

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
CURSO DE CIÊNCIA DA COMPUTAÇÃO

ÉDERSON DE VARGAS VIEIRA

**Otimização da Arquitetura de um Serviço
de E-mails em Produção**

Monografia apresentada como requisito parcial
para a obtenção do grau de Bacharel em Ciência
da Computação.

Prof. Dr. Alexandre da Silva Carissimi
Orientador

Porto Alegre, Julho de 2014

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Vieira, Éderson de Vargas

Otimização da Arquitetura de um Serviço de E-mails em Produção / Éderson de Vargas Vieira. – Porto Alegre: Graduação em Ciência da Computação da UFRGS, 2014.

66 f.: il.

Trabalho de Conclusão (bacharelado) – Universidade Federal do Rio Grande do Sul. Curso de Ciência da Computação, Porto Alegre, BR–RS, 2014. Orientador: Alexandre da Silva Carissimi.

I. Carissimi, Alexandre da Silva. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos Alexandre Netto

Vice-Reitor: Prof. Rui Vicente Oppermann

Pró-Reitor de Graduação: Prof. Sérgio Roberto Kieling

Diretor do Instituto de Informática: Prof. Luís da Cunha Lamb

Coordenador do Curso de Ciência da Computação: Prof. Raul Fernando Weber

Bibliotecária-Chefe do Instituto de Informática: Beatriz Regina Bastos Haro

"The cake is a lie."

— APERTURE SCIENCE SUBJECT

AGRADECIMENTOS

Primeiramente, gostaria de agradecer ao meu pai, José Almaides Vieira. Agradeço por todo o apoio que me foi dado desde antes do ingresso na universidade até o presente momento. Sem esse apoio, provavelmente, eu sequer teria começado a estudar na UFRGS. Agradeço também a minha mãe, Cacilda de Vargas Vieira, por todos os lanches que ela trouxe enquanto eu estava em meu quarto, de cabeça baixa, estudando para as cadeiras do curso. Esses lanches nunca passaram despercebidos e, com certeza, foram muito importantes, por mais que a minha reação tenha sido um “obrigado” bem sem graça. Agradeço também aos meus irmãos Anderson e Leandro. Saber que eles estavam tão ansiosos quanto eu, esperando o resultado do vestibular, foi uma surpresa e, com certeza, fiquei muito feliz com o fato. Agradeço, ainda, à minha namorada, Aline Flor, pela companhia que é e por ter me aguentado durante todo o curso de graduação, sendo compreensiva nos momentos em que tive que dar prioridade às atividades da universidade.

Agradeço também ao Luís Otávio Soares, chefe dos laboratórios do Instituto de Informática, pela oportunidade de trabalhar com ele durante cinco dos seis anos e meio que passei dentro da universidade. Graças a ele e toda a equipe AdmRede - que sempre esteve por perto - aprendi muito sobre responsabilidades e ganhei muita experiência na área que sempre quis trabalhar, redes de computadores. Agradeço muito ao Prof. Dr. Alexandre da Silva Carissimi, por ter me aceitado como orientando e por todo o auxílio e atenção que ele destinou a mim e meu trabalho durante esse último ano. Foi, absolutamente, um privilégio realizar esse trabalho sob sua orientação e estou muito feliz por ter batido em sua porta, procurando por um orientador.

Por último - e não menos importante - agradeço a todos os amigos e colegas que me acompanharam durante todos os últimos anos. Em especial, agradeço ao grupo “Pelados em POA”, que por diversas vezes me animaram e fizeram rir, mesmo durante o mais terrível dos finais de semestre.

SUMÁRIO

RESUMO	9
ABSTRACT	11
LISTA DE FIGURAS	13
LISTA DE TABELAS	15
LISTA DE ABREVIATURAS E SIGLAS	17
1 INTRODUÇÃO	19
1.1 Objetivo	20
1.2 Organização do Trabalho	20
2 SERVIÇO DE CORREIO ELETRÔNICO	21
2.1 Mail User Agent(MUA)	21
2.2 Mail Transfer Agent (MTA)	22
2.3 Local Delivery Agent (LDA)	22
2.3.1 Formato Mbox	22
2.3.2 Formato Maildir	23
2.4 Protocolos de Acesso e Envio de Mensagens	24
2.4.1 Post