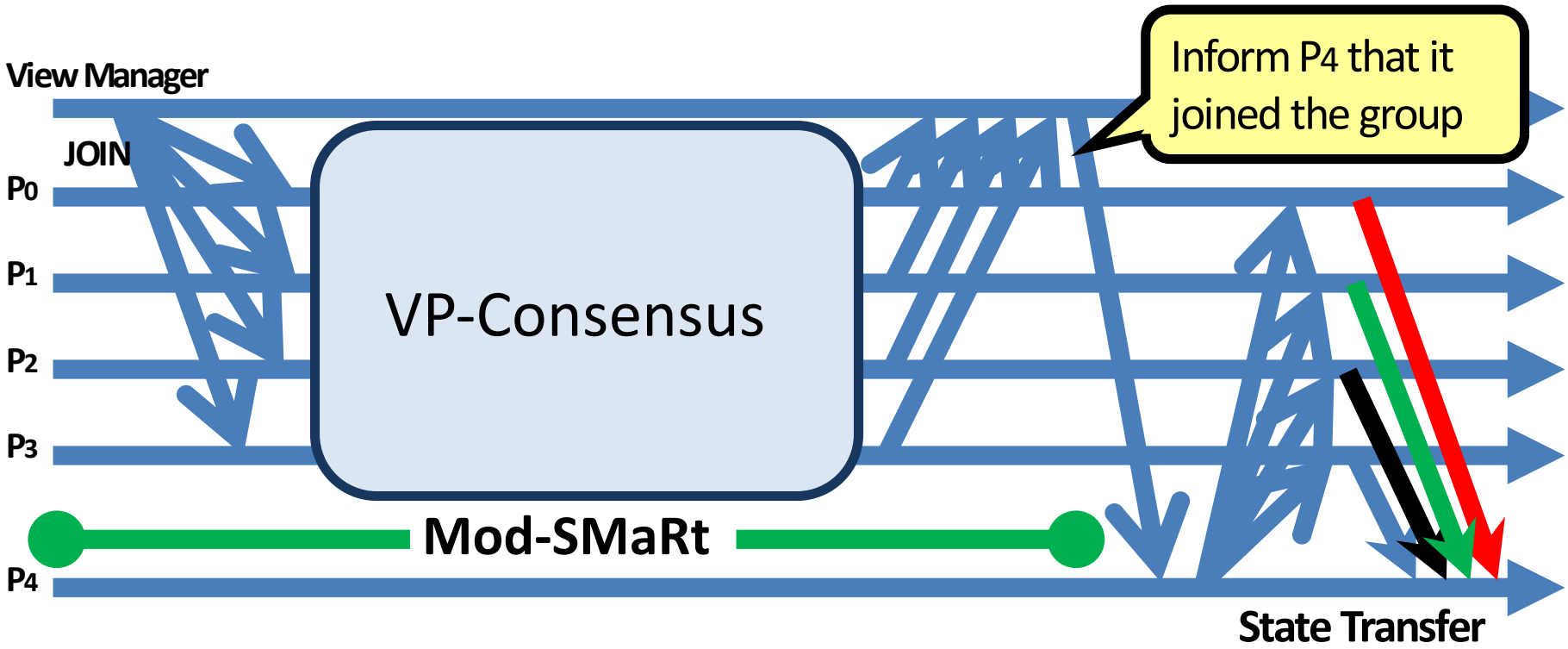
[Bessani et al. USENIX ATC'13]

- Techniques for efficient durability
  - Parallel Logging
  - Sequential checkpoints
  - Collaborative state transfer



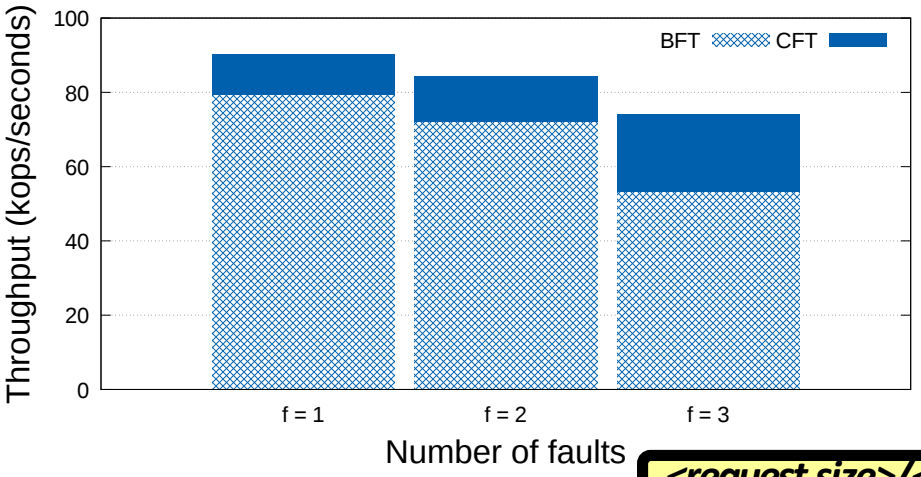
BFT-SMaRt Reconfiguration

Initiated by the View Manager - a trusted client used by system administrators that adds/removes replicas

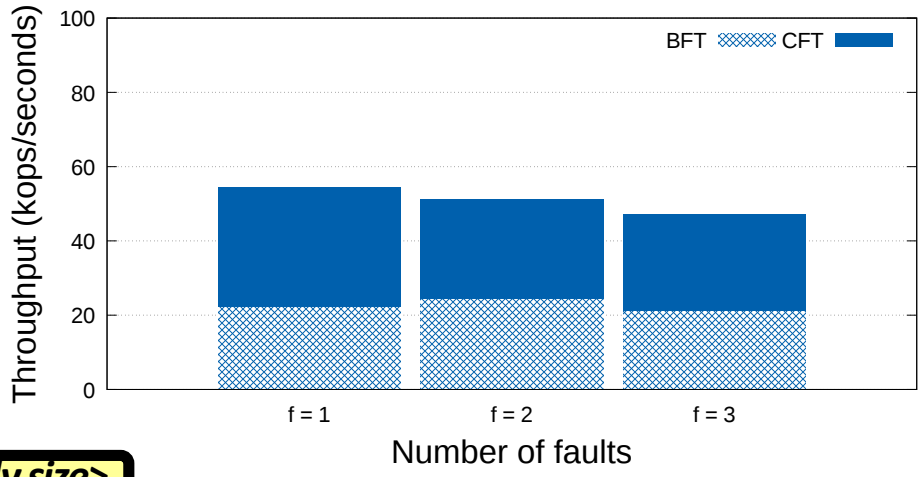


BFT-SMaRt Performance

(gigabit Ethernet, no disks)

