

Análise do Overhead da Observabilidade na Computação em Névoa

Breno Costa, Aletéia P. F. Araújo

¹Departamento de Ciencia da Computação – Universidade de Brasília (UnB)
Brasília – DF – BraSil

brenogscosta@gmail.com, aleteia@unb.br

Resumo. *A observabilidade em ambientes de Computação em Névoa tem grande importância para a manutenção dos SLAs acordados com os usuários dos serviços. Este estudo tem como objetivo avaliar estratégias para reduzir o overhead causado pelo uso de ferramentas que proporcionem um aumento da observabilidade na plataforma. Utilizando o cenário prático do Mobile IoT-RoadBot, um aplicativo de cidade inteligente, avaliamos o overhead de se aumentar a observabilidade em um ambiente de teste composto por dispositivos IoT e um nó de Névoa, empregando ferramentas de código aberto como Prometheus e a pilha ELK para gerenciar métricas, logs e traces. A análise do overhead revelou que, embora o impacto nos dispositivos IoT seja mínimo, nos nós de Névoa o uso de CPU e memória é significativamente maior, ressaltando a necessidade de estratégias eficazes para reduzir o volume de dados e otimizar a configuração das ferramentas. Os resultados alcançados representaram uma diminuição de 80% no volume total de dados de observabilidade armazenados no nó de Névoa, sem que houvesse diminuição do nível de observabilidade. O acréscimo causado pelo aumento da observabilidade no volume de dados gerenciado pela Mobile IoT Roadbot foi menor que 1%.*

Abstract. *Observability in Fog Computing environments is of great importance for guaranteeing the SLAs agreed with the users. This study aims to evaluate strategies to reduce the overhead caused by the use of tools that provide increased observability in the platform. Using the practical scenario of Mobile IoT-RoadBot, a smart city application, we evaluated the overhead of increasing observability in a test environment composed of IoT devices and a Fog node, employing open-source tools such as Prometheus and the ELK stack for management, logging, and tracing. An analysis of the overhead revealed that, although the impact on IoT devices is minimal, in Fog nodes the CPU and memory usage is significantly higher, highlighting the need for effective strategies to reduce the volume of data and improve the configuration of the tools. The results achieved represented an 80% increase in the total volume of observability data stored in the Fog node, without increasing the level of observability. The increase caused by the increase in observability in the volume of data managed by the Mobile IoT Roadbot was less than 1%.*

1. Introdução

A observabilidade em sistemas distribuídos é definida como a capacidade de compreender o estado interno de um sistema a partir de seus dados externos [Kalman 1960]. Esta capacidade é crucial para a rápida análise e resolução de problemas encontrados durante a execução, auxiliando na manutenção dos níveis de serviço acordados. A observabilidade é particularmente importante na Computação em Névoa, que, ao fornecer recursos computacionais próximos aos usuários, enfrenta desafios como a restrição de recursos e a heterogeneidade dos dispositivos e conexões [Bonomi et al. 2012].

Em trabalho anterior [Costa and Araújo 2023], definimos a observabilidade na Computação em Névoa em termos da quantidade de domínios de instrumentação (ID), explicitando os principais IDs: métricas, logs e traces. Definimos ainda, um ciclo de vida dos dados de observabilidade (ODLC) na Névoa, proposto em seis fases essenciais: 1) Geração de dados, onde métricas, logs e traces são

gerados pelo sistema; 2) Armazenamento local de dados nos dispositivos IoT; 3) Transmissão de dados até o ponto de agregação e análise, um dispositivo de Névoa [Bachiega et al. 2021]; 4) Agregação de dados, onde os dados de diversos dispositivos são consolidados; 5) Consulta aos dados para tomada de decisão, com foco nos dados mais recentes para respostas mais relevantes; 6) Armazenamento de dados de longo prazo no ambiente de Nuvem.

Apresentamos um cenário motivador, o Mobile IoT-RoadBot [Banerjee et al. 2024], um aplicativo de cidade inteligente que coleta vídeos das calçadas e os envia continuamente para a Nuvem para identificar e registrar a necessidade de manutenção de abrigos de ônibus e placas de trânsito. Este aplicativo foi implantado em caminhões de serviço de coleta de lixo (WCSTs) em Brimbank, Austrália, e serve como um ambiente de teste para avaliar o desempenho da rede 5G. Os desafios enfrentados incluem a incerteza da rede, falta de espaço para armazenar vídeos, mau funcionamento da câmera e travamento do aplicativo, destacando a importância da observabilidade para a rápida identificação e resolução de problemas.

Por fim, definimos um ambiente de testes para avaliar a observabilidade na computação em Névoa, consistindo em quatro unidades Raspberry Pi 4 (dispositivos IoT), um servidor de borda TBOX-128X0 (nó de Névoa) e uma máquina virtual no Google Cloud, representando a Nuvem. Este ambiente reproduz os dados reais gerados por quatro caminhões de coleta de lixo em serviço nas ruas da Austrália em 2022. Para o gerenciamento de métricas, foi selecionado o Prometheus, e para o gerenciamento de logs, a pilha ELK, composta por Elasticsearch, Logstash e outros componentes, foi escolhida. Estas ferramentas de código aberto foram selecionadas para caracterizar os desafios de aumentar o nível de observabilidade neste ambiente. O ambiente e as ferramentas que o compõem podem ser visualizados na Figura 1.

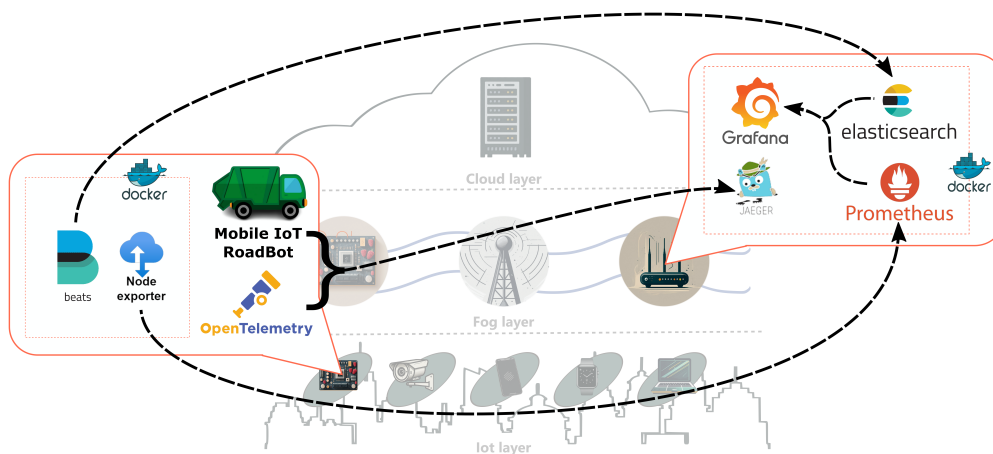


Figura 1. Sistema de observabilidade de Névoa composto por ferramentas de código aberto

2. Overhead da Observabilidade na Computação em Névoa

A observabilidade na Computação em Névoa não é um fim em si mesma. O resultado positivo que um nível aumentado de observabilidade pode oferecer deve ser equilibrado com a quantidade de recurso computacional necessário (o overhead) para alcançar esse nível. Embora isso seja óbvio de forma genérica, essa afirmação ganha mais relevância em um ambiente de Computação em Névoa devido às características de restrição de recursos e incerteza de rede. Esse resultado pode ser modelado conforme a Equação 1: $\text{Resultado} = \text{Observabilidade} / \text{Overhead}$, onde Overhead é um número não nulo entre 0 e 100, representando a porcentagem de recursos (ou seja, CPU, memória, largura de banda da rede) consumidos ao executar o ODLC. Quanto maior o overhead, menor poderá ser o resultado, considerando que a restrição de recursos é uma característica da Névoa.

