

Chapter 48

Comparative Analysis of Efficient Platforms: Scalable Algorithms and Parallel Paradigms for Large Scale Image Processing

Khawaja Tehseen Ahmed

Bahauddin Zakariya Univeristy, Pakistan

Mazhar Ul-Haq

*NUST School of Electrical Engineering and
Computer Science, Pakistan*

Arsalaan Ahmed Shaikh

*NUST School of Electrical Engineering and
Computer Science, Pakistan*

Raihan ur Rasool

*NUST School of Electrical Engineering and
Computer Science, Pakistan*

ABSTRACT

With the advancement of technology we are heading towards a paperless environment. But there are still a large numbers of documents that exist in paper format in our daily lives. Thus the need to digitize these paper documents, archive them and view them at all times has arisen. The number of documents of a small organization may be in thousands, millions or even more. This chapter presents comparative analysis of different programming languages and libraries where it is intended to parallel process a huge stream of images which undergo unpredictable arrival of the images and variation in time. Since the parallelism can be implemented at different levels, different algorithms and techniques have also been discussed. It also presents the state of the art and discussion of various existing technical solutions to implement the parallelization on a hybrid platform for the real time processing of the images contained in a stream. Experimental results obtained using Apache Hadoop in combination with OpenMP have also been discussed.

1. INTRODUCTION

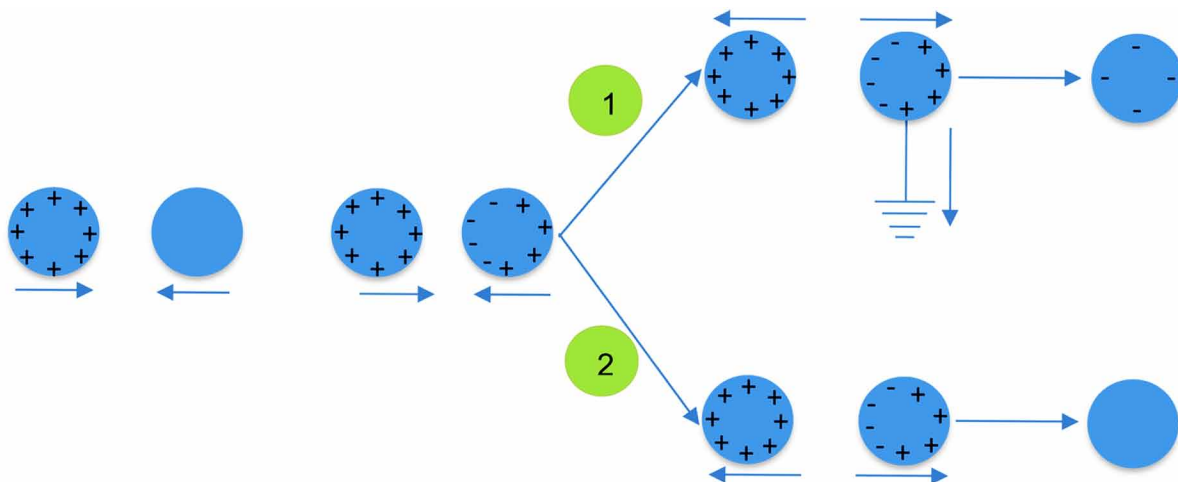
This research is part of an ongoing project that aims to offer a service of dematerialization of documents to very small enterprises. Dematerialization is being done in order to archive documents that exist in a paper form, and efficiently retrieve them on demand from anywhere. When this service will be operational, scanners located in very small firms will be able to digitize paper documents before sending these

DOI: 10.4018/978-1-5225-0983-7.ch048

so obtained numeric images to a server. Then, this server will process the received images in order, on one hand, to improve the quality of and, on the other hand, to reduce the volume of the digitized images. This process has been described in Figure 1.

Finally, the processed images will be communicated to an archiving center which will allow the owner of an archived document, to consult its digital image. For this purpose data and task parallelism is required for High Performance Computing and the processing of every image may contain tasks which could be run in parallel on the various cores of each computational node. Image processing is widely applied in many applications such as medical image processing, non-photorealistic rendering, remote sensing, optical sorting and many more. The processing of the flow of images is difficult to characterize. Some applications require to process images that are of very large size and require that all the processing can be done in a very short interval of time. The throughput of arrival of the images may undergo great and unpredictable variations according to the time and the duration of the processing of an image can also undergo great and unpredictable variations according to the image type and size. The overall process can be divided into three portions scanning and transmitting, processing and archival. The focus of this study is towards the processing module which receives images and processes them. Our aim is to study the problems that may arise in this area and technologies that address them, and propose and test a solution. Following are some areas of concern: Processing of images that are contained in a stream. The number of images to process at any given time may vary to a great extent The processing time is dynamic and depends on each individual image Because of the varying amount of load on the processing module, a scalable platform is needed to be identified and technologies that provide parallelism have to be studied so that the system efficiently processes high number of images in a certain amount of time. This chapter describes various aspects of different technologies related to image processing on hybrid platform. The objective of this research is to analyze various technical solutions in parallel image processing, where different techniques are compared to each other, and the best suited techniques are suggested for parallel image processing on large scalable infrastructure. Section 2 analyzes different platforms and argues

Figure 1. Conversion and archival of documents



18 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/comparative-analysis-of-efficient-platforms/164646

Related Content

NT-ODP Theoretical Model Extension: Toward an Efficiency-Based New Technologies Use of the Organization Digital Patrimony

Mustapha El Hamzaoui and Faycal Bensalah (2021). *International Journal of Smart Security Technologies* (pp. 1-15).

www.irma-international.org/article/nt-odp-theoretical-model-extension/272098

Online Spatial HIV/AIDS Surveillance and Monitoring System for Nigeria

Peter Adebayo Idowu (2017). *Biometrics: Concepts, Methodologies, Tools, and Applications* (pp. 448-479).

www.irma-international.org/chapter/online-spatial-hiv-aids-surveillance-and-monitoring-system-for-nigeria/164615

Dynamical Analysis of Drug Efficacy and Mechanism of Action Using GFP Reporters

Jianping Hua, Chao Sima, Milana Cypert, Edward R. Dougherty, Jeffery M. Trent and Michael L. Bittner (2017). *Biometrics: Concepts, Methodologies, Tools, and Applications* (pp. 1145-1174).

www.irma-international.org/chapter/dynamical-analysis-of-drug-efficacy-and-mechanism-of-action-using-gfp-reporters/164642

An Adaptive Magnetic Field Source for Magnetic Drug Fixation

Alexandru Mihail Morega, Cristina Savastru and Mihaela Morega (2013). *International Journal of Monitoring and Surveillance Technologies Research* (pp. 20-33).

www.irma-international.org/article/an-adaptive-magnetic-field-source-for-magnetic-drug-fixation/101963

Flow-Based Anomaly Detection Using BNN for Attack Mitigation on SDN

Nang May Phu Lwin and Su Thawda Win (2022). *International Journal of Smart Security Technologies* (pp. 1-17).

www.irma-international.org/article/flow-based-anomaly-detection-using-bnn-for-attack-mitigation-on-sdn/304072