

HDF Group Past, Present & Future



Proprietary and Confidential. Copyright 2016, The HDF Group.



Outline

Who we are and where
we come from

Mission and
commitments to
community

Future directions for
HDF Group & HDF
R&D

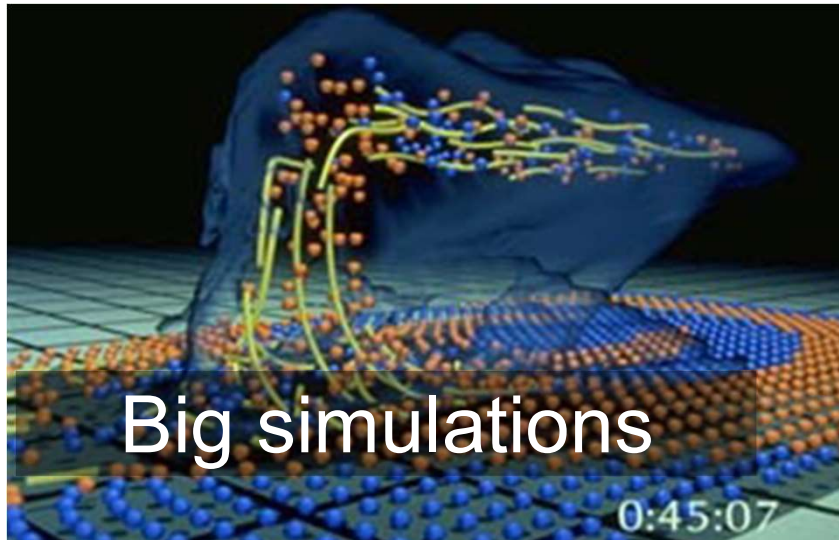
HDF
Group



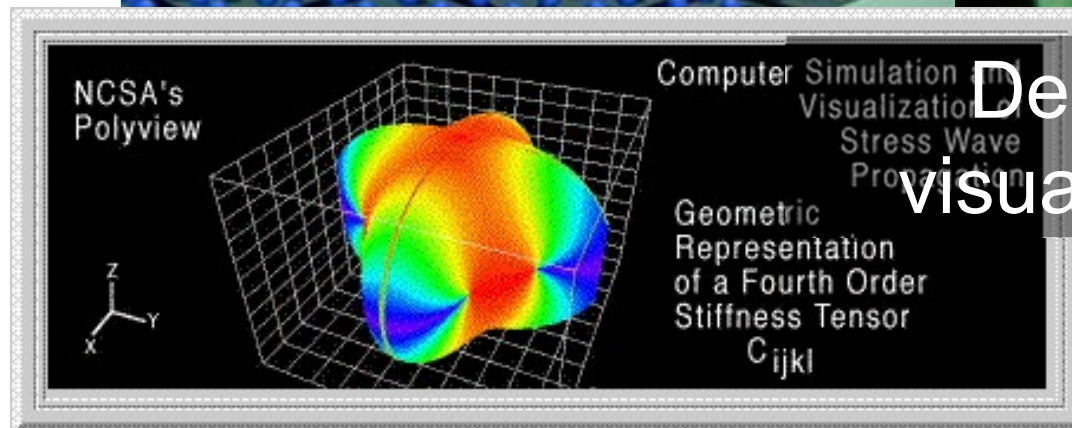
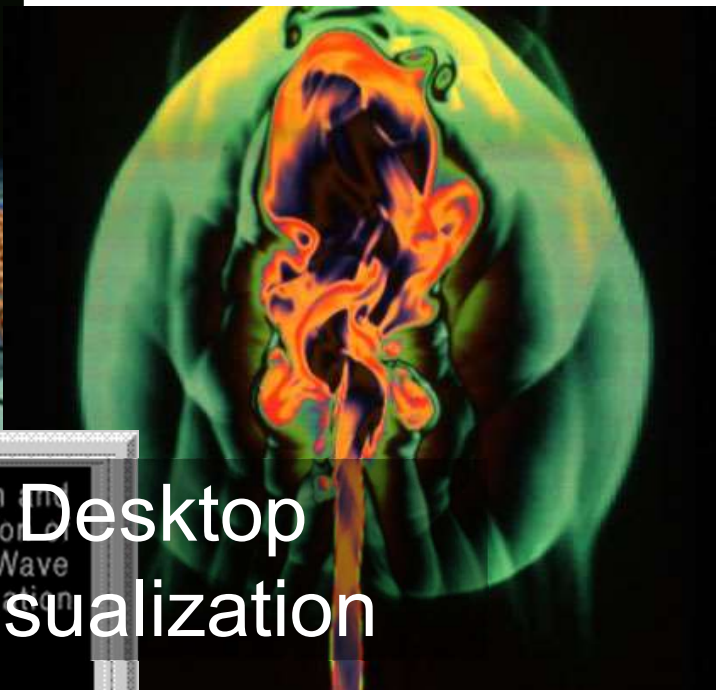


Who we are and where we come from

National Center for Supercomputing Applications - 1987



Big simulations



Desktop
visualization

The birth of HDF

Graphics Foundations Task Force
Minutes of Meeting 10/28/87

Attending:

lwhite, marrott, tkrauskopf, mkrogh, bmihalas, dhe

Chair & Reporting:

lwhite

Regular meeting date:

Thursday's 9:30 to 11:00 room 173

Discussion:

Cooperation with NSF & DOE centers.

