Modeling Data-Plane Power Consumption of Future Internet Architectures

Chen Chen
ETH Zurich

David Barrera
ETH Zurich

Adrian Perrig
ETH Zurich

Abstract—With current efforts to design Future Internet Architectures (FIAs), the evaluation and comparison of different proposals is an interesting research challenge. Previously, metrics such as bandwidth or latency have commonly been used to compare FIAs to IP networks. We suggest the use of *power consumption* as a metric to compare FIAs. While low power consumption is an important goal in its own right (as lower energy use translates to smaller environmental impact as well as lower operating costs), power consumption can also serve as a proxy for other metrics such as bandwidth and processor load.

Lacking power consumption statistics about either commodity FIA routers or widely deployed FIA testbeds, we propose models for power consumption of FIA routers. Based on our models, we simulate scenarios for measuring power consumption of content delivery in different FIAs. Specifically, we address two questions: 1) which of the proposed FIA candidates achieves the lowest energy footprint; and 2) which set of design choices yields a power-efficient network architecture? Although the lack of real-world data makes numerous assumptions necessary for our analysis, we explore the uncertainty of our calculations through sensitivity analysis of input parameters.

I. Introduction

The current Internet requires a considerable amount of power, consuming nearly 1% of annual electricity production worldwide [24]. Around 50GW of power is consumed by network equipment, and this number is expected to double by 2020 [40]. Increased power consumption not only implies greater monetary cost, but also exerts an expanding environmental impact such as carbon footprint [1], [36] and pollution [8]. Reversing the trend is both imperative and rewarding. In fact, a 10% reduction of global network power consumption could eliminate the need for 5 recent large nuclear reactors today [6].

The past five years have seen significant effort from the research community to design Future Internet Architectures (FIAs). The underlying goals and designs of FIAs can vary drastically. For example, Named Data Networking (NDN [25]) treats content (rather than end hosts) as the principal entity and enables efficient content distribution. Mobility First [32] treats mobile devices as first-class citizens on the network. NEBULA [13] provisions a highly-available and extensible core network interconnecting data centers. XIA [12] enables evolution of the network's underlying protocol stacks. SCION [44] enables highly-available communication. With such diverse design goals, evaluating and comparing FIAs under a common framework is a difficult task.

Although metrics such as bandwidth or latency have been used to compare the IP network to FIAs [20], and power consumption of Content Centric Network (CCN) has been evaluated [31], to our knowledge power consumption as a

metric has not yet been considered for comparison of different FIAs. Consequently, answers to key design questions such as "does fetching content directly from a remote server require less power than retrieving the same content from a nearby cache?" and "which is more power efficient: packet-carried state or routing table lookups?" remain largely unknown and yield some counter-intuitive answers as we show in this paper.

In certain cases, the power implication of some FIA design choices is straightforward. For example, some architectures require routers to conduct cryptographic operations for security reasons [13], [44], which inevitably increases router computation, and thus increases router power consumption. However, the power consumption implications of other FIA design choices are difficult to pre-determine because the design choices introduce trade-offs. For example, architectures that use Packet-Carried State (PCS) do not require routing tables, which reduces the routers' power consumption. However, using PCS requires embedding extra forwarding information in packet headers, which increases the number of bits that must be transmitted, and therefore increases the power consumed to forward the packets.

As a first step towards analyzing the power consumption of the IP network and FIAs, we focus on the power consumption of the data-plane. Since data-plane traffic consumes 83% of the total power consumed by the Internet (compared to 17% consumed by control-plane [16]), we believe our analysis covers the largest component of power consumption in IP networks and FIAs.

Three main challenges exist in analyzing power consumption of FIAs. First, modeling FIA router forwarding behavior is more complex than modeling today's IP routers which are themselves non-trivial to model. Levels of abstraction for router models vary, ranging from gate-level modeling [41], to microarchitecture-level modeling [30], to router-level modeling [24]. Choosing the correct level of abstraction is necessary to preserve the model's simplicity, yet still make it useful to highlight the differences of various FIA designs.

To address the first challenge, we present a generic FIA router model that captures the commonality of IP and FIA routers as well as the peculiarities of FIA routers (Section III). We use a hybrid method that combines two models at different levels: 1) a high-level model to characterize the power consumed by the common behaviors shared between IP routers and FIA routers; and 2) a low-level model to characterize the power consumed by the behaviors unique to FIA routers.

The second challenge is that power consumption analysis of network data-planes requires analysis of router forwarding behaviors in the context of real-world network topologies and

workloads. For example, NDN [25] requires routers to host content caches, which increase routers' power consumption. However, content caches reduce the number of network links the queries and responses traverse. Thus, the overall power consumed to process these queries may be reduced. Whether the reduced power can compensate for the power expended to operate the content cache largely depends on how routers inter-connect and the temporal locality of the workload.

To address the second challenge, we conduct a large-scale simulation of content delivery traffic across multiple autonomous systems (ASes) to analyze power consumption of different network architectures (Section IV). Particularly, we focus on using the simulation to investigate the influence of content caches and packet-carried forwarding information on power consumption.

The third challenge is to define a common comparison framework for different FIAs. The framework should provide useful clues about power-efficient FIA designs, independent of specific design choices of particular architectures (e.g., the hash function used). For our comparison framework, we generalize common FIA design choices. We evaluate the influence of packet forwarding techniques and cache placement strategies on power consumption of individual routers and on the entire network.

In this paper, we make the following contributions:

- We present the first work comparing the IP network and FIAs using power consumption as a metric. Our comparison framework allows us to identify power-efficient FIAs as well as guide designs of power-efficient FIAs.
- We propose a generic model to characterize the forwarding behaviors of FIA routers and conduct a large-scale simulation based on our router model to analyze the power consumption of network architectures.
- We evaluate the influence of two architectural design choices (related to packet forwarding and cache placement) on power consumption, and find that packet-carried state is generally more power efficient. We also find that caching, while advantageous in reducing latency, does not offer substantial reductions in power consumption.

II. BACKGROUND

In this section, we describe two architectural dimensions that enable different design choices that are common to many FIA designs: 1) forwarding technique, i.e., making forwarding decisions by routing table lookup versus packet-carried state; 2) cache placement, i.e., caching content by pervasive caching versus edge caching.

A. Forwarding technique: Routing Table Lookup versus Packet-carried State

To route a packet through the Internet, routing state can be kept either in routing tables constructed and maintained by individual routers or carried in packets themselves. In the latter case, packet headers contain information about paths that these packets traverse. We denote these two methods of making forwarding decisions as Routing Table Lookup (RTL) and Packet-Carried State (PCS), respectively.

RTLs relieve hosts from tasks such as path management and keep packet headers small. In fact, RTLs prevail in intra- and inter-domain routing protocols today. However, RTL consumes considerable amounts of power. BGP routers maintain and search routing tables containing more than 500K entries [5]. Large routing tables mandate large RAM size to store the tables. In order to match the forwarding speed with increasing link speeds, Application-Specific Integrated Circuits (ASIC) with Ternary Content-Addressable Memory (TCAM) are installed on line cards. TCAM chips are expensive and power-hungry. In fact, routing table lookups consume about 32% of the entire power consumption of IP routers [16].

Among FIAs, NDN adopts RTL to make forwarding decisions. The NDN routers use Forwarding Information Base (FIB) to route users' interests for contents. FIB conducts longest-prefix match using content names, instead of IP addresses. We expect NDN FIB to be much larger than routing tables in IP routers given the number of possible content names. Perino and Varvello [31] estimate that a NDN FIB contains up to 20M records.

In comparison, some FIA designs adopt PCS for making forwarding decisions. PCS allows end hosts to control the paths that packets traverse. PCS simplifies and speeds up packet forwarding on routers, since searching for matches in routing tables is not required. An open question is whether PCS helps reduce overall power consumption. PCS reduces routers' power consumption by removing the need for expensive routing table lookups, but PCS enlarges the size of each packet (by adding path information in packet headers), and thus requires additional power to transmit the extra bits. Moreover, FIAs like SCION and NEBULA mandate cryptographic operations when processing the state carried in packets, which also adds to the power consumed by routers.

