(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 640 191**

(51) Int. CI.:

G06F 17/30 (2006.01)
G06F 11/07 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **11.08.2004 PCT/US2004/026186**

(87) Fecha y número de publicación internacional: **03.03.2005 WO05020001**

(96) Fecha de presentación y número de la solicitud europea: **11.08.2004 E 04786501 (9)**

(97) Fecha y número de publicación de la concesión europea: **14.06.2017 EP 1661047**

(54) Título: **Sistemas y métodos para soporte informático automatizado**

(30) Prioridad:

11.08.2003 US 494225 P
11.08.2004 US 916800

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
02.11.2017

(73) Titular/es:

NEHEMIAH SECURITY (100.0%)
8330 Boone Boulevard Suite 200
Tyson's, VA 22182, US

(72) Inventor/es:

HOOKS, DAVID, EUGENE

(74) Agente/Representante:

ELZABURU, S.L.P

DESCRIPCIÓN

Sistemas y métodos para soporte informático automatizado

CAMPO DE LA INVENCIÓN

La presente invención se refiere de manera general a sistemas y métodos para soporte informático automatizado.

ANTECEDENTES

A medida que la tecnología de la información continúa aumentando en complejidad, los costes de gestión de problemas escalarán a medida que se eleva la frecuencia de los incidentes de soporte y los requisitos del conjunto de habilidades de los analistas humanos llegan a ser más exigentes. Las herramientas convencionales de gestión de problemas están diseñadas para reducir los costes aumentando la eficiencia de los seres humanos que realizan estas tareas de soporte. Esto se logra típicamente automatizando al menos parcialmente la captura de información de tiques de problemas y facilitando el acceso a bases de conocimiento. Aunque es útil, este tipo de automatización ha alcanzado el punto de disminuir los retornos, en la medida que no responde a la debilidad fundamental en el modelo de soporte en sí mismo, su dependencia de los seres humanos.

La Tabla 1 ilustra la distribución de los costes de mano de obra asociados con la resolución de incidentes en el modelo de soporte convencional basado en seres humanos. Los datos mostrados se proporcionan por Motive Communications, Inc. de Austin, Texas (www.motive.com), un importante proveedor de software de servicio de soporte técnico. Las partidas de costes más altas son aquéllas asociadas con tareas que requieren análisis y/o interacción humana (por ejemplo, Diagnóstico, Investigación, Resolución).

Tabla 1

Tareas de Soporte	% Coste de Mano de Obra
Problemas Simples y Repetitivos (30%)	
Configuración de Escritorio (Infligidos por el usuario)	4%
Entorno de Escritorio (Malfuncionamiento del software)	9%
Interconexión y Conectividad	7%
Cómo (preguntas)	10%
Problemas Complejos y Dinámicos (70%)	
Clasificación (Identificar derechos del usuario y del soporte)	7%
Diagnóstico (Analizar estado de la máquina)	11%
Investigación (Encontrar la fuente del problema)	35%
Resolución y Reparación (Caminar con el usuario a través de la reparación)	18%

Las soluciones convencionales de software para la gestión automatizada de problemas se esfuerzan por disminuir estos costes y añadir valor a través de una amplia gama de niveles de servicio. Forrester Research, Inc. de Cambridge, MA (www.forrester.com) proporciona una caracterización útil de estos niveles de servicio. Forrester Research divide las soluciones convencionales de soporte informático automatizado en cinco niveles de servicio, incluyendo: (1) La Curación en Masa - resolver incidentes antes de que ocurran; (2) Autocuración - resolver incidentes cuando ocurren; (3) Autoservicio - resolver incidentes antes de que un usuario llame; (4) Servicio Asistido - resolver incidentes cuando un usuario llama; y (5) Visita del Lado del Soporte Técnico - resolver incidentes cuando todo lo demás falla. Según Forrester, el coste por incidente usando un servicio convencional de autocuración es menos de un dólar. Sin embargo, el coste escala rápidamente, alcanzando más de trescientos dólares por incidente si finalmente se requiere una visita del soporte técnico.

El objetivo de la Curación en Masa es resolver incidentes antes de que ocurran. En sistemas convencionales, este objetivo se logra haciendo que todas las configuraciones de PC sean iguales o, como mínimo, asegurándose de que un problema encontrado en un PC no pueda ser replicado en cualesquiera otros PC. Los productos convencionales asociados típicamente con este nivel de servicio consisten en herramientas de distribución de software y herramientas de gestión de configuración. Productos de seguridad tales como exploradores antivirus, sistemas de detección de intrusión y comprobadores de integridad de datos también se consideran parte de este nivel, dado que se centran en prevenir que ocurran incidentes.

Los productos convencionales que intentan abordar este nivel de servicio operan limitando la población gestionada a un pequeño número de configuraciones buenas conocidas y detectando y eliminando un número relativamente pequeño de configuraciones malas conocidas (por ejemplo, firmas de virus). El problema con este planteamiento es que supone que: (1) todas las configuraciones buenas y malas se pueden conocer antes de tiempo; y (2) una vez que se conocen que se mantengan relativamente estables. A medida que aumenta la complejidad de los sistemas informáticos y de interconexión, la estabilidad de cualquier nodo particular en la red tiende a disminuir. Es probable que cambie frecuentemente tanto el hardware como el software en cualquier nodo particular. Por ejemplo, muchos productos de software son capaces de actualizarse automáticamente a sí mismos usando parches de software

accedidos sobre una red interna o Internet. Dado que hay un número infinito de configuraciones buenas y malas y dado que cambian constantemente, estos productos convencionales de auto-curación nunca pueden ser más que parcialmente eficaces.

Además, los autores de virus continúan desarrollando virus cada vez más inteligentes. El software convencional de detección y erradicación de virus depende de la capacidad para identificar un patrón conocido para detectar y erradicar un virus. Sin embargo, a medida que aumenta el número y la complejidad de los virus, los recursos requeridos para mantener una base de datos de virus conocidos y correcciones para esos virus combinados con los recursos requeridos para distribuir las correcciones a la población de nodos en una red llegan a ser abrumadores. Además, un PC convencional que utiliza un sistema operativo Microsoft Windows incluye más de 7.000 archivos de sistema y más de 100.000 claves de registro, todas ellas de múltiples valores. Por consiguiente, a todos los efectos prácticos, puede existir un número infinito de estados buenos y un número infinito de estados malos, haciendo más complicada la tarea de identificación los estados malos.

El objetivo del nivel de Auto-Curación es detectar y corregir automáticamente los problemas antes de que den como resultado una llamada al servicio de soporte técnico, idealmente antes de que el usuario incluso sea consciente de que existe un problema. Las herramientas y las utilidades convencionales de Auto-Curación han existido desde finales de los 80 cuando Peter Norton introdujo un juego de herramientas de diagnóstico y reparación de PC (www.Symantec.com). Estas herramientas también incluyen herramientas que permiten a un usuario restaurar un PC a un punto de restauración establecido antes de la instalación de un nuevo producto. Sin embargo, ninguna de las herramientas convencionales funciona bien bajo condiciones del mundo real.

Un problema fundamental de estas herramientas convencionales es la dificultad en la creación de un modelo de referencia con suficiente alcance, granularidad y flexibilidad para permitir que “normal” sea distinguido con fiabilidad de “anormal”. Agravando el problema está el hecho de que la definición de “normal” debe cambiar constantemente a medida que se despliegan nuevas actualizaciones y aplicaciones de software. Este es un desafío técnico formidable y uno que aún tiene que ser conquistado por cualquiera de las herramientas convencionales.

El objetivo del nivel de Autoservicio es reducir el volumen de llamadas al servicio de soporte técnico proporcionando una colección de herramientas automatizadas y bases de conocimiento que permitan a los usuarios finales ayudarse a sí mismos. Los productos convencionales de Autoservicio consisten en bases de conocimiento de “cómo” y colecciones de soluciones de software que automatizan funciones de soporte repetitivo, de bajo riesgo, tales como restablecer contraseñas olvidadas. Estas soluciones convencionales tienen una desventaja significativa en que aumentan la probabilidad de daño auto-infligido. Por esta razón se limitan a tipos específicos de problemas y aplicaciones.

El objetivo del nivel de Servicio Asistido es mejorar la eficacia humana proporcionando una infraestructura automatizada para gestionar una solicitud de servicio y proporcionando capacidades para controlar remotamente un ordenador personal e interactuar con los usuarios finales. Los productos convencionales de Servicio Asistido incluyen software de servicio de soporte técnico, materiales de referencia en línea, y software de control remoto.

Aunque los productos en este nivel de servicio son quizás los más maduros de los productos y soluciones convencionales descritos en la presente memoria, aún no satisfacen plenamente los requisitos de usuarios y organizaciones. Específicamente, la capacidad de estos productos para diagnosticar automáticamente problemas está gravemente limitada tanto en términos de los tipos de problemas que se pueden identificar correctamente, así como de la precisión del diagnóstico (a menudo de elección múltiple).

Una Visita del Lado del Soporte Técnico llega a ser necesaria cuando todo lo demás falla. Este nivel de servicio incluye cualesquiera actividades de “acceso directo” que puedan ser necesarias para restaurar un ordenador que no pueda ser diagnosticado/reparado remotamente. También incluye el seguimiento y la gestión de estas actividades para asegurar una resolución oportuna. De todos los niveles de servicio, lo más probable es que este nivel requiera tiempo significativo de recursos humanos altamente capacitados y, por lo tanto, caros.

Los productos convencionales en este nivel consisten en herramientas de diagnóstico y productos de software especializados que rastrean y resuelven los problemas de los clientes con el tiempo y potencialmente a través de múltiples representantes de servicio al cliente.

De esta manera, lo que se necesita es un cambio de paradigma, que es necesario para reducir significativamente los costes de soporte. Este cambio se caracterizará por la aparición de un nuevo modelo de soporte en el que las máquinas servirán como agentes primarios para tomar decisiones e iniciar acciones

COMPENDIO DE LA TÉCNICA ANTERIOR

El documento US 2003/0028825 A1 (Hines) - 6 de febrero de 2003 se refiere a un método de corrección de errores y depuración y describe:

> “el sistema de gurú de servicio funciona automáticamente para procesar una imagen de un sistema informático para identificar, en su caso, qué condiciones previas del problema se satisfacen (es decir, el caso proactivo) y luego identifica problemas particulares de este conjunto más pequeño que coinciden con una descripción precisa del síntoma del problema (es decir, el caso reactivo)”

El método descrito se limita a corregir un único sistema informático.

El documento US 2002/0194550 A1 (LOPKE) - 19 de diciembre de 2002 se refiere a un sistema de diagnóstico de usuario final y describe:

> “Como se muestra en la FIG. 1, el inspector 40 está enlazado con el registro 26 del sistema. El inspector 40 obtiene datos de configuración del registro 26. Los datos de configuración se refieren a información requerida para configurar los componentes de software y hardware que definen el sistema 20 informático. Preferiblemente, el inspector 40 obtiene datos de configuración que reflejan información de configuración en tiempo real.”

El documento EP 1 172 732 A1 (HITACHI) - 16 de enero de 2002 se refiere a “[0014] un objeto de la presente invención es proporcionar un sistema informático en el que un ordenador puede adquirir información de fallo incluso en el caso donde ocurre en el ordenador un fallo que deshabilita un OS de la ejecución del procesamiento del fallo.”

El documento US 2003/0110248 A1 (RITCHE) - 12 de junio de 2003 se refiere a “un sistema informático y un método asociado, con una herramienta para procesar alertas de error emitidas durante la distribución de paquetes de aplicaciones a dispositivos cliente de red” [resumen]

COMPENDIO DE LA INVENCIÓN

Según la presente invención, un método de soporte informático automatizado comprende:

> i) recibir una instantánea desde una pluralidad de ordenadores:
> ii) comparar la instantánea con una base de datos de estados de ordenador; y,
> iii) identificar una anomalía, el método además que comprende:
> iv) recibir instantáneas desde una pluralidad de ordenadores dentro de una población de ordenadores, en donde las instantáneas individuales incluyen datos que indican un estado de un ordenador respectivo;
> v) almacenar la pluralidad de instantáneas en un almacén de datos;
> vi) crear automáticamente un modelo de referencia adaptativo basado al menos en parte en las instantáneas y que comprende un conjunto de reglas personalizado a las características de la población de ordenadores, el conjunto de reglas que se desarrolla identificando patrones entre las instantáneas de la pluralidad de ordenadores de manera que el modelo de referencia adaptativo sea indicativo de estados normales en los ordenadores dentro de la población;
> vii) comparar instantáneas de al menos uno de la pluralidad de ordenadores con el modelo de referencia adaptativo; y
> viii) determinar, el método además, si al menos una anomalía está presente en el estado del al menos uno de los ordenadores.

Realizaciones adicionales del método, un sistema organizado y dispuesto para efectuar el método y un medio legible por ordenador sobre el que está codificado el código de programa para efectuar el método de la presente invención, se exponen en las reivindicaciones adjuntas 2 a 15.

Las realizaciones de la presente invención proporcionan sistemas y métodos para soporte informático automatizado. Un método según una realización de la presente invención comprende recibir una pluralidad de instantáneas desde una pluralidad de ordenadores, almacenar la pluralidad de instantáneas en un almacén de datos, y crear un modelo de referencia adaptativo basado al menos en parte en la pluralidad de instantáneas. El método comprende además comparar al menos una de la pluralidad de instantáneas con el modelo de referencia adaptativo, e identificar al menos una anomalía basada en la comparación. En otra realización, un medio legible por ordenador (tal como, por ejemplo, una memoria de acceso aleatorio o un disco de ordenador) comprende un código para llevar a cabo tal método.

Estas realizaciones se mencionan no para limitar o definir la invención, sino para proporcionar ejemplos de realizaciones de la invención para ayudar a la comprensión de la misma. Realizaciones ilustrativas se discuten en la Descripción Detallada, y se proporciona allí una descripción adicional de la invención. Las ventajas ofrecidas por las diversas realizaciones de la presente invención se pueden comprender aún más examinando esta especificación.

BREVE DESCRIPCIÓN DE LAS FIGURAS

Estas y otras características, aspectos y ventajas de la presente invención se comprenden mejor cuando se lee la siguiente Descripción Detallada con referencia a los dibujos anexos, en los que:

> la Figura 1 ilustra un entorno ejemplar para la implementación de una realización de la presente invención;

la Figura 2 es un diagrama de bloques que ilustra un flujo de información y acciones en una realización de la presente invención;
la Figura 3 es un diagrama de flujo que ilustra un proceso global de detección de anomalías en una realización de la presente invención; y
la Figura 4 es un diagrama de bloques que ilustra componentes de un modelo de referencia adaptativo en una realización de la presente invención;
la Figura 5 es un diagrama de flujo que ilustra un proceso de normalización de información de registro en un agente en una realización de la presente invención;
la Figura 6 es un diagrama de flujo que ilustra un método para identificar y responder a una anomalía en una realización de la presente invención;
la Figura 7 es un diagrama de flujo que ilustra un proceso para identificar ciertos tipos de anomalías en una realización de la presente invención;
la Figura 8 es un diagrama de flujo que ilustra un proceso para generar un modelo de referencia adaptativo en una realización de la presente invención;
la Figura 9 es un diagrama de flujo, que ilustra un proceso para detección proactiva de anomalías en una realización de la presente invención;
la Figura 10 es un diagrama de flujo, que ilustra un proceso reactivo para detección de anomalías en una realización de la presente invención;
la Figura 11 es una captura de pantalla de una interfaz de usuario para crear un modelo de referencia adaptativo en una realización de la presente invención;
la Figura 12 es una captura de pantalla de una interfaz de usuario para gestionar un modelo de referencia adaptativo en una realización de la presente invención;
la Figura 13 es una captura de pantalla de una interfaz de usuario para seleccionar una instantánea a usar para la creación de un filtro de reconocimiento en una realización de la presente invención;
la Figura 14 es una captura de pantalla de una interfaz de usuario para gestionar un filtro de reconocimiento en una realización de la presente invención;
la Figura 15 es una captura de pantalla que ilustra una interfaz de usuario para seleccionar un “sistema excelente” para uso en una plantilla de políticas en una realización de la presente invención; y
la Figura 16 es una captura de pantalla de una interfaz de usuario para seleccionar activos de plantilla de políticas en una realización de la presente invención.

