FLY ELEPHANT

Infrastructure for the work of Data Scientists

Dmitry Spodarets

Who am I

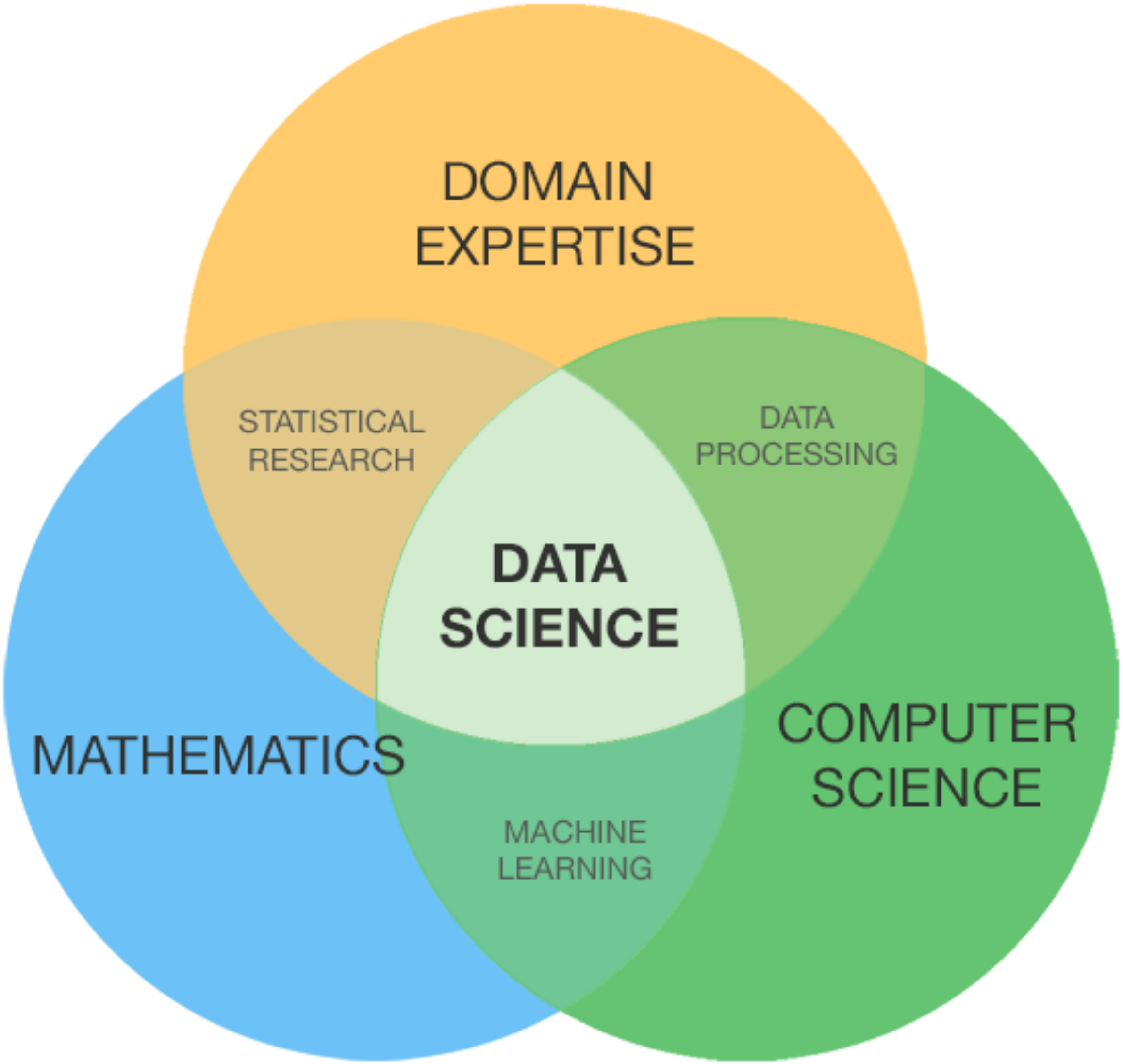
Dmitry Spodarets

- CEO and Founder at FlyElephant
- Lecturer at Odessa Polytechnic University
- Organizer of technical conferences about AI, BigData, HPC, JS, FOSS …

Agenda

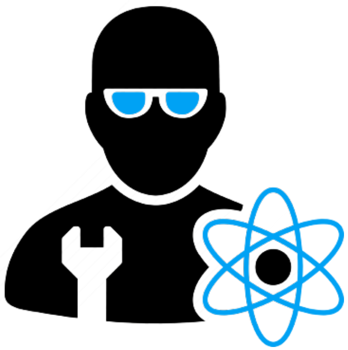
- In production
- Database
- Notebook / IDE
- Software
- Deep Learning Tools
- Programming Languages
- Visualization
- Computing power
- Architecture
- Docker
- Cloud Services
- FlyElephant

