Data Streaming Processing Window Joined With Graphics Processing Units (GPUs)

Shen Lu
University of South Florida, USA

Richard S. Segall
Arkansas State University, USA

INTRODUCTION

A massive wave of changes in industries are created by data. Business models, applications and even data driven economy depend on data to overcome the stress of the competition and even ensure the certainty of the survival. Data can represent not only the development trends but also the values of the changes. With the time passing, the values of data decreases, which makes the usage of time-series time more important.

Big Data is Large-scale data and can be either discrete or continuous. This article entails research that discusses the continuous case of Big Data often called "Data Streaming". More and more businesses will depend on being able to process and make decisions on streams of data. This article utilizes the algorithmic side of data stream processing often called "stream analytics" or "stream mining".

In a streaming data system, the data is made available at the moment a client application needs it. As stated by Warski (2016) processing data in a streaming fashion has become more and more popular over the "traditional" way of batch processing of Big Data sets available as a whole. According to Warski (2016), the focus has shifted in the industry of "it's no longer that important how big is your data but rather it's much more important of how fast you can analyze it and gain insights."

Data stream processing incorporates a notion of a Window on a stream as a way to convert an infinite stream into a finite relation in order to apply relational operations. The data stream can be based on either time or the tuple. Time-based execution provides a way to model simultaneity, and tuple-based execution provides a way to react to primitive events. To process the two different kinds of streams, one would need to find a convergence language in which the simple difference between the two streams would be resolved. For this data streaming Window Join, one would need to focus on time-based data stream.

BACKGROUND

A Central Processing Unit (CPU) is the electronic circuitry within a computer that carries out the instructions of a computer program by performing the basic arithmetic, logical, control and input/output (I/O) operations specified by the instructions. The computer industry has used the term "central processing unit" at least since the early 1960s. Traditionally, the term "CPU" refers to a processor, more specifically to its processing unit and Control Unit (CU), distinguishing these core elements of a computer from external components such as main memory and I/O circuitry. (Wikipedia (2018e)).

