1

Topic Level Scoring:

2

3

1. AWS Fundamentals: 100%

4

5

2. Designing and Developing: 100%

6

7

3. Deployment and Security: 87%

8

9

4. Debugging: 100%

10

11

My last minute notes - taken from the AcloudGuru lectures only. I was hoping to get 100% but not sure which 2 questions I got wrong out of 55 (scratching my head :))

12

13

http://169.254.169.254/latest/meta-data/

14

15

SDK -

16

17

IOS, Android, Browser (Java scripts)

18

19

Java, .NET,

20

21

Node.js, PHP, Python, Ruby

22

23

Go, C++

24

25

- SQS: - message oriented API

26

27

SQS - Message can contain upto !256KB of text, billed at 64KB chunks,

28

29

Single request can have 1 to 10 messages upto maximum of 256KB payload

30

31

Even though there is one message of 256Kb its basically 4 request for billing since (4 \* 64KB)

32

33

NO ORDER - SQS messages can be delivered multiple times in any order

34

35

Design - you can have 2 priority queues for priority based message one for higher and other for lower priority

36

37

EC2 instances always poll for messages from the queue (pull from the queue and not push)

38

39

Visibility timeout always start from when the application instance polled the message.

40

41

Great design - Visibility timeout expires that means there is a failure somewhere since that message was polled but not processed and hence not deleted so other some other process will poll the message again and visibility timeout starts again.

42

43

Visibility timeout by default is 30 Seconds up to 12 hour maximum (ChangeMessageVisibility) / maximum visibility

44

45

Maximum long polling timeout 20 seconds (http://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/sqs-long-polling.html) —ReceiveMessageWaitTimeSeconds

46

47

Messages in the Queue can be retained for up to 14 days

48

49

First 1 million request ares free, then $0.50 PER EVERY MILLION REQUESTS

50

51

SNS

52

53

works on a publish - subscribe model, SNS notifies the message, and hence push based approach. Inexpensive pay as you go

54

55

CloudWatch or Autoscaling triggers SNS

56

57

SNS can notify to Email, Text / SMS, SQS or any HTTP end point.

58

59

protocols: HTTP, HTTPS, EMAIL, EMAIL-JSON, SQS or Application - messages can be customized for each protocol

60

61

SNS messages are stored redundantly to multiple AZs

62

63

SNS Dataformat - JSON (Subject, Message, TopicArn, MessageId, unsubscribeURL etc..)

64

65

$0.50 per 1 million SNS request

66

67

Different price for different recipient types

68

69

to HTTP: $0.06 / 100,000 notifications deliveries

70

71

to EMAIL: $2 / 100,000 notifications deliveries

72

73

to SMS: $0.75 / 100 notifications deliveries

74

75

http://docs.aws.amazon.com/sns/latest/dg/mobile-push-send-devicetoken.html (CreatePlatformEndpoint API)

76

77

SWF - task oriented API

78

79

Simple Work flow - human interaction to complete order or collection of services to complete a work order.

80

81

Workers - interact with SWF to get task, process received task and return the result

82

83

Deciders - program that co-ordinates the tasks, i.e. - ordering, concurrency and scheduling

84

85

Workers and Deciders can run independently

86

87

TASK is only assigned ONCE and NEVER DUPLICATED (key difference from SQS where messages can be processed multiple times)

88

89

SWF Domain - think of it as a container for the work flow. you can register a domain by Console or API

90

91

Maximum workflow processing time can 1 year (equivalent seconds) - SQS is 12 hours processing time

92

93

CloudFormation

94

95

Use of CFT, Beanstalk and Autoscaling are free but you pay for the AWS resources that these services create.

96

97

Fn::GetAtt - values that you can use to return result for an AWS created resource or used to display in output

98

99

By Default - rollback everything on error

100

101

Infrastructure as a code, Version controlled, declarative and flexible

102

103

ElasticBeanstalk

104

105

Its uses ASG,ELB,EC2,RDS,SNS and S3 to provision things.

