A Scalable and Virtualized Testbed for IoT Experiments

Ruichun Gu[1,2], Hongyu Zhang[1], Dong Pei[1] and Junxing Zhang[1,*]

[1]Inner Mongolia University, College of Computer Science, Hohhot, China
[2]Inner Mongolia University of Science and Technology, Baotou, China
reachcool@163.com, junxing@imu.edu.cn (*Corresponding author)

Abstract With the rapid development of information technology, the Internet of Things (IoT) is getting more attentions, which has promoted a new wave of information and industrial tide. In order to reduce the CAPEX, OPEX and TIEX of IoT systems, before the deployment, people often evaluate IoT architectures, protocols, as well as configurations on testbeds. As physical resources of any testbed are limited, it is challenging to conduct large-scale IoT experiments. In this paper, we apply the virtualization technology to optimize an existing network testbed, adopt VMNet to emulate wireless sensors in the IoT experiment, and improve the scalability of the experiment to support large-scale IoT assessments. Our scheme leverages a multi-host collaborative with static multi-sink architecture to solve the bottleneck problem of the large-scale IoT emulation and multiple interpolation algorithms to supplement the time continuity and spatial integrity of the sensed data for enhanced fidelity of the IoT experiment. According to our experiments, the virtualized IoT testbed not only reduced the TIEX of IoT emulation sharply, but also enhanced the scalability of IoT experiments.

Keywords: IoT; WSN; Testbed; Virtualization; Scalability

1 Introduction

The rapidly developing of information technology has driven the widely applying of the Internet of Things (IoT). IoT is one of the latest information process technologies, the objects of which have the features of connecting with each other, distributed in widely area, and processed concentrated. Through the wireless and mobile networks, we could get the information about the environment by IoT. The development of the IoT technology has benefited from the continuous progress of wireless network and information processing technology. With the developing of the wireless and mobile network, these intelligent objects could be connected all the time. The data collected from the IoT is also increasing, therefore, we need to analyze them more efficiency for the decision support of reality production and life [1, 2].

As the sensor plays a major role in the IoT architecture, the emergence of the IoT paradigm has augmented the scope of WSNs demand, further cultivating the ongoing research in this field. An estimated 50 billion devices are envisioned to be connected to the network by 2020 [3]. WSNs are expected to play a significant role in IoT, since the sensor nodes are the main building blocks of this concept [4].

A general WSN node is composed with the sensor unit, the communication unit, the power unit and the process unit, it must be small and lightweight for different deploy demands, therefore, the ability of computing, power supply, storage and communicate have been constrained strictly. A WSN has special structure--many of nodes with high density, nodes failure rate is high and the topology may change at any time--once had been deployed, it cannot be modified easily. As a result, for the purpose of reducing the Capital Expenditure (CAPEX), Operating Expenditure

TRIDENTCOM 2017, September 28-29, Dalian, People's Republic of China

DOI 10.4108/eai.28-9-2017.2273310

(OPEX) and Time Expenditure (TIEX), before the deployment, it is necessary to emulate or simulate the architecture, protocols, and parameters of an actual WSN. In the emulation, there should be thousands of nodes to satisfy the need of real WSN, which determines the design of a perfect WSN emulation testbed is very difficult. Serval publicly available WSN testbeds, which could solve some emulation needs, but still have problems such as: 1), many of them use real sensor nodes for emulation, one set of sensor nodes could not emulate multiple experiments at same time, and it need a large dedicate space to deploy nodes, this is more difficult; 2), some emulation data is discrete, the emulation effect is poor, the results cannot meet the needs of real environment; 3), if the number of emulation nodes increases or the topology changes, the whole experiment must be redeployed, and cannot respond quickly. An ideal WSN testbed could postpone the actual deployment to the last minute, so as to design a WSN testbed, which is controllable, easy usable, flexible and could meet variety of experimental environment needs, is having the practical significance and research value.

