Scalability Patterns & Solutions for Dynamic high-load Java Websites

Beurs van Berlage, Damrak 243, Amsterdam, 20/06/2014
Ard Schrijvers, a.schrijvers@onehippo.com, ard@apache.org

HIPPO

goto; conference


goto;
conference
Please evaluate this talk via the mobile app!
INTERNATIONAL
SOFTWARE DEVELOPMENT
CONFERENCE
Follow us @gotoamst
HIPPO
goto;
conference

What Hippo does / sells

Traditionally Hippo used to sell a CMS capable of **managing content** and a customer specific site implementation.

Hippo strictly separates the **editing** process from the **presentation** logic.

Content is stored in a **generic format**, allowing it to be **reused** across multiple pages and/or channels.

No longer just a CMS

No longer are we a CMS that is just about putting content or web pages at the conceptual center.
Today our real strength is the fact that we have the **Visitor** as the focus, and on a technical level, our delivery tier that interacts with that visitor to serve out relevant pages by really listening to the visitor.

Implications

1. Every page is rendered live from the application taking the visitor into account
2. Serving html from a reverse **caching** proxy (squid/varnish/mod_cache) is not an option

Note that offloading css, js, images, etc to reverse caching proxies or some CDN is still our common practice

Requirements for Hippo's delivery tier framework

1. support many concurrent visitors
2. instantly reflect frequently changing content
3. runtime adding sites and/or changing URL's of existing sites
4. runtime changing the appearance of sites
5. search including authorization
6. faceted navigation requiring authorized counts
7. personalization of pages
8. storing of visitor data

Amazon EC2 performance test results

Serving **personalized** pages and **storing** all request data and accumulated visitor characteristics, a single Hippo cluster node already saturated the available Amazon bandwidth

8.6 million pages per hour

2389 pages per second

6398400000 (6 billion) pages per month

A brief history

I am working at Hippo since 2001

Lead developer Hippo's delivery tier (framework)

Apache committer of Jackrabbit and Cocoon

Biggest mistake

Back in 2001, XML / XSLT was buzzing and bleeding edge

We needed a time tracking system at Hippo

.... so I built one by storing **one** XML in **one** access db **blob** and a XSLT to transform it into a time tracking system...with ASP.

Around 2003 we started using Cocoon

Cocoon: XML and XSLT publishing Open Source Java framework built around the concepts of **separation of concerns**

CMS and delivery tier built in Cocoon

Slide (XML Content Repository) accessed over WebDAV

Lessons learned

Apache and community!

Separation of concerns : Content and presentation

Request matching and the reverse: Link rewriting references between content to URLs.

Cocoon / XSLT was (and is) too slow

Lessons learned

Reverse caching proxies (mod_cache, squid, varnish, ssi tricks)

Indexing content with Apache Lucene (around 2003 that was version 1.2)

Many caching strategies and their problems / difficulties (for developers)

Cache invalidation mechanisms (JMS eventing)

Lessons learned

Authorization and fast search results hard to combine

Using remote repositories is too slow if you require many sources

Around 2005 integrated Apache Jetspeed

Apache Jetspeed: Open Source Enterprise Portal framework and platform

- ★ native integration of the CMS
- ★ portal used as delivery tier
- ★ combining portlets, content and 3rd party services in one solution

Hippo Portal

Lessons learned

Multi webapp state sharing is complex

Multi webapp orchestration of services

Writing cross webapp shared APIs

HMVC pattern for the delivery tier

2007 start Hippo CMS 7

CMS: Stateful AJAX based webapp written in Wicket

Delivery tier framework (HST) written from scratch

Hippo Repository: a JCR compliant repository on top of Apache Jackrabbit

Some CMS 7 Customers

Ministry of Foreign Affairs

Crossing Borders Helping Abroad

- Integrated Translations
- 100+ Domains
- 300+ Topical Microsites
- Travel Safety Mobile App


WELEDA
"We chose Hippo for one reason:
Targeting content to consumers is paramount."
Laurent Vonach, Online Manager International Marketing, Weleda

Dutch police : From 400 web sites to 1

"With Hippo, we rolled out the mobile site together with the desktop site. That's the advantage of having a central Content Management System that serve content to all channels."

http://www.cmscritic.com/how-open-source-software-transformed-a-nations-police-force/

Multilingual content delivery in twelve languages

Real time updates to design models

Dynamic content from Amazon Cloud to millions of global customers

Content updated and distributed via 28 AppStores managed by Hippo CMS

"By managing our content with Hippo in the Amazon Cloud, we now roll out content updates to all our customers worldwide in just a few minutes. The solution gave us the agility we needed to transition our business to a subscription model."

- Targeting based on entry page, geo location and click behavior
- Enterprise Forms with ‘killer questions’ to pre-qualify applications
- Responsive Design
- Multibrand: TempoTeam and Randstad.nl

http://www.ns.nl

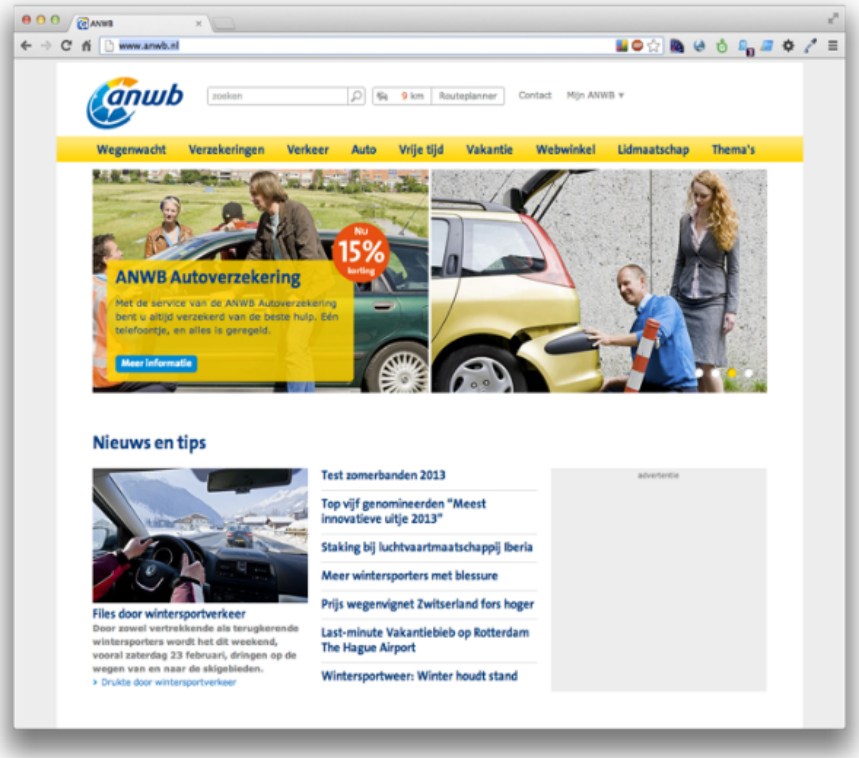
Local Collaboration Powers Innovation

20 mio unique visitors / month
10,000 employees

- Web, mobile, tablet & kiosk
- Content as a Service (REST)
- Transactional System integration

Dutch railway

-Web, mobile, tablet & kiosk
-Integration with back office & e-Commerce
-Integration with Train - information management System
-On-board live sync system



- Centralized Content for a Decentralized Organization
- 200 forms and 68 applications
- MyANWB portal
- Content reuse in 16 mobile apps and 7 publications
- 120 content editors


GULP
DERMICIS
Ziggo
TALMER
BANK AND TRUST
innovam
GROEP
Rijksoverheid
Ministry of Foreign Affairs
randstad
STUDIECENTRUM RECHTSPLEGING
TRAINING AND STUDY CENTRE FOR THE JUDICIARY
DUO
bugaboo
PC Hooft Groep
VERZEKERAAR EN VERZORGER VAN UITVAARTEN
CVZ
College voor zorgverzekeringen
Nederlands Forensisch Instituut
Ministerie van Veiligheid en Justitie
BROCACEF
Zorglogistiek
WELEDA
research now
1&1
Oogziekenhuis
Zonnestraal
VEIKKAUS
Universiteit Leiden
Couchbase
HetWaterschapshuis
tempo-team
YACHT
a Randstad company
Ministerie van Algemene Zaken
avans
hogeschool
NetwerkVSP
VISIT
BRUSSELS
BellAliant
Hogeschool van Amsterdam
acm
Association for
Computing Machinery
Advancing Computing as a Science & Profession
MOBILE
Come In
WE'RE
OPEN
HIPPO
goto;
conference

What all customers have in common

Most have high volume sites

They all use Hippo **differently** to deliver
(personalized) content
to different channels

