Energy-Aware Routing based on Power Profile of Devices in Data Center Networks using SDN

Tran Manh Nam, Nguyen Huu Thanh, Ngo Quynh Thu and Hoang Trung Hieu
Hanoi University of Science and Technology
Hanoi, Vietnam

Stefan Covaci
Technical University of Berlin
Berlin, Germany

***Abstract* — ICT is currently responsible for 4% of the EU's carbon emission. In this industry, data centers are growing quickly to supply the escalating services and application demands. The design of high performance and energy-efficient data center networks is an important issue while still being confronted with many challenges such as scalability, reliability, etc. On the other hand, the network devices' difference in power consumption as well as property make the researching process more difficult to construct a routing algorithm that effectively works with various devices. This paper presents a novel energy-saving scheme that can flexibly control and route traffic relying on the difference of network devices' energy-profile. By using Open Flow mechanism and Net-FPGA cards based on Software Defined Networks, we successfully deploy energy-aware data center network, and its experimental results show that the novel scheme improves the power saving level and energy-efficiently works based on the switches' energy-profile.**

Keywords—Energy-profiling; OpenFlow; Data center network; Fat-tree; Software defined network; Net-FPGA;

I. INTRODUCTION

Nowadays, the Internet services are booming with diverse networked applications such as social networks, multimedia softwares as well as cloud computing services. Besides, today's networks are mainly configured statically using a fully constant performance. So, many notable redundant devices including routers, switches, and links are required to maintain the systems with high availability and reliability. As a result, the systems constantly consume large amount of power even in the absence of traffic demand. Consequently, the energy consumption of data centers has caused many problems such as power cost and carbon emission. The total yielding electricity use by data centers in 2010 occupied about 1.3% of total electricity used in the world, and 2% of total electricity used in the United State [1]. Emissions from the ICT sector are estimated to keep rising significantly over the coming years – from 0.53 billion tons (Gt) carbon dioxide equivalent (CO2e) in 2002 to 1.43 Gt CO2e in 2020 under BAU growth [2].

It is extremely necessary to design a data center with high performance and energy-efficient networking system. It requires a new flexible and reconfigurable network architecture that helps a data center network to be more energy-aware. Software-Defined Networking (SDN) [3, 4] is an emerging networking paradigm that can help us make the network satisfy the above requirements by decoupling the data plane and control plan.

978-1-4799-7961-5/15/$31.00 ©2015 IEEE

In this paper, the main contribution of our work is to construct the energy-aware network as well as a novel energy-saving scheme by using OpenFlow, a software defined networking technology. The scheme makes use of the energy-profile of network devices and switchport's power consumption under various link rates. By adding network energy-profiles to the topology optimization and routing processes, the scheme can save up to 11% energy consumption in comparison to *power scaling* approach [9,12]. The energy-profiling-based energy efficient method also makes our scheme flexibly work with heterogeneous devices of different vendors.

This article is organized as following: Section II describes the previous works related to the authors' finding; Section III presents the problem statements; Section IV describes our energy-saving scheme; and section V shows the performance evaluation. Finally, conclusions and references are drawn in section VI and VII.