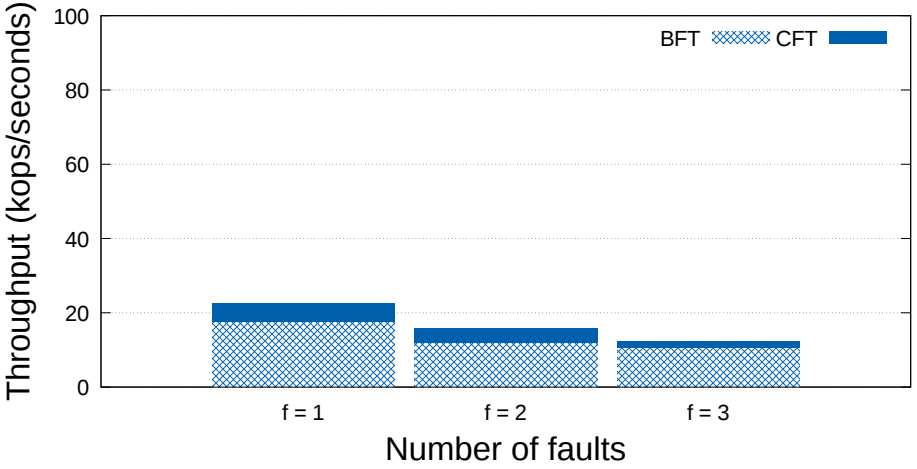


(a) 0/0

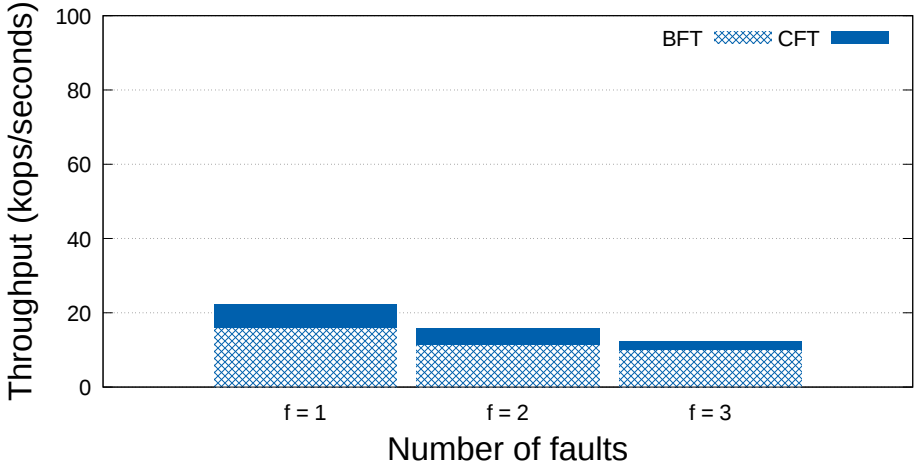


(b) 0/1024

<request size>/<reply size>

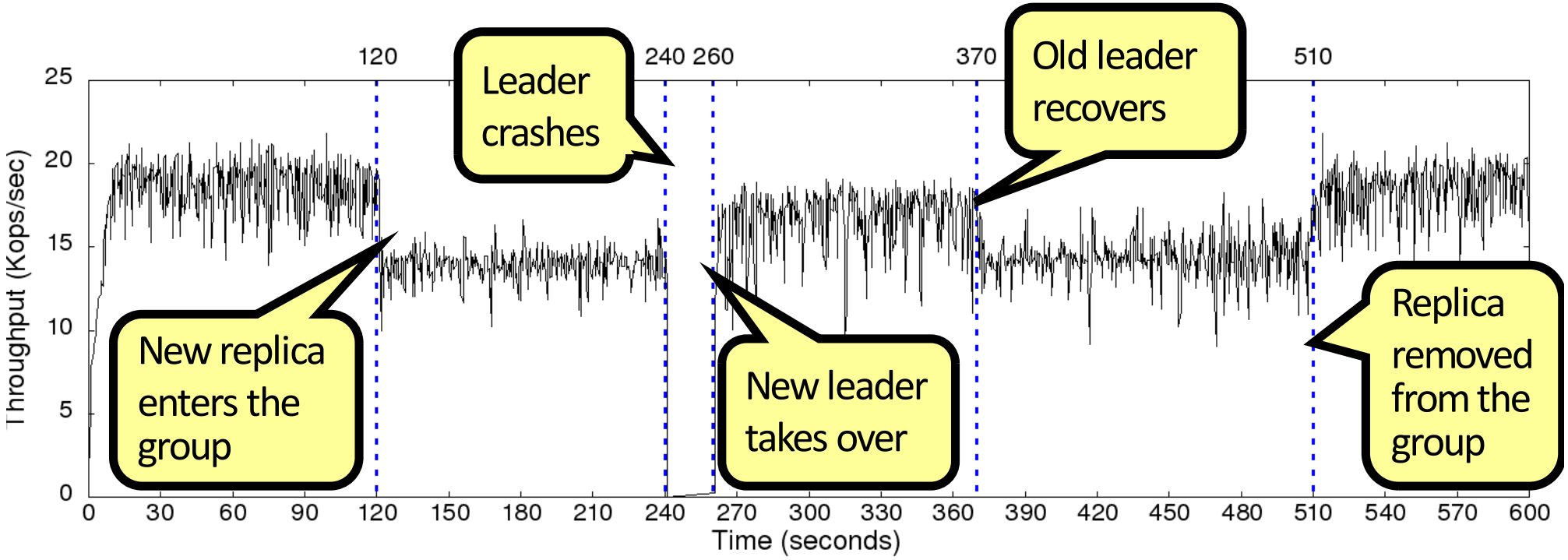


(c) 1024/0



(d) 1024/1024

Performance under “sporadic” events



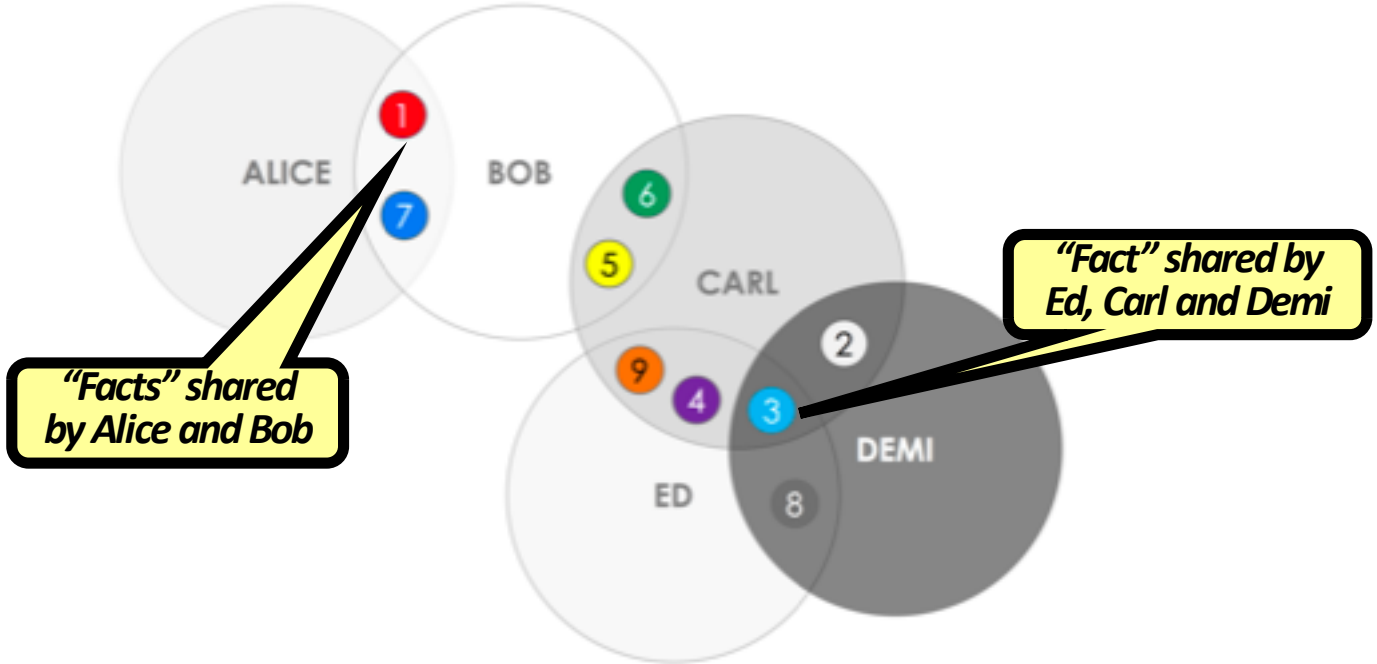
Part 1

BFT-SMART IN PERMISSIONED LEDGERS

- Startup from NY with 40+ people
- Technology
  - Smart contracts on top of state machine replication
  - BFT-SMaRt ported to Go
- Our involvement
  - Never saw the code
  - We talked a lot about collaboration, but just helped them understand the code and debug the synchronization phase of the protocol

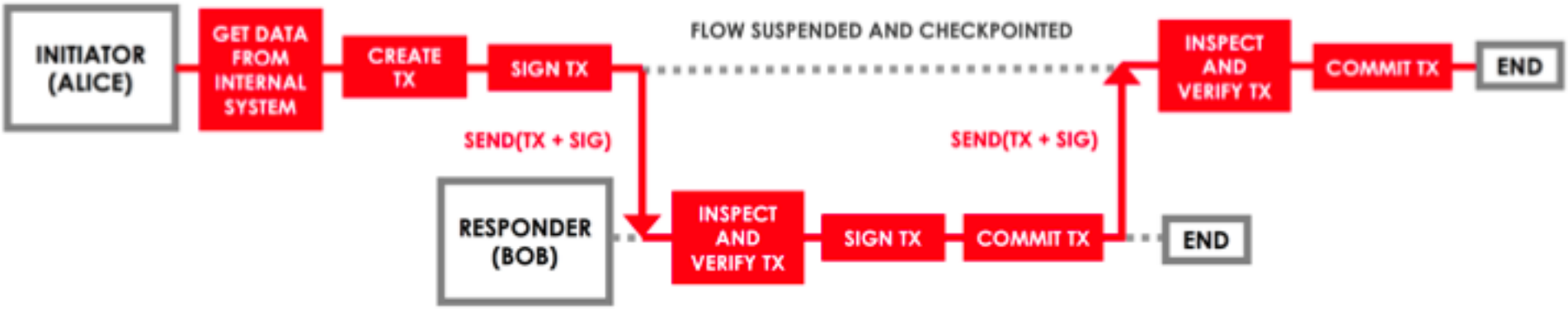
c·rda

- Open-source blockchain project targeting (at least initially) the financial market
- Key idea: **there is no shared global ledger**
  - Instead, **there are many distributed ledgers**



