

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
4 February 2010 (04.02.2010)

(10) International Publication Number
WO 2010/014200 A1

(51) International Patent Classification:
G06F 9/455 (2006.01) *G06F 9/46* (2006.01)

(74) Agent: MUNYON, Dean, M.; Meyertons, Hood, Kivlin, Kowert & Goetzl, P.C., P.O. Box 398, Austin, TX 78767-0398 (US).

(21) International Application Number:
PCT/US2009/004349

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(22) International Filing Date:
28 July 2009 (28.07.2009)

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(25) Filing Language: English
(26) Publication Language: English
(30) Priority Data:
61/084,008 28 July 2008 (28.07.2008) US

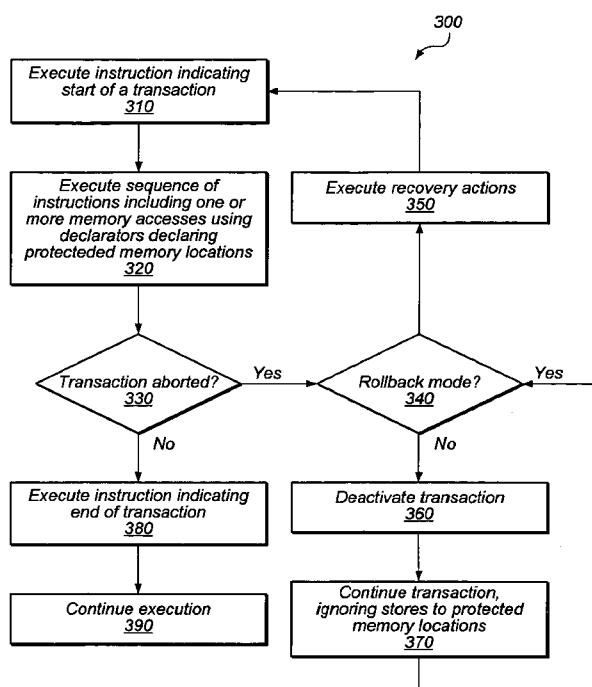
(71) Applicant (for all designated States except US): ADVANCED MICRO DEVICES, INC. [US/US]; One Amd Place, P.o. Box 3453, Sunnyvale, CA 94088 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): CHRISTIE, David, S. [CA/US]; 6201 Needham Lane, Austin, TX 78739 (US). HOHMUTH, Michael, P. [DE/DE]; Bettina Street 14, 01099 Dresden (DE). DIESTELHORST, Stephan [DE/DE]; Ostra-allee 29, 01067 Dresden (DE).

[Continued on next page]

(54) Title: VIRTUALIZABLE ADVANCED SYNCHRONIZATION FACILITY



(57) Abstract: A system and method for executing a transaction in a transactional memory system is disclosed. The system includes a processor of a plurality of processors coupled to shared memory, wherein the processor is configured to execute a section of code, including a plurality of memory access operations to the shared memory, as an atomic transaction relative to the execution of the plurality of processors. According to embodiments, the processor is configured to determine whether the memory access operations include any of a set of disallowed instructions, wherein the set includes one or more instructions that operate differently in a virtualized computing environment than in a native computing environment. If any of the memory access operations are ones of the disallowed instructions, then the processor aborts the transaction.

FIG. 3



Published:

— *with international search report (Art. 21(3))*

TITLE: VIRTUALIZABLE ADVANCED SYNCHRONIZATION FACILITY**BACKGROUND**

[0001] Shared-memory computer systems allow multiple concurrent threads of execution to access shared memory locations. Unfortunately, writing correct multi-threaded programs is difficult due to the complexities of coordinating concurrent memory access.

[0002] Traditionally, to ensure program correctness, programmers have used locks or other mutual exclusion mechanisms for coordinating access to shared memory locations. For example, using traditional locks, a thread may be configured to acquire and hold a lock on each memory location to which it needs exclusive access. While the thread holds the lock, no other thread may acquire the lock, and therefore, no other thread may access the memory location protected by that lock. However, traditional locking techniques are vulnerable to various pitfalls, including dead-lock, race conditions, priority inversions, software complexity, and performance limitations.

[0003] An alternative approach to concurrency control is transactional memory. In a transactional memory programming model, a programmer may designate a section of code (i.e., an execution path or a set of program instructions) as a “transaction” which should be executed atomically with respect to other threads of execution. For example, if the transaction includes two memory store operations, then the transactional memory system ensures that all other threads may only observe either the cumulative effects of both memory operations or of neither, but not the effects of only one. In addition to a simplified programming model, transactional memory systems can also increase application performance since they may allow finer grained memory sharing than do traditional locks.

[0004] Various transactional memory systems have been proposed in the past, including those implemented in software, in hardware, or in a combination thereof. However, many previous concepts and implementations are bound by various limitations. For example, software-based transactional memory systems (STMs) suffer an undesirable performance overhead while hardware proposals (HTMs) may be prohibitively complex to implement.

25

SUMMARY

[0005] A system and method for executing a transaction in a transactional memory system is disclosed. The system includes a processor of a plurality of processors coupled to shared-memory. The processor may be configured to execute a section of code, which includes a plurality of memory access operations to the shared-memory, as an atomic transaction relative to the execution of the plurality of processors.

[0006] According to some embodiments, the processor may be configured to determine whether the memory access operations include any of a set of disallowed instructions, wherein the set of disallowed instructions includes one or more instructions that operate differently in a virtualized computing environment than in a native computing environment. If any of the memory access operations are ones of the disallowed instructions, then the processor aborts the transaction.

35

BRIEF DESCRIPTION OF DRAWINGS

[0007] FIG. 1 is a block diagram illustrating components of a multi-processor computer system configured to implement ASF, according to various embodiments.

[0008] FIG. 2 is a block diagram illustrating a more detailed view of components comprising a processor, according to one embodiment.

[0009] FIG. 3 is a general flow diagram illustrating a method for executing a transaction using ASF, according to various embodiments.

[0010] FIG. 4 is a block diagram illustrating an example of a nested transaction, according to one embodiment.

5 [0011] FIG. 5 is a flow diagram illustrating an implementation of a SPECULATE instruction to begin a speculative section in rollback mode, according to some embodiments.

[0012] FIG. 6 is a flow diagram illustrating the steps of executing a BEGIN instruction to begin a speculative section in store-conditional mode, according to some embodiments.

10 [0013] FIG. 7 is a table summarizing a set of rules defining how various ASF implementations may handle data contention, according to some embodiments.

[0014] FIG. 8 is a flow diagram illustrating a method by which, according to various embodiments, ASF mechanisms may execute a transaction in rollback mode.

[0015] FIG. 9 illustrates a method for committing a transaction, such as by executing a COMMIT instruction, according to some embodiments.

15 [0016] FIG. 10 is a flow diagram illustrating a method for executing a transaction in store-conditional mode, according to some embodiments.

[0017] FIG. 11 is a flowchart illustrating a method for configuring ASF-based transactions to interoperate with lock-based code, according to one embodiment.

20 [0018] FIG. 12 illustrates one embodiment of a computer system configured to implement various embodiments of ASF, as described herein.

[0019] While the invention is described herein by way of example for several embodiments and illustrative drawings, those skilled in the art will recognize that the invention is not limited to the embodiments or drawings described. It should be understood that the drawings and detailed description hereto are not intended to limit the invention to the particular form disclosed, but on the contrary, the invention is to cover all modifications, equivalents and alternatives falling within the spirit and scope of the present invention as defined by the appended claims. Any headings used herein are for organizational purposes only and are not meant to limit the scope of the description or the claims. As used herein, the word "may" is used in a permissive sense (i.e., meaning having the potential to) rather than the mandatory sense (i.e. meaning must). Similarly, the words "include", "including", and "includes" mean including, but not limited to.

30 [0020] Transactional memory systems may allow software threads in multi-threaded systems to access (read and/or write) a set of shared memory locations atomically with respect to other threads, without requiring the overhead and programming complexity associated with traditional synchronization mechanisms such as mutual-exclusion using locks. However, further techniques are necessary for increasing the efficiency of transactional memory systems.

[0021] According to some embodiments, various hardware mechanisms may be used to implement efficient transactional memory mechanisms, as described herein. In some embodiments, computationally inexpensive primitives (e.g., program instructions) may be defined for instructing the hardware to perform various synchronization functions, which may be used to synthesize higher-level synchronization mechanisms. The

hardware mechanisms and/or program instructions may collectively be referred to herein as the Advanced Synchronization Facility (ASF).

[0022] In some embodiments, an existing processor architecture (e.g., x86) may be augmented to implement ASF mechanisms. For clarity of explication, the remainder of this disclosure describes how the known 64-bit x86 architecture AMD64 may be augmented with ASF mechanisms to implement transactional memory. However, these example embodiments are not intended to limit ASF mechanisms to this architecture alone and given the benefit of this disclosure, implementations of ASF mechanisms in other processor architectures will become evident to those skilled in the art.

[0023] FIG. 1 is a block diagram illustrating components of a multi-processor computer system configured to implement ASF, according to various embodiments. According to the illustrated embodiment, computer system 100 may include multiple processors, such as processors 110a and 110b. In some embodiments, processors 110 may comprise multiple physical or logical (e.g., SMT) cores and be coupled to each other and/or to a shared memory 150 over an interconnect, such as 140. In various embodiments, different interconnects may be used, such as a shared system bus or a point-to-point network in various topographies (e.g., fully connected, torus, etc.).

[0024] According to the illustrated embodiment, each processor 110 may include one or more levels of memory caches 130. Levels of memory caches may be hierarchically arranged (e.g., L1 cache, L2 cache, L3 cache, etc.) and may be used to cache local copies of values stored in shared memory 150.

[0025] In various embodiments, memory caches 130 may include various cache-coherence mechanisms 132. Cache-coherence mechanisms 132 may, in one embodiment, implement a cache coherence communication protocol among the interconnected processors to ensure that the values contained in memory caches 130 of each processor 110 are coherent with values stored in shared memory and/or in the memory caches of other processors. Several such protocols exist (including the MESI (i.e., Illinois protocol) and MOESI protocols), and may be implemented in various embodiments. Cache coherence protocols may define a set of messages and rules by which processors may inform one another of modifications to shared data and thereby maintain cache coherence. For example, according to the MESI protocol, each block stored in a cache must be marked as being in one of four states: modified, exclusive, shared, or invalid. A given protocol defines a set of messages and rules for sending and interpreting those messages, by which processors maintain the proper markings on each block. Depending on the state of a given cache block, a processor may be restricted from performing certain operations. For example, a processor may not execute program instructions that depend on a cache block that is marked as invalid. Cache coherence mechanisms may be implemented in hardware, software, or in a combination thereof, in different embodiments. Cache coherence messages may be communicated across interconnect 140 and may be broadcast or point-to-point.

[0026] According to the illustrated embodiment, each processor 110 may also include various ASF mechanisms for implementing transactional memory, as described herein. In various embodiments, more processors 110 may be connected to interconnect 140, and various levels of cache memories may be shared among multiple such processors and/or among multiple cores on each processor.

[0027] FIG. 2 is a block diagram illustrating a more detailed view of components comprising a processor, such as processors 110, according to one embodiment. According to the illustrated embodiment of FIG. 2, processor 200 comprises register file 210, which may include various registers, each of which may be of any size (e.g., 16-bit, 32-bit, 64-bit, 128-bit, etc.) For example, register file 210 may include various known x86 registers, such as rSP register 212 (stack pointer), rIP register 214 (instruction pointer), rAX register 216 (accumulator register), and/or

rFLAGS register 218 (flags register indicating processor state). In some embodiments, register file 210 may further comprise any number of general purpose registers 220 and/or floating point registers 222. In some embodiments, register file 210 may include one or more 128-bit registers, such as XMM registers 224. In various embodiments, register file 210 may comprise any number of other registers 226, which may be of various sizes.

5 [0028] According to the illustrated embodiment, processor 200 may include memory caches 230 for storing local copies of values in shared memory and cache-coherence mechanisms 232 for maintaining the consistency of those values across various copies dispersed within shared memory, the cache, and other caches. Processor 200 may also include ASF mechanisms 240 for implementing transactional synchronization mechanisms, as described herein. ASF mechanisms 240 may include the data structures and/or logic to implement memory transactions as described herein, according to various embodiments. In some embodiments, ASF mechanisms 240 may include a locked line buffer 242, which may be used in case of an abort to roll back memory changes made by a partially executed transaction, as described below. In some embodiments, ASF mechanisms 240 may include a nesting level register 244, which may hold a value indicating the depth to which the current transaction is nested in other transactions, as described below in more detail. In some embodiments ASF mechanisms may include a flag to indicate whether a transaction is active, such as transaction active flag 246. In some embodiments, ASF mechanisms may include a flag indicating a transactional mode, such as store-conditional mode flag 248, which may be set to indicate whether an executing transaction is executing in store-conditional mode or in another mode. In further embodiments, other flags and/or registers may be implemented in ASF mechanisms 240 in support of transactional execution.

10

15

[0029] FIG. 3 is a general flow diagram illustrating a method for executing a transaction using ASF, according to various embodiments. According to the illustrated embodiment, ASF may allow software (e.g., a thread) to begin a transaction (i.e., a *critical section* of code) by executing a given instruction (or multiple instructions) indicating the start of a transaction, as in 310. As discussed later, in various embodiments, ASF instructions such as SPECULATE or BEGIN may be used to indicate the start of a critical section.

