



(12) **Offenlegungsschrift**

(21) Aktenzeichen: **10 2020 128 959.6**  
(22) Anmeldetag: **03.11.2020**  
(43) Offenlegungstag: **10.06.2021**

(51) Int Cl.: **G06F 13/00** (2006.01)  
**G11C 5/00** (2006.01)  
**H01L 23/538** (2006.01)

(30) Unionspriorität:  
**16/701,739**                      **03.12.2019**      **US**

(71) Anmelder:  
**Intel Corporation, Santa Clara, Calif., US**

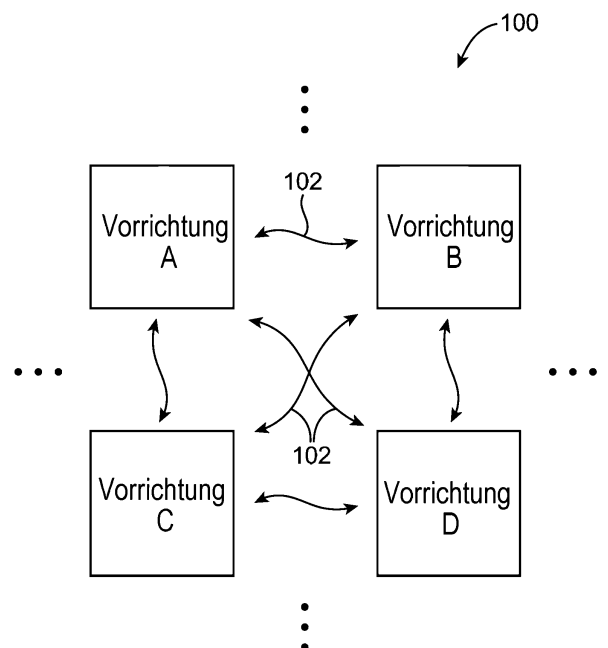
(74) Vertreter:  
**BOEHMERT & BOEHMERT Anwaltspartnerschaft  
mbB - Patentanwälte Rechtsanwälte, 28209  
Bremen, DE**

(72) Erfinder:  
**Sethuraman, Saravanan, Santa Clara, CA, US;**  
**Morris, Tonia, Santa Clara, CA, US;** **Lai, Siaw  
Kang, Santa Clara, CA, US;** **Lim, Yee Choong,  
Santa Clara, CA, US;** **Ong, Yu Ying, Santa Clara,  
CA, US**

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.**

(54) Bezeichnung: **VERFAHREN UND VORRICHTUNGEN ZUR VERWALTUNG DES THERMISCHEN VERHALTENS  
IN MEHRCHIP-KAPSELUNGEN**

(57) Zusammenfassung: Eine integrierte Schaltung kann Schaltkreise zum Zugreifen auf einen zugeordneten Die-Stapel umfassen. Die Schaltkreise können Temperaturinformationen sowie mehrere Betriebsparameter empfangen, die dabei helfen, zu bestimmen, ob es wünschenswert sein kann, Zugriffsbefehle oder Anforderungen zu einem oder mehreren Dies in dem Stapel umzurouten. Die Schaltkreise können einen intelligenten Crossbarschalter umfassen, der eine Adressenübersetzung oder Hashing-Funktion implementiert, um dabei zu helfen, die logische Benutzeradresse auf einen physischen Adressenraum abzubilden. Das Durchführen von thermisch bewusster Verkehrsverwaltung auf diese Weise kann sicherstellen, dass in dem System akzeptable Timingreserven aufrechterhalten werden, um die Wahrscheinlichkeit von Fehlern zu minimieren.



## Beschreibung

### Hintergrund

**[0001]** Das Vorliegende betrifft allgemein integrierte Schaltungen und insbesondere integrierte Schaltungskapselungen mit mehr als einem integrierten Schaltungs-Die.

**[0002]** Eine integrierte Schaltungskapselung umfasst typischerweise einen integrierten Schaltungs- bzw. IC-Die und ein Substrat, auf dem der Die angebracht ist. Der integrierte Schaltungs-Die ist mittels Bonddrähten oder Lötunkten mit dem Substrat gekoppelt. Signale von dem integrierten Schaltungs-Die breiten sich durch die Bonddrähte oder Lötunkte zum Substrat aus.

**[0003]** Da die Anforderungen an integrierter Schaltungstechnologie weiterhin selbst die Gewinne übertreffen, die durch immer weiter abnehmende Vorrichtungsdimensionen gewährt werden, verlangt eine zunehmende Anzahl von Anwendungen eine gekapselte Lösung mit mehr Integration, als in einem Silizium-Die möglich ist. Um diesem Bedarf gerecht zu werden, kann mehr als ein IC-Die in einer integrierten Schaltungskapselung (d. h. einer Mehrchip-Kapselung) platziert werden. Da verschiedene Arten von Vorrichtungen verschiedenen Arten von Anwendungen dienen, können in einigen Systemen mehr IC-Dies erforderlich sein, um die Anforderungen von Hochleistungsanwendungen zu erfüllen. Um bessere Leistungsfähigkeit und höhere Dichte zu erhalten, kann dementsprechend eine Mehrchip-Kapselung mehrere lateral entlang derselben Ebene angeordnete Dies umfassen oder kann mehrere übereinander gestapelte Dies umfassen.

**[0004]** Bei solchen Mehrchip-Kapselungslösungen umfasst eine Mehrchip-Kapselung typischerweise einen Hauptlogik-Die und einen Hochbandbreitenspeicher- bzw. HBM-Stapel, die beide auf einem gemeinsamen Interposer-Substrat in der Mehrchip-Kapselung montiert sind. Der HBM-Stapel könnte vier vertikal übereinandergestapelte Speicher-Dies umfassen. Der HBM-Stapel wird typischerweise in dichter Nähe zum Hauptlogik-Die oder einem anderen Beschleuniger-Die in derselben Kapselung gebildet und hat somit einzigartige thermische Beschränkungen.

**[0005]** Wenn die Betriebstemperatur des HBM-Stapels eine bestimmte Schwelle übersteigt, verschlechtert sich die HBM-Stapel-Leistungsfähigkeit aufgrund von timingbezogenen Problemen. Mit ansteigender Temperatur nimmt die Timingreserve tendenziell ab, was zu einer Zunahme von Speicherfehlern führt. In einem Prüfungsvorgang wurde entdeckt, dass bei 60% von Bauteilen beim Auslesen von Daten aus einem HBM-Stapel, der geprüft wird, Bitfehler entstanden. Die Menge an Bitfehlern wird sich nur verschlimmern,

wenn mehr Speicher-Dies zu dem HBM-Stapel hinzugefügt werden und wenn Anwendungen weiter auf Betrieb mit höherer Bandbreite drängen, um dadurch das Risiko von Hochtemperatur-Timingverstößen zu vergrößern.

**[0006]** In diesem Kontext entstehen die hier beschriebenen Ausführungsformen.

### Figurenliste

**Fig. 1** ist eine Darstellung eines beispielhaften Systems integrierter Schaltungsvorrichtungen, die zur Kommunikation miteinander betreibbar sind, gemäß einer Ausführungsform.

**Fig. 2** ist eine Querschnitts-Seitenansicht einer beispielhaften Mehrchip-Kapselung gemäß einer Ausführungsform.

**Fig. 3** ist eine Darstellung eines beispielhaften Systems, das einen mit einem zugeordneten Speicher-Die-Stapel gekoppelten Haupt-Die und einen Hostprozessor umfasst, gemäß einer Ausführungsform.

**Fig. 4** ist eine Darstellung eines beispielhaften Pseudokanals, der mehrere Speicherbänke umfasst, gemäß einer Ausführungsform.

**Fig. 5** ist eine Darstellung einer beispielhaften Speichercontrollerschaltung, die dafür ausgelegt ist, thermische Drosselung auf der Basis mehrerer Betriebsparameter durchzuführen, gemäß einer Ausführungsform.

**Fig. 6** ist eine Darstellung, die zeigt, wie Temperaturinformationen von mehr als einem Die in einem Die-Stapel erhalten werden können, gemäß einer Ausführungsform.

**Fig. 7** ist eine Darstellung von beispielhaften intelligenten Schaltnetzwerken, die dafür ausgelegt sind, Adressenübersetzung/-hashing durchzuführen, gemäß einer Ausführungsform.

**Fig. 8** ist eine Darstellung, die zeigt, wie ein Adressenraum in eine statische Region und eine dynamische Region aufgeteilt werden kann, gemäß einer Ausführungsform.

**Fig. 9** ist ein Flussdiagramm beispielhafter Schritte zum Betrieb der in Verbindung mit **Fig. 2-7** gezeigten Schaltungen gemäß einer Ausführungsform.

### Ausführliche Beschreibung

**[0007]** Die vorliegenden Ausführungsformen betreffen verschiedene Verfahren zur Verwaltung der thermischen Eigenschaften in einer Mehrchip-Kapselung, die einen Haupt-Die umfasst, der betreibbar ist, um über entsprechende Kanäle mit einem zugeordneten Die-Stapel zu kommunizieren. Bei einer geeig-

neten Anordnung kann eine thermische Steuerungsschaltung auf dem Haupt-Die dafür ausgelegt sein, mehrere Betriebsparameter von einem oder mehreren Dies im Die-Stapel zu empfangen und selektiv Zugriffe auf äußeren Dies zu drosseln. Ein Die kann als „heiß“ angesehen werden, wenn bestimmt wird, dass seine Temperatur größer als eine vorbestimmte Temperaturschwelle ist (z. B. wenn die Temperaturmessung 30°C, 40°C, 50°C, 60°C, 70°C, 80°C, 90°C, 100°C oder mehr überschreitet). Ein Die kann als „kalt“ oder nicht durch Timingverstöße gefährdet angesehen werden, wenn die Temperatur unter der vorbestimmten thermischen Schwelle liegt. Zum Beispiel kann die Steuerung die Anzahl der Befehle, die zu jedem Die innerhalb eines gegebenen Zeitfensters gesendet werden, schätzen und die Fälle, bei denen sie wahrscheinlich eine vorbestimmte thermische Grenze überschreiten würde, kennzeichnen.

**[0008]** Die thermische Steuerungsschaltung kann außerdem mit Hilfe von Adressenübersetzung (z. B. durch Snooping der an die verschiedenen Kanäle gehenden Adressen und entsprechendes Befähigen der Timer) und intelligente Crossbar-Swizzling-Techniken einen beschäftigten Verkehrskanal auf einen weniger beladenen Kanal umrouten, um ein akzeptables thermisches Profil auf dem Die-Stapel aufrechtzuerhalten. Die Adressenübersetzung kann die physische Adresse an Schreibvorgängen umabbilden und kann diese Umabbildung bis zum Schreibvorgang in dieselbe Adresse bewahren.

**[0009]** Die Steuerungsschaltung kann auch dafür ausgelegt sein, die Temperatur eines oder mehrerer Dies in dem Die-Stapel zu erhalten oder anderweitig genau zu schätzen/vorherzusagen. Das Abrufen der Temperatur nur von dem untersten Die in dem Die-Stapel reicht nicht dafür aus, richtige Entscheidungen für den gesamten Die-Stapel zu treffen. Somit kann die Steuerung Temperaturinformationen nicht nur von dem untersten Die, sondern auch vom obersten Die in dem Die-Stapel abrufen, um einen entsprechenden Temperaturgradienten zu erhalten, um dabei zu helfen, die Temperatur beliebiger der zwischen dem obersten und dem untersten Die sitzenden mittleren Dies zu schätzen. Nach dem Detektieren des Temperaturmesswerts für jeden der Dies in dem Die-Stapel können die Befehle so geschwenkt werden, dass öfter auf kalte Dies als auf die heißeren Dies zugegriffen wird. Im Fall eines gespiegelten Speichersubsystems kann die Steuerung nur Lesevorgänge aus der gespiegelten Kopie, statt der Hauptkopie, durchführen, um dabei zu helfen, Temperaturzunahmen in der Hauptkopie zu vermeiden, statt den Lesevorgang über andere Dies im Stapel zu verteilen.

**[0010]** Auf der Systemebene sollten die Anwendungen wissen, wo zukünftige Schreibvorgänge platziert werden sollten, um thermische Probleme in den mehreren Die-Stapeln im System zu vermeiden. Die ther-

mische Steuerung kann außerdem eine gewisse Art von Befehls-Neueinteilungsmechanismus (z. B. eine Reigenbewegung) implementieren, um zu verhindern, dass irgendein Die in dem HBM sich zu schnell erhitzt.

**[0011]** Wenn es auf diese Weise konfiguriert ist und betrieben wird, kann ein Systemabsturz vermieden werden (d. h. durch Entfernen der Hauptgrundursache von Fehlern aufgrund des Anstiegs der Temperatur), während missionskritische Anwendungen auf Systemen laufengelassen werden, die sich sowohl für existierende als auch zukünftige Produkte auf die Verwendung von Kapselungen mit gestapelten Dies verlassen könnten. In einem Beispiel kann der Die-Stapel Teil eines Datenzentrums sein, das vielfältige verschiedene Anforderungen verarbeitet. Das Datenzentrum kann zum Beispiel eine Datenverarbeitungsanforderung zum Durchführen von Verschlüsselung, Entschlüsselung, Maschinenlernen, Videoverarbeitung, Spracherkennung, Bilderkennung, Datenkomprimierung, Datenbank-Sucheinstufung, Bioinformatik, Netzwerksicherheitsmuster-Identifikation, räumliche Navigation oder eine gewisse andere spezialisierte Aufgabe empfangen.

**[0012]** Für Fachleute ist erkennbar, dass die vorliegenden beispielhaften Ausführungsformen ohne einige oder alle dieser spezifischen Details praktiziert werden können. In anderen Fällen wurden wohlbekannte Operationen nicht im Detail beschrieben, um die vorliegenden Ausführungsformen nicht unnötigerweise zu verschleiern.

**[0013]** Fig. 1 ist eine Darstellung eines beispielhaften Systems **100** miteinander verbundener elektronischer Vorrichtungen. Das System miteinander verbundener elektronischer Vorrichtungen kann mehrere elektronische Vorrichtungen aufweisen, wie etwa Vorrichtung A, Vorrichtung B, Vorrichtung C, Vorrichtung D, sowie Verbindungsressourcen **102**. Die Verbindungsressourcen **102**, wie etwa leitfähige Leitungen und Busse, optische Interconnect-Infrastruktur oder verdrahtete und drahtlose Netzwerke mit optionalen Zwischen-Schaltnetzwerken, können verwendet werden, um Signale von einer elektronischen Vorrichtung zu einer anderen elektronischen Vorrichtung zu senden oder um Informationen von einer elektronischen Vorrichtung zu mehreren anderen elektronischen Vorrichtungen rundzusenden. Zum Beispiel kann ein Sender in Vorrichtung B Datensignale zu einem Empfänger in Vorrichtung C senden. Ähnlich kann Vorrichtung C einen Sender verwenden, um Daten zu einem Empfänger in Vorrichtung B zu senden.

**[0014]** Die elektronischen Vorrichtungen können eine beliebige geeignete Art von elektronischer Vorrichtung sein, die mit anderen elektronischen Vorrichtungen kommuniziert. Beispiele für solche elektronischen Vorrichtungen wären integrierte Schaltungen

mit elektronischen Komponenten und Schaltungen wie Anlogschaltungen, Digitalschaltungen, Mischsignalschaltungen, in einer einzigen Kapselung gebildete Schaltungen, in verschiedenen Kapselungen untergebrachte Schaltungen, Schaltungen, die auf einer PCB (Leiterplatte) miteinander verbunden sind, usw.

**[0015]** Da sich die integrierte Schaltungstechnologie in Richtung kleinerer Vorrichtungsdimensionen skaliert, verbessert sich die Vorrichtungsleistungsfähigkeit weiter auf Kosten eines vergrößerten Stromverbrauchs. Als Bemühung, den Stromverbrauch zu reduzieren, kann mehr als ein Die in einer einzigen integrierten Schaltungskapselung (d. h. einer Mehrchip-Kapselung) platziert werden. Da verschiedene Arten von Vorrichtungen verschiedenen Arten von Anwendungen dienen, können in einigen Systemen mehr IC-Dies erforderlich sein, um die Anforderungen von Hochleistungsanwendungen zu erfüllen. Um bessere Leistungsfähigkeit und höhere Dichte zu erhalten, kann dementsprechend eine Mehrchip-Kapselung mehrere lateral entlang derselben Ebene angeordnete Dies umfassen oder kann mehrere übereinander gestapelte Dies umfassen.

**[0016]** Fig. 2 ist eine Querschnitts-Seitenansicht einer beispielhaften Mehrchip-Kapselung **200**, die ein Kapselungssubstrat **206**, einen auf dem Kapselungssubstrat **206** angebrachten Interposer-Die **204** und mehrere über dem Interposer **204** angebrachte Dies umfasst. Der Interposer kann **204** kann eine aktive Vorrichtung (z. B. ein Die mit Transistoren oder anderen aktiven Komponenten, die mit Strom versorgt werden können) oder eine passive Vorrichtung (z. B. ein Die, der keine aktiven Komponenten enthält) sein. Ein Haupt-Die **202** (z. B. eine anwendungsspezifische integrierte Schaltung, eine programmierbare integrierte Schaltung, wie etwa ein am Einsatzort programmierbares Gatearray, eine Grafikverarbeitungseinheit oder ein anderer Prozessor) und ein oder mehrere Die-Stapel **220** können auf dem Interposer **204** angebracht werden. Der Haupt-Die **202** wird manchmal als der Hauptlogik-Die bezeichnet.

**[0017]** In dem Beispiel von Fig. 2 können mindestens ein erster Die-Stapel **220-1** und ein zweiter Die-Stapel **220-2** auf dem Interposer **204** angebracht sein. Der erste Die-Stapel **220-1** umfasst zwei integrierte Schaltungs-Dies vertikal mit Bezug aufeinander gestapelt, während der zweite Die-Stapel **220-2** vier integrierte Schaltungs-Dies vertikal mit Bezug aufeinander gestapelt umfasst. Diese Anordnung ist lediglich veranschaulichend. Im Allgemeinen kann eine beliebige Anzahl von Die-Stapeln **220** auf dem Interposer **204** angebracht und mit dem Haupt-Die **202** gekoppelt werden, wobei jeder Die-Stapel **220** mindestens zwei vertikal gestapelte Dies, mindestens drei vertikal gestapelte Dies, mindestens vier vertikal gestapelte Dies, 4-8 vertikal gestapelte Dies, 8-

12 vertikal gestapelte Dies, 12-16 vertikal gestapelte Dies oder mehr als 16 vertikal gestapelte Dies umfassen kann. Die verschiedenen Dies in jedem Die-Stapel können über manchmal als Silizium-Durchkontaktierungen (TSV) bezeichnete Kommunikationspfade oder andere geeignete Verbindungsstrukturen miteinander kommunizieren.

**[0018]** An der oberen Oberfläche des Interposers **204** können als Schnittstelle mit dem Haupt-Die **202** und den Die-Stapeln **220** Mikrohögel **212** gebildet werden. An der unteren Oberfläche des Interposers **204** können Hügel wie Hügel **214** der kontrollierten Kollaps-Chipverbindung (C4) (manchmal als Flip-Chip-Hügel bezeichnet) gebildet werden. Im Allgemeinen weisen C4-Hügel **214** (z. B. zur Schnittstelle mit Komponenten außerhalb der Kapselung verwendete Hügel) verglichen mit den Mikrohögel **212** (z. B. zur Schnittstelle mit anderen Dies in derselben Mehrchip-Kapselung verwendeten Hügel) wesentlich größere Größe auf. Die Anzahl der Mikrohögel **212** ist auch im Allgemeinen größer als die Anzahl der Flip-Chip-Hügel **214** (z. B. kann das Verhältnis der Anzahl von Mikrohögel **212** zur Anzahl von C4-Hügel größer als 2:1, 5:1, 10:1 usw. sein).

**[0019]** Es werden hier manchmal Vorrichtungskonfigurationen als Beispiel beschrieben, bei denen ein oder mehrere Die-Stapel **220** in der Mehrchip-Kapselung ein Speicherstapel (z. B. ein Stapel von Speichervorrichtungen) ist, wie etwa der DRAM-Speicherstapel mit Hochbandbreitenspeicher (HBM) gemäß JEDEC JESD235. Dies ist jedoch lediglich veranschaulichend. Die hier beschriebenen Techniken und Verbesserungen können nicht nur auf gestapelte Speicherchips angewandt werden, sondern auch auf gestapelte Sendeempfeängerchips, gestapelte Vernetzungsadapterchips, gestapelte diskrete Komponenten, gestapelte Konfigurationsvorrichtungen, gestapelte Debugging-Chips und/oder andere Arten von gestapelten Speicherungs- und Verarbeitungskomponenten in einer Mehrchips-Kapselung erweitert werden.

**[0020]** Wie im Hintergrundabschnitt erwähnt, kommt es bei Produkten, die Die-Stapel, wie etwa HBM-Stapel, umfassen, oft zu timingbezogenen Problemen, wenn die Temperatur eine bestimmte Schwelle überschreitet. Wenn die Temperatur zu hoch ist, verschlechtert sich die Timingreserve und das System beginnt, Speicherfehler zu erfahren. Herkömmliche Ansätze zur Minderung von thermischen Problemen umfassen nur das Lesen der Temperatur aus nur dem untersten Basis-Die im HBM-Stapel, selektives Herunterfahren des zugeordneten Kommunikationskanals (was unerwünscht ist), Verwürfeln der Speicherdaten und Verwenden von Kühlkörperlösungen. Da die Stapeldichte weiter zunimmt und Hersteller weiter mehr Dies hinzufügen, um die Stapelhöhe zu vergrößern, werden die herkömmlichen Methoden

nicht nur in der Lage sein, solche thermisch bezogenen Probleme effizient zu mindern, wenn der HBM-Stapel Spitzenverkehr erfährt oder wenn wiederholt durch den Speichercontroller auf denselben Die im Stapel zugegriffen wird. Die vorbekannten Ansätze unternehmen auch nichts, wenn sich einer der mittleren Dies im HBM-Stapel unter Stress befindet.

**[0021]** Fig. 3 ist eine Darstellung eines beispielhaften Systems, das einen Haupt-Die **202** umfasst, der mit einem zugeordneten Speicher-Die-Stapel **220'** und einem Hostprozessor **330** gekoppelt ist, gemäß einer Ausführungsform. Der Speicher-Die-Stapel **220'** kann ein HBM-Stapel, ein Die-Stapel nur mit nichtflüchtigen Speicherchips, ein Die-Stapel nur mit flüchtigen Speicherchips, ein Die-Stapel sowohl mit flüchtigen als auch nichtflüchtigen Speicherchips oder ein Die-Stapel mit anderen Arten von integrierten Schaltungsvorrichtungen sein. Wie in Fig. 3 gezeigt, kann der HBM-Stapel **220'** mehrere Pseudokanäle (PC) umfassen. Zum Beispiel kann der Speicherstapel **220'** einen ersten Pseudokanal PC1, einen zweiten Pseudokanal PC2, ... und einen n-ten Pseudokanal PCn umfassen, wobei n gleich 2, 4, 8, 16, 2-16, größer als 16 oder eine beliebige geeignete ganze Zahl sein kann. Bei einer geeigneten Anordnung können die zwei Pseudokanäle einem gegebenen Speicher-Die im HBM-Stapel zugeordnet sein (z. B. können PC1 und PC2 einem der Dies im Stapel zugeordnet sein, wie etwa durch den gestrichelten Kasten **221** angegeben). Im Allgemeinen kann mindestens ein Pseudokanal einem gegebenen Die im HBM-Stapel zugeordnet sein, oder es können mehr als zwei Pseudokanäle einem gegebenen Die im HBM-Stapel zugeordnet sein.

