Caltech

CS/IDS 142: Lecture 10.2
Bitcoin Properties and Analysis

Richard M. Murray

4 December 2019

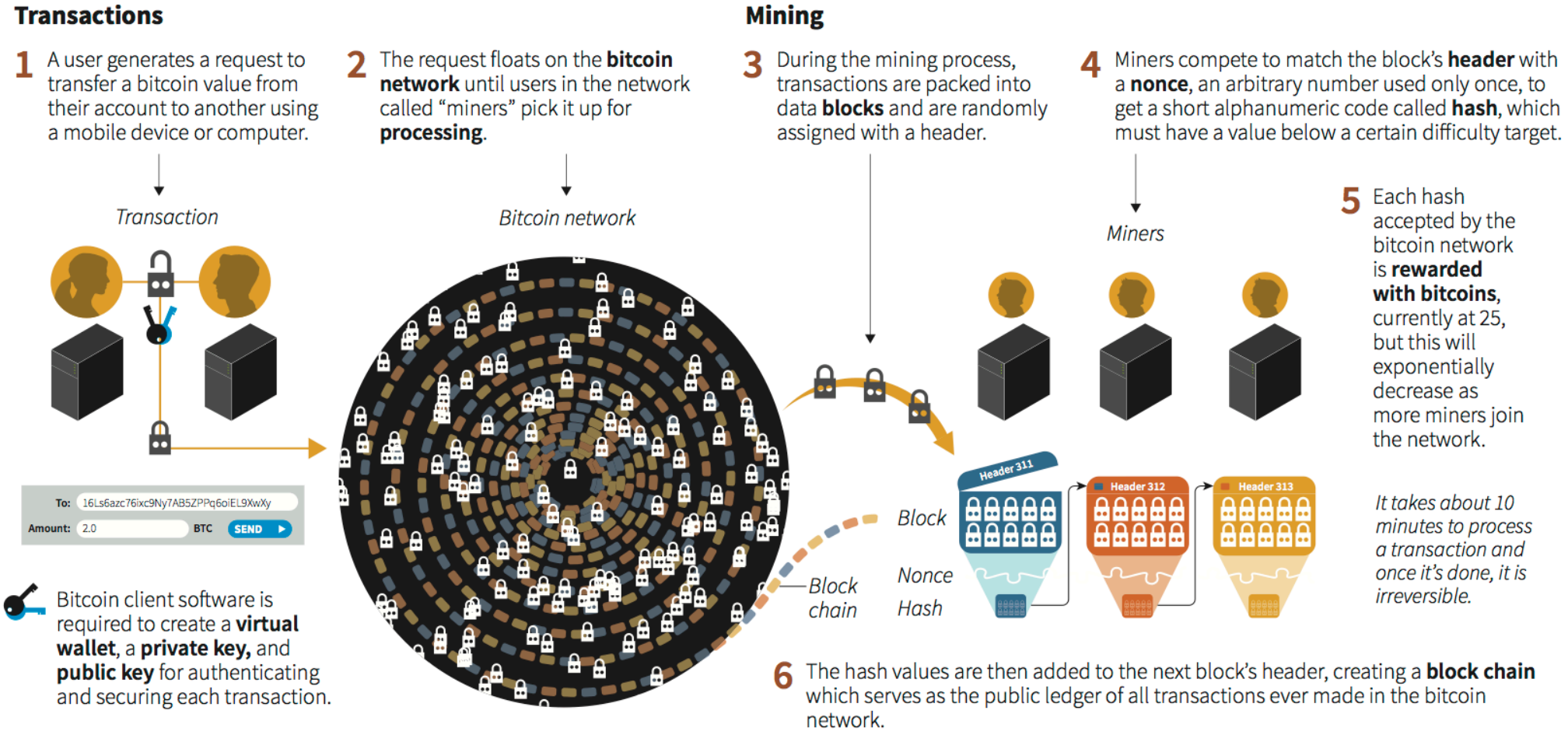
Goals:

- Describe what is known about correctness of Bitcoin
- Analyze double spent attacks and “orphan races”

Reading:

- Bitcoin: A peer-to-peer electronic cash system, Satoshi Nakamoto. http://bitcoin.org/bitcoin.pdf, 2008.
- A. Narayanan, J. Bonneau, E. Felten, A. Miller, S. Goldfeder, *Bitcoin and Cryptocurrency Technologies*. Princeton University Press, 2017. Chapter 1 (optional) and Chapter 2. http://bitcoinbook.cs.princeton.edu
- [optional] J. Garay and A. Kiayias and N. Leonardos, The Bitcoin Backbone Protocol: Analysis and Applications, Cryptology ePrint Archive, Report 2014/765, https://eprint.iacr.org/2014/765 (revised 1 Jul 2019)

