HIGH PERFORMANCE PRIVATE CLOUD FOR SATELLITE DATA PROCESSING – ENGINEERING IN CLOUD

Raghavendra Kune[1], Ankit Chaudhri[2], Pramod Kumar Konugurthi[3], Geeta Varadan[4]

Advanced Data Processing Research Institute (ADRIN),
Department of Space, Government of India,
203 Akbar Road, Tarbund, ManovikasNagar (P.O), Secunderabad-9,
INDIA
{[1]raghav,[3]pramod}@adrin.res.in
[2]chaudhriankit@gmail.com
[4]geeta@isro.gov.in

***Abstract*— In this paper we present a private cloud for satellite data processing and expect to see increased throughput, significant cost reduction in the compute infrastructure, easy management of the system and low maintenance over a conventional multi core desktop satellite data processing system. In Satellite data processing; "Data product generation" is an engineering activity, often used to process the satellite data and produce the data products, called "value added data products", such data products are good quality images improved in radiometry and geometry over the raw/unprocessed data. These, data products are often used in analysing the earth observation systems such as object detection, object change analysis, urban image mapping etc. In Remote Sensing field, the satellite images are acquired at the ground stations from the various satellite sensors in the earth orbit; and are archived at Data Centres for data product generation and application specific analysis. Here, we discuss one such existing satellite data processing systems based on MPI implementation for distributed processing on multi core desktop systems in a data centre and systematically analyses the systems limitations such as dynamic scaling of the computational resources, limitations in MPI (Message Passing Interface) based data product processing, low utilization of the physical compute resources and show the experimental results over processing on physical Vs virtual systems. Also, here we illustrate the technologies used in the private cloud for satellite data processing.**

***Keywords*— Cloud Computing; Virtualization; MPI; Data Products; Aneka Programming Model; Data Grid; OGC Services; Engineering in Cloud.**

1. Introduction

Remote Sensing involves various ground stations setup across the globe to acquire the images from the satellite sensors in the earth orbit. Due to the visibility limitations of the satellite imaging, the ground stations are geographically distributed. However, such received images are at the ground stations are accumulated in the data storage areas at respective reception centers constitute the data grid[1].The information extracted from these remotely sensed images is being applied for various activities like weather forecasting and climate science, cartography and applications like agriculture, town planning, natural resource management, disaster management etc[2]. The complexity involved in data analysis and modeling these data sets has been augmented over the last four decades due to the continuous growth in the capabilities of instrumentation of remote sensing satellites. The demand for various data products (in the form of value additions) generated from remotely sensed data is also growing due to increase in the number of user groups. In order to meet these demands, more satellites are placed in the orbit, which makes processing, archiving, and retrieval of data, a complex activity. Satellite images are often large in data size and current remote sensing systems generate hundreds of gigabytes of raw sensor data per day to produce useful data products. The process involved in the generation of these products is time consuming and demands large computational and infrastructural resources.

Cloud computing has emerged as a next generation computing in IT and uses service oriented architectures for delivering resources such as compute, storage and software as services. Cloud computing facilitates in reducing the infrastructure cost and effective utilization of the resources. Cloud infrastructure can be classified as public, private, or hybrid based on the model of deployment [3] and services offered to its' users. In a public cloud 3rd party manages cloud infrastructure and services which are available on subscription basis (pay as you go model). In a private cloud deployment a company/organization manages its own cloud infrastructure for internal and/or partners use. A hybrid cloud takes shape when a private cloud is supplemented with computing capacity from public clouds [4]. The clouds offer services at IaaS (Infrastructure as a Service), PaaS (Platform as a service) and SaaS (Software as a Service) layers. Offering virtualized resources (computation, storage and communication) on demand is known as Infrastructure as a Service. PaaS offers a

Copy Right 2012, This paper appeared in the Proceedings of the First International Conference on Advances in Cloud Computing (ACC-2012), July 26-28, 2012 Bangalore, India. Universities Press, Rajiv Ranjan, Raj Buyya and Anirban Basu(Editors)

higher level of abstraction to make a cloud easily programmable. A cloud platform offers an environment on which developers create and deploy applications and do not necessarily need to know the number of processors or amount of memory required for using the application. Software as a Service (SaaS) allows users to access applications via web services, thus eliminating the need to install applications on their local systems and other maintenance issues.

In this paper we described a model to transform the existing data center satellite data processing system to a private cloud, enabling infrastructure virtualization, resource provisioning and applications as services. The reasons for moving to the cloud are; easy management of compute and storage infrastructure, hiding data and process complexities, better utilization of the resources and achieving high throughput and higher return on investment.

The paper is organized into four sections. Section 2 describes the related work, section 3&4 depicts satellite data processing in the traditional existing model, section 5 describes design and architecture of satellite data processing in private cloud and section 6 depicts the experimental results.