Selection of Standard Graphics File Format.

Some of our needs in standard file format:

extensible data storage mechanism

speed

options for compression

clear capability for grided data storage

our formats (nor will they ever).

4. It's format spec is readily available to anyone wanting it.
5. It can store data other than just raster data, such as text and basic geometry.



Reasons for having an AEHOO level format:

1. It is possible to store data in a method that has not been possible to do with other formats.
2. It should encompass every conceivable future type of data (expandability).
3. If and when it works (well) others may adopt it.
4. It makes interesting research.
5. It could be a machine independent format.
6. All other data formats should be included as subsets of an AEHOO file format.



Version 1.0

Draft - for comment and review



HIERARCHICAL DATA FORMAT

Version 1.0 HDF Specification

Draft version 3.0
April 1988

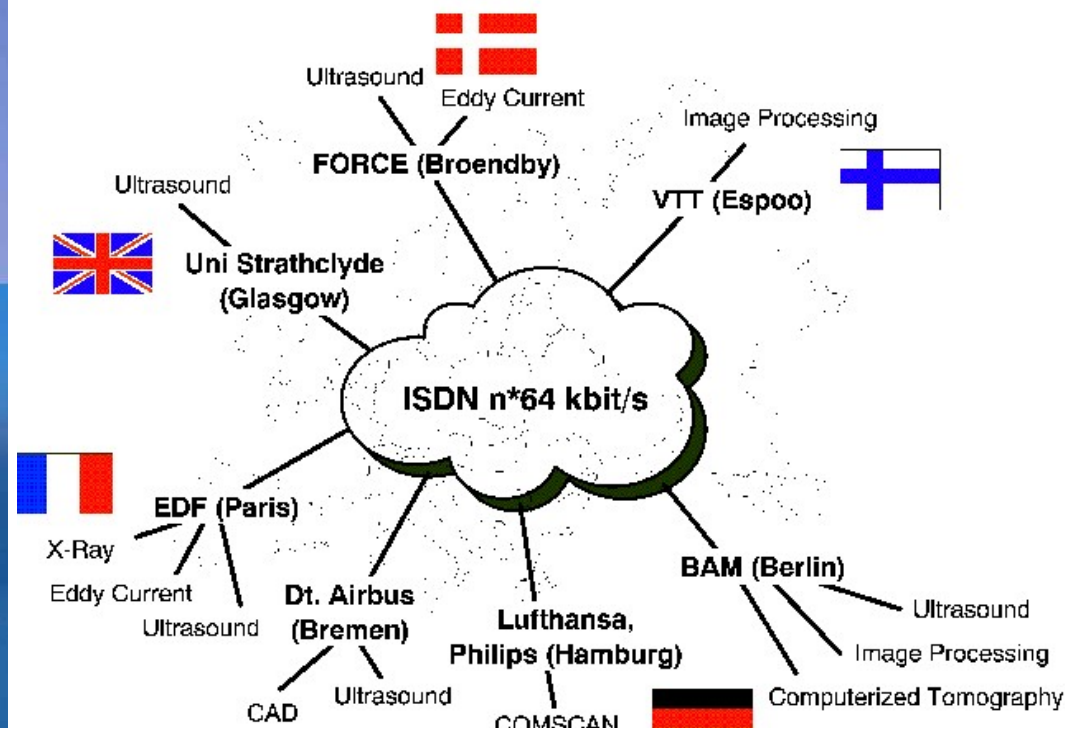
Tim Krauskopf
Gaige B. Paulsen

National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign
Copyright © 1988 Board of Trustees of the University of Illinois

NASA Earth Observing System

- 6,700 Data Products
- 12 Data Archive Centers
- 16 terabytes per day
- Distributed to 1.7 million end users worldwide





Non-destructive testing - TRAPPIST

- Limits on object & file size (<2GB)
- Limited number of objects (<20K)
- Rigid data models
- I/O performance
- Code complexity

Shortcomings of HDF (1996)

