

# Catching Failures of Failures at Big-Data Clusters: A Two-Level Neural Network Approach



**Andrea Rosà<sup>\*</sup>, Lydia Y. Chen<sup>†</sup>, Walter Binder<sup>\*</sup>**

<sup>\*</sup>Università della Svizzera italiana, Faculty of Informatics, Switzerland

<sup>†</sup>IBM Research Zurich Lab, Cloud Server Technologies Group, Switzerland

IEEE/ACM IWQoS 2015

June 16, 2015

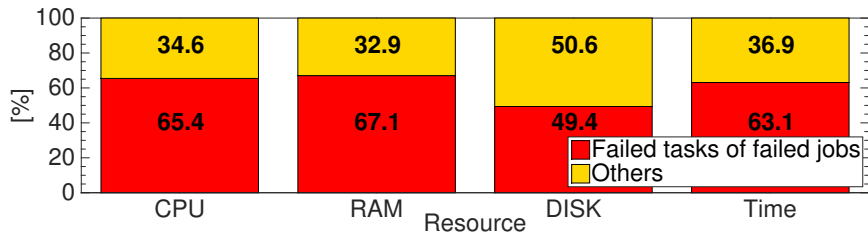
This work has been supported by the Swiss National Science Foundation (project 200021\_141002) and EU commission under FP7 GENiC project (608826).

---

# Big-Data Systems

- ▶ **Big-data** is becoming a key requirement for many applications
  - ▶ Used for large-scale simulations, scientific computations, sensor networks, ...
- ▶ Systems have large **scale** and are very **complex**
- ▶ A lot of job and task **failures**...
- ▶ ...that potentially turn into significant **resource waste**

# Resource Waste



Field data: Google cluster trace [1]

- ▶ Failed tasks of failed jobs (**failures of failures**)
  - ▶ 25% of tasks...
  - ▶ ... waste 60% of resources
- ▶ It is essential to **predict** failures of failures and **mitigate** resulting resource waste

[1] J. Wilkes, More Google cluster data, Google research blog. Nov 2011.

## Contributions

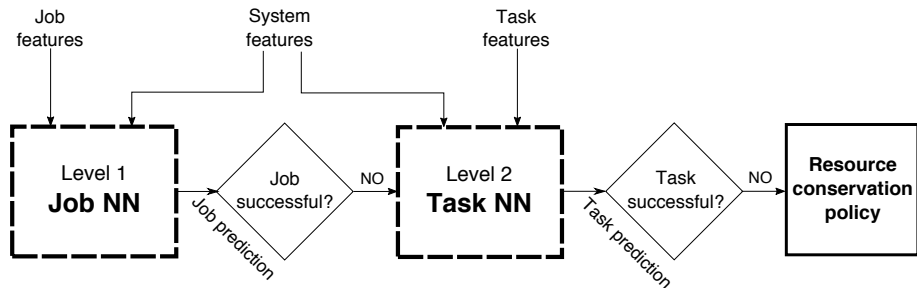
- ▶ Development of **on-line prediction methodology** for **job and task outcomes**
  - ▶ Prediction upon job and task arrival
- ▶ Proposal of **resource conservation policy**
  - ▶ Terminates failures of failures after a grace period
  - ▶ Idea: misclassified tasks still have chance to complete successfully
- ▶ Goal: maximize **resource conservation** and minimize **harmful terminations**

# Data Description

- ▶ Google cluster trace [1] used as case study
  - ▶ 29 days of workload
  - ▶ 672k+ **jobs**, 25M+ **tasks**

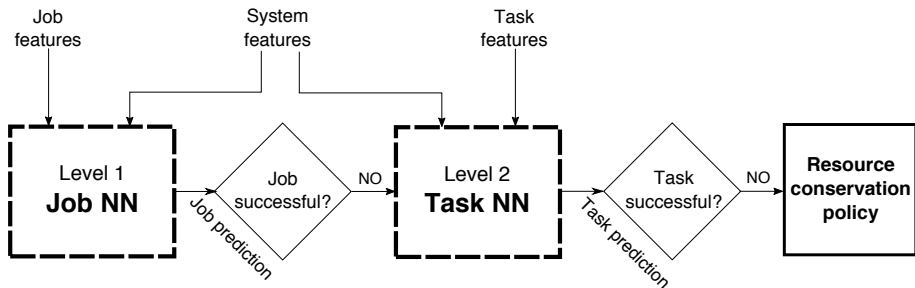
[1] J. Wilkes, More Google cluster data, Google research blog. Nov 2011.

