F2c2C-DM: A Fog-to-cloudlet-to-Cloud Data Management Architecture in Smart City

Amir Sinaeepourfard *, John Krogstie *, Sobah Abbas Petersen *, and Dirk Ahlers

* Department of Computer Science
‡ Department of Architecture and Planning
*‡ Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
{a.sinaee, john.krogstie, sobah.a.petersen, dirk.ahlers}@ntnu.no

Abstract—Smart city solutions make a high-level technological innovation in a city to expand their citizens’ quality of life by several different technological management strategies (such as resource management and data management) among end-users, city planners, and technological devices (e.g., sensors, IoT devices, etc.). Currently, data management architectures have been offered by some researchers to organize obtained data in smart cities, including Centralized Data Management (CDM) and Distributed-to-Centralized Data Management (D2C-DM). In addition, the D2C-DM architecture can provide several advantages from the combined advantages of distributed (e.g., Fog) and centralized (e.g., Cloud) technologies, such as reducing network traffic and their latencies, upgrading the security levels and so on. In this paper, we propose several novel contributions. First, we design a D2C-DM with Fog, cloudlet, and Cloud technologies to organize huge amounts of the data production in smart city scenarios, from physical devices (including IoT and sensors devices) to non-physical devices (including third-party applications, and other databases). Second, we tailor our proposed D2C-DM architecture to the Zero Emission Neighborhoods (ZEN) scenario to manage its different data types, including context, research, and Key Performance Indicator (KPI) data. Finally, the advantages of this D2C-DM architecture are discussed.

Keywords— Smart City; Sensors Data; Data Management; Distributed-to-Centralized Data Management; Fog-to-Cloud Data Management; Fog-to-cloudlet-to-Cloud data management

I. INTRODUCTION AND MOTIVATION

Smart city solutions have been proposed by multiple researchers in recent decades to enhance their citizens’ quality of life by various technological management strategies (such as resource management and data management) between end-users, city planners, and technological devices (e.g., sensor) [1-3]. As a result, data management solutions play a key role in the smart city scenarios to organize all data produced (including sensors data, simulated data, etc.) by a vast number of the technological devices.

Traditional data management strategies in smart cities have been offered by centralized facilities based on Cloud computing technologies [2, 3] or partly decentralized or federated systems. There are several benefits which can be gained by using Cloud computing technologies, e.g., (almost) unlimited resources capacity [2-4]. However, there are several disadvantages through forwarding all data and services to Cloud technologies, including data quality and security issues, high communication latencies, etc. [5, 6]. Therefore, D2C-DM strategies are an emerging technology to contribute between the edge of networks (as a place that is the nearest place to the data sources) and centralized schema (as a place that is the furthest place to the data sources) through a unified strategy [7-10].

In this paper, first, we draw a novel D2C-DM architecture (from Fog-to-cloudlet-to-Cloud technologies) in the context of a smart city. The proposed architecture provides both the potential of distributed (such as Fog and cloudlet) and centralized (such as Cloud) technologies. Second, we tailor a proposed D2C-DM architecture to the ZEN center scenarios to manage their different data types, including context, research, and KPI data. Finally, the most obvious advantages of this D2C-DM architecture are explained.

The rest of the paper is structured as follows. Section II introduces the related work of different technological management strategies with a focus on data management strategies in the smart city scenarios. Section III describes the smart city scenarios and then discusses the main insights related to D2C-DM, including Fog-to-Cloud data management (F2C-DM) and cloudlet-to-Cloud data management (c2C-DM) architecture. Section IV describes particular smart city scenarios in Norway (ZEN center) and their related data management requirements. Section V tailors the smart city scenario to the ZEN center scenarios. Then, we design a Fog-to-cloudlet-to-Cloud data management (F2c2C-DM) architecture for smart cities. Section VI mentions the most obvious benefits of this D2C-DM architecture. Finally, Section VII concludes the paper and shows our future research map.

II. RELATED WORK

Smart city environments have been highlighted by several distinct technological management strategies (such as resource management and data management) between end-users to city planners and technological devices (e.g., sensor) to develop their citizens’ quality of life [1-6].