DESCRIPCIÓN DETALLADA

Las realizaciones de la presente invención proporcionan sistemas y un método para soporte informático automatizado. Con referencia ahora a los dibujos en los que números similares indican elementos similares a lo largo de las diversas figuras, la Figura 1 es un diagrama de bloques que ilustra un entorno ejemplar para la implementación de una realización de la presente invención. La realización mostrada incluye una facilidad 102 de soporte automatizado. Aunque la facilidad 102 de soporte automatizado se muestra como una facilidad única en la Figura 1, puede comprender múltiples facilidades o ser incorporada en el emplazamiento donde reside la población gestionada. La facilidad de soporte automatizado incluye un cortafuegos 104 en comunicación con una red 106 para proporcionar seguridad a los datos almacenados dentro de la facilidad 102 de soporte automatizado. La facilidad 102 de soporte automatizado incluye también un componente 108 Colector. El componente 108 Colector proporciona, entre otras características, un mecanismo para transferir datos dentro y fuera de la facilidad 102 de soporte automatizado. La rutina de transferencia puede usar un protocolo estándar tal como el protocolo de transferencia de archivos (FTP) o el protocolo de transferencia de hipertexto (HTTP) o puede usar un protocolo propietario. El componente Colector también proporciona la lógica de procesamiento necesaria para descargar, descomprimir y analizar sintácticamente las instantáneas entrantes.

La facilidad 102 de soporte automatizado mostrada también incluye un componente 110 Analítico en comunicación con el componente 108 Colector. El componente 110 Analítico incluye hardware y software para implementar el modelo de referencia adaptativo descrito en la presente memoria y almacenar el modelo de referencia adaptativo en un componente 112 de Base de Datos. El componente 110 Analítico extrae modelos de referencia adaptativos e instantáneas de un componente 112 de Base de Datos, analiza la instantánea en el contexto del modelo de referencia, identifica y filtra cualesquiera anomalías, y transmite el agente o los agentes de respuesta cuando sea adecuado. El componente 110 Analítico también proporciona la interfaz de usuario para el sistema.

La realización mostrada también incluye un componente 112 de Base de Datos en comunicación con el componente 108 Colector y el componente 110 Analítico. El componente 112 de Base de Datos proporciona unos medios para almacenar datos de los agentes y para los procesos realizados por una realización de la presente invención. Una función primaria del componente de Base de Datos puede ser almacenar instantáneas y modelos de referencia adaptativos. Incluye un conjunto de tablas de bases de datos, así como la lógica de procesamiento necesaria para gestionar automáticamente esas tablas. La realización mostrada incluye solamente un componente 112 de Base de Datos y un componente 110 Analítico. Otras realizaciones incluyen muchos componentes 112, 110 de Base de Datos y Analítico. Una realización incluye un componente de Base de Datos y múltiples componentes Analíticos, permitiendo que múltiple personal de soporte comparta una única base de datos mientras que realiza tareas analíticas en paralelo.

Una realización de la presente invención proporciona soporte automatizado a una población 114 gestionada que puede comprender una pluralidad de ordenadores 116a, b cliente. La población gestionada proporciona datos a la facilidad 102 de soporte automatizado a través de la red 106.

En la realización mostrada en la Figura 1, un componente 202 de Agente se despliega dentro de cada máquina 116a, b monitorizada. El componente 202 de Agente reúne datos del cliente 116. A intervalos programados (por ejemplo, una vez al día) o en respuesta a un comando del componente 110 Analítico, el componente 202 de Agente toma una instantánea detallada del estado de la máquina en la que reside. Esta instantánea incluye un examen detallado de todos los archivos del sistema, archivos de aplicaciones designadas, el registro, contadores de rendimiento, procesos, servicios, puertos de comunicación, configuración de hardware, y archivos de registro. Los resultados de cada exploración se comprimen entonces y se transmiten en forma de una Instantánea a un componente 108 Colector.

Cada uno de los servidores, ordenadores, y componentes de red mostrados en la Figura 1 comprenden procesadores y medios legibles por ordenador. Como es bien conocido por los expertos en la técnica, una realización de la presente invención se puede configurar de numerosas formas combinando múltiples funciones en un único ordenador o, alternativamente, utilizando múltiples ordenadores para realizar una única tarea.

Los procesadores utilizados por una realización de la presente invención pueden incluir, por ejemplo, procesadores lógicos digitales capaces de procesar la entrada, ejecutar algoritmos y generar la salida en la medida que sea necesario en soporte de procesos según la presente invención. Tales procesadores pueden incluir un microprocesador, un ASIC, y máquinas de estado. Tales procesadores incluyen, o pueden estar en comunicación con, medios, por ejemplo medios legibles por ordenador, que almacenan instrucciones que, cuando se ejecutan por el procesador, hacen que el procesador realice los pasos descritos en la presente memoria.

Realizaciones de medios legibles por ordenador incluyen, pero no se limitan a, un dispositivo de almacenamiento o de transmisión electrónico, óptico, magnético u otro capaz de proporcionar un procesador, tal como el procesador en comunicación con un dispositivo de entrada sensible al tacto, con instrucciones legibles por ordenador. Otros ejemplos de medios adecuados incluyen, pero no se limitan a, un disquete, CD-ROM, disco magnético, chip de memoria, ROM, RAM, un ASIC, un procesador configurado, todos los medios ópticos, todas las cintas magnéticas u otros medios magnéticos, o cualquier otro medio desde el cual un procesador informático pueda leer instrucciones. También, otras diversas formas de medios legibles por ordenador pueden transmitir o transportar instrucciones a un ordenador, incluyendo un encaminador, una red privada o pública, u otro dispositivo o canal de transmisión, tanto cableado como inalámbrico. Las instrucciones pueden comprender código de cualquier lenguaje de programación informático, incluyendo, por ejemplo, C, C#, C++, Visual Basic, Java y JavaScript.

La Figura 2 es un diagrama de bloques que ilustra un flujo de información y acciones en una realización de la presente invención. La realización mostrada comprende un componente 202 de Agente. El componente 202 de Agente es la parte del sistema que se despliega dentro de cada máquina monitorizada. Puede realizar tres funciones principales. En primer lugar, puede ser responsable de reunir datos. El componente 202 de Agente puede realizar una exploración extensiva de la máquina 116a, b cliente a intervalos programados, en respuesta a un comando del componente 110 Analítico, o en respuesta a eventos de interés detectados por el componente 202 de Agente. Esta exploración puede incluir un examen detallado de todos los archivos de sistema, archivos de aplicación designados, el registro, contadores de rendimiento, configuración de hardware, registros, tareas en ejecución, servicios, conexiones de red y otros datos relevantes. Los resultados de cada exploración se comprimen y se transmiten sobre la red 106 en forma de una “instantánea” al componente 108 Colector.

En una realización, el componente 202 de Agente lee cada byte de archivos a ser examinado y crea una firma digital o una comprobación aleatoria para cada archivo. La firma digital identifica el contenido exacto de cada archivo en lugar de simplemente proporcionar metadatos, tales como el tamaño y la fecha de creación. Algunos virus convencionales cambian la información de la cabecera del archivo en un intento de engañar a los sistemas que confían en metadatos para la detección. Tal realización es capaz de detectar con éxito tales virus.

La exploración del cliente por el componente 202 de Agente puede ser intensiva en recursos. En una realización, una exploración completa se realiza periódicamente, por ejemplo, diariamente, durante un tiempo cuando el usuario no está usando la máquina cliente. En otra realización, el componente 202 de Agente realiza una exploración delta de la máquina cliente, registrando solamente los cambios desde la última exploración. En otra realización, las exploraciones por el componente 202 de Agente se ejecutan bajo demanda, proporcionando una herramienta valiosa para un técnico o persona de soporte que intenta remediar una anomalía en la máquina cliente.

La segunda función principal realizada por el agente 202 es la de bloqueo de comportamiento. El agente 202 monitoriza constantemente (o sustancialmente constantemente) el acceso a los recursos del sistema de claves, tales como los archivos del sistema y el registro. Es capaz de bloquear selectivamente el acceso a estos recursos en tiempo real para evitar daños de software malicioso. Mientras que ocurre la monitorización del comportamiento sobre una base continua, el bloqueo de comportamiento se habilita como parte de una acción de reparación. Por ejemplo, si el componente 110 Analítico sospecha de la presencia de un virus, puede descargar una acción de reparación

para hacer que el cliente bloquee el acceso del virus a los recursos de información claves dentro del sistema gestionado. El componente 202 cliente proporciona información del proceso de monitorización como parte de la instantánea.

La tercera función principal realizada por el componente 202 de Agente es proporcionar un entorno de ejecución para agentes de respuesta. Los agentes de respuesta son componentes de software móviles que implementan procedimientos automatizados para abordar diversos tipos de condiciones problemáticas. Por ejemplo, si el componente 110 Analítico sospecha de la presencia de un virus, puede descargar un agente de respuesta para hacer que el componente 202 de Agente elimine los activos sospechosos del sistema gestionado. El componente 202 de Agente puede ejecutarse como un servicio u otro proceso en segundo plano en el ordenador que se monitoriza. Debido al alcance y la granularidad de la información proporcionada por una realización de la presente invención, la reparación se puede realizar con más precisión que con sistemas convencionales. Aunque se describe en términos de un cliente, la población 114 gestionada puede comprender PC, estaciones de trabajo, servidores o cualquier otro tipo de ordenador.

La realización mostrada también incluye un componente 206 de modelo de referencia adaptativo. Un reto técnico difícil en la construcción de un producto de soporte automatizado es la creación de un modelo de referencia que se pueda usar para distinguir entre estados normales y anormales del sistema. El estado del sistema de un ordenador moderno se determina por muchas variables de múltiples valores y, en consecuencia, hay virtualmente un número casi infinito de estados normales y anormales. Para empeorar las cosas, estas variables cambian con frecuencia a medida que se despliegan nuevas actualizaciones de software y a medida que se comunican los usuarios finales. El modelo 206 de referencia adaptativo en la realización mostrada analiza las instantáneas de muchos ordenadores e identifica estadísticamente patrones significativos usando un algoritmo genérico de extracción de datos o un algoritmo propietario de extracción de datos diseñado específicamente para este propósito. El conjunto de reglas resultante es extremadamente rico (cientos de miles de reglas) y se personaliza a las características únicas de la población gestionada. En la realización mostrada, el proceso de construcción de un nuevo modelo de referencia es completamente automático y se puede ejecutar periódicamente para permitir que el modelo se adapte a los cambios deseables tales como el despliegue planificado de una actualización de software.

Dado que el modelo 206 de referencia adaptativo se usa para el análisis de patrones estadísticamente significativos de una población de máquinas, en una realización, se analiza un número mínimo de máquinas para asegurar la precisión de las medidas estadísticas. En una realización, se prueba una población mínima de aproximadamente 50 máquinas para lograr patrones sistemáticamente relevantes para el análisis de las máquinas. Una vez se establece una referencia, se pueden usar muestras para determinar si está ocurriendo algo anormal dentro de la población entera o cualquier miembro de la población.

En otra realización, el componente 110 Analítico calcula un conjunto de métricas de madurez que permiten al usuario determinar cuándo se ha acumulado un número suficiente de muestras para proporcionar un análisis preciso. Estas métricas de madurez indican el porcentaje de relaciones disponibles en cada nivel del modelo que han cumplido criterios predefinidos correspondientes a diversos niveles de confianza (por ejemplo, Alto, Medio y Bajo). En una realización tal, el usuario monitoriza las métricas y asegura que se han asimilado suficientes instantáneas para crear un modelo maduro. En otra realización tal, el componente 110 Analítico asimila muestras hasta que alcanza una meta de madurez predefinida establecida por el usuario. En cualquiera de tales realizaciones, no es necesario asimilar un cierto número de muestras (por ejemplo, 50).

La realización mostrada en la Figura 2 también comprende un componente 208 de Plantilla de Políticas. El componente 208 de Plantilla de Políticas permite al proveedor de servicios insertar manualmente reglas en forma de “políticas” en el modelo de referencia adaptativo. Las políticas son combinaciones de atributos (archivos, claves de registro, etc.) y valores que cuando se aplican a un modelo, anulan una parte de la información generada estadísticamente en el modelo. Este mecanismo se puede usar para automatizar una variedad de actividades comunes de mantenimiento tales como la verificación de cumplimiento de las políticas de seguridad y la comprobación para asegurar que se han instalado las actualizaciones de software adecuadas.

Cuando algo va mal con un ordenador, a menudo impacta a una serie de activos de información diferentes (archivos, claves de registro, etc.). Por ejemplo, un “Troyano” podría instalar archivos maliciosos, añadir ciertas claves de registro para asegurar que se ejecutan esos archivos y abrir puertos para la comunicación. La realización mostrada en la Figura 2 detecta estos cambios indeseables como anomalías comparando la instantánea de la máquina infectada con la norma incorporada en el modelo de referencia adaptativo. Una anomalía se define como un activo presente de manera inesperada, un activo ausente de manera inesperada, o un activo que tiene un valor desconocido. Las anomalías se comparan frente a una biblioteca de Filtros 216 de Reconocimiento. Un Filtro 216 de Reconocimiento comprende un patrón particular de anomalías que indica la presencia de una condición de causa raíz particular o una clase genérica de condiciones. Los Filtros 216 de Reconocimiento también asocian las condiciones con una indicación de gravedad, una descripción textual y un enlace a un agente de respuesta. En otra realización, un Filtro 216 de Reconocimiento se puede usar para identificar e interpretar anomalías benignas. Por ejemplo, si un usuario añade una nueva aplicación que el administrador confía que no causará ningún problema, el sistema según la presente invención seguirá informando la nueva aplicación como un conjunto de anomalías. Si la

aplicación es nueva, entonces la notificación de los activos que añade como anomalías es correcta. Sin embargo, el administrador puede usar un Filtro 216 de Reconocimiento para interpretar las anomalías producidas añadiendo la aplicación como benigna.

En una realización de la presente invención, ciertos atributos se refieren a procesos continuos. Por ejemplo, los datos de rendimiento están compuestos por diversos contadores. Estos contadores miden la aparición de diversos eventos durante un período de tiempo particular. Para determinar si el valor de tal contador es normal a través de una población, una realización de la presente invención calcula una desviación media y una estándar. Se declara una anomalía si el valor del contador cae más de un cierto número de desviaciones estándar lejos de la media.

En otra realización, un mecanismo maneja el caso en el que el modelo 206 de referencia adaptativo asimila una instantánea que contiene una anomalía. Una vez que un modelo logra el nivel de madurez deseado se somete a un proceso que elimina las anomalías que se puedan haber asimilado. Estas anomalías son visibles en un modelo maduro como excepciones aisladas a relaciones fuertes. Por ejemplo, si el archivo A aparece en conjunto con el archivo B en 999 máquinas, pero en 1 máquina el archivo A está presente, pero el archivo B está ausente, el proceso asumirá que la relación posterior es anómala y se eliminará del modelo. Cuando el modelo se usa posteriormente para comprobar, cualquier máquina que contenga el archivo A, pero no el archivo B, se marcará como anómala.

La realización de la invención mostrada en la Figura 2 también incluye una biblioteca 212 de agentes de respuesta. La biblioteca 212 de agentes de respuesta permite al proveedor de servicios autorizar y almacenar respuestas automatizadas para condiciones problemáticas específicas. Estas respuestas automatizadas se construyen a partir de una colección de secuencias de comandos que se pueden despachar a una máquina gestionada para realizar acciones como sustituir un archivo o cambiar un valor de registro. Una vez que se ha analizado una condición problemática y se ha definido un agente de respuesta, se debería corregir automáticamente cualquier aparición posterior de la misma condición problemática.

La Figura 3 es un diagrama de flujo que ilustra un proceso global de detección de anomalías en una realización de la presente invención. En la realización mostrada, el componente (202) de Agente realiza una instantánea de una forma periódica, por ejemplo, una vez al día 302. Esta instantánea implica recoger una cantidad masiva de datos y puede tardar en cualquier sitio desde unos pocos minutos hasta horas en ejecutar, dependiendo de la configuración del cliente. Cuando la exploración está completa, los resultados se comprimen, formatean y transmiten en forma de una instantánea a un servidor seguro conocido como el componente 304 Colector. El componente Colector actúa como un repositorio central para todas las instantáneas que se envían desde la población gestionada. Cada instantánea entonces se descomprime, se analiza sintácticamente y se almacena en diversas tablas en la base de datos por el componente Colector.

La función (218) de detección usa los datos almacenados en el componente (206) del modelo de referencia adaptativo para comprobar los contenidos de la instantánea frente a cientos de miles de relaciones estadísticamente relevantes que se conoce que son normales para esa población 308 gestionada. Si no se encuentra 310 ninguna anomalía, el proceso termina 324.

Si se encuentra 310 una anomalía, se consultan los Filtros (210) de Reconocimiento para determinar si la anomalía coincide con cualesquiera condiciones 312 conocidas. Si la respuesta es sí, entonces la anomalía se notifica según la condición que se ha diagnosticado 314. De otro modo, la anomalía se notifica como una anomalía 316 no reconocida. El Filtro (216) de Reconocimiento también indica si se ha autorizado o no una respuesta automatizada para ese tipo particular de condición 318.