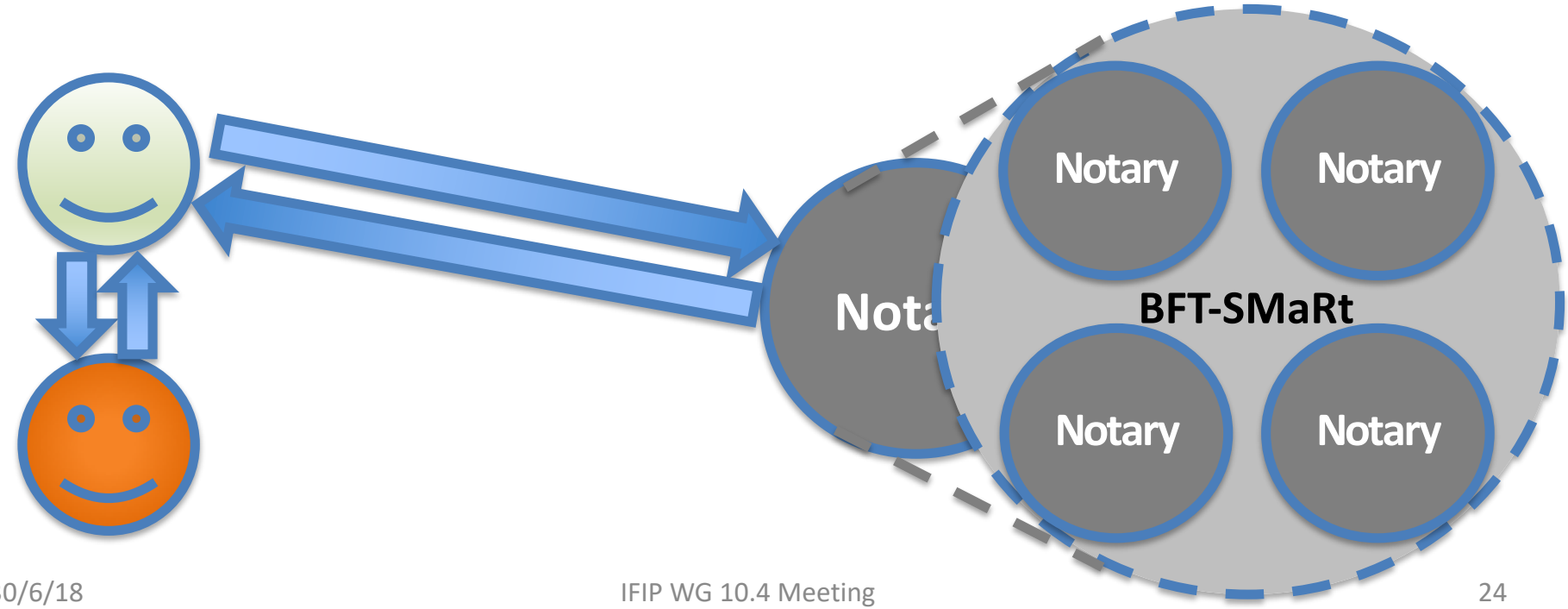
corda

- Only involved participants have to *execute* and *validate* the transaction
- A transaction is *committed* only if it achieve
  - **Validity consensus**: all involved participants need to validate and sign the transaction
  - **Uniqueness consensus**: requires a notary service



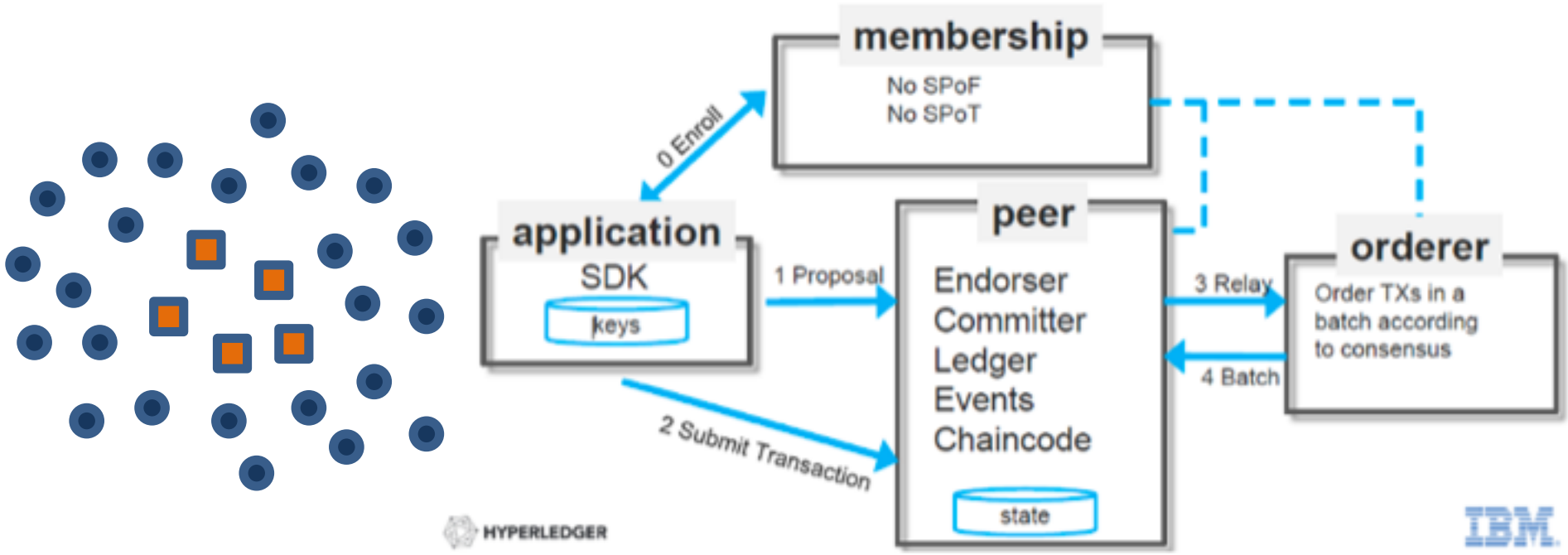
corda

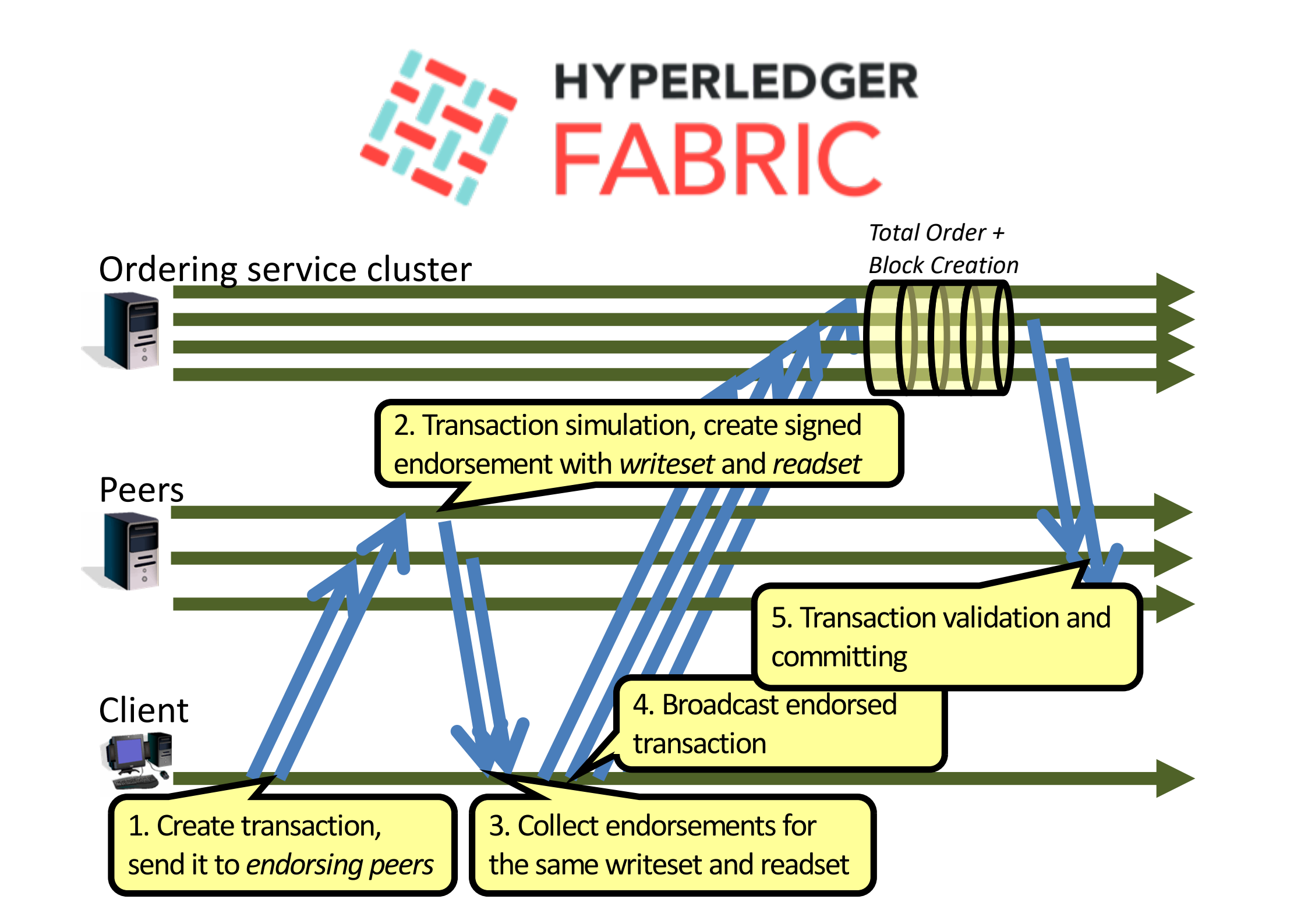
- Notary implements an insert-only key-value store that register all state “consumptions”
- Some specific transaction validation might be executed
- Multiple notaries might be used



