RCMOS – Revista Científica Multidisciplinar O Saber.
ISSN: 2675-9128. São Paulo-SP.

Ano VII, v.1 2026 | submissão: 14/02/2026 | aceito: 17/02/2026 | publicação: 20/02/2026

Arquiteturas 5g core nativas em nuvem cloud-native 5g core architectures
Cloud-native 5g core architectures
Arquitecturas de núcleo 5G nativas en la nube

Danyell Teixeira Soares
Engenheiro de telecomunicações com 16 anos de experiência. MBA em Gestão Empresarial pela Fundação Getulio Vargas (FGV-SP, 2015). Graduado em Engenharia Elétrica com ênfase em Telecomunicações pela Pontifícia Universidade Católica de São Paulo (PUC-SP, 2008).

Resumo: Este artigo analisa a transformação técnica e estrutural do core móvel nas redes 5G Standalone (5G SA), focando na transição de elementos legados para uma Arquitetura Baseada em Serviços (SBA) e funções nativas em nuvem (CNFs). Explicam-se as interações entre AMF, SMF, UDM, AUSF, PCF, NRF e UPF sob o protocolo HTTP/2, a orquestração via Kubernetes em modelos stateless e as estratégias de segurança orientadas a Zero Trust e Service Mesh. O estudo adota como metodologia a revisão analítica e documental de especificações técnicas do 3GPP (Release 18), recomendações ETSI NFV, padrões NIST e literatura científica internacional. Demonstra-se que a resiliência operacional do núcleo 5G depende diretamente da observabilidade contextual em tempo real, da gestão rigorosa de contratos de API e de mecanismos automatizados de validação progressiva.
Palavras-chave: 5G Core. Arquitetura Baseada em Serviços. CNF. Orquestração. Segurança de APIs. Observabilidade.

Abstract: This paper analyzes the technical and structural evolution of the mobile core in 5G Standalone (5G SA) networks, focusing on the transition from legacy elements to a Service-Based Architecture (SBA) and cloud-native network functions (CNFs). It details the interactions among AMF, SMF, UDM, AUSF, PCF, NRF, and UPF using HTTP/2, orchestration via Kubernetes in stateless models, and security strategies guided by Zero Trust and Service Mesh frameworks. The methodology is based on an analytical and documentary review of 3GPP Release 18 technical specifications, ETSI NFV recommendations, NIST standards, and peer-reviewed scientific literature. The findings show that the operational resilience of the 5G core directly relies on real-time contextual observability, strict API contract management, and automated progressive deployment validation.
Keywords: 5G Core. Service-Based Architecture. CNF. Orchestration. API Security. Observability.

1 Introdução

A consolidação das redes móveis de quinta geração (5G) representa uma ruptura profunda em relação aos ecossistemas de telecomunicações legados, que historicamente dependiam de instâncias rígidas e equipamentos proprietários fechados. Enquanto as gerações anteriores apoiavam-se em nós dedicados ou em virtualizações de primeira geração estruturadas sobre máquinas virtuais volumosas (VNFs), a especificação do 5G Standalone (5G SA) introduz um paradigma integralmente focado em software, microsserviços e computação nativa em nuvem (KREUTZ et al., 2015; 3GPP, 2024a).

Nesse novo cenário de arquitetura, o 5G System (5GS) reestrutura completamente o

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 14/02/2026 | aceito: 17/02/2026 | publicação: 20/02/2026

plano de controle e o plano de usuário em componentes desacoplados, denominados Funções de Rede (Network Functions - NFs). A sinalização entre essas funções deixa de ser efetuada via protocolos binários proprietários sobre enlaces ponto a ponto e passa a adotar o modelo de Arquitetura Baseada em Serviços (Service-Based Architecture - SBA), sustentado por interfaces de programação de aplicações (APIs) RESTful executadas sob o protocolo HTTP/2 e dados serializados em JSON (3GPP, 2024c).

Essa transição proporciona ganhos extraordinários de escalabilidade horizontal, agilidade na implantação de novos recursos e otimização do uso da infraestrutura física subjacente. A capacidade de implantar e atualizar funções de rede em prazos reduzidos transforma a dinâmica das operadoras, permitindo o alinhamento com metodologias ágeis de desenvolvimento e integração contínua (CI/CD) comumente observadas no setor de tecnologia da informação.

