How Can Architecture Help to Reduce Energy Consumption in Data Center Networking?

László Gyarmati  
High Speed Networks Laboratory  
Dept. of Telecommunication and Media  
Informatics  
Budapest Univ. of Technology and Economics  
gyarmati@tmit.bme.hu

Tuan Anh Trinh  
High Speed Networks Laboratory  
Dept. of Telecommunication and Media  
Informatics  
Budapest Univ. of Technology and Economics  
trinh@tmit.bme.hu

ABSTRACT
Reducing the power consumption of data centers has recently been received much attention from the research community and the industry alike. It is because the benefits are manifold: for the data centers themselves and towards the environment. A number of issues concerning with the consumption of the servers or the switches have been investigated. However, the energy requirement of the data center architecture has not yet been sufficiently addressed. In this paper, we address this issue and make two key contributions. Before delving into the details of possible solutions for reducing power of existing data center architectures, we first characterize the impact of architectural parameters on the power consumption of data centers by presenting preliminary results on the power consumption of state-of-the-art data center structures, namely BCube [8], DCell [9], and Fat-tree [1, 11, 7]. Secondly, based on the insights and lessons learnt from the analysis, we present our vision on possible practical solutions to reduce the energy usage of data center networking.

Categories and Subject Descriptors
C.2.1 [Network Architecture and Design]: Network communications; C.4 [Performance of Systems]: Design studies

General Terms
Design, Reliability, Management

Keywords
data center, power consumption, energy-efficient architecture

1. INTRODUCTION
Data center networking has received a special attention from the networking research community in the recent years.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee.

e-Energy '10, Apr 13-15, 2010, Passau, Germany
© ACM 2010 ISBN: 978-1-4503-0422-1/10/04...

The revolutionary network-based services, e.g., cloud computing, online social networks, video broadcasting services, of the last decade required the development of special network architectures capable to interconnect tens or even hundreds of thousands of servers with high inter-server connectivity.

The key goal was to design a network structure that is agile, resilience, and most importantly provides as much throughput as possible. The overall properties of the architectures is an effect of the structure of the network, i.e., how the servers are connected to switches, the deployed devices, e.g., how fast the switching fabric can operate, and the routing protocol applied throughout the system.

Moreover, there are two major design paradigms: creating a data center in a centralized or distributed way. Centralized data centers, on the one hand, are designed to be structurally resilient in terms of power, cooling, and network connections; the data center operator may have only a single site to provide its services from. These data centers are called more often mega data centers referring to their size. On the other hand, a distributed data center consists of a number of small, or medium size data centers at multiple locations, usually they are distributed geographically as well. The design of these data centers are simplified as there is no need to be the infrastructure resilience; failures are handled using replication mechanisms implemented in software.

In order to propose possible solution, which decreases the power consumption of the data centers, first it is compulsory to understand the power consumption of current data center architectures, which structural factors have an impact on the energy consumption. Although the designs of data center networks as well as their costs have been addressed recently, the energy consumption of the data center architectures has not been analyzed in details. Therefore, in this paper, we devote special attention to the energy requirement of the structure of data centers by investigating the power consumption of the state-of-the-art data center architectures. The paper tries to reveal the trade-off between the throughput and the power consumption of data center designs and motivate system architects to create more energy efficient data center structures by proposing some promising design patterns.

The power consumption of the data centers have been identified as an amortized cost, which is 15 percent of the total cost of a data center [6]. Moreover, energy related costs are faced in terms of infrastructure costs as well; providing reliable power supply to a large-scale data center is
expensive.
We deal with the structural power consumption of data centers, however, there exist several other parts of data centers where more energy efficient solutions can be reached, including reducing the power consumption and heat dissipation of servers and switches [10]. Energy-efficient replication strategies may reduce the internal traffic of data centers, and resource allocation strategies that allow power-off currently underutilized parts of data center throughout virtualization techniques [2].

The paper is structured as follows. First, we review the recent data center architectures. Afterwards, we present several metrics that describe the provided connectivity and the energy consumption of the structures. Then, we quantify these properties based on simulation results. Based on the insights of the simulation results we discuss our vision on possible solutions for reducing the power consumption of the architectures.

2. DATA CENTER ARCHITECTURES AND PROPERTIES

The DCell data center architecture [9] is created out of commodity mini-switches to scale-out. The construction of DCell is recursive. The smallest building block is called DCell which consists of an n port switch and n servers, which are connected to the switch. The higher level DCell structures are made out of gk = tk−1 + 1 DCellk−1s, while DCell0 has tk = gk tk−1 servers. These recursive expressions scale up rapidly, therefore, DCell can have enormous number of servers with small structural levels (k) and switch ports (n). The DCell structure contains switches only in the lowest hierarchy level, therefore, the servers are actively involved in the routing process; each server has k + 1 ports.

