

CORD: Content Oriented Routing with Directories

J.J. Garcia-Luna-Aceves,^{†*} Qian Li,^{*} and Turhan Karadeniz^{*}

^{*}Computer Engineering Department

University of California, Santa Cruz, CA 95064

[†]PARC, Palo Alto, CA 94304

Email: { jj, liqian, turhan }@soe.ucsc.edu

Abstract—Recent proposals for content routing in information-centric networks (ICN) require the use of content-based routing tables listing routes to name prefixes or individual named data objects (NDO), and a single naming space for NDOs. We present CORD (Content Oriented Routing with Directories) as an alternative to content routing in ICNs. CORD eliminates the need for large content-based routing tables by establishing routes to directories using distance-vector signaling and by mapping name prefixes or names of NDOs to directories using publish-subscribe mechanisms. Simulation experiments using the topology of a real ISP network are used to compare CORD with name-based content routing approaches based on link-state and distance information. The results show that CORD attains comparable data delivery and end-to-end delays, but incurs orders of magnitude less control overhead. In addition, CORD supports multipath forwarding of content requests and content.

I. INTRODUCTION

Several information centric network (ICN) architectures have been proposed [1], [2], [9] as alternatives to the current Internet architecture to address the fact that Internet usage is dominated by peer-to-peer communication and user-generated content. All ICN architectures aim at accessing content and services by name, independently of their location, in order to improve system performance and end-user experience.

Section II summarizes the prior work on routing schemes for ICN architectures. Most ICN architectures rely on name-based routing of content, which integrates name resolution and content routing. Routers advertise or compute routes to named data objects (NDO) or name prefixes, and content requests for specific NDOs are forwarded towards the nearest routers storing those NDOs. As our review of prior work in Section II indicates, prior content routing approaches assume that the entire ICN use the same naming space for NDOs and that routing tables list routes to NDOs or name prefixes, which incurs more overhead than routing to address ranges.

Section III presents CORD (*Content Oriented Routing with Directories*), which is an approach to content routing within autonomous systems in which directory nodes act as intermediaries to establish *virtual cords* linking consumers of content with content producers or caching sites. The primary objective of using directories between content producers or caches and the consumers of content is to reduce control overhead in the ICN. Instead of having routing tables listing routes to individual NDOs or name prefixes, they only list routes to the directories that maintain the mappings between name prefixes or NDO names and the locations where their

copies reside. This type of indirection in routing is inspired by McQuillan's work [13] on message addressing capabilities in the early days of the ARPANET.

CORD consists of three main elements: (a) maintaining multiple loop-free routes to directories that maintain the mappings from NDOs and name prefixes to the addresses of sites storing the content; (b) maintaining loop-free routes from directories to destination nodes nearby; and (c) publish-subscribe mechanisms for publishers and consumers of content to advertise and request content.

Section IV describes the results of simulation experiments used to compare the performance of CORD to the link-state approach advocated in NLSR [12] and OSPFN [19], and a loop-free routing approach based on distance vectors in which all the replicas of each NDO or name prefix are known. The protocols are compared using the AT&T network topology. The impact that network size, traffic load, and opportunistic caching have on performance is examined using packet-delivery ratio, average end-to-end delay, control plane overhead, and data plane overhead as the performance metrics. The results from these experiments show that CORD incurs orders of magnitude less control plane overhead than routing schemes that require all content replicas to be known while attaining the same packet delivery ratios and similar average delays.

II. RELATED WORK

The ICN architectures proposed recently advocate various ways to accomplish name resolution and routing, and all of them use on-path caching of content [2], [9]. Due to space limitations we only mention very few of these to contrast them to CORD.

Several ICN projects advocate using a link-state routing approach for intra-domain content routing, and adding content prefixes to BGP for inter-domain content routing (e.g., [5], [6], [4], [15], [18]). NLSR [12] and OSPFN [19] are two protocols for name-based routing of content within an autonomous-system. Routers exchange topology information by flooding two types of link states advertisements (LSA). LSAs can describe the state of physical links just as it is done in traditional link-state routing protocols. In addition, routers flood LSAs about prefixes for which they have copies. Gritter and Cheriton [8] proposed the name-based routing protocol (NBRP) as an extension of BGP. In essence, name-prefix reachability is advertised among content routers, and path information is used to avoid permanent loops. The routing

supports name-based content routing with routers having to know only how to reach directories, and directories having to know how to reach some routers.

