How can we leverage software foundations for DOE software sustainability?

Todd Gamblin

PESO Co-PI

Lawrence Livermore National Laboratory

What are software foundations

- Nonprofit corporations, typically one of:
  - **501(c)(3): Charitable (religious/educational/scientific) organization**
    - Serves the public
    - Cannot lobby or influence legislation
    - Exempt from federal, state, and local taxes

  - **501(c)(6): Organization of people with common business interests**
    - Serves its members' common interest
    - Seeks to improve business conditions
    - Exempt from only federal income taxes
    - Can lobby (but funds used for lobbying are taxed)

"THE APACHE WAY TO ME…"

"community beyond commercial offerings"
— Mei Long, 18 year user of Apache Zeppelin, Spark, Hadoop, Lucene, Hive, Maven, Cassandra, contributor to Apache Zeppelin

WHAT MAKES THE APACHE WAY SO HARD TO DEFINE?

The Apache Way is a living, breathing interpretation of one's experience with our community-led development process. Apache projects and their communities are unique, diverse, and focused on the activities needed at a particular stage of the project's lifetime, including nurturing communities, developing great code, and building awareness. What is important is that they embrace:

- **Earned Authority:** all individuals are given the opportunity to participate, but their influence is based on publicly earned merit – what they contribute to the community. Merit lies with the individual, does not expire, is not influenced by employment status or employer, and is non-transferable (merit earned in one project cannot be applied to another). More on merit.

- **Community of Peers:** individuals participate at the ASF, not organizations. The ASF's flat structure dictates that roles are equal irrespective of title, votes hold equal weight, and contributions are made on a volunteer basis (even if paid to work on Apache code). The Apache community is expected to treat each other with respect in adherence to our Code of Conduct. Domain expertise is appreciated; Benevolent Dictators For Life are disallowed. More on individual participation.

- **Open Communications:** as a virtual organization, the ASF requires all communications related to code and decision-making to be publicly accessible to ensure asynchronous collaboration, as necessitated by a globally-distributed community. Project mailing lists are archived, publicly accessible, and include:

  - dev@ (primary project development)
  - user@ (user community discussion and peer support)
  - commits@ (automated source change notifications)
  - occasionally supporting roles such as marketing@ (project visibility)

…as well as restricted, day-to-day operational lists for Project Management Committees. Private decisions on code, policies, or project direction are disallowed; off-list discourse and transactions must be brought on-list. More on communications and the use of mailing lists.

Our Mission

The mission of NumFOCUS is to promote open practices in research, data, and scientific computing by serving as a fiscal sponsor for open source projects and organizing community-driven educational programs.

NumFOCUS is a 501(c)(3) public charity in the United States.

Our Vision

NumFOCUS envisions an inclusive scientific and research community that utilizes actively supported open source software to make impactful discoveries for a better world.

How We Support Open Source

NumFOCUS provides a stable, independent, and professional home for the open source projects powering contemporary scientific inquiry and business processes. We aim to ensure that funding and resources are available to **sustain projects in the scientific data stack** over the long haul.

Learn more about our Fiscal Sponsorship Program.

What's In a Name?

Our first Treasurer, Anthony Scopatz, came up with the name "NumFOCUS" as an abbreviation for "Numerical Foundation for Open Code and Useable Science." (This is why you will sometimes see us referred to as "NumFOCUS Foundation.") While officially the NumFOCUS name isn't an abbreviation for anything, our mission supports open code as the foundation for useable and reproducible science—in other words, Open Code = Better Science.

The Linux Foundation Method

A proven and repeatable way to scale project communities via a comprehensive portfolio of support programs for aspiring industry leading projects.

Neutral home for code and collaboration

We aim to democratize code and scale adoption, for all projects.

Ecosystem curation and community building

We strive to create new technology categories by identifying trends, accelerating the growth of nascent technologies, and removing barriers to adoption.

Enterprise ready, the OSS way

We provide turnkey technology and support programs for developer enablement, business operations, training & certification, marketing and events, and membership development to help projects scale fast.