**[0022]** Jeder Pseudokanal kann seinen eigenen dedizierten Controller aufweisen, wie etwa die HBM-Steuerschaltung **306** (siehe z. B. den PC1 zugeordneten HBM-Controller **306-1**, den PC2 zugeordneten HBM-Controller **306-2**, ... und den PCn zugeordneten HBM-Controller **306-n**), der auf dem Haupt-Die **202** gebildet ist. Bei anderen Ausführungsformen haben mehrere Pseudokanäle ihren eigenen geteilten/dedizierten Speichercontroller. In dem Szenario, bei dem sich zwei Pseudokanäle einen gemeinsamen Speicherchip teilen, können Befehle von den zwei Pseudokanälen auf alternierenden Zyklen eingeteilt werden. Der HBM-Stapel **220'** kann über eine Schnittstellenkomponente der Bitübertragungsschicht (manchmal als „PHY“ abgekürzt) **304** im Haupt-Die **202** über den Pfad **302** zwischen Dies mit den entsprechenden Controllern **306** kommunizieren. Die PHY-Komponente **304** kann PMD-Subschichtblöcke (Physical Medium Dependent) (z. B. Sendeempfängerblöcke), PCS-Blöcke (Physical Coding Sublayer) (z. B. eine Subschicht, die bestimmt, wann eine funktionale Strecke hergestellt wurde, und Codierung wie 64b/66b-Codierung durchführt) und PMA-Blöcke (Physical Medium Attachment) umfassen, die als Schnitt-

stelle zwischen den PMD-Blöcken und den PCS-Blöcken dienen.

**[0023]** Die Speichercontroller **306** können über jeweilige On-Chip-Hochgeschwindigkeits-Interconnect-Schnittstellenbusse, wie etwa die AXI-Schnittstelle (Advanced eXtensible Interface) (siehe z. B. die AXI-Schnittstellenbusse **309-1**, **309-2**, ..., **309-n**), mit einem Prozessorkern **310** im Haupt-Die **202** gekoppelt sein. Obwohl hier manchmal die AXI-Schnittstelle als Beispiel beschrieben wird, können andere Arten von Hochfrequenz-Mehrfach-Raster/Mehrfach-Slave-On-Chip-Schnittstelle oder protokollgesteuerter Hochgeschwindigkeits-Intra-Die-Kommunikationsschnittstelle verwendet werden. Der Prozessorkern **310** kann ein Soft-Prozessorkern (z. B. ein Prozessorkern, der statisch oder dynamisch umkonfigurierbar ist) oder ein Hart-Prozessorkern (z. B. ein Prozessorkern, der nicht umkonfigurierbar und fest verdrahtet ist, um einen gegebenen Umfang von Funktionen auszuführen) sein. Bei einigen Ausführungsformen kann der Prozessorkern **310** über die intelligenten Schaltnetzwerke **308** mit dem HBM-Controllern gekoppelt sein. Die intelligenten Schaltnetzwerke **308**, die zwischen dem Prozessorkern und den Speichercontrollern angeordnet sind, können dafür ausgelegt sein, auf der Basis von Informationen, die von dem HBM-Subsystem empfangen werden, sowie sie gegebenenfalls über den Signalpfad **312** bereitgestellt werden, selektiv Verkehr zwischen verschiedenen Pseudokanälen innerhalb eines gegebenen Dies im HBM-Stapel oder zwischen verschiedenen Dies innerhalb des Speicherstapels umzurouten.

**[0024]** Der Haupt-Die **202** kann ferner eine Eingabe-Ausgabe-Komponente umfassen, wie etwa eine Hostschnittstelle **320**, die dafür ausgelegt ist, über den Pfad **332** mit einer Host-CPU (Zentralverarbeitungseinheit) zu kommunizieren. Der Pfad **332** kann eine kohärente Computerbusschnittstelle sein, wie etwa UPI (UltraPath Interconnect), QPI (QuickPath Interconnect), CXL (Compute Express Link), CCIX (Cache Coherent Interconnect for Accelerators), Gen-Z, (OpenCAPI) (Open Coherent Accelerator Processor Interface), IAL (Intel Accelerator Link), die NVLink von Nvidia oder andere Computerbusschnittstellen. Der Pfad **332** kann auch eine nichtkohärente Computerbusschnittstelle umfassen, wie etwa die PCIe-Schnittstelle (Peripheral Component Interconnect Express). Im Allgemeinen können die Hostschnittstelle(n) **320** unter Verwendung mehrerer Datenspuren (z. B. mindestens 2 Spuren, mindestens 4 Spuren, mindestens 8 Spuren, mindestens 16 Spuren, mindestens 32 Spuren, mindestens 64 Spuren usw.), einer einzigen Datenspur, eines parallelen Datenbusses, eines seriellen Datenbusses oder anderer Computerbusstandards, die Datentransferraten von mindestens 250 MBps (Megabyte pro Sekunde), 500 MBps, 1 GBps (Gigabyte pro Sekunde), 5 GBps, 10

GBps, 16 GBps, 32 GBps, 64 GBps oder mehr unterstützen können, implementiert werden.

**[0025]** Das Host-Betriebssystem (OS) kann auf die Host-CPU **330** (die manchmal einfach als der „Host“ oder „Host-Prozessor“ bezeichnet wird) geladen werden. Das Host-OS kann einen Hypervisor **334** implementieren, der die Verwendung einer oder mehrerer virtueller Maschinen (z. B. virtueller Maschinen VM1, VM2, ..., VMn) auf dem Host-Prozessor **330** ermöglicht. Virtuelle Maschinen sind selbstständige virtualisierte Partitionen, die eine unabhängige Hardware-Datenverarbeitungsressource simulieren. Jede virtuelle Maschine ist deshalb in der Lage, eine oder mehrere Anwendungen (siehe App) laufen zu lassen. Gegebenenfalls könnte eine einzelne Anwendung auf zwei oder mehr virtuellen Maschinen laufen. Der Hypervisor **334** kann Teil der auf dem Hostprozessor **334** laufenden Software oder Firmware sein und kann als Virtuellmaschinen-Überwachungsrichtung (manchmal als VMM oder Virtual Machine Manager bezeichnet), die Hardwareressourcen des Systems verwaltet, so dass sie effizient auf die virtuellen Maschinen (VM) im System verteilt werden.

**[0026]** Jede virtuelle Maschine kann als eine „Gast-Maschine“ bezeichnet werden, auf der ihr eigenes Gast-Betriebssystem (OS) läuft. Jede virtuelle Maschine kann zum Laufenlassen einer oder mehrerer Benutzeranwendungen verwendet werden. Der Hypervisor **334** präsentiert dem Gast-OS der VM eine virtuelle Betriebsplattform und verwaltet die Ausführung der Gast-Betriebssysteme, während virtualisierte Hardwareressourcen geteilt werden. Der Hypervisor **334** kann nativ auf der Hardware des Hosts (als Hypervisor des Typs **1** „blankes Metall“) laufen oder kann über einem existierenden Host-Betriebssystem (als „gehosteter Hypervisor des Typs **2**) laufen. Gegebenenfalls können zusätzliche Virtualisierungstreiber und -werkzeuge (nicht gezeigt) verwendet werden, um jeder virtuellen Gastmaschine dabei zu helfen, effizienter mit der zugrundeliegenden physischen Hardware der Host-CPU **330** zu kommunizieren. Im Allgemeinen kann der Prozessor **330** dafür ausgelegt sein, mindestens zwei VM, zwei bis zehn VM, mehr als zehn VM, hunderte VM, tausende VM oder eine beliebige geeignete Anzahl virtueller Maschinen zu hosten.

**[0027]** Im Allgemeinen kann die auf der Host-CPU **330** laufende Software unter Verwendung von Softwarecode implementiert werden, der auf nichttransistorischen computerlesbaren Speicherungsmedien (z. B. greifbaren computerlesbaren Speicherungsmedien) gespeichert wird. Der Softwarecode kann manchmal als Software, Daten, Programmanweisungen, Anweisungen, Script oder Code bezeichnet werden. Die nichttransistorischen computerlesbaren Speicherungsmedien können nichtflüchtigen Speicher, wie etwa nichtflüchtigen Direktzugriffsspeicher (NV-

RAM), eine oder mehrere Festplatten (z. B. magnetische Laufwerke oder Halbleiterlaufwerke), ein oder mehrere wechselbare Flash-Laufwerke oder andere wechselbare Medien oder dergleichen umfassen. Auf den nichttransistorischen computerlesbaren Speicherungsmedien gespeicherte Software kann auf den Verarbeitungsschaltkreisen des Hostprozessors **330** ausgeführt werden.

**[0028]** Viele thermische Verwaltungsfähigkeiten können in den Speichercontrollern **306** bereitgestellt werden. Vor der Beschreibung der verschiedenen Parameter, die durch die Speichercontroller zum Drosseln der Speicherbefehle in Betracht gezogen werden könnten, ist es hilfreich, zuerst zu verstehen, wie Pseudokanäle organisiert sind. **Fig. 4** ist eine Darstellung eines beispielhaften Pseudokanals **400**, der mehrere Speicherbänke umfasst. Wie in **Fig. 4** gezeigt, kann der Pseudokanal **400** mehrere Speicherbänke enthalten (z. B. eine erste Speicherbank **402-1**, eine zweite Speicherbank **402-2**, ... und eine m-te Speicherbank **402-m**). Jede Speicherbank **402** kann mehrere Zeilen oder Seiten **404** umfassen. Jede Bank **402** kann nur eine einzige offene Seite unterhalten. Um auf eine andere Seite in einer Speicherbank zuzugreifen, muss somit die existierende Seite vor dem Öffnen der gewünschten Seite geschlossen werden.

**[0029]** **Fig. 5** ist eine Darstellung einer Speichercontrollerschaltung **306**, die dafür ausgelegt ist, auf der Basis eines oder mehrerer von Speicherstapel empfangener Betriebsparameter (siehe z. B. den Pfad **312-1** in **Fig. 3**) thermische Drosselung durchzuführen. Wie in **Fig. 5** gezeigt, kann der Speichercontroller **306** thermische Schätzungs- und Drosselungsschaltkreise **500** umfassen, die dafür ausgelegt sind, die mehreren Betriebsparameter/-metriken (die manchmal kollektiv als thermisch bewusste Informationen oder Statistiken bezeichnet werden) zu empfangen.

**[0030]** Bei einer geeigneten Ausführungsform kann der Speichercontroller verschiedene Arten von statistischen oder Überwachungsinformationen empfangen, die einen Speicherbankstatus, eine Anzahl von Befehlen (z. B. die Anzahl von Leseanforderungen, Schreibanforderungen, Auffrischungen, Vorladungen, Aktivierungen, Leerlaufbefehlen usw.), eine Zeilen-/Seiten-/Bandbreiten-Auslastungsmetrik, einen oder mehrere Schwellenwerte, wie etwa einen Auslastungsschwellenwert und einen Speicherfehlerschwellenwert, wie etwa einen Schwellenwert für den ECC (Fehlerkorrekturcode) und Temperaturdaten umfassen können. Diese Liste soll auf keinerlei Weise erschöpfend sein. Auf der Basis dieser Parameter kann der Speichercontroller Verkehr in oder unter den verschiedenen Pseudokanälen eines einzelnen Dies im Speicherstapel umrouten. Gegebenenfalls können diese Parameter auch zu den in-

telligenten Schaltnetzwerken **308** (siehe z. B. Pfad **312-2** in **Fig. 3**) weitergeleitet werden, so dass die intelligenten Schaltnetzwerke **308** Speicherzugriffe auf den gesamten Speicherstapel verteilen können, um die verfügbare Bandbreite voll zu nutzen.

**[0031]** Bei einer anderen geeigneten Ausführungsform können die empfangenen Schwellenwerte auch einen Zeilen- oder Seitenebenen-Schwellenzählwert umfassen, der die Anzahl aufeinanderfolgender Zugriffe auf eine gegebene Zeile/Seite aufzeichnet. Eine Situation, in der kontinuierlich auf eine Seite zugegriffen wird, wird manchmal als „Zeilenhammer“-Problem bezeichnet, das auftreten kann, wenn das System attackiert wird oder während eines gewissen anderen ungewöhnlichen Szenarios. Durch Überwachen des Seitenzählwerts ist der Speichercontroller in der Lage, das Zeilenhammerproblem zu vermeiden und kann auch thermischen Stress an dem attackierten Die lindern.

**[0032]** Bei noch einer anderen geeigneten Anordnung kann der Speichercontroller Verkehr auf der Basis eines Musters einer Region, auf die häufig zugegriffen wird, umrouten (z. B. kann es wünschenswert sein, Verkehr zu Bereichen des Dies zu routen, auf die weniger häufig zugegriffen wird). Zum Beispiel kann der Speichercontroller der auf der entsprechenden virtuellen Maschine laufenden Anmeldung Rückmeldung darüber geben, welche Region besser zum Schreiben ist, indem ein codiertes Flag als Teil von Leistungsfähigkeits-Überwachungsvorrichtungen gesendet wird oder durch Exponieren dieser Informationen im Hinblick auf einen Adressenbereich.

