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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 September 2014 (18.09.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/140697 A1

(51) International Patent Classification:
G06F 9/50 (2006.01)

(21) International Application Number:
PCT/IB2013/052098

(22) International Filing Date:
15 March 2013 (15.03.2013)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **ASELSAN ELEKTRONIK SANAYI VE TICARET ANONIM SIRKETI** [TR/TR]; Mehmet Akif Ersoy Mahallesi, 296. Cadde No:16, Yenimahalle, 06370 Ankara (TR).

(72) Inventors: **AKGUN, Toygar**; Aselsan Elektronik Sanayi ve Ticaret Anonim Sirketi, Akyurt Tesisleri (MGEO Grup Baskanligi) P.K.30, Etlik, 06011 Ankara (TR). **ÖZDEMİR, Serhat**; Aselsan Elektronik Sanayi ve Ticaret Anonim Sirketi, Akyurt Tesisleri (MGEO Grup Baskanligi) P.K.30, Etlik, 06011 Ankara (TR). **ASLANGÜL, Süleyman Alpay**; Aselsan Elektronik Sanayi ve Ticaret Anonim Sirketi, Akyurt Tesisleri (MGEO Grup Baskanligi) P.K.30, Etlik, 06011 Ankara (TR). **KARAGÖZ, Mehmet Fatih**; Aselsan Elektronik Sanayi ve Ticaret Anonim Sirketi, Akyurt Tesisleri (MGEO Grup Baskanligi) P.K.30, Etlik, 06011 Ankara (TR).

(74) Agent: **ANKARA PATENT BUREAU LIMITED**; Bestekar Sokak No: 10, Kavaklidere, 06680 Ankara (TR).

(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, BN, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

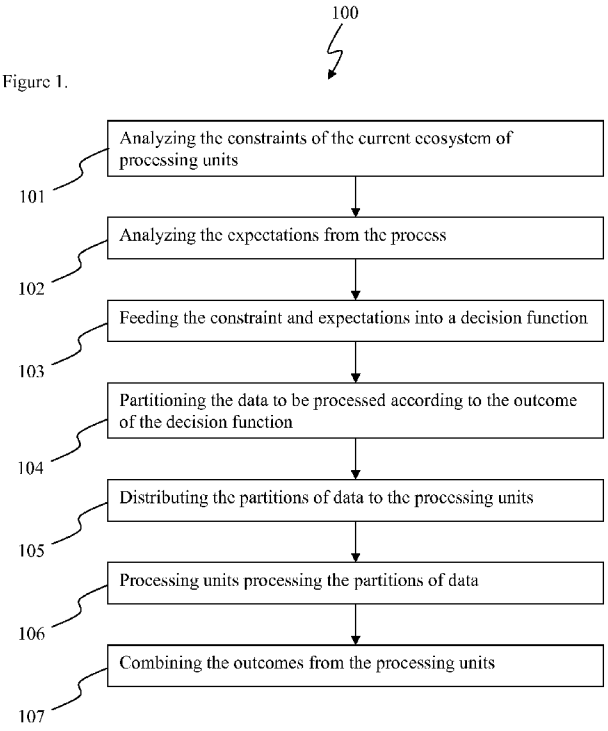
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,

[Continued on next page]

(54) Title: A METHOD FOR DATA PARTITIONING AND ASSIGNMENT

(57) Abstract: The present invention relates to a method for data partitioning and assignment (100) for optimal multi-processor processing that can be utilized for heterogeneous systems, where processors can be stand-alone parallel processors or cores within such stand¬ alone processors, and where partitions of data are assigned according to constraints.

Figure 1.



WO 2014/140697 A1

TM), European (AL, 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, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

DESCRIPTION

A METHOD FOR DATA PARTITIONING AND ASSIGNMENT

Field of the invention

The present invention relates to a method for data partitioning and assignment for optimal multi-processor processing where processors can be stand-alone parallel processors or cores within such stand-alone processors.

