

PRAIRIE VIEW A&M UNIVERSITY
COLLEGE OF ENGINEERING
PRAIRIE VIEW, TEXAS

**A SCALABLE AND DISTRIBUTED SEISMIC DATA TOOLKIT ON
BIG DATA ANALYTICS PLATFORM**

A Thesis
Submitted to the Graduate School
In Partial Fulfillment of the Requirements for
The Degree of

MASTER OF SCIENCE
IN
COMPUTER SCIENCE

Submitted By:

CHAO CHEN

Certificate of Approval:

(Advisor's name)
Chairman
Student Advisory Committee

(Department Head's name)
Department Head
Computer Science Department

(Dean's name)
Dean
College of Engineering

(Graduate School Dean's name)
Acting Dean
Graduate School

December 2016

ABSTRACT

A Scalable and Distributed Seismic Data Toolkit on Big Data Analytics Platform

December 2016

Chao Chen, B.S., southwest university of science and technology;

Chair of Advisory Committee: Dr. Lei Huang

Seismic volume is a basic data structure that has been widely used in geophysical computing, as well as in big data analytics applications of petroleum industry. It has been well studied in High Performance Computing (HPC) platforms. However, a little study has been done in big data analytics platforms. In this paper, we present an efficient Distributed Seismic Data Analytics Toolkit that works on the popular Apache Spark big data analytics platform. The toolkit supports different ways of seismic volume data distributions, repartition, transposing, access, and data parallelism with a variety of parallel execution templates. This work studies the performance characteristics of these seismic volume data operations with different configurations, current status of this big data analytics cloud at PVAMU as well as our research plan for future work. The scalability and parallelism are the main characteristics of this platform.

ACKNOWLEDGEMENTS

I would like to thank my advisor, professors, colleagues and friends who helped me a lot in the past two years. I cannot make it through and finish my thesis without your kind help.

First I want to thank my advisor Dr. Lei Huang, for giving me the chance to work in CloudComputing Lab and guiding my study and research in cloud computing. And thanks to Dr. Lin Li and Dr. Frizell for spending your time to refine my thesis. Special thanks to Mr Ted Clee, for sharing his deep knowledge of seismic data analytics.

I also appreciate all my colleagues and classmates worked in the lab in the past two years, for sharing so many interesting things from different cultures, and, of course, for helping and encouraging each other in research work.

Finally I want to say thank you to my friend Yuzhong Yan, for always being the guy helps me out at so many tough time in my life. I have learned so much from you not only as an engineer but also as a kind and wise man.

Best wishes to all of you and your families.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	vi
LIST OF TABLES	viii
1. INTRODUCTION	1
1.1 Background	1
1.2 Motivations	2
1.2.1 Big Data and Scalability	2
1.2.2 Complicated Workflow	3
1.3 Objectives	4
2. RELATED WORK	5
2.1 Researches in Petroleum Industry	5
2.2 Apache Hadoop and Spark	6
3. IMPLEMENTATION	9
3.1 Architecture	9
3.2 Interfaces and Functionalities	11
3.2.1 Seismic Volume Data Loading, Distribution and Saving	12
3.2.2 Volume Data Accessing	15
3.2.3 Volume Data 3D Transposing	16
3.2.4 Aggregation and Overlapping	17
3.2.5 User Defined Function Mapping	20
3.3 Application Utilities	21
3.3.1 Parallel Templates	22
3.3.2 Data Server and Remote Web Visualization	23
3.3.3 Web-based Workflow Platform	25

4. EXPERIMENTS, RESULTS AND ANALYSIS	28
4.1 3D Volume Transposing	28
4.1.1 Use Case	28
4.1.2 Statistics and Analysis	28
4.2 3D Stencil Application	33
4.2.1 Use Case	33
4.2.2 Statistics and Analysis	34
5. CONCLUSIONS AND FUTURE WORK	39
REFERENCES	40

LIST OF FIGURES

FIGURE	Page
1.1 Reflection Seismology [10] [3]	2
3.1 Software Stack of Seismic Data Analytics Platform	10
3.2 Framework of Seismic Data Analytics SDK	11
3.3 Main APIs of Seismic Data Analytics SDK	12
3.4 Reflection seismology Survey and Seismic Volume Data [15] [9]	13
3.5 Seismic Volume Data Distribution Flow	14
3.6 Seismic Volume Data Dimensions	15
3.7 The Indexing for Resolving 3D Transposing Problem	17
3.8 Default Distribution Fashion of Volume, planesPerMap=1, overlap=0	18
3.9 Aggregated Distribution of Volume, planesPerMap=3, overlap=0 . . .	19
3.10 Overlapped distribution of Volume, planesPerMap=1, overlap=1 . . .	20
3.11 Sample Code for Applying User-defined function in Parallel	21
3.12 Sample Code of Line Parallel Template	23
3.13 Sample Code of Sub-volume Parallel Template	24
3.14 Sample Code of Executing Sub-volume Parallel Template	24
3.15 The Framework of Remote Web Visualization Service	25
3.16 Visualization Web Interface	26
3.17 Workflow Web Interface	27
4.1 Transposing Experiment Statistics on Cluster with 288(576) cores. . .	29
4.2 Transposing Time on CPU Cores without Aggregation	29

4.3	Transposing Time on CPU Cores with Aggregation.	30
4.4	Transposing Time on Aggregated Planes(ppm: planesPerMap).	31
4.5	CPU utilization statistics from NMONVisualier.	32
4.6	DRAM utilization statistics from NMONVisualier.	32
4.7	The Transposing Performance Statistics on XSEDE Cluster	33
4.8	The Performance Scalability of Transposing on XSEDE Cluster . . .	33
4.9	The Data Distribution and Input of Overlaop Template	35
4.10	The Speedup of Parallel Template Codes with 28 Cores to Sequential Codes	37
4.11	The Speedup of Parallel Template Codes with 224 Cores to Sequential Codes	37
4.12	The Best Speedup of Parallel Templates for Stencil Computation . .	38

LIST OF TABLES

TABLE	Page
-------	------

CHAPTER 1

INTRODUCTION

Petroleum is a traditional industry where massive seismic data sets are acquired for exploration using land-based or marine surveys. Huge amount of seismic data has already been generated and processed for several decades in the industry, although there was no the big data concept at that time. High Performance Computing (HPC) has been heavily used in the industry to process the pre-stack seismic data in order to create 3D seismic property volumes for interpretation.

1.1 Background

Currently, reflection seismology (or seismic reflection) is the most important method used in the petroleum industry to estimate the properties of the Earth's subsurface from reflected seismic waves and explore geophysics using the principles of seismology. The complete processing flow of this method involves data acquisition, data processing, data interpretation and attributes analysis [8].

As shown in Figure 1.1, data acquisition is performed by seismic sources such as dynamite or air gun generate and spread out seismic waves, which are reflected back from encountered materials underground and then recorded by the receiving sensors. To get data-scientists-usable 2D/3D seismic dataset, the collected seismic wavelet needs to be pre-processed by data processing methods including deconvolution, common-midpoint (CMP) stacking and migration. After data processing, the seismic events are geometrically re-located in either space or time to the location the event occurred in subsurface and create a complete image of subsurface [8].

The goal of seismic data interpretation and attributes analysis is to locate the

potential petroleum reservoirs from processed seismic reflections map. It involves deep knowledge of seismic attributes, geophysics, and intensive collaborations between data scientists and geophysicists. This thesis focus on develop a scalable and distributed toolkit to facilitates seismic data attributes analytics.

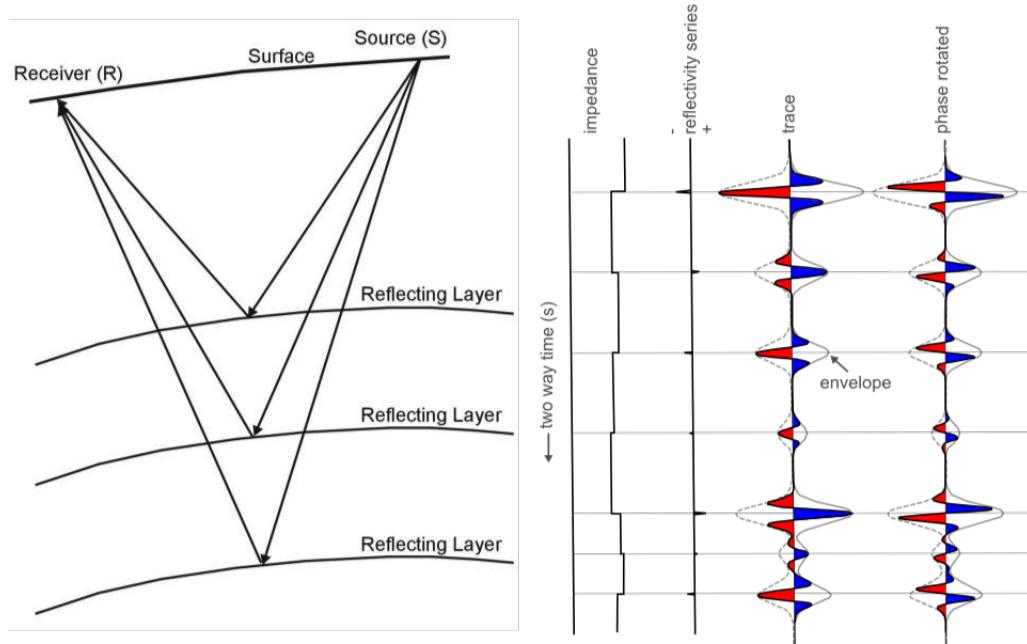


Figure 1.1: Reflection Seismology [10] [3]

1.2 Motivations

1.2.1 Big Data and Scalability

The emerging challenges in petroleum domain are the burst increase of the volume size of acquired data and high-speed streaming data from sensors in wells that need to be analyzed on time. Many types of captured data are used to create models and

images of the Earth’s structure and layers 5,000-35,000 feet below the surface and to describe activities around the wells themselves, such as machinery performance, oil flow rates and pressures. With approximately one million wells currently producing oil and/or gas in the United States alone, and many more gauges monitoring performance, this dataset is growing daily [19]. Moreover, not only the real world datasets, but also the dimensions and complexity of datasets themselves increase dramatically, such as 4D even 5D and high density seismic data. The traditional HPC solutions are able to improve the performance of many computation-intensive models, however, most processing models are also data-intensive which are still the bottleneck for the whole workflow.

In many the data- and technology-driven industries, big data analytics platforms and cloud computing technologies have made great progress in recent years toward meeting the requirements of exploring the valuable information from fast-growing data volumes and varieties. Hadoop and Spark are currently the most popular open source big data platforms that provide scalable solutions to store and process big data, which deliver dynamic, elastic and scalable data storage and analytics solutions to tackle the challenges in the big data era. These platforms allow data scientists to explore massive datasets and extract valuable information with scalable performance. Many technologies advances in statistics, machine learning, NoSQL database, and in-memory computing from both industry and academia continue to stimulate new innovations in the data analytics field.

