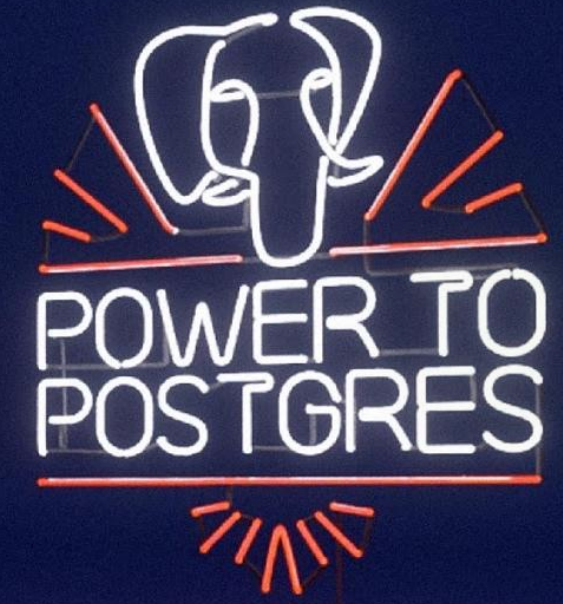
POSTGRES 500



Tom Rieger

Senior Solutions Engineer

EDB

What is ‘Postgres 500’?

Live benchmark so everyone can see ‘behind the curtain’

HammerDB — Transactional Workload - Consistent workload, scale and concurrency

Azure

PostgreSQL Azure

aws

Amazon Aurora PostgreSQL aws

RDS PostgreSQL aws

History of database benchmarking and sizing

- General workloads
  - Transactional
  - Analytical
- Vendor-specific workloads
  - Database-specific
  - Vendor-specific – SAP SD
- More fit-for-purpose
  - Load testing
  - Performance testing
  - IO performance
- Historically benchmarks were focused or
  performance NOT economics

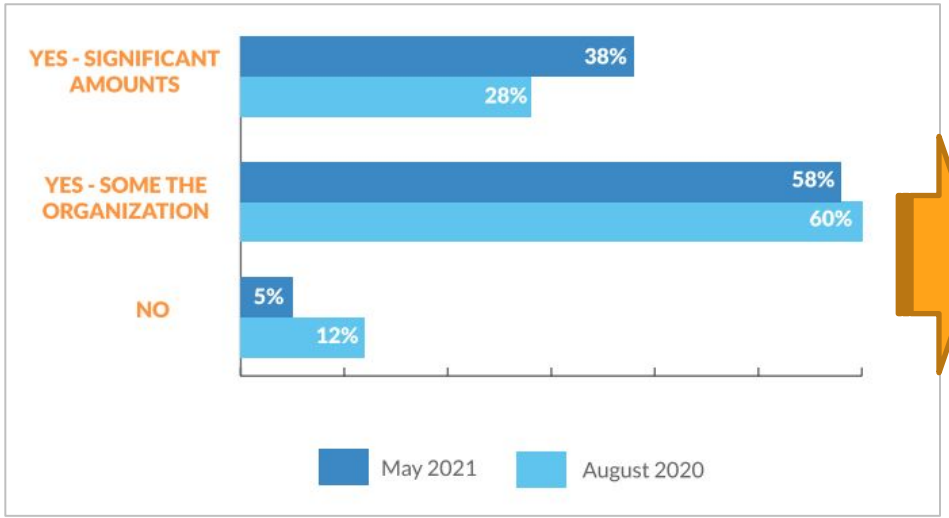
TPC®

$$

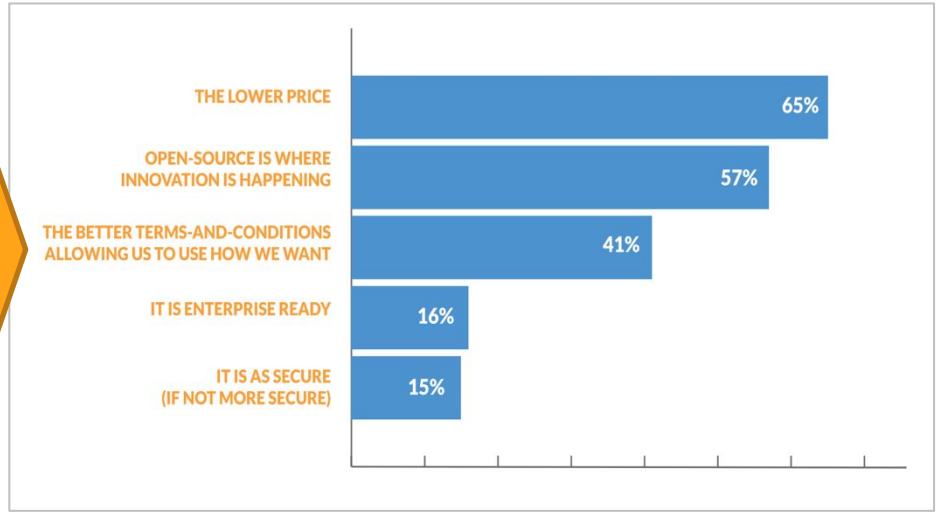
Trends and Joint research

"Driver and Impediments to Corporate Technology Change – 2021 and Beyond

"Do you think there is wasted money in keeping old technology alive in your organization?"



"Why do you think open-source technology is thriving?"

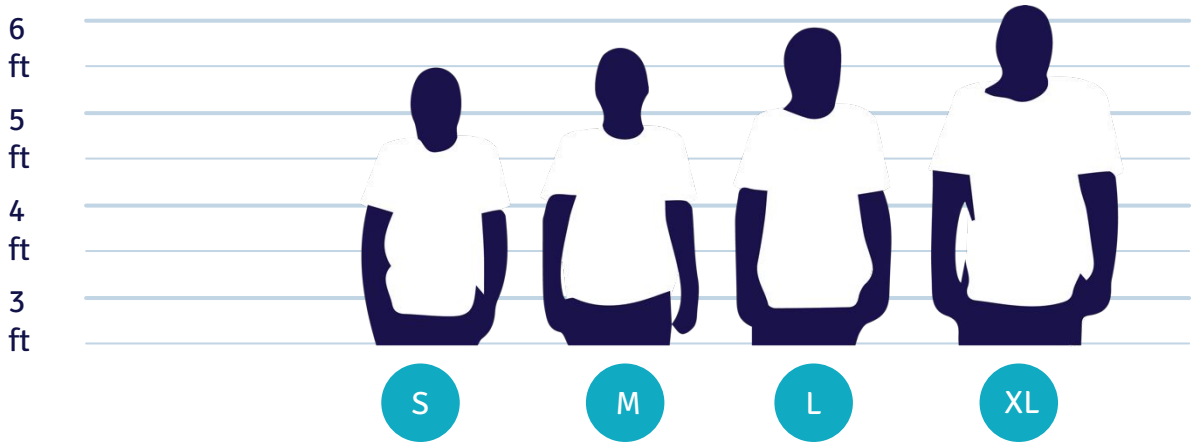


The IT world is no longer a ‘tailored suit’

