# An Investigation of Advanced Information Centric Model for Efficient Bandwidth Management

## Muchelule Yusuf

Lecturer, School of Business and Technology Umma University, Kenya

*Abstract:* **Information centric model is a new approach to networking which enables Efficient network software-based control and application independent information caching. However, in a transmission network, there exist redundant bytes in packets that are cached at content routers hence exhausting bandwidth and occasioning such problems as bandwidth glitches, low throughput, and denial of service among others. This study developed an investigation of an advanced redundancy elimination mechanism which takes into account minimal network memory operations leading to optimum bandwidth management and network power consumption reduction. The main aim of the research was to come up with an Investigation of Advanced information centric model to identify and eliminate redundancy so as to mitigate network glitches such as bandwidth glitches, low throughput, and denial of service among others. The research used purposive sampling to select sample network sites, collected data and evaluated our model using both real time network traffic data and synthetic traffic. The research analyzed bytes in packets and found out that there were 46% redundant bytes. The research compared various file types in transit and found that 97% of traffic was binary files such as video, audio and pictures while 3% constituted text files. This prompted the research to find out file organization of binary files which was found that they hardly partially intersect therefore was no need for chunking them. It was found that available bandwidth will be optimized by eliminating redundant bytes and reducing number of shim headers. Implementation of the Investigation of Advanced bandwidth management model will make internet applications be more reliable due to optimized bandwidth and reduced processing load.**

*Keywords:* **Information centric model, bandwidth management and synthetic traffic.**

## 1. INTRODUCTION

In the world today, computer networks are playing a central role in work, business, education, entertainment and social life. Organizations and individuals are benefiting greatly as computer networks have made communication faster and reliable through applications such as emails and video conferencing among others. Moreover, Tanenbaum and Wetherall (2011) argue that nowadays it is easier to control machines remotely, share peripheral devices, data and programs. Therefore it is important to efficient ly manage the way data is transmitted over the network in order not to overload the available bandwidth. Halgren (2012) agrees in his study that even though the capacity and speed of the network are constantly increasing and its associated costs are declining, it is still not a good reason for users to ignore the additional investments and efforts needed to optimize bandwidth management.

More studies reveal that investments in bandwidth optimization are the ones that can contribute to a reduction in total cost of ownership, specifically in respect to efficiency gains and maximized resources (Martin, 2010; Munir, 2014). This research has discussed more about bandwidth management which this chapter addresses and derives objectives that are used to achieve the solution to the problem.

**Background of the Study:**

The Internet interconnects millions of computers, providing a global communication, storage, and computation infrastructure. Moreover, the Internet is currently being integrated with mobile and wireless technology, ushering in an impressive array of new applications (Kurose & Ross, 2013). For instance, Bonaventure (2011) has pointed out that internet has allowed distributed applications such as remote login, electronic mail, Web surfing, instant messaging, audio, video streaming, Internet telephony, distributed games, peer-to-peer (P2P) file sharing, and much more running on its end systems to exchange data with each other.
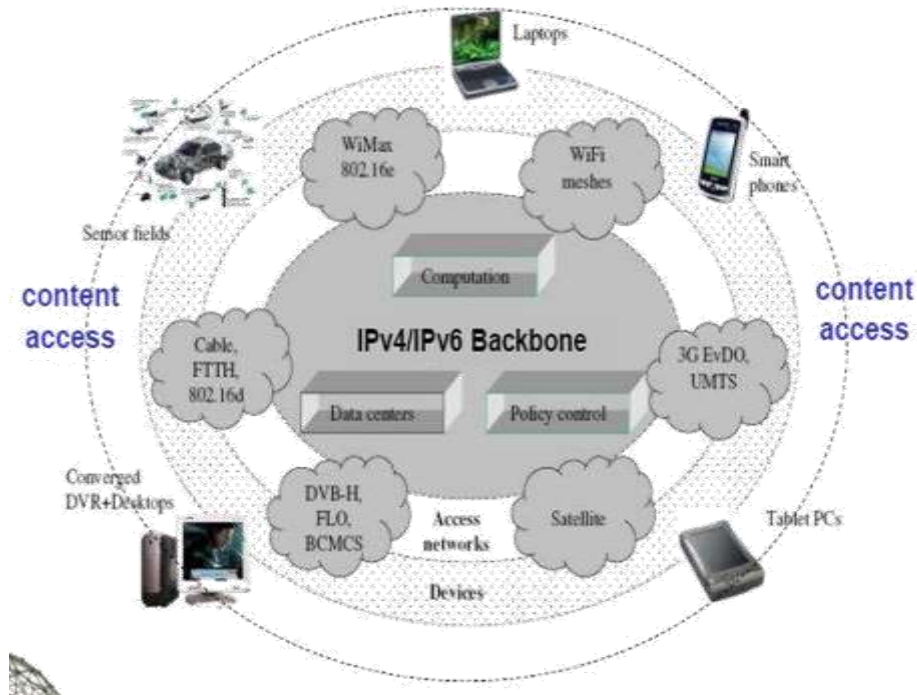


**Figure 1: Internet Architecture. (Source: Pavlou, 2011)**

This end system must communicate either through wired or wireless medium which has a laid down internet protocol, a good example is shown in figure 1. For instance, protocols in routers determine a packet's path from source to destination; hardware-implemented protocols in the network interface cards of two physically connected computers control the flow of bits on the wire between the two network interface cards; congestion-control protocols in end systems control the rate at which packets are transmitted between sender and receiver (Pavlou, 2011).
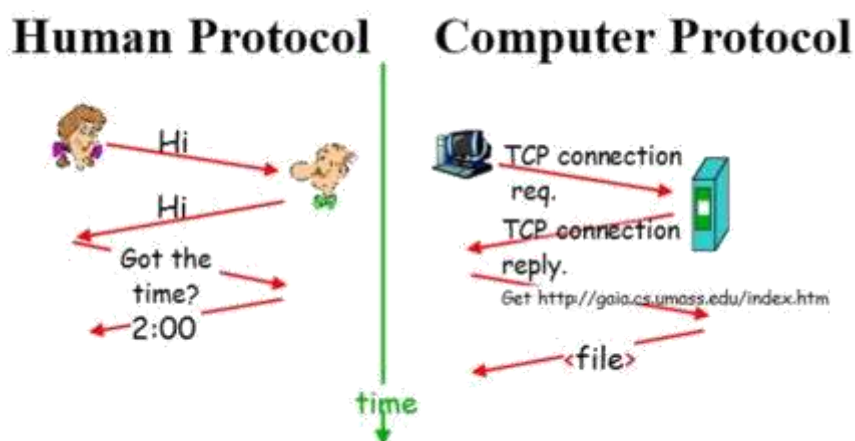


**Figure 2; Human and Network Protocol. (Source: Tanenbaum & Wetherall, 2011)**

Communication between sender and receiver can be easily understood using human analogies, since us humans communicate all of the time. Consider what you do when you want to ask someone for the time of day. A typical communication between two human being is shown in Figure 2. This analogy has been borrowed in computer protocols and has been used to set rules for network communication. In this transmission, long messages of data are broken into smaller messages called packets and routed to their destination (Tanenbaum & Wetherall, 2011).

