



INTRODUCTION TO RAPIDS

The RAPIDS suite of open source software libraries gives you the freedom to execute end-to-end data science and analytics pipelines entirely on NVIDIA GPUs.

The RAPIDS suite of open source software libraries gives you the freedom to execute end-to-end data science and analytics pipelines entirely on NVIDIA GPUs.

Most libraries are also SDKs so as to be used as the backend to other tools

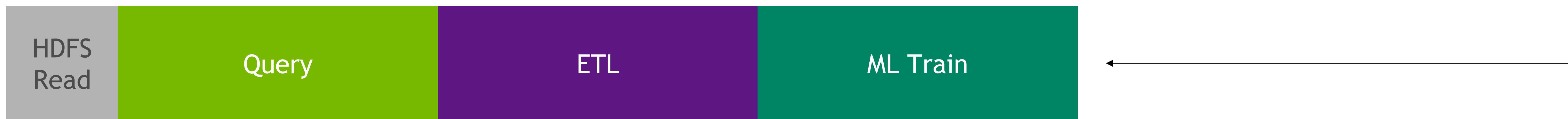
THE EVOLUTION OF DATA PROCESSING

FASTER DATA ACCESS, LESS DATA MOVEMENT

Hadoop Processing, Reading from Disk



CPU-Based Spark In-Memory Processing



25-100x Improvement
Less Code
Language Flexible
Primarily In-Memory

Traditional GPU Processing



5-10x Improvement
More Code
Language Rigid
Substantially on GPU

RAPIDS



50-100x Improvement
Same Code
Language Flexible
Primarily on GPU



Why Use GPUs

GPGPUs are built for intensive parallel processing. As datasets continue to grow, data scientists are limited by the sequential nature of CPU compute. GPUs provide the power and parallelism necessary for today's data science.