Para avaliar o overhead de alcançar um nível mais alto de observabilidade em um ambiente de Computação em Névoa, usamos o testbed descrito. Este testbed foi empregado para medir: 1) O

overhead de cada ferramenta nos dispositivos IoT; 2) O overhead do componente servidor de cada ferramenta rodando em um nó de Névoa; 3) O volume de dados de observabilidade coletados, armazenados e transmitidos ao longo do ciclo de vida dos dados. Enquanto o Mobile IoT-RoadBot estava em execução em um Raspberry Pi, cada agente foi implantado no mesmo dispositivo, e o overhead foi medido em termos de uso de CPU e memória por 4 horas. Para obter a média de uso de CPU e memória a cada 5 minutos, foi utilizado o System Activity Reporter (SAR). Após registrar essas informações, a ferramenta foi desinstalada e o processo repetido com as outras ferramentas (Node Exporter, Filebeat e OpenTelemetry SDK). Um procedimento similar foi realizado em relação ao lado do servidor das ferramentas (Prometheus, Elasticsearch e Jaeger). Após medir o overhead, o Mobile IoT-RoadBot foi executado por 9 horas (um dia de trabalho regular para cada caminhão) para avaliar o volume de dados coletados por cada domínio de instrumentação e transmitidos através das camadas da arquitetura do testbed. Finalmente, para avaliar os benefícios de alcançar um nível mais alto de observabilidade em ambientes de Névoa, o output real de 4 WCSTs foi reproduzido por 9 horas no ambiente de testes. Cada saída real de WCST foi reproduzido por um Raspberry Pi que também executava agentes de observabilidade, enviando dados através do ODLC dos Dispositivos IoT para o servidor na Nuvem.

3. Avaliação dos resultados

A avaliação do overhead associado à implementação de ferramentas de observabilidade em ambientes de Computação em Névoa revelou aspectos importantes sobre o impacto dessas ferramentas no desempenho dos dispositivos IoT e nos nós de Névoa. As figuras 2a e 2b ilustram o overhead nos dispositivos IoT. Foi observado que o overhead de CPU e memória é negligenciável, com o uso agregado de CPU por todos os três componentes principais (NodeExporter, Filebeat e OpenTelemetry SDK) mantendo-se abaixo de 12% em média. Quanto ao uso de memória, foi necessário menos de 150MiB de RAM de forma agregada. Este resultado indica que o impacto das ferramentas de observabilidade nos dispositivos IoT é mínimo, permitindo que eles continuem a operar eficientemente sem comprometer seus recursos limitados.

No entanto, a situação muda quando consideramos o overhead nos nós de Névoa, apresentado nas figuras 2c e 2d. Devido ao fato de que os nós de Névoa lidam com um volume de dados quatro vezes maior, enviado pelos dispositivos IoT, o uso de CPU é consequentemente maior, embora ainda seja considerado justificável, com uma média de uso de CPU de menos de 25%. Este aumento no uso de CPU reflete a carga adicional de processamento, recebimento e armazenamento de dados imposta aos nós de Névoa. Em termos de uso de memória, as ferramentas de observabilidade apresentaram diferentes consumos nos nós de Névoa, com o Prometheus e o Jaeger consumindo quantidades moderadas de RAM, enquanto o Elasticsearch exigiu uma quantidade significativamente maior, quase 4,5GB. Este aspecto destaca a necessidade de monitorar e gerenciar cuidadosamente o overhead de memória, especialmente em aplicações que geram grandes volumes de dados ou apresentam comportamento de rajadas.

Esses resultados sublinham a importância de uma seleção cuidadosa e configuração das ferramentas de observabilidade em ambientes de Névoa, para equilibrar a necessidade de visibilidade e controle do sistema com o imperativo de manter o overhead em níveis gerenciáveis. A limitação do volume de dados armazenados a uma semana e a escolha de ferramentas com consumo de memória menores são estratégias eficazes para gerenciar o overhead sem comprometer o aumento da observabilidade. A análise detalhada do overhead associado a cada ferramenta fornece uma base sólida para tomar decisões informadas sobre a implementação de práticas de observabilidade em ambientes de Computação em Névoa, garantindo que o desempenho do sistema não seja adversamente afetado.

