

Exam Questions Professional-Cloud-DevOps-Engineer

Google Cloud Certified - Professional Cloud DevOps Engineer Exam

<https://www.2passeasy.com/dumps/Professional-Cloud-DevOps-Engineer/>



NEW QUESTION 1

Your company follows Site Reliability Engineering principles. You are writing a postmortem for an incident, triggered by a software change, that severely affected users. You want to prevent severe incidents from happening in the future. What should you do?

- A. Identify engineers responsible for the incident and escalate to their senior management.
- B. Ensure that test cases that catch errors of this type are run successfully before new software releases.
- C. Follow up with the employees who reviewed the changes and prescribe practices they should follow in the future.
- D. Design a policy that will require on-call teams to immediately call engineers and management to discuss a plan of action if an incident occurs.

Answer: B

NEW QUESTION 2

You encounter a large number of outages in the production systems you support. You receive alerts for all the outages that wake you up at night. The alerts are due to unhealthy systems that are automatically restarted within a minute. You want to set up a process that would prevent staff burnout while following Site Reliability Engineering practices. What should you do?

- A. Eliminate unactionable alerts.
- B. Create an incident report for each of the alerts.
- C. Distribute the alerts to engineers in different time zones.
- D. Redefine the related Service Level Objective so that the error budget is not exhausted.

Answer: A

Explanation:

Eliminate bad monitoring : Unactionable alerts (i.e., spam) <https://cloud.google.com/blog/products/management-tools/meeting-reliability-challenges-with-sre-principles>

agree with kyubiblaze about having to remove unactionable items aka spam: "good monitoring alerts on actionable problems" @ <https://cloud.google.com/blog/products/management-tools/meeting-reliability-challenges-with-sre-principles>

NEW QUESTION 3

You support a high-traffic web application and want to ensure that the home page loads in a timely manner. As a first step, you decide to implement a Service Level Indicator (SLI) to represent home page request latency with an acceptable page load time set to 100 ms. What is the Google-recommended way of calculating this SLI?

- A. Bucketize the request latencies into ranges, and then compute the percentile at 100 ms.
- B. Bucketize the request latencies into ranges, and then compute the median and 90th percentiles.
- C. Count the number of home page requests that load in under 100 ms, and then divide by the total number of home page requests.
- D. Count the number of home page requests that load in under 100 m
- E. and then divide by the total number of all web application requests.

Answer: C

Explanation:

<https://sre.google/workbook/implementing-slos/>

In the SRE principles book, it's recommended treating the SLI as the ratio of two numbers: the number of good events divided by the total number of events. For example: Number of successful HTTP requests / total HTTP requests (success rate)

NEW QUESTION 4

You are running an application on Compute Engine and collecting logs through Stackdriver. You discover that some personally identifiable information (PII) is leaking into certain log entry fields. All PII entries begin with the text userinfo. You want to capture these log entries in a secure location for later review and prevent them from leaking to Stackdriver Logging. What should you do?

- A. Create a basic log filter matching userinfo, and then configure a log export in the Stackdriver console with Cloud Storage as a sink.
- B. Use a Fluentd filter plugin with the Stackdriver Agent to remove log entries containing userinfo, and then copy the entries to a Cloud Storage bucket.
- C. Create an advanced log filter matching userinfo, configure a log export in the Stackdriver console with Cloud Storage as a sink, and then configure a log exclusion with userinfo as a filter.
- D. Use a Fluentd filter plugin with the Stackdriver Agent to remove log entries containing userinfo, create an advanced log filter matching userinfo, and then configure a log export in the Stackdriver console with Cloud Storage as a sink.

Answer: B

Explanation:

<https://medium.com/google-cloud/fluentd-filter-plugin-for-google-cloud-data-loss-prevention-api-42bbb1308e7>

NEW QUESTION 5

You support an application running on App Engine. The application is used globally and accessed from various device types. You want to know the number of connections. You are using Stackdriver Monitoring for App Engine. What metric should you use?

- A. flex/connections/current
- B. tcp_ssl_proxy/new_connections
- C. tcp_ssl_proxy/open_connections
- D. flex/instance/connections/current

Answer: A

Explanation:

https://cloud.google.com/monitoring/api/metrics_gcp#gcp-appengine

NEW QUESTION 6

You support a popular mobile game application deployed on Google Kubernetes Engine (GKE) across several Google Cloud regions. Each region has multiple Kubernetes clusters. You receive a report that none of the users in a specific region can connect to the application. You want to resolve the incident while following Site Reliability Engineering practices. What should you do first?

