PERCONA

Reduce PostgreSQL Costs in the Cloud

Jobin Augustine

PostgreSQL Escalation Specialist

Thanks to

Original Contents

- Daniel Guzman Burgos, PMM Tech Lead, Percona
- Peter Zaitsev, Founder, Percona
- Experience shared by users and customers
- Personal views

PERCONA

Cloud - Experience

ZDNET

tomorrow
belongs to those who embrace it
today

trending innovation home & office business finance education security

/ business

Home / Business / Cloud

Cloud computing: More costly, complicated and frustrating than expected - but still essential

A new report by Capita shows that UK businesses are growing disillusioned by their move to the cloud. It might be because they are focusing too much on the wrong goals.

Written by **Daphne Leprince-Ringuet,** Contributor on Feb. 27, 2020

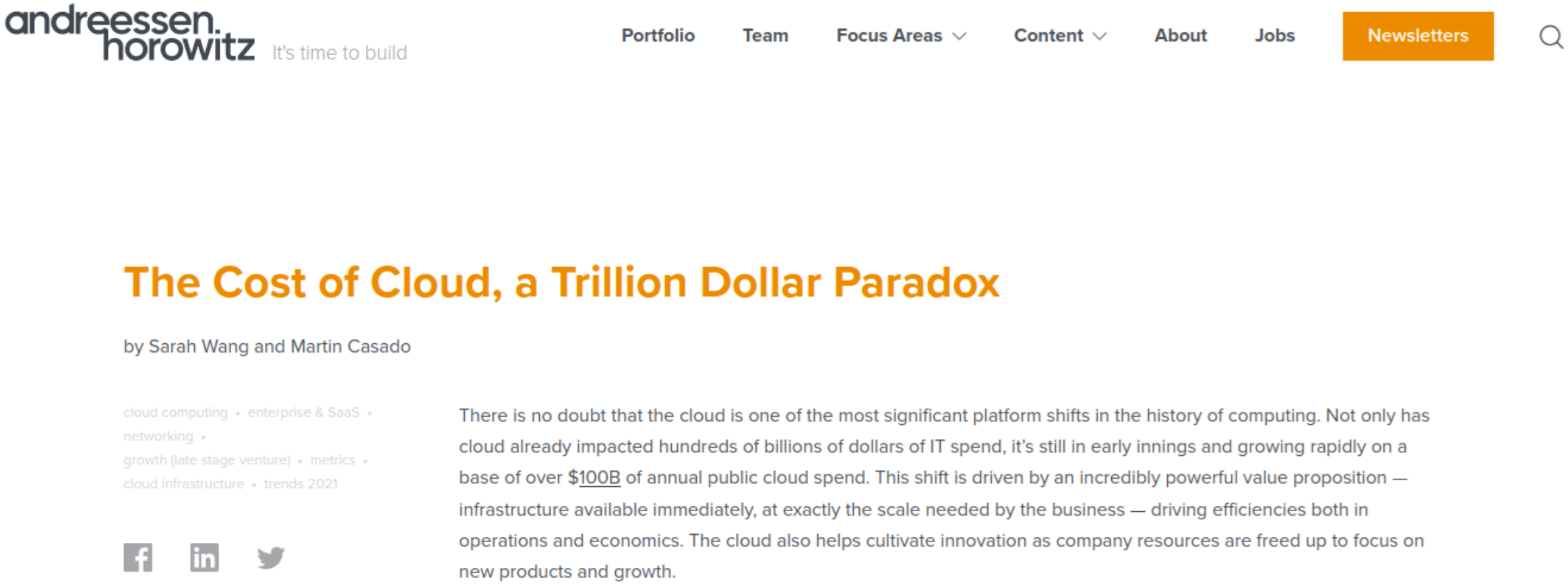
https://www.zdnet.com/article/cloud-computing-more-costly-complicated-and-frustrating-than-expected-but-still-essential/

Cloud Experience

"There has been a sort of hype about cloud in the past few years. Those who have started migrating really focused on cost saving and rushed in without a clear strategy. Now, a high percentage of enterprises have not seen the outcomes they expected."

"Up to 58% of organisations said that moving to the cloud has been more expensive than initially thought"

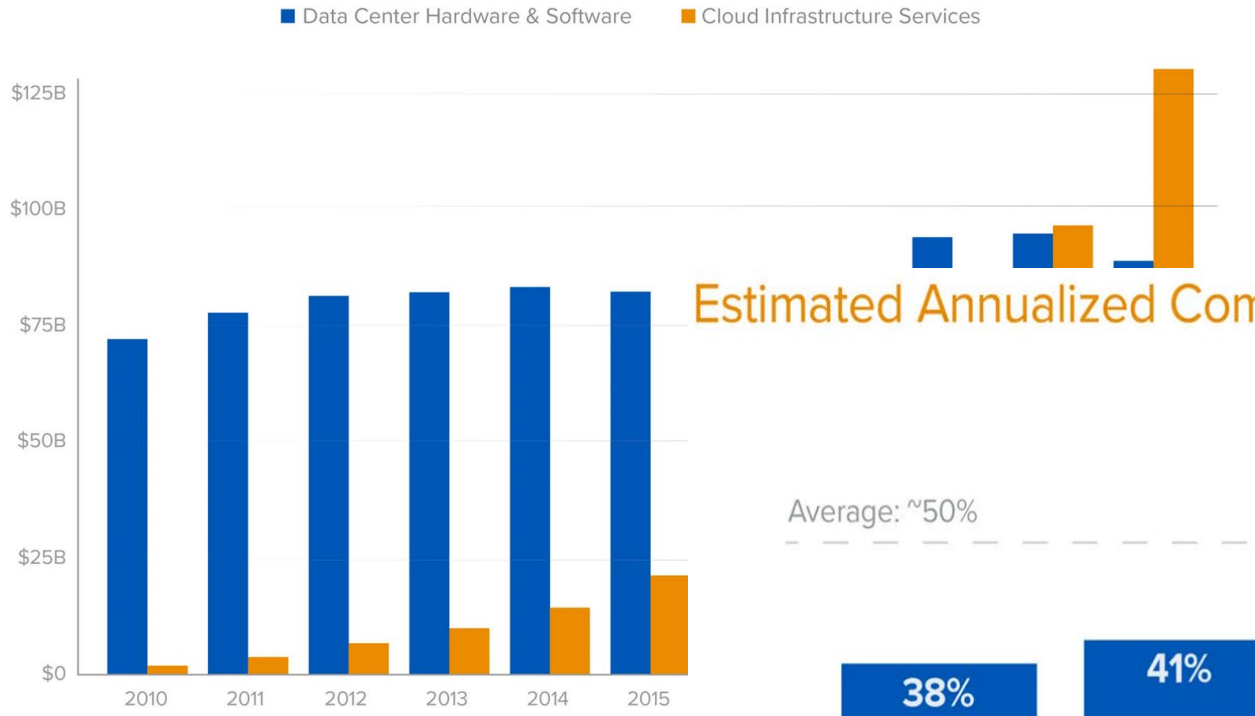
"80% of organisations will overshoot their cloud infrastructure budgets because of their failure to manage cost optimisation."