Kurose and Ross (2010) have discussed more on network routing mechanisms in which they point out that each of packets in a transmission media needs to traverse from source to destination over routers and communication links. Packet routing approach has been used to ensure successful TCP client-server communication by involving a three way handshake between the sender and receiver before TCP connection is established (Schmidt, 2008) as shown in figure 2. Once a packet is received in its destination, its kept temporary in content router's cache before it's erased from memory (Kurose & Ross, 2010).

In particular, routing techniques have been evolving in order to adjust to the current networking bandwidth optimization needs (Bonaventure, 2011). Like in early packet routing approaches, Media Access addresses (MAC) of routers was used to uniquely identify packet source and destination (Kurose & Ross 2010; Stalling, 2007). However current routing approaches have shifted to information centric networking where information is viewed as first class entity of a network infrastructure (Parvlou, 2011). This has made it possible for network applications such as Bit torrent, YouTube, Google Video, Over the top Video, social networks and photo sharing sites to be more scalable, resilient and Efficient (Tarnauca, 2011).

Despite major shift from host to host to information centric networks there is still network congestions due to overfilling network links to capacity (Halgren, 2012). This is due to rise of the amount of data travelling on the network. For this reason, investments in bandwidth optimization techniques have concentrated on techniques that can contribute to a reduction in total cost of ownership, especially in respect to efficiency gains and maximized network resources (Munir, 2004).

In many business organizations all over the world, bandwidth is treated as a critical resource which need managed. In the recent past many researchers have experimented on Bandwidth management techniques such as throttling, middle boxes Redundancy Elimination have been suggested (Paulo, 2010; Munir, 2014) in a bid to manage bandwidth resource. For instance, Munir(2014) explains bandwidth throttling as an intentional limiting of network resource to users in an attempt to regulate network traffic and minimize network congestion. On the other hand Paulo (2010) point out that Middle box devices are used to analyze and reduce redundancy of incoming and outgoing content in networks. But these techniques can cause denial of service or sources of network failure (Paulo, 2010) hence the need for an advanced mechanism to mitigate the problems experienced by the current methods (Stoylar & Alexander, 2010). For example, in Kenyan universities, research and collaborative work are becoming regular processes prompting network administrators to manage bandwidth as a critical resource that is limited, expensive and of high demand (Martin, 2010).

**Statement of the Problem:**

In a telecommunication network, there is duplication of bytes that are common to different packets cached, routed and forwarded over information centric network leading to bandwidth wastage and reduction of quality of service. Current network setups use bandwidth throttling, middle boxes, packet aware routing Redundant Elimination and byte level redundancy elimination mechanisms for bandwidth resource management.

Bandwidth throttling, blocks some links in networks in order to conserve bandwidth, leading to denial of service, slow loading, skipping and stuttering experiences. On the other hand, Middlebox redundancy elimination devices are expensive, complex to manage, and creates new failure modes for the networks that use them. Constructing redundancy aware routes is challenging since it is not be economically viable to deploy redundancy elimination in every link. Also to preserve end to end performance and control signaling costs, routes cannot be determined on a per link basis, routes have to be determined independent of packet content. Byte level redundancy elimination mechanism currently used in most of information centric networks saves more bandwidth than packet aware routing RE but it has high processing demand due to high rate of processor cycles involved during chunking process. Therefore this research has developed an Investigation of Advanced information centric model that detects and eliminates redundant bytes in packets cached in

**ISSN 2394-7322**

**International Journal of Novel Research in Marketing Management and Economics**
Vol. 3, Issue 2, pp: (92-106), Month: May - August 2016, Available at: www.noveltyjournals.com

routers, while reducing cost of memory operations. This ensures that there is no overfilling of the network links to capacity reducing such problems as denial of service, deadlocks and bottlenecks.

**Problem Justification:**

It is desirable to introduce redundancy elimination mechanism in order to reduce network costs incurred through initial setup costs experienced when buying middle box devices. On the other hand, redundant bytes eliminated will reduce wastage of bundles purchased from ISP. In addition network throughput will be high as compared to the current network setups as more packets will be delivered at a given time frame making applications such as e-banking, e-commerce and video conferencing to be relied upon by end users. Furthermore, network and system administrators will benefit a lot by implementing this mechanism as one of the bandwidth management mechanism.

**Objectives of the Study:**

The general objective of the study was to research on existing bandwidth management techniques in order to develop an investigation of an advanced information centric model for efficient bandwidth management in telecommunication networks.

The specific objectives of the study were to:

i.  Analyze the network routing and caching optimization needs.

ii.  Investigate current network content redundancy elimination techniques used in bandwidth optimization.

iii. Develop an Investigation of Advanced information centric model that ensures Efficient bandwidth management and reduces network processing load based on traffic type.

iv. Evaluate the performance of the Investigation of Advanced information centric model by measuring percentage of bandwidth saving and processing load that it will provide as compared to existing model.

# 3.   RESEARCH METHODOLOGY

**Research Design:**

This study employed experimental research design to measure efficiency of information centric model designed. It was found to be appropriate in our study since it is concerned with deliberately changing of one or more process variables in order to observe the effect the changes have on one or more response variables. It was also found to be an Efficient procedure for planning experiments so that the data obtained can be analyzed to yield valid and objective conclusions (Nassiuma 2000). In our study packet samples were captured from different network setups their headers and contents analyzed to find percentage of repetitive contents.

**Sample Size and Sampling Design:**

An important aspect of designing an experiment is to know how many observations are needed to make conclusions of sufficient accuracy and with sufficient confidence. This research used small, medium and large enterprises network environment that had proxy server, content routers, switches and gateways. We classified the enterprises as small, medium or large based on the number of internal host IP addresses they accommodated (less than 50, 50-100, and 100-250, respectively) in the entire trace at each of these sites. While this classification was somewhat arbitrary, we used this division to study if the benefits depended on the size of an enterprise. Note that the total amount of traffic in each trace was approximately correlated to the number of host IP addresses, though there was a large amount of variation from day to day. Typical incoming traffic numbers for small enterprises varied from 0.3-10GB/day, for medium enterprises from 2-12GB/day and for large enterprises from 7-50GB/day. The access link capacities at these sites varied from a few Mbps to several tens of Mbps. The total size of traffic we captured (including inbound/ outbound traffic and headers) was approximately 3.5TB.

