



Real-Time Stream Processing with HDP

Hortonworks. We do Hadoop.

Spring 2014

Version 1.0

Agenda

- **CDR Monitoring Use Cases**
- **Solution Architecture**
- **Demo**
- **Code walk-through**
- **Next steps**

CDR Monitoring Use Cases

- **CDR (Call Detailed Record) current generated in CSV files in a directory**
 - Data can arrive few minutes late
 - Plans to change date capture in the future
- **Two use cases, analyzing CDR data in Real Time**
 - Use case 1: Network Type Change
 - Use case 2: Dropped calls

CDR Monitoring Use Cases

- **Use Case 1: Network Type Change**

1. Group CDR rows from same “session-id”
2. Compare each row with previous row (from same “session-id”) and check if “network-type” has changed,
3. If yes compute: “cell-id” from previous row, “cell-id” from new row, “change-type” (example: from type 1 to type 2, or from type 2 to type 1)
4. Compute top “cell-id” with network type changes from past 30 minutes (rolling window)

SessionID	SimCardID	Phone Number	Date Time	Duration	Cell ID	Network Type
change-9b409c6b-c285-436d-809f-9daedfe67d41	SIM-00008	PHONE-00004	24/04/2015 17:51:20	214	cell J	1
change-9b409c6b-c285-436d-809f-9daedfe67d41	SIM-00008	PHONE-00004	24/04/2015 17:51:20	214	cell A	2
change-0b29b10c-7a7f-4dac-8cb6-90effe07c662	SIM-00008	PHONE-00004	24/04/2015 17:51:21	77	cell F	2
change-0b29b10c-7a7f-4dac-8cb6-90effe07c662	SIM-00008	PHONE-00004	24/04/2015 17:51:21	77	cell F	1