Summary: Distributed Ledger and Bitcoin



Sources: Bitcoin.org; Bitcoin Ladder

C. Inton, staff, 09/12/2015

REUTERS

- Bitcoin implements a consensus protocol to agree on valid transactions
- Priority in proposed blockchains is determined by length of proposed blockchain
- Authenticated signatures => no forgery, but can still have "double spend" attacks

Formal Analysis of Distributed Ledger (1 of 2)

No formal analysis of the correctness of the Bitcoin protocol is yet available

- Safety properties are available, but not guarantee of progress
- Most proofs focus on showing the prefix of honest peers is stable

Problem specification (following Garay, Kiayias, Leonardos)

- Safety: All honest peers will have the same prefix for some depth k
- Progress: A conflict-free transaction will eventually be deeply confirmed in the blockchain of an honest peer

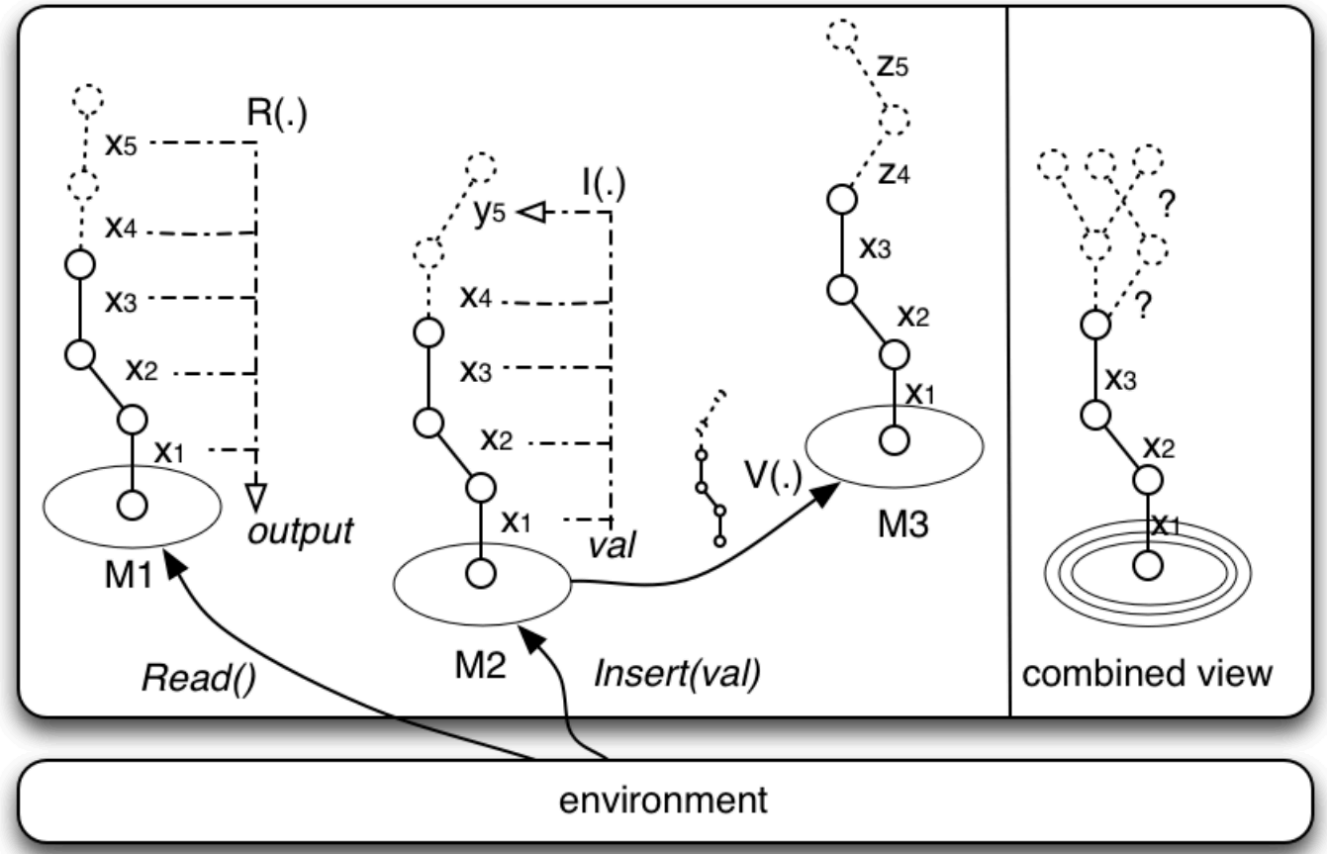
Formal definitions and analysis:

- Let C be a blockchain and let $C^{\lceil k}$ be the chain with the last k blocks removed
- Let n = number of players, t = number of traitors, $\mu = t/(n-t)$
- Common-Prefix Property: For any two honest players P1 and P2 adopting chains C1, C2 and round r1 = r2, it holds that $C1^{\lceil k} = C2^{\lceil k}$
- Chain Quality Property: For any host party P with chain C, it holds that for any ℓ consecutive blocks of of C, the ratio of adversarial blocks is at most μ
- Chain Growth Property: For any honest party P with chain C it holds that for any s rounds there are at least τ s blocks added to the chain (τ = chain growth parameter)

Formal Analysis of Distributed Ledger (2 of 2)

Bitcoin backbone protocol

- Read instruction [M1]: return content of chain
- Insert instruction [M2]: extend chain, solve proof-of-work, and broadcast extended chain to all
- Validate instruction [M3]: receive newly extended chain and adopt if better than local chain
- Note: miners agree on prefix to chain, but not on latest transactions



Properties of the protocol

- Likelihood that common prefix not present drops exponentially in length of chain
- Exponentially unlikely that adversary contributed to chain as the chain gets longer

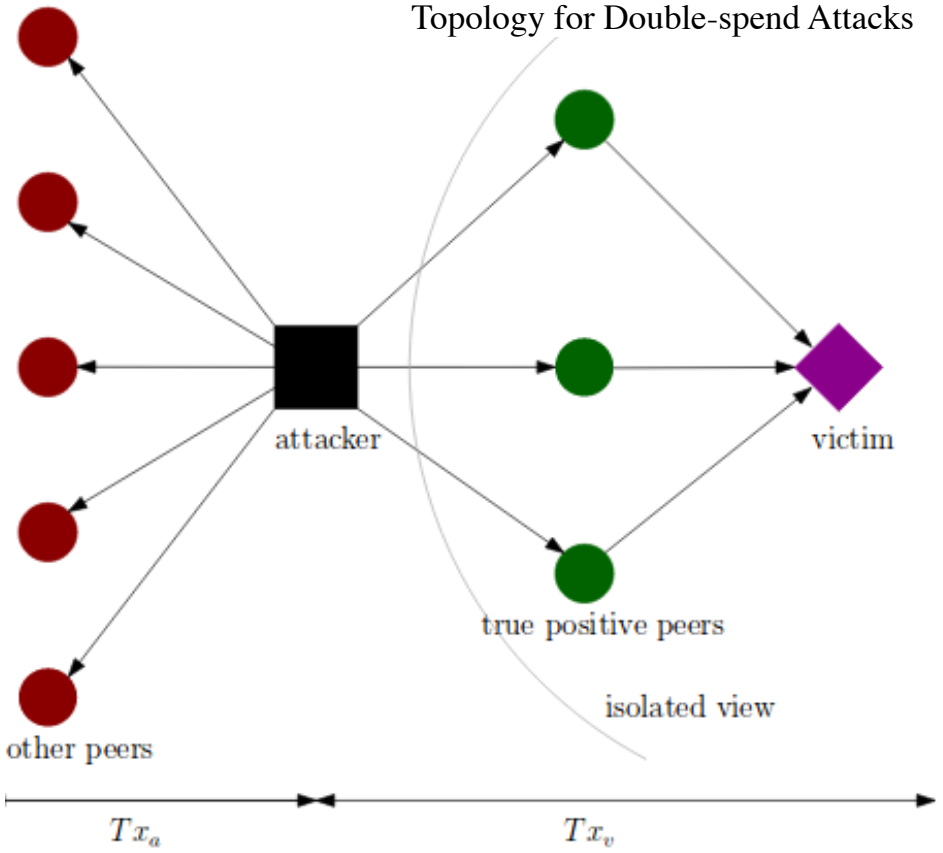
Note: all properties are in terms of probabilities…

Double Spend Attacks

Steps in a double spend attack

- Broadcast actual transaction with merchant that we want to attack
- Broadcast fraudulent transaction or secretly mine branch that builds on latest block w/ conflicting transaction
- Wait until transaction has ben confirmed by the merchant
- If fraudulent transaction is in longest chain before merchant gets enough confir-mations ⇒ don't deliver service
- If merchant receives confirmations, extend secretly mined branch until it is longer than public branch
- Broadcast secretly mined branch, and since it has the longest chain, it is accepted

Source: M. Lei, Exploiting Bitcoins
Topology for Double-spend Attacks



Txv = transaction that seemingly confirms payment for service claimed, but is then invalidated.

