Abstract—The Internet has many backbone components on top of which the whole world is connected. It is important to make these components, like Border Gateway Protocol (BGP), Domain Name System (DNS), and Public Key Infrastructure (PKI), secure and work without any interruption. All of the aforementioned components have vulnerabilities, mainly because of their dependence on the centralized parties, that should be resolved.

Blockchain is revolutionizing the concept of today’s Internet, primarily because of its degree of decentralization and security properties. In this paper, we discuss how blockchain provides nearly complete solutions to the open challenges for these network backbone components.

Keywords—Blockchain, Internet, BGP, DNS, PKI.

Anzahl Sätze—Parça zinciri ile PKI, BGP ve DNS iyileştirmeleri: Sistematik Bir İnceleme

Parça zinciri ile PKI, BGP ve DNS iyileştirmeleri: Sistematik Bir İnceleme

I. INTRODUCTION

The first design of the Internet was presented as a centralized single entity. With time, the Internet was divided into sub-systems (e.g., DNS, BGP, PKI). Still, these components were built on a centralized architecture. This introduces security, privacy, and performance issues such as a single point of failure, trust issues, high latencies, and storage [1]. This results in these centralized services getting hacked frequently. Efforts were put in to make the services distributed [2]. These improved the Internet by solving the above-mentioned challenges but introduced new types of issues like scheduling, resource allocation, coordination, device management, scalability, security, trust, and multiple weak points of contact for attackers [2]. The overall performance was decreased because of the replication of work, backups, and the communication of distributed parts of the overall system [2]. Recently, blockchain-based solutions were introduced, with the goal being improving security while keeping the speed, cost and correctness comparable to the legacy components presented in Table 1. We can see that the current protocols take less time and cost (as most of the vendors have already implemented them) and achieve the desired (correct) results. For example, DNS protocol will always give the IP (Internet Protocol) address of the domain name, unless it malfunctions or is attacked. But the security of these components is vulnerable to the attacks and should be improved as described in sections III to V.

A blockchain is a public distributed ledger that can record transactions that are connected using a cryptographic hash function [3]. The basic functionality provided by a blockchain is a secure mechanism for storing and obtaining data, ordered by the timestamp of each record in the data, in a publicly verifiable and immutable manner. For that reason, in most of the system architectures, blockchain provides a storage mechanism for data collection and consensus among participants. In general, blockchain can help in (1) decentralization, (2) provenance and immutability of data, (3) security, and (4) heterogeneity and programmability.

Our contributions: In this paper, we first overview the blockchain technology, then provide a discussion of three widely-employed Internet components (PKI, BGP, DNS) and their security vulnerabilities, afterward showing a detailed explanation of the available blockchain-based solutions, and conclude with a summary and open issues.

II. BLOCKCHAIN TECHNOLOGY

A blockchain is a decentralized, distributed, and public digital ledger that records data in the form of transactions across multiple devices to enforce immutability, except with a very

Table I: Legacy Protocol Performance

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Time per query</th>
<th>Security</th>
<th>Correctness</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNS</td>
<td>0.048s ÷</td>
<td>Needs improvement</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>BGP</td>
<td>8.38s for 100% propagation</td>
<td>Needs improvement</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>PKI</td>
<td>Within few milliseconds</td>
<td>Needs improvement</td>
<td>High</td>
<td>Low</td>
</tr>
</tbody>
</table>

1https://wp-rocket.me/blog/test-dns-server-response-time-troubleshoot-site-speed
2http://www.circleid.com/posts/how_a_routing_prefix_travels_through_the_internet
3https://blogs.technet.microsoft.com/option_explicit/2012/04/19/validating-a-certificate/
small probability of the adversary controlling a large fraction of the processing power, stake, etc. A blockchain can be split into network, consensus, storage, view, and side planes [4], enabling researchers to work on a single idea while improving the overall blockchain infrastructure.

Hash Function: Blockchain is built upon collision-resistant deterministic hash functions that map an arbitrary-length input to a fixed-length n-bit output. The hash function should have the property of collision resistance, meaning that an adversary cannot find two different inputs mapping to the same output in polynomial time. This property is important for the integrity and immutability of blockchain.