Hippo's business model

GOLD

CGI

GOLD

CGI

GOLD

hinttech dayon

GOLD

Hinttech-Dayon

GOLD

THE FACTOR E

GOLD

theFactor.e

GOLD

FINALIST open IT oplossingen

GOLD

Finalist

GOLD

NORTHPOINT

GOLD

Northpoint

GOLD

Capgemini CONSULTING.TECHNOLOGY.OUTSOURCING

GOLD

Capgemini

SILVER

oshyn POWERING INNOVATION

SILVER

Oshyn

SILVER

Satzmedia

SILVER

Satzmedia

SILVER

netpioneer

SILVER

Netpioneer

SILVER

thoughtmatrix

SILVER

Thought Matrix

SILVER

AUTHX

SILVER

AuthXConsulting

SILVER

youngsoft INC

SILVER

Youngsoft

SILVER

FINITY

SILVER

Affinity Systems

SILVER

edoras Multichannel Implementation Expert

SILVER

edoras

SILVER

iSDC

SILVER

ISDC

SILVER

kernpunkt optimiert

SILVER

Kernpunkt

HIPPO

goto; conference


SILVER
DigitasLBi
SILVER
DigitasLBi
SILVER
Design UK
SILVER
DesignUK
SILVER
CC group
SILVER
CCGroup
SILVER
nubizz ecommerce
SILVER
Nubizz
SILVER
BLUETRADE
Make Digital Impact
SILVER
Bluetrade
SILVER
Key Consulting Groupe
SILVER
Key Consulting Group
SILVER
ORDINA
SILVER
Ordina
SILVER
osudio
SILVER
Osudio
SILVER
KILLNINE
SILVER
Killnine
SILVER
TRIFORK.
SILVER
Trifork
SILVER
{xmachina
advancing e-business
SILVER
xmachina
SILVER
Indivirtual
SILVER
Indivirtual
SILVER
comSysto
sense and respond
SILVER
comSysto
SILVER
Smile
OPEN SOURCE SOLUTIONS
SILVER
Smile
SILVER
ESISS
SILVER
ESISS
SILVER
INCENTRO
INFORMATION BECOMES INSPIRATION
SILVER
Incentro
MOBILE
Come In
WE'RE
OPEN
HIPPO
goto;
conference

Open Source stack:

Hippo's stack

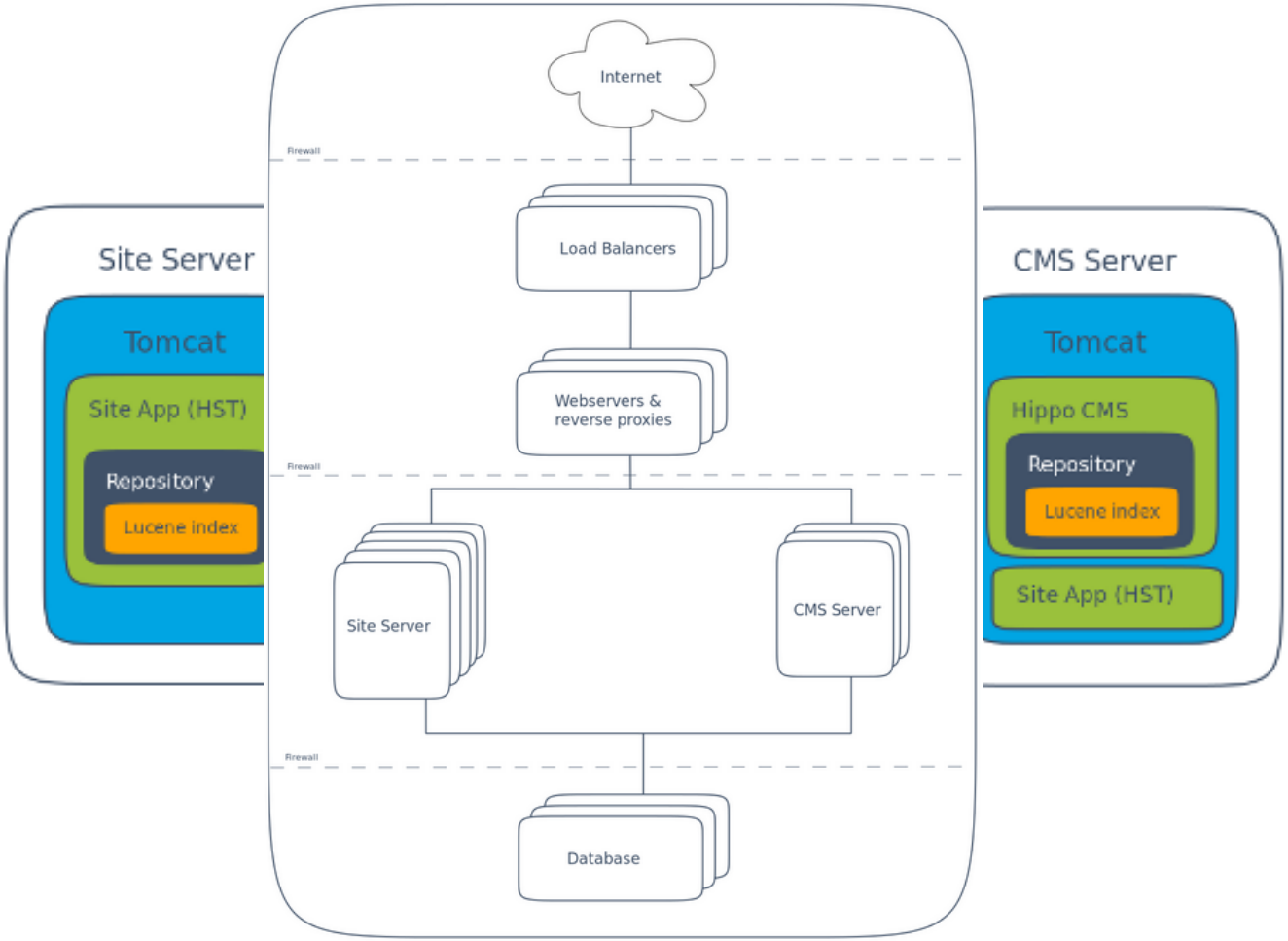
Apache License Version 2.0

except some enterprise modules on the **periphery** of our stack

Used Open Source licenses

Apache License Version 2.0
Day Specification License (JCR)
Python-2.0
BSD-2 / BSD-3
MIT / X11
EDL 1.0
EPL 1.0
MPL 1.1 / 2.0
W3C Software License
GPLv3 under Sensha OS Exception for Application/Development (ExtJS)
Indiana University Extreme! Lab Software License Version 1.1
CDDL 1.0 / 1.1
CPL 1.0
CC-A 2.5/3.0
CC-BY 2.5
ICU
SIL OFL 1.1
Public Domain
WTFPL 2.0