1.2.2 Complicated Workflow

Since the seismic data processing flow involves deep knowledges of geophysics, data science and computer science, it requests intensive collaborations among the scientists and developers from many different fields. This situation has long been

another bottleneck for the whole system. For an instance, most scientists are using MATLAB code to build their models, which is usually hard to translate to MPI codes. In most cases, the program needs to be reconstructed and parallelized by software engineers. For this situation, it is already a big challenge to both geoscientist and software engineers to understand each other's work, not to mention to maintenance or optimize the huge amount of legacy code in this industry.

1.3 Objectives

Geophysicists need an ease-to-use and scalable platform that allows them to advance the seismic data exploration process, design more intelligent algorithms to increase the drilling success rate. By Incorporating the latest big data analytics technology with the geoscience domain knowledge will speed up their innovations in the exploration/interpretation phase.

Although there are some big data analytics platforms available in the market, they are not widely deployed in the petroleum industry since there is a big gap between these platforms and the special needs of the industry. For example, the seismic data formats are not supported by any of these platforms, and the machine learning algorithms need to be integrated with geology and geophysics knowledge to make the findings meaningful.

The objectives of the work are to develop a seismic data analytics software development kits (SDK) to enable geophysicists to easily leverage the latest big data analytics technology to improve the seismic data exploration.

CHAPTER 2

RELATED WORK

The most famous definition of big data, comes from Gartner analyst Doug Laney, specifies the 3Vs characteristics: volume, velocity and variety [21]. By which volume means the amount of data, velocity stands for the real-time speed of data in and out, and variety is the range of data types and sources. As mentioned in previous chapter, the burst increase of volume size, high-speed real-time streaming data from sensors, and various types of structured, unstructured and semi-structured data coming from different stages of seismic data processing all together matches the 3Vs definition. It determines seismic data is costly to store, access and manage in traditional methodology. Therefore, new technology should be adopted to address these problems appropriately.

2.1 Researches in Petroleum Industry

Although lots of motivation exist in petroleum companies to adopt big data solutions to improve efficiency and reduce cost, only a few of them have deployed big data solutions. This situation may due to some technique barriers such as lack of technology knowledge, big data solution are not applicable in some steps of traditional workflow, and the cost and risk to convert legacy software to new platform etc. Moreover, there are lots of concerns of business-wise, such as the cost of infrastructure, data security issue (business or political restrictions on data accessing).

In [16], it mentioned working on the application of big data analytics in the Oil & Gas industry is in the experimental stage. Only a handful companies have adopted Big Data in the field:

1. Chevron proof-of-concept using Hadoop (IBM BigInsights) for seismic data processing;
2. Shell piloting Hadoop in Amazon Virtual Private Cloud (Amazon VPC) for seismic sensor data;
3. Cloudera Seismic Hadoop project combining Seismic Unix with Apache Hadoop;
4. PointCross Seismic Data Server and Drilling Data Server using Hadoop and NoSQL;
5. University of Stavanger data acquisition performance study using Hadoop.

Much of the software innovation that is key to the digitization of big oil is happening at oil service contracting companies, such as Halliburton and Schlumberger, and big IT providers including Microsoft, IBM, Oracle and Open Source Projects [16].

2.2 Apache Hadoop and Spark

Since Google released its white paper series of big data processing technologies in 2004, the landscape of big data development has been changed profoundly. Many big data projects were inspired and developed based on MapReduce and Google File System framework. Hadoop and Spark is two most widely used open source big data solutions for many business and industry applications in recent years.

A year after the publication of MapReduce and Google File System framework, Doug Cutting and Mike Cafarella created an open source project Apache Hadoop, which has been used in many industries for a huge variety of tasks that all share the common theme of variety, volume and velocity of structured and unstructured

data [1]. The core of Apache Hadoop consists of a storage part, known as Hadoop Distributed File System (HDFS), and a processing part called MapReduce [5]. Distributed file system is fundamental to many main stream big data platforms as it is able to store data across number of storage devices of a cluster. Compare to traditional file system which holds sequential data on one device, HDFS provides far better scalability and support for parallel IO processing mode. Meanwhile, another open source big data project Apache Spark provides programmers with an application programming interface centered on a data structure called the resilient distributed dataset (RDD), a fault-tolerant collection of elements that can be operated on in parallel [11]. Since Spark itself does not provide distributed file system, it is usually installed on top of Hadoop, by which Spark could utilize HDFS interface to handle distributed data storage and access. The most different part between Hadoop and Spark is the parallel processing interface. MapReduce writes the result back to the storage after each reduction, while Spark utilizes RDD to handle most its operations and result in memory. This leads to up to 100 times performance improvement compare to Hadoop in certain circumstances [11]. Moreover, Spark provides more advanced features such as real-time streaming processing interface and machine-learning library.

However, all of these big data platforms are designed for general purpose applications and focus on distributing the data, computation and IO overloads. Most MapReduce based framework do not have, or only have limited communication mechanisms between different maps, which is important to resolve the data and logical dependence problems in many complex applications. When it comes to the field of seismic data processing and analysis, the problem is more complicated. Since most scientists and researchers in petroleum industry do not have big data related knowledge or even computer science background, how to hide parallelism from them

and let them easily deploy their works on new platform is a big challenge for all the researchers.

CHAPTER 3

IMPLEMENTATION

The main goal of Seismic Data Analytics SDK is to develop a scalable and distributed software development tool to enable scalable computation and analytics of seismic volume datasets. This chapter will present the software architecture and the main functionalities of SDK, as well as some utilities we have built for user to deploy their applications on this big data platform.

3.1 Architecture

Seismic Data Analytics SDK is built upon Apache Hadoop and Spark. Figure 3.1 shows the software stack of a workable seismic data analytics platform. As shown in this diagram, the gray part is the OS layer, the elements with green color stands for the infrastructure layer of this big data platform, and on top of that, the components with blue color is the SDK. At the bottom of infrastructure layer, there is Hadoop Distributed File System (HDFS) that stores the big seismic data files by utilizing the large number of local disks. The Cassandra as a NoSQL database is also used to store seismic data, intermediate results and meta data. YARN and Mesos are used for resources management. Apache Spark is the data distribution and parallel execution engine based on the innovative idea of Resilient Distributed DataSets (RDD) concept. MLlib is included in the Spark as the machine learning package to enable machine learning based data analytics algorithms. OpenCV is the widely used image processing package that is used to provide image processing capability. Breeze is the numerical processing package including linear algebra, signal processing, statistics, and other numerical computation and optimizations written in Scala.

We have developed the seismic data RDD on top of Spark as the base distributed seismic datasets to enable parallel operations and machine learning algorithms. Geophysicists and data scientists can use the Seismic Data Analytics SDK to develop their own algorithms and leverage the capability of Apache Spark provides, as well as image processing, numerical computation, and deep learning packages.

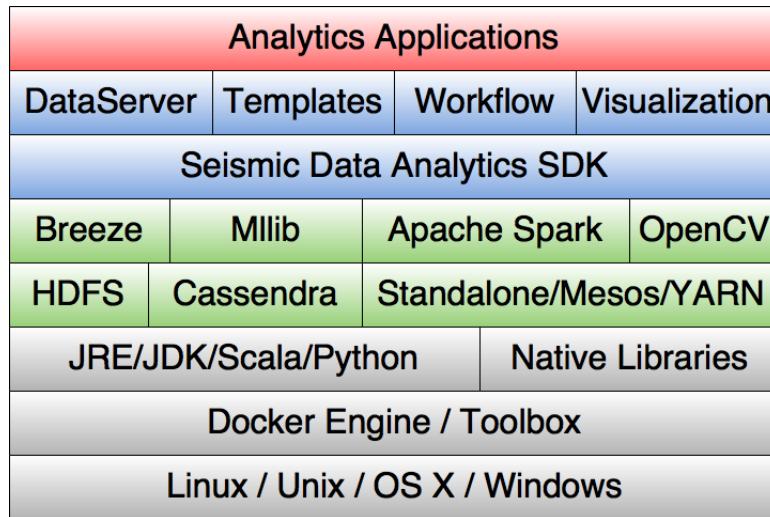


Figure 3.1: Software Stack of Seismic Data Analytics Platform

Figure 3.2 simplifies the development efforts for scalable and distributed computing and analytics of seismic datasets. It is built on top of the Apache Hadoop and Spark. The Hadoop provides a distributed file system(HDFS) and resource management system (YARN and Mesos), while Spark provides a high-level distributed data representation via Resilient Data Sets (RDD) and a data-parallelism execution engine. Seismic Data Analytics SDK provides configurable data distribution fashions for seismic volume data, as well as a configurable parallel execution interface to simplify the parallel programming efforts. Based on the functionality of SDK, we developed two useful utilities, parallel templates and data server, to facilitate SDK

for users to easily deploy their applications. Moreover, since Hadoop and Spark provide faults tolerance and task scheduling utilities, the toolkit inherits from them to provide fault tolerance and dynamic task scheduling for better reliability and task management.

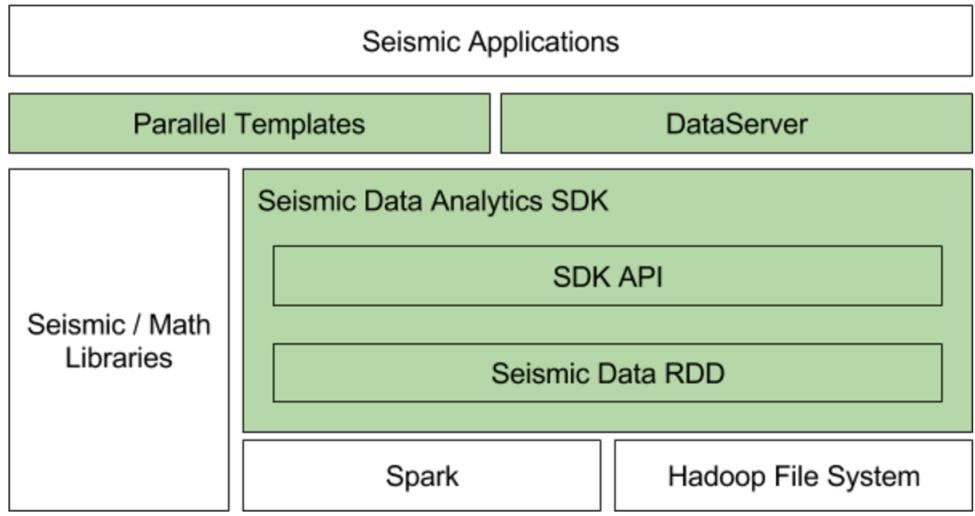


Figure 3.2: Framework of Seismic Data Analytics SDK

3.2 Interfaces and Functionalities

Figure 3.3 shows the main functionalities of Seismic Data Analytics SDK, including data loading/saving, configurable distribution, data accessing, 3D transpose and user-defined function mapping. It provides a single public class *SeismicVolume*, which integrates all the APIs of SDK. Developers are able to create *SeismicVolume* instances for specified seismic dataset, access data by configurable grain, perform 3D transposing and apply user function to the distributed data instance, and finally save the result to distributed file system through save API. The invalid input data format include 3D binary data and SEG-Y file [17] which is one of the most widely

used industrial standard format for seismic data.

