securing

© D.Fletcher for CloudTweaks.com

Jakub Kaluzny

Big problems with big data – Hadoop interfaces security

ZeroNights, Moscow, 2015

whoami

Sr. IT Security Consultant at SecuRing

- Consulting all phases of development
- penetration tests
- high-risk applications and systems

Researcher

- Hadoop, FOREX, MFP printers, proprietary network protocols

Agenda

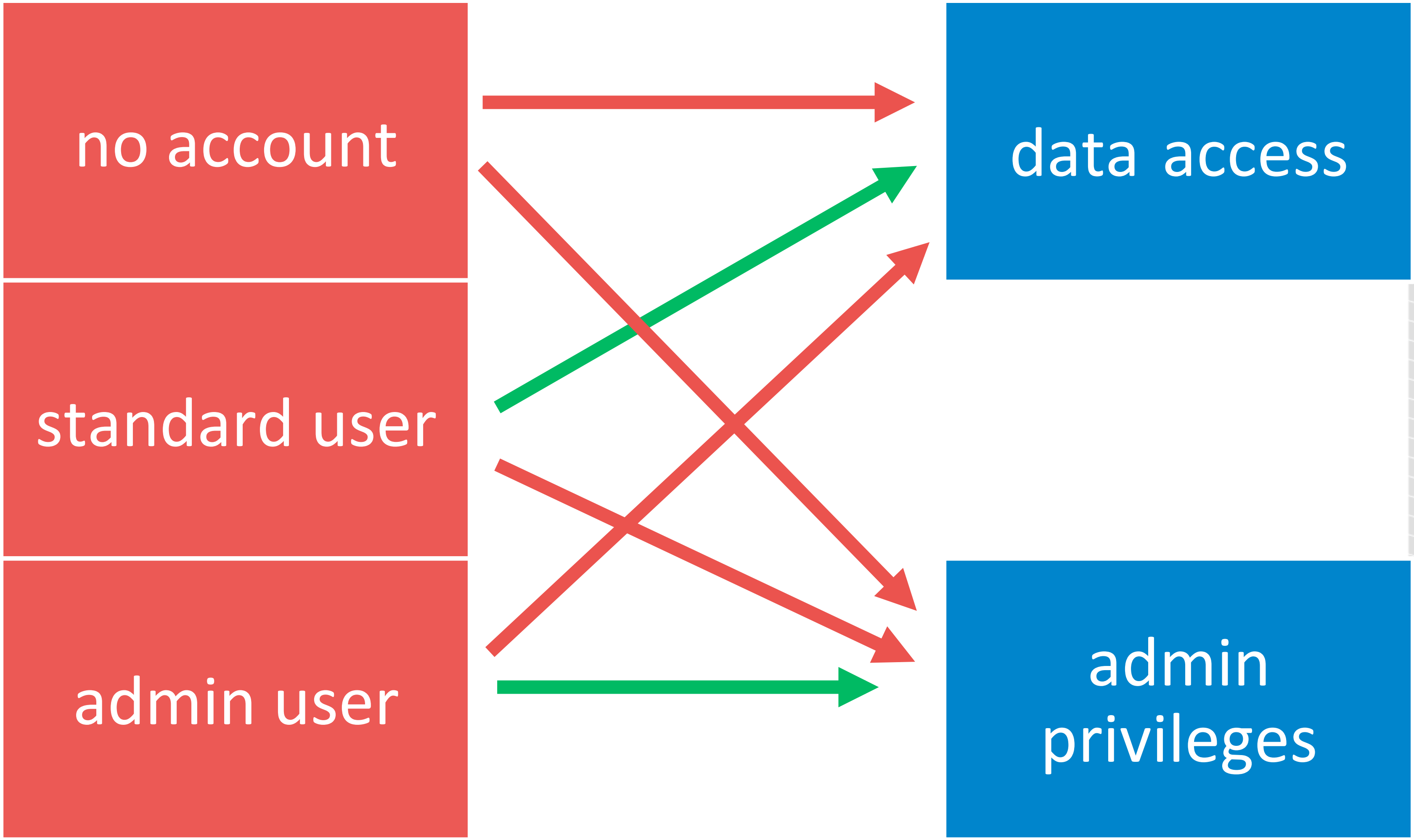
Big data nonsenses

Crash course on hacking Hadoop installations

Ways to protect big data environments

Expect some CVEs

Results summary



no account

standard user

admin user

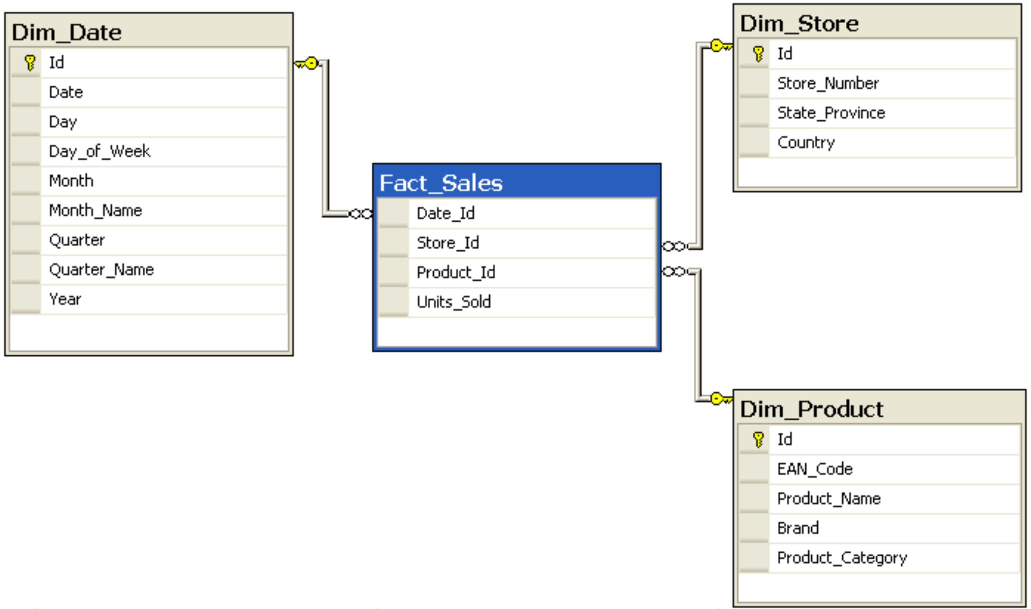
data access

admin privileges

WHAT IS HADOOP?

Know your target

Normal database



Dim_Date
- Id
- Date
- Day
- Day_of_Week
- Month
- Month_Name
- Quarter
- Quarter_Name
- Year

Fact_Sales
- Date_Id
- Store_Id
- Product_Id
- Units_Sold

Dim_Store
- Id
- Store_Number
- State_Province
- Country

Dim_Product
- Id
- EAN_Code
- Product_Name
- Brand
- Product_Category

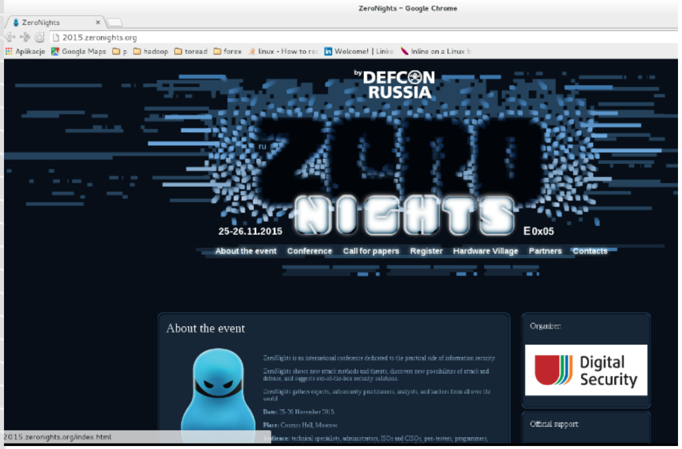
Users

Roles

Data

Model

Normal database architecture



About the event

ZeroNights is an international conference dedicated to the practical side of information security.

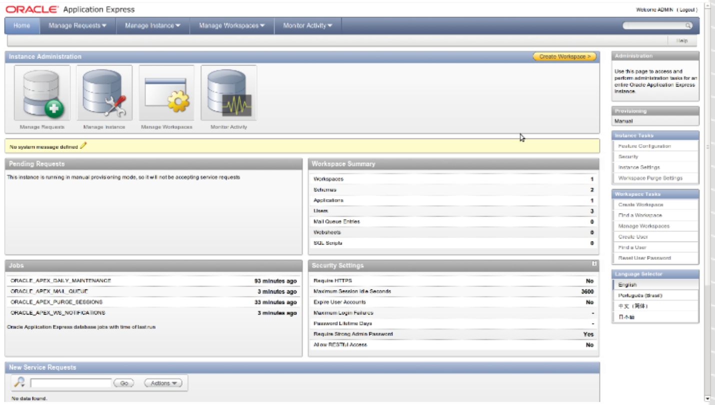
ZeroNights shows new attack methods and threats, discovers new possibilities of attack and defense, and suggests out-of-the-box security solutions.

ZeroNights gathers experts, infosecurity practitioners, analysts, and hackers from all over the world.

Date: 25-26 November 2015

Place: Cosmos Hall, Moscow

Audience: technical specialists, administrators, ISOs and CISOs, pen-testers, programmers,



Instance Administration

No system message defined

Pending Requests

This instance is running in manual provisioning mode, so it will not be accepting service requests

Workspace Summary

Workspaces	1
Schemas	2
Applications	1
Users	3
Mail Queue Entries	0
Websheets	0
SQL Scripts	0

Jobs

ORACLE_APEX_DAILY_MAINTENANCE	93 minutes ago
ORACLE_APEX_MAIL_QUEUE	3 minutes ago
ORACLE_APEX_PURGE_SESSIONS	33 minutes ago
ORACLE_APEX_WS_NOTIFICATIONS	3 minutes ago

Oracle Application Express database jobs with time of last run

Security Settings

Require HTTPS	No
Maximum Session Idle Seconds	3600
Expire User Accounts	No
Maximum Login Failures	-
Password Lifetime Days	-
Require Strong Admin Password	Yes
Allow RESTful Access	No

New Service Requests

No data found.

Administration

Use this page to access and perform administration tasks for an entire Oracle Application Express instance.

Provisioning

Manual

Instance Tasks

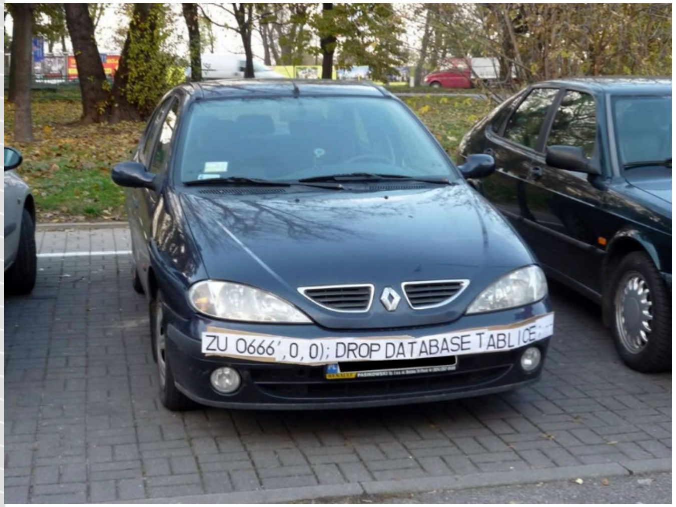
- Feature Configuration
- Security
- Instance Settings
- Workspace Purge Settings

Workspace Tasks

- Create Workspace
- Find a Workspace
- Manage Workspaces
- Create User
- Find a User
- Reset User Password

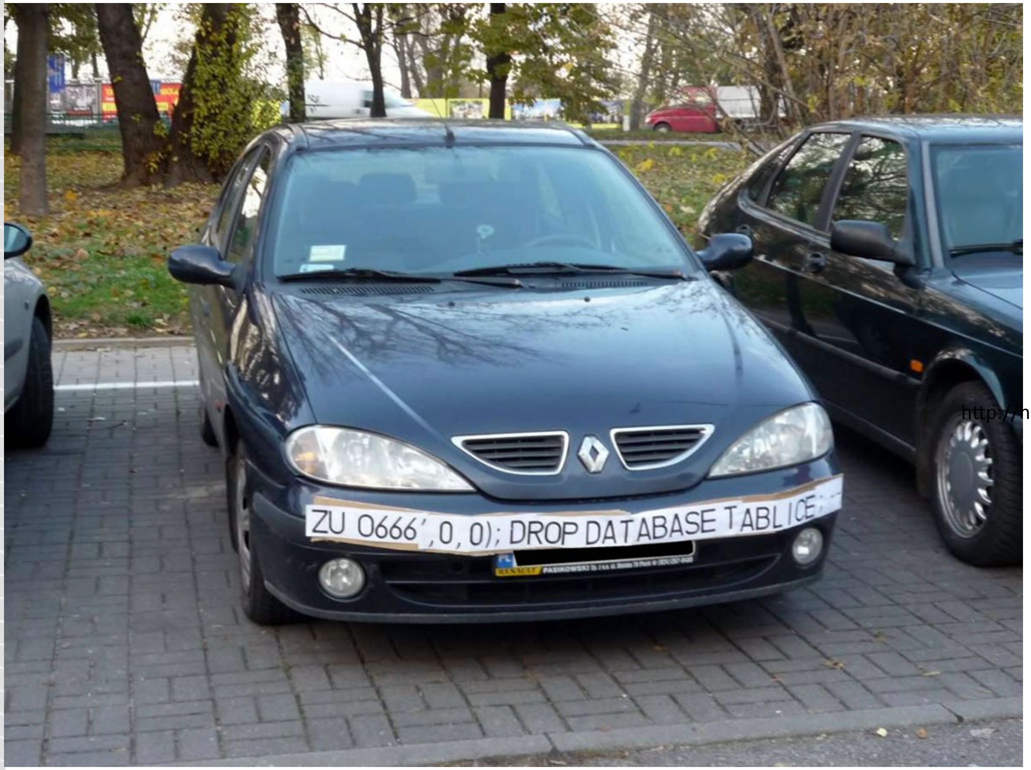
Language Selector

- English
- Português (Brasil)
- 中文（简体）
- 日本語



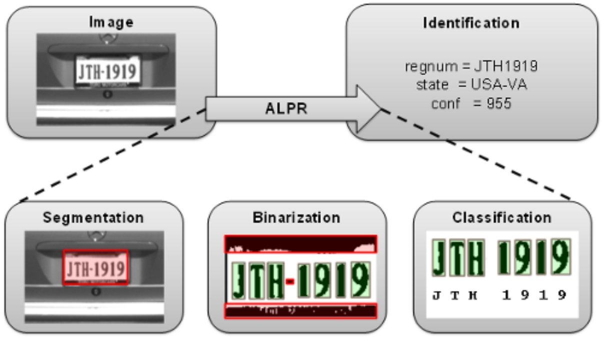
http://hackaday.com/2014/04/04/sql-injection-fools-speed-traps-and-clears-your-record/

Still normal database scenario



http://hackaday.com/2014/04/04/sql-injection-fools-speed-traps-and-clears-your-record/

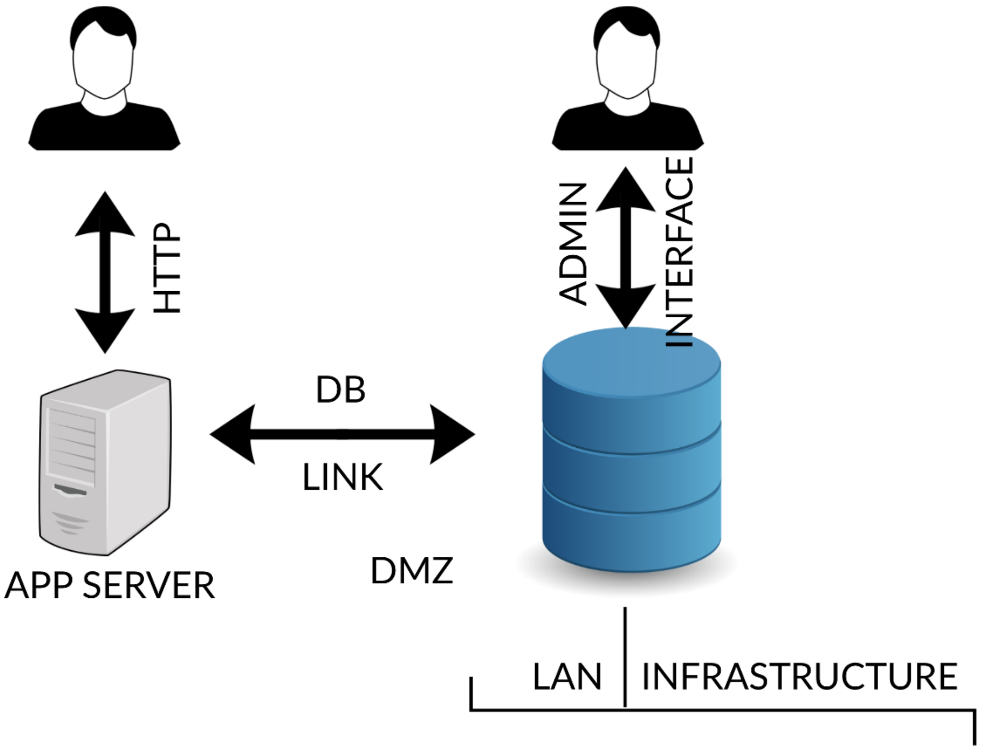
http://hococonnect.blogspot.com/2015/06/red-light-cameras-in-columbia.html



http://8z4.net/images/ocr-technology

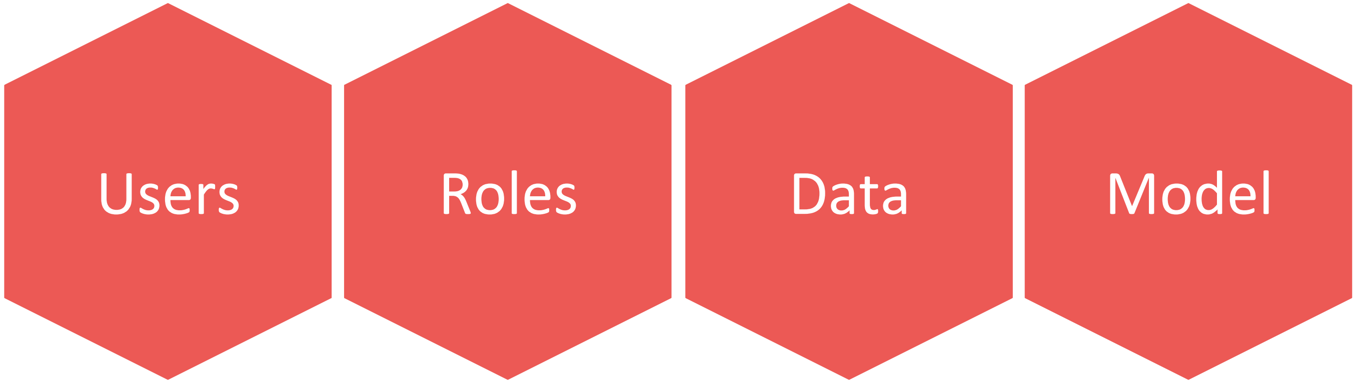
CWE-xxx: SQL Injection through license plate

Normal database injection points



Normal database

Clear rules



Users

Roles

Data

Model

Clear target

Anegdote

user db,

a lot of clients

critical banking data,
one supplier

Only one common table

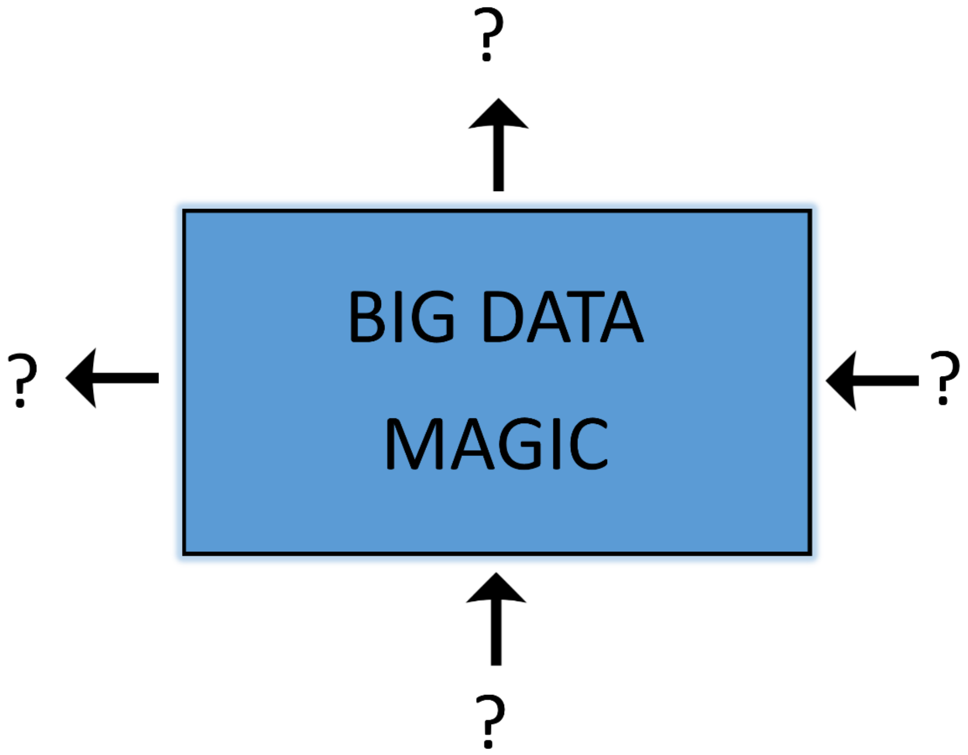
Q: Why don’t you split it into 2 dbs with a db link?
A: Too much effort and we want to have fast statistics from all data.

What is Hadoop?