B. Updating Routing Information

CORD uses a distance-vector routing approach to maintain routes to directories. To guarantee loop-free routes, CORD uses sequence numbers that restrict the selection of next hops towards a given directory, such that only those neighbors with shorter distances to the directory or with a more recent sequence number reported by the directory can be considered as successors. Algorithm 1 (DSU) is used to update routing information for directories.

Algorithm 1 DSU: Directory Status Update

```

1: Input:  $DT^i$ ,  $NT^i$ ;
2: if  $\exists q \in NT^i \mid sn_{cq}^i > sn_c^i$  then
3:   if  $(v = s_c^i) \wedge (sn_{cv}^i > sn_c^i) \wedge (d_{cv}^i = \infty)$  then
4:      $d_c^i = \infty$ ;  $sn_c^i = sn_{cv}^i$ ;
5:   else
6:      $d_c^i = \text{Min}\{d_{cf}^i + 1 \mid (f \in NT^i) \wedge$ 
7:        $(sn_{cf}^i = \text{Max}\{sn_{cv}^i \mid v \in NT^i\})\}$ ;
8:      $s_c^i = j \mid (j \in NT^i) \wedge (d_{cj}^i = d_c^i - 1)$ ;
9:      $sn_c^i = \text{Max}\{sn_{cv}^i \mid v \in NT^i\}$ ;
10:   end if
11: else
12:    $d_c^i = \text{Min}\{d_{cf}^i + 1 \mid (f \in NT^i) \wedge (sn_{cf}^i = sn_c^i) \wedge (d_{cf}^i < d_c^i)\}$ ;
13:    $s_c^i = j \mid (j \in NT^i) \wedge (d_{cj}^i = d_c^i - 1)$ ;
14: end if

```

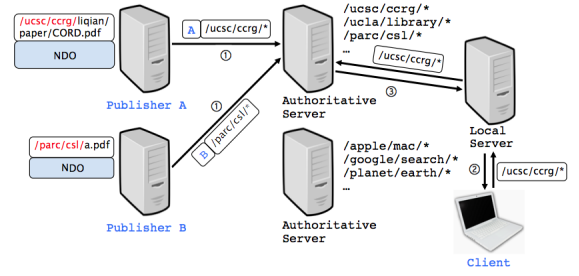
Let N^i be the set of one-hop neighbors of node i . Node i updates DT_j^i as a result of HELLOs from neighbor $j \in N^i$ or the loss of connectivity to neighbor j . If node i loses connectivity to node j , the entries in DT_j^i are deleted. The node, which is predefined as a directory, is the only one that can change the sequence number for its own entry in directory table updates sent in HELLOs.

If node i receives a HELLO from j or a link failure occurs that makes it update DT_j^i for entry $c \neq i : \{nid_{cj}^i, d_{cj}^i, sn_{cj}^i\}$, node i updates its entry for c in DT^i according to Algorithm 1 (DSU), which forces node i to propagate a reset update or to select a successor to directory c that is either closer to c or has reported a more recent sequence number from c .

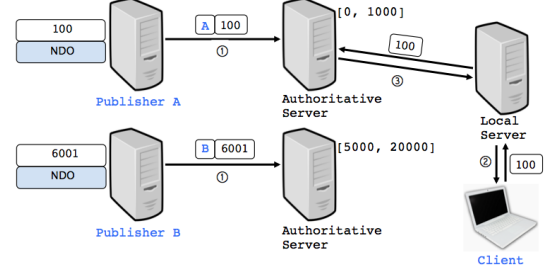
C. Mapping of Content to Directories

All routers maintain multiple loop-free routes to directories using sequence-numbered distances. The number of directories in the network is related to the network size and traffic load, which means the larger network and heavy traffic load the more directories are needed in the network to support scalability and provide highly efficient query processing. With the utilization of consistent hashing, changes of directories size do not affect the correctness and efficiency in CORD. CORD allows directories to announce to the entire network the name prefixes or intervals of NDOs for which they serve as anchors through the HELLO messages they send periodically. A router with content to publish extracts the name space of the NDO, and finds the corresponding anchor from its DT^i .

