



[12] 发明专利申请公开说明书

[21] 申请号 03153072.9

[43] 公开日 2004年3月17日

[11] 公开号 CN 1482548A

[22] 申请日 2003.8.12 [21] 申请号 03153072.9

[30] 优先权

[32] 2002. 8. 12 [33] US [31] 10/218,950

[71] 申请人 国际商业机器公司

地址 美国纽约州

[72] 发明人 小埃弗里特·A·科尔
奥尔加·戈洛瓦尼斯基
克拉克·D·杰弗里斯
小罗伯特·F·坎特纳
兰德尔·W·沃泽拉
阿夫拉哈姆·齐哈维

[74] 专利代理机构 北京市柳沈律师事务所

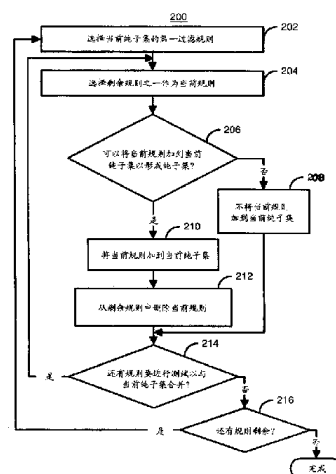
代理人 蒲迈文 黄小临

权利要求书4页 说明书9页 附图4页

[54] 发明名称 对多重搜索实行的过滤器规则进行划分的方法和系统

[57] 摘要

公开了一种用于在计算机系统中存储多条过滤器规则的方法和系统。所述多条过滤器规则在至少一个量纲中使用至少一个值范围。该方法和系统包括将所述多条过滤器规则中的一部分分离为多个过滤器规则纯子集。所述多个过滤器规则纯子集均包括所述多条过滤器规则中的至少一条。该方法和系统还包括组合所述多个过滤器规则纯子集中的一部分以提供组合子集，并且判定所述多条过滤器规则的使用效率是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善。该方法和系统还包括，如果所述多条过滤器规则的搜索效率得到改善，则存储所述多条过滤器规则，包括作为组合子集存储所述多个过滤器规则子集中的一部分。



1. 一种用于在计算机系统中存储多条过滤器规则的方法，所述多条过滤器规则在至少一个量纲中使用至少一个值范围，所述方法包括如下步骤：
- 5 (a) 将所述多条过滤器规则中的一部分划分为多个过滤器规则纯子集，所述多个过滤器规则纯子集均包括所述多条过滤器规则中的至少一条；
- (b) 组合所述多个过滤器规则纯子集中的一部分以提供组合纯子集；
- (c) 判定所述多条过滤器规则的使用效率是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善；以及
- 10 (d) 如果所述多条过滤器规则的使用效率得到改善，则存储所述多条过滤器规则，包括作为组合子集存储所述多条过滤器规则中的一部分。
2. 如权利要求1所述的方法，还包括如下步骤：
- (e) 如果所述多条过滤器规则的使用效率没有得到改善，则重新组合所述多个过滤器规则纯子集中的第二部分。
- 15 3. 如权利要求1所述的方法，其中，存储步骤(d)导致多个存储子集，所述方法还包括如下步骤：
- (e) 组合所述多个存储子集中的一部分，以提供第二组合子集；
- (f) 判定所述多条过滤器规则的使用效率是否通过组合所述多个存储过滤器规则子集中的一部分而得到改善；以及
- 20 (g) 如果所述多条过滤器规则的使用效率得到改善，则存储所述多个存储子集中的一部分，包括作为第二组合子集存储所述多个存储子集中的一部分。
4. 如权利要求3所述的方法，还包括如下步骤：
- (h) 重复步骤(e)、(f)和(g)，直到所述多条过滤器规则的使用效率不再得到改善。
- 25 5. 如权利要求3所述的方法，还包括如下步骤：
- (h) 重复步骤(e)、(f)和(g)，直到所述多条过滤器规则的使用效率达到特定级别。
6. 如权利要求1所述的方法，其中，组合步骤(b)还包括如下步骤：
- (b1) 组合所述多个过滤器规则纯子集中的一对。
- 30 7. 如权利要求1所述的方法，其中，效率判定步骤(c)还包括如下步骤：
- (c1) 判定实行所述多条过滤器规则的速度是否通过组合所述多个过滤器

规则纯子集中的一部分而得到改善。

8. 如权利要求1所述的方法, 其中, 效率判定步骤(c)还包括如下步骤:

(c1) 判定存储所述多条过滤器规则的代价是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善。

5 9. 如权利要求1所述的方法, 其中, 效率判定步骤(c)还包括如下步骤:

(c1) 判定实行所述多条过滤器规则的速度与存储所述多条过滤器规则的代价的组合是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善。

10 10. 一种计算机可读介质, 包含一种用于在计算机系统中存储多条过滤器规则的程序, 所述多条过滤器规则在至少一个量纲中使用至少一个值范围, 所述程序包括如下指令:

(a) 将所述多条过滤器规则中的一部分划分为多个过滤器规则纯子集, 所述多个过滤器规则纯子集均包括所述多条过滤器规则中的至少一条;

(b) 组合所述多个过滤器规则纯子集中的一部分以提供组合纯子集;

15 (c) 判定所述多条过滤器规则的使用效率是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善; 以及

(d) 如果所述多条过滤器规则的使用效率得到改善, 则存储所述多条过滤器规则, 包括作为组合子集存储所述多条过滤器规则中的一部分。

20 11. 如权利要求10所述的计算机可读介质, 其中, 所述程序还包括如下指令:

(e) 如果所述多条过滤器规则的使用效率没有得到改善, 则重新组合所述多个过滤器规则纯子集中的第二部分。

12. 如权利要求10所述的计算机可读介质, 其中, 存储指令(d)导致多个存储子集, 所述程序还包括如下指令:

25 (e) 组合所述多个存储子集中的一部分, 以提供第二组合子集;

(f) 判定所述多条过滤器规则的使用效率是否通过组合所述多个存储过滤器规则子集中的一部分而得到改善; 以及

(g) 如果所述多条过滤器规则的使用效率得到改善, 则存储所述多个存储子集中的一部分, 包括作为第二组合子集存储所述多个存储子集中的一部分。

30 13. 如权利要求12所述的计算机可读介质, 其中, 所述程序还包括如下指令:

(h) 重复指令(e)、(f)和(g),直到所述多条过滤器规则的使用效率不再得到改善。

14. 如权利要求12所述的计算机可读介质,其中,所述程序还包括如下指令:

5 (h) 重复指令(e)、(f)和(g),直到所述多条过滤器规则的使用效率达到特定级别。

15. 如权利要求10所述的计算机可读介质,其中,组合指令(b)还包括如下指令:

(b1) 组合所述多个过滤器规则纯子集中的一对。

10 16. 如权利要求10所述的计算机可读介质,其中,效率判定指令(c)还包括如下指令:

(c1) 判定实行所述多条过滤器规则的速度是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善。

15 17. 如权利要求10所述的计算机可读介质,其中,效率判定指令(c)还包括如下指令:

(c1) 判定存储所述多条过滤器规则的代价是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善。

18. 如权利要求10所述的计算机可读介质,其中,效率判定指令(c)还包括如下指令:

20 (c1) 判定实行所述多条过滤器规则的速度与存储所述多条过滤器规则的代价的组合是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善。

19. 一种用于在计算机系统中存储多条过滤器规则的系统,所述多条过滤器规则在至少一个量纲中使用至少一个值范围,所述系统包括:

25 处理器,用于将所述多条过滤器规则中的一部分划分为多个过滤器规则纯子集,所述多个过滤器规则纯子集均包括所述多条过滤器规则中的至少一条,并且组合所述多个过滤器规则纯子集中的一部分以提供组合子集;

性能评价工具,用于判定所述多条过滤器规则的使用效率是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善;以及

30 存储器,用于如果所述多条过滤器规则的使用效率得到改善,则存储所述多条过滤器规则,包括作为组合子集存储所述多个过滤器规则子集中的一

部分。

20. 如权利要求 19 所述的系统，其中，如果所述多条过滤器规则的使用效率没有得到改善，则处理器还重新组合所述多个过滤器规则纯子集中的第二部分。

5 21. 如权利要求 19 所述的系统，其中，存储器包括多个存储子集，并且处理器还组合所述多个存储子集中的一部分，以提供第二组合子集，其中，性能评价工具还判定所述多条过滤器规则的使用效率是否通过组合所述多个存储过滤器规则子集中的一部分而得到改善，并且如果所述多条过滤器规则的使用效率得到改善，则存储器还存储所述多个存储子集中的一部分，包括
10 作为第二组合子集存储所述多个存储子集中的一部分。

22. 如权利要求 21 所述的系统，其中，处理器还重新组合所述多个存储过滤器规则子集中的另一部分，并且性能评价工具还重新判定所述多条过滤器规则的使用效率是否通过重新组合所述多个存储过滤器规则子集中的其他部分而得到改善，直到所述多条过滤器规则的使用效率不再得到改善。

15 23. 如权利要求 22 所述的系统，其中，处理器还重新组合所述多个存储过滤器规则子集中的另一部分，并且性能评价工具还重新判定所述多条过滤器规则的使用效率是否通过重新组合所述多个存储过滤器规则子集中的其他部分而得到改善，直到所述多条过滤器规则的使用效率达到特定级别。

24. 如权利要求 19 所述的系统，其中，处理器通过组合所述多个过滤器
20 规则纯子集中的一对来组合所述多条过滤器规则中的一部分。

25. 如权利要求 19 所述的系统，其中，性能评价工具通过判定实行所述多条过滤器规则的速度是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善，判定所述多条过滤器规则的使用效率。

26. 如权利要求 19 所述的系统，其中，性能评价工具通过判定存储所述
25 多条过滤器规则的代价是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善，判定所述多条过滤器规则的使用效率。

27. 如权利要求 19 所述的系统，其中，性能评价工具通过判定实行所述
30 多条过滤器规则的速度与存储所述多条过滤器规则的代价的组合是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善，判定所述多条过滤器规则的使用效率。

对多重搜索实行的过滤器规则进行
划分的方法和系统

5

技术领域

本发明涉及计算机系统，特别涉及一种用于更高效搜索过滤器规则以与键进行匹配的方法和系统。

10

背景技术

图 1 示出可以连接到因特网 30 的传统网络 10 和 20。网络 10 包括主机 12、14 和 16，而网络 20 包括主机 22 和 24。各网络 10 和 20 还分别包括交换机 18 和 26，并且可以分别包括一个或多个服务器如服务器 17、19 和 28。另外，网络 10 和 20 还可以分别包括一个或多个接到因特网 30 的网关 13 和 25。路由器还可以控制网络 10 和 20 上的通信消息(traffic)的网络 10 和 20 的其他部分未被显式示出，它们将被视作分别由交换机 18 和 26 所固有描述的以及更一般的网络 10 和 20。

图 2 示出典型交换机 50 的一部分，它可以用于交换机 18 和 26 以及/或者路由器(未示出)。交换机 50 包括网络处理器 52 和存储器 54。交换机 50 典型地还包括其他组件(未示出)。网络处理器 52 管理交换功能，包括使用下述规则对分组 (packet) 进行分类。存储器 54 保存与规则相关的数据。

