



# Data exploration at the speed of thought

Lessons learned from inside Google

Nico Gaviola

Head of Healthcare and Lifesciences UKIE

[nicogaviola@google.com](mailto:nicogaviola@google.com)

 Google Cloud Platform



Google's mission is to organize the world's information and make it universally accessible and useful.



Sundar Pichai  
CEO, Google

The image shows a large server room with rows of server racks. The Google logo is superimposed in the center. Below the logo is a white search bar, and below that are two buttons: "Google Search" and "I'm Feeling Lucky".

# Google

Google Search

I'm Feeling Lucky

# Google computing scale



*uploads per minute*



**500hrs**



*users*



**1B+**



*search index*



**100PB+**



*query response time*



**0.25s**

# Hitting the limits, early on...



**The Anatomy of a Large-Scale Hypertextual Web Search Engine**  
Sergey Brin and Lawrence Page  
Computer Science Department, Stanford University, Stanford, CA 94305

**Abstract**  
This paper describes the design and implementation of a large-scale, distributed, web search engine. The engine is designed to handle the large volume of data and the high rate of change in the web. It is implemented as a distributed system, with the search engine and the database spread across multiple machines. The system is designed to be scalable and to handle the large volume of data and the high rate of change in the web. The system is implemented as a distributed system, with the search engine and the database spread across multiple machines. The system is designed to be scalable and to handle the large volume of data and the high rate of change in the web.

**1. Introduction**  
The web is a vast and ever-changing source of information. It is a source of information that is constantly growing and changing. The web is a source of information that is constantly growing and changing. The web is a source of information that is constantly growing and changing. The web is a source of information that is constantly growing and changing.

The Anatomy of a Large-Scale Hypertextual Web Search Engine

1996, Sergey Brin and Lawrence Page  
Computer Science Department, Stanford University, Stanford,  
CA 94305