En una realización, los Filtros (216) de Reconocimiento pueden reconocer y consolidar múltiples anomalías. El proceso de coincidencia de Filtros de Reconocimiento con anomalías se realiza después de que se haya analizado la instantánea entera y se hayan detectado todas las anomalías asociadas con esa instantánea. Si se encuentra una coincidencia entre un subconjunto de anomalías y un Filtro de Reconocimiento, el nombre del Filtro de Reconocimiento se asociará con el subconjunto de anomalías en el flujo de salida. Por ejemplo, la presencia de un virus podría generar un conjunto de anomalías de archivo, anomalías de proceso y anomalías de registro. Un Filtro de Reconocimiento se podría usar para consolidar estas anomalías, de modo que el usuario simplemente vería un nombre descriptivo relacionando todas las anomalías con una causa común probable, es decir, un virus.

Si se ha autorizado la respuesta automatizada, entonces la biblioteca (212) de agentes de respuesta descarga los agentes de respuesta adecuados a la máquina 320 afectada. El componente 202 de Agente en la máquina afectada entonces ejecuta la secuencia de las secuencias de comandos necesaria para corregir la condición 322 problemática. El proceso mostrado entonces termina 324.

Las realizaciones de la presente invención reducen sustancialmente el coste de mantenimiento de una población de ordenadores personales y servidores. Una realización logra este objetivo detectando y corrigiendo automáticamente las condiciones problemáticas antes de que se escalen al servicio de soporte técnico y proporcionando información

de diagnóstico para acortar el tiempo requerido para que un analista de soporte resuelva cualesquiera problemas no abordados automáticamente.

Cualquier cosa que reduzca la frecuencia con la que ocurren incidentes tiene un impacto positivo significativo en el coste del soporte informático. Una realización de la presente invención monitoriza y ajusta el estado de una máquina gestionada de modo que sea más resistente a las amenazas. Usando Plantillas de Políticas, los proveedores de servicios pueden monitorizar rutinariamente la postura de seguridad de cada sistema gestionado, ajustando automáticamente los ajustes de seguridad e instalando las actualizaciones de software para eliminar vulnerabilidades conocidas.

En un modelo de soporte basado en seres humanos, las condiciones problemáticas se detectan por los usuarios finales, se notifican a un servicio de soporte técnico y se diagnostican por expertos humanos. Este proceso acumula costes de una serie de formas. En primer lugar, hay costes asociados con la pérdida de productividad mientras el usuario final espera la resolución. También, existe el coste de la recogida de datos, normalmente realizada por el personal del servicio de soporte técnico. Además, existe el coste del diagnóstico, que requiere los servicios de un analista de soporte formado (caro). Por el contrario, un modelo de soporte basado en máquina implementado según la presente invención detecta, notifica y diagnostica automáticamente muchas condiciones problemáticas relacionadas con el software. La tecnología del modelo de referencia adaptativo permite la detección de condiciones anómalas en presencia de extrema diversidad y cambian con una sensibilidad y precisión imposibles anteriormente.

En una realización de la presente invención, para evitar falsos positivos, el sistema se puede configurar para operar en diversos niveles de confianza, y las anomalías que se conocen que son benignas se pueden filtrar usando Filtros de Reconocimiento. Los Filtros de Reconocimiento también se pueden usar para alertar al proveedor de servicios sobre la presencia de tipos específicos de software no deseado o malicioso.

En sistemas convencionales, los incidentes informáticos se resuelven normalmente por seres humanos a través de la aplicación de una serie de acciones de reparación de ensayo y error. Estas acciones de reparación tienden a ser de la variedad “maza”, es decir, soluciones que afectan mucho más que las condiciones problemáticas que estaban destinadas a corregir. Los procedimientos de reparación de múltiples opciones y las soluciones de maza son una consecuencia de una comprensión inadecuada del problema y una fuente de costes innecesarios. Debido a que un sistema según la presente invención tiene los datos para caracterizar completamente el problema, puede reducir el coste de reparación de dos formas. En primer lugar, puede resolver automáticamente el incidente si se ha definido un Filtro de Reconocimiento que especifica la respuesta automatizada requerida. En segundo lugar, si la reparación automática no es posible, las capacidades de diagnóstico del sistema eliminan las conjeturas inherentes en el proceso de reparación basado en seres humanos, reduciendo el tiempo de ejecución y permitiendo una mayor precisión.

La Figura 4 es un diagrama de bloques que ilustra componentes de un modelo de referencia adaptativo en una realización de la presente invención. La Figura 4 es meramente ejemplar.

La realización mostrada en la Figura 4 ilustra un modelo 402 de referencia adaptativo de múltiples capa, de silo único. En la realización mostrada, el silo 404 comprende tres capas: la capa 406 de valor, la capa 408 de agrupación y la capa 410 de perfil.

La capa 406 de valor rastrea los valores de pares activo/valor proporcionados por el componente (202) de Agente descrito en la presente memoria a través de la población (114) gestionada de la Figura 1. Cuando se compara una instantánea con el modelo 402 de referencia adaptativo, la capa 406 de valor del modelo 402 de referencia adaptativo evalúa la parte de valor de cada par activo/valor contenido en la misma. Esta evaluación consiste en determinar si un valor de activo en la instantánea viola un patrón estadísticamente significativo de valores de activos dentro de la población gestionada como se representa por el modelo 402 de referencia adaptativo.

Por ejemplo, un Agente (116b) transfiere una instantánea que incluye una firma digital para un archivo de sistema particular. Durante el proceso de asimilación (cuando se está construyendo el modelo de referencia adaptativo), el modelo registra los valores que encuentra para cada nombre de activo y el número de veces que se encuentra ese valor. De esta manera, para cada nombre de activo, el modelo conoce los valores “legales” que ha visto en la población. Cuando se usa el modelo para comprobación, la capa 406 de valor determina si el valor de cada atributo en la instantánea coincide con uno de los valores “legales” en el modelo. Por ejemplo, en el caso de un archivo, es posible una serie de valores “legales” debido a que diversas versiones del archivo podrían existir en la población gestionada. Se declararía una anomalía si el modelo contuviese uno o más valores de archivo que eran estadísticamente coherentes y la instantánea contuviese un valor de archivo que no coincidía con ninguno de los valores de archivo en el modelo. El modelo también puede detectar situaciones donde no hay un valor “legal” para un atributo. Por ejemplo, los archivos de registro no tienen un valor legal dado que cambian frecuentemente. Si no existe ningún valor “legal”, entonces el valor del atributo en la instantánea se ignorará durante la comprobación.

En una realización, el modelo 402 de referencia adaptativo implementa criterios para asegurar que una anomalía es verdaderamente una anomalía y no sólo una nueva variante de archivo. Los criterios pueden incluir un nivel de

confianza. Los niveles de confianza no detienen que un archivo único sea notificado como una anomalía. Los niveles de confianza limitan las relaciones usadas en el modelo durante el proceso de comprobación a aquellas relaciones que cumplan ciertos criterios. Los criterios asociados con cada nivel están diseñados para lograr una cierta probabilidad estadística. Por ejemplo, en una realización, los criterios para el alto nivel de confianza están diseñados para alcanzar una probabilidad estadística mayor que el 90%. Si se especifica un nivel de confianza inferior, entonces se incluyen en el proceso de comprobación relaciones adicionales que no son tan fiables estadísticamente. El proceso de consideración de relaciones viables, pero menos probables, es similar al proceso humano de especulación cuando necesitamos tomar una decisión sin toda la información que nos permitiría estar seguros. En un entorno que cambia continuamente, el administrador puede desear filtrar las anomalías asociadas con niveles de confianza bajos, es decir, el administrador puede desear eliminar tantos falsos positivos como sea posible.

En una realización que implementa el nivel de confianza, si un usuario informa que algo está mal con una máquina, pero el administrador es incapaz de ver cualesquiera anomalías en el nivel de confianza por defecto, el administrador puede reducir el nivel de confianza, permitiendo que el proceso de análisis considere relaciones que tienen menor significación estadística y se ignoran en niveles de confianza más altos. Reduciendo el nivel de confianza, el administrador permite que el modelo 402 de referencia adaptativo incluya patrones que pueden no tener suficientes muestras para ser estadísticamente significativos, pero podrían proporcionar pistas en cuanto a cuál es el problema. En otras palabras, el administrador está permitiendo a la máquina especular.

En otra realización, la capa 406 de valor elimina automáticamente los valores de activos del modelo 402 de referencia adaptativo si, después de asimilar un número especificado de instantáneas, los valores de activos no han presentado ningún patrón estable. Por ejemplo, muchas aplicaciones generan archivos de registro. Los valores de los archivos de registro cambian constantemente y rara vez son los mismos de una máquina a otra. En una realización, estos valores de archivo se evalúan inicialmente y entonces, después de un número especificado de evaluaciones, se eliminan del modelo 402 de referencia adaptativo. Eliminando estos tipos de valores de archivo del modelo 402, el sistema elimina comparaciones innecesarias durante el proceso de detección 218 y reduce los requisitos de almacenamiento de la base de datos reduciendo la información de bajo valor.

Una realización de la presente invención no está limitada a eliminar valores de activos del modelo 402 de referencia adaptativo. En una realización, el proceso también se aplica a los nombres de activos. Ciertos nombres de activos son “únicos por naturaleza”, es decir, son únicos para una máquina particular, pero son un subproducto de la operación normal. En una realización, un proceso separado maneja nombres de activos inestables. Este proceso en tal realización identifica nombres de activos que son únicos por naturaleza y les permite permanecer en el modelo de modo que no se notifiquen como anomalías.

La segunda capa mostrada en la Figura 4 es la capa 408 de agrupación. La capa 408 de agrupación rastrea las relaciones entre los nombres de activos. Un nombre de activo puede aplicarse a una variedad de entidades, incluyendo un nombre de archivo, un nombre de clave de registro, un número de puerto, un nombre de proceso, un nombre de servicio, un nombre de contador de rendimiento o una característica de hardware. Cuando un conjunto particular de nombres de activos está presente de manera general en tándem en las máquinas en una población (114) gestionada, la capa 408 de agrupación es capaz de marcar una anomalía cuando está ausente un miembro del conjunto de nombres de activos.

Por ejemplo, muchas aplicaciones en un ordenador que ejecuta un sistema operativo Microsoft Windows requieren una multitud de bibliotecas de vínculos dinámicos (DLL). Cada DLL dependerá a menudo de una o más DLL. Si la primera DLL está presente, entonces las otras DLL también deben estar presentes. La capa 408 de agrupación rastrea esta dependencia y si una de las DLL falta o se altera, la capa 408 de agrupación alerta al administrador de que ha ocurrido una anomalía.

La tercera capa en el modelo 402 de referencia adaptativo mostrado en la Figura 4 es la capa 410 de perfil. La capa 410 de perfil en la realización mostrada detecta anomalías basadas en violaciones de relaciones de agrupación. Hay dos tipos de relaciones, asociativas (las agrupaciones aparecen juntas) y excluyentes (las agrupaciones nunca aparecen juntas). La capa 410 de perfil permite que el modelo de referencia adaptativo detecte activos que faltan no detectados por la capa de agrupación, así como conflictos entre activos. La capa 410 de perfil determina qué agrupaciones tienen fuertes relaciones asociativas y excluyentes entre sí. En tal realización, si no se detecta una agrupación particular en una instantánea donde normalmente se esperaría en virtud de la presencia de otras agrupaciones con las que tiene fuertes relaciones asociativas, entonces la capa 410 de perfil detecta la ausencia de esa agrupación como una anomalía. Del mismo modo, si se detecta una agrupación en una instantánea donde no se esperaría normalmente debido a la presencia de otras agrupaciones con las cuales tiene fuertes relaciones excluyentes, entonces la capa 410 de perfil detecta la presencia de la primera agrupación como anomalía. La capa 410 de perfil permite al modelo 402 de referencia adaptativo detectar anomalías que no serían detectables en niveles inferiores del silo 404.

El modelo 402 de referencia adaptativo mostrado en la Figura 4 se puede implementar de diversas formas que son bien conocidas por los expertos en la técnica. Optimizando el procesamiento del modelo 402 de referencia adaptativo y proporcionando recursos de procesamiento y almacenamiento suficientes, una realización de la

presente invención es capaz de soportar un número ilimitado de poblaciones gestionadas y clientes individuales. Tanto la asimilación de un nuevo modelo como el uso del modelo en la comprobación implican la comparación de cientos de miles de nombres y valores de atributo. Realizar estas comparaciones usando las cadenas de texto para los nombres y valores es una tarea de procesamiento muy exigente. En una realización de la presente invención, cada cadena única en una instantánea entrante se asigna a un identificador entero. Las comparaciones se realizan entonces usando los identificadores enteros más que las cadenas. Debido a que los ordenadores pueden comparar enteros mucho más rápido que las largas cadenas asociadas con nombres de archivo o nombres de clave de registro, la eficiencia de procesamiento se mejora extremadamente.

El modelo 402 de referencia adaptativo depende de datos del componente (202) de Agente. La funcionalidad del componente (202) de Agente se ha descrito anteriormente, que es un resumen funcional de la interfaz de usuario y el componente (202) de Agente en una realización de la presente invención.

Una realización de la presente invención es capaz de comparar entradas de registro a través de las máquinas cliente en una población gestionada. Una dificultad en la comparación de las claves de registro a través de diferentes máquinas que ejecutan un sistema operativo Microsoft Windows se deriva del uso de un Identificador Global Único ("GUID"). Un GUID para un elemento particular en una máquina puede diferir del GUID para el mismo elemento en una segunda máquina. Por consiguiente, una realización del presente sistema proporciona un mecanismo para normalizar los GUID con propósitos de comparación.

La Figura 5 es un diagrama de flujo que ilustra un proceso de normalización de información de registro de un cliente en una realización de la presente invención. En la realización mostrada, los GUID se agrupan primero en dos grupos 502. El primer grupo es para los GUID que no son únicos (duplicados) a través de las máquinas en la población gestionada. El segundo grupo incluye los GUID que son únicos a través de las máquinas, es decir, la misma clave tiene un GUID diferente en diferentes máquinas dentro de la población gestionada. Las claves para el segundo grupo se clasifican 504 a continuación. De esta forma, se puede identificar la relación entre dos o más claves dentro de la misma máquina. La intención es normalizar tales relaciones de una forma que las permitirá que sean comparadas a través de múltiples máquinas.

La realización mostrada a continuación crea una comprobación aleatoria para los valores en las claves 506. Esto crea una firma única para todos los nombres, nombres de caminos y otros valores contenidos en la clave. La comprobación aleatoria se sustituye entonces por el GUID 508. De esta manera, la singularidad se mantiene dentro de la máquina, pero la misma comprobación aleatoria aparece en cada máquina de modo que se pueda identificar la relación. La relación permite al modelo de referencia adaptativo identificar anomalías dentro de la población gestionada.

Por ejemplo, los virus convencionales a menudo cambian las claves de registro de modo que la máquina infectada ejecutará el ejecutable que propaga el virus. Una realización de la presente invención es capaz de identificar los cambios al registro en una o más máquinas de la población debido a su capacidad para normalizar las claves de registro.

La Figura 6 es un diagrama de flujo que ilustra un método para identificar y responder a una anomalía en una realización de la presente invención. En la realización mostrada, un procesador, tal como el componente (108) Colector, recibe una pluralidad de instantáneas desde una pluralidad de ordenadores 602. Aunque la siguiente discusión describe el proceso mostrado en la Figura 6 como que se realiza por el componente (110) Analítico, cualquier procesador adecuado puede realizar el proceso mostrado. La pluralidad de instantáneas puede comprender tan pocas como dos instantáneas de dos ordenadores. Alternativamente, la pluralidad de instantáneas puede comprender miles de instantáneas. Las instantáneas comprenden datos acerca de los ordenadores en una población para ser examinados. Por ejemplo, la pluralidad de instantáneas se puede recibir desde cada uno de los ordenadores en comunicación con una red de área local de la organización. Cada instantánea comprende una colección de pares de activo/valor que representan el estado de un ordenador en un punto particular en el tiempo.

A medida que el componente (108) Colector recibe las instantáneas, las almacena 604. Almacenar las instantáneas puede comprender almacenarlas en un almacén de datos, tal como en la base de datos (112) o en memoria (no mostrada). Las instantáneas se pueden almacenar temporal o permanentemente. También, en una realización de la presente invención, la instantánea entera se almacena en un almacén de datos. En otra realización, solamente se almacenan las partes de la instantánea que han cambiado desde una versión anterior (es decir, una instantánea delta).

