Cloud Computing Architecture for Artificial Intelligence
(人工智慧雲端運算架構)

Host: 閔立中, TKU AI創智社社長
Time: 18:00-20:00, May 24, 2022 (Tuesday)
Place: TKU All Club, MS Teams

戴敏育 副教授
Min-Yuh Day, Ph.D, Associate Professor

國立臺北大學 資訊管理研究所
Institute of Information Management, National Taipei University

https://web.ntpu.edu.tw/~myday

2022-05-24
戴敏育 博士
(Min-Yuh Day, Ph.D.)

國立臺北大學 資訊管理研究所 副教授
中央研究院 資訊科學研究所 訪問學人

國立臺灣大學 資訊管理 博士

Publications Co-Chairs, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013- )

Program Co-Chair, IEEE International Workshop on Empirical Methods for Recognizing Inference in TExt (IEEE EM-RITE 2012- )

Publications Chair, The IEEE International Conference on Information Reuse and Integration for Data Science (IEEE IRI)
Outline

• Cloud Computing Architecture for AI
• AWS Certifications Roadmap for Machine Learning
• AWS Serverless Architecture
Cloud Computing Architecture
Gartner Magic Quadrant for Cloud Infrastructure and Platform Services

- Amazon Web Services
- Microsoft
- Google

Software as a service

Multi-tier client-server architecture

Service-oriented Architecture

Client 1 -- Web Server -- Service gateway -- S1, S2, S3, S4, S5, S6 -- Client 2, Client 3, Client ...

VM

Virtual web server

Virtual mail server

Hypervisor

Host OS

Server Hardware

Container

User 1

Container 1

Application software

Server software

Container manager

Host OS

Server Hardware

User 2

Container 2

Application software

Server software

Everything as a service

Infrastructure as a service (IaaS)

Platform as a service (PaaS)

Software as a service (SaaS)

Cloud data center

Storage Network

Cloud management

Monitoring

Photo editing

Logistics management

Database Software development

Computing Virtualization

Artificial Intelligence (AI)
AI, NLP, ML, DL

Text Analytics and Text Mining

Source: Ramesh Sharda, Dursun Delen, and Efraim Turban (2017), Business Intelligence, Analytics, and Data Science: A Managerial Perspective, 4th Edition, Pearson
Evolution of Decision Support, Business Intelligence, and Analytics

The timeline in Figure 1.8 shows the terminology used to describe analytics since the 1970s. During the 1970s, the primary focus of information systems support for decision making focused on providing structured, periodic reports that a manager could use for decision making (or ignore them). Businesses began to create routine reports to inform decision makers (managers) about what had happened in the previous period (e.g., day, week, month, quarter). Although it was useful to know what had happened in the past, managers needed more than this: They needed a variety of reports at different levels of granularity to better understand and address changing needs and challenges of the business. These were usually called management information systems (MIS). In the early 1970s, Scott-Morton first articulated the major concepts of DSS. He defined DSSs as “interactive computer-based systems, which help decision makers utilize data and models to solve unstructured problems” (Gorry and Scott-Morton, 1971). The following is another classic DSS definition, provided by Keen and Scott-Morton (1978):

Decision support systems couple the intellectual resources of individuals with the capabilities of the computer to improve the quality of decisions. It is a computer-based support system for management decision makers who deal with semistructured problems.

Note that the term decision support system, like management information system and several other terms in the field of IT, is a content-free expression (i.e., it means different things to different people). Therefore, there is no universally accepted definition of DSS.

During the early days of analytics, data was often obtained from the domain experts using manual processes (i.e., interviews and surveys) to build mathematical or knowledge-based models to solve constrained optimization problems. The idea was to do the best with limited resources. Such decision support models were typically called operations research (OR). The problems that were too complex to solve optimally (using linear or nonlinear mathematical programming techniques) were tackled using heuristic methods such as simulation models. (We will introduce these as prescriptive analytics later in this chapter and in a bit more detail in Chapter 6.)

In the late 1970s and early 1980s, in addition to the mature OR models that were being used in many industries and government systems, a new and exciting line of models had emerged: rule-based expert systems. These systems promised to capture experts’ knowledge in a format that computers could process (via a collection of if–then–else rules or heuristics) so that these could be used for consultation much the same way that one
# The Rise of AI

## Origin & Definition of AI

Artificial intelligence (AI) is not new. The term was coined in 1956 by John McCarthy, a Stanford computer science professor who organized an academic conference on the topic at Dartmouth College in the summer of that year. The field of AI has gone through a series of boom-bust cycles since then, characterized by technological breakthroughs that stirred activity and excitement about the topic, followed by subsequent periods of disillusionment and disinterest known as ‘AI Winters’ as technical limitations were discovered. As you can see in figure 1, today we are once again in an ‘AI Spring’.