- A. Reroute the user traffic from the affected region to other regions that don't report issues.
- B. Use Stackdriver Monitoring to check for a spike in CPU or memory usage for the affected region.
- C. Add an extra node pool that consists of high memory and high CPU machine type instances to the cluster.
- D. Use Stackdriver Logging to filter on the clusters in the affected region, and inspect error messages in the logs.

Answer: A

Explanation:

Google always aims to first stop the impact of an incident, and then find the root cause (unless the root cause just happens to be identified early on).

NEW QUESTION 7

You have a set of applications running on a Google Kubernetes Engine (GKE) cluster, and you are using Stackdriver Kubernetes Engine Monitoring. You are bringing a new containerized application required by your company into production. This application is written by a third party and cannot be modified or reconfigured. The application writes its log information to `/var/log/app_messages.log`, and you want to send these log entries to Stackdriver Logging. What should you do?

- A. Use the default Stackdriver Kubernetes Engine Monitoring agent configuration.
- B. Deploy a Fluentd daemonset to GK
- C. Then create a customized input and output configuration to tail the log file in the application's pods and write to Stackdriver Logging.
- D. Install Kubernetes on Google Compute Engine (GCE) and redeploy your application
- E. Then customize the built-in Stackdriver Logging configuration to tail the log file in the application's pods and write to Stackdriver Logging.
- F. Write a script to tail the log file within the pod and write entries to standard output
- G. Run the script as a sidecar container with the application's pod
- H. Configure a shared volume between the containers to allow the script to have read access to `/var/log` in the application container.

Answer: B

Explanation:

<https://cloud.google.com/architecture/customizing-stackdriver-logs-fluentd>

Besides the list of default logs that the Logging agent streams by default, you can customize the Logging agent to send additional logs to Logging or to adjust agent settings by adding input configurations. The configuration definitions in these sections apply to the fluent-plugin-google-cloud output plugin only and specify how logs are transformed and ingested into Cloud Logging. <https://cloud.google.com/logging/docs/agent/logging/configuration#configure>

NEW QUESTION 8

Your application runs on Google Cloud Platform (GCP). You need to implement Jenkins for deploying application releases to GCP. You want to streamline the release process, lower operational toil, and keep user data secure. What should you do?

- A. Implement Jenkins on local workstations.
- B. Implement Jenkins on Kubernetes on-premises
- C. Implement Jenkins on Google Cloud Functions.
- D. Implement Jenkins on Compute Engine virtual machines.

Answer: D

Explanation:

Your application runs on Google Cloud Platform (GCP). You need to implement Jenkins for deploying application releases to GCP. You want to streamline the release process, lower operational toil, and keep user data secure. What should you do?

<https://plugins.jenkins.io/google-compute-engine/>

NEW QUESTION 9

You support a high-traffic web application with a microservice architecture. The home page of the application displays multiple widgets containing content such as the current weather, stock prices, and news headlines. The main serving thread makes a call to a dedicated microservice for each widget and then lays out the homepage for the user. The microservices occasionally fail; when that happens, the serving thread serves the homepage with some missing content. Users of the application are unhappy if this degraded mode occurs too frequently, but they would rather have some content served instead of no content at all. You want to set a Service Level Objective (SLO) to ensure that the user experience does not degrade too much. What Service Level Indicator (SLI) should you use to measure this?

- A. A quality SLI: the ratio of non-degraded responses to total responses
- B. An availability SLI: the ratio of healthy microservices to the total number of microservices
- C. A freshness SLI: the proportion of widgets that have been updated within the last 10 minutes
- D. A latency SLI: the ratio of microservice calls that complete in under 100 ms to the total number of microservice calls

Answer: B

Explanation:

<https://cloud.google.com/blog/products/gcp/available-or-not-that-is-the-question-cre-life-lessons>

NEW QUESTION 10

You are using Stackdriver to monitor applications hosted on Google Cloud Platform (GCP). You recently deployed a new application, but its logs are not appearing on the Stackdriver dashboard.

You need to troubleshoot the issue. What should you do?

