Consensus

Vincent Botbol, Marc Beunardeau,
Mathias Bourgoin, Julien Tesson

8 octobre 2019

Plan du cours

7 semaines + projet

1. Aperçu général
2. Cryptographie
3. Algorithmes de consensus
4. Privacy
5. Systèmes distribués en milieu adversarial
6. Protocoles économiques
7. Soutenance de projet

Réplication pour la tolérance aux pannes

Contexte

Duplication des calculs pour continuer à progresser en cas de défaillance d'un processeur.

Toutes les machines doivent se mettre d'accord (atteindre un consensus) afin d'être sûres qu'elles ont obtenu la même valeur.

Applications

- Processus tolérant aux panne
- Systèmes de fichiers distribués
- Bases de données distribuées

Le coup de la panne

Modèles de défaillances

- Défaillance totale (Fail-stop)
- Défaillance Byzantine

Modèle de réseau

- Synchrone
  - Tous les messages arrivent en un temps borné
- Asynchrone : Les messages peuvent être
  - Dupliqués, retardés, perdus

Sureté vs progression

La sureté (safety) et la progression (vivacité - liveness)

Théorème FLP (1985 - Fischer, Lynch and Patterson)
In an asynchronous setting, where only one processor might crash, there is no distributed algorithm that solves the consensus problem.

Paxos - (1982 - Lamport, Shostak and Pease)

Protocole pour un consensus distribué asynchrone non-Byzantin

Rôles

- Client (produit une valeur)
- Proposer (propose une valeur au consensus)
- Acceptor (vote pour une valeur)
- Learner (Traite la valeur acceptée par le consensus)
- Leader (Proposer selectionné)

Paxos - (1982 - Lamport, Shostak and Pease)

Protocole pour un consensus distribué asynchrone non-Byzantin

Propriétés

- Progression non-garantie (pour obtenir la sûreté)
  - L'algorithme assure la sureté en présence de plusieurs leader, pas la progression.
  - L'algorithme assure la sureté et la progression si un seul proposer pense être leader.
- N^2 messages pour N validateurs

Défaillance byzantine

Défaillance Byzantine

allégorie

Des généraux de l'armée byzantine campent autour d'une cité ennemie. Ils ne peuvent communiquer qu'à l'aide de messagers et doivent établir un plan de bataille commun (attaque ou retraite), faute de quoi la défaite sera inévitable. Cependant un certain nombre de ces généraux peut s'avérer être des traîtres, qui essayeront donc de semer la confusion parmi les autres. Le problème est donc de trouver un algorithme pour s'assurer que les généraux loyaux arrivent tout de même à se mettre d'accord sur un plan de bataille.

Défaillance Byzantine

Définition
Défaillance d'un ou plusieurs participants au consensus dans laquelle les participants défaillants peuvent paraître sains mais ne suivent pas le protocole.

Inclus

- Programmes défaillants
- Attaquant malicieux

PBFT (1999 - Castro, Liskov)

Practical Byzantine Fault Tolerance

- Max (n-1)/3 processus défaillants
- Sûreté assurée pour un réseau asynchrone
- Progression assurée pour un réseau inéluctablement synchrone
- N^2 Messages

Ouverture du réseau

Réseaux ouverts - permissionless

- les communications ne sont pas authentifiée
- Les nœuds vont et viennent
- Le nombre de participants est fluctuant

Conséquence
Il est à présent possible de créer un grand nombre de noeuds malicieux à la volées, et donc d'obtenir la majorité (*Sybil attack*).

