O uso do versionamento de dados para rastreabilidade de parâmetros

Autor(res)

Arthur Abreu Rodrigues

Categoria do Trabalho

Trabalho Acadêmico

Instituição

FACULDADE ANHANGUERA

Introdução

Nos últimos anos, a indústria passou a lidar com um cenário de produção cada vez mais detalhado e dependente de informação. Não basta mais apenas fabricar uma peça; é preciso saber como ela foi feita, quais parâmetros mudaram ao longo do caminho e quais versões ficaram registradas. Esse tipo de acompanhamento é o que chamamos de rastreabilidade de parâmetros.

De forma geral, existem dois modos de registrar essas informações. O primeiro é o modelo convencional, que consiste em criar versões a partir de cópias das versões anteriores, por exemplo, se faço uma alteração em um parâmetro, as novas versões vão ser a nova alteração, mais uma copia dos dados das alterações feitas anteriormente, gerando dessa forma redundância de dados e até células vazias. Já no segundo modelo, temos o versionamento por snapshots, que funciona como uma fotografia, apenas o que mudou é salvo enquanto os dados antigos são aproveitados, e não duplicados.

A diferença entre esses dois métodos pode parecer pequena em um primeiro momento, mas se torna gigantesca quando pensamos em escala industrial. Imagine uma fábrica que produz dezenas de milhares de peças por ano, cada uma passando por várias revisões. O espaço de armazenamento e os custos de manter tantas versões completas podem se tornar um problema real.

É nesse ponto que o modelo por snapshots ganha relevância. Além de reduzir o consumo de memória, ele garante que o histórico completo da peça seja preservado sem a necessidade de duplicar dados que não mudaram. Esse ganho técnico se traduz em economia financeira, mas também em maior clareza na auditoria e na gestão de qualidade.

O objetivo deste artigo é analisar de forma prática os impactos de cada abordagem, mostrando como o versionamento por snapshots pode oferecer vantagens concretas para empresas que precisam lidar com grandes volumes de informação.

Objetivo

Este trabalho analisa dois modelos de versionamento de dados no setor industrial: o convencional, que replica todo o conteúdo a cada alteração, e o de snapshots, que guarda só o que muda, como um retrato fiel do instante.

A proposta é avaliar, em cenários de grande volume, qual opção reduz custos, otimiza memória e reforça a rastreabilidade.

Material e Métodos

O estudo foi conduzido como uma espécie de ensaio comparativo, simulando o desempenho de dois modelos de versionamento de dados no contexto do gerenciamento de peças industriais. O primeiro, que poderíamos chamar de tradicional, segue a lógica de copiar tudo a cada nova versão. Nesse arranjo, cada parâmetro ocupa 1 GB e cada peça carrega seis parâmetros, o que resulta num custo de 6 GB por versão. Ao longo de seis versões, a conta chega a nada menos que 36 GB — um verdadeiro peso morto quando pensamos em escala. O segundo modelo, baseado em snapshots, adota outra filosofia: registrar apenas o que realmente mudou, preservando o restante como se fosse uma fotografia fiel do instante. Admitindo-se que, em média, apenas um parâmetro sofra alteração, cada versão consome 1 GB, acumulando 6 GB após seis versões. Os metadados adicionais, como IDs e timestamps, entram nessa equação quase como meros figurantes, já que ocupam espaço insignificante perto de arquivos de 1 GB.

Para aproximar a análise da realidade do chão de fábrica, definiu-se um cenário de simulação em que o volume anual de produção alcança 50.000 peças, cada uma delas com seis versões. O tamanho por parâmetro foi fixado em 1 GB e o ambiente de armazenamento escolhido foi o AWS S3 Standard. Aqui vale um parêntese: os preços da nuvem também entram como protagonistas. Consideraram-se os valores de setembro de 2025, variando de US$ 0,023 por GB/mês (até 50 TB) a US$ 0,021 por GB/mês (acima de 500 TB).

Os cálculos são reveladores: no modelo convencional, o montante acumulado ao final de um ano chega a 1.800.000 GB, enquanto no modelo por snapshots o total cai drasticamente para 300.000 GB. Em suma, é quase como assistir a uma balança pender nitidamente para um lado, revelando de forma clara qual modelo traz mais eficiência e economia quando o assunto é armazenamento em larga escala.

Resultados e Discussão

Os resultados da simulação evidenciaram diferenças expressivas entre os dois modelos de versionamento analisados. No cenário convencional, em que todos os parâmetros são replicados a cada versão, o consumo anual de armazenamento atingiu cerca de 1.800 terabytes. Já no modelo por snapshots, considerando a hipótese de apenas um parâmetro alterado por versão, o volume caiu drasticamente para 300 terabytes por ano. A economia absoluta foi de 1.500 terabytes, o que equivale a uma redução percentual em torno de 83,3%. Em termos práticos, isso significa que, a cada seis versões de uma mesma peça, o gasto de memória pode ser reduzido de seis vezes o tamanho integral para apenas o necessário, como se fosse substituir o hábito de reescrever um livro inteiro pela simples anotação de uma correção em suas margens.