**[0033]** Weiterhin mit Bezug auf **Fig. 5** können die dynamischen Schätzungs- und Drosselungsschaltkreise **500** eine Befehls-Umordnungswarteschlange **502** anweisen, die Befehle auf der Basis der thermisch bewussten Statistiken selektiv umzuordnen. Man betrachte zum Beispiel ein Beispiel, bei dem ein HBM-Stapel Dies **D0**, **D1**, **D2** und **D3** umfasst, wobei **D2** aktuell heiß und **D0** kühl ist. Wenn die ausgegebenen Befehle **D0**, **D2**, **D2**, **D2**, **D2**, **D0**, **D1** und **D3** in dieser Reihenfolge anvisieren, können die zugeordneten Speichercontroller ihre Befehls-Umordnungswarteschlangen **502** so drosseln, dass die für die Dies **D0**, **D1** und **D3** bestimmten Befehle zuerst ausgeführt werden, um es dem Die **D2** zu erlauben, sich abzukühlen, bevor auf ihn zugegriffen wird. Gegebenenfalls kann die Befehls-Umordnungswarteschlange auch im HBM-Modul **220'** (**Fig. 3**) gebildet werden, so dass Befehle lokal im Speicherstapel in Warteschlangen eingereicht und umgeordnet werden.

**[0034]** Bei noch einer anderen geeigneten Ausführungsform kann der Speichercontroller die vom HBM-Stapel empfangenen Temperaturdaten analysieren und kann die Verkehrsaktivität für einen oder mehrere Dies im Stapel abhängig von den Temperaturda-

ten selektiv begrenzen. **Fig. 6** ist eine Darstellung, die zeigt, wie Temperaturinformationen von mehr als einem Die in dem Die-Stapel **220'** erhalten werden können. Wie in **Fig. 6** gezeigt, kann der Die-Stapel **220'** eine beliebige gewünschte Anzahl gestapelter Dies **602** umfassen (z. B. mindestens zwei 3D-gestapelte Dies, mindestens drei 3D-gestapelte Dies, mindestens vier 3D-gestapelte Dies, 4-12 gestapelte Dies, mehr als 12 gestapelte Dies usw.).

**[0035]** Der unterste Die **602-B** im Die-Stapel kann einen Temperatursensor **604** aufweisen, der Temperaturinformationen für diesen Die den Speichercontrollern und/oder den intelligenten Schaltnetzwerken auf dem Haupt-Die bereitstellt. Außerdem kann der unterste Die **602-T** im Die-Stapel auch einen Temperatursensor **604** aufweisen, der Temperaturinformationen für diesen Die den Speichercontrollern und/oder den intelligenten Schaltnetzwerken auf dem Haupt-Die bereitstellt. Mit diesen Informationen können die thermischen Schätzungsschaltkreise **500** im Speichercontroller (siehe **Fig. 5**) einen Temperaturgradienten berechnen, um die Temperatur für einen beliebigen der Zwischen-Dies, wie etwa den Die **602-I** im Die-Stapel zu interpolieren. Gegebenenfalls kann der Zwischen-Die **602-I** auch seinen eigenen dedizierten Temperatursensor **604'** umfassen, um einen genaueren Messwert als die interpolierten Ergebnisse bereitzustellen. Bei Konfiguration auf diese Weise kann der Speichercontroller Befehle an einen oder mehrere Dies im Die-Stapel drosseln (indem z. B. nur erlaubt wird, dass 50% oder eine andere Teilmenge von Befehlen ausgeführt werden), wenn die Temperatur- oder anderen Informationen, wie etwa ein Aktivitäts-Tracker, eine bestimmte Schwelle oder thermische Grenze erreichen oder überschreiten.

**[0036]** Gemäß noch einer anderen geeigneten Ausführungsform können die intelligenten Schaltnetzwerke **308** mit einer Adressenübersetzungsschaltung **700** versehen sein, die dafür ausgelegt ist, logische (Benutzer-) Adressen auf physische (Speicher-) Adressen abzubilden. Die Adressenübersetzungsschaltkreise **700** können entweder statische Adressenübersetzung (wobei die Beziehung von logischen und physischen Adressen fest ist) oder dynamische Adressenübersetzung (wobei die Abbildung von logischen und physischen Adressen dynamisch auf der Basis eines gewissen vorbestimmten Algorithmus justiert werden kann) implementieren. Bei dynamischer Adressenübersetzung bildet die Schaltung **700** die physische Adresse dynamisch neu auf Schreibvorgänge ab und bewahrt diese Umabbildung bis zum nächsten Schreiben in dieselbe Adresse. Das Durchführen von Umabbildung auf diese Weise kann dabei helfen, die Speicherauslastung über verschiedene Ebenen (d. h. über verschiedene Regionen eines gegebenen Dies oder über verschiedene Dies in einem Die-Stapel) zu verteilen.

**[0037]** Gegebenenfalls kann die durch die Adressenübersetzungsschicht bereitgestellte Umabbildung auch in jedem Speichercontroller auf der Systemebene implementiert werden. Zusätzliche Adressenbit können verwendet werden, um der Anwendung anzugeben, welche Abbildung verwendet oder erwartet wurde. Die intelligenten Schaltnetzwerke **308** können (beispielsweise) als ein Crossbarschalter implementiert werden, so dass die Adressen Umabbildung manchmal als Crossbar-Schwenkung (Swizzling) bezeichnet werden können.

**[0038]** Die verschiedenen Kanäle (darunter Pseudokanäle) können dem Benutzer als ein kontinuierlicher Adressenraum erscheinen. **Fig. 8** ist eine Darstellung, die zeigt, wie ein Adressenraum **800** in eine statische Region **802** und eine dynamische Region **804** aufgeteilt werden kann, gemäß einer Ausführungsform. Die Größe der statischen Region **802** und dynamischen Region **804** kann auf der Basis der Bedürfnisse der Anwendung geändert werden. Wie in **Fig. 8** gezeigt, kann die Region **802** in m statische Teilregionen aufgeteilt werden, während die Region **804** in n dynamische Teilregionen aufgeteilt werden kann. Die Anzahl der Regionen innerhalb der Regionen **802** und **804** kann geändert werden.

**[0039]** Bei einer geeigneten Anordnung kann die physische Adresse Kanaladressenbit, eine einen ausgewählten Die im Stapel identifizierende Stapelkennung, Bankadressenbit, Zeilenadressenbit, Spaltenadressenbit usw. umfassen. Die Abbildung von der logischen Benutzeradresse auf die physische Speicheradresse kann auf einigen wenigen vordefinierten oder benutzerdefinierten Hash-Funktionen basieren und ist typischerweise eine injektive Abbildung. Die Inhalte der Adressentabelle **800** können in dem Adressentabellenspeicher **702** gespeichert werden, der als eine Adressenübersetzungs-Nachschlagetabelle dient, um zu bestimmen, welche Adresse wohin gehen muss, um die gewünschte thermische Stabilität im Speicherstapel aufrechtzuerhalten.

**[0040]** Der Tabellenspeicher **702** (der manchmal auch als Seitentabelle bezeichnet wird) kann auch einen Schwellenzähler unterhalten, um zu prüfen, ob Zugriff auf eine bestimmte Adresse ihre Grenzen erreicht hat, oder um zu bestimmen, ob ein gegebener Die oder eine gegebene Seite von einer spezifizierten Attacke der Zeilenhammerart betroffen ist. Wenn zum Beispiel der Speichercontroller oder die intelligenten Schaltnetzwerke bestimmen, dass neuere Zugriffe immer eine kleine Anzahl von Bänken anviesieren, sollte sich eine neue dynamische Hashing-Funktion auf das Verteilen des Zugriffs auf mehrere Kanäle, Pseudokanäle oder Bänken konzentrieren. Die Adressentabelle kann in einem On-Chip-Speicher (z. B. On-Chip-SRAM-Block), auf dem HBM-Stapel oder sowohl auf dem On-Chip-Speicher als auch auf dem HBM-Stapel gespeichert werden. Gegebenenfalls

nenfalls kann die auf dem Hostprozessor laufende Anwendung mit dem Speichercontroller kommunizieren, um die Adressenorte effektiv zu verwalten, um einen Leseadressenbereich zu exponieren, indem zukünftige Schreibvorgänge erwünscht sind, um den Adressenraum auf der Basis der Temperatur und anderer Statistiken dynamisch zuzuteilen oder die Zuteilung aufzuheben usw.

**[0041]** Wie oben beschrieben, ist eine Art von Adressenübersetzung, die durch die Schaltung **700** unterstützt werden kann, Adressen-Hashing. Man betrachte zum Beispiel ein Szenario, in dem die physische Speicheradresse durch Nehmen des Exklusiv-OR (XOR) der logischen Benutzeradresse und eines vorbestimmten Hash-Werts berechnet wird. Wenn die Benutzeradresse „0x0000“ und die Hash-Funktion „0x5555“ ist, wäre somit die entsprechende physische Adresse gleich „0x5555“. Wenn als ein anderes Beispiel die Benutzeradresse „0x0001“ und die Hash-Funktion „0x5555“ ist, wäre die entsprechende physische Adresse gleich „0x5554“. Dieses Beispiel ist lediglich veranschaulichend. Im Allgemeinen kann jede Art von statischer/dynamischer Adressenübersetzung oder Hashing-Funktion verwendet werden, um Befehle auf verschiedene Teile eines Die-Stapels umzuverteilen.

**[0042]** Im Allgemeinen können die intelligenten Schaltnetzwerke **308** dafür ausgelegt sein, Verkehr zu und von dem Die-Stapel umzurouten, um Zugriffe/Befehle über den gesamten Die-Stapel zu verteilen, um eine gleichmäßige Temperaturverteilung auf einer Systemebene aufrechtzuerhalten (z. B. um sicherzustellen, dass kein einzelner Die übermäßig heißer als ein anderer wird).

**[0043]** Zum Beispiel wird eine grafikintensive Arbeitslastbandbreite etwa 85% der Maximalkapazität betragen und wird leseintensiv mit wenigen Seitentreffern sein. Solche Arten von Arbeitslast würden die Stromentnahmeanforderung vergrößern, die ihrerseits die Temperatur aufgrund der Notwendigkeit, jede Seite/Zeile nach jedem Lesen vorzuladen und aufzufrischen, signifikant erhöht. Man betrachte ein Szenario, in dem eine 8 GB-HBM-Vorrichtung einen Stapel von acht Dies hoch mit 4 GB pro Kanal (z. B. jeder Die hat zwei Pseudokanäle) und adressiert unter Verwendung von 14 Zeilenadressenbit, 6 Spaltenadressenbit, 4 Bankadressenbit, umfasst und wobei die Gesamtseitengröße 2 KB ist (z. B. 1 KB für jeden Pseudokanal).

**[0044]** Nach dem Analysieren eines oder mehrerer vom Die-Stapel empfangener Parameter (siehe z. B. die empfangenen Parameter/Metriken in **Fig. 5**, die auch parallel den intelligenten Schaltnetzwerken zugeführt werden können) oder nach dem Berechnen der Temperatur eines oder mehrerer Dies im Stapel können die intelligenten Schaltnetzwerke die Adres-



senübersetzungs-/Hashing-Schaltung benutzen, um die Pseudokanäle zu schwenken, indem die HBM-Ebenen-Adresse und die Adressenbit höherer Ordnung auf der Crossbarschaltebene betrachtet werden, so dass die Befehle gleichmäßig verteilt werden. Der Ausdruck „schwenken“ kann als Umordnung, Umverteilung, Mischung, Randomisierung oder anderweitiges Umrouten der Speicherbefehle/-anforderungen definiert werden.

**[0045]** Beispielsweise können die intelligenten Schaltnetzwerke in Verbindung mit einem oder mehreren Speichercontrollern Befehle durch jeden Die im Stapel in einer Reigenbewegung zyklieren lassen, so dass die Anforderung gleichmäßig über den Stapel verteilt werden, statt wieder und wieder denselben Die zu adressieren. Dies lässt sich durch Verwendung eines Adressen-Hashing-Algorithmus erzielen, bei dem die LSB (niedrigwertigen Bit) die Die-Adressenbit sind. Man betrachte ein Szenario, in dem ein HBM-Stapel 8 vertikal gespeicherte Speicherchips umfasst, so dass eine Chipkennung (ID) mit drei Bit ausreichen würde, um die 8 Dies zu adressieren. In diesem Szenario kann es sich bei der 3-Bit-Chip-ID um die LSB handeln, wobei die Bit höherer Ordnung Zeilenadressenbit, Spaltenadressenbit, Bankadressenbit, Rangbit usw. umfassen könnten. Auf diese Weise konfiguriert würde ein Adressenzähler, der die LSB inkrementiert, effektiv die Befehle in der Reigen-Weise effektiv auf verschiedene Dies schwenken. Diese Reigenbewegung ist lediglich veranschaulichend. Der Benutzer kann gegebenenfalls auch die Adressen neu zuweisen, so dass die Schwenkung in einer anderen vorbestimmten oder zufälligen Weise ausgeführt wird.