Background of the invention

Regardless of the type, in order for the data to be used, it has to be processed. For example, data in a conventional hard disk drive consist of ones and zeros. These ones and zeros are processed so that the type of file they form can be understood. Let's say one group of ones and zeros form a music file. In order to convert these ones and zeros to sound, again, the data has to be processed. Data processing is used in a very broad list of fields. Security systems use data processing for analyzing the audiovisual feedbacks from the devices lying around in order to detect any security breaches. Data processing is also used in analyzing connection parameters in phone systems. While data processing can be realized by a single processing unit, harder tasks require more processing units in order to complete the required work in a shorter time. For this reason, multiple-cored processing units are developed and being used in the personal computer systems. However, some data files are so large that they take a long time to be processed by a typical processing unit. As the need for processing large data grows, so does the number of processors to complete the processing in a reasonable time.

One might know that projects such as SETI@home, Einstein@home etc. have such huge data files that no single supercomputer can process the data in a reasonable time. In order to overcome this problem, the data is divided into smaller chunks and these chunks are distributed to volunteers all over the world to be processed. These chunks

get processed in the computers of the volunteers and the results are sent back to a central server. This approach is often called distributed computing.

State-of-the-art processing platforms are increasingly multi-core and heterogeneous due to the power efficiency and computation power advantages provided by such architectures. Furthermore, state-of-the-art heterogeneous multi-core processing platforms typically include GPUs (Graphics Processing Units) and FPGAs (Field Programmable Gate Arrays) along with other multi-core processors such as ARM (Advances RISC Machines) processors, DSPs (Digital Signal Processors), CPUs (Central Processing Units) and CGRAs (Coarse Grain Reconfigurable Arrays). The inherently multi-core nature of state-of-the-art heterogeneous processing platforms is best leveraged by hardware aware massive multi-threading APIs such as OpenCL (Open Computing Language, currently developed by the Khronos Group) and CUDA (Compute Unified Device Architecture, owned and developed by NVIDIA Corporation). OpenCL supports a wide array of multi-core processors including ARM, CPU, GPU, DSP and FPGA, whereas CUDA is dedicated to NVIDIA GPUs. Using these massive multi-threading APIs engineers can program FPGAs and GPUs to execute massively multi-threaded (64 – 1536 threads) programs and gain substantial performance increases over single threaded programs or simple multi-threaded programs with a small number of threads (2 – 16 threads) typically observed for CPUs, ARMs or DSPs.

Multi-threaded execution (with 2 – 16 threads) of generic algorithms and processing tasks on CPUs, ARM processors and DSPs can provide performance gains that roughly scale with the number of available cores. However, there exists a wide group of algorithms and processing tasks that allow a very high level of parallelism. For these embarrassingly parallel algorithms and processing tasks, massive multi-threading (64 – 1536 threads) provides substantial performance gains over simple multi-threading. GPUs and FPGAs currently present on state-of-the-art heterogeneous multi-core processing platforms are designed for massive multi-threaded execution, and for the algorithms that allow a high level of parallelism these processors are the best choice of execution.

State-of-the-art FPGAs and GPUs can both be used for massive multi-threaded program execution. However, due to their distinctive architectural differences they provide different trade-offs in terms of raw performance, power efficiency, programmability, total execution time and handling thread branching. Current approaches are based on conducting trial and error based studies for a given algorithm or processing task that will be massively multi-threaded and based on these studies picking GPU or FPGA as the final platform of execution. Once the decision is made, the algorithm is mapped to a specific processing platform and executes on it. There also exist more involved algorithm-to-processor mapping based approaches such as elastic computing disclosed in **WO2012142069**, which assigns each processing task or algorithm to a specific processor based on a detailed analysis.

The international patent application numbered **WO2011079942** discloses a method for compiling code for a multi-core processor, comprising: detecting and optimizing a loop, partitioning the loop into partitions executable and mappable on physical hardware with optimal instruction level parallelism, optimizing the loop iterations and/or loop counter for ideal mapping on hardware, chaining the loop partitions generating a list representing the execution sequence of the partitions..

The international patent application numbered **WO2012142069** discloses elastic computing tools which analyze available devices and resources (e.g., cores, GPUs, FPGAs, etc.) and current run-time parameters, and then transparently select from numerous pre-analyzed implementation possibilities to optimize for performance, power, energy, size, or any combination of these goals.

