Multidisciplinary Sciences Reports

Artigo Original

ARQUITETURA EM NUVEM COM AWS: UMA PROPOSTA ESCALÁVEL, SEGURA E DE BAIXO CUSTO

CLOUD ARCHITECTURE WITH AWS: A SCALABLE, SECURE AND LOW-COST PROPOSAL

1 Igor Resende igorpeli12@gmail.com

1- Universidade Vila Velha (UVV)

RESUMO

Este trabalho apresenta uma proposta de arquitetura em nuvem utilizando os serviços da Amazon Web Services (AWS), com foco em escalabilidade, segurança, disponibilidade da aplicação e otimização do custo operacional. A proposta busca demonstrar como ambientes baseados em nuvem podem oferecer vantagens em relação a arquiteturas tradicionais on-premises, especialmente quanto à flexibilidade de recursos, agilidade de implantação e redução da complexidade operacional. Reconhece-se, porém, que a eficácia da computação em nuvem depende da qualificação dos profissionais responsáveis pelo planejamento e pela configuração dos serviços. A ausência de conhecimento técnico pode comprometer aspectos críticos como segurança, disponibilidade da aplicação e controle de custos. O método consistiu na proposição de uma arquitetura em três camadas: balanceamento, aplicação e banco de dados, com uso de Virtual Private Cloud (VPC), sub-redes públicas e privadas, Amazon Elastic Container Service (Amazon ECS) com Fargate, Amazon Relational Database Service (Amazon RDS), Amazon Simple Storage Service (Amazon S3), Identity and Access Management (IAM) e AWS CloudFormation. A análise técnica indica que, quando documentada e padronizada por Infrastructure as Code (IaC), a arquitetura se torna mais replicável, compreensível e adequada a equipes com menor experiência.

PALAVRAS-CHAVE: Arquitetura em Nuvem. AWS. Escalabilidade. Infraestrutura como Código. Computação em Nuvem.

ABSTRACT

This paper presents a cloud architecture proposal using Amazon Web Services (AWS), focusing on scalability, security, application availability, and operational cost optimization. The proposal aims to show how cloud-based environments may offer advantages over traditional on-premises infrastructures, especially regarding resource flexibility, deployment agility, and reduced operational complexity. However, cloud computing effectiveness depends on the qualifications of the professionals responsible for planning and configuring services. A lack of technical knowledge may compromise critical aspects such as security, application availability, and cost control. The method consisted of proposing a three-layer architecture: load balancing, application, and database, using Virtual Private Cloud (VPC), public and private subnets, Amazon Elastic Container Service (Amazon ECS) with Fargate, Amazon Relational Database Service (Amazon RDS), Amazon Simple Storage Service (Amazon S3), Identity and Access Management (IAM), and AWS CloudFormation. The technical analysis indicates that, when documented and standardized through Infrastructure as Code (IaC), the architecture becomes more replicable, understandable, and suitable for teams with less experience.

KEYWORDS: AWS. Scalability. Infrastructure as Code. Cloud Computing.

INTRODUÇÃO

Tradicionalmente, a computação era realizada em infraestrutura interna, na qual a própria organização detinha responsabilidade pelo gerenciamento de hardware, software, recursos computacionais, segurança e disponibilidade da aplicação. Nesse modelo, chamado de *on-premises*, cabia à empresa comprar, instalar, gerir e manter seus recursos. Contudo, essa abordagem envolve investimento inicial elevado, escalabilidade limitada, maior complexidade de gerenciamento e maior exposição a falhas operacionais ou desastres naturais.

A computação em nuvem, segundo o *National Institute of Standards and Technology* (1), é um modelo que permite acesso remoto, sob demanda e flexível a um conjunto compartilhado de recursos computacionais, como redes, servidores, armazenamento e aplicações, que podem ser rapidamente alocados e liberados com mínimo esforço de gerenciamento ou intervenção do provedor.

Tal arquitetura tornou-se possível graças à virtualização, tecnologia que permite criar representações virtuais de recursos físicos, como máquinas, contêineres, bancos de dados e redes. Com essa abordagem, é possível executar múltiplas máquinas virtuais em um único hardware físico, mantendo isolamento

lógico entre os ambientes. Essa técnica possibilita melhor aproveitamento da infraestrutura física, aumento de eficiência, redução de desperdícios operacionais e ampliação da escalabilidade (2).