Artificial intelligence can be defined as human intelligence exhibited by machines; systems that approximate, mimic, replicate, automate, and eventually improve on human thinking. Throughout the past half-century a few key components of AI were established as essential: the ability to perceive, understand, learn, problem solve, and reason. Countless working definitions of AI have been proposed over the years but the unifying thread in all of them is that computers with the right software can be used to solve the kind of problems that humans solve, interact with humans and the world as humans do, and create ideas like humans. In other words, while the mechanisms that give rise to AI are ‘artificial’, the intelligence to which AI is intended to approximate is indistinguishable from human intelligence. In the early days of the science, processing inputs from the outside world required extensive programming, which limited early AI systems to a very narrow set of inputs and conditions. However since then, computer science has worked to advance the capability of AI-enabled computing systems.

Board games have long been a proving ground for AI research, as they typically involve a finite number of players, rules, objectives, and possible moves. This essentially means that games – one by one, including checkers, backgammon, and even Jeopardy! to name a few – have been taken over by AI. Most famously, in 1997 IBM’s Deep Blue defeated Garry Kasparov, the then reigning world champion of chess. This trajectory persists with the ancient Chinese game of Go, and the defeat of reigning world champion Lee Sedol by DeepMind’s AlphaGo in March 2016.

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1956</td>
<td>Dartmouth conference led by John McCarthy coins the term &quot;artificial intelligence&quot;</td>
</tr>
<tr>
<td>1957</td>
<td>Edward Feigenbaum develops the first Expert System, giving rebirth to AI</td>
</tr>
<tr>
<td>1964</td>
<td>Eliza, the first chatbot is developed by Joseph Weizenbaum at MIT</td>
</tr>
<tr>
<td>1975</td>
<td>IBM’s Watson Q&amp;A machine wins Jeopardy!</td>
</tr>
<tr>
<td>1980</td>
<td>Apple integrates Siri, a personal voice assistant into the iPhone</td>
</tr>
<tr>
<td>2011</td>
<td>2016 AlphaGo defeats Lee Sedol</td>
</tr>
<tr>
<td>2014</td>
<td>YouTube recognizes cats from videos</td>
</tr>
</tbody>
</table>

### The Rise of AI Timeline

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1950</td>
<td>AI is born</td>
</tr>
<tr>
<td>1956</td>
<td>Focus on specific intelligence</td>
</tr>
<tr>
<td>1970</td>
<td>Expert systems &amp; knowledge</td>
</tr>
<tr>
<td>1980</td>
<td>Deep learning: pattern analysis &amp; classification</td>
</tr>
<tr>
<td>1990</td>
<td>Machine learning</td>
</tr>
<tr>
<td>2000</td>
<td>Focus on specific problems</td>
</tr>
<tr>
<td>2010</td>
<td>Big data: large databases</td>
</tr>
<tr>
<td>2020</td>
<td>High-speed networks and connectivity</td>
</tr>
</tbody>
</table>

**AI Winter I**

- Limited computer processing power
- Limited database storage capacity
- Limited network ability
- Real-world problems are complicated
- Combinatorial explosion
- Disappointing results: failure to achieve scale
- Collapse of dedicated hardware vendors

**AI Winter II**

Definition of Artificial Intelligence (A.I.)
Artificial Intelligence

“... the science and engineering of making intelligent machines”

(John McCarthy, 1955)
Artificial Intelligence

“... technology that thinks and acts like humans”

Artificial Intelligence

“... intelligence exhibited by machines or software”

### 4 Approaches of AI

<table>
<thead>
<tr>
<th>Thinking Humanly</th>
<th>Thinking Rationally</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acting Humanly</td>
<td>Acting Rationally</td>
</tr>
</tbody>
</table>

4 Approaches of AI

<table>
<thead>
<tr>
<th>2. Thinking Humanly: The Cognitive Modeling Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Thinking Rationally: The “Laws of Thought” Approach</td>
</tr>
<tr>
<td>1. Acting Humanly: The Turing Test Approach (1950)</td>
</tr>
<tr>
<td>4. Acting Rationally: The Rational Agent Approach</td>
</tr>
</tbody>
</table>

AI Acting Humanly:
The Turing Test Approach
(Alan Turing, 1950)

• Knowledge Representation
• Automated Reasoning
  • Machine Learning (ML)
    • Deep Learning (DL)
  • Computer Vision (Image, Video)
  • Natural Language Processing (NLP)
• Robotics

Artificial Intelligence (AI)

Machine Learning (ML)

- Supervised Learning
- Unsupervised Learning
- Deep Learning (DL)
  - CNN
  - RNN LSTMs GRUs
  - GAN
- Semi-supervised Learning
- Reinforcement Learning

Source: https://leonardoaraujosantos.gitbooks.io/artificial-intelligence/content/deep_learning.html
3 Machine Learning Algorithms

Machine Learning (ML)

Source: https://www.mactores.com/services/aws-big-data-machine-learning-cognitive-services/
Cloud Computing Architecture for Artificial Intelligence
Cloud Computing Architecture for AI Machine Learning