II. RELATED WORK

A. Software defined network

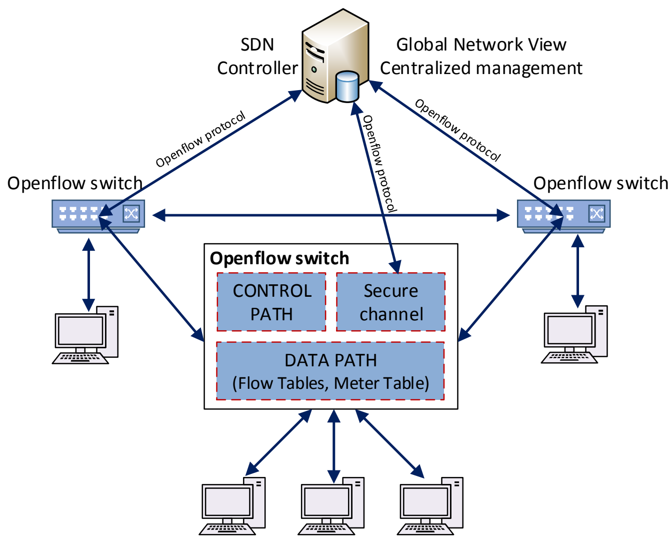

Fig. 1. OpenFlow controller and switches

The energy-efficient approaches must be deployed on the energy-aware data center networks (DCN) that require programmable network architecture. Software Defined Network (SDN) is a network architecture that is suitable for the requirements of energy-aware DCN. The architecture provides monitoring and controlling capabilities such as centralized management, energy consumption monitoring and traffic prediction. In comparison with traditional architecture, SDN decouples the control plane and data plane of networking devices and enables the control plane to become directly programmable. The control plane is placed in SDN controller that maintains the global view of a network and controls the network's activity as a single, local networking device.

The OpenFlow [5] component which is the fundamental element in the SDN architecture gives secure access to the data plane of devices such as routers or switches from remote controller over the network. The switches and controller communicate via the defined OpenFlow protocol messages such as *packet-received, send-packet-out, modify-forwarding-table,* and *get-stats.*

B. Data Center Network

1) Fat-tree topology

In our work, we build a DCN with the fat-tree topology [6, 7] that has three layers: core, aggregation, and edge. The main benefit of this topology is decreasing in the oversubscription ratio, removing the single point of failure of the hierarchical architecture and reducing a construction cost.

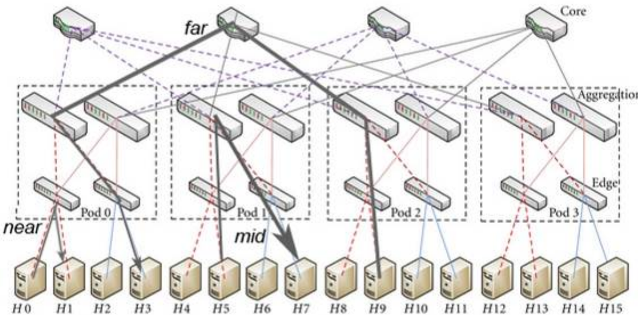

Fig. 2. K=4 fat-tree topology with different traffic scenarios

A k fat-tree is architecture of DCN using similar k-port switches. There are k PODs (Performance Optimized Data Centers), each POD contains $k/2$ aggregation switches and $k/2$ edge switches. The number of core switches is $k^2/4$ and each core switch has one link to each POD. In this topology, each access switch is connected to $k/2$ servers. Appropriately, a fat-tree with k PODs uses $5k^2/4$ k-port switches and can support $k^3/4$ servers. In order to increase the reliability of a data center network, we maintain the minimum working rate of all network devices. In this case, the connections among all switches always remain in minimum speed 10Mbps and the connections among switches and server constantly operate maximum speed 1Gbps.

2) Traffic scenarios: Near, Middle, Far

We implement our novel method in the fat-tree topology under four traffic scenarios: near, middle, far and mix. In near traffic scenario, the source and destination of any flow are connected to the same edge switch (Fig.2), the exchanged traffic traverses over only edge switch. In the middle traffic scenario, a source and destination pair of any flow resides in the same POD (Fig.2), so in this scenario all flows traverse over edge and aggregation switches. In the far traffic scenario, the source and destination of every flow residing in different PODs (Fig.2). Therefore, in this situation, all flows traverse over edge, aggregation and core switches. The mix traffic scenario is the combination of three above scenarios. In this work we regularly use the mix traffic scenario.

3) Network architecture

We deploy the Data Center Network that uses a model (Fig.3) extended from the Elastic-tree model [8]. The DCN consists of three main modules: *optimizer, routing* and *power control*. The *optimizer module*, with its input information such as traffic flows, topology and energy-profile of switch, must find the most energy-efficient subnet that satisfies current traffic demand. After this calculation, the optimizer module outputs the active topology to the routing module and power control module. Afterwards, the power control module changes the power states of switches, line cards, and interfaces whereas the routing module chooses the paths for all flows.

