A quel point Cardano est-il sécurisé ?

Traduction du blog-post de @belowsearcher “How secure is Cardano?” paru ici le 26 Janvier 2019 par @psychomb

Avec une attaque réussie sur la blockchain Ethereum Classic (ETC) récemment, les attaques de type 51% sont revenues en première page, laissant plusieurs communautés en plein questionnement quant à la sécurité réelle de leur cryptomonnaie favorite. Cet article va essayer de donner un aperçu du modèle de sécurité s’appliquant à la couche de règlement de Cardano, aussi appelé le 'Cardano Settlement Layer'.

Afin de comprendre ce qui rend une cryptomonnaie sécurisée, il convient en premier lieu de se tourner vers Bitcoin, la cryptomonnaie par laquelle tout a commencé. Les deux propriétés suivantes sont deux caractéristiques cruciales pour tout registre distribué, Bitcoin compris :

- **La persistance.** Les transactions passées inscrites dans le registre doivent être inaltérables. C’est la notion d’immutabilité des transactions.

- **La vivacité.** Toute nouvelle transaction doit pouvoir être inscrite dans le registre sans retard injustifié. C'est ce qui rend le registre résistant à la censure.

Chez Bitcoin, la persistance est essentiellement assurée en combinant deux techniques:

- **La preuve de travail ou 'Proof of Work' (PoW).** Les noeuds de consensus utilisent leur puissance de calcul pour résoudre un problème cryptologique. Le premier y parvenant gagne le droit de créer le prochain bloc de la chaine. La résolution du problème par un noeud dépend donc pour beaucoup de sa puissance de calcul, même si un peu de chance rentre aussi en jeu.

- **La règle de la chaine la plus longue et de la chaine la plus lourde.** Lorsque plusieurs chaines de blocs coexistaient sur le réseau, par exemple après une attaque visant à modifier le registre, la chaine contenant le plus grand nombre de blocs était sélectionnée comme l'authentique chaine Bitcoin. Puisque les blocs ne peuvent être crées que par la preuve de travail, la chaine la plus longue était considérée comme celle ayant nécessité le plus de travail. Toutefois, il a récemment été introduit la notion de chaine la plus lourde, visant à quantifier plus précisément les calculs nécessaires à sa création, levant ainsi toute ambiguïté potentielle.

Lorsqu'un noeud de consensus gagne la course au calcul, il produit son nouveau bloc et se trouve récompensé pour cela en Bitcoin spécialement crées pour l'occasion. On parle alors de Bitcoins ou de pièces 'minées'. Mais ce n'est pas tout, il récupère aussi les frais des transactions qu'il aura eu à traiter. Si vous avez bien suivi jusque là, vous ne serez pas surpris d'apprendre que ce dernier point assure donc la vivacité de la blockchain, puisque cette rémunération supplémentaire incite les noeuds à traiter et à inclure des transactions.

Mais au fait, qu'est ce qu'une attaque 51% ?

Lorsqu'une seule personne ou un groupe de personnes détient plus de la moitié de la puissance de calcul (les fameux 51%), l'hypothèse d'une majorité honnête ne tient plus et une attaque 51% peut être menée sur la blockchain. Dans le cas des blockchains de type PoW, cela veut dire que le ou les attaquants sont capables de résoudre le problème cryptologique plus vite que les autres participants et peuvent donc produire le bloc suivant. Cet avantage permet de créer une version alternative du registre dans laquelle l'historique des transactions peut être réécrit, éliminant la persistance (immutabilité) et la vivacité (résistance à la censure) de la blockchain.

Bien que l'attaquant ne puisse pas réécrire les soldes des comptes des autres utilisateurs, ni inventer des transactions, les transactions précédentes peuvent en revanche être annulées. Plus l'attaquant a de la puissance de calcul, plus il peut remonter dans le temps. Ainsi, l'attaquant pourrait dépenser des Bitcoins (ou autre cryptomonnaie PoW) qu'il possède pour acquérir un actif du monde physique tel que de l'or, puis annuler à postériori cette transaction dans le registre. Le pirate possède alors à la fois l'or pour lequel il a échangé ses Bitcoins et les Bitcoins qu'il avait initialement dépensés pour acquérir cet or. Cela s'appelle une double dépense.

