[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
H04L 9/00 (2006.01)

[12] 发明专利申请公布说明书

[21] 申请号 200680034493.4

[43] 公开日 2009 年 10 月 28 日

[11] 公开号 CN 101569129A

[22] 申请日 2006.7.27
[21] 申请号 200680034493.4
[30] 优先权
[32] 2005. 7.29 [33] US [31] 11/193,292
[32] 2005. 7.29 [33] US [31] 11/193,295
[32] 2005. 7.29 [33] US [31] 11/194,075
[32] 2005. 7.29 [33] US [31] 11/193,291
[32] 2005. 7.29 [33] US [31] 11/194,078
[86] 国际申请 PCT/US2006/029714 2006.7.27
[87] 国际公布 WO2007/016478 英 2007.2.8
[85] 进入国家阶段日期 2008.3.19
[71] 申请人 BIT9 公司
地址 美国马萨诸塞
[72] 发明人 托德·F·布伦南 艾伦·希拉里 约翰·汉拉蒂
[74] 专利代理机构 中国国际贸易促进委员会专利商标事务所
代理人 杜 娟
权利要求书 22 页 说明书 45 页 附图 12 页

[54] 发明名称

网络安全系统和方法

[57] 摘要

一种安全系统，防御已知和未知的病毒、蠕虫、间谍件、黑客和不需要的或者未知的软件。 所述系统可以实现集中的策略，其允许管理员批准、阻止、隔离或者记录文件行为。 所述系统在主机和在服务器中保存文件元信息。 主机检测可以引起对于文件内容或者文件名的改变的文件操作，并且作为结果更新主机和/或服务器元信息。 服务器元信息上的改变可以被主机获得。

1. 一种在系统中使用的方法，所述系统具有服务器和一个或多个相关联的主计算机（主机），所述方法包括:

在服务器上对于多个文件保存一组服务器元信息，其中对于每个唯一文件内容签名包括文件的内容的签名、文件或者签名被第一次报告到服务器的日期、以及用于指示是否和在什么条件下可以由主机对于文件执行特定的文件操作的状态数据;

在主机上对于多个文件保存一组元信息，其中，对于每个文件包括状态数据和文件内容的签名;

在主机上检测对于文件内容或者名称的可能改变，并且更新主机和/或服务器元信息；以及

服务器向主机提供服务器元信息中的改变。

2. 按照权利要求 1 的方法，其中，所述内容的签名包括文件的内容的一个或多个加密散列值的结果。

3. 按照权利要求 1 的方法，还包括：对于每个主机，维护具有文件名和状态数据的独立名称高速缓冲存储器。

4. 按照权利要求 3 的方法，其中，响应于对于文件的文件操作的请求，主机访问名称高速缓冲存储器，以确定是否允许文件操作，并且如果在所述名称高速缓冲存储器中没有指示是否允许所述文件操作，则使得所述文件内容散列化，并且将所述文件的散列值与主机高速缓冲存储器中的元信息比较，以确定是否允许文件操作。

5. 按照权利要求 4 的方法，其中，如果主机高速缓冲存储器没有用于确定是否允许所述文件操作的数据，则所述主机查询所述服务器。

6. 按照权利要求 5 的方法，其中，所述服务器响应于上载的文件或者所述文件的内容的签名，使得执行一个或多个分析，其中包括反病毒或者反间谍件扫描。

7. 按照权利要求 6 的方法，其中，所述文件操作被允许来在所

述分析完成之前进行。

8. 按照权利要求 6 的方法，其中，如果所述分析指示不应当允许特定的文件操作，则服务器使得主机更新它们的元信息以指示对于随后的尝试，对于那个文件不允许特定的文件操作。

9. 按照权利要求 6 的方法，其中，如果所述分析指示应当允许特定文件操作，则服务器使得主机更新它们的元信息，以指示对于随后的尝试，对于那个文件允许特定的文件操作。

10. 按照权利要求 6 的方法，其中，如果未确定是否应当允许特定的文件操作，则主机更新它们的元信息以指示仅仅在特定的条件下，对于随后的尝试，对于那个文件允许特定的文件操作。

11. 按照权利要求 2 的方法，还包括：对于在文件中的嵌入的有效内容，从所述文件提取所述内容，执行所述内容的加密散列，并且保存所述有效内容的散列值。

12. 按照权利要求 11 的方法，其中，在执行加密散列之前，在有效的格式化的文件中提供所述有效内容，以产生缩小的文件。

13. 按照权利要求 12 的方法，其中，所述有效内容包括宏，并且所述有效的格式化的文件包括字处理文件。

14. 按照权利要求 11 的方法，其中，所述主机将所述有效内容与从其提取有效内容的文件相关联。

15. 按照权利要求 1 的方法，其中，响应于主机接收到对于在文件上的文件操作的请求，主机访问主机高速缓冲存储器以确定状态，并且如果存在高速缓冲存储器未命中，则所述主机延迟所述文件操作，同时系统执行进一步的分析。

16. 按照权利要求 15 的方法，其中，所述文件操作包括执行、文件读取和文件写入。

17. 按照权利要求 15 的方法，其中，所述进一步的分析包括下述的一个或多个：反病毒扫描、间谍件扫描、与已知文件或者内容的列表相比较、以及与由其它服务器提供的结果相比较。

18. 按照权利要求 1 的方法，其中，所述状态数据包括是否根据

文件的内容来禁止文件操作、根据文件的路径和名称来禁止文件操作、批准文件操作、通过主机本地批准文件操作、或者等待进一步的分析来允许或禁止文件操作。

19. 按照权利要求 1 的方法，其中，所述服务器还保存用于指示何时服务器第一次看到文件或者文件内容的散列值的数据。

20. 按照权利要求 1 的方法，其中，所述服务器还保存用于指示文件被修改的最后时间的数据。

21. 按照权利要求 1 的方法，其中，所述状态数据包括用于指示对于一些主机禁止而对于其它主机允许文件操作的数据。

22. 按照权利要求 1 的方法，其中，所述服务器保存多少主机具有文件的拷贝的计数。

23. 按照权利要求 22 的方法，其中，所述服务器使用多少主机具有文件的拷贝的计数，将所述计数与阈值相比较，并且在所述计数已经超过所述阈值后禁止主机进行文件操作。

24. 按照权利要求 23 的方法，其中，所述服务器保存用于指示分析结果和推荐状态的数据。

25. 按照权利要求 1 的方法，其中，在主机中保存的元信息还包括文件路径名称、第一次看到的日期和最后的修改日期。

26. 按照权利要求 1 的方法，其中，分析在主机存储器上的所有文件，就像在每个文件上发生了改变，自动触发所述分析以将主机状态与服务器状态同步。

27. 按照权利要求 1 的方法，其中，在服务器上的每个文件的元信息包括关于文件的名称的数据。

28. 按照权利要求 1 的方法，其中，所述服务器通过以主机可访问的方式发布元信息中的改变而向主机提供元信息中的改变，并且所述主机访问所发布的改变，并且修改主机上的所述元信息。

29. 一种在系统中使用的方法，所述系统具有服务器和相关联的主计算机（主机），所述方法包括:

在服务器上对于多个文件保存一组元信息，其中对于每个唯一文

件内容签名包括文件的内容的签名、用于指示是否和在什么条件下可以由主机对于文件执行特定的文件操作的状态数据、以及当第一次看到所述文件或者签名时的时间;

在主机上对于多个文件保存一组元信息，其中，对于每个文件包括状态数据、文件内容的签名和文件路径名称;

主机检测对于文件内容或者名称的可能改变，并且更新主机和/或服务器元信息; 以及

服务器向主机提供服务器元信息中的改变。

30. 按照权利要求 29 的方法，其中，如果服务器还未具有所述文件的元信息，则所述文件被上载到所述服务器。

31. 按照权利要求 29 的方法，其中，所述主机访问基于文件名而存储信息的第一高速缓冲存储器和基于文件内容的散列值而存储信息的第二高速缓冲存储器。

32. 按照权利要求 30 的方法，其中，所述服务器执行所述文件的内容的一个或多个散列，并且将所述一个或多个散列与由主机执行的散列相比较。

33. 一种在系统中使用的方法，所述系统具有服务器和相关联的主计算机（主机），所述方法包括:

在服务器上对于多个文件保存一组元信息，其中对于每个唯一文件内容签名包括状态数据，所述状态数据用于指示是否和在什么条件下禁止、允许、或者还没有完全确定与文件相关联的特定操作;

在主机上对于多个文件保存一组元信息，其中，对于每个文件包括所述状态数据;

在主机上检测对于文件内容或者名称的可能改变，并且更新主机和/或服务器元信息;

服务器向主机提供服务器元信息中的改变; 以及

响应于在服务器中没有所述文件的输入项或者存在还没有完全确定的状态，则所述服务器执行所述文件的分析。

34. 按照权利要求 33 的方法，其中，所述服务器使得进行一个

或多个反病毒或者反间谍件扫描。

35. 按照权利要求 34 的方法，其中，所述服务器以后确定是否应当对于特定文件禁止或者允许文件操作，所述服务器使得在状态中的这个改变被传播到主机。

36. 按照权利要求 35 的方法，其中，所述服务器通过发布元信息中的改变而传播该改变，以便主机访问和检索所述改变。

37. 一种系统，包括:

服务器;

多个与服务器相关联的主计算机（主机）;

所述服务器具有服务器存储器，用于对于多个文件保存一组元信息，其中对于每个文件包括关于文件的名称的数据、文件的内容的签名、和状态数据，所述状态数据用于指示是否和在什么条件下禁止、允许或者还没有确定与文件相关联的特定操作;

每个主机具有本地存储器，用于对于多个文件保存一组元信息，其中对于每个文件包括状态数据和签名;

在主机上检测对于文件内容或者名称的可能改变，并且更新主机和/或服务器元信息;

所述服务器使得向主机提供所述服务器元信息中的改变。

38. 按照权利要求 37 的方法，其中，所述服务器通过以主机可访问的方式发布改变而提供所述改变，并且允许主机访问所发布的改变，以便主机修改主机上的元信息。

39. 按照权利要求 37 的系统，其中，内容的签名包括文件的内容的一个或多个加密散列值的结果。

40. 按照权利要求 37 的系统，其中，每个主机具有带有文件名和状态数据的独立的名称高速缓冲存储器。

41. 按照权利要求 37 的系统，其中，响应于主机接收到对于在文件上的文件操作的请求，主机访问主机高速缓冲存储器以确定状态，并且如果存在高速缓冲存储器未命中，则所述主机延迟所述文件操作，同时系统执行进一步的分析。

42. 按照权利要求 37 的系统，其中，所述服务器还保存用于指示何时服务器第一次看到文件或者文件内容的散列值的数据。

43. 按照权利要求 37 的系统，其中，所述服务器还保存用于指示文件被修改的最后时间的数据。

44. 按照权利要求 37 的系统，其中，所述服务器保存多少主机具有文件的拷贝的计数。

45. 按照权利要求 37 的系统，其中，在主机中存储的元信息还包括文件路径名称、第一次看到的日期和最后的修改日期。

46. 一种用于计算机系统的方法，所述计算机系统包括多个主计算机（主机）和与所述主机相关联的服务器，所述方法包括:

服务器向主机传播与文件操作相关联的策略和策略选项的主集，所述策略选项至少表示是否和在什么条件下这样的操作被允许或者禁止，并且

服务器向所述主机传播值;

在所述主机上存储的值表示在所述策略和策略选项的主集中的策略和策略选项的哪个子集要在主机上实现;

所述主机实现由所述值指示的文件操作策略。

47. 按照权利要求 46 的方法，其中，每个策略具有单个配置参数，其指示策略选项之一，被传播的值用于选择多个策略的每个的策略选项。

48. 按照权利要求 46 的方法，所述主集包括策略和选项的列表，被转播的所述值用于选择所述列表之一。

49. 按照权利要求 46 的方法，其中，所述策略至少具有三个选项，它们构成限制的有序集，所述限制递增地增加或者减少主机执行文件操作的能力。

50. 按照权利要求 46 的方法，其中，所述服务器响应于管理员的人工改变而改变所述值。

51. 按照权利要求 46 的方法，其中，所述服务器自动改变所述值，而不使用来自管理员的输入。

52. 按照权利要求 51 的方法，其中，所述自动服务器值改变响应于所检测的安全事件或者 SNMP 消息或者 syslog 消息或者报告或者网络消息或者电子邮件消息。

53. 按照权利要求 46 的方法，其中，所述主机自动改变所述值，而不使用来自人的输入或者不使用来自服务器的命令。

54. 按照权利要求 46 的方法，其中，所述主机响应于在同一主机上的策略报告或者响应于在主机上检测的事件或者响应于在主机上执行的命令而自动改变所述值。

55. 按照权利要求 46 的方法，其中，所述策略选项包括：自动允许具有相关联的元信息状态的文件的执行和/或读取操作，所述状态指示批准这样的操作。

56. 按照权利要求 46 的方法，其中，所述策略选项包括：阻止具有相关联的元信息状态的文件的执行和/或读取操作，所述状态指示禁止这样的操作；并且/或者，向服务器发送报告。

57. 按照权利要求 46 的方法，其中，所述策略选项包括：阻止具有相关联的待决元信息状态的文件的执行和/或读取操作，所述状态指示还没有确定允许或者禁止这样的行为。

58. 按照权利要求 46 的方法，其中，所述策略选项包括：允许具有相关联的待决元信息状态的文件的执行和/或读取操作，所述状态指示还没有完全确定允许或者禁止这样的行为。

59. 按照权利要求 46 的方法，其中，所述策略选项包括：在对于对具有相关联的待决元信息状态的文件的操作的请求的情况下，向服务器发送报告，所述状态指示还没有完全确定允许或者禁止这样的行为。

60. 按照权利要求 46 的方法，其中，所述策略选项包括：检测和跟踪具有相关联的待决和/或禁止的元信息状态的新文件的建立或者修改或者第一次执行。

61. 按照权利要求 46 的方法，其中，所述策略选项包括：阻止具有相关联的待决元信息状态的文件的建立或者修改。

62. 按照权利要求 46 的方法，其中，所述策略选项包括：自动删除或者移动具有相关联的禁止的元信息状态的文件。

63. 按照权利要求 46 的方法，其中，所述策略选项包括自动将新建立或者修改的文件的主机元信息状态设置为批准。

64. 按照权利要求 46 的方法，其中，所述策略选项包括：自动将在主机上新建立或者修改的文件的服务器元信息状态设置为批准。

65. 按照权利要求 46 的方法，其中，所述主机保存主机上的每个文件的元信息，所述元信息包括具有至少三个可能值的状态，所述可能值是批准、禁止和待决。

66. 按照权利要求 65 的方法，其中，允许指示允许不进一步监控地进行操作。

67. 按照权利要求 65 的方法，其中，所述策略选项包括：当主机和/或服务器正在分析文件时，延迟文件操作。

68. 按照权利要求 67 的方法，其中，当服务器确定其没有与那个文件相关联的元信息时，服务器将待决状态与文件相关联。

69. 按照权利要求 67 的方法，其中，当主机确定其没有与那个文件相关联的元信息时，主机将待决状态与文件相关联。

70. 按照权利要求 46 的方法，其中，所述服务器保存用于指示何时由任何主机第一次看到文件的元信息。

71. 按照权利要求 46 的方法，其中，所述策略和策略选项的至少一些指示基于文件的名称的行为。

72. 按照权利要求 46 的方法，其中，所述策略和策略选项的至少一些表示基于文件的内容的行为。

73. 按照权利要求 46 的方法，其中，所述策略和策略选项的至少一些表示基于文件的名称和内容的组合的行为。

74. 按照权利要求 46 的方法，其中，所述服务器保存关于文件的元信息，其中包括文件的内容的散列值。

75. 按照权利要求 74 的方法，其中，所述文件的内容的散列值是在文件内的感兴趣的内容的散列值。

76. 按照权利要求 46 的方法，其中，所述策略包括新文件的执行、对于文件的写访问和对于文件的读取，其中，所述选项包括使用进一步的监控来允许所述行为发生、禁止所述行为或者批准所述行为。

77. 按照权利要求 76 的方法，其中，所述进一步监控包括：记录和提供报告的一个或多个。

78. 按照权利要求 46 的方法，其中，所述值之一禁止所有的新可执行部分。

79. 按照权利要求 46 的方法，其中，所述值之一允许所有的文件操作。

80. 按照权利要求 46 的方法，其中，所述服务器通过以主机可访问的方式发布新值来改变所述值，所述主机访问所述新值，将所述新值与所述主机具有的值相比较，并且将其值改变为新值。

81. 按照权利要求 80 的方法，其中，存在值和相关联的策略和策略选项的有序集，所述主机通过其它的中间值而递增地改变到所述新值。

82. 按照权利要求 46 的方法，其中，每个主机被布置到多个主机组之一中，所述服务器改变至少一个但是不是全部主机组的值，以便主机组具有不同的值。

83. 按照权利要求 46 的方法，其中，所述服务器通过向主机发送所述新值而改变所述值。

84. 按照权利要求 54 的方法，其中，主机可以获得对于服务器文件元信息状态的改变，并且/或者，所述改变传播到主机。

85. 一种计算机系统，包括：

多个主计算机（主机）；以及

服务器，用于向主机传播与文件操作相关的策略和策略选项的主集，所述策略选项用于至少指示是否和在什么条件下允许或者禁止这样的操作，并且

所述服务器还用于向主机传播一个值，以存储在主机上；

在所述主机上存储的值表示在策略和策略选项的主集中的策略和策略选项的哪个子集要在主机上实现;

所述主机实现由所述值指示的文件操作策略。

86. 按照权利要求 85 的系统，其中，由服务器传播的信息包括用于指示多个不同的策略的每个的策略选项集的值。

87. 按照权利要求 86 的系统，其中，所述主机被组织为多个主机组，所述服务器向一个或多个主机组但是不是全部主机组传播在所述值上的改变。

88. 按照权利要求 85 的系统，其中，所述文件操作包括对于文件的写访问和文件的执行，所述选项包括在限制的有序集中的多个选项，所述限制递增地增加或者减少主机执行文件操作的能力。

89. 按照权利要求 85 的系统，其中，所述主集包括策略和策略选项的列表，所述服务器提供包括用于指示所述列表之一的值的信息。

90. 按照权利要求 85 的系统，其中，所述服务器在主机可访问的位置发布信息，并且所述主机访问所述信息并且更新它们的值。

91. 一种用于系统中的方法，所述系统具有服务器和多个主计算机，所述主计算机与服务器相关联，所述方法包括:

服务器指定文件的元信息查询;

向一个或多个主机组分发所述元信息查询;

所述主机从在存储器中存储的本地主机元信息进行所述元信息查询;

所述主机向服务器发送来自所述元信息的查询的结果，所述结果包括关于主机上的文件的信息;

所述服务器接收和存储来自主机的结果。

92. 按照权利要求 91 的系统，其中，所述服务器根据规则集设置安全策略，所述服务器响应于从主机接收的查询的结果而自动改变适用于至少一些主机的规则。

93. 按照权利要求 92 的方法，其中，所述服务器响应于所述结

果而自动触发安全警告。

94. 按照权利要求 91 的方法，其中，所述服务器将从主机接收的结果合并，以产生统一的报告。

95. 按照权利要求 91 的方法，其中所述服务器向每个主机发送所述查询。

96. 按照权利要求 91 的方法，其中，所述服务器发布所述查询以由每个主机访问，并且其中，每个主机获得由所述服务器发布的查询。

97. 按照权利要求 91 的方法，其中，可以对于一组主机查询的文件的元信息包括下面的内容的一个或多个:

文件名的正则表达式模式规格，

文件路径的正则表达式模式规格，

文件的感兴趣的内容的散列值，

当主机第一次看到文件或者文件的散列值时的时间范围，

主机的名称，

主机的 IP 地址，

文件的类型，

来自一组至少三个状态的、与文件相关联的一个或多个主机文件状态，所述三个状态是批准、禁止、待决分析，

是否已经由主机对于文件执行了特定的文件操作，以及

主机组。

98. 按照权利要求 97 的方法，其中，所述查询用于具有被识别的文件名的文件。

99. 按照权利要求 97 的方法，其中，所述查询用于具有被识别的文件路径的文件。

100. 按照权利要求 97 的方法，其中，所述查询用于具有其内容的被识别的散列值的文件。

101. 按照权利要求 97 的方法，其中，所述查询是用于具有被识别的当主机第一次看到所述文件时的时间范围的文件。

102. 按照权利要求 97 的方法，其中，所述查询用于具有文件操作的被识别状态的文件，所述状态表示是否已经批准或者禁止了所述文件操作。

103. 按照权利要求 97 的方法，其中，所述查询包括项目（1）-（6）的两个或者多个。

104. 按照权利要求 97 的方法，其中，所述查询包括项目（1）-（6）的三个或者多个。

105. 按照权利要求 91 的方法，其中，由主机识别到服务器的每个文件的结果包括:

文件名，

文件路径，

文件的感兴趣的内容的散列值，

当主机第一次看到文件或者文件的散列值时的时间，

主机的名称，

主机的 IP 地址，

文件的类型，

来自一组至少三个状态的、与文件相关联的一个或多个主机文件状态，所述三个状态是批准、禁止、待决分析，

是否已经由主机对于文件执行了特定的文件操作，以及

主机组。

106. 按照权利要求 91 的方法，其中，所述服务器保存元信息的存储库，服务器向主机提供更新以改变在主存储器中存储的元信息。

107. 按照权利要求 96 的方法，其中，所述主机使用最后已知的修改的元信息时间来轮询，并且服务器发回对于元信息的本地主机存储库的更新是否是待决的指示。

108. 按照权利要求 96 的方法，其中，所述主机元信息被存储在核心和用户空间中的多个永久高速缓冲存储器中。

109. 按照权利要求 96 的方法，其中，在文件被服务器第一次看到后经过限定时段，文件的元信息和/或所述文件被删除。

110. 按照权利要求 96 的方法，其中，在服务器中保存的元信息包括内容签名、由一个或多个主机组第一次看到的日期/时间、和近来分析结果和时间的历史。

111. 按照权利要求 110 的方法，其中，在服务器中保存的元信息还包括近来状态改变的历史、改变的原因、和元信息最后改变的时间。

112. 一种计算机系统，包括:

多个主计算机;

与所述主计算机相关联的服务器;

每个主计算机具有元信息数据存储库，其具有名称信息、内容信息、内容的散列值、和多个文件的每个的安全信息，所述主计算机响应于来自服务器的查询，根据所定义的标准而搜索所述元信息，并且提供满足所述标准的文件的列表。

113. 按照权利要求 112 的系统，其中，通过管理接口向所述服务器提供查询。

114. 按照权利要求 112 的系统，其中，所述主计算机定期查看所述服务器以获得元信息更新。

115. 按照权利要求 112 的系统，其中，可以对于一组主机可查询的文件的元信息包括下面的内容的一个或多个:

文件名的正则表达式模式规格，

文件路径的正则表达式模式规格，

文件的感兴趣的内容的散列值，

当主机第一次看到文件或者文件的散列值时的时间范围，

主机的名称，

主机的 IP 地址，

文件的类型，

来自一组至少三个状态的、与文件相关联的一个或多个主机文件状态，所述三个状态是批准、禁止、待决分析，

是否已经由主机对于文件执行了特定的文件操作，以及

主机组。

116. 按照权利要求 112 的系统，其中，所述查询用于具有被识别的文件名的文件。

117. 按照权利要求 112 的系统，其中，所述查询用于具有被识别的文件路径的文件。

118. 按照权利要求 112 的系统，其中，所述查询用于具有感兴趣的内容的被识别的散列值的文件。

119. 按照权利要求 112 的系统，其中，所述查询是用于具有被识别的当主机第一次看到所述文件或文件散列值时的时间范围的文件。

120. 按照权利要求 112 的系统，其中，所述查询用于具有文件操作的被识别状态的文件，所述状态表示在特定条件下是否已经批准或者禁止了所述文件操作。

121. 按照权利要求 112 的系统，其中，所述查询包括项目（1）-（6）的两个或者多个。

122. 按照权利要求 112 的系统，其中，所述查询包括项目（1）-（6）的三个或者多个。

123. 一种用于服务器和一组相关联的主机的方法，包括:

在服务器中存储与在主机上看到的文件相关的元信息状态，所述元信息包括文件的内容的签名;

对于每个签名存储初始时间;

在与初始时间相关的限定时段，执行所述文件的至少一个安全分析或者所述文件内容的签名的分析; 并且

改变所述文件状态，并且向主机提供与所改变的状态相关联的信息。

124. 按照权利要求 123 的方法，其中，所述执行处理包括执行反病毒和反间谍件扫描之一。

125. 按照权利要求 123 的方法，其中，所述签名包括内容的一个或多个加密散列值，所述服务器存储具有唯一散列值的每个文件的

单个拷贝。

126. 按照权利要求 123 的方法，还包括存储用于指示是否和在什么条件下可以通过主机对于文件执行特定的文件操作，其中，所述执行处理包括查看对于其已经在特定条件下批准或者禁止了特定的文件操作的签名的一个或多个列表。

127. 按照权利要求 123 的方法，其中，与初始时间相关的限定时段是具有在所限定的范围内的所述文件或者签名被主机和/或服务器第一次看到的日期的全部文件。

128. 按照权利要求 123 的方法，还包括：存储用于指示是否和在什么条件下可以通过主机对于文件执行特定的文件操作的状态，其中，所述状态包括在特定条件下对于特定的文件操作禁止、允许或者允许进一步的监控。

129. 按照权利要求 128 的方法，其中，所述进一步监控包括记录或者提供报告。

130. 按照权利要求 123 的方法，还包括对于在服务器上的每个元信息记录存储与每个分析相关联的附加记录。

131. 按照权利要求 130 的方法，其中，所述附加记录包括与执行分析的人员、分析时间和分析结果相关的数据。

132. 按照权利要求 131 的方法，其中，所述附加记录还包括推荐的状态和分析结果信息串。

133. 按照权利要求 123 的方法，还包括响应于所述分析而产生 syslog 消息和/或统计更新和/或警告。

134. 按照权利要求 123 的方法，其中，所述分析引起下述的一个或多个：内容转发、元信息转发、分析转发和在安全规则上的改变。

135. 按照权利要求 123 的方法，其中，作为对于另一方的主计算机的外包服务，执行权利要求 1 的行为。

136. 按照权利要求 135 的方法，其中，服务器使用被设置为禁止或者批准的特定行为自动向其它网络装置通知对于文件的列表的改变。

137. 按照权利要求 135 的方法，其中，所述服务基于年代和分析结果来向其它服务器和网络装置传送近来的元信息。

138. 按照权利要求 123 的方法，其中，所述安全分析包括将近来的分析结果与其它服务器的近来分析结果相比较。

139. 按照权利要求 123 的方法，其中，所述安全分析包括用户定义的内容分析。

140. 按照权利要求 123 的方法，其中，所述服务器具有基于初始时间在多个时间执行的一组多个分析。

141. 按照权利要求 140 的方法，其中，所述多个时间之一小于在初始时间后的一天，并且所述多个时间的另一个大于在初始时间后的一天。

142. 按照权利要求 140 的方法，其中，所述多个时间之一小于在初始时间后的一天，并且所述多个时间的另一个大于在初始时间后的一个星期。

143. 按照权利要求 123 的方法，其中，所述初始时间与当主机在系统中第一次看到文件或者文件内容的散列值时的时间相关联。

144. 按照权利要求 123 的方法，其中，所述初始时间与当服务器在系统中第一次看到文件或者文件内容的散列值时的时间相关联。

145. 按照权利要求 123 的方法，其中，服务器向主机传播一组策略，所述服务器也提供用于指示要通过主机在哪些条件下实现哪些策略的信息。

146. 按照权利要求 123 的方法，还包括：主机接收第一文件，提取在第一文件内的感兴趣的内容，将感兴趣的内容重新打包为有效的格式化类型的文件，以产生缩小的文件，并且向所述缩小的文件应用散列。

147. 按照权利要求 123 的方法，其中，执行文件的至少一个安全分析的处理包括：向另一个计算机系统提供文件，并且从所述另一个计算机系统接收分析结果。

148. 按照权利要求 123 的方法，还包括：存储用于指示是否和

在什么条件下可以通过主机对于文件执行特定文件操作的状态，当所述状态中有改变时，所述服务器向主机传播元信息。

149. 按照权利要求 148 的方法，其中，所述服务器通过发布元信息而传播该元信息，以便主机访问和获得所述元信息。

150. 按照权利要求 123 的方法，其中，所述服务器执行所述内容的散列，并且将所述散列与至少一个主机执行的散列相比较。

151. 按照权利要求 150 的方法，其中，所述服务器执行多个散列。

152. 按照权利要求 123 的方法，其中，主机查询服务器以确定在向服务器上载签名和/或文件内容之前所述服务器是否具有用于特定文件的元信息。

153. 按照权利要求 123 的方法，其中，所述服务器向远程网络装置查询以确定是否签名或者名称对应于被批准的文件的白名单。

154. 按照权利要求 123 的方法，其中，所述服务器向远程网络装置查询以确定是否签名或者名称对应于被禁止的文件的黑名单。

155. 按照权利要求 123 的方法，其中，所述服务器向远程网络装置查询以确定是否签名或者名称对应于新或者未分类的文件的待决列表。

156. 按照权利要求 123 的方法，其中，所述服务器存储将文件和签名与其它文件和签名的组相关联的信息。

157. 按照权利要求 123 的方法，其中，所述服务器存储将文件和签名与关于在其它服务器上存储的信息的统计相关联的信息。

158. 按照权利要求 123 的方法，其中，所述签名包括内容的一个或多个加密散列值，服务器向远程网络装置查询以查找相对于已知产品分类数据库的散列值，以便识别相关的产品和对应于与文件散列值相关联的产品的其它文件。

159. 按照权利要求 123 的方法，其中，所述服务器执行下述的一个或多个：内容查询、元信息查询、分析结果查询和安全规则查询，并且所述服务器查询网络上的另一个装置。

160. 一种计算机系统，包括:

服务器，其包括存储器，用于存储与在与服务器相关联的主机上看到的文件相关的安全相关元信息，所述元信息对于每个文件包括用于指示是否和在什么条件下主机对于文件执行特定的文件操作的状态;

所述服务器在限定时段使得进行文件的至少一个安全分析，所述限定时段基于当主机和/或服务器已经接收到文件或者文件的签名时的初始时间; 并且

响应于至少一些分析，改变所述状态，并且向主机提供与所改变的状态相关联的信息。

161. 按照权利要求 160 的系统，还包括: 与所述服务器相关联的一组主机。

162. 按照权利要求 161 的系统，其中，所述服务器存储用于指示是否和在什么条件下主机对于文件执行特定的文件操作的状态，其中，所述执行处理包括查看对于其已经批准或者禁止了文件操作的文件的一个或多个列表。

163. 按照权利要求 160 的系统，其中，所述服务器自动向其它网络装置通知对于被设置为禁止或者批准的行为的列表的改变。

164. 按照权利要求 161 的系统，其中，所述服务器根据年代和分析结果来向其它的服务器和网络装置发送近来的元信息。

165. 按照权利要求 160 的系统，其中，所述初始时间与当服务器在系统中第一次看到文件或者文件内容的散列值时的时间相关联。

166. 按照权利要求 161 的系统，其中，所述服务器向主机传播策略集，所述服务器也提供用于指示主机在什么条件下要实现所述策略的哪些的信息。

167. 按照权利要求 161 的系统，其中，所述服务器存储用于指示是否和在什么条件下主机对于文件执行特定的文件操作的文件状态，当在状态中有改变时，所述服务器向主机传播所述元信息。

168. 按照权利要求 167 的系统，其中，所述服务器通过发布元

信息而传播所述元信息，以便通过主机访问和获得元信息。

169. 按照权利要求 161 的系统，其中，所述签名包括所述内容的一个或多个加密散列值，所述服务器执行所述内容的散列，并且将所述散列与由至少一个主机执行的散列相比较。

170. 按照权利要求 169 的系统，其中，所述服务器执行多个散列以获得签名。

171. 一种计算机实现的方法，包括:

接收第一文件;

提取在第一文件内的感兴趣的内容;

将所述感兴趣的内容重新打包为一个或多个有效格式化类型的文件，以产生第一缩小的文件;

向缩小的文件应用签名; 以及

存储所述签名，并且将所述签名与第一文件相关联。

172. 按照权利要求 171 的方法，还包括: 对于第二文件执行权利要求 1 的行为，并且将第一缩小文件的签名与第二缩小文件的签名相比较，以确定是否它们相同。

173. 按照权利要求 172 的方法，其中，所述签名包括文件内容的一个或多个加密散列值。

174. 按照权利要求 171 的方法，其中，在主计算机上执行权利要求 1 的行为，以及向与所述主计算机相关联的服务器传送所述第一缩小文件，以用于进一步的内容分析。

175. 按照权利要求 171 的方法，其中，所述感兴趣的内容包括宏。

176. 按照权利要求 175 的方法，其中，所述感兴趣的内容被重新打包为字处理文件。

177. 按照权利要求 171 的方法，其中，存储散列值还包括存储当服务器第一次看到所述缩小的文件或者其签名时的日期。

178. 按照权利要求 171 的方法，其中，所述第一文件是具有演示幻灯片的演示文件。

179. 按照权利要求 171 的方法，还包括：作为电子邮件附件发送所述第一缩小文件。

180. 按照权利要求 178 的方法，其中，向具有电子邮件附件的病毒扫描能力的网关发送电子邮件。

181. 按照权利要求 171 的方法，其中，感兴趣的内容包括可执行部分、脚本、档案文件和安装程序之一。

182. 按照权利要求 171 的方法，其中，所述感兴趣的内容被重新打包为字处理文件或者档案文件。

183. 按照权利要求 171 的方法，其中，所述签名包括文件内容的一个或多个加密散列值。

184. 一种计算机系统，包括：

存储器，用于存储关于文件的元信息，其中对于每个文件包括内容的签名和用于指示是否和在什么条件下可以执行特定的指定文件操作的状态信息；

主计算机，响应于接收到第一文件来提取所述第一文件内的感兴趣的内容，

将所述感兴趣的内容重新打包为一个或多个有效的格式化类型的文件，以产生第一缩小文件，

向第一缩小文件应用签名或者散列值；以及

存储所述签名或者散列值，并且将所述签名或者散列值与所述第一文件相关联。

185. 按照权利要求 184 的系统，其中，所述签名包括文件内容的一个或多个加密散列值。

186. 按照权利要求 184 的系统，其中，存在多个主计算机和与所述多个主计算机相关联的一个服务器，所述主计算机向所述服务器提供签名以存储。

187. 按照权利要求 184 的系统，其中，所述元信息对于每个文件包括：文件名称、缩小内容的签名、和用于指示是否和在什么条件下可以执行特定的指定文件操作的批准或者禁止或者未知状态信息

的至少一个。

188. 按照权利要求 187 的系统，其中，所述元信息具有用于不同操作的多个批准或者禁止状态，所述不同操作包括执行和写入操作。

189. 按照权利要求 188 的系统，其中，所述服务器在数据存储库中存储元信息，并且所述主计算机在数据存储库中存储数据的高速缓冲存储器，所述服务器使得向所述主计算机高速缓冲存储器提供更新。

190. 按照权利要求 188 的系统，其中，所述服务器存储内容散列值、文件名、批准和禁止的状态和服务器第一次看到文件或者文件的散列值的日期。

191. 按照权利要求 190 的系统，其中，所述服务器可以存储待决状态，其中，等待进一步分析来允许操作以将状态改变到允许或者禁止。

192. 按照权利要求 186 的系统，其中：

所述主计算机通过检测可以改变文件的内容和/或名称的操作来接收第一文件；

所述主计算机通过在本地元信息存储库上的文件名或者文件名的签名来查找文件元信息；

如果未找到所述文件名或者签名，则根据内容的散列值向服务器提供查询；

如果找到文件名或者签名，则访问所述元信息以确定是否和在什么条件下可以执行特定的指定文件操作。

193. 按照权利要求 192 的系统，其中，所述签名包括一个或多个文件内容的加密散列值。

194. 按照权利要求 192 的系统，其中，所述主机定期查看服务器以获得元信息更新。

195. 按照权利要求 186 的系统，还包括：在第一次看到文件后经过限定时段自动执行文件的进一步的分析。

196. 按照权利要求 189 的系统，其中，在服务器中保存的元信息还包括近来状态改变的历史和改变的原因、以及最后改变元信息的时间。

197. 按照权利要求 194 的系统，其中，根据服务器元信息，从主机向服务器自动传送文件。

网络安全系统和方法

背景技术

大企业具有大的信息技术（IT）安全预算和分层的IT安全系统，但是网络损害、来自病毒和蠕虫的损害和间谍件问题是常有的。当前的IT安全技术维护起来很昂贵，并且不提供针对新的或者未知的威胁的保护，而新的威胁正在以提高的速率被分布、检测和报告。

位于网络周界的安全解决方案（诸如防火墙）具有限于直接通过它们的网络流量的可见度。诸如电子邮件病毒、网页浏览器利用、无线访问、VPN、即时消息传送和文件共享的输入向量产生绕过这些技术的越来越能够渗透的周界。难于限定提供足够的控制和可见度的、在现代网络中的周界。许多攻击仅仅在它们已经损害了机器或者网络后产生网络流量。例如，在病毒开始从在网络内的机器发出电子邮件时之前，那个机器已经被损害。为了在攻击执行之前停止攻击，一般需要保护文件，而不仅仅是网络流量。

可以通过主机代理来提供可见度和保护，所述主机代理是软件，有时与硬件相结合地使用，所述主机代理工作在网络内的多个独立的计算机（“主机”）上。主机代理一般并行工作，使用所述主机的资源的一些来在后台执行安全功能。通过可能访问主机的所有重要的内部功能，主机代理可以在理论上在任何损害发生之前检测和停止在主机上的威胁。主机代理安全系统有时被称为端点安全系统，因为它们工作在网络的“端部”上。

当前的企业端点安全系统经常试图使用已知的位模式（诸如反病毒（AV）扫描和反间谍件（AS）扫描）来检测和阻止攻击。模式扫描使用被预先识别为坏的模式的黑名单。类似地，一些安全系统使用所检测的已知行为简档，其可以被描述为坏行为模式的黑名单。在这两种情况下，黑名单永久过期，不能响应于新的或者未知的攻击。黑

名单也相对于诸如新病毒的攻击无效，所述新病毒可以比用于得出、测试和分布黑名单更新的能力更快地传播。在每个星期发现许多新的病毒的情况下，所有种类的黑名单变得越来越无效。行为模式开发和测试起来很复杂，结果，它们具有高误报警率；即，当事实上一个行为无危险时，它们错误地得出所述行为坏的结论。随着新的攻击演化，行为改变，导致丢失检测的错误。通过等待直到诸如病毒的攻击显示坏的行为，被影响的机器可能已经被损害。总之，黑名单试图跟踪已知为错误的内容，而错误的内容总是在改变。

另一种企业端点技术是异常检测。这可以被看作行为黑名单记录，通过随着时间观察行为来以统计方式确定所述行为黑名单。除了继承行为黑名单的缺点，当以统计方式估算好和坏行为时，异常检测还增加新的错误模式，因此肯定存在估算错误。这个过程经常导致不可接受的高误报警和丢失检测率。

另一类端点安全系统将执行仅仅限于在白名单上的程序。所述白名单是已知的好程序的模式的列表。如果在所述列表中不包括一个程序，则其将不运行。这样的系统对于典型的现代企业不够灵活，并且，难于维护所产生的白名单。例如，大多数大企业部署定制的程序，其在内部被开发，并且经常地改变。而且，这些程序可以包括不能向第三方曝光的敏感的知识产权和安全风险。白名单提供商不可能及时地访问以预先批准这个软件。其它示例是操作系统和其它更新。同样，没有任何中央数据库或者中央授权证书来验证特定程序或者更新对于所有的企业是良好的。白名单系统的故障模式是严重的，阻止了对于关键的、但是未被批准的应用和商业功能的访问。

结果，在中央将文件内容访问划分为仅仅一个或者两个状态（批准和禁止）的系统具有关于竞争（定时）条件的问题。大量的软件不明确地适合于任何类别，并且没有对于在一个企业内的所有软件普遍地被信任的任何中央授权。即使这不是一个因素，也可能需要时间来划分中间件。在新的病毒的情况下，需要 6-48 小时或者更多时间来将新的病毒分类为坏，但是在此之前，爆发会是传染性的。因此即使对

于从主机向中央批准授权的强网络连接，也会需要比几分钟更长的时间来检测和分析新的软件。为了透明地将这个基于内容的授权加到后台的操作系统中，延迟通常必须小于一分钟，否则，所述文件系统将超时，并且发生伪访问阻止错误。