With focus on data management strategies, two main proposals of the data management architecture exist in the current literature of smart cities, CDM-DM and D2C-DM. The CDM-DM architectures suggested by several studies [2, 3], mention that data management schemes are centralized in one
place. This idea highlights that the data is produced from many different data sources spread across the city but data is stored and reached from a centralized platform, mostly using Cloud computing technologies [2, 3]. In addition, D2C-DM architectures are suggested by a few groups in the smart cities [7-10], which, uses a distributed schema for data management, using some high level of distributed technologies and ideas such as Fog Computing [6, 11] or cloudlet technologies [12]. The D2C-DM architecture has been used in both potentials of Cloud and distributed (such as Fog, cloudlet, etc.) technologies at the same time. However, almost all of D2C-DM architectures have been proposed through Fog and Cloud technologies (F2C-DM architecture) [7-10].

We note the following limitations in the existing data management strategies:

- Only a few proposals have been suggested for the D2C-DM architecture;
- Almost all of the D2C-DM architectures have been designed through Fog-to-Cloud technologies;
- Almost all of the D2C-DM architectures have been focused to cover the sensor data in smart cities; therefore, there are no proposed architectures to add other data types and formats to the D2C-DM architecture.

For all of the reasons mentioned above, we propose a hierarchical distributed F2c2C-DM architecture by using Fog, cloudlet, and Cloud technologies for smart city scenarios. This architecture can address all the above limitations. Eventually, this architecture can be fitted to the idea of the ZEN center.

### III. SMART CITY SCENARIO

Smart cities are discussed as one of the most essential and critical research areas nowadays. The smart cities concepts focus on the main changes in people's lifestyle from traditional to technological solutions. Technological solutions (e.g., smart energy) are being proposed to increase citizen's quality of life. Smart cities deal with the deployment of a network of devices (including IoT, mobile, and sensors devices) [2, 3, 13]. The network devices produce the IoT data sources in smart cities. In addition, IoT data sources can connect to different data sources in the smart cities, including third-party applications and other related databases. Therefore, traditional smart city architecture is divided into different layers. The layers are a sensing layer (data production), a network layer (data movement), a middleware layer (data management) and an application layer (products and services creation) [2, 3, 13].

In Fig. 1, the conceptual levels model around a neighborhood and the smart city is illustrated [14]. Fig. 1 showed that there are three primary levels in this model, including micro, meso, and macro. First, the micro level includes each element of the city (for instance, a building, a house, etc.). Second, the meso level covers the neighborhoods in the city (constituting buildings, houses, streets, and their neighbors). Lastly, the macro level consists of all neighborhoods in the city.

Data are such primary feeds for smart cities and play a vital role for the smart cities developments. The data provide a massive opportunity for a city to be smart and agile through the different types of services for a city. Moreover, the services use the data to build their service scenarios for the smart city stakeholders’ requirements by the appropriate information according to the contextual state, or some high-level knowledge discovered from different data analysis techniques. Therefore, we highlight that data management is one of the most important and challenging concepts in smart cities.

![Fig. 1. The conceptual levels model around a neighbourhood and smart city](image)

This section is organized into two main subsections. First, we describe the Data movement and types in smart cities (real-time, historical, and last-recent data). Second, we explain the idea of D2C-DM data management architectures in smart city.

#### A. Data Movements and Types

As we mentioned earlier, the technological architecture of smart city scenarios is addressed into four-tier architecture layers, consisting of the sensing (data production), the network (data movement), the middleware (data management), and the application layers (products and services creation) [2, 3, 13].

With the focus on the data production scenario, the sensing layer covers physical and non-physical data sources across the city. Those main data sources produce a large amount of data in the smart city. Therefore, as shown in Fig. 2, we first characterize the data produced by physical (e.g., sensors data) and non-physical (e.g., data extracted from social media or other third-party applications) data sources according to its age, ranging from real-time to historical data.

- **Real-time data**: The real-time data is produced in the nearest place to the end-users by physical data sources. It means that real-time data is used when produced basically in critical shallow latency applications. In addition, as shown in Fig. 2, the real-time data can be produced in the micro and meso level.

- **Historical data**: Data is considered as historical data (older data) as it is added and saved on data repositories. In this case, historical data can be considered to be a remote place from the end-users and data sources in the city because accessing data from the Cloud computing technologies (as an example of the centralized place technologies) has a high level of latency. Moreover, as shown in Fig. 2, the historical data can be stored at the macro level.
• **Last-recent data:** The last recent data is not close to the end-users and data sources, and not as far as the historical data is. Moreover, as shown in Fig. 2, the last-recent data can be generated in the meso level.

