Alibaba Cloud

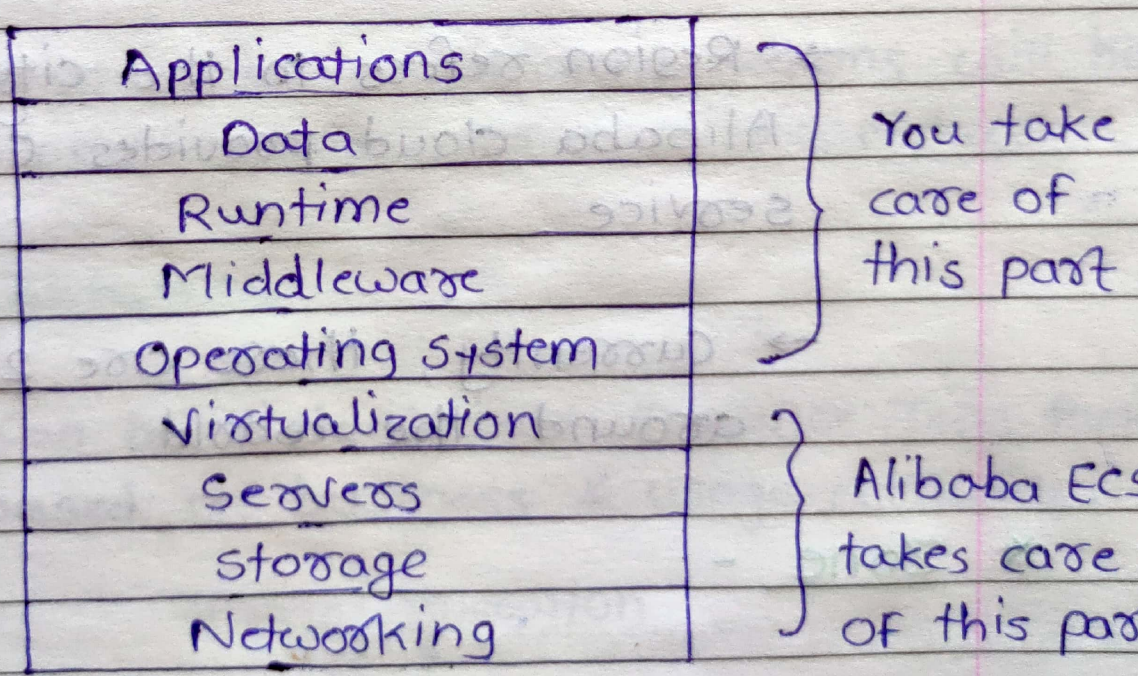
Page No.
Date: / /

Chapter 1
Manage & Operate Ecs

* **ECS (Elastic Compute Service) -**

→ Computing service with flexible processing capacity.

→ It's an IaaS offering from Alibaba.



→ Data Centers
↳ Server Racks
↳ Virtualization (Xen / KVM)
↳ ECS

* Apsara -

→ ECS runs on top of Apsara Distributed system.

→ Its virtualization technology is based on XEN/KVM & utilize Apsara Distributed File System Pangu for storage.

* Region -

→ Region refers to the city/area where Alibaba cloud provides Cloud computing service.

→ Currently, there are 21 regions around the world.

* Zone -

→ Zone refers to a physical data center within a region with independent power supply & network.

→ Creating resources in multiple zones can facilitate high availability.

→ Creating resources in the same zone facilitates low communication latency.

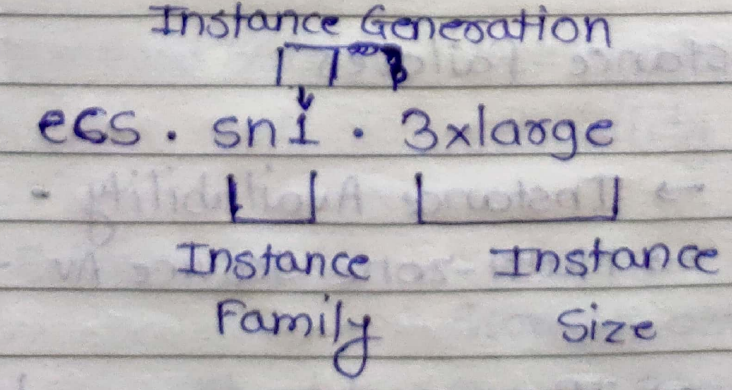
→ Regions contain betn 1 & 3 AZs.

