Overview of Swirlds Hashgraph
스월즈 사 해시그래프 개요 한글 번역본- EMD

비코인(Bitcoin)이나 이더리움(Ethereum)의 원천기술인 블록체인(blockchain)에 대항하는 분산원장기술(Distributed Ledger Technology)로 1년 전에 발표된 해시그래프(Hashgraph) 백서(whitepaper) 개요의 한글(Hangul) 번역이 없어 국내 궁금한 분들 위해 저의 번역본을 포스팅합니다. 스월즈 본사와도 정확한 번역 인증을 요청할 예정이나 독자분들의 의견도 매우 환영합니다. 분산형 합의 알고리즘(distributed consensus algorithm)에 대해 관심있는 분들에게는 흥미로운 글이 될거라 생각됩니다.
This is a Korean translation of a white paper on Hashgraph, a distributed ledger technology introduced a year ago as an alternative to Blockchain, the technology behind Bitcoin and Ethereum. I will be asking Swirlds to confirm the accuracy of this translation, but I welcome any comments. This will be a good read for anyone interested in distributed consensus algorithms.

Overview of Swirlds Hashgraph
스월즈 사 해시그래프 개요

The hashgraph data structure and Swirlds consensus algorithm provide a new platform for distributed consensus. This paper gives an overview of some of its properties, and comparisons with the Bitcoin blockchain. In this paper, the term “blockchain” will generally refer to the system used in Bitcoin, rather than the large number of variants that have been proposed.
해시그래프의 데이터 구조(data structure) 및 스월즈 합의 알고리즘(consensus algorithm)은 분산형 합의(distributed consensus)에 도달할 수 있는 새로운 플랫폼이다. 본 문서는 해시그래프의 특징을 설명하며 비트코인 블록체인(blockchain)과의 차이점을 기술한다. 여기서 “블록체인”을 비트코인이 도입한 시스템으로만 국한하며 추진 중인 다양한 블록체인 파생 솔루션은 다루지 않는다.

The goal of a distributed consensus algorithm is to allow a community of users to come to an agreement on the order in which some of them generated transactions, when no single member is trusted by everyone. In this way, it is a system for generating trust, when individuals do not already trust each other. The Swirlds hashgraph system achieves this along with being fair, fast, provable, Byzantine, ACID compliant, efficient, inexpensive, timestamped, DoS resistant, and optionally non-permissioned. This is what those terms mean:
분산형 합의 알고리즘(distributed consensus algorithm)의 목표는 특정 사용자에 대한 절대적인 신뢰 없이 커뮤니티에 속한 사용자들간에 발생한 거래(transaction)의 순서에 대해 합의를 도출하는데 있다. 이로 인해, 본 시스템은 개별 사용자가 서로를 신뢰하지 않는 상황에서 신뢰를 구축하기 위해 존재한다. 스월즈 해시그래프 시스템은 신뢰를 구축함과 동시에 공정하고 신속하며 증명이 가능(provable)하고 비잔티움 (장애 허용) 보안 수준, 데이터의 ACID (원자성, 일관성, 고립성, 지속성) 준수, 효율성, 저렴성,

타임스탬프(timestamp) 형식, 서비스 거부 (DoS) 공격 저항성, 및 공개형 지정(optionally non-permissioned)이 모두 가능하다. 위 특징들에 대해 부연 설명하자면 다음과 같다:

The hashgraph is fair, because no individual can manipulate the order of the transactions. For example, imagine a stock market, where Alice and Bob both try to buy the last available share of a stock at the same moment for the same price. In blockchain, a miner might put both those transactions in a single block, and have complete freedom to choose what order they occur. Or the miner might choose to only include Alice's transaction, and delay Bob's to some future block. In the hashgraph, there is no way for an individual to affect the consensus order of those transactions. The best Alice can do is to invest in a better internet connection so that her transaction reaches everyone before Bob's. That's the fair way to compete. Alice won't be able to bribe the miner to give her an unfair advantage, because there's no single person responsible for the order.
해시그래프는 특정 사용자가 거래의 순서를 변경할 수 없기 때문에 공평하다. 주식 시장의 예를 들어, 만약 엘리스와 밥이 동시에 동일한 가격으로 남은 마지막 주식 하나를 매수한다고 가정해보자. 블록체인 시스템에서는 채굴자(miner)가 두 거래 모두 동일한 블록에 넣어 진행한다면 순서를 임의로 정할 수 있다. 또는 엘리스의 거래만 포함시키고 밥의 거래를 다음 블록에 포함시킬 수도 있다. 해시그래프에서는 이 경우 한 개인이 거래의 순서를 변경할 수 없다. 여기서 엘리스가 분산형 시스템 내 모든 사용자에게 밥의 거래 보다 먼저 전파할 수 있는 최선의 방법은 인터넷 속도에 투자하는 것이며 이로 공정한 경쟁을 도모할 수 있다. 거래 처리 권한이 특정 인물에게 위임되지 않았기에 엘리스는 채굴자에게 뇌물을 주며 부당한 혜택을 꾀할 수 없다.

The hashgraph is also fair in another way, because no individual can stop a transaction from entering the system, or even delay it very much. In blockchain, a transaction can be delayed by one or two mining periods, if many of the miners are refusing to include it. In alternatives to blockchain based on leaders, this delay can be extremely long, until the next change of leader. But in the hashgraph, attackers cannot stop a member from recording a transaction in any way other than cutting off their internet access.
아울러, 해시그래프에서는 특정 사용자가 시스템에 입력되는 거래를 저지하거나 멈출 수 없으므로 공정하다. 블록체인 시스템에서는 만약 다수의 채굴자들이 특정 거래를 블록에 포함시키는 것을 거부할 경우, 한개 또는 두개의 채굴 기간(mining period)까지 처리가 지연될 수 있다. 블록체인의 대안으로 리더형(leader-based)이 있는데, 여기서는 다음 지도자가 선정될 때까지 지연 시간이 굉장히 오래걸릴 수 있다. 그러나 해시그래프에서는 인터넷 연결을 차단하는 방법외에는 사용자의 거래를 막을 수 없다.

The hashgraph is fast. It is limited only by the bandwidth. So if each member has enough bandwidth to download 4,000 transactions per second, then that is how many the system can handle. That would likely require only a few megabits per second, which is a typical home broadband connection. And it would be fast enough to handle all of the transactions of the entire Visa card network, worldwide. The Bitcoin limit of 7 transactions per second can clearly be

improved in various ways. Though some ways of improving it, such as a gigantic block size, could actually make the fairness of the system even worse.
해시그래프는 빠르다. 해시그래프의 속도에 영향을 줄 수 있는 것은 대역폭(bandwidth)의 크기 뿐이다. 만약 각 회원이 1초에 4,000건의 거래(transaction) 다운로드가 가능한 대역폭을 사용한다면 그것이 해시그래프의 처리 속도다. 이 상황에서는 초당 몇 메가비트(Mbit)만으로도 가능하며 이는 일반 자택에 설치되는 광대역 (broadband) 연결 수준이다. 이 정도으로도 전세계의 비자(Visa) 카드 네트워크 내 모든 거래 처리가 가능할 만큼 신속하다. 현재 비트코인의 초당 7건 거래처리률은 분명 다양한 방법으로 개선해야 할 것이다. 개선 사항으로 논의되고 있는 블록 크기의 거대화 (gigantic block size) 방법은 시스템 공정성을 더욱더 훼손시킬 수 있는 단점이 있다.

