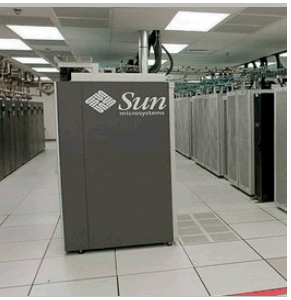
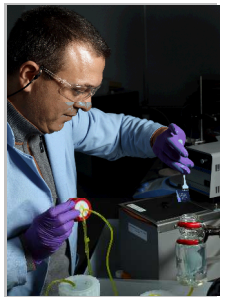
SAND2015-6041PE

Lessons learned from implementing the Hadoop Distributed File System as a server backend





By Daniel Houston

Sandia National Laboratories

Exceptional service in the national interest

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

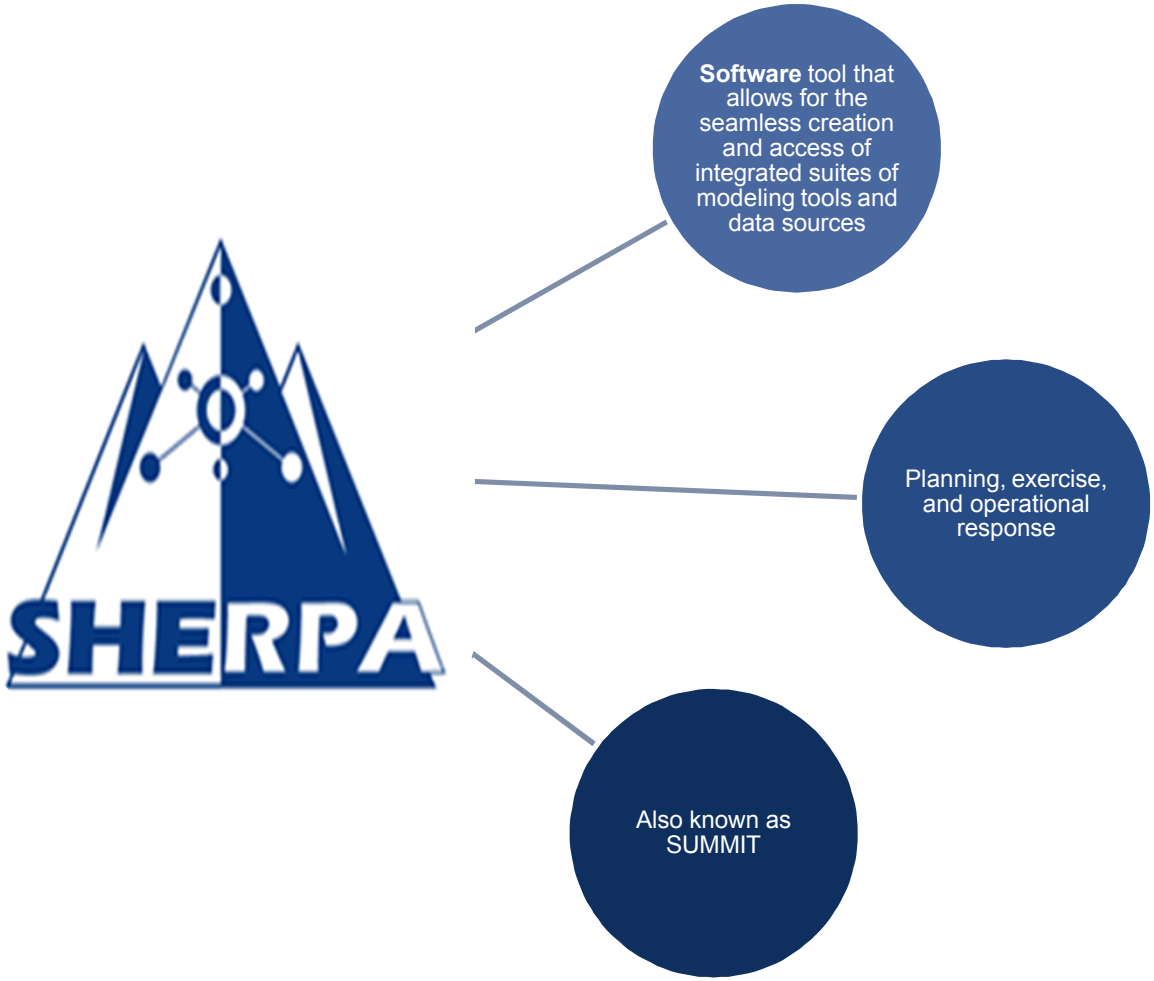
About Me

Sandia National Laboratories

- Junior at the University of Utah