106

107

Environment Tier - Webserver, Worker

108

109

Predefined Configurations - IIS, Node.JS, PHP, Python, Ruby, Tomcat, Go, .NET,

110

111

preconfigured docker: Glassfish, Python or generic docker

112

113

Environment URL - has to be unique

114

115

Dashboard - Recent events, Monitor, Logs, Alarms, Upload and Deploy and Configurations

116

117

Configuration - Scaling, Instances (DIRTMCG instance types, key pair), Notifications, Software configuration (e.g. PHP.ini), Networking tier (ELB, VPC config), Data tier(RDS)

118

119

Environment properties (Access key and secret key as parameters)

120

121

DynamoDB

122

123

fast - flexible No sql database - single digit ms latency, fully managed, supports document and key-value (web, gaming, ad-tech, IOT)..

124

125

Table, Item (row), attribute (key - value)

126

127

Eventual Consistent Reads vs Strongly Consistent Reads

128

129

Read Capacity Units, Write Capacity Units (can be scaled up) - push button scalability

130

131

Writes are written to 3 different location / facilities/ datacenter (synchronous) - Amazon DynamoDB synchronously replicates data across three facilities in an AWS Region, giving you high availability and data durability.

132

133

Two types of primary key -

134

135

(1) Single Attribute (think unique id) - Partition Key (Hash Key) composted of 1 attribute (no nesting allowed here) - Partition key will help determine the physical location of data.

136

137

(2) Composite key (think unique id and range) - Partition Key(Hash Key) & Sort Key (Range key - e.g date) - composed of 2 attributes - if two data have same partition key (same location) it must have a different sort key, and they will be stored together on single location.

138

139

Secondary Indexes

140

141

(1) Local Secondary Index - Same Partion Key + Different Sort Key ( can only be created while creating the table, cannot be added/removed or modified later)

142

143

(2) Global Secondary Index - Different Partition Key + Different Sort Key ( can be created during the table creation or can be added later or removed / modified later)

144

145

DynamoDB Streams

146

147

use to capture any kinda modification to the dynamo db table, Lambda can capture events and push notifications thru SES

148

149

Table can be exported to csv (either select all items )

150

151

Query vs Scan

152

153

Query operation finds item in a table using only primary key attribute values , must provide partition attribute name and the value to search for, you can optionally provide a sort key attribute name and value to refine search results (e.g. all the forums with this ID since last 7 days). By default Query returns all the data attributes for those items with specified primary keys. You can further use ProjectionExpression parameter to only return a selected attributes.

154

155

Query results are always sorted by the sort key (ascending for both numbers and string by default). To reverse the sort order set the ScanIndexForward parameter to false

156

157

By Default Queries are going to be Eventually consistent but can be changed to StronglyConsistent.

158

159

Scan operation is basically examines every item - e.g. dumping the entire table, by default Scan returns all the data attributes but we could use ProjectionExpression parameter to only return a selected attributes.

160

161

Query operation is more efficient than scan operation

162

163

For quick response time design your table in a way that you can use Query Get or BatchGetItem API (read multiple items - can get upto 100 items or up to 1MB of data) ,

164

165

Alternatively design your application to use scan operation in a way that minimize impact of your table’s request rate since it can use up the entire table’s provisioned throughput in a single scan operation

166

167

DynamoDB Provisioned Throughput calculations

168

169

Items == rows

170

171

Read Provisioned Throughput

172

173

All units are rounded up to 4KB increments

174

175

Eventual Consistent reads (default) consist of 2 reads per second

176

177

Strongly Consistent reads consist of 1 read per second

178

179

( Size of Read Rounded to nearest 4KB Chunk / 4 KB \* no of items ) / 2 <— if eventual consistency

180

181

( Size of Read Rounded to nearest 4KB Chunk / 4 KB \* no of items ) / 1 <— if strongly consistency