The hashgraph is provable. Once an event occurs, within a couple of minutes everyone in the community will know where it should be placed in history. More importantly, everyone will know that everyone else knows this. At that point, they can just incorporate the effects of the transaction, and then discard it. So in a minimal crypto currency system, each member (each “full node” in blockchain terminology) needs only to store the current balance of each wallet that isn’t empty. They don’t need to remember any old blocks. They don’t need to remember any old transactions. That shrinks the amount of storage from Bitcoin’s current 60 GB to a fraction of a single gigabyte. That would even fit on a typical smartphone.
해시그래프는 증명이 가능하다. 거래 발생 후 단 몇분 안에 커뮤니티 내의 모든 사용자들은 그 거래가 어디에 기록되는 지 알 수 있다. 더욱 중요한 사실은 모든 사용자들이 똑같은 정보를 공유하게 된다는 점이다. 사용자들간의 합의가 성립된 후에는 거래를 최종 처리한 후 다 잊으면 된다. 예를 들어 가장 기본적인 암호화폐(crypto-currency) 시스템에서 각 회원들은 (블록체인에서는 “풀노드(full node)”라 칭함) 활성화된 지갑 (wallet) 들의 현 잔고만 저장하고, 예전의 블록이나 거래는 기억할 필요가 없게 된다. 이 결과, 비트코인의 60GB 저장공간에 비해 해시그래프는 1GB미만의 저장공간만으로도 사용이 가능하며 이는 일반 스마트폰에서도 구동이 가능하다.

The hashgraph is Byzantine. This is a technical term meaning that no single member (or small group of members) can prevent the community from reaching a consensus. Nor can they change the consensus once it has been reached. And each member will eventually reach a point where they know for sure that they have reached consensus. Blockchain does not have a guarantee of Byzantine agreement, because a member never reaches certainty that agreement has been achieved (there’s just a probability that rises over time). Blockchain is also nonByzantine because it doesn’t automatically deal with network partitions. If a group of miners is isolated from the rest of the internet, that can allow multiple chains to grow, which conflict with each other on the order of transactions. It is worth noting that the term “Byzantine” is sometimes used in a weaker sense. But here, it is used in its original, stronger sense that (1) every member eventually knows consensus has been reached (2) attackers may collude and (3) attackers even control the internet itself (with some limits). Hashgraph is Byzantine, even by this stronger definition.

해시그래프의 보안 수준은 비잔티움 (장애 허용)이다. 이는 기술적인 용어로 커뮤니티가 합의를 도출하는데 어느 누구도 (소규모 집단 포함) 이를 방지할 수 없음을 뜻한다. 또한 한번 합의가 도출된 이후부터는 그 내용을 수정할 수도 없다. 각 회원은 모든 회원간의 도출된 합의 내용에 대해 확신을 갖을 수 있는 시점에 도달하게 된다. 블록체인 시스템에서는 회원간 합의에 대해 확신(certainty)할 수 없으므로 (시간의 흐름에 따라 완전한 합의 도달 확률이 증가할 수 있는 있지만) 이는 비잔티움 합의가 될 수 없다. 아울러, 블록체인 시스템은 네트워크 파티션 (network partition)에 자동으로 대응하지 않기 때문에 비잔티움 수준이 될 수 없다. 만약 특정 집단이 다른 회원들과 인터넷 상에서 분리되어 있는 경우 이는 여러개의 블록들을 생성해 낼 수 있으며 ("forking"이라 칭함), 이는 거래의 순서에 있어 불필요한 혼란을 초래할 수 있다. 비잔티움이라는 용어는 실제 의미보다 약한 의미로 사용되기도 하는데 해시그래프는 본래의 엄격한 수준의 정의를 적용하고 있다. 해시그래프는 (1) 모든 회원들이 도출된 합의에 대해 알고 있어야 한다는 점, (2) 잠정 위해자들이 (attackers) 공모를 할 수 있다는 점, 그리고 (3) 위해자들이 인터넷 자체를 통제할 수도 있다는 (물론 제한된 의미에서의) 점 등을 인식하고 있기 때문에 엄격한 의미에서의 비잔티움을 보장할 수 있다.

