International Journal of Computer Engineering and Information Technology

VOL. 8, NO. 10, October 2016, 189–195
Available online at: www.ijceit.org
E-ISSN 2412-8856 (Online)

Enhanced Hybrid Asymmetric-Multicast Hash-Routing for Information Centric Networks

Munaf Hamza Kareem[1], Hazem M. El-Bakry[2] and Mervat Abu-Elkheir[3]

[1, 2, 3] Information Systems Department, Faculty of Computer & Information Sciences, Mansoura University, Mansoura, Egypt

ABSTRACT

Information-Centric Networking methods (ICN) present fundamental resilience to a lot of users in order to collect data. It is one of the substantial sharing features of ICN designs are the global caching. They are exceedingly admissible that the in-network cache will enhance the execution. Besides, there was no totally abridgement on the mobility to plan an effective caching program in ICN network. In our research we discuss the modality of enhancing the Hybrid Asymmetric-Multicast Hash-Routing strategy by finding the best location for storage regarding the nodes which are participating directly in the search for content achieves higher cache hit ratio.

Keywords: *Information Centric Networks; Caching; Hash-Routing.*

1. INTRODUCTION

Information Centric Networking (ICN) is a relatively emerging field of research that stems from the development of using networks telecommunication and IP-based networks to publish large volumes of contents that are addressed to more than one recipient. This involves contents that were created for public (for example, by big entities such as news agency, movie studio, and meteorological office and etc)[1], as well as content available to the publication only among limited groups of recipients. ICN concentrates on getting and sending data to end users instead of connecting end hosts which are exchanging information. The IP paradigm is one-to-one communications despite permanent enhancements over the years.

There are four systems that are actively developed to support ICN:

- Data-Oriented Network Architecture (DONA)[2].
  The Data Oriented Network structure from UC Berkeley's is one of the early completed ICNs structure, as it primarily alters naming through substituting the structural URLs with flat names. Unlike URLs which are restricted to specified places via their DNS components, the flat names in Data Oriented Network structure can be permanent, even if the data changes. This permits data to be stashed and reproduced at the network layers, accordingly raising data accessibility. At the end, names in Data Oriented Network structure permits users to assure of the validity that the collected information is similar to a needed name through cryptographic methods. From the other side, Data Oriented Network structure holds IP address and route, either internationally or nationally, using a name resolution method as an covering that describes its flat names to the identical data.
- Content-Centric Network, presently in the Named Data Networking program [3].
  The Named Data Networking (NDN) Conventional communication between a pair of networked machines has developed into a screenplay where new services create unmatched quantities of content (e.g. video traffic and cloud access), and simultaneously permit multilink mobile devices to reach that data via various connection chances. The current host-centric structured has been cobbled to evolve content-oriented methods such as Content Delivery Networks, peer-to-peer overlays and HTTP proxies, utilized over the current structure. Enhancing these enhancements to reach content, network operators use a series of management regulations governing a complicated diversified environment, via the accumulative employment of methods such as Quality of Services (QoS), network policies, balancing of load, failback actions, and over the-air up to date, among others. The focus on content distribution has led up to the term of Information-Centric Network permitting content to be handled by name and not by place or end-point address. However, there has been very little concern of management sides of ICN structures, which are primary for enhancing network operations.