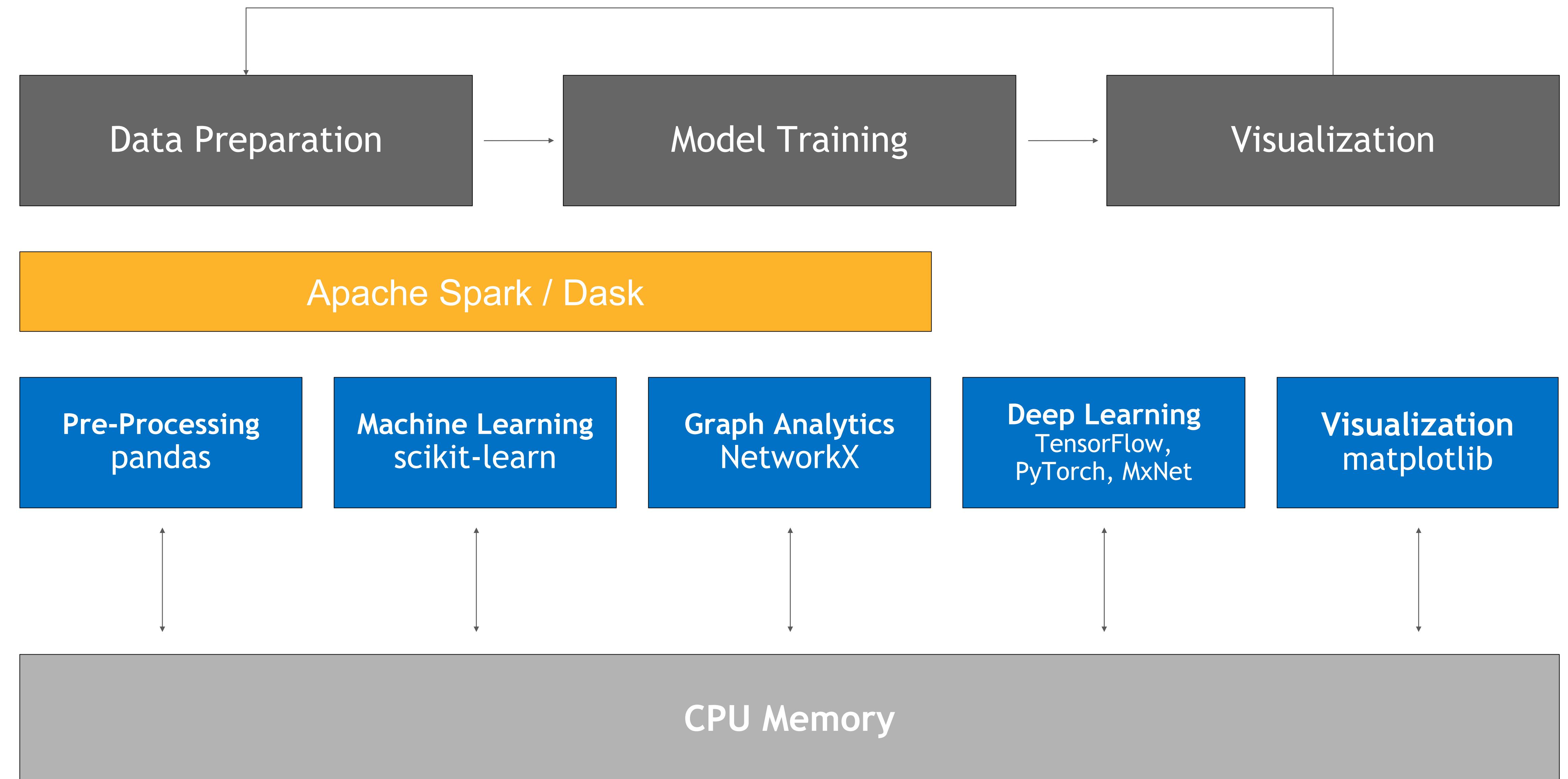
OPEN SOURCE TOOLS HAVE DEMOCRATIZED DATA SCIENCE

HIGHLY ACCESSIBLE, EASY TO USE TOOLS ABSTRACT COMPLEXITY

With the rise of data science, businesses, researchers, and individuals are using data to solve a wide variety of challenges across many problem domains.

Tools like pandas, Apache Spark, and scikit-learn have made applying data science easier, facilitating further data-driven evolution.

While these tools have fostered accessibility in data science, their reliance on traditional CPU processing creates bottlenecks increasing cycle time, operational cost, and overhead.