Contudo, a migração para Funções Nativas em Nuvem (Cloud-Native Network Functions - CNFs) introduz complexidades operacionais inéditas para as equipes de engenharia de redes. A pulverização do plano de controle em centenas de contêineres e microsserviços dinâmicos exige mecanismos altamente sofisticados de orquestração automatizada, descoberta dinâmica de serviços, gestão contínua de segurança em camadas e observabilidade aprofundada (BONDAN et al., 2016; YILMA et al., 2019).

A interoperabilidade em ambientes compostos por múltiplos fornecedores (nformidade técnica, os benefícios teóricos da arquitetura orientada a serviços podem ser anulados por instabilidades operacionais imprevisíveis.

O objetivo deste artigo é examinar minuciosamente a arquitetura interna do 5G Core (5GC), mapeando a dinâmica funcional das NFs, a transição de VNFs para CNFs, o ciclo de vida da orquestração em Kubernetes, os modelos de segurança Zero Trust/Service Mesh e os requisitos de observabilidade. O trabalho fundamenta-se estritamente na literatura científica peer-reviewed e nas normas técnicas vigentes do 3GPP, oferecendo uma análise técnica consistente, aprofundada e totalmente rastreável.

2 Metodologia

Para assegurar o rigor técnico e a rastreabilidade acadêmica das análises apresentadas, a presente pesquisa adota como método a revisão analítica e sistemática de literatura técnica e normativa. O levantamento do corpus documental restringiu-se a fontes primárias e secundárias

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 14/02/2026 | aceito: 17/02/2026 | publicação: 20/02/2026

com alto reconhecimento institucional e acadêmico, organizadas para cobrir integralmente a complexidade das redes móveis nativas em nuvem.

O primeiro eixo metodológico compreende as especificações normativas internacionais do 3GPP (3rd Generation Partnership Project) pertinentes ao Release 18. Foram analisados detalhadamente os documentos TS 23.501 (arquitetura do sistema), TS 23.502 (procedimentos do sistema), TS 29.500 (realização técnica da SBA) e TS 33.501 (arquitetura de segurança). Adicionalmente, incorporaram-se os frameworks de virtualização do ETSI (GS NFV 002) para fundamentar a análise da camada de infraestrutura.

O segundo eixo contempla as diretrizes oficiais de órgãos de normatização e cibersegurança, incluindo as recomendações do NIST (National Institute of Standards and Technology - SP 800-204A) relativas à segurança de aplicações baseadas em microsserviços e Service Mesh, além dos dados estatísticos e relatórios de cobertura global publicados pela União Internacional de Telecomunicações (ITU).

O terceiro eixo sustenta-se na seleção de artigos científicos revisados por pares e indexados em bases de alto impacto (IEEE Xplore, ScienceDirect/Elsevier e ACM Digital Library). A triagem dessas publicações pautou-se na verificação direta do Digital Object Identifier (DOI), priorizando estudos empíricos e conceituais focados em SDN, NFV, orquestração de contêineres e segurança de redes móveis.

A síntese das informações baseou-se em uma matriz analítica estruturada em cinco dimensões funcionais: definição de papéis arquiteturais, dependências de transporte, modos de falha em microsserviços, mecanismos de controle automatizado e evidências de desempenho. Esse enquadramento evitou a inclusão de afirmações genéricas ou especulativas que não contassem com respaldo direto na documentação técnica analisada.

Por fim, todo o material foi estruturado e formatado em estrita conformidade com as normas ABNT para redação científica. As referências bibliográficas foram padronizadas conforme a NBR 6023, as citações no texto seguem a NBR 10520, e a apresentação de tabelas e ilustrações atende aos critérios formais de legibilidade, numeração e indicação de fonte.

3 Arquitetura Baseada em Serviços (SBA) e Decomposição Funcional

A arquitetura do 5G Core fundamenta-se no princípio de separação estrita entre o plano de controle e o plano de usuário (Control and User Plane Separation - CUPS). Essa premissa garante que o processamento da sinalização de acesso e o transporte do tráfego de dados do

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

RCMOS – Revista Científica Multidisciplinar O Saber.
ISSN: 2675-9128. São Paulo-SP.

