Introduction of NDN with Comparison to Current Internet Architecture based on TCP/IP

Dnyanada P. Arjunwadkar
Computer Engineer
University of Pune
Maharashtra, India

ABSTRACT
Over the years number of people using Internet has escalated. The primary motive of the Internet has been revised. To deal with this drastic change, network architecture has to be redesigned considering future trends. The future architecture of Internet can be developed only after scrutinizing various aspects of current Internet architecture. Thus study of strengths and weaknesses of existing architecture will guide the learners to build a robust future architecture of Internet. This paper represents comparative study of TCP/IP model of existing architecture and Name Data Networking approach of Content Centric Networking of proposed model. This paper discusses fundamental factors such as approaches to current and future architectural model, Packet formats and differences in security mechanisms of these models.

General Terms
Network Architecture and Design, Named Data Network

Keywords
Named Data Network, Content Centric Networks, Future Internet Architecture

1. INTRODUCTION
Initially, when evolution in computing had begun computing resources were limited. Use of computers was restricted to high priority military operations and research institutions. Computing resources such as tape drives, storage disks or non computing resources like important documents, scientific findings and research papers often had to be shared. In order to share such resources, communication between the two computing machines was essential. With this objective a complex network of computers located across different geographic locations was developed. Success of sharing the data using a network of computers and ease of access to computing resources encouraged the Universities and Corporate Companies to develop their own networks. These heterogeneous physical networks could communicate with each other using the TCP/IP. Inter-connection of these networks lead to the invention of the Internet.

The use of Internet has drastically changed from mere communication to information dissemination. Staggering amount of data is generated every day. People are using Internet to store their content online to be able to access their content from anywhere in the world. People use Internet to find information about job vacancies, news articles, work related data, online tutorials etc in text, video or audio formats. A recent article in Forbes sheds light on how the popular TV series 'Game of Thrones' was downloaded 1.5 million times. Which lead to transfer of 2,000 terabytes of data within 12 hours after its telecast [1].It is a known fact that once a video goes viral, it receives large number of requests from different parts of the world simultaneously. To be able to satisfy such high number of requests simultaneously, implementation of latest multicast algorithms is required. NDN architecture handles such situations effectively as it is multipoint to multipoint protocol unlike TCP/IP. Internet model of TCP/IP is based on the conversation between two machines. This conversation requires addresses of source and destination machine. Today's internet model is not best suited for information dissemination. The proposed model of NDN is implemented using Content Centric Networking approach. In this architecture the content is named. The objective of NDN architecture is to efficiently handle the issues that current internet architecture faces. This future architecture promises better utilization of bandwidth, aims to improve throughput and decrease the network traffic generated during the transfer of popular web content. In this paper detail comparison of existing TCP/IP model and proposed NDN model is given.

This paper will aid the reader to comprehend fundamental differences between current internet architecture and future internet architecture. This Paper represents a comparison of two architectures in a simplified manner. It includes basic functioning approach and architectural components of both the systems. It emphasizes on distinguishing factors between TCP/IP and NDN model. Packet formats and Security implementations are also discussed in this paper. By the end of this paper reader will have a clear understanding of how these architectures differ.

2. COMPARITIVE STUDY

Table 1. Comparison

NDN	TCP/IP
Future Internet Architecture	Current Internet Architecture
Information Distribution	Information Sharing
Information Centric Network	Conversation Oriented
Content Centric	Address Centric
Elimination of DNS	Can't Function without DNS
Not Host Centric	Host Centric
Multipoint to Multipoint	Point to Point
Large Scale Information Dissemination	Inefficient Information Dissemination
Router Content Cache	No Router

