



The **Apache Software Foundation**

<http://www.apache.org/>

Apache James: more than emails in the cloud

Ioan Eugen Stan

Berlin Buzzwords 2012



About myself



- Technology and Free Software Enthusiaste
- Apache James Committer
- Debian contributor and fan
- Fresh CTO for Axemblr - a company focused on tools for the cloud
- former Fire Fighter Officer

Project overview and info



Apache James Project - provide a set of (Java) libraries related to e-mail communication that build into an advanced enterprise mail server.

Use Java (or JVM language) and Mail => try
Apache James components



Apache James Project

- Apache James Hupa - rich webmail client
- Apache James Protocols - lightweight framework for implementing mail protocols (IMAP4, POP3, SMTP, etc.)
- Apache James Mailets - mail processing components
- Apache James Mime4j - library for parsing MIME documents (including email)
- Apache James Mailbox - flexible mailbox storage
- Apache James jSPF - implements SPF (Sender policy framework)
- Apache James jSieve - implements Sieve mail filtering language
- Apache James jDKIM - implements DKIM
- Apache James Server - the "One (component) to rule them all"
- Other: MPT, Postage

Apache James Mailets



Maillet = email processing agent

Maillet = matcher + maillet

Mailets are:

- flexible and powerful
- alternative to Procmail
- James Server = Maillet container
- most of email processing is done in Mailets
- you can access the envelope

Apache James Mailets



There are several groups of Mailets:

- Mailet Base - useful for developing
- Standard Mailets
- Crypto Mailet
- jSieve Mailets

You can chain mailets to form complex logic.

Apache James Mailets



Samle Camel DSL for mailet processing:

- You define 'processors'
- Test for condition and dispatch the mail to the right processor:

```
<!-- If infected go to virus processor -->
```

```
<mailet match="HasMailAttributeWithValue=org.apache.james.infected, true" class="ToProcessor">
```

```
<processor>virus</processor>
```

```
</mailet>
```

(see mailetcontainer-template.conf file)

Apache James Hupa



- A rich webmail app written in GWT
- uses IMAP to access email
- access any IMAP account with Hupa (James, GMail, Yahoo)
- [Hupa Evolution](#) - GSoC 2012 project to improve Hupa
- functional email client
- lacks some common features (address book)
- Live Demo <http://james.zones.apache.org/>

Apache James Hupa



Login as: *manuel.carrasco@gmail.com* [Contacts](#) [Logout](#)

[New](#) [Rename](#) [Delete](#)

[New](#) [Delete](#) [Delete all messages](#) [Mark seen](#) [Mark unseen](#) [Refresh](#) 25

Select: [All](#) [None](#) 1 - 25 of 8927 << First < Previous Next > Last >>

<input type="checkbox"/>		Ioan Eugen Stan	Re: work on Apache CMS site migration	23.May.
<input type="checkbox"/>		Stefano Bagnara	Re: work on Apache CMS site migration	23.May.
<input type="checkbox"/>		Ioan Eugen Stan	Re: work on Apache CMS site migration	23.May.
<input type="checkbox"/>		Eric Charles	Re: work on Apache CMS site migration	23.May.
<input type="checkbox"/>		Ioan Eugen Stan	Re: work on Apache CMS site migration	23.May.
<input type="checkbox"/>		Ioan Eugen Stan	Re: work on Apache CMS site migration	22.May.
<input type="checkbox"/>		Ioan Eugen Stan	Re: work on Apache CMS site migration	21.May.
<input type="checkbox"/>		Eric Charles (JIRA)	[jira] [Created] (JAMES-1413) AMQ blob-transfer paths should be relative	19.May.
<input type="checkbox"/>		Eric Charles (JIRA)	[jira] [Created] (JAMES-1412) Operation cannot be performed while a Tra	19.May.
<input type="checkbox"/>		Eric Charles (JIRA)	[jira] [Updated] (JAMES-1411) PathLocker may cause deadlocks	19.May.
<input type="checkbox"/>		Eric Charles (JIRA)	[jira] [Created] (JAMES-1411) PathLocker may cause deadlocks	19.May.
<input type="checkbox"/>		Stefano Bagnara	Re: svn commit: r1339519 - in /james/app/trunk: pom.xml src/main/resources/ME	17.May.
<input type="checkbox"/>		Ioan Eugen Stan	Re: svn commit: r1339519 - in /james/app/trunk: pom.xml src/main/resources/ME	17.May.
<input type="checkbox"/>		Ioan Eugen Stan	Re: svn commit: r1339519 - in /james/app/trunk: pom.xml src/main/resources/ME	17.May.
<input type="checkbox"/>		Stefano Bagnara	Re: svn commit: r1339519 - in /james/app/trunk: pom.xml src/main/resources/ME	17.May.
<input type="checkbox"/>		Ioan Eugen Stan (JIRA)	[jira] [Commented] (JAMES-1374) Migrate James web sites to Apache CMS	16.May.