- Publish-Subscribe Internet Routing Paradigm (PSIRP), nowadays in the Publish-Subscribe Internet Technology (PURSUIT) projects [4].
  The Publish Subscribe Internet Routing Paradigm project and its permanence the Publish Subscribe Internet Technology project, Whether financed by the EU Framework 7 agenda, produced a structure that totally substitutes the IP protocol bundle with a publish-subscribe protocols The Publish Subscribe Internet Technology project architecture composed of three independent processes: rendezvous, topology administration and forwarding. When the rendezvous process corresponds to a subscription to a publication, it guides the topology managerial function to originate a path between the publishers and the subscribers. This path is at the end utilized by the forwarding process to enforcement of the real transmit of information.
- Information network from the Future Internet's design (4WARD) agenda [5], presently in the Internet Solutions (SAIL) project that can be scaled and developed.
  An information-centric structure has fundamental features in large-scale data spreading ways like the allocation of Web pages and multimedia content. Moreover, it offers boost for sporadic connectivity, information access in (partially) discontinuous scenarios, and mobility scenarios. The NetInf architecture also conciliates non-dissemination applications, including streaming and interpersonal communications. Beyond those use cases, NetInf also boosts the combination of the real world and relevant data and services in the Internet, which we call the Augmented Internet scenario. NetInf expands the term of the identifier/locator split with another level of indirection to separate objects from their store place(s). This information-centric security approach qualifies a better level of safety and confidence than actually found in the Internet. Besides, it is substantial to indicate that Network of Information is autonomous from the implicit transformation protocol and is, moreover, in a unified position to get benefit of other technologies ranging from new transportation methods to virtualization to in-network administration.
  The structure has a number of core components to facilitate content-based delivery. These components are:
- Naming: Named Data Objects (NDO) important for (ICN), because all kinds of objects we have stored is reached her by the name, even while transferring these objects to another place. The structure of the name assigned piece of information (or service) that can be transmitted across the network is one of the primary features of each suggested ICN structure. In all ICN infrastructures, information names are place-independent. On the other hand, based on the approach, the names may extend from flat to hierarchic land can be or not be human legible [6].
- Routing and Forwarding — It is responsible for taking the necessary path for routing user request to the content provider and then directing the content to the user. There are two generic ways in ICNs to address routing, both strongly based on the features of the object namespace, particularly, if the names are aggregated or not. We, besides, differentiate between two routing phases:
  - Routing of NDO requests
  - Routing of NDO back to the requester
  - The first method uses a name resolution service (NRS) that hoard bindings from object names to topology-based locators pointing to similar store places in the network. This approach has three fanciful routing methods:
  - Conveying the request message to the responsible NRS node where the object name is interpreted into one or multiple source addresses
  - Conveying the requested message to the source address(es)
  - Conveying the data from the source to the requester (s)

  All stages can probably utilize various routing algorithms. A *name-based routing* method might be utilized, in especial for the first stage. The second and third stages may utilize topology-based routing such as Internet Protocol (IP) [7].
- Caching: The ICN model gives the notion to handle and reach content in a direct way instead of accessing the locations where the contents are situated. This notion also makes accessing cached copies of content more practical than before. It can also increase energy effectiveness by minimizing traffic throughout the network. However, one of the remaining questions is where to put caches and how to administer their entries. We differentiate between on-track and off-track caching. In on-path caching the network utilizes information cached along the track needs by name resolution requests, while in off-track caching the network utilizes data cached outside that track. There is a broad domain of potential choices. It is in especial paramount to define whether only caches on a direct path from a client to a server should be utilized (on-track caching) or if it is likely to transfer from the direct track to select from other possible caches (off-track caching) [8].
- Mobility: Mobility of subscribers is fundamentally enhanced in ICN structures, since mobile users can transmit all new subscriptions for information after the handoff. Publisher mobility is more troublesome to boost, because the system name resolutions (in the associated approach) or routing tables(in a separate approach) need to be updated [9].

- Security: This feature is relevant to the naming structure. From the other side, human- being readable names demand trusted agents or trust relationships with the name resolution system to assure of the returned information similar to the requested name. On the other side, flat names can enhance self-accreditation, but are not-human being readable, thus demanding another trusted system to map human being-readable names to flat names [10].

