Chapter 8
Software-Defined Networking (SDN)-Based Network Services for Smart Learning Environment

Kannan Govindarajan, Vivekanandan Suresh Kumar, David Boulanger, Jérémie Seanosky, Jason Bell, Colin Pinnell, Kinshuk and Thamarai Selvi Somasundaram

Abstract This research paper introduces a state of the art for coupling software-defined networking (SDN) capability in smart learning environment (SLE). Smart Competence LEarning Analytics platform (SCALE) is the learning analytics platform, which has been developed to analyze and measure the learners data (Big data) generated in SLE for understanding the various learners competence measures. The learning analytics involves various processes such as distributing, analyzing, and merging the results across the clusters. It requires a large amount of bandwidth and computation cycles in an on-demand manner to reduce the total computation time and computation cost. Hence, in this research paper, we introduce a SDN-based networking principle with SCALE, which provision the network paths and allocates the bandwidth in an on-demand manner. The experiment results were conducted to analyze the performance and impact of SDN in SCALE platform.

Keywords Smart learning environment · Big data · Learning analytics · Smart Competence LEarning Analytics platform · Software-defined networking · OpenFlow · Computation time · Computation cost

K. Govindarajan (✉) · V.S. Kumar · D. Boulanger · J. Seanosky · J. Bell
C. Pinnell · Kinshuk
Athabasca University, Edmonton, Canada
e-mail: kgovindarajan@athabascau.ca

T.S. Somasundaram
Madras Institute of Technology, Anna University, Chennai, India
e-mail: stselvi@annauniv.edu

Y. Li et al. (eds.), *State-of-the-Art and Future Directions of Smart Learning*,
Lecture Notes in Educational Technology, DOI 10.1007/978-981-287-868-7_8

8.1 Introduction

Smart learning environment (SLE) [1] is a powerful learning tool that provides on-demand and adaptive support to the learners based on the analysis of their learners needs. In recent years, it has became more popular and a large number of learners, academic institutions, and industries started to use across the worldwide that generates large amount of learners' data. Smart Competence LEarning Analytics platform (SCALE) [2] is an in-house learning analytics platform designed and developed to track the fine-level learning experiences and translates those experiences into opportunities for customized feedback, reflection, and regulation. It has witnessed that, it has to deal with a large amount of learners data also known as Big data. Hence, it is essential to extract, process, and analyze those learners data in a rapid and an efficient manner to understand the learners various competence measures. Ku [3] analyzed the problems in conventional e-learning environment and stated that "smart learning is not just providing e-learning services through mobile devices, it should have an intelligent customized learning services optimized for the learners' on a personalized mobile device." Map Reduce [4] and Dryad [5] are the most popular distributed processing frameworks used in the market to process the Big data. These frameworks split the Big data and distribute those data in parallel across the Hadoop [6] clusters to analyze, transfer, and merge the results. The Big data processing requires vast amount of computation cycles and bandwidth across the clusters and between the analytics engine and the clusters in on-demand manner.

However, SCALE is currently using the conventional networking principle of subscription-based method. Conversely, the main drawback in the conventional networking is allocation of bandwidth in an on-demand manner that requires lots of manual intervention and time-consuming process and involves more laboring cost that increases the total computation time and computation cost. Hence, in this research paper, we propose to marry the software-defined networking (SDN)-based [7] networking principle with SCALE platform, which will create the network paths, inject the flows, and allocate the bandwidth based on the demand. OpenFlow [8] protocol is used to establish the communication between the networking devices. The research works [9–11] have investigated the advantages of SDN for Big data application processing to reduce the processing time, dynamically creating network paths, etc. The article discussed [12] the capability of SDN to build a large, intelligent network that is capable to handle both structured and unstructured data as part of Big data analytics, and it eliminated the pain of manual administration, which exists in the conventional networking resources. In summary, the main contributions of research work are summarized below:

(A) Designing a SDN-based network services for SCALE platform for performing learning analytics in a Smart Learning Environment.
(B) Provisioning of network paths, injecting flows, and allocating the bandwidth in a dynamic manner for reducing the total computation time and total computation cost.

(C) Creating a SDN-based Hadoop infrastructure with OpenFlow switches and OpenFlow protocols.
(D) Analyzing the impact of SDN-based networking services in the SCALE platform.

