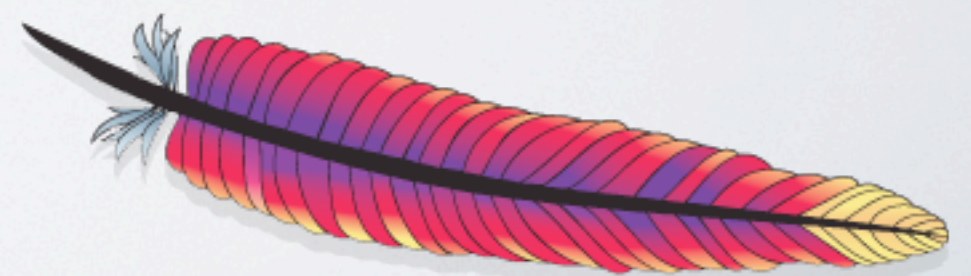# WELCOME!

ΛPΛCHE:

BIG_DATA

EUROPE

Up Next: STATE OF THE FEATHER

by: Shane Curcuru, Vice President, Brand Management
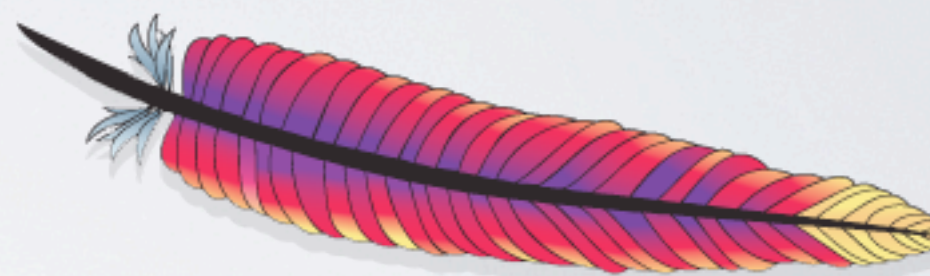The Apache Software Foundation

Good morning, Budapest!
My name is Shane Curcuru, and I'd like to welcome everyone to Apache: Big Data Europe. We are very excited to have you all here and thank you all for attending.

We'd also like to thank all the sponsors for our event here in the lovely Corinthia Budapest, without whom this event wouldn't have been possible.

Our Diamond Sponsors are: HortonWorks, IBM, and Jahia, our Platinum Sponsor: Pivotal, and our Silver Sponsors, Criteo Labs and Red Hat. Thank you event sponsors!
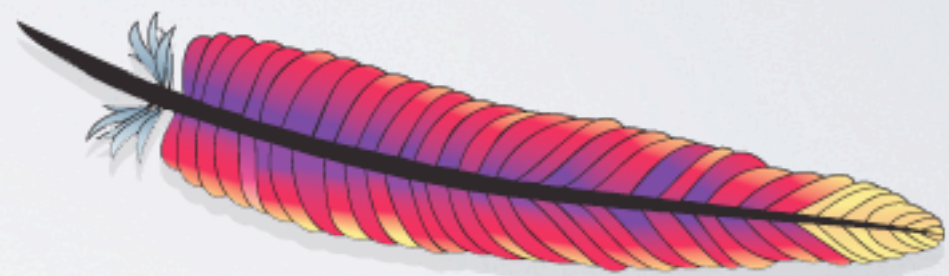
STATE OF THE FEATHER

Shane Curcuru

Vice President, Brand Management
The Apache Software Foundation

Hi - I'm Shane Curcuru, and I'm the Vice President of Brand Management for the Apache Software Foundation. For our first keynote, we have the State of the Feather - a very brief overview of the **growth and successes of the ASF over the past year**, to help you understand what Apache as a whole is doing these days.

But first, I want to make one small correction on this slide... **See that title**? It's really not that important. The Apache Way is to allow everyone who productively contributes to the project a voice and a vote - we have a very flat organization.

