womENcourage™ Computing Connecting Everyone TRONDHEIM NORWAY 2023

UiT Norges arktiske universitet

A Self-Configuration and Healing Controller To Analyze Misconfigurations of Clusters and IoT Edge Devices

Areeg Samir and Håvard Dagenborg UiT – The Arctic University of Norway

PROBLEM

Complex computational environments running modern online services and IoT applications are vulnerable to security breaches and information leakage due to misconfigurations.

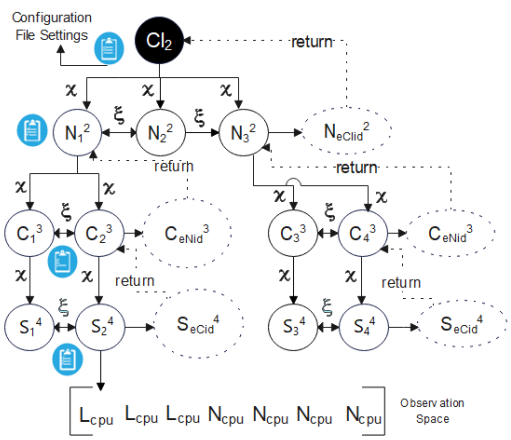
OBJECTIVE

Propose a self-configurable and healing controller that detects, identifies, and recovers from various misconfigurations in clusters and edge components.

METHODOLOGY

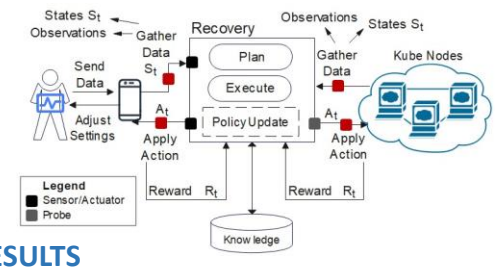
Classify misconfigurations and graphically map the hidden misconfiguration settings of the system to the observed performance metrics (CPU, Memory, Network) using HHMM. The hierarchy with the lowest probabilities is considered vulnerable misconfigured states.

For each misconfigured state, key-value pairs were extracted based on the Kind and API. The controller iterated through the settings to check each pair against centrally managed correct configuration settings given a confidence score. Improperly configured state setting is checked against (hidden) predefined misconfiguration type to identify the category of the detected misconfiguration. The controller records the new characteristics of the misconfiguration type.



Misconfiguration Path	$Cl_2 > N_2^2 > N_3^2 > C_3^3 > S_3^4$
Misconfigured Component	N_2^2
Misconfiguration Type	$Conf_{cc2}$

For identified settings, the controller selects an optimal recovery policy and observes the environment's performance after applying the action and receiving a reward. Successfully applied action indicates the recovery of misconfiguration. The controller continuously updates the policy to find an optimal policy until the difference between the update becomes marginal.



RESULTS

Models	Recall	Accuracy
HHMM	95.01%	94%
CRFs	92.86%	92%
AOR	97.66 %	

The controller is trained on some misconfigurations that allow privilege escalation to the host. The performance of the proposed detection is better than the CRF, with 95% recall and 94% accuracy. We observed that the length of observation sequences significantly impacts the detection accuracy.

The successfully recovered components to the total number of misconfigured components (AOR) are measured after multiple runs. The results stated that the controller performed better with increased training dataset size and observations.