Data Science



Data Science: people

DevOps

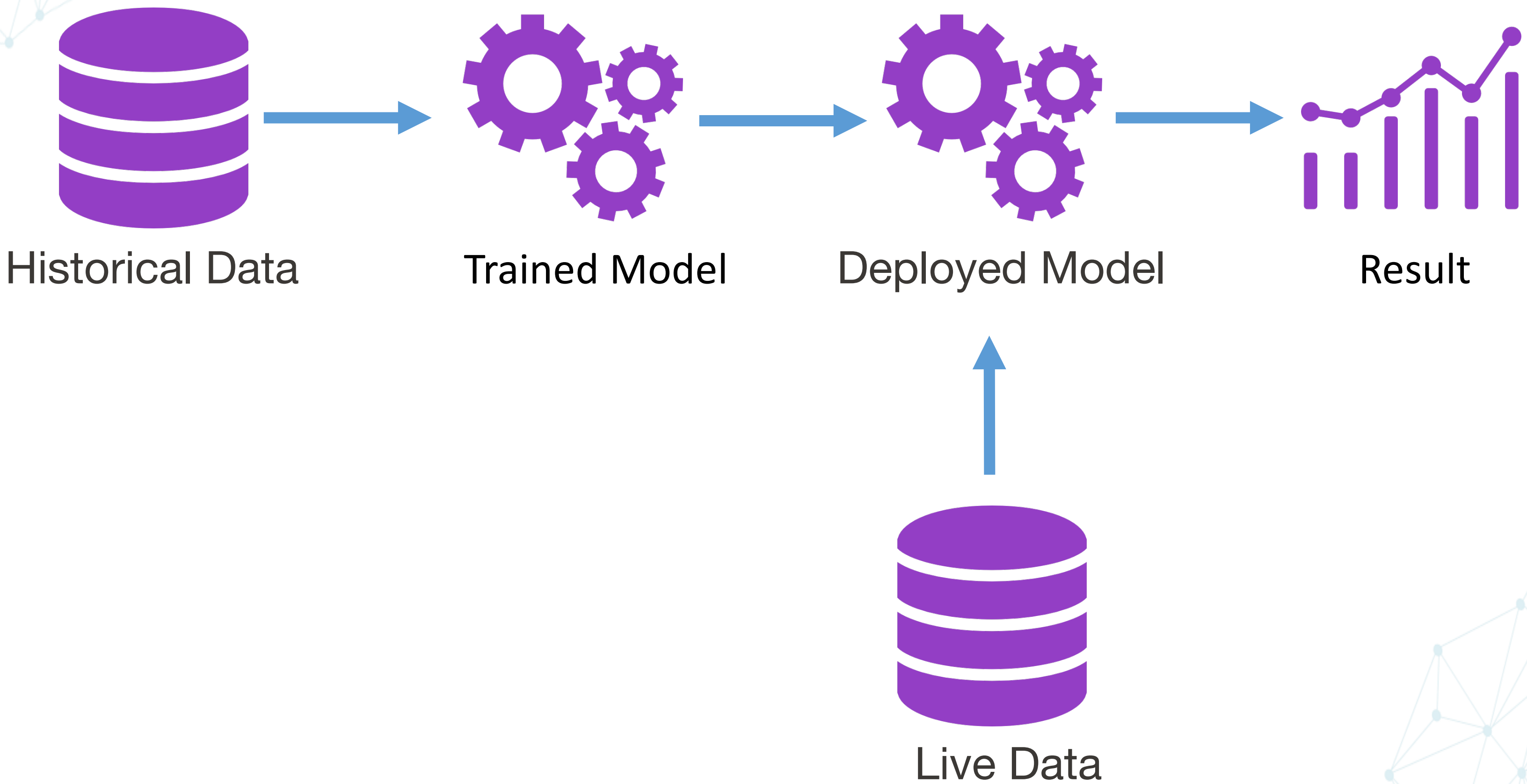


Engineer

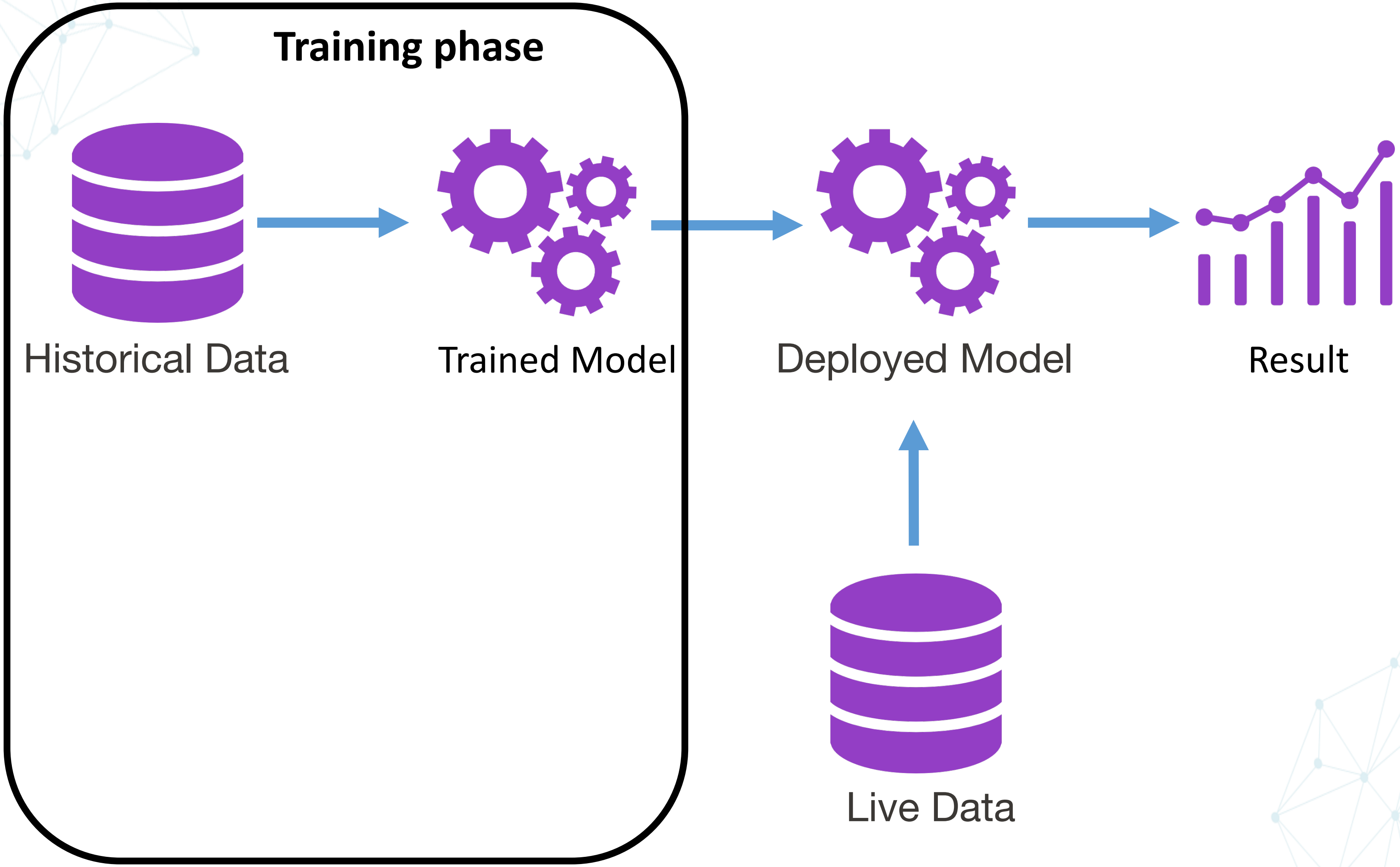
Data Scientist

Business

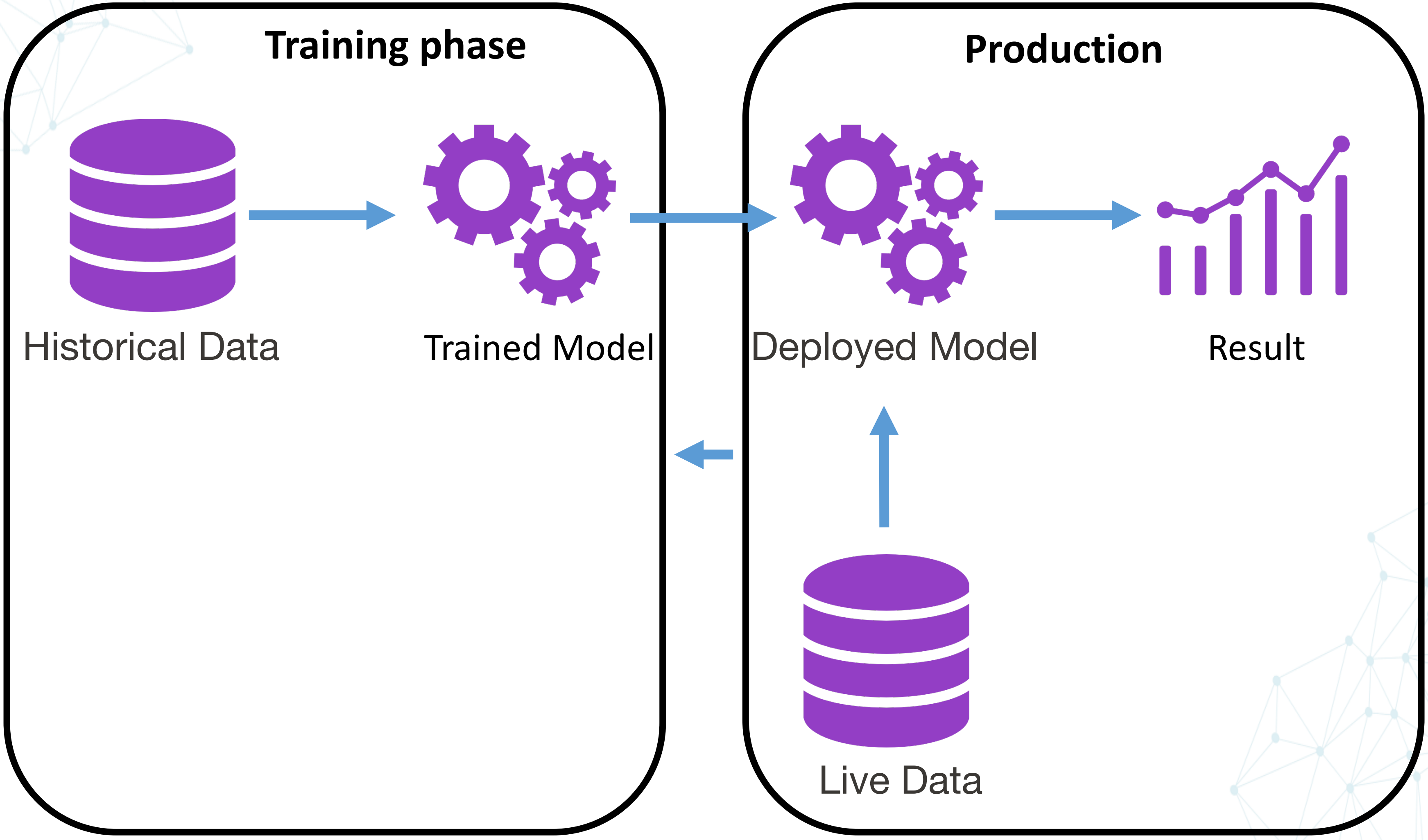
In production



In production



In production



Database

cassandra

APACHE HBASE

HIVE

HDFS

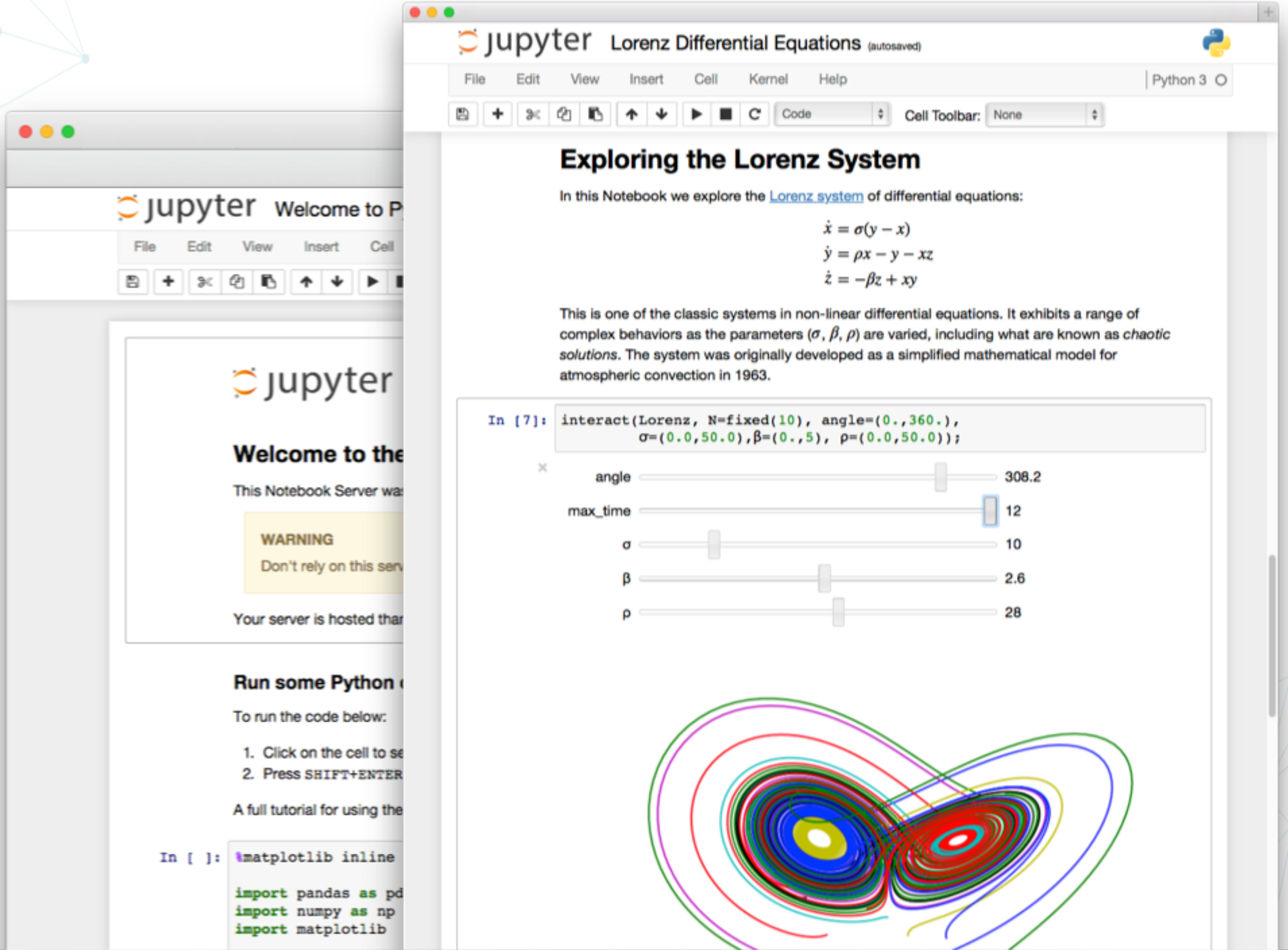
mongoDB®

MySQL®

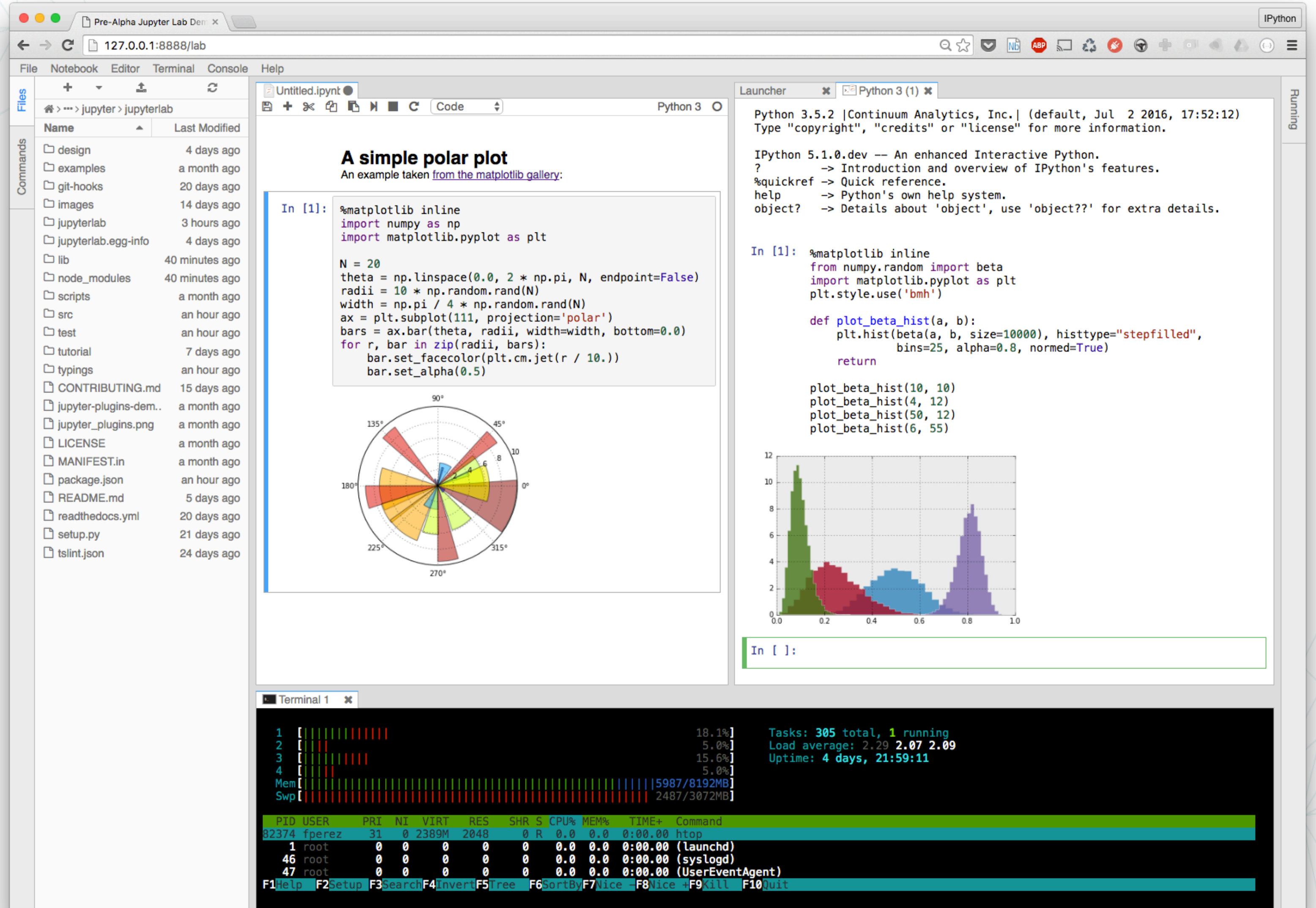
PostgreSQL

Notebook / IDE

Jupyter Notebook



Jupyter Lab



Pre-Alpha Jupyter Lab Dem
127.0.0.1:8888/lab
File Notebook Editor Terminal Console Help
Untitled.ipynb
Launcher
Python 3 (1)
A simple polar plot
An example taken from the matplotlib gallery:
In [1]:
%matplotlib inline
import numpy as np
import matplotlib.pyplot as plt
N = 20
theta = np.linspace(0.0, 2 * np.pi, N, endpoint=False)
radii = 10 * np.random.rand(N)
width = np.pi / 4 * np.random.rand(N)
ax = plt.subplot(111, projection='polar')
bars = ax.bar(theta, radii, width=width, bottom=0.0)
for r, bar in zip(radii, bars):
bar.set_facecolor(plt.cm.jet(r / 10.))
bar.set_alpha(0.5)
Python 3.5.2 |Continuum Analytics, Inc.| (default, Jul 2 2016, 17:52:12)
Type "copyright", "credits" or "license" for more information.
IPython 5.1.0.dev -- An enhanced Interactive Python.
%matplotlib inline
from numpy.random import beta
import matplotlib.pyplot as plt
plt.style.use('bmh')
def plot_beta_hist(a, b):
plt.hist(beta(a, b, size=10000), histtype="stepfilled", bins=25, alpha=0.8, normed=True)
return
plot_beta_hist(10, 10)
plot_beta_hist(4, 12)
plot_beta_hist(50, 12)
plot_beta_hist(6, 55)
Terminal 1
Tasks: 305 total, 1 running
Load average: 2.29 2.07 2.09
Uptime: 4 days, 21:59:11

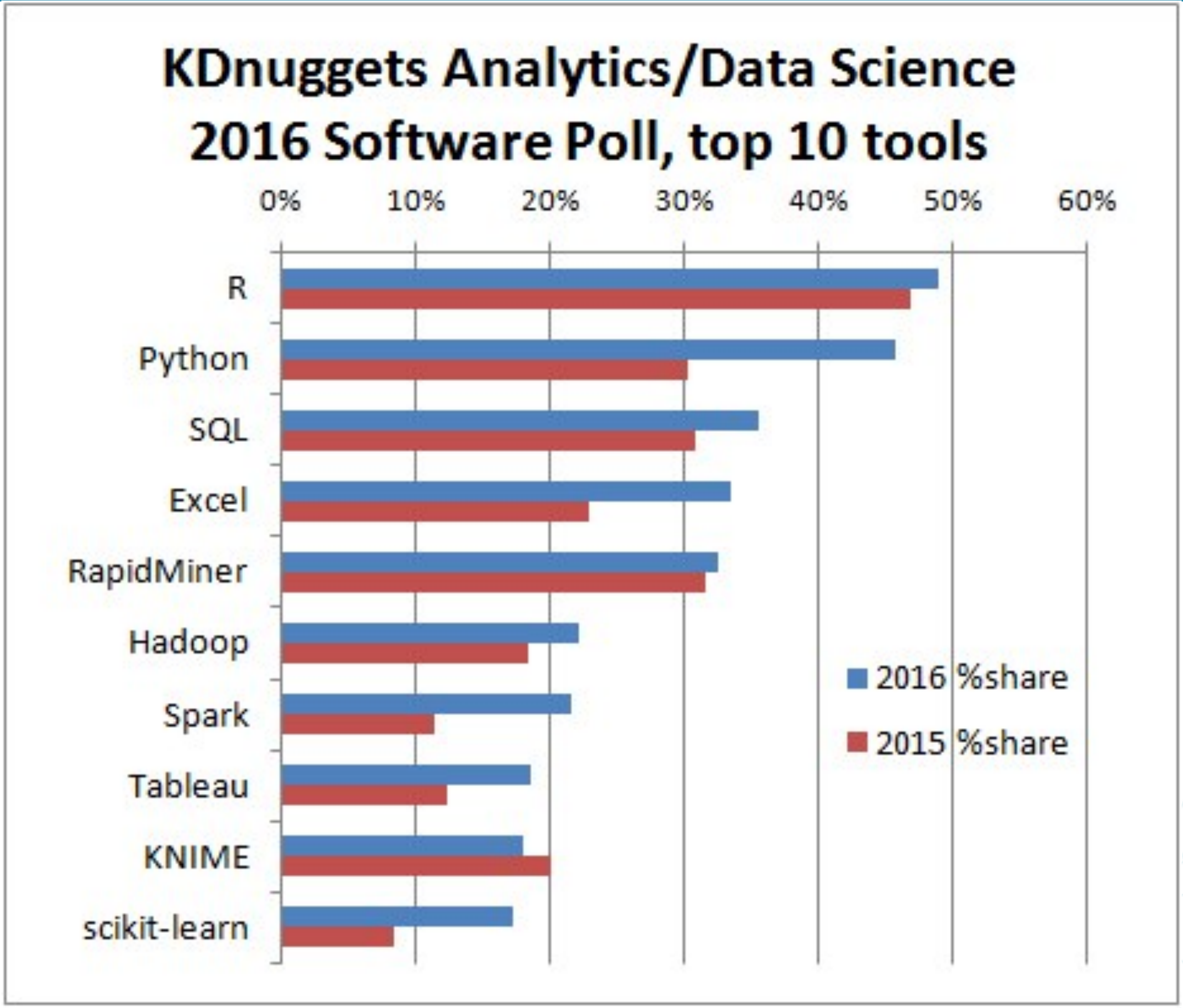
Jupyter Lab

- JupyterLab is the natural evolution of the Jupyter Notebook user interface
- JupyterLab is an IDE: *Interactive* Development Environment
- Flexible user interface for assembling the fundamental building blocks of interactive computing
- Modernized JavaScript architecture based on npm/webpack, plugin system, model/view separation
- Built using PhosphorJS (http://phosphorjs.github.io/)
- Design-driven development process

https://github.com/jupyter/jupyterlab

http://blog.jupyter.org/2016/07/14/jupyter-lab-alpha/

Software



http://www.kdnuggets.com/2016/06/r-python-top-analytics-data-mining-data-science-software.html

Software

New (in this poll) tools that received at least 1% share votes in 2016 were:

- **Anaconda**, 16%
- **Microsoft other ML/Data Science tools**, 1.6%
- **SAP HANA**, 1.2%
- **XLMiner**, 1.2%

http://www.kdnuggets.com/2016/06/r-python-top-analytics-data-mining-data-science-software.html

Software

- Tools with the highest growth (among tools with at least 15 users in 2015) were

Tool	% change	2016 %share	2015 %share
Dato (turi)	377%	2.4%	0.5%
Dataiku	292%	7.8%	2.0%
MLlib	253%	11.6%	3.3%
H2O	233%	6.7%	2.0%
Amazon Machine Learning	171%	1.9%	0.7%
scikit-learn	107%	17.2%	8.3%
IBM Watson	99%	4.2%	2.1%
Splunk/ Hunk	98%	2.2%	1.1%
Spark	91%	21.6%	11.3%
Scala	79%	6.2%	3.5%

http://www.kdnuggets.com/2016/06/r-python-top-analytics-data-mining-data-science-software.html

Deep Learning Tools

- **Tensorflow**, 6.8%
- **Theano ecosystem (including Pylearn2)**, 5.1%
- **Caffe**, 2.3%
- **MATLAB Deep Learning Toolbox**, 2.0%
- **Deeplearning4j**, 1.7%
- **Torch**, 1.0%
- **Microsoft CNTK**, 0.9%
- **Cuda-convnet**, 0.8%
- **mxnet**, 0.6%
- **Convnet.js**, 0.3%
- **darch**, 0.1%
- **Nervana**, 0.1%
- **Veles**, 0.1%
- **Other Deep Learning Tools**, 3.7%

http://www.kdnuggets.com/2016/06/r-python-top-analytics-data-mining-data-science-software.html

Programming Languages

- **R**, 49.0 % share (was 46.9), 4.5% increase
- **Python**, 45.8% share (was 30.3%), 51% increase
- **Java**, 16.8% share (was 14.1%), 19% increase
- **Unix shell/awk/gawk** 10.4% share (was 8.0%), 30% increase
- **C/C++**, 7.3% share (was 9.4%), 23% decrease
- **Other programming/data languages**, 6.8% share (was 5.1%), 34.1% increase
- **Scala**, 6.2% share (was 3.5%), 79% increase
- **Perl**, 2.3% share (was 2.9%), 19% decrease
- **Julia**, 1.1% share (was 1.1%), 1.6% decrease
- **F#**, 0.4% share (was 0.7%), 41.8% decrease
- **Clojure**, 0.4% share (was 0.5%), 19.4% decrease
- **Lisp**, 0.2% share (was 0.4%), 33.3% decrease

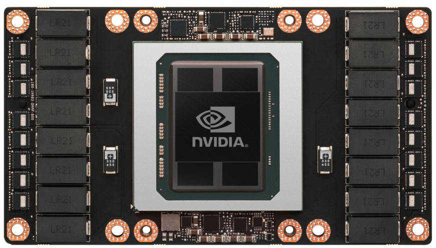
http://www.kdnuggets.com/2016/06/r-python-top-analytics-data-mining-data-science-software.html

Visualization

- **Tableau**
- **Zoomdata**
- **Qlik**
- **Plotly**
- **Matplotlib 2.0** (for Python)
- **Ggplot2** (for R)
- **D3.js**

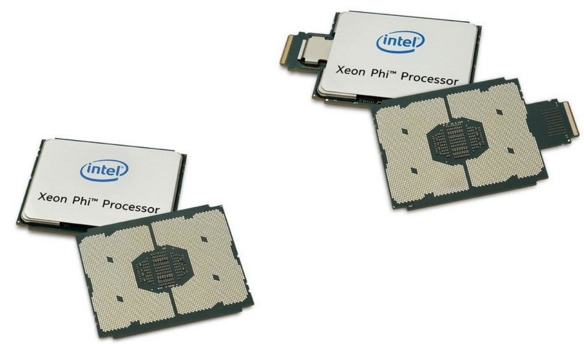
Computing power

Computing power



nvidia tesla p100
~5 TeraFLOPS

NVIDIA DGX-1 Deep Learning Supercomputer
170/3 TFLOPS (GPU FP16 / CPU FP32)



intel xeon phi processor
~3 TeraFLOPS

Computing power (NVIDIA)