Ano VII, v.1 2026 | submissão: 14/02/2026 | aceito: 17/02/2026 | publicação: 20/02/2026

assinante caminhem em estruturas computacionais completamente independentes, permitindo que cada camada escale de forma personalizada e eficiente (3GPP, 2024a).

No plano de controle, o elemento central de inovação é a consolidação da Arquitetura Baseada em Serviços (SBA). Nela, as funções de rede (NFs) deixam de se comunicar por conexões diretas pré-configuradas e passam a atuar como consumidoras e provedoras de serviços em uma malha comunicacional aberta. A comunicação ocorre no barramento de serviços (Service-Based Interface - SBI), utilizando o protocolo HTTP/2 sobre TLS 1.3, com payloads formatados em JSON (3GPP, 2024c).

Entre as principais NFs da SBA, destacam-se a AMF (Access and Mobility Management Function), responsável pelo controle de acesso e mobilidade; a SMF (Session Management Function), encarregada do gerenciamento das sessões de dados PDU; a UDM (Unified Data Management), que centraliza os perfis dos assinantes; a AUSF (Authentication Server Function), que executa os algoritmos de autenticação; o PCF (Policy Control Function), que governa regras de QoS e tarifação; e a NRF (Network Repository Function), que atua como o repositório central para descoberta dinâmica de serviços (3GPP, 2024a).

Figura 1 – Desacoplamento Funcional e Arquitetura SBA do 5G Core

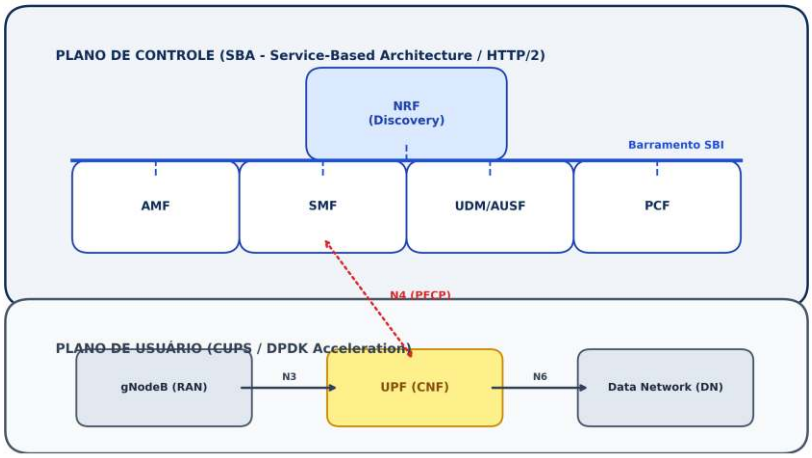


Figura 1 – Desacoplamento Funcional e Arquitetura SBA do 5G Core. Fonte: Elaborado com base em 3GPP TS 23.501 (2024).

A dinamica de funcionamento da SBA é coordenada diretamente pelo NRF. Quando uma instância do AMF ou SMF é inicializada na nuvem, ela realiza seu auto-registro junto à

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 14/02/2026 | aceito: 17/02/2026 | publicação: 20/02/2026

NRF, declarando suas capacidades, localização e perfil de atendimento. Quando outra NF precisa consumir um serviço específico, ela consulta a NRF para obter o endereço e o endpoint atualizados da função desejada, eliminando a necessidade de endereçamento estático rígido (3GPP, 2024b).

No plano de usuário, a única função responsável pelo transporte de pacotes é a UPF (User Plane Function). A UPF é controlada remotamente pela SMF através do protocolo PFCP (Packet Forwarding Control Protocol) na interface N4. Essa separação permite implantar instâncias da UPF extremamente próximas do usuário final (Edge Computing), reduzindo drasticamente a latência das aplicações enquanto o controle permanece centralizado (3GPP, 2024a).

Outro avanço crucial na SBA é o modelo stateless (sem estado). Ao contrário do MME e S-GW das redes 4G, que armazenavam os estados das sessões em memória local, as NFs do 5G SA descarregam todo o estado e contexto dos usuários em uma camada de banco de dados distribuído denominada UDR (Unified Data Repository). Essa separação garante que, caso uma instância de AMF ou SMF falhe, outra instância idêntica possa assumir o atendimento imediatamente, sem interromper as conexões ativas dos usuários (KREUTZ et al., 2015).