HYPERLEDGER FABRIC

- Open-source, modular, permissioned [EuroSys'18]
- Architecture: not all "peers" are equal

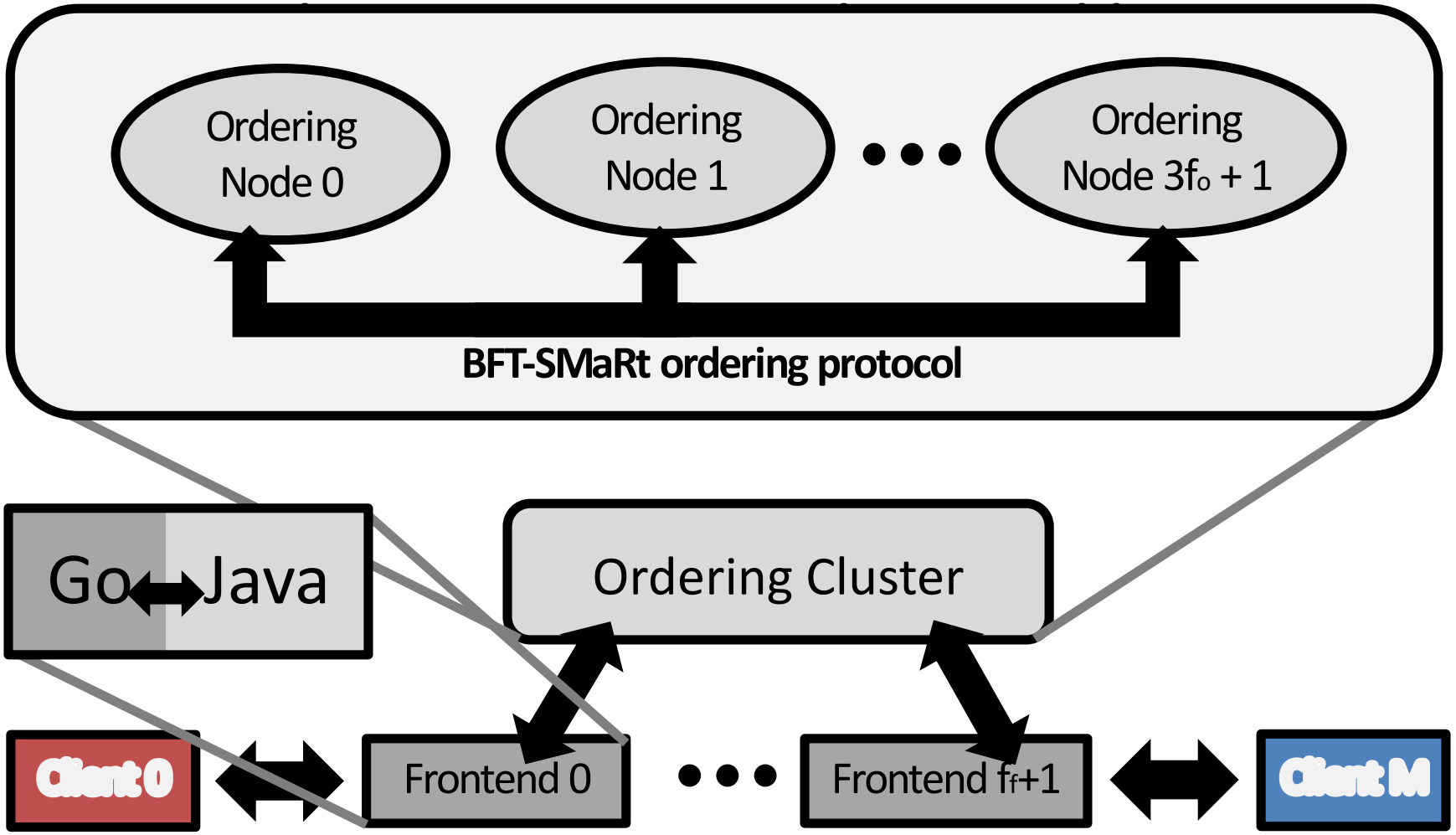



HYPERLEDGER
FABRIC
Ordering service cluster
Total Order +
Block Creation
Peers
Client
2. Transaction simulation, create signed endorsement with *writeset* and *readset*
5. Transaction validation and committing
4. Broadcast endorsed transaction
1. Create transaction, send it to *endorsing peers*
3. Collect endorsements for the same writeset and readset

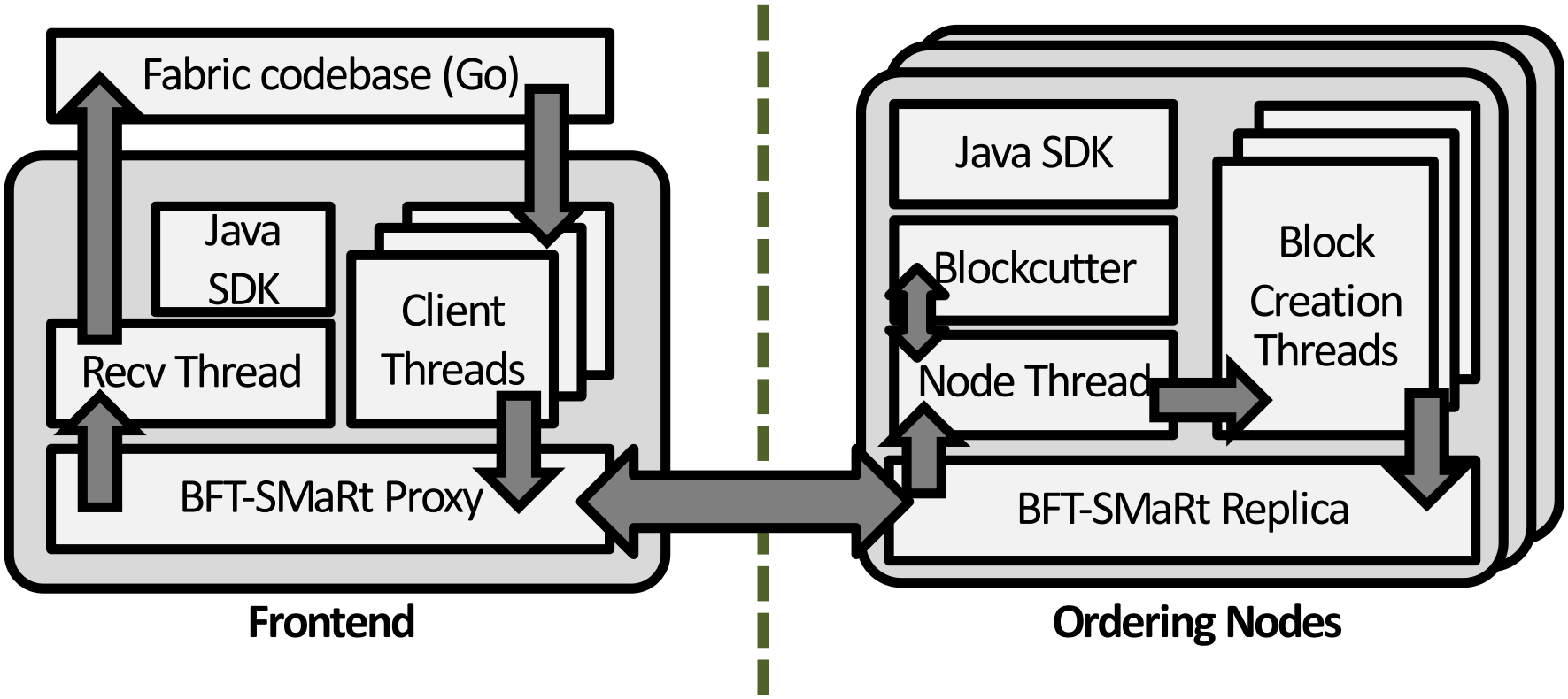
- Fabric supports different ordering services modules for different types of consensus
- Current release (v1.2.0) provides two:
  - Centralized module (***Solo***)
  - Apache Kafka-based module (***Kafka***)
- No module for Byzantine consensus

