ApacheCon
Europe 2009

Conference Program

[Updated 20090313]

23-27 March • Amsterdam

Leading the Wave of Open Source

http://www.eu.apachecon.com

Presented by: The Apache Software Foundation
Produced by: Stone Circle Productions, Inc.

SPONSORED BY

HIPPO
Gold

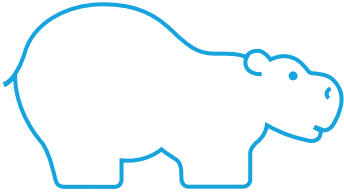cloudera   Google   hp invent   spring source   YAHOO! DEVELOPER NETWORK

Silver and Bronze

**2009: Ten Years of The Apache Software Foundation
Celebrating a Decade of Open Source Leadership**

Thank You to Our Sponsors

Gold Sponsor

HIPPO

Silver Sponsors

Google™

hp invent

spring source

YAHOO! DEVELOPER NETWORK

Bronze Sponsors

cloudera

Joost™

t thawte
it's a trust thing™

Community Sponsors and Exhibitors

The Apache Software Foundation
http://www.apache.org/

HALO Worldwide

Hotwax Media

lucid IMAGINATION

Progress FUSE
fusesource.com

topicus

Community and Media Partners

CrowdVine Social Networks

CAcert

Feather Cast
http://feathercast.org/

German Unix User Group

WHERE JAVA PEOPLE MEET
June 22-25, 2009 Zurich
JAZOON

JTEAM

jWeekend
Training, Consulting, Development

LINUX MAGAZINE

MANNING PUBLICATIONS

Methods & Tools

NLUUG

WifiSoft.org

Produced By

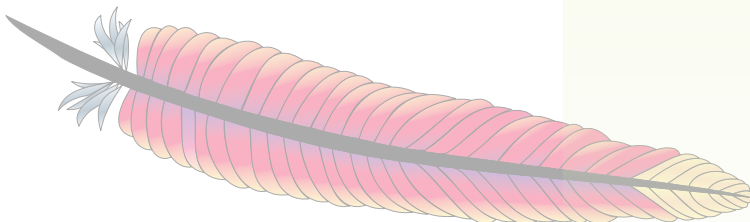STONECIRCLE PRODUCTIONS, INC.

Welcome to ApacheCon Europe 2009

Celebrating the 10th Anniversary of The Apache Software Foundation

Congratulations to The Apache Software Foundation on ten years of Open Source Leadership! It is an honor and a privilege working with all of you in the Apache Community. We, at Stone Circle Productions, look forward to another ten years of creating and producing conferences that support and promote the Apache Way.

From the team at Stone Circle Productions,

Charel, Delia, Susan, Kirsten, David, Luke, Nate, Richard, Neil and Ping

STONE CIRCLE
PRODUCTIONS, INC.

Sponsors

Hippo
Gold Sponsor + Exhibitor

Hippo develops the Apache licensed Hippo CMS and Hippo Portal application. Hippo is an avid believer of the Apache way of developing, and participates pro-actively in ASF Projects. Hippo's developers are stimulated to get involved in Apache in order to become committed in relevant projects, sharing knowledge with others in true Apache spirit. You can find code initiated by Hippo people in many projects, amongst others: Apache Cocoon, Apache Jetspeed and Apache Slide. Hippo is a member of the Java Community Process and its developers help define new Java specifications.

Hippo CMS is already widely used in the Netherlands and across Europe. With the acquisition of the assets of Blue Sunrise, another great company active in the Apache community, Hippo obtained a strong threshold into the USA. Hippo is now well on the way to become a household name in the US market space, and is rolling out vastly to set the global standard in open source content management and portal applications.

For more information, visit the company website at http://www.onehippo.com. or get involved in the Hippo community at http://www.hippocms.org.

Google
Silver Sponsor

Google is a proud user and supporter of open source software and development methodologies. Google contributes back to the Open Source community in many ways, including source code, project hosting on Google Code, projects for students including Google Summer of Code and the Google Highly Open Participation Contest, and support for a wide variety of projects, LUGS, and events around the world. Learn more at code.google.com/opensource/

Hewlett-Packard
Silver Sponsor

HP Open Source and Linux solutions are designed to solve your mission-critical IT problems today! We sell over a million Linux servers a year, delivering high performance, scalability, security, storage, and backed by our commitment to the open source community. Learn more about our market leading Linux platforms and Open Source Middleware Stacks at www.hp.com/go/linux. Also be sure to check out our new FossBazaar and Fossology initiatives which are aimed at further encouraging the adoption of Open Source in the enterprise. Our solutions are delivered and supported by over 6500 Open Source and Linux professionals across 160 countries. Come visit us at our booth to hear how HP and the Open Source community are working together.

SpringSource
Silver Sponsor + Exhibitor

SpringSource builds Java infrastructure software which eliminates the complexity of enterprise Java. SpringSource created Spring, the de facto standard platform to build, run and manage enterprise Java applications. SpringSource also employs the leading committers for Apache Tomcat and is the leading Tomcat support provider. SpringSource software has been designed from the ground up for complete portability and scalability for today's data center and for next-generation virtualized, grid and cloud deployments. Nearly half of the Global 2000, including many of the world's leading retail, financial services, manufacturing, healthcare, technology and public sector clients are SpringSource customers. For more information visit: www.springsource.com.

Yahoo!
Silver Sponsor

Yahoo! Inc. is a leading global Internet brand and one of the most trafficked Internet destinations worldwide. The Yahoo! Developer Network offers Web Services and APIs that make it easy for developers to build applications and mashups that integrate data sources in new ways, making the web a more useful and fun place for everyone. Yahoo! is headquartered in Sunnyvale, California.

Cloudera
Bronze Sponsor

Cloudera provides commercial support for Apache Hadoop clusters running in the cloud or on-site. Our team includes Hadoop-core developers and founders responsible for some of the world's largest clusters. We're hiring.

Joost
Bronze Sponsor + Hackathon

Founded by Janus Friis and Niklas Zennström, Joost brings video and social interaction together online. All of the video on Joost is professionally produced and streams on-demand to people around the world for free in a high-quality format. Today, Joost has more than 57,000 program elements, including more than 400 television series and 1,200 movie and short film titles, and 18,000 music videos. For more information, visit www.joost.com.

Sponsors

thawte
Bronze Sponsor + Hackathon

thawte is a leading global Certificate Authority (CA). Our SSL and code signing digital certificates are used globally to secure servers, provide data encryption, authenticate users, protect privacy and ensure online identities go through stringent authentication and verification processes. Our SSL offerings include Extended Validation (EV), Server Gated Cryptography (SGC), SANS, and wildcard certificates. The focus of *thawte* is to allow companies to establish a trusted relationship with anyone, anywhere on the Internet. *thawte* is committed to the egalitarian ethos of the Internet and is dedicated to ensuring world-class multi-lingual customer service, and providing outstanding value for money. For more information, go to https://www.thawte.com

Community Sponsors + Exhibitors

The Apache Software Foundation
Community Sponsor

Established in 1999, the all-volunteer Foundation oversees more than sixty-five leading Open Source projects, including Apache HTTP Server — the world's most popular Web server software. Through The ASF's meritocratic process known as "The Apache Way", nearly 300 individual Members and 2,000 Committers successfully collaborate to develop freely available enterprise-grade software, benefiting millions of users worldwide: thousands of software solutions are distributed under the Apache License; and the community actively participates in ASF mailing lists, mentoring initiatives, and ApacheCon, the Foundation's official user conference, trainings, and expo. The ASF is funded by individual donations and corporate sponsors including Google, HP, Microsoft, Progress Software, SpringSource, and Yahoo! For more information, visit http://www.apache.org/

HALO Worldwide
Community Sponsor

HALO Worldwide is a consulting firm specializing in promotion, publicity, and brand integration programs that span strategic communications and positioning, brand planning and translation, public relations and outreach, and image management. HALO principals have gained a reputation for being among the most competent in the business, leveraging more than 30 years' combined out-of-the-box thinking, intuitive marketing, and media relations expertise to develop highly effective campaigns across a range of media resources. With a clientele that ranges from startups to blue chip organizations, we are passionate about managing reputations, driving brand awareness and recognition, and developing valued relationships with credibility and gravitas. http://www.haloworldwide.com/

Hotwax Media
Community Sponsor

As the global leader in Apache Open For Business (OFBiz), HotWax Media provides web-based enterprise resource planning (ERP), e-commerce, and business automation consulting services. From enterprise-level e-commerce to warehouse, manufacturing, and financial management, HotWax Media and OFBiz help optimize business processes. Open source is the smart choice today for businesses ranging from small-and medium-sized organizations to global Fortune 100 companies. Using Apache OFBiz, HotWax Media delivers top quality enterprise software for a fraction of the price of outdated, proprietary vendors. For more information visit www.hotwaxmedia.com.

Lucid Imagination
Community Sponsor

Lucid Imagination is the first commercial entity exclusively dedicated to Apache Lucene/Solr open source technology. As an active participant in the enormous community using Lucene/Solr, Lucid Imagination offers certified distributions of Lucene and Solr, commercial-grade support, training, high-level consulting and value-added software extensions. The company's web site will serve as a knowledge portal for the Lucene community, with information and resources to help developers build and deploy Lucene-based solutions in a more efficient and cost-effective manner. The Lucid Imagination founding team consists of several key contributors and committers to the Lucene project, as well as experts in enterprise search application development. A privately held company, Lucid Imagination is based in San Mateo, CA, USA.

Progress Software
Exhibitor

The FUSE Community (www.fusesource.com) was created to help users adopt and deploy Apache ServiceMix, Apache ActiveMQ, Apache Camel, and Apache CXF in enterprise environments. The FUSE distributions are tested, certified versions of these popular Apache projects and are backed by world-class support and services. Progress Software employs many of the key committers to these projects.

topicus
Community Sponsor

Topicus is an innovative software house with about 150 employees comfortably situated in the historic center of Deventer, the Netherlands. Topicus specializes in supply chain integration, building Software as a Service applications and process management in finance, healthcare and education. Within each sector Topicus delivers new service concepts utilizing Apache technology such as Wicket, Tomcat, Maven and httpd to the fullest. The last five years Topicus was amongst the 50 fastest growing technology companies in The Netherlands and received the Fast 50 Sustainable Grower award in 2008.

Sponsors

CAcert
Community Partner

CAcert.org is a community-driven certificate authority that issues free public key certificates to the public (unlike other certificate authorities which are commercial and sell certificates). CAcert has over 110,000 verified users and has issued over 350,000 certificates as of July 2008. These certificates can be used to sign and encrypt email, authenticate and authorize users connecting to websites and secure data transmission over the Internet.

CrowdVine
Community Partner

CrowdVine, founded in 2007, builds simple and powerful social software to help people connect. While we believe our products are very good, the killer feature of social software is people. That's why when you go to our features page, you'll see a list of reasons to use our software, not just a list of tools. Everything we do is aimed at making it easier for people to meet. Once they've done that, it's our job to get out of the way.

GUUG
Community Partner

The German Unix User Group, founded 1984, is a non-profit organization for professional system and network administrators, security experts and system programmers preferring to use a Unix operating system. This includes free systems like Linux, BSD and OpenSolaris variants as well as commercial Unix systems. Important to us are open and well defined interfaces and APIs, the compliance with internet and other open standards and last but not least: the chance to have a look deep into the system. Since more than 20 years GUUG organizes national and international conferences where well known experts present cutting edge topics and attendees can get professional training. http://www.guug.de/

Feather Cast
Media Partner

Created in 2006 by Apache Software Foundation (ASF) Members/co-hosts Rich Bowen and David Reid, FeatherCast highlights the Foundation's various people, projects, and plans. The independently-run series of podcasts draws developers, users, Open Source enthusiasts, and general IT audiences to hear perspectives on activities relevant to the Apache Community. More than 40 leading voices have contributed their insight on popular topics such as ASF-incubated projects, ApacheCon, ActiveMQ, and Google's Summer of Code. With episodes added regularly, catch future FeatherCasts on new ASF Members, intellectual property issues, WS-*, and more at FeatherCast.org and via iTunes.

Jazoon
Community Partner

Jazoon'09 The International Conference on Java Technology will take place in Zurich, Switzerland from 22 to 25 June 2009 and will again attract famous Java experts and geeks to Zurich. In contrast to other events, Jazoon places emphasis on the active participation of the community and in turn invites everyone to contribute lectures on interesting research results, developments and practical experiences.

JTeam
Community Partner

JTeam creates customized, business-critical software based on light weight Java and Open Source for the Benelux market. We specialize in Enterprise Search, E-commerce and CMS solutions, applying Agile principles and a focus on skills-transfer and business-case driven implementation for diverse clients such as Ilse Media, Randstad, The National Tax Authority, Zenggi.com and ABN Amro. JTeam is the sister company of SpringSource, and is known worldwide for its expertise on JEE, Spring, Hibernate, DWR, Maven, Compass, Lucene, Solr, GWT, Grails, Hippo and other Open Source technologies and products.

For more information please visit us at: www.jteam.nl or stop by JTeam's Apache Solr Case Study on Thursday evening at 17:30 in the SOA Track.

jWeekend
Community Partner

jWeekend provides consultancy, bespoke software development and training in OO and Java technologies including Wicket, JPA, Spring and Swing - with some weekend and out of hours courses available! We serve organisations ranging from the largest investment banks, telcos and government departments to small software houses, consultancies and family businesses. Our people are renowned in the industry for real world commercial development as well as presentation and mentoring acumen. We are based in London and work with clients globally to deliver high quality systems and training solutions. Our new Global Training Partner Scheme announced this month adds a new dimension and will give more developers around the world the opportunity to experience our unique courses: http://www.jWeekend.com

GET HIPPOTIZED

The Hippo ECM stack gives you all the tools you need to develop large scale content driven web sites, portals and intranets. Based on the Apache Java stack*, and available under the Apache Software License, it gives you all the freedom you need. And if you or your customer needs any help, call Hippo. We'll support you.

* Jackrabbit - Wicket - Jetspeed - Cocoon - Lucene - MyFaces - Tomcat

HIPPO

Sponsors

Linux Pro Magazine
Media Partner

In Linux Pro Magazine you'll find the tools, tutorials, reviews, and concrete technical discussions you'll need to unlock the secrets of Linux - and not just from the server side. Linux Pro delivers solutions for real users with Linux on real desktops. http://www.linuxpromagazine.com/

Manning Publications
Community Partner

Manning Publication publishes computer books for professionals--programmers, system administrators, designers, architects, managers and others. Our focus is on computing titles at professional levels. We care about the quality of our books. We work with our authors to coax out of them the best writing they can produce. We consult with technical experts on book proposals and manuscripts, and we may use as many as two dozen reviewers in various stages of preparing a manuscript. The abilities of each author are nurtured to encourage him or her to write a first-rate book.

