

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

EVERTON SOUZA WIENCE

**AVALIAÇÃO DE DESEMPENHO DO SERVIDOR NGINX COMO PROXY
REVERSO COM CACHE PARA APLICAÇÕES WEB**

GUARAPUAVA

2025

EVERTON SOUZA WIENCE

**AVALIAÇÃO DE DESEMPENHO DO SERVIDOR NGINX COMO PROXY
REVERSO COM CACHE PARA APLICAÇÕES WEB**

**Performance evaluation of NGINX server as reserve proxy for web
applications**

Proposta de Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Tecnólogo em Sistemas para Internet do Curso de Tecnologia em Sistemas para Internet da Universidade Tecnológica Federal do Paraná.

Orientador: Prof^a. Dr^a. Sediane Carmen Lunardi
Hernandes

Coorientador: Prof. Dr. Hermano Pereira

GUARAPUAVA

2025



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

LISTA DE FIGURAS

Figura 1 – Ilustração de um servidor web (FOROUZAN; MOSHARRAF, 2013). . . .	6
Figura 2 – Comparação entre <i>proxy</i> direto e reverso.	7

SUMÁRIO

1	INTRODUÇÃO	3
1.1	Objetivos	3
1.1.1	Objetivo geral	3
1.1.2	Objetivos específicos	4
1.2	Justificativa	4
1.3	Estrutura do trabalho	5
2	REFERENCIAL TEÓRICO	6
2.1	Servidores web	6
2.1.1	O Servidor NGINX	6
2.2	<i>Proxy e proxy reverso</i>	7
2.3	Mecanismos de <i>cache</i> em aplicações web	7
3	PROPOSTA	9
4	CONSIDERAÇÕES FINAIS	10
	REFERÊNCIAS	11

1 INTRODUÇÃO

As aplicações web têm se tornado parte essencial da vida contemporânea, sendo utilizadas em diversos setores, como comunicação, entretenimento, comércio eletrônico e sistemas corporativos. Com a crescente digitalização da sociedade, tais aplicações passaram a desempenhar um papel crítico não apenas no cotidiano dos usuários, mas também no funcionamento de empresas e instituições. Nesse cenário, aspectos como desempenho, escalabilidade e disponibilidade são fundamentais para garantir experiências satisfatórias e confiáveis por parte dos usuários.

Entre as técnicas disponíveis, destaca-se a utilização de *proxy* reverso aliado ao uso de mecanismos de *cache* em servidores web. O *proxy* reverso atua como intermediário entre os clientes e um ou mais servidores de aplicação, trazendo benefícios como balanceamento de carga, proteção contra acessos indevidos e armazenamento temporário de conteúdos em *cache* (CLOUDFLARE, 2024b). O *cache*, por sua vez, evita o processamento repetido de requisições idênticas, reduzindo a latência percebida pelo usuário e otimizando o uso de *Central Processing Unit* (CPU) e memória do servidor (AIVALIOTIS, 2016).

Nesse contexto, o servidor NGINX tem ganhado destaque devido à sua ampla adoção em ambientes corporativos e acadêmicos, além de sua eficiência no tratamento de requisições simultâneas. Reconhecido por sua capacidade de atuar tanto como servidor web quanto como *proxy* reverso, o NGINX oferece recursos avançados de caching que permitem melhorar o desempenho de aplicações web em diferentes cenários.

Diante disso, este trabalho busca investigar o impacto da utilização de *proxy* reverso com cache no servidor NGINX em aplicações web estáticas e dinâmicas, avaliando seus efeitos em métricas como tempo de resposta, consumo de memória e uso de CPU. O estudo tem como objetivo fornecer subsídios práticos e científicos que auxiliem acadêmicos, profissionais de Tecnologia da Informação e pequenas empresas na adoção de soluções eficientes para otimização de desempenho em serviços web.

1.1 Objetivos

1.1.1 Objetivo geral

O objetivo geral do trabalho consiste em realizar um estudo dos impactos quanto ao uso do servidor NGINX como *proxy* reverso com cache em relação ao desempenho de aplicações web estáticas e dinâmicas.

