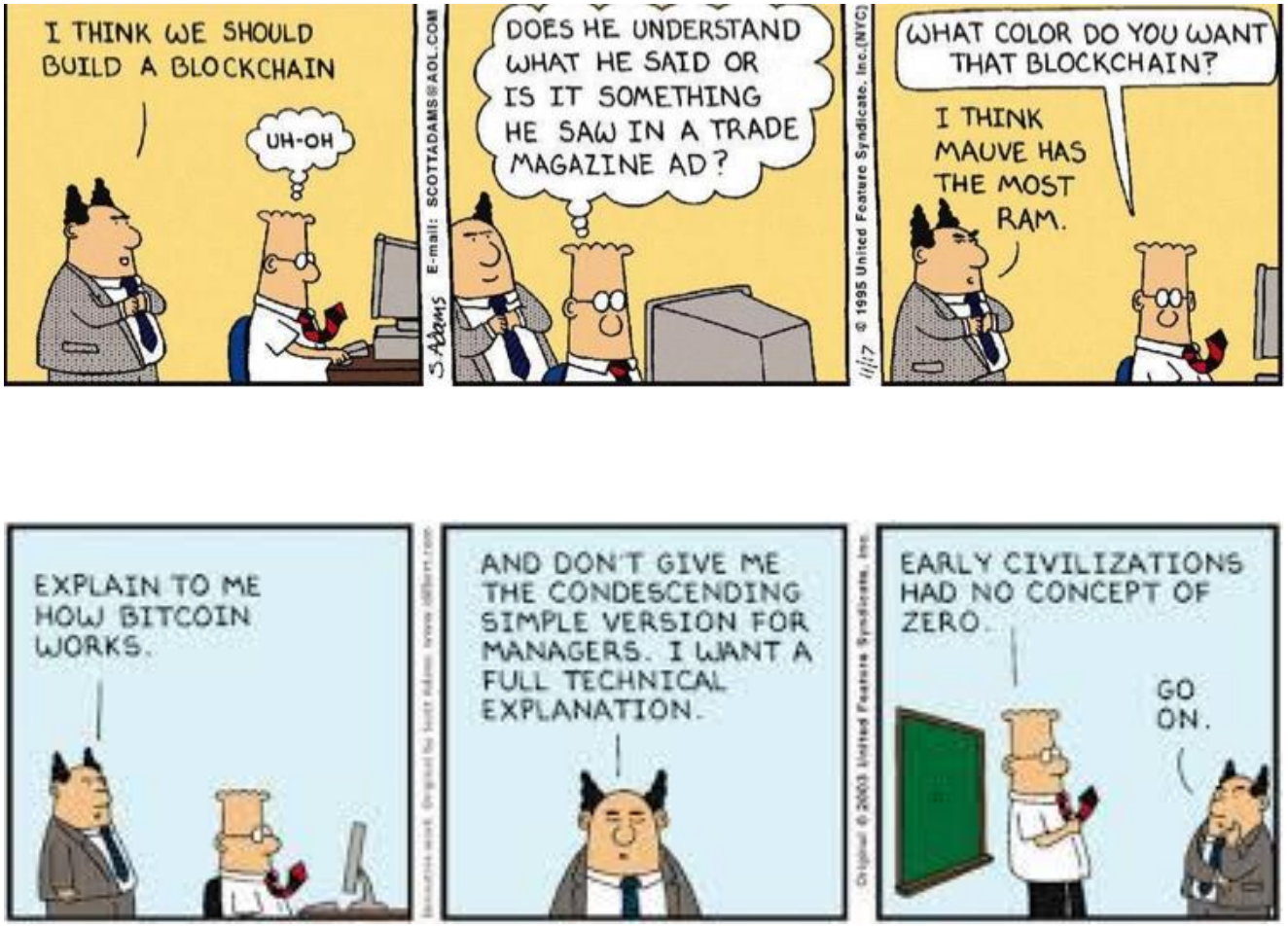
REDOCS 2017

Gestion de la confidentialité des données dans une architecture de type blockchain