- Computer Science B.S. est. Dec 2016
- From Salt Lake City
- Previously worked at a financial analytics company (Visible Equity)
  - Large data manipulation and storage
  - Data visualization
- Interests are Big Data Management, Operating Systems, Networks, and Computer Languages

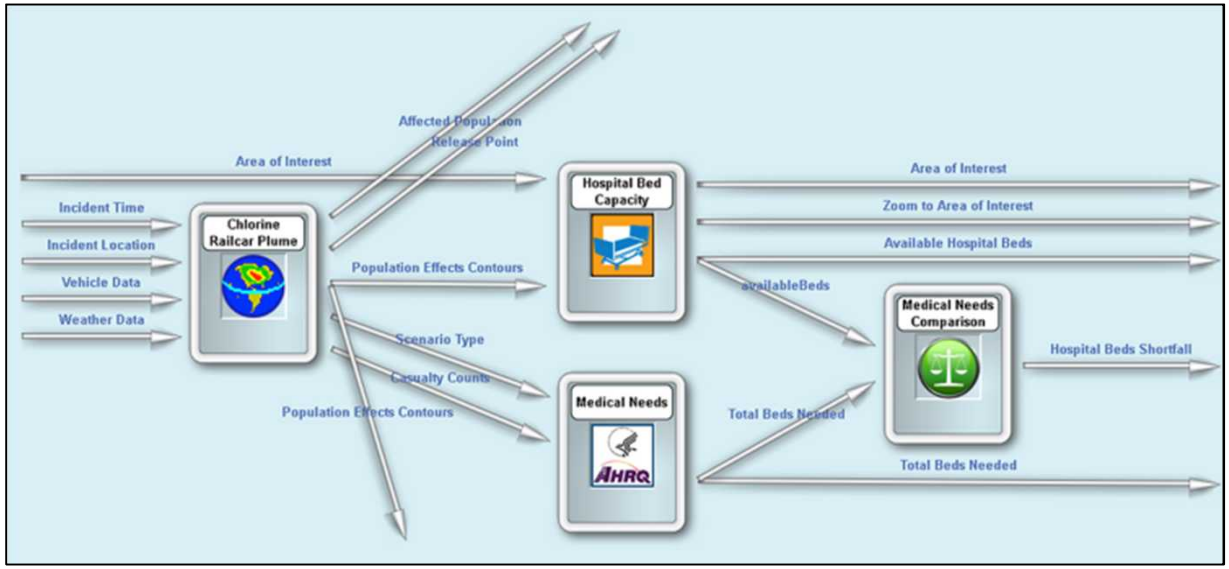
Sherpa Introduction

SHERPA – the **S**UMMIT for **H**omeland **E**mergency **R**esponse, **P**lanning, and **A**nalysis