发明内容

在此所述的安全系统允许管理员检测、监控、定位、识别和控制在大计算机网络上安装的文件。所述系统可以防御已知和未知的病毒、蠕虫、间谍件、黑客、未批准的/不需要的软件（例如不符合商业使用策略的软件应用）和社会工程攻击。当新的可执行部分、脚本和嵌入脚本出现和传播到联网系统时，管理员可以访问关于所述新的可执行部分、脚本和嵌入脚本的详细信息和统计。所述系统可以实现集中的策略，其允许管理员批准、阻止、隔离或者记录文件行为。所述系统也可以收集对于诊断和定位问题文件或者攻击有益的详细信息。所述系统提供用于大计算机站的可见度、控制和保护。

所述系统架构最好包括代理软件，其运行在每个被保护的主机和服务器（被称为“服务器”）上，所述服务器提供集中策略管理、事件监控、代理协调和病毒扫描。所述服务器可以被实现为装置（appliance）（其一般暗示更有限功能的装置）。单个装置可以支持多个主机，例如 10000 个主机。另一个服务器或者装置（有时被称为“超级服务器”）可以监控多个装置。

在每个被保护的主计算机上运行的代理软件分析文件系统行为，并且根据在服务器上配置的策略来采取行动。在一种实现方式中，当主机试图打开或者写入文件时，所述代理软件计算所述文件的内容的散列值（hash）以唯一地向所述系统标识所述文件。所述代理软件使用这个散列值来查找所述文件的状态或者策略。根据这个信息，代理软件可能阻止操作，记录事件，隔离文件或者采取一些其它的指定行为。

所述系统也包括许多其它的特征，其可以组合地或者独立地有

益，其中包括在此所述的从文档提取文件的能力、从文件提取宏的能力、集中内容跟踪和分析以及“查找文件”功能。

在此所述的系统可以使用至少两个附加状态：待决，其表示中间的未限定的威胁级；局部批准，其是批准一个主机，但是不必然批准中央授权（因此所有的其它主机）。后者允许主机略微地脱离基线。待决状态允许主机根据各种威胁级和企业使用策略来阻止或者允许对于新的内容的访问。虽然使用普通的二进制批准术语批准和禁止，但是将批准划分为 3-4 个状态导致每个独立状态的不同的改善能力。一般地，新的还没有被分类的软件是待决的。软件的传统二进制访问状态（禁止/批准）不足够灵活，这样的分类系统是不可升级的。

作为新/待决的软件的指定是有益的。大多数企业具有某种形式的“禁止新的可执行部分”策略，诸如“不允许雇员从因特网下载和运行未经批准的软件”。可是，企业当新的软件传播时不能检测所述新的软件，直到太晚，不知道何时它们的策略被违反，并且没有有效地实施它们的策略的手段。通过当待决的新程序正在被修改/写入文件系统时跟踪所述待决的新程序，主机代理可以当新的内容从几乎任何手段（无论是电子邮件、即时消息传送器、下载、USB 密钥、移动笔记本计算机等）进入网络时实时地检测和报告所述新的内容。通过识别待决的程序，一些简单的、可升级的、有效的策略是可能的，诸如“允许但是当主机运行新的可执行部分时警告”或者“不能通过这组主机来安装或者运行新的未批准的程序”或者“当在 24 小时内在超过 N 个主机上出现同一新的未批准的程序时警告”。因此，可以安全地定位、跟踪或者在阻止的同时分析新的程序。其它被批准的商业软件继续运行。新的被批准的软件可以被安装和运行，诸如 AV 更新或者安全补丁。这种方法是主动的响应，在允许生产率的同时保护防止未知的可能恶意的软件，并且赢得分析时间而不需要任何时间紧要的黑名单或者白名单更新。

现有的文件白名单和黑名单系统趋于在本质上是全局的，因为在中央保留许多独立的名单（每个主机上一个）是困难的。如在此所述，

主机可以保留它们自己的名单，其可以从中央名单分出。特别是，对于本地批准和待决状态可以是这种情况，并且对于基于名称的状态（诸如 NameBan 和 NameApprove）而言经常如此。因为“name（名称）”是本地属性，因此这些状态可以从中央受控的状态分出。例如，如果文件“foo”具有特定的 hash = x 和中央服务器状态待决，则在主机上，所述文件可以是本地批准或者名称禁止的或者名称批准的，后两者依赖于在主机上的文件的本地名称。在此所述的系统允许被同时应用到在每个主机上的每个文件的几千个名称属性的有效地管理和策略实现。名称批准（NameApprove）基于其在主机上建立文件允许灵活的本地批准和中央批准能力。与主机组相结合，这允许精确地灵活有效地指定在哪里和在哪些主机上批准了新的内容。

即使对于这种新的灵活策略系统，企业通常需要对于不同的角色和情况实施不同的策略。例如，IT 管理员和内部软件开发者可能需要仔细地运行新的软件，而其它的雇员仅仅需要小标准套件的相对静态的应用。当在攻击下时，这种情况可以迅速地改变。例如，如果在超过 N 个主机上检测到病毒，则扩展“禁止新的可执行部分”策略有意义。与不能适应于在企业内并且在不同的条件下的各种策略的刚性系统相比较，这种灵活性和递增响应是在此所述的“参数内容控制”的优点。“参数内容控制”允许灵活的锁定模式，其可以根据网络和主机条件被中央地管理，并且迅速地被改变。这允许递增文件内容和/或基于文件名称的限制和批准。

不像处理主机用户证书、处理标识符、数据源（URL）、目录结构和操作系统安全描述符的其它端点安全技术那样，在此所述的系统不必使用这些因素来作为主机策略的一部分。在主机上，这些因素会是不可靠的，并且会易于受到损害攻击，并且它们会妨碍可升级性。这些因素导致不可升级的策略，因为精细粒度的策略可以以复杂的方式在多个主机上交互。即使所述操作系统被损害并且一个攻击获得管理特权和所有相关联的安全描述符，在此所述的“禁止新的可执行部分”将提供实质的保护。

“内容跟踪”系统使用诸如待决的附加状态来当新的内容通过网络移动时监控和分析所述新的内容。当前的技术不允许实时地在大量主机上的每个新的可执行文件的全局中央可见度和跟踪。依赖于文件系统扫描的端点系统（诸如 AV 扫描器）和主机应用目录（诸如 Tripwire）定期地和缓慢地爬过大文件系统以查找新的或者改变的软件。这通常对于主机是破坏性的，会需要较多时间，并且通常至多每天一次地被调度。通过聚焦在新的内容并且将那个信息存储在存储器中，内容跟踪系统更具有可升级性和响应性。因为很少有从未被在大组 N 中的任何主机见过的新的软件到达，并且更少有许多主机 M 使得新的软件在短时间内出现，因此通过这个区别而便利了报告、响应和分析。

一旦检测到新的软件，以及时的方式来定位和识别它会是有益的。如果某个新的软件原来是新的攻击并且正在传播，则期望很快地响应。此外，当前的技术可以在几分钟到几小时的时间量程上在网络上的单个主机上定位单个新文件。即使在单个主机上，通过名称或者内容查找很新的文件需要 15-60 分钟，这将在查询正在被处理的同时不利地影响主机的盘性能。在过去 20 年中，硬盘已经在字节存储空间上变得更大，但是还没有在速度上成比例地增大。“分布式元信息查询”特征加速了在大量（几千）的主机上在几秒内定位和识别关键文件属性，并且具有中央指定的查询，中央报告的结果，较少或者没有主机盘影响。不像跟踪所有文件的传统的跟踪技术（包括还没有改变的那些）那样，在此的本发明当文件正在改变时跟踪在存储器中的文件改变，这提供了对于主机查询来自存储器的文件元信息的有效手段。中央地处理这个信息第一次提供了在主机文件系统的集合上独立文件的移动的响应全局视图。最后，作为安全服务，主机连接到、发送到中央服务器和从中央服务器查询是重要的。这是本发明的重要部分，其允许主机通过一个或多个防火墙或者 NAT 装置从服务器分离，并且避免了保证在接受/监听模式中的附加主机网络插口的困难问题。

使用内容分析的当前的端点主机代理系统对于更新主机代理具

有问题。例如，对于更有效的 AV 扫描器，它们应当在使得可以获得的更新的几个小时或者几分钟内被更新。带有滞后的 AV 的任何主机具有风险，并且许多 AV 系统被不正确地配置，导致更新滞后。因为它们不有效地跟踪文件改变，因此 AV 扫描器通常需要较长的时间来响应于被写入到文件系统的新的内容。而且，当前的主机内容分析技术不必重新分析文件，不考虑安全因素。例如，更重要的是，新的内容越新，就越经常地分析所述新的内容。如果文件在网络中完全未变达到 2 年，则有可能不必每 10 分钟扫描所述文件。但是，如果新的文件在 10 分钟之前开始通过网络传播，则经常是前两天扫描所述新的文件有意义。一般地，随着时间过去，存在越来越少的关于新的恶意可执行文件的新信息。“集中化的定时分析”特征能够处理这些问题。仅仅需要更新一个分析代理，即中央的那个，所有的主机立即受益。不太可能主机配置与内容分析更新干扰。通过仅仅跟踪新的文件，并且通过基于被暴露到网络的年代（时间）的调度分析，可以有效地和更快地定位和识别新的坏内容。最后，诸如 AV 的许多端点内容分析技术与操作系统紧密地集成。结果，会难于将来自不同提供商的几个内容查看代理布置在一个主机上。分析技术的多样性改善了检测和分类精度。此外，本发明通过使用中央服务器在必要时向不同的服务器发出分析来解决这个问题。

可执行内容（exe 文件）和嵌入的宏（在微软 Office 文件中嵌入的宏）趋向于以簇或者组传播。字处理文件可能包含 10 个宏，并且在大小上超过 30MB，但是所述宏仅仅占用那个空间的一部分。大的安装包会在大小上具有几百 MB，但是其内部文档的可执行部分通常占用总的大小的一小部分。病毒经常作为文档附件（诸如 zip 文件）通过电子邮件传播，以避免检测。在这些文档内，病毒有效负荷可能小。对于所有这些情况，较大的“容器”文件会掩蔽可能不需要的新代码的传播。“内容提取器”特征通过保护（嵌套的）容器关系来处理多个当前限制，并且同时便利：内容的跟踪、类似容器的跟踪、跟踪处理关联性、最小化不必要的重新分析、最小化文件传送带宽和通过将

内容重新打包为其它已知的文件类型而保护与其它分析技术的兼容性。新的内容的中央存储和跟踪、以及相对于内容的第一次出现时间的分析的中央调度在安全、全局可见度、企业管理系统集成和未来扩展上提供了强大的优点。

虽然已经将在此所述的系统与其它系统相区别，但是这样的区别不意味着否认权利要求对于这些系统的涵盖。在此所述的系统和特征可以组合地或者分离地被提供，并且在许多情况下，可以被集成到现有已知的系统中，其中包括如上所述的那些系统。

通过下面的附图、详细说明和权利要求，其它特征和优点将变得清楚。

附图说明

图 1 是示出在此所述的安全系统的概览的方框图。

图 2 是示出在图 1 中的系统的部件的更详细的方框图。

图 3 是图解用于执行分析的处理的流程图。

图 4-图 5 是由所述系统执行的处理的示意图。

图 6 是示出定时分析的示例的图。

图 7 是在定时分析期间执行的步骤的流程图。

图 8 是内容提取处理的示意图。

具体实施方式

参见图 1，也被称为数字抗体系统（DAS）10 的系统允许管理员监控、理解和控制在大的计算机网络上安装的文件，并且可以防御已知和未知的病毒、蠕虫、间谍件、黑客和社会工程攻击以及未被批准的软件（例如不用于商务用途的文件共享软件）。所述系统包括一个或多个服务器，其中之一在此被示出为服务器 14（装置）。这个服务器提供了集中的策略管理、事件监控、代理协调和内容分析（例如间谍件和病毒扫描）。单个服务器可以支持多个主机 12，例如几百或者几千主机。服务器也维护与分析相关联的元数据的数据库，所述元数

据为诸如相对于文件和程序的扫描历史和批准状态。这个元数据被称为每个文件和程序的“抗体”。

每个被保护的主机 12 具有主机代理 16，其优选地被实现为软件。它分析文件系统行为，并且根据在服务器上配置的策略来采取行动。这些策略（在下面更详细地被描述）识别是否阻止、记录、允许或者隔离诸如文件访问或者可执行部分的执行的行为。每个主机代理 16 具有：本地“抗体”存储器 16，其是与文件相关联的元信息的高速缓冲存储器；以及，参数策略引擎 20，用于实现来自服务器 14 的策略。

服务器 14 具有多个功能和接口。所述接口包括：主机通信接口 22，用于与主机通信；基于网页的图形用户接口（GUI），用于与网页浏览器管理控制台 26 通信；报告接口 26，用于作为到企业管理系统 28 的接口；以及远程分析接口 30，用于与内容分析服务 32（例如病毒和间谍件扫描器）通信。服务器 14 也包括分析块 34 和主抗体存储器 36，所述主抗体存储器 36 与抗体分析服务 38 通信并存储相关联的主机的抗体的主列表。服务 38 可以包括非现场（off-site）证书授权，其具有与抗体相关联的附加信息，诸如作为诸如微软 Office 的特定产品包的成员的抗体类别。

图 2 示出了所述系统及其部件的放大图，所述部件包括服务器 14、具有用户和核心部分的主机 12 和其它网络和网页服务 40。如在此所示，服务器包括：新的文件处理和文件池块 42，其包括已经在网络上出现的近来的文件的拷贝；预定分析引擎 44，用于识别要分析的文件和散列值；内容签名器 46，用于使用诸如 MD5 和 SHA-1 之类的算法来建立内容的加密散列；主抗体存储器 36；配置管理 50；记录和报告 52。服务器与网络和网页服务 40 交互，所述网络和网页服务 40 包括分析 54、AV（或者其它内容）扫描器 56 和管理服务 57。

主机 12 的用户部分 60 具有：抗体高速缓冲存储器 64，用于按照名称和数据来保存来自数据库 34 的更新；文件和事件处理 66；分析引擎 68；内容提取器 70，用于提取在包中的感兴趣的内容和独立内容的相关联的分组；内容签名器 72，用于建立内容的加密散列；服

务器元信息（MI）状态解析器 74，用于查看抗体的抗体高速缓冲存储器 64，并且对于服务器查看抗体；以及文件状态解析器 76，用于查看向服务器的内容上载进程，并且对于服务器查看上载的证书。

主机 12 的核心部分 80 具有：高速缓冲存储器 82，用于存储通过文件名称组织的抗体；近来文件操作和文件信息的高速缓冲存储器 84。所述核心也具有截取/阻挡功能 86，其接收和截取文件操作请求，并且向全状态过滤器 88 提供这些请求，所述全状态过滤器 88 首先查看近来文件操作的高速缓冲存储器 84。如果没有匹配，则它查看保存安全策略的触发器和行为块 90。这个块 90 耦接到："defcon"块 92，所述"defcon"块 92 具有用于指示系统的安全级的值；策略引擎 94，其控制块 82、90 和 92 以控制各种文件操作，其中包括执行、文件读取、文件写入和其它行为。所述触发器和行为块 90 与抗体高速缓冲存储器 82 通信，所述抗体高速缓冲存储器 82 根据文件的名称来查找关于文件的元信息。策略引擎 94 也控制行为，诸如阻止、报告或者允许文件操作和向用户报告。

所述系统包括用于设置这个安全系统的多个方法和方面，其中许多被单独使用或者与其它组合使用。下面更详细地说明这些方法和方面。

一个方面是使用中央扫描来查看文件和可执行部分，并且保持用于指示是否已经预先查看了数据的散列值。所述散列值可以被存储在数据库中，并且也被缓存在本地主机中。

另一个方面在于使用中央设置的参数，有时被称为"D"或者"Defcon"，其控制主机的策略。这个中央策略和参数可以被应用到所有的主机，或者被应用到所选择的主机组。所述参数可以被操作员人工设置，或者可以不用人为干预而（通常响应于某个事件）被系统调整。所述策略可以包括阻止或者允许特定行为，或者可以使得一个行为待决，这使得其被允许进行进一步的监控，诸如记录。待决状态具有多个益处，包括考虑在系统中的等待时间，以及实现不适用于传统的二进制批准/禁止模型的策略。这些等待时间包括在识别有害代码之

前的时间，在系统中的误操作期间或者当主机与网络断开的时间。

在另一个方面，中央服务器可以指定元信息的查询，并且将那个查询分布到所有或者所选择的主机组。这些主机执行来自元信息的本地存储器的查询，并且向服务器发回结果，可以使得所述服务器调整参数。

在另一个方面，所述系统包括一种用于防止可以在其它文件中嵌入的宏病毒传播的方法。这种功能可以用于 Visual Basic 宏，但是所述方法可以应用到除了 Visual Basic 之外的任何其它宏语言。

在另一个方面，新文件的所有拷贝被保存在服务器 42 中的特定目录中。可以根据定时器来执行进一步的分析，并且可以在文件被第一次看到后几天后执行所述进一步分析。在文件第一次出现后的某段时间（例如 30 天）后，所述文件可以被重新扫描以查看病毒、间谍件或者其它问题，并且所述系统可以根据结果而采取行动。例如，用于指示在文件中包含病毒的分析将使得这个文件的对应的抗体数据库 36 输入项（entry）包括禁止状态。这种改变以及其它的抗体数据库改变将被传播到主机。

中央设置的参数和参数内容策略

在系统中的安全基于在每个服务器中定义并且通过推和/或拉技术被传播到所有相关联的主机或者主机组的策略。这些策略涉及：可以对于可执行部分和文件所做的内容，诸如读取、执行和写入；当它们被主机建立和改变时要做的内容；扫描如何进行；如何进行记录；以及许多其它的功能，并且对于每个策略（例如对于新看到的可执行部分可以进行什么操作），可以有多个策略选项（诸如禁止、允许或者允许和记录）。所述策略可以基于在文件中的内容（数据）或者文件的名称或者组合。可以通过签名（诸如一个或多个加密散列值）来定义所述内容。采样策略的非专有列表包括：

1. 新的可执行部分和独立脚本（例如*.exe 或者*.bat）的阻止/记录执行

2. 新的嵌入内容（例如在*.doc 中的宏）的阻止/记录读取/

执行

3. 网页内容的阻止/记录安装/修改（在*.html 或者*.cgi 文件中的内容的改变）

4. 允许诸如上述的（3）的策略的更新

5. 自动批准通过两个病毒扫描的文件（例如将对应的文件状态设置为批准）

6. 被管理员特别禁止的文件的阻止/记录安装/执行

7. 通过数据来隔离/删除/记录被感染的文件

8. 通过名称来隔离/记录被感染的文件

9. 在管理地定义的“类”中的新文件的阻止/记录执行；例如管理员可能期望阻止屏幕保护程序*.scr，但是不是全部类别的可执行部分*.exe、*.dll、*.sys 等...