182

183

Write Provisioned Throughput

184

185

All units are rounded up to 1KB increments

186

187

All writes consist of 1 write per second

188

189

( Size of write in KB \* no of items ) / 1

190

191

When you exceed your maximum allowed provisioned throughput for a table or one or more global secondary index you will get 400 HTTP Status code - ProvisionedThroughputExceededException

192

193

AssumeRolewithWebIdentity role

194

195

Idempotent conditional write

196

197

Atomic counters - always need to increment so its not idempotent

198

199

if data is critical and no margin of error then must use Idempotent conditional write.

200

201

http://docs.aws.amazon.com/amazondynamodb/latest/developerguide/Limits.html#limits-tables

202

203

Only Tables(256 table per region) and ProvisionedThroughput(80 K read, 80K write per account for US east, 20K for other regions) limits can be increased

204

205

http://docs.aws.amazon.com/amazondynamodb/latest/developerguide/QueryAndScanGuidelines.html (Reduce Page Size for Scan operation and Isolate Scan Operation)

206

207

S3

208

209

secure, durable, highly scalable object store (1 byte to 5TB), universal namespace (must be unique bucket - regardless of regions),object based key value store, VersionID, Metadata, ACL

210

211

The total volume of data and number of objects you can store are unlimited. Individual Amazon S3 objects can range in size from 1 byte to 5 terabytes. The largest object that can be uploaded in a single PUT is 5 gigabytes. For objects larger than 100 megabytes, customers should consider using the Multipart Upload capability. it mean the largest single file into S3 is 5G, but after the 5G files are in S3, they can be assembled into a 5T file,

212

213

You can use a Multipart Upload for objects from 5 MB to 5 TB in size (Exam question, scenario where more than 5GB file needs to be uploaded)

214

215

object based storage vs block based Storage (EFS)

216

217

data is spread out in multiple facilities, you can loose two facilities and still have access to files

218

219

For PUTS of New Objects (Read after Write Consistency), For Overwrite PUTS and DELETE (Eventual Consistency)

220

221

http://docs.aws.amazon.com/general/latest/gr/awsservicelimits.html#limits\_s3 ( Number of S3 bucket limit per account — 100)

222

223

---Storage Tiers/ Class

224

225

S3 Standard - Durability (11 9s), Availability (99.99 %) - reliable regular for just about everything

226

227

S3 IA (Infrequent Access) - Durability (11 9s), Availability (99.9 %) - accessed every 1 month to 6 months or so (infrequent) but rapid access and low retrieval time (few ms)

228

229

S3 RRS(Reduced Redundant Storage)- Durability (99.99%), Availability (99.99 %) - less durability (data that can easily be regenerated - e.g thumbnails) - cheapest of all s3, less fault tolerant then the other two since you are willing to loose the data, reproducible data

230

231

Glacier - for archival only (3 to 5 hours restore time)

232

233

S3 price - charged for Storage, number of requests, data transfer (tiered so more you use less charge)

234

235

bucket name has to be all lowercase letters

236

237

S3 for static website hosting (Static Website Hosting > Enable website hosting) - no dynamic

238

239

Any time you create a bucket nothing is publicly accessible / Any time you add an object to a bucket its private by default (you will get 403) > Make the files public (even for public hosting)

240

241

every object inside the bucket can have different storage class (S3 standard, S3-IA, S3-RRS) and you can turn on server side encryption (AES - 256)

242

243

regular bucket link: https://s3-eu-west-1.amazonaws.com/ankittest <— https

244

245

bucket with Static website hosting: http://ankittestsite.s3-website-eu-west-1.amazonaws.com <— http (has to be for static hosting), you can turn it into SSL / https with cloudfront though

246

247

CORS (CROSS ORIGIN RESOURCE SHARING) - to avoid the use of proxy

248

249

Versioning - once enable you cannot disable versioning / although it can be suspend it , if you want to turn it off delete the bucket and recreate (version id)

250

251

Once you delete the delete marker, you can get the file back that you have deleted while versioning on

252

253

every version is stored separately in the bucket for each version / might not be a good choice for cost perspective for large media files., multiple updates use case also not ideal for versioning.