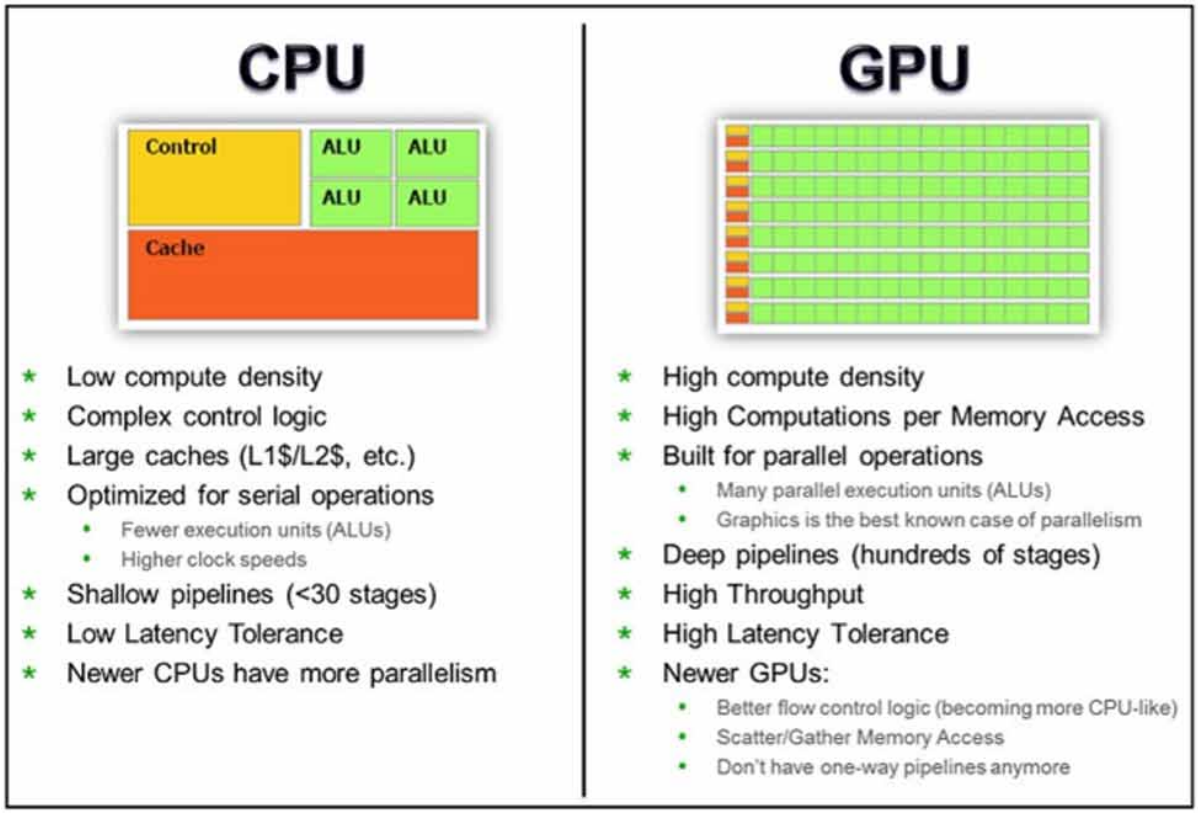
DOI: 10.4018/978-1-7998-3479-3.ch043

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

A Graphics Processing Unit (GPU) is a specialized electronic circuit designed to rapidly manipulate and alter memory to accelerate the creation of images in a frame buffer intended for output to a display device. GPUs are used in embedded systems, mobile phones, personal computers, workstations, and game consoles. Modern GPUs are very efficient at manipulating computer graphics and image processing, and their highly parallel structure makes them more efficient than general-purpose CPUs for algorithms where the processing of large blocks of data is done in parallel. In a personal computer, a GPU can be present on a video card, or it can be embedded on the motherboard or—in certain CPUs—on the CPU die. (Wikipedia (2018d).

Figure 1 below shows the highlights of comparisons between Central Processing Unit (CPU) versus Graphics Processing Unit (GPU). Note that a CPU has a low compute density and a GPU has a high compute density. A CPU has a low latency tolerance and a GPU has a high latency tolerance, where latency refers to a measure of the time delay required for information to travel across a network. A CPU has shallow pipelines in contrast to a GPU that has high throughput. Hence this article discusses experiments using GPUs that are extremely more efficient than the traditional CPU framework.

Figure 1. Comparison of CPU (Central Processing Unit) vs. GPU (Graphics Processor Units (GPUs)



Below is a table of highlights of research advances from 2002 to 2018 in data streaming related to the data streaming processing of Window Joins with Graphics Processing Units (GPUs) and topic of processing algorithms for data analysis of streaming data.

20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/data-streaming-processing-window-joined-with-graphics-processing-units-gpus/260217

Related Content

Acquiring Competitive Advantage through Effective Knowledge Sharing
Iman Raeesi Vanani, Davood Qorbaniand Babak Sohrabi (2015). *Encyclopedia of Information Science and Technology, Third Edition (pp. 4980-4988).*
www.irma-international.org/chapter/acquiring-competitive-advantage-through-effective-knowledge-sharing/112946

Application Research of Constructing Large Networks in Tourist Attraction Based on Leaf Ridge Network Architecture
Li Liu, Ziyan Liu, Huizhen Long, Jing Shiand Liang Wang (2026). *International Journal of Information Technologies and Systems Approach (pp. 1-17).*
www.irma-international.org/article/application-research-of-constructing-large-networks-in-tourist-attraction-based-on-leaf-ridge-network-architecture/404753

Multimodality Medical Image Fusion using M-Band Wavelet and Daubechies Complex Wavelet Transform for Radiation Therapy
Satishkumar S. Chavanand Sanjay N. Talbar (2015). *International Journal of Rough Sets and Data Analysis (pp. 1-23).*
www.irma-international.org/article/multimodality-medical-image-fusion-using-m-band-wavelet-and-daubechies-complex-wavelet-transform-for-radiation-therapy/133530

Review and Brief History of Collaborative Systems: Taxonomy, Services and Classification
Nuria Lloret Romero (2012). *Systems Science and Collaborative Information Systems: Theories, Practices and New Research (pp. 139-160).*
www.irma-international.org/chapter/review-brief-history-collaborative-systems/61289

On the Suitability of Soft Systems Methodology and the Work System Method in Some Software Project Contexts
Doncho Petkov, Steven Alter, Olga Petkovaand Theo Andrew (2013). *International Journal of Information Technologies and Systems Approach (pp. 22-34).*
www.irma-international.org/article/on-the-suitability-of-soft-systems-methodology-and-the-work-system-method-in-some-software-project-contexts/78905