

# Cross-Layer Impact Characterization of Faults in Cloud-Native 5G Core Networks

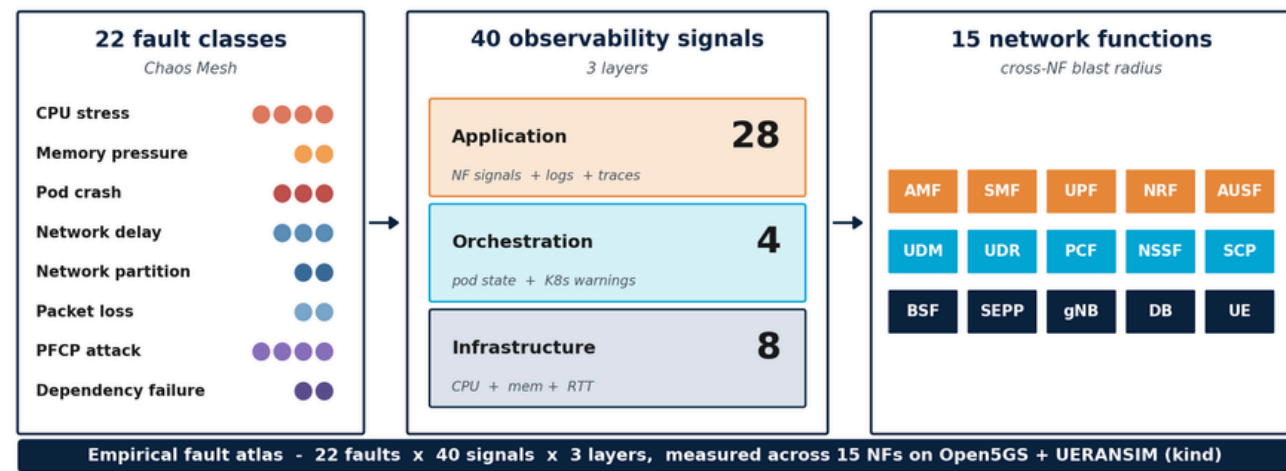
Boyan Bonev (5950856, b.i.bonev-1@student.tudelft.nl)

Supervisor: Sehan Samarakoon Mudiyansele, Responsible professor: Nitinder Mohan, Examiner: Dr. Jérémie Decouchant

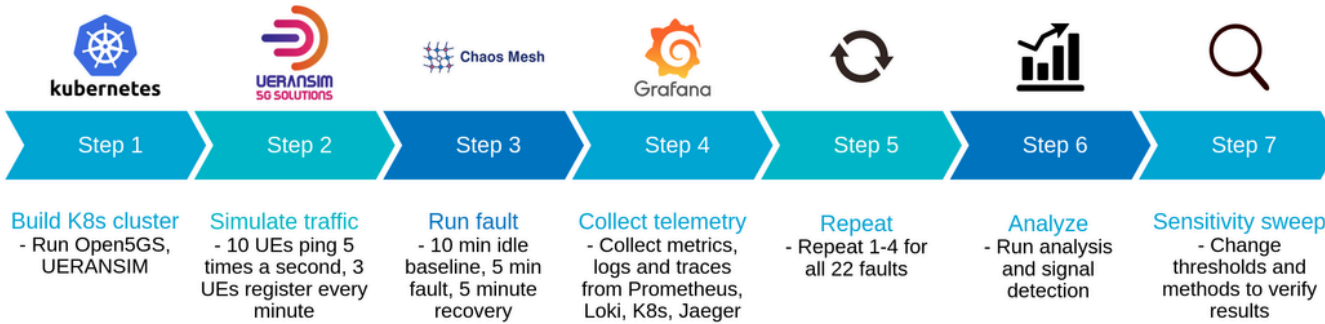
Final poster - 17 June 2026

## Background & Motivation

- 5G Cores are going cloud-native → cross-layer faults
- Microservice fault taxonomies (Zhou 2018, Silva 2022) do not show where the faults manifest and are not 5G oriented
- 5G specific work only measures degradation in signal quality under stress (Moreira 2025)
- Contribution - empirical fault atlas mapping:



## Experimental Setup



## Signal Detection Methodology

### Z-rule: metric signals

fired when both conditions hold for  $\geq 2$  consecutive (rolling) windows

$$|\bar{x}_w - \mu| > \max(Z\sigma, \rho|\mu|, f, \varepsilon)$$

AND (ratio guard, when set)

$$|\bar{x}_w| \geq r|\mu| \text{ OR } |\bar{x}_w| \leq |\mu|/r$$

Sensitivity sweep: parameters varied (baseline in bold)

Z	z-score threshold	2.5	<b>3.0</b>	3.5
persist	consecutive anomalous (rolling) windows	1	<b>2</b>	3
window	rolling-window length	15s	<b>30s</b>	45s
f	per-signal absolute floor (scaled)	$\times 0.5$	$\times 1$	$\times 2$
r	per-signal ratio guard (scaled)	$\times 0.5$	$\times 1$	$\times 2$
$\rho$	relative-size guard $\rho \mu $ ( $\rho = 0.30$ , scaled)	$\times 0.5$	$\times 1$	$\times 2$

Atlas conclusions remain unchanged under every perturbation ( $\geq 97\%$  cell agreement, first-layer attribution is kept). Replacing the mean/ $\sigma$  baseline with median/MAD leaves 95.9% of cells unchanged.

## References

- [1] X. Zhou et al., "Fault analysis and debugging of microservice systems: Industrial survey, benchmark system, and empirical study," *IEEE Trans. Softw. Eng.*, vol. 47, no. 2, Feb. 2021, doi: 10.1109/TSE.2018.2875312.
- [2] F. G. S. Filho, V. Lelli, I. de Sousa Santos, and R. M. C. Andrade, "Towards a fault taxonomy for microservices-based applications," in *Proc. XXXVI Brazilian Symp. Software Engineering\**, 2022, doi: 10.1145/3555228.3555245.
- [3] J. Flora, P. Gonçalves, M. Teixeira, and N. Antunes, "A study on the aging and fault tolerance of microservices in Kubernetes," *IEEE Access*, vol. 10, pp. 132786–132799, 2022, doi: 10.1109/ACCESS.2022.3231191.
- [4] N. Fu, G. Cheng, G. Dai, H. Mei, X. Qiu, and Y. Teng, "A failure analysis framework to provide pure anomalous data using multi-source data of fault-sensitive microservices," *J. Syst. Softw.*, vol. 230, art. 112513, 2025, doi: 10.1016/j.jss.2025.112513.
- [5] R. Moreira, L. F. R. Moreira, and F. de Oliveira Silva, "Performance evaluation and threat mitigation in large-scale 5G core deployment," *arXiv:2507.17850 [cs.CR]*, 2025.
- [6] D. A. Abongo, A. Kehila, S. Chérif, R. Bouraoui, and Z. Labidi, "A new framework for failure classification in 5G core network," in *Proc. IEEE ComNet\**, 2024, doi: 10.1109/COMNET64071.2024.10987290.

