Performance Analysis of In-Network Caching in Content-Centric Advanced Metering Infrastructure

Nour El Houda Ben Youssef  
SAGE-LATIS/ENISO University of Sousse  
ENSI University of La Manouba  
WEVIOO  
Tunisia

Yosra Barouni, Sofiane Khalfallah, Jaleedhine Ben Hadj Slama  
SAGE-LATIS/ENISO University of Sousse  
Tunisia

Khaled Ben Driss  
WEVIOO  
Tunisia

Abstract—In-network caching is a key feature of content-centric networking. It is however a relatively costly mechanism with hardware requirements besides placement/replication strategies elaboration. As content-centric networking is proposed in the literature to manage smart grid (SG) communications, we aim, in this research work, to investigate the cost effectiveness of in-network caching in this context. We consider, in particular, the Advanced Metering Infrastructure (AMI) service that comes into prominence since its outputs are imperative inputs of most smart grid applications. In this research work, AMI communication topology and data traffic are characterized. Corresponding simulation environment is then built. Thereafter, various placement and replacement strategies are compared in a simulation study to be further able to propose a suitable cache placement and replacement combination for AMI in Smart Grid.

Keywords—caching; placement; replacement; content-centric networking; Named Data Networking; Advanced Metering Infrastructure; Smart Grid

I. INTRODUCTION

As a heterogeneous, distributed and large scale system, the Smart Grid Communication System (SGCS) is the subject of many research works trying to propose a well-tailored communication layer that guarantees smart grid requirements. Among many solutions, CCN is proposed to offer real time data transmission with inherent security levels and competitive latency. This research work belongs to a series of articles aiming at investigating content-centric networking adequacy for smart grids. In [1], CCN performances were compared to the Internet Protocol stack while managing smart grid communications. A deeper qualitative and quantitative analysis is pursued in [2] studying CCN support for renewable energy resources integration into SGs. The aforementioned works allowed us to take a position in favor of CCN as an eligible communication solution for this system. In this article, the focus is granted to caching as one of the most important building blocks of CCN. In fact, besides host decoupling and content-based routing, in-network caching is considered as one of the most relevant CCN features. The main goal is to study the cost effectiveness and the performance of this mechanism in smart grids. Advanced Metering Infrastructure (AMI), a service responsible for the generation of a huge amount of data such as power consumption and electrical parameters data, is particularly considered. This smart grid application can particularly avail of in-network caching since many data flows disseminate the same content to many nodes like price signals or maintenance commands. Previous work [3] proposes CCN for enabling AMI service but the focus is mostly on proposing a naming scheme to elaborate CCN-AMI. A relatively poor performance assessment is presented as only bandwidth consumption metric is studied. To the best of our knowledge, no other research work dealt with content-centric in-network caching impact on AMI performance.

This article leads off by an overview of the advanced metering infrastructure detailing this service and its major stakeholders. Afterward, part II exposes related works by exploring research works investigating CCN use in smart grids. Then content-centric networking is presented with a particular focus on its in-network caching feature in part III. Next, simulation environment and scenarios description is presented in part IV. The last part of this article exposes and analyses the results obtained by the conducted simulations.

II. STATE OF THE ART: CCN IN SMART GRIDS

Content-centric networking (CCN) in smart grids is being in the center of an active research effort where researchers and industrials are collaborating to propose a content-centric communication layer for smart grids. The first endeavor to enable SGCS using CCN was in [4] where a content-centric overlay network is deployed to exchange smart grid data traffic. It, however, used geographic routing preventing it to be considered as a pure content-centric solution. Then, the authors of [5] studied Information-Centric Networking (ICN) performance in Real Time State Estimation (RTSE) of smart grids. This work has been expanded in [6] and an information-centric platform baptized C-DAX is proposed to RTSE in active power distribution grids. Other research works dealt with different smart grid subsystem, we noticed particularly the adoption of ICN in home area networks for residential energy management [7]. ICN is also investigated to enable smart city services [8]; in particular vehicle-to-grid communications has been studied. The aforementioned research works analyzed ICN performance in selected smart grid subsystems but, despite its importance, no particular attention is paid to caching policies assessment. Due to its deep impact on networking performances, caching mechanisms requires thorough analysis. The present research work elaborates under the umbrella of the PASRI-MOBIDOC (www.pasri.tn) project in Tunisia funded by the European Union and WEVIOO (www.wevioo.com).
work aims at filling this gap by probing into in-network caching impact on AMI communications.

