Accepted Manuscript

Prototyping a GPGPU neural network for deep-learning big data analysis

Alcides Fonseca, Bruno Cabral

PII: S2214-5796(16)30040-5
DOI: http://dx.doi.org/10.1016/j.bdr.2017.01.005
Reference: BDR 57

To appear in: *Big Data Research*

Received date: 17 April 2016
Revised date: 14 November 2016
Accepted date: 20 January 2017

Please cite this article in press as: A. Fonseca, B. Cabral, Prototyping a GPGPU neural network for deep-learning big data analysis, *Big Data Res.* (2017), http://dx.doi.org/10.1016/j.bdr.2017.01.005

Prototyping a GPGPU Neural Network for Deep-Learning Big Data Analysis

Alcides Fonseca[1] and Bruno Cabral[1]

University of Coimbra, Portugal
{amaf,bcabral}@dei.uc.pt

Abstract. Big Data concerns with large-volume complex growing data. Given the fast development of data storage and network, organizations are collecting large ever-growing datasets that can have useful information. In order to extract information from these datasets within useful time, it is important to use distributed and parallel algorithms.
One common usage of big data is machine learning, in which collected data is used to predict future behavior. Deep-Learning using Artificial Neural Networks is one of the popular methods for extracting information from complex datasets. Deep-learning is capable of more creating complex models than traditional probabilistic machine learning techniques.
This work presents a step-by-step guide on how to prototype a Deep-Learning application that executes both on GPU and CPU clusters. Python and Redis are the core supporting tools of this guide. This tutorial will allow the reader to understand the basics of building a distributed high performance GPU application in a few hours. Since we do not depend on any deep-learning application or framework — we use low-level building blocks — this tutorial can be adjusted for any other parallel algorithm the reader might want to prototype on Big Data. Finally, we will discuss how to move from a prototype to a fully blown production application.

Keywords: Big-Data, Deep-Learning, Prototyping, GPGPU, Cluster, Distributed, Parallel Programming

1 Introduction

Deep Learning [1] refers to the usage of Artificial Neural Networks (ANN or NN) with several hidden layers used for data with a high dimensionality. A common example and benchmark for Deep Learning is image classification from the ImageNet dataset [2]. ANNs can be used for classification tasks [3], with several applications in industry, business and science [4]. Examples of applications include character recognition in scanned documents [5], predicting bankruptcy [6] or health complications [7]. Autonomous driving also makes heavy use of ANNs [8].

An ANN begins with random weights, practically deciding everything at random. By training the ANN with several existing instances of the problem,