

# Abordagem em Borda de Rede para Coleta e Processamento Contínuo de Telemetria em Larga Escala na Indústria 4.0

Alann Rodrigues Ferreira, Alisson Vasconcelos Brito, Fernando Menezes Matos

<sup>1</sup>Centro de Informática – Universidade Federal da Paraíba (UFPB)  
Cep 58058-600 – Campus Mangabeira – João Pessoa – PB – Brazil

alanncz45@gmail.com, fernando@ci.ufpb.br, alisson@ci.ufpb.br

**Abstract.** *In industrial environments, a small amount of data can be managed and analyzed in a simple way, but when it gains new proportions, growing steadily, the task of managing this data becomes more complex. This paper presents a Network Edge approach to establish a continuous stream of collection and processing of telemetric data from industrial machines. Called ThreeVS-industry, this approach uses OPC-UA as an interoperability standard in the industrial automation space, as, collected data is sent to a Big Data environment. The goal of this approach is to integrate machines and automated systems to help managers gain information and insights, which will serve as the basis for continuous process improvements and strategic planning. In order to integrate a device failure detection mechanism into the approach, a simulated experiment was implemented in Cloud Computing and Edge Computing environments. In all runs, using Edge Computing proved superior, with gains of up to 76.18% in fault detection time.*

**Resumo.** *Em ambientes industriais, uma pequena quantidade de dados pode ser gerenciada e analisada de forma simples, mas quando ganham novas proporções, crescendo de forma constante, torna-se mais complexo a tarefa de gerir esses dados. Este trabalho apresenta uma abordagem em Borda de Rede para estabelecer um fluxo contínuo de coleta e processamento de dados telemétricos de máquinas industriais. Chamada de ThreeVS-industry, essa abordagem utiliza OPC-UA como padrão de interoperabilidade no espaço de automação industrial, à medida que, dados coletados são enviados para um ambiente de Big Data. O objetivo dessa abordagem é integrar máquinas e sistemas automatizados para ajudar gestores a obter informações e insights, que servirão de base para aprimoramentos contínuos em processos e planejamentos estratégicos. Com o intuito de integrar um mecanismo de detecção de falhas nos dispositivos à abordagem, um experimento simulado foi implementado em ambientes de Computação em Nuvem e Computação de Borda. Em todas as execuções, utilizar Computação em Borda mostrou-se superior, com ganhos de até 76.18% em relação ao tempo de detecção de falha.*

## 1. Introdução

Várias empresas de manufatura aplicam em suas linhas de produção dispositivos industriais com o intuito de automatizar tarefas, aumentar a segurança do trabalhador e aumentar a produção em geral, reduzindo com isto os custos operacionais [Jardine et al. 2006]. Com

estes dispositivos tornando-se mais predominantes nos ambientes de produção, a indústria manufatureira investe forte em instalações físicas por meio das tecnologias de informação e comunicação para obter informações sobre o processo produtivo de forma detalhada e precisa. Assim, os gestores podem tomar decisões mais assertivas para aprimorar seus produtos e consequentemente, melhor satisfazer as expectativas dos consumidores.

Nas palavras de Klaus Schwab [Tessarini 2018] “estamos no início de uma revolução que está mudando fundamentalmente como vivemos, trabalhamos e nos relacionamos um com o outro”. Trata-se da Indústria 4.0: um novo modelo de produção, fortemente relacionado com a Internet das Coisas (IoT), onde máquinas, ferramentas e processos estarão conectados à Internet através de sistemas ciberfísicos, interagindo entre si e com a capacidade de operar, tomar decisões e se auto-corrigir. IoT é uma maneira de transformar dados brutos, obtidos do ambiente de sistemas físicos, em dados úteis que podem ser usados para fins de monitoramento por meio da Internet, utilizando por exemplo, a Computação em Nuvem, que emergiu como uma tecnologia essencial no contexto da Indústria 4.0 [Gori and Gori 2022].

Outra tecnologia, conhecida como Computação de Borda, vem atraindo interesse pelo seu potencial de satisfazer requisitos que não são atendidos por um modelo centralizado em Nuvem [Coutinho et al. 2016]. Este paradigma estende os recursos computacionais disponíveis na Nuvem para a Borda de rede para apoiar as soluções IoT. Apesar das potenciais vantagens do modelo baseado em Nuvem para Internet das coisas, alguns cenários não são favoráveis à sua aplicação. Os sistemas de Computação em Nuvem são altamente centralizados, onde a maioria da computação ocorre em poucos e grandes centros de dados [Coutinho et al. 2016]. Em ambientes IoT onde os dispositivos estão próximos uns dos outros, e implementam soluções com requisitos de baixa latência, torna-se ineficiente transmitir dados através da Internet para núcleos de processamento distantes. Por exemplo, uma planta industrial em funcionamento com vários robôs operando, no qual, é de suma importância monitorar dados telemétricos dessas máquinas para prever e evitar falhas. Nesse cenário, eventuais sobrecargas e atrasos na rede tornarão essa solução impraticável devido à enorme quantidade de dados produzidos conforme novas máquinas são adicionadas ao ambiente. Essa tendência para concentração dos dados nas bordas da rede é uma realidade contemporânea e estimulada pelo atual crescimento da IoT [Coutinho et al. 2016]. Portanto, minimizar a quantidade de dados enviados para Nuvem é fundamental na redução dos custos envolvidos e no tempo de resposta das aplicações.

