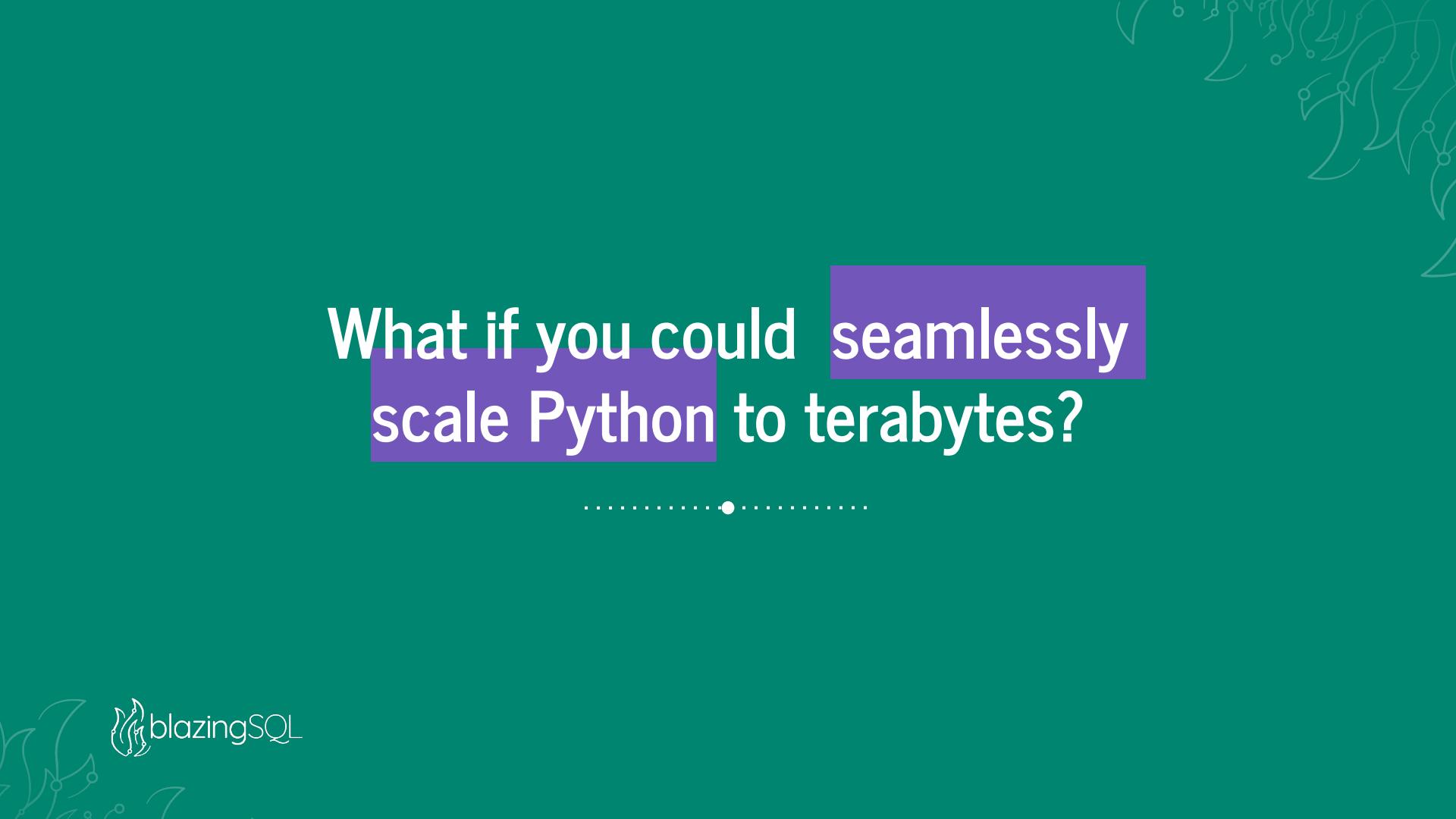




Rodrigo Aramburu

rodrigo@blazingsql.com

@blazingsql



What if you could seamlessly scale Python to terabytes?

.....•.....

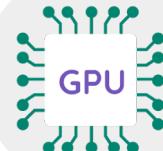
GPU Data Science with RAPIDS

A Python ecosystem for running complete data science workloads on GPU-accelerated packages.



- **Open-Source:**

For standards, interoperability, and peace of mind.



