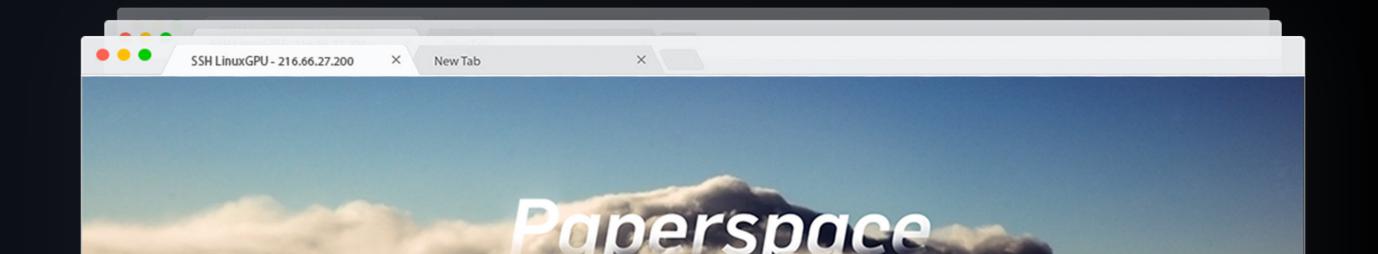
Paperspace

www.paperspace.com



Paperspace

Serverless Al for the future of intelligence.



Introduction

Deep Learning platform built for developers.

Infrastructure automation and software layer to build intelligent applications.



A new generation of Al developers require a rethinking of tooling and workflows.

Why this matters

Developers spend 75% of their time managing infrastructure.

So what's the underlying problem?

The cloud was built for a different use-case (web servers) and a different audience (DevOps).

The DevOps ecosystem is rich

Traditional Web Services













+ 100s more

storage, CDN, deploy, monitor, VPC, load balance, IPsec, CI/CD, DNS ...

Deep Learning



data, notebooks, train, visualize, collaborate, version, hyperparameters ...

The key to solving this problem is finding the right layer of abstraction.

Put Uber/Facebook-grade Al platform in the hands of every developer

There is a huge disconnect between modern business objectives and the DL tools that can fulfill them.

Business Objective



Infrastructure

- BI
- Prediction
- Optimization
- Recommender systems
- Any heuristic

Closing the Gap

- GPUs
- Datastore
- Algorithms (CNNs, RNNs, ...)
- Frameworks (Pytorch, TensorFlow, etc)

Paperspace abstracts powerful infrastructure behind a simple software layer making cloud ML as easy as modern web services.