Of the aforementioned components, caching is researched extensively, as it facilitates content delivery while maintaining a balance between storage capacity and delivery delay. Caching when there is a request from the medium node and there were copies exist in this node will be returned in response locally. This leads to a number of problems, including the caching effective, and how it will be deportation optimized for caching, and the perfect place to store the content. To help raising the caching we have proposed to enhance the Hybrid Asymmetric-Multicast Hash-Routing strategy. This strategy is originally a hybrid strategy within the strategies of Hash- routing, which depends on using the strategies of Asymmetry and Multicast , the essential job of these strategies is Asymmetric, when there is an order, it will move from one node to another to search for the user's order , when it find this order, it then turned it back through choosing the shortest route , on the other hand, the Multicast strategy also looks for the user's order, as in the Asymmetric system, but the difference lies here after choosing the shortest route and returning back, it then throw a copy of the order in the path, in which it was actually in during the search , throw the copy in the first place, the hybrid strategy select the best of these strategies. We choose this, strategy because it allow us to select the best place to throw in, as it can calculate the number of nodes, that have been navigating, therefore we can choose the best place in it, which is suitable for the rest of the servers, and it also can be accessed with minimal movement number between nodes. The amendment we conducted on this strategy is to selection for the middle node in this path to be close to all servers, which represent the best site for storage.

2. RELATED WORK

One of the most paramount common matters of ICN designs is the global caching. Caching of data in computer networks has attracted the concern since a long period. Beginning with web caching and continuous with caching in peer-to-peer networks. The ICN paradigm offers the notion to handle and reach content directly instead of reaching the place where the contents are located. This notion also makes reaching cached copies of content more practical than before. It may also increase energy effectiveness by decreasing traffics throughout the network. However, the question remains where to place caches and how to administer their entries..

In order to enhance the caching gain and preclude the uncertainties in the performance of the much simple random caching strategy, W. Chai et al. [11].Suggested a centrality-focused caching algorithm by utilizing the term of (ego network) between centrality and a parataxis of it (EgoBetw) for scalable and distributed realization in dynamic network environments where the full topology cannot be known in a prior way. To reduce the network postponement and publisher capacity for ICN network and overcome its negatives, Z. Ming et al. [12].Suggested an age-based cooperative cache scheme as the distance from the server may take a comparatively long time to recover their contents. I. Psaras et al. [13].Suggested an algorithm based on path lengths named ProbCache to approximate the capacity of paths to cache contents, multiplexes content flows accordingly, utilize resources effectively, and minimize caching redundancies and in turn, network traffic redundancies. S. Borst et al. [14].Indicated that Caching strategies offer an efficient way for alleviating these huge bandwidth needs by proliferating the most popular contents closer to the network edge, rather than saving it in a primary site, and reduce the traffic load lessens the required transportation capability and capital expenditures, and alleviates execution dilemmas. S. Eum et al. [15].Proposed an ICN structure which is known as Cache Aware Target identification (CATT) and inserting the structure by focusing on two of its fundamental components: routing and content caching. Their operations were explained in the context of ICN to approve how the Cache Aware Target identification architecture could attain its various design objectives, namely availableness, adaptation, diversification, and robustness. In addition, the PBR provides a method to choose a content file based not only on vicinity but also on the quality of the content, which makes the selection process rather diversified . Y. He et al. [16].Defined Content Gradually Tend to Important Node (CGTIN) as a cache decision strategy which does it best to push more widespread contents to more paramount nodes and extend the survival time of these more popular contents and conclude that LCE obviously is not a great option due to is inability to differentiate efficiently the significance of nodes on content delivery paths.

We can summarize the ICN caching techniques and algorithms as follows:

Leave Copy Everywhere (LCE): In this methodology a duplicate of a substance is reproduced at any reserve on the way between serving hub and recipient. No caching (NO_CACHE), briefest way directing: This relates to the

movement in a typical TCP/IP system with no CDNs or overlay storing, where every single substance solicitation are served by the first source.

Symmetric hash-routing (HR_SYMM): In this technique, the forwarding and the arrival ways along which content requests and delivery take place are symmetric [17].

Asymmetric hash-routing(HR_ASYMM): In this technique, the forwarding and the arrival ways are away.