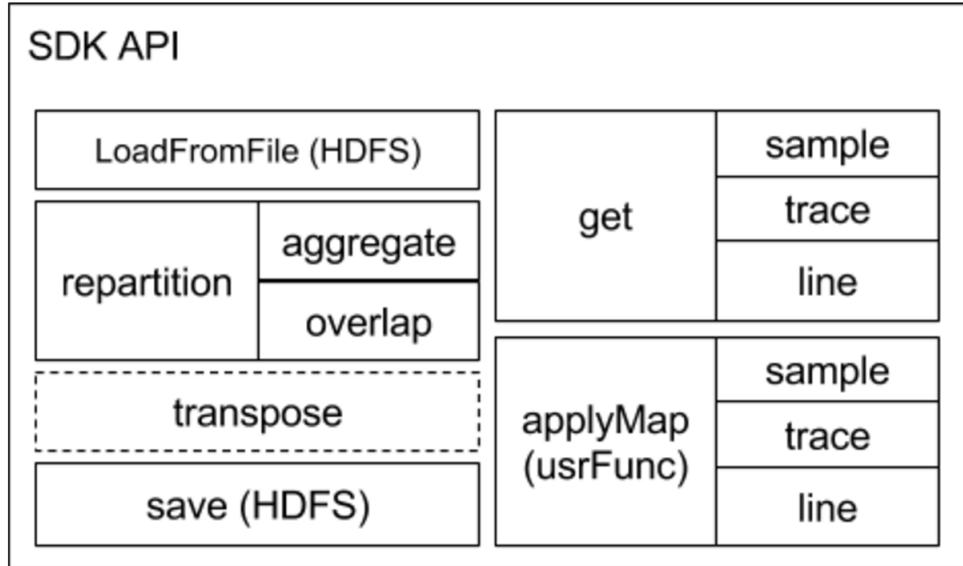


Figure 3.3: Main APIs of Seismic Data Analytics SDK

3.2.1 Seismic Volume Data Loading, Distribution and Saving

As mentioned in chapter 1, seismic 3D volume data is a collection of estimated property values of the Earth's subsurface, obtained through seismic reflection survey and organized in 3D spacing form. It's widely used in energy companies for geophysics analysis, which could conduct more accurate subsurface exploration and exploit. As shown in Figure 3.4, the seismic volume data is defined through 3 different directions in 3D space: inline, crossline and timeline. The industry usually stored the data slice by slice along crossline direction and each slice is a inline section, which is also the default data organization format of Seismic Data Analytics SDK. In this case, each inline slice is a single split of the whole dataset.

The public class *SeismicVolume* provides an API *loadFromFile()* to load seismic

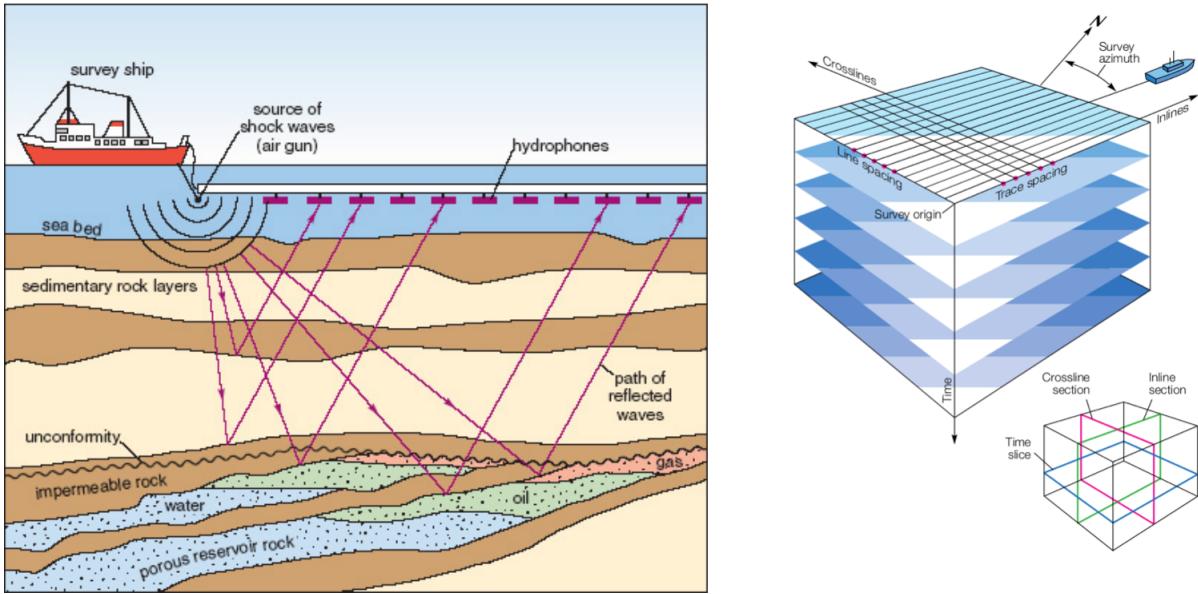


Figure 3.4: Reflection seismology Survey and Seismic Volume Data [15] [9]

data from HDFS and distribute them over Spark RDD according to user's configurations. This API generates a *SeismicVolume* instance which contains a SeismicRDD with float/byte as internal binary data types. The SeismicRDD is a derived class from Spark RDD class with a variety of distributed fashions of seismic volume data. In addition, it also provides some optional parameters for advanced users who have already familiar with distributed system to specify the advanced data distribution fashions.

Figure 3.5 shows the flow of distributing an seismic volume file through the Hadoop filesystem and Spark RDD. It assumes the file has already been uploaded to Hadoop file system, which is able to support the distributed IO accessing for Spark to load data in parallel. After user get the *SeismicVolume* instance of a specified file, all Spark RDD operations can be applied to the inside SeismicRDD object. Utilizing the RDD methods provided by Spark, developer could perform various of data operations and calculations on the dataset in parallel. By default, SDK distributes the

volume in inline format slice by slice, which means each partition contains one single inline slice. The distribution direction could also be configured to crossline/timeline by given parameters in other APIs, this function will be present in the following Volume Data 3D Transposing section. User could configure the slices count of each distribution to manage the data distribution grain, thus to efficiently tuning the performance of applications.

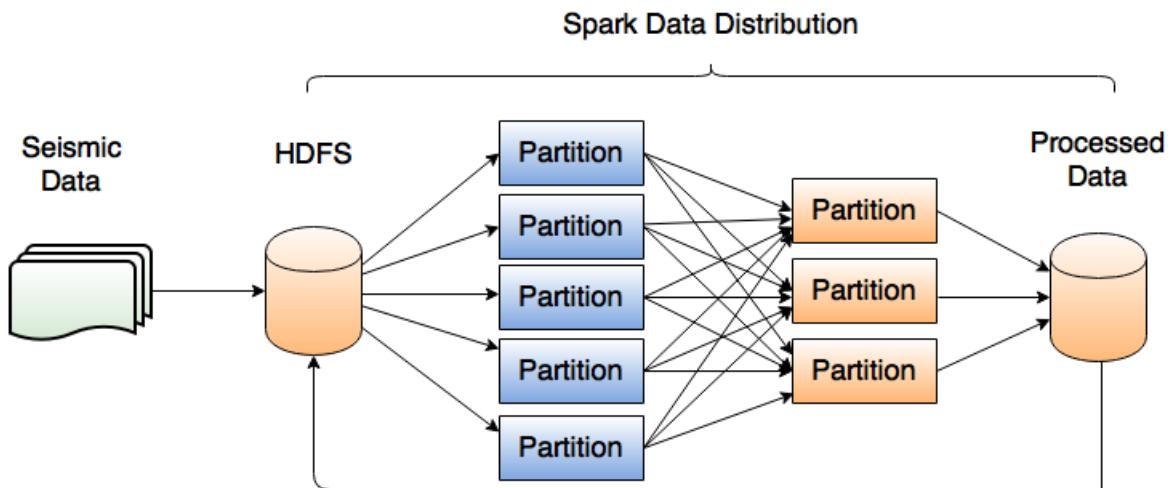


Figure 3.5: Seismic Volume Data Distribution Flow

SeismicVolume class also provides `save()` API to allow user to store the data of *SeismicRDD* back to the HDFS. Although this operation is not recommended since it could introduce performance issues caused by data shuffling, it is still necessary when user need to backup the runtime data or apply the runtime data to traditional sequential workflows.

3.2.2 Volume Data Accessing

Seismic Data Analytics SDK allows users to access any slice/trace/sample data in any direction of the volume through *SeismicVolume* APIs $\text{getLine}(\text{direction:} \text{Int}, \text{idx:} \text{Int}) : \text{Array}[T]$, $\text{getTrace}(\text{dir:} \text{Int}, \text{i:} \text{Int}, \text{j:} \text{Int}) : \text{Array}[T]$ and $\text{getSample}(\text{i:} \text{Int}, \text{j:} \text{Int}, \text{k:} \text{Int}) : T$. For the convenience of addressing the data organization, we defined the three dimension of seismic volume as I (Dimension of inline slices), J (Dimension of crossline slices) and K (Dimension of timeline slices), as shown in Figure 3.6 and Figure 3.7. Users can specify any one of I, J and K directions to access data for visualization or computation purpose. Since Spark does not provide arbitrary data access, the APIs used IndexedRDD, a more efficient key-value management than native Spark API [6], to implement and speedup queries of slice/trace/sample data from any distributions to the master node.

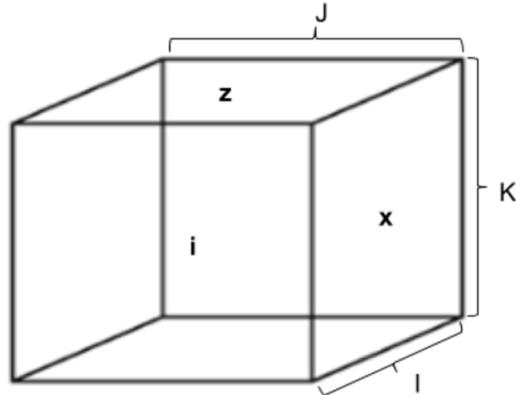


Figure 3.6: Seismic Volume Data Dimensions

3.2.3 Volume Data 3D Transposing

Since the volume data could only be stored in file following one specific direction(I, J or K), developer could not access slices of the other two directions directly. In traditional solutions, if user needs data in other directions, organized as cross-line slice or time-line slice, the data fetching program must perform lots of seek operations between different file offsets to collect the data of a single cross-line or time-line slice. This tedious procedure will dramatically slow down the whole software performance. To resolve this problem and to achieve reasonable performance, SDK handles the transposing of the 3D volume data inside the *SeismicVolume* APIs and caches all three directions format SeismicRDD in *SeismicVolume*.