Up to now:

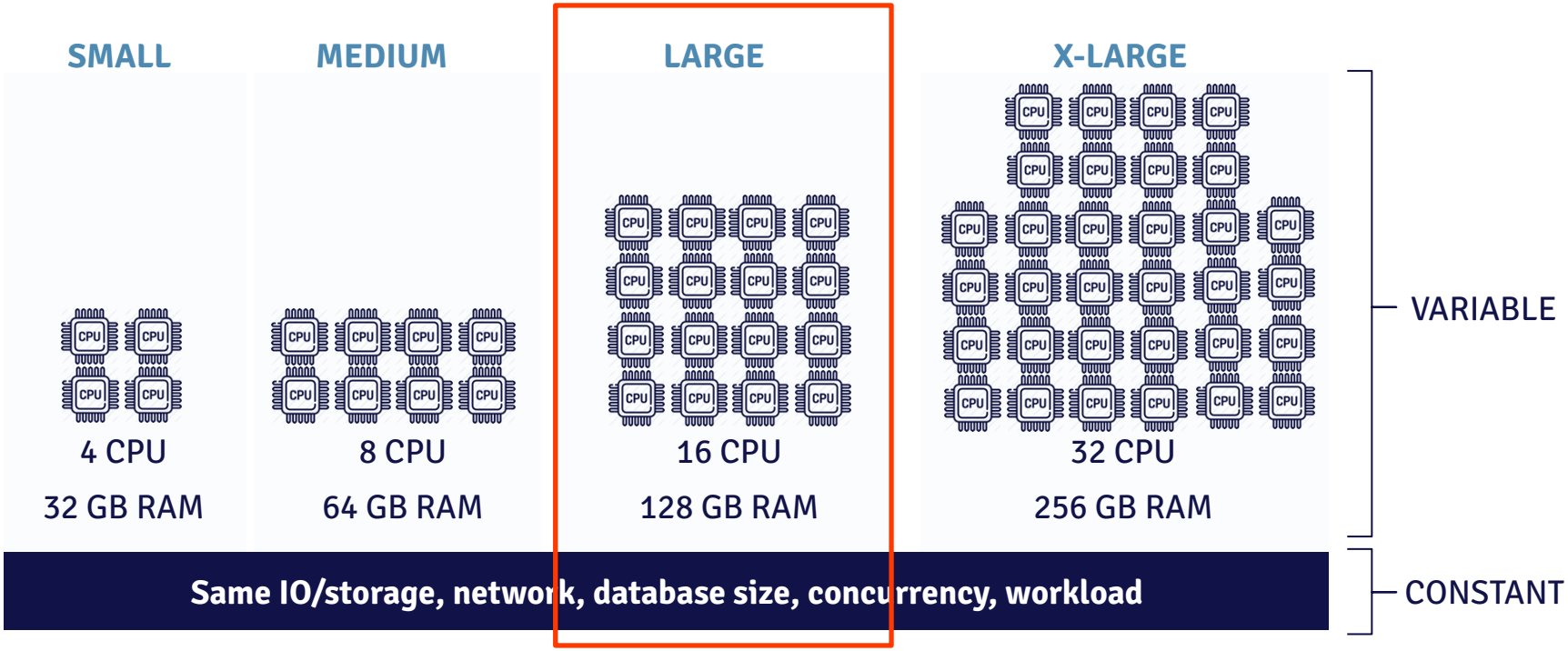
‘Educated guesses’ and ‘release note analysis’ to understand sizing and cost.

The cloud allows a more elastic ‘t-shirt’ sizing



	S	M	L	XL
Name	Hecto	Jo	Dan	Marcus
Height	5, 4	5, 7	5, 10	6, 1
Weight	125	160	170	225
Chest	33 lbs	33 lbs	35 lbs	42 lbs
Waist	32	32	36	38

Think of cloud provisioning as ‘t-shirt’ sizes



LET'S START THE RACE

The effort at hand

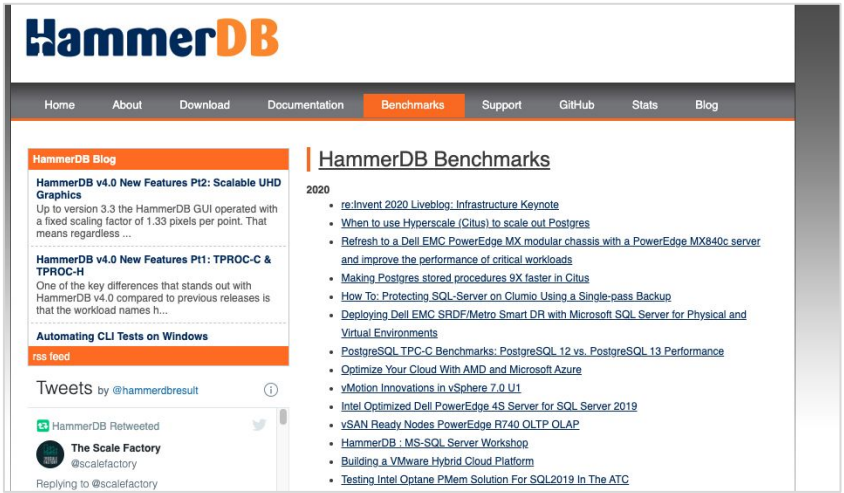
- Evaluate the optimal mix of CPU, RAM, storage and network AND comparable DBaaS
- Use a defensible, standards-based workload that clients can also utilize in their own environment
- Run side-by-side benchmark
- Review the technical and economic result

HammerDB TPC®

Workloads for this effort – Why?