More screenshots at: <http://james.apache.org/hupa/screenshots.html>

Apache James jSieve



- implements Sieve mail filtering language (RFC 3028), plus extensions
- Sieve = limited programming language to filter emails
- used in Sever to safely execute user scripts
- provides:
 - jSieve library to use in your JVM platform code
 - jSieve Mailet - enables server side filtering
 - jSieve utilities - useful, related stuff

Apache James jSPF



- A Java SPF implementation ([RFC4408](#))
- SPF (Sender Policy Framework) is designed to detect email spoofing (spam from yourself)
- used by Server component
- ready for your code!

Apache James jDKIM



- Used to fight SPAM and phishing
- java DKIM (Domain Keys Identified Mail) library - thank you Captain Obvious
- provides verification and signing for email messages in your app
- provides Mailets for use in James Server
- pluggable - you can tweel almost every behaviour

Apache James Mime4j



- java parsing library for email message streams
- supports rfc822 and MIME format
- similar to java-mail but much MORE tolerant with illegally formatted messages
- more lightweight than java-mail
- two parsing API's for your convenience:
 - event driven using MimeStreamParser (SAX style)
 - iterative using MimeTokenStream

Apache James Protocols



- lightweight, generic framework for developing mail protocols
- not bound to any socket implementation
- has a fast Netty implementation
- implementation for
 - SMTP with hooks for user extension code
 - POP3
 - IMAP4
 - LMTP
- easy to implement your own line based protocol



Apache James Mailbox

- Implements a Email Storage
- Used by James Server to store email
- Can be used by other applications to implement [Mailbox Pattern](#)
- 1 API, 4 implementations:
 - JPA, for database storage (MySQL, Postgres, [Apache Derby](#) - default)
 - Maildir - *NIX systems only
 - JCR - Java Content Repository ([Apache Jackrabbit](#))
 - Apache HBase backend
- Integration with Lucene (IMAP SEARCH)



Mailbox HBase details

- HBase client - supply an hbase-settings.xml config and you are ready to go
- Simple schema:
 - JAMES_MAILBOXES
 - JAMES_MESSAGES
 - JAMES_SUBSCRIPTIONS
- [GSoc 2012 project: Distributed mailbox indexing](#) (Lucene HBase integration) to provide search for messages stored in HBase mailbox

Mailbox HBase schema

Mailbox shema design

Table: MAILBOXES	
Row Key	Columns
Mailbox unique ID (UUID)	Count, last UID, name, namespace, user, uid validity

Table: SUBSCRIPTIONS	
Row Key	Columns
user_name@domain	Mailbox id's that the user is subscribed to.

Table: MESSAGES	
Row Key	Columns
mailboxUID:mail UID* *stored in reverse: newer messages first	Meta-data: flags, media type, date ... (in memory)
	Message headers
	Message body - split into fix sized chunks