Packet samples were captured between 1.00 am and 6.30 pm since it was found that this is the time networks are utilized by users.

**ISSN 2394-7322**

**International Journal of Novel Research in Marketing Management and Economics**
Vol. 3, Issue 2, pp: (92-106), Month: May - August 2016, Available at: www.noveltyjournals.com

**Table 1: Network sites used for the study**

| Trace Site | Unique Client IPs | Dates(total days) | Size(TB) |
|---|---|---|---|
| Small enterprise | 27-37 | 10/11/2014- 17/11/2014(7 days) | 0.3 |
| Medium enterprise | 61-90 | 16/11/2014-22/11/2014(6 days) | 1.1 |
| Large enterprise | 101-210 | 05/12/2014-14/12/2014(9 days) | 2 |

Table1 represents content stores of sites which reflected the traffic typically observed at the access points. The stores consisted of YouTube video files, music files, images, and executable files; totaling to 3.4 TB.

**Experimental Setup:**

It is important to understand the set up environment where the research is carried out in order to make sure that the effects observed when manipulating the experimental setup are not just random effects due to chance. Our set up consists of the transmitter that is used to send data on the network. It also contains a receiver that accepts the transmitted information. The network monitoring module is connected to this network so as to capture and analyze the transmitted data, determine its source and destination as shown in figure 3. The experiment was set in a manner it could accommodate both wired (CAT5E Ethernet cable) and wireless devices (access point 802.11g enabled). All setups were connected to internet service provider via a router.

**Laptop reception:**



**Figure 3: Experimental setup**

**Data Collection Methods:**

The research relied on data from both real traces and secondary data from other researchers which could be vital in our research. Full packet traces were collected in LAN access links for six busy sites with high uploads and downloads. We used wireshark packet traffic analyzer module(Laura, 2014) to collect data that included network parameters such as maximum throughput, source and destination address, the type of protocol implemented, the time to live of packets, bandwidth and bit rates, Input Output graph, frame sizes , packet content and traffic data types. The collection consisted of snapshot recorded morning and evening hours from 10/11/2014 to December 14$^{th}$, 2014.

**Data analysis and presentation:**

It was found appropriate to inspect data streaming particularly http data by use of a parser program. A parser is a program that receives input in the form of sequential source program instructions, interactive online commands, markup tags, or some other defined interface and breaks them up into parts that can give useful information about the input data. This parser could be important as it could enable us to add a bypass for binary data in our chunking algorithm.

Using PHP we designed HTTP header parser which inspected the initial replies arriving from the server, recorded the file type field in the response message and stored its value in memory. The parser developed was platform independent since it could execute in both windows and linux environment. We used the *file_get_contents()* function to read file into a string while the *filesize()* function was used to return the size of the specified file. On the other hand the *get_mime_type ()* function returned the file type of streamed data.

**ISSN 2394-7322**

**International Journal of Novel Research in Marketing Management and Economics**
Vol. 3, Issue 2, pp: (92-106), Month: May - August 2016, Available at: www.noveltyjournals.com

The collected data was analyzed and presented using descriptive statistics such as frequencies, percentages, flow charts and graphs. The data traces were analyzed to find out the percentage of different file types they represent. The study found out that 97% of total volume constituted binary files while 3% represented text files. This prompted us further to present a table to compare percentage of redundant contents in these file types that is in text, sound, movies and images. Graph to compare bandwidth saving of different bandwidth optimization techniques with our Investigation of Advanced bandwidth optimization model was shown. Tabular and Graphical representation of costs of bandwidth before and after redundancy elimination across different ISP was presented. It was also vital to find out network power consumption rates of different bandwidth optimization mechanism. Graphical presentation comparing overall redundancy elimination to different packet chunk sizes was presented. Finally a flow chart representing our Investigation of Advanced bandwidth optimization model was presented.

## 3. RESULTS AND DISCUSSION

This chapter gives an explanation for the results that were obtained from the field study that was carried out. The probable explanation of the observed results and their analysis forms the basis of this chapter. The chapter is divided into the following key areas: Investigation of Advanced information centric model ,results and discussion from total volume of data captured; evaluation of percentage of bandwidth saving of different bandwidth management mechanism with our Investigation of Advanced bandwidth optimization algorithm, comparison of processing load of different redundancy elimination algorithm with our Investigation of Advanced Redundancy Elimination algorithm; comparison of throughput of current Redundancy elimination algorithms with our Investigation of Advanced redundancy elimination model; discussion of the benefits of the Investigation of Advanced redundancy elimination model.

**Investigation of Advanced Information Centric model:**

It is evident that the price to pay for detecting redundancy in low binary files is high prompting the research to utilize this observation to increase the performance of existing RE algorithms, such as the ones presented by Munir (2014) and Diego (2012).

Our algorithm utilizes the fact that the probability of binary files partial intersection is minimal therefore utilizes this to reduce network processing load. It is a two way selection circuit evaluation process. In figure 4, a variable that analyzes the content to determine file type is introduced before the chunking process is done.
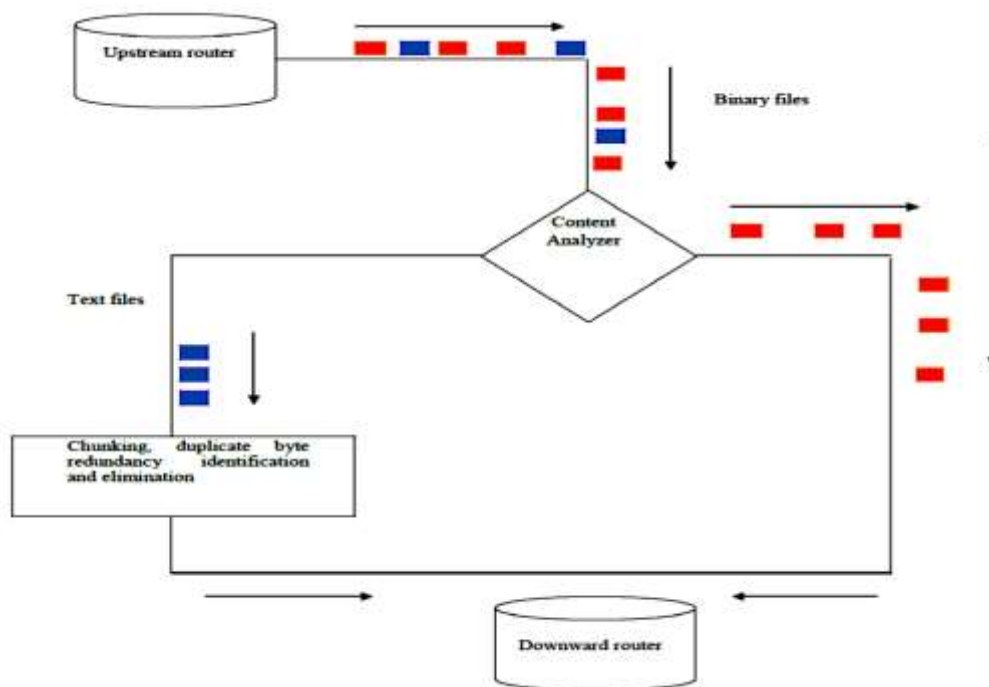