* Resources within a same region can communicate over a Private Network.

* Resources in different regions will have to communicate over a public Network.

* ECS Instances -

→ Can be divided into Instance Type families based on business & usage scenarios.



For Instance Metadata

$ curl http://100.100.100.200/latest/~~metadata~~/

(*) Purchasing Methods -

(1) Pay-As-You-Go (PAYG)
↳ By seconds

(2) Subscription
↳ Significant discount (up to 85%) compared to PAYG.
↳ Monthly / Yearly subscriptions

(3) Preemptible Instance
↳ price varies depending on resource demands & supplies.
↳ Low prices compared to PAYG.
↳ Hourly bidding.

(4) Reserved Instances
↳ Discount coupon on billing of PAYG.

(*) Instance failover

→ Instance Availability - 99.975% per month
→ Multi-zone service Av. - 99.995% pm

→ If a machine goes down, it gets replaced by another machine in same zone

for User Data

$curl http://100.100.100.200/latest/user-data

* ECS storage

→ ECS utilizes cloud Disk as the system & data disk.

cloud Disk - (Block Storage)

→ Based on Apsara Dist. File System Pangu

→ 3 redundant copies stored separately on different physical servers under different switches.

→ Mount / Unmount is supported.

3 Types

① Ultra Disk - High cost-effective, high data reliability, medium random IOPS.

② Standard SSD - High random IOPS, high data reliability, high performance.

③ Enhanced SSD - Next gen Distributed Block Storage arch, 25 Gigabit Ethernet, Remote Direct Memory Access (RDMA). 1 million random IOPS (upto)

→ Cloud Disk can be separately purchased.

→ can be mounted to any instance in the same zone, does not support cross-zone mounting.

→ can be mounted to only one instance at a time.

→ Users can decide to release the disk together with the instance or not.

→ Supports volume-based payment (charged hourly) only.

⊛ Disk Snapshot

→ Can be stored in OSS

→ can be manual / automatic (policies).

Auto Snapshot

↳ Max of 100 Auto snapshot policies in each region.

→ Each user can create 64 snapshots for each disk.

→ Snapshot fee is charged based on the size of storage space & time length it is kept.

→ Snapshots are incremental. only the changed parts beth two snapshots will be recorded.

→ Changed parts are indexed with application numbers.

(*) ECS Image -

→ Template for ECS instances.
→ Merely a snapshot of system disk.

Types

① System Image - officialy provided by Alibaba.
② Marketplace Image - Third-party images
③ User-Defined Image (custom) - Created from system disk snapshot of instance.
④ Shared Image - shared among Alibaba cloud accounts.

* Ecs Networking -

(*) Alibaba cloud VPC

→ Virtual Network within Alibaba Cloud based on SDN technology.

→ Users can customize their own network topology.

→ Can Integrate existing Data Center through dedicated line or VPN to form a Hybrid Cloud.

Key Features

(1) Security Isolation

→ VPC adopts VxLAN protocol. Each VPC is assigned with an independent tunnel ID

→ VPC controls L2 of ARP broadcast domains within a single vNIC.

(2) User Defined Network

→ Users can customize VPC addresses as well as private IPs of Ecs instances.

③ Dedicated Line / VPN Access
→ connect on-prem DCs.

④ Elastic public network IP (EIP)
→ can bind to any ECS instance in the same region.

* VPC components

① vSwitch

→ A basic network device of a VPC & used to connect different cloud instances in a subnet with a VPC.

→ You can segment your VPC into subnets by adding one or more vSwitches.

→ A VPC can have max. of 24 vSwitches.