[0030] After beginning a transaction, the thread may execute a series of instructions comprising the transaction body, as in 320. Such instructions may include a number of memory access (read and/or write) operations, some number of which may designate that the accessed memory location should be *protected*. A thread may designate that a given memory location should be protected by using one or more special *declarator* instructions provided by ASF. ASF mechanisms may ensure that access to protected memory (as designated by declarator instructions) occur atomically with respect to all other concurrently executing threads in the system (i.e., all at once or not at all).
20
25

30 For example, if the transaction includes multiple protected writes to memory, then ASF mechanisms may ensure that no other thread may observe the result of only a subset of those protected writes to memory. In another example, according to some embodiments, if the transaction includes one or more protected memory read operations, then ASF mechanisms may ensure that the transaction completes successfully only if no other thread has modified the protected read locations before the transaction has completed. In various embodiments, a “memory location” protected by a declarator operation may be of different sizes, such as that of a system dependent cache block or of another size.

35

[0031] According to such embodiments, ASF may be configured to protect only memory lines that have been specified using designated declarator instructions. In such embodiments, all other memory locations may remain unprotected and may be modified inside a critical section using standard x86 instructions. These modifications to unprotected memory may become visible to other CPUs immediately upon execution, for example, in program

40

order. In some embodiments, a transaction body (e.g., instructions executed in 320) may comprise one or more other transactions (i.e., nested transactions).

[0032] According to various embodiments, if one or more conditions of the transactional execution have been violated, then the transaction may abort (as indicated by the affirmative exit from 330). In various embodiments, 5 transactional attempts may be aborted at any point because of contention with other processors, far control transfers (such as those caused by interrupt and faults), execution of explicit software ABORT instructions, insufficient hardware resources, other implementation-specific conditions, etc.

[0033] In various embodiments, when a transactional attempt aborts, different recovery actions may be taken depending on the mode of transactional execution. In various embodiments, for each transaction, ASF may support 10 a *rollback* mode of execution and/or a *store-conditional* mode of execution. In some embodiments, the mode of execution may determine what actions are performed in response to a transactional attempt being aborted. For example, in some embodiments, aborting a transaction in rollback mode may cause execution to be “rolled back” to the start of the transaction while aborting a transactional attempt in store-conditional mode may cause the transaction to continue but with transactional store operations not being performed (i.e., execution of transactional 15 stores may be conditional on the transaction not having been aborted).

[0034] According to the illustrated embodiment, when a transaction executing in rollback mode aborts (as indicated by the affirmative exit from 340), ASF mechanisms may be configured to execute recovery actions, as in 350, which may include discarding modifications to the contents of the protected lines. By discarding such modifications, as in 350, ASF mechanisms may cause the modifications to be unobservable to other threads in the 20 system, thereby complying by the atomicity property of the transaction. However, ASF mechanisms may be configured to not roll back modifications to unprotected memory, such as those performed using conventional x86 memory instructions. In some embodiments, the application programmer may provide software for accommodating these unprotected modifications, such as software recovery routines configured to reenter an initialization sequence leading up to the critical section.

[0035] In various embodiments, the recovery actions of 350 may be configured to roll back only a subset of the 25 system registers (e.g., rIP and rSP) rather than all registers. In such embodiments, software may therefore be written to not rely on the content of various registers when entering transactional execution (e.g., by ignoring the initial contents of some registers after an abort event and/or to not modifying various registers during transactional execution).

[0036] In some instances, before an interrupt or exception handler returns, operating system code or other 30 processes may have executed in the interim. Furthermore, in some instances, other processes may have even executed ASF transactions that inspected and/or modified locations targeted by the interrupted transaction. In some embodiments, ASF mechanisms may obviate these concerns by not maintaining any ASF-related state across context switches. Instead, in such embodiments, when the interrupted thread returns to the processor, ASF 35 mechanisms may be configured to automatically abort and reattempt the transaction.

[0037] According to the illustrated embodiment, after executing some recovery action or actions, the thread may then reattempt the transaction, as indicated by the feedback loop from 350 to 310.

[0038] In some embodiments, an abort of a transaction executing in store-conditional mode (as indicated by the 40 negative exit from 340) may be handled differently from an abort of a transaction executing in rollback (as indicated by the affirmative exit from 340). For example, while an abort of a rollback mode transaction may automatically reset execution flow to the beginning of the critical section (or to other recovery code), an abort of a

transaction in store-conditional mode may be handled by undoing or otherwise discarding changes made by the critical section to values in protected memory locations and then “deactivating” the transaction, as in 360. In various embodiments, deactivating the transaction may include setting a status code indicating that an abort has occurred and/or that the transaction is no longer active. In store-conditional mode, after the transaction is deactivated (as in 360), the system may continue to execute the critical section without executing any store operations to protected memory locations (as in 370). For example, after the transaction is deactivated, subsequent store operations to protected memory locations may be treated as no-ops.

[0039] In some embodiments, when the aborted transaction completes, a status code may be set in a register, such as rAX register 216 to indicate that the transaction was aborted. In some embodiments, the thread may be configured to detect that an abort status code has been set and perform one or more recovery actions, as indicated by the feedback loop from 370 to 350. In various embodiments, the recovery actions of 350 may be different for transactions that were aborted in store-conditional mode (i.e., from 370) than for those that were aborted in rollback mode (i.e., from 340). A more detailed discussion of transactions in rollback and store-conditional modes is provided below.

[0040] According to the illustrated embodiment, once the transaction body has been completed, the thread may execute one or more instructions indicating the end of the transaction to the ASF mechanisms, as in 380, such as a COMMIT instruction as discussed below. After committing the transaction, as in 380, the thread may continue execution, as in 390.

[0041] In some embodiments, ASF may support nesting of one or more transactions within one or more other transactions. For example, after a transaction is started as in 310 (e.g., by executing a SPECULATE command), the instructions of the transaction body being executed in 320 may begin another transaction (e.g., by executing another SPECULATE command) before the first transaction completed. In this case, the second (“child”) transaction may be said to be “nested” within the first (“parent”) transaction. A transaction that is not nested within any other transaction may be referred to herein as a “top-level” transaction.

[0042] FIG. 4 illustrates an example of a nested transaction, according to one embodiment. According to FIG. 4, top-level parent transaction 400 is a rollback mode transaction that comprises two nested transactions: nested parent transaction 410, which is also in rollback mode and nested transaction 440, which is store-conditional mode. According to the illustrated embodiment, nested parent transaction 400 contains two nested transactions, including nested transaction 420 and nested transaction 430, which are in rollback and store-conditional modes respectively.

[0043] In some embodiments, ASF mechanisms may be configured to flatten nested transactions (e.g., composed of multiple SPECULATE-COMMIT pairs) into a single transaction. In some embodiments, nested transactions may share ASF hardware resources. That is, in some embodiments, during execution of a nested transaction, the memory locations protected by the parent transaction remain protected in the child transaction. Furthermore, in some embodiments, memory locations protected in a nested transaction may remain protected in the parent transaction, even after the nested transaction has completed. Thus, in some embodiments, parent transactions may need to continue to use protected memory operations when dealing with memory locations protected by a child transaction. In such embodiments, use of a regular memory write operation by the parent on a memory location protected by a child transaction may result in a general protection fault being raised (i.e., #GP exception).

[0044] In some embodiments, one or more store-conditional transactions may be nested within one or more rollback transactions. In some embodiments, due to the flattening of the nested transactions, ASF mechanisms may be configured to respond to an abort of a nested store-conditional transaction by performing the recovery actions of

the top-level parent transaction (as in 350). However, according to some embodiments, a nested transaction may not be permitted inside of a store-conditional transaction.

[0045] To enable transactional execution, ASF mechanisms may define and expose various coherency control instructions that can be invoked by software threads. While some instructions may be named differently in different embodiments, various embodiments may expose coherency control instructions matching or analogous to one or more of the following.

[0046] FIG. 5 illustrates an implementation of a SPECULATE instruction, according to some embodiments. As discussed above, in some embodiments, the SPECULATE instruction may be executed by a thread to indicate the start of a transaction in rollback mode. According to the illustrated embodiment, executing the SPECULATE operation may comprise checking whether the transaction is nested beyond a maximum nesting depth supported by the system, as in 505. In various embodiments, different registers and/or memory locations may be used to hold the depth of the currently executing transaction. If the transaction is deeper than the system-defined maximum nesting depth, as indicated by the affirmative exit from 505, then ASF mechanisms may be configured to raise a general protection fault (#GP[0]), as in 515.

[0047] According to the illustrated embodiment, executing the SPECULATE instruction may further comprise determining whether the transaction is nested within a store-conditional transaction, as in 510. In some embodiments wherein transactions cannot be nested within store-conditional transactions, in response to detecting that the parent transaction is a store-conditional transaction, as indicated by the affirmative exit from 510, then the ASF mechanisms may be configured to raise a general protection fault (or take some other action), as in 515. In various embodiments, a register or memory location used to indicate the current mode of execution may be read to determine whether a parent transaction is executing in store-conditional mode. In further embodiments, various other checks for the transaction's validity may be performed and a GP fault may be raised if one or more are violated.

[0048] According to the illustrated embodiment, if the transaction is valid (e.g., does not exceed a maximum nesting depth and is not nested within a store-conditional transaction), then the rAX register (or other suitable register) may be reset, such as by setting it to a zero value. In some embodiments, the rAX register may hold various values and codes indicative of transactional execution status, as is described below.

[0049] According to the illustrated embodiment, the execution of SPECULATE may then determine whether the transaction is a top level transaction, as in 530. In some embodiments, this determination may comprise checking the value of the incremented nesting level to determine whether the current transaction is nested. If the transaction is a top-level transaction, as indicated by the affirmative exit from 530, then executing SPECULATE may comprise recording a checkpoint including information usable by the processor to return execution to a given rollback point if the transaction is aborted. In some embodiments, recording the checkpoint may include recording the values that the instruction pointer (e.g., rIP 214) and stack pointer (e.g., rSP 212) will have immediately after the SPECULATE instruction has been executed, as in 535 and 540. Thus, in case of an abort, the recorded checkpoint may be used to transfer control to the instruction immediately following the SPECULATE instruction. In some embodiments, executing SPECULATE may further comprise setting or modifying values in one or more other registers, such as in rFLAGS register 218 to indicate processor states.

[0050] In some embodiments, while a SPECULATE instruction may begin a transaction in rollback mode, a different instruction (e.g., BEGIN) may begin a transaction in store-conditional mode.

[0051] FIG. 6 illustrates the steps of executing a BEGIN instruction, according to some embodiments. Method 600 may begin by determining one or more validity conditions for the transaction. In some embodiments, these conditions may be analogous to those checked when executing a SPECULATE transaction. For example, in method 600, executing BEGIN comprises determining whether a maximum nesting level has been reached (as in 605) and determining whether the current transaction is being nested within another store-conditional section (as in 610). If either condition is true (as indicated by the affirmative exits from 605 and 610 respectively), ASF mechanisms may be configured to raise a GP or other fault, as in 615. Otherwise, according to method 600, executing BEGIN may include resetting the rAX or other register (e.g., to zero), as in 620, incrementing the nesting level (as in 625), and setting a flag indicating that the current transaction is in store-conditional mode (as in 630) and proceeding with transactional execution of the transaction body (as in 635). In various embodiments, the flag set in 630 may be used to determine whether the current transaction is in store-conditional mode, such as in steps 510 and 610.

[0052] In some embodiments, ASF mechanisms may define various memory-reference instructions, called *declarators*, for designating which memory locations should be protected as part of a transaction (i.e., the memory locations for which atomic access is desired). For example, in some embodiments, ASF mechanisms may expose a declarator memory access instruction for reading memory that is analogous to the standard x86 MOV instruction. Like the traditional MOV instruction, an ASF-defined LOCK MOV instruction may be used for loading values from memory. However, according to some embodiments, if a thread reads a memory location inside of a transaction using the LOCK MOV instruction, then ASF mechanisms may add the memory cache block containing the first byte of the referenced memory location to the set of protected cache blocks. In some embodiments, software mechanisms may ensure that unaligned memory accesses do not span both protected and unprotected lines. According to some embodiments, a declarator instruction referencing an already protected cache block may behave like a regular memory reference and not change the protected status of the block. In some embodiments, declarators may not be permitted outside of a critical section (e.g., an exception may be raised). In some 20 embodiments declarators outside of critical section may be treated like regular memory operations. This may be the case for some embodiments wherein store-conditional mode is available.

[0053] In some embodiments, ASF mechanisms may define declarators other than LOCK MOV instructions. For examples, instructions analogous to x86 PREFETCH and/or PREFETCHW may be used (e.g., LOCK PREFETCH, LOCK PREFETCHW). Like their x86 analogs, LOCK PREFETCH and LOCK PREFETCHW may be used to 30 fetch a value from memory into cache for reading (PREFETCH) or for writing (PREFETCHW). However, unlike standard prefetches without a LOCK prefix, LOCK PREFETCH and LOCK PREFETCHW may make a memory location protected. In addition, in some embodiments, LOCK PREFETCH and LOCK PREFETCHW may also check the specified memory address for translation faults and memory-access permission (read or write, respectively) and generate a page fault if unsuccessful. In some embodiments, LOCK PREFETCH and LOCK 35 PREFETCHW may generate a #DB exception when they reference a memory address for which a data breakpoint has been configured.