Aline Gouget, Advanced Cryptography
October, 2017

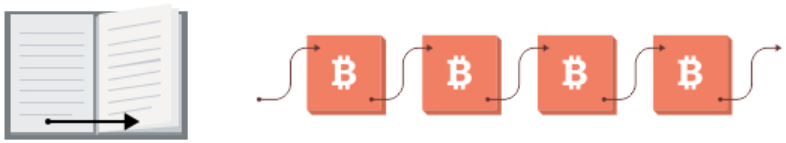


http://dilbert.com/

gemalto

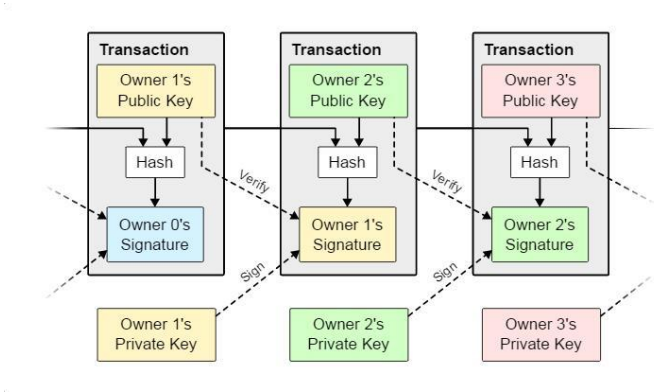
Technologies Blockchain : qu'est-ce que c'est ?

- Une technologie de « grand livre » distribué, a.k.a DLT



- Permet de maintenir une liste croissante d'enregistrements (records)

- Chaque bloc est horodaté et relié aux blocs précédents



- Toutes les transactions sont signées et vérifiées de manière décentralisée

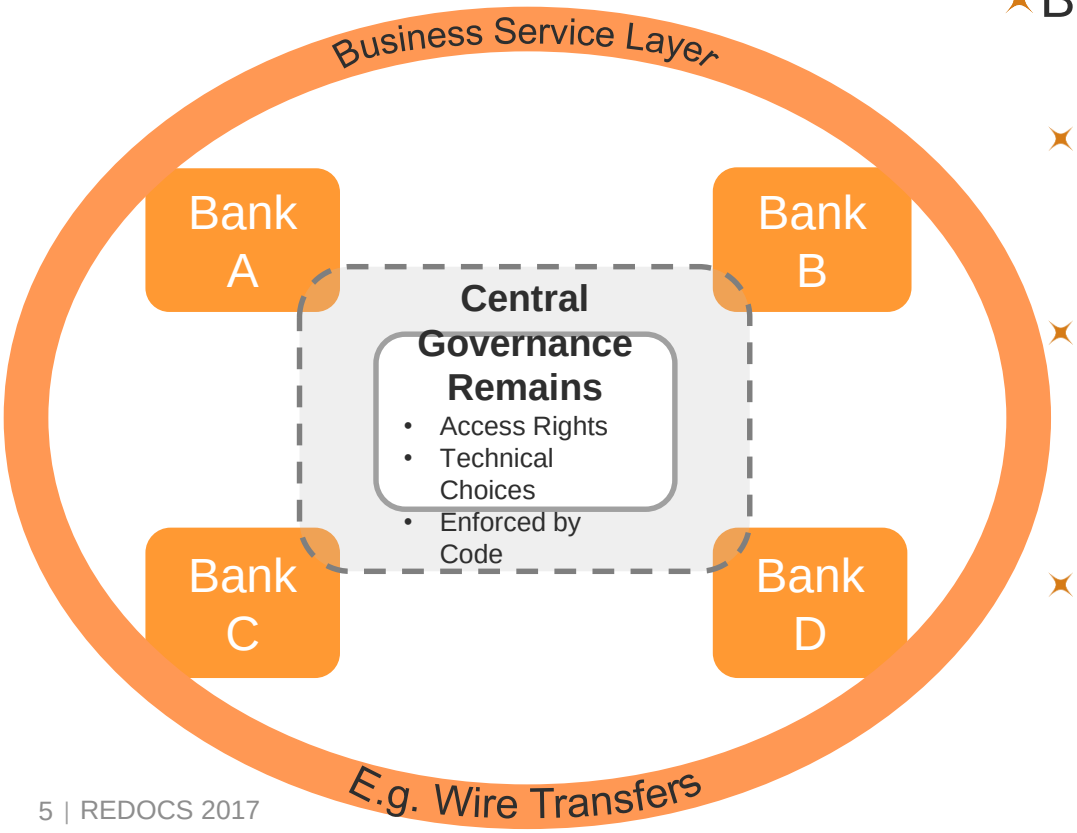
gemalto

“Transaction”, à prendre au sens large

- Nombreux sens possibles, en fonction des cas d’usages
  - Monnaie électronique type Bitcoin avec des pseudonymes
  - Transaction financières avec authentification du payeur
  - Echange d’actifs, contrats intelligents
  - Données brutes ou hash de données brutes

Pourquoi les industriels s'y intéressent ?