Methods and Tools
Media Partner

Methods and Tools is a free magazine providing practical knowledge for the software developer, tester and project manager http://www.methodsandtools.com/

NLUUG
Community Partner

NLUUG is the association of (professional) Open Source and Open Standards users in the Netherlands. Since the late seventies, the NLUUG has brought together the community of systems administrators, programmers, researchers and IP network professionals. The primary goal of the NLUUG is to extend the application of, and knowledge about, open systems and UNIX.

WifiSoft.org
Community Partner

The WifiSoft.org Foundation is a not-for-profit organisation that initiates and stimulates local and international cooperation between open communities. WifiSoft.org achieves and facilitates cooperation between Open Communities via the Internet and the organization of events. We organize and facilitate lectures and hands-on workshops.
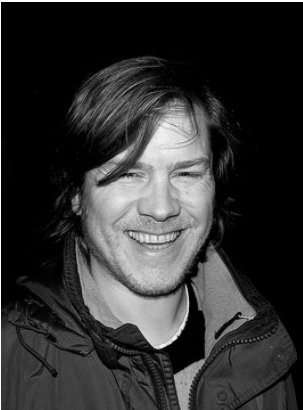
Keynotes

Wednesday, 25 March – 09:30 | Raghu Ramakrishnan

Data Management in the Cloud

We are in the midst of a computing revolution. As the cost of provisioning hardware and software stacks grows, and the cost of securing and administering these complex systems grows even faster, we're seeing a shift towards computing clouds. Clouds are essentially services accessed over a network, and offer developers scalable, robust computing infrastructure on a "pay as you go" basis, with the ability to dynamically adjust the amount of "rented" resources, and thereby, the bill. For cloud service providers, there is efficiency from amortizing costs and averaging usage peaks. Internet portals like Yahoo! have long offered application services, such as email for individuals and organizations. Companies are now offering services such as storage and compute cycles, enabling higher-level services to be built on top. In this talk, I will discuss Yahoo!'s vision of cloud computing, and describe some of the key initiatives, highlighting the technical challenges involved in designing hosted, multi-tenanted data management systems.

Thursday, 26 March – 17:30 | James Governor

Open Sourcing The Analyst Business – Turning Proprietary Knowledge Inside Out. Or... How And Why We Became Osi-Compliant

James will talk about changes in media-driven business models, the move to a service economy for analysis, licensing issues, internet coordination and collaboration, how RedMonk grew its business by learning from, consulting with, and championing open source technology and approach.

Friday, 27 March – 11:30 | Moderator: Danese Cooper

Apache Pioneer's Panel – 10 Years of The Apache Software Foundation

2009 marks the 10th Anniversary of The Apache Software Foundation. What started as a formalization of methods of collaboration for the world's most popular Web server has grown into a very popular watering hole on the savanna of Web-aware software development. Join several pioneering ASF members in this very special panel where we'll hear stories from the early days of the Apache Way.

thawte™
it's a trust thing™

Apache and *thawte* go way back…
to the beginning to be exact

Back in 1995, *thawte* modified the recently-released Apache to enable the global use of strong encryption, helping companies establish trusted relationships with anyone, anywhere on the Internet.

We're still committed to the same egalitarian ethos that's made Apache and the open source movement great. With global, multi-lingual customer service and outstanding value for the money, *thawte* is for those who know technology.

www.*thawte*.com

Sessions

Opening Plenary and State of the Feather
Lars Eilebrecht & Jim Jagielski
9:00 | Matterhorn 2

Welcome from the ApacheCon Conference Chair and ASF Chairman, including event highlights and a current look at what is going on at The Apache Software Foundation.

Opening Keynote - Data Management in the Cloud
Raghu Ramakrishnan
9:30 | Matterhorn 2

We are in the midst of a computing revolution. As the cost of provisioning hardware and software stacks grows, and the cost of securing and administering these complex systems grows even faster, we're seeing a shift towards computing clouds. Clouds are essentially services accessed over a network, and offer developers scalable, robust computing infrastructure on a "pay as you go" basis, with the ability to dynamically adjust the amount of "rented" resources, and thereby, the bill. For cloud service providers, there is efficiency from amortizing costs and averaging usage peaks. Internet portals like Yahoo! have long offered application services, such as email for individuals and organizations. Companies are now offering services such as storage and compute cycles, enabling higher-level services to be built on top. In this talk, I will discuss Yahoo!'s vision of cloud computing, and describe some of the key initiatives, highlighting the technical challenges involved in designing hosted, multi-tenanted data management systems.

Introducing Mahout: Apache Machine Learning
Grant Ingersoll
10:30 | Matterhorn 1
D **N**

With the massive rise in both structured and unstructured data comes the need to better understand what is in the data. Machine Learning algorithms are often designed to make sense of large volumes of data with minimal human input. In this talk, we cover some of the basic algorithms of machine learning (without all the gory math) and show how the new Apache Mahout project is attempting to solve these problems in a scalable way by utilizing Hadoop.

Introduction to Hadoop
Owen O'Malley
10:30 | Matterhorn 2
D **N**

Hadoop is an Apache project that provides a framework for running applications that processes large amounts of data (hundreds of terabytes) on large clusters (thousands of computers) of commodity hardware. The Hadoop framework transparently provides applications both reliability and data motion. Hadoop implements a distributed file system, similar to GFS, and MapReduce. This presentation presents the motivation and approach for Hadoop, an overview of the components and architecture, and an overview of the tools built on top of Hadoop, such as Pig and Hive.

Panel: The Business of Open Source - Power, Prestige, and Propulsion
Moderator: Sally Khudairi
Ajay Anand — Director of Product Management for Grid Computing, Yahoo!
Mark Brewer — Senior Vice President and GM of Enterprise Delivery, SpringSource
Arjé Cahn — Chief Technology Officer, Hippo
Steve Loughran — Senior Researcher, HP Laboratories, Bristol UK
10:30 | Matterhorn 3
B **O**

The ApacheCon Business Track addresses key business, marketing and legal/licensing issues in Open Source. A panel of influencers will address key topics in The Business of Open Source that include: Prestige Users - implementation opportunities, deployment challenges, and best practices; Propulsion Providers - business (re)generation, product development, and fulfillment; and Power Players - market shapers, money makers, and business model-shakers. Moderator Sally Khudairi will create a lively, interactive dialogue by including questions and comments from the audience throughout the 50-minute session.

Embrace OSGi - A Developer's Quickstart
Carsten Ziegeler
10:30 | St Gallen
D **N**

The first choice for highly modular, dynamic, and extensible applications is the OSGi technology. However the theory sounds very tempting but what about the real world? Starting with the basics of OSGi this session is focused on practical examples, tools, and procedures for a rapid adaption of OSGi in own projects. Learn how to avoid the typical traps and how to use the advantages of OSGi.

Lucene Case Studies
Erik Hatcher
11:30 | Matterhorn 1
D **N**

How are the Lucene technologies being used in the real world? This fast-paced talk presents many applications leveraging lucene.apache.org technologies, including Java Lucene, Solr, Nutch, and Mahout. We'll see how Lucene powers more library searches, sells more shoes, rents more videos, books more hotels, than high priced closed-source alternatives. We'll delve deeply into Lucid Imagination's search system for the Lucene community, how we continuously index web, wiki, mail, code, and database data sources and present them elegantly.

Sessions

Hadoop Map-Reduce: Tuning and Debugging
Arun Murthy
11:30 | Matterhorn 2   **D** **E**

As Apache Hadoop, and Hadoop Map-Reduce become widely adopted, especially in real-world applications which drive revenue, it becomes increasingly important to get the most out of Hadoop installations and the Map-Reduce applications. Also, distributed debugging and profiling Map-Reduce applications is hard, but critical for success. This talk will cover several ways to peer into Map-Reduce applications as they crunch terabytes of data. This wide-ranging discussion will also cover topics such as using debuggers/profilers on your applications, using Map-Reduce Counters, other simple ways to tune your applications, and how to avoid common pitfalls. We will also talk about the critical 'data-path' for application data as processed data flows from the map-step to the reduce-step and how to tune it to get optimal performance for user applications.

Adopting Open Source in the Enterprise
Adrian Trenaman
11:30 | Matterhorn 3   **B** **O**

For some organizations, the decision to adopt open source software is not taken lightly, and, once made, can require significant shifts in thinking and organization culture. This session discusses the issues organizations face when strategically adopting open source, including how to select open source projects, how to match open source involvement & contribution with corporate strategy, how to select & manage professional open source software vendors, what needs to be done in terms of licensing and legal, and what kind of infrastructure you need to support your open source strategy.

Tales from the OSGi trenches
Bertrand Delacretaz
11:30 | St Gallen   **D** **N**

In this talk we share our experience with the Apache Felix OSGi framework, used for a major rewrite of Day's family of content management products. After more than two years working with OSGi, the impact on our products, developers, customers and service people is very high, in a positive way. OSGi is no silver bullet either. The extreme modularization and dynamic service deployment features of OSGi make our products much more robust and maintainable, but the costs associated with changing people's way of thinking about code and modules, and with testing and debugging highly dynamic systems, must not be underestimated. Based on real-life code samples, we will show how OSGi is used at several levels in our products, from low-level interactions with the framework to very simple creation of (compiled or scripted) services. We will also present some of the automated testing techniques used in our project. Sharing our experience will help you decide if OSGi is for you, and more importantly at which level you should use it.

Chew & Chat | Behind the Scenes of The Apache Software Foundation
Lars Eilebrecht
12:30 | Matterhorn 2   **B** **O**

This presentation will give you everything you always wanted to know about the Apache Software Foundation (ASF), but were afraid to ask. It will show you that there is more than just the Apache web server, and provide you with information on how the ASF works. The difference between membership and committership, who decides what, how elections take place, the technical infrastructure, project management committees, and the philosophy behind the incubator. Come and see behind the scenes of the Apache Software Foundation and its many projects.

Advanced Indexing Techniques with Apache Lucene
Michael Busch
14:00 | Matterhorn 1   **D** **E**

In this presentation we will talk about advanced indexing and search techniques using the Payloads feature and interesting use cases such as efficient numeric search, faceted search, term occurrence boosting, part-of-speech tagging etc. The Lucene developers are currently also working on Flexible Indexing. This is one of the biggest and most powerful new features ever added to Lucene. This session will cover the current state and present new developments in the Flexible Indexing area, such as the new TokenStream API, a configurable indexing chain and pluggable indexing consumers. These are very exciting new additions to Lucene and everyone who wants to stay up-to-date with Lucene should attend this session.

Pig - Making Hadoop Easy
Olga Natkovich
14:00 | Matterhorn 2   **D** **N**

Hadoop provides a powerful platform that enables scalable, fault-tolerant, data-centric computing. However, as a user-facing programming paradigm it is too low-level and requires users to write significant amount of custom code, re-implement common processing primitives (e.g. join), and worry about chaining jobs together. Pig is a Hadoop sub-project that provides a higher level programming language to describe parallel computation. It takes care of implementing common relational operators such as join and filter, operator pipelining, and job chaining while providing ways to incorporate custom user code via user defined functions and streaming. The result is much simpler and more compact code, increased user productivity, and reduced maintenance time. At

Orientation: **B** Business **T** Technical **D** Developer   **Difficulty:** **O** Overview **N** Novice **E** Experienced

Sessions

the same time, unlike SQL databases which rely on a query optimizer to determine the execution strategy for a user program, Pig stays faithful to the spirit of map-reduce whereby a user program specifies a simple sequence of steps for the system to obey. The talk will introduce Pig and its programming model, contrast it with Hadoop¹s model and provide motivation to use Pig as the preferred programming paradigm for most applications. The performance tradeoffs will also be discussed.

Open Source for Interoperability
Paul Fremantle
14:00 | Matterhorn 3   **B** **O**

Interoperability is the crux of many new technologies: Open Source has been joined by Open Standards, and the combination is creating a whole new wave of distributed and cloud computing possible. But interoperability is always hard, and takes longer to achieve than we expect. In this talk, we look at how Open Source can be used to crack this nut: how to build interoperability using the techniques of Open Source and Apache: meritocracy, distributed development, continuous testing and shared ownership. Open Standards and Open Source are actually completely different - Open Source projects allow coders to collaborate. Open Standards organizations are a no man's land where competitors can agree the minimum required to compete on a level playing field. In this talk, Paul Fremantle will address how to use the collaborative approach of Open Source to enhance Open Standards. Using several examples from several standards initiatives, Paul will how this approach has succeeded, and how others can replicate this success.

OSGi as Framework for Building a Product Line: Experience and Best Practices
Ruwan Linton
Afkham Azeez
14:00 | St Gallen   **D** **N**

In this talk we present the experience of using OSGi as the framework for building a componentized line of SOA middleware products. Some of the challenges include the proper sizing & architecture of OSGi bundles, managing dynamic bundle dependencies, dealing with OSGi implementation incompatibilities and scalability. However the net result has been tremendously successful in creating a platform that allows us to extend / replace functionality almost trivially. "

HBasics: An introduction to Hadoop's Big Data database
Michael Stack
15:00 | Matterhorn 1   **D** **N**

HBase is the Hadoop database. Its an open-source, distributed, column-oriented store modeled after Google's Bigtable. Just as Bigtable leverages the distributed data storage provided by the Google File System, HBase provides Bigtable-like capabilities on top of Hadoop. The HBase project's goal is the hosting of very large tables -- billions of rows X millions of columns -- on top of clusters of commodity hardware. This talk will introduce HBase and present on the current state of development. HBase user deployments -- schemas, cluster hardware and size -- will also be discussed.

Running Hadoop in the Cloud
Tom White
15:00 | Matterhorn 2   **D** **E**

The combination of cloud infrastructure and Hadoop makes a very compelling story for on-demand data analysis. This talk gives a practical run-through of how to run Hadoop in the cloud (on various providers), and examines the trade offs between dynamic clusters, data locality and Hadoop performance.

Sustainable software: the community way
Gianugo Rabellino
15:00 | Matterhorn 3   **B** **O**

This session aims to look at software from an economics angle, showing how Open Development and community-based software can be powerful drivers towards building sustainable software and making money while at it. Now that Open Source is fully gone mainstream, there a still quite a few open questions on the table, such as how can Open Source beat real world economics rules (i.e., "who's footing the bill"), and whether the Open Source community will be able to build durable ecosystems that change the industry for good, proving it is not just a passing fad. The community model is paramount to building sustainable software, as this session will demonstrate by looking at hard industry numbers. Case studies from leading communities (Apache, Eclipse, Mozilla, Debian, and more) will be paired with the evolution of old-school business models from leading IT players, providing a reality check on the nascent Open Source industry, and how Open Development can be your business' best friend.

Administering Apache Geronimo 2.x
David Jencks
15:00 | St Gallens   **T** **N**

Apache Geronimo 2.x is certified Java Enterprise Edition 5.0 container suitable for everything from a development environment to enterprise-level deployments. Geronimo leverages many Apache projects, such as Tomcat, OpenEJB, ActiveMQ, Derby. This session will cover administering Geronimo from two points of view: configuring an existing Geronimo server directly through the powerful web based administration console or editing configuration files, and secondly through Geronimo plugins

Sessions

containing configuration information either used in a custom server assembly or installed into an existing server.