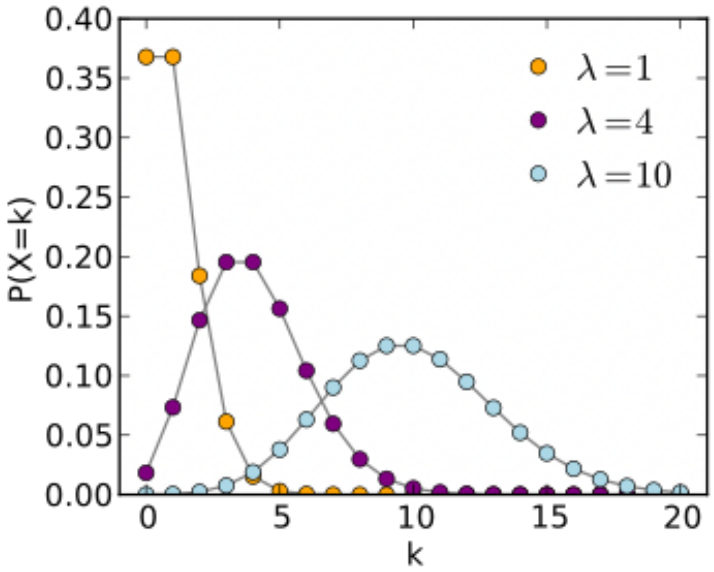
Txa = transaction that attacker broadcasts to other peers and is included in blockchain in the end.

Goal: determine likelihood that an attacker can succeed in *k* rounds of blocks

Analyzing Double Spent Attacks (1 of 2)

Modeling a solo miner

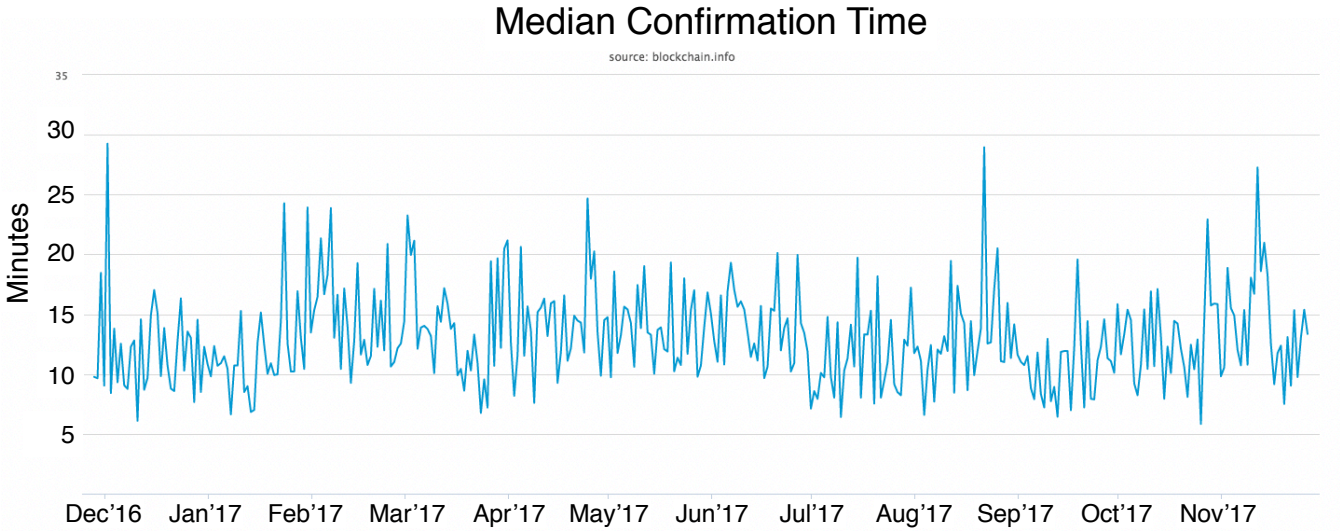
- Difficulty *D* determines the difficulty of finding a valid block
- Hash-rate *h* ⇒ *ht* hashes in time *t*
- Binomial distribution in discrete space
- Probability of any hash satisfying PoW condition is small ($\frac{1}{2^{32}D}$)
- Approximate with Poisson distribution, $\lambda = \frac{ht}{2^{32}D}$
- Time between consecutive blocks is exponentially distributed (general property of Poisson process)



$$\Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!},$$

Defeating Double Spends

- Merchant waits for a few confirmations (about 6) before delivering service
- Media confirmation time = 10 min ⇒ ~60 minutes