[0054] In some embodiments, once a memory location has been protected using a declarator, it may be modified again speculatively, but not nonspeculatively, within the transaction. For example, in some embodiments, after a 40 memory location has been read using a LOCK MOV read instruction, the value stored in the memory location may be speculatively modified using an ASF-defined LOCK MOV store instruction. According to embodiments, such speculative updates may become visible only when the transaction is committed. According to such embodiments,

if the transactional attempt aborts, then speculative updates are rolled back and/or otherwise discarded. In some embodiments, during transactional execution, a memory location that has been protected using a declarator may only be modified using a speculative store instruction. In such embodiments, if a thread in transactional mode attempts to modify the value of a protected memory location using conventional non-speculative store operations,

5 ASF mechanisms may raise an exception/fault (e.g., #GP(0)).

[0055] In some embodiments, if a speculative store instruction (e.g., LOCK MOV) is executed outside of a critical section, then an exception may be raised. In other embodiments, such as those wherein store-conditional mode is available, speculative stores outside of a critical section may be treated as a no-op.

[0056] In various embodiments, declarator memory instructions may participate in a system's cache coherence protocol. For example, if a LOCK MOV or LOCK PREFETCH instruction for reading a memory location misses in cache, it may send a non-invalidating probe to other processors, as dictated by the system's cache coherence protocols. In another example, if a given cache line does not already reside in the local cache in exclusive/owned status, then modifying it using a LOCK MOV instruction may result in sending an invalidating probe to other processors, as dictated by the system's cache coherence protocols. In some embodiments, executing a LOCK PREFETCHW instruction may also result in sending an invalidating probe, etc.

[0057] Various ASF mechanisms may be used to monitor and/or enforce protected memory locations (e.g., protected cache blocks). In some embodiments, ASF mechanisms may include an extension to one or more of the system's caches (e.g., memory caches 230), to indicate which cache lines (i.e., blocks) are protected. For example, in one embodiment, each line in a given cache may comprise a "protected" flag, which is set if the cache line is protected in a currently executing transaction and unset if it is not protected. In some embodiments, the protected flag may comprise one or more bits.

[0058] In some embodiments, ASF mechanisms for tracking protected cache lines may comprise a *locked line buffer*. In such embodiments, when a value in a protected memory location is modified (e.g., using a LOCK MOV instruction), an entry may be made into the locked line buffer to indicate the cache block and the value it held before the modification. In such embodiments, in the event of an abort of the transaction, the entries of the locked line buffer may be used to restore the pre-transaction values of each protected cache line to the local cache. In such embodiments, the locked line buffer may participate in a cache coherence protocol of the system, as described below.

[0059] In other embodiments, instead of using a locked line buffer to undo memory stores as described above, 30 various ASF implements may instead prevent store instructions to protected memory locations in a critical section from being written to cache and/or memory before the transaction is committed. For example, ASF may be configured to keep all memory modifications in an internal store buffer and forward buffered values to subsequent load operations in the transaction. In such embodiments, once the transaction commits, ASF mechanisms may allow the buffered store operations in the internal store buffer to be written back to the cache.

[0060] In some embodiments, due to the fixed capacity of various ASF hardware components, various limitations 35 may exist on the number of memory locations that may be simultaneously protected during a transaction (or set of nested transactions). For example, as discussed above, according to some embodiments, an implementation may require that all protected memory locations simultaneously reside in a data cache (e.g., memory caches 230) for the duration of the critical section. In such an embodiment, if a protected line is evicted from the cache during a transaction (e.g., due to capacity constraints), the critical section may be aborted. For example, a critical section 40 that happened to reference N+1 memory locations that all mapped to the same set in an N-way associative data

cache may fail to complete since at least one protected cache line would be evicted from cache when protected memory location N+1 is accessed. However, if a transaction performs a more distributed reference pattern, then it may be able to concurrently protect more memory locations than N before any one cache index is exceeded and a protected line is displaced from the data cache.

5 [0061] In various embodiments, capacity limitations other than cache associativity may exist. For example, in embodiments in which a locked line buffer is used, the maximum number of concurrently protected modified memory locations may be determined by the capacity of the locked line buffer. In another example, in embodiments utilizing a store-buffering scheme, ASF hardware capacity may be dependent on the capacity of the store buffer (i.e., the maximum number of outstanding stores supported by the system pipeline).

10 [0062] In various embodiments, ASF mechanisms may guarantee that a critical section will not fail due to insufficient hardware capacity as long as the number of protected locations does not exceed a given minimum guaranteed capacity. In various embodiments, this guarantee may be made regardless of where in the cacheable address space protected memory locations reside. For example, in embodiments that require that all protected memory locations simultaneously reside in a data cache, the minimum guaranteed capacity may be dependent upon 15 the data cache's associativity (i.e., size of associativity sets). In various embodiments, if a transaction exceeds the hardware capacity, then the transactional attempt may abort.

[0063] In some embodiments, ASF mechanisms may allow a thread executing a transaction to remove an unmodified protected memory location from the transaction's set of protected memory locations. In some embodiments, the thread may accomplish this by executing an explicit RELEASE instruction provided by ASF. In 20 some embodiments, when a protected memory location is released (e.g., using the RELEASE instruction), then it is no longer monitored for contention with other threads. For example, in embodiments wherein a protected flag is utilized, the value of the protected flag associated with the released cache block may be modified to indicate that the block is no longer protected. Thus, by removing an unmodified protected memory location from the set of protected memory locations, a thread may avoid unnecessary data conflicts with other threads and/or exceeding an 25 ASF implementation's hardware capacity, which may lead to transactional aborts. In some embodiments, a RELEASE instruction may or may not guarantee that the specified protected memory location will be released. In some embodiments, only protected cache lines that have not been modified may be released.

[0064] As described above, in some embodiments, if a transaction attempts to protect (using declarators) more memory locations (e.g., cache blocks) than the ASF implementation's capacity can support then the transaction may 30 be aborted. In other embodiments, the abort may be executed in response to a fault being raised.

[0065] In various embodiments, transactional aborts may also be caused by conditions other than insufficient hardware capacity. For example, an abort may be caused by memory contention, that is, interference caused by another processor attempting to access protected memory locations. In various embodiments, ASF mechanisms may be configured to detect various cache coherence messages (e.g., invalidating and/or non-invalidating probes) 35 that may be relevant to one or more protected cache lines and determine whether the probe indicates that a data conflict exists. In response to detecting a data conflict, the ASF-mechanism may abort a transactional attempt. For example, consider a first thread executing in transactional mode and protecting a memory location (i.e., reading a memory location using a declarator instruction). If a second thread subsequently attempts a store to the protected memory location, then the processor executing the second thread may send an invalidating probe to the processor 40 executing the first thread, in accordance with the particular cache coherence protocol deployed by the system. If the first processor receives the invalidating probe while the memory location is still protected (e.g., before the first

thread commits its transaction or otherwise releases the memory location) then a data conflict may exist. In such an instance, ASF mechanisms may be configured to detect that the invalidating probe is relevant to a protected memory location and in response, abort the first thread's transaction.

[0066] According to various embodiments, a transaction may be aborted if ASF mechanisms detect that an invalidating probe relevant to a protected cache line is received. In some embodiments, a transaction may also be aborted if ASF mechanisms detect that a non-invalidating probe relevant to a modified protected cache line is received.

[0067] FIG. 7 is a table summarizing a set of rules defining how various ASF implementations may handle data contention, according to some embodiments. The table of FIG. 7 describes the outcomes when a first thread executing on a first processor (CPU A) performs a given memory access operation on a given memory location, while a second thread on a second processor (CPU B) is executing a transaction that has protected that location. As described above, various cache coherence protocols allow memory caches to cache memory values in different "states", such as "shared" state for read-only access, or in "owned" or "exclusive" state for write access. The "Protected Shared" and "Protected Owned" columns correspond to a protected cache line being in shared and owned state respectively.

[0068] For example, the top entry in the table of FIG. 7 details how to handle a situation, according to one embodiment, where CPU A is executing in any mode (inside or outside of a critical section) and performs a read operation to a memory location that CPU B is protecting in either shared or owned state. According to the illustrated embodiment, if CPU B is protecting the location in owned state, then the transaction of CPU B aborts and otherwise, the transaction of CPU B does not abort. Similar outcomes would result if CPU A had executed a prefetch operation in any mode or if it had executed a LOCK MOV or LOCK PREFETCH operation from inside of a transaction. However, according to the illustrated embodiment, if CPU A were to perform a write or PREFETCHW operation to the memory location in any mode or a LOCK PREFETCHW operation to the memory location in transactional mode, then CPU B would abort the transaction regardless of whether it held the memory location in shared or owned states.

[0069] In some embodiments, a transaction may be aborted explicitly using an ASF-defined ABORT instruction. In some embodiments, a transaction may be aborted because it attempts to execute a disallowed instruction, such as one that results in a far control transfer. In various embodiments, far control transfers may include instructions that transfer control to a location in another segment, such as by changing the content of the CS register in x86 implementations. Far control transfers may include traps, faults, exceptions, NMIs, SMIs, unmasked and nondeferred interrupts, disallowed instructions converted into exceptions, etc. In some embodiments, disallowed instructions may include privileged instructions, such as those that must be executed at an elevated privilege level (e.g., CPL=0), instructions that cause a far control transfer or an exception, and any instructions that may be intercepted by a secure virtual machine (SVM) hypervisor. In various embodiments, disallowed instructions may include:

- FAR JMP, FAR CALL, FAR RET
- SYSCALL, SYSRET, SYSENTER, SYSEXIT
- INT, INTx, IRET, RSM
- BOUND, UD2
- PUSHF, POPF, PAUSE, HLT, CPUID, MONITOR, MWAIT, RDTSC, RDTSCP, RDPMC
- IN, OUT

- All privileged instructions
- All SVM instructions

[0070] As used herein, the term SVM instructions may refer to any instructions that a virtual machine monitor and/or virtual machine may use to interact across the boundary of the virtual machines. In various embodiments, such instructions may include, but are not limited to, VMRUN (i.e., run a virtual machine), VMLOAD/VMSAVE (i.e., load/save various virtual machine state into a processor and/or to a save area in memory), and/or VMMCALL (i.e., to execute a system call to a virtual machine monitor).

5 [0071] In various embodiments, a virtual machine monitor may prevent execution of a configurable set of instructions on a processor, such as by intercepting those operations. Such instructions may be referred to herein as “interceptable”. According to some embodiments, various or all SVM instructions may be interceptable. In some embodiments, the execution of any interceptable instruction inside of a transaction may cause the transaction to be aborted.

10 [0072] In some embodiments, ASF mechanisms may prohibit instructions within a transaction that operate differently in a virtualized environment (e.g., virtual machine) than in a native environment. By imposing such 15 restrictions, embodiments may be fully virtualizable and can be used within a virtual machine without suffering from the unpredictable or incorrect behavior that such instructions may cause when executed in a virtual environment. For example, in some embodiments, all interceptable instructions may be prohibited inside of a transaction. In some embodiments, virtual machine specific instructions and/or privileged instructions may be prohibited.

20 [0073] In some embodiments, attempting to execute disallowed or far control transfer causing instructions inside of a critical section may generate an exception (e.g., #GP exception), which may cause the transactional attempt to be aborted. In some embodiments, the far control transfer may be executed after the abort. In such embodiments, upon return from the far control transfer (or the fault handler invoked by the exception caused by the disallowed transaction), a software recovery routine may be executed.

25 [0074] As described above, in various embodiments, ASF mechanisms may abort a transactional attempt due to hardware capacity limitations, memory contention with another thread, the thread executing a disallowed instruction (e.g., far control transfer), and/or if the thread executes an explicit ABORT instruction.

30 [0075] FIG. 8 is a flow diagram illustrating a method by which, according to various embodiments, ASF mechanisms may execute a transaction in rollback mode. According to method 800, a thread may begin a transaction in rollback mode by first executing a SPECULATE instruction. As detailed in FIG. 5, executing SPECULATE may include checkpointing (i.e., saving) the instruction and stack pointer values for later use. However, in some embodiments, if the transaction is nested within another transaction, then a checkpoint may already exist and the SPECULATE instruction would not result in another checkpoint being recorded, as in FIG. 5.

35 [0076] After executing the SPECULATE instruction, the thread may determine whether an abort status has been set, as in 810, and if so, jump to and execute a recovery routine, such as 815. In some embodiments, software may determine whether an abort status code has been set by examining one or more registers (e.g., rAX register) where various abort status codes may be placed. In some embodiments, different abort status codes may indicate whether and for what reason the previous transactional attempt was aborted. In various embodiments, different abort status codes may be set, such as those indicating that the previous transactional attempt was not aborted or was aborted 40 due to contention, capacity constraints, execution of a disallowed instruction, a far control transfer, or other reasons. In some embodiments, recovery routine(s) 815 may take different recovery actions depending on the determined

abort status code. For example, in some instances, recovery routine(s) 815 may determine that the thread should not reattempt transactional execution using ASF.