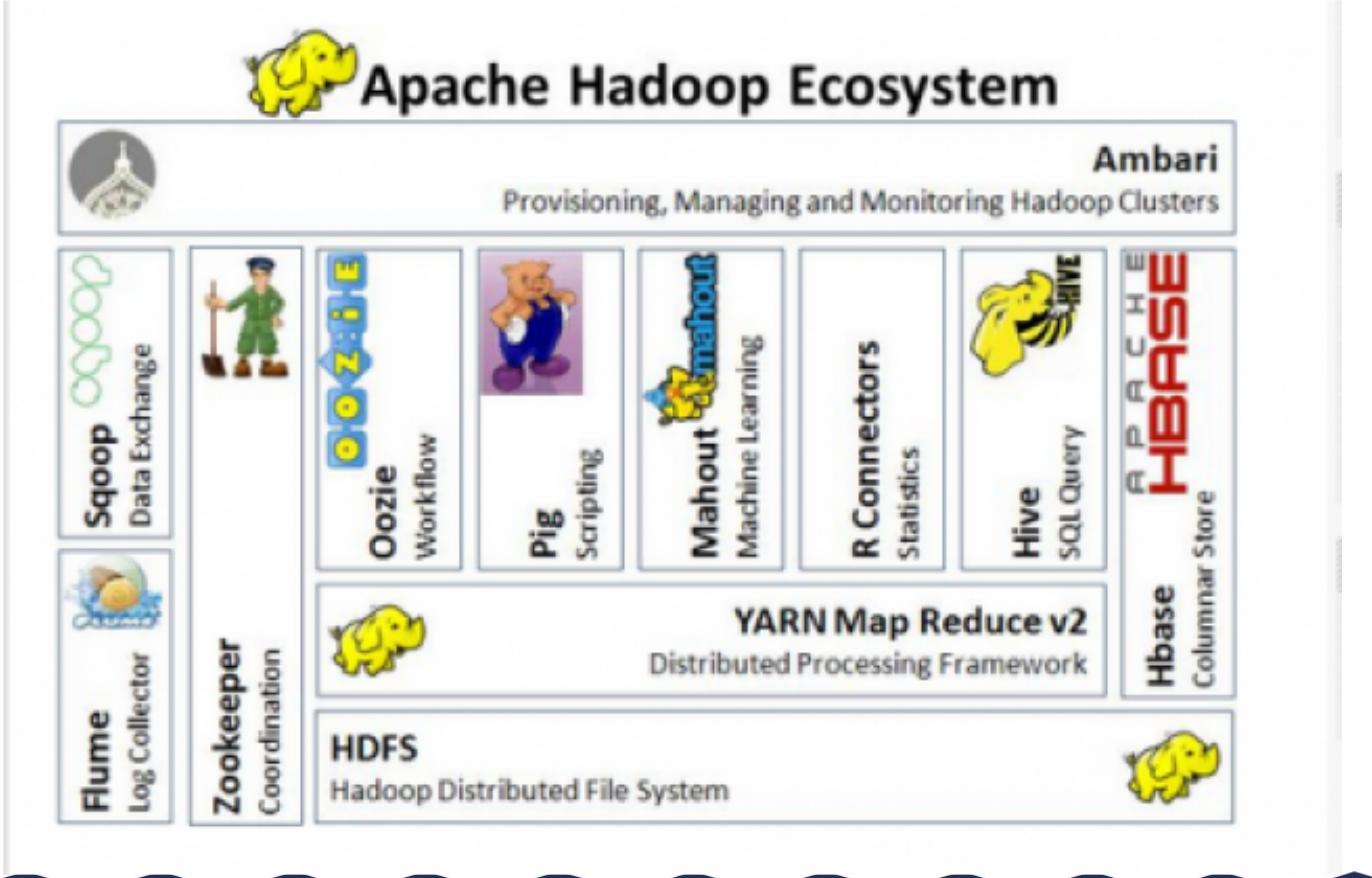
http://fiveprime.org/blackmagic.cgi?id=7007203773

https://www.flickr.com/photos/photonquantique/2596581870/

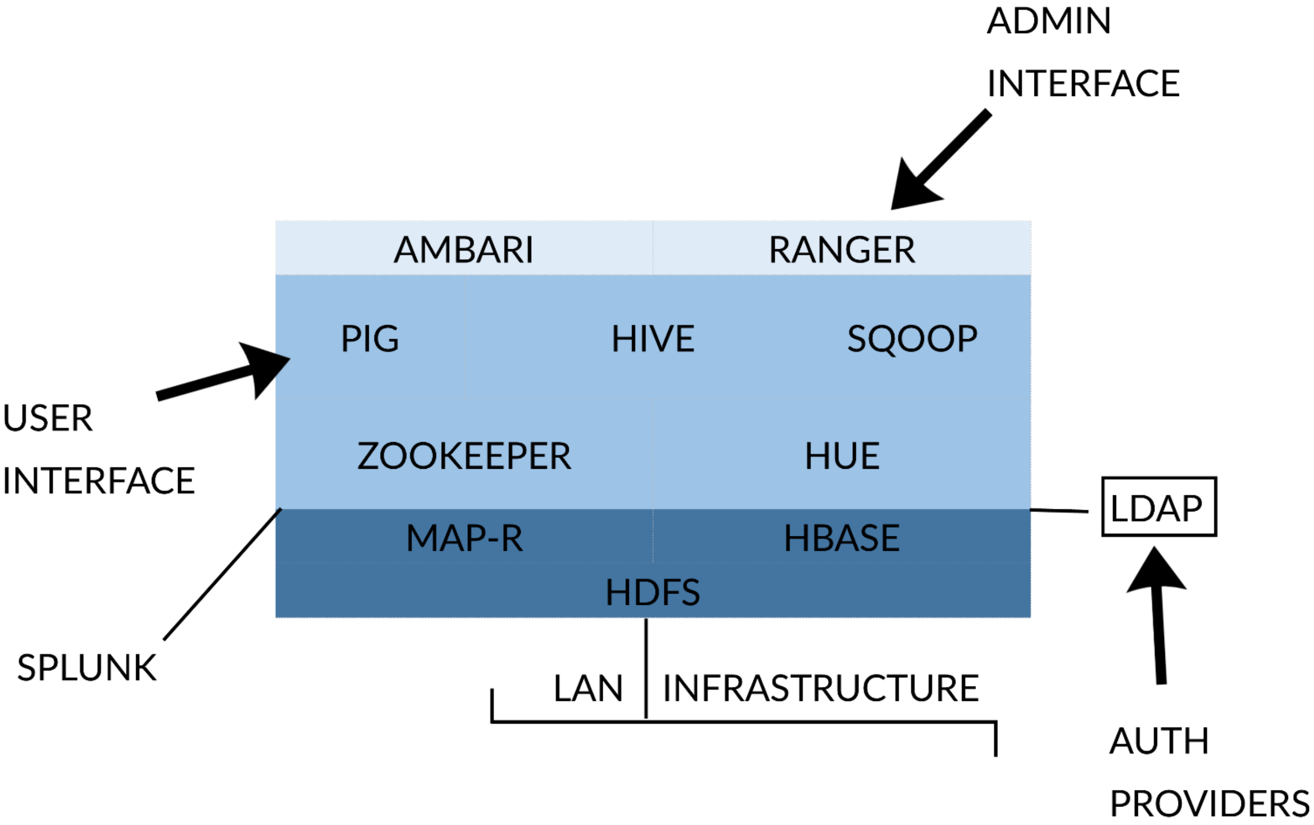
Hadoop architecture schema



More on Hadoop



Hadoop injection points



ADMIN INTERFACE

AMBARI | RANGER

PIG | HIVE | SQOOP

ZOOKEEPER | HUE

MAP-R | HBASE

HDFS

USER INTERFACE

SPLUNK

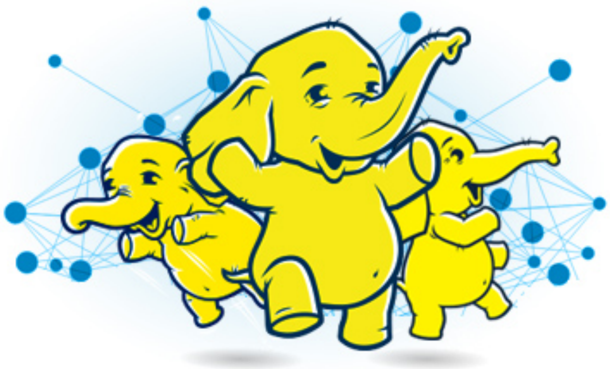
LDAP

LAN | INFRASTRUCTURE

AUTH PROVIDERS

Hadoop scenario

https://www.flickr.com/photos/mattimattila/8349565473



http://bigdataanalyticsnews.com/tag/hortonworks/

https://en.wikipedia.org/wiki/Moneygami

What is a lot of data?

facebook

21 PB of storage in a single HDFS cluster

2000 machines

12 TB per machine (a few machines have 24 TB each)

1200 machines with 8 cores each + 800 machines with 16 cores each

32 GB of RAM per machine

15 map-reduce tasks per machine

What is a lot of data?

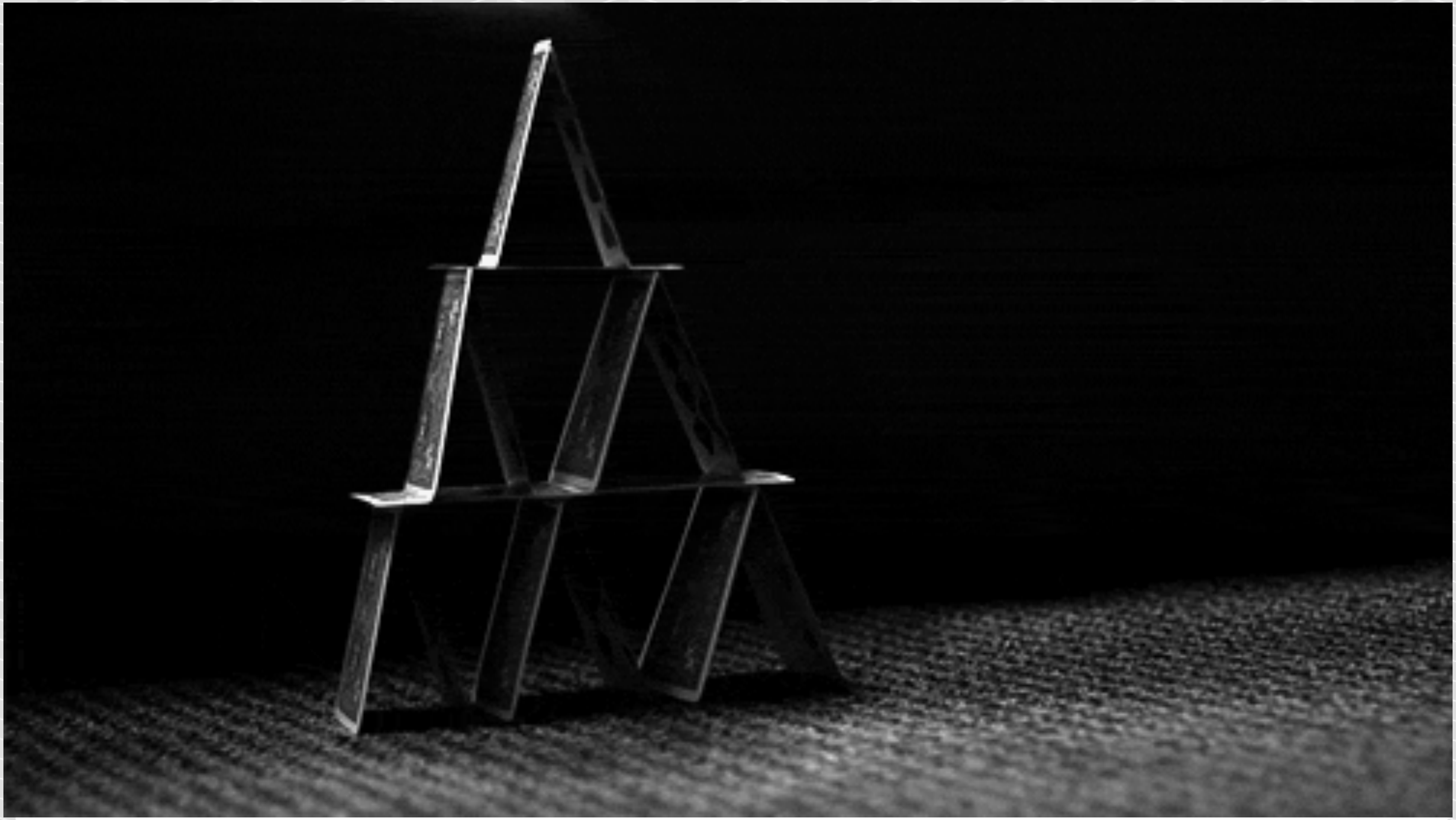
Our latest assessment:

- 32 machines, 8 cores each
- 24TB per machine
- 64 GB of RAM per machine

http://mrrobot.wikia.com/wiki/E_Corp

- Almost 1 PB disk space and 2TB of RAM

Attacker perspective



https://plus.google.com/+Magiccardtrickszonetips

RISK ANALYSIS

Know your threats

Risk analysis

Who

How

What

Who?

Business perspective: competitor, script-kiddies, APT

Technical perspective:

External attacker

- Anonymous
- Ex-employee

Insider

- Exployee (with some rights in Hadoop): user, admin
- Infected machine, APT

Risk analysis

Who
How
What

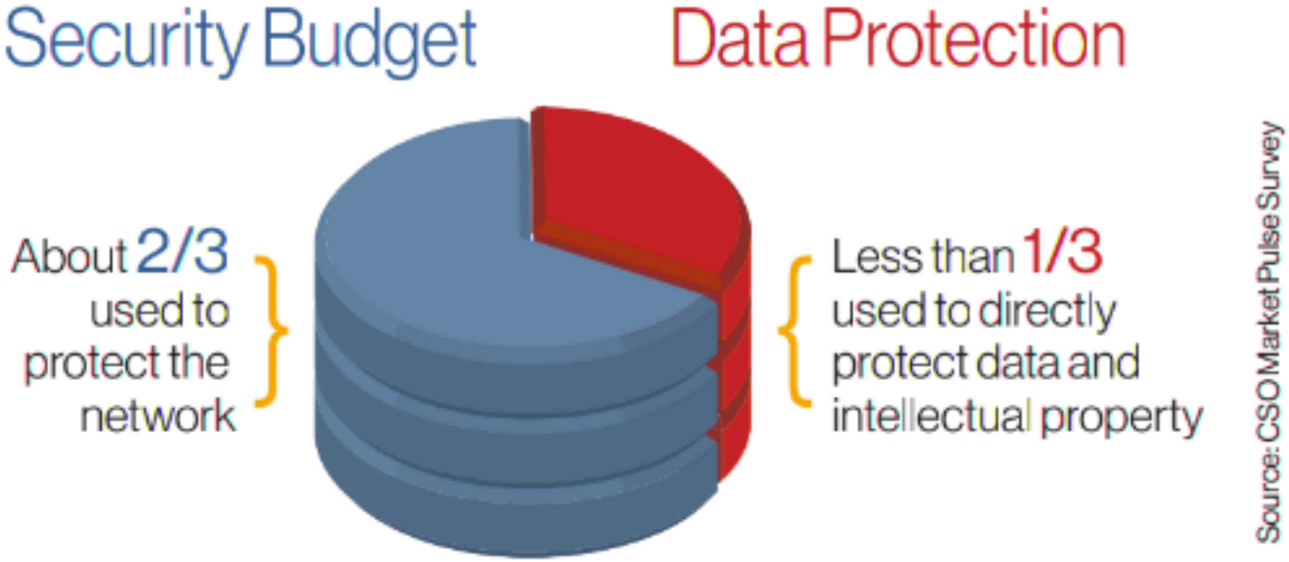
Full compromise

Wojciech Dworakowski

IT Security Expert, Owner at SecuRing, OWASP Poland Chapter Leader

Online banking owned by single attacker

Data safety vs. data security



Security Budget

Data Protection

About 2/3 used to protect the network

Less than 1/3 used to directly protect data and intellectual property

Source: CSO Market Pulse Survey

For what?

Q: What will be stored? A: „We do not know what data will be stored!”

Typical bank scenario

All transaction data

All sales data

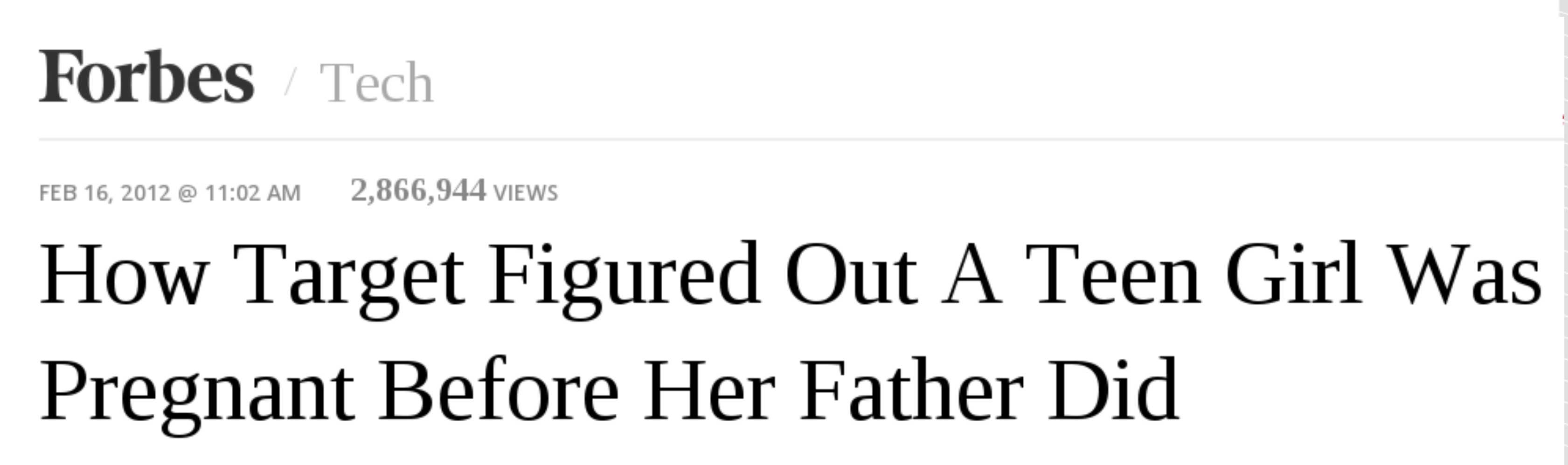
All client data

https://www.reddit.com/r/gifs/comments/37aara/calculations_intensify/

http://thewondrous.com/julia-gunthel-worlds-most-flexible-secretary/

Bigdata analytic says: „People who bought a dashcam are more likely to take a loan for a new car in the next month”

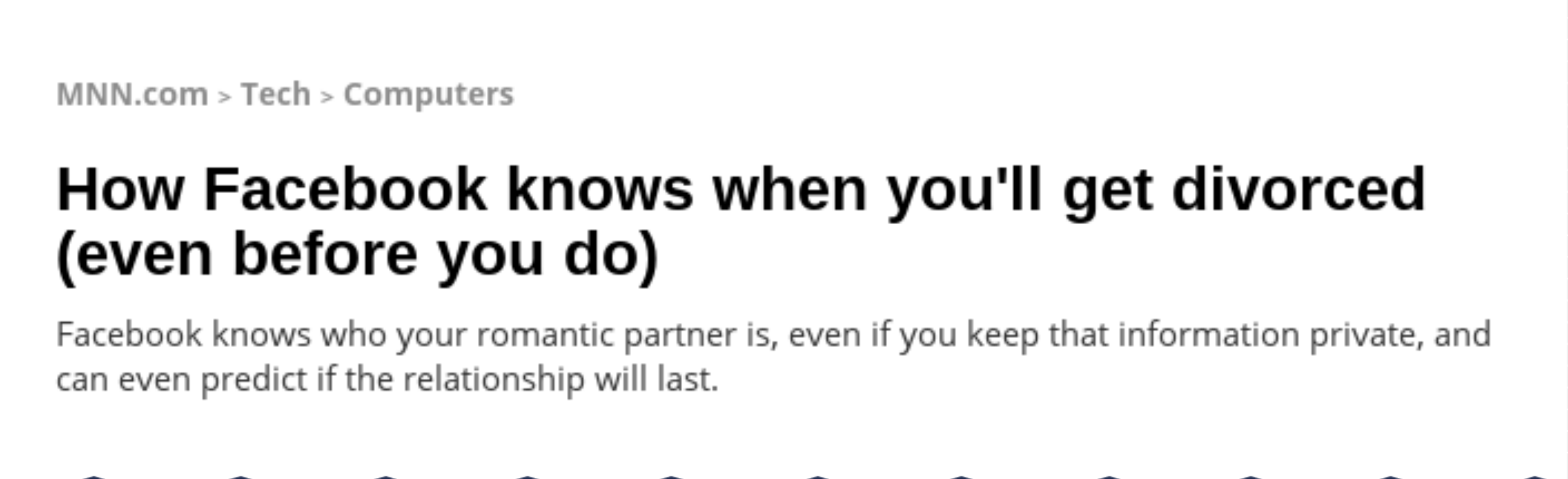
For what? Data theft



Forbes / Tech

FEB 16, 2012 @ 11:02 AM 2,866,944 VIEWS

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did



MNN.com > Tech > Computers

How Facebook knows when you'll get divorced (even before you do)

Facebook knows who your romantic partner is, even if you keep that information private, and can even predict if the relationship will last.

Other

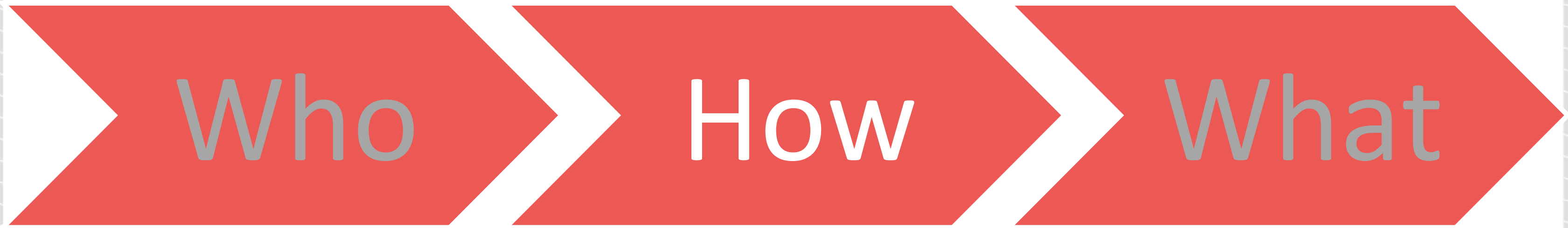
Privilege escalation

- Authentication bypass

Abuse

- DoS
- Data tampering

Risk analysis



Who

How

What

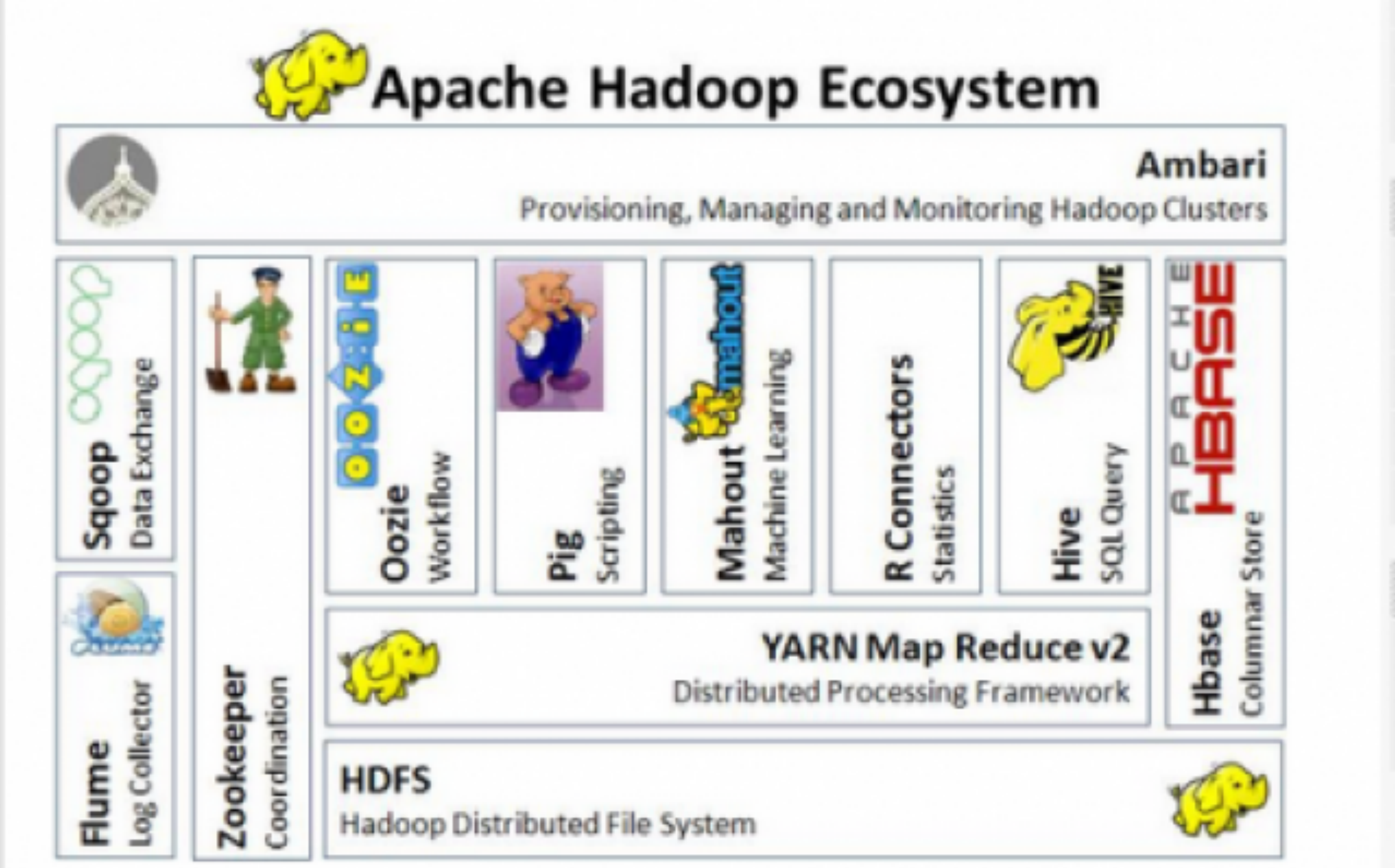
How?