To explain the implementation clearly, we denote the seismic volume data as shown in Figure 3.6, in which i means I slice, x means J slice and z stands for K slice. The data is stored in iSlices format by default. To resolve the transposing problem in each distribution evenly, we split the volume to I of iSlices and distribute them over SeismicRDD. Each iSlice is a 2D matrix. As shown in Figure 3.7, each iSlice matrix consists of J of iTraces which have the length of K. An iSlice matrix could be iterated iTrace by iTrace. Since in 3D spacing, each iTrace is also the trace of xSlice, for example, the iTraces(0) is the xTrace of the 0th xSlice, the iTraces(1) is the xTrace of 1st xSlice, etc. Thus, we implement a map function to index all iTraces of the volume. The new index is combined by index of iTrace and index of iSlice. After indexing the map, we got a volume RDD with new (iTraceIndex)(iSliceIndex) index as the key, trace data as the value. As shown in Figure 3.7, to get a xSlice, we need group all the traces with the same iTraceIndex by utilizing the *group* operations of Spark RDD. After grouping, the xSlices data collection is generated in the new SeismicRDD distribution map. To organize them as a xSlices volume, all we need

to do is sorting them by iTraceIndex. So far, the data in requested direction has already been stored in the *SeismicVolume* instance, therefore user could access and manipulate seismic data in any direction by specifying the direction parameter of related *SeismicVolume* APIs.

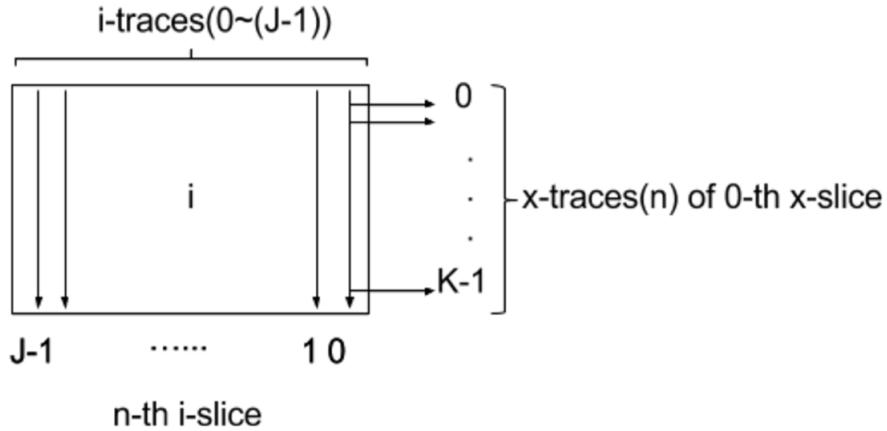


Figure 3.7: The Indexing for Resolving 3D Transposing Problem

3.2.4 Aggregation and Overlapping

By default, as shown in Figure DefDist, SDK distributes the volume in one specific direction slice by slice. In this case, each slice is a single split of the whole dataset. However, it will cause performance problems for some applications if we only support one distribution. For an instance, the transposing solution as mentioned in previous section needs to do lots of data exchanges between different data splits for the *group* operation to shuffle the dataset to expected arrangement. Since data shuffle in Spark RDD relies on lots of physical storage access and network transmissions in each worker node, it will become the bottleneck of the transposing performance if the

number of partitions is too big. Obviously, to speedup the transposing operation, user need to reduce the number of partitions thus to reduce the data communications between worker nodes. Therefore, user need a way to configure the data split size and partition number so they can tuning the performance of their programs.

Utilizing the *Seismic Volume API repartition(planesPerMap:Int,overlapPlanes:Int):SeismicVolume* we can change the distribution layout to the aggregated and overlapped fashions, as shown in Figure 3.9 and Figure 3.10.

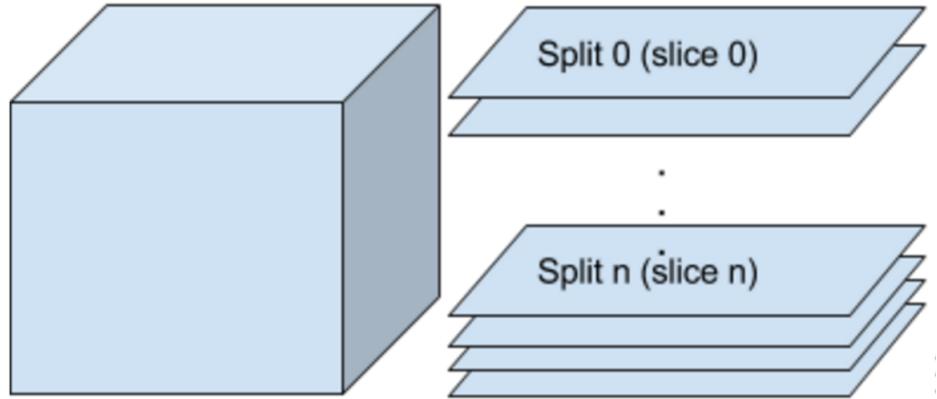


Figure 3.8: Default Distribution Fashion of Volume, $\text{planesPerMap}=1$, $\text{overlap}=0$

As shown in Figure 3.9, aggregated data distribution could be achieved by utilizing the RDD group operations we mentioned in Volume Data 3D transposing section. The solution of this problem is to re-indexing all the slices in parallel through RDD map function by arranging an unique key to multiple data splits, then use RDD *groupByKey()* API to repartition the dataset. An aggregated dataset could have multiple slices in a single data split.

This method not only lets user change the size of distributed splits, more powerfully, it allows developer to set the overlapped data areas between splits and to access

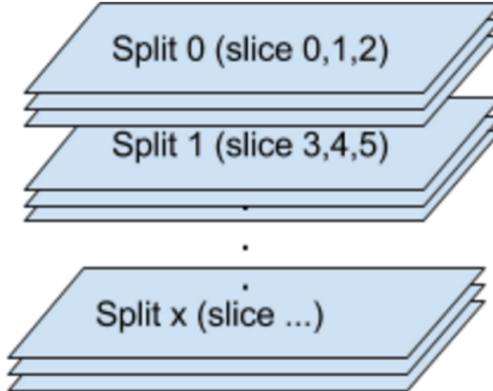


Figure 3.9: Aggregated Distribution of Volume, planesPerMap=3, overlap=0

the overlapped parts in each split. Practically, some applications may have strong data dependences in their logic which is very hard to parallelize the solutions. For example, the stencil computation needs lots of data communications between neighbor units in each step, which is impossible to run it in parallel with other MapReduce frameworks.

To resolve this problem, `repartition()` API generates the head and tail boundaries RDD on top of the aggregated SeismicRDD and indexing them according to the related partition numbers. Finally, the boundaries are appended to each related data split through RDD `zip()` API to generate a new *SeismicVolume* instance which contains overlapped SeismicRDD.

Therefore, this method is not only capable of tuning the performance of distributed tasks, but also simplifies the stencil-style computation requiring neighbor communication, which is not easy to parallelize in MapReduce programming mode. We have developed a more complicated 3D stencil use case with 3D spacing overlapping which will be presented in detail in following Experiments chapter. It could be used for resolving seismic 3D attributes computation problem.

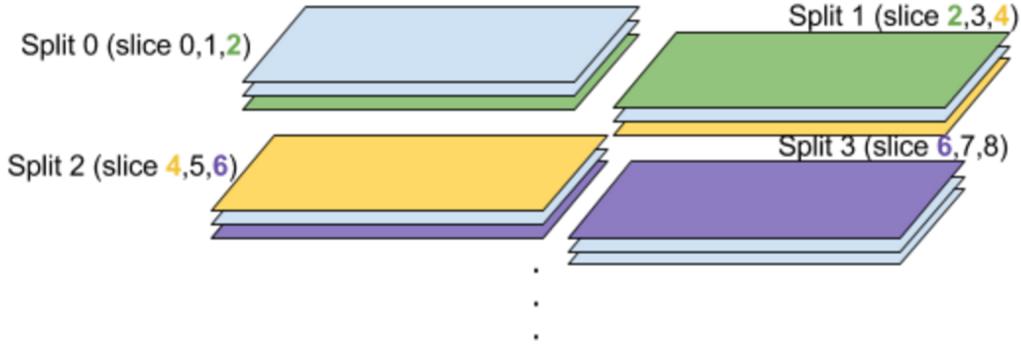


Figure 3.10: Overlapped distribution of Volume, planesPerMap=1, overlap=1

3.2.5 User Defined Function Mapping

As mentioned in Introduction chapter, one of the project objective is to provide an easy-to-use solution for geophysicists or data scientists to apply their programs on big data platform without concerning the parallelism and code reconstruction. To achieve this, *SeismicVolume* provides *applyMap()* API to allow developer to apply user-defined functions on any direction of the SeismicRDD in parallel.

The prototype of this API is *applyMap(direction:Int, f:(T-U))*. The first parameter *direction* indicates the direction that user would like to apply the function on. The other parameter *f:(T-U)* is a standard spark RDD key-value pairs operation callback function, which feeds the function distributed volume data with key-value forms in parallel. The data length in each key-value function depends on the specified distribution parameters of the target *SeismicVolume*. User could apply any operation or computations on the given input data, and an output key-value pairs is required for the return value. This callback function will be executed along with the related data distribution in parallel. After execution, it will generate a new *SeismicVolume* object containing the processed and distributed dataset output by

user-defined function. Figure 3.11 shows the example code for users to apply their function.

```

def fftFunc[T:ClassTag](e:(Long,Array[T])):(Long,Array[T]) = {
    val k = e._1;
    val v = e._2;

    var w = kVar;
    var h = jVar;
    if (dir == 2) {
        w = kVar;
        h = iVar;
    }
    else if (dir == 4) {
        w = jVar;
        h = iVar;
    }

    val planeSize = w * h;
    val num = v.length/(planeSize);
    var res = new Array[T](0);
    for(i <- 0 until num) {
        val line = new DenseMatrix(h, w, v.slice(planeSize*i,planeSize*i+planeSize));
        val lined= line.mapValues(a=>a.asInstanceOf[Double]);
        val fft = fourierTr(lined);
        val tmp = fft.mapValues(a=>a.abs.asInstanceOf[T]).data;
        res = res ++ tmp;
    }
    (k,res)
}

```

Figure 3.11: Sample Code for Applying User-defined function in Parallel

3.3 Application Utilities

Data distribution plays an important role in distributed parallel programs to achieve scalable performance. However, most scientists and researchers in petroleum industry do not have enough big data background knowledge to support them to convert their work to Spark applications. To overcome the usability barriers, we developed some utilities on top of Seismic Data Analytics SDK to simplify the way for deploying general sequential applications on Spark-Hadoop platform, which include several parallel templates for user-defined procedure, a general purpose data server

for remote data access, as well as the web interfaces for remote data visualization and user-defined workflow. With these tools, developers could facilitate their works with big data platform in minimal efforts without concerning all the parallelism details.

