Continuous Microservice Granularity Management with Saturation Signals in an Industrial Context

Yan Justino[1], **Rafael Batista Duarte**[1], **Carlos Eduardo da Silva**[2]

[1]Centro de Estudos e Sistemas Avançados do Recife (CESAR)
Av. Cais do Apolo, 77 - 50030-220 - Recife – PE – Brazil

[2]School of Computing and Digital Technologies - Sheffield Hallam University
Sheffield – UK

{ylj,rbd}@cesar.school, C.DaSilva@shu.ac.uk

Abstract. *This paper presents an industry experience with Granulify, an evidence-based approach for continuous microservice granularity management applied during a 13-month modernisation initiative at a large financial institution in Brazil. The approach integrates the Granularity Classification Spectrum for qualitative boundary diagnosis and the Granularity Saturation Method for quantitative assessment. The longitudinal study revealed that periods without systematic management exhibited over-decomposition and architectural degradation, while Granulify enabled the team to detect saturation signals and perform strategic consolidation with measurable improvements.*

1. Context

Properly defining the size and scope of a microservice remains one of the main challenges faced by development teams [Bogner et al. 2021, Hassan et al. 2020]. Granularity decisions directly impact modularity, performance, and maintainability. Inadequate decomposition can lead to an excess of components, intensifying operational complexity, coordination effort, and failure probability [Behrad et al. 2021].

This challenge aligns with the Grand Research Challenges in Information Systems in Brazil (GranDSI-BR 2016–2026) [Araujo et al. 2017], under the themes of *IS Complexity* and *IS Development*. Interoperability among distributed systems introduces complexity dimensions that require systematic approaches for continuous architectural management.

In practice, teams face recurring concerns: (i) decisions based on subjective judgement without tooling support [Driessen et al. 2024]; (ii) the need for dynamic, evidence-based assessment [Hassan et al. 2020]; (iii) a scarcity of operationalisable and repeatable procedures [Vera-Rivera et al. 2021]; and (iv) limited longitudinal industrial evidence [Bogner et al. 2021]. A systematic mapping of 34 studies (2017–2024) revealed that only 5.88% address continuous management, 17.65% analyse longitudinal data, and none defines a saturation indicator [Justino et al. 2025].

The industrial context is an investment management platform at one of the largest private financial institutions in Brazil, handling over 40 million monthly transactions and 250 thousand users. The migration from a monolith to microservices generated recurring challenges in defining and revising service boundaries, motivating the development of *Granulify*.

2. Adopted Process

Granulify is an evidence-based approach for continuous granularity management in distributed systems, developed through *Design Science Research* operationalised via *Action Research* over 13 months (Sep/2023 to Sep/2024) on the financial platform described above. Its three principles are: (i) *granularity as a spectrum*; (ii) *continuous monitoring*; and (iii) *evidence-based decisions*.

The macro-process (Figure 1) comprises three iterative phases: **(1) Granularity Spectrum Mapping**, positioning the architecture along a continuum via the GCS model; **(2) Saturation Measurement**, applying the Saturation Method to assess quantitative signals; and **(3) Architecture Redefinition**, documenting split, merge, or reorganisation decisions via Architecture Decision Records (ADRs).

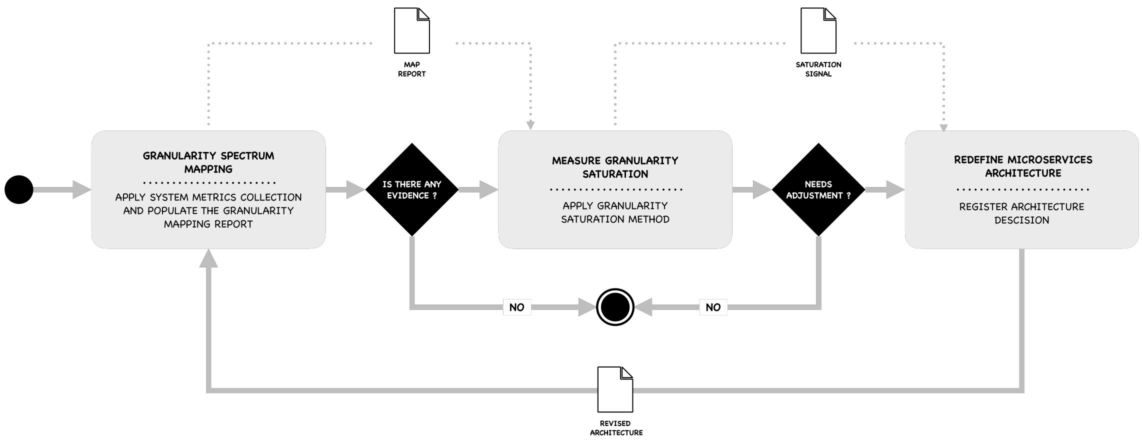

Figure 1. Granulify macro-process. *Source: Authors.*

Granularity Classification Spectrum (GCS). Organises architectural styles into seven levels: monoliths (**G0**), modular monoliths (**G1**), service-based architectures (**G2**), extended microservices (**G3**), intermediate microservices (**G4**), focused microservices (**G5**), and nanoservices/FaaS (**G6**). The classification employs eight heuristic criteria: deployment unit, internal modularity, deployment independence, data ownership, coupling, business alignment, functional scope, and state management.