10. 当指定文件被复制到可装卸介质时记录

11. 除了在特定目录中之外，新的可执行部分、脚本和嵌入内容的阻止/记录执行，即允许用户在特殊的目录中建立新的脚本或者可执行部分，但是保护文件系统的其余部分

12. 当离线、远程连接或者本地连接时的不同的主机策略

13. 通过数据或者通过名称而列出包含指定文件的主机/路径

14. 列出具有被阻止的可执行部分、脚本和嵌入脚本的主机

15. 列出具有被感染或者被禁止的文件的主机/路径

16. 自动批准来自被定义的更新服务（例如来自可信来源）的文件

17. 由管理员对于特定的主机组（即存在多个组）特别禁止的文件的阻止/记录执行

18. 由于性能原因或者测试而完全地去活主机系统

19. 在一段时间（用户可配置）后自动批准文件

20. 允许新的文件被安装/执行多达 x 次（用户可配置）。阻止任何更多的安装和/或执行直到被批准

21. 当新的文件被写入时本地批准所述新的文件

22. 当新的文件被写入时中央地批准所述新的文件

服务器可以保存每个主机组的一个或多个策略，并且每个策略按照被中央地设置并且指示策略的选项的参数来可变地被实施。这些策略和选项可以被逻辑地组织为二维阵列，其中，有效的参数沿着一维移动以选择各种策略的策略选项。这个参数在此被称为 D 值。所有的主机可以具有 D 的一个值，或者主机的逻辑子组具有它们子集的 D 的值；例如，在销售部门中的主机可以被分配 D = 1，并且在市场部门中的主机可以同时被分配 D = 2。在一种实现方式中，主机查看（轮询）服务器以查看是否 D 值已经改变。当每个主机发现 D 已经改变时，它们每个开始“移动”到新的 D 值。这种移动可以逐步地进行。可以作为网络消息从主机向服务器提供这些轮询。D 值控制策略行为。对于给定的策略（例如“禁止新的可执行部分”或者“禁止新的脚本”），D = 2 阻止策略侵害行为（在这种情况下，执行“新的可执行部分”），D = 4 警告（向服务器无声警告）但是允许，D = 6 允许而根本不警告。无论是否 D = 2、4 或者 6，主机最好当新的可执行部分被写入时继续注意和记录所述新的可执行部分。当在此的示例使用 D 的数值时，D 可以具有以字母、字或者字母和数字的任何组合表达的“值”。

所述 D 值也控制策略激活。对于给定的策略（例如“禁止新的可执行部分”或者“禁止新的脚本”），D = 1 使能“写保护”策略，因此，根本不能写入新的可执行部分，而 D = 8 完全禁止所有的策略，D = 2、4 和 6 情况可以像如上所述设置的那样。在这种情况下，D = 8 可以甚至禁止注意何时新的可执行部分被写入到文件系统的策略。

当可以在服务器中中央地设置 D 值时，其被本地地实现在主机上。可以由管理员利用连接到服务器的浏览器通过在管理控制台上的图形用户接口（GUI）、或者经由简单网络管理协议（SNMP）而将其设置。D 值被当做“目标”值；主机试图移动到尽可能地接近这个值，这可能需要几秒或者几分钟。在一些情况下，主机可以从由服务器指定的目标值本地地分出。可以在主机上调用命令行程序，或者用户可以被提示 D 的特定值，并且可以改写 D 的目标值。例如在个人的机

器需要禁止安全（D＝8）并且没有与服务器的网络连接时，这个特征是有益的。特定的行为可以自动改变在主机上的 D 值，诸如检测来自授权程序的更新（例如反病毒更新）。

所述策略反映了在安全和可使用性之间的折衷。在上述的示例中，D＝8 最为有用，并且最不安全——没有策略被激活，主机代理被有效地禁止阻止和跟踪。当 D 向最大安全（D＝1）移动时，越来越多的限制策略被激活，并且当策略被违反时执行的行为变得越来越严重。有序的状态是所期望的，因为它们更容易可视化和测试（一般，可以仅仅测试需要测试的端点，诸如 D＝1 和 D＝8）。当所述值提高或者降低时，使用有序的状态，文件和用户的数量变得连续地更容易访问或者更受限制。这些有序的状态自然地反映安全和可使用性之间的折衷。

当 D 在活动的系统上改变时，会发生竞争条件。基本问题是：如果在安装程序时 D 的值要从 8→1 改变，则多个文件的安装会变得“半阻止”或者“半安装”。结果，特定的 D 转换可以触发文件抗体状态重新分析和文件抗体块状态变换。

本地 D 改变有时可以被本地策略触发器引起。通常，在服务器上中央地设置 D。但是有时，触发本地主机策略，其然后使得本地主机 D 值改变。这有益于例如完成在被锁定的系统上的安装（D＝2）。继续这个示例，在 D＝2 安装打印机驱动器可能导致问题，因为一些打开的新的安装文件需要执行以完成所述安装。而且，不同的主机机器可能需要打开和执行不同的程序以完成所述安装（例如 Windows 2000 和 Windows XP）。在这种情况下，一个特定抗体文件类型——被批准的程序“printer_setup.exe”——的执行将主机的本地 D 从 2→3，这是略弱的状态，其自动本地仅仅批准这些新的安装文件和它们的后代。

可以根据连接类型来改变 D 值，不论是否为本地（在有线 LAN 上）、远程（诸如通过电话调制解调器或者虚拟专用网络（VPN））或者完全断开。主机代理因此存储这些类型的连接的一组指定的 D

值，然后自动响应于改变（例如当用户将主机与 LAN 断开时）而从所述组中选择。而且，不同的 D 值可以导致在报告、记录和跟踪细节上的减少或者增加。

也可以从中央服务器（有时被称为“超级服务器”）设置策略，所述中央服务器可以控制许多服务器/服务器。假定每个服务器控制 2000 个主机，并且存在 1000 个超级服务器，则不可能用于设置 D = 1 的超级服务器命令将对于所有的 2000000 个主机适当。相反，超级服务器可以命令所有的服务器和主机具有本地允许的尽可能强的 D。因此，一些服务器和它们所连接的主机将到达它们的极限例如 D = 2。其它服务器可以进行到 D = 1，但是然后也许它们的主机组的一些限于 D = 4，因此那些主机将走强，但是不强于 D = 4。同一限制对于上述范围的另一端为真。如果超级服务器命令 D = 8，则一些服务器和主机仅仅可能取而代之走到 D = 6。因为 D 是有序状态，因此这些限制是简单的整数范围（最小值和最大值）。

D 的值可以根据某个事件的检测而改变，诸如传播文件。如果新的文件的太多的拷贝正在服务器的多个主机上传播，则服务器可以选用地将 D 提高以停止所述传播（例如走到 D = 2）。这种事件可以被指定为特定名称的太多（例如按照名称列出的前 10 个）或者按照唯一内容的太多（例如按照数据的散列值列出的前 10 个）。

所述值也可以响应于由服务器感知的新的事件而根据服务器请求改变，所述事件诸如新来到的文件或者可能的病毒攻击。在大多数情况下，是管理员（一个人）按照所计划的用户操作或者根据特定文件事件的观察而启动 D 的改变。可以自动改变 D，例如在操作的进行期间，在这种情况下，主机/服务器在操作终止后将 D 的值返回到其原始水平。外部触发器可以改变 D 的值，诸如 SNMP。

另一个响应是服务器自动批准在少于特定阈值数量的主机上的内容，但是当超过所述主机数量时自动禁止对于那个内容的访问。这样的策略可以用于限制在网络中的任何内容或者文件的拷贝的数量。而且，这样的策略可以用于仅仅报告超过特定主机数量的内容。

服务器可以对于每个主机逻辑组（诸如销售主机、市场主机和工程主机）分别保存策略集。策略集可以具有与抗体版本号码类似的唯一识别号码。差别在于一旦被部署，则策略集变得“只读”以协调策略集的以后的问题，并且取消问题部署。也可以使用与 Unix 实用程序“diff”和“patch”类似的技术对于差别配置和其它更新如此进行。主机可以向服务器查询它们的组的当前策略集的 ID 号，并且如果存在不匹配，则它们可以向所述服务器发送“GetPolicySet”查询。

策略集可以包括多个策略，诸如“新的可执行部分”策略或者“新的脚本”策略。每个策略可以处于活动（接通）、不活动（断开）或者测试模式（其中，允许阻止，但是向服务器发送“将阻止”消息）中。每个策略可以具有多个规则，每个规则具有基本的“触发和行为”模型。触发器是被测试的模式。如果模式匹配，则执行所产生的行为。例如，可以将“在 D = 2 的新的可执行部分的阻止执行”指定如下：

Trigger=(D=2 & FileOp=Execute & State=Pending & FileExtensionClass=ExecutableClass)，其中，ExecutableClass =（*.exe|*.sys|*.dll|...）

Action=(Block & Report & Notify(P))，其中，“Block”停止操作，“Report”向服务器发送通知，“Notify”使用参数集 P 来警告用户。

使用这种结构，除了在核心抗体高速缓冲存储器更新、D 更新和策略集更新的情况下，所述核心可以不用与用户空间交互而实施所有的策略。策略集仅仅需要被存储在一个位置，并且在这种实现方式中它们仅仅需要在核心中被解释。策略集可以被认证和存储在一个安全的背景（所述核心）中，结果产生对于篡改的更强的安全性。

通过 D 来对策略和行为进行参数化，因为 D 允许不同的规则匹配不同的触发器。具有特定状态的文件可以阻止特定的操作。这些状态可以是名称和数据属性的组合。这些状态在用户空间中被确定，在核心空间中被镜像，并且最后，所述状态被服务器确定。一种有益的策略是阻止被禁止的文件，并且在一些 D 值的情况下，阻止待决（新）文件的执行。

所述策略可以作为策略的一组列表而被提供在具有访问性和安全性的折衷的范围上。所述服务器可以然后提供信息以使得主机选择所述列表之一。通过使得所述多个列表存在于所述主机上并且允许所述主机使用“拉”手段来更新策略，所述主机可以方便地在服务器的控制下更新安全策略。

下面的表格示出了所述 D 值如何可以影响在组策略集中的各种策略的一个示例，其中，行是在主集内的策略，列是行为，并且所述单元具有用于指示所述行为的 D 的数值范围。在所述表中指定的行为和其它细节被汇总如下：

策略名称 \D 值	D = 10 全局批准	D = 8 禁止保护	D = 6 仅仅跟踪	D = 4 无声警告	D = 3 本地批准	D = 2 锁定	D = 1 写保护
新/待决的可执行部分*.exe, *.sys, ...	自动全局批准新的，报告	允许	允许	允许执行，报告	自动本地批准新的，报告	阻止执行，通知，报告	阻挡写入/执行，通知，报告
新/待决的单独脚本*.vbs, *.bat, ...	自动全局批准新的，报告	允许	允许	允许执行，报告	自动本地批准新的，报告	阻止执行，通知，报告	阻止写入/执行，通知，报告
在*.doc, *.xls, ... 中的新/待决的嵌入脚本	自动全局批准新的，报告	允许	允许	允许执行，报告	自动本地批准新的，报告	阻止执行，通知，报告	阻止写入/执行，通知，报告
新的网页内容*.html, *.asp, ...	自动全局批准新的，报告	允许	允许	允许写入，报告	允许写入，报告	写入保护，报告	写入保护，报告
批准（散列值和/或名称）可执行部分/脚本/嵌入	允许	允许	允许	允许	允许	允许	允许
禁止/未批准的（通过散列值）可执行部分/脚本/嵌入	允许	允许	阻止执行，通知，报告	阻止执行，通知，报告	阻止执行，通知，报告	阻止执行，通知，报告	阻止执行，通知，报告
禁止/未批准的（通过名称）可执行部分/脚本/嵌入	允许	允许	阻止执行，通知，报告	阻止执行，通知，报告	阻止执行，通知，报告	阻止执行，通知，报告	阻止执行，通知，报告
内容改变和内容建立跟踪	跟踪，报告	允许	跟踪，报告	跟踪，报告	跟踪，报告	跟踪，报告	跟踪，报告

（1） 允许：允许操作，否则无声

（2） 阻止：阻止操作，否则无声

（3） 跟踪：跟踪操作和所产生的内容（如果所述内容是待决的或者被禁止的），否则无声。一般不跟踪被批准的内容。

（4） 报告：向服务器发送通知。

（5） 通知：向主机端点用户指示为什么操作被阻止/中断。

（6） 自动本地批准：具有本地主机状态＝待决的新的主机文件和/或新的内容被本地设置为：当建立/修改文件/内容时，主机状态＝批准或状态＝仅仅在本地主机上本地批准。

（7） 自动全局批准：具有本地状态＝待决的新的主机文件和/或新的内容被全局地设置为：当建立/修改文件/内容时，在服务器上的服务器状态＝批准。

抗体引入，文件元信息

具体参见图 2，对于被允许的行为，在系统中的服务器包括抗体数据库 36，其主要用于跟踪文件扫描历史和每个文件的批准状态。抗体是关于文件的数据块（即元数据或者元信息），其可以包括一些或者全部的下面的字段：

● 第一次看到的时间。当文件或者散列值被主机第一次看到并且被报告到服务器的时间。

● 文件 ID。文件的唯一标识符，包括诸如 MD5、SHA-1 和 OMAC 的内容的一个或多个散列值。

● 文件类型。文件类别（例如可执行、脚本，办公文件、文档等）。其是当第一次看到文件（见下）时从文件名称得到的，并且也从文件内容的分析得到。

● 状态。当前文件的状态，包括批准、待决或者禁止。

● 方法，服务器得知文件的方式（自动，人工等）。

● 文件名。被第一次看到和报告到服务器的文件的名称。其可以不是文件的当前名称，而仅仅是在网络上看到的第一实例的名称。

● 文件路径。被第一次看到和报告到服务器的文件的路径。

● 当第一次被看到/报告时的主机文件名称/路径/扩展。

● 当最后被看到/报告时的主机文件名称/路径/扩展。

● 第一次看到/报告的主机 IP 地址文件。

● 第一次看到的主机，在其上第一次看到和报告文件或者散列值的主机的名称。

● 分析结果。最新的扫描或者其它分析的结果。

● 第一次分析。文件的第一次扫描/分析的时间。

● 最后的分析。文件被最后扫描/分析的时间。

● 最后更新。文件状态被最后修改的时间。

● 父容器。到已经与文件相关联的其它文件的链接。

● 父容器属性。文件名称、第一次看到的时间、第一次看到的主机、文件路径、产品类别和一个相关联的容器文件的状态。

● 根容器。到已经与所述文件相关联的其它文件的链接。根容器是未在另一个容器内包含的容器。

● 根容器属性。文件名称、第一次看到的时间、第一次看到的主机、文件路径、产品类别和一个相关联的根容器文件的状态。

● 参考父文件容器，如果已知的话。这些用于保存包含关联，诸如“在散列值 = x 的文档文件中观察到这个散列值 = y 的文件”。

● 文件内容类型（通过内容分析而确定），诸如可执行、脚本文件、嵌入的宏。

服务器具有系统的抗体的全集。当每个主机在用户高速缓冲存储器 64 中和在核心高速缓冲存储器 82 中包括抗体的本地子集时，服务器是用于设置和改变到特定状态的主管方（authority）。例如，服务器是中央地发起和（向主机）传播改变的主管方，所述改变包括从待决向批准或者禁止的状态过渡（这三种状态最好与内容散列值相关联），而主机是唯一可以将状态设置为本地批准的主管方。在数据库 36 中的每个输入项是永久的，并且最好容易使用文件数据散列值索引被访问。数据库可以选用地由其它关键字加索引，诸如文件名、第一次看到的日期、状态、分析结果、主机 ID 或者主机计数，以便管理

员可以容易地浏览抗体数据库。

当具有抗体的数据库被描述为正处于服务器之内或者之上时，应当明白，这表示所述数据库与服务器相关联。其可以物理地驻留在同一机箱和服务器的处理功能内，或者其可以驻留在不同的机箱或者甚至在远程位置中。如果远程的话，则应当存在适当的有线或者无线的连接来获得数据。

抗体（AB）跟踪引入

当建立新的文件或者修改现有的文件时，可以触发跟踪策略，由此引起一系列文件和抗体分析事件。首先，主机执行一系列步骤来确定是否已经有对于内容的重大修改，所述内容对应于已经被分析的内容，并且对于所述内容而言，已经在主机高速缓冲存储器中存储了抗体。如果内容抗体未在主机高速缓冲存储器中，则向服务器查询以确定是否所述服务器已经分析了内容。如果服务器没有对应的抗体，则所述内容可以被上载到服务器以进行进一步的分析。直到所述服务器可以确实地确定状态，与所述内容相关联的状态被设置到待决或者还没有被确定。可以限制对于待决内容的随后的访问。服务器根据自从所述内容在服务器上被第一次看到的时间来对于内容进行分析。根据所述分析或者其它外部确定，服务器可以确实地确定在状态上的改变。这些改变可以被主机指示来用于以后的检索，因此主机可以使用所改变的状态来更新它们的抗体高速缓冲存储器。

主抗体跟踪

参见图 3，主机截取文件操作（501），包括执行、读取、重新命名或者写入，并且向全状态文件操作过滤器（502）提供所述操作。如果所述文件名称不在核心高速缓冲存储器中，并且存在核心高速缓冲存储器未命中（miss）（510），并且如果已经有可能的文件或者内容修改（511），则所述状态无效。所述文件然后去往内容提取器，内容提取器提取感兴趣的有效内容（503）以产生减小的文件（如下更详细所述），并且向内容签名器提供所述减小的文件（504）。内容签名器向减小的文件应用诸如 MD5 之类的加密散列。这个散列与

文件和文件名相关联。可以当散列和其它分析（高速缓冲存储器未命中解析）完成时延迟/停止文件操作。

主机也根据散列值内容来进行本地查找，以试图获得状态（505）。如果内容和状态未被找到，则所述状态被设置为待决。这可以表示文件操作被允许进行，虽然诸如记录之类的另外的监控也可以发生。如果所述内容被找到，则名称、内容、容器（包含有效内容的文件）和状态全部关联在一起（507）。否则，主机向服务器请求在其存储器中查找内容（506）。如果在那里找到了，则名称、内容、容器（包含有效内容的文件）和状态全部关联在一起（507）。如果内容和状态未被找到，则所述状态被设置为待决，并且所述内容被上载到服务器（508），服务器确认所述上载（509）。服务器也可以查看与多个服务器相关联的“超级服务器”。容器关系被存储，并且与文件和其它容器相关联。容器信息也被发送到服务器和主机以及被发送来用于分析。“根”容器是未被另一个容器包含的容器。通过它们相关联的文件以及通过加密散列值来识别容器。

一般地，向文件内容的“有效”部分或者整个文件内容的散列值或者签名分配抗体状态。因此，一般地，HASH（文件数据/内容）→状态。这映射数据→状态。状态（S）可以包含多个信息，诸如“批准”（白名单）或者“禁止”（黑名单）或者“待决”（诸如还没有被完全分析的新看到的文件的“灰名单”）。

这个系统的优点是将名称状态与内容状态组合。例如，服务器可以指定和存储多个名称禁止，诸如*msblast.exe。服务器将名称状态策略存储为正则表达式和相关联的元信息的列表。与所述正则表达式匹配的任何文件驱动器/路径/名称/扩展然后继承名称元信息。每当文件名被改变或者名称元信息指定改变时更新这个名称元信息。从服务器向主机传播名称状态和策略。例如，通过加上*msblast.exe→NameBan，服务器将感测到新的策略/状态，并且将向主机传播那个规格。主机然后对于它们的名称元信息高速缓冲存储器搜索与*msblast.exe 的匹配，并且那些匹配的文件将继承 NameBan

状态。主机文件状态是名称和数据状态的叠加：例如，如果temp_msblast.exe 具有内容状态 = 待决，则其组合状态是禁止，因为NameBan 相对于待决具有优先权。以类似的方式来处理名称批准状态。

抗体被分层地存储在数据库中。存在如上所述的抗体的四种主要存储位置。在主机代理中，核心抗体高速缓冲存储器 82 映射文件NAME（名称）→抗体 STATE（状态）。例如，NAME = c:\windows\bar.exe→STATE=approved。简写的话，这种映射是N→S。核心可以并且实际上根据所述状态来实施策略，而不需要访问文件内容。当文件可以在核心中被加密但是以未加密以上的形式可见时，这是有益的。所述核心直接访问名称，但是不访问散列值。由于可以有较长延迟（秒、分钟、小时、天），核心高速缓冲存储器可以较弱地与其它的高速缓冲存储器一致，并且最后与服务器一致。

主机代理具有用户抗体名称高速缓冲存储器（UN）和用户抗体数据高速缓冲存储器（UD）60。UN 将文件名映射到文件内容（数据）的散列值，即 UN 映射 N→数据。并且类似地，UH 将数据映射到状态 Data（数据）→S。一般地，N→数据的映射是多对一，UN 镜像本地文件系统的结构。所述数据→S 的映射一般是一对一的，因为对于优选使用的强散列值（诸如 MD5），散列值冲突是很少见的。UN 和 UD 高速缓冲存储器也与服务器弱一致，但是 UN 和 UD 与本地主机文件系统强一致，就像核心高速缓冲存储器那样。UN 和 UD 可以被组合如下：N→数据→S = N→S。

服务器具有已经被其任何主机报告的一般每个唯一的散列值的抗体数据库 34，并且超级服务器（如果存在一个）具有已经在任何其服务器上看到的一般每个唯一的散列值的抗体数据库。限于唯一散列值限制了存储和处理，虽然可以使用在存储和处理上的进一步的改善来存储更多。而且，限于唯一散列值导致更有效的分析和更低的网络流量。

