ISSN (Online) : 0974-5645
ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 8(S5), 59–62, March 2015* DOI : 10.17485/ijst/2015/v8iS5/61480

Study on the Impact of Big Data Traffic Caused by the Unstable Routing Protocol

Hyoung woo Park*, Il Yeon Yeo, Haengjin Jang and NamGyu Kim

Supercomputing Center of Korea Institute of Science and Technology Information, 245 Darhak-ro, YuSeong-Gu, Daejeon-Si, 305-806, South Korea; hwpark@kisti.re.kr, ilyeon@kisti.re.kr, hjjang@kisti.re.kr, ssgyu@kisti.re.kr

Abstract

Thousands of CPUs and storages are required for the analysis of scientific peta-byte scale data. The network interfaces of these systems are connected on a data center network for cluster computing and Grid computing for high throughput. Therefore, routing configuration of these network interfaces is usually set by internal routing protocols and set by having the only single gateway as a last resort or a default routing path. This style of routing configuration is the most typical configuration of routing protocols for the network of the legacy data center. And, as the volume of data is getting dramatically increasing due to the popularization of big data services, data center rapidly grow to have hundreds of thousands servers. Therefore, real routing path frequently changes by the fault of network devices or systems in large-scale data center environment even though data center networks works well apparently. The inefficiency of this kind of routing scheme caused by the instability of the internal routing protocol begins to appear as a considerable problem in big data center network management lately. In this paper, we showed the serious decrease of the throughput due to very short flapping time of the internal routing protocol by NS2 simulation.

Keywords: Default Ip Routing, Domestic Route Flapping, Large-Scale Data, Network Throughput

1. Introduction

Big data analysis requires lots of computing resource. For example, CERN LHC (Large Hadron Collider) produces multi-peta (10^{15}) byte of experimental data every year. These data analysis need tens of thousands of CPUs. Big data center for the efficient allocation of CPUs provides cluster computing[1] service and Grid computing[2] service. Therefore, most of data centers have multi-tiered network architecture for lots of computer connection and use IGP (Interior Gateway Protocol) as a local traffic control. Figure 1 shows tiered network architecture of legacy data centers. Typical IGPs that used in data centers are RIP, OSPF, and IGRP. Data traffic management in data center[3,4] is getting important because IGPs are configured very simply by the administrator of data centers for the internal tired network of big data center. The typical configuration of IGPs is accomplished by only 2 lines of routing configuration statements. One is for the declaration of local network. The other is for the definition of default gateway. This simple configuration of IGPs has a strong tendency to be applied to the all of CPUs in data center. To make matters worse, thousands of CPUs are configured in the same subnet because the administrator of data centers like to put lots of CPUs in a same network for the dynamic resource allocation.

2. Timer Based Routing Information Exchange for Data Center Tiered Network

The basic mechanism of internal routing protocol is the exchange of routing information between routing systems such as routers, hosts, and so on. The exchange of routing information is controlled by the timer. The timer sets the

*Author for correspondence

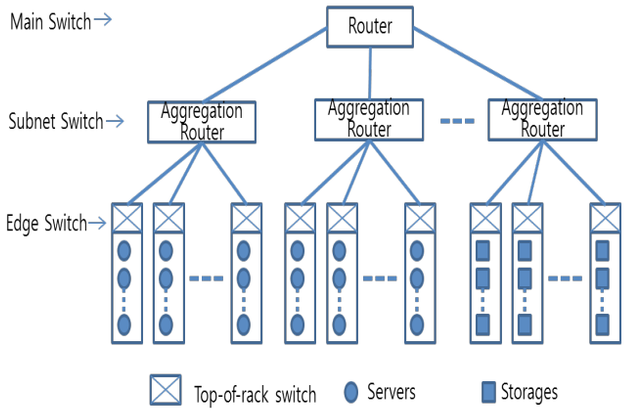

Figure 1. Tiered network architecture.

period of routing information by network administrator but most of routers and hosts are configured by default parameters of timers. There are update timer, invalid timer, hold down timer, and flush down timer for timers for the exchange of internal routing information.

2.1 Update Timer

The update timer controls the interval between two gratuitous Response Messages. By default the value is 30 seconds in the case of RIP. The response message is broadcast to all RIP enabled interfaces.