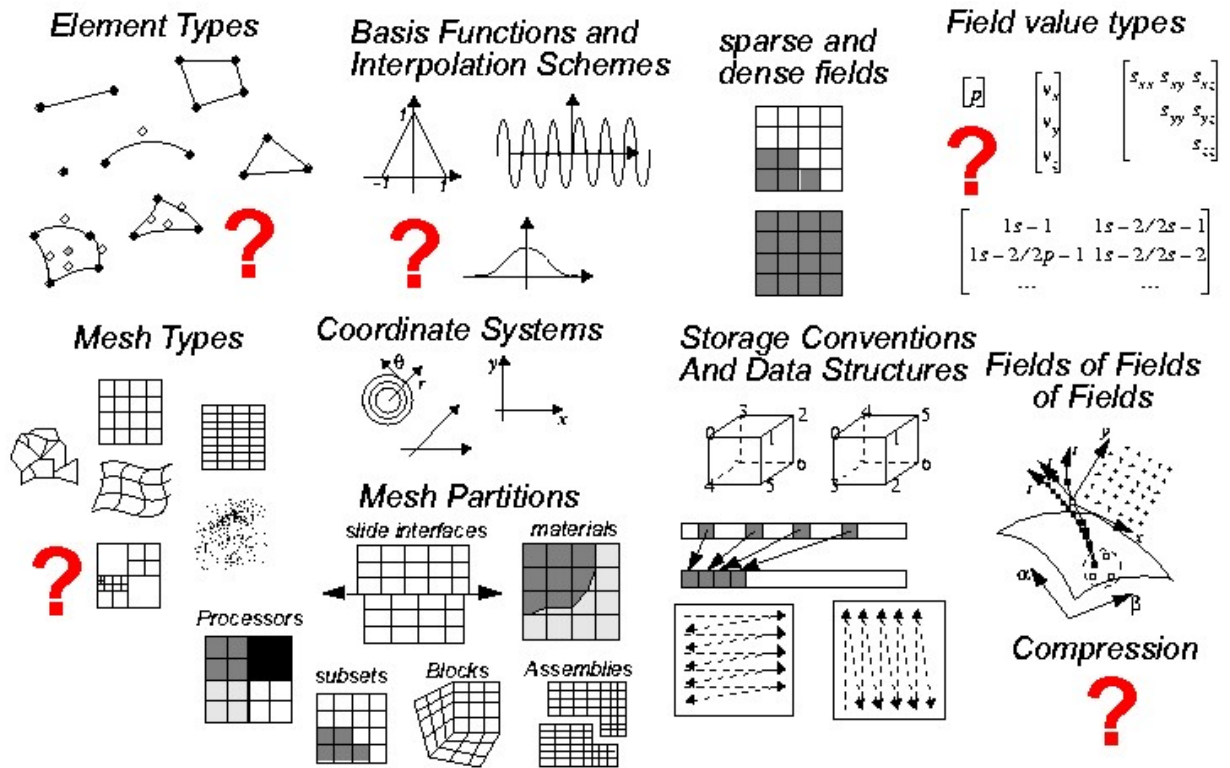
**Accelerated
Strategic
Computing
Initiative**



How to maintain a nuclear stockpile without testing?



Data can be complicated



The birth of HDF5



A black and white photograph of a woman in a striped shirt working in a large archive. She is standing between tall metal shelving units filled with microfilm reels. She is holding a stack of reels on a small metal cart. The shelves are filled with rows of microfilm reels, creating a sense of a vast collection. The woman is looking up at the shelves, reaching for a reel. The overall scene conveys the scale and manual nature of long-term data preservation.