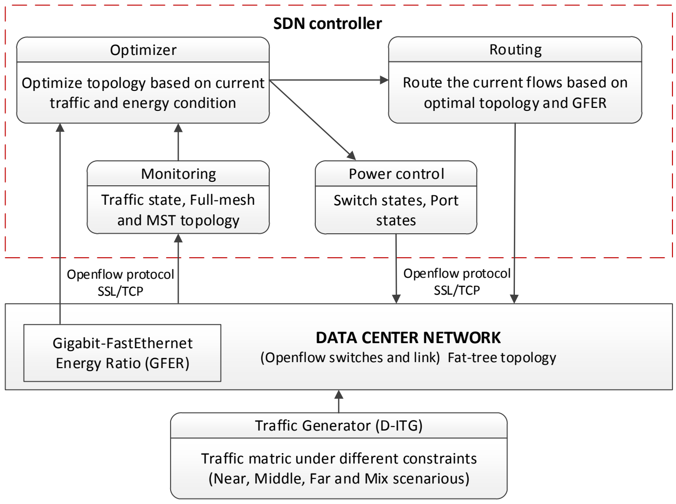

Fig. 3. System architecture

In additional, we develop this Elastic-tree model by building extended blocks such as *monitoring module* and *traffic generator*. These modules are described in more detail in Section IV

4) Power scaling approach

We use methods focusing on energy consumption proportionally adapted to the amount of data that flow via the network. Regarding Data Center Network (DCN), the *dynamic adaptation* approaches [9, 10, 11, 12] attracts much attention from the research communities recently. One of the approaches used in this work is *Power Scaling* which allows dynamically reducing the working rate of processing engines or of link interfaces, such as reducing the operating clocks of the devices. This approach does not only allow us to greatly reduce energy consumption but also to make the topology optimization and routing of the network more flexible. For example, the switch port of a commercial OpenFlow-enabled Pronto switch [13] consumes 63mW, 260mW and 913mW in their working rates of 10Mbps, 100Mbps and 1Gbps respectively. As a result, with a low traffic demand, we can save energy consumption by using link rate of 10Mbps or 100Mbps instead of constantly full capability speed of 1Gbps.

III. PROBLEM STATEMENT

The power consumption of DCN is calculated as the total power consumption of the network devices such as routers or switches. [14] shows that the power consumed by each device depends on followings factors:

- The number of active ports;
- The capacityrates at which each port operates. Normally 1Gbps port consumes the largest amount of power while the 10Mbps consumes the least [15];
- The amount of traffic that goes through a port does not have any significantly effect on its power consumption;
- The power consumption also depends on the specific devices which are individually different based on their power profiles.

In consequence, the power consumption of an DCN depends on the number of active links and switches as well as routing algorithm applied. For instance, [13] shows that the switchport of a commercial OpenFlow-enabled Pronto switch consumes 63mW, 260mW and 913mW in their working rates of 10Mbps, 100Mbps and 1Gbps, respectively. The energy consumption ratio of 1Gbps port to 100Mbps port is approximately 3.5. In contradiction to a common consensus that energy consumption of a network can be saved by turn off links and switches as many as possible, we prove that this argument is not always true. From the above example, in order to save energy we can use three 100Mpbs ports when traffic throughput is lower than 300Mbps, or 1Gbps port in case the throughput is higher than 300Mbps. In other words, instead of accumulating the amount of300Mbps throughput in an 1Gbps link, a routing algorithm can be performed to distribute the traffic to three 100Mbps ports, so that more energy can be saved. As a result, the topology optimization as well as routing algorithm must be flexible and aware of the power profile of network deivces as energy consumption ratio also depends on each device individually. In [15] we measure the power consumption of OpenFlow-enabled NetFPGA switch under several working states (*idle, 10Mbps, 100Mpbs and 1Gbps*). Results show that energy consumption of switch port is 112mW and 1080mW for 100Mbps and 1Gbps speeds respectively. That is, the energy consumption ratio of 1Gbps port to 100Mbps is approximately 9.5 times, that is much greater than that of Pronto. This consumed energy difference will have an impact on the routing and topology optimization process. From this point of view, the routing and topology optimization process must be based on the energy profile of devices.