Dans le cas récent de l'attaque 51% sur la blockchain Ethereum Classic (ETC), il s'agissait en réalité d'une cible assez facile et cela pour deux raisons. Premièrement, ETC utilise le même algorithme que Ethereum (ETH), la blockchain-mère dont il est issu, mais fonctionne avec une puissance de calcul moindre. Bien que l'on ne sache pas qui a exécuté l'attaque, il est tout à fait possible qu'un acteur d'Ethereum ait simplement redirigé sa puissance de

calcul sur ETC. L'autre vulnérabilité d'ETC réside dans le fait qu'il ne coûte que 4106 $ de l'heure pour qui voudrait réaliser une attaque 51%. La puissance de calcul requise pourrait donc être entièrement louée sur une plateforme spécialisée telle que NiceHash, ce qui signifie qu'un attaquant n'aurait même pas besoin d'acquérir de matériel à ses propres frais.

Si l'on revient à Bitcoin, la distribution de la puissance de calcul au sein des différents groupes de mineurs est souvent un sujet de discussion. Par exemple, en juillet 2014, GHash.io contrôlait plus de 51% de la puissance de calcul de Bitcoin, créant ainsi un point de défaillance unique, heureusement sans conséquences alors. La raison probable de cela est que lorsque les participants tentent d'optimiser leurs gains, les simulations montrent que les groupes de mineurs ont tendance à former un groupe unique, synonyme de centralisation. C'est ce qu'on appelle la tragédie des biens communs : même si les participants accordent une grande valeur à la décentralisation en tant que concept, aucun d'entre eux n'est effectivement prêt en supporter individuellement le fardeau.

En quoi les systèmes de preuve d'enjeu (PoS) diffèrent ?

Il existe des inconvénients potentiels ou avérés au PoW. Par exemple, l'utilisation intensive de puissants calculateurs par les mineurs consomme beaucoup d'énergie. Bien sûr, l'efficacité de ces équipements s'améliore sans cesse, les mineurs tendent à privilégier des sources d'énergie peu coûteuses (voire renouvelables) et l'on peut toujours soutenir que la valeur ajoutée de Bitcoin à la société justifie une consommation élevée d'énergie. Toutefois, un système pouvant atteindre un niveau de sécurité similaire tout en consommant moins serait une amélioration indéniable, ne serait ce que d'un point de vue écologique.

Dans les cryptomonnaies PoW avec offre limitée (au hasard… Bitcoin !), le nombre de nouvelles pièces pouvant être minées diminue avec le temps et seuls les frais de transactions constitueront à terme les récompenses pour la création de blocs. Bien que l'on parle ici d'un futur lointain, il n'est pas clair que ces frais de transactions agrégés constitueront des récompenses à la hauteur des coûts des mineurs. Voudront-ils alors toujours faire ce travail ?

Enfin, les utilisateurs classiques ou détenteurs de pièces qui ne possèdent pas de nœud de consensus 'minier' - un calculateur digne de ce nom - ne peuvent pas directement participer à la gouvernance du réseau alors même que leurs intérêts ne sont pas toujours alignés avec ceux des développeurs et des mineurs.

Les systèmes PoS tentent d'améliorer ces points. Dans les systèmes PoW, les participants au mécanisme de consensus mettent essentiellement de l'argent en jeu en le convertissant en matériel informatique et en électricité pour effectuer les calculs requis. Dans les systèmes PoS, les participants utilisent la monnaie native de la blockchain pour prouver qu'ils ont bien quelque chose à

perdre : un enjeu (d'où le nom d'ailleurs). En conséquence, aucun calcul intensif n'est nécessaire chez les PoS, ce qui réduit la consommation d'énergie. Dans certains cas, cet enjeu donne aussi le droit de participer à la gouvernance du réseau par le vote.

Tout cela ressemble à une solution parfaite, direz vous. Oui, mais les systèmes PoS ont aussi leurs petits problèmes de conception. Explications :

Étant donné qu'aucune ressource matérielle lourde n'est nécessaire pour produire des blocs dans un système PoS, il est possible de créer un ou plusieurs historiques alternatifs du registre et de créer ainsi sans frais plusieurs chaînes concurrentes.

Certains systèmes, comme le futur protocole d'Ethereum 'Casper', ont proposé de résoudre ce problème en gelant l'enjeu et en punissant les nœuds de consensus si un comportement malveillant était observé. Vous aurez compris que la punition en question ici est la confiscation de toute ou partie de l'enjeu qui a été gelé. Bien que cela puisse dissuader les nœuds d'agir de manière malveillante, cela limite également la capacité des participants honnêtes à dépenser leurs cryptomonnaie quand ils le veulent et peut même les exposer à une punition plus qu'injuste en cas d'attaque 51% réussie ; l'attaquant devenu majoritaire considèrera en effet les nœuds honnêtes comme des tricheurs méritant d'être punis.

