

SUSTAINABLE MINING AND GEOINFORMATION

Prof. Mukunda Dev Behera

Centre for Ocean, River, Atmosphere and Land Sciences (CORAL)

Indian Institute of Technology Kharagpur

Week – 02

Lecture 09: Cloud Computing

Welcome to the 9th lecture. Today, let us discuss cloud computing. In the last eight lectures, we have discussed the basics of geoinformation and an introduction to geoinformation. We discussed different datasets as far as remote sensing sensors are concerned. The optical, microwave, lidar, thermal, and also we discussed 3D modeling and GNSS. Today, in the 10th lecture, let us talk about the 9th and 10th lectures.

Let us talk about two advanced topics as far as geo-information is concerned. In today's 9th lecture, we are going to discuss cloud computation or cloud computing. So, how does this differ from GIS? We already understand that a geographic information system integrates both spatial and non-spatial data. Here, we move one step up, where we try to address some constraints in terms of where the data will be stored.

Because we are dealing with a lot of data, much of which is available in the public domain or open sources. So, how do we deal with the data? How do we pull the data stored in another place, another server, or another data server? This eliminates another constraint: creating space in terms of hardware. This, in turn, is the life of an interpreter, of a user.

So in that situation or in this context, let us try to understand, let us try to learn about cloud computing, or what we call cloud computation, and its application in the mining industry. The concepts we are going to cover include an overview of cloud platforms for geospatial data processing. So, I have chosen three cloud computing platforms for today's discussion. We will talk about Google Earth Engine (GEE), which all of us know, as well as Microsoft and Amazon. We will also discuss the different benefits of cloud computing for big data analysis in the mining sector.

CONCEPTS COVERED

- Overview of Cloud Platforms for Geospatial Data Processing (e.g., Google Earth Engine, MS/ Amazon)
- Benefits of Cloud Computing for Big Data Analysis in Mining
- Applications: Large-scale Environmental Monitoring and Temporal Analysis of Mining Regions



As far as applications are concerned, we will also explore large-scale environmental monitoring and temporal analysis in the mining sector, and how cloud computation is becoming widely used and beneficial. Cloud computing provides on-demand access to configurable resources with minimal management. So, what does this mean? It provides on-demand access. The demand must come from the user, whether from the mining industry or the mining sector.

Cloud Computing

- Cloud Computing Provides On-demand Access to Configurable Resources With Minimal Management
- It Uses Remote Servers and Internet Connections, Reducing the Need for Local Infrastructure
- Cloud Applications Require No Bulky Installations, Allowing Global Access through the Internet
- Based On Grid Computing, Cloud Services Offer Dynamic, On-Demand Resource Allocation
- Cloud Computing is Cost-Effective, With Low Setup Costs and Reduced Operating Expenses



(Chauha and Prasad, 2016)



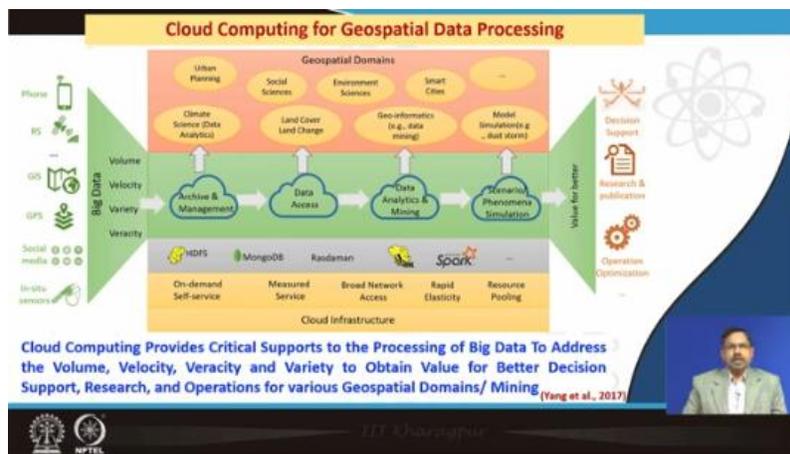
Then, we configure the resources, which are available in terms of data here and there. We configure them for use. So, it is a user-driven or demand-based computation. It uses remote servers and internet connections. These two things are essential.