[0077] According to the method 800, the thread may then begin to execute the transaction body, which may include protecting some number of protected memory locations, as in 820. In some embodiments, protecting memory locations may be performed using various declarator instructions as described above, such as LOCK MOV. In some embodiments, a protected flag corresponding to the cache block of each protected memory location may be set to indicate that the location is protected. According to the illustrated example of FIG. 8, executing the transaction body may further include writing to one or more protected memory locations, as in 825. In some embodiments, this may be performed by using various speculative store operations as described above, such as 10 LOCK MOV. In embodiments utilizing a locked line buffer, writing to a protected memory location may include storing the unmodified value to the locked line buffer and then performing the write operation to a local cache.

[0078] According to method 800, an abort condition may be encountered during execution, as indicated by the affirmative exit from 830. As described above, in different instances, an abort condition may be caused by contention with other threads, capacity limitations, far control transfers, disallowed instructions, and/or by the 15 thread executing an explicit ABORT instruction. If an abort condition is detected (affirmative exit from 830) then ASF mechanisms may be configured to discard any modifications made by the partially executed transaction to all protected memory locations.

[0079] In various embodiments, discarding modifications to all protected memory locations, as in 835, may be accomplished in different manners, dependent on the particular speculative execution mechanism being used. For 20 example, in some embodiments wherein ASF utilizes a store buffer to delay the write-back phase of store operations to protected memory as described above, discarding modifications (as in 835) may comprise discarding any store operations to protected memory locations that are waiting in the store buffer. In some embodiments in which a locked line buffer is used, discarding modifications (as in 835) may comprise writing the old values of each memory location (values of each memory location before it was modified by the transaction), stored in the locked 25 line buffer, back to the local cache.

[0080] According to method 800, aborting the transaction in rollback mode may further comprise releasing all protected memory locations, as in 840, such that they are no longer protected. For example, releasing the protected memory locations may include unsetting one or more protected flags associated with each memory location protected by the aborted transaction. In embodiments utilizing locked line buffers, releasing the protected memory 30 locations may further include removing the entries of the locked line buffer corresponding to the protected memory locations modified in the transaction.

[0081] According to the illustrated embodiment, ASF mechanisms may then determine the cause of the abort to 35 determine an appropriate abort status code. In some embodiments, ASF mechanisms may communicate the abort status code to software mechanisms by encoding the code into a status register (e.g., rAX register 216), as in 845.

[0082] In some embodiments, a status register (e.g., the rAX register) may be used to simultaneously hold a plurality of status codes. For example, some subset of the status register bits may be used to hold an abort status code while one or more other subsets may hold additional information. For example, if the abort was caused by the thread executing an ABORT instruction, a subset of the bits of status register may hold one or more values passed to the ABORT instruction by the thread when invoked. In some embodiments, a subset of the bits of the status 40 register may also hold an indication of the nesting depth of the current transaction. In further embodiments, different subsets of bits in the status register may hold various other sets of status information.

[0083] According to method 800, ASF mechanisms may then roll back execution to the beginning of the transaction, such as by rolling back the instruction and stack pointers to the checkpointed values, as in 850. In some instances, if the transaction is a top level transaction, then the checkpoint may have been recorded as part of executing the SPECULATE instruction (as in 805) and rolling back execution may result in returning control flow to the start of the transaction body, as indicated by the feedback loop from 850 to 810. In other embodiments, if the transaction is a nested transaction, then the checkpoint may have been recorded in the top level transaction and rolling back execution may result in returning control flow to the beginning of the top-level transaction's body. Thus, in some embodiments, aborting a nested transaction may comprise aborting all parent transactions in which the nested transaction is nested.

5 [0084] According to method 800, the transaction body may be attempted until it completes without an abort condition being detected. If the transactional attempt completes without abort, as indicated by the negative exit from 830, then the thread may commit the transaction, as in 855. In some embodiments, committing the transaction may comprise executing a COMMIT instruction.

10 [0085] FIG. 9 illustrates a method for committing a transaction, such as by executing a COMMIT instruction, according to some embodiments.

15 [0086] FIG. 9 is a flow diagram illustrating a method by which ASF mechanisms may abort a transaction that is executing in store-conditional mode, according to various embodiments. According to the illustrated embodiment, method 900 may begin by determining if a transaction is active, as in 905. If a transaction is active, as indicated by the affirmative exit from 905, then the transaction did not abort during execution. Thus, in some embodiments, a code indicating a successful execution may be stored in a status register. For example, in some embodiments, the rAX register may be cleared as in 910, such as by storing a 0 value in the register.

20 [0087] According to method 900 committing the transaction may further comprise setting a flag to indicate that execution is not in store-conditional mode, as in 915, and decrementing a counter that indicates the nesting level of the active transaction, as in 920. In some embodiments, if the transaction being committed is a nested transaction, as indicated by the negative exit from 925, then the transaction may be ended, as in 940. Thus, in such embodiments, speculative stores executed inside of a nested transaction may remain speculative (and susceptible to contention) even after the nested transaction has committed.

25 [0088] According to the embodiment of FIG. 9, if the transaction is a top-level transaction, as indicated by the affirmative exit from 925, then ASF mechanisms may commit all modifications to protected memory locations so that they are globally visible to other threads in the system, as in 930. In various embodiments, committing modifications to protected memory locations may involve different actions, which may depend on the speculative mechanism used. For example, in embodiments where speculative store operations to protected memory are delayed in the store buffer until the transaction commits, committing the modifications may imply performing a write-back of these values to memory (e.g., to local cache and/or to main memory). In various embodiments, committing a top level transaction may further comprise releasing all protected memory locations, as in 935, such that they are no longer protected by ASF mechanisms. For example, in embodiments wherein a protected flag is used to indicate protected cache blocks, the value of each flag associated with each protected cache block may be set to indicate that the cache block is no longer protected. Thus, ASF mechanisms need not ensure atomicity with respect to released memory locations.

30 [0089] In some embodiments, a thread attempts to commit a transaction that is not active, as indicated by the negative exit from 905, then the behavior may depend on whether or not the mode of execution is store-conditional

or rollback. If the execution mode is rollback, as indicated by the negative exit from 945, then ASF mechanisms may be configured to raise an exception, such as a #GP exception. Otherwise, if the execution mode is store-conditional (as indicated by the affirmative exit from 945), then ASF mechanisms may be configured to store the abort status code in a status register, such as rAX (as in 955) and jump to a recovery routine, as in 960. The transaction may then be reattempted as needed.

5 [0090] FIG. 10 is a flow diagram illustrating a method for executing a transaction in store-conditional mode, according to some embodiments. According to method 1000, executing a transaction in store-conditional mode may begin by executing a BEGIN instruction, as in 1005. In some embodiments, executing a BEGIN instruction may include executing method 600 of FIG. 6. In various embodiments, the BEGIN instruction may set one or more flags that may be used to determine that a transaction executing in store-conditional mode is active.

10 [0091] According to the illustrated embodiment, the transaction may then protect one or more memory locations using declarator instructions, as in 1010. In some instances, the transaction may also attempt to write to one or more of the memory locations protected in 1010, as in 1015.

15 [0092] According to method 1000, if ASF mechanisms do not detect an abort condition (e.g., contention, capacity limits exceeded, far control transfer, disallowed transaction, etc.), as in the negative exit from 1025, then the thread may commit the transaction, as in 1040. In some embodiments, committing the transaction, as in 1025, may include executing a commit instruction, such as that illustrated by FIG. 9.

20 [0093] In some embodiments, if an abort condition is detected during execution, as indicated by the affirmative exit from 1025, then the transaction may be aborted, as in 1030. In some embodiments, aborting the transaction may include undoing or otherwise discarding any modifications to protected memory locations made by store operations of the transaction. In some embodiments, an abort status code indicating the cause of the abort (e.g., conflict, capacity, disallowed transaction, etc.) may be recorded as part of aborting the transaction, as in 1030.

25 [0094] According to the illustrated embodiment, after the transaction is aborted (as in 1030), the control flow may remain in the transaction body and the transaction may continue to be executed with the exclusion of speculative stores to protected memory locations, as in 1035. In some embodiments, after the transactional attempt is aborted, declarator instructions appearing in the transaction body may be executed as regular memory references (e.g., read instructions that do not set a protected flag for the memory location) and/or accesses to protected memory locations may be ignored (e.g., treated as a no-op instruction).

30 [0095] According to some embodiment, a thread executing a transaction in store-conditional mode may be configured to verify whether or not the transaction has been aborted. For example, in some embodiments, the thread may execute a VALIDATE instruction, as in 1040, to determine whether a transaction is still active (i.e., not yet aborted). In some embodiments, the VALIDATE instruction may be executable to copy the current abort status code into a program readable status register (e.g., into the rAX register) where the value may be read by software. If the transaction has not been aborted, then the VALIDATE instruction may be executable to set the status register to a value indicating that the transaction is still active (e.g., set rAX to 0 value).

35 [0096] According to the illustrated embodiment, if the thread executes a VALIDATE instruction, it may then check the status (e.g., rAX) register to determine whether the transaction has been aborted. If the transaction is not active, as indicated by the negative exit from 1045, then the thread may execute some recovery actions, as in 1055. According to various embodiments, executing the recovery actions may include releasing any protected memory locations and/or returning program control to the start of the transaction for re-execution. If the transaction is still

active, as indicated by the affirmative exit from 1045, then the thread may continue executing the transaction and/or commit the transaction, as in 1050.

[0097] According to the illustrated embodiment, once the thread finishes executing the transaction, it may attempt to commit the transaction, as in 1050. As indicated above, committing the transaction may include executing a commit instruction, such as that illustrated in FIG. 9. As part of executing the commit instruction, ASF mechanisms may detect whether the transaction is active (as in 905) and if not, store an abort status in a software readable register (e.g., in rAX as in 955) and jump to a recovery routine (as in 960), which may be analogous to or the same as recovery actions 1055. Executing the recovery routine may include releasing any protected memory locations and/or returning program control to the start of the transaction for re-execution. In some embodiments, if the section is nested within one or more rollback mode transactions, those parent transactions may be aborted and control may be returned to the checkpoint taken by the highest level SPECULATE instruction, as described above in reference to FIG. 8.

[0098] In various embodiments, a programmer may compose applications that utilize a transactional memory programming model for ASF-enabled systems. In such embodiments, a programmer may configure a program to begin a transaction by executing a transactional start instruction (e.g., SPECULATE, BEGIN, etc), execute one or more accesses to protected memory (e.g., LOCK MOV, etc.), and then execute one or more instructions to commit the transaction (e.g., COMMIT).

[0099] In some embodiments, a program may be written to utilize an ASF-enabled transactional memory programming model in conjunction with other concurrency control mechanisms, such as non-ASF, lock-based code.

20 For example, consider a data structure such as a B-tree. Concurrent threads in a shared-memory application may perform frequent insert and delete operations to the B-tree in a transactional, lock-free manner using ASF-based transactional memory. Occasionally the B-tree may need to be rebalanced for efficiency, an operation that may include atomically operating on many memory locations of the B-tree. However, since this operation may include protecting many memory locations, attempting to complete it as a transaction may result in frequent aborts. For 25 example, conflict aborts may result from conflicts with other concurrently executing threads that perform delete and/or insert operations on the B-tree. In another example, protecting so many memory locations simultaneously may be beyond the capacity of the ASF implementation, therefore causing capacity aborts of the transaction. In such situations, rather than using transactional memory to execute the rebalancing operation, the operation may be more efficient if configured to use traditional, lock-based mutual exclusion. For example, a lock may be associated 30 with the entire B-tree and may be acquired for mutually exclusive access to the B-tree. In this example, the rebalancing operation may begin by acquiring lock for the B-tree, then executing the rebalancing operation, and then releasing the lock. However, in such instances, care must be taken to ensure that code that utilizes lock-based mutual exclusion semantics interoperates safely with code that utilizes ASF-based transactional memory semantics.

[0100] FIG. 11 is a flowchart illustrating a method for configuring ASF-based transactions to interoperate with 35 lock-based code, according to one embodiment. According to the illustrated embodiment, a thread may begin a transaction, as in 1110, such as by executing a SPECULATE or BEGIN instruction, as described above.

[0101] The set of memory locations that are protected inside of a transaction may be referred to herein as the 40 *read-set* of the transaction. In the presence of lock-based code, the read-set of a transaction may include one or more memory locations that are associated with locks. Any memory location associated with and protected by a lock may be referred to herein as *lock-protected*.

[00102] According to the illustrated embodiment, a thread may be configured to atomically read (e.g., using a declarator operation) the value of each lock associated with each lock-protected memory location in the transaction's read-set, as in 1120. Thus, by atomically reading the values of the locks, the thread adds the locks to the read-set of the transaction (i.e., ASF mechanisms monitor the locks and protect them as part of the transaction).

5 For example, in the B-tree example above, a transactional insert operation may begin by first atomically reading the value of a lock associated with the entire B-tree structure. In some embodiments, atomically reading the value of the lock may include reading the lock value using a declarator operation.

[00103] According to the illustrated embodiment, the thread may then determine if any of the locks read in step 1120 are held. If any are held, as indicated by the affirmative exit from 1130, then the thread may retry the 10 transaction. In some embodiments, the thread may continue to reread the values of each lock until all of the locks are free, as indicated by the negative exit from 1130. In some embodiments, if any of the locks are held, then the thread may abort the transaction (e.g., by executing an ABORT instruction) and then reattempt executing the 15 transaction. In other embodiments, if any of the locks are held, the thread may attempt to roll back any modifications that the transaction has made to values in protected memory locations and then to commit the transaction.