Some of the topics covered are:

• First steps in getting Geronimo ready for a production environment,
• Administering embedded Tomcat, ActiveMQ, Derby.
• Deploy/start/stop of applications & other server components.
• Creating Database pools, Security realms, etc.
• Installing plugins for Geronimo.
• Managing keys, digital certificates and configuring SSL.
• Running a cluster of multiple server instances from the same installation.

Content storage with Apache Jackrabbit
Jukka Zitting
16:30 | Matterhorn 1   D N

Hierarchical database or a transactional file system? Apache Jackrabbit combines some of the best features of relational databases and traditional file systems to implement a flexible high level storage solution for a wide range of applications. Jackrabbit, a fully conforming implementation of the Content Repository for Java Technology API (JCR), comes packed with features like full text search, versioning, and transactions. Built-in HTTP mappings with WebDAV extensions make content stored in Jackrabbit easily accessible in network environments. This presentation introduces you to the key concepts of JCR and shows how to use Apache Jackrabbit and related projects to build various types of content applications like wiki and blog engines, email archives or image galleries. Special emphasis is placed on a data-first approach to content design that helps make your applications extensible with little or no extra code.

Configuring Hadoop for Grid Services
Allen Wittenauer
16:30 | Matterhorn 2   T N

This session will share the experiences of the Yahoo! Grid Computing team in deploying and maintaining several large, shared Hadoop clusters used by researchers and engineers. Information will include how the team deploys and maintains the Hadoop software, configuring one of the new scheduling systems, uses of the new security features, and managing the single points of failure.

Apache license as a business model: Challenges and opportunities
Paul Fremantle
16:30 | Matterhorn 3   B O

Conventional wisdom in the open source business community is that the formula for successful open source businesses is GPL + dual license. The Apache

license in particular is viewed as "too challenging" to monetize due to its "overly friendly" nature. In this talk we examine the challenges and opportunities of running a business based on the Apache license.

Administering Apache Geronimo through Custom Server Assemblies
David Jencks
16:30 | St Gallens   T E

Simple administration tasks are often easy using the Geronimo admin console or direct editing of configuration files. However this is not particularly amenable to reproducibility or auditing. An alternate approach is to assemble customized servers that are preconfigured exactly as needed including the applications and services required. I'll show how to do this in a controllable workflow amenable to source control and how this can be used for easy cluster administration.

Rapid JCR applications development with Sling
Bertrand Delacretaz
17:30 | Matterhorn 1   D E

Apache Sling is a scriptable applications layer, built on the Apache Felix OSGi framework, that provides a RESTful interface to a JCR content repository. In this talk, we'll see how Sling enables rapid development of JCR-based content applications, by leveraging the JSR 223 scripting framework along with the rich set of OSGi components provided by Sling. We will create a simple application from scratch in a few minutes, and explain a more complex multimedia application that does a lot with few lines of code. This talk will help you get started with Sling and understand how the different components fit together.

Dynamic Hadoop Clusters
Steve Loughran
17:30 | Matterhorn 2   D E

This presentation looks at the challenges of bringing up Hadoop clusters on dynamically allocated server —real or virtual. It shows how to bring up the cluster, verify its health, get data into the file system, submit work, and retrieve the results. Once you can bring up a cluster dynamically, you can start to use Hadoop in interesting ways —inside a unit test, as part of a workflow, or on spare machines at night. It also lets you deploy to dynamically allocated servers, real or virtual, and keep an eye on the live machines. This talk will use Hadoop, SmartFrog, VMWare and perhaps Amazon EC2 by way of the Typica and Restlet libraries.

Sessions

Distributed Team building
Andreas Gies
17:30 | Matterhorn 3   `T` `N`

In this session we will have a closer look at developing a distributed application based on an ESB. As the components and services are developed quite indepently from each other, the specification for each component and it's documentation must be fit for retrieval and reuse. New challenges for development teams include knowledge sharing and knowledge reuse as well as shorter and more agile development iterations. Teams also have to overcome their fears when it comes down to changing their way of working. We will have a look at some strategies to address those challenges and overcome the fears in order to compose a team working successfully in an agile and distributed environment.

Apache Felix on Androids
Marcel Offermans
17:30 | St Gallens   `D` `N`

Apache Felix is an implementation of the OSGi specification, a dynamic module system for Java that can be used on systems ranging from embedded to enterprise. Google Android is a platform for mobile devices that uses an application component model based on a Java dialect. Combining both models allows the creation of modular, dynamically updateable desktop applications that can adapt to their environment. In this talk we will show that Felix already runs out of the box on Android and how it can be used to create applications out of components, dealing with issues like location awareness, languages and general updates. Based on a plugin store we will demonstrate dynamic installation and update of components.
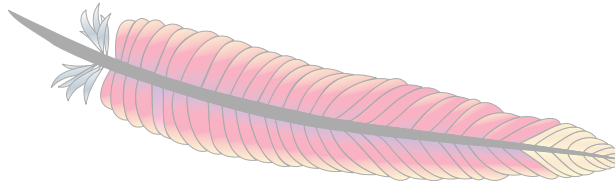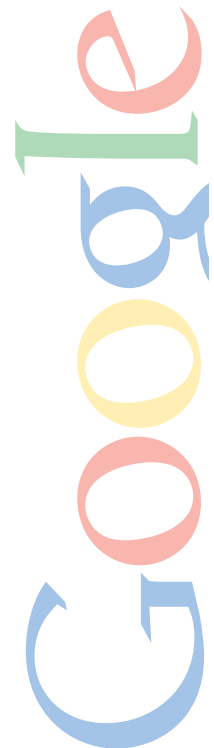
Sessions

Securing Apache Tomcat for your Environment
Mark Thomas
9:00 | Matterhorn 1 **T** **E**

   A default Apache Tomcat installation is secure but each installation environment is different and may have additional security requirements. This presentation will examine the security configuration options available in Apache Tomcat, when to use them (and when not to use them) and the threats they might help mitigate. The rationale behind having resource passwords (eg for database access) in clear text in server.xml will also be discussed.

Achieving Scalability and High Availability for clustered Web Services using Apache Synapse
Ruwan Linton
9:00 | Matterhorn 2 **D** **N**

   This will give an insight on how you can use Apache Synapse mediation framework to guarantee the scalability and availability of hosted services. It requires a cluster of services to be available and Synapse will balance the load and fail over route the messages among the cluster of services to provide scalability and availability. Apache Synapse is designed to be a simple, lightweight and high performance Enterprise Service Bus (ESB). Since it is based on a small asynchronous core, it can handle 2500 concurrent connections on a standard server hardware. The Synapse engine is configured with a simple XML format and comes with a set of ready-to-use transports on which it can be used as a load balancer and a fail over router. Synapse supports static and dynamic load balancing, static load balancing can be used when you have a set of known hosts in the cluster where as dynamic load balancing could be used for an environment where the nodes on the service cluster is dynamic for example when auto scaling of the service cluster is active. In the case of dynamic load balancing the load balance endpoint gets the list of nodes which provides this service from the underlying Group Communication Framework of the cluster. Apache Synapse can be used for general mediation tasks while providing the scalability and availability to a set of services as well. You may use Synapse to log the messages coming into your network or to throttle the messages to your services.

Open Source Collaboration Tools are Good For You!
Bertrand Delacretaz
9:00 | Matterhorn 3 **T** **O**

   What are the core requirements for a set of team collaboration tools? Looking at how ASF project communities collaborate online, we have identified four core drivers that help these projects succeed. We will show how the collaboration tools used by the ASF can allow any project team to move from an "ask around the office" collaboration model to our efficient "distributed self service information" model, while

focusing on those core drivers to avoid being distracted by the tools themselves. Our analysis will help you estimate the effort and expected benefits of such a move.

What's new in Servlet 3.0
Filip Hanik
10:00 | Matterhorn 1 **D** **O**

   Servlet 3.0 is a long awaited specification coming out of the JCP. The Servlet specification has not been updated in a long time, and the Servlet containers have become more diverse in order to keep up with market demand for new features. The Servlet 3.0 specification is trying to address some of these demands as well coming up with some new features. This session will give you an overview of what the specification has to offer and what the plans are for Tomcat 7.

Event Driven Architectures with Apache Synapse
Paul Fremantle
10:00 | Matterhorn 2 **D** **E**

   Event Driven Architecture (EDA) is a highly scalable approach for building large scale distributed applications and integrating across large organizations. In an EDA, many systems can manage their own destiny in connecting to the rest of the applications, and this approach means that new applications can be much more self-reliant. EDA has been widely adopted in the financial sector, but is gaining ground in many other areas. This presentation will introduce EDA and the benefits, challenges and best practices for building EDA, and also show how to architect real-life EDA systems. In addition, we will explore Complex Event Processing, which adds a whole new set of capabilities to find patterns, react in real time to data, and correlate across high volume data feeds. In addition, we will look in detail at how the Apache Synapse ESB can be used to build an EDA, both by itself and in conjunction with other projects such as Apache Qpid and ActiveMQ. We will examine how to build topic spaces, create routing and subscription patterns, and how to add Complex Event Processing to an EDA. This talk is aimed at coders and system architects who want to understand how event architectures can improve their systems and how to use Apache technology to build them.

Putting it All Together
Justin Erenkrantz
10:00 | Matterhorn 3 **B** **O**

   In most open-source projects, often left unsaid is how to effectively contribute within the accepted "societal norms" of a project. Contributions can take a wide range of forms: adding new features, bug fixes, documentation, and often just general feedback. In this session, we will review how an individual

Sessions

developer can contribute constructively to an open source project - with a specific eye on helping you contribute to Apache projects.

Architectures for the cloud
Steve Loughran
10:00 | St. Gallens   **T** **O**

This presentation looks at the hardware and software issues for building the infrastructure of "the cloud": network infrastructure evolution, hardware trends in the servers, virtualization and private VLANs, storage, and the implications for applications. It argues that we have to start treating a large datacentre installation as a single machine, and an application brought up in such a datacentre as a single application. It will show how Hadoop and layers on top of Hadoop form the ideal platform for much of the system, how other Apache products fit in to other parts of the datacentre, and where there are gaps in the portfolio. Everything from Log4J to Tomcat needs to adapt to this world —this talk looks at the issues and the opportunities.

Performance Tuning Apache Tomcat for Production
Filip Hanik
11:30 | Matterhorn 1   **T** **E**

Apache Tomcat committer Filip Hanik will in this interactive session discuss performance tuning Apache Tomcat for your production environment. We will focus on Tuning Tomcat and the JVM to correctly handle your application, usage patterns, hardware and network topology in your production environemnt. You'll learn when and how to apply the different tuning and configuration options as well as understanding load balancers and how they can impact your configuration settings. Also discussed: the impact of clustering and replication on your environment. Session will be interactive with the audience, bring your configuration questions, your current performance problems, your network topology, your JVM settings, and we will configure YOUR Tomcat to perform in an optimal way.

Tales from the Front Line: how ServiceMix, ActiveMQ & CXF are being used to solve real problems
Adrian Trenaman
11:30 | Matterhorn 2   **D** **N**

In this session we discuss how application integration and middleware-focussed Apache projects - like ServiceMix, ActiveMQ, Camel & CXF - are being used to build innovative solutions that add real value to organizations in health-care, government, airline-solutions and finance. For each use-case, we discuss both logical and deployment architecture concerns. We also discuss the motives for adopting open-source in these solutions, and why these Apache projects where chosen above other open source alternatives.

Documentation: get it right!
Niels van Kampenhout
11:30 | Matterhorn 3   **T** **O**

Good documentation is vital for the health of any open source project, yet the challenges that this fact presents are not always overcome. Let's face it: open source software has a reputation for poorly written and organized documentation. What causes this reputation? Is it really that bad? Is it really that difficult? How can we do better? And why do we need documentation anyway, can't you just look at the code?! This talk explores an often overlooked aspect of open source software development. It explains what makes documentation good documentation, and how it can benefit a project. It shows that, with a bit of common sense and clever use of resources, it's actually not that hard to produce useful documentation. A rough guide to getting your documentation right, while strengthening your community at the same time.

Django - giving your webapp a pony!
Nick Burch
11:30 | St. Gallens   **D** **N**

Django is an open source web application framework written in Python, which has recently released version 1.0. It describes itself as "A Web framework for perfectionists with deadlines", which its fast growing userbase agree with. Unlike many other frameworks, you can get started in 5 minutes, but 3 months later you're not fighting with it every step of the way! Django powers a growing list of sites, is behind Google App Engine, and when coupled with Jython you can turn you whole webapp into a .war for enterprise deployment. The talk will cover what Django offers, how to get started, how to keep going, how its philosophy keeps you coding and happy, and some great things you can achieve with it. We'll also do a whistle stop tour of some interesting things such as migrations, the different deployment options, json support, database abstraction and snippets. Oh, and Django's mascot is a pony!

Chew & Chat | Apache 101: Sponsoring the ASF at the Corporate and Individual Level
Jim Jagielski
13:00 | St. Gallens   **B** **O**

As an independent, non-profit organization, The Apache Software Foundation (ASF) does not receive government funding. We rely each year on the generous support from corporations, foundations, and private individuals to help offset day-to-day operating expenses such as bandwidth and connectivity, servers and hardware, legal and accounting services, marketing and public relations, general office expenditures, and support staff. Through the ASF Sponsorship Program, we can bolster continued growth across the ASF's 59 Top-Level Projects, 24 projects in the Apache Incubator, 20 initiatives in

Orientation: **B** Business   **T** Technical   **D** Developer   **Difficulty:**   **O** Overview   **N** Novice   **E** Experienced

Sessions

Apache Labs, community relations activities, and more. This session will provide participants with an overview of the ASF Sponsorship Program, the best methods for an organization or individual to support The ASF, and how you can get involved by "giving back" to one of the most compelling communities in Open Source.

mod_proxy versus mod_jk Clustering with HTTP Server as front-end
Jean-Frederic Clere
14:00 | Matterhorn 1   **T** **E**

There is often the question among users what should I use: mod_proxy or mod_jk? Mod_jk looks a kind of toolbox for integrating Tomcat application in web server and mod_proxy looks more than a straight forward loadbalancer. This presentation will try to help to choose the best module depending on the application structure. The goal of the two reverse proxy modules are different therefore they will continue existing and exchanging features. Now everyone will be able to choose the technology that fits their needs.

Tooling for Service Mix 4
Andreas Gies
14:00 | Matterhorn 2   **D** **N**

This presentation will walk you through the development and deployment of a sample application based on Apache Service Mix 4. After a short overview of the Service Mix 4 architecture and principles we will use the Eclipse based FUSE Integration Designer to design and implement a sample workflow. We will then package this workflow up into OSGi bundles and demonstrate how to deploy those via a central Maven repository into a given Service Mix instance instance.

How to become a project at the ASF
Martijn Dashorst
14:00 | Matterhorn 3   **B** **O**

Many projects aspire to join the Apache Software Foundation and are anxious to take the plunge. The ASF has instituted a special project for guiding and mentoring aspiring projects to become a successful project under the Apache umbrella: the incubator. Any new project needs to go through the incubator before it can be established as an Apache project. Many projects have passed through the incubator and some got stranded there. Learn what makes a project successful and what makes a project fail. You will learn from real projects that successfully graduated or failed for various reasons. Learn how to enter the incubator, foster your project under guidance of the incubator's mentors, how graduation works and what comes next.