![Image](70x557 to 272x686)

**Fig. 2. Data types in smart cities through the conceptual levels model**

### B. Data Management

Smart city environments have been proposed by several distinct data management architectures between end-users to city planners and technological devices to produce useful services for their citizens [2, 3, 6, 11]. In this subsection, we focus on possible architectures for D2C-DM in the smart city.

This subsection is categorized into two main subsections. First, we explain the F2C-DM architecture for the smart cities. Then, we give some details about the c2C-DM architecture in smart cities scenarios.

1) **F2C-DM Architecture:** Fog computing technologies have emerged through Cisco [12]. Fog computing can contribute with devices in the edge of networks to make facilities for resources and data management in the Internet of Things (IoT) [6, 11, 15]. By this management plan, data should not be forwarded to a central place (usually in Cloud) [7-10]. Therefore, the network traffic and latencies can be optimized. Moreover, this F2C-DM architecture offers a hierarchical distributed architecture, where the Cloud technologies positioned in the architecture for deep storage and processing, and the Fog computing is used for the extension of the Cloud computing in the edge of the network [7-10].

In [7-10], the F2C-DM architectures are assumed by a variable number of layers from Fog to Cloud technologies with respect to the business and data model requirements. In addition, a basic data management architecture is suggested by a three-layered architecture. The Fog-layer-1 includes a set of Fog nodes (at the edge of the network) to perform the first level of processing and storage. The Fog-layer-2 is considered for second level of computing and storage with respect to the capabilities of the available devices in the city. Finally, the Cloud layer is the highest level in the F2C-DM architecture. The highest level of computing and storage are applied in this layer because the Cloud technologies have almost unlimited resources with the highest level of the latencies.

2) **c2C-DM Architecture:** Cloudlet technologies have been offered by CMU [12]. Fog and cloudlet technologies follow the same principals. However, the main difference is the capacity of resources in Fog and cloudlet technologies. On the one side, the Fog technologies use the potential of the data sources in the smart cities, such as traffic lights, smartphones and so on. On the other side, the cloudlet technologies prepare a “data center in a box” at the edge of the network. Therefore, in fact, this idea brings the Cloud services closer to the users and data sources [12].

There is no focus regarding data management through cloudlet technologies so far, but the c2C-DM can follow the same strategies of the F2C-DM as mentioned earlier in [7-10]. Consequently, the c2C-DM architecture may propose a different number of layers from data sources to Cloud technologies. As an example, the next section (Section V) will present our proposed F2cC-DM for the smart city scenario (as shown in Fig. 5).

### IV. ZEN CENTER SCENARIO

The ZEN center continues the idea by the European Union that all buildings should reach the “nearly zero energy” shortly in 2020 [16]. The ZEN center’s objectives scale from buildings to neighborhood for the “zero energy” idea. Therefore, the ZEN center focuses on zero emission neighborhoods in smart cities [16] but the ZEN center is focused strongly on the neighborhood level (including the constituting elements within it such as buildings, streets, transport systems, infrastructure, etc.). As we mentioned in Section III, Fig. 1 illustrates that the neighborhood is positioned at the meso-level between the micro and macro levels of building and city. Finally, the ZEN center covers eight pilot projects in different cities in Norway [16]. The locations of the pilot projects are Bode, Trondheim, Steinkjer, Evenstad, Elverum, Oslo, Barum and Bergen.

Data management and monitoring requirements play an important concept in the ZEN center. The data management and requirement tasks are defined under the work package 1 (Analytical framework for the design and planning of ZEN) of the ZEN center [16]. In addition, the “ZEN Data Management and Monitoring: Requirements and Architecture” report [14] mentioned that the ZEN center is looking for an efficient data management and monitoring for their pilot projects. Therefore, the data must be managed by transparent and reproducible solutions with sufficient data quality to meet monitoring, analysis, learning, and research requirements.

This section is organized into two main subsections. First, we show different data types in the ZEN center. Second, we explain our previous work to design the D2C data management architecture for the ZEN center concerning Fog and Cloud technologies.

#### A. Data Movements and Types

In the “ZEN Data Management and Monitoring: Requirements and Architecture” report [16], all different data types of ZEN center are categorized. Those data types are namely context, research, and KPI data as described below:

- **Context data:** Normally all these different data types have been generated by IoT and sensors devices or come from external systems or repositories, such as data of weather, environmental conditions, energy systems, urban planning data, etc. This type of data...
supports the interpretation of results and other data but is not of interest solely on their own.