There has to be good high speed internet connection so that it can be connected to servers which are remotely located across different continents of the world. And that's how it reduces the need, the requirement for local level infrastructure, which many of us or many of these labs are at individual level. We may not be able to afford. And the point is different data we are pulling from different servers which are located remotely. So that also matters.

helps us in pulling data from different sources from our purpose instead of downloading or calling it or first transferring it to our own hardware or the computer. So cloud applications require no bulky installation. You have a laptop, that's enough. Use that at your own place, but you have to have a good internet connection. So then it allows us for global access using the internet.

So we do not need any bulky installation in terms of hardware. Based on grid computing, cloud services offer dynamic on demand resource allocation. So it is based on grid computation. So as now we can understand that it is cost effective. Cost effective in the sense in one server, in one data server it is there and we do not have to use so many for this or at least at the user end it is not involving any bulk installation.

So, it is cost-effective with low setup cost, and thereby it reduces the operational expenditures at the user end. So, the depiction below shows a kind of cloud computation where you, the user, have a laptop or a PC, but you are connected through connected to remote servers through high-speed internet connections. That is the requirement as far as cloud computation is concerned. So now, let us see what the critical supports are in terms of big data processing.

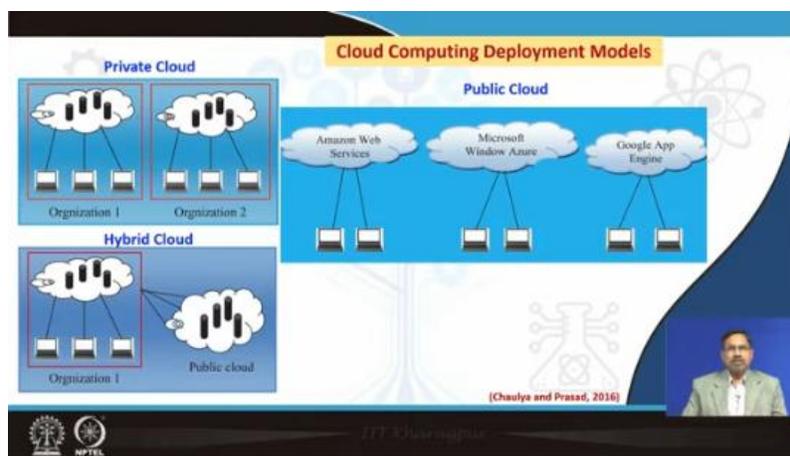


So, this cloud computation for geospatial or big data computation can be clubbed and shown in this way. It addresses—the cloud computation addresses volume, veracity, velocity, and variety of data. So, that means we can characterize big data using four Vs: voluminous data, the speed of the data, the variety—that means the type of the data—and veracity, how huge a database it involves. And then, to obtain value for better decision support, research, and operations.

Now, on the left-hand side, we can see this big data may come from various sources, starting from mobile phones, remote sensors or remote sensing data, different platforms using GIS, even from GNSS platforms, social media, and also ground or in-situ sensor-based observations. So, the data can come from various sources and be stored in different servers, which are remotely connected through high-speed internet. It can have various applications, as far as our mining sector is concerned; it can have various applications. And then, once this big data is stored in different servers, it has to be managed in terms of archival.

