



(10) **DE 11 2020 004 992 B4** 2025.01.02

(12) **Patentschrift**

(21) Deutsches Aktenzeichen: **11 2020 004 992.2**
(86) PCT-Aktenzeichen: **PCT/IB2020/059628**
(87) PCT-Veröffentlichungs-Nr.: **WO 2021/074802**
(86) PCT-Anmeldetag: **14.10.2020**
(87) PCT-Veröffentlichungstag: **22.04.2021**
(43) Veröffentlichungstag der PCT Anmeldung
in deutscher Übersetzung: **30.06.2022**
(45) Veröffentlichungstag
der Patenterteilung: **02.01.2025**

(51) Int Cl.: **G06F 21/56 (2013.01)**

Innerhalb von neun Monaten nach Veröffentlichung der Patenterteilung kann nach § 59 Patentgesetz gegen das Patent Einspruch erhoben werden. Der Einspruch ist schriftlich zu erklären und zu begründen. Innerhalb der Einspruchsfrist ist eine Einspruchsgebühr in Höhe von 200 Euro zu entrichten (§ 6 Patentkostengesetz in Verbindung mit der Anlage zu § 2 Abs. 1 Patentkostengesetz).

(30) Unionspriorität:
16/655,372 **17.10.2019** **US**

(73) Patentinhaber:
**International Business Machines Corporation,
Armonk, NY, US**

(74) Vertreter:
**LifeTech IP Spies & Behrndt Patentanwälte PartG
mbH, 80687 München, DE**

(72) Erfinder:
**Bastide, Paul, Cambridge, MA, US; Ranchal,
Rohit, Austin, TX, US; Khan, Shakil Manzoor, San
Jose, CA, US; Bakthavachalam, Senthil, San Jose,
CA, US**

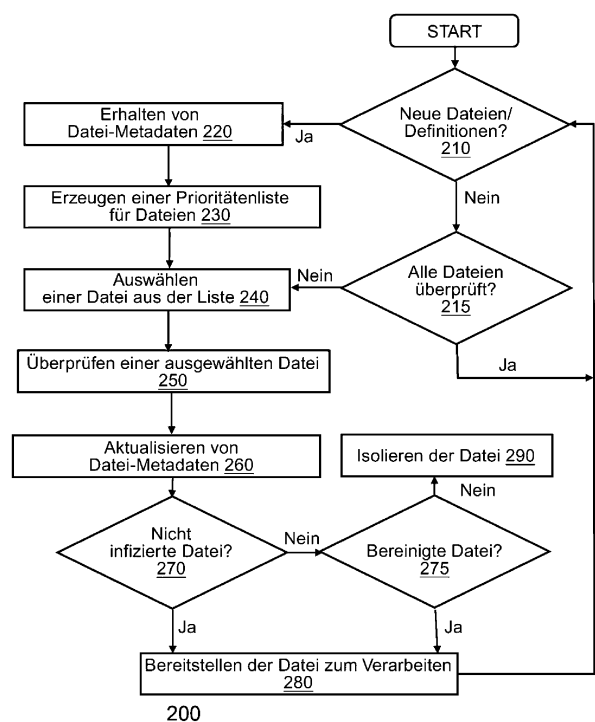
(56) Ermittelter Stand der Technik:

US	8 302 193	B1
EP	2 002 640	B1
EP	2 686 803	B1
WO	2005/ 119 489	A2

(54) Bezeichnung: **AUFRECHTERHALTEN DER SICHERHEIT EINES SYSTEMS**

(57) Hauptanspruch: Durch einen Computer (150) implementiertes Verfahren (200) zum Aufrechterhalten der Sicherheit eines Datenverarbeitungssystems, wobei das Verfahren (200) aufweist:

Empfangen (220) von Metadaten aus einer Metadaten Speichereinheit durch einen oder mehrere Computerprozessoren (154), wobei die Metadaten mindestens einem Teil einer Dateidatei zugehörig sind und wobei die Metadaten einen oder mehrere Indikatoren für Viren aufweisen;
Erzeugen (230) einer Priorität durch den einen oder die mehreren Computerprozessoren für den mindestens einen Teil der Datendatei entsprechend den Metadaten;
Durchführen einer Überprüfung (250) des mindestens einen Teiles der Datendatei durch den einen oder die mehreren Computerprozessoren (154) entsprechend der Priorität;
Aktualisieren (260) der Metadaten des mindestens einen Teiles der Datendatei entsprechend der Überprüfung durch den einen oder die mehreren Computerprozessoren; und
Bereitstellen (280) der Datendatei für eine Weiterverarbeitung durch den einen oder die mehreren Computerprozessoren (154),
wobei den Metadaten eine Kontrollsumme zugeordnet wird, die kontinuierlich parallel zum Strömen der Datendatei in einen Datenspeicher bestimmt wird.



Beschreibung

Technisches Gebiet

[0001] Die Offenbarung betrifft allgemein ein Aufrechterhalten der Sicherheit eines Datenverarbeitungssystems und der Integrität von Daten in einem großen Datenspeichersystem. Die Offenbarung betrifft insbesondere ein Aufrechterhalten von Systemsicherheit und Datenintegrität durch Überprüfen auf böswilligen Code und bösartige Daten.

HINTERGRUND

[0002] Aktuelle Multi-Tenant-Systeme, beispielsweise Systeme im Gesundheitswesen, im Einzelhandel, in der Telekommunikation und im Finanzdienstleistungssektor, können mehrere Anbieter, Kunden und Dienstleistungsorganisationen umfassen, die mit einem einzigen logischen Datenverarbeitungssystem verbunden sind. Das logische Datenverarbeitungssystem kann cloudbasiert oder lokal sein und sowohl kontinuierliche Datenströme als auch Massendatenpakete empfangen, die sich in einem Datenspeicher zu Gigabytes und Terabytes an Daten anhäufen.

[0003] Eingehende Daten müssen auf Viren, Malware und andere bösartige Daten überprüft werden, um die Sicherheit des Systems und seiner Benutzer aufrechtzuerhalten. Vorhandene Daten müssen erneut überprüft werden, da Virusdefinitionen aktualisiert werden, wenn neue Zero-Day-Exploits entdeckt werden. Zum Überwachen eingehender Daten wird ein Abwehrsystem gegen Angriffe von außen (intrusion prevention system, IPS) verwendet. Das IPS kann integriert sein und Echtzeit-Daten überwachen, wenn sie von externen Quellen im Datenstrom zum Speichern an den Datensee übertragen werden, oder es kann extern sein, wobei Kopien der an das IPS weitergeleiteten Daten verarbeitet werden. Eingehende Daten können auch redundante Daten umfassen, die die Arbeitslast des IPS erhöhen, da doppelte Inhalte überprüft werden müssen.

[0004] Die Prüfaktivitäten des IPS müssen ohne Unterbrechung des Kundenzugangs zu den Daten und ohne Unterbrechung der kontinuierlichen Ströme eingehender Daten ausgeführt werden. Die Datenströme müssen ohne Unterbrechung im System landen, um normale Arbeitsgänge fortzusetzen, Kosten für Zugang/Abgang zu senken, eine Verschwendung von System-Ressourcen zu verringern oder zu beseitigen und einen vollständigen Zugang zu Systemdiensten für Benutzer und andere Teile der Systemplattform aufrechtzuerhalten.

[0005] Auf diesem technischen Gebiet gibt es bereits eine Reihe von Veröffentlichungen. Das Dokument EP 2 686 803 B1 beschreibt Systeme und Verfahren zum Nachschlagen von Meta-Daten zum Schutz vor Schadprogrammen. Dazu werden Verfahren und Systeme beschrieben, die vorhersagen, wann Anti-Malware-Meta-Daten benötigt werden. Auf dieser Basis wird dann ein Nachschlagen von Schadprogrammen (Malware) priorisiert. Das Dokument US 8 302 193 B1 beschreibt zusätzlich Verfahren und Systeme zum Durchsuchen von Dateien nach Schadprogrammen. Dazu werden in den Dateien erste Schadprogramm-Kennzeichner ermittelt, die einem ersten Schadprogramm zugeordnet sind. Ähnlich wird mit einem zweiten Schadprogramm-Kennzeichner verfahren. Weiterhin beschreibt das Dokument WO 2005 / 119 489 A2 ein E-Mail-Verteilersystem unter Nutzung von Meta-Daten für E-Mails, um einen virtuellen Speicher zu verwalten. Dabei weisen die Meta-Daten auf einen Status der Verarbeitung von E-Mails hin. Schließlich beschreibt das Dokument EP 2 002 640 B1 eine Transkodierung für eine Datei ein verteiltes Dateisystem. Dazu werden auch Prüfsummen gebildet, welche bestätigen helfen, dass Daten in bestimmten Abschnitten nicht durch zufällige Hardware-Fehler, wie sie in komplexen elektronischen Systemen vorkommen können, korruptiert wurden.

[0006] Trotz dieser bereits gemachten Fortschritte besteht in der Technik weiterhin ein Bedarf dafür, das vorgenannte Problem zu lösen, um einen weiter verbesserten Schutz vor Schad-Software zu bieten.

KURZDARSTELLUNG

[0007] Diese Aufgabe wird durch die Gegenstände der unabhängigen Patentansprüche gelöst. Weitere Ausgestaltungen sind durch die jeweils abhängigen Patentansprüche beschrieben. Unter einem weiteren Aspekt stellt die vorliegende Erfindung ein durch einen Computer implementiertes Verfahren zum Aufrechterhalten der Sicherheit eines Datenverarbeitungssystems bereit, wobei das Verfahren aufweist: Empfangen von Metadaten aus einer Metadaten Speichereinheit durch einen oder mehrere Computerprozessoren, wobei die Metadaten mindestens einem Teil einer Datendatei zugehörig sind; Erzeugen einer Priorität für den mindestens

einen Teil der Datendatei durch den einen oder mehrere Computerprozessoren entsprechend den Metadaten; und Durchführen einer Überprüfung der Datendatei entsprechend der Priorität durch den einen oder mehrere Computerprozessoren; wobei die Metadaten einen oder mehrere Indikatoren für Viren aufweisen.

[0008] Unter einem weiteren Aspekt stellt die vorliegende Erfindung ein durch einen Computer implementiertes Verfahren zum Aufrechterhalten der Sicherheit eines Systems bereit, wobei das Verfahren aufweist: Empfangen einer aktualisierten Virusdefinition durch einen oder mehrere Computerprozessoren; Erhalten von Metadaten, die mindestens einer Datendatei zugehörig sind, durch einen oder mehrere Computerprozessoren aus einer Metadaten Speichereinheit; Erzeugen einer Priorität für die mindestens eine Datendatei durch den einen oder mehrere Computerprozessoren entsprechend den Metadaten und der aktualisierten Virusdefinition; Auswählen einer Datendatei entsprechend der Priorität durch den einen oder mehrere Computerprozessoren; Durchführen einer Überprüfung der Datendatei durch den einen oder mehrere Computerprozessoren; Aktualisieren der Metadaten der Datendatei durch den einen oder mehrere Computerprozessoren entsprechend der Überprüfung; und Bereitstellen der Datendatei durch den einen oder mehrere Computerprozessoren zum weiteren Verarbeiten.

[0009] Unter einem weiteren Aspekt stellt die vorliegende Erfindung ein durch einen Computer implementiertes Verfahren zum Aufrechterhalten der Sicherheit eines Systems bereit, wobei das Verfahren aufweist: Erhalten von Metadaten, die mindestens einer Datendatei zugehörig sind, durch einen oder mehrere Computerprozessoren aus einer Metadaten Speichereinheit; Erzeugen einer Priorität für die mindestens eine Datendatei durch den einen oder mehrere Computerprozessoren entsprechend den Metadaten; Auswählen einer Datendatei entsprechend der Priorität durch den einen oder mehrere Computerprozessoren; Durchführen einer Überprüfung der Datendatei durch den einen oder mehrere Computerprozessoren; Ermitteln einer Bewertung allgemeiner Schwachstellen- und Gefährdungen (common vulnerabilities and exposures, CVE) für die Datendatei durch den einen oder mehrere Computerprozessoren; Aktualisieren der Metadaten der Datendatei durch den einen oder mehrere Computerprozessoren entsprechend der Überprüfung; und Bereitstellen der Bewertung durch den einen oder mehrere Computerprozessoren für einen Systembenutzer zur Auswertung.