https://en.wikipedia.org/wiki/Dowsing#Rods

WHAT HADOOP REALLY IS

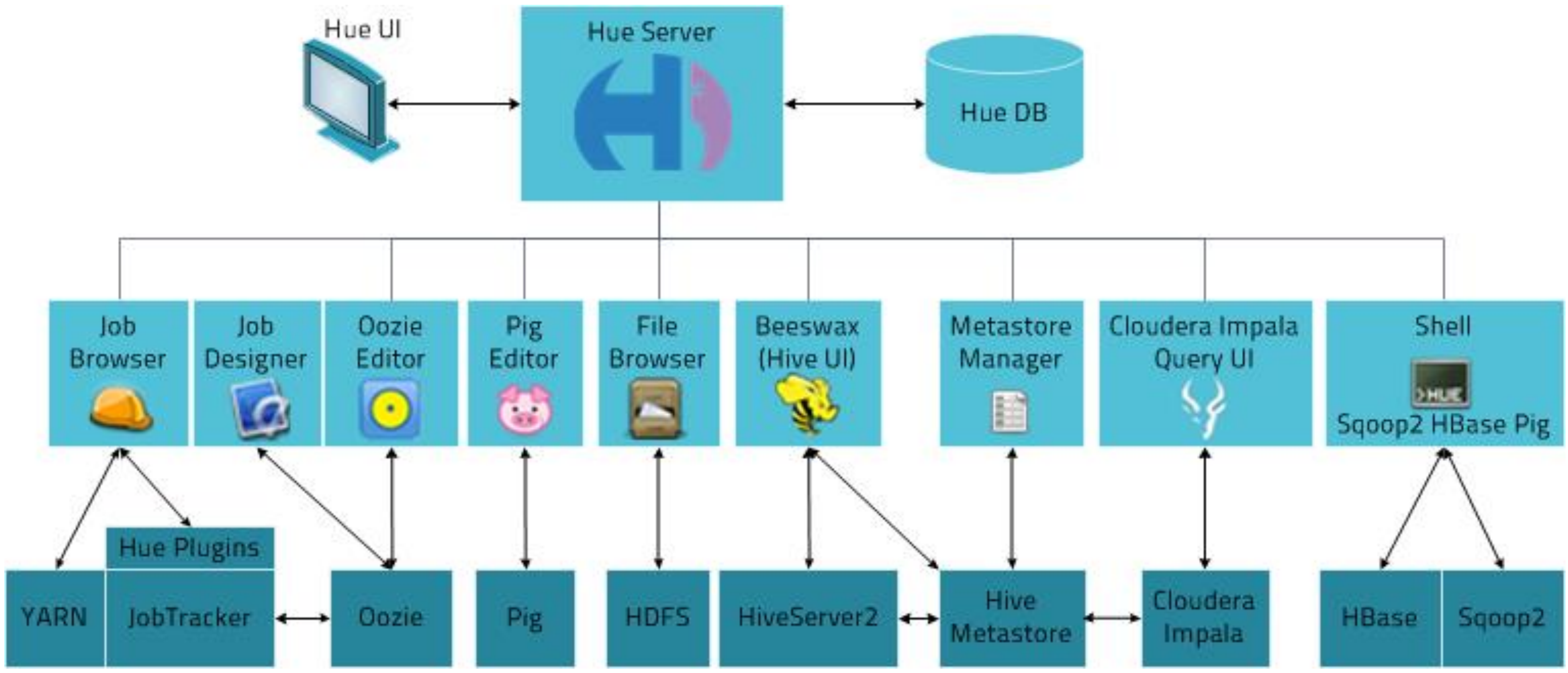
under sales-magic-cloud-big-data cover

Typical architecture



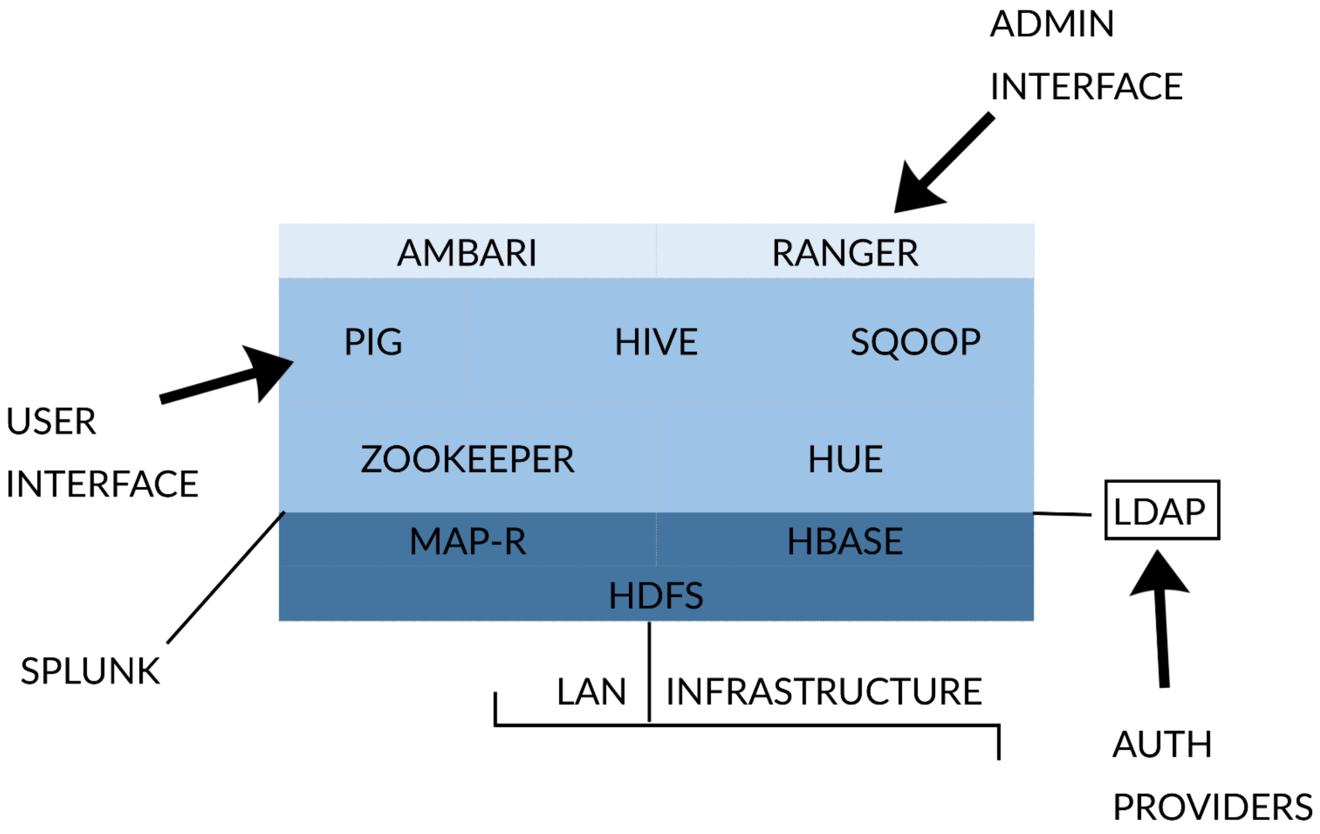
http://thebigdatablog.weebly.com/blog/the-hadoop-ecosystem-overview

Apache Hue



http://techbusinessintelligence.blogspot.com/2014/11/tableau-software-cloudera-hadoop.html

Hadoop injection points



ADMIN INTERFACE

AMBARI
RANGER

PIG
HIVE
SQOOP

USER INTERFACE

ZOOKEEPER
HUE

MAP-R
HBASE
HDFS

LDAP

SPLUNK

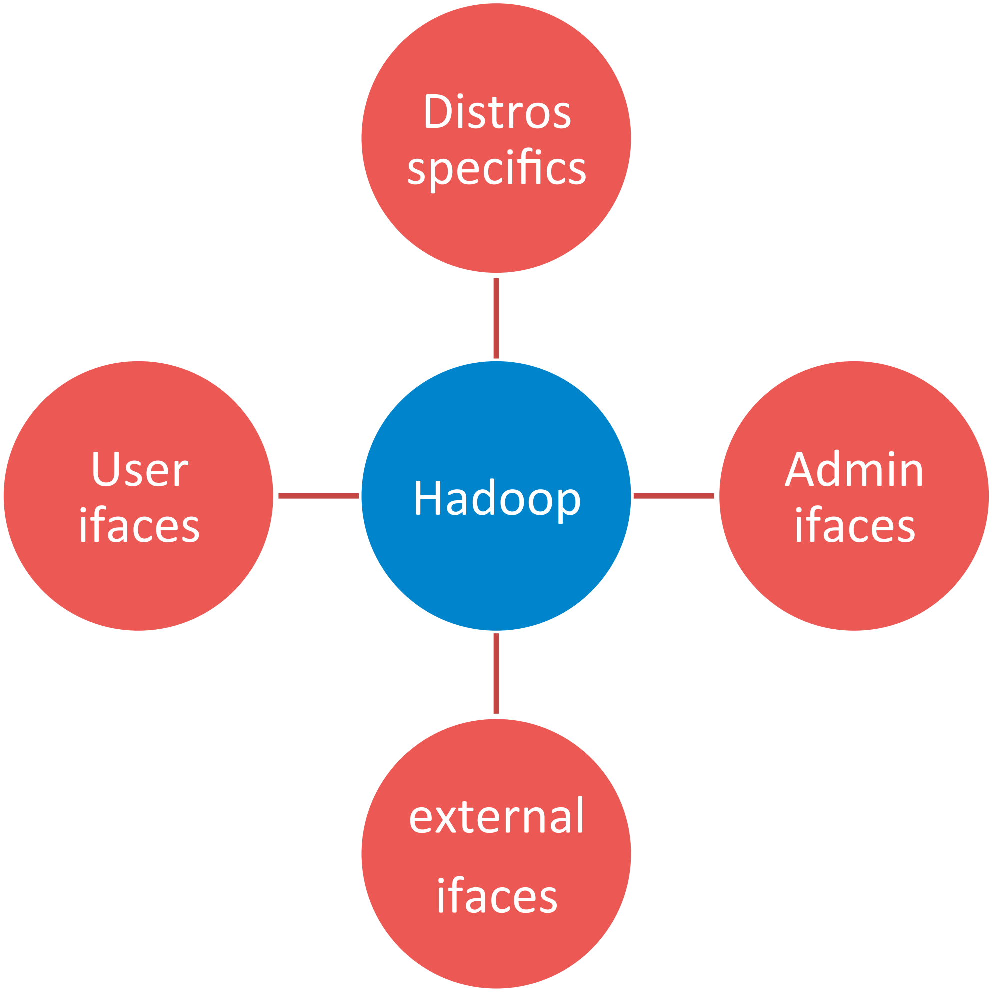
LAN
INFRASTRUCTURE

AUTH PROVIDERS

Differs much amongst distros

INTERFACES

Interfaces



Distros specifics

User ifaces

Hadoop

Admin ifaces

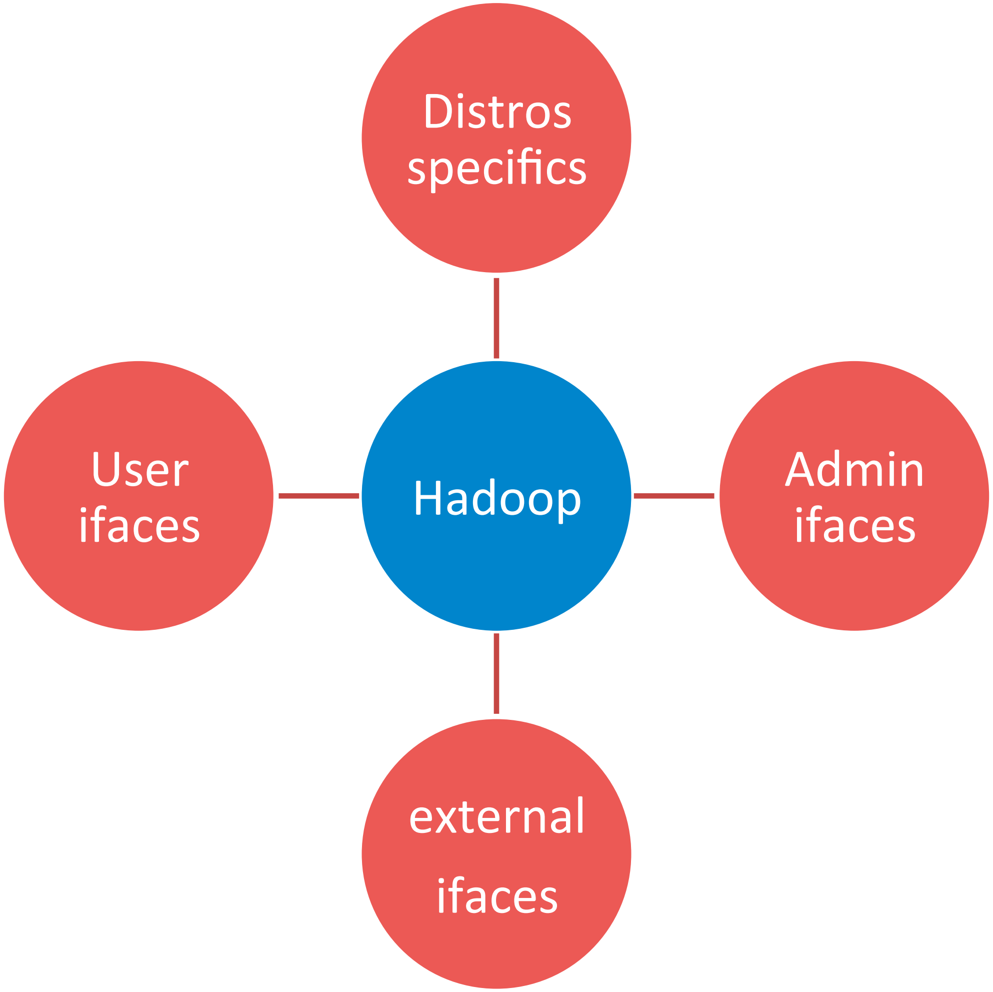
external

ifaces

OUR STORY WITH BIG DATA ASSESSMENT

a.k.a. crash course on hacking big data environments

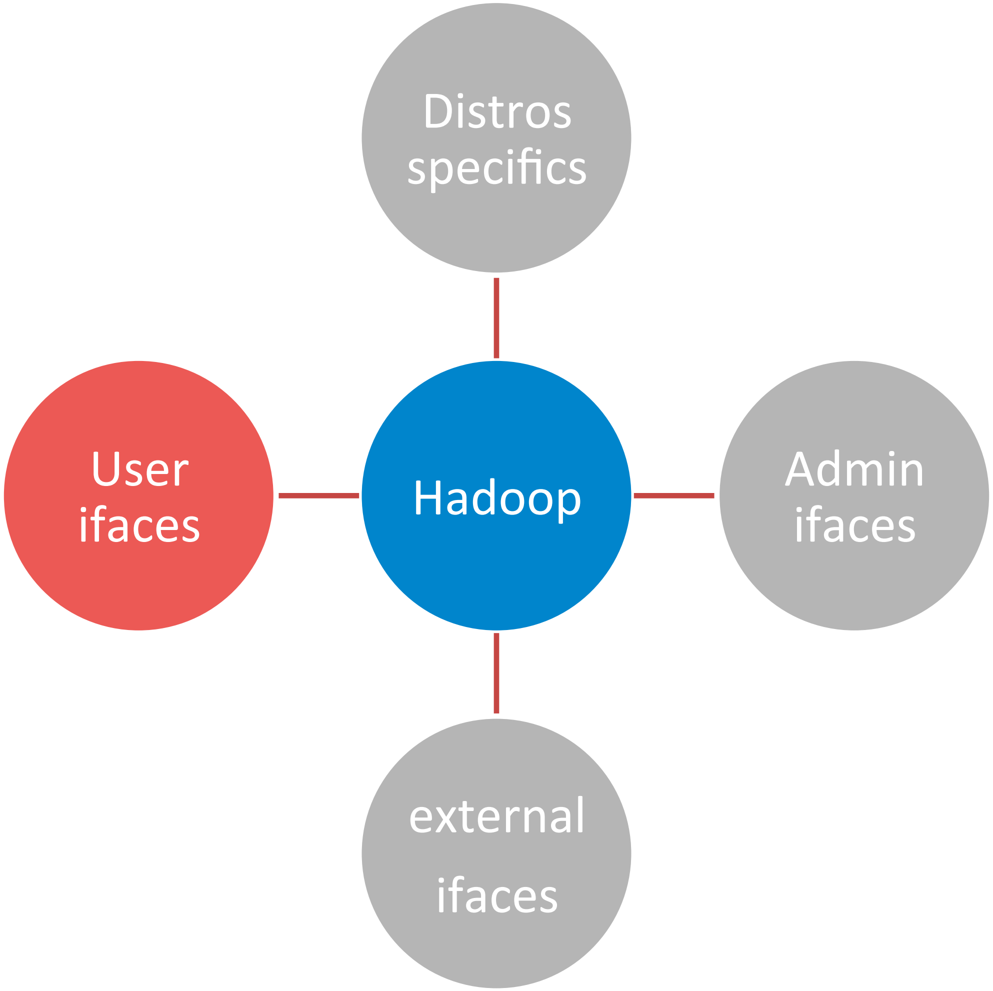
Interfaces


Distros specifics
User ifaces
Hadoop
Admin ifaces
external ifaces

USER INTERFACES

for employees and applications

User interfaces



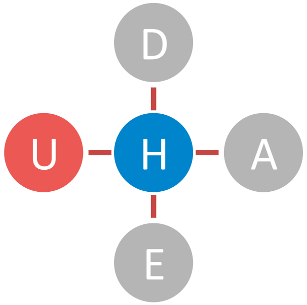
User interfaces

Apache Hue

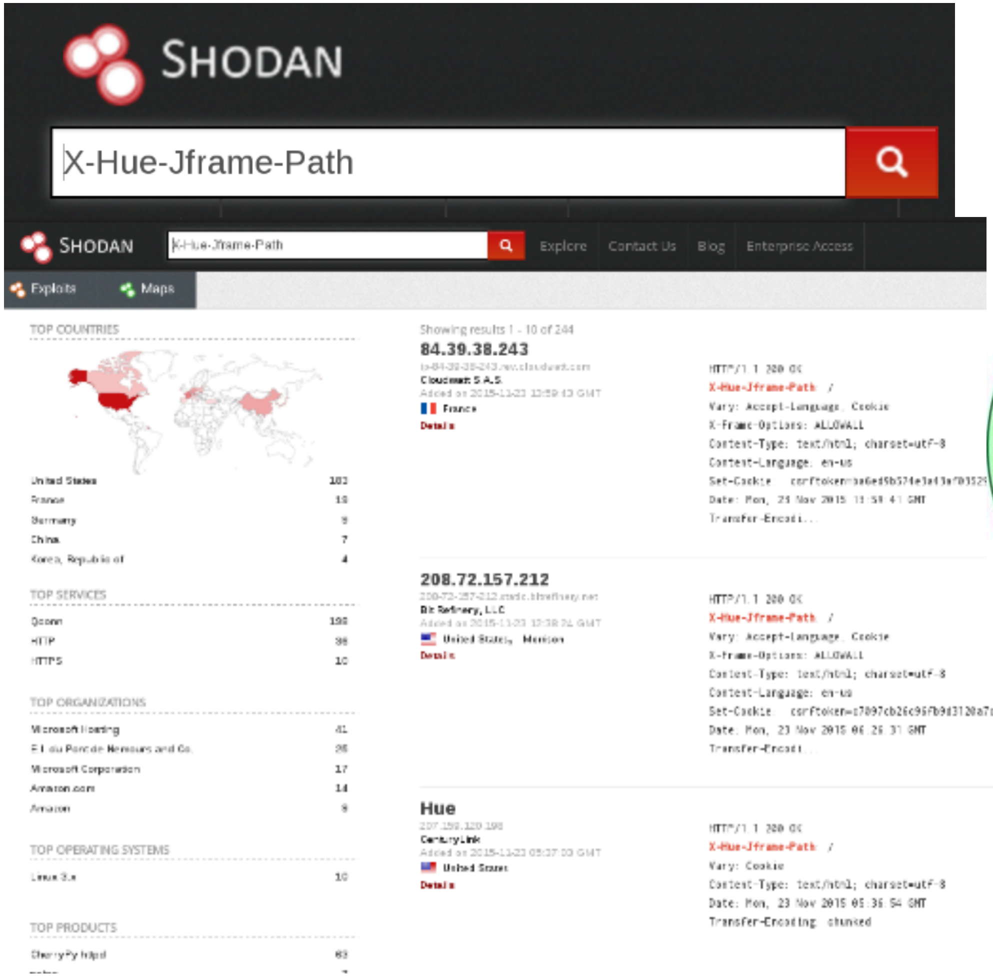
- Pig, Hive, Impala, Hbase, Zookeeper, Mahout, Oozie

Other

- Tez, Solr, Slider, Spark, Phoenix, Accummulo, Storm



Is Hue an internal interface?



SHODAN

X-Hue-Jframe-Path

TOP COUNTRIES

United States	183
France	18
Germany	9
China	7
Korea, Republic of	4

TOP SERVICES

Qconn	198
HTTP	38
HTTPS	10

TOP ORGANIZATIONS

Microsoft Hosting	41
E.I. du Pont de Nemours and Co.	26
Microsoft Corporation	17
Amazon.com	14
Amazon	9

TOP OPERATING SYSTEMS

Linux 3.x	10

TOP PRODUCTS

CherryPy httpd	63

Showing results 1 - 10 of 244

84.39.38.243
ip-84-39-38-243.rev.cloudwatt.com
Cloudwatt S.A.S.
Added on 2015-11-23 13:59:43 GMT
France
Details

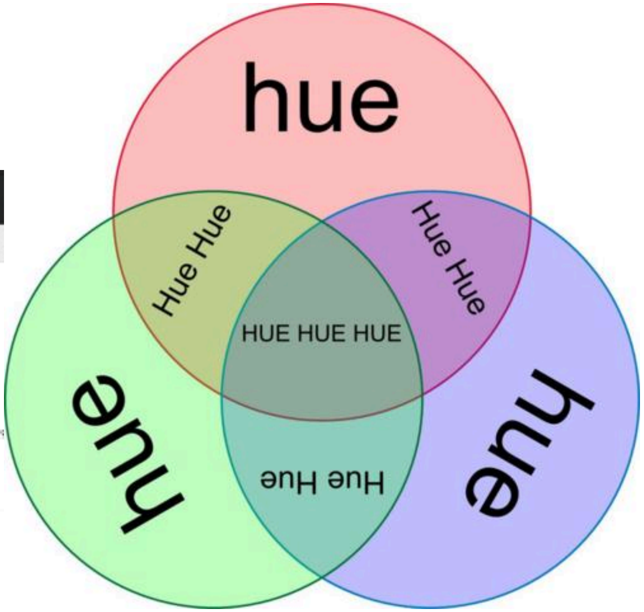
```
HTTP/1.1 200 OK
X-Hue-Jframe-Path: /
Vary: Accept-Language, Cookie
X-Frame-Options: ALLOWALL
Content-Type: text/html; charset=utf-8
Content-Language: en-us
Set-Cookie:  csrftoken=ba6ed5b574e3a43af03529
Date: Mon, 23 Nov 2015 13:59:41 GMT
Transfer-Encodi...
```

208.72.157.212
208-72-157-212.static.bitrefinery.net
Bit Refinery, LLC
Added on 2015-11-23 12:38:24 GMT
United States, Meridian
Details

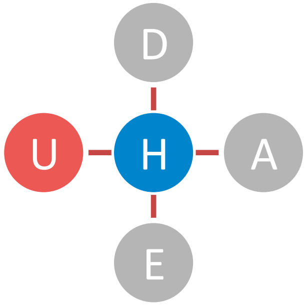
```
HTTP/1.1 200 OK
X-Hue-Jframe-Path: /
Vary: Accept-Language, Cookie
X-Frame-Options: ALLOWALL
Content-Type: text/html; charset=utf-8
Content-Language: en-us
Set-Cookie:  csrftoken=c7097cb26c96fb9d3120a7c6c94825c5; Max-Age=31449600; Path=/
Date: Mon, 23 Nov 2015 06:26:31 GMT
Transfer-Encodi...
```

Hue
207.159.120.198
CenturyLink
Added on 2015-11-23 05:37:03 GMT
United States
Details

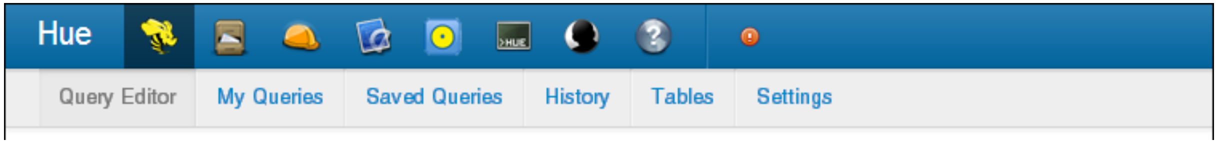
```
HTTP/1.1 200 OK
X-Hue-Jframe-Path: /
Vary: Cookie
Content-Type: text/html; charset=utf-8
Date: Mon, 23 Nov 2015 05:36:54 GMT
Transfer-Encoding: chunked
```