Il ressort de tout cela que la règle de la chaîne la plus longue observée dans les PoW n'est pas directement applicable dans les PoS. Les nœuds qui joignent le réseau pour la première fois ou après avoir été déconnectés pendant un certain temps doivent donc faire confiance aux informations qu'ils reçoivent des autres nœuds. C'est ce que l'on appelle le problème de l'amorçage. Ce dernier problème accroît la vulnérabilité du réseau face aux attaques dites 'à longue portée'. En effet, lors d'une attaque à longue portée, une version alternative de la chaîne de blocs est proposée par un attaquant à un nœud. Ce nœud est naïf et ne dispose que d'une information limitée, voire d'aucune information récente lui permettant de déterminer s'il s'agit de la bonne version du registre.

Si le PoS n'est pas un concept nouveau, aucun système de PoS n'a jusqu'à ce jour permis de relever tous ces défis de conception afin d'atteindre le même niveau de sécurité que Bitcoin.

Comment Cardano tente-t-il de les résoudre ?

Lorsque le papier traitant de Bitcoin est paru le 31 octobre 2008 et que le réseau a été mis en ligne le 3 janvier 2009, il s'agissait essentiellement d'une expérience. Alors que le concept était clairement très, très bien pensé, sa sécurité n'a été mathématiquement prouvée qu'en 2015.

Contrairement à Bitcoin, pour lequel la pratique a précédé la théorie formelle, l'un des objectifs de Cardano est de formaliser avant de passer à la pratique. Il s'agit donc de prouver que chaque revendication de sécurité est mathématiquement correcte avant de la mettre en œuvre sous forme de code.

Cardano est constitué en deux couches ; la couche de règlement sur laquelle les transactions monétaires sont exécutées et une couche de calcul utilisée pour les contrats intelligents. Lorsqu'on discute de la sécurité de Cardano, les deux couches doivent évidemment être prises en compte. Toutefois, cet article se concentre sur la sécurité de la couche de règlement de Cardano, tandis que la sécurité de sa couche de contrats intelligents (comparée à Ethereum, par exemple) fera l'objet d'un prochain article.

L'écosystème Cardano a été lancé par Input Output Hong Kong (I.O.H.K.), Emurgo et la Fondation Cardano. C'est à I.O.H.K. qu'il incombe de développer la technologie de la blockchain elle-même. Depuis son lancement en 2015, plus de 40 articles universitaires sur Cardano ont été publiés, dont certains peuvent être consultés dans la bibliothèque de recherche sur le site web de I.O.H.K. Le programme de recherche du mécanisme de consensus de la couche de règlement s'intitule 'Ouroboros'.

Ouroboros Classic

La première version d'Ouroboros, Ouroboros Classic, s'est concentrée sur la sécurité dans un environnement synchrone. Il s'agit d'un environnement dans lequel les nœuds sont toujours en ligne et prêts à produire des blocs en cas de besoin, et leurs horloges sont toutes synchronisées.

Dans Ouroboros, une époque est une période divisée en 21 600 tranches de 20 secondes chacune, ce qui signifie que chaque époque dure exactement 5 jours. Chaque tranche représente une fenêtre temporelle de 20 secondes pendant laquelle le 'chef de tranche' (un nœud sélectionné) peut créer un bloc. Avant le début d'une époque, tous les chefs de tranches pour cette époque sont élus au hasard.

Pour ce faire, Ouroboros Classic utilise la méthode appelée 'Suivez le Satoshi', inventée par le créateur de Litecoin, Charlie Lee, en 2012. Pour faire court, chaque Lovelace (1 Lovelace = 0.000001 ADA) d'un enjeu est une forme de billet de loterie permettant de gagner le droits de créer un bloc. Cela signifie que n'importe qui peut participer avec n'importe quelle mise d'enjeu (1 Lovelace suffit). Les chances de gagner sont proportionnelles au nombre de Lovelace misés. Plus l'enjeu est important, plus les chances d'être élu sont grandes.