**Figure 4: UML presentation of the research Model**

**The research algorithm has two major stages:**

The First Stage, data inspection: it inspects each packet going through the RE engine for signatures that indicate the content type of the upcoming flow. This is applied to both HTTP and P2P or flash stream content. The HTTP OK response from the server is tracked and parsed for "mime-type" header field. This field typically indicates to the client software the type of the payload contained in the response flow. The mimetypes that we used for our algorithm are binary files ("audio", "video", images ISO files and executable files) and text files (word, html, pdf). When any of these binary mime-types are found in an HTTP OK response, that flow is marked as binary flow and is treated differently than other TCP flows. However, persistent HTTP connections pose a different challenge as they transfer multiple contents over the same TCP flow. To address this, the Content Detection stage also parses the HTTP OK packet for the "content-length" header field.

This field indicates the length of the content that will be transferred before the next GET request is processed.

The Second Stage, RE Bypass: This step of the algorithm utilizes the information from the first stage to reduce the processing load, maintenance, and cache lookup time at the network element. This step is basically a modified version of a typical RE algorithm, such as CombiHeader, which bypasses the packets from the flow marked as binary-flow. This simple choice of bypassing ensures that we do not spend valuable processing power for chunking traffic which will most probably not match with any later chunks. The completely matched binary files, packet level redundancy aware algorithm. With this simple modification in RE algorithms operation, not only huge processing power is saved, but also the practical throughput increases due to less delay on the network element. If content analyzer has identified non- binary files fingerprinting takes place using four stages.

Step one is to divide data stream into small chunks using combination of static, content-defined and file-based chunking approaches (Brinkmann, 2009). Step two is to calculate hash function for each chunk using SHA-3 fingerprinting algorithm to calculate a Hash Value of each of the substring of size.

Step three is to detect duplicate content by comparing hash results with already stored index. Step four is to update indexes and store data

The research model in figure 4 can also be represented by use of algorithm. Our algorithm introduced a variable (*Binary-flow*) to solve the problem as follows:

**Figure 4: Investigation of Advanced Redundancy Elimination algorithm**



```
While Steiner algorithm applies do

WHILE packet received do
If GET Request then
Set binary-flow←0
Set content-length←0
Else if OK Response then
If MIME-Type is in[Video ,audio, application] then
Set binary-flow←1 for the flow
Set content-length←Content length field
End if
Else
Find the flow of the packet
If content-length >0 and binary-flow=1 then
Pass packet without chunking or caching
Else
If content-length=0 and binary-flow=1 then
Binary-flow←0
End if
mask←0x0008A3110583080 {48 Byte window; 8KB chunks}
longval←0 {has to be 64 bits}
for all byte∈ stream do
```

**Content mix and order of files captured in the experiment:**

With the increasing and speed of internet speed, users are increasingly downloading and uploading content. This prompted us to study and analyze the sample size to determine percentage of file types streaming through the internet. It also prompted us to do a comparison between duplicated content and associated file type. We designed HTTP header parser as shown in table 2 which inspects the initial replies arriving from the server, the parser looks for the file type field in the response message and stores its value.

**Table 2 Showing sizes of different files captured from content store**

| File NO. | File Name | File Size(MB) |
|---|---|---|
| 1 | 00Illusion.mp3 | 102 |
| 2 | Brandy.avi | 67003 |
| 3 | Chiquitita.mp3 | 510 |
| 4 | Collection of text files | 101 |
| 5 | 5.jpg | 301 |
| 6 | collection of text files | 241 |
| 7 | Collection of text files | 6182 |
| 8 | 11.mp3 | 654 |
| 9 | 15.mp3 | 815 |
| 10 | AVSEQ03.DAT | 8189 |
| 11 | Collection of text files | 557 |
| | **TOTAL** | **84655** |

The research collected different file types from the content store and found out that 97% of total volume constituted binary files such as audio, video and pictures. 3% of file the volume constituted text files such as pdf, Microsoft office word, PowerPoint and html files. This was made possible by use of Wireshark packet analyzer and http parser tool.



**Figure 5: Sample of captured packets**

Figure 5, shows sample packets in transit, statistical aspects such as number of packet transmitted, time taken to transmit and comparison of packet contents.

**Evaluation of bandwidth management techniques**

The key to making most of network links efficient and reliable lies in grooming and reducing the traffic that travels across them, and in avoiding as many potential sources of delay as possible. Therefore it is important to evaluate the performance of each bandwidth techniques. Since each bandwidth management technique had its own way of testing bandwidth optimization we had to set their experiment separately and record their findings as shown in section 4.3.1, 4.3.2, 4.3.3 and 4.3.4. This research summarized the findings from each bandwidth management technique and produced the graph in figure 6.
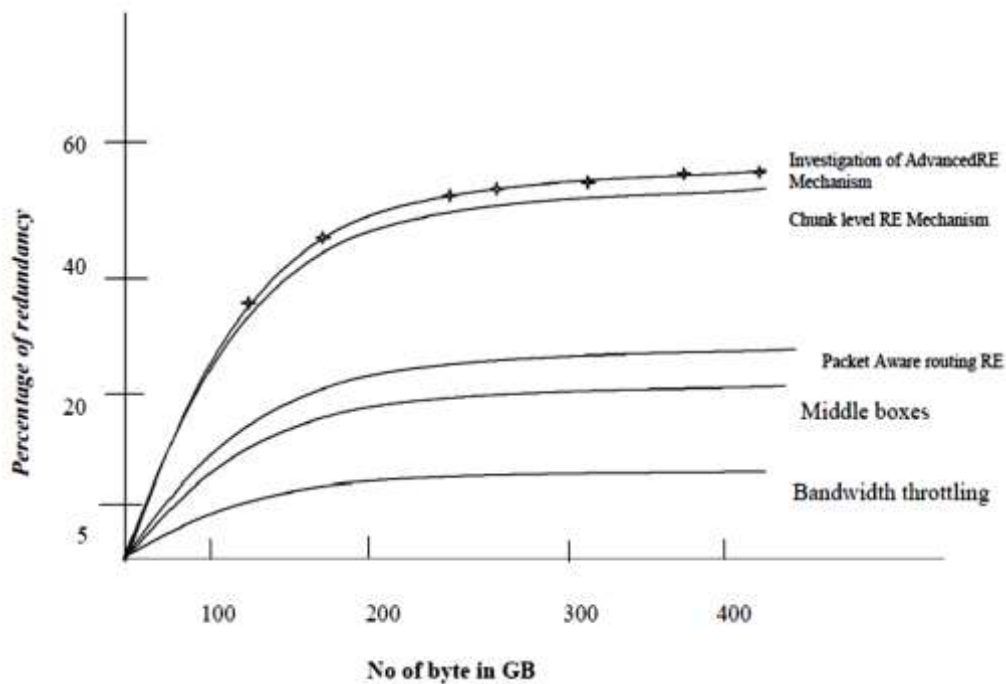