- A. Confirm that the Stackdriver agent has been installed in the hosting virtual machine.
- B. Confirm that your account has the proper permissions to use the Stackdriver dashboard.
- C. Confirm that port 25 has been opened in the firewall to allow messages through to Stackdriver.
- D. Confirm that the application is using the required client library and the service account key has proper permissions.

Answer: A

Explanation:

<https://cloud.google.com/monitoring/agent/monitoring/troubleshooting#checklist>

NEW QUESTION 10

You support an application deployed on Compute Engine. The application connects to a Cloud SQL instance to store and retrieve data. After an update to the application, users report errors showing database timeout messages. The number of concurrent active users remained stable. You need to find the most probable cause of the database timeout. What should you do?

- A. Check the serial port logs of the Compute Engine instance.
- B. Use Stackdriver Profiler to visualize the resources utilization throughout the application.
- C. Determine whether there is an increased number of connections to the Cloud SQL instance.
- D. Use Cloud Security Scanner to see whether your Cloud SQL is under a Distributed Denial of Service (DDoS) attack.

Answer: B

NEW QUESTION 12

You need to define Service Level Objectives (SLOs) for a high-traffic multi-region web application. Customers expect the application to always be available and have fast response times. Customers are currently happy with the application performance and availability. Based on current measurement, you observe that the 90th percentile of latency is 120ms and the 95th percentile of latency is 275ms over a 28-day window. What latency SLO would you recommend to the team to publish?

- A. 90th percentile – 100ms 95th percentile – 250ms
- B. 90th percentile – 120ms 95th percentile – 275ms
- C. 90th percentile – 150ms 95th percentile – 300ms
- D. 90th percentile – 250ms 95th percentile – 400ms

Answer: C

Explanation:

<https://sre.google/sre-book/service-level-objectives/>

NEW QUESTION 14

You are working with a government agency that requires you to archive application logs for seven years. You need to configure Stackdriver to export and store the logs while minimizing costs of storage. What should you do?

- A. Create a Cloud Storage bucket and develop your application to send logs directly to the bucket.
- B. Develop an App Engine application that pulls the logs from Stackdriver and saves them in BigQuery.
- C. Create an export in Stackdriver and configure Cloud Pub/Sub to store logs in permanent storage for seven years.
- D. Create a sink in Stackdriver, name it, create a bucket on Cloud Storage for storing archived logs, and then select the bucket as the log export destination.

Answer: D

Explanation:

<https://cloud.google.com/logging/docs/routing/overview>

NEW QUESTION 16

You support a high-traffic web application that runs on Google Cloud Platform (GCP). You need to measure application reliability from a user perspective without making any engineering changes to it. What should you do?

Choose 2 answers

- A. Review current application metrics and add new ones as needed.
- B. Modify the code to capture additional information for user interaction.
- C. Analyze the web proxy logs only and capture response time of each request.
- D. Create new synthetic clients to simulate a user journey using the application.
- E. Use current and historic Request Logs to trace customer interaction with the application.

Answer: CE

Explanation:

<https://cloud.google.com/architecture/adopting-slos?hl=en>

NEW QUESTION 17

You support a service with a well-defined Service Level Objective (SLO). Over the previous 6 months, your service has consistently met its SLO and customer satisfaction has been consistently high. Most of your service's operations tasks are automated and few repetitive tasks occur frequently. You want to optimize the balance between reliability and deployment velocity while following site reliability engineering best practices. What should you do? (Choose two.)

- A. Make the service's SLO more strict.
- B. Increase the service's deployment velocity and/or risk.
- C. Shift engineering time to other services that need more reliability.
- D. Get the product team to prioritize reliability work over new features.

E. Change the implementation of your Service Level Indicators (SLIs) to increase coverage.

Answer: BC

Explanation:

(<https://sre.google/workbook/implementing-slos/#slo-decision-matrix>)

NEW QUESTION 19

You are on-call for an infrastructure service that has a large number of dependent systems. You receive an alert indicating that the service is failing to serve most of its requests and all of its dependent systems with hundreds of thousands of users are affected. As part of your Site Reliability Engineering (SRE) incident management protocol, you declare yourself Incident Commander (IC) and pull in two experienced people from your team as Operations Lead (OLJ) and Communications Lead (CL). What should you do next?