The hashgraph is ACID compliant. This is a database term, and applies to the hashgraph when it is used as a distributed database. A community of members uses it to reach a consensus on the order in which transactions occurred. After reaching consensus, each member feeds those transactions to that member's local copy of the database, sending in each one in the consensus order. If the local database has all the standard properties of a database (ACID: Atomicity, Consistency, Isolation, Durability), then the community as a whole can be said to have a single, distributed database with those same properties. In blockchain, there is never a moment when you know that consensus has been reached. But if we were to consider 6 confirmations as achieving "certainty", then it would be ACID complaint in the same sense as hashgraph.
해시그래프는 ACID를 준수한다. 이는 데이터베이스 용어로서, 해시그래프를 분산형 데이터베이스(distributed database)로 운용할 경우 적용된다. 커뮤니티 내 회원들은 거래(transaction)가 발생한 순서에 대해 합의를 구할때 본 지침을 적용한다. 합의가 도출된 후, 각 회원이 갖고 있는 데이터베이스의 로컬 복사본 (local copy)에 합의된 추가 거래 내용을 입력하게 된다. 만약 모든 로컬 데이터베이스가 ACID (원자성, 일관성, 고립성, 지속성) 표준을 준수할 경우, 커뮤니티 전체가 하나의 분산형 데이터베이스를 보유하고 있다고 명명할 수 있다. 블록체인에서는 완전한 합의가 도출된 시점을 알 수 있는 방법이 없다. 여섯번 정도의 확인과정을 통해 알아내는 정도가 해시그래프와 비슷한 수준에서의 ACID 준수라고 볼 수 있을 것이다.

The hashgraph is 100% efficient, as that term is used in the blockchain community. In blockchain, work is sometimes wasted mining a block that later is considered stale and is discarded by the community. In hashgraph, the equivalent of a "block" never becomes stale.
해시그래프는 블록체인 커뮤니티에서 언급되는 100% 효율성을 갖고 있다.

블록체인에서는 추후에 탈락(stale block) 되거나 커뮤니티에서 채택하지 않을 블록때문에 시간을 허비하는 경우가 있다. 해시그래프에서는 블록이 절대 탈락되지 않는다.

The hashgraph is inexpensive, in the sense of avoiding proof-of-work. In Bitcoin, the community must waste time on calculations that slow down how fast the blocks are mined. As computers become faster, they'll have to do more calculations, to keep the rate slow. The calculations don't have any useful purpose, except to slow down the community. This requires the serious miners to buy expensive, custom hardware, so they can do this work faster than their competitors. But hashgraph is 100% efficient, no matter how fast its "blocks" are mined. So it doesn't need to waste computations to slow itself down. (Note: there are blockchain variants that also don't use proof-of-work; but Bitcoin does require proof-of-work).
해시그래프는 작업-증명(proof-of-work) 절차가 필요 없기 때문에 저렴하다. 비트코인에서는 불필요한 계산과정때문에 시간이 지연되는 경우가 있다. 앞으로 컴퓨터들의 성능이 좋아질 수록 이러한 과정들은 추가가 될 것이다. 이런 것들은 일의 처리를 지연시킬 뿐 별다른 도움이 되지 않는다. 이로인해 서로간의 경쟁에서 이기고 싶은 채굴자들은 고가의 특별 주문제작된 하드웨어를 구매해야 할 지도 모른다. 하지만 해시그래프는 블록들이 처리되는 속도가 빨라져도 영향을 받지 않기 때문에 효율적이다. 다시 말해 일처리의 속도를 줄이기 위해 컴퓨터 연산을 낭비할 필요가 없다는 의미다 (참고: 블록체인 중에는 작업-증명을 사용하지 않는 플랫폼도 존재하나; 비트코인은 작업-증명 작업을 요구한다.)

The hashgraph is timestamped. Every transaction is assigned a consensus time, which is the median of the times at which each member first received it. This is part of the consensus, and so has all the guarantees of being Byzantine and provable. If a majority of the participating members are honest and have reliable clocks on their computer, then the timestamp itself will be honest and reliable, because it is generated by an honest and reliable member, or falls between two times that were generated by honest and reliable members. This consensus timestamping is useful for things such as smart contracts, because there will be a consensus on whether an event happened by a deadline, and the timestamp is resistant to manipulation by an attacker. In blockchain, each block contains a timestamp, but it reflects only a single clock: the one on the computer of the miner who mined that block.
해시그래프는 타임스탬프(timestamp) 형식이다. 모든 거래에는 합의 시간(consensus time)이 배정되며 이는 각 회원들이 해당 거래를 처음 수신한 시간들의 중앙값(median)이다. 합의 도출을 위한 절차인 타임스탬프 작업은 비자티움 수준을 보장하며 또한 증명이 가능(provable)하다. 만약 참여하는 회원들의 대다수(majority)가 정직하고 신뢰할 수 있는 컴퓨터 시간을 입력 했다면 정직하고 신뢰할 수 있는 회원들이 생성한 공동의 시간표이므로 (참 시간이) 타임스탬프(timestamp) 내 한개 또는 두개 시점(times) 사이에 존재한다고 볼 수 있고 그 결과 정확성과 신뢰성이 보장된다. 합의 타임스탬프 형식은 마감시간 내에 사건이 실제로 발생했는지에 대한 합의가 정립되며 공격에 취약하지 않기 때문에 스마트 계약(smart contracts) 등의 환경을 개발하는데에