2. Related Work

Remote Sensing data processing involves processing the large data sets to deliver the good accuracy data products. In the recent past works are conducted to process the data using the technologies; cluster, grid and recently cloud. Christopher et al. [5] presented their work on facilitating NASA earth science Data processing using Nebula Cloud Computing and OpenStack[6]. An eScience application Data Re-projection and reduced pipeline on MODIS satellite data over Windows azure [7] demonstrated satellite data processing with a goal of hiding data complexities from end users. Environmental data processing and products derivation of NOAA satellite and information service to domestic and foreign users is demonstrated at OSDPD [8]. Ammar et al. [9] discussed the basic issues in data usage and processing on cloud and their limitations over public clouds. Karmen et al. [10] addressed a satellite based cloud computing and virtualized information resources and the impact of satellite cloud computing on the development of satellite information clouds as open access public resources. Natavia et al. [11] presented their experience with in the development of the Grid Systems for satellite data processing in the Space Research Institute of NASA-NSAU and the conceptual foundations for the development of grid systems that aimed for satellite data processing. The applicability of Cloud Computing to a large-scale satellite ground systems and architectures that illustrate how several Cloud Computing implementation approaches apply to ground systems and aspects of satellite ground system architectures are presented [12]. Golpayegahi et al. [13], presented a work on processing a multiyear remote sensing data using Hadoop Distributed File System and MapReduce Parallel Computing Framework over the cloud of high end computer clusters. Other works conducted on Satellite product generation such as Service Oriented Utility Grid for Remote Sensing Applications [14] for data product generation, and the 3-Dimensional Visualization for the satellite images [15] are works based on Rigorous sensor Model [16], are ideally suited for Computational Grids.

Clouds are the next generation computing in IT for guaranteed service delivery. Amazon Elastic Compute Cloud(Amazon EC2)[17] is a public cloud ; offers web-scale computing and storage environment with a variety of operating systems running on Amazon's computing infrastructure and on Amazon Simple Storage Server Amazon S3[18]; this infrastructure is used for public domain, still a challenge for best data security and large scale data processing. Windows Azure [19] offering on-demand compute, storage, scale and manage web applications on the internet through Microsoft data centers. Microsoft Hyper-V [20] on Microsoft Windows Server 2008 offers windows based virtual machines. Ubuntu Enterprise cloud [21] is used to build private cloud using KVM as the virtualization Technology and Walrus for storage. The machine instances are compatible to Amazon EC2 AMI and storage is compatible to Amazon S3. Ubuntu Enterprise cloud is a choice for the enterprises to build private cloud which are compatible to Amazon EC2 public cloud. Eucalyptus [22] is cloud software for enterprises and government agencies to build and own private/hybrid cloud computing environment. Aneka [23] is the platform (PaaS) for design and deployment of .NET based services on private/public/hybrid clouds. Aneka offers programming models to build applications as services for the cloud, enabling efficient use of the computing infrastructure with dynamic resource provisioning and scheduling mechanisms. One such private cloud for satellite data product generation was demonstrated in our earlier work presented at the CCGRID 2010[24] using Aneka task based programming model. Aneka uses resource pools configured with various hypervisors. Xen[25] and Vmware [26] hypervisors configure dynamically scalable virtual machines running on the bare minimum hardware.

Combination of open source softwares such as GeoServer[27] etc, and OpenLayers address presentation of remote sensing data on a web client using OGC's WMS and WFS standards.

The current work describes a private cloud framework for Remote Sensing applications to derive the optimal resources for satellite product generation.

3. Satellite Data Processing

Satellite data processing consists of various methods to correct the radiometric errors and geometric distortions in the basic data generated by the sensor; this data is termed as Level-0. The procedures like georeferencing and registration [28] applied on the Level-0 data to generate the products such as;

Copy Right 2012, This paper appeared in the Proceedings of the First International Conference on Advances in Cloud Computing (ACC-2012), July 26-28, 2012 Bangalore, India. Universities Press, Rajiv Ranjan, Raj Buyya and Anirban Basu(Editors)

Level 1 – Radio metrically corrected and geometrically corrected only for earth rotation (Browse product)

Level 2 – Both radiometric and geometrically corrected (Standard product)

Level-0 data under goes through a Rigorous sensor Model [16] (A series of transformation to convert a pixel to ground coordinates) shown in Fig. 1, this process is termed as data product generation, such data product is either systematically or geometrically corrected. The generation process is a compute intensive task based on scale (map) or coverage (Region of interest) and here each task is independent and considered to be a BOT (Bag of Tasks) for processing.

