Confidential Containerized Federated Learning for Distributed Security

Md Muzammal Hoque, Andrea Gentili, Ijaz Ahmad, Andrea Dalla Costa, Jani Suomalainen, Markku Kylänpää
VTT Technical Research Centre of Finland
firstname.lastname@vtt.fi

- Federated learning is a prominent AI solution for private critical data, but vulnerable to malicious aggregating nodes.
- We demonstrated two privacy-enhancing approaches: Differential Privacy and Trusted Execution Environments.
- Their costs for resource consumption and AI accuracy were explored with Regular and Confidential Containers on the Intel and ARM platforms.

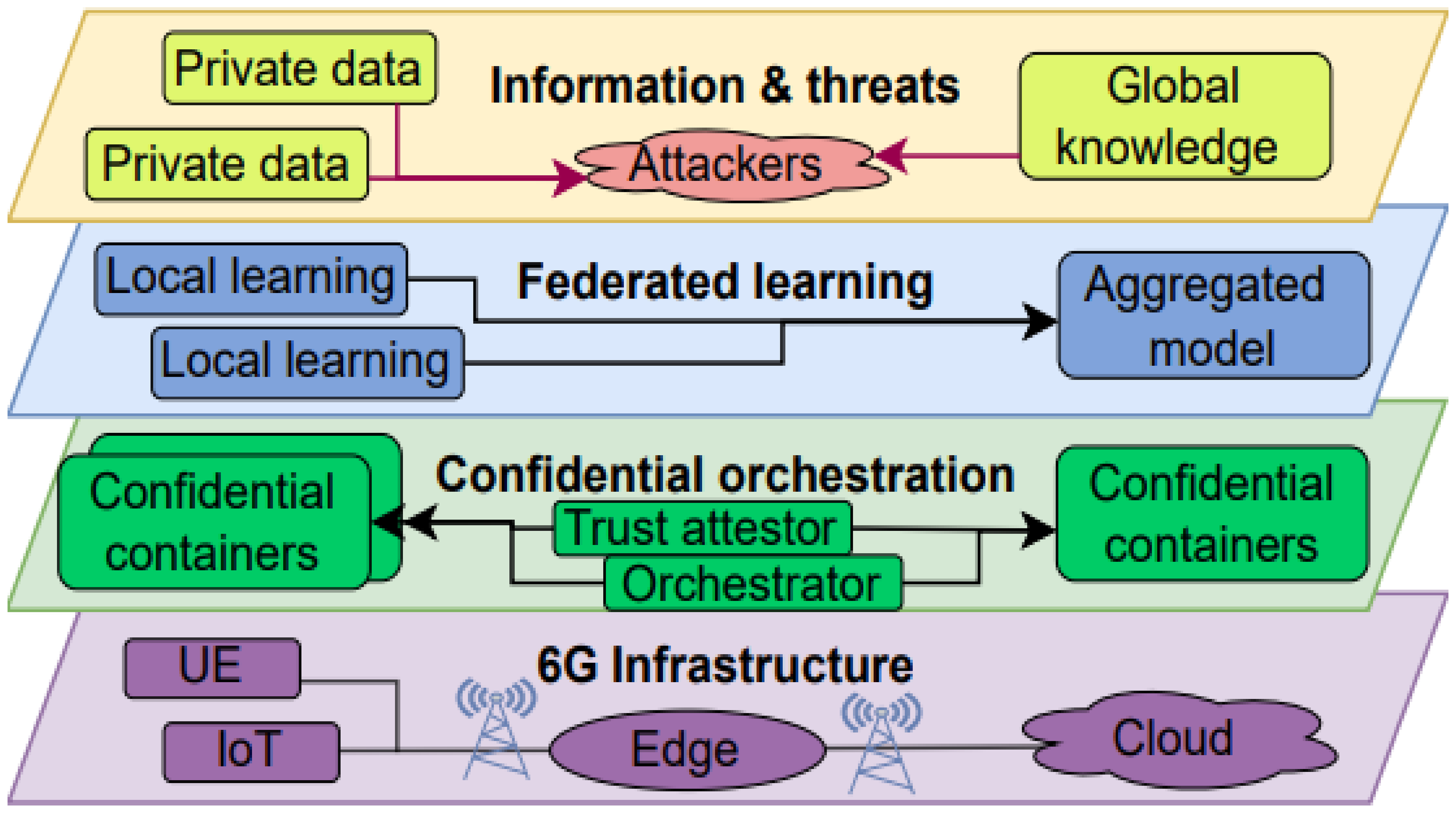

Fig 1: A conceptual model of containerized confidential learning on 6G outlines infrastructure, orchestration, learning, and information threats in distinct layers

Experimental Setup

- The experiment was conducted using FederatedScope [1] FL platform.
- The experiment involves an FL server and 3 distributed clients; each deployed as a container on a Kubernetes cluster that included 3 worker nodes: 2 Raspberry Pi 4B and 1 VM running in our CyberRange. The cluster is configured to use both docker containers or Confidential Containers. The FL server and client 1 were deployed in the VM, client 2 and 3 on one Raspberry Pi each.
- The work utilized a two-layer CNN model (ConvNet2) for image classification with the MNIST [2] image dataset.
- For attacking the FL model, we used Improved Gradient Leakage attack (iDLG) [3].