10,000 foot view Hippo CMS 7



Hippo Repository on top of Jackrabbit

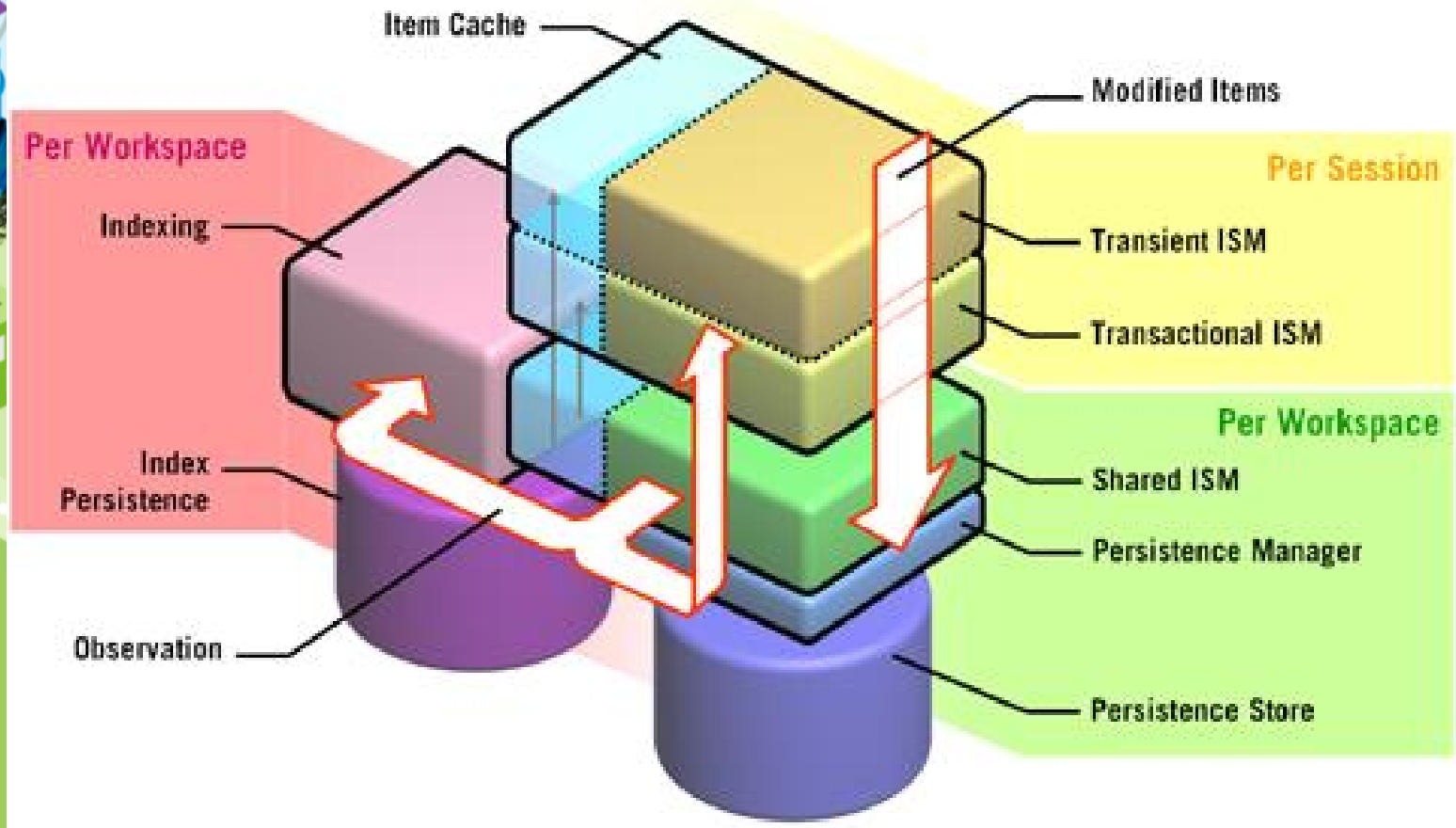
Jackrabbit is a reference implementation of Java Content Repository (JSR-170/JSR-283)

A content repository is a hierarchical content store with support for structured and unstructured content, full text search, versioning, transactions, observation, and more.

JCR in a nutshell

```
public interface Node {
      Node getNode(String relPath);
      Node addNode(String relPath);
      Property getProperty(String name)
      Property setProperty(String name,
                           Value value);
}
```

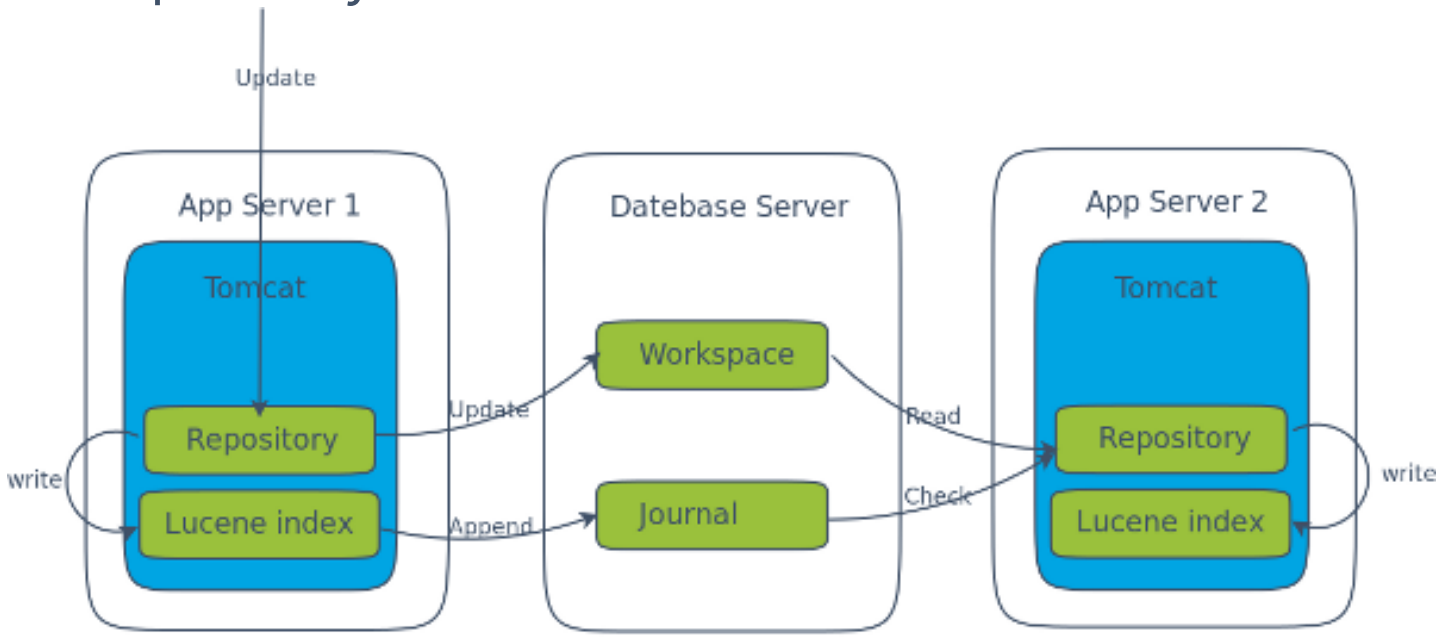
Jackrabbit architecture



Source: http://jackrabbit.apache.org/how-jackrabbit-works.html

Jackrabbit clustering

Always have a repository embedded in the containers for the webapps that require a repository and do not use remote



How to query the repository

1. A subset of XPath (JSR-170)
2. A subset of SQL (JSR-170)
3. JCR-SQL2 (JSR-283)
4. JCR-JQOM (JSR-283)

Complex XPath query

```
/jcr:root/nodes//element(*,my:type)
        [jcr:contains(.,'jsr') and
         my:subnode/@jcr:primaryType='my:html']
         /my:body[jcr:contains(.,'170')]
```

Jackrabbit (Lucene) index

Challenges:

1. Hierarchical queries cannot be mapped easily to Lucene
2. After Session#save() instant reflection of search results required (real-time search) but at the time of JSR-170 Lucene was at version 1.4.
3. Lucene indexes always need to be local: You cannot bring the data to the computation!!
4. Search results should return only **authorized** hits

Jackrabbit (Lucene) index

Challenge 1:

Hierarchical queries cannot be mapped easily to Lucene

Solution 1:

Just try to avoid them even though Adobe (Day) developers did an amazing job

Jackrabbit (Lucene) index

Challenge 2:

After Session#save() instant reflection of search results required (real-time search)

Solution 2:

A set of Lucene indexes instead of a single one. Again Adobe (Day) developers did an amazing job...with Lucene 1.4!!

Jackrabbit (Lucene) index

Challenge 3:

Lucene indexes always need to be local: You cannot bring the data to the computation!!

Solution 3:

Every Jackrabbit cluster node has a local Lucene (multi-) index.

Jackrabbit (Lucene) index

Challenge 4:

Search results should return only **authorized** hits

Solution 4:

Hippo chose for an authorization model on top of JCR that could be mapped to Lucene queries and could be **AND-ed** with every normal query

