A Thing

Peoples
Places
Things

ABSTRACT

In recent years, Information-centric Networking (ICN) has received much attention from both academic and industry participants. ICN offers a data-centric means of inter-networking that is radically different from today's host-based IP networks. Security and privacy issues in ICN have become increasingly important as ICN technology gradually matures and nears real-world deployment. As is well known, in today's Internet, security and privacy features were originally not present and had to be incrementally and individually retrofitted (with varying success) over the last 35 years. In contrast, since ICN-based architectures (e.g., NDN, CCNx, etc.) are still evolving, it is both timely and important to explore ICN security and privacy issues as well as devise and assess possible mitigation techniques.

This report documents the program and outcomes of the Dagstuhl Seminar 16251 "Information-centric Networking and Security." The goal was to bring together researchers to discuss and address security and privacy issues particular to ICN-based architectures. Attendees represented diverse areas of expertise, including: networking, security, privacy, software engineering, and formal methods. Through presentations and focused working groups, attendees identified and discussed issues relevant to security and privacy, and charted paths for their mitigation.

1. INTRODUCTION

Dagstuhl seminar 16251 "Information-centric Networking and Security" was a short workshop held June 19-21, 2016. The goal was to bring together researchers with different areas of expertise relevant to ICN to discuss security and privacy issues particular to ICN-based architectures. These problems have become increasingly important as ICN technology gradually matures and nears real-world deployment.

Threat models are distinct from IP. Differentiating factors between the two include new application design patterns, trust models and management, as well as a strong emphasis on object-based, instead of channel-based, security. Therefore, it is both timely and important to explore ICN security and privacy issues as well as devise and assess possible mitigation techniques. This was the general purpose of the Dagstuhl seminar. To that end, the attendees focused on the following issues:

- What are the relevant threat models with which ICN must be concerned? How are they different from those in IP-based networks?
- To what extent is trust management a solved problem in ICN? Have we adequately identified the core elements of a trust model, e.g., with NDN trust schemas?
- How practical and realistic is object-based security when framed in the context of accepted privacy measures used in IP-based networks?
- Are there new types of cryptographic schemes or primitives ICN architectures should be using or following that will enable (a) more efficient or secure packet processing or (b) an improved security architecture?

The seminar answered (entirely or partially) some of these questions and fueled discussions for others. To begin, all participants briefly introduced themselves. This was followed by several talks on various topics, ranging from trust management and identity to privacy and anonymity. Subsequently, the attendees split into working groups to focus more intensely on specific topics. Working group topics included routing on encrypted names, ICN and IoT, non-privacy-centric aspects of ICN security, as well as trust and identity in ICN. Once the working group sessions were over, a representative from each presented outcomes to all attendees. (These are documented in the remainder of this report.) The major takeaways from the seminar were as follows.

First, the ICN community still does not have a clear answer for how to handle namespace and identity management. While trust management in ICN can be distributed and function without a global PKI, it seems difficult to break away from this model for namespace management and arbitration. This has strong implications on how names are propagated in the routing fabric. Can any producer application advertise any name, anywhere in the network? If not, how can name prefix advertisements be constrained or limited?

Second, given that ICN focuses on object security, the need for and use of transport protocols that provide forward secrecy should be deferred to higher layers. Attendees found

that while most ICN-based architectures do not preclude forward secrecy, it should not be a requirement at the network layer.

Third, there is still deep uncertainty about whether ICN should embrace a content locator and identifier split. Names in architectures such as NDN and CCN serve as both a locator and identifier of data, though there are extensions that permit explicit locators (e.g., through the use of NDN LINK objects). This distinction is necessary under the common understanding that routing should concern itself with topological names. Finding data through non-topological names should not be in the data plane as part of the global routing space. However, if we revert to a distinction between topological locators and identifiers, then features unique to ICN become much more limited. One facet that is certainly unique to ICN is how software is written. Specifically, we have the opportunity to move beyond the mental model of a fixed address space and re-design existing network stacks and APIs.

Fourth, privacy seems difficult to achieve without major architectural changes to ICN-based systems. In particular, since data names reveal a great deal of information to the passive eavesdropper, privacy demands that names and payloads have no correlation. However, achieving this seems infeasible without the presence of an upper-layer service akin to one that would resolve non-topological identifiers to topological names.

Lastly, there are no compelling reasons to apply esoteric (and often untested) cryptographic techniques in ICN, at least at the network layer. Computationally bounded and "boring" cryptographic primitives, such as digital signatures, hash functions, etc., should be the extent of per-packet cryptographic processing done by routers. Anything more would become fodder for Denial-of-Service attacks that could render the entire infrastructure ineffective. However, architecture designs should not restrict themselves to specific algorithms. In other words, there must be flexibility in accommodating multiple (and evolving) cryptographic primitives. This could be useful if, for example, post-quantum digital signature schemes become necessary for the longevity of content authenticators.

