A Survey of Information-Centric Networking Approaches

Dyon Buitenkamp, buiten.d@wustl.edu (A paper written under the guidance of Prof. Raj Jain)

Download

Abstract:

The demand for mobile data traffic is growing exponentially and will soon surpass what can be supplied with the traditional host-centric architecture. A promising solution to these problems leaves the legacy end-to-end communication behind and takes an information-centric approach. This novel architecture is known as Information-Centric Networking (ICN) and is based on a receiver-driven content retrieval model. The user will retrieve information based on the requested content and not based on a specific location where the content is stored. In this paper, the many important concepts in ICN are discussed, which include the different architectures and projects surrounding ICN. Challenges will also be discussed related to naming, routing, and caching in Information-Centric Networking.

Keywords:

Information-Centric Networking, ICN, Data-Oriented Network Architecture, DONA, Named Data Networking, NDN, Publish-Subscribe Internet Technology, PURSUIT, Scalable and Adaptive Internet Solutions, SAIL, naming, routing, caching, scalability.

Table of Contents

1. Introduction

With a growing society comes a need for growth in technology as well. In particular, the demand for mobile data traffic is growing exponentially and will soon surpass what can be supplied with the traditional host-centric architecture. The Cisco Visual Networking Index forecast predicts that by the year 2020, mobile and wireless devices will account for 71 percent of the total IP traffic, which leads to a sevenfold increase in mobile data traffic globally between 2017 and 2022 [CISCO19]. Part of the cause for mobile data traffic growth is the rising use of video traffic. By 2020, 82 percent of all IP traffic will be IP video traffic, which will be a fourfold increase from 2017.

These demands are being attempted to be met by many architectures, with the main one being the Content Delivery Network (CDN). CDN is a network that is not scalable to the growth of users and will be a problem in the near future. A promising solution to the problems of legacy networks leaves the traditional host-centric end-to-end communication behind, such as CDN, and takes a more information-centric approach. This novel architecture is known as Information-Centric Networking (ICN). In 1999, Stanford University introduced a name-based architecture called Translating Relaying Internet Architecture, integrating Active Directories (TRIAD). This was the beginning of the concept of ICN [Cheriton00]. ICN is based on a receiver-driven content retrieval model, where the traditional host-based IP addresses are replaced by an information-based scheme [Fang18]. The user will retrieve information based on the requested content and not based on a specific location.

This paper will introduce the concept of ICN, as well as the history of multiple ICN projects that have been proposed over the years. This will also include the different challenges the paradigm faces in terms of naming, routing, cache, security, mobility, and applications.

2. ICN Concepts

Information-Centric Networks uses a new internet communication approach that is data-oriented and does not focus on who is delivering its contents. Since this scheme is data-based, the paradigm relies on content naming to create an ICN, as well as routing and caching like in traditional networks. To fully understand what the benefits of ICN are, one must know the difference between the traditional networks and ICN.

2.1 Traditional vs. ICN

CDN is one of the main network architectures currently in use. The basic concept of CDN is having a network of servers close to the network edge with up to date content to optimize availability and latency [Sarkar16]. As the video traffic demand increase, as well as becomes more mobile, the placing of network edge servers and non-feasibility of 100% replications of the main servers, CDN is just not sufficient enough for the growing demand due to many users requesting the same server for the same content, which causes traffic overload. Not only is this relating to mobile devices, but the increase in IoT devices will add additional difficulty to the performance and reliability of the CDNs.

In CDN, a user may enter a web address in their browser, such as www.wustl.edu, and the browser sends this request to the Domain Name System (DNS). The IP address associated with the WUSTL server will be returned, and the user sends a request for the HTML page from that server. The server returns the HTML as well as the URL for the edge servers, which the DNS will translate into the corresponding IP addresses so that the user can direct their requests to the edge server. Although, if the edge server does not have the data that has been requested, it will redirect the request to the original server and cache the response for later requests. The usefulness of this paradigm depends on the storage space and the rate of requests, as well as the uniqueness of the requests.

The traditional network consists of hosts being named by their IP addresses, which is based on their location in the network. Therefore, since the name is bound to the address, and the address is based on location, the address will change when the physical location is changed. With a rise in mobile devices, this will be an increasing problem. With ICN, this is no longer a problem, because ICN does not request content based on the address of the host, but rather based on the content itself. Instead of needing to know a host, in ICN, just the name of the content will do.

2.2 Naming Approaches

There are many approaches to retrieving data in a name-based manner, which differs from the traditional architecture by not requiring a host's IP address to obtain content. This approach eliminates identification and location as the primary parameters when requesting data. To use the naming scheme, various parameters determine if the naming scheme is efficient enough. These are split into primary and secondary parameters. The primary parameters consist of identification, routing assistance, readability, security, and scalability [Krishna19].

Identification, also known as uniqueness, means that the naming scheme has to provide a unique name to all content that is not the same. This is to prevent collision with the contents being requested by a user.

Routing Assistance refers to the naming scheme making the data objects routable. In a network where content is requested based on content names, without contacting a server, the content will need to be easily reachable.

Readability pertains to the need for humans to be able to read the name of the content the user is querying. The naming scheme should, therefore, provide a name that is readable to a human as well as descriptive of what the contents hold.

Security is an important factor to verify the contents requested. The naming scheme will have to certify the data objects through self-certification or attestation by a third party. This is so that the user can authenticate the contents it requested without it being corrupted or falsified.

Scalability is one of the most important requirements of a naming scheme. The move from the traditional host-centric network to a receiver-driven network is precisely because of scalability issues with more mobile data traffic. If the naming scheme is not scalable, staying with CDN might be the better option.

The secondary parameters consist of name resolution, implementation complexity, caching effect, the effect on existing names, naming scheme translation ease, and ease of addition of existing content [Krishna19].

Name Resolution relates to the primary parameter of readability. Content should be able to be identified by its name. Depending on the naming scheme, the lookup that maps the name to the data object might differ.

Implementation Complexity refers to compatibility with legacy architectures. As with IPv4, one cannot disregard all IPv4 addresses and only use protocols to read IPv6 addresses. The naming scheme must be complex enough to implement the new paradigm, but not complex enough to not work along with traditional networks.

