Where is in a name? A Survey of Mobility in Information-Centric Networks

Gareth Tyson¹, Nishanth Sastry², Ruben Cuevas³, Ivica Rimac⁴, and Andreas Mauthe⁵

¹Queen Mary, University of London, UK,
²King’s College London, UK
³Universidad Carlos III de Madrid, Spain
⁴Bell Labs, Alcatel-Lucent, Germany,
⁵Lancaster University, UK

1. INTRODUCTION

Host mobility has been a long standing challenge in the current Internet architecture. Huge proportions of traffic are now attributed to mobile devices [1]; however, despite this prominence, mobility often remains a badly handled concept. Some have recently argued that the main reason for this lies in its choice of what to name [2]. The Internet Protocol (IP) names hosts based on their topological network location. Through this, it intrinsically binds the what (the name) to the where (the address). Consequently, a mobile host moving its physical location is often required to change its name creating numerous problems.

Observations such as this have led to a flurry of research looking at how the future Internet could be re-designed. A prominent example is that of information-centric networks (ICNs) [2][3][4]. ICNs propose a key paradigm shift, which involves replacing the Internet’s existing host-based naming scheme with an information-based one instead. This article therefore chooses to follow Shakespeare’s advice and ask “what’s in a name?”, rather than IP’s approach of “where is in a name?”.

Through this principle, an ICN becomes an infrastructure that revolves around the provision of uniquely identified content¹ to consumers, rather than the routing of data between device pairs. By removing the use of host-centric naming, it is therefore hoped that it will

¹We will use the terms information and content interchangeably.

be possible to seamlessly change a host’s physical and topological location without needing to perform the types of complex network management that host-centric networks require (e.g. creating forwarding between home and foreign addresses [5]).

In this article we aim to explore and review these concepts and ideas. We first explore what an ICN is, before investigating some of the key benefits of designing a network around the concept of information. From this, we then present some prominent ICN proposals before using these to identify important remaining challenges.

2. DEFINING AN ICN

In essence, an ICN is a network that has the primary purpose of distributing information. As such, it exposes a content request style abstraction unlike the existing Socket API. This is because a host-centric network (e.g. the Internet) is designed to route packets from a source to a destination, whilst an information-centric network is designed to deliver information from a provider to a consumer.

Its roots lie in previous attempts to build infrastructures centred on the dissemination of information. Of most note are publish/subscribe mechanisms [6], as well as peer-to-peer content delivery systems [7]. Both use overlay architectures to allow publishers to make information (e.g. data, files etc.) available to consumers. Importantly, however, these various systems are disparate with applications typically utilising specific infrastructures and protocols for their own needs (c.f. [8]). In contrast, an ICN attempts to underpin these applications through the ubiquitous support of information dissemination as an explicit network layer concept, rather than something simply built over it. Consequently, ICNs introduce new network components that unify such things as information naming, routing, security and management within a single architecture. Through this, a number of key differences between traditional host-centric networking can be identified:

— *Naming*: Host-centric networks utilise names that identify a host by its topological position. In contrast ICNs name unique items of content [9], which could exist in many places throughout the network.

— *Routing*: Host-centric networks route between hosts using pairs of topological identifiers (e.g. IP addresses). In contrast, ICNs route (or bind) between points of consumption and ‘optimal’ content sources.
— **Security:** Host-centric networks attempt to secure communication channels between hosts. In contrast, ICNs attempt to secure the integrity of individual content objects, regardless of their delivery mechanism.

— **API:** Host-centric networks expose APIs that allow data to be sent to a given location. In contrast, ICNs expose APIs that allow content to be published and consumed.

The rest of this article now explores the benefits of the above differences, specifically from the viewpoint of improving node mobility.

### 3. WHAT ARE THE BENEFITS OF ICN FOR MOBILITY?

The proposed improvement in mobility support is achieved by re-focussing network routing on content objects, rather than hosts. Consequently, in an ICN, changes in a node’s physical location do not necessarily need changes in its related network information (e.g. routing state). This high-level concept therefore opens up many potential benefits. This section looks at some of the possible advantages that could be gained if the theoretical principles of ICN were realised.