Example Authorization Query

```
(+_:HIPPO_PT_FACET:13109076:templatetype) (+_:HIPPO_PT_FACET:13109076:namespace)
(+_:HIPPO_PT_FACET:13109076:namespacefolder) (+_:HIPPO_PT_FACET:13109076:field)
(+_:HIPPO_PT_FACET:13109076:nodetype) (+_:HIPPO_PT_FACET:7275975:templatequery)
(+_:HIPPO_PT_FACET:14608509:templateset) (+_:HIPPO_PT_FACET:13109076:prototypeset)
(+HIPPOSORTABLE::hipposysedit:prototype) (+_:HIPPO_PT_FACET:14697776:facetresult)
(+_:HIPPO_PT_FACET:16174620:derivedefinition) (+(_:HIPPO_PT_FACET:16174620:propertyreference
_:HIPPO_PT_FACET:16174620:builtinpropertyreference _:HIPPO_PT_FACET:16174620:
relativepropertyreference
_:HIPPO_PT_FACET:16174620:resolvepropertyreference)) (+_:HIPPO_PT_FACET:16174620:
securityfolder)
(+_:HIPPO_PT_FACET:14697776:handle) (+_:HIPPO_PT_FACET:16174620:applicationfolder)
(+HIPPOSORTABLE::liveuser +(_:HIPPO_PT_FACET:16174620:user _:HIPPO_PT_FACET:16174620:
externaluser))
(+_:HIPPO_PT_FACET:14697776:facetselect) (+_:HIPPO_PT_FACET:16174620:queryfolder)
(+_:HIPPO_PT_FACET:16174620:configuration) (+_:HIPPO_PT_FACET:14219914:report)
(+_:HIPPO_PT_FACET:16174620:propertyreferences) (+_:HIPPO_PT_FACET:16762557:root)
(+_:HIPPO_PT_FACET:7275975:translations) (+7275975:HIPPOFACET:holder:liveuser)
(+_:HIPPO_PT_FACET:16174620:facetsubsearch) (+_:HIPPO_PT_FACET:16174620:userfolder)
(+_:HIPPO_PT_FACET:14697776:translation) (+_:HIPPO_PT_FACET:7275975:templates)
(+_:HIPPO_PT_FACET:14697776:facetsearch) (+_:HIPPO_PT_FACET:5688619:unstructured)
(+_:HIPPO_PT_FACET:16174620:derivativesfolder) (+(+MatchAllDocsQuery -HIPPOSORTABLE::
hipposysedit:prototype)
+((+MatchAllDocsQuery -_:FACET_PROPERTIES_SET:14697776:availability)
14697776:HIPPOFACET:availability:live) +(_:HIPPO_PT_FACET:14697776:document
_:HIPPO_PT_FACET:14093235:config _:HIPPO_PT_FACET:9867704:exampleAssetSet
_:HIPPO_PT_FACET:9867704:exampleImageSet _:HIPPO_PT_FACET:9867704:imageset
_:HIPPO_PT_FACET:9867704:stdAssetGallery _:HIPPO_PT_FACET:9867704:stdImageGallery
_:HIPPO_PT_FACET:9867704:stdgalleryset _:HIPPO_PT_FACET:7275975:directory
_:HIPPO_PT_FACET:7275975:document _:HIPPO_PT_FACET:7275975:folder
_:HIPPO_PT_FACET:7275975:gallery _:HIPPO_PT_FACET:7275975:space
_:HIPPO_PT_FACET:13109076:nodetype _:HIPPO_PT_FACET:14219914:report
_:HIPPO_PT_FACET:11431386:basedocument _:HIPPO_PT_FACET:11431386:newsdocument
_:HIPPO_PT_FACET:11431386:textdocument)) (+_:HIPPO_PT_FACET:5688619:versionLabels)
(+_:HIPPO_PT_FACET:5688619:version) (+_:HIPPO_PT_FACET:5688619:versionHistory)
(+_:HIPPO_PT_FACET:16762557:system) (+_:HIPPO_PT_FACET:5688619:frozenNode)
```

Example Authorization Query Continued

```
(+_:HIPPO_PT_FACET:5688619:versionedChild) (+_:HIPPO_PT_FACET:16762557:versionStorage)
(+_:HIPPO_PT_FACET:12208518:item) (+_:HIPPO_PT_FACET:12208518:folder)
(+_:HIPPO_PT_FACET:1000430:allowedSingleWhitespaceElement) (+_:HIPPO_PT_FACET:1000430:
cleanupElement)
(+_:HIPPO_PT_FACET:1000430:cleanup) (+_:HIPPO_PT_FACET:1000430:serializationElement)
(+_:HIPPO_PT_FACET:1000430:serialization) (+_:HIPPO_PT_FACET:1000430:config)
(+_:HIPPO_PT_FACET:16174620:modulefolder) (+_:HIPPO_PT_FACET:16174620:module)
(+_:HIPPO_PT_FACET:7776938:workflow) (+_:HIPPO_PT_FACET:1717184:request)
(+_:HIPPO_PT_FACET:11744324:triggers) (+_:HIPPO_PT_FACET:11744324:trigger) (+_:HIPPO_PT_FACET:
16174620:type)
(+_:HIPPO_PT_FACET:16174620:workflow) (+_:HIPPO_PT_FACET:16174620:ocmqueryfolder)
(+_:HIPPO_PT_FACET:16174620:workflowcategory) (+_:HIPPO_PT_FACET:14697776:request)
(+_:HIPPO_PT_FACET:16174620:workflowfolder) (+_:HIPPO_PT_FACET:16174620:types)
(+_:HIPPO_PT_FACET:14697776:query) (+_:HIPPO_PT_FACET:7776938:clusterfolder)
(+_:HIPPO_PT_FACET:7776938:application) (+((+MatchAllDocsQuery -_:FACET_PROPERTIES_SET:0:
cluster.name)
(+MatchAllDocsQuery -0:HIPPOFACET:cluster.name:hst-editor))
+_:HIPPO_PT_FACET:7776938:plugin +(+MatchAllDocsQuery
-0:HIPPOFACET:plugin.class:org.hippoecm.frontend.plugins.reviewedactions.
PublishAllShortcutPlugin)
+((+MatchAllDocsQuery -_:FACET_PROPERTIES_SET:0:cluster.name)
(+MatchAllDocsQuery -0:HIPPOFACET:cluster.name:cms-dev)) +(+MatchAllDocsQuery
-0:HIPPOFACET:plugin.class:org.hippoecm.frontend.plugins.cms.admin.AdminPerspective)
+((+MatchAllDocsQuery -_:FACET_PROPERTIES_SET:0:cluster.name)
(+MatchAllDocsQuery -0:HIPPOFACET:cluster.name:cms-tree-views/configuration)))
(+_:HIPPO_PT_FACET:7776938:pluginCluster) (+_:HIPPO_PT_FACET:7776938:pluginconfig)
```

Can such a to-be-AND-ed query perform?

Results of the Authorization Query

Also users with little read access have instant authorized searches

Correct total hit size from Lucene

Correct instant faceted navigation authorized counts

Requirements for Hippo's delivery tier framework

1. support many concurrent visitors
2. instantly reflect frequently changing content
3. runtime adding sites and/or changing URL's of existing sites
4. runtime changing the appearance of sites
5. search including authorization
6. faceted navigation requiring authorized counts
7. personalization of pages
8. storing of visitor data

Requirements for Hippo's delivery tier framework

1. support many concurrent visitors
2. **instantly reflect frequently changing content**
3. runtime adding sites and/or changing URL's of existing sites
4. runtime changing the appearance of sites
5. search including authorization
6. faceted navigation requiring authorized counts
7. personalization of pages
8. storing of visitor data

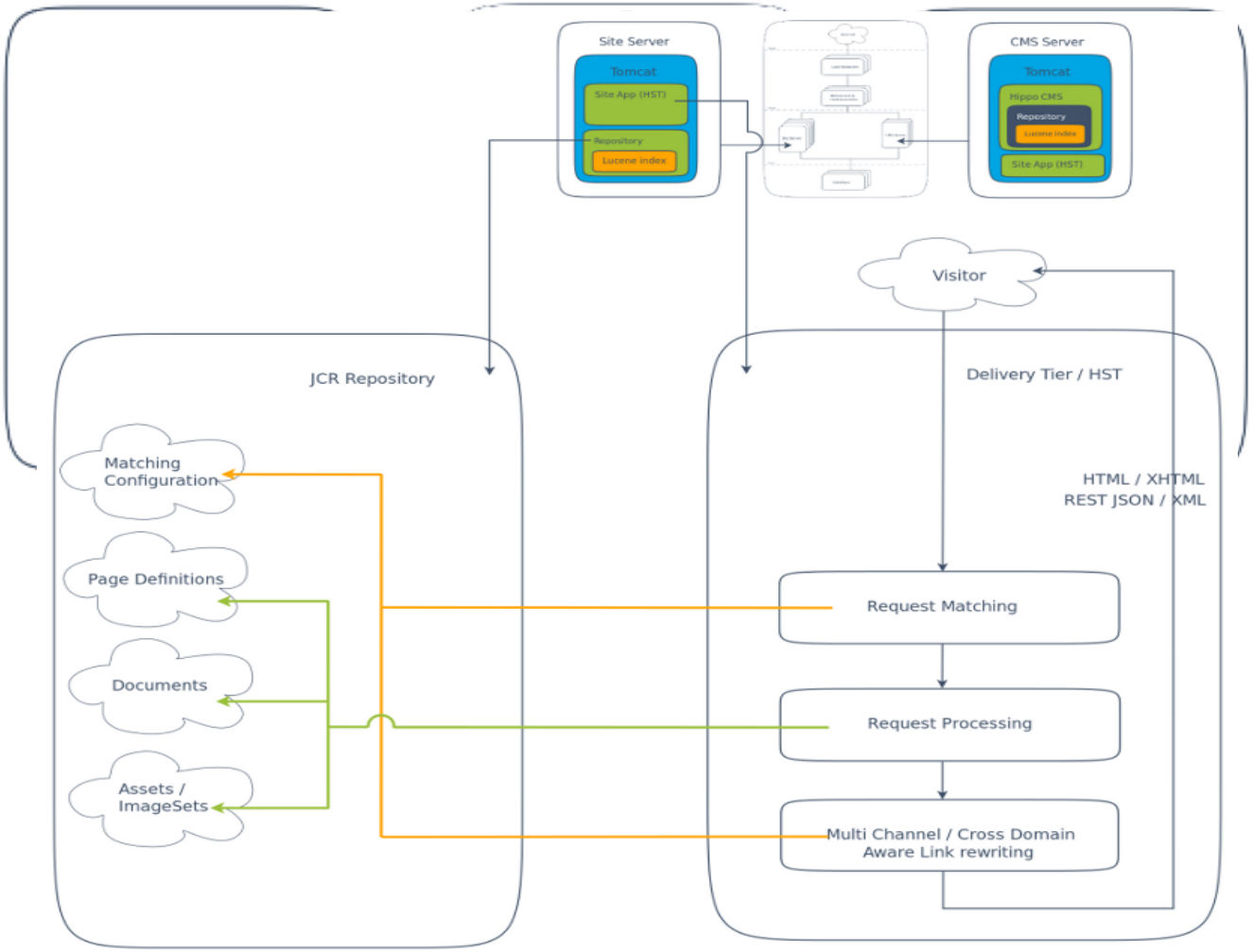
Requirements for Hippo's delivery tier framework