Cependant, une loterie a besoin de plus que de simples tickets ; elle a également besoin d'une méthode pour sélectionner un gagnant au hasard. Dans Ouroboros Classic, le nombre aléatoire aidant à déterminer les chefs de tranche de la période suivante est générée à l'aide d'un schéma cryptologique

appelé PVSS. Chaque fois qu'un bloc est créé, les nœuds jouent à pile ou face afin de générer un nombre aléatoire et utilisent PVSS pour en chiffrer les résultats sur la chaîne de blocs, et ainsi tout le monde peut le vérifier. À la fin de l'époque, ces nombres sont combinés afin de produire un nombre aléatoire final que tous les participants utilisent pour choisir des chefs de tranche pour la période suivante. Étant donné que les nombres aléatoires créés durant une époque aident à former l'époque suivante, une boucle se forme. C'est pourquoi le protocole a été baptisé Ouroboros, d'après le serpent mythique se mordant la queue.

Ouroboros Classic a été le premier protocole PoS dont la capacité à garantir la persistance et la vivacité dans un environnement synchrone a été mathématiquement prouvée - bien entendu, en supposant une majorité honnête de participants. Cependant en conditions réelles, les nœuds peuvent être déconnectés accidentellement (panne d'électricité, panne d'ordinateur) ou intentionnellement (le gestionnaire de nœud décide de tout arrêter). De plus, les horloges sur Internet ne sont pas toujours toutes synchronisées. Cela signifie qu'en conditions réelles, l'hypothèse de synchronie retenue pour le protocole n'a aucune chance d'exister. De plus, la sélection des chefs de tranche est totalement transparente dans Ouroboros Classic et ils sont connus à l'avance, ce qui n'est pas idéal pour qui veut établir un haut niveau de sécurité. C'est la raison pour laquelle une deuxième version du protocole, 'Ouroboros Praos', avait pour but d'assurer la sécurité dans un environnement semi-synchrone et à masquer le processus de sélection des chefs de tranche.

Ouroboros Praos

En grec ancien, 'praos' signifie relâché. En effet, dans ce nouveau protocole les participants ne sont pas sous pression pour être en ligne en permanence avec une horloge synchronisée. Pour y parvenir, plusieurs techniques sont combinées. Premièrement, la méthode PVSS (cf. paragraphe précédent) a été remplacée par une autre fonction appelée VRF. Les VRF ont été inventés par Silvio Micali, lauréat du prix Turing, actuellement professeur au MIT et travaillant sur une cryptomonnaie appelée Algorand. Au cours de chaque époque, les nœuds participants utilisent trois éléments dans l'élection des chefs de tranche : une photo instantanée de la distribution des enjeux de l'époque avant qu'elle ne commence, un nombre aléatoire calculé en fonction de l'époque précédente et enfin, la fonction VRF présente dans le code exécuté par chaque nœud.

La photo instantanée de la distribution des enjeux est en soi un principe assez simple. Avant le début d'une nouvelle époque, une image de tous les Lovelaces (0.000001 ADA) faisant partie des enjeux en cours et des nœuds contrôlant leurs droits d'enjeu est prise. Étant donné que cette photo instantanée est utilisée pour élire tous les chef de tranche pour une époque entière, les pièces des enjeux ne sont jamais gelées et restent donc utilisables à tout moment par leurs propriétaires.

Dans l’époque courante, les noeuds utilisent la photo instantanée de la distribution des enjeux et le nombre aléatoire déterminé au cours de l’époque précédente comme valeurs d’entrée pour leur fonction VRF. Cette fonction leur permet de générer un nombre pseudo-aléatoire qui détermine si ils ont remporté l’élection. Le nœud vainqueur crée le bloc et publie ce numéro pseudo-aléatoire dans l'en-tête du bloc. Tous les autres noeuds utilisent leur propre VRF pour vérifier si ce numéro présent dans le bloc correspond bien au noeud qui a remporté l’élection. Les nœuds ne recherchent donc pas qui a remporté l’élection avant que le bloc ne soit signé. Cela signifie également que si un nœud est en mesure de créer un bloc mais qu'il est hors ligne à ce moment-là, la possibilité de créer le bloc ne fait que passer et les autres nœuds ne découvrent jamais qui était supposé le faire. Le bloc ne peut pas être créé par un autre nœud (par exemple un attaquant), car il serait immédiatement reconnu comme invalide par les autres.