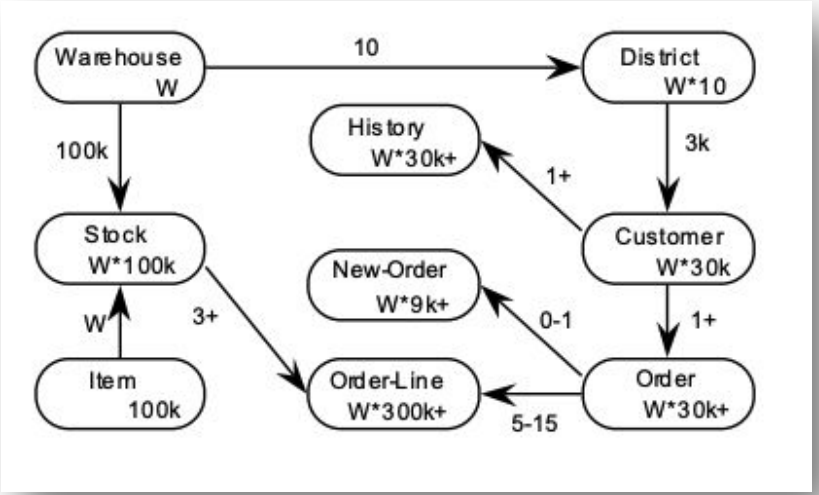
- HammerDB is open-source – evolving, community and free
- Standards based
  - TPC-C for transactional
  - TPC-H for analytical
- Works with all major databases – Oracle, SQLServer, DB2, MySQL, Postgres

https://hammerdb.com/



Quick Anatomy of TPC-C

- Order Processing Use-case
- Typical Transactional use case
  - Simple SQL
  - Large in volume
- Two primary ‘knobs’ to size the effort
  - Warehouses – database size
  - User Concurrency
- **Resulting measure:**
  Transactions-per-minute (TPM)



Action Type	Mix
SELECT	75%
INSERT	8%
UPDATES	16%
DELETE	1%

EDB

amazon web services EC2

Example: Amazon brings a lot of choices

44 different choices with '8 cores'

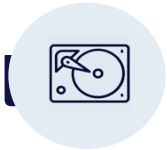
Purpose	instance Name	CPU	RAM	EBS bandwidth	Network bandwidth	Storage	PER MONTH
General Purpose	a1.2xlarge	8	16		Up to 10	EBS Only	$93.81
General Purpose	t4g.2xlarge	8	32	Up to 2,780	Up to 5 GBps	EBS Only	$123.08
General Purpose	c6g.2xlarge	8	16	Up to 4,750	Up to 10	EBS Only	$125.12
General Purpose	t3a.2xlarge	8	32		Up to 5 GBps	EBS Only	$137.68
General Purpose	m6g.2xlarge	8	32	Up to 4,750	Up to 10	EBS Only	$141.04
General Purpose	c6gd.2xlarge	8	16	Up to 4,750	Up to 10	1 x 474 NVMe SSD	$141.25
General Purpose	c5a.2xlarge	8	16	Up to 3,170	Up to 10	EBS Only	$141.62
General Purpose	t3.2xlarge	8	32		Up to 5 GBps	EBS Only	$152.28
General Purpose	c5.2xlarge	8	16	Up to 4,750	Up to 10	EBS Only	$156.22
General Purpose	c6gn.2xlarge	8	16	Up to 9.5	Up to 25 Gbps	EBS Only	$159.21
General Purpose	t2.2xlarge	8	32		Moderate	EBS Only	$167.90
General Purpose	c5d.2xlarge	8	16	Up to 4,750	Up to 10	1 x 200 NVMe SSD	$176.66
General Purpose	m5.2xlarge	8	32	Up to 4,750	up to 10Gbps	EBS Only	$176.66
General Purpose	m4.2xlarge	8	32	1,000	High	EBS Only	$180.89
General Purpose	c4.2xlarge	8	15	1,000	High	EBS Only	$183.96
Memory Optimized	r6g.2xlarge	8	64	Up to 4,750	Up to 10	EBS Only	$185.42
General Purpose	c5n.2xlarge	8	16	Up to 4,750	Up to 25 Gbps	EBS Only	$198.56
Memory Optimized	r5a.2xlarge	8	64	Up to 2,880	Up to 10	EBS Only	$208.05
General Purpose	m5d.2xlarge	8	32	Up to 4,750	Up to 10Gbps	1 x 300 NVMe	$208.05
Memory Optimized	r6gd.2xlarge	8	64	Up to 4,750	Up to 10	1 x 474 NVMe SSD	$211.92
General Purpose	m5n.2xlarge	8	32	Up to 4,750	Up to 25 Gbps	EBS Only	$219.00
Storage Optimized	h1.2xlarge	8	32		Up to 10 Gigabit	1 x 2,000GB HDD	$232.14
Memory Optimized	r5.2xlarge	8	64	Up to 4,750	Up to 10 Gbps	EBS Only	$232.14
Memory Optimized	r5ad.2xlarge	8	64	Up to 2,880	Up to 10	1 x 300 NVMe SSD	$240.90
Memory Optimized	r4.2xlarge	8	61		Up to 10	EBS Only	$245.28
General Purpose	m5dn.2xlarge	8	32	Up to 4,750	Up to 25 Gbps	1 x 300 NVMe	$250.39
Memory Optimized	r5d.2xlarge	8	64	Up to 4,750	Up to 10 Gbps	1 x 300 NVMe	$264.99
Accelerated Computing	inf1.2xlarge	8	16	Up to 4.75 Gbps	Up to 25 Gbps	EBS Only	$268.64
Accelerated Computing	f1.2xlarge	8	122		Up to 10	470 GB	$268.64
Memory Optimized	r5n.2xlarge	8	64	Up to 4,750	Up to 25 Gbps	EBS Only	$273.75
Memory Optimized	r5b.2xlarge	8	64	Up to 10,000	Up to 10	EBS Only	$274.12
Memory Optimized	r5b.2xlarge	8	64	Up to 10,000	Up to 10 Gpbs	EBS Only	$274.12
General Purpose	m5zn.2xlarge	8	32	3,170	Up to 25 Gbps	EBS Only	$303.83
Memory Optimized	r5dn.2xlarge	8	64	Up to 4,750	Up to 25 Gbps	1 x 300 NVMe	$307.33
Storage Optimized	i3.2xlarge	8	61		Up to 10 Gbps	1 x 1,900 NVMe	$312.44
Memory Optimized	z1d.2xlarge	8	64		Up to 10 Gigabit	1 x 300 NVMe SSD	$342.27
Accelerated Computing	g4dn.2xlarge	8	32		Up to 25 Gbps	225 GB	$346.02
Storage Optimized	i3en.2xlarge	8	64		Up to 25 Gbps	2 x 2,500 NVMe SSD	$449.68
Storage Optimized	d3.2xlarge	8	64	1,700	Up to 15	6 x 2TB HDD	$459.17
Storage Optimized	d2.2xlarge	8	61		High	6 x 2000 HDD	$586.92
Memory Optimized	x1e.2xlarge	8	122	Dedicated 500 Mbps	Up to 10 Gbps	EBS + 120GB SSD	$750.44
Accelerated Computing	p3.2xlarge	8	61	14 Gbps	Up to 25 Gbps	EBS Only	$1,524.24
General Purpose	c5ad.2xlarge	8	16	Up to 3,170	Up to 10	1 x 300 NVMe SSD	??
Storage Optimized	d3en.2xlarge	8	32	1,700	Up to 25 Gbps	4 x 14TB HDD	??