## Results

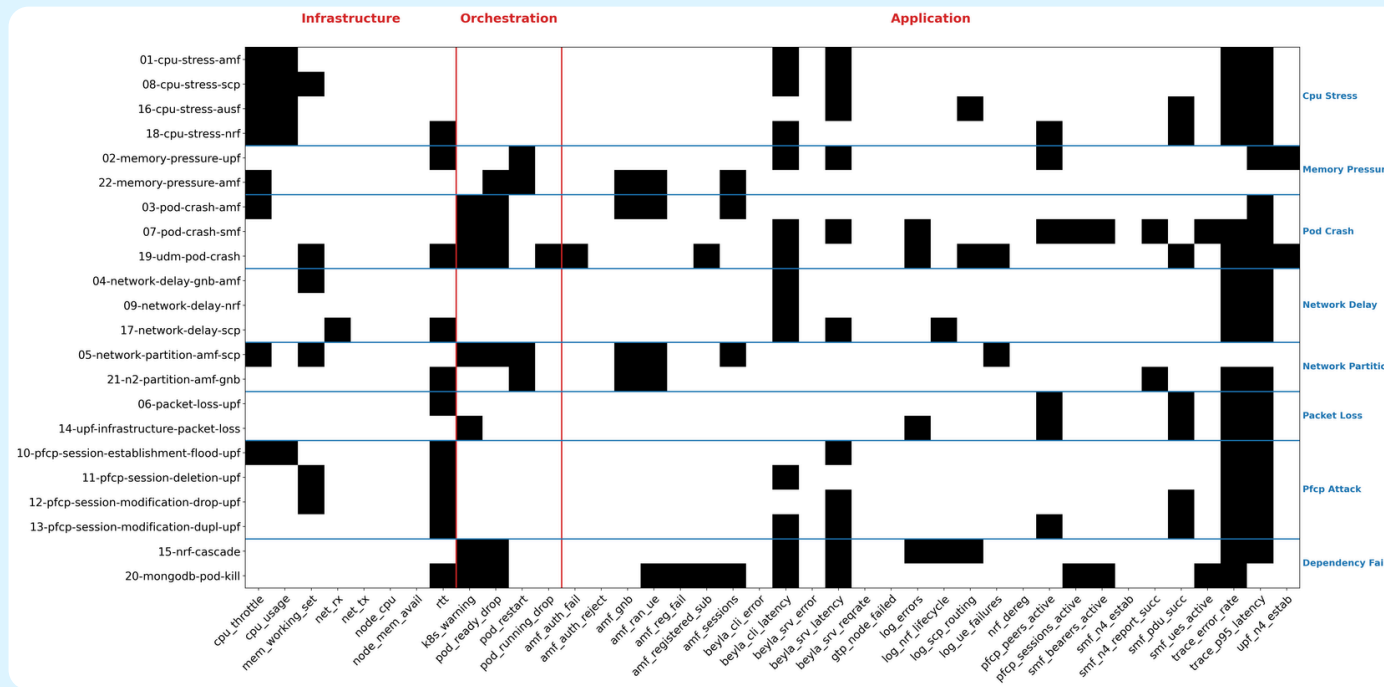


Figure 1: The fault atlas: which of the 40 observability signals (columns, grouped by layer) each fault (rows, grouped by class) manifested (black). Vertical red lines separate the three layers, and horizontal blue lines separate fault classes.