El componente (110) Analítico utiliza los datos en la pluralidad de instantáneas para crear un modelo 606 de referencia adaptativo. Cada una de las instantáneas comprende una pluralidad de activos, que comprenden una pluralidad de pares de nombres de activos y valores de activos. Un activo es un atributo de un ordenador, tal como un nombre de archivo, un nombre de clave de registro, un parámetro de rendimiento, o un puerto de comunicación. Los activos reflejan un estado de un ordenador, real o virtual, dentro de la población de ordenadores analizados. Un valor de activo es el estado de un activo en un punto particular en el tiempo. Por ejemplo, para un archivo, el valor

puede comprender una comprobación aleatoria MD5 que representa los contenidos del archivo; para una clave de registro, el valor puede comprender una cadena de texto que representa los datos asignados a la clave.

El modelo de referencia adaptativo también comprende una pluralidad de activos. Los activos del modelo de referencia adaptativo se pueden comparar con los recursos de una instantánea para identificar anomalías y con otros propósitos. En una realización, el modelo de referencia adaptativo comprende una colección de datos acerca de diversas relaciones entre activos que caracterizan uno o más ordenadores normales en un punto particular en el tiempo.

En una realización, el componente (110) Analítico identifica una agrupación de nombres de activos. Una agrupación comprende uno o más grupos no solapados de nombres de activos que aparecen juntos. El componente (110) Analítico también puede intentar identificar relaciones entre las agrupaciones. Por ejemplo, el componente (110) Analítico puede calcular una matriz de probabilidades que predice, dada la existencia de una agrupación particular en una instantánea, la probabilidad de la existencia de cualquier otra agrupación en la instantánea. Las probabilidades que se basan en un número grande de instantáneas y que son o bien muy altas (por ejemplo, mayores que el 95%) o bien muy bajas (por ejemplo, menores que el 5%) se pueden usar por el modelo para detectar anomalías. Las probabilidades que se basan en un número pequeño de instantáneas, (es decir, un número que no es estadísticamente significativo) o que no son ni muy altas ni muy bajas no se usan para detectar anomalías.

El modelo de referencia adaptativo puede comprender un criterio de confianza para determinar cuándo se puede usar una relación para probar una instantánea. Por ejemplo, el criterio de confianza puede comprender un umbral mínimo para un número de instantáneas contenidas en el modelo de referencia adaptativo. Si no se excede el umbral, la relación no se usará. La referencia adaptativa puede comprender, también o en su lugar, un umbral mínimo para un número de instantáneas contenidas en el modelo de referencia adaptativo que incluyen la relación, utilizando la relación solamente si se excede el umbral. En una realización, el modelo de referencia adaptativo comprende un umbral máximo para una proporción del número de valores de activos diferentes al número de instantáneas que contienen los valores de activos. El modelo de referencia adaptativo puede comprender uno o más umbrales mínimo y máximo asociados con valores de activos numéricos.

Cada uno de la pluralidad de activos en el modelo de referencia adaptativo o en una instantánea se puede asociar con un tipo de activo. El tipo de activo puede comprender, por ejemplo, un archivo, una clave de registro, una medida de rendimiento, un servicio, un componente hardware, un proceso de ejecución, un registro y un puerto de comunicación. También se pueden utilizar otros tipos de activos por las realizaciones de la presente invención. Con el fin de conservar espacio, se pueden comprimir los nombres de activos y los valores de activos. Por ejemplo, en una realización de la presente invención, el componente (108) Colector identifica la primera aparición de un nombre de activo o valor de activo en una de la pluralidad de instantáneas y genera un identificador asociado con esa primera aparición. Posteriormente, si el componente (108) Colector identifica una segunda aparición del nombre de activo o valor de activo, el componente (108) Colector asocia el identificador con el segundo nombre de activo y valor de activo. El identificador y nombre de activo o valor de activo entonces se puede almacenar en un índice, mientras que solamente el identificador se almacena con los datos en el modelo de referencia adaptativo o instantánea. De esta forma, se minimiza el espacio necesario para almacenar nombres o valores de activos repetidos frecuentemente.

El modelo de referencia adaptativo se puede generar automáticamente. En una realización, el modelo de referencia adaptativo se genera automáticamente y luego se revisa manualmente para contar con el conocimiento de personal de soporte técnico u otros. La Figura 11 es una captura de pantalla de una interfaz de usuario para crear un modelo de referencia adaptativo en una realización de la presente invención. En la realización mostrada, un usuario selecciona las instantáneas a ser incluidas en el modelo moviéndolas desde la ventana de Menú 1102 de Selección de Máquina a la ventana 1104 de Máquinas en Tarea. Cuando el usuario completa el proceso de selección y pulsa el botón 1106 Terminar se crea una tarea automática que hace que el modelo sea generado. Una vez se ha creado el modelo, el usuario puede usar otra pantalla de interfaz para gestionarlo. La Figura 12 es una captura de pantalla de una interfaz de usuario para gestionar un modelo de referencia adaptativo en una realización de la presente invención.

Con referencia de nuevo a la Figura 6, una vez que se ha creado el modelo de referencia adaptativo, el componente (110) Analítico compara al menos una de la pluralidad de instantáneas con el modelo 608 de referencia adaptativo. Por ejemplo, el componente (108) Colector puede recibir y almacenar en el componente (112) de Base de Datos cien instantáneas. El componente (110) Analítico usa las cien instantáneas para crear un modelo de referencia adaptativo. El componente (110) Analítico entonces comienza comparando cada una de las instantáneas en la pluralidad de instantáneas con el modelo de referencia adaptativo. En algún momento más tarde el componente (108) Colector puede recibir 100 nuevas instantáneas desde los componentes de Agente, que entonces se pueden usar por el componente Analítico para generar una versión revisada del modelo de referencia adaptativo y para identificar anomalías.

En una realización de la presente invención, la comparación de una o más instantáneas con un modelo de referencia adaptativo comprende examinar relaciones entre nombres de activos. Por ejemplo, la probabilidad de existencia de un primer nombre de activo puede ser alta cuando está presente un segundo nombre de activo. En una realización, la comparación comprende determinar si todos los nombres de activos en una instantánea existen dentro del modelo de referencia adaptativo y son coherentes con una pluralidad de relaciones de alta probabilidad entre nombres de activos.

Con referencia aún a la Figura 6, en una realización, el componente (110) Analítico compara la instantánea con el modelo de referencia adaptativo con el fin de identificar cualesquiera anomalías que puedan estar presentes en un ordenador 610. Una anomalía es una indicación de que alguna parte de una instantánea se desvía de lo normal como se define por el modelo de referencia adaptativo. Por ejemplo, un nombre o valor de activo puede desviarse del nombre de activo normal y valor de activo esperado en una situación particular como se define por un modelo de referencia adaptativo. La anomalía puede señalar o no que un problema o condición problemática conocida o nueva existe sobre o en relación con el ordenador con el cual está asociada la instantánea. Una condición es un grupo de anomalías que están relacionadas. Por ejemplo, un grupo de anomalías se puede relacionar porque surgen de una causa raíz única. Por ejemplo, una anomalía puede indicar la presencia de una aplicación particular sobre un ordenador cuando la aplicación no está presente de manera general en los otros ordenadores dentro de una población dada. El reconocimiento de anomalías también se puede usar para funciones tales como balanceo de capacidad. Por ejemplo, evaluando las medidas de rendimiento de varios servidores, el componente (110) Analítico es capaz de determinar cuándo desencadenar el despliegue y la configuración automáticos de un nuevo servidor para abordar las demandas cambiantes.

Una condición comprende un grupo de anomalías relacionadas. Por ejemplo, un grupo de anomalías pueden estar relacionadas debido a que surgen de una causa raíz única, como la instalación de un programa de aplicaciones o la presencia de un “gusano”. Una condición puede comprender una clase de condición. La clase de condición permite que diversas condiciones sean agrupadas unas con otras.

En la realización mostrada en la Figura 6, si se encuentra una anomalía, el componente (110) Analítico intenta hacer coincidir la anomalía con un filtro de reconocimiento con el fin de diagnosticar una condición 612. La anomalía se puede identificar como una anomalía benigna con el fin de eliminar el ruido durante el análisis, es decir, con el fin de evitar oscurecer condiciones problemáticas reales debido a la presencia de anomalías que son el resultado de procesos operativos normales. Una comprobación es una comparación de una instantánea con un modelo de referencia adaptativo. Se puede realizar automáticamente una comprobación. La salida de una comprobación puede comprender un conjunto de anomalías y condiciones que se han detectado. En una realización, la anomalía se hace coincidir con a una pluralidad de filtros de reconocimiento. Un filtro de reconocimiento comprende una firma de una condición o de una clase de condiciones. Por ejemplo, un filtro de reconocimiento puede comprender una colección de pares de nombres y valores de activos que, cuando se toman juntos, representan la firma de una condición que es deseable reconocer, tal como la presencia de un gusano. Un filtro de reconocimiento genérico puede proporcionar una plantilla para crear filtros más específicos. Por ejemplo, un filtro de reconocimiento que está adaptado para buscar gusanos en general se puede adaptar para buscar un gusano específico.

En una realización de la presente invención, un filtro de reconocimiento comprende al menos uno de: un nombre de activo asociado con la condición, un valor de activo asociado con la condición, una combinación de nombre de activo y valor de activo asociados con la condición, un umbral máximo asociado con un valor de activo y con la condición, y un umbral mínimo asociado con un valor de activo y con la condición. Los pares nombre/valor de activo de una instantánea se pueden comparar con los pares nombre/valor del filtro de reconocimiento para encontrar una coincidencia y diagnosticar una condición. La coincidencia de nombre/valor puede ser exacta o el filtro de reconocimiento puede comprender un comodín, que permite que un valor parcial sea introducido en el filtro de reconocimiento y luego hecho coincidir con la instantánea. Un nombre y/o un valor de activo particular se pueden hacer coincidir con una pluralidad de filtros de reconocimiento con el fin de diagnosticar una condición.

Se puede crear un filtro de reconocimiento de diversas formas. Por ejemplo, en una realización de la presente invención, un usuario copia las anomalías de una máquina donde está presente la condición de interés. Las anomalías se pueden presentar en un resumen de anomalías a partir del cual se pueden seleccionar y copiar al filtro. En otra realización, un usuario introduce un carácter comodín en una definición de filtro. Por ejemplo, una parte de programa espía llamado Gator genera miles de claves de registro que comienzan con la cadena “hklm\software\gator\”. Una realización de la presente invención puede proporcionar un mecanismo de comodín para tratar eficazmente esta situación. El carácter de comodín puede ser, por ejemplo, el signo de porcentaje (%), y se puede usar antes de una cadena de texto, después de una cadena de texto o en medio de una cadena de texto. Continuando con el ejemplo de Gator, si el usuario introduce la cadena “hklm\software\gator\%” en el cuerpo del filtro, entonces cualquier clave que comience con “hkml\software\gator” se reconocerá por el filtro. El usuario puede desear construir un filtro para una condición que aún no se ha experimentado en la población gestionada. Por ejemplo, un filtro para un virus basado en información públicamente disponible en Internet en lugar de un caso real del virus dentro de la población gestionada. Para abordar esta situación, el usuario introduce la información relevante directamente en un filtro.

La Figura 13 es una captura de pantalla de una interfaz de usuario para seleccionar una instantánea a usar para la creación de un filtro de reconocimiento en una realización de la presente invención. Un usuario accede a la captura de pantalla mostrada para seleccionar las instantáneas a ser usadas para crear el filtro de reconocimiento. La Figura 14 es una captura de pantalla de una interfaz de usuario para crear o editar un filtro de reconocimiento en una realización de la presente invención. En la realización mostrada, los activos de la instantánea seleccionada en la interfaz ilustrada en la Figura 13 se muestran en la ventana 1402 Fuente de Datos. El usuario selecciona estos activos y los copia en la ventana 1404 Fuente para crear el filtro de reconocimiento.

En una realización, la coincidencia entre un filtro de reconocimiento y un conjunto de anomalías está asociada con una medida de calidad. Por ejemplo, una coincidencia exacta de todos los nombres de activos y valores de activos en el filtro de reconocimiento con nombres de activos y valores de activos en el conjunto de anomalías puede estar asociada con una medida de calidad más alta que una coincidencia de un subconjunto de los nombres de activos y valores de activos en el filtro de reconocimiento con nombres de activos y valores de activos en el conjunto de anomalías.

El filtro de reconocimiento puede comprender también otros atributos. Por ejemplo, en una realización, el filtro de reconocimiento comprende una marca de control para determinar si incluir el nombre de activo y el valor de activo en el modelo de referencia adaptativo. En otra realización, el filtro de reconocimiento comprende una o más descripciones textuales asociadas con una o más condiciones. Aún en otra realización, el filtro de reconocimiento comprende un indicador de gravedad que indica la gravedad de una condición en términos de, por ejemplo, cuánto daño puede causar, cuán difícil puede ser de eliminar, o alguna otra medida adecuada.

El filtro de reconocimiento puede comprender campos que son de naturaleza administrativa. Por ejemplo, en una realización, el filtro de reconocimiento comprende un identificador de filtro de reconocimiento, un nombre de creador y una fecha-hora de actualización.

Aún con referencia a la Figura 6, el componente (110) Analítico a continuación responde a la condición 614. Responder a la condición puede comprender, por ejemplo, generar una notificación, tal como un correo electrónico a un técnico de soporte, enviar un tique de problema a un sistema de gestión de problemas, solicitar permiso para tomar una acción, por ejemplo, pedir la confirmación de un técnico de soporte para instalar un parche, y eliminar la condición de al menos uno de la pluralidad de ordenadores. Eliminar la condición puede comprender, por ejemplo, hacer que se ejecute un agente de respuesta en cualquiera de la pluralidad de ordenadores afectados por la condición. La condición puede estar asociada con una respuesta automática. Los pasos de diagnóstico 612 y de respuesta a las condiciones 614 se pueden repetir para cada condición. También, el proceso de encontrar anomalías 610 se puede repetir para cada instantánea individual.

En la realización mostrada en la Figura 6, el componente (110) Analítico determina a continuación si han de ser analizadas 616 instantáneas adicionales. Si es así, los pasos de comparar la instantánea con el modelo 608 de referencia adaptativo, encontrar anomalías 610, hacer coincidir las anomalías con un filtro de reconocimiento para diagnosticar una condición 612, y responder a la condición 614 se repiten para cada instantánea. Una vez que han sido analizadas todas las instantáneas, el proceso termina 618.

En una realización de la presente invención, una vez que el componente (110) Analítico ha identificado una condición, el componente (110) Analítico intenta determinar cuáles de la pluralidad de ordenadores dentro de una población están afectados por la condición. Por ejemplo, el componente (110) Analítico puede examinar las instantáneas para identificar un conjunto particular de anomalías. El componente (110) Analítico entonces puede hacer que una respuesta a la condición sea ejecutada en nombre de cada uno de los ordenadores afectados. Por ejemplo, en una realización, un componente (202) de Agente reside en cada uno de la pluralidad de ordenadores. El componente (202) de Agente genera la instantánea que se evalúa por el componente (110) Analítico. En una realización tal, el componente (110) Analítico utiliza el componente (202) de Agente para ejecutar un programa de respuesta si el componente (110) Analítico identifica una condición en uno de los ordenadores. En el diagnóstico de una condición, el componente (110) Analítico puede ser capaz de identificar o no una causa raíz de una condición.

La Figura 7 es un diagrama de flujo que ilustra un proceso para identificar ciertos tipos de anomalías en una realización de la presente invención. En la realización mostrada, el componente (110) Analítico evalúa instantáneas para una pluralidad de ordenadores 702. Estas instantáneas pueden ser instantáneas base que comprenden el estado completo del ordenador o instantáneas delta que comprenden los cambios en el estado del ordenador desde la última Instantánea base. El componente (110) Analítico usa las instantáneas para crear un modelo 704 de referencia adaptativo. Señalar que cuando se usan instantáneas delta para este propósito, el componente Analítico debe reconstituir primero el equivalente de una instantánea base aplicando los cambios descritos en la instantánea delta a la instantánea base más reciente. El componente (110) Analítico recibe posteriormente una segunda instantánea (base o delta) para al menos uno de la pluralidad de ordenadores 706. La instantánea se puede crear basada en diversos eventos, tales como el paso de una cantidad de tiempo predeterminada, la instalación de un nuevo programa, o algún otro evento adecuado.