In this paper, the virtualization technology is introduced into the IoT testbed, which could improve the resources utilization, enhance the scalability and flexibility, and meet the variety needs of large-scale IoT emulations. This paper presents a virtualization based emulation platform, whose purpose is to provide a high scalability and flexibility of real-time, distributed IoT emulation platform. The main contributions of this paper are: 1) adopt the VMNet [5] technology, a PC-based WSN emulation environment, to improve the scalability of the testbed; 2) find out the bottleneck of the emulation experiment, and propose a model of multi-host collaborative with static multi-sink architecture for this problem; 3) use appropriate algorithms to supplement the time continuity and spatial integrity of the emulation data.

The remainder of this paper is organized as follows: Section 2 discusses related work; Section 3 describes the proposed system architecture of virtualization IoT testbed; Section 4 presents the analysis and design of the virtualized IoT testbed in detail; we describe our experimental results and evaluation in Section 5; the conclusions and future work are in the Section 6.

2 Related Work

The large-scale experiments often need large quantity of nodes, but the number of sensor nodes in the prototype of WSN testbed is limited. As the IoT has a huge topology and large number of nodes, the scalability of the testbed has become a main challenge, some countermeasures are required.

To solve the problem of scalability, the basic way of most existing testbeds is to add more nodes into the emulation. This method is bound to cause more CAPEX, OPEX and TIEX. The number of nodes, after all, is limited, so it cannot resolve the problem of scalability completely. The other way is to combine other testbeds for the scalability, which needs to provide kinds of interfaces for the collaboration, as different testbeds have different structures and characters, this method always hard to be implemented.

Several testbeds use virtualization software for more virtual nodes, the Ohio State University's Kansei and the testbed of the University of Washington have more meaningful about virtualization.

The Kansei testbed is intended for facilitating the study of large-scale sensor network applications. A hybrid modeling solution is used, it connects the sensor array with host, and then turns up TOSSIM [6] simulators on the host, the applications are running on the simulator. When begin to sensing and communicating, pause the simulator, save the status of simulator, and then, use the real nodes connected with the host to complete the sensing and communicating job.

The testbed of the Washington University in St. Louis virtualized the MAC layer [7]. They had installed the GloMoSim on the testbed, and extended the structure of GloMoSim for the full emulation model.

The FIT IoT-Lab testbed of OneLab is an open testbed composed lots of low-power wireless nodes and mobile robots available for experimenting with large-scale wireless IoT technologies. The IoT-Lab testbed is deployed in 6 sites across France, each site features a different node and hardware capabilities, and all sites are interconnected and available through the web portal, common REST interfaces and consistent CLI tools [8].

Emulab of the University of Utah is an open network testbed, giving researchers a wide range of environments in which to develop, debug, and evaluate their systems [9]. Emulab has integrated the emulation, simulation and real-world networks with a variety of experimental environments. It's a time and space shared public facilities, which mainly used for network and distributed systems research, has a large number of users and installation sites around the world. Emulab is open source software, so anyone can get the source code for further development, and build their customized installation. In order to design a set of experimental platforms to meet the requirements of the IoT experiments, there are already many research and optimizations based on Emulab. GAO's work [10], which is a simple IoT testbed, had modified the communication interfaces of Emulab, and added the support for IoT nodes. For the host scalability research, Hibler [11] had proposed a node virtualization and optimization solution for the host scalability in the network work experiments. They extended the FreeBSD Jails [12] in the Emulab OS for the node virtualization.

The VMNet of the Hong Kong University of Science and Technology is a single sink WSN emulator that aims at enabling realistic performance evaluation for WSN applications. It is a PC-based emulation environment on which WSN applications can be directly run for debugging, testing, and performance evaluation. Additionally, VMNet has a highly modularized architecture for assembling virtual hardware components, studies on the lower layers (e.g., hardware drivers, OS, and networking) as well as cross-layer techniques can also be done in this environment by plugging the target modules into the emulator. However, VMNet could be running on just one host for each experiment, when the nodes are accumulating, the host resources are exhausting, and thus it could not emulate large-scale IoT experiments.

In this paper, we apply the virtualization technology for the scalability of the IoT testbed based on the Emulab platform and the PC-based WSN emulation environment VMNet. Owing to the advantages of Emulab, it greatly improves the operability of the IoT experiments; through deeply analyzed, the bottleneck of the IoT experiments is found out, and a multi-host collaborative with static multi-sink virtual emulation architecture is proposed, which solves the bottleneck problem of emulation performance, thus improves the scalability of the IoT testbed; In the aspect of the time continuity and spatial integrity of the emulation data, the appropriate algorithms are used to enhance the authenticity and reliability of the emulation experiment.