In this work, we build an energy-saving scheme that can flexibly implement routing and topology optimization as well as effectively work with different network devices. To achive this target, we build the DCN based on the Software Defined Networking concept that allows energy-aware routing and optimization. The architecture is described more details in the next sections.

IV. ENERGY-SAVING SCHEME

A. Extending energy-aware DCN:

The Topology-Aware-Heuristic originally proposed by Heller et al. [8] is extended in our work by adding two modules as the following:

- *Monitoring module:* the module that resides in the OpenFlow controller monitors all network device statistics, such as switch state, link state, active topology, traffic utilization and so forth. The information is exchanged between switches and controllers via OpenFlow messages.
- *Traffic generator:* in our work, traffic is generated according to realistic traffic distributions. The *Distributed Internet Traffic Generator (D-ITG)* [16] is used in our test to generate Internet traffic among the servers in the fat-tree architecture.

B. Energy consumption of varios devices

1) 4-Gigabit-port NetFPGA card [15,17]: It is the low-cost reconfigurable hardware platform optimized for high-speed networking. The Net-FPGA includes all of the logic resources, memories, and 4 Gigabit Ethernet interfaces which are necessary to build a complete programmable switch, router, and/or security device. Because the entire data path is implemented in hardware, the system can support back-to-back packets at full Gigabit line rates and has a processing latency measured in only a few clock cycles.

TABLE I. ENERGY PROFILE OF NET-FPGA 4X1GBPS [15]

Net-FPGA configuration		Power (mW)
Static (baseline)		4496
FPGA Core	At least one 1Gbps port	2694
	Not all ports running at 1Gbps (10/ 100Mbps)	1394
	Idle	81
One Ethernet port	Idle	23
	10Mbps	52
	100Mbps	112
	1000Mbps	1080

Based on the results of power consumption (Table 1) of switches, ports and links, the energy saving of the whole network is estimated by using the power model [6] of each switch.

$$P_{sw} = P_{static} + (N_0 * P_0 + N_{10} * P_{10} + N_{100} * P_{100} + N_{1000} * P_{1000}) + P_{FPGA} \quad (1)$$

Where:

- *N0, N10, N100, N1000:* number of ports in link state *Idle*, *10Mbps*, *100Mbps*, *1Gbps*, respectively.
- *P0, P10, P100, P1000*: power consumption of ports in link state *Idle*, *10Mbps*, *100Mbps*, *1Gbps*, respectively.
- *P_{FPGA}: power consumption of FPGA Core*

2) Pronto 3240: As we mentioned before, [13] shows that the power consumption can be determined with Equation (2) and summarized in Table 2

$$P_{sw} = P_{static} + N_{10}{*}P_{10} + N_{100}{*}P_{100} + N_{1000}{*}P_{1000} \qquad (2)$$

TABLE II. POWER SUMMARY FOR A 48-PORT PRONTO 3240

Configuration	Power (mW)
P_{static}	67.700
$P_{10Mpbgs}$ per port	63
$P_{100Mpbgs}$ per port	260
$P_{1000Mpbgs}$ per port	913

C. Gigabit-FastEthernet Energy Ratio (GFER)

From the above consumed energy results of links, switches, we can see that the difference of power consumption between link rates 100Mbps and 1Gbps is remarkable. The routing and topology optimization algorithm must be based on this difference in order to make the DCN more energy-efficient. In this article we propose an index called the *Gigabit-FastEthernet Energy Ratio (GFER)*, which is rounded down the energy consumption ratio P_{1Gbps} to $P_{100Mbps}$. In our SDN energy-aware DCN, the controller uses this index as an input value to calculate the energy-efficient path and subnet.

$$GFER = \frac{P_{1Gbps}}{P_{100Mbps}} \qquad (3)$$

We use the two aforenemtioned *GFERs* as the examples in this work. The GFER_{FPGA} of OpenFlow-enable Net-FPGA is 9 and $GFER_{PRONTO}$ of OpenFlow-enable Pronto 3240 is 3.