El componente (110) Analítico compara la segunda instantánea con el modelo de referencia adaptativo para intentar y detectar anomalías. Pueden existir diversos tipos de anomalías en un ordenador. En la realización mostrada, el componente (110) Analítico primero intenta identificar nombres de activos que estén ausentes inesperadamente 710. Por ejemplo, todos o sustancialmente todos los ordenadores dentro de una población pueden incluir un archivo particular. La existencia del archivo se señala en el modelo de referencia adaptativo mediante la presencia de un nombre de activo. Si el archivo está ausente inesperadamente de uno de los ordenadores dentro de la población, es decir, no se encuentra el nombre de activo, alguna condición puede estar afectando al ordenador en el que falta el archivo. Si el nombre de activo está ausente inesperadamente, la ausencia se identifica como una anomalía 712. Por ejemplo, una entrada que identifica el ordenador, la fecha y un activo ausente inesperadamente se puede introducir en un almacén de datos.

El componente (110) Analítico intenta a continuación identificar los nombres de activos que están presentes inesperadamente 714. La presencia de un nombre de activo inesperado, tal como un nombre de archivo o entrada de registro, puede indicar la presencia de una condición problemática, como un gusano informático. Un nombre de activo está presente inesperadamente si nunca se ha visto antes o si nunca se ha visto antes en el contexto en el que se encuentra. Si el nombre de activo está presente inesperadamente, la presencia se identifica como una anomalía 720.

El componente (110) Analítico intenta a continuación identificar un valor 718 de activo inesperado. Por ejemplo, en una realización, el componente (110) Analítico intenta identificar un valor de activo de cadena que es desconocido para el nombre de activo asociado con él. En otra realización, el componente (110) Analítico compara un activo numérico con los umbrales mínimo o máximo asociados con el nombre de activo correspondiente. En realizaciones de la presente invención, los umbrales se pueden ajustar automáticamente basados en la desviación media y estándar para valores de activos dentro de una población. Según la realización mostrada, si se detecta un valor de activo inesperado, se identifica como una anomalía 720. El proceso termina entonces 722.

Aunque el proceso en la Figura 7 se muestra como un proceso en serie, la comparación de una instantánea con el modelo de referencia adaptativo y la identificación de anomalías pueden ocurrir en paralelo. Además, cada uno de los pasos representados se puede repetir numerosas veces. Además, o bien las instantáneas delta o bien las instantáneas base se pueden comparar con el modelo de referencia adaptativo durante cada ciclo.

Una vez que se ha completado el análisis, el componente (110) Analítico puede generar un resultado, tal como un informe de anomalía. Este informe se puede proporcionar además a un usuario. Por ejemplo, el componente (110) Analítico puede generar una página web que comprende los resultados de una comparación de una instantánea con un modelo de referencia adaptativo. Las realizaciones de la presente invención pueden proporcionar unos medios para realizar auditorías de seguridad automatizadas, comprobación de integridad de archivos y registros, detección de virus basada en anomalías, y reparación automatizada.

La Figura 8 es un diagrama de flujo que ilustra un proceso para generar un modelo de referencia adaptativo en una realización de la presente invención. En la realización mostrada, el componente (110) Analítico accede a una pluralidad de instantáneas desde una pluralidad de ordenadores a través del componente de Base de Datos. Cada una de las instantáneas comprende una pluralidad de pares de nombres de activos y valores de activos. El componente (110) Analítico crea automáticamente un modelo de referencia adaptativo que se basa, al menos en parte, en las instantáneas.

El modelo de referencia adaptativo puede comprender cualquiera de una serie de atributos, relaciones, y medidas de los diversos nombres y valores de activos. En la realización mostrada en la Figura 8, el componente (110) Analítico encuentra primero uno o más nombres de activos únicos y entonces determina el número de veces que ocurre cada nombre de activo único dentro de la pluralidad de instantáneas 804. Por ejemplo, puede ocurrir un archivo para un controlador de sistema operativo básico sustancialmente en todos los ordenadores dentro de una población. El nombre de archivo es un nombre de activo único; aparecerá solamente una vez dentro de una instantánea, pero probablemente ocurrirá sustancialmente en todas las instantáneas.

En la realización mostrada, el componente (110) Analítico determina a continuación los valores de activos únicos asociados con cada nombre de activo 806. Por ejemplo, el activo de nombre de archivo para el controlador descrito en relación con el paso 804 probablemente tendrá el mismo valor para cada aparición del activo de nombre de archivo. Por el contrario, el valor de archivo para un archivo de registro probablemente tendrá tantos valores diferentes como apariciones, es decir, un archivo de registro en cualquier ordenador particular contendrá un número diferente de entradas de todos los otros ordenadores en una población.

Dado que la población puede ser muy grande, en la realización mostrada en la Figura 8, si el número de valores únicos asociados con un nombre de activo excede un umbral 808, se detiene 810 la determinación. En otras palabras, en el ejemplo del archivo de registro descrito anteriormente, si el ordenador está o no en un estado normal no depende de un archivo de registro que tenga un valor coherente. Los contenidos de archivo de registro se espera que varíen en cada ordenador. Señalar, sin embargo, que la presencia o ausencia del archivo de registro se puede almacenar en el modelo de referencia adaptativo como una indicación de normalidad o de una anomalía.

En la realización mostrada en la Figura 8, el componente (110) Analítico determina a continuación los valores de activos de cadena únicos asociados con cada nombre de activo 812. Por ejemplo, en una realización, solamente hay dos tipos de valores de activos, cadenas y números. Los valores de comprobaciones aleatorias de archivos y claves de registro son ejemplos de cadenas; un valor de contador de rendimiento es un ejemplo de un número.

El componente (110) Analítico determina a continuación una medida estadística asociada con valores numéricos únicos asociados con un nombre de activo 814. Por ejemplo, en una realización, el componente (110) Analítico captura una medida de rendimiento, tal como una búsqueda de memoria. Si un ordenador en una población a menudo busca en la memoria, puede ser una indicación de que un programa malicioso está ejecutándose en segundo plano y requiriendo recursos de memoria sustanciales. Sin embargo, si todos o un número considerable de ordenadores de una población a menudo buscan en la memoria, puede indicar que los ordenadores generalmente carecen de recursos de memoria. En una realización, el componente (110) Analítico determina una desviación media y una estándar para valores numéricos asociados con un nombre de activo único. En el ejemplo de la memoria, si la medida de la búsqueda de memoria para un ordenador cae lejos fuera de la media estadística para la población, se puede identificar una anomalía.

En una realización de la presente invención, el modelo de referencia adaptativo se puede modificar aplicando una plantilla de políticas. Una plantilla de políticas es una colección de pares de activo/valor que se identifican y se aplican a un modelo de referencia adaptativo para establecer una norma que refleja una política específica. Por ejemplo, la plantilla de políticas puede comprender una pluralidad de pares de nombres de activos y valores de activos que se espera que estén presentes en un ordenador normal. En una realización, la aplicación de la plantilla de políticas comprende modificar el modelo de referencia adaptativo de modo que los pares de nombres de activos y valores de activos presentes en la plantilla de políticas parecen haber estado presentes en cada una de la pluralidad de instantáneas, es decir, parecen ser el estado normal de un ordenador en la población.

La Figura 15 es una captura de pantalla que ilustra una interfaz de usuario para seleccionar un “sistema excelente” para uso en una plantilla de políticas en una realización de la presente invención. Como se ha descrito anteriormente, el usuario selecciona primero el sistema excelente en el que se ha de basar la plantilla de políticas. La Figura 16 es una captura de pantalla de una interfaz de usuario para seleccionar activos de plantilla de políticas en una realización de la presente invención. Como con la interfaz de usuario para crear filtros de reconocimiento. El usuario selecciona activos desde una ventana 1602 de Fuente de Datos y los copia a una ventana de contenidos, la ventana 1604 de contenido de Plantilla.

La Figura 9 es un diagrama de flujo, que ilustra un proceso para la detección proactiva de anomalías en una realización de la presente invención. En la realización mostrada, cuando ocurre un análisis, el componente (110) Analítico establece una conexión con la base de datos (112) que almacena las instantáneas a ser analizadas 902. En la realización mostrada, se utiliza solamente una base de datos. Sin embargo, en otras realizaciones, se pueden analizar datos de múltiples bases de datos.

Antes de que se ejecuten las comprobaciones de diagnóstico, se crean 904 uno o más modelos de referencia. Los modelos de referencia se actualizan periódicamente, por ejemplo, una vez por semana, para asegurar que la información que contienen permanece actual. Una realización de la presente invención proporciona un programador de tareas que permite que la creación del modelo sea configurada como un procedimiento completamente automatizado.

Una vez que se ha creado un modelo de referencia, se puede procesar de diversas formas para permitir diferentes tipos de análisis. Por ejemplo, es posible definir una plantilla 906 de políticas como se ha descrito anteriormente. Por ejemplo, una plantilla de políticas puede requerir que todas las máquinas de una población gestionada tengan un software antivirus instalado y operativo. Una vez que se ha aplicado una plantilla de políticas a un modelo, las comprobaciones de diagnóstico frente a ese modelo incluirán una prueba de cumplimiento de políticas. Las plantillas de políticas se pueden usar en una variedad de aplicaciones incluyendo auditorías de seguridad automatizadas, comprobación de umbral de rendimiento y gestión de actualizaciones de Windows. Una plantilla de políticas comprende el conjunto de activos y valores que serán forzados dentro del modelo como norma. En una realización, el proceso de edición de plantillas se basa en un planteamiento de “sistema excelente”. Un sistema excelente es uno que presenta los activos y valores que un usuario desea incorporar dentro de la plantilla. El usuario localiza la instantánea que corresponde al sistema excelente y luego selecciona cada par de activo/valor que el usuario desea incluir en la plantilla.

En el proceso mostrado en la Figura 9, la plantilla de políticas se aplica entonces a un modelo para modificar su definición de normal 908. Esto permite que el modelo sea formado de formas que le permita comprobar el cumplimiento frente a las políticas definidas por el usuario como se describe en la presente memoria.

Un modelo también se puede convertir 910. El proceso de conversión altera un modelo de referencia. Por ejemplo, en una realización, el proceso de conversión elimina del modelo cualesquiera activos de información que sean únicos, es decir, cualesquiera activos que ocurran en una y sólo una instantánea. Cuando se ejecuta una

comprobación frente a un modelo convertido, todos los activos de información únicos se notificarán como anomalías. Este tipo de comprobación es útil en emerger condiciones problemáticas previamente desconocidas que existen en el momento en que se instalan por primera vez los componentes de Agente. Los modelos convertidos son útiles en el establecimiento de una línea base inicial dado que exponen características únicas. Por esta razón los modelos convertidos algunas veces se llaman modelos de línea de base en las realizaciones de la presente invención.

En otra realización, el proceso de construcción del modelo elimina del modelo cualesquiera activos de información que coincidan con un filtro de reconocimiento, garantizando que las condiciones problemáticas conocidas no se incorporen al modelo. Cuando el sistema se instala por primera vez, la población gestionada contiene bastante a menudo una serie de condiciones problemáticas conocidas que aún no han sido notificadas. Es importante descubrir estas condiciones y eliminarlas del modelo, dado que, de otro modo, se incorporarán al modelo de referencia adaptativo como parte del estado normal de una máquina.

El componente (202) de Agente toma una instantánea del estado de cada máquina gestionada de una forma programada 924. La instantánea se transmite y se introduce en la base de datos como una instantánea. También se pueden generar instantáneas bajo demanda o en respuesta a un evento específico, tal como la instalación de una aplicación.

En el proceso proactivo de gestión de problemas mostrado, se realiza 912 una comprobación periódica de las últimas instantáneas frente a un modelo de referencia actualizado. La salida de una comprobación periódica es un conjunto de anomalías, que se muestran a un usuario como resultados 914. Los resultados también incluyen cualesquiera condiciones que se identifican como resultado de la coincidencia de las anomalías con los filtros de reconocimiento. Los filtros de reconocimiento se pueden definir como se ha descrito anteriormente 916. Las anomalías se pasan a través de los filtros de reconocimiento para su interpretación dando como resultado un conjunto de condiciones. Las condiciones pueden oscilar en gravedad de algo tan benigno como una actualización de Windows a algo tan grave como un Troyano.

Las condiciones problemáticas que pueden ocurrir en un ordenador cambian a medida que evolucionan los componentes de hardware y de software que componen ese ordenador. En consecuencia, hay una necesidad continua de definir y compartir nuevos filtros de reconocimiento a medida que se descubren nuevas combinaciones de anomalías. Los filtros de reconocimiento se pueden considerar como una forma muy detallada y estructurada de documentar las condiciones problemáticas y, como tales, representan un mecanismo importante para facilitar la colaboración. La realización mostrada comprende un mecanismo para exportar filtros de reconocimiento a un archivo XML e importar filtros de reconocimiento desde un archivo XML.

Una vez que se identifican las condiciones, se generan 920, informes que documentan los resultados de una comprobación proactiva. Los informes pueden comprender, por ejemplo, una descripción resumida de todas las condiciones detectadas o una descripción detallada de una condición particular.

La Figura 10 es un diagrama de flujo, que ilustra un proceso reactivo en una realización de la presente invención. En el proceso mostrado en la Figura 10, se supone que ya se ha creado un modelo de referencia adaptativo. El proceso mostrado comienza cuando un usuario llama a un servicio de soporte técnico para notificar un problema 1002. En el paradigma tradicional de servicio de soporte técnico, el siguiente paso sería recopilar verbalmente información acerca de los síntomas que se experimentan por el usuario. Por el contrario, en la realización de la presente invención mostrada, el siguiente paso es ejecutar una comprobación de diagnóstico de la máquina sospechosa frente a la instantánea 1003 más reciente. Si esto no produce un diagnóstico inmediato de una condición problemática, pueden existir tres posibilidades: (1) La condición ha ocurrido desde que se tomó la última instantánea; (2) la condición es nueva y no está siendo reconocida por sus filtros; o (3) la condición está fuera del alcance del análisis, por ejemplo, un problema de hardware.

Si se sospecha que la condición problemática ha ocurrido desde que se tomó la última instantánea, entonces el usuario puede hacer que el componente (202) de Agente en la máquina cliente tome otra instantánea 1006. Una vez que está disponible la instantánea resultante, se puede ejecutar 1004 una nueva comprobación de diagnóstico.

Si se sospecha que la condición problemática es nueva, el analista puede ejecutar una función de comparación que proporciona un desglose de los cambios en el estado de una máquina sobre una ventana de tiempo específica tales como nuevas aplicaciones que pueden haber sido instaladas 1008. El usuario también puede ver una representación detallada del estado de una máquina en diversos puntos en el tiempo 1010. Si el analista identifica una nueva condición problemática, el usuario puede identificar el conjunto de activos como un filtro de reconocimiento para análisis 1012 posteriores.

Aunque los productos convencionales se han centrado en mejorar la eficacia del modelo de soporte basado en seres humanos, las realizaciones de la presente invención están diseñadas alrededor de un paradigma diferente, un modelo de soporte basado en máquina. Esta diferencia fundamental en el planteamiento se manifiesta a sí misma más profundamente en las áreas de recopilación y análisis de datos. Dado que una máquina en lugar de un ser humano realizará gran parte del análisis de los datos recogidos, los datos recogidos pueden ser voluminosos. Por

ejemplo, en una realización, los datos recogidos de una única máquina, conocidos como “comprobación de salud” o instantánea para la máquina, incluyen valores para cientos de miles de atributos. La capacidad de recoger un gran volumen de datos proporciona realizaciones de la presente invención con una ventaja significativa sobre los sistemas convencionales en términos del número y variedad de condiciones que se pueden detectar.

Otra realización de la presente invención proporciona una poderosa capacidad analítica. La base para el análisis de alto valor en tal realización es la capacidad de distinguir con precisión entre condiciones normales y anormales. Por ejemplo, un sistema según la presente invención sintetiza su modelo de referencia automáticamente extrayendo relaciones estadísticamente significativas a partir de los datos de instantáneas que recoge de sus clientes. El modelo de referencia “adaptativo” resultante define qué es normal para esa población gestionada particular en ese momento particular en el tiempo.

Una realización de la presente invención combina la recopilación de datos y las características de análisis adaptativo descritas anteriormente. En tal realización, las capacidades superiores de recopilación de datos combinadas con la potencia analítica del modelo de referencia adaptativo se traducen en una serie de ventajas competitivas significativas, incluyendo la capacidad de proporcionar protección automática frente a amenazas de seguridad dirigiendo auditorías de seguridad diarias y comprobando las actualizaciones de software para eliminar vulnerabilidades. Tal realización también puede ser capaz de explorar proactivamente todos los sistemas gestionados de una forma rutinaria para encontrar problemas antes de que den como resultado una pérdida de productividad o llamadas al servicio de soporte técnico.