**[0046]** Die intelligenten Schaltnetzwerke (z. B. ein Crossbarschalter) kann auch Schwenkung hauptsächlich auf der Basis der Temperatur (der ein höherer Gewichtungsfaktor oder Kredit gegeben werden kann), des Bankstatus jedes Dies, der Bandbreitenauslastung (z. B. der Befehlsrate innerhalb des zulässigen Taktzykluszyklus) jedes der Pseudokanäle, der Latenzpriorität jedes Befehls usw. durchführen. Gegebenenfalls können Daten auch optional in den intelligenten Schaltnetzwerken verwürfelt werden, um Zufälligkeit aufrechtzuerhalten, statt Daten auf der Kanalebene zu verwüfeln. Durchführen von Datenverwürfelung auf der Crossbarschaltebene kann dabei helfen, verbesserte Flexibilität zu gewährleisten. Bei noch anderen geeigneten Ausführungsformen können auf dem Speicherstapel gespeicherte Daten gespiegelt werden (z. B. es werden sowohl eine Hauptkopie als auch eine Duplikatkopie gespeichert), und wenn es erwünscht ist (dieses Stück Daten zuzugreifen, kann nur eine der Hauptkopie oder der Spiegelkopie gelesen werden, je nachdem, welche kühler ist. Bei einigen Ausführungsformen könnten Pseudokanäle in mehrere virtuelle Ränge innerhalb der HBM-Kanäle umgewandelt werden, so dass die virtuellen

Ränge einzeln gesteuert werden können, um Aktionen wie Spiegelung oder Rangverschachtelung zu unterstützen, um den Durchsatz für hochintensive Arbeitslasten wie Vernetzungs- und AI-Anwendungen (künstliche Intelligenz) zu verbessern. Die Speichersteuerschaltungen (wie z. B. die HBM-Controller) und die intelligenten Schaltnetzwerke, die auf diese Weise konfiguriert werden, können manchmal kollektiv als Aktivitäts-/thermische Überwachungsschaltkreise, thermisch bewusste Verkehrs-Umroutungsschaltkreise oder Steuerschaltkreise bezeichnet werden.

**[0047]** Fig. 9 ist ein Flussdiagramm von beispielhaften Schritten zum Betrieb der in Verbindung mit Fig. 2-7 gezeigten Schaltungen. In Schritt 900 können Speichercontroller und/oder die intelligenten Schaltnetzwerke Speicherzugriffsstatistiken, Temperaturdaten und andere Parameter von dem Die-Stapel (z. B. vom HBM-Die-Stapel oder anderen Arten von Die-Stapel) empfangen. Die Speicherzugriffsstatistiken können Bankstatus, Befehlsauslastung, Anzahl der Befehle an jeden Die, einen Zeilenschwellenzähler und Bandbreite, Fehlerschwelle, Muster von Regionen, auf die häufig zugegriffen wird, und/oder andere Metriken umfassen, die dabei helfen könnten, das thermische Verhalten auf dem Die-Stapel zu schätzen.

**[0048]** In Schritt 902 können der Speichercontroller und/oder die intelligenten Schaltnetzwerke bestimmen, ob die empfangenen Parameter/Statistiken vorbestimmte Schwellenwerte überschreiten. Wenn nicht, kann das System mit dem normalen gewünschten Betrieb voranschreiten (Schritt 912). Wenn bestimmt wird, dass die empfangenen Parameter die vorbestimmten Schwellenwerte überschreiten, kann Verarbeitung zu Schritt 904 schreiten, zu bestimmen, ob der Speicher gespiegelt ist. Wenn der Speicher gespiegelt ist, wird nur die kältere Kopie gelesen (Schritt 906).

**[0049]** Wenn der Speicher nicht gespiegelt ist, können die intelligenten Schaltnetzwerke und/oder der bzw. die Speichercontroller statische/dynamische Adressenübersetzung oder Hashing, Befehls-umordnung, intelligentes Crossbar-Schalten, erweiterte thermische Drosselungstechniken (z. B. Drosselungsoperationen, die Drosselung nicht nur auf der Basis der Basis-Die-Temperatur durchführen, sondern mehrere zusätzlich Parameter berücksichtigen) und andere geeignete Umroutungsmethoden durchführen, um eine gleichmäßige Verteilung von Die-Anforderungen sicherzustellen, um zu verhindern, dass sich irgendein Die im Stapel zu schnell erhitzt. Gegebenenfalls können die intelligenten Schaltnetzwerke und/oder der/die Speichercontroller der auf der Host-CPU laufenden Anwendung zusätzliche Rückmeldung bereitstellen (z. B. um einen Adressenbereich für zukünftige Schreibvorgänge zu exponieren

usw.). Danach kann Verarbeitung in Schritt **912** mit Normalbetrieb voranschreiten.

**[0050]** Die oben beschriebenen Beispiele, bei denen die Speichercontroller **306** und/oder die intelligenten Schaltnetzwerke **308** dafür ausgelegt sind, statistische Informationen von dem Speicherstapel **220'** zu empfangen, sind lediglich veranschaulichend und sollen den Schutzzumfang der vorliegenden Ausführungsform nicht begrenzen. Bei noch anderen geeigneten Anordnungen können der Speichercontroller und/oder die intelligenten Schaltnetzwerke in der Lage sein, die interessierenden Statistiken (z. B. die Anzahl von Seitenöffnungen/-schließungen) und andere Daten, die dabei helfen können, die relativen thermischen Bedingungen der IC-Chips im Die-Stapel zu bestimmen, unabhängig zu verfolgen.

**[0051]** Obwohl die Verfahren von Operationen in einer spezifischen Reihenfolge beschrieben werden, versteht sich, dass zwischen beschriebenen Operationen andere Operationen ausgeführt werden können, beschriebene Operationen justiert werden können, so dass sie zu etwas Zeiten auftreten, oder beschriebene Operationen in einem System verteilt werden können, das das Auftreten von Verarbeitungsoperationen in verschiedenen Intervallen erlaubt, die der Verarbeitung zugeordnet sind, solange die Verarbeitung der Overlay-Operationen auf eine gewünschte Weise ausgeführt werden.

Beispiele:

**[0052]** Die folgenden Beispiele betreffen weitere Ausführungsformen.

Beispiel 1 ist eine integrierte Schaltung, umfassend: eine Schnittstellenschaltung, ausgelegt zum Kommunizieren mit einem externen Die-Stapel, der mehrere vertikal gestapelte Dies aufweist; und Steuerschaltkreise, ausgelegt zum Analysieren von dem externen Die-Stapel zugeordneten Informationen und zum Verteilen von Zugriffen auf die mehreren vertikal gestapelten Dies, um zu verhindern, dass irgendwelche der mehreren vertikal gestapelten Dies eine Temperatur aufweisen, die eine vorbestimmte thermische Grenze übersteigt.

Beispiel 2 ist die integrierte Schaltung nach Beispiel 1, wobei die analysierten Informationen Parameter umfassen, die aus der folgenden Gruppe ausgewählt sind: ein Speicherbankstatus, eine Anzahl von Befehlen, eine Speicherbandbreiten-Auslastungsmetrik, mindestens ein Schwellenwert und eine Anzahl von Seitenöffnungen und -schließungen.

Beispiel 3 ist die integrierte Schaltung nach irgendeinem der Beispiele 1-2, wobei die ana-

lysierten Informationen einen Speicherfehler-Schwellenwert umfassen.

Beispiel 4 ist die integrierte Schaltung nach irgendeinem der Beispiele 1-3, wobei die analysierten Informationen ein Regionenmuster häufigen Zugriffs eines gegebenen Dies in dem externen Die-Stapel umfassen.

Beispiel 5 ist die integrierte Schaltung nach irgendeinem der Beispiele 1-4, wobei die analysierten Informationen Temperaturdaten umfassen, die mehr als einem der vertikal gestapelten Dies in dem externen Die-Stapel zugeordnet sind.

Beispiel 6 ist die integrierte Schaltung nach Beispiel 5, wobei die Temperaturdaten eine erste Temperatursensormessung von einem obersten Die in dem externen Die-Stapel und eine zweite Temperatursensormessung von einem untersten Die in dem externen Die-Stapel umfassen.

Beispiel 7 ist die integrierte Schaltung nach Beispiel 6, wobei die Steuerschaltkreise ferner ausgelegt sind zum Berechnen eines Temperaturgradienten auf der Basis der ersten und zweiten Temperatursensormessung zur Schätzung eines Temperaturwerts für einen zwischen dem obersten und untersten Die in dem externen Die-Stapel angeordneten Zwischen-Die.

Beispiel 8 ist die integrierte Schaltung nach irgendeinem der Beispiele 1-7, wobei die Steuerschaltkreise Speichercontrollerschaltungen umfassen, die dafür ausgelegt sind, für heiße Dies in dem Die-Stapel bestimmten Verkehr zu drosseln.

Beispiel 9 ist die integrierte Schaltung nach irgendeinem der Beispiele 1-8, wobei die Steuerschaltkreise Speichercontrollerschaltungen umfassen, die dafür ausgelegt sind, Verkehr von einem gestauten Kanal auf einen weniger gestauten Kanal umzurouten, um ein thermisches Profil über den Die-Stapel auszugleichen.

Beispiel 10 ist die integrierte Schaltung nach irgendeinem der Beispiele 1-9, wobei die Steuerschaltkreise Schaltnetzwerke umfassen, die dafür ausgelegt sind, Zugriffe auf den externen Die-Stapel so zu schwenken, dass auf kalte Dies in dem Die-Stapel häufiger als auf heiße Dies in dem Die-Stapel zugegriffen wird.

Beispiel 11 ist die integrierte Schaltung nach Beispiel 10, wobei die Schaltnetzwerke eine Adressenübersetzungsschaltung umfassen, die dafür ausgelegt ist, logische Adressen auf entsprechende physische Adressen abzubilden.

Beispiel 12 ist die integrierte Schaltung nach Beispiel 11, wobei die Adressenübersetzungsschaltung dafür ausgelegt ist, statische Adressenabbildung durchzuführen.

Beispiel 13 ist die integrierte Schaltung nach Beispiel 11, wobei die Adressenübersetzungsschaltung dafür ausgelegt ist, dynamische Adressenabbildung durchzuführen.

Beispiel 14 ist die integrierte Schaltung nach irgendeinem der Beispiele 1-13, wobei die externen Daten dafür ausgelegt sind, eine Hauptkopie gegebener Daten und eine gespiegelte Kopie der gegebenen Daten zu speichern, und wobei die Steuerschaltkreise dafür ausgelegt sind, nur auf eine ausgewählte der Hauptkopie und der gespiegelte Kopie mit der niedrigeren Temperatur zuzugreifen.

Beispiel 15 ist ein Verfahren zum Betrieb eines integrierten Schaltungs-Dies, umfassend: mit einer Schnittstellenschaltung in dem integrierten Schaltungs-Die: Kommunizieren mit einem externen Chip-Stapel; mit Steuerschaltkreisen in dem integrierten Schaltungs-Die: Empfangen und Analysieren mehrerer verschiedener Arten statistischer Informationen von dem externen Chip-Stapel; und als Reaktion auf Empfang und Analysieren der mehreren verschiedenen Arten statistischer Informationen, Verwenden der Steuerschaltkreise zum Aufrechterhalten eines ausgeglichenen thermischen Profils in dem externen Chip-Stapel durch Umrouten von Anforderungen zu dem externen Chip-Stapel.

Beispiel 16 ist das Verfahren nach Beispiel 15, wobei Empfangen der mehreren verschiedenen Arten statistischer Informationen Empfangen von Temperaturmesswerten von mindestens zwei Dies in dem externen Chip-Stapel umfassen.