The caching effect in a naming scheme goes hand-in-hand with various caching schemes in ICNs. To improve efficiency in a network, caching is an important factor and extensive research is being done, which is beyond the scope of this paper. Since the naming scheme is a major factor in caching, it needs to be helping and not hurt the caching techniques.

The effect on existing names should be flawless. Multiple naming schemes should be able to coexist in a network without causing conflicts. This includes legacy networks as well as different ICN naming schemes. Since these schemes will need to coexist, another parameter, **naming scheme translation ease**, is an important factor for naming schemes to be able to translate easily between each other.

Ease of addition of existing content to a new naming scheme is very important. Content that is currently uploaded needs to be able to be used in an ICN and might be a challenge since it cannot be removed.

There are many factors to take into consideration when choosing or creating a naming scheme, and there are trade-offs, which results in multiple different schemes. The naming schemes to identify content in ICNs are categorized as follows: flat naming, hierarchical naming, attribute-based naming, and hybrid naming.

Flat Naming uses a cryptographic hash function to create a flat content identifier [Koponen07]. In flat naming, a string consisting of multiple characters is created and used as the output of the hash function, which is a globally unique principal field, *P*. This is then bound to the data object referred to as the Named Data Object (NDO), which has a unique object label, *L* [Ahlgren12]. This is then put together in a *P* : *L* form.

Hierarchical Naming consists of concatenating multiple components, after which a unique identifier can be created and assigned. This is very similar to the structure of Uniform Resource Identifiers (URI), which means that the current mechanisms for handling IP addresses can easily be made to handle hierarchical names. This format also makes it more user-friendly and is scalable because names can be aggregated [Dannewitz10].

Attribute-Based Naming does not request content by a name as the other two naming schemes but instead uses attribute-value pairs (AVP) to identify contents. For example, the AVP could look something like: [class="alert", severity=6, device-type="web-server", alert-type="hardware failure"], which would match a set of constraints that would identify the right contents [Carzaniga03].

Hybrid Naming is an attempt at an all-benefit solution to naming schemes. It tries to take advantage of each of the other three naming schemes and make the disadvantages have a lesser impact [Zhang14]. In this naming scheme, the first part of the name will be hierarchical, and the second part of the name will be flat naming. Then the third part will add attributes to the name based on the attribute-based naming scheme. Having the first part hierarchical gives the benefit of easier aggregations, while flat naming of the second part will limit the length, which improves query speed for length-based algorithms.

Multiple projects have been in the works over the years, focusing on efficient ICNs, which includes efficient naming schemes. Table 1 [Krishna19] illustrates the different advantages and disadvantages of each naming scheme that contributes to the decision of what scheme to use in various ICN architectures.

Table 1: Comparison of Naming schemes and their parameters

	Flat	**Hierarchical**	**Attribute-Based**	**Hybrid**
Identification	Yes, unique NDO to name mapping	Unable to guarantee	Unable to guarantee	Yes, due to the flat portion
Routing Assistance	Cannot assist with routing	Yes, due to tree-like nature	Cannot assist with routing	Yes, due to the hierarchical portion
Readability	No, unreadable to humans	Is possible	Yes, names humans relate to	Dependent on attributes
Security	Yes, due to the hash functions	Unable to validate content	Does not help security	Yes, due to the flat portion
Scalability	Not scalable	Scalable	Undetermined	Undetermined
Name Resolution	Good	Can be easy and difficult	Not good. Lack of uniqueness	Good
Implementation Complexity	Need for third party	High (undesirable)	High (undesirable)	High (undesirable)

Caching effect	Neither hinders or helps caching	Negatively impacts	May negatively impact	Helps caching
Effect on existing names	Does not affect positively or negatively	Does not affect positively or negatively	Needs separate mechanism	Does not affect positively or negatively
Naming scheme translation ease	Easy, with third party	Difficult	Difficult	Easy
Existing content addition ease	Very easy	Challenges in differentiating content	Difficult to generate rule, naming is easy	Resource consuming and challenging

Flat, hierarchical, attribute-based, and hybrid naming schemes cannot just be chosen based on their benefits by themselves, but one must also consider how these naming schemes affect the routing in the network.

2.3 Routing Techniques

In an ICN, the requested content is requested by name and will, therefore, need to be routed to the correct place. Since the requester does not know the location of the contents, efficient routing has to be done, and the network will need an existing knowledge of where valid copies of the content are. Therefore, there are four important characteristics regarding content or name-based routing: content-oriented, robustness, efficiency, and scalability [Velloso13]. With these characteristics, intuitively it is known that packets should not have any location information; routing should quickly recover from faults, control information should have low impact, and routing should work well in all scenarios and topologies, respectively. These are the characteristics, but the properties of routing mechanisms that are important to keep in mind, although outside the scope of this paper, are: scalability, content state, discovery of closest copy, resolution and retrieval locality, discovery guarantee, network-level deployment, and security. The ICN routing techniques can also be grouped into two categories: non-hierarchical and hierarchical routing.

Non-Hierarchical Routing does not have any structures in place to store routing information and is, therefore, also referred to as unstructured routing. It is non-hierarchical because it does not organize routers in a hierarchical structure, which leads to the lack of a central/root node where the routing information would have been stored. The routing information would have to be calculated at each node as the packets flow through the network.

Hierarchical Routing, as the name suggests, organizes the routers in a hierarchical structure, and this hierarchical relationship can be used to have a deterministic flow of data going through the network. Therefore, *hierarchical routing* is also referred to as *structured routing*. In ICN, there are two main architectures of hierarchical routing, which are tree-based and distributed hash tables (DHT). DHT uses a hash table that is distributed among the nodes in the network and provides an efficient way for routing by making queries for cryptographic keys in an overlay network [Ganesan04]. Tree-based topologies need knowledge of affiliation, parity, superiority or inferiority to determine the location of the desired destination node. This is done by having parent nodes act as gateway nodes between their children. The request only goes up to the highest hierarchical level it needs to, based on the subtree the destination node belongs to. This is similar to the concept of Destination-Oriented Directed Acyclic Graphs (DODAGs).