ПРОДУКТЫ TESLA ДЛЯ ЛЮБЫХ ЗАДАЧ

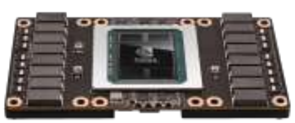
СУПЕРКОМПЬЮТЕР ДЛЯ DEEP LEARNING

DGX-1

Полностью интегрированное решение для DL

МАСШТАБИРУЕМЫЕ ВЫЧИСЛЕНИЯ

Tesla P100 SXM2



Вычислительные центры для приложений которые хорошо масштабируются на GPU

СМЕШАННЫЕ ВЫЧИСЛЕНИЯ

Tesla P100 PCIE

Вычислительные центры имеющие CPU и GPU приложения

HYPERSCALE ВЫЧИЛСЕНИЯ

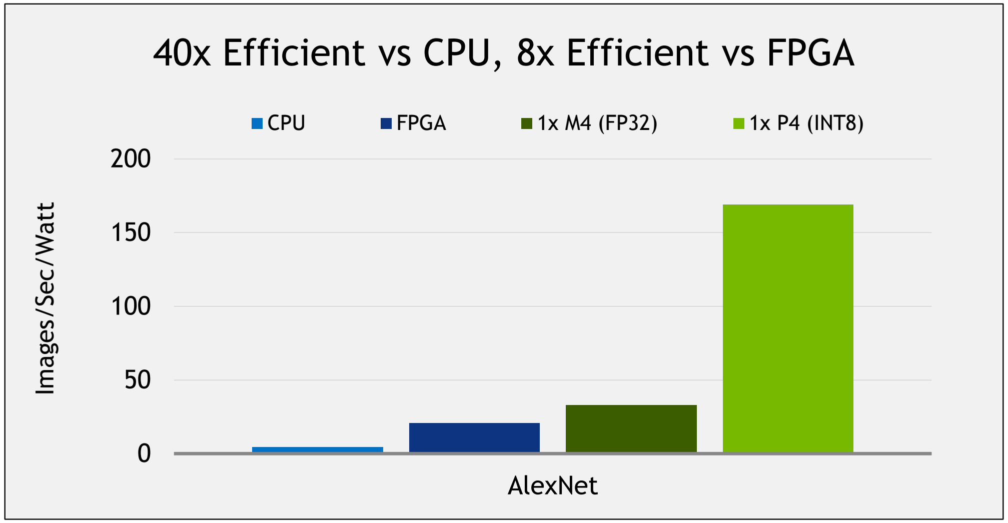
Tesla P4, P40

Создание DL-сервисов, обработка видео и изображений, обучение нейросетей

Computing power (NVIDIA)

TESLA P4

Максимальная эффективность для масштабируемых серверов



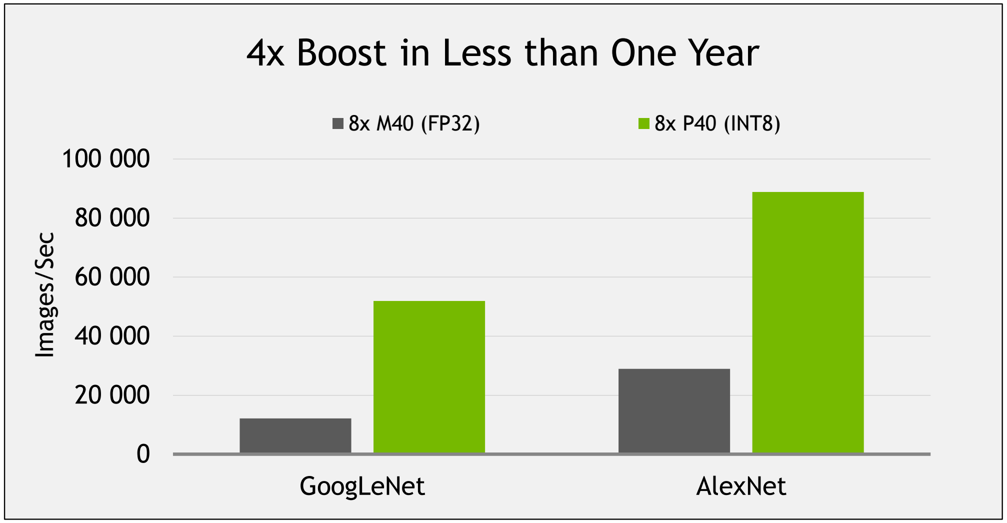
P4	
# of CUDA Cores	2560
Peak Single Precision	5.5 TeraFLOPS
Peak INT8	22 TOPS
Low Precision	4x 8-bit vector dot product with 32-bit accumulate
Video Engines	1x decode engine, 2x encode engine
GDDR5 Memory	8 GB @ 192 GB/s
Power	50W & 75 W

AlexNet, batch size = 128, CPU: Intel E5-2690v4 using Intel MKL 2017, FPGA is Arria10-115
1x M4/P4 in node, P4 board power at 56W, P4 GPU power at 36W, M4 board power at 57W, M4 GPU power at 39W, Perf/W chart using GPU power

Computing power (NVIDIA)

TESLA P40

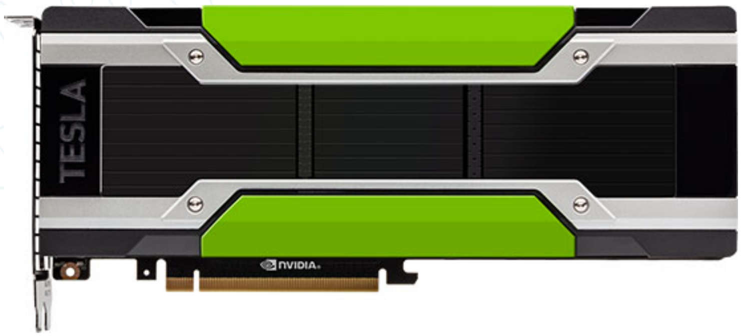
Максимальная пропускная способность для масштабируемых серверов

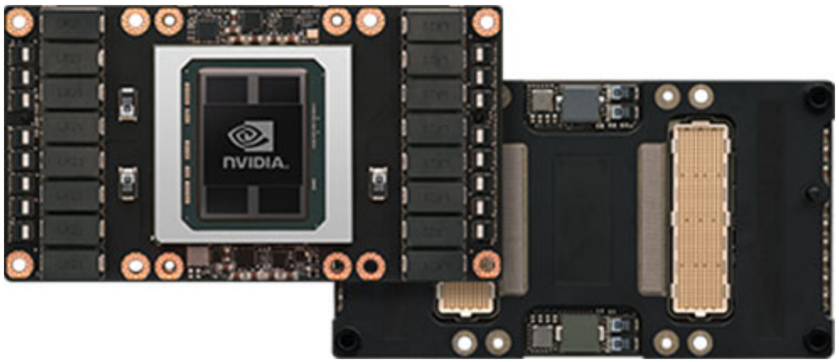


	P40
# of CUDA Cores	3840
Peak Single Precision	12 TeraFLOPS
Peak INT8	47 TOPS
Low Precision	4x 8-bit vector dot product with 32-bit accumulate
Video Engines	1x decode engine, 2x encode engines
GDDR5 Memory	24 GB @ 346 GB/s
Power	250W

GoogLeNet, AlexNet, batch size = 128, CPU: Dual Socket Intel E5-2697v4

Computing power (NVIDIA)





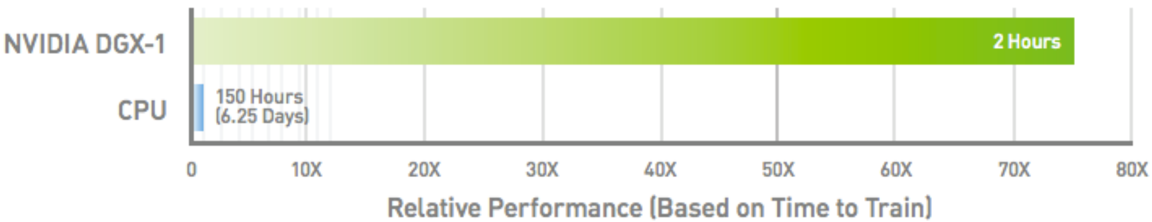
	P100 for PCIe-Based Servers	P100 for NVLink-Optimised Servers
Double-Precision Performance	4.7 TeraFLOPS	5.3 TeraFLOPS
Single-Precision Performance	9.3 TeraFLOPS	10.6 TeraFLOPS
Half-Precision Performance	18.7 TeraFLOPS	21.2 TeraFLOPS
NVIDIA NVLink™ Interconnect Bandwidth	-	160 GB/s
PCIe x16 Interconnect Bandwidth	32 GB/s	32 GB/s
CoWoS HBM2 Stacked Memory Capacity	16 GB or 12 GB	16 GB
CoWoS HBM2 Stacked Memory Bandwidth	720 GB/s or 540 GB/s	720 GB/s
Enhanced Programmability with Page Migration Engine	✓	✓
ECC Protection for Reliability	✓	✓
Server-Optimised for Data Centre Deployment	✓	✓

Computing power (NVIDIA)

SYSTEM SPECIFICATIONS

GPUs	8x Tesla P100
TFLOPS (GPU FP16 / CPU FP32)	170/3
GPU Memory	16 GB per GPU
CPU	Dual 20-core Intel® Xeon® E5-2698 v4 2.2 GHz
NVIDIA CUDA® Cores	28672
System Memory	512 GB 2133 MHz DDR4
Storage	4x 1.92 TB SSD RAID 0
Network	Dual 10 GbE, 4 IB EDR
Software	Ubuntu Server Linux OS DGX-1 Recommended GPU Driver
System Weight	134 lbs
System Dimensions	866 D x 444 W x 131 H (mm)
Packing Dimensions	1180 D x 730 W x 284 H (mm)
Maximum Power Requirements	3200W
Operating Temperature Range	10 - 30°C

