CTNDCI: Identifying the Challenges Towards a distributed Nano Data Center Infrastructure

MELANIE HAUSER, LMU Munich, Germany
DIANA IRMSCHER, LMU Munich, Germany
KATRIN KOLB, LMU Munich, Germany
MENGCHU LI, LMU Munich, Germany
KATHARINA RUPP, LMU Munich, Germany
ANDREAS SCHOLZ, LMU Munich, Germany

In this paper we identify the challenges currently preventing nano data centers from becoming the dominant form of content provision on the internet. With the global increase in IP traffic the question of how to provide and deliver data is becoming increasingly important. Monolithic data centers, as they are used today, pose several problems, such as high energy consumption and lack of scalability. An alternative solution mitigating the problems of monolithic data centers has been proposed in the form of a distributed nano data center infrastructure. Research has shown this to be a superior solution. However, no widespread solution based on a nano data center infrastructure has been implemented as of yet. By identifying the main challenges nano data centers are facing steps can be taken to overcome these challenges in a more focused way, leading to a more economic data distribution.

CCS Concepts: • **Computer systems organization** → **Embedded systems**; *Redundancy*; Robotics; • **Networks** → Network reliability;

Additional Key Words and Phrases: Green IT; Nano data center; Energy consumption; Security; Availability; Scalability; Data distribution

1 INTRODUCTION

2 STATE OF THE ART

2.0.1 ECHOS. ECHOS introduces a concept for Nano Data Center that can or should completely replace monolithic data centers [5]. The authors call it a radical solution for data management and provision. According to this concept, so-called "boxes" are set up at the edges of the network, eg. in home gateways (see [6]). These boxes communicate with each other via a peer-to-peer system. The peer-to-peer system as well as the bandwidth is controlled by a central unit, such as the ISP. However, the approach of networking boxes via a peer-to-peer system, and thus providing or sharing content, requires some conditions. So it is first necessary to provide a distributed hosting edge infrastructure. Furthermore, there are still some problems added. In ECHOS these are listed as follows [5]:

- "Lack of service guarantees due to uncontrolled interface between different application [...]."
- "Inefficient use of network's and other peer's resources and consequently supoptiomal performance [...]."
- "Even if sufficient status information is in place, still P2P is inherently unable to use it as it was designed around selfish user behavior and free-riding prevention mechanim [...]."
- "Absence of security and control make it impossible to guarantee the integrity and security of content."

Authors' addresses: Melanie Hauser, LMU Munich, Munich, Germany, melanie.hauser@campus.lmu.de; Diana Irmscher, LMU Munich, Munich, Germany, d.irmscher@campus.lmu.de; Katrin Kolb, LMU Munich, Munich, Germany, katrin.kolb@campus.lmu.de; Mengchu Li, LMU Munich, Munich, Germany, mengchu.li@yahoo.com; Katharina Rupp, LMU Munich, Munich, Germany, katharina.rupp@web.de; Andreas Scholz, LMU Munich, Munich, Germany, andreas.scholz@campus.lmu.de.

3 CHALLENGES TOWARDS A DISTRIBUTED NANO DATA CENTER INFRASTRUCTURE

3.1 Political challenges

3.2 Technical challenges

3.2.1 Energy Consumption Model for Nano Data Centers. While nano data centers are motivated by the energy consumption problem [7], current research reveals that the advantage of the nano data centers in energy efficiency relies on certain technical foundations. To give an overview of the technical challenges towards nano data center development, we first need to figure out the energy consumers of the nano data center applications.

Considering the nano data center (NaDa) platform proposed in [7], users will host tiny managed nano servers on their end-user devices such as Triple-Play gateways and DSL/cable modems, and communicate with each other following a Peer-to-Peer (P2P) philosophy. Thus, suppose a user (client) wants to access the content stored in a nano server hosted by another user (manager), the energy consumption for this process consists of [4]:

- the energy consumed by the client for requesting the content, denoted as E_{req};
- the energy consumption of the content transportation process, denoted as E_{trans}; and
- the energy consumed by the manager for storing the content and processing the request, denoted as E_{serv}.

