



SCHWEIZERISCHE EIDGENOSSENSCHAFT
BUNDESAMT FÜR GEISTIGES EIGENTUM

⑤ Int. Cl.³: G 06 F 13/00
G 11 C 9/06

Erfindungspatent für die Schweiz und Liechtenstein

Schweizerisch-liechtensteinischer Patentschutzvertrag vom 22. Dezember 1978

⑫ **PATENTSCHRIFT** A5

⑪

622 633

⑫ Gesuchsnummer: 5668/77

⑫ Anmeldungsdatum: 06.05.1977

⑫ Priorität(en): 11.06.1976 DE 2626295

⑫ Patent erteilt: 15.04.1981

⑫ Patentschrift
veröffentlicht: 15.04.1981

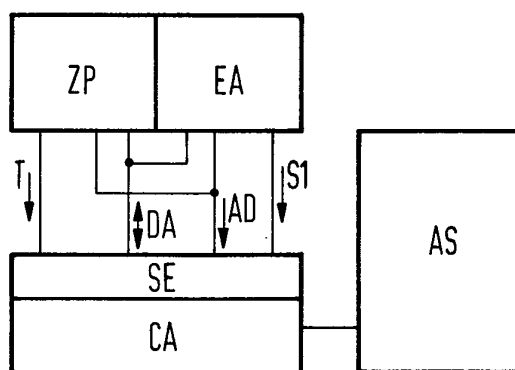
⑫ Inhaber:
Siemens Aktiengesellschaft, Berlin und München,
München 2 (DE)

⑫ Erfinder:
Wolfgang Feissel, München 70 (DE)

⑫ Vertreter:
Siemens-Albis Aktiengesellschaft, Zürich

⑫ Datenverarbeitungsanlage mit einem sowohl einem Zentralprozessor als auch einem Ein-/Ausgabeprozessor zugeordneten Pufferspeicher.

⑫ Die Datenverarbeitungsanlage enthält einen Zentralprozessor (ZP) und einen Ein-/Ausgabeprozessor (EA), die über einen gemeinsamen Pufferspeicher (CA) Daten mit einem Arbeitsspeicher (AS) austauschen. Dem Pufferspeicher (CA) ist ferner eine Steuereinheit (SE) zugeordnet, die aufgrund eines Ersetzungsalgorithmus eine Freigabe von wechselnden Teilbereichen des vollen Pufferspeichers für noch nicht zwischengespeicherte Daten in der Weise steuert, dass der Ersetzungsalgorithmus während einer Ein-/Ausgabeoperation blockiert und nur der vor dem Blockieren bereits festgelegte Teilbereich freigegeben wird.



PATENTANSPRÜCHE

1. Datenverarbeitungsanlage mit einem Zentralprozessor, einem Ein-/Ausgabeprozessor, einem Arbeitsspeicher mit grosser Speicherkapazität und grosser Zugriffszeit, einem Pufferspeicher mit kleiner Speicherkapazität und kleiner Zugriffszeit, über den die Prozessoren mit dem Arbeitsspeicher Daten austauschen, sowie mit einer dem Pufferspeicher zugeordneten Steuereinheit, die unter anderem auch aufgrund eines Ersetzungsalgorithmus ein Freigeben von wechselnden Teilbereichen des vollen Pufferspeichers für noch nicht zwischengespeicherte Daten steuert, dadurch gekennzeichnet, dass die dem Pufferspeicher (CA) zugeordnete Steuereinheit (SE) Schaltglieder (F1 bis Fn, U1 bis Un) enthält, die während einer Ein-/Ausgabeoperation den Ersetzungsalgorithmus blockieren, mit dem Daten im Pufferspeicher ersetzt werden, so dass nur der durch den Ersetzungsalgorithmus vor dem Blockieren bereits festgelegte Teilbereich des Pufferspeichers zum Zwischenspeichern der Ein-/Ausgabedaten freigegeben ist.

2. Datenverarbeitungsanlage nach Anspruch 1, bei der der Pufferspeicher nach dem Kongruenzklassenprinzip organisiert und aus einer Mehrzahl von in Klassen gegliederten Speicherbänken aufgebaut ist, dadurch gekennzeichnet, dass die Steuereinheit (SE) des Pufferspeichers (CA) derart ausgebildet ist, dass sie während einer Ein-/Ausgabeoperation das Einspeichern von Daten nur in einen einzigen der Speicherbänke (SB1 bis SBn) freigibt.