Atualmente, observa-se um número crescente de organizações que buscam soluções tecnológicas baseadas em arquiteturas de computação em nuvem, recorrendo a provedores como a *Amazon Web Services* (AWS). Essa tendência está apoiada na necessidade de redução de custos operacionais, facilidade de manutenção, disponibilidade da aplicação e escalabilidade. Apesar das vantagens oferecidas, a adoção desse modelo também apresenta desafios relacionados a custos inesperados, segurança, integração com sistemas legados e escassez de profissionais qualificados (3).

Com a crescente adoção da computação em nuvem, evidencia-se que a simples migração para a nuvem não garante, por si só, segurança, disponibilidade da aplicação ou eficiência operacional. Entre os principais desafios estão custos inesperados, complexidade de integração com sistemas existentes, riscos de segurança cibernética e necessidade de planejamento adequado. Assim, a efetividade da nuvem depende da compreensão dos serviços utilizados, dos modelos de cobrança e das boas práticas de arquitetura (4).

A escassez de profissionais capacitados em arquitetura de nuvem reforça a necessidade de propostas documentadas e padronizadas. Estudos recentes apontam que 98% das organizações enfrentam dificuldades para encontrar profissionais devidamente qualificados em nuvem (3). Esse cenário indica que a adoção da nuvem depende não apenas da disponibilidade tecnológica, mas também da capacidade técnica das equipes responsáveis por sua implantação, manutenção e governança.

Parte-se da hipótese de que uma arquitetura bem documentada, baseada em serviços gerenciados e em *Infrastructure as Code* (IaC), pode reduzir a complexidade operacional e facilitar a adoção da nuvem por equipes com menor experiência. Essa hipótese orienta a construção de uma proposta arquitetural organizada, compreensível, replicável e alinhada a boas práticas.

O objetivo deste artigo é propor uma arquitetura em nuvem utilizando serviços

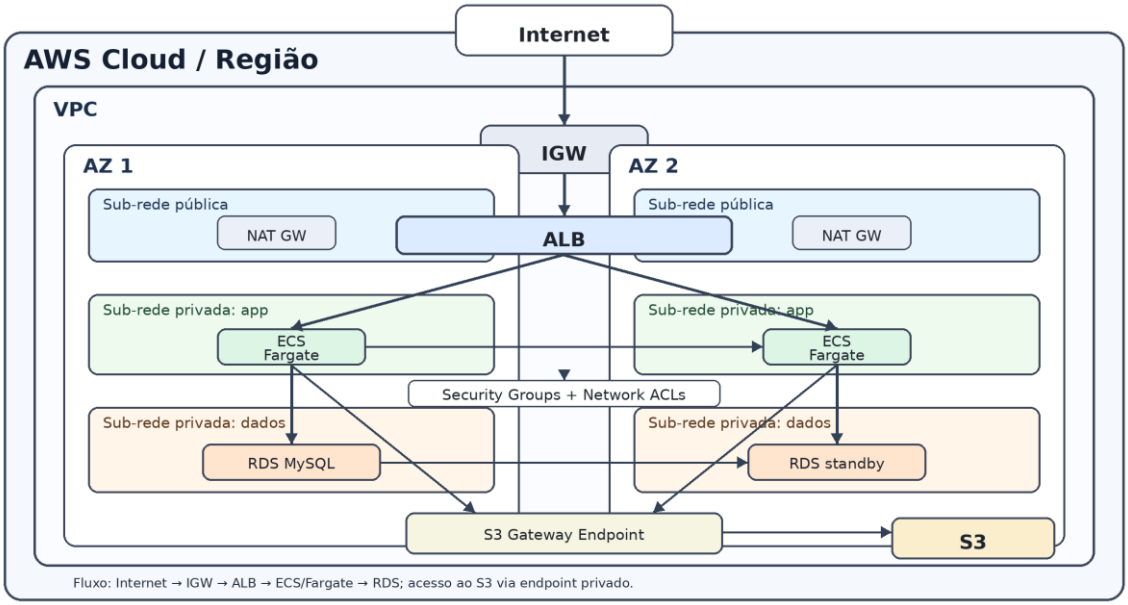
da AWS que ofereça segurança, disponibilidade da aplicação e otimização de custo operacional, com base em boas práticas, de forma que possa ser replicada por diferentes organizações com ajustes mínimos conforme suas particularidades.

MATERIAL E MÉTODOS

A metodologia adotada baseia-se na proposição técnica de uma arquitetura em nuvem estruturada em três camadas e planejada para uso na AWS por meio de serviços gerenciados. A proposta foi concebida para demonstrar, de forma aplicada, como diferentes componentes podem ser integrados para favorecer segurança, disponibilidade da aplicação e otimização do custo operacional.

