REGCENTRIC

FUTURE OF DATA IN CLOUD

NOVEMBER **2020**

Considering the future of data in your cloud migration strategy

Cloud based IT services is certainly a trend that has further accelerated since the COVID-19 crisis. In this report, we explore some of the key considerations around the future of data (a critical asset) in cloud for any organisation that is defining their cloud migration strategy.

The need for Innovation during a crisis

The economic impacts of COVID19 which was preceded by an extremely competitive business environment has resulted in some extraordinary challenges for organisations to survive in the industry. The crisis has also created some unexpected shifts in consumer behaviors and expectations. For example, remote working and social distancing has created radical shifts in demand from one industry/distribution channel to another such as significant increase in online shopping. In times like this, access to high quality data for decision making and a well-established digital technology-based business model is crucial for survival and growth for many industries.

> # Innovation distinguishes between a leader and a follower
> ~ Steve Jobs

Historically, any established organisations that failed to embrace innovation and emerging market trends get superseded by those organisations who are not afraid to constantly reinvent themselves. Organisations such as Borders, Blockbusters and Kodak are examples for how once successful and leading organisations that failed to adapt to industry trends, became extinct.

Migration to Cloud based services is one such emerging trend that is being adopted by many organisations. A further acceleration of this trend can be anticipated in the coming years as the longer-term impacts of the economic crisis unfolds and organisations become more cost conscious and aim for flexible technology services models. Whilst this is a great step in the right direction, it is important for organisations to take a holistic approach to the overall transformation to get the maximum return on investment.
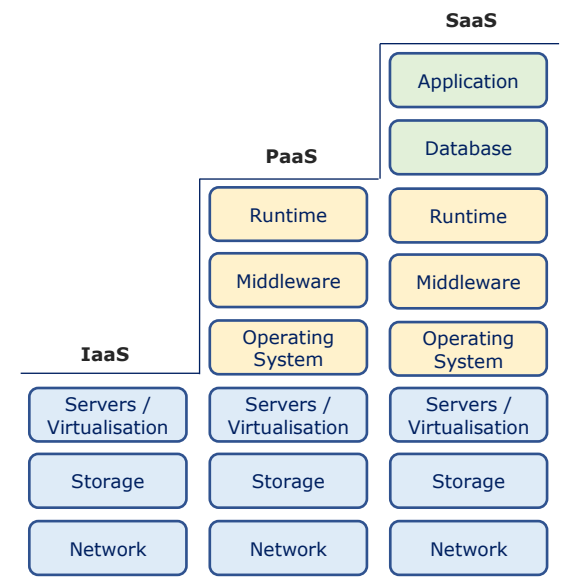
An Overview of Cloud Services

There are a 3 major cloud-based services that every organisation that is considering cloud enabled digital transformation must evaluate as part of their strategy. These are Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). The common theme across all of them is their pay-as-you-consume cost model which can be quite cost effective. The diagram provides an overview of Cloud offerings.

Infrastructure as a Service (IaaS): Provisioning of core IT infrastructure elements such as computing resources, servers, network, storage via cloud. This offering enables management of infrastructure as a self-service option on the cloud. IaaS takes away the need for organisations to purchase hardware and maintain data centers. Instead, organisation can commission infrastructure on demand both efficiently as well as in a cost-effective manner.

SaaS

IaaS	PaaS	SaaS
		Application
		Database
	Runtime	Runtime
	Middleware	Middleware
	Operating System	Operating System
Servers / Virtualisation	Servers / Virtualisation	Servers / Virtualisation
Storage	Storage	Storage
Network	Network	Network

Typically, IaaS is delivered via virtualisation technology and made available through APIs or dashboards. Amazon Web Services (AWS), Google Compute Engine, Microsoft Azure are some of the leading IaaS offerings. IaaS provides a cost-effective alternative to organisations looking to reduce total cost of ownership and create flexibility to scale up or down based on changing demand.

Platform as a Service (PaaS): Provisioning of a development framework allowing developers to efficiently develop applications on the cloud. Organisations have the choice to manage the infrastructure themselves or use a third party however the infrastructure is also hosted on the cloud to enable development.