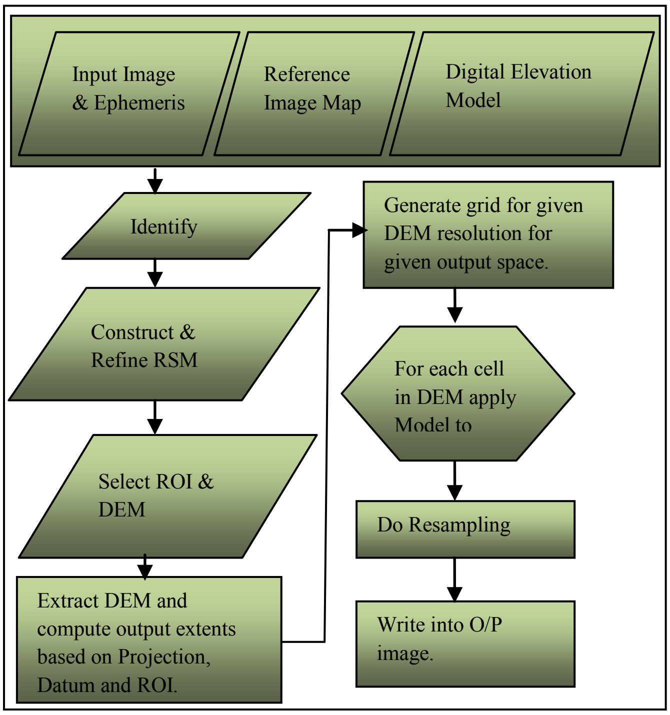

Fig. 1 Precision product generation using RSM

4. Existing Data Product generation System - Data Center Model

Currently the data products are generated on the physical desktops and generation process is tightly coupled with the interactive elements and computing elements. The computing process is based on MPI (Message Passing Interface) [29]; executes the jobs on the available physical desktop systems; in the sequence one after another on the first come first serve basis. The major drawback of the current system is lack of features such as load balancing, failed jobs detection and re-scheduling of the failure jobs, which are very important to achieve higher throughput and reliability. Another important limitation of the existing system is lower resource utilization i.e., even though the jobs need less resources for computation, entire system resources are locked. The existing system does not scale the computing resources when the demand increases for processing the more data products. To overcome the issues, in the next section, we present a private cloud model for dynamic scaling of the infrastructure, job execution in parallel with load balancing, job monitoring and failure request handling.

The paper addresses issues such as transformation from high-end data center to the private cloud setup, transforming the existing data product generation utility/application as the software service, effective scheduling of the jobs, and dynamic expansion of the computing resources for the optimized resources utilization for remote sensing data product generation.

5. High Performance Satellite Data Processing - Private Cloud

The architecture of the proposed private cloud is shown in Fig. 2. It is a three layered architecture; bottom layer forming the infrastructure layer (IaaS); middle layer is the development layer (PaaS) for application development and dynamic resource provisioning, top layer is the application layer (SaaS) where applications are deployed as services for the end users. Platform layer is interfaced with the Data Grid/Data Centre(s) layer to serve Data as a Service (DaaS) in the proposed cloud. Data in remote sensing domain is expensive and often the user ends up paying for the data even though he/she does not use or require, since minimum extent changes from vendor to vendor. Hence here we introduce DaaS model, where flexibility can be provided to the users to pick and choose between the sensors based on many parameters such as cost, spatial resolution, temporal resolution (when the image was acquired etc). However in the present work only one cost model is used for DaaS. IaaS layer can be configured with hypervisors to provide a scalable environment for storage, compute and network.

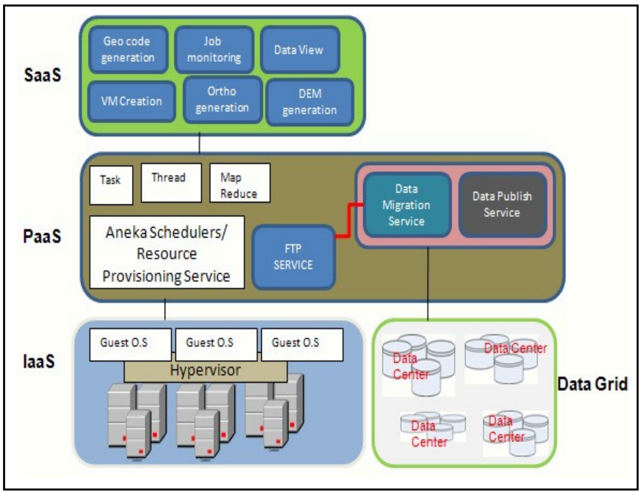

Fig. 2 Architecture of the private cloud

Copy Right 2012, This paper appeared in the Proceedings of the First International Conference on Advances in Cloud Computing (ACC-2012), July 26-28, 2012 Bangalore, India. Universities Press, Rajiv Ranjan, Raj Buyya and Anirban Basu(Editors)

