

(BUILDING AN) AI PLATFORM ON HTCONDOR

Motivations, lessons learnt and Next
Steps

Cedalion standing on the shoulders of Orion by Nicolas
Poussin, 1658



AGENDA

Motivations

Design Guidelines

Platform Abstractions

Platform Architecture

Demo

Platform Roadmap

Summary

Questions



Team's background was in Hadoop, Spark, Mesos and other 'Big Data' Technologies that came out of the valley



Worked with highly regulated industries like Healthcare and Finance trying to help leverage their data to answer hard questions



Was not satisfied with the restrictions that available technologies were placing along with lack of hybrid cloud support.



Lack of a truly End to End AI/ML platform was also a concern.



SAFE AI was usually non-existent or an after-thought

Security, Assurance, Fairness and Effectiveness (SAFE)

MOTIVATIONS



Do not re-invent
the wheel



Learn from the
mistakes of
others.



Newer does not
always mean
better



Measure twice
and cut once



Security should
not be an
afterthought



Provide freedom
of choice

DESIGN GUIDELINES

PLATFORM ABSTRACTIONS

Datasets

Upload and reference Data

Executables

Binary files, scripts

Libraries

Dependencies to Executables

Workflows

Abstractions on top of
Pegasus DAX files

Notebooks

Jupyter Notebook Support

Dashboards

Visualization dashboards

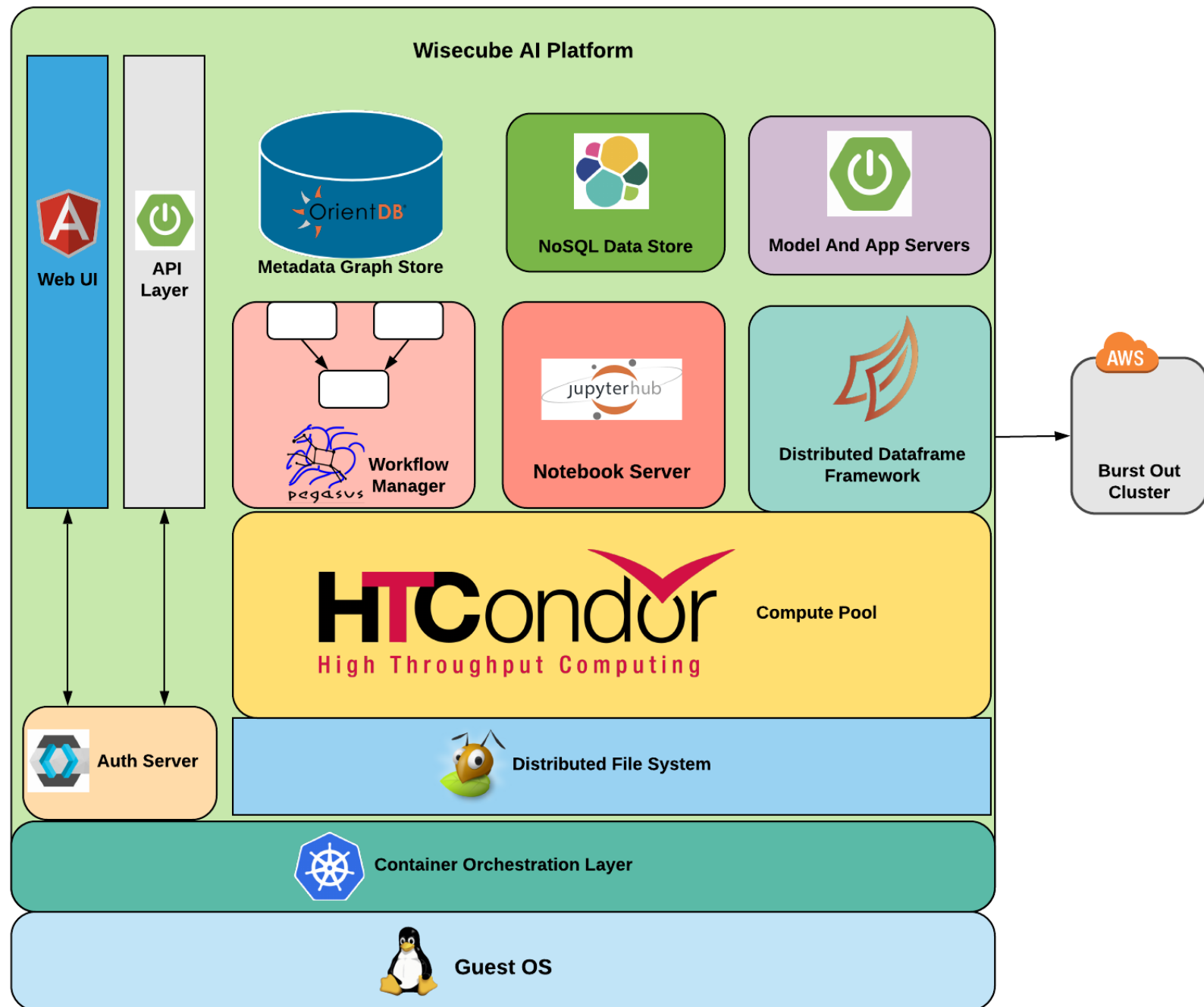
Clusters*

Loosely translates to Condor
pools

Models*

Handles Model file and API

PLATFORM ARCHITECTURE



problem has been detected and Windows has been shut down to prevent damage to your computer.

The problem seems to be caused by the following file: SPCMDCON.SYS

PAGE_FAULT_IN_NONPAGED_AREA

If this is the first time you've seen this Stop error screen, restart your computer. If this screen appears again, follow these steps:

Check to make sure any new hardware or software is properly installed. If this is a new installation, ask your hardware or software manufacturer for any Windows updates you might need.

If problems continue, disable or remove any newly installed hardware or software. Disable BIOS memory options such as caching or shadowing. If you need to use Safe Mode to remove or disable components, restart your computer, press F8 to select Advanced Startup Options, and then select Safe Mode.

Technical information:

*** STOP: 0x00000050 (0xFD3094C2,0x00000001,0xFBFE7617,0x00000000)

*** SPCMDCON.SYS Address FBF57617 base at FBF55000 DateStamp 3d6dd63c

PLATFORM ROADMAP

End to End ML Features – (Models, Apps)

SAFE AI Feature set

Cloud bursting Support (condor-annex looks very promising)

Grid Universe support

Open Source Core platform (Apache 2.0) Looking for potential contributors/users

SUMMARY

Building a Data Science/AI/ML Platform is never fun

Open source tools like Pegasus, HTCondor make it a lot easier

Still a challenge to pick the right set of technologies

Still very use-case dependent

Get ready to do a lot of Devops (and then some more)

K8s has a lot of cool tools to help bundle complex platforms (helm charts etc)

Since all of these tools keep evolving, you are never done !



Wisecube

QUESTIONS?

vishnu@wisecube.ai

www.wisecube.ai