Apache Droids – an intelligent standalone robot framework
Thorsten Scherler
14:00 | St. Gallens   **D** **O**

In the session we give an introduction into Apache Droids, Droids aims to be an intelligent standalone robot framework that allows to create and extend existing droids (robots). In the future it will offer an administration application to manage and control the different droids. Apache Droids is not designed for a special usecase, it is a framework: Take what you need, do what you want. Droids makes it very easy to extend existing robots or write a new one from scratch, which can automatically seek out relevant online information based on the user's specifications and invoke arbitrary business logic. If a project needs special processing for a crawled url the developer can easily write different kind of plugins to implement the functionality. The project started as an Apache Labs project and entered the Apache Incubator with Apache HttpComponents and Apache Lucene as mentoring project.

Introduction to NIO.2 (Asynchronous I/O) and how you can benefit from being asynchronous!
Jeanfrancois Arcand
15:00 | Matterhorn 1   **D** **N**

This session will introduce NIO.2 API and Concepts and demonstrate how Apache Projects like Tomcat can take advantage of the new API, both in term of code clarity and performance. First, the new Asynchronous I/O API will be discussed in details. Second, a new Tomcat Connector will be written to demonstrate some of the new API like AsynchronousServerSocket, NetworkChannel, etc. Next, the fundamental difference between NIO.1 and NIO.2 will be discussed by demonstrating how much simple is to write application using this NIO.2. Finally, we will discuss the future of NIO.1 frameworks like MINA and Grizzly and how they can take advantages of the NIO.2 concepts.

Service Mix 4 topologies
Andreas Gies
15:00 | Matterhorn 2   **D** **N**

In this talk we will look at the most common deployment topologies for ESB applications based on Apache Service Mix. To illustrate the topologies we will deploy a sample application into various constellations and discuss the pro's and con's of each approach. Especially we will look at fail over and load balancing capabilities, throughput, deployment complexity and discuss the packaging of an application feasible for distributed deployment.

Schedule at-a-glance

Data Mining and Search Technologies Matterhorn 1	Geeks for Geeks (Hadoop) Sponsored by cloudera Matterhorn 2	Business and Community Matterhorn 3	OSGi and Geronimo St Gallen
09:00-09:30	**Opening Plenary & State of the Feather in Matterhorn 2**		
	Lars Eilebrecht & Jim Jagielski		
9:30-10:20	**Opening Keynote - Data Management in the Cloud in Matterhorn 2**		
	Raghu Ramakrishnan		

| 10:30-11:20 | **Introducing Mahout: Apache Machine Learning** Grant Ingersoll **D N** | **Introduction to Hadoop** Owen O'Malley **D N** | **Panel: The Business of Open Source - Power, Prestige, and Propulsion** Sally Khudairi **B O** | **Embrace OSGi - A Developer's Quickstart** Carsten Ziegeler **D N** |
| 11:30-12:20 | **Lucene Case Studies** Erik Hatcher **D N** | **Hadoop Map-Reduce: Tuning and Debugging** Arun Murthy **D E** | **Adopting Open Source in the Enterprise** Adrian Trenaman **B O** | **Tales from the OSGi trenches** Bertrand Delacretaz **D N** |

12:30 -14:00	**Lunch Break Sponsored by** YAHOO! DEVELOPER NETWORK
12:30-14:00	**Chew & Chat	Behind the Scenes of The Apache Software Foundation** Lars Eilebrecht **B O**

| 14:00-14:50 | **Advanced Indexing Techniques with Apache Lucene** Michael Busch **D E** | **Pig - Making Hadoop Easy** Olga Natkovich **D N** | **Open Source for Interoperability** Paul Fremantle **B O** | **OSGi as Framework for Building a Product Line: Experience and Best Practices** Ruwan Linton **D N** |
| 15:00-15:50 | **HBasics: An introduction to Hadoop's Big Data database** Michael Stack **D N** | **Running Hadoop in the Cloud** Tom White **D E** | **Sustainable software: the community way** Gianugo Rabellino **B O** | **Administering Apache Geronimo 2.x** David Jencks **T N** |

16:00-16:30	**Coffee Break Sponsored by** YAHOO! DEVELOPER NETWORK

| 16:30-17:20 | **Content storage with Apache Jackrabbit** Jukka Zitting **D N** | **Configuring Hadoop for Grid Services** Allen Wittenauer **T N** | **Apache license as a business model: Challenges and op-portunities** Paul Fremantle **B O** | **Administering Apache Geronimo through Custom Server Assemblies** David Jencks **T E** |
| 17:30-18:20 | **Rapid JCR applications develop-ment with Sling** Bertrand Delacretaz **D E** | **Dynamic Hadoop Clusters** Steve Loughran **D E** | **Distributed Team building** Andreas Gies **T N** | **Apache Felix on Androids** Marcel Offermans **D N** |

18:30-20:00	**Welcome Reception and ASF Anniversary Kickoff - Atrium Area - Sponsored by** hp invent HIPPO YAHOO! DEVELOPER NETWORK

Orientation: **B** Business **T** Technical **D** Developer   **Difficulty:** **O** Overview **N** Novice **E** Experienced

Special Events

Hackathon | 23-24 March

Sponsored by [Joost logo] **and** [thawte logo]

Previously closed to the general public, this year the Apache community welcomes everyone to join in the two-day codefest called Hackathon! If you are interested and have the desire and skills to code and develop Open Source software, please consider joining us. The Hackathon will start at about 10:00 on Monday.

Apache MeetUps | Evenings of 23-24 March

Sponsored by [HIPPO logo] [lucid IMAGINATION logo] [topicus logo]

MeetUps are informal sessions on Apache projects and are organized by the members of the project community itself. With an open agenda, for which anyone can propose a talk, covering a wide range of subjects, MeetUps bring together a diverse crowd of committers, users and people interested in a specific project. Check the Wiki for the current schedule.

BarCampApache | 23 March

Sponsored by [Progress FUSE fusesource.com logo] [Google logo] [Joost logo]

BarCampApache is an Apache-focused unconference. It will gather people interested in The ASF and it's various projects to share and learn in a completely open environment. BarCampApache will be intense with discussions, demos and interaction from participants packed into an ad-hoc, participant-driven event.

Welcome Reception and ASF Anniversary Kickoff | 25 March

Sponsored by [hp invent logo] [HIPPO logo] [YAHOO! DEVELOPER NETWORK logo]

Join us Wednesday evening to relax and meet and geek with the people you have seen all day but have been too busy with the conference to really sit down and talk to. Have a bit of food and a glass of beer or wine and enjoy yourself!

The Big Data Hadoop Track | 25 March

Sponsored by [cloudera logo]

A full day of Hadoop sessions is being presented in the "Geeks for Geeks Track" on Wednesday.

Lightning Talks | 26 March

This lively session has become a fun and expected ApacheCon tradition! Thursday evening, after the day's sessions are done, we will gather in the main hall and the beer will flow. As a community, taking time to laugh together is as important as taking time to code together. The Lightning Talks are spontaneous and each year has its own unique highlights. Topics can be of any sort, as long as they relate to the ASF or to ASF projects.

ApacheCon Raffle | 27 March

Sponsored by [logo]

Join us at the Closing Plenary on Friday, where randomly-selected individuals receive fabulous giveaways such as cool gadgets, T-shirts, timepieces, electronics, and more! The catch? You have to be there to win!

Schedule at-a-glance

	Tomcat for Developers and Administrators Matterhorn 1	Service-Oriented Architecture Matterhorn 2	Business and Community Matterhorn 3	Geeks for Geeks St Gallen
09:00-09:50	**Securing Apache Tomcat for your Environment** Mark Thomas [T] [E]	**Achieving Scalability and High Availability for clustered Web Services using Apache Synapse** Ruwan Linton [D] [N]	**Open Source Collaboration Tools are Good For You!** Bertrand Delacretaz [T] [O]	**TBD**
10:00-10:50	**What's new in Servlet 3.0** Filip Hanik [D] [O]	**Event Driven Architectures with Apache Synapse** Paul Fremantle [D] [E]	**Putting it All Together** Justin Erenkrantz [B] [O]	**Architectures for the cloud** Steve Loughran [T] [O]
11:00-11:30	**Coffee Break**			
11:30-12:20	**Performance Tuning Apache Tomcat for Production** Filip Hanik [T] [E]	**Tales from the Front Line: how ServiceMix, ActiveMQ & CXF are being used to solve real problems.** Adrian Trenaman [D] [N]	**Documentation: get it right!** Niels van Kampenhout [T] [O]	**Django - giving your webapp a pony!** Nick Burch [D] [N]
12:30-14:00	**Lunch Break**			
13:00-14:00	**Chew & Chat | Sponsoring the ASF at the Corporate and Individual Level** Jim Jagielski [B] [O]			
14:00-14:50	**mod_proxy versus mod_jk. Clustering with HTTP Server as front-end.** Jean-Frederic Clere [T] [E]	**Tooling for Service Mix 4** Andreas Gies [D] [N]	**How to become a project at the ASF** Martijn Dashorst [B] [O]	**Apache Droids - an intelligent standalone robot framework** Thorsten Scherler [D] [O]
15:00-15:50	**Introduction to NIO.2 (Asynchronous I/O) and how you can benefit from being asynchronous!** Jeanfrancois Arcand [D] [N]	**Service Mix 4 topologies** Andreas Gies [D] [N]	**Eclipse IAM, Maven integration for Eclipse** Carlos Sanchez [D] [N]	**Best of breed - httpd, forrest, solr and droids** Thorsten Scherler [D] [E]
16:00-16:30	**Coffee Break**			
16:30-17:20	**What the Bayeux? Understanding, Using and Developing with the Bayeux Protocol** Filip Hanik [D] [N]	**Apache Solr - A Case Study** Uri Boness [D] [N]	**Munging URLs with mod_rewrite** Rich Bowen [T] [N]	**Using MINA 2.0 in Real Life** Emmanuel Lecharny [D] [E]
17:30-18:20	**Keynote — Open Sourcing The Analyst Business – Turning Proprietary Knowledge Inside Out. Or... How And Why We Became Osi-Compliant In Matterhorn 2** James Governor			
18:30-19:30	**Lightning Talks in Matterhorn 2**			

The Apache Software Foundation

would like to thank...

The Apache Software Foundation could not exist without the continued generous support from the community. We would like to take this opportunity to acknowledge and thank our Sponsors.

Platinum Sponsors

Google™

Microsoft®

YAHOO!®

Gold Sponsor

hp® invent

Silver Sponsors

spring source

Progress FUSE™
fusesource.com

Bronze Sponsors

Joost
AirPlus International
Two Sigma Investments

Tetsuya Kitahata
Matt Mullenweg

Infrastructure Support

Oregon State University Open Source Lab
SURFnet

Schedule at-a-glance

	HTTP Server Administration Matterhorn 1	Builds and Clouds Matterhorn 2	Java Development Matterhorn 3	Geeks for Geeks St Gallen
09:00-09:50	TBD [T] [N]	AntUnit Your Functional Test Jimmy Zhao [T] [O]	Wicket in Action Martijn Dashorst [D] [N]	Becoming a Tomcat super user Mark Thomas [T] [E]
10:00-10:50	Advanced Topics in Apache HTTP Server 2.2 Jim Jagielski [T] [O]	Managing Your Repositories With Archiva Maria Odea Ching [T] [O]	Scripting your Java Application with BSF 3.0 Felix Meschberger [D] [N]	High Availability != High cost Norman Maurer [T] [E]
11:00-11:30	Coffee Break			
11:30 - 13:00	Apache Pioneer's Panel - 10 years of The Apache Software Foundation in Matterhorn 2 Moderated by Danese Cooper (90 minute panel)			
13:00-14:30	Lunch Break			
13:30-14:20	Chew & Chat	The Apache Way J Aaron Farr			[B] [O]
14:30-15:20	New in HTTP Server 2.4: Session support with mod_session Graham Leggett [T] [E]	Apache Continuum - Build, Test, and Release Wendy Smoak [D] [N]	What's new with Apache POI Nick Burch [D] [O]	Improving mod_proxy : Example of an "home made" balancer: Mod_cluster Jean-Frederic Clere [T] [E]
15:30-16:20	On-the-fly Subversion replication Norman Maurer [T] [E]	Enterprise build and test in the cloud Carlos Sanchez [T] [N]	Apache DS 2.0 : What's new ? Emmanuel Lecharny [T] [O]	Pragmatic Models: building systems with domain specific languages Paul Fremantle [D] [E]
16:30-17:20	Deciphering mod_ssl: Using SSL with the Apache HTTP Server Joe Orton [T] [N]	Autoscaling Axis2 Web services on Amazon EC2 Afkham Azeez [D] [E]	Shindig for Blogs and Wikis Dave Johnson [D] [N]	Deploying to the Edge with CouchDB J. Chris Anderson [D] [E]
17:30	Coffee, Tea, Closing Remarks and a RAFFLE!!!! Sponsored by			

SAVE THE DATE

ApacheCon US 2009
Oakland, CA
2-6 November 2009
http://www.us.apachecon.com

ApacheCon Europe 2010
Spring 2010

ApacheCon US 2010
Atlanta, GA
1 -5 November 2010

ApacheCon North America 2011
Vancouver BC, Canada
7 – 11 November 2011

REGIONAL GATHERINGS

BarCampApacheOxford 2009
Oxford, UK

4-5 April 2009 | University Club, Oxford UK

http://www.barcamp.org/
BarCampApacheOxford

BarCamp | MeetUp Asia
China | Sri Lanka
Late Summer 2009

Sessions

Eclipse IAM, Maven integration for Eclipse
Carlos Sanchez
15:00 | Matterhorn 3   D N

Eclipse IAM (Eclipse Integration for Apache Maven), formerly "Q for Eclipse", is an Open Source project that integrates Apache Maven and the Eclipse IDE for faster, more agile, and more productive development. The plugin allows you to run Maven from the IDE, import existing Maven projects without intermediate steps, create new projects using Maven archetypes, synchronize dependency management, search artifact repositories for dependencies that are automatically downloaded, view a graph of dependencies and more! Join us to discover how to take advantage of all these features, as well as how they can help you to improve your development process.

Best of Breed – HTTP Server, Forrest, Solr, and Droids
Thorsten Scherler
15:00 | St. Gallens   D E

This session shows how to use best of breed Apache products to develop high traffic pages where performance is the determine factor. Each day the Junta de Andalucia is publishing the official bulletin on the governmental web page http://www.juntadeandalucia.es/boja/. The daily bulletin is generated with Apache Forrest using Adobe Indesign files as base. In the process of the static export of the content of the bulletin the different laws are crawled with Apache Droids which then uses Apache Solr to index them. Finally an Apache httpd server is used to answer the high volume requests.

