In-network Caching Scheme in Information Centric Networking

Anjali
Research Scholar, Department of Computer Science and Applications,
Panjab University, Chandigarh 160014, India

Abstract— This Paper presents Information Centric Networking (ICN) a.k.a Named Data Networking (NDN) or Content Centric networking (CCN) which is the advance research on networks and also it is a Future Internet Architecture (FIA) for data dissemination in a very efficient manner. This project initiated in 2010 and more care about data rather than host. So ICN concentrates on What content is required by the user rather than Where that content resides. IP Architecture works on Host-Centric Network architecture and ICN works on Data-Centric Network architecture. For the Data Dissemination, IP used only one type of packet i.e. IP Packet and in ICN it uses two different packets, one for request i.e. Interest Packet and other for response i.e. Data Packet. ICN have several features like Naming scheme is different from the existing architecture, More Security is given to the content rather than the location where data resides, Ubiquitous Caching technique for faster retrieval of data, Different Routing and Forwarding Algorithm, Leave Copy Everywhere concept and many more. This Paper basically includes the Architecture of ICN, Information related Different Packets, Caching technique i.e. where to cache the data, Centrality method, Content Replacement Algorithm.

Keywords— Named-Data Networking (NDN), Content Centric Networking (CCN), Information Centric Networking (ICN), Caching, Future Internet Architecture (FIA), Leave Copy Everywhere (LCE).

I. INTRODUCTION

(CDN) But as traffic rising very speedily so more efficient architecture needed. Thus ICN has been proposed for Future Internet Architecture. This is a single model that replace Where with What, it means user care about What content is required not Where that data is located. ICN doesn't work on the basis of address (Source Address and Destination Address), But it works on the scheme of Naming and Named Content data. It has changed the whole scenario of the network, because ICN is a Data-Centric Architecture not a Host-Centric Architecture. IP handles same type of traffic but conversely NDN handles heterogeneous type of traffic. In Routing and Forwarding, request sent by the consumer in the Interest Packet, routers firstly check the Content Store(CS), if CS does not have content then request goes to Pending Interest Table (PIT), which is used to give direction to the Response Packet i.e. Data Packet towards the source, which have the data. In the Preceding section, we will study about the Architecture of NDN, Detail about Data Packet and Interest Packet, Working of Both Packets, Caching Scheme used for Content Replacement, Content Placement, Centrality based In-Network caching scheme, Comparison between ICN and IP.

II. ARCHITECTURE OF ICN

Nowadays Internet has eccentric and outstanding worldwide success story. In 1960s and 70s, Current Internet Architecture TCP/IP was evolved, through which point to point conversation done between two entities. But at that time, usage and accessing of the internet was less as compared to present scenario because lot of new data and applications continuously accommodate by Internet today. According to user perspective, Data Fetching with high efficiency, Secure content accessing, Quality of service (QoS) and many more. For designing a better network architecture, that should have:

1) High Security to Data.
2) Improve Quality of Service like High Speed transmission, Less Latency.
3) Having good flexibility and resilience.
4) Speedily retrieve the content not always from the actual host.

To the best of our knowledge, not even the single architecture satisfies all the goals concurrently. Overlay Network Architecture came and took initiative to improve the data dissemination like peer to peer (P2P), Content Delivery Network

ICN is a new Future Internet Architecture. Both ICN and IP have hour glass shape shown in Figure [1] but difference is at network layer, ICN uses content chunks rather than IP. ICN works on Pull model. Requestor or Consumer starts the communication through the interchange of two types of packets- Content/Data Packet and Interest Packet. These two types of packets have name by which data can be identified [1]. A requestor puts the name i.e. Named data object (NDO) of the content in the Interest packet and sends it to the network [2]. In the network, a router observes the Interest Packet and forwards it. When Interest Packet arrives at NDO containing router, then the Data packet that contains the name and signature of data returns to the requestor. The name can be in human readable form. The signature attached in data object shows the name data integrity and authenticity for the security measures.