If hierarchical names are used, directories send updates about the list of the name prefixes for which they are anchors. For instance, a directory could announce being the anchor for



(a) Mapping of Hierarchical Name



(b) Mapping of Flat Name

Fig. 2: Mapping of Content to Directory

the name prefix “/ucsc/ccrg/*” and a content request for the NDO with name “/ucsc/ccrg/liqian/paper/CORD.pdf” would be sent to that anchor.

If flat names are used, directories announce the range of NDO identifiers they serve. Assuming an NDO has flat name 100, a directory could advertise a range from 0 to 1,000. Fig. 2 illustrates the consistent mapping to directories using NDO’s hierarchical name and flat name when publishing and subscribing to content.

D. Publish-Subscribe Mechanisms

1) *Publishing Content*: Publishing in CORD consists of having a few local directories know the routes to a given node storing the NDOs in a name prefix or a given range of flat names, and having an anchor directory know the mapping from an NDO range or name prefix to a list of local directories.

Algorithm 2 Publishing to Directories

```

1: Input:  $o^k$ ,  $DT^i$ ,  $LDL^i$ ,  $CST^i$ ,  $NT^i$ ;
2: if  $o^k \notin CST^i$  then
3:   /* data object  $o^k$  is new to node  $i$  */
4:    $CST^i \leftarrow \{o^k\}$ ;
5:   if  $i$  is origin of  $o^k$  then
6:      $i$  publish  $o^k$  to  $a_k$  with  $(i \leftarrow \{o^k\}, LDL^i)$ ;
7:     /*  $a_k$  is the anchor directory for  $o^k$  */
8:   end if
9:    $i$  publish  $o^k$  to  $l_k = \text{hash}(o^k)$  with  $(d \leftarrow \{o_d^k\}, i, LDL^i)$ ;
10:  /*  $l_k$  is the selected local directory from  $LDL^i$  for  $o^k$  */
11: end if

```

Each non-directory node i publishes itself with the k directories listed in its *local directory list* (LDL^i). If node i is attached to a content producer, it also publishes the existence of the content with one or more anchor directories, and with its local directories. The local directories in LDL^i are within r hops of node i and serve as the “landmarks” for other nodes to reach node i , given that nodes far away from node i do not have routes to node i . Accordingly, a local directory for

node i must maintain updated routes to node i , and it also maintains the mapping ($i \leftarrow \{o_i^1, \dots, o_i^n\}, LDL^i$), so that it can find alternate ways to reach node i if its route to i fails, and it can resolve subscription requests for content stored at i . The anchor directories are needed for nodes far away from content to obtain the mappings between the content and the local directories of the original content producer as a way to obtain the content. Algorithm 2 explains when and how to publish content to directories. For simplicity, we assume that a single anchor directory is selected for any one original content producer.

The forwarding of a publication message from a node to its local directories is done by the exchange of HELLOs. The routes maintained by local directories to nearby nodes are refreshed periodically based on a HELLO interval. A node that is the original source of an NDO publishes the existence of the NDO by sending a publication message to the anchor directory known to be in charge of the name prefix or NDO range to which the NDO belongs. An anchor directory stores the mapping from the identifier of a node i to the name prefixes or NDO ranges corresponding to NDOs stored at node i , and the list of local directories for node i (LDL^i). Nodes caching an NDO do not publish the NDO with the anchor directory of the NDO; they simply inform their local directories.

The submission of a publication message from node i to an anchor directory regarding NDO o_i^k is done by node i sending the message with the mapping ($i \leftarrow \{o_i^k\}, LDL^i$) towards anchor a . Each node v in the route from node i to directory a forwards the publication packet towards a and caches the mapping. Hence, the anchor directory and each node processing a publication message is able to redirect nodes sending subscription requests for NDO o_i^k to the local directories of node i .