3.3.1 Parallel Templates

Based on Seismic Data Analytics SDK APIs, we developed several parallel templates to make SDK be easily used by domain algorithm designer other than computer scientists. These templates defines the data distributions and parallel computation so that users can simply select the right templates for their algorithms without handling the data distribution and parallelism details. Three templates currently include: Trace, Line and Sub-volume. Each template can handle one or more volumes, and will output one or more volumes. Trace template is simple, in which the input is a 2D array (dimension 1 for number of volumes and dimension 2 for 1D trace data), and output is also a 2D array. Line template defines a 3D array as input and a 3D array as output respectively (dimension 1 for number of volumes and dimension 2 for 2D slice data, line is the petroleum terminology). Sub-volume template is a powerful solution to handle data 3D data distribution with overlaps, in which both input and output are 4D array (dimension 1 for number of volumes and dimension 2 for 3D overlapped sub-volume data). The Sub-volume template outputs data without any overlapping. Users can specify parameters about how to distribute data as well as the overlapped areas.

Figure 3.12 shows an example of the Line Template and Figure 3.13 shows the example of Sub-volume Template. Both of them are straight forward and require very few knowledge about parallel computing. The only thing user should be aware of is the distribution grain of input/output data, which is specified in the distribution parameter when user creates the template instance. As shown in Figure 3.14, it

demonstrates the execution of an Sub-volume Template instance, in which the distributed input sub-volume size in dimension I and J (for each parallel *proc()* callback function) is set to 26 x 21(The sub-volume size in dimension K is always the complete length of iTrace/xTrace, which is the minimal unit for seismic computation), and the overlapping size is 4 in both I and J direction (Each data split extends to 30 x 25 in I x J).

```

class LineApp(override val inputList: Array[SeismicVolume[Float]], override val distPrm: DistParams)
  extends SeismicAppLineTmpl[Float, Float](inputList, distPrm)
{
  override def proc(input: Array[Array[Array[Float]]]): Array[Array[Array[Float]]] = {
    val len = input.length;
    val output = new Array[Array[Array[Float]]](len);
    for(i <- 0 until len) {
      output(i) = input(len-1-i);
    }
    output;
  }
}

```

Figure 3.12: Sample Code of Line Parallel Template

3.3.2 Data Server and Remote Web Visualization

In petroleum industry, an important application scenario is data visualization, which allows user to browse and analyze seismology features of seismic data in 3D spacing. With visualization tools, computer renders the seismic data to 3D graphic views and allows user to browse and manipulate the graph along any direction, which inspired the geophysicists and data scientists to develop various of useful models. However, traditional visualization tools in industry are only capable of handling small datasets, or render only few segments of the big datasets at a time. The performance of big dataset visualization has long been the critical bottleneck of regular workflow of industry.

```

class TraceOlpFilter(override val inputList: Array[SeismicVolume[Float]], override val distPrm: OverlapDistParams)
  extends SeismicAppTraceOverlapTmpl[Float, Float](inputList, distPrm)
{
  override def proc(inputData: Array[Array[Array[Array[Float]]]],
                    olpWindow: TraceOverlapWindow): Array[Array[Array[Array[Float]]]] = {

    val len = inputData.length;
    val outData = new Array[Array[Array[Array[Float]]]](len);

    for (i <- 0 until len) {
      val centerVolume = new Array[Array[Array[Float]]](olpWindow.olpCenterI);

      /*extract center data from olp window*/
      for (o1 <- olpWindow.olpHeadI until (olpWindow.olpHeadI + olpWindow.olpCenterI)) {
        val centerSlice = new Array[Array[Float]](olpWindow.olpCenterJ);

        for (o2 <- olpWindow.olpHeadJ until (olpWindow.olpHeadJ + olpWindow.olpCenterJ)) {
          val trace = inputData(i)(o1)(o2);
          centerSlice(o2 - olpWindow.olpHeadJ) = trace;
        }
        centerVolume(o1 - olpWindow.olpHeadI) = centerSlice;
      }
      outData(i) = centerVolume;
    }
    outData;
  }
}

```

Figure 3.13: Sample Code of Sub-volume Parallel Template

```

/*create an overlap layout:
I x J: 26 x 21 in center and with 4 traces overlapping in each of I and J direction
*/
val olpSpec = new OverlapDistParams(26, 21, 4, 4);

val olpTmpl = new TraceOlpFilter(invols, olpSpec);

// Run the filter to produce filtered volume
val fvol = olpTmpl.exec(INLINE);
val res = fvol(0);

```

Figure 3.14: Sample Code of Executing Sub-volume Parallel Template

To resolve this problem, we developed a web-based remote data visualization service which is able to load and render the seismic data for 3D visualization in real-time. Figure 3.15 shows the framework of this service. This solution is developed on top of Seismic Data Analytics SDK APIs and FEI Digital Rock Visualization Packages. We design and implement the DataServer service on top of Seismic Data Analytics SDK, which could load big seismic datasets from HDFS, transpose and store them to SeismicRDD in backends then feed the requested data in any direction

to FEI visualization service in real-time. The FEI Digital Rock Visualization package is developed by FEI, the company which has been focus on featuring Digital Rock technology and solutions many years [4]. Finally, the rendered data view is presented through web interface and allows user to browse and manipulate in any direction. With this solution, user do not need to install complicated visualization tools and packages since the platform handles everything on server-side. More importantly, the performance of rendering big seismic data is improved and scalable after facilitated by Seismic Data Analytics SDK. Figure 3.16 show s the web interface of this solution.

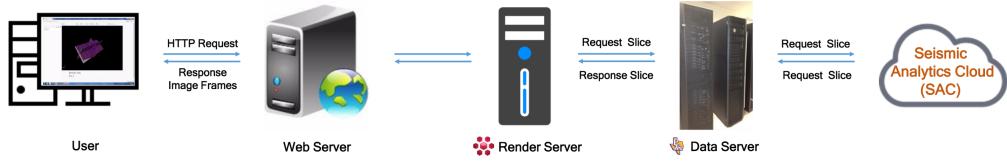


Figure 3.15: The Framework of Remote Web Visualization Service

3.3.3 Web-based Workflow Platform

Another service we developed is a web-based workflow platform, which provides a friendly web interface to make cloud platform easy to use without programming, with which users could create workflow with drag and drop, could run the created workflow and check the results through visualization model. The Workflow Web Service is implemented on an open source project Clowdflows [2], which provides a Django based framework to develop a customized widget and to manage widgets conveniently. As a free and open source web application framework written in Python, Django follows the Model, View and Controller (MVC) architectural pattern, so it is suitable for interact with both cloud service and web client.

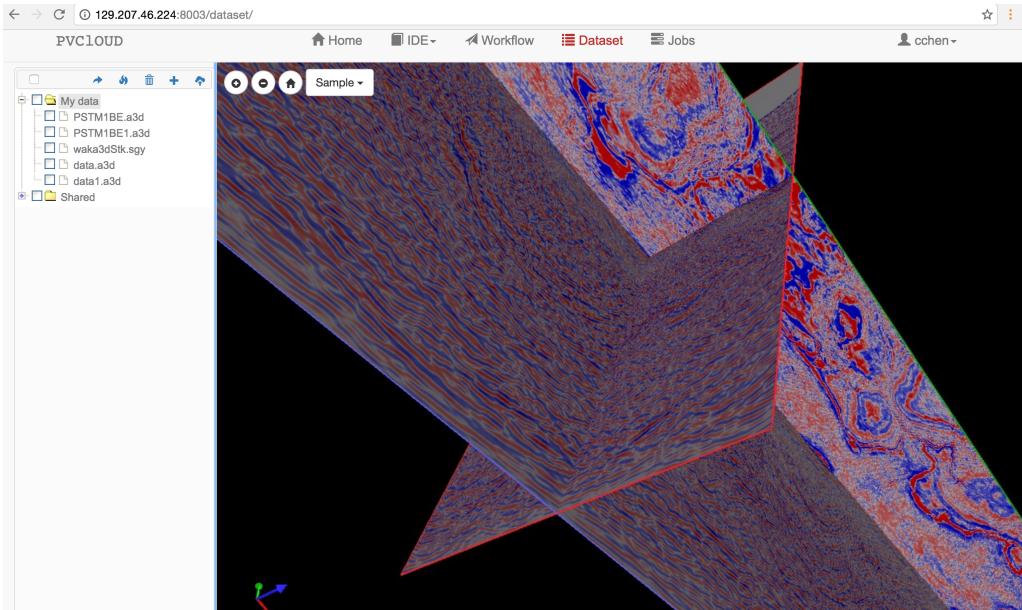


Figure 3.16: Visualization Web Interface

The client side view of workflow is shown as Figure 3.17. Users could select widgets and specify seismic data files as input to construct and customize their workflows. Each widget in the workflow view is an independent application component which could be implemented on top of Seismic Data Analytics SDK. A widget acquires at least one port as input or output thus multiple widgets are able to be combined to a workflow by multiple pipelines which connect the ports of all widgets. Each pipeline in the workflow view indicates a data communication, which transports data between widgets and drives the execution of whole workflow. The workflow framework implemented by Python code includes Django views (GUI), Django models (widgets data management) and topological sorting algorithm (connections check). The data files are stored in HDFS of the cloud platform and could be browsed and selected from the navigation trees on the editor page.