[0010] Unter einem weiteren Aspekt stellt die vorliegende Erfindung ein Computersystem zum Aufrechterhalten der Sicherheit eines Systems bereit, wobei das Computersystem aufweist: einen oder mehrere Computerprozessoren; eine oder mehrere durch einen Computer lesbare Speichereinheiten; und gespeicherte Programmanweisungen auf der einen oder mehreren durch einen Computer lesbaren Speichereinheiten zum Ausführen durch den einen oder mehrere Computerprozessoren, wobei die gespeicherten Programmanweisungen aufweisen: Programmanweisungen zum Empfangen von Metadaten, die mindestens einer Datendatei zugehörig sind, aus einer Metadaten Speichereinheit; Programmanweisungen zum Erzeugen einer Priorität für die mindestens eine Datendatei entsprechend den Metadaten; und Programmanweisungen zum Durchführen einer Überprüfung der Datendatei entsprechend der Priorität; wobei die Metadaten einen oder mehrere Indikatoren für Viren aufweisen.

[0011] Unter einem weiteren Aspekt stellt die vorliegende Erfindung ein Computerprogrammprodukt zum Aufrechterhalten der Sicherheit eines Systems bereit, wobei das Computerprogrammprodukt ein durch einen Computer lesbares Speichermedium aufweist, das von einer Verarbeitungsschaltung gelesen werden kann und Anweisungen zum Ausführen durch die Verarbeitungsschaltung speichert, um ein Verfahren zum Ausführen der Schritte der Erfindung auszuführen.

[0012] Unter einem weiteren Aspekt stellt die vorliegende Erfindung ein Computerprogramm bereit, das auf einem durch einen Computer lesbaren Medium gespeichert ist und in den internen Speicher eines digitalen Computers geladen werden kann, wobei das Computerprogramm Software-Codeteile aufweist, um beim Ausführen des Programms auf einem Computer die Schritte der Erfindung auszuführen.

[0013] Unter einem weiteren Aspekt stellt die vorliegende Erfindung ein Computerprogrammprodukt zum Aufrechterhalten der Sicherheit eines Systems bereit, wobei das Computerprogrammprodukt eine oder mehrere durch einen Computer lesbare Speichereinheiten und gespeicherte Programmanweisungen auf der einen oder mehreren durch einen Computer lesbaren Speichereinheiten aufweist, wobei die gespeicherten Programmanweisungen aufweisen: Programmanweisungen zum Empfangen von Metadaten, die mindestens einer Datendatei zugehörig sind, aus einer Metadaten Speichereinheit; Programmanweisungen zum Erzeugen einer Priorität für die mindestens eine Datendatei entsprechend den Metadaten; und Programmanweisungen zum Ausführen einer Überprüfung der Datendatei entsprechend der Priorität; wobei die Metadaten einen oder mehrere Indikatoren für Viren aufweisen.

[0014] Im Folgenden wird eine Kurzdarstellung dargelegt, um ein grundlegendes Verständnis einer oder mehrerer Ausführungsformen der Offenbarung zu vermitteln. Diese Kurzdarstellung ist nicht dazu gedacht, wichtige oder kritische Elemente zu kennzeichnen oder einen Schutzbereich der einzelnen Ausführungsformen oder einen Schutzbereich der Ansprüche abzugrenzen. Ihr einziger Zweck besteht darin, Konzepte in vereinfachter Form als Auftakt zu der später folgenden ausführlicheren Beschreibung darzustellen. In einer oder mehreren hierin beschriebenen Ausführungsformen ermöglichen Einheiten, Systeme, durch einen Computer implementierte Verfahren, Vorrichtungen und/oder Computerprogrammprodukte ein Priorisieren und Überprüfen von Datendateien auf bösartigen Code, bevor Benutzern ein Zugreifen auf die Daten gestattet wird.

[0015] Unter einem Aspekt umfasst die Erfindung Verfahren, Systeme und durch einen Computer lesbare Medien im Zusammenhang mit einem Aufrechterhalten der Sicherheit eines Systems durch kontinuierliches Überwachen von Datendateien, wenn diese im Datenstrom oder in großen Mengen empfangen werden, bevor die Dateien verarbeitet werden oder anderweitig von Benutzern darauf zugegriffen wird. Unter diesem Aspekt gehen die Systeme, Verfahren und Computerprogrammprodukte so vor, dass sie Metadaten, die mindestens einem Teil einer Datendatei zugehörig sind, aus einer Metadaten Speichereinheit empfangen, eine Priorität für den mindestens einen Teil einer Datendatei entsprechend den Metadaten erzeugen und eine Überprüfung des Teils der Datendatei durchführen. Die Metadaten umfassen einen oder mehrere Indikatoren für Viren. Nach dem Empfangen werden die Daten Benutzern ohne Unterbrechung zur Verfügung gestellt.

[0016] Unter einem Aspekt umfasst die Erfindung Verfahren, Systeme und durch einen Computer lesbare Medien im Zusammenhang mit einem Aufrechterhalten der Sicherheit eines Systems, da sich Virusdefinitionen im Laufe der Zeit ändern. Unter diesem Aspekt gehen die Systeme, Verfahren und Computerprogrammprodukte vor, indem sie eine aktualisierte Virusdefinition empfangen, Metadaten, die mindestens einer Datendatei zugehörig sind, aus einer Metadaten Speichereinheit erhalten, eine Priorität für die mindestens eine Datendatei entsprechend den Metadaten und der aktualisierten Virusdefinition erzeugen, eine Datendatei entsprechend der erzeugten Priorität auswählen, eine Überprüfung der Datendatei durchführen, die Metadaten der Datendatei entsprechend der Überprüfung aktualisieren und die Datendatei zum weiteren Verarbeiten bereitstellen.

[0017] Unter einem Aspekt umfasst die Erfindung Verfahren, Systeme und durch einen Computer lesbare Medien im Zusammenhang mit einem Aufrechterhalten der Sicherheit eines Systems mit dem zusätzlichen Vorteil, dass sie einem Benutzer oder einem automatisierten Empfangssystem die Bewertung allgemeiner Schwachstellen und Gefährdungen für jede überprüfte Datei bereitstellen. Der Benutzer oder das Empfangssystem ermittelt sodann, ob die überprüfte Datei zum weiteren Verarbeiten und Verwenden in ein Datenverarbeitungssystem geladen werden soll. Die Erfindungen unter diesem Aspekt gehen so vor, dass Metadaten, die mindestens einer Datendatei zugehörig sind, aus einer Metadaten Speichereinheit erhalten werden, eine Priorität für die mindestens eine Datendatei entsprechend den Metadaten erzeugt wird, eine Datendatei entsprechend der Priorität ausgewählt wird, eine Überprüfung der Datendatei durchgeführt wird, eine Bewertung allgemeiner Schwachstellen und Gefährdungen für die Datendatei ermittelt wird, die Metadaten der Datendatei entsprechend der Überprüfung aktualisiert werden und die Bewertung einem Systembenutzer zur Auswertung bereitgestellt wird.

KURZE BESCHREIBUNG DER ZEICHNUNGEN

[0018] Durch die ausführlichere Beschreibung einiger Ausführungsformen der vorliegenden Offenbarung in den beigefügten Zeichnungen werden die obigen und andere Aufgaben, Merkmale und Vorteile der vorliegenden Offenbarung offensichtlicher, wobei sich die gleiche Bezugszahl im Allgemeinen auf die gleichen Komponenten in den Ausführungsformen der vorliegenden Offenbarung bezieht.

Fig. 1 stellt eine schematische Veranschaulichung eines Systems gemäß einer Ausführungsform der Erfindung dar.

Fig. 2 stellt einen Ablaufplan dar, der eine Verarbeitungsfolge gemäß einer Ausführungsform der Erfindung zeigt.

Fig. 3 zeigt eine Cloud-Computing-Umgebung gemäß einer Ausführungsform der Erfindung.

Fig. 4 zeigt Abstraktionsmodellsschichten gemäß einer Ausführungsform der Erfindung.

AUSFÜHRLICHE BESCHREIBUNG

[0019] Einige Ausführungsformen werden unter Bezugnahme auf die beigelegten Zeichnungen ausführlicher beschrieben, in denen die Ausführungsformen der vorliegenden Offenbarung veranschaulicht werden. Die vorliegende Offenbarung kann jedoch auf verschiedene Weise umgesetzt werden und sollte daher nicht so ausgelegt werden, dass sie auf die hierin offenbarten Ausführungsformen beschränkt ist.

[0020] In einer Ausführungsform können eine oder mehrere Komponenten des Systems Hardware und/oder Software einsetzen, um hochtechnische Probleme zu lösen (z.B. ein Priorisieren von Dateien anhand von Metadaten und ein Überprüfen priorisierter Dateien usw.). Diese Lösungen sind nicht abstrakt und können aufgrund der Verarbeitungsfähigkeiten, die beispielsweise zum Ermöglichen einer Code-Überprüfung und -Fehlerbehebung erforderlich sind, nicht als eine Reihe von geistigen Handlungen eines Menschen ausgeführt werden. Darüber hinaus können einige der ausgeführten Prozesse von einem spezialisierten Computer ausgeführt werden, um bestimmte Aufgaben im Zusammenhang mit Speichervorgängen durchzuführen. Zum Beispiel kann ein spezialisierter Computer eingesetzt werden, um Aufgaben im Zusammenhang mit einer Code-Überprüfung und -Fehlerbehebung und dergleichen auszuführen.

[0021] Bei verteilten Big-Data-Verarbeitungssystemen stellen sich besondere Probleme der Systemsicherheit. Diese Systeme empfangen und verteilen sehr große Datendateien, die in Gigabyte und Terabyte gemessen werden. Die empfangenen Daten stammen aus einer Vielzahl von Quellen, von denen einige zuverlässiger sind als andere und von denen einige möglicherweise nicht überprüfen oder sicherstellen, dass die bereitgestellten Daten frei von Viren, Malware und anderen eingebetteten Problemen mit Daten sind. Systembenutzer benötigen einen ununterbrochenen Zugriff auf Systemdaten in einer verteilten Umgebung, in der ein Anhalten von Diensten oder Datenzugriff zum Überprüfen oder Beheben von Datenproblemen keine praktikable Option ist. Systembenutzer können Datenblöcke zum weiteren Verarbeiten und Verwenden herunterladen. Eine Angabe des Sicherheitsprofils der Daten hilft Benutzern beim Ermitteln, welche Blöcke sicher sind und einem Datenverarbeitungssystem zum Verwenden hinzugefügt werden können. Solche Systeme können auch doppelte Dateien aus einer einzigen Quelle oder aus mehreren Quellen empfangen. Ein Verarbeiten doppelter Datensätze aus Sicherheitsgründen vergeudet knappe Systemressourcen.

[0022] In einer Ausführungsform umfasst das System einen Datenspeicher zum Speichern großer (Gigabyte und Terabyte großer) Dateien. Das System umfasst außerdem eine Metadatenspeichereinheit, die eine oder mehrere Metadatentabellen zwischenspeichert. Die Metadatenspeichereinheit kann lokal oder cloudbasiert sein. Die Tabellen umfassen Metadatenätze für die großen Dateien, die im Datenspeicher zwischengespeichert sind. In dieser Ausführungsform speichern die Tabellen Metadatenfelder, darunter einige oder alle der folgenden: Daten in Bezug auf eine Dateikennung, Status der Dateiverarbeitung - ist die Datei für nachgelagerte Verwendungen verfügbar usw., Zeitpunkt der Dateiverarbeitung - einen Eintrag des letzten Zeitpunkts, zu dem die Daten durch das Verfahren verarbeitet wurden, eine CVE-Bewertung für die Datei, in der Datei erkannte Viren, ausführliche Daten zur Virusbeschreibung, Informationen zum Dateiformat und Daten in Bezug auf nachgelagerte Verarbeitungsanforderungen für die Datei. Verschiedene Dateiformate sind enger korreliert mit datengebundenen Problemen. Dateien, die für eine nachgelagerte Verarbeitung vorgesehen sind, erfordern im Gegensatz zu einer langfristigen Speicherung mehr Aufmerksamkeit und ein sofortiges Überprüfen.