Multicast hash- routing (HR_MULTICAST): In this technique, if there is a reserve miss, when content parcels return in the area, the bundle is multicast, one duplicate being sent to the definitive store and the other to the beneficiary. On the off chance that the store is on the way from source to beneficiary, this technique acts as a typical symmetric hash-steering methodology.

Hybrid Asymm-Multicast hash- routing (HR_HYBRID_AM): In this technique, if there is a reserve miss, when content parcels return in the space, the bundle is conveyed to the beneficiary after the most limited way. In the event that the extra number of bounces required to send a duplicate to the legitimate reserve is beneath a particular division of the system breadth, then one duplicate is sent to the definitive store too. In the event that the store is on the way from source to beneficiary, this system carries on as an ordinary symmetric hash-directing procedure.

Hybrid Symm-Multicast hash- routing (HR_HYBRID_SM): In this application, the edge switch accepting a substance parcel chooses whether to pass on the bundle utilizing multicast or symmetric hash-steering taking into account the aggregate expense for conveying the Data to both store and recipient as far as bounces.

Cache less for more (CL4M): In this technique, on the way storing and examine if reserving in just a subset of hubs along the conveyance way can satisfy better execution regarding store and server hit rates.

Prob Cache (PROB_CACHE): This system reserves content protests probabilistically on a way with a likelihood relying upon numerous components, including separation from source and destination and storing space accessible on the path [18].

Leave Copy Down (LCD): According to this methodology, one duplicate of a substance which s imitated just in the storing hub you jump far from the serving hub toward the receiver [19].

Random choice (RAND_CHOICE) reserve in one irregular reserve on way: This procedure stores the served content precisely in one single reserve on the way from serving hub to recipient chose haphazardly.

Random Bernoulli (RAND_BERNOULLI) reserve haphazardly in stores on way: In this procedure, a substance is arbitrarily embedded in a store on the way from serving hub to collector with likelihood.

3. MODIFIED HYBRID ASYMMETRIC-MULTICAST HASH-ROUTING

Hash-Routing consists of three embodiments, which are: Symmetric, Asymmetric, and Multicast, it proposed two initial hybrid strategies combine Multicast and Symmetric, the second combine Multicast and Asymmetric.

We use the second method, which works through utilizing the best in these two methods, where the asymmetric method , when an order has been sent, we conduct searches within the nodes , and after receiving the order, it turned it back through choosing the shortest route to the user, or by using minimum movement number between nodes. As for the Multicast, it takes the same route as the previous method, yet it may differ from the previous method, it throw on the first node to the user, then turn back from the second short route to deliver the order to the user, we chose this strategy and categorized it , as we mentioned earlier, because of the possibility of finding the best location for throwing, through calculating the number of nodes , that have been navigating through it, through which, we can calculate the best close place to all the rest servers, as shown in the following fees, that placed strategy mechanism after developing it.. As it is shown in below:

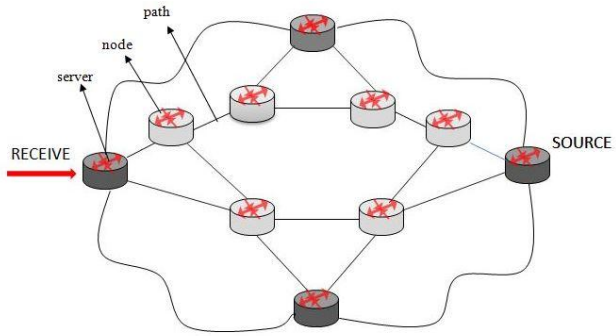

Fig. 1. Illustration of the nodes of servers before any request.

In Fig 1 illustrates the case of the contract before entering the demand and that demand after entering the edges begin work hash function and which directs the definition of the contract is responsible and which contain a copy of the request.