3 System Architecture

Based on the Emulab platform, we combine a virtualization WSN emulation environment named VMNet to implement an IoT testbed with high usability and scalability. The architecture of the proposed system is depicted in Fig. 1.

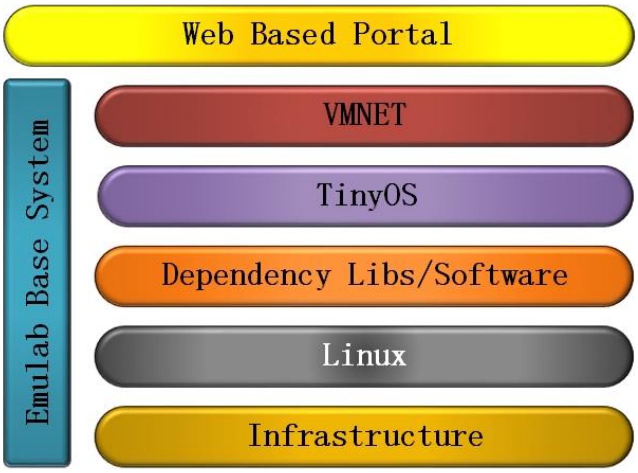

Figure 1 The Architecture of Virtualized IoT Testbed

As described in the figure, the configuration of this emulation environment is from bottom to up. First, based on the infrastructures, the OS of the nodes in Emulab must be one of the Linux distributions (we use Ubuntu 14.04 here). Second, install the dependency libs and software required by the TinyOS, on which the VMNet is running. Third, configure and compile the TinyOS, install the VMNet emulator, and then the IoT emulation could be carried out on this platform.

For the usability of the emulation testbed, we defined a proprietary image suitable for the execution of IoT experiments. This image includes the dependency libs and software, TinyOS, TinyDB and VMNet emulator. During beginning or swapping in an experiment, the nodes could boot from the PXE and get the proprietary OS image from the Frisbee server of the Emulab boss node, the parameters such as node's OS, host IP addresses, network performance metrics, and topology are defined in a NS liked script in the initial step of the experiment.

The users and resources of the whole testbed is managed by Emulab, a user could login the management platform through a web portal to operate the detail of emulation experiment. After logged in, in addition to configuring the emulation environment, it is necessary to configure the corresponding parameters of the WSN, such as the node type and topology, the configuration is writing in a XML file. When an experiment is swapped in, VMNet will read the parameters from the configured XML file, if no failures is reported, the emulation will start.

4 Analysis and Design

4.1 The performance bottleneck of emulation

For the emulation experiments of IoT, the more details of emulation, the more host resources are consumed. When the emulation requirements of the experimenter are higher and more complex, the emulation performance often becomes the focus. With the increase of the number of nodes, the emulation performance will decline. The emulation performance and scale become two interacted restriction factors. This is explained by the fact that the larger the scale, the more the host resources consumed, and thus the lower the performance. Moreover, as the network size increases, also, the number of hops are required to transfer sensor data to the sink node will increase, the more details will be in addition to the emulation, the consumption of system resources will increase as well, thus the emulation performance will decline accordingly.

In order to find the bottleneck factors that restrict the emulation performance, we have made an analysis of the resources utilizations when running emulation experiments, the results are shown in table 1. In this experiment, we have emulated 100 virtual nodes on a host server (Lenovo ThinkServer RS140). As described in the table, the memory and disc I/O do not restrict the performance, however, the CPU utilization is up to 100%, and the whole experiment performance have declined.

Table 1 Statistics of the host resource usage

Resources	Utilization Status
CPU	100%
Memory	12.69%
Disk I/O (logs and sensor data)	Avg. 20kb/s

To further validate the capabilities of a single host, we had emulated 1-100 virtual nodes on a single host with a single sink, the utilization of the host CPU is shown in Fig. 2. When the nodes number is up to 80, the CPU resource is exhausted. It could not emulate a large-scale IoT experiment more than 80 nodes.