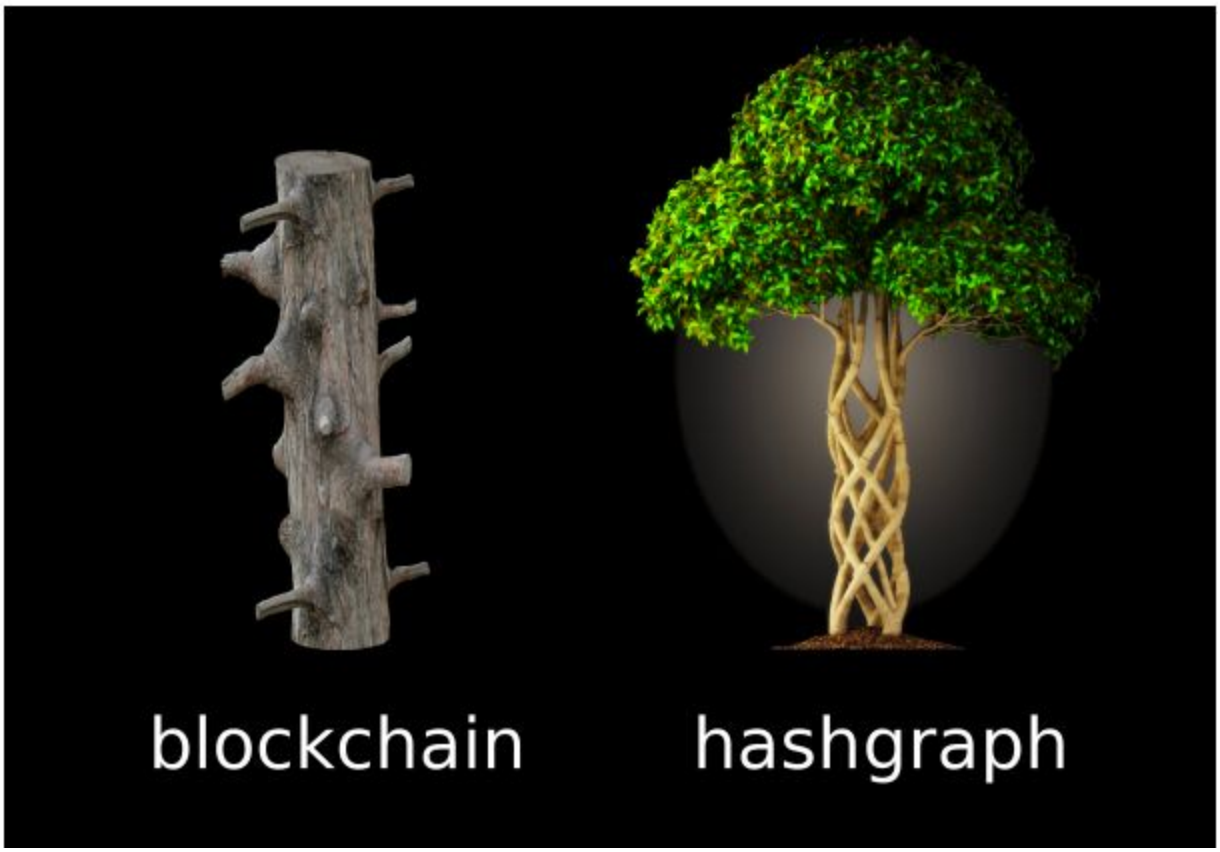
유용하다. 블록체인 시스템은 각 블록에 타임스탬프가 있지만 채굴을 처리한 컴퓨터 시계만 반영된다.