1) Detail Regarding Interest Packet and Data Packet ::

Interest Packet   In Interest Packet the name can be send

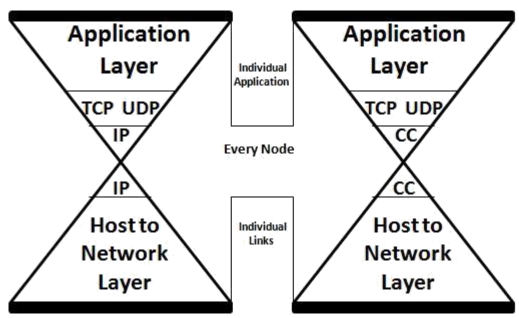


Fig. 1. Hour Glass Shape of IP and ICN Architecture

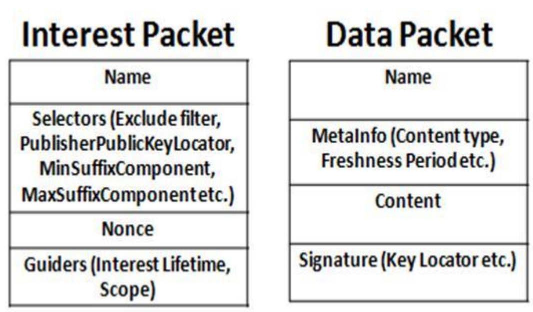


Fig. 2. Different Fields of Packet

whenever the consumer wants some data. Each Interest Packet has field like Name, Selectors, Nonce and Guiders/InterestLifetime. See Figure [2]

Name: ICN Name is a hierarchical name of variable length, which contains a sequence of name components. Selectors: It is an optional field which include Exclude filter (specify the component, when requested content come in Data Packet, the specified component exclude from the data packet), MinSuffixComponents(allows a data consumer to indicate the minimum number of component in the name), MaxSuffixComponents(allows a data

consumer to indicate the maximum number of component in the name)[3], PublisherPublicKeyLocator (contain the hash of the producers public key for the requested content)[4] etc.
Nonce: It carries a randomly-generated 4-octet long byte-string[3]. The Name and Nonce combination should uniquely identify an Interest packet. For discarding the same data Nonce are useful.
InterestLifetime: It is also optional field which includes InterestLifetime (time for the life of Interest Packet) and scope (limit the number of hops an Interest Packet can propagate)

Data Packet When source sends content to the requestor, it sends the data or required information in Data packet. Each Data Packet has field like Name, MetaInfo, Content and Signature[5]. See Figure [2]
Name: Same as the Name in the Interest Packet.
MetaInfo: MetaInfo contained the information like Content

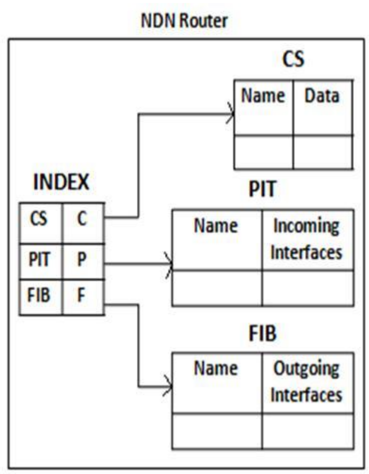


Fig. 3. Different Components in NDN Router

type, Freshness period etc.
Content: Content field includes the data. With each name there is some data associated.
Signature: Signature computed over the entire Content Packet and through the signature receiver assured that the data is correct. Security feature is directly related with this field. It includes key locator (the key needed to verify the Content Signature).

2) Detail Regarding Components of NDN: See Figure [3]
Content Store (CS): Used for fast data fetching and data caching.
Pending Interest Table (PIT): The information is stored regarding the incoming interfaces of Interest Packet that are not satisfied yet [6]. It stores the information so that Data Packet can be delivered back to same path.
Forwarding Information Base (FIB): FIB is a table which consists of name prefixes and corresponding outgoing interfaces. Similarly in case of IP-Based Network, FIB also exists but in that IP address is given instead of name and next hop is given instead of outgoing interfaces..