In-Network Storage	Content Cache No In-Network Storage
Optimization of Bandwidth Congestion Reduction Improved throughput	No Optimization of Bandwidth Often Congestion Occurs
Stateful Data Plane, Adaptive Forwarding	Stateless Data Plane Non Adaptive Forwarding by Router
3 Entities Maintained	1 Entity
FIB, PIT, CS	FIB
FIB Stores Multiple Hop Status, Performance Information	FIB Stores Only Next Hop Information
Existing Routing Protocols Propagation based On Name Prefix	Existing Routing Protocols Propagation Based on IP prefix
Security is Provided to Content Itself Not Using Abstractions	End to End Channel is Secured like SSL
Interest Initiated Model	Client Server Model for Interaction
Content Distribution (many users REQUESTING SAME DATA AT DIFFERENT TIME) , Multicast(SAME TIME) both Handled efficiently	Inefficient Content Distribution

2.1 Components of Future Internet Architecture

•**Name:** It represents the interest expressed by the consumer specifying file name and format.

•**Content:** It is the requested file.

•**Consumer:** One who requests for content.

•**Producer:** One who generates the content.

•**Interest:** It is a request for a specific file by the consumer. Consumer requests for a content using Name.

•**Data Packet:** It contains the content requested along with the name of that content.

•**Node:** A device in the network implementing NDN concepts.

•**Interface**: Connection of node to link.

•**Router:** In NDN functioning of routers is more than routing a packet from consumer to producer. It has to keep track of incoming interests of consumers, data packet fetched to respond to the incoming interests and maintaining the cache, ephemeral in-network storage.

•**FIB:** Forwarding Information base is maintained by each node in a network. It has information about route entries based on name prefixes.

•**CS:** Content Store is an ephemeral cache maintained at each node in a network. It is capable of storing recent responses (data packets) and interests. The size of this storage could be different in different routers.

•**PIT:** Pending Interest table. If the cache maintained does not have the data for the expressed interest, it stores the data name requested and the information about the interface where the interest arrived.

2.2 Approach in NDN model

NDN is the future Internet architecture developed to keep up with the ever increasing amount of content being generated and distributed over the Internet. NDN is an instance of Content Delivery Network. It emphasizes on the 'What' data the consumer is interested in and not 'Where' it is stored. Data could be stored where it was generated or in the cache of nearest node in a network. The physical address of the nodes where data is stored is not required in NDN unlike TCP/IP. Since the physical address is not essential for making communication possible, there is no need of DNS to map names to IP addresses. In this architecture there are 2 types of packets: Data packet and Interest packet.

When the consumer expresses his interest by specifying the name of the file i.e. Content name, this interest is forwarded in the network based on the name of the content. Each intermediate node in the network has 3 entities associated with it as mentioned earlier. On receiving interest, node first performs the name-based lookup of the requested content in the Content Store. If this node has the name of the content requested, then it responds with the data packet right away. This in-network storage permits the node to satisfy the request locally. If there are multiple consumers requesting for the same content (often when a video goes viral on the Internet, multiple requests for the same content are received by the server which hosts that content) an intermediate node which has the copy responds. Thus there is no need to send multiple requests for the same content along the common channel up to the node which holds the content. By avoiding sending similar requests upstream, bandwidth is minimized. This optimization of link will reduce downstream latency [2]. But if content store does not have the copy of the requested content, then PIT is checked. If PIT already has entry requesting for the content then intermediate node records information about the interest arriving interface.

Whenever a response is received, data packet is forwarded to all such requesting interfaces and corresponding entries are deleted from PIT. Existing PIT entry suggests that interest has already been forwarded upstream by an intermediate node. Therefore, duplicate interest is not forwarded. If PIT does not contain an entry for expressed interest then incoming interest interface along with the name of the content and outgoing interface where Interest is forwarded are recorded in PIT and then these interests are routed in a network based on name prefix without the knowledge of source or destination address.FIB table is similar to IP routing table. It has information about name prefixes and interfaces where it can be forwarded. These interfaces lead to source which has the desired data. Multiple interfaces can be present for single name prefix. Thus an interest is forwarded to all possible paths and data can be retrieved from multiple paths [3].

2.3 Components of Current Internet Architecture

Protocols involved in TCP/IP: TCP, IP, UDP, ICMP.