D. Algorithms

1) Power scaling algorithm

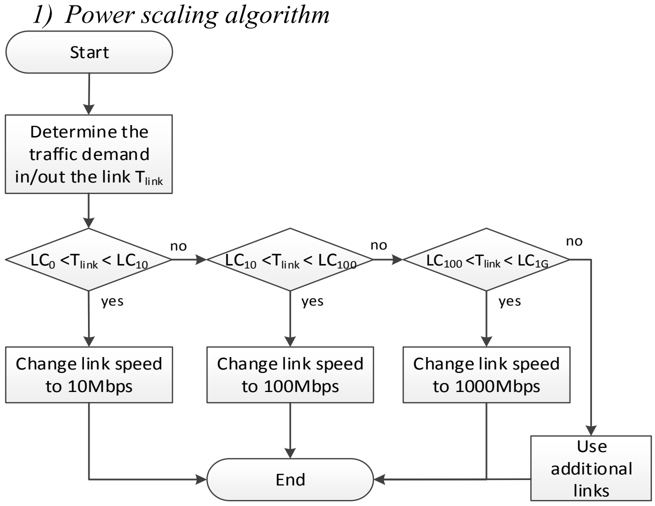

Fig. 4. Power scaling algorithm

The role of the *optimizer* component is to find a network subset which satisfies current traffic condition. Its input includes topology, network traffic utilization, the power profiles of switches, and the desired fault tolerance properties [8]. In our work, we deploy the *power scaling* approach that allows reducing energy consumption greatly by adaptively changing the working rate of the processing engines or links such as reducing operating clock of devices or decreasing the link rate of a switchport. The diagram in Fig.4 shows the operation of the proposed *power scaling* algorithm [18]. Based on the traffic state measured by the *monitoring* module, the *optimizer* determines which state must be used on this link then sends the results to the *power control* and *routing* modules. The speeds of link can be changed to *10Mbps, 100Mbps and 1Gbps* adaptively to traffic demands.

In this flowchart, we use:

- *TLink*: Traffic demand on link
- *LC* - Link Capability: 0Mbps, 10Mbps, 100Mbps, 1Gbps

2) Routing and topology optimization algorithm

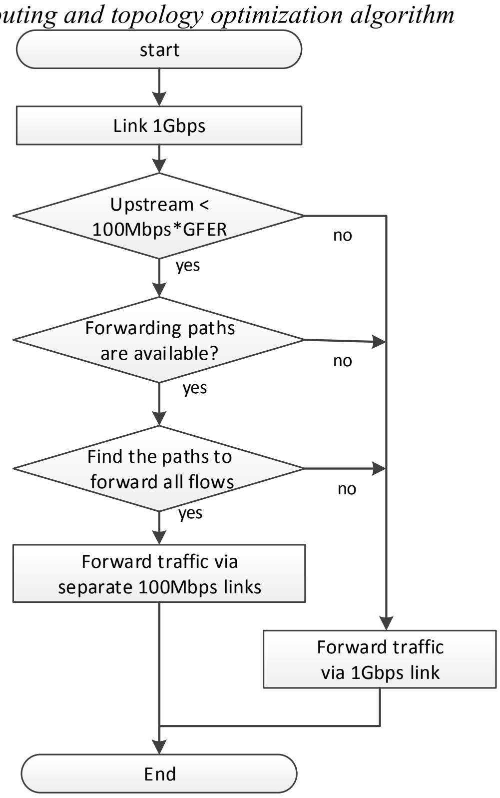

Fig. 5. Routing and topology optimization algorithm

We further develop the scheme by adding the topology optimization algorithm and device energy-profiling method. By using this, the scheme can automatically determine in which conditions the traffic will be routed via separate 100Mbps links instead of one 1Gbps link. The diagram in Fig.5 shows an algorithm for a link between edge and aggregation layers (and the algorithm for a link between aggregation and core layers is similar in fat-tree topology). For each *1Gbps* link, we run the algorithm to check the following criteria:

- $Traffic_{Link} < 100Mpbs * GFER$: The traffic demand on this link is smaller than the multiplication of 100Mbps and GFER
- Whether the forwarding paths for all 100Mpbs links are available to the destinations. In this condition, the routing algorithm tries to find new paths for all flows

in this link. With each flow, we can find out a source-destination edge switch. After determining the new route, the number of 100Mbps links that must be turned on are calculated and then routes for the flows in the link are modified.