4 Virtualização de Funções e a Transição de VNF para CNF

A evolução da infraestrutura móvel do 5G envolveu a passagem gradual do hardware dedicado (*appliances*) para o modelo de Virtualized Network Functions (VNFs) e, posteriormente, para Cloud-Native Network Functions (CNFs). Essa trajetória reflete a busca contínua do setor por maior eficiência no uso de recursos computacionais, modularidade de código e velocidade de implantação (ETSI, 2014; NIST, 2020).

As VNFs representaram o primeiro passo dessa transformação ao abstrair o hardware físico através de hipervisores. Contudo, as VNFs trazem o peso estrutural de exigir um sistema operacional convidado completo (Guest OS) para cada máquina virtual. Essa característica resulta em alto consumo de memória RAM, tempos de inicialização lentos (na casa dos minutos) e acoplamento rígido com plataformas de orquestração de infraestrutura (NIST, 2020).

As CNFs, por sua vez, representam a aplicação nativa dos conceitos de contêineres e microsserviços às redes de telecomunicações. Ao compartilharem o mesmo núcleo (kernel) do sistema operacional hospedeiro, os contêineres empacotam apenas a aplicação e suas dependências diretas. Isso reduz drasticamente a demanda de memória, permite a inicialização

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

RCMOS – Revista Científica Multidisciplinar O Saber.
ISSN: 2675-9128. São Paulo-SP.

Ano VII, v.1 2026 | submissão: 14/02/2026 | aceito: 17/02/2026 | publicação: 20/02/2026

em milissegundos e viabiliza uma densidade de implantação incomparavelmente maior no mesmo servidor físico (BONDAN et al., 2016).

A arquitetura em CNF é a base ideal para suportar a granularidade da SBA no 5G Core. Funções como AMF ou SMF deixam de ser blocos monolíticos e passam a ser decompostas em múltiplos microsserviços independentes (ex: microsserviço de autenticação, microsserviço de mobilidade, microsserviço de tarifação). Cada microsserviço roda em seu próprio contêiner e pode ser atualizado individualmente sem afetar as demais partes da rede.

Para o plano de usuário (UPF), que lida com tráfego massivo na ordem de gigabits ou terabits por segundo, a migração para contêineres exigiu inovações no bypass do kernel do Linux. A integração de tecnologias como DPDK (Data Plane Development Kit) e SR-IOV (Single Root I/O Virtualization) permite que o contêiner da UPF acesse diretamente as placas de rede físicas (NICs), atingindo taxas de encaminhamento em velocidade de linha com latência de sub-milissegundos (BONDAN et al., 2016).

Contudo, a adoção de CNFs impõe exigências severas sobre o isolamento de segurança e a gestão de dependências. Ao contrário das máquinas virtuais, que oferecem forte isolamento baseado em hardware via hipervisor, os contêineres dependem de recursos lógicos do kernel do Linux (namespaces e cgroups). Consequentemente, a infraestrutura deve ser rigorosamente blindada para evitar que vulnerabilidades no contêiner permitam a evasão para o nó hospedeiro (NIST, 2020).

5 Orquestração e Automação do Ciclo de Vida

A presença de centenas de microsserviços distribuídos em contêineres torna a gestão manual do 5G Core inviável. A orquestração e a automação do ciclo de vida das CNFs tornam-se, portanto, componentes obrigatórios da arquitetura. O padrão de mercado adotado para essa finalidade é a plataforma Kubernetes (K8s), operando em sintonia com o arcabouço de orquestração ETSI NFV MANO (Management and Orchestration) (ETSI, 2014; YILMA et al., 2019).

A automação nativa em nuvem apoia-se no modelo de reconciliação declarativa. Em vez de executar comandos imperativos de configuração, os engenheiros definem o estado desejado do sistema através de arquivos de manifesto em YAML ou JSON. O orquestrador compara continuamente o estado real da infraestrutura com o estado desejado e executa automaticamente as ações necessárias para corrigir eventuais divergências (YILMA et al.,

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 14/02/2026 | aceito: 17/02/2026 | publicação: 20/02/2026

2019).