NEBULA and SCION, adopt PCS to forward packets. We assume that NEBULA uses ICING [29] as its data plane. In ICING, packet headers contain Proof-of-Consents (PoCs), which certify the providers' consent to carry the packets, and Proof-of-Provence (PoPs), which allow upstream nodes to prove to the downstream nodes that the upstream nodes indeed transmit the packets. In SCION, packet headers carry a chain of Hop Fields (HFs). HFs carry the border routers' decisions for routing packets, but HFs are only meaningful to the routers that generate them. In both NEBULA and SCION, processing the packets requires routers to process symmetric cryptographic operations and compute Message Authentication Codes (MACs).

B. Cache Placement: Edge Caching versus Pervasive Caching

Caching content closer to the consumer to reduce network latency and bandwidth cost is a common practice today [11]. This type of content caching is usually organized as a dedicated network of content servers, each of which resides in the edge network to serve local consumers' content requests. DNS redirection is leveraged to re-direct content requests to nearby content servers. We refer to this type of content caching as *edge caching*.

The research community has proposed to install content caches directly on routers, which provides additional opportunities to further reduce latency and bandwidth overhead [25]. Upon receiving a content request, a router can immediately reply with content if the content is cached locally. If the

queried content cannot be served locally, the router can forward the request towards a different cache through some routing protocol. We refer to this type of content caching as *pervasive caching* since the content cache could exist in both core and edge networks.

Among FIAs, NDN proposes pervasive caching as one of its fundamental design principles. Each NDN router includes a *content store* which caches and serves content. Upon cache misses, NDN forwards the packets to a nearby router that may cache the content, and thus can reduce the length of the content delivery path.

Although edge caching and pervasive caching have been compared using latency and bandwidth as metrics [20], the difference in power consumption of edge caching and pervasive caching has not yet been explored. Pervasive caching reduces the average length of the paths that content queries and replies traverse, and thus reduces the power consumed to transmit the packets. However, this caching mode requires routers to host content caches. In order to match the speed of the line cards, routers need additional processing to rapidly search the existing cached content. Additional storage coupled with additional processing will inevitably consume more power. We explore these issues further in Sections III and IV.

III. MODELING POWER CONSUMPTION

In this section, we first introduce our general power consumption model (Section III-A). Next, we use a high-level model to characterize the baseline power consumption of a router (Section III-B). Then we present models for the forwarding-decision-making module (Section III-C) and content caching module (Section III-D), to capture the influence of FIA design choices on power consumption.

A. Overview

To model power consumption, we propose a generic router model that captures the forwarding behaviors of both IP routers and FIA routers, as Figure 1 shows. According to the two design approaches (forwarding method and caching method) that we investigate, we separate the content cache module and the forwarding-decision-making module from other router components. Table I summarizes the design choices that FIAs use for these two different modules.

	Forwarding Type		Cache Type
Architecture	PCS	RTL	
IP		TCAM	Edge
NDN		SRAM-BF	Pervasive
NEBULA	PoC& PoP		Edge
SCION	Hop Field		Edge

TABLE I
METHODS USED BY NETWORK ARCHITECTURES FOR MAKING FORWARDING DECISIONS AND CACHING CONTENT.

We group the rest of the router components, such as queue management and switching fabric in Figure 1, which are common components for both IP and FIA routers, and treat the power consumption of these components as a baseline for our analysis. Because precise power analysis of all the components is impractical, we make the simplifying assumption that the baseline power consumption of FIA routers is the same as that of an IP router.
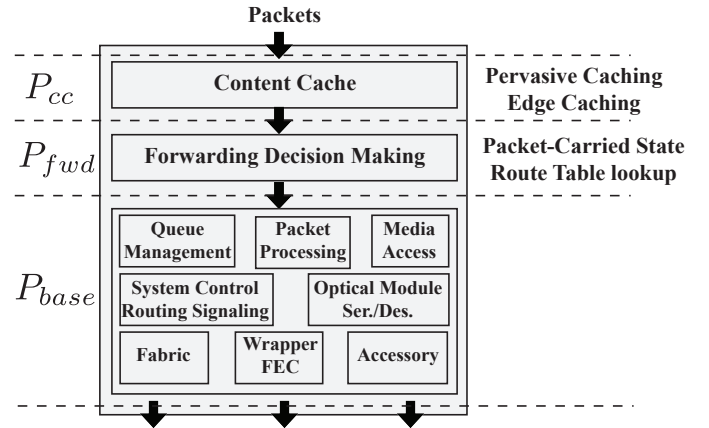

Fig. 1. Abstraction of the forwarding behavior of a FIA router. We present a similar router-component dissection as Tamm et al. [38]. In the figure, "FEC" is short for Forward Error Correction, "Ser./Des." is short for Serialization and Deserialization modules and "Accessory" includes fans, power supplies, shelves, step-up converter, etc.

We denote the total power consumed by an IP or FIA router to forward packets as P^{arch}, the power consumption of local content caching system as P_{cc}^{arch}, the power consumption of making forwarding decisions as P_{fwd}^{arch}, and the baseline power consumption of all the other components as P_{base}. The super-script "*arch*" can be substituted by IP, NDN, SCION, or NEBULA. Thus, $P^{arch} = P_{base} + P_{fwd}^{arch} + P_{cc}^{arch}$. All power consumption is measured in Watts.

B. Modeling Baseline Power Consumption

To model the baseline power consumption (P_{base}), we adopt the methodology of Lee et al. [26] which is capable of deriving P_{base} for heterogeneous Internet routers in core and access networks. We assume that FIA routers use the same technology for the components shared with today's IP routers. Thus, FIA routers consume the same amount of baseline power as their IP router counterparts.

We denote the power consumption of a router when idle as P_{idle}, the power consumption inscribed on the nameplate of the router as P_N, the maximal throughput as I_{max} and the actual throughput as I. We can express P_{base} as $P_{base} = P_{idle} + \alpha(P_N - P_{idle})$, where $\alpha = \frac{I}{I_{max}}$ is a factor characterizing the link utilization. A general observation is that a core router is more power efficient than an edge router. For example, a core router CRS-1 has nameplate power 16.8 kW, and 6.40 Tbps throughput [17]. In contrast, an edge router ARS-1013's nameplate power is 4.0 kW, but only has 0.28 Tbps bandwidth.

C. Modeling Forwarding Decision Making Power Consumption

1) Modeling P_{fwd}^{IP} for IP routers: Various methods exist for an IP router to find the next interface to forward a packet [33]. In this paper, we consider the common hardware-based approach using Ternary Content-Addressable Memory (TCAM). A TCAM can perform a longest-prefix match over the entire routing table with a single access. However, it is known to consume at least three times more energy than Static Random-Access Memory (SRAM) [31] and has a larger chip

size. TCAMs are used in commodity routers such as the Cisco Catalyst 6500 [10].

To model the power consumption of a TCAM-based routing table, we assume that the power consumption of the TCAM is related to its size and its lookup rate. Let E_{TCAM} be the power consumed by TCAM per bit per lookup, s be the size of a prefix record, N be the number of all prefixes stored in the FIB, and r be the average number of packets processed. We can express the P_{fwd}^{IP} of an IP router as $P_{fwd}^{IP} = r \cdot s \cdot N \cdot E_{TCAM}$.

2) Modeling P_{fwd} for NDN routers: TCAMs are inadequate to accommodate the Forwarding Information Base (FIB) for NDN routers [31]. As suggested by Perino et al. [31], we analyze a scheme called "Longest Prefix Match with Bloom Filters" (LPM-BF [18]) as the FIB lookup method for a NDN router instead of TCAM-based lookups.

LPM-BF uses Bloom Filters stored in on-chip SRAM for the task of longest-prefix match in line cards. Compared with longest prefix matching using TCAM, LPM-BF only requires SRAM and DRAM, which are cheaper, smaller in chip size, and larger in capacity. Table II shows the power consumption of different storage mediums. It demonstrates the advantage of using SRAM and DRAM instead of TCAM.