**Figure 6:** Graph comparing percentage of redundancy elimination produced by bandwidth management techniques due to the probability that part of the blocked portion contains redundant data. Middle boxes achieve 17% redundancy elimination since they are only placed at the end of network points. The result is also influenced by the fact that middle boxes track redundant traffic between devices connected to them and cannot track mobile or devices using wireless network. With packet routing redundancy aware routing we were only able to achieve 21% redundancy elimination better improvement as compared to bandwidth throttling. But with byte level without bypass we were able to achieve 43% redundancy elimination which was as a result of chunking individual packets. Byte level redundancy elimination with bypass mechanism however produced high level of redundancy elimination (46%) and was optimal in terms of processing load making our algorithm to be the best in bandwidth optimization as compared to the rest.

**Bandwidth throttling:**

This was the first experiment that we carried out to test the percentage of bandwidth saving that it could produce. We applied a proxy server to filter in and outbound messages. Our goal was to block some sites when bandwidth is exhausted specifically to heavy bandwidth usage sites such as video downloading links.. Among the test was to check if the quality of service has improved or not.

In doing this, we connected workstations to our Local Area Networks. After the connection was established, we tried to download an "iso" file from www.ubuntu.com/download and check the speed. The downloading started automatically and the speed was within the specified rate i.e. 512Kbps. files with ".iso" extension are members of bad_ extensions acl on the squid delay pool configuration.
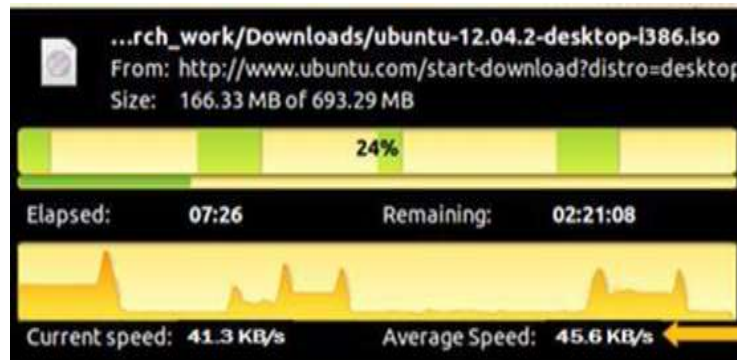
**Figure 7: downloading speed before applying bandwidth throttling.**
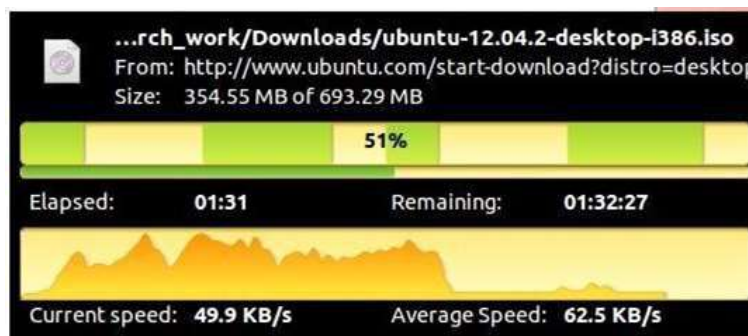


**Figure 8: Downloading speed after applying bandwidth throttling.**

Test carried out before the bandwidth throttling configuration shows that the downloading speed is low(24%) as compared to when bandwidth throttling approach is applied(51%).Moreover, despite the fact that few users were on the network at the time of testing, the speed was still poor. Comparing the network that has applied bandwidth throttling with one that has not shows high level of quality of service. Despite great improvement, a user in the network wanted to access a video tutorial but the site was temporary unavailable which inconvenienced him a lot. From the analyzed results the method achieved 5% redundancy elimination and this was due to the probability that part of the blocked portion contains redundant data.

**Packet routing redundancy aware elimination mechanism:**

We consider network design problems for information networks where routers can replicate data but cannot alter it. This functionality allows the network to eliminate data redundancy in traffic, thereby saving on routing costs.
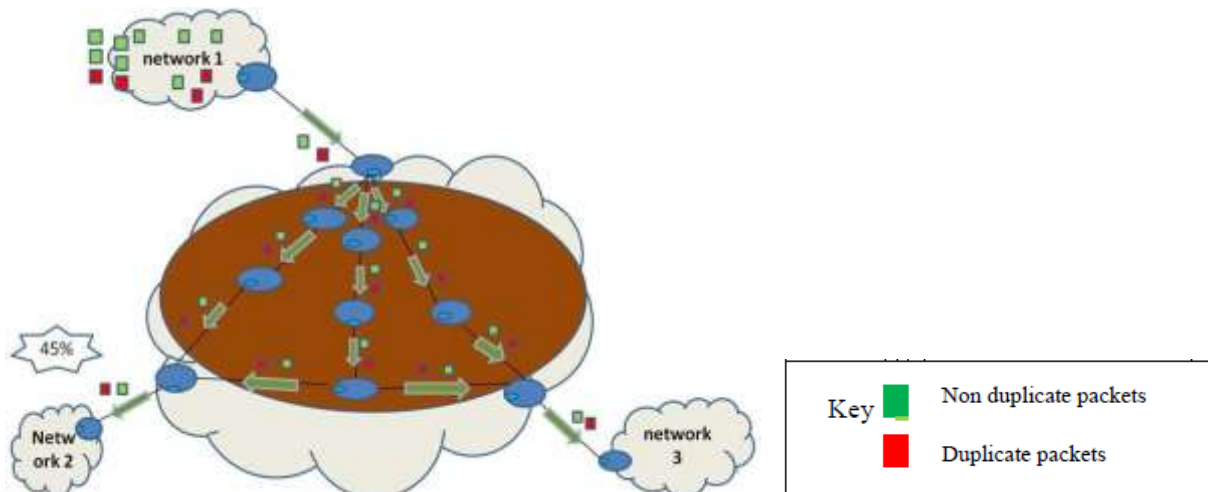


**Figure 9: Simulation of packet routing**

Assuming 12 packets are being sent from router in network 1 to router in network 3 and 2, and then applying redundancy aware routing we will have only transmitted 10 packets instead of 12 at a given route saving 20% bandwidth. Consider network design problems for information networks where edges can replicate data but cannot otherwise alter it. In this setting our goal is to exploit the redundancy in the given traffic matrix to save on routing costs. Formally, we are given a graph over a single server and many clients.

