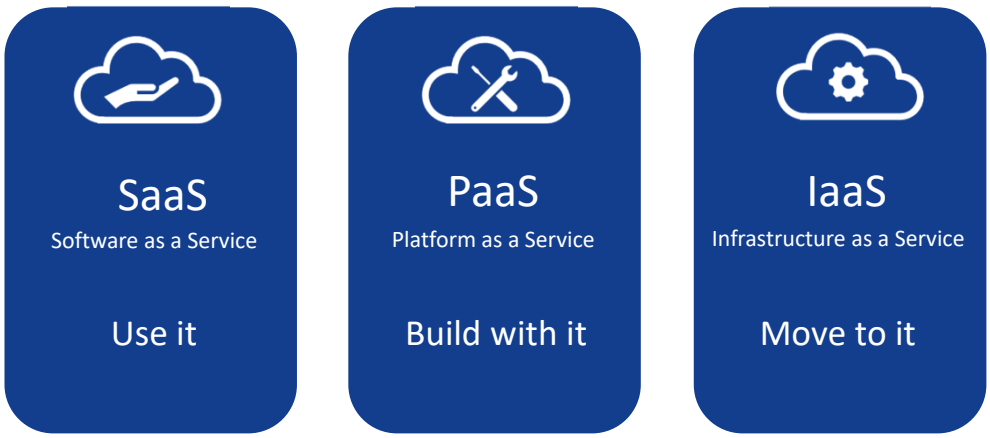
University of South Australia

Big Data in the Cloud

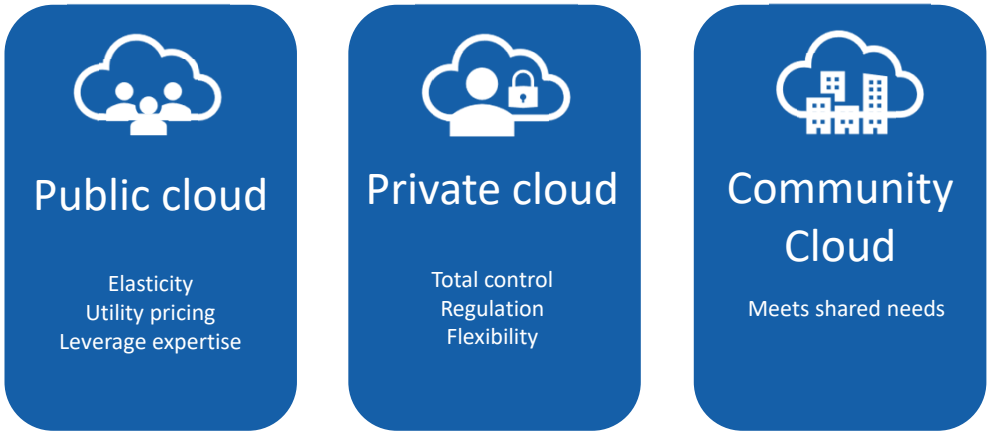
1

Cloud service models



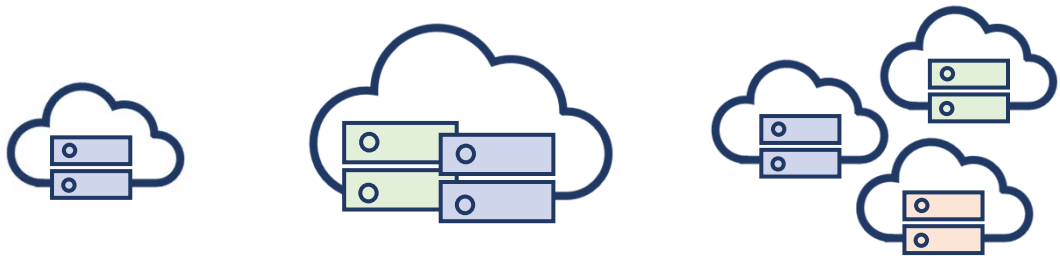
2

PaaS deployment types



3

Deployment options



4

Cloud in the enterprise

81% of enterprises have at least one application in the cloud, or a portion of their computing infrastructure already in the cloud

2020 IDG Cloud Computing study

5

Why public cloud?