• Lots of data
• Computation / Algorithms
• Knowledge

Artificial Intelligence and Machine Learning: AWS vs Azure vs GCP

Artificial Intelligence and Machine Learning: AWS vs Azure vs GCP

Artificial Intelligence and Machine Learning: AWS vs Azure vs GCP

Artificial Intelligence and Machine Learning: AWS vs Azure vs GCP

Machine learning building block services

• Speech to text and text to speech
• Chatbots
• Translation
• Text Analytics
• Document Analysis
• Image and video analysis
• Anomaly detection
• Personalization

Machine learning platforms

• Guided model development
• Full ML workbench
• MLOps
• Augmented AI
  • AWS Augmented AI (Amazon A2I)
    • The power of real, living, breathing humans to help improve your machine learning service

Machine Learning Infrastructure

• Hardware
  • AWS: Habana Gaudi ASIC instances, custom processor AWS Trainium, optimized for model training. Inferentia for machine learning inferences.
  • Azure: FPGA-based virtual machines tuned specifically for machine learning workloads.
  • GCP: Tensor Processing Unit (TPU), ASIC-optimized for the TensorFlow framework.

• Machine learning explainability and bias
  • AWS: SageMaker Clarify
  • Azure: Responsible ML and Fairlearn SDK
  • GCP: AI Explanations

Machine Learning on AWS

Innovate faster with the most comprehensive set of AI and ML services

Make accurate predictions, get deeper insights from your data, reduce operational overhead, and improve customer experience with AWS machine learning (ML). AWS helps you at every stage of your ML adoption journey with the most comprehensive set of artificial intelligence (AI) and ML services, infrastructure, and implementation resources.

Build with a proven leader
Tailor ML to your business needs
Accelerate your ML adoption

Source: https://aws.amazon.com/
Amazon SageMaker is an ML service enabling data scientists, data engineers, MLOps engineers, and business analysts to build, train, and deploy ML models for any use case, regardless of ML expertise.

Source: https://aws.amazon.com/machine-learning/
Amazon SageMaker for Business Analysts

Amazon SageMaker Canvas

Build machine learning models without writing code

Collaborate and send models to data scientists using Amazon SageMaker Studio for review and feedback

1. **Browse, import, and join data**
   - Import data from disparate sources and create new unified datasets for training prediction models

2. **Pick target**
   - Select values you want to predict

3. **Prepare and analyze data**
   - Automatically detect errors, cleanse, and analyze data to help determine if your data is prepared for ML

4. **Create model**
   - Create ML models with the click of a button

5. **Generate and understand predictions**
   - Generate single or bulk predictions and understand predictions

Amazon SageMaker for Data Scientists

Amazon SageMaker Studio Notebooks

**Access ML Data**
Connect to many data sources such as Amazon S3, Apache Spark, Amazon Redshift, CSV files, and more.

**Prepare Data**
Transform data to browse data sources, explore metadata, schemas, and write queries in popular languages.

**Build ML Models**
Optimized with 150+ popular open source models, and frameworks such as TensorFlow and PyTorch.

**Deploy and Fine Tune at Scale**

**Deploy and Monitor Results**
Create, automate, and manage end-to-end ML workflows to improve model quality.

**Track ML Models**

**Browse Active Experiments**

**Search Previous Experiments**

**Compare All Results**

**Capture Debug Metrics**

**Train and Tune ML Model**
Correct performance problems in real time.

Easily add intelligence to your applications
No machine learning skills required

Explore AWS AI services

AWS pre-trained AI Services provide ready-made intelligence for your applications and workflows. AI Services easily integrate with your applications to address common use cases such as personalized recommendations, modernizing your contact center, improving safety and security, and increasing customer engagement. Because we use the same deep learning technology that powers Amazon.com and our ML Services, you get quality and accuracy from continuously-learning APIs. And best of all, AI Services on AWS don't require machine learning experience.

Source: https://aws.amazon.com/machine-learning/ai-services/
AWS Machine Learning AI Services

Computer vision

- **Analyze images and videos**
  Catalog assets, automate workflows, and extract meaning from your media and applications.
  
  **Amazon Rekognition »**

- **Detect defects and automate inspection**
  Identify missing product components, vehicle and structure damage, and irregularities for comprehensive quality control.
  
  **Amazon Lookout for Vision »**

- **Utilize computer vision at the edge**
  Improve operations with automated monitoring to find bottlenecks and assess manufacturing quality and safety.
  