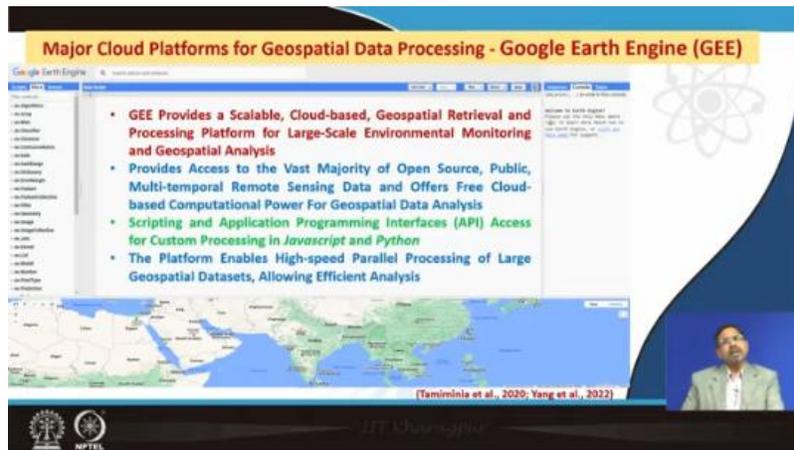
And then it has to be accessed the data using data access protocol, data analytics and mining particularly it needs to be accessed for analysis as far as mining sector as mining industry is concerned. Then the scenario or phenomena simulation also can happen. Once you get a pattern, a trend, you can simulate for any kind of projection or any kind of situation. So with this, you are ready with a value added product that enables for better decision making, for better research or even for operational optimization. So, broadly this demonstration, this picture, figure talks about the cloud computing for geospatial or geoinformation data processing in mining sector.

As far as the deployment models are concerned, it can be of three types. The private, public and hybrid. In terms of private, you have your own servers or the data storage places where you are the own owner. You are not connecting it to the internet. You may have connection as far as the internet or the LAN is concerned.



And another is a hybrid kind of cloud where you have your own private cloud connected to a public cloud. So it has both connection as far as the private and the public. And a typical public cloud we are put here in terms of Amazon Web Services, Microsoft Window

Azure and the Google App Engine which we will discuss in detail. This depiction has come from the Chaulia Prasad publication given in 2016. Let us look at these three one by one.



The Google Earth Engine is one of the major cloud computational platforms for geo-information data processing. So, the Google Earth Engine, or GEE, interface looks like this, which is before us now: Google Earth Engine. On the left-hand side, you have different applications, and the names are written there. And you have a graphical space. The Google Earth Engine provides a scalable, cloud-based geo-information, or geospatial, retrieval and processing platform.

It offers scalability, meaning it handles different scales and types of data organized at various levels. It is cloud-based. It also provides a retrieval and processing platform. So, it also offers a form of retrieval. You retrieve data by calling it, giving a signal, writing code, or simply requesting it.

Then you process it, visualize it, and use it for large-scale environmental monitoring applications, which are useful to the mining sector. Google Earth Engine also provides access to a vast majority of open-source, public, multi-temporal remote sensing data. Since it is a public platform, it gives us access to a lot of open-source data available in the public domain. This includes multi-temporal satellite data available in the public domain, offering free, cloud-based computational power for geospatial data analysis. Look at its power—since it is offered for free, it has many applications in the mining sector and for resources, particularly for students who cannot afford expensive computational hardware. For them, it is truly a boon.

So the third point is Google Earth Engine scripting and API, the application programming interface access for custom processing in JavaScript and Python. So you need to have

understanding or knowledge of JavaScript and Python programming. So once you have this, you can write the codes, you can give the commands and accordingly you can access the data, you can analyze the data, you can process the data, you can come up with final analysis in geospatial domain which has mining applications, application in mining industry. The Google Earth platform thereby enables high-speed parallel processing of large geospatial datasets, allowing efficient analysis. So it thereby offers, it depends on what is our objective.

based on that we can have high speed parallel processing of large geospatial data which helps us in efficient analysis as far as different mining applications are concerned. Now let us look at the second one. The next one is Microsoft Azure, A-Z-U-R-E. So friends, this Microsoft Azure, I have on the left-hand side, I have put some of the depictions that stands for this particular Microsoft Azure package. So this is a cloud computing platform and infrastructure service.

