P.O.Box 8718, Beijing 100080, China	Journal of Software, Jan. 2007,18(1):117-126
E-mail: jos@iscas.ac.cn	ISSN 1000-9825, CODEN RUXUEW, CN 11-2560/TP
http://www.jos.org.cn	Copyright © 2007 by *Journal of Software*

基于有状态Bloom filter引擎的高速分组检测

叶明江, 崔 勇, 徐 恪, 吴建平

叶明江, 崔 勇, 徐 恪, 吴建平
(清华大学 计算机科学与技术系,北京 100084)
作者简介: 胡叶明江(1982－),男,湖南怀化人,博士生,主要研究领域为网络安全,计算机网络体系结构,P2P覆盖网络.崔勇(1976－),男,博士,助理研究员,主要研究领域为计算机网络体系结构,协议的仿真和测试,多目标优化的路由算法及其评价.徐恪(1974－),男,博士,副教授,主要研究领域为新一代互联网体系结构,交换机和路由器体系结构,P2P与Overlay网络.吴建平(1953－),男,教授,博士生导师,CCF高级会员,主要研究领域为计算机网络体系结构,计算机网络协议测试.
联系人: 叶明江 Phn: +86-10-62785822, Fax: +86-10-62788109, E-mail: yemingjiang@csnet1.cs.tsinghua.edu.cn

Received 2005-11-08; Accepted 2006-04-03

Abstract
More and more network security applications depend on inspecting the content of the packets to detect the malicious attacks. To detect these attacks online, packet inspection demands exceptionally high performance. A lot of research works have been done in this field, and yet there is still significant room for improvement in throughput and scalability. This paper proposes a fast packet inspection algorithm based on state-based Bloom filter engines (SABFE). To achieve high throughput, parallel design is adopted when searching in one Bloom filter engine and between multiple Bloom filter engines. In addition, specific lookup table and prefix register heap are constructed in SABFE to keep the state of the matched prefix for the sake of detecting long patterns. The analysis and the evaluation show that the high throughput of the algorithm can satisfy the wire speed detection requirement when the low resource consumption in hardware resource further improves the scalability of SABFE.

Ye MJ, Cui Y, Xu K, Wu JP. Fast packet inspection using state-based Bloom filter engine. *Journal of Software*, 2007,18(1):117-126.
DOI: 10.1360/jos180117
http://www.jos.org.cn/1000-9825/18/117.htm

摘要
越来越多的网络安全技术通过分析网络分组中的内容来检测报文中是否含有恶意攻击代码.为了能够在线检测攻击,部署在路由器中的分组检测模块对于分组检测的速度也提出了越来越高的要求.虽然在这个领域已有很多研究工作,然而在性能、可扩展性和适用性方面还有很多可研究的空间.提出了一种基于有状态Bloom filter引擎的高速分组检测方法State-Based Bloom filter engine(SABFE).通过并行查找Bloom filter和前缀寄存器堆,以及利用多个并行的Bloom filter引擎进行流并行检测,达到了较高的吞吐性能.同时,利用快速查找表和前缀寄存器堆保存当前子串的匹配状态来检测长的规则.分析和模拟实验表明:该方法在规则长度增加时依然保持了较高的吞吐性能,可以实现线速的分组检测,同时,极大地减少了硬件资源开销,提高了可扩展性.

基金项目：Supported by the National Natural Science Foundation of China under Grant Nos.60473082, 60403035 (国家自然科学基金); the National Grand Fundamental Research 973 Program of China under Grant No.2003CB314801 (国家重点基础研究发展规划(973))

References:

[1] Moore D, Paxson V, Savage S, Shannon C, Staniford S, Weaver N. Inside the slammer worm. IEEE Security and Privacy, 2003, 1(4):33-39.

[2] Moore D, Shannon C. Code-Red: A case study on the spread and victims of an Internet worm. In: Proc. of the 2002 ACM SICGOMM

Internet Measurement Workshop. Marseille, 2002. 273-284. http://portal.acm.org/citation.cfm-id=637244&dl=ACM&coll=&CFID=15151515&CFTOKEN=6184618

[3] Kim HA, Karp B. Autograph: Toward automatic distributed worm signature detection. In: Proc. of the USENIX Security Symp. Diego, 2004. 271-286. http://www.usenix.org/events/sec04/tech/full_papers/kim/kim.pdf

[4] Singh S, EstanC, Varghese G, Savage S. Automated worm fingerprinting. In: Proc. of the 6th ACM/USENIX Symp. on Operating System Design and Implementation (OSDI). San Francisco, 2004. 45-60. http://www.usenix.org/events/osdi04/tech/full_papers/singh/singh.pdf

[5] Axelsson. Intrusion detection systems: A survey and taxonomy. Technical Report, 99-15, Chalmers University, 2000.

[6] Bloom B. Space/Time trade-offs in Hash coding with allowable errors. Communications of the ACM, 1970,13(7):422-426.

[7] Dharmapurikar S, Krishnamurthy P, Sproull T, Lockwood J. Deep packet inspection using parallel Bloom filters. In: Proc. of the Symp. on High Performance Interconnects (HotI). Stanford, 2003. 44-51. http://www.hoti.org/archive/Hoti11_program/papers/hoti11_07_dharmapurikar_s.pdf

[8] Dharmapurikar S, Attig M, Lockwood J. Design and implementation of a string matching system for network intrusion detection using FPGA-based Bloom filters. Technical Report, WUCSE-2004-12, St. Louis: Washington University, 2004.

[9] Song HY, Dharmapurikar S, Turner J, Lockwood J. Fast hash table lookup using extended Bloom filter: An aid to network processing. In: Proc. of the ACM SIGCOMM 2005. Philadelphia, 2005. 20-26. http://portal.acm.org/citation.cfm-id=1080114&dl=ACM&coll=&CFID=15151515&CFTOKEN=6184618

[10] Yu F, Katz RH, Lakshman TV. Gigabit rate packet pattern-matching using TCAM. In: Proc. of the 12th IEEE Int'l Conf. on Network Protocols (ICNP 2004). Berlin, 2004. 174-183. http://portal.acm.org/citation.cfm-id=1025890&dl=GUIDE&coll=GUIDE

[11] Lakshminarayanan K, Rangarajan A, Venkatachary S. Algorithms for advanced packet classification with ternary CAMs. In: Proc. of the ACM SIGCOMM 2005. Philadelphia, 2005. 193-204. http://portal.acm.org/citation.cfm-id=1080115&dl=ACM&coll=&CFID=15151515&CFTOKEN=6184618#

[12] Taylor DE. Survey and taxonomy of packet classification techniques. Technical Report, WUCSE-2004-24, St. Louis: Washington University, 2004.

[13] Baeza-Yates R. Algorithms for string searching: A survey. SIGIR Forum, 1989,23(3-4):34-58.

[14] Bakerand ZK, Prasanna VK. Time and area effificient pattern matching on FPGAs. In: Proc. of the 2004 ACM/SIGDA 12th Int'l Symp. on Field Programmable Gate Arrays (FPGA 2004). Monterey, 2004. 223-232. http://portal.acm.org/citation.cfm-coll=GUIDE&dl=GUIDE&id=968312#

[15] Tuck N, Sherwood T, Calder B, Varghese G. Deterministic memory-efficient string matching algorithms for intrusion detection. In: Proc. of the IEEE Infocom Conf. Hong Kong, 2004. 333-340. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp-arnumber=1354682