Long-term
data
preservation

HDF keys to success

Needs

Open source durability

Complex uses

Complex formatting

Long term storage

HDF
Group

Institutional support

User support

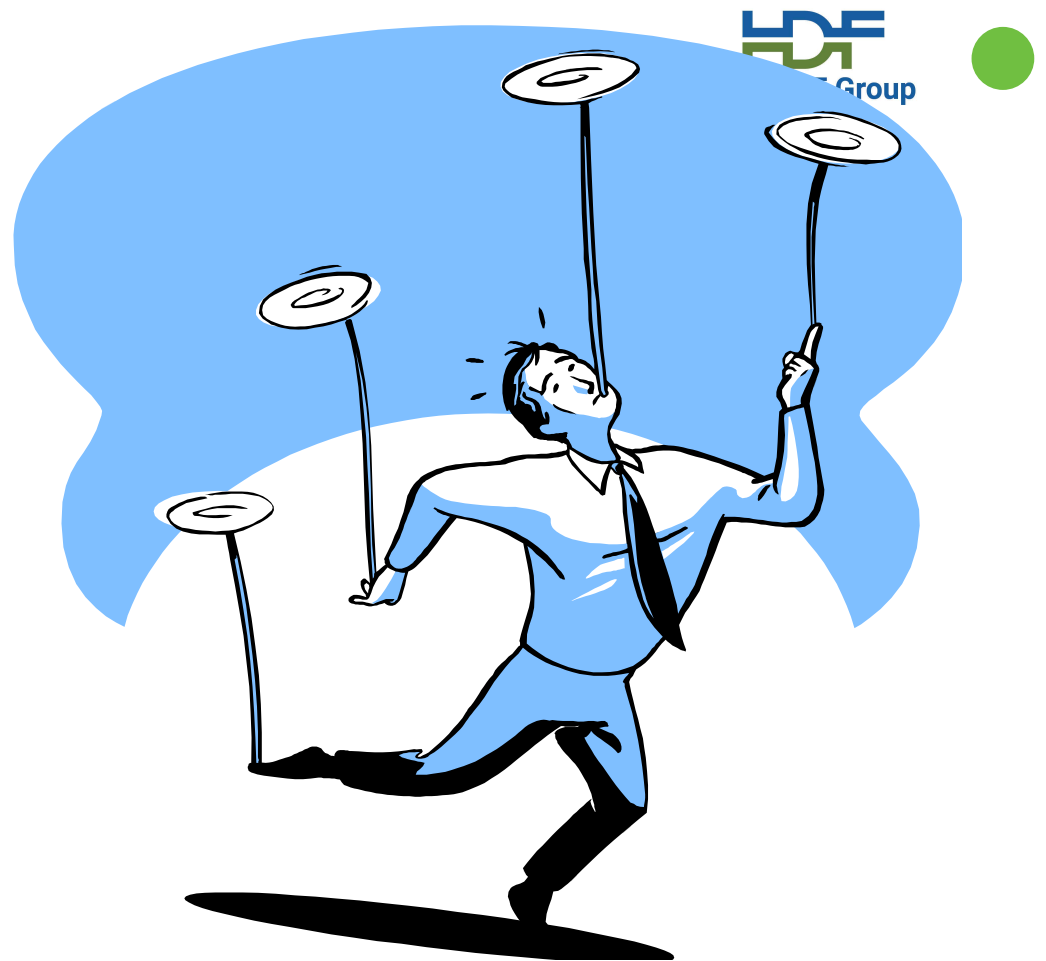
API, library, utilities,
documentation

Software for the long term.



Spinning off

- Business model options
- For-Profit or not?
- How to make money?
- Intellectual property



Who is the HDF Group?



Small company
~ 40 employees
focus on High
Performance Computing
and Scientific Data
Offices in Champaign, IL
+ 7 other locations



Our flagship platform –
HDF5 – is at the heart of our
open source ecosystem.
Tens of thousands use
HDF5 every day.
1500+ projects on Github



Work with industry,
research institutions,
government
to improve HDF5,
build HDF5 solutions,
provide support.