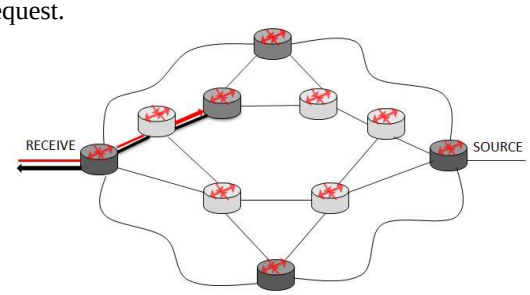

Fig. 2. Illustration in the case of finding demand in node nearby on the receiver.

In Fig 2 illustrates mode of action in the event of finding a copy of the user's request in a node nearby on the server where be take a copy of this node to the user and there is no need to go to the source.

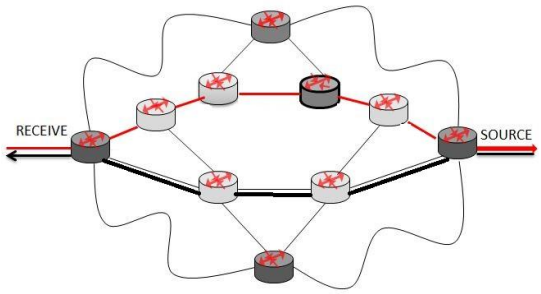

Fig. 3. Illustration go to the source and then choose the shortest way to return.

In Fig 3 illustrates the state of search in all nodes on the road, and that is not find a copy of the user's request, will goes to the source. After obtaining the copy that are required choose the shortest roads that lead to the user, and this is the first part played by this hybrid strategy.

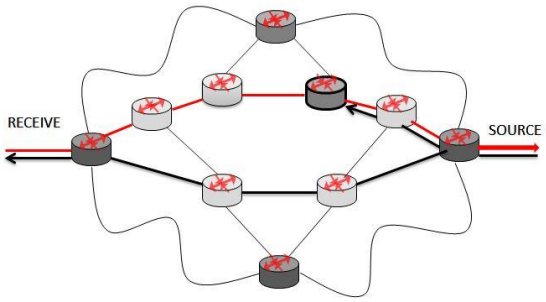

Fig. 4. Illustration to store a copy on the path before choosing the shortest route.

In Fig 4 illustrates the second part of the work of this hybrids strategy which works to add a copy of the user's request, which was obtained from the source in one of the nodes on the first track, which has been searching in all this nodes, in that path did not be find copy of the user request and then arrived to the source, after arriving are choosing the middle node in this path and opt for the middle node because of the amendment, which has been on this hybrids strategy, the purpose of this method is to be a copy in node nearby of all servers in this path, which is useful in not having to go to the source again.

4. PERFORMANCE EVALUATION &DISCUSSION

In this paper, simulations using Icarus [20], a demonstrator based on the Fast Network Simulation Setup (FNSS) [21]. Emulations have been executed on four real properties: GEANT (European academic network), GARR (Italian academic network), WIDE (Japanese academic network) and Tiscali (pan-European commercial ISP). Network routers have been allocated constant cache sizes. Content requests have been sculptured as Poisson processes while content popularity has been sculptured with a Zipf allocation. These results have been collected by conducting a set of simulations each one of them is consisting of 500 thousand requests, generated following a Poisson distribution. In conclusion, such a topology includes 63 nodes of which one is a source, 30 are caches and 32 are receivers.

5. RESULTS

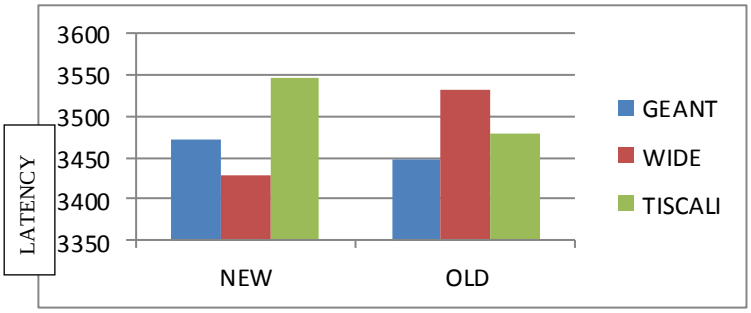

Fig. 5. Illustration of the results for latency.