- **Research data:** These data types have been generated by third-party applications, such as simulation or planning data, or sensor streams. This data which has been collected and used can come from pilots or entities/installations/buildings within pilot areas prototypes from live building and energy data from buildings. Some examples of these data types are occupant behavior data, energy data, etc.

- **KPI data:** A number of KPI in a set of themes have been defined by the ZEN center [16], including emissions, economy, spatial qualities for project monitoring. They come from surveys, simulations, and direct measurements.

To sum up, three data types are defined by the ZEN center, including context, research, and KPI data. As depicted in Fig. 3, we matched these three ZEN data types with the conceptual levels model (as shown in Fig. 1 [14]). Therefore, first, the context data is accessible on all levels (including micro, meso, and macro). Second, the research data can exist on meso and macro levels. Lastly, in a similar way to the context data, the KPI data is available in all the levels.

**B. Data Management**

In our previous work for the ZEN center [7, 8], we proposed the F2C-DM architecture based on the Fog and Cloud technologies, and discussed how the IoT data sources (from sensors) can be managed in the proposed architecture. In addition, we faced several additional data complexities in our previous proposed architecture [7, 8]. First, this architecture must cover eight different pilots and cities (in a specific size and location) in Norway. Then, the geographical distribution of the ZEN center scenario is larger than one smart city. Second, this architecture must deal with several distinct data sources (by types, formats etc.) of the ZEN center.

**V. Our Proposal: Tailoring Two Scenarios**

This section aims to tailor scenarios of smart cities and the ZEN center. To do this, the section is structured as two main subsections. First, we match the different data types in the smart city and ZEN center. Second, we design a novel hierarchal D2C-DM architecture based on Fog, cloudlet, and Cloud technologies for the ZEN center.

**A. Data Movements and Types**

As we discussed in the sections above about data types in smart cities and ZEN center, on the one hand, we mentioned that there are three main different data types in smart cities, including real-time data (physical data sources), last-recent data (non-physical data sources), and historical data (all data sources). On the other hand, the ZEN center has three main distinct data types, including context, research, and KPI data. Therefore, Fig. 4 tailored the smart city and ZEN center data types to the conceptual levels model as shown below:

- **Micro Level:** The real-time data produced near to the end-users in the city by physical data sources (such as sensors). In addition, the real-time data can be matched with the definition of the "context data" and "KPI data" for the ZEN center.

- **Meso Level:** The meso layer covers both real-time (physical data sources in the neighborhood), and last-recent (non-physical data sources) as details shown below:
  - **Real-time data:** there are a number of physical data sources in the neighborhood (such as vehicular network, traffic lights, etc.). Therefore, these data sources can generate the real-time data in this layer. In addition, the real-time data can be fitted with the “context data” and “KPI data” for the ZEN center.
  - **Last-recent data:** is located in the meso layer between macro and micro level. Moreover, the last-recent data can be tailored to “research data” and “KPI data” for the ZEN center.

- **Macro Level:** The macro level includes the historical data (all data sources) which is farther away to the end-users in the city. Moreover, the historical data can be tailored to the definition of the "context data", "research data", and "KPI data" for the ZEN center.

**B. Our Proposed F2c2C-DM Architecture**

As shown in Fig.5, we have designed a fully hierarchal distributed (from sensors and IoT devices) to centralized (Cloud computing technologies) data management architecture for smart cities. The main base of the proposed architecture is composed throughout the two main axes, Time and Location. These axes demonstrate our idea about data management in smart cities through the concept of the “data types”, “data management architecture”, and “technology layers”.

---

Fig. 3. Data types in ZEN center through the conceptual levels model

Fig. 4. Tailoring different data types of the smart cities and ZEN center
As depicted in Fig. 5, a hierarchy data management architecture (F2c2C-DM) includes the following three layers of architecture: (Fog-Layer, cloudlet-Layer, and Cloud layer) and each are described below.