PaaS layer is configured with Aneka which provides a cloud application development platform (CAP) for developing and running compute and data intensive applications. As a platform it provides users with both a runtime environment for executing applications and a set of API's that allow you to build new applications. PaaS also addresses scheduling of tasks and dynamic resource provisioning. The Data Grid Layer forms the network of data centre(s) where the satellite data is archived and the data transfer takes place via FTP. The Data Grid hosts satellite data received from various sensors. SaaS layer consists of OGC [31] compliant web processing Services (WPS) [31] for processing satellite data.
The deployed private cloud consists of a Cloud portal which is the single point of entry for users. The cloud portal allows the data providers and researchers to publish their data and remote sensing applications through the cloud portal. The research analysts can use the published data and applications for analysis. In the following sections we discuss Infrastructure Virtualization, Application Services, Dynamic Resource Provisioning, Results and Conclusions.

A. Infrastructure Virtualization

The servers used to set up a scalable cloud environment have the configuration as shown in Table I. This infrastructure can be virtualized as 128 independent systems comprising a single core each, which enables us to provide a large desktop infrastructure using six servers. The IaaS layer is configured using Xen Hypervisor, and the resources are managed using Xen Center.

TABLE I Server Configuration

Server	Quantity	No. of CPUs(Cores)	Total CPUs (Cores)
Dell Blade	4	24	96
HP proliant	2	16	32

A resource pool is created to club all the resources for better utilization and management. Templates are created so that virtual machines can be cloned as and when required.

TABLE II Template Configuration

Template	OS	RAM(MB) (MIN)	DISK(GB) (MIN)	Processor	No. of Cores
Geo	Win XP Sp2	128	10	Intel Xeon 64bit,3GHz	1
Ortho	Win XP Sp2	256	10	Intel Xeon 64bit, 3GHz	1
DEM	Win XP Sp2	256	10	Intel Xeon 64bit, 3GHz	1

A Template is a self contained operating system image having configuration settings necessary to process the demand. The templates are configured on the local systems, tested and then imported to the pool. Dynamic resource provisioning mechanisms of Aneka is used to scale the virtual machines depending on the load. The configuration of the templates is shown in Table II.

B. Data as a Service

Data collected from various satellites is archived in data centre(s) / ground station, which are geographically distributed across the globe. Due to the visibility constraints of satellites, the data transmitted by satellite sensors is received at different ground stations which constitute the data grid as shown in Fig. 3.
Data grid/cloud is an emerging technology for managing large amounts of distributed data. This Technology is highly anticipated by scientific communities, such as satellite data processing, where research communities work with outsized collection of data. Data in these data centre(s) is replicated (mirrored) across various sites, and in some cases similar data may be received at more than one centre for high availability. Data replication technique is useful for redundancy and also in selecting the site which can give better data transfer rate in case of network congestion; the decision can be made based on the site for which communication network is less congested over the nearby proximity site. Data Grid/Cloud is interfaced with cloud as shown in Fig 3. Data is made available as a service to users. It allows them to view process and analyze data without having it transferred on to their local systems and thus eliminating all the data management issues.

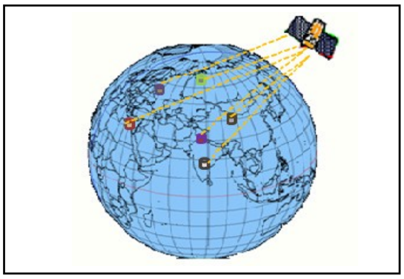

Fig. 3 Data Grid/Cloud

Satellite image metadata is very significant to understand image itself. Here we propose publishing of metadata information in the form of XML which is validated against an external xml schema definition. Users can make use of the metadata publishing service to access metadata of any desired image. Also, every image is delivered with a metadata file which describes the image. Even though the satellite data is acquired in various formats and resolutions, but they can be put into a single agreed format. Here, we used the Geonetwork [31] metadata format according to the ISO 19115 standard. Metadata consists of various parameters like date of acquisition, satellite, sensor, corner coordinates and cost per square kilometer, the structure is shown in Table III. In the schema the information under <strip> </strip> denotes the information about each strip (data).

Publisher: owner of the data.
Date: Date on which the data is recorded.
Satellite: The unique satellite identification number.
Sensor: Sensor by which the data is recorded (like Panchromatic, Multispectral).
Segment: Scene identification in the corresponding orbit.
Orbit: The unique orbit path of the satellite.
File: The directory path and file to be identified
Lat#, Long#: The four corners of the recorded data.
Cost: Amount for processing per Square kilometer.