一般地，新的文件响应于“新文件”或者“脏文件”事件而从主机向

服务器向超级服务器传播，并且新计算的抗体状态相反地以抗体更新的形式从超级服务器向服务器向主机核心传播。以这种方式，中央地控制、管理和验证抗体。所述服务器“拥有”和证明所述抗体，并且所述服务器提供所述抗体还没有被改变或者伪造的认证。主机保存它们自己的抗体，其一般（但是不必要）对应于在服务器上的那些抗体。因此，损害的或者误操作的主机不能使得服务器或者超级服务器抗体集变差，损害的主机也不能使得其它主机的抗体变差。

在主机上，抗体状态最好被存储，以便其不与散列值/数据相关联，而是通过名称相关联。核心分析、解译和实施策略，并且按照名称来查找所述文件的状态。可以明白，优选的实现方式在核心中实施策略，但是其它的实现方式可以在用户空间中实施策略。当查找状态时，在用户空间或者核心中，其实际是确定结果产生的状态的混合体。例如，如果 foo.exe 的数据抗体待决，但是名称抗体根据其名称而被禁止，则 GetABState（foo.exe）返回“按照名称禁止”的结果。存在用于阻止具有抗体状态 = NameBan 的文件的执行的独立策略。那个策略的行为被 D 的值如上参数化。一个差别是阻止“按照名称禁止”的策略在较低的 D 安全设置有效。例如在 D = 4，“待决”文件将执行（使用无声警告），但是被禁止的文件不执行。

名称禁止被表示为正则表达式的列表，并且可以在服务器上包括通配符（*），例如“*oo.exe”或者“*msblast.exe”。这些列表具有版本号码。当主机轮询时，它们查看它们的版本号码。当主机检测到不匹配时，其然后从服务器发送 GetNameBans 查询（即主机从服务器拉新的禁止数据）。然后，对于名称抗体重新评估这些正则表达式。名称禁止是状态的属性，并且仅仅当名称禁止列表改变时或者当文件名改变时被重新计算。不必在每个文件操作比较通配符列表。因此，数据抗体和名称抗体的双特性有益。而且，成百上千的名称正则表达式可以同时有效，而不对于每个文件操作要求在核心中进行几千个正则表达式匹配计算，这会是极为昂贵的。

文件内容跟踪

向回参见图 2，截取/阻止功能 86 可以截取并且读取文件访问请求。其可以在获得策略信息的同时中止请求，根据在核心内的策略来阻止请求，并且对于被阻止的请求返回适当的错误代码。功能 86 从文件访问请求读取请求处理名称、所述请求的本地系统时间、所请求的数据（包括全路径）和所请求的行为（例如读取、写入或者执行）。在一个实施例中，功能 86 向“全状态过滤器”88 提供所有的文件访问请求，每个操作被阻止，直到过滤器 88 返回标记，所述标记指示所述操作或者被阻止或者被允许。

过滤器 88 从功能 86 截取文件访问请求，并且对于大多数文件访问请求返回“阻止”或者“允许”的行为。不能与已经被批准的文件访问请求相关联的任何文件访问请求被转发到核心触发器和行为模块 90，其返回“阻止”或者“允许”的行为。这个行为被过滤器 88 存储，并且最好对于任何随后的相关联的类似文件访问请求被发送到功能 86。

过滤器 88 维护已经打开的文件的高速缓冲存储器 84（通过核心范围的唯一标识符来索引；例如在 Windows NT 中的核心文件句柄）。每个高速缓冲存储器输入项包含文件标识符（核心文件句柄），并且阻止或者允许对于读取、写入或者执行的许可。

如果多个处理访问同一文件，则每个将具有其本身的高速缓冲存储器输入项。如果给定的处理尝试新的文件访问，则全状态过滤器将经历那个文件的高速缓冲存储器未命中，这将使得其向所述触发器和行为模块提交文件访问请求。如果模块 90 允许，则所请求的操作（读取、写入或者执行）的标记应当被设置为“允许”。否则，其应当被设置为“阻止”。如果仅仅获得一种许可（例如读取）的处理然后尝试另一种处理（例如写入），则模块 90 将再一次被接触。

其使用期限超过特定值（例如 60 秒）的高速缓冲存储器输入项可以被删除。这允许剪除由于某些原因未被去除的输入项。这也允许通过模块 90 定期重新查看文件。

在这个示例中，通过用于文件“foo.exe”的主机代理核心程序（HK）来在阻止垫片（shim）86 中的核心中捕获文件写入操作。在

D = 4 的值下，文件操作（在此为文件写入操作）被激活的“脏跟踪”策略捕获，并且这从主机核心程序向主机代理用户空间程序（HU）发出“脏”事件。这个事件指定文件名和脏操作。对于这个操作不参考核心高速缓冲存储器 82，因为脏跟踪策略使得那个字段为空。

HU 然后对于 foo.exe 执行在文件和事件处理 66 和分析引擎 68 中的多个本地分析操作。首先，查看 foo.exe 以看是否其存在，是否是可读，是否其确实是可执行部分。可以执行其它的操作，诸如在过滤器 88 中提取“感兴趣的数据”；例如，如果文件是 foo.bat，则可以去除脚本说明。所提取的 foo.exe 的数据然后被以加密方式散列化，并且这个散列值用于尝试在 HU 抗体高速缓冲存储器 60 中的查找表。如果名称和数据已经存在，则不进行任何其它事情。如果名称是新的，但是所述数据是已知的，则在 UN 高速缓冲存储器中建立新的名称抗体。这个处理是被称为“阶段 1 分析队列”项目的所有部分。可以排队在主机上在阶段 1 队列中等待散列化的许多文件。阶段 1 队列仅仅具有名称抗体和元信息，因为所述数据还没有被知道或者分析。

如果主机已经看到这个文件数据和散列值，则那个散列值的对应的已知元信息与那个文件的主文件元信息相关联，所述元信息是从 UD 本地存储器或者本地盘存储器以那个顺序被检索出来的。如果主机还没有看到这个数据，则 UD 高速缓冲存储器“未命中”。所述散列值被置于阶段 2 分析队列。实际上，存在数据抗体，即，逻辑地跟踪数据的状态（诸如“批准”、“禁止”或者“待决”），并且也存在名称抗体，例如“按照名称禁止”。例如，如果服务器禁止“oo.exe”，则 foo.exe 的名称抗体将指示“NameBan”，并且名称禁止策略可以据其来阻止。因此即使高速缓冲存储器可能知道已经（通过名称）禁止了 foo.exe，脏跟踪解析仍然继续。名称和数据抗体的这种区别在范围上对于独立的主机是本地的，但是其对于 FindFile（查找文件）功能（下述）和对于策略实施变得重要。所述数据抗体因此被置于阶段 2 队列中。

阶段 2 分析尝试首先从存储器高速缓冲存储器、然后从基于本地盘的数据存储器、然后从服务器解析本地状态信息。如果服务器被连

接，则阶段 2 队列将当解析元信息时为空。当从这个队列去除 foo.exe 时，如果未本地找到那个散列值，则向服务器查询是否它已经看到这个数据散列值。如果回答是否，则 foo.exe 及其散列值和其它元信息被置于阶段 3 队列中以上载到服务器。另外，如果服务器还没有看到所述散列值或者如果服务器分析还没有完全结束以确定其它状态，服务器将向主机发送默认的抗体状态，即“待决”。如果服务器已经计算了有效的抗体和状态，则其返回这个抗体元信息。如果服务器从未看到 foo.exe 的这个数据，则在所述服务器的经历中的所有机器从未看到这个文件的意义上来说其是新的。

当从阶段 3 队列去除 foo.exe 时，其使用加密的单向传送被上载到服务器。即，使用 FTPS（安全文件传送协议）和只写服务器目录，文件可以被上载到服务器，但是不能被下载。当成功地完成上载时，主机向服务器通知传送了 foo.exe。这种传送通过散列值被引用，以便最小化信息泄漏和用于附加的安全性。

当服务器知道上载了 foo.exe 时，像主机那样，其通过经由几个阶段分析文件而开始。在这种情况下建立新的抗体，所述服务器使用其被同步的被验证时钟来记录其首次出现的时间。而且，执行所述提取和散列化，并且那些结果取代所述主机的相应内容。

服务器分析按照在服务器上指定和存储的时间表。这个时间表与文件或者其散列在服务器上的第一次出现时间相关。例如，如果文件在中午到达并且时间表是“在+0 散列查找并且在+0 AV 扫描并且在+2 小时 AV 扫描”，则在中午，将使用外部散列查找服务来计算和查找文件散列。然后，执行 AV 扫描。两个小时后，在下午两点，执行那个文件的另一个 AV 扫描。用于描述所述时间表的另一个方式是其与“在服务器上的文件使用期限”相关联。

当抗体在服务器上改变状态时，递增的计数值被写入到所述抗体中。这个计数用于仅仅选择自从任何特定的主机或者超级服务器登录起已经改变的抗体的范围。例如，如果前一个抗体改变是从待决→批准过渡的 glorp.bat 并且全局抗体版本计数是 277，则对应于 glorp.bat

的散列的服务器抗体将获得版本号 277，并且计数将是 278。因此，对应于抗体 foo.exe 的版本号是 278，并且计数是 279。

当主机定期轮询时，它们提供它们最后的抗体版本号，并且服务器将发送自从最后的轮询起已经改变的所有抗体。优选的是，所述服务器发送当前的编号，并且当主机意识到不匹配时，其向服务器查询抗体更新，并且返回数据抗体的列表。这些然后被合并为主机抗体，并且也向核心内下发改变。虽然主机可以获取并且存储从未看到的数据的一些抗体，一般仅仅对应于现有的主机文件的那些抗体被合并。通常丢弃其它的抗体。服务器捕获最后几分钟的更新，以最小化对于每个主机定制所有的更新的效果。此外，因为主机通常获得比它们需要的更多的抗体，并且因为新的抗体很少见，因此这个流量是有限的。抗体更新小，大多数其它消息都如此。

抗体可以以类似的方式保持与超级服务器同步。在此，超级服务器可以轮询服务器，并且获得抗体更新列表。超级服务器可以合并它们，并且发出每个服务器的定制的更新。这些更新全部在它们会滞后几分钟或者几天上弱一致，但是必须有互锁（interlock）和保护以避免在更新中的“空洞”。

存在与抗体的合并相关联的其它方面和特征。例如，一些服务器可能不从超级服务器接受特定的抗体更新。而且，主机将不允许特定的本地状态改变到特定的服务器指定状态。

一个问题是关于高速缓冲存储器的初始状态和初始策略。服务器高速缓冲存储器可以预先安装已知的良好和坏的散列抗体，或者其可以是空的，都是可以的。但是，主机必须偶尔“点滴式充电（trickle charge）”。例如，当主机首次连接到特定服务器时，检测到这个事实，并且主机将执行点滴式充电，其中，在主机文件系统上的每个单个感兴趣的文件被插入到阶段 1 队列中。在此处理期间使用 D 的特定值以保证不确定的高速缓冲存储器将不引起问题。抗体一般全部以状态“待决”开始，并且它们缓慢地与服务器同步。而且，所有的主机抗体和队列信息和相关联的全局符定期地被保存，并且过度重新启动。

核心高速缓冲存储器一致性

在主机代理的启动或者其它初始化时，对于具有有效元信息的每个已知的现有主机文件，所述核心被安装从用户空间已知的每个有效抗体。一些抗体更新当它们被从服务器或者从在用户空间中的分析队列接收时被发送到核心中。但是，某些更新是核心高速缓冲存储器未命中的结果。如果确定一个策略有效，并且如果需要抗体状态而如果那个状态不可获得，则所述核心一般将所述操作延迟某个时间，并且向用户空间发送核心未命中事件。即使不立即需要抗体，一些事件也可以被延迟。这是当一个策略通过与用户接口（弹出消息框）交互（例如点击是（yes）来覆盖（override）被阻止的待决操作并且使得随后的受限操作接续，而不阻止某个时间）来允许主机用户覆盖受限状态（待决）的情况。

在一个示例中，安装程序打开被称为 inst.exe 的新的程序，然后将其重新命名和执行。通过在执行所述分析的同时延迟所述重新命名并且延迟所述执行，核心将避免暂时的不一致。结果产生的抗体从用户空间异步地被发送，然后一旦完成了所述异步更新，则待决操作解除阻止，并且使用所需要的状态信息来评估策略。

所述核心高速缓冲存储器在初始化时包含在文件系统中的几乎所有文件的抗体。可以在核心高速缓冲存储器中留下空洞或者其它不一致的操作（即使是短时间的）被延迟并且互锁，以便保持一致性。用户空间高速缓冲存储器被优化以便以很低的延迟来解析核心未命中。当核心和用户空间高速缓冲存储器对于服务器侧延迟很不敏感时，核心高速缓冲存储器对于互锁和正确的持久性敏感。

查找文件（FindFile）

因为 UN 和 UD 高速缓冲存储器最好被优化以用于低延迟的查找，因此这些高速缓冲存储器可以被用作来自服务器的分布式抗体查询的一部分（在此被称为“FindFile”功能），以产生什么文件在什么主机上的视图。通过经由在服务器或者超级服务器上的网页接口提交网页浏览器形式，管理员可以指定 FindFile 请求。例如，可以联合地

指定下面的限定符:

（1） 文件名的正则表达式模式指定，

（2） 文件路径的正则表达式模式指定，

（3） 文件的感兴趣的内容的散列值，

（4） 与文件相关联的容器的散列值或者其它 ID，

（5） 当主机第一次看到文件或者文件的散列值时的时间范围，

（6） 主机的名称，

（7） 主机的 IP 地址

（8） 文件的类型

（9） 来自一组至少三个状态（批准、禁止、待决分析）的、与文件相关联的一个或多个主机文件状态。例如，一组 AllBanned=(NameBan, BanByHash)。

（10）是否主机对于文件执行了特定的文件操作，以及

（11）主机组

参见图 4，完成的 FindFile 请求与电子邮件的相似之处在于：服务器发布（post）对于由指定主机进行随后的检索的请求。当主机登录（check in）时，它们获知是否存在来自服务器的 FindFile 消息在等待它们。当一个主机获知其具有未完结的 FindFile 请求时，其使用获得查找文件请求（GetFindFileRequests）来检索所述请求，GetFindFileRequests 如图 4 中的线（1）所示。换句话说，所述请求最好被实现为从服务器的“拉”。这允许更安全的实现，而不需要监听的主机插口。

所连接的主机的每个通过访问来自它们的抗体高速缓冲存储器的适用数据而处理它们的 FindFile 请求，并且将结果列表发布到结果数据库，被示出为发布查找文件结果（PostFindFileResults）（图 4 中的线（2）），其中包括下面的所返回的每个文件的信息的一些或者全部:

（1） 文件名，

（2） 文件路径，

（3） 文件的感兴趣的内容的散列值，

（4） 当主机第一次看到文件或者文件的散列值时的时间，

（5） 主机的名称，

（6） 主机的 IP 地址，

（7） 文件的类型，

（8） 文件的容器信息，

（9） 来自一组至少三个状态（批准、禁止、待决分析）的、与文件相关联的一个或多个主机文件状态，

（10）是否已经由主机对于文件执行了特定的文件操作，以及

（11）主机组

在一种实现方式中，通过主机首先连接到服务器并且发送一个或多个网络消息，并且在断开之前接收对于主机消息的服务器答复，完成所有的主机服务器通信（不仅仅是 FindFile）。同样，这具有更安全的优点：因为不需要监听主机插口。存在另一个优点：仅仅需要维护服务器寻址和路由，而不维护主机寻址、路由和减少发现这样的主机信息的必要。

服务器混合和建立来自主机的 FindFile 结果列表的主列表。这些列表的并集是完整的 FindFile 请求响应，并且其随着时间建立，通常以少于一分钟的时间完成。因为本地主机处理仅仅访问抗体高速缓冲存储器，而不访问主机文件系统，因此这些查询会很快。双重名称和数据抗体关联系统和高速缓冲存储器允许这一点。服务器然后例如通过网页接口向管理员输出结果。而且，特定的 FindFile 结果可以影响和触发 SNMP、syslog、警告和其它通知系统。

超级服务器也可以以类似的方式发布要由服务器访问的请求，或者超级服务器可以直接向服务器递交 FindFile 请求。然后，服务器可以向超级服务器返回合并的结果，超级服务器然后可以将这些合并为较大的主结果。这类似于当处理 FindFile 请求时在服务器和主机之间的关系。

定时器触发的中央分析

参照图 5，服务器可以根据事件来执行分析，例如每次主机上载内容时的分析，或者系统可以根据时间来执行这些分析。如上所述，新的内容可以被上载到服务器，并且使用外部和/或内部分析代理来执行分析以建立被存储在数据库中的元数据或者元信息。例如在当上载新的内容时相对于文件的第一次观察的特定时间间隔后，所述系统可以然后查看另外的所计划的分析。服务器和超级服务器可以执行许多类型的另外的基于时间的分析。

参见图 6，当首次看到一个文件并且其抗体被加到服务器数据库时，效果是好像对于每个文件启动定时器。因此，例如，时间间隔可以是（在第一次看到或者向服务器的报告后，t = 0 = 立即，t = 12 小时之后，t = 2 天之后，t = 30 天之后），并且可以基于服务器的时钟。除了一次的定时行为之外，可以指定周期的行为。如在此所示，可以在不同时间执行反病毒（AV）和反间谍件（AS）扫描，并且可以执行其它分析。对于以后的时段，这可以是与可能已经查看了所述文件的其它服务器的比较。通常，以后的分析将基于在某个时段内首次看到的所有文件。例如，在一个小时时间内首次看到的所有文件将在所述时段内的最后文件起的 12 小时获得 12 小时分析。

参见图 7，系统选择要分析的文件，并且发送文件以执行指定的分析。可以对于每个时间间隔指定不同的操作。因为文件在服务器上被保存一段时间，这些时间启动的分析可以进行，而不论原始主机是否仍然连接。可以执行的中央定时服务器分析的示例包括:

（1） 计算替代散列值（例如使用 MD5 或者 SHAI 算法），验证被报告的散列值，并且存储所有的散列值。

（2） 使用服务器证书或者其它第三方证书来认证和签名内容。

（3） 本地或者经由另一个服务器的查询而查找相对于已知的坏数据库（黑名单）的散列值。

（4） 本地或者经由另一个服务器的查询而查找相对于已知的好数据库（白名单）的散列值。

（5） 查找相对于已知产品分类数据库的散列值以识别对应于文件散列值的产品（和其它信息）。

（6） 发送用于病毒扫描的文件（例如通过作为例如 MIME 附件的 FTP 或者 SMTP），或者本地执行。

（7） 像在（4）中那样发送用于间谍件扫描的文件，或者本地执行。

（8） 像在（4）中那样发送用于与站点特别定制分析的文件，或者本地执行。

（9） 向在网络文件服务器上的特殊受限网络访问子目录（例如被认证的 samba 或者 FTPS）输出文件。

（10） 发送新文件需要分析的 SNMP 陷阱，并且指定它们的位置。

（11） 发送新的文件需要分析的 Syslog 或者电子邮件消息，并且指定它们的位置。

（12） 查看特定的目录以看是否另一个系统已经批准或者未批准文件。

（13） 对于服务器执行定制分析。

（14） 自动执行以第一分析的结果为条件的第二分析。

（15） 从外部分析系统接收包含分析结果的被认证的网络消息。

上述分析的结果在服务器上被汇总，服务器更新元信息存储库（124）中的状态，特别是要广播到主机的状态。服务器对于是否应当批准或者禁止文件作出建议。信息被汇总，以便管理员可以使用一个网页浏览器行为来批准或者禁止文件组。可选地，来自上述分析的结果可以用于自动使用特定的抗体来批准或者禁止文件。服务器可以提供报告、警告或者其它信息，并且可以改变所有或者一个或多个组的主机的参数 D 值。服务器标记状态改变以便以后通过更新（130）来分发，优选的是以主机从服务器拉（pull）更新的方式。

抗体分析/批准服务

因为系统聚焦在新的文件上，因此，可以使得外来的文件分析服务实用和有益。这些服务可以是自动化的（例如使用 SOAP/网页服务调用）或者人工的（沿着到服务提供商的服务器的已认证的链路）。可以使用远程服务器本地或者在装置外部执行的这些服务可以包括:

（1） 人工地输入散列值或者沿着预先计算的网页链路以获得已知的好和坏数据库查找项的查询结果。诸如公司的实体可能希望保存全局白名单或者全局黑名单。后者不对于散列值起作用，因为它们太多。前者不起作用，因为不同的公司具有用于限定“好”程序的不同策略。这些服务处理如下所述的白名单/黑名单/灰名单和表决。