Objects of the invention

The object of the invention is to provide a method for data partitioning and assignment for optimal multi-processor processing where processors can be stand-alone parallel processors or cores within such stand-alone processors

The object of the invention is to provide a method for data partitioning and assignment where the partitions of data are assigned to processors according to constraints.

Another object of the invention is to provide a method for data partitioning and assignment, which can be utilized for heterogeneous systems.

Detailed description of the invention

A method for data partitioning and assignment in order to fulfill the objects of the present invention is illustrated in the attached figures, where:

Figure 1 is the flowchart of the method for data partitioning and assignment.

A method for data partitioning and assignment (100) comprises the steps of;

- analyzing the constraints of the current ecosystem of processing units (101),
- analyzing the expectations from the process (102),
- feeding the constraint and expectations into a decision function (103),
- partitioning the data to be processed according to the outcome of the decision function (104),
- distributing the partitions of data to the processing units (105),
- processing units processing the partitions of data (106),
- combining the outcomes from the processing units (107)

In the method for data partitioning and assignment (100), the constraints of the current ecosystem of processing units are analyzed (101). These constraints and capabilities include but are not limited to power consumption, thread branching capabilities, workloads per thread, memory access patterns, thread synchronization requirements, processor functionality checks (to see a specific processor is still functional), etc. After determining the constraints and capabilities of the current processing units ecosystem, expectations from the process is also analyzed (102). The

expectations include but are not limited to maximum power consumption, maximum process time, the quality of the outcomes etc. After determining the expectations, the capabilities, constraints and expectations are fed to a decision function (103). The decision function is a function, which uses operations such as comparisons, mathematical operations or any other operations, by itself or in combination with other operations, in order to calculate the optimum set of processing units, taking the expectations and constraints of every different application into consideration. The decision function yields a set of processing units for the data to be fed. The data is then partitioned in order to meet the set of processing units yielded by the decision function (104). The partitions of data are distributed to different processing units in order to be processed (105). The processing units then process the partitions and yield outcomes (106). As the outcomes of the processing units represent the partitioned data that is fed to the processing units, the outcomes are unified in order to form the non-partitioned data (107).

In one embodiment of the invention, the decision function has learning capability. It keeps an internal set of parameters, which pertain to linear or non-linear decision and/or update function(s), which it updates upon every batch of input values. This decision function is named decision function with learning.

In one embodiment of the invention, the constraints of the process remain constant throughout the process. Thus, running the decision function once and processing the data according to the outcome of this decision function is sufficient. This decision function is named off-line decision function.

In another embodiment of the invention, the constraints of the process are subject to change throughout the process. These changes include, but are not limited to remaining battery power, number of different types of processing units, changing time constraints etc. For example, if the system on which the method for data partitioning and assignment (100) runs is a standalone system, the remaining battery power can be taken into consideration. If the number of processing units can change in a system during the execution of the method for data partitioning and assignment

(100) (because of burnt processors etc.), the decision function can recalculate the set of processing units for the data to be fed. In order to achieve this, the method for data partitioning and assignment (100) is executed occasionally. This decision function is named on-line decision function.

In one embodiment of the invention, there may be long periods of times when the constraints of the process remain constant throughout the process with occasional periods when the constraints of the process are subject to change throughout the process. For such use cases, there will be an additional block right before the decision function, named the steady state finder. This block will determine if the system is in steady state and thus the off-line decision function should be used or if the system is not in steady state and thus the on-line decision function should be used. Such conjured aggregate decision function is named hybrid decision function.

In one embodiment of the invention, the input data may be a color image in planar format consisting of red, green and blue color planes. Typically green color plane is sampled twice more densely compared to red and blue color planes and contains more spatial detail. As a result, green color plane is typically targeted for heavier spatial processing. In such a scenario, each of the color planes can be processed on different processing unit. For example, green plane can be processed on a GPU while red and blue planes can be processed on FPGAs or vice versa.

In one embodiment of the invention, the input data may be a color image in planar format consisting of Luminance (Y), and chrominance (Cb, Cr) planes. Due to the nature of the YCbCr color space, most of the spatial details are present in the Y plane. Furthermore, based on this observation Cb and Cr planes are down-sampled in the spatial domain to save space. As a result, Y color plane is typically targeted for heavier spatial processing. In such a scenario, the luminance plane can be processed on a GPU and chrominance (Cb, Cr) planes can be processed on FPGAs or vice versa.