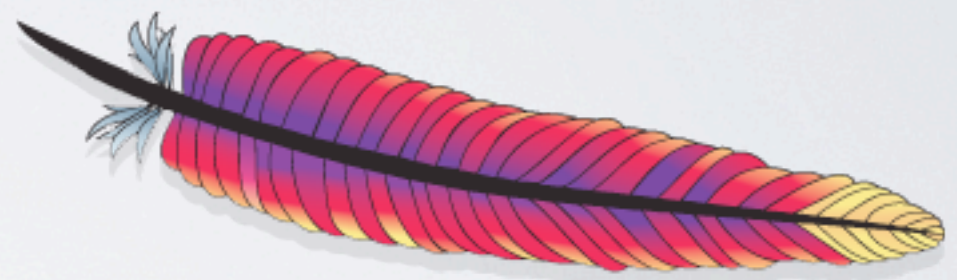**So let's make that slide say what I really am:**

STATE OF THE FEATHER

Shane Curcuru

Volunteer

A Volunteer. **A key differentiator** for the ASF and our Apache way is that all the governance positions and all the project direction decisions are done by volunteers within each community.
The i**ndividuals doing the work** in any one particular project - or within the foundation as an organization - are the ones setting the direction for that project.
Obviously many committers are being paid by their employer as an investment in various project's technologies, but from the Apache point of view, we're all volunteers.

The Apache Software Foundation (ASF), a
non-profit public charity provides
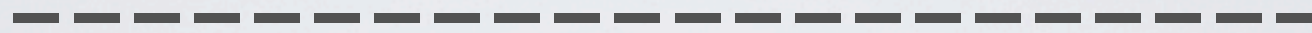**software at no charge**
for the **public good**.

The ASF does this by **providing support** to
our many **independently governed** Apache
project **communities of volunteers**.

What is Apache's true mission? Our mission is to provide software at no charge, for the public good. We **provide useful software** for people - to help the world. Software Apache produces and gives away for free is used by people everywhere around the world, every day, to help humans do more work with computers. And it's critical to **ensure this software is freely available** and is truly useful to all users, everywhere.

Now the thing that many people don't think about is that the Foundation itself doesn't directly work on this mission It's the many Apache projects that create each of the various software products, We accomplish our mission by providing support to the many independently governed Apache project communities.
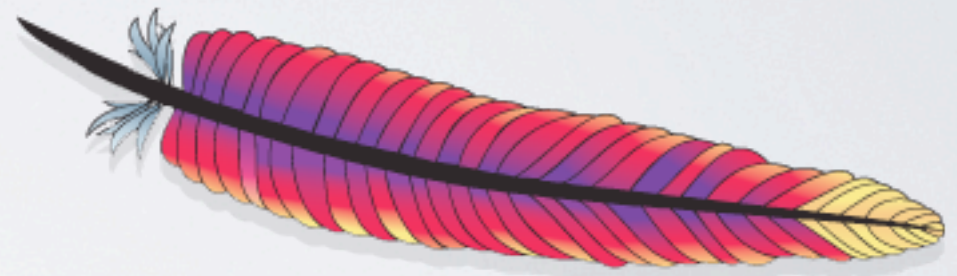
The ASF is here to take care of the paperwork, licenses, trademarks, and provide services to our communities. We ensure the **legal framework is taken care** of so they can concentrate on the code.   We don't guide or direct the projects - we allow the projects to govern themselves and choose what they want to work on, and provide an environment where they can do that work.

COMMUNITY

CODE

We consider ourselves not simply a group of projects sharing a server, but rather a **community of developers and users**.

Project communities set their own direction. The ASF exists to support them.

We have a saying: **Community Over Code.** Think of this as the **secret sauce** of open source at Apache -The Apache Way of doing things.
For Apache, the most important thing in a project is the community, and every project community is different. Ensuring that the individuals actually doing the work in a project have the say over project direction ensures that their goals and actions are aligned.

Now in a volunteer-based organization, we should be sure to give lots of **thanks**.

THANK YOU SPONSORS!