参照图 1 和 2，为了管理网络如网络 10 或 20 中的通信，使用过滤器规则。过滤器规则典型地由交换机、路由器和用来执行分组分类的网络其他部分使用。各过滤器规则用来对正在通过网络传输的分组进行分类，从而确定如何处理分组以及执行什么服务。例如，过滤器规则可以用于测试从外部源进入网络的分组以确保可以阻止闯入网络的企图。例如，可以测试从因特网 30 进入网络 10 的通信消息，从而确保拒绝来自未经授权源的分组进入。类似地，可以防止来自网络一部分的分组访问网络的另一部分。例如，可以防止来自某些主机 12、14 或 16 的分组访问服务器 17 或服务器 19。还可以记录主机试图连接服务器的情况，从而网络所有者可以采取适当的操作。这些过滤器规则还可以用来根据分组的优先级传输通信消息。例如，即使当来自主机 14 或

16 的分组可能丢失时,也可以传输来自特定主机如主机 12 的分组,因为该分组具有更高优先级。过滤器规则还可以用来确保当非常拥挤时即使传输来自已建会话的通信消息也不允许开始新会话。还可以根据过滤器规则实现其他功能。

5 过滤器规则典型地还具有优先级。过滤器规则还可以根据各过滤器规则的优先级来交互。过滤器规则的优先级可以用来确定当键与两条或多条过滤器规则的范围匹配时所采取的操作。在这种情况下,具有较高优先级的过滤器规则控制所采取的操作。例如,第一规则可以是处理大多数情况的缺省规则。第二规则可以是第一规则的例外。第二规则典型地将具有比第一规则更高的优先级,以确保在分组与第一和第二规则都匹配的情况下将实行第二规则。在传统系统中,所有过滤器规则根据它们的优先级排在一个列表中。另外,在传统系统中,各过滤器规则的不同优先级由其在列表中的位置来反映。因此,在传统系统中,优先级数与过滤器规则数相同。因此,在传统系统中,过滤器规则的优先级数目大。

15 为了判定特定规则是否将作用于特定分组,对键进行测试。典型所用的键是分组的网际协议(IP)首标和分组的其他首标。IP 首标典型地包含 IP 源地址(SA)和 IP 目的地址(DA)以及协议类型(P)字段。IP 分组有效荷载典型地封装另一首标和有效荷载结构,并且内部结构典型地是传输控制协议(TCP)、用户数据报协议(UDP)或其他标准结构。TCP、UDP 或其他标准结构的首标又包含源端口(SP)和目的端口(DP)值。这些字段典型地排序为 SA、DA、SP、DP 和 P,并且分别为 32、32、16、16 和 8 位。规则典型地作用于这些字段中的一个或多个。例如,根据源和/或目的地址,规则可以判定来自特定主机的分组是否允许到达特定目的地址。

25 此外,键经常还包含不同于上述字段的另外位。例如,开始会话的 TCP 同步分组可以具有与现有会话的 TCP 数据分组不同的特征。该特征使用上述字段之外的位来体现。该另外位可以由管理网络上的通信消息的过滤器规则使用。例如,当网络拥挤时,过滤器规则可以主动丢弃 TCP 同步分组,而传输现有会话的 TCP 分组。这些操作允许网络继续工作并且帮助减轻拥挤。然而,为了执行该功能,规则检查 TCP 首标中的同步位以判定分组是否是 TCP 同步过程的一部分。因此,过滤器规则典型地使用包括分组的 IP 首标和其他首标的至少一些字段并且可以包括另外位的键来工作。

在测试键中,判定是否对特定分组实行过滤器规则从而对分组进行分类。键通过判定键的某些字段是否在规则的范围内来测试。各规则在一个或多个量纲中包含一个值范围或一个确切值。每个量纲对应于键的一个字段(典型地是IP首标或其他首标)。一种过滤器规则字段范围可以是单个值。在这种情况下,键将必须确切匹配规则值以作用于分组。其他规则具有某一字段范围,该范围由带有以固定最高有效位开始的特定数目连续位和以任意最低有效位结束的其他连续位的所有二进制值组成。在这种情况下,固定位称作前缀。前缀是包含一定数目的1和0(1或0)的二进制数,其后面是占位符或通配符(*)。范围的下界通过用0替换所有通配符来获得。范围的上界通过用1替换所有通配符来确定。其他规则具有任意范围。任意范围是不能使用单个前缀表达的范围。然而,任意范围可以使用多个前缀来表达。

各种机制可以用于搜索过滤器规则以与键进行匹配。例如,如果过滤器规则的特定量纲要求确切匹配,则搜索可以包括搜索过滤器规则哈希值(hash value)的表。搜索其他过滤器规则以与键进行匹配可能要求遍历二元决策树或其他机制。根据用来搜索过滤器规则的机制以及过滤器规则,搜索效率可能不同。

因此,需要一种改善搜索过滤器规则以与键进行匹配的效率的系统和方法。本发明就是解决这一需要。

20 发明内容

本发明提供一种用于在计算机系统中存储多条过滤器规则的方法和系统。所述多条过滤器规则在至少一个量纲中使用至少一个值范围,其中,量纲也称作字段。该方法和系统包括将所述多条过滤器规则中的一部分分离为多个过滤器规则纯子集。所述多个过滤器规则纯子集均包括所述多条过滤器规则中的至少一条。该方法和系统还包括组合所述多个过滤器规则纯子集中的一部分以提供组合子集,并且判定所述多条过滤器规则的使用效率是否通过组合所述多个过滤器规则纯子集中的一部分而得到改善。如果结果是所得到子集的实行速度和/或存储效率仍可以作进一步的改善,则另外重新组合过滤器规则的纯子集或非纯子集。该方法和系统还包括如果搜索和/或存储所述多条过滤器规则的效率得到改善,则存储包括经过组合的所述多个过滤器规则子集中的一部分的所述多条过滤器规则。