III. WORKING OF PACKETS

1) *Working of Interest Packet::* Request coming from the consumer is forwarded to the router enclosed in an Interest packet. NDN router primarily checks the Content Store (CS)
[6] for the data, if present it inserts the content in the Data packet and send Data Packet to Requestor. If not existed then forward the request to PIT .If the same data has already been requested by other consumer, then incoming interface is added in front of the existing named content. This request is not sent further as the same name content request has already exists in the network. Reduction in network

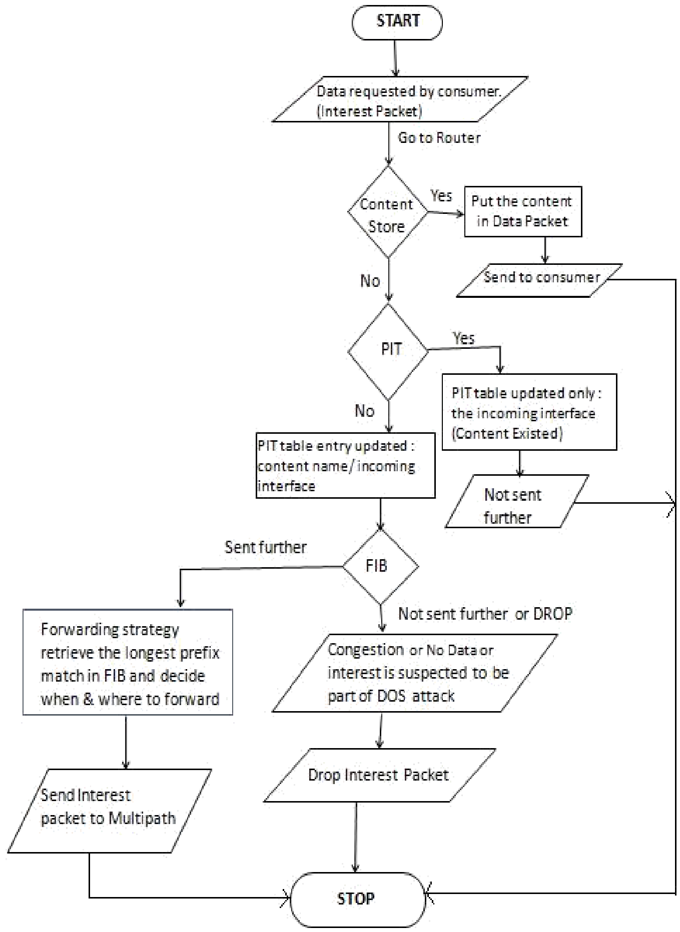


Fig. 4. Flowchart of Interest Packet

traffic leads to increased performance. If data is not found in the PIT table, new entry gets registered i.e. insertion of a new row (content name and incoming interface) in the table. It sends request to the Forwarding Information Base (FIB). FIB decides whether to send request further or not. If the network is congested at that time or Interest is suspected to be the part of Denial of service (DoS) [2] then the router will not send the Interest Packet further, i.e. it drops the Interest Packet. If the network is not congested at that time, then matches the name prefix with the longest name in the FIB table .On the basis of forwarding strategy, it is decided when and where to forward the request. The Interest Packet is sent to multiple paths so that router can quickly response to the request. Figure [4] show the flowchart of the Interest Packet.

Working of Interest Packet in the form of pseudocode are given below:

1) Start.
2) Data is requested by Consumer, the request go to Router.
3) If Content Store has data, Put the content in Data Packet and send to requestor and stop, otherwise go to step 4.