In the LPM-BF scheme, the FIB is organized by a hash table and stored in off-chip DRAM. Bloom filters, each of which is responsible to test matches for prefixes with a specific length, are stored in on-chip SRAM. For each address, all possible prefixes are simultaneously matched against the Bloom filters until a longest-prefix match is found. Then the FIB hash table in DRAM is used to find the next hop for the matched prefix.

We divide the power consumption of an ASIC implementing LPM-BF into two parts: computation and storage. For computation, we primarily consider the computation for the Bloom filters, denoted as P_c^{LPM-BF} in ASICs. For storage, we compute the power consumption of the SRAM and the DRAM required by the longest prefix match tasks in typical line cards, denoted as P_s^{LPM-BF}. We describe the power consumption for forwarding decisions for NDN routers as $P_{fwd}^{NDN} = P_c^{LPM-BF} + P_s^{LPM-BF}$.

Let B be the number of Bloom filters, M be the total number of bits in the on-chip SRAM, and N be the number of all prefixes. According to Dharmapurikar et al. [18], one basic configuration satisfies that $k = \frac{M}{N} \ln 2$ and $f = \left(\frac{1}{2}\right)^k$ where k is the number of hash function, and f is the false-positive rate.

Let E_{hash} be the power consumed to compute a hash function, r be the number of packets requiring longest prefix matching per second. P_c^{LPM-BF} can be expressed as $P_c^{LPM-BF} \sim (Bk + Bf + 1) r \cdot E_{hash}$, where Bk is the number of hashes mandated by the Bloom filters, Bf is the number of hashes caused by the false positives of the Bloom filters, and the additional hash is required to index the hash tables in DRAM. A typical FPGA bitcoin miner today consumes 1 Joule per 20 MHash [2]. Thus, we choose $E_{hash} = 50nJ/Hash$.

Technology	Power (Watt/bit)	Max. Size	Typical Frequency
TCAM	$\sim 3\mu W$	$\sim 32Mb$	$\sim 360MHz$
SRAM	$\sim 40nW$	$\sim 200Mb$	~ 633 MHz
DRAM	$\sim 250pW$	$\sim 64GB$	~ 1333 MHz
Flash	$\sim 0.3pW$	$\sim 2TB$	N.A.

TABLE II
OVERVIEW OF STORAGE TECHNOLOGIES.

Let s be the size of one record in the FIB, β be the load factor of the hash table in DRAM for FIB. The size of DRAM used $O = \frac{S \cdot N}{\beta}$. Let E_{SRAM} be the power consumption of SRAM per bit per access, and E_{DRAM} be the power consumption of DRAM per bit and r_{max} is its maximum frequency. We assume that α is the proportion of the power consumption by DRAM accesses, e.g., read and write. $(1 - \alpha)$ is the proportion of DRAM's activation and background power consumption. A typical value of α for DDR3 DRAM is 46% [?]. P_s^{LPM-BF} can be expressed as $P_s^{LPM-BF} = r \cdot M \cdot B \cdot k \cdot E_{SRAM} + \frac{r \cdot (Bf+1)}{r_{max}} \alpha \cdot O \cdot E_{DRAM} + (1-\alpha) O \cdot E_{DRAM}$. $r \cdot M \cdot B \cdot k \cdot E_{SRAM}$ stands for the power consumed by the on-chip SRAM. $\frac{r \cdot (Bf+1)}{r_{max}} \alpha \cdot O \cdot E_{DRAM}$ stands for the power consumed by DRAM accesses, and $(1 - \alpha) O \cdot E_{DRAM}$ is DRAM's activation and background power consumption.

3) Modeling P_{fwd} for NEBULA and SCION routers: Unlike IP and NDN, both NEBULA and SCION use packet-carried state for finding the interface to forward a packet. In other words, the forwarding decisions resides in the packet header and no route table needs to be stored on routers. The lack of routing tables (and thus lack of relatively expensive table lookup operations) helps reduce the power consumption of packet forwarding. However, both NEBULA and SCION routers use cryptographic primitives to verify the integrity of the routing decisions embedded in the packet headers, which add to P_{fwd}.

Since the verification of the routing decisions carried in packets is the only computation-intensive operation in the forwarding process for NEBULA and SCION, we only consider the computation of cryptographic verification when modeling P_{fwd}. Let E_{verif}^{FIA} be the power consumed to verify the routing decision carried in the packet. We express P_{fwd} as $P_{fwd}^{FIA} = r \cdot E_{verif}^{FIA}$.

In NEBULA, the verification process involves verifying the "Proof of Consent" (PoC) and "Proof of Provenance" (PoP) carried in the packets as well as generating new PoPs to prove provenance. Let l_{AS} be the average AS-level path length. The average energy consumed on a NEBULA border router to verify packet-carried routing decisions E_{verif}^{NEBULA} can be expressed as $E_{verif}^{NEBULA} = E_{hash} + (l_{AS}^2 + l_{AS} + 2)E_{AES}$. We choose $l_{AS} = 4.4$ as measured by Kuhne and Asturiano [4].

In comparison, the verification process on a SCION border router requires only one AES-MAC computation to verify that the HF was generated by the border router itself. As a result, the energy consumed by a SCION border router to verify the packet-carried state E_{verify}^{SCION} can be expressed as $E_{verif}^{SCION} = E_{AES}$. A single 128-bit AES operation on an Intel CPU with AESNI technology consumes 4.8 cycles/byte on i7-980X with frequency 3.3GHz and 12 threads [22]. The maximal power consumption of i7-980X CPU is 130W. Accordingly, we choose $E_{AES} = 250nJ/AESop$.

Comparison of P_{fwd} among FIAs Figure 2 graphs the power consumption of forwarding decision making for a line card plugged in a border router of different FIAs when link speed varies from 1Gbps to 40Gbps. For the FIB size of NDN, Perino and Varvello have suggested 20 million entries [31]. Accordingly, we vary NDN's FIB size from 500K entries to 50M entries to demonstrate the influence of routing-table size

on routers' power consumption. Because we assume that the power consumption of a router's TCAM is only related to the TCAM's size, the power consumed by an IP router is constant when link speed changes.

In general, making forwarding decisions using packet-carried state consumes less power than that using routing table lookup. Depending on the complexity of the packet-carried state verification, the power consumed can vary significantly. In NEBULA which requires several crypto-operations to forward each packet, the performance improvement over IP is 1 order of magnitude. Since SCION only requires a single AES-MAC to verify the packet-carried state, the performance of making a forwarding decision is more than 3 orders of magnitude more efficient than that of an IP border router.

On the other hand, because LPM-BF consumes less power at low link speed, routing table lookups in NDN with 500K records in the route table consumes 10X less power than that using TCAM in an IP border router. However, the advantage becomes less apparent as the routing table size grows. As a matter of fact, the routing-table lookup using LPM-BF in NDN at 40Gbps with 50M records in each route table consumes up to 12 times more power than that using TCAM in an IP border router.
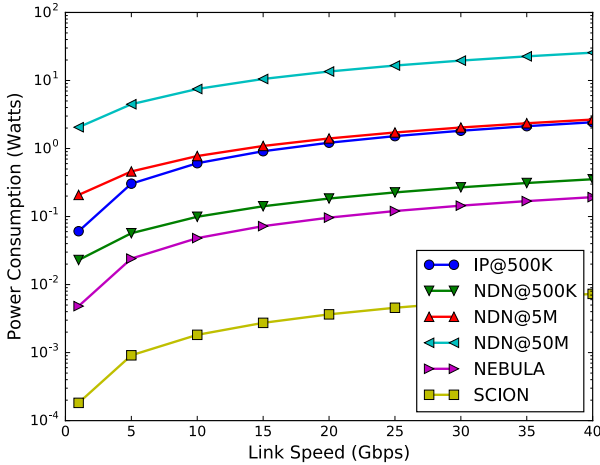

Fig. 2. P_{fwd} under different link speed for border routers. For NDN, we evaluate routing tables containing 500K, 5M, and 50M entries.