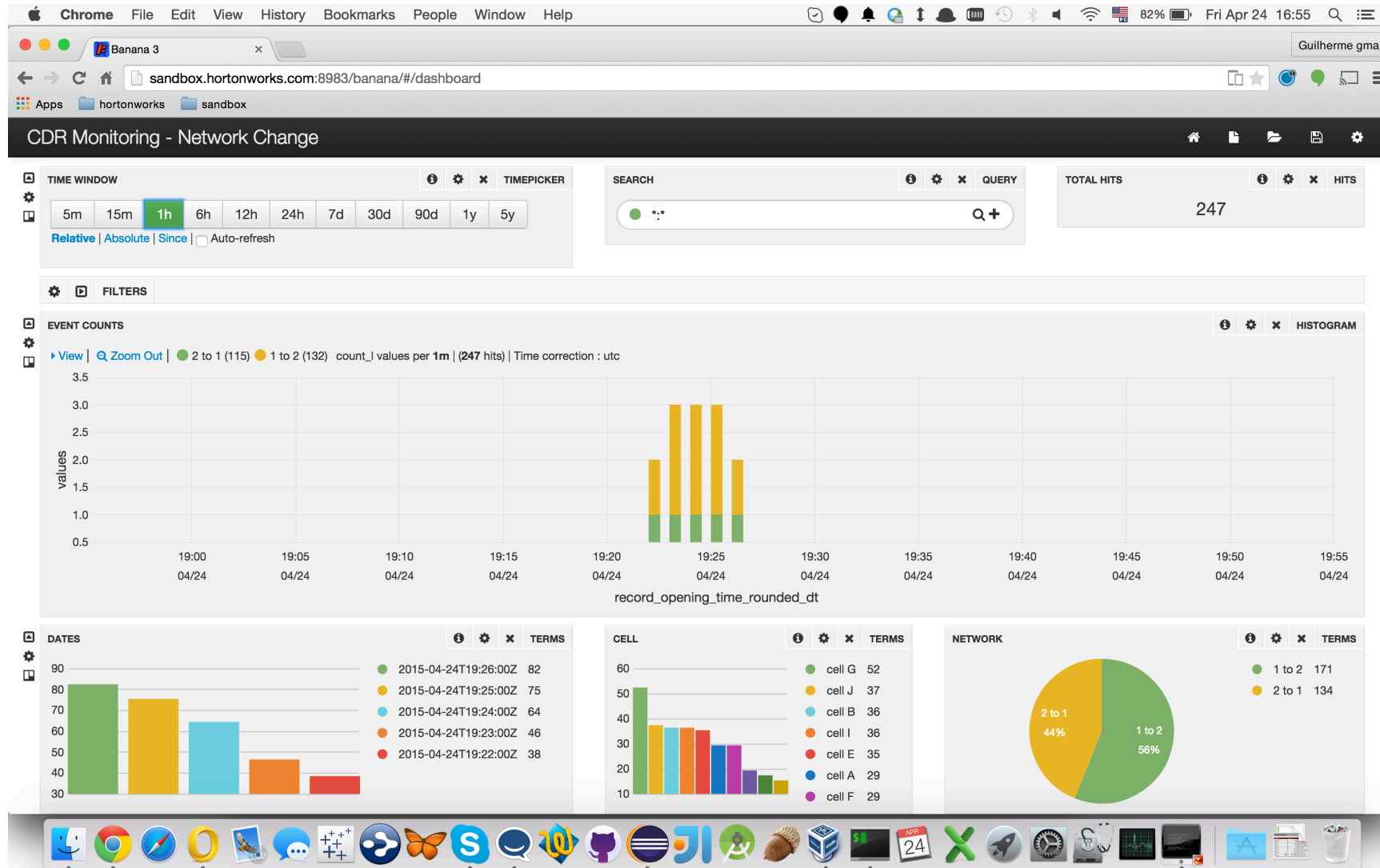
CDR Monitoring Use Cases

- **Use Case 2: Dropped Calls**

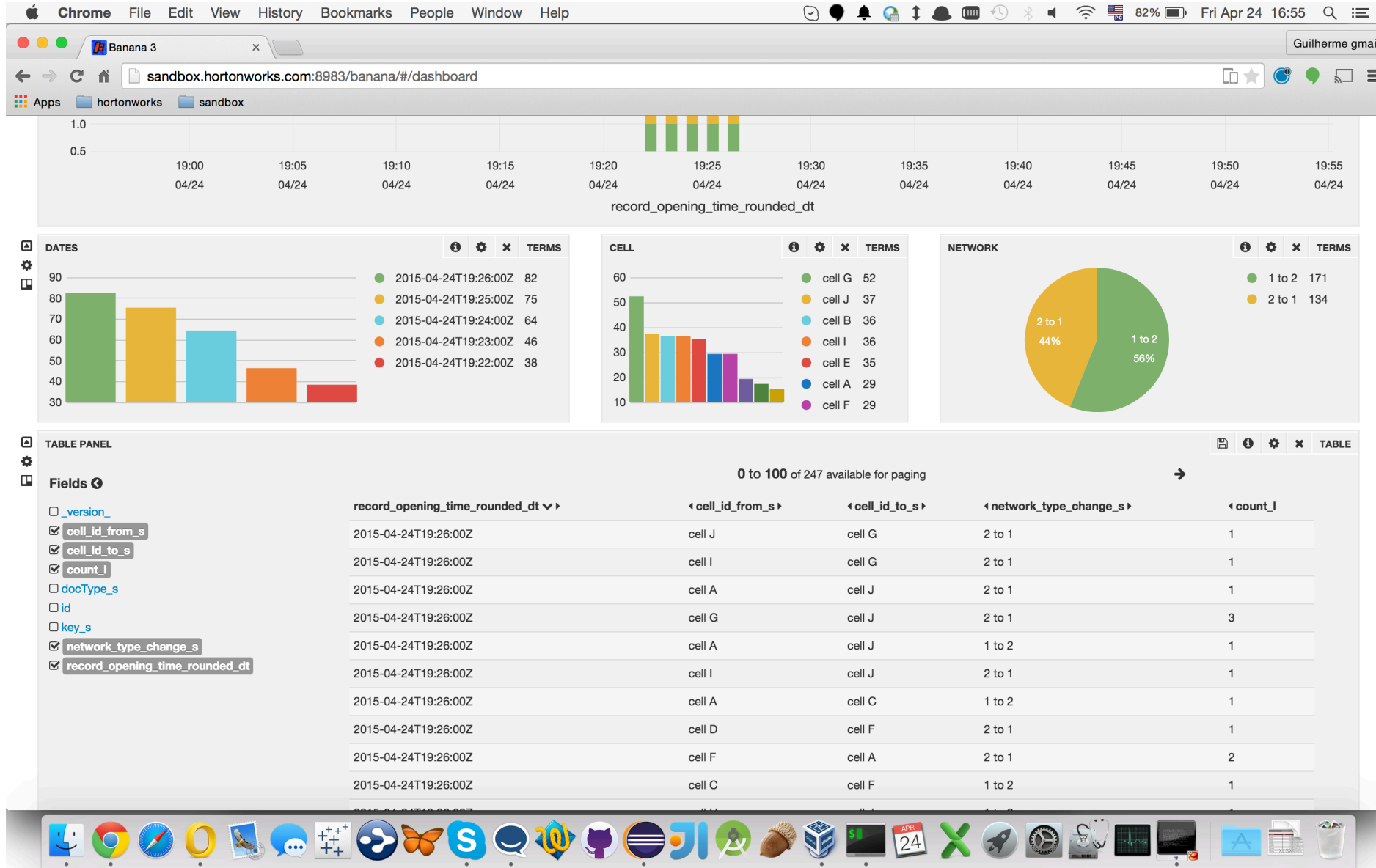
1. Group CDR rows from same “sim-card-id” + “phone number”
2. Compare each row with previous row (from same “sim-card-id” + “phone number”) and check if there is a new call within less than 1 minute from last call (“start time” + duration)
3. If yes compute: “cell-id” where call has been dropped and “drop reason”
4. Compute top “cell-id” with dropped calls from past 30 minutes (rolling window)