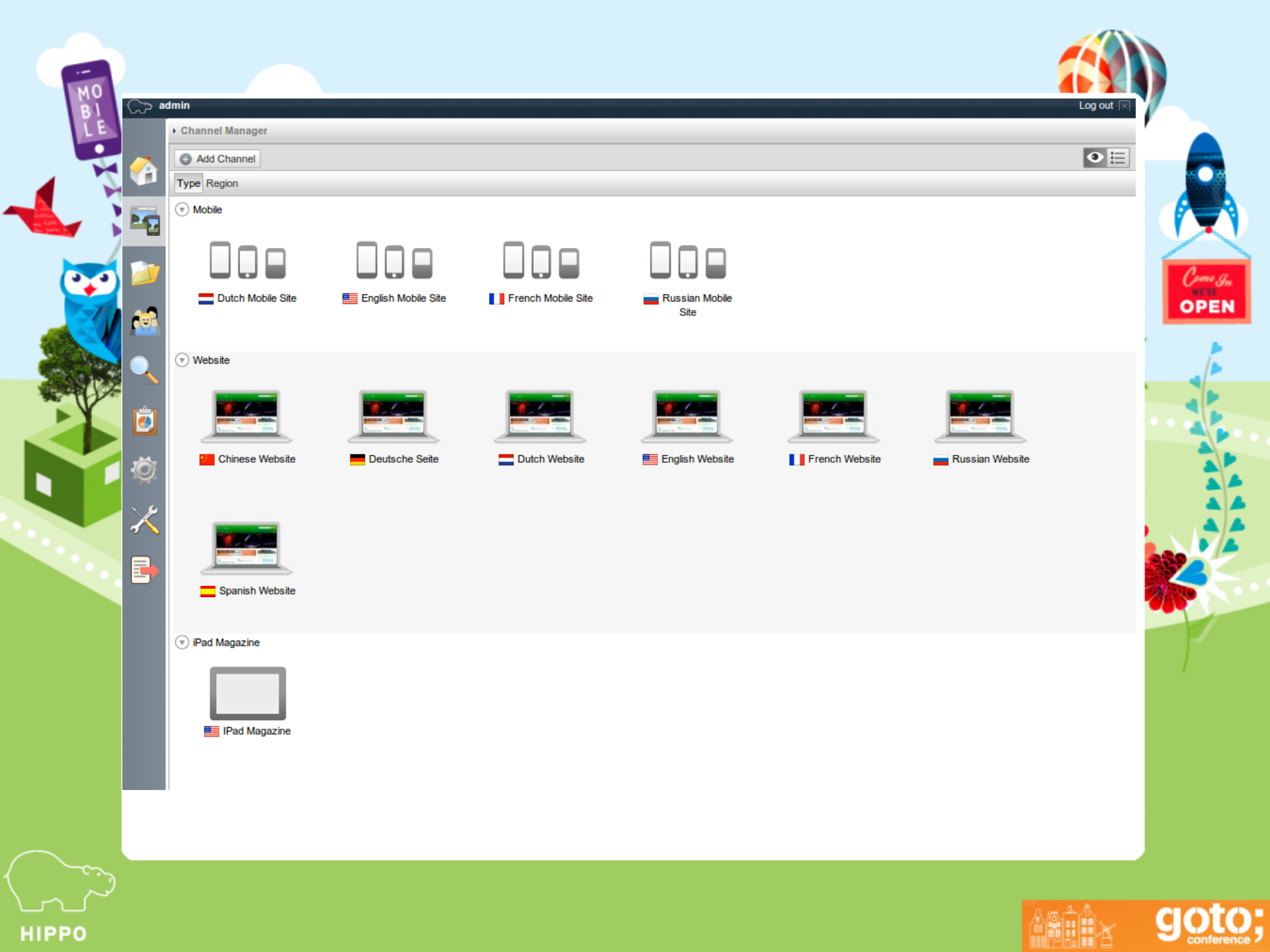
1. support many concurrent visitors
2. **instantly reflect frequently changing content**
3. runtime adding sites and/or changing URL's of existing sites
4. runtime changing the appearance of sites
5. **search including authorization**
6. **faceted navigation requiring authorized counts**
7. personalization of pages
8. storing of visitor data

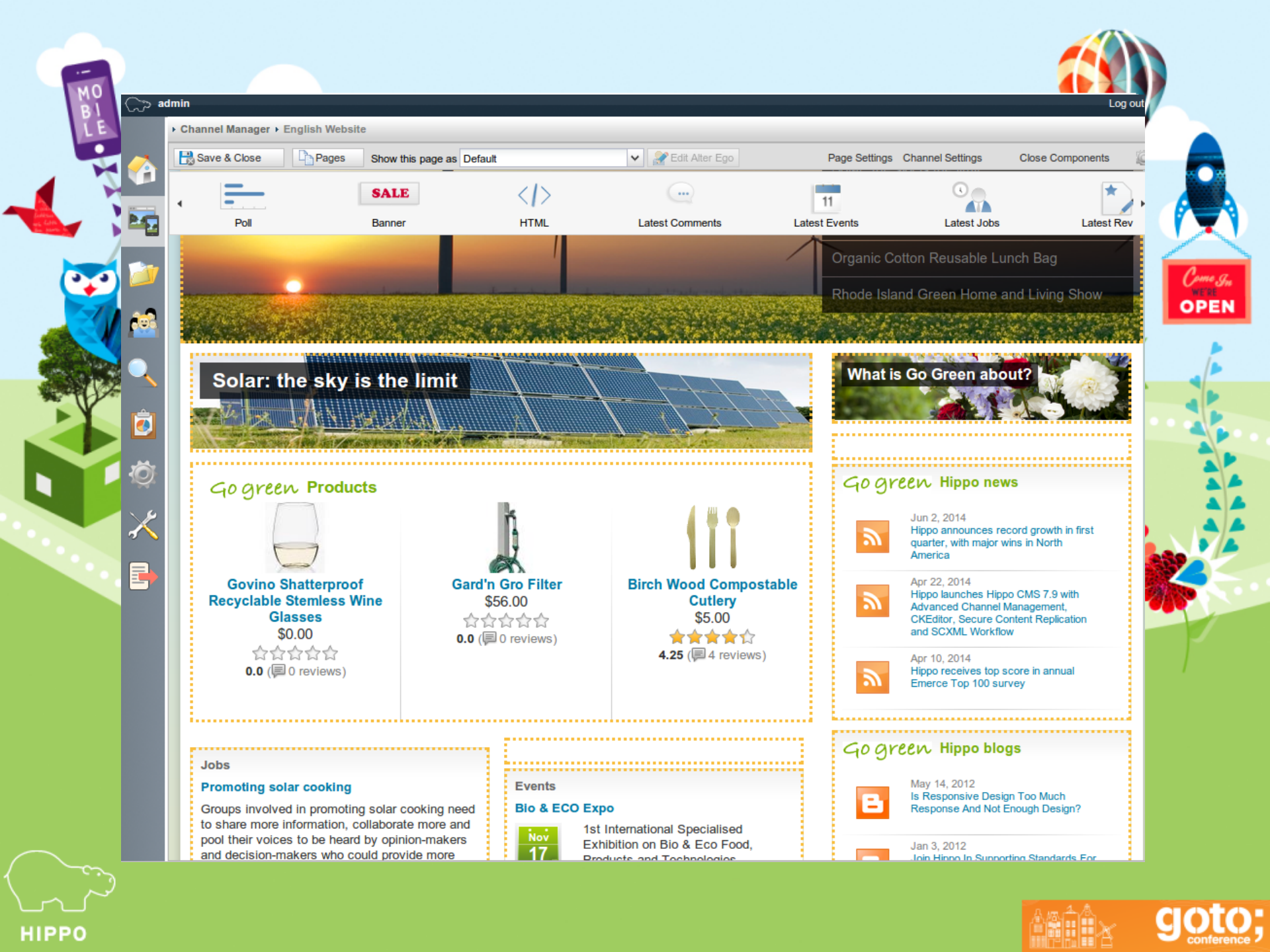
Hippo's delivery tier in a nutshell

1. Open Source (Apache License Version 2.0)
2. Acronym: HST
3. It's not a toolkit but a framework
4. Pluggable container which is using Spring Framework configurations
5. Its main phases can be divided in
   a. A matching & link rewriting phase
   b. A processing phase (default a HMVC pattern)
6. The configuration for (5) is stored in the repository and runtime modifiable
7. The HST keeps an in memory model for (6)
8. It's primarily content driven, not page driven: Hippo CMS manages content & page definitions, not pages.