93.81/ month

$

$1,524.24/ month

The real abstract topic - STORAGE



Google Cloud

- No performance control
- Quoted for **600GB**

Type	MAX 'sustained' IOPS	MAX 'sustained' Throughput
Standard	*Read: 450 Write: 900*	*Read: 72 MB/sec Write: 72 MB/sec*
Balanced	*Read: 3,600 Write: 3,600*	*Read: 72 MB/sec Write: 72 MB/sec*
SSD	*Read: 15,000 Write: 15,000*	*Read: 240 MB/sec Write: 240 MB/sec*

https://cloud.google.com/compute/docs/disks/?&_ga=2.62767211.-658045011.1598996595#pdperformance

aws

- All about which level and what you are willing to pay – **600GB sizing**

Type	MAX 'sustained' IOPS	MAX Throughput
Optimized HHD (st1)		*MAX 147 MB/s max*
General Purpose SSD (gp2)	*3000*	
General Purpose SSD (gp3)	3000	500 MB/sec
Provisioned IOPS SSD (io1)	MAX 30,000	
Provisioned IOPS SSD (io2)	MAX 64,000	

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-volume-types.html?icmpid=docs_ec2_console

Azure

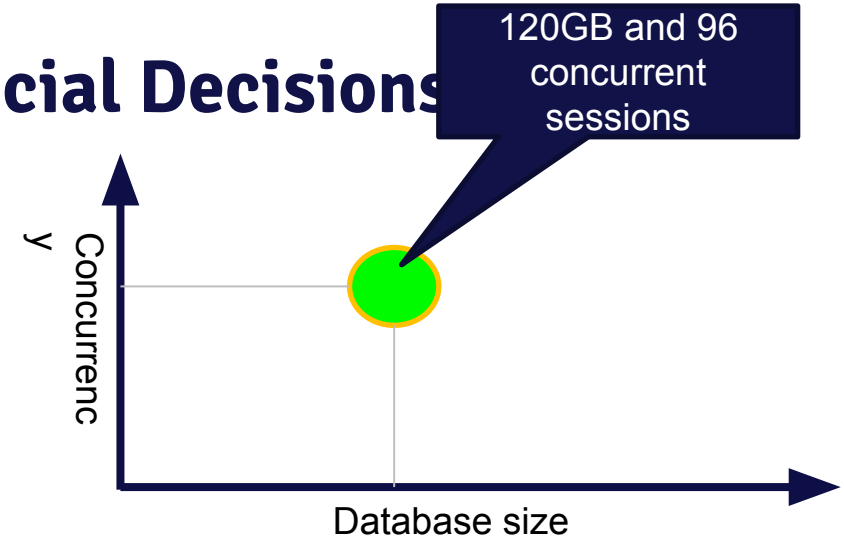
- Azure has 'binary' pricing = 128GB, 256, 512GB, etc.
- **Assume 1024 GB**

Type	MAX 'sustained' IOPS	MAX Throughput
Standard HDD	*500*	*60 MB/sec*
Standard SSD	*500*	*60 MB/sec*
Premium SSD	*5000*	*200 MB/sec*
Ultra Disk	51,200*	768 MB/sec*

https://docs.microsoft.com/en-us/azure/virtual-machines/disks-types

NEXT STEP: Technical/Financial Decisions

- Two Dimensions – database size and concurrent sessions



- What virtual machines to use
- What storage to use and how to tune it (where possible)

Storage in the cloud – and price/performance

Example:

Microsoft Azure

((# of GB * $0.15)+(# of IOPS * $0.06)+(# of MBps * $1.23))*12

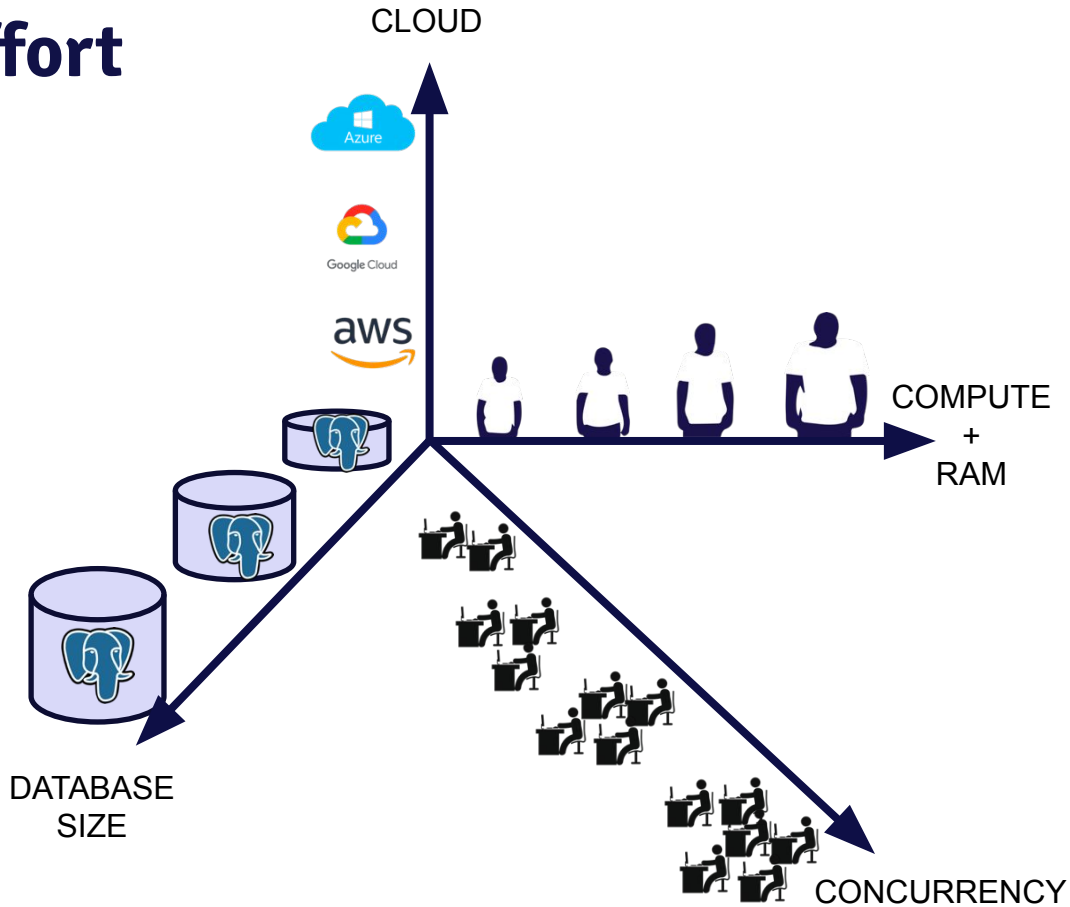
- Top performing storage: **Ultra Disk**
- You pay for it in discrete increments: 256GB, 512GB, 1024GB, etc
  - i.e. Even if you want a volume of 300GB you pay for 512GB
- You configure:
  - IOPS = operations/sec
  - Throughput = MBps
- Each VM has published 'limits'

- Pricing
  - $0.15/mon per GB
  - $0.06/mon per IOPS
  - $1.23/mon per MBps
- For our sizing efforts we used the

Name	Size	"Max uncached disk throughput IOPS/MBps"
E4ds_v4	4 CPU/32 GB	6400/96
E8ds_v4	8 CPU/64 GB	12800/192
E16ds_v4	16 CPU/128 GB	25600/384
E32ds_v4	32 CPU/256 GB	51200/768

The dimensions of the effort

- HammerDB 3.3 running on its own 4-way server
- Ran over 370 separate benchmark runs across AWS, Azure and Google
- Generated over 34TBs of data/ across different sized instances
- Evaluated different storage, Postgres configurations and more



Storage in the cloud – and price/performance

Example:
Microsoft Azure

- MEDIUM T-Shirt
- 120GB database
- 96 concurrent sessions
- 5 minute ‘warm-up’ and 30 minute run
- Started with a fresh database each time
- ONLY VARIABLE CHANGE WAS DISK PERFORMANCE!

IOPS	MBps	Result: TPM	Monthly Storage Cost	$/TPM
35,000	2,000	**303,163**	**$9,378**	$0.42
20,000	1,000	**293,734**	**$5,071**	$0.26
15,000	750	**285,979**	**$3,841**	$0.21
12,800	500	**293,383**	**$2,954**	$0.17
12,800	200	**288,399**	**$2,214**	$0.14
10,000	200	**286,854**	**$1,870**	$0.13
9,000	100	**285,301**	**$1,624**	$0.12
7,000	100	**292,939**	**$1,256**	$0.10
6,000	80	**287,503**	**$1,084**	$0.09

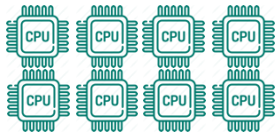
Storage in the cloud – and price/performance

AWS EC2 and
Io2 disk

- MEDIUM T-Shirt
- 120GB database
- 96 concurrent sessions
- 5 minute ‘warm-up’ and 30 minute run
- Started with a fresh database each time
- ONLY VARIABLE CHANGE WAS DISK PERFORMANCE!

IOPS	Result: TPM	Monthly Storage Cost	$/TPM
35,000	**352,863**	**$2,291**	$0.12
30,000	**351,260**	**$2,025**	$0.11
20,000	**357,431**	**$1,375**	$0.09
15,000	**354,238**	**$1,050**	$0.08
10,000	**353,212**	**$725**	$0.07
7,000	**355,717**	**$530**	$0.06
5,000	**353,743**	**$400**	$0.05
4,000	**349,428**	**$335**	$0.05
3,500	**340,057**	**$302**	$0.05
3,000	**327,774**	**$270**	$0.06

AWS RDS – IOPS, tuning and cost



8 CPU

64 GB RAM

IO1 IOPS	Storage Cost	Total Annual Cost	Out of the Box Results (TPM)	$/TPM	EDB Best Efforts Results (TPM)	$/TPM
30,000	$36,900	$45,531	204,145	$0.22	314,926	$0.14
20,000	$24,900	$32,331	209,761	$0.15	325,910	$0.10
15,000	$18,900	$25,731	209,485	$0.12	333,219	$0.08
10,000	$12,900	$19,131	169,537	$0.11	314,832	$0.06
7,000	$9,300	$15,171	126,055	$0.12	316,606	$0.05
5,000	$6,900	$12,531	104,271	$0.12	291,526	$0.04
4,000	$5,700	$11,211	87,239	$0.13	277,107	$0.04

EDB

▼ Show calculations

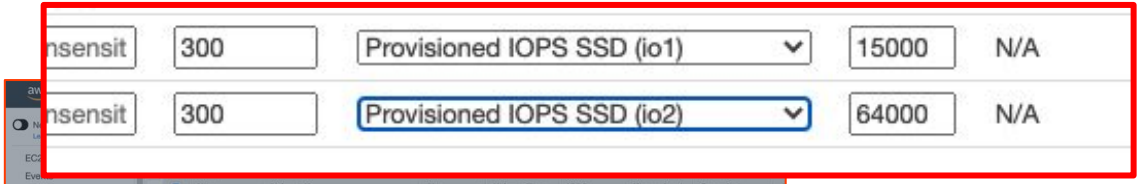
600 GB x 0.10 USD = 60.00 USD (Storage Cost)
10,000 Reads/Second + 10,000 Writes/Second = 20,000 Number of I/Os per second
20,000 I/Os per second x 730 hours x 60 minutes x 60 Seconds = 52,560,000,000 Number of I/Os per month
52,560,000,000 x 0.0000002 USD = 10,512.00 USD (I/O Rate Cost)
60.00 USD + 10,512.00 USD = 10,572.00 USD