The server has a universe of data packets available, and each client desires a subset of the packets. The goal is to determine a collection of paths, one from the source to each client, such that the total cost of routing is minimized. Here the cost of routing on an edge is proportional to the total size of the distinct packets that the edge carries. For example, if the edge belongs to two paths that each carry the same packet, then the edge only needs to route the packet once and not twice.

However, constructing redundancy aware routes is challenging since it is not be economically viable to deploy redundancy elimination in every link. Also to preserve end to end performance and control signaling costs, routes cannot be determined on a per link basis, routes have to be determined independent of packet content.

**Byte level redundancy elimination:**

If chunking can be applied further to the captured packet, redundancy will be identified as much as possible across the entire content at the source. To accomplish this, we can re-use any of finger printing techniques (Alexander 2010). The chunking process proceeds as follows: for each file, of size L, a window of size w is moved from the beginning of the file to the end of the file. This window w represents the minimum size of the redundant string (or contiguous sequence of bytes) we would like to identify, and it is usually a number between 32 and 64. Each of the total $L - w + 1$ strings are used to compute a Rabin fingerprint generating $L-w+1$ fingerprints.

Using this method bandwidth savings can be represented as function of the minimum chunk size C for the network traces, considering cache size S ranging from 0.25 to 1.5Gbytes.
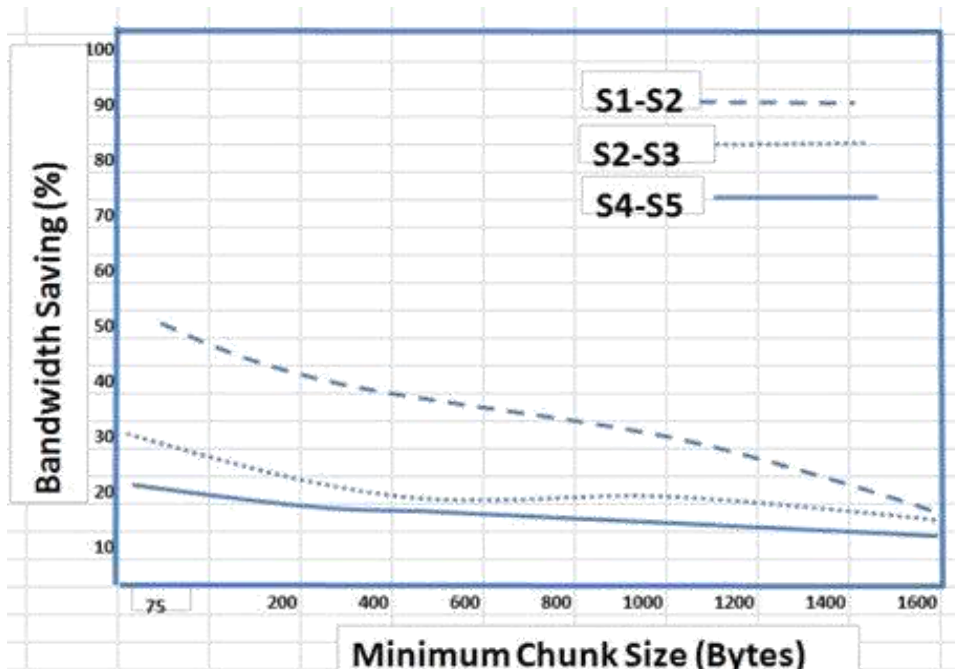


**Figure 10: Bandwidth saving versus minimum chunk size**

Figure 10 shows that for both traces the peak of the bandwidth savings results with the minimum chunk size (C=75Bytes).Increasing the value of C has two effects. First, the bandwidth savings are reduced, e.g., only 10% bandwidth saving when C=1425Bytes; this happens because the efficiency of the RE scheme reduces with large chunk sizes, as the probability to find redundant data decreases. Therefore the smaller the chunk size the higher redundancy is eliminated.

**ISSN 2394-7322**

**International Journal of Novel Research in Marketing Management and Economics**
Vol. 3, Issue 2, pp: (92-106), Month: May - August 2016, Available at: www.noveltyjournals.com

However when chunking was done and redundant profiles applied to each file, similarity of the files showed that the binary files were either completely similar to each other, or not similar at all (no visible data points between 0% and 100% similarity). This finding lead us to the observation that applying RE algorithm over binary files does not pay back well, and bypassing all such flows will allow us to achieve virtually same compression rate with much lighter processor load

Therefore if typical binary files such as music, videos and executable binary can be exempted from redundancy elimination operations memory operations will be reduced. These will be taken into account by introducing a variable to identify and exempt non-text files.

**Evaluation of processing load in redundancy Elimination mechanisms:**

Current redundant elimination mechanism removes unnecessary content from network by stripping off chunks of data from upstream routers which is already in downstream routers. Later reconstruction is done by downstream routers based on headers inserted by upstream routers. In the process every packet has to be chunked which requires massive processing power. Applying our Investigation of Advanced redundancy elimination mechanism means multimedia files (binary files) will be bypassed from chunking meaning network processing load will be reduced. Using results from table 4.2 which has 11 different files with a total of 84655MB shows that 97% of our sample size contains multimedia files while text files are represented by 3%. Applying minimum chunk size of 75byte derived from figure 4.7 to our content we will have 1,128,733,330 processor cycles using existing redundancy elimination mechanism making an assumption that all the packets were chunked. However, applying our investigation of advanced redundancy mechanism will result into 338619999 processor cycle representing a reduction of 91.74% of network memory accesses as shown in figure 11.
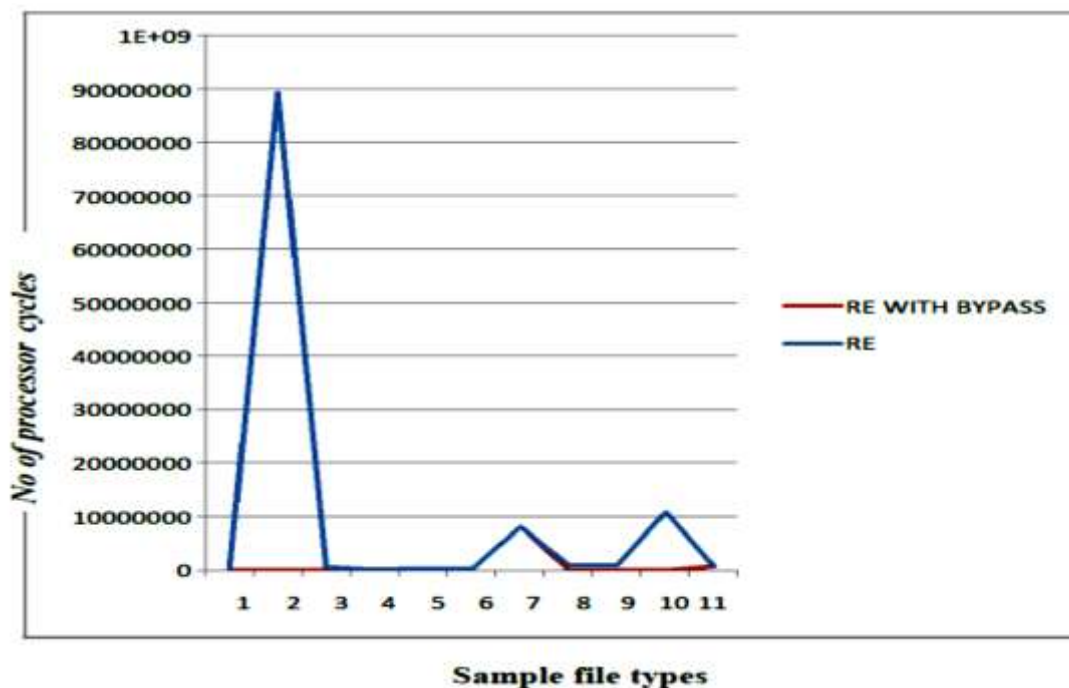