2.2 Invalid Timer

The invalid timer specifies how long a routing entry can be in the routing table without being updated. This is also called as expiration Timer. By default, the value is 180 seconds in the case of RIP. After the timer expires, the hop count of the routing entry will be set to 16, marking the destination as unreachable.

2.3 Hold-down Timer

Hold-down timer is started per route entry, when the hop count is changing from lower value to higher value. This allows the route to get stabilized. During this time no update can be done to that routing entry. This is not part of the RFC 1058. This is Cisco's implementation. The default value of this timer is 180 seconds.

2.4 Flush-down Timer

The flush timer controlling the time between the routes is invalidated or marked as unreachable and removal of entry from the routing table. By default the value is 240 seconds in the case of RIP. This is 60 seconds longer than Invalid timer. So for 60 seconds the router will be advertising about this unreachable route to all its neighbors. This timer must be set to a higher value than the invalid timer.

3. The Cause and the Impact of Instability of IGPs in Big Data Center

3.1 The Impact by Default IP Routing

Tables 1 and 2 show default values of timers those were set in routing protocols in legacy data center. In Table 2, Hello Timer is similar to Update Timer in Table 1 and Dead Interval and Invalid Timer is alike.

The impact of default routing can be enumerated by explaining the extension of the damaged data volume by the failure of a server's NIC and single failure that caused by the malfunction of a certain default route. The unit used in Table 1 and 2 is a second. Therefore, IGPs have to wait at least 180 second to recognize the problem of the network. When 10G Ethernet is used, 2Tb can be affected. The dynamic exchange of routing information can be inappropriate for the data center in the era of big data. Therefore, researches on internal routing architecture for big data center come to again as a new issue[5–7].

3.2 The Impact by Route Flapping of Internal Routing Protocol

The transmission delay time is less than 10ms in a data center network. The default value of timers mentioned-

Table 1. Default parameters of routing protocol timer (type 1)

	Update Timer	Invalid Timer	Hold down Timer	Flush Down Timer
RIP	30	180	180	240
IGRP	90	270	280	630

Table 2. Default parameters of routing protocol timer (type 2)

	Hello Timer	Dead Interval
OSPF	10	40
EIGRP	5	15

section 3.1 is much longer than 10ms. It is approximately more than ten thousand times of data transmission time. Therefore, route flapping can happen frequently within the default value of timers of IGPs. The invisible route flapping caused by the algorithm of internal routing protocols such as distance vector algorithm can reduce the throughput of the network of data center. The magnitude of this reduction can be magnified by big data. We tried to show this problem using by NS2 simulation[10].

4. Simulations and Result

The configuration of simulation for the demonstration of the problem of internal route flapping is consisted of 6 nodes. All of links showed in Figure 2 are set by 10ms for the simulation of local network. The scenario of simulation is that node n0 sends data to node n5 by FTP. TCP packet size is set by 9Kbyte for jumbo frame. TCP type for node n0 is new RENO. For route flapping, we make the link between node n1 and node n4 drop periodically. In these simulations, we made drops three times for the simulation of route flapping. It is showed at Table 3. In Table 3, unit is second and the value means time after simulation begins. A simulation is carried for 20 seconds. Other parameters in these simulations used default parameter of NS2. Table 4 showed throughput difference between the network with route flapping and the normal route network. Table 5 showed the result when node n0 is configured as remote node by adjusting the value of link delay between node n0 and node n1 to have 200ms. In the Table 4 and Table 5, the value is number of packets are received at node n5. Figure 3 showed the overall result of these simulations. From Figure 3, we can read the throughput with route flapping is less than from 10% to 30% than that of normal routing state

The default value of timers in IGPs is very high compared to the link delay in the network of big data center. Short period of route flapping in the big data center can invoke loss of cost for the management of big data center.

Table 3. Up and down time of routing protocol of the link between node n1 and n4 for the simulation of route flapping

Routing Protocol	Up Time	Down Time
Distance Vector Algorithm	1.0	4.0
	7.0	10.0
	13.0	16.0

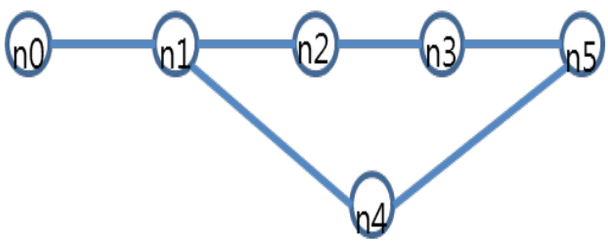

Figure 2. Simulation diagram of route flapping for big data impact measuring.