Analyzing Double Spent Attacks (2 of 2)

Probability that an attacker can catch up

- p = probability that an honest node mines a block
- q = probability that the attacker mines a block
- q_z = probability that attackers catches up if he is z blocks behind
- Analysis similar to "Gambler Ruin" problem
  - Exponential drop in probability as the gap increases

$$q_z = \begin{cases} 1 & p \leq q \\ (q/p)^z & p > q \end{cases}$$

Probability of a Double Spend succeeding

- We assumed the honest chain is z blocks ahead.
- Expected number of blocks attacker has mined = Poisson distribution with $\lambda = z\ (q/p)$
- Probability of attacker catching up with k blocks inserted up to time z =
  - probability that attacker inserted k blocks up to time z
  - probability that attacker catches up if he is z-k blocks behind
- Sum over all possible number of blocks that attacker could have inserted

$$\sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \cdot \begin{cases} (q/p)^{(z-k)} & \text{if } k \leq z \\ 1 & \text{if } k > z \end{cases} \longrightarrow 1 - \sum_{k=0}^{z} \frac{\lambda^k e^{-k}}{k!} (1 - (q/p)^{(z-k)})$$

Numerical calculations / HW example

10% computing power

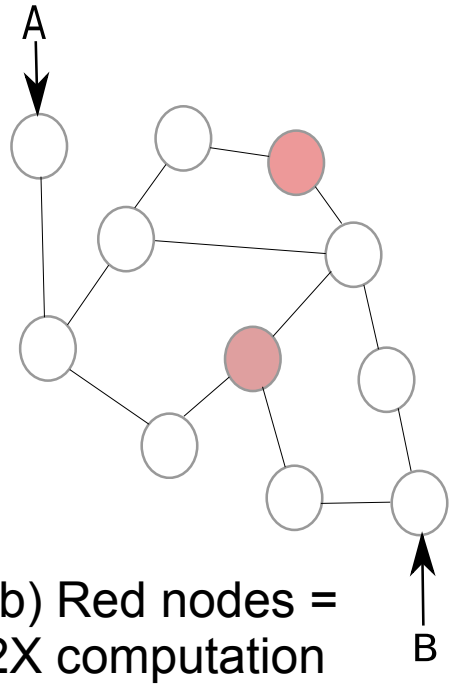
- z=0 P=1.0000000
- z=1 P=0.2045873
- z=2 P=0.0509779
- z=3 P=0.0131722
- z=4 P=0.0034552
- z=5 P=0.0009137
- z=6 P=0.0002428
- z=7 P=0.0000647
- z=8 P=0.0000173
- z=9 P=0.0000046
- z=10 P=0.0000012
- Need 5 confirmations to be 99.9% confident

30% computing power

- z=0 P=1.0000000
- z=5 P=0.1773523
- z=10 P=0.0416605
- z=15 P=0.0101008
- z=20 P=0.0024804
- z=25 P=0.0006132
- z=30 P=0.0001522
- z=35 P=0.0000379
- z=40 P=0.0000095
- z=45 P=0.0000024
- z=50 P=0.0000006
- Need 24 confirmations to be 99.9% confident

HW: use a simpler model of attack

- No secret chain, single attack, two broadcasts
- Graph = network topology (=> who gets the message first)
- (a) What is probability of A versus B in chain