As fábricas tradicionais precisam de recursos de monitoramento para controle de produção automatizada para produzir de forma eficiente e lucrativa, contudo, a integração entre o sistema real e as representações do sistema virtual na produção tradicional é precária [Gori and Gori 2022]. A execução dos serviços de manutenção em máquinas e equipamentos industriais, sempre foram um dos grandes gargalos nos processos produtivos [Santos et al. 2018]. No que concerne a engenharia, o monitoramento de equipamentos industriais pode ser definido como uma coleta sistemática de dados e uma análise de informações coletadas de como o objeto monitorado opera no contexto de operação fabril [Fabricio 2018]. Significa que as informações são coletadas durante toda a operação e que a análise dessas informações contribui para que os usuários interessados acompanhem os resultados alcançados e verifiquem a necessidade de ajustes nas ações para que os objetivos da operação sejam obtidos. Visto isso, o monitoramento de equipamentos

industriais tem por objetivo o acompanhamento do processo de produção, a avaliação da “saúde” do equipamento frente às suas especificações técnicas, condições, e a detecção precoce de processos de deterioração física que causem avarias [Han and Song 2003].

Uma pequena quantidade de dados pode ser gerenciada e analisada de forma simples, mas quando dados ganham novas proporções, crescendo de forma constante, torna-se mais complexo a tarefa de gerir esses dados. Contudo, [Ribeiro 2014] observa que a gestão desses recursos possibilita a resolução de problemas que nem sabemos que existem, e que não podemos prescindir de ferramentas, pois a capacidade do ser humano de analisar dados e informações em larga escala é limitada. Neste cenário, pode-se apontar uma dificuldade a ser enfrentada pelas empresas de manufatura: a geração em alta velocidade de uma quantidade variada e extensa de dados, denominada Big Data [Ayub et al. 2021], gerando desafios relacionados às formas de extração e assimilação de conhecimento, além da instigação da capacidade humana para gerir estes dados disponíveis, e assim gerar valor sobre essas informações. Logo, são necessários instrumentos que auxiliem a executar tarefas de extração e análise de dados de forma eficiente e prática.

A mudança do modelo de fábrica tradicional para o modelo industrial 4.0 é uma questão urgente, pois essa transformação estimula um efeito profundo e duradouro na produção futura em todo o mundo. A principal forma de implementar este modelo é estabelecer uma Fábrica Inteligente (Smart Factory), pois ela representa uma manufatura onde máquinas e produtos interagem entre si sem controle humano, conduzindo as empresas a novos padrões de desempenho [Gori and Gori 2022]. Este artigo propõe uma abordagem IoT apoiada por Computação de Borda para coleta e processamento contínuo de dados de chão de fábrica aderente a Indústria 4.0, chamada ThreeVS-industry. A abordagem utiliza OPC-UA (Open Platform Communications Unified Architecture) como padrão de interoperabilidade para troca de dados no espaço de automação industrial, e um ambiente Big Data para extrair dados gerados durante o processo e transformá-los em informações úteis de forma eficiente e prática. Portanto, os objetivos da abordagem incluem a transição dos processos maquinários nas fábricas tradicionais para um modelo de Smart Factory e prover facilidade aos gestores e desenvolvedores a criar e integrar sistemas em seus processos de manufatura. Em resumo, as principais contribuições deste trabalho podem ser destacadas da seguinte forma: i) introdução de uma abordagem inteligente que integra máquinas industriais a sistemas automatizados, ii) permitir acesso a informações do nível de produção no chão de fábrica para uma melhor tomada de decisões, iii) facilitar integração de sistemas que sejam vantajosos para o processo de manufatura.

Com o intuito de avaliação, um experimento simulado de mecanismo de predição de falhas em robôs industriais foi integrado à ThreeVS-industry. Com base nos resultados obtidos, a abordagem possibilitou a extração e processamento contínuo de dados telemétricos em máquinas industriais, e ao tentar prever uma falha, apresentou taxa média de sucesso de 68,60% e taxas de desempenho até 76,18% superiores em comparação a métodos tradicionais em Nuvem.

As próximas seções deste artigo estão organizadas como se segue. A segunda seção apresenta os trabalhos relacionados a esta pesquisa. A terceira seção descreve a abordagem ThreeVS-industry. A quarta seção apresenta a metodologia de avaliação dos resultados, por fim, a quinta seção apresenta as conclusões e trabalhos futuros.

## 2. Trabalhos Relacionados

O OPC Classic tem sido utilizado na indústria permitindo a troca de dados entre dispositivos, entretanto, os sistemas baseados no OPC Classic, têm escalabilidade limitada devido à dependência de plataforma [Schwarz and Börcsök 2013]. O OPC-UA surge para superar essas limitações e está sendo usada em muitas aplicações industriais com capacidade abrangente de modelagem de informações, permitindo comunicação padronizada máquina a máquina [Givehchi et al. 2017].

Há uma significativa produção bibliográfica relacionada à concepção de aplicações relacionadas a conectividade de dispositivos industriais. Dentro desse universo estão os trabalhos voltados para a proposição de arquiteturas de monitoramento e controle distribuído do chão de fábrica baseado no acesso OPC-UA. Para uma leitura mais aprofundada no assunto, [Schleipen et al. 2016] descreve diferentes cenários e casos de usos implantados, para afirmar o papel do OPC-UA como tecnologia capacitadora para produção flexível, adaptável e transparente dentro do contexto da Indústria 4.0.

[Fumagalli et al. 2021] propõe a implantação do protocolo OPC-UA em uma célula didática de Manufatura Integrada por Computador (CIM), para possibilitar a aquisição de dados e controle de um braço robótico modelo Scorbobot ER9 em conjunto com um centro de usinagem CNC (Controle Numérico Computadorizado) EMCO PC MILL 155.

[Silva and Michel 2021] integra rede de sensores sem fio de longo alcance, utilizando a tecnologia LoRa, o protocolo LoRaWAN e uma rede OPC-TCP com padrão OPC-UA. Esta proposta permite a criação de espaços ciberfísicos onde todos os elementos podem se comunicar de forma segura, adaptável e escalável, oferecendo uma solução de conectividade IIoT —(Industrial Internet of Things), útil para a Indústria 4.0.