The efficiency and safety of routing are very important and limiting the amount of routing that needs to be done will ultimately help with those characteristics. One way to limit the amount of routing is by caching. Caching in ICN is a key component since it will limit how far requests will need to go.

2.4 Caching in ICN

In the traditional CDNs, caching is based on content popularity, which means that if there is a large amount of request for certain content, it will be cached closer to the edge of the network. Caching techniques in ICNs work a little bit different because it does not have centralized management that CDNs do. ICN uses local information alone to determine what to cache. Any node in the network can act as an edge-node cache at any time, which means that the existing network can be made into content distribution networks since the content requests do not rely on location or host.

3. ICN Architectures

In the previous section, multiple components of ICNs were discussed. The differences in all these schemes are put together in many different ways, which is why multiple projects have risen to improve the concept of having an efficient content-based network. TRIAD was the first name-based transportation-layer architecture introduced by Stanford University in 1999 [Cheriton00]. It allowed for better naming support and opened the door for future development. Data-Oriented Network Architecture (DONA) was one of the first ICN architectures that were based on clean-slate concepts and was presented in 2007. Many new projects followed suit over the years, both U.S. funded projects and EU funded projects.

As shown in Figure 1 [Fang18], projects dedicated to ICN have come and gone. The U.S. funded projects such as Content-Centric Networking (CCN), Named Data Networking (NDN), MobilityFirst, and ICE-AR, while the EU funded Publish-Subscribe Internet Technology (PURSUIT), Scalable & Adaptive Internet Solutions (SAIL), Content Mediator Architecture for Content-Aware Networks (COMET), CONVERGENCE, and ANR Connect. Other collaboration projects include GreenICN and ICN-2020 [Tourani17].

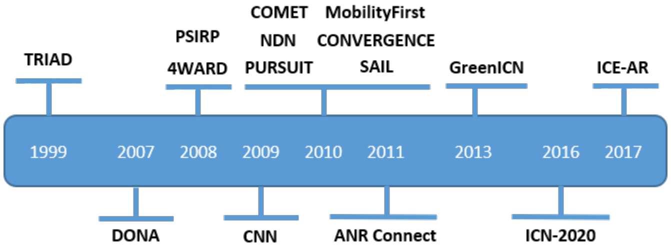


Figure 1: Timeline of ICN projects

Although these architectures all bring different elements to the table, many are evolving from an older project into newer projects. These architectures will be discussed in more detail in this paper, starting with the first completer ICN architecture, DONA.

3.1 Data-Oriented Network Architecture

Data-Oriented Network Architecture (DONA) was the first ICN architecture to be completed. This project was brought forth by UC Berkeley and built on TRIAD's design principles, where the hierarchical URLs are replaced by flat names [Koponen07]. With DONA, information availability increases since it is not relying on just one host to provide the contents.

As mentioned, DONA uses a flat naming scheme which consists of a cryptographic hash of the publisher's key, as well as an object ID assigned by the publisher. The name is unique within the domain of the publisher. It is suggested that publishers use a hash of the object as the identifier. The way DONA is composed can be observed in Figure 2 [Xylomenos13]. There is a network of interconnected resolution handler (RH) entities, which will retrieve and publish objects in a hierarchical fashion.