Project insights and management tools

We help projects streamline operations and boost community engagement with cloud-based, collaborative tooling, contributor and participation analytics, and infrastructure management.
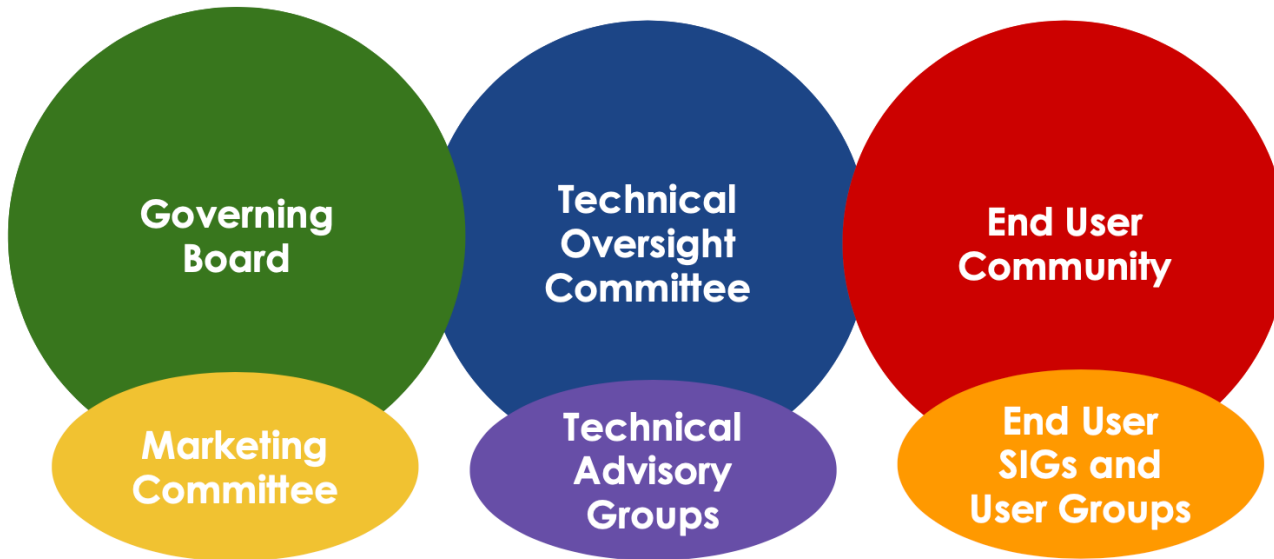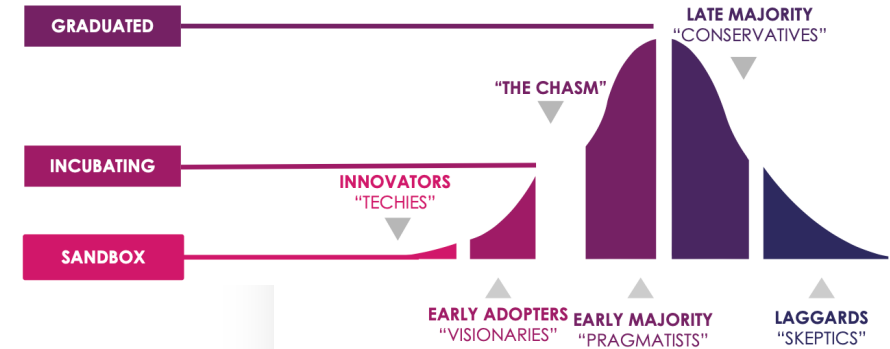
CLOUD NATIVE COMPUTING FOUNDATION

Who is CNCF?

The CNCF hosts critical components of the global technology infrastructure. CNCF brings together the world's top developers, end users, and vendors and runs the largest open source developer conferences.

GRADUATED

INCUBATING

SANDBOX

"THE CHASM"

INNOVATORS "TECHIES"

EARLY ADOPTERS "VISIONARIES"

EARLY MAJORITY "PRAGMATISTS"

LATE MAJORITY "CONSERVATIVES"

LAGGARDS "SKEPTICS"

Governing Board

Marketing Committee

Technical Oversight Committee

Technical Advisory Groups

End User Community

End User SIGs and User Groups

- Mainly vendors
- Fund the organization
- Marketing and strategic direction

- 11 top technical architects
- Admit new projects
- Acts as a resource to projects