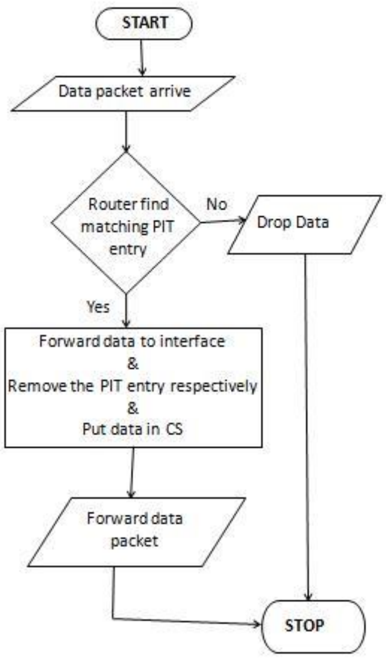


Fig. 5. Flowchart of Data Packet

4) Check the PIT, if entry is already there, then put the interface name in front of the requested content name and request not sent further.
5) If entry is not present in PIT, then content name and incoming interface are put in the PIT table. Request sent to FIB.
6) Request is not sent further, if there is congestion or data is a part of DOS attack otherwise using some forwarding strategy and match in FIB table and decide where to forward.
7) Send Interest Packet to Multipath.
8) Stop.

2) *Working of Data Packet::* The Data Packet coming from one of the interfaces, router firstly finds the matching PIT entry with respect to named content. If the named content is not available in PIT table then the content is stored in Content Store and drop the Data Packet otherwise forward data to the incoming interfaces and remove the PIT entry respectively. Also save the data in the Content Store for increasing performance (response time less), so that if further the same request is coming from another requestor then there is no need to find the data from the producer. Figure [5] show the flowchart of the Data Packet.

Working of Data Packet in the form of pseudocode are given below:

1) Start.
2) Data Packet arrives from the source.
3) Router find the PIT entry, if it is not existed then save the data in CS and drop the Data Packet and stop. Otherwise go to step 4.
4) Forward data to interface and simultaneously remove the PIT entry and put the data in the CS for further caching process.
5) Stop.

The Interest Packet is forwarded hop by hop and Data Packet is also coming via the same path. There is more redundancy in NDN because it copies the content at each intermediate interface and for every request generated there is always one response. Overall NDN not only avoids network congestion and conflicts, it deals with content authenticity and integrity, get rid of dependence of end to end connection, Multihoming, mobility, flow balance/load balancing, security which improve performance and efficiency of larger scale content distribution.

IV. CACHING

Caching is one of the advance features of NDN, through which efficiency of data dissemination increases. In the Conventional Architecture, caching concept was also used. In IP, Multiple URL was given when multiple replica of similar object resides at multiple servers. So if two URL for the similar objects was used, then it treated as a two different objects. But in ICN, hierarchical naming scheme i.e. Named data object (NDO) is given to each data rather than IP address, so this problem is being solved. When any Interest Packet is coming, the NDN router first analyses the Content Store [2]. If the data is present in the Content Store (acts as buffer memory), then data will be sent back to the consumer. So here In-Network Caching concept is used for faster retrieval of data. Some new features of NDN caching are Fine Granularity, Ubiquity and Transparency.

Fine Granularity: Divide the large file in to small granules (chunks), through this rather than saving the large file in the cache, save the chunks at the distinct nodes. Through this performance increases and space utilization also increased. But there is some open issue like neither analytical nor experimental study on the chunk level object popularity.

Ubiquity: In ICN, Cache data is present everywhere and the point where the cache is done is no longer fixed.

Transparency: ICN uses cache shared infrastructure scheme means different type of traffic class use same cache space. But in IP, cache space is used for only particular type of traffic.

V. CONTENT PLACEMENT/CONTENT MANAGEMENT

For faster retrieval of data, Where to store the content. Following are some strategy for content placement like Leave Copy Everywhere (LCE), Move Copy Down(MCD), Leave Copy Down (LCD), Randomly copy one(RCone), Probabilistic Cache.

Leave Copy Everywhere (LCE): From the name of LCE, everyone understand that copy the content at each router through which request going from requestor to source. With the help of this the content are speedily and easily accessible to other user. For example: Suppose there are 1 to m nodes and source is present or cache hit at n^{th} node (n<=m), then the content will place from 1 to n nodes. This as a result, content redundancy increases so lot of redundant data at each node but performance is increase.