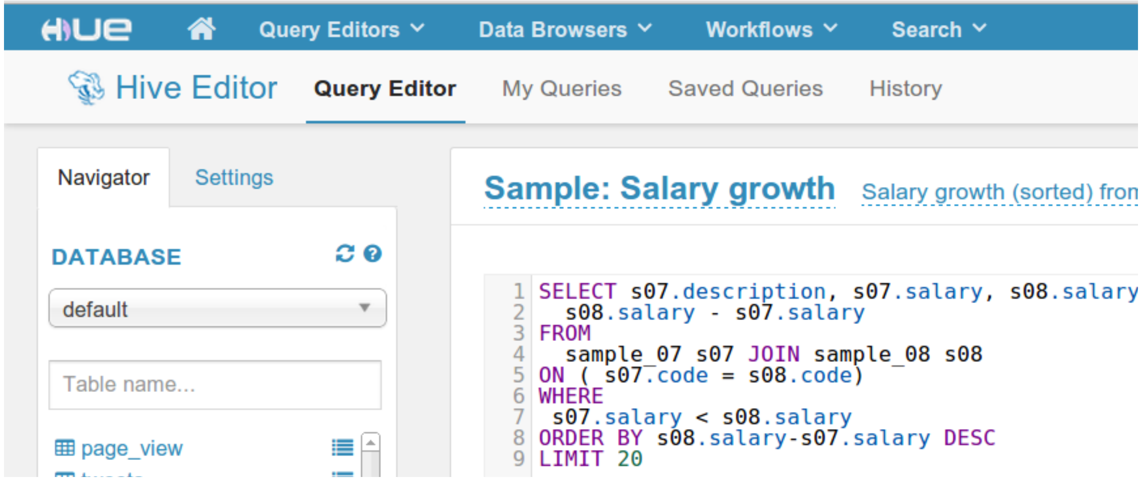


http://9gag.com/gag/awrwVL1/hue-hue-hue

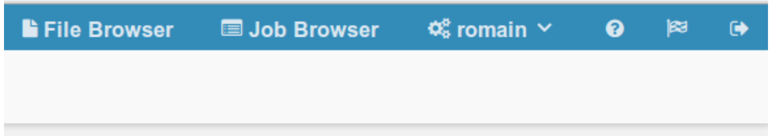


Apache Hue overview

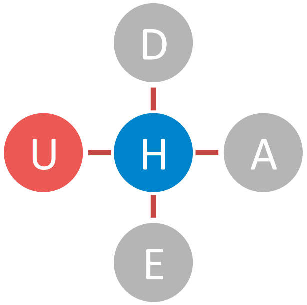




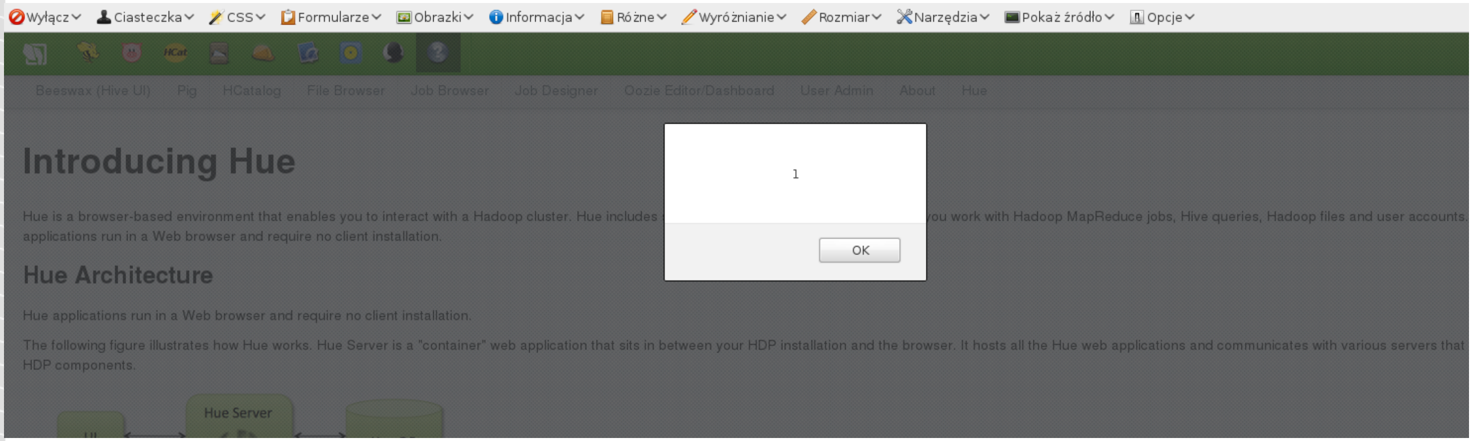
```
SELECT s07.description, s07.salary, s08.salary
  s08.salary - s07.salary
FROM
  sample_07 s07 JOIN sample_08 s08
ON ( s07.code = s08.code)
WHERE
 s07.salary < s08.salary
ORDER BY s08.salary-s07.salary DESC
LIMIT 20
```



http://gethue.com/



Apache Hue DOM XSS



Introducing Hue

Hue is a browser-based environment that enables you to interact with a Hadoop cluster. Hue includes ou work with Hadoop MapReduce jobs, Hive queries, Hadoop files and user accounts. applications run in a Web browser and require no client installation.

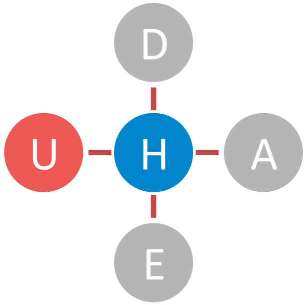
Hue Architecture

Hue applications run in a Web browser and require no client installation.

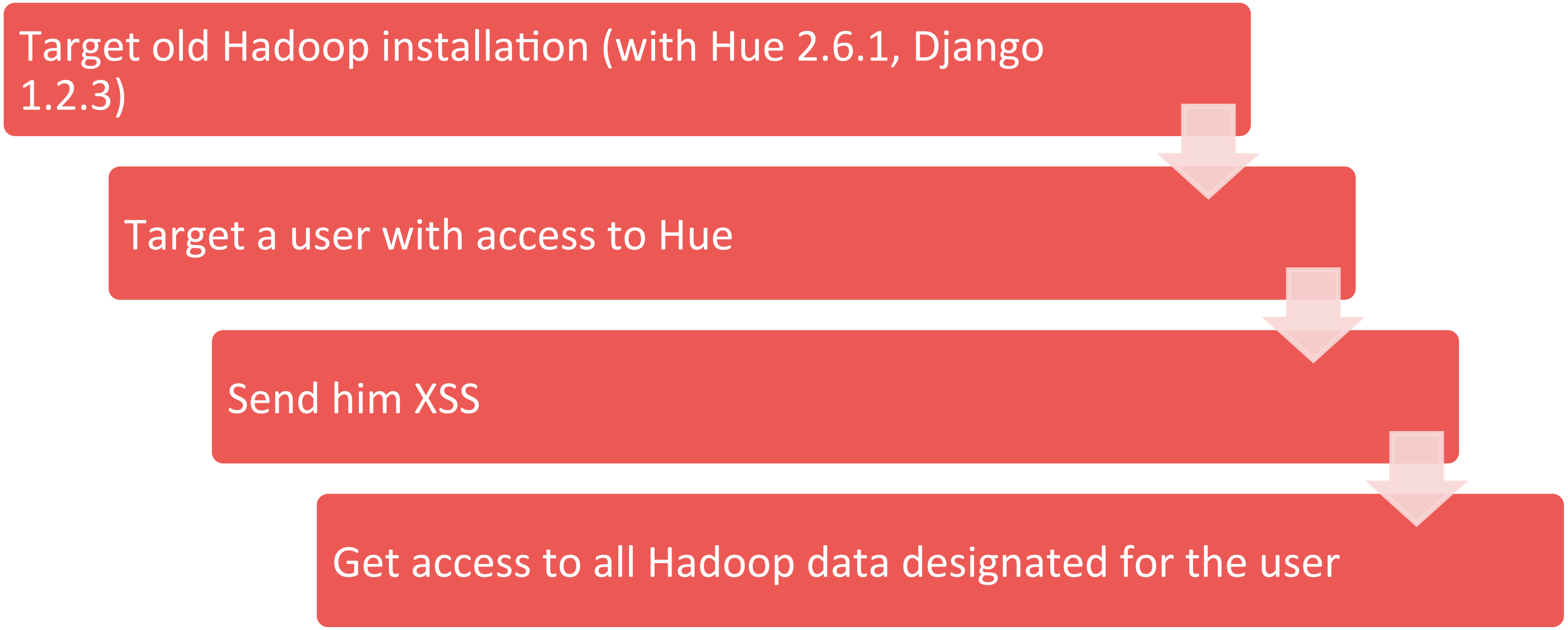
The following figure illustrates how Hue works. Hue Server is a "container" web application that sits in between your HDP installation and the browser. It hosts all the Hue web applications and communicates with various servers that HDP components.

```
var _anchor = $("a[name='" +
decodeURIComponent(window.location.hash.subs
tring(1)) + "']").last();
```

Payload: URL/help/#<img src="x" onerror="alert(1)">



Apache Hue attack scenario



Target old Hadoop installation (with Hue 2.6.1, Django 1.2.3)

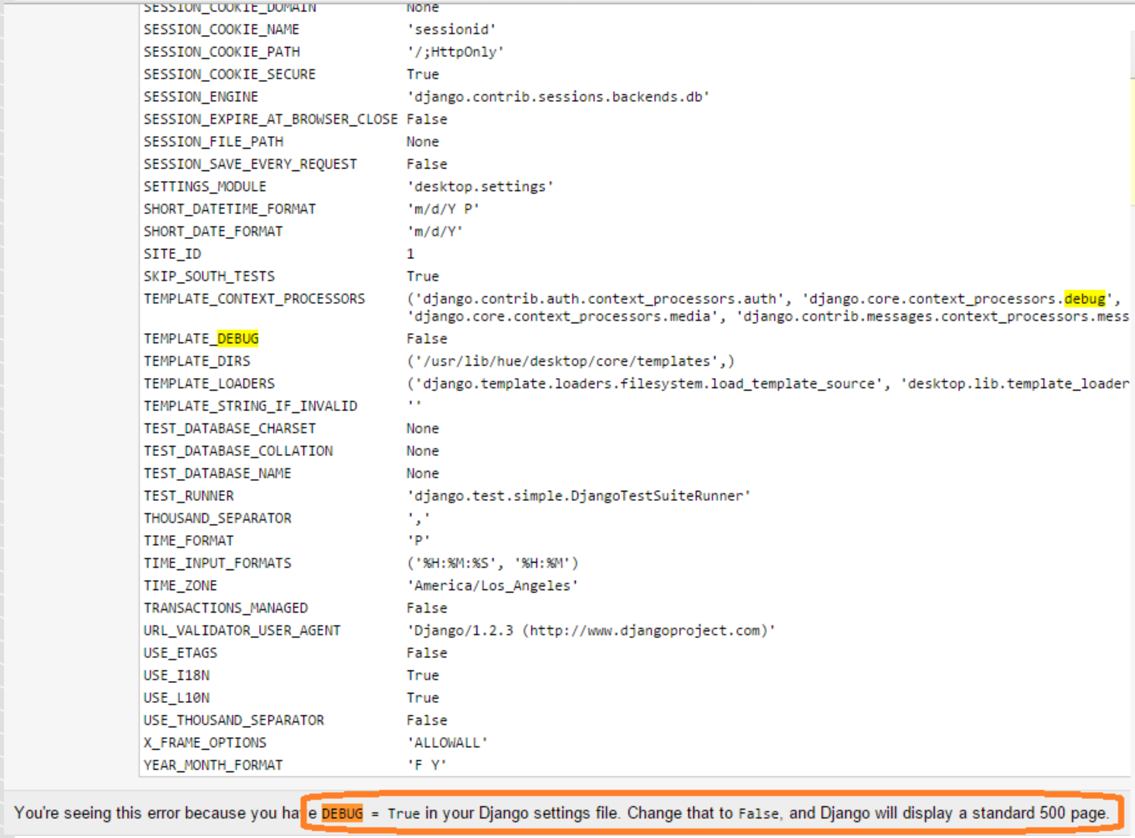
Target a user with access to Hue

Send him XSS

Get access to all Hadoop data designated for the user

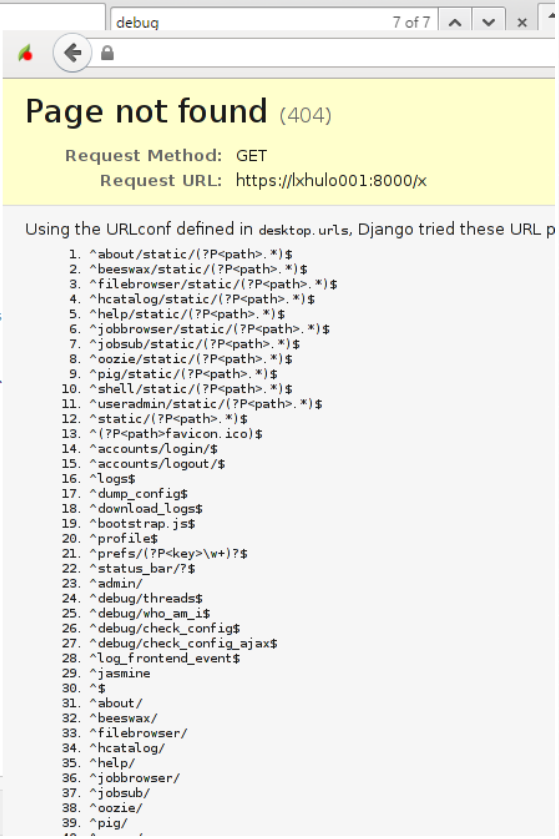
Default configurations sucks

X-Frame-Options:ALLOWALL



```
SESSION_COOKIE_DOMAIN               None
SESSION_COOKIE_NAME                 'sessionid'
SESSION_COOKIE_PATH                 '/;HttpOnly'
SESSION_COOKIE_SECURE               True
SESSION_ENGINE                      'django.contrib.sessions.backends.db'
SESSION_EXPIRE_AT_BROWSER_CLOSE     False
SESSION_FILE_PATH                   None
SESSION_SAVE_EVERY_REQUEST          False
SETTINGS_MODULE                     'desktop.settings'
SHORT_DATETIME_FORMAT               'm/d/Y P'
SHORT_DATE_FORMAT                   'm/d/Y'
SITE_ID                             1
SKIP_SOUTH_TESTS                    True
TEMPLATE_CONTEXT_PROCESSORS         ('django.contrib.auth.context_processors.auth', 'django.core.context_processors.debug',
                                    'django.core.context_processors.media', 'django.contrib.messages.context_processors.mess
TEMPLATE_DEBUG                      False
TEMPLATE_DIRS                       ('/usr/lib/hue/desktop/core/templates',)
TEMPLATE_LOADERS                    ('django.template.loaders.filesystem.load_template_source', 'desktop.lib.template_loader
TEMPLATE_STRING_IF_INVALID          ''
TEST_DATABASE_CHARSET               None
TEST_DATABASE_COLLATION             None
TEST_DATABASE_NAME                  None
TEST_RUNNER                         'django.test.simple.DjangoTestSuiteRunner'
THOUSAND_SEPARATOR                  ','
TIME_FORMAT                         'P'
TIME_INPUT_FORMATS                  ('%H:%M:%S', '%H:%M')
TIME_ZONE                           'America/Los_Angeles'
TRANSACTIONS_MANAGED                False
URL_VALIDATOR_USER_AGENT            'Django/1.2.3 (http://www.djangoproject.com)'
USE_ETAGS                           False
USE_I18N                            True
USE_L10N                            True
USE_THOUSAND_SEPARATOR              False
X_FRAME_OPTIONS                     'ALLOWALL'
YEAR_MONTH_FORMAT                   'F Y'
```

You're seeing this error because you have `DEBUG = True` in your Django settings file. Change that to `False`, and Django will display a standard 500 page.



Page not found (404)

Request Method: GET
Request URL: https://lxhulo001:8000/x

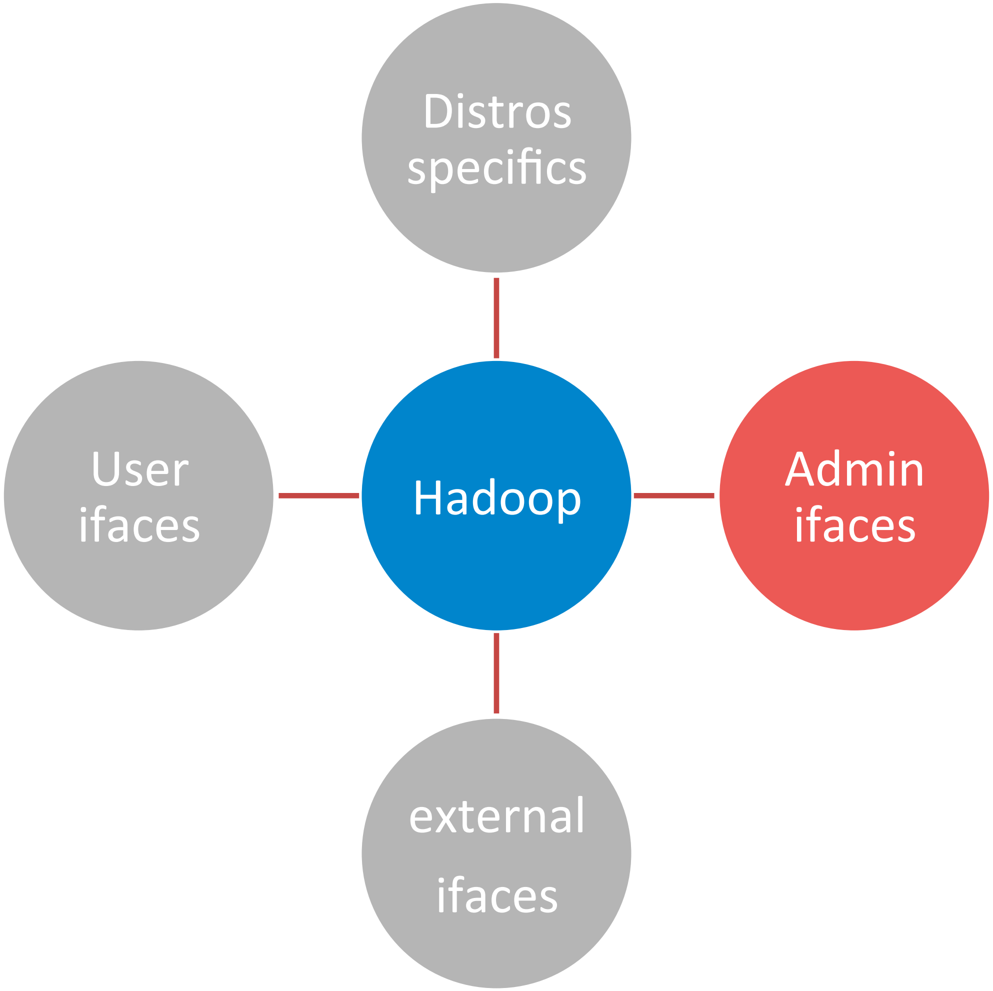
Using the URLconf defined in `desktop.urls`, Django tried these URL p

```
 1. ^about/static/(?P<path>.*)$
 2. ^beeswax/static/(?P<path>.*)$
 3. ^filebrowser/static/(?P<path>.*)$
 4. ^hcatalog/static/(?P<path>.*)$
 5. ^help/static/(?P<path>.*)$
 6. ^jobbrowser/static/(?P<path>.*)$
 7. ^jobsub/static/(?P<path>.*)$
 8. ^oozie/static/(?P<path>.*)$
 9. ^pig/static/(?P<path>.*)$
10. ^shell/static/(?P<path>.*)$
11. ^useradmin/static/(?P<path>.*)$
12. ^static/(?P<path>.*)$
13. ^(?P<path>favicon.ico)$
14. ^accounts/login/$
15. ^accounts/logout/$
16. ^logs$
17. ^dump_config$
18. ^download_logs$
19. ^bootstrap.js$
20. ^profile$
21. ^prefs/(?P<key>\w+)?$
22. ^status_bar/?$
23. ^admin/
24. ^debug/threads$
25. ^debug/who_am_i$
26. ^debug/check_config$
27. ^debug/check_config_ajax$
28. ^log_frontend_event$
29. ^jasmine
30. ^$
31. ^about/
32. ^beeswax/
33. ^filebrowser/
34. ^hcatalog/
35. ^help/
36. ^jobbrowser/
37. ^jobsub/
38. ^oozie/
39. ^pig/
```

ADMIN INTERFACES

for admins and maintenance

Admin interfaces



Admin interfaces

Apache Ambari

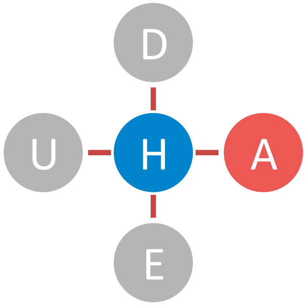
- Provisioning, monitoring

Apache Ranger

- Security: authorization, authentication, auditing, data encryption, administration

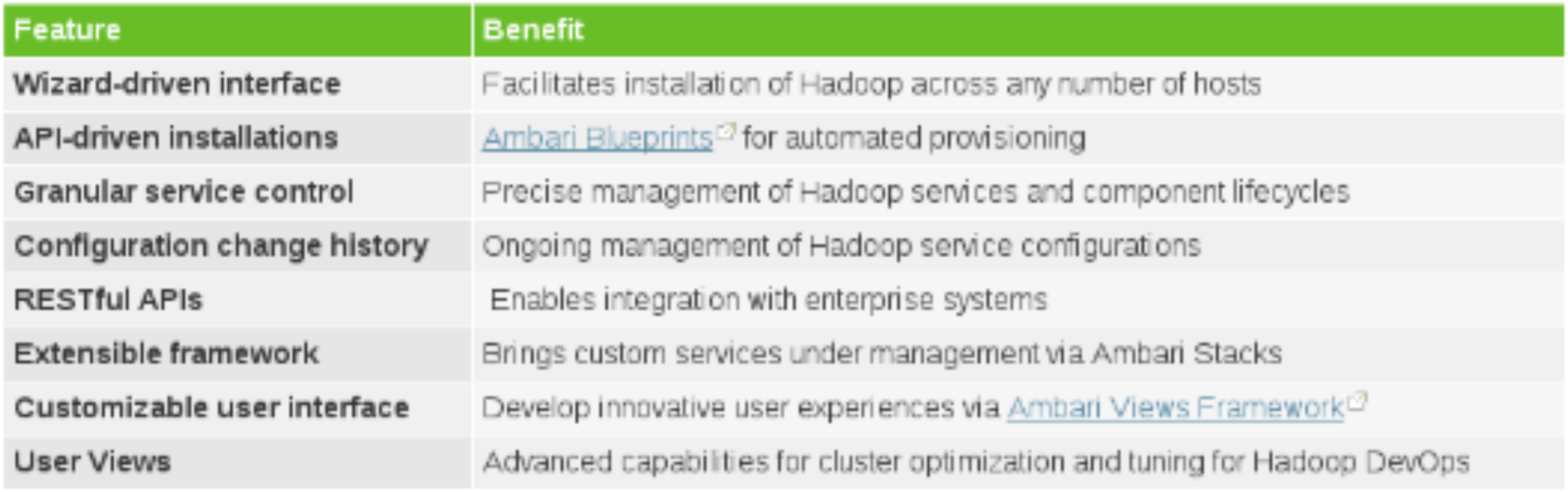
Other

- Knox, Cloudbreak, Zookeeper, Falcon, Atlas, Sqoop, Flume, Kafka



Apache Ambari

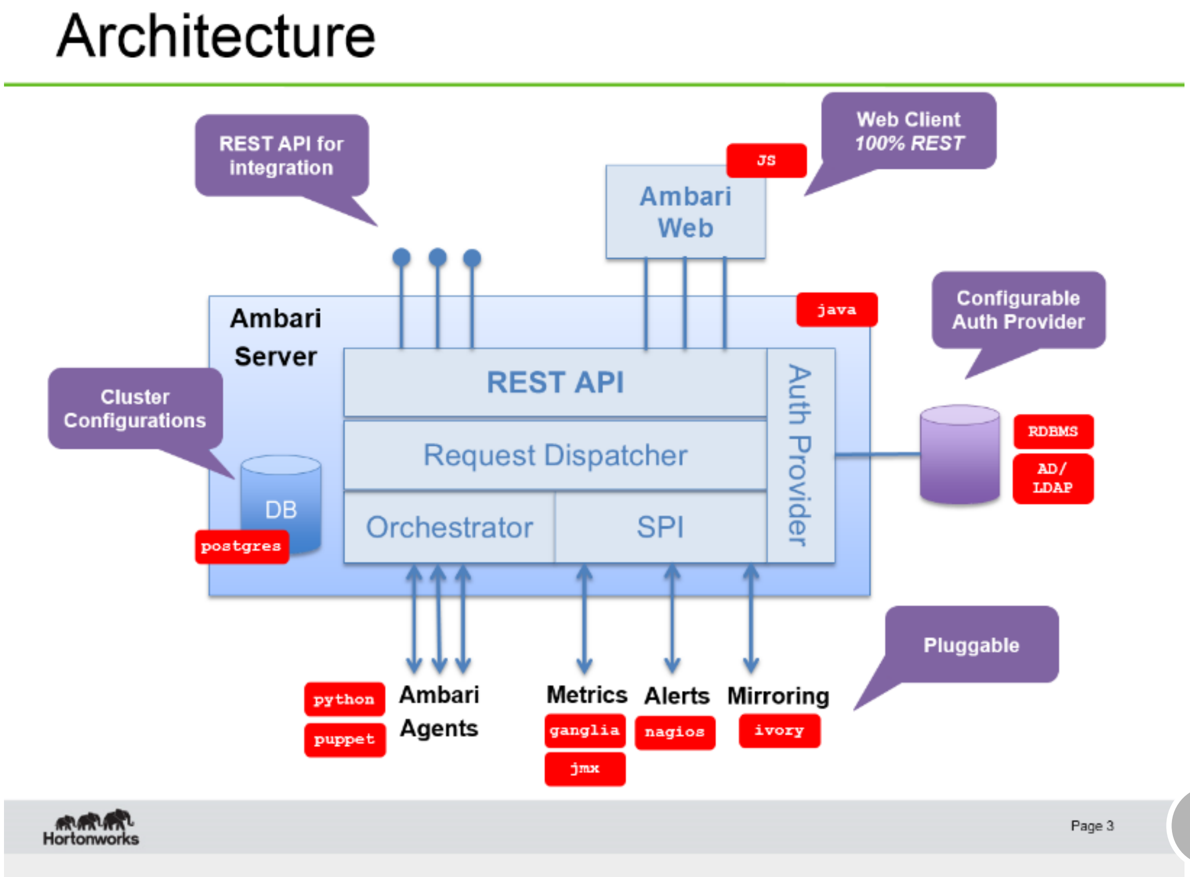
Trochę o Ambari



Feature	Benefit
Wizard-driven interface	Facilitates installation of Hadoop across any number of hosts
API-driven installations	Ambari Blueprints for automated provisioning
Granular service control	Precise management of Hadoop services and component lifecycles
Configuration change history	Ongoing management of Hadoop service configurations
RESTful APIs	Enables integration with enterprise systems
Extensible framework	Brings custom services under management via Ambari Stacks
Customizable user interface	Develop innovative user experiences via Ambari Views Framework
User Views	Advanced capabilities for cluster optimization and tuning for Hadoop DevOps

http://www.slideshare.net/hortonworks/ambari-using-a-local-repository?next_slideshow=1