Hippo's delivery tier




admin
Log out
Channel Manager
Add Channel
Type
Region
Mobile
Dutch Mobile Site
English Mobile Site
French Mobile Site
Russian Mobile Site
Website
Chinese Website
Deutsche Seite
Dutch Website
English Website
French Website
Russian Website
Spanish Website
iPad Magazine
IPad Magazine
MOBILE
Come In WE'RE OPEN
HIPPO
goto; conference


admin
Log out
Channel Manager ▸ English Website
Save & Close
Pages
Show this page as
Default
Edit Alter Ego
Page Settings
Channel Settings
Close Components
Poll
SALE
Banner
HTML
Latest Comments
11
Latest Events
Latest Jobs
Latest Rev
Organic Cotton Reusable Lunch Bag
Rhode Island Green Home and Living Show
Solar: the sky is the limit
What is Go Green about?
Go green Products
Govino Shatterproof Recyclable Stemless Wine Glasses
$0.00
0.0 (0 reviews)
Gard'n Gro Filter
$56.00
0.0 (0 reviews)
Birch Wood Compostable Cutlery
$5.00
4.25 (4 reviews)
Go green Hippo news
Jun 2, 2014
Hippo announces record growth in first quarter, with major wins in North America
Apr 22, 2014
Hippo launches Hippo CMS 7.9 with Advanced Channel Management, CKEditor, Secure Content Replication and SCXML Workflow
Apr 10, 2014
Hippo receives top score in annual Emerce Top 100 survey
Jobs
Promoting solar cooking
Groups involved in promoting solar cooking need to share more information, collaborate more and pool their voices to be heard by opinion-makers and decision-makers who could provide more
Events
Bio & ECO Expo
Nov
17
1st International Specialised Exhibition on Bio & Eco Food,
Go green Hippo blogs
May 14, 2012
Is Responsive Design Too Much Response And Not Enough Design?
Jan 3, 2012
Come In WE'RE OPEN
HIPPO
goto; conference

Challenge

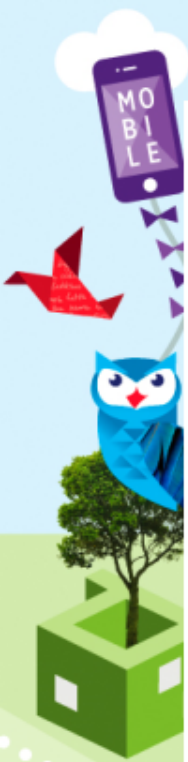
Having many concurrent visitors while runtime adding sites and/or changing URL's of existing sites and changing the appearance (requiring model reloads) while supporting 500+ channels including cross domain (site) link rewriting

General pattern to get around this

Use a lazy append-only (immutable) in memory model tied to a request combined with request bound flyweights and be stateless (by default)

Immutability : Vertical scaling
Stateless : Horizontal scaling

CQRS (Command Query Responsibility Segregation) pattern to write changes to the model without requiring the query (read) model



Requirements for Hippo's delivery tier framework

1. support many concurrent visitors
2. **instantly reflect frequently changing content**
3. runtime adding sites and/or changing URL's of existing sites
4. runtime changing the appearance of sites
5. **search including authorization**
6. **faceted navigation requiring authorized counts**
7. personalization of pages
8. storing of visitor data

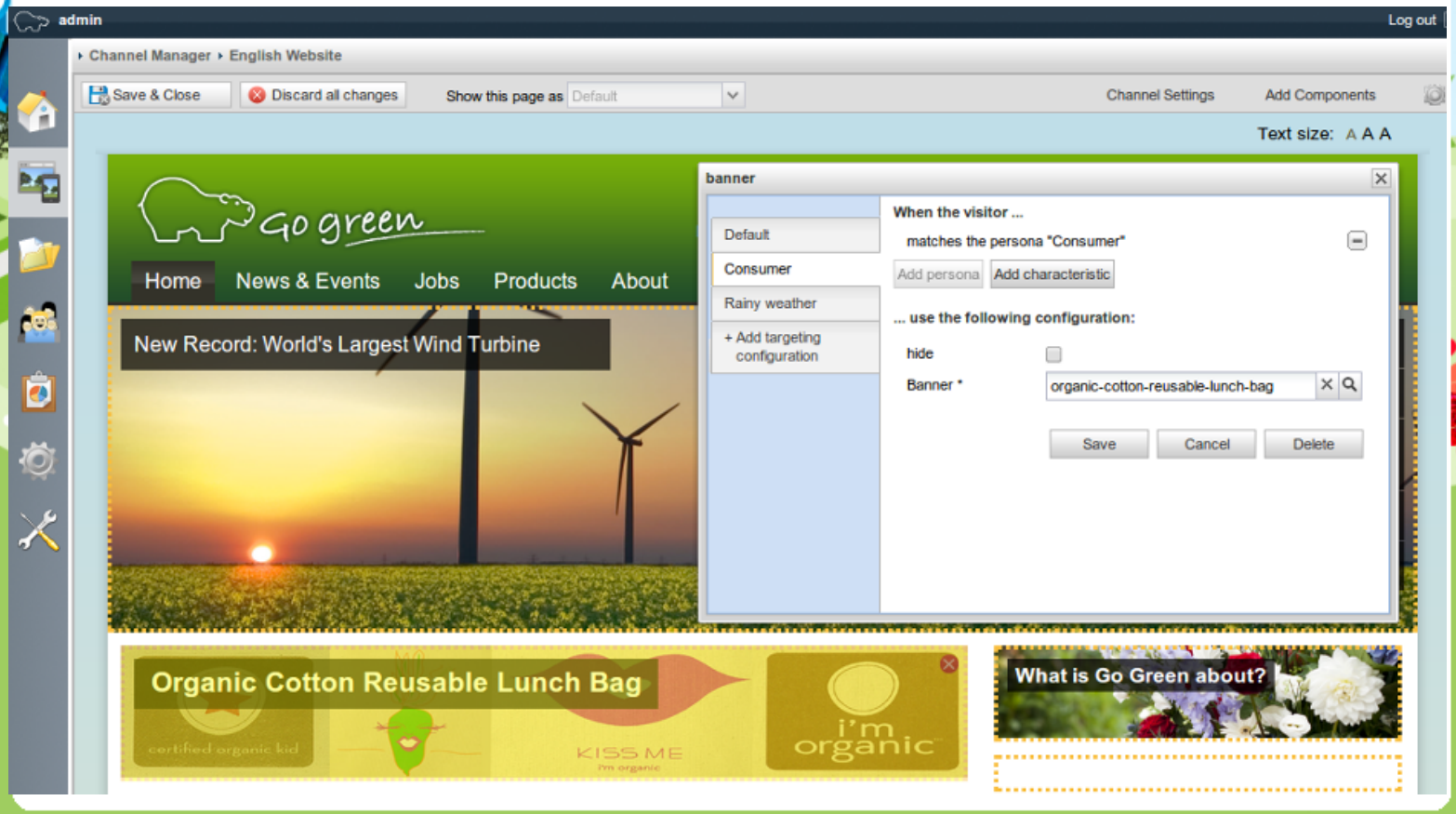
Requirements for Hippo's delivery tier framework

1. **support many concurrent visitors**
2. **instantly reflect frequently changing content**
3. runtime adding sites and/or changing URL's of existing sites
4. runtime changing the appearance of sites
5. **search including authorization**
6. **faceted navigation requiring authorized counts**
7. personalization of pages
8. storing of visitor data

