Future Generation Computer Systems 37 (2014) 203–211

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Exploiting mean field analysis to model performances of big data architectures

Aniello Castiglione [a,*], Marco Gribaudo [b], Mauro Iacono [c], Francesco Palmieri [d]

[a] Dipartimento di Informatica, Università degli Studi di Salerno, Via Ponte Don Melillo, I-84084 Fisciano, Italy
[b] Dipartimento di Elettronica, Informatica e Bioingegneria, Politecnico di Milano, Via Ponzio 34/5, I-20133 Milano, Italy
[c] Dipartimento di Scienze Politiche, Seconda Università degli Studi di Napoli, Viale Ellittico 31, I-81100 Caserta, Italy
[d] Dipartimento di Ingegneria Industriale e dell'Informazione, Seconda Università degli Studi di Napoli, Via Roma 29, I-81031 Aversa, Italy

HIGHLIGHTS

- The paper presents a set of methods for approximate inference of probabilistic models.
- The proposed approach limits the excessive classical state-space growth problem.
- The MFA can be applied to modeling problems with very large scale stochastic systems.
- The MFA is effective and reliable in evaluating the performance of very large big data.
- The MFA is able to model performance of big data architectures indices in a bounded time.

ARTICLE INFO

Article history:
Received 9 March 2013
Received in revised form
13 June 2013
Accepted 17 July 2013
Available online 5 August 2013

Keywords:
Multi-formalism modeling
Big data
Mean field analysis
Performance evaluation
Map-reduce architectures

ABSTRACT

Big data processing systems are characterized by a relevant number of components that are used in parallel to run multiple instances of the same tasks in order to achieve the needed performance levels in applications characterized by huge amounts of data. Such a number of components depend on the dimension of the involved data, so that new resources (e.g., processing or storage servers) are usually added as the working database grows. A reliable performance evaluation of these systems is at the same time crucial, in order to enable administrators and developers to keep the pace with data growth, and extremely difficult, due to the intrinsic complexity of these architectures. Notwithstanding, the available literature does not yet offer sufficient experiences, nor significant methodologies, in such a direction.

This paper presents a novel modeling approach, based on mean field analysis, a set of methods for approximate inference of probabilistic models, derived from statistical physics, for performance evaluation of big data systems. This approach, by containing the excessive state space growth characterizing more traditional modeling methodologies, also requires a significantly reduced effort with respect to simulation based ones.

1. Introduction

The ever increasing need for modern network-based ubiquitous applications and processing a huge quantity of data in short times is fostering the massive development of architectures commonly referred to as big data processing systems. Such systems are advocated as the most flexible and promising solution to handle the massively parallel computations needed to benefit from the continuously growing, dynamic information generated by users on the Internet or diffuse wide-scale sensing or data collection/tracking systems. For example, the data access and processing demand characterizing applications like Online Social Networks (OSNs), High Energy Physics experiments, Business Intelligence and genomic data can consist of many petabytes of data, increasing in some cases on a daily basis at a fast pace, and posing a serious challenge for the scalability and timeliness of the required computation.

Big data processing architectures face such a challenge by exploiting, by means of proper paradigms, interfaces and cooperation models, a huge number of relatively inexpensive computing or storage resources/nodes, that may be composed by several processing units and equipped with a significant number of disks,

* Corresponding author. Tel.: +39 3808800044; fax: +39 089969821.
E-mail addresses: castiglione@acm.org, castiglione@ieee.org (A. Castiglione), gribaudo@elet.polimi.it (M. Gribaudo), mauro.iacono@unina2.it (M. Iacono), fpalmier@unina.it (F. Palmieri).

0167-739X/$ – see front matter
http://dx.doi.org/10.1016/j.future.2013.07.016

distributed throughout the Internet and connected by high speed communication links. Traditional relational database management systems (RDBMS) are not fit for such applications, because of the dimensions and the nature of relevant data, that consist of simple, enormous lists rather than complexly interrelated and articulated database schemas. Big data is rather organized in chunks of such lists, often replicated on multiple different sites according to several strategies [1], each potentially handled by a different node. The distributed cooperation paradigm used on these systems is known as *map-reduce*, and is based on the replication of a common task on each computing node, properly designed to execute the desired algorithm on its chunk, followed by a phase in which all results are collected and properly processed. The number of nodes depends on the number of lines of such lists and on the capabilities of each node, but it can be roughly said that it can continuously grows with the arrival of new data.