Table 4. Throughput comparison between route flapping network and normal routing network within local domain

Case	Normal Routing(A)	Unstable Routing(B)	Ratio (B/A)
10M/s	8103	7335	90.52%
100M/s	18656	13740	73.65%
1G/s	19476	16435	84.39%
10G/s	19596	16395	83.67%

Table 5. Throughput comparison between route flapping network and normal routing network beyond local domain

Case	Normal Routing(A)	Unstable Routing(B)	Ration (B/A)
10M/s	966	909	94.10%
100M/s	1070	993	92.80%
1G/s	1057	984	93.09%
10G/s	1078	987	91.56%

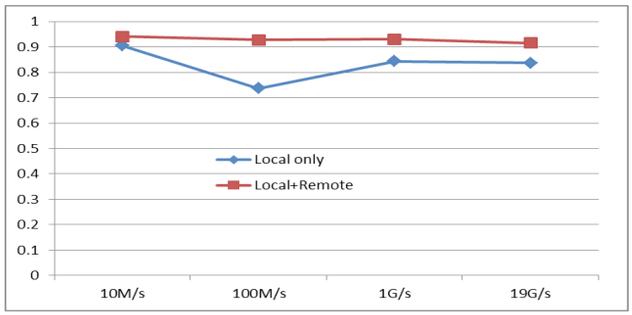

Figure 3. Throughput ratio between internal route flapping network and normal routing network.

5. Conclusion

As we mentioned above, routing configuration based on default parameters can threat the network stability of data center according to the increase of big data service. And route flapping within the domain routing can also cause the reduction of network throughput. We showed the impact degree with simulation that indicated the reduction with the range of from 10% to 30%. Big data centers that have hundreds of thousands of systems should have special monitoring systems that can check the change the route of IP packets. Until now, we ignore the asymmetry characteristics of IP route. We just concern the reachability of packets. Path that packet traverse is not a matter of concern. But, in the era of big data, we should monitor the stability of internal IP route[8,9] for the efficient management of the throughput for the local network of big datacenter, once again.

6. Acknowledgement

This early work was supported by the program of the Construction and Operation for Large-scale Science Data Center, 2014 funded by the KISTI and by the program of the Global hub for Experiment Data of Basic Science, 2014 funded by the NRF (N-14-NM-IR06).

7. References

1. Cloud computing - data center strategy: architecture, and solutions, point of view white paper for U.S. public sector. 2014. Available from: http://www.cisco.com/web/strategy/docs/gov/CiscoCloudComputing_WP.pdf
2. Foster I. What is the grid? A three point checklist. 2002; Available from: http://www.mcs.anl.gov/~itf/Articles/WhatIsTheGrid.pdf
3. Benson T, Anand A, Akella A, Zhang M. MicroTE fine grained traffic engineering for data centers. ACM CoNEXT; 2011.
4. Benson T, Akella A, Maltz A. Network traffic characteristics of data centers in the wild. IMC'10; 2010.
5. Zahavi E, Keslassy I, Kolodny A. Distributed adaptive routing for big-data applications running on data center networks. ANCS'12; 2012.
6. Soran A, Akdemir MF, Yuksel M. Parallel routing on multi-core routers for big data transfers. CoNEXT; 2013.
7. Park H, Yeo I, Jang H. Impact on the network stability of big data center by the configuration of default IP routing. ICCT; 2014.
8. Liu B, Sun Y, Cheng J, Zhang Y. Generic fault-avoidance routing protocol for data center networks. IETF; 2014. Report No.: draft-s1-rtgwg-far-dcn-00.
9. Francois P, Filsfils C, Evans J, Bonaventure O. Achieving sub-second IGP convergence in large IP networks. ACM SIGCOMM Computer Communication Review. 2005; 35(3):35–44.
10. NS2 simulator. 2014; Available from: http://www.isi.edu/nsnam/ns/