Table 1: Illustration table of the results for latency.

TISCALI	WIDE	GEANT	LATENCY
3546.304	3427.994	3472.92	NEW RESULTS
3479.995	3532.336	3447.262	OLD RESULTS

Meaning of Latency: is the quantity of time a message takes to cut out a system. In a computer network, it is a term of how much time it takes for a package of data to get from one specified point to another. It is sometimes measured as the time needed for a package to be returned to its sender. We notice in fig 5 that the introduction of our modification to the Hybrid Asymmetric Multicast Hash Routing strategy achieved good results regarding the latency. We find that latency is less in WIDE, and the highest for GEANT and TISCALI because the number of nodes in this topology which must moving on her has increased, and the size of these large topology led to the rise of latency. Table (1) illustrates these results.

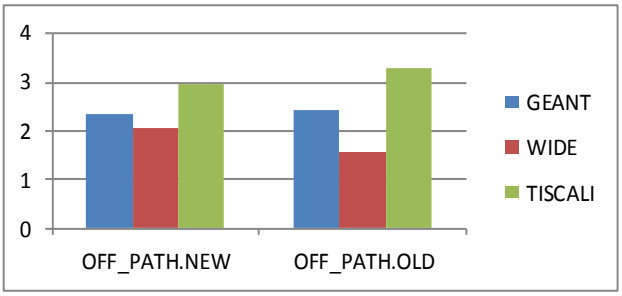

Fig. 6. Illustration the results finding cache in outside path.

Table 2: Illustration table the results finding cache in outside path.

TISCALI	WIDE	GEANT	CACHE_HIT_RATIO
2.973547	2.048167	2.359738	OFF_PATH.NEW
3.27579	1.553338	2.418363	OFF_PATH.OLD

Cache-hit ratio: is about how many times the database found something in cache divided by the number of times it looked for some object in the cache. The higher this ratio, the more efficient the cache is at enhancing performance. In fig 6 illustrates change in this strategy achieved high results in WIDE Which led to an increases cache-hit ratio, and the cause of decreasing in GEANT and TISCALI due to finding the cache on-path more of off-path, and this increases the number of searches ,so will decrease the cache-hit ratio in this topology. Table (2) illustrates these results.

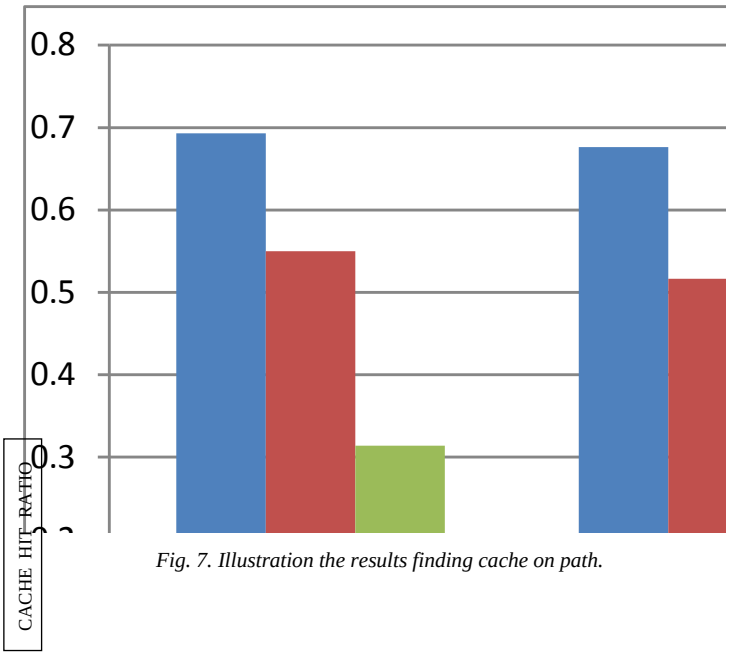

Fig. 7. Illustration the results finding cache on path.