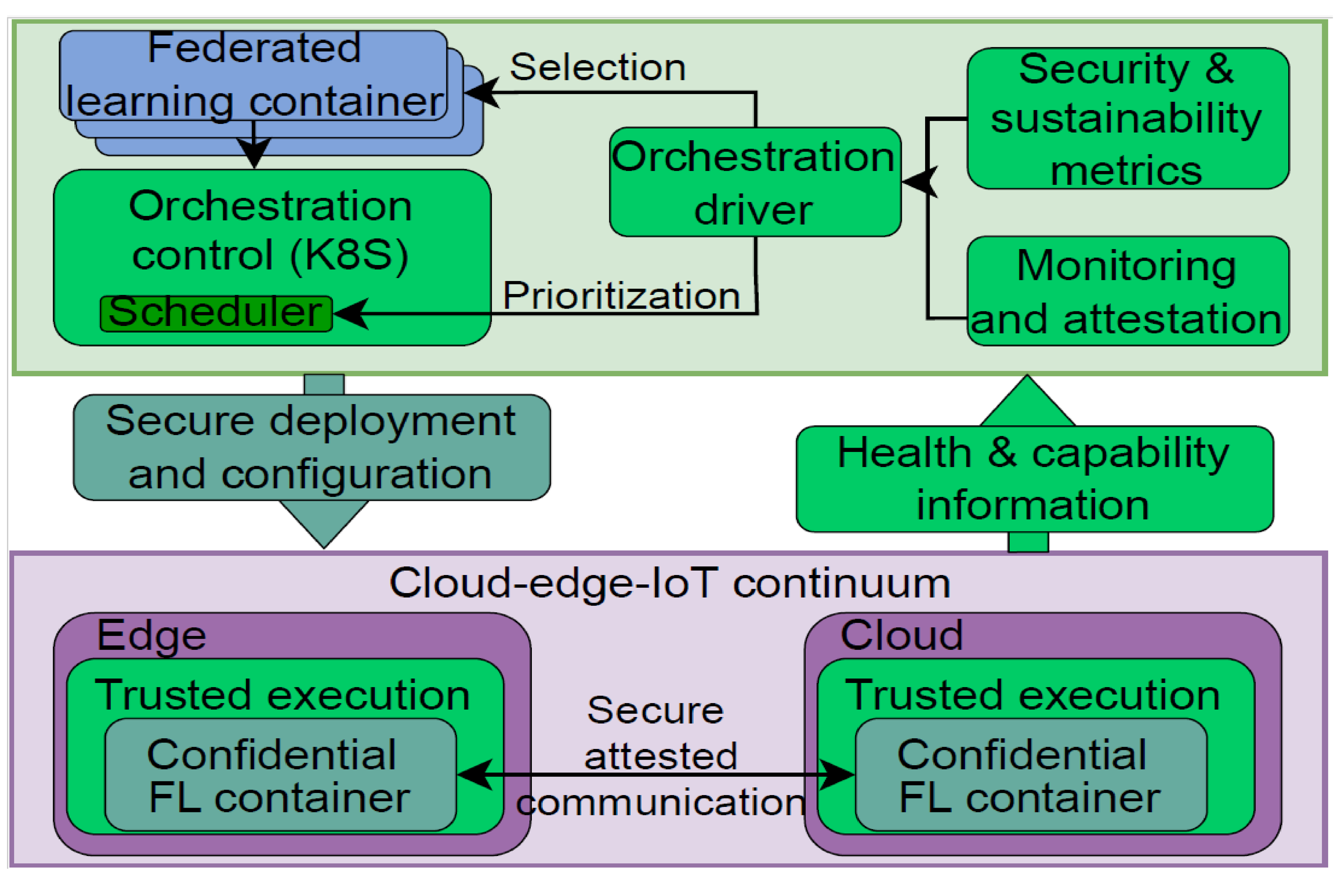

Fig 2: An architecture for the security and sustainability-driven concept

[1] Xie, Y., Wang, Z., Gao, D., Chen, D., Yao, L., Kuang, W., Li, Y., Ding, B., & Zhou, J. (2023). FederatedScope: A flexible federated learning platform for heterogeneity. Proceedings of the VLDB Endowment, 16(5), 1059–1072. https://doi.org/10.14778/3579075.3579081.
[2] L. Deng, "The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web]," in IEEE Signal Processing Magazine, vol. 29, no. 6, pp. 141-142, Nov. 2012, doi: 10.1109/MSP.2012.2211477.
[3] Ding, X., Liu, Z., You, X., Li, X., & Vasilakos, A. V. (2024). Improved gradient leakage attack against compressed gradients in federated learning. Neurocomputing, 608, 128349.