Many advantages compared to traditional on-premises data centres:

- Removes many of the barriers
- Eliminates many of the hardware and labour costs
- Enables complete flexibility, agility and scalability

6

Some hard lessons for early adopters

Without a well defined 'Enterprise Data Strategy':

- New data and analytics silos can get created
- Harder to manage
- Costly to operate

7

More than a ploy to reduce IT costs

60%	Increased business agility
51%	Cost reduction or flexibility
38%	Accelerated innovation
37%	Ability to access, analyse and act on data
35%	Better system reliability

8

Ongoing challenges to realising these goals

- **40%** Integration complexity
- **37%** Diverse skill sets required
- **36%** Service integration and management
- **26%** Regulatory compliance
- **27%** Lack of standarisation

9

Which platform do Data Scientists use?

48.2% amazon web services

35.3% Azure

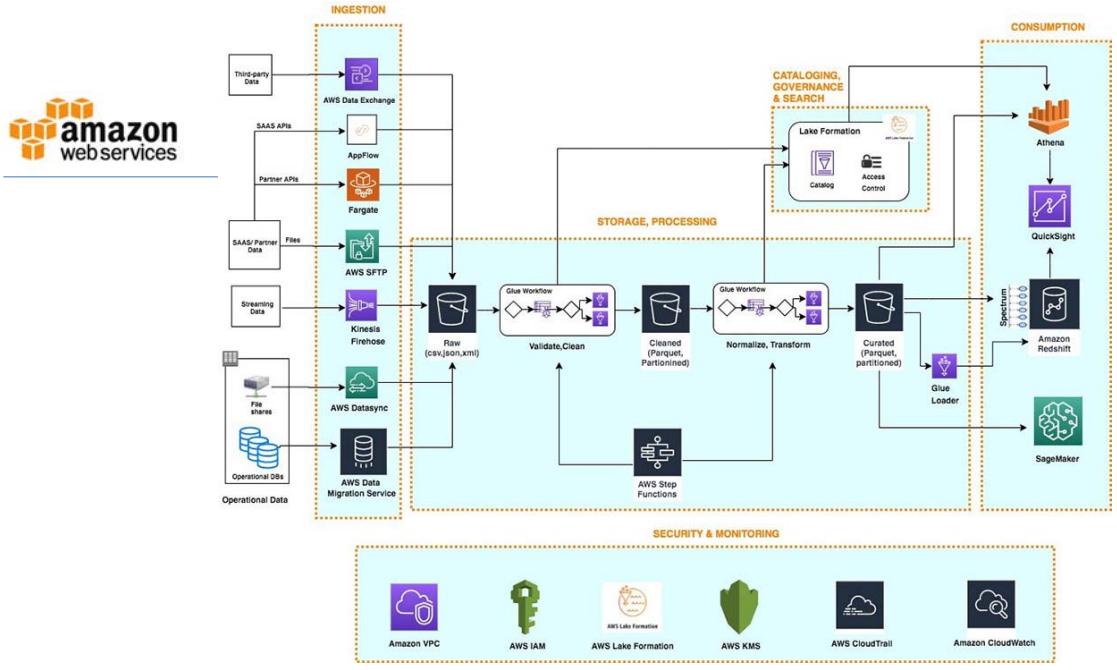
29.4% Google Cloud

10

Clear leaders

Gartner report, ID: G00441742
1 September 2020

11



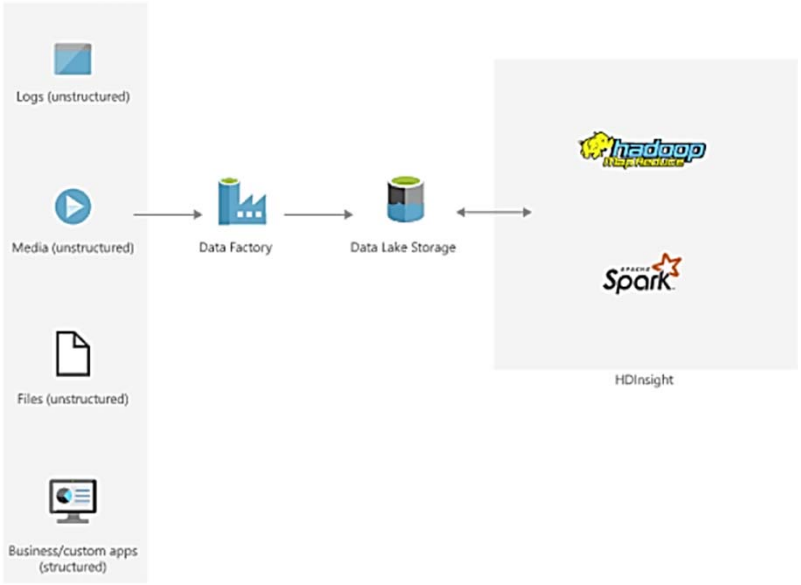
