← Google Discussions

Exam Professional Data Engineer All Questions
View all questions & answers for the Professional Data Engineer exam

Go to Exam

📄 **EXAM PROFESSIONAL DATA ENGINEER TOPIC 1 QUESTION 13 DISCUSSION**

Actual exam question from Google's Professional Data Engineer

Question #: 13

Topic #: 1

[All Professional Data Engineer Questions]

You want to process payment transactions in a point-of-sale application that will run on Google Cloud Platform. Your user base could grow exponentially, but you do not want to manage infrastructure scaling.

Which Google database service should you use?

    A. Cloud SQL

    B. BigQuery

    C. Cloud Bigtable

    D. Cloud Datastore

Show Suggested Answer

by 👤 **jvg637** at *March 11, 2020, 6:22 p.m.*

Comments

Type your comment...

Submit

☐ 👤 **DeepakKhattar** Highly Voted 👍 4 years, 11 months ago
Initially, thinking D is the best answer but when question is re-re-read, A seems to be correct answer for following reasons
1. Is payment TRANSACTION -- DB should able to perform full blown transaction (updating inventory, sales info etc. though

1. Its payment-TRANSACTION — DB should able to perform full blown transaction (updating inventory, sales info etc, though not specified) , not just ATOMIC which DataStore provides
2. Its point-of-sale application, not ONLINE STORE where HIGH number of concurrent users ordering stuff.
3. User Base could grow exponentially - again more users does mot mean concurrent users and more processing power. Its only about storage.
4. Do not want to Manage infrastructure scaling. - Cloud SQL can scale in terms of storage.
5. CloudStore is poor selection for OLTP application
- Each property is index - so higher latency

Not sure, during exam 2 min is enough to think on various point..
I may be wrong or wrong path ... lets brainstrom..

👍 ↩ 🚩 upvoted 78 times

👤 **canon123** 3 years, 6 months ago

CloudSql does not auto scale.

👍 ↩ 🚩 upvoted 10 times

> 👤 **BigQuery** 3 years, 5 months ago
>
> https://cloud.google.com/architecture/elastically-scaling-your-mysql-environment#objectives
>
> Please read. It can be configured for autoscaling.
>
> 👍 ↩ 🚩 upvoted 3 times
>
> > 👤 **hendrixlives** 3 years, 4 months ago
> >
> > That link explains how to set MySQL autoscaling with Google Compute Engine instances (you install and manage MySQL on the VM). This can not be applied to Cloud SQL (managed service). In Cloud SQL, only the storage can be automatically increased, and changing the Cloud SQL instance size requires a manual edit of the instance type.
> >
> > 👍 ↩ 🚩 upvoted 4 times
> >
> > > 👤 **MisuLava** 2 years, 6 months ago
> > >
> > > yes. and that is ok since this is a point of sale. an exponential increase in number of clients still means reduced parallel processing (how many customers can buy in the very same time) so an increase in memory and CPU is very unlikely to be necessary. yes, an exponential increase in the number of customers means more memory, and more storage, which in cloud SQL increases automatically.
> > >
> > > 👍 ↩ 🚩 upvoted 3 times
> >
> > > 👤 **nkunwar** 2 years, 7 months ago
> > >
> > > C SQL doesn't AUTO SCALE, you need to manually edit , Please show where does it says AUTO SCALING
> > >
> > > 👍 ↩ 🚩 upvoted 2 times

👤 **Blobby** 3 years, 8 months ago

Can't online be considered PoS? CloudSQL does have constraints for scaling and Google seem to specifically be selling Datastore for transactional use cases so going with D:
https://cloud.google.com/datastore/docs/concepts/transactions

👍 ↩ 🚩 upvoted 3 times

> 👤 **Blobby** 3 years, 7 months ago
>
> Based on a re-read of the above comments and other later questions agree with A.
> pls ignore my first answer.
>
> 👍 ↩ 🚩 upvoted 5 times