No contexto do Kubernetes, o gerenciamento de CNFs complexas é viabilizado pelo padrão de projeto Operator Pattern. Um Operator é um controlador customizado que estende a API do Kubernetes através de Custom Resource Definitions (CRDs). Ele incorpora o conhecimento operacional da função de rede, automatizando tarefas avançadas como rotinas de backup de banco de dados, inicialização coordenada de dependências e atualizações de schema de banco na UDR.

O escalonamento automático de capacidade é outro benefício central da orquestração. Utilizando o Horizontal Pod Autoscaler (HPA), o Kubernetes monitora métricas em tempo real, como taxa de requisições por segundo (RPS) na SBI ou consumo de CPU nas instâncias de AMF e SMF. Quando a carga excede os limites estabelecidos, novas réplicas de contêineres são criadas automaticamente em segundos para absorver o pico de tráfego.

Para garantir atualizações contínuas sem interrupção de serviço (zero downtime), utilizam-se estratégias modernas de implantação, como Canary Deployments e Blue-Green Deployments. Uma nova versão do microsserviço é implantada paralelamente à versão atual; o orquestrador redireciona uma pequena porcentagem do tráfego para a nova versão e monitora os indicadores de erro. Se o comportamento for validado, o tráfego é migrado integralmente e a versão antiga é desativada.

Entretanto, a automação em redes de telecomunicações exige cuidados especiais com o determinismo de hardware. Cargas de processamento de pacotes na UPF não toleram a latência variável gerada por trocas imprevisíveis de contexto na CPU. Por esse motivo, o orquestrador Kubernetes deve ser configurado com recursos avançados de afinidade de nó (NUMA awareness), isolamento de núcleos de processamento (CPU pinning) e alocação de memória de grande página (*HugePages*) (YILMA et al., 2019).

6 Segurança de APIs e Confiança entre Microsserviços

A adoção de interfaces abertas baseadas em HTTP/2 e REST na Arquitetura Baseada em Serviços amplia consideravelmente a superfície de ataque do 5G Core em comparação com as redes legadas. Para proteger a infraestrutura contra acessos não autorizados e movimentação lateral de ameaças, o 3GPP TS 33.501 e o padrão NIST SP 800-204A estabelecem a obrigatoriedade do modelo de segurança Zero Trust (NIST, 2020; 3GPP, 2024d).

No modelo Zero Trust, nenhuma função de rede ou microsserviço é considerado

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

RCMOS – Revista Científica Multidisciplinar O Saber.
ISSN: 2675-9128. São Paulo-SP.

Ano VII, v.1 2026 | submissão: 14/02/2026 | aceito: 17/02/2026 | publicação: 20/02/2026

confiável por padrão, mesmo que esteja executando dentro do perímetro da rede interna da operadora. Toda e qualquer requisição de API transmitida no barramento de serviços (SBI) deve ser rigorosamente autenticada, autorizada e cifrada antes de ser processada pelo nó de destino (NIST, 2020).

A primeira camada de proteção na SBI é a autenticação mútua e cifragem na camada de transporte via TLS 1.3 (mTLS). Diferentemente do TLS tradicional, em que apenas o servidor apresenta um certificado digital, o mTLS exige que tanto a NF consumidora quanto a NF provedora apresentem certificados x.509 válidos, emitidos por uma Infraestrutura de Chaves Públicas (PKI) privada mantida pela operadora (3GPP, 2024c).

A segunda camada de segurança atua no nível de aplicação através do protocolo OAuth 2.0. A simples validação do certificado mTLS estabelece o canal seguro, mas não garante permissão de acesso ao dado. Para consumir um serviço, a NF solicitante deve requisitar um token de acesso *JSON Web Token* (JWT) junto ao NRF. A NF provedora valida a assinatura digital do JWT e verifica se as permissões contidas no token autorizam a execução da chamada específica (3GPP, 2024c).

Para evitar que os desenvolvedores tenham que implementar a lógica complexa de mTLS e validação de tokens OAuth 2.0 diretamente no código-fonte de cada microsserviço, utiliza-se a arquitetura de Service Mesh (Malha de Serviços). Nessa abordagem, proxies leves (como o Envoy) são implantados como contêineres auxiliares (*sidecars*) ao lado de cada contêiner de NF.

