



# How Hadoop clusters break in the real world

Ari Rabkin (UC Berkeley and Cloudera, inc.) Randy Katz (UC Berkeley)

## Problem

System builders don't know which failure patterns are common, so can't avoid. Results in brittle systems.

## Approach

Look at real failure data for Hadoop ecosystem to understand what goes wrong. Data includes HBase, Pig, Hive, Zookeeper, Flume, Oozie, etc.

## Observations

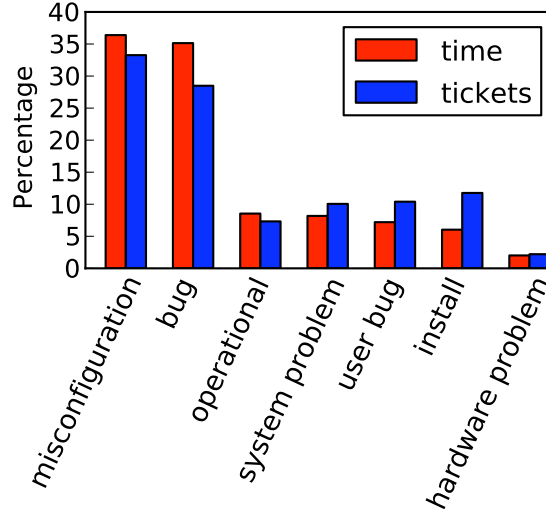
Classified 293 Cloudera support tickets, in period from Feb. to July 2011.

Used ticket count + reported supporter time spent as metrics.

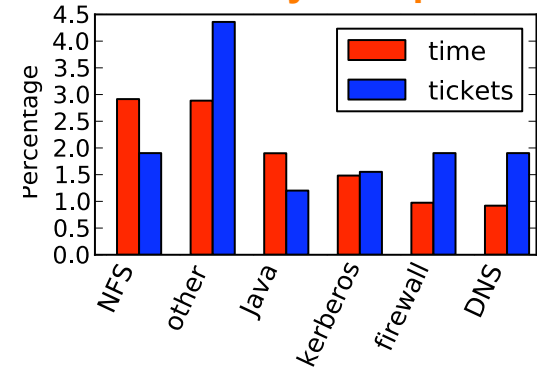
Classification was by ultimate fix: Ex: 'bug' means support case required patch to close.

Most other problems could also have been fixed by some re-design. (These are not counted as bugs.)

## Breakdown of support time and tickets



## Detail on system problems



## Observation

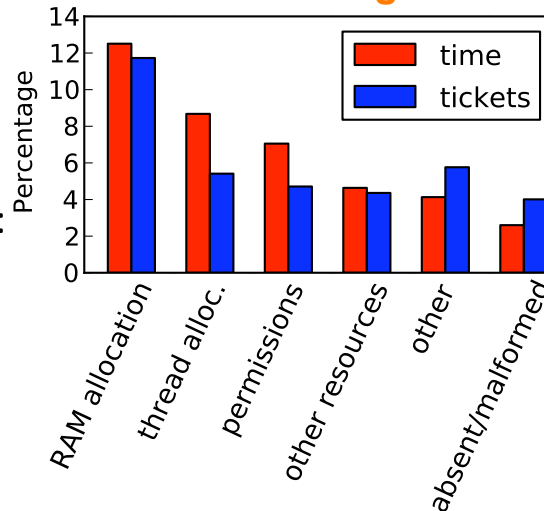
Incompatibly-set options cause all thread allocation and most RAM allocation problems.

Config. options should be aligned with sensible scaling dimensions.

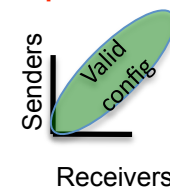
Potentially prevents most misconfigurations

Hadoop requires balanced # of send and receive threads for MR shuffle. But users specify each separately.

## Detail on misconfigurations



## Error-prone current design



## Better alternative