- A. Look for ways to mitigate user impact and deploy the mitigations to production.
- B. Contact the affected service owners and update them on the status of the incident.
- C. Establish a communication channel where incident responders and leads can communicate with each other.
- D. Start a postmortem, add incident information, circulate the draft internally, and ask internal stakeholders for input.

Answer: A

Explanation:

<https://sre.google/sre-book/managing-incidents/>

NEW QUESTION 24

Your application services run in Google Kubernetes Engine (GKE). You want to make sure that only images from your centrally-managed Google Container Registry (GCR) image registry in the altostrat-images project can be deployed to the cluster while minimizing development time. What should you do?

- A. Create a custom builder for Cloud Build that will only push images to gcr.io/altostrat-images.
- B. Use a Binary Authorization policy that includes the whitelist name pattern gcr.io/altostrat-images/.
- C. Add logic to the deployment pipeline to check that all manifests contain only images from gcr.io/altostrat-images.
- D. Add a tag to each image in gcr.io/altostrat-images and check that this tag is present when the image is deployed.

Answer: B

NEW QUESTION 27

Your team uses Cloud Build for all CI/CO pipelines. You want to use the kubectl builder for Cloud Build to deploy new images to Google Kubernetes Engine (GKE). You need to authenticate to GKE while minimizing development effort. What should you do?

- A. Assign the Container Developer role to the Cloud Build service account.
- B. Specify the Container Developer role for Cloud Build in the cloudbuild.yaml file.
- C. Create a new service account with the Container Developer role and use it to run Cloud Build.
- D. Create a separate step in Cloud Build to retrieve service account credentials and pass these to kubectl.

Answer: A

Explanation:

<https://cloud.google.com/build/docs/deploying-builds/deploy-gke> <https://cloud.google.com/build/docs/securing-builds/configure-user-specified-service-accounts>

NEW QUESTION 31

Your company experiences bugs, outages, and slowness in its production systems. Developers use the production environment for new feature development and bug fixes. Configuration and experiments are done in the production environment, causing outages for users. Testers use the production environment for load testing, which often slows the production systems. You need to redesign the environment to reduce the number of bugs and outages in production and to enable testers to load test new features. What should you do?

- A. Create an automated testing script in production to detect failures as soon as they occur.
- B. Create a development environment with smaller server capacity and give access only to developers and testers.
- C. Secure the production environment to ensure that developers can't change it and set up one controlled update per year.
- D. Create a development environment for writing code and a test environment for configurations, experiments, and load testing.

Answer: D

NEW QUESTION 35

You are deploying an application that needs to access sensitive information. You need to ensure that this information is encrypted and the risk of exposure is minimal if a breach occurs. What should you do?

- A. Store the encryption keys in Cloud Key Management Service (KMS) and rotate the keys frequently
- B. Inject the secret at the time of instance creation via an encrypted configuration management system.
- C. Integrate the application with a Single sign-on (SSO) system and do not expose secrets to the application
- D. Leverage a continuous build pipeline that produces multiple versions of the secret for each instance of the application.

Answer: A

Explanation:

<https://cloud.google.com/security-key-management>

NEW QUESTION 37

You are creating and assigning action items in a postmodern for an outage. The outage is over, but you need to address the root causes. You want to ensure that your team handles the action items quickly and efficiently. How should you assign owners and collaborators to action items?

- A. Assign one owner for each action item and any necessary collaborators.
- B. Assign multiple owners for each item to guarantee that the team addresses items quickly
- C. Assign collaborators but no individual owners to the items to keep the postmortem blameless.
- D. Assign the team lead as the owner for all action items because they are in charge of the SRE team.

Answer: A

Explanation:

<https://devops.com/when-it-disaster-strikes-part-3-conducting-a-blameless-post-mortem/>

NEW QUESTION 39

Your product is currently deployed in three Google Cloud Platform (GCP) zones with your users divided between the zones. You can fail over from one zone to another, but it causes a 10-minute service disruption for the affected users. You typically experience a database failure once per quarter and can detect it within five minutes. You are cataloging the reliability risks of a new real-time chat feature for your product. You catalog the following information for each risk:

- Mean Time to Detect (MTTD) in minutes
- Mean Time to Repair (MTTR) in minutes
- Mean Time Between Failure (MTBF) in days
- User Impact Percentage

The chat feature requires a new database system that takes twice as long to successfully fail over between zones. You want to account for the risk of the new database failing in one zone. What would be the values for the risk of database failover with the new system?