[de Oliveira and Álvares 2018] propõe o planejamento e a implementação de um framework para monitoramento e teleoperação de máquinas-ferramenta CNC, tendo como elemento de teste e validação um centro de torneamento CNC da marca Romi modelo Galaxy 15M, provido com controlador Fanuc 18i-Ta. A implementação tem como resultado um sistema com arquitetura cliente/servidor baseada na Web (CyberDNC), combinado com elementos aderentes a Indústria 4.0, como um serviço MTConnect para acesso a dados de CNC, serviço OPC via Web para interação com a máquina através de funções de PMC/PLC (Programmable machine control/Programmable logic controller).

[Kim and Sung 2018] propõe o projeto de um wrapper OPC-UA autônomo e discute seu desempenho por meio de extensos experimentos usando uma implementação de protótipo. O wrapper consiste em dois componentes principais, ou seja, servidor UA e cliente Classic, que se comunicam entre si via memória compartilhada e semáforo. Uma característica importante do projeto é que ele é independente de plataforma. Ele emprega uma biblioteca de tempo de execução DCOM implementada em Java. Assim, torna possível construir um sistema de wrapper de custo competitivo usando soluções não-Windows COTS (Commercial Off-The-Shelf). Com a modificação de algumas chamadas de função para comunicação entre processos, o wrapper pode ser portado para outras plataformas sem muita dificuldade.

Com base na análise dos trabalhos relacionados, é possível observar que, todos tem foco em monitoramento industrial sem considerar a persistência do estado dos dados

monitorados, e os que consideram, não apresentam uma abordagem prática e padronizada para armazenamento e pós-processamento para transformá-los em informações úteis.

**Tabela 1. Trabalhos Relacionados.**

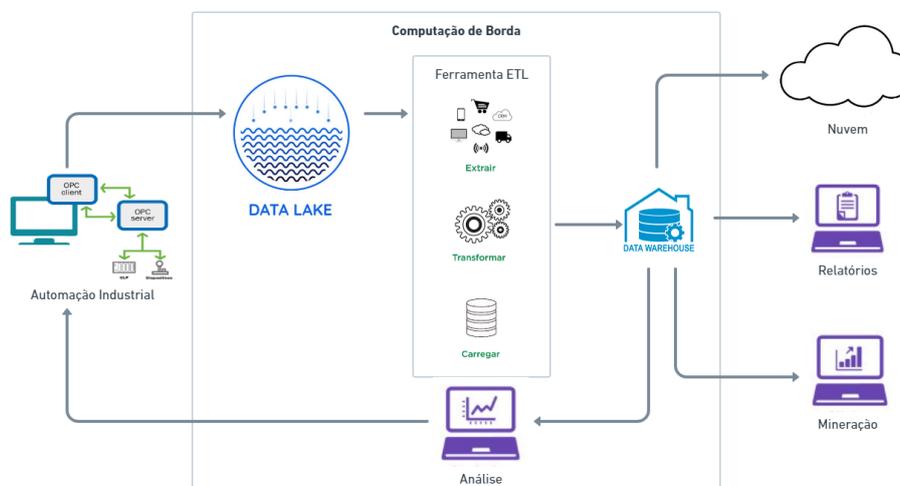
REFERÊNCIA	INTEROPERABILIDADE	PERSISTÊNCIA DE DADOS	AMBIENTE BIG DATA	OBJETIVO
Fumagalli et al. 2021	OPC-UA	SIM	NÃO	Aquisição de dados e controle de braço robótico
Silva and Michel 2021	OPC-UA	NÃO	NÃO	Integrar rede de sensores sem fio de longo alcance
de Oliveira and Alvares 2018	MTConnect	NÃO	NÃO	Framework para monitoramento e teleoperação de máquinas-ferramenta CNC.
Kim and Sung 2018	OPC-UA	NÃO	NÃO	Propõe o projeto de um wrapper OPC-UA autônomo e discute seu desempenho
ThreeVS-industry	OPC-UA	SIM	SIM	Integrar máquinas e sistemas automatizados, realizando persistência de dados para obter informações e insights

### 3. ThreeVS-industry

A abordagem ThreeVS-industry considera aos trabalhos relacionados analisados anteriormente, a persistência do fluxo de dados que representa o estado da máquina industrial em determinado momento. Portanto, é composta por cinco componentes; Cliente OPC-UA, Servidor OPC-UA, Data Lake, Processo de ETL e Data Warehouse. [Rodrigues and dos Santos Mello 2022] define Data Lake como repositórios de grandes volumes de dados variados mantidos em seus formatos brutos (estruturados ou não) para ser manipulados e transformados livremente conforme as necessidades de recuperação e análises da organização. [Ferreira et al. 2010] definem ETL como um processo para extrair dados de um sistema de Bases de Dados (BD), sendo esses dados processados, modificados, e posteriormente inseridos numa outra BD. [Monteiro et al. 2004] definem Data Warehouse como um repositório de dados provenientes dos dados operacionais (dos Sistemas OLTP), onde se cria um ambiente homogêneo e padronizado, com finalidade de proporcionar análises de negócio concentradas em um só local.