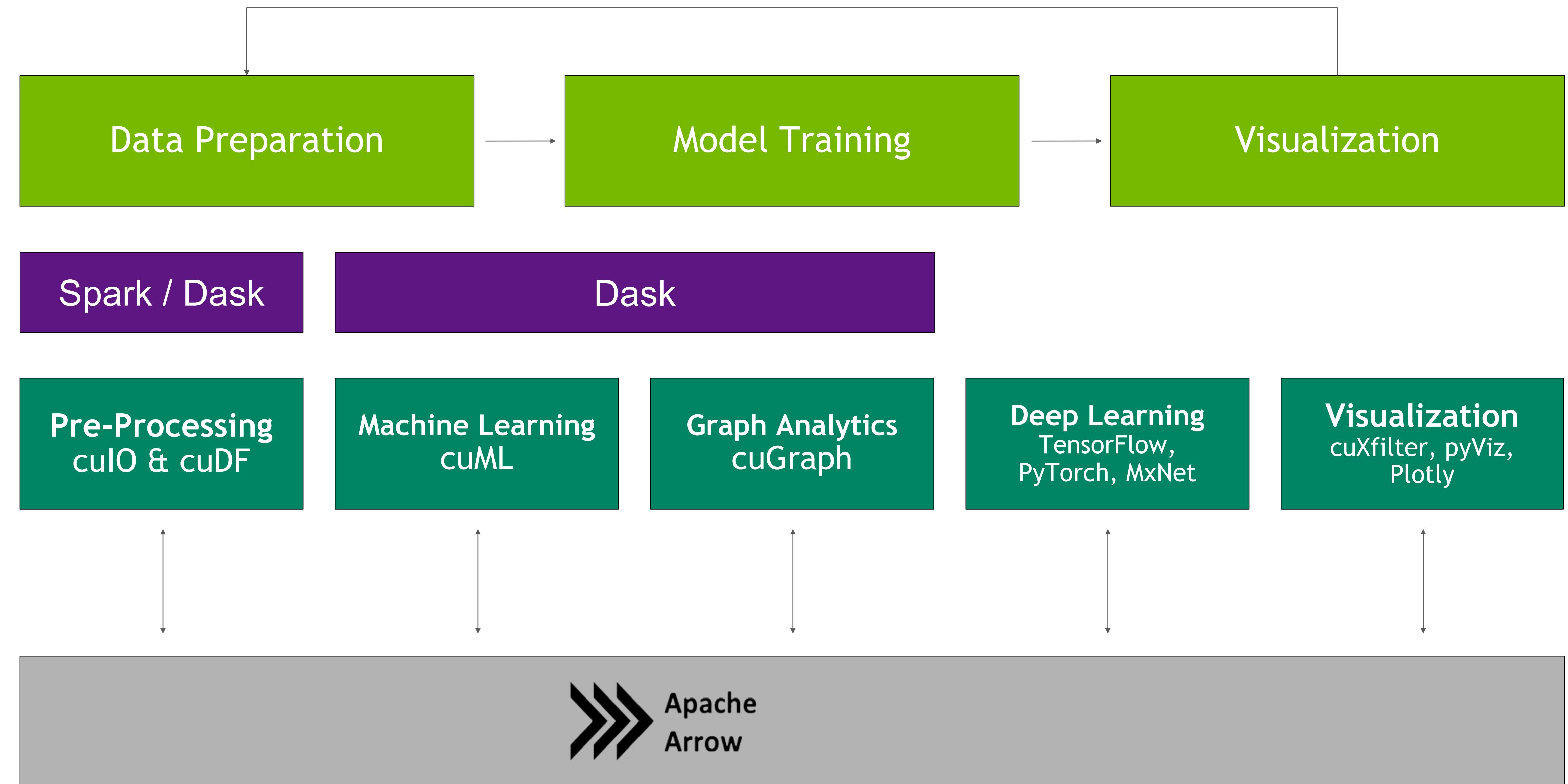
ACCELERATED DATA SCIENCE WITH RAPIDS

POWERING POPULAR DATA SCIENCE ECOSYSTEMS WITH NVIDIA GPUS

RAPIDS abstracts the complexities of accelerated data science by building upon popular Python and Java libraries, enabling users to see benefits immediately.

RAPIDS utilizes NVIDIA CUDA primitives for low-level compute optimization and exposes GPU parallelism and high-bandwidth memory speed through user-friendly interfaces like pandas, scikit-learn, Apache Spark or Dask.

With Apache Spark or Dask, RAPIDS can scale out to multi-node, multi-GPU cluster to power through big data processes.





A Perfect Match for Data Science

Columnar layout leverages GPU strengths.

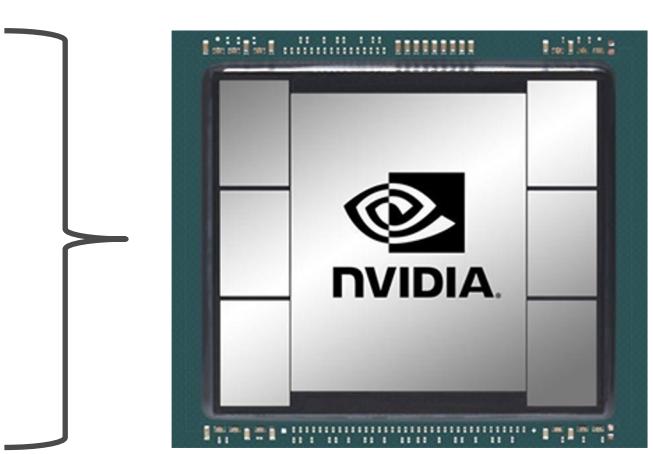
CPU Single Instruction Multiple Data (SIMD)
instructions also take advantage of it.

Consistency with CPU version simplifies
development and conversion.