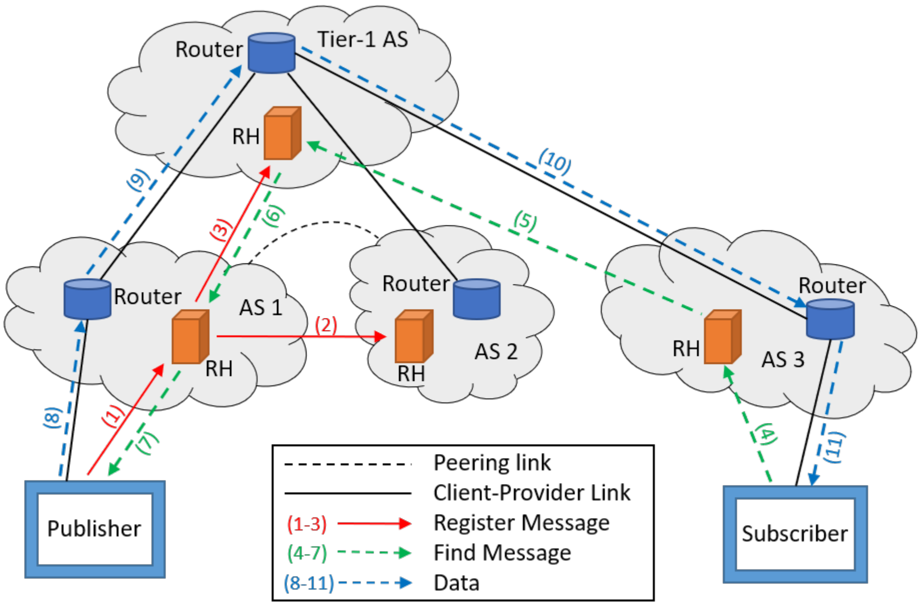


Figure 2: DONA Architecture

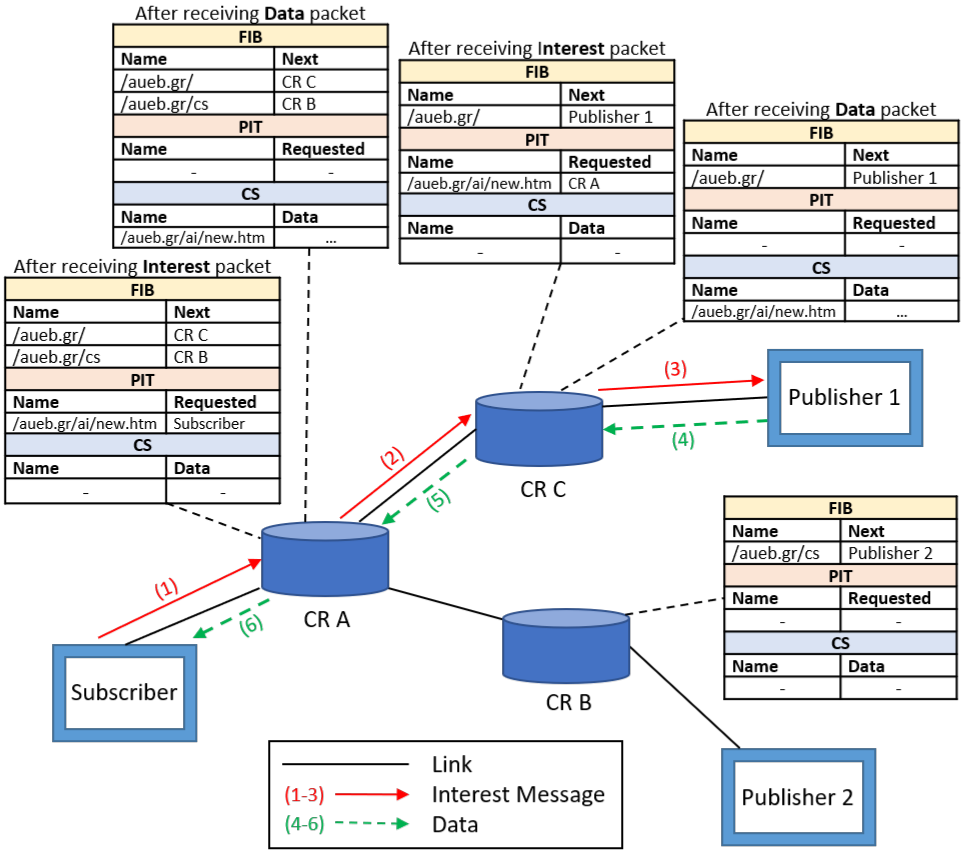
Following the process in Figure 2, to publish content, the publisher sends a Register message to its local RH (1). The local RH then forwards this message to its parent and peering RHs (2-3), who will store a mapping of the pair containing the local RH and the contents name. When a subscriber is interested in the contents, it will request it by sending out a Find message. This Find message will be sent to the subscriber's local RH (4). The message will propagate through the hierarchical tree of RHs until it finds a match mapping the contents to a local RH (4-5). With found mapping, propagation will take place to send the requests from that RH to the publisher (6-7). Tier-1 providers are at the top of the hierarchical tree and therefore know all existing objects in the network. DONA offers two ways of responding to the Find message: the data is sent using regular IP routing and forwarding, or the data is sent using domain-local IP addresses.

This second option is depicted in Figure 2, where it uses the path-labels gathered by the Find message and backpropagates from RH to RH until it reaches the subscriber (8-11).

3.2 Named Data Networking

Researchers at the Palo Alto Research Center first introduced CCN in 2009. CCN is a data named communications architecture with packets being addressed with content names rather than location [Jacobson09]. The design principles of CCN were quickly applied to other projects, of which one was NDN, which was chosen as one of four projects focusing on the future of the internet to receive funding from the U.S. National Science Foundation (NSF) [Zhang10]. In NDN (as well as in CCN), a client requests content by sending an Interest packet into the network with the content's name.

NDN uses hierarchical naming, which is a naming scheme very similar to URLs, as mentioned before. Although similar, NDN names are not the same as URLs, as hierarchical naming does not have DNS names or IP addresses. Figure 3 [Xylomenos13] shows a simplified example of how NDN works.



After receiving **Data** packet

FIB	
Name	Next
/aueb.gr/	CR C
/aueb.gr/cs	CR B
PIT	
Name	Requested
-	-
CS	
Name	Data
/aueb.gr/ai/new.htm	...

After receiving **Interest** packet

FIB	
Name	Next
/aueb.gr/	Publisher 1
PIT	
Name	Requested
/aueb.gr/ai/new.htm	CR A
CS	
Name	Data
-	-

After receiving **Data** packet

FIB	
Name	Next
/aueb.gr/	Publisher 1
PIT	
Name	Requested
-	-
CS	
Name	Data
/aueb.gr/ai/new.htm	...

After receiving **Interest** packet

FIB	
Name	Next
/aueb.gr/	CR C
/aueb.gr/cs	CR B
PIT	
Name	Requested
/aueb.gr/ai/new.htm	Subscriber
CS	
Name	Data
-	-

FIB	
Name	Next
/aueb.gr/cs	Publisher 2
PIT	
Name	Requested
-	-
CS	
Name	Data
-	-

Figure 3: NDN Architecture

To request data objects, subscribers in an NDN network sends out an Interest packet. The Interest is sent to a Content Router (CR) and will be forwarded hop-by-hop across CRs. Each CR maintains three data structures: *Forwarding Information Base* (FIB), *Pending Interest Table*

(PIT), and *Content Store* (CS). The FIB stores pairs of names to forwarding output direction. This is so that Interest packets can be forwarded to the correct place based on the requested content name. The PIT stores pairs of names to which interface requested the content. This is then used to backpropagate data objects back to the subscriber. Lastly, the CS is, as the name suggests, storage for content that travels through the interface. This is a local cache. Once an Interest packet arrives at the CR, the FIB, PIT, and CS in the CR are updated.

The CRs in the example in Figure 3 have existing content names mapped to output interfaces. The subscriber sends out an Interest packet for content with the name */aueb.gr/ai/new.htm* (1). CR A receives the packet and checks the FIB where to forward the request, which is CR C. The Interest packet is then forwarded to CR C (2). The PIT in the CR A stores the name of the request and where it came from for when the data comes back. In CR C, the same process takes place, but this FIB has an output interface directed to Publisher 1. The Interest packet is, therefore, forwarded to Publisher 1 (3). The Interest message is now disregarded, and a Data message is returned to CR C (4). The CR C checks its PIT to determine where to send the Data message based on where the request came from. Once the Data message is sent (5), the CS is updated to store the data in case it is requested again. The CR A continues this process, and the subscriber receives the requested data (6). The names in the FIB and the requested data-name might differ, yet still be valid. This is because the names in the FIB may be prefixes. The content name used in the example was */aueb.gr/ai/new.htm*, which can still be forwarded under prefixes such as */aueb.gr/* or */aueb.gr/ai/*.

