Converging insightful, data-led applications with traditional web applications

Michael McCune
Senior Software Engineer

Steven Pousty
Lead Developer Evangelist, OpenShift Online
GOALS

● Brief introduction to how things have evolved
● Talk about what we have now
● Demo of this in practice
● Have fun!
ASSUMPTIONS

- You have written a web app at some point in your life
- You know some basics of Spark and such
WHERE ARE WE COMING FROM?
Log

Command ===>

http://www.sas.com/service/techsup/intro.html

Online Documentation/Manuals are available at URL:
http://support.sas.com/onlinedoc/913/docMainpage.jsp

NOTE: No options specified.
NOTE: 33554432 bytes were available above the line at initialization.
NOTE: 4259848 bytes were available below the line at initialization.
NOTE: 330390144 bytes were available above the line after adjustment for MKHLEAVE=524288.
NOTE: The initialization phase used 0.08 CPU seconds and 9659K.
NOTE: The address space has used a maximum of 712K below the line and 10480K above the line.

---

Program Editor

Command ===>

00001
00002
00003
00004
00005
00006
00007
00008
00009
00010
00011
00012
00013
00014
DATA temp;
  input @1 subj 4.
  @6 f_name $11.
  @18 l_name $6.
  +3 height 2.
  +5 wt_date mmddyy8.
  +1 calorie comma5.;
format wt_date mmddyy8. calorie comma5.;
DATALINES:
1024 Alice Smith 1 65 125 12/1/95 2,036
1167 Maryann White 1 68 140 12/01/95 1,800
1168 Thomas Jones 2 190 12/2/95 2,302
1201 Benedictine Arnold 2 68 190 11/30/95 2,432
1302 Felicia Ho 1 63 115 1/1/96 1,972
; RUN;

PROC PRINT data = temp;
  title 'Output dataset: TEMP';
  id subj;
RUN;
WHAT ABOUT APPS?
NEXT ERA
WELCOME TO PARADISE
WE HAVE HAD SOME GREAT DEVELOPMENTS

- People figured out how to make commodity hardware do great things
- Distributed software has gotten better
- Programming languages have gotten easier
- CS has put effort into statistical libraries
- The browser has grown up
WHAT DOES CONVERGENCE LOOK LIKE?

OPENSHIFT
ARCHITECTING YOUR APPLICATION

MySQL
Kafka
ActiveMQ
Python
PostgreSQL
HDFS
MongoDB
Node.js
Spark
Ruby
HTTP
HAMMERING YOUR APP INTO SHAPE

Node.js → Python → Spark → MongoDB → HTTP

Ingest → Process → Publish
BUILDING AS A TEAM

Node.js  Python  Spark  MongoDB  HTTP
TAKING IT TO THE NEXT LEVEL

Project Front-end
- Node.js
- Mock-server

Project REST-server
- Python
- Spark
- Spark
WHAT DOES IT DO?
WHAT IS WORD2VEC?

\[ v(\text{"madrid"}) - v(\text{"spain"}) + v(\text{"france"}) \approx v(\text{"paris"}) \]
DEMO
THINGS THAT HELPED OUR COLLABORATION
THINGS THAT REQUIRE GREATER COORDINATION

- API construction
- Compute resource affinity
- Persistent storage
WHAT ABOUT THE DEVOPS?
IMPROVING THE PROCESS
TEST DRIVE IT!

More projects, tutorials and examples can be found at

RADANALYTICS.IO
CONTACT US

Michael McCune
- elmiko on Github and IRC
- @FOSSjunkie on Twitter

Steven Pousty
- TheSteve0 on Twitter, IRC, SmugMug, Github, Ingress, Instagram and Skype
RED HAT SUMMIT

LEARN. NETWORK. EXPERIENCE OPEN SOURCE.