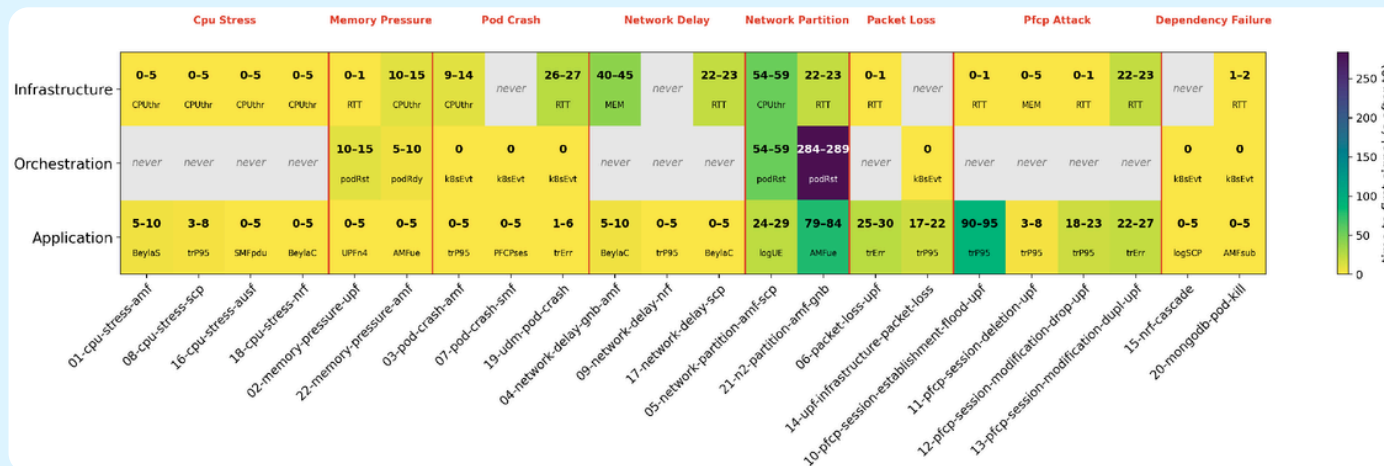


Figure 2: Time to first signal per layer (seconds after t0) for each fault. Each cell gives the onset bucket and first signal, and "never" marks a layer with no signal. Resolution is 5 s for Prometheus/Loki/Jaeger/pod gauges, 1 s for RTT, and sub-second for Kubernetes events.