ThreeVS-industry adota o paradigma SOA (Service-Oriented Architecture), definindo processos em conjunto de unidades bem definidos, cada qual com sua funcionalidade descrita [Erl 1900]. funciona da seguinte forma: para melhorar o desempenho e minimizar a quantidade de dados que venham a ser enviados para a Nuvem ou qualquer outro sistema externo, os serviços são definidos com Computação de Borda. Estabelece a comunicação dos dispositivos Industriais aos serviços por meio da tecnologia OPC-UA como padrão de interoperabilidade para troca de dados contínua, segura e confiável no espaço de automação industrial. Os dados coletados são enviados e armazenados em um sistema Data Lake, independente de tipo, forma e estrutura. A medida que são persistidos, tem-se uma ferramenta de ETL extraindo os dados brutos e transformando-os em

estruturas e informações padronizadas que podem ser melhor interpretadas. Ainda no processo de ETL, as estruturas resultantes são armazenadas em um Data Warehouse que fica disponível para aplicações internas e externas que necessitem dessa informação, por exemplo, o serviço de Análise presente no bloco de Computação de Borda, que executa análises dos dados contidos no Data Warehouse e comunica os resultados para o bloco de automação industrial. Na Figura 1, representando o fluxo de processos, podemos observar o bloco de Computação de Borda agregando os principais serviços da abordagem, localizados bem próximos ao ambiente de automação industrial com OPC-UA. Esse posicionamento geográfico dos serviços desempenha um papel significativo na ingestão e processamento de dados [Coutinho et al. 2016].



**Figura 1. Arquitetura da Abordagem ThreeVS-industry.**

Com a definição dessa estrutura, conseguimos lidar com o imenso volume de dados gerados diariamente pelas máquinas. No bloco de Computação de Borda, temos um serviço de análise dos dados já transformados e inseridos em um Data Warehouse, executando o mais próximo possível de onde são produzidos, sendo a forma mais adequada para tratar desafios relacionados ao volume, variedade e velocidade de dados gerados. Isso possibilita, por exemplo, acelerar o tempo de resposta de aplicações com requisitos de processamento em tempo real, como uma análise preditiva nos dados para detectar uma possível falha no dispositivo. Essa detecção preventiva, definida como “Alerta”, seria uma informação de alto valor para os operadores das máquinas, possibilitando evitar a falha com operações de manutenção. Fora do bloco de Computação de Borda, temos uma vasta possibilidade de aplicações úteis para os gestores, dependendo do que eles querem e precisam. Alguns exemplos seriam aplicações de geração de relatórios, mineração de dados e serviços diversos na Nuvem.

#### 4. Metodologia

Foram realizados experimentos por simulações de Robôs UR (Universal Robots) enviando dados de telemetria para ambientes de Big Data por meio de dois micros serviços, cliente e servidor OPC-UA. Os algoritmos foram implementados na linguagem Python 3.10 e conectados pelo protocolo HTTP (Hypertext Transfer Protocol) a dois ambientes distintos; um ambiente com Computação de Borda, e outro com computação em Nuvem.

Para o experimento em Nuvem, a empresa AI Robots <sup>1</sup>, que hoje contribui com os processos industriais desenvolvendo soluções de robotização, aceitou em colaborar com a pesquisa e forneceu ambiente e registros de metadados temporais necessários para avaliarmos a abordagem proposta. Para o experimento de Computação em Borda, foi utilizado duas máquinas locais, um computador de mesa e um notebook.

Para o processo de coleta foram consideradas cinco rotinas de execução com quantidades de robôs (1, 2, 50, 100 e 150 robôs) diferentes enviando dados simultaneamente para o ambiente em Nuvem e para o ambiente com computação de Borda. Em cada rotina, em um dos robôs foi inserido um comportamento intencional que configura uma possível falha devido à alta temperatura do robô. O intuito desse experimento é avaliar a capacidade em ambas as abordagens de ambientes, de prever uma possível falha antes que ela aconteça. Além disso, em uma conversa com alguns funcionários da AI Robots, os quais trabalham na área de chão de fábrica, apontaram que o tempo padrão de coleta de telemetria usados para um robô UR (Universal Robots) seria a cada 10 segundos para não sobrecarregar o sistema do robô. Para teste de validação da abordagem foi configurado a coleta para ser feita a cada 5 segundos.

As métricas utilizadas para avaliação foram: i) Taxa de sucesso; ii) tempo de processamento; e iii) Desempenho. **Taxa de sucesso:** Representa o aproveitamento da abordagem em detectar uma falha antes que ela ocorra no intervalo de tempo de início de coleta de dados até o momento exato da falha. Por exemplo: considere uma execução de coleta em um dispositivo com:  $T_i = 0$ ,  $T_p = 5$  e  $T_f = 10$ , onde,  $T_i$  é o tempo inicial da coleta,  $T_p$  é o momento em que foi previsto a falha e  $T_f$  o tempo exato da falha do dispositivo, neste caso temos uma taxa de sucesso de 50%. Se  $T_p = 10$ , temos 0% de sucesso em detectar a falha, pois ela foi detectada no seu exato momento, impedindo de tomar decisões para evitar o problema. O caso mais extremo é quando temos taxas negativas, por exemplo; caso  $T_p = 15$  temos uma taxa de sucesso de -50%, pois a falha foi detectada, mas ocorreu momentos depois ao seu evento. **Tempo de processamento:** Representa o tempo gasto para extrair dados, processar dados, prever uma falha e emitir alerta para o setor de automação. Esta medida influencia diretamente a taxa de sucesso, quanto menor o tempo necessário para processamento, maiores são as taxas de sucesso. **Desempenho:** Representa o comportamento final em tempo de predição de falha da ThreeVS-industry comparado aos métodos tradicionais em Nuvem. Quanto menor for o tempo de processamento e consequentemente taxas de sucesso melhores, teremos um quadro de desempenho melhor.