If we denote the total energy consumption as E_{total}, we can derive the following formula:

$$E_{\text{total}} = E_{\text{req}} + E_{\text{trans}} + E_{\text{serv}}. \tag{1}$$

To detail the formula from technical aspects, we need to understand the internet protocol (IP) network. [1] modeled the IP network as the combination of three domains: the access network, the metropolitan and edge networks and the core network. For centralized data center applications, [1] visualized the IP network model as shown in Figure 1. The access network connects each end-user to the metropolitan and edge network, which serves as the interfaces to the core network. Centralized data centers are usually directly connected with the core network, but for nano data centers, since nano servers are hosted in end-user devices, the data has to traverse the access network twice [3].

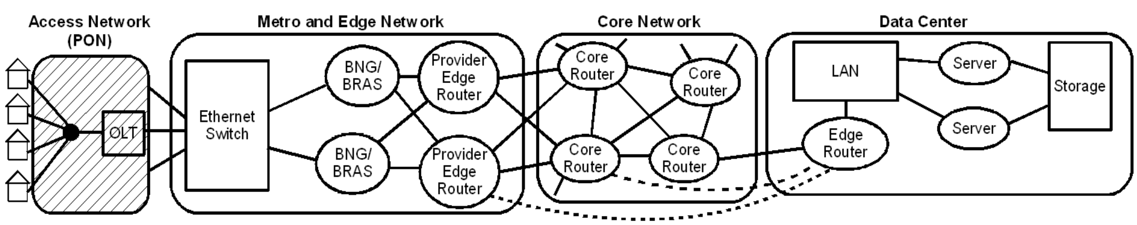

Fig. 1. IPTV network model for centralized data centers [1]

We can now detail formula (1) and adapt it to the energy consumption model proposed in [4] as the following:

$$E_{\text{req}} = E_c + E_{\text{access}}, \tag{2}$$

$$E_{\text{trans}} = E_{\text{edge}} \cdot h_{\text{edge}} + E_{\text{core}} \cdot h_{\text{core}}, \tag{3}$$

$$E_{\text{serv}} = E_{\text{access}} + E_m, \tag{4}$$

where E_c represents the energy consumed in the end-user device of the client, E_{access}, E_{edge}, and E_{core} represent the energy consumed in the access network, edge network and core network, respectively, h_{edge} and h_{core} represent the number of hops in the edge and core networks, and E_m represents the energy consumed in the end-user device of the manager.

In the following, we will introduce four technical challenges that we regard as fundamental and argument our selections with respect to the proposed energy consumption model. These challenges are: the activation of nano servers; the selection of the access network that the nano servers are attached to; the location of nano servers; and the number of data replication strategy.

3.2.2 Activation of Nano Servers. A nano data center platform is constructed with end-user devices as nano servers. We refer a nano server as *active*, when it is on and the user is accessing the nano data center service, and we refer a nano server as *idle*, when it is on but the user is not accessing the nano data center service. Whenever a nano server is on, no matters it is active or idle, it consumes energy. [7] proposed a thorough study on how the activation status of nano servers affects the total energy consumption. [7] denoted the active times of a nano server as t_{act} and the idle time of the nano server as t_{idle}, and introduced a coefficient R that represents the ratio of the active time of nano servers to the whole duration when nano servers are on:

$$R = \frac{t_{\text{act}}}{t_{\text{act}} + t_{\text{idle}}}. \tag{5}$$

As proposed in [7], R is involved in the calculation of both E_c and E_m in the energy consumption model. As a result, the energy consumption of nano data centers correlates with R as shown in Figure 2, where five different active time ratios (0.01, 0.05, 0.2, 0.5, 1) are chosen for comparison, and the energy consumption of a centralized data center is also shown as a reference. We can see that the energy consumption of nano servers increases as the active time ratio decreases, and when the ratio is large than 0.2, the energy consumption of nano servers surpasses the energy consumption of the centralized data center, i.e. the nano data center becomes less energy efficient.