III. ADVANCED METERING INFRASTRUCTURE

A. Overview

Advanced metering infrastructure (AMI) is a smart grid potential application on the consumer side, responsible for metering services between electricity utility companies and their customers. It is an integrated system of smart meters, communication channels and meter data management systems. Many relevant smart grid subsystems like Advanced Distribution Infrastructure (ADI), Advanced Transmission Infrastructure (ATI) and Asset Management (AM) rely on the advanced metering infrastructure in order to properly achieve their functions. AMI has many benefits to both customers and power providers. On the consumer side, it allows him to be well informed regarding prices and services which provide him with the ability of managing its consumption patterns and costs. On the power utility side, AMI has an impact on two major services: the billing and Distribution/Transmission operations. Receiving time stamped power consumption information helps establishing an efficient billing system. In addition, customer information processing and mining may enable the utility offerings improvements. Moreover, accurate reporting of various electrical parameters is essential for the smart grid main purpose which is a reliable power transmission and distribution.

B. Actors, roles and interactions

Realizing the power grid modernization relies on several cooperating subsystems. Deploying an advanced metering infrastructure can be considered as one imperative step toward the smart grid since its outputs are the feeds of advanced distribution and transmission infrastructures [9]. Most relevant AMI functions are the following:

- Power data consumption readings,
- Electrical parameters reporting (phase, voltage, power factor),
- Time-based pricing and billing.

Achieving the aforementioned functionalities involves four main actors (see Fig.1):

- **Smart meters**: a device deployed at the customer premises responsible for periodic power consumption data readings. It also contributes at transmission and distribution operations by reporting some electrical parameters like phase, voltage, etc. Smart meters have to meet few requirements; communication skill can be considered as the most imperative one as it allows the device to achieve its major role. Smart meters must provide users by an ergonomic display feature. It also should support remote control commands to be executed by distribution management systems in order to maintain a reliable power grid.

- **Electrical properties sensors**: Voltage sensors (VS) and phase measurement units (PMU) are electrical field devices responsible for reporting real time electrical properties used to monitor the power distribution system.

- **Communication layer**: Each smart meter deployed in the smart grid has IN/OUT data traffic to receive or send. It has, in fact, to periodically send data about power consumption or electrical parameters status to remote collecting and processing nodes. In addition, it receives operational commands and price signals from operation centers. A communication layer managing all these data flows is then an important AMI subsystem. The richness of communication technologies landscape allows many design choices. Wireless and wired technologies can be combined to build a well tailored communication layer considering bandwidth and latency requirements at each level of the grid.

- **Data management system**: It is an integrated system for managing the huge amount of data received from customers at the utility company premises. Data collectors are first deployed in neighborhoods to collect data from smart meters in Home Area networks (HANs). It includes then meter data management system (MDMS) that can be seen as a database responsible for gathering and storing metering data from data collectors in neighborhood area networks (NANs). This system benefits, nowadays, from big data technologies and data centers expansion. The data management system also includes a billing system and a customer information system and has several interfaces with distribution operators such as distribution management system. Indeed this latter needs power quality information sent by smart devices deployed throughout the grid in order to monitor the power distribution system and trigger operational commands if needed.

To achieve advanced metering functions, the aforementioned actors need to interact and exchange data. The most relevant data flows are power consumption information sent periodically by smart meters and collected in data collectors deployed in neighborhoods. Power consumers receive also real time power pricing allowing them to have an
active role regarding their consumption patterns. AMI data traffic description provided in Table 1 was elaborated after consulting various research articles surveying smart grid services [10][11].

IV. CACHING IN CONTENT-CENTRIC NETWORKING

A. Content-centric networking in a nutshell

Till now, communication solutions proposed to manage computer networks are mostly based on addressing hosts and maintaining a bi-host connection used to exchange data. However, new communication patterns are content dissemination oriented which lead to the imperative need of a recast of existing networking paradigms. This fact has enriched the literature, since the early 2000s, by several research works [12][13][14] proposing the recast of the Internet and the adoption of new communication tenets. Although content-centric networking has first been elucidated in [15] to cater for the Internet recast, it has been explored by many researchers to grant required networking performances in diverse use cases such as social networking, vehicular networking, smart cities, etc. This attraction is due to many reasons. Building the whole communication process around the content unlike traditional paradigms where hosts are in the center of communications can be considered as the major attractive trait of CCN. Indeed we pay more attention on retrieving the desired content then knowing the host providing it. To do so content needs to be identified instead of addressing hosts this guaranteed by a naming strategy: the second pillar of CCN. A content-based routing protocol is also used to manage content travel from its producer to its consumer. In-network caching, an important building block of CCN, ardently seduces researchers as it promises enhanced networking performances.