Beispiel 17 ist das Verfahren nach irgendeinem der Beispiele 15-16, wobei Empfangen der mehreren verschiedenen Arten von statistischen Informationen Empfangen von Temperaturmesswerten und anderen Arten von Parametern, die von den Temperaturmesswerten verschieden sind, umfasst.

Beispiel 18 ist das Verfahren nach Beispiel 17, wobei die anderen Arten von Parametern Metriken umfassen, die aus der folgenden Gruppe ausgewählt sind: ein Statussignal, eine Anzahl von Befehlen, eine Bandbreiten-Auslastungsmetrik und ein Schwellenwert.

Beispiel 19 ist das Verfahren nach irgendeinem der Beispiele 15-18, ferner umfassend: Vergleichen der statistischen Informationen mit einer vorbestimmten Schwelle, um zu bestimmen, ob Befehlsschwenkung oder Drosselung notwendig ist.

Beispiel 20, ist ein System, umfassend: ein Kapplungssubstrat; einen Interposer auf dem Kapplungssubstrat; einen Speicherstapel mit mehreren auf dem Interposer angebrachten vertikal

gestapelten Dies; und einen Haupt-Die, der auf dem Interposer angebracht ist und der thermisch bewusste Überwachungs- und Schaltnetzwerke umfasst, die dafür ausgelegt sind, sicherzustellen, dass die Temperatur jedes der vertikal gestapelten Dies in dem Speicherstapel unter der vorbestimmten thermischen Grenze gehalten wird, um beim Zugreifen auf den Speicherstapel Timingverstöße zu verhindern.

**[0053]** Zum Beispiel können alle optionalen Merkmale der oben beschriebenen Vorrichtungen auch mit Bezug auf das Verfahren oder den Prozess, die hier beschrieben werden, implementiert werden. Das Obige veranschaulicht lediglich die Prinzipien der vorliegenden Offenbarung, und Fachleute können verschiedene Modifikationen vornehmen. Die obigen Ausführungsformen können einzeln oder in beliebiger Kombination implementiert werden.

### Patentansprüche

1. Integrierte Schaltung, umfassend: eine Schnittstellenschaltung, ausgelegt zum Kommunizieren mit einem externen Die-Stapel, der mehrere vertikal gestapelte Dies aufweist; und Steuerschaltkreise, ausgelegt zum Analysieren von dem externen Die-Stapel zugeordneten Informationen und zum Verteilen von Zugriffen auf die mehreren vertikal gestapelten Dies, um zu verhindern, dass irgendwelche der mehreren vertikal gestapelten Dies eine Temperatur aufweisen, die eine vorbestimmte thermische Grenze übersteigt.
2. Integrierte Schaltung nach Anspruch 1, wobei die analysierten Informationen Parameter umfassen, die aus der folgenden Gruppe ausgewählt sind: ein Speicherbankstatus, eine Anzahl von Befehlen, eine Speicherbandbreiten-Auslastungsmetrik, mindestens ein Schwellenwert und eine Anzahl von Seitenöffnungen und -schließungen.
3. Integrierte Schaltung nach einem der Ansprüche 1-2, wobei die analysierten Informationen einen Speicherfehler-Schwellenwert umfassen.
4. Integrierte Schaltung nach einem der Ansprüche 1-3, wobei die analysierten Informationen ein Regionenmuster häufigen Zugriffs eines gegebenen Dies in dem externen Die-Stapel umfassen.
5. Integrierte Schaltung nach einem der Ansprüche 1-4, wobei die analysierten Informationen Temperaturdaten umfassen, die mehr als einem der vertikal gestapelten Dies in dem externen Die-Stapel zugeordnet sind.
6. Integrierte Schaltung nach Anspruch 5, wobei die Temperaturdaten eine erste Temperatursensormessung von einem obersten Die in dem externen

Die-Stapel und eine zweite Temperatursensormessung von einem untersten Die in dem externen Die-Stapel umfassen.

7. Integrierte Schaltung nach Anspruch 6, wobei die Steuerschaltkreise ferner ausgelegt sind zum Berechnen eines Temperaturgradienten auf der Basis der ersten und zweiten Temperatursensormessung zur Schätzung eines Temperaturwerts für einen zwischen dem obersten und untersten Die in dem externen Die-Stapel angeordneten Zwischen-Die.

8. Integrierte Schaltung nach einem der Ansprüche 1-7, wobei die Steuerschaltkreise Speichercontrollerschaltungen umfassen, die dafür ausgelegt sind, für heiße Dies in dem Die-Stapel bestimmten Verkehr zu drosseln.

9. Integrierte Schaltung nach einem der Ansprüche 1-7, wobei die Steuerschaltkreise Speichercontrollerschaltungen umfassen, die dafür ausgelegt sind, Verkehr von einem gestauten Kanal auf einen weniger gestauten Kanal umzurouten, um ein thermisches Profil über den Die-Stapel auszugleichen.

10. Integrierte Schaltung nach einem der Ansprüche 1-9, wobei die Steuerschaltkreise Schaltnetzwerke umfassen, die dafür ausgelegt sind, Zugriffe auf den externen Die-Stapel so zu schwenken, dass auf kalte Dies in dem Die-Stapel häufiger als auf heiße Dies in dem Die-Stapel zugegriffen wird.

11. Integrierte Schaltung nach Anspruch 10, wobei die Schaltnetzwerke eine Adressenübersetzungsschaltung umfassen, die dafür ausgelegt ist, logische Adressen auf entsprechende physische Adressen abzubilden.

12. Integrierte Schaltung nach Anspruch 11, wobei die Adressenübersetzungsschaltung dafür ausgelegt ist, statische Adressenabbildung durchzuführen.

13. Integrierte Schaltung nach Anspruch 11, wobei die Adressenübersetzungsschaltung dafür ausgelegt ist, dynamische Adressenabbildung durchzuführen.

14. Integrierte Schaltung nach einem der Ansprüche 1-13, wobei die externen Daten dafür ausgelegt sind, eine Hauptkopie gegebener Daten und eine gespiegelte Kopie der gegebenen Daten zu speichern, und wobei die Steuerschaltkreise dafür ausgelegt sind, nur auf eine ausgewählte der Hauptkopie und der gespiegelte Kopie mit der niedrigeren Temperatur zuzugreifen.

15. Verfahren zum Betrieb eines integrierten Schaltungs-Dies, umfassend:  
mit einer Schnittstellenschaltung in dem integrierten Schaltungs-Die: Kommunizieren mit einem externen Chip-Stapel;

mit Steuerschaltkreisen in dem integrierten Schaltungs-Die: Empfangen und Analysieren mehrerer verschiedener Arten statistischer Informationen von dem externen Chip-Stapel; und

als Reaktion auf Empfang und Analysieren der mehreren verschiedenen Arten statistischer Informationen, Verwenden der Steuerschaltkreise zum Aufrechterhalten eines ausgeglichenen thermischen Profils in dem externen Chip-Stapel durch Umrouten von Anforderungen zu dem externen Chip-Stapel.

16. Verfahren nach Anspruch 15, wobei Empfangen der mehreren verschiedenen Arten statistischer Informationen Empfangen von Temperaturmesswerten von mindestens zwei Dies in dem externen Chip-Stapel umfassen.

17. Verfahren nach einem der Ansprüche 15-16, wobei Empfangen der mehreren verschiedenen Arten von statistischen Informationen Empfangen von Temperaturmesswerten und anderen Arten von Parametern, die von den Temperaturmesswerten verschieden sind, umfasst.

18. Verfahren nach Anspruch 17, wobei die anderen Arten von Parametern Metriken umfassen, die aus der folgenden Gruppe ausgewählt sind: ein Statussignal, eine Anzahl von Befehlen, eine Bandbreiten-Auslastungsmetrik und ein Schwellenwert.

19. Verfahren nach einem der Ansprüche 15-18, ferner umfassend:  
Vergleichen der statistischen Informationen mit einer vorbestimmten Schwelle, um zu bestimmen, ob Befehlsschwenkung oder Drosselung notwendig ist.

20. System, umfassend:  
ein Kapselungssubstrat;  
einen Interposer auf dem Kapselungssubstrat;  
einen Speicherstapel mit mehreren auf dem Interposer angebrachten vertikal gestapelten Dies; und  
einen Haupt-Die, der auf dem Interposer angebracht ist und der thermisch bewusste Überwachungs- und Schaltnetzwerke umfasst, die dafür ausgelegt sind, sicherzustellen, dass die Temperatur jedes der vertikal gestapelten Dies in dem Speicherstapel unter der vorbestimmten thermischen Grenze gehalten wird, um beim Zugreifen auf den Speicherstapel Timingverstöße zu verhindern.

21. System nach Anspruch 20, wobei die thermisch bewussten Überwachungs- und Schaltnetzwerke ferner ausgelegt sind zum Erhalten von Temperaturdaten, die mehr als einem der vertikal gestapelten Dies in dem Speicherstapel zugeordnet sind.

22. System nach Anspruch 21, wobei die thermisch bewussten Überwachungs- und Schaltnetzwerke ferner ausgelegt sind zum Berechnen eines Temperaturgradienten auf der Basis der Temperaturdaten

zur Schätzung eines Temperaturwerts für einen zwischen einem obersten Die und einem untersten Die in dem Speicherstapel angeordneten Zwischen-Die.

23. System nach einem der Ansprüche 20-22, wobei die thermisch bewussten Überwachungs- und Schaltnetzwerke ferner ausgelegt sind zum Drosseln von für heiße Dies in dem Speicherstapel bestimmtem Verkehr.

24. System nach einem der Ansprüche 20-23, wobei die thermisch bewussten Überwachungs- und Schaltnetzwerke ferner ausgelegt sind zum Umrou-ten von Verkehr von einem gestauten Kanal auf einen weniger gestauten Kanal zum Ausgleichen eines thermischen Profils in dem Speicherstapel.

25. System nach einem der Ansprüche 20-24, wobei die thermisch bewussten Überwachungs- und Schaltnetzwerke ferner ausgelegt sind zum Schwen-ken von Zugriffen auf den Speicherstapel dergestalt, dass häufiger auf kalte Dies in dem Die-Stapel als auf heiße Dies in dem Die-Stapel zugegriffen wird.