② vRouter

→ Its a hub in VPC that connects all vSwitches in VPC & also serves as a Gateway to connect VPC to other networks.

⊛ Security Groups -

→ Specifies one or more firewall rules, including protocol, port & source IP that are allow to access.

→ These rules are effective to all instances within the Security Group.

→ Every instance belongs to atleast one security Group.

Every SG can have 2 ways to authorize access :

1. Specify Source IP address field.
2. Specify Source SG ID. (only works in Intranet.)

⊛ Every user can have 100 SG at max.

⊛ Every SG can contain 1000 instances at most.

⊛ A single instance can join up to 5 SG

⊛ Every SG can set 100 rules at most.

⊛ System create a Default SG for each user. It allows public network connection.

Chapter 2
Manage & Operate RDS

* **RDS** - (PaaS Service)

→ Managed, highly available Relational Database service provided by Alibaba.

→ Supports:
- ↳ MySQL
- ↳ SQL Server
- ↳ Postgre SQL
- ↳ Maria DB

→ Fully managed host & OS
- ↳ No access to db host OS
- ↳ Limited ability to modify configs

→ Fully managed storage
- ↳ Upto 20000 IOPS & 2TB storage for each instance.

→ Grab-and-go service with gurantee of up to 99.99% business availability.

→ Automatic failover.

Page No.
Date: / /

* RDS Instance -

→ A user can decide the database engine, CPU, memory, storage, Network used by the RDS instance.

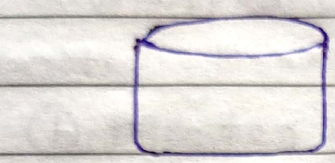
RDS provides 3 editions of instances :

① Basic Edition
- → Uses the storage- computing isolated arch. & a single computing node
- → High cost-effectiveness

② High-Availability Edition
- → classic HA arch. with one master node & one slave node.
- → Uses ephemeral SSD storage.

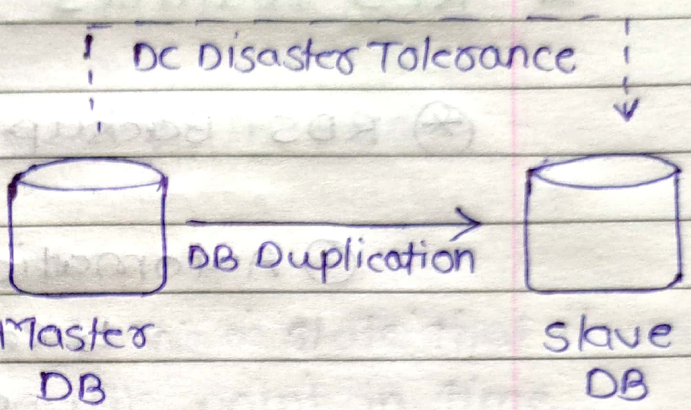
③ Finance Edition
- → Uses 3 node arch with one Master node & 2 slave nodes.
- → Data consistency through synchronizing multiple log copies.
- → Provides financial level data reliability & cross-IDC disaster tolerance.
- → Free SQL auditing.

Basic Edition



Master Database

High-availability Edition



→ HA Edition runs the service on 2 instances in the same or different zones with automatic data synchronization.

→ When primary instance is not available, system will automatically initiate failover to the standby or secondary instance.

* A single-zone RDS instance can withstand server or rack failures.
A multi-zone RDS instance can survive failure of an entire Data center.

→ Currently, no extra charge for multi-zone RDS instances. Can purchase directly or convert single-zone to multi-zone by using Inter-zone migration.

Page No.
Date: / /

* RDS Features -

⊛ RDS Backup -

① Automatic Backup
↳ configure a backup policy to adjust the cycles of Data & log backup.
↳ supports physical backup only.

② Manual Backup
↳ supports both physical & logical backup.

→ Instance Backup File
↳ If Total backup data < 5120 MB
↳ No. extra charge

⇒ Users can obtain temporary links to download backup data as needed.