[0023] Die offenbarten Erfindungen sind auf einzelne oder mehrere Anwendungssysteme anwendbar, die fest auf einem einzigen Server installiert oder über mehrere Cloud-Mietverträge verteilt sind. Die Erfindungen sind anwendbar auf eine nahezu in Echtzeit erfolgende Datenverarbeitung von Datenströmen sowie auf eine Stapel- und Massendatenverarbeitung.

[0024] In einigen Ausführungsformen umfassen die Metadaten auch Kontrollsummendaten, die den Datendateien zugehörig sind. Für Stapel- oder Massendatendateien, die extern empfangen werden, berechnet das Verfahren eine Kontrollsumme, beispielsweise den Hash-Wert (SHA256), für die Datei. Selbst geringfügige Änderungen an der Datei führen zu einem anderen Kontrollsummenwert, was darauf hinweist, dass die Datei geändert wurde. Die Kontrollsumme wird als der Datei zugehörige Metadaten gespeichert. In einigen Ausführungsformen kann die Datei mit einem Kontrollsummenwert versehen werden. In diesen Ausführungsformen vergleicht das Verfahren die berechnete Kontrollsumme mit dem bereitgestellten Wert, um zu ermitteln, ob an der Datei Änderungen vorgenommen wurden. Die Kontrollsumme kann eine Kontrollsummenteknik wie etwa SHA256, MD5 oder ein anderes Verfahren zum Berechnen einer Kontrollsumme unter Verwenden der Byte-Darstellung der Datendatei verwenden.

[0025] In einer Ausführungsform berechnet das Verfahren Kontrollsummenwerte für Datenstromdateien sowie für extern empfangene Massendateien. Für Datenströme berechnet das Verfahren kontinuierlich Kontrollsummenwerte parallel zu den im Datenspeicher des Systems ankommenden Datenströmen. Das Verfahren speichert die während des Streaming berechneten Kontrollsummenwerte als Metadaten in einer Nachschlagetabelle. Nachdem die Datendatei vollständig angekommen ist, wird der endgültige Kontrollsummenwert für die Datei in Verbindung mit der Streaming-Datei in der Metadaten Speichereinheit gespeichert. In einer Ausführungsform unterteilt das Verfahren die Datendatei in Teile wie Blöcke oder Segmente der Datei. In dieser Ausführungsform berechnet das Verfahren eine Kontrollsumme für jeden Teil der Datendatei. Wenn jeder Teil vollständig empfangen worden ist, wird die Kontrollsumme für jeden Teil der Datendatei in Form von Metadaten gespeichert, die diesem Teil der Datei zugehörig sind. In einer Ausführungsform überlappen sich die Teile, um sicherzustellen, dass Viren nicht zwischen benachbarten Blöcken aufgeteilt und übersehen werden. Segmente können so aufgefüllt werden, dass jedes Segment einen Teil des vorherigen Segments umfasst, zum Beispiel kann jedes 100-Megabyte-Segment so aufgefüllt werden, dass es die letzten 100 Kilobyte des vorherigen Segments umfasst. In einer Ausführungsform werden 100-Megabyte-Segmente um +/- 500 Kilobyte aufgefüllt, um sicherzustellen, dass aufgrund des Segmentierens der Datendatei beim Überprüfen der Segmente keine Lücken entstehen. In dieser Ausführungsform würden die Segmente zwischen 0 und 1,5 MB, 1 und 2,5 MB, 2 und 3,5 MB usw. liegen, wodurch die Wahrscheinlichkeit verringert wird, dass ein Teil von bösartigem Code einer Entdeckung entgeht, indem er zwischen zwei benachbarten Segmenten aufgeteilt wird.

[0026] In einer Ausführungsform vergleicht das Verfahren die Kontrollsummenwerte mit gespeicherten Metadatenwerten, die vorhandenen Datendateien und Segmenten zugehörig sind. Durch den Vergleich werden Duplikate der aktuellen Stapel- oder Streaming-Datei/des aktuellen Segments erkannt. Das Verfahren vergleicht fortlaufend berechnete Kontrollsummen, wenn Dateien in den Datenspeicher gestreamt werden, um potenziell ähnliche oder identische Dateien zu erkennen. In dieser Ausführungsform verwendet das Verfahren die endgültige Kontrollsumme, um alle potenziell identischen Dateien als tatsächlich identisch zu bestätigen.

[0027] In einer Ausführungsform leitet das Verfahren als Reaktion auf ein Abgleichen einer Kontrollsumme in den Metadaten eine Bereinigungsroutine für die Datei oder das Dateisegment auf der Grundlage der CVE-Bewertung und Virendetails der abgeglichenen Datei/des abgeglichenen Segments ein. In dieser Ausführungsform sind die abgeglichenen Metadaten einer Datei mit bekannten Viren zugehörig. In dieser Ausführungsform bereinigt das Verfahren die Datei oder das Segment und aktualisiert die Metadaten, um anzuzeigen, dass die Bereinigung stattgefunden hat. Ein Bereinigen der Dateien kann ein Entfernen bekannter bösartiger Codeteile, die einem erkannten Virus zugehörig sind, aus der Datei oder ein Wiederherstellen der Datei unter Verwenden einer Sicherungskopie der Datei umfassen. Die bereinigte Datei wird einer Liste für eine Priorisierung hinzugefügt. In dieser Ausführungsform berechnet das Verfahren eine neue Kontrollsumme für die bereinigte Datei und fügt die neue Kontrollsumme den Metadaten hinzu.

[0028] In einer Ausführungsform überprüft das Verfahren die Metadaten Speichereinheit, um zu ermitteln, ob neue Datenspeicher-Stapeldateien geladen wurden, oder um zu ermitteln, ob neue Streaming-Datendateien geladen werden. In dieser Ausführungsform überprüft das Verfahren die allen neuen Dateien zugehörigen Metadaten. Das Verfahren überwacht dann kontinuierlich die Metadaten Speichereinheit auf Hinweise auf neue Datendateien. In einer Ausführungsform ruhen das Verfahren und die Systeme der Erfindung, wenn keine neuen Dateien oder Virusdefinitionen erkannt werden.

[0029] In einer Ausführungsform erhält das Verfahren die Metadaten für alle kürzlich geladenen oder aktuellen Streaming-Dateien. Das Verfahren erstellt eine priorisierte Liste neuer Datendateien unter Verwenden der Metadaten. Die Dateien werden im Hinblick auf Indikatoren für Viren oder auf die Wahrscheinlichkeit, dass eine Datei Malware enthält, priorisiert. Das Verfahren korreliert die Metadaten aller neuen Dateien, um die priorisierte Liste zu erstellen. Das Verfahren berücksichtigt Kontrollsummenwerte bezüglich zuvor geprüfter Dateien im Hinblick auf zuvor gefundene Viren. Das Verfahren berücksichtigt auch die Virusdefinition, die bei einer früheren Überprüfung einer Datei oder eines Duplikats verwendet wurde, und ob die Virusdefinition auf dem neuesten Stand ist. Überprüfungen, die mit einer veralteten Virusdefinition durchgeführt wurden, erhalten eine höhere Priorität für eine erneute Überprüfung. Datendateien, die ein nachgelagertes Verarbeiten benötigen, wird eine höhere Priorität zugeteilt als Dateien, die dies nicht benötigen. Dateien mit einer Kontrollsummenabweichung im Vergleich zu einer bereitgestellten Kontrollsumme oder einer Kontrollsumme aus einer früheren Überprüfung wird eine höhere Priorität für eine erneute Überprüfung zugeteilt. In dieser Ausführungsform berücksichtigt das Verfahren die Metadaten des Satzes neuer und vorhandener Dateien im Hinblick auf etwaige Änderungen an den Virusdefinitionen. Das Verfahren erstellt eine Rangfolge des Satzes

von Datendateien und priorisiert die Dateien von der höchsten (am meisten eine Überprüfung benötigend) zur niedrigsten (am wenigsten eine Überprüfung benötigend). In dieser Ausführungsform überprüft das Verfahren neue, unverarbeitete Dateien - solche ohne übereinstimmende Dateikennung oder Dateikontrollsumme, stuft diese Dateien mit hoher Priorität ein und erstellt eine neue Metadatenliste in der Metadaten Speichereinheit für die Dateien. Das Verfahren blockiert außerdem einen Benutzerzugriff auf unverarbeitete Dateien zumindest so lange, bis eine Überprüfung der Dateien abgeschlossen ist. In dieser Ausführungsform betrachtet das Verfahren doppelte Dateien, die dieselbe Dateikennung und Dateikontrollsumme wie eine vorhandene Datei aufweisen, als niedrig priorisiert.

[0030] In einer Ausführungsform wählt das Verfahren eine Datei aus der erzeugten priorisierten Liste aus und überprüft die Datei auf Viren, Malware und andere datengebundene Probleme. In dieser Ausführungsform wählt das Verfahren die Datendatei entsprechend ihrer Priorität aus, beginnend mit der höchsten Priorität. Jede Datei der Liste wird der Reihe nach entsprechend einer absteigenden Priorität geprüft. In dieser Ausführungsform wählt das Verfahren die Datei mit der höchsten Priorität auf der Liste aus. Dateien werden zusammen mit ihrer Priorität der Liste hinzugefügt, so dass eine neue Datei mit hoher Priorität vor einer älteren Datei mit niedriger Priorität geprüft werden kann.

[0031] In einer Ausführungsform überprüft das Verfahren die Dateien auf Viren unter Verwenden der detaillierten aktuellen Virusdefinitionen. In dieser Ausführungsform ermittelt das Verfahren eine CVE-Bewertung für jede Datei sowie eine Liste gefundener Viren für jede Datei. Der Zeitpunkt der Überprüfung, die verwendete Virusdefinition, die gefundenen Viren, die Anzahl der gefundenen Viren, die Dateikontrollsumme, das Dateiformat und der Datei zugehörige nachgelagerte Verarbeitungsanforderungen werden alle als Metadaten für die Datei in der Metadaten Speichereinheit aktualisiert.

[0032] In einer Ausführungsform werden Dateien in Blöcke oder Segmente unterteilt, um ein schnelleres Überprüfen zu ermöglichen und eine Dateispeicherung zu optimieren. In dieser Ausführungsform ermittelt und speichert das Verfahren Metadaten für jedes Segment sowie für die gesamte Datei. In dieser Ausführungsform schließt das Verfahren zusammengehörige Segmentdaten als Teil der Metadaten ein. Ein Überprüfen eines Segments einer Gruppe zusammengehöriger Segmente löst ein Überprüfen der benachbarten zusammengehörigen Segmente aus. In dieser Ausführungsform überschneiden sich die Überprüfungen der zusammengehörigen Segmente, was die Wahrscheinlichkeit erhöht, Viren und Malware zu finden. In dieser Ausführungsform kann die Datei als eine Reihe von logisch zusammengehörigen Segmenten - Segment0, Segment1, Segment2 usw. - unterteilt, gespeichert und überprüft werden. In einer Ausführungsform überprüft das Verfahren binäre Versionen von Dateien. In dieser Ausführungsform behält das Verfahren die Dateien so lange als Binärdateien bei, bis die Überprüfung abgeschlossen ist und die Datei als Teil des verfügbaren Datenspeichers für ein nachgelagertes Verarbeiten und einen Benutzerzugriff geeignet ist. In dieser Ausführungsform verwendet das Verfahren Extrahieren-Transformieren-Laden oder einen anderen Vorgang eines Datenverarbeitungsmittels, um die Daten aus der Datei zum Verwenden als Teil des Datenspeichers abzurufen.

[0033] In einer Ausführungsform wertet das Verfahren nach Überprüfen einer Datei oder eines Dateisegments die Metadaten aus. Das Verfahren stellt die Datei sodann zum nachgelagerten Verarbeiten bereit, überträgt sie an eine Viren-/Malware-Bereinigungsroutine oder isoliert die Datei, wenn keine Bereinigungsroutinen verfügbar sind. In einer Ausführungsform ermöglicht das Verfahren Benutzern einen Zugriff auf die CVE-Bewertung jeder Datei. In dieser Ausführungsform bestimmt der Benutzer ein akzeptables Risikoniveau in Verbindung mit CVE-Bewertungen und legt einen Schwellenwert für CVE-Bewertungen fest. Dateien mit einer CVE-Bewertung unter dem Schwellenwert werden in das Datenverarbeitungssystem des Benutzers zum Verwenden durch Datenabnehmer aufgenommen. Dateien mit Bewertungen oberhalb des Schwellenwerts werden abgewiesen. In einer Ausführungsform lädt das Verfahren Dateien ohne gefundene Viren oder mit akzeptabel niedrigen CVE-Bewertungen zum nachgelagerten Verwenden hoch. Das Verfahren aktualisiert die Metadaten in der Metadaten Speichereinheit für alle geprüften Dateien.