An Overview of Cloud Services

PaaS (contd.): Essentially this offering enables development to be carried out on the web hence eliminating any platform requirements such as operating system, or infrastructure etc. In addition, it also simplifies development due to the nature of the development frameworks containing several reusable development components. AWS Elastic Beanstalk, OpenShift and Google App Engine are some of the major providers of PaaS.
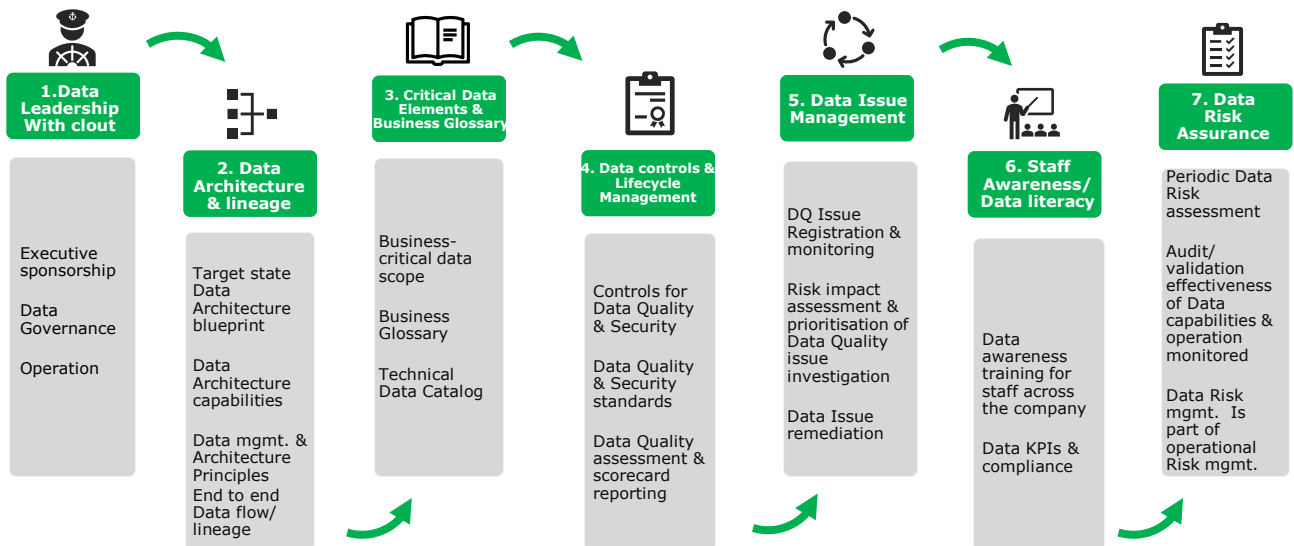
Software as a Service (SaaS): Provisioning of software applications via the cloud is SaaS. This is one of the most common forms of cloud services out there. Many software vendors have adopted this approach to provide an alternative offering to their customers. Most software's hosted on the cloud are accessible via web browser. Popular examples of SaaS offerings are Microsoft Office365, Google Apps, Cisco WebEx, Zoom etc.

Considering the future of your data in cloud

Migrating to Cloud without embedding the right Data architecture and Data Management capabilities can prove to be an expensive endeavor resulting in limited value. Considering the variety, volume and speed of Big Data, the data quality issues and risks will be multi-fold if appropriate controls are not put in place.

> Data Management in Cloud introduces a new set of challenges to an already challenging discipline as data grows exponentially in this digital age.

RegCentric has defined Seven Data Disciplines to deliver effective Data Management for organisations in a practical and sustainable way, with people, technology and process controls, to securely harness and enhance the data to unlock its full value for business outcomes. The Data disciplines includes consideration on how data is stored, validated, accessed, and managed throughout its lifecycle, across the organisation, for its maximum value with minimum cost and risk.

| 1. Data Leadership With clout	2. Data Architecture & lineage	3. Critical Data Elements & Business Glossary	4. Data controls & Lifecycle Management	5. Data Issue Management	6. Staff Awareness/ Data literacy	7. Data Risk Assurance
Executive sponsorship	Target state Data Architecture blueprint	Business-critical data scope	Controls for Data Quality & Security	DQ Issue Registration & monitoring	Data awareness training for staff across the company	Periodic Data Risk assessment
Data Governance	Data Architecture capabilities	Business Glossary	Data Quality & Security standards	Risk impact assessment & prioritisation of Data Quality issue investigation	Data KPIs & compliance	Audit/ validation effectiveness of Data capabilities & operation monitored
Operation	Data mgmt. & Architecture Principles End to end Data flow/ lineage	Technical Data Catalog	Data Quality assessment & scorecard reporting	Data Issue remediation		Data Risk mgmt. Is part of operational Risk mgmt.

