Sharpening Kubernetes Audit Logs with Context Awareness

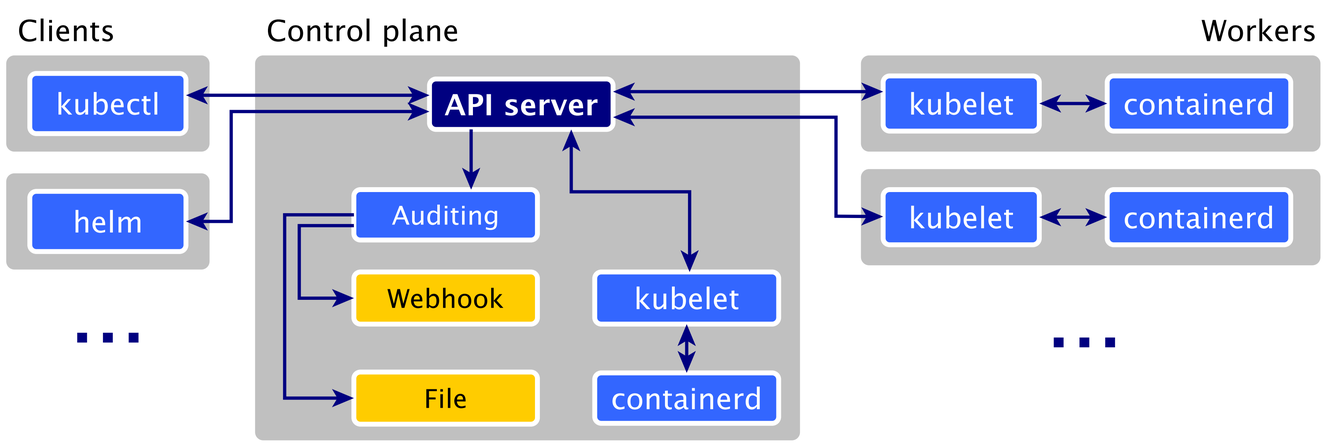
Matteo Franzil[1,2] matteo.franzil@unitn.it
Valentino Armani[2] varmani@fbk.eu
Luis Augusto Dias Knob[2] l.diasknob@fbk.eu
Domenico Siracusa[1,2] domenico.siracusa@unitn.it

[1] Department of Information Engineering and Computer Science, University of Trento, Italy
[2] Center for Cybersecurity, FBK - Fondazione Bruno Kessler, Trento - Italy

Introduction

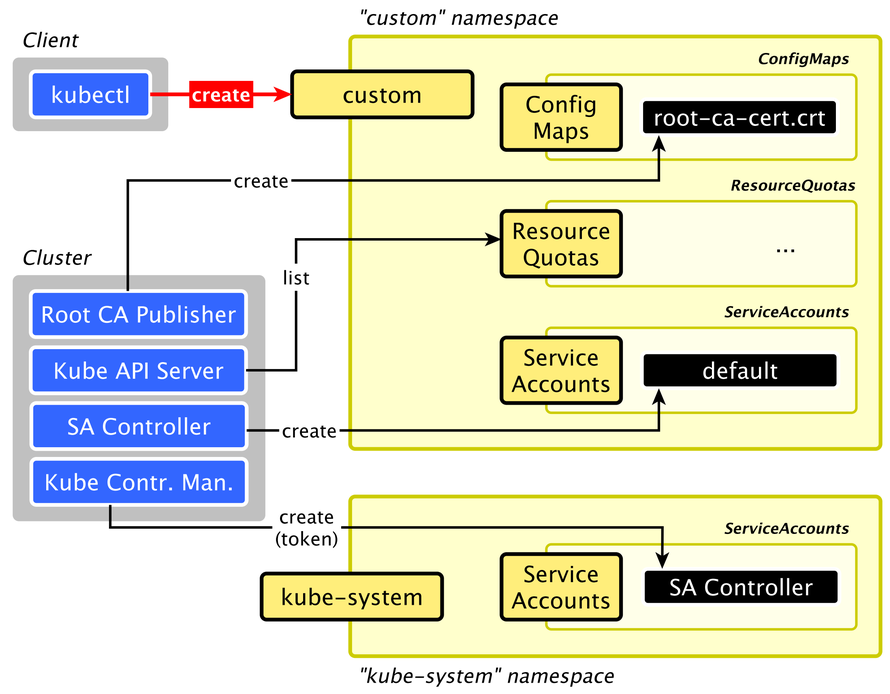
Kubernetes is the *de-facto* standard for container orchestration. Its **Auditing** component is a powerful tool for tracking API activity, providing a detailed and chronological record of every action performed in the cluster.
Audit logs are powerful, but very verbose and lack *context*: actions in the cluster are scattered and hard to correlate with each other. To address these setbacks, we introduce **K8NTEXT**.