Modeling and evaluating the performances of big data architectures is a really difficult task, due to their intrinsic complexity in terms of structure and number of components, whereas the capability to predict the behavior of such a system would be an important support for capacity planning and management, both if the system is physically owned by stakeholders or if it is deployed on a cloud computing facility. Thus, the paradigm shifts from RDBMS-based applications to map-reduce-driven big data applications, necessarily introduced because of the explosion of new data-centric network-based architectures and computing models, is far from being a silver bullet, and the involved amount of data could dramatically enhance the problems caused by wrong decisions or data storage and access strategies, since "big data" does not mean "infinite capacity". Anticipating the effects of the deployment of new resources/nodes, being able to explore the scalability limits of implemented applications, supporting the comparison between different future strategies, is of paramount importance to design a correct evolution roadmap of each big data architecture, and keep potential scalability problems under control. While literature reports several results in benchmarking big data systems, a few works seem to have been conducted about simulating or analytically modeling them for performance forecasting purposes.

In this paper we propose an approximate evaluation approach for big data systems, based on mean field analysis. Mean field analysis allows stochastic modeling of processes in which a big number of instances from a finite number of classes of model elements interact with each other. This technique allows a representation of such processes in terms of classical modeling languages, such as Petri nets or Markov chains, that are then used to generate a complete and reliable system model that can be solved bypassing the state space explosion problem typical of traditional modeling techniques.

The paper is organized as follows: after this introduction, big data modeling problems and experiences are introduced in Section 2, together with the main ideas behind mean field analysis. Section 3 presents the general approach, used in Section 4 to evaluate a case study taken from a real experience. Section 5 presents the main results obtained, followed by the conclusions in Section 6.

2. Related works

Performance analysis and modeling for large-scale Big Data processing architectures is widely recognized to be an extremely challenging task. This is essentially due to the large number of variables involved, starting from complex applications potentially composed by hundreds or thousands of cooperating tasks running on many different real or virtual machines, providing runtime resources, ideally scattered throughout the Internet, where each different entity is characterized by its own properties, performance details and operating constraints, often not easily representable through a traditional high level model (see [2–5]). Accordingly, several different approaches have been used to model and predict the overall system performance and evolution trends, ranging from simulation, diagnostics observations to more analytical approaches, such as resource usage/demand characterization, predictive modeling, etc.

Several simulation efforts, such as [6–9], try to predict application performance in MapReduce architectures. Other simulations, presented in [10], have been driven by logs collected from Microsoft large-scale clusters.

Alternatively, several diagnostic tools developed for Hadoop and Dryad, such as Chukwa [11], Kahuna [12] and Artemis [13], focus on mining the system logs to detect performance behavior and problems. In particular, the analysis of very large Hadoop logs on a 400-nodes system has been presented in [14] whereas other data made available by Google, concerning their cloud backend, are available from [15,16].

While these approaches can lead to significant quantitative performance information and good indications about the distributed system behavior they are affected by an extremely serious drawback: simulating or assessing huge architecture, working with very large datasets can be prohibitively time-demanding.

For many distributed data-intensive architectures/applications, the study of data transfers across the network may lead to extremely effective performance indications. The experiences presented in [17] characterize the communication between realistic applications in data intensive parallel systems as a basis for modeling and predicting their future performance. On the other hand, the proposal in [18] leverages specific templates reflecting the structure of the involved application and algorithms, by combining them with the specific hardware features to reliably estimate how the system performance evolves over time.

The works presented in [19,20] face with the problems of performance evaluation and forecasting by respectively modeling the performance by using "Resource Usage Equations" and developing a methodology for characterizing applications performance from their evolution trends in terms of large datasets usage and system configurations. Other approaches, such as [21] model complex systems behavior by using mathematical equations and variables, that can be easily used for performance forecasting.

In [22] a sophisticated performance prediction framework specifically developed for scientific applications has been presented.

On the other hand, the authors in [23] propose an accurate, predictive analytical model for large scale applications encompassing the performance and scaling characteristics of SAGE, a multidimensional hydrodynamics code with adaptive mesh refinement. Linear prediction models for large distributed applications have been carefully examined in [24] by determining that such models can be able to accurately predict load performance over a wide range of conditions. Finally, in [25] compositional models for distributed data-intensive Web applications have been studied in order to analyze performance scaling.

While mean field analysis has been recently applied to the analysis of very large communication systems [26,3], to the best of our knowledge this is the first approach based on multiformalisms and mean field analysis for effectively modeling and analyzing the performance of very complex architectures such the ones characterizing the emerging data-intensive distributed applications. We envision that this approach can be more time-efficient than simulation and log analysis-based ones, giving at the same time more precise and realistic results than traditional analytic or prediction models.