Table 3: Illustration table the results finding cache on path.

TISCALI	WIDE	GEANT	CACHE_HIT_RATIO
0.31412	0.5498	0.693	ON_PATH.NEW
0.310378	0.516532	0.67627	ON_PATH.OLD

In fig 7 here we note the high results in all topologies, the reason is due to selection the best location to store a copy of the demand in this path , which was in the middle this way by choose code calculate the number of nodes in this path, and divisible on two to get the best site for storage. Table (3) illustrates this results.

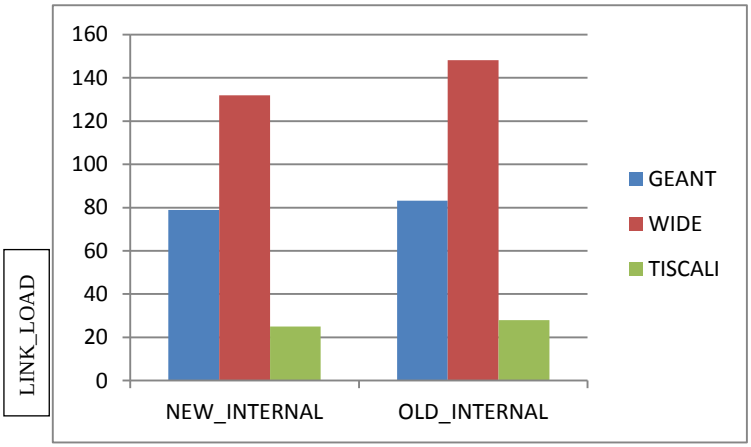

Fig. 8. Illustration of the results of links load internal.

Table 4: Illustration schedule for the number of links internal for node.

TISCALI	WIDE	GEANT	LINK_LOAD
25.03295	131.8343	78.94491	NEW_INTERNAL
27.92136	148.1583	83.15267	OLD_INTERNAL

LINK_LOAD: is the bandwidth that a traffic stream takes from the total link capacity, it has no relation with the link latency. Here we note in fig 8 a decrease in the number links internal ,the reason for this is that it has become easier access because he was chosen place of storage at a site close to all servers, which led to a decrease needed to number of links internal to find demand. Table (4) illustrates this results.

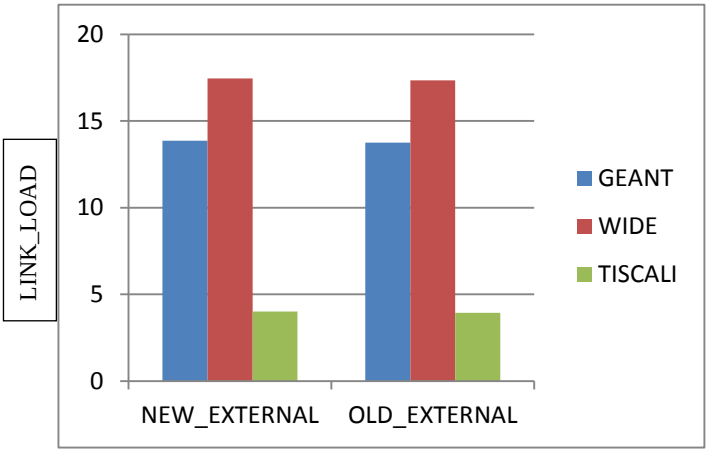

Fig. 9. Illustration of the results of links load external.

Table 5: Illustration schedule for the number of links external for node.

TISCALI	WIDE	GEANT	LINK_LOAD
4.015978	17.44415	13.86223	NEW_EXTERNAL
3.929826	17.34303	13.75726	OLD_EXTERNAL

Here we note in fig 9 a rise in the number of links external, the reason for this is that it has become get copies more because it has been selected the appropriate node for storage and in a nearby location on all servers. Table (5) illustrates this results.