[0034] Hier ein Beispiel:

Eine Client-Anwendung stellt eine Verbindung mit einer Verfahrrensplattform unter Verwenden einer gegenseitigen Authentifizierung über Transport Layer Security (Transport Layer Security Mutual Authentication) her.

[0035] Die Client-Anwendung beginnt mit dem Hochladen von Gesundheitsdaten.

[0036] Der Client hat die Größe des Inhalts nicht angegeben, und die Plattform unterstützt maximal Dateien mit einer Größe von mehreren Terabyte.

[0037] Die Plattform streamt die Daten in den Datenspeicher und startet die Kontrollsumme unter Verwenden von SHA256.

[0038] Die Plattform hat die Daten noch nicht zum Herunterladen markiert, und die nachgeschalteten Abnehmer können die neuen Daten noch nicht nutzen.

[0039] Die Plattform beginnt mit dem Bilden einer Kontrollsumme für die Teile der Datei. {Teil1: 0 bis 1M, Teil 2: 1 bis 2M, Teil 3: 2 bis 3M}.

[0040] Die Client-Anwendung beendet das Hochladen der Datei.

[0041] Die Plattform aktualisiert die Kontrollsumme - 01234567A in der Spalte für die Metadatenunterstützung zusammen mit den berechneten Kontrollsummen für Teile von Dateien.

[0042] Parallel dazu überprüft der Virenprüfdienst der Plattform die Metadatenpalte.

[0043] Die Metadaten enthalten die folgenden Informationen für die hochgeladene Datei:

DATEIKENNUNG	1-2-1-1
VERARBEITET	N
Kontrollsumme	01234567A
VERARBEITUNGSZEIT	NULL
CVE-Bewertung	NULL
Virus und Virendetails	NULL

Zusammengehörige Segmente (vorherige/nächste)	NULL
---	------

[0044] Eine Verfahrenskoordinations-einrichtung überprüft, ob die Kontrollsumme bereits vorhanden ist. Ist dies der Fall, priorisiert die Erfindung diejenigen, die keine relevanten oder ähnlichen Kontrollsummen haben.

[0045] Wenn die Koordinations-einrichtung keine relevanten Kontrollsummen findet, extrahiert sie die Datei auf der Grundlage der Zeit.

[0046] Die Koordinations-einrichtung überprüft die Datei und gibt die Ergebnisse zurück - gefundener Virus, Virus-Metadaten und markiert die Datei als verarbeitet, mit Virusinhalt und CVE-Bewertung.

DATEIKENNUNG	1-2-1-1
VERARBEITET	N
Kontrollsumme	01234567A
VERARBEITUNGSZEIT	2019-02-01 07:56:00AM UTC
CVE-Bewertung	CVSS v2.0 Schweregrad und Maßzahlen: Basisbewertung: 5.0 MEDIUM Vektor: (AV:N/AC:L/Au:N/C:N/I:P/A:N) (V2 legend) Auswirkung Teilbewertung: 2.9 Verwertbarkeit Teilbewertung: 10.0
Virusbewertung und -inhalt	Eicar
Zusammengehörige Segmente (vorherige/nächste)	Keine

[0047] Fig. 1 stellt eine schematische Darstellung von beispielhaften Netzwerkressourcen bereit, die einem Ausführen der offenbarten Erfindungen zugehörig sind. Die Erfindungen können in den Prozessoren eines beliebigen der offenbarten Elemente ausgeführt werden, die einen Befehlsstrom verarbeiten. Wie in der Figur gezeigt wird, ist eine vernetzte Client-Einheit 110 drahtlos mit einem Server-Teilsystem 102 verbunden. Eine Client-Einheit 104 ist drahtlos über ein Netzwerk 114 mit dem Server-Teilsystem 102 verbunden. Die Client-Einheiten 104 und 110 weisen Clients für Virenprüfprogramme (nicht gezeigt) zusammen mit ausreichenden Datenverarbeitungsressourcen (Prozessor, Speicher, Netzwerkdatenübertragungs-Hardware) zum Ausführen des Programms auf. Wie in Fig. 1 gezeigt wird, weist das Server-Teilsystem 102 einen Server-Computer 150 auf. Fig. 1 zeigt ein Blockschaubild von Komponenten des Server-Computers 150 in einem vernetzten Computersystem 1000 gemäß einer Ausführungsform der vorliegenden Erfindung. Es sollte klar sein, dass Fig. 1 nur eine Veranschaulichung einer Implementierung darstellt und keine Einschränkungen in Bezug auf die Umgebungen andeutet, in denen verschiedene Ausführungsformen implementiert werden können. Es können viele Modifikationen an der dargestellten Umgebung vorgenommen werden.

[0048] Der Servercomputer 150 kann einen oder mehrere Prozessoren 154, einen Speicher 158, einen dauerhaften Speicher 170, eine Datenübertragungseinheit 152, eine oder mehrere Ein/Ausgabe- (E/A)-Schnittstellen 156 und eine Datenübertragungsstruktur 140 umfassen. Die Datenübertragungsstruktur 140 stellt Datenübertragungsvorgänge zwischen einem Cache 162, dem Speicher 158, dem dauerhaften Speicher 170, der Datenübertragungseinheit 152 und der/den Ein/Ausgabe- (E/A)-Schnittstelle(n) 156 bereit. Die Datenübertragungsstruktur 140 kann mit jeder Architektur implementiert werden, die für ein Weiterleiten von Daten und/oder Steuerinformationen zwischen Prozessoren (wie etwa Mikroprozessoren, Datenübertragungs- und Netzwerkprozessoren usw.), Systemspeicher, peripheren Einheiten und beliebigen anderen Hardware-Komponenten innerhalb eines Systems ausgelegt ist. Die Datenübertragungsstruktur 140 kann zum Beispiel mit einem oder mehreren Bussen implementiert werden.

[0049] Bei dem Speicher 158 und dem dauerhaften Speicher 170 handelt es sich um durch einen Computer lesbare Speichermedien. In dieser Ausführungsform umfasst der Speicher 158 einen Direktzugriffsspeicher (RAM) 160. Im Allgemeinen kann der Speicher 158 jedes geeignete flüchtige oder nichtflüchtige durch einen Computer lesbare Speichermedium umfassen. Der Cache 162 ist ein schneller Speicher, der die Leistung des Prozessors/der Prozessoren 154 verbessert, indem er Daten, auf die kürzlich zugegriffen wurde, und Daten in der Nähe von Daten, auf die kürzlich zugegriffen wurde, aus dem Speicher 158 zwischenspeichert.

[0050] Programmanweisungen und Daten, die zum Ausführen von Ausführungsformen der vorliegenden Erfindung verwendet werden, z.B. das Virenprüfprogramm 175, werden im dauerhaften Speicher 170 zur Ausführung und/oder zum Zugriff durch einen oder mehrere der jeweiligen Prozessoren 154 des Server-Computers 150 über den Cache 162 gespeichert. In dieser Ausführungsform umfasst der dauerhafte Speicher 170 ein magnetisches Festplattenlaufwerk. Alternativ oder zusätzlich zu einem magnetischen Festplattenlaufwerk kann der dauerhafte Speicher 170 einen Halbleiterdatenträger, eine Halbleiterspeichereinheit, einen Nur-Lese-Speicher (ROM), einen löschbaren programmierbaren Nur-Lese-Speicher (EPROM), einen Flash-Speicher oder jedes andere durch einen Computer lesbare Speichermedium umfassen, das in der Lage ist, Programmanweisungen oder digitale Informationen zu speichern.

[0051] Die vom dauerhaften Speicher 170 verwendeten Medien können auch austauschbar sein. Beispielsweise kann eine austauschbare Festplatte für den dauerhaften Speicher 170 verwendet werden. Zu anderen Beispielen gehören optische und magnetische Festplatten, Speichersticks und Chipkarten, die in eine Datenträger-Aufnahme eingelegt werden, um auf ein anderes durch einen Computer lesbares Speichermedium übertragen zu werden, das ebenfalls Teil des dauerhaften Speichers 170 ist.

[0052] Die Datenübertragungseinheit 152 schafft in diesen Beispielen die Voraussetzung für Datenaustauschvorgänge mit anderen Datenverarbeitungssystemen oder -einheiten, darunter Ressourcen der Client-Datenverarbeitungseinheiten 104 und 110. In diesen Beispielen umfasst die Datenübertragungseinheit 152 eine oder mehrere Netzwerkschnittstellenkarten. Die Datenübertragungseinheit 152 kann Datenaustauschvorgänge durch Verwenden von physischen oder drahtlosen Datenübertragungsverbindungen ermöglichen. Softwareverteilungsprogramme und andere Programme und Daten, die für eine Implementierung der vorliegenden Erfindung verwendet werden, können über die Datenübertragungseinheit 152 in den dauerhaften Speicher 170 des Server-Computers 150 heruntergeladen werden.

[0053] Die E/A-Schnittstellen 156 ermöglichen eine Eingabe und Ausgabe von Daten bei anderen Einheiten, die mit dem Server-Computer 150 verbunden werden können. Beispielsweise können eine oder mehrere E/A-Schnittstellen 156 eine Verbindung zu einer oder mehreren externen Einheiten 190 wie etwa einer Tasta-

tur, einem Tastenfeld, einem berührungsempfindlichen Bildschirm, einem Mikrofon, einer Digitalkamera und/oder einer anderen geeigneten Eingabeeinheit herstellen. Zu einer oder mehreren externen Einheiten 190 können auch tragbare durch einen Computer lesbare Speichermedien gehören, zum Beispiel Speichersticks, tragbare optische oder magnetische Festplatten und Speicherkarten. Software und Daten, die zum Ausführen von Ausführungsformen der vorliegenden Erfindung verwendet werden, z.B. das Virenprüfprogramm 175 auf dem Server-Computer 150, können auf solchen tragbaren durch einen Computer lesbaren Speichermedien gespeichert und über die E/A-Schnittstelle(n) 156 in den dauerhaften Speicher 170 geladen werden. Die E/A-Schnittstelle(n) ist/sind auch mit einer Anzeige 180 verbunden.

[0054] Die Anzeige 180 stellt einen Mechanismus zum Anzeigen von Daten für einen Benutzer bereit, und es kann sich dabei beispielsweise um einen Computermonitor handeln. Die Anzeige 180 kann auch als berührungsempfindlicher Bildschirm fungieren, beispielsweise als Anzeige eines Tablet-Computers.

[0055] Fig. 2 stellt einen Ablaufplan 200 bereit, der beispielhafte Vorgänge im Zusammenhang mit der Anwendung der Offenbarung veranschaulicht. Nach dem Programmstart überprüft das Verfahren des Programms 175 bei 210 auf neue Stapel- oder Streaming-Dateien und/oder neue Virusdefinitionen. Stapeldateien können bereits im Datenspeicher vorhanden sein und Streaming-Dateien können im Datenspeicher ankommen. Neue Virusdefinitionen können von der Netzwerkdatenübertragungseinheit 152 an ein System zur Abwehr von Eindringversuchen (Intrusion Prevention System) übermittelt werden. Bei 220 überprüft das Verfahren, ob neue Datei-Metadaten in einer Metadaten Speichereinheit des dauerhaften Speichers 170, in einem System-Speicherbereichsnetz oder einem anderen Systemspeicher vorhanden sind. Bei neuen Dateien ohne gespeicherte Metadaten erzeugt das Verfahren des Programms 175 einen neuen Metadaten-satz und beginnt mit dem Erfassen von Datei-Metadaten. Die Metadaten werden sowohl aus der Datei selbst als auch aus den der Datei zugehörigen Verfahrensaktivitäten erfasst. Bei neuen Virusdefinitionen sucht das Verfahren nach Dateien, die noch nicht mit den neuen Virusdefinitionen geprüft wurden.