Antes de executar o processo de coleta, foi feito um experimento de carga para o processo de ETL realizado pelo Databricks. O intuito desse teste era validar a escalabilidade e tempo de processamento com diferentes cargas presente no Data Lake. Foram feitas quatro rotinas de execução. Na primeira rotina considerou-se 20 robôs com 10 arquivos de telemetria presentes no Data Lake, totalizando 200 arquivos para serem processados. Na segunda rotina, considerou-se 100 robôs com 10 arquivos de telemetria presentes no Data Lake, totalizando 1.000 arquivos para serem processados, e por último, na terceira e quarta rotinas considerou-se 100 robôs com 100 arquivos de telemetria presentes no Data Lake, totalizando 10.000 arquivos para serem processados, onde a terceira rotina usou menos recursos de hardware do que a quarta. A abordagem foi comparada ao

---

<sup>1</sup><https://www.airobots.com.br>

ambiente computacional tradicional em Nuvem. Devido ambiente em Nuvem ter seus recursos controlados pelo Azure, para tornar justa a comparação, foi estabelecido o máximo possível de recursos de hardware semelhantes entre os dois ambientes. As Tabelas 2 e 3 a seguir, mostram recursos e propriedades de cada ambiente.

**Tabela 2. Ambiente local com computação de borda - Serviços e recursos.**

Computação em borda		
Dispositivo	Recursos	Função
Máquina A	I5-1135G7 4.20GHz, 8 Smart Cache 4 Núcleos e 8 threads, 12GB RAN	Cliente e Servidor OPC-UA, Serviço de análise de dados para predição de falhas.
Máquina B	I3 12100F 4.30 GHz, 12 Smart Cache 4 Núcleos e 8 threads, 24GB RAN	Serviços Big Data compostos por Data Lake, Databricks (ETL), Data Warehouse e Mongo.

**Tabela 3. Ambiente em nuvem Azure - Serviços e recursos.**

LOCAL		
Dispositivo	Recursos	Função
Máquina A	I5-1135G7 4.20GHz, 8 Smart Cache 4 Núcleos e 8 threads, 12GB RAN	Servidor OPC-UA
Máquina B	I3 12100F 4.30 GHz, 12 Smart Cache 4 Núcleos e 8 threads, 24GB RAN	Cliente OPC-UA
Nuvem Azure		
Serviço	Recursos	Função
storageaccount	14GB RAN, 4 núcleos, 8 Threads (Dinâmica escalonável dependendo da demanda)	Data Lake onde todos os arquivos de telemetria serão armazenados.
databricks	14GB RAN, 4 núcleos, 8 Threads (Dinâmica escalonável dependendo da demanda)	Instancia da ferramenta Databricks como processo de ETL para tratar os dados armazenados no Data Lake.
dw	14GB RAN, 4 núcleos, 8 Threads (Dinâmica escalonável dependendo da demanda)	Data Warehouse com instancia da ferramenta Azure Synapse Analytics como processo de análise de dados.
mongo	Taxa de transferência total: 8000 RU/s (Unidades de Solicitação por segundo)	Instancia de banco de dados Mongo que foi adotado para ao experimento.

## 5. Análise dos Resultados

A Tabela 4 apresenta quatro estados diferentes com dados de telemetria coletados pela ThreeVS-industry. Pode-se notar que, à medida que o número de robôs aumenta, o Data Lake será povoado mais rápido e terá mais arquivos de telemetria a serem processados, aumentando a carga de trabalho do processo. Os dois primeiros cenários, com 200 e 1.000 arquivos respectivamente, operam de forma estatisticamente semelhante. Já na configuração com 10.000 arquivos nota-se uma melhoria de desempenho de aproximadamente 50% no tempo de processamento, à medida que o número de CPUs aumenta. Conclui-se que o desempenho referente ao tempo de processamento depende diretamente da quantidade de arquivos a serem processados, que depende da quantidade de robôs presentes no sistema. Desta forma, recursos de software e de hardware precisam ser adicionados conforme mais robôs são integrados ao fluxo de coleta para garantir uma boa eficiência de escalabilidade.

**Tabela 4. Processos de ETL - Teste de carga de processamento de dados.**

Robôs	Arquivos processados	Tempo de processamento	Recursos
20	200	00h:04m:27s	14GB RAM, 4 núcleos, 8 Threads
100	1000	00h:20m:56s	14GB RAM, 4 núcleos, 8 Threads
100	10.000	01h:08m:14s	14GB RAM, 4 núcleos, 8 Threads
100	10.000	00h:36m:17s	14GB RAM, 8 núcleos, 16 Threads

A Figura 2 mostra o comportamento da temperatura em função do tempo em uma coleta de telemetria do robô. No Gráfico, podemos verificar que existe um superaquecimento do dispositivo, chegando a valores de 85 °C, 90 °C e até 100 °C. Este trabalho considerou a norma ISO 13732-1:2006 EN ISO 13732-1:2008 [2006/42/CE] <sup>2</sup> presente no manual no robô UR3 para simular picos de temperatura durante a execução do robô e coletar esses valores. Os recursos computacionais em todos os testes foram os mesmos apresentados nas Tabelas 2 e 3 deste trabalho.