The BCube data center architecture [8] is designed to be applied in container based, modular data centers, which have a few thousands of servers. BCube is defined recursively, a BCubek is structured from n BCube1−k1 and n k n-port switches. This redundant structure can have n k+1 servers, there are multiple edge-disjoint paths between any two servers; therefore, BCube is able to distribute the load efficiently.

One of the design principles of fat-tree topologies [1, 11, 7] is to build a data center network using small, commodity switches. The fat-tree topology, also known as Clos topology, consists of three structural layers. At the top, the core level contains (n/2)2 n-port switches. The medium layer has n pods, each containing two layers of n/2 switches, i.e., one pod has n n-port switches. The higher level switches are connected to the core switches, while the lower switches have connection to the servers; a single pod contains n hosts. Accordingly, a fat-tree topology supports (n3/4) servers.

As a reference, we incorporate into our analyzes the balanced tree structure as well. A balanced tree distributes its leaves as evenly as possible between its branches. A balanced tree has a single switch in the core, which has n ports, similarly, the switches in the intermediate levels has n ports too. The servers are located at the leaves, connecting to one of the switches. If the balanced tree has k levels the architecture can have n k servers.

We illustrate the abovementioned data center structures in Figure 1. The servers are shown with rectangles, while switches are presented with circles.

Next, we present metrics based on which we investigate the properties of the introduced data center structures. We refer to the sum of the energy consumed by the parts of the data center fabric as the total power consumption, measured in Watts. In particular, the total power consumption incorporates the energy requirement of the switches and the extra energy consumed at the servers that have multiple ports. However, we do not consider the power requirement of the servers and the devices that are related to other operational devices, e.g., the cooling facility. Thus, the power consumption of the fabric denotes the net energy consumption of the data center architecture.

In terms of connectivity, the data center structures can be characterized based on diameter, average shortest path length, maximal betweenness centrality, and maximal edge betweenness; the last two metrics suggest the load balancing capability of the data center architecture. We are aware of the fact that each data center architecture has its own, specialized routing algorithm, however, we think these general properties reflect the transmission capabilities of the architectures. We express analytically the power consumption and the diameter of the structures in Table 1; the energy consumptions of a single server and switch are denoted as E and Esw. In all cases, the power consumption of the architecture highly depends on the parameters of the architecture, namely the number of ports, denoted by n, that a server or switch can have and the number of structural levels, denoted by k.

3. SIMULATION RESULTS

We carried out simulations in order to investigate the trade-off between the power consumption and transmission capabilities of data center architectures; we present the results in this section. The simulator is implemented in Python; we make the simulator available to the research community [5]. Several switching fabric are used in the simulations; to be realistic, we applied power consumption values of switches currently available on the market. In particular, the power consumptions of the 8-port (2960-STC-L), 24-port (2960-24TC-L), and 48-port (2960-48TC-L) Cisco switches are 12W, 27W, and 39W [3], while the 5-port (DGS-2205) D-Link switch consumes 5.12W [4]. We assumed that the power consumption of a server port is 3W.

First, we show the energy requirement of several data center architectures in Figure 2 as a function of the number of servers in the structure. The figures present the same values, however, on a different scale; thus, structures with less than 1000 servers can also be analyzed. All of the data center architectures (DCell, BCube, Fat-tree, and balanced tree) are generated with exactly the same number of servers that the structure enables with the given parameters. Accordingly, the presented results are the best-case scenarios, e.g.,

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Power consumption</th>
<th>Diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balanced tree</td>
<td>Enk + Es + Esw ∑(i=0 to k−1) n i</td>
<td>2k</td>
</tr>
<tr>
<td>Fat-tree</td>
<td>E(n/2)2 + n2</td>
<td>n</td>
</tr>
<tr>
<td>DCell</td>
<td>(E + Es + Esw)(n + 1)4</td>
<td>2k+1 − 1</td>
</tr>
<tr>
<td>BCube</td>
<td>Enk+1 + Es ∑(i=0 to k−1) n i</td>
<td>k+1</td>
</tr>
</tbody>
</table>

Table 1: Power consumption and diameter of data center architectures (n ports, k levels)
structures with fewer servers require almost the same energy than the presented ones, except the power consumption of the ports of the servers. Not surprisingly, the balanced tree structure consumes the less power for a given number of servers, however, the transmission capabilities of balanced tree structures are limited, i.e. the server in the root would be a throughput bottleneck in a real-world application.