Emphasis on zero-copy and shallow-copy
operations minimizes a core bottleneck.

	user_id	timestamp	source_ip
row_1	6939800	2021-04-12 05:31	42.155.123.142
row_2	4015666	2021-04-12 05:47	67.132.212.125
row_3	5456236	2021-04-12 05:52	81.205.129.121

	Row-wise data layout	Columnar data layout
row_1	6939800 2021-04-12 05:31 42.155.123.142	user_id 6939800 4015666 5456236
row_2	4015666 2021-04-12 05:47 67.132.212.125	timestamp 2021-04-12 05:31 2021-04-12 05:47 2021-04-12 05:52
row_3	5456236 2021-04-12 05:52 81.205.129.121	source_ip 42.155.123.142 67.132.212.125 81.205.129.121



MINOR CODE CHANGES FOR MAJOR BENEFITS

ABSTRACTING ACCELERATED COMPUTE THROUGH FAMILIAR INTERFACES

CPU

pandas

```
In [1]: import pandas as pd
In [2]: df =
pd.read_csv('filepath')
```

CPU Spark

```
spark.sql("""
select
    order
    count(*) as
order_count
from
    orders""")
```

scikit-learn

```
In [1]: from
sklearn.ensemble import
RandomForestClassifier
In [2]: clf =
RandomForestClassifier()
In [3]: clf.fit(x, y)
```

NetworkX

```
In [1]: import networkx
as nx
In [2]:
page_rank=nx.pagerank(gra
ph)
```

GPU

cuDF

```
In [1]: import cudf
In [2]: df =
cudf.read_csv('filepath')
```

GPU Spark

```
spark.conf.set("spark.rapids
.enabled","true")

spark.sql("""
select
    order
    count(*) as
order_count
from
    orders""")
```

cuML

```
In [1]: from cuml.ensemble
import
RandomForestClassifier
In [2]: cuclf =
RandomForestClassifier()
In [3]: cuclf.fit(x, y)
```

cuGraph

```
In [1]: import cugraph
In [2]:
page_rank=cugraph.pageran
k(graph)
```