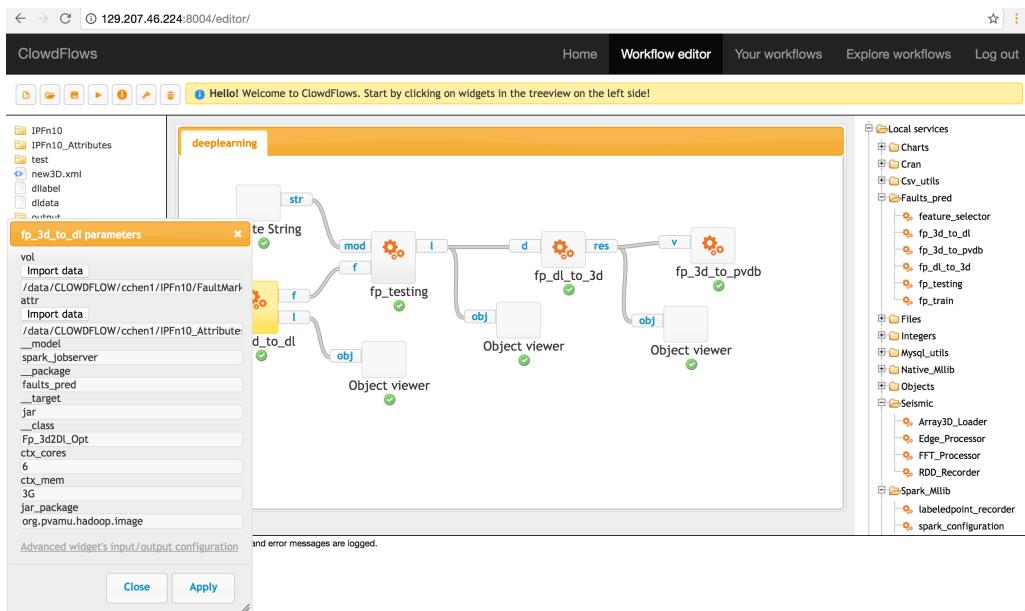


Figure 3.17: Workflow Web Interface

CHAPTER 4

EXPERIMENTS, RESULTS AND ANALYSIS

To verify and demonstrate the scalability that the user application could achieve through utilizing DMAT, we conducted a series of experiments, including data transposing and 3D stencil calculation, an overlapping-calculation application.

4.1 3D Volume Transposing

4.1.1 Use Case

This section presents the experiments of the 3D transposing problem mentioned in previous chapter. The dataset we used for transposing (from I to J direction) experiment is a 300GB seismic 3D volumetric data, which is 31017 x 97223 x 31 in I x J x K direction with float data type. We design the experiment to verify the performance of transposing is scalable and is mainly affected by the data distribution configuration and the amount of available hardware resource(the total count of cores).

4.1.2 Statistics and Analysis

4.1.2.0.1 Scalability to the Number of CPU Cores We conducted the transposing experiment on a cluster with 24 nodes, each node has 12 cores(24 cores in Hyper-threading) and 48GB available DRAM. The total CPU cores is 288(576 in Hyper-threading). Figure 4.1 shows the performance metrics of this experiment. The x in *Transpose[x]* stands for the aggregation parameter we applied on the dataset. As mentioned in previous chapter, the value of aggregation parameter x determines the size of each data split.

From the metrics, we got two bar chart which shows the time consuming for the transposing of aggregated and non-aggregated data. It clearly demonstrates that the

Cores	LoadFromFile(sec)	Transpose[1](sec)	Transpose[16](sec)	Transpose[64](sec)	Transpose[128](sec)
36	711.02	30246.483	604.083	444.867	433.9
72	531.849	15941.077	334.019	283.037	263.478
144	445.16	8865.005	210.677	178.212	168.96
288	404.325	5134.525	211.871	136.755	142.68
576	398.696	3914.722	179.026	131.292	141.675

Figure 4.1: Transposing Experiment Statistics on Cluster with 288(576) cores.

performance scalability on dimension of the count of CPU cores is promised, as shown in Figure 4.2 and Figure 4.3. The performance difference between the transposing on aggregated and non-aggregated seismic volume data will be addressed in the following paragraph.

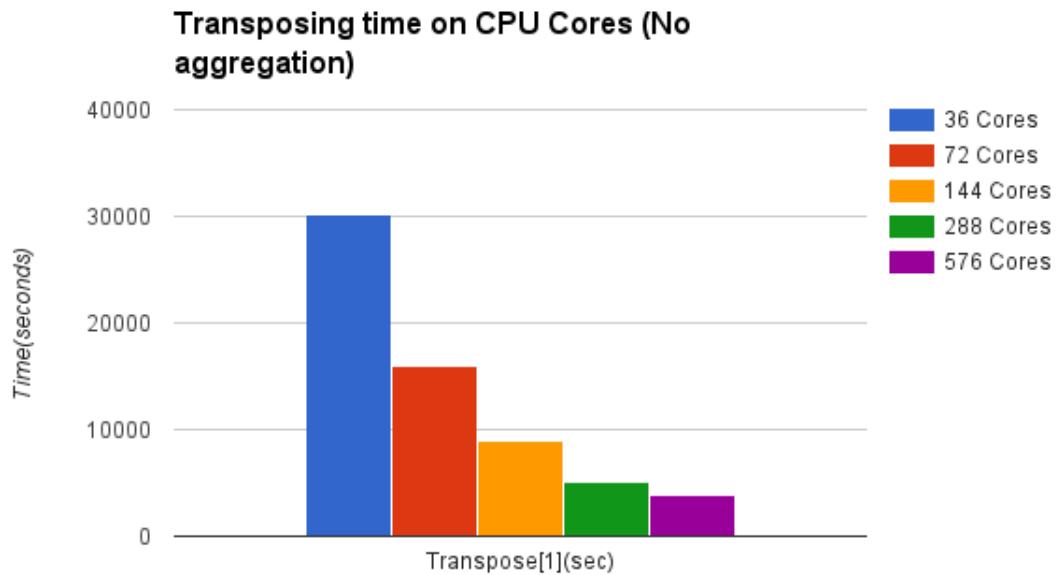


Figure 4.2: Transposing Time on CPU Cores without Aggregation

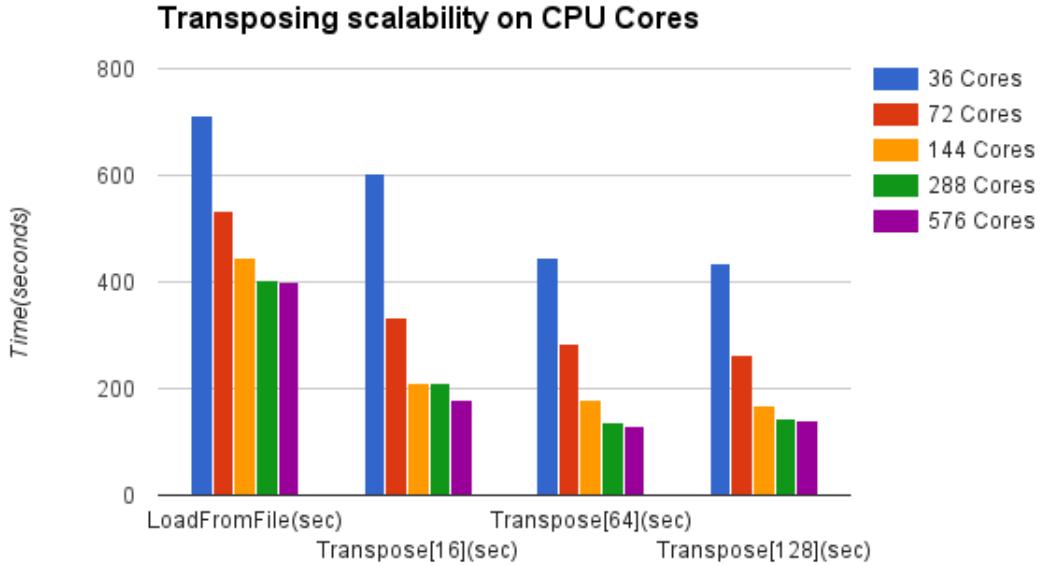


Figure 4.3: Transposing Time on CPU Cores with Aggregation.

4.1.2.0.2 Scalability to the Fashion of Data Distributions There is another important factor, the data distribution fashion, will also affect the performance of transposing. As we mentioned in previous section, the transposing program will perform some shuffle operation on volumetric RDD. Shuffle operations [12] is Spark's mechanism for re-distributing data so that it could change the data organization across partitions. This typically involves copying data across executors and nodes, which makes the shuffle a complex and costly operation. To reduce the shuffle costs, we could aggregate the SeismicRDD data splits to reduce the number of partitions thus to decrease the amount of data need to exchange during 3D transposing on each data split. Figure 4.4 shows the performance improvement when we increasing the aggregated planes number (x in $\text{Transpose}[x]$) of each data distribution.

However, the curve of performance trends to flat when the aggregation increases

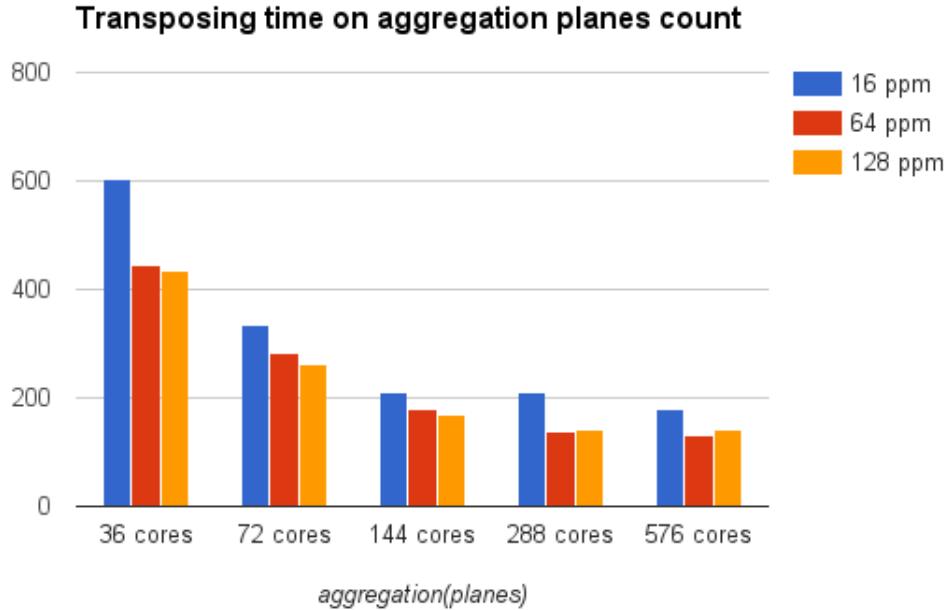


Figure 4.4: Transposing Time on Aggregated Planes(ppm: planesPerMap).

to a higher level. That is caused by the performance trade-off when reducing the distribution partitions, which will lead to a lower cluster CPU cores utilizing rate. Therefore, to achieve a reasonable performance for transposing operation, developer needs to considerate the trade-off between distribution configurations and data shuffling.

Figure 4.5 and Figure 4.6 show the CPU utilization and memory usage of each node when the transposing experiment was running on the cluster. These system statistic visualization views were generated by a free software NMONVisualizer [7], which is a Java GUI tool for analyzing Nmon performance files.