根据在此公开的系统和方法，本发明提供一种划分过滤器规则子集的机制，它可以改善用于与键进行匹配的过滤器规则的存储效率和/或实行速度。

附图说明

- 5 图 1 是其中可以使用本发明的网络计算机系统的图；
图 2 是其中可以使用本发明的交换机的图；
图 3 是示出用于改善实行效率的存储过滤器规则的本发明系统的一个实施例的方框图；
图 4 是示出用于在计算机系统中存储与过滤器规则相关的数据的本发明
10 方法的一个实施例的高层流程图；
图 5 是示出用于将计算机系统中的过滤器规则划分为纯子集的本发明方法的一个实施例的流程图；
图 6 是示出用于对过滤器规则纯子集进行排序的本发明方法的一个实施例的流程图；
15 图 7 是示出用于测试过滤器规则的组合以改善效率的本发明方法的一个实施例的流程图。

具体实施方式

- 20 本发明涉及分组分类的一种改进。下面描述让本领域的普通技术人员能够实现和使用本发明，并且它是在专利申请及其要求的上下文中提供的。对优选实施例的各种修改对于本领域的技术人员而言将是明显的，并且在此所述的一般原理可以应用于其他实施例。因此，本发明不限于所示实施例，而是具有与在此所述的原理和特性一致的最广范围。

- 25 本发明提供一种用于在计算机系统中存储多条过滤器规则的方法和系统。所述多条过滤器规则在至少一个量纲或字段中使用至少一个值范围。该方法 and 系统包括将所述多条过滤器规则中的一部分分离为多个过滤器规则纯子集。所述多个过滤器规则纯子集均包括所述多条过滤器规则中的至少一条。该方法和系统还包括组合所述多个过滤器规则纯子集中的一部分以提供组合子集，并且判定所述多条过滤器规则的使用效率是否通过组合所述多个过滤器
30 规则纯子集中的一部分而得到改善。如果结果是所得到子集的实行速度和/或存储效率仍可以作进一步的改善，则另外重新组合过滤器规则的纯子集或

非纯子集。该方法和系统还包括如果搜索和/或存储所述多条过滤器规则的效率得到改善，则存储包括经过组合的所述多个过滤器规则子集中的一部分的所述多条过滤器规则。

5 本发明将按照一组特定规则来描述。然而，本领域的普通技术人员应该清楚，该方法和系统将有效用于使用其他范围的其他规则。本领域的普通技术人员还应该清楚，该方法和系统有效用于使用其他任务规则的其他系统。然而，绝不反对将该方法和系统应用于其他规则。

为更具体阐述本发明的方法和系统，现在参照图 3，图 3 示出了利用三元内容可寻址存储器(TCAM)来存储与过滤器规则相关的数据的本发明系统
10 100 的一个实施例的方框图。系统 100 最好用于网络如网络 10 和 20 中。更具体地说，系统 100 最好用于如交换机 18 或 26 的组件中。然而，系统 100 可以用于计算机系统的其他部分中。这样，系统 100 与过滤器规则一起使用。过滤器规则在一个或多个量纲中具有范围。这些量纲对应于键的字段，包括 IP 首标或其他首标的字段。各范围可以通过一个或多个前缀来描述。

15 系统 100 包括处理器 102、存储器 104 和性能评价工具(“PET 106”)。存储器 104 最好类似于存储器 54，并且用于存储与过滤器规则相关的数据。处理器 102 最好执行下述功能如将过滤器规则划分为纯子集以及帮助组合过滤器规则的纯子集。过滤器规则的纯子集是带有标为 $i=1, 2, 3, \dots, N$ 的 N (典型为 5)个字段的过滤器规则集。对于每个字段 i ，必须只占三种可能性中的一种。
20 各规则中的字段 i 可以是对于各规则完全一致的可能不同值；各规则中的字段 i 可以完全是通配符；或者各规则中的字段 i 可以是同一公共前缀。例如，具有下面字段的下列过滤器规则一起形成一个纯子集。

表 1

规则	源地址	目的地址	源端口	目的端口	协议
R1	1.2.3.4	5.6.7.8	80	*	6
R2	1.2.3.4	5.6.7.20	80	*	6
R3	1.2.3.4	5.6.7.52	80	*	6
R4	1.2.3.4	5.6.7.79	80	*	6
R5	1.2.3.4	5.6.7.113	80	*	6

25 过滤器规则 R1、R2、R3、R4 和 R5 具有相同的源地址、源端口和协议。

R1、R2、R3、R4 和 R5 的这些字段还要求确切匹配。过滤器规则 R1、R2、R3、R4 和 R5 的目的地址全部通过同一前缀 5.6.7 来描述。所有过滤器规则 R1、R2、R3、R4 和 R5 的目的端口可以为任何值。因此，过滤器规则 R1、R2、R3、R4 和 R5 形成一个纯子集。

5 PET 106 可以用来评价使用特定过滤器规则集的系统如交换机的性能。PET 106 最好考虑对过滤器规则执行的搜索类型、正被搜索的过滤器规则数目以及用来执行搜索的机制如二元决策树或 TCAM。PET 106 用来如下所述评估过滤器规则不同组合的效率。在一个优选实施例中，PET 106 测量存储过滤器规则的代价和/或实行过滤器规则的速度。

