Introduction

• Demand for cloud-based processing frameworks continues to grow
• Cloud applications is becoming increasingly mission-critical and deadline sensitive, especially in shared clusters
• Cloud providers seek efficient techniques to meet the SLA
• A few slow tasks, called stragglers, could significantly impact job execution time

Our work proposes Shed, an optimization framework that leverages dynamic cloning to jointly maximize jobs’ Probability of Completion before Deadline (PoCD) by fully utilizing the available resources.

System Architecture

Proposed Online Algorithm

Algorithm 1: Proposed Online Algorithm
1. Upon submission of a new job:
2. Kill all jobs which missed their deadlines
3. $J = \{j_1, j_2, j_3, \ldots\}$
4. if $|J| = 1$ then
5. $r_{\max} = \frac{\lambda \cdot m - N_j - |J|}{N_j}$
6. $r_j = r_{\max}$
7. else
8. $r_j = 0 \forall j$
9. $\omega = 0$
10. $\kappa = \lambda \cdot m - \sum_{j=1}^{J} N_j - |J|$
11. Calculate $R_j \forall j$
12. while $J \neq \{0\}$ do
13. $j' = \arg \min_j \{R_j\}$
14. if $N_{j'} + \omega > \kappa$ then
15. $J = J - \{j'\}$
16. else
17. $r_{j'} = r_{j'} + 1$
18. $\omega = \omega + N_{j'}$
19. Calculate $R_j \forall j$
20. end if
21. end while
22. end if

Joint PoCD Optimization

maximize $\sum_{j} U(p_j)$
subject to $\sum_{j=1}^{N_j} (r_j + 1) + |J| \leq \lambda \cdot m$
$p_j = R_j(r_j), \forall j$
$r_j \geq 0, \forall j$

Conclusion

In this work, we propose Shed, an optimization framework that leverages dynamic cloning to jointly maximize PoCD and cluster utilization. We also present an online scheduler that dynamically optimizes resources upon new job arrival. Our solution includes an online greedy algorithm to find the optimal number of clones needed for each job. Our results show, in some cases, that Shed can achieve 100% PoCD compared to Dolly and Hadoop with speculation. The proposed algorithm is able to achieve more than 90% utilization of available cloud resources, whereas Dolly and Hadoop achieves only about 22%.

Reference

• J. Dean et al, 2008
• G. Ananthanarayanan et al, 2013

Evaluation

Theorems

$R_{on} = \left[ 1 - \left( \frac{r_{\min}}{D_j} \right)^{\beta(r_j + 1)} \right]^{N_j}$
$R_{ru} = 1 - \left( \frac{1 - \eta_{\min}}{D_j - r_j} \right)^{\beta(r_j + 1)}^{N_j}$