Leave Copy Down (LCD): It is a Content Placement strategy which removes the problem of LCE. From the name of LCD, it shows that content is placed at the adjacent or immediate node of the source rather than the full path. For example: Suppose there are 1 to m nodes and source is present or cache hit at n^{th} node (n<=m), then the content will place only $(n-1)^{th}$ node. From this, redundancy decreases as compared to LCE.

Move Copy Down (MCD): It is a content management strategy in which removes the content from the source and save the data to the adjacent or immediate nodes i.e. straight downstream node and it reduces the content redundancy. For Example: Suppose there are 1 to m nodes and source is present or cache hit at nth node (n<=m), then in MCD data removes from the n^{th} node and save at $(n-1)^{th}$ node.

Randomly Copy One (RCone): In this strategy, copy the object at one node, randomly selected along the returning path when cache hit occurs. For Example: Suppose there are 1 to m nodes and source is present or cache hit at nth node (n<=m) then in RCone strategy content will place at any one of the node from 1 to n nodes.

Probabilistic Cache: In this strategy, the requested object copied at every node according to their probability. Probability (P) is inversely proportional to the distance (d) from this node to the requestor node. Probability decreases as distance increases and vice versa.

VI. CONTENT REPLACEMENT

When the cache space is over, then there is a need of content replacement. On the basis of some strategy, content will be replaced with new content. Some of the content replacement strategies are First-in First-out (FIFO), Least Recently Used (LRU), Least Frequently Used (LFU), Content Popularity and many more.

First-in First-out (FIFO): Content are replaced by FIFO Strategy i.e. data which are entered or in firstly, then replace or out firstly. But the problem is that FIFO doesn't consider the Content Popularity.

Least Recently Used (LRU): Replacement are done on the basis of access-time-pattern based policy, the data which are least recently used are replaced by the new content. But the problem is that LRU doesn't consider the Content Popularity and also the new content will replace the existing data, even if it is not in use.

Least Frequently Used (LFU): Replacement are again done on the basis of access-time-pattern based policy, the data which are least frequently used are replaced by the new content. It is not a good technique because some temporarily content (which are not used now) have more frequency than the new content, so that content are not replaced. With this new content doesn't have that much space for saving their content.

Recent studies show that LRU and Random Replacement Algorithm have similar Performance. There is some open areas in Content Replacement i.e. how content popularity for particular content to be find and on the basis of that which replacement algorithm takes place.

VII. CENTRALITY BASED IN-NETWORK CACHING SCHEME

Lot of issues are created because of ubiquitous caching scheme. The basic reason is that caching space is smaller as compared to overall population of the content to be cached. So on the basis of some Centrality scheme, Less Replacement will do and more Cache Hit Ratio will come.

Degree Centrality: Degree Centrality is based on the degree of the graph. Degree refers to the number of edges attached to the node. In order to find out the standardized score, divide each degree of particular node by (n-1), if n is number of nodes.

Closeness Centrality: Closeness Centrality is based on the closeness of the graph. Closeness refers to the length of the average shortest path between a vertex and all vertexes in the graph. In order to find out the standardized score, divide each closeness of particular node by (n-1), if n is number of nodes.

Betweenness Centrality: Betweenness Centrality is based in the Betweenness of the graph. Betweenness refers to the number of the shortest paths from all the vertices to all other that pass through that node.

Example of all the centrality in Figure [6] and the highlighted nodes are the nodes which have high centrality from all the nodes. For Degree Centrality, C and E nodes are selected. For Closeness Centrality, D node is selected. For Betweenness Centrality, C and D nodes are selected.

VIII. COMPARISON BETWEEN IP AND NDN

Both architecture using hour glass shape for the architecture, but in the NDN IP is replaced by Content Chunks[15]. IP address is of fixed length (32 bits in IPv4 and 128 bits in IPv6), but in NDN the Name prefixes used which are of variable length arranged in hierarchy delimited by /. Figure [7] ,differentiate these two Architecture.