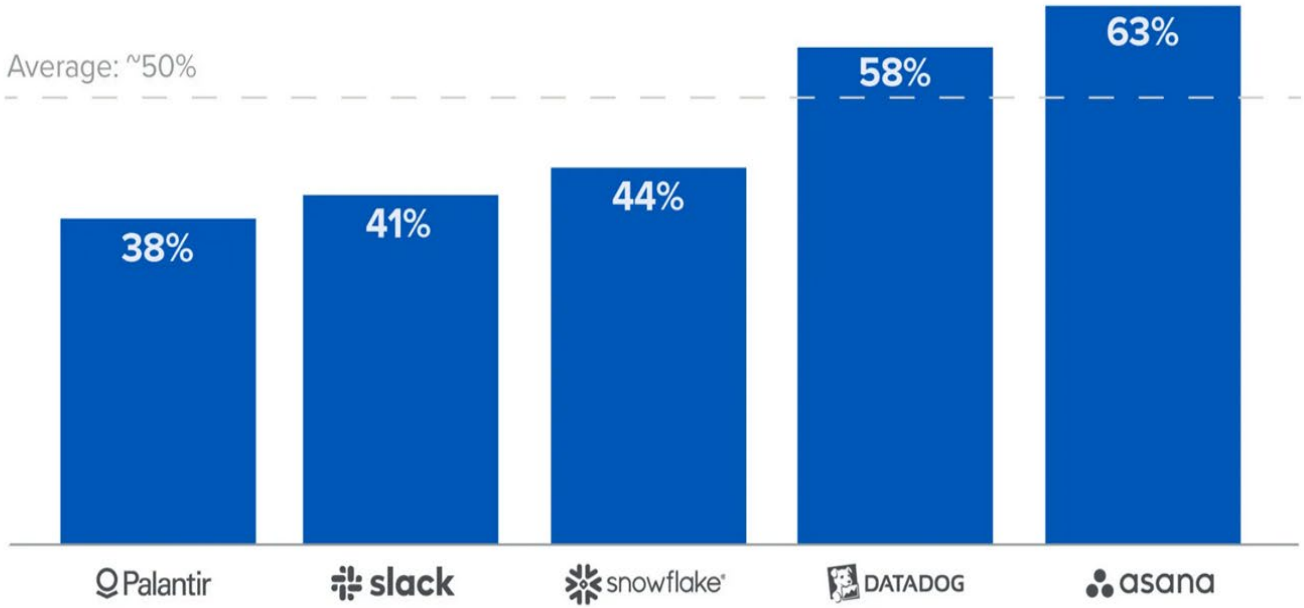
"the pressure it puts on margins can start to outweigh the benefits, **as a company scales and growth slows**. Because this shift happens ***later*** in a company's life, it is difficult to reverse as it's a result of years of development focused on new features"

https://a16z.com/2021/05/27/cost-of-cloud-paradox-market-cap-cloud-lifecycle-scale-growth-repatriation-optimization/

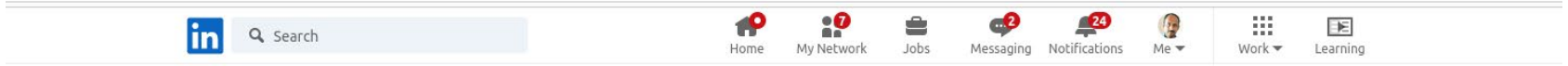
Worldwide Enterprise Spending on Cloud and Data Centers



Estimated Annualized Committed Cloud Spend as % of Cost of Revenue



Cloud spend amounted to 81% of COR is reported



Images may be subject to copyright

The Cloud: How did it get so expensive?