A complete platform for modern deep learning

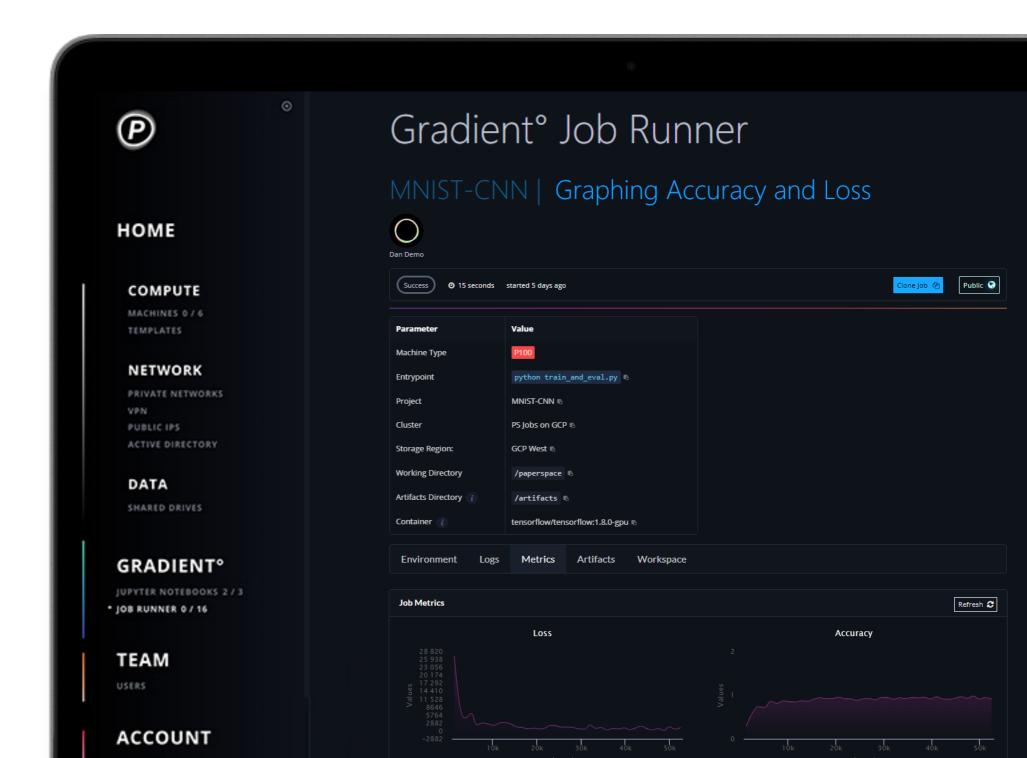
Ingest → Train →

Analyze → Deploy

٠

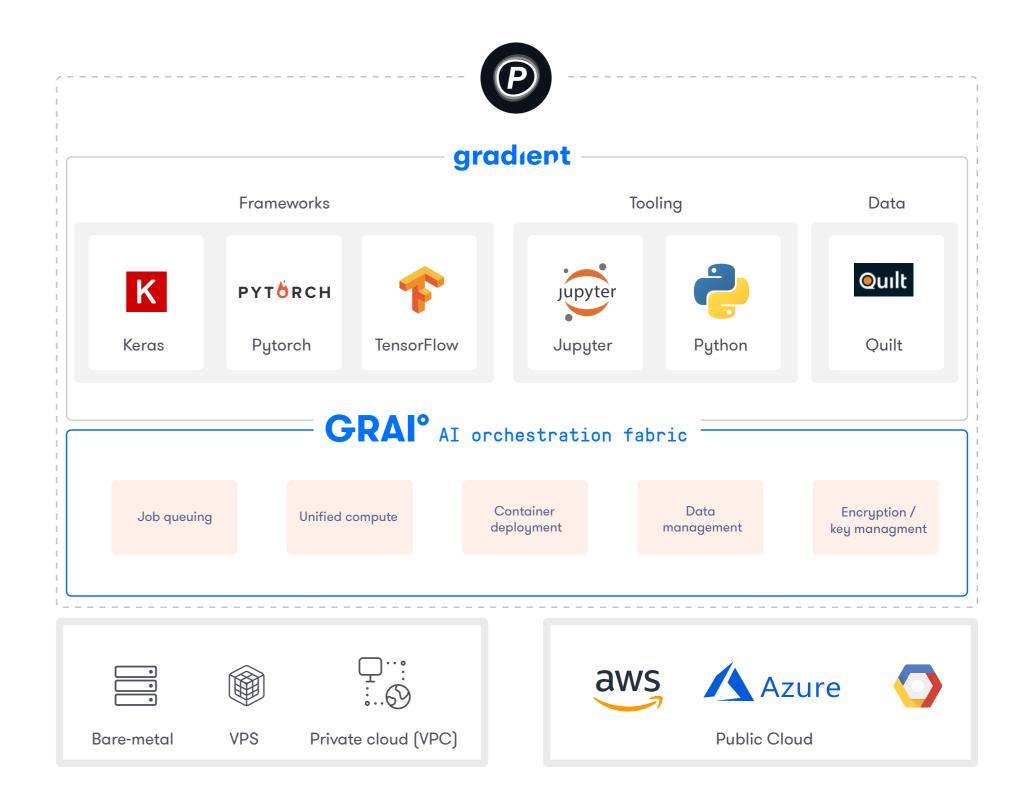
Manage, collaborate, share

- Fully-managed GPU infrastructure
- Unified dev experience
- 1 click Jupyter Notebooks/Lab
- API & language integrations
- ACL/team controls



GRAI° Model building AI orchestration fabric

- Job queuing / management
- Cloud agnostic
- Accelerator architecture native
 (GPU, FPGA, ASIC, TPU, etc)
- Unified compute
- Extensible
- Built on best practices
 (containers, kubernetes, and data policies)

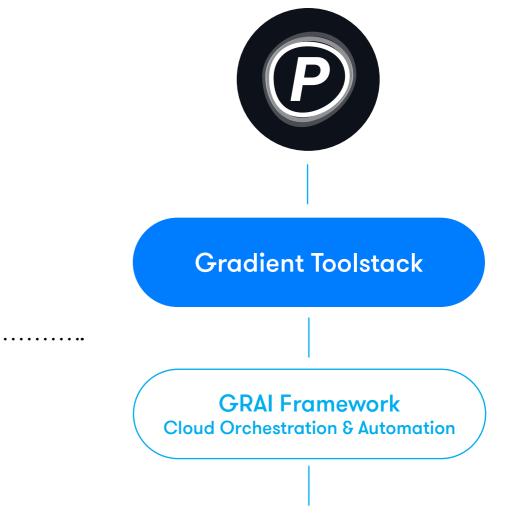


Cloudscale with a single line of code

```
> import paperspace as ps
# Run job on GPU cluster
> ps({'Type': 'TPU', 'container': 'TensorFlow' ... })
```

Connecting modern ML and the cloud by converting infrastructure into code.

Raw compute is not sufficient.



