Do Predictors for Resource Overcommitment Even Predict?

Georgia Christofidi, Thaleia Dimitra Doudali
@EuroMLSys ‘24

IMDEA Software Institute, Madrid, Spain
1. **Traditional Resource Allocation**

Asks for \(X\) resources

\[
\text{Application allocated to a machine with capacity } K \geq X
\]

2. **Resource Allocation with Overcommitment**

Applications allocated to a machine with capacity \(K < X+Y+Z\)

- **Benefit**: Improve resource and cost efficiency
- **Risk**: Resource usage underestimations may cause competition for resources

Accurate predictions will increase benefit and reduce risk.

How? Predict resource usage!
Do Current Predictors Even Predict?

**Existing Predictors**

1. Borg
   - 90% * U

2. Resource Central
   - sum of the k-th %-ile

3. N-Sigma
   - U + N * std(U)

**Future Work**

- build a lightweight, practical predictor that **accurately predicts** and closes the existing gap in attainable resource savings.

**Our Insight**

- TITTL has an average prediction error of 568%. Prediction > Resource Limit **94%** of the times (No Overcommitment!)

- Maximum feasible resource savings (61%) with an **accurate** predictor

- Usable resource savings (2%)

**Future Work**

- Scan for code and paper: