Aplicações de Ledger Distribuída e os Mecanismos de Base para o Consenso

Allan Edgard Silva Freitas

Professor Titular (IFBA)

allan@ifba.edu.br

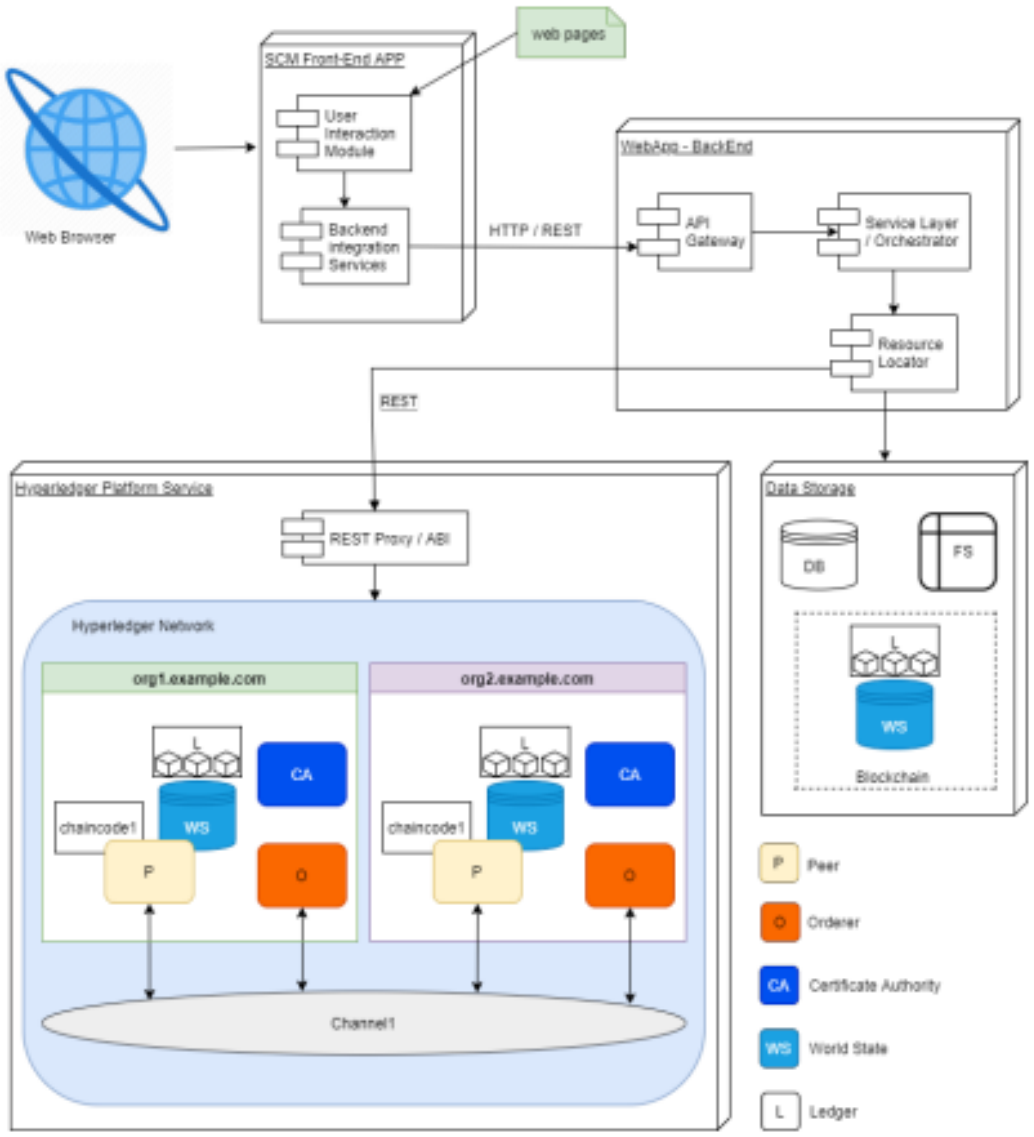
INSTITUTO FEDERAL
Bahia

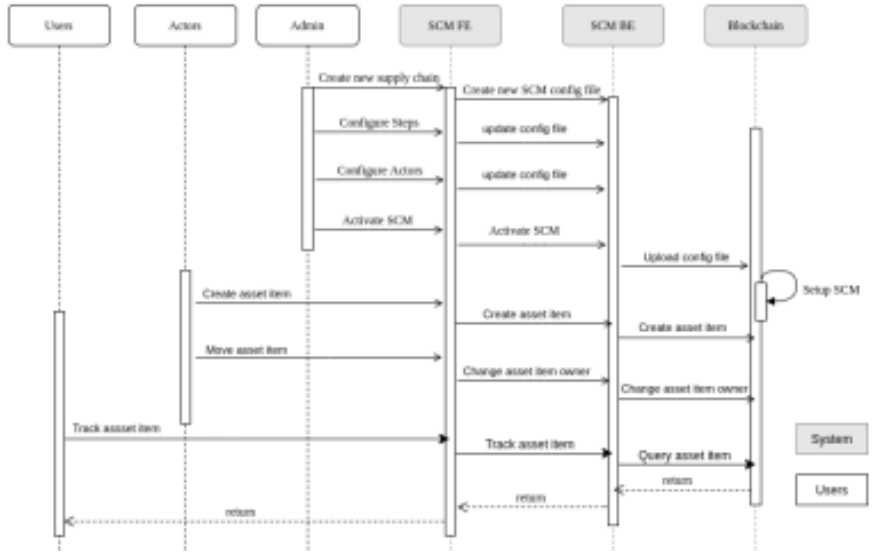
PROGRAMA DE
PÓS-GRADUAÇÃO EM
ENGENHARIA DE SISTEMAS E PRODUTOS

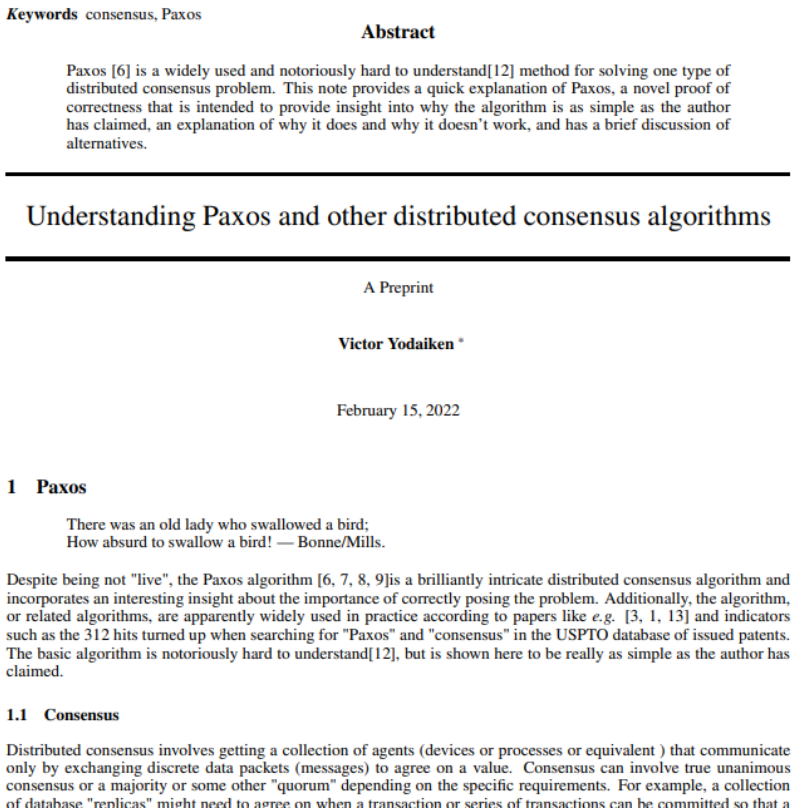
Aplicações de Ledger Distribuída

- Uso em rastreabilidade de commodities (framework desenvolvido)
- Uso em eleições (trabalho em curso, combinando técnicas de segurança e aliando auditoria com transparência x confidencialidade)

Gerenciamento de commodities por meio de um arcabouço desenvolvido sobre o Hyperledger Fabric







Keywords consensus, Paxos

Abstract

Paxos [6] is a widely used and notoriously hard to understand[12] method for solving one type of distributed consensus problem. This note provides a quick explanation of Paxos, a novel proof of correctness that is intended to provide insight into why the algorithm is as simple as the author has claimed, an explanation of why it does and why it doesn't work, and has a brief discussion of alternatives.