2) *Subscribing to Content*: When node t needs to request NDO o^k , it first sends a subscription request to a local directory in LDL^t , which is selected using a hash function that computes $hash(o^k) = l_t$, where $l_t \in LDL^t$, and sends its subscription request towards l_t .

If directory l_t has received publication from local nodes regarding o^k before, it replies with the identifier of a node p where o^k exists, as well as the local directories of p . If l_t does not know about o^k , it sends a negative reply to node t . At that point, node t sends its request to the anchor directory a_k for o^k , based on the name prefixes and NDO ranges advertised by anchor directories. In turn, anchor directory a_k responds with the mapping ($d \leftarrow \{o_d^k\}, LDL^d$) towards the nearest local directory of node t selected from LDL^t . The answer is redirected to t by either the selected directory l_t^j or the first relay node along the path from a_k to directory l_t^j with a route to t . Hence, node t obtains a subscription response from either one of its local directories or the anchor directory of o^k . Node t is then able to send a content request according to the data-plane mechanisms defined for the ICN in which it operates.

IV. PERFORMANCE COMPARISON

We implemented CORD and other content routing protocols using the discrete event simulator QualNet [21] (version 5.0). We ran simulation experiments using the 154-node AT&T topology, which is well known. We compared the performance

of CORD with that of NLSR, which is based on the link-state approach, and a loop-free distance-vector approach to content routing. The distance vector approach uses sequence numbers to ensure that routes to destination nodes are loop-free, and nodes learn about all the replicas of each NDO in the ICN.

We use end-to-end delay, control plane overhead, data plane overhead, and packet delivery ratio as our performance metrics. The control plane overhead is the average number of control packets generated by the routing protocols, and the data plane overhead is the average number of data plane packets generated by the routing protocols, including subscription requests sent from data consumers to producers and forwarded content packets. We evaluated the three protocols in wired networks. The three protocols used the same time period to refresh their routing structures. For CORD we used a maximum distance of 3 hops to select local directories. Each simulation ran for 10 different seed values.

A few scenarios were used to evaluate the performance of three protocols in ICN by means of simulation experiments. We randomly selected a few groups in the AT&T topology, and within each group nodes are connected with each other through different paths. Nodes in selected groups can be content consumers requesting content. Each requested NDO exists in the network in such way that it is generated by one original content producer, but may also be cached anywhere else in the network. We investigated the impact of different scenarios on the performance of these three content routing protocols.

A. Impact of Number of Content-Requesting Group

In this scenario we evaluated the impact of increasing the number of consumer groups on the performance of the protocols. We started from 5 groups up to 10 groups, and the number of nodes in each group varied from 10 to 20. To minimize the influence from other parameters, there was only one content consumer from each group in this scenario. In addition, the number of content producers varied from 1 to 10, so that the number of content flows grew with the number of consumer groups at the same pace to avoid content sparsity in the network.

The results of this scenario with performance metrics used in comparison are shown in Fig. 3, where “1-data” means only one content producer existing in the network, and “5-group” means there are five content requesting groups. Given that the three approaches we simulated attain close to 100% delivery in all cases, so we do not show that metric in our results due to space limitations.

Fig. 3-a shows that CORD attains similar end-to-end delays of delivered content when the protocols have the same delivery ratios.

Fig. 3-b shows the average control plane overhead induced by the protocols. CORD incurs much smaller overhead and contrasts with the overhead induced by NLSR, which experiences a steep overhead increase for 10 groups. CORD incurs limited and fairly constant control overhead, because only unicast publish-subscribe requests to directories are sent other than HELLOs. By contrast, NLSR needs to flood link state advertisements (LSA) regarding the existence of new copies