TABLE III Sample Metadata

```
<STRIPS>
    <STRIP>
        <PUBLISHER>adrin</PUBLISHER >
        <DATE>14-MAR-07</DATE >
        <SATELLITE>IRS-P5</ SATELLITE >
        <SENSOR>PAN</ SENSOR >
        &lt;SEGMENT&gt;6042332&lt;/ SEGMENT &gt;
        <ORBIT>10047</ ORBIT >
        <FILE>20070314_604233.Img </FILE>
        <LAT1>36.106702</ LAT1 >
        < LON1 >94.296127</ LON1>
        < LAT2>36.048473</ CORNER2_LAT >
        < LON2>94.626013</ CORNER2_LON >
        < LAT3 >35.780754</ CORNER3_LAT >
        < LON3 >94.550916</ CORNER3_LON>
        < LAT4>35.83888</ CORNER4_LAT >
        <LON4 >94.222143</CORNER4_LON>
        <COST>5.0</ COST >
    </STRIP>
</STRIPS>
```

When data is selected from the grid, the required region of interest only can be extracted and transferred instead of full file. This methodology has many significant advantages for the user and Cloud community; such as better utilization of the network bandwidth, reducing the data transfer time, and providing “pay for what you use” model. Currently FTP is used for transferring files between the master and executor nodes.

Data Query
A web interface is proposed and provided to users which allow them to query for datasets available on the data grid. The query results are overlaid graphically on a world map which makes it easier for the users to select the required data. The world map provides users the contextual information about the location itself. The present implementation is configured using simple Geoserver [27] with Blue marble and world cities layer, which can be augmented/replaced by more content rich cites like Bhuvan [32], Google maps [33].

The selected data is then overlaid as vector layer, using an in house developed overlay tool as show in Fig. 4. For the selected data users can place demand for Geo, Ortho or DEM services either by marking regions of interest on the coverage or by selecting map sheets falling within the coverage extents. These requests are then finally submitted to the cloud for processing.

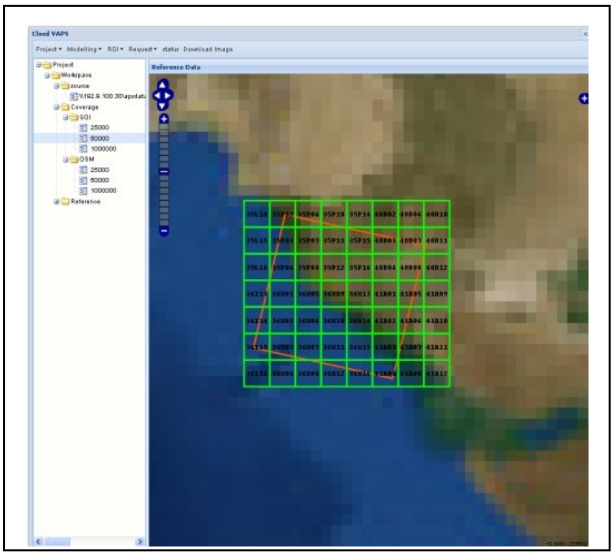

Fig. 4 Selected coverage overlaid on the world map

C. Platform and Application Services

Application services are implemented as OGC complaint web services. The product generation process is made available as a cloud service to users for processing data. VAPS (Value Added Product System) [24] is an in house customized desktop tool which enables users to:

- view and select regions of interest
- select and refine ground control points (GCP) to do RSM for better geometric accuracy [16] and
- To generate GEO, ORTHO or DEM (data elevation model) products in different scales mentioned in Table IV.

The cloud application services offer all the mentioned above features of VAPS to users. It enables the users to view and select the data of their choice for product generation. Each selection can be treated as an independent task. Hence, task based programming model of Aneka is used to execute multiple tasks on the cloud infrastructure simultaneously. Aneka as described previously is a cloud application development platform that leverages scalable cloud infrastructure for application development and execution. It provides a runtime environment and a set of APIs for programming in .NET environment; that allow developers to build .NET applications leveraging cloud. Aneka, while mostly designed to exploit the computing power of windows based machines, which are most common within an enterprise

environment, is portable over different platform and operating systems by leveraging implementations like Mono. Currently, in this work we experimented the application to run on windows based Desktop PCs. The virtual machines are created using Xen API's and data transfer and scheduling is addressed by Aneka FTP and Aneka Clock priority based scheduling algorithms. After successful execution the data can be downloaded to the client. Aneka comes with the various scheduling algorithms like FIFO, Clock Rate Priority, etc. for optimal execution of tasks on the virtual infrastructure. The jobs submitted to the scheduler are monitored via web based console which uses the job monitoring service. The status of the jobs can be: *un- submitted, submitted, running, finished and failed*. The description of job's status is given below.
Un-submitted: Job submitted to the scheduler but not scheduled to any virtual machine resource.
Submitted: Job is assigned to the virtual machine for execution.
Running: Job is under execution
Finished: Job finished processing; ready to view and download.
Failed: Job failed; should be rescheduled on other virtual machine.
The existing legacy desktop based applications in C/C++ are transformed to .NET services by leveraging PInvoke (Platform Invocation Services) [23]. These services are published in Aneka container [24] for data product generation. As and when the tasks are submitted by various users, the scheduler automatically invokes and does the service migration to the executor nodes for computation.