Understanding Paxos and other distributed consensus algorithms

A Preprint

Victor Yodaiken *

February 15, 2022

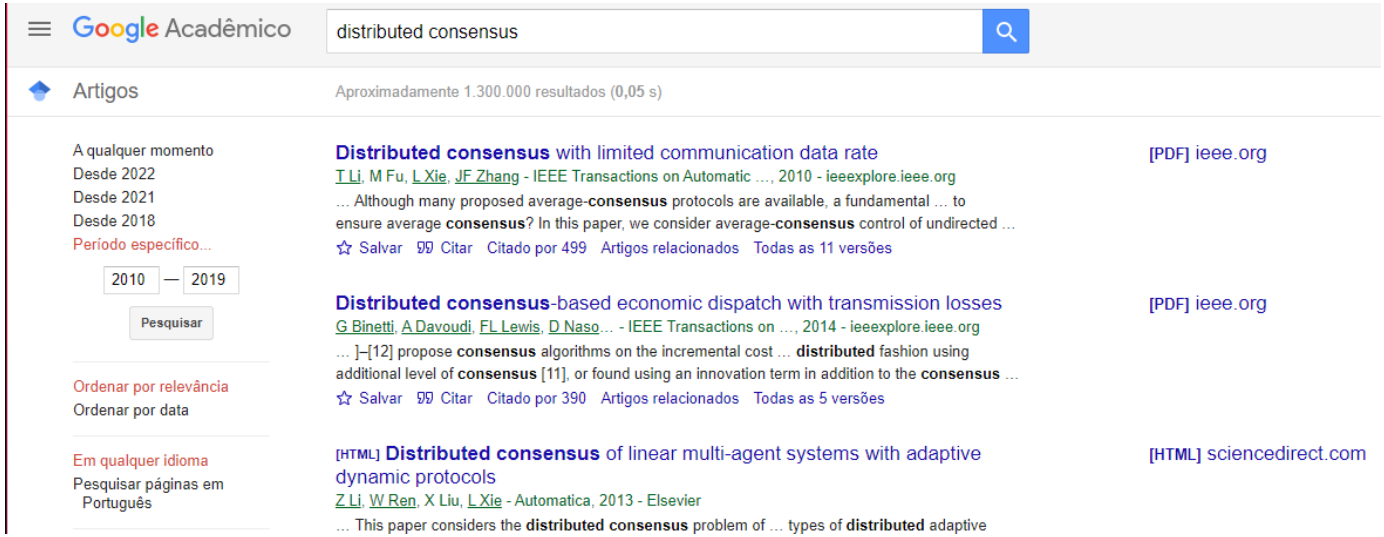
1 Paxos

There was an old lady who swallowed a bird;
How absurd to swallow a bird! — Bonne/Mills.

Despite being not "live", the Paxos algorithm [6, 7, 8, 9]is a brilliantly intricate distributed consensus algorithm and incorporates an interesting insight about the importance of correctly posing the problem. Additionally, the algorithm, or related algorithms, are apparently widely used in practice according to papers like *e.g.* [3, 1, 13] and indicators such as the 312 hits turned up when searching for "Paxos" and "consensus" in the USPTO database of issued patents. The basic algorithm is notoriously hard to understand[12], but is shown here to be really as simple as the author has claimed.

1.1 Consensus

Distributed consensus involves getting a collection of agents (devices or processes or equivalent) that communicate only by exchanging discrete data packets (messages) to agree on a value. Consensus can involve true unanimous consensus or a majority or some other "quorum" depending on the specific requirements. For example, a collection of database "replicas" might need to agree on when a transaction or series of transactions can be committed so that a