→ You can also dump backup files to an OSS Archive Storage for cheaper & steady offline storage.

⊛ RDS Recovery -

3 Types of Data recovery in RDS -

① **Master node roll-back**
- ↳ Restore a node to a state that it was in at a specific point in time.

② **Slave node repair**
- ↳ Automatically create a new slave node to reduce risks when an irreparable failure occurs to the slave node.

③ **Temporary Instance (Recommended)**
- ↳ Create a temp instance with your backup & migrate data to master instance.
- ↳ Inherits account & password of backup file but keep network type of current instance.
- ↳ Uses its instance name as password.
- ↳ Only 1 temp instance can be created at the same time.
- ⊛ ↳ A Temp instance is valid for 48 Hours.

(*) Read-only Instance -

→ ApsaraDB for MySQL allows Read-only instances to be directly attached to RDS in order to distribute the read pressure on the Master instance.

→ Each Read-only instance has an independant connⁿ string & read pressure can be automatically distributed on the application side.

(*) RDS Diversified Data Storage -

(1) Memcache & Redis

↳ Provides high response speed for cached 'hot data.

↳ Cache area can support more higher Requests Per Second than the RDS.

(2) OSS

↳ Helps reduce the storage pressure on RDS.

↳ For unstructured data.

* RDS Monitoring

① Real-Time Monitoring

→ Crucial info. like CPU Utilization, IOPS, connections & disk space utilization in real-time.

② ApsaraDB for RDS Manager

→ Schedule inspections & customize inspection metrics.

→ Expert suggestions on SQL statmnt optimization & diagnostic report on performance of your instances.

→ CloudMonitor :

↳ Displays RDS operation status based on 4 metric items:

↳ Disk Usage

↳ IOPS usage

↳ Connection Usage

↳ CPU usage

(*) RDS Security -

→ RDS provides a 3-level security Defence system to protect DB.

(1) Anti-DDoS
- ↳ Real-time traffic monitoring at network entry point.
- ↳ If any high-traffic attack is identified, source IPs are either cleaned or blacklisted.

(2) IP Whitelist configuration
- ↳ Supports configuration of upto 1000 IP addresses.

(3) Protection of DB from various attacks
- ↳ Intercepts SQL injection, brute force attacks, etc.
- ↳ SQL Audits.

→ RAM
- → You can use Alibaba cloud Resource Access Mgmt (RAM) to control who can perform actions on RDS.

→ Access control

→ RDS requires IPs to be whitelisted before clients running in them can access any DB instance.

→ RDS allows setting up SSL connⁿs from ECS to RDS via the LBs. This allows end-to-end encryption.

⊛ RDS Migration Service

→ Migrate data from a local DB to Apsara DB or migrate ApsaraDB to another instance.

Data Transfer service (DTS) Tool - 3 modes

① Structural migration -

↳ Tables, views, triggers, stored procedures & stored functions.

② Full Migration -

↳ All data in source DB to target instance.

③ Incremental Migration

↳ DTS synchs data changes made in migration process to target instance.

Chapter 3
Manage & operate OSS

* What is OSS -

Types of storage

① File Storage
② Block Storage } Traditional

③ Object Storage } Cloud Storage

Object Storage Service (OSS)

→ Encrypted & secure cloud storage service which stores, processes & accesses massive amount of data from anywhere in the world.

Features of OSS

① Scalability
→ Massive volume of 50 PB for a single bucket.
→ Massive no. of objects.

② Simplicity -

→ RESTFUL API

→ Flexible object size varying from 0-48.8 TB

→ Flexible upload : Normal / Multipart / Append.

③ Security -

→ Multi layer protection & Anti-DDoS

→ Multi-user isolation

④ Availability -

→ 3 replica in 3 AZs.

→ 99.9999999 % data reliability.

→ 99.9 % Service Availability.

→ Auto failover with cross-region replication & Auto. service expansion.