If the *1Gbps* link satisfies all of the criteria, we try to find new paths for all flows in this link. With each flow, we can find out a source/destination edge switches and forward it via 100Mbps. With the GFER of each switch, the algorithm can energy-efficiently work with any devices' factor.

V. PERFORMANCE EVALUATION

A. *Case study*

With the power-profile-aware topology optimization algorithm, the DCN scheme adaptively works and routes data based on the traffic demands and devices' energy profiles. With the purpose of improving the reliability as well as decreasing the packet loss rate and delay, in this work we always maintain the minimum working speed of all switches (all ports work at 10Mbps) in the network even though the port is idle. In this section we are going to present a case study with two aforementioned types of switches in the same network topology and traffic condition. The topology in this case study is *K=8* Fat-tree, which can support 128 servers and the testing flow is a 350Mbps far-traffic from one POD to another (from server A to server B).

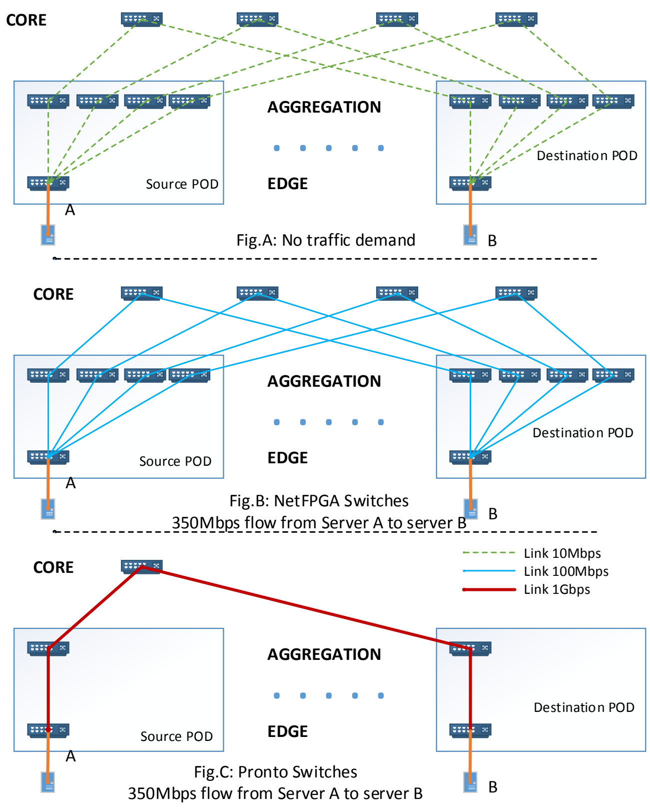

Fig. 6. Topology optimization and routing in a DCN.

Due to the large size topology, we only describe the related part of the DCN instead of whole DCN. Fig.6.A shows a part of the DCN that relates to devices in two PODs where the server A and server B connect to. The default working speed between switches is 10Mbps. We can see in the Fig.6.B where the DCN works with OpenFlow-enable Net-FPGA switches, the $GFER_{FPGA}$ is nine (100Mbps*$GFER_{FPGA}$ is greater than the demand flow – 350Mbps), the traffic is routed via separate 100Mbps links. In contrast, Fig.6.C shows the DCN with Pronto switches (the $GFER_{Pronto}$ is three and 100Mbps* $GFER_{Pronto}$ is lower than the demand flow), the traffic is routed via one 1Gbps link in the same traffic conditions with previous Net-FPGA case.

The case study illustrates the flexibility of the algorithms. With the same topology and traffic conditions, the data traffic is served in different routes based on the energy-profiling of switches.

B. *Performance comparison*

1) *Comparison with power scaling algorithm*

In this session, we measure the energy consumption of energy-aware DCN with two different energy-efficient strategies: *Power Scaling (PS)* and *Power Scaling with Energy-Profile-aware (PSnEP).* Under the similar traffic condition, we combine the energy-saving level of each algoritm. The traffic scenario is mix (ratio among *near:middle:far* is 1:1:1).