Platinum Sponsors

Since we're about community, let's take the time to thank an important part of our community - our sponsors.
These are all our **Platinum** sponsors, who sign up to donate a fixed amount annually to our budget. We're lucky to have many **recurring sponsors** who continue to come back and provide us with a stable funding base, both many large corporations, as well as at least one personal individual (in the upper left).
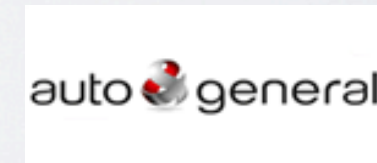
Sponsors also give us **resources within our communities** - most of these organizations also assign staff members to contribute to Apache projects because they are using many Apache software products.
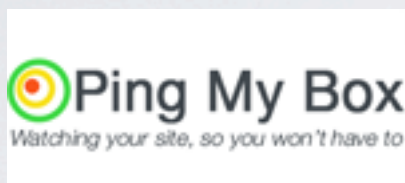
THANK YOU SPONSORS!

Gold

Silver

Bronze: Cloudsoft Corporation, Accor, Basis Technology, Bluehost, Talend, Samsung, Twitter

Gold, Silver, and Bronze Sponsors

Thank you to our Gold, Silver, and Bronze sponsors, all of whom donate funds for our income. Even though the ASF only does **organic fundraising** with volunteers, we've been seeing a good growth in the number of **new or upgraded sponsors**

What's interesting here is we're starting to see sponsors sign up who are **not traditional software shops**. We have a small but growing number of companies and individuals who just want to give back in recognition of all the value they've seen from **using Apache software project**s.
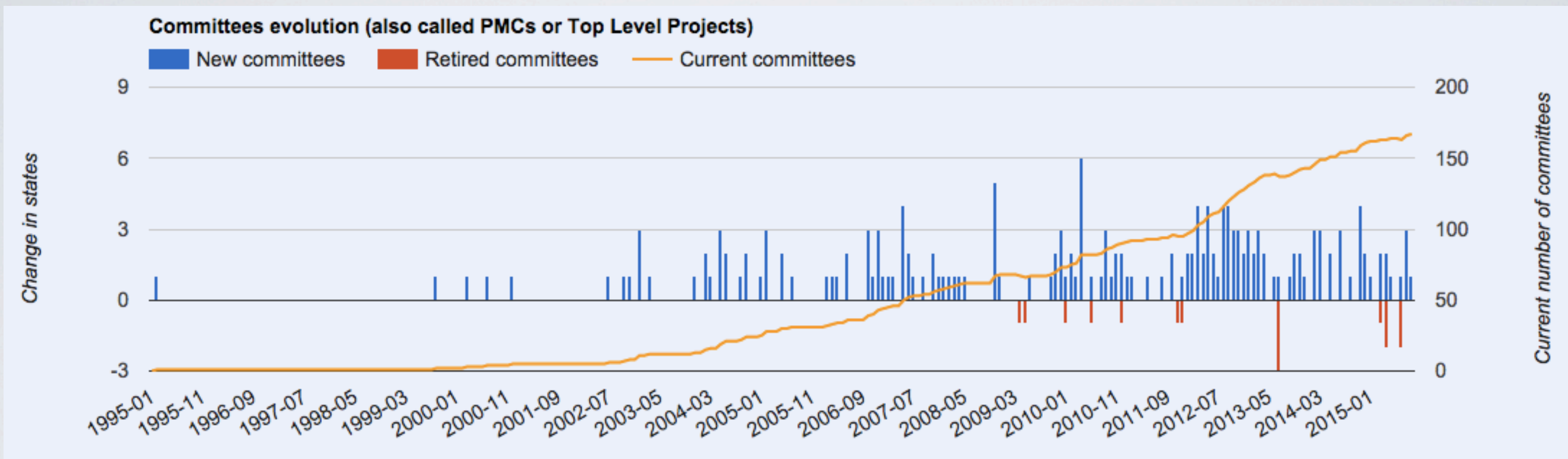
THANK YOU SPONSORS!

Infrastructure
Sponsors

Plus: thanks to the many **volunteers**
who help our crack infrastructure team!