Una realización de la presente invención que implementa las capacidades del modelo de referencia adaptativo también es capaz de detectar condiciones problemáticas previamente desconocidas. Además, tal realización se sintetiza y mantiene automáticamente, requiriendo pocas o ningunas actualizaciones del proveedor para ser eficaz. Tal realización se personaliza automáticamente para una población gestionada particular, permitiéndole detectar modos de fallo únicos para esa población.

Una ventaja adicional de una realización de la presente invención es que en el caso de que una condición problemática no se pueda resolver automáticamente, tal realización puede proporcionar una cantidad masiva de información técnica estructurada para facilitar el trabajo del analista de soporte.

Una realización de la presente invención proporciona la capacidad de reparar automáticamente un problema identificado. Tal realización, cuando se combina con el modelo de referencia adaptativo de la realización descrita previamente, es capaz únicamente de reparación automatizada debido a su capacidad de identificar todos los aspectos de una condición problemática.

Las realizaciones de la presente invención también proporcionan muchas ventajas sobre los sistemas y métodos convencionales en términos de los niveles de servicio descritos en la presente memoria. Por ejemplo, en términos del nivel de servicio de Curación en Masa, es considerablemente menos costoso prevenir un incidente que resolver un incidente una vez que ha ocurrido el daño. Las realizaciones de la presente invención aumentan sustancialmente el porcentaje de incidentes que se pueden detectar/prevenir sin necesidad de intervención humana y de una manera que abarque la naturaleza diversa y dinámica de los ordenadores en entornos del mundo real.

Además, una realización de la presente invención es capaz de dirigir el nivel de servicio de Auto-Curación detectando y reparando automáticamente tanto anomalías conocidas como desconocidas. Una realización que implementa el modelo de referencia adaptativo descrito en la presente memoria está adecuada de forma única a la detección y reparación automáticas. El servicio y la reparación automáticos también ayudan a eliminar o al menos minimizar la necesidad de Autoservicio y de Visitas del lado del Soporte Técnico.

Las realizaciones de la presente invención proporcionan ventajas en el nivel de Servicio Asistido proporcionando capacidades superiores de diagnóstico y recursos de información extensos. Una realización recopila y analiza cantidades masivas de datos de usuario final, facilitando una variedad de necesidades asociadas con el modelo de soporte basado en seres humanos incluyendo: auditorías de seguridad, auditorías de configuración, gestión de inventario, análisis de rendimiento y diagnóstico de problemas.

REIVINDICACIONES

1. Un método de soporte informático automatizado que comprende:

i) recibir (602) una instantánea desde un ordenador (116a, b);
ii) comparar (608) la instantánea con una base de datos de estados de ordenador; y,
iii) identificar una anomalía basada en el resultado de la comparación;

caracterizado por:

iv) recibir (602) instantáneas desde una pluralidad de ordenadores (116a, b) dentro de una población de ordenadores, en donde las instantáneas individuales incluyen datos que indican un estado de un ordenador respectivo;
v) almacenar (604) las instantáneas en un almacén de datos;
vi) crear (606) automáticamente un modelo (206, 402) de referencia adaptativo basado al menos en parte en las instantáneas y que comprende un conjunto de reglas personalizado a las características de la población de ordenadores, estando desarrollado el conjunto de reglas identificando patrones entre las instantáneas de la pluralidad de ordenadores de manera que el modelo de referencia adaptativo es indicativo de estados normales en los ordenadores dentro de la población;
vii) comparar (608) instantáneas de al menos uno de la pluralidad de ordenadores con el modelo de referencia adoptivo; y
viii) determinar (610), basado en el resultado de la comparación, si está presente una anomalía (720) en el estado del al menos uno de los ordenadores.

2. El método de la reivindicación 1, que comprende además hacer coincidir (612) al menos una anomalía (720) con al menos un filtro (216) de reconocimiento para diagnosticar una condición en al menos uno de la pluralidad de ordenadores (116a, b).

3. El método de la reivindicación 2, en donde el filtro (216) de reconocimiento comprende un patrón particular de anomalías que indica la presencia de una condición de causa raíz particular o una clase genérica de condiciones.

4. El método de la reivindicación 3, que comprende responder (614) a la condición mediante al menos uno de:-

i) generar una notificación;
ii) enviar un tique de problema a un sistema de gestión de problemas;
iii) solicitar permiso para tomar una acción;
y,
iv) eliminar la condición de al menos uno de la pluralidad de ordenadores (116a, b).

5. El método de la reivindicación 4, en donde la eliminación de la condición comprende hacer que un programa de reparación sea ejecutado en al menos uno de la pluralidad de ordenadores (116a, b) afectados por la condición.

6. El método de cualquiera de las reivindicaciones 3 a 5, que comprende además:

i) determinar cuáles de la pluralidad de ordenadores (116a, b) están afectados por la condición;
y,
ii) causar una respuesta (614) a la condición a ser ejecutada en nombre de cada uno de la pluralidad de ordenadores afectados por la condición.

7. El método de cualquiera de las reivindicaciones 3 a 6, en donde el diagnóstico (612) de la condición comprende identificar una causa raíz de al menos una anomalía (720).

8. El método de cualquiera de las reivindicaciones 2 a 7, en donde al menos un filtro (216) de reconocimiento está asociado con una respuesta (214) automatizada para la condición.

9. El método de cualquiera de las reivindicaciones 2 a 8, en donde la condición es una clase que comprende un grupo de condiciones.

10. El método de cualquiera de las reivindicaciones 2 a 9, que además comprende determinar una calidad de una coincidencia (612) entre al menos un filtro (216) de reconocimiento y al menos una anomalía (720).

11. El método de cualquiera de las reivindicaciones 1 a 10, en donde el modelo (402) de referencia adaptativo comprende una pluralidad de activos cada uno asociado con un tipo de activo que comprende uno de:-

i) un archivo,
ii) una clave de registro,

iii) una medida de rendimiento,
iv) un servicio,
v) un componente de hardware,
vi) un proceso en ejecución,
vii) un registro,
y
viii) un puerto de comunicación.

12. El método de cualquiera de las reivindicaciones 1 a 11, en donde comparar (912) al menos una de la pluralidad de instantáneas con el modelo (402) de referencia adaptativo comprende:-

i) generar un resultado (914);
y,
ii) proporcionar (920) el resultado a un usuario.

13. El método de cualquiera de las reivindicaciones 1 a 12, en donde la pluralidad de instantáneas se crea mediante un agente de software que reside en cada uno de la pluralidad de ordenadores (116a, b).

14. Un sistema (102) de soporte informático automatizado que comprende:

i) un componente (108) Colector; y,
ii) un componente (110) analítico en comunicación con el componente colector;

caracterizado por que:

iii) el sistema (102) de soporte automatizado está organizado y dispuesto para efectuar las reivindicaciones del método en cualquiera de las reivindicaciones 1 a 13.

15. Un medio legible por ordenador en el cual se codifica un código de programa que comprende instrucciones ejecutables por ordenador para efectuar el método de soporte informático automatizado de cualquiera de las reivindicaciones 1 a 13.

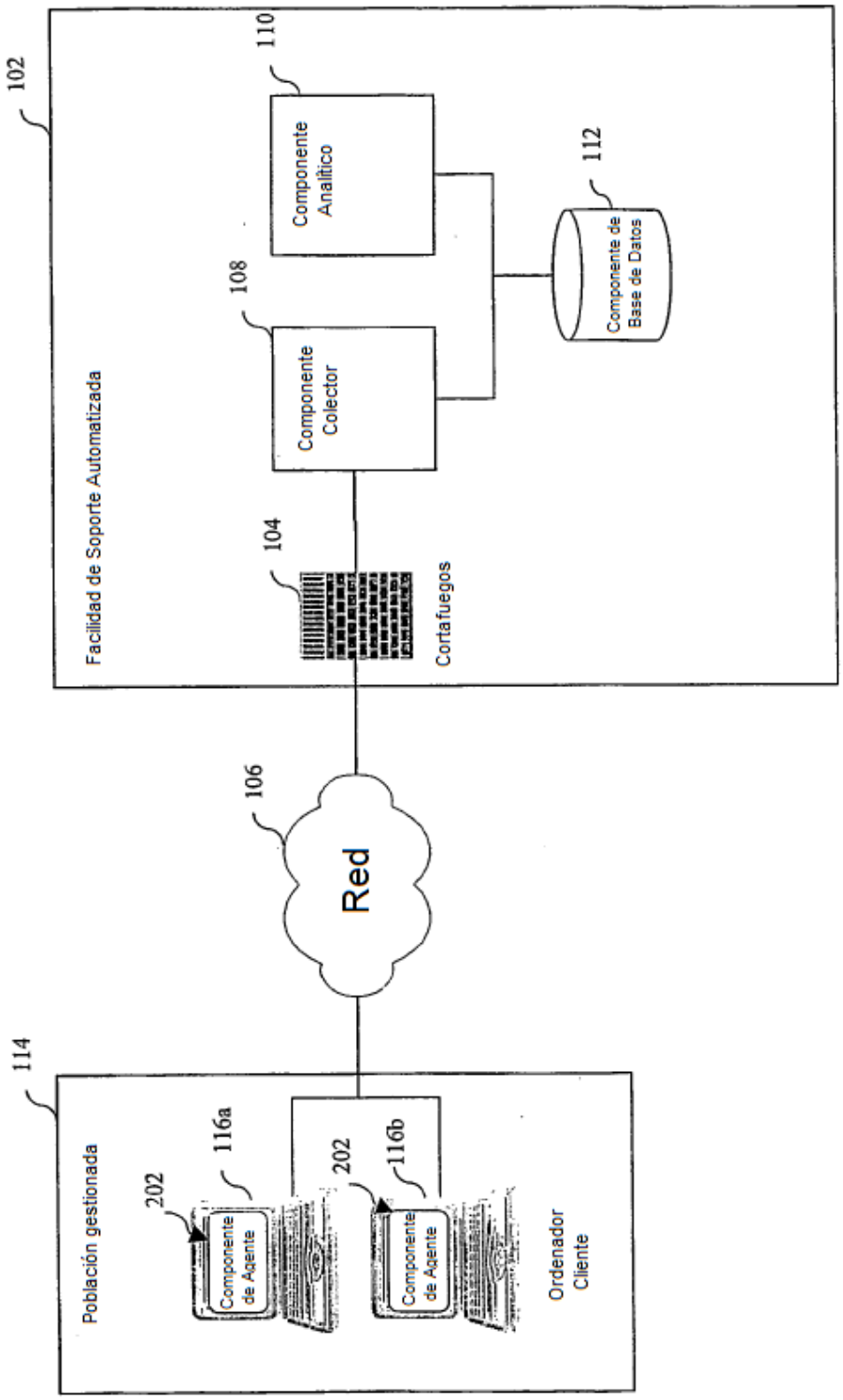
102
Facilidad de Soporte Automatizada
110
Componente Analítico
112
Componente de Base de Datos
108
Componente Colector
104
Cortafuegos
106
Red
114
Población gestionada
116a
202
Componente de Agente
202
116b
Componente de Agente
Ordenador Cliente

FIG. 1

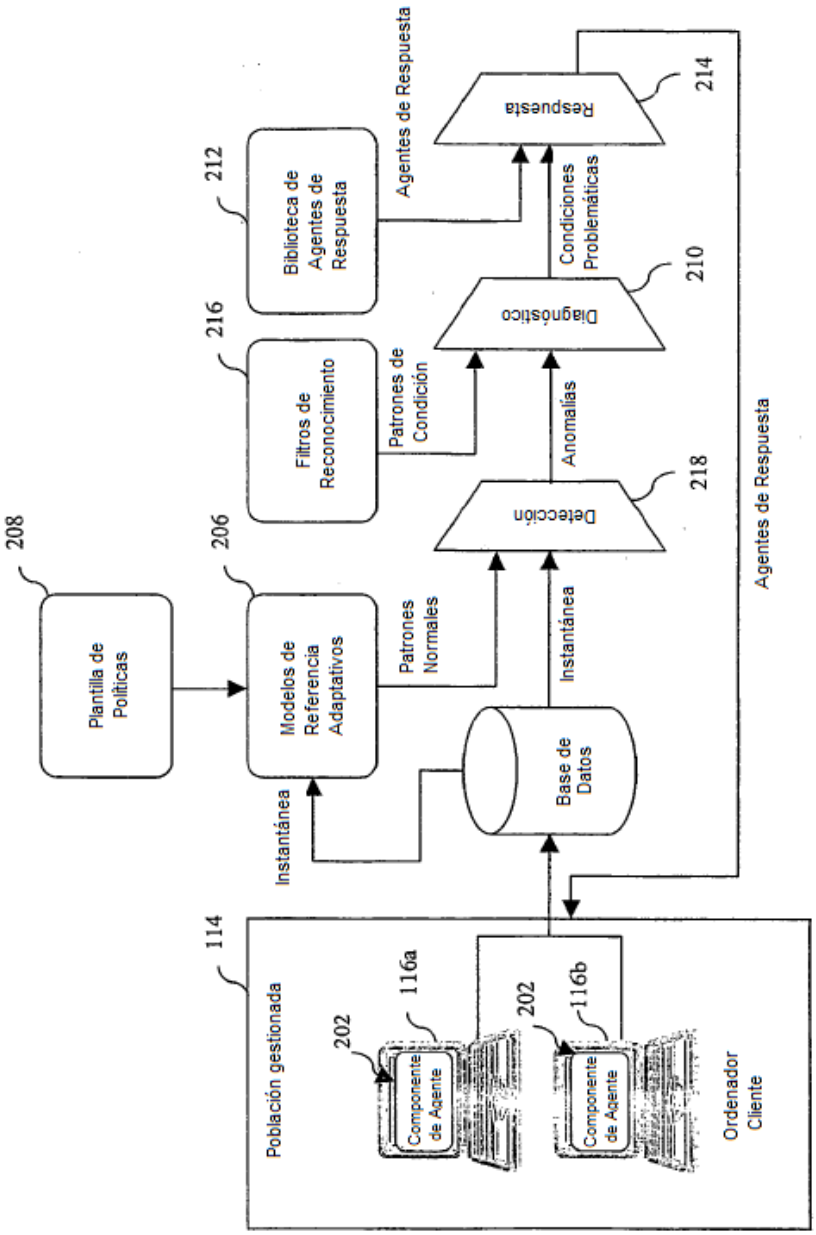
Población gestionada
114
Componente de Agente
202
116a
Componente de Agente
202
116b
Ordenador Cliente
Instantánea
Base de Datos
Instantánea
Patrones Normales
Modelos de Referencia Adaptativos
206
Plantilla de Políticas
208
Detección
218
Anomalías
Patrones de Condición
Filtros de Reconocimiento
216
Diagnóstico
210
Condiciones Problemáticas
Biblioteca de Agentes de Respuesta
212
Agentes de Respuesta
Respuesta
214
Agentes de Respuesta