- **Fog-Layer** is a nearby layer to the end-users and IoT devices in the smart city and ZEN center pilots. This layer covers various numbers of the IoT-Sources (such as sensors, smart-phones, etc.) in the Fog-Areas (constituting different types of buildings and their neighborhoods) and Fog-Device (the most powerful node for the processing and storage among the IoT-Sources). This layer can handle several duties for data management tasks as shown below.
  - **Technology solutions**: this layer uses the Fog technologies to handle the IoT data sources.
  - **Data Type**: Fog-Layer deals with the real-time data. In addition, this layer supports the “context” and “KPI” data for the ZEN center.
  - **Data Management architecture**: this layer is a part of the distributed data management architecture.

- **cloudlet-Layer** is a mid-layer between Fog and Cloud layers. In addition, this layer is located in the same city of the pilot, but it is not as near to IoT devices like the Fog-Layer. This layer connects all IoT data sources with other data types in the same city of the pilot, including simulated data, third-party applications and so on. This layer can manage different data management tasks as shown below.
  - **Technology solutions**: this layer handles the cloudlet technologies to manage all obtained data sources in the same city of the pilot.
  - **Data Type**: The layer supports the last-recent data. Moreover, this layer covers the “context”, “research”, and “KPI” data for the ZEN center.
  - **Data Management architecture**: this layer is a part of the distributed data management architecture.

- **Cloud layer** is in the grand position of the architecture. Cloud technology includes the most solid resources concerning processing and storage. This layer has several data management tasks as described below.
  - **Technology solutions**: this layer supports the Cloud technologies to organize all received data from everywhere.
  - **Data Type**: this layer saves the historical data and also supports all ZEN data types, including “context”, “research”, and “KPI” data.
  - **Data Management architecture**: this layer is considered as a CDM architecture.

As shown in Fig.5, the bottom-up architecture is designed based on from very small scale to very large scale of the city. In addition, the proposed F2c2C-DM architecture highlights that the bottom layer (Fog-Layer) across the city has the low latency of the network communications and the limited resources capacities in terms of the processing and storage facilities. Moving forward to the upper layer provides undesired level of the latency for the network communications and improves the processing and storage levels by upper level of the resources.

Many applications/services are envisaged for the ZEN center and their related pilots concerning their requirements. Then, the ZEN applications/services use a variety of data sources and types of our proposed architecture. As shown in Fig. 5, we proposed data according to its age, ranging from real-time to historical data and the data location (from bottom to up).

VI. ADVANTAGES OF THE NOVEL F2c2C-DM ARCHITECTURE

The most obvious benefit of this F2c2C-DM architecture is to combine both advantages of centralized and distributed data management in one unified data management architecture by using the Cloud, cloudlet, and the Fog computing technologies. This makes a high level of the computing and storage capacities from the Cloud layer, reduced network traffic, and communication latencies from the Fog layers. However, some additional benefits can be obtained from this hierarchical and distributed model, as listed below:

- The F2c2C-DM architecture can manage multiple data types and formats from physical and non-physical data sources in the smart city.
- Real-time and last-recent data accesses are much quicker than in a CDM architecture. Therefore, there are two times of the data transfer from edge of the network to Cloud technologies through the same path (send and receive data).
- By decreasing the data transmission length, the security and communication failure are improved.
- By having real-time data at distributed layers, the network load is decreased because we can create some applications at the edge of the network accessing these data locally.
- The D2C-DM provides huge facilities to set distinct types of policies (including appraising data quality, applying data security, updating frequency mechanism, etc.) in each cross-layer (from sensors to Cloud technologies) concerning city planner policies and business models.

VII. CONCLUSION AND FUTURE MAPS

The main contribution of this paper is to design an effective data management architecture for the ZEN center focused around the smart city context. We used the potential of D2C-DM, which is tailored to both advantages of centralized and distributed data management in one unified data management architecture (F2c2C-DM). Therefore, we present a novel F2c2C-DM architecture through multiple technologies (Fog, cloudlet, Cloud technologies) first. Then, we tailor this proposed data management architecture to the ZEN center scenarios to handle the ZEN data management complexities from data creation to data consumption. Finally, main benefits of the F2c2C-DM architecture are described (from supporting all data types to applying distinct policies).

As part of our future map, we will search for more options related to continue developing the proposed data management architecture concerning the ZEN center requirements.
ACKNOWLEDGMENT

This paper has been written within the Research Centre on Zero Emission Neighborhoods in smart cities (FME ZEN). The authors gratefully acknowledge the support from the ZEN partners and the Research Council of Norway.

REFERENCES