Accordingly, in this paper, we had deemed that the CPU is the bottleneck of emulation performance.

In the VMNet, one experiment could only run on one host with a single sink, the CPU of the host limited the scalability of the IoT testbed. In order to solve this problem, we have optimized the VMNet, and proposed a multi-host collaborative with static multi-sink architecture virtualized IoT testbed, which extends the experiment from one host to a multi-host collaboration environment, and each sink node of each host is connected to the base station with wired links, the virtual nodes are connected with the sink node of local domain host with emulated wireless links. Through the collaboration of more CPUs of the multi-host to offloading the CPU pressure, so as to expand the number of virtual nodes, and to extend the scalability of the IoT testbed. The presented architecture is illustrated in Fig. 3.

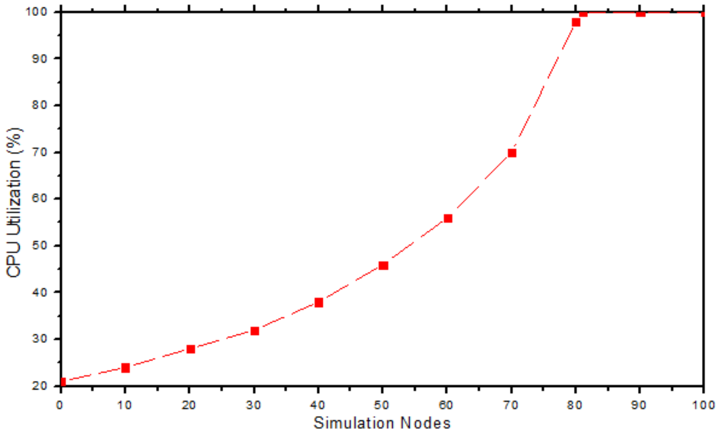

Figure 2 The Statistic of CPU utilization

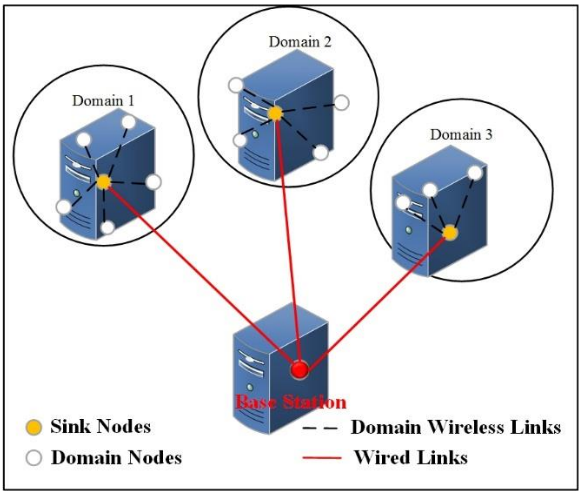

Figure 3 The structure of multi-host collaborative with static multi-sink emulation environment

4.2 The perfection of emulation data

Sensor data is an important part of the IoT experiments. The sensor data with higher facticity not only improve the fidelity of the emulation, but also help to validate the application and the network service. Due to the limited number of sensor nodes in the prototype of the IoT system, it cannot provide adequate sensor data for emulation; however, we could generate more realistic sensing data in other ways. The strategy adopted in this paper is based on the Intel Lab data set and the GDI data set collected in the real sensor network, and they are improved and standardized with appropriate algorithms for the emulation.

If the sample data set is used directly for emulation, although, it can guarantee the authenticity, but it cannot meet the diverse experimental requirements. In one hand, the sampling period of the sensing data in the sample data set is fixed. On the other hand, when the sampling period set by the experimenter is not consistent with the sampling period of sample data, at some time points the sensing data is absent.

According to the data set of time continuity and the spatial integrity, after have learned a variety of algorithms, we determine to use the Hermite interpolation algorithm to compensate the time continuity of sensor data, in the aspect of supplement the spatial integrity, for the Intel Lab data sets, we use inverse distance weighting interpolation algorithm, and Cubic Spline Interpolation algorithm for the GDI data set.