A arquitetura proposta foi organizada em três camadas principais: camada de balanceamento, camada de aplicação e camada de banco de dados. Cada camada foi distribuída entre sub-redes e zonas de disponibilidade, refletindo uma estratégia voltada à resiliência. A Figura 1 apresenta a organização geral da solução e a relação entre os principais componentes.

Figura 1 – Arquitetura geral da solução na AWS, com VPC (Virtual Private Cloud), ALB (Application Load Balancer), ECS (Elastic Container Service), RDS (Relational Database Service), S3 (Simple Storage Service), IAM (Identity and Access Management), NAT (Network Address Translation), IGW (Internet Gateway) e AZ (Availability Zone). Fonte: Próprio autor.



Na camada de balanceamento e acesso externo, a proposta prevê o posicionamento do *Application Load Balancer* (ALB) nas sub-redes públicas, conforme indicado na Figura 1. O ALB distribui as requisições recebidas pela internet entre as tarefas da aplicação localizadas nas sub-redes privadas, permitindo exposição controlada apenas do ponto de entrada da aplicação.

A base de rede da proposta é a *Virtual Private Cloud* (VPC), que representa uma rede virtual logicamente isolada dentro da AWS. A VPC é organizada em sub-redes públicas e privadas distribuídas em zonas de disponibilidade distintas. As sub-redes públicas A e B correspondem aos segmentos destinados ao balanceamento e aos *gateways*, enquanto as sub-redes privadas A e B correspondem aos segmentos destinados à aplicação e ao banco de dados, conforme a Figura 1 (5).

As sub-redes públicas são destinadas aos componentes que precisam de comunicação direta com a internet, como o balanceador de carga e os *gateways*. As sub-redes privadas são reservadas a componentes sensíveis, como contêineres da aplicação e banco de dados. Esse isolamento reduz a superfície de exposição da infraestrutura e é coerente com a recomendação de segmentar recursos e controlar os caminhos de comunicação na VPC (5).

Para permitir acesso controlado à internet, a proposta prevê um *Internet Gateway* associado à VPC. No caso das sub-redes privadas, o acesso de saída pode ser realizado por meio de um *Network Address Translation Gateway* (NAT Gateway), permitindo que recursos privados acessem serviços externos sem receber conexões iniciadas diretamente pela internet (5).

Na camada de aplicação, propõe-se a utilização de contêineres em modo *Fargate no Amazon Elastic Container Service* (Amazon ECS). O Amazon ECS é um serviço gerenciado de orquestração de contêineres, enquanto o *Fargate* permite executar esses contêineres sem necessidade de provisionar, configurar ou escalar diretamente servidores ou clusters de máquinas virtuais (6).

Para garantir o balanceamento de carga entre as tarefas em execução, a proposta utiliza um *Application Load Balancer* (ALB). O ALB opera de forma redundante quando associado a sub-redes públicas em múltiplas zonas de disponibilidade, distribuindo o tráfego entre as tarefas disponíveis da aplicação e reduzindo a dependência de um único ponto de atendimento.

Na camada de banco de dados, propõe-se o uso do *Amazon Relational*

Database Service (Amazon RDS) como banco de dados gerenciado. O acesso ao banco deve ocorrer apenas a partir da camada de aplicação, por meio de regras específicas em *Security Groups*, evitando exposição direta à internet. Em cenários que exigem maior disponibilidade da aplicação, o RDS pode ser planejado com configuração Multi-AZ, de acordo com os requisitos do ambiente (7).

O *Amazon Simple Storage Service* (Amazon S3) é utilizado como repositório de objetos, como arquivos estáticos, imagens e documentos da aplicação. Embora apareça relacionado à arquitetura na Figura 1, o S3 não é posicionado dentro da VPC nem em sub-rede pública, pois é um serviço regional da AWS acessado por *endpoints*. Para acesso privado a partir da VPC, a proposta prevê o uso de VPC *Gateway Endpoint* para Amazon S3 (8).

Também são previstas *Identity and Access Management Roles* (IAM Roles) para permitir que serviços em execução, como os contêineres hospedados no Amazon ECS com Fargate, acessem outros recursos da AWS sem uso de credenciais fixas no código da aplicação. As IAM Roles representam identidades com permissões específicas e temporárias, assumidas por serviços autorizados conforme a política definida (9).

Para complementar a segurança da VPC, a proposta combina *Security Groups* e *Network Access Control Lists* (Network ACLs). Enquanto os *Security Groups* atuam no nível dos recursos e controlam tráfego de entrada e saída com estado, as Network ACLs atuam no nível das sub-redes e controlam regras de tráfego de forma sem estado. A combinação desses mecanismos permite definir camadas complementares de controle de comunicação (5).