👤 **jvg637** `Highly Voted 👍` 5 years, 1 month ago

D seems to be the right one. Cloud SQL doesn't automatically scale

👍 ↩ 🚩 upvoted 38 times

> 👤 **BigQuery** 3 years, 5 months ago
>
> Cloud SQL does scale automatically. THERE IS A SETTING WHERE YOU DEFINE INCREASE MEMORY SPACE WHEN IT REACHED 70%.
>
> https://cloud.google.com/sql/docs/features#features_3
>
> Here it say's
> -> Fully managed SQL Server databases in the cloud.
> -> Custom machine types with up to 624 GB of RAM and 96 CPUs.
> -> Up to 64 TB of storage available, with the ability to automatically increase storage size as needed.
>
> 👍 ↩ 🚩 upvoted 2 times
>
> > 👤 **hendrixlives** 3 years, 4 months ago
> >
> > Storage scale is automatic (e.g. you begin with a 50GB disk and it grows automatically as needed), but the instance size (CPU/memory) will be the same. The questions states that the user base may increase exponentially. Even if you

have enough disk space to store all your user data, the increase in users will cause problems if your instance (CPU/memory) is too small, since the instance will not be able to process all the queries at the required speed.

👍 ↩ 🚩 upvoted 5 times

　　⊟ 👤 **imsaikat50** 2 years, 4 months ago

　　I believe the key point is it's a POS, not an e-commerce. Keeping that in mind, exponential user increase in POS might not mean concurrent user increase, which could be a huge consideration in case of it is being e-commerce.

　　I would rather go with 'Cloud SQL' as the best answer.

　　👍 ↩ 🚩 upvoted 2 times

⊟ 👤 **fassil** `Most Recent ⊘` 3 weeks, 4 days ago

`Selected Answer: D`

You don't need to go far, guys. Cloud SQL does not support auto scale; cut the shit here based on the requirement—the question specifically says "you do not want to manage infrastructure scaling."

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **Parandhaman_Margan** 1 month, 3 weeks ago

`Selected Answer: D`

POS app needing scalable DB without managing infrastructure. Cloud Datastore or Firestore? But options are Cloud SQL, BigQuery, Bigtable, Datastore. Since it's for transactions and scale, maybe Cloud Datastore (D) or Cloud Bigtable. But Cloud Bigtable is for high throughput. The question says not to manage scaling. Cloud Datastore is serverless. So D.

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **monyu** 1 month, 3 weeks ago

`Selected Answer: D`

B and C are discarded Since we are dealing with transactional data.
A is discarded since we require to deal with Infrastructure scaling in case the base of users increase and also (not necessarily) the concurrent transactions.

Users base is gonna grow and we DO NOT WANT TO DEAL with infrastructure scaling. D is the most appropriated

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **cqrm3n** 3 months, 2 weeks ago

`Selected Answer: D`

The answer is D, Cloud Datastore (now Firestore in Datastore mode), because it supports auto scaling and low latency reads and writes. Cloud SQL is not the correct answer because it requires more active management for scaling.

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **GHill1982** 6 months, 2 weeks ago

`Selected Answer: C`

For handling payment transactions in a point-of-sale application with potential exponential growth and without the need to manage infrastructure scaling, Cloud Bigtable would be the best choice.

👍 ↩ 🚩 upvoted 3 times

⊟ 👤 **Radhika7983** 7 months, 1 week ago

D seems to be the answer. This is what I think based on my analysis below.
POS is OLTP system but now a days NOSQL with ACID properties also are used for OLTP,
Cloud sql is good for relational database and it would have been an option here but it clearly says that "you do not want to manage infrastructure scaling". In cloud SQL, which is managed service and not server less, you need to manually do vertical scaling(scale up and scale down).
Hence I believe CLOUD SQL is not the option here.
I also tried creating a datastore using google cloud console and it gives 2 options now that is cloud firestore in native mode and cloud firestore in data store mode. automatic scaling is available in both where there is no manual scaling up or down is required. Also, both firestore in native and datastore provides ACID properties. Also, firestore is now optimized for OLTP.
Please see below
https://cloud.google.com/solutions/building-scalable-apps-with-cloud-firestore
Though the question only talks about datastore, I am just providing additional information.
Considering all what I read through, D is the answer.

👍 ↩ 🚩 upvoted 13 times

　　⊟ 👤 **awssp12345** 3 years, 10 months ago

　　I agree. I think people are missing the part of the question that mentions they don't want to maintain the DB.

　　👍 ↩ 🚩 upvoted 1 times

　　⊟ 👤 **awssp12345** 3 years, 10 months ago

　　This should be accepted and the highest voted answer.