[0056] Bei 230 erzeugt das Verfahren des Programms 175 eine priorisierte Liste aller neuen Dateien entsprechend ihrem Bedarf an einer Virenprüfung. Das Verfahren bestimmt die Dateipriorität unter Verwenden von einem oder mehreren berechneten Dateikontrollsummenwerten und eines Korrelationsindikators für Viren aus den Metadaten der neuen Dateien. Höhere Prioritäten werden Dateien mit einer höheren Dringlichkeit für ein Überprüfen zugewiesen. (Dateien, auf die für eine nachgelagerte Verarbeitung zugegriffen werden muss, Dateien ohne historische Metadaten, Dateien mit Kontrollsummenproblemen usw.)

[0057] Bei 240 wählt das Verfahren des Programms 175 den Teil der Datendatei mit der höchsten Priorität aus der erzeugten Liste zum Überprüfen aus. Bei 250 überprüft das Verfahren die Datei unter Verwenden der aktuellen Virusdefinitionen. Bei 260 aktualisiert das Verfahren die Metadaten der Datei, darunter die Uhrzeit und das Datum der Überprüfung, die verwendeten Virusdefinitionen, die Anzahl und die Details von gefundenen Viren, die neue CVE-Bewertung für die Datei, alle zugehörigen Segmentinformationen usw. Bei 270 markiert das Verfahren die Datei als frei von bekannten Viren zum Prüfzeitpunkt. In dieser Ausführungsform stellt das Verfahren nachgeschalteten Datenabnehmern, Datennutzern und Administratoren die CVE-Bewertung zur Verfügung. Die Empfänger der CVE-Bewertung ermitteln, ob die Datei ausreichend sauber ist, um sie Abnehmern zur Verfügung zu stellen.

[0058] Das Verfahren leitet Dateien, die Viren enthalten, an Dateibereinigungsprozesse weiter. In Schritt 275 ermittelt das Verfahren, ob die Datei durch den Prozess bereinigt wurde. Nicht infizierte Dateien werden an Schritt 280 weitergeleitet. Viren enthaltende Dateien, die nicht bereinigt wurden, werden an Schritt 290 weitergeleitet und von einer Verwendung durch Datenabnehmer und vom Bereitstellen zum Zugriff durch Datenabnehmer ausgeschlossen.

[0059] Bei 280 werden nicht infizierte Dateien dem Datenspeicher hinzugefügt, so dass sie für Datenabnehmer nutzbar sind. Nach dem Schritt 280 kehrt das Verfahren zum Schritt 210 zurück, um nach neuen Dateien/Virusdefinitionen zu suchen. Wenn keine neuen Dateien oder Definitionen gefunden werden, geht das Verfahren weiter zum Schritt 215 und ermittelt, ob alle Dateien auf der erzeugten Prioritätsliste mit den aktuellen Virusdefinitionen überprüft wurden. Das Verfahren kehrt zum Schritt 240 zurück und wählt die nächste priorisierte Datei aus, wenn noch nicht alle Dateien überprüft wurden. Wenn alle Dateien überprüft wurden, kehrt das Verfahren zum Schritt 210 zurück und fährt mit dem Überprüfen auf neue Dateien und Virusdefinitionen fort.

[0060] Zwar umfasst diese Offenbarung eine ausführliche Beschreibung von Cloud-Computing, es sollte jedoch von vornherein klar sein, dass Implementierungen der hierin wiedergegebenen Lehren nicht auf eine

Cloud-Computing-Umgebung beschränkt sind. Vielmehr können Ausführungsformen der vorliegenden Erfindung in Verbindung mit einem beliebigen anderen Typ von gegenwärtig bekannter oder zu einem späteren Zeitpunkt entwickelter Datenverarbeitungsumgebung implementiert werden.

[0061] Cloud-Computing ist ein Dienstleistungsmodell, das einen komfortablen und bedarfsorientierten Netzwerkzugang zu einem gemeinsam genutzten Pool von konfigurierbaren Datenverarbeitungs-Ressourcen ermöglicht (z.B. Netzwerke, Netzwerk-Bandbreite, Server, Verarbeitung, Speicher, Speicherplatz, Anwendungen, virtuelle Maschinen und Dienste), die mit minimalem Verwaltungsaufwand bzw. minimaler Interaktion mit einem Anbieter des Dienstes schnell zur Verfügung gestellt und freigegeben werden können. Dieses Cloud-Modell kann mindestens fünf Charakteristika, mindestens drei Dienstmodelle und mindestens vier Bereitstellungsmodelle umfassen.

[0062] Bei den Charakteristika handelt es sich um die folgenden:

Bedarfsorientierte Selbstbedienung (on-demand self-service): Ein Cloud-Kunde kann je nach Bedarf einseitig Datenverarbeitungs-Ressourcen wie Server-Zeit und Netzspeicher automatisch in Anspruch nehmen, ohne dass eine menschengeführte Interaktion mit dem Anbieter des Dienstes erforderlich ist.

Allgemeiner Netzzugriff (broad network access): Ressourcen sind über ein Netzwerk verfügbar und über Standardmechanismen zugänglich, die eine Verwendung durch heterogene Thin- oder Thick-Client-Plattformen (z.B. Mobiltelefone, Laptops und PDAs) unterstützen.

Ressourcen-Bündelung (resource pooling): Die Datenverarbeitungs-Ressourcen des Anbieters werden gebündelt, um mehreren Kunden zu dienen, wobei ein Multi-Tenant-Modell verwendet wird, bei dem verschiedene physische und virtuelle Ressourcen entsprechend dem Bedarf dynamisch zugewiesen und erneut zugewiesen werden. Es besteht insofern eine gewisse Unabhängigkeit vom Standort, als der Kunde im Allgemeinen keine Kontrolle über den bzw. Kenntnis vom genauen Standort der bereitgestellten Ressourcen hat, den Standort jedoch auf einer höheren Abstraktionsebene (z.B. Region, Staat oder Datenzentrum) festlegen kann.

Schnelle Anpassungsfähigkeit (rapid elasticity): Ressourcen können rasch und anpassungsfähig, in einigen Fällen automatisch, bereitgestellt werden, um schnell eine Erweiterungsmöglichkeit (scale out) zu bieten, und rasch wieder freigegeben werden, wodurch eine schnelle Reduzierung (scale in) geboten wird. Aus Sicht des Kunden scheinen die zur Bereitstellung verfügbaren Ressourcen oftmals unbegrenzt zu sein und können in beliebiger Menge zu einem beliebigen Zeitpunkt erworben werden.

Dienst für nutzungsabhängige Abrechnung (measured service): Cloud-Systeme steuern und optimieren die Ressourcen-Nutzung automatisch durch Einsatz einer Messfunktion auf einer Abstraktionsebene, die für den Typ des Dienstes (z.B. Speicherplatz, Verarbeitung, Bandbreite sowie aktive Benutzerkonten) geeignet ist. Die Ressourcen-Nutzung kann überwacht, gesteuert und aufgezeichnet werden, wobei sowohl für den Anbieter als auch für den Abnehmer des verwendeten Dienstes Transparenz gegeben ist.

[0063] Bei den Dienstmodellen handelt es sich um die Folgenden:

Software als Dienstleistung (Software as a Service, SaaS): Die für den Kunden bereitgestellte Funktionalität besteht in der Nutzung der in einer Cloud-Infrastruktur ausgeführten Anwendungen des Anbieters. Auf die Anwendungen kann von verschiedenen Client-Einheiten durch eine Thin-Client-Schnittstelle wie einen Web-Browser (z.B. auf dem Web beruhende eMail) zugegriffen werden. Der Kunde verwaltet bzw. steuert nicht die zugrunde liegende Cloud-Infrastruktur, darunter Netzwerk, Server, Betriebssysteme, Speicherplatz oder sogar einzelne Anwendungsmöglichkeiten, mit der möglichen Ausnahme begrenzter benutzerspezifischer Anwendungskonfigurationseinstellungen.

[0064] Plattform als Dienstleistung (Platform as a Service, PaaS): Die für den Kunden bereitgestellte Funktionalität besteht in der Verwendung von vom Kunden erstellten oder angeforderten Anwendungen, die unter Verwenden von vom Anbieter unterstützten Programmiersprachen und Hilfsprogrammen erzeugt werden, in der Cloud-Infrastruktur. Der Kunde verwaltet oder steuert nicht die zugrundeliegende Cloud-Infrastruktur, darunter Netze, Server, Betriebssysteme oder Speicherplatz, hat jedoch die Kontrolle über die eingesetzten Anwendungen und möglicherweise über Konfigurationen der Hosting-Umgebung der Anwendung.

[0065] Infrastruktur als Dienstleistung (Infrastructure as a Service, IaaS): Die für den Kunden bereitgestellte Funktionalität besteht in der Bereitstellung von Datenverarbeitung, Speicherplatz, Netzwerken und anderen grundlegenden Datenverarbeitungs-Ressourcen, wobei der Kunde eine beliebige Software einsetzen und

ausführen kann, die Betriebssysteme und Anwendungen umfassen kann. Der Kunde verwaltet oder steuert nicht die zugrunde liegende Cloud-Infrastruktur, hat jedoch die Kontrolle über Betriebssysteme, Speicherplatz, eingesetzte Anwendungen und möglicherweise eine begrenzte Kontrolle über die Auswahl von Netzwerkkomponenten (z.B. Host-Firewalls).

[0066] Bei den Einsatzmodellen handelt es sich um die Folgenden:

Private Cloud: Die Cloud-Infrastruktur wird einzig und allein für eine Organisation betrieben. Sie kann durch die Organisation selbst oder einen Dritten verwaltet werden und kann sich in den eigenen Räumen oder in fremden Räumen befinden.

Gemeinschaftliche Cloud (community cloud): Die Cloud-Infrastruktur wird von mehreren Organisationen gemeinsam genutzt und unterstützt eine spezielle Benutzergemeinschaft mit gemeinsamen Interessen (z.B. Zielsetzung, Sicherheitsanforderungen, Strategie- und Konformitätsüberlegungen). Sie kann durch die Organisationen oder einen Dritten verwaltet werden und kann sich in den eigenen Räumen oder fremden Räumen befinden.

Öffentliche Cloud (public cloud): Die Cloud-Infrastruktur wird der Allgemeinheit oder einer großen Industrie-Gruppe zur Verfügung gestellt und ist im Besitz einer Organisation, die Cloud-Dienste verkauft.

Kombinierte Cloud (hybrid cloud): Die Cloud-Infrastruktur setzt sich aus zwei oder mehreren Clouds (privat, gemeinschaftlich oder öffentlich) zusammen, die eigenständige Einheiten bleiben, jedoch durch eine standardisierte oder firmeneigene Technologie, die die Übertragbarkeit von Daten und Anwendungen ermöglicht, miteinander verbunden sind (z.B. Cloud-Zielgruppenverteilung für eine Auslastungsverteilung zwischen Clouds).

[0067] Eine Cloud-Computing-Umgebung ist dienstleistungsorientiert mit Fokus auf Statusunabhängigkeit, loser Kopplung, Modularität und semantischer Interoperabilität. Den Kern des Cloud-Computing bildet eine Infrastruktur, die ein Netzwerk aus miteinander verbundenen Knoten aufweist.

[0068] Unter Bezugnahme auf **Fig. 3** wird nun eine veranschaulichende Cloud-Computing-Umgebung 50 abgebildet. Wie gezeigt ist, weist die Cloud-Computing-Umgebung 50 einen oder mehrere Cloud-Computing-Knoten 10 auf, mit denen von Cloud-Nutzern verwendete lokale Datenverarbeitungseinheiten wie der elektronische Assistent (PDA, personal digital assistant) oder das Mobiltelefon 54A, der Desktop-Computer 54B, der Laptop-Computer 54C und/oder das Automobil-Computer-System 54N Daten austauschen können. Die Knoten 10 können miteinander Daten austauschen. Sie können physisch oder virtuell in ein oder mehrere Netzwerke wie private, gemeinschaftliche, öffentliche oder kombinierte Clouds gruppiert werden (nicht gezeigt), wie vorstehend beschrieben wurde, oder in eine Kombination daraus. Dies ermöglicht es der Cloud-Computing-Umgebung 50, Infrastruktur, Plattformen und/oder Software als Dienstleistung anzubieten, für die ein Cloud-Nutzer keine Ressourcen auf einer lokalen Datenverarbeitungseinheit vorhalten muss. Es sei darauf hingewiesen, dass die Arten von in **Fig. 3** gezeigten Datenverarbeitungseinheiten 54A bis N lediglich veranschaulichend sein sollen und dass die Datenverarbeitungsknoten 10 und die Cloud-Computing-Umgebung 50 über eine beliebige Art Netzwerk und/oder über eine beliebige Art von über ein Netzwerk aufrufbarer Verbindung (z.B. unter Verwendung eines Web-Browsers) mit einer beliebigen Art von computergestützter Einheit Daten austauschen können.