[00104] According to the illustrated embodiment, once the thread detects that no lock read in 1120 is held, as indicated by the negative exit from 1130, the thread may continue to execute the transaction body, as in 1140. In the illustrated embodiment, since the memory locations of various locks are protected using a declarator operation, any change to the value of the lock (e.g., a different thread acquiring the lock) may cause a conflict abort of the 20 transaction. Therefore, an abort condition is detected during transactional execution (e.g., a conflict abort due to a lock being acquired), then ASF-mechanisms and/or the thread may abort the transaction and execute some recovery actions (as in 1160) and reattempt the transaction (as indicated by the feedback loop from 1160 to 1120). For example, continuing the B-tree example from above, if a different thread attempts to acquire the lock associated 25 with the B-tree (e.g., as part of performing a rebalancing operation), then a cache coherence message may be sent to the executing thread, causing a conflict abort of the insert operation transaction.

[00105] According to the illustrated embodiment, if no abort condition is detected (as indicated by the negative exit from 1150), then the thread may commit the transaction, as in 1160.

[00106] FIG. 12 illustrates one embodiment of a computer system configured to implement various embodiments 30 of ASF, as described herein. Computer system 1200 may be any of various types of devices, including, but not limited to, a personal computer system, desktop computer, laptop or notebook computer, mainframe computer system, handheld computer, workstation, network computer, a consumer device, application server, storage device, a peripheral device such as a switch, modem, router, etc., or in general any type of computing device.

[00107] Computer system 1200 may include a plurality of processors 1270, each of which may include multiple 35 cores, any of which may be single or multi-threaded (e.g., simultaneous multi-processing, HyperthreadingTM, etc.).

In some embodiments, processors 1270 may include dynamic scheduling mechanisms 1272, such as those capable 40 of delaying speculative stores in load/store queues 1274 for implementing a speculative store buffer, as described herein. In some embodiments, processors 1270 may include various load, store, and/or load/store queues, such as 1274, for holding in-flight memory operations in the processor pipeline. Processors 1270 may also include registers 1276, such as rAX, rFLAGS or other special purpose registers, such as ones used for recording nesting depth, transactional mode, or status (active vs. inactive) of a transaction. In some embodiments, processors 1270 may include any number of ASF hardware transactional memory mechanisms 1278, as described herein. For example,

ASF mechanisms 1278 may include a locked line buffer and/or hardware logic for monitoring memory locations protected by an active transaction. In some embodiments, processors 1270 may also include various cache coherence mechanisms 1279, which may be used to implement different cache coherence protocols (e.g., MESI, MOESI, etc.) and may be used by ASF mechanisms to detect conflict aborts, as described herein.

5 [00108] The computer system 1200 may also include one or more persistent storage devices 1260 (e.g. optical storage, magnetic storage, hard drive, tape drive, solid state memory, etc), which may store files, for example, by using a file system, such as file system 1262. Computer system 1200 may include one or more memories 1210 (e.g., one or more of cache, SRAM, DRAM, RDRAM, EDO RAM, DDR 10 RAM, SDRAM, Rambus RAM, EEPROM, etc.) shared by the multiple processors. Various embodiments may include fewer or additional
10 components not illustrated in FIG. 12 (e.g., video cards, audio cards, additional network interfaces, peripheral devices, a network interface such as an ATM interface, an Ethernet interface, a Frame Relay interface, etc.)
[00109] Processors 1270, storage device(s) 1260, and shared memory 1210 may be coupled to system interconnect 1250. One or more of the system memories 1210 may contain program instructions 1220. Program instructions 1220 may be encoded in platform native binary, any interpreted language such as JavaTM byte-code, or in any other
15 language such as C/C++, JavaTM, etc or in any combination thereof.

20 [00110] Program instructions 1220 may include program instructions executable to implement one or more ASF-based transactional memory applications 1222. Each application 1222 may be multithreaded and may be configured to utilize ASF transactional memory mechanisms. In some embodiments, one or more of applications 1222 may be configured to operate using both ASF transactional memory instructions as well as mutual exclusion locks, as described herein. In such embodiments, shared memory 1210 may include various data structures and variables 1230, any of which may be associated with one or more mutual exclusion locks 1240.

25 [00111] In some embodiments, program instructions 1220 and/or ASF-based transactional memory applications 1222 may be provided as an article of manufacture that may include a computer-readable storage medium having stored thereon instructions that may be used to program a computer system (or other electronic devices) to perform a process according to various embodiments. A computer-readable storage medium may include any mechanism for storing information in a form (e.g., software, processing application) readable by a machine (e.g., a computer). The machine-readable storage medium may include, but is not limited to, magnetic storage medium (e.g., floppy diskette); optical storage medium (e.g., CD-ROM); magneto-optical storage medium; read only memory (ROM); random access memory (RAM); erasable programmable memory (e.g., EPROM and EEPROM); flash memory; 30 electrical, or other types of tangible medium suitable for storing program instructions. In addition, program instructions may be communicated using intangible media—optical, acoustical or other form of propagated signal (e.g., carrier waves, infrared signals, digital signals, etc.).

35 [00112] Although the embodiments above have been described in considerable detail, numerous variations and modifications will become apparent to those skilled in the art once the above disclosure is fully appreciated. It is intended that the following claims be interpreted to embrace all such variations and modifications.

Industrial Applicability

[00113] This invention may generally be applicable to shared-memory computer systems.

WHAT IS CLAIMED:

1. An apparatus, comprising:

5 a processor coupleable to a shared memory that is shared by one or more other processors, wherein the processor is configured to execute a section of code that includes a plurality of memory access operations to the shared memory as an atomic transaction relative to the one or more other processors;

10 wherein the processor is configured to determine whether the plurality of memory access operations includes at least one of a set of disallowed instructions, wherein the set of disallowed instructions includes one or more instructions that operate differently in a virtualized computing environment than in a native computing environment; and

wherein the processor is configured to abort the transaction in response to one of the plurality of memory access operations including at least one of the set of disallowed instructions.

2. The apparatus of claim 1, wherein the set of disallowed instructions includes at least one

15 instruction that executes on the processor differently depending on whether the at least one instruction is within a virtual machine or not.

3. The apparatus of claim 1, the set of disallowed instructions includes a far control transfer

20 instruction, wherein the far control transfer instruction is executable to transfer program control to a memory segment other than the one containing the instruction by modifying a value in a segment register of the processor.

4. The apparatus of claim 1, wherein the set of disallowed instructions includes one or more

instructions that may be executed only at a highest privilege level of the processor.

25 5. The apparatus of claim 1, wherein the processor is configured to abort the transaction by discarding any modifications to one or more values in the shared memory made by the execution of one or more of the plurality of memory access operations.

6. A method, comprising:

30 a first of a plurality of processors that is coupled to a shared memory executing a section of code including a plurality of memory access operations to shared memory as an atomic transaction with respect to the plurality of processors;

35 the first processor determining whether the plurality of memory access operations includes at least one of a set of disallowed instructions, wherein the set of disallowed instructions includes one or more instructions that operate differently depending whether or not the instruction is executed within a virtual machine; and

in response to the plurality of memory access operations including at least one of the set of disallowed instructions, the first processor aborting the transaction.

7. The method of claim 6, the set of disallowed instructions comprises a far control transfer

40 instruction, wherein the far control transfer instruction is executable to transfer program control to a memory segment other than the one containing the instruction by modifying a value in a segment register of the processor.

8. The method of claim 6, wherein the set of disallowed instructions comprises instructions that may be executed only at a highest privilege level of the first processor.

5 9. The method of claim 6, wherein said aborting comprises discarding any modifications to one or more values in the shared memory made by the execution of one or more of the plurality of memory access operations.

10 10. A system, comprising:

a shared memory; and

a plurality of processors coupled to the shared memory;

wherein a first of the plurality of processors is configured to execute a section of code that includes a plurality of memory access operations to the shared memory, wherein the first processor is configured to execute the plurality of memory access operations as an atomic transaction relative to remaining ones of the plurality of processors;

15 wherein the first processor is configured to determine whether the plurality of memory access operations includes at least one instruction in a set of disallowed instructions, wherein execution of at least one instruction in the set of disallowed instructions is based on a privilege level of the first processor; and

20 wherein the first processor is configured to abort the transaction in response to one of the plurality of memory access operations including at least one of the set of disallowed instructions.

11. The system of claim 10, wherein the set of disallowed instructions includes an instruction that executes on the processor differently depending on whether the instruction is executed within a virtual machine.

25 12. The system of claim 10, wherein the set of disallowed instructions includes privileged instructions.

30 13. The system of claim 10, wherein the set of disallowed instructions includes a far control transfer instruction.

14. The system of claim 10, wherein the first processor is configured to abort the transaction by discarding any modifications to one or more values in the shared memory made by the execution of one or more of the plurality of memory access operations.

35 15. The system of claim 10, wherein the set of disallowed instructions includes one or more instructions associated with a hypervisor.