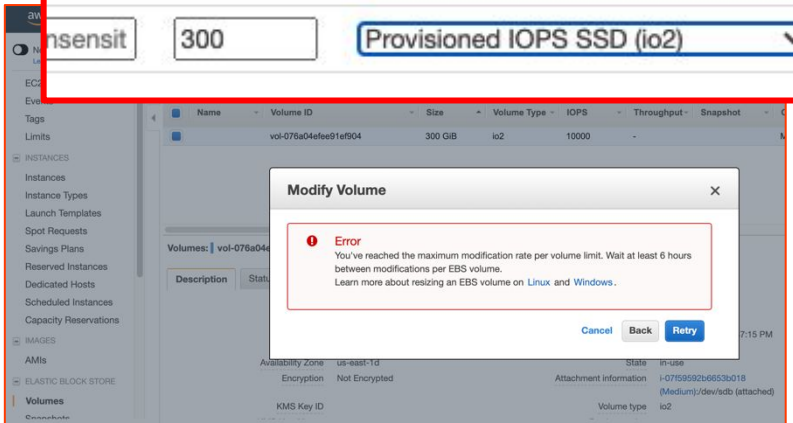
Storage pricing (monthly): 10,572.00 USD

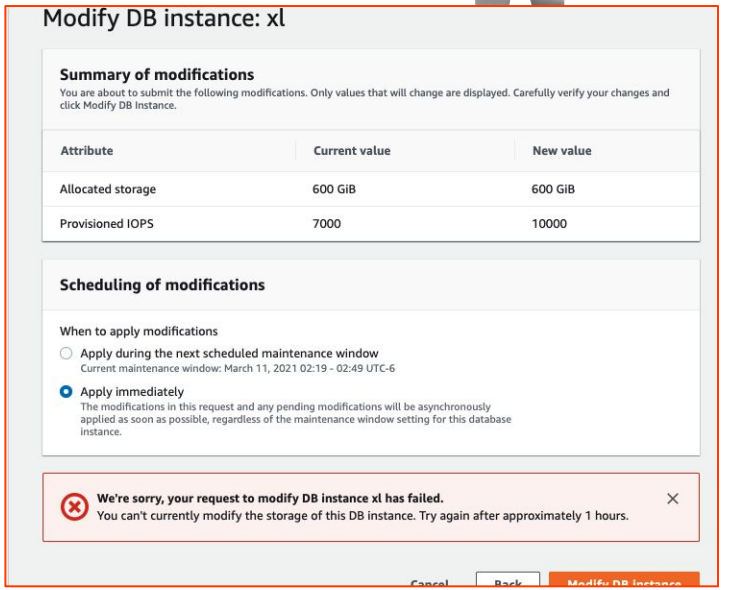
aws

Other Observations - AWS

- RDS - 180 Postgres Conf parameters **cannot** be changed
- Limited in how often you can change the instance
- Death-by-monitoring tools – Cloud Watch, Performance Insights….$$$$



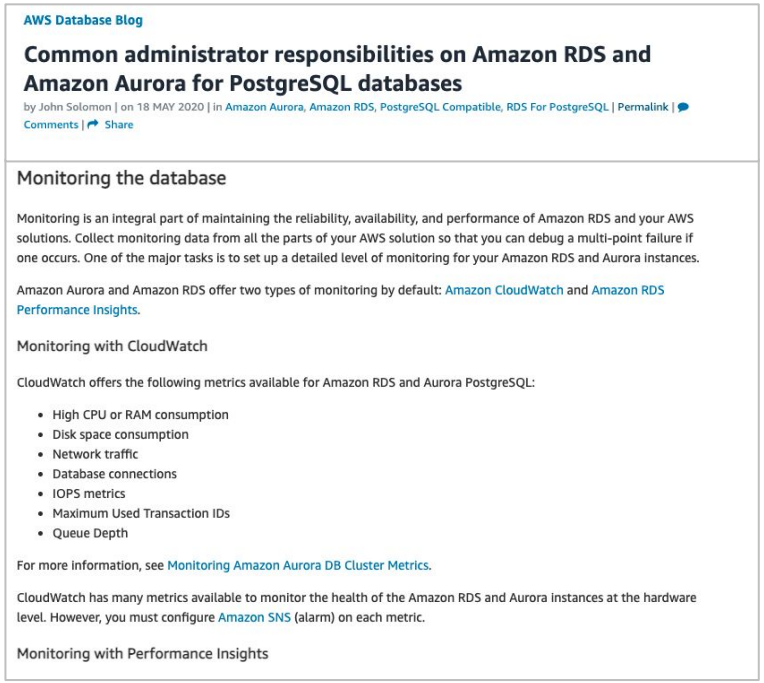




Provisioned IOPS SSD volumes can range in size from 4 GiB to 16 TiB. You can provision from 100 IOPS up to 64,000 IOPS per volume on Instances built on the Nitro System and up to 32,000 on other instances. The maximum ratio of provisioned IOPS to requested volume size (in GiB) is 50:1 for `io1` volumes, and 500:1 for `io2` volumes. For example, a 100 GiB `io1` volume can be provisioned with up to 5,000 IOPS, while a 100 GiB `io2` volume can be provisioned with up to 50,000 IOPS. On a supported instance type, the following volume sizes allow provisioning up to the 64,000 IOPS maximum:

- `io1` volume 1,280 GiB in size or greater (50 × 1,280 GiB = 64,000 IOPS)
- `io2` volume 128 GiB in size or greater (500 × 128 GiB = 64,000 IOPS)

Other interesting observations



AWS Database Blog

Common administrator responsibilities on Amazon RDS and Amazon Aurora for PostgreSQL databases

by John Solomon | on 18 MAY 2020 | in Amazon Aurora, Amazon RDS, PostgreSQL Compatible, RDS For PostgreSQL | Permalink | Comments | Share

Monitoring the database

Monitoring is an integral part of maintaining the reliability, availability, and performance of Amazon RDS and your AWS solutions. Collect monitoring data from all the parts of your AWS solution so that you can debug a multi-point failure if one occurs. One of the major tasks is to set up a detailed level of monitoring for your Amazon RDS and Aurora instances.