# Single Node to Cluster



GFS

2002



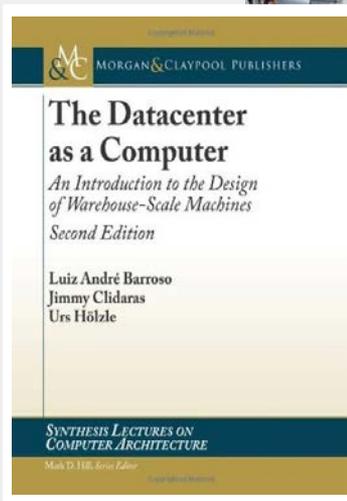
MapReduce

2004



BigTable

2006

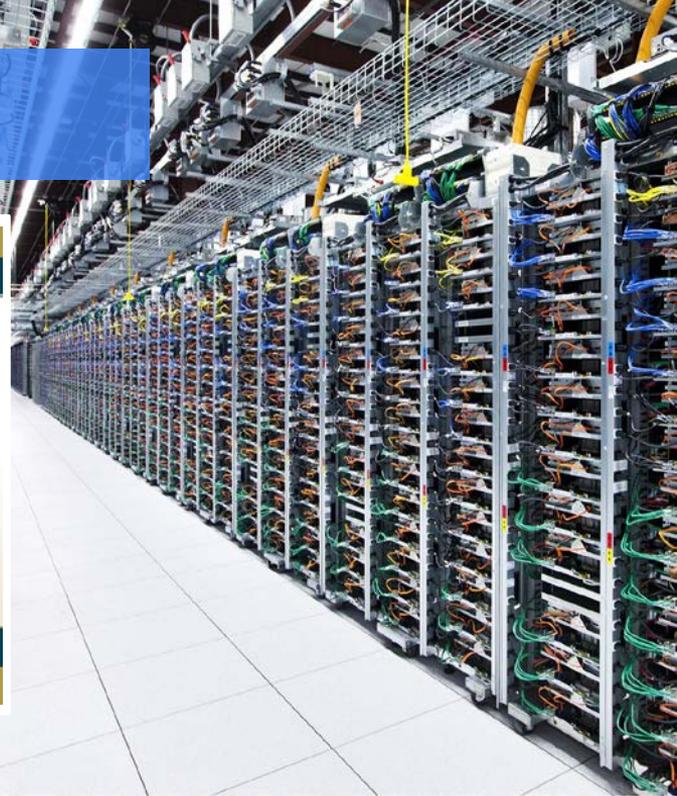


2008

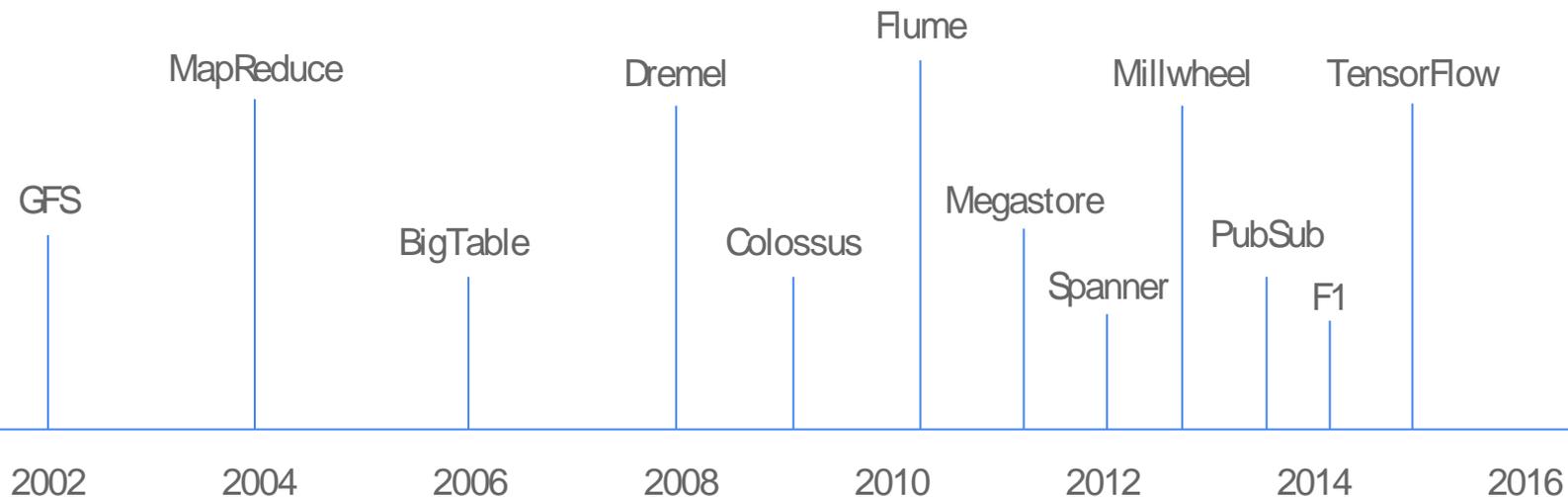
2010

2012

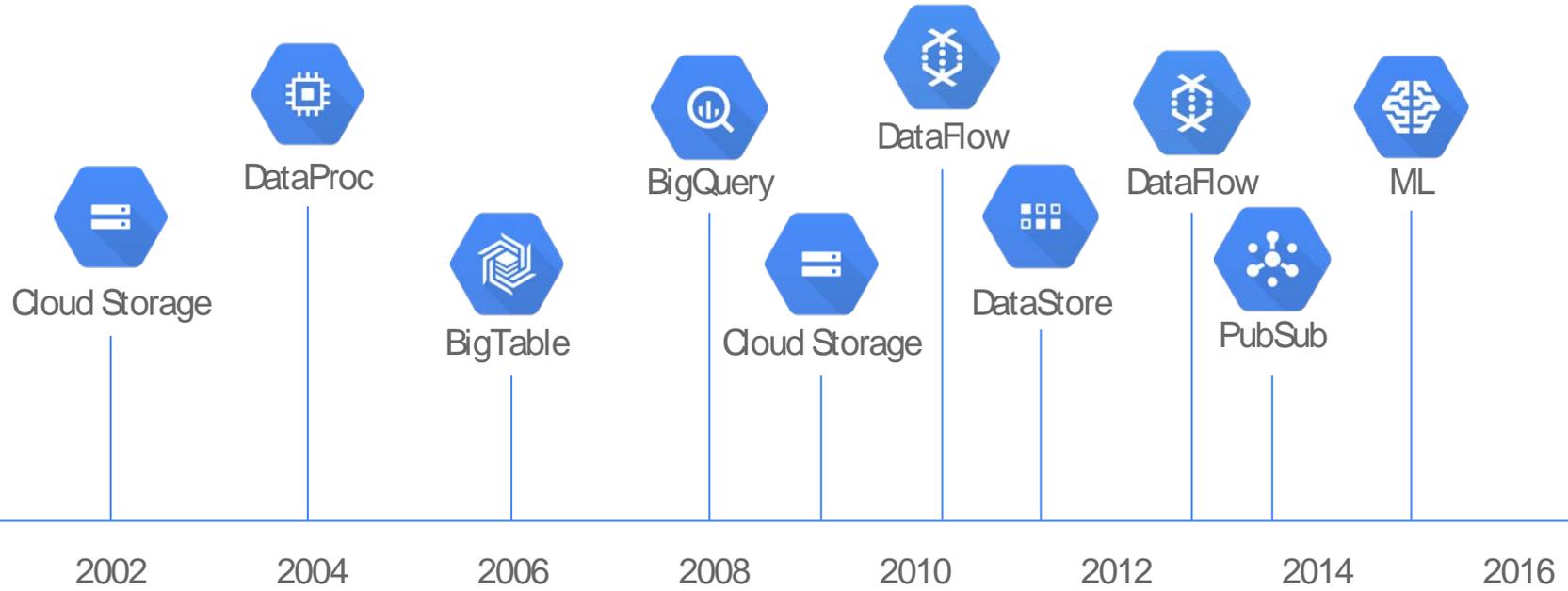
2013