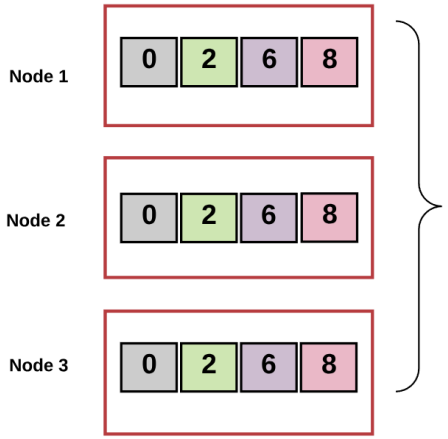


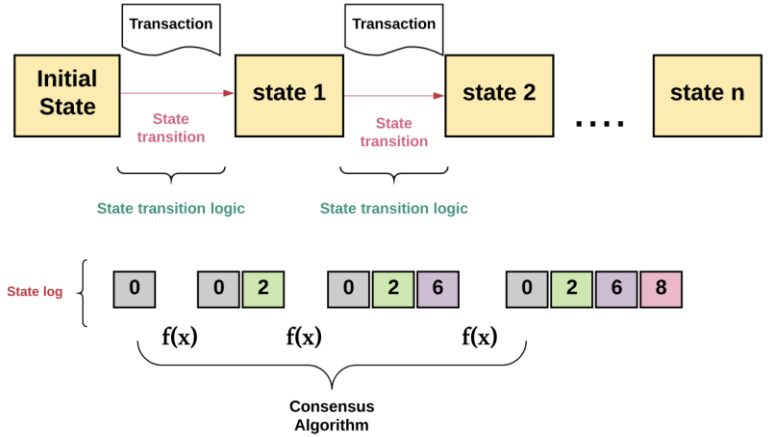
Sobre os mecanismos de base do consenso

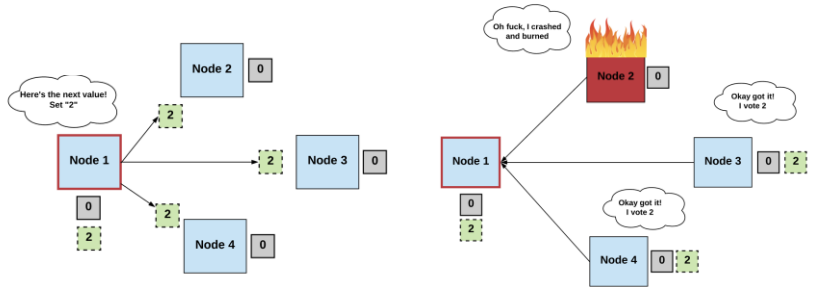
- Uma pesquisa simples no Google Scholar indica que um aumento significativo de trabalhos que falem de *distributed consensus* nos últimos anos
- - Anos 90: 111K
- - Anos 00: 339K
- - Anos 10: 1.300K
- - Anos 20: 134K em 2 anos e pouco...

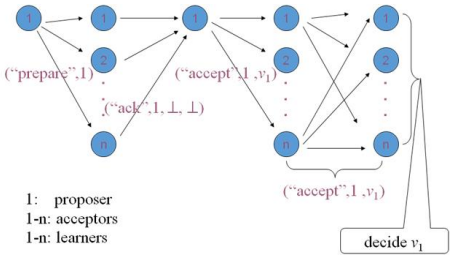
Consenso é a base de várias aplicações de computação distribuída

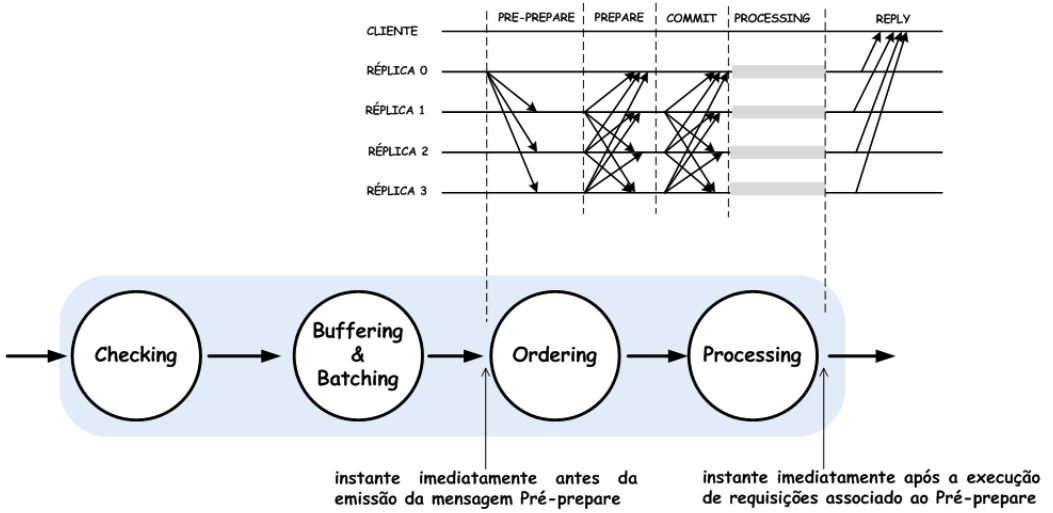
- Depende do modelo de sistema: de síncrono ao assíncrono, passando por diversas possibilidades de parcialmente síncrono...
- Depende do modelo de falhas, em geral, CFT (crash fault-tolerant) ou BFT (bizantine fault-tolerant)...