Joshua Ferry
3x Founder | Venture Advisor | Investor | Forbes Contributor

4 articles + Follow

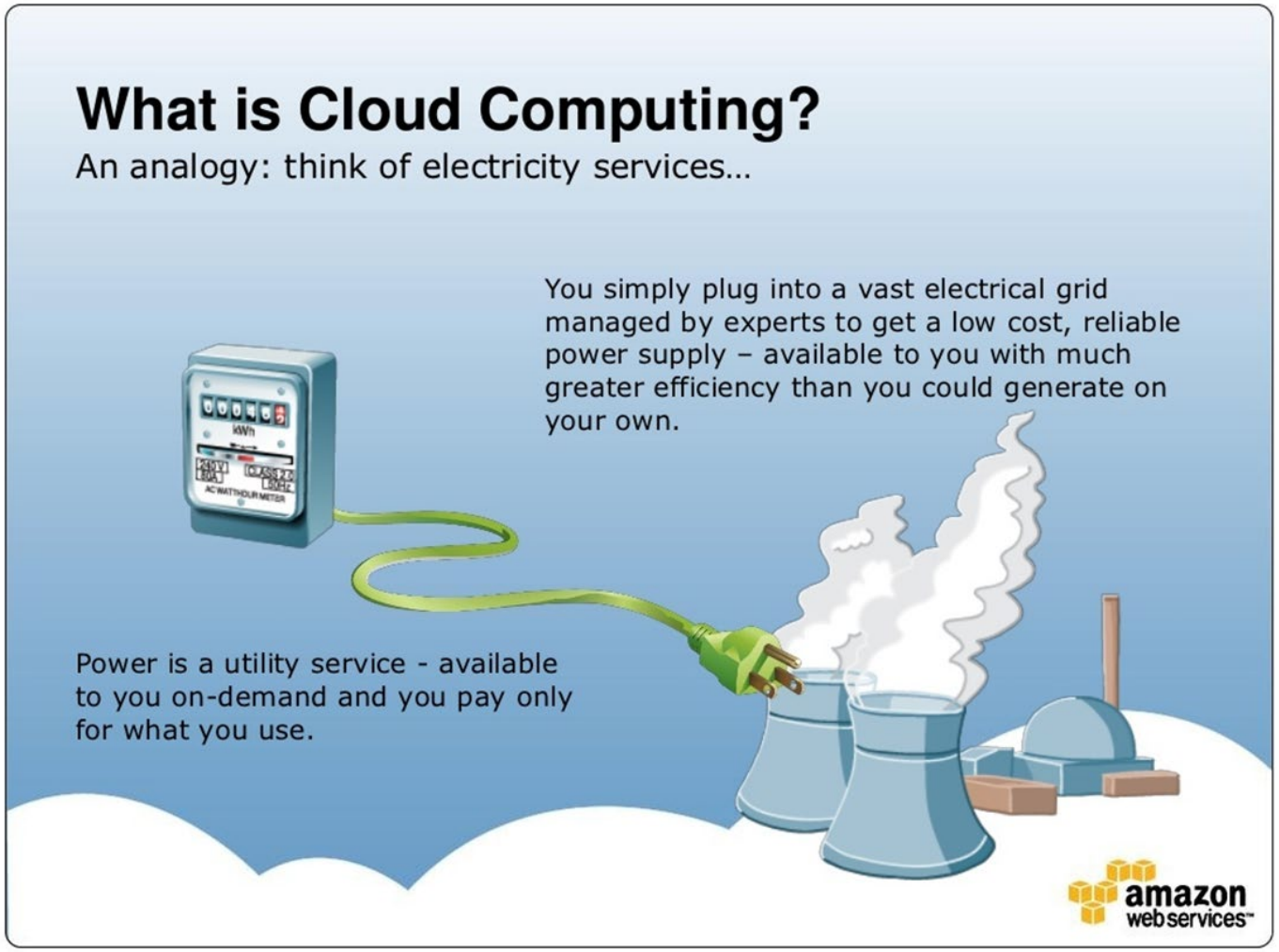
May 31, 2018

CloudInsyte: a Research and Advisory company specializing in Cloud, Cyber Security and Global Data Center procurement

May 30

If you've migrated to a Cloud-based platform such as AWS, GCP or Azure, there's a good chance you've realized that it's starting to cost you more than you initially estimated before making the switch. (If you haven't realized it yet,

How it started?



(Early AWS Presentations promoting the cloud)

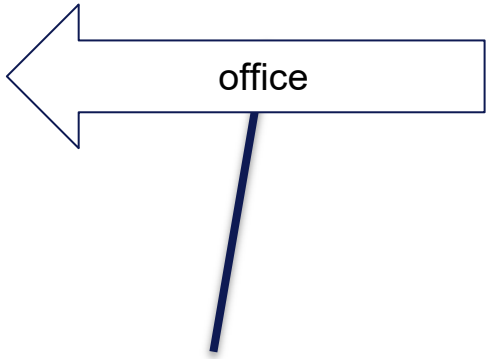
Programmable, instantly accessible infrastructure comes with so much flexibility. It allows new deployment and management approaches we could not even have imagined in the “old world.”

Big Gains

- **Commoditized** the Datacenter, Easily accessible for everyone just like Electricity, Mobile Network
- Reduced the **provisioning time** to negligible.
- Reduced the **Capex**
- Acquire **only when it is required** and let it go.

Long term

- **Commoditization** brings easy competition - Eg. Web hosting
- Reduced margin for providers
- Providers promotes **proprietary services**
- **Increased Opex** for users. Bad times becomes more painful.
- Acquired, But sticky and not easy to release



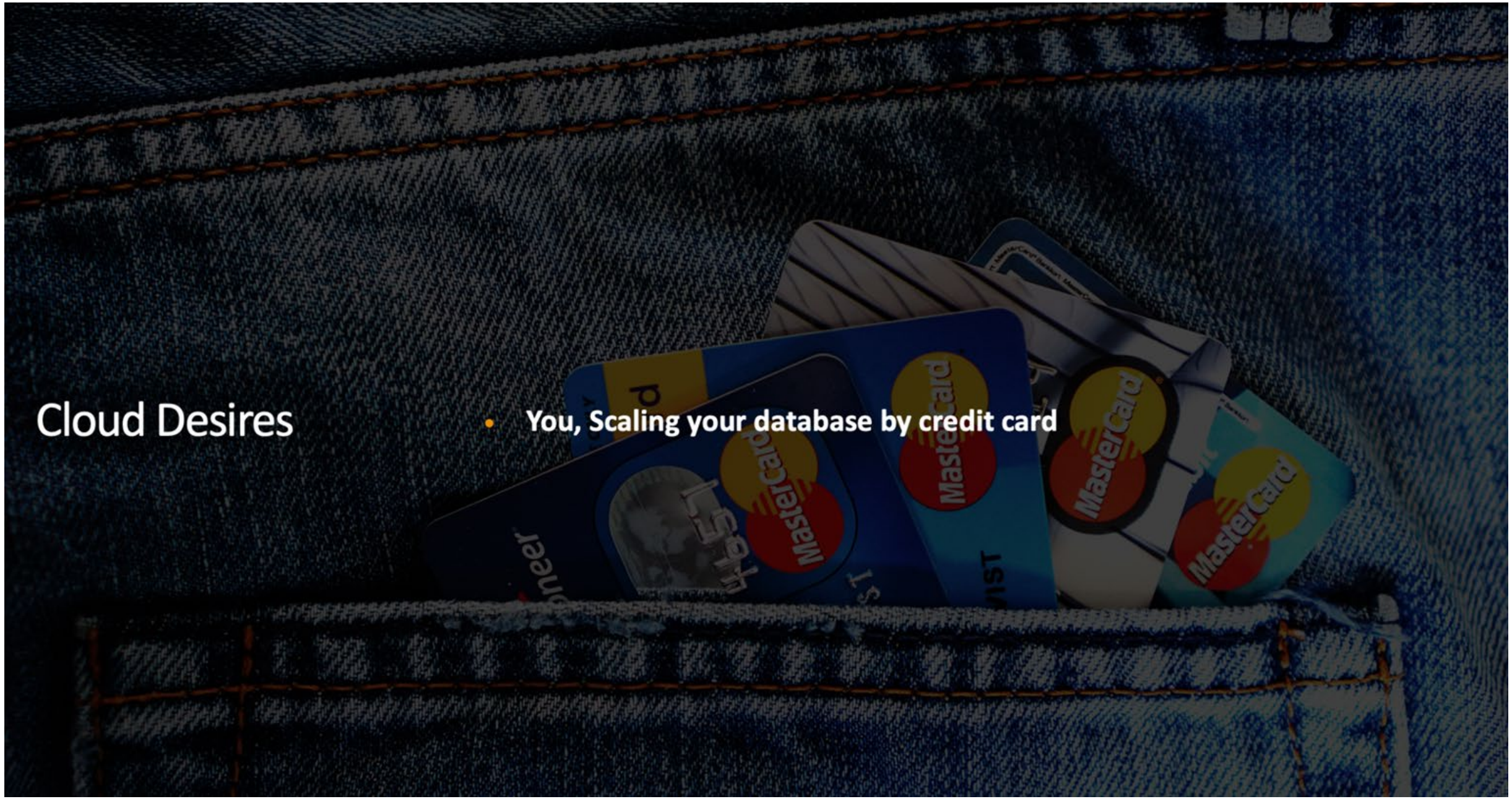
Marketing **proprietary software** and Services

- “We have a better software better than PostgreSQL, We call it XYZPostgreSQL”
  - “Community Edition”, “Vanila PostgreSQL”
- “We have XXX times faster than PostgreSQL”
- “Its fully managed, So you need less people/resources to manage”

Certifications / Courses are used for proprietary software/service marketing.