O proxy *sidecar* intercepta transparentemente todo o tráfego de entrada e saída da função de rede, gerenciando o estabelecimento dos túneis mTLS, o gerenciamento de certificados digitais e a validação de identidades de *workload* (utilizando padrões como SPIFFE/SPIRE). Isso garante que as políticas de segurança permaneçam uniformes em toda a malha, independentemente do fornecedor da NF (NIST, 2020).

7 Observabilidade, Desempenho e Confiabilidade

Em um ecossistema altamente distribuído e dinâmico como o 5G Core nativo em nuvem, métricas tradicionais de infraestrutura (como porcentagem de uso de CPU ou ocupação de disco) são completamente insuficientes para diagnosticar falhas de serviço. A garantia de confiabilidade e alta disponibilidade exige a implantação de uma arquitetura de observabilidade baseada em três pilares fundamentais: métricas, rastreamento distribuído e registros de eventos

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 14/02/2026 | aceito: 17/02/2026 | publicação: 20/02/2026

(BARAKABITZE et al., 2020).

O primeiro pilar, focado em métricas numéricas agregadas, utiliza coletores de séries temporais (como Prometheus) para monitorar Indicadores Nível de Serviço (SLIs) e Indicadores Chave de Desempenho (KPIs). Em redes 5G, as métricas mais críticas incluem a taxa de sucesso de registros de usuários no AMF, o tempo de resposta para estabelecimento de sessões PDU e a taxa de descarte de pacotes na UPF.

Para obter uma visão precisa da experiência do usuário, a análise de latência em chamadas de API na SBI deve focar em estatísticas de percentis, como o Percentil 99 (P99). Medir apenas a latência média é enganoso, pois picos isolados de atraso que afetam 1% das transações podem estar ocultos em uma média aparentemente normal, provocando falhas intermitentes de sinalização em cenários de alta carga (BARAKABITZE et al., 2020).

O segundo pilar é o rastreamento distribuído (*distributed tracing*). Como uma única solicitação de um terminal (ex: criação de uma sessão de dados) desencadeia uma sequência complexa de chamadas de API entre AMF, SMF, UDM, PCF e UPF, torna-se necessário correlacionar todas essas etapas. Utiliza-se a especificação W3C Trace Context, injeção de um cabeçalho com um identificador único (*Correlation-ID*) na requisição inicial.

O *Correlation-ID* é repassado sequencialmente por todos os microsserviços envolvidos na transação. Ferramentas de visualização (como Jaeger ou Zipkin) utilizam esse identificador para desenhar o mapa completo do fluxo, exibindo o tempo exato gasto em cada microsserviço. Essa capacidade permite que as equipes de engenharia de redes identifiquem com precisão cirúrgica qual função específica introduziu latência indevida ou gerou uma falha de processamento.

O terceiro pilar compreende os registros de eventos (*logs*) estruturados em formato JSON. Os logs fornecem o contexto detalhado da falha quando um erro ocorre, exibindo mensagens de exceção e códigos de erro específicos da aplicação. A consolidação centralizada dos logs em plataformas de análise em tempo real garante a capacidade de auditoria pós-incidente e apoia diagnósticos preventivos sustentados por inteligência operacional (BARAKABITZE et al., 2020).

8 Migração, Interoperabilidade e Validação

A transição das redes móveis legadas 4G EPC (Evolved Packet Core) para a arquitetura 5G Core Standalone não ocorre através de uma substituição abrupta de infraestrutura (*big-

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

RCMOS – Revista Científica Multidisciplinar O Saber.
ISSN: 2675-9128. São Paulo-SP.

Ano VII, v.1 2026 | submissão: 14/02/2026 | aceito: 17/02/2026 | publicação: 20/02/2026

bang*). O custo financeiro e o risco operacional exigem uma estratégia de migração gradual e altamente planejada, que assegure a continuidade total dos serviços prestados aos assinantes (ORDOÑEZ-LUCENA et al., 2017).