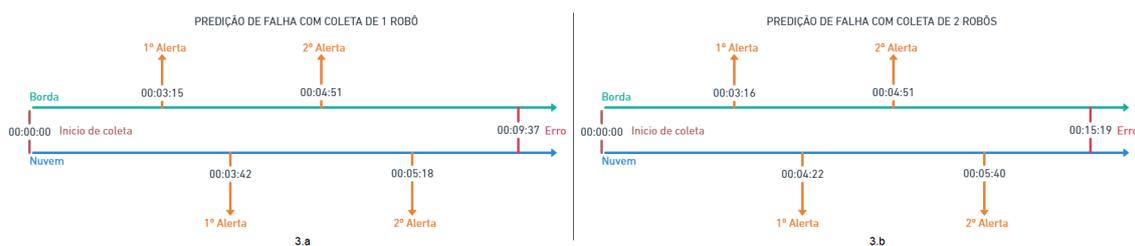


**Figura 2. Variação de temperatura do Robô UR (Universal Robots).**

No Gráfico, as áreas circunscritas em laranja representam indicadores para uma possível falha no dispositivo devido a picos altos em sua temperatura, onde, os dois primeiros círculos da esquerda para a direita representam um padrão indicativo de falha, e o terceiro a confirmação deste padrão. O círculo vermelho representa a falha ao estourar o limite máximo de temperatura e consequentemente a parada de funcionamento do dispositivo. Essa simulação foi adotada em todos os testes seguintes neste trabalho, onde, esperou-se que, através da abordagem ThreeVS-industry, possam ser emitidos alertas de predição de falhas no seguinte fluxo; i) após perceber o padrão de dois picos, um primeiro alerta é emitido indicando possível falha; ii) após obter a confirmação do primeiro indicador com o terceiro pico de temperatura, um segundo alerta é emitido.

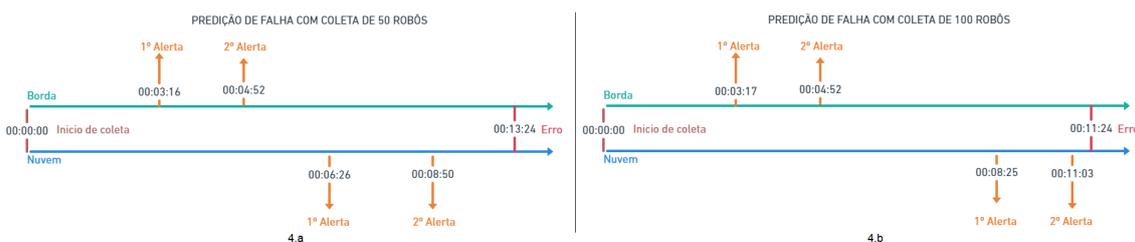
<sup>2</sup><https://www.universal-robots.com/>

A Figura 3 demonstra o momento em que a abordagem gera alertas para execuções de coletas de telemetria em comparação com métodos tradicionais em Nuvem. No gráfico 3.a temos a execução de coleta de apenas um robô, e no gráfico 3.b temos a execução simultânea de dois robôs, onde, apenas um simula indicadores de falha. Podemos observar na execução apresentada, que, em ambas as execuções e em ambos os ambientes, a abordagem apresentou comportamento semelhante com tempos de resposta (alerta preditivo) bem próximos, o que mostra que, com poucos dispositivos integrados ao fluxo, a abordagem ThreeVS-industry se equipara aos métodos tradicionais.



**Figura 3. Alertas preditivos com coleta de 1 e 2 robôs em execução.**

Seguindo o mesmo fluxo de simulação apresentado anteriormente, a Figura 4 demonstra a execução com 50 e 100 robôs nos Gráficos 4.a e 4.b respectivamente. Podemos observar que, com Computação de Borda nas duas execuções, a variação de desempenho não apresentou mudanças significativas. Já no ambiente tradicional em Nuvem, o tempo de resposta ficou mais próximo do momento exato da falha, o que mostra que, com o hardware atual, abordagens tradicionais começam a apresentar uma queda de desempenho computacional em paralelo com a inclusão de novos dispositivos ao fluxo, apresentando uma detecção de falha um pouco mais próxima do momento exato da mesma.

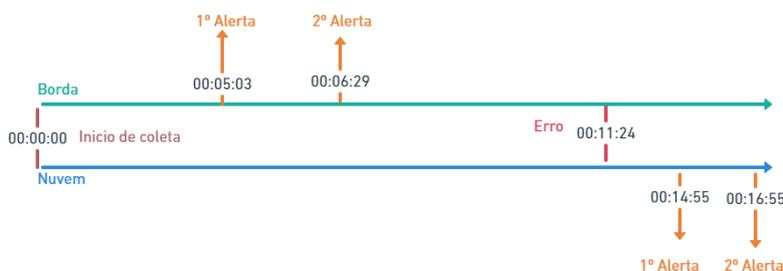


**Figura 4. Alertas preditivos com coleta de 50 e 100 robôs em execução.**

A última execução foi realizada com a coleta simultânea de 150 robôs, seguindo os mesmos parâmetros das execuções anteriores, a Figura 5 demonstra essa execução. Podemos observar que, a Computação de Borda continuou a apresentar um tempo de resposta satisfatório com pequena variação devido à carga de trabalho, já no ambiente em Nuvem apresentou um baixo desempenho, mostrando que, adotando um modelo tradicional nas configurações atuais de hardware não atinge o objetivo de prever falhas, emitindo os alertas após a falha acontecer.

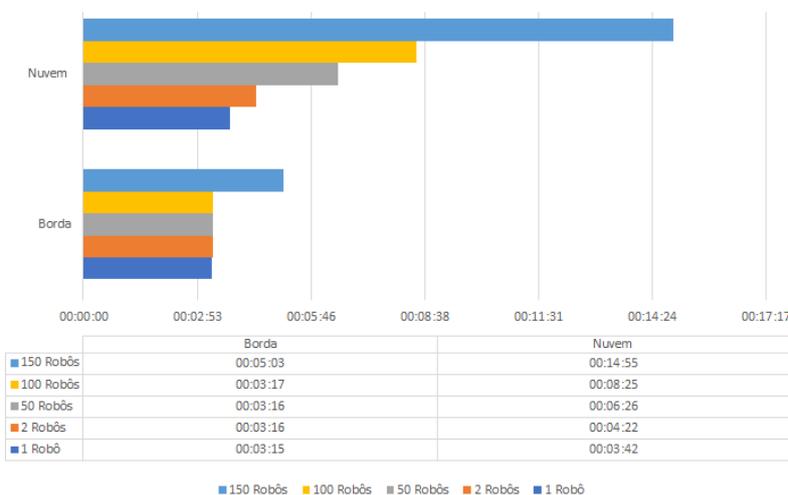
Nos dois cenários, a abordagem consegue prever falhas, mas tem seu desempenho reduzido à medida que novos robôs e, conseqüentemente mais carga de trabalho, são integrados ao fluxo. Nas configurações de hardware utilizadas nos testes, implementar uma

abordagem com Computação de Borda tem uma maior taxa de sucesso e um melhor desempenho, tendo um tempo de resposta semelhante em todas as execuções e apresentando uma variação pouco significativa. Por outro lado, a solução em Nuvem apresentou uma baixa na taxa de sucesso e consequentemente uma degradação no desempenho à medida que novos dispositivos são integrados ao fluxo.