FIG. 2

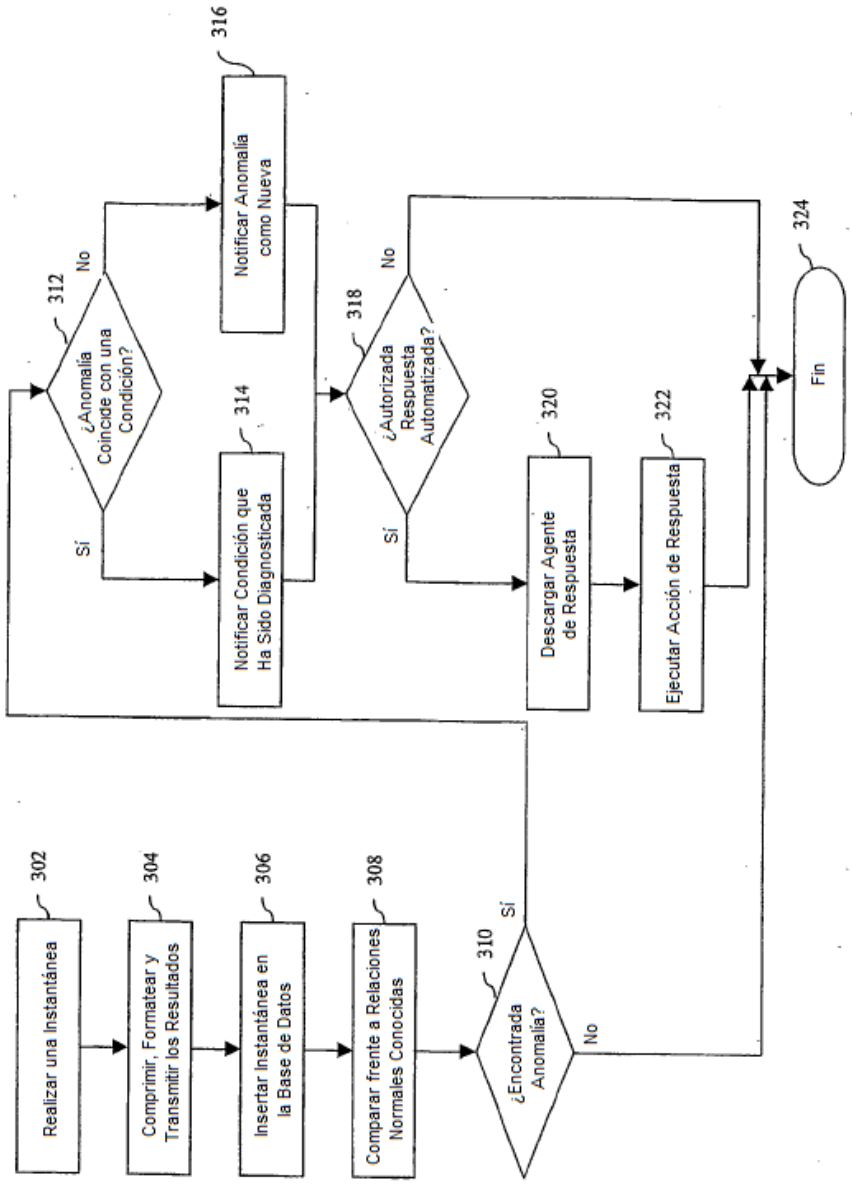
302
Realizar una Instantánea
304
Comprimir, Formatear y Transmitir los Resultados
306
Insertar Instantánea en la Base de Datos
308
Comparar frente a Relaciones Normales Conocidas
310
¿Encontrada Anomalía?
Sí
No
312
¿Anomalía Coincide con una Condición?
Sí
No
314
Notificar Condición que Ha Sido Diagnosticada
316
Notificar Anomalía como Nueva
318
¿Autorizada Respuesta Automatizada?
Sí
No
320
Descargar Agente de Respuesta
322
Ejecutar Acción de Respuesta
324
Fin

FIG. 3

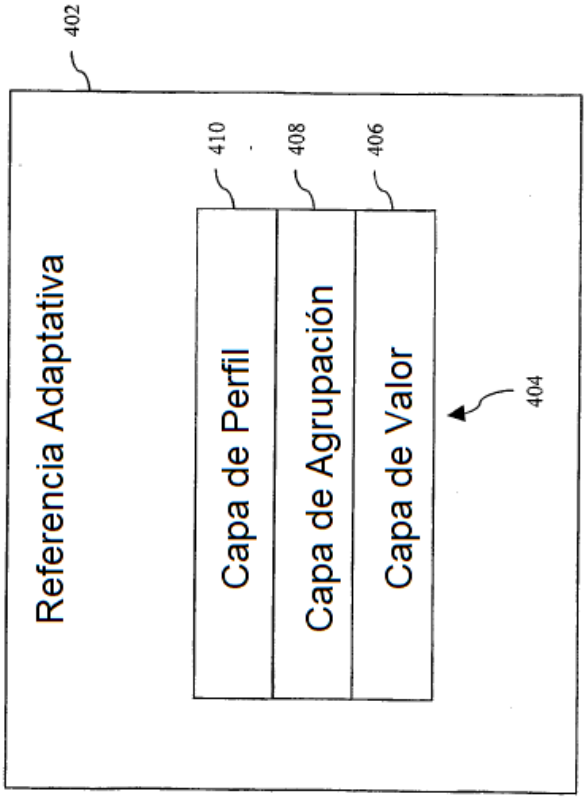
402
Referencia Adaptativa
410
408
406
Capa de Perfil
Capa de Agrupación
Capa de Valor
404

FIG. 4

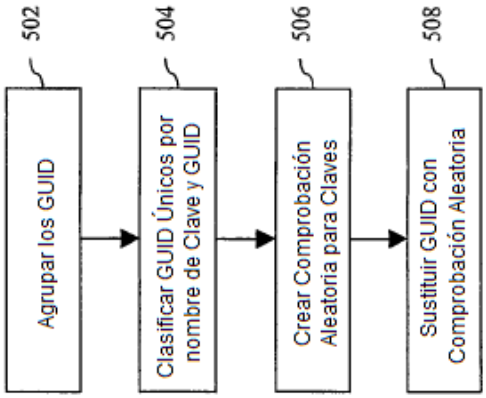
502
Agrupar los GUID
504
Clasificar GUID Únicos por nombre de Clave y GUID
506
Crear Comprobación Aleatoria para Claves
508
Sustituir GUID con Comprobación Aleatoria

FIG. 5

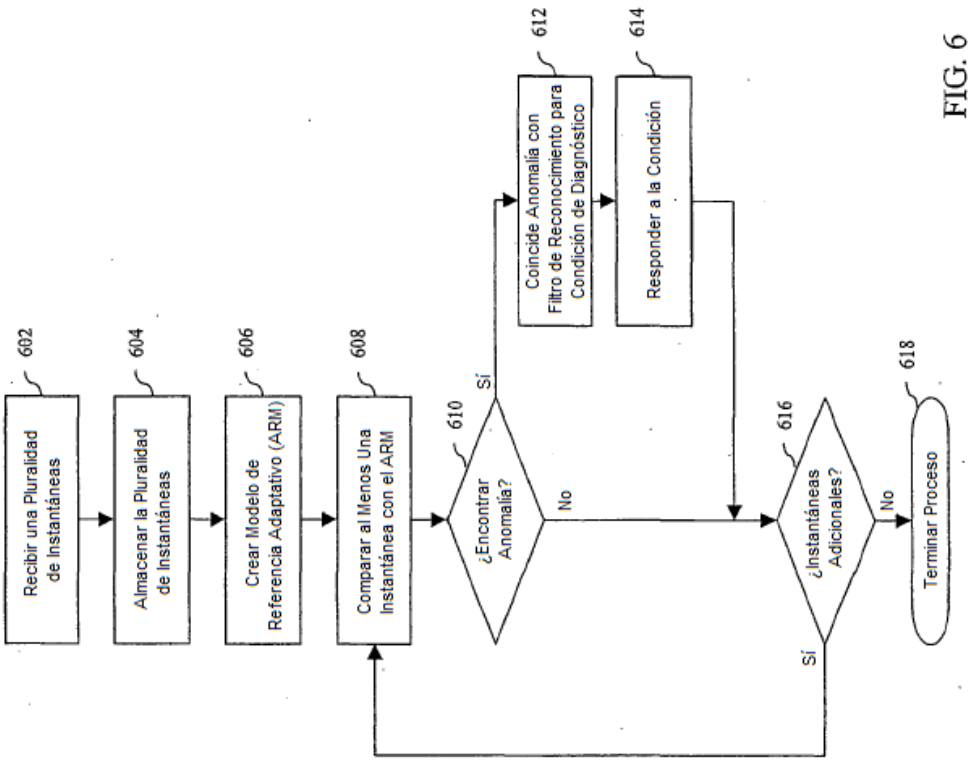
602
Recibir una Pluralidad de Instantáneas
604
Almacenar la Pluralidad de Instantáneas
606
Crear Modelo de Referencia Adaptativo (ARM)
608
Comparar al Menos Una Instantánea con el ARM
610
¿Encontrar Anomalía?
Sí
No
612
Coincide Anomalía con Filtro de Reconocimiento para Condición de Diagnóstico
614
Responder a la Condición
616
¿Instantáneas Adicionales?
Sí
No
618
Terminar Proceso

FIG. 6

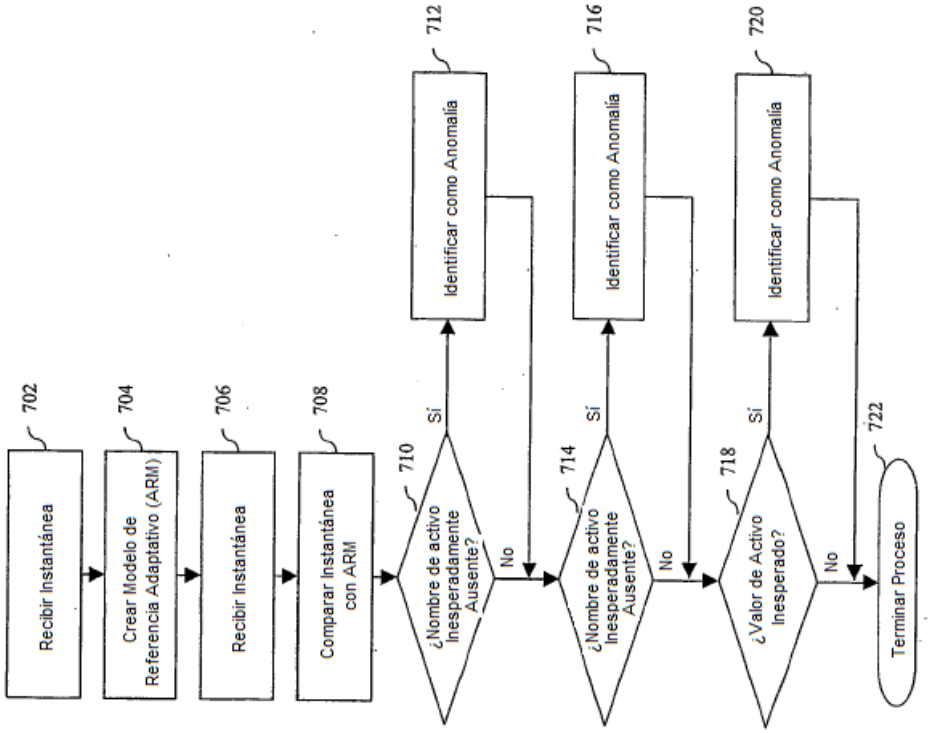
Recibir Instantánea
Crear Modelo de Referencia Adaptativo (ARM)
Recibir Instantánea
Comparar Instantánea con ARM
¿Nombre de activo Inesperadamente Ausente?
Sí
No
Identificar como Anomalía
¿Nombre de activo Inesperadamente Ausente?
Sí
No
Identificar como Anomalía
¿Valor de Activo Inesperado?
Sí
No
Identificar como Anomalía
Terminar Proceso
702
704
706
708
710
712
714
716
718
720
722

FIG. 7

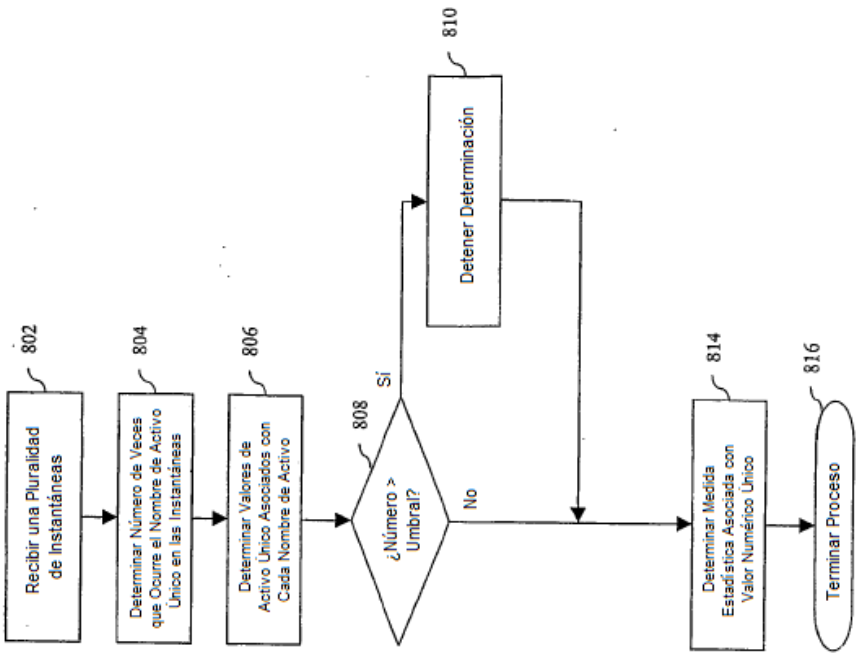
802
Recibir una Pluralidad de Instantáneas
804
Determinar Número de Veces que Ocurre el Nombre de Activo Único en las Instantáneas
806
Determinar Valores de Activo Único Asociados con Cada Nombre de Activo
808
¿Número > Umbral?
Sí
No
810
Detener Determinación
814
Determinar Medida Estadística Asociada con Valor Numérico Único
816
Terminar Proceso

FIG. 8

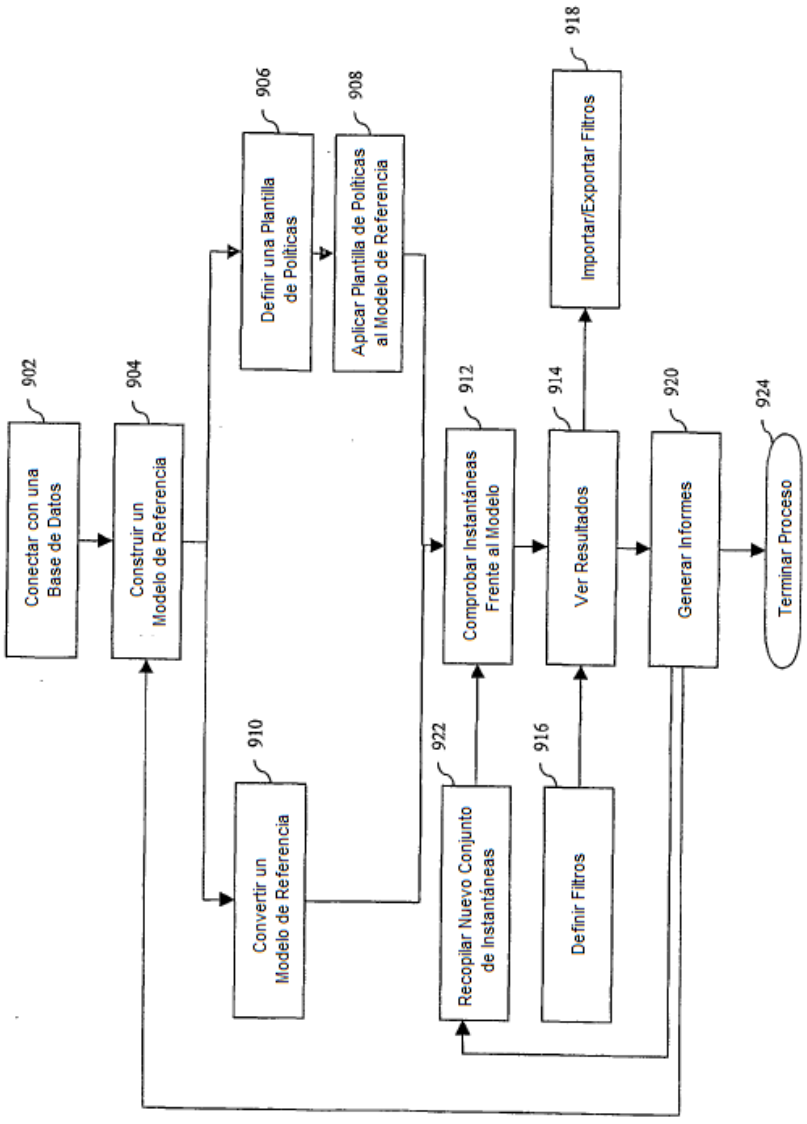
902
Conectar con una Base de Datos
904
Construir un Modelo de Referencia
906
Definir una Plantilla de Políticas
908
Aplicar Plantilla de Políticas al Modelo de Referencia
910
Convertir un Modelo de Referencia
912
Comprobar Instantáneas Frente al Modelo
914
Ver Resultados
916
Definir Filtros
918
Importar/Exportar Filtros
920
Generar Informes
922
Recopilar Nuevo Conjunto de Instantáneas
924
Terminar Proceso

