

Building A Cloud Data Analytics Platform

**Creating Data Tools Analyts And Data Scientists Can Easily
Use**

CONTENTS

3	EXECUTIVE SUMMARY
4	INTRODUCTION
5	THE PROBLEM AND SOLUTIONS
6	PROS AND CONS OF USING CLOUD
	PROS OF CLOUD STORAGE
	CONS OF CLOUD STORAGE
7	TOP REASONS FOR CLOUD MIGRATION
8	BEST CLOUD SERVICE COMPANIES
9-10	SWITCHING DIFFERENT INFRASTRUCTURES AND BENEFITS
	EC2 FROM SERVERS
	RDS FROM STANDARD MYSQL OR POSTGRES
	REDSHIFT FOR DATA WAREHOUSING
	COGNITO VS HAVING TO CODE AUTHORIZATION AND AUTHENTICATION 7
	LAMBDA INSTEAD OF CREATING AN ENTIRE API 7
11	COST SAVINGS OF CLOUD 8
	HARDWARE 8
	DEVELOPERS/MAINTENANCE 8
	RELIABILITY AND SCALABILITY 8
12	CONCLUSION

EXECUTIVE SUMMARY

The goal of this paper is to outline the benefits of switching or developing your companies technical infrastructure to cloud.

We will go over the pros and cons as of investing in cloud infrastructure. From there we will discuss the major reasons to pick cloud and some of the big players in the industry.

We will also be outline the tools cloud service providers like Amazon and Microsoft offer. Like EC2, Lambda, RDS and more. This will help provide insights into what services you may need as a small or medium size business owner.



PHOTO BY AUSTIN
DISTEL ON UNSPLASH

INTRODUCTION



PHOTO BY TAYLOR VICK ON UNSPLASH

In Internet-based computing, cloud storage is a trend that utilizes a shared computing system involving numerous computers that work on a specific network to accomplish a particular task on demand. Moreover, cloud storage is used to store end-users' data within the cloud without using a local system; through network connectivity, this data can be accessed anywhere and client services can be provided.

Cloud computing is one of today's most exciting technologies because of its capacity to lessen costs associated with computing while increasing flexibility and scalability for computer processes.

During the past few years, cloud computing has grown from being a promising business idea to one of the fastest-growing sectors of the IT industry.

PROS AND CONS OF CLOUD

PROS

Simplicity of administration: The support of the programming, general infrastructure, and hardware used to buttress stockpiling is definitely improved by an application in the cloud. Applications that take the favorable cloud-based form are generally far less demanding to set up and maintain than a proportionate level of administration installed on the premises.

Cost-effectiveness: Cloud storage is useful for mitigating ownership fees. Removing costly systems and the requirement for the client to maintain them normally gives organizations noteworthy cost reserves that more than counterbalance the charges for cloud storage.

The methodology of storing information in remote cloud servers is known as cloud storage. Storing on the cloud is far better than other conventional storage strategies. A portion of the explanation behind that is: - Companies do not have to establish physical storage devices at their own server farm or workplace; -

Lower impact failures and upgrades: Cloud computing typically delivers cost-effective storage hardware redundancies. This results in uninterrupted service during a scheduled or unplanned breakdown. This also applies to hardware upgrades, which will no longer be visible to the client;

Simplified layout: Cloud storage solutions free up the capacity for the IT director of Detailed Planning. Flexible cloud-based storage solutions are provided as required, eliminating the need for more storage that can be required to accommodate them.

CONS

Leaks and data access: without permission between virtual devices operating on the same server;

Code Changes: Errors on the part of a cloud supplier in handling the correct management and saving of sensitive data;

Off-Premise: the cloud service may be unavailable for extended periods of time due to errors and system crashes; Of course, this also can occur when you own servers and your team has to fix it.

Reliability issues: Truth be told, when you use the cloud, you are stuck relying on a third-party for your companies infrastructure. This means if things go down, then you just have to wait until said third-party brings them back up. However, this issue is often more common on-premise vs in the cloud.

TOP REASONS FOR CLOUD MIGRATION

Control and reduce costs

Virtualization in the cloud allows business owners to control costs on installation, hardware and devices, and frequent upgrades. It is also invaluable for small businesses to save money on frequent IT repairs. Without large IT budgets SMBs are often forced to rely on pay-to-fix (break-fix) repair services and patches that don't resolve the real challenges.