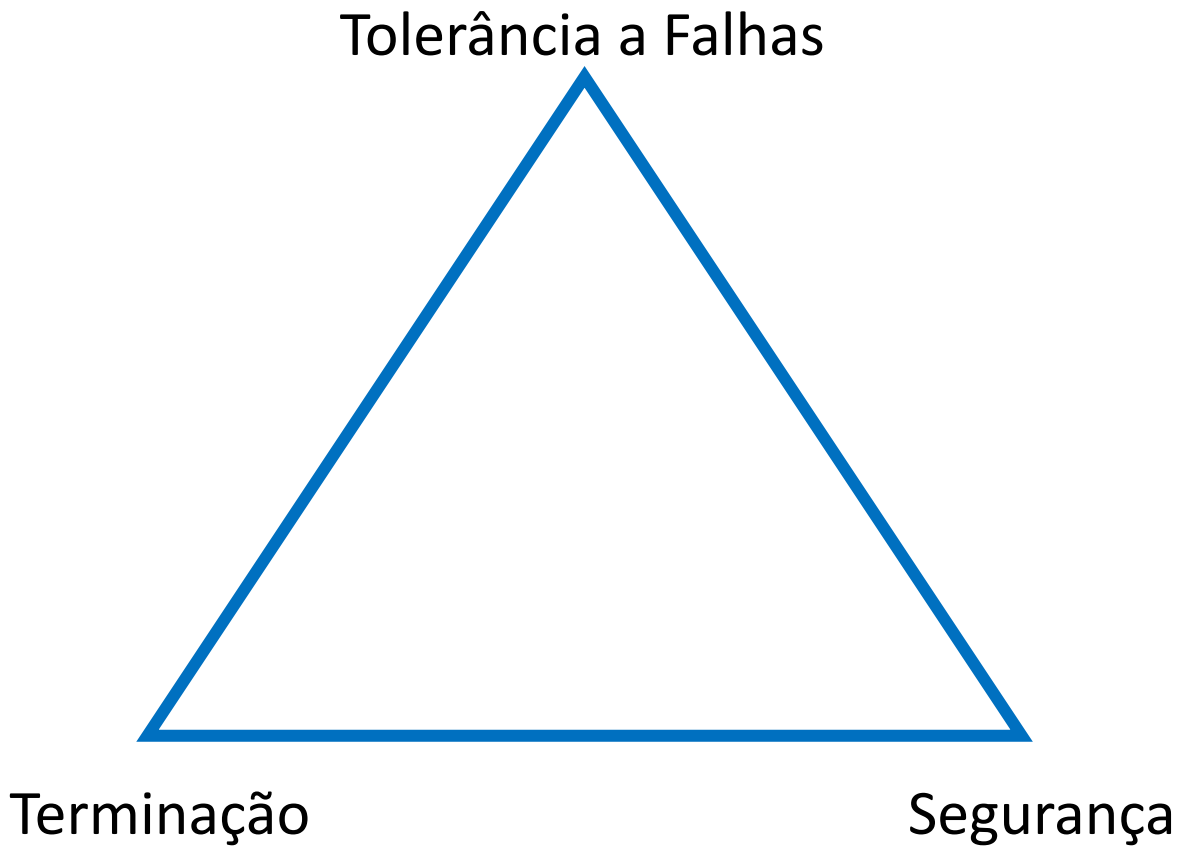






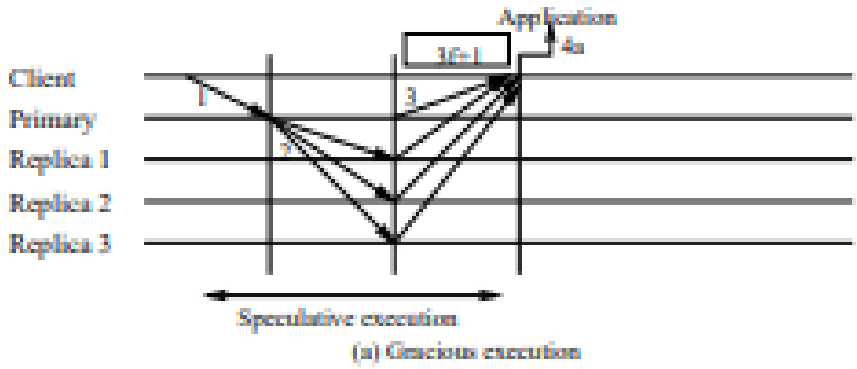
Relembrem FLP!