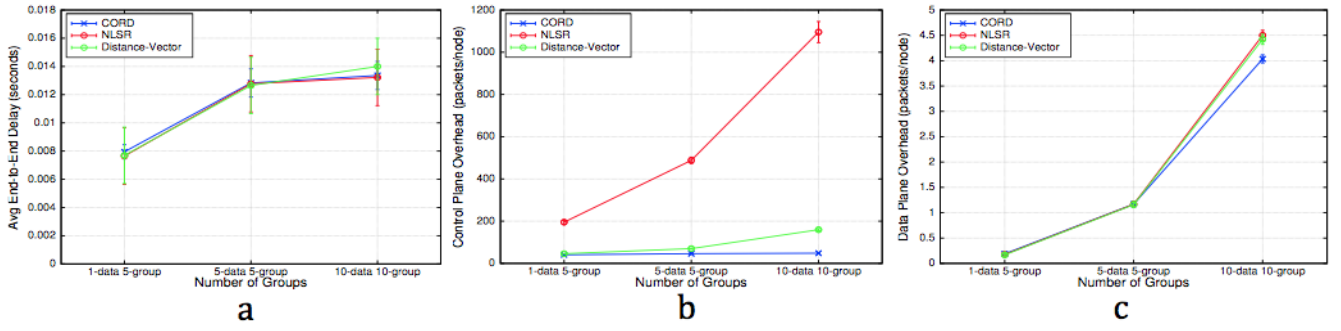


Fig. 3: Impact of increasing number of content requesting groups: (a) End-to-end delay of delivered data packets, (b) average number of control plane packets sent per node, (c) average number of data plane packets sent per node.

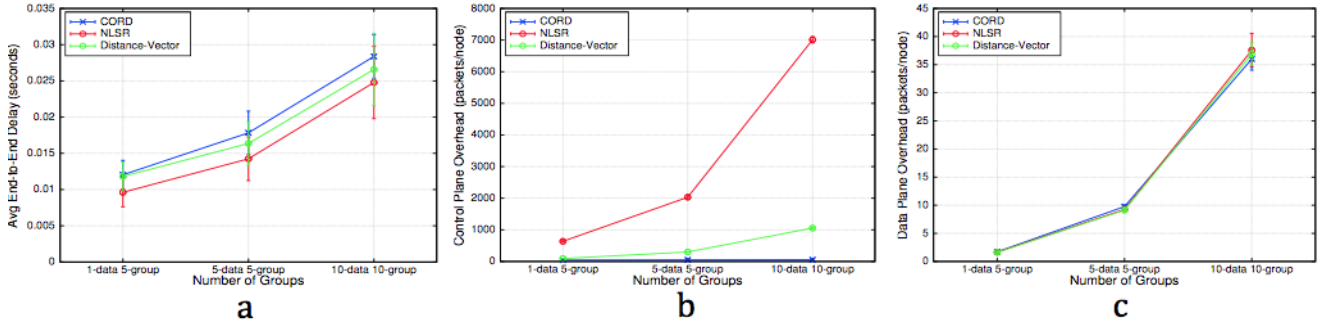


Fig. 4: Impact of increasing number of flows: (a) End-to-end delay of delivered data packets, (b) average number of control plane packets sent per node, (c) average number of data plane packets sent per node.

of NDOs, and the distance-vector routing approach needs to flood the network about the existence of new copies of NDOs at different nodes. Fig. 3-c shows the data plane overhead induced by three protocols. We can see that as the number of groups increases, the other two protocols incur more data plane overhead than CORD.

B. Impact of Increasing Number of Flows

In this scenario we increased the number of content consumers, which is the other way to increase the number of content flows, to see the impact on performance. We used the same selected consumer groups; however, nodes in each group are all content consumers and request NDOs at the same time. Fig. 4 shows the results in this scenario.

Fig. 4-a shows that CORD incurs slightly higher end-to-end delays to deliver NDOs than the other two protocols, which is mostly due to the fact that, in some cases, packets may take routes that are slightly longer than the shortest paths attained with the other two approaches.

Fig. 4-b shows the control plane overhead induced by the protocols. CORD incurs very limited and fairly constant control overhead, which contrasts with the overhead incurred by other two approaches for 10 groups with 10 data producers in the network.

Fig. 4-c shows that the data plane overhead induced by the three protocols is similar, even with an increasing number of flows. The key reason why CORD is more efficient than the other approaches is that CORD eliminates the need to communicate information about the network or content replicas.