Cloud Desires

- You, Scaling your database by credit card



When you are running at **scale** in cloud

- You can almost guarantee that there'll be some **underutilized or forgotten resources**.
- When faced with the apparently limitless resources in cloud, most engineers will tend to go for the **larger resources**.
- Making that choice often keeps them from **getting paged in the middle of the night.**

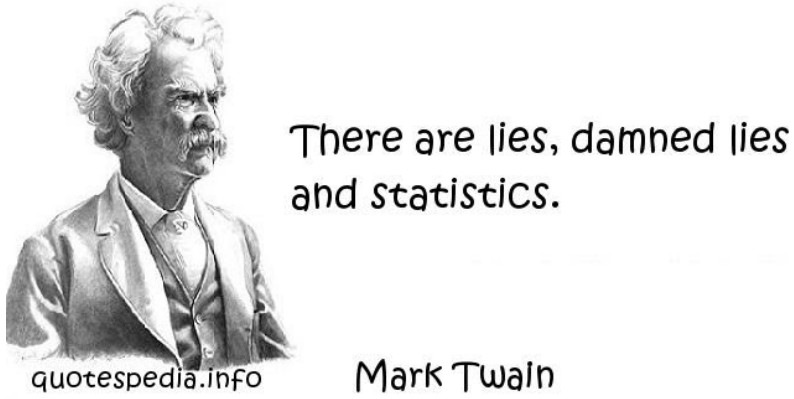
Usage Reduction Challenges

- It can be difficult to know which resources need reviewing.
- It requires that teams have the **right metrics** to understand their application's requirements.

What can we do?

Stage 1. Approach to Cloud

- Critical about claims
  - Test the waters
- Aware about the marketing and conflict of interests
- Benchmark the own workload.
  - Load / Performance testing.



Synthetic benchmarks

Approach to Cloud

- There is not one-time fix
- Requires a cultural Shift
- Engineers to think of cost

Stage 2. Identifying the wastage / over-spending

- Monitoring - usage vs unused
- Assessing cost vs value of proprietary softwares / solutions.
- Engaging the experts.

Are we locked up ourselves?

Stage 3. Identifying lock-ins

- Lock-in with Cloud Vendor
- Use Proprietary Solutions
- Highly Differentiated Cloud
- Hostage
- No Vendor Choice

- Freedom to Run Anywhere
- Use Open Source
- Cloud Is Commodity
- Customer
- Choice of Vendors

Stage 4. Addressing the Wastage

- Decommissioning under utilized
- Multi-Cloud approach - Are we ready?
- Hybrid approach - Are we ready?

Jailbreaking & Right-sizing

Jailbreaks

- IaaS vs DBaaS
- Architectures
- Custom disk configurations
- Baremetal , VMs , Containers/Kubernetes
- Multiple cloud vendors
- Gaining full control over your assets

Evaluation of Options

- Cloud vendors
- IaaS vs DBaaS
- x86 vs ARM
- Intel vs AMD
- Instance types
- Storage Options
- Tablespaces
- Splitting of I/O

Example : ARM based processors in cloud

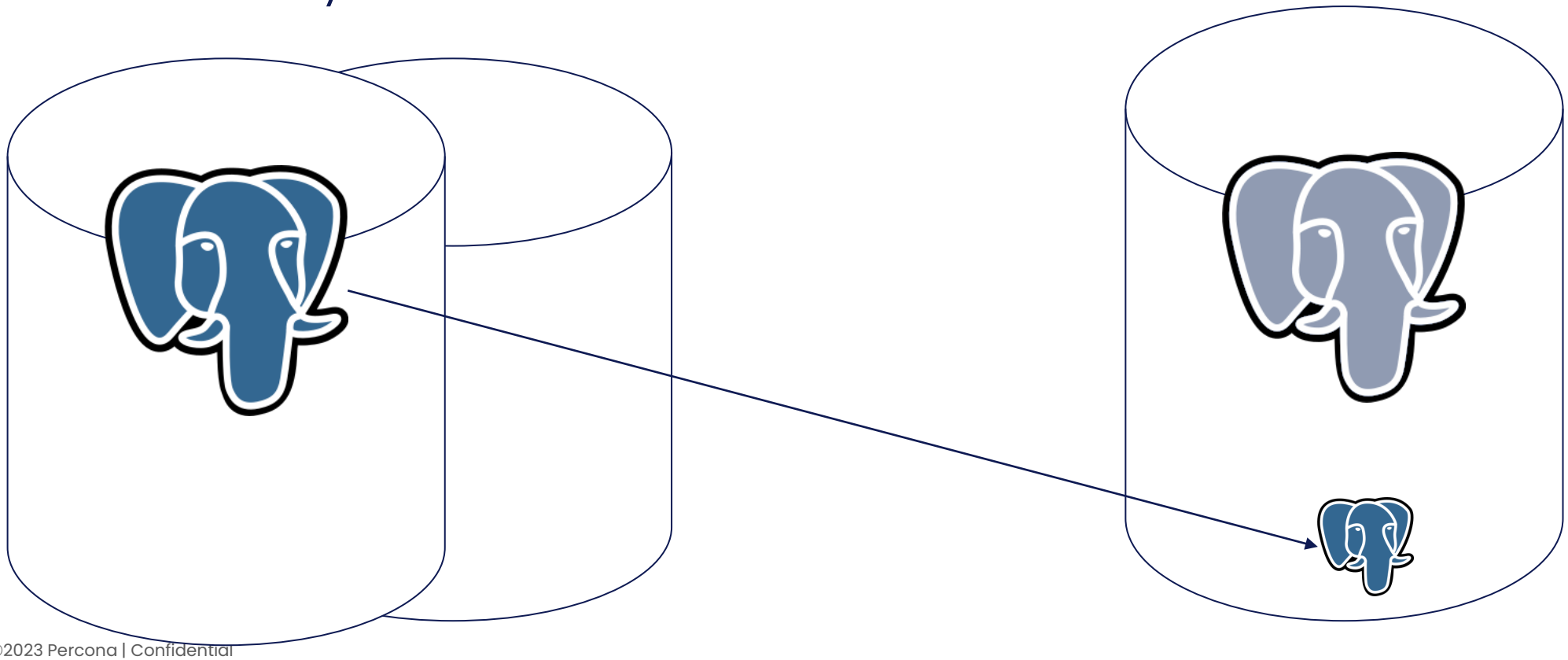
Graviton

Ampere Altra A1

https://www.percona.com/blog/postgresql-on-arm-based-aws-ec2-instances-is-it-any-good/