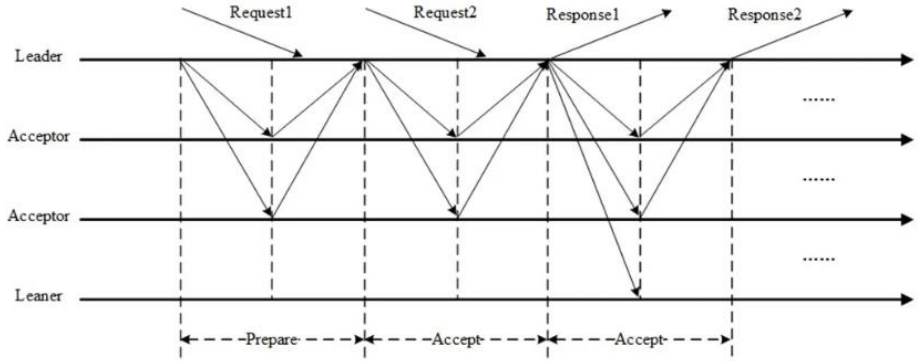
- Falhas por crash não são percebidas em sistemas puramente assíncronos
- Limites temporais são necessários para terminação do consenso na presença de falhas...



Refletindo sobre a natureza do consenso

- Consenso, em especial com **acordo uniforme**, requer uma quantidade considerável de mensagens
- Mesmo em CFT, é uma primitiva custosa
- Consenso BFT é ainda um pouco mais custoso
- Técnicas especulativa/otimistas promovem variações no consenso CFT (caso de Multi Paxos) ou no BFT (caso de Zyzzyva), mas precisamos sempre que algo bom aconteça!





O que seria algo bom?

- Paxos, RAFT, PBFT etc. dependem de um líder, por exemplo:
  - Fase 1 do Paxos é a tomada da liderança
  - RAFT tem uma eleição de líder clara
  - PBFT tem uma rotação do líder/primário

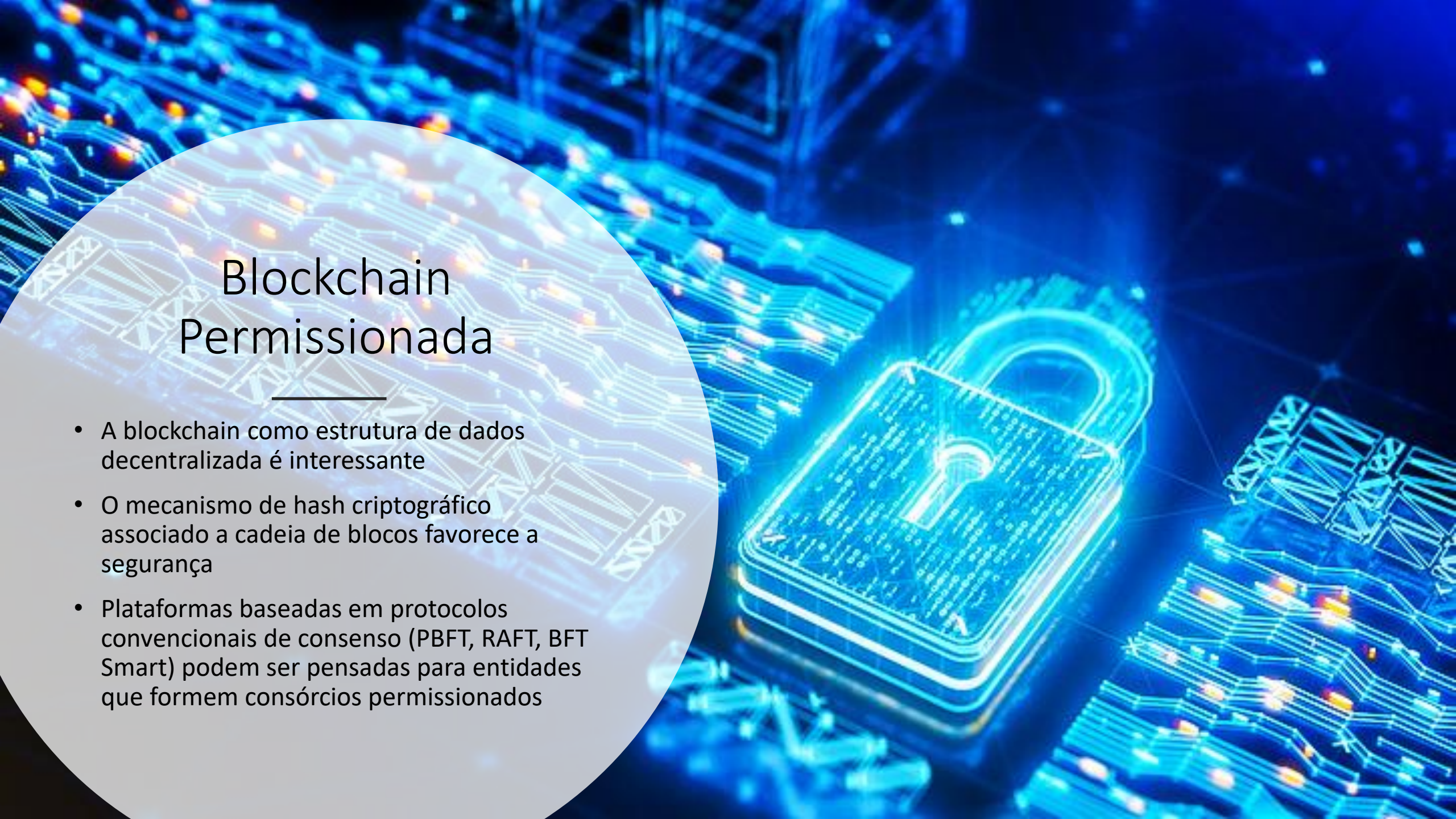
Abordagens especulativas/otimistas assumem que não há disputa de poder para tomar vantagem nas rodadas