- A. MTTD: 5MTTR: 10MTBF: 90Impact: 33%
- B. MTTD:5 MTTR: 20MTBF: 90Impact: 33%
- C. MTTD:5 MTTR: 10MTBF: 90Impact 50%
- D. MTTD:5 MTTR: 20MTBF: 90Impact: 50%

Answer: B

Explanation:

<https://www.atlassian.com/incident-management/kpis/common-metrics> <https://linkedin.github.io/school-of-sre/>

NEW QUESTION 40

Your organization recently adopted a container-based workflow for application development. Your team develops numerous applications that are deployed continuously through an automated build pipeline to the production environment. A recent security audit alerted your team that the code pushed to production could contain vulnerabilities and that the existing tooling around virtual machine (VM) vulnerabilities no longer applies to the containerized environment. You need to ensure the security and patch level of all code running through the pipeline. What should you do?

- A. Set up Container Analysis to scan and report Common Vulnerabilities and Exposures.
- B. Configure the containers in the build pipeline to always update themselves before release.
- C. Reconfigure the existing operating system vulnerability software to exist inside the container.
- D. Implement static code analysis tooling against the Docker files used to create the containers.

Answer: D

Explanation:

<https://cloud.google.com/binary-authorization>

Binary Authorization is a deploy-time security control that ensures only trusted container images are deployed on Google Kubernetes Engine (GKE) or Cloud Run. With Binary Authorization, you can require images to be signed by trusted authorities during the development process and then enforce signature validation when deploying. By enforcing validation, you can gain tighter control over your container environment by ensuring only verified images are integrated into the build-and-release process.

NEW QUESTION 42

You are running an experiment to see whether your users like a new feature of a web application. Shortly after deploying the feature as a canary release, you receive a spike in the number of 500 errors sent to users, and your monitoring reports show increased latency. You want to quickly minimize the negative impact on users.

What should you do first?

- A. Roll back the experimental canary release.
- B. Start monitoring latency, traffic, errors, and saturation.
- C. Record data for the postmortem document of the incident.
- D. Trace the origin of 500 errors and the root cause of increased latency.

Answer: A

NEW QUESTION 43

Your application images are built and pushed to Google Container Registry (GCR). You want to build an automated pipeline that deploys the application when the image is updated while minimizing the development effort. What should you do?

- A. Use Cloud Build to trigger a Spinnaker pipeline.
- B. Use Cloud Pub/Sub to trigger a Spinnaker pipeline.
- C. Use a custom builder in Cloud Build to trigger a Jenkins pipeline.
- D. Use Cloud Pub/Sub to trigger a custom deployment service running in Google Kubernetes Engine(GKE).

Answer: B

Explanation:

<https://cloud.google.com/architecture/continuous-delivery-toolchain-spinnaker-cloud> <https://spinnaker.io/guides/user/pipeline/triggers/pubsub/>

NEW QUESTION 48

Your organization wants to implement Site Reliability Engineering (SRE) culture and principles. Recently, a service that you support had a limited outage. A manager on another team asks you to provide a formal explanation of what happened so they can action remediations. What should you do?

- A. Develop a postmortem that includes the root causes, resolution, lessons learned, and a prioritized list of action item
- B. Share it with the manager only.
- C. Develop a postmortem that includes the root causes, resolution, lessons learned, and a prioritized list of action item
- D. Share it on the engineering organization's document portal.
- E. Develop a postmortem that includes the root causes, resolution, lessons learned, the list of people responsible, and a list of action items for each person
- F. Share it with the manager only.
- G. Develop a postmortem that includes the root causes, resolution, lessons learned, the list of people responsible, and a list of action items for each person
- H. Share it on the engineering organization's document portal.

Answer: B

NEW QUESTION 52

Your team has recently deployed an NGINX-based application into Google Kubernetes Engine (GKE) and has exposed it to the public via an HTTP Google Cloud Load Balancer (GCLB) ingress. You want to scale the deployment of the application's frontend using an appropriate Service Level Indicator (SLI). What should you do?