10 图 4 是示出用于在计算机系统中存储与过滤器规则相关的数据的本发明方法 150 的一个实施例的高层流程图。方法 150 最好使用系统 100 来实现。对要用于计算机系统例如交换机 18 或 26 中的过滤器规则集执行方法 150。通过步骤 152，将过滤器规则划分为纯子集。步骤 152 使用处理器 102 来执行。通过步骤 154，尝试性地组合一些纯子集来提供组合子集。在一个优选实施例中，在步骤 154 尝试性地组合两个过滤器规则子集。步骤 154 最好首先组合规则(元素)最多的子集。通过步骤 156，判定过滤器规则的使用效率是否通过组合得到改善。在一个优选实施例中，过滤器规则的使用效率通过存储过滤器规则的代价或实行过滤器规则的速度来给出。过滤器规则的使用效率也可以是存储过滤器规则的代价和实行过滤器规则的速度组合。如果过滤器规则的使用效率得到改善，则通过步骤 158，以在步骤 154 获得的组合存储过滤器规则。否则，不存储组合。然后，通过步骤 160，可以再次执行步骤 154-158 以进一步改善过滤器规则的搜索性能。在一个优选实施例中，重复步骤 154-158，直到组合过滤器规则各部分不再改善性能为止。然而，在另一个实施例中，可以使用另一目标。例如，可以重复步骤 154-158，直到达到特定性能级别或者进一步组合所实现的性能改善处于或低于特定级别为止。因此，重复步骤 154-158，直到达到期望性能目标。

25 因此，使用方法 150 和系统 100，将过滤器规则划分为纯子集。搜索纯子集典型地更简单和更高效。例如，在表 1 的规则纯子集中，哈希处理可以忽略除目的地址字段之外的所有键字段。另外，联合通过重复应用方法 150 而形成的子集(形成某对的合并)，直到进一步的性能改善不可能或者满足某种其他标准。如果子集组合不提高过滤器规则搜索效率，则在方法 150 中不存储

组合。然而，如果组合提高效率，则一起存储从而一起搜索组合子集。因此，改善使用方法 150 和/或系统 100 的计算机系统如交换机 18 或 26 的性能。

图 5 是示出用于将计算机系统中的过滤器规则划分为纯子集的本发明方法 200 的一个实施例的流程图。通过步骤 202，选择一条过滤器规则作为当前纯子集的一部分。步骤 202 最好还包括从计算机系统的过滤器规则列表中删除该过滤器规则。通过步骤 204，选择剩余过滤器规则之一作为当前规则。通过步骤 206，判定当前规则是否加到当前纯子集以形成一个新的纯子集。因此，步骤 206 可以包括，测试当前规则，以确保当前过滤器规则的字段要求确切匹配且与当前子集中过滤器规则的相应字段相同，可以通过与子集中过滤器规则的相应字段的前缀相同的前缀来描述，或者可以与子集中过滤器规则的相应字段一样全部采用通配符来描述。

如果判定当前过滤器规则不能加到当前纯子集并且形成新的纯子集，则通过步骤 208 不将当前规则加到当前纯子集。如果当前过滤器规则可以加到当前纯子集来形成新的纯子集，则通过步骤 210 将当前规则加到当前纯子集。另外，通过步骤 212，从规则列表中删除当前规则。

通过步骤 214，判定是否存在任何剩余规则要进行测试以与当前纯子集合并。如果是，则返回到步骤 204，并且选择另一条当前过滤器规则。如果测试完毕所有过滤器规则以包括在当前纯子集中，则通过步骤 216 判定是否存在任何过滤器规则剩余(未被删除)。如果否，则方法 200 结束。否则，返回到步骤 202，并且形成另一个纯子集。

因此，使用方法 200，可以将过滤器规则划分为纯子集。因此，方法 200 最好用来执行图 4 所示的方法 150 的步骤 152。回到图 5，方法 152 将计算机系统的过滤器规则划分为考虑计算机系统的所有过滤器规则所需的多个纯子集。一旦将过滤器规则划分为纯子集，则可以如下所述组合纯子集。

图 6 是示出用于对过滤器规则的纯子集进行排序的本发明方法 230 的一个实施例的流程图。在一个优选实施例中，方法 230 用来准备过滤器规则的纯子集以在图 4 所示的方法 150 的步骤 154 进行组合。回到图 6，通过步骤 232，确定各纯子集中的元素数。然后，通过步骤 234，可以丢弃空子集。通过步骤 236，以非递增元素数依次列出纯子集。因此，使用方法 230，以条目最多的纯子集为开始且以条目最少的纯子集为结束列出各纯子集。结果，当结合方法 230 执行步骤 154 或下述方法 250 时，首先将测试元素最多的纯子

集以与其他子集合并。因此，方法 230 可以允许方法 150 和 250 更快速地导致提高存储效率或执行过滤器规则以与键进行匹配的搜索速度(使用效率)的纯子集组合。

图 7 是示出用于测试过滤器规则的组合以改善效率的本发明方法 250 的一个实施例的流程图。在一个优选实施例中，方法 250 可以用来执行图 4 所示的方法 150 的步骤 154-160。回到图 7，方法 250 用于组合纯子集对。然而，绝不反对本发明的方法和系统的另一个实施例以不同数目的子集形成组合。通过步骤 252，确定在纯子集中搜索过滤器规则的代价。在步骤 252 确定的代价可以当作基线。通过步骤 254，选择一对子集，并且形成它们的合并，以尝试性地生成其中少一个子集的新子集组。通过步骤 256 确定相对于先前子集的搜索子集对(组合)的速度和/或存储新子集组的代价。步骤 256 最好使用 PET 106 来执行，并且考虑可能使用的不同类型的搜索引擎。例如，两个纯子集均可以使用全匹配算法来搜索。然而，由于两个纯子集的组合可能不是纯子集，因此，组合可能要求更复杂的搜索机制如二元决策树。如果搜索代价和/或存储纯子集对合并的效率小于搜索总代价和/或存储两个纯子集的效率，则通过步骤 258 以纯子集的组合代替单独纯子集。步骤 258 最好对所有纯子集对都执行。在一个优选实施例中，在首次执行步骤 258 之后，当与特定其他纯子集放在一起时更高效搜索的过滤器规则纯子集得到组合。然后，通过步骤 260，重复步骤 252-258 以反复组合子集，直到进一步组合不改善搜索速度和/或存储代价。在随后重复中，步骤 252-258 作用于子集组合而不是纯子集对。例如，当首次重复步骤 252-258 时，在前一趟保存的纯子集对可以与另一纯子集或另一纯子集对组合。下次重复步骤 252-258 时，可以组合子集对的对。该过程最好继续直到进一步组合不再可以改善搜索代价为止。

