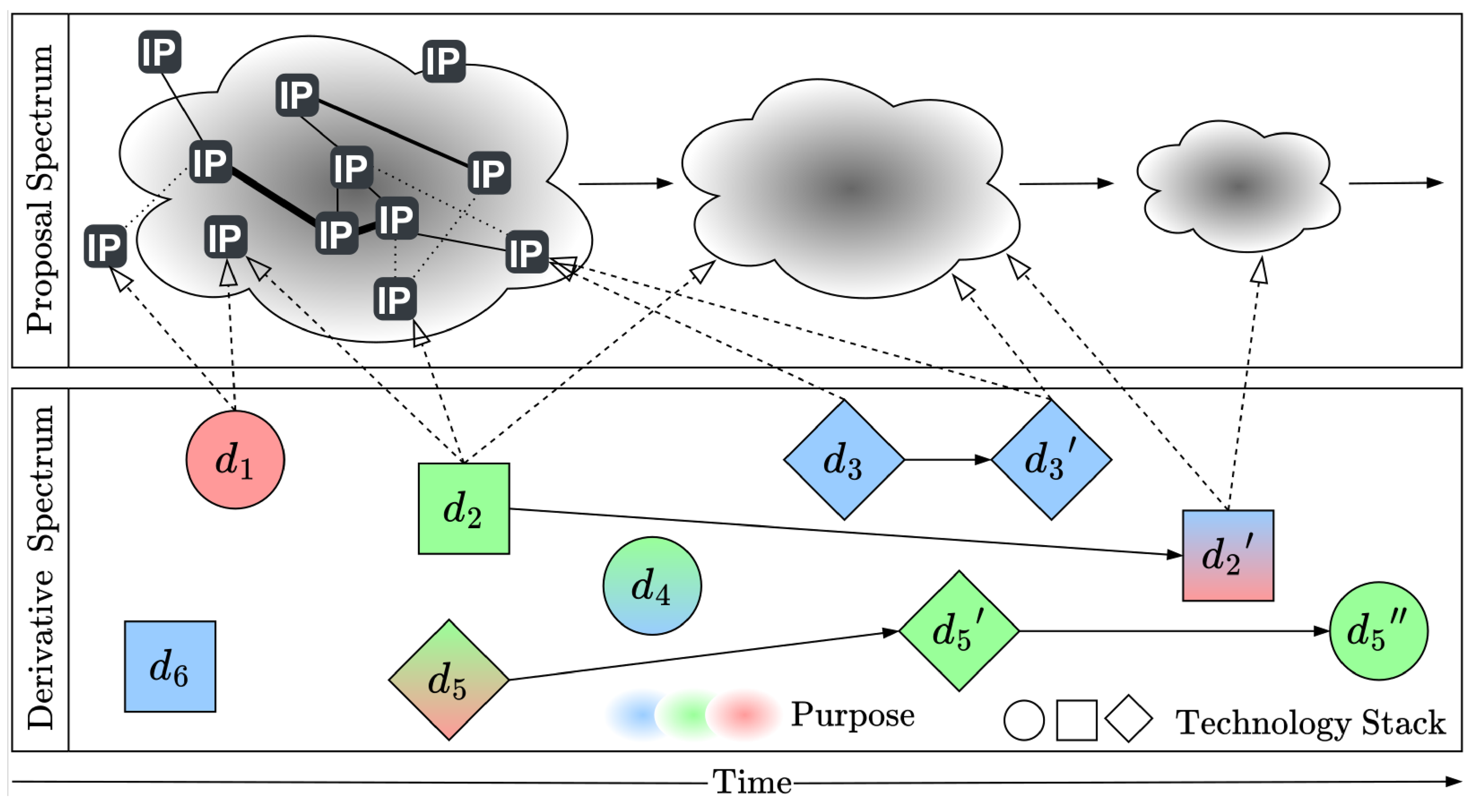
Towards Large-Scale Variability Investigations in Nostr

Michael Kaiser, Software Engineering Group (SEG), University of Bern, Switzerland

Introduction



Within this project, we laid the groundwork to investigate a recently described software variability paradigm termed Community-Driven Variability (CDV) [Bö25] in the decentralized messaging relay system called Nostr (Notes and Other Stuff Transmitted by Relay) [nob]. In contrast to classical variability-intensive systems such as, e.g., software product lines (SPLs) [Ap13; Ke21], CDV exhibiting ecosystems are not centrally managed to guide evolution. Instead, the focus lies on achieving interoperability within the software family through the continuous effort to shape an open set of specification documents, referred to as improvement proposals or implementation possibilities (IPs). Based on this set of IPs, developers within the community independently derive their own software variants by selecting and implementing a desired subset of IPs. The variability induced in CDV ecosystems consequently unfolds across additional dimensions compared to classical software family paradigms, reaching beyond features and configurations to entirely different software target use cases and even different technology stacks.