What the Bayeux? Understanding, Using and Developing with the Bayeux Protocol
Filip Hanik
16:30 | Matterhorn 1   D N

While AJAX and now the Comet technique are becoming more popular and mainstream, there is still a lot of uncertainty about what the Bayeux protocol offers, why you need it and causing confusion on how to use it, less implement it. In less than an hour we will demystify the protocol, lay out its structure and the reasoning behind it. You will learn the different frameworks, both client and server side, how to use them and how they differ. And finally we will also show you how to implement the protocol itself or simply build a dynamic web application using one or more of the existing Bayeux frameworks.

Apache Solr - A Case Study
Uri Boness
16:30 | Matterhorn 2   D N

This session will introduce the attendees to Solr by a real world example. We will show how Solr enabled us to replace an existing commercial search engine in one of the most popular online company directories in The Netherlands. We'll discuss the decision making process that led the company to explore open source alternatives to their search backend in general and why Solr was chosen in particular. We will also show how using Solr extensible infrastructure enabled us to implement non-trivial search functionality such as geo-location search and complex ranking rule schemes.

Munging URLs with mod_rewrite
Rich Bowen
16:30 | Matterhorn 3   T N

It's the bane of every website administrator, and most of us have done it wrong more often than we've done it right. mod_rewrite mangles and munges URLs in order to make them easy to type, easy to remember, and attractive to put on your brochures. In this session, Rich covers the basics of mod_rewrite syntax, a little bit of the regular expression lexicon you need to know, and steps through a few recipes for common scenarios. Rich is the author of 'The Definitive Guide to Apache mod_rewrite'

Using MINA 2.0 in Real Life
Emmanuel Lecharny
16:30 | St. Gallens   D E

MINA 2.0 is a network framework written in Java. Having unique features and extension points, writing applications based on MINA is way simpler than doing so using the Java API. Many projects already use MINA : - Apache Directory Server (LDAP, LDAPS, Kerberos, NTP, using both TCP and UDP) is based on MINA unique capabilities to handle thousands of simultaneous clients, supporting more than 10 thousands requests per seconds. The encoder/decoder is stateful. - Vysper is an Apache Lab, demonstrating that one can write a XMPP server based on MINA - FtpServer is a complete server, based on MINA - AsyncWeb is an Http server written on top of MINA This presentation will demonstrate how easy it is to develop an application based on MINA, and we will use a XML based protocol to show that MINA 2.0 helps the server developers to focus more on the server's functionalities than on the network layer.

Keynote — Open Sourcing the Analyst Business – Turning Proprietary Knowledge Inside Out. Or... How and Why We Became OSI-compliant
James Governor
17:30- 18:20

James will talk about changes in media-driven business models, the move to a service economy for analysis, licensing issues, internet coordination and collaboration, how RedMonk grew its business by learning from, consulting with, and championing open source technology and approach.

Sessions

To Be Announced
9:00 | Matterhorn 1

AntUnit Your Functional Test
Jimmy Zhao
9:00 | Matterhorn 2   **T** **O**

There are many powerful test tools now in the world. And many of them are also open source. But in my last two years experiences, my colleagues and I found a super simple way to test our work. And the tool is Ant based. Simple but powerful. And then we tried this way and it works! A few month later we found Apache released AntUnit. It's so closed to our solution, but I can say most of people haven't aware of the power of AntUnit. I would like to introduce this powerful tool to our open source fans.

Wicket in Action
Martijn Dashorst
9:00 | Matterhorn 3   **D** **N**

Apache Wicket is a component oriented Java web application framework that brings back object oriented programming to the Web. With just HTML and Java, Wicket gives designers and programmers a tool they can work together in. In this session you'll be introduced to Wicket's programming model, and we'll quickly go from simple components, form processing and providing feedback to integrating with Spring and Guice, securing, testing and deploying your application. A lot of ground to cover but the best way to learn everything you wanted to know about Wicket.

Becoming a Tomcat super user
Mark Thomas
9:00 | St. Gallens   **T** **E**

Have you ever wanted to to do more than just download Tomcat and install some web applications? Do you want to learn more about Tomcat internals and how everything fits together? Do you want to be able to apply patches for bugs to your Tomcat installation without having to wait for the next release? If you answered yes to any of the above questions then this presentation is for you. Starting with how the Tomcat subversion repository is structured, this presentation will show you how to build each of the major Tomcat versions from source, how to use your local build to generate binary patches for specific bugs and how to extend Tomcat functionality for your environment. After an overview of the Tomcat request processing architecture, a request will be examined as it passes through Tomcat using remote debugging. Key classes and decision points will be highlighted to provide attendees with potential starting points when debugging their own issues. To finish the presentation, and to give a practical example of how the knowledge gained may be put to use, a current bug will be examined, the root cause identified and a patch to fix it generated.

Advanced Topics in Apache HTTP Server 2.2
Jim Jagielski
10:00 | Matterhorn 1   **T** **O**

An overview of what is new and improved in the 2.2.x version of the Apache Web Server. With that as a baseline, we will discuss advanced topics, such as increasing performance and reliability, scalability options and monitoring performance and resource allocations.

Managing Your Repositories With Archiva
Maria Odea Ching
10:00 | Matterhorn 2   **T** **O**

Repositories have become a key component in software development nowadays. They are used for project source versioning, for project deployment and more basically for storing artifacts. Having this set up introduces problems like bulging repositories, corrupted files, download problems, access controls, and many more. So, introducing Apache Archiva.. Archiva is a build artifact repository manager which has the capabilities of hosting and serving an artifact repository. This session will focus on how to best use and utilize Archiva's key features to work with Maven. The session would be divided into four sections namely, introduction, strategies for managing your Maven repositories, administration and a short demonstration. The first section will be a brief overview of what Archiva is and what it does while the second section will cover repository configuration and set up, hooking up Maven with Archiva, and grouping your repositories. Security and maintenance will be tackled on the third section then followed by a short demonstration.

Scripting your Java Application with BSF 3.0
Felix Meschberger
10:00 | Matterhorn 3   **D** **N**

One very important functionality of modern extensible applications is support for developing such extensions in any scripting languages. Many scripting

Orientation: **B** Business   **T** Technical   **D** Developer   **Difficulty:**   **O** Overview   **N** Novice   **E** Experienced

Sessions

languages available today provide some sort of Java integration but each integration is different making it very difficult for the vendor of the application to support more than one scripting language. Enter the Java Script API as defined in JSR-223. This API provides support for standardized integration of scripting languages in Java applications. Bindings already exist for a number fo scripting languages such as Groovy, JavaScript, Python, Ruby, Tcl. This session will show how easy it is to add scripting support to a Java application using the Java Scripting API and thus support whatever scripting language the user of the application likes to use. Practical demonstrations using Apache BSF 3.0 as the Java Scripting API implementation and Apache Sling as a Java application to be scripted will show how easy it is to add scripting support and to add scripting languages quickly and at runtime without even restarting the application.

High Availability != High cost
Norman Maurer
10:00 | St. Gallens   **T** **E**

In today's world it's getting more important to serve users on a 24x7 period with as little downtime as possible. With Drbd and Heartbeat its possible to build a low-cost Master/Slave Cluster on Linux based Operation Systems without to much hustle. Get a brief overview how-to setup a Master/Slave-Cluster providing whatever service you need using "low-cost" hardware and Open-Source software.

"2009 marks the 10th Anniversary of the Apache Software Foundation. What started as a formalization of methods of collaboration for the world's most popular web server has grown into a very popular watering hole on the savanna of web-aware software development. Join several pioneering ASF members in this very special panel where we'll hear stories from the early days of the Apache Way"

Apache Pioneer's Panel - 10 years of The Apache Software Foundation
Moderator: Danese Cooper
11:30 | Matterhorn 2

2009 marks the 10th Anniversary of The Apache Software Foundation. What started as a formalization of methods of collaboration for the world's most popular Web server has grown into a very popular watering hole on the savanna of Web-aware software development. Join several pioneering ASF members in this very special panel where we'll hear stories from the early days of the Apache Way.

Chew & Chat | The Apache Way
J Aaron Farr
13:00 | Matterhorn 2   **B** **O**

Apache prides itself on its emphasis of community driven development. But in practice, how does this work? Aaron Farr shares lessons learned from the

inner working of the Apache Software Foundation -- from its principles and practices, to the people and businesses around it.

New in HTTP Server 2.4: Session support with mod_session
Graham Leggett
14:30 | Matterhorn 1   **T** **E**

Sessions are a standard feature of web application servers, however the interoperability of sessions between different web application servers is generally limited. This talk introduces the mod_session modules collection in httpd v2.4 that attempts to create a unified session for httpd and web application servers. It is aimed at people interested in practical single sign-on, as well as for people trying to bring sanity to mixed architecture environments. We will start by introducing some of the history behind mod_session, why it was developed, and what problems mod_session is trying to solve. We will then continue by introducing what a session is within mod_session, and cover some of the ways in which a session might be created and where a session might be stored. We will then move on to cover the options available for securing sessions using encryption, and handling sessions on highly loaded or distributed environments. The focus will then shift to web applications and web application servers, and how they might read from and write to a session without having to care how the session is implemented or configured. We will then show some examples of how sessions might be used in practice, with examples in secure and in highly loaded environments. The talk will conclude by looking at some of the future development planned for mod_session, and how people can get involved in further development before httpd v2.4 is finally released.

Apache Continuum - Build, Test, and Release
Wendy Smoak
14:30 | Matterhorn 2   **D** **N**

Apache Continuum is an enterprise-ready continuous integration server with features such as role based security, automated builds, release management, and integration with popular build tools and source control management systems. Whether you have a centralized build team or want to put control of releases in the hands of developers, Continuum can help you improve quality and maintain a consistent build environment. This presentation will explain the benefits of continuous integration and show you how to get the most out of Continuum. We'll start with installation and configuration, show how to add projects and restrict access, how to set up build schedules and notifications, and finally we'll explore Continuum's integration with both Apache Maven and Subversion with a demonstration of releasing a project.

Sessions

What's new with Apache POI
Nick Burch
14:30 | Matterhorn 3   D O

Apache POI is a long standing project to provide pure Java access to Microsoft Office file formats. However, it's not a static project, and over the last 18 months its features have once again grown dramatically. The session will start with a whistle-stop tour of the older features, then dive into what's new. One such topic will be reading and writing old style and new xml based (ooxml) Excel files with the same code. Another will be the advanced text extraction for an ever growing list of file formats (9 and counting). We'll see how to render powerpoint files to images, and build a powerpoint file from Java. We'll cover reading excel files in a memory efficient manner, and see how the new Office Open XML file formats have and haven't changed everything.

Improving mod_proxy : Example of an "home made" balancer: Mod_cluster
Jean-Frederic Clere
14:30 | St. Gallens   T E

Summary: When using Mod_proxy + Mod_proxy_balancer Apache httpd only 2 balancer modes and reduced management support is possible. Anyway it is possible to write custom balancers to support different type of clusters. A Complete example of a balancer to front-end a dynamic cluster of Tomcat's will presented.Mod_proxy allows to * Abstract: The goal of the presentation is to demonstrate how easy it is to improve mod_proxy interaction with a cluster of Tomcats. There are different session replications mode in clusters. Full replication, Buddy replication, passive nodes for High Availability. Load balancing and failure recovery should be based on a dynamic view of the cluster not on a static description in a configuration file. Recent improvements in Mod_proxy allows the standard Mod_proxy to front-end a cluster of servers but actual ASF mod_proxy_balancer offers only 3 balancer modes (byrequest, bytraffic and bybusyness) and reduced managing support (via mod_status). What is needed to have a completely dynamic cluster: On Tomcat side a Listener to collect and transmit node informations to the balancer module and on Httpd side a balancer module able to use those informations to route incoming requests according to different loadbalancing and failure recovery policies. How it is done: A quick overview of the Tomcat listener of the example and an overview of the balancer module. Explanation the features of mod_proxy that are available to write a balancer. The mod_cluster solution: Protocol, Logic and Tomcat integration.

On-the-fly Subversion replication
Norman Maurer
15:30 | Matterhorn 1   T E

Do it like the ASF Infrastructure team! Learn from our experience - how-to setup Subversion on-the-fly replication to spread the load over as many servers as you like. Use Open-Source software including Apache HTTPD 2.2.x (mod_proxy, mod_dav_svn) and Subversion 1.5.x. Learn about the pros and cons of this setup. Learn about common problems and how-to fix them.

Enterprise build and test in the cloud
Carlos Sanchez
15:30 | Matterhorn 2   T N

Building and testing software can be a time and resource consuming task. Cloud computing / on demand services like Amazon EC2 allow a cost-effective way to scale applications, and applied to building and testing software can reduce the time needed to find and correct problems, meaning a reduction also in time and costs. Properly configuring your build tools (Maven, Ant,...), continuous integration servers (Continuum, Cruise Control,...), and testing tools (TestNG, Selenium,...) can allow you to run all the build/testing process in a cloud environment, simulating high load environments, distributing long running tests to reduce their execution time, using different environments for client or server applications,... and in the case of on-demand services like Amazon EC2, pay only for the time you use it. In this presentation we will introduce a development process and architecture using popular open source tools for the build and test process such as Apache Maven or Ant for building, Apache Continuum as continuous integration server, TestNG and Selenium for testing, and how to configure them to achieve the best results and performance in several typical use cases (long running testing processes, different client platforms,...) by using he Amazon Elastic Computing Cloud EC2, and therefore reducing time and costs compared to other solutions.

Apache DS 2.0 : What's new ?
Emmanuel Lecharny
15:30 | Matterhorn 3   T O

Apache Directory Server 2.0 is currently being finalized, and will be ready by ApacheCon EU this spring. So what's new in this version ? This presentation will expose the new features and improvements we have done since the last versions, and we also will compare ADS to other LDAP servers. Here is a short list of those features : - Java based, and embedable server - Support LDAP and Kerberos - Multi-Master replication - Stored-Procedures and triggers - ChangeLog for rollbacking operations at will - Multi-backend support - Better performances

Orientation: B Business   T Technical   D Developer   Difficulty: O Overview   N Novice   E Experienced

Sessions

- OpenGroup certified as a LDAP compliant server
This version comes hand to hand with a new set of
graphical tools, Apache Directory Studio.

Pragmatic Models: building systems with domain specific languages

Paul Fremantle
15:30 | St. Gallens   **D** **E**

Model Driven Architecture (MDA) is a love it or
hate it concept. Some people love the approach, giving
them the right abstractions to design first, and others
think its a waste of time and just holds them back
from coding. In this presentation, Paul Fremantle will
look at building real systems efficiently using Domain
Specific Languages (DSLs). By building languages
that really capture the overall problem space, we can
avoid the abstractness of MDA, while gaining the real
effectiveness of coding. The presentation will cover a
real system - a complex integration of legacy products
including a master data repository, and show how
building models in a one or more domain specific
languages made the whole system faster to code, more
transparent, and easier to customize and maintain. It
will also show how DSLs can be used to glue together
existing languages including scripting languages,
SQL, BPEL, XSLT and others. Finally, the talk will
look at effective ways of creating DSLs, including
the use of new languages and scripting frameworks,
together with code-generation tools. The talk is aimed
at anyone looking to build complex systems faster and
more effectively.