D. Modeling Content Caching Power Consumption

1) Edge caching P_{cc}: To estimate the power consumption of edge caching, we consider a state-of-the-art of content cache from Netflix [7]. Netflix is the largest video distribution service in the world, accounting for as much as 34% of worldwide network traffic during peak hours [34]. As one of the leaders in the space, Netflix has strong incentives to use powerful yet power-efficient hardware. Netflix's content routers have storage capacities of 108TB, support 10Gbps network throughput and consume 600 Watts [7]. For comparison purposes in this paper, we use 600W/10Gbps as our baseline P_{cc} for networks with edge caching.

2) Pervasive caching P_{cc}: In current FIAs, the design and implementation of content caches (particularly caches in high-speed core routers) remains an open problem. Content cache

systems in FIAs generally have the following design goals: 1) rely upon inexpensive hardware to encourage massive deployment; 2) allow provisioning of higher storage capacity as link speeds increase; and 3) serve contents at high speeds, ideally close to the arrival rates of packets.

Since pervasive caching routers are still a relatively new concept, measurement data on power consumption of real-world devices is, to the best of our knowledge, not yet available. Thus, we consider two key-value object store schemes (HashCache [14] and SILT [27]) that fulfill the design requirements. We envision content cache in FIAs to use HashCache or SILT, whichever consumes the least power.

Key-value store. Both HashCache and SILT implement a two-layer architecture: 1) an underlying storage layer in large and relatively slow medium storing the actual objects, and 2) an indexing layer in a small and relatively fast medium to efficiently handle the queries and locate the corresponding content. A typical setup for these key-value stores includes an SSD-based storage layer indexed by a DRAM-based indexing layer.

There are three key parameters characterizing the key-value store schemes: κ, determining the number of bytes needed in the index layer for each object in the storage layer, and two amplification factors, A_{rd} and A_{wr}, determining the number of reads or writes required for the storage layer when there is a read hit or a write hit in the storage layer. Table III lists the typical values for both two key-value store schemes.

We use a conservative estimate about the power consumption of key-value stores by only accounting for the power consumption of the underlying storage medium. We leave out the power consumed by hash computations needed, because the hash functions can be computed very efficiently [3].

Let C_{st} be the storage capacity of the storage layer, E_{st} be the power consumed to store each bit by the storage medium supporting the storage layer, and E_{idx} be the power consumed to store each bit by the storage medium implementing the index layer. We derive $P_{cc} = (E_{st} + \kappa E_{idx}) C_{st}$.

Choosing storage mediums. Selecting storage mediums for both the storage and index layers involves considering both the storage capacity and transaction rates desired for the key-value store and those offered by current technologies. Specifically, we take into account two categories of limitations: *storage-capacity limitation* and *transaction-rate limitation*.

Let (C_1^{max}, R_1^{max}) be the maximal storage capacity and maximal transaction rate for the storage medium used by the index layer, and (C_{st}, R_{st}) be those for the storage medium used by the storage layer. We express the *storage-capacity limitation* by: 1) $\kappa C_{st} \leq C_1^{max}$, and 2) $C_{st} \leq C_2^{max}$.

Let λ_{in} be the arrival rates of content-distribution-relevant packets, α be the percentage of the content queries (the other packets are data packets), r_{hit} be the percentage of cache hit rate for the content queries, r_{mod} be the probability of writing to add new cached objects. We express

Method	κ (Bytes/Object)	A_{read}	A_{write}
SILT	1	1.01	4
Hashcache(SetMem)	11/8	1	1
Hashcache(logLRU)	(15∼47)/8	1	1

TABLE III
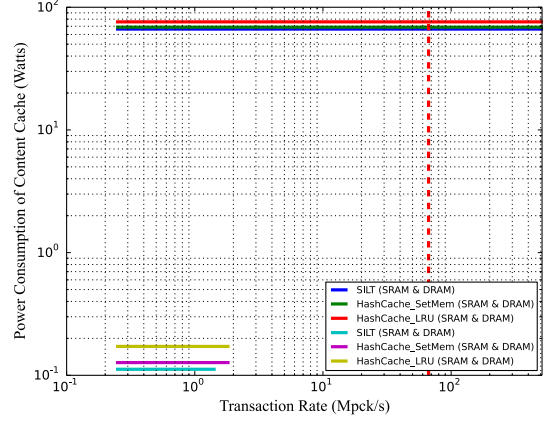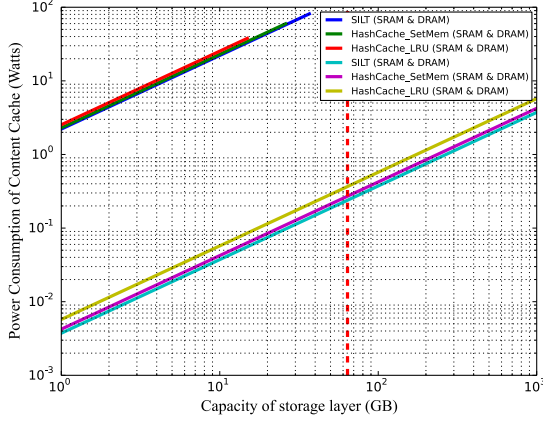KEY PARAMETERS FOR CONTENT STORE ALGORITHMS.

Fig. 4. (a) The power consumed by content caches of various sizes. Either SRAM-DRAM or DRAM-SSD is chosen as the combination of storage mediums. The red dashed line shows the maximal size of DRAM. (c) The power consumed by content caches for different transaction rates with parameters α=0.5, r_{hit} = 0.1, r_{mod} = 0.01. The red dashed line shows the maximal transaction rate of DRAM.
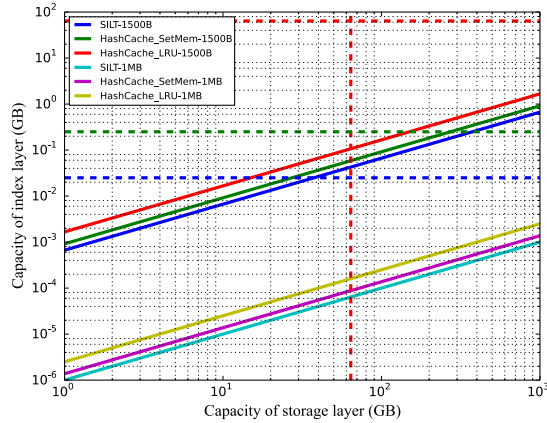

Fig. 3. The size of the index layer as a function of the size of the storage layer for different key-value stores. The blue, green, and red dashed line shows the maximal sizes of SRAM, RLDRAM and DRAM, respectively.

the *transaction-rate limitation* as: 1) $\lambda_{in} \leq R_1^{max}$, and 2) $(\alpha r_{hit}A_{read} + (1-\alpha)r_{mod}A_{write}) \leq R_2^{max}$.

Figure III-D2 shows the size of the index layer as a function of the size of the storage layer for different key-value stores. When the object size is small (e.g., 1500 bytes), the size of the index layer is the major bottleneck to build a content cache with large capacity and high transaction rates. In contrast, when the object size is larger (e.g., 1MB) such as multimedia content, the size of the DRAM becomes the bottleneck.

Figures 4(a) and 4(b) show the power consumed by content cache with different sizes and capable of handling different transaction rates. In general, an SRAM-DRAM combination could offer 300X higher transaction rates while consuming 600X more power than a DRAM-SSD combination. For a router with high link speed ($\geq$ 10Gbps), an SRAM-DRAM combination could be leveraged to implement the content cache, while DRAM-SSD combination is suitable to implement a power-efficient content cache for a router with low link speed ($\leq$ 2Gbps).

As for the comparison between different key-value stores, SILT benefits from its smaller value of κ in two aspects. First,

SILT has a smaller index layer. As the index layer is more power hungry, this feature renders SILT to be 3% and 15% more efficient compared to HC-SetMem and HC-LogLRU, respectively. Second, when implementing packet-level caches using the SRAM-DRAM combination, SILT could implement a content cache with a 1.3 to 2 times larger storage layer.

On the other hand, with larger amplification factors, the transaction rate handled by SILT is lower when the storage layer speed becomes the bottleneck. This is the case when DRAM-SSD is chosen to implement content cache of larger capacity, where the maximal transaction rate of SILT is 25% lower than those of HC-SetMem and HC-LogLRU.