- Real end users of these technologies
- Communicate back requirements
- and good and bad experiences
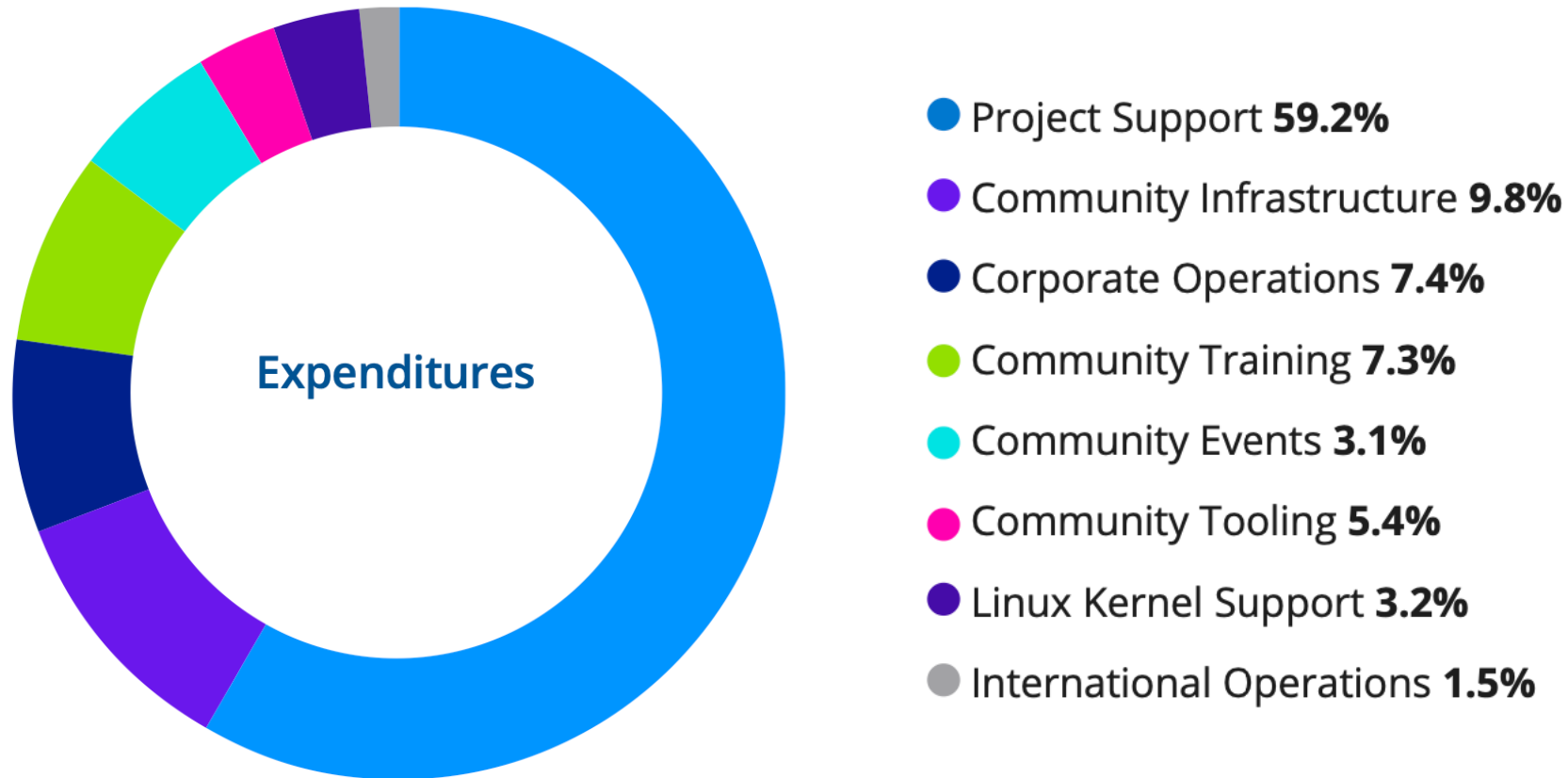
How big are foundations?

LLVM Foundation	~$800,000 / yr
Python Software Foundation	~$2,100,000 / yr
Apache Software Foundation	~$2,200,000 / yr
NumFOCUS	~$5,600,000 / yr
Eclipse Foundation	~$26,100,000 / yr
Linux Foundation (including CNCF)	~$248,000,000 / yr

(mix of 2020-2022 numbers)

Unclear how creating a new foundation can match these influence and leverage

What does the money go towards?