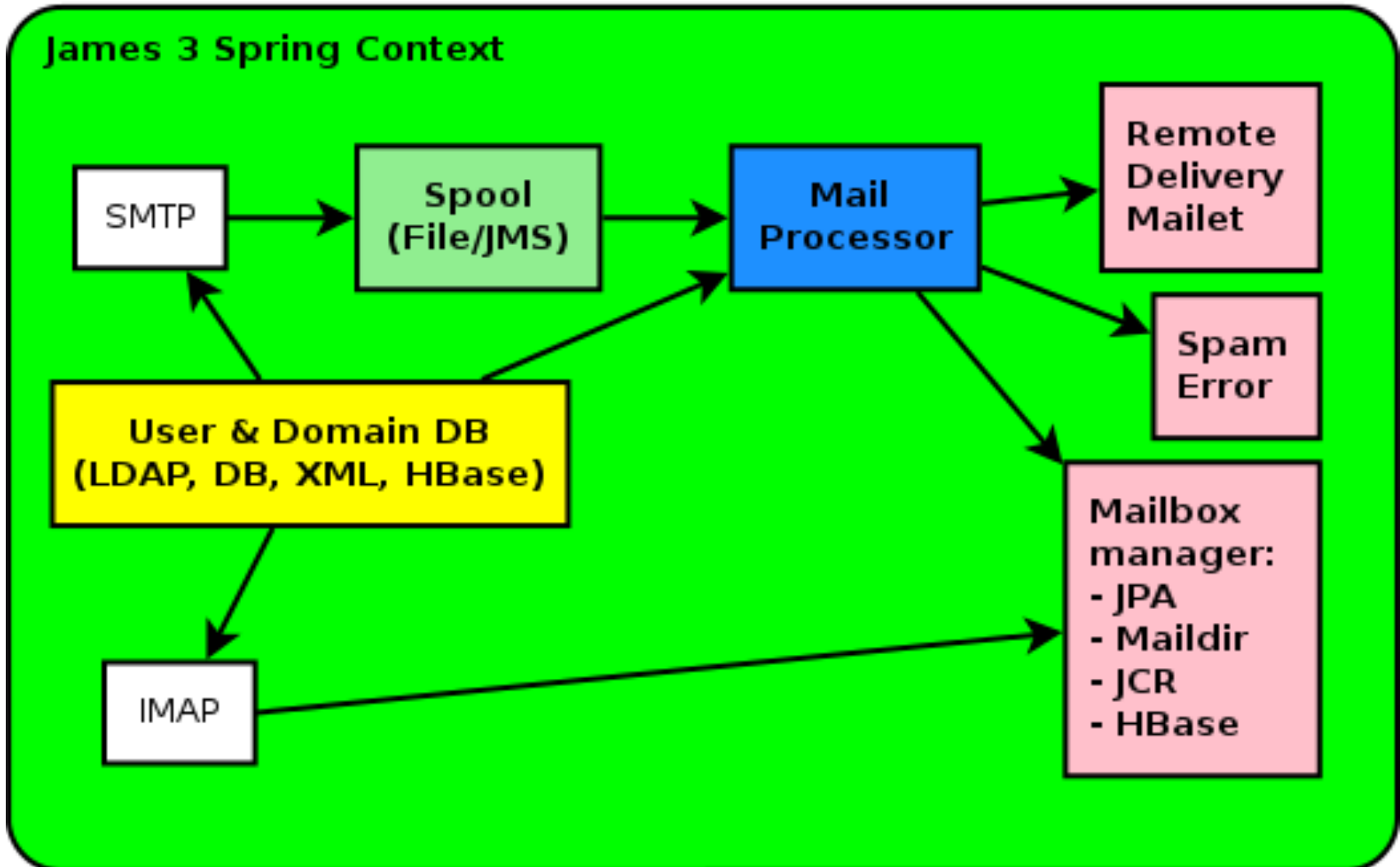
Apache James Server

- The Apache James Project's "Flag Ship"
- Uses all other James artifacts and more
- Spring based
- Available in binary forms:
 - WAR
 - ZIP standalone app
 - Debian *fat* package (in progress)
- works out of the box with Apache Derby as mailbox and user/domain store
- unpack, basic config, start sending email (5-10 min set-up)

James 3 architecture



James 3 Simplified Architecture





How to scale James

Scale storage:

- JPA mailbox with sharding ?!?
- JCR mailbox with clustering ?!?
- use HBase mailbox implementation

Scale processing:

- use [GreyMatter](#) - Akka~Maillet integration
- use multiple James instances*
- combine the above approaches

* requires distributed mailbox locking and UID generation - Zookeeper based implementation progress

Anatomy of Email - RFC 5322



From: John Doe <jdoe@machine.example>
To: Mary Smith <mary@example.net>
Subject: Saying Hello
Date: Fri, 21 Nov 1997 09:55:06 -0600
Message-ID: <1234@local.machine.example>

This is a message just to say hello.
So, "Hello".

Find 5 similarities



Ioan Eugen Stan
@ieugen222
<http://ieugen.blogspot.com/>

Edit your profile

18 TWEETS
40 FOLLOWING
10 FOLLOWERS

My subscriptions

People subscribed to my sent-mailbox

Sent messages

Machine learning on Mailboxes

Tweets

Following

Followers

Favorites

Lists

Similar to you

Alex Parvulescu @alexparvulescu
Following

Nearyoo Test @nytest
Follow

Stephan Huez @stephanhuez
Follow

Ioan Eugen Stan @ieugen222
at #bbuzz keynote - time, concurrency, star navigation and Google
Expand

Ioan Eugen Stan @ieugen222
@normanmaurer they are recorded so they will probably be. Update as soon as I know details.
View conversation

Ioan Eugen Stan @ieugen222
@lazershark please visit axemblr.com
Expand

Ioan Eugen Stan @ieugen222
@alexsisu I'm at bbuzz, we should meet. what are you attending?
View conversation

Emilian Losneanu @losneanu
#coderetreat @cegeka with @mihaisoloi, @drcomaneci @edimoraru @ieugen222 facilitated by @adibolb and @talboomerik
Retweeted by Ioan Eugen Stan
Expand

© 2012 Twitter, Inc. All rights reserved.
Blog Status Apps Security Jobs
Advertisers Businesses Media Developers

Final notes



Respect the firefighters!

Final slide



Thank you !
Questions/Comments?
Applause !?

Contact: ieugen@apache.org