Cloud Infra



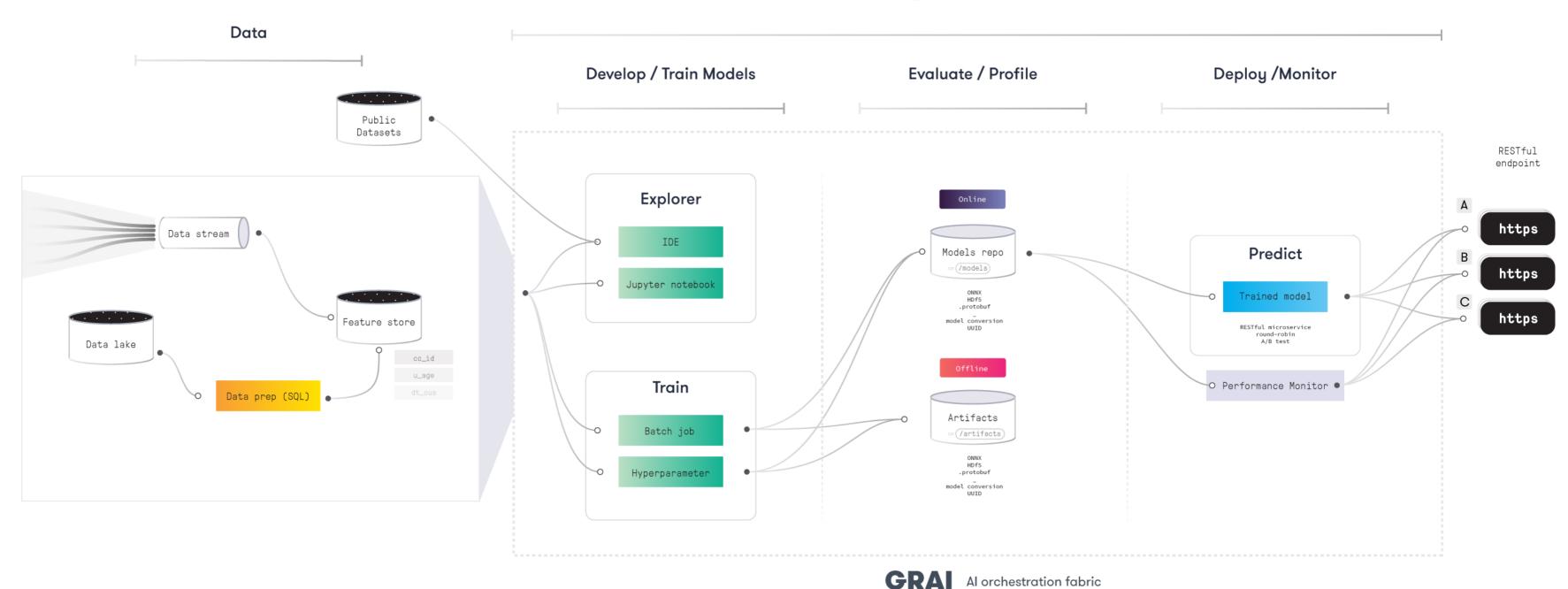




Network . Storage . Compute

The Pipeline

gradient



Remarks from the trenches

Trends:

- 1. Chip renaissance
- 2. Evolution of ML/Al in practice
- 3. Consolidation around best practices

Chip renaissance

- Graphcore
- Cerebras
- Nervana
- Wave
- Google TPU

. . .

The big question today is whether accelerator architectures will follow commodity CPU x86 or lead to a golden era for high-end, use-specific hardware.



The evolution of ML/AI in practice

2016 2018

Consumable API \rightarrow Refit the Model \rightarrow Model as core IP

- Clarifai
- AWS Rekognition
- Google Cloud Vision
- MS cognitive services

- Paperspace
- Algorithmia
- FloydHub
- ClusterOne

Consolidation around best practices

- Containerization
- Jupyter
- Job runner architecture
- Pipeline
- etc.

Jupyter Notebook Example



Run some Python code!

To run the code below:

- 1. Click on the cell to select it.
- Press SHIFT+ENTER on your keyboard or press the play button (►) in the toolbar above.

A full tutorial for using the notebook interface is available here.

Feel free to open new cells using the plus button (+), or hitting shift-enter while this cell is selected.

Behind the scenes, the software that powers this is <u>tmpnb</u>, a Tornado application that spawns <u>pre-built Docker containers</u> and then uses the <u>jupyter/configurable-http-proxy</u> to put your notebook server on a unique path.

Feel free to open new cells using the plus button (+), or hitting shift-enter while this cell is selected.

Behind the scenes, the software that powers this is <u>tmpnb</u>, a Tornado application that spawns <u>pre-built Docker containers</u> and then uses the <u>jupyter/configurable-http-proxy</u> to put your notebook server on a unique path.

hello@paperspace.com www.paperspace.com (718) 619 4325

Thank you.