Mission and commitments to OSS and community

The HDF Group

Log In

Create Free Account

Downloads

Documentation

Support Portal

Register

Twitter icon, Search icon

HDF Group
Website

HDF5 European Workshop for Science and Industry

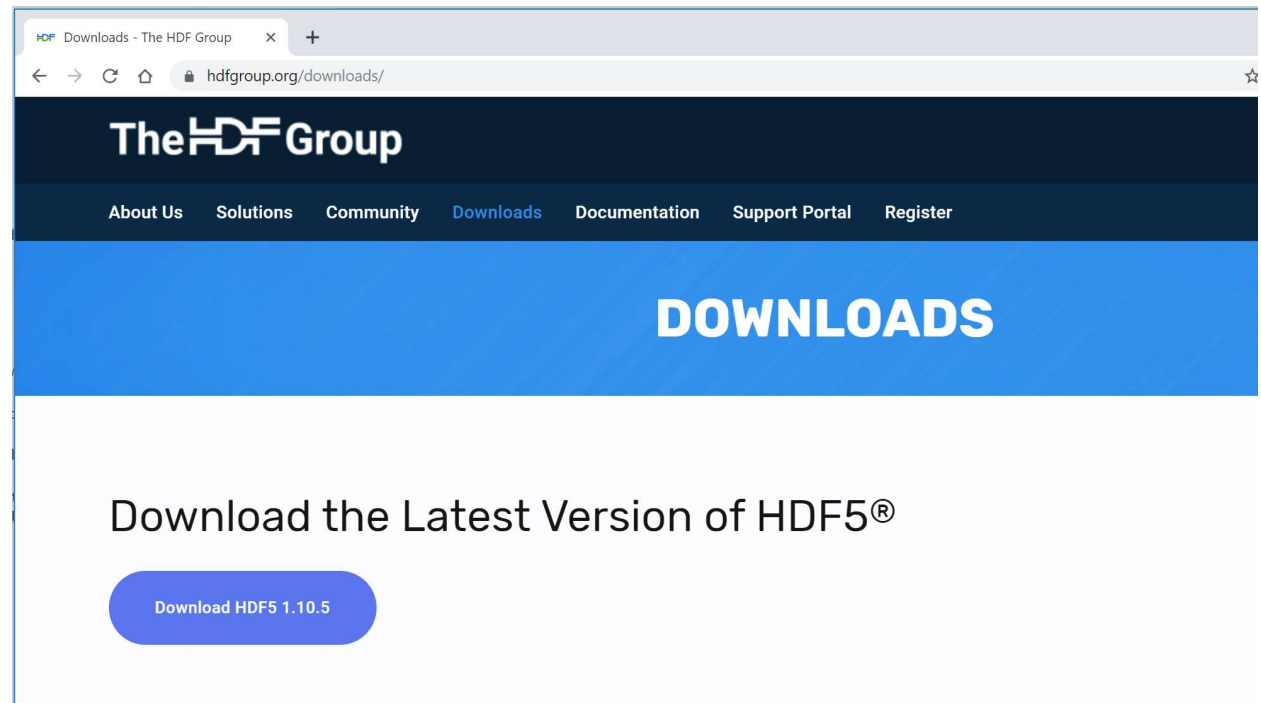
17-18 September 2019

Submit Abstract

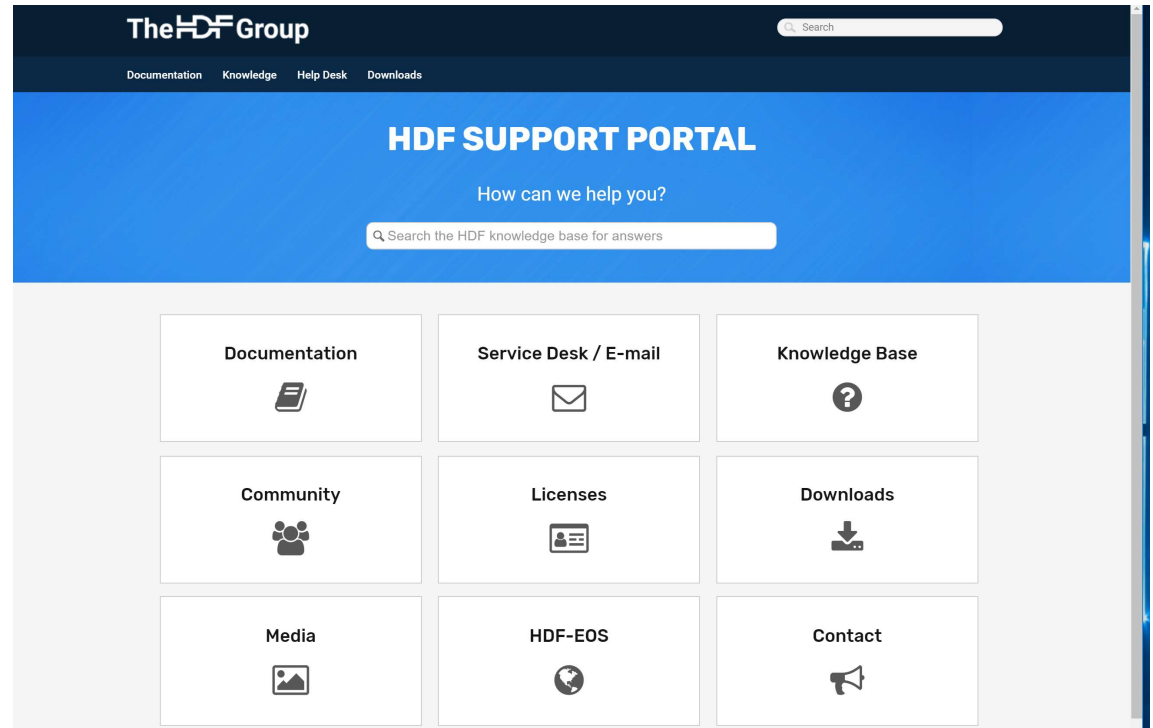
More Info

Register to Attend

Software

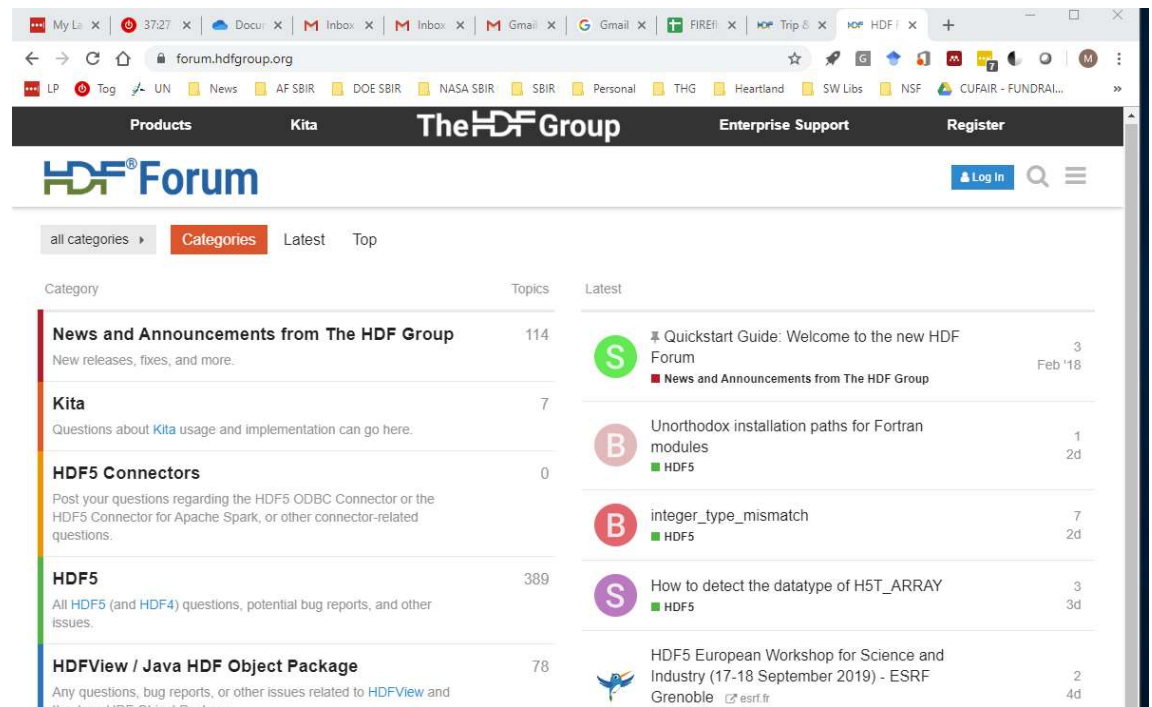


HDF Support portal












The screenshot displays the HDF Support Portal website. At the top, the 'The HDF Group' logo is on the left, and a search bar is on the right. Below the logo, a navigation bar contains links for 'Documentation', 'Knowledge', 'Help Desk', and 'Downloads'. The main header is a blue banner with the text 'HDF SUPPORT PORTAL' and 'How can we help you?'. Below this is another search bar with the placeholder text 'Search the HDF knowledge base for answers'. The main content area features a 3x3 grid of service tiles, each with an icon and a title: 'Documentation' (book icon), 'Service Desk / E-mail' (envelope