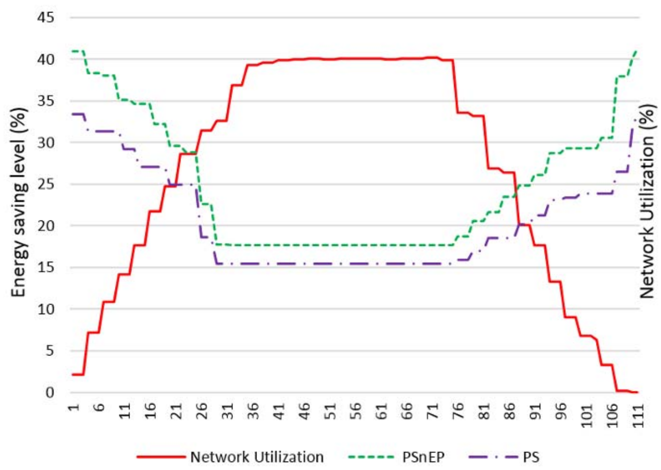

Fig. 7. Fat-tree with K = 6, mix scenario, Net-FPGA switch

The network utilization (NU) is the current data in the DCN, $NU = \frac{\sum_i \sum_j t_{ij}}{LinkSpeed * Server\_link}$, where t_{ij} is the traffic from server *i* to server *j*, *Server_link* is the number of link to servers (the DCN with fat-tree topology can support $k^3/4$ servers) and the *LinkSpeed* is 1Gbps. The *energy-saving* level is the ratio between amount of saved energy when applying new schemes and energy of fullmesh topology (all switches and ports are working at maximum speed – 1Gbps). In case of fullmesh, the energy consumption level is 100% and the energy-saving level is 0%. By using the traffic generator D-ITG, we sussefully build and experiment the energy-aware DCN based on SDN, the *energy-saving level* is inveserly proportional to the *Network Utilization*. As seen in Fig. 7, the difference of energy-saving level between *Power Scaling* and our algorithm – *Power Scaling with Energy-Profiling-Aware* routing is significant.

2) Comparison among different topology's size

We evaluate the energy-saving level of the DCN in various sizes of fat-tree topology (up to *K*=16 topology that can support 1024 servers). In cases of a small size topology that is up to *K*=6 (support 54 servers), we use the real testbed that is addressed in [18], while in large topology we use the *Reliable Analyzer for Energy-Saving* simulation tool [19]. Table III shows the average energy-saving level ratio of our *Power Scaling with Energy-Profiling-aware* algorithm to the conventional *Power-Scaling*.

TABLE III. ENERGY-SAVING RATIO OF *PSnEP* TO *PS* ALGORITHM

Fat-tree topology	*K*=4	*K*=6	*K*=8	*K*=12	*K*=16
$Average(\frac{Power_{PSnEP}}{Power_{PS}}\%)$	90.06	94.43	95.16	96.48	97.14

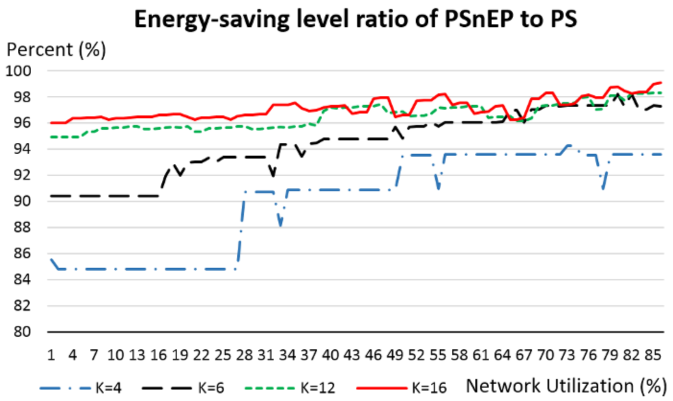

Fig. 8. Energy-saving level ratio of PSnEP to PS algorithm

The energy-saving level ratio in various sizes of topology shows that our novel algorithm outperforms the conventional *Power Scaling* algorithm for small size of DCN. When the network utilization of the DCN or the size of topology increases, the energy-saving ratio gradually decreases.

