**Call for Book Chapter**

****

**Distributed Computing in Big Data Analytics**

***Concepts, Technologies and Applications***

**Editors**

**Sourav Mazumder**

IBM Analytics lab Service

San Francisco, USA

Email: smazumder@us.ibm.com

**Robin Sing Bhadoria**

Discipline of Computer Science & Engineering

Indian Institute of Technology, Indore, India

Email: [robin19@ieee.org](mailto:robin19@ieee.org)

**Dr. Ganesh Chandra Deka**

Ministry of Skill Development & Entrepreneurship, India

Email: ganeshdeka2000@gmail.com

Introduction

Big Data Analytics is becoming a very important tool in decision making in Business, Health Care and even walks life of individuals. The Big Data Technologies are used for Storage, Retrieval & Processing of high volume of data of various types to achieve any type of Analytics in a fast and predictable way. Distributed Computing is the key to Big Data Analytics. Distributed Computing Algorithms and Tools enables Big Data Storage, Access, Transfer and Visualization of huge volume of data in multiple low cost machines within stipulated time with minimum possible cost.

This book aims to cover the gap in the current space of books/literature available for Big Data Analytics using Distributed Computing technology. The chapters in this book will cover the key concepts of Distributed Computing which are important and widely used in Big Data Analytics. The proposed book will also cover the key technologies which support Distributed Processing in Big Data Analytics worlds. Finally, this book will also cover key Applications of Big Data Analytics and highlight how principles of Distributed Computing are used in those cases.

Following would be the salient features of the book:

* The key concepts and patterns of Distributed Computing which practitioners need to keep in mind while designing Big Data Analytics use cases.
* Details on how exactly different Big Data Technologies leverage those key concepts and patterns of Distributed Computing.
* The most novel part would be the discussion of the key Big Data Analytics applications, widely relevant in current world, in the light of Distributed Computing principles and technologies. The types of applications being covered are – IoT, Cognitive Analytics, Social Media Analytics and Scientific Data Analytics.

Publisher

This book will be published by Springer in **Scalable Computing and Communications** book series. For more information, please visit: <http://www.springer.com/series/15044>

Submission Deadlines

**December 30**, **2016**: Submission of full chapter

**January 10**, **2017**: Acceptance/Rejection notification

**January 31**, **2017**: Submission of revised Chapter

Topics

Interested researcher and academician are invited to send their proposed chapter abstract with tentative table of content (ToC). Following chapters could be contributed for above mention book title:

Chapter 1- Introduction of Distributed Computing

* General Introduction to Distributed Computing
* History and Motivation behind Distributed Computing
* A simple model for Distributed Computing
* Key characteristics of Distributed Computing
* Examples of various technologies utilizing concepts of Distributed Computing
* Introduction to Big Data Analytics and its importance in current and future use cases
* The role of Distributed Computing in Big Data Analytics

**Part 1: Distributed Computing Concepts in Big Data Analytics**

Chapter 2- Fundamental Concepts of Distributed Computing used in Big Data Analytics

* Multi-threading and Multi processing
* Computing Architectures - SIMB, MIMD, Array vs Vector Processors
* Scale up vs Scale out systems
* Queuing Network Model for Distributed Computing –Workloads, Service Centers and SLAs
* Application of CAP Theorem
* Characteristics of Computing Clusters - Bottlenecks and Challenges
* Quality of Service (QoS) requirements in Big Data Analytics implemented using Fundamental Concepts of Distributed Computing - Performance, Interoperability (with multiple systems and interfaces), Fault Tolerance, Consistency, Availability, Security, Elasticity, Manageability