- Depends, but here is an example for Linux Foundation:

Expenditures

- Project Support **59.2%**
- Community Infrastructure **9.8%**
- Corporate Operations **7.4%**
- Community Training **7.3%**
- Community Events **3.1%**
- Community Tooling **5.4%**
- Linux Kernel Support **3.2%**
- International Operations **1.5%**

What's project support?

- Infrastructure (CI, testing, Slack, etc.)
- Project metric tracking (adoption, users, etc.)
- Web development
- Meeting Facilitation
- Documentation
- Legal services
- Marketing
- Certifications and Training
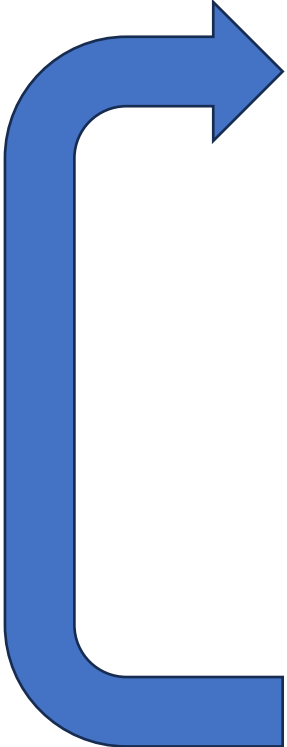- Outreach, Events, Inclusivity initiatives

LF has a very large event organization team
Events like KubeCon attract 10k+ people

What do Foundations NOT fund?

- **Most major feature development!**
  - Foundations are not typically developers for contract
  - *Could* subcontract to other organizations for focused development
  - They're typically experts at community organizing

- **Typically, the *work* of projects is done through in-kind effort**
  - Foundation members are contributors
  - Aim is to coordinate efforts and effectively govern and grow project
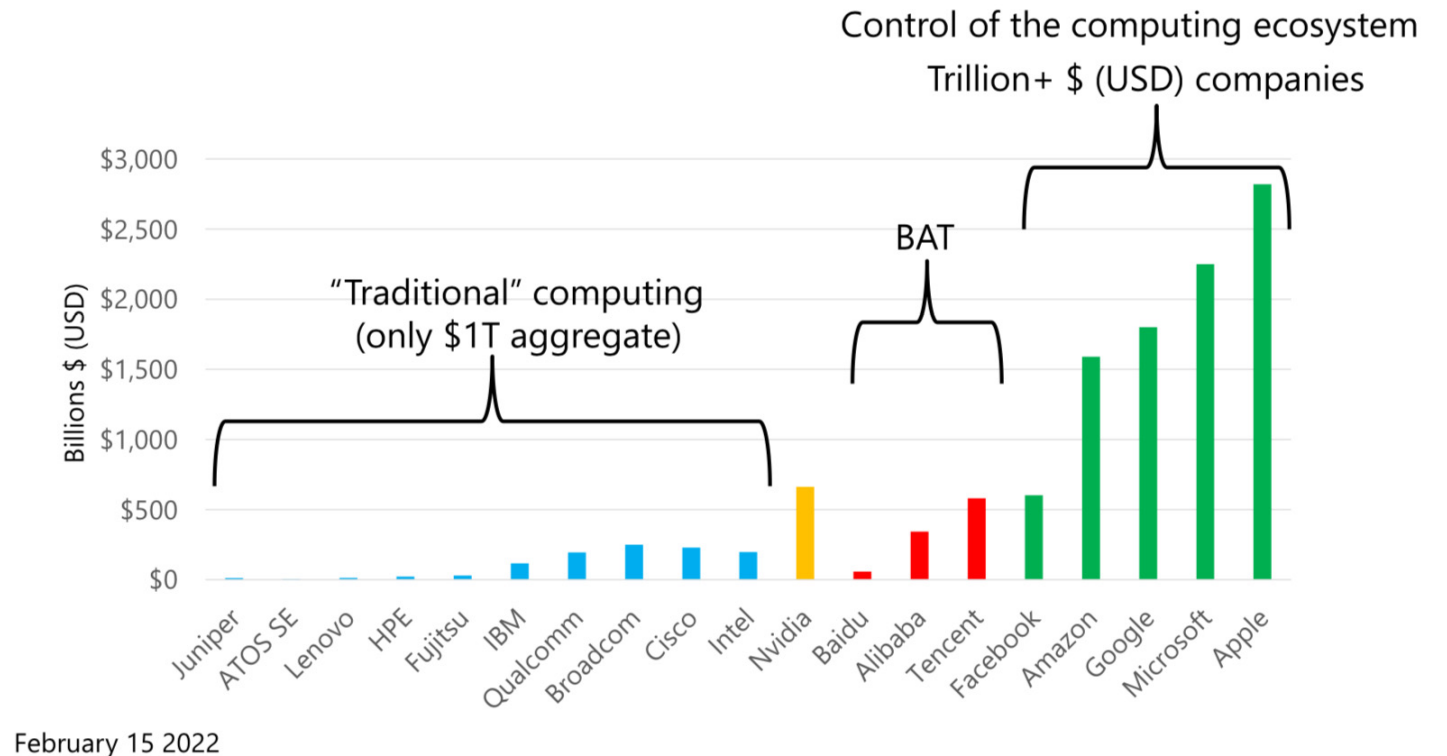 - Sustainment comes from growth