In one embodiment of the invention, the input data may be consisting of separate

frequency bands. In such a scenario, each of the frequency bands can be processed on different processor. For example, higher frequency bands can be processed on GPUs and lower frequency bands can be processed on FPGAs or vice versa.

Although the present invention has been described and illustrated in detail, it is clearly understood that the same is by way of illustration and example only and is not to be taken by way of limitation, the spirit and scope of the present invention being limited only by the terms of the appended claims.

CLAIMS

1. A method for data partitioning and assignment (100) characterized in that it comprises the steps of;
   - analyzing the constraints of the current ecosystem of processing units (101),
   - analyzing the expectations from the process (102),
   - feeding the constraint and expectations into a decision function (103),
   - partitioning the data to be processed according to the outcome of the decision function (104),
   - distributing the partitions of data to the processing units (105),
   - processing units processing the partitions of data (106),
   - combining the outcomes from the processing units (107).

2. A method for data partitioning and assignment (100) according to claim 1 characterized in that the decision function is decision function with learning.

3. A method for data partitioning and assignment (100) according to claim 1 characterized in that the decision function is off-line decision function.

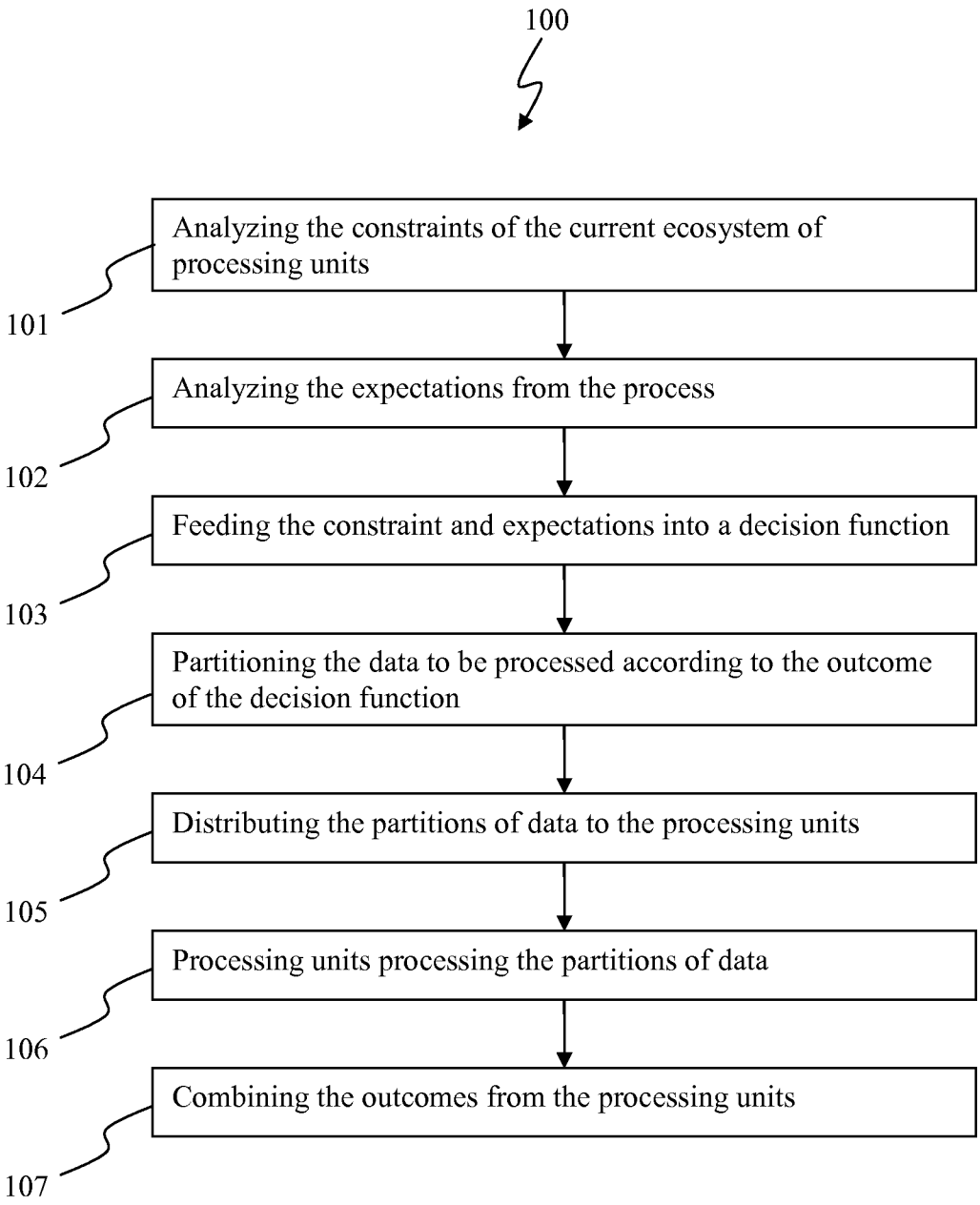
4. A method for data partitioning and assignment (100) according to claim 1 characterized in that the decision function is on-line decision function.