Transactions and Blocks: The 'transaction' term was first used by Bitcoin [3], where a transaction contains the amount of Bitcoin value transferred between entities and information of the sender and the receiver. Generally, a transaction is data or information of the variant type and can be created by any participant.

Blocks are created by the verifiers (miners). A block is a set of approved transactions, along with a timestamp and a hash pointer to the previous block. The first block of a blockchain is called the Genesis Block. The genesis block is almost always hardcoded with a verifiable universal fact and does not refer to a previous block. As shown in Figure 1 hashes of all transactions are kept in a Merkle tree for efficient memory management [5]. For example Hash0 is the hash of transaction Tx0, Hash1 is the hash of transaction Tx1, Hash01 is the hash of Hash0 and Hash1, and eventually we have the Merkle root hash Tx_Root. The block also contains the hash of the previous block (prev_hash). Thus, it is computationally infeasible to modify or tamper with the contents of the previous blocks, as this would require finding the hash of all of the remaining blocks to keep the chain connected.

Types of Blockchains. In a blockchain, entities can be readers or writers (e.g., Hyperledger [7]), enabling researchers to work on a single idea while improving the overall blockchain infrastructure.

Consensus Protocols. Blockchain presents a solution for the environment where the parties do not have to trust each other and collaborate. As there is no universal trusted third party, each blockchain has to have a consensus protocol for reliability and consistent state of the network.

Proof of Work (PoW): In PoW, a node can get its block accepted if it can solve a cryptographic puzzle (hash) and spend some computational resources in the process. It was first implemented by Bitcoin [3].

Proof of Stake (PoS): A node is randomly selected depending upon the stake/resources (ether in Ethereum [8]) she has. Then its block is accepted to be appended to the chain.

Byzantine Fault Tolerance (BFT): In Practical BFT (PBFT) [9], there is a leader election (where each entity participates) to elect an entity that has authority to add a new transaction in the chain. This protocol assumes that there more than 2/3 of the honest participants. In Delegated BFT (DBFT) [10], participants, by voting, pick the delegate they support. The selected delegates, through the BFT algorithm, reach a consensus and generate new blocks.

Problems. Blockchain presents a new perspective on Internet security, but the traditional Bitcoin blockchain (which is used as the solutions discussed in the upcoming sections show) still has issues regarding huge power, energy, storage, and communication requirements [11] (current Bitcoin blockchain is around 200 GB). Moreover, the Bitcoin blockchain is secure assuming that more than 50% of the hashing power belongs to the honest parties in the system. There are further attacks on the incentive mechanism, such as selfish mining [12].

III. PKI

Earlier, anyone could pretend to be anybody over the Internet as none of the Internet layers verifies the identity of the entity over the network. This created privacy and trust issues. As a solution, Secure Socket Layer (SSL) and Transport Layer Security (TLS) were introduced. The concept is to provide data integrity, confidentiality, and authenticity using a public/private key pair. But these keys can be compromised. Then another idea was introduced to create cryptographic identities known as digital certificates. Digital certificates contain the identity and public keys of the entity to be used for encryption and authentication. The problem remained that it is very easy to create a digital certificate by self-signing it. There was a need of having a trusted third party that can provide these certificates, known as a Certificate Authority (CA). CA is a trusted entity that issues digital certificates that verify a digital entity’s identity over the Internet. The infrastructure to manage, store and distribute these certificates is called the Public Key Infrastructure (PKI).

A CA signs a certificate to bind the public key of a server to its identity. Then SSL/TLS uses these certificates to authenticate the web-server. Trusting the CA, the browser obtains the server’s public key to establish a secure connection. There are two types of certificate authorities ROOT-CA and SUB-CA. The certificate is trusted if it is signed by ROOT-CA. The browsers or operating systems come with many ROOT-CA public keys stored in their databases. As ROOT-CAs might
be limited in number and become bottle-neck when there is a lot of demand, there are SUB-CAs that can sign the certificates. PKI works on Chain of Trust. To authenticate a certificate, the browser (or any other entity) checks whether or not the certificate is signed by a valid ROOT-CA. If the signer is a SUB-CA, the validation continues in a chain up to the ROOT-CA. Recent research indicates that CAs can be dishonest, get attacked, or can be using faulty or outdated cryptographic algorithms [13]. Which affects the security of PKI.