Figure 2 presents the GCS model, which serves as the basis for qualitative architectural diagnosis and for communication between technical teams and managers. The model is operationalised through an 8-item questionnaire, each mapped to the heuristic criteria, that guides the classification of the architecture into one of the seven levels.

Granularity Saturation Method (GSM). The GSM processes time series of metrics to quantify how granularity changes relate to systemic variations. The orthogonal metrics model comprises 9 metrics across 4 dimensions: Development (coordination per service, effort variation), Deployment (change volume, failure rate), Code (test coverage, *code smell* density, cyclomatic complexity density), and Operational (infrastructure cost, high-risk change rate). Each metric has a polarity $p_j \in \{-1, +1\}$ indicating whether higher values are desirable or undesirable.

The method follows four steps: **(1) Per-metric variation:** for each metric j in period i, the percentage variation is computed with polarity adjustment and automatic *clipping* to contain extreme values. **(2) Dimension aggregation:** variations are aggregated by

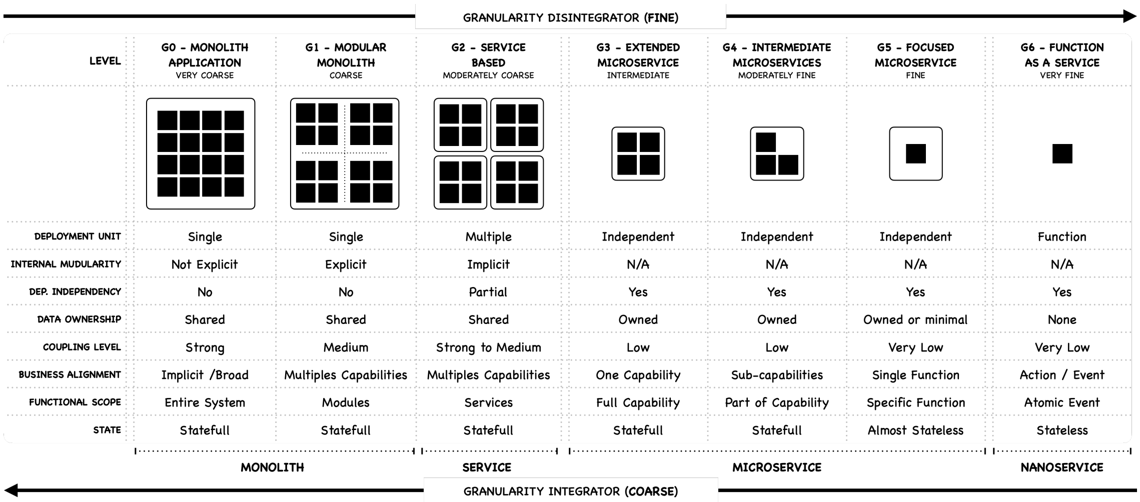

Figure 2. Granularity Classification Spectrum (GCS). ***Source: Authors.***

semantic dimension using the median, which provides robustness against *outliers* (50% breakdown point). **(3) Global indices:** two complementary indicators are computed per period: the *Net Benefit* (NB_i), the signed median of dimensional variations, which captures direction (improvement vs. degradation); and the *Total Variation* (TV_i), the median of absolute values, which captures the magnitude of structural change regardless of direction. **(4) Saturation Score and Classification:** efficiency is calculated as $\mathcal{E}_i = NB_i/TV_i$, normalised to the score $SS_i = (\mathcal{E}_i + 1)/2$, yielding a $[0, 1]$ scale: values close to 1 indicate efficient change, close to 0.5 indicate empirical saturation, and close to 0 indicate net deterioration.

The regime classification combines the granularity direction (CNS_i increasing or decreasing) with score bands of SS_i calibrated by empirical quartiles (q_{25} and q_{75}), subject to an evidence floor ($TV_i \geq \tau$). The resulting regimes are: *efficient decomposition* ($CNS\uparrow$, $SS \geq q_{75}$), *empirical saturation* ($CNS\uparrow$, $q_{25} < SS < q_{75}$), *over-decomposition* ($CNS\uparrow$, $SS \leq q_{25}$), *efficient consolidation* ($CNS\downarrow$, $SS \geq q_{75}$), *harmful consolidation* ($CNS\downarrow$, $SS \leq q_{25}$), and *low evidence* ($TV < \tau$). Figure 3 visually illustrates the saturation zones.

In the visual metaphor, the containers represent three scenarios: the **Low** zone ($SS \leq q_{25}$) corresponds to periods where structural change predominantly causes degradation; the **Neutral** zone ($q_{25} < SS < q_{75}$) represents empirical saturation, where changes yield neither consistent benefit nor harm; and the **High** zone ($SS \geq q_{75}$) indicates efficient change, where the observed variation translates into net architectural improvement. The resulting saturation signal — a time series of regimes — enables longitudinal analysis: for example, a sequence *efficient decomposition* $\rightarrow$ *empirical saturation* $\rightarrow$ *over-decomposition* indicates progressive degradation, signalling the need to halt fragmentation or initiate consolidation.