No contexto do Mobile IoT-RoadBot, um cenário de uso real foi implementado para avaliar a eficácia das estratégias de redução do volume de dados. A análise dos dados de observabilidade coletados durante os experimentos revelou insights importantes sobre o gerenciamento eficaz do volume

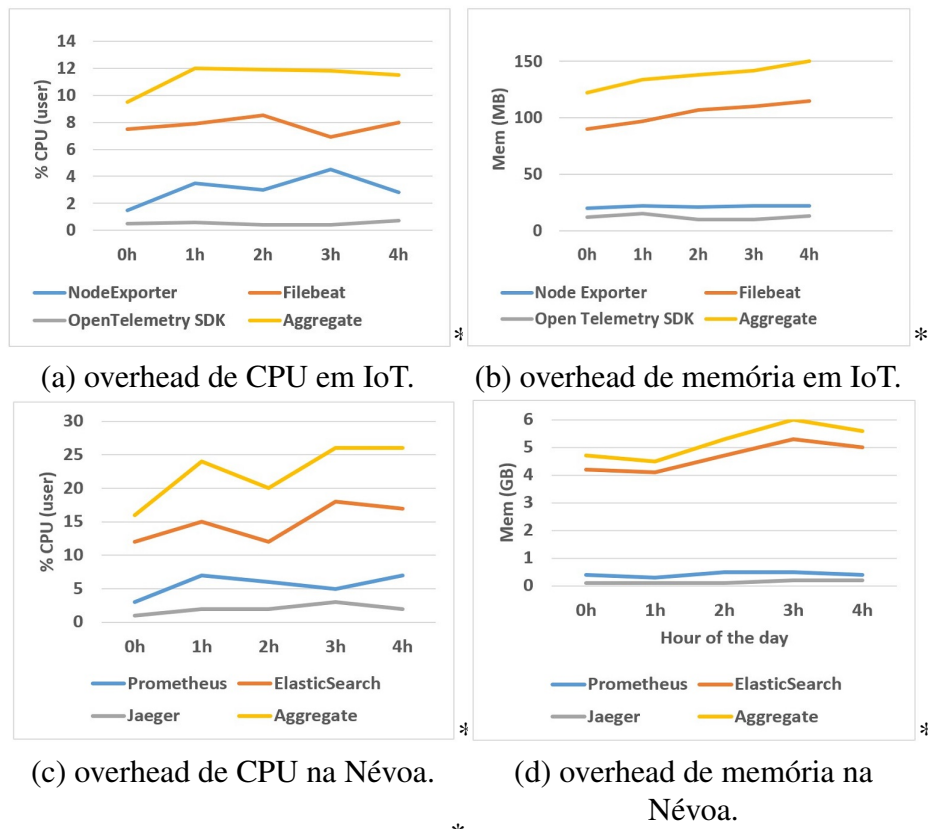


Figura 2. Avaliação do overhead de ferramentas de observabilidade em: dispositivos IoT (a) e (b); nós de Névoa (c) e (d).

de dados em ambientes de Computação em Névoa. Um resumo dos volumes coletados e transmitidos pelo OLDC, com a configuração padrão das ferramentas, pode ser encontrado na Tabela 1. Foi observado que, apesar do NodeExporter ser uma ferramenta leve em termos de consumo de CPU e memória, o volume de dados gerado não é insignificante. A configuração padrão desta ferramenta resulta na exposição de um conjunto de métricas que, embora totalize apenas 65KB, por ser coletado pelo Prometheus a cada 5s por padrão, acumula um volume substancial de dados. Considerando que há quatro dispositivos IoT expondo métricas, o volume de dados transmitidos e armazenados no nó Fog é de cerca de 8,75 GB no intervalo de uma semana, período em que esses dados estarão disponíveis para tomada de decisão e outras análises na camada da Névoa. Após atingir uma semana de idade, as informações são removidas do nó da Névoa e enviadas para a Nuvem para armazenamento de longo prazo e análise histórica. Como estimativa, o volume na Nuvem atingirá 75 GB após 2 meses de operação.

Usamos como estratégias para diminuir o overhead de dados: 1. remover o texto de ajuda no NodeExporter, que respondia por 20% do volume total e não tinha utilidade; 2. alterar a configuração do Node Exporter para expor apenas métricas de interesse (CPU, memória, disco, rede e fonte de alimentação), e 3. diminuir a frequência de coleta pela metade no Prometheus (a cada 10 segundos). O volume de dados de métricas no nó da Névoa diminuiu em 87%, o que também afetou positivamente o uso de CPU e memória pelo Prometheus. Através da configuração apropriada do Filebeat para coletar apenas os logs relevantes, foi possível manter o volume de dados de logs em um nível aceitável de 0.67 GB ao longo de uma semana. O mesmo ocorreu com os dados de traces, que têm volume muito baixo para essa aplicação.

