COACH CLOUD EDITION · 2026 | VERIFIED BY CLOUDEXAMPRO.COM

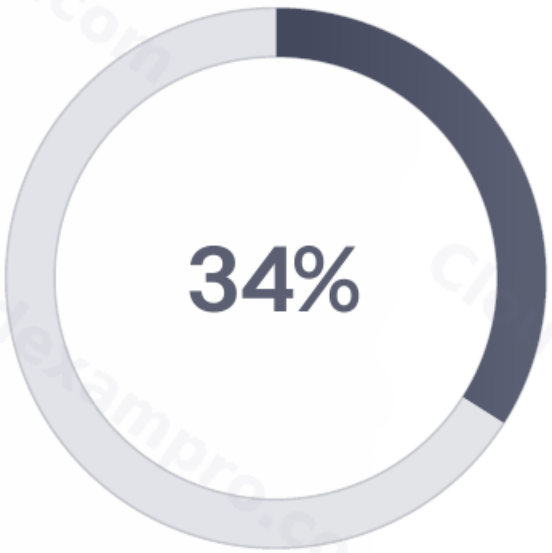
DEA-C01 Premium Cheat Sheet

Scan it in 10 minutes. Walk in. Win. — Coach Cloud | **Cloudexampro.com**

This guide maps 10 high-impact technical modules directly to the 4 official AWS exam domains. Every module is built around the exam's core philosophy: **least operational burden, cost optimization, and serverless-first design.**

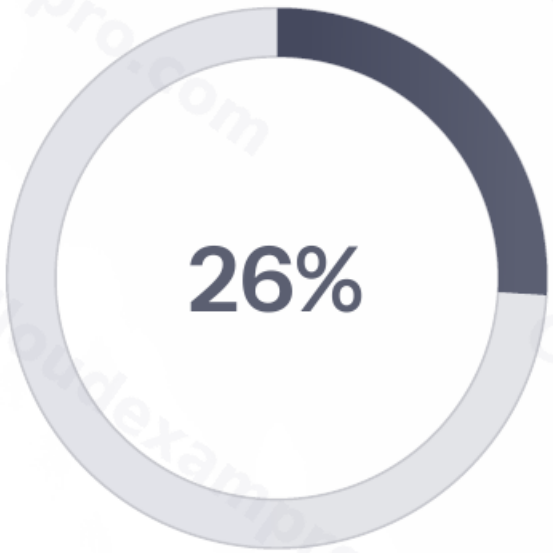
EXAM BLUEPRINT

Official Exam Domains & Coach's 7 Golden Rules



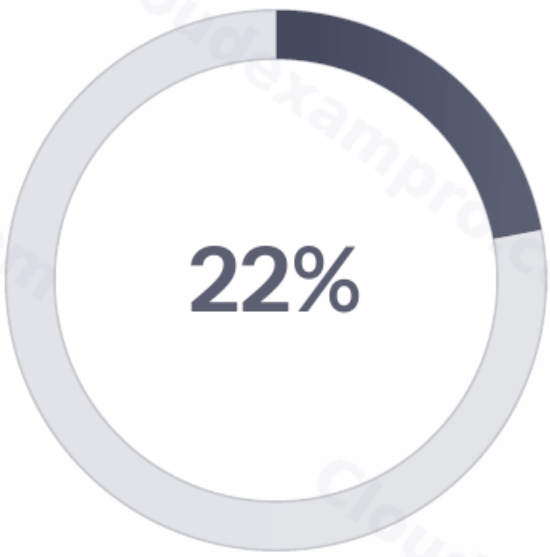
Domain 1: Ingestion

Largest domain — streaming, batch, and ETL



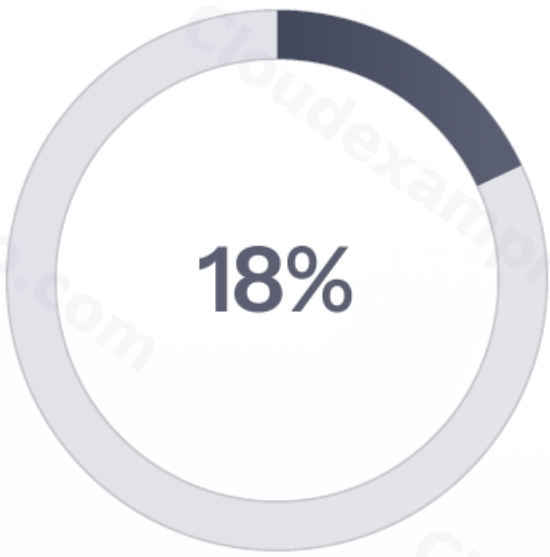
Domain 2: Storage

S3, EBS, Glacier — consistency and cost



Domain 3: Operations

Monitoring, orchestration, deployment



Domain 4: Security

IAM, KMS, compliance — every question

Coach's 7 Golden Rules — Lock These In First

1. Least Operational Burden

Always pick Serverless / Managed over self-managed.