2. SUMMARY OF PLENARY TALKS

The main seminar material was driven by plenary talks on a variety of topics ranging, such as trust management, namespace management, privacy, and anonymity. These talks began with a discussion of threat models and their importance in ICN. In particular, recall that ICNs attempt to diverge from IP with respect to the central abstraction of hosts and point-to-point communication between them. The ICN emphasis on named data and object security instead channel security is one clear differentiating factor contributing to this parity. To quantify the degree by which secret is improved (or worsened), threat models are needed. In general, they must capture a particular design challenge in ICN, such as infrastructure protection, user-friendly key distribution and trust management and enforcement, and content protection and access control. And given the wide gap between IP and ICN, there is a great need for common threat models to use in the design phrase.

One particular threat revolves around consumer anonymity, which was also the topic of discussion at the seminar. The subject(s) and contents of a packet are not the only facets that should be considered with respect to privacy. Origin and destination details (e.g., geographical location or position within a network topology), as well as identity information (e.g., consumer identifying information), can sometimes harm network users. As shown by [], ICN caching and interest-collapse mechanisms make ICN itself inherently vulnerable to the possibility for an adversary to locate consumers. Moreover, an approach similar to the one to violate consumer (location) privacy, might be used also to detect eavesdropper. Therefore, the threat model must consider this vulnerability and adversaries capable of exploiting it successfully.

Another design challenge unique to ICN is how to develop scalable object-based access control mechanisms. A variety of encryption techniques have been used in the past to protect access to confidential network data [**?**]. Many design approaches, particularly in CCN and NDN, exist well above the network layer. In contrast, publish-subscribe ICNs such as ENCODERS [] integrate access control into the network. It uses multi-authority attribute-based encryption to allow content access to be scoped to selected nodes in the system. Since the system is completely decentralized, peers serve as brokers that match content from publishers with interests expressed by subscribers. In order to perform such a match, an intermediate node must be authorized to see both the relevant content tags and subscriber interests. One talk at the seminar focused on the observation that access control policies applied to the metadata (content tags and subscriber interests) effectively create reachability constraints that are independent from the one defined by the routing protocols. Consequently, this security-routing interaction must be treated carefully during policy definition.

3. PARALLEL GROUP WORK

3.1 ICN and IoT

The Internet of Things (IoT) is connecting billions of smart devices (e.g., sensors) and is growing very fast. We expect more than 1 million networked "things" per square kilometer in 2030. In this group, we tried to explore how much data density we can afford and how we communicate with the "things" (say, directly to the sensors, or indirectly through the cloud or gateways). We discussed the potential of implementing ICN for the IoT. For instance, the ICN routers connecting to sensors can cache sensor data to improve the performance of data dissemination. Users can obtain data directly from the sensors and the ICN routers, without going

through the cloud. This raised several security concerns:

- How are sensors securely configured at the time of initialization?
- How can software updates be performed securely?
- How can we handle ICN mobility for IoT? For example, each sensor may have a unique publisher identity. How do mobility and naming affect the scalability of routing?

We also discussed the potentials and concerns of caching data at the sensors. First, sensors are resource limited devices. However, memory resources may increase and the price will go down. Second, it is advantageous to retrieve data directly from the sensors in some use cases (e.g., to control home lighting without going through the cloud). Third, when using cryptography on sensors, the encryption time could be long and cause a delay in data retrieval. Lastly, sensors have to be always on to listen to the interests, which may consume a lot of energy. Scheduling or adaptive duty cycles might be considered to mitigate this.

Based on these observations, our summary and future plan is as follows. First, we plan to come up with sample IoT use cases, which allow us to understand more about the security needs and communication patterns in ICN for IoT. Second, we will aim at answering the following questions: How does IoT benefit from ICN? How does one securely configure and bootstrap sensors? And what is the cost of providing security for IoT data?

3.2 Namespace and Identity Management

In an ideal ICN architecture, applications should be able to express their trust preferences or policies and let the "middleware" enforce them. This raises two important questions: (1) what is the minimum set of policies that can be factored out of all trust models, and (2) what is the middleware that does this enforcement? The trust schemas pioneered by the NDN architecture [**?**] are exemplary of the common rules that can be used to express most trust models. Among other things, they specify what keys are allowed to sign what data. Since both keys and data are named resources in NDN and other ICN architectures, this means that a schema allows for arbitrary hierarchical trust models. It remains to be seen if other non-hierarchical trust models will be applicable to ICN.