- Le modèle « consortium » ou « permissioned »
  - Ensemble de partenaires business sans « chef » commun
  - Un petit sous-ensemble des membres ne doit pas pouvoir empêcher le bon fonctionnement du système



- Bénéfices attendus
  - Réduction des coûts et/ou complexité
  - Amélioration des processus d'auditabilité (automatisation, confiance)
  - Augmentation de la transparence tout en gardant le contrôle

gemalto

Le projet

- Focus sur le modèle « consortium »

- Sujet: Protection des données lorsque plusieurs cercles de confiance partagent le même grand livre distribué

- Exemple
  1. Une crypto-monnaie privée gérée par des entités financières
     - Les données de transactions sont partagées uniquement au sein du cercle de confiance
     - L'horodatage doit s'ancrer dans le grand livre distribué commun
     - L'auditabilité doit pouvoir se faire en divulguant tout ou partie des données

  2. Production/consommation d'énergie tout en protégeant la vie privée des utilisateurs
     - Les factures sont établies par une entité de facturation spécifique, en utilisant des données insérées dans le ledger
     - Chacun des acteurs gérant une partie du réseau d'énergie peut faire les statistiques utiles lui permettant d'adapter ses équipements (e.g. re-dimensionnement de la partie du réseau dont il a la charge)

- Spécification d'un prototype sur Hyperledger Fabric

Protection des données, que peut-on vouloir faire ?

- Authentifier la source fournissant la donnée
- Authentifier les demandes d'accès en lecture à une donnée
- Protéger la confidentialité & l'authenticité de la donnée pendant le transport
- Stockage sécurisée de la donnée, potentiellement pour une durée longue
- Garantir différentes propriétés de « privacy » : pseudonymat, anonymat, non-reliabilité

Phase 1: Hyperledger Fabric

- https://github.com/hyperledger/fabric

- Implémentation de la technologie blockchain qui propose une architecture modulable permettant de plugger diverses implémentations de fonctions

- Comprendre, analyser l'architecture et les modèles de sécurité associés
  - Quelles sont les entités impliquées: utilisateur final, client, entité de validation, service de membership, auditeur de membership,…
  - Point clé: qui gère les clés de qui/quoi? Pour quel modèle de confiance? Qui stocke les données

- Comprendre les fonctionnalités permettant de protéger la confidentialité des données via l'utilisation de membership service ou de contrat confidentiels

- Spécification d'un premier exemple pouvant être implémenté sur Hyperledger Fabric
  - Proposer une architecture + flux de transaction impliquant 1 cercle de confiance

Phase 2: proposition d'améliorations

- Que peut-on améliorer dans la version actuelle d'hyperledger fabric?

- L'objectif principal est de réduire autant que possible le niveau de confiance nécessaire dans chacune des entités du système
  1. Distribuer la fonction de service de membership?
  2. Chiffrement des données, comment faire?
     - On voudrait par exemple que certaines données soit accessible au sein d'un cercle de partenaires business sans pour autant que les valideurs aient nécessairement accès à toutes les clés de déchiffrement
  3. Quel protocole de consensus est le plus adapté?

- Point clé: qui gère les clés de qui/quoi? Pour quel modèle de confiance? Qui stocke les données

gemalto

Phase 3: Spécification d'un Proof-Of-Concept

- Protection des données lorsque plusieurs cercles de confiance partagent le même grand livre distribué
- L'idée est d'identifier ce que l'on peut faire de mieux avec l'architecture Hyperledger Fabric actuelle sur un premier exemple
- Et d'ajuster avec vos propositions d'amélioration, en spécifiant ce qu'il faudrait alors implémenter
- Et finalement,
  - Quelles sont les propriétés de sécurité garanties? En particulier concernant la protection des données
  - Quel niveau de confiance doit-on avoir dans quelle entité?

Pour commencer, pour aller plus loin

- IBM blockhain for dummies
  - https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=XIM12354USEN&
- Cryptography and Protocols in Hyperledger Fabric (Slides) Real-World Cryptography Conference 2017
  - https://www.zurich.ibm.com/~cca/talks/20170106-blockchain-rwc.pdf
- Welcome to Hyperledger Fabric
  - http://hyperledger-fabric.readthedocs.io/en/latest/
- Blockchain Protocol Analysis and Security Engineering 2017
  - https://cyber.stanford.edu/blockchainconf

Merci de votre attention!

Questions?