IX. SUMMARY AND CONCLUSION

Recently Proposed ICN Architecture shows number of ad-vantages as compared to existing Conventional Architecture. Security is increased on the data not on the location. ICN remove the IP Address problem, it uses hierarchical naming scheme. For performance increasing issue, it uses the concept of In-Network Caching having different centrality algorithm and retrieval process is faster. Instead of that lot of open issues in ICN like Leave Copy Everywhere methods for caching, method used for hierarchical scheme, how much size of cache, which centrality scheme is best and many more. So in the future, there is a need to work on these issues and challenges and find a justifiable solution for these issues.

REFERENCES

[1] Yi, Cheng, et al. "Adaptive forwarding in named data networking." ACM SIGCOMM computer communication review 42.3 (2012): 62-67.

[2] Yi, Cheng, et al. "A case for stateful forwarding plane." Computer Communications 36.7 (2013): 779-791.

[3] "NDN Packet Format Specification, Interest Packet, Named Data Networking. Available: http://named-data.net/doc/ndn-tlv/interest.html. [Accessed: February 28, 2016].

[4] Gasti, Paolo, et al. "DoS and DDoS in named data networking."Computer Communications and Networks (ICCCN), 2013 22nd International Conference on. IEEE, 2013.

[5] "NDN Packet Format Specification, Data Packet, Networking. Available: http://named-data.net/doc/ndn-tlv/data.html. [Accessed: February 28, 2016].

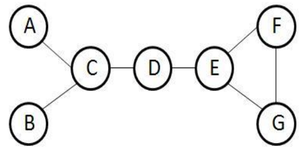
[6] L. Zhang, A. Afanasyev, J. Burke, V. Jacobson, K. Claffy, P. Crowley, C. Papadopoulos, L. Wang, and B. Zhang, " Named Data Networking," Available: http://named-data.net/publications/techreports/tr001ndn-proj/. [Accessed: February 28, 2016].

[7] Choi, Jaeyoung, et al. "A survey on content-oriented networking for efficient content delivery." Communications Magazine, IEEE 49.3 (2011): 121-127.

Centrality Example:

For node C, here n=7, calculate

1) Degree =3 , D_c=3/6 (Degree/n-1)
2) Closeness= (c-a)+(c-b)+(c-d)+(c-e)+(c-f)+(c-g)
   =1+1+1+2+3+3 =11
   C_c=6/11 (n-1/closeness)



3) Betweenness=δ(a-b)+ δ(a-d)+ δ(a-e)+ δ(a-f)+ δ(a-g)+ δ(b-d)+ δ(b-e)+ δ(b-f)+ δ(b-g)=9
   B_c=9/15 (Betweenness/{(n-1)(n-2)/2})

Name of Router	Degree Centrality		Closeness centrality		Betweenness centrality	
	Degree	D_c	Closeness	C_c	Betweenness	B_c
A	1	1/6	16	6/16	0	0/15
B	1	1/6	16	6/16	0	0/15
C	3	3/6	11	6/11	9	9/15
D	3	2/6	10	6/10	9	9/15
E	3	3/6	11	6/11	8	8/15
F	2	2/6	15	6/15	0	0/15
G	2	2/6	15	6/15	0	0/15

Fig. 6. Example of Degree, Closeness and Betweenness Centrality

NAMED DATA NETWORKING	CONVENTIONAL ARCHITECTURE(IP)
Future Internet Architecture.	Current Internet Architecture.
Data Centric Network Architecture.	Host Centric network Architecture.
"What".	"Where".
Content Store is present.	Content Store is not present.
Address Space and NAT Demolish.	Address Space and NAT used.
3 Entities Maintained i.e. PIT, FIB, CS.	1 Entity Maintained i.e. FIB.
FIB Stores multiple hop Status information.	FIB Store only next hop Status information.
Content Popularity scheme used for caching.	No Content Popularity scheme used.
New Routing Protocol NLSR based on name prefix.	Existing Routing Protocols based on IP Address.
Name prefixes are of variable length arranged in hierarchy delimited by '/'.	IP address is of fixed length.
Secure the data.	Secure the container.
In-network Caching and Cache shared infrastructure.	No Concept of In-network Caching and Cache shared infrastructure.

Fig. 7. Comparison Between IP And NDN