3. Solution

The application of Granulify produced results in two phases: retrospective and prospective.

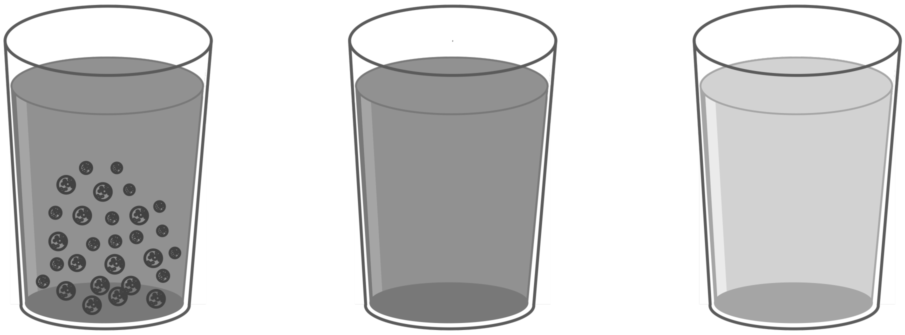

Figure 3. Saturation zones based on the SS score. The liquid volume represents the total variation (TV) and the colour intensity reflects efficiency. *Source: Authors.*

Retrospective Analysis (Sep/2023 – Feb/2024) Without systematic management, the architecture evolved in an *ad hoc* manner. The method identified *over-decomposition* in Nov/2023 ($NB = -0.21$), with higher *code smell* density (12.7/KLOC) and costs rising from \$17k to \$31k. In Dec/2023, it detected *empirical saturation* ($NB = -0.02$), with actual effort 4.5× the estimate and costs of \$35k. These signals were validated against qualitative observations from retrospectives and architectural reviews.

Prospective Monitoring (Mar/2024 – Sep/2024) With Granulify in operation, the team made proactive adjustments: in Apr/2024, three added microservices resulted in *efficient decomposition* ($NB = 0.16$; $SS = 0.99$); in Aug/2024, five services triggered *over-decomposition* ($NB = -0.05$); and in Sep/2024, strategic consolidation of two services achieved *efficient consolidation* ($NB = 0.37$; $SS = 0.99$).

Organisational Benefits (i) *Architectural visibility*: the GCS provided a shared vocabulary between technical teams and managers; (ii) *Data-driven decisions*: saturation signals replaced subjective judgements; (iii) *Cost reduction*: early detection of over-decomposition prevented service proliferation; (iv) *Predictability*: planning efficiency improved during periods with systematic management.

Contribution to Practice Granulify is the first approach that combines continuous management, longitudinal evidence collection, operational instruments, and temporal saturation signals. The approach is replicable and adaptable to different domains and organisational contexts. As a limitation, the study was conducted in a single context, and generalisation depends on future replications in other domains.

References

Araujo, R., Suzana, R., and Boscarioli, C. (2017). I GranDSI-BR – Grandes Desafios de Pesquisa em Sistemas de Informação no Brasil 2016 a 2026. Technical report, Comissão Especial de Sistemas de Informação (CE-SI), Sociedade Brasileira de Computação (SBC).

Behrad, S., Espes, D., Bertin, P., and Phan, C.-T. (2021). Impacts of service decomposition models on security attributes: A case study with 5g network repository function. In *2021 IEEE 7th International Conference on Network Softwarization (NetSoft)*. IEEE.

Bogner, J., Fritzsch, J., Wagner, S., and Zimmermann, A. (2021). Industry practices and challenges for the evolvability assurance of microservices: An interview study and systematic grey literature review. *Empirical Software Engineering*, 26(5).

Driessen, F., Ferreira Pires, L., Moreira, J. L. R., Verhoeven, P., and van den Bosch, S. (2024). A quantitative assessment method for microservices granularity to improve maintainability. In Sales, T. P., de Kinderen, S., Proper, H. A., Pufahl, L., Karastoyanova, D., and van Sinderen, M., editors, *Enterprise Design, Operations, and Computing. EDOC 2023 Workshops*, pages 211–226, Cham. Springer Nature Switzerland.

Hassan, S., Bahsoon, R., and Kazman, R. (2020). Microservice transition and its granularity problem: A systematic mapping study. *Software: Practice and Experience*, 50(9):1651–1681.

Justino, Y., Silva, C., and Duarte, R. (2025). Continuously managing microservice granularity: An evidence-based industrial approach. In *Anais do XXXIX Simpósio Brasileiro de Engenharia de Software*, pages 226–236, Porto Alegre, RS, Brasil. SBC.

Vera-Rivera, F. H., Gaona, C., Astudillo, H., and and (2021). Defining and measuring microservice granularity—a literature overview. *PeerJ Computer Science*, 7:e695.