Kubernetes Audit Logging



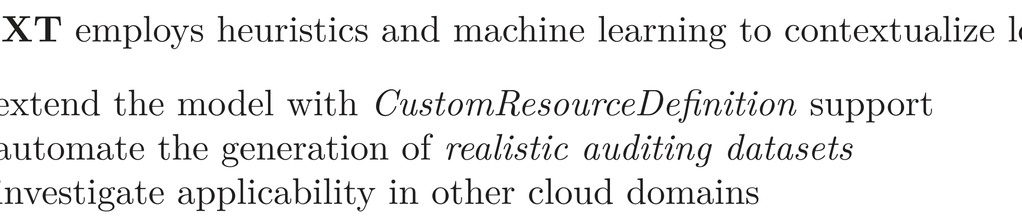
Every single **API call** is recorded in Kubernetes, from users listing Pods to leases being renewed. Kubernetes allows granular logging of each request's stages, parameters, and interactions. **How much of this information is really necessary?**

Creating a namespace

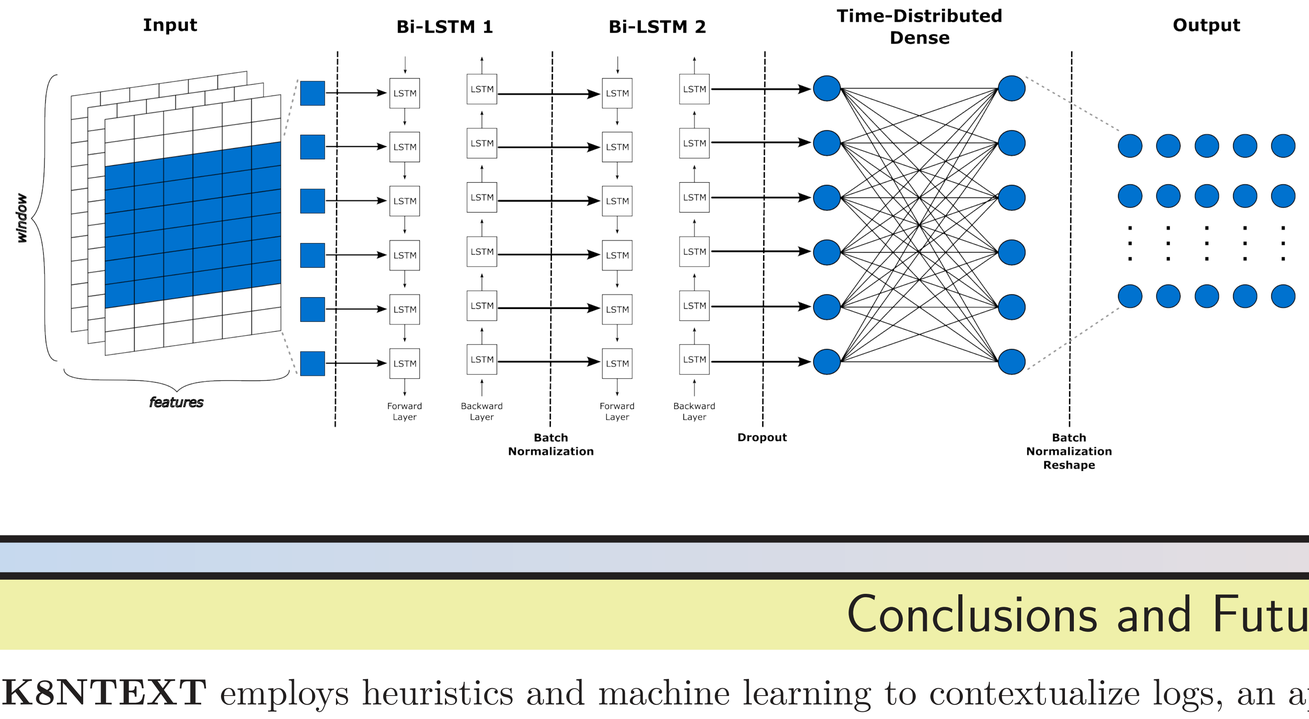


K8NTEXT

Labeling the actions:



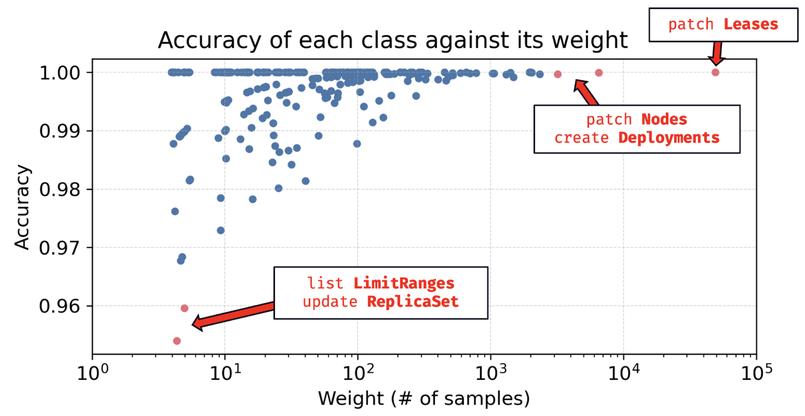
Predicting the labels:



Experimental results

W	Time (s)	F1 score
5	177.50 ± 27.45 s	0.8103 ± 0.0240
10	195.15 ± 42.63 s	0.9267 ± 0.0173
20	240.10 ± 41.85 s	0.9681 ± 0.0093
30	238.35 ± 36.24 s	0.9791 ± 0.0099
40	261.55 ± 51.51 s	0.9808 ± 0.0068
50	300.70 ± 43.77 s	0.9866 ± 0.0047
60	319.60 ± 43.94 s	0.9820 ± 0.0082

Performance of the model with different window sizes.



We constructed a realistic, yet unbalanced dataset. Underrepresented classes suffered from worse accuracy, but still in an acceptable range.

Conclusions and Future Work

K8NTEXT employs heuristics and machine learning to contextualize logs, an approach rarely seen in the field so far. We plan:

- to extend the model with *CustomResourceDefinition* support
- to automate the generation of *realistic auditing datasets*
- to investigate applicability in other cloud domains