254

255

Versioning’s MFA Delete Capability can be used to provide additional layer of security.

256

257

Cross Region Replication - (requires versioning enabled on source and destination buckets)

258

259

you can enable - need source and destination bucket (create a new bucket, source bucket will not show up on drop down of destination)

260

261

Existing objects will not be replicated, only new objects will be replicated across the region

262

263

Lifecycle management in S3

264

265

(1) when versioning is disabled

266

267

Transition to IA S3 - min 30 days and has a 128KB minimum of object size

268

269

Archive to Glacier - min 1 day if IA is not checked, min 60 day if Transition to IA S3 is checked

270

271

Permanently Delete - min 2 day if IA is not checked and 1 is selected for Glacier, min 61 day if IA is selected 30, Glacier is selected 60.

272

273

(2) when versioning is enabled you have lifecycle management options to take action on previous version as well as current version.

274

275

Security and Encryption in S3

276

277

by default newly created buckets are private

278

279

Access control using Bucket Policies (entire bucket) and ACL(individual objects and folders)

280

281

access logs - all the request made to S3 buckets, to another bucket or another account’s S3 bucket

282

283

Encryption

284

285

(1) In Transit - SSL / TLS

286

287

(2) Data at rest

288

289

Server Side Encryption

290

291

SSE- S3 Server Side Encryption with S3 managed keys, (amazon AES 256 handled for you) - click on the object and encrypt

292

293

SSE - KMS - AWS Key management services , managed keys - additional charges / audit trail of keys, amazon manage keys

294

295

SSE - C - Server side encryption with Customer provided keys - you manage encryption keys

296

297

Client Side Encryption

298

299

you encrypt the data on client side and upload to s3

300

Every non-anonymous request to S3 must contain authentication information to establish the identity of the principal making the request. In REST, this is done by first putting the headers in a canonical format, then signing the headers using your AWS Secret Access Key.

301

302

You can insert a presigned url into a webpage to download private data directly from S3.

303

304

The object creation REST APIs (see Specifying Server-Side Encryption Using the REST API) provide a request header, x-amz-server-side-encryption that you can use to request server-side encryption.

305

306

S3 Transfer Acceleration

307

308

Utilize local edge locations to upload content to S3 - incur extra cost

309

310

further away you are the more benefit you get (faster)

311

312

GateWay

313

314

(1) Gateway stored volumes - entire dataset is stored onsite and asynchronously backed up to S3

315

316

(2) Gateway cached volumes - Most frequently used data is stored onsite and entire dataset is stored on S3

317

318

(3) Gateway Virtual Taped library - Used for backup if you don’t want to use Tapes, like Netbackups etc..

319

320

Import Export

321

322

Import / Export Disk

323

324

Import to S3, EBS, Glacier

325

326

export from S3

327

328

Import / Export Snowball

329

330

Import to S3

331

332

Export to S3

333

334

S3 stored data in alphabetical / lexigraphical order. so if you want to spread the load across S3, filename should not be similar (Optimize performance)

335

336

CloudFront

337

338

Content Delivery Network - edge locations, reduced latency, traffic serves from the closest nodes

339

340

Edge locations - content will be cached (over 50), different from region / AZ. TTL (speed of image // media is quicker - first user suffers the performance), can be not only read only (you can write it)

341

342

Origin can be - S3, EC2, ELB, Route53 also NON AWS origin server ,

343

344

Distribution - name given to the CDN consist of collection of Edge locations

345

346

(1) Web Distribution -

347

348

(2) RTMP (media streaming / flash) Distribution - for Adobe flash files only

349

350

you can have multiple origins of a distribution

351

352

Path Pattern (\*)