E. Summary

With the models constructed in Sections III-B, III-C, and III-D, we can compute the power consumed to forward packets for each type of FIA routers. To normalize the results, we calculate the energy used for forwarding a single bit on each FIA router.
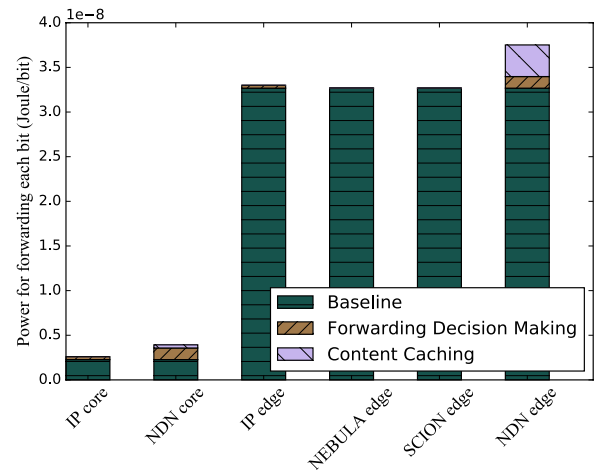

Fig. 5. Power consumption of FIA core and edge routers. "core" means core routers, and "edge" means edge routers. We omit SCION and NEBULA core routers' results as forwarding decision making only happens on edge routers.
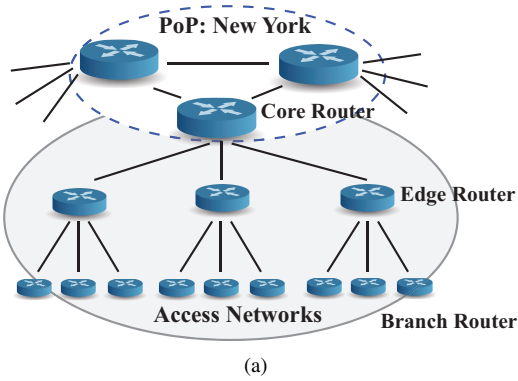
Fig. 6. Access tree connected to the New York PoP. The access tree shown is a complete tree with depth=3, arity=3.

Figure 5 shows the Joule per bit for each FIA router with a specific configuration. For NDN, each core router is equipped with 1TB DRAM-based content cache managed by SILT with SRAM as the index layer and each edge router is equipped with 256GB content cache using the same key-value store setup. IP, NEBULA, and SCION routers do not perform any content cache. All interfaces of NDN routers use ASICs based on LPM-BF to make forwarding decisions. We assume the number of entries in NDN's FIB is 20M and the number of entries in the IP routing table is 512K.

According to Figure 5, NDN routers consume more power than IP and other FIA routers due to their content cache and larger FIB. The increase in the power budget for a NDN router ranges from 72% for core routers to 15% for edge routers compared to a SCION router, which consumes the least power due to its efficient forwarding.

Though SCION and NEBULA are more power efficient than NDN when we only consider traffic forwarding, content caches may reduce the overall number of bits that are transmitted. In the following section, we further explore this trade-off by performing large-scale simulations.

IV. SIMULATION

Based on the model of FIA routers in Section III, we now compare the power consumption of each FIA in content distribution scenarios. We conduct our experiments by simulating the forwarding behaviors of the IP network and FIAs when used for content distribution.

A. Simulation Setup

Topology. The topology used in our simulations is based on education backbone networks and the Rocketfuel topology [37]. We obtain router-level information for two education/research backbones: Geant and Abilene. We also extract router-level information for six different ISPs: Telstra (AS1221), Sprint (AS1239), NTT (AS2914), Verio (AS3257), Level3 (AS3356), AT&T (AS7018). Routers are grouped according to their Points of Presence (PoPs). The PoPs are then annotated with city-level location information.

We follow the methods proposed by Fayazbakhsh et al. [20] to approximate access networks by trees appended to each PoP. The internal nodes of the trees are edge routers. We use complete trees with varying depths and arities.

Traffic Patterns. Content distribution traffic represents a significant amount of overall Internet traffic. Thus, our primary goal is to evaluate the power consumption of FIAs in content distribution scenarios. For traffic access patterns, previous work has suggested that a Zipf distribution closely approximates real world content access from end hosts [20]. The key parameter α in a Zipf distribution decides the relative popularity of different contents. A larger α means that popular content queries constitute a larger proportion of all queries, which also means more temporal locality in the content access pattern.

We use synthesized content-access traces with $\alpha = 0.99$ (which approximates US users' behaviors) as suggested by Fayazbakhsh et al. [20]. For the query distribution, we employ the population of the city for each PoP to distribute the queries across the access networks belonging to each PoP. For simplicity, the queries are assumed to only enter the network through leaf nodes of each access network.

B. Routing-table lookup versus packet-carried state

We begin by evaluating the power consumption with respect to the first design choice (i.e., routing-table lookup versus packet-carried state) for making forwarding decisions. For architectures leveraging routing-table lookups, we consider current IP networks and NDN. For those using packet-carried state, we consider SCION and NEBULA. To prevent measurement noise induced by caching, we intentionally remove the content caching module for NDN in the simulation. We will add it back in Section IV-C.

As discussed in Section III-C, the primary sources of routing table power consumption are routing table maintenance and routing table lookup operations. We assume TCAMs are used for the IP network and LPM-BF is used for NDN, because LPM-BF can search a larger NDN routing table and consume less power. In the case of IP networks, we set the number of prefixes to be 500K according to the FIB size for BGP from RouteViews [9]. In case of NDN, we choose the number of (content name) prefixes to be 20M which can be supported by a 200Mbit on-chip SRAM.

The main sources of power consumption for packet-carried state forwarding are state verification and transmission of extra bits in packet headers. For the verification of packet-carried state, we assume 128-bit AES as the pseudo-random permutation to construct multiple crypto-primitives in NEBULA and SCION. For the packet payload size, we select 1350 bytes [35] for content responses and 40 bytes for content queries.

We focus on inter-domain forwarding decisions because BGP routing tables are several orders of magnitude larger than intra-domain routing tables. Furthermore, we assume only the network layer protocols differ among FIAs, but all the other layers (transport layer, application layer, etc.) in the network stack remain the same.

Figure 7 demonstrates the power consumed by the synthesized content-access traces in different AS topologies. Across all AS topologies, *making forwarding decisions by using packet-carried state is 15% more efficient than by doing routing table lookups*. The reason, as partially described in Section III-C, is that maintaining and searching routing-tables
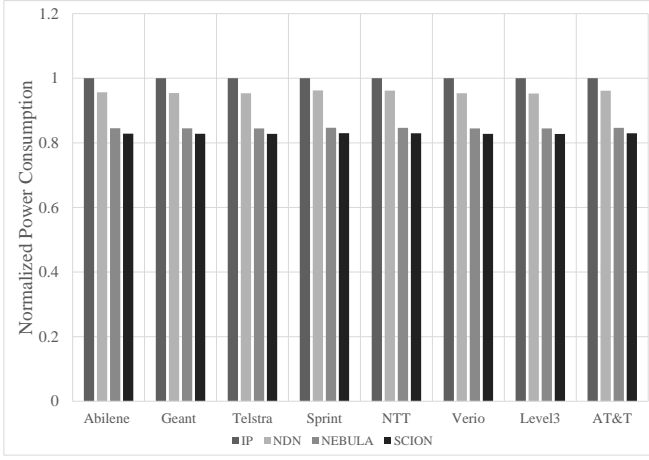
Fig. 7. Power consumption of IP, NDN, NEBULA and SCION. There is no edge cache in the network. For each topology, the results are normalized by baseline results of the IP network.
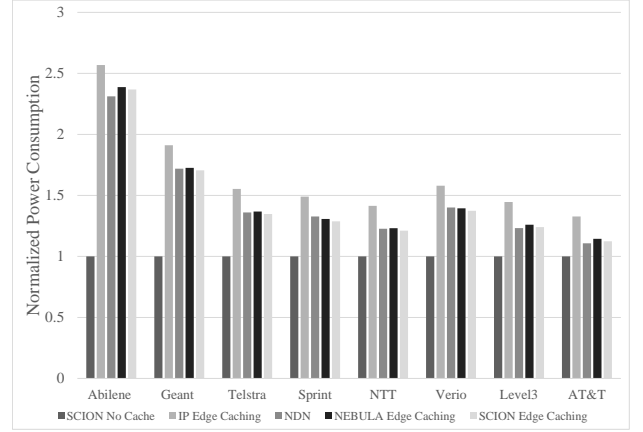

Fig. 8. Power consumption of end-to-end communication, edge caching and pervasive caching with capacity ratio $c = 0.05$. All results are normalized by SCION with no content caching.

on individual routers consumes more power than including forwarding decisions within each packet.