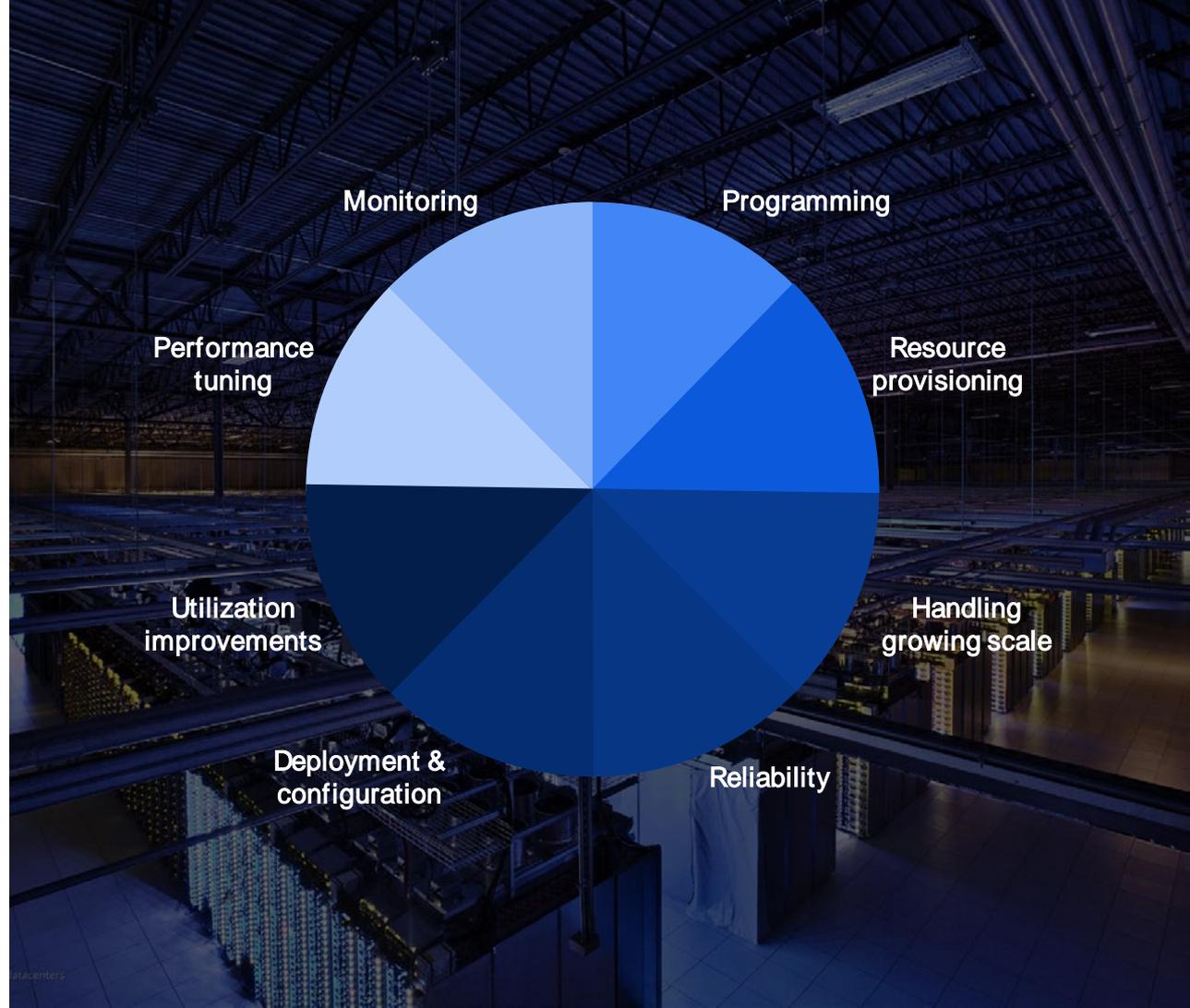
# Google's Data Research



# Google's Data Products



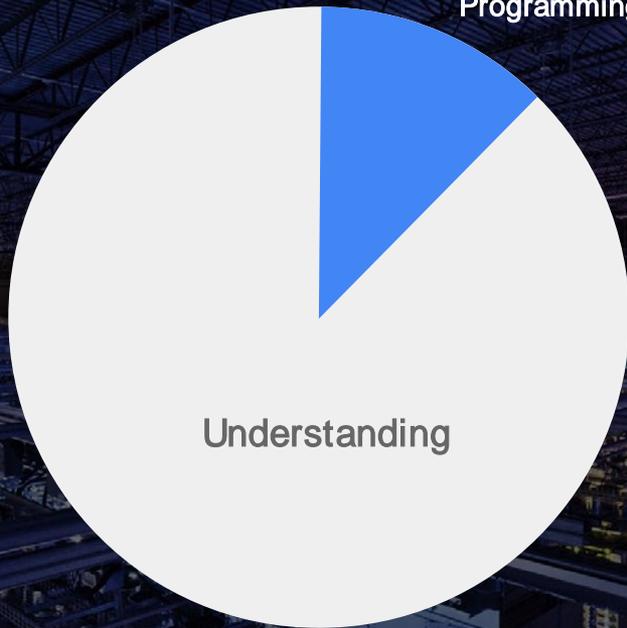