icon), 'Knowledge Base' (question mark icon), 'Community' (group of people icon), 'Licenses' (license document icon), 'Downloads' (download arrow icon), 'Media' (image icon), 'HDF-EOS' (globe icon), and 'Contact' (megaphone icon).

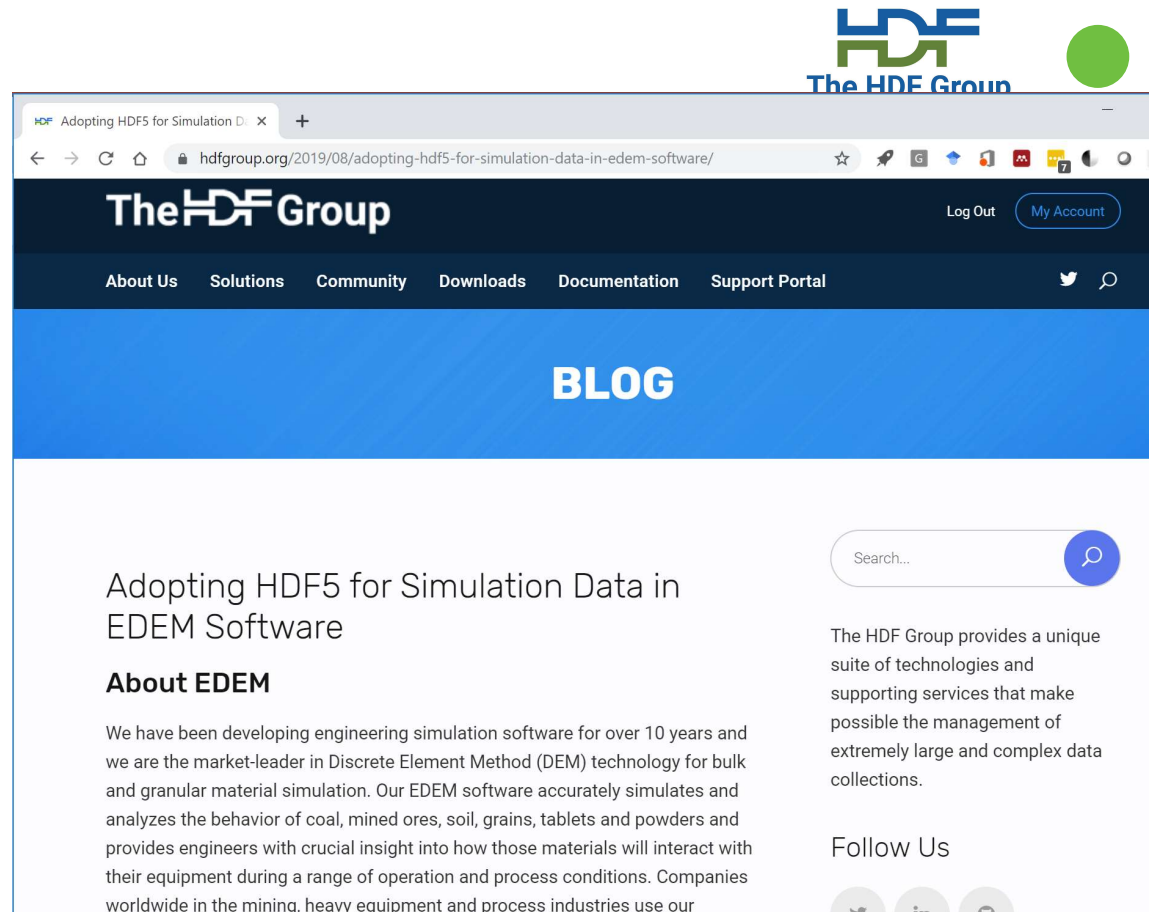
HDF Forum



The screenshot shows the HDF Forum website interface. At the top, there's a navigation bar with links for Products, Kita, The HDF Group, Enterprise Support, and Register. Below this is the forum header with the HDF Forum logo and a Log In button. The main content area is divided into two columns. The left column lists categories with their respective topic counts: News and Announcements from The HDF Group (114), Kita (7), HDF5 Connectors (0), HDF5 (389), and HDFView / Java HDF Object Package (78). The right column displays a list of recent forum posts, each with a topic icon, title, category, and date. The posts include a Quickstart Guide, a discussion on Fortran modules, an integer_type_mismatch issue, a question about H5T_ARRAY datatype, and an announcement for the HDF5 European Workshop.

