RAPIDS

Open GPU Data Science

@RAPIDSai

https://github.com/rapidsai

https://rapids-goai.slack.com/join

RAPIDS

https://rapids.ai

General Purpose and Domain-Specific Libraries

cuDF

- GPU-accelerated ETL functions
- Tracks Pandas and other common PyData APIs
- Dask + UCX integration for scaling

RAPIDS for Apache Spark

- RAPIDS accelerator for Apache Spark

RAPIDS ML

- GPU-native cuML library, plus XGBoost, FIL, HPO, and more

cuGraph

- GPU graph analytics, including Louvain, PageRank, and more
- Multi node Multi GPU features

cuxfilter

- GPU-accelerated cross-filtering

Viz integration

- pyViz: Plotly Dash, Bokeh, Datashader, HoloViews, hvPlot
- Node-RAPIDS bindings for node.js

Domain-Specific Libraries

CLX + Morpheus
Cyber log processing + anomaly detection

cuStreamz
Streaming analytics

cuSignal
Signals processing

cuCIM
Computer vision & image processing primitives

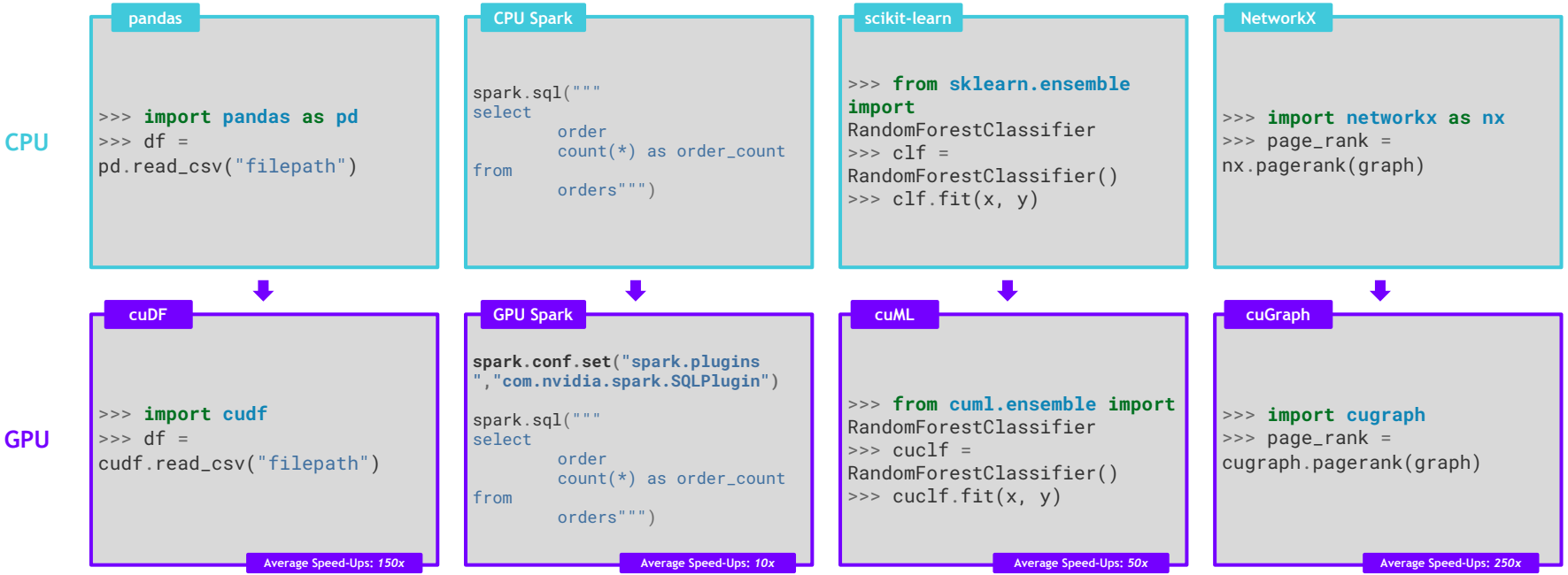
...and more!