• **IP:** Internet Protocol (IP) provides necessary information for routing of packets in a network. IP is not a reliable

protocol. To provide reliability, it has to team up with TCP. IP performs fragmentation, only if a network has defined restrictions on the size of datagram.

- **TCP:** Transmission Control Protocol (TCP) is Connection oriented protocol. It guarantees the delivery of the TCP packets, making it a reliable approach.TCP protocol is implemented in process-to-process communications.TCP packets are encapsulated with additional information about addresses and IP datagram is constructed.TCP participates in a Three-way Handshaking Process. In order to ensure delivery of Datagram it keeps track of sequencing and acknowledgments.

- **UDP:** User Datagram Protocol (UDP) is connectionless transport protocol. It does not guarantee delivery of packets. UDP is an unreliable approach. Thus it is used in applications such as video streaming where failing to deliver a single packet will not affect the application. It does not perform Three-way handshaking process which results in less overhead. It does not keep track of Sequencing and acknowledgments.

- **ICMP:** Internet Control Message Protocol (ICMP) is implemented to send control messages between devices in a network. Control messages such as Host not reachable, Port not reachable, Redirect messages to control traffic, Source Quench message to control incoming traffic, message to notify expiry of Time to Live etc.

Other components include:

- **IP routers:** they have only one data structure i.e. FIB.

- **IP FIB:** Forward information base is useful to take switching decisions based on an IP address prefix match. It contains the information about the single outgoing interface i.e. Next hop information only.

- **Buffer memory of IP:** as soon as packet is forwarded, it is flushed out of buffer memory using MRU technique.

2.4 Approach in TCP/IP Model

The 4 layer TCP/IP model has IP: Internet Protocol placed in its network layer. It deals with packaging data into datagram that will travel independently in a network. IP does not keep a track of such datagram. Addressing and routing of these datagram is carried out in the network layer of TCP/IP. With the knowledge of the destination's IP address, these datagram take different routes to reach the destination. Datagram reached at destination may be out-of-order or duplicated. IP is connectionless and unreliable protocol. For reliability IP is paired with reliable protocols such as TCP [4]. Transmission Control Protocol from the transport layer of TCP/IP model looks after the sequencing of all such datagram. TCP protocol establishes one-to one connection, providing reliable service. It sends the acknowledgement upon receiving the datagram so lost datagram can be detected and sent again. Traffic is controlled by TCP. The application layer of TCP/IP supports higher protocols such as DNS which maps the name to its physical address. As without the physical address the connection cannot be established in TCP/IP.DNS uses TCP for critical and bulk queries .IP handles the datagram routing based on this physical address and datagram routing. While TCP would be responsible for higher level functions such as providing reliability and error detection. This combination became known as TCP/IP.

2.5 NDN Packet Format

In TCP/IP point to point path is established based on knowledge of source and destination addresses. Path connecting the two points is the path along which packets are delivered [5]. But in NDN approach any intermediate node which has the copy of requested content can respond. (It emphasizes on the 'What' data the consumer is interested in and not 'Where' it is stored.) Thus in NDN, concept of point to point data delivery does not exist as these end points involved in the connection cannot be determined beforehand.

NDN packets do not have fixed length headers. It helps to minimize processing cost of packets. Thus packets of very small size can be transferred without the overhead. It gives the packets, flexibility. Instead of fixed length headers the design uses the TLV format to provide the flexibility of adding new types. This feature is capable of dealing with the scenario in the future where older types may get discontinued as the protocol evolves over time. This is an added advantage over TCP/IP.TLV stands for Type Length Value. Based on type field Data packets and Interest packets are distinguished.