Nostr

Nostr [nob] is a decentralized social-networking protocol built around relays that distribute user messages. Designed to stay as generic as possible – some even frame it as a “social layer” in the OSI stack – it has grown since 2020 into a lively ecosystem of relays, clients, and services. Today it powers Twitter/X alternatives, blogs, chats, and even micro-payment tools [al25; ap25]. As Nostr relies on asymmetric cryptography, users can access all these heterogenous applications with a single key pair. The functioning of the protocol itself, including mandatory and optional features within it, is defined in so-called Nostr Implementation Possibilities (NIPs) [noa], which are maintained by independent contributors.

Research Goals

1) Deploy our own Nostr relay and announce it to the network.

   Motivation: *Acquire better Nostr understanding, provide clean and controllable implementation, capable of fast adaption to new emerging challenges.*

2) Leverage the relay as a watchtower to collect network data, Nostr messages, and foreign relay profiles.

   Motivation: *Have a long-term data collection and analysis platform to study short-lived or rotated data/logs.*

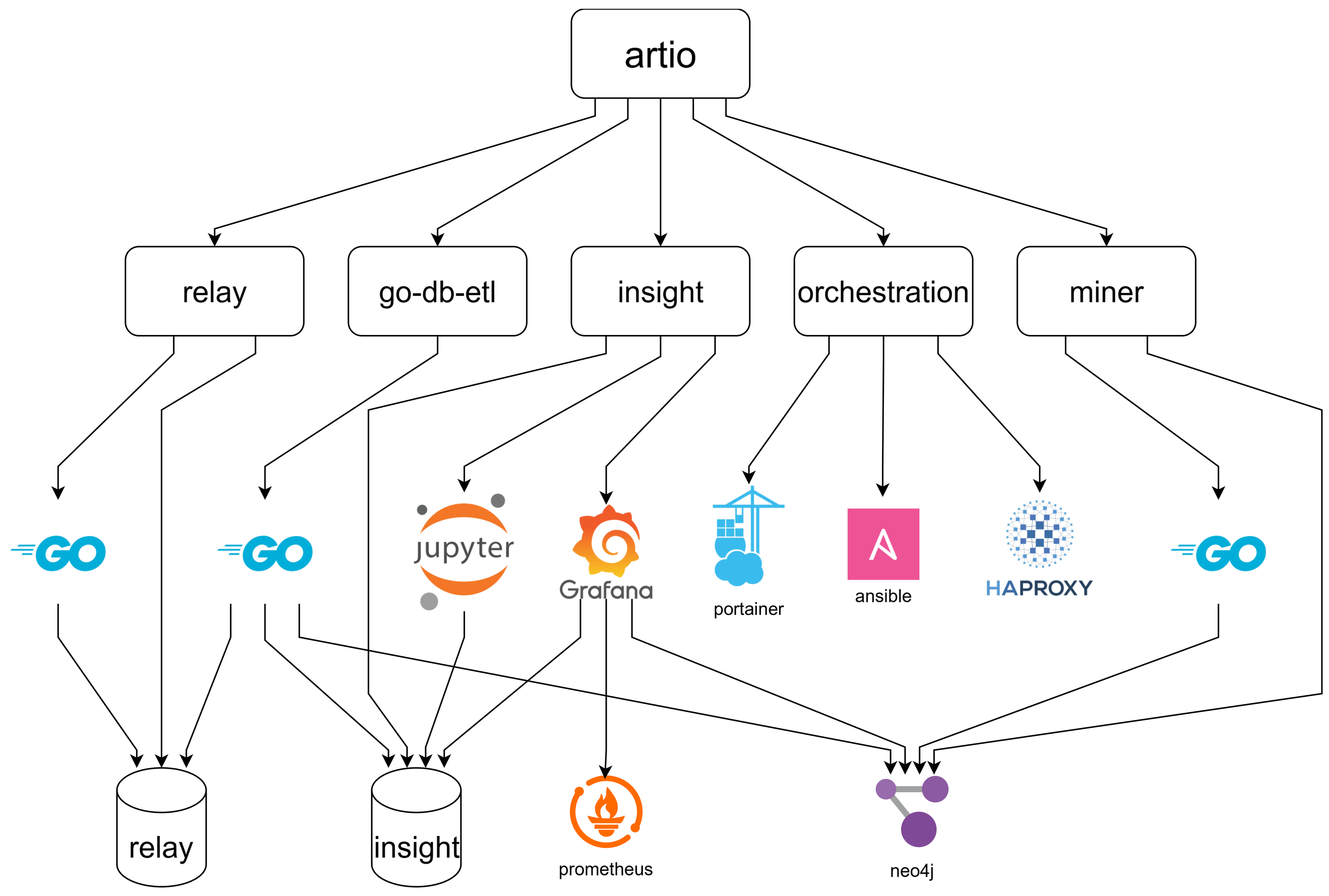
3) Target interesting interoperability facets of the protocol and the network becoming visible through the watchtower.

