可扩展路由器控制平面的高性能通信模型

徐 恪, 吴 鲲, 王青青

徐 恪1, 吴 鲲1, 王青青2
1(清华大学 计算机科学与技术系,北京 100084)
2(兰州大学 信息科学与工程学院,甘肃 兰州 730000)
作者简介: 徐恪(1974－),男,江苏洪泽人,博士,副教授,主要研究领域为新一代互联网体系结构,高性能路由器体系结构.吴鲲(1976－),男,博士,主要研究领域为高性能路由器体系结构.王青青(1984－),女,硕士生,主要研究领域为Overlay网络,高性能路由器.
联系人: 徐 恪 Phn: +86-10-62795818 ext 6856, E-mail: xuke@tsinghua.edu.cn, http://network.cs.tsinghua.edu.cn/teacher/xuke/

Abstract

The inter-nodes communication bottleneck has been a key factor that restricts the large-scale expanding of scalable router's software architecture. To solve the problem, this paper introduces a transmission adapted sub-layer in the supporting model of traditional software architecture. Through feature extracting of the up-going data stream and pattern matching with registered task, data stream can then be classified and divided based on the content of information to increase the effective communication rate. The paper further analyzes the model's performance from task's three characteristics: Distribution rate, spread number and traffic rate, and provides an optimized task dispatching reference. It shows that the introduced transmission adapted sub-layer can reduce inter-layer redundant flow and extensibility bottleneck of communication. Finally, the presented experiment verifies the theoretical analysis.

摘要

可扩展路由器控制平面节点间通信的瓶颈问题是制约软件体系结构大规模扩展的关键因素.针对此问题,在传统的软件体系结构的支撑模型中引入了传输适配子层的结构,上行的数据流经特征抽取与已注册的任务进行模式匹配,从而完成了对控制信息流基于内容的分类与分流,提高了其有效通信率.进一步根据任务的分布率、分散数和流量率这3个特征对模型进行了性能分析,表明了适配层的引入可以消除面间冗余流量和通信的可扩展瓶颈.最后通过实验验证了理论分析的正确性.

References:

[1] Schaller RR. Moore's law: Past, present and future. IEEE Spectrum, 1997,34(6):52-59.

[2] Gupta P. Algorithms for routing lookups and packet classification [Ph.D. Thesis]. Stanford: Stanford University, 2000.

[3] Zhang XZ, Lu XC, Zhu PD, Peng W. A synchronization framework and critical algorithm maintaining single image of IP forwarding tables among cluster router's nodes. Journal of Software, 2006,17(3):445-453 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/17/445.htm

[4] Iyer S, McKeown N. Analysis of the parallel packet switch architecture. IEEE/ACM Trans. on Networking, 2003,11(2):314-324.

[5] Keslassy I, Chuang ST, Yu K, Miller D, Horowitz M, Solgaard O, Mckeown N. Scaling Internet routers using optics. In: Proc. of the 2003 Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communications. New York: ACM Press, 2003. 189-200.

[6] Chao HJ, Deng K, Jing Z. PetaStar: A petabit photonic packet switch. IEEE Journal on Selected Areas in Communications, 2003, 21 (7):1096-1112.

[7] Cheyns J, Develder C, Colle D, De Truck F, Lagasse PM, Demeester P. Clos lives on in optical packet switching. IEEE Communications Magazine, 2004,42(2):114-121.

[8] Kohler E, Morris R, Chen B. The click modular router. ACM Trans. on Computer Systems, 2000,18(3):263-297.

[9] Mosberger D, Peterson L. Making paths explicit in the scout operating system. In: Proc. of the 2nd USENIX Symp. on Operating System Design and Implementation (OSDI). 1996. 153-168.

[10] Decasper D, Dittia Z, Parulkar G, Plattner B. Router plugins: A software architecture for next generation routers. IEEE/ACM Trans. on Networking, 2000,8(1):2-15.

[11] Gottlieb Y, Peterson L. A comparative study of extensible routers. In: Proc. of the IEEE Open Architectures and Network Programming. 2002. 51-62.

[12] Wu K, Wu JP, Xu K. A tree-based distributed model for BGP route processing. In: Porc. of the Int'l Con. on High Performance Computing and Communications 2006. LNCS 4208, 2006. 119-128.

附中文参考文献:
[3] 张晓哲,卢锡城,朱培栋,彭伟.一种集群路由器转发表同步框架及关键算法.软件学报,2006,17(3):445-453. http://www.jos.org.cn/1000-9825/17/445.htm