SessionID	SimCardID	Phone Number	Date Time	Duration	Cell ID	Network Type	Drop Reason
drop-95997adb-259e-43f8-a43f-7684a866da97	SIM-00001	PHONE-00005	24/04/2015 17:51:13	6	cell J	1	A
drop-95997adb-259e-43f8-a43f-7684a866da97	SIM-00001	PHONE-00003	24/04/2015 17:52:00	189	cell J	1	B
drop-aadc97a8-e459-4c37-8890-258d3bc14e0c	SIM-00002	PHONE-00007	24/04/2015 17:51:14	9	cell C	2	A
drop-aadc97a8-e459-4c37-8890-258d3bc14e0c	SIM-00002	PHONE-00007	24/04/2015 17:51:39	123	cell C	2	A

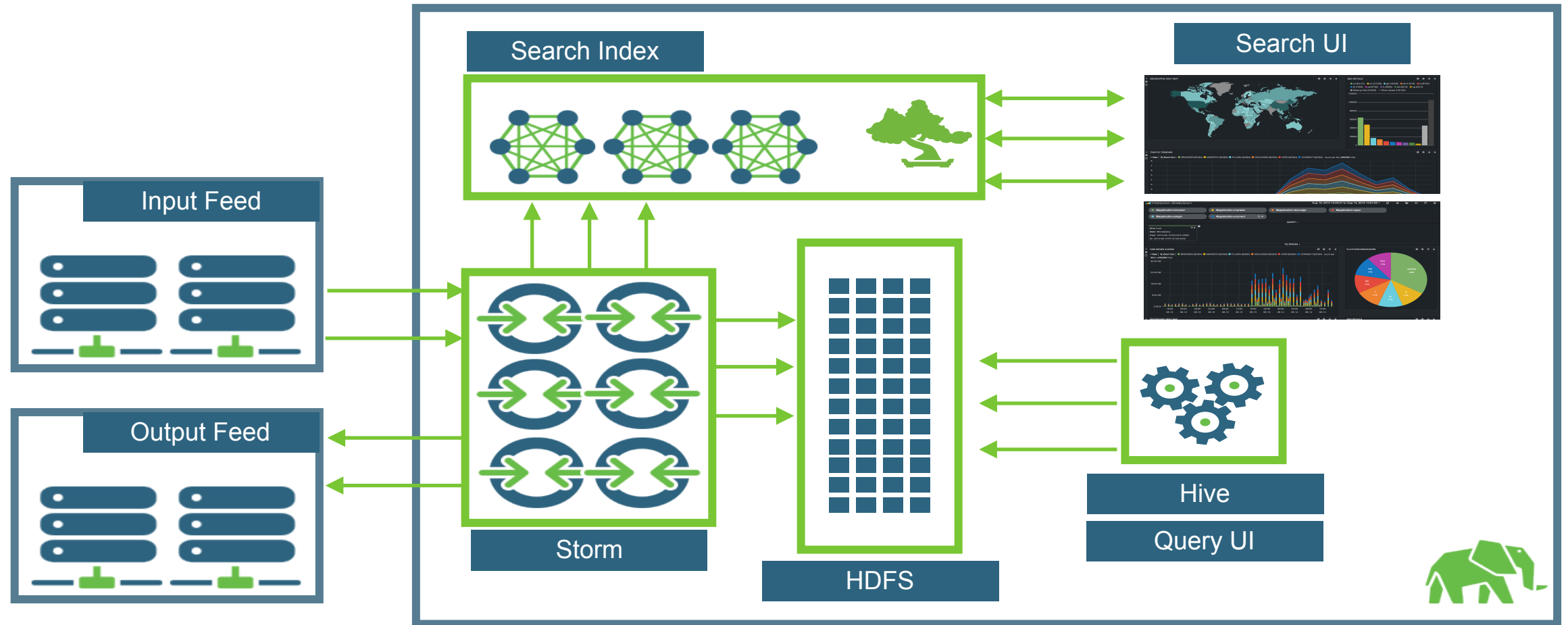
CDR Monitoring Use Cases



CDR Monitoring Use Cases



Data Flow and Architecture



Input Harness Generating JMS Messages

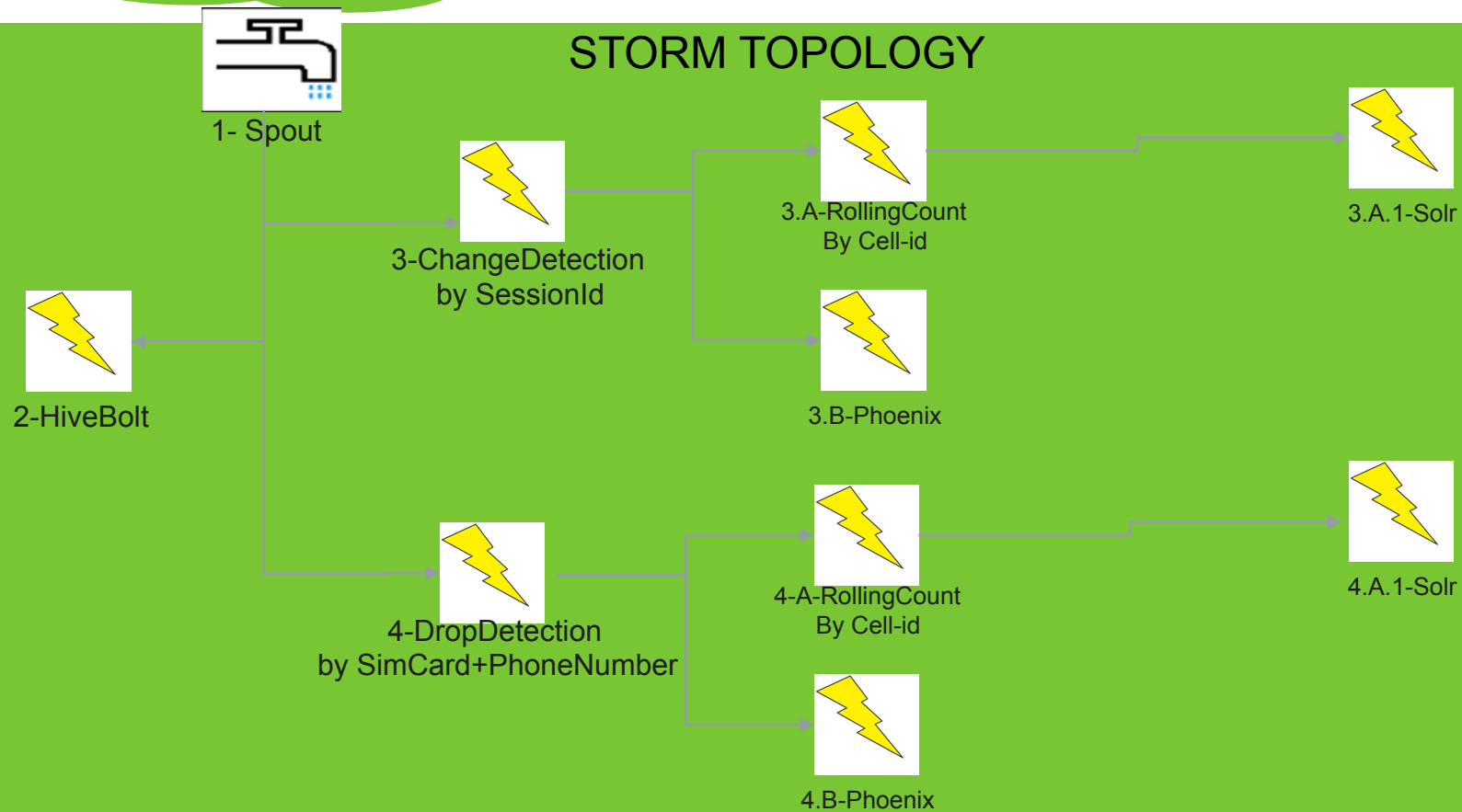
Solution Architecture



Flume

KAFKA (messages)

STORM TOPOLOGY



Solution Architecture