Com o uso das estratégias descritas, foi alcançada uma redução significativa no volume de dados agregados de observabilidade, de 10GB para 2GB, representando uma diminuição de 80%. Os quatro caminhões cujos dados de observabilidade são reproduzidos pelos dispositivos IoT neste

Tabela 1: Avaliação dos dados do Mobile IoT-Roadbot para cada domínio de observabilidade.

Ferramenta	Domínio	Dados Gerados	Frequência	Volume por Hora	Armaz. Local	Agregação	Vol. Névoa (1 semana)	Armaz. Nuvem	Vol. Nuvem (2 meses)
Node Exporter	Metrics	65KB	cada 5s	46 MB	Não	Sim	8.75 GB	Sim	75 GB
Filebeat	Logs	1KB	cada 1s	3.50 MB	Sim	Sim	0.67 GB	Sim	5.77 GB
Open Telemetry	Traces	4KB	cada 15s	1 MB	Não	Sim	0.2 GB	Sim	1.54 GB

experimento transmitiram 291 GB de dados de vídeo usando a rede 5G em uma semana de operação no mundo real. Portanto, os dados de observabilidade (2 GB) representariam um overhead de menos de 1% neste caso de uso.

Essa eficiência demonstra a viabilidade de alcançar um alto nível de observabilidade em ambientes de Computação em Névoa com um overhead mínimo, desde que a coleta e gestão de dados de observabilidade sejam adequadamente administradas. No entanto, há que se considerar que casos de uso que difiram consideravelmente do que foi avaliado nesse estudo, podem não ter resultados similares aos alcançados.

4. Considerações Finais

A observabilidade na Computação em Névoa traz consigo o desafio de gerenciar o overhead adicional que as ferramentas de observabilidade podem introduzir no sistema. Este overhead é particularmente crítico em ambientes de Computação em Névoa devido à limitação de recursos e à incerteza da rede.

A definição de observabilidade proposta neste contexto leva em consideração os desafios específicos da Computação em Névoa, incluindo a necessidade de coletar, armazenar e analisar dados de observabilidade de forma eficiente em um ambiente com restrições de recursos. A utilização de ferramentas de código aberto para implementar a observabilidade em um testbed de Computação em Névoa real demonstrou ser viável, com um overhead mínimo, após a adoção de estratégias para reduzir o volume de dados e otimizar a configuração das ferramentas. Esses resultados sublinham a viabilidade de alcançar um alto nível de observabilidade em ambientes de Névoa, o que é essencial para a gestão eficaz de aplicações IoT complexas e distribuídas.

Em resumo, a observabilidade na Computação em Névoa é uma área de pesquisa emergente que aborda a necessidade crítica de visibilidade e controle em ambientes distribuídos e com restrições de recursos. Através da implementação e avaliação de ferramentas de código aberto em um ambiente de testes real, este trabalho contribui para o entendimento de como a observabilidade pode ser alcançada efetivamente na Computação em Névoa, destacando tanto os desafios quanto as estratégias para minimizar o overhead associado.

References

- Bachiega, J., da Costa, B. G. S., and Araujo, A. P. F. (2021). Computational perspective of the fog node. *22nd International Conference on Internet Computing & IoT*.
- Banerjee, A., Costa, B., Forkan, A. R. M., Kang, Y.-B., Marti, F., McCarthy, C., Ghaderi, H., Georgakopoulos, D., and Jayaraman, P. P. (2024). 5g enabled smart cities: A real-world evaluation and analysis of 5g using a pilot smart city application. *Internet of Things*, 28:101326.
- Bonomi, F., Milito, R., Zhu, J., and Addepalli, S. (2012). Fog computing and its role in the internet of things. *MCC '12*, pages 13–16, New York, NY, USA. ACM.
- Costa, B. and Araújo, A. P. (2023). Observabilidade no ambiente de computação em névoa. In *Anais da VI Escola Regional de Alto Desempenho do Centro-Oeste*, pages 10–14. SBC.
- Kalman, R. E. (1960). On the general theory of control systems. In *Proceedings First International Conference on Automatic Control, Moscow, USSR*, pages 481–492.