3.3 Publish-Subscribe Internet Technology

PURSUIT is an EU funded project and has the Publish-Subscribe Internet Routing (PSIRP) architecture as its predecessor [Tarkoma09]. This architecture has a complete publish-subscribe protocol stack, instead of the traditional IP protocol stack. PURSUIT has three core functions: *rendezvous*, *topology management*, and *forwarding*. Rendezvous refers to a collection of Rendezvous Nodes (RNs) known as Rendezvous Network (RENE). The topology management is made up of Topology Managers (TMs), while the forwarding function is controlled by the Forwarding Nodes (FNs).

The naming of data objects in the PURSUIT architecture is a more generalized version of the DONA naming scheme. The naming scheme is composed of a unique pair of IDs called scope ID and rendezvous ID. The scope ID relates to the grouping of contents that are related, while the rendezvous ID is a unique ID within the group it is in. As with the previous architectures, in Figure 4 [Xylomenos13], an example is given of how the publish-subscribe architecture works.

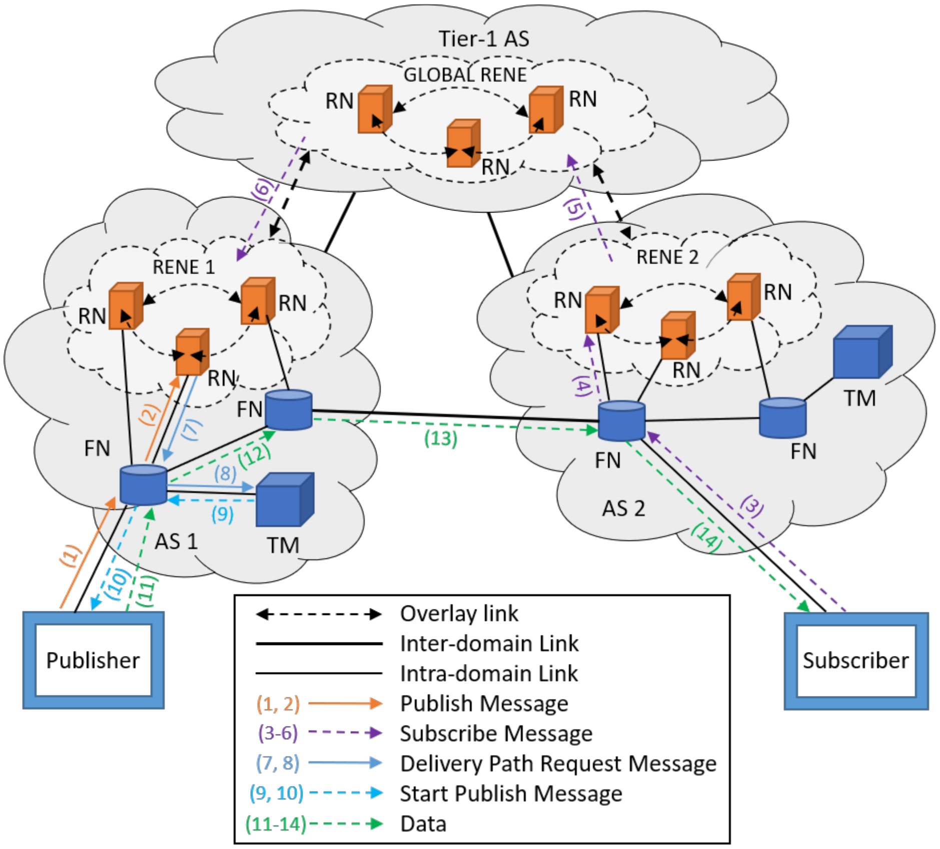


Figure 4: PURSUIT Architecture

As the figure suggests, PURSUIT has a hierarchical routing structure and uses the DHT routing architecture. When a publisher wants to make its contents known to the network, it will send out a Publish message to its local RN (1-2), which is directed there based on the scope ID by the DHT. Then, if a subscriber wants this content, it will issue a Subscribe message for this content to its local RN, which will use DHT to route the message to the publisher's local RN (3-6). The RN then sends a Delivery Path Request message to the TM, and, as the name suggests, the TM then manages the topology by creating a route between the publisher and the subscribe (7-8). The TM creates a routing Bloom filter for each flow and adds it to the packet's header. The Bloom filter takes the tags of the links, which are based on hash functions, and aggregates these. The TM then sends a Start Publish message to the publisher with the route it created (9-10), which the publisher then uses to send the data through a set of FNs back to the subscriber (11-14).

3.4 Scalable and Adaptive Internet Solutions

PSIRP and PURSUIT were not the only projects funded by the EU Framework 7 Programme. At the same time, as when PSIRP was introduced in 2008, the Architecture and Design for the Future Internet (4WARD) was introduced as another promising project. A few years later, the 4WARD architecture was developed into a new and improved architecture, Scalable and Adaptive Internet Solutions. This architecture combines two of the previously discussed projects, NDN and PURSUIT. SAIL combines elements of these two architectures and can also act as a hybrid [SAIL01]. Since NDN uses hierarchical naming and PURSUIT uses flat naming, the naming scheme for SAIL can be seen as "flat-ish." Names in the SAIL architecture do not include location, but rather names consisting of an authority part and a local part that is linked to the authority. These can be in hash form and, therefore, not readable to humans. For comparison purposes, a flat naming scheme is used but can be considered hierarchical when used for routing because the longest prefix matching can be used. To further explain how the SAIL architecture functions, a simple example is shown in Figure 5 [Xylomenos13].