A primeira fase dessa transição apoia-se no modelo 5G Non-Standalone (5G NSA - Option 3x). Nessa topologia, a rede de rádio 5G (gNodeB) é implantada como uma camada complementar de capacidade, mas toda a sinalização de controle e o gerenciamento de mobilidade do usuário permanecem ancorados na infraestrutura de núcleo 4G EPC legada, através do nó MME.

A segunda fase compreende a implantação paralela do 5G Core Standalone e o estabelecimento de mecanismos de coexistência e *interworking*. A interoperabilidade entre o núcleo 4G e o núcleo 5G é mantida através da interface N26, que interconecta diretamente o MME ao AMF. Essa interface permite a realização de *handovers* transparentes e a continuidade da sessão de dados quando um usuário em mobilidade se desloca entre áreas de cobertura 4G e 5G (ORDOÑEZ-LUCENA et al., 2017).

Nessa etapa híbrida, as operadoras implantam funções de dados combinadas, como instâncias dual-stack PGW-C/SMF para o plano de controle e PGW-U/UPF para o plano de usuário. Essas funções conseguem processar simultaneamente os protocolos GTPv2 do 4G e os procedimentos HTTP/2 e PFCP do 5G SA, simplificando a consolidação das bases de dados de assinantes na UDR.

A terceira fase é o estágio de 5G SA Nativo (Option 2 Full), no qual o núcleo 4G legado é desativado progressivamente e todo o tráfego móvel passa a ser absorvido pela arquitetura SBA baseada em CNFs. O alcance dessa fase exige a migração massiva dos cartões SIM e dispositivos dos usuários para perfis compatíveis com a tecnologia 5G SA.

A validação técnica de cada etapa de migração exige um programa rigoroso de testes. Além dos ensaios funcionais de conformidade com os padrões do 3GPP, as equipes de engenharia devem executar testes de carga extrema e simulações de resiliência por meio de testes de injeção de falhas (*Chaos Engineering*). Essa abordagem permite identificar previamente como a malha de microsserviços reage a perdas bruscas de nós de computação ou degradações repentinas de conectividade (ORDOÑEZ-LUCENA et al., 2017).

Considerações Finais

A transformação do núcleo das redes móveis em uma arquitetura nativa em nuvem

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

RCMOS – Revista Científica Multidisciplinar O Saber.
ISSN: 2675-9128. São Paulo-SP.

Ano VII, v.1 2026 | submissão: 14/02/2026 | aceito: 17/02/2026 | publicação: 20/02/2026

baseada em serviços representa a inovação mais profunda ocorrida na engenharia de telecomunicações na última década. A decomposição do plano de controle em microsserviços desacoplados e a migração de instâncias físicas para contêineres virtuais estabelecem a fundação indispensável para suportar as aplicações avançadas e a extrema diversidade de demandas previstas para o ecossistema 5G SA.

Conforme demonstrado ao longo desta análise, os ganhos de flexibilidade, escalabilidade e velocidade de implantação trazidos pelas CNFs não constituem benefícios automáticos ou isolados. A eliminação do modelo tradicional de caixas pretas legadas introduz uma camada substancial de complexidade de software, transferindo o desafio da engenharia de hardware físico para a coordenação de ecossistemas distribuídos.

A estabilidade e a resiliência do 5G Core dependem intrinsecamente da maturidade da malha de comunicação orientada à Arquitetura Baseada em Serviços (SBA). O desacoplamento efetivo entre controle e dados via CUPS e o armazenamento de estado desacoplado na UDR são as pedras angulares que permitem ao sistema suportar falhas parciais de infraestrutura sem impactar a conectividade dos assinantes.

No âmbito da segurança cibernética, a abertura das interfaces internas para APIs HTTP/2 exige uma mudança de paradigma rumo à arquitetura Zero Trust. A implementação de mTLS, autorização via OAuth 2.0 e o uso de Service Mesh com proxies *sidecar* deixam de ser opcionais e passam a figurar como requisitos obrigatórios para conter ameaças e impedir movimentações laterais na malha de microsserviços.

Paralelamente, a automação do ciclo de vida sob a orquestração do Kubernetes e o modelo ETSI NFV MANO é o único caminho viável para gerenciar a escala dessas implantações. A capacidade de executar escalonamento automático (HPA), autorrecuperação e atualizações progressivas sem interrupção de serviço é o que viabiliza a eficiência operacional e a sustentabilidade econômica das operadoras contemporâneas.