TABLE IV

Scale	Approximate Area(in Km)
1 million	193600
250,000	12100
50,000	756.25
25,000	189.0625

6. Results

The model is generic and is applicable for public/private/hybrid clouds. The private cloud proposed is configured with the DELL Blades X64 systems with the Xen[25] from Citrix for virtualization of the resources , Aneka [23] for resource provisioning, and finally we present data publishing, product generation applications as OGC (Open Geospatial Consortium) [27] compliant Software As A Service.
In the proposed private cloud we conduct two sets of experiments; first experiment is to compute optimal resources required for processing satellite data products like GEO, ORTHO and DEM and the second compares the task execution times between desktop processor and virtual machines on the cloud.

A. *Experiment 1: Optimal Resource Configuration*

The table V below depicts the average time taken to generate each of the data products in different RAM configurations. The results show that systems with 512 MB RAM is sufficient enough to generate the data products for GEO, ORTHO and DEM. Hence, templates are created with the resources of 512MB, Windows XP operating system and 10 GB disk storage space.

TABLE V

Type	RAM(MB)	Avg Time (sec)
Geocode(10 Tasks)		
	128	159.75
	256	50.0072
	512	41.7919
Ortho (10 Tasks)		
	256	211.3333
	512	140.3333
DEM (10 Tasks)		
	256	260
	512	205

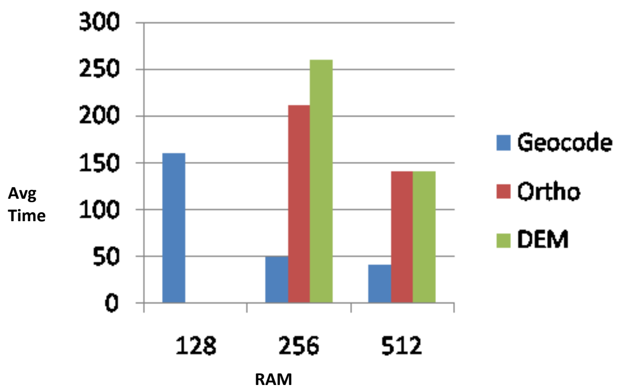

Fig. 5 Times lines for Table V

The above shown experiments are conducted on different data sets and for various sensors. The results depict, however increase in the RAM from 512 MB to 1024 MB will not achieve a significant improvement in reducing average time of execution. Hence virtual resource with 512MB RAM is configured and imported in the resource pool as template for dynamic resource provisioning using Aneka in Xen.

B. *Experiment 2: Comparison of Execution between Desktop and a Cloud System*

We conducted the experiment to compare the execution timings between desktop processor and virtual machines in cloud. Here we demonstrate an experiment to achieve high throughput using cloud infrastructure. The major drawback of this desktop application is that only one product can be generated at a time whereas in cloud we can dynamically provision resources and each task can be scheduled on a different virtual machine. For instance, if there is a requirement of hundred products, in a desktop application these hundred requests are processed one by one in a sequential order while in cloud we can dynamically provision hundred virtual machines and then each product request executes simultaneously on each of the hundred virtual machines which in turn increases throughput. Tables VI and VII display the results obtained from this experiment.

TABLE VI Timings for geocoded product

Scale	Desktop(Single Task) – seconds	Cloud(100 Tasks)- seconds
One Million	126	132
250,000	9	21
50,000	2	17
25,000	2	18

These results demonstrate the time required to execute hundred tasks in cloud and is nearly equal to process a single product on the desktop. With this it is evident that the proposed private cloud is capable of handling large number of demands to achieve high performance and high through by leveraging compute resources effectively. In cloud environment we can configure each core in our resource pool as an individual system and this system would be capable of handling each job request (product generation job).