   Motivation: *Be capable to target new facets quickly due to the rapid evolving nature of the ecosystem.*

Methodology

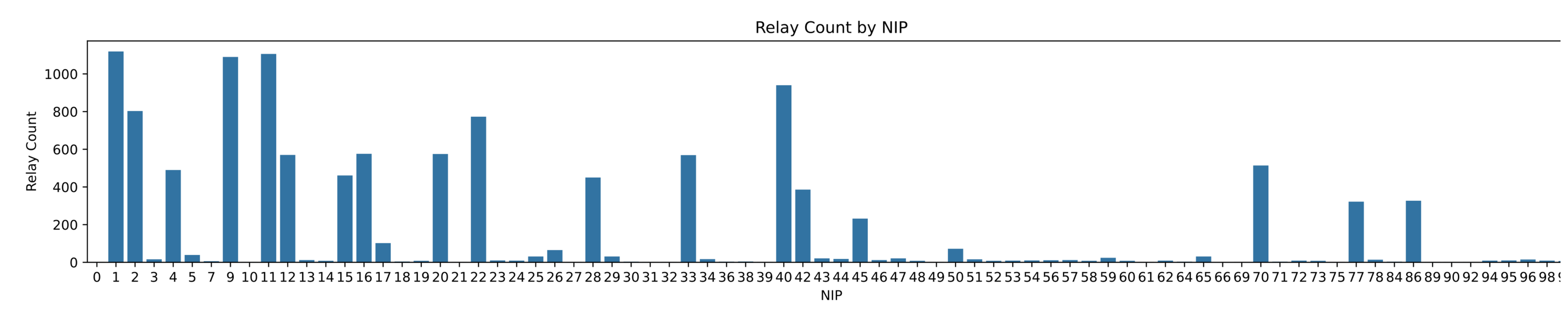
Throughout the groundwork and the emerging follow-up work the architecture visible next, has emerged and proven to be a meaningful solution to analysing large scale variability. It contains (i) a precise view of the minimal NIP set required for the relay communication and (ii) a controllable extensible basis for experimenting with additional NIP features relevant to future studies.

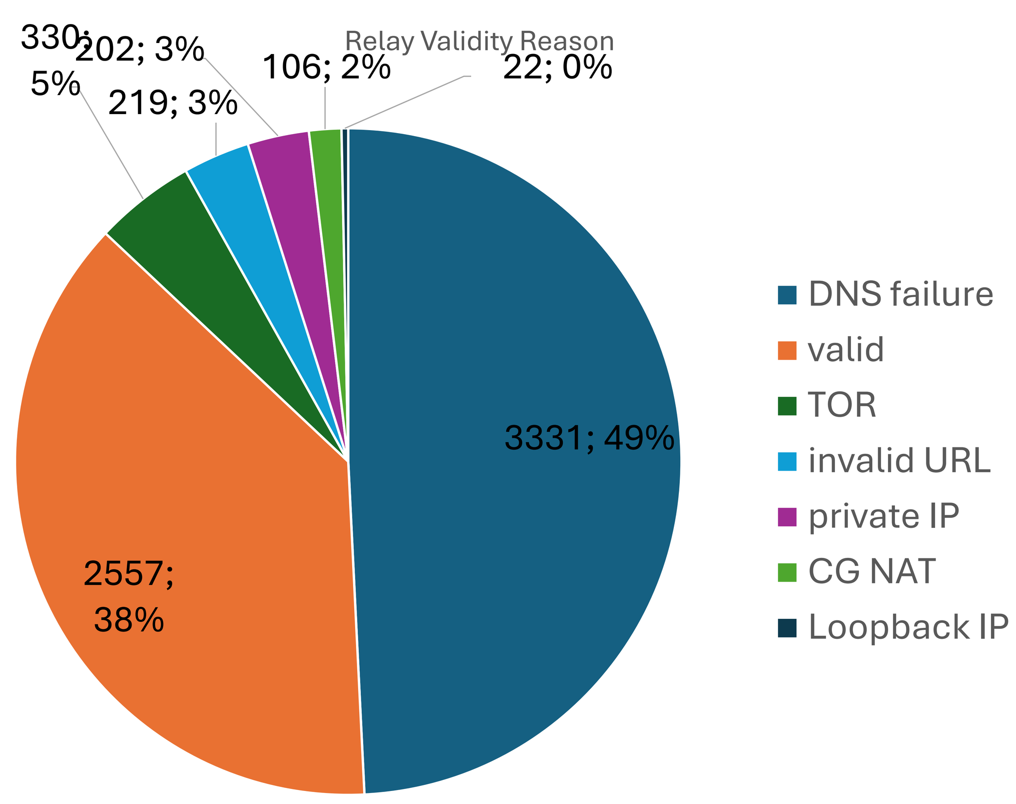
Secondly, we designed a data vault [DV] that periodically harvests all traffic handled by the relays together with telemetry data in a time-series database (Prometheus). These data sets are then used within the analysis platform called artio-insight. The main visualisation component is Grafana with other the raw database schema designed in a way that enables other forms of visualizations such as Jupyter notebooks.