**Figure 11: Comparison of RE with our Investigation of AdvancedRE.**

This shows that our Investigation of Advanced bandwidth management mechanism is Efficient    in terms of memory usage (94413333) as compared to existing RE mechanism (1,128,733,330).

**Evaluation of Throughput of bandwidth management techniques:**

Network throughput refers to the amount of useful data passing through a network. In a network, both useful data and signaling information (that ensures proper reception of the transmitted information) are being transmitted. Of importance to the network administrator is the amount of useful data that his network can accommodate. Networks with greater throughputs are said to be Efficient    networks. This is because they indicate the proper utilization of the network bandwidth. Therefore if redundant content can be minimized it means there will be high network throughputs.
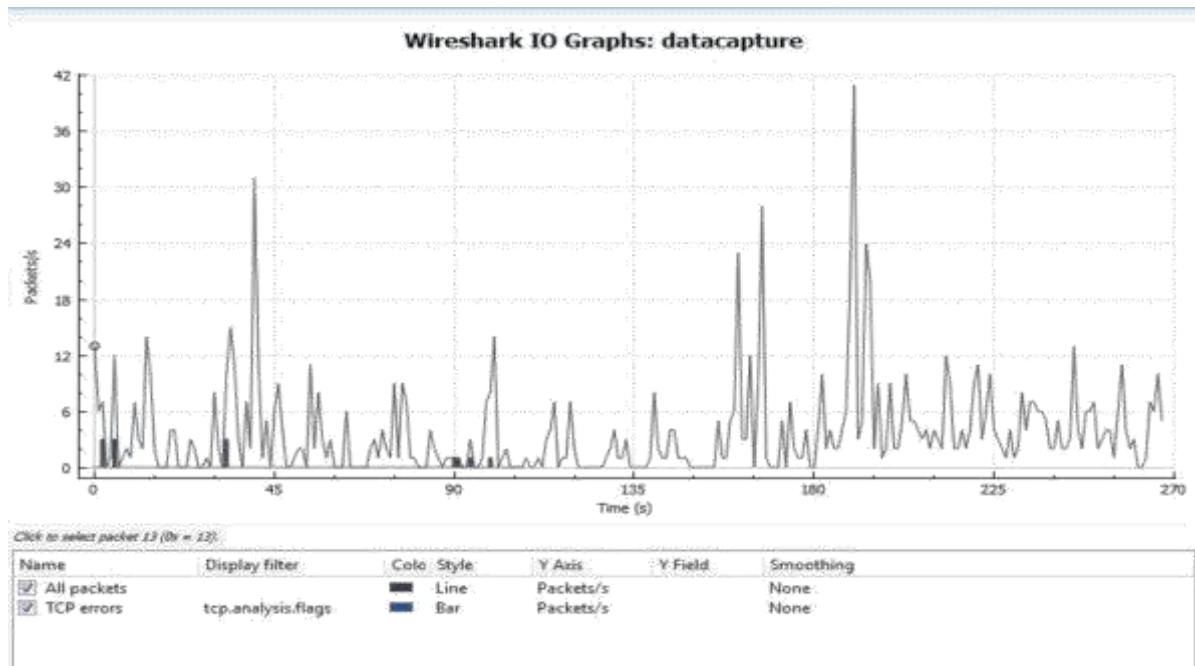
**ISSN 2394-7322**

**International Journal of Novel Research in Marketing Management and Economics**
Vol. 3, Issue 2, pp: (92-106), Month: May - August 2016, Available at: **www.noveltyjournals.com**

**Figure 12: Input output data capture graph**

The results captured in figure 12 shows that network throughput varies over time, with the throughput being higher at one time and very low at other times. For example, 30 S - 45s, 150S-180S there is heavy throughput, while between 90S and 92S, the network throughput is zero. Therefore there is need for coming up with a mechanism that will utilize bandwidth well to ensure maximum throughput.

It is also good to consider the packet window size that will produce best result when deciding network throughput by applying throughput formula, where in the receive window size and in the Round-Trip Time, it is possible to get the best minimum chunk size that can yield high throughput or maximum transmitted data that can pass through a channel without causing deadlocks.

**Benefits of the Investigation of Advanced Information Centric Model:**

Internet users who purchase bundles from ISPs will benefit by saving costs due to Overall bandwidth saving provided by minimal redundancy mechanism. Initial network costs required to setup an organization network will also be reduced as this mechanism will be implemented as a protocol eliminating purchasing of middle-box devices.

Minimal data redundancy will lead to elimination of traffic jams and denial of service which will make internet applications to be downloading and uploading content at a faster rate. Applications such as Bit torrent, YouTube, Google Video, over the top Video, social networks and photo sharing sites will benefit from this protocol as they require synchronous network setup. These applications will be uploading and downloading more contents at a given time as compared to current information centric networks as traffic jams will be reduced.

Quality of service (QoS) will be high as there will be reduced error rates, better bandwidth saving, increased throughput and minimal transmission delay. These will be important for the transport of traffic with special requirements such as video streaming applications, telephone networks for audio conversations, as well as supporting new applications with even stricter service demands.