Improve employee morale and collaboration

Smart business owners know the critical role that employee morale and interactions play in a company's success. Implementing the cloud encourages users to get out of the office to network, to join meetings, and to work remotely. In the cloud, employees have secure access to all the resources, tools and systems necessary for their job without the risk of failure, loss or intrusion. The cloud also helps attract the best talent for the company, regardless of geographic location. The same access, tools and communication features that benefit existing employees are equally important to bring the right talent onboard without exorbitant relocation costs or losing the talent completely.

Time and ease of migration

An IT analyst can get a company situated in the cloud with almost no downtime. This is because most of the work is done online, and large scale overhaul isn't necessary to migrate. From the time a company decides to migrate to the cloud to the moment they go live, the switch can take as little as four weeks. As a company grows, scalability of the cloud service becomes a strategic advantage as you configure your cloud setup to meet ever-changing demands, priorities, and resources.

TOP CLOUD PROVIDERS

Amazon Web Services is a market leader in cloud services. As one of the first providers to not only spearhead cloud computing itself but also really transform it into much more than just virtual machines and data storage. They have also developed the concept of serverless models. AWS is clearly ahead of its competitors in this category. AWS Lambda, Amazon API Gateway, Amazon S3, and Amazon Redshift are just a few of the core products AWS offers.

Azure is Microsoft's Cloud arm. It is slowly catching up with AWS thanks to fast moving developers and partnering with Accenture to further sell their services. Microsoft's cloud seems to follow a similar paradigm as their operating system. Develop easy to use products that don't allow for too much flexibility but don't require massive technical know-how.

GCP is Google's cloud platform which got in early with the other cloud services has many similar options to AWS and often can be compared one-to-one. For example, AWS has Redshift, GCP has Big Table and Big Query. Where Google has Big Compute Engine, AWS has EC2. Although in general AWS tends to offer more bells and whistles, both are pretty comparable for a large portion of your cloud needs.

Oracle Cloud Although late to the game, Oracle Cloud offers many similar services to Azure and AWS. However, Oracle more closely focuses on the data side of their cloud services. Offering advanced data management options like their autonomous database that utilizes machine learning to help manage your systems.