  **AWS Panorama »**

<table>
<thead>
<tr>
<th>Extract text and data</th>
<th>Acquire insights</th>
<th>Control quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull valuable information from millions of documents at speed.</td>
<td>Maximize the value of unstructured text with natural language processing (NLP).</td>
<td>Add humans to the review process to ensure accuracy and compliance of sensitive data.</td>
</tr>
</tbody>
</table>

*Amazon Textract →*

*Amazon Comprehend →*

*Amazon A2I →*
AWS Machine Learning AI Services
Language AI

**Build chatbots & virtual agents**
Create automated conversation channels to improve customer service.

*Amazon Lex »*

**Automate speech recognition**
Enhance your applications and workflows with automatic speech recognition.

*Amazon Transcribe »*

**Give your apps a voice**
Convert text into life-like speech, improving user experience and accessibility.

*Amazon Polly »*

AWS Machine Learning AI Services
Improve customer experience

Find accurate information faster
Enhance websites and applications with natural language speech to help users quickly search for what they need.
Amazon Kendra »

Personalize online experiences
Use ML to customize applications and websites to each individual user.
Amazon Personalize »

Engage audiences in every language
Expand your reach and accessibility with fast, accurate, and customizable translation.
Amazon Translate »

Source: https://aws.amazon.com/machine-learning/ai-services/
AWS Machine Learning AI Services

Business metrics

**Forecast business metrics**

Harness unique data types and time series data to create accurate end-to-end prediction models.

Amazon Forecast »

**Detect online fraud**

Stop adversaries and identify potential attacks with technology honed through years of use on Amazon.com.

Amazon Fraud Detector »

**Identify data anomalies**

Detect and identify root causes of unexpected changes in metrics such as revenue and retention.

Amazon Lookout for Metrics »

<table>
<thead>
<tr>
<th>Improve application availability</th>
<th>Automated code reviews</th>
<th>Eliminate costly inefficient code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simplify operational performance measurement and reduce application downtime.</td>
<td>Detect bugs and assess critical issues and vulnerabilities fast for higher quality code.</td>
<td>Use runtime behavior analysis to improve application performance and decrease compute costs.</td>
</tr>
</tbody>
</table>

*Amazon DevOps Guru ➞ Amazon CodeGuru Reviewer ➞ Amazon CodeGuru Profiler*

**AWS Machine Learning AI Services**

**Industrial AI**

- **Detect abnormal machine conditions**
  
  Automatically detect abnormal machine conditions by analyzing sensor data

  *Amazon Lookout for Equipment »*

- **Predictive maintenance**
  
  End-to-end predictive maintenance system that includes sensors, gateway, anomaly detection service and end-user application

  *Amazon Monitron »*

AWS Machine Learning AI Services

Healthcare

Store and analyze health data

Securely store, transfer, query, and analyze health data to offer a complete view of a patient’s medical history

Amazon Healthlake »

Extract health data

Extract information from unstructured medical text accurately and quickly

Amazon Comprehend Medical »

Source: https://aws.amazon.com/machine-learning/ai-services/
AWS Products and Services

Analytics
Business Applications
End User Computing
Media Services
Robotics
Application Integration
Game Tech
Migration & Transfer
Satellite
Compute
Customer Engagement
Internet of Things
Security, Identity & Compliance
AR & VR
Internet of Things
Mobile
Machine Learning
AWS Cost Management
Database
Networking & Content Delivery
Storage
Blockchain
Developer Tools
Management & Governance
Quantum Technologies
Source: https://aws.amazon.com/
AWS Services

- Amazon EC2
  - Virtual servers in the cloud
- Amazon Simple Storage Service (S3)
  - Scalable storage in the cloud
- Amazon Aurora
  - High performance managed relational database
- Amazon DynamoDB
  - Managed NoSQL database
- Amazon RDS
  - Managed relational database service for MySQL, PostgreSQL, Oracle, SQL Server, and MariaDB

Source: https://aws.amazon.com/
AWS Services

• AWS Lambda
  • Run code without thinking about servers

• AWS Elastic Beanstalk
  • Run and manage web apps

• Amazon VPC
  • Isolated cloud resources

• Amazon Lightsail
  • Launch and manage virtual private servers

• Amazon SageMaker
  • Build, train, and deploy machine learning models at scale

Source: https://aws.amazon.com/
Available AWS Certifications

**Professional**
- Two years of comprehensive experience designing, operating, and troubleshooting solutions using the AWS Cloud

**Associate**
- One year of experience solving problems and implementing solutions using the AWS Cloud

**Foundational**
- Six months of fundamental AWS Cloud and industry knowledge

Specialty
- Technical AWS Cloud experience in the Specialty domain as specified in the exam guide

https://aws.amazon.com/certification/
Available AWS Certifications

1. CLF
   - 六個月的基礎 AWS 雲端和產業知識
   - 雲端從業人員

2. SAA
   - 一年使用 AWS 雲端解決問題和實作解決方案的經驗
   - 架構師

3. SAP
   - 兩年使用 AWS 雲端設計、操作及難排解方案的廣泛經驗

- 专家级
  - 考試指南中指定的專業領域中的 AWS 雲端技術經驗

- DA
  - Advanced Networking
  - Data Analytics

- ML
  - Database
  - Machine Learning

- Security

https://aws.amazon.com/certification/
AWS Certification

Validate technical skills and cloud expertise to grow your career and business.