To address the second question, we had to agree upon what the network is responsible for enforcing. (This is discussed at length in [**?**].) First and foremost, network layer "trust enforcement" should not prohibit or prevent other application-layer trust models. This means that the network functionality must be simpler than that which is supported by the middleware. Currently, this is comprised of (at most) digital signature and content object hash verification. Behaviors such as certificate chain resolution or key retrieval should not be part of the core network functionality. This means that in the general "network," routers are only responsible for single signature or hash verification. All other network nodes (e.g., consumers and producers) contain the middleware responsible for handling the remaining trust enforcement steps.

After addressing network trust, we turned to identity and discussed the following major questions:

- How are names registered and managed in ICN?
- How can names possibly be location agnostic (without aids such as the NDN LINK)? Is there always a discovery or locator service?

Namespace ownership is intrinsically tied to identity. Thus, namespace advertisements under different namespaces or in different networks must be authenticated with respect to the claimed owner's identity. In this context, an identity is a public and private key pair. We struggled with issues about namespace scale and the practicality of a global namespace. Questions such as, "how do NATs work in a global namespace?," drove the discussion. No consensus or common understanding about how namespaces and identities should be managed was reached.

3.3 Routing on Encrypted Names

This group started with a discussion about routing on encrypted names but ended up being an exploration of name privacy and the necessary conditions for it to be possible in different ICN architectures. In this context, we defined name privacy to be the property where a so-called "network name," i.e., the name encoded in a packet, has no correlation or connection with the corresponding content object. Specifically, name privacy means that a network name reveals nothing about the data inside the content object. Ideally, names should reveal no more than what is currently revealed by an IP address and port. After settling on this definition, we laid out our assumptions to use when discussing name privacy, including:

- There is no name discovery process or search engine.
- Content may be requested by an identifier (ID) such as its cryptographic hash digest. Moreover, revealing the content ID does not compromise privacy.
- Consumers know the public key of the producer with which they want to communicate.
- Network names have an implicit routable prefix and application-specific suffix. By default, consumers do not know the locator and identifier split in a name.
- Requests may specify the ID of (1) a signature verification key or (2) the expected content.

To begin, there are fundamentally two ways to request content: (1) with and (2) without a content ID. In case (1), a request name needs to only contain a routable prefix that will

move the request to some cache or producer which can return the corresponding content. These locators can be completely separate from the desired content and, therefore, this approach can satisfy our name privacy goal. However, without implicit knowledge about the locator for some desired content, an upper-layer service is necessary to obtain said information.

In case (2), the application-specific suffix of a name must not reveal anything about the data. To achieve this, it must be encrypted. Name encryption introduces a number of other questions, such as how to obtain the routable prefix, what key to use for encryption, and how to "protect" the result. Assume that the routable prefix is known and that the producer public key is used for name suffix encryption. If the resultant content payload is not encrypted then one may be able to infer the name from its contents. Therefore, the content response itself must also be encrypted. This requires requests to carry a consumer-generated key that is protected in a CCA-secure envelope. Otherwise, eavesdroppers could replay requests with the same encrypted name but their own key to obtain a decrypted response.

In all cases, we concluded that to achieve name privacy then one needs some upper-layer service. Whether its role is to provide the routable prefix for a name, encrypt the response, or to separate a content ID and locator via some other means is an orthogonal discussion. Also, name privacy seems to, in most cases, invalidate the utility of shared caches, which puts it at odds with the primary feature of many ICN-based architectures. Thus, our conjecture is that name privacy is not possible natively in the network.

3.4 Locators and Identifiers

This working group discussed locators and fetching data with non-topological names (or even topological names that are cached off-path). Routing should, it seems, only concern itself with topological names or addresses. Finding data (objects) with non-topological names should not be done in the data plane. It should be done via a service.

In CCN, the service could resolve a named root manifest to then resolve locator names by hash. In NDN, it resolves the link routing hints to allow off-path interest forwarding. In TagNet, there is a distinction between Locator names and Descriptor names. Locator names have a strong binding between their name and a point of attachment. Descriptor (hash) names, on the other hand, are free-form and could be present anywhere. One resolves a tag query (of either type) to a topological locator and then does data transfer on that locator.

This lead to a discussion on locator and identifier split. Should CCN embrace this, or continue on with its mixed use of the name? For example, if there is a clear locator field and then a clear identifier tuple (name, [keyed restriction], [hash restriction]), one would get full matching expressivity with the functionality of nameless object locators. A similar approach could be done in NDN, though with a different tuple. There was no consensus on this idea, though it is worth exploring.