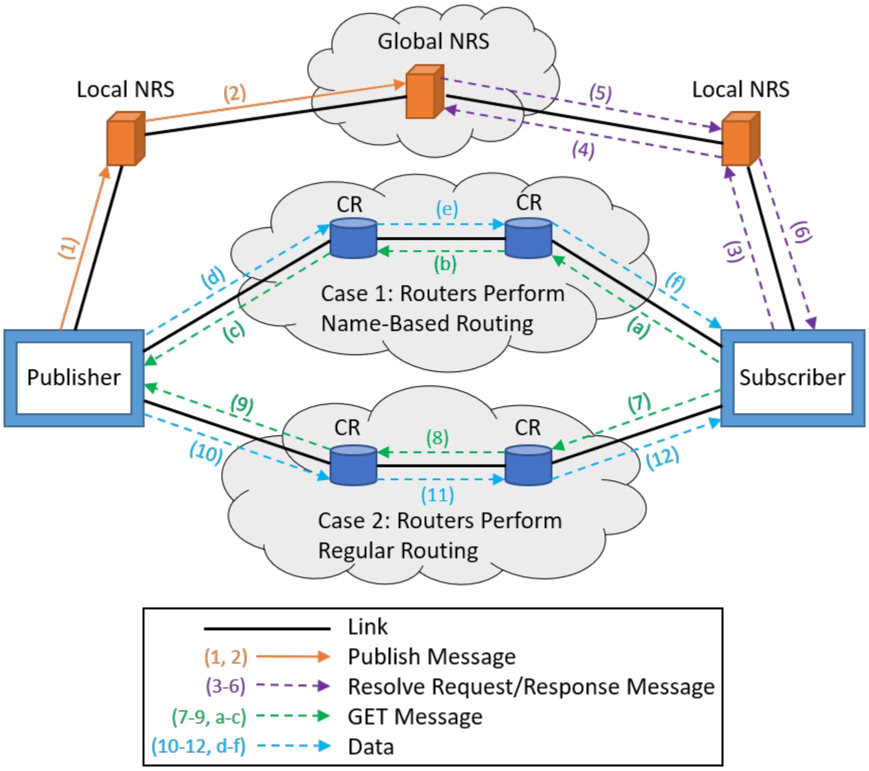


Figure 5: SAIL Architecture

The local and global Name Resolution Systems (NRS) are used to map data object names to locators. Locators are used to find corresponding data and reach it, similar to IP addresses. As described in the naming scheme of SAIL, there is a local and authority part. The local NRS handles the local part of the naming resolution, and the global NRS handles the authority part. Say a publisher makes its contents available by sending a Publish message to the local NRS (1). The local NRS will receive this message, which will also include a locator, which will be stored. Any local part that belongs to the same authority will be aggregated into a Bloom filter in the local NRS. The Publish message is then sent to the Global NRS (2). The global NRS then stores all the mappings and replaces any old mappings. If a subscriber wants to request content, it will issue a Get message to its local NRS, which forwards it to the global NRS (3-4). The global NRS will return the locator it has mapped back through the local NRS and to the subscriber (5-6). Now that the subscriber has the locator, it can send the Get message to the publisher. When the name resolution and data routing in the SAIL architecture are decoupled, the subscriber sends the Get message through *Case 2* (7-9), and the publisher will respond with the data in a Data message (10-12). Although, when the name resolution and data routing in the SAIL architecture are coupled, it is similar to NDN where Content Routers (CRs) are used for routing. In *Case 2*, a

Get message is sent from the subscriber to its local CR (a), which forwards the message until the publisher is reached (b-c). The contents will be sent back by the publisher in the form of a Data message that will propagate back the same way the Get message came (d-f).

3.5 Others

There are many other interesting projects out there offering different key components.

The Network of Information (NetInf) was started in the 4WARD architecture [Ahlgren08]. NetInf uses a publish-subscribe scheme as well as flat naming, which maps names to locators. Content is retrieved by having the publisher first publish its data objects to the network, and NRS will store the name and locator mapping. Data can be stored in cache in many places, and there can, therefore, be many locators. Once a subscriber requests the contents, the routing forwarders will deliver the data.

MobilityFirst was funded by NSF in 2010, as with NDN, which includes an architecture that focuses on mobility, wireless links, multicast, multihoming, security, and in-network caching [Seskar11]. In other words, as the name suggests, in this architecture, mobility comes first. It uses a globally unique identifier (GUID), which a publisher requests from the naming service. The name is then registered with a global name resolution service (GNRS). Subscribers use the Name Certification Service (NCS) to get the data object name. A request for contents will then contain the GUID and is then mapped to an actual address by the GNRS. Routers will continuously contact the GNRS to get an updated destination address since this architecture focuses on mobility. The publisher can then send back the contents once the Get message arrives.

Content Mediator Architecture for Content-Aware Networks (COMET) was funded by the same EU Framework 7 Programme as PURSUIT and SAIL [COMET01]. This architecture uses mapping of information to optimize the selection of data sources as well as their distribution. COMET uses a Content Mediation Plane (CMP) that works as a coordinator between the network providers and the content servers. This creates a network and content awareness. COMET has two different architectures: a coupled design and a decoupled design. The coupled design is called Content-Ubiquitous Resolution and Delivery Infrastructure for Next Generation Services (CURLING), which has naming resolution and data routing coupled. The decoupled design does not fundamentally change the underlying internet, yet improves information delivery.

Content-Centric Inter-Network (CONET) is an architecture that proposes a CONET layer to provide network access to content by name instead of to a host [Detti11]. CONET used a unique Network Identifier (NID) consisting of a name based on data or service. CONET also focuses on local caching, which has an in-network functionality to support efficient routing.

CONVERGENCE is, again, also funded by the EU Framework 7 Programme, but focuses on user access to data [CONV01]. This access includes everything from digital data to real-world objects and is represented by a Versatile Digital Item (VDI). It uses a CONET component to have publishers make VDIs available for subscribers to show an interest. CONVERGENCE reuses existing functionalities to make the transition from IP easier.

ANR-Connect, **GreenICN**, **ICN-2020**, and **ICE-AR** are four different architectures that are worth mentioning. GreenICN is working on bridging the gap between the ICN network and devices so they can be highly scalable and energy-efficient [GREENICN01]. ICN-2020 is developing a concept that focuses on video delivery, IoT features, and cloud services [ICN2020]. Lastly, the ICE-AR is developing a new wireless network architecture focusing on the limitations of emerging applications such as augmented reality [ICEAR01].