2. Cost Optimization

Parquet + Partitions / Spot / gp3 / Gateway Endpoints.

3. Encryption

KMS CMK (customer-managed) always — enables CloudTrail audit.

4. Access Control

IAM Roles NEVER Users or Access Keys for services.

5. Private AWS Access

VPC Gateway Endpoints — S3 + DynamoDB = FREE. All others = Interface Endpoints (cost).

6. Real-time on Aurora/RDS

Zero-ETL → Redshift. No Glue. No Firehose. Seconds latency.

7. S3 Consistency

STRONG for ALL ops since 2020. "Eventual" = DEAD. ELIMINATE IT.

MODULE 0

Medallion Architecture + Zero-ETL

Zero-ETL is AWS's biggest architectural push for 2026. The exam **will** test this — know every detail.

🏅 Medallion Architecture — The Data Lake Blueprint

1 → 2 → 3

🥉 Bronze — Raw Layer

WHERE: S3 Standard/IA
WHAT: Land as-is, any format, zero transform

🥈 Silver — Cleaned Layer

WHERE: S3 + Glue ETL Jobs
WHAT: Validate, deduplicate, convert to Parquet/ORC

🥇 Gold — Business Layer

WHERE: Redshift / Athena (Iceberg)
WHAT: Aggregated, query-optimized, partitioned

Zero-ETL — The 2026 Exam Favorite

Direct AWS-managed replication: Source DB → Redshift. No pipeline. Seconds latency.

- Aurora MySQL → Redshift ✅ PRIMARY
- Aurora PostgreSQL → Redshift ✅ PRIMARY
- RDS MySQL → Redshift ✅ GA
- DynamoDB → Redshift / OpenSearch ✅ GA

Zero-ETL Eliminates ❌

- Glue ETL jobs to build and maintain
- Data Firehose delivery streams
- S3 intermediate staging buckets
- Custom Lambda transformation functions
- Error retry and reprocessing logic

✅ AWS manages ALL replication automatically. Near-real-time data in Redshift (seconds). Automatic schema propagation.

⚠️ ⚠️ Exam Trap: Zero-ETL is NOT universal. Aurora/RDS → Redshift is the primary use case. Complex multi-source transforms → AWS Glue still wins.

MODULE 1 + 2

Compute & Storage — EC2 · EBS · S3 · Glacier

🖥️ Compute Quick Reference

Persistent Storage

EBS (gp3) — survives stop/terminate

Ephemeral Storage

Instance Store — DIES on stop. Never for data!

Lowest Cost Compute

Spot Instances — up to 90% off

Predictable Workload

Reserved (1–3 yr) — up to 72% off

Containers, Zero Servers

ECS + Fargate — "least burden" answer

⚠ ⚠️ Instance Store = PERMANENTLY LOST on stop/terminate. EBS = Persistent. Always.

💾 EBS — gp3 is the 2026 Default

Volume Type	Status	Best For
gp3 ✅	Current Default	General purpose — everything
gp2 ❌	Legacy — retire it	Nothing new
io2 Block Express	Premium	Mission-critical DBs
st1	Current	Big data, sequential
sc1	Current	Cold HDD, lowest cost

gp3 is ~20% cheaper than gp2. IOPS + throughput configured **independently**. If gp2 appears as an option → eliminate immediately.

🪣 S3 Storage Classes & Glacier Tiers

Tier	Retrieval	Min Duration	Exam Trigger
S3 Standard	Milliseconds	None	Hot data, active pipeline
Intelligent-Tiering	Milliseconds	None	Unknown access pattern
Standard-IA	Milliseconds	30 days	Monthly access, fast retrieval
Glacier Instant Retrieval	**Milliseconds**	90 days	Archive + instant access
Glacier Flexible	Minutes–Hours	90 days	Backup, hours acceptable
Glacier Deep Archive	12–48 Hours	180 days	7–10yr compliance — CHEAPEST

✓ 💰 Athena Cost Formula: Parquet/ORC + Date Partitions + Compression = **90% cost reduction**. Intelligent-Tiering has NO retrieval fee but HAS a monitoring fee per object.