BFT-SMaRt Ordering Service

[Sousa et al, DSN'18]



BFT-SMaRt Ordering Service

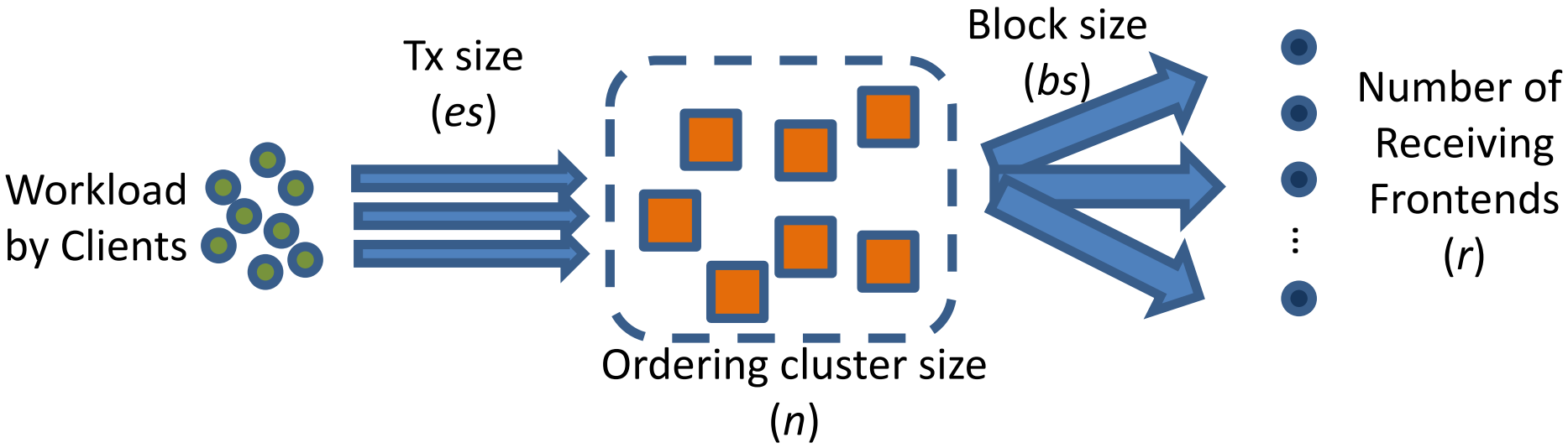


BFT-SMaRt Ordering Service

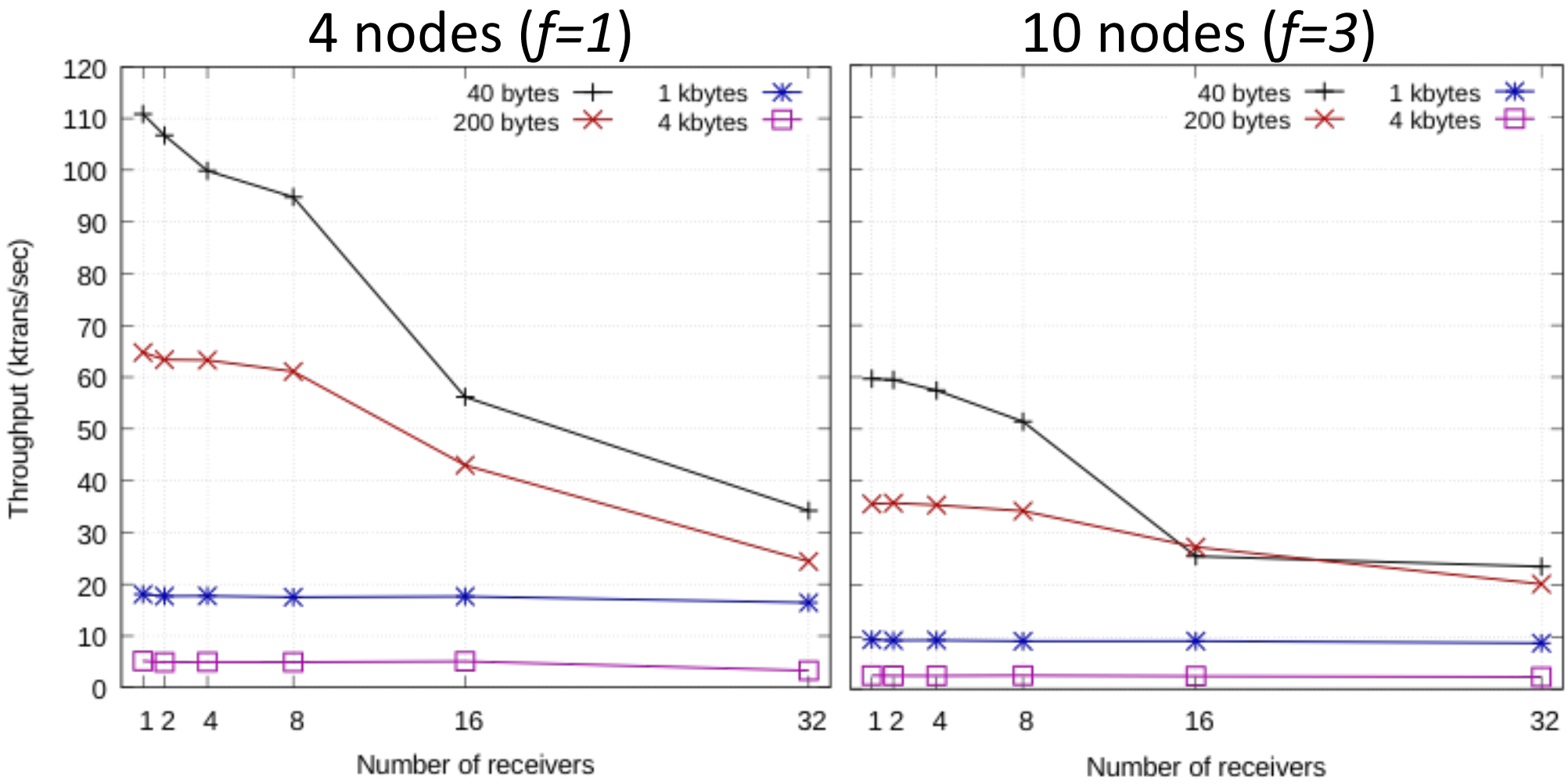
- Node state (to be persisted and transferred):
  - the ordered transactions still in the blockcutter,
  - header of the last generated block, and
  - latest configuration block
- Blocks can be validated and signed in parallel without incurring in non-determinism
- Frontends collect *2f+1* matching blocks signed from different ordering nodes

Evaluation

- Factors at play:



LAN Evaluation



Some takeaways

- (LAN) Even with blocks of 100 4kB-txs, 32 frontends and a cluster of 10 nodes, the service orders ~2200 txs/sec
  - This is considered a big network for Fabric
  - (illustrative) 2x more than Ethereum's theoretical peak of 100 txs/sec, and vastly superior to Bitcoin's 7 txs/sec
- (WAN) 5 sites in 4 continents can order 1kB-txs in < 400ms (w/ a load of 1000 txs/s)