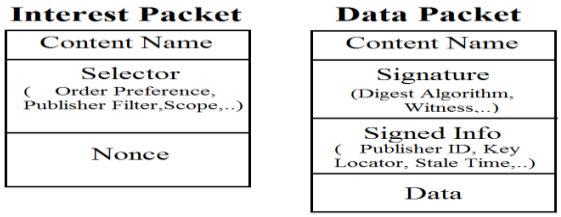

Fig 1: NDN Packets [6]

Interest packets are composed of two essential components which are **Content Name** and **Nonce**. Content Name has specified name of requested data. Interest packets are uniquely identified by the combination of Name and Nonce. Nonce is generated at random by the consumer. It is used to distinguish between two different consumers requesting for the same content. At times it may happen that consumer is sending his interest repeatedly. This can also be identified with the help of Nonce. Re-issuing the interest by consumer suggests that the interest has not responded yet. Thus next time router receives same interest from same consumer, router forwards this interests on different interfaces. Nonce also helps to identify an interest earlier forwarded which has looped back.Thus looping Interests are destroyed.

In addition to these essential fields there exist multiple optional fields such as **Selectors, scope, Interestlifetime** which define the behaviour of interest packets. For instance selector field is used for discovering and selecting the Data that matches best to expressed interest. The scope field defines how far interest packet can travel. The combination of Name and Nonce should uniquely identify an Interest packet. This is used to detect looping.

Data packet consists of Content Name, metainfo, Content (data), Signature.

Content Name: NDN content has a hierarchical name consisting of a sequence of name components. Naming conventions in NDN are followed in a way that only globally used entities are required to have globally unique names. Otherwise locally identifiable entities can have local names for local context. A two level nested TLV is used to represent a **Name**. Name is first element and signature is last element of Data Packet.

Metainfo field contains additional information about content type, freshness period, finalblockid etc. Content type can be set to default (=0), LINK (=1), and KEY (=2). The default type suggests that actual data bits are identified by name of data. The next type of content LINK relates to another name which is also used to identify actual data content. The next type of content, KEY is a public key. Freshness period is denoted by nonnegative number. It is an optional field. It is useful for replacement in content store if storage runs out of memory space. If freshness period is expired then corresponding data is marked as stale. Stale data is also a valid form of data. The expiration of freshness period takes into consideration a possibility of generation of newer version of the same data. The finalblockid is an optional field which gives information about the final block in the sequence of fragments.

Data: The Data packet represents some arbitrary binary data (held in the Content element) together with its Name.

Signature: With the help of Signature, generated content is linked to producer of that content. Signature provides more info about the publisher. If the source producer can be verified and trusted, each data packet signed by that producer can also be trusted. Therefore when data packets are retrieved from nearest router cache, they can be trusted based on their signatures. Thus instead of securing the connection between source and destination NDN attempts to secure each individual Data Packet by signing it.Signature in NDN is defined as two successive TLV blocks which are **signatureinfo** and **signaturevalue**. Signatureinfo tells more about the description of signature, information that could be used to acquire parent certificate. Signatureinfo is included in signature calculation. Signaturevalue is not included in signature calculation. It represents actual bits of the signature.

2.6 TCP/IP Packet Format

TCP is one of the widely used protocols in current internet architecture. It moves the data in a continuous byte stream which is suitable for bulk data transfer over the network. Full Duplex and reliable service is also useful for interactive data applications. IP header provides all the information that is helpful for routing. It supplies **source** and **destination IP Addresses**, **Time to Live** so that the undeliverable datagram are destroyed, type of service to be provided which is used to decide handling of the datagram based on factors like precedence, delay during transport.

IP headers also include field such as **Protocol** which indicates which other protocol IP is paired with. It is used to indicate TCP, UDP or ICMP protocol that is used in transport of Datagram. As mentioned earlier IP performs fragmentation if there is a restriction on size of datagram. In such situations fields like **Offset** and **Total Length** is useful.