The slide features the Microsoft Azure logo at the top left. The main content is a list of services and capabilities:

- Microsoft Azure is a Cloud Computing Platform and Infrastructure Service from Microsoft, Providing a Range of Services like **Infrastructure As A Service (IaaS)**, **Platform As A Service (PaaS)**, And **Software As A Service (SaaS)**
- The Monitoring Capabilities of Azure Services Includes: **Cloud Services**, **Mobile Services** (Metrics for Mobile Endpoint Availability only), **SQL Databases**, **Storage**, **Virtual Machines**, **Websites** (Metrics For Web Endpoint Availability Only)

(Collier and Shahan, 2015)

A Versatile Cloud Platform Offering Various Tools for Geospatial Data Integration and Processing

Azure Maps for Geospatial Data Visualization and Services Like Azure Synapse and Databricks for Big Data Analytics

Supports Machine Learning, Real-time Analytics, and IoT Integration, Making it Effective for Analysing Real-time Spatial Data

Useful for Integrating IoT Data in Mining Planning, Real-Time Monitoring, and Developing AI-driven Geospatial Insights

A small video inset in the bottom right corner shows a man speaking.

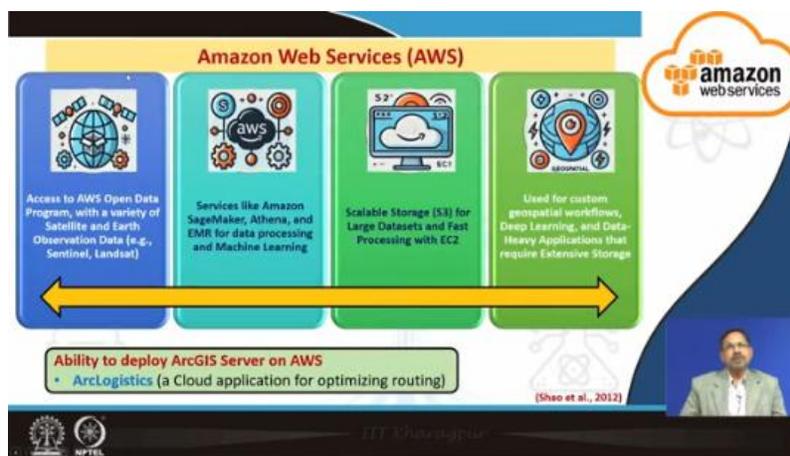
So it is a computational platform. So that means it gives you a platform and infrastructure service from Microsoft. So thereby it provides us a range of services like three important one is IaaS, PaaS and SaaS. So I stands for infrastructure as a service, B stands for platform as a service, S stands for software as a service. So infrastructure, platform and software as a service, IaaS, PaaS and SaaS.

The monitoring capability of Azure services include cloud services, mobile services and also database storage using SQL or PLSQL, whatever the query language, storage, virtual machines and websites. So, this Microsoft Azure is a versatile cloud platform that offers various tools for geospatial data integration and processing. for geospatial data visualization and services like azure synapse and databricks for big data analytics so within the azure maps we have azure synapse and databricks there are different operational

interfaces that helps us in analyzing big data and this microsoft azure It also supports the machine learning algorithm, lot of high end multivariate statistics, real time data analytics. If we are getting some data on real time basis, it helps us in analyzing or taking the data in terms of real time.

And also IoT-based integration. We are going to talk in the 10th lecture more about the IoT, Internet of Things. So this Microsoft Azure has capability to integrate the real-time data and IoT integration. making it effective for analyzing real-time spatial data. So that way it has a slight advantage or difference as far as Google Earth Engine platform is concerned.

And this Microsoft Azure, it is useful for integrating IoT data in mining, planning, real-time monitoring and developing AI artificial intelligence driven geospatial insights. So, the third one as example is AWS, the Amazon Web Services. The Amazon Web Services, it has access, it gives access to a kind of open data program with a variety of satellite and earth observation platforms such as Sentinel, Landsat, all those things which like the previous two also offers. The AWS, the Amazon Web Service, it also offers the services like Amazon SageMaker, Athena, EMR for data processing and machine learning.