À chaque époque (environ aux ~ 3/4 des 5 jours), tous les numéros pseudo-aléatoires présents dans les en-têtes de bloc sont combinés. Tous les nœuds utilisent cette valeur combinée pour calculer localement le nombre aléatoire commun qui sera utilisé pour l’époque suivante. Puisque tous les nœuds retrouvent les numéros pseudo-aléatoires dans la même blockchain et utilisent la même méthode pour les combiner, tous les résultats correspondent, même si les nœuds les calculent localement. Ce nombre aléatoire nouvellement créé et le nouvel instantané de la distribution des enjeux sont ensuite utilisés pour l’époque suivante, créant ainsi un cycle sans fin. Dans Ouroboros Praos, des preuves mathématiques ont montré que la persistance et la vivacité peuvent être garanties même dans un cadre semi-synchrone et toujours dans l’hypothèse d’une majorité honnête de participants. Cependant, le problème de l’amorçage n’était pas encore résolu, ce qui est donc devenu le centre d’attention de la troisième version du protocole : Ouroboros Genesis.

Ouroboros Genesis

Comme décrit précédemment, lorsqu'un nouveau nœud ou un nœud qui est hors ligne depuis un certain temps rejoint le réseau, il doit pouvoir faire confiance aux informations fournies par d'autres nœuds concernant la version de la blockchain qui représente la vérité. Dans les systèmes PoW, cela peut être fait en utilisant la règle de blockchain la plus longue (et/ou lourde), puisque la majorité du temps de calcul a été consacrée à sa création. Elle est donc considérée comme la vraie version du registre - en supposant que la majorité des les mineurs sont honnêtes.

Les protocoles PoS utilisent généralement des méthodes alternatives mais ces dernières ne fonctionnent que dans un environnement synchrone où les nœuds sont toujours en ligne. On a déjà vu qu’une telle hypothèse est quasiment impossible à tenir dans le monde réel. Dans l’article décrivant Ouroboros Genesis, les auteurs concluent même que dans un tel contexte, aucun des systèmes PoS existants ne peut produire un registre pleinement fonctionnel comme le fait Bitcoin.

Pour résoudre ce problème d'amorçage, une nouvelle règle de sélection de chaîne appelée 'règle de plénitude' est proposée dans Ouroboros Genesis. Bien que les preuves mathématiques décrites dans le document de 64 pages soient difficiles à saisir pour les non-cryptographes, les auteurs démontrent que, juste après leur point de divergence, les chaînes concurrentes produites par des attaquants présentent une densité moindre que la véritable chaîne. Autrement dit, la chaîne de l'attaquant contiendra moins de blocs peu de temps après le point de divergence, même si elle peut contenir potentiellement plus de blocs au total et être plus longue.

Par conséquent, lorsque plusieurs chaînes de même longueur sont disponibles, la règle de plénitude recherche le point de divergence des chaînes et détermine pour quelle version la distribution de blocs après le point de divergence est la plus dense. En raison de cette règle, les nœuds qui sont nouveaux sur le réseau ou qui ont été hors-ligne pendant un certain temps peuvent rejoindre le réseau et avoir la garantie de télécharger la version correcte de la blockchain entière, tant qu'il y a suffisamment de parties honnêtes. Cela résout le problème de l'amorçage et empêche les attaques à longue portée.

Il convient de noter que la règle de plénitude ne fonctionne que dans un protocole comme Ouroboros à cause des propriétés suivantes : (i) le temps est divisé en intervalles, (ii) les chefs de tranche sont élus à l'avance pour l'ensemble de l'époque et (iii) tous les nœuds peuvent vérifier que chaque bloc a été créé par le nœud approprié. Cette combinaison de fonctionnalités permet de garantir que personne ne peut se tromper en créant un bloc pendant la tranche de quelqu'un d'autre. En conséquence, il est impossible pour un seul nœud de créer une fausse chaîne sans disposer d'un nombre incalculable d'emplacements vides. Une telle chaîne serait automatiquement ignorée en raison de la règle de plénitude, car elle serait bien moins dense.

Depuis la version Genesis, Ouroboros est le premier protocole PoS mathématiquement démontré comme garantissant la persistance et la vivacité dans un environnement à la fois synchrone et semi-synchrone - sous l'hypothèse d'une majorité honnête de participant, exactement comme Bitcoin. Par conséquent, il est plus sécurisé que d'autres protocoles PoS qui nécessitent au moins 2/3 de participants honnêtes, comme Casper (Ethereum) ou Algorand. Ouroboros est tout aussi sécurisé que Bitcoin, avec toutefois une dépense d'énergie bien inférieure et de meilleures performances.

Dans la seconde partie de ce très long article, nous aborderons un désavantage d' Ourobouros par rapport à Bitcoin. Cela est lié à sa nature de preuve d'enjeu, et nous verrons si cela représente un réel danger pour l'écosystème Cardano...