The rest of this paper is organized as follows: Sect. 8.2 describes the research works, which are closely related and provides background knowledge to support our research work. Section 8.3 discusses the problem statement and the proposed solution to solve the problem. Section 8.4 describes the system architecture for implementing the proposed work. Section 8.5 discusses the performance measures of the proposed system. Section 8.6 concludes the paper with recommendations and future work to be explored further.

8.2 Related Works

Ku [3] analyzed the base technologies for the SLE and examined the problems in conventional e-learning environment. The study has concluded that, it is essential to include context aware technology, Big data technology, and cloud computing in traditional e-learning services to build a SLE. Das et al. [10] proposed a network management framework named FlowComb, which is helpful for Big data application processing to achieve high utilization and low data processing times. Their proposed system is capable to predict the application transfer time based on the software agents installed on application servers. Their experimental results are evident that their proposed system reduces the Big data processing time in a remarkable manner. Benson et al. [13] proposed a framework named as Cloud Networking as a Service (CloudNaaS). It extends the self-provisioning model of providing network devices in an on-demand manner in addition to compute and store devices. Mahenge and Mwangoka [14] designed a cost-effective mobile-based content delivery system to deliver the contents in a resource and bandwidth constraint environment. They claimed that the proposed system has the potential to reduce the bandwidth cost, the server workload, and the Internet usage overhead. They improved the quality of experience and learners' participation using their proposed system.

8.3 Problem Statement and Solution

SCALE requires a large amount of computation cycles and network bandwidth to process and analyze the collected data in an on-demand manner. Figure 8.1 represents the working mode of SCALE using the conventional networking structure. Figure 8.2 represents the schematic representation of the solution to solve the above-said problem.

SCALE communicates with SDN controller through SDN connector to initiate learners data analytics process. The controller node and cluster nodes should be

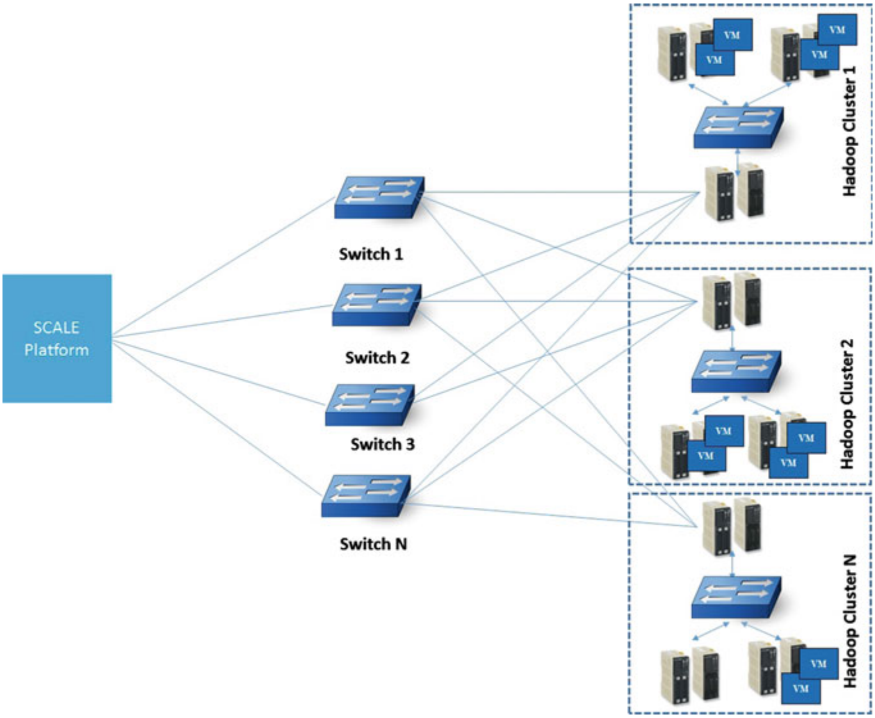

Fig. 8.1 Problem representation

configured with Open vSwitch (OVS) [15] software that can act as the physical switch, as well as the virtual switch. The queues (Q1–QN) are created in OVS and each queue is assigned with a specific bandwidth value. The virtual network V1 and V2 has been created, and it is allocated with a link bandwidth value of B1 and B2, respectively; it can be modified dynamically based on the demand.