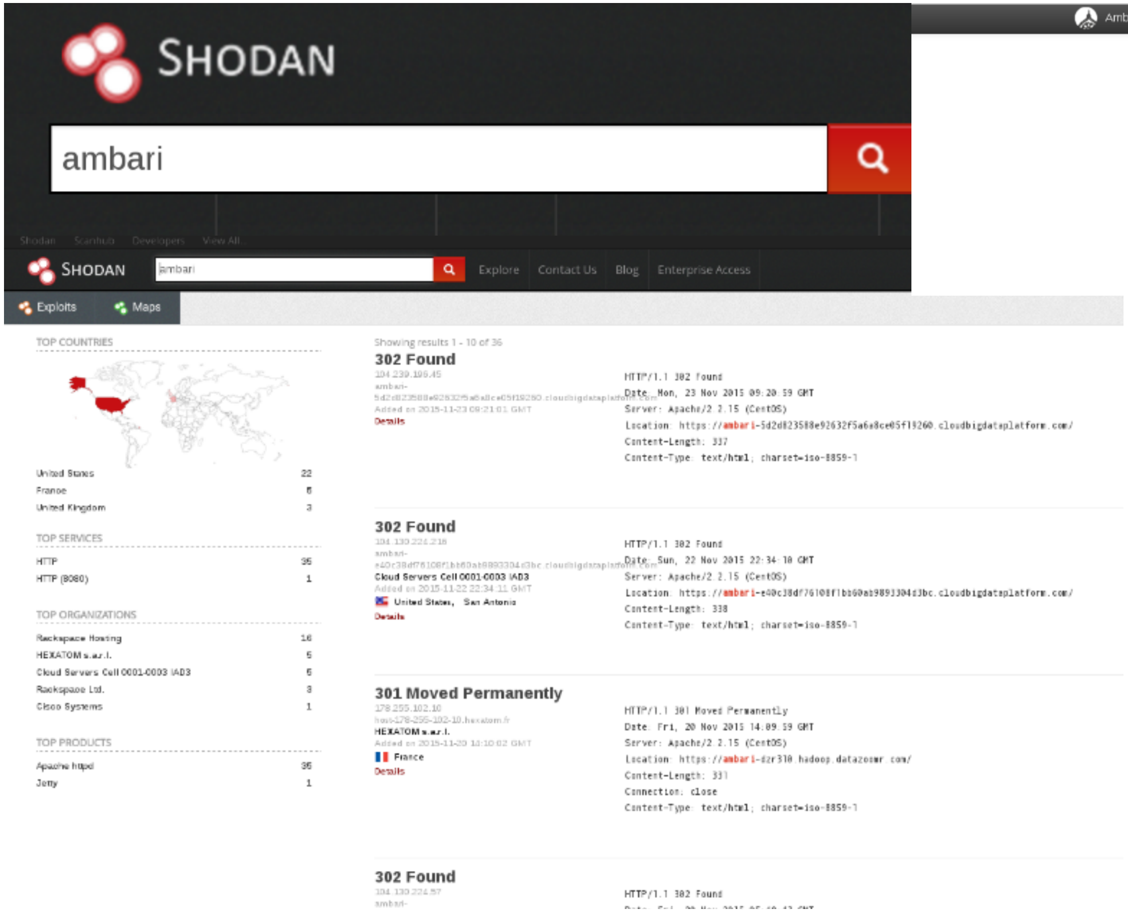
Apache Ambari



Architecture

http://www.slideshare.net/hortonworks/ambari-using-a-local-repository?next_slideshow=1

Is Ambari an internal interface?



TOP COUNTRIES

United States	22
France	6
United Kingdom	3

TOP SERVICES

HTTP	35
HTTP (8080)	1

TOP ORGANIZATIONS

Rackspace Hosting	16
HEXATOM s.a.r.l.	5
Cloud Servers Cell 0001-0003 IAD3	5
Rackspace Ltd.	3
Cisco Systems	1

TOP PRODUCTS

Apache httpd	35
Jetty	1

Showing results 1 - 10 of 36

302 Found
104.239.196.45
ambari-5d2d823588e92632f5a6a8ce05f19260.cloudbigdataplatform.com
Added on 2015-11-23 09:21:01 GMT
Details

```
HTTP/1.1 302 Found
Date: Mon, 23 Nov 2015 09:20:59 GMT
Server: Apache/2.2.15 (CentOS)
Location: https://ambari-5d2d823588e92632f5a6a8ce05f19260.cloudbigdataplatform.com/
Content-Length: 337
Content-Type: text/html; charset=iso-8859-1
```

302 Found
104.130.224.218
ambari-e40c38df76108f1bb60ab9893304d3bc.cloudbigdataplatform.com
Cloud Servers Cell 0001-0003 IAD3
Added on 2015-11-22 22:34:11 GMT
United States, San Antonio
Details

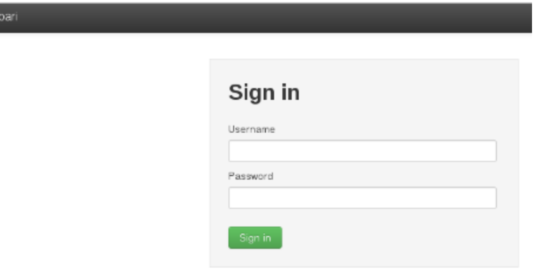
```
HTTP/1.1 302 Found
Date: Sun, 22 Nov 2015 22:34:10 GMT
Server: Apache/2.2.15 (CentOS)
Location: https://ambari-e40c38df76108f1bb60ab9893304d3bc.cloudbigdataplatform.com/
Content-Length: 338
Content-Type: text/html; charset=iso-8859-1
```

301 Moved Permanently
178.255.102.10
host-178-255-102-10.hexatom.fr
HEXATOM s.a.r.l.
Added on 2015-11-20 14:10:02 GMT
France
Details

```
HTTP/1.1 301 Moved Permanently
Date: Fri, 20 Nov 2015 14:09:59 GMT
Server: Apache/2.2.15 (CentOS)
Location: https://ambari-dzr310.hadoop.datazoomr.com/
Content-Length: 331
Connection: close
Content-Type: text/html; charset=iso-8859-1
```

302 Found
104.130.224.57
ambari-

```
HTTP/1.1 302 Found
```

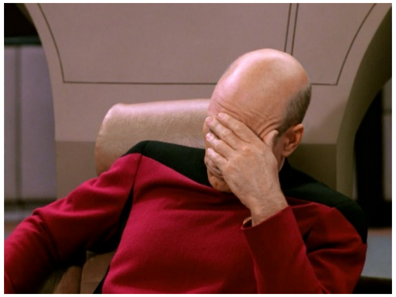


Sign in

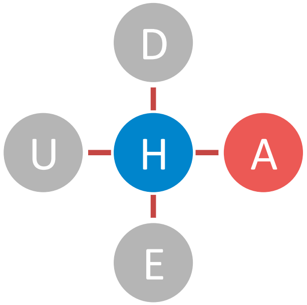
Username

Password

Sign in

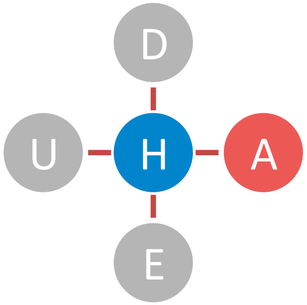


http://knowyourmeme.com/memes/facepalm



Apache Ambari

- Standard users can sign into Ambari (WHY?)
- Low hanging fruits: directory listing by default, no cookie flags, no CSRF protection
- Interesting proxy script ->



D

U H A

E

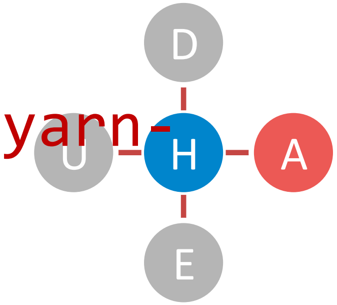
Apache Ambari REST API proxy

Standard request:

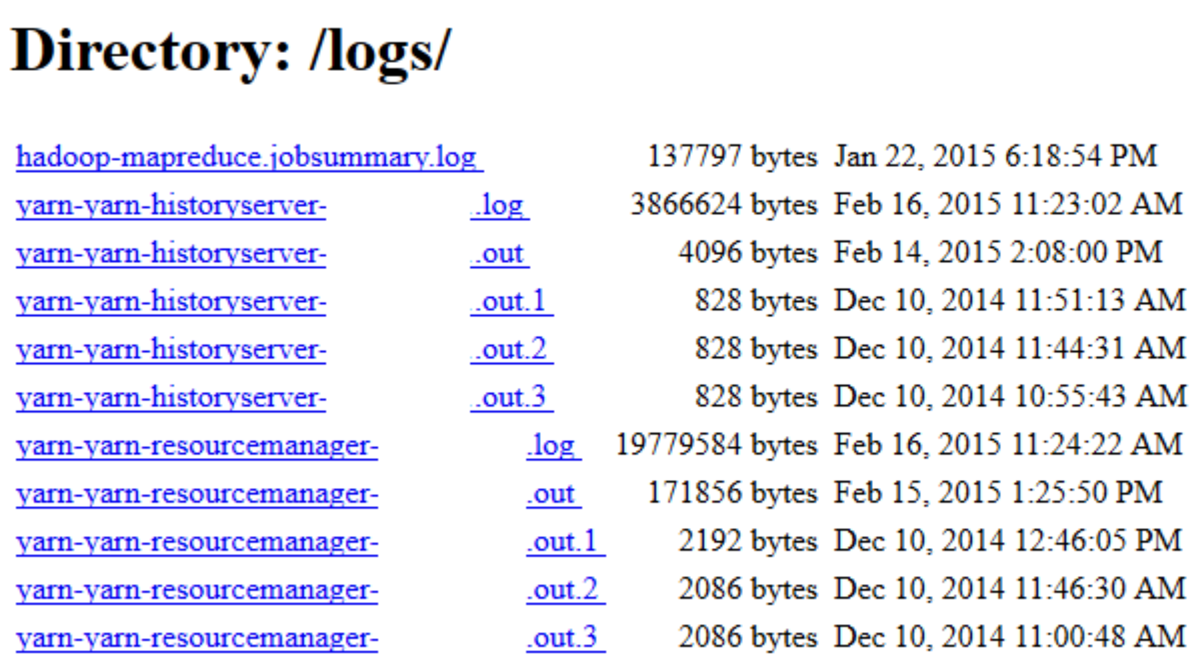
```
/proxy?url=http://XXXXXXXXX:8188/ws/v1/
timeline/HIVE_QUERY_ID?
limit=1&secondaryFilter=tez:true&_=142418001
6625
```

Tampered request (logs accessible only from DMZ):

```
/proxy?url=http://google.com
/proxy?url=http://XXXXXXX:8088/logs
/proxy?url=http://XXXXXXX:8088/logs/yarn-
yarn-resourcemanager-XXXXXXX.log
```



Apache Ambari Server Side Request Forgery

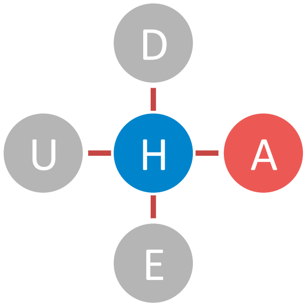


Directory: /logs/

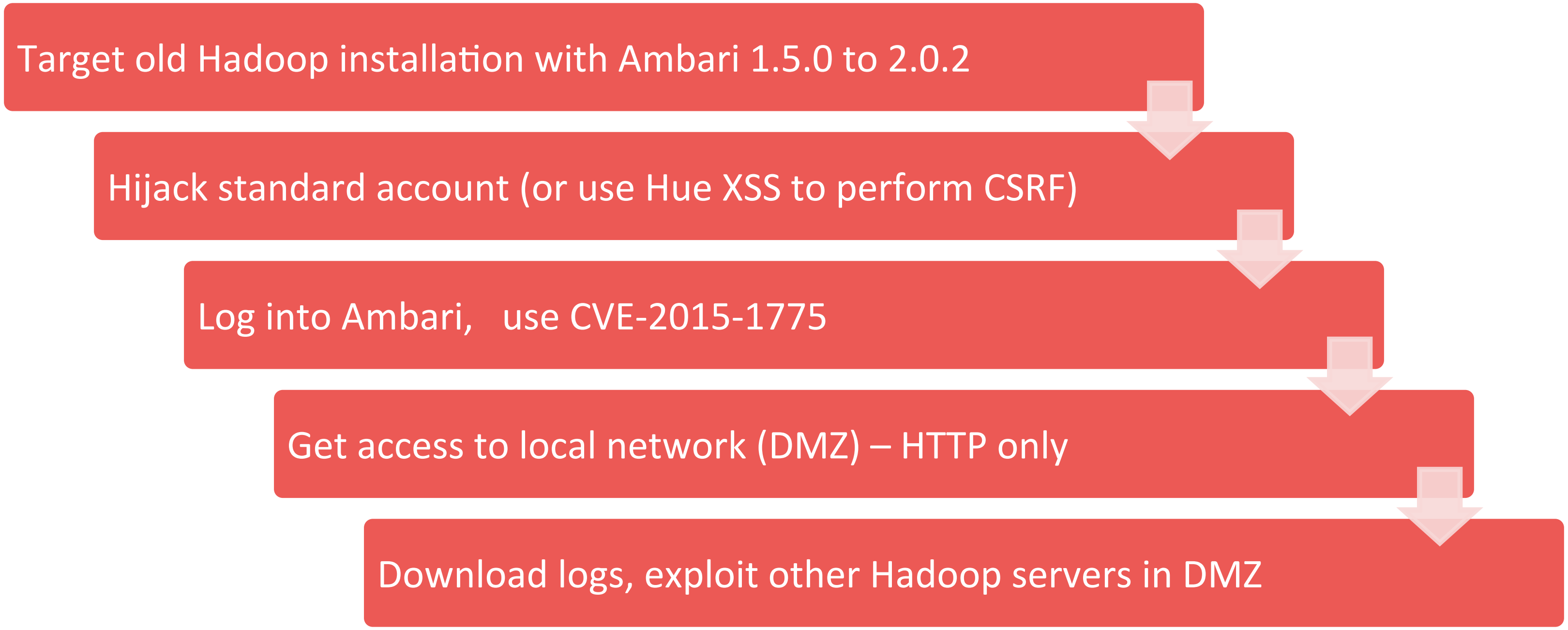
```
hadoop-mapreduce.jobsummary.log            137797 bytes  Jan 22, 2015 6:18:54 PM
yarn-yarn-historyserver-          .log    3866624 bytes  Feb 16, 2015 11:23:02 AM
yarn-yarn-historyserver-          .out       4096 bytes  Feb 14, 2015 2:08:00 PM
yarn-yarn-historyserver-          .out.1      828 bytes  Dec 10, 2014 11:51:13 AM
yarn-yarn-historyserver-          .out.2      828 bytes  Dec 10, 2014 11:44:31 AM
yarn-yarn-historyserver-          .out.3      828 bytes  Dec 10, 2014 10:55:43 AM
yarn-yarn-resourcemanager-          .log 19779584 bytes  Feb 16, 2015 11:24:22 AM
yarn-yarn-resourcemanager-          .out   171856 bytes  Feb 15, 2015 1:25:50 PM
yarn-yarn-resourcemanager-          .out.1   2192 bytes  Dec 10, 2014 12:46:05 PM
yarn-yarn-resourcemanager-          .out.2   2086 bytes  Dec 10, 2014 11:46:30 AM
yarn-yarn-resourcemanager-          .out.3   2086 bytes  Dec 10, 2014 11:00:48 AM
```

PWNED!

CVE-2015-1775



Apache Ambari attack scenario



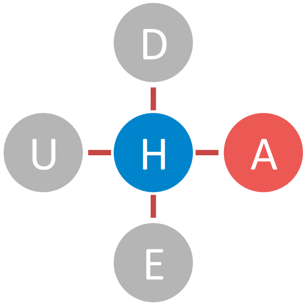
Target old Hadoop installation with Ambari 1.5.0 to 2.0.2

Hijack standard account (or use Hue XSS to perform CSRF)