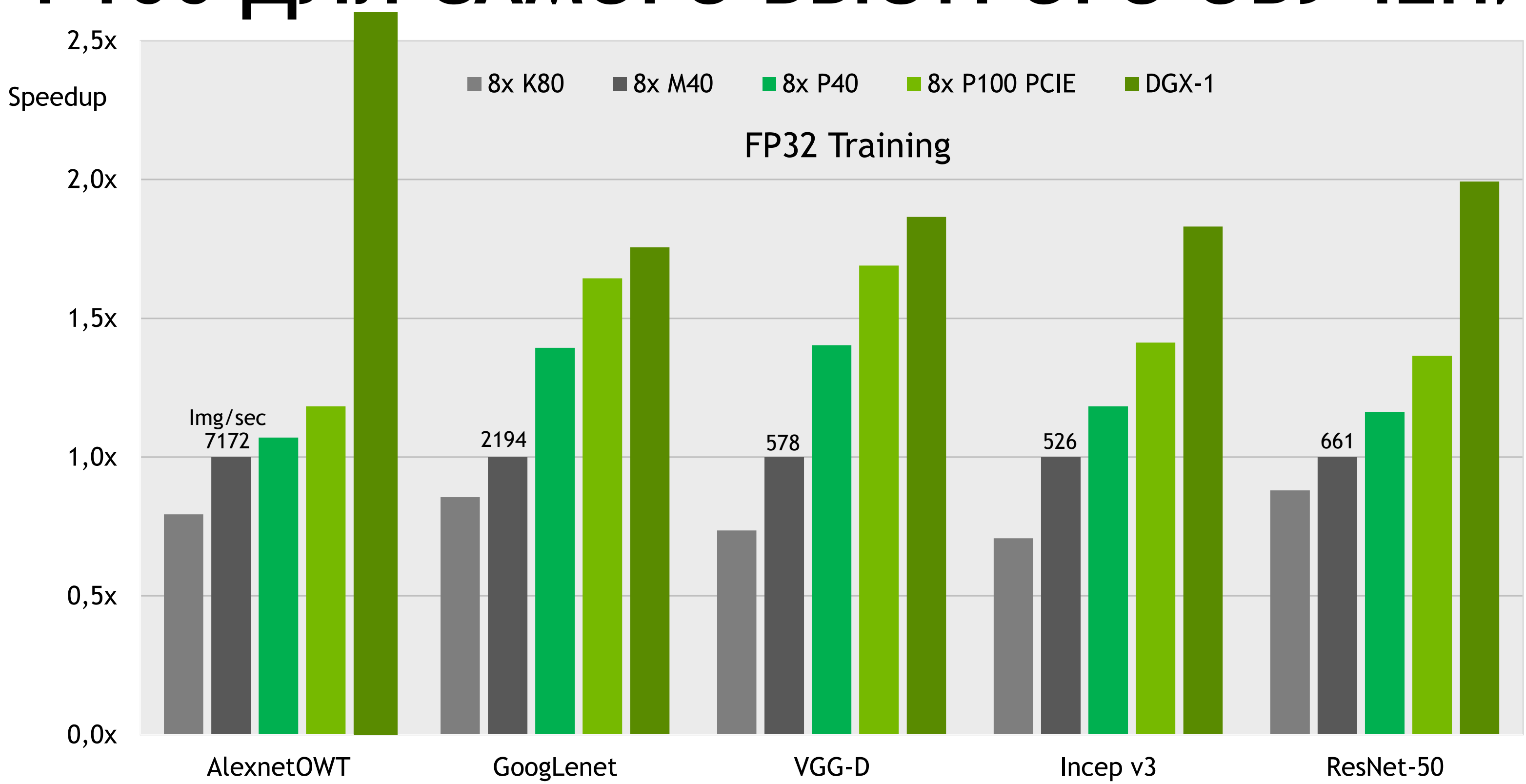
NVIDIA DGX-1 Delivers 75X Faster Training



Note: Caffe benchmark with AlexNet, training 1.28M images with 90 epochs | CPU server uses 2x Xeon E5-2697 v3 CPUs.

Computing power (NVIDIA)

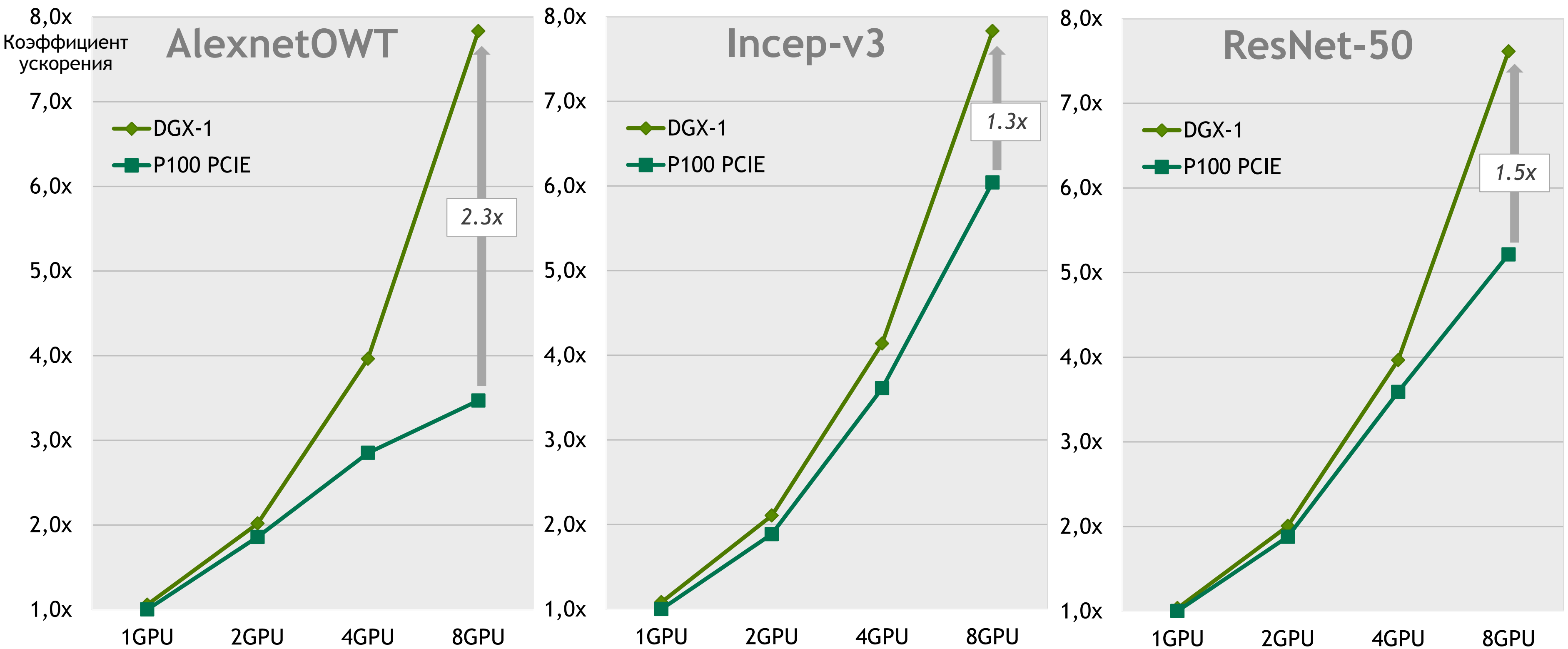
P100 ДЛЯ САМОГО БЫСТРОГО ОБУЧЕНИЯ



Deepmark test with NVCaffe. AlexnetOWT/GoogLenet use batch 128, VGG-D uses batch 64, Incep-v3/ResNet-50 use batch 32, weak scaling
K80/M40/P100/DGX-1 are measured, P40 is projected, software optimization in progress, CUDA8/cuDNN5.1, Ubuntu 14.04

Computing power (NVIDIA)

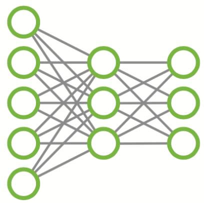
NVLINK: ЛИНЕЙНОЕ МАСШТАБИРОВАНИЕ



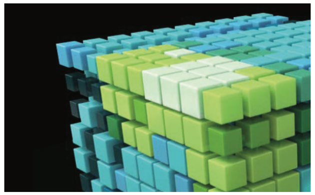
Deepmark test with NVCaffe. AlexnetOWT use batch 128, Incep-v3/ResNet-50 use batch 32, weak scaling,
P100 and DGX-1 are measured, FP32 training, software optimization in progress, CUDA8/cuDNN5.1, Ubuntu 14.04

NVIDIA Deep Learning SDK

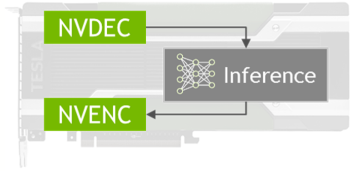
cuDNN



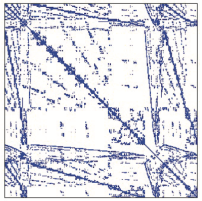
cuBLAS



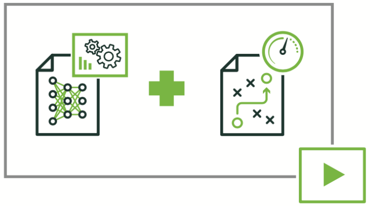
DeepStream SDK



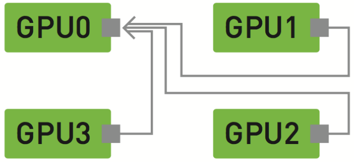
cuSPARSE



TensorRT



NCCL



Caffe Chainer DL4J Deeplearning4j KERAS Microsoft CNTK MatConvNet

MINERVA mxnet Purine TensorFlow theano torch

Computing power (NVIDIA)

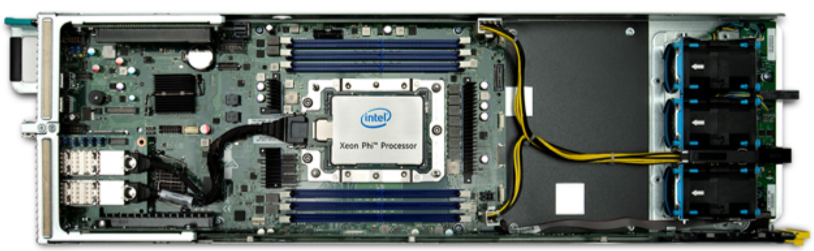
Jetson TX1

	JETSON TX1
GPU	1 TFLOP/s 256-core Maxwell
CPU	64-bit ARM A57 CPUs
Memory	4 GB LPDDR4 | 25.6 GB/s
Video decode	4K 60Hz
Video encode	4K 30Hz
CSI	Up to 6 cameras | 1400 Mpix/s
Display	2x DSI, 1x eDP 1.4, 1x DP 1.2/HDMI
Wifi	802.11 2x2 ac
Networking	1 Gigabit Ethernet
PCIE	Gen 2 1x1 + 1x4
Storage	16 GB eMMC, SDIO, SATA
Other	3x UART, 3x SPI, 4x I2C, 4x I2S, GPIOs

Computing power (Intel)

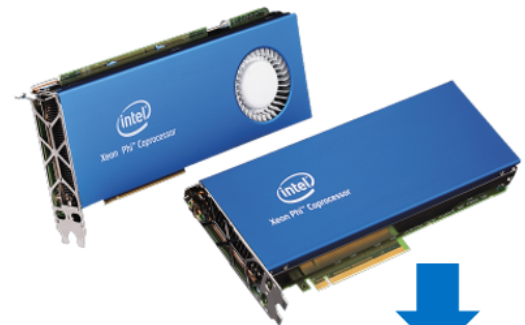
Intel® Xeon Phi™ Processor

with integrated
Intel® Omni-Path Fabric



Host Processor in Groveport Platform
Self-boot Intel® Xeon Phi™ processor

Intel® Xeon Phi™ Coprocessor x200



Ingredient of Grantley Platforms
Requires Intel® Xeon® processor host

Computing power (Intel)