因此，使用方法 250，典型地组合纯子集，然后如果性能改善，则重新组合。如果判定搜索和/或存储两个子集合并比单独搜索和/或存储一对子集更高效，则以子集合并代替两个单独子集。对于其他子集对和更高级别的子集组合，重复该过程。结果，所保存的规则最终组合具有由 PET 106 测量的更低存储代价，更快速度或某种其他改善。因此，可以改善使用过滤器规则的计算机系统的性能。

上面公开了一种用于划分和存储过滤器规则的方法和系统。根据本发明编写的软件要存储在某种形式的计算机可读介质如存储器、CD-ROM 中，或

者通过网络传输，并且由处理器执行。因此，计算机可读介质包括例如可以通过网络传输的计算机可读信号。虽然本发明是根据所示实施例来描述的，但是本领域的普通技术人员应该清楚，对这些实施例可以进行改变，并且这些改变将在本发明的精神和范围内。因此，在不脱离所附权利要求的精神和范围的情况下，本领域的普通技术人员可以进行各种修改。

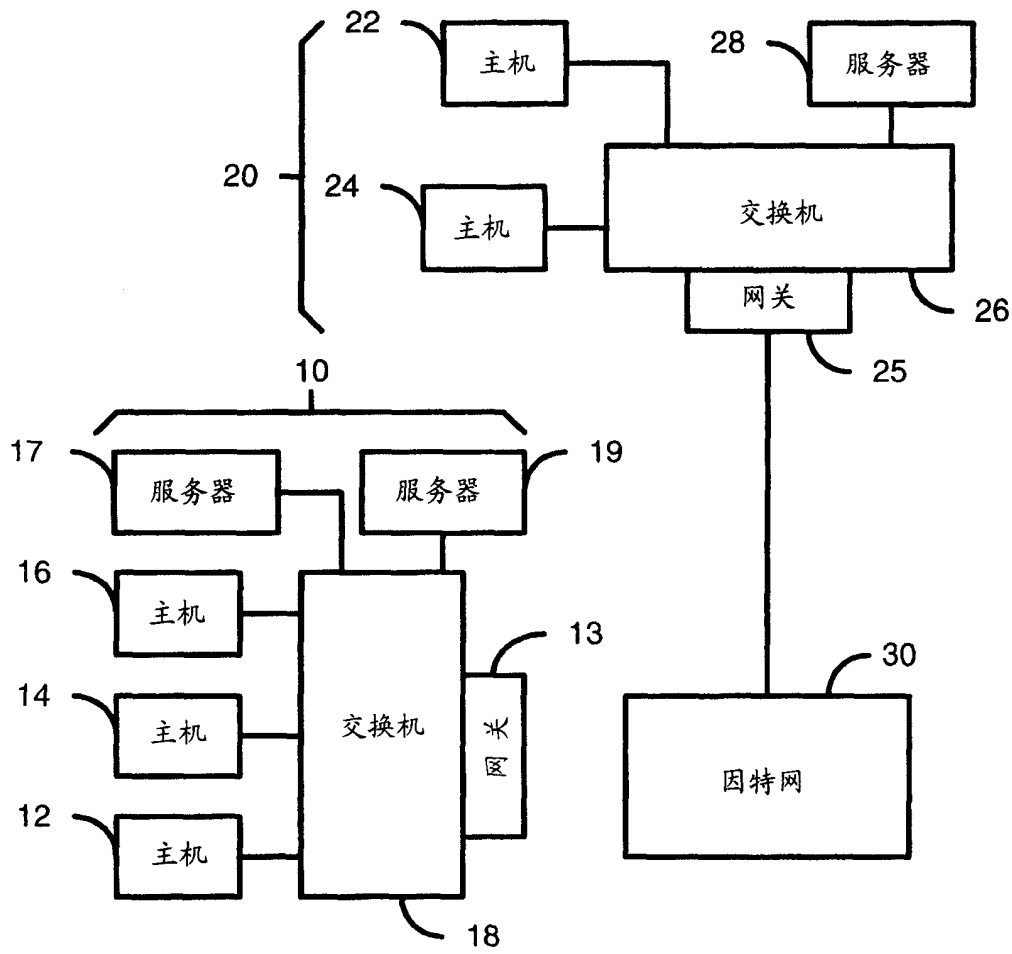


图 1

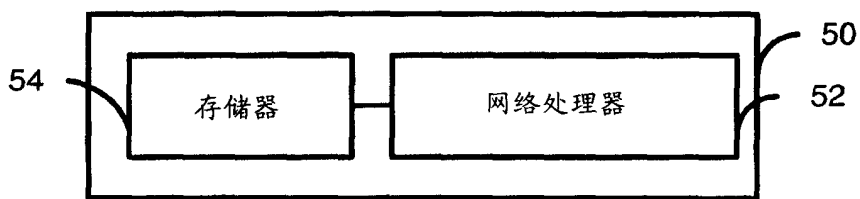