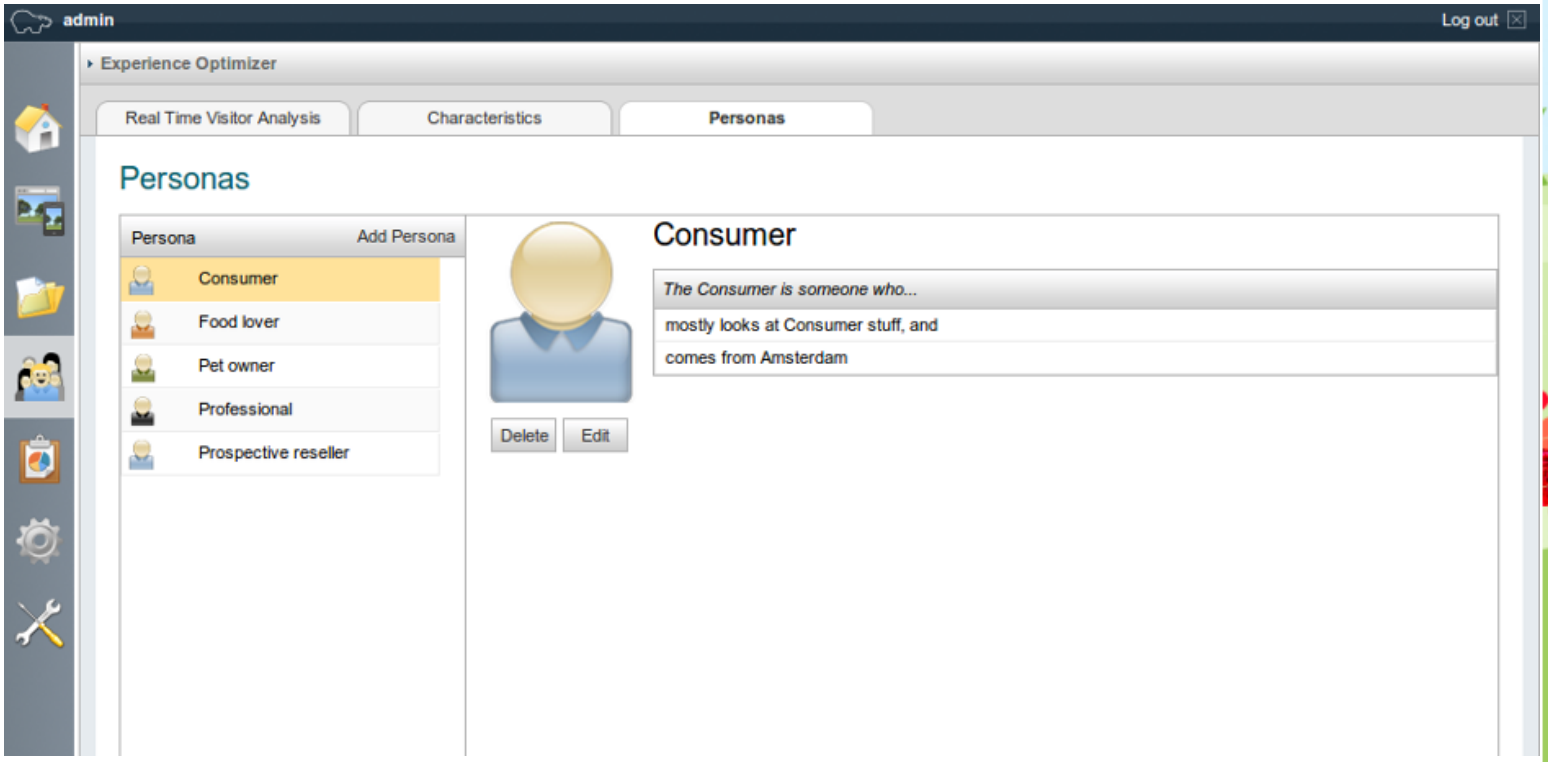
Requirements for Hippo's delivery tier framework

1. **support many concurrent visitors**
2. **instantly reflect frequently changing content**
3. **runtime adding sites and/or changing URL's of existing sites**
4. **runtime changing the appearance of sites**
5. **search including authorization**
6. **faceted navigation requiring authorized counts**
7. personalization of pages
8. storing of visitor data

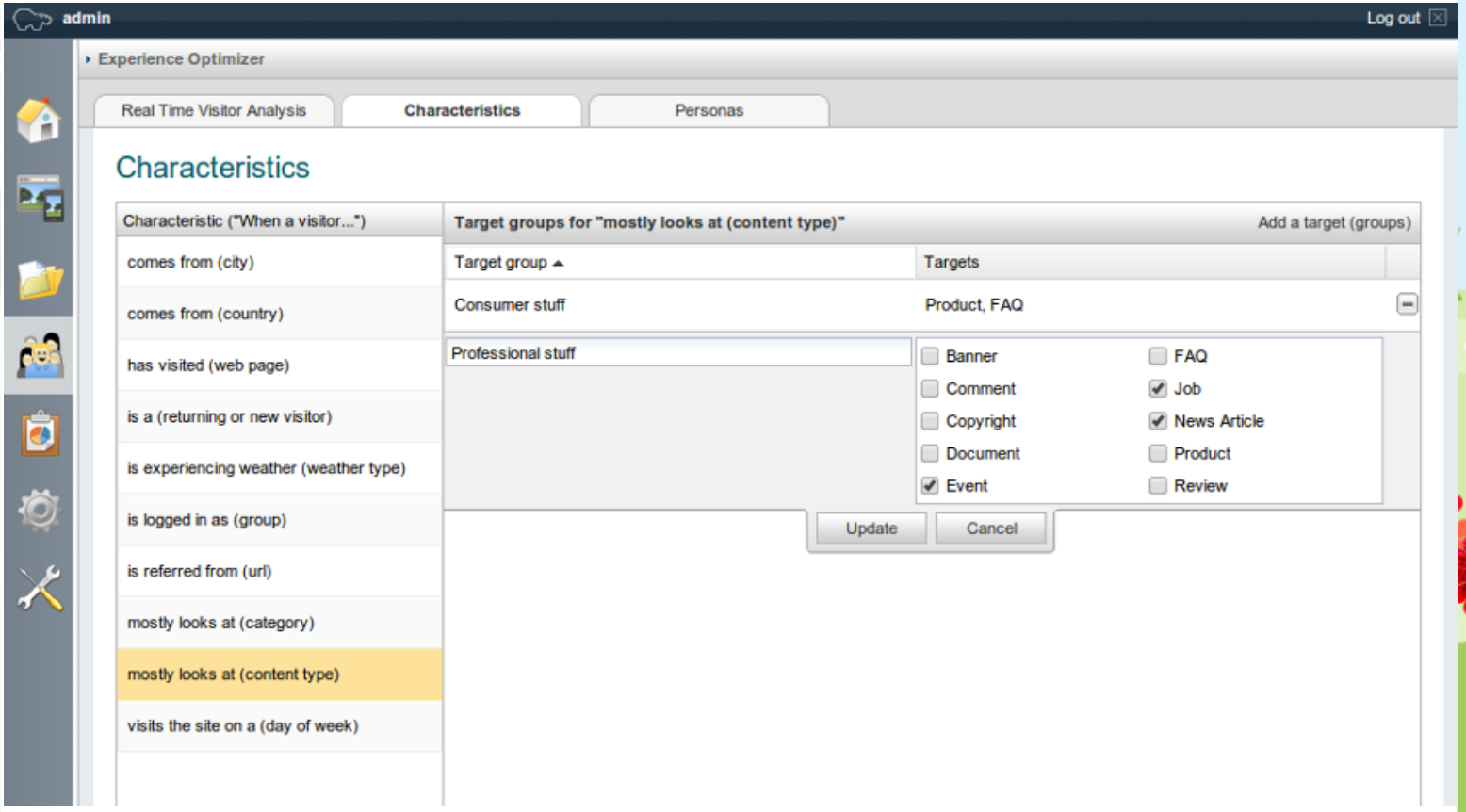
Next Challenge: Deliver different pages to different visitors