1 / 12

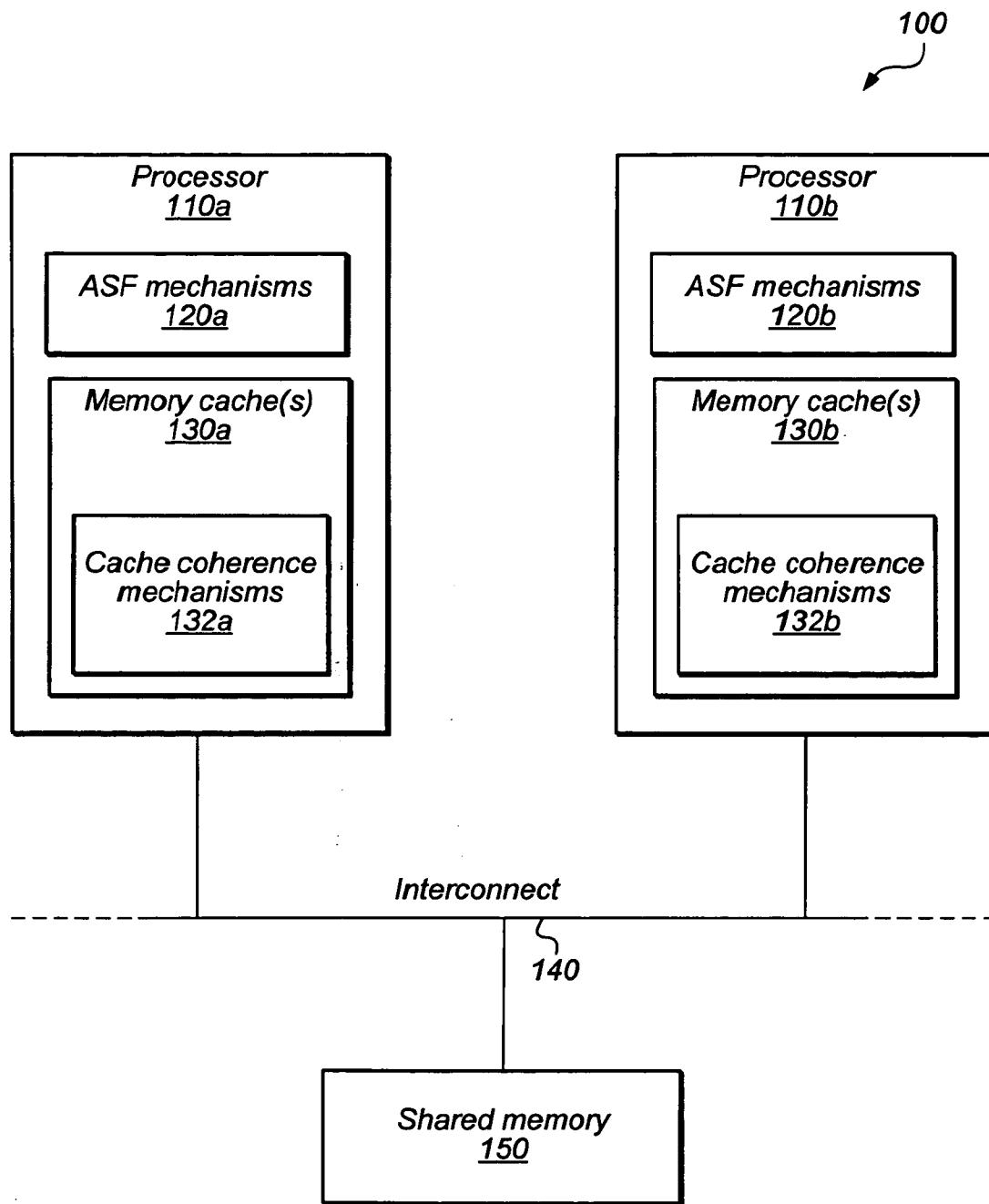


FIG. 1

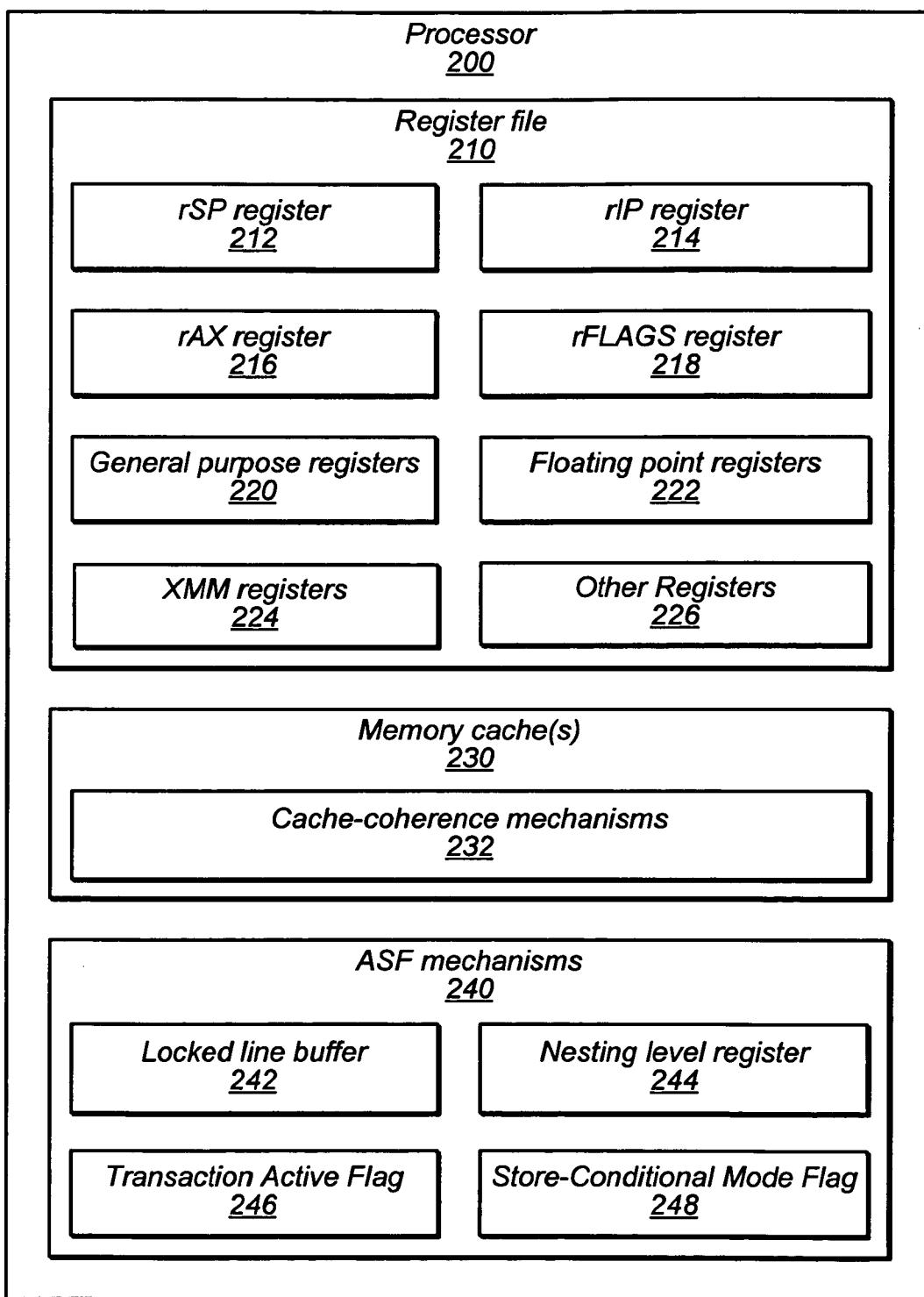


FIG. 2

3 / 12

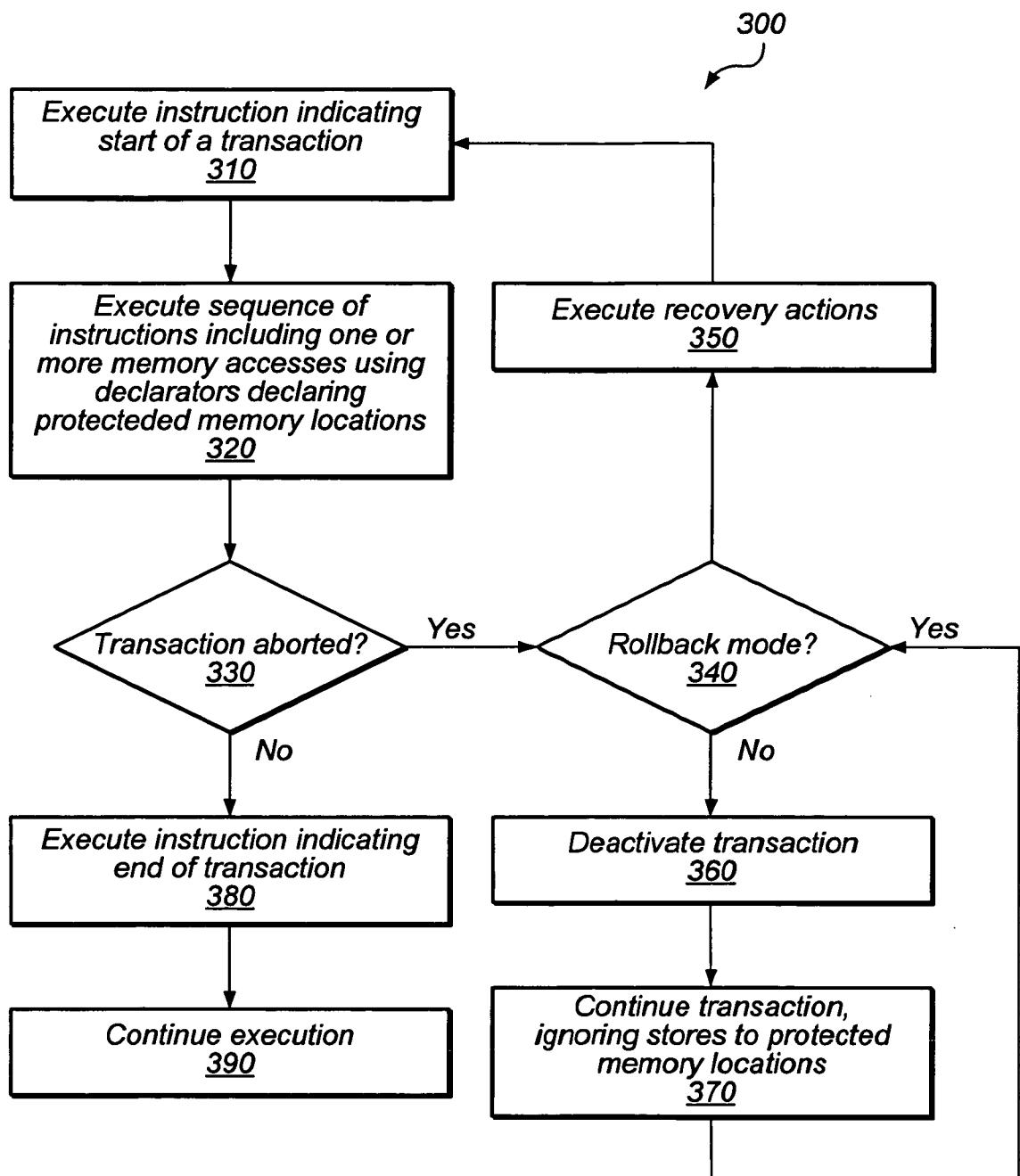


FIG. 3

4 / 12

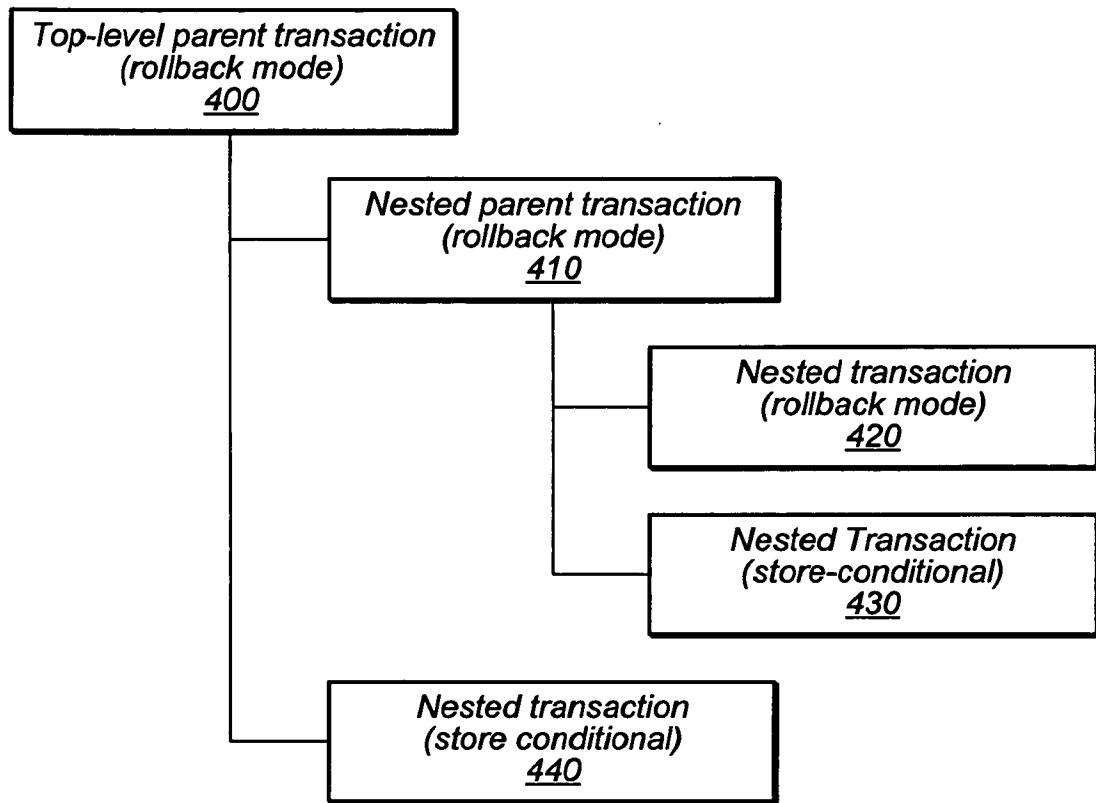


FIG. 4

5 / 12

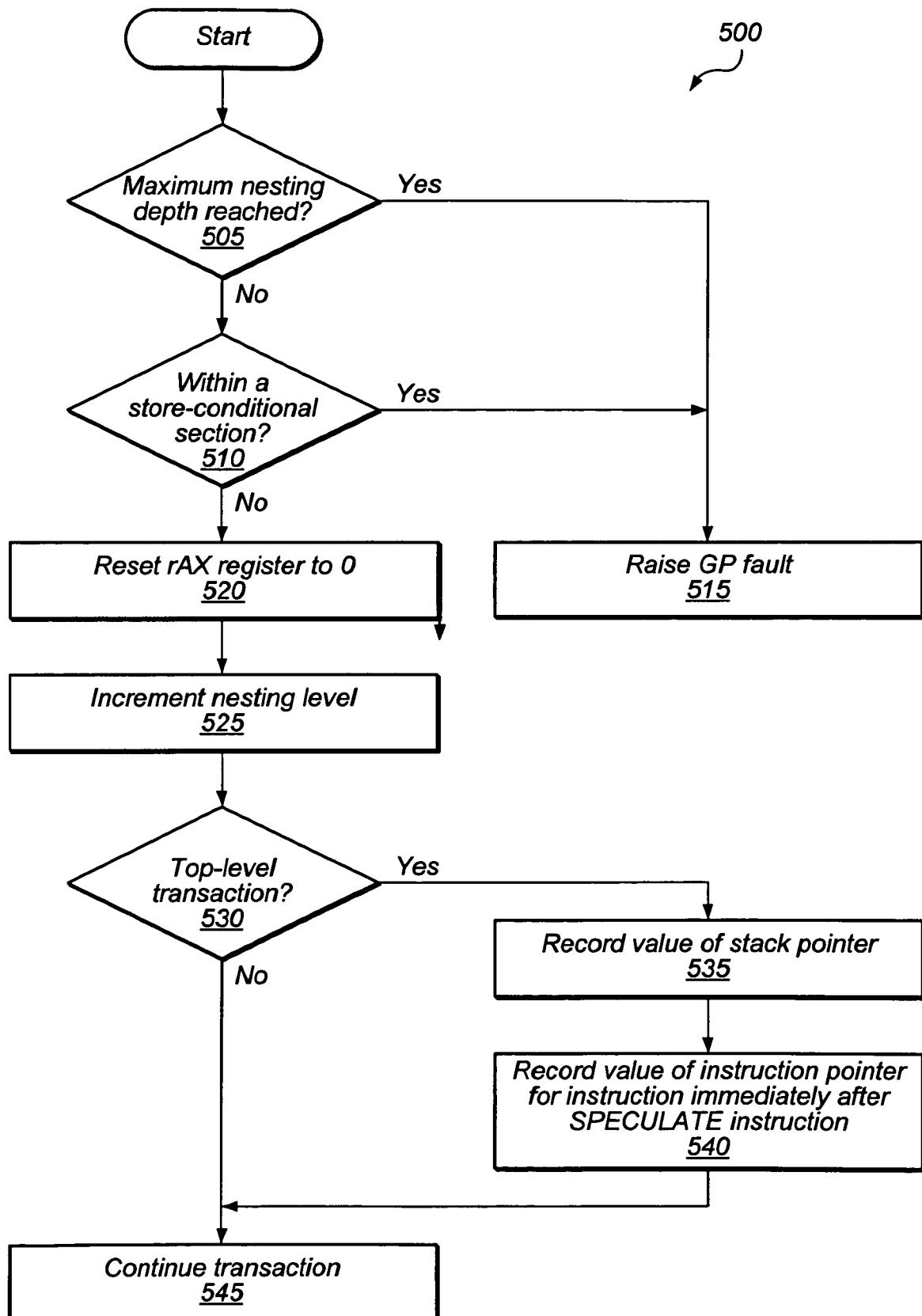


FIG. 5

6 / 12

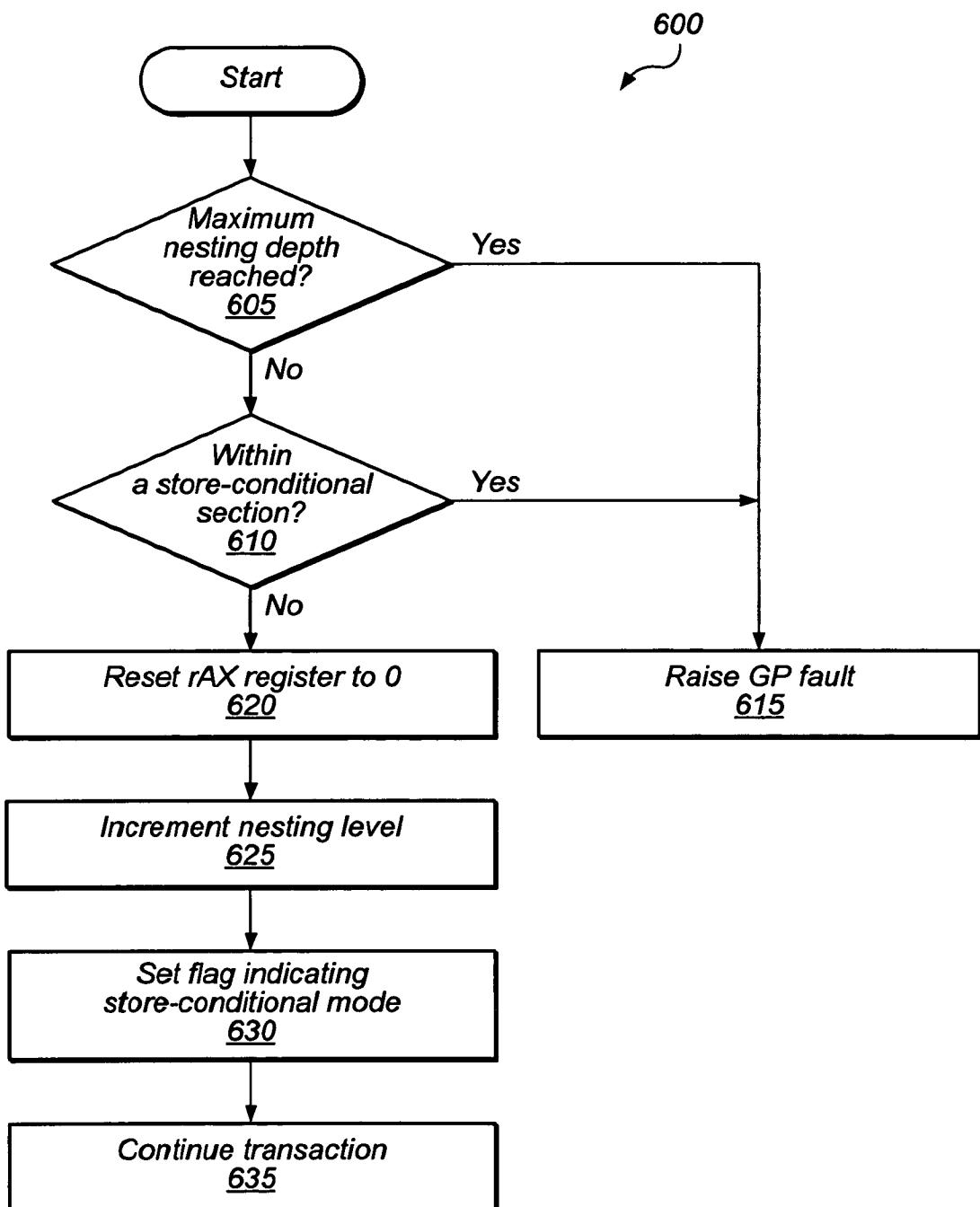


FIG. 6

7 / 12

CPU A mode	CPU A operation	CPU B cache-line state	
		Protected Shared	Protected owned*
Any	<i>Read operation</i>	OK	<i>B aborts</i>
Any	<i>Write operation</i>	<i>B aborts</i>	<i>B aborts</i>
Any	<i>Prefetch operation</i>	OK	<i>B aborts</i>
Any	<i>PREFETCHW</i>	<i>B aborts</i>	<i>B aborts</i>
Critical section	<i>LOCK MOVx (load)</i>	OK	<i>B aborts</i>
Critical section	<i>LOCK PREFETCH</i>	OK	<i>B aborts</i>
Critical section	<i>LOCK PREFETCHW</i>	<i>B aborts</i>	<i>B aborts</i>