There was also some discussion on the benefit of ICN if one still needs to do an external name to address lookup. Why bother if one still needs a DNS-like function? One partial answer is that in the non-global routing space (i.e., data center, maybe IoT, some internal applications), one could inject all names into the internal routing protocol and realize the full benefit of application-specific names. Another argument is that it improves how one writes software to not have to deal with IP addresses and host-based networking. One could also see benefits from a re-designed network stack beyond sockets.

3.5 Security, Not Privacy

The "No Privacy" security working group sought to answer the following question: is an ICN security architecture easier to devise if the designs fundamentally make privacy hard to achieve? In particular, the group discussed:

- What ICN entities (content consumers, hosts, routers, content creators) need identities?
- What entities can simply operate with a public and private key pair but no formal name?
- Does splitting routing out as an application help?
- Do interests need to be authenticated at each router?

The group achieved a simple security model. Members of the group hope to write up the result as a short paper.

4. OPEN PROBLEMS

In this section we highlight several open problems that loomed large at the seminar. They should be considered in future research endeavors.

4.1 Privacy

What is the future of privacy for ICN? To what, or whom, is ICN privacy related? Existing architectures leak a significant amount of information by default, including: who requests information, whose information is requested, when content is requested, and other miscellaneous properties, e.g., data contents, name, size, etc. As of yet, we have not adequately addressed these privacy problems. Beyond ICN packets, ICN programmers will generate linked data, e.g., FLIC [1]. We should provide techniques, services, and recipes for programmers that make transport privacy a trivial. Before doing so, however, we must first define what is transport privacy in the context of ICN.

4.2 Forward Object Secrecy?

We revisit the question of securing the data, not the pipe. This has been the running mantra for security in ICN. Is forward secrecy possible in NDN and CCN group access control by encryption? Is the "take what you want" model of

group access control, dependent on long-lived keys, realistic for future networks? Is it desirable? Also, is there any role for perimeter security or is encryption enough? In this talk, we pose these questions and others to the group to stimulate a wider discussion.

Forward secrecy is the property that exposure of a principal's long-term secret keys does not compromise the secrecy of their previously generated ephemeral (session) keys. This is a useful property to have in the presence of eavesdropping attackers intercepting and logging traffic. It minimizes data and key compromise windows and therefore reduces the overall attack surface. However, it requires protocols and techniques for deriving ephemeral keys and then updating them regularly. The single request-response model of many ICN-based architectures does not lend itself to the establishment of forward secrecy without building a higher-layer protocol, such as CCNxKE [1], or involving more exotic cryptographic schemes. Consequently, the majority of work on ICN object security has ignored this property, which puts ICN at odds with best practice techniques used in IP-based protocols. In this talk, we seek to raise awareness of this issue and seek answers to the following important questions. First, under what conditions does transport security require forward secrecy? Second, can object encryption subsume transport security? And lastly, is forward secrecy in ICN needed?

4.3 Names and Routing

ICN names are user-generated content in FIBs. In effect, FIBs serve as a (globally) replicated name set wherein any name owner can write into the set. The complexity of this state is influenced by the fact that prefix owners can always de-aggregate and create arbitrary names, even if prefixes are restrictively assigned. However, this raises questions of resource exhaustion attacks on FIBs and general complexity attacks (e.g., hash collisions). Newer attacks try to leak information from the FIB contents to target the forwarding plane. This talk outlines the severity of these problems in hopes of discussing potential solutions.

4.4 Coping Network Services

The Internet has a history of adapting the existing law system to new business paradigms. One such paradigm is in-network processing, which, in recent years, has expanded to aid and impact routing, forwarding, packet replication, packet splitting and merging, quality of control, caching, and others. The relationship between these services and existing laws has been a continual tussle. When and how does caching affect copyright laws? When do other services violate the Secrecy of Correspondence (SoC) statute? Moreover, Deep Packet Inspection on SSL/TLS connections, while technically feasible, may violate the SoC statute and various other privacy rules. Thus, there has been a recent push for all-or-nothing secrecy, which unfortunately stifles network business opportunities. In this talk, we advocate for controllable privacy that allows secrecy preferences to be expressed in packet headers. We claim that ICN packet headers should be constructed to allow privacy and secrecy preferences to be expressed by their senders. This is one area where ICN can innovate to allow in-network processing to continue without violating existing laws.

5. OUTLOOK

XXX

6. CONCLUSION

This paper described in detail the events and outcomes of the Dagstuhl 16251 seminar on ICN security and privacy. Despite significant research over the past half decade, there are still many open problems whose solutions seem difficult if not impossible to achieve with the existing architectures. Are we too invested in the current architectures to make deep design changes to patch these problems? Is there something to be gained by sacrificing properties such as privacy in favor of features such as object security? If so, is this the right tradeoff to make today? Only future research and development will tell.