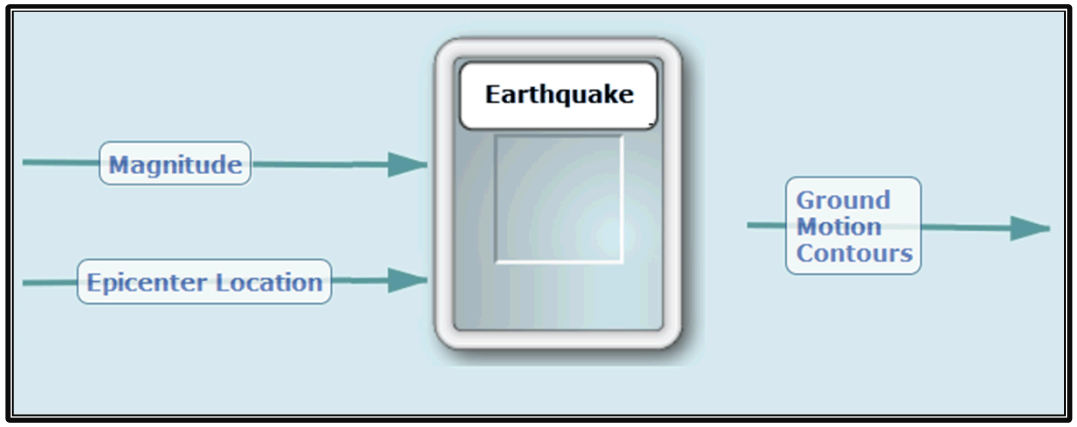
SUMMIT Process

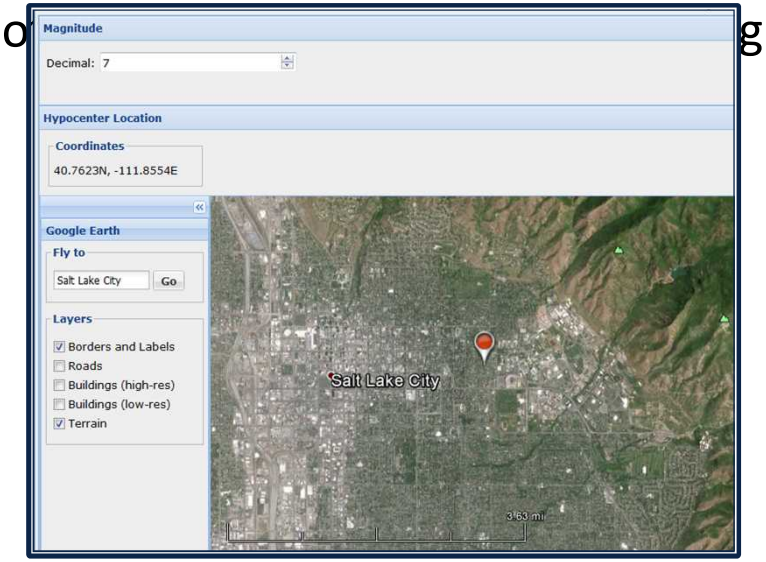
- Models
- Model Wrappers
- Templates
- Template Runs