- **GPU-Accelerated:**

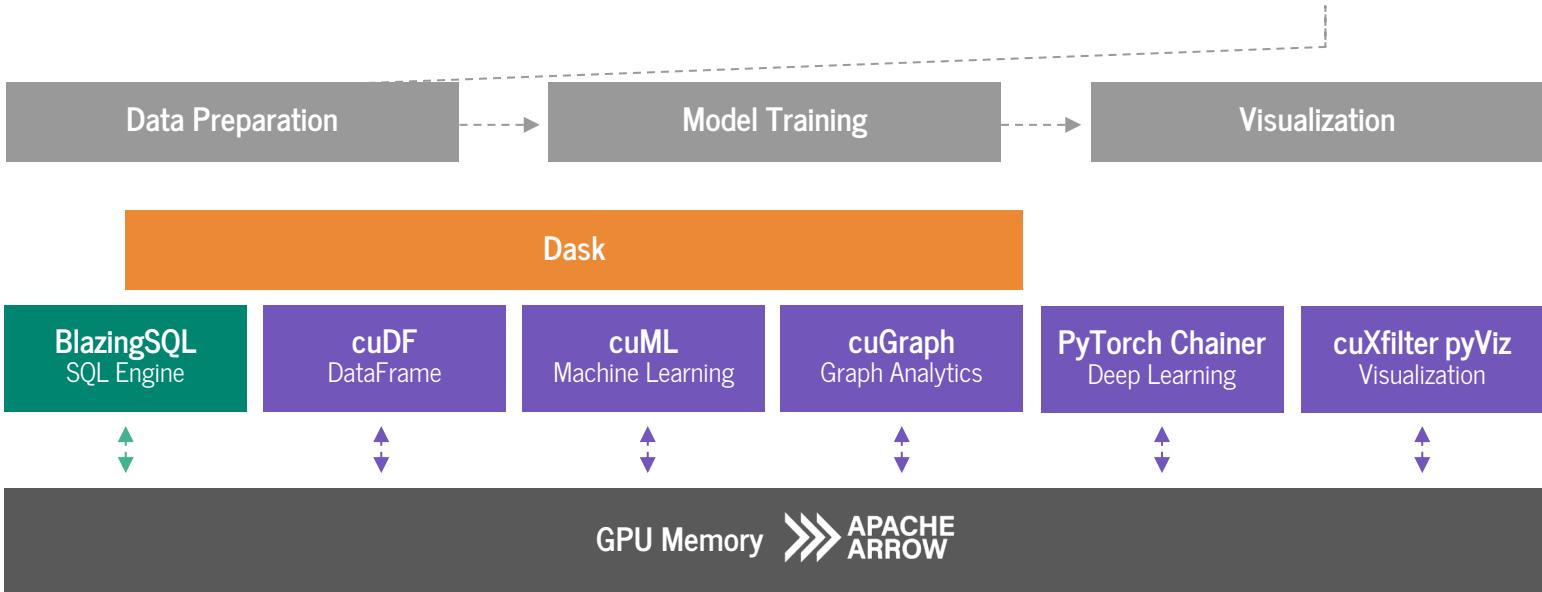
For blazing speeds at a fraction of the cost.



- **Data Science Packages:**

For an easy transition with familiar APIs.

RAPIDS Ecosystem



BENCHMARKS

TPCx-BB (1 TB)

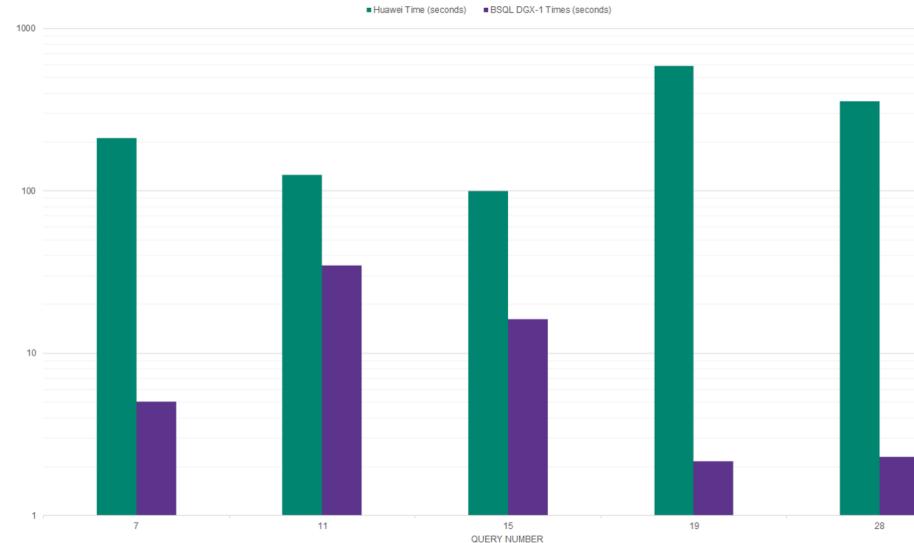
The *Big Data Benchmark (BB)* is used to compare analytics platforms capable of processing large scale data analytics.