Amazon Aurora and Amazon RDS offer two types of monitoring by default: Amazon CloudWatch and Amazon RDS Performance Insights.

Monitoring with CloudWatch

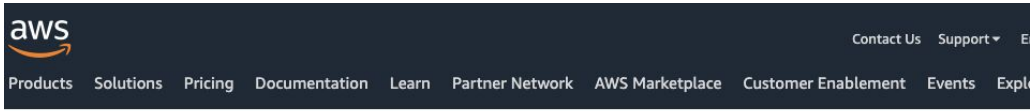
CloudWatch offers the following metrics available for Amazon RDS and Aurora PostgreSQL:

- High CPU or RAM consumption
- Disk space consumption
- Network traffic
- Database connections
- IOPS metrics
- Maximum Used Transaction IDs
- Queue Depth

For more information, see Monitoring Amazon Aurora DB Cluster Metrics.

CloudWatch has many metrics available to monitor the health of the Amazon RDS and Aurora instances at the hardware level. However, you must configure Amazon SNS (alarm) on each metric.

Monitoring with Performance Insights



aws

Contact Us Support

Products Solutions Pricing Documentation Learn Partner Network AWS Marketplace Customer Enablement Events Expl

How can I decrease the total provisioned storage size of my Amazon RDS DB instance?

Last updated: 2020-06-22

I want to decrease the total allocated storage size of my Amazon Relational Database Service (Amazon RDS) DB instance. How can I do this?

Short description

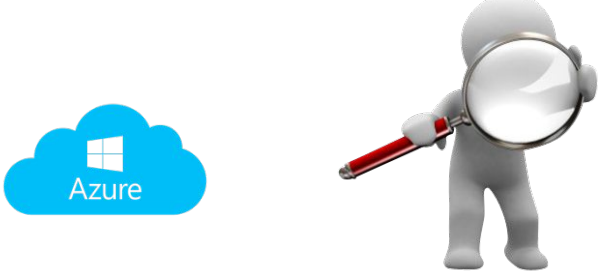
After you create an Amazon RDS DB instance, you can't modify the allocated storage size of the DB instance to decrease the total storage space it uses. To decrease the storage size of your DB instance, create a new DB instance that has less provisioned storage size. Then, migrate your data into the new DB instance using one of the following methods:

- Use the database engine's native dump and restore method.
  Note: This method causes some downtime.
- Use AWS Database Migration Service (AWS DMS) for minimal downtime.