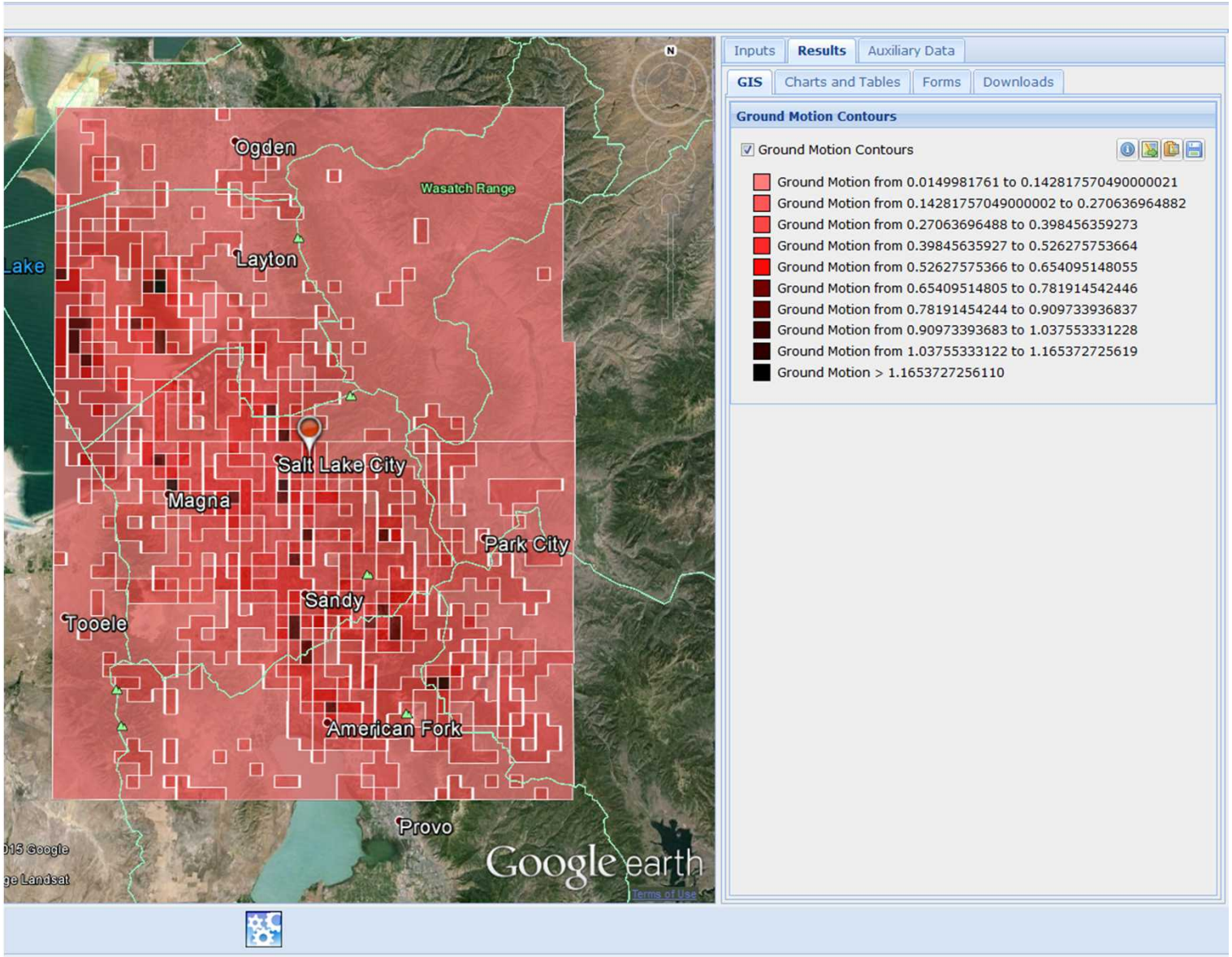


Open Quake

- Earthquake scenario simulation
- Inputs:
  - Magnitude: 7
  - Location: Salt Lake City
    - (40.7623, -111.8554)
- Outputs:
  - Ground mo ground acceleration (***g***)






Ogden
Wasatch Range
Layton
Salt Lake City
Magna
Park City
Sandy
Tooele
American Fork
Provo
Google earth
Inputs
Results
Auxiliary Data
GIS
Charts and Tables
Forms
Downloads
Ground Motion Contours
Ground Motion Contours
Ground Motion from 0.0149981761 to 0.142817570490000021
Ground Motion from 0.14281757049000002 to 0.270636964882
Ground Motion from 0.27063696488 to 0.398456359273
Ground Motion from 0.39845635927 to 0.526275753664
Ground Motion from 0.52627575366 to 0.654095148055
Ground Motion from 0.65409514805 to 0.781914542446
Ground Motion from 0.78191454244 to 0.909733936837
Ground Motion from 0.90973393683 to 1.037553331228
Ground Motion from 1.03755333122 to 1.165372725619
Ground Motion > 1.1653727256110