Currently, there are two types of approaches for the security of PKI. Log Based PKI: Highly available servers are appointed for publishing and secure monitoring of the certificates to ensure that CAs do not behave maliciously. Still, there are issues with log-based solutions like revocation explained in [14]. Web of Trust (WoT): is a decentralized approach. Users can put their trust in another entity by signing their certificate. Then each trusted entity keeps a certificate that contains signatures of the users that trusted it, in addition to its public key.

Limitations of PKI security and the advantages of using blockchain to enhance the security and efficacy of PKI is given by [15]. The architecture assumes the existence of blockchain-backed PKI and uses it to secure the critical (rich) credentials. A privacy-aware PKI system based on blockchain was presented by [16]. The paper claims that there are many instances (like anonymous social forums), where the entity does not want to reveal its identity. Current PKI leaks this information by knowing which key is used in the protocol. The paper uses blockchain to have online and offline keys and encryption to hide the identity of the user. [17] proposes Etherium-based blockchain technology to build secure PKI systems, resolving the issues of log-based PKI and the WoT approaches. Blockchain resolves the single point of failure issue and the need for a newcomer to prove its trustworthiness.

Certificate Transparency (CT) was introduced by [18], using an append-only public log, to improve the accountability of CAs. As certificates are publicly recorded in the log servers, a fraudulent certificate can be detected, and the countermeasures can be taken to handle the potential attack. Though this removes the central authenticating entity [18], considering the increasingly huge number of current certificates, it may introduce computation and communication burden on both clients and servers [19]. Moreover, it is shown that split-world attack can be performed on CT [20], where the attacker presents different views of the log to successfully impersonate as the victim. Another issue with CT is that it is not a privacy-preserving scheme [21]. Certificate Revocation is another phenomenon that reduces the efficacy of these solutions as stale certificates can be used by attackers [20].

[22] gives an approach to detect the man-in-the-middle (MITM) attack happened/happening to a victim client. The concept of notary nodes is used, where the server connects and requests the observation of its certificate. In [23] public notaries are used by the client, hence replacing the dependence on the web browsers/operating systems to validate the certificates. In [22], the MITM can be detected but cannot be prevented, and in [23], the users still have to rely on the notaries (which can all be compromised).

[13] proposes a blockchain-based solution to construct certificate transparency. The certificates are published as transactions in a global blockchain by the web servers, which is downloaded by browsers. To verify a certificate, the browser just has to see if the certificate is in the blockchain. As the certificates of the authorities are also published, the CAs are also publicly accountable. This creates more trust and decreases the possibility of having a fraudulent CA. Each certificate has a period of validity and can be revoked at will by omitting it from the next block and hence putting into a certificate revocation list (CRL). In this solution, compromised CA can be detected two ways. One is that other CAs will not approve of its transactions. Secondly, the target server will only publish the certificates from valid CAs and the fraudulent certificates will not be appended to the global chain. As the attacker fails to impersonate the target server, a wrong certificate will not be publicized even if the Publishing Key Pair is compromised. In the case both PK and CA are compromised, the countermeasures are taken within a period before the certifiers certify the change. Still, the attacker may prevent the browser from obtaining the blockchain, which is difficult in such a distributed setting. Second, there can be forks introduced but unless more than 1/3 of the certifiers are malicious, then this is not possible either [13]. However, this scheme has the following drawbacks identified in [20]:

1) An adversary can use unexpired transactions of the revoked certificate to impersonate the victim server; this is a type of a man-in-the-middle attack.
2) The proposal is inefficient in terms of storage and has large headers.
3) The proposal depends on the CAs to publish revocation information of the certificate to the blockchain but the compromised CA might not issue Certificate Revocation information to the public.