Average Speed-Ups: 150x

Average Speed-Ups: 10x

Average Speed-Ups: 50x

Average Speed-Ups: 250x

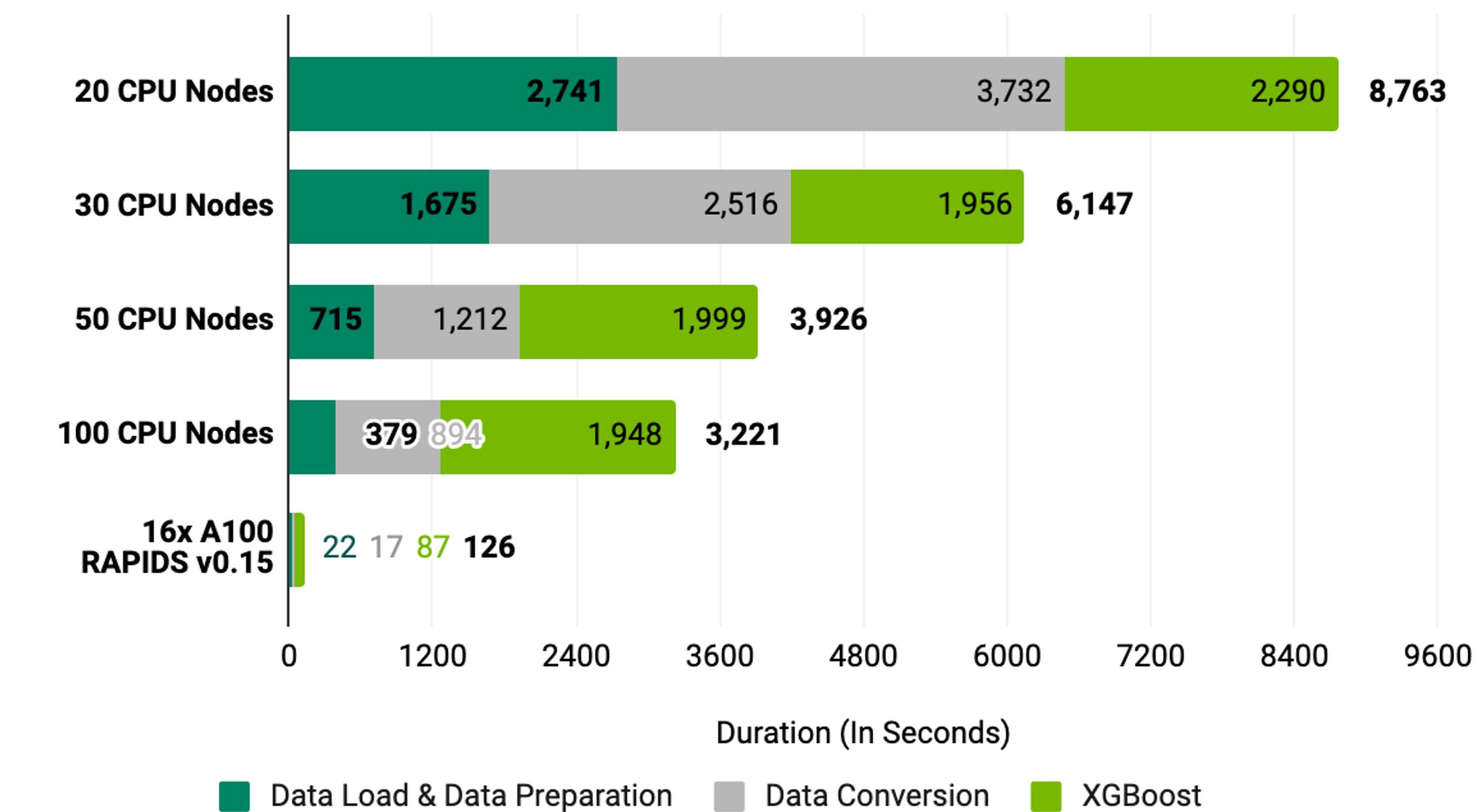
LIGHTNING-FAST END-TO-END PERFORMANCE

REDUCING DATA SCIENCE PROCESSES FROM HOURS TO SECONDS

RAPIDS delivers massive speed-ups across the end-to-end data science lifecycle. Conducting benchmarks in a commercial cloud environment, we're able to get incredible performance running a common ML model training pipeline.

Between loading and cleansing data, engineering features, and training a classifier using a 200GB CSV dataset, a RAPIDS-based pipeline completed these operations in *just over two minutes*. The same process takes two and half hours on a similar CPU-configuration.

RAPIDS End-to-End Workflow Runtimes



16

A100s Provide More Power than 100 CPU Nodes

70x

Faster Performance than Similar CPU Configuration

20x

More Cost-Effective than Similar CPU Configuration

*CPU approximate to n1-highmem-8 (8 vCPUs, 52GB memory) on Google Cloud Platform. TCO calculations-based on Cloud instance costs.

IMPROVING PERFORMANCE OVER TIME

HOLISTIC DEVELOPMENT YIELDS LONG-TERM RESULTS

As RAPIDS has matured, performance has continued to improve, drastically reducing cycle time and costs for data science operations. Since its initial release in fall 2018, RAPIDS has improved performance *nearly 3x* by thinking about software and hardware holistically.

Through whole-stack innovation, RAPIDS continues to improve as NVIDIA hardware and software advance. As new, more powerful GPU architectures, RAPIDS makes it simple for data scientists to realize the benefits of innovative hardware.