FIG. 9

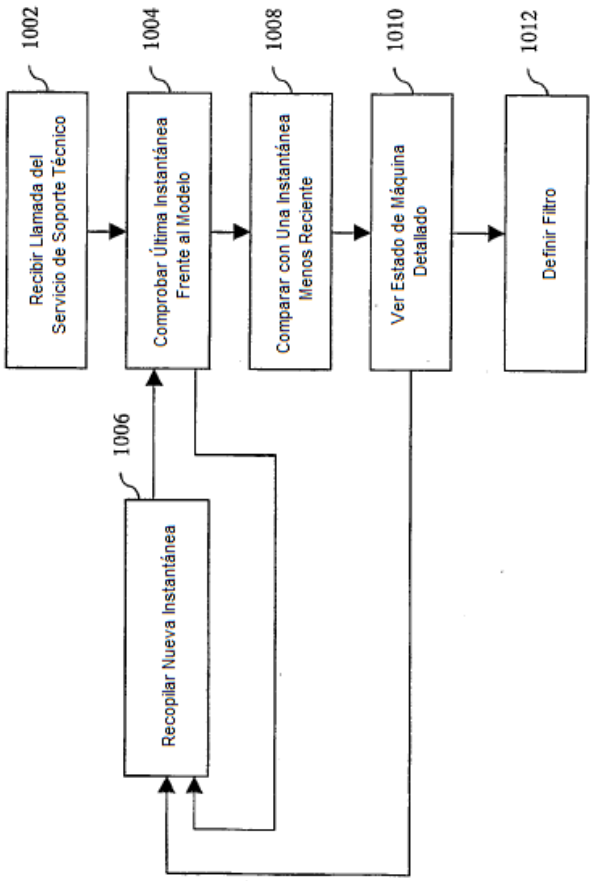
1002
Recibir Llamada del Servicio de Soporte Técnico
1004
Comprobar Última Instantánea Frente al Modelo
1008
Comparar con Una Instantánea Menos Reciente
1010
Ver Estado de Máquina Detallado
1012
Definir Filtro
1006
Recopilar Nueva Instantánea

FIG. 10

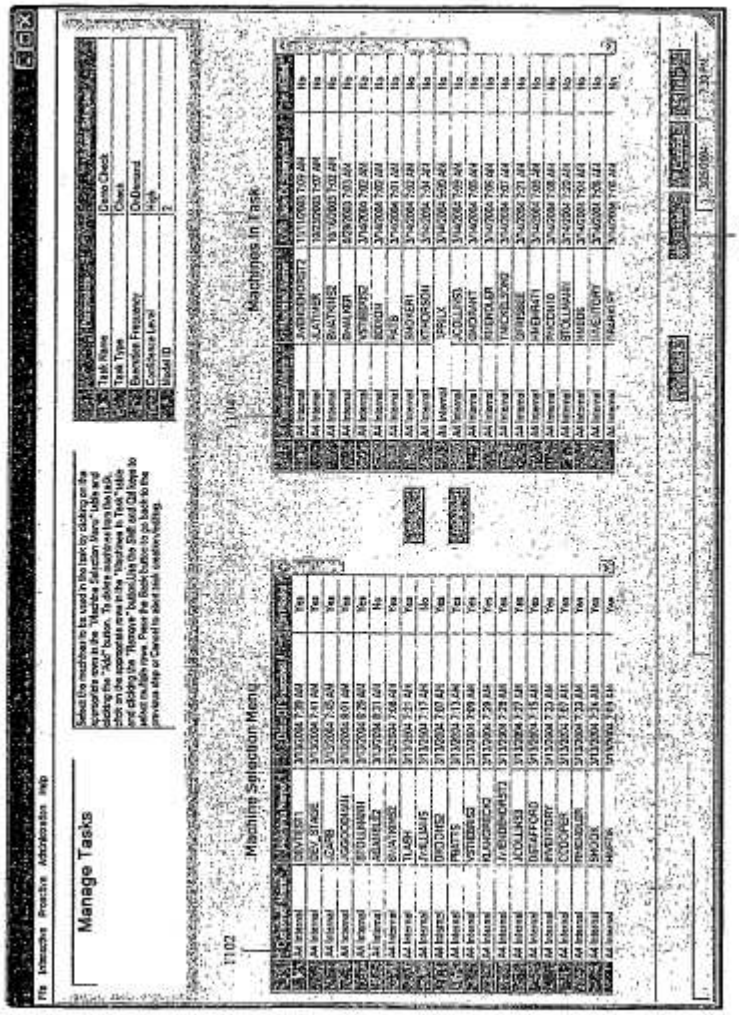
Manage Tasks
Machine Selection Menu
Machines in Task
1102
1104
1106

FIG. 11

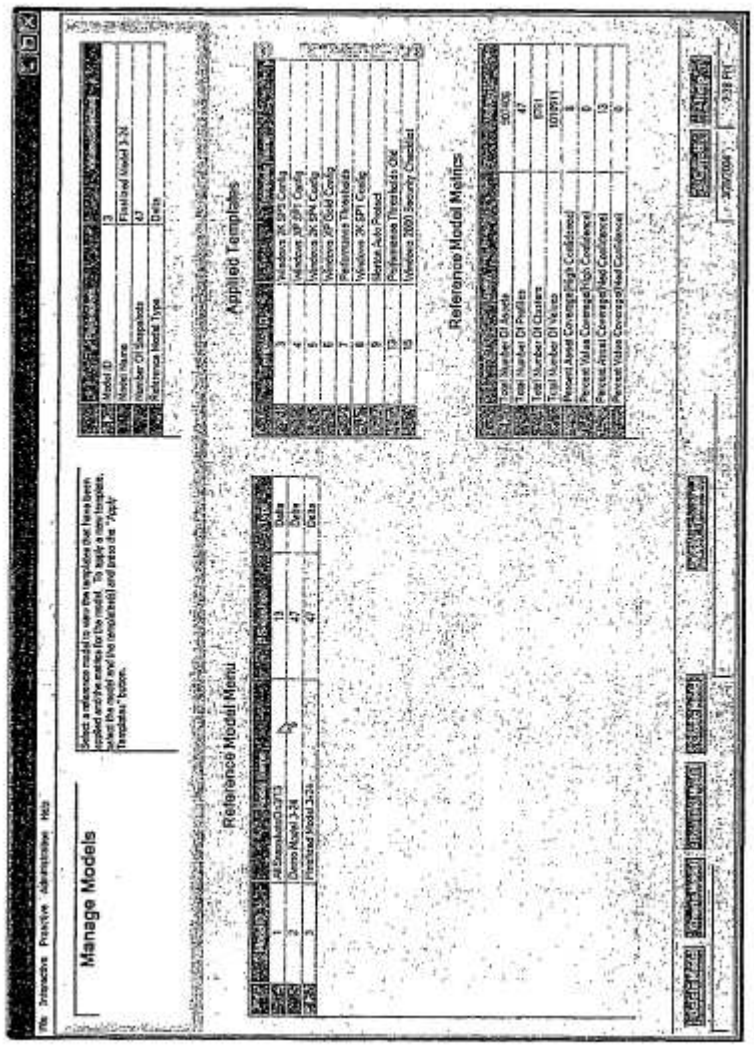
Manage Models
Reference Model Menu
Applied Templates
Reference Model Metrics

FIG. 12

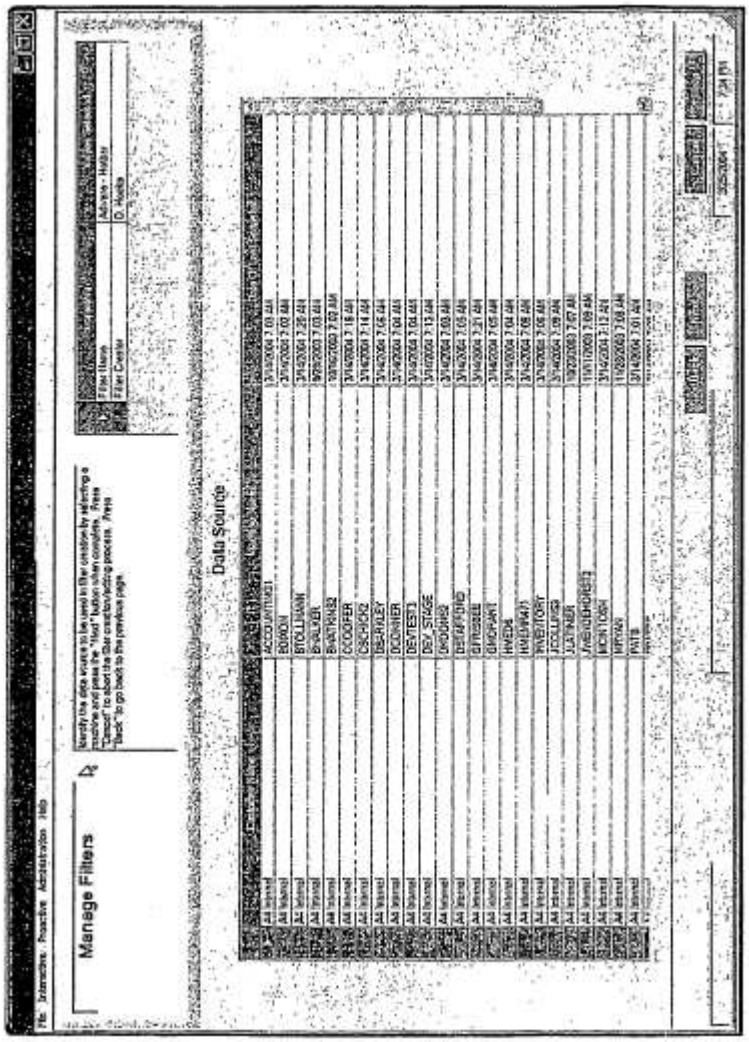

FIG. 13

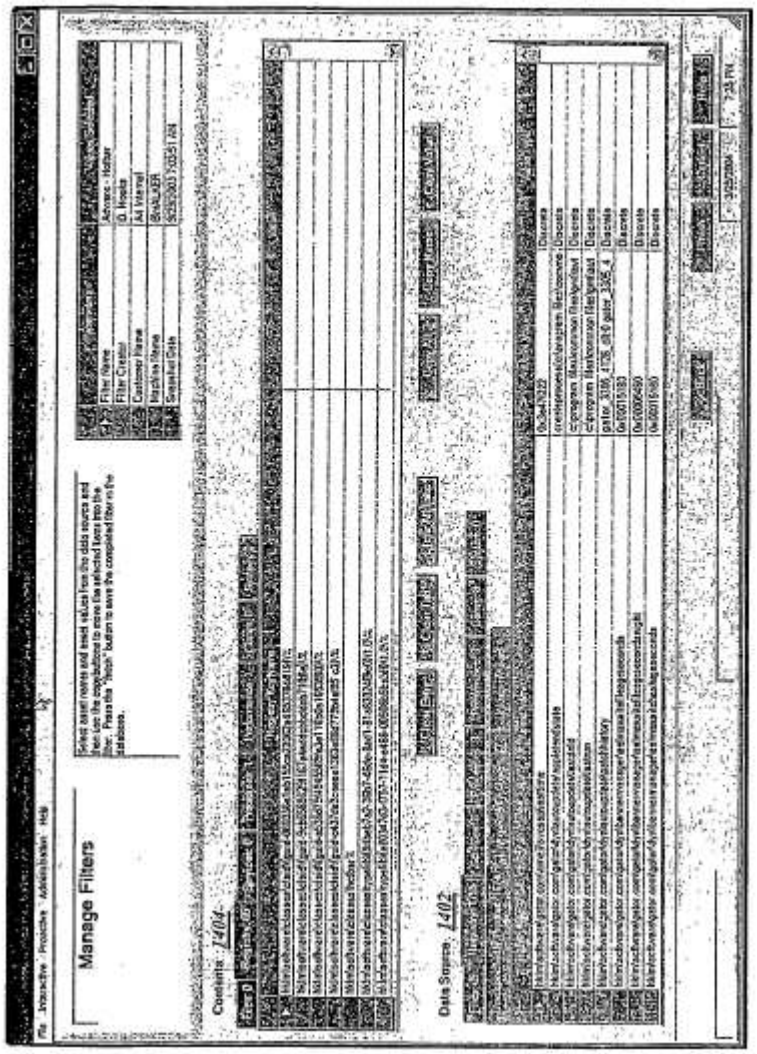

FIG. 14

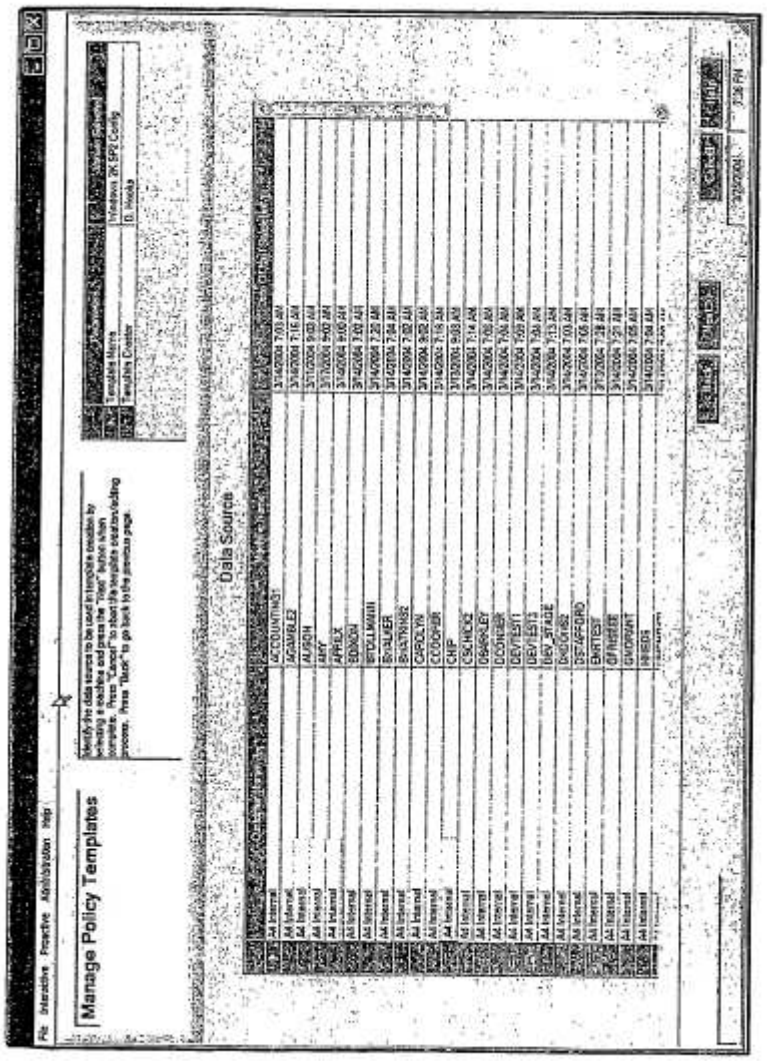

FIG. 15

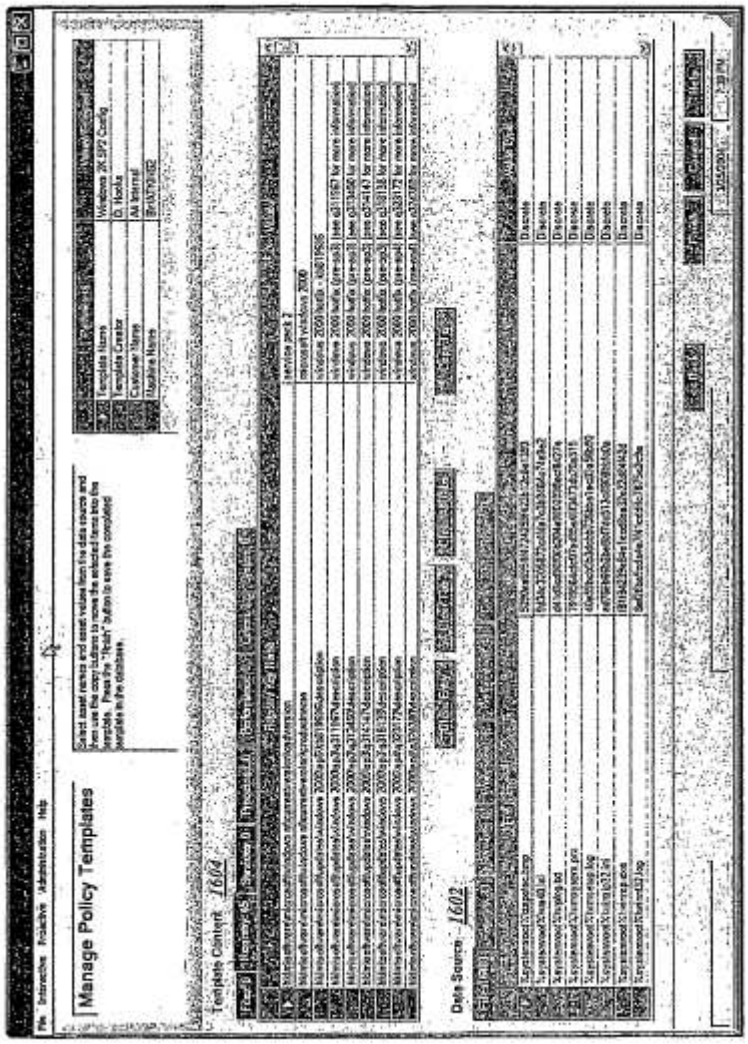
Manage Policy Templates
Template Content
1604
Data Source
1602

FIG. 16