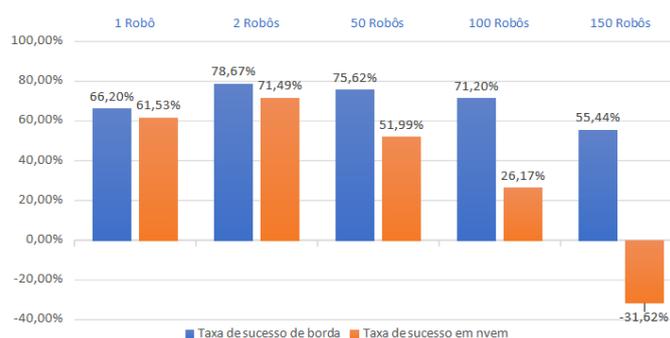
**Figura 5. Alertas preditivos com coleta de 150 robôs em execução.**

Na Figura 6 temos a representação do tempo gasto em processamento para alertar em cada ambiente de computação. Podemos observar que na Computação em Borda, o tempo necessário apresentou variação significativa, porém sem comprometer significativamente seu desempenho, quando as execuções simultâneas continuam 150 robôs. A Computação em Nuvem apresentou alta variação no tempo de processamento e alerta quando as execuções simultâneas já continham 50 robôs, chegando a comprometer seu desempenho com 100 robôs executando simultaneamente.



**Figura 6. Tempo de Processamento e Alerta.**

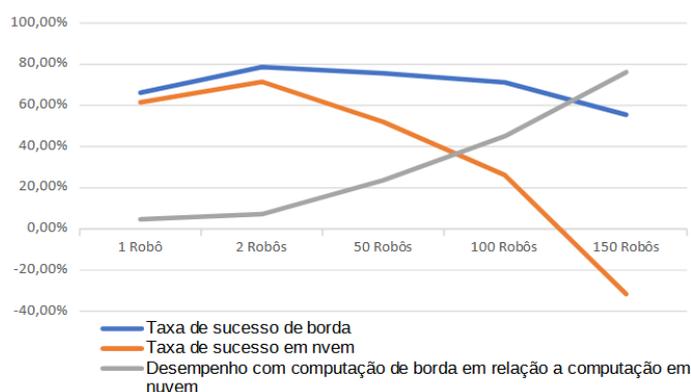
A Figura 7 ilustra a taxa de sucesso da abordagem nos dois cenários computacionais, no Gráfico, podemos observar os dois cenários computacionais apresentando taxas satisfatórias com execuções até 100 robôs para a Computação em Nuvem e 150 robôs para Computação de Borda. Na execução com apenas 1 robô, as duas abordagens apresentaram taxas melhores que na execução com 2 robôs. Isso se dá pelo fato de a primeira execução os serviços internos dos sistemas precisarem gerar estrutura de dados, visões e rotinas computacionais no Data Lake, Mongo e Data Warehouse.



**Figura 7. Taxa de Sucesso - Computação de Borda X Computação em Nuvem.**

A taxa de sucesso com Computação em Nuvem apresenta uma queda elevada a partir de 50 robôs executando simultaneamente, o que é explicado por dois fatores: i) a coleta de dados para o ambiente em Nuvem funciona de maneira assíncrona, portanto, os dados de telemetria ficam desorganizados dentro do Data Lake, sendo necessário uma etapa de ordenação nos dados antes dos processos de ETL; ii) maior latência de rede intrínseca de ambientes em nuvem. O aumento elevado e constante de arquivos a serem processados influencia diretamente no tempo gasto para transferência e ordenação destes dados. Como resultado, podemos observar no gráfico 7.a uma taxa de sucesso negativa de -31,62% no cenário de Computação em Nuvem com coleta de 150 robôs. Estes dois fatores não estão presentes no cenário de Computação de Borda.

A Figura 8 compara os resultados e mostra a porcentagem de desempenho no uso de Computação de Borda em relação à Computação em Nuvem. No gráfico, temos a representação da projeção da taxa de sucesso da abordagem nos dois cenários computacionais e a projeção do desempenho em utilizar Computação de Borda ao invés de Computação em Nuvem. Pode-se observar que, em ambas as abordagens, há um declínio na taxa de sucesso como já apresentado anteriormente. No entanto, esse declínio é menor usando Computação de Borda. Ao analisar o teste em Nuvem com 50 robôs, a taxa de sucesso é de 26,17%, enquanto na Computação de Borda a taxa alcança 71,20%. Neste mesmo caso, pode-se observar que o uso de Computação de Borda mostra um desempenho superior de 45,03% em relação à Computação em Nuvem, e, no caso mais extremo, com 150 robôs, temos uma melhoria de 76,18%.