B. In-network caching

CCN is characterized by in-network caching of data that improves network quality of services especially delivery latency. Satisfying data requests is not obligatory through locating the original data source but can also be done from multiple data stores. We find, in the literature, many caching policies that can be categorized in in-path caching or off-path caching. In the first one, only data replicas found along the path taken by the name resolution request are exploited. In contrast, the second one allows exploiting caches outside this path [17]. We can also differentiate caching policies based on router cooperation while placing content cache. Cooperative or non-cooperative strategies are mostly known [16].

In order to achieve in-network caching, placement and replacement algorithms need to be deployed. First of all, one needs to decide where to cache content; would it be in each node or in some selected ones. This decision leads us to the elaboration of a cache placement strategy like Leave Copy Everywhere (LCE) or Leave Copy Down (LCD), etc [18]. It is also common to place caches probabilistically or based on content proximity to its consumer [19]. A cache replacement strategy is also required since we wouldn’t like to exceed the cache size. Least Recently Used (LRU), Least Frequently Used (LFU) and First In First Out (FIFO) are the most usual replacement policies in the literature.

V. PERFORMANCE ANALYSIS OF CACHING FOR ADVANCED METERING INFRASTRUCTURE

A. Simulation tools

Being in the heart of an intensive research effort, content-centric networking is the subject of many research projects which lead us to many implementations of this paradigm. Among many, Named Data Networking (NDN) project (https://named-data.net/) is taking the lead by producing a large panel of deliverables such as core solution implementation, periodic technical reports, testbeds and simulation modules. In the present research work, we use NDNSIM, the NDN simulation module under NS-3, to simulate content-centric communications. Network simulator-3 (NS-3): https://www.nsnam.org/) is an event-based simulator widely used by the research community due to the richness of its libraries. We first built a networking topology according to figure 1 in order to reproduce AMI communication environment. Since no consensus has been made on the communication technologies to use while building the smart grid communication system, we have abstracted from the physical layer. In fact, the networking landscape offers many alternatives to enable smart grid communications; for instance home area networks (HANs) can exploit narrowband power line communication as wired technology or Bluetooth or Wi-Fi as wireless technology. The common aspect at this level of the grid is a low data rate. As for neighborhood area networks (NANs), cellular networks of broadband power line communications can be adopted. Finally the wide area network (WAN) can be achieved by optical fiber technology or fourth and fifth cellular network generations such as LTE [20]. Since assessing physical technologies performances in smart grids requires a dedicated research work, we excluded this task from our scope and therefore decided to use point to point links in NS-3 with well studied characteristics according to the requirements of each smart grid communication level.
TABLE I. AMI Data Flows

<table>
<thead>
<tr>
<th>Id</th>
<th>Source Actor</th>
<th>Exchanged Information</th>
<th>Destination Actor</th>
<th>Packet size (bytes)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Smart meter</td>
<td>Power consumption data</td>
<td>Meter data management system (MDMS)</td>
<td>200</td>
<td>Every hour</td>
</tr>
<tr>
<td>2</td>
<td>Billing system</td>
<td>Real time prices</td>
<td>Smart meter</td>
<td>210</td>
<td>Every 15 minutes</td>
</tr>
<tr>
<td>3</td>
<td>Distribution management system (DMS)</td>
<td>Operational commands</td>
<td>Smart meter</td>
<td>150</td>
<td>Triggered by events</td>
</tr>
<tr>
<td>4</td>
<td>Voltage sensor (smart meter)</td>
<td>Voltage, power factor</td>
<td>Distribution management system (DMS)</td>
<td>250</td>
<td>Every 5 minutes</td>
</tr>
<tr>
<td>5</td>
<td>Voltage sensor (transformer)</td>
<td>Voltage, power factor</td>
<td>Distribution management system (DMS)</td>
<td>250</td>
<td>Every 5 minutes</td>
</tr>
<tr>
<td>6</td>
<td>Phase measurement unit (transformer)</td>
<td>Voltage Phase</td>
<td>Distribution management system (DMS)</td>
<td>1536</td>
<td>Every 5 minutes</td>
</tr>
</tbody>
</table>

B. Simulation topology

To conduct simulations, the AMI topology has been built under NS-3 with the restriction of our scope to 5 Km radius area. Based on a prior research work[11] describing a smart grid communication infrastructure in Canada, it has been noticed that the communication infrastructure generally follows the electrical one which implies that the smallest branch of the adopted topology contains 10 smart meters according the north American norms [18]. In fact, within 5 km, 50 substations are deployed and 120 transformers are connected to each substation. Then, 10 homes will be served by each transformer leading to 10 smart meters wired to each one.