#### 3.1. Host Multihoming

A long standing challenge in host-centric networks is allowing mobile hosts to exploit multiple network interfaces (e.g. Bluetooth, UMTS, WiFi etc.). This is because typically most protocols rely on establishing individual connections using each host’s address. However, because an address is bound to a specific network interface, it is difficult to easily switch between them. For example, a HTTP GET request is always received over a single TCP connection from a single source address. Consequently, during mobile hand-offs, it is difficult to exploit multiple potential network interfaces that might be available when using HTTP.

In contrast, the concept of an ICN detaches itself from host-to-host connections. Instead, communications within an ICN are typically based around a request/reply model. As such, requests can easily be multiplexed over multiple interfaces. This means that applications running on a multihomed ICN node could seamlessly exploit these different interfaces without needing to understand which interface has actually been used.

3.2. Network Address Consistency

Currently, many mobility mechanisms attempt to maintain consistency in the node’s network address. This is vital for many applications that may utilise a node’s IP address for long-term usage. A typical example is BitTorrent, which will see a node’s IP address being registered with a tracker for future discovery. Mobile IP [5], for instance, introduces the concept of a Home Agent to allow hosts to change their physical address, whilst still maintaining a constant public address. This, however, creates undesirable overheads due to the need to tunnel data through this Home Agent. Unfortunately, the alternative requires greater intelligence in applications to make them aware of mobility, thereby allowing them to update their location information.

In contrast, the concept of an ICN does not force applications to take on host-centric information. Instead, it detaches the application from such concerns. This allows the application to abstractly publish or consume content, without the need to store (or even know) its own network-layer address. In essence, it promotes content, which is already an explicit application-layer element, to an explicit network-layer entity as well, thereby requiring the application to only maintain knowledge that does not deviate from its own traditional knowledge base.

3.3. Removal of Connection-Oriented Sessions

A key problem with mobility in host-centric networks is their frequent dependency on connection-oriented protocols (e.g. TCP). Thus, mobility can often require the re-establishment of these connection-oriented sessions so that both parties are aware of the up-to-date network addresses, as well as any pertinent parameters. Generally, TCP sessions are used in host-centric networks to establish reliability parameters (e.g. sequence numbers) and configure flow/congestion control (e.g. window size). This is necessary because the network stack does not have an explicit understanding of the data it is sending/receiving, therefore requiring bi-lateral cooperation to ensure that a receiver receives the right data at an appropriate speed.

In contrast, in an ICN, communications are made explicit within the network stack: when a node sends a request for an object, it can understand if that request has been satisfied. As such, the communications model becomes receiver-driven, without the need for cooperation.
from the sender to achieve in-order reliability. Through this, it also becomes possible to perform flow/congestion control by simply altering the frequency of requests. Therefore, sessions established between specific parties become less necessary.

3.4. Scoping of Content and Location

Currently, consumers are generally identified by their location (IP address). Often, however, this is incorrectly used for scoping purposes, i.e. incorrect information is interpreted from the address. For instance, the BBC iPlayer service can only be accessed from UK IP addresses; consequently, this makes mobility difficult for legitimate UK residents who may temporarily utilise connectivity abroad. A similar problem emerges in CDNs when attempting to utilise IP addresses for selecting optimal content replicas. This is because (at request time) the CDN will utilise a node’s location to resolve an optimal source, even though the node may later change its location.

In contrast, an ICN makes an explicit separation between the what (the user or content) and the where (their location). Thus, a node’s location can seamlessly change whilst still maintaining a consistent name (and profile) for the user. Through this, it would not be necessary to (incorrectly) interpret things from changing location-based addresses; instead, such information could be encapsulated within separate node descriptions that the network could then exploit (e.g. for access control).

3.5. Resilience through Replication

Information exchange in a host-centric network is usually based on some concept of location (e.g. a URL). As such, if the host identified in the URL fails or, alternatively, if any of the intermediate routers fail, the content will become unavailable (this is particularly prevalent in MANETs [10] and DTNs [11]).

In contrast, an ICN does not bind content to specific locations through the use of host identifiers; instead, content is the key addressable entity. This allows content to be stored anywhere, potentially allowing local copies to be retrieved. On the one hand, this can improve performance [12]. However, beyond this, the effects of network failures can also be mitigated [13]. This is because ICN caching can increase the number of potential end points for each request, thereby adding redundancy.
4. INFORMATION-CENTRIC PROPOSALS

This section briefly reviews some prominent ICNs, alongside their approaches to handling mobility.