**Figura 8. Desempenho - Computação de Borda X Computação em Nuvem.**

## 6. Conclusão e Trabalhos Futuros

Este trabalho apresentou a abordagem ThreeVS-industry para integrar dispositivos industriais e sistemas automatizados para manter uma coleta e processamento de dados de forma contínua. A abordagem se mostrou útil e eficiente, permitindo coletar dados de chão de fábrica e armazenar essas informações para trabalhos futuros. Através dessa abordagem de integração contínua, pode-se automatizar processos de transformação e análise de dados, gerar informações e insights, que servirão de base para aprimoramentos contínuos em processos e planejamentos estratégicos.

O mercado hoje exige análises e respostas rápidas, a Computação em Nuvem não consegue acompanhar essa demanda devido à latência de rede gerada pela distância a fonte de dados e a necessidade de ordenação nos dados coletados. Aumentar o poder de processamento e largura de banda pode resolver o problema, no entanto, a quantidade de dados gerados e o custo para enviar para a nuvem e ordenar, podem chegar a níveis impraticáveis. Os resultados de simulação em dois cenários computacionais mostraram que, utilizar a abordagem ThreeVS-industry implementada em Nuvem, gera problemas de desempenho em casos de integração de dispositivos em excesso, enquanto a implementação com Computação de Borda se mostra promissora e eficiente com ótimas taxas de sucesso e desempenho para tratar dados em larga escala gerados pelos processos de manufatura. Para aqueles com poucos dispositivos industriais e com um cenário que se precisa manter os dados e transformá-los para operações e serviços futuros como relatórios, adotar uma abordagem em Nuvem se torna viável até uma certa quantidade de dispositivos, entretanto, a necessidade de análise preditiva deve ser controlada. Como trabalhos futuros, pode-se adicionar uma terceira etapa, compostas por técnicas de predição de falhas, relatórios, dashboards ou qualquer outro processo que gere valor para os gestores que administram processos industriais.

## Referências

- Ayub, D., de Assis Silva, W., Marques, M. A. M., Kleina, M., and Drozda, F. O. (2021). Análise de big data por meio de estatísticas multivariadas na indústria 4.0: uma revisão da literatura. *Revista Competitividade e Sustentabilidade*, 8(1):3–16.
- Coutinho, A., Carneiro, E. O., and Greve, F. G. P. (2016). Computação em névoa: Conceitos, aplicações e desafios. *Minicursos do XXXIV SBRC*, pages 266–315.
- de Oliveira, L. E. S. and Álvares, A. J. (2018). Development of a system for monitoring and teleoperation cnc machine tools through the internet adherent to industry 4.0. *Revista Produção E Desenvolvimento*.
- Erl, T. (1900). *Service-oriented architecture*. Pearson Education Incorporated Upper Saddle River.
- Fabricio, M. A. (2018). Monitoring of industrial electrical equipment using iot. *Repositório Institucional PUC-Campinas*.
- Ferreira, J., Miranda, M., Abelha, A., and Machado, J. (2010). O processo etl em sistemas data warehouse. In *INForum*, pages 757–765.
- Fumagalli, M. A., Branco, A. A. C., and Dias, W. S. (2021). Implementação do protocolo opc ua para controle de uma célula de manufatura, utilizando framework web voltado para indústria 4.0. *FTT Journal of Engineering and Business*, 1(6).

- Givehchi, O., Landsdorf, K., Simoens, P., and Colombo, A. W. (2017). Interoperability for industrial cyber-physical systems: An approach for legacy systems. *IEEE Transactions on Industrial Informatics*, 13(6):3370–3378.
- Gori, R. S. L. and Gori, D. D. L. (2022). Smart factory e a indústria 4.0: uma revisão sistemática de literatura. *Revista Sítio Novo*, 6(2):141–155.
- Han, Y. and Song, Y. (2003). Condition monitoring techniques for electrical equipment—a literature survey. *IEEE Transactions on Power Delivery*, 18(1):4–13.
- Jardine, A. K., Lin, D., and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, 20(7):1483–1510.
- Kim, W. and Sung, M. (2018). Standalone opc ua wrapper for industrial monitoring and control systems. *IEEE Access*, 6:36557–36570.
- Monteiro, A. V. G., Pinto, M. P. O., and da Costa, R. M. E. M. (2004). Uma aplicação de data warehouse para apoiar negócios. *Cadernos do IME-Série Informática*, 16:48–58.
- Ribeiro, C. J. S. (2014). Big data: os novos desafios para o profissional da informação. *Informação & Tecnologia*, 1(1):96–105.
- Rodrigues, J. X. and dos Santos Mello, R. (2022). Um estudo sobre arquiteturas e metadados em data lakes. In *Anais do XVII Escola Regional de Banco de Dados*, pages 131–134. SBC.
- Santos, B. P., Alberto, A., Lima, T. D. F. M., and Charrua-Santos, F. M. B. (2018). Indústria 4.0: Desafios e oportunidades. *Revista Produção e Desenvolvimento*, 4(1):111–124.
- Schleipen, M., Gilani, S.-S., Bischoff, T., and Pfrommer, J. (2016). Opc ua industrie 4.0 - enabling technology with high diversity and variability. *Procedia CIRP*, 57:315–320. Factories of the Future in the digital environment - Proceedings of the 49th CIRP Conference on Manufacturing Systems.
- Schwarz, M. H. and Börcsök, J. (2013). A survey on opc and opc-ua: About the standard, developments and investigations. In *2013 XXIV International Conference on Information, Communication and Automation Technologies (ICAT)*, pages 1–6. IEEE.
- Silva, A. G. and Michel, H. C. C. (2021). Gateway lorawan com conexão opc ua para dispositivos iot industriais. In *Simpósio Brasileiro de Automação Inteligente-SBAI*, volume 1.
- Tessarini, Geraldo e Saltorato, P. (2018). Impactos da indústria 4.0 na organização do trabalho: uma revisão sistemática da literatura. *Revista Produção Online*, 18(2):743–769.