Huawei:

- Cloudera for Apache Hadoop 5.9.0
- 16 servers with system cost @ **\$282,762**

RAPIDS:

- BlazingSQL + RAPIDS Stack
- 1 DGX-1 server with system cost @ ~**\$120,000**
- 8 V100 GPUs (256GB VRAM)



*Vertical axis log scale

**Picked 5 queries out of 30 (Best 2, Worst 2, 1 Middle of the Road)



Over 2X
Cheaper



44x Faster
on Average



BLAZINGSQL

SQL on DataFrames

BlazingSQL runs SQL queries on cuDF dataframes. Now you can transform the same dataframe with SQL and pandas-like (cuDF) operations at scale.

```
df = cudf.read_csv('../data/sample_taxi.csv')
bc.create_table('taxi', df)
```

```
bc.sql('select count(*) from taxi')
```

COUNT(*)

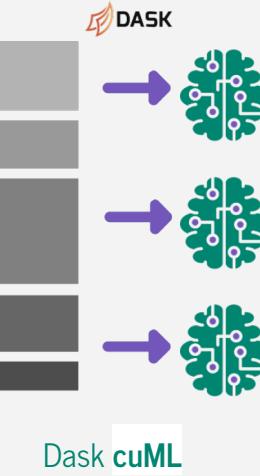
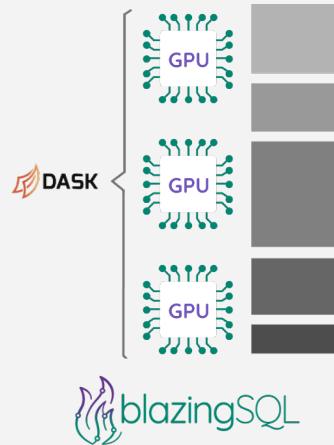
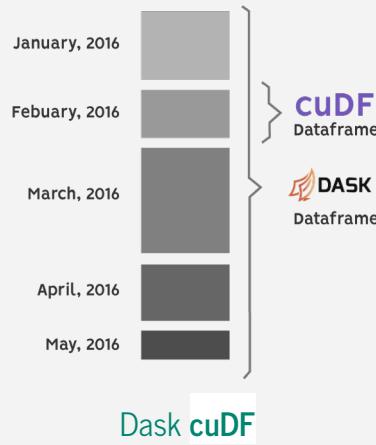
0	743660
---	--------

Demo Time

Recreate all the demos here at

<https://app.blazingsql.com>





DISTRIBUTED
COMPUTE

Dask

BlazingSQL and RAPIDS are able to scale from one to many GPUs by leveraging **Dask**.





More Demo Time



COMMUNITY

Stronger Together

Follow and get involved in our different channels for updates and guide the direction of the ecosystem.



○ Github:

- github.com/blazingdb
- github.com/rapidsai
- github.com/dask



○ Twitter:

- [@blazingsql](https://twitter.com/blazingsql)
- [@rapidsai](https://twitter.com/rapidsai)



○ Blog:

- blog.blazingdb.com
- medium.com/rapids-ai



○ RAPIDS Academy:

- [Security Webinar](#)
- [learnrapids.com](#)

Get Started Today



- **Anaconda:**

Easily install the ecosystem on the laptop, desktop, or server of your choice.

[Install Instructions](#)



- **BSQL Notebooks:**

High-performance Jupyter notebooks with pre-installed packages and a Tesla T4 GPU.

app.blazingsql.com



- **AWS Marketplace:**

Launch your own GPU data science workstation in the cloud with BlazingSQL + RAPIDS.

[AWS Marketplace](#)





—
@blazingsql

www.blazingsql.com

rodrigo@blazingsql.com