Todas as configurações de infraestrutura são propostas para descrição por meio do AWS *CloudFormation*, consolidando a abordagem de *Infrastructure as Code* (IaC). Na AWS, o *CloudFormation* permite declarar recursos em modelos escritos em JSON ou YAML, possibilitando padronização, controle de versão e repetição do ambiente em diferentes contextos (10,11).

Com essa abordagem, recursos como VPCs, sub-redes, gateways, grupos de segurança, serviços do Amazon ECS, Amazon RDS e Amazon S3 podem ser declarados em arquivos reutilizáveis. A finalidade não é afirmar que houve implantação produtiva com mensuração de desempenho, mas apresentar uma proposta arquitetural replicável e tecnicamente organizada.

RESULTADOS E DISCUSSÕES

A análise técnica da arquitetura proposta resultou em um modelo estruturado em três camadas: balanceamento, aplicação e banco de dados. Essa separação organiza os componentes de acordo com suas responsabilidades e níveis de exposição, mantendo o ponto de entrada da aplicação nas sub-redes públicas e os recursos sensíveis, como aplicação e banco de dados, nas sub-redes privadas.

Como resultado da modelagem lógica da rede, a VPC foi organizada com sub-redes públicas e privadas, tabelas de rotas específicas e componentes de conectividade adequados a cada finalidade. As sub-redes públicas foram associadas ao Internet Gateway para permitir a exposição controlada do ALB, enquanto as sub-redes privadas foram direcionadas à camada de aplicação e ao banco de dados, evitando exposição direta à internet.

Essa organização reforça o princípio de segmentação de rede e reduz a superfície de exposição da infraestrutura. Ao separar componentes públicos e privados, a arquitetura preserva banco de dados e aplicação em segmentos restritos, permitindo que o acesso externo seja concentrado no balanceador de carga e controlado por regras de segurança.

Na camada de aplicação, a utilização do Amazon ECS com Fargate é adequada para reduzir atividades operacionais, pois elimina a necessidade de gerenciamento direto de servidores. Assim, a execução dos contêineres passa a depender de uma camada gerenciada pela AWS, favorecendo a replicabilidade da proposta e permitindo que a equipe concentre esforços na configuração da aplicação e das políticas de acesso (6).

A camada de banco de dados foi estruturada com o Amazon RDS em sub-redes privadas, permitindo que o acesso ao banco ocorra apenas a partir da camada de aplicação, conforme representado na Figura 1. Essa escolha reduz riscos associados à exposição direta do banco de dados. O Amazon S3, por sua vez, foi mantido como serviço regional externo à VPC e acessado logicamente pela aplicação por meio de *endpoint*, evitando sua representação como recurso interno de sub-rede.

A utilização de VPC *Gateway Endpoint* para acesso ao Amazon S3 contribui para evitar que determinados fluxos dependam da internet pública ou do NAT Gateway, uma vez que o tráfego entre a VPC e o S3 pode ser roteado por um caminho privado suportado pela AWS (8). Essa característica é relevante para a proposta por reduzir dependências de conectividade externa em fluxos relacionados ao armazenamento de objetos.

Em relação à segurança, a combinação entre *Security Groups* e *Network* ACLs possibilita controles complementares de tráfego. Os *Security Groups* controlam acessos diretamente nos recursos, enquanto as Network ACLs atuam nas sub-redes. Essa combinação permite estabelecer regras distintas para cada camada da aplicação e reforça a segmentação indicada na arquitetura (5).

Outro ponto relevante da proposta é a padronização da infraestrutura por meio do AWS *CloudFormation*. A utilização de *Infrastructure as Code* permite descrever recursos em arquivos reutilizáveis, facilitando versionamento, revisão e replicação do ambiente. Essa característica é especialmente importante para equipes com menor experiência, pois reduz a dependência de configurações manuais no console da AWS (10,11).

A análise da arquitetura também evidencia que a adoção da nuvem não garante, isoladamente, segurança, economia ou disponibilidade da aplicação. Esses benefícios dependem da forma como os serviços são configurados, documentados e governados. Uma arquitetura em nuvem mal planejada pode gerar exposição indevida de recursos, custos desnecessários e dificuldade de manutenção; por outro lado, a adoção de boas práticas, segmentação de rede, serviços gerenciados e automação tende a reduzir a complexidade operacional (4).