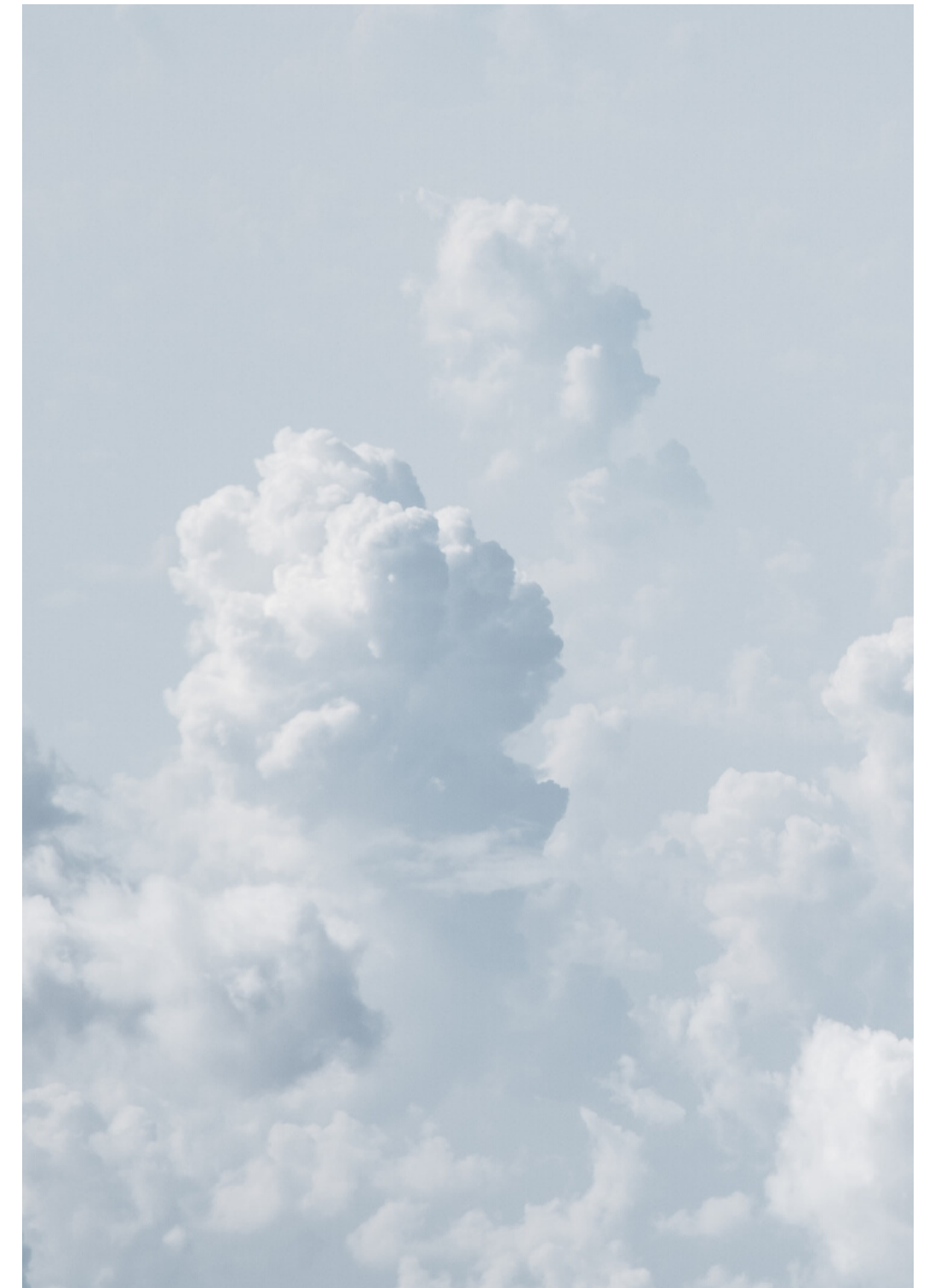


PHOTO
BY ELCARITO ON UNSPLASH 07

RDS Instead Of MySQL or Postgres

Amazon Aurora is a high performance, highly available, and enterprise-grade database built for the cloud. Leveraging Amazon Aurora can result in better performance and greater availability than other open-source databases and lower costs than most commercial grade databases. Although there might be some need to migrate, our team of cloud consultants have migrated many systems and databases and have experience setting up new

Redshift For Data warehousing

Amazon Redshift has changed how enterprises think about data warehousing by dramatically lowering the cost and effort associated with deploying data warehouse systems without compromising on features and performance.

Amazon Redshift is a fast, fully managed, petabyte-scale data warehousing solution that makes it simple and cost-effective to analyze large volumes of data using existing business intelligence (BI) tools. With Amazon Redshift, you can get the performance of columnar data warehousing engines that perform massively parallel processing (MPP) at a tenth of the cost. You can start small for \$0.25 per hour with no commitments and scale to petabytes for \$1,000 per terabyte per year.

EC2 For Servers

Technology changes continuously, and you often have no choice about updating or upgrading. That can make it tempting to delay moving cloud applications that are functioning fine where they are. However, as we learned in our study, by keeping your workloads in EC2 instances based on older architecture, you're leaving value on the table.

It's been found between the lower cost and improved performance of next-generation EC2 instances based on Intel Xeon Scalable Processors using AVX-512, you could double your performance per dollar by moving some applications and re-optimizing them.

For some applications still using instructions for Intel AVX2, moving to the new instances could boost your performance per dollar by 21.9 percent.⁷ These numbers make it clear that moving your workloads is a good idea.

Cognito Instead Of Developing Authorization

Amazon Cognito is Amazon Web Services' service for managing user authentication and access control. Although it was originally associated with AWS's mobile backend-as-a-service offering (MBaaS), it has recently gained the attention of the serverless crowd, who are looking for ways to offload user management concerns to a service provider.

Cognito solves this problem by providing a fully managed, scalable and cost-effective sign-up/sign-in service — but at the cost of a steep learning curve. One of the reasons for this is because Cognito is comprised of two services — User Pools and Identity Pools (a.k.a. Federated Identities) — that are similar on the surface but different under the hood. These two services solve the same problem (i.e. authentication and authorization) but do so in very different ways. They can also be used separately or together, providing both flexibility and a source of confusion at first.

