

Internal distribution code:

- (A) [-] Publication in OJ
(B) [-] To Chairmen and Members
(C) [-] To Chairmen
(D) [X] No distribution

**Datasheet for the decision
of 31 July 2018**

Case Number: T 1335/16 - 3.5.06
Application Number: 03774703.7
Publication Number: 1550032
IPC: G06F9/38, G06F12/06, G06F12/08
Language of the proceedings: EN

Title of invention:

METHOD AND APPARATUS FOR THREAD-BASED MEMORY ACCESS IN A
MULTITHREADED PROCESSOR

Applicant:

Qualcomm Incorporated

Headword:

Thread-based memory access/QUALCOMM

Relevant legal provisions:

EPC 1973 Art. 56, 84
EPC Art. 123(2)

Keyword:

Inventive step over documents relied upon in the decision
under appeal - after amendment (yes)
Remittal to the department of first instance

Decisions cited:

Catchword:



Beschwerdekammern
Boards of Appeal
Chambres de recours

Boards of Appeal of the
European Patent Office
Richard-Reitzner-Allee 8
85540 Haar
GERMANY
Tel. +49 (0)89 2399-0
Fax +49 (0)89 2399-4465

Case Number: T 1335/16 - 3.5.06

D E C I S I O N
of Technical Board of Appeal 3.5.06
of 31 July 2018

Appellant: Qualcomm Incorporated
(Applicant) 5775 Morehouse Drive
San Diego, CA 92121 (US)

Representative: Strehl Schübel-Hopf & Partner
Maximilianstrasse 54
80538 München (DE)

Decision under appeal: Decision of the Examining Division of the
European Patent Office posted on 15 December
2015 refusing European patent application No.
03774703.7 pursuant to Article 97(2) EPC

Composition of the Board:

Chairman W. Sekretaruk
Members: M. Müller
S. Krischer

Summary of Facts and Submissions

I. The appeal is against the decision of the examining division, despatched with reasons dated 15 December 2015, to refuse European patent application No. 03 774 703.7 for lack of inventive step over document

D1: US 2002/103990 A1

in view of common knowledge, exemplified by a document which the board refers to as "D3":

D3: Kogge P M, "The Architecture of Pipelined Computers", McGraw-Hill, 1981, pages 39-44.

A further document was cited in the decision but not relied upon in the reasons, namely:

D2: WO 00/68821 A.

II. Notice of appeal was filed on 25 February 2016, the appeal fee being paid on the same day. A statement of grounds of appeal was received on 25 April 2016. The appellant requested that the decision be set aside and that a patent be granted on the basis of claims according to a main request or auxiliary requests 2 to 6, all as filed with the grounds of appeal, the other application documents being:

description, pages

4-11 as published,

1-3 as received on 22 June 2011, and
drawings, sheets

1-3 as published.

III. In an annex to a summons to oral proceedings, the board informed the appellant of its preliminary opinion that the claimed invention lacked inventive step over D1, Article 56 EPC 1973. A number of issues relating to Article 84 EPC 1973 and to claim construction were also addressed.

IV. In response to the summons, by letter of 29 June 2018 the appellant filed new sets of claims according to auxiliary requests 7 and 8.

V. Oral proceedings were held on 31 July 2018, during which the appellant filed amended claim 1, based on the former main request, and requested that the decision under appeal be set aside and that a patent be granted on the basis of claim 1 of the new request, with further claims and a description to be adapted.

VI. Claim 1 reads as follows:

"A method for accessing a memory (600) by a multithreaded processor (102), the method comprising:

determining a thread identifier associated with a corresponding processor thread of the multithreaded processor (102) the thread identifier corresponding to the output of a thread counter in the multi-threaded processor, the thread counter output identifying a particular thread being executed, whereby the processor is configured such that multiple threads are processed in a round robin order; and

selecting a particular portion of the memory (600) to be accessed by the corresponding processor thread, characterized in that:

selecting the particular portion of the memory (600) comprises utilizing, by selection circuitry of the multithreaded processor, at least a portion of the

thread identifier to select the particular portion of the memory (600) to be accessed by the corresponding processor thread,

wherein utilizing at least a portion of the thread identifier comprises:

utilizing a first portion (604) of the thread identifier and the selection circuitry to select one of a plurality of multiple-bank memory elements (M0-M3) within the memory (600), and

utilizing a second portion (606) of the thread identifier by applying it as a select signal to the selection circuitry to select one of a plurality of memory banks (B0, B1) within the selected one of the multiple-bank memory elements (M0-M3)."