图 2

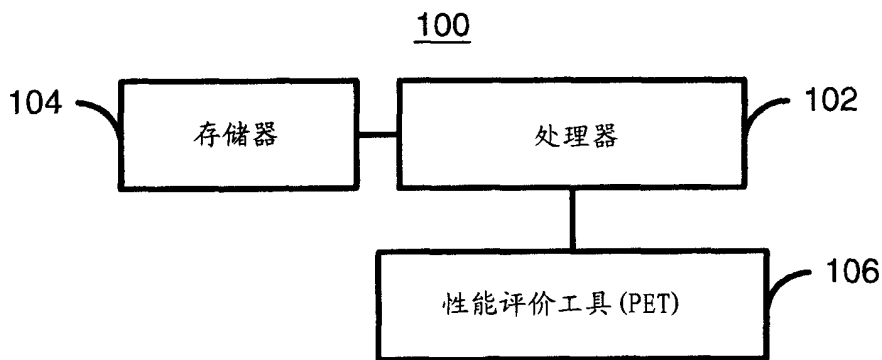


图 3

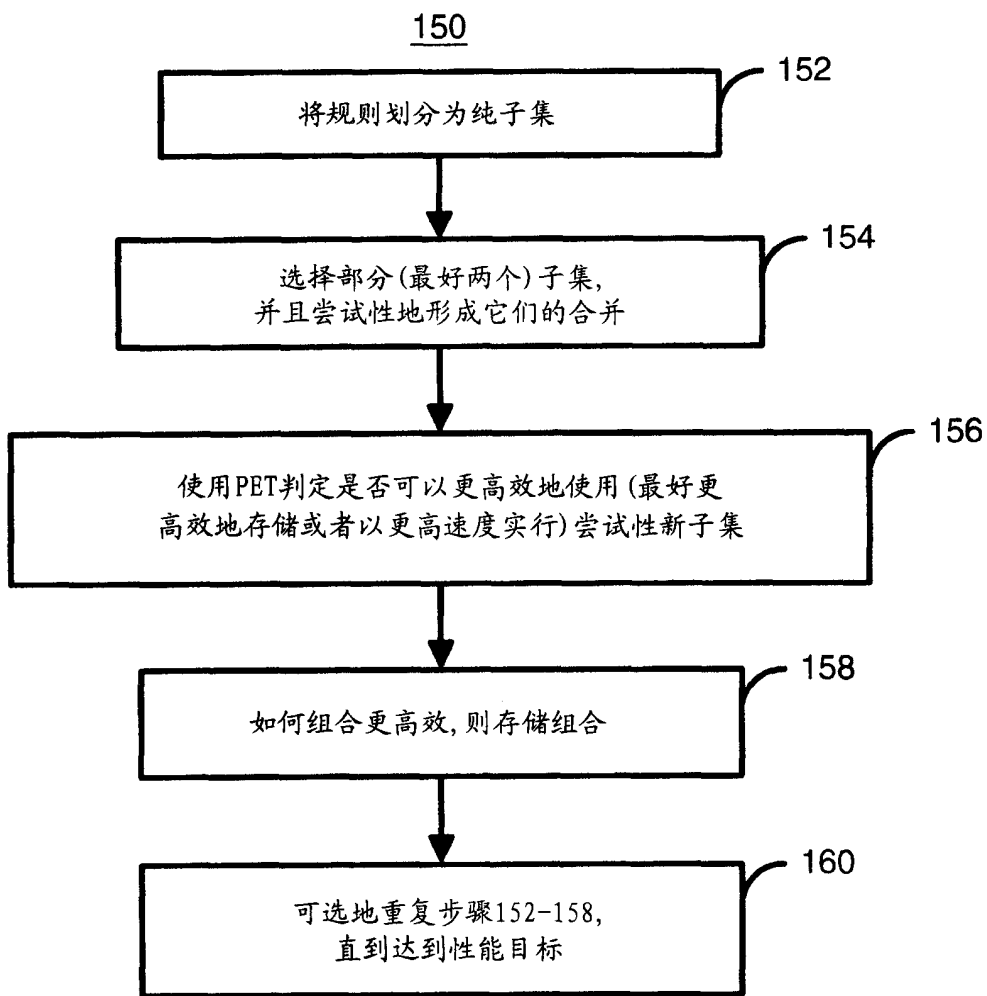


图 4

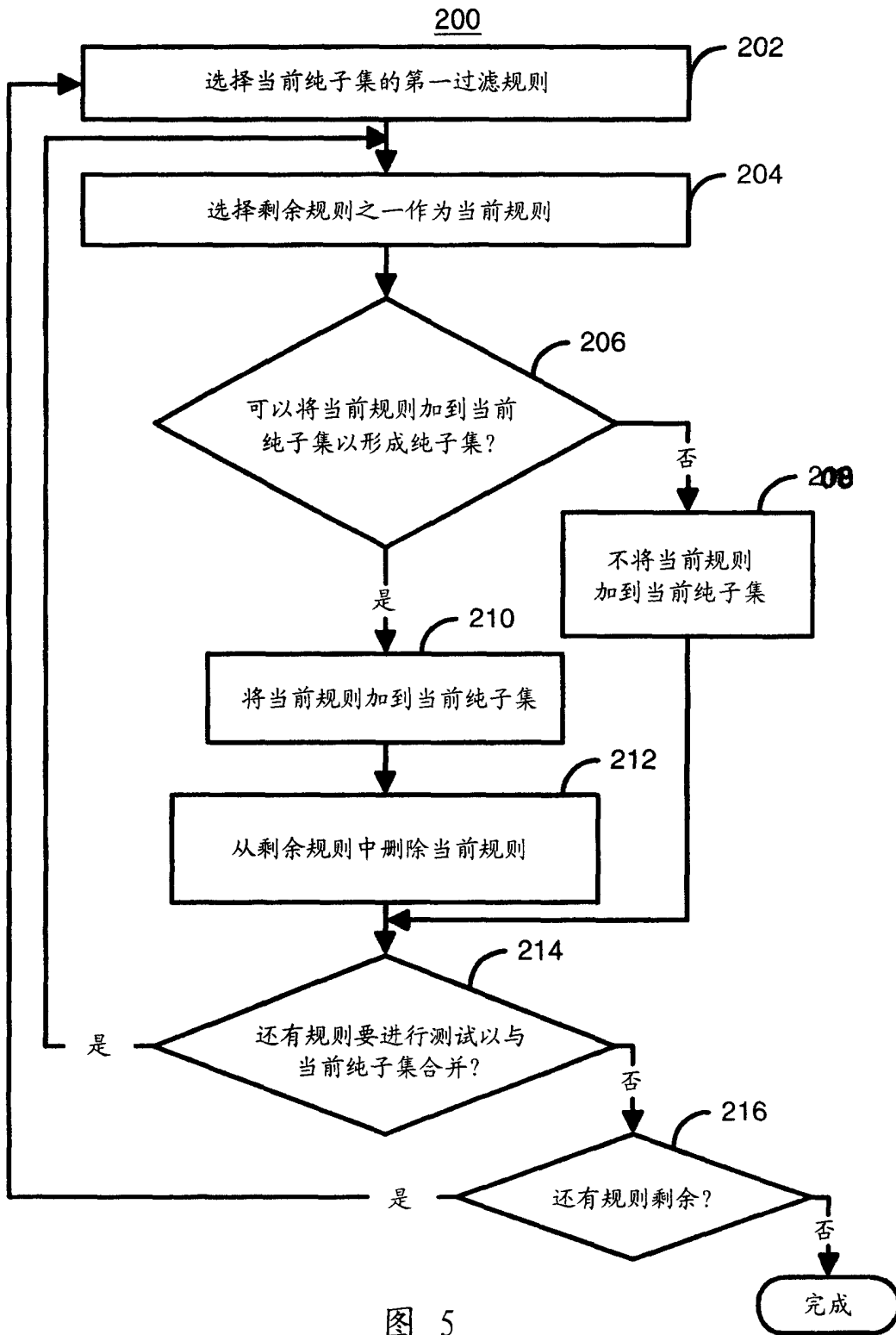


图 5

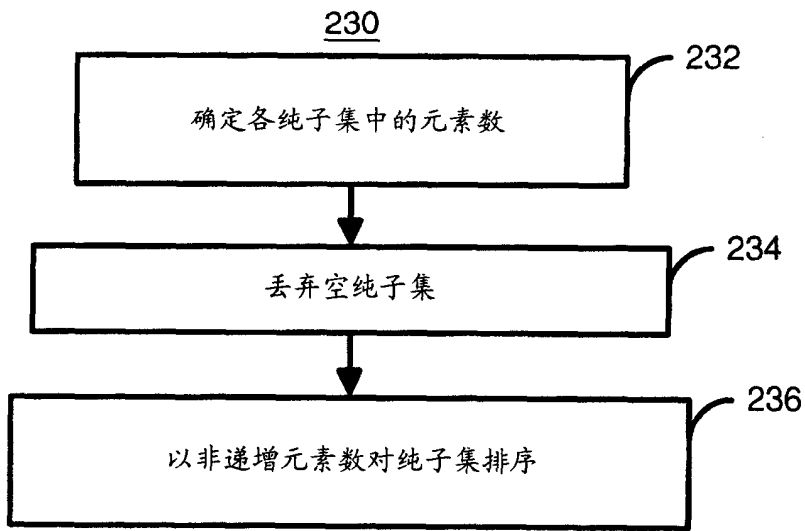


图 6

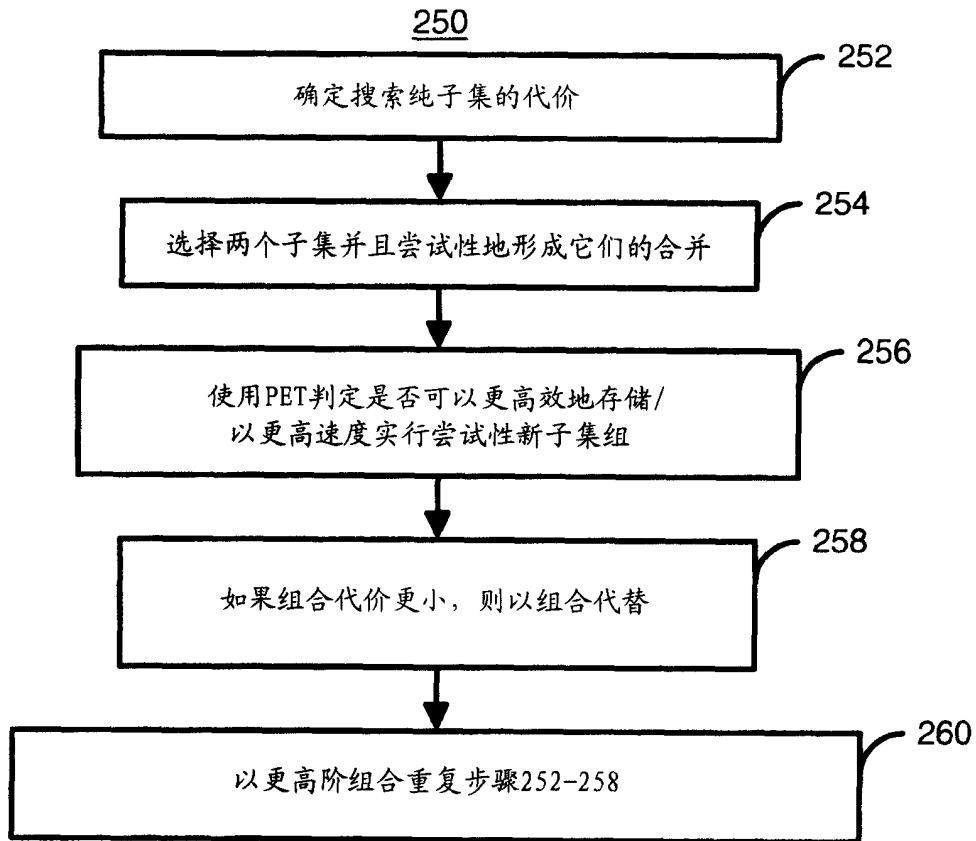


图 7