Multi Data Center

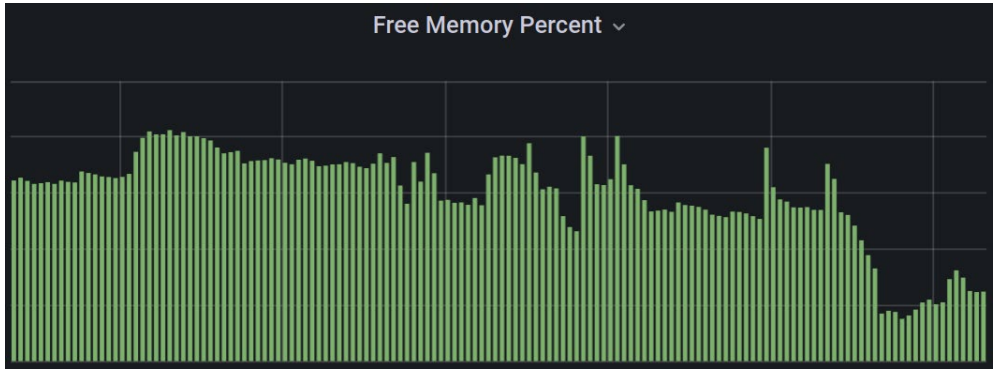
- Multi - DC deployments are costly
- Standby on another datacenter

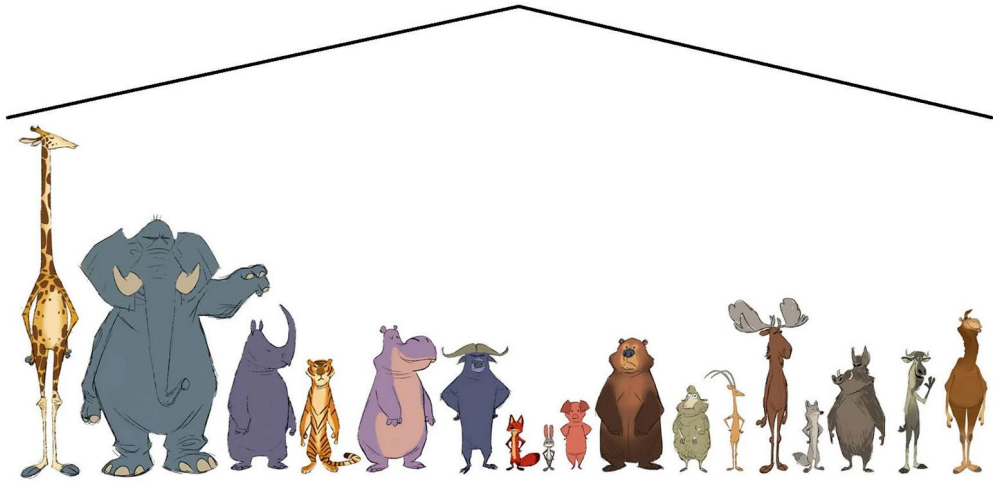


RIGHT SIZING DATABASE INFRA

- Identifying bottlenecks and fixing
- Identifying inefficiencies.
- Identifying wastages
- Improving the schema
- Tuning at different levels.
  - Host hardware
  - Host OS for database workload
  - PostgreSQL Instance
  - Sessions and connection management
  - Schema
  - SQL statements

Load - Spike vs Spreading





- Procuring for the peak load = wastage
- Cloud charges are independent of server resource usage.

- Backups
- ETL jobs
- **Table Vacuum and analyze**

Example : Scheduled vacuum

```
WITH cur_vaccs AS (SELECT split_part(split_part(substring(query from '.* \..*'),'.',2),' ',1) as tab FROM
pg_stat_activity WHERE query like 'autovacuum%')
SELECT 'VACUUM FREEZE "'|| n.nspname ||'"."'|| c.relname ||'";'
 FROM pg_class c
 JOIN pg_namespace n ON c.relnamespace = n.oid
 LEFT JOIN pg_class t ON c.reltoastrelid = t.oid and t.relkind = 't'
 WHERE c.relkind in ('r','m') AND NOT EXISTS (SELECT * FROM cur_vaccs WHERE tab = c.relname)
ORDER BY GREATEST(age(c.relfrozenxid),age(t.relfrozenxid)) DESC
LIMIT 100;
\gexec
```

```
20 11 * * * /full/path/to/psql -X -f /path/to/vacuumjob.sql > /tmp/vacuumjob.out 2>&1
```

https://github.com/jobinau/pgscripts/blob/main/vacuumjob.sql

https://www.percona.com/blog/importance-of-postgresql-vacuum-tuning-and-custom-scheduled-vacuum-job/

TCO - Measure the cost

- Number of people to manage
- Cost of migration till upgrade
- Specialized skills required.
- Time to fix problems
- Freedom to tune
- Cloud Opex cost

Effective Meaning

“fully managed” = “Limited freedom. you can't improve it / fix it”

“Pay as you use” = “Going to pay heavy as the usage increases”

Managed Databases & Unmanageable Troubles

1. PostgreSQL Upgrade

- No more pg_upgrade

2. No more Tablespaces

- All partitions on same disk
- No specific file system possible
  - No compression on storage

3. Segregation of Disks not possible

- Dedicated WAL mount not possible
- PG logs

4. Connection pooler on DB Host not possible

- No option for additional components on DB host

5. Limited Extensions

- No columnar store
- Sharding solutions
- Timescale

6. More time and Effort required

- Example, Parameter change

7. Development scope

- pl/Python, C, C++

8. Disaster Recovery Site

- No switchover and Switchback

9. Logical replication slot continuity

- Easy to do with Patroni

10. Customization & In-depth troubleshooting

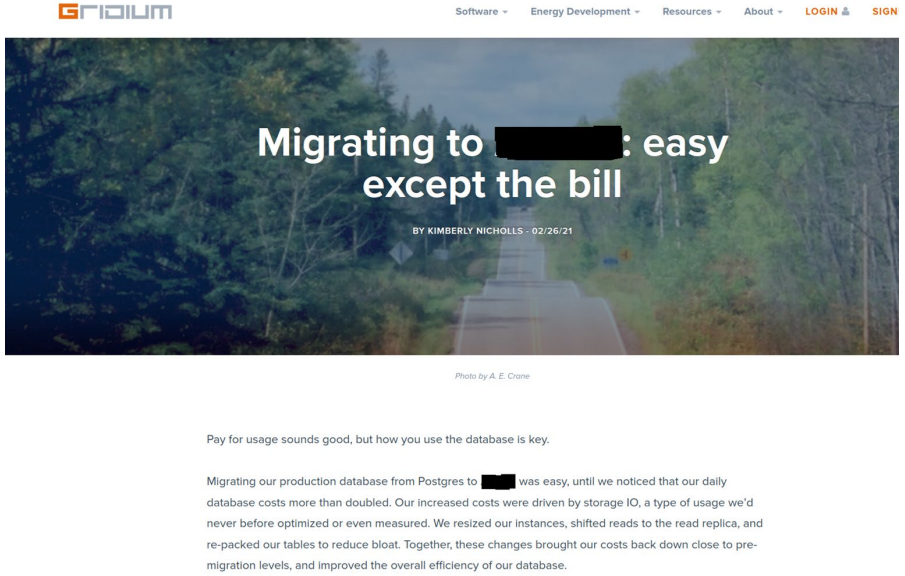
- Custom Dictionary for text search

DAVID HEINEMEIER HANSSON

October 19, 2022

Why we're leaving the cloud

Basecamp has had one foot in the cloud for well over a decade, and HEY has been running there exclusively since it was launched two years ago. We've run extensively in both Amazon's cloud and Google's cloud. We've run on bare virtual machines, we've run on Kubernetes. We've seen all the cloud has to offer, and tried most of it. It's finally time to conclude: Renting computers is (mostly) a bad deal for medium-sized companies like ours with stable growth. The savings promised in reduced complexity never materialized. So we're making our plans to leave.