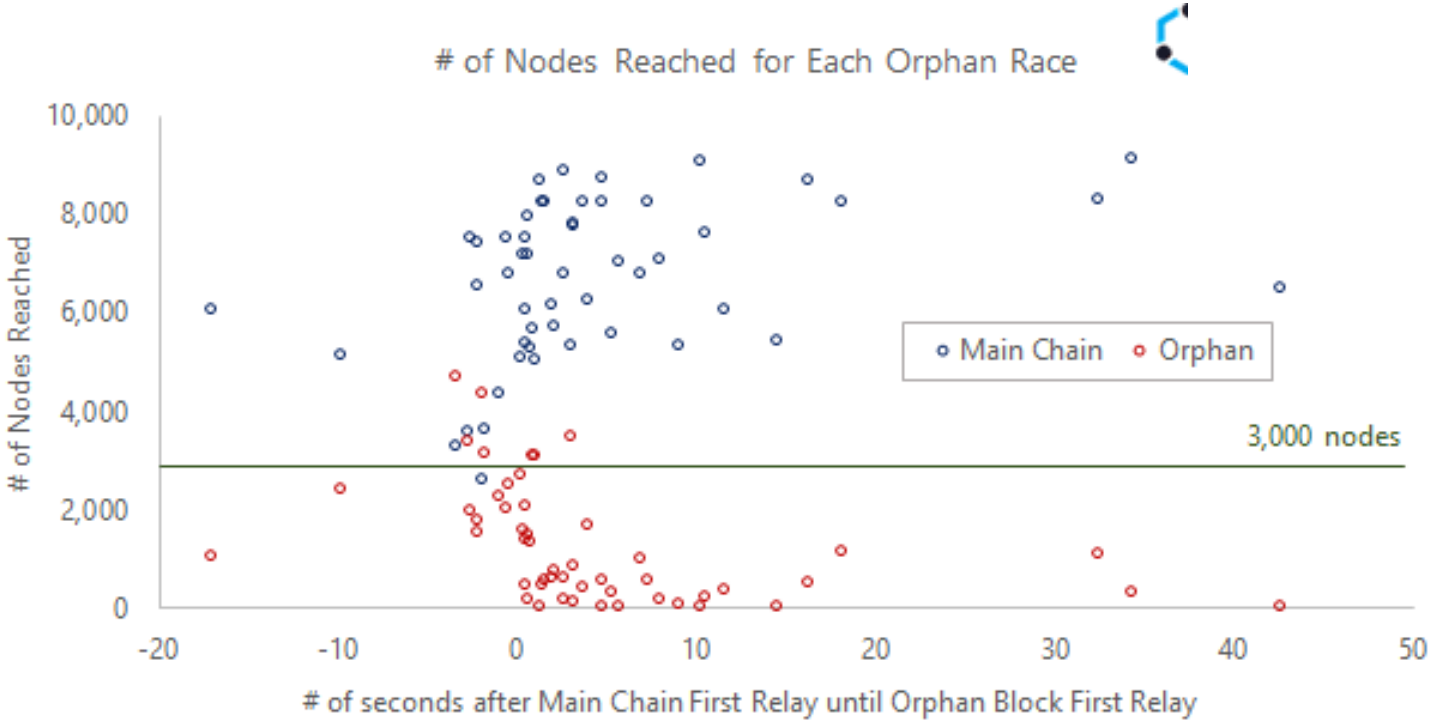


- (b) Red nodes = 2X computation

"Orphaned" Blocks

How can attackers (or miners) make sure their blocks are included in case of tie

- Orphaned blocks created when two miners produce blocks at similar times
- Alternatively, can also be caused by attacker (eg, via double spend attempt)
- "Orphan" blocks do have parents, but parent not part of longest chain
- If you are a miner/attacker, you want to "win" orphan races as often as possible
- Look at data from what blocks actually get incorporated into Bitcoin blockchain