Deciphering mod_ssl: Using SSL with the Apache HTTP Server

Joe Orton
16:30 | Matterhorn 1   **T** **N**

mod_ssl is one of the most complex modules
shipped with the Apache HTTP Server. This
presentation will start by explaining how to set up an
SSL server using httpd and mod_ssl, then move on to
cover use of advanced features from client certificate
authentication and fine-grained access control through
to session caching and performance.

Autoscaling Axis2 Web services on Amazon EC2

Afkham Azeez
16:30 | Matterhorn 2   **D** **E**

Fault tolerance, high availability & scalability are
essential prerequisites for any enterprise application
deployment. One of the major concerns of enterprise
application architects is avoiding single points of
failure. There is a high cost associated with achieving
high availability & scalability. We will look at an
economical approach towards automatically scaling
Apache Axis2 Web service applications while
maintaining the availability & scalability guarantees at
an optimum economical cost.

Shindig for Blogs and Wikis

Dave Johnson
16:30 | Matterhorn 3   **D** **N**

Blogs, wikis and feeds helped to make the web
more social by making it easy for folks to read,
write and have conversations on the web; and now
social networking technolgies are making the web
and even more social. In this session you'll learn
about OpenSocial, a new standard for interacting
with social networking data via Web Service and via
JavaScript Gadgets that can be embedded into social
networking sites. You'll learn about Apache Shindig
(incubating), which is the reference implementation
of OpenSocial, and how it can be used to add support
for social networking and gadgets to existing web
applications and specifically Apache Roller and Apache
JSPWiki. The session will cover: - Quick introduction
to OpenSocial and Shindig - Overview of products/
services that leverage OpenSocial - Benefits of social
networking in blogs and wikis - How to support
Google Gadgets in Roller and JSPWiki via Shindig -
How to enable social features in Roller and JSPWiki
via Shindig - How to add comprensive social graph
support to Roller and JSPWiki via SocialSite - How to
create an OpenSocial Applications that access Roller
and JSPWiki

Deploying to the Edge with CouchDB

J. Chris Anderson
16:30 | St. Gallens   **D** **E**

Standalone CouchDB applications can be deployed
directly to the client. This allows us to harness the
power of full-stack view-source, along with the
freedom of peer-based replication. We're building
a vibrant ecosystem of code, developers and users.
I'll share some of the challenges and implications
of deploying to the edge of the network, and maybe
even change the way you think about the web. I will
focus on the affordances provided by CouchDB's
replication, for building applications that take
advantage of portable code and data. As a result, users
can run powerful web-apps on their local machines
(or remote CouchDB nodes) without any change in
code. View-source on standalone Couch Apps gives
users access to "server side" components as well as
browser code. Accessing the full application stack
locally has positive repercussion for open source. More
importantly, it enables network effects as people build
and share CouchDB applications. There are important
constraints to follow when you're building apps that
can deploy to any running CouchDB instance. I'll
demo the CouchApp framework, which makes it easy
to work with the grain of CouchDB, ensuring that
your applications are portable across nodes, as well
as scalable over large clusters. Applications deployed
to the edge don't have to run in a vacuum. I'll also
describe how CouchDB replication flows can be used
to setup publishing, messaging, and other social
platforms.

29

Trainings

A Day of REST
J Aaron Farr
Monday Full-Day Training

In this workshop, we cut through the haze and the buzz and examine working use cases of web services and how they can improve your application design and bottom line. We'll explore how to create services that work in harmony with the web using a REST style approach. Learn how to turn your application into a platform by creating web APIs. Both client and server side programming will be covered using technologies ranging including Ajax, Ruby, Python and Java.

Learning CouchDB: Next-Generation Web Storage At Any Scale
J. Chris Anderson
Monday Morning Half-Day Training

Apache CouchDB can serve pure Ajax applications, using no middle-tier application server. Standalone CouchDB applications can be very quick and easy to develop and maintain, as well as being easily portable. They are ideal for in-house or education projects, as well as opening up novel opportunities using CouchDB's peer-based replication. We will build relationship tracking groupware using only HTML, JavaScript, and CouchDB. Users will be able to manage contact information, as well as add to do items tagged with different people and contacts. We'll design the application for sharing, so that a workgroup can maintain individual copies of the data, but also circulate updates to other members of their group, either on an ad-hoc p2p basis or via a central server. Participants will learn the fundamentals of data management with CouchDB, as well as the capabilities CouchDB makes available due to its HTTP/REST interface. We'll cover CouchDB's schema free JSON document model, include its implementation of MVCC. We'll learn to use CouchDB's MapReduce views to make orderly queries from heterogeneous data. We'll demonstrate CouchDB's peer-based replication, and understand it's implications for data portability. Upon completing the workshop, participants will be ready to use CouchDB in real-world applications. Using CouchDB's HTTP API via Ajax applications is great preparation for working with CouchDB in more "heavyweight" deployment scenarios. We also hope to show participants that the standalone model may be the best way to build applications for in-house deployment, education, or other scenarios where users are relatively trusted.

High Performance CouchDB
J. Chris Anderson
Monday Afternoon Half-Day Training

Apache CouchDB is a document-based, rather than relational, database, which provides a simple JSON/HTTP API. It can be used for web apps, but also has a sweet spot for giant data stores and high-transfer rate applications. In this workshop we'll cover what you need to know, when you are pushing your hardware to the limits in a high-performance environment. We'll cover: interfacing to CouchDB directly via Erlang; techniques for minimizing disk seeks; optimizing Map Reduce views for application data access patterns; providing reliable transactional saves using replication, redundant nodes, and update notifications.

Lucene Boot Camp
Grant Ingersoll
Monday & Tuesday Two-Day Training

Apache Lucene is a high-performance, cross-platform search API in production use in a large number of applications. This tutorial will be an in-depth, hands-on class targeted at developers who wish to integrate Lucene Java into their applications. The two day tutorial will cover the key concepts behind Lucene, along with code examples, documentation and resources. Concepts covered are: Lucene setup; indexing content, indexing customization; searching, including how to use the many Lucene Query APIs to achieve better search results. Additionally, we will cover advanced topics like Lucene performance and threading issues, Span Queries, filters, term vectors, and sorting, all of which can help solve common problems in search applications. The class will gain hands-on experience with Lucene and leave with the know-how to build a Lucene-based application. No prior Lucene knowledge is required, but participants must be able to code in the Java programming language and come prepared to code. Concepts learned can also be applied to Solr. Students are encouraged to bring their own content to index/search. Please refer to the Lucene Boot Camp website (http://www.lucenebootcamp.com) for more information on the class.

Apache HTTP Server - Nuts to Bolts
Jim Jagielski
Rich Bowen
Monday & Tuesday Two-Day Training

Similar to the training at AC US 08 in New Orleans, this training provides a Nuts to Bolts Guide to Apache HTTP Server 2.2. This includes the basic design and architecture of the server, an overview of the API, deep dives into various useful modules. It also looks at some of the deficiencies of the current design and what is on the horizon for 2.4/3.0.

Hadoop Tools and Tricks for Data Processing Pipelines
Aaron Kimball, Christophe Bisciglia & Tom White
Part One Monday Morning Half-Day

There's more to Hadoop than just getting your cluster running. In this session, we will look closely at data processing pipelines, and teach participants how to leverage Hadoop and related tools to conduct complex data analysis tasks.

Trainings

We will assume a functioning Hadoop cluster and properly configured supporting tools such as Hive, Pig, Streaming, etc. We will go over, in depth, at least one data processing pipeline. This will include data collection, formatting, processing and presentation. Participants will see examples of code, and the glue necessary to start with the data you have, and end with the results you want. Cloudera will provide this cluster free of charge to training participants for the duration of the conference, and provide support for user issues.

Hadoop Tools and Tricks for Data Processing Pipelines - advanced
Aaron Kimball, Christophe Bisciglia & Tom White
Part Two Monday Afternoon Half-Day

This session is intended for intermediate Hadoop users. Users need to have used Hadoop and be familiar with the basics. Users should be comfortable with Java programming and scripting in their language of choice. We will not cover basics or cluster configuration, but will focus on enabling people with the need to process large amounts of data to leverage Hadoop and related tools such as Hive, Pig, and Streaming. Target users may have large volumes of log files or user data that they would like to explore.

Media & Analyst Training
Sally Khudairi
Monday Morning Half-Day

Prepare yourself for a variety of media opportunities with communication skills that resonate with editors and analysts and help generate positive coverage. Learn how to craft, control, and bridge messages and gain confidence during the interviewing process. Present your position clearly and become more persuasive to effectively target editors and influence the analyst community. Hone your presentation techniques -- from public speaking to analyst briefings -- through group coaching, individual interviews, and breakout practice sessions. Takeaways include identifying common mistakes and tips on how to avoid them, press do's and don'ts, analyst relationship management, developing thought campaigns, positioning previews and launch tactics, and more. A must-attend for those interested in promoting their Open Source project and becoming a resource for the media community at large.

Intermediate Media & Analyst Training
Sally Khudairi
Monday Afternoon Half-Day

Perception is reality! Step up your basic media skills to become a more effective spokesperson and take control of the interviewing process. Handle sensitive topics, aggressive interviews, and misquotes with gravitas and confidence. Present yourself credibly in a variety of scenarios - press conferences, analyst briefings, communications crises, and more. Class

format includes group coaching, individual interviews, and breakout practice sessions. Hands-on, tactical tips on improving your written media skills to mitigate the reality of overworked reporters and make your efforts more credible to analysts and the community at large. NEW! guidelines on working with the ASF Public Relations Committee. NOTE: Restricted class size; limited to participants who have already completed basic media/analyst training.

Everything Tomcat - Administering, Tuning, Troubleshooting and Developing
Mark Thomas
Monday & Tuesday Two-Day Training

The fastest way for you to learn everything you need to know about Tomcat. If you are an administrator, developer, release engineer, or perhaps a bit of all, then this training is for you. Whether you are an expert or a beginner, you are guaranteed to learn something new in this class. The curriculum spans from basic Tomcat configuration to advanced tuning of both Tomcat and the JVM. We'll go over several different deployment methodologies, with and without web server connectors, load balanced or not, and how to maximize the performance of Tomcat in your infrastructure. You'll learn how to troubleshoot Tomcat, how to classify different types of error messages and how to correctly reconfigure your server based on the actual error. Towards the end we will take a deep dive into the core of Tomcat, learning both architecture and code, and how to write your own Tomcat

Introduction to Wicket
Martijn Dashorst
Monday Full-Day Training

Learn how to use Apache Wicket to create web applications on your own from the masters. This hands-on lab will provide a quick introduction to the Wicket framework and we'll start with coding right away. At the basis for this course lies the Wicket in Action book, written by the course leader. We'll start with setting up our project, move on from a simple hello world application to implementing an online cheese store. We'll learn to connect it to services delivered by Spring and a back end served with a JPA provider (Hibernate or OpenJPA). During this course we'll cover the end-to-end basics of web application development: unit testing, writing maintainable code, internationalization, security and deployment.

Solr Boot Camp
Erik Hatcher
Tuesday Full-Day Training

Solr has emerged as THE way to add search services to applications written in any language. Lucene powers Solr's RESTful web services. This all-day tutorial will begin with a Solr overview, and move through

31

Trainings

installation and configuration. The bulk of session will be spent diving deeply into integrating Solr with various technologies. Solr features caching, replication support for load balancing, faceting, highlighting, more-like-this, simple HTTP protocol supporting XML, JSON, and other formats. Erik will discuss his Solr experiences on the applications he has built with it.

Maven Training for Junior & Medior Java Developers
Zeger Hendrikse
Tuesday Morning Half-Day Training

Overview: In this training, you will learn how to organize your product lifecycle management with Maven. Participants are introduced to both Maven as a concept and a tool. In addition to a number of exercises, some demonstrations, tips & tricks and best practices will be given.

This training is geared towards junior and medior Java developers interested in Java/JEE development with Maven. Exposure to web application development with is recommended, as well as some familiarity with basic Java frameworks such as JUnit and Log4j.

Objectives:
• A concise overview of Maven
• How to get started with Maven
• Integration with Eclipse IDE
• Hands-on experience with simple Maven projects
• Understanding of the benefits of Maven
Preparation

Lab Work: It is necessary to bring your laptop. In order to maximize your time available for the exercises, it is recommended to have Eclipse 3.4 (Ganymede) preinstalled on your laptop together with the Maven plug-in: http://m2eclipse.sonatype.org/update

Maven Training for Senior Java developers
Zeger Hendrikse
Tuesday Afternoon Half-Day Training

Overview: In this workshop, you will learn how to organize your product lifecycle management with Maven. Participants are introduced to both Maven as a concept and a tool. In addition to a number of exercises, some demonstrations, tips & tricks and best practices will be given. Exercises for getting acquainted with continous integration are also provided. This workshop is geared towards senior developers who have experience in developing JEE web applications. You should be reasonably comfortablewith frameworks such as Hibernate, Spring, JUnit and Log4j, as webapplication development with Maven is a substantial part of this workshop.

Objectives
• An concise overview of Maven
• Integration with Eclipse IDE
• Hands-on experience with Maven projects
• Generation of a Maven project site
• Multimodule projects
• Getting started with continuous integration
• Understanding of the benefits of Maven
General information

Lab Work: It is necessary to bring your laptop. In order to maximize your time available for the exercises, it is recommended to have Eclipse 3.4 (Ganymede) preinstalled on your laptop together with the Maven plug-in: http://m2eclipse.sonatype.org/update

Behavior-Driving Your Apache Wicket Application: Making the Most of Webdriver and JDave-Wicket
Timo Rantalaiho, Markus Hjort & Marko Sibakov
Tuesday Full-Day Training

Apache Wicket is a top-quality component-oriented web user interface framework for Java. JDave is a powerful Behavior-Driven Development (BDD) framework for those willing to switch Test-Driven Development (TDD) on Java to a higher gear. WebDriver is an excellent functional web testing tool for testing a deployed web application with a real browser. This hands-on training session will teach you how to get good test coverage by expressive, executable specifications on your Wicket code. The trainers have extensive experience on delivering complex, ajaxified real-world applications using these technologies.

First JDave, WebDriver and BDD philosophy will be briefly introduced, then practical examples of jdave-wicket and using jdave and WebDriver with Wicket demonstrated. The bulk of this training session will consist of hands-on programming assignments of applying jdave-wicket and WebDriver for testing and adding features to an existing Wicket application.

For the participants, a good command of Java is required, and knowledge on Apache Wicket and some unit testing or BDD framework for Java (such as JUnit, TestNG, JBehave or JDave). Experience on Wicket, WebDriver, Selenium, jMock and Hamcrest matchers is a plus. Eclipse or IntelliJ IDEA is preferred as an IDE. Participants should have a laptop, but can also work in pairs.