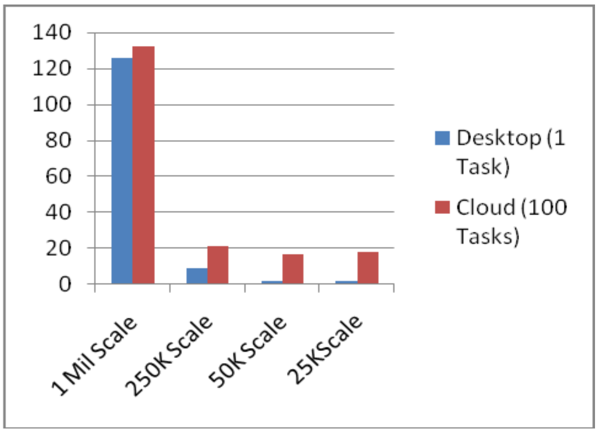

Fig. 6 Time lines for Desktop / Cloud – geocoded Product

TABLE VII Timings for ortho product

Scale	Desktop Timings(Single Task)(sec)	Cloud Timings(100 Tasks)(sec)
1,000,000	1540	1557
250,000	72	90
50,000	7	21
25000	6	19

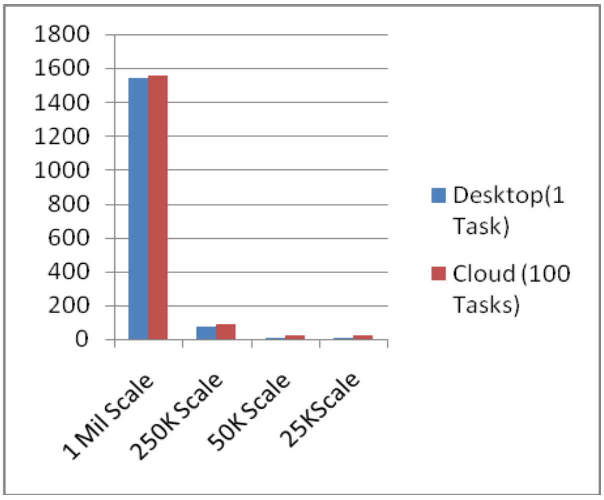

Fig . 7 Time lines for Desktop / Cloud- Ortho Product

7. Conclusions

The model “High Performance Private Cloud for Satellite Data Processing” demonstrated a Cloud Computing Framework for Satellite Data Product Generation to achieve high throughput/performance over a standalone bag of tasks processing in desktop grid.
Experiment I described in results section, clearly indicated that if optimal resource configuration for any application can be arrived at, then one could use cloud resources efficiently thus increasing VM density leading to better throughput, performance and better resource utilization. Here, we presented a work for transforming the desktop based interactive applications as .NET services and using dynamically scalable computing infrastructure to achieve the high performance using the Aneka Task based programming model and schedulers.
Effective utilization of the computing resources, load balancing, failure job management and high throughput is demonstrated in the experiment II, using the virtual desktop and PaaS infrastructure.

8. Future work

The paper presents the work focused on private cloud, but transforming it to a public cloud / hybrid cloud poses many challenges to be explored in the areas like privacy, security, vendor lock-in. Hybrid cloud model can be explored in the case of a need for more compute resources. Currently a single service provider for data is considered, it is proposed to extend the same for multiple service providers where each service provider may follow different cost models for data as well as software services. Presently budgeting models for estimating the cost for the processing of the data is not considered, a model can be worked out to compute the cost based on the various parameters like delivery time of the product, area in square kilometer, vicinity of the data product etc.

Acknowledgement

We express our sincere thanks to Prof. Rajkumar Buyya, University of Melbourne & CEO Manjrasoft Pty Ltd, Australia for the support extended to us in understanding the Aneka Programming model. We are thankful to Christian Simone Vecchiola and Dileban karuna Moorthy, university of Melbourne for the support in interfacing Aneka and Xen for dynamic resource provisioning in the cloud. We express our thanks to our colleague Sri A.Akilan, Scientist; ADRIN helping us in interfacing the Desktop MPI applications with .NET based Data Product Generation Service.

References

[1] A.Chervenak, I.Foster, C.Kessleman, C.Salisbury, S.Tuecke , "The Data Grid : Towards an Architecuture of the Distributed Management and Analysis of Large Scientific Data Sets" *Network Symposium, Seattle 1999*.

[2] Paul M.Mather , *Computer processing of remotely sensed images- An introduction* , John Wiley Sons,1987.

[3] Rajkumar Buyya, Chee Shin Yeo, Srikumar Venugopal, James Broberg, and Ivona Brandic, "Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility" *Future Generation Computer Systems, Volume 25, Number 6, Pages: 599-616*, ISSN: 0167-739X, Elsevier Science, Amsterdam, The Netherlands, June 2009.

[4] B.Sotomayor, R.S. Montero, I.M.Llorente, and I.Foster, "Virtual Infrastructure management in private and hybrid clouds" *IEEE Internet Computing*, 13(5):14-22, September/October 2009.