Das eleições

- Em geral, a eleição (e as demais fases) destes algoritmos é baseada em quóruns
- Manter quóruns de forma adequada pode requerer blocos como *membership* e detecção de defeitos...

Das possibilidades

- A eleição silenciosa permite uma decentralização muito maior sem haver gerenciamento de membros
- E em caso de colisão? O mecanismo de cadeia mais longa gerencia...
  - Caso haja o fork, uma das subcadeias vai ser naturalmente abandonada com o tempo pelos demais nós
  - Na Bitcoin, um bloco é considerado persistido 6 blocos depois
  - A subcadeia abandonada deve ter suas transações ressubmetidas

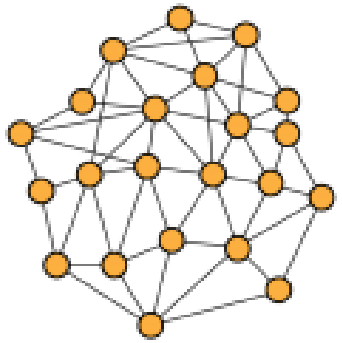


Blockchain Permissionada

- A blockchain como estrutura de dados decentralizada é interessante
- O mecanismo de hash criptográfico associado a cadeia de blocos favorece a segurança
- Plataformas baseadas em protocolos convencionais de consenso (PBFT, RAFT, BFT Smart) podem ser pensadas para entidades que formem consórcios permissionados

C : Consistency

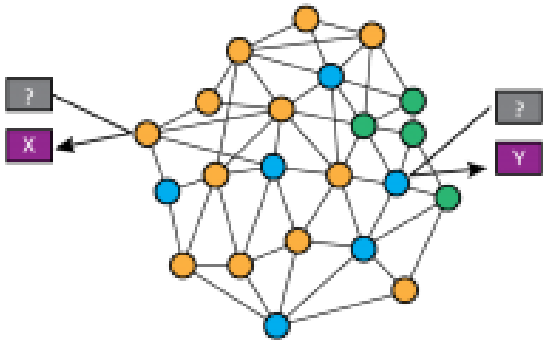
At any given time, all nodes in the network have exactly the same (most recent) value.



= Value: X @ 2018-05-03T08:52:40

A : Availability

Every request to the network receives a response, though without any guarantee that returned data is the most recent.