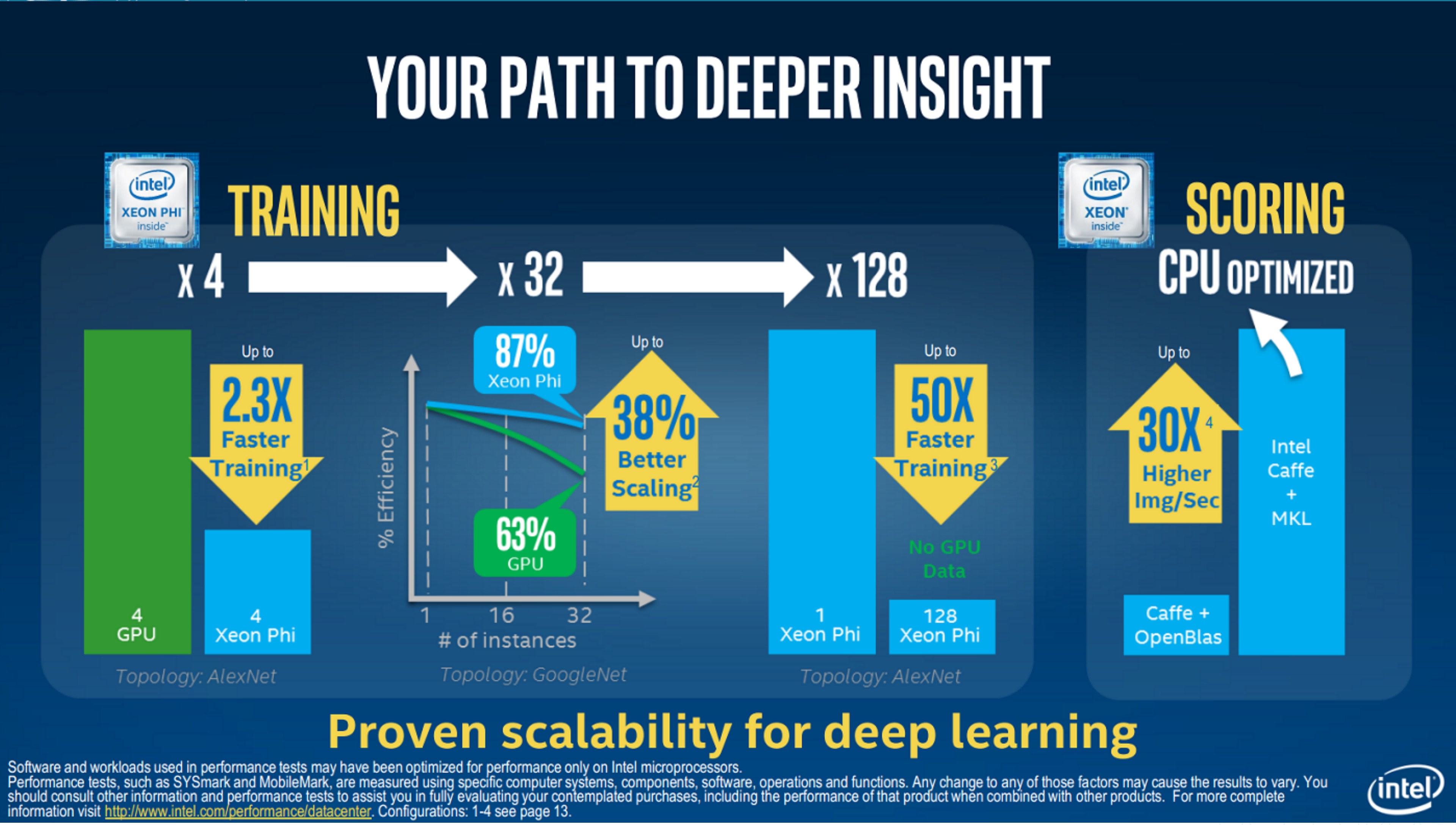
CHOOSE YOUR OPTIMIZATION POINT

	CORES	GHZ	MEMORY (INTEGRATED)	FABRIC (INTEGRATED)	DDR4	POWER[2]	RECOMMENDED CUSTOMER PRICING[3]
7290[1] Best Performance/Node	72	1.5	16GB 7.2 GT/s	Yes	384GB 2400 MHz	245W	$6254
7250 Best Performance/Watt	68	1.4	16GB 7.2 GT/s	Yes	384GB 2400 MHz	215W	$4876
7230 Best Memory Bandwidth/Core	64	1.3	16GB 7.2 GT/s	Yes	384GB 2400 MHz	215W	$3710
7210 Best Value	64	1.3	16GB 6.4 GT/s	Yes	384GB 2133 MHz	215W	$2438

[1] Available beginning in September [2] Plus 15W for integrated fabric

[3] Pricing shown is for parts without integrated fabric. Add additional $278 for integrated fabric versions of these parts. Integrated fabric parts available in October.

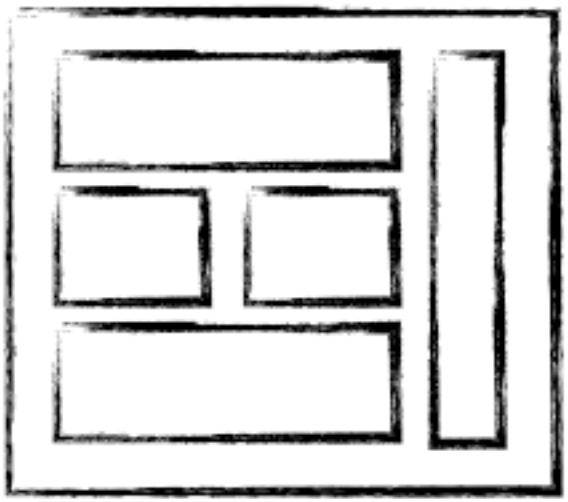
Computing power (Intel)



Computing power (Intel)

- **Intel Math Kernel Library (Intel MKL)**
  Natively supports C, C++ and Fortran Development.
  Cross-language compatible with Java, C#, Python and other languages.
- **Intel Data Analytics Acceleration Library (Intel DAAL)**
  Includes Python, C++, and Java APIs and connectors to popular data sources including Spark and Hadoop.
- **Intel MPI Library**
  Natively supports C,C++ and Fortran development

Architecture



MONOLITHIC/LAYERED

MICRO SERVICES

Docker

Cloud Services

Azure Machine Learning

Amazon Machine Learning

Google Machine Learning

DOMINO

FLY ELEPHANT

DataRobot

ALGORITHMIA

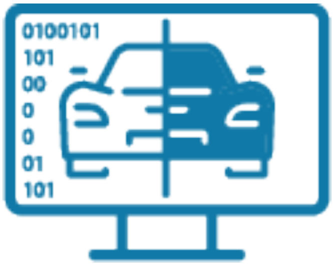
FLY ELEPHANT

Your Home for High Performance Computing

Compute • Collaborate • Manage

Solutions for

Data Science



Engineering Simulation

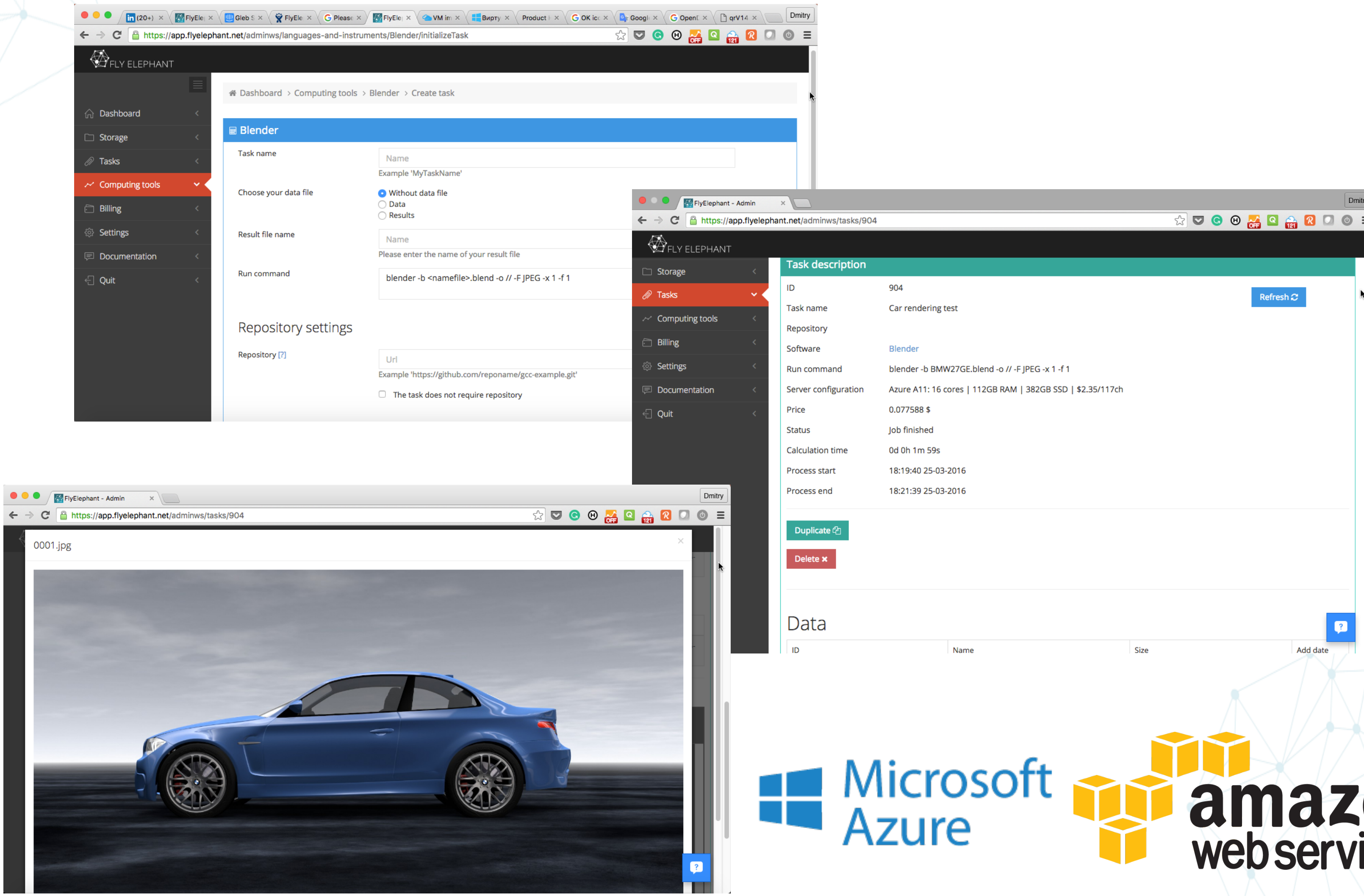
Rendering

Academia

FlyElephant Platform

- **Computing resources.** You can get quick access to different Clouds or HPC clusters from one place.
- **Ready-computing infrastructure.** With one click you get your software calculation infrastructure and the computing resources that you need.
- **Collaboration & Sharing.** Create projects, invite colleagues to join them and share the calculation results within the projects.
- **Fast Deployment.** Deploy your computing tasks as APIs without engineering or DevOps.
- **Expert Community.** Our partner companies and individual experts will help solve any of your issues very quickly.

Cloud computing



Microsoft Azure

amazon web services™

Tools

GCC

R

Scala

python™

Scilab

OpenFOAM

GROMACS FAST. FLEXIBLE. FREE.

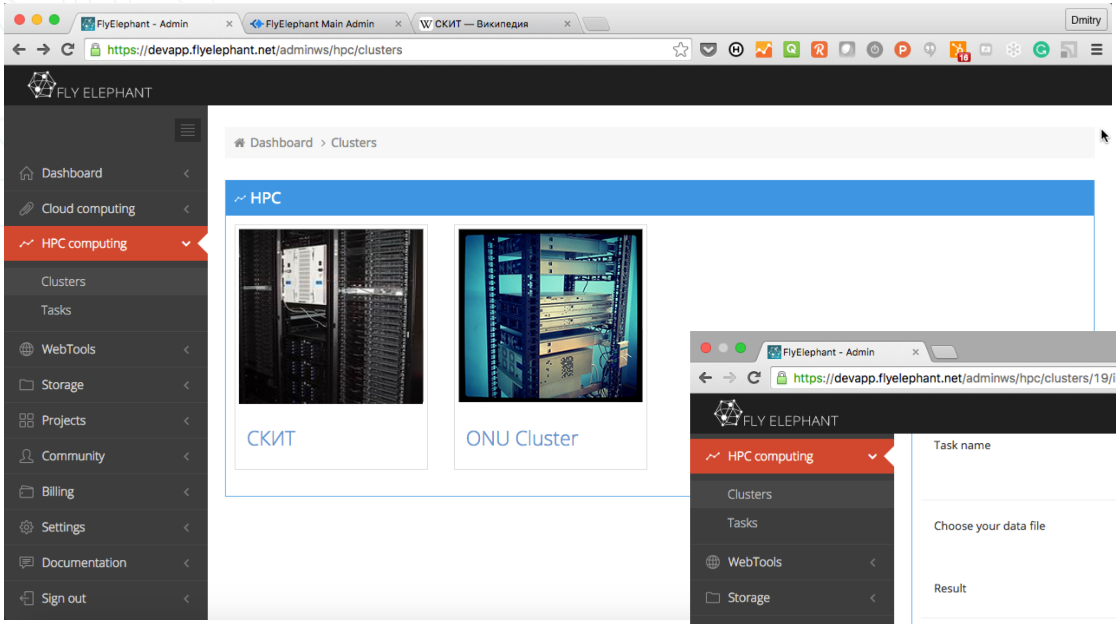
Java

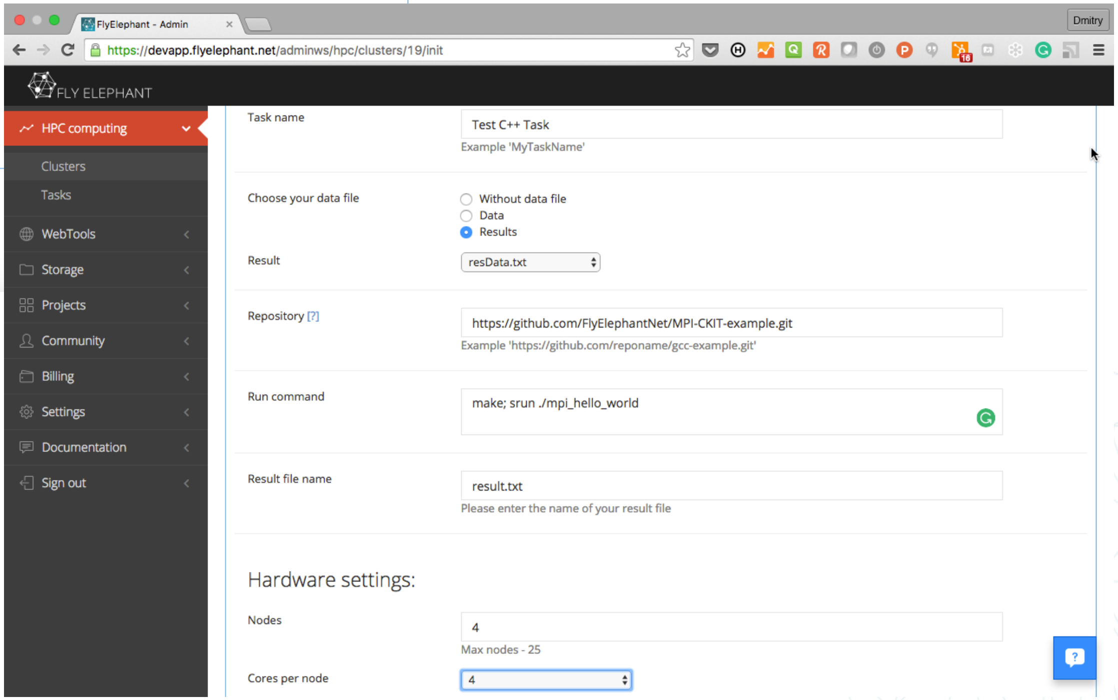
julia

OpenDroneMap

ANACONDA

Computing on a HPC cluster





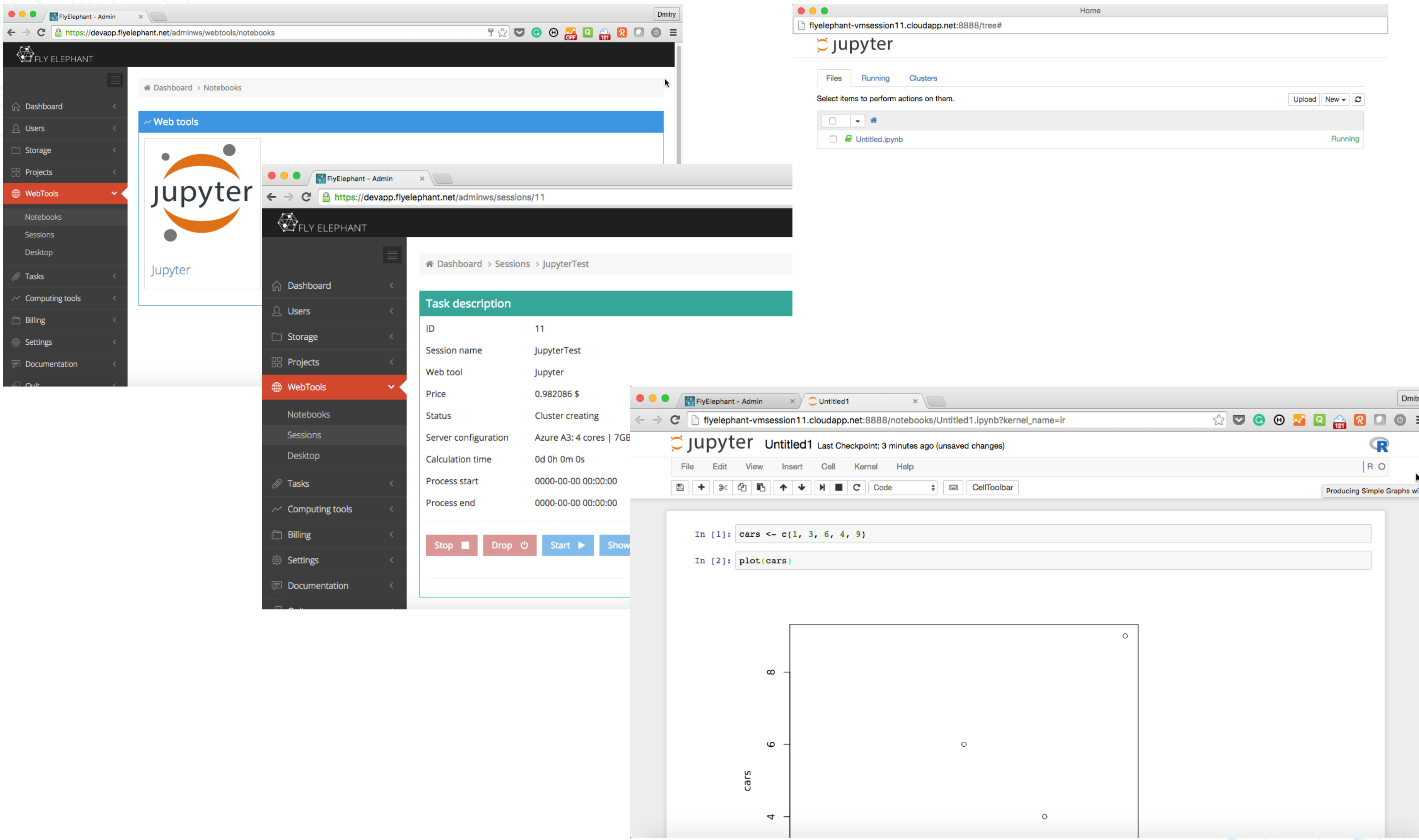
Support of Docker

Hardware settings:

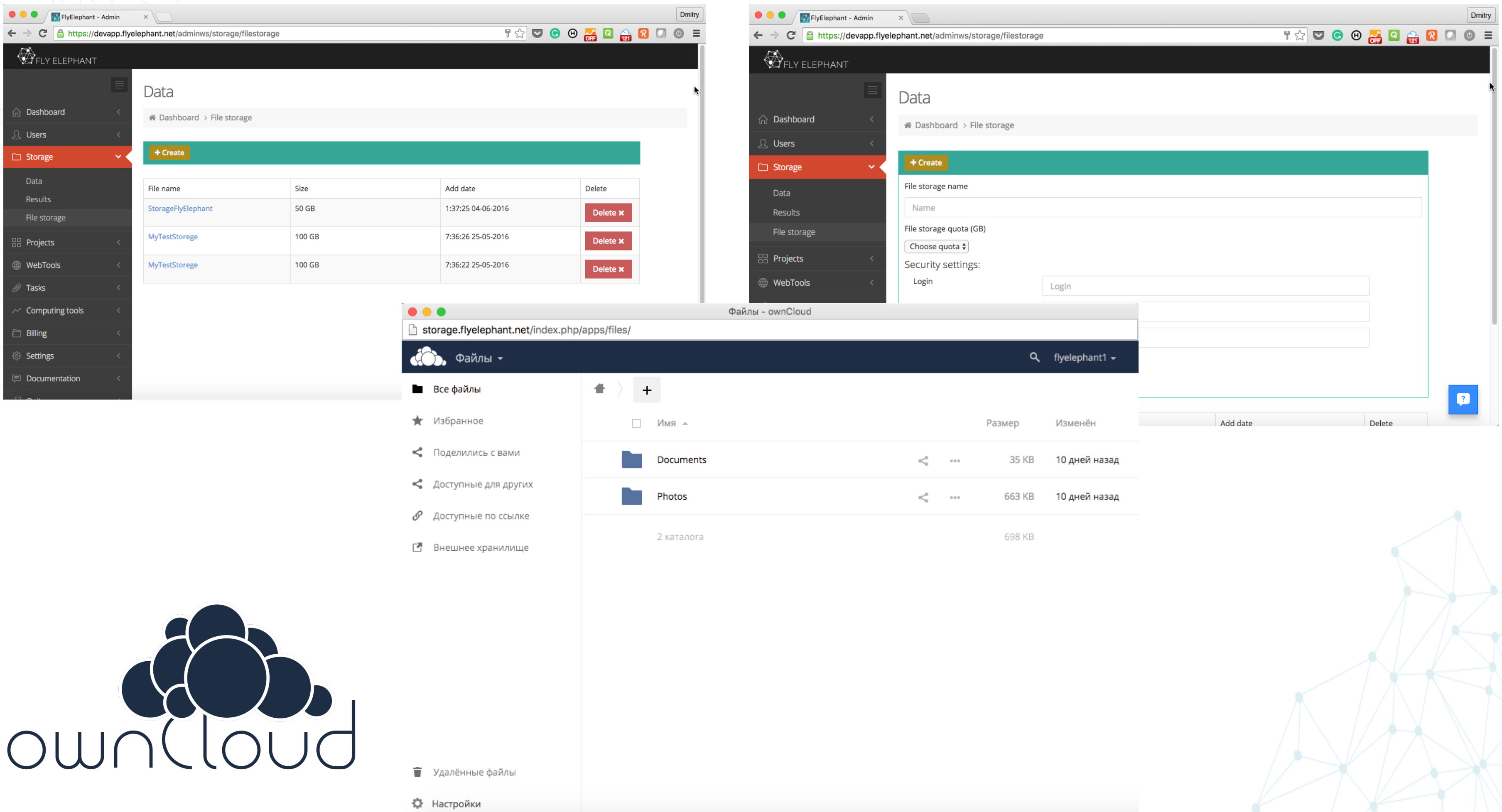
Environment type	Docker
The number of cores to compute	32
Amount of RAM	112 GB

docker

Jupyter Notebooks

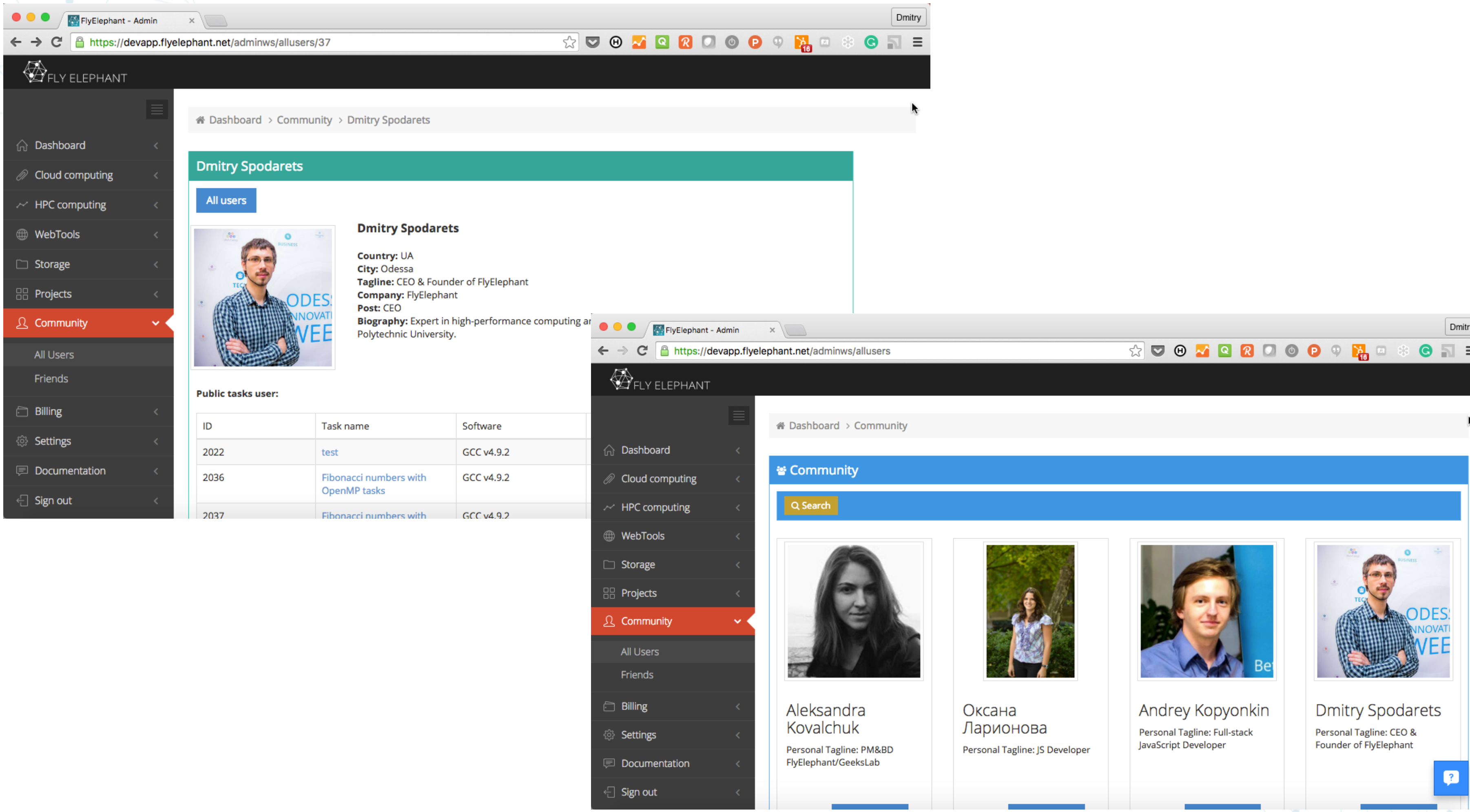


Storage

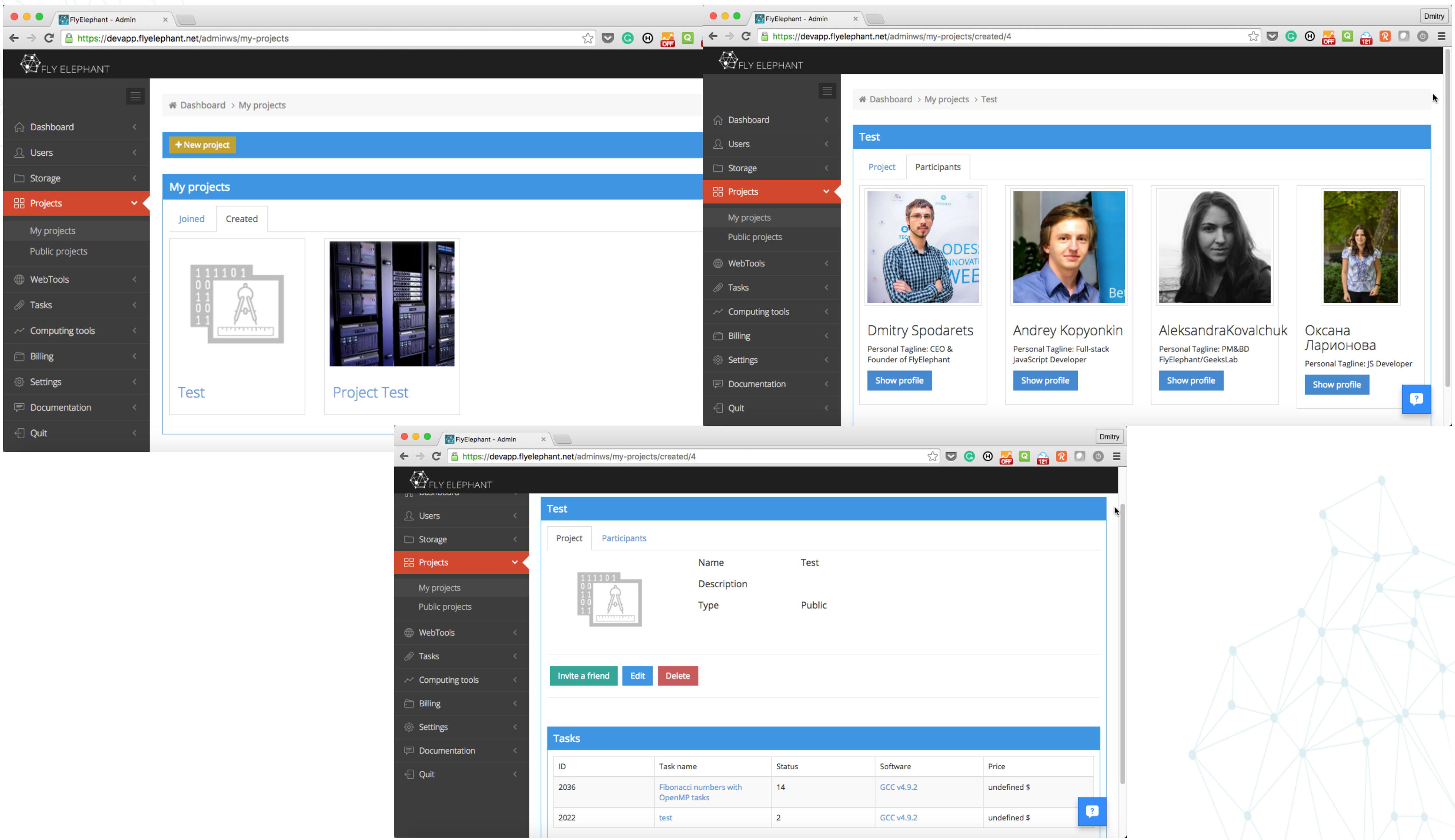


ownCloud

Community


FlyElephant - Admin
https://devapp.flyelephant.net/adminws/allusers/37
FLY ELEPHANT
Dashboard
Cloud computing
HPC computing
WebTools
Storage
Projects
Community
All Users
Friends
Billing
Settings
Documentation
Sign out
Dashboard › Community › Dmitry Spodarets
Dmitry Spodarets
All users
Dmitry Spodarets
Country: UA
City: Odessa
Tagline: CEO & Founder of FlyElephant
Company: FlyElephant
Post: CEO
Biography: Expert in high-performance computing a
Polytechnic University.
Public tasks user:
ID | Task name | Software
2022 | test | GCC v4.9.2
2036 | Fibonacci numbers with OpenMP tasks | GCC v4.9.2
FlyElephant - Admin
https://devapp.flyelephant.net/adminws/allusers
FLY ELEPHANT
Dashboard › Community
Community
Search
Aleksandra Kovalchuk
Personal Tagline: PM&BD FlyElephant/GeeksLab
Оксана Ларионова
Personal Tagline: JS Developer
Andrey Kopyonkin
Personal Tagline: Full-stack JavaScript Developer
Dmitry Spodarets
Personal Tagline: CEO & Founder of FlyElephant

Projects


FlyElephant - Admin
https://devapp.flyelephant.net/adminws/my-projects
FLY ELEPHANT
Dashboard
Users
Storage
Projects
My projects
Public projects
WebTools
Tasks
Computing tools
Billing
Settings
Documentation
Quit
Dashboard > My projects
+ New project
My projects
Joined
Created
Test
Project Test
https://devapp.flyelephant.net/adminws/my-projects/created/4
Dmitry
Dashboard > My projects > Test
Test
Project
Participants
Dmitry Spodarets
Personal Tagline: CEO & Founder of FlyElephant
Show profile
Andrey Kopyonkin
Personal Tagline: Full-stack JavaScript Developer
Show profile
AleksandraKovalchuk
Personal Tagline: PM&BD FlyElephant/GeeksLab
Show profile
Оксана Ларионова
Personal Tagline: JS Developer
Show profile
Name
Test
Description
Type
Public
Invite a friend
Edit
Delete
Tasks
ID
Task name
Status
Software
Price
2036
Fibonacci numbers with OpenMP tasks
14
GCC v4.9.2
undefined $
2022
test
2
GCC v4.9.2
undefined $


API

Roadmap
(November)

- **FlyElephant Box (Closed Beta Preview)**
- AWS
- OpenStack
- Your Docker images
- MPI cluster
- Hadoop/Spark cluster

www.flyelephant.net

slack.flyelephant.net

We are looking for partners

FLY ELEPHANT

Dmitry Spodarets

d.spodarets@flyelephant.net
www.flyelephant.net