2.5

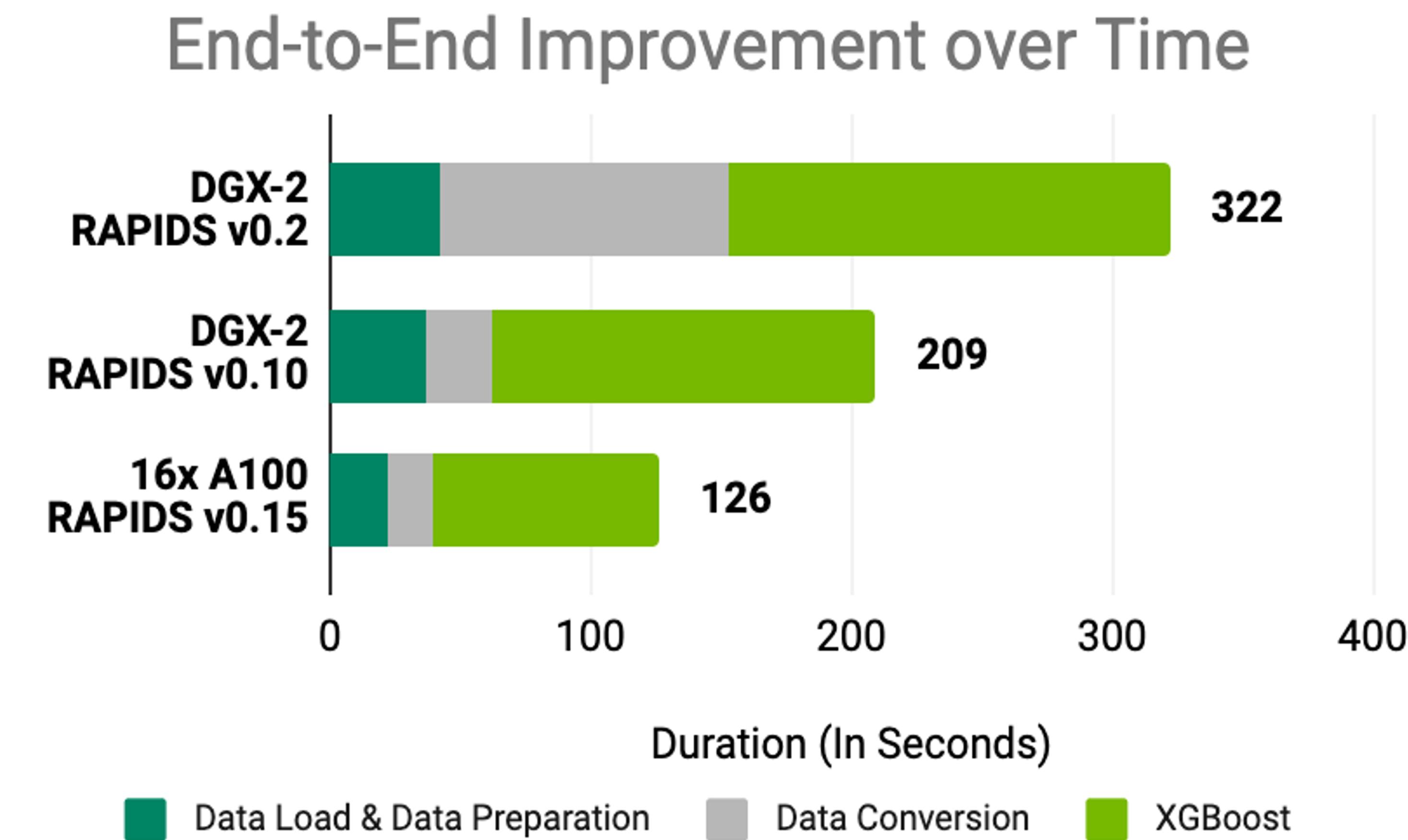
Year between initial and
latest RAPIDS release

3X

Faster Performance than
Similar CPU Configuration

7X

Better TCO with Improved
Compute Performance



THE RAPIDLY GROWING RAPIDS ECOSYSTEM SUPPORTED, USED, & EXTENDED BY A WIDE VARIETY OF PARTNERS

CONTRIBUTORS



ADOPTERS



Booz | Allen | Hamilton



OPEN SOURCE



nuclio



DEPLOY RAPIDS *EVERWHERE*

EASILY ACCELERATED YOUR WORKLOADS IN YOUR ENVIRONMENT

GET SOFTWARE



[GITHUB](#)



[ANACONDA](#)



[DOCKER](#)