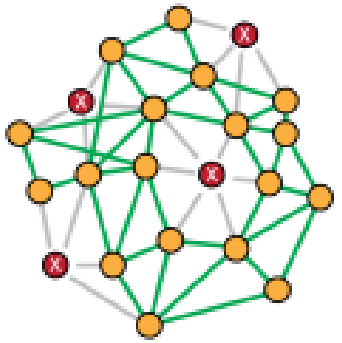
= Value: X @ 2018-05-03T08:52:40

= Value: Z @ 2018-05-03T08:32:58

= Value: Y @ 2018-05-03T07:12:12

P : Partition tolerance

The network continues to operate, even if an arbitrary number of nodes are failing.



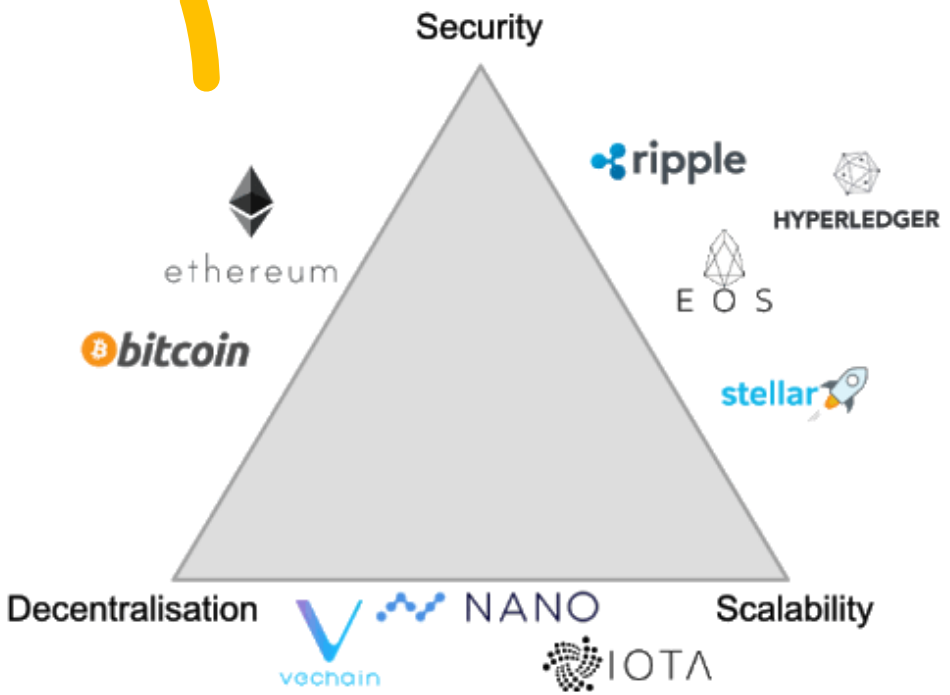
Nem tudo são flores: ou o Teorema CAP

- Blockchains não permissionadas têm que lidar com particionamentos. Bitcoin, por exemplo, vai favorecer a disponibilidade, prejudicando a consistência:
- - Ou seja, haverá *forks* e em algum momento espera-se que haja consistência a termo (eventually consistency).

https://paulkernfeld.com/2016/01/15/bitcoin-cap-theorem.html

O Trilema da Blockchain...

- O trilema da Blockchain não é o teorema CAP. Não há a priori uma impossibilidade teórica de obter, mas uma constatação prática.
- Exemplo: a bitcoin processa 7 TPS mas tem ´~ 83K nós
- Há uma proposta em construção do Ethereum 2.0 por meio de Shard chains para aumentar a escala de transações (2023 ou 2024 e prorrogando...)



Possibilidades fora da caixa

- Redes de canais de pagamento operam como um contrato de gaveta: decide entre partes interessadas e persiste somente mais a frente
- Para a bitcoin: Lighting Network
- Aumenta a escala, mas lembremos do trilema

Reflexões

- As propostas da Academia têm sido construídas de forma sofisticada e rodado em ambientes reais com grandes desafios e vários problemas/questões como o trilema da escala, segurança e disponibilidade
- Consenso tradicional favorece consistência ao custo de menor número de nós participantes, com membership e de eleição explícita de líder
- Blockchains não permissionadas lidam com particionamentos e utilizam estratégias que favorecem a disponibilidade, prejudicando a consistência (a termo)
- Estamos trabalhando nos limites destas ‘impossibilidades’ para pensar de forma disruptiva em novas formas de consenso.

Questões?

allan@ifba.edu.br