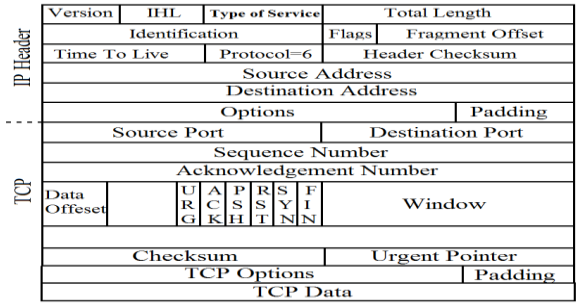

Fig 2: TCP/IP Packet [7]

As stated earlier TCP provides a reliable connection using Three-way handshake also with the help of sequence numbers and acknowledgements. Thus there exists number of fields that support various services such as sequencing, acknowledgement of delivered packet, establishment of end to end channel etc. TCP is process to process communication protocol. Thus Ports help to identify processes that will communicate with each other. **Sequence Number** will help to packets to be delivered in sequence and **Acknowledgement Number** will ensure they have reached destination. **RST** and **FIN** used for tear down process. Checksum field is useful for error checking.TCP packets are encapsulated with IP datagram as shown in the diagram.

2.7 Security in NDN

One of the important aspects where current and future internet architecture differs is the security. In NDN dynamic Content Cache having a copy of requested content responds to the interest. Source of content received could be different from where it was initially produced. Thus it is very important to provide security to content. Attempt is made to secure the content itself instead of securing the connection through which content travels. Security NDN is content centric. The content can be secured from unintended audience by implementing encryption mechanism. Encryption will help to maintain confidentiality. Encrypted data can be decrypted with valid keys. Without these keys no intermediate node can gain access to data. Thus Access to data is controlled. Security of content has the following properties [8].

- Provenance: It will determine origin or source of data.
- Validity: this property will address concerns such as whether the received copy corrupted? Is the copy received complete? Etc.
- Relevance: It will determine if data is relevant to interest expressed.

The relevance of the name of the content and content itself is essential. The consumer should get the content that they have requested. Thus there is the mechanism of keys which bind the content to its name by signing it. Sometimes this signature includes information about content producer. This additional information is in the form of key locator. It helps to determine provenance i.e. Origin of the data. It provides trust in data. User can rely on the signed data received. Integrity of data can be trusted if content is signed. Therefore, it is said that this signature securely binds together the tuple –

< Name, Content, Publisher's Key > – authenticating that the data is what its name purports it to be [9]. Verifying signature in pieces of content can be time consuming. Thus verification of all content objects is expensive.

2.8 Security in TCP/IP

When evolution of the Internet began, the main motive was to connect two heterogeneous networks. Security was never a concern as the number of users was limited. Today goal of providing security over the Internet is of prime importance. Today internet users are exposed to multiple security threats. To tackle with them following mechanisms are implemented.

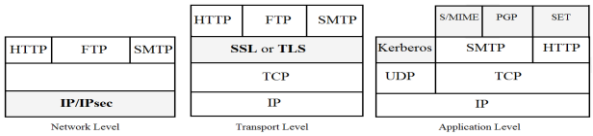

Fig 3: TCP/IP packet

• **IPsec** IPsecurity is the suite of protocols providing security at the network layer. This is applied between host-to-host, network-to-network or host-to network to enhance security. With the help of security protocols such as Authentication Header and Encapsulating Security Payload it offers stronger authentication and encryption techniques. IPsec operates in two modes transport and tunnel mode. Transport mode focuses on an end to end security by protecting payload. Tunneling mode focuses on VPN by providing protection to payload, header and routing information.

AH protocol furnishes authentication, data integrity and protects against relay attacks. But it does not provide Confidentiality. In addition to what AH offers, ESP provides confidentiality. It uses various cryptographic techniques to achieve data integrity and authentication. Security Association (SA) is integral component of IPsec architecture. Security Association contains data needed for IPsec to function. It includes IP address of source, authentication keys, encryption key, key lifetime etc. Security association is unidirectional i.e. Two separate associations are required for inbound and outbound packet transfer. Key management is done by protocols like Internet Key Exchange. IPsec was designed for ipv6 but can be used for systems using ipv4.