⑤ Inclusiveness -

→ Integration with other services

→ EMR, MaxCompute, Hybrid DB, Hadoop FileSystem

→ EBS snapshot, RDS snapshot, Docker Image

→ RAM, STS, SLS, Cloud Monitor

* OSS Concepts -

(*) Buckets

Serve as containers for objects stored in OSS.

→ Every object must be kept in a bucket.
→ Bucket name is globally unique & cannot be changed.
→ No limit on no. of objects in each bucket.
→ An appn can have one or more buckets.

(*) Objects

Fundamental entities stored in OSS Buckets.

→ consist of data & metadata
→ Object is uniquely identified within a bucket by a key (name)
→ Each file of a user is an object.
→ File size limitation
  ↳ via PutObject mode, object size cannot exceed 5 GB
  ↳ via Multipart mode, object size cannot exceed 48.8 TB
→ object includes key, data & metadata.

Page No.
Date: / /

(*) Keys

Objects unique identifier for an object within a Bucket.

→ Every object in a bucket has exactly one key
→ combination of bucket, key.

e.g.

http://my-bucket.oss.ap-southeast-1.aliyuncs.com/algo.jpg.

* Bucket Operations

① Create Bucket

→ choose an existing region
→ Bucket name is globally unique & cannot be changed after creation.
→ Bucket name must:
↳ contain lowercase letters, nos. & hyphen
↳ Begin & end with lowercase letters/nos.
↳ 3-64 characters in length.

② Configure Bucket ACL

→ can be configured during bucket creation or can be modified after creation.

→ only bucket owner can change ACL.
→ 3 Types of Bucket ACL :
↳ Private
↳ Public Read
↳ Public Read Write

③ check Bucket Info
→ On OSS mgmt. console.

④ Delete Bucket
→ To prevent delete by mistake, users are not allowed to delete a non-empty bucket.

* Object operations -

① Create Directory & upload files
→ Directory naming rules:
↳ Use valid UTF-8 characters
↳ Slashes (/) are not allowed.
↳ Sub directories with names of ".." are not allowed.
↳ Directory name must be between 1-254 characters.

② Access File

→ Users can get access URLs from OSS console to access the file.

③ Delete File

→ 3 methods:

↳ Delete a single file at a time

↳ Delete multiple files at a time. (supports 1000 objects at a time)

↳ Auto-delete using pre-defined life-cycle policies.

* Image Processing

→ Alibaba Cloud OSS Image Service (IMG) is an image processing service with massive capacity, high security, low costs & high reliability.

→ offers image processing APIs.

http(s):// img-demo.oss.ap-southeast-1.aliyuncs.com / example.jpg ? x-oss-process = ...

Bucket name — object — Query string

Parameters mode
e.g. image/resize,w_200

Style mode
e.g. style/style.name

* Website Hosting & Monitoring -

→ You can host a static website on OSS

→ Monitoring of all of the key metrics as part of Cloud Monitor.

* Bind a custom Domain & Anti-leech -

→ You can bind a custom domain name to your OSS bucket in OSS console.

→ Add a CNAME record in DNS

→ After CNAME resolution, OSS automatically processes access requests to custom domain.

Anti-leech -

→ OSS supports anti-leech based on the field referer in the HTTP header.

→ You can use OSS console or API to configure referer whitelist for a bucket or whether to allow access by requests where referer is blank.

→ Referer field supports wildcards * & ?

→ when the whitelist is empty, system checks if referer field is Null (otherwise all requests get rejected.)

→ Anti-leech verification is performed only when users access objects through URL signatures or anonymously.

when request header contains the "Authorization" field, anti-leech verification is not performed.

* OSS Security Features

① HTTPS endpoint

② Access & control
→ Bucket & object level ACL & policies.

③ Server Side Encryption
→ SSE-KMS & SSE-OSS.

④ Identity Authentication (RAM & STS)
→ Cross-account authorization
→ Temporary access auth.

Page No.
Date: / /

Chapter 4
Manage & operate SLB

* SLB

→ Server Load Balancer (SLB) automatically distributes incoming application traffic across multiple applications, microservices & containers hosted on ECS instances.