Lambda Instead Of An API

AWS Lambda is an event-driven, serverless compute service that extends other AWS services with custom logic, or creates other backend services that operate with scale, performance, and security. Lambda can automatically run code in response to multiple events, such as HTTP requests through Amazon API Gateway, modifications to objects in Amazon S3 buckets, table updates in Amazon

Amazon S3 Instead Of SFTP

Amazon's Simple Storage Service (Amazon S3) is a cloud storage service that allows you to interface with your stored objects using REST and SOAP. S3 provides access to a storage system that is fast, reliable, scalable, and inexpensive data storage infrastructure. Several client types, big or small, can make use of its services to storing and protecting data for different use cases.

Amazon S3 offers an object (which are essentially files) storage service with features for integrating data, easy-to-use management and everything else cloud often offers. It can essentially act as a type of file server that can manage your company's content for your website like videos and photos or be used to develop a data layer for your analytics. The platform makes data organization and configuration flexible through adjustable access controls to deliver tailored solutions. Overall, one of the biggest reasons many companies turn to S3 is because of its cost.

CLOUD COST SAVINGS

Hardware

An advantage of cloud computing is the reduction in hardware cost. Instead of purchasing in-house equipment, hardware needs are left to the vendor. For companies that are growing rapidly, new hardware can be a large, expensive, and inconvenience. Cloud computing alleviates these issues because resources can be acquired quickly and easily. Even better, the cost of repairing or replacing equipment is passed to the vendors.

Along with purchase cost, off-site hardware cuts internal power costs and saves space. Large data centers can take up precious office space and produce a large amount of heat. Moving to cloud applications or storage can help maximize space and significantly cut energy expenditure.

Developers/Maintenance

Cloud solutions can also lead to a dramatic decrease in labor and maintenance costs. As a result of the hardware being owned by vendors and stored in off-site locations, there is less demand for in-house IT staff. If servers or other hardware need repairs or upgrades, it is the responsibility of the vendor and doesn't cost your company any time or money. Eliminating routine maintenance can free your IT staff to focus on important initiatives and development. In some cases, this could even mean reducing staff size. For companies lacking the resources for an in-house IT staff, the cloud will help eliminate costly third party hardware repair bills.

Reliability and Scalability

Cloud computing helps IT enterprises to optimize and secure application performance in a cost effective manner. Besides security, cloud providers are also responsible for reliability and availability, because all users expect the highest level of QoS (Quality of Service). The cloud providers use some solutions such as partitioning to achieve maximum performance. But according to whether the cloud is based on public, private, or hybrid, the management and control of these performance parameters from RAS viewpoint will vary. Cloud-based applications are based on network software running on a virtual machine in a virtualized environment.

Cloud providers often have several powerful servers and resources that provide appropriate services for their users, but the cloud is at risk to a degree similar to that of other Internet-based technologies. Unfortunately, there are some attacks for which no perfect defense exists such as a powerful DoS attack. But as paper discussed in occurrence of DoS attacks, cloud may be a good solution or mitigation because cloud providers can use mirrors or devote more resources to protecting against attacks. However, this solution's performance depends on provider facilities.



One of the biggest benefits of cloud technology is in the ability to take advantage of billion dollar level technology without hiring five engineers and spending hundreds of thousands on servers.

By having cloud-based software, like RDS, EC2, and remote data storage, you get the advantage of seamless transitions from computer to computer.

Log into any computer, and your data is readily available. With remote storage, you've taken the burden of data recovery off your shoulders and put it onto a more experienced provider, whose bottom line depends on your quick recovery.

Running your business in the cloud gives you an edge in your space. Your applications run faster, are accessible from anywhere, and you are freed from the maintenance of updating your aging infrastructure. With cloud technology, you can scale quickly based on your needs, giving you an advantage over your non cloud competitors.

Services

CLOUD ARCHITECTURE AND DEPLOYMENT

Utilizing the cloud allows you to take advantage of the same technology that major tech companies are using at a fraction of the cost.

ANALYTICS AND AUTOMATION

Our team can help you automate workflows and provide analytics to help optimize your business and provide key insights.

DATA WAREHOUSING AND MACHINE LEARNING

Your business has data in all different shapes and forms. Creating a centralized place where all that data can live and you can answer questions allows you to develop analytics and machine learning models quickly and easily



Work with us

WEBSITE

www.logitanalytics.com

EMAIL

info@logitanalytics.com