No que diz respeito à disponibilidade da aplicação, a distribuição dos recursos entre diferentes zonas de disponibilidade reduz pontos únicos de falha. Alguns serviços, como o Amazon S3, apresentam resiliência regional própria, enquanto outros exigem decisões específicas de configuração, como RDS em Multi-AZ e distribuição das tarefas do Amazon ECS em múltiplas zonas. Dessa forma, a proposta organiza componentes para favorecer continuidade do serviço, sem afirmar medição experimental de disponibilidade.

Quanto à replicabilidade, o uso do AWS *CloudFormation* representa um dos principais elementos da proposta. Ao documentar a infraestrutura como código, a solução deixa de depender exclusivamente de configurações manuais e passa a ser reproduzível em outros ambientes. Isso torna a arquitetura mais adequada a organizações que precisam de padronização e previsibilidade na criação de ambientes em nuvem.

Dessa forma, a proposta atende ao objetivo de organizar serviços da AWS em uma arquitetura escalável, segura e replicável. O artigo não busca comprovar redução financeira por meio de métricas de custo, mas apresentar uma estrutura que pode contribuir para a redução de complexidade operacional ao priorizar serviços gerenciados, documentação arquitetural, controle de acesso e infraestrutura declarativa.

CONCLUSÃO

A arquitetura proposta demonstra que serviços gerenciados da AWS podem ser organizados em um modelo de nuvem estruturado, com foco em escalabilidade, segurança, disponibilidade da aplicação e otimização do custo operacional. A proposta não apresenta validação quantitativa de desempenho ou custos financeiros, mas oferece uma organização arquitetural fundamentada em boas práticas e adequada à replicação em diferentes contextos.

O principal aspecto que viabiliza a proposta para equipes com menor experiência é a combinação entre documentação arquitetural e *Infrastructure as Code*. Ao descrever a infraestrutura de forma declarativa, padronizada e reutilizável, a equipe reduz a necessidade de configurações manuais sucessivas, diminui ambiguidades de implantação e facilita a revisão dos componentes antes da criação do ambiente.

Conclui-se que a proposta cumpre o objetivo de apresentar uma arquitetura em nuvem baseada em boas práticas, adequada a cenários que exigem segurança, disponibilidade da aplicação e controle operacional. Também se reforça que a eficiência da nuvem depende diretamente do planejamento, da padronização e da

correta configuração dos serviços, especialmente em organizações com limitações de equipe e conhecimento técnico.

CONFLITO DE INTERESSES

Declara-se não haver conflito de interesses relacionado ao conteúdo apresentado neste artigo.

REFERÊNCIAS

1. Mell P, Grance T. The NIST Definition of Cloud Computing. Gaithersburg: National Institute of Standards and Technology; 2011. NIST Special Publication 800-145. Disponível em: https://csrc.nist.gov/pubs/sp/800/145/final . Acesso em: 13 maio 2026.

2. Marinescu DC. Cloud Computing: Theory and Practice. 3rd ed. Cambridge: Morgan Kaufmann; 2022.

3. SoftwareOne. Cloud Skills Report. 2023. Disponível em: https://www.softwareone.com/en/media-releases/2023/11/08/softwareone-publishes-cloud-skills-report . Acesso em: 13 maio 2026.

4. Amazon Web Services. AWS Well-Architected Framework. 2024. Disponível em: https://docs.aws.amazon.com/wellarchitected/latest/framework/welcome.html . Acesso em: 13 maio 2026.

5. Amazon Web Services. Amazon Virtual Private Cloud: User Guide. Disponível em: https://docs.aws.amazon.com/vpc/latest/userguide/what-is-amazon-vpc.html . Acesso em: 13 maio 2026.

6. Amazon Web Services. Architect for AWS Fargate for Amazon ECS. Disponível em: https://docs.aws.amazon.com/AmazonECS/latest/developerguide/AWS_Fargat

e.html . Acesso em: 13 maio 2026.

7. Amazon Web Services. Creating Amazon RDS resources with AWS CloudFormation. Disponível em: https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/creating-resources-with-cloudformation.html . Acesso em: 13 maio 2026.

8. Amazon Web Services. Gateway endpoints for Amazon S3. Disponível em: https://docs.aws.amazon.com/vpc/latest/privatelink/vpc-endpoints-s3.html . Acesso em: 13 maio 2026.

9. Amazon Web Services. IAM roles. Disponível em: https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles.html . Acesso em: 13 maio 2026.

10. Amazon Web Services. What is CloudFormation? Disponível em: https://docs.aws.amazon.com/AWSCloudFormation/latest/UserGuide/Welcome.html . Acesso em: 13 maio 2026.

11. Verma A. Infrastructure as Code. Birmingham: Packt Publishing; 2023.