（2） 查找与特定抗体相关联的抗体（例如与同一应用或者类似的应用相关联的文件组）。

（3） 识别与散列值相关联的提供商和应用。

（4） 查明多少公司和计算机具有那个文件并且多长时间具有那个文件。这些公司将不通过名称而被识别，仅仅被计数。服务提供商将作为这个服务的一部分秘密地收集这个信息。服务提供商建立结果和服务的双盲（double-blind）数据库。

（5） 查明多少公司已经禁止或者批准了文件，并且它们与所述文件一起批准了哪些文件。同样，从终端用户的视点来看，这些全部是盲的并且通过散列值来进行。服务提供商不必收集或者存储文件名或者文件数据，仅仅收集以抗体形式的元信息。事实上，文件名并且当然文件本身应当被考虑为专有信息。

（6） 基于上述查询的结果以及也基于服务器侧的分析的自动化的服务器侧批准。

内容提取器（CE）

内容通常形成内容的组或者包。其示例包括可执行程序和在.zip 文件中的病毒或者在微软 Office 文件（例如 Word、Excel 和 Powerpoint 文件）中的宏或者在安装包中的文件，诸如微软.msi 文件。参见图 8，接收文件，并且内容提取器查找嵌入的内容类型，诸如在 Office 文件中的宏。优选的是，仅仅提取这样的“有效”类型的内容。

在检测到可能的文件修改（600）或者未知状态后，提取器获取所提取的部分，并且将它们转换为有效的内容文件类型（诸如没有文本或者附图的 Word（.doc）文件），以重新打包它们。这个处理被图解为步骤 600-605。结果产生的重新打包的文件一般比原始文件（“容器”）小得多，并且被称为“缩小品”。所述缩小品的散列值被计算（603），并且将所述缩小散列值与容器散列值相关联（604）。容器可以被嵌套，并且具有关联的那些也被跟踪。以后，如果内容需要被上载，仅仅上载所述缩小品。作为选用，可以根据提取的分析的结果来上载容器文件及其元信息。可以根据提取的分析的结果来上载根容器和它们的元信息。例如，文件 setup.exe 包含文件 main.cab，其继而包含文件 install.exe。相对于 install.exe，main.cab 是 install.exe 的父容器，并且 setup.exe 是 install.exe 的根容器以及 main.cab 的父容器。所有这些关联被存储，优选地被存储为在独立文件的散列值之间的关系。

这个处理减少了分析阶段的网络流量和覆盖区（footprint），并且其允许仅仅跟踪嵌入内容，而不跟踪与其它文件（例如继承的文件模板）相关联的宏。这对于在宏装载时截取它们的方法不适用。提取器允许位置独立的嵌入宏检测和跟踪。

所述缩小品被重新打包为其它有效文件类型具有下述优点：所述缩小品与第三方分析系统兼容，例如被重新打包为小 Word 文件的宏可以作为电子邮件附件被发送到病毒扫描电子邮件网关。另一个示例是 zip 文件，temp.zip，其包含 5 个文件，仅仅其中之一是有效的，foo.exe。temp.zip 的缩小品可以是 zip 文件 foo.zip，其中仅仅有 foo.exe；或者，缩小品可以是 foo.exe 本身。Foo.zip 的签名或者 foo.exe 的签名最好被相关联为对应于 temp.zip 的签名。缩小品可以再次被以电子邮件发送到 AS 扫描电子邮件网关。一些容器缺少有效内容，因此可能不被跟踪。在跟踪缩小品上有效率高的优点，也有仅仅检测和分析新内容的优点。以这种方式，可以产生更多的精确的统计、警告和分析。诸如待决状态文件的未分类内容的自动和特定的早期检测允

许强大的策略和内容管理。

服务器用户接口

服务器的用户接口提供了多个“面板”，其中每个允许配置和管理系统的不同方面。在这个部分中，术语“用户”用于指示有权使用服务器用户接口的管理员。可以经由 SSL 加密的连接通过标准的网页浏览器来访问所述用户接口。认证和访问控制被提供来保持服务器的完整性，并且确定特定用户的特权级。

当用户第一次访问系统时，用户被认证并基于这个认证被分配特权级。这个特权级确定是否用户被允许无限制的访问或者只读的访问；也可以提供访问的更细的粒度。按照用户名和时间来跟踪和记录用户行为。在服务器上安装的证书可以用于控制和加密对于用户接口的访问，并且也提供被返回到主机的信息的签名和可能的加密。这些证书可以在维护面板中被安装和更新。对于接口的所有输入应当被正确地验证以保证服务器正向在它们的配置中的主机提供正确的信息。

网络状态接口提供运行的系统的概览，其中包括近来的事件和相关联的信息，所述相关联的信息包括唯一文件标识符、事件时间戳、事件类型、事件优先权、文件类型和名称以及通过名称和唯一标识符识别的主机系统。所述接口也提供关于在特定时段（例如最后一个小时、最后一天）期间的系统的状态的汇总信息。在统计面板中可以获得更详细的信息。在此显示的信息包括所进行的新的可执行部分的编号、所检测的新的脚本、具有新的嵌入内容的文件、未批准的文件和被感染的文件。

统计面板显示由系统收集的更详细的统计。这个信息包括在各种时段（例如，最后一个小时、最后 24 个小时、最后一个星期）中的下述事件的数量。它可以包括例如在网络上看到的新的可执行部分的数量、具有新的嵌入内容的文件、新的网页文件（HTML，ASP 等）、还没有人工或者通过扫描被批准的文件、通过扫描处理批准的文件、人工或者经由自动批准而批准的文件、扫描失败的文件、已知被感染并且已经被阻止的文件、被禁止并且阻止的可执行部分、自从事件被

第一次安装起由服务器处理的全部事件以及自从最后一次重启起的事件。

伴随每种类别的统计，用户可以观看一个项目的“前 10 个列表”，高亮在由服务器管理的所有主机上的每个的最常看到的实例。前 10 个列表的示例包括按照多少主机至少具有文件的一个拷贝的计数而排序的前 10 个近来发现的文件，这个列表的变体包括按照唯一的散列值的计数、按照唯一的文件名的计数、按照散列值被禁止的计数、按照名称被禁止的计数、近来被禁止的计数、近来更新/修改的计数、按照唯一组/容器/根容器/产品的计数。经由 SNMP 来更新和报告前 10 列表。配置面板可以用于根据前 10 计数和其它更新变量来配置警告和自动响应。警告包括记录报告、SNMP 陷阱、syslog 消息、电子邮件通知和其它网络消息。响应包括禁止文件、批准文件、改变一个或多个主机组的参数 D、改变一个或多个主机组的策略、改变一个或多个主机的主机组分配和分析文件。

统计面板也包括关于系统的整体信息，其中包括：被划分为有效和无效（无效主机是近来未接触服务器的主机）的、由这个服务器服务的主机的总数；在服务器数据库中的抗体的总数；以及正常运行时间，即系统自从最后一次重启已经正常运行了多长时间。

也可以经由向服务器的 SNMP（简单网络管理协议）查询来获得在这个面板上显示的统计信息，允许与网络管理系统交互。

绘图面板允许用户绘制和打印近来的行为的图形和图表。这个面板可以与统计面板组合。也可以以 XML 格式来获得绘制信息以在外部应用中显示。可以被绘制的图形的示例包括在近来的时段（按照分钟计的一个小时、按照小时计的一周等）上的行为或者“前 10 个列表”图形显示。

由于对于由服务器保留的统计信息的限制，对于可以获得的多个绘图有一些限制。当管理员正在使用 SNMP 管理系统时，其也能够在组织内提供已经在使用的格式的统计绘图。

抗体数据库面板允许用户直接地与在服务器上存储的抗体数据

库交互。数据库的内容被显示，并且用户可以选择通过不同的标准来分类所述显示，或者通过选择过滤模式而限制所述显示。用户也可以与抗体本身交互；这些操作被详细说明如下。

服务器可以使用辅助信息数据库，其包括在主抗体数据库中不要求的字段。在这个数据库中的字段的一个示例可以是第一个看到的文件名或者初始文件类。

对于每个文件，在这个面板上显示下面的信息：

- 第一次看到的时间。文件或者散列值被主机第一次看到并且报告到服务器的时间。
- 文件 ID。文件的唯一标识符，包括内容的一个或多个散列值，诸如 MD5、SHA-1 和 OMAC。
- 文件类型。网络类别（例如可执行、脚本、office 文件、档案文件等）。其是当其被第一次看到时从文件名得到（见下），并且也从文件内容的分析得到。
- 状态。包括批准、待决、禁止的当前文件状态。
- 方法。服务器得知文件（自动、人为等）的方式。
- 文件名。文件当第一次被看到并且被报告到服务器时的名称。其可能不是文件的当前名称，而仅仅是在网络上看到的第一示例的名称。
- 文件路径。文件当第一次被看到并且被报告到服务器时的路径。
- 第一次看到的主机。文件或者散列值被首次看到和报告所在的主机的名称。
- 分析结果。最新的扫描或者其它分析的结果。
- 第一次分析。文件的第一次扫描/分析的时间。
- 最后一次分析，最后扫描/分析文件的时间。
- 最后更新。最后修改文件状态的时间。
- 父容器。到已经与文件相关联的其它文件的链接。
- 父容器属性。一个相关联的容器文件的文件名、第一次被

看到的时间、第一次被看到的主机、文件路径、产品类别和状态。

● 根容器。到已经与文件相关联的其它文件的链接。根容器是未在另一个容器中被包含的容器。

● 根容器属性，一个相关联的根容器文件的文件名、第一次被看到的时间、第一次被看到的主机、文件路径、产品类别和状态。

可以对于从列表选择的文件执行下面的操作:

● 文件细节。这从抗体数据库提供了关于文件的附加信息，包括批准或者禁止所述文件的接口用户（其中，所述文件被首次看到）、用户加上的任何评述。

● 批准。明确地批准当前选择的文件。这个选择应向用户提供足够的警告，因为其将在所有主机上批准所述文件。

● 不批准。明确地不批准已经被批准的文件，最好过渡到待决的状态。

● 禁止。明确地禁止文件。这使得文件在所有的主机上被禁止。

● 分析/病毒扫描。强制所选择的文件的分析/扫描的计划。

● 删除。去除关于这个文件的信息。这将使得服务器下一次看到所述文件时将所述文件当作新的文件。

● 在主机上查找文件。这个操作链接到文件查找器，用于将所选择的文件名提供为输入。

● 查找容器。查找文件的可能容器和那些容器的信息。

● 查找根容器。查找文件的可能根容器和那些容器的信息。

● 查找网页服务信息。查询各种其它网络服务器以找到关于所述文件和/或其容器/产品的附加信息。

文件查找器面板允许用户启动在所有的被管理的主机上找到特定文件的位置的尽力处理。因为这个处理可能是耗时的，因此用户在启动新的搜索之前被通知。可能不在产品的所有版本中实现文件查找器。可以在查询的部分完成期间报告 **FindFile** 处理。

也以通过选择一个或多个特定文件来从抗体数据库面板启动所

述处理（参见第 0 部分），这然后将用户带到文件查找器面板，并且自动填充适当的信息。

这个处理需要与服务器通信的所有的主机异步地返回状态，因此，所述面板将打开新的视图以动态地显示被接收的结果。如果用户启动另一个搜索，则将终止当前的搜索。可以在未来的版本中允许多个文件搜索。

主机组面板允许服务器已知的主机与特定的逻辑组相关联。在接口的初始版本中不可获得全组功能，在这种情况下，这个屏幕将显示关于由这个服务器支持的单个组的信息。

所述面板支持操纵组成员资格，包括:

- 加上新的组。
- 去除现有的组，当去除一个组时，不从服务器的数据库去除主机，而是将其重新分配到默认组。
- 从一个组向另一个移动主机。

在这个面板上显示关于每个主机的下面的信息:

- 主机。主机的 DNS 名称。
- 唯一 ID。主机的唯一标识符。
- IP 地址。这个主机的最后一个已知的 IP 地址。
- 状态。主机的在线状态。
- 最后看到。主机向服务器登录的最后时间。
- 操作系统。主机的操作系统版本。
- 版本。在主机上的操作系统的版本。

文件类面板允许观看和编辑被映射到每个类的文件扩展。通过扩展来定义下面的一些类。通过内容的分析来确定其它类。通过扩展和分析来确定一些类。这些扩展是只读的。

一些预先定义的扩展类是:

- 可执行部分。包括 exe、com、dll、pif、scr、drv 和 ocx 的扩展。
- 脚本。包括 vbs、bat 和 cmd 的扩展。

- 嵌入的宏内容。包括 doc、dot、xls、xla、xlt、xlw、ppt、pps 和 pot 的扩展。
- 网页内容。包括 htm、html、asp 和 cgi 的扩展。

策略面板是服务器的配置的核心。用户可以显示和编辑通过主机组分组的、在所有的被管理的主机上实施的策略。这个面板也显示当前选择的组的当前全局 D 设置。

这个部分允许用户定义当前所选择的组的全局 D 级。当选择新的 D 级时，不立即应用所述改变，但是必须明确地选择所述改变。选择新的所提出的 D 级改变策略信息和行为的显示以示出这个新级的那些信息和行为。从面板导航离开将不应用所述改变。

策略列表显示在特定文件类（例如可执行部分、脚本等）上的特定 D 级的各种行为和效果。可以使能或者禁止策略，但是不可以编辑策略。在所述列表中包括下面的策略:

- 新的可执行部分
- 新的单独脚本
- 新的嵌入脚本
- 新的网页内容
- 未批准的网页内容
- 忽略更新代理（自动批准来自特定更新源/处理/位置的新的内容）
- 病毒/间谍件感染的文件

每当禁止策略时，那个类的文件的跟踪仍然继续，但是被影响的主机系统不采取行动。

对于每个策略，显示行为网格。所述网格指示在当前所选择的 D 级上应用哪些策略设置。

- 行为
- 阻止执行。这个文件类的执行要被阻止吗?
- 阻止写入。向这个文件类的文件的写入将被阻止吗?这个设置仅仅用于网页内容和未批准的文件。其仅仅用于严格地受控的系

统，而不用于正常的操作。

● 隔离。将隔离这类的文件吗?可以通过阻止读取，而不是移动到独立的目录来隔离文件。在病毒感染的文件的情况下，这些可以被写入，但是随后被删除，但是这个功能也可以不在初始被实现。

● 记录。将记录对于这个类的文件的访问吗?

● 批准

● 隐含的批准。文件在这个 D 级上被隐含地批准了吗?在适当的扫描和等待时间后，隐含的批准改变文件的被批准状态。

● 明确的批准。在这个 D 级上明确地批准文件吗?

类似于上述的那个的行为网格与预先制订的策略组合地向用户示出了特定 D 级的效果的表示。下面的表格示出了在各种 D 级（0-7）上的行为和预先进行的策略的组合的示例。

通知参数

当阻止内容访问时，主机用户被通知。对于在列表上的每个策略，并且对于每个主机组，可以获得下面的设置:

● 所显示的消息。在用户交互对话上显示的文本。在列表框中列出多个消息。

● 按钮文本。在用户交互对话上的单个按钮上显示的文本。

● 超时。对话将向用户显示多长时间。零超时表示一直到用户接受，并且对话保持无限期地被显示。

● 作为选用，对于 D 的特定值，用于覆盖内容限制一段时间的按钮。

● 具有关于策略的更多信息的 ULR 链接。

所述通知参数也包括全局设置，用于与通知消息一起限定在主机上显示的图像。这些设置对于每个预先制定的策略单独可配置。在服务器管理接口中编辑通知参数。那些参数与继而被分配到主机组的策略相关联，并且当策略改变时被传播到主机。

扫描年代参数

这个部分允许用户配置在当第一次看到和批准文件时（自动批准

扫描）、进行第二次（批准）扫描的时间和发生第三次（重复）扫描的时间之间的时间。可以如图 7 中那样指定更多的扫描和时间。

维护

维护部分允许用户配置服务器本身的全局设置。

- 系统配置。与服务器与本地网络和主机系统的交互相关联的设置。
- IP 地址和子网掩码。子网掩码允许将主机分类为远程和本地类型。远程主机具有更受限的通信，以保存带宽。主机组可以具有不同的策略集和 D 参数设置，其被对于每个连接类型远程、本地或者断开指定。远程主机将产生较小的网络流量，例如较少的服务器报告。并且远程主机最好向服务器报告新的内容的散列值，但是不上载所述内容。
- IP 取路由信息。
- 密码。设置或者重新设置用于访问服务器接口的密码。
- 证书。从可装卸的介质（并且作为选用，从网络）安装证书。这些被主机使用来验证服务器的身份，并且也用于到服务器的 SSL 接口。
- SNMP。设置 SNMP 服务器的列表以接收陷阱，并且被允许查询服务器的配置。
- SNMP 陷阱选择。选择哪种事件类型引起哪些陷阱，以及所述陷阱将被发送到哪种 SNMP 服务（并且与优先权相关的、高的、中间的、低的、信息的等...）
- Syslog。设置服务器的列表以对于各种事件类型和优先权经由 syslog 接收记录信息。
- NTP 时间同步服务器。设置用于时间同步的服务器列表。在服务器上的时间在启动时从其内部时钟被获得，然后与这个外部 NTP 时间源同步。服务器将跟踪与服务器时间的主机偏离。
- 系统状态（服务器）
- 正常运行时间。显示自从最后一次系统重启起的时间。

- 软件版本。显示服务器软件的版本信息。
- 盘空间。显示服务器的本地盘和存储统计。
- 病毒/间谍件签名更新
- 最后签名更新。最后签名更新的时间。
- 更新服务配置。配置被安装的反病毒软件的更新服务，包括下载位置和时间表。
- 更新扫描器。更新病毒扫描器软件。
- 更新签名。强制病毒签名的更新。
- 服务器软件更新
- 当前版本。显示当前服务器软件版本。
- 重启。使用当前安装的镜像来重启服务器。
- 安装新的镜像。从可装卸介质或者网络（例如经由 FTP）向服务器安装新的软件镜像。
- 返回到前一个版本，返回到先前使用的软件镜像。
- 外部服务配置
- 内容扫描系统的网络地址、服务类型和批准授权。
- 元信息共享服务的网络地址、服务类型和批准授权。
- 外部内容传送和用户定义的分析的外部文件服务器地址、协议、登录和目录。
- SNMP、syslog、电子邮件和新的内容的 SOAP 通知的外部内容通知服务配置。
- 备份。向可装卸介质（并且也向网络）备份和恢复所述配置。
- 存储配置和数据库。存储所述配置和抗体数据库（例如经由 XML）
- 安装配置和数据库。安装所述配置和抗体数据库（例如以 XML）。

服务器包括处理能力，诸如可编程的微处理器、数字信号处理器（DSP）或者应用相关的处理和存储器。主机可以包括个人计算机或者类似的计算机或者其它处理装置，其中包括手持装置、PDA 或者在网络上的其它装置。