Category	Topics	Latest
News and Announcements from The HDF Group New releases, fixes, and more.	114	 Quickstart Guide: Welcome to the new HDF Forum  News and Announcements from The HDF Group Feb '18
Kita Questions about Kita usage and implementation can go here.	7	 Unorthodox installation paths for Fortran modules  HDF5 1 2d
HDF5 Connectors Post your questions regarding the HDF5 ODBC Connector or the HDF5 Connector for Apache Spark, or other connector-related questions.	0	 integer_type_mismatch  HDF5 7 2d
HDF5 All HDF5 (and HDF4) questions, potential bug reports, and other issues.	389	 How to detect the datatype of H5T_ARRAY  HDF5 3 3d
HDFView / Java HDF Object Package Any questions, bug reports, or other issues related to HDFView and the Java HDF Object Package.	78	 HDF5 European Workshop for Science and Industry (17-18 September 2019) - ESRF Grenoble 2 4d

Blogs



The screenshot displays the website of The HDF Group. At the top right, the logo for 'The HDF Group' is visible, featuring a stylized 'HDF' in blue and green. Below the logo, the website's navigation bar includes links for 'About Us', 'Solutions', 'Community', 'Downloads', 'Documentation', and 'Support Portal'. A 'Log Out' button and a 'My Account' link are also present. The main content area is titled 'BLOG' in large white letters on a blue background. Below this, the featured article is 'Adopting HDF5 for Simulation Data in EDEM Software'. The article's introduction states: 'We have been developing engineering simulation software for over 10 years and we are the market-leader in Discrete Element Method (DEM) technology for bulk and granular material simulation. Our EDEM software accurately simulates and analyzes the behavior of coal, mined ores, soil, grains, tablets and powders and provides engineers with crucial insight into how those materials will interact with their equipment during a range of operation and process conditions. Companies worldwide in the mining, heavy equipment and process industries use our'. To the right of the article, there is a search bar with the placeholder text 'Search...' and a magnifying glass icon. Below the search bar, a brief description of The HDF Group's services is provided: 'The HDF Group provides a unique suite of technologies and supporting services that make possible the management of extremely large and complex data collections.' At the bottom right, there is a section titled 'Follow Us' with icons for social media platforms.