Next, we compare FIAs using the same forwarding method. For IP and NDN, both of which use routing-table lookup, NDN routers consume 4% less power than IP due to efficent forwarding-decision making. Though we assume NDN routers have larger routing tables that contain 20M entries each, NDN routers consume less power because LPM-BF allows NDN routers to store and search routing tables more efficiently.

For NEBULA and SCION, both of which use packet-carried state, SCION routers consume 3% less power than NEBULA routers. This is partially due to packet-carried state verification being more expensive in NEBULA routers than in SCION routers and partially due to NEBULA's larger packet headers.

C. Edge caching versus pervasive caching

In this section, we consider the influence of different caching methods on the power consumption of packet forwarding in different network architectures. Particularly, we consider IP, NEBULA, SCION with edge caching and NDN which inherently supports pervasive caching. For completeness, we also evaluate IP, NEBULA and SCION without content caching. Because SCION without caching was shown to be the most power efficient in the previous section, we use it as the baseline result to normalize the results of other architectures. While definitions of edge caching may vary, our simulation follows the one used by Fayazbakhsh et al. [20], i.e., only leaf nodes in access networks cache contents.

Cache Budget Ratio. Cache capacity is the primary factor that impacts cache performance and the power consumption of the caching device. We define cache budget ratio in order to fairly compare edge caching and pervasive caching. Let R be the number of routers capable of caching, C be the average cache capacity of each router, O be the total number of individual contents in the network, and s be the average size of each content. We define the cache budget ratio c as $c = \frac{R \times C}{O \times s}$.

We choose $c = 5\%$ as a baseline which is observed as a relationship between CDN cache provisioning and the total

requested objects seen by the cache each day [20]. We assume that cache capacity is uniformly distributed among router.

Cache Replacement Strategy. Cache replacement strategies are important for content routers to exploit the locality of content accesses. Therefore, cache replacement strategies are expected to influence the performance of content routers in power consumption. In our simulation, we select the Least-Recently Used (LRU) method as our baseline strategy. Note that designing an optimal or high-performance cache replacement strategy is out of scope for this paper.

Cache Discovery Strategy. For edge caching used in IP, NEBULA and SCION, we simply assume that content request is only served by each standalone cache server. Infrastructure to coordinate the cache servers [11] is not provisioned. We call this strategy *simple edge caching*. For pervasive caching, we assume *on-path cache discovery*, in which only content cached in the on-path routers would be served.

Figure 8 shows the power consumed by different network architectures with caching. We use the baseline capacity ratio $c = 0.05$, and we also evaluate *simple edge caching* as the cache discovery strategy for networks with edge caching and *on-path cache discovery* as the strategy for networks with pervasive caching. We use a synthesized content-access trace following a Zipf distribution with $\alpha = 0.99$, which is the α computed from a real-world content-access trace [20].

Surprisingly, network architectures with caching enabled tend to consume 15-100% more power than those that do not cache. Because a smaller α value implies less locality in the content access pattern, caching becomes less efficient, which results in caching consuming more power in the access pattern scenario of our simulation. In other words, *networks with end-to-end communication without caching seems to consume less overall power compared to those that use caching*.

Regarding the comparison among network architectures with caching, NDN with pervasive caching only saves on-average 2% power compared to SCION with edge caching, which is the most power-efficient among IP, NEBULA, and SCION with edge caching. Compared to IP, which also leverages routing-table lookup for making forwarding decisions, NDN consumes up to 16% less power. The result implies

that *pervasive caching helps reduce power consumption, but the power budget cut is limited*. The reason is two-fold: 1) multiple-layer caching or cooperative caching provide limited improvement to single-layer caching, as indicated by previous works [42], [20]; 2) pervasive caching requires more power-consuming caching devices, which further reduces the small advantage in power consumption by having shorter length.

Finally, *SCION with no caching consumes the least amount of power*. According to our previous analysis, SCION with no caching benefits from two design choices: 1) efficient verification of packet-carried state for making forwarding decisions, 2) end-to-end design without caching.

Sensitivity analysis. We also conduct sensitivity analysis with respect to different cache budget ratios c, different content access patterns, and multiple cache discovery strategies. Results show that the observations remain true for various combinations of parameters. We document our evaluation method and results of sensitivity analysis in detail in our technical report [?].

D. Summary of Key Observations

1) Network architectures that use packet-carried state instead of routing-table lookups exhibit lower power consumption. This observation holds even in the presence of larger packet headers.
2) FIAs without caching consume less overall power compared to those that use caching.
3) The use of pervasive caching results in marginal reductions in power consumption.
4) Among the studied FIAs, SCION with no caching consumes the least amount of power.

E. Sensitivity Analysis

In order to further understand the influence of various parameters on our results, we conduct a one-dimensional sensitivity analysis: for each analysis, we vary only one of the parameters while fixing all the other parameters.

1) Cache budget ratio: For analyzing the influence of cache budget ratio on our observations, we vary the cache budget ratio from $c = 0$, where no content cache is present in the network, to $c = 1$, where all content in the Internet can be cached in the network. Note that the current cache budget for major CDN provider is around 5% ($c = 0.05$). Thus our analysis covers a wide range of values for the cache budget ratio parameters.

Figure 9(a) graphs the power consumption of different networks with caching under different cache budget ratios. With increasing cache budget, the power consumption of different networks all increases. When the cache budget reaches 1, NDN with pervasive caching, which consumes least amount of power among the candidates, still consumes 150% more power than SCION with no caching.

On the other hand, the power consumption for NDN with pervasive caching increases at a slower rate compared to IP, NEBULA, and SCION. NDN with pervasive caching consumes the same amount of power as SCION with edge caching when the cache budget is 0.76. NDN with pervasive caching reduces up to 30% the power budget in comparison to edge caching when the cache budget reaches 1. However,

because the current cache budget for large CDN providers is around 0.05 [20], which is far smaller than 0.76, we expect our observation 3 still holds in the near future.

2) Content Access Locality: For this analysis, we measure power consumption for different network architectures when varying Zipf distribution parameter α. Larger α value indicates more locality in content access. With increasing locality in the content accesses, the both edge caching and pervasive caching will reduce energy footprints.

In Figure 9(b), we show that the energy footprints for IP, NEBULA, SCION with edge caching are still higher than that of SCION without cache even with high locality in the access pattern ($\alpha=1.5$ compared to the case when $\alpha=0.99$ in content accesses from US). On the other hand, for NDN, its power efficiency will catch up with that of SCION without caching when α surpasses 1.1. With high locality ($\alpha=1.5$) NDN's power efficiency outperforms SCION without caching by 25%. Therefore, our observation 4 will not hold when the Zipf distribution parameter is over 1.1. We note, however, that current α values are 0.99 for US, 0.92 for Europe, and 1.04 for Asia. Thus, for the time being, SCION appears to be the most power efficient architecture.

3) Cache Discovery Strategy: For pervasive caching, we compare two cache discovery strategies: *on-path cache discovery*, in which only content cached by the on-path routers are served, and *nearest cache discovery*, in which the request of the content is redirected to the nearest router caching the content.

For edge caching, we compare two cache discovery strategies: *simple edge caching*, in which each caching server serves a content request only from local caches, and *cooperative cache discovery*, in which each caching server redirects content requests to the nearest servers that cache the content.