©RegCentric

Seven Data Disciplines

In addition to managing data through right disciplines, governance & controls, organisations must also consider emerging technology trends around how data is being captured, processed, retained and accessed. "Lifting and Shifting" solutions from legacy, as-is, on to Cloud platform defeats the business objective of cloud enabled transformation and the expected return on investment.

Key Tech Trends to Fast-track value from Data in the Cloud

With the significant digital advancement, there are many new technology and data trends emerging in the industry to help organisations achieve maximum value from Data and Analytics. Of which, following are five key trends associated with data, that are fast becoming part of the mainstream, and are fundamentally turning the dial on organisation's productivity and opportunity, and business value being created from its data.

1. Streaming & Real-time Data enrichment and Calculations
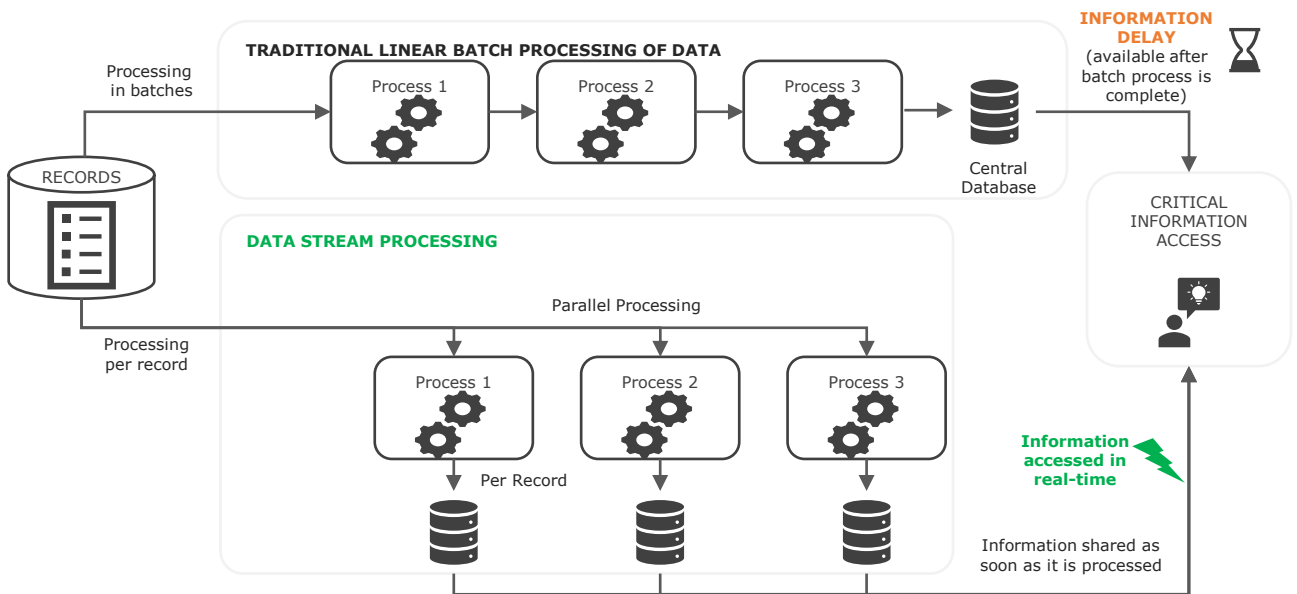
The legacy approach to advanced data processing requirements such as risk and financial calculations, based on batch processing with monolithic applications takes hours to run, and takes several months to implement changes. This is going to be increasingly challenging to sustain in the current volatile environment with increasing risks and regulatory scrutiny.

Organisations should leverage from Cloud Platform's that are built from the ground up using microservices architecture and containerised development framework, leveraging latest technologies that can empower creation of smaller IT teams to build, package and deploy highly flexible, scalable and change agile modularised calculation/enrichment applications. It will significantly increase the throughput, at a much lower cost. Utilising real-time streaming capability to perform data calculations/enrichment and analytics will significantly reduce the overall time latency in generating insights.