5 Evaluation

In this paper, the architecture of Emulab prototype is optimized, add the support of the IoT emulation environment, the bottleneck of the emulation performance is found out, and the scalability problem of the testbed is solved by virtualization technology of the multi-host collaborative with static multi-sink architecture. In order to verify the effectiveness of the optimization solution, some experiments are carried out for evaluation.

5.1 The experimental environment

We use the Lenovo ThinkServer RS140 (with the equipment of CPU: Xeon E3-1226 v3, 3.3GHZ; Memory: 4GB; Storage: 1TB) for the administration nodes and experimental nodes of testbed. We use 3 physical nodes in the multi-host collaborative experiments for the CPU pressure offload.

In the figures below, the dash line named Single-host Single-Sink with red square symbol indicate the emulation on a single host with one sink node, and the solid line named Multi-host Multi-Sink with blue round symbol indicate the proposed architecture of multi-host collaborative with static multi-sink, with one sink on each host.

5.2 The statistic of emulation time

We statistic the real time of emulating 10 seconds with 10-100 virtual nodes on the single host and multi-host, respectively, and the results are shown in Fig. 4. The virtual nodes are distributed evenly in the matrix with the separation distance of 5 meters. And the query statements running in these two emulations are the same as follows:

Query: SELECT temp FROM sensors SAMPLE PERIOD 2048

As shown in the figure, the emulation time of the multi-host multi-sink mode is significantly lower than the single-host single-sink mode, with the increasing of the emulation nodes, the time of the multi-host mode is increasing also, but comparing with the single host mode, the emulation time is decreasing significantly.

The experiment illustrates that when conducting a large-scale IoT emulation the deployment and test can be completed in a relatively short time with the multi-host mode. It could reduce the experimental TIEX, and improve the efficiency.

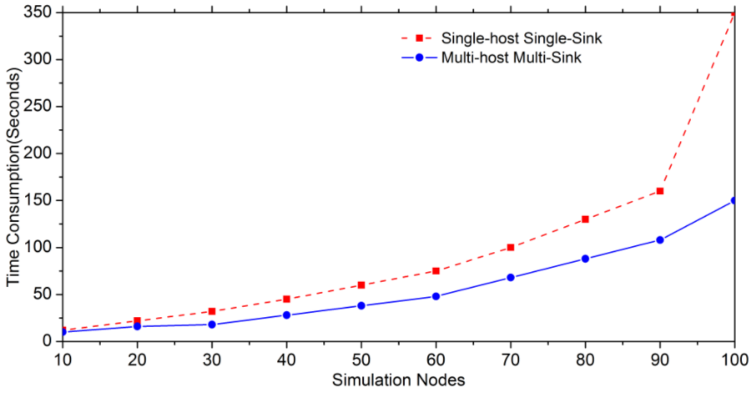

Figure 4 The compare of Emulation Time consumption

5.3 The statistic of CPU utilization

In order to find out the bottleneck of the virtualized IoT emulation, we had emulated the fluctuating of CPU utilization with the nodes number accumulating under the single-host single-sink mode in last section. And in this section, we do this test again for the statistic of CPU utilization under the multi-host mode, and the compare that with the last test, the results are shown in Fig. 5.

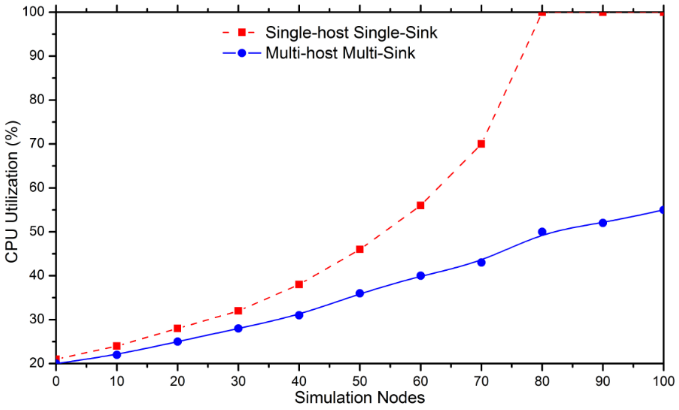

Figure 5 the Compare of CPU Utilization

As described in Fig. 5, in the single-host mode, when the number of emulation nodes is up to 80, the CPU utilization is achieved to 100%. And under the mode of multi-host mode, the advantage of CPU utilization is not obvious with few nodes. With the increasing of the nodes number, the CPU utilization rate has increased, but the slope of the increasing line is gradually flat, and the advantages are significant more and more. From here it can be seen that the virtualization technology with the multi-host collaborative with static multi-sink architecture, can improve the scalability of the IoT emulation platform effectively.