The hashgraph is DoS resistant. Both blockchain and hashgraph are distributed in a way that resists Denial of Service (DoS) attacks. An attacker might flood one member or miner with packets, to temporarily disconnect them from the internet. But the community as a whole will continue to operate normally. An attack on the system as a whole would require flooding a large fraction of the members with packets, which is more difficult. There have been a number of proposed alternatives to blockchain based on leaders or round robin. These have been proposed to avoid the proof-of-work costs of blockchain. But they have the drawback of being sensitive to DoS attacks. If the attacker attacks the current leader, and switches to attacking the new leader as soon as one is chosen, then the attacker can freeze the entire system, while still attacking only one computer at a time. Hashgraph avoids this problem, while still not needing proof-of-work.
해시그래프는 서비스 거부 공격(Denial of Service: DoS) 저항(resistant)이 가능하다. 블록체인과 해시그래프 모두 서비스 거부 공격을 저항할 수 있도록 분산되어 있다. 한명의 회원 또는 채굴자에게 데이터 패킷(packet)을 플러딩(flooding)하여 인터넷 연결을 임시로 차단할 수는 있으나 커뮤니티 전체는 정상적으로 진행될 것이다. 시스템 전체를 공격하기 위해서는 상당수의 회원들을 대상으로 패킷 플러딩을 해야 하는데 이는 실행하기가 까다롭다. 한때 블록체인의 대안으로 리더 (leader) 또는 라운드-로빈 (round robin) 방식을 기반으로 한 시스템들이 출시된 적이 있었다. 이것들은 블록체인의 작업-증명 비용을 줄이기 위해 출시 되었으나 서비스 거부 공격에 (DoS) 취약하다는 단점이 있었다. 만약 공격자가 현재 리더를 공격하고 새로운 리더가 선정되자 마자 그 공격을 다음 리더에게 이동시킨다면 한명만 공격하면서도 시스템 전체를 마비시킬 수 있다. 해시그래프는 작업-증명 단계 없이도 이 문제를 피할 수 있다.

The hashgraph is optionally non-permissioned, while still avoiding the cost of proof-of-work. A permissioned system is one where only trusted members can participate. An open system is not permissioned, and allows anyone to participate. Standard blockchain can be open if it uses proof-of-work, but variants such as proof-of-stake typically have to be permissioned in order to be secure. A hashgraph system can be designed to work in a number of different ways. One of the more interesting is to use proof-of-stake, allowing members to vote proportional to their ownership of a particular cryptocurrency. A good cryptocurrency might be widely used, so that it is difficult for an attacker to corner the market by owning a large fraction of the entire money supply. If a large fraction of the currency owners all participate in a hashgraph system, then proof-of-stake will make it safe from Sybil attacks, which are attacks by hordes of sock-puppet fake accounts. Such a system would be secure even if it were not permissioned, while still avoiding the cost of proof-of-work.
해시그래프는 공개형 (non-permissioned) 시스템으로 설정이 되어있고 작업-증명의 비용을 없앴다. 허가형 (permissioned) 시스템은 승인을 받은 회원만 참여가 가능하고 공개형 시스템 (open system)은 승인절차를 요구하지 않고 누구든지 참여가 가능한 커뮤니티이다. 일반적인 블록체인은 작업-증명 (proof-of-work)을 사용하면 공개화로의

전환이 가능하지만 지분-증명 (proof-of-stake)과 같은 파생 솔루션들은 허가형 (permissioned)으로 운용 되어야 안전하다. 해시그래프 시스템은 다양한 적용이 가능하도록 설계되었다. 그 중 흥미로운 점은 지분-증명을 사용하는 것인데 이는 특정 암호화폐의 소유자들에게 보유하고 있는 화폐량에 비례하는 투표권한을 부여하는 구조이다. 새로운 화폐가 생성되는 과정에서 시장의 상당 지분을 통제하여 전체 시스템을 공격하기 어렵도록 가급적 많은 사람들이 사용하는 암호화폐(cryptocurrency)가 좋은(good) 암호화폐이다. 만약 다수의 화폐 소유자가 해시그래프 시스템에 참여한다면 지분-증명의 도입을 통해 허위의 다중계정(sock-puppet fake accounts) 집단이 전개하는 시벌(Sybil) 공격으로부터 보호 받을 수 있다. 이런 시스템은 공개(non-permissioned)형으로 전환해도 보안상 안전하며 작업-증명의 비용을 없앨 수 있다.