Get started

Certification exam guides and details
Explore our role-based certifications and our Specialty certifications in specific technical areas. Select an exam to learn more:

Select ▼

Options for taking your exam
Curious about how to take your AWS Certification exam? We offer flexible, convenient options for taking exams so you can select what works best for you.

Learn more ▸

Access your AWS Certification Account
Schedule and take exams from your AWS Certification Account. You can also access your AWS Certified benefits, such as your digital badge and 50% discount voucher for a future exam.

https://aws.amazon.com/certification/
AWS Certification

**Professional**
Two years of comprehensive experience designing, operating, and troubleshooting solutions using the AWS Cloud

**Associate**
One year of experience solving problems and implementing solutions using the AWS Cloud

**Foundational**
Six months of fundamental AWS Cloud and industry knowledge

**Specialty**
Technical AWS Cloud experience in the Specialty domain as specified in the exam guide

https://aws.amazon.com/certification/
AWS Certification

Format
65 questions; either multiple choice or multiple response

Type
Foundational

Delivery Method
Testing center or online proctored exam

Time
90 minutes to complete the exam

Cost
100 USD (Practice Exam: 20 USD)

Language
Available in English, Indonesian (Bahasa), Japanese, Korean, and Simplified Chinese

https://aws.amazon.com/certification/
AWS Certification

Format
65 questions; either multiple choice or multiple response

Type
Associate

Delivery Method
Testing center or online proctored exam

Time
130 minutes to complete the exam

Cost
150 USD (Practice exam: 20 USD)

Language
Available in English, Japanese, Korean, and Simplified Chinese

https://aws.amazon.com/certification/
AWS Certification

Professional

Two years of comprehensive experience designing, operating, and troubleshooting solutions using the AWS Cloud

Format
75 questions; either multiple choice or multiple response

Type
Professional

Delivery Method
Testing center or online proctored exam

Time
180 minutes to complete the exam

Cost
300 USD (Practice Exam: 40 USD)

Language
Available in English, Japanese, Korean, and Simplified Chinese

https://aws.amazon.com/certification/
AWS Certification

**Format**
65 questions; either multiple choice or multiple response

**Type**
Associate

**Delivery Method**
Testing center or online proctored exam

**Time**
130 minutes to complete the exam

**Cost**
150 USD (Practice exam: 20 USD)

**Language**
Available in English, Japanese, Korean, and Simplified Chinese

[https://aws.amazon.com/certification/](https://aws.amazon.com/certification/)
AWS Certification

Format
65 questions; either multiple choice or multiple response

Type
Associate

Delivery Method
Testing center or online proctored exam

Time
130 mins to complete the exam

Cost
150 USD (Practice exam: 20 USD)

Language
Available in English, Japanese, Korean, and Simplified Chinese

https://aws.amazon.com/certification/
AWS Certification

Professional

Two years of comprehensive experience designing, operating, and troubleshooting solutions using the AWS Cloud

Format
75 questions; either multiple choice or multiple response

Type
Professional

Delivery Method
Testing center or online proctored exam

Time
180 minutes to complete the exam

Cost
300 USD (Practice exam: 40 USD)

Language
Available in English, Japanese, Korean, and Simplified Chinese

https://aws.amazon.com/certification/
AWS Certification

Format
65 questions; either multiple choice or multiple response

Type
Specialty

Delivery Method
Testing center or online proctored exam

Time
180 minutes to complete the exam

Cost
300 USD (Practice exam: 40 USD)

Language
Available in English, Japanese, Korean, and Simplified Chinese

https://aws.amazon.com/certification/
AWS Certified Cloud Practitioner

- Level: **Foundational**
- Length: **90 minutes** to complete the exam
- Cost: **100 USD**
- Format: **65 questions**; either multiple choice or multiple response
- Delivery method: **Pearson VUE** and PSI; testing center or online proctored exam

https://aws.amazon.com/certification/certified-cloud-practitioner/
AWS Certified Solutions Architect – Associate

• Level: Associate
• Length: 130 minutes to complete the exam
• Cost: 150 USD
• Format: 65 questions; either multiple choice or multiple response
• Delivery method: Pearson VUE and PSI; testing center or online proctored exam

AWS Certified SysOps Administrator - Associate

- Level: Associate
- Length: 180 minutes to complete the exam
- Cost: 150 USD
- Format: 65 scoring opportunities that may be multiple choice, multiple response, or exam lab
- Delivery method: Pearson VUE and PSI; testing center or online proctored exam

https://aws.amazon.com/certification/certified-sysops-admin-associate
AWS Certified Solutions Architect – Associate

- Level: Associate
- Length: 130 minutes to complete the exam
- Cost: 150 USD
- Format: 65 questions; either multiple choice or multiple response
- Delivery method: Pearson VUE and PSI; testing center or online proctored exam

https://aws.amazon.com/certification/certified-developer-associate
AWS Certified Solutions Architect - Professional