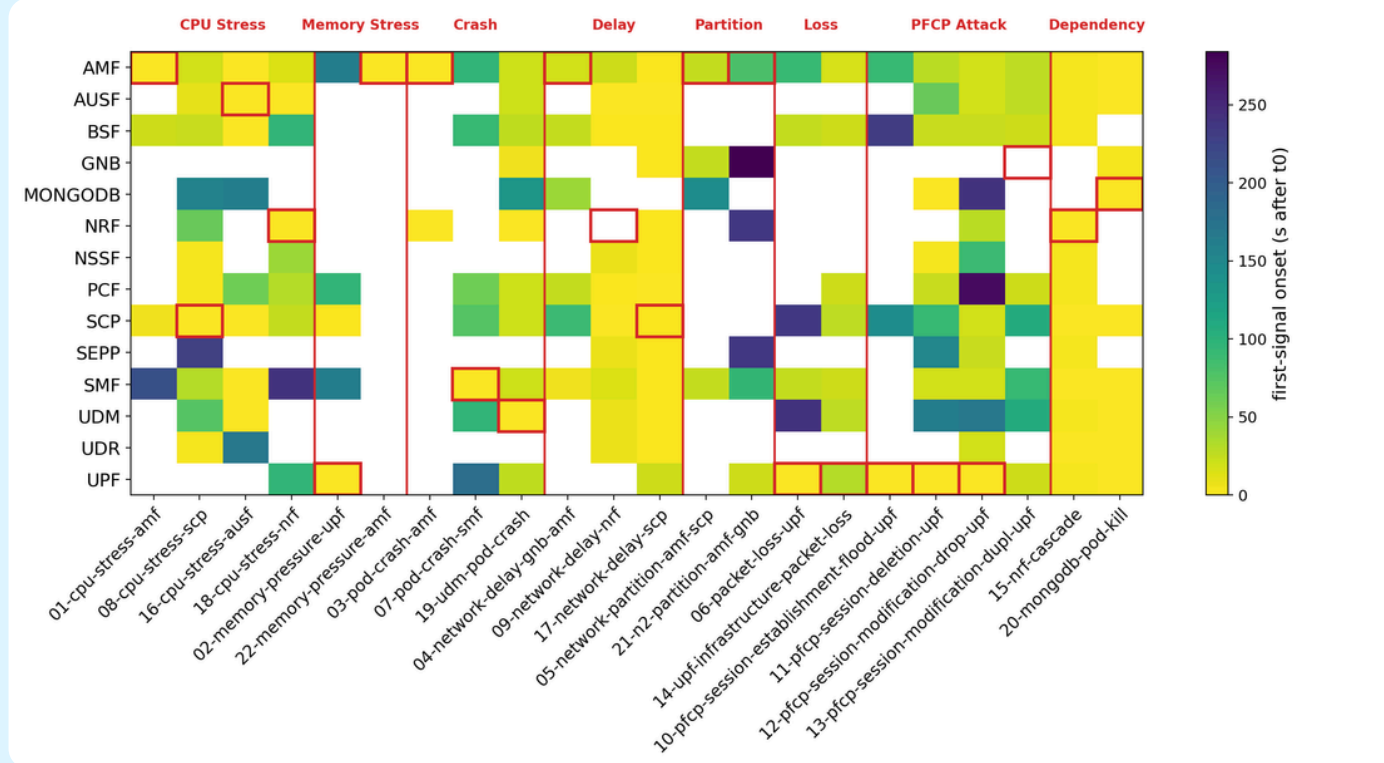


Figure 3: Network-function blast radius: first-signal onset (seconds after t0, colour) for every function under every fault. White marks an

## Second Run Verification

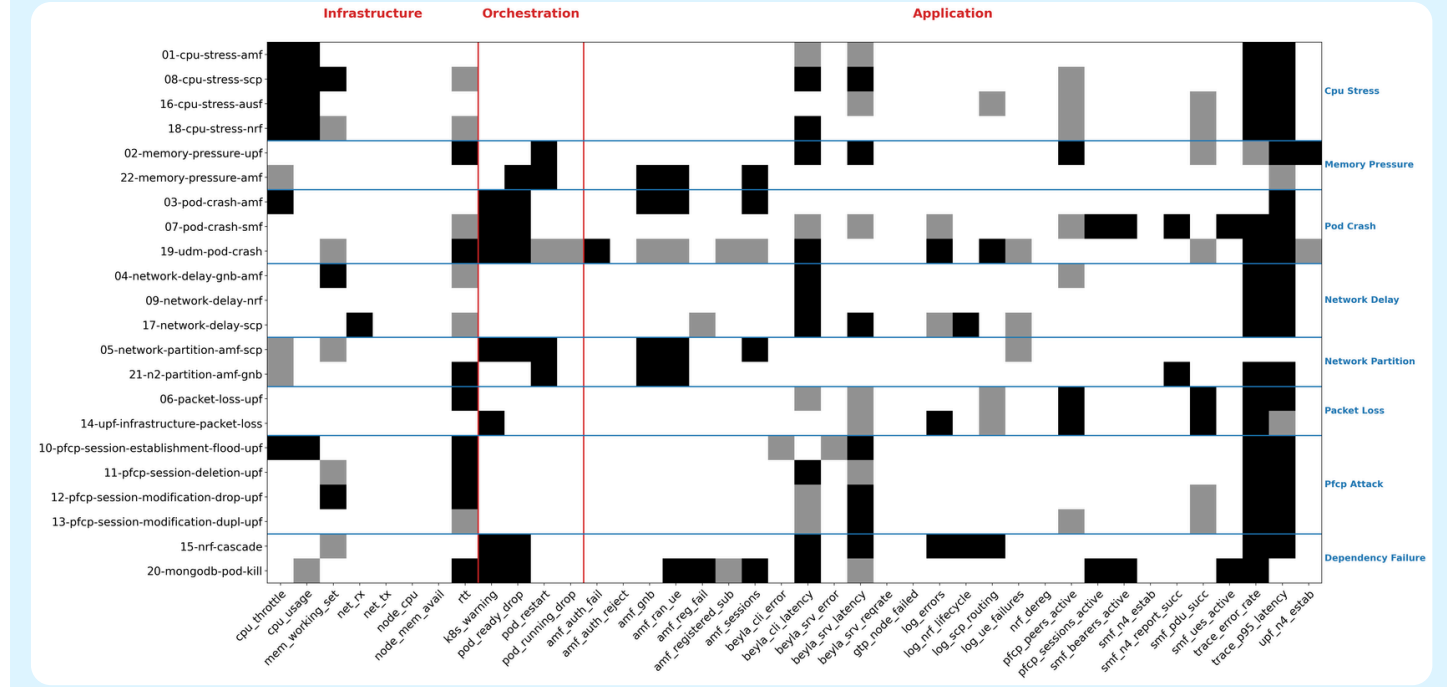


Figure 4: Cross-run reproducibility atlas (canonical run versus the independent second run). Cell shade is the fraction of the two runs in which each (fault, signal) cell fired: black = both runs, grey = one run only, white = neither. Red vertical lines separate the layers and blue horizontal lines separate fault classes, as in Figure 1. The two runs agree on 93.1% of cells (Cohen's  $\kappa = 0.768$ ), and the grey cells are the balanced, near-threshold disagreements.