The last set of sponsors we need to thank don't give us money, but something just as important: **bandwidth**.
All of these service providers donate free bandwidth, hosting, system administration, and site design and maintenance services, along with **free cloud hosting credits**, for the Foundation and to various projects.
These sponsors provide the services that allow our small paid staff to **leverage** their skills to provide services to all projects.
We also want to thank the many **Apache volunteers** who also help build and run many of our project systems - they add a great deal of expertise and work to our small core of paid staff.
Let's have a round of applause for all of our sponsors.

PROJECTS GROWTH

Committees evolution (also called PMCs or Top Level Projects)
New committees | Retired committees | Current committees

The ASF provides services for the entire project lifecycle:

43 Podlings (upcoming projects)

167 active Apache projects

25 dormant Attic projects

So what has Apache been doing this past year? We've been continuing our **steady growth in new projects** for the last ten years in terms of total software projects governed. Each **blue line** denotes a new top level project that came to Apache, and the yellow line shows the total number of Apache projects over time.

And very importantly for our mission, Apache We cover all parts of **project lifecycle**:
New podlings - or new initiatives or projects **applying to join the Incubator** - are where we provide mentoring and work with project communities on independent governance. What you might not realize is that the primary work in the incubator is **stewardship of the community**, not code completion.

Once a podling or new project shows they have a strong and diverse community, we announce them as a Tope Level Project, or TLP.

Mature technologies not being actively developed are moved to the attic: the non-profit nature of the Foundation will ensure that the project's assets will be h**osted and publicly available for the long future.**

TECHNOLOGY DIVERSITY

• Big Data

• Cloud API & management

• Content Management

• Database & NoSQL

• Web & mobile development

• Messaging

• Geospatial / Scientific

• Dev / Build Tools

• Network servers

• OSGi

If you have the community, Apache is happy to host **any** technology!

Here at Apache: Big Data conference, it's important to remember that Apache projects span virtually all software technology areas today.
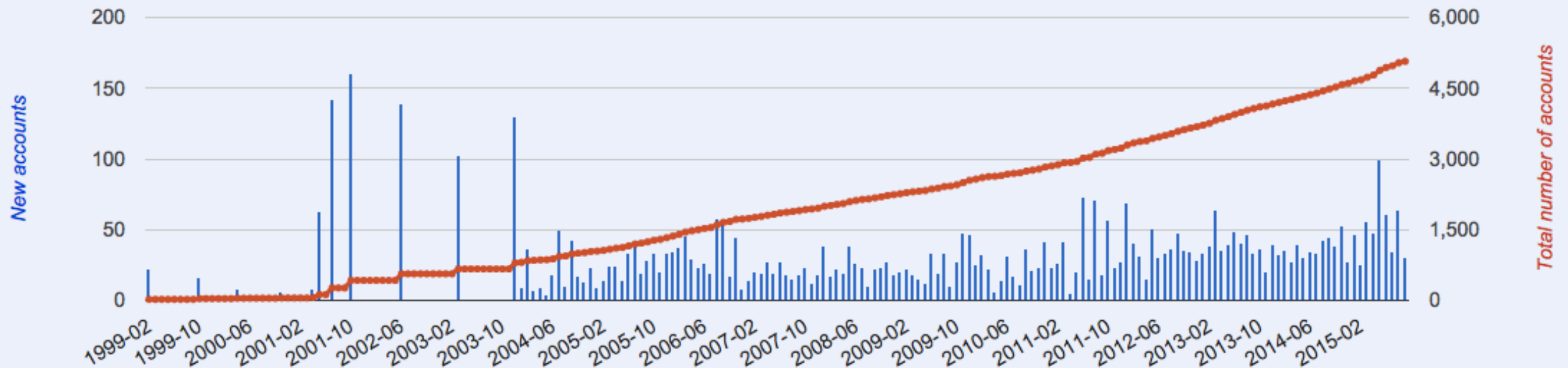
Along with more than a **dozen major big data projects**, Apache has multiple projects in all of these technology areas - in many cases, industry leading technologies. We also have a broad array of big data and other podlings, or new project initiatives.

 The **expertise of volunteers** in our many Apache project communities spans software - and beyond.

COMMITTER GROWTH

Account creation timeline

Apache Committers are:

Software Engineers, Sysadmins, Database Administrators, Network Architects, Developer Advocates - all the technical titles you might imagine, plus...