4.1.2.0.3 Cross-Verification on the Third-party Cloud Platform: XSEDE Cluster
 To further verify the scalability, we also setup the same experiment on the XSEDE supercomputing cluster [14]. We request 44 nodes from XSEDE cluster, which has

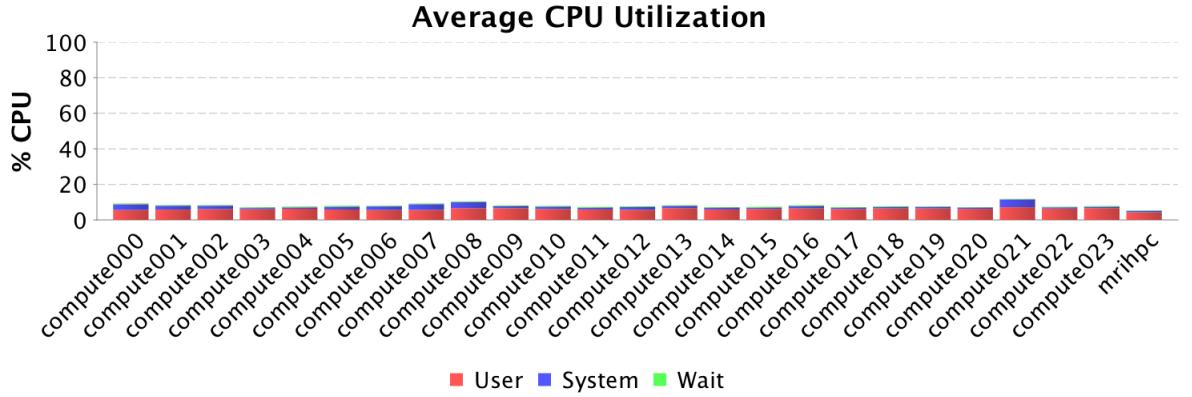


Figure 4.5: CPU utilization statistics from NMONVisualier.

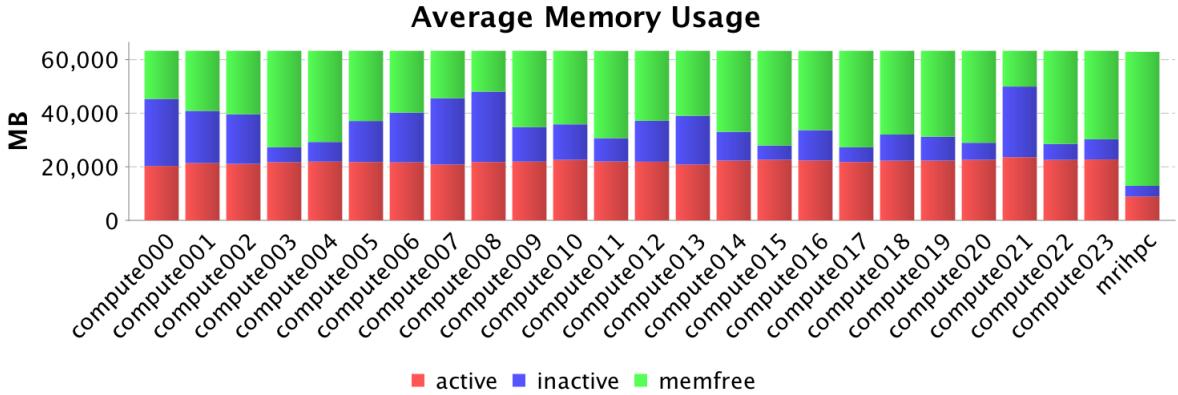


Figure 4.6: DRAM utilization statistics from NMONVisualier.

12 cores in each node. Figure 4.7 shows the result statistics of this experiment. As shown in Figure 4.8, we conduct the experiment to test the scalability of case transposing the same volumetric data from I to J direction, aggregation is 16 planes, and the result also verified the scalability of this distributed application.

Cores	Transpose Time(sec)
192 cores	494.434
240 cores	431.174
288 cores	417.077
336 cores	331.337
384 cores	286.763
432 cores	268.772
480 cores	259.086
528 cores	238.074

Figure 4.7: The Transposing Performance Statistics on XSEDE Cluster

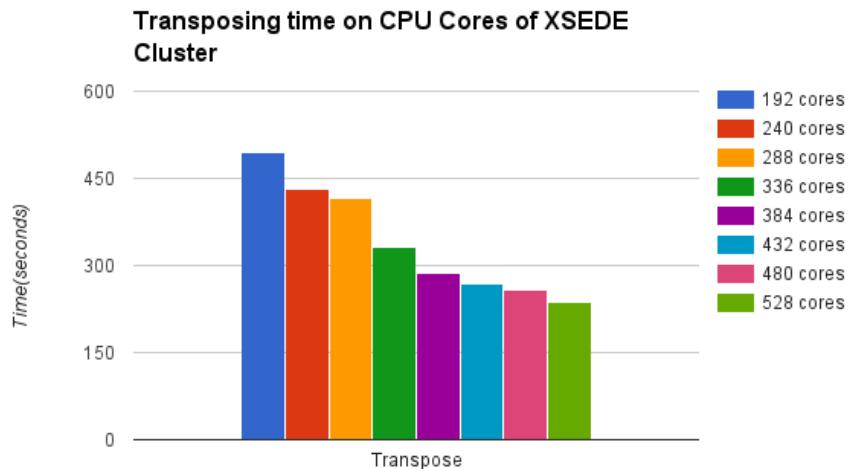


Figure 4.8: The Performance Scalability of Transposing on XSEDE Cluster

4.2 3D Stencil Application

4.2.1 Use Case

Stencil computations are most commonly used in context of scientific and engineering applications such as signal and image processing, computer simulations etc. Stencil itself represents an iterative kernel that updates array elements according to fixed patterns [13]. The optimization on stencil computation has been well studied in

[20],[22] and [18]. However, most of these optimizations focus on single node implementation with GPU or multi-core CPUs. In [25], [24] and [23], the authors provide a parallel implementation with Spark RDD, which gives a scalable solution for big data that can not host on a single node. In this experiment, we use the defined parallel templates in DMAT to implement stencil computations, and test their performance as well as scalability.

4.2.2 Statistics and Analysis

The dataset we choose to conduct our experiment is called Penobscot dataset, which is actual seismic image data with 3D dimension size 600x481x1501. The cluster consists of 8 nodes, in which one is management node and other 7 nodes are computation nodes. Each node was equipped with Intel Xeon E5-2690 Sandy Bridge CPU (2.9 GHz, 16 Cores or 32 Cores with Hyper-threading support), 128GB DDR3 memory and are inter-connected with 1GB ethernet. JDK 1.8.0_40, Hadoop 2.5.1 and Spark 1.6.1 are used for compiling and running applications. Wall clock is used to get the running time, and Nmon/Spark Web UI are used for performance analysis.

For the algorithm, we use a variant of Jacobi iteration, which uses a 3D subvolume with dimension size of 3x3x3 as input, and in the computation, each new output value at (i, j, k) is the average value of 26 surrounding samples plus itself. In the case of 3x3x3 subvolume, the overlap area is 1. For the sequential codes, we just split the big 3D data file into small partition and each partition includes several 2D planes (the overlap between partition is one 2D plane), and then use 3 nested loops to compute the average value. For the parallel codes using DMAT, we use the overlapped template, which specify parameters both in I and J directions. Different configurations of cores and numbers of planes in each partition are set to check performance and scalability. The numbers of cores (28, 56, 112, 224) are used for

each test case respectively to verify the scalability of parallel codes. Within each configuration of cores, we use different combinations of dimension size (1, 2, 4, 8) in I and J directions. For instance, I(4) and J(2) mean that the dimension size of input subvolume 6x4, which comes from $(4+2*1)x(2+2*1)$ in the case of overlap is 1, which is shown in Figure 4.9

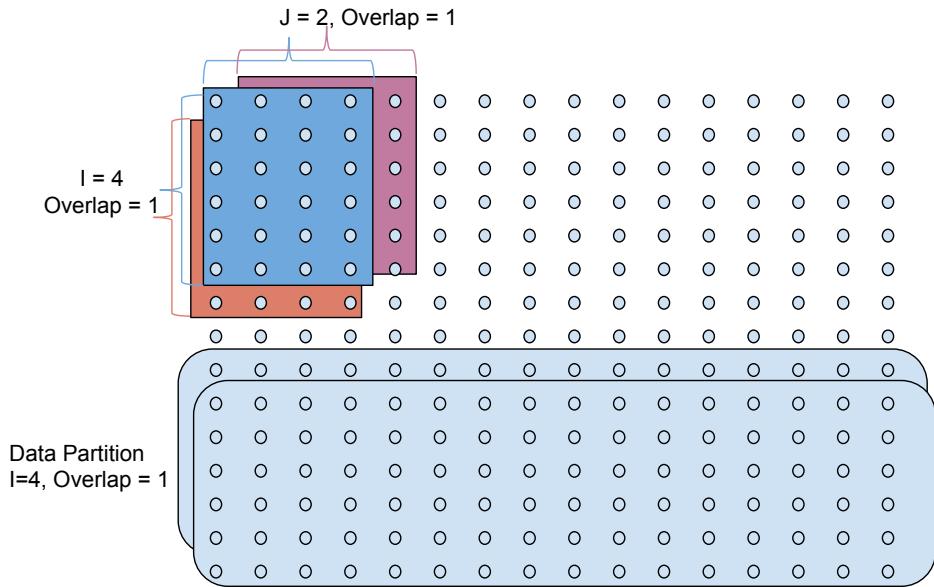


Figure 4.9: The Data Distribution and Input of Overlap Template

Figure 4.10 and Figure 4.11 show the speed up of parallel codes on Spark with 28 cores and 224 cores to the sequential codes respectively. From these two figures, the changes on number of J have little impact on the performance, because it does not change the number of planes in each partition and the number of partitions. In the template implementation, two nested loops are used for feeding the input of each

stencil kernel. However, the change on number of I tells how the SDK distributes the data and the number of partitions, thus will determine how many tasks are need to finish that stage of in the job, and each task need to be assinged one thread or one core to undertake the computation. In the case that size of I is 1, it gets the best speed up in all test cases. It seems to be abnormal that the performance decreases from 1 plane of I to 2 planes of I . Increasing I will enlarge the size of each partition as shown in Figure 4.9, however, the amount of computations keeps constant. In this stencil computation, we iterate several times to reach the balance. At the beginning of each loop, the data need to be repartitioned based on the overlap parameter, since the latest edge data need to be updated to each partition for next iteration. In the process of repartition, it needs to get planes at the left and right edge, sort them by key and zip with original latest results, which trigger data shuffle in Spark. The bigger the size of partitions is, the more time it takes to shuffle them. The 'Shuffle Read Blocked Time' increases drastically from 1 plane of I to 2 planes of I , which accounts for the performance decreasing. Figure 4.12 shows the best speedup with different configurations of cores, in which the scalability is obvious while increasing the number of cores.