6. CONCLUSION

We improved the performance of temporary storage in the ICN structure and this by improving the performance of Hybrid Asymmetric-Multicast Hash-Routing strategy which lead to the increase of cache hit ratio and this by comparing the Current results with the previous ones for the same strategy before the adjustment and the results of this simulation Shown that the performance of this strategy improved and it may be useful with the rest of strategies regarding the temporary storage in the ICN structure.

REFERENCES

[1] K. Pentikousis. et al., "INFORMATION-CENTRIC NETWORKING,"IEEE Commun. Mag, pp. 1-4, July. 2012.

[2] T. Koponen et al., "A Data-Oriented (and Beyond) Network Architecture,"proc. SIGCOMM 07,Kyoto, Japan, Aug. 2007.

[3] V. Jacobson et al., "Networking Named Content," Proc. CoNEXT, Rome, Italy, pp. 1-12, 2009.

[4] D. Trossen et al., "D2.3 – Architecture Definition, component Descriptions, and Requirements," Deliverable, PSIRP, Feb, 2013.

[5] B. Ahlgren et al., "Second NetInf Architecture Description," 4WARD EU FP7 Project, Deliverable D-6.2 v2.0, Apr. 2010.

[6] B. Ahlgren et al., " Design Considerations for a Network of Information," IEEE/ACM Trans., pp. 1-6, Dec. 2008.

[7] B. Ahlgren et al., " A Survey of Information-Centric," IEEE Commun. Mag, pp. 1-11, July. 2012.

[8] M. Dr¨axler, H. Karl., " Efficiency of On-Path and Off-Path Caching Strategies in Information Centric Networks," IEEE Commun. Mag, pp. 1-7, 2012.

[9] G. Xylomenos et al., " Caching and Mobility Support in a Publish-Subscribe Internet Architecture," IEEE Commun. Mag, pp. 1-7, July. 2012.

[10] A. Ghodsi et al., " Naming in Content-Oriented Architectures," SIGCOMM ICN'11, Toronto, Ontario, Canada. Aug 19, 2011.

[11] W. Chai et al., " Cache ''less for more'' in information-centric networks," Elsevier B.V. London, pp. 1-13, Jan. 2013.

[12] Z. Ming et al., " Age-based Cooperative Caching in Information-Centric Networks," IEEE INFOCOM, pp. 1-6, 2012.

[13] I. Psaras et al., " Probabilistic In-Network Caching for Information-Centric Networks," IEEE/ACM Trans., pp. 1-6, Aug. 2012.

[14] S. Borst et al., " Distributed Caching Algorithms for Content Distribution Networks," IEEE INFOCOM, pp. 1-9, 2010.

[15] S. Eum et al., " CATT: Cache Aware Target Identification for ICN," IEEE Commun. Mag, pp. 1-8, Dec. 2012.

[16] Y. He et al., " A Cache Strategy in Content-centric Networks Based on Node's Importance," ITJ., pp. 588.592, 2014.

[17] L. Saino et al., " Hash-routing Schemes for Information Centric Networking," IEEE/ACM Trans., pp. 1-6, Aug. 2013.

[18] I. Psaras et al., " In-Network Cache Management and Resource Allocation for Information-Centric Networks," IEEE Computer Society., pp. 2920-2931, Oct. 2014.

[19] N. Laoutaris et al., " The LCD interconnection of LRU caches and its analysis," Elsevier Science Publishers B. V. Jou, pp. 609-634, July. 2006.

[20] L. Saino et al., " Icarus: a Caching simulator for Information Centric Networking," ICST., pp. 66-75, Mar. 2014.

[21] L. Saino et al., " A Toolchain for Simplifying Network Simulation Setup," ICST., pp. 1-10, Mar. 2013.