• **HTTPS** Hyper text transfer Protocol is an application layer protocol which works with Secure Socket Layer protocol of transport layer. This internet security standard was subsequently knows as Transport Layer Security (TLS). Whenever HTTPS Protocol is applied a secured connection is formed which is noticeable from the URL 'https://'. Such secured connections are useful for bank transaction or where data protection is priority. HTTPS encrypts data flow in communication between client and web server. It uses public key encryption to secure the path which guarantees message integrity. When user wants to initiate a data sensitive operation, it sends a request to web server. Web server invokes SSL. Client and web server agree upon certain security parameters by participating in handshaking process. Web server then authenticates client by sending certificate and if the client trusts server that process continues. The communication path is secured using a symmetric key which is generated by encrypting session key with public key of server. Session key is generated on client side [10]. This protocol protects data in transition by securing path only. But once data reaches destination responsibility to protect data depends on other processes. S-HTTP is another protocol used for encrypting web communication. It is used when part of data needs to be encrypted and HTTPS is used when most of the information is to be securely transmitted. In such cases HTTPS encrypts the entire communication channel. HTTPS is widely deployed to provide security over internet.

• **SET** Secure Electronic Transaction was developed for conducting card transactions for VISA, MasterCard in a safe and secured environment. It includes authentication of customer and merchant. Transactions are carried out without revealing card details. This is possible because of Dual Signature. Dual Signature has information about order information for merchant and payment information for banks etc. in SET confidentiality is provided using encryption based on DES. SET ensures integrity of data transferred by using RSA signatures and SHA-1 hash codes. X.509v3 digital certificates are used for authentication in SET. Privacy of participants is maintained. It does not disrupt the functioning of other security protocols such as IPsec, SSL etc.

3. CONCLUSION

The idea behind this paper was to present the comparison of two Internet architectures. In this paper differences in current and future Internet architecture are given which will help the reader to learn concepts of new architecture by comparing it with current architecture. This comparison will help the reader to understand the future Internet architecture effectively. Future Internet architecture is developed by eliminating flaws of current Internet architecture. Therefore to study the future architecture it is recommended to know the fundamentals of existing architecture provided in this paper. This information will be useful for learners studying different architectures of Internet.

4. ACKNOWLEDGMENTS

Thanks to Dr. Parikshit N. Mahalle for his valuable suggestions and advice for this paper.

5. REFERENCES

[1] http://www.forbes.com/sites/jaymcgregor/2014/06/17/game-of-thrones-season-finale-becomes-most-pirated-show-in-history/s.

[2] Van Jacobson, Diana K. Smetters, James D. Thornton, Michael F. Plass, Nicholas H. Briggs, Rebecca L. Braynard , **'**Networking Named Content' (Palo Alto Research Center, Palo Alto, CA, USA)

[3] Mishari Almishari, Paolo Gastiz, Naveen Nathan, Gene Tsudik, 'Optimizing Bi-Directional Low-Latency Communication in Named Data Networking'.

[4] TCP/IP Protocol Suite fourth edition by Behrouz A. Forouzan.

[5] "NDN specification Documentation, Release 0.1a2, NDN Project Team, March 27, 2014."

[6] http://named-data.net/wp-content/uploads/ndn_packet.png.

[7] http://www.cisco.com/web/about/ac123/ac147/ac174/ac196/about_cisco_ipj_archive_article09186a00800c8417.html

[8] Diana Smetters and Van Jacobson, Palo Alto Research Center, 'Securing Network Content'.

[9] Katie Shilton, University of Maryland, College Park. Jeff Burke University of California, Los Angeles. Kc claffy CAIDA/UC, San Diego .Charles Duan University of Colorado Law School .Lixia Zhang, University of California, Los Angeles, 'A World on NDN:Affordances & Implications of the Named Data Networking Future Internet Architecture'

[10] 'ALL IN ONE – CISSP, EXAM GUIDE' fourth edition by Shon Harris. (CISSP, MCSE).