[0069] Unter Bezugnahme auf **Fig. 4** wird nun ein Satz von funktionalen Abstraktionsschichten gezeigt, die durch die Cloud-Computing-Umgebung 50 (**Fig. 3**) bereitgestellt werden. Es sollte von vornherein klar sein, dass die in **Fig. 4** gezeigten Komponenten, Schichten und Funktionen lediglich veranschaulichend sein sollen und Ausführungsformen der Erfindung nicht darauf beschränkt sind. Wie abgebildet ist, werden die folgenden Schichten und entsprechenden Funktionen bereitgestellt:

[0070] Eine Hardware- und Software-Schicht 60 enthält Hardware- und Software-Komponenten. Zu Beispielen für Hardware-Komponenten gehören: Mainframe-Computer 61; auf der RISC- (Reduced Instruction Set Computer) Architektur beruhende Server 62; Server 63; Blade-Server 64; Speichereinheiten 65; und Netzwerke sowie Netzwerkkomponenten 66. In einigen Ausführungsformen umfassen Software-Komponenten eine Netzwerk-Anwendungsserver-Software 67 und eine Datenbank-Software 68.

[0071] Eine Virtualisierungsschicht 70 stellt eine Abstraktionsschicht bereit, aus der die folgenden Beispiele für virtuelle Einheiten bereitgestellt werden können: virtuelle Server 71; virtueller Speicher 72; virtuelle Netzwerke 73, darunter virtuelle private Netzwerke; virtuelle Anwendungen und Betriebssysteme 74; und virtuelle Clients 75.

[0072] In einem Beispiel kann eine Verwaltungsschicht 80 die nachfolgend beschriebenen Funktionen bereitstellen. Die Ressourcen-Versorgung (resource provisioning) 81 stellt dynamisches Beschaffen von Datenverarbeitungs-Ressourcen sowie anderen Ressourcen bereit, die zum Durchführen von Aufgaben innerhalb der Cloud-Computing-Umgebung verwendet werden. Die Gebührenerfassung und Preisermittlung (metering and pricing) 82 stellt eine Kostenverfolgung beim Verwenden von Ressourcen innerhalb der Cloud-Computing-Umgebung sowie eine Abrechnung oder Rechnungsstellung für die Inanspruchnahme dieser Ressourcen bereit. In einem Beispiel können diese Ressourcen Anwendungssoftware-Lizenzen aufweisen. Die Sicherheit (security) stellt eine Identitätsüberprüfung für Cloud-Kunden und Aufgaben sowie Schutz für Daten und andere Ressourcen bereit. Das Benutzerportal 83 stellt den Zugang zu der Cloud-Computing-Umgebung für Nutzer und für Systemadministratoren bereit. Die Dienstgüteverwaltung (service level management) 84 stellt eine Zuordnung und Verwaltung von Cloud-Computing-Ressourcen bereit, so dass die Anforderungen an die Dienstgüte erfüllt werden. Die Planung und Erfüllung der Dienstgütevereinbarung (SLA, Service Level Agreement planning and fulfillment) 85 stellt die Vorbereitung für und die Beschaffung von Cloud-Computing-Ressourcen bereit, für die gemäß einer SLA ein künftiger Bedarf vorausgesehen wird.

[0073] Eine Arbeitslastenschicht (workloads layer) 90 stellt Beispiele einer Funktionalität bereit, für welche die Cloud-Computing-Umgebung verwendet werden kann. Zu Beispielen für Arbeitslasten und Funktionen, die von dieser Schicht bereitgestellt werden können, gehören: Abbildung und Navigation 91; Software-Entwicklung und Lebenszyklusverwaltung 92; virtuelles Klassenzimmer 93 als Bildungsangebot; Datenanalyseverarbeitung 94; Transaktionsverarbeitung 95; und ein Virenprüfprogramm 175.

[0074] Bei der vorliegenden Offenbarung kann es sich um ein System, ein Verfahren und/oder ein Computerprogrammprodukt mit jedem möglichen technischen Integrationsgrad handeln. Die Erfindung kann in jedem System, einzeln oder parallel, das einen Befehlsstrom verarbeitet, vorteilhaft angewandt werden. Das Computerprogrammprodukt kann (ein) durch einen Computer lesbare(n) Speichermedium (oder -medien) mit durch einen Computer lesbaren Programmanweisungen darauf umfassen, um einen Prozessor dazu zu veranlassen, Aspekte der vorliegenden Offenbarung auszuführen.

[0075] Bei dem durch einen Computer lesbaren Speichermedium kann es sich um eine physische Einheit handeln, die Anweisungen zum Verwenden durch eine Einheit zum Ausführen von Anweisungen behalten und speichern kann. Bei dem durch einen Computer lesbaren Speichermedium kann es sich zum Beispiel um eine elektronische Speichereinheit, eine magnetische Speichereinheit, eine optische Speichereinheit, eine elektromagnetische Speichereinheit, eine Halbleiterspeichereinheit oder jede geeignete Kombination daraus handeln, ohne auf diese beschränkt zu sein. Zu einer nicht erschöpfenden Liste spezifischerer Beispiele des durch einen Computer lesbaren Speichermediums gehören die Folgenden: eine tragbare Computerdiskette, eine Festplatte, ein Direktzugriffsspeicher (RAM), ein Nur-Lese-Speicher (ROM), ein löschbarer programmierbarer Nur-Lese-Speicher (EPROM bzw. Flash-Speicher), ein statischer Direktzugriffsspeicher (SRAM), ein tragbarer Kompaktspeicherplatte-Nur-Lese-Speicher (CD-ROM), eine DVD (digital versatile disc), ein Speicher-Stick, eine Diskette, eine mechanisch codierte Einheit wie zum Beispiel Lochkarten oder erhabene Strukturen in einer Rille, auf denen Anweisungen gespeichert sind, und jede geeignete Kombination daraus. Ein durch einen Computer lesbares Speichermedium soll in der Verwendung hierin nicht als flüchtige Signale an sich aufgefasst werden, wie zum Beispiel Funkwellen oder andere sich frei ausbreitende elektromagnetische Wellen, elektromagnetische Wellen, die sich durch einen Wellenleiter oder ein anderes Übertragungsmedium ausbreiten (z.B. ein Lichtwellenleiterkabel durchlaufende Lichtimpulse) oder durch einen Draht übertragene elektrische Signale.

[0076] Hierin beschriebene, durch einen Computer lesbare Programmanweisungen können von einem durch einen Computer lesbaren Speichermedium auf jeweilige Datenverarbeitungs-/Verarbeitungseinheiten oder über ein Netzwerk wie zum Beispiel das Internet, ein lokales Netzwerk, ein Weitverkehrsnetz und/oder ein drahtloses Netzwerk auf einen externen Computer oder eine externe Speichereinheit heruntergeladen werden. Das Netzwerk kann Kupferübertragungskabel, Lichtwellenübertragungsleiter, drahtlose Übertragung, Leitwegrechner, Firewalls, Vermittlungseinheiten, Gateway-Computer und/oder Edge-Server aufweisen. Eine Netzwerkadapterkarte oder Netzwerkschnittstelle in jeder Datenverarbeitungs-/Verarbeitungseinheit empfängt durch einen Computer lesbare Programmanweisungen aus dem Netzwerk und leitet die durch einen Computer lesbaren Programmanweisungen zum Speichern in einem durch einen Computer lesbaren Speichermedium innerhalb der entsprechenden Datenverarbeitungs-/Verarbeitungseinheit weiter.

[0077] Bei durch einen Computer lesbaren Programmanweisungen zum Ausführen von Arbeitsschritten der vorliegenden Offenbarung kann es sich um Assembler-Anweisungen, ISA-Anweisungen (Instruction-Set-Architecture), Maschinenanweisungen, maschinenabhängige Anweisungen, Mikrocode, Firmware-Anweisun-

gen, zustandssetzende Daten, Konfigurationsdaten für einen integrierten Schaltkreis oder entweder Quellcode oder Objektcode handeln, die in einer beliebigen Kombination aus einer oder mehreren Programmiersprachen geschrieben werden, darunter eine objektorientierte Programmiersprache wie Smalltalk, C++ o.ä. sowie prozedurale Programmiersprachen wie die Programmiersprache „C“ oder ähnliche Programmiersprachen. Die durch einen Computer lesbaren Programmanweisungen können vollständig auf dem Computer des Benutzers, teilweise auf dem Computer des Benutzers, als eigenständiges Software-Paket, teilweise auf dem Computer des Benutzers und teilweise auf einem fernen Computer oder vollständig auf dem fernen Computer oder Server ausgeführt werden. In letzterem Fall kann der entfernt angeordnete Computer mit dem Computer des Benutzers durch eine beliebige Art Netzwerk verbunden sein, darunter ein lokales Netzwerk (LAN) oder ein Weitverkehrsnetz (WAN), oder die Verbindung kann mit einem externen Computer hergestellt werden (zum Beispiel über das Internet unter Verwendung eines Internet-Dienstanbieters). In einigen Ausführungsformen können elektronische Schaltungen, darunter zum Beispiel programmierbare Logikschaltungen, vor Ort programmierbare Gatter-Anordnungen (FPGA, field programmable gate arrays) oder programmierbare Logikanordnungen (PLA, programmable logic arrays) die durch einen Computer lesbaren Programmanweisungen ausführen, indem sie Zustandsinformationen der durch einen Computer lesbaren Programmanweisungen nutzen, um die elektronischen Schaltungen zu personalisieren, um Aspekte der vorliegenden Erfindung durchzuführen.

[0078] Aspekte der vorliegenden Erfindung sind hierin unter Bezugnahme auf Ablaufpläne und/oder Blockschaubilder von Verfahren, Vorrichtungen (Systemen) und Computerprogrammprodukten gemäß Ausführungsformen der Erfindung beschrieben. Es wird darauf hingewiesen, dass jeder Block der Ablaufpläne und/oder der Blockschaubilder sowie Kombinationen von Blöcken in den Ablaufplänen und/oder den Blockschaubildern mittels durch einen Computer lesbare Programmanweisungen ausgeführt werden können.

[0079] Diese durch einen Computer lesbaren Programmanweisungen können einem Prozessor eines Universalcomputers, eines Spezialcomputers oder einer anderen programmierbaren Datenverarbeitungseinheit bereitgestellt werden, um eine Maschine zu erzeugen, so dass die über den Prozessor des Computers bzw. der anderen programmierbaren Datenverarbeitungsvorrichtung ausgeführten Anweisungen ein Mittel zum Umsetzen der in dem Block bzw. den Blöcken der Ablaufpläne und/oder der Blockschaubilder festgelegten Funktionen/Schritte erzeugen. Diese durch einen Computer lesbaren Programmanweisungen können auch auf einem durch einen Computer lesbaren Speichermedium gespeichert sein, das einen Computer, eine programmierbare Datenverarbeitungseinheit und/oder andere Einheiten so steuern kann, dass sie auf eine bestimmte Art funktionieren, so dass das durch einen Computer lesbare Speichermedium, auf dem Anweisungen gespeichert sind, ein Herstellungsprodukt aufweist, darunter Anweisungen, welche Aspekte der/des in dem Block bzw. den Blöcken des Ablaufplans und/oder der Blockschaubilder angegebenen Funktion/Schritts umsetzen.

[0080] Die durch einen Computer lesbaren Programmanweisungen können auch auf einen Computer, eine andere programmierbare Datenverarbeitungsvorrichtung oder eine andere Einheit geladen werden, um das Ausführen einer Reihe von Prozessschritten auf dem Computer bzw. der anderen programmierbaren Vorrichtung oder anderen Einheit zu verursachen, um einen auf einem Computer ausgeführten Prozess zu erzeugen, so dass die auf dem Computer, einer anderen programmierbaren Vorrichtung oder einer anderen Einheit ausgeführten Anweisungen die in dem Block bzw. den Blöcken der Ablaufpläne und/oder der Blockschaubilder festgelegten Funktionen/Schritte umsetzen.