So it also has capability in terms of data processing including using machine learning tools. AWS offers scalable storage S3 for large data sets and fast processing with EC2. We will talk about this sometime later. So it offers a scalable and large or you say fast data processing capability. so that is why it is used for custom geospatial workflow so we can have a customized geospatial workflow for deep learning data heavy applications that require extensive storage so it has little important means that is a difference as far as dealing with the data using machine learning deep learning or ai tools

So the ability to deploy ArcGIS or AWS, you say Amazon Web Services on web service because it has linkage with the ArcGIS in terms of Arc Logistics. Arc Logistics is a cloud application for optimizing routing. So this web service offered by Amazon has the ability to deploy it with ArcGIS server. Now let us look at the cloud computing vendors, their services and products. So this PaaS, SaaS and IaaS we just discussed in terms of platform, infrastructure and the software as a service.

Vendor	Service Type	Product
Google Inc.	PaaS	Google App Engine
Microsoft Corp.	PaaS	Windows Azure
Salesforce.com	PaaS	Force.com
Google Inc.	SaaS	Google Gmail
Microsoft	SaaS	Hotmail
Amazon Web Services LLC	IaaS	Amazon EC2, Amazon S3
Flexiant Ltd.	IaaS	Flexiscale
Servepath	IaaS	GoGrid Cloud Hosting
Rackspace	IaaS	Rackspace Cloud

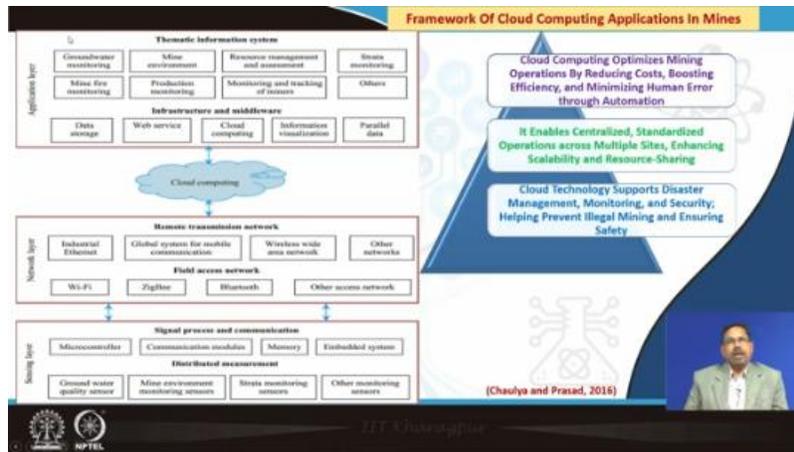
(Chaulya and Prasad, 2016)

So as far as different vendors are concerned, look at the left hand side column. The Google, INC, Microsoft Corporation, Salesforce.com, they all offer a kind of PS platform as a service. The products could be Google App Engine, the Windows, Azure and Force.com. Whereas this Google INC also use software as a service in terms of Google Gmail. Look at the red wash regions, the Microsoft and Amazon Web Service LLC, both of them use the service types like SaaS and IaaS.

So Microsoft uses SaaS, and Amazon Web Services LLC uses IaaS, Infrastructure as a Service. So, as far as Microsoft is concerned, the Software as a Service example is Hotmail, which we all use. Amazon Web Services LLC provides Infrastructure as a Service, such as Amazon EC2 and Amazon S3. Now, some other vendors like Flexient, Servepath, and Rackspace—all three of them—use Infrastructure as a Service as far as cloud computing is concerned. And the products could be FlexiScale, GoGrid Cloud Hosting, and Rackspace Cloud, respectively.

This has been taken from the publication by Chaulya and Prasad, which appeared in 2016. Let us take a critical look at the framework of cloud computing applications in mining. So, as far as thematic information is concerned, it may include mine fire monitoring, mining environment, production monitoring, resource mining resources monitoring and assessment, and monitoring and tracking of miners, etc. And the infrastructure and the mid-

level—or you could say the mid-layers—include data storage, web services, cloud computing, information visualization, and parallel data. Now, we have different network layers and sensing layers.