We realize that achieving nearest cache discovery and cooperative cache discovery require sophisticated mechanisms which in turn consume additional unaccounted power. However, by intentionally omitting the power consumed to distribute information about cached content to neighbour routers, we treat the nearest cache discovery and cooperative cache discovery as the optimal cases to characterize the boundaries of cache discovery strategy space.

Figure 9(c) shows the power consumed by architectures with different caching strategies. We are interested in the differences in power consumption that are caused by the various cache discovery strategies. In general, the differences between different cache discovery strategies for both networks with edge caching and NDN with pervasive caching lie within 2%. Our observation 2, 3, and 4 hold with various cache strategies.

4) Summary: When varying the cache budget, all our observations hold. We see similar results when varying the cache discovery strategy. When varying the locality in access patterns, we note that certain values for content access distribution weaken our observations. Specifically, with high locality, NDN has the potential to be more power efficient than SCION. In addition, these high locality scenarios show a more positive effect on the benefits in power consumption due to pervasive caching.
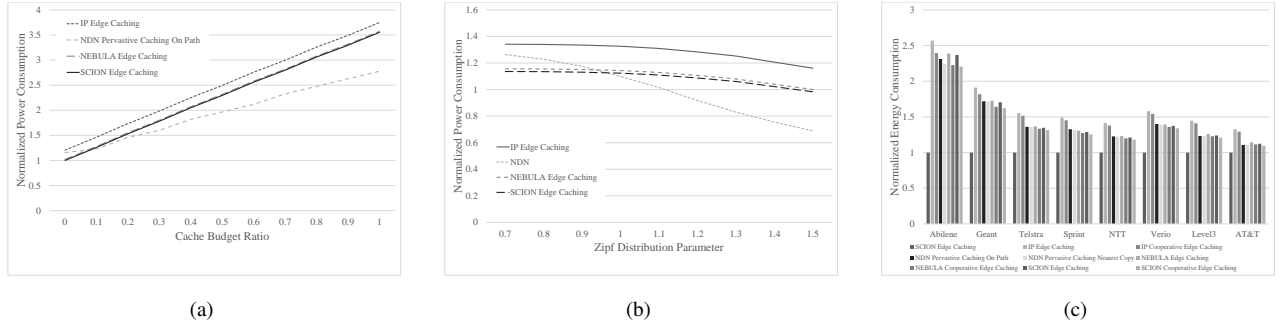
Fig. 9. (a) Power consumption of edge caching and pervasive caching with different cache budget ratios. All results are normalized by the power consumption of SCION with no cache. (b) Power consumption of edge caching and pervasive caching with different Zipf distribution parameters. All results are normalized by the SCION without cache. (c) Power consumption of edge caching and pervasive caching with different cache discovery strategy. For edge caching, simple edge caching and cooperative edge caching are considered. For pervasive caching, on-path cache discovery and nearest-copy discovery are evaluated. All results are normalized by SCION with edge caching.

Observation	CB	LAP	CDS
1. PCS consumes less power than lookups	✓	✓	✓
2. End-to-end consumes less power than caching	✓	✓	✓
3. Pervasive caching offers marginal power consumption improvements	✓	✗	✓
4. SCION consumes least amount of power	✓	✗	✓

TABLE IV

SUMMARY OF THE SENSITIVITY ANALYSIS RESULTS. CB STANDS FOR CACHE BUDGET, LAP STANDS FOR LOCALITY IN ACCESS PATTERN, CDS STANDS FOR CACHE DISCOVERY STRATEGY. A ✓ MEANS THAT OUR OBSERVATION REMAINS THE SAME WHEN THE FACTOR FOR THE CURRENT COLUMN VARIES IN THE FULL RANGE THAT WE CONSIDERED. A ✗ MEANS THAT OUR OBSERVATION REMAINS THE SAME WHEN THE FACTOR FOR THE CURRENT COLUMN IS RESTRICTED IN A CERTAIN RANGE.

V. DISCUSSION

In this paper, we concentrate on data-plane traffic power consumption in content delivery scenarios. Bolla et al. [16] estimate that data-plane traffic consumes 83% of the total Internet power compared to 17% consumed by control-plane traffic. Furthermore, content delivery applications constitute a majority of the Internet traffic. For example, Netflix and Facebook together account for 47% of the downstream traffic today [34]. Our analysis assumes that the FIAs' control-plane still consumes a small proportion of the total power, and content delivery applications' traffic still constitutes the majority of the Internet traffic. Since implementation details regarding the FIAs' control-plane behaviors are not yet fully specified [25], [32], [13], [44], in-depth analysis of control-plane power consumption behavior remains an open problem. In addition, the analysis presented herein does not capture power consumption behavior of real-time traffic, such as Skype communication. We defer the analysis of real-time traffic power consumption to future work, but we expect the results to be consistent with the observations in this paper.

IP forwarding techniques. In our analysis of the fowarding-decision-making module, we have chosen TCAM as the underlying technology for IP routers, because it has been widely used by ASICs in commodity routers [10]. Admittedly, there are many alternative methods for searching routing-tables [23]. For example, Cisco builds ASICs in the CRS-1 router for making forwarding decisions based on a treebit map with reduced-latency DRAM (RLDRAM) [19]. The detailed comparison of different forwarding hardware is out of scope for this paper.

Content caches on routers. We assume NDN routers use key-value stores to build content caches. Key-value stores provide high query rates while minimizing power consumption. However, the key-value stores we analyze (HashCache and SILT), are both built for persistent storage. We expect that a key-value store built exclusively for caching can consume less power, and thus further reduces power consumption in NDN when using a pervasive caching layer.

VI. RELATED WORK

Power consumption of routers, the Internet, and FIAs. Power consumption of the Internet infrastructure has been well studied at various levels of granularity. Ye, Micheli, and Benini theoretically model the power consumption of the switching fabric in routers by their electrical components, such as capacitors [43]. Baliga et al. [15] and Tucker et al. [39] model the power consumption of optical IP networks based based on the power consumption of individual heterogeneous routers, switches, etc. Lee et al. [26] estimate the power consumption of CCNs. Our paper adds to this body of work by analyzing both the power consumption of FIA routers themselves, and large simulated FIA networks. Compared to previous work, our general power consumption model spans across two levels of granularity: 1) the power consumption of the computation and storage needed by different FIA router components; and 2) overall FIA network power consumption under various caching strategies and workloads.

FIA evaluation and metrics. Previous work mainly focuses on the evaluation of content-centric networks (CCNs) or on the evaluation of specific CCN subsystems. Fricker et al. [21] evaluate the caching performance of CCNs under the influence of network traffic compositions. Muscariello, Carofiglio and Gallo [28] evaluate the performance of bandwidth and shared storage in CCN designs. Fayazbakhsh et al. [20] use network latency, network congestion and origin server load to compare different caching strategies in CCNs. Perino and Varvello [31] use router throughput, monetary costs, and energy efficiency to estimate the feasibility of deployment of CCN routers. The main objective of our work is not limited to the exploration of power consumption of each individual FIA. Rather, we concentrate on presenting an evaluation framework to explore the power implications of

adopted FIA design principles. The results obtained herein can help guide designers toward power-efficient network designs.

VII. Conclusion and Future Work

In this paper we have modeled and compared the power consumption of future Internet architectures. We performed experiments at multiple levels of granularity ranging from per-bit power consumption of router components, to network-wide power consumption under the use of different caching strategies. From our analysis, we were able to draw several observations: 1) the use of packet-carried state is more power-efficient than routing table lookups; 2) based on our workload assumptions, end-to-end communication consumes less power than using in-network caches; and 3) there is no substantial difference between energy footprints of networks with edge caching as compared to ones with pervasive caching.

We propose power consumption as a general unified metric to optimize networks, as lower energy translates into smaller amounts of work performed. Thus, power minimization also optimizes the amount of equipment used, network performance, and environmental impact. We hope that our approach serves as a useful step toward making power analysis a common evaluation mechanism for network architectures.

VIII. Acknowledgments

References

[1] Berkeley lab study highlights growing energy impact of internet video streaming. http://newscenter.lbl.gov/2014/06/02/berkeley-lab-study-highlights-growing-energy-impact-of-internet-video-streaming/.