C. Impact of Caching Scheme

To evaluate our caching scheme as well as compare the performance with the impact of caching, this scenario took in-network caching into consideration. We considered “path caching,” where all nodes cache content opportunistically when they forward NDOs to requesting nodes; and “edge caching,” in which only those nodes that request NDOs cache them. We evaluated the impact of these two caching schemes using the same scenarios described above.

Fig. 5, 6 show the results of using “path caching”, and Fig. 3, 4 present the results of using “edge caching”. In Fig. 5, there is only one content consumer in each group, whereas nodes are all content consumers in each group in Fig. 6. When we compare Fig. 3, 4 with Fig. 5, 6, “path caching” helps the three protocols to attain slightly lower end-to-end delays to deliver NDOs; however the difference is very small. When we compare the control plane overhead and data plane overhead in the figures, we find the difference by using two caching schemes is very small. Hence, the results indicate that “edge caching” provides most of the advantages of “path caching” with far less storage overhead.

V. CONCLUSION

We introduced CORD (Content Oriented Routing with Directories) as an alternative to content routing in ICNs. CORD eliminates the need for content-based routing tables by establishing routes to directories using distance-vector signaling and by mapping name prefixes or names of NDOs to directories using publish-subscribe mechanisms. CORD

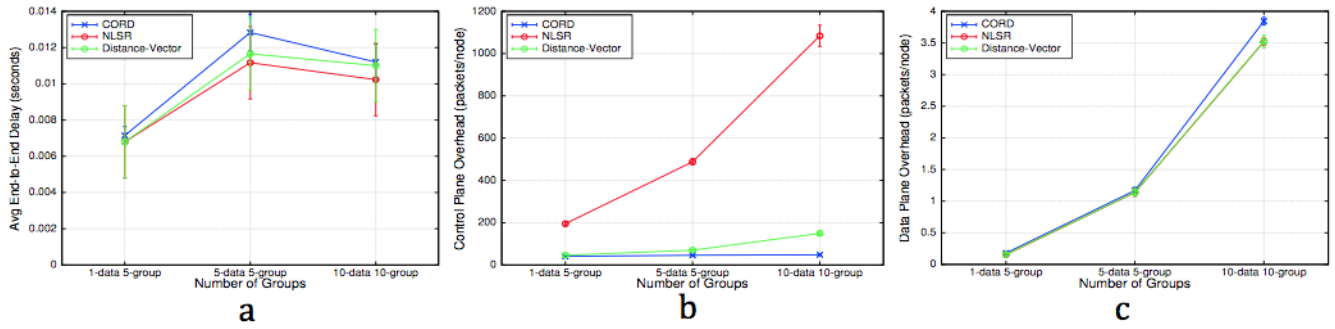


Fig. 5: Impact of caching, one consumer each group: (a) End-to-end delay of delivered data packets, (b) average number of control plane packets sent per node, (c) average number of data plane packets sent per node.

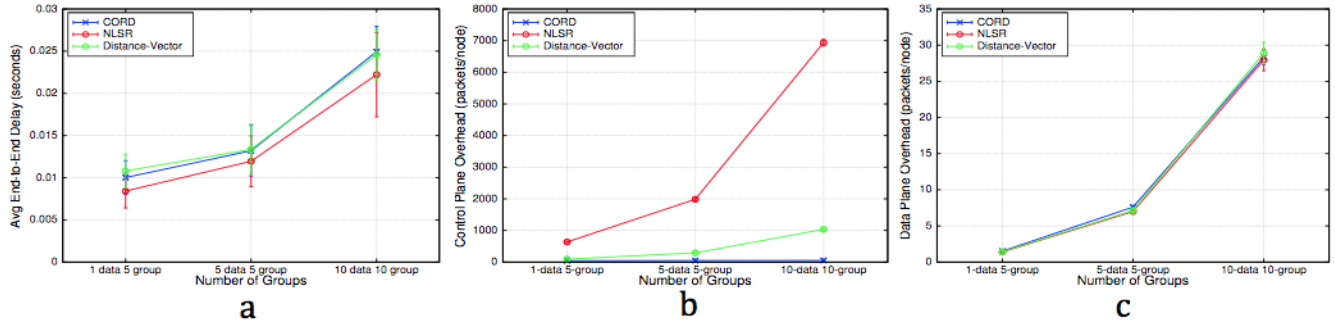


Fig. 6: Impact of caching, nodes are all consumers in each group: (a) End-to-end delay of delivered data packets, (b) average number of control plane packets sent per node, (c) average number of data plane packets sent per node.

constitutes the first approach for name-based content routing based on distance information to directories.