MODULE 3

Networking — VPC · Direct Connect · Transit Gateway

VPC Components — Fast Reference

Component	Function	Key Fact
Security Groups	Instance-level firewall	Stateful — return traffic auto-allowed
NACLs	Subnet-level firewall	Stateless — define BOTH directions
NAT Gateway	Private → internet (outbound)	Managed — use over NAT Instance
VPC Peering	Two VPCs, private IPs	NOT transitive
S3 Gateway Endpoint	Private S3 access	FREE ✅ Deploy everywhere
DynamoDB Gateway EP	Private DynamoDB access	FREE ✅ Deploy everywhere
Interface Endpoints	All other AWS services	Hourly cost 💰

🔌 Direct Connect vs. VPN

Direct Connect	Low, consistent latency. Dedicated bandwidth. Higher cost. Primary connection.
Site-to-Site VPN	Variable latency (internet). Shared bandwidth. Lower cost. Backup to DX.

🔄 Transit Gateway

Hub-and-spoke replaces complex peering mesh. Connects ALL VPCs + on-premises via one managed hub. Supports inter-region peering. Share across accounts via AWS RAM.

🎯 Pro-Tip: "50 VPCs need connectivity to on-premises." → Transit Gateway + Direct Connect Gateway. One hub. Zero mesh complexity.

⚠️ ⚠️ Exam Trap: VPC Peering is NOT transitive. A↔B + B↔C ≠ A↔C. Need transitive → Transit Gateway. S3 Gateway Endpoint = FREE. Eliminates NAT Gateway costs ($0.045/GB saved).

MODULE 4

Serverless & Streaming — Lambda · Step Functions · Kinesis · MSK

⚡ AWS Lambda — Key Facts

Max Execution

15 minutes — hard limit

Memory

128 MB → 10 GB

CPU Scaling

Proportional to memory

Cold Start Fix

Provisioned Concurrency

⚠️ ⚠️ Job > 15 minutes → Glue (ETL) / EMR (big data) / ECS (containers). NEVER Lambda for long-running jobs.

🔄 Step Functions — Orchestration

Type	Duration	Model	Best For
Standard	1 year	Exactly-once	Long pipelines, audit trail
Express	5 minutes	At-least-once	High-volume event processing

States: Task, Choice, Parallel, Map, Wait, Succeed, Fail, Pass. Built-in retry with exponential backoff.

📡 Kinesis — 2026 Official Names & Decision Guide

ℹ️ ⚡ CRITICAL: "Kinesis Data Firehose" is now officially **Amazon Data Firehose**. Old name = wrong on exam.

Kinesis Data Streams

Manual shard management. Custom consumers, Lambda integration. You manage scaling.

Amazon Data Firehose

Fully managed, auto-scales. Auto-delivers to S3/Redshift/OpenSearch. Built-in JSON → Parquet conversion.

Kinesis Data Analytics

Serverless. Real-time SQL or Apache Flink stream processing. Managed.

Amazon MSK

Fully managed Apache Kafka. 100% API compatible. MSK Serverless = zero capacity planning (2026 least-burden).

MODULE 5

Database — RDS · Aurora · DynamoDB · Redshift

⚠ 🚨 #1 Most-Tested Trap — NEVER Mix These Up: Multi-AZ = HIGH AVAILABILITY (automatic failover, synchronous, ZERO reads served). Read Replicas = READ SCALABILITY (NOT automatic failover, asynchronous, serves read traffic).

🔑 DynamoDB — Know These Cold

Feature	What It Does	Cost
DAX	In-memory cache → microseconds	Extra
TTL	Auto-delete expired items	FREE ✅
Streams	Item-level CDC → Lambda	Per read
Global Tables	Multi-region active-active	Extra
PITR	Restore any second, 35 days	Extra

ⓘ GSI = create anytime. LSI = table creation ONLY. Tested directly — don't get it backwards.

📊 Redshift — Performance Keys

EVEN Distribution

No joins, balanced load. Default.

KEY Distribution

Large table joins — fixes slow JOINs by co-locating matching rows.

ALL Distribution

Small dimension tables — full copy on all nodes.

AUTO Distribution

Redshift decides. Recommended default.

Spectrum = query S3 WITHOUT loading. COPY = fastest bulk load INTO Redshift. Zero-ETL = Aurora → Redshift (no pipeline).

Aurora Highlights

5× Faster

Than MySQL | 3× faster than PostgreSQL

Auto-Scaling

10 GB → 128 TB — never provision manually

Global Database

<1 second cross-region replication

Zero-ETL

Aurora → Redshift — 2026 exam favorite

MODULE 6

Analytics — Glue · Athena · EMR · Lake Formation · QuickSight

⚙️ AWS Glue — Serverless ETL

Source → Crawler → Data Catalog → ETL Job → Destination (S3/Redshift). Bronze → Silver/Gold. Athena queries immediately.