✦ CIO of a major financial company

✦ Airline Pilot of a national airline

✦ Technical Writer at software vendor

✦ Space Flight Dynamicist at aerospace vendor

✦ Principal Scientist at NASA

But we're not just growing in projects, we're growing in volunteers. Again, we can see a **steady growth** in new Apache committers working on one of our projects, all the way back to 2000.
We're now around **5000 committers -** or people who have a direct say and write access to one or more Apache projects.
While plenty of our committers are coders as you might expect, we have a growing number of committers who's **skills are not the traditional coding or software positions.**

WHERE WE ARE

Apache is a truly global organization. Not only is our software used around the world - as well as in several spacecraft out in the solar system today! - but our committers come from around the world.

Global - yes, that includes **New Zealand**. Also note strong growth in **India/China** as open source development catches up their to their existing open source use.

Note: we still don't have global coverage in South America or Africa - we can still do better, especially with non-native english speakers, and with developers who might be building a different kind of application than we might here.
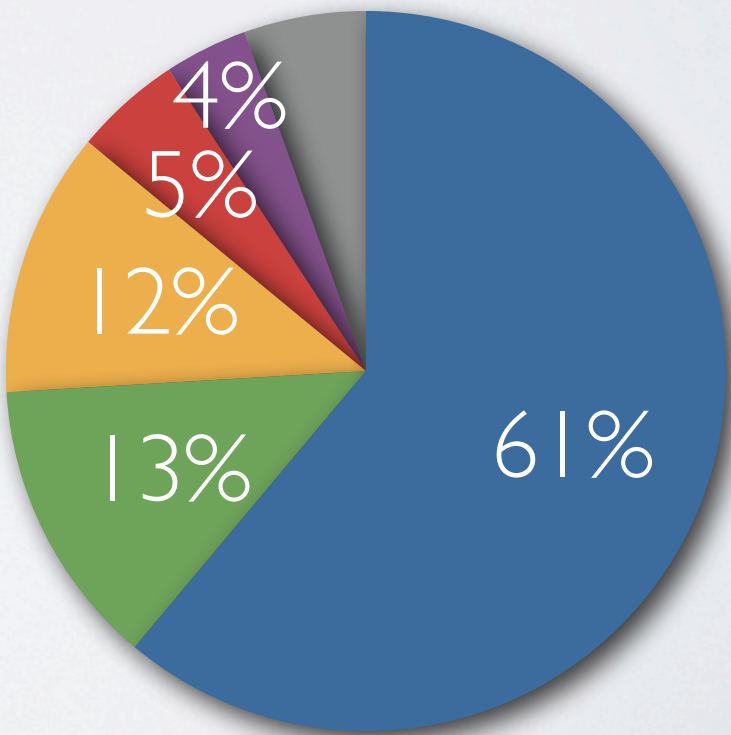
ANNUAL BUDGET

Budgeted Income 2015

Sponsors	$815,000
Individual Donors	$60,000
Other Income	$17,000

Year to Year Total Budget

2013	$774,000
2014	$811,000
2015	$1,186,000
2016	$1,043,000

(Figures rounded to thousands)

Budgeted Expenses 2015

🔵	Infrastructure	$636,000
🟢	Administration	$135,000
🟡	Media & Publicity	$124,000
🔴	Travel Assistance	$50,000
🟣	Brand	$39,000
⚪	Other	$57,000

Pie chart: 61%, 13%, 12%, 5%, 4%

So - how much does the foundation spend to **keep all this running**? Here's a quick breakdown of our current fiscal year budget. Our **primary income** is annual donations from our formal sponsorship program, although we have a good number of smaller one-time individual donors of many kinds.

Our **primary spending** - as you might expect - is on infrastructure, which powers the servers and services to keep our projects running. We also provide a number of **other paid services** for projects that ask for them. Now these are **real numbers:** our current annual budget is just over one million dollars US.