4.1. NDN

NDN [2] is a prominent design (also known as CCN and CCNx); Figure 1 provides an overview of its operation. Content naming is based on a flexible hierarchical structure, allowing a variety of namespaces. In NDN, a content request is issued by sending an Interest packet, which is routed through the network to an instance of the content. Routing is performed using similar mechanisms to current IP infrastructure, utilising longest prefix matching. Therefore, to maintain scalability, the naming hierarchy is exploited to aggregate address space together in routing tables. Thus, in NDN each request is only resolved to a specific location during the final stages of the routing process (i.e. at the last hop). Following this, if available, the source responds with a Data packet, which follows the reverse path back to the requester using ‘breadcrumbs’ left in a Pending Interest Table on each router (the Data packets are also cached on each router). Requests are therefore performed on packet-sized objects.

Consumer mobility in NDN is intrinsic due to its consumer-driven nature. When a consumer re-locates, it can re-issue any previously sent Interest packets that have not been satisfied yet. This can occur seamlessly because there is no need to perform any new registrations etc. (although re-sending Interests obviously has overheads). Through this, it has been shown that NDN can still handle up to 97% of requests even during high mobility [14]. Provider mobility, however, is more challenging as, practically speaking, there is no separation between content identifier and routing locator. As such, to ensure route aggregation, the content item’s naming hierarchy must also reflect the underlying topology that it is routed over. Moving individual content items to different locations could therefore undermine this aggregation and create a state explosion in the core of the network. Consequently, it is better for domains of content objects to move as one, rather than having individual items of content move independently.
DONA [15] introduces ICN in the form of a replacement (or supplement) to DNS. Content names are of the form \(P:L\), where \(P\) is the cryptographic hash of the publisher’s public key and \(L\) is a label that identifies the content. DONA requires each domain to deploy servers called Resolution Handlers (RH) that index content stored by authorised storage points. RHs are then structured into a tree topology that represents the BGP topology of the network, as shown in Figure 2. Lookups are performed by querying a consumer’s local RH; if no reference is found, the query is forwarded up the tree until a source is discovered. The request is then forwarded to the source and an out-of-band delivery is established by the source. Importantly, however, due to the overhead of routing each request, DONA is unlikely to use packet-sized objects like NDN does (to minimise load).

DONA handles consumer mobility by changing a host’s RH to that of the new network. If necessary, any existing requests can then simply be re-issued to the new RH to locate the new optimal source. Unlike some other designs, however, DONA relies on out-of-band deliveries of content; although not stipulated, this would likely take place over TCP/IP. Consequently, unlike NDN, this would require session re-establishment after consumer relocation (either to re-establish the same connection or one with a newly selected source), thereby complicating the process. Provider mobility is also supported by allowing nodes to re-publish their content with the new network’s RH. Clearly, however, in this situation,
any active transfers would then need to be re-established by the consumer or, alternatively, continued using a mechanism like Mobile IP.

4.3. NetInf

NetInf primarily relies on a Name Resolution (NR) service. Using the NR services, providers publish Named Data Objects (NDOs) alongside their locators (termed routing hints) for later discovery by consumers or intermediate routers forwarding requests [16]. The resolution process therefore simply maps each NDO’s self-certifying identifier [17] to a set of locators. The NR service is underpinned by a Multi-level DHT (MDHT) [18], allowing global content lookups, whilst also supporting local resolution. This process is shown in Figure 3 with recursive querying of the MDHT. Requesters therefore perform lookups, which are responded to with either a list of potential sources or a selected optimal source. Content can then be delivered from a source using any supported delivery protocol, including those which allow in-router caching on the intermediate path (e.g. [19]).

Consumer mobility in NetInf is achieved through its indirection between identifiers and locators. The exact details of this vary based on the chosen locator selector mode [18]. In the requester-controlled mode, a requester is provided with a list of potential sources, thereby allowing a node to select a new optimal source following re-location. In contrast, the MDHT-controlled mode results in a consumer only receiving a single source on each request, mandating a re-located node to contact the NR service again. Regardless of this, both modes should enable mobility, assuming fast lookups. Provider mobility is more challenging as it

Fig. 3. Overview of NetInf

requires the NR service to be updated and all consumers to re-bind to the new location; however, it is claimed that updates can be scalably handled. It is important to note, however, that this would only maintain content availability for new requests — existing requests would either need to be re-sent or continued though a mechanism similar to Mobile IP (as with DONA).