In terms of the network layers, we can have industrial Ethernet, which includes remote transmission networks. So, let us take a look at the right-hand side. We have cloud computing optimization in terms of mining operations, which can be achieved by reducing costs, boosting efficiency, thereby improving the efficiency of the mining operation. And minimizing human error through automation. So, cloud computing offers or helps us optimize mining operations by reducing costs.

It thereby boosts efficiency as far as production in the mining industry is concerned and also minimizes errors, particularly human errors, through many automatic applications or automations. This cloud computing enables centralized, standardized operations across multiple sites, thereby enhancing scalability and resource sharing. Cloud technology supports disaster management, monitoring, and security, helping prevent illegal mining and ensuring safety. So, friends, in a sense, when we have a lot of data and real-time connectivity, we can implement much automation using cloud computing. In turn, we can make our environment—in this case, the mining environment—more friendly as far as human operations are concerned.

We can provide a safer environment, thereby minimizing losses and optimizing efficiency. So, the benefits of cloud computing for big data analysis in mining can be summarized under these five broad points: real-time data analysis. Once data is available in real time, it helps us assess, manage, or support asset management. Cloud scalability enables real-time analysis of large datasets and supports asset management through IoT tracking. We will discuss IoT, the Internet of Things, in the next class.

Benefits of Cloud Computing for Big Data Analysis in Mining

- ❖ **Real-Time Data Analysis:** Cloud Scalability enables Real-Time Analysis of Large Datasets and supports Asset Management through IoT Tracking
- ❖ **Access to Enterprise Resource Planning (ERP):** The cloud provides ERP software access for managing Sales, Inventory, logistics, and Payroll from a Single Interface
- ❖ **Collaboration Tools:** Cloud-based Customer Relationship management systems facilitate Teamwork, Document Management, and Real-Time File-Editing with Automated Backups
- ❖ **Data Security:** IaaS ensure secure, redundant data centers with high uptime and protection against outages
- ❖ **Environmental Benefits:** Cloud virtualization increases energy efficiency and helps companies reduce their carbon footprint and energy costs (Chauha and Prasad, 2016)

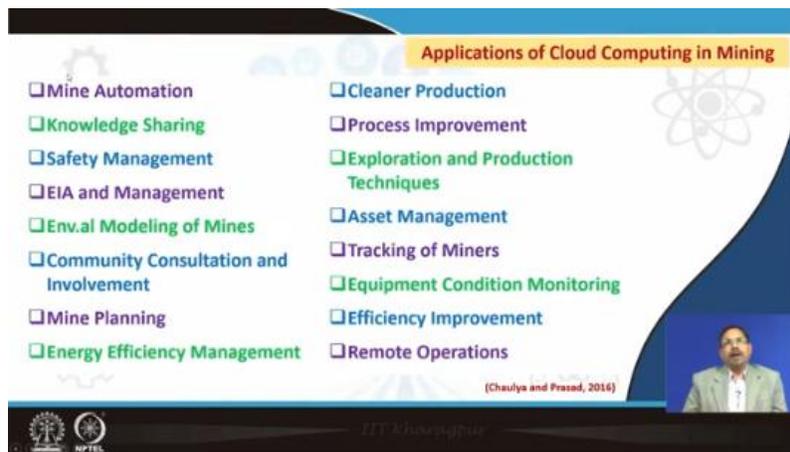
Cloud computing helps us assess ERP, Enterprise Resource Planning. It is a package—the Enterprise Resource Planning software. The cloud provides ERP software access for managing sales, inventory, logistics, and payroll from a single interface. If we are connected to cloud computing and have a lot of real-time data linked through high-speed internet across different servers, we can use this data in our ERP to manage activities like sales, inventory, stock, and products. Logistics and even payroll can be managed from a single interface.

And it has in big way as far as collaboration tool is concerned. So cloud-based customer relationship management systems facilitate teamwork, document management, And real time file editing with automated backups. So look at the thing that real time file editing, it edits the file on real time basis and also goes on keeping an automated backup. So it helps us in facilitating and making our life very, very easy and comfortable.