Now compare the million dollars we spend on infrastructure to the total value of Apache project code in just your organization. Pretty big, huh? Imagine that for everyone in your industry, and around the world.   **Bang for buck is incredibly high.** We say Apache is an innovation machine, and we turn a small amount here into a big value for the world.

One important note: Apache **never spends budget on project code development.** All the code our great Apache projects produce is created by our volunteer committers..

You can see our total spending budget has been increasing year over year to keep up with growth - **kudos to our infrastructure** team for finding the right mix of cloud services versus hardware to actually slightly reduce our budget this year while still providing more services.

INFRASTRUCTURE

- 194 VMs / servers in 7 colos & 8+ clouds

- 120 different technical services provided to projects & to public

- **5 paid staff to manage everything**

  - 10,000 INFRA tickets closed in past 12 years

- 26TB <u>weekly</u> traffic (**not** including mirror downloads)

  - 13,600,000 <u>weekly</u> unique visitors

- ~1 billion emails sent <u>annually</u>

http://status.apache.org/

How do we support all of this code?
Our amazing infrastructure team - of only 5 paid staff - provides almost **200 machines** and VMs and over 100 services. Services here include everything to manage a software project, in some cases including multiple versions of tools like bugtrackers and cloud environments.
The **26TB of weekly traffic** does not include software downloads, which are handled by our two mirror systems.

And that **almost billion emails** annually is just ham, not spam!

But those are just big numbers - what do they mean?

INFRASTRUCTURE

ON **ONE DAY** LAST WEEK,
Apache websites & services served
2,245,423 unique users
from more than
11,454 cities in
206 countries/states *

* all 196 UN states,
plus territories

How can we measure a **direct impact on the public good** from what Apache does?

One of our **infrastructure ponies** counted up all the server logs from one day last week and found:
In just one day last week, over two million unique users accessed data from the Foundation or an Apache project from every single UN recognized country on the planet. Not only does the software Apache gives away power half of the internet, developers and users alike around the world use our project websites for reference and help every single day.

These are people **learning about or working** on Apache projects - remember, this doesn't show how many people are *using* Apache software everyday.

MEDIA RELATIONS

- Total Press releases: 12
- Informal Announcements: 230
- Industry Events, Awards, and Outreach liaison: 9
- Media Requests: 46
- Analyst Briefings: 51
- Analyst Reports: 202

Plus: new Annual & Quarterly Reports,
major site redesign for www.apache.org

Our Marketing & Press team assists all Apache projects with publicity and analyst relations, as well as fielding and routing queries for the Foundation itself. Along with assisting projects with press releases and making contact with analysts, the press team has done an amazing job on our **annual report**, as well as modernizing and i**mproving the overall foundation website.** Note that this is one of the few paid positions within the ASF. Providing timely responses as well as assistance **translating between the many engineers** of our projects and the analyst and press community is a critical service that we can **provide to our projects**.
Note: since most of our projects are made of engineers, it's key to have someone who can help translate for the press and analysts.

OTHER SERVICES

- **Travel Assistance Committee**
- **Legal** and licensing support
- Brand management & **trademark** registration
- Executive Assistant & corporate Secretary
- Foundation-wide **fundraising**
- Apache **swag** (giveaways) at events to build community
- Volunteer-provided services:
  - Community Development
  - Incubator **Mentoring**
  - Security Team Coordination

The Foundation provides a number of other housekeeping services - the kinds of things that are important to get done on time to keep the organization running.

Travel assistance committee - pays to bring in newer project contributors to help strengthen our communities.
...
Hidden value- our volunteer services - not in our budget because they're not paid - include highly expert mentoring and community development for many different podlings and projects.

WANT REGULAR UPDATES?