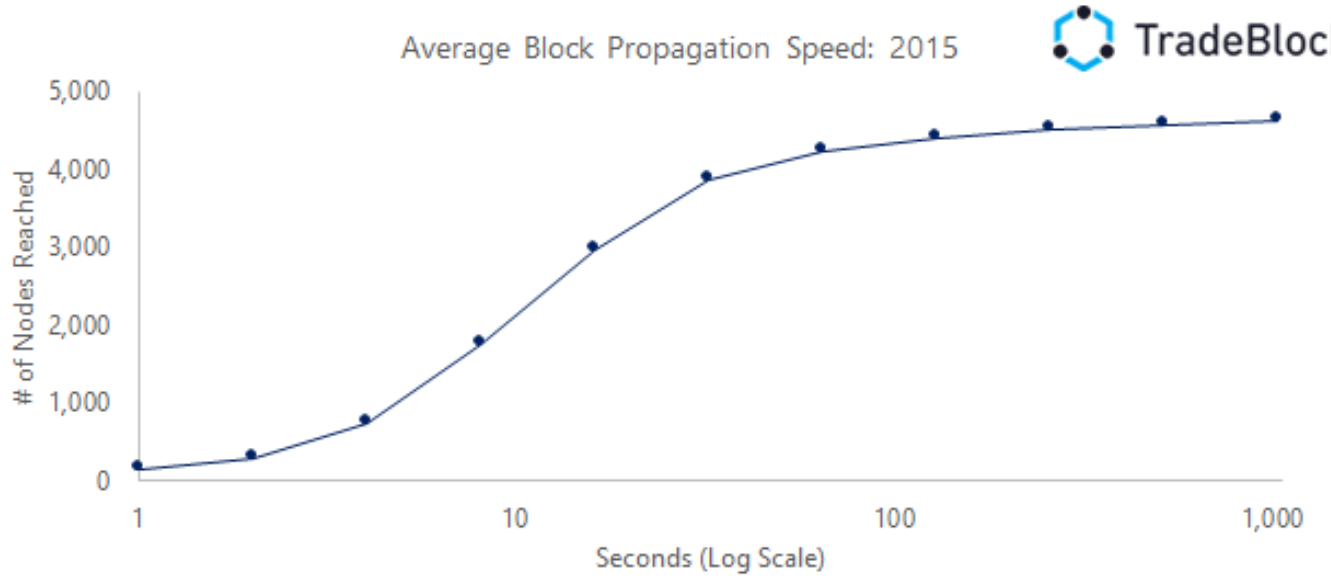


- Blocks that get relayed to many other nodes are more likely to get included in the chain
- Blocks that arrive quickly at their first relay are more likely be be included in the chain
- Lesson: helps to be in a well-connected, fast part of the network…

https://tradeblock.com/blog/bitcoin-network-capacity-analysis-part-6-data-propagation

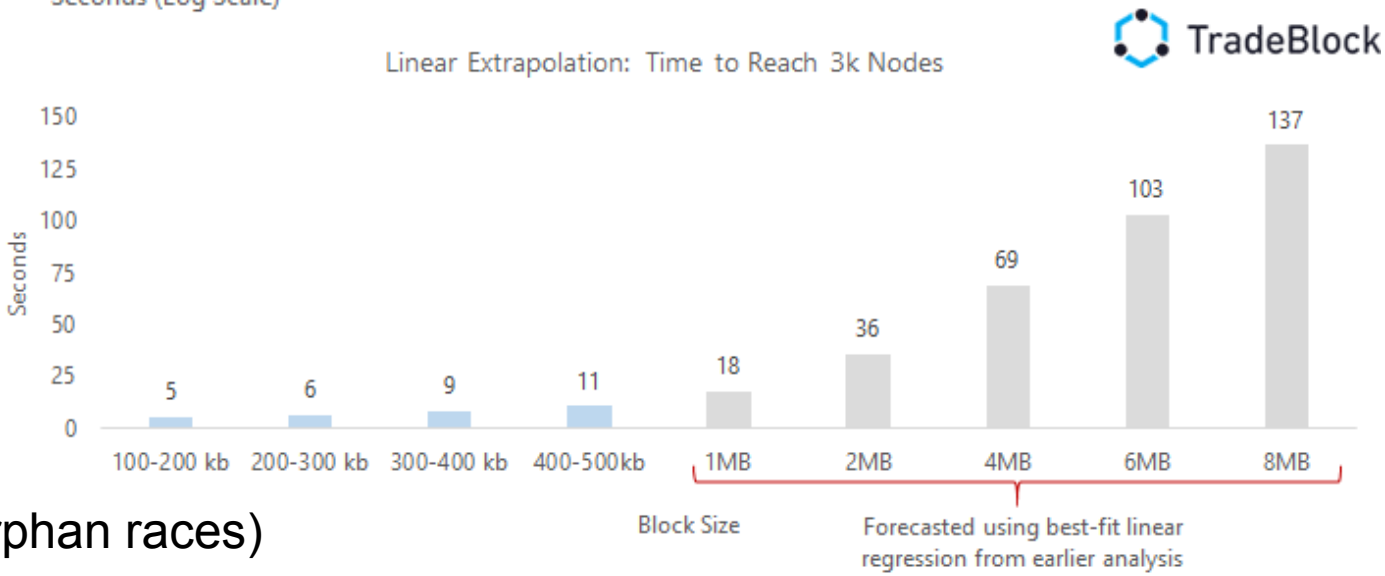
Propagation Speed and Orphan Races

- Orphan rate is roughly 1% (1.3 blocks/day)
- As blocks get larger, more latency => need to optimize # of transactions/block