3. Datenverarbeitungsanlage nach Anspruch 1, dadurch gekennzeichnet, dass die Steuereinheit (SE) derart ausgebildet ist, dass sie während einer Ein-/Ausgabeoperation das Einspeichern der Daten nur in eine einzige Zelle des Pufferspeichers (CA) freigibt.

4. Datenverarbeitungsanlage nach Anspruch 2, dadurch gekennzeichnet, dass die Zellen im Pufferspeicher (CA) an einer vorgegebenen Stelle eine Markierungsstelle für ein Binärzeichen aufweisen, die anzeigt, ob der zugehörige Dateneintrag in dieser Zelle vom Zentralprozessor (ZP) angefordert worden ist und damit die nächste Speicherbank festlegt, die, bezogen auf die Klasse, bei Bedarf zu ersetzen ist und dass die Steuereinheit (SE) während jeder Ein-/Ausgabeoperation verhindert, dass durch ein solches Binärzeichen (R1 bis Rn) nach einem Zugriff des Ein-/Ausgabeprozessors (EA) zu einem Eintrag dieser als benutzt ausgewiesen wird.

Die Erfindung bezieht sich auf eine Datenverarbeitungsanlage gemäss dem Oberbegriff des Patentanspruches 1.

Es ist bereits allgemein bekannt, bei leistungsfähigen, kommerziellen Datenverarbeitungsanlagen zwischen einem Arbeitsspeicher mit grosser Speicherkapazität und verhältnismässig grosser Zugriffszeit und einem Zentralprozessor einen Pufferspeicher mit kleiner Speicherkapazität und sehr kleiner Zugriffszeit anzuordnen. Der Pufferspeicher, der auch als Cache bezeichnet wird, hat in erster Linie die Aufgabe, möglichst viele Lesezugriffe des Zentralprozessors zum Arbeitsspeicher zu befriedigen. Bei einem Lesezugriff gibt der Zentralprozessor zunächst die Adresse der gewünschten Daten an das Cache. Falls das Cache das gewünschte Datenwort enthält, wird dieses sofort an den Zentralprozessor weitergegeben. Anderenfalls muss die Adresse an den Arbeitsspeicher weitergeleitet werden und die Daten werden zunächst vom Hauptspeicher an das Cache und danach von diesem zum Zentralprozessor übertragen. Da Daten bei Verarbeitungsprozessoren häufig für eine bestimmte Zeit mehr als einmal benutzt werden, lässt sich die Zugriffszeit im Mittel verkürzen.

Falls jedoch der Zentralprozessor über den Pufferspeicher

und der Ein-/Ausgabeprozessor, also mehr als eine Verarbeitungsrichtung direkt zum Arbeitsspeicher Zugriff haben, können Schreibkonfliktsituationen entstehen, die darin bestehen, dass unter der gleichen Adresse im Arbeitsspeicher und im Cache verschiedene Daten stehen, so dass die Datenverarbeitungsanlage falsch arbeitet. Ausserdem muss bei einer solchen Struktur eines Speichersystems eine Möglichkeit vorgesehen werden, bei vollem Pufferspeicher Teile freizugeben, damit neu angeordnete Daten in den Pufferspeicher aufgenommen werden können. Wenn mehrere Prozessoren zu dem Speichersystem Zugriff haben, kann eine Verarbeitungseinrichtung dann aber benachteiligt werden, wenn häufiger zugreifende Prozessoren deren noch benötigte Daten verdrängen und selbst den Pufferspeicher voll belegen.

Um dies zu verhindern, ist aus der DE-OS 22 61 586 eine Speichereinrichtung bekannt, zu der ein Zentralprozessor, ein Wartungsfeld und ein Kanalwerk bzw. daran angeschlossene periphere Einheiten über einen Arbeitsspeicherkoordinator Zugriff haben und die aus einem grossen Arbeitsspeicher und einem vorgeschalteten kleinen Pufferspeicher besteht. Der Arbeitsspeicherkoordinator stellt dabei eine Steuereinheit des Arbeitsspeichersystems dar, der grundsätzlich jeden zunächst an den Pufferspeicher gerichteten Zugriff vermittelt, der seinerseits bei Bedarf zum Arbeitsspeicher zugreift.