5. A method for data partitioning and assignment (100) according to claim 1 characterized in that the decision function is hybrid decision function.

6. A method for data partitioning and assignment (100) where data partitioning and assignment is based on color plane characteristics.

7. A method for data partitioning and assignment (100) where data partitioning and assignment is based on frequency bands.

Figure 1.


100
Analyzing the constraints of the current ecosystem of processing units
101
Analyzing the expectations from the process
102
Feeding the constraint and expectations into a decision function
103
Partitioning the data to be processed according to the outcome of the decision function
104
Distributing the partitions of data to the processing units
105
Processing units processing the partitions of data
106
Combining the outcomes from the processing units
107

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2013/052098

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F9/50
ADD.

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 practicable, search terms used)
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BINGSHENG HE ET AL: "Mars", PROCEEDINGS OF THE 17TH INTERNATIONAL CONFERENCE ON PARALLEL ARCHITECTURES AND COMPILATION TECHNIQUES, PACT '08, 29 October 2008 (2008-10-29), page 260, XP055111303, New York, New York, USA DOI: 10.1145/1454115.1454152 ISBN: 978-1-60-558282-5 abstract page 260, right-hand column, line 9 - line 24 page 261, right-hand column, line 14 - line 34 page 263, left-hand column, line 1 - line 22; figure 2 Section 3.4.2, paragraphs 1-3:; page 264	1-7

-/--

[X] 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 application or patent 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	Date of mailing of the international search report
8 April 2014	14/04/2014

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Ross, Christopher

4

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2013/052098

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	COLBY RANGER ET AL: "Evaluating MapReduce for Multi-core and Multiprocessor Systems", 2007 IEEE 13TH INTERNATIONAL SYMPOSIUM ON HIGH PERFORMANCE COMPUTER ARCHITECTURE, 1 January 2007 (2007-01-01), pages 13-24, XP055111222, DOI: 10.1109/HPCA.2007.346181 ISBN: 978-1-42-440804-7 abstract; figure 1 Starting with section 2.2,; page 14, right-hand column - page 15, right-hand column, line 31 Section 3.2.1:; page 16 Section 3.2.4:; page 18 page 19, left-hand column, line 16 - line 27 page 20, left-hand column, line 15 - line 21 -----	1-7
X	YI SHAN ET AL: "FPMR", PROCEEDINGS OF THE 18TH ANNUAL ACM/SIGDA INTERNATIONAL SYMPOSIUM ON FIELD PROGRAMMABLE GATE ARRAYS, FPGA '10, 21 February 2010 (2010-02-21), page 93, XP055111216, New York, New York, USA DOI: 10.1145/1723112.1723129 ISBN: 978-1-60-558911-4 page 93, right-hand column, paragraph 1 page 93, right-hand column, paragraph 3 - page 94, left-hand column; figure 1 Starting at section 2.1:; page 94, right-hand column - page 95, left-hand column, last line; figure 3 Starting at section 2.3:; page 95, right-hand column - page 96, left-hand column, line 14 -----	1-7

4