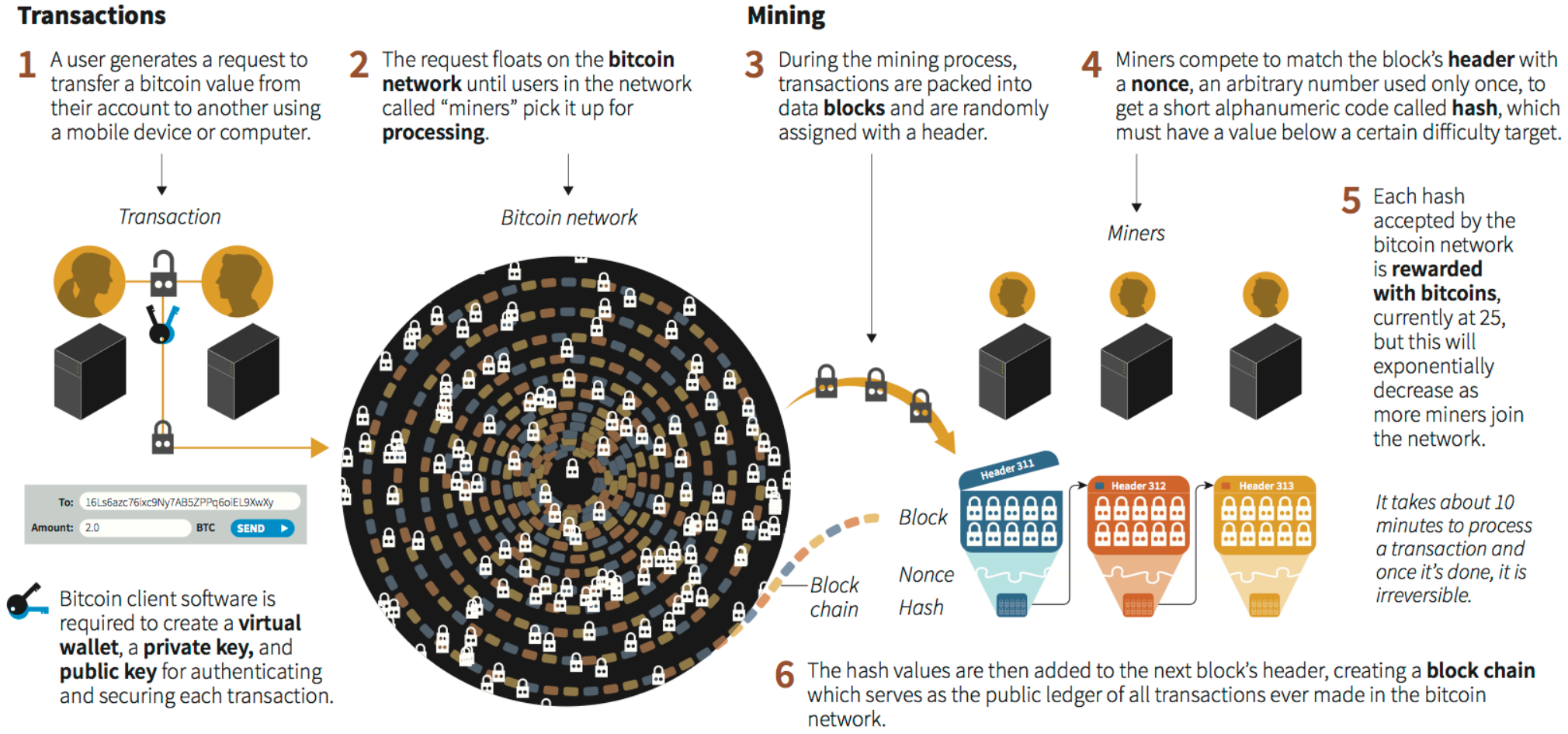


Block size implications

- Longer block sizes may be required as number of transactions increases
- But miners are not incentivized to create large blocks (will lose orphan races)



Summary: Distributed Ledger and Bitcoin



Sources: Bitcoin.org; Bitcoin Ladder

C. Inton, staff, 09/12/2015

REUTERS

Rest of the week:

- Fri: final exam review

CS 142 - Distributed Computing

Instructors: Richard Murray and Mani Chandy

PICK UP HANDOUTS AT LECTURE HALL ENTRANCES

Announcements

- HW #8 is due 6 Dec (Fri) at 5 pm; extensions until 8 Dec (Sun), 5 pm
- Final exam: out on 6 Dec (Fri) at 9 am; due on 13 Dec (Fri) at 5 pm
  - Same format as midterm (open book/notes, 2-3 hrs, take home)
  - Piazza will be frozen on 10 Dec (Tue) at ~6 pm
  - Solutions to HW #8 will be posted by 7 Dec (Tue) at ~6 pm (NLT 8 pm)
- Recitation sections this week and next
  - 2 Dec (Mon), 5-6 pm in 243 ANB
  - 3 Dec (Tue), 5-6 pm in 243 ANB
  - 5 Dec (Thu), 5-6 pm in 243 ANB
  - 9 Dec (Sun), 5-6 pm in 106 ANB
  - 10 Dec (Mon), 5-6 pm in 243 ANB
  - 11 Dec (Tue), 5-6 pm in 243 ANB