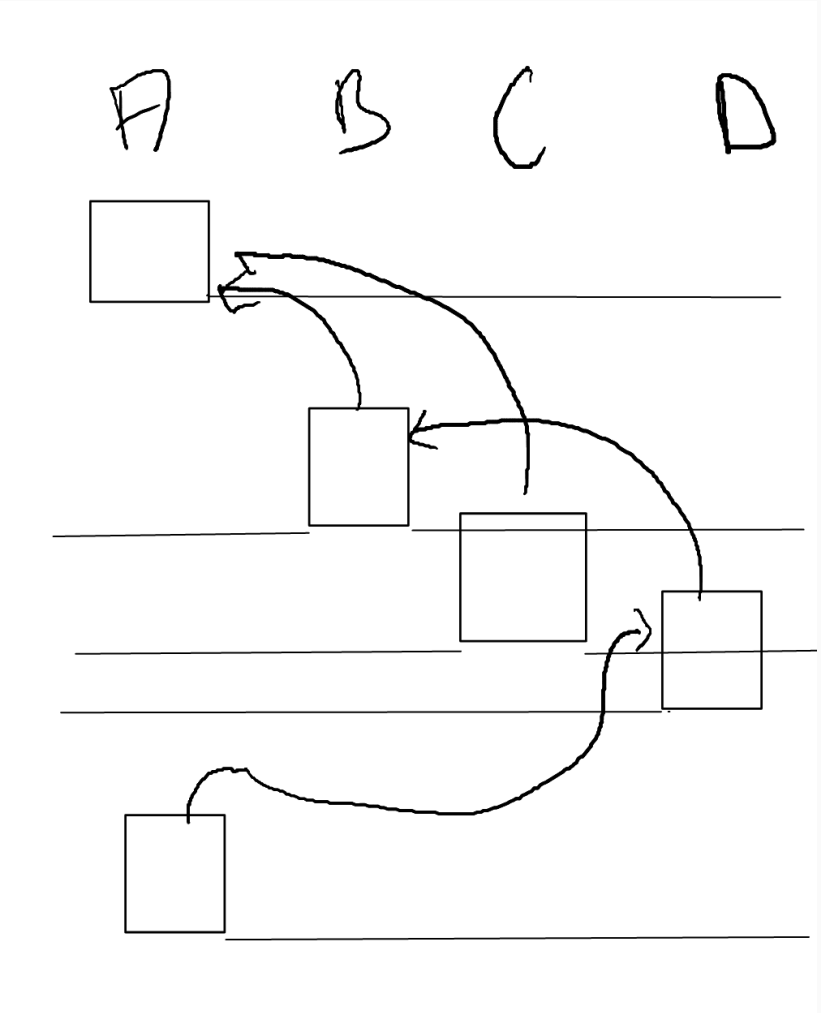
Il faudra donc s'assurer que casser le consensus est trop coûteux.

Preuve de travail (Proof of Work- PoW)

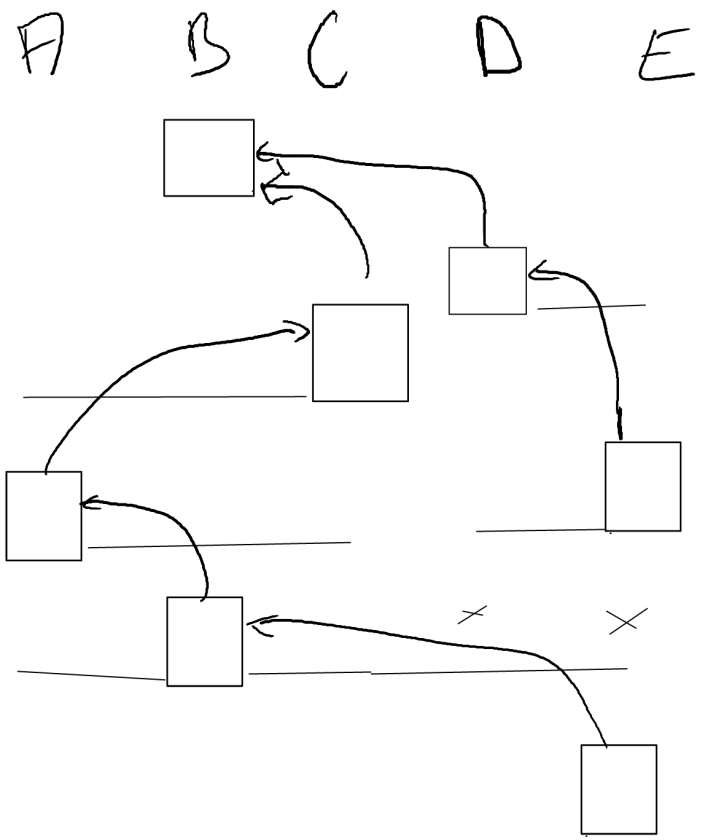
Bitcoin, Ethereum

- Leader $\rightarrow$ *Premier* à créer la meilleur proposition *valide*
- Accepteur $\rightarrow$ créé une proposition au-dessus de la proposition précédente pour devenir le leader suivant.