353

354

Restrict viewer access by signed URL or Signed Cookies

355

356

Restrict content based on geo location (whitelist and blacklist)

357

358

Create invalidate - invalidate TTL (you pay for it) like purge in Akamai

359

360

VPC - logical datacenters in AWS

361

362

Can span multiple AZ, but can’t span multiple regions, PEER VPC, but no Transitive Peering

363

364

Custom VPC has to be /16 can’t go higher then that /8 is not allowed

365

366

When you create Custom VPC it creates default security group, default network ACL and default route table., it doesn’t create default Subnet

367

368

One Subnet == one AZ, you can have security group spanning multiple AZ, ACL’s span across AZ (assign sg and ACL to two different subnets)

369

370

any CIDR block 5 reserved IPs (.0, .1, .2, .3, .255)

371

372

so for CIRD block /24: 2^8 - 5 = 256 - 5 = 251 available IP address space

373

374

when you create internet gateway, by default its detached, attach it to VPC then, only 1 IGW per VPC

375

376

When you create a VPC Default Routetable(Main Routable) is created where the default Routes are,

377

378

10.0.0.0/16 Local <— all subnets inside VPC will be able to talk to each other

379

380

Don’t touch Main route table

381

382

Create another routetable for route out to internet (0.0.0.0/0 IGW) <— route out to the internet

383

384

Last thing you associate this new route table to one of the subnet which will make it public. (you can enable auto assign public IP for the public subnet)

385

386

1 subnet can have 1 routetable

387

388

ICMP is for ping / monitor

389

390

NAT instance and NAT gateway

391

392

NAT Instance - disable source / destination check., always behind security group, must be in public subnet, must have an EIP, ,must be a route out of the private subnet to NAT

393

394

Increase the instance size if bottleneck

395

396

Change the main route table - add a route (0.0.0.0/0 NAT Instance target)

397

398

NAT Instance is a single point of failover (put it behind a ASG),

399

400

NAT gateway - released in 2016 - amazon handled

401

402

Amazon maintains it for you, no need to handle yourself. (security patches applied by AWS)

403

404

You can just create the gateway and assign EIP (put it in public subnet) (automatically assigned)

405

406

Change the main route table - add a route (0.0.0.0/0 NAT gateway target)

407

408

No need for disable source/destination check or no need to put it behind a security group - it handles it for you.

409

410

Highly available / redundancy no need for ASG.NAT gateways are little bit costly - always use it in production scale automatically up to 10Gbps

411

412

ACL vs SG

413

414

Security groups are statefull - any inbound rule , applies to outbound as well (Only Allow rules)

415

416

by default all inbound deny, all outbound allow

417

418

can span across AZ

419

420

ACL are stateless -

421

422

For default ACL, all inbound and outbound rules are allowed by default - associated with all subnets in VPC by default

423

424

for Custom ACL, all inbound and outbound traffic is denied by default - not associated with any subnet

425

426

1 subnet is only associated with ACL. granular rules for ACLs, numbered rules (recommended steps of 100)

427

428

rule no. 99 takes precedence over rule no. 100 (if 99 is blocked and 100 is allowed) 99 will be executed.

429

430

Can SPAN across AZ

431

432

Ephemeral port - 1024 - 65535 should be allowed to take traffic.

433

434

if you want to BLOCK IP address then must use ACL, because security group doesn’t have deny

435

436

Bastion - keep it in public subnet to allow SSH / RDP into instances into private subnets (High availability - Bation in two public subnets and also ASG - Route 53 running Health checks on those Bastion)

437

438

VPC Flowlogs: to capture all the traffic information into logs - logs everything (create IAM role and create cloud watch log group - and log stream)

439

440

VPC Cleanup: can’t delete VPC if you have active running instance or ELB is running

441

442

Thank you so much ACLOUDGURU team, you guys rock, wouldn't have been possible without you.

443