Es folgen 9 Seiten Zeichnungen

Anhängende Zeichnungen

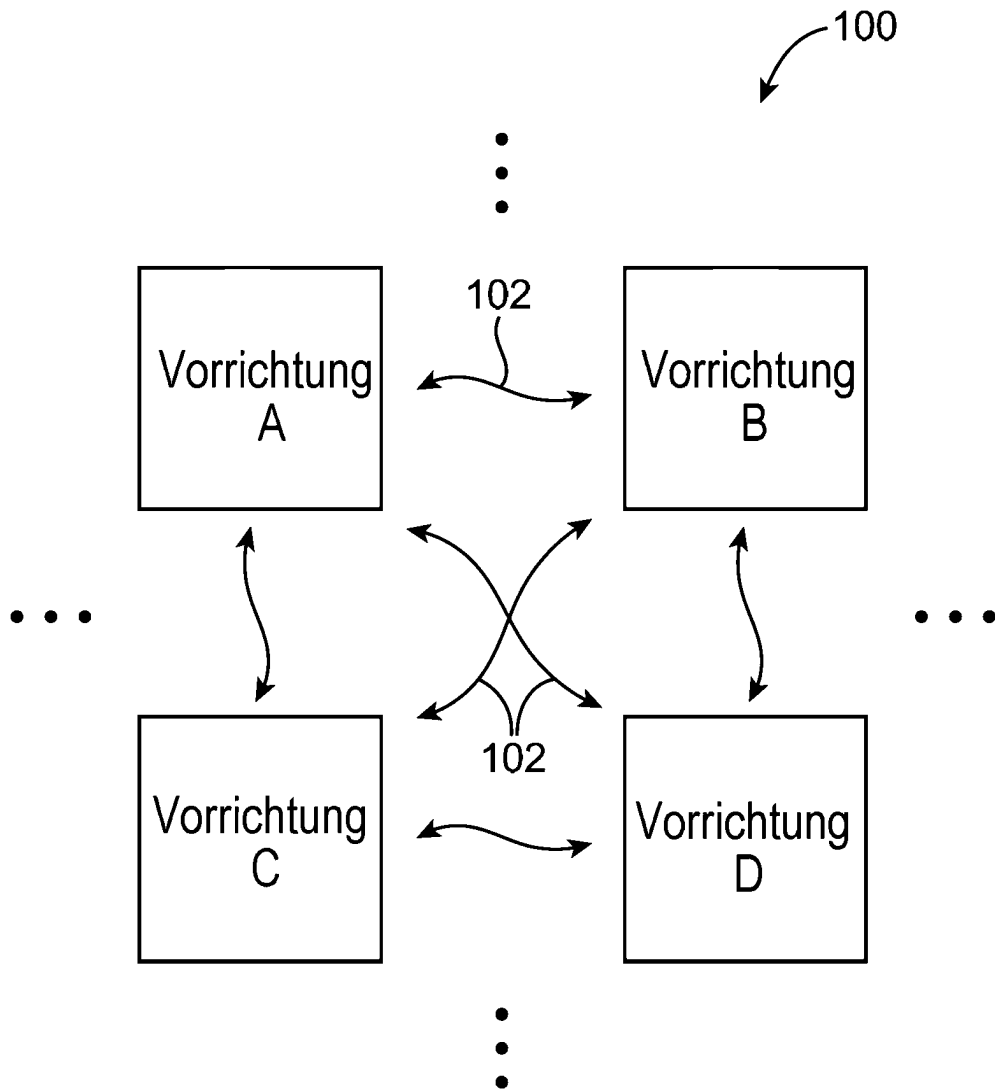


FIG. 1

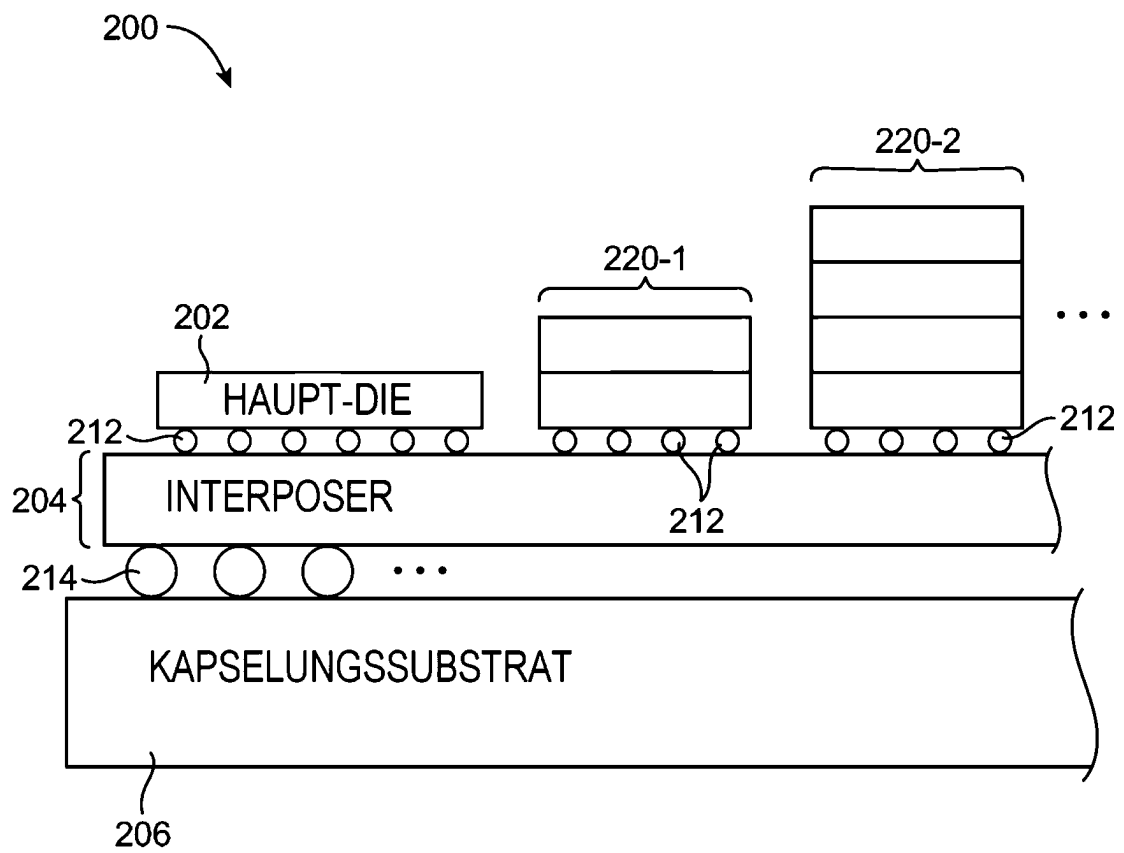


FIG. 2

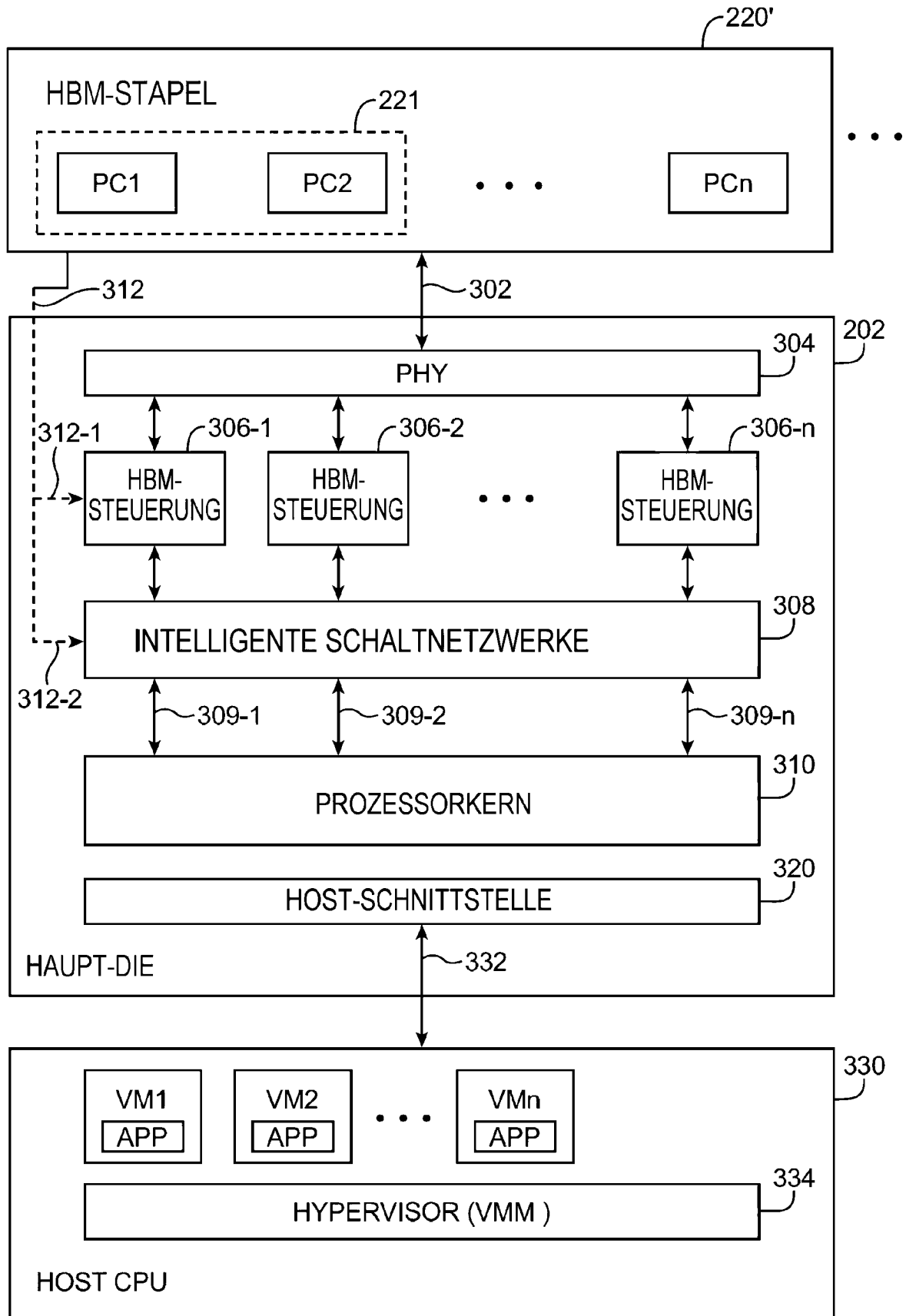


FIG. 3



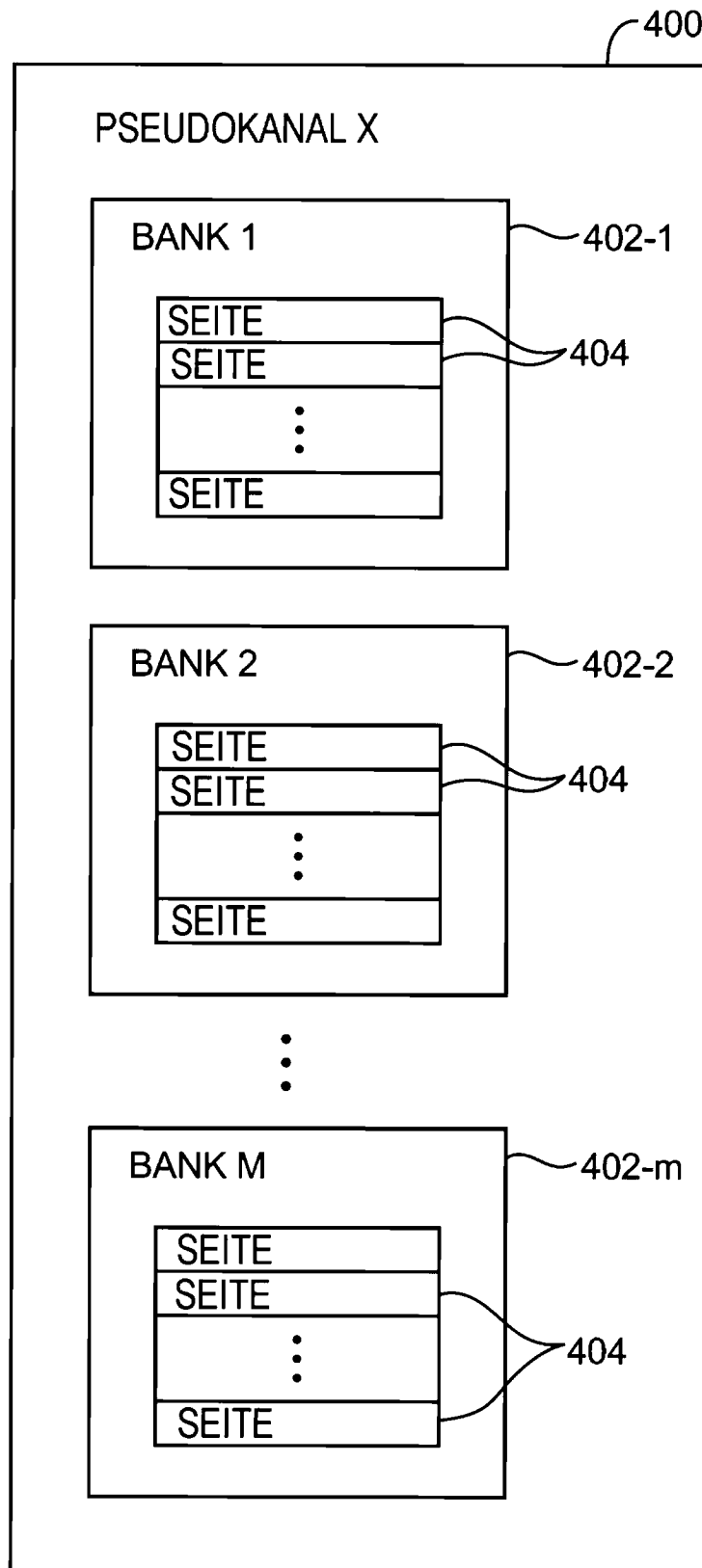


FIG. 4

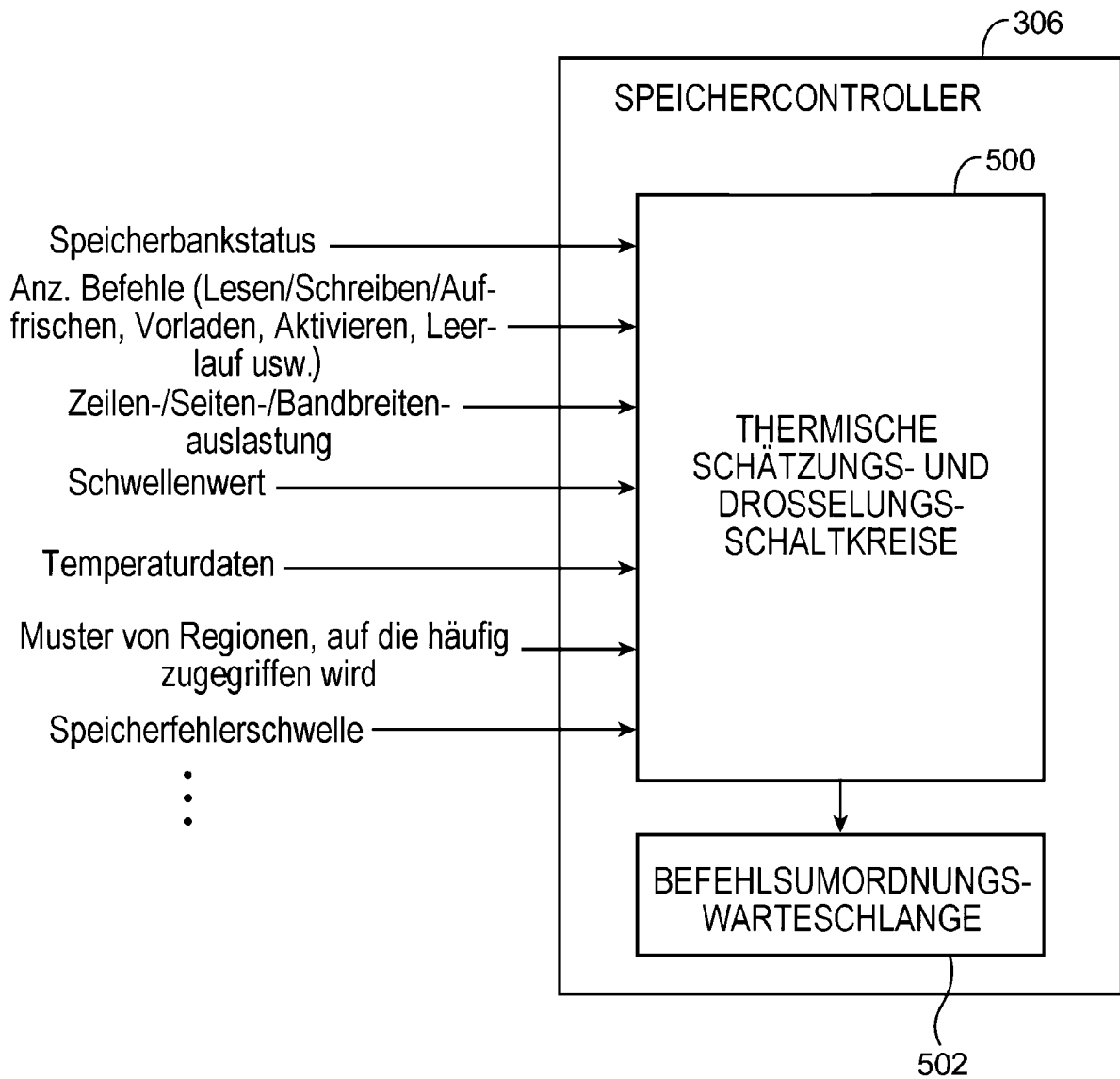


FIG. 5

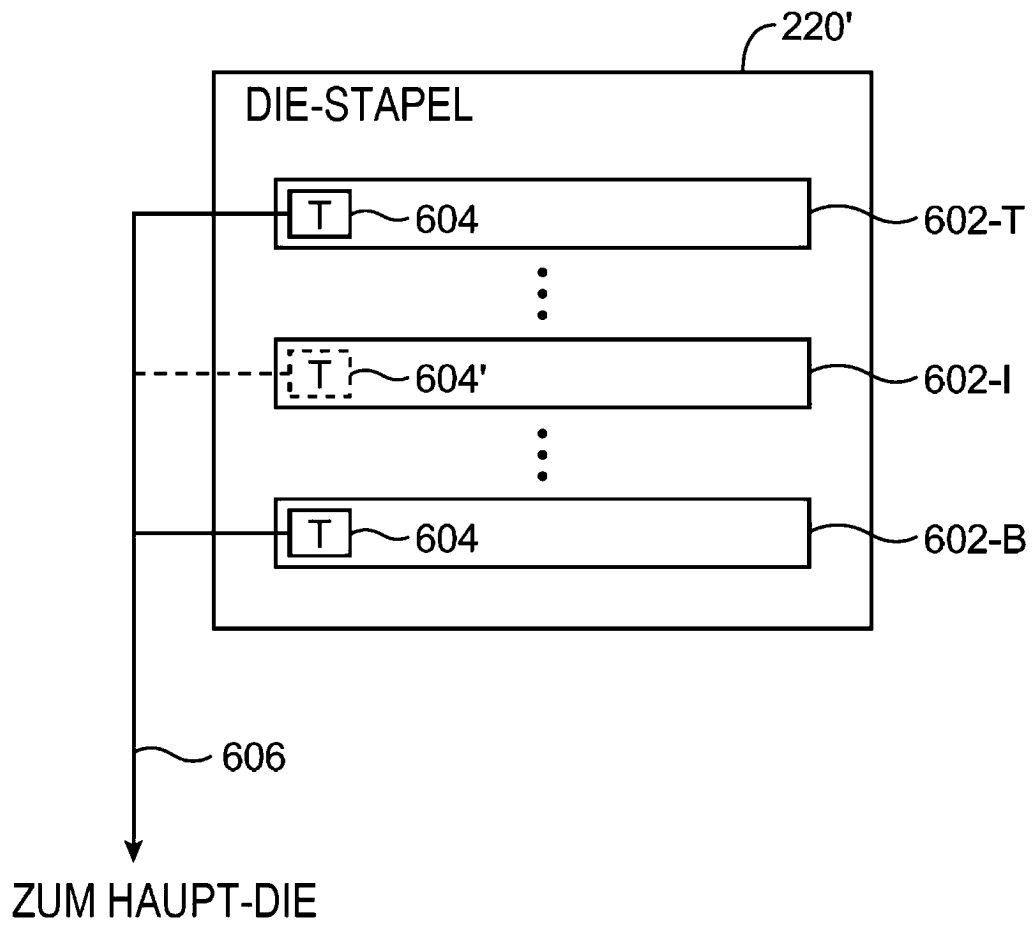


FIG. 6

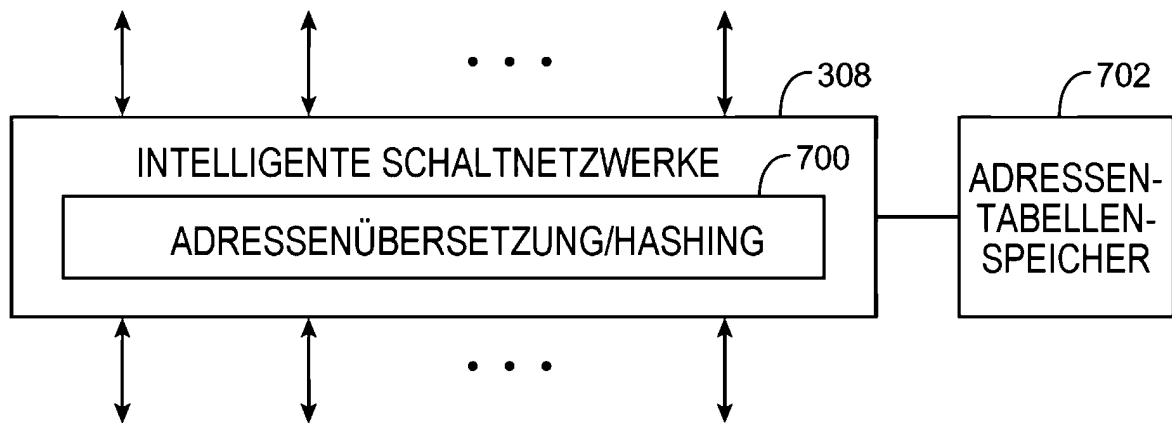