[0081] Die Ablaufpläne und die Blockschaubilder in den Figuren veranschaulichen die Architektur, die Funktionalität und den Betrieb möglicher Ausführungen von Systemen, Verfahren und Computerprogrammprodukten gemäß verschiedenen Ausführungsformen der vorliegenden Erfindung. In diesem Zusammenhang kann jeder Block in den Ablaufplänen oder Blockschaubildern ein Modul, ein Segment oder einen Teil von Anweisungen darstellen, die eine oder mehrere ausführbare Anweisungen zum Ausführen der bestimmten logischen Funktion(en) aufweisen. In einigen alternativen Ausführungen können die in dem Block angegebenen Funktionen in einer anderen Reihenfolge als in den Figuren gezeigt stattfinden. Zwei nacheinander gezeigte Blöcke können zum Beispiel in Wirklichkeit im Wesentlichen gleichzeitig ausgeführt werden, oder die Blöcke können manchmal je nach entsprechender Funktionalität in umgekehrter Reihenfolge ausgeführt werden. Es ist ferner anzumerken, dass jeder Block der Blockschaubilder und/oder der Ablaufpläne sowie Kombinationen aus Blöcken in den Blockschaubildern und/oder den Ablaufplänen durch spezielle auf Hardware beruhende Systeme umgesetzt werden können, welche die festgelegten Funktionen oder Schritte durchführen, oder Kombinationen aus Spezial-Hardware und Computeranweisungen ausführen.

Patentansprüche

1. Durch einen Computer (150) implementiertes Verfahren (200) zum Aufrechterhalten der Sicherheit eines Datenverarbeitungssystems, wobei das Verfahren (200) aufweist:
Empfangen (220) von Metadaten aus einer Metadaten Speichereinheit durch einen oder mehrere Computerprozessoren (154), wobei die Metadaten mindestens einem Teil einer Datendatei zugehörig sind und wobei die Metadaten einen oder mehrere Indikatoren für Viren aufweisen;
Erzeugen (230) einer Priorität durch den einen oder die mehreren Computerprozessoren für den mindestens einen Teil der Datendatei entsprechend den Metadaten;
Durchführen einer Überprüfung (250) des mindestens einen Teiles der Datendatei durch den einen oder die mehreren Computerprozessoren (154) entsprechend der Priorität;
Aktualisieren (260) der Metadaten des mindestens einen Teiles der Datendatei entsprechend der Überprüfung durch den einen oder die mehreren Computerprozessoren; und
Bereitstellen (280) der Datendatei für eine Weiterverarbeitung durch den einen oder die mehreren Computerprozessoren (154),
wobei den Metadaten eine Kontrollsumme zugeordnet wird, die kontinuierlich parallel zum Strömen der Datendatei in einen Datenspeicher bestimmt wird.
2. Durch einen Computer (150) implementiertes Verfahren (220) nach Anspruch 1, wobei die Priorität entsprechend einer Korrelation der Metadaten der Datendatei erzeugt wird.
3. Durch einen Computer (150) implementiertes Verfahren (200) nach Anspruch 1, das außerdem ein Übergeben der Datendatei durch den einen oder die mehreren Computerprozessoren (154) an eine Bereinigungsroutine aufweist.
4. Durch einen Computer (150) implementiertes Verfahren (200) nach einem der vorhergehenden Ansprüche, wobei das Verfahren (200) außerdem aufweist:
Abgleichen einer Kontrollsumme für einen Teil der Datendatei mit einer Kontrollsumme in der Metadaten Speichereinheit durch den einen oder die mehreren Computerprozessoren (154); und
Einleiten einer Datenbereinigungsroutine durch den einen oder die mehreren Computerprozessoren (154) als Reaktion auf das Abgleichen.
5. Durch einen Computer (150) implementiertes Verfahren (200) nach einem der vorhergehenden Ansprüche, wobei das Verfahren (200) außerdem ein Aktualisieren (260) von einem Teil der Datendatei zugehörigen Kontrollsummen-Metadaten durch den einen oder mehrere Computerprozessoren (154) aufweist.
6. Durch einen Computer (150) implementiertes Verfahren (200) gemäß einem der vorangegangenen Ansprüche, wobei das Empfangen (220) der Metadaten aufweist:
Empfangen (220) einer aktualisierten Virusdefinition durch einen oder die mehreren Computerprozessoren (154); und
Erhalten der Metadaten, die den mindestens einen Teil der Datendatei zugehörig sind, aus einer Metadaten Speichereinheit durch den einen oder die mehreren Computerprozessoren (154);
wobei das Erzeugen (230) einer Priorität für die mindestens eine Datendatei ein Erzeugen der Priorität entsprechend den Metadaten und einer aktualisierten Virusdefinition aufweist; und
wobei das Durchführen der Überprüfung (250) aufweist:
Auswählen des mindestens einen Teiles der Datendatei durch den einen oder mehrere Computerprozessoren entsprechend der Priorität; und
Durchführen einer Überprüfung der Datendatei durch den einen oder die mehreren Computerprozessoren.
7. Durch einen Computer (150) implementiertes Verfahren (200) gemäß einem der Ansprüche 1 bis 5, wobei das Empfangen der Metadaten aufweist:
Erhalten (220) von Metadaten, die der mindestens einen Datendatei zugehörig sind, aus einer Metadaten Speichereinheit durch den einen oder mehrere Computerprozessoren; und
wobei das Erzeugen (230) der Priorität ein Erzeugen der Priorität für den mindestens einen Teiles der Datendatei entsprechend der Metadaten aufweist; und
wobei das Durchführen der Überprüfung (250) aufweist:
Auswählen des mindestens eines Teiles der Datendatei durch den einen oder die mehreren Computerprozessoren (154) entsprechend der Priorität;
Durchführen einer Überprüfung (250) des mindestens einen Teiles der Datendatei durch den einen oder die

mehreren Computerprozessoren (154);

Ermitteln einer Bewertung allgemeiner Schwachstellen und Gefährdungen für die Datendatei durch den einen oder die mehreren Computerprozessoren (154); und
Übermitteln der Bewertung durch den einen oder die mehreren Computerprozessoren (154) an einen Systembenutzer zur Auswertung.

8. Computersystem (150) zum Aufrechterhalten der Sicherheit eines Systems, wobei das Computersystem (150) aufweist:

einen oder mehrere Computerprozessoren (154);

eine oder mehrere durch einen Computer lesbare Speichereinheiten (158, 170); und

gespeicherte Programmanweisungen auf der einen oder mehreren durch einen Computer (150 lesbaren Speichereinheiten (158, 170) zum Ausführen durch den einen oder die mehreren Computerprozessoren (154), wobei die gespeicherten Programmanweisungen Programmanweisungen aufweisen zum:

Empfangen von Metadaten, die mindestens einer Datendatei zugehörig sind, aus einer Metadaten Speichereinheit;

Erzeugen (230) einer Priorität für die mindestens eine Datendatei entsprechend den Metadaten; und

Durchführen (250) einer Überprüfung mindestens eines Teiles der Datendatei entsprechend der Priorität;

Aktualisieren (260) der Metadaten des mindestens einen Teiles der Datendatei; und

Zurverfügungstellung (280) der Datendatei zu einer Weiterverarbeitung,

wobei die Metadaten einen oder mehrere Indikatoren für Viren aufweisen, und

wobei den Metadaten eine Kontrollsumme zugehörig ist, die kontinuierlich parallel zum Strömen der Datendatei in einen Datenspeicher bestimmt wird.

9. Computersystem (150) nach Anspruch 8, wobei die Priorität entsprechend einer Korrelation der Metadaten des mindestens einen Teiles der Datendatei erzeugt wird.

10. Computersystem (150) nach Anspruch 8 oder 9, das außerdem Programmanweisungen aufweist, um die Datendatei an eine Bereinigungsroutine zu übergeben.

11. Computersystem (150) nach einem der Ansprüche 8 bis 10, das außerdem aufweist:

Programmanweisungen zum Abgleichen einer Kontrollsumme für einen Teil der Datendatei mit einer Kontrollsumme in der Metadaten Speichereinheit durch den einen oder mehrere Computerprozessoren (154); und

Programmanweisungen zum Einleiten einer Datenbereinigungsroutine als Reaktion auf das Abgleichen der Kontrollsumme.

12. Computersystem (150) nach einem der Ansprüche 8 bis 11, das außerdem Programmanweisungen aufweist, um Kontrollsummen-Metadaten, die einem Teil der Datendatei zugehörig sind, zu aktualisieren.

13. Computerprogrammprodukt zum Aufrechterhalten der Sicherheit eines Systems (1000), wobei das Computerprogrammprodukt aufweist:

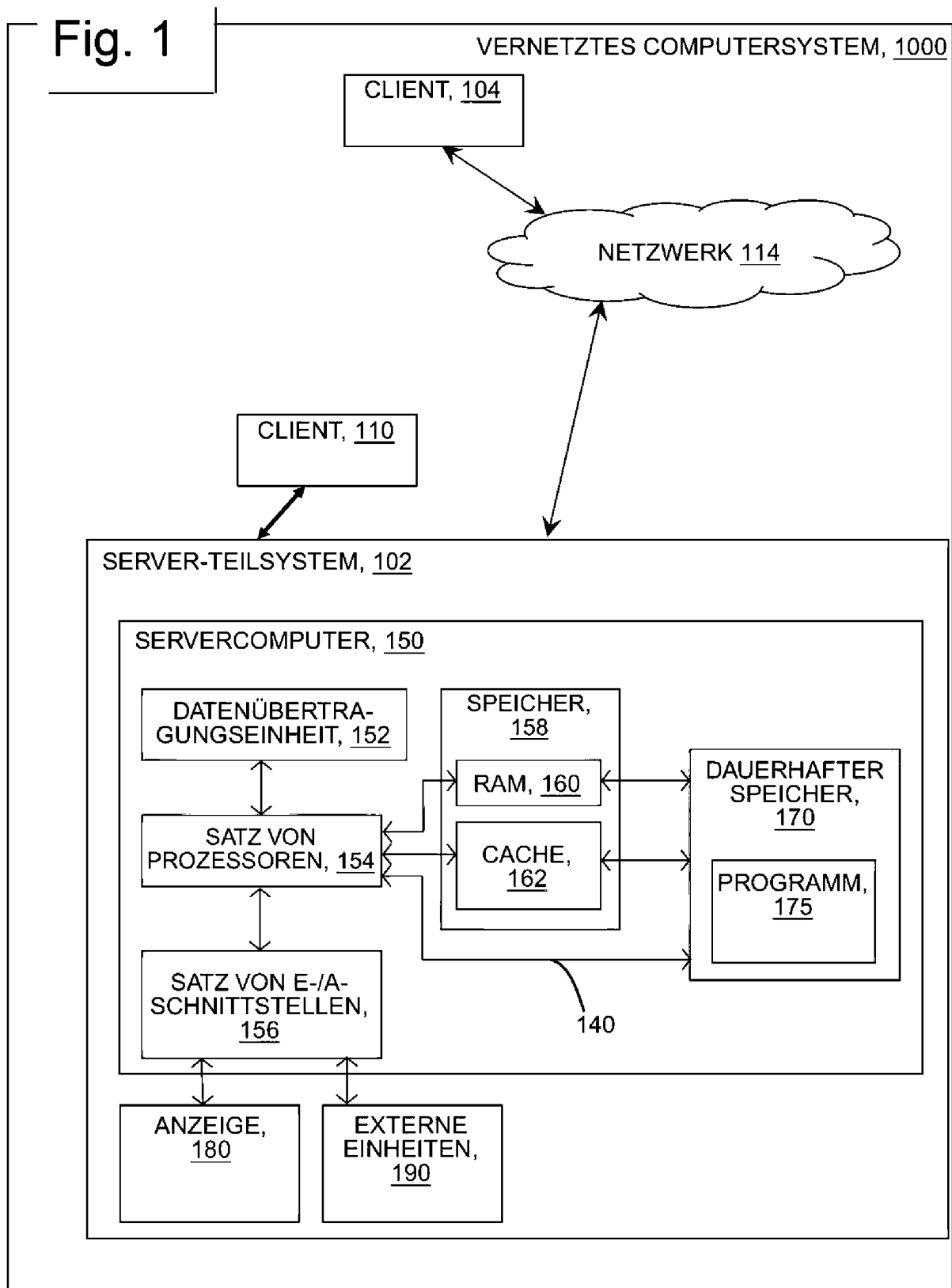
ein durch einen Computer (150) lesbares Speichermedium, das durch eine Verarbeitungsschaltung lesbar ist und Anweisungen zum Ausführen durch die Verarbeitungsschaltung speichert, um ein Verfahren nach einem der Ansprüche 1 bis 7 auszuführen.