VII. At the end of the oral proceedings, the chairman announced the board's decision.

Reasons for the Decision

The appellant's right to be heard

1. In the decision under appeal (point 10.3 of the reasons), the examining division referred to an "attached extract from a standard textbook of 1981", but no such extract was attached. The appellant challenged the decision on this basis (see grounds of appeal, page 8, sentence 2).

1.1 The examining division had used the same phrase in its summons to oral proceedings (see point 4) and in that case had not failed to attach a document, namely D3. Introduction of this document with the summons is also

mentioned in the decision under appeal (see point 5 of the facts and submissions).

- 1.2 Moreover, "the extract" according to D3 is discussed in the decision (see point 12 of the reasons), and no further prior-art document is relied upon in the decision.
- 1.3 The board thus considers that it should have been evident at the time that the sentence in point 10.2 of the reasons was meant to refer to D3 and that the suggestion that it be "attached" to the decision was caused by a cut-and-paste error from the summons. The appellant's right to be heard under Article 113(1) EPC 1973 is not affected by this deficiency.

The invention

2. The application relates to techniques for memory access by a multithreaded processor (see abstract).
- 2.1 More specifically, the application considers processors with a fixed number of threads which are scheduled in a predetermined, preferably round-robin, order. Such processors use a thread identifier, typically held in a suitable register. If there are eight threads, the thread identifier will need three bits. For the round-robin case, it is disclosed that the thread identifiers are generated by a suitable "thread counter" (see figures 3 and 6; page 3, paragraph 2; page 6, last paragraph; page 9, paragraphs 1 and 2; and page 10, paragraph 4).
- 2.2 In this context, the application is concerned with improving memory access so as decrease memory access times and, in particular, to reduce the stalling of

threads waiting for the required memory to become available, with a view to hardware simplicity (see page 2, lines 12-16; and page 6, lines 8-9).

2.3 As a solution it is proposed that the memory is arranged in such a way that each thread accesses only a particular, dedicated memory bank. For instance, if there are eight threads, four "multiple-bank memory elements" with two banks each may be provided (see for instance page 6, paragraph 2; page 9, paragraph 1; and figure 6). Alternatively, there may be two memory elements with four banks each or, in principle, any other such combination (see the paragraph bridging pages 9 and 10).

2.4 It is further proposed that the bits of the thread identifier are used as select signals on circuit level to access the memory bank associated with any particular thread. It is disclosed that "first" and "second" "portions" of the thread identifier are used to select, respectively, a memory element and a bank within it. The size of the two portions depends on how many memory elements and banks there are. For instance, if each memory element contains two banks, the "second portion" will be a single bit; if each memory element comprises four banks, the second portion will be two bits; and the general case is also discussed. Preferably, the first and second portions are, respectively, the most and least significant bits of the thread identifier (see again page 9, paragraph 2, to page 10, paragraph 1; page 10, last paragraph; and figure 6).

Article 123(2) EPC

3. Claim 1 is generally based on original claim 1 in combination with figure 6. That the thread identifier is generated by a thread counter used by the processor for round-robin scheduling is disclosed on page 10 as originally filed, paragraph 4. That the corresponding selection circuitry is part of the processor is disclosed on page 9, lines 14-16, and page 10, lines 8-12; that the thread identifier is used as a memory select signal is disclosed on page 9, lines 14-19; and that the specific case illustrated in figure 6 can be generalised to different numbers and sizes of memory elements and banks is discussed in the paragraph bridging pages 9 and 10. That the two portions of the thread identifier need not be limited to most and least significant bits is disclosed in claims 1 and 2 as originally filed. The board is therefore satisfied that claim 1 complies with Article 123(2) EPC.