Part 1

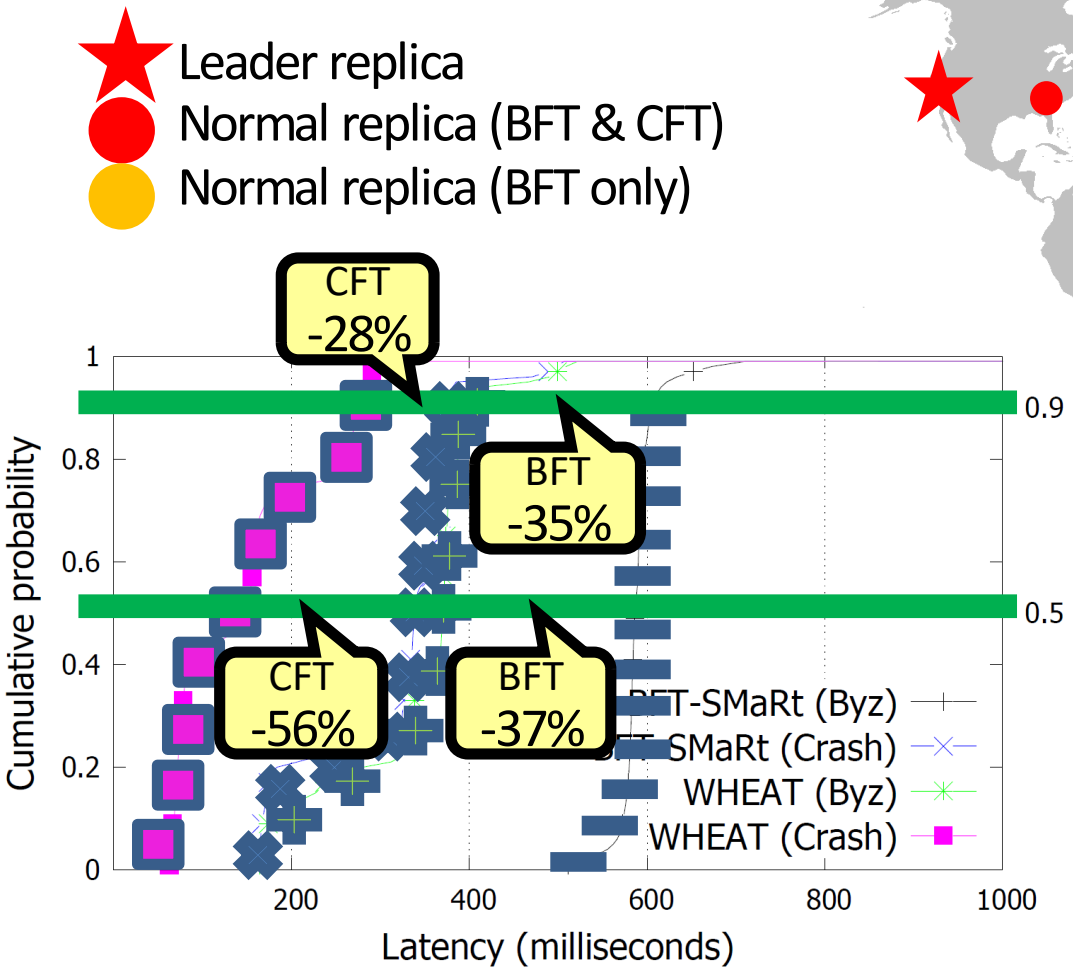
BEYOND BFT-SMART

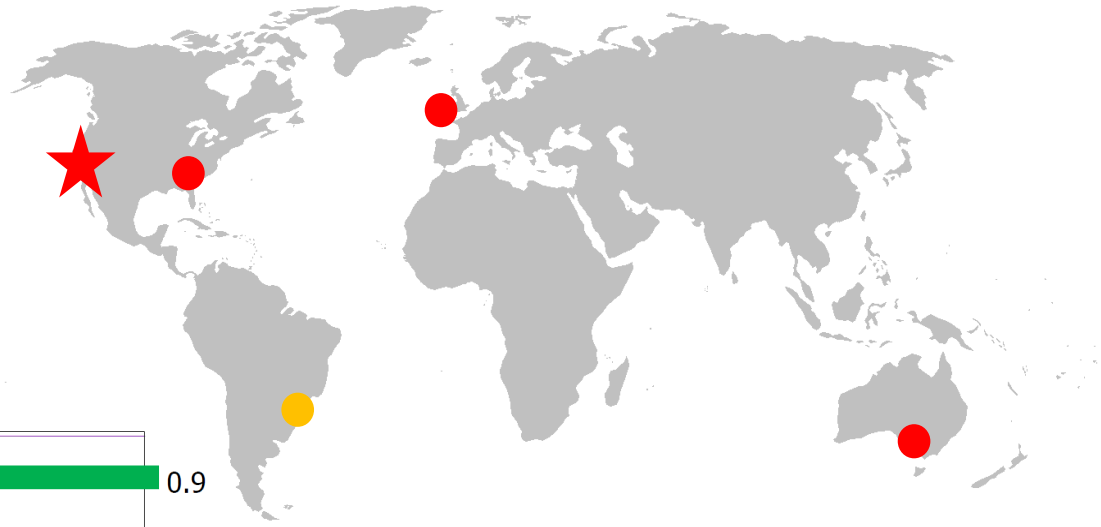
Our Research Agenda

- **Robust BFT replication library**
  - Maintain a good basic implementation
- **Geo-replication**
  - Key BFT application: distributed trust
- **Scalability & Elasticity**
  - Increase performance dynamically w/ additional replicas
- **Diversity and Fault Independence**
  - How to withstand *f* malicious faults
- **Design a simple blockchain "platform"**
  - How to go from BFT SMR to a Blockchain

Geo-Replicated State Machines

[Sousa & Bessani. SRDS'15]