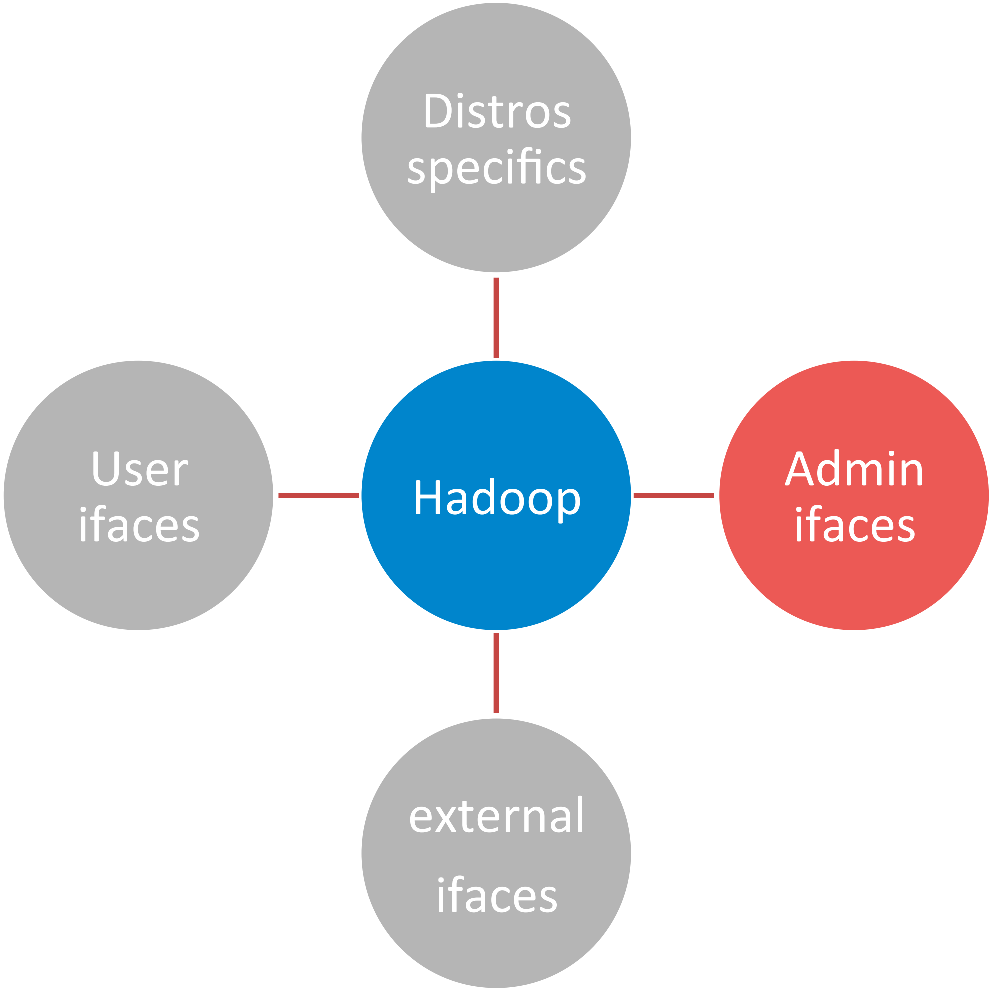
Log into Ambari,   use CVE-2015-1775

Get access to local network (DMZ) – HTTP only

Download logs, exploit other Hadoop servers in DMZ



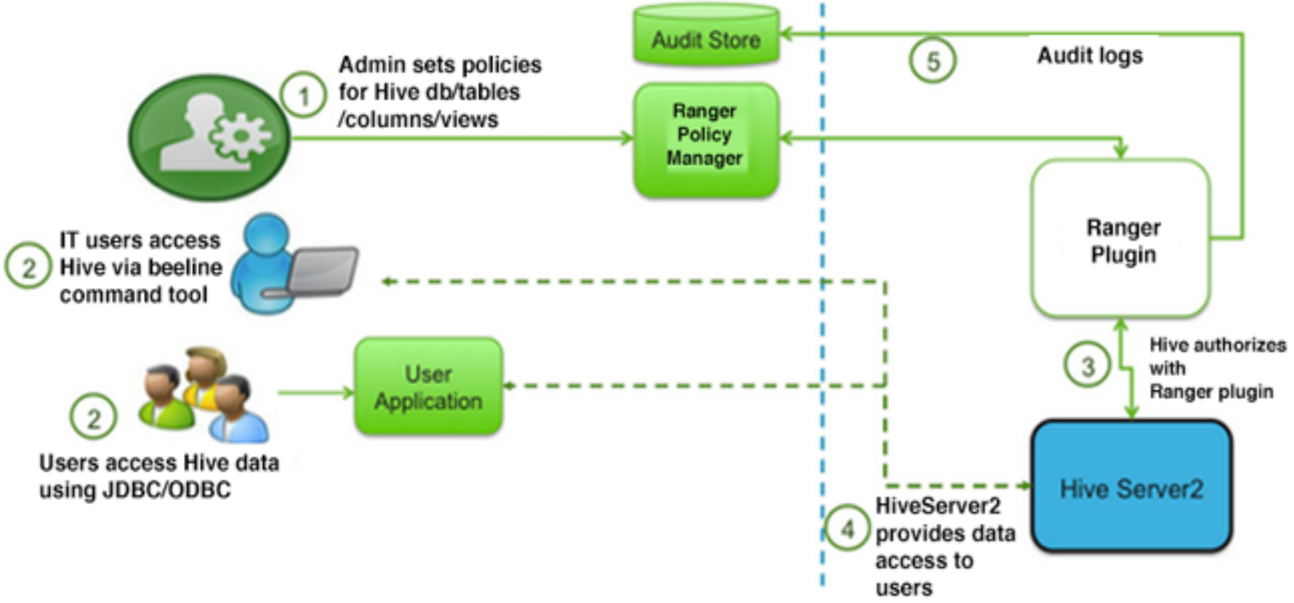
Admin interfaces



Apache Ranger overview

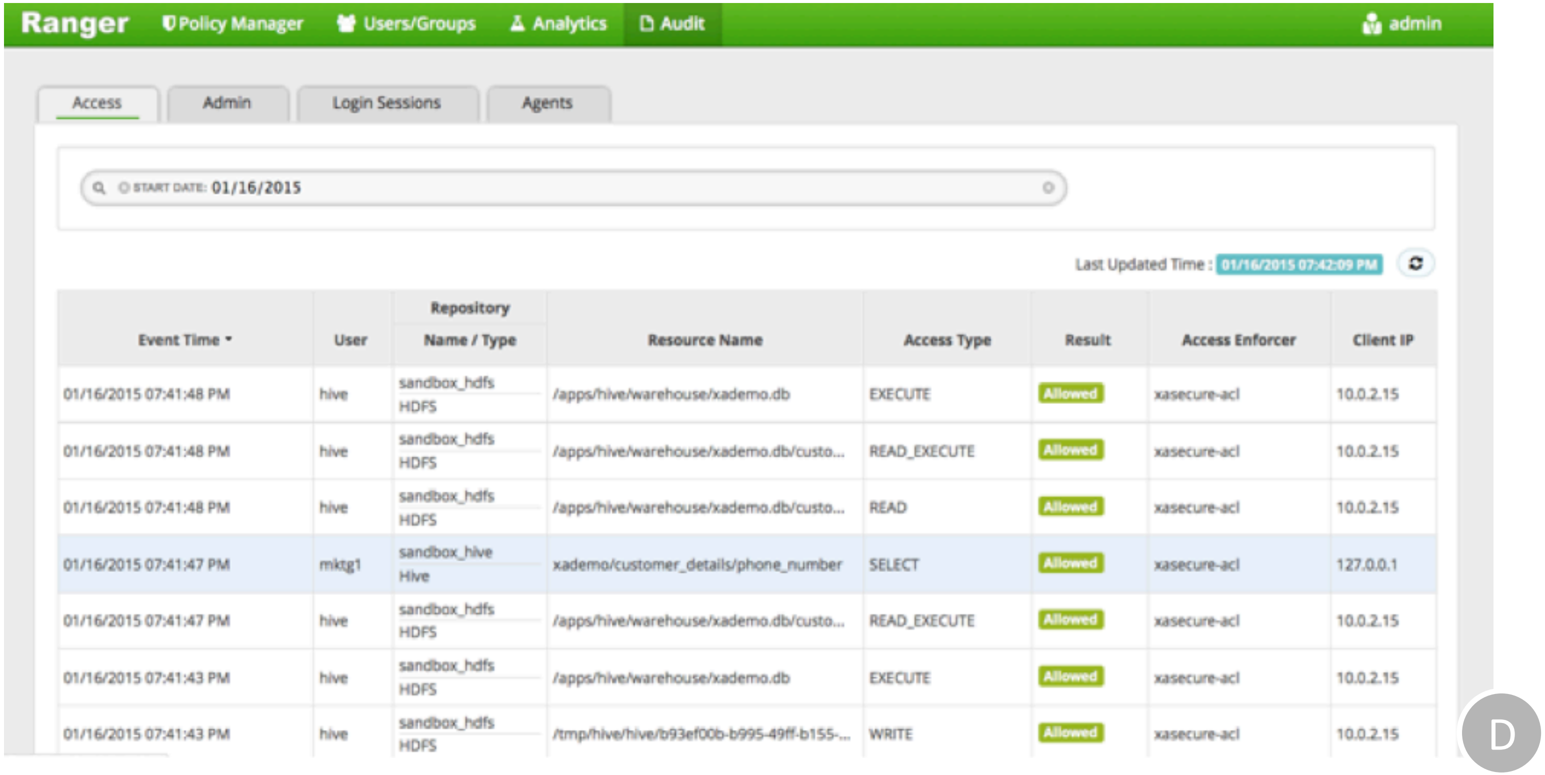
Previously: Apache Argus, XA-Secure

Provides central administration for policies, users/groups, analytics and audit data.



http://docs.hortonworks.com/HDPDocuments/HDP2/HDP-2.3.2/bk_Sys_Admin_Guides/content/ref-746ce51a-9bdc-4fef-85a6-69564089a8a6.1.html

Apache Ranger overview

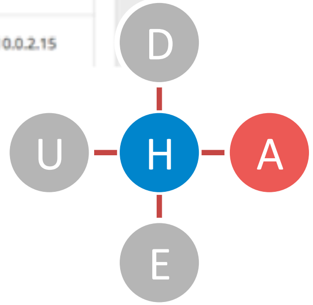


START DATE: 01/16/2015

Last Updated Time : 01/16/2015 07:42:09 PM

Event Time	User	Repository Name / Type	Resource Name	Access Type	Result	Access Enforcer	Client IP
01/16/2015 07:41:48 PM	hive	sandbox_hdfs HDFS	/apps/hive/warehouse/xademo.db	EXECUTE	Allowed	xasecure-acl	10.0.2.15
01/16/2015 07:41:48 PM	hive	sandbox_hdfs HDFS	/apps/hive/warehouse/xademo.db/custo...	READ_EXECUTE	Allowed	xasecure-acl	10.0.2.15
01/16/2015 07:41:48 PM	hive	sandbox_hdfs HDFS	/apps/hive/warehouse/xademo.db/custo...	READ	Allowed	xasecure-acl	10.0.2.15
01/16/2015 07:41:47 PM	mktg1	sandbox_hive Hive	xademo/customer_details/phone_number	SELECT	Allowed	xasecure-acl	127.0.0.1
01/16/2015 07:41:47 PM	hive	sandbox_hdfs HDFS	/apps/hive/warehouse/xademo.db/custo...	READ_EXECUTE	Allowed	xasecure-acl	10.0.2.15
01/16/2015 07:41:43 PM	hive	sandbox_hdfs HDFS	/apps/hive/warehouse/xademo.db	EXECUTE	Allowed	xasecure-acl	10.0.2.15
01/16/2015 07:41:43 PM	hive	sandbox_hdfs HDFS	/tmp/hive/hive/b93ef00b-b995-49ff-b155-...	WRITE	Allowed	xasecure-acl	10.0.2.15

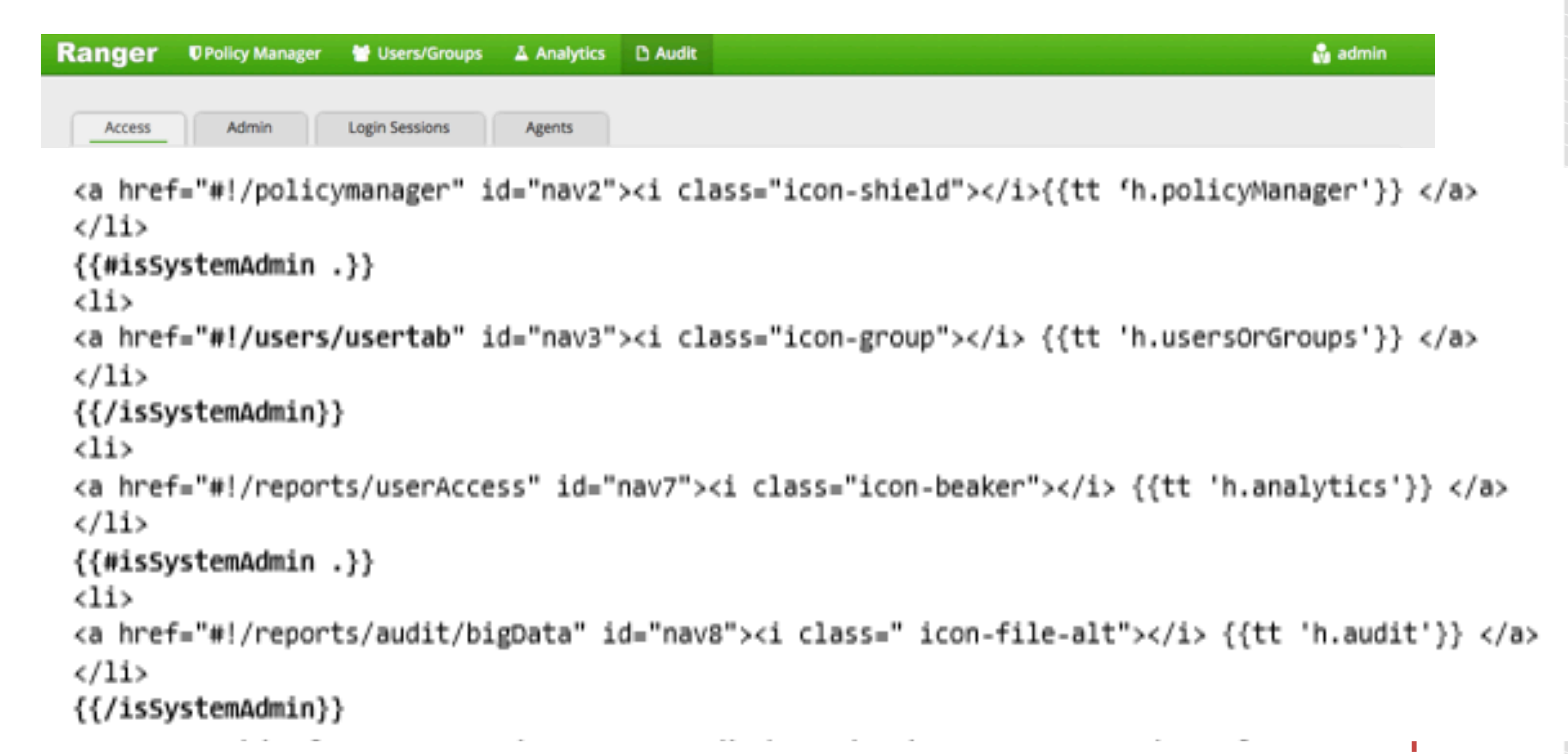
http://hortonworks.com/blog/best-practices-for-hive-authorization-using-apache-ranger-in-hdp-2-2/



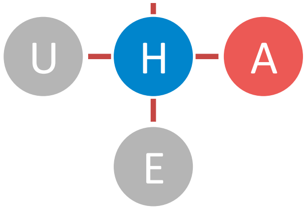
Apache Ranger

- Low hanging fruits: no HTTP hardening, SlowHTTP DoS
- Standard users can log into Ranger but have no permissions
- Interesting function level access control ->

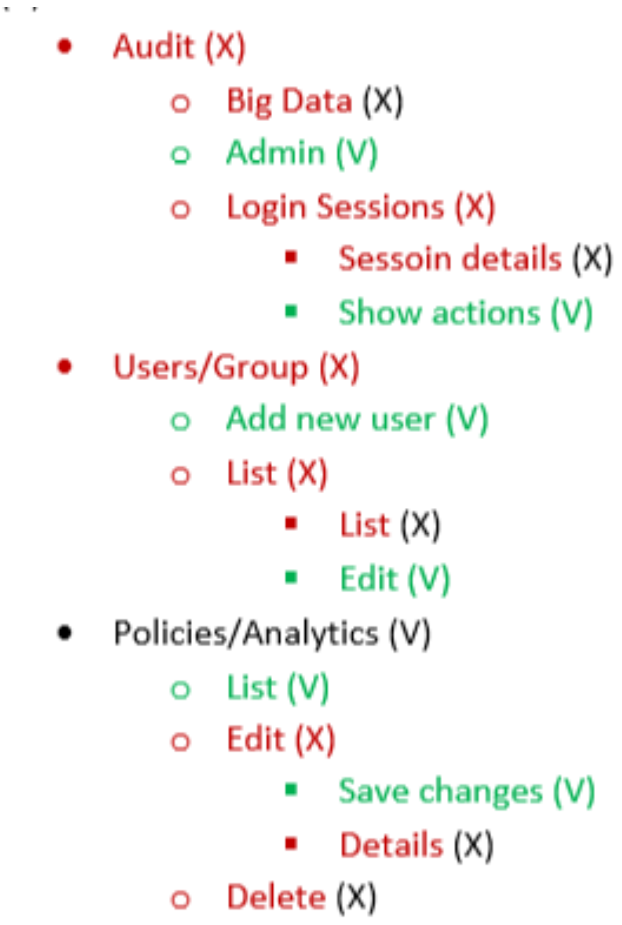
Apache Ranger vulnerabilities



```
<a href="#!/policymanager" id="nav2"><i class="icon-shield"></i>{{tt 'h.policyManager'}} </a>
</li>
{{#isSystemAdmin .}}
<li>
<a href="#!/users/usertab" id="nav3"><i class="icon-group"></i> {{tt 'h.usersOrGroups'}} </a>
</li>
{{/isSystemAdmin}}
<li>
<a href="#!/reports/userAccess" id="nav7"><i class="icon-beaker"></i> {{tt 'h.analytics'}} </a>
</li>
{{#isSystemAdmin .}}
<li>
<a href="#!/reports/audit/bigData" id="nav8"><i class=" icon-file-alt"></i> {{tt 'h.audit'}} </a>
</li>
{{/isSystemAdmin}}
```

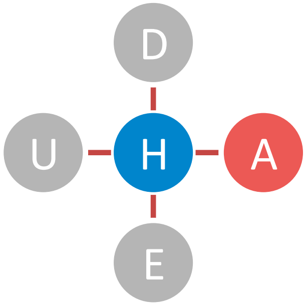


Missing function level access control

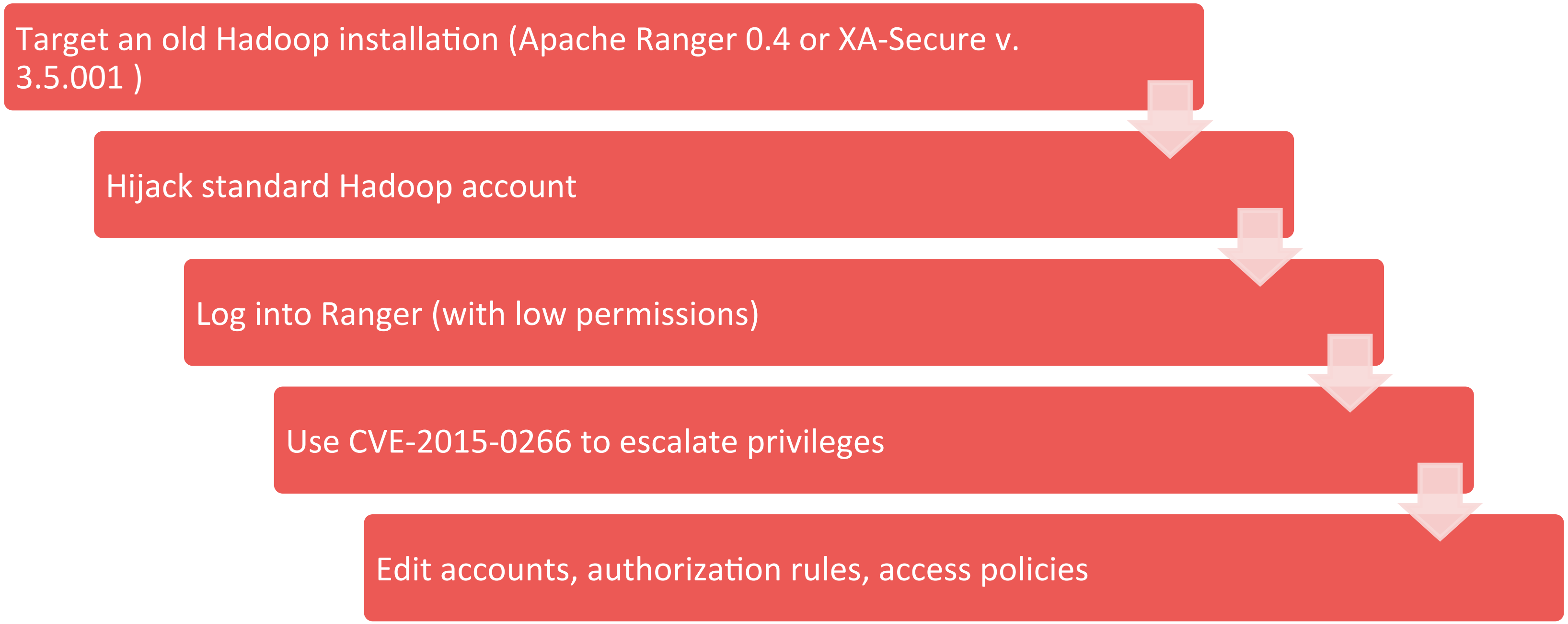


- Audit (X)
    - Big Data (X)
    - Admin (V)
    - Login Sessions (X)
        - Sessoin details (X)
        - Show actions (V)
- Users/Group (X)
    - Add new user (V)
    - List (X)
        - List (X)
        - Edit (V)
- Policies/Analytics (V)
    - List (V)
    - Edit (X)
        - Save changes (V)
        - Details (X)
    - Delete (X)

CVE-2015-0266



Apache Ranger attack scenario



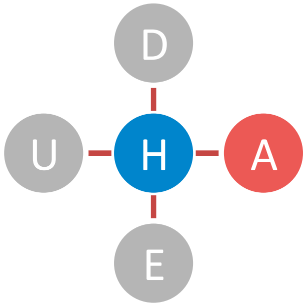
Target an old Hadoop installation (Apache Ranger 0.4 or XA-Secure v. 3.5.001)

Hijack standard Hadoop account

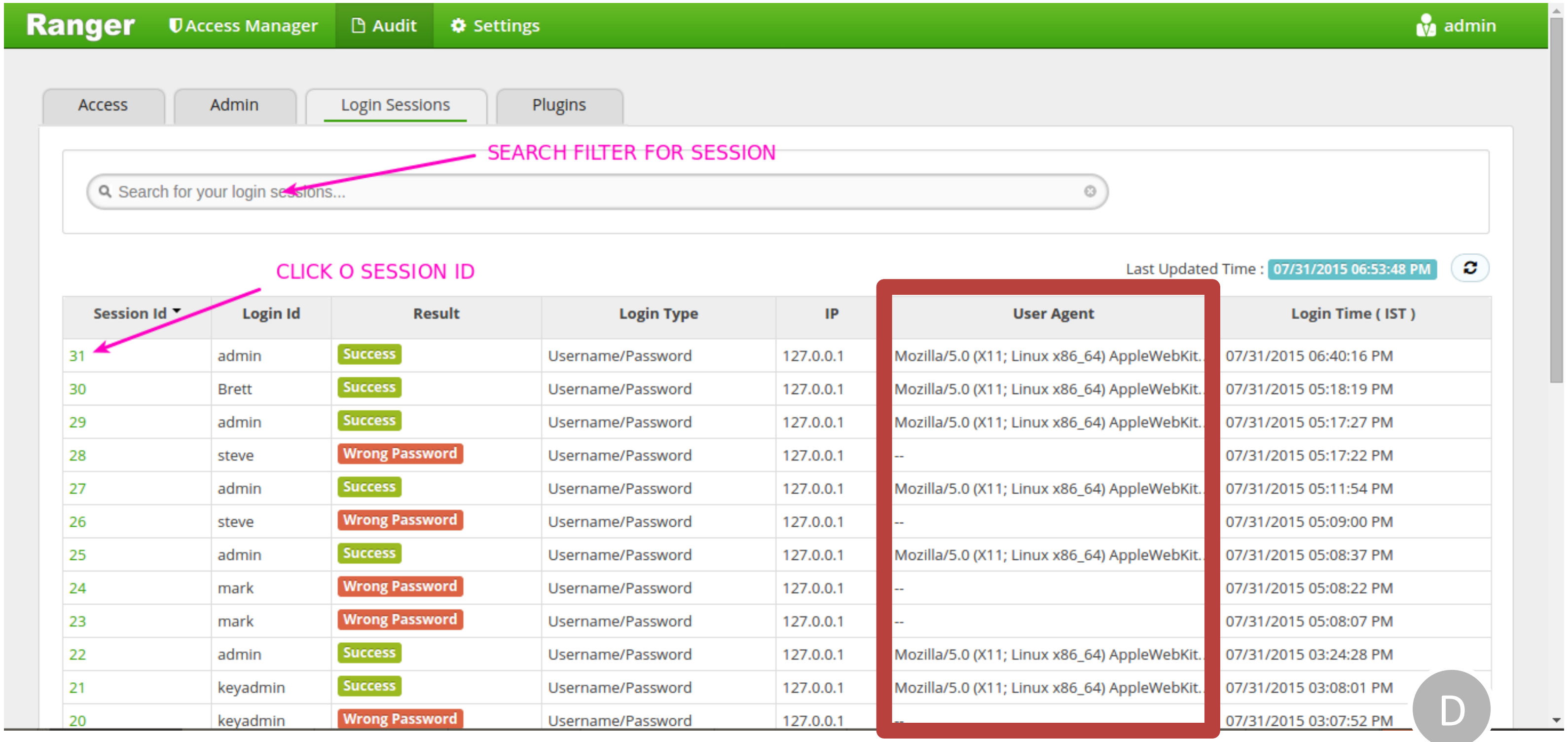
Log into Ranger (with low permissions)

Use CVE-2015-0266 to escalate privileges

Edit accounts, authorization rules, access policies



Apache Ranger vulnerabilities



SEARCH FILTER FOR SESSION

CLICK O SESSION ID

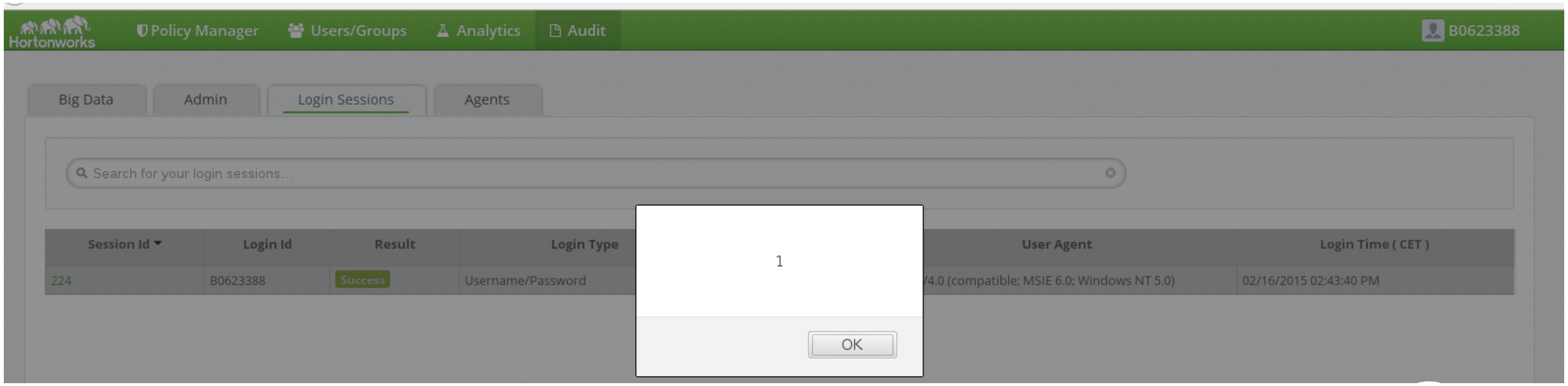
Last Updated Time : 07/31/2015 06:53:48 PM

Session Id	Login Id	Result	Login Type	IP	User Agent	Login Time (IST)
31	admin	Success	Username/Password	127.0.0.1	Mozilla/5.0 (X11; Linux x86_64) AppleWebKit...	07/31/2015 06:40:16 PM
30	Brett	Success	Username/Password	127.0.0.1	Mozilla/5.0 (X11; Linux x86_64) AppleWebKit...	07/31/2015 05:18:19 PM
29	admin	Success	Username/Password	127.0.0.1	Mozilla/5.0 (X11; Linux x86_64) AppleWebKit...	07/31/2015 05:17:27 PM
28	steve	Wrong Password	Username/Password	127.0.0.1	--	07/31/2015 05:17:22 PM
27	admin	Success	Username/Password	127.0.0.1	Mozilla/5.0 (X11; Linux x86_64) AppleWebKit...	07/31/2015 05:11:54 PM
26	steve	Wrong Password	Username/Password	127.0.0.1	--	07/31/2015 05:09:00 PM
25	admin	Success	Username/Password	127.0.0.1	Mozilla/5.0 (X11; Linux x86_64) AppleWebKit...	07/31/2015 05:08:37 PM
24	mark	Wrong Password	Username/Password	127.0.0.1	--	07/31/2015 05:08:22 PM
23	mark	Wrong Password	Username/Password	127.0.0.1	--	07/31/2015 05:08:07 PM
22	admin	Success	Username/Password	127.0.0.1	Mozilla/5.0 (X11; Linux x86_64) AppleWebKit...	07/31/2015 03:24:28 PM
21	keyadmin	Success	Username/Password	127.0.0.1	Mozilla/5.0 (X11; Linux x86_64) AppleWebKit...	07/31/2015 03:08:01 PM
20	keyadmin	Wrong Password	Username/Password	127.0.0.1	--	07/31/2015 03:07:52 PM

https://cwiki.apache.org/confluence/display/RANGER/Apache+Ranger+0.5+-+User+Guide