Particular nodes of the AMI application described in figure 1 have been added to our topology. Data collectors are deployed at neighborhood area network (NAN) while meter data management systems (MDMS) and distribution management system (DMS) were added at wide area network (WAN). To sum up the smallest branch of our topology counts a total of 14 nodes: 10 smart meters, 1 transformer, 1 substation, 1 MDMS and 1 DMS.

![Fig. 2. Downstream interest processing in NDN](image)

C. Simulation scenarios

Ndnsm used to simulate content-centric networking implements Named Data Networking [21] components under NS-3. It reproduces all its important building blocks: NDN core protocol, pending interest table (PIT), forwarding information base (FIB), Content Store (CS) and applications. The following scenario exposes their interaction to manage communications. From a content-centric point of view, exchanging content occurs in two steps: a publisher producing content and a subscriber expressing interest in receiving it. Each network node has three data structures: a content store used for in-network caching, a pending interest table used to store received interests if the node can’t satisfy them by a cache hit and a forwarding information base that stores possible forwarding paths to process incoming interests. When an interest is received, the Content Store (CS) is firstly checked. If the desired content is stored locally in the CS, then the node send it through the interest incoming face otherwise the Pending Interest Table is checked. If an entry with the same interest is found in the PIT, the interest incoming face is concatenated to existing faces and the interest itself is dropped. Interest existence in the PIT means that the node already received an interest in retrieving the same content and is waiting for a response. Once the desired content is received, it will be send to each face in the corresponding PIT entry. The final step is to check the Forwarding Information Base (FIB); it is carried out only if the interest does not exist in the PIT. Two scenarios are possible. Whether we find a corresponding entry for the interest which means that the node will forward it to the nodes found in the FIB and waits for them to send back the desired content. Or no entry is found in the FIB, this is the worst case since it means that the node has no idea how to find the corresponding content and will simply drop the interest (see figure 2). The upstream data processing is illustrated by figure 3. A node receiving a data packet first checks its content store to execute the caching policy. Of course whether to cache the content or not depends on the adopted caching placement strategy. For instance, if LCE (Leave Copy Everywhere) strategy is chosen, every node receiving a data packet appends it to its own CS. A replacement policy is also needed to define the way a CS is refreshed when the threshold size is reached. This policy describes which content to replace by the new incoming one; note that various replication and replacement strategies have been developed above in part III.

![Fig. 3. Upstream Data processing in NDN](image)

To assess in-network caching performances for AMI service, we deployed consumer and producer applications in the aforementioned topology nodes in order to simulate AMI data traffic described in Table 1. The first simulation scenario
aims at globally evaluating caching benefit for AMI. To do so, metrics were compared after enabling and disabling the caching functionality. Further simulation scenarios dig deeper into caching replacement policy and cache replication strategy evaluation. Simulations were run for 7200 seconds allowing to all data traffic to occur and in-network caching algorithms (placement and replacement) to take effect.

D. Simulation Results

1) In-network caching impact on global performances
The results of the first simulation show a reduced communication delay at each node of the topology and a reduced throughput (see fig 4). We remind that in this first simulation, the same AMI data traffic has been generated and two scenarios were compared; first with disabled in-network caching and second after enabling this feature. The delay represents the amount of time needed to satisfy an interest, it includes the queuing delay and the propagation delay. The reduction of communication delay is due to faster retrieval of data packets since closer nodes are satisfying data requests thanks to their content stores. The throughput is reduced significantly in comparison to that of the case without in-networking caching. Indeed less packets are travelling toward and backward source nodes since many interests are being served by content store hits.

2) Replacement strategy evaluation
After globally noticing in-network caching benefit while managing advanced metering communications, our goal is furthermore to investigate replacement strategy for AMI. Metrics evaluated at this level are:

- Hop count: the number of hops a data packet needs to travel in order to reach its requester.
- Cache hits: specifies the number of interests that were satisfied from the content store (CS).