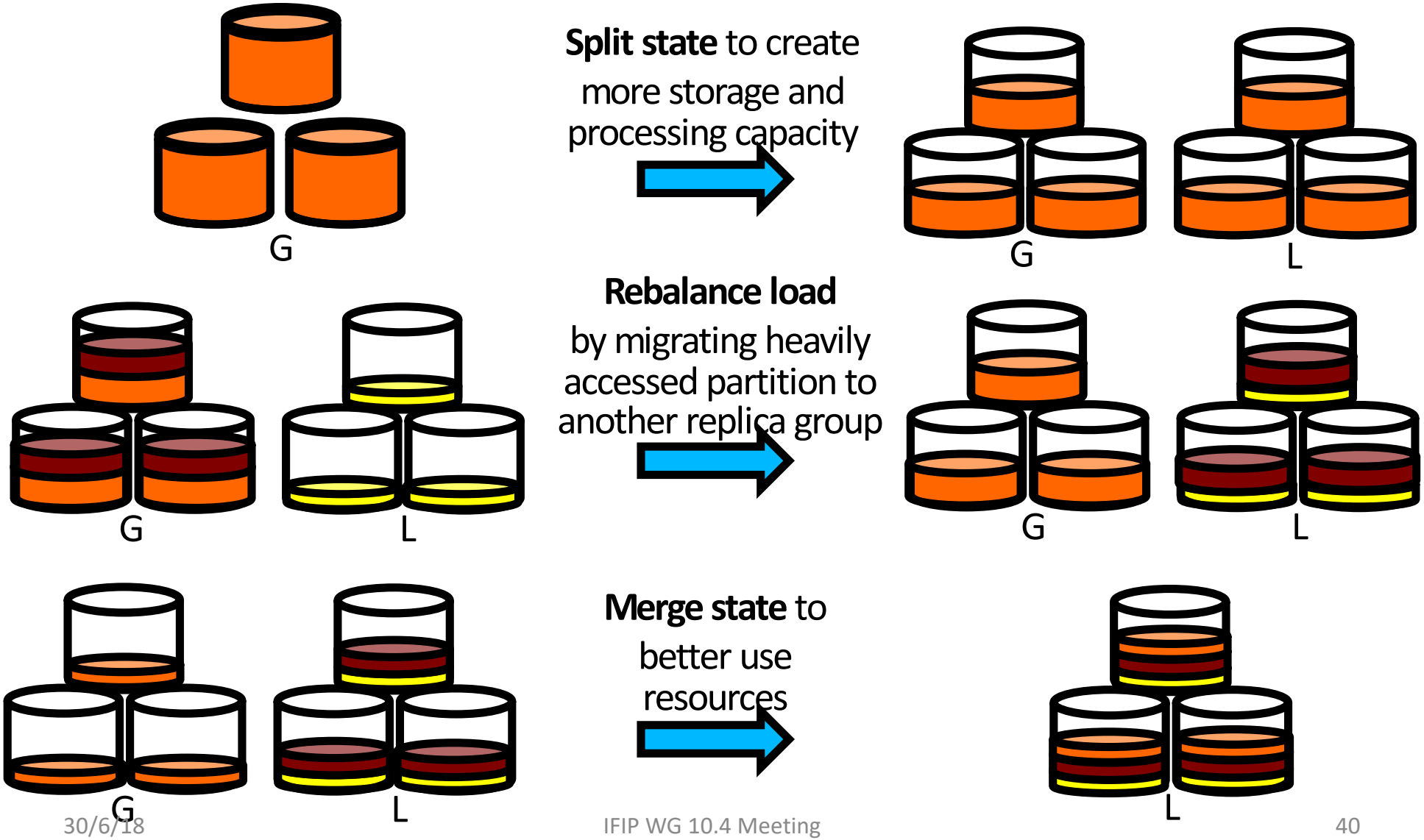




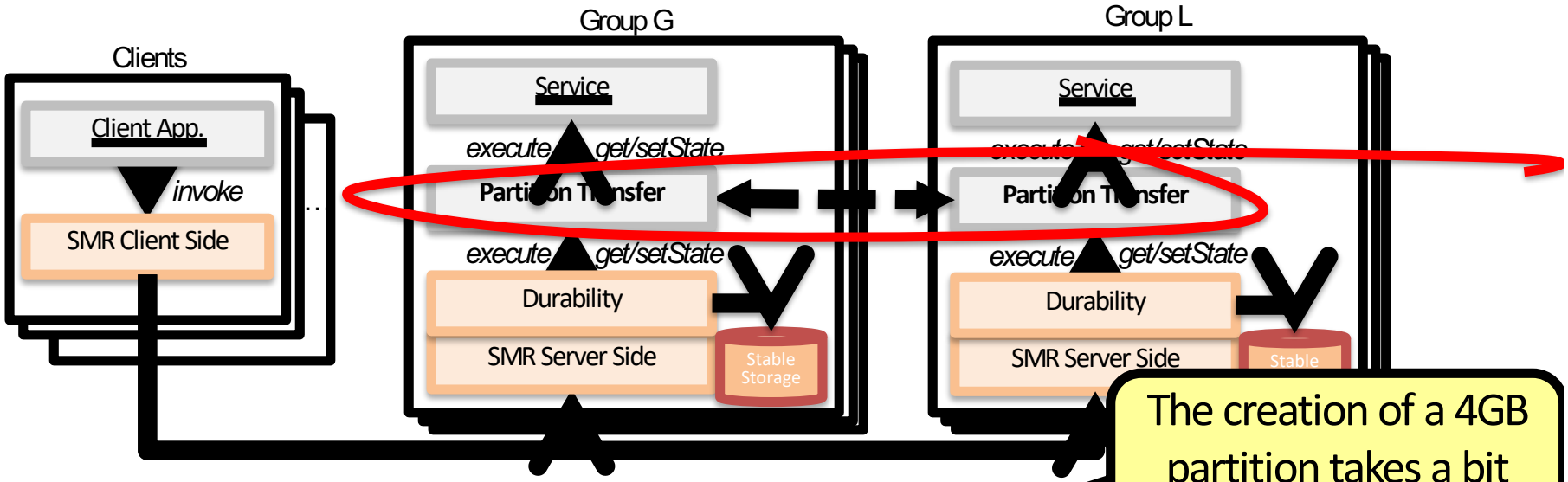
- Key techniques:
  - More replicas
  - Weighted replication
  - Tentative execution

Elastic State Machine Replication

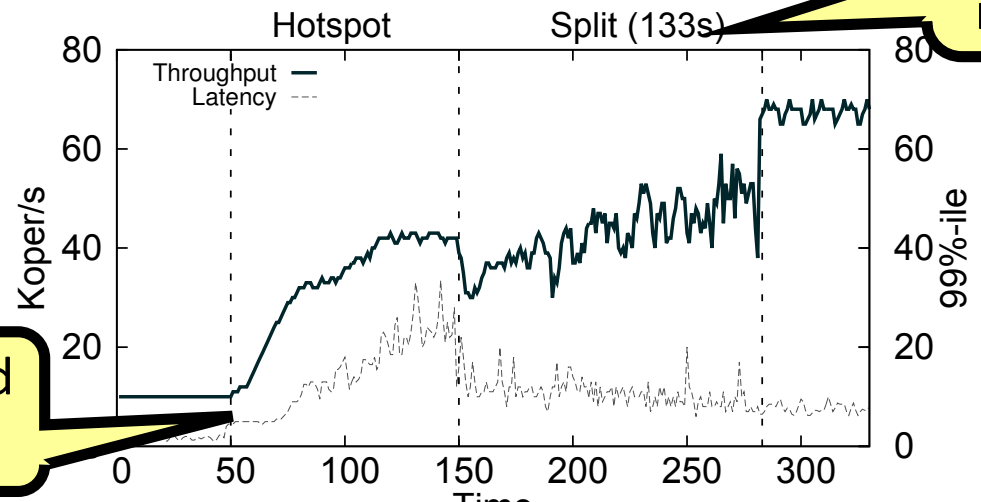
[Nogueira et al. IEEE TPDS'17]



Elastic State Machine Replication

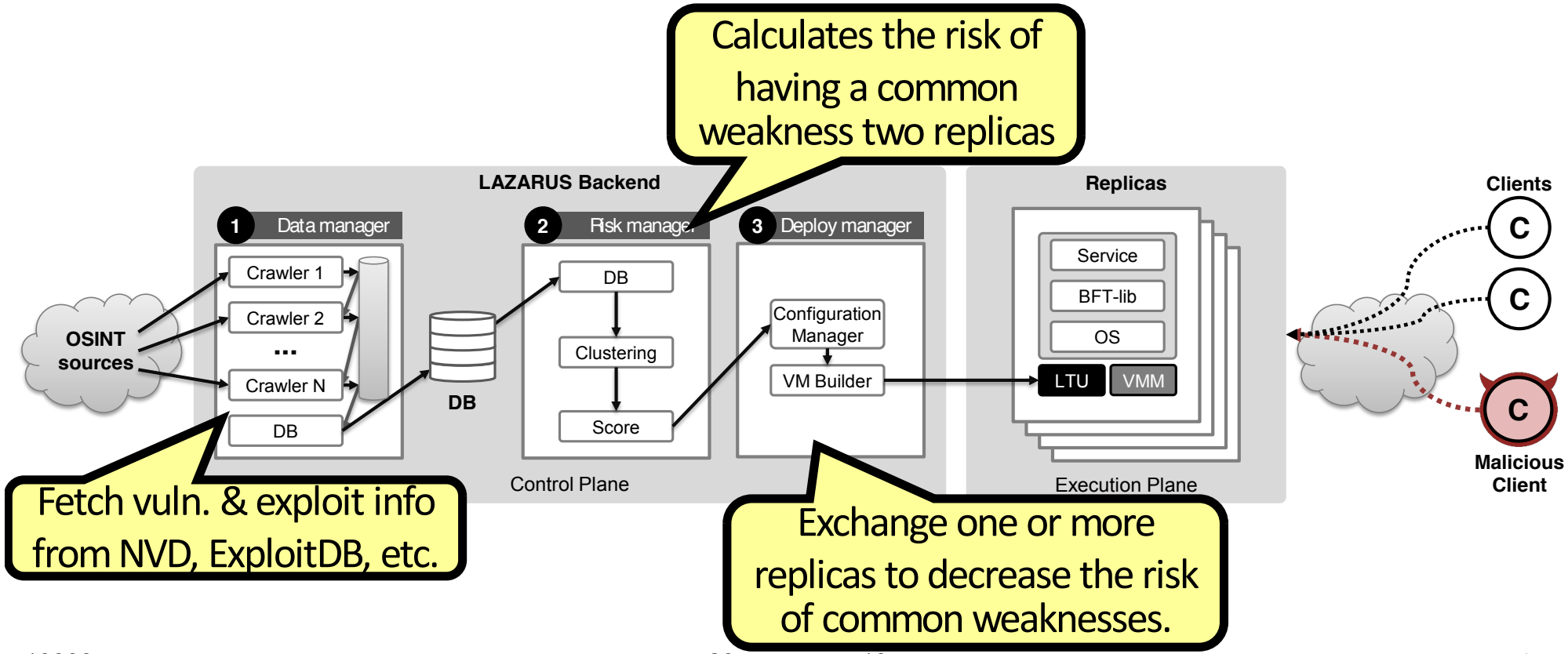


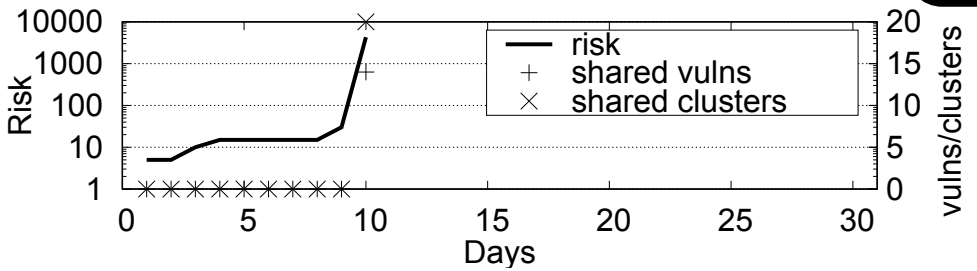
The creation of a 4GB partition takes a bit more than 2 minutes