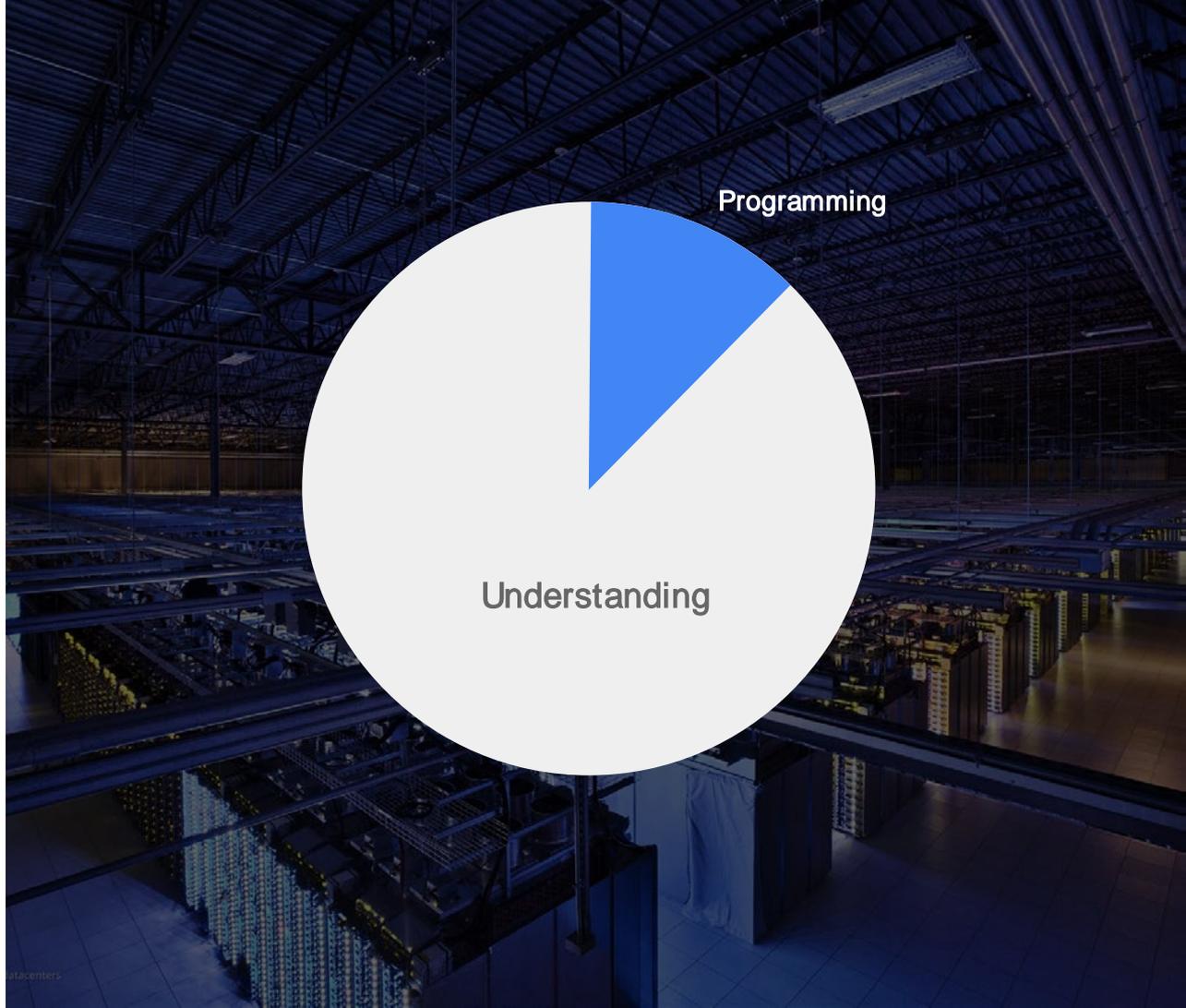
# Typical Big Data Jobs





# Big Data with Google

Focus on insights.  
Not infrastructure.