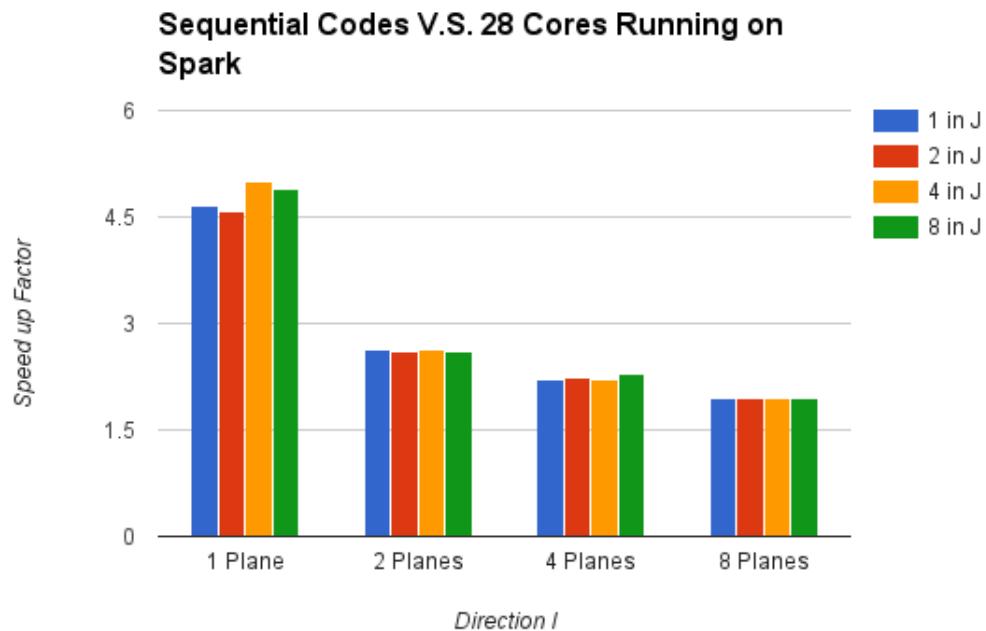


Figure 4.10: The Speedup of Parallel Template Codes with 28 Cores to Sequential Codes

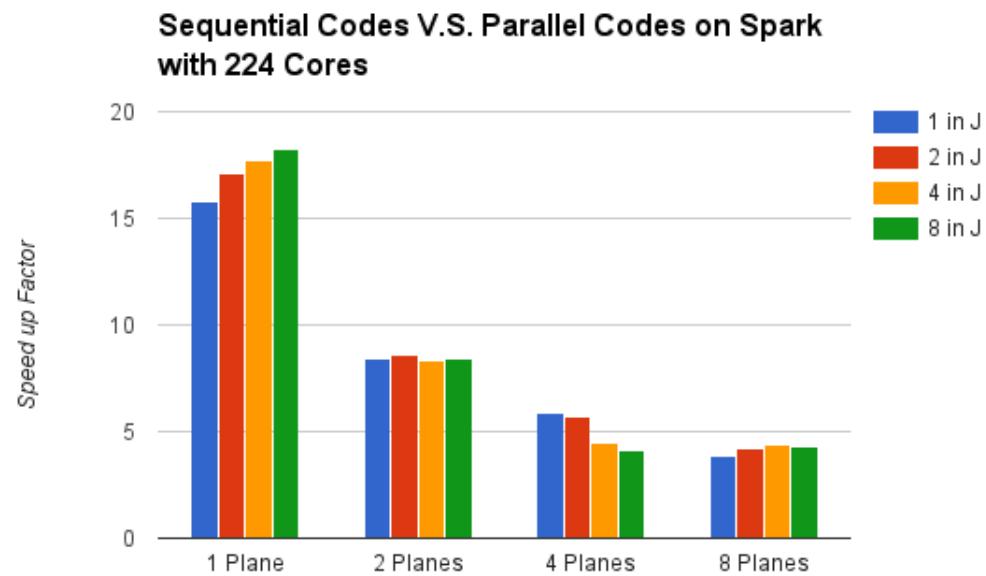


Figure 4.11: The Speedup of Parallel Template Codes with 224 Cores to Sequential Codes

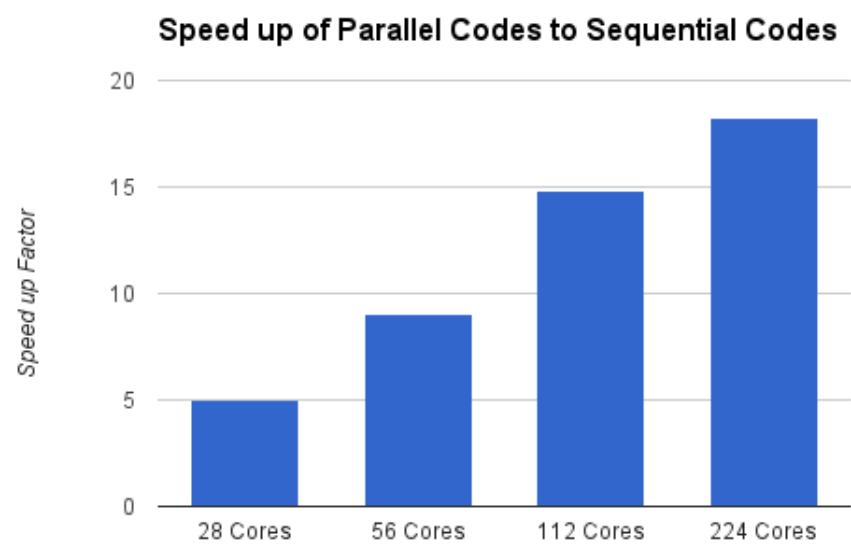


Figure 4.12: The Best Speedup of Parallel Templates for Stencil Computation

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

The paper presents a scalable distributed Seismic Data Analytics toolkit that is implemented on top of Apache Hadoop and Spark big data analytics platforms. It is an attempt to apply HPC optimizations to big seismic data analytics applications, and simplify the parallelism efforts. The experiments and related analysis given in this paper also shows that this toolkit provides promised scalability, by which performance enhancement can be achieved by increasing cluster hardware resources or tuning distribution parameters without refactoring the application code.

In the future, advanced data distribution strategies such as tiling and bricking with 3D overlapping will be implemented, as well as the application utilities and web service for data browsing and analytics purposes will be developed to better facilitate the works of scientists. Both computation and visualization will be put under consideration to design scalable data processing solutions on big data analytics platforms.

REFERENCES

- [1] 5-google-projects-changed-big-data-forever.
- [2] Clowdfloss. <https://github.com/xflows/clowdfloss/wiki>. [Retrieved: June, 2016].
- [3] E is for envelope. <http://agilegeoscience.squarespace.com/journal/2011/3/23/e-is-for-envelope.html>. [Retrieved: July, 2015].
- [4] FEI Digital Rock. <https://www.fei.com/oil-gas/>. [Retrieved: June, 2016].
- [5] Hadoop Introduction. <http://hadoop.apache.org/>. [Retrieved: January, 2014].
- [6] IndexedRDD for Apache Spark.
- [7] NMONVisualizer. <http://nmonvisualizer.github.io/nmonvisualizer/>. [Retrieved: June, 2016].
- [8] Reflection seismology. http://en.wikipedia.org/wiki/Reflection_seismology. [Retrieved: June, 2015].
- [9] Seismic Inline. http://www.glossary.oilfield.slb.com/Terms/i/in_line.aspx. [Retrieved: July, 2015].
- [10] Seismic Reflection Methods. http://www.epa.gov/esd/cmb/GeophysicsWebsite/pages/reference/methods/Surface_Geophysical_Methods/Seismic_Methods/Seismic_Reflection_Methods.htm. [Retrieved: July, 2015].

- [11] Spark Lightning-fast cluster computing. <http://spark.incubator.apache.org/>. [Retrieved: January, 2014].
- [12] Spark RDD Shuffle Operation. <http://spark.apache.org/docs/latest/programming-guide.html#shuffle-operations>. [Retrieved: June, 2016].
- [13] Stencil code. https://en.wikipedia.org/wiki/Stencil_code. [Retrieved: June, 2016].
- [14] The Extreme Science and Engineering Discovery Environment. <https://www.xsede.org/overview>. [Retrieved: June, 2016].
- [15] What is AOV. <http://www.agilegeoscience.com/blog/2011/6/6/what-is-avo.html>. [Retrieved: July, 2015].
- [16] Baaziz A. How big data is changing the oil and gas industry, December 2013.
- [17] SEG Technical Standards Committee. SEG-D, Rev 2.1. http://www.seg.org/documents/10161/77915/seg_d_rev2.1.pdf, January 2006.
- [18] Kaushik Datta, Mark Murphy, Vasily Volkov, Samuel Williams, Jonathan Carter, Leonid Oliker, David Patterson, John Shalf, and Katherine Yelick. Stencil computation optimization and auto-tuning on state-of-the-art multicore architectures. In *Proceedings of the 2008 ACM/IEEE conference on Supercomputing*, SC '08, pages 4:1—4:12, Piscataway, NJ, USA, 2008. IEEE Press.
- [19] Adam Farris. How big data is changing the oil and gas industry. <http://analytics-magazine.org/how-big-data-is-changing-the-oil-a-gas-industry/>, November 2012.
- [20] Dongni Han, Shixiong Xu, Li Chen, and Lei Huang. PADS: A Pattern-Driven Stencil Compiler-Based Tool for Reuse of Optimizations on GPGPUs. In *the*

ICPADS Conference: The 17th IEEE International Conference on Parallel and Distributed Systems. IEEE.

- [21] Andrea De Mauro, Marco Greco, and Michele Grimaldi. A formal definition of big data based on its essential features. *Library Review*, 65(3):122–135, 04 2016.
- [22] Anthony Nguyen, Nadathur Satish, Jatin Chhugani, Changkyu Kim, and Pradeep Dubey. 3.5-D Blocking Optimization for Stencil Computations on Modern CPUs and GPUs. In *Proceedings of the 2010 ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis*, SC ’10, pages 1–13, Washington, DC, USA, 2010. IEEE Computer Society.
- [23] Y. Yan, C. Chen, and L. Huang. A productive cloud computing platform research for big data analytics. In *2015 IEEE 7th International Conference on Cloud Computing Technology and Science (CloudCom)*, pages 499–502, Nov 2015.
- [24] Y. Yan, L. Huang, and L. Yi. Is apache spark scalable to seismic data analytics and computations? In *Big Data (Big Data), 2015 IEEE International Conference on*, pages 2036–2045, Oct 2015.
- [25] Yuzhong Yan. A template-based seismic data analytics cloud platform. Master’s thesis, Prairie View A&M University, 100 University Dr, Prairie View, TX 77446, United States, 12 2015.