Several vendors are currently investing heavily on such Cloud enabled real-time calculator applications to generate risks and regulatory measures and continuously updating the rules in alignment with changing regulations. Financial institutions should look to either leverage from such vendor solutions or build real-time calculation/ enrichment applications on PaaS, based on their IT resource skills and capacity.

2. Artificial Intelligence and Machine Learning

Artificial intelligence (AI) refers to a broader concept whereby machines can perform tasks independently and intelligently learning like humans without human support, using data, applying algorithms and making smart decisions to act upon. Machine learning (ML) is an application of artificial intelligence (AI) that allows machines to learn themselves using data. ML helps machines to create the algorithms completely independently (unsupervised ML) or with help from people (supervised/semi-supervised) depending on the complexity of the expected outcome.

2. Artificial Intelligence and Machine Learning

Below are some real-world examples where organisations are maximising the combination of real-time processing and AI/ML capabilities:

Fraud Detection: American Express analyses real-time credit card transactions, registering fraudulent characteristics for fraud detection algorithms using ML techniques with a wide breadth of data. It enables accurate Fraud detection in real-time, saving significant costs in remedial actions and avoid false positives, minimising impact on customer experience and confidence.

Risk Management: HSBC has launched ML based 'country-by-country liquidity reporting' on a Cloud platform. HSBC CIO Darryl West states the liquidity reporting algorithm gives HSBC a country-by-country view in "minutes not hours", providing the bank with a significant productivity boost.

Customer experience: Applying AI/ML pattern recognition algorithms, across a variety of customer data (incl. social media) helps to understand customer preferences and provide tailored services to their specific needs. United Overseas Bank (UOB), Singapore, has launched an AI-based simpler and smarter digital banking service for its customers to manage their finances.

3. Augmented Data Management

> By the end of 2022, data management tasks will be reduced by 45 percent through augmented data management.
>
> ~Analyst house Gartner, Inc

Augmented data management refers to technical and data solutions that consists of in-built AI/ML functionality to intuitively automate Data management capabilities such as Data Quality issue detection, Data remediation, Metadata management and Master Data Management, with algorithms that can self-learn over time. It significantly improves efficiency of these Data management activities which currently have heavy reliance on limited individual subject matter experts and are labour intensive.

Examples of Augmented Data management includes:

❑ **Data Catalogue** tools creating data dictionary inferring from data usage patterns, technical metadata that are automatically extracted from database schemas etc. using AI/ML algorithms.

❑ **Data Standardisation, cleansing and match-merging** using algorithms that can create matching rules based on data content, self-learn and refine logics over time

❑ **Developing and maintaining data lineage** automatically based on data integration logic, operational logs etc.

This allows federation of data management activities in a consistent way, to end users with less technical skills to self-service, based on well-defined standards and delivery patterns while maximising the highly technical resources in innovation, establishing new patterns and strategic opportunities.

4. Data Fabric

Data fabric is a combination of data integration and data management technology designed to ease the complexities of integrating a variety of data at different velocity across different platforms (on premise and cloud), with automated data management controls end to end. It provides frictionless access and sharing of data in a distributed data environment – i.e. It enables streamlined and consistent data management framework, allowing seamless data access, ingestion, integration and sharing data by design, across otherwise siloed storage.

Data Fabric includes packaged connectors to data sources, removing coding effort to extract data. Its built-in data management capabilities such as unified data semantics/business glossary, data cleansing, standardising, and preparation, have ML enabled smarts to infer from repeated patterns (i.e. Augmented Data management). Hence significantly reduces the time, cost and number of highly specialised IT resources required for data integration and transformation initiatives. Through API enabled functionalities, it supports data distribution across internal and external consumers.

5. Graph Databases

Traditionally, relational databases have been used to deploy data entity relationships with defined logical and physical data model. This requires data designers to understand data, its structure, entity relationships based on use-cases, and types of questions the end-user will ask to retrieve data. With wide breadth and distribution of Big Data, and ever-changing environment where end-users can't easily envisage the questions they should ask, there is a need to test and learn to help them zone in and know how to maximise the value of data. Relational databases and data model designs will not be effective or flexible enough to meet these types of emerging requirements around making the most of data.