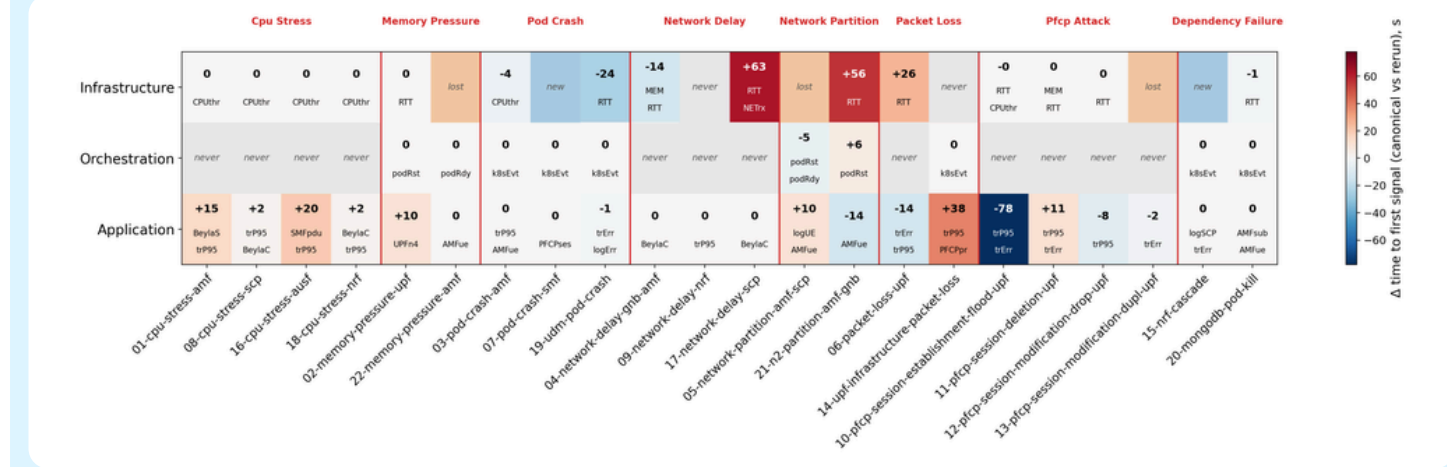


Figure 5: Cross-run timing reproducibility: per fault (column, grouped by class) and layer (row), the signed difference in time-to-first-signal between the second run and the canonical run,  $\Delta = t_{\text{second}} - t_{\text{canonical}}$ . White is an identical onset, blue means the second run was faster and red slower; lost marks a layer that fired only in the canonical run and new one that fired only in the second run, while never is silent in both. Each cell gives  $\Delta$  in seconds and the first signal (both runs' signals are stacked when they differ). The layout matches Figure 2. Most cells are near zero and the disagreements concentrate on the same near-threshold RTT and trace signals as the cross-run atlas.

## Findings & Discussion

- Faults are cross-layer - 21/22 faults are detected in  $>1$  layer
- Fault classes have a stable layer-firing order
- Orchestration layer detects only 10/22 faults (the ones leading to crashes), confirming [3]
- AMF is most affected NF, firing a signal in all 22 faults
- The atlas is robust under changing thresholds, estimators and a second run
- Future work: repeat the experiments 10+ times to establish per-cell confidence, reproduce on a different core to confirm findings and generate higher user and control plane load to simulate realistic traffic