→ With Server Guard, SLB can defend DDoS attacks including: CC, SYN Flood, etc.

components:

① SLB Instances

→ To use SLB, you must create SLB instance with atleast 1 Listener & 2 ECS instances configured.

② Listeners

→ checks client requests & forward requests to backend servers. Also performs health check on Backend server.

③ Backend Servers

* SLB Features -

⊛ SLB Architecture

→ SLB can provide L4 SLB (TCP, UDP) & L7 SLB (HTTP, HTTPS) according to protocol used by different applications.

→ SLB system consists of 3 parts:
  ↳ Layer-4 SLB
  ↳ Layer-7 SLB
  ↳ control system

→ Layer-4 SLB uses customized open source LVS (Linux Virtual Server).

→ Layer-7 SLB uses open source software Tengine.

→ control system is used to configure & monitor the SLB system.

Layer-4
  → Supports TCP & UDP with SSL
  → operates at Transport layer
  → No header modification
  → can only make limited routing decisions by inspecting first few packets in the TCP stream.

Layer-7

→ Supports HTTP & HTTPS
→ connection terminated at the LB & pooled to the server.
→ Headers may be modified.
→ X-Forwarded-For header contains client IP address.

⊛ SLB Security

→ 5 Gbps DDoS protection.
→ All traffic from Internet first goes through Anti-DDoS Basic & then arrives at LB.
→ can defend attacks like SYN Flood, UDP flood, ACK flood, ICMP flood, HTTP Flood, DNS Query flood, etc
→ sets the cleaning Threshold & Black hole threshold according to public nw bandwidth configured for SLB.

Flow cleaning - When incoming traffic exceeds cleaning threshold or matches certain attack traffic method, Flow cleaning is triggred

Black Hole - When traffic exceeds default traffic threshold, gets triggred & discards further traffic.

Page No.
Date: / /

* SLB components

(*) SLB Instances

① Public Network SLB : Pay-As-You-Go (PAYG)

② Private Network SLB : Free
- → can only be used in private NW env.
- → Recommended for a multi-tier arch.

(*) Listeners

→ Define the protocol & port on which the LB listens for incoming connections

→ Each LB needs atleast 1 listener to accept incoming traffic.

→ Routing (forwarding) rules are defined on listeners.

→ Session stickiness is set at listener level.

→ Health check configurations are done at listener level.

→ Peak BW can be done at listened level.

(*) Upto 50 listeners can be added to a SLB instance. Each listener corresponds to an appn deployed on the ECS.

✱ Backend servers

→ SLB forwards external requests to backend servers to process.

→ can be registered with same listener using multiple ports.

→ SLB doesn't support cross-region deployment. Ensure that the region of SLB & Backend ECS instances is same.

→ SLB does not limit OS used.

✱ Listeners : Forwarding Rules

→ Support 3 scheduling algorithms :

① Round Robin - (RR)

→ Requests are distributed across the group of backend ECS instances sequentially.

② Weighted Round Robin - (WRR)

→ You can set weight for each Backend server.

→ Servers with higher weights receive more requests.

③ Weighted least connections - (WLC)

→ In addition to the weight set, the no. of connⁿ to client is also considered.

✳ SLB Additional Settings

✳ SLB Multi-zone Disaster Tolerance

→ Using the primary/backup zone feature in SLB.

✳ SLB Cross-Region Disaster Tolerance

→ Using DNS.

✳ SLB Auto Scaling

→ Detects impaired Ecs instances

→ Repairs those instances automatically.

→ Automatically scales your Ecs Fleet.

Chapter 5

Manage & Operate Auto Scaling

* **Auto Scaling** -

→ Management service that allows users to automatically adjust elastic computing resources according to business needs.

Modes:

① **Elastic Scale-up**

→ Adds additional computing resource to the pool during peak hours.

② **Elastic scale-down**

→ Releases ECS resources when requests decrease.

③ **Elastic Self-Healing**

→ When an unhealthy instance has been detected, automatically replace it with new one.