Ajax on Struts 2: How a Second Generation Web Application Framework Meets the Demands of RIA
Chad Michael Davis
Tuesday Full-Day Training

Struts 2 is a brand new, second generation web application framework. Two criteria by which all new frameworks should be judged are ease of development and compatibility with RIA techniques. If a new

Trainings

framework is worth it's hype, it should ease your carpal tunnel syndrome (less typing!) while simultaneously making Ajax, and other RIA techniques, as natural and convenient as traditional models of web application development. This training will try to provide two things. First, we'll give everyone a hands-on introduction to Struts 2 development by coding a "Hello World on steroids". In addition to demonstrating the Struts 2 development flow, this pumped-up Hello World will also demonstrate how to do Ajax on Struts 2. You'll get a peak at a day in the life of a Struts 2 developer, and you'll also see how the framework facilitates clean Ajax development. It's part intro, part best practices. Bring your laptop and be ready to code. Ideally, have a Java IDE and a servlet container ready to go. Team programming will be encouraged ; teams of two should work best -- one typist and one "listener". We're going to move fast, but it should work.

Apache CXF - Developing and Deploying Open Source SOA Endpoints
Adrian Trenaman
Tuesday Full-Day Training

This one-day tutorial will show how to design, build and deploy RESTful and Web-based web services using Apache CXF. The session will cover code-first and WSDL first services design for JAX-WS, and show how to use JAX-RS for restful services. We'll show how to swap different transports (HTTP, JMS) with different payload types (XML, SOAP, JSON). And, we'll show how CXF's support for JavaScript can be used to create client- and serverside components for AJAX applications. This workshop will combine a mix of lectures, whiteboarding, and hands-on exercise sessions.
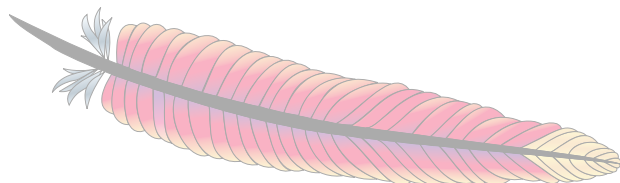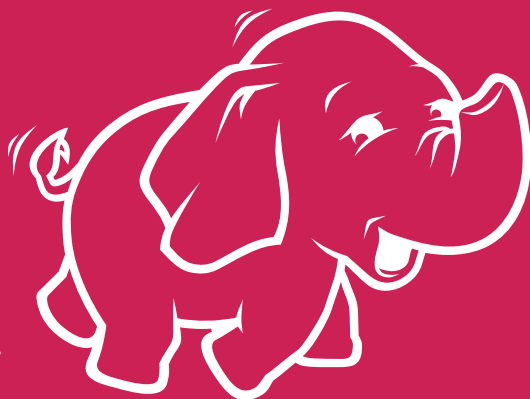
Speakers

Ajay Anand

Ajay Anand is Director of Product Management for Grid Computing at Yahoo!. Ajay was product manager of Sun's first high availability file and database servers and then worked in product and marketing management roles in the areas of storage management, middleware, and identity management. Previously he was Director of Product Management for SGI's storage products and Aspect's customer management middleware. Ajay holds an MS in Computer Engineering and an MBA from the University of Texas at Austin, and a BSEE from the Indian Institute of Technology.

J. Chris Anderson

J. Chris Anderson is co-founder of Grabb.it, a Portland-based web startup which provides Web scale analytics for the music industry. He specializes in web search and large scale CouchDB deployment, both for Grabb.it and as an independent consultant. He is a CouchDB Committer and O'Reilly author.

Jeanfrancois Arcand

Jean-Francois Arcand works for Sun Microsystems. He currently leads project Grizzly, an extended NIO based framework used in multiples projects. He also works on Web 2.0 topics like Ajax performance and leads the Comet activities at Sun. Jean-Francois lives and works from home in Prevost, a very small city in Québec where life is perfect.

Afkham Azeez

Afkham is a committer & Project Management Committee (PMC) member of the Apache Web Services project. He works as a software architect at WSO2 and is the Product Manager of WSO2 Web Services Application Server (WSAS). He also leads the WSO2 Carbon framework project. His areas of interest include distributed computing, SOA, Highly Available & Highly Scalable applications & J2EE technologies. His current research area is cloud computing and he is currently trying to make Web services autoscale on Amazon EC2. He is the author of the clustering implementation for Apache Axis2. He holds a B.Sc. first class honors degree in Computer Science & Engineering from the University of Moratuwa, Sri Lanka, and is currently reading for a Masters Degree in Computer Science specializing in Software Architecture.

Christophe Bisciglia

Christophe Bisciglia joins Cloudera from Google, where he created and managed their Academic Cloud Computing Initiative. Starting in 2007, he began working with the University of Washington to teach students about Google's core data management and processing technologies - MapReduce and GFS. This quickly brought Hadoop into the curriculum, and has since resulted in an extensive partnership with the National Science Foundation (NSF) which makes Google-hosted Hadoop clusters available for research

and education worldwide. Beyond his work with Hadoop, he holds patents related to search quality and personalization, and spent a year working in Shanghai. Christophe earned his degree, and remains a visiting scientist, at the University of Washington.

Uri Boness

Uri Boness is a software engineer and architect with more than eight years experience developing Java applications. Currently, Uri serves as the chief scientist at JTeam. Over the years, Uri has contributed and initiated several open source projects, including FacetSearch—a lucene based faceted classification framework. This project was eventually stopped in favor Solr. In the last few years Uri has delivered several Solr and Lucene based projects and gave several training sessions on these subjects. His main technology interests are focused on search and RIA, which he considers to be the real backbone of the new age of enterprise applications.

Rich Bowen

Rich is a member of the Apache Software Foundation, and a contributor to the Apache Web Server documentation effort. He's the author of Apache Cookbook, and of The Definitive Guide to

mod_rewrite. By day, Rich is the Director of IT at ClearMyRecord.com, helping ex-offenders get a second chance at life.

Mark Brewer

Mark Brewer is Senior Vice President and GM of Enterprise Delivery at SpringSource. Brewer is a seasoned executive with more than 20 years of experience within the technology and software industry. Mark was previously CEO of Covalent Technologies, which SpringSource acquired in early 2008. Mark joined Covalent Technologies, the leading provider of support and services for Apache Software Foundation (ASF) open source projects, in late 2001 as Vice President of Field Operations and took over as CEO in early 2004. In this position, Mark was responsible for the day-to-day management of the company, including overseeing worldwide sales, marketing and business development initiatives. Prior to Covalent, Mark served as Vice President of Worldwide Sales and Business Development for Banter Technologies where he built successful business models for selling eCRM solutions to large financial institutions and enabling technologies to CRM software vendors.

Previously, Mark was Vice President of Sales at Andromedia helping to define and build the sales model and partner strategies for the leading provider of enterprise web site analysis and personalization solutions provider. Mark was instrumental in the successful merger between Andromedia and Macromedia (now Adobe)in late 1999.

Before Andromedia, Mark held a variety of sales management and executive positions with Powersoft (PWRS), which was acquired by Sybase (SYBS) in 1995, during the company's rapid growth phase in the early and mid 1990s.

Mark holds a B.S. in Biology with minors in Chemistry and Computer Science from Western Washington University.

Nick Burch

Apache Member PMC Chair for POI

Michael Busch

Michael Busch is a Software Architect at the IBM Silicon Valley Lab in San Jose, California. He has been working for 3+ years on Enterprise Search applications, such as IBM's OmniFind and eDiscovery products. He has a Bachelors Degree in Computer Science from the University of Paderborn, Germany. Michael is a committer on the Apache Lucene Java project.

Arjé Cahn

Arjé Cahn is CTO of open source content management developer Hippo (Amsterdam, The Netherlands). He is an Apache Cocoon committer, Apache Software Foundation Member and a regular speaker on the subjects of Open Source, XML

and content management. He has written for both technical and business publications on Apache and open source.

Maria Odea Ching

Maria Odea Ching has been a committer at the Apache Software Foundation since 2006. She is currently serving as chair of the Apache Archiva project management committee and is also involved in the Apache Continuum and Apache Maven projects. She is a Software Engineer at Exist Global, Inc.

Jean-Frederic Clere

Jean-Frederic has spent more than 20 years writing client/server software. He is committer in APR, Jakarta, Httpd and Tomcat and he likes complex projects where different languages and machines are involved. Borne in France, Jean-Frederic lived in Barcelona (Spain) for 14 years. Since May 2006 he lives in Neuchatel (Switzerland) where he works for RedHat in the JBoss division.

Danese Cooper

Danese Cooper has a 20-year history in the software industry and has long been an advocate for transparent development methodologies. Danese worked for six years at Sun Microsystems, Inc. on the inception and growth of the various open source projects sponsored by Sun (including OpenOffice.org, java.net and blogs.sun.com). She was Sun's Chief Open Source Evangelist and founded Sun's Open Source Programs Office. She has unique experience implementing open source projects from within a large proprietary company. She joined the OSI Board in December 2001 and currently serves as Secretary & Treasurer. She is also a Member of the Apache Software Foundation and serves on a special Advisory Board for Mozilla.org. In 2005 Danese joined Intel to advise on open source projects, investment and support. In 2009 she joined REvolution Computing, her commercial open source startup. She speaks internationally on Open Source and Licensing issues.

Martijn Dashorst

Martijn Dashorst is a PMC member for the Apache Wicket project and a committer for the Incubator project. Martijn has presented Wicket at several conferences, including JavaOne, and JavaPolis. He works for Topicus, an appliation service provider for health care, financial and education institutions based in the Netherlands.

Chad Michael Davis

Chad Michael Davis, of Blackdog Software, Inc., is a software consultant specializing in Java Enterprise Edition development. In addition to building Java web applications, he exercises his academic and engineering inclinations by writing about software and software development. His publications include a chapter on open source software development for the 2005

Speakers

edition of the yearly Advances in Computers, and the book Struts 2 in Action (Manning Publications, 2008). He has an MS in Computer Science (University of Arkansas at Little Rock) and an MA in English (Colorado State University).

Bertrand Delacretaz

Bertrand Delacretaz works as a senior developer in Day Software's R&D group (http://www.day.com), using open source tools to create world-class content management systems. Bertrand is a member and director of the Apache software foundation and has been or is involved in the Apache Cocoon, FOP, Solr, Sling, Tika and ESME projects.

Lars Eilebrecht

Lars is co-founder and member of The Apache Software Foundation, and started contributing to the Apache web server project in 1997. In addition, he is the Vice President of the Conference Planning Committee, a member of the Apache security team, and the Apache public relations committee. He has a degree in computer engineering from the University of Siegen, Germany, where he wrote his first book about the Apache web server. He held various senior engineering, consulting and management positions at various ISPs, mobile network providers and software development companies. Currently he works as a freelance IT consultant specializing in IT security and Open Source.

Justin Erenkrantz

Justin R. Erenkrantz is currently President of The Apache Software Foundation. He is also a Senior Software Engineer at Joost and is completing his PhD at the University of California, Irvine. He has been a long-time contributor to the Apache HTTP Server, Subversion, APR, Serf, mod_mbox, and flood.

J Aaron Farr

J Aaron Farr currently serves on the Apache Software Foundation board of directors and as the ASF's treasurer. He has contributed to the Avalon, Excalibur, Incubator and Labs projects. After developing J2EE and Eclipse RCP software for Sony Electronics and Siemens, he founded The JadeTower Corporation. Aaron currently resides in Hong Kong.

Paul Fremantle

Paul Fremantle is CTO of WSO2, where he works on Open Source projects in Apache, including the Apache Synapse and Incubator projects. He has contributed to Apache since the first Apache SOAP project. While at IBM, he was instrumental in starting up the Apache WSIF, and Apache Woden projects, as well as being heavily involved in the AxisC/C++ initiative, where he led IBM's involvement. Paul was a Senior Technical Staff Member in IBM, where he was the lead architect and co-creator of IBM's Web Services Gateway. Paul is the co-chair of the OASIS WS-RX technical committee and lead the JSR110

committee (JWSDL). Before joining IBM, Paul worked as a consultant in the pharmaceutical industry. Publications include co-authoring "Building Web Services in Java, 2nd Edition", articles on Web Services and SOA, and a redbook - "The XML Files: Using XML and XSL in WebSphere". Paul has presented at ApacheCon, Colorado Software Summit, XML Europe, Software Architecture and other industry conferences. Paul has an M.A. in Mathematics and Philosophy and an M.Sc in Computation from Oxford University.

Andreas Gies

Andreas started his consulting career in the mid 80s even before received his diploma from the University of the German Forces in Munich. Right after his studies he dove into the middleware space, beginning on IBM mainframes using MQ Series working his way through CORBA, J2EE, Web Services and ESB based architectures. After leaving the German Air Force he worked for Sterling Software, IONA Technologies and the last 7 years for PROGRESS Software as Principal Architect specialized on large scale deployments. When he is not hip deep in software architectures he enjoys his family life with a wife and 2 daughters in Bavaria, teaching Yoga and hiking in the Bavarian Forrest.

James Governor

James Governor is Principal Analyst and co-founder of RedMonk, the first analyst firm built on open source principles and license styles. He works with technology vendors and enterprises, advising on community development and platform selection strategies. Clients range from the biggest to the smallest firms- from IBM (350k employees) to Reductive Labs (4 employees). James makes extensive use of social networks, notably Twitter, in driving and delivering his research agenda and supporting the community that makes it possible. James is also co-author of the forthcoming O'Reilly publication – Web 2.0 Architectures: What entrepreneurs and information architects need to know. James is also deeply interested in sustainability issues, and drove the formation of RedMonk's new consulting and advisory service Greenmonk. He also acts as an advisor to Akvo.org, the open source for water knowledge.

Filip Hanik

Filip is a Senior Software Engineer for SpringSource and a key participant in SpringSource's Apache Tomcat initiatives. Filip brings 12 years of extensive experience in architecture, design and development of distributed application frameworks and containers and is recognized for his top-quality system development skills and continuous participation of Open Source development projects. Filip is a committer to the

Apache Tomcat project where he is a leading authority on Tomcat clustering and a key contributor to the core of the platform. He also authored the online book Hitch-Hiker's Guide to Tomcat. Prior to SpringSource, Filip was a Senior Software Enginieer for Covalent Technologies, which was acquired by SpringSource in early 2008. Previously, Filip worked as a Senior Software Engineer/Architect for La Quinta Corporation. Filip has also made contributions to software initiatives for Walmart.com, Sony Music, France Telecom and has held a variety of senior software engineering positions with technology companies in both the United States and Sweden. He received his education at Chalmers University of Technology in Gothenburg, Sweden where he majored in Computer Science and Computer Engineering.

Erik Hatcher

Erik Hatcher is an Apache Software Foundation member, and an active committer on the Lucene and Solr projects. Erik has co-authored the award-winning book Java Development with Ant (Manning) and the well reviewed Lucene in Action (Manning). Erik has spoken frequently at industry conferences, including JavaOne, ApacheCon, OSCON, and the No Fluff, Just Stuff symposium circuit. Erik co-founded and member of the technical staff of Lucid Imagination, a full service Lucene/Solr shop.