14. Computerprogramm, das auf einem durch einen Computer (150) lesbaren Speichermedium gespeichert und in den internen Speicher eines digitalen Computers ladbar ist, wobei das Computerprogramm Software-Codeteile aufweist, um bei Ausführen des Programms in einem Computer das Verfahren nach einem der Ansprüche 1 bis 7 auszuführen.

Es folgen 4 Seiten Zeichnungen

Anhängende Zeichnungen

Fig. 1



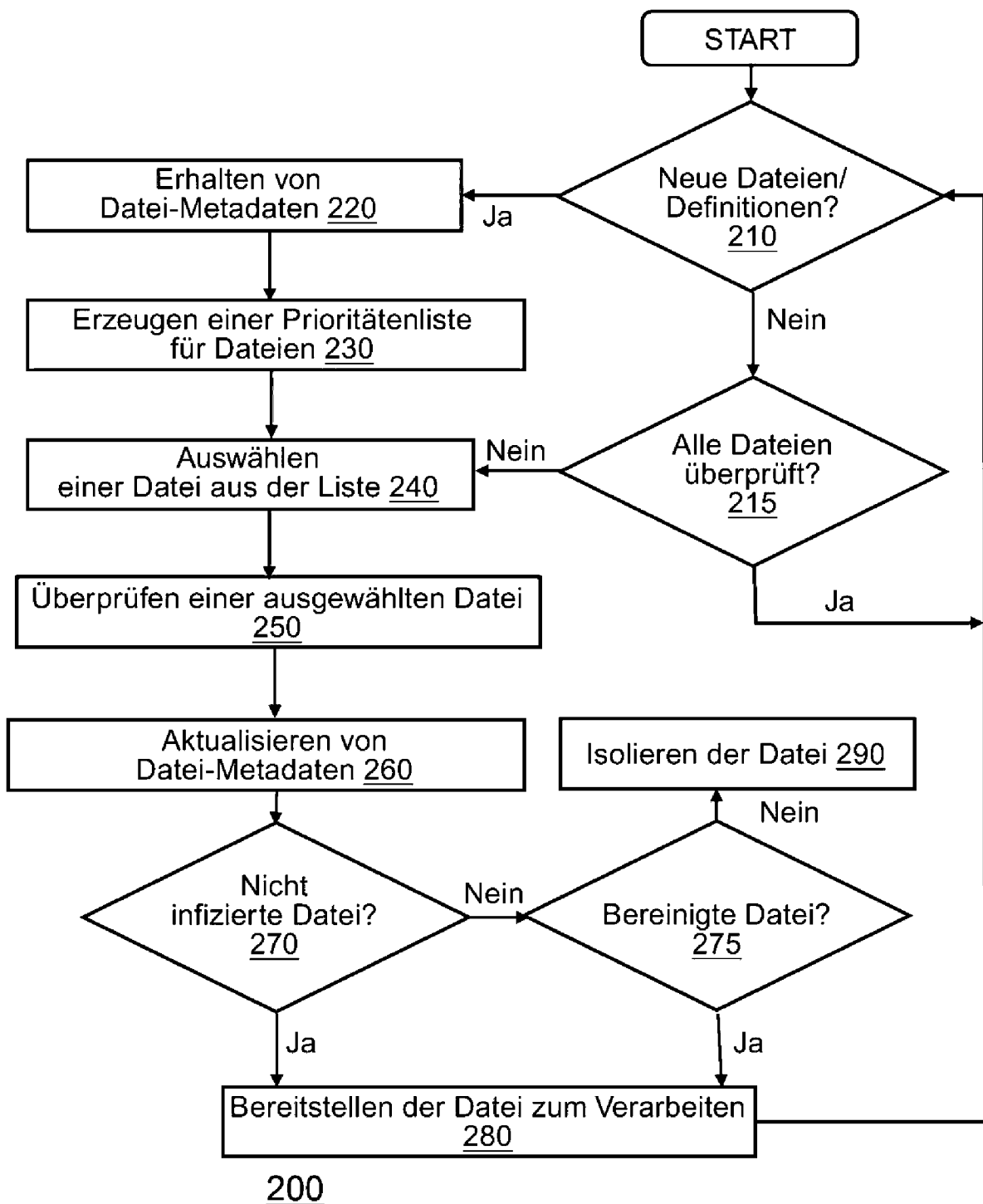


Fig. 2

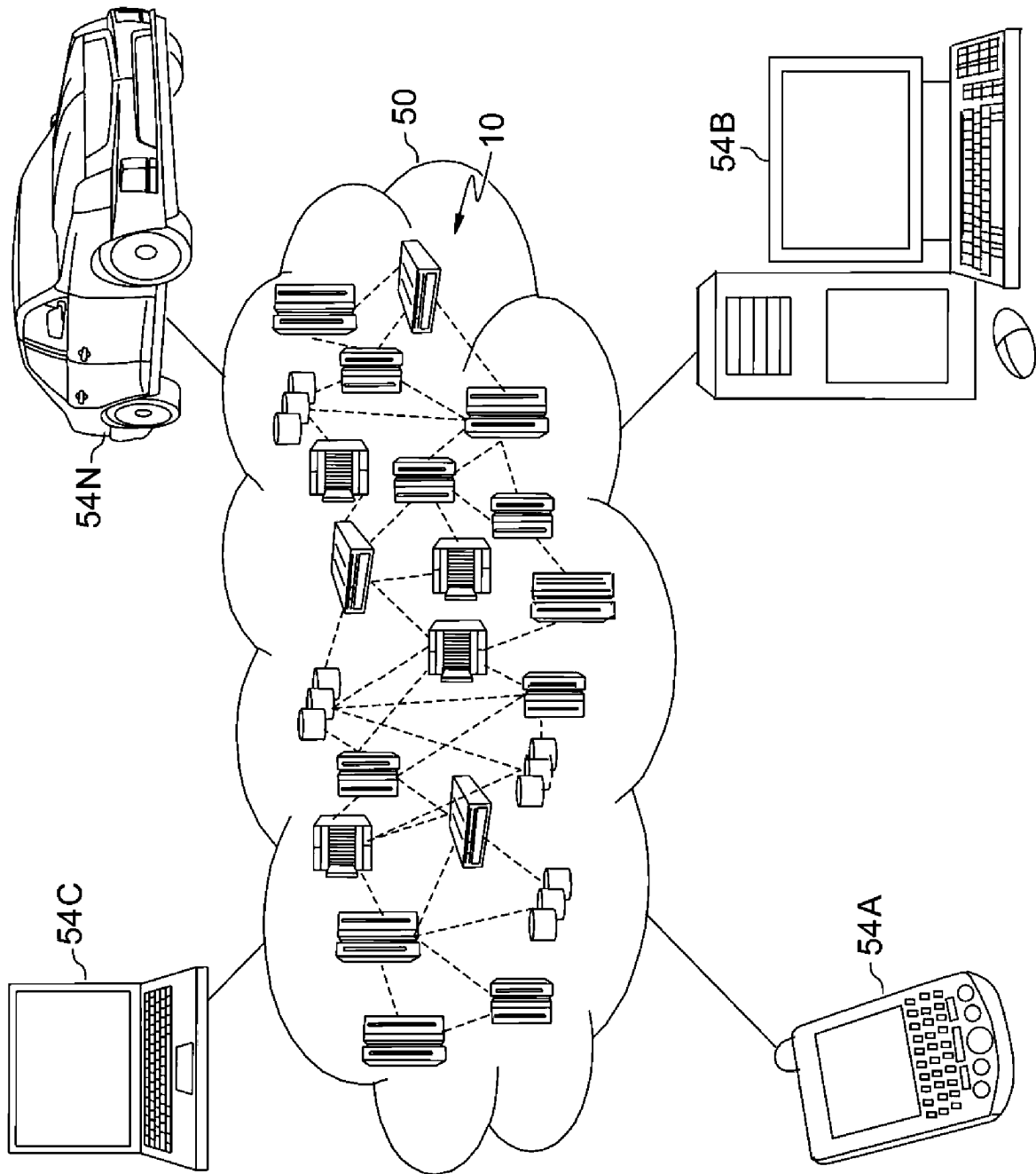


FIG. 3

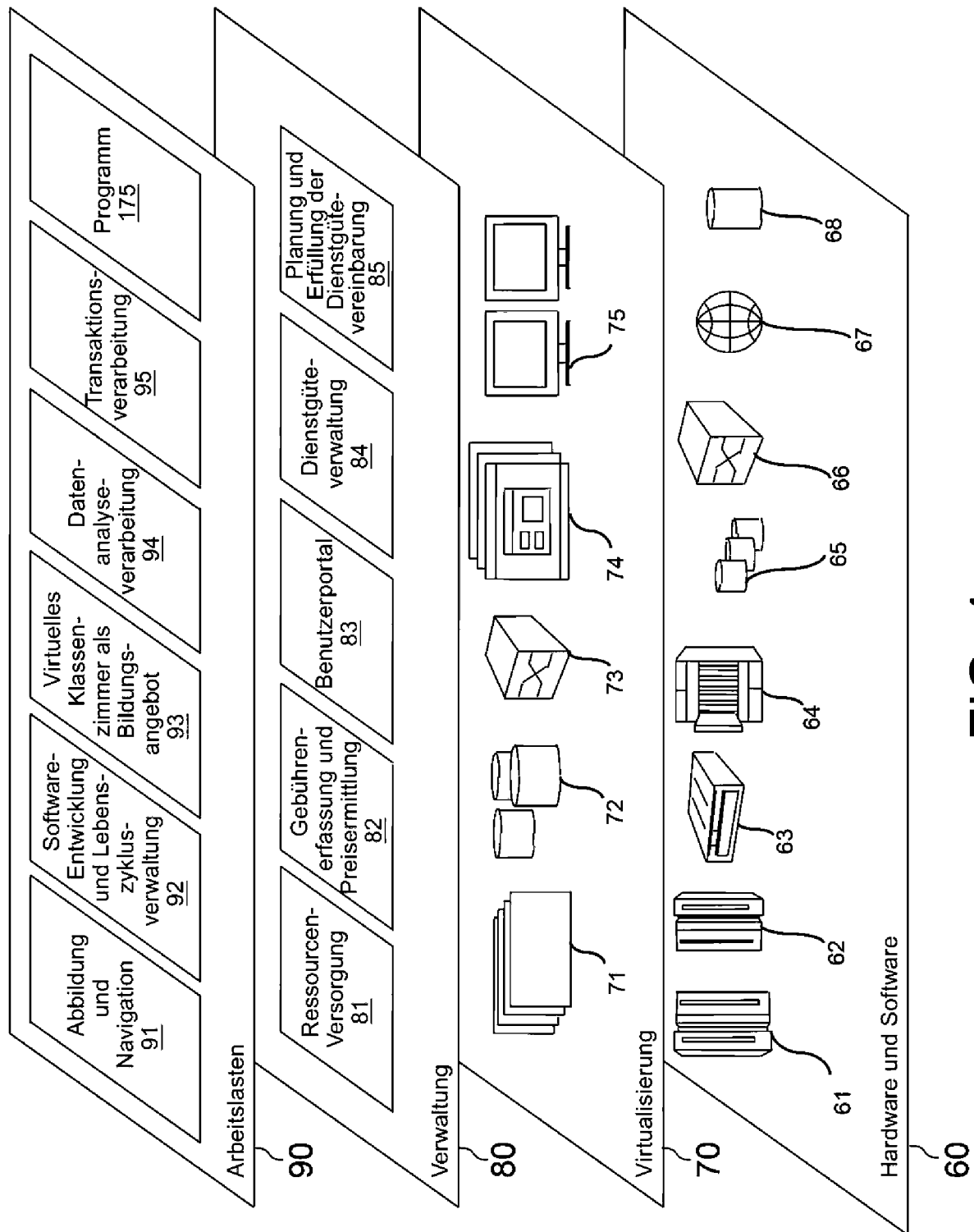


FIG. 4