Resource Consumption and Performance Analysis

Device	Train Accuracy	Val Accuracy	Test Accuracy	Power (mWh)*	Time (Min)	Peak Memory (MB)
Baseline containerized FL						
Client 1	1	.94	.99		15.93	680.35
Client 2	1	.92	.95	775	15.22	548.55
Client 3	1	.96	.93		15.00	548.03
Server	-	.94	.96		17.01	629.58
Differential Privacy (NBAFL, epsilon: 10, mu: 0.01)						
Client 1	1	.95	.99		18.53	699.30
Client 2	1	.94	.94	973	17.80	549.19
Client 3	1	.94	.93		17.57	549.12
Server	-	.95	.93		19.98	632.90
Confidential Containerized FL						
Client 1	1	.94	.99		16.41	683.80
Client 2	1	.96	.93	786	14.93	541.66
Client 3	1	.92	.95		16.50	541.66
Server	-	.94	.96		17.55	627.05

* Power consumption was measured with Otii Arc Pro

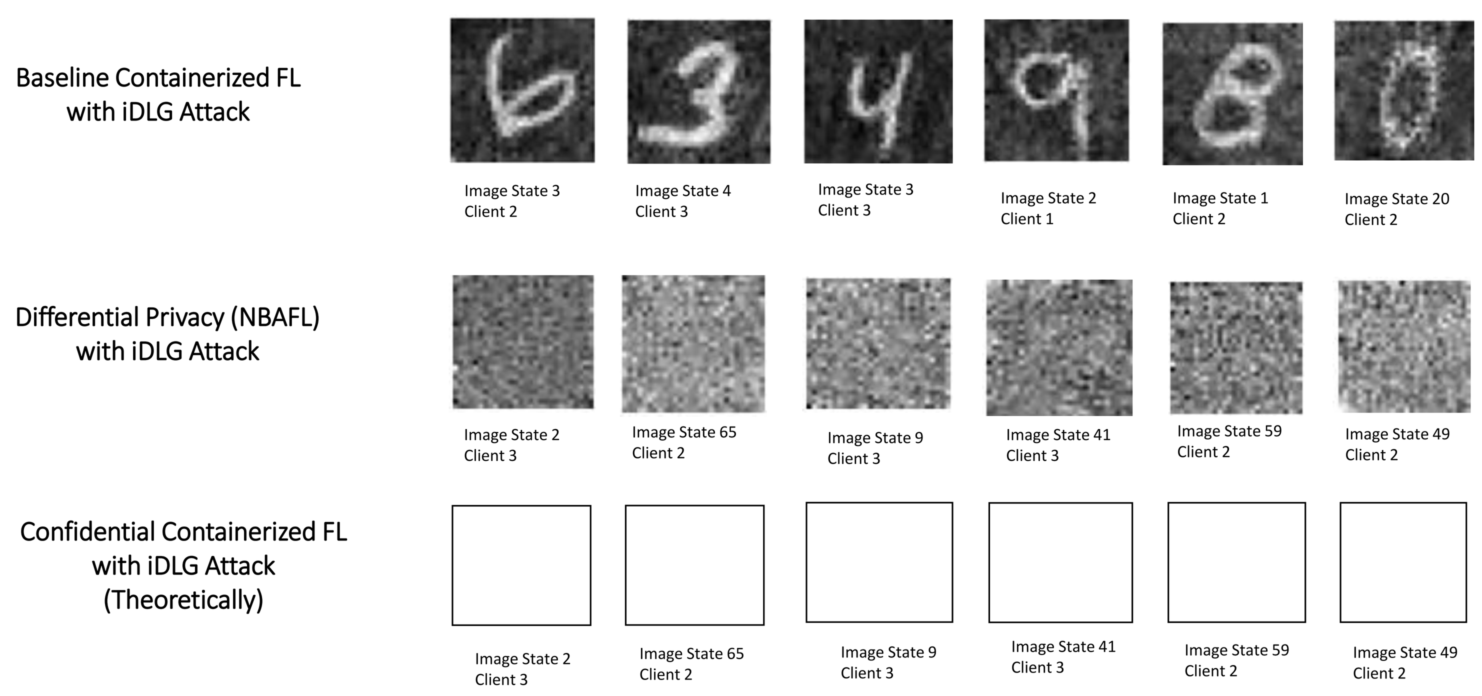

Fig 3: iDLG Attacks on Baseline, DP, and TEEs

Takeaways

- **CoCo excels by providing application agnostic security without impacting ML accuracy but comes with small memory, processing and time penalty**
- **DP works in any HW platform but impacts slightly accuracy and significantly energy use**
- **Overall, deploying FL in CoCo offers a balanced trade-off between sustainability and security.**

beyond the obvious

www.vttresearch.com

Contact: Muzammal Hoque, Andrea Gentili, and Andrea Dalla Costa
Email: firstname.lastname@vtt.fi

Acknowledgements: Confidential 6G, SUNSET-6G and VTT are thanked for funding the work.