6 Conclusions and Future Work

The work of this paper is to optimize the IoT architecture prototype testbed, through virtualization technology to get rid of the limitations of hardware resources, and use the multi-host cooperation mode to offload the CPU pressure, effectively solves the scalability problem of the IoT testbed platform.

First of all, we have combined the Emulab testbed with the virtualization WSN emulator—VMNet to create a highly scalable virtualized IoT testbed. Second, we have measured the use of system resources in the emulations and found out the bottleneck of emulation performance is the CPU limitation of the single host on which the VMNet is running. In order to solve the bottleneck problem, we have proposed the multi-host collaborative static multi-sink architecture for the IoT emulation. Third, we have applied the appropriate algorithm to supplement the time continuity and spatial integrity of the emulation data in order to improve the fidelity of the IoT emulation experiment. In the last, we had verified the feasibility and effectiveness of the proposed solution. As the results shown, the scalability of the IoT testbed has been improved significantly by the proposed scheme.

However, there are still issues that need us to make further discussion on: first, we use more physical servers in the multi-host collaboration architecture to offloading the CPU pressure, but the other resources, such as memory and disk volume, did not make the utilization to the full; second, during the experiments, we should stop and restart them when the configuration is changed, it cannot be modified dynamically.

Therefore, in the future work, we will run the VMNET on the virtualization platform such as OpenStack, for better resource utilization. And will try to propose another architecture for a comprehensive solution for the dynamic resource schedule in the virtualized IoT testbed, which will modify the experiments configuration dynamically without stop them.

Acknowledgment: This work was supported in part by the National Natural Science Foundation of China (Grant No.61261019), the Inner Mongolia Autonomous Region Natural Science Foundation (Grant No.113113), and the Program of Higher-level talents of Inner Mongolia University.

References:

[1]. Atzori, L., A. Iera and G. Morabito, The Internet of Things: A survey. Computer Networks, 2010. 54(15): p. 2787-2805.

[2]. Daniel Giusto, et al., The Internet of Things. 2010: Springer New York.

[3]. Syst., C. and S. Jose, Fog Computing and the Internet of Things: Extend the Cloud to Where the Things Are, 2016.

[4]. Kobo, H.I., A.M. Mahfouz and G.P. Hancke, A Survey on Software-Defined Wireless Sensor Networks: Challenges and Design Requirements. IEEE Access, 2017. PP(99): p. 1-1.

[5]. Wu, H., et al., VMNet: Realistic Emulation of Wireless Sensor Networks. Parallel & Distributed Systems IEEE Transactions on, 2007. 18(2): p. 277-288.

[6]. Stevens, C., et al. Simulating Mobility in WSNs: Bridging the Gap between ns-2 and TOSSIM 2.x. in IEEE/ACM International Symposium on Distributed Simulation and Real Time Applications. 2009.

[7]. Blum, B., T. He and Y. Pointurier. MAC layer abstraction for simulation scalability improvements. 2007.

[8]. Adjih, C., et al. FIT IoT-LAB: A large scale open experimental IoT testbed. IEEE World Forum on Internet of Things. 2015.

[9]. Emulab, Emulab-Network Emulation Testbed Home. 2017.

[10]. Gao, H., A PRELIMINARY STUDY ON NETWORK TESTBED FOR EXPERIMENTS OF THE INTERNET OF THINGS. 2014, Inner Mongolia University.

[11]. Hibler, M., et al. Large-scale virtualization in the Emulab network testbed. Usenix Technical Conference, Boston, Ma, USA, June 22-27, 2008.

[12]. Kamp, P.H., R.N.M. Watson and R.F. Org. Jails: Confining the omnipotent root. 2010.