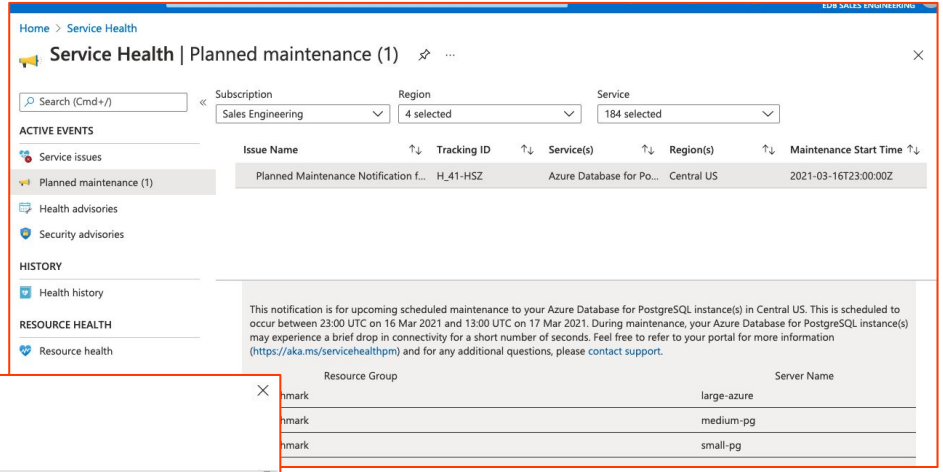
Resolution

DB dump and restore

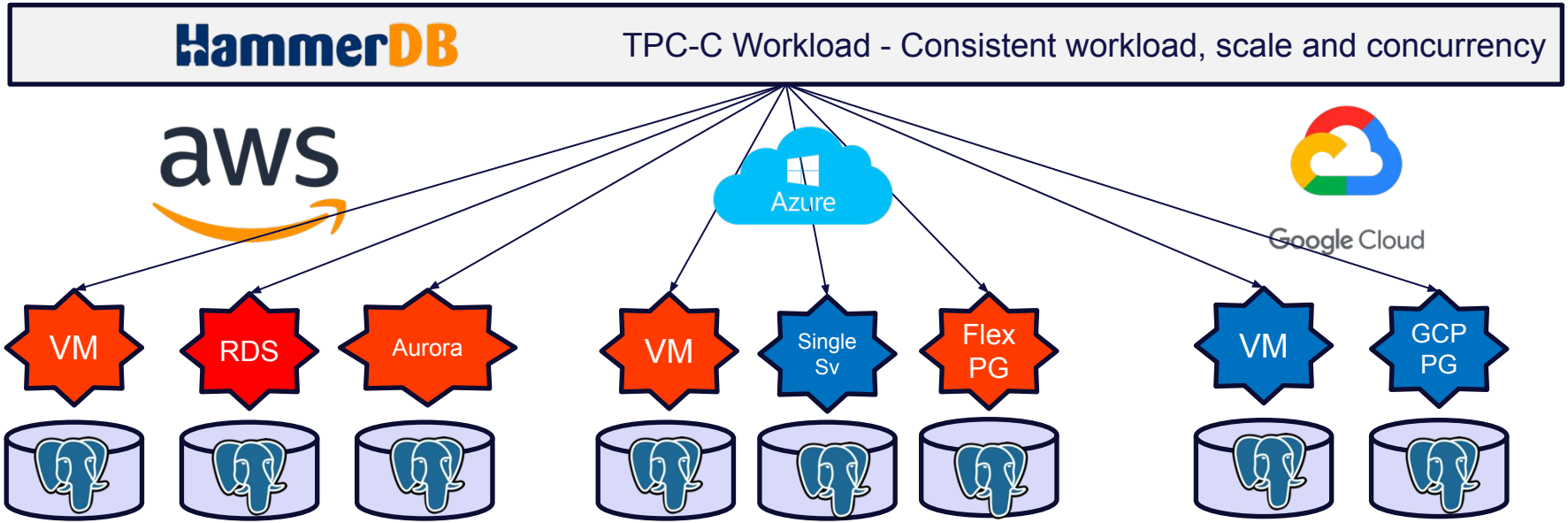
Other Observations - Azure



- 155 Postgres Conf parameters **cannot** be changed
- **Single Server** is built on:
  - Old 7 year old Broadwell CPU
  - 3 IOPS/GB storage performance
  - Postgres 11.6
  - Built on Windows Server
- **Flexible Server** is built on:
  - Not sure – you don't get to choose
  - 4-5 IOPS/GB storage performance
  - Postgres 13.3
- You can scale HP – but



Multi-Cloud Sizing and Benchmarking – to help you decide



Deliverables to you

Make it all more predictable and prescriptive

"How to" – Legacy workload evaluation documents and reproducible scripts

"eBook" - Best practices around cloud setup and database configuration

GitHub Scripts

- To allow you to execute your own 'race' and sizing effort.

Calculator - Financial business case formulas and documents to help speak in 'debit/credit' lexicon

Key Takeaways

- It is difficult to technically compare cloud database offerings – but **cost-per-transaction** is the most important measure.
- Most enterprises inadvertently **over-provision and over-pay** for databases-in-the-cloud.
- Cloud-as-a-service offerings **restrict** your flexibility to tune, configure and optimize based on your unique workloads.
- Some DBaaS offerings are built on obsolete hardware and older releases of Postgres
- ***Let us know if you want to have a specific conversation by cloud and need***
- **MORE TO COME!!**

HANDS ON COMPARISON: Are the cloud Postgres database-as-a-service all they are cracked up to be?

POSTGR
VISION 2021

Register N

POSTGR

JUNE 22 - 2

postgresvision.com

Wednesday, June 23 • 10:15am - 11:00am

Back To Schedule

HANDS ON COMPARISON: Are the cloud Postgres database-as-a-service all they are cracked up to be?

https://sched.co/j8M4 Tweet Share

How do you measure what makes sense from the cloud today? EDB has taken a standard workload and completed extensive comparisons-and-contrasts of AWS, Azure and Google Cloud and want to make sure you know how they stack up. It is easy to think they have turned Postgres into 'a service', but this session will review the technical and economics of each of these options - with a resulting $/transaction you can take to the bank.

Speakers

Tom Rieger

Senior Sales Engineer, EDB

Tom Rieger is a senior solutions engineer for EDB and brings over 30 years of technology experiences across the data landscape always focused on relational technologies. Working in information technologies for companies like Textron and Labatt (yes - the beer) along with vendors like... Read More →

Mark Wong

Performance Engineer, EDB

Mark Wong is currently employed by EDB as a Performance Engineer and is a PostgreSQL Major Contributor. His background is in database systems solutions and performance. He first introduced himself to thePostgreSQL community in 2003 with open source benchmarking kits and performance... Read More →

nd customers

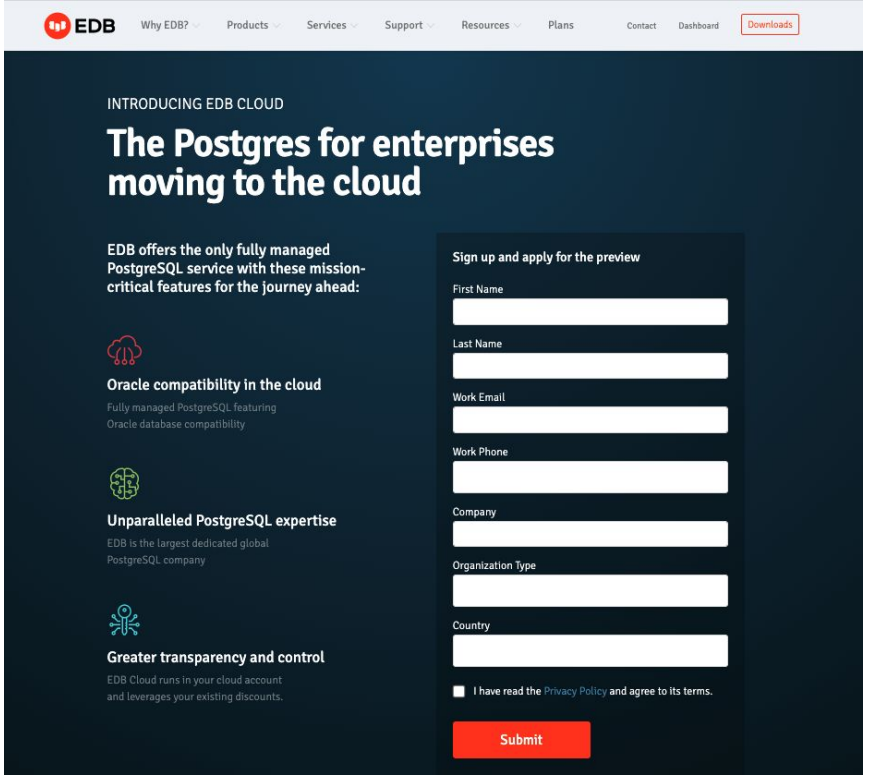
gues

PostgreSQL community connection

You are invited to try something

EDB brings you the best of both worlds

- EDB brings to you a ‘preview’ of what a real database-as-a-services offering should be
- Enterprise focus on the database
- Greater transparency and control – to optimize your cost/transaction!
- www.enterprisedb.com



What you can expect after the webinar

An email from me with…..

- Access to a recording of this webinar
- Access to the slides presented today
- Where you can find the Github library which include repeatable scripts
- Fundamental HammerDB how-to guide to try this yourself
- Where to find the research paper referenced
- Answer any questions

In closing

Questions

Please make sure if you have questions to type them into the tool

Exit survey

There is an exit survey that will pop up – please tell us your thoughts on this webinar

Contact us

Tom Rieger – tom.rieger@enterprisedb.com

CALL ME – 952-221-6514

<LINKEDIN>