## 4. CONCLUSION

In summary, previous researchers have provided valuable and detailed insights into optimization of bandwidth in Information Centric Networks. However, this research recommended deployment of byte level redundant content elimination with bypass to exclude multimedia content from chunking reducing network processing load. This was due the results that showed that our investigation of advanced redundant elimination mechanism would eliminate 46% of

redundant traffic a great improvement from the existing mechanisms such as bandwidth throttling (5%) middle boxes (17%), packet routing redundancy aware routing (21%) and byte level RE without bypass (43%). Our mechanism also proofed to be efficient   in terms of memory usage as it reduced 97% of memory operation from existing RE method. Finally this research has presented the promising features of Investigation of Advanced information centric networks design for the new Internet architecture which are naming, name resolution and data routing, caching, mobility and security. While this work shows benefits of our Investigation of advanced information centric model, it is our future endeavor to explore more on the challenges of implementing it as IP layer protocol.

## REFERENCES

[1] Alexander A.(2010).*Winnowing: Local Algorithms for Document Fingerprinting.*

[2] SIGMOD international conference on Management of data, Chicago: University of Illinois.

[3] Anand A., Sekar, V., & Akella, A.(2009). SmartRE: an architecture for coordinated Network wide redundancy elimination. *ACM SIGCOMM Computer Communication Review. SIGCOMM, 39* (4).

[4] Bengt, A. (2010). Network of Information: information Centric Networking Architecture. German: *Proceedings of the 2008 ACM CoNEXT Conference.*

[5] Bichanga, O., & Wario, Y. (2014). Effects of E-Banking on Growth of Customer Base in Kenyan Banks. *International Journal of Research in Management & Business Studies*. Nairobi: Kenyatta University.

[6] Bonaventure, O.(2011). *Computer Networking Principles Protocols and Practice. Louvain: Université catholique de Louvain.*

[7] Brinkmann, A.(2008). Theoretical Aspects of storage Systems. 12th International Conference, *OPODIS 2008, Luxor, Egypt, December 15-18, 2008. Proceedings.*

[8] Choi, J., Han, J., Cho E., Kwon T., & Choi Y. (2010). Survey on Content-Oriented Networking for Efficient Content Delivery. *IEEE Comm, 49* (3).

[9] Diego, P.(2012). Redundancy Elimination for Information Centric Networks. *ICN '12 Proceedings of the second edition of the ICN.*

[10] Dina, B., & David, J. (2003). *Duplicate Record Elimination in Large Data Files*. Madison: University of Wisconsin-Madison.

[11] Dolvara, G.(2013) Recent information Centric Networking Approaches. Retrieved from *http://www.cnsm-conf.org/ 2013/documents/CNSM2013-Keynot2-G-Pavlou.pdf.*

[12] Halgren, B.(2012). Content, Connectivity and Cloud; Ingredients for the future Network. *in Proc. of CoNEXT'09 - 5th International Conference on Emerging Networking Experiments and Technologies, Rome, Italy.*

[13] Justine, S., Shaddi, H., & Arvind, K.(2013).*Making Middleboxes Someone Else's Problem*: Network Processing as a Cloud Service. Washington. University of Washington.

[14] Koponen T., Chawla M.( 2007).*A Data-Oriented (and Beyond) Network Architecture*:Kyoto, Japan: John wileys and sons.

[15] Kurose, F., & Ross, W.(2013). *Computer Networks:* A Top Down Approach Featuring the Internet. 6th edition. Brooklyn: Pearson Education printing press

[16] Laura C.(2014). *Wireshark open source software*.

[17] Retrieved from https://www.wireshark.org.

[18] Martin, B.(2010). Bandwidth Management and Optimization: Policy development workshop, Nairobi: KENET.

[19] Mironov, I.(2005). *Hash functions: Theory, attacks and applications.* Microsoft Research, *TechReport, Publication no. MSR-TR-2005-187.Chicago,US.*

**ISSN 2394-7322**

**International Journal of Novel Research in Marketing Management and Economics**
Vol. 3, Issue 2, pp: (92-106), Month: May - August 2016, Available at: www.noveltyjournals.com

[20] Mochizuki, .K., Shimazaki, S., Hanawa, D., & Oguchi, K.( 2012.) "Proposal of new traffic control method in the next generation home network," in Proc.

[21] Telecommunications and Signal Processing, 35th International Conference on Sci. & Technol., Japan: Seikei Univ., Musashino,.

[22] Nassiuma, K.(2000). *Survey sampling: theory and methods*; Nairobi University press Neilson, T.( 2006). *Photonics for switching and routing*. IEEE Journal of Selected Topics in Quantum Electronics. 19(2).

[23] Oso, W.Y, & Onien D.(2005): *A general guide to writing research proposal and report:* Kampala: Makerere University.

[24] Paulo, M., & Rute, S.(2010). *The role of Information Centric Networking paradigm on the future Internet architectures.* Tokyo: API Interest Group

[25] Pavlou, G. (2011). *Information Centric Networking overview, current state and key challenges:* London, Pearson printing press.

[26] Rabin, M.(1999). Fingerprinting by Random Polynomials. Technical Report; CRCTTR-15-81, Harvard: Harvard University.

[27] Schmidt, T. (2008). Routing and Packet Forwarding. Finding the Shortest Path. Retrieved from http://www14.in.tum.de/konferenzen/Ferienakademie08/talks/tina_schmidt/prese ntation_schmidt_tina_routing_and_packet_forwarding.pdf.

[28] Shirley, R.(2014). *The Cryptographic Hash Algorithm Family*: Revision of The Secure Hash Standard And Ongoing Competition For New Hash Algorithms.India: National Institute of Standards and Technology.

[29] Stalling, W.(2007). *Data and Computer Communications*. 8th ed. New York: Pearson/Prentice Hall.

[30] Stolyar, T., & Alexander R. (2010). Shadow-routing based control of flexible multi-server pools in overload. Operations Research, vol3, issue no. 030.

[31] Tanenbaum, S., & Wetherall, D. (2003). *Computer Networks. 5th edition.* Boston Columbus. Prentice Hall:

[32] Tarnauca, B., & Nechifor, S. (2010). *Netinf evaluation*, EC FP7-ICT-4WARD Project, deliverable D-6.3.

[33] Vaidya, R.(2011). Emerging Trends on Functional Utilization of Mobile Banking in Developed Markets in Next 3-4 Years. *International Review of Business Research Papers 7*(1).

[34] Yang, L., & Park Y.(2009) Algorithms and Architectures for parallel processing. 14th International Conference, ICA3PP 2014, Dalian, China.

[35] Zafar, Q., Cheng-Chun, T., & Rui M. (2013). Practical and Incremental Convergence between SDN and Middleboxes. Retrieved from http://nsl.cs.usc.edu/~rmiao/publications/Zafar13_.pdf.