- **Crawlers:** Auto-discover schema → Data Catalog
- **Data Catalog:** Central metadata shared by Athena, Spectrum, EMR, Lake Formation
- **Job Bookmarks:** Track processed data — prevents reprocessing
- **DataBrew:** Visual no-code cleaning (250+ transforms)

⚠️ ⚠️ Glue = batch/micro-batch (~1–2 min startup). NOT real-time. Real-time → Lambda + Kinesis.

🔍 Athena — Serverless SQL on S3

Pricing: $5 per TB scanned. Formats: Parquet, ORC, CSV, JSON, Avro. Federated Query → RDS, DynamoDB via Lambda connectors.

💰 Athena Cost Optimization Formula

→ Convert CSV/JSON → Parquet or ORC → 87% less data scanned

→ Add DATE PARTITIONS → Skip entire partitions

→ Compress (Snappy or ZSTD) → Smaller, faster, cheaper

→ Partition Projection → Skip Glue Catalog lookups

✅ ✅ Result: Up to 90% cost reduction. Standard Athena tables = NO UPDATE/DELETE. Need DML? → Apache Iceberg tables.

EMR · Lake Formation · QuickSight

🐘 Amazon EMR

Big data Spark/Hadoop. Master = On-Demand. Core = On-Demand. Task nodes = Spot (60–70% savings). EMRFS → S3 persistent storage. EMR Serverless = zero cluster management.

🏛️ AWS Lake Formation

Column-level + Row-level security on data lake tables. LF-Tags → attribute-based access control at scale. IAM alone CANNOT do column-level access on S3 data — Lake Formation CAN.

📊 Amazon QuickSight

Serverless BI. Pay-per-session for readers. SPICE = in-memory cache → sub-second queries. ML Insights = anomaly detection + forecasting. SPICE = cached, NOT live.

MODULES 7·8·9

Security, Deployment & Migration — Universal Standards

🔐 Security — Three Pillars (Apply to Every Service)

Pillar 1: KMS CMK

Encryption at rest on ALL data stores. SSE-KMS = CloudTrail audit trail. Enable at creation — cannot add later. Customer-managed CMK for compliance.

Pillar 2: IAM Roles

NEVER Users or Access Keys for services. EC2 → Instance Profile. Lambda → Execution Role. Glue → Service Role. Temporary credentials, auto-rotated.

Pillar 3: VPC Endpoints

S3 + DynamoDB Gateway Endpoints = FREE. Deploy everywhere. Traffic on AWS backbone — NO internet, NO NAT fees. Interface Endpoints cost money.

🔑 IAM Policy Evaluation — Memorize This

01

Explicit DENY → WINS ALWAYS. No exceptions. Ever.

02

Explicit ALLOW → Allowed (only if no explicit deny exists)

03

No Statement → Implicit DENY (default — nothing allowed)

SCPs = guardrails ONLY (RESTRICT). IAM policies GRANT. Need both together.

🚀 Migration Tool Selector

On-prem files → S3/EFS	AWS DataSync
Same DB engine	AWS DMS
Different engines	AWS SCT + DMS
200 TB, internet too slow	AWS Snowball Edge
Exabyte-scale	AWS Snowmobile
SaaS → S3, no code	Amazon AppFlow

Snow decision rule: Data Size ÷ Bandwidth = Transfer Time. Internet > 1 week? → Snowball Edge.

🔒 Secrets Manager = auto-rotation built-in. SSM Parameter Store = manual rotation. DB passwords → always Secrets Manager. CloudTrail = WHO did WHAT. VPC Flow Logs = network traffic. GuardDuty = threats. Macie = PII in S3.