Apache Ranger XSS through UserAgent

```
User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.0) <script>alert(1);</script>
```

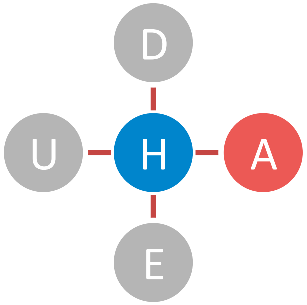


Session Id	Login Id	Result	Login Type	User Agent	Login Time (CET)
224	B0623388	Success	Username/Password	/4.0 (compatible; MSIE 6.0; Windows NT 5.0)	02/16/2015 02:43:40 PM

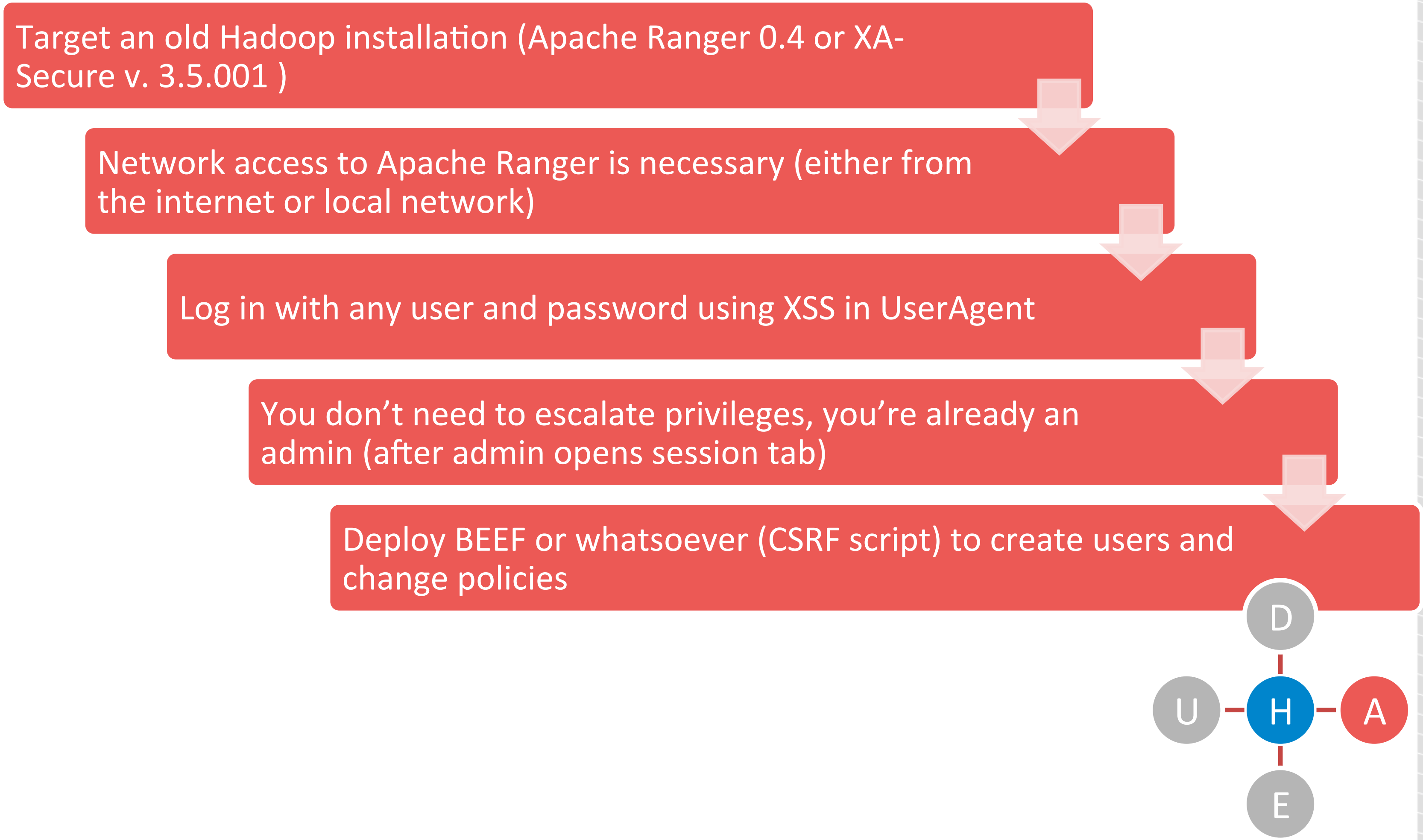
1

PWNED!

CVE-2015-0265



Apache Ranger attack scenario



Target an old Hadoop installation (Apache Ranger 0.4 or XA-Secure v. 3.5.001)

Network access to Apache Ranger is necessary (either from the internet or local network)

Log in with any user and password using XSS in UserAgent

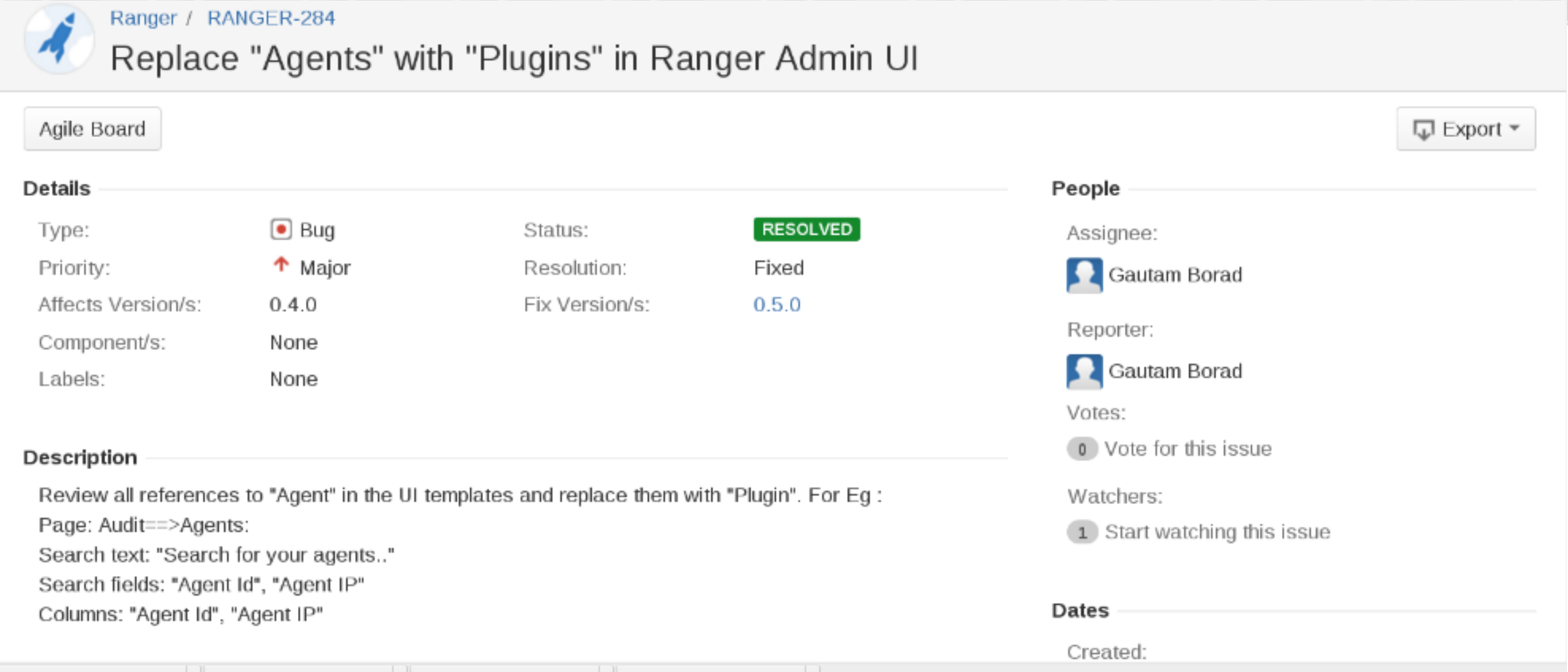
You don’t need to escalate privileges, you’re already an admin (after admin opens session tab)

Deploy BEEF or whatsoever (CSRF script) to create users and change policies

Apache Ranger patched

- Affected version: Apache Ranger v 0.4.0, XA Secure v. 3.5.001
- Both vulnerabilities patched in Ranger v 0.5.0
- For a while developers did a self-full-disclosure ->

RANGER-284 in public Jira now



Ranger / RANGER-284

Replace "Agents" with "Plugins" in Ranger Admin UI

Details

Type:	Bug	Status:	RESOLVED
Priority:	Major	Resolution:	Fixed
Affects Version/s:	0.4.0	Fix Version/s:	0.5.0
Component/s:	None		
Labels:	None		

Description

Review all references to "Agent" in the UI templates and replace them with "Plugin". For Eg :
Page: Audit==>Agents:
Search text: "Search for your agents.."
Search fields: "Agent Id", "Agent IP"
Columns: "Agent Id", "Agent IP"

People

Assignee:
Gautam Borad

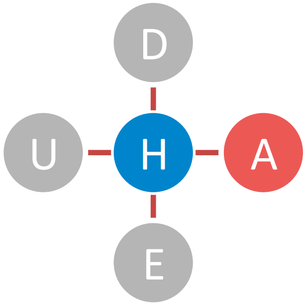
Reporter:
Gautam Borad

Votes:
0 Vote for this issue

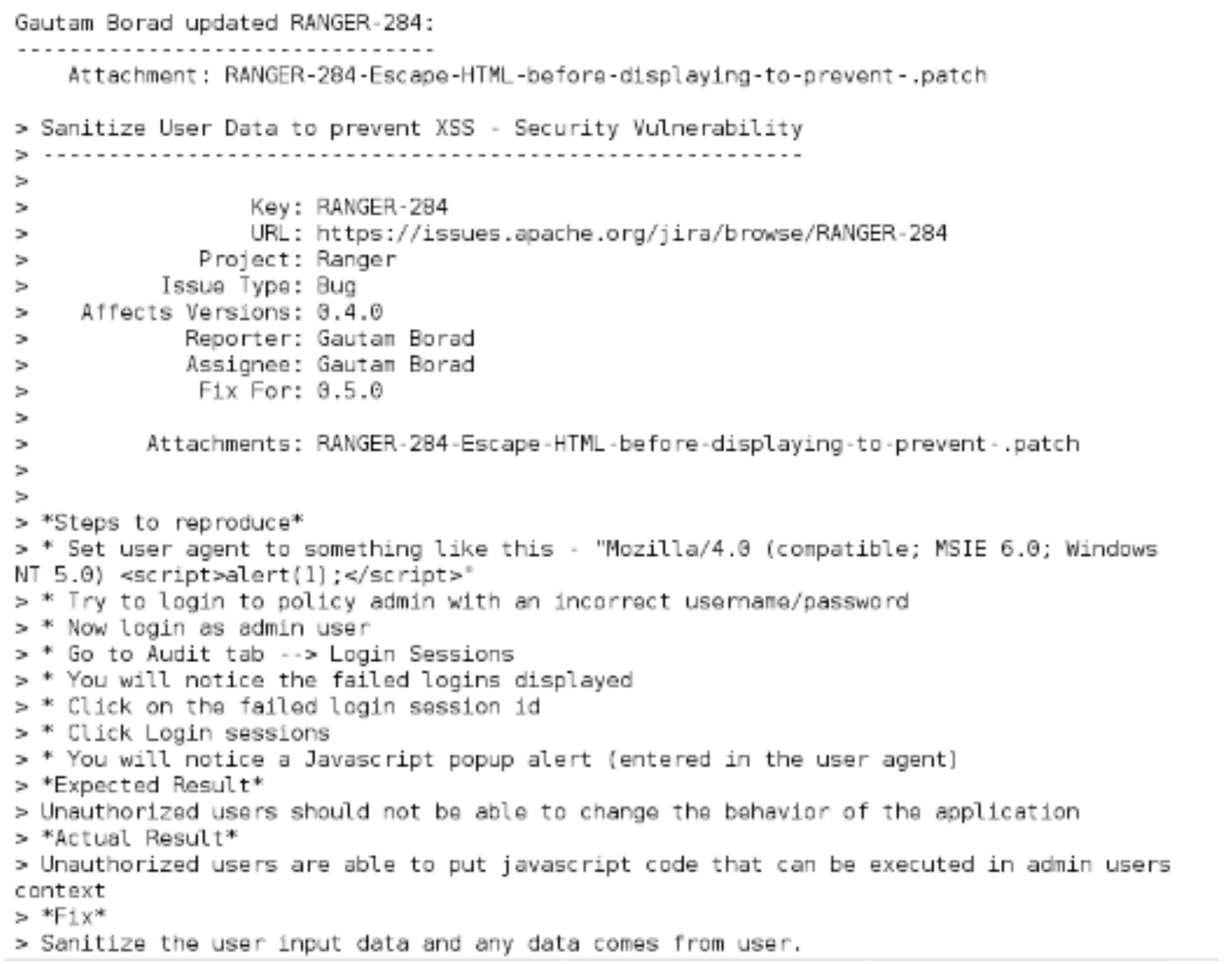
Watchers:
1 Start watching this issue

Dates

Created:



RANGER-284 shortly after vendor contact



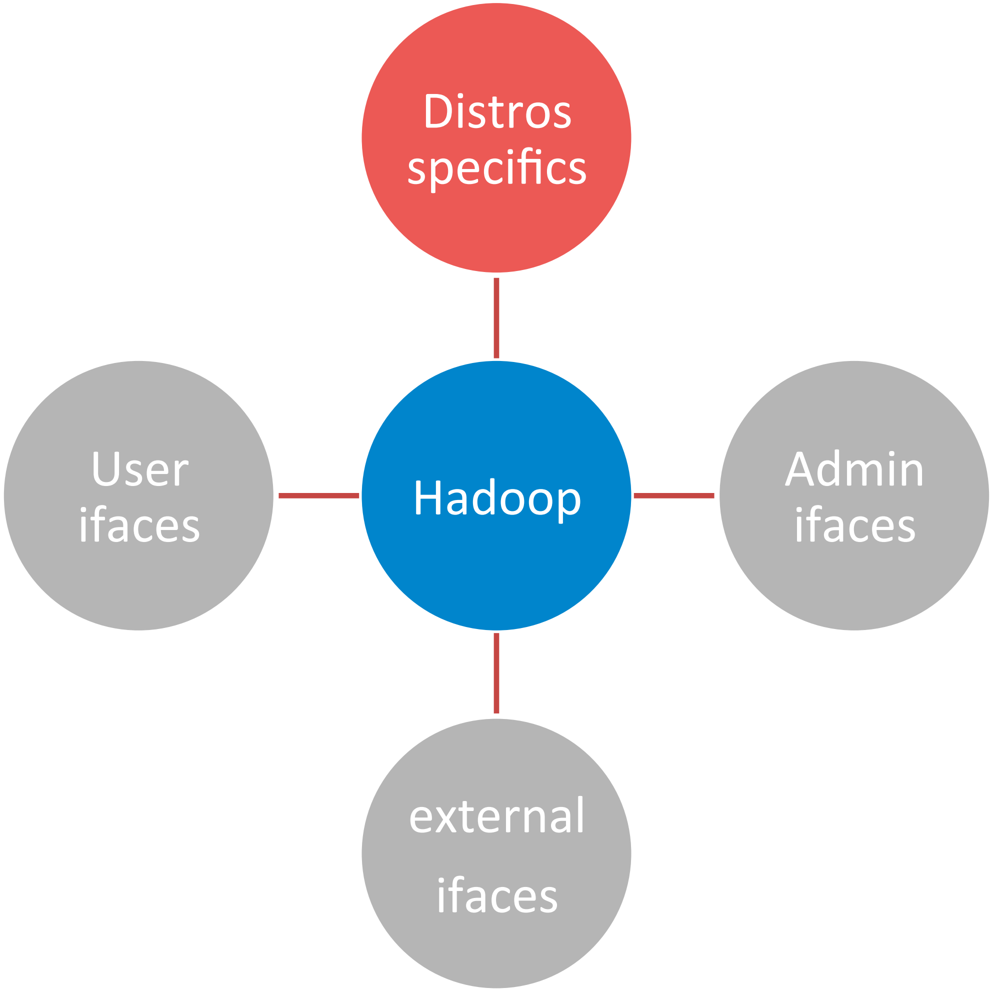
```
Gautam Borad updated RANGER-284:
--------------------------------
    Attachment: RANGER-284-Escape-HTML-before-displaying-to-prevent-.patch

> Sanitize User Data to prevent XSS - Security Vulnerability
> ----------------------------------------------------------
>
>                 Key: RANGER-284
>                 URL: https://issues.apache.org/jira/browse/RANGER-284
>             Project: Ranger
>          Issue Type: Bug
>    Affects Versions: 0.4.0
>            Reporter: Gautam Borad
>            Assignee: Gautam Borad
>             Fix For: 0.5.0
>
>         Attachments: RANGER-284-Escape-HTML-before-displaying-to-prevent-.patch
>
>
> *Steps to reproduce*
> * Set user agent to something like this - "Mozilla/4.0 (compatible; MSIE 6.0; Windows
NT 5.0) <script>alert(1);</script>"
> * Try to login to policy admin with an incorrect username/password
> * Now login as admin user
> * Go to Audit tab --> Login Sessions
> * You will notice the failed logins displayed
> * Click on the failed login session id
> * Click Login sessions
> * You will notice a Javascript popup alert (entered in the user agent)
> *Expected Result*
> Unauthorized users should not be able to change the behavior of the application
> *Actual Result*
> Unauthorized users are able to put javascript code that can be executed in admin users
context
> *Fix*
> Sanitize the user input data and any data comes from user.
```

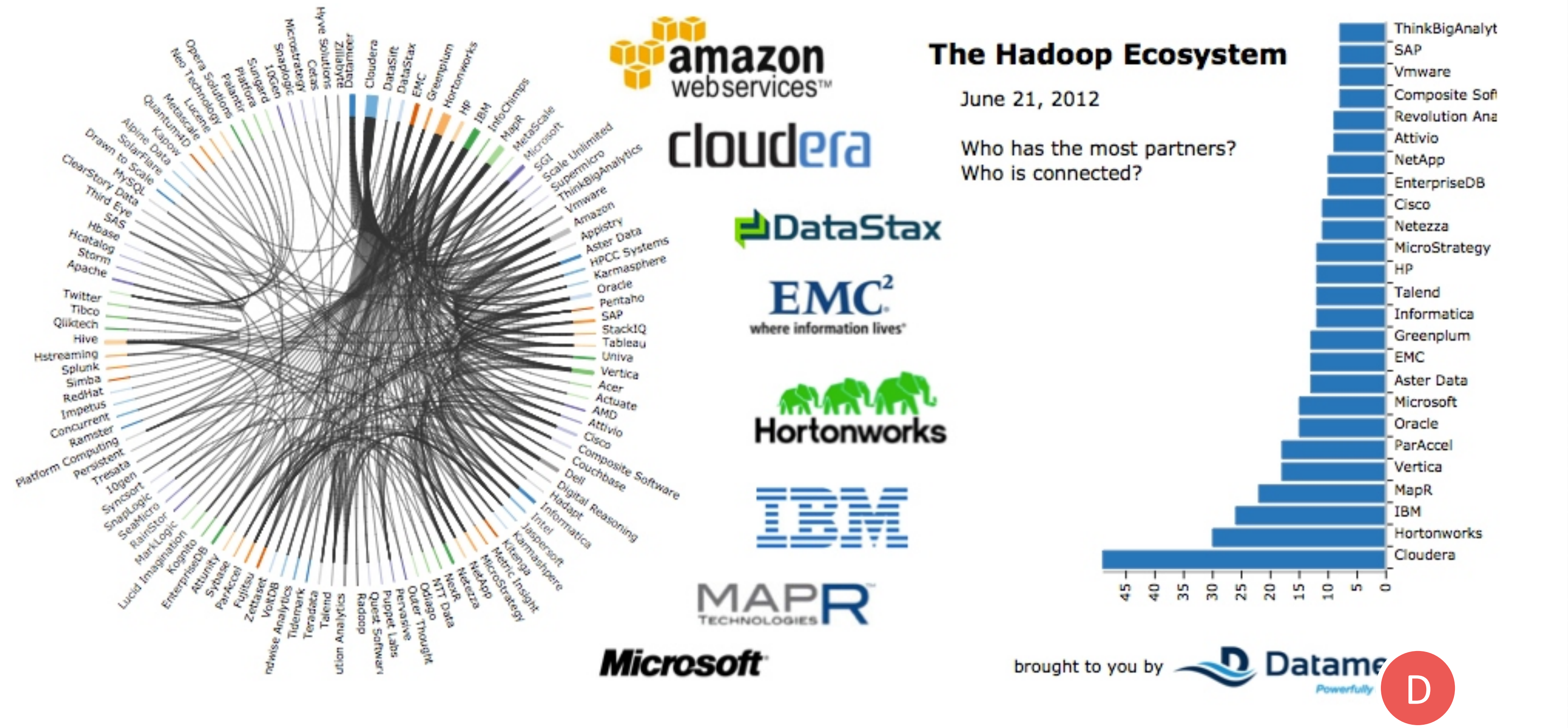
DISTRIBUTIONS SPECIFICS

not in every environment

Distribution specifics



Distros



http://blog.cloudera.com/blog/2012/07/the-hadoop-ecosystem-visualized-in-datameer/

D
U H A
E

Basic distinction

cloud based

hosted locally

Distros

How long does it take to create a new distro version?

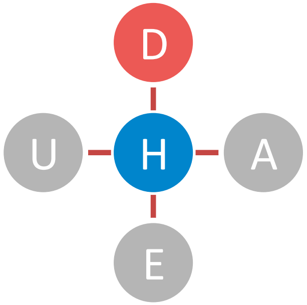
How many components are outdated at that time?

How long does it take to deploy a new distro at a company?

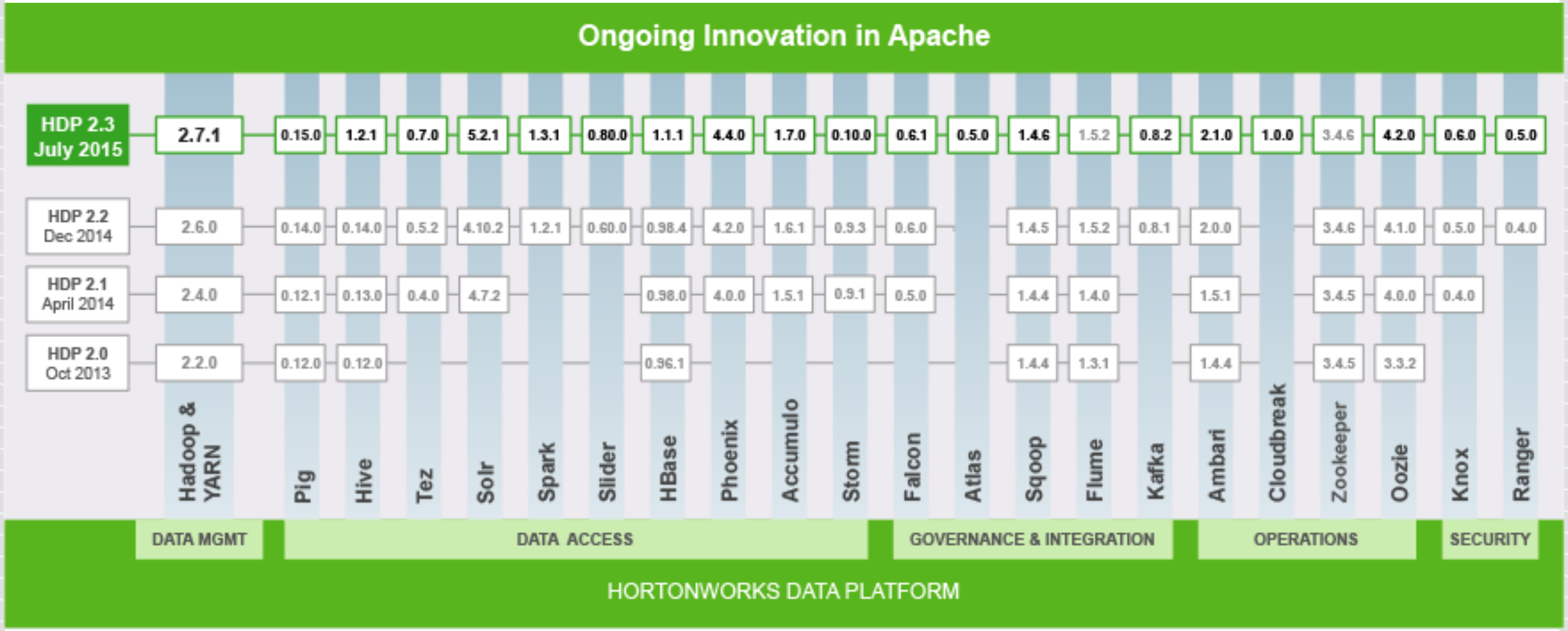
How many components are outdated at that time?