8.4 Proposed System Architecture

The system architecture for implementing the proposed research work is shown in Fig. 8.3.

The data collector module collects the learners' data from Moodle, MI-LATTE, CODEX, and Hackystat sensors that consists of both structured and unstructured data. It stores the data in Hadoop Distributed File System (HDFS) clusters through data storage manager. The SDN connector acts as the interface between the OpenFlow switches and SCALE platform using the OpenFlow protocol. The data preprocessor periodically queries, fetches and, preprocesses the learners' data that remove the unwanted fields, name, and id of the learners for privacy purpose.

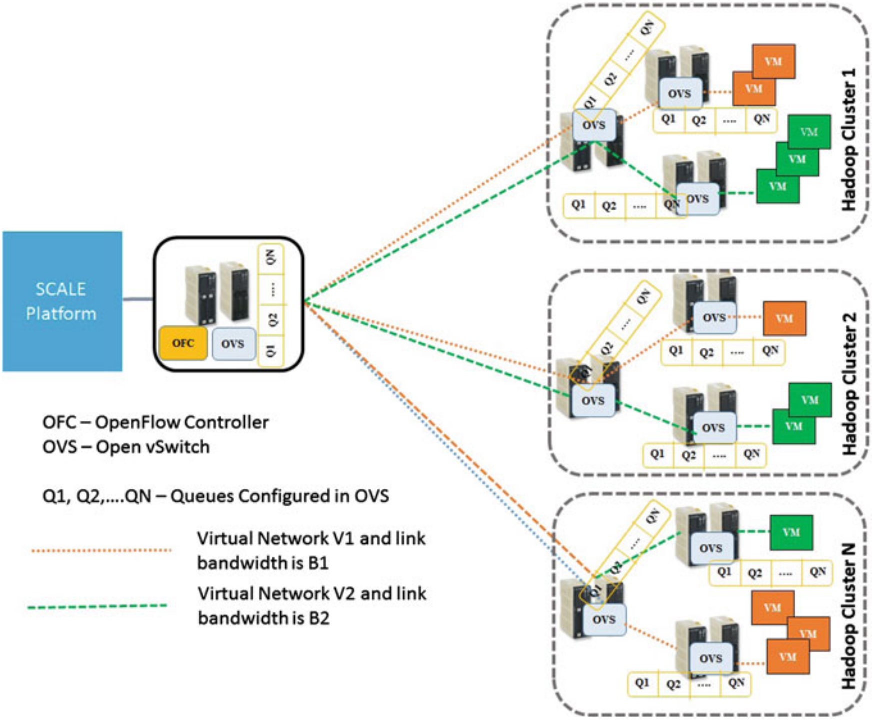

Fig. 8.2 Schematic representation of solution

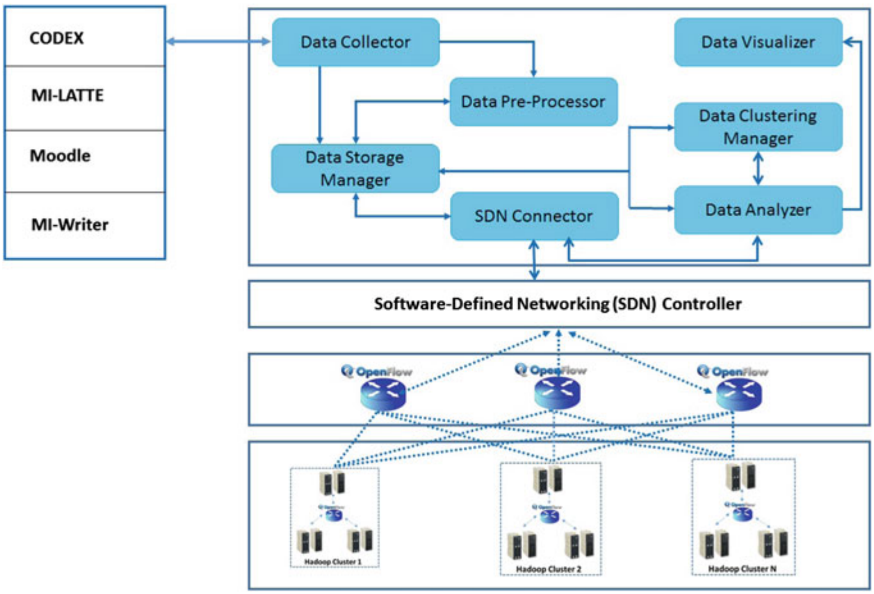

Fig. 8.3 System architecture

The preprocessed data are sent to the data clustering manager which is embedded with the particle swarm optimization (PSO) [16] clustering algorithm. It group the similar learners in an optimal manner based on the learners efficiency, quality, and accuracy. The data analyzer component is integrated with various machine learning and statistical algorithms that initiates the learning analytics process across the Hadoop clusters. The SDN controller provisions the network paths in the OpenFlow switches in a dynamic manner and allocates the required bandwidth to network links in each cluster by injecting the flows for performing learning analytics process. Finally, the analyzed learners' data are fed into the data visualizer to visualize the various competence measures in terms of graphical charts and statistical values.

8.5 Performance Measures

The performance of a SDN-based learning analytics platform is determined by analysis of total computation time and computation cost, which is calculated using Eqs. (8.1) and (8.2). To perform this experiment, we have generated the learners data, which is based on the traces of real-time learners data. The data size ranges