已经在此描述了本发明的实施例，显然在不脱离所要求保护的本发明的范围的情况下，可以进行修改。

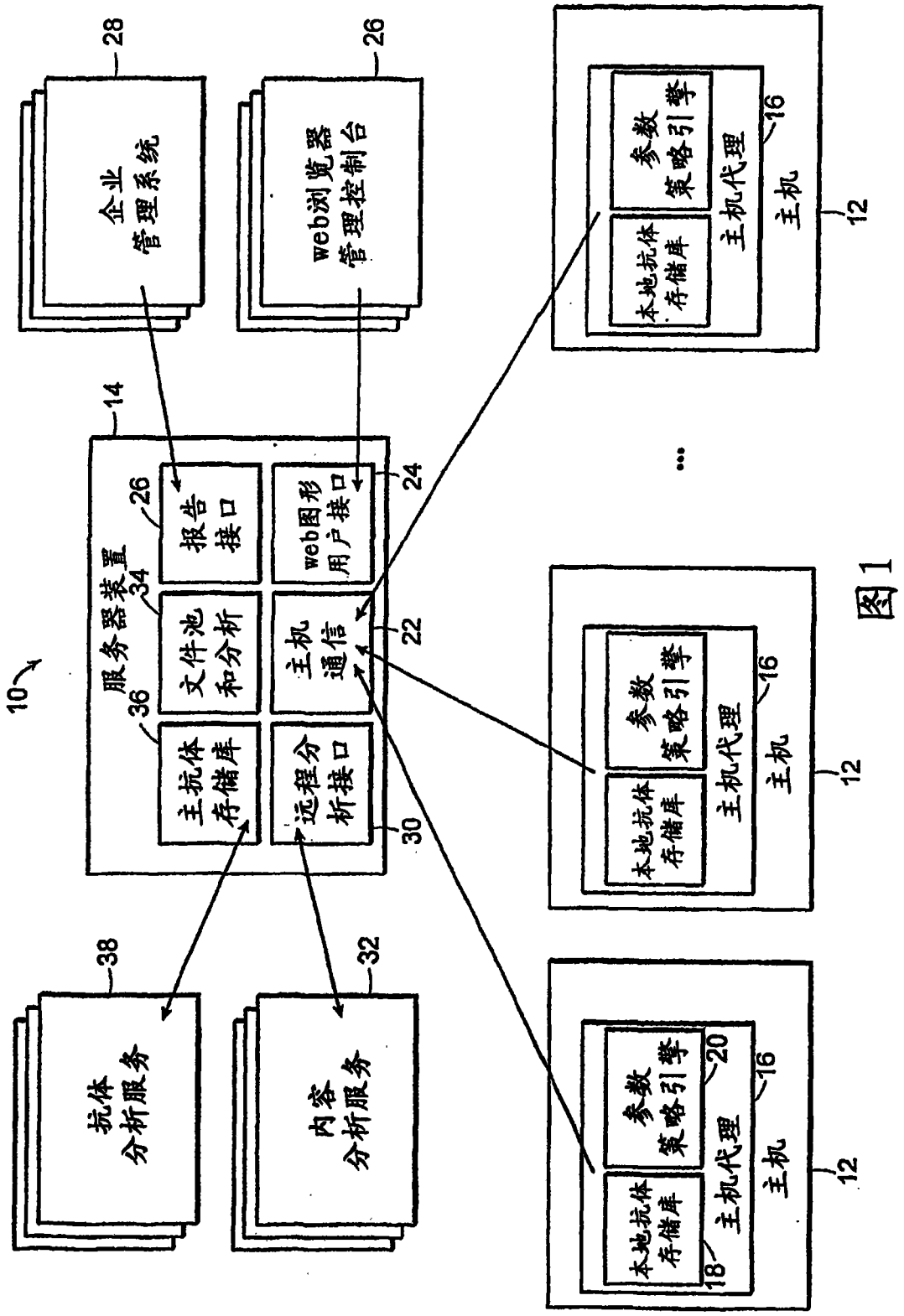
10
服务器装置
14
报告接口
26
web图形用户接口
24
文件池和分析
34
主机通信
22
主抗体存储库
36
远程分析接口
30
企业管理系统
28
web浏览器管理控制台
26
抗体分析服务
38
内容分析服务
32
主机
12
主机代理
16
参数策略引擎
20
本地抗体存储库
18

图 1

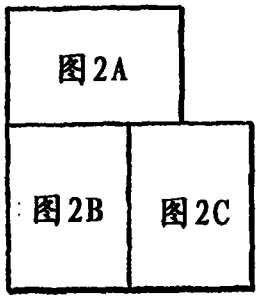

图2

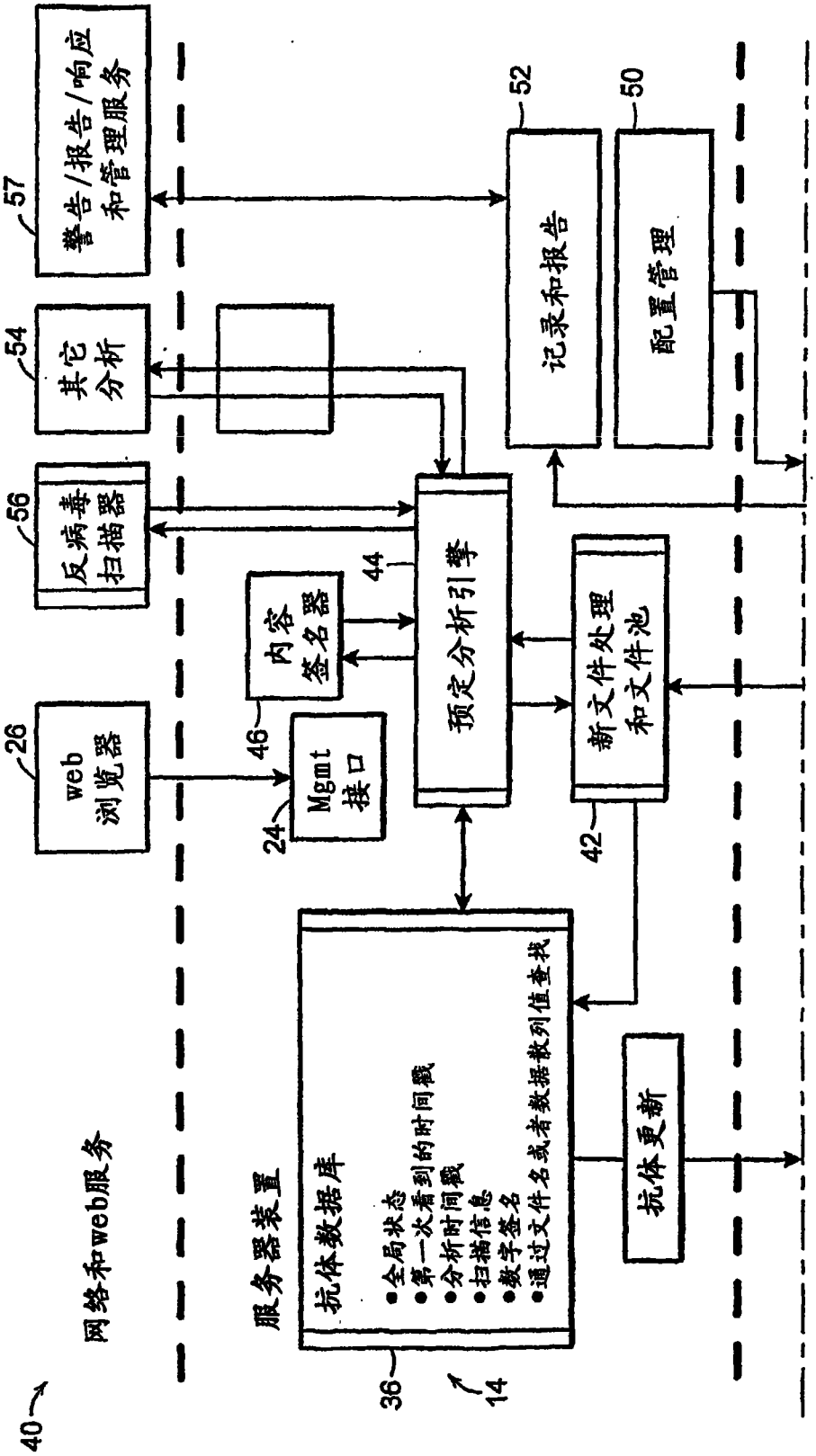
40
网络和web服务
服务器装置
14
抗体数据库
•全局状态
•第一次看到的时间戳
•分析时间戳
•扫描信息
•数字签名
•通过文件名或者数据散列值查找
36
抗体更新
42
新文件处理和文件池
预定分析引擎
44
内容签名器
46
Mgmt接口
24
web浏览器
26
反病毒扫描器
56
其它分析
54
警告/报告/响应和管理服务
57
记录和报告
52
配置管理
50

图2A

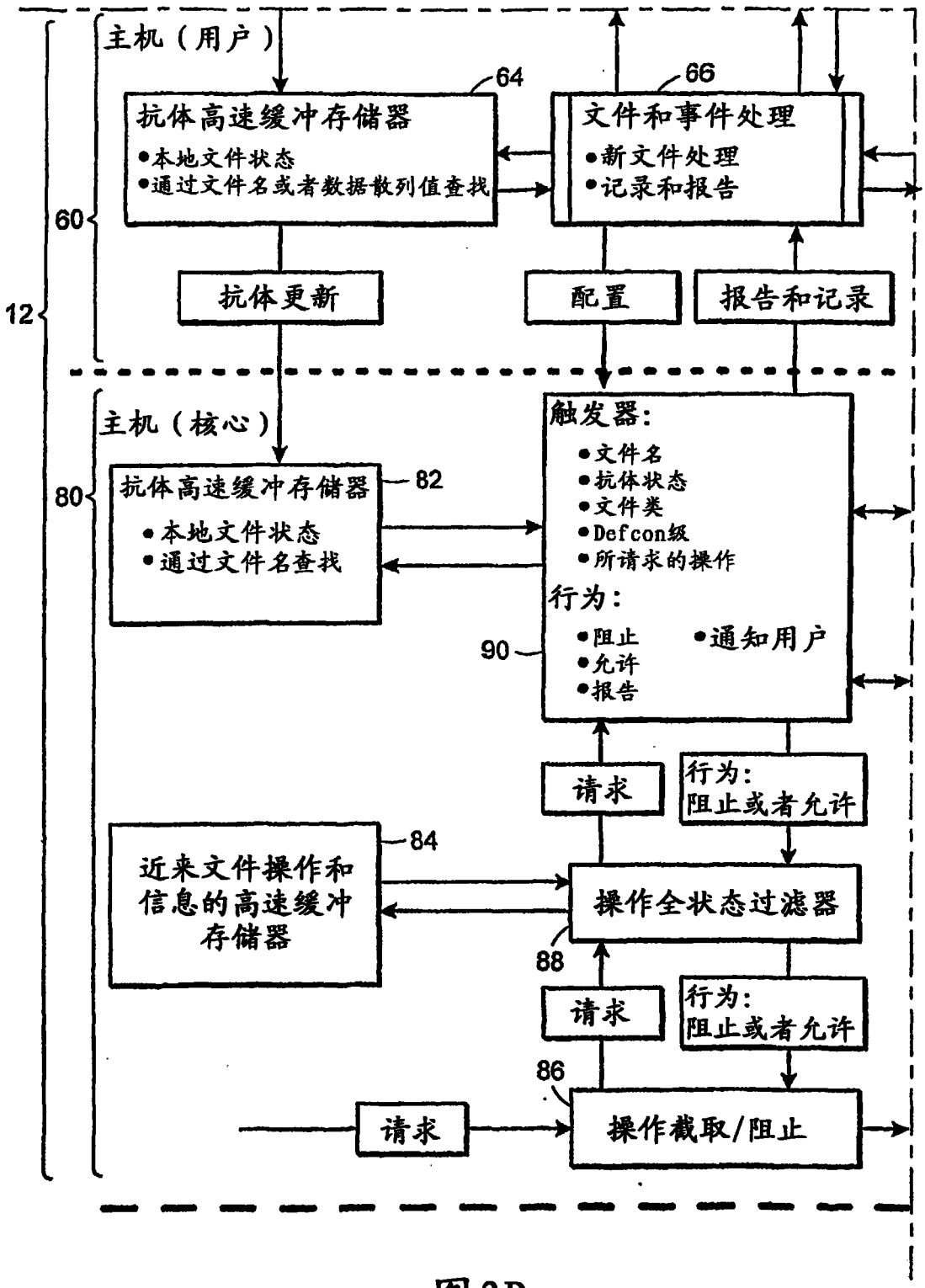
主机（用户）
64
抗体高速缓冲存储器
•本地文件状态
•通过文件名或者数据散列值查找
66
文件和事件处理
•新文件处理
•记录和报告
60
12
抗体更新
配置
报告和记录
主机（核心）
触发器：
•文件名
•抗体状态
•文件类
•Defcon级
•所请求的操作
行为：
•阻止
•允许
•报告
•通知用户
80
抗体高速缓冲存储器
•本地文件状态
•通过文件名查找
82
90
请求
行为：
阻止或者允许
84
近来文件操作和信息的高速缓冲存储器
操作全状态过滤器
88
请求
行为：
阻止或者允许
86
请求
操作截取/阻止

图2B

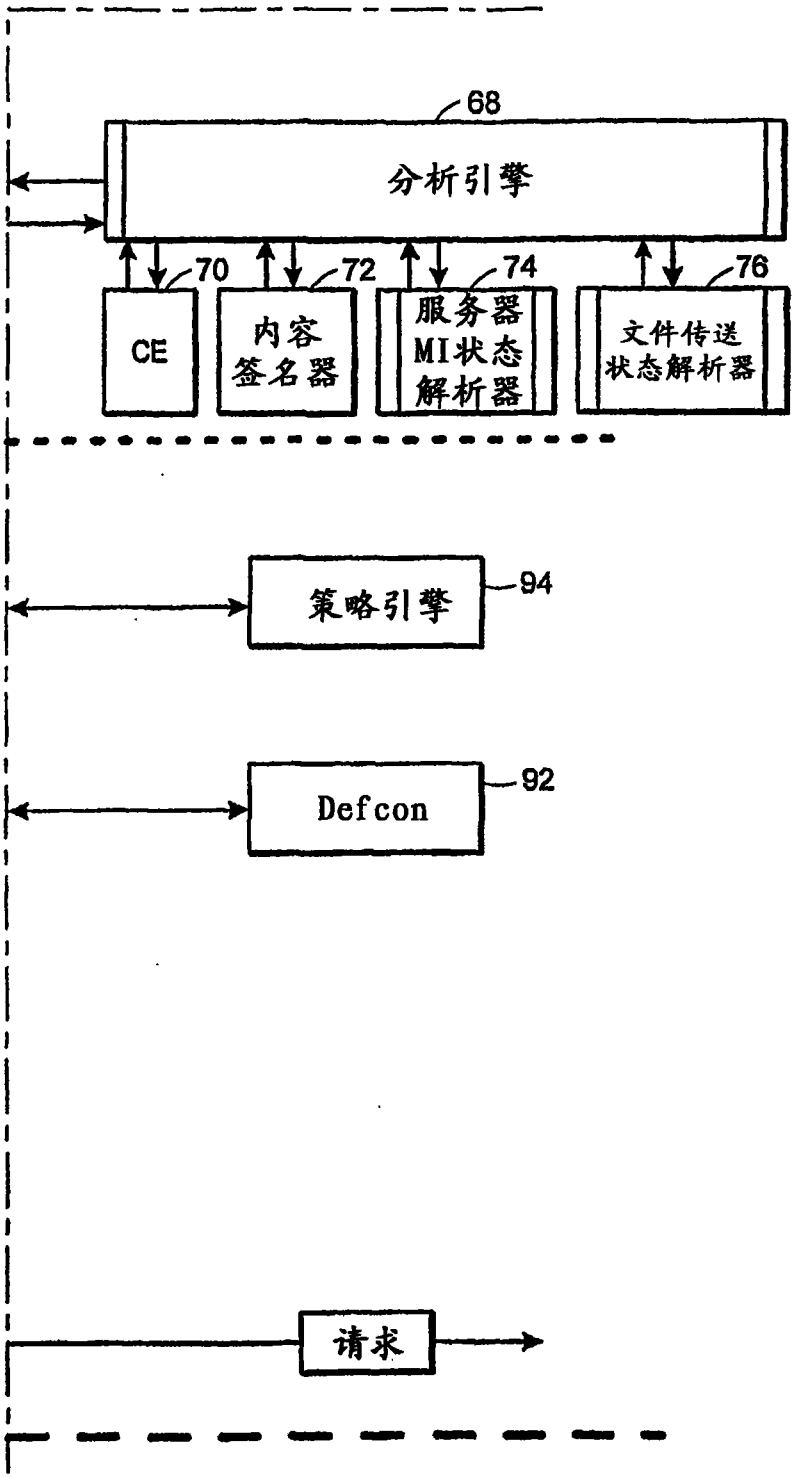
68
分析引擎
70
72
74
76
CE
内容
签名器
服务器
MI状态
解析器
文件传送
状态解析器
策略引擎
94
Defcon
92
请求

图2C

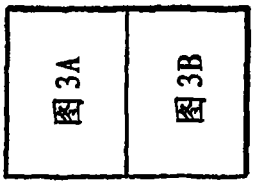
图3A
图3B

图3

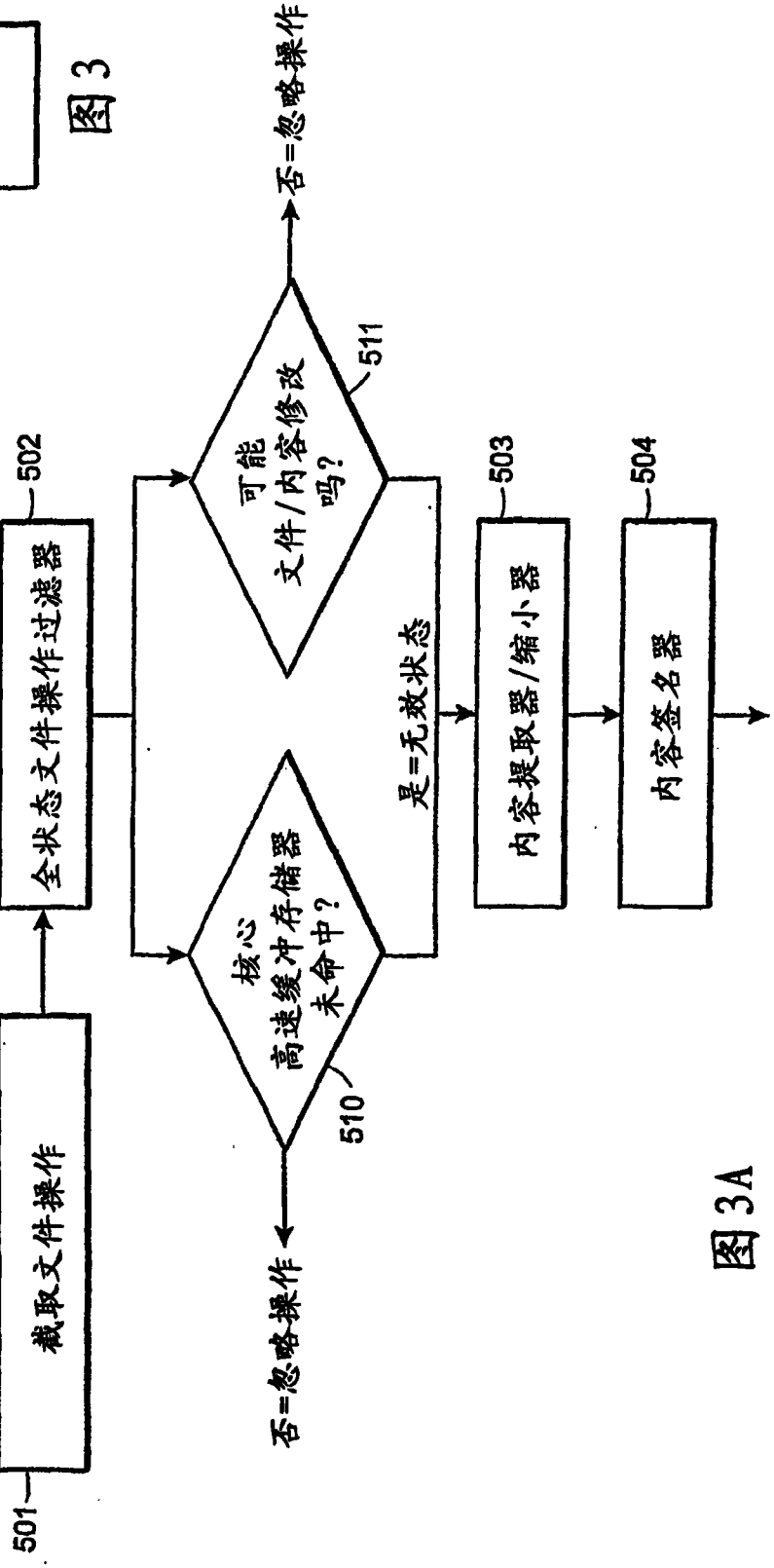
截取文件操作
501
全状态文件操作过滤器
502
核心
高速缓冲存储器
未命中?
510
否=忽略操作
可能
文件/内容修改
吗?
511
否=忽略操作
是=无效状态
内容提取器/缩小器
503
内容签名器
504