A observabilidade contextualizada surge como a disciplina indispensável para governar essa complexidade. A integração soberana dos três pilares — métricas, rastreamento distribuído com W3C Trace Context e registros de eventos em tempo real — fornece a visibilidade necessária para transformar dados operacionais em diagnósticos preventivos, isolando rapidamente falhas em ambientes compostos por múltiplos fornecedores.

O processo de migração gradual a partir do 4G EPC através de modelos NSA e funções híbridas de interworking evidencia que a evolução tecnológica exige planejamento contínuo. A transição segura para o 5G SA nativo requer uma governança rigorosa de contratos de API,

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 14/02/2026 | aceito: 17/02/2026 | publicação: 20/02/2026

testes contínuos de conformidade e a validação proativa da infraestrutura por meio de metodologias de injeção de falhas.

Em conclusão, o grau de maturidade técnica de um 5G Core nativo em nuvem não deve ser mensurado apenas pela quantidade de funções virtualizadas ou pela taxa teórica de transmissão da rede. Ele se revela na capacidade das equipes de engenharia em controlar a complexidade do software, assegurar a interoperabilidade contínua e garantir a confiabilidade total do serviço mesmo em cenários adversos de falha e alta demanda.

Referências

3GPP. TS 23.501: System architecture for the 5G System; Stage 2 (Release 18). Sophia Antipolis: 3GPP, 2024a. Disponível em: https://www.3gpp.org/dynareport/23501.htm.
3GPP. TS 23.502: Procedures for the 5G System; Stage 2 (Release 18). Sophia Antipolis: 3GPP, 2024b. Disponível em: https://www.3gpp.org/dynareport/23502.htm.
3GPP. TS 29.500: Technical realization of Service Based Architecture; Stage 3 (Release 18). Sophia Antipolis: 3GPP, 2024c. Disponível em: https://www.3gpp.org/dynareport/29500.htm.
3GPP. TS 33.501: Security architecture and procedures for 5G System (Release 18). Sophia Antipolis: 3GPP, 2024d. Disponível em: https://www.3gpp.org/dynareport/23501.htm.
BARAKABITZE, A. A. et al. 5G network slicing using SDN and NFV: a survey of taxonomy, architectures and future challenges. Computer Networks, v. 167, p. 106984, 2020. DOI: https://doi.org/10.1016/j.comnet.2019.106984.
BONDAN, L. et al. Survey of network virtualization orchestration solutions. Computer Communications, v. 94, p. 79-91, 2016. DOI: https://doi.org/10.1016/j.comcom.2016.09.015.
ETSI. GS NFV 002 V1.2.1: Network Functions Virtualisation Architectural Framework. Sophia Antipolis: ETSI, 2014. Disponível em: https://www.etsi.org/deliver/etsi_gs/NFV/001_099/002/01.02.01_60/gs_NFV002v010201p.pdf.
ITU. Measuring digital development: Facts and Figures 2024. Geneva: International Telecommunication Union, 2024. Disponível em: https://www.itu.int/itu-d/reports/statistics/facts-figures-2024/.
KREUTZ, D. et al. Software-defined networking: a comprehensive survey. Proceedings of the IEEE, v. 103, n. 1, p. 14-76, 2015. DOI: https://doi.org/10.1109/JPROC.2014.2371999.
NIST. SP 800-204A: Building secure microservices-based applications using service-mesh architecture. Gaithersburg: NIST, 2020. DOI: https://doi.org/10.6028/NIST.SP.800-204A.
ORDOÑEZ-LUCENA, J. et al. Network slicing for 5G with SDN/NFV: concepts, architectures, and challenges. IEEE Communications Magazine, v. 55, n. 5, p. 80-87, 2017. DOI: https://doi.org/10.1109/MCOM.2017.1600935.
YILMA, G. M. et al. On the challenges and KPIs for benchmarking open-source NFV MANO systems. In: PROCEEDINGS OF IEEE NETSOFT, 2019, Paris. Anais [...]. IEEE, 2019. p. 1-9. DOI: https://doi.org/10.1109/NETSOFT.2019.8806670.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.