“owned*” --Modified or owned

FIG. 7

8 / 12

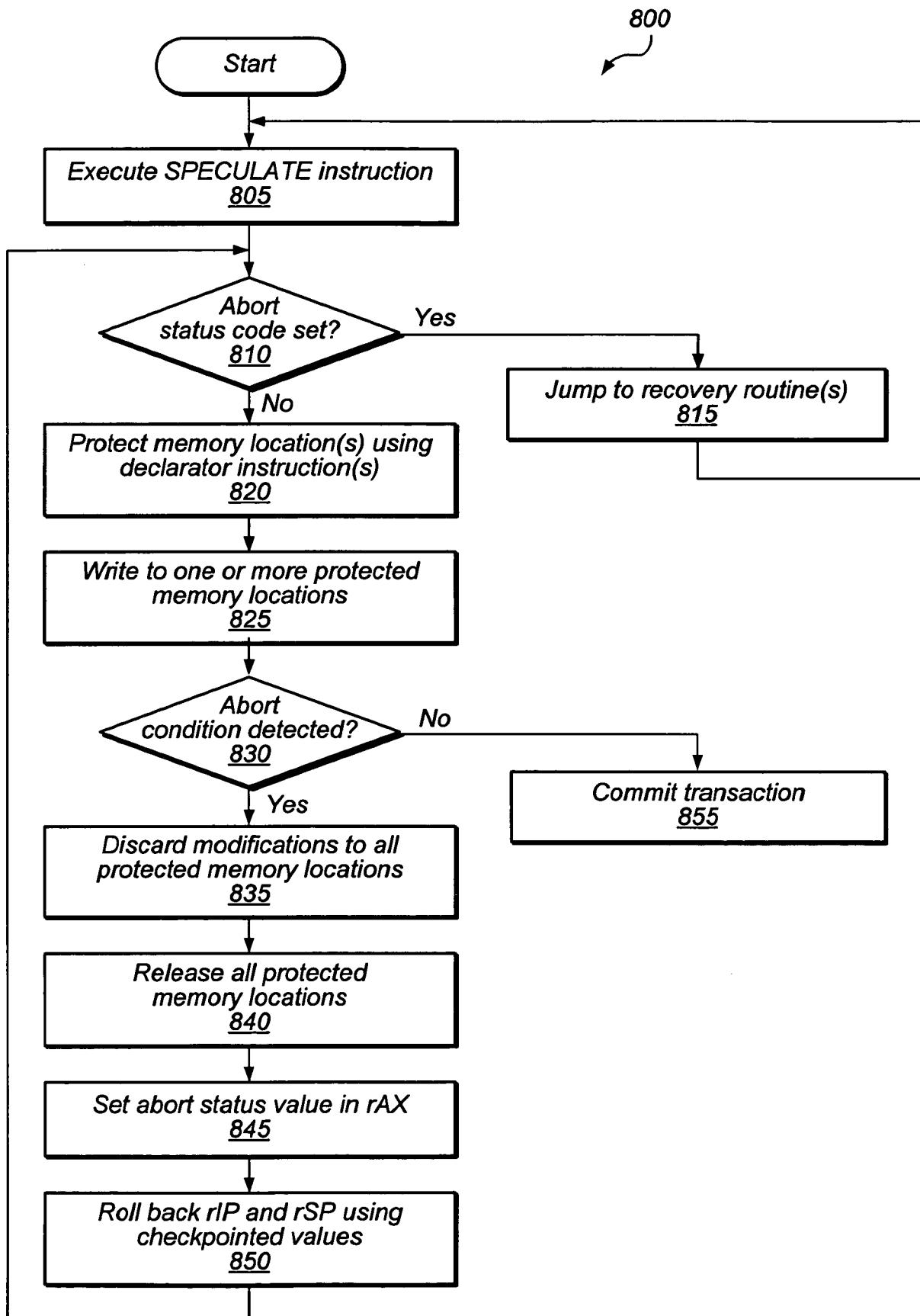


FIG. 8

9 / 12

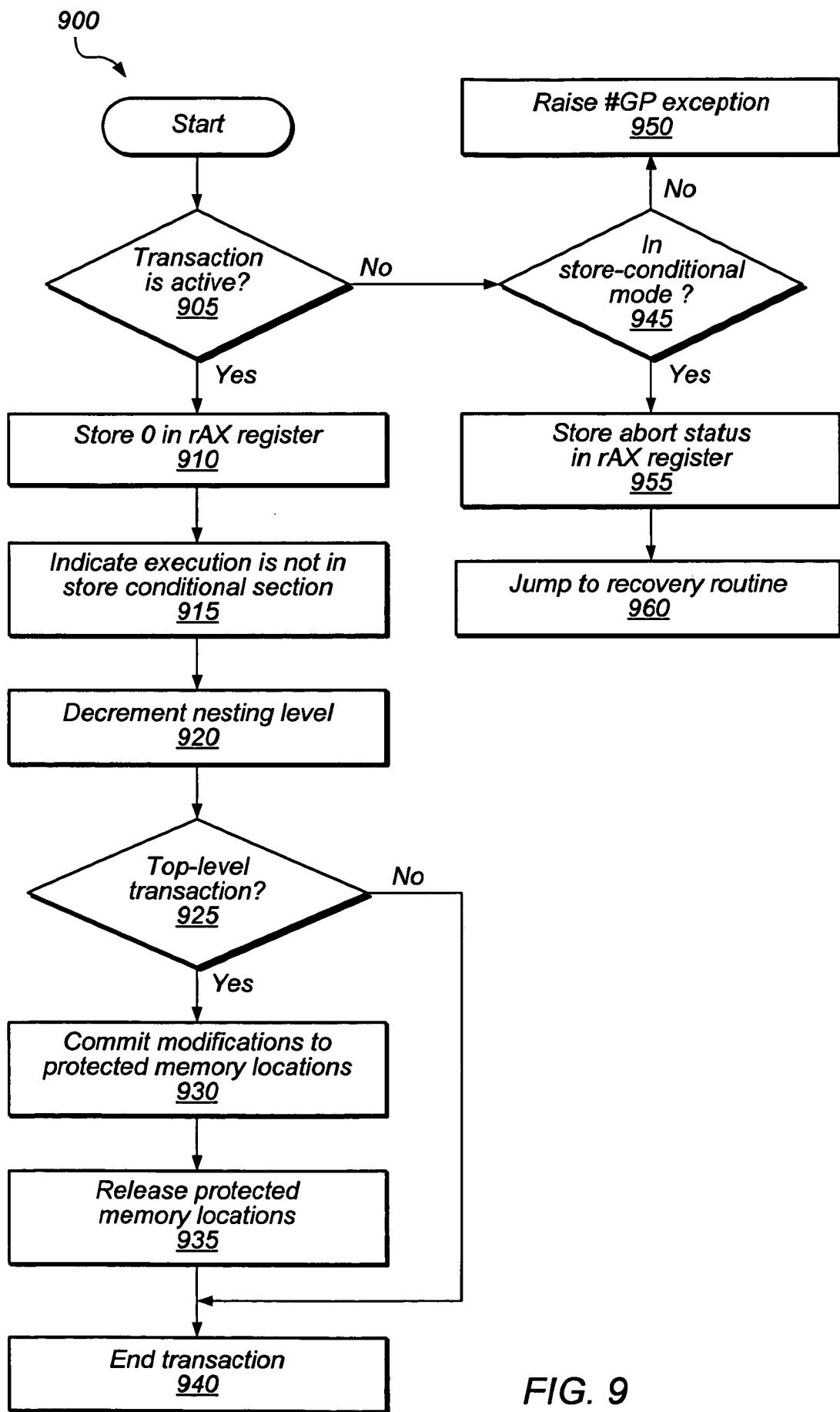


FIG. 9

10 / 12

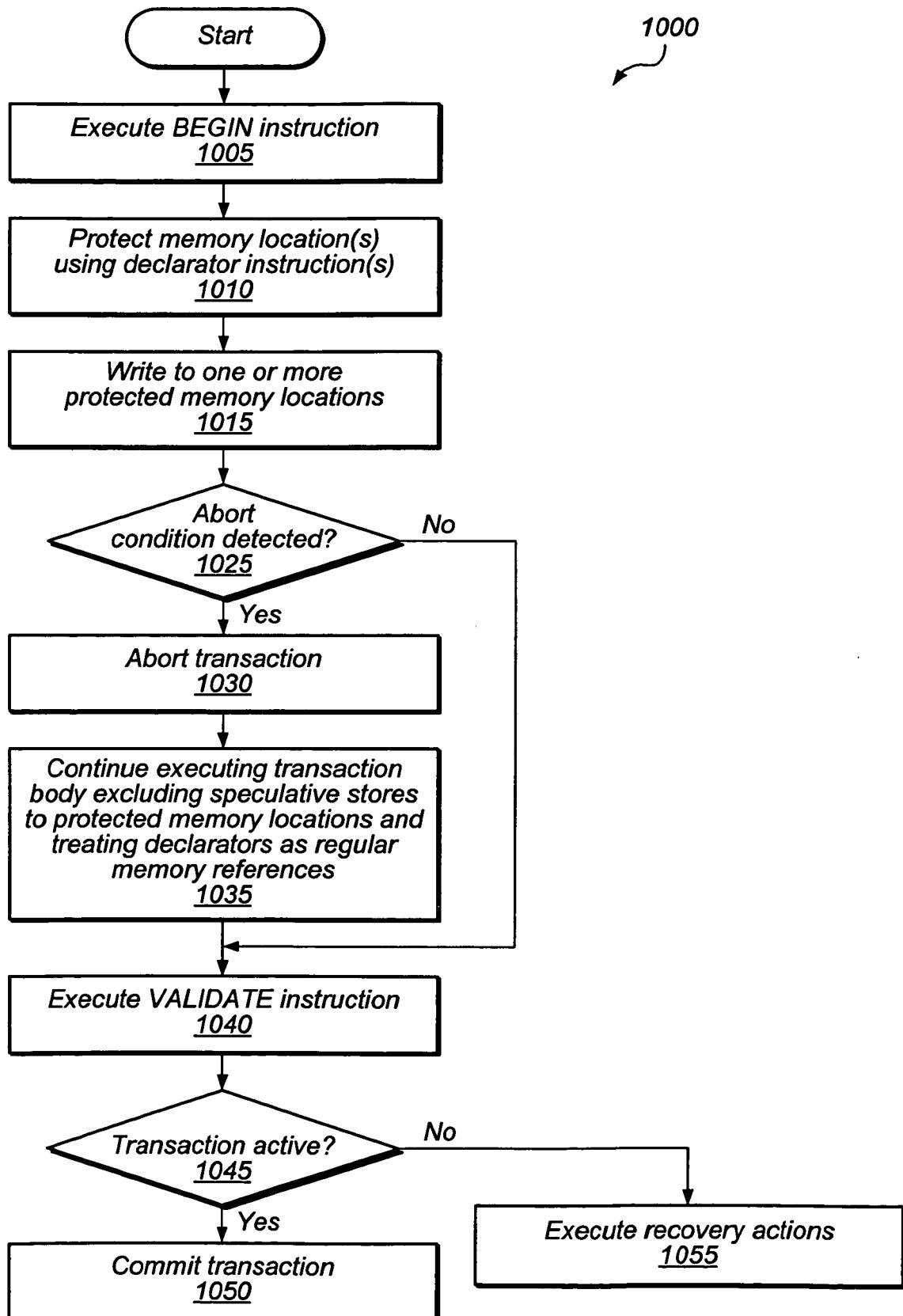


FIG. 10

11 / 12

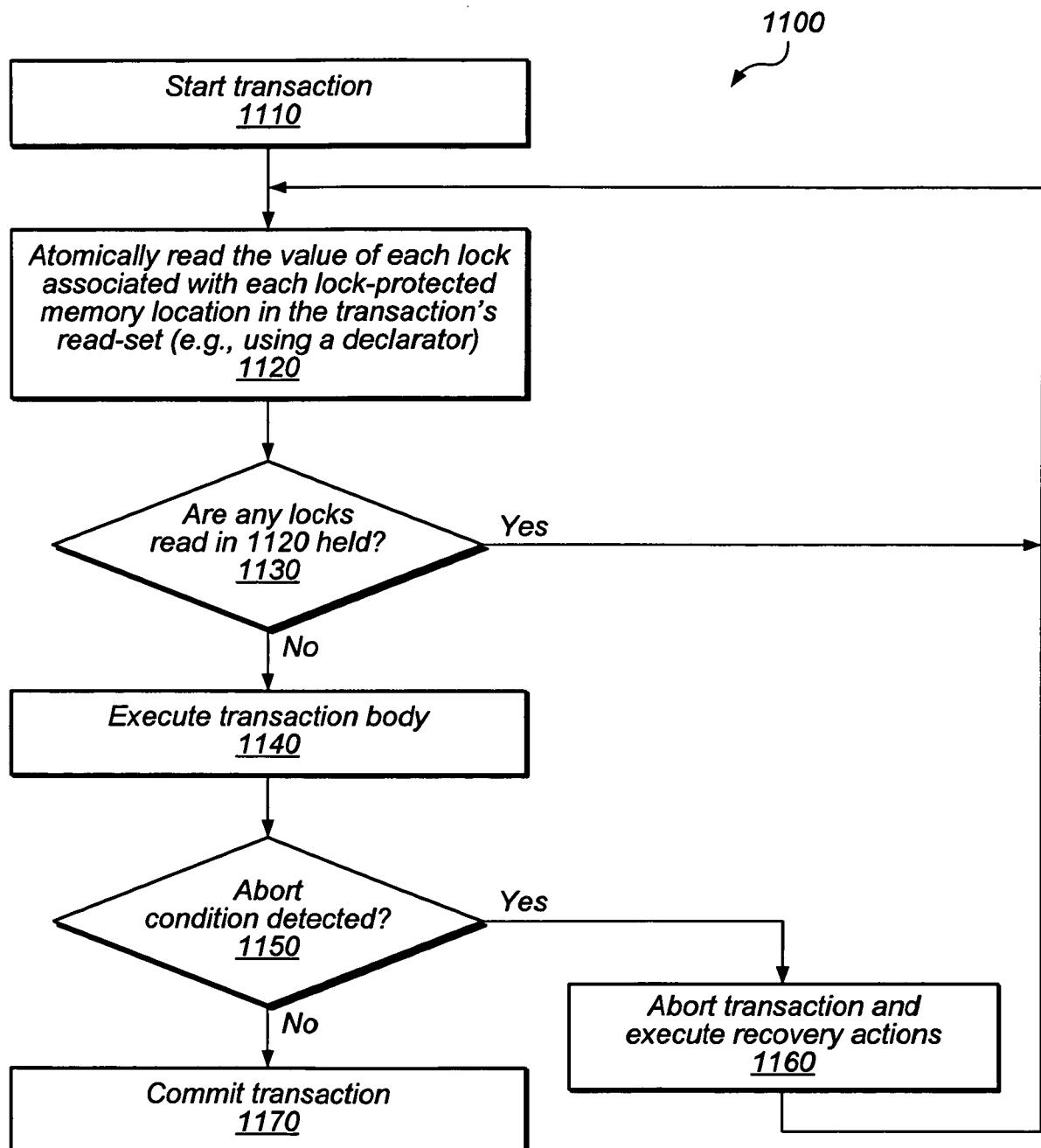


FIG. 11

12 / 12

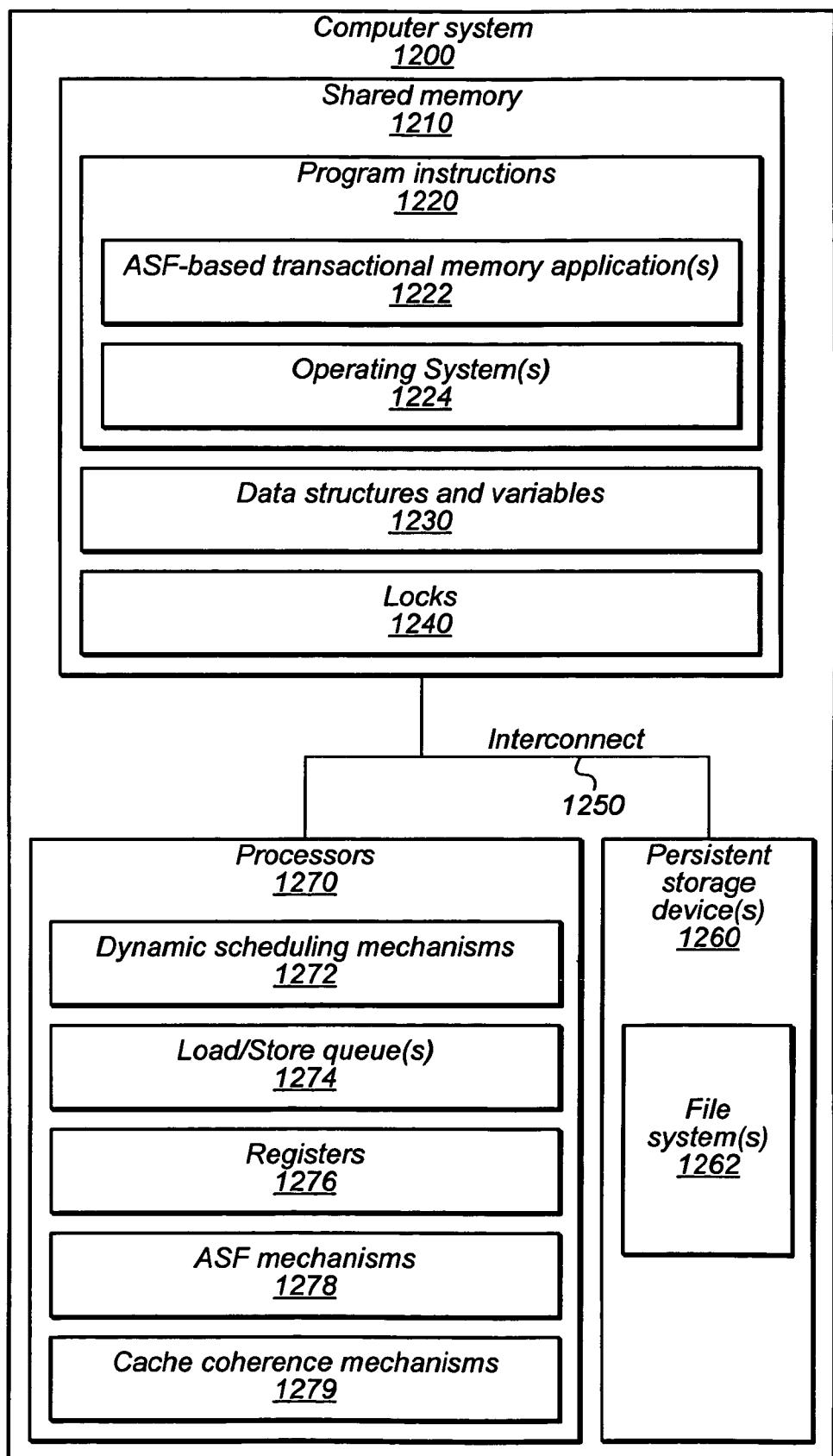


FIG. 12

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2009/004349

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F9/455 G06F9/46

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, INSPEC, IBM-TDB, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2007/239942 A1 (RAJWAR RAVI [US]; AKKARY HAITHAM H [US]; LAI KONRAD [US]) 11 October 2007 (2007-10-11) paragraphs [0011], [0029], [0030], [0032]; figure 4 -----	1-15
A	RAJWAR R; HERLIHY M; KONRAD LAI: "Virtualizing Transactional Memory" COMPUTER ARCHITECTURE, 2005. ISCA '05. PROCEEDINGS. 32ND INTERNATIONAL SYMPOSIUM ON MADISON, WI, USA 04-08 JUNE 2005, 20050604 - 20050608 PISCATAWAY, NJ, USA, IEEE, 4 June 2005 (2005-06-04), pages 494-505, XP010807930 the whole document ----- -/-	1-15

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

29 September 2009

Date of mailing of the international search report

09/10/2009

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Thibaudeau, Jean

INTERNATIONAL SEARCH REPORT

International application No PCT/US2009/004349

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2008/005504 A1 (BARNES JESSE [US]; RAJWAR RAVI [US]) 3 January 2008 (2008-01-03) paragraphs [0018] - [0020], [0037], [0038], [0042] -----	1-15
A	GARFINKEL T; ROSENBLUM M: "Virtual Machine Monitors: Current Technology and Future Trends" COMPUTER, IEEE SERVICE CENTER, LOS ALAMITOS, CA, US, vol. 38, no. 5, 1 May 2005 (2005-05-01), pages 39-47, XP011132222 ISSN: 0018-9162 page 41, right-hand column, line 21 - line 32 -----	1-4,6-8, 10-13

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2009/004349

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
US 2007239942	A1 11-10-2007	CN 101059783	A	24-10-2007
US 2008005504	A1 03-01-2008	CN 101097544	A	02-01-2008
		DE 112007001171	T5	30-04-2009
		KR 20090025295	A	10-03-2009
		WO 2008005687	A2	10-01-2008