Bitcoin Group - Alex Owen, Matt Gigliotti, Eric Walker

Data Analysis

We ran tests on initial sync from to the bitcoin network from a Mid-2014 MacBook Pro, and then with a Raspberry Pi. In both cases, an SSD connected with SATA was used. For the MacBook Pro, we used a dbcache of 4GB, and for the raspberry Pi, we used a dbcache of 100MB. This is what was suggested online as the best configurations.

We averaged out the data points and organized them by year. 2009 was the beginning of bitcoin and we associate that with a very small UTXO set. Our 2016 data was from January through February. We mainly went with small sample sets to average for the Macbook Pro because the full set would have taken more than a day to process, as our best process would have been manually chopping up bits of the data, which was close to 3 gigabytes of text. The file itself was so big that our computers would stall out to even slice it.

After looking at our data, we realized that the LevelDB cache was smaller than the what we had allocated for it on the MacBook Pro. The cache size when we stopped the test was under 3GB. This would make sense that it did not really reduce the the average access time for the MacBook Pro. For the Raspberry Pi, we noticed an increase in UTXO access time between 2009 (small history of transactions) compared to 2016. This confirmed our suspicions that the larger the UTXO set, the longer it would take to access each transactions.

	Total Accesses	Average Time (µs)
🟦 2009	2,887	1.873
🟨 2009	2,887	1.138
🟦 2010	573,348	0.649
🟦 2015	2,581 (partial sample)	0.552
🟨 2016	5,970,634	1.883

🟦 Macbook Pro	🟨 Raspberry Pi

The goal of on-chain scaling is to reach Visa Scale, which is 40,000 transactions per second. Many believe that Bitcoin is a technologically superior form of money, and many times a better technology sees higher usage at its peak than the previous one. An example of this is Whatsapp and Facebook messenger, which has 3 times the volume in messages compared to SMS, and the proliferation of the internet is still not at its peak. So a somewhat conservative estimate on transaction necessities is 100,000 transaction per second because bitcoin has the potential to really help people in developing companies where Visa does not really participate as much.

So let's go through the numbers. There is a minimum of 1 UTXO per each transaction. If microtransaction outputs needed to be bundled, it would be significantly higher. So let's call it 1.5 UTXO accesses per transaction. That would mean it would take 1.883 *1.75 = 3.3 microseconds just for memory access for each transaction, which would allow for roughly 300,000 transactions per second. It is worth noting that we were only able to get to 2016 on our tests, and the UTXO set is twice as large in size now, with 50% more UTXO's. This would likely hurt the access times even further, so 300,000 could serve as an absolute upper bound to this problem.

This transaction capacity assumes 0 forks or invalid blocks, which cannot be guaranteed. Our experiments did not account at all for cryptographic verification and other processing for the transactions. The cryptographic processing is very computationally heavy in bitcoin. As we learned in class, memory access is a bottleneck to processing, and at least ⅓ of the total time spent processing within a bottleneck seems too much to complete on time. We know that the database accesses can be multithreaded, but this still seems too high given other issues mentioned above.

Therefore, we conclude the equilibrium point for a raspberry pi already synced would be much less transactions than the 100,000 likely necessary to scale bitcoin as the world's payment system. We believe that the memory access patterns must be improved significantly to have a chance at achieving the on-chain throughput desired while maintaining the current decentralization. Additionally, achieving this throughput would mean that the time to sync a node would increase exponentially, and that a raspberry pi would not be able to ever finish this process. Therefore, it seems almost necessary for 2nd-layer technologies, like the lightning network to take a significant amount of transactions off the chain.