CertLedger [20] provides a solution that is resilient to split-world attacks, does not depend on CA for the certificate revocation, and preserves the privacy of the clients.

Open Issues: There are still some issues like what is the incentive for the certifiers? What if honest nodes are compromised after joining? What happens when the key pair of the server is compromised?

IV. BORDER GATEWAY PROTOCOL (BGP)

Autonomous Systems (AS) are responsible for the routing among their network typologies. Typically, an AS represents a collection of IP prefixes to which data is routed [24]. Sometimes these ASes require a flow of data among themselves to reach the destination not present within individual networks. This data flow is peer-to-peer (P2P) in its nature, and is done through the Border Gateway Protocol (BGP) [25]. In BGP, an AS announces the IP prefix of all the IPs reachable through itself, together with path delay metrics. All the other Internet Service Providers (ISPs) behind ASes update their routing tables according to the BGP announcement. Although efficient in practice, ASes assume that their neighboring ASes behave honestly and propagate correct routing information (without having the global knowledge). However, the interests of ASes
may conflict, and this weak notion of trust can be breached by malicious ASes (e.g. prefix hijacking [24]).

Previous efforts have developed many solutions like Secure BGP (S-BGP) [26], Secure Origin BGP (SoBGP) [27], Inter-domain Route Validation (IRV) [28] and Path-End Validation [29]. Each of them depends upon a central trusted entry, which has a huge cost of management and is complex. Also, these solutions depend upon the PKI, which has the vulnerabilities described in the previous section. Another problem is the lack of adoption of these protocols as they introduce costly additional infrastructure for their operation. To avoid the cost of adoption, ASes and ISPs are reluctant to migrate towards these solutions despite the known security threats and their clear benefits.

BGP is open to different kinds of attacks. In BGP route manipulation attack, an adversary manages to change the BGP table to disrupt the traffic of the Internet. Whereas in BGP route hijacking, an attacker AS announces the prefixes belonging to the victim. As a result, the traffic is re-routed to or through the attacker AS. An attacker can also send malicious or faulty BGP traffic to a victim. The victim exhausts its resources to handle the traffic and is left incapable of processing valid BGP traffic. In Route Leak, as the result of a potential attack or the AS malfunction, an AS issues incorrect information about the IP addresses on their network. This results in inefficient routing and failures for the traffic.

BGP route hijacking is most dangerous and can be classified into two types: partial attack and complete attack [30]. The partial attack occurs when an adversarial AS announces an identical IP prefix as that of the victim AS. Attack on Youtube in 2008 [24] is an example of partial hijacking. In a complete attack, the adversary AS announces more specific prefixes than the target AS. Since the default forwarding is based on the longest prefix matching, ASes switch to more specific prefixes and start sending the packet through that route. [30]. Figure 2 explains the difference between the attacks. In the partial attack, an attacker would announce the 208.65.153.0/24, which is already announced by AS1. Since the two announcements are the same, when any other AS receives the announcement, it can either switch to it or continue with the old routing path. In the complete attack, an attacker would announce 208.65.153.128/25. This IP has a longer prefix match than 208.65.153.0/24 in the respective finger (for example see finger table of AS3 in Fig. 2b) tables so other ASes would switch to this route. Interestingly, Youtube used the same concept as a legitimate way to get back the traffic in the 2008 attack [30]. Some BGP attack examples include a global route leak in November 2017, a country-wide Internet outage in Japan due to BGP issues in August 2017, and possible financial traffic malfunction, an AS issues incorrect information about the IP addresses on their network. This results in inefficient routing and failures for the traffic.

Blockchain-based Solutions: One effort towards securing BGP was to use the PKI presented by IANA [31] to sign the routes. This scheme is problematic as to effectively use PKI, ASes set up a route assigning authority called Routing Origin Authorization (ROA) which is very costly to implement. Also, as we constructed that PKI infrastructure has security vulnerabilities (section III), signed BGP update messages based on PKI signatures in BGPSec do not result in a secure path verification protocol.