MODULE 10 | 🏆 FINAL WEAPON

Exam Command Center — Top 20 Traps + 60-Second Drill

Coach says: *"This table is worth 20 points on your exam. Read it twice."*

🚨 Top 20 Exam Traps — Pure Gold

#	The Trap ❌	The Truth ✅
1	"S3 eventual consistency for overwrites/deletes"	DEAD WRONG — S3 = STRONG consistency for ALL ops since Dec 2020.
2	"Multi-AZ = Read Scaling"	Multi-AZ = HA/Failover ONLY. Standby serves ZERO reads.
3	"Instance Store = Persistent storage"	Instance Store DIES on stop/terminate. EBS = Persistent. Always.
4	"Kinesis Data Streams auto-scales"	Data Streams = manual shard management. Amazon Data Firehose = auto-scales.
5	"It's called Kinesis Firehose"	REBRANDED → Official 2026 name = Amazon Data Firehose.
6	"Lambda handles 2-hour jobs"	Lambda hard max = 15 MINUTES. Longer → Glue / EMR / ECS.
7	"LSI can be added anytime"	LSI = TABLE CREATION ONLY. GSI = anytime.
8	"gp2 is the EBS default"	gp3 is the 2026 default. ~20% cheaper. Independent IOPS.
9	"Glue = real-time processing"	Glue = batch/micro-batch only. Real-time → Kinesis + Lambda.
10	"SCPs grant permissions"	SCPs = guardrails only (RESTRICT). IAM policies GRANT.
11	"Athena: UPDATE/DELETE on all tables"	Standard tables = NO DML. Need ACID → Apache Iceberg tables.
12	"DMS converts schema automatically"	DMS = migrates DATA. SCT = converts SCHEMA. Heterogeneous → need BOTH.
13	"SSE-S3 provides key audit trail"	SSE-KMS = CloudTrail key-level audit. SSE-S3 = ZERO key audit.
14	"VPC Peering is transitive"	NOT transitive. A↔B + B↔C ≠ A↔C. Need Transit Gateway.
15	"Zero-ETL works for any source/destination"	Aurora/RDS → Redshift (primary) + DynamoDB → Redshift/OpenSearch.
16	"Multi-AZ standby serves read traffic"	Multi-AZ standby = ZERO reads. Exists ONLY for automatic failover.
17	"Glacier Instant Retrieval = slow"	Glacier INSTANT = MILLISECONDS. The name tells you everything.
18	"Gateway Endpoints cost money"	S3 + DynamoDB Gateway Endpoints = 100% FREE.
19	"Secrets Manager = SSM Parameter Store"	Secrets Manager = auto-rotation built-in. DB passwords → Secrets Manager.
20	"CloudTrail logs network traffic"	CloudTrail = API calls (who did what). Network traffic = VPC Flow Logs.

⚡ 60-Second Pre-Exam Final Drill

S3
STRONG consistency ALL ops. Eventual = DEAD ❌

Glacier
Instant=ms | Flexible=min-hrs | Deep=12hrs (cheapest)

Athena
$5/TB. Parquet+Partitions = 90% savings ✅

EBS
gp3 default. ~20% cheaper. Independent IOPS ✅

Firehose
"Amazon Data Firehose" — NEW 2026 NAME ✅

Redshift
KEY distribution = fix slow JOINs ✅

Glue
Batch ETL ONLY. Crawler→Catalog→Job. NOT real-time ✅

Lambda
15-min MAX. Stateless. Roles NOT keys. Ever. ✅

DynamoDB
GSI=anytime. LSI=create ONLY. DAX=microsec. TTL=free ✅

Zero-ETL
Aurora/RDS→Redshift. No Glue. No Firehose. Seconds ✅

LakeForm
Column+Row security. IAM ALONE cannot do this ✅

StepFn
Orchestrate pipelines. Retry built-in. Least burden ✅

Secrets
Auto-rotate credentials. KMS encrypted. NOT SSM ✅

CloudTrl
WHO+WHAT+WHEN. GuardDuty=threats. Macie=PII in S3 ✅

IAM
Roles>Users ALWAYS. Least Privilege. DENY always wins ✅

Multi-AZ
HA/Failover ONLY. NOT read scaling. NEVER mix up ✅

S3 GW EP
FREE. Eliminates NAT costs. Deploy everywhere ✅

MSK
Managed Kafka. Drop-in compatible. MSK Serverless=zero mgmt ✅

EMR
Spot Task nodes=60-70% savings. EMRFS=S3 persistent ✅

VPC EP
Gateway (S3+DDB)=FREE | Interface=costs | DUAL WIN ✅

Coach Cloud's Final Word: "Two questions win EVERY scenario on this exam: **Question 1:** What does AWS WANT you to choose? **Question 2:** What requires the LEAST operational burden? Answer those two questions right — and that DEA-C01 badge is YOURS. Serverless over servers. Managed over self-managed. Roles over users. KMS on everything. Gateway Endpoints everywhere — they're FREE. Zero-ETL when Aurora meets Redshift. gp3 over gp2. Strong consistency always. Parquet + partitions = Athena savings. Firehose = Amazon Data Firehose. Say it right. You've done the work. You know the traps. You have the weapon. Now walk in there and get certified. I'll see you on the other side." — Coach Cloud, World-Class AWS Technical Coach, **Cloudexampro.com**