Clarity, Article 84 EPC

4. The invention rests on the general idea of re-using a thread identifier, which the multithreaded processor happens to use anyway, to address a dedicated memory bank per thread. In its summons (point 4.2), the board referred to this as the "dual use" of the thread identifier.
 - 4.1 Some of the board's clarity objections related to the fact that the claims then on file did not make this dual use clear. The previous claim language specified only that the thread identifier was "associated with a corresponding processor thread" and "utilized" to "select" a memory bank (see points 4.2 and 4.5 of the

summons). It is now clearly specified that the thread identifier has a specific - arguably primary - use for scheduling in the multithreaded processor and that the thread identifier representation is used directly as memory select signals.

- 4.2 The board's concern that it was not specified where the method was carried out (see point 4.1 of the summons) is overcome by the express statement that it is carried out by the multithreaded processor.
- 4.3 Further clarity objections related to previous auxiliary requests and are immaterial for present claim 1 (see points 4.3 and 4.4 of the summons).
- 4.4 The board has no other clarity objections and thus concludes that claim 1 complies with Article 84 EPC 1973.

The prior art

- 5. D1 discloses what is called a "precession machine" (see the title and paragraphs 9 and 10), a multithreaded processor with a fixed number of threads such as 4, 8 or 16 (see figures 2, 3a, 3b and 5).
- 5.1 Thread contexts are stored in dedicated register sets (see paragraph 10). Processor time is allocated to threads in turn (see in particular figures 2 and 5), following the sequence in which the corresponding thread identifiers are listed in a "cycle allocation table" (see paragraph 14). By listing a particular thread identifier more than once, the corresponding thread may receive a larger share of processor time (see paragraphs 16 and 49).

- 5.2 It is disclosed that each thread has "its own program counter, register set and stack space" but threads share the "code space, data space, and various operating resources" with "peer threads" (see paragraphs 27 and 49 and figure 4).
- 5.3 In order to increase the throughput of the "precession machine", the document discloses using banked memory and "alternating memory accesses between different banks" (see paragraph 46, in particular lines 3 to 5). In general, this is said to have the effect that consecutive memory accesses are "likely to go to a different bank" (see paragraph 46, lines 19 and 20).
- 5.4 It is also disclosed that "each of the memory banks" may be "associated with one or more of the active threads". In these embodiments, the skilled person would understand that consecutive memory accesses *from different threads* would be not only likely but in fact guaranteed to go to a different memory bank (in contrast to the appellant's position, see grounds of appeal, page 5, paragraph 2). It is specifically discussed in D1 that there may be two or four memory banks for four threads (see paragraph 47, in particular the last three sentences).
6. D3 is an extract from a textbook on computer architecture. It discloses that the memory access rate can be increased by what is called "memory interleaving". For that purpose, memory addresses are distributed among several memory modules so that "sequential memory addresses" tend to end up in different memory modules. In this way, access to different modules can overlap in time and the memory access rate is increased (see page 40, paragraph 3, and figures 2-19 and 2-20; see in particular the timing

diagram in figure 2-20). It is also disclosed that several memory modules (e.g. $N=2^k$) can be accessed in parallel at the same addresses (see address lines a1-am), the read words then being serialised in a selected order (see page 40, paragraph 4). And it is disclosed that both techniques can be and have been combined by providing several "basic memory modules", each of them being "interleaved" (see page 43, last paragraph).

Inventive step vis-à-vis D1 and D3

7. D1 discloses that each individual thread may access a different, dedicated memory bank (see paragraph 47, in particular lines 2 to 5 and the last three sentences).

7.1 D1 does not disclose

- (i) that the memory banks may be spread over several, individually "selectable" memory "elements", or
- (ii) that the memory bank associated with a thread is "selected" on circuit level, and how.