Quando os dados foram projetados para um ambiente de nuvem AWS S3, os contrastes tornaram-se ainda mais claros. No modelo convencional, o custo mensal médio de armazenamento foi estimado em aproximadamente 38.350 dólares, resultando em um gasto anual de 460.200 dólares no cenário conservador, em que todo o volume final é considerado durante os 12 meses. No cenário realista, em que se admite o crescimento gradual dos dados ao longo do ano e se utiliza o volume médio como referência, o custo anual caiu para 233.400 dólares. Em contrapartida, no modelo por snapshots, o custo mensal foi estimado em apenas 6.650 dólares, representando um total de 79.800 dólares no cenário conservador e de 40.200 dólares no cenário realista. Dessa forma, a economia anual obtida com snapshots atingiu 380.400 dólares no cenário A e 193.200 dólares no cenário B, valores que,

quando acumulados em períodos plurianuais, podem significar a diferença entre um gargalo operacional e uma gestão sustentável de dados.

A análise de sensibilidade reforçou a robustez da abordagem por snapshots, ainda que seus ganhos diminuam à medida que cresce o número de parâmetros alterados por versão. Quando apenas um parâmetro sofre modificação, a economia é da ordem de 83%. Se dois parâmetros são alterados, a economia cai para 66%; com três, para 50%; com quatro, para 33%; e com cinco, para 17%. Já no limite de seis parâmetros modificados simultaneamente, os custos se equiparam ao modelo convencional. Na prática, esse comportamento indica que snapshots são especialmente vantajosos em ambientes industriais onde as versões tendem a ajustar detalhes pontuais, como tolerâncias dimensionais, acabamentos superficiais ou revisões de inspeção, sem a necessidade de redefinir integralmente todos os parâmetros de uma peça.

Além da expressiva redução em custos e volumes de dados, os snapshots oferecem benefícios qualitativos que vão além da contabilidade de gigabytes. Por funcionarem como verdadeiras fotografias imutáveis do estado da peça em determinado momento, tornam a rastreabilidade mais simples e confiável, aspecto fundamental em auditorias ou na comprovação de conformidade regulatória. Outro ponto é a simplificação nas consultas, já que versões anteriores podem ser recuperadas sem a necessidade de filtrar registros redundantes ou artificiais, prática comum em bancos de dados baseados em replicações integrais. Há ainda a questão da escalabilidade: com menores volumes acumulados, operações de backup, replicação e análise distribuem-se de forma mais ágil, contribuindo para a eficiência do sistema como um todo.

Contudo, essa vantagem não vem sem desafios. O modelo por snapshots exige maior rigor no controle da integridade referencial, uma vez que cada versão depende de vínculos entre diferentes tabelas. Se tais vínculos forem corrompidos ou mal gerenciados, corre-se o risco de perda de consistência histórica, o que pode comprometer a confiabilidade dos registros. Para mitigar esse risco, é essencial adotar boas práticas de indexação, mecanismos de validação automática e rotinas de auditoria periódicas, garantindo que os benefícios do modelo não sejam anulados por falhas de gestão.

Conclusão

O estudo demonstrou que o versionamento por snapshots reduz drasticamente consumo de armazenamento e custos em larga escala, mantendo rastreabilidade e versões imutáveis. Mesmo com alterações em alguns parâmetros, a economia permanece significativa. Recomenda-se sua adoção em sistemas industriais de médio e grande porte, acompanhada de boas práticas de governança e políticas de retenção, garantindo eficiência, escalabilidade e confiabilidade histórica.

Referências

AWS. Amazon Simple Storage Service (S3) Pricing. Disponível em: https://aws.amazon.com/s3/pricing/ . Acesso em: set. 2025.

BERNARDES, M. M.; GONÇALVES, R. A. Sistemas de gerenciamento do ciclo de vida do produto (PLM). Revista Produção Online, v. 14, n. 4, 2014.

CHAPMAN, A.; WILES, A. Data versioning in industrial environments: ensuring traceability and compliance. Journal of Industrial Information Integration, v. 22, 2021.

COUTINHO, P. C.; SOUZA, R. M. Auditoria e rastreabilidade em sistemas industriais digitais. Revista de Engenharia de Produção, v. 28, n. 2, 2020.

JACOBSON, I.; NGUYEN, P. Product lifecycle management in the era of Industry 4.0. Springer, 2019.

MENDES, R. T.; ALMEIDA, V. J. Gestão de dados industriais: práticas e desafios em ambientes digitais. Anais do Congresso Brasileiro de Engenharia de Produção, 2022.

SANTOS, H. F.; OLIVEIRA, J. A. Estratégias de armazenamento em nuvem para sistemas industriais. Journal of Cloud Computing, v. 12, 2023.