Most cases:

- MAJOR – ca. 1 year
- MINOR – ca. 3 months
- PATCH – ca. 1-2 months (differs much)



Hortonworks HDP components by version



Ongoing Innovation in Apache

	DATA MGMT	DATA ACCESS									
	Hadoop & YARN	Pig	Hive	Tez	Solr	Spark	Slider	HBase	Phoenix	Accumulo	Storm
HDP 2.3 July 2015	2.7.1	0.15.0	1.2.1	0.7.0	5.2.1	1.3.1	0.80.0	1.1.1	4.4.0	1.7.0	0.10.0
HDP 2.2 Dec 2014	2.6.0	0.14.0	0.14.0	0.5.2	4.10.2	1.2.1	0.60.0	0.98.4	4.2.0	1.6.1	0.9.3
HDP 2.1 April 2014	2.4.0	0.12.1	0.13.0	0.4.0	4.7.2			0.98.0	4.0.0	1.5.1	0.9.1
HDP 2.0 Oct 2013	2.2.0	0.12.0	0.12.0					0.96.1			

	GOVERNANCE & INTEGRATION					OPERATIONS				SECURITY	
	Falcon	Atlas	Sqoop	Flume	Kafka	Ambari	Cloudbreak	Zookeeper	Oozie	Knox	Ranger
HDP 2.3 July 2015	0.6.1	0.5.0	1.4.6	1.5.2	0.8.2	2.1.0	1.0.0	3.4.6	4.2.0	0.6.0	0.5.0
HDP 2.2 Dec 2014	0.6.0		1.4.5	1.5.2	0.8.1	2.0.0		3.4.6	4.1.0	0.5.0	0.4.0
HDP 2.1 April 2014	0.5.0		1.4.4	1.4.0		1.5.1		3.4.5	4.0.0	0.4.0	
HDP 2.0 Oct 2013			1.4.4	1.3.1		1.4.4		3.4.5	3.3.2		

HORTONWORKS DATA PLATFORM

http://hortonworks.com/hdp/whats-new/

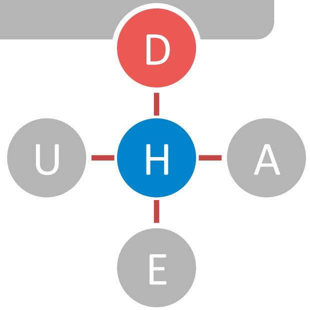
Distros

Old components with known issues

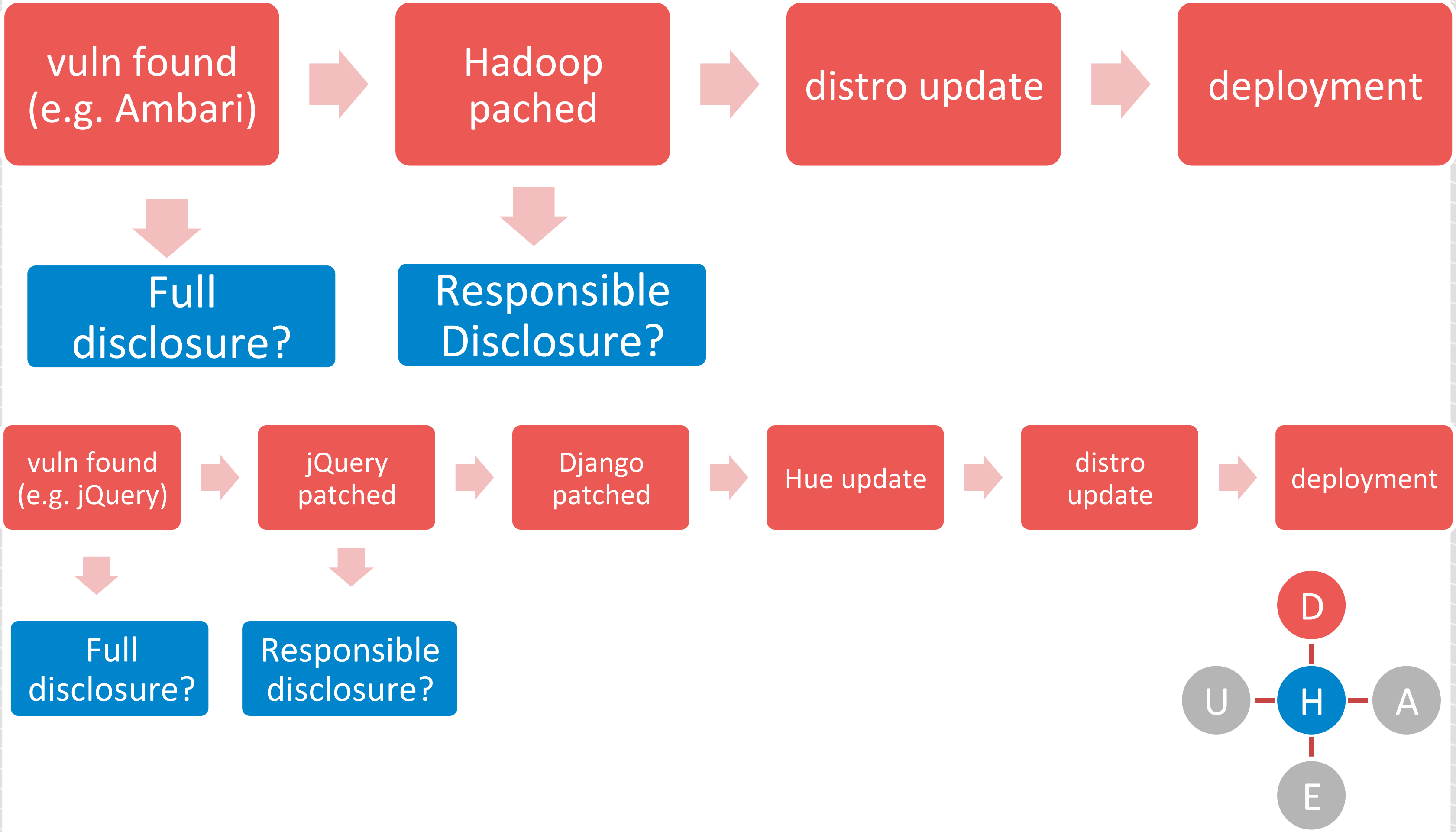
- Old OS components (java, php, ruby, etc.)
- Old OS components (e.g. old tomcat used by Oozie and HDFS)
- Old Hadoop components (e.g. old Hue, Ambari, Ranger)

Default passwords

Default configuration



Vulnerability timeline



vuln found (e.g. Ambari)

Hadoop pached

distro update

deployment

Full disclosure?

Responsible Disclosure?

vuln found (e.g. jQuery)

jQuery patched

Django patched

Hue update

distro update

deployment

Full disclosure?

Responsible disclosure?

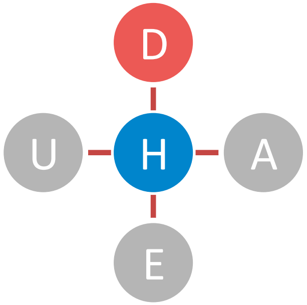
Distros

Old components with known issues

Default passwords

- SSH keys configured but default passwords still work
- Default mysql passwords, NO mysql passwords

Default configuration



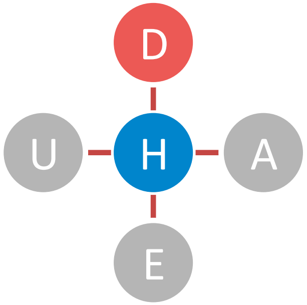
Distros

Old components with known issues

Default passwords

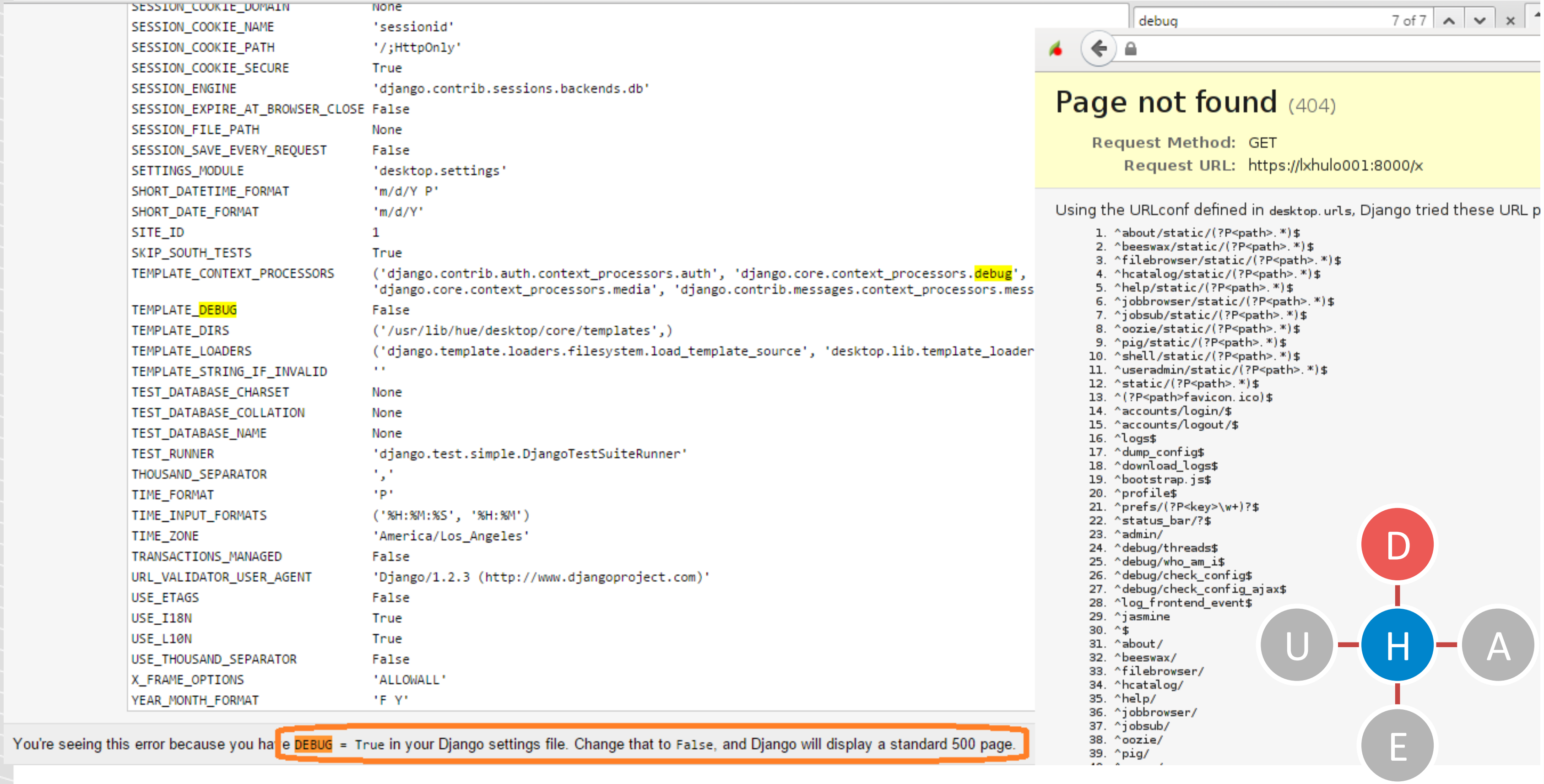
Default configuration

- No network level hardening
- No HTTP hardening (clickjacking, session mgmt, errors)
- Hue uses Django with DEBUG turned on by default
- „Hacking virtual appliances” by Jeremy Brown



Default configurations sucks

X-Frame-Options:ALLOWALL



```
SESSION_COOKIE_DOMAIN               None
SESSION_COOKIE_NAME                 'sessionid'
SESSION_COOKIE_PATH                 '/;HttpOnly'
SESSION_COOKIE_SECURE               True
SESSION_ENGINE                      'django.contrib.sessions.backends.db'
SESSION_EXPIRE_AT_BROWSER_CLOSE     False
SESSION_FILE_PATH                   None
SESSION_SAVE_EVERY_REQUEST          False
SETTINGS_MODULE                     'desktop.settings'
SHORT_DATETIME_FORMAT               'm/d/Y P'
SHORT_DATE_FORMAT                   'm/d/Y'
SITE_ID                             1
SKIP_SOUTH_TESTS                    True
TEMPLATE_CONTEXT_PROCESSORS         ('django.contrib.auth.context_processors.auth', 'django.core.context_processors.debug',
                                    'django.core.context_processors.media', 'django.contrib.messages.context_processors.mess
TEMPLATE_DEBUG                      False
TEMPLATE_DIRS                       ('/usr/lib/hue/desktop/core/templates',)
TEMPLATE_LOADERS                    ('django.template.loaders.filesystem.load_template_source', 'desktop.lib.template_loader
TEMPLATE_STRING_IF_INVALID          ''
TEST_DATABASE_CHARSET               None
TEST_DATABASE_COLLATION             None
TEST_DATABASE_NAME                  None
TEST_RUNNER                         'django.test.simple.DjangoTestSuiteRunner'
THOUSAND_SEPARATOR                  ','
TIME_FORMAT                         'P'
TIME_INPUT_FORMATS                  ('%H:%M:%S', '%H:%M')
TIME_ZONE                           'America/Los_Angeles'
TRANSACTIONS_MANAGED                False
URL_VALIDATOR_USER_AGENT            'Django/1.2.3 (http://www.djangoproject.com)'
USE_ETAGS                           False
USE_I18N                            True
USE_L10N                            True
USE_THOUSAND_SEPARATOR              False
X_FRAME_OPTIONS                     'ALLOWALL'
YEAR_MONTH_FORMAT                   'F Y'
```

You're seeing this error because you ha e `DEBUG = True` in your Django settings file. Change that to `False`, and Django will display a standard 500 page.

Page not found (404)

Request Method: GET
Request URL: https://lxhulo001:8000/x

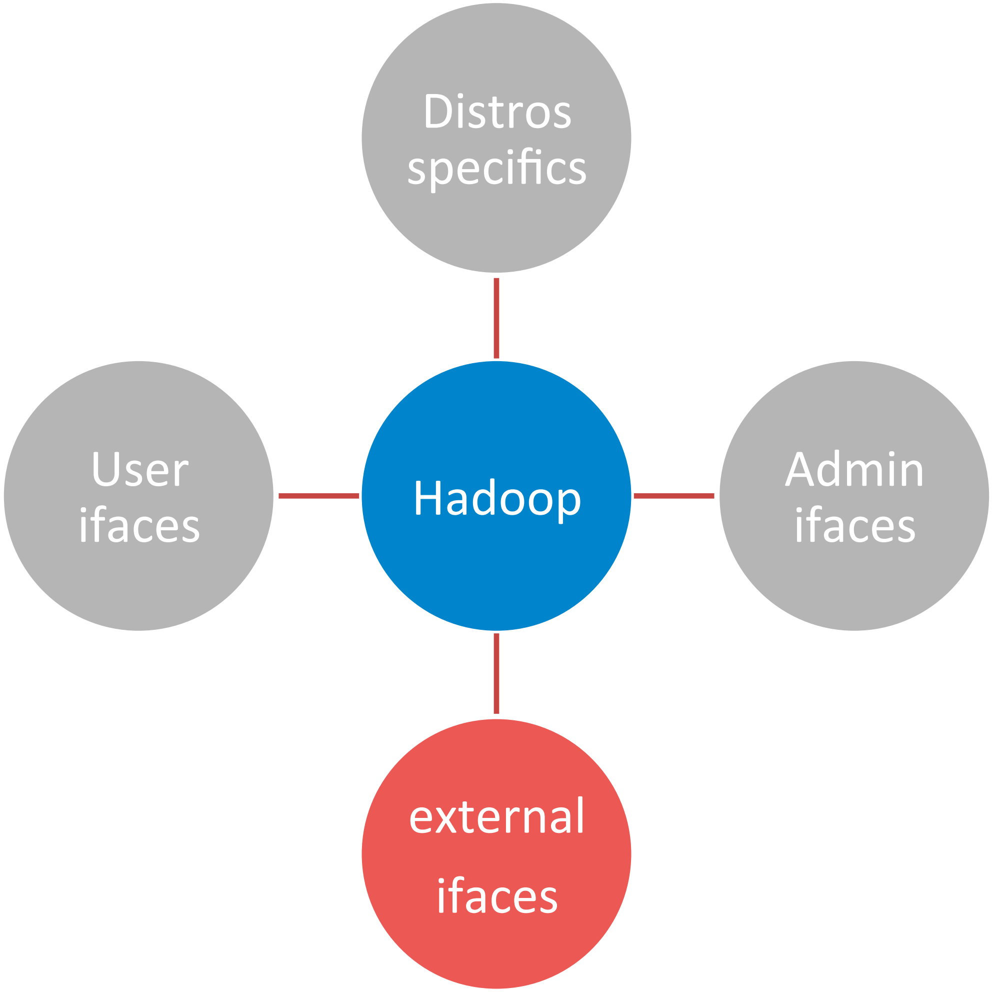
Using the URLconf defined in `desktop.urls`, Django tried these URL p

```
 1. ^about/static/(?P<path>.*)$
 2. ^beeswax/static/(?P<path>.*)$
 3. ^filebrowser/static/(?P<path>.*)$
 4. ^hcatalog/static/(?P<path>.*)$
 5. ^help/static/(?P<path>.*)$
 6. ^jobbrowser/static/(?P<path>.*)$
 7. ^jobsub/static/(?P<path>.*)$
 8. ^oozie/static/(?P<path>.*)$
 9. ^pig/static/(?P<path>.*)$
10. ^shell/static/(?P<path>.*)$
11. ^useradmin/static/(?P<path>.*)$
12. ^static/(?P<path>.*)$
13. ^(?P<path>favicon.ico)$
14. ^accounts/login/$
15. ^accounts/logout/$
16. ^logs$
17. ^dump_config$
18. ^download_logs$
19. ^bootstrap.js$
20. ^profile$
21. ^prefs/(?P<key>\w+)?$
22. ^status_bar/?$
23. ^admin/
24. ^debug/threads$
25. ^debug/who_am_i$
26. ^debug/check_config$
27. ^debug/check_config_ajax$
28. ^log_frontend_event$
29. ^jasmine
30. ^$
31. ^about/
32. ^beeswax/
33. ^filebrowser/
34. ^hcatalog/
35. ^help/
36. ^jobbrowser/
37. ^jobsub/
38. ^oozie/
39. ^pig/
```

EXTERNAL INTERFACES

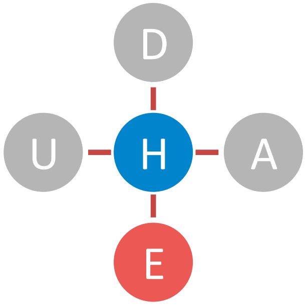
For clients or whatsoever

External interfaces

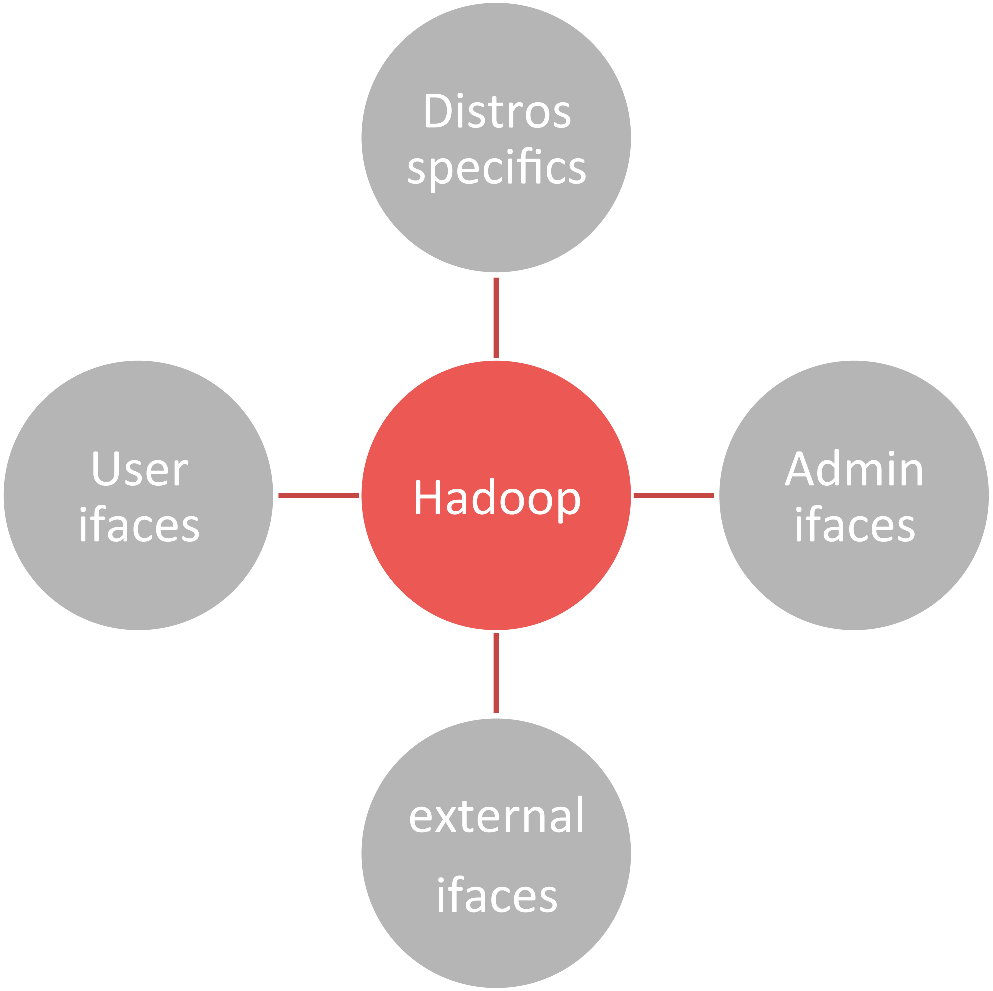


External

- More than 25 internal Apache apps/modules
- Vendor/distro specific apps/interfaces
- Popular monitoring: Ganglia, Splunk
- Auth providers: LDAP, Kerberos, OAuth
- Many apps, many targets



Hadoop



Distros specifics

User ifaces

Hadoop

Admin ifaces

external

ifaces

SUMMARY

ways to protect your big data environment

Ways to protect your Hadoop environment

Excessive network access

- Keep it super tight!

Excessive user pesmissions

Typical web vulnerabilities

Obsolete software

Distros dependent vulnerabilities

External system connections

Ways to protect your Hadoop environment

Excessive network access

Excessive user permissions

- Map business roles to permissions

Typical web vulnerabilities

Obsolete software

Distros dependent vulnerabilities

External system connections

Ways to protect your Hadoop environment

- Excessive network access
- Excessive user permissions
- Typical web vulnerabilities
  - Pentest it! Introduce application independent security countermeasures
- Obsolete software
- Distros dependent vulnerabilities
- External system connections

Ways to protect your Hadoop environment

Excessive network access

Excessive user permissions

Typical web vulnerabilities

Obsolete software

- Make a list of all components. Monitor bugtracks and CVEs.

Distros dependent vulnerabilities

External system connections

Ways to protect your Hadoop environment

Excessive network access

Excessive user permissions

Typical web vulnerabilities

Obsolete software

Distros dependent vulnerabilities

- A pentest after integration is a must. Demand security from software suppliers.

External system connections

Ways to protect your Hadoop environment

Excessive network access

Excessive user permissions

Typical web vulnerabilities

Obsolete software

Distros dependent vulnerabilities

External system connections

- Make a list of all external system connections. Do a threat modeling and pentest corresponding systems.

securing

MORE THAN SECURITY TESTING

Thank you

Contact me for additional materials

@j_kaluzny
jakub.kaluzny@securing.pl