# Google's Big Data Vision

**RUN QUERY**

Pay \$5 per TB

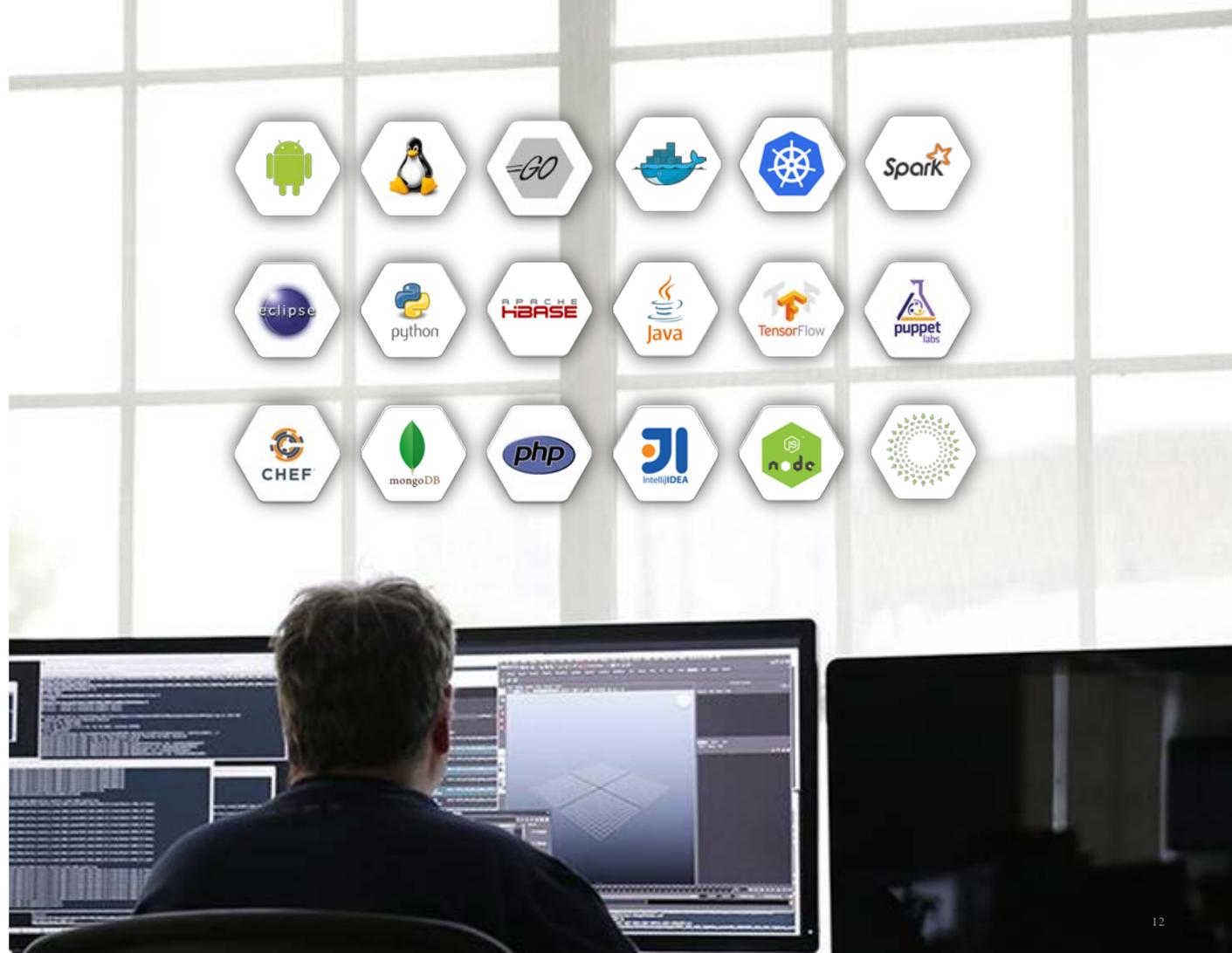


# Open Source & APIs

Active contributor to numerous OSS projects

Make migrations easier with open APIs

Customers should use us because they love us, not because they are unable to move off