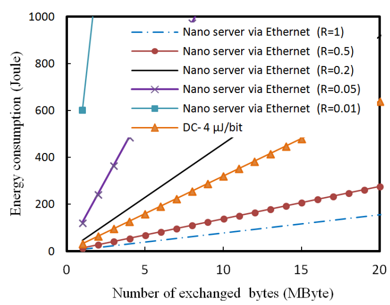

Fig. 2. Energy consumption of nano servers with different active time ratio, and of a centralized data center [7]

.

A disappointing fact is that we cannot simply assume the active time ratio R to be higher than 0.2 most of the time. Taking the widely-used video delivered by Internet protocol (IPTV) as an example: according to [7] and [2], the IPTV user activity shows large variation throughout the day, as shown in Figure 3. Even the in the peak hour, fewer than 20% of customers are active (i.e. $R < 0.2$); as for in the midnight, less than 5% of customers are active (i.e. $R < 0.05$). On average, the active ratio R is around 0.07, which means that if all nano servers are on in the whole day, it is energy **in**efficient to apply nano data centers.

Therefore, to make a nano data center platform energy-efficient, it is necessary to increase the active time ratio. From our view, a possible solution is to turn off the nano server time to time to reduce the idle time of the nano servers. But for the nano servers attached to DSL-modems or other functional devices, such as the NaDa model proposed in [7], turning off nano servers can lead to side-effects on the usage of other internet services of users, and thus the implementation of this solution requires further study. Another possible solution is to modify

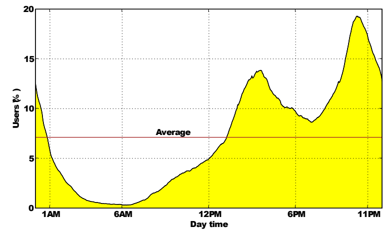

Fig. 3. User activity of the IPTV service [7]

.

the applications for the nano data center, such as run multiple applications that have different peak hours, aiming to maximize the active time of the nano servers. So far we are unaware of any research tackling these problems, and thus we list the activation of nano servers as one of the four technical challenges towards nano data center development.

3.2.3 Access Network.

3.2.4 Location of Nano Servers.

3.2.5 Data Replication.

4 EVALUATION/RESULTS

includes research results and interview results etc.

5 DISCUSSION

6 ACHIEVEMENTS

7 CONCLUSION AND FUTURE WORK

ACKNOWLEDGMENTS

REFERENCES

[1] Jayant Baliga, Robert Ayre, Kerry Hinton, and Rodney S. Tucker. 2011. Architectures for Energy-Efficient IPTV Networks. *IEEE Systems Journal.* 5 (09 2011), 518–527. https://doi.org/10.1109/JSYST.2011.2165601

[2] Meeyoung Cha, Pablo Rodriguez, Jon Crowcroft, Sue Moon, and Xavier Amatriain. 2008. Watching television over an IP network. ACM, 71–48. https://doi.org/10.1145/1452520.1452529

[3] Anja Feldmann, Andreas Gladisch, Mario Kind, Christoph Lange, Georgios Smaragdakis, and Fritz-Joachim Westphal. 2010. Energy Trade-offs among Content Delivery Architectures. IEEE. https://doi.org/10.1109/CTTE.2010.5557700

[4] Fatemeh Jalali, Robert Ayre, Arun Vishwanath, Kerry Hinton, Tansu Alpcan, and Rodney S. Tucker. 2014. Energy Consumption of Content Distribution from Nano Data Centers versus Centralized Data Centers. *SIGMETRICS Performance Evaluation Review* 42 (2014), 49–54.

[5] P. Rodriguez N. Laoutaris and L. Massoulie. 2008. ECHOS: Edge Capacity Hosting Overlays of Nano Data Centers. *SIGCOMM Comput. Commun. Rev.* 38 (01 2008), 51–54. https://doi.org/10.1145/1341431.1341442

[6] Technicolor. 2011. Technicolor Research & Innovation - Nano Data Center. Online video. (01 2011). https://www.youtube.com/watch?v=uvkBUKqFbkI

[7] Vytautas Valancius, Nikolaos Laoutaris, Laurent Massoulié, Christophe Diot, and Pablo Rodriguez. 2009. Greening the internet with nano data centers. In *CoNEXT*. ACM, 37–48.

A QUESTIONNAIRE

B INTERVIEW