• Monthly board reports are posted publicly
https://www.apache.org/foundation/board/calendar

• Subscribe to announce@apache.org

• Read https://blogs.apache.org/foundation/

• Follow @TheASF

Apache projects have their own blogs, mailing
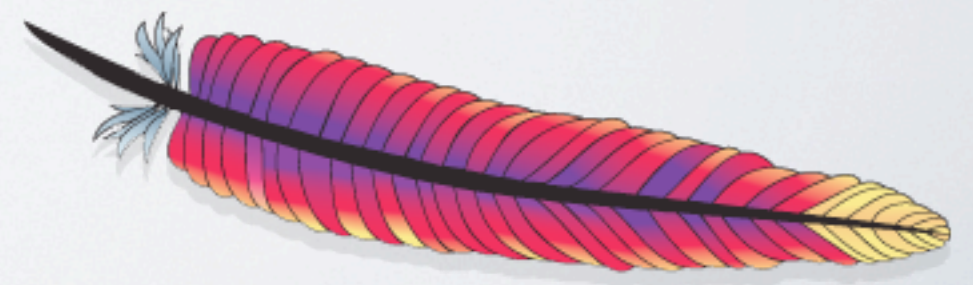lists, and twitter handles too!

If you'd like to keep tabs on what the foundation is doing, we have plenty of ways to do that.

Remember - the Foundation provides support to our hundreds of projects - **each of which has their own ways of keeping in touch.** The best way to learn about your favorite project is to join their mailing list!

LEARNING THE APACHE WAY

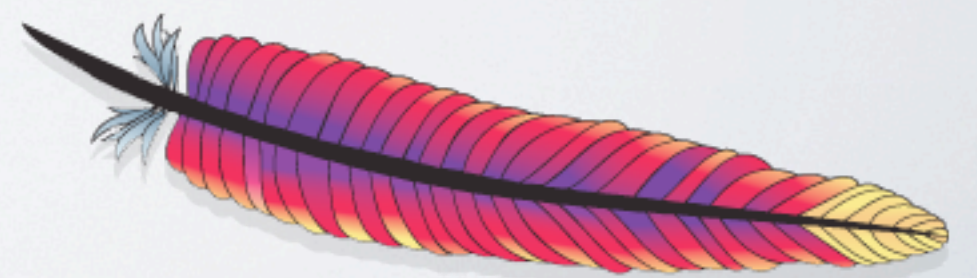Want to learn more about the secret sauce of the Apache Way?

- Wednesday, 11AM: Profiting From Apache Projects Without Losing Your Soul - Shane Curcuru

- Thursday, 10:30AM: The Apache Way
  Rich Bowen

- Thursday, 11:30AM: Open Source At Scale
  Bertrand Delacretaz

- Community Development project:
  http://community.apache.org/

If you want to learn more about the Apache Way and our **secret sauce f**or effective open source, here are starting points this week.

APACHE STAYS THE COURSE

OUR GOAL IS to continue to provide services and mentoring to our collaborative open source projects, for the **next 50 years**.
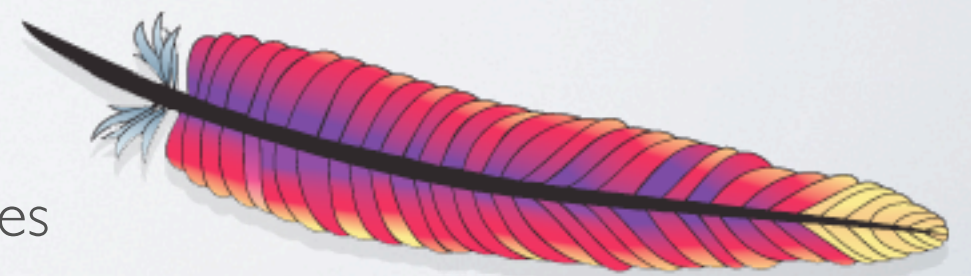
Our mission is to be here in **50 years**, supporting our projects. That's not a typo: we are truly building an organization to continue to provide for the public good for decades to come.

THANK YOU!

APACHE:

BIG_DATA

EUROPE

Up Next:
BIG SCIENCE AND BIG DATA AT CERN

by: Dirk Duellmann, Deputy Leader for Data & Storage Services
CERN

Our next speaker is Dirk Duellmann, deputy leader of the data and storage services group in the IT department at CERN.
Dirk will be discussing CERN's large scale data management and analysis infrastructure for physics data at the Large Hadron Collider,
comparing the established scientific workflow with more recent deployments of Apache Hadoop ecosystem components at CERN.
Please welcome Dirk.