1.1.2 Objetivos específicos

- Estudar e analisar como funciona o *proxy* reverso e o mecanismo de cache do servidor NGINX, bem como a relação entre *proxy* reverso e *caching*.
- Realizar levantamento de aplicações/serviços web de diferentes portes para compor os cenários de testes experimentais.
- Sintetizar os dados coletados em testes de carga realizados em ambientes com e sem *proxy* reverso com cache para medir parâmetros como latência, tempo de resposta e uso de recursos (p.e., memória e processador).
- Avaliar os resultados coletados propondo boas práticas de *proxy* reverso para os diferentes tipos de aplicações.

1.2 Justificativa

Com a crescente dependência de serviços web em diversos setores, o desempenho dos servidores tornou-se um fator crítico para a experiência do usuário. Páginas lentas não apenas impactam a satisfação e a retenção de usuários, mas também podem gerar prejuízos financeiros significativos. O tempo de resposta elevado, especialmente em requisições de conteúdo estático (como imagens, arquivos CSS e JavaScript) é um problema. Sem um *proxy* reverso, o servidor de origem precisa processar cada solicitação, gerando latência e sobrecarga. Em contrapartida, o uso do *cache* no NGINX resolve esse problema. Ele armazena uma cópia do conteúdo na primeira requisição e, a partir da segunda, entrega a resposta diretamente do *cache*. Esse mecanismo tem como objetivo reduzir o tempo de espera do cliente, oferecendo uma experiência de navegação mais rápida e fluida.

Além disso, servidores web sem *cache* precisam de mais recursos de CPU e memória para lidar com o tráfego, o que pode levar a um gargalo de desempenho. A arquitetura orientada a eventos do NGINX, combinada com o *cache*, libera o servidor de origem de grande parte das requisições repetidas. Assim, com o NGINX servindo as respostas em *cache*, o servidor original pode focar apenas em processar novas requisições. Com isso, a carga computacional do servidor diminui e tem-se uma otimização dos recursos existentes.

Nesse contexto, a investigação do uso do NGINX como *proxy* reverso com *cache* justifica-se por seu potencial em solucionar problemas práticos de desempenho e escalabilidade. O NGINX, em sua configuração com *cache*, atua como uma camada intermediária que otimiza o fluxo de tráfego, servindo como uma solução eficiente e de baixo custo para ambientes de alta demanda.

Além disso, este estudo contribui para a literatura e para a prática profissional. Ao oferecer uma análise aplicada e a demonstração dos benefícios do NGINX, o trabalho se posiciona

como um estudo de caso valioso. A relevância no mercado é evidente, uma vez que a técnica é amplamente adotada por grandes empresas de tecnologia.

1.3 Estrutura do trabalho

Este trabalho está organizado da seguinte forma. O Capítulo 1 apresenta a introdução, objetivos e justificativa. O Capítulo 2 aborda o referencial teórico sobre servidores web, NGINX, *proxy* reverso e *cache*. A proposta deste Trabalho de Conclusão de Curso é descrita no Capítulo 3 e o Capítulo 4 traz as considerações finais.

2 REFERENCIAL TEÓRICO

No referencial teórico, serão discutidos os principais conceitos que embasam este estudo, começando com uma introdução sobre **servidores web**. Em seguida, serão exploradas as definições e funções de **proxy reverso** e **mecanismos de caching**, destacando a importância dessas tecnologias para a otimização de recursos. Por fim, o capítulo abordará as características do servidor **NGINX**, ressaltando sua relevância para o contexto da pesquisa.

2.1 Servidores web

Os servidores web são *softwares* responsáveis por receber e processar requisições do *Hyper Transfer Protocol* (HTTP) ou do *Hypertext Transfer Protocol Secure* (HTTPS) realizadas por clientes (navegadores ou outras aplicações) e retornar os recursos solicitados, como páginas *Hypertext Markup Language* (HTML), imagens, arquivos ou dados estruturados (AIVALIOTIS, 2016). Tradicionalmente, os servidores web desempenham papel central na arquitetura cliente-servidor da web, atuando como ponto de contato entre o usuário e os sistemas hospedados (AIVALIOTIS, 2016). A Figura 1 ilustra a ideia de um servidor servindo a dois usuários na Internet.