[2] Bitcoin mining hardware comparison. http://en.bitcoin.it/wiki/Mining_hardware_comparison.

[3] How fast are hash functions. http://cr.yp.to/talks/2008.06.05/slides.pdf.

[4] Interesting graph: As path length over time. http://labs.ripe.net/Members/mirjam/interesting-graph-as-path-lengths.

[5] Internet routers hitting 512k limit, some become unreliable. http://arstechnica.com/security/2014/08/internet-routers-hitting-512k-limit-some-become-unreliable/.

[6] List of nuclear reactors. http://en.wikipedia.org/wiki/List_of_nuclear_reactors.

[7] Netflix Open Connect. https://www.netflix.com/openconnect/hardware.

[8] Power, Pollution and the Internet. http://www.nytimes.com/2012/09/23/technology/data-centers-waste-vast-amounts-of-energy-belying-industry-image.html?pagewanted=all&_r=0.

[9] Routeview project. http://www.routeviews.org/.

[10] Cisco Catalyst 6500. http://www.cisco.com/c/en/us/products/switches/catalyst-6500-series-switches.

[11] Akamai. http://www.akamai.com.

[12] Ashok Anand et al. XIA: An architecture for an evolvable and trustworthy Internet. In *ACM Workshop on Hot Topics in Networks*, 2011.

[13] Tom Anderson et al. The NEBULA Future Internet Architecture. In *The Future Internet*, pages 16–26. Springer, 2013.

[14] Anirudh Badam, KyoungSoo Park, Vivek S. Pai, and Larry L. Peterson. HashCache: Cache Storage for the Next Billion. In *USENIX NSDI*, pages 123–136, 2009.

[15] Jayant Baliga, Robert Ayre, Kerry Hinton, Wayne V Sorin, and Rodney S Tucker. Energy consumption in optical IP networks. *IEEE Journal of Lightwave Technology*, 27(13):2391–2403, 2009.

[16] Raffaele Bolla, Roberto Bruschi, Franco Davoli, and Flavio Cucchietti. Energy efficiency in the future internet: a survey of existing approaches and trends in energy-aware fixed network infrastructures. *IEEE Communications Surveys & Tutorials*, 13(2):223–244, 2011.

[17] Cisco. Cisco router products. http://www.cisco.com/en/US/products/hw/routers/.

[18] Sarang Dharmapurikar, Praveen Krishnamurthy, and David E Taylor. Longest prefix matching using bloom filters. In *ACM Conference on Applications, technologies, architectures, and protocols for computer communications*, pages 201–212, 2003.

[19] Will Eatherton, George Varghese, and Zubin Dittia. Tree bitmap: hardware/software IP lookups with incremental updates. *ACM SIGCOMM Computer Communication Review*, 34(2):97–122, 2004.

[20] Seyed Kaveh Fayazbakhsh, Yin Lin, Amin Tootoonchian, Ali Ghodsi, Teemu Koponen, Bruce Maggs, KC Ng, Vyas Sekar, and Scott Shenker. Less pain, most of the gain: incrementally deployable ICN. In *ACM SIGCOMM*, pages 147–158, 2013.

[21] Christine Fricker, Philippe Robert, James Roberts, and Nada Sbihi. Impact of traffic mix on caching performance in a content-centric network. In *Computer Communications Workshops (INFOCOM Workshops)*, pages 310–315, 2012.

[22] Shay Gueron. Intel aes new instructions (aesni): white paper. 2012.

[23] Pankaj Gupta and Nick McKeown. Algorithms for packet classification. *IEEE Network*, 15(2):24–32, 2001.

[24] Kerry Hinton, Jayant Baliga, Michael Z Feng, RWA Ayre, and Rodney S Tucker. Power consumption and energy efficiency in the Internet. *IEEE Network*, 25(2):6–12, 2011.

[25] Van Jacobson, Diana K. Smetters, James D. Thornton, Michael F. Plass, Nicholas H. Briggs, and Rebecca L. Braynard. Networking named content. In *Conference on Emerging Network Experiment and Technology*, pages 1–12, 2009.

[26] Uichin Lee, I. Rimac, D. Kilper, and V. Hilt. Toward energy-efficient content dissemination. *IEEE Network*, 25(2):14–19, 2011.

[27] Hyeontaek Lim, Bin Fan, David G Andersen, and Michael Kaminsky. SILT: A memory-efficient, high-performance key-value store. In *ACM SOSP*, pages 1–13, 2011.

[28] Luca Muscariello, Giovanna Carofiglio, and Massimo Gallo. Bandwidth and storage sharing performance in information-centric networking. In *ACM SIGCOMM Workshop on Information-centric Networking*, pages 26–31, 2011.

[29] Jad Naous et al. Verifying and enforcing network paths with icing. In *ACM CoNEXT*, page 30. ACM, 2011.

[30] Li-Shiuan Peh and William J Dally. A delay model for router microarchitectures. *IEEE Micro*, 21(1):26–34, 2001.

[31] Diego Perino and Matteo Varvello. A reality check for content centric networking. In *ACM SIGCOMM Workshop on Information-centric Networking*, pages 44–49, 2011.

[32] Dipankar Raychaudhuri, Kiran Nagaraja, and Arun Venkataramani. MobilityFirst: a robust and trustworthy mobility-centric architecture for the future Internet. *ACM SIGMOBILE Mobile Computing and Communications Review*, 16(3):2–13, 2012.

[33] Miguel Á Ruiz-Sánchez, Ernst W Biersack, and Walid Dabbous. Survey and taxonomy of IP address lookup algorithms. *IEEE Network*, 15(2):8–23, 2001.

[34] Sandvine. Global Internet Phenomena Report. http://www.sandvine.com/downloads/general/global-internet-phenomena/2013, 2013.

[35] Stefan Saroiu, Krishna P Gummadi, Richard J Dunn, Steven D Gribble, and Henry M Levy. An analysis of internet content delivery systems. In *USENIX OSDI*, pages 315–328, 2002.

[36] Arman Shehabi, Ben Walker, and Eric Masanet. The energy and greenhouse-gas implications of internet video streaming in the United States. *Environmental Research Letters*, 9(5), 2014.

[37] Neil Spring, Ratul Mahajan, and David Wetherall. Measuring ISP topologies with Rocketfuel. *ACM SIGCOMM Computer Communication Review*, 32(4):133–145, 2002.

[38] Oliver Tamm, Christian Hermsmeyer, and Allen M Rush. Eco-sustainable system and network architectures for future transport networks. *Bell Labs Technical Journal*, 14(4):311–327, 2010.

[39] R Tucker, Jayant Baliga, Robert Ayre, Kerry Hinton, and W Sorin. Energy consumption in ip networks. In *ECOC Symposium on Green ICT*, page 1, 2008.

[40] Willem Vereecken, Ward Van Heddeghem, Didier Colle, Mario Pickavet, and Piet Demeester. Overall ICT Footprint and Green Communication Technologies. In *International symposium on Communications, Control and Signal Processing*, 2010.

[41] Hang-Sheng Wang, Li-Shiuan Peh, and Sharad Malik. A power model for routers: Modeling Alpha 21364 and InfiniBand routers. In *High Performance Interconnects*, pages 21–27, 2002.

[42] Alec Wolman, M Voelker, Nitin Sharma, Neal Cardwell, Anna Karlin, and Henry M Levy. On the scale and performance of cooperative web proxy caching. In *ACM SIGOPS Operating Systems Review*, volume 33, pages 16–31. ACM, 1999.

[43] Terry Tao Ye, Giovanni De Micheli, and Luca Benini. Analysis of power consumption on switch fabrics in network routers. In *ACM Annual Design Automation Conference*, pages 524–529, 2002.

[44] Xin Zhang, Hsu-Chun Hsiao, Geoffrey Hasker, Haowen Chan, Adrian Perrig, and David G. Andersen. SCION: Scalability, Control, and Isolation on Next-Generation Networks. In *IEEE Symposium on Security and Privacy*, pages 212–227, 2011.