Batch Runs

Sandia National Laboratories

- One template
- Set of different inputs
- Run template repeatedly with different combinations of inputs
- Add two integers: A + B = C
  - Input A = {1, 2, 3} Input B = {10, 20}
  - 10+1
  - 10+2
  - 10+3
  - 20+1
  - 20+2
  - 20+3
- Looking for trends and most-likely scenarios

Batch Runs

Sandia National Laboratories

- System with 5 inputs at 5 values per input
  - 5^5 Runs
- Batch Runs = 5^5 runs with storage per run at about 0.5 MB = ~1.5 GB of data per batch run
- Very quickly creates large amounts of data

Large Amounts of Data

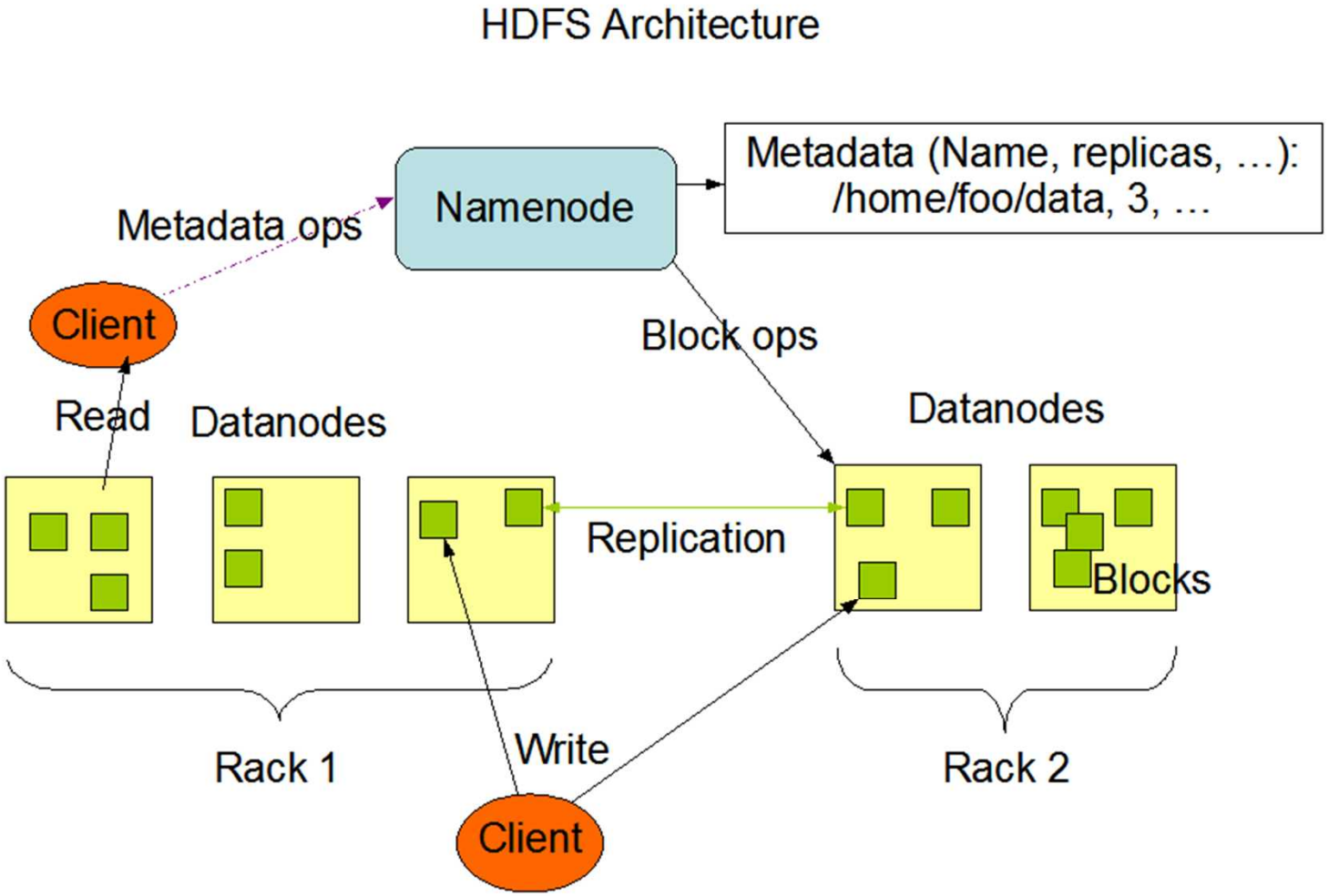
Sandia National Laboratories

- “The system should support an average high profile event with at least 50 GB of disk space…System will require 220 GB of disk space per year for saving template runs, and must provide a data capacity to accommodate a growth rate by at least 20%”
  - Summit Systems Design Document

Sandia National Laboratories

THE HADOOP DISTRIBUTED FILE SYSTEM (HDFS)

The Hadoop Distributed File System (HDFS)



The Hadoop Distributed File System (HDFS)

Sandia National Laboratories

- Highly Scalable
- Keep same File System structure
- Parallelizable = high throughput

- Requires more setup
- Learning Curve

Sandia National Laboratories

LESSONS LEARNED / CHALLENGES

HDFS is not a Database

Sandia National Laboratories

- Common misconception
- It is a file system
- Speed and security are secondary concerns
- There are tools built on HDFS that provide an SQL-like abstraction
  - Hbase
  - Hive
    - Tools are still in development, documentation very limited
- Use HDFS if scalability and availability are your main concerns

Getting Access to a cluster

Sandia National Laboratories

- Lots of Hardware can be expensive
- Large virtual cluster
- Virtual Machine Cluster
  - Needs lots of memory and lots of disk space
- Only have access to one Virtual Node
- Solution: Run in pseudo-distributed mode

Connecting remotely to HDFS

Sandia National Laboratories

- Java client not built for remote connections
- Nasty Stacktrace, unknown cause

```
java.io.IOException: File /user/ubuntu/pies could only be replicated to 0 nodes, instead of 1