Zeger Hendrikse

Zeger is a focused and driven professional with extensive experience in various projects at industrial and governmental organizations. His activities range from application development to consultancy, including knowledge transfer by motivating and coaching junior and medior developers and teaching courses product lifecycle management. He is known to be a good team player, a strong motivator and an enthousiastic teacher.

Markus Hjort

Methodology Specialist at Reaktor Innovations, Markus has been in the industry for over nine years with experience from various technologies. Certified Scrum Master with extensive experience on agile methods and a long-time active participant in process improvement wherever he's worked at, Markus kick-started the local Coding Dojo events in Finland back in 2005. He is one of the pioneers of the Finnish agile community, speaks frequently at international conferences, and has years of experience on using Wicket in large scale projects.

Grant Ingersoll

Grant Ingersoll is a committer on the Lucene Java, Solr and Mahout projects and a member of the Lucene PMC. Grant has used Lucene and Solr for implementing many different search and text-based applications, including cross-language information retrieval systems, question answering, sentiment analysis, collection analysis and customization applications. Grant is also a co-creator of the Apache Mahout machine learning project.

Jim Jagielski

Jim's been active on the 'Net since the early 80's, starting as editor of the A/UX FAQ. He worked on the NCSA server and joined the Apache Group (as it was called back then) at a very early stage. He actively contributes on HTTPD, APR and Tomcat, but also hacks on other projects (ASF and others) as well in addition to mentoring many ASF incubator podlings. In addition to being a charter and core member of the ASF Jim serves as Director and Chairman for the foundation. His real job is as Chief Architect and Principal Software Engineer for the Covalent Products Division of SpringSource. Previously, Jim served at CTO for Covalent Technologies, which was acquired by SpringSource in early 2008.

David Jencks

After studying mathematics in college and graduate school, David Jencks spent 14 years building baroque musical instruments. Returning to an early interest in programming, he has been involved with open source projects since around 2000. A founding member of Apache Geronimo, he is also a committer on ActiveMQ, Directory, Portals, and OpenEJB at Apache and Jaybird, Jencks, Mojo, and Tranql elsewhere. He has been an IBM employee since 2005.

Dave Johnson

Dave Johnson is a software developer, blogger and web geek from Raleigh, North Carolina. He is a member of the Apache Software Foundation working on Apache Roller open source blog server project. And, until recently, worked as a social software architect at Sun Microsystems, Inc. working on Roller, SocialFish and Project SocialSite.

Sally Khudairi

Active in the Web since 1993, Sally Khudairi is the publicist behind some of the industry's most prominent standards and organizations. The former deputy to Sir Tim Berners-Lee and long-time champion of collaborative innovation, she became involved with The Apache Software Foundation in 1999, and was elected its first female and non-technical member. Sally is chief executive of luxury brand communications consultancy HALO Worldwide.

Aaron Kimball

Aaron Kimball has been working with Hadoop since early 2007. Aaron has worked with the NSF and several other universities nationally and internationally to advance education in the field of large-scale data-intensive computing. He helped create and deliver academic course materials first used at the University of Washington, which were later adopted by many

Speakers

other academic institutions, as well as Hadoop training materials used by several industry partners. Aaron has also worked as an independent consultant focusing on Hadoop and Amazon EC2-based systems. Aaron holds a B.S. in Computer Science from Cornell University, and an M.S. in Computer Science and Engineering from the University of Washington.

Emmanuel Lecharny

Emmanuel is the current Apache Directory Server Chairman, and an ASF member. With more than 20 years in the field, he still likes to develop, which is one of the reason he participates in the Directory project. System architecture and optimizing existing applications is what feeds him on a daily basis...

Graham Leggett

Graham is a member of the Apache Software Foundation and started contributing to the httpd and subsequently the APR project from 1999. He has a degree in electrical engineering from the University of the Witwatersrand in South Africa, and is currently a consultant in the field of security and software engineering.

Ruwan Linton

Ruwan Linton is a Senior Software Engineer working for WSO2 Inc. He is also a committer of Apache Software Foundation and currently working on the Apache Synapse and WS projects as an active developer as well as a PMC member on Apache Synapse. He has successfully completed a degree of B.Sc. Engineering in University of Moratuwa in Sri Lanka, specialized in Computer Science and Engineering. I am a committer & Project Management Committee (PMC) member of the Apache Web Services project. I work as a software architect at WSO2 and am the Product Manager of WSO2 Web Services Application Server (WSAS). I also lead the WSO2 Carbon framework project. My areas of interest include distributed computing, SOA, Highly Available & Highly Scalable applications & J2EE technologies. My current research area is cloud computing and am currently trying to make Web services autoscale on Amazon EC2. I am the author of the clustering implementation for Apache Axis2. I hold a B.Sc. first class honors degree in Computer Science & Engineering from the University of Moratuwa, Sri Lanka, and am currently reading for a Masters Degree in Computer Science specializing in Software Architecture.

Steve Loughran

Steve Loughran is a senior researcher at HP Laboratories, exploring the challenges of large-scale systems. He is an Apache Member and a longstanding Committer on the Ant project, and is the author of Ant in Action. Since August 2008 he's been a committer on the Hadoop project, and is currently busy getting Hadoop to run in HP's datacenters.

Norman Maurer

Norman Maurer has been working for almost 7 Years as Unix System Engineer. At the moment he work for an Internet Service Provider as Senior Unix System Engineer. This Job includes implementing business critical LINUX / UNIX Systems based on Open-Source and commercial software. In the past he worked most time for Internet Service Providers. He joined the Apache Software Foundation in 2006 as committer for the JAMES project and was elected as PMC member later. After some volunteering he joined the Apache Software Foundation Infrastructure Team in 2008. Finally he was voted as Member. When he has some free time, he try to travel to nice places all over the world.

Felix Meschberger

Felix Meschberger works as a senior developer for Day Management AG (http://www.day.com) creating content management systems using open source and internal tools. Felix is a committer to the Apache Jackrabbit and Felix projects and contributed Apache Sling to the Incubator in September 2007.

Arun Murthy

Arun C Murthy is a long-term contributor and committer on the Apache Hadoop project. He is also a member of the Apache Hadoop PMC. http://people.apache.org/~acmurthy/

Olga Natkovich

Olga Natkovich leads Pig development team at Yahoo. She has over 10 years of development experience building large distributed systems and holds several patents in this area. She has Master Degree in Computer Science from University of Illinois at Urbana-Champaign

Owen O'Malley

Owen O'Malley is a Software Architect in Yahoo's Grid Computing team and is the chair of the Hadoop PMC. He received his PhD in Software Engineering from University of California, Irvine.

Marcel Offermans

Marcel is a software architect at luminis with broad experience in different operating systems, languages and applications, ranging from data-oriented enterprise applications to embedded and distributed systems. At luminis iQ products, he is responsible for the architecture and development of an OSGi based software provisioning product. Furthermore a PMC member of the Apache Felix open source project and interested in all Java and OSGi related development.

Joe Orton

Speakers

Joe Orton has been writing free software for ten years and is the author of a number of open source WebDAV tools such as "cadaver". He works in the UK for Red Hat as maintainer of web-server packaging, from Apache httpd to PHP and beyond. Joe is a member of the Apache Software Foundation, and is a committer on the httpd and APR projects.

Gianugo Rabellino

I'm Gianugo Rabellino, Open Source enthusiast, passionate golfer, foodie and geek at heart. I'm the founder and CEO of Sourcesense, a European Open Source consultancy and systems integrator, a member of the Apache Software Foundation and a guy who's been paying his bills with Open Source for more than 15 years now. Did I mention I'm a lawyer?

Raghu Ramakrishnan

Raghu Ramakrishnan is Chief Scientist for Audience and Cloud Computing at Yahoo!, and is a Research Fellow, heading the Web Information Management group in Yahoo! Research. He is Professor of Computer Sciences at the University of Wisconsin-Madison (on leave), and was founder and CTO of QUIQ, a company that pioneered question-answering communities, powering Ask Jeeves' AnswerPoint as well as customer-support for companies such as Compaq. Ramakrishnan's research is in the area of database systems, with a focus on data mining, query optimization, and web-scale data management, and has influenced query optimization in commercial database systems and the design of window functions in SQL:1999. His paper on the Birch clustering algorithm received the SIGMOD 10-Year Test-of-Time award, and he has written the widely-used text "Database Management Systems" (with Johannes Gehrke).

He is Chair of ACM SIGMOD, on the Board of Directors of ACM SIGKDD and the Board of Trustees of the VLDB Endowment, and has served as editor-in-chief of the Journal of Data Mining and Knowledge Discovery, associate editor of ACM Transactions on Database Systems, and the Database area editor of the Journal of Logic Programming. Ramakrishnan is a Fellow of the Association for Computing Machinery (ACM) and the Institute of Electrical and Electronics Engineers (IEEE), and has received several awards, including the ACM SIGKDD Innovations Award, the ACM SIGMOD Contributions Award, a Distinguished Alumnus Award from IIT Madras, a Packard Foundation Fellowship in Science and Engineering, and an NSF Presidential Young Investigator Award.

Timo Rantalaiho

Timo Rantalaiho is a senior software designer at Reaktor Innovations, with extensive experience on applying agile methodology and practices in demanding, business-critical, Java-based systems. An Apache Wicket committer and JDave contributor, he's passionate about constantly improving code quality in his projects.

Carlos Sanchez

Carlos Sanchez has been involved in Open Source for nearly ten years. He has specialized in solving business challenges in a wide variety of industries, including e-commerce, financial services, telecommunications, and software development. A member of the Apache Maven Project Management Committee (PMC), Carlos is one of the maintainers of the central Maven repository and a co-author of "Better Builds with Maven," the first book about Maven 2. He is currently Director at Exist Global in Los Angeles. Carlos received his Computer Engineering degree at the University of Coruña, Spain, and enjoys traveling and discovering new places.

Thorsten Scherler

Thorsten is a committer and PMC member on the Apache Cocoon, Forrest and Lenya project. He is the principal investigator of Apache Droids. Thorsten uses ASF products professionally for customer projects and teaching classes.

Marko Sibakov

Marko Sibakov is Senior Consultant at Reaktor Innovations. Marko has many years of experience about planning and implementing successful software projects, making heavy use of Wicket and JDave in the last years and lately also of WebDriver. As a Certified Scrum Master Marko's passion is to use Agile methods and best possible engineering practices in every project he does. Right now Marko is working with his team for the Finland Post (Itella).

Wendy Smoak

Wendy Smoak is a member of The Apache Software Foundation and a committer on several projects, where she focuses on user support, documentation, and infrastructure. She is a consultant with Exist Global, where her work centers around enterprise adoption of Apache Maven and related technologies.

Michael Stack

Michael is an HBase committer and a member of the Hadoop PMC. He works for Microsoft Live Search.

Mark Thomas

Mark has been using and developing Tomcat for over five years. He first got involved in the development of Tomcat when he needed better control over the SSL configuration than was available at the time. After fixing that first bug, he started working his way through the remaining Tomcat bugs and is still going. Along the way Mark has become a Tomcat committer and PMC member, volunteered to be the Tomcat 4 release manager, created the Tomcat security pages, become a member of the ASF and

Speakers

joined the Apache Security Committee. He also helps maintain the ASF's Bugzilla instances. Mark has a MEng in Electronic and Electrical Engineering from the University of Birmingham, UK and is currently a Senior Software Engineer with the Covalent Products Division of SpringSource.

Adrian Trenaman

Ade is a Distinguished Consultant at Progress Software, where he currently specializes SOA and ESB based integration using the FUSE distribution of Apache-based open source - ActiveMQ, ServiceMix, CXF & Camel. Ade architects, designs and codes solutions built on FUSE; he also mentors and coaches FUSE users so that they can get the most out of open-source.

Niels van Kampenhout

Niels is responsible for software quality, documentation and training at Hippo, vendor of open source content management and portal software. When not chasing developers to have them write documentation, he can be found mountain biking in the hills of Northern California.

Tom White

Tom White has been an Apache Hadoop committer since February 2007, and is a member of the Apache Software Foundation. He is writing "Hadoop: The Definitive Guide" for O'Reilly. He works for Cloudera, a new startup providing Hadoop support. Previously he worked as an independent consultant specializing in Hadoop. Tom has a Bachelor's degree in Mathematics from the University of Cambridge, and a Master's degree in History and Philosophy of Science from the Universities of Leeds, UK, and Florence, Italy.

Allen Wittenauer

Allen Wittenauer is the Senior Systems Architect for Yahoo!'s Grid Computing group.

Jimmy Zhao

I'm the project lead of China Java User Group. Currently working for HP as a project manager of an global automated test infrastructure.

Before join HP I'm a software engineer in the Financial and Telecoms industries.

My favorite project is Apache Ant, and this is something I want share with people attend my talk. Enjoy the open source life everyday!

Feel free to talk to me if you are interested in Ant or other Open Source projects, Agile and medium-small project management.

Carsten Ziegeler

Carsten Ziegeler is senior developer for JEE and portal applications at Day Software. He is a member of the Apache Software Foundation and has been participating for more than fifteen years in several open source projects. Carsten is a member of several Apache communities and project management committees like Cocoon, Felix, Sling, Excalibur and Portals.

Jukka Zitting

Jukka Zitting has worked with open source content management for over ten years and has been focusing on content repositories and related technologies after joining the Apache Jackrabbit project three years ago. Jukka is a member of the Apache Software Foundation and serves currently as the chair of the Apache Jackrabbit PMC.
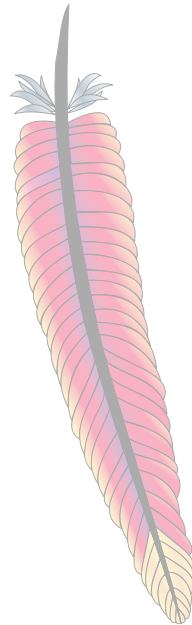
ApacheCon
Europe 2009
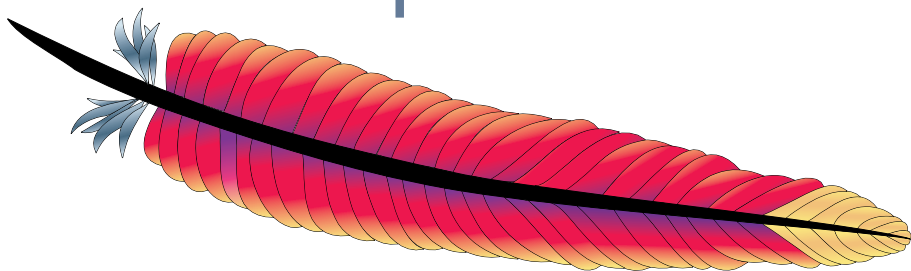
**The Official User Conference of
The Apache Software Foundation**

Thank you for helping us make ApacheCon Europe a success!

We look forward to welcoming you to our future

ApacheCon US 2009

Oakland, CA ▪ 2-6 November, 2009

http://www.us.apachecon.com

BarCamp|MeetUp Asia

China | Sri Lanka ▪ Fall 2009

http://www.apache.org/

ApacheCon Europe 2010

Spring 2010

http://www.eu.apachecon.com