The real power of graph is that the in-built functionality and the database capability can build the data relationships across structured and unstructured data without a lot of effort in logical/physical data model design.

Graph database nodes:
- Label : Customer, Property : Mary
- Label : Account, Property : 1234
- Label : Customer, Property : John
- Label : Collateral, Property : Home
- Label : Banker, Property : Harry

Relationships: Connected To; Is primary customer of; Is primary owner of; Manages relationship of

When querying data, relational data base will have performance challenges where there is wide range of related data sets to JOIN, whereas graph database performs such queries much faster, even with extremely high volume and breadth of data.

With graph's built-in intelligence and ability to handle relationships at scale, greater flexibility, efficiency, and ability to navigate across complex data landscape, it provides robust, real-time capability for use-cases such as below

❑ Fraud detection using large volume of transaction data

❑ Track sensitive data for Data privacy and Confidentiality regulatory compliance

❑ Manage Identity and Access across platforms for different roles, rules, and resources
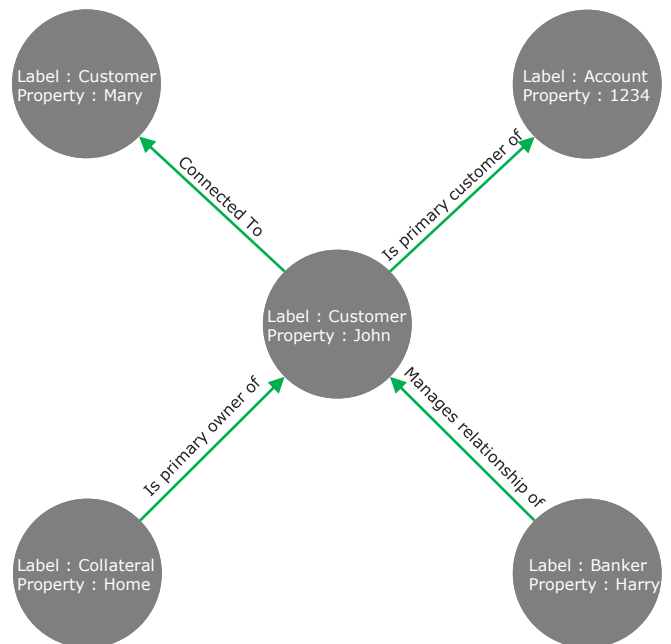
Conclusion

Data and digital transformation aligning to rapidly evolving digital disruption, is inevitable for organisations to keep up with changing customer needs, regulatory expectations, managing cost pressures and to survive in highly competitive and volatile environment.

Current COVID-19 with its unprecedented economic impact, organisations are struggling to survive the tide, facing high risks and extreme challenges to sustain revenue through existing products and service offerings. Innovation and cost efficiency have become extremely critical to see the light at the end of the tunnel.

Cloud migration with the adoption to new data and technology trends will generate greater efficiency and flexibility to the organisation and will release capacity and greater enablement for innovation in data, analytics, and digitisation. This will open opportunities for organisations to re-imagine their products, service offerings and business model, to sustain and scale its growth in this challenging environment.

About Us

RegCentric delivers innovative services and solutions that leverage the latest advances in technology to increase efficiency, provide insights, reduce risks, and ensure regulatory compliance. We provide strategic advice and design, build, and implement technology solutions that deliver the best outcome in a technology-agnostic way. Our team consists of data, risk & regulatory technology experts with decades of experience supporting transformative initiatives for regulators as well as the regulated industry.

How RegCentric can assist

RegCentric can assist organisations gain control over their data by implementing DataTech solutions that leverage latest advances in technology as well as by providing subject matter expertise on proven methodologies for managing exponential growth of data.

Please get in touch with us if you would like to know more about our services:

Advisory Services: Data Strategy, Data Architecture and Data Management transformation roadmap definition, Data Governance & Controls, Technology recommendation for managing data.

Implementation Services: Take advantage of next gen technology to quickly and cost effectively implement Cloud, Data and Analytics capabilities delivering high performance, flexibility & agility.

www.regcentric.com

LinkedIn

info@regcentric.com

02 8091 7187

Level 24, Three International Towers | 300 Barangaroo Ave | Sydney, NSW 2000