Der Pufferspeicher ist nun aus zwei oder mehreren, sich nicht überlappenden Teilbereichen aufgebaut. Diese Teilbereiche sind den verschiedenen, Speicherzugriffe anfordernden Einrichtungen der datenverarbeitenden Anlage fest zugeordnet. Beispielsweise ist ein erster Bereich vorgesehen, der für Ein-/Ausgabedaten reserviert ist, die über die Selektorkanäle übertragen werden, der Speicherplatz in einem zweiten Bereich wird dann dem Zentralprozessor und den Multiplexkanälen zur Verfügung gestellt. Neben dieser nur beispielhaft genannten Unterteilung des Pufferspeichers und ihrer Zweckbestimmung sind aber auch andere Unterteilungen denkbar.

Allen solchen Formen für mögliche Unterteilungen des Pufferspeichers in Teilbereiche, die jeweils fest einer anfordernden Verarbeitungseinrichtung zugeordnet sind, ist aber der Nachteil der mangelnden Flexibilität in der Verteilung der Pufferspeicherkapazität gemeinsam. So lässt sich die an sich bereits geringe Speicherkapazität nicht optimal ausnutzen. Zwar wird erreicht, dass keine der anfordernden Verarbeitungseinrichtungen auf Kosten einer anderen den Pufferspeicher in beliebigem Umfang belegen kann, andererseits aber das Konzept des Pufferspeichers als eines wahlfreien eingeschränkt.

Der Erfindung liegt daher die Aufgabe zugrunde, eine Datenverarbeitungsanlage gemäss dem Oberbegriff des Patentanspruches 1 zu schaffen, bei der der in Teilbereiche unterteilte Pufferspeicher ohne strenge Zuordnung einzelner Teilbereiche zu den Prozessoren im Datenverkehr mit diesen Prozessoren so betrieben wird, dass dem tätigen Ein-/Ausgabeprozessor bei vollem Pufferspeicher nur ein maximaler, physikalisch dann aber nicht festgelegter Anteil der Speicherkapazität zur Verfügung steht, so dass Schreibkonfliktsituationen zwischen den beiden Prozessoren vermieden werden.

Diese Aufgabe wird erfindungsgemäss durch die im Kennzeichen des Patentanspruches 1 beschriebenen Merkmale gelöst.

Mit einfachen Mitteln gelingt es so, auch für den Ein-/Ausgabeprozessor den Pufferspeicher nutzbar zu machen, ohne in seiner Wirksamkeit für den Zentralprozessor beeinträchtigt zu sein. Es wäre an sich vorstellbar, eine eigene Speicherplatzverwaltung einzuführen, die auch dieses Ziel erreicht, jedoch wäre dies viel aufwendiger. Da es sehr wichtig ist, im Pufferspeicher häufig benutzte Daten möglichst lange halten zu können, wird durch das Einspeichern der Daten bei einer Ein-/Ausgabeoperation in vorgegebene Bereiche des Pufferspeichers verhindert, dass bei schnellen und lange andauernden Ausgabeoperationen

die vom Zentralprozessor benötigten Daten im Pufferspeicher durch die auszugebenden Daten gelöscht werden.

Die Wirksamkeit der erfindungsgemässen Lösung lässt darüber hinaus auch auf den Arbeitsspeicher bezogene Strukturen des Pufferspeichers zu. So kann der nach dem Kongruenzklassenprinzip organisierte Pufferspeicher aus einer Mehrzahl von gleichartigen, jeweils in Klassen gegliederte, Speicherbänken aufgebaut sein. Wenn in diesem Falle beispielsweise der Ersetzungsalgorithmus für die Daten darin besteht, diejenigen Daten, die zuerst eingespeichert wurden, auch als erste wieder zu überschreiben, ist es von Vorteil, wenn die Steuereinheit gemäss dem Kennzeichen des Patentanspruchs 2 ausgebildet ist.

Ein besonders geringer Anteil der vom Zentralprozessor benötigten Daten im Pufferspeicher wird bei Ein-/Ausgabeoperationen gelöscht, wenn die Steuereinheit derart ausgebildet ist, dass sie während einer Ein-/Ausgabeoperation das Einspeichern der Daten nur in eine einzige Zelle des Pufferspeichers freigibt.