- A. Configure the horizontal pod autoscaler to use the average response time from the Liveness and Readiness probes.
- B. Configure the vertical pod autoscaler in GKE and enable the cluster autoscaler to scale the cluster as pods expand.
- C. Install the Stackdriver custom metrics adapter and configure a horizontal pod autoscaler to use the number of requests provided by the GCLB.
- D. Expose the NGINX stats endpoint and configure the horizontal pod autoscaler to use the request metrics exposed by the NGINX deployment.

Answer: C

Explanation:

<https://cloud.google.com/kubernetes-engine/docs/tutorials/autoscaling-metrics>

NEW QUESTION 55

You deploy a new release of an internal application during a weekend maintenance window when there is minimal user traffic. After the window ends, you learn that one of the new features isn't working as expected in the production environment. After an extended outage, you roll back the new release and deploy a fix. You want to modify your release process to reduce the mean time to recovery so you can avoid extended outages in the future. What should you do? Choose 2 answers

- A. Before merging new code, require 2 different peers to review the code changes.
- B. Adopt the blue/green deployment strategy when releasing new code via a CD server.
- C. Integrate a code linting tool to validate coding standards before any code is accepted into the repository.
- D. Require developers to run automated integration tests on their local development environments before release.
- E. Configure a CI server
- F. Add a suite of unit tests to your code and have your CI server run them on commit and verify any changes.

Answer: BE

NEW QUESTION 60

You support a multi-region web service running on Google Kubernetes Engine (GKE) behind a Global HTTP'S Cloud Load Balancer (CLB). For legacy reasons, user requests first go through a third-party Content Delivery Network (CDN), which then routes traffic to the CLB. You have already implemented an availability Service Level Indicator (SLI) at the CLB level. However, you want to increase coverage in case of a potential load balancer misconfiguration, CDN failure, or other global networking catastrophe. Where should you measure this new SLI? Choose 2 answers

- A. Your application servers' logs
- B. Instrumentation coded directly in the client
- C. Metrics exported from the application servers
- D. GKE health checks for your application servers
- E. A synthetic client that periodically sends simulated user requests

Answer: BE

NEW QUESTION 61

You need to deploy a new service to production. The service needs to automatically scale using a Managed Instance Group (MIG) and should be deployed over multiple regions. The service needs a large number of resources for each instance and you need to plan for capacity. What should you do?

- A. Use the n1-highcpu-96 machine type in the configuration of the MIG.
- B. Monitor results of Stackdriver Trace to determine the required amount of resources.
- C. Validate that the resource requirements are within the available quota limits of each region.
- D. Deploy the service in one region and use a global load balancer to route traffic to this region.

Answer: C

Explanation:

https://cloud.google.com/compute/quotas#understanding_quotas <https://cloud.google.com/compute/quotas>

NEW QUESTION 62

You support a service that recently had an outage. The outage was caused by a new release that exhausted the service memory resources. You rolled back the release successfully to mitigate the impact on users. You are now in charge of the post-mortem for the outage. You want to follow Site Reliability Engineering practices when developing the post-mortem. What should you do?

- A. Focus on developing new features rather than avoiding the outages from recurring.
- B. Focus on identifying the contributing causes of the incident rather than the individual responsible for the cause.
- C. Plan individual meetings with all the engineers involved
- D. Determine who approved and pushed the new release to production.
- E. Use the Git history to find the related code commit
- F. Prevent the engineer who made that commit from working on production services.

Answer: B

NEW QUESTION 65

You support a user-facing web application. When analyzing the application's error budget over the previous six months, you notice that the application has never consumed more than 5% of its error budget in any given time window. You hold a Service Level Objective (SLO) review with business stakeholders and confirm that the SLO is set appropriately. You want your application's SLO to more closely reflect its observed reliability. What steps can you take to further that goal while balancing velocity, reliability, and business needs? (Choose two.)

- A. Add more serving capacity to all of your application's zones.
- B. Have more frequent or potentially risky application releases.
- C. Tighten the SLO to match the application's observed reliability.
- D. Implement and measure additional Service Level Indicators (SLIs) from the application.
- E. Announce planned downtime to consume more error budget, and ensure that users are not depending on a tighter SLO.

Answer: DE

Explanation:

<https://sre.google/sre-book/service-level-objectives/>

You want the application's SLO to more closely reflect its observed reliability. The key here is error budget never goes over 5%. This means they can have additional downtime and still stay within their budget.