Dentre as características mais relevantes de um servidor web, destacam-se: capacidade de lidar com múltiplas conexões simultâneas, suporte a protocolos modernos (como HTTP/2 e HTTP/3), escalabilidade, segurança e compatibilidade com diferentes tecnologias de aplicação. O desempenho do servidor web é um fator determinante para a experiência do usuário final e para a eficiência no consumo de recursos computacionais (BUTLER, 2017).

Alguns exemplos de servidores web são o *Apache HTTP Server*, o *Lighttpd* e o *NGINX*.

2.1.1 O Servidor NGINX

O NGINX é um servidor web de alto desempenho que se consolidou como uma das principais ferramentas utilizadas mundialmente em razão de sua capacidade de lidar com grande volume de conexões simultâneas com baixo consumo de recursos (AIVALIOTIS, 2016). Sua ar-

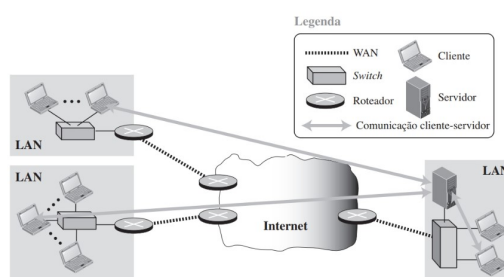


Figura 1 – Ilustração de um servidor web (FOROUZAN; MOSHARRAF, 2013).

quitetura baseada em eventos assíncronos permite escalabilidade superior quando comparada a modelos tradicionais baseados em *threads*.

Além de atuar como servidor web, o NGINX também é amplamente empregado como *proxy* reverso e balanceador de carga. O *NGINX Cookbook* (BUTLER, 2017) destaca casos práticos de configuração que demonstram ganhos de desempenho significativos em aplicações que lidam com tráfego intenso e requisições repetidas.

2.2 Proxy e proxy reverso

Um *proxy* é um *software* intermediário entre clientes e servidores. No contexto de redes e aplicações web, ele pode desempenhar funções como filtragem de tráfego, anonimização, restrição de acesso ou *caching* de conteúdos. Já o *proxy* reverso atua na direção oposta: posicionado entre os clientes e os servidores de aplicação, intercepta requisições destinadas aos servidores e as distribui conforme necessário (CLOUDFLARE, 2024b).

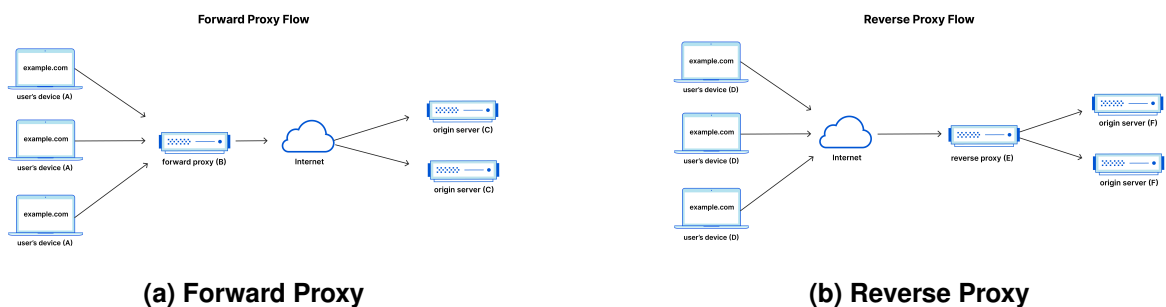


Figura 2 – Comparação entre *proxy* direto e reverso.

Entre os principais benefícios do uso de *proxy* reverso, destacam-se (CLOUDFLARE, 2024b):

- Balanceamento de carga, permitindo que requisições sejam distribuídas entre múltiplos servidores.
- Segurança, protegendo os servidores de aplicação contra ataques diretos.
- *Caching*, armazenando temporariamente conteúdos e reduzindo a necessidade de processamento repetitivo.

2.3 Mecanismos de *cache* em aplicações web

Um *cache* é um local de armazenamento temporário para cópias de arquivos ou outros tipos de dados que busca oferecer às aplicações acesso mais rápido (CLOUDFLARE, 2024a). Em aplicações web, o mecanismo de *cache* também é utilizado.

Desta forma, existem diferentes níveis de cache para aplicações web:

1. *Cache* no navegador, armazenando recursos no dispositivo do cliente.
2. *Cache* em proxies, interceptando tráfego e guardando conteúdos frequentemente requisitados.
3. *Cache* no servidor de aplicação, otimizando respostas dinâmicas.