Considering the mentioned problems and the properties of blockchain discussed in Section II, [31] gives a structure of a system that uses blockchain for the better security and performance of BGP. Blockchain helps in the following ways:

1) All transactions occur between peers without any intermediary (like ROA). Also, there is no need for a third party to authentication, eliminating the possibility of tampering or spoofing by a malicious entity.
2) Provides announcement immutability and re-traceability of the chain of BGP routes. Also, a route is validated by multiple parties and is more trustworthy.
3) The authors argue that the blockchain should be different from the Bitcoin blockchain and its properties should depend upon the nature of the use case.

The paper does not provide an implementation and do not discuss the possible attacks and the prevention.

The authors in [24] present a clique-based BGP architecture, RouteChain, to secure BGP against both complete and partial attacks. The method distributes the ASes into subgroups. The system has a global blockchain, and each subgroup has its own private-permissioned blockchain. The main purpose of this sharding is to reduce the storage overhead of having only one global chain and to decrease the transaction validation which is critical for the timely detection of a potential attack. The ASes are grouped based on their geographical proximity for low delays. All groups select a leader to randomly announce the local routes to the global chain. They use local collaboration among ASes to prevent the complete attack. Whenever there is an update, all the ASes in the group check if their path changes with the update. If it does, they observe the original path and its corresponding prefix. Next, they locate the true owner of the prefix through the global blockchain. If the new update does not belong to the true owner, then the update is discarded. The paper claims that the consensus in the partial attack scenario is achieved in 200 milliseconds, and it is 54.23 seconds for the complete attack. Comparing this timing with the attack on Youtube hijacking incident (in which within 20 minutes, 97 ASes were hijacked) RouteChain asserts that the system will notify the ASes about the attack while it is in its initial stages [24]. The protocol runs on top of the current BGP architecture, which makes it adaptable and economical.

Open Issues: In BGP, different ASes have different policies

for sharing the route with the other AS or not. For example, if AS1 does not want its traffic to go through AS2, then AS1 will advertise the path that does not include AS2 even if the resultant path is longer. How can blockchain solutions cater to these policies? Can different blockchain architectures be used to solve the BGP security issues?

V. Domain Name System (DNS)

The domain name system is used for the resolution of the global domain names. It has a distributed hierarchical design with one root server and 13 specialized servers operated by agencies within a few parts of the world. This design made the system simple, scalable, and flexible. The (in)security of DNS leads to many advanced attacks on DNS including DNS spoofing/cache poisoning, DNS hijacking, and DNS rebinding. A DNS DDoS attack in October 2016 brought down many of the websites including Netflix, Twitter, and CNN [39].

DNSSec is crucial for providing origin authentication and message integrity to communication between the user and the name server. In DNSSec, the root server sends the certificate containing the public key of the next name server along with the IP address of the next server and the hash of the entire message (for data integrity). The next server does the same until the IP address for the domain name is found. This protocol solves the problem of authentication and integrity but still suffers from various attacks namely IP fragmentation and DDoS. The most major problem is its slow adoption.

Blockchain-based DNS Alternatives: Namecoin [34] was built with the motivation of removing managing domains to avoid trust in a single entity. Namecoin uses different prefixes to store and map other types of name-value pairs. For example, the “d/” prefix is used for domain names and “id/” is used to register identities. It further uses the virtual .bit top-level domain name that is not officially registered in the current DNS system. This means Namecoin is isolated from the DNS system and users have to install additional resolving software for resolution of the .bit domain names. Just like DNS, Namecoin provides complete functionalities for registering, renewing, and transferring a domain. The developers modified Bitcoin blockchain to store name-value data like transactions, still utilizing the PoW-based mining mechanism for consensus. Namecoin has shown to have some security flaws: for instance, it was found that a single miner consistently had more than 51% of the total computing power on the Namecoin network [33]. In another instance, a Namecoin bug allowed people to steal names from anyone [33]. Performance-vise, [33] experienced a latency spike and throughput drop due to software issues of Namecoin.