Chapter 3- Distributed Computing Patterns useful in Big Data Analytics

* Distributed Computing Primitives – Lock, Synchronization, Queue, Knowledge, Logical Clocks etc.
* Communication Protocols – Timer based, Balanced Sliding
* Various communication patterns - Synchronous, Asynchronous and Pseudo Synchronous Communications
* Various distributed processing patterns - MPI, MPP, Scatter Gather, Shared memory systems
* Message Ordering Patterns
* Distributed Shared Data – Partitions, Lineage, Indices , Global States, Data Locality, Predicates, etc.
* Deadlocks – various scenarios and solutions
* Distributed Transaction Management
* Common Algorithms used in Distributed Computing - Routing Algorithms, Election Algorithms, Wave Algorithms, Traversal Algorithms, Graph Algorithms etc.
* Use of various Distributed Computing patterns in Big Data Analytics Use Cases

**Part 2: Technologies**

Chapter 4- Distributed Computing in Core Technologies used in Big Data Analytics

* Parallel RDBMS
* Data warehousing appliances
* Hadoop Ecosystem
* NoSQL Databases
* Fast Data Analytics Systems (Spark, Flink)
* Distributed Streaming Technologies
* Distributed Search
* Distributed Content Management technologies
* Distributed Machine Learning
* Reporting and Visualization Technologies

Chapter 5- Distributed Computing in Infrastructure Technologies used in Big Data Analytics

* Meta Data Management Technologies
* Distributed Messaging systems
* Resource Management Technologies
* Security Technologies
* High Availability Technologies
* Management technologies for big Clusters
* Distributed Caching Technologies
* Software and Hardware Load balancing and Failover Technologies
* Cloud Computing

**Part 3: Industry Applications**

Chapter 6- Distributed Computing in Internet of Things

* QoS Requirements of Distributed Computing relevant to IoT – High Data Volume, Fast Data Crunching, Interoperability (Supporting multiple Interfaces for data ingestion and data consumption at very fast rate), Elasticity, Security, Priority, etc.
* Mapping various Distributed Computing Patterns, Algorithms to address the QoS requirements of IoT
* Mapping various Distributed Computing Technologies used in IoT
* 2 Examples/Case Studies (e.g. Smart Cities, Connected Cars, etc.) related to IoT elaborating the QoS, Patterns and Algorithms to be used, Choice of Technologies, etc.

.

Chapter 7 – Distributed Computing in Cognitive Analytics

* QoS Requirements of Distributed Computing relevant to Cognitive Analytics – High Data Volume, Fast Data Crunching, Data Availability, Elasticity, Security, Priority etc.
* Mapping various Distributed Computing Patterns, Algorithms to address the QoS requirements of Cognitive Analytics
* Mapping various Distributed Computing Technologies used in Cognitive Analytics
* 2 Examples/Case Studies (e.g. Operational Optimization, Next Best Customer Action, etc.) related to Cognitive Analytics elaborating the QoS, Patterns and Algorithms to be used, Choice of Technologies, etc.

Chapter 8 – Distributed Computing in Social Media Analytics

* QoS Requirements of Distributed Computing relevant to Social Media – High Data Volume, Very High Concurrent Request, Low latency Response, Fast Data Crunching, Graph Processing, Elasticity, Security, High Availability, etc.
* Mapping various Distributed Computing Patterns, Algorithms to address the QoS requirements of Social Media
* Mapping various Distributed Computing Technologies used in Social Media Analytics
* 2 Examples/Case Studies (e.g. Sentiment Analysis, Influencer Identifications etc.) related to Social Media elaborating the QoS, Patterns and Algorithms to be used, Choice of Technologies, etc.

Chapter 9 – Distributed Computing in Scientific Data Analytics

* QoS Requirements of Distributed Computing relevant to Scientific Data Analytics – High Data Volume for high resolution, Ability to implement Mathematical Simulations in iterative and Fast way, Security, High Availability, etc.
* Mapping various Distributed Computing Patterns, Algorithms to address the QoS requirements of Scientific Data Analytics
* Mapping various Distributed Computing Technologies used in Scientific Data Analytics
* 2 Examples/Case Studies (e.g. Climate Analytics, Genomics, etc.) related to Social Media elaborating the QoS, Patterns and Algorithms to be used, Choice of Technologies, etc.