　　👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **kishanu** 7 months, 1 week ago

D is the hero here.
Though Cloud SQL has an upper hand when it comes to transactions(OLTP), it does not autoscale its computing capabilities as compared to datastore.
Do visit: https://cloud.google.com/datastore/docs/concepts/overview#what_its_good_for

👍 ↩ 🏳 **upvoted 2 times**

⊟ 👤 **hendrixlives** 7 months, 1 week ago

D is correct: Datastore (currently Firestore in native or datastore mode). It is a fully managed and serverless solution that allows for transactions and will autoscale (storage and compute) without the need to manage any infrastructure.
A is wrong: Cloud SQL is fully a managed transactional DB, but only the storage grows automatically. As your user base increases, you will need to increase the CPU/memory of the instance, and to do that you must edit the instance manually (and the questions specifically says "you do not want to manage infrastructure scaling")
B is wrong: Bigquery is OLAP (for analytics). NoOps, fully managed, autoscales and allows transactions, but it is not designed for this use case.
C is wrong: Bigtable is a NoSQL database for massive writes, and to scale (storage and CPU) you must add nodes, so it is completely out of this use case.

👍 ↩ 🏳 **upvoted 4 times**

⊟ 👤 **kuik** 3 years, 4 months ago

May be some history can help to decide which is best answer.
Datastore built by Google uses BigTable as it's storage, while the company who built FireStore uses Cloud Spanner as it's storage. Google decided that they like the FireStore technology and acquired it.
If Cloud Spanner is an option I would choose it. So, D for me, although it's json storage format, but the Cloud Spanner it uses as storage fits all the requirements.

👍 ↩ 🏳 **upvoted 2 times**

⊟ 👤 **axantroff** 7 months, 1 week ago

B - not an option
C - lack of ACID transactions
A - lack of resource automatic scalability
D - (correct, IMHO) support ACID, suitable for OLPT and scalable enough

👍 ↩ 🏳 **upvoted 1 times**

⊟ 👤 **rocky48** 7 months, 1 week ago

B - not an option
C - lack of ACID transactions
A - lack of resource automatic scalability
D - (correct, IMHO) support ACID, suitable for OLPT and scalable enough

👍 ↩ 🏳 **upvoted 3 times**

⊟ 👤 **TVH_Data_Engineer** 7 months, 1 week ago

Cloud Datastore (now part of Google Cloud Firestore in Datastore mode) is designed for high scalability and ease of management for applications. It is a NoSQL document database built for automatic scaling, high performance, and ease of application development. It's serverless, meaning it handles the scaling, performance, and management automatically, fitting your requirement of not wanting to manage infrastructure scaling.

Cloud SQL, while a fully-managed relational database service that makes it easy to set up, manage, and administer your SQL databases, is not as automatically scalable as Datastore. It's better suited for applications that require a traditional relational database.

👍 ↩ 🏳 **upvoted 1 times**

⊟ 👤 **baimus** 7 months, 2 weeks ago

This actually is A. I was initially in the D camp, and have spent considerable time reading about it now (circa 1 hour). D is explicitly not suitable for payment transactions, as Datastore supports ACID transactions, but only within entity groups, which are small, localized sets of data. This restriction means that transactions are not suitable for scenarios requiring multi-entity consistency across the entire database.
The only two products recommended for payments in the Google ecosystem are Cloud Spanner and Cloud SQL. Is Cloud SQL managed? I'd say it wasn't really, due to the need to configure instances, but that is trumped by the fact it is the only choice suitable for a payment transaction system.

👍 ↩ 🏳 **upvoted 2 times**

⊟ 👤 **SatyamKishore** 8 months, 3 weeks ago

Cloud Datastore is a fully managed, NoSQL document database that is highly scalable and designed to automatically handle

large increases in traffic without requiring manual intervention. It's well-suited for applications with a rapidly growing user base.

👍 ↩ 🚩 upvoted 1 times

☐ 👤 **iooj** 9 months, 1 week ago

Selected Answer: D

Firestore extension of Datastore can handle acid transactions and allows autoscaling

👍 ↩ 🚩 upvoted 1 times

☐ 👤 **jamalkhan** 10 months, 1 week ago

Selected Answer: A

A. Requires Transactions.

👍 ↩ 🚩 upvoted 1 times

Load full discussion...