# Google Security Model & You!



You own your data and remain Data Controller



You can delete or remove your data at any time



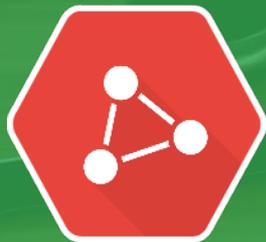
Google does not share your content or personal information  
[google.com/privacy](https://google.com/privacy)



Strict Internal Policies :  
all accesses to customer or consumer data applications are logged



Internal data access auditing tracks Googlers



Example

# LLOYDS BANKING GROUP



“Right at the start of the partnership we were able to **reduce time to insight from 96 hours to 30 minutes** by using BigQuery”

Gary Sanders  
Head of Digital Analytics

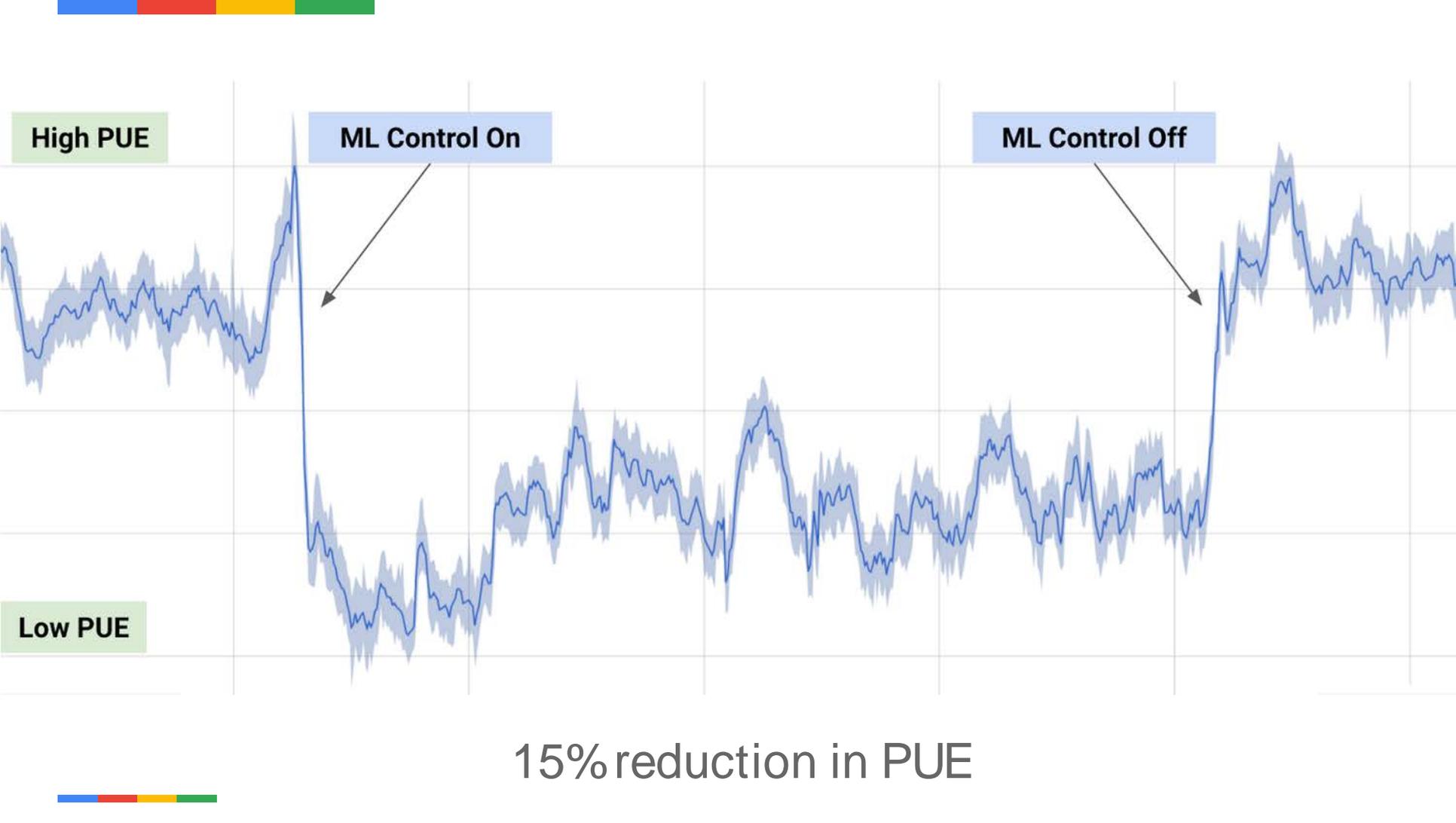
# What's Next?



“Machine learning is a core,  
transformative way by which we’re re-  
thinking how we’re doing everything”

Sundar Pichai  
CEO, Google





High PUE

ML Control On

ML Control Off

Low PUE

15% reduction in PUE



# Fully trained, easy to use Machine Learning models



Cloud  
Translate



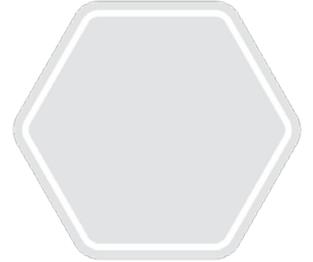
Cloud  
Vision



Cloud  
Speech



Cloud  
Natural Language



Stay tuned...



# Use your own data to train models



Cloud Storage



BigQuery



Cloud Datalab



Cloud Machine Learning



Develop, Model,  
Train, Test

# One more thing



# Free training courses coming near you!

Region	Country	City	Start Date	Course	Language	Training Partner	Trainer	Registration	Admin	Capacity	Registered	Waitlisted
EMEA	United Kingdom	London	9/23/2016	CPB100: Google Cloud Platform Big Data & Machine Learning Fundamentals	English	Google	Lak Lakshmanan	<a href="#">REGISTER</a>	<a href="#">Admin</a>	70	<a href="#">70</a>	
EMEA	United Kingdom	London	10/18/2016	CPB100: Google Cloud Platform Big Data & Machine Learning Fundamentals	English	ROI Training	Grant Moyle	<a href="#">REGISTER</a>	<a href="#">Admin</a>	45	<a href="#">39</a>	
EMEA	United Kingdom	London	10/20/2016	CPB101: Serverless Data Analysis with BigQuery and Cloud Dataflow	English	ROI Training	Grant Moyle	<a href="#">REGISTER</a>	<a href="#">Admin</a>	45	<a href="#">25</a>	
EMEA	Netherlands	Amsterdam	10/26/2016	CPB100: Google Cloud Platform Big Data & Machine Learning Fundamentals	English	gCompany	Koen Maes	<a href="#">REGISTER</a>	<a href="#">Admin</a>	40	<a href="#">6</a>	
EMEA	Spain	Barcelona	10/27/2016	CPB100: Google Cloud Platform Big Data & Machine Learning Fundamentals	English	Extrema-sistemas	Jose L Ugia Gonzalez	<a href="#">REGISTER</a>	<a href="#">Admin</a>	30	<a href="#">8</a>	
EMEA	France	Paris	11/7/2016	CPB100: Google Cloud Platform Big Data & Machine Learning Fundamentals	French	SFEIR	Didier Girard	<a href="#">REGISTER</a>	<a href="#">Admin</a>	60	<a href="#">31</a>	
EMEA	Belgium	Brussels	11/17/2016	CPB100: Google Cloud Platform Big Data & Machine Learning Fundamentals	English	gCompany	Koen Maes	<a href="#">REGISTER</a>	<a href="#">Admin</a>	25	<a href="#">4</a>	

# Thank you!