Characteristics:

→ On-demand
→ Automatic
→ Flexible
→ Intelligent

* Auto scaling Concepts -

(*) Auto Scaling Functions -

(1) Scales up instances
↳ Scheduled scaling - Specified time
↳ Dynamic scaling - Based on cloud Monitor metrics.

(2) Supports SLB configuration.

(3) Supports RDS Access whitelist.

(*) Auto scaling concepts -

(1) Scaling Group
→ collection of ECS instances with similar configuration.
→ Defines min & max no. of instances in the group, associated SLB & RDS instances & other attributes.

(2) Scaling Configuration
→ Defines config. info of ECS instances.

(3) Scaling Rule
→ Defines scaling actions.

④ Scaling Activity

→ When a scaling rule is successfully triggered, a scaling activity is generated.

⑤ Scaling Trigger Task

→ Task that trigger a scaling rule.

⑥ cool-down period

→ Period during which Auto Scaling cannot execute any new scaling activity after another scaling activity is executed successfully in a scaling group.

④ Auto Scaling Procedure

① Create a Scaling Group
② Create Scaling Configuration
③ Enable scaling Group
④ Create scaling rule
⑤ Create scheduled task
⑥ Create an alarm task (Cloud Monitor API PutAlarm Rule)

Page No.
Date: / /

(*) Relationship between Auto Scaling concepts :

→ Scaling Group includes :
- ↳ Scaling Configuration
- ↳ Scaling Rules
- ↳ Scaling Activities

→ Removing a Scaling Group also removes all 3 associated concepts.

→ Scaling Trigger Task has 2 categories :
- ↳ Scheduled Task
- ↳ Cloud Monitor Alert Task

These tasks are independent of Scaling group. Deleting the scaling group does not delete the task.

→ Only 1 scaling activity can be executed at a time in Scaling Group.

→ A scaling activity cannot be interrupted.

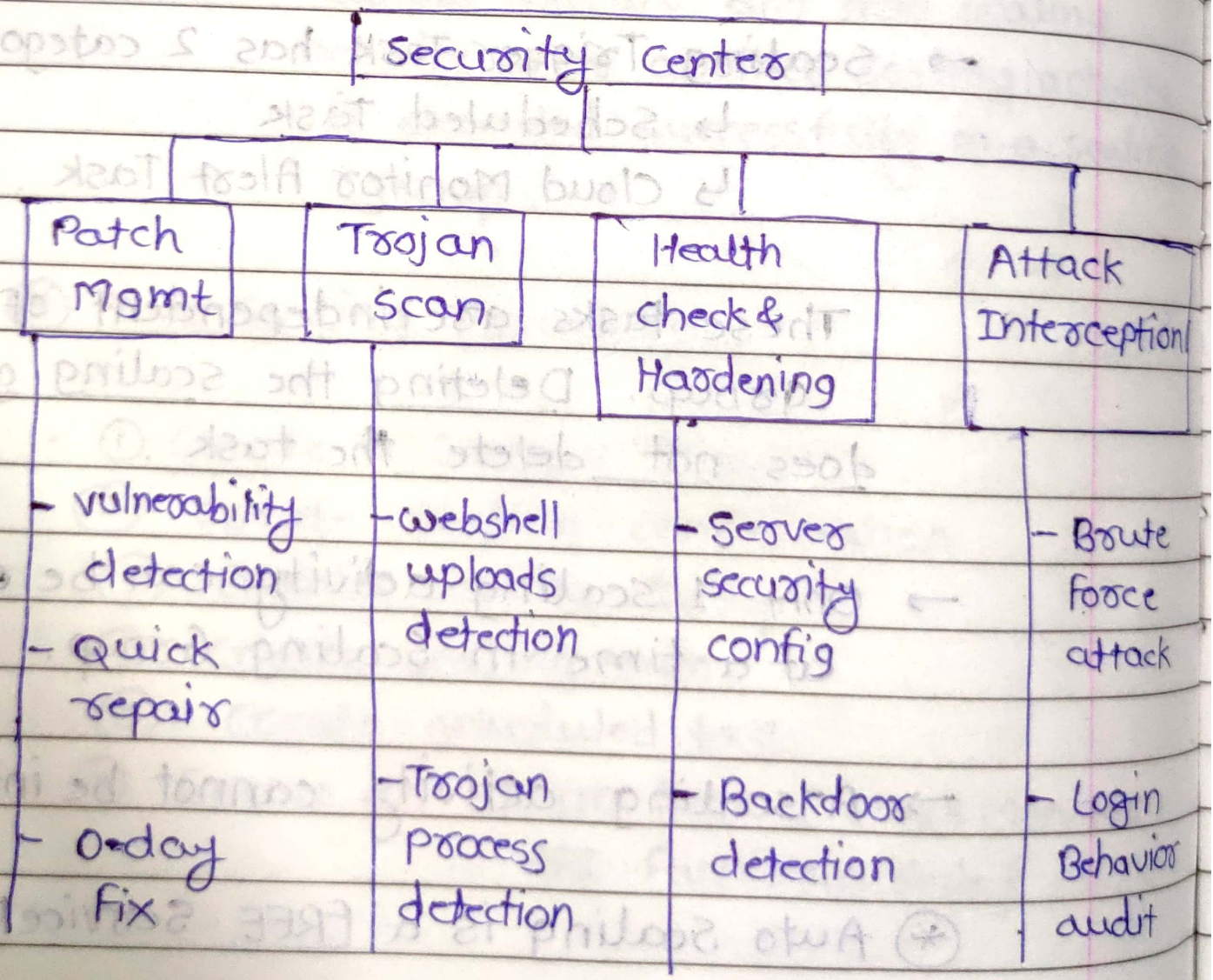
(*) Auto Scaling is a FREE Service!