NEW QUESTION 68

You are performing a semiannual capacity planning exercise for your flagship service. You expect a service user growth rate of 10% month-over-month over the next six months. Your service is fully containerized and runs on Google Cloud Platform (GCP), using a Google Kubernetes Engine (GKE) Standard regional cluster on three zones with cluster autoscaler enabled. You currently consume about 30% of your total deployed CPU capacity, and you require resilience against the failure of a zone. You want to ensure that your users experience minimal negative impact as a result of this growth or as a result of zone failure, while avoiding unnecessary costs. How should you prepare to handle the predicted growth?

- A. Verify the maximum node pool size, enable a horizontal pod autoscaler, and then perform a load test to verify your expected resource needs.
- B. Because you are deployed on GKE and are using a cluster autoscaler
- C. your GKE cluster will scale automatically, regardless of growth rate.
- D. Because you are at only 30% utilization, you have significant headroom and you won't need to add any additional capacity for this rate of growth.
- E. Proactively add 60% more node capacity to account for six months of 10% growth rate, and then perform a load test to make sure you have enough capacity.

Answer: A

Explanation:

<https://cloud.google.com/kubernetes-engine/docs/concepts/horizontalpodautoscaler>

The Horizontal Pod Autoscaler changes the shape of your Kubernetes workload by automatically increasing or decreasing the number of Pods in response to the workload's CPU or memory consumption

NEW QUESTION 70

You support a web application that runs on App Engine and uses CloudSQL and Cloud Storage for data storage. After a short spike in website traffic, you notice a big increase in latency for all user requests, increase in CPU use, and the number of processes running the application. Initial troubleshooting reveals: After the initial spike in traffic, load levels returned to normal but users still experience high latency. Requests for content from the CloudSQL database and images from Cloud Storage show the same high latency.

No changes were made to the website around the time the latency increased. There is no increase in the number of errors to the users.

You expect another spike in website traffic in the coming days and want to make sure users don't experience latency. What should you do?

- A. Upgrade the GCS buckets to Multi-Regional.
- B. Enable high availability on the CloudSQL instances.
- C. Move the application from App Engine to Compute Engine.
- D. Modify the App Engine configuration to have additional idle instances.

Answer: D

Explanation:

Scaling App Engine scales the number of instances automatically in response to processing volume. This scaling factors in the `automatic_scaling` settings that are provided on a per-version basis in the configuration file. A service with basic scaling is configured by setting the maximum number of instances in the `max_instances` parameter of the `basic_scaling` setting. The number of live instances scales with the processing volume. You configure the number of instances of each version in that service's configuration file. The number of instances usually corresponds to the size of a dataset being held in memory or the desired throughput for offline work. You can adjust the number of instances of a manually-scaled version very quickly, without stopping instances that are currently running, using the Modules API `set_num_instances` function. <https://cloud.google.com/appengine/docs/standard/python/how-instances-are-managed>

<https://cloud.google.com/appengine/docs/standard/python/config/appref>

`max_idle_instances` Optional. The maximum number of idle instances that App Engine should maintain for this version. Specify a value from 1 to 1000. If not specified, the default value is automatic, which means App Engine will manage the number of idle instances. Keep the following in mind: A high maximum reduces the number of idle instances more gradually when load levels return to normal after a spike. This helps your application maintain steady performance through fluctuations in request load, but also raises the number of idle instances (and consequent running costs) during such periods of heavy load.

NEW QUESTION 71

.....

THANKS FOR TRYING THE DEMO OF OUR PRODUCT

Visit Our Site to Purchase the Full Set of Actual Professional-Cloud-DevOps-Engineer Exam Questions With Answers.

We Also Provide Practice Exam Software That Simulates Real Exam Environment And Has Many Self-Assessment Features. Order the Professional-Cloud-DevOps-Engineer Product From:

<https://www.2passeasy.com/dumps/Professional-Cloud-DevOps-Engineer/>

Money Back Guarantee

Professional-Cloud-DevOps-Engineer Practice Exam Features:

- * Professional-Cloud-DevOps-Engineer Questions and Answers Updated Frequently
- * Professional-Cloud-DevOps-Engineer Practice Questions Verified by Expert Senior Certified Staff
- * Professional-Cloud-DevOps-Engineer Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- * Professional-Cloud-DevOps-Engineer Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year