Bei einer Steuerung des Ersetzungsalgorithmus durch den Zentralprozessor ist eine Lösung gemäss dem Kennzeichen des Patentanspruchs 4 besonders vorteilhaft.

Im folgenden wird ein Ausführungsbeispiel der Datenverarbeitungsanlage gemäss der Erfindung anhand von Zeichnungen erläutert. Es zeigen:

Fig. 1 ein Blockschaltbild der Datenverarbeitungsanlage,

Fig. 2 ein Blockschaltbild einer Steuereinheit im Pufferspeicher.

Das in Fig. 1 dargestellte Blockschaltbild zeigt einen Zentralprozessor ZP und einen Ein-/Ausgabeprozessor EA, der integrierter Bestandteil des Zentralprozessors ZP sein kann. Der Zentralprozessor ZP und der Ein-/Ausgabeprozessor EA sind über eine Pufferspeicher CA, auch Cache genannt, mit einem Arbeitsspeicher AS verbunden. Bei einem Lesezugriff gibt der Zentralprozessor an das Cache zunächst ein Adresswort AD ab, das die Adresse eines gewünschten Datenwortes angibt. Falls das gewünschte Datenwort im Cache CA enthalten ist, wird es durch Signale DA an den Zentralprozessor abgegeben. Andernfalls wird das Adresswort an der Arbeitsspeicher AS abgegeben. Anschliessend wird vom Arbeitsspeicher AS zum Cache CA ein Datenblock übertragen, der aus mehreren Datenwörtern besteht und das gewünschte Datenwort enthält. Danach wird das gewünschte Datenwort zum Zentralprozessor übertragen.

In ähnlicher Weise werden Daten vom Arbeitsspeicher über das Cache CA zum Ein-/Ausgabeprozessor EA übertragen. Bei schnellen und lange andauernden Ein-/Ausgabeoperationen ist es jedoch wichtig, dass vom Zentralprozessor ZP benötigte Daten im Cache CA nicht durch Ein-/Ausgabedaten gelöscht werden, da es zum Erreichen einer grossen Treffer rate im Cache CA äusserst wichtig ist, häufig benutzte Daten möglichst lange im Cache CA zu halten.

Im Cache CA ist eine Steuereinheit SE vorgesehen, die den Ersetzungsalgorithmus festlegt, nachdem im Cache CA vorhandene Daten durch neue Daten ersetzt werden. Um sicherzustellen, dass Ein-/Ausgabedaten die vom Zentralprozessor ZP benötigten Daten nicht löschen, sorgt die Steuereinheit SE dafür, dass die Ein-/Ausgabedaten jeweils nur in vorgegebenen Bereichen des Caches CA eingespeichert werden. Beispielsweise kann die Steuereinheit SE derart ausgebildet sein, dass die Ein-/Ausgabedaten jeweils nur in eine einzige Zelle des Caches CA eingespeichert werden. Falls das Cache CA beispielsweise aus einer Mehrzahl von Speicherbänken aufgebaut ist und die Daten zum Arbeitsspeicher AS in das Cache CA jeweils nach dem bekannten Kongruenzklassenprinzip eingespeichert werden, kann es auch vorteilhaft sein, wenn die Ein-/

Ausgabedaten jeweils nur innerhalb einer einzigen Bank eingespeichert werden. Im ungünstigsten Fall werden dann vom Zentralprozessor ZP benötigte Daten nur jeweils in einer einzigen Speicherbank gelöscht. In beiden Fällen gibt der Ein-/Ausgabeprozessor EA ein Signal S1 an die Steuereinheit SE ab, mit dem der jeweilige Betriebszustand eingestellt wird.

Eine weitergehende Möglichkeit zur Steuerung des Ersetzungsalgorithmus besteht darin, dass an einer vorgegebenen Stelle im Cache CA, dem sogenannten Flagfeld eines jeden Puffereintrages ein Binärzeichen eingeführt wird, aus dem zu ersehen ist, ob der zugehörige Eintrag im Cache CA vom Zentralprozessor ZP benutzt worden ist. Bei Lesedatenzugriffen kann dann verhindert werden, dass bei Benutzung eines Eintrages dieser im Flagfeld über das Binärzeichen als benutzt ausgewiesen wird.