图3A

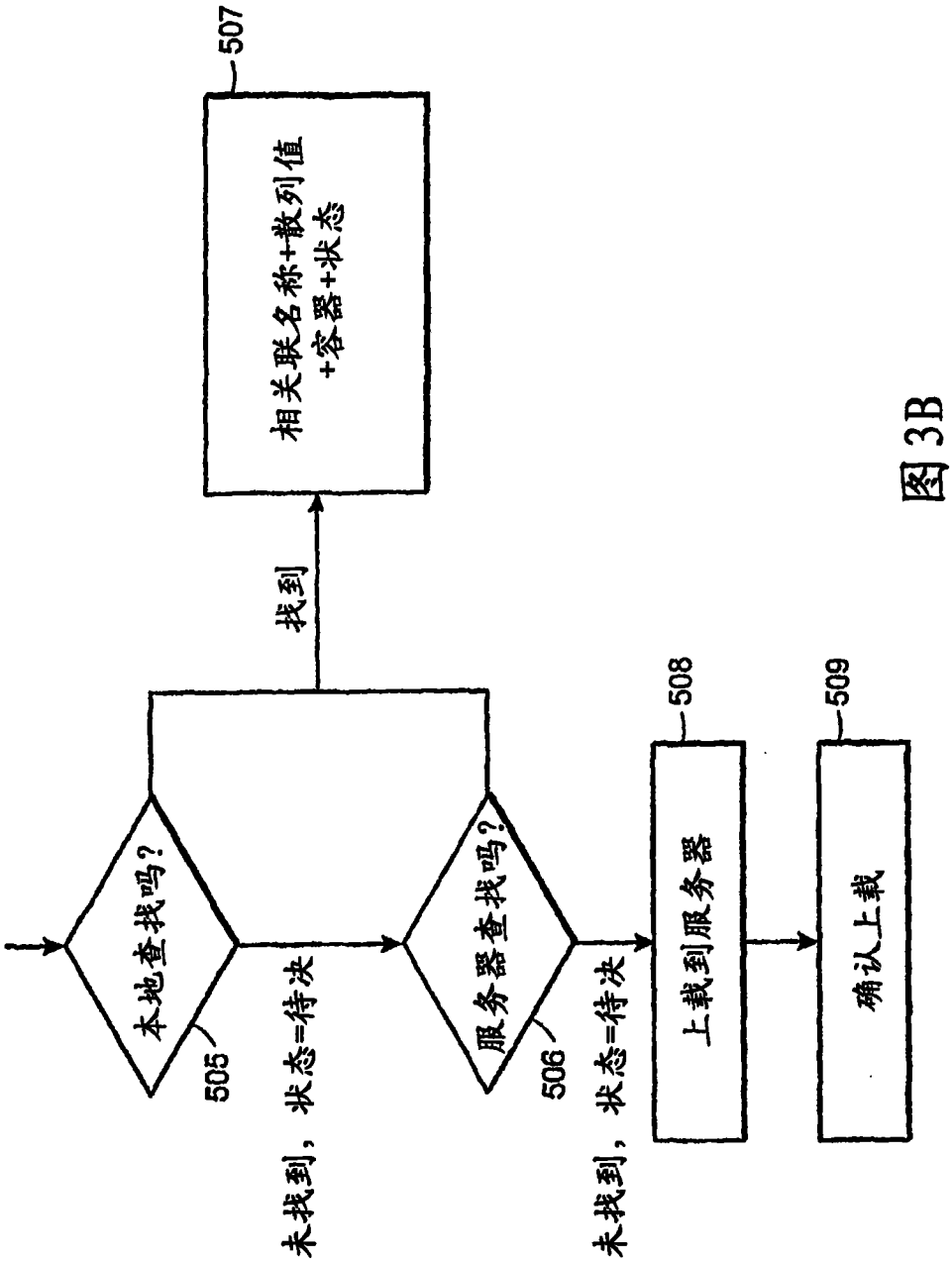
本地查找吗?
505
找到
507
相关联名称+散列值
+容器+状态
未找到，状态=待决
服务器查找吗?
506
未找到，状态=待决
508
上载到服务器
509
确认上载

图 3B

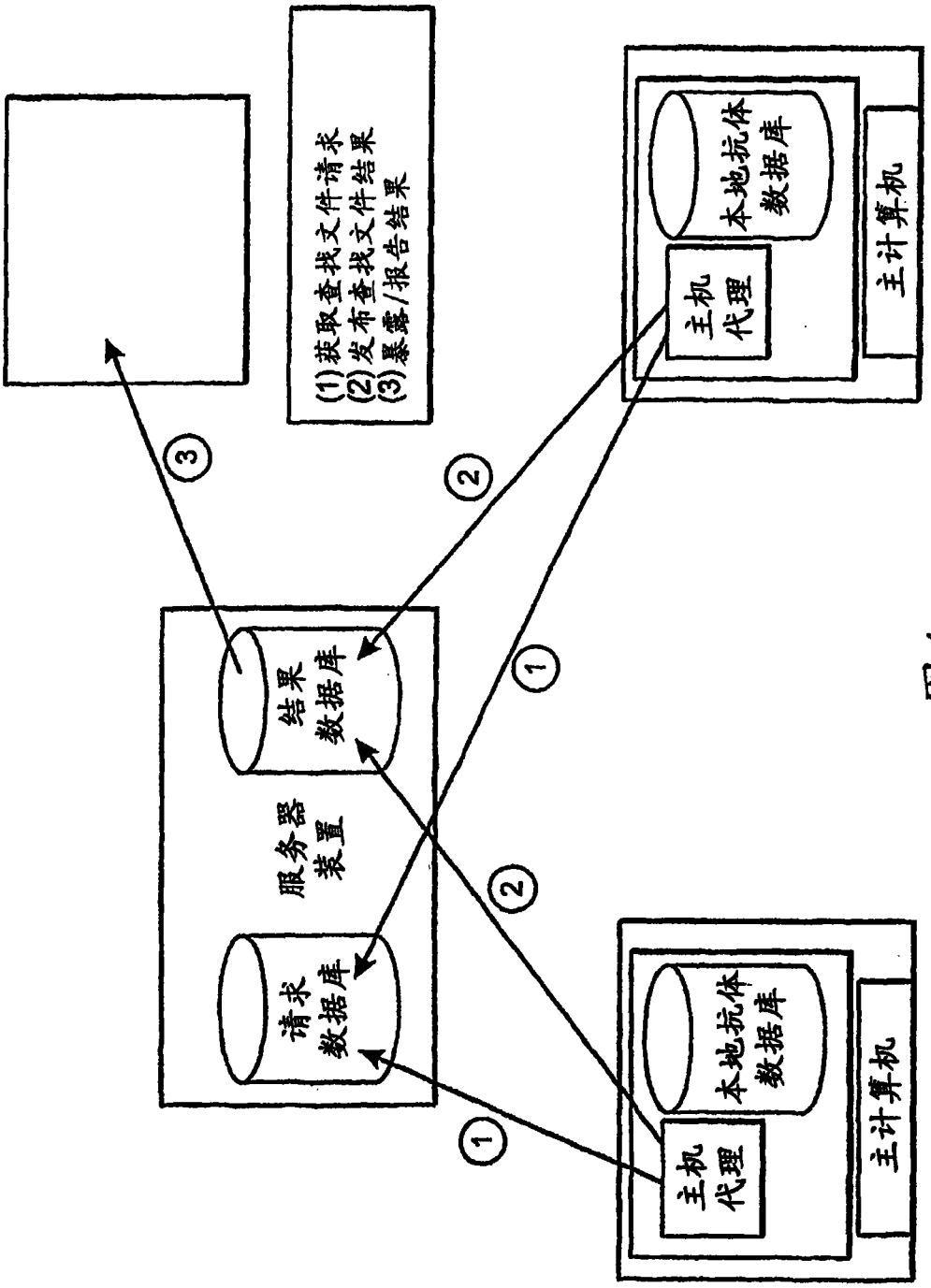
(1)获取查找文件请求
(2)发布查找文件结果
(3)暴露/报告结果
结果数据库
服务器装置
请求数据库
主机代理
本地抗体数据库
主计算机
主机代理
本地抗体数据库
主计算机
①
②
③

图 4

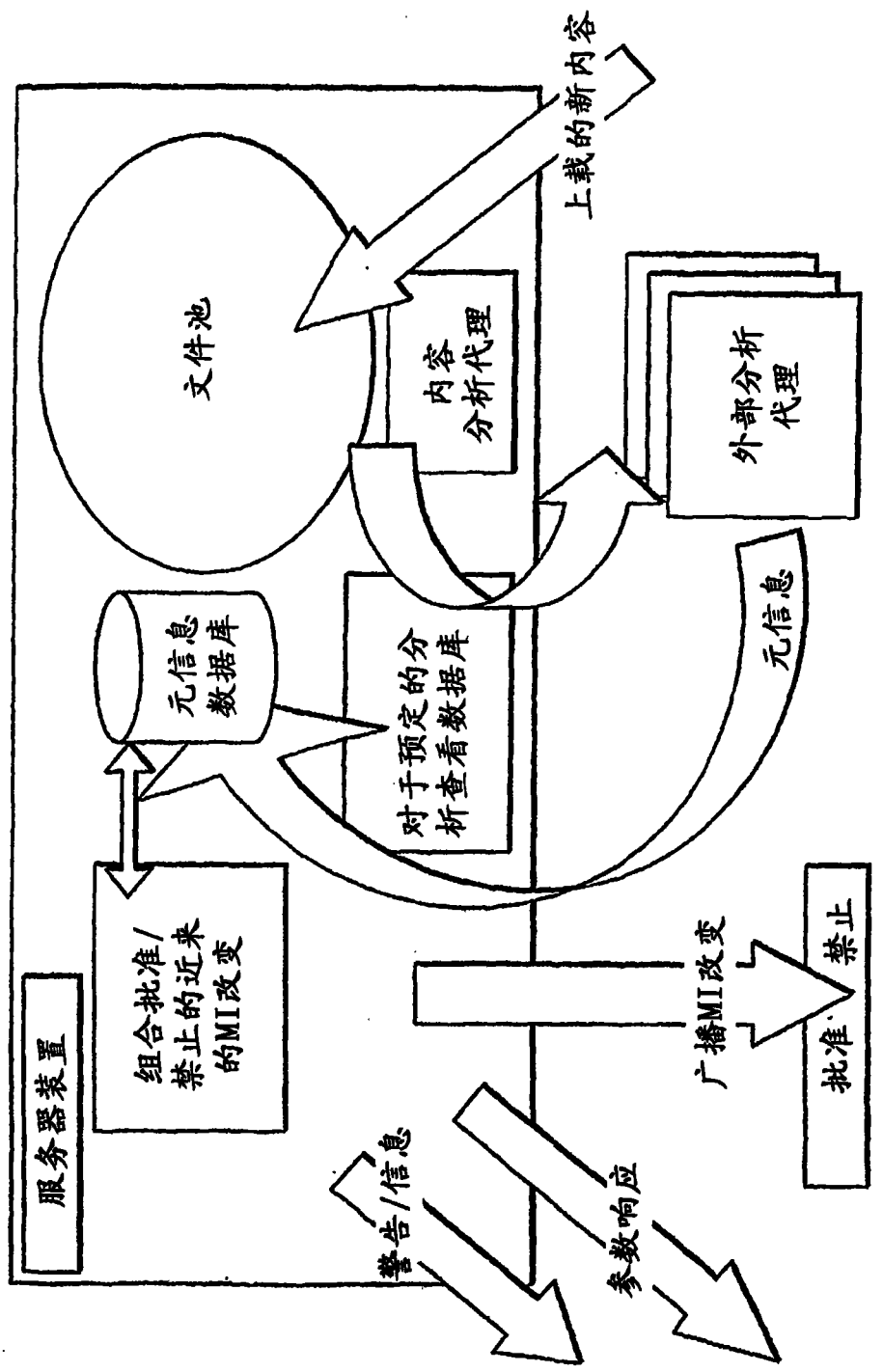
服务器装置
组合批准/禁止的近来的MI改变
元信息数据库
文件池
内容分析代理
对于预定的分析查看数据库
外部分析代理
上载的新内容
元信息
广播MI改变
批准
禁止
警告/信息
参数响应

图5

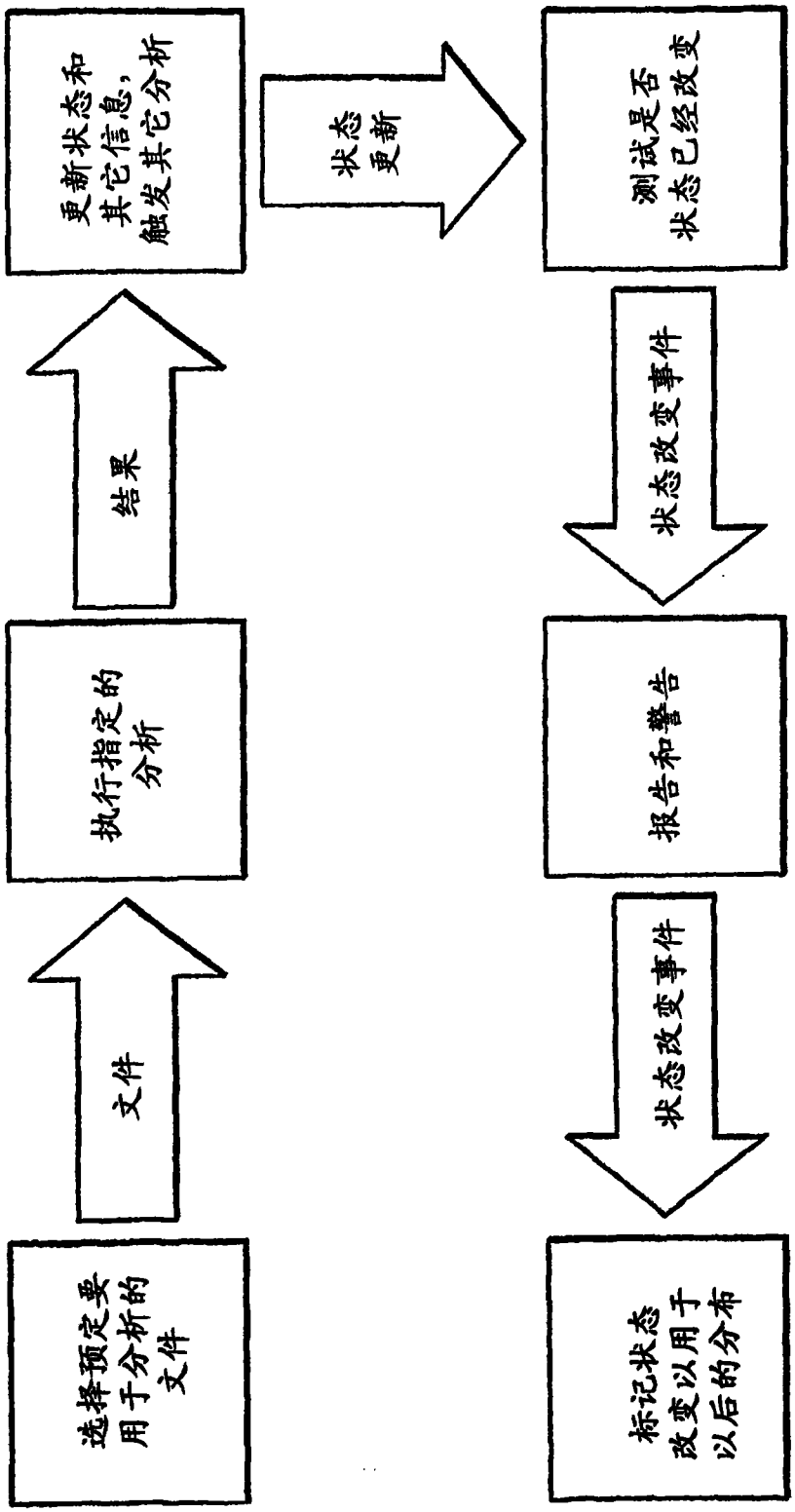
选择预定要用于分析的文件
文件
执行指定的分析
结果
更新状态和其它信息，触发其它分析
状态更新
测试是否状态已经改变
状态改变事件
报告和警告
状态改变事件
标记状态改变以用于以后的分布

图6

$\Delta t = t_{当前} - t_{在网络上第一次看到文件}$

Δt=0	Δt=12小时	Δt=2天	Δt=30天
散列验证			
AV扫描#1 AV扫描#2	AV扫描#1 AV扫描#2	AV扫描#1 AV扫描#2	AV扫描#1 AV扫描#2
AS扫描#1 AS扫描#2	AS扫描#1 AS扫描#2	AS扫描#1 AS扫描#2	AS扫描#1 AS扫描#2
其它分析#1 （分析新的）			
		其它分析#2 （分析永久的）	其它分析#2 （分析永久的）

图7

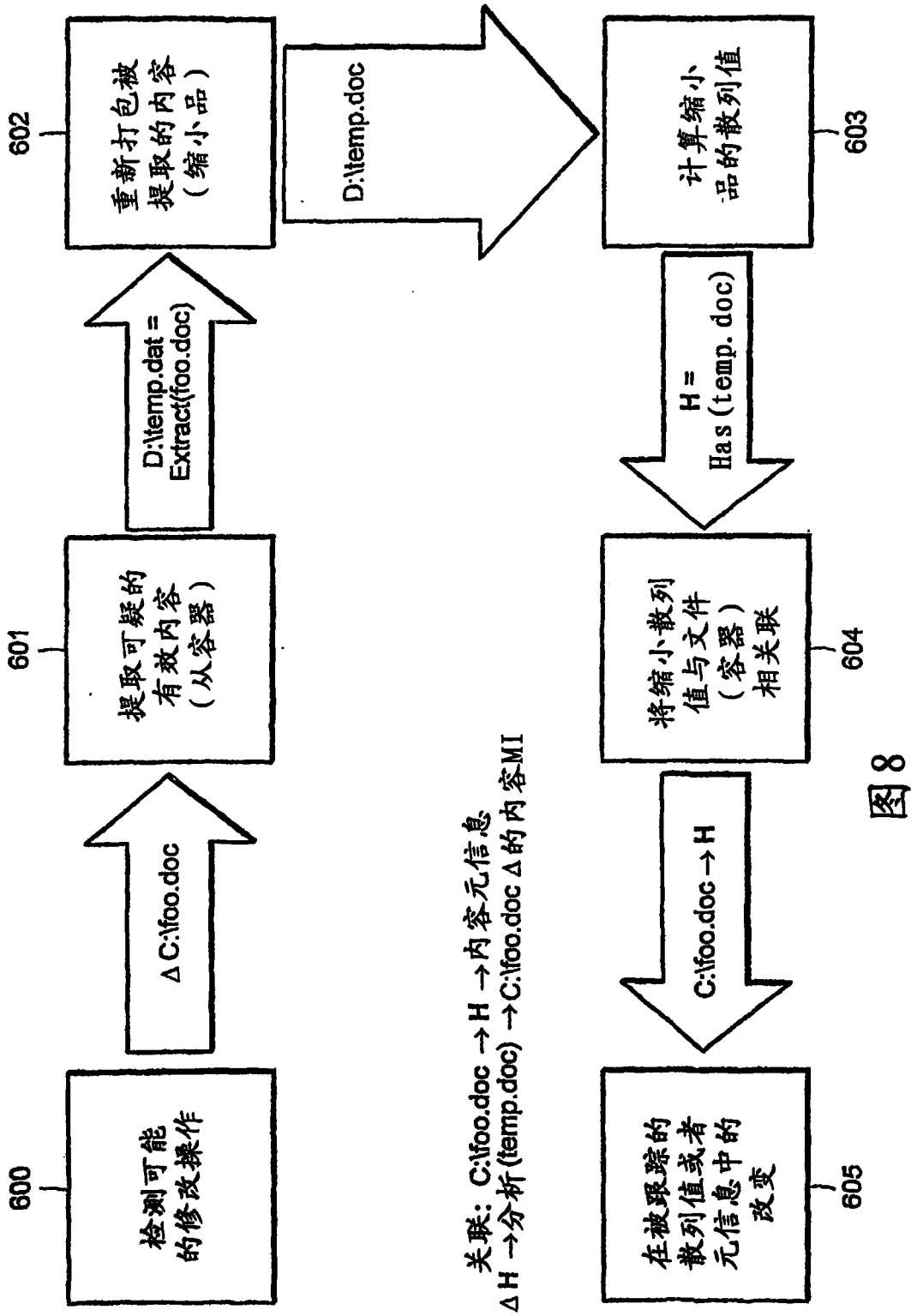
600
检测可能的修改操作
ΔC:\foo.doc
601
提取可疑的有效内容（从容器）
D:\temp.dat = Extract(foo.doc)
602
重新打包被提取的内容（缩小品）
D:\temp.doc
603
计算缩小品的散列值
H = Has(temp.doc)
604
将缩小散列值与文件（容器）相关联
C:\foo.doc → H
605
在被跟踪的散列值或者元信息中的改变
关联：C:\foo.doc → H → 内容元信息
ΔH → 分析(temp.doc) → C:\foo.doc Δ的内容MI

图8