CLOUD PROVIDERS

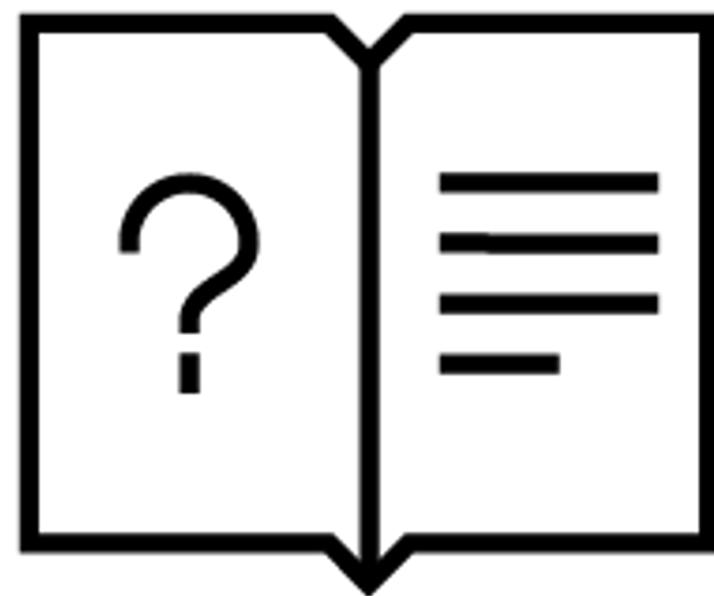


CLOUD SERVICES



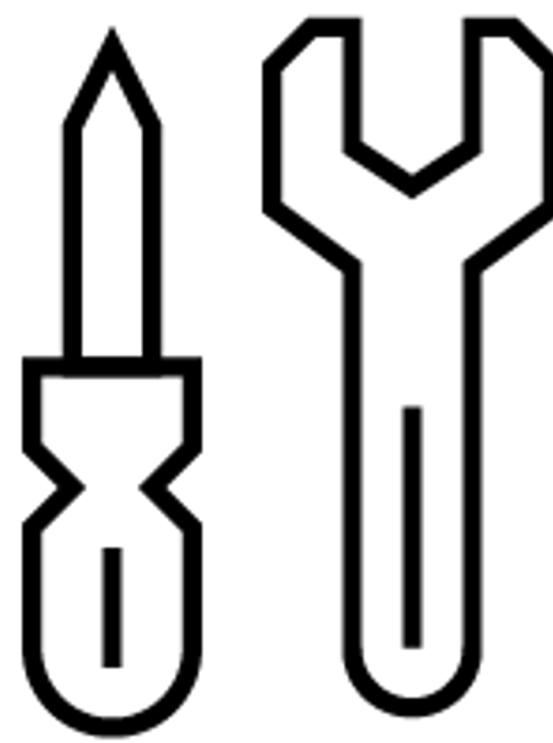
HOW TO GET STARTED WITH RAPIDS

A VARIETY OF WAYS TO GET UP & RUNNING



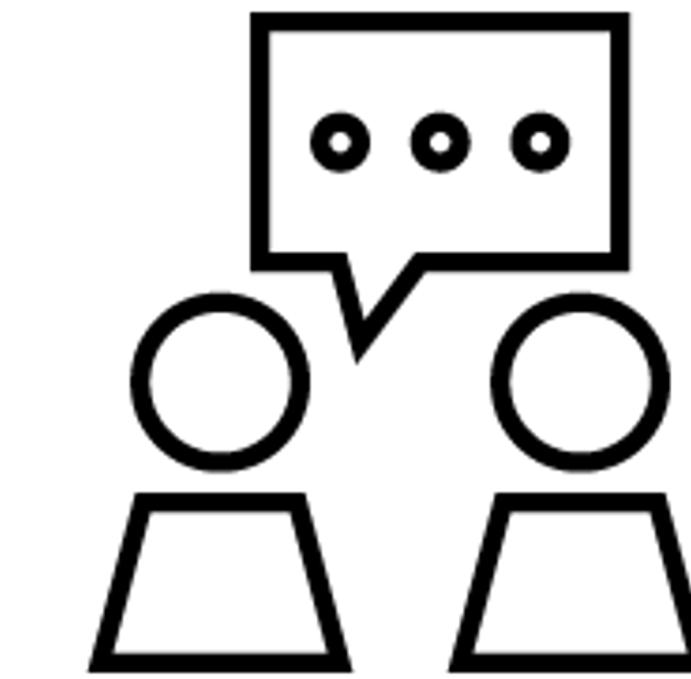
More about RAPIDS

- Learn more at [RAPIDS.ai](#)
- Read the [API docs](#)
- Check out [the RAPIDS blog](#)
- Read the [NVIDIA DevBlog](#)



Self-Start Resources

- Get started with [RAPIDS](#)
- Deploy on [the Cloud today](#)
- Start with [Google Colab](#)
- Look at [the cheat sheets](#)



Discussion & Support

- Check the [RAPIDS GitHub](#)
- Use the [NVIDIA Forums](#)
- Reach out on [Slack](#)
- Talk to [NVIDIA Services](#)