- Level: Professional
- Length: 180 minutes to complete the exam
- Cost: 300 USD
- Format: 75 questions; either multiple choice or multiple response
- Delivery method: Pearson VUE and PSI; testing center or online proctored exam

https://aws.amazon.com/certification/certified-solutions-architect-professional
AWS Certified DevOps Engineer - Professional

• Level: Professional
• Length: 180 minutes to complete the exam
• Cost: 300 USD
• Format: 75 questions; either multiple choice or multiple response
• Delivery method: Pearson VUE and PSI; testing center or online proctored exam

https://aws.amazon.com/certification/certified-devops-engineer-professional
AWS Certified Security - Specialty

- Level: Specialty
- Length: **170 minutes** to complete the exam
- Cost: **300 USD**
- Format: **65 questions**; either multiple choice or multiple response
- Delivery method: **Pearson VUE** and PSI; testing center or online proctored exam

AWS Certified Database - Specialty

- Level: Specialty
- Length: 180 minutes to complete the exam
- Cost: 300 USD
- Format: 65 questions; either multiple choice or multiple response
- Delivery method: Pearson VUE and PSI; testing center or online proctored exam

https://aws.amazon.com/certification/certified-database-specialty
AWS Certified Machine Learning – Specialty

- **Level**: Specialty
- **Length**: **180 minutes** to complete the exam
- **Cost**: **300 USD**
- **Format**: **65 questions**; either multiple choice or multiple response
- **Delivery method**: **Pearson VUE** and PSI; testing center or online proctored exam

AWS Certified Data Analytics - Specialty

• Level: Specialty
• Length: 180 minutes to complete the exam
• Cost: 300 USD
• Format: 65 questions; either multiple choice or multiple response
• Delivery method: Pearson VUE and PSI; testing center or online proctored exam

AWS Certified Advanced Networking - Specialty

• Level: Specialty
• Length: 170 minutes to complete the exam
• Cost: 300 USD
• Format: 65 questions; either multiple choice or multiple response
• Delivery method: Pearson VUE and PSI; testing center or online proctored exam

https://aws.amazon.com/certification/certified-advanced-networking-specialty
AWS Certifications Roadmap

- AWS Certified Cloud Practitioner
- AWS Certified Solutions Architect - Associate
- AWS Certified Developer - Associate
- AWS Certified SysOps Administrator - Associate
- AWS Certified DevOps Engineer - Professional
- AWS Certified Solutions Architect - Professional
- AWS Certified Security - Specialty
- AWS Certified Advanced Networking - Specialty
- AWS Certified Database - Specialty
- AWS Certified Data Analytics - Specialty
- AWS Certified Machine Learning - Specialty

https://aws.amazon.com/certification/
AWS Certifications Roadmap

Cloud Architect

https://aws.amazon.com/certification/
AWS Certifications Roadmap

Cloud Developer

https://aws.amazon.com/certification/
AWS Certifications Roadmap

DevOps

https://aws.amazon.com/certification/
AWS Certifications Roadmap

Data Analytics / Machine Learning

https://aws.amazon.com/certification/
AWS Certifications Roadmap

Security

https://aws.amazon.com/certification/
AWS Certified Cloud Practitioner (CLF-C01)

Source: https://aws.amazon.com/certification/certified-cloud-practitioner/
AWS Certified Cloud Practitioner

• This certification provides individuals in a larger variety of cloud and technology roles with a way to validate their AWS Cloud knowledge and enhance their professional credibility.

• This exam covers four domains, including cloud concepts, security, technology, and billing and pricing.

https://aws.amazon.com/certification/certified-cloud-practitioner/
AWS Certified Solutions Architect – Associate (SAA-C02)

Source: https://aws.amazon.com/certification/certified-solutions-architect-associate
AWS Certified Solutions Architect – Associate

• This certification validates your ability to effectively demonstrate knowledge of how to architect and deploy secure and robust applications on AWS technologies.

• This exam is for anyone with at least one year of hands-on experience designing available, cost-efficient, fault-tolerant, and scalable and distributed systems on AWS.