The architectures offer various benefits, but also come with many disadvantages. In the next section, the improvements in ICNs that are still being researched will be discussed.

4. ICN Challenges

Many challenges come with creating a new network. In the traditional network, the major issue is the exponential increase in mobile data traffic. Soon, the legacy network will not be able to supply what is being demanded, which is where the concept of ICNs started. The challenges that come with the Information-Centric Networks will briefly be discussed in the following subsections.

4.1 Naming

Naming in ICN is an important component, as seen in previous sections. The properties of content names should include global uniqueness, security, location independence, and human-friendliness. None of the previously mentioned architectures satisfy all of these properties but come a lot closer to fulfilling some of these than legacy networks do. Naming is difficult to evaluate properly since different naming schemes provide different advantages depending on the architecture the scheme is implemented in.

4.2 Routing

One important factor in routing has yet to be explored in the ICN architectures: management of the content state. Although there are different routing mechanisms proposed in the various ICN architectures, as with naming, they all provide different advantages. The main challenge with routing is still that there is no way to delete content or a way to update metadata of content. These are just some of the many routing issues in ICNs that do not satisfy the properties routing mechanisms should have: scalability, content state, discovery of closest copy, resolution and retrieval locality, discovery guarantee, network-level deployment, and security.

4.3 Caching

In ICNs, one main concept of storage is to cache in all routers in the network. This is to improve latencies and performance. Storage of content has always been an issue, and there are, therefore many research areas within the topic of cache: analytical models for network caches, replacement policies, and storage policies. Some of the challenges in these areas consist of the complexity of large-scale network caches, discarding the right content when new content arrives at the cache, and efficiently storing contents in the cache.

5. Summary

With a growing society, the demand for mobile data traffic has been growing exponentially and will soon surpass what can be supplied with the traditional host-centric architecture. Between 2017 and 2022, there will be a sevenfold increase in mobile data traffic. ICN looks to improve this problem by replacing the traditional host-based IP addresses with an information-based scheme. The new communications network has three core concepts: naming, routing, and caching.

Naming schemes consist of a flat naming, hierarchical naming, attribute-based naming, or hybrid naming. The performance of these schemes can be evaluated based on their parameters, where the primary parameters are: identification, routing assistance, readability, security, and scalability. As for routing, techniques are grouped into non-hierarchical and hierarchical routing, and this is one of the least explored aspects of ICNs. Caching, as the third core concept, uses only local information to determine what contents to cache, unlike traditional networks which have centralized management.

There are many ICN architectures that all provide different ideas toward what the future network should look like, such as CCN, NDN, MobilityFirst, PURSUIT, SAIL, COMET, CONVERGENCE, ANR Connect, GreenICN, ICE-AR, and ICN-2020. The ones that were primarily discussed in this paper were DONA, NDN, PURSUIT, and SUIT. DONA was the first complete ICN architecture proposed in 2007 and set a baseline for the following architectures. NDN was a continuation of CCN and used a hierarchical naming scheme, which is unlike DONA's use of flat naming. EU funded projects have also taken off, with one being PURSUIT. PURSUIT consists of RNs in a RENE, with TMs and FNs. The last one discussed in great detail was SAIL, which is an architecture that combines aspects from both the NDN and PURSUIT architectures.

ICNs still have many improvements that still need to be explored and implemented, which includes areas of naming, routing, and caching. It is a big challenge to satisfy all the different parameters that can determine the efficiency of the naming scheme, so this becomes a balance game. Not only does it have to work with certain architectures, but also other components of the network such as routing and caching. Routing and caching will bring another set of challenges, where there is a tradeoff between efficiency and complexity.

6. References

1. [Ahlgren08] B. Ahlgren, M. D'Ambrosio, I. Marsh, C. Dannewitz, B. Ohlman, K. Pentikousis, O. Strandberg, R. Rembarz, and V. Vercellone, "Design Considerations for a Network of Information," 2008 ACM Conference on Emerging Network Experiment and Technology, 2008, https://www.researchgate.net/publication/221325314_Design_considerations_for_a_network_of_information

2. [Ahlgren12] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, and B. Ohlman, "A survey of information-centric networking," IEEE Communications Magazine, vol. 50, no. 7, 2012, pp. 26-36, https://ieeexplore.ieee.org/document/6231276
3. [Carzaniga03] A. Carzaniga and A. L. Wolf, "Forwarding in a Content-Based Network," SIGCOMM '03 Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications, 2003, pp. 163-174, https://conferences.sigcomm.org/sigcomm/2003/papers/p163-carzaniga.pdf
4. [Cheriton00] D. R. Cheriton and M. Gritter, "TRIAD: A Scalable Deployable NAT-based Internet Architecture," Stanford University, 2000, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.77.191&rep=rep1&type=pdf
5. [CISCO19] Cisco, "Cisco Visual Networking Index: Forecast and Trends, 2017-2022," 2019, https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html
6. [CONV01] "The Convergence Project," http://www.ict-convergence.eu/ [Includes publications of the CONVERGENCE architecture as well as technical publications.]
7. [Dannewitz10] C. Dannewitz, J. Golic, B. Ohlman, and B. Ahlgren, "Secure Naming for a Network of Information," 2010 INFOCOM IEEE Conference on Computer Communications Workshops, 2010, pp. 1-6, https://ieeexplore.ieee.org/document/5466661
8. [Detti11] A. Detti, N. Melazzi, S. Salsano, and M. Pomposini, "CONET: A Content-Centric Inter-Networking Architecture," ICN '11 Proceedings of the ACM SIGCOMM workshop on Information-centric networking, 2011, pp. 50-55, https://conferences.sigcomm.org/sigcomm/2011/papers/icn/p50.pdf
9. [Fang18] C. Fang, H. Yao, Z. Wang, W. Wu, X. Jin, and F. R. Yu, "A Survey of Mobile Information-Centric Networking: Research Issues and Challenges," IEEE Communications Surveys & Tutorials, vol. 20, no. 3, 2018, pp. 2353-2371, https://ieeexplore.ieee.org/document/8303694
10. [Ganesan04] P. Ganesan, K. Gummadi, and H. Garcia-Molina, "Canon in G major: designing DHTs with hierarchical structure," 24th International Conference on Distributed Computing Systems, 2004, pp. 263-272, https://ieeexplore.ieee.org/document/1281591
11. [GREENICN01] "GreenICN," http://www.greenicn.org/ [Includes publications of the GreenICN architecture as well as technical publications.]
12. [ICEAR01] "ICN-Enabled Secure Edge Networking with Augmented Reality (ICE-AR)," http://ice-ar.named-data.net [Includes publications of the ICE-AR architecture as well as technical publications.]
13. [ICN2020] "icn2020," http://www.icn2020.org [Includes publications of the ICN-2020 architecture as well as technical publications.]
14. [Jacobson09] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard, "Networking Named Content," CoNEXT '09 Proceedings of the 5th international conference on Emerging networking experiments and technologies, 2009, pp. 1-12, https://conferences.sigcomm.org/co-next/2009/papers/Jacobson.pdf
15. [Koponen07] T. Koponen, M. Chawla, B. Chun, A. Ermolinskiy, K. H. Kim, S. Shenker, and I. Stoica, "A Data-Oriented (and beyond) Network Architecture," SIGCOMM '07 Proceedings of the 2007 conference on Applications, technologies, architectures, and