We used simulation experiments to compare its performance with that of name-based routing of content using the link-state approach advocated in NLSR [12] for ICN and a loop-free distance-vector approach. CORD achieves the same high data delivery, attains comparable delays to deliver NDOs, and incurs substantially less control plane overhead than the alternatives. The key reason why CORD outperforms the other name-based routing approaches is that it eliminates the need to maintain topology information or routing information for all the replicas of the same content.

VI. ACKNOWLEDGMENTS

This research was sponsored in part by the Jack Baskin Chair of Computer Engineering at UCSC.

REFERENCES

- [1] B. Ahlgren et al., "A Survey of Information-centric Networking," *IEEE Commun. Magazine*, July 2012, pp. 26–36.
- [2] M.F. Bari et al., "A Survey of Naming and Routing in Information-Centric Networks," *IEEE Commun. Magazine*, July 2012, pp. 44–53.
- [3] A. Carzaniga et al., "A Routing Scheme for Content-Based Networking," *Proc. IEEE Infocom '04*, 2004.
- [4] A. Detti et al., "CONET: A Content-Centric Inter-networking Architecture," *Proc. ACM ICN '12*, 2012.
- [5] Content Centric Networking Project (CCN) [online]. <http://www.ccnx.org/releases/latest/doc/technical/>
- [6] Content Mediator Architecture for Content-aware Networks (COMET) Project [online]. <http://www.comet-project.org/>
- [7] J.J. Garcia-Luna-Aceves, "System and Method for Discovering Information Objects and Information Object Repositories in Computer Networks," U.S. Patent 7,162,539, 2007.
- [8] M. Gritter and D. Cheriton, "An Architecture for Content Routing Support in The Internet," *Proc. USENIX Symposium on Internet Technologies and Systems*, 2001.
- [9] G. Xylomenos et al., "A Survey of Information-centric Networking Research," *IEEE Communication Surveys and Tutorials*, July 2013.
- [10] V. Jacobson et al., "Networking Named Content," *Proc. IEEE CoNEXT '09*, Dec. 2009.
- [11] T. Koponen et al., "A Data Oriented (and Beyond) Network Architecture," *Proc. ACM SIGCOMM 07*, 2007.
- [12] A.K.M. Mahmudul-Hoque et al., "NLSR: Named-Data Link State Routing Protocol," *Proc. ACM ICN '13*, Aug. 2013.
- [13] J.M. McQuillan, "Enhanced Message Addressing Capabilities for Computer Networks," *Proceedings of the IEEE*, Vol. 66, No. 11, Nov. 1978.
- [14] Mobility First project [online]. <http://mobilityfirst.winlab.rutgers.edu/>
- [15] NDN Project [online]. <http://www.named-data.net/>
- [16] Publish Subscribe Internet Technology (PURSUIT) Project [online]. <http://www.fp7-pursuit.eu/PursuitWeb/>
- [17] J. Raju, J.J. Garcia-Luna-Aceves and B. Smith, "System and Method for Information Object Routing in Computer Networks," U.S. Patent 7,552,233, June 23, 2009.
- [18] Scalable and Adaptive Internet Solutions (SAIL) Project [online]. <http://www.sail-project.eu/>
- [19] L. Wang et al., "OSPFN: An OSPF Based Routing Protocol for Named Data Networking," Technical Report NDN-0003, 2012.
- [20] C. Dannewitz, "NetInf: An information-centric design for the future internet", 2009.
- [21] S. N. Technologies. Qualnet. Available: <http://www.scalable-networks.com/>
- [22] C. Tsilopoulos, D. Makris, and G. Xylomenos, "Bootstrapping a Publish/Subscribe Information Centric Network," to appear In *Future Network and Mobile Summit 2011*, 2011.