https://aws.amazon.com/blogs/big-data/aws-serverless-data-analytics-pipeline-reference-architecture/

12

Microsoft Azure

HDInsight

Customizable service for open-source analytics



13

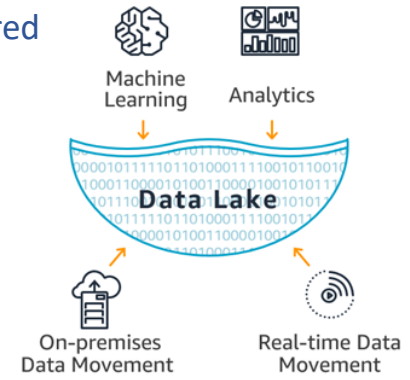
Microsoft Azure

Connectors

14

Data stored in a data lake

- Centralized repository for storing structured and unstructured data at any scale
- Data is stored as-is (schema-on-read)
- Different types of analytics:
  - Dashboards and visualizations
  - Big Data processing
  - Real-time analytics
  - Machine learning



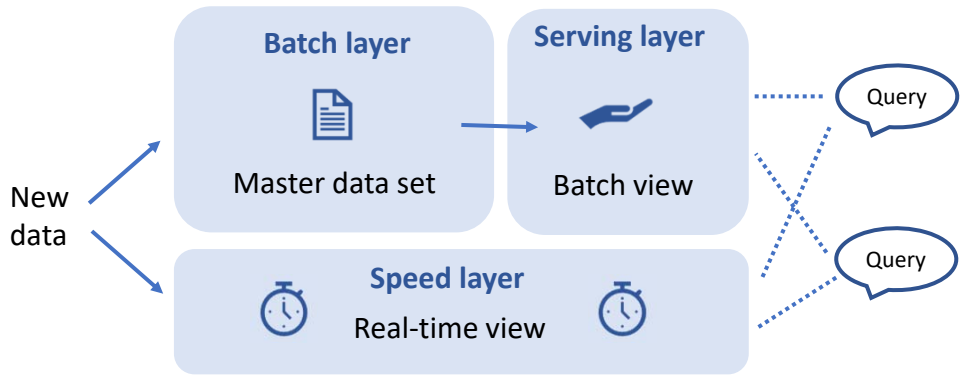
15

Serverless lambda architecture

- A large-scale, distributed data processing system:
  - Flexible and extensible
  - Fault-tolerant against hardware failures and human mistakes

16

Lambda architecture has three layers



17

Benefits of lambda architectures

No server management

Flexible scaling

Automated high availability

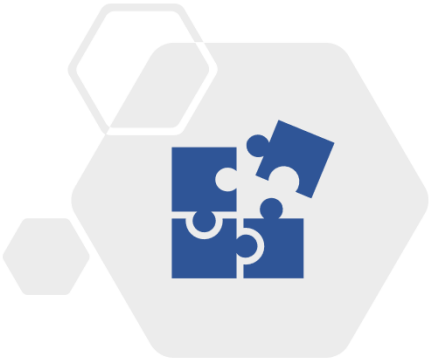
Business agility

18

Challenges with lambda architectures

Complexity:

- Typically, two separate code bases for batch and streaming layers



19

20