Das in Fig. 2 dargestellte Schaltbild der Steuereinheit SE ist derart ausgebildet, dass bei Zugriffen des Zentralprozessors ZP zum Cache CA jeweils diejenigen Daten als erste gelöscht werden, die als erste in das Cache CA eingespeichert worden sind. Es wird angenommen, dass das Cache CA aus einer Mehrzahl von Speicherbänken SB1 bis SBn gebildet wird, von denen in Fig. 2 jeweils nur die Tag- und Flagfelder und nicht die Datenfelder dargestellt sind.

Die Steuereinheit besteht aus einem Verstärker V, einer der Anzahl der Speicherbänke entsprechenden Anzahl von Verknüpfungsgliedern U1 bis Un und Flipflops F1 bis Fn.

Zunächst wird angenommen, dass Daten zwischen dem Arbeitsspeicher AS und dem Zentralprozessor ZP übertragen werden. Das Signal S1 hat in diesem Fall den Binärwert 0. Ein Signal am invertierenden Ausgang des Verstärkers V hat damit den Binärwert 1 und jeweils die linken UND-Glieder in den Verknüpfungsgliedern U1 bis Un sind freigegeben. Vom Arbeitsspeicher AS werden die Daten in ein Datenfeld einer der Speicherbänke SB1 bis SBn eingetragen. Gleichzeitig wird in ein zugehöriges Tagfeld durch die Signale AD die Seitenadresse der jeweiligen Daten eingespeichert. Im Flagfeld, in dem eine Mehrzahl von Binärzeichen gespeichert werden können, wird beispielsweise durch eines der Binärzeichen V1 bis Vn angezeigt, ob die zugehörigen Daten gültig sind. Mit Hilfe eines weiteren Binärzeichens R1 bis Rn werden die jeweils zuletzt für den Zentralprozessor ZP eingetragenen Daten gekennzeichnet. Beispielsweise möge in der letzten Speicherzelle der Speicherbank SB1 das Binärzeichen R1 den Binärwert 1 haben. Mit einem Übernahmeimpuls T, der vom Zentralprozessor ZP abgegeben wird und der die Reihenfolge zwischen speichert, nach der die Einträge im Cache CA für den Zentralprozessor ZP erfolgt waren, nimmt das Flipflop F1 den Binärwert 1 an und zeigt damit an, dass der nächstfolgende Eintrag in die Speicherbank SB2 eingespeichert wird und das Binärzeichen R2 nimmt den Binärwert 1 an. Auf diese Weise werden die Daten nacheinander in die Speicherbänke SB1 bis SBn eingespeichert.

Falls durch das vom Ein-/Ausgabeprozessor EA abgegebene Signal S1 eine Ein-/Ausgabeoperation angezeigt wird, nimmt das Signal am nichtinvertierenden Ausgang des Verstärkers V den Binärwert 1 an und die jeweils rechten UND-Glieder der Verknüpfungsglieder U1 bis Un werden freigegeben. Bei einer Einspeicherung von Daten wird damit nur diejenige Speicherbank SBx belegt, bei der das Binärzeichen R(x-1) gerade den Binärwert 1 hat. Auf diese Weise wird sichergestellt, dass bei schnellen und andauernden Ein-/Ausgabeoperationen nur die Daten in einer Zelle einer Bankklasse, d.h. nur in einer Speicherbank, überschrieben werden, während die übrigen, vom Zentralprozessor ZP gegebenenfalls benötigten Daten in der jeweiligen Klasse nicht verändert werden.

Fig. 1

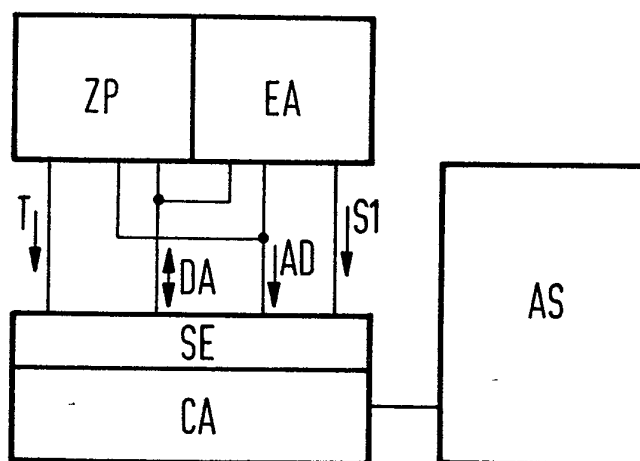


Fig. 2