Blockstack [33] combines a DNS system with PKI and purely works with the Bitcoin blockchain. To improve the efficiency of the Bitcoin blockchain for handling a large amount of name-value pairs, a separate logical layer, namely virtual chain, is proposed that works on top of the blockchain to maintain the naming system while the underlying blockchain is only used for achieving consensus on the state of the DNS (or any naming system in general) and the integrity of the name-value data records. Blockstack has significant improvements over Namecoin as it increases the data storage capacity considerably and the virtual chain improves the maintenance of the system.

In [39], authors use blockchain to improve the security and performance of DNSSec. The solution decreases the number of keys needed for DNSSec for easy key management and reduction in DNSSec response size. They use the X509 Cloud blockchain network that is used to store X.509 certificates. This allows us to speed up the verification process.

DecDNS [35] gives a blockchain based data storage model for DNS. They also build multiple DNS nodes for further decentralization and addressing single-point-of-failure in DNS resolution. They report 0.006025s response for parallel domain name resolution, which satisfies the DNS performance requirements and shows the potential of using the architecture in the real network for domain name resolutions.

Open Issues: There should be compatibility between the traditional and the blockchain-based DNS architectures for the systems using the conventional methods for DNS, as this leaves performance and security concerns. For example, a query from the blockchain-backed DNS system can be misinterpreted or discarded by a system using a legacy DNS mechanism. How can blockchain-based and traditional DNS mechanisms inter-operate and integrate? This area of the combination of blockchain-based and traditional decentralization mechanisms for name resolution requires more research.

VI. Comparison

We finish with an analysis of the mentioned blockchain-based solutions using three main parameters. Security: The level of integrity, confidentiality, and availability introduced by the proposed solutions. Performance: The performance of these solutions compared to the legacy solutions. Resource Constraints: The speed, storage, and cost constraints involved. Table IIa gives an overview of the proposed solutions per-protocol regarding the analysis parameters.

In Table IIb, we identified the key security aspects between blockchain in general and the currently deployed protocols. Blockchain provides Data integrity by the use of hashing in block construction backed by the public key cryptography (as discussed in Section II) to create trust between trustless entities under the same system. The blockchain excels in availability as the ledger is distributed globally on different nodes offering robustness against the single point of failure problem. Further, blockchain provides fault tolerance, i.e., even if some participants leave the system, fail, or get attacked, the blockchain system is not affected as each participant has (more or less) the same copy of the ledger.

Looking at the drawbacks, we can see that the contents of the records in blockchain are transparent and costs confidentiality to the users considering to opt for blockchain solutions. Furthermore, as discussed in [11], blockchain requires lots of storage and computational resources, which come at a high cost and decreased scalability. Furthermore, the blockchain can be attacked (for example using selfish mining [12]). With the ever-increasing use of the Internet, this issue needs to be solved using better optimizations of currently available blockchain systems, such as LightChain [42].
Table II: Overall analysis of blockchain based solutions

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Solutions</th>
<th>Hashed Data</th>
<th>Security</th>
<th>Performance</th>
<th>Resources</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>PKI</td>
<td>[32], [33], [34], [35], [20]</td>
<td>TLS Certificates</td>
<td>Improved</td>
<td>Comparable</td>
<td>Storage and computational costs</td>
<td></td>
</tr>
<tr>
<td>DNS</td>
<td>[36], [37], [38], [39]</td>
<td>Domain Names</td>
<td>Improved</td>
<td>Comparable</td>
<td>Same as of blockchain</td>
<td></td>
</tr>
<tr>
<td>BGP</td>
<td>[40], [31], [41]</td>
<td>Domain Routes</td>
<td>Improved</td>
<td>Comparable</td>
<td>Storage and speed</td>
<td></td>
</tr>
</tbody>
</table>

VII. CONCLUSION

Blockchain is an emerging solution for improving the security of the overall structure of the Internet to ensure the interruption-free performance of the network. In this paper, we presented blockchain-based solutions for three core network components and systems: PKI, BGP, and DNS, and identified open problems. We plan to further extend our cryptographic discussion with the knowledge we gained from [43].

ACKNOWLEDGEMENTS

The authors acknowledge TÜBİTAK (the Scientific and Technological Research Council of Turkey) 119E088 grant.

REFERENCES

[34] “Namecoin,” https://namecoin.org/.