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getAdditionalBlock(FSNamesystem.java:1448)
at org.apache.hadoop.hdfs.server.namenode.NameNode.addBlock(NameNode.java:690)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:39)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:25)
at java.lang.reflect.Method.invoke(Method.java:597)
at org.apache.hadoop.ipc.WritableRpcEngine$Server.call(WritableRpcEngine.java:342)
at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:1350)
```

- Solution: WebHDFS or HTTPfs

```
public static void main(String[] args) {
    Configuration conf = new Configuration();
    conf.set("fs.defaultFS", "webhdfs://123.456.789.101:8888/");

    try {
        FileSystem fs = FileSystem.get(conf);

        // Do Stuff
    } catch (IOException e) {

    }
}
```

Documentation is limited

Sandia National Laboratories

- Open source project, still needs time to build up documentation
- Documentation that exists is very basic
- Online blogs can be a start
- Lots of community support
- Hadoop: The Definitive Guide by Tom White

Tools built on top of HDFS

Sandia National Laboratories

- Map Reduce
- SQL Engine:
  - Apache Hive
- Other DB-like abstractions:
  - Hbase
- Yarn
- Spark
- Pig
- Sqoop
- WebHDFS
- Hbase

Other Challenges

Sandia National Laboratories

- Cluster configuration
- File System abstraction slightly different
- Security is a secondary concern
- Online tutorials
- Hadoop: The Definitive Guide

Value Added

Sandia National Laboratories

- Allows scalable data storage for batch run data
- Data stored on HDFS, which means it is easily accessible by numerous big-data aggregation and analysis tools (ie Map Reduce, Spark)
- Abstract implementation allows for flexibility in back end for different needs

Take Away

Sandia National Laboratories

- HDFS has a learning curve.
- HDFS is a file system, not a database
- Documentation is limited, so requires digging and ‘floundering’
- If speed is not your concern, HDFS is a great back end storage option
- Community support is there and will continue to grow
- Need hardware availability for eventual utilization