More specifically, D1 does not disclose

- (ii') that the thread identifier is used by the multithreaded processor for round-robin scheduling with the help of a thread counter and re-used directly as a memory select signal, or
- (ii'') that, for this purpose, the thread identifier is split into two "portions".

7.2 As regards feature (i) the board had assumed - without reference to written prior art - that it had been commonly known that individual memory elements might

contain several memory banks (point 10 of the summons). This assumption was not challenged by the appellant. The board wishes to add, however, that the relative size of available memory chips and required memory banks directly implies how many memory banks fit into one memory element or how many memory elements are required to provide a single memory bank. The board thus considers that feature (i) alone therefore cannot support an inventive step.

- 7.3 In the summons, the board had interpreted the term "thread identifier" broadly as any number suitable for identifying a thread - or, in the language of claim 1, "associated with a corresponding processor thread" - and thus identified the numbers 1 to 4 used in paragraph 47 of D1 with the claimed thread identifier. The board had further argued why it would have been obvious in general to express the mapping from threads to memory banks in terms of "portions" of the thread identifier in binary form (this is, essentially, feature (ii')); see points 4.2, 8.2, 11.2 and 11.3 of the summons).
- 7.4 Due to the amendments made by the appellant, claim 1 is now explicit about the fact that the thread identifier is not just any number but a number generated and used by the multithreaded processor for round-robin scheduling. In other words, the above-mentioned "dual use" is now specifically claimed. Accordingly, features (ii') and (ii'') are now intimately connected, and so the inventive step of feature (ii') cannot be considered separately from that of feature (ii'').
- 7.5 D1 does not disclose the claimed processor circuitry and therefore cannot on its own suggest its re-use for memory bank selection on circuit level.

- 7.6 The examining division found that it had been "notorious knowledge" to "organis[e] memories in banks of multiple modules by grouping together memory modules in an effort to increase parallelism, while at the same time reducing the complexity of the memory controller" (see the decision, reasons 10.3), and that this view was supported by D3.
- 7.7 It may be left open whether this was actually "notorious" knowledge and whether D3 disclosed this general teaching because, either way, D3 does not disclose or suggest feature (ii') or (ii'').
- 7.8 One embodiment of D3 discloses interleaved memory, in which "a stream of reads" accesses different memory elements in subsequent processor cycles so that at least some of the accesses will execute in parallel (see figure 2-20), and another discloses parallel access to several memory modules at the same address and the subsequent sequential output of the retrieved values (figure 2-19). Neither of the embodiments discloses or suggests thread-selective access to just one memory bank in a memory element or the use of a thread identifier as memory select signals.
- 7.9 The board therefore concludes that the subject-matter of claim 1 is not rendered obvious by D1, on its own or in view of D3, Article 56 EPC 1973.

Remittal for further prosecution

8. The amendments to claim 1 were based on features taken from the description.

- 8.1 Even though the board takes the view that they relate to a hardware embodiment which should have been searched, it cannot be sure that it actually was searched.
- 8.2 Moreover, D2, marked in the supplementary European search report as "X" ("particularly relevant if taken alone"), was not assessed during examination, and the international search report mentions two documents marked as "Y" ("particularly relevant if combined").
- 8.3 For both reasons, the board is not in a position to order the grant of a patent on the basis of its positive finding vis-à-vis D1 and D3.
- 8.4 The board therefore exercises its discretion under Article 111(1) EPC 1973 and remits the case to the examining division for further prosecution so that the further prior art on file can be considered and, if need be, an additional search can be carried out.
- 8.5 Moreover, the appellant will have to be given an opportunity to adapt the description (see for instance page 10, lines 17-21) and to complete the set of claims. In particular, the appellant may want to provide an independent system claim corresponding to present claim 1.

Order

For these reasons it is decided that:

1. The decision under appeal is set aside.
2. The case is remitted to the examining division for further prosecution.

The Registrar:

The Chairman:



B. Atienza Vivancos

W. Sekretaruk

Decision electronically authenticated