AWS Academy and Certifications

- AWS Academy Cloud Foundations (ACF)
  - AWS Certified Cloud Practitioner (CLF-C01)
    - https://aws.amazon.com/certification/certified-cloud-practitioner/

- AWS Academy Cloud Architecting (ACA)
  - AWS Certified Solutions Architect – Associate (SAA-C02)

https://aws.amazon.com/training/awsacademy/
# AWS Certified Cloud Practitioner (CLF-C01)

<table>
<thead>
<tr>
<th>Domain</th>
<th>% of Examination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain 1: Cloud Concepts</td>
<td>26%</td>
</tr>
<tr>
<td>Domain 2: Security and Compliance</td>
<td>25%</td>
</tr>
<tr>
<td>Domain 3: Technology</td>
<td>33%</td>
</tr>
<tr>
<td>Domain 4: Billing and Pricing</td>
<td>16%</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

## AWS Certified Solutions Architect – Associate (SAA-C02)

<table>
<thead>
<tr>
<th>Domain</th>
<th>% of Examination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain 1: Design <strong>Resilient</strong> Architectures</td>
<td>30%</td>
</tr>
<tr>
<td>Domain 2: Design <strong>High-Performing</strong> Architectures</td>
<td>28%</td>
</tr>
<tr>
<td>Domain 3: Specify <strong>Secure</strong> Applications and Architectures</td>
<td>24%</td>
</tr>
<tr>
<td>Domain 4: Design <strong>Cost-Optimized</strong> Architectures</td>
<td>18%</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

AWS Certified Cloud Practitioner (CLF-C01)

Source: https://aws.amazon.com/certification/certified-cloud-practitioner/
AWS Certified Cloud Practitioner (CLF-C01)

• Domain 1: Cloud Concepts
  • 1.1 Define the AWS Cloud and its value proposition
  • 1.2 Identify aspects of AWS Cloud economics
  • 1.3 List the different cloud architecture design principles

Source: https://aws.amazon.com/certification/certified-cloud-practitioner/
• Domain 2: Security and Compliance

• 2.1 Define the AWS shared responsibility model
• 2.2 Define AWS Cloud security and compliance concepts
• 2.3 Identify AWS access management capabilities
• 2.4 Identify resources for security support

Source: https://aws.amazon.com/certification/certified-cloud-practitioner/
• Domain 3: Technology

• 3.1 Define methods of deploying and operating in the AWS Cloud

• 3.2 Define the AWS global infrastructure

• 3.3 Identify the core AWS services

• 3.4 Identify resources for technology support

Source: https://aws.amazon.com/certification/certified-cloud-practitioner/
• Domain 4: Billing and Pricing

  • 4.1 Compare and contrast the various pricing models for AWS
  • 4.2 Recognize the various account structures in relation to AWS billing and pricing
  • 4.3 Identify resources available for billing support

Source: https://aws.amazon.com/certification/certified-cloud-practitioner/
AWS Certified Solutions Architect – Associate (SAA-C02)

Source: https://aws.amazon.com/certification/certified-solutions-architect-associate
• Domain 1: Design Resilient Architectures
  • 1.1 Design a multi-tier architecture solution
  • 1.2 Design highly available and/or fault-tolerant architectures
  • 1.3 Design decoupling mechanisms using AWS services
  • 1.4 Choose appropriate resilient storage
AWS Certified Solutions Architect – Associate (SAA-C02)

• Domain 2: Design High-Performing Architectures
  • 2.1 Identify elastic and scalable compute solutions for a workload
  • 2.2 Select high-performing and scalable storage solutions for a workload
  • 2.3 Select high-performing networking solutions for a workload
  • 2.4 Choose high-performing database solutions for a workload

Source: https://aws.amazon.com/certification/certified-solutions-architect-associate
AWS Certified Solutions Architect – Associate (SAA-C02)

• Domain 3: Design Secure Applications and Architectures
  • 3.1 Design secure access to AWS resources
  • 3.2 Design secure application tiers
  • 3.3 Select appropriate data security options

Source: https://aws.amazon.com/certification/certified-solutions-architect-associate
• Domain 4: Design Cost-Optimized Architectures
  • 4.1 Identify cost-effective storage solutions
  • 4.2 Identify cost-effective compute and database services
  • 4.3 Design cost-optimized network architectures
## AWS Certifications

### Exam Pricing

<table>
<thead>
<tr>
<th>Exam Type</th>
<th>Price in USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foundational</td>
<td>$100</td>
</tr>
<tr>
<td>Associate</td>
<td>$150</td>
</tr>
<tr>
<td>Professional</td>
<td>$300</td>
</tr>
<tr>
<td>Specialty</td>
<td>$300</td>
</tr>
</tbody>
</table>
Web Application with AWS Core Services
fb.com on AWS

AWS Application Services

AWS Development and DevOps Services

AWS Serverless Architecture
AWS Serverless Airline Booking

Source: https://github.com/aws-samples/aws-serverless-airline-booking
AWS Serverless Airline Booking Stack

UI/UX
- Quasar framework
- Vue.js
- AWS Amplify
- Stripe Elements

Data/Lang
- Amazon DynamoDB
- Python
- TypeScript
- JavaScript

API/Auth
- AWS AppSync
- Amazon API Gateway
- Amazon Cognito

Messaging
- Amazon SNS
- AWS Step Functions

Source: https://github.com/aws-samples/aws-serverless-airline-booking
AWS Serverless Airline Booking
High level infrastructure architecture

Source: https://github.com/aws-samples/aws-serverless-airline-booking
AWS Serverless Architecture

AWS Operational Responsibility Models

Source: Heitor Lessa (2019), How to build a full stack serverless airline ticketing web app, https://www.youtube.com/watch?v=MyoOEHToZag
Build a Serverless Web Application
Build a Serverless Web Application

Overview
In this tutorial, you’ll create a simple serverless web application that enables users to request unicorn rides from the Wild Rydes fleet. The application will present users with an HTML based user interface for indicating the location where they would like to be picked up and will interface on the backend with a RESTful web service to submit the request and dispatch a nearby unicorn. The application will also provide facilities for users to register with the service and log in before requesting rides.