"It's like paying a quarter of your house's value for earthquake insurance when you don't live anywhere near a fault line"

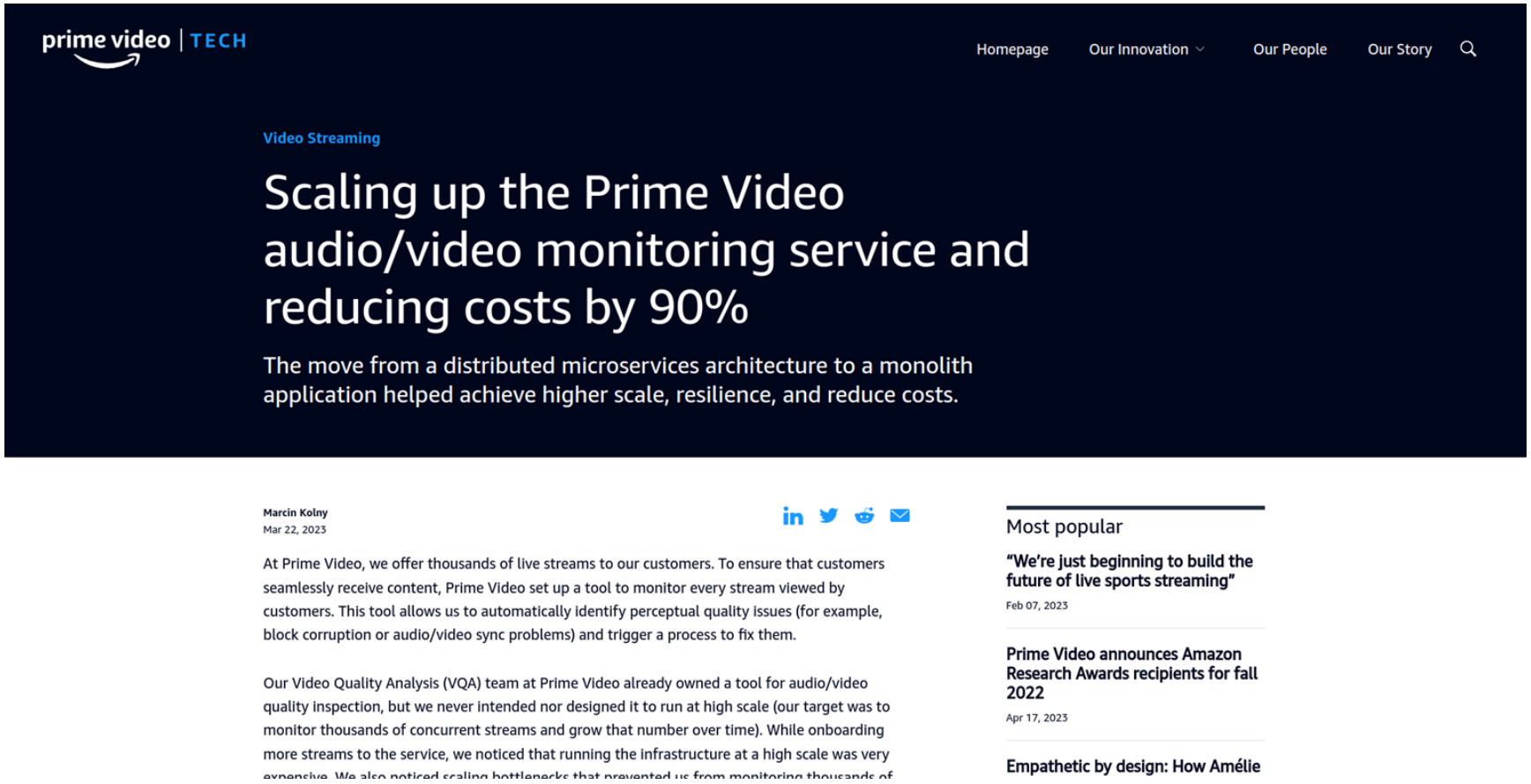
"I've yet to hear of organizations at our scale being able to materially shrink their operations team, just because they moved to the cloud."

"We're paying over half a million dollars per year for database and search services... Do you know how many insanely beefy servers you could purchase on a budget of half a million dollars"

https://world.hey.com/dhh/we-stand-to-save-7m-over-five-years-from-our-cloud-exit-53996caa