Réconciliation - rejet de fork



Réconciliation - rejet de fork



La meilleur des chaînes

fitness
La règle de consensus consiste à choisir la chaîne de blocs valides la plus longue.

Implicitement, il s'agit de celle qui a nécessité le plus de travail.

Sûreté et progression en PoW

Propriétés probabilistes, si la majorité de la puissance de calcul est honnête :

sûreté
Les blocs sont construits plus rapidement sur une chaine valide. Un bloc non valide deviendra rapidement orphelin.

Progression
La probabilité qu'un bloc soit inclus est lié à la longueur de la chaîne sur laquelle il est construit.

Les proposeurs sont donc incités à inclure les blocs proposés par d'autres.

Preuve d'enjeux (Proof of Stake - PoS)

Emmy+, Tendermint, Avalanche

- Leader désigné de manière pseudo-aléatoire
- proportionnellement à la quantité de jetons possédée
- sous-ensemble pseudo-aléatoire d'accepteurs (endorsers dans Emmy)

Alignement des intérêts : producteurs de blocs doivent posséder beaucoup de jeton, si la chaîne disfonctionne, leurs jetons perdent de la valeur.

Sureté et progression en PoS

Gestion des forks

Punir la participation à plusieurs chaînes :

- Dépôt de garantie pour la création d'un bloc
- Dépôt détruit ou redistribué si construction de blocs sur deux chaînes différentes.

Progression

Priorités pseudo-aléatoire pour désigner une liste de leaders auxquels sont assigné des priorités.si plusieurs leader injectent un block, la chaîne comportant le bloc du leader le plus prioritaire est retenu. Les porposeurs sont donc incités à construire sur cette chaine plutôt que de proposer un bloc de même niveau avec une plus faible priorité.

Delegated PoS

Délégation
Les possesseur de jetons peuvent déléguer leur droit de production de blocs.

Avantage

- nombre de proposeurs/validateurs réduits
- infrastructure plus légère

Modèle du consensus

Apply
Applique les transactions d'un bloc sur l'état courant.

Vérifie la validité du bloc.

Fitness ou FCR (Fork Choice Rule)
Calcul un score pour un bloc. Permet de choisir le bloc gagnant lors d'une divergence.

Modèle d'état du registre

UTXO - unspent transaction output
Transaction =

- des transactions non-dépensées sont déverrouillées
- de nouvelles transactions sont verrouillée (éventuellement par différentes clés publiques)

Account

- Un solde par compte
- transaction = virement signé par la clé privée du compte source

Contrat automatiques (Smart-contracts)

Ébauche dans bitcoin (bitcoin's scripts), popularisé par ethereum.

- Transaction vers compte exécute un script lié
- Programme ré-exécuté sur chaque nœud
- notion de durée d'exécution maximale (Gas)

Preuve d'autorité (Proof of authority- PoA)

See PBFT

(permissionned blockchains with authorized proposers)

Sharding

- Découpage des adresses en sous-groupes.
- Petits groupes peuvent établir plus rapidement un consensus
- Transactions inter-groupes compliquées.

Annexe

Dictionnaire

Pseudo-aléa(pseudo-randomness)
Processus produisant une valeur statistiquement aléatoire, de manière déterministe (prédictible) à partir d'une graine d'aléa (seed).

Sans la graine, la valeur obtenue est quasiment aussi difficile à prédire que le résultat d'un tirage aléatoire ; avec la graine on peut toujours prédire la valeur.