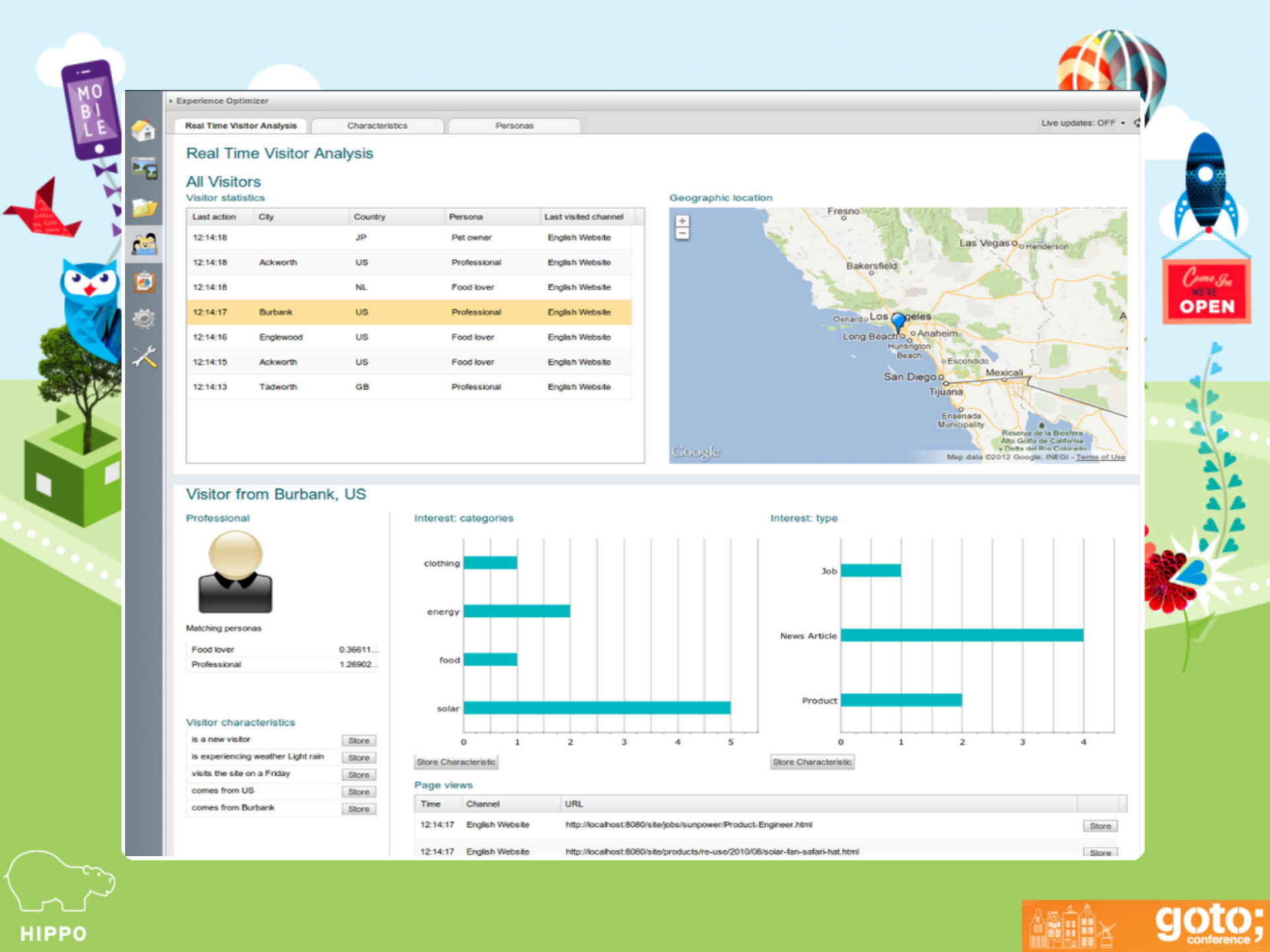


Persona Consumer example



Characteristics




Experience Optimizer
Real Time Visitor Analysis
Characteristics
Personas
Live updates: OFF
Real Time Visitor Analysis
All Visitors
Visitor statistics
Last action | City | Country | Persona | Last visited channel
12:14:18 | | JP | Pet owner | English Website
12:14:18 | Ackworth | US | Professional | English Website
12:14:18 | | NL | Food lover | English Website
12:14:17 | Burbank | US | Professional | English Website
12:14:16 | Englewood | US | Food lover | English Website
12:14:15 | Ackworth | US | Food lover | English Website
12:14:13 | Tadworth | GB | Professional | English Website
Geographic location
Fresno
Las Vegas
Henderson
Bakersfield
Oxnard
Los Angeles
Long Beach
Anaheim
Huntington Beach
Escondido
Mexicali
San Diego
Tijuana
Ensenada Municipality
Google
Map data ©2012 Google, INEGI - Terms of Use
Visitor from Burbank, US
Professional
Matching personas
Food lover 0.36611..
Professional 1.26902..
Visitor characteristics
is a new visitor Store
is experiencing weather Light rain Store
visits the site on a Friday Store
comes from US Store
comes from Burbank Store
Interest: categories
clothing
energy
food
solar
0 1 2 3 4 5
Store Characteristic
Interest: type
Job
News Article
Product
0 1 2 3 4
Store Characteristic
Page views
Time | Channel | URL
12:14:17 | English Website | http://localhost:8080/site/jobs/sunpower/Product-Engineer.html | Store
12:14:17 | English Website | http://localhost:8080/site/products/re-use/2010/08/solar-fan-safari-hat.html | Store
MOBILE
Come In WE'RE OPEN
HIPPO
goto; conference

Technical requirements

Having many concurrent visitors while

1. serving relevant (personalized) pages*
2. storing their request logs
3. storing their accumulated visitor data
4. computing visitor profiles
5. tracking cluster wide visitor statistics
6. staying stateless (by default)

* The relevance module is part of Hippo enterprise support

Statistics required to be able to support:

“facts that happen less frequently are more important when they happen”

For this we require cluster wide averages. More precisely, we use cluster wide *exponential moving averages*.

Storage solutions

1. Store request log as json in Couchbase
2. Store (and retrieve) visitor accumulated data as json in Couchbase
3. Use Couchbase Map and Reduce Views for statistics

Relevant (personalized) page creation

Context Aware Page Cache

Including thundering herd protection

And 100% personalized parts?

Built-in support for async **AJAX/ESI/SSI**

Recap Hippo's delivery tier

You do not need to tune it to make it fast.

However a fast framework
does not guarantee
a fast/snappy site

Delivery tier diagnostics

1. Possible to switch on/off in production
2. Dissects a request through the framework and monitors time spend in different parts
3. Output to log or some storage like ElasticSearch and inspect it with Kibana

Requirements for Hippo's delivery tier framework

1. support many concurrent visitors
2. instantly reflect frequently changing content
3. runtime adding sites and/or changing URL's of existing sites
4. runtime changing the appearance of sites
5. search including authorization
6. faceted navigation requiring authorized counts
7. personalization of pages
8. storing of visitor data

Diagnostics

```
- HstFilter (124ms): {hostName=www.demo.test.onehippo.com, uri=/site/, query=null}
  |- Host and Mount Matching (0ms): {}
  |- Sitemap Matching (0ms): {}
  `- Pipeline processing (124ms): {pipeline=null}
     |- Targeting update valve (4ms): {}
     |- HstComponentInvokerProfiler (4ms): {method=doBeforeRender, window=homepage, componen
     |- HstComponentInvokerProfiler (0ms): {method=doBeforeRender, window=main, component=co
     |- HstComponentInvokerProfiler (0ms): {method=doBeforeRender, window=home-boxes-left, c
     |- HstComponentInvokerProfiler (4ms): {method=doBeforeRender, window=latestjobs, compon
     |   `- HstQuery (4ms): {query=//*[(@hippo:paths='a5a24dd3-04a0-4ed1-a59a-ffc960ae69f2')
((@jcr:primaryType='hippogogreen:job'))] order by @hippogogreen:closingdate descending}
     |- HstComponentInvokerProfiler (8ms): {method=doBeforeRender, window=latestcomments, co
     |   `- HstQuery (8ms): {query=//*[(@hippo:paths='a9868436-8f6a-4da0-9b67-7eb6e013c885')
((@jcr:primaryType='hippogogreen:comment'))] order by @hippostdpubwf:lastModificationDate de
     |- HstComponentInvokerProfiler (0ms): {method=doBeforeRender, window=home-boxes-right,
     |- HstComponentInvokerProfiler (4ms): {method=doBeforeRender, window=latestevents, comp
     |   `- HstQuery (4ms): {query=//*[(@hippo:paths='392850eb-ba46-472f-a1d5-e69aeccfb65d')
xs:dateTime('2012-12-03T14:40:46.638+01:00')) and ((@jcr:primaryType='hippogogreen:event'))]
     |- HstComponentInvokerProfiler (8ms): {method=doBeforeRender, window=latestreviews, com
     |   `- HstQuery (8ms): {query=//*[(@hippo:paths='6896b028-bfaf-48fd-bd35-6fe38cea758b')
((@jcr:primaryType='hippogogreen:review'))] order by @hippogogreen:closingdate descending}
     |- HstComponentInvokerProfiler (0ms): {method=doBeforeRender, window=home-boxes-promo,
```

We are hiring!

HIPPO

goto; conference


goto;
conference
Please evaluate this talk via the mobile app!
INTERNATIONAL
SOFTWARE DEVELOPMENT
CONFERENCE
Follow us @gotoamst
HIPPO
goto;
conference