protocols for computer communications, 2007, pp. 181-192, http://ccr.sigcomm.org/online/files/fp177-koponen1.pdf

16. [Krishna19] M. Bala Krishna, "User-Centric and Information-Centric Networking and Services," CRC Press, 2019, ISBN: 9781138633322
17. [COMET01] "COMET Project," http://www.comet-project.org/ [Includes publications of the COMET architecture as well as technical publications.]
18. [SAIL01] "SAIL," http://www.sail-project.eu/ [Includes publications of the SAIL architecture as well as technical publications.]
19. [Sarkar16] D. Sarkar, N. Rakesh, and K. K. Mishra, "Content delivery networks: Insights and recent advancement," 2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), 2016, pp. 1-5, https://ieeexplore.ieee.org/document/7913113
20. [Seskar11] I. Seskar, K. Nagaraja, S. Nelson, and D. Raychaudhuri, "MobilityFirst Future Internet Architecture Project," AINTEC '11 Proceedings of the 7th Asian Internet Engineering Conference, 2011, pp. 1-3, https://www.researchgate.net/publication/220785899_MobilityFirst_Future_Internet_Architecture_Project
21. [Tarkoma09] S. Tarkoma, M. Ain, and K. Visala, "The Publish/Subscribe Internet Routing Paradigm (PSIRP): Designing the Future Internet Architecture," Towards the Future Internet, 2009, pp. 102-111, https://pdfs.semanticscholar.org/a50c/a85640ef204a89551f96ce58e9bcd4149d83.pdf
22. [Tourani17] R. Tourani, S. Misra, T. Mick, and G. Panwar, "Security, Privacy, and Access Control in Information-Centric Networking: A Survey," IEEE Communications Surveys & Tutorials, vol. 20, no. 1, 2018, pp. 566-600, https://ieeexplore.ieee.org/document/8027034
23. [Velloso13] P. B. Velloso and G. M. de Brito, "Information-Centric Networks: A New Paradigm for the Internet," John Wiley & Sons, 2013, ISBN: 9781848214491
24. [Xylomenos13] G. Xylomenos, C. Ververidis, V. Siris, N. Fotiou, C. Tsilopoulo, X. Vasilakos, K. Katsaros, and G. Polyzos, "A Survey of Information-Centric Networking Research," IEEE Communications Surveys & Tutorials, vol. 16, no. 2, 2014, pp. 1024-1049, https://ieeexplore.ieee.org/document/6563278
25. [Zhang10] L. Zhang, D. Estrin, J. Burke, V. Jacobson, J. D. Thornton, D. K. Smetters, B. Zhang, G. Tsudik, K. Claffy, D. Krioukov, D. Massey, C. Papadopoulos, T. Abdelzaher, L. Wang, P. Crowley, and E. Yeh, "Named Data Networking (NDN) Project," NDN, Technical Report NDN-0001, 2010, http://named-data.net/techreport/TR001ndn-proj.pdf
26. [Zhang14] H. Zhang, W. Quan, J. Guan, C. Xu, and F. Song, "Uniform Information with a Hybrid Naming (HN) Scheme," IETF, 2015, Retrieved November 5, 2019, https://tools.ietf.org/pdf/draft-zhang-icnrg-hn-00.pdf

7. Acronyms

4WARD	Architecture and Design for the Future Internet
AS	Autonomous System
AVP	Attribute-Value Pairs
CCN	Content-Centric Networking

CDN	Content Delivery Network
CMP	Content Mediation Plane
COMET	Content Mediator Architecture for Content-Aware Networks
CONET	Content-Centric Inter-Network
CR	Content Router
CS	Content Store
CURLING	Content-Ubiquitous Resolution and Delivery Infrastructure for Next Generation Services
DHT	Distributed Hash Tables
DNS	Domain Name System
DONA	Data-Oriented Network Architecture
FIB	Forwarding Information Base
FN	Forwarding Node
GNRS	Global Name Resolution Service
GUID	Globally Unique Identifier
ICN	Information-Centric Networking
NCS	Name Certification Service
NDN	Named Data Networking
NDO	Named Data Object
NetInf	Network of Information
NID	Network Identifier
NRS	Name Resolution Systems
NSF	National Science Foundation
PIT	Pending Interest Table
PSIRP	Publish-Subscribe Internet Routing
PURSUIT	Publish-Subscribe Internet Technology
RENE	Rendezvous Network
RH	Resolution Handler
RN	Rendezvous Node
SAIL	Scalable & Adaptive Internet Solutions
TM	Topology Manager
TRIAD	Translating Relaying Internet Architecture integrating Active Directories
URI	Uniform Resource Identifiers
VDI	Versatile Digital Item

Last Modified: December 10, 2019
This and other papers on latest advances in computer networking are available on line at
http://www.cse.wustl.edu/~jain/cse570-19/index.html

Back to Raj Jain's Home Page