FIG. 7

800 

		Benutzer-Adressen- bereich	Physischer Adressen- bereich
Statische Region	Statische Region 1	0x0000-0x1FFF	0x0000-0x1FFF
	Statische Region 2	...	...
	...	...	...
	Statische Region M	...	...
Dynamische Region	Dynamische Region 1	...	...
	Dynamische Region 2	...	...
	...	...	...
	Dynamische Region N	...	...

802 {  
804 {

FIG. 8

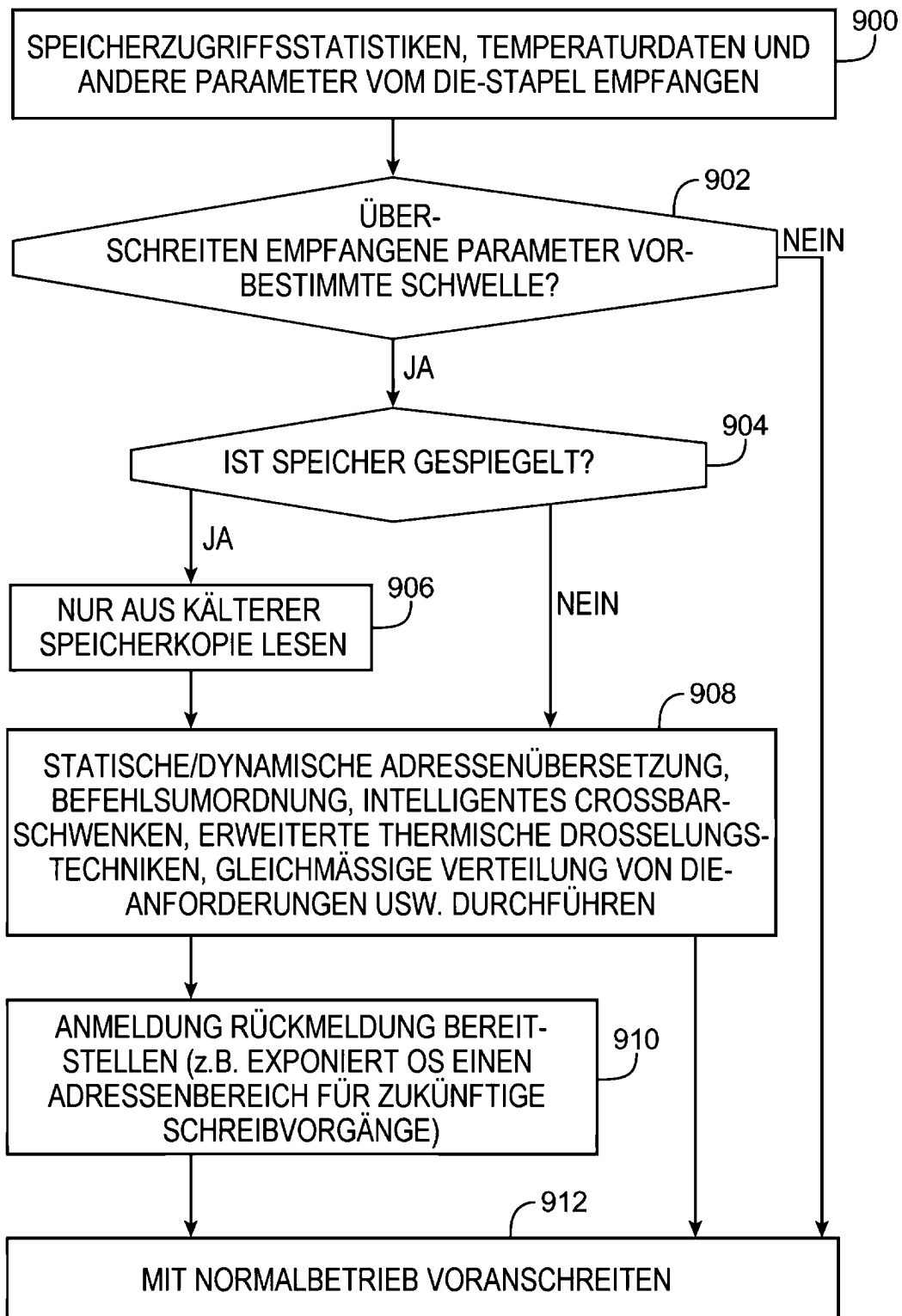


FIG. 9