from 2 to 32 GB and the OpenFlow switches are configured with a maximum date rate of 200 Mbps. For the first experiment (without SDN), a single queue is created and allocated a bandwidth value of 100 Mbps. The second experiment (with SDN) is conducted by creating multiple queues, which ranges from Q0 to Q4 and bandwidth is allocated in the range of 20–100 Mbps. The proposed work allocates the bandwidth to the learning analytics process of various data sizes in an on-demand manner based on the required bandwidth in SCALE platform. Figures 8.4 and 8.5 represent the total computation time and total computation cost to perform the learning analytics process for the collected learners data. The total computation time (TC_{Time}) is calculated using Eq. (8.1). It is the summation of task distribution time (TD_{Time}), data analyzing time (DA_{Time}), results merging time

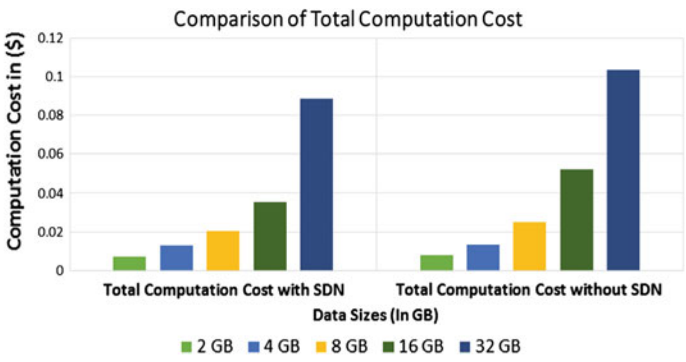

Fig. 8.4 Comparison of total computation time

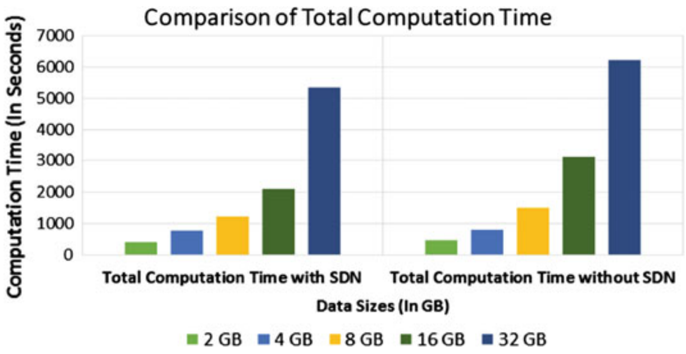

Fig. 8.5 Comparison of total computation cost

(RM_{Time}), and data transfer time (DT_{Time}). The total computation cost (T_{Cost}) is calculated using the Eq. (8.2), which is based on the usage of bandwidth value for the learning analytics process of various data sizes.

$$TC_{Time} = TD_{Time} + DA_{Time} + RM_{Time} + DT_{Time} \tag{8.1}$$

$$TC_{Cost} = (TD_{Time} + RM_{Time} + DT_{Time}) * BW\_Usage * \$/h \tag{8.2}$$

8.6 Discussion and Recommendation

SLE is capable of providing an on-demand and adaptive support for the learners based on their needs. SCALE is the learning analytics platform to analyze and understand the learners competence measures that requires a huge amount of computation cycles and bandwidth in an on-demand manner. In this research paper, we have investigated the effect of coupling SDN-based networking principle in learning analytics process. From this study and the experimental results, the following things have been identified and recommended as follows:

- The SLE generates a huge volume of learners' data known as Big data. The total computation time required to process those data are high that increase the total computation cost.
- The conventional networking is unsuccessful for provisioning of bandwidth in an on-demand manner due to various limitations. However, SDN has the capability to provision the network paths, allocating the bandwidth to the provisioned network paths in an on-demand manner.
- The SDN-based network services in Smart Competence LEarning Analytics (SCALE) platform reduces the total computation time and computation cost of learning analytics process.

The future work intends to extend this research study further: (i) to develop an intelligent scheduling mechanism to identify the suitable computation and network nodes that increase the efficiency of learning analytics process. (ii) to extend the SDN-based SCALE platform for video streaming of online lectures and improving the Quality of service.

References

1. Hwang, G.-J. (2014). Definition, framework and research issues of smart learning environments—a context-aware ubiquitous learning perspective. *Smart Learning Environments*, *1*, 4. doi:10.1186/s40561-014-0004-5
2. Kumar, V., Boulanger, D., Seanosky, J., Kinshuk, Panneerselvam, K., & Somasundaram, T. S. (2014). Competence analytics. *Journal of Computer Education, 1*(4), 251–270.
3. Ku, J.-H. (2014). Analysis of base technologies for building smart learning service environment. *Advanced Science and Technology Letters, 49*, 46–49. doi:10.14257/astl.2014.49.11
4. Dean, J., & Ghemawat, S. (2004, December). MapReduce: Simplified data processing on large clusters. In *OSDI'04: Sixth Symposium on Operating System Design and Implementation*. San Francisco, CA.
5. Isard, M., Budiu, M., Yu, Y., Birrell, A., & Fetterly, D. (2007). Dryad: distributed data-parallel programs from sequential building blocks. In *European Conference on Computer Systems (EuroSys)*.
6. Hadoop (2015). https://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-common/ClusterSetup.html
7. Software-Defined Networking (SDN). The new norm for networks. https://www.opennetworking.org/images/stories/downloads/openflow/wpsdn-newnorm.pdf
8. OpenFlow (2015). http://archive.openflow.org/wp/learnmore/
9. Monga, I., Pouyoul, E., & Guok, C. (2012, November) Software-defined networking for big-data science—architectural models from campus to the WAN. In *High Performance Computing, Networking, Storage and Analysis (SCC), 2012 SC Companion* (pp. 1629, 1635). November 10–16, 2012. doi:10.1109/SC.Companion.2012.341
10. Das, A., Lumezanu, C., Zhang, Y., Singh, V., Jiang, G., & Yu, C. (2013).Transparent and flexible network management for big data processing in the cloud. In *Proceedings of 5th USENIX Conference on Hot Topics Cloud Computing* (pp. 1–6). San Jose, CA, USA.
11. Qin, P., Dai, B., Huang, B., & Xu, G. (2014). Bandwidth-aware scheduling with sdn in hadoop: A new trend for Big data. *CoRR, abs/1403.2800*.
12. Big data blog (2014). http://www.ingrammicroadvisor.com/big-data/how-software-defined-networking-can-increase-your-big-data-roi
13. Benson, T., Akella, A., Shaikh, A., & Sahu, S. (2011). CloudNaaS: A cloud networking platform for enterprise applications. In *Proceedings of the 2nd ACM Symposium on Cloud Computing, SOCC'11* (pp. 1–13).
14. Mahenge, M. P. J., & Mwangoka, J. W. (2014, December). Mobile-based system for cost-effective e-learning contents delivery in resource and bandwidth constrained learning environments. *Journal, Knowledge Management and E-Learning*, *6*(4), 449.
15. Open vSwitch (OVS) (2015). http://openvswitch.org/releases/openvswitch-2.3.1.tar.gz
16. Govindarajan, K., Somasundaram, T. S., Kumar, V. S., Kinshuk, (2013). Particle swarm optimization (PSO)-based clustering for improving the quality of learning using cloud computing. In *2013 IEEE 13th International Conference on Advanced Learning Technologies (ICALT)* (pp. 495, 497). July 15–18, 2013. doi:10.1109/ICALT.2013.160