Simulations were run using the same AMI topology and data traffic described in previous sections. Three replacement strategies are proposed by NDNSIM to insert new content packets once the content store size has been reached: FIFO (First In First Out), LRU (Least Recently Used) and LFU (Least Frequently Used). We vary cache size and notice that cache hits is proportional to cache size. It is a rational result since nodes with bigger cache size are able to store more data packets allowing, then, to satisfy more interests from CS. This fact is the reason behind the hop count decrease as the CS size increases. Indeed, as more interests are served from caches, shorter paths are taken by data packets to reach their requesters. Regarding the eviction strategies comparison, FIFO is having the lowest interest number satisfied from content stores (see fig 5). As for LRU and LFU, they are closely competing with a slight advantage for LFU strategy. On the other hand, the highest hop count reduction is observed with the LRU replacement strategy leaving FIFO and LFU behind. We remind that, if required, LRU removes content on recency of use basics; most recently used content is kept longer in CS. As for LFU, popularity of retrieval is what governs packet data eviction. We notice also, based on the curve allure, that the effectiveness of replacement strategies is more obvious when the content store (CS) size is beneath 50 packets. Thus, high content store size (beyond 50 packets) dissolves the replacement strategy impact on cache hits and hop count. Indeed, data packets won’t get evicted from the CS unless storage threshold reached.

3) Placement strategy evaluation
The last stage of our simulation study aims at assessing various placement strategies. Placement strategy refers to a set of rules deciding where to cache content. In our context, four placement policies have been considered:

Fig. 4. Caching impact on delay and throughput
Fig. 5. Cache replacement policy evaluation

- **LCE (Leave Copy Everywhere):** this strategy allows the insertion of content in every node along its path toward its requester. Each content packet will be then stored in the content store of all nodes crossed while traveling back to the consumer.

- **Probabilistic placement strategy:** A normal distribution decides whether a content is stored in the content store or not.

- **Probcache placement strategy [22]:** Content is cached with a probability depending on two factors: remaining caching capacity on the downloading path and cache weight depending on content closeness to its requester. Probcache is considered as a cooperative caching mechanism since caching decision depends on factors provided by all routers along the path.

- **Leave copy in none-constrained devices (LCNCD):** Since we are assessing caching for a particular smart grid service, we take into consideration, in this placement strategy, the fact that smart meters are constrained devices. In this last placement strategy we only deploy content stores in none-constrained devices which exclude smart meters from in-network caching.

Fig. 6 illustrates the results obtained after running the simulations described above with wider topology size. We varied branches number of the tree topology shown in fig.1 to reach a maximum of 5 data collectors and 50 smart meters. We observe a trivial impact of the cache placement strategy on the cache hits. The incidence on hop count is however more significant, especially with an important disadvantage of the probabilistic placement strategy that needs more hop number to satisfy interests. The most substantial result at this level is revealed by the superposition of LCE and LCNCD (Leave Copy in None-Constrained Devices) curves which imply that disabling the caching feature in the smart meters does not degrade the system performances. This result is due to the tree topology used in our simulations. Positioning smart meters at the edge of the network alters caching usefulness at their level. Probabilistic placement strategy is also less beneficial from delay point of view. Fig.7 shows, in fact, high packet transmission delays when normal probabilistic distribution governs cache placement. It strengthens, also, the fact that disabling cache in constrained field devices doesn’t affect network performance as the transmission delay is not influenced with LCNCD placement policy.

VI. CONCLUSION

Advanced metering infrastructure in smart grids is responsible for sensing, measuring, collecting and sending consumption data and electrical parameters. This key role brings up AMI design weight on smart grid effectiveness. In the meanwhile, content-centric networking with its ubiquitous in-network caching feature is increasingly gaining attention as the future networking trend. It’s premised that caching content on the delivery path can improve data delivery. Our goal was to investigate this assumption in handling AMI communication requirements. A simulation-based analysis reproducing AMI communication topology and data traffic showed reduced data delivery delay once in-network caching activated. It showed also inadequacy of FIFO replacement strategy for this context and better performances for content-popularity based strategies. Placement strategies analysis allowed us to show that disabling caching in smart meters does not affect the SGCS performance. It consolidates then our position on the adequacy of CCN to smart grids. This work concludes our series of articles investigating content-centric networking for smart grid communications. Our perspectives are the design of a content-centric smart grid communication infrastructure providing SGCS requirements such as latency, data transmission delay, quality of service and interoperability.
Fig. 6. Cache placement policy evaluation

REFERENCES


Fig. 7. Impact of placement strategy on transmission delay