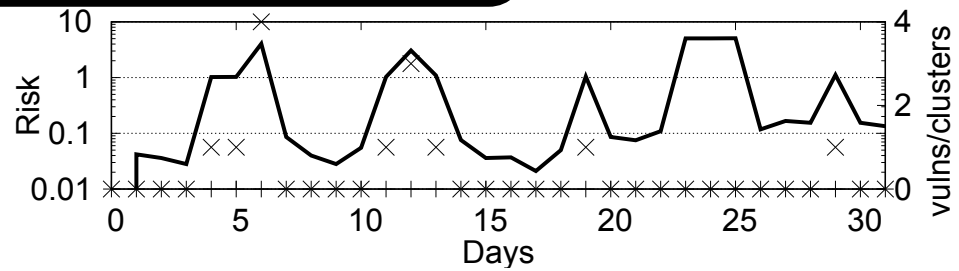
Hotspot starts! Workload increases 10x

Diversity Management



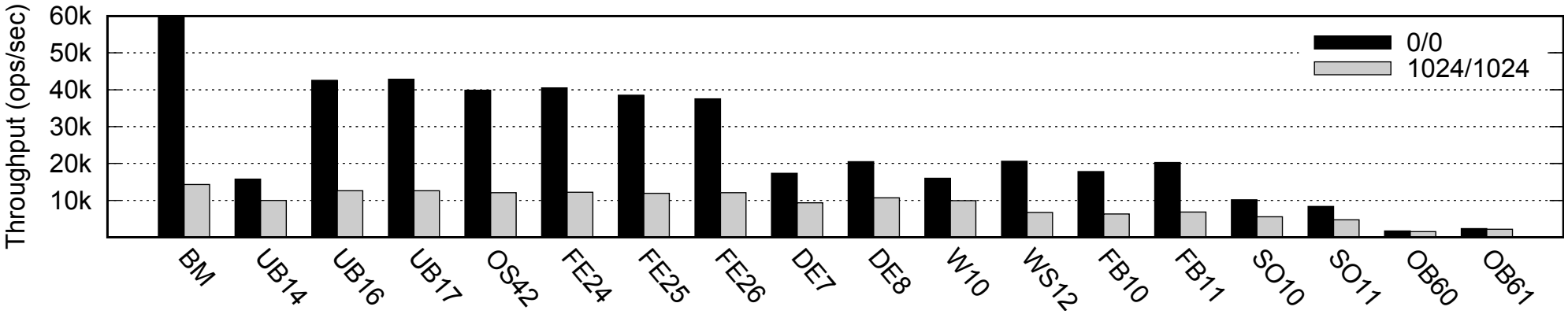


(a) Random

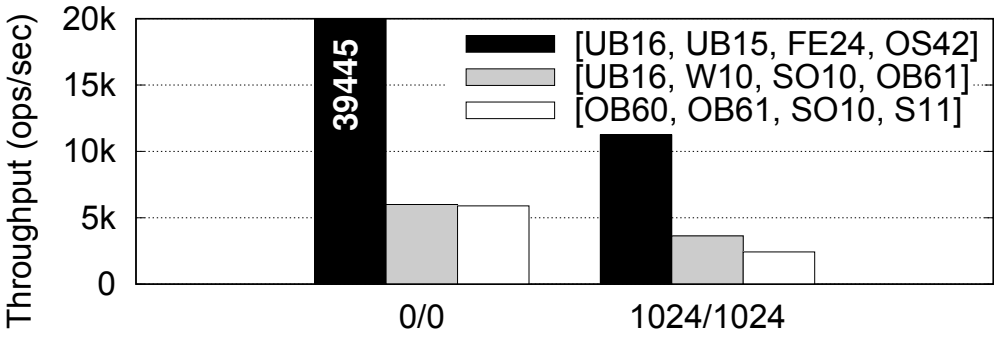


(b) LAZARUS

Diversity Management

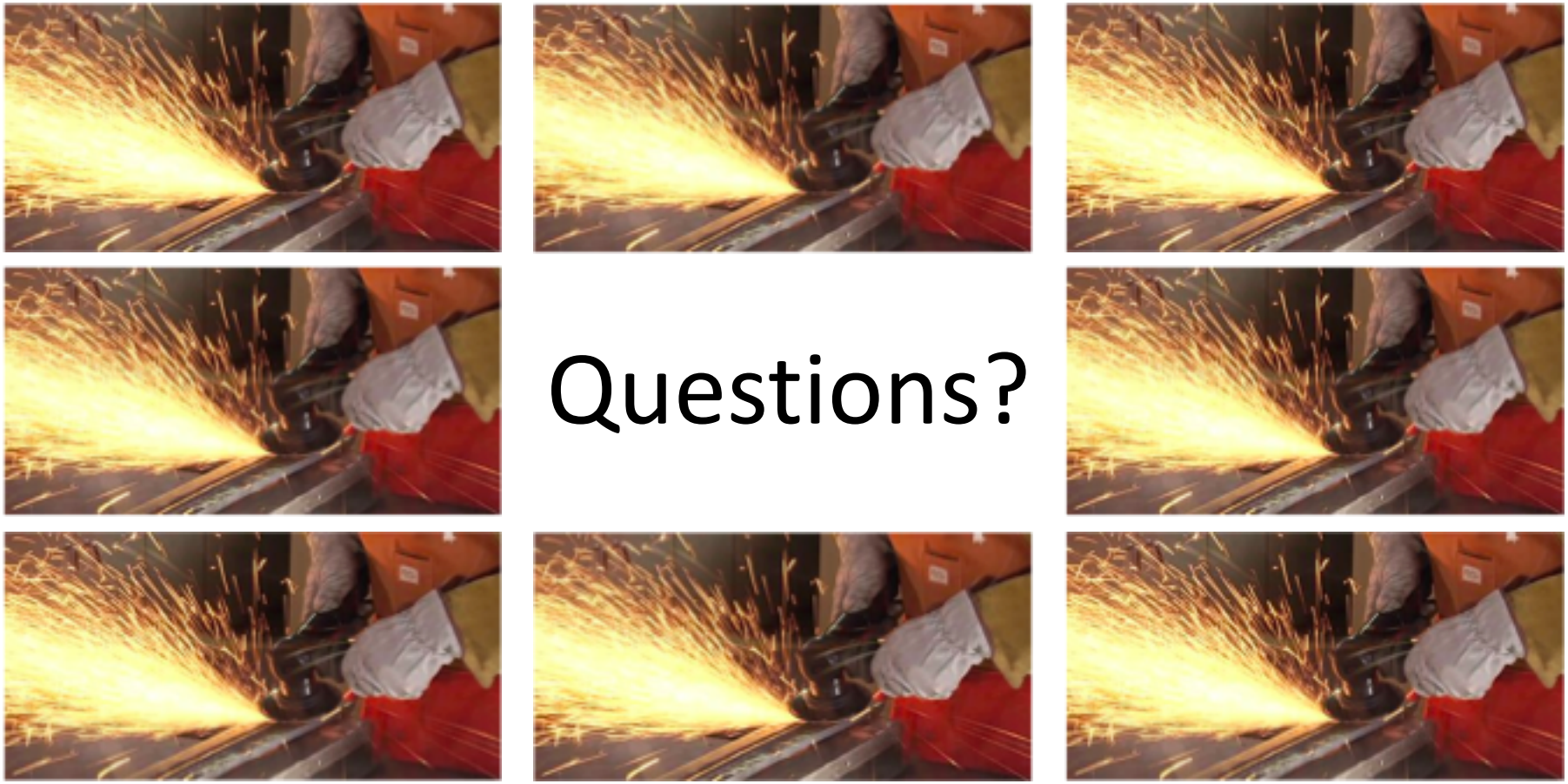


ID	Name	Cores	JVM	Mem.
UB14	Ubuntu 14.04	4	Java Oracle 1.8.0_144	15GB
UB16	Ubuntu 16.04	4	Java Oracle 1.8.0_144	15GB
UB17	Ubuntu 17.04	4	Java Oracle 1.8.0_144	15GB
OS42	OpenSuse 42.1	4	Openjdk 1.8.0_141	15GB
FE24	Fedora 24	4	Openjdk 1.8.0_141	15GB
FE25	Fedora 25	4	Openjdk 1.8.0_141	15GB
FE26	Fedora 26	4	Openjdk 1.8.0_141	15GB
DE7	Debian 7	4	Java Oracle 1.8.0_151	15GB
DE8	Debian 8	4	Openjdk 1.8.0_131	15GB
W10	Windows 10	4	Java Oracle 1.8.0_151	1GB
WS12	Windows Server 2012	4	Java Oracle 1.8.0_151	1GB
FB10	FreeBSD 10	4	Openjdk 1.8.0_144	15GB
FB11	FreeBSD 11	4	Openjdk 1.8.0_144	15GB
SO10	Solaris 10	1	Java Oracle 1.8.0_141	15GB
SO11	Solaris 11	1	Java Oracle 1.8.0_05	15GB
OB60	OpenBSD 6.0	1	Openjdk 1.8.0_72	1GB
OB61	OpenBSD 6.1	1	Openjdk 1.8.0_121	1GB



BFT-SMaRt as a Blockchain

- What to change?
  - Durable Logging -> Blockchain
  - State machine service-> smart contract
  - BFT reconfiguration -> Churn/committee management
  - VP-consensus -> Scalable VP-consensus



Questions?

- To know more:
  - BFT-SMaRt: http://bft-smart.github.io/library/
  - **Bessani et al.** ***State Machine Replication for the Masses with BFT-SMaRt.*** **IEEE/IFIP DSN'14.**
  - **Bessani et al.** ***On the Efficiency of Durable State Machine Replication.*** **USENIX ATC'13.**
  - **Sousa, Bessani.** ***Separating the WHEAT from the Chaff: An Empirical Design for Geo-replicated State Machines.*** **IEEE SRDS'15.**
  - **Sousa et al.** ***A Byzantine Fault-Tolerant Ordering Service for Hyperledger Fabric Blockchain Platform.*** **IEEE/IFIP DSN'18.**