Chapter 6
Cloud Security Intro

* Host Security

Security Center

→ Server security O&M manager offered by Alibaba cloud Security.



→ Patch mgmt, Health check & Hardning

* Network Security

Anti-DDoS (Basic, Pro & Premium)

→ Security service to safeguard your data & applications from DDoS attacks.

Premium Version

→ Exclusive IP for services deployed in non-China mainland.
→ Global near source mitigation using BGP anycast technology.
→ Unlimited traffic protection.
→ Insurance mode for basic service.

	Anti-DDoS Basic	Anti-DDoS Pro (china) & Premium (overseas)
Free of charge	Yes	No
Protect attacks above 5 Gb/s	No	Yes
can protect IDC outside Alibaba Cloud	No	Yes
Black Hole Policy	Yes	No

* Application Security

Web Application Firewall (WAF)

→ It is an appliance, server plugin or filter that applies a set of rules to HTTP traffic.

→ Protection against OWASP Top 10 threats.

Features

→ Data Breach Protection
→ Crawler Protection
→ Human-machine identification
→ Big data & threat intelligence.
→ Oday vulnerability hotfix
→ Accurate access block.

Chapter 7
Alibaba Cloud china Gateway Service

* Alibaba china connect

→ Service to resolve 3 challenges when building a china ready website.

3 components :
1. Complete webhosting service
2. ICP license
3. china Alibaba Resources

ICP License

→ An Internet Content Provider (ICP) license is a legal & mandatory requirement from the Ministry of Industry & IT (MIIT) for all websites hosted on a server in china.

→ 6 Step process
1. Register an account
2. Submit company info
3. Alibaba cloud verification
4. Provide photo verification
5. MIIT verification
6. success .

* Alibaba Cloud Support Services -

Consulting Services
- → Readiness Assessment
- → Solution Architecture
- → Security Architecture
- → Service mgmt consulting

Managed services
- → Aftersales Support } International
- → Event management }
- → Guided / Frontline / Platform Operation

System Integration services
- → Initial Deployment
- → Migration Implementation
- → Load Performance Test } Intl
- → Architecture optimization services

Training Services
- → Alibaba cloud Product Training
- Intl { → Alibaba cloud Arch. design Training
- Intl { → Alibaba cloud Product operation Training
- → Transformation Assistance