No contexto do NGINX, o cache integrado ao *proxy* reverso permite que respostas a requisições repetidas sejam entregues rapidamente sem necessidade de acessar o servidor de aplicação, diminuindo o uso de *Central Processing Unit* (CPU) e memória.

3 PROPOSTA

Este trabalho se propõe a estudar e analisar os impactos do servidor NGINX como *proxy* reverso com *cache* quanto ao desempenho para aplicações web. Logo, o trabalho consistirá na instalação, configuração e análise de ambientes web utilizando o servidor NGINX com e sem *proxy* reverso com *cache*. Serão realizadas comparações de desempenho em aplicações estáticas (conteúdo fixo) e dinâmicas (com processamento de dados), de modo a identificar ganhos, limitações e melhores práticas.

A metodologia envolverá a criação de cenários de teste, a execução de testes de carga utilizando ferramentas específicas (como *Apache Benchmark*, *wrk* ou *k6*) e a coleta de métricas de desempenho. Os resultados obtidos serão analisados de forma crítica e comparativa. Desta forma, acadêmicos, profissionais de tecnologia e pequenas empresas podem utilizar este estudo como base para otimizar o desempenho de aplicações web sem recorrer a soluções proprietárias de alto custo.

Espera-se demonstrar, por meio de experimentação, os benefícios da utilização de *proxy* reverso com *cache* no NGINX, fornecendo evidências quantitativas de redução de latência e melhor aproveitamento de recursos. Ao final, pretende-se propor um conjunto de recomendações práticas que possam ser aplicadas em diferentes contextos.

4 CONSIDERAÇÕES FINAIS

Este projeto de Trabalho de Conclusão de Curso tem como foco a investigação dos efeitos da implementação de *proxy* reverso com *cache* utilizando o servidor NGINX em aplicações web, tanto estáticas quanto dinâmicas. A análise será conduzida por meio de métricas de desempenho como tempo de resposta, uso de memória e carga de CPU, com o intuito de oferecer fundamentos práticos e científicos que auxiliem na escolha de soluções eficazes para a otimização de serviços online.

A importância desta pesquisa aplicada está diretamente ligada à crescente demanda por sistemas web mais ágeis, escaláveis e confiáveis, especialmente em um contexto onde organizações dependem fortemente de plataformas digitais. Com isso, espera-se que os resultados beneficiem tanto o meio acadêmico — ao consolidar práticas recomendadas — quanto profissionais da área de tecnologia, ao fornecer orientações técnicas aplicáveis no cotidiano.

Entre os principais resultados almejados, destaca-se a comprovação de que o uso de *proxy* reverso com *cache* no NGINX pode diminuir significativamente a latência e melhorar o aproveitamento dos recursos computacionais, promovendo ambientes mais eficientes. O impacto dessa melhoria pode se refletir em uma experiência de usuário mais fluida e em maior competitividade para empresas que adotam tais estratégias.

Ainda assim, o projeto reconhece a existência de desafios, como a criação de cenários de teste que representem fielmente o uso real, a escolha criteriosa das métricas de avaliação e a consideração das variações de desempenho conforme o tipo de aplicação. Esses aspectos exigirão atenção especial ao longo da pesquisa para assegurar a precisão e a confiabilidade dos resultados obtidos.

REFERÊNCIAS

AIVALIOTIS, D. **Mastering NGINX - Second Edition**. Birmingham: Packt Publishing, 2016.

BUTLER, T. **NGINX Cookbook: Over 70 recipes for real-world configuration, deployment, and administration**. Birmingham: Packt Publishing, 2017.

CLOUDFLARE. **O que é caching?** 2024. Acesso em: 10 set. 2025. Disponível em: <https://www.cloudflare.com/pt-br/learning/cdn/what-is-caching>.

CLOUDFLARE. **O que é um proxy reverso?** 2024. Acesso em: 02 set. 2025. Disponível em: <https://www.cloudflare.com/pt-br/learning/cdn/glossary/reverse-proxy/>.

FOROUZAN, B. A.; MOSHARRAF, F. **Redes de Computadores**. Porto Alegre: Grupo A, 2013. E-book. ISBN 9788580551693.