[5] Christopher Lynnes, Michael Theobald, Esfandiari Asghar, Jane campino, Bruce Vollmer – Facilitating NASA earth Science Data Processing using Nebula Cloud Computing. UCAR SEA software Engineering conference 2012 , page 1. NASA/GMU CSISS.

[6] Openstack : OpenSource Cloud Computing platform, http://www.openstack.org (Last accessed on 16th May 2012)

[7] Jieli, DebAgarwal,Marty Humphray, Catharine Van Ingen, Keith Jackson, Young Ryu, "eScience in the Cloud: A MODIS satellite Data Reprojection and Reduction Pipeline in the Windows azure platform", *International Parallel and Distributed Processing Symposium(IPDPS) 2010.19-23 April 2010.*

[8] OSDPD : Office of Satellite Data Processing and Distribution : www.ospd.noaa.gov/ml/index.html(Last accessed on 16th May 2012)

[9] Ammar Khalid, Husnain Mujtaba, "Data Processing issues in cloud computing", *Second International conference on machine Vision*.

[10] Karmen Kanev, Nikolay Mirenkov, Proceedings of 2011 *IEEE workshops of international conference on Advanced information networking and applications pages 147-152.WAINA 2011*.

[11] Natavia Kussal, Andrii Shelestov,Mykhailo Korbakov,Oleksi Kravchenko, Serhiy Shakun, Mykola Llin, Alina Rudakova, Volodymyr Pasechnik, "Grid Infrastructure for Satellite Data Processing in Ukraine", *ITech 2007, Information Research Applications.*

[12] Richard Anthony , John Fritz , Doug Barnhat, "Cloud Computing Applications for Large-Scale Satellite ground systems", *2011 Military Communications for Large Scale Satellite ground Systems*

[13] N.Golpayegahi,M.Halem, "Cloud Computing for Satellite Data Processing on High end compute clusters", *IEEE international conference on cloud computing , 2009. p.p.88-92.*

[14] P.Deepak,K.Pramod Kumar,Geeta Varadan,"A Service Oriented Utility Grid for Data Parallel Remote Sensing Applications",*2009 High Performance Computing & Simulation(HPCS).*

[15] P.Deepak,K.PramodKumar,GeetaVaradan, "Serive Oriented Utility grid for 3-Dimensional Topographic Visualization from Satellite Images", *4th International Conference 7-12 Dec'2008, IndianaPolis,Indiana,USA*.

[16] Radhadevi .P.V,S. S. Solanki, 2008, "In-flight Geometric Calibration of Different Cameras of IRS-P6 Using a Physical Sensor Model". *The Photogrammetric Record 23(121): 69–89.*

[17] Amazon EC2: http://aws.amazon.com/ec2/ (accessed May'2012)

[18] Amazon S3: http://aws.amazon.com/s3/ (accessed May'2012)

[19] Windows Azure Platform : http://www.microsoft.com/windowsazure/ (accessed nov'2011)

[20] Microsoft Window Hyper-V Technology : http://www.microsoft.com/windowsserver2008/en/us/hyperv-main.aspx (accessed May'2012)

[21] Ubuntu cloud: http://www.ubuntu.com/cloud (accessed May'2012)

[22] Eucalyptus : Building the hybrid clouds ; www.eucalyptus.com (accessed nov'2011)

[23] Christian Vecchiola, Xingchen Chu, and Rajkumar Buyya, Aneka: A Software Platform for .NET-based Cloud Computing, High Speed and Large Scale Scientific Computing, 267-295pp, W. Gentzsch, L. Grandinetti, G. Joubert (Eds.), ISBN: 978-1-60750-073-5, IOS Press, Amsterdam, Netherlands, 2009.

[24] Raghavendra K , Akilan A, Ravi K, Pramod Kumar K, Geeta Varadan, "Satellite Data Product Generation on Aneka .NET Cloud" . *Research Product Demonstration at CCGRID 2010, MelbourneAustralia. URL:www.manjrasoft.com/ccgrid2010*

[25] Citrix Xen http://www.xen.org. (accessed May'2012)

[26] VmWare : Virtualization Technology ; www.vmware.com

[27] GeoServer , http://geoserver.org (accessed May'2012)

[28] George Joseph, *Fundamentals of Remote Sensing* , 2nd ed.,India: Universities Press.

[29] Intel MPI Library for building parallel processing applications on Windows and Linux ; http://software.intel.com/sites/products/collateral/hpc/cluster/mpi_indepth_40.pdf (accessed May'2012)

[30] ANEKA market oriented Cloud Computing- www.manjrasoft.com (accessed May'2012)

[31] http://www.opengeospatial.org (accessed May'2012)

[32] http://bhuvan.nrsc.gov.in/bhuvan_links.html(accessed May'2012)

[33] https://maps.google.co.in/ (accessed May'2012)