

# Sustaining Research Software: Commercial Perspective

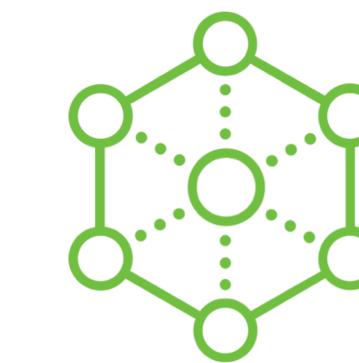
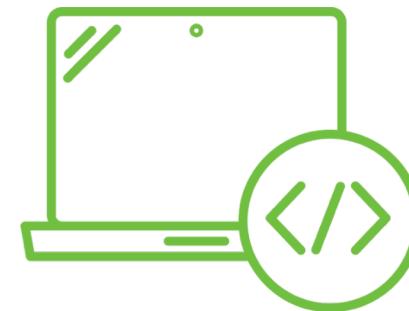
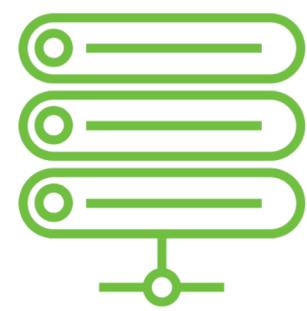
November 14, 2018

Dave Pearah  
CEO

The  Group

# Who is the HDF Group?

We are the creator, official maintainer, and copyright holder of HDF<sup>®</sup>:  
“de-facto standard for scientific computing”



HDF Group has developed open source solutions for Big Data challenges for over 32 years

Small business with focus on High Performance Computing and Exascale Technical Data

Our flagship platform – HDF5 – is the heart of our open source ecosystem. Thousands use + build on HDF5 every day (1000+ projects on Github)

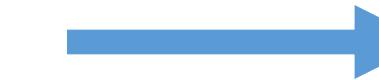
# Unsustainable Software: How did we get here?

We have particularly acute (often unique) challenges with software sustainability because we're not optimizing for the same things that “normal” software people do:

	“Scientific - HPC” Universe	“Regular” Software Universe
Target goal	Fastest code RUN time [Formula 1 lap time]	Fastest code AUTHOR time [Toyota Camry assembly time]
Constraint [expensive / scarce]	HPC cycles [hardware]	Programmers [people]
Timeframe	Project / paper	Lots of users... forever

# HDF Sustainability 1.0

We primarily  
did this...



Consulting & Engineering

... to eek out  
free cycles to  
sustain this

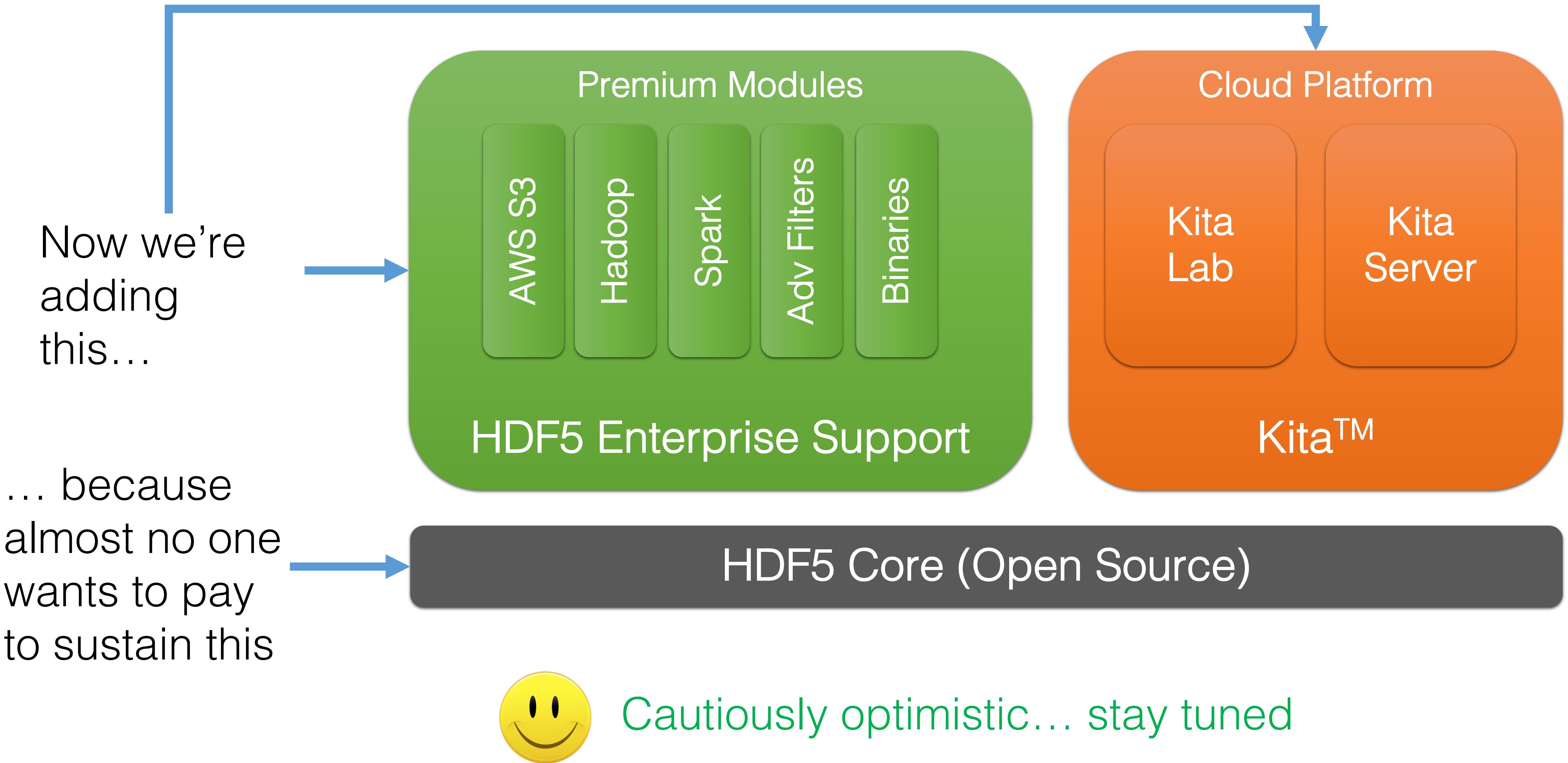


HDF5 Core (Open Source)



Spoiler: doesn't work at scale

# HDF Sustainability 2.0



# Recommendation: 5% for Sustainability

- **Create sustainable economy for software that aligns funding to value**
  - For each funding call, require 5% of funding to go to sustainable software
  - Specifically: if 3<sup>rd</sup>-party open source software is used, then share that 5% with the people / organizations tasked with maintaining that code
  - Funding will be aligned with usefulness
  - A little goes a long way, particularly for popular codes
- Much better than one-time, spot programs that focus on sustainability... which is just a lottery system / popularity contest