So what can foundations do for DOE?

Growth

1. Growing stakeholders:
   - Open governance allows new stakeholders to help steer theproject
     - Feel some ownership of the project
   - Neutral home allows more companies to engage

2. Growing contributors
   - Outreach, training, and marketing lead to more users
   - Broader user base leads to more contributors

3. More adoption drives more interest and investment

Broader adoption can drive interest from big players

- Adoption has value
  - Customers ask for adopted/used software

- Our software benefits vendor customers.
  - Cost of reinvention is high

- The bigger community you can build, the more likely cloud providers will take interest

Control of the computing ecosystem
Trillion+ $ (USD) companies

"Traditional" computing (only $1T aggregate)

BAT

Billions $ (USD)

February 15 2022

Myth: "There are only a few people in the world who can contribute to our software"
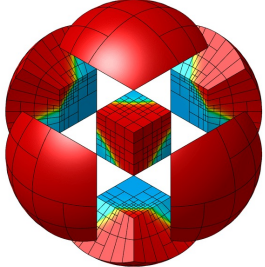
- The open source community is huge, and casting a wide net can *find* you more contributors

- One of the libraries I've heard this claim for is MFEM
  - MFEM had 100+ attendees at a user group meeting
  - Community continues to grow

- AWS recently built PALACE, an open source quantum code, *on top of our MFEM library*
  - AWS developers have contributed some of their work back to MFEM

HPC wire

Search... Go

AWS Debuts Palace – Finite Element Analysis Suite Fine-Tuned for the Cloud
By John Russell

February 21, 2023

In an interesting twist on quantum-inspired work making its way into traditional HPC – and in this case a step further into cloud-based HPC – AWS today introduced Palace, short for PArallel, Large-scale Computational Electromagnetics, a parallel finite element code for full-wave electromagnetics simulations. Palace was first developed at the AWS Center for Quantum Computing to perform large-scale 3D simulations of complex electromagnetics models in the design of quantum computing hardware. While Palace can be used in quantum hardware design, AWS expects it to be used in a wide range of simulations.

Announced in a blog today, AWS researchers[i] wrote, "We are making Palace[ii] freely available on GitHub as an open-source project for electromagnetic modeling workloads, not limited to

MFEM

Foundations can exist alongside PESO as an outreach and community building arm

Contribute to and influence project communities

Drive industry and community interest

Manage ASCR, agency, funding for core project effort

Manage engagement w/CoPs

Final thought: An umbrella for HPC

- HPC in the cloud is growing at 17% CAGR
  - Compare to HPC CAGR of 7%
- Cloud HPC is a *growth area* for industry
- We have the expertise in HPC
  - Accelerated computing
  - "100X" performance
- Making it work *in the cloud* grows our communities and contributor base

- **CNCF's mission is to create a portable reference stack for cloud-native services**

- ***We* could be the portable reference stack for cloud HPC**