https://world.hey.com/dhh/how-to-recover-from-microservices-ce3803cc

Reality check



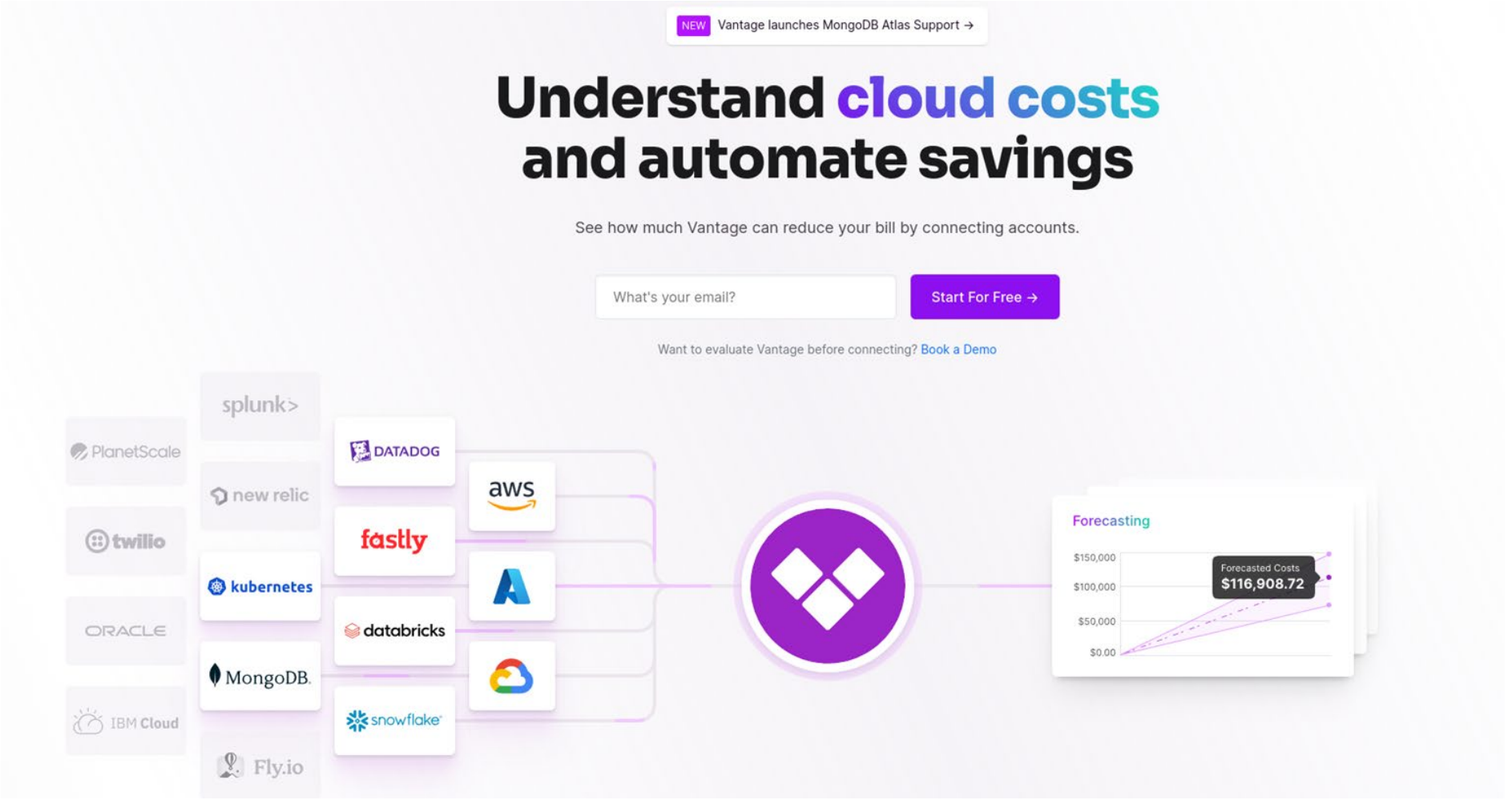
"moved all components into a single process to keep the data transfer within the process memory, which also simplified the orchestration logic"

https://www.primevideotech.com/video-streaming/scaling-up-the-prime-video-audio-video-monitoring-service-and-reducing-costs-by-90

Everything in cloud is **not** too costly

- Storage services are cheap
- Local storage is good.
- Assessing the value becomes more important
- Avail the cost saving features offered by cloud providers
  - Spot instances
  - Long term contracts - Reserved
  - Severless for occasional load.
- Cloud Offerings with I/O billable
  - Higher Instances may become effectively cheaper - active dataset fits in memory
  - Database optimization pays off
- Multi-tenancy on same host

Services for cost optimization.



Summary

- Analyzing the bill
- workload analysis
- Right-sizing exercise
- PostgreSQL is easy to setup and maintain.
- Resource optimization at different levels.
  - Underutilized servers
  - Making informed decisions.
  - Query optimization.
  - Schema optimization
  - Connection management
  - Application behavior analysis.

Important Takeaways

Beware of lock-ins

Strong Marketing

Cost - Opex

Keep multi-cloud always open

IaaS may give better value

Monitoring and assess for optimization

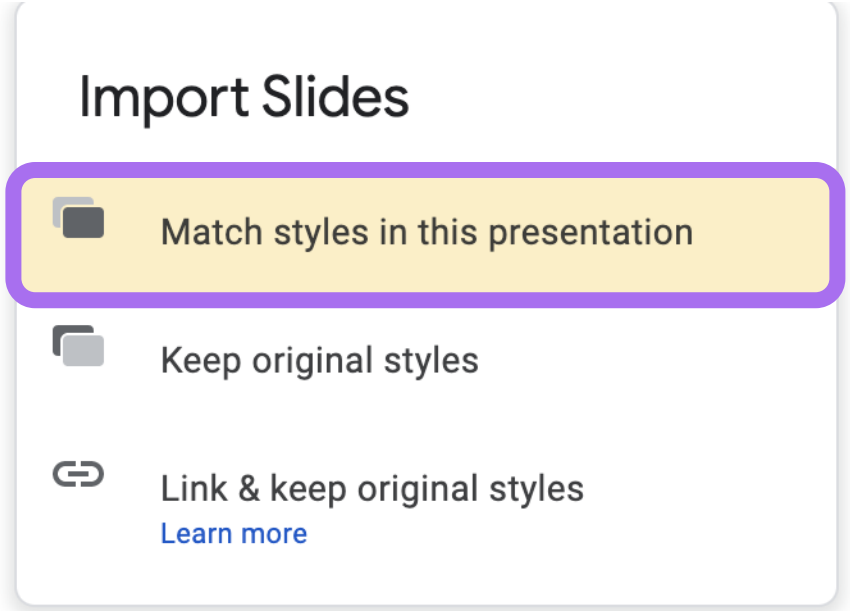
THANK YOU!

percona.com


Once a Perconian,
always a Perconian
WE ARE
Hiring!
VISIT US AT THE BOOTH

Importing slides from another slide deck

When importing another slide or slides to this presentation, please choose the option: **Match styles in this presentation**



Colour/Color Palette

#0e1a53	#5e6686
#ff7e1a	#ffc496
#f24500	#f9a98a
#1486ff	#93c7ff
#f0b336	#f8dca3
#30d1b2	#a0eadc
#a86fef	#d7bdf8
#697793	#bac0cd

Poppins Bold 25

Poppins Semibold Size 12

Poppins Normal size 13
Poppins Normal size 13
Poppins Normal size 13
Poppins Normal size 13
Poppins Normal size 13

- Poppins Normal size 13
- Poppins Normal size 13
- Poppins Normal size 13
- Poppins Normal size 13

Poppins Bold 25

Poppins Semibold Size 12

Poppins Normal size 13
Poppins Normal size 13
Poppins Normal size 13
Poppins Normal size 13
Poppins Normal size 13

- Poppins Normal size 13
- Poppins Normal size 13
- Poppins Normal size 13
- Poppins Normal size 13

Poppins Bold 25

Poppins Semibold Size 12

Poppins Normal size 13
Poppins Normal size 13
Poppins Normal size 13
Poppins Normal size 13
Poppins Normal size 13

- Poppins Normal size 13
- Poppins Normal size 13
- Poppins Normal size 13
- Poppins Normal size 13