# Methodology



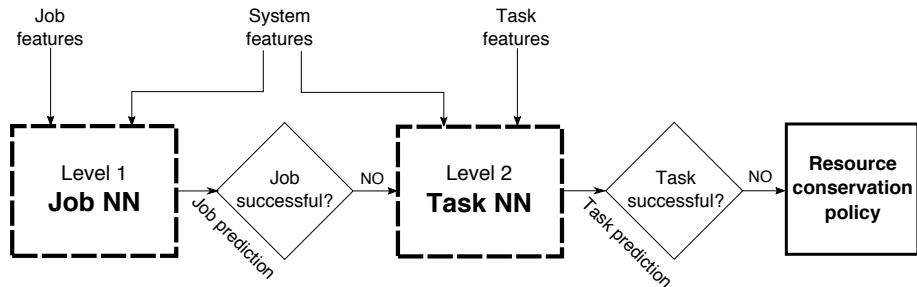
- ▶ **2-level** prediction model
  - ▶ Job classes: **successful**, **failed**
  - ▶ Task classes: **finish**, **eviction**, **fail**, **kill**
  - ▶ Based on **Neural Networks (NN)**
    - ▶ Input: job/task/system features
    - ▶ Output: prediction

# Methodology



- ▶ Resource conservation policy
  - ▶ Only on failed tasks of failed jobs
  - ▶ **Grace period** depends on predicted class
  - ▶ AVG execution time: **fail** = 100min; **eviction** = 156min

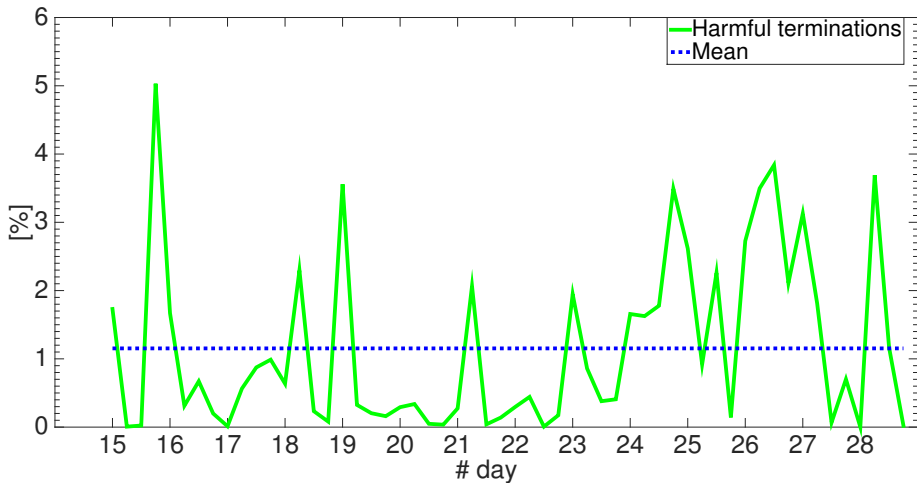
# Methodology



- ▶ More details in the paper...
  - ▶ On-line prediction
  - ▶ Prediction confidences
  - ▶ Learning window size
  - ▶ Grace period computation
  - ▶ Feature set
  - ▶ ...

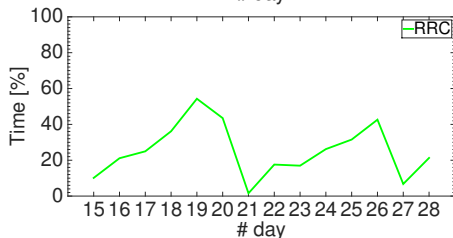
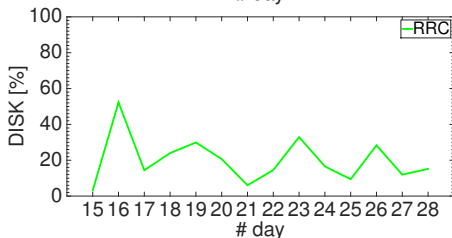
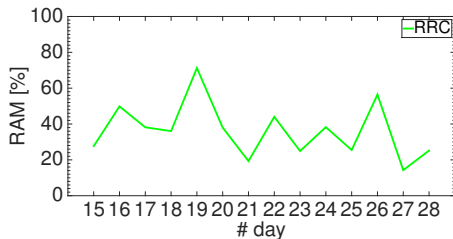
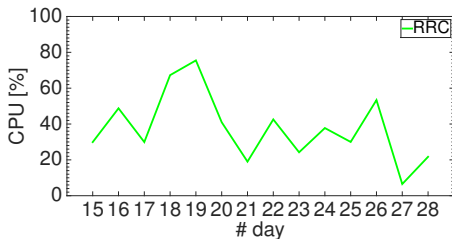


## Testing Phase - Harmful Terminations



► AVG harmful terminations = 1.15%

# Testing Phase - Resource Conservation



- ▶ RCC = Reduction of Resource Consumption
- ▶ Total *RRC* = 49% (CPU), 45% (RAM), 18% (DISK), 26% (time)

## Conclusion

- ▶ We developed an **on-line prediction methodology** for **job and task outcome**
- ▶ We developed a **delay-based resource conservation policy**
- ▶ High **resource conservation** and few **harmful task terminations**

## Conclusion

- ▶ We developed an **on-line prediction methodology** for **job and task outcome**
- ▶ We developed a **delay-based resource conservation policy**
- ▶ High **resource conservation** and few **harmful task terminations**



**Andrea Rosà**



andrea.rosa@usi.ch



<http://www.inf.usi.ch/phd/rosaa>



Università della Svizzera italiana (USI)  
Faculty of Informatics

Università  
della  
Svizzera  
italiana

Faculty  
of Informatics