In small-size data centers, BCell and DCube have roughly the same energy requirements. As the number of the servers increases DCell consumes less power than BCube. For example, the DCell_{4}(8) has 5256 servers and consumes 55.2 kW while BCube_{3}(8) has a higher power consumption (73.7 kW) with less servers (4096).

The Fat-tree data center structure stands between DCell and BCube in terms of power consumption, e.g., a Fat-tree with 24-port switches provides connectivity to 3456 server consuming 29.8 kW energy. Fat-tree architecture has only one parameter, the number of ports of the switches that, on the one hand, results a simple design but, on the other hand, limits the granularity of the structure, e.g., the 48-port architecture provides transmission capabilities up to 27648 servers.

The relation between the power consumption and the transmission capabilities are presented in more details in Figure 3. We illustrated several different data center architectures with exactly the same number of servers (1000). The values next to the illustrating shapes refer to the structural parameters (n,k) of the data center architectures.

In terms of the average shortest path of the structures (Fig. 3(a)), architectures with high port number switches have the smallest paths; in addition, they have also moderate power consumption. DCell provides longer paths than the other architectures in the presented scenarios; however, the energy requirement of DCell is much less than the others. In case of BCube, the average shortest path length increases linearly as more power is consumed, however, the maximal load of a server is much lower in this case (Fig. 3(b)). We only present a single fat-tree structure in the figures, as the power consumption of the 48-port fat-tree architecture is significantly higher (102 kW) than the others. This increased power consumption is not justified with the transmission capabilities, as the average length of the shortest paths is 3.65.

The maximal betweenness centrality values of the data center networks are shown in Figure 3(b). A data center architecture has smaller throughput bottleneck if the value of this metric is small. Accordingly, despite the small power consumption of the DCell structures, this structure may not be always an optimal structure due to the high load on a single point, e.g., applications with large throughput require-ment may have bottleneck problems. In case of BCube, the transmission capability of the architecture does not improve, after 20 kWs, even though the power consumption of the structure is increased. The Fat-tree architecture has favorable throughput bottleneck properties, i.e. both the servers and the links among them are used evenly.

The load of the link is expressed with the maximal edge betweenness of the structure (Fig. 3(c)). In case of DCell, the decrease of power consumption, which means fewer hierarchy levels, decreases the load on the link. This is not the case at BCube, where the power consumption decline does not affect significantly the load on the edges. There is an order of magnitude difference between the edge betweenness values of the architectures—as the vertical axis is logarithmic—which refers to diverse structural properties.

4. LESSONS LEARNT AND OUR VISION

Current data center architectures have only several parameters, which influence the structure and the number of servers in it; therefore, the size of the data center cannot be set on a fine-grained scale. Accordingly, the energy consumption of data centers, which are designed for larger number of servers, is much larger than it would be necessary. This phenomenon was shown in the simulation results, where architectures capable to interconnect much more servers than the analyzed 1000-server case required more power.

Our vision is based on the insights learned from the previous part, where we compared the power consumption of existing data center approaches. One possible solution is to design data center architectures where additional parameters are included, e.g., the data centers may include different type of switching fabrics in terms of the number of ports. The resulting heterogeneous data center architectures would consume only as much power as required for the given data center size.

The design pattern of recent data center architectures stick to symmetrical structures, which results an artificial constrain on the data center design. Symmetrical network architectures can be implemented in small sized networks, however, if a network has thousands of vertices the symmetry causes network over-provisioning, i.e., the structure may provide connectivity for more servers than required. The biological networks such as metabolic, protein structure and signaling networks are asymmetric, which may also warns us that symmetric architectures may not always be the best possible solutions. Asymmetric parts in data centers include increasing number of servers in small portions, not every server has a direct Internet connection. In our vision, the data centers of the future have asymmetric structures, they are more likely meshed networks than symmetric, tree-like.
architectures.

5. CONCLUSIONS

In this paper, we have addressed the issue of power consumption of data center architectures and proposed possible future design paradigms. We have shown that the power consumption of the structures shows diverse characteristics similarly to the transmission capabilities. We presented simulation results that support our initial assumption that there exists a trade-off between the power requirement and the transmission capabilities. In our vision, the future data center architectures will be designed using heterogeneous switching devices in an asymmetric way based on the insights learned from the simulation results. We hope our paper will initiate a discussion about the power consumption of data center architectures and drive research effort to design energy efficient data center structures.

6. ACKNOWLEDGEMENT

This paper has been partially supported by HSNLab, Budapest University of Technology and Economics, http://www.hsnlab.hu.

7. REFERENCES