Application Architecture

AWS Experience: Beginner
Time to complete: 2 hours
Cost to complete: Each service used in this architecture is eligible for the AWS Free Tier. If you are outside the usage limits of the Free Tier, completing this tutorial will cost you less than $0.25.

Source: https://aws.amazon.com/getting-started/projects/build-serverless-web-app-lambda-apigateway-s3-dynamodb-cognito/
Build a Serverless Web Application with Amazon S3, AWS Lambda, Amazon API Gateway, Amazon DynamoDB, and Amazon Cognito

Source: https://aws.amazon.com/getting-started/projects/build-serverless-web-app-lambda-apigateway-s3-dynamodb-cognito/
Build a Serverless Web Application with Amazon S3, AWS Lambda, Amazon API Gateway, Amazon DynamoDB, and Amazon Cognito

Source: https://aws.amazon.com/getting-started/projects/build-serverless-web-app-lambda-apigateway-s3-dynamodb-cognito/
Static Web Hosting

Amazon S3 hosts static web resources including HTML, CSS, JavaScript, and image files which are loaded in the user's browser.

Source: https://aws.amazon.com/getting-started/projects/build-serverless-web-app-lambda-apigateway-s3-dynamodb-cognito/
Build a Serverless Web Application with Amazon S3, AWS Lambda, Amazon API Gateway, Amazon DynamoDB, and Amazon Cognito

Source: https://aws.amazon.com/getting-started/projects/build-serverless-web-app-lambda-apigateway-s3-dynamodb-cognito/
User Management

Amazon Cognito provides user management and authentication functions to secure the backend API.
Build a Serverless Web Application

with Amazon S3, AWS Lambda, Amazon API Gateway,
Amazon DynamoDB, and Amazon Cognito

Source: https://aws.amazon.com/getting-started/projects/build-serverless-web-app-lambda-apigateway-s3-dynamodb-cognito/
Serverless Backend

Amazon DynamoDB provides a persistence layer where data can be stored by the API's Lambda function.

Source: https://aws.amazon.com/getting-started/projects/build-serverless-web-app-lambda-apigateway-s3-dynamodb-cognito/
Build a Serverless Web Application
with Amazon S3, AWS Lambda, Amazon API Gateway,
Amazon DynamoDB, and Amazon Cognito

Source: https://aws.amazon.com/getting-started/projects/build-serverless-web-app-lambda-apigateway-s3-dynamodb-cognito/
Build a Serverless Web Application with Amazon S3, AWS Lambda, Amazon API Gateway, Amazon DynamoDB, and Amazon Cognito

RESTful API
JavaScript executed in the browser sends and receives data from a public backend API built using Lambda and API Gateway.

Source: https://aws.amazon.com/getting-started/projects/build-serverless-web-app-lambda-apigateway-s3-dynamodb-cognito/
Terminate resources

Resource Cleanup

You will terminate an Amazon S3 bucket, an Amazon Cognito User Pool, an AWS Lambda function, an IAM role, a DynamoDB table, a REST API, and a CloudWatch Log.

It is a best practice to delete resources you are no longer using to avoid unwanted charges.
Summary

• Cloud Computing Architecture for AI
• AWS Certifications Roadmap for Machine Learning
• AWS Serverless Architecture
References

- [https://aws.amazon.com/certification/](https://aws.amazon.com/certification/)
- [https://www.aws.training/](https://www.aws.training/)
- [https://aws.amazon.com/training/awsacademy/](https://aws.amazon.com/training/awsacademy/)
- [https://aws.amazon.com/education/awseducate/](https://aws.amazon.com/education/awseducate/)
- AWS Certified Cloud Practitioner
- AWS Certified Solutions Architect – Associate
- AWS Cloud Practitioner Essentials (Second Edition)
  - [https://aws.amazon.com/training/course-descriptions/cloud-practitioner-essentials/](https://aws.amazon.com/training/course-descriptions/cloud-practitioner-essentials/)
- Architecting on AWS
  - [https://aws.amazon.com/training/course-descriptions/architect/](https://aws.amazon.com/training/course-descriptions/architect/)
Cloud Computing Architecture for Artificial Intelligence

Q & A

Min-Yuh Day, Ph.D, Associate Professor

Institute of Information Management, National Taipei University

https://web.ntpu.edu.tw/~myday

2022-05-24