Poppins Medium Size 34

Poppins Semibold Size 18

Poppins Normal size 13

Poppins Normal size 13

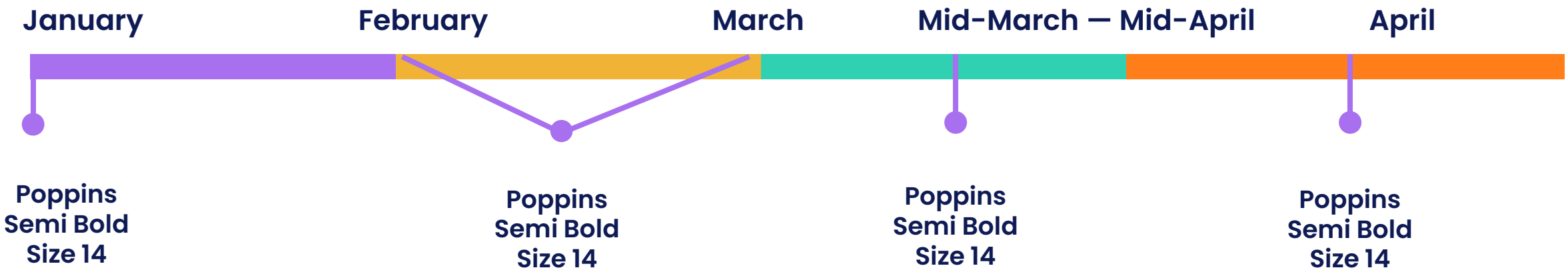
Poppins Normal size 13

Poppins Normal size 13

Poppins Normal size 13

Poppins Normal size 13

Timeline



PERCONA

Databases run better with Percona

Poppins Medium S48

Poppins Normal Size 20

Poppins Bold Size 17

Poppins Normal 12

Poppins Bold Size 17

Poppins Normal 12

Poppins Bold Size 17

Poppins Normal 12

Poppins Bold Size 17

Poppins Normal 12

Poppins Bold 24

Poppins Normal 16

Poppins Medium Size 34

Poppins Semibold Size 18

Poppins Normal size 13

Poppins Normal size 13

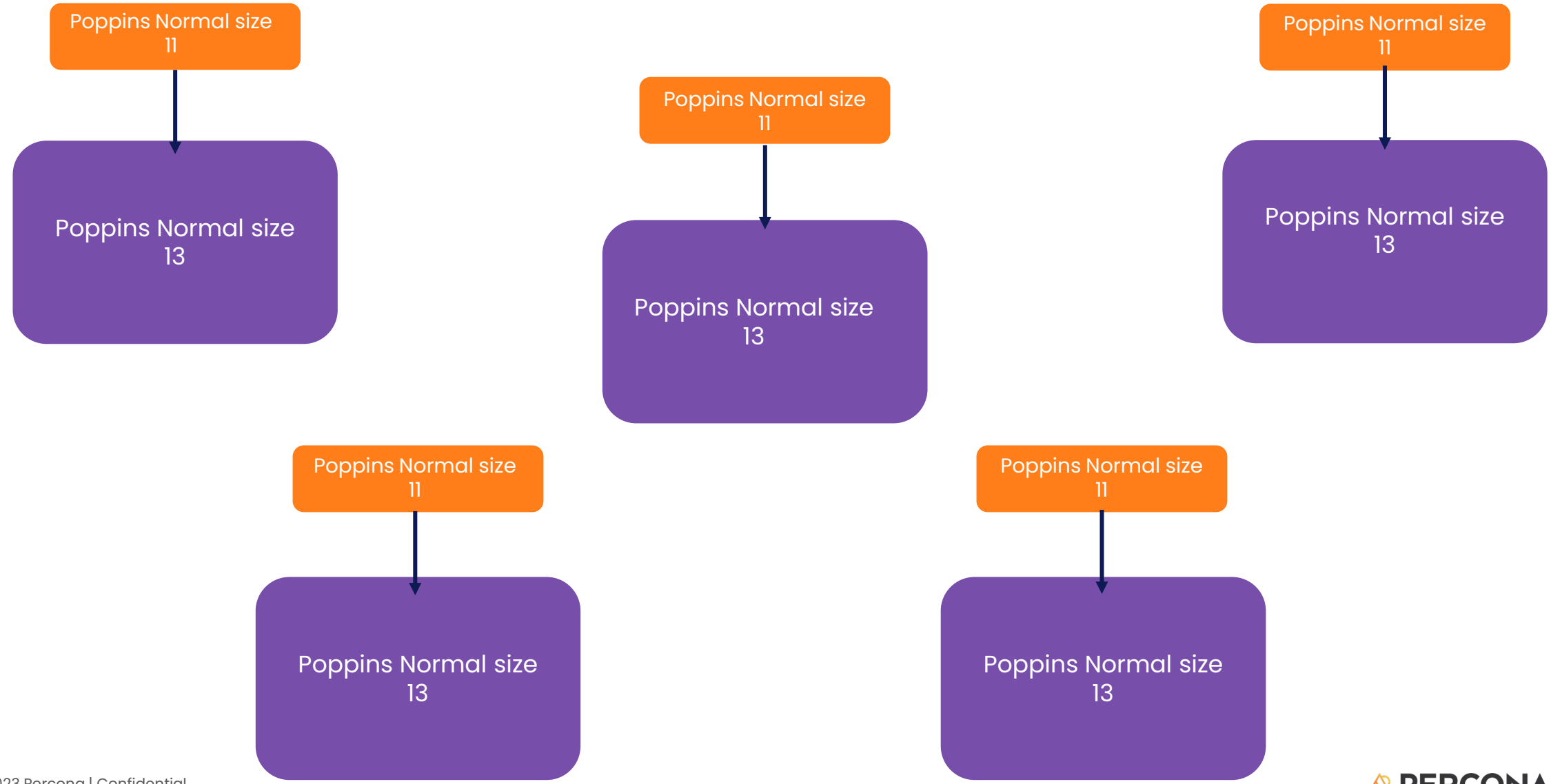
Poppins Normal size 13

Poppins Normal size 13

Poppins Normal size 13

Poppins Normal size 13

Poppins Medium Size 34



Title here - Poppins Medium size 34 colour #oe1a53

Poppins Normal size 20 - colour #oe1a53

Title here - Poppins Medium size 34 colour #oe1a53

Poppins Normal size 20 - colour #oe1a53

Poppins Medium S48

Poppins Normal Size 20

Title - Poppins Medium size 20

Text - Poppins Normal size 18

Title - Poppins Medium size 20

Text - Poppins Normal size 18

Title here - Poppins Medium size 34 colour #oe1a53

Subtitle Poppins Medium Size 20

Text - Poppins Normal size 18

Subtitle Poppins Medium Size 20

Text - Poppins Normal size 18

Poppins Medium size 18

- Poppins Normal size 18
- Poppins Normal size 18
- Poppins Normal size 18

Poppins Medium size 18

- Poppins Normal size 18
- Poppins Normal size 18
- Poppins Normal size 18

Add title here - Poppins Medium size 34 colour #oe1a53

Poppins Normal size 20 - colour #oe1a53

Poppins Medium Size 42

Poppins Medium size 30

Poppins Medium Size 42

THANK YOU!

percona.com