4.4. PURSUIT

PURSUIT [3] proposes the use of a publish/subscribe abstraction, as opposed to the synchronous get used by most other approaches. Within PURSUIT, significant focus is given to the decomposition of network functionality into three key components: Rendezvous, Topology Management and Forwarding. Each of these could be potentially implemented in different ways, however, here we focus on their current realisation for global networking. When providers wish to publish content, they register it with the Rendezvous System using both a Scope and Rendezvous Identifier (SI and RI). These are flat identifiers that are interconnected by a tree structure, in which SIs are inner nodes that aggregate RIs together (as leaf nodes). The Rendezvous System is therefore a lookup service (e.g. a DHT) that can map an identifier to a data source, as shown in Figure 4. Once a source is discovered, the Topology Manager is used to construct a path to the source, which then results in a Forwarding Identifier (FI) being generated. Within the prototype, the FI is a bloom filter, which encodes the hops that any data must take through the network to reach the subscriber from the provider, i.e. source routing (c.f. LIPSIN [20]).
Consumer mobility in PURSUIT is relatively straightforward to achieve. When a consumer re-locates, it re-subscribes to the content being accessed. This results in a new FI being computed for the host’s new location. Clearly, the efficiency of consumer mobility is therefore dependent on the speed at which new FIs can be generated and mapped to RI/SIs. To alleviate this, the architecture proposes the use of explicit caches that providers can continue to stream to whilst consumers are switching between access points [21]. It is claimed that PURSUIT can lead to 50% less packet loss during mobility compared with Mobile IPv6 [22]. Provider mobility would have a higher overhead as it would require updating information in the Rendezvous System. More importantly, it would also invalidate the existing FIs a provider was using. Consequently, new routes would need to be computed for all subscribers. The speed of this process would therefore largely define the hand-off delay; it is important to note, however, that re-location within the same domain would allow the majority of the existing path to be reused, thereby increasing speed.

4.5. Juno

Juno [4] proposes the placement of information-centric functionality in the middleware layer, shown in Figure 5. Content is based on flat self-certifying identifiers that are indexed on a DHT called the Juno Content Discovery Service (JCDS). Content identifiers are therefore resolved rather than routed to. Unlike other designs, however, Juno focusses on achieving backwards-compatibility by performing software re-configuration to interoperate with any sources that might offer the content, regardless of their delivery protocols. To achieve this, Juno attempts to discover as many content sources as possible by also probing third party indexing services such as eMule. Once a set of sources have been discovered, Juno’s Delivery Framework retrieves the content by utilising dynamically attachable protocol plug-ins that
each has the capability to interact with a given source/protocol. For instance, if a HTTP source were located, a HTTP plug-in would be dynamically attached to retrieve the content. Importantly, Juno attempts to intelligently re-configure between the use of different sources based on the higher level needs of the application (e.g. performance, resilience, monetary cost etc.)

Consumer mobility is easily achieved in Juno by simply re-selecting sources after host re-location. This can be done locally as Juno keeps a full list of sources from the resolution process. The hand-off delay, however, will be defined by the bootstrap time of the delivery protocols being used; for example, connecting to a BitTorrent swarm will take longer than establishing a HTTP connection. Provider mobility is similarly possible by simply updating the JCDS; like NetInf, the performance of this depends on the DHT.

5. RESEARCH CHALLENGES

Clearly, an increasingly large research effort is being invested in ICN. However, a number of key challenges remain, particularly in the mobility domain. This section now explores the most prominent of these.

5.1. Provider Mobility

Broadly speaking, consumer mobility is a well handled phenomenon due to the consumer-driven nature of most ICN designs. However, a larger challenge is maintaining routing consistency during provider mobility. This is because whenever a provider re-locates, it is clearly necessary to update (potentially global) locator information. This is heavily exac-
erbated by the obvious increase in the number of content objects when compared to hosts (an ICN must be able to deal with at least $10^{12}$ objects [23]). The effects of this can be mitigated via caching and replication but less frequently requested content is still likely to suffer if high speed provider hand-offs cannot be achieved.