cuSpatial
Spatial analytics

NVTabular
RecSys ETL library

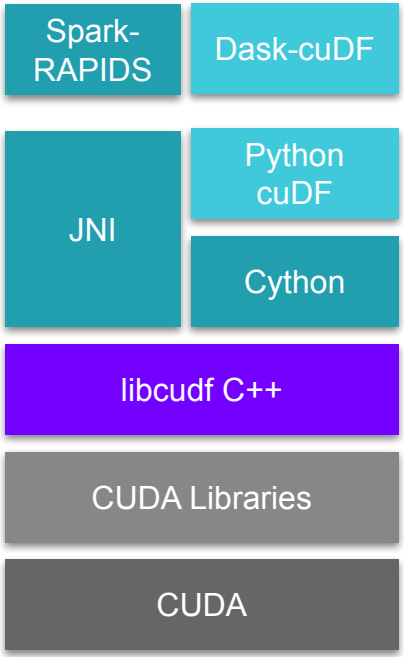
Minor Code Changes for Major Benefits

Abstracting Accelerated Compute through Familiar Interfaces



What is RAPIDS ETL

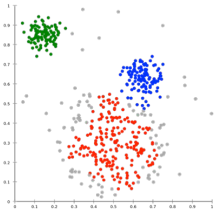
Expandable platform for GPU data science

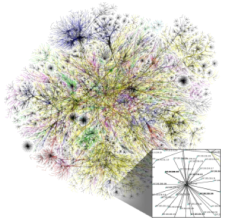


- **libcuDF**
  - High-performance C++ layer with GPU-optimized CUDA kernels, data types, operations, and primitives
- **cuDF**
  - Familiar pandas-like Python API
- **Dask-cuDF**
  - Multi-node, multi-GPU scaling for Dask DataFrames and SQL with easy deployment for Python users
- **Spark-RAPIDS**
  - Zero-code-change Apache Spark and Apache Spark SQL acceleration

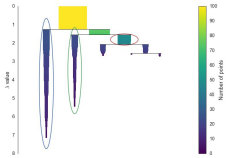
Algorithms

Accelerated Data Science





dmlc XGBoost



- **cuML**
  - Classification / Regression
  - Clustering / Dimensionality Reduction
  - Time Series
  - Decision Trees / Random Forests
- **cuGraph**
  - Sampling
  - Tree / Structure
  - Community
- **XGBoost**
  - “XGBoost is all you need”
- **HDBSCAN**
  - Soft clustering
  - Approximate predict

RAPIDS Everywhere

Integrating where data science is done

- RAPIDS now has over 100 integrations, covering both open source and commercial software
- Easily accelerate the AutoML, interpretability, plotting or domain-specific modeling tools of your choice
- Both RAPIDS-sponsored and community integrations growing every day
- For projects already using Pandas / Scikit-learn / NetworkX, integration is often a few lines of code
- Have you integrated RAPIDS into a project we haven't mentioned? Join our Slack and tell us about it! (https://rapids.ai/slack-invite)

Expanding Access to RAPIDS

New platforms, new communities

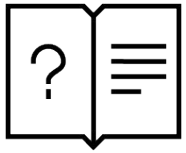
- Windows Subsystem for Linux (WSL) support is now GA

- ARM SBSA support is now GA

- RAPIDS can now be installed via Pip, Python's standard package manager (Experimental)

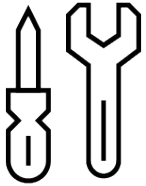
How to Get Started with RAPIDS

A Variety of Ways to Get Up & Running



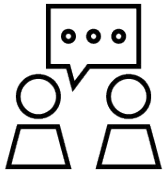
More about RAPIDS

- 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 SageMaker Studio Lab
- Look at the cheat sheets



Discussion & Support

- Check the RAPIDS GitHub
- Use the NVIDIA Forums
- Reach out on Slack
- Talk to NVIDIA Services

Get Engaged

@RAPIDSai

https://github.com/rapidsai

https://rapids-goai.slack.com/join

https://rapids.ai