The following figure illustrates why hashgraph has these desirable properties.
다음 그림은 어떻게 해시그래프가 이렇게 바람직한 특성들을 갖췄는지 보여준다.



Why does hashgraph have these properties? Because it's like a tree that's braided, not pruned.
왜 이런 특성을 갖고 있나? 해시그래프는 가지치기(pruned)를 하지 않고 줄기가 잘 묶여진(braided) 나무와 같기 때문이다.

In both blockchain and hashgraph, any member can create a transaction, which will eventually be put into a container (the "block"), and will then spread throughout the community. In blockchain,

those containers are intended to form a single, long chain. If two miners create two blocks at the same time, the community will eventually choose one to continue, and discard the other one. It's like a growing tree that is constantly having all but one of its branches chopped off.
블록체인과 해시그래프 시스템 모두 특정 회원이 거래를 생성할 수 있고 이는 컨테이너(container) ("블록")에 저장되어 커뮤니티 전체에게 전파된다. 블록체인에서의 컨테이너 개념은 한개의 긴 사슬(a single, long chain)을 형성하도록 되어있다. 만약 채굴자들이 두개의 블록을 동시에 생성("forking"이라 칭함)한다면 커뮤니티는 둘 중 하나만 선택하고 나머지를 버려야 한다. 비유하자면 하나를 제외한 모든 가지들을 계속 쳐내야 하는 나무를 키우는 것과 같다.

In hashgraph, every container is used, and none are discarded. So all the branches continue to exist forever, and eventually grow back together into a single whole. This is more efficient. Furthermore, blockchain fails if the new containers arrive too quickly, because new branches are sprouting faster than they can be pruned. That is why blockchain needs proof-of-work or some other mechanism to artificially slow down the growth. But in hashgraph, nothing is thrown away. So there is no harm in the structure growing quickly. Every member can create transactions and containers whenever they want. So it is very simple, and tends to be very fast.
해시그래프에서는 버리는 것 없이 모든 컨테이너가 사용된다. 그 결과, 나무의 모든 가지는 영원히 공존하며 전체를 하나로 묶도록 설계되어 있다. 이게 더 효율적이다. 아울러, 블록체인에서는 만약 새로운 컨테이너가 너무 빨리 도착하면 가지를 치기 전에 새로운 가지가 생성되므로 시스템이 붕괴(fail)된다. 이러한 이유로 블록체인은 작업-증명 또는 비슷한 원리를 사용해 인위적으로 성장 속도를 늦춰야 한다. 하지만 해시그래프에서는 버려지는 것이 없다. 따라서, 시스템이 빠르게 성장하면서 생기는 위험이 없다. 모든 회원은 언제든 필요시 새로운 거래 및 컨테이너를 생성할 수 있기 때문에 매우 단순하면서 빠른 속도를 자랑할 수 있다.

Finally, because the hashgraph doesn't require pruning, it is simpler, which allows more powerful mathematical guarantees, such as Byzantine agreement and fairness. Distributed databases such as Paxos are Byzantine, but not fair. Blockchain is neither Byzantine nor fair. But the Swirlds hashgraph is both Byzantine and fair.
마지막으로 해시그래프에서는 가지치기가 필요 없기 때문에 더 단순하고, 이로 인해 비잔티움 합의(Byzantine agreement)와 공정성과 같은 강력한 수학적 보증(mathematical guarantee)을 가능케 한다. Paxos와 같은 분산형 데이터베이스는 비잔티움이긴 하나 공정하지 못하다. 블록체인은 비잔티움도 아니며 공정하지도 않다. 하지만 스월즈 해시그래프는 비잔티움 합의와 공정성을 모두 갖고 있다.

출처: http://www.swirlds.com/downloads/Overview-of-Swirlds-Hashgraph.pdf