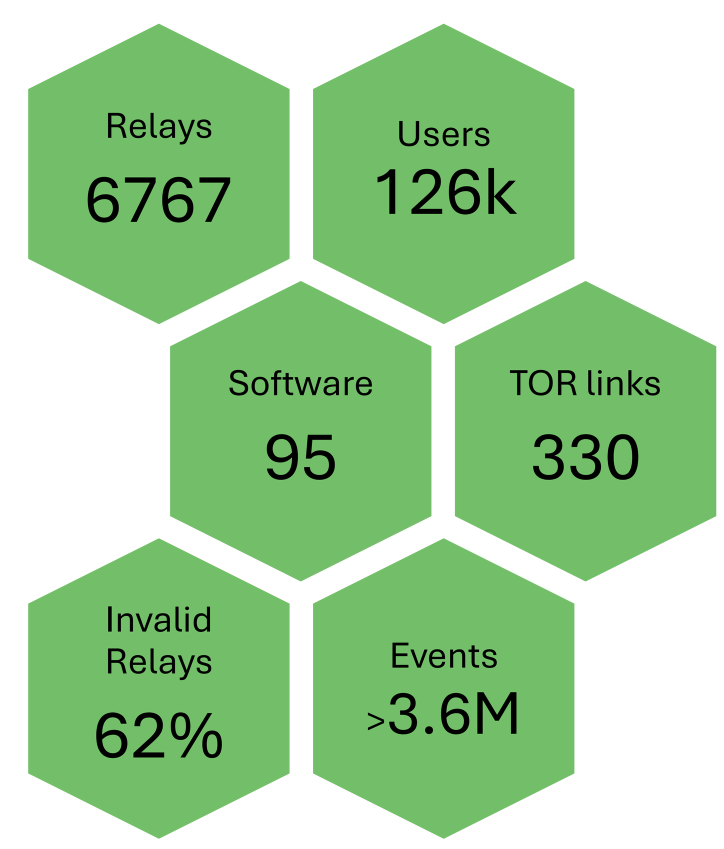


Results

First results are currently under active analysis. Early findings uncovered that protocol violations are not a rarity but rather a frequent occurrence. They manifest in different ways depending on the NIP that we are looking at. For example, only about 40% of the relays announced by NIP-65 messages are valid relays considering local networks, DNS resolution failures and many more. So far, we could discover 2’605 valid relays throughout the ecosystem operating on over 90 different software stacks. Overall, we were able to uncover around 126k unique public keys that represent the users of Nostr.







Future Work

This project established the conceptual and technical foundations for investigating CDV in the Nostr ecosystem at scale. The ensuing Master’s thesis will advance these efforts in three directions:

i. extending the presented tools and conducting a broader investigation of the Nostr relay implementation spectrum,

ii. analysing network topology to identify user and relay clustering based on NIP adoption, and

iii. investigating interoperability issues arising from NIP-driven implementation freedom and developing suitable countermeasures to mitigate them.

[Bö25] Bögli, R. et al.: Beyond Software Families: Community-Driven Variability. In: FSE Companion '25, S. 571–575, 2025, https://doi.org/10.1145/3696630.3728501.
[ap25] nostr apps: Explore Nostr Apps, 2025, https://nostrapps.com/, Stand: 14. 07. 2025.
[al25] aljazceru: Aljazceru/Awesome-Nostr, 2025, https://github.com/aljazceru/awesome-nostr, Stand: 14. 07. 2025.
[Ap13] Apel, S. et al.: Feature-Oriented Software Product Lines. 2013.
[noa] nostr-protocol: Nostr Implementation Possibilities (NIPs), https://github.com/nostr-protocol/nips, Stand: 04. 10. 2024
[nob] nostr-protocol: Nostr-Protocol/Nostr, nostr-protocol, https://github.com/nostr-protocol/nostr, Stand: 16. 12. 2024.
[Ke21] Kehrer, T. et al.: Bridging the Gap Between Clone-and-Own and Software Product Lines. In. S. 21–25, 2021.
[DV] Linstedt D., Olschinke M. Building a scaleable Data Warehouse with Data Vault 2.0 SBN: 978-0-12-802510-9.