VI. CONCLUSION

In this paper, we successfully deploy a novel energy-efficient DCN scheme by using the Software Defined Networking technology. The scheme optimize network topology and route traffic based on the energy profiles of network devices. Results show that PSnEP outperforms the conventional PS in terms of energy consumption. For future work, we are going to verify the QoS and fault-tolerance of the data center network and aim at proposing trade-off between energy saving level and QoS.

VII. REFERENCE

[1] J. Koomey, Growth in Data Center Electricity Use 2005 to 2010, Analytics Press, 2011.

[2] Global e-Sustainibility Initiative (GeSI), — SMART 2020: Enabling the low carbon economy in the information age, 2012

[3] N. Mckeown, "How SDN will Shape Networking," October 2011. [Online]. Available: http://www.youtube.com/watch?v=c9-K5O qYgA

[4] S. Schenker, "The Future of Networking, and the Past of Protocols," October 2011. [Online]. Available: http://www.youtube.com/watch?v= YHeyuD89n1Y

[5] OpenFlow Switch Specification – v1.3, Open Network Foundation, September 27, 2013. Available: https://www.opennetworking.org

[6] M. Al-Fares, A. Loukissas, A. Vahdat, A scalable, commodity data center network architecture, in: ACM SIGCOMM, 2008.

[7] R.N. Mysore, A. Pamboris, N. Farrington, N. Huang, P. Miri, S. Radhakrishnan, V. Subramanya, A. Vahdat, Portland: a scalable fault-tolerant layer 2 data center network fabric, SIGCOMM Comput. Commun. Rev. 39 (4) (2009) 39–50.

[8] B. Heller, et al , "Elastic tree: saving energy in data center networks", the 7th USENIX Symposium on Networked System Design and Implementation (NSDI), San Jose, CA, USA, April 2010

[9] S. Nedevsky, et al., "Reducing network energy consumption via sleeping and rate adaptation", Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation (NSDI'08), 2008.

[10] R. Bolla, R. Bruschi, A. Carrega, F. Davoli, "Green network technologies and the art of trading-off", Proceedings of the IEEE INFOCOM 2011 Workshop on Green Communications and Networking, Shanghai, China, April 2011, pp. 301–306.

[11] R. Bolla, R. Bruschi, et al, "Cutting the energy bills of Internet Service Providers and Telecoms through power management: an impact analysis", Comput. Networks 56 (10) (2012) 2320–2342. July.

[12] K. Christensen et al, "IEEE 802.3az: the road to energy efficient Ethernet", IEEE Commun. Magz. 48 (11) (2010) 50–56.

[13] Xiaodong Wang, et al, "CARPO: Correlation-aware power optimization in data center networks", INFOCOM, 2012 Proceedings IEEE, March 2012

[14] P. Mahadevan, P. Sharma, S. Banerjee, and P. Ranganathan, "A Power Benchmarking Framework for Network Devices," in Proceedings of IFIP Networking, May 2009

[15] Pham Ngoc Nam, Nguyen Huu Thanh, Vu Quang Trong, Tran Hoang Vu, Truong Thu Huong, Phuoc Tran-Gia and Christian Schwartz, "A New Power Profiling Method and Power Scaling Mechanism for Energy-Aware NetFPGA Gigabit Router", in Elsevier's Journal of Computer Networks, Special Issue on Green Communications (doi:10.1016/j.comnet.2014.10.036)

[16] Traffic generator, D-ITG [Online]. www.grid.unina.it/software/ITG/

[17] V. Sivaraman A. Vishwanath Zhi Zhao, C. Russell, Profiling perpacket and per-byte energy consumption in the NetFPGA Gigabit router, in: IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2011 vol., no. 10–15 April 2011, pp. 331–336.

[18] T. Nguyen Huu et al., Modeling and experimenting combined smart sleep and power scaling algorithms in energy-aware data center networks, Simulat. Modell. Pract. Theory (2013), http://dx.doi.org/10.1016/j.simpat.2013.05.011

[19] Tran Manh Nam, Truong Thu Huong, Nguyen Huu Thanh, Pham Van Cong, Ngo Quynh Thu, Pham Ngoc Nam, "A Reliable Analyzer for Energy-Saving Approaches in Large Data Center Networks", in Proceedings of ICCE 2014 – July 2014