The precise focus of this challenge varies with the different naming and discovery techniques employed. NDN, for instance, uses hierarchical naming and route aggregation to improve scalability. However, because names are also used for routing locators, they must efficiently map to topological locations. This creates significant challenges when re-locating providers to different topological positions because it clearly undermines the hierarchy of the address space; in fact, using any content that is cached off-path introduces a similar challenge. Unfortunately, line-speed switching relies heavily on this aggregation, meaning that provider mobility will introduce significant scalability challenges. Regardless of this, clearly, routing information will need to be disseminated during provider mobility leading to convergence delays.

Challenges also arise in resolution approaches such as NetInf or Juno. This is because any provider mobility must be reported to the resolution service; clearly, high levels of mobility could result in phenomenal loads. For instance, in [18], the authors discussed the handling of 1% of churn in registrations, however, mobility could increase this greatly. Similarly, when this occurs, potentially all related consumers will need to be notified of changes to allow them to re-bind to the new source via the resolution service (opposed to NDN, which does not require this). This is particularly problematic if subsets of larger objects cannot be independently requested, leading to the need to request the whole object again. As previously mentioned, to address this, approaches similar to Mobile IP have begun to emerge, allowing messages to be forwarded to a mobile provider’s current location [24] (thereby not requiring resolution/routing updates in the core). Clearly, these solutions re-introduce some of the problems (e.g. tunneling overheads) that ICNs wished to move away from. A key point to observe, however, is that during provider mobility, it becomes possible for consumers to re-bind to alternate sources (e.g. a cache), thereby mitigating hand-off delays. Despite this, a prominent remaining challenge is to design mechanisms that can elegantly allow provider mobility without the need for any complicated or high overhead procedures.
5.2. Response Routing

Due to the (attempted) removal of location from the concept of ICN, many approaches utilise pre-defined pairwise hop-by-hop knowledge (e.g. breadcrumbs) to ensure that data can find its way back to consumers without needing host-centric routing. Unfortunately, however, this can be challenging in a mobile network because paths could change frequently. In NDN, for instance, Data packets always follow the reverse paths of their equivalent Interest packets; thus, if a host changes its location, the response path will change, leaving a window of potentially many data packets being routed to an out-of-date location (in TCP, for instance, window scaling allows windows of approximately a gigabyte [25]). Interestingly, some protocols designed for handling this network dynamism (e.g. [26]) still rely on reverse path routing. Alternatively, other solutions avoid multi-hop routing and simply rely on one-to-one opportunistic connections [27], thereby losing access to content that is multiple hops away. A related situation also arises in PURSUIT through its use of per-hop source routing. If, for instance, the computed route changes due to mobility, this per-hop knowledge will become out-dated. Encoding redundant virtual links [20] can mitigate this but even these could become invalid due to mobility. Clearly, if ICNs are to be deployed in mobile environments, handling these physical path changes is an extremely important research issue to address. This will therefore likely involve creating a compromise between both host-centric and information-centric routing.

5.3. Discovering Local Cached Content

One of the key benefits of an ICN is the ability to deploy ubiquitous caching. This, however, can create significant challenges in mobile environments, particularly MANETs and DTNs, due to the potential cost of managing cached replicas.

The specifics of this challenge will vary with the particular approach employed. Approaches such as NDN would likely suffer heavily due to the increased overhead of maintaining routing information for cached content sources. In fact, to mitigate this, some recent mobile routing algorithms (e.g. CHANET [28]) do not require mobile nodes to advertise their cached content, instead only allowing opportunistic on-path caching. In contrast, approaches such as DONA and Juno, which utilise resolution services, would also suffer due to the need for resolution updates for every cached instance (this is analogous to extremely
high levels of provider mobility). Further, when considering ad hoc environments, things could be even worse if mobility meant that these resolution services somehow became inaccessible. Interestingly, promising information-centric MANET routing protocols have begun to emerge, addressing some problems (e.g. LFBL [29], Slinky [30]). However, these are still yet to be extensively tested in terms of performance, scalability etc.

A prominent research challenge that therefore remains is to build and evaluate naming, resolution and routing schemes that can handle this type of unpredictable ad hoc re-location of content. A particularly important challenge is achieving this for unpopular content that doesn’t benefit from caching (i.e. accessed only once). For instance, when dealing with smaller MANETs (e.g. <300), Varvello et al. [13] found that the performance benefits of using more sophisticated structured routing protocols (e.g. GHT [31]) were dwarfed by their overheads due to the presence of unpopular content.

5.4. Real-time Hand-Off Delays

Mobility during time insensitive network interactions is a relatively easy issue to handle in many ways. This is because there is no real constraint on the hand-off delay. In contrast, mobility during real-time communications (e.g. video conferencing) is far more difficult because hand-offs must be in the order of milliseconds.

Typically, the main benefit of using an ICN for mobility is that cached or replicated copies of the content could potentially mitigate any hand-off delays. However, many real-time communications have little potential for caching (e.g. a voice call). Further, as some real-time communications are multi-directional, all parties behave as both consumers and providers, thereby increasing the network load of mobility. Practically speaking, systems that use content resolution rather than routing would likely perform better because a resolution service would only requires a small number of centralised updates. However, the exact performance is yet to be understood. A prominent remaining challenge is therefore ensuring that real-time multimedia can be fully supported in an ICN. On the one hand, this refers to handling mobility, but it also extends to many other issues including QoS and QoE.
5.5. Privacy and Security

Privacy and security in open mobile systems has been a long-term challenge, with many possible attacks. Unsurprisingly, a key research challenge is therefore handling these types of concerns in an ICN. In principle, ICNs primarily focus on securing the content itself through its unique name, i.e. guaranteeing a content item is what it claims to be. This approach, however, can obviously introduce privacy challenges because it requires nodes to expose their interests to the network. Thus, if third parties can map identifiers to content items (which often will be possible), a user’s privacy could be heavily undermined.

Beyond this, alternate security issues relating to such things as routing are yet to be adequately explored. In theory, any node can publish the ability to serve an item of content, thereby empowering malicious nodes to manipulate routing. This can be an issue in traditional fixed infrastructure; however, it is particularly prevalent in networks such as MANETs, which have extremely open routing policies (e.g. black hole routing).

5.6. Practical Challenges

A further challenge that perhaps exceeds all others is the question of practical deployment. Clearly, the discussed benefits can only be gained if a node connects to (and moves between) domains that support ICN. Unfortunately, however, in practice, it is likely that any near-future ICN deployments will be incremental overlay structures, making it impossible to quantify the effectiveness of mobility support without concrete details of their configuration. The existence of islands of ICNs connected via tunnels, for instance, could severely damage mobility support if nodes were required to ‘dial-in’ via VPN-like services after every hand-off. Even low-delay access services within the domain of the node might create unacceptable delays if complex authentication were required. Interestingly, this problem becomes even more challenging when considering the diversity of ICN proposals, likely leading to the need for complicated inter-networking technologies.

Despite this, clearly, the long-term goal of ICN would be to move towards a native deployment (e.g. dual stack). This would, of course, be an incremental process in which the benefits increase with each new domain’s uptake (c.f. IPv6 deployment [32]). However, an intelligent deployment strategy could help mitigate any potential problems. Specifically, using the lessons learnt from previous overlay technologies, many problems could be avoided.
by integrating underlay knowledge (e.g. locality awareness [33]). This still does, however, leave open practical questions such as how nodes might discover and connect to ICN-enabled domains, including various auto-configuration challenges.

6. CONCLUSIONS

This article has explored the world of ICN, looking at how a shift from host-centric to information-centric design principles could support greater mobility in the future Internet. Clearly, this is a hugely important topic, as we observe more and more traffic being generated by mobile hosts [1]. However, despite the clear potential of ICN, a number of challenges remain. Of particular importance is the ability to handle the scalability challenges of increasing numbers of content items (in the form of router entries) and providers (in the form of router caches). Whereas this is a significant problem in fixed infrastructure, it is even more challenging in mobile environments.

It is important, however, to note that these are not necessarily weaknesses in ICN. Instead, they are exciting topics deserving future attention. Such promising research has already begun to develop, however, it is evident that the diversity of mobile content access means that any ‘one-size-fits-all’ approach will fall short. Consequently, we believe the key future research challenge is building flexible general purpose architectures that can handle all the situations discussed within this article.

REFERENCES