Now coming to the data security so in terms of infrastructure as a service IaaS the data security ensure secure redundant data centers with high uptime and protection against the outages. So if we have a lot of data and the security of the data is ensured through infrastructure as a service, thereby the data redundancy with high uptime and protection is taken care against the different outages. Now the last one is different types of environmental benefits. Cloud virtualization increases energy efficiency and helps companies to reduce their carbon footprint energy cost look at this once we have the virtualization yeah this this increases the energy efficiency and helps the companies the mining companies to reduce their carbon footprint in terms of

pollution or in terms of the climate and the energy costs. So that is how the climate the cloud visualization is really really important for us. Now let us have a look at the

applications of cloud computing in mining. So we can have different applications as far as the cloud computing is concerned in the mining sector. We can have mine automation.



If we have a program or code written to pull different data from different sources and we also have a connection for real-time data, then a lot of automation can be done. In that sense, it minimizes human error and optimizes productivity as far as the mining sector is concerned. Knowledge sharing through cloud computing allows knowledge to be shared across the internet space. Safety management: once we have the database and analytics, we can do a lot of good work. Management practices concerning safety, environmental impact assessment, and environmental modeling of mines can also benefit from cloud computation, which has the power to connect and consult communities.

Or to involve the community. So the feedback and inputs from the community, in terms of consultation and involvement, can also serve as a layer of information in this cloud computation. Thereby, it helps in mining planning and increases energy efficiency. And thereby, by increasing energy efficiency, it may also lead to a cleaner environment by minimizing CO₂ or methane emissions. So, pollution can be minimized, in turn giving us cleaner production.

The environment or the products we get are from clean sources. Process improvement can also be done using cloud computation because different data sets help us improve various processes. Independently, cumulatively. Then exploration and production techniques, assessment, asset management, and tracking of miners. If miners are connected through GPS or GNSS, they can be well-tracked, along with their environment and surroundings.

Then equipment conditions also can be monitored. If the different equipments are connected to different sensors, the conditions can be monitored periodically and some

action wherever it is required can be taken in advance before any disaster is met with. So in a sense, it helps in remote operation, which is very, very important as far as mining, both opencast and underground mining are concerned. So in that sense, the cloud computation has a very critical and important role to play in mining sector in terms of efficient management or improving the management efficiency. So, these are the six references which have been used to prepare these slides for you.



And at the conclusion, we understood that cloud computing transforms the mining industry by enabling efficient handling of large data, large database. The cloud computation provides on-demand access to computing sources. So, it is we need, it is the demand comes from us. So, it provides on-demand access. So, once the demand is there and you ask the database, the server that yes, I need this.



So, the demand is raised through the user so that it is accessed. And then reducing the reliance on local infrastructure. You don't need to build a huge infrastructure, particularly the database server infrastructure at your end, at the local end. So key platforms we have

been very well using Google Earth Engine, the Microsoft Azure, the Amazon Web Server. So these three that's why we selected to discuss with you.

And these key platforms help in terms of enhancing environmental monitoring, data processing, and also using machine learning tools. So, this cloud computing supports rapid deployment and global accessibility without extensive hardware. So, hardware cost is totally nullified at the local level or at the user end. Deployment models—yes, that could be both private and public, and if both of them are there, it could be in hybrid cloud computational mode. So, these deployment models offer flexible resource management.

You have some resources you want to share, and some resources you don't want to share. So, you have control, you have a clever or intelligent way of managing your resources and sharing them. And it has a connection as far as real-time data collection, generation, and analysis are concerned, as well as data security and streamlined operations. As I just mentioned, you can have data security with you, the data you want to share and the data you do not want to share. And in a sense, cloud computing helps in EIA or environmental impact assessment studies, thereby supporting sustainable mining practices.

So, in totality, cloud computing facilitates improved decision-making and community engagement. So, this is one of the advanced tools—an advanced geo-information application domain—we are using for the management or improvement of mining industry management. Thank you very much.