The HDF Group

Adopting HDF5 for Simulation Data in EDEM Software

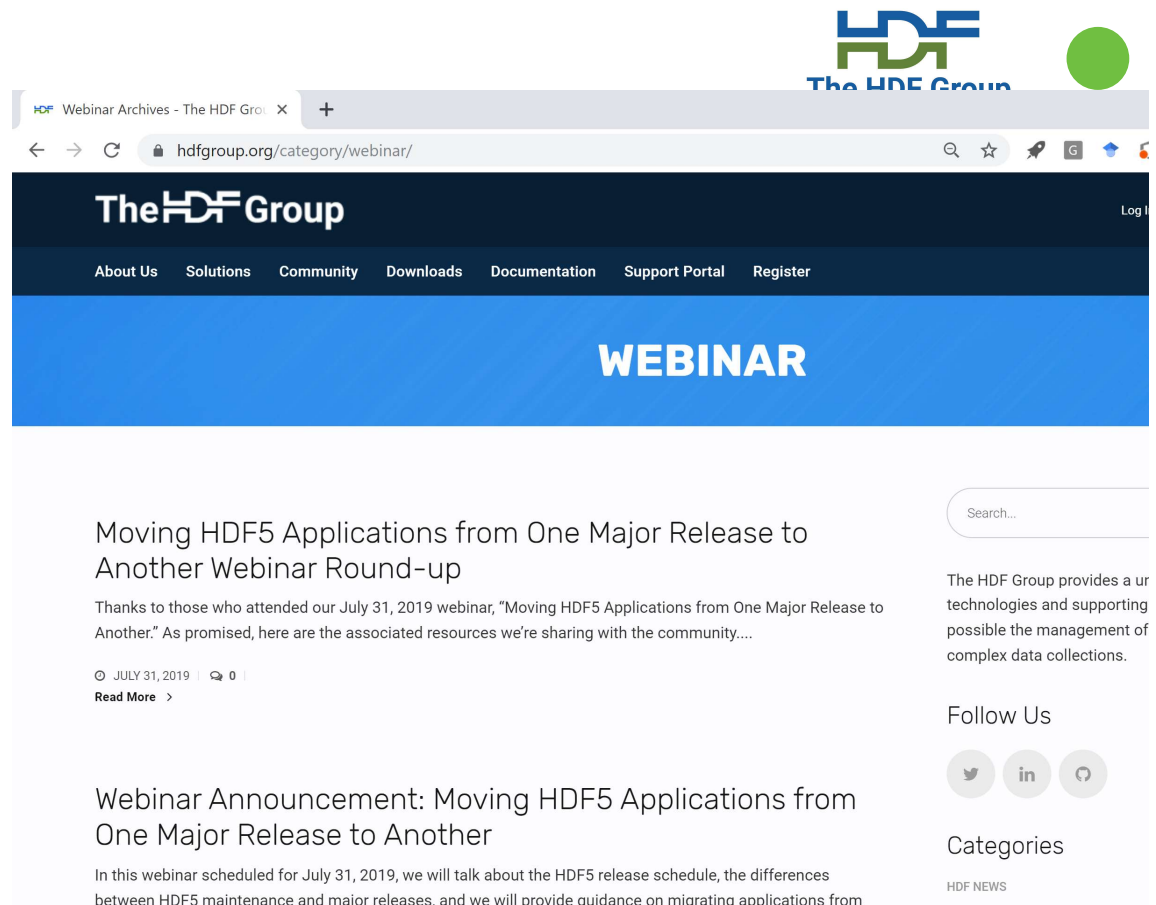
About EDEM

We have been developing engineering simulation software for over 10 years and we are the market-leader in Discrete Element Method (DEM) technology for bulk and granular material simulation. Our EDEM software accurately simulates and analyzes the behavior of coal, mined ores, soil, grains, tablets and powders and provides engineers with crucial insight into how those materials will interact with their equipment during a range of operation and process conditions. Companies worldwide in the mining, heavy equipment and process industries use our

The HDF Group provides a unique suite of technologies and supporting services that make possible the management of extremely large and complex data collections.

Follow Us

Webinars



The screenshot displays the 'Webinar Archives' page on the HDF Group website. The browser's address bar shows the URL 'hdfgroup.org/category/webinar/'. The website's header includes the 'The HDF Group' logo and a navigation menu with links: 'About Us', 'Solutions', 'Community', 'Downloads', 'Documentation', 'Support Portal', and 'Register'. A prominent blue banner with the word 'WEBINAR' in white capital letters is positioned below the navigation menu. The main content area features a featured webinar titled 'Moving HDF5 Applications from One Major Release to Another Webinar Round-up'. Below the title, a short paragraph describes the webinar, followed by the date 'JULY 31, 2019' and a 'Read More >' link. To the right of the main content, there is a search bar, a brief description of the HDF Group's services, a 'Follow Us' section with social media icons for Twitter, LinkedIn, and YouTube, and a 'Categories' section with a link to 'HDF NEWS'.

The HDF Group

WEBINAR

Moving HDF5 Applications from One Major Release to Another Webinar Round-up

Thanks to those who attended our July 31, 2019 webinar, "Moving HDF5 Applications from One Major Release to Another." As promised, here are the associated resources we're sharing with the community...

JULY 31, 2019 | 0 |
[Read More >](#)

Webinar Announcement: Moving HDF5 Applications from One Major Release to Another

In this webinar scheduled for July 31, 2019, we will talk about the HDF5 release schedule, the differences between HDF5 maintenance and major releases, and we will provide guidance on migrating applications from

Search...

The HDF Group provides a ur technologies and supporting possible the management of complex data collections.

Follow Us

[Twitter](#) [LinkedIn](#) [YouTube](#)

Categories

[HDF NEWS](#)



Technology Advisory Board

Get engaged



Present a Webinar
about your project

HDFqI in Nov/Dec 2019

Write a blog

About your project or some aspect of HDF5

Contact

lori.cooper@hdfgroup.org

Future directions for the HDF Group and HDF R&D

1998-2007 - Maturation

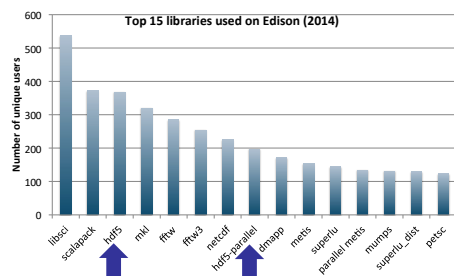
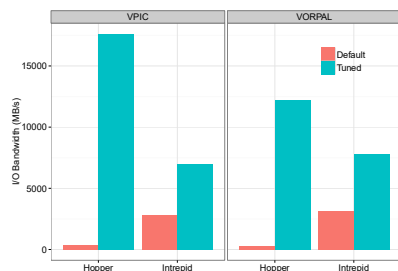
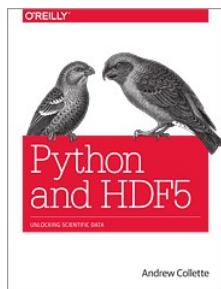
- Matured the HDF5 data model and customized I/O including parallel I/O
- Community standards NeXUS, HDF-EOS, etc. emerged





2008-2019 - Expanding to non-HPC communities and Exascale computing

- Widespread use
- Improved I/O performance
- Growth of software ecosystem
- New features to support particle accelerators, exascale architectures, observational data, etc.





Future directions

Continue...



Support legacy library versions

Adapt to OS changes, etc.

Improve library, tools, docs

Increase community involvement

Leverage new memory, computing and storage architectures



- Deep memory hierarchy
- Alternate storage and I/O
- Non-traditional HDF5 uses

Facilitate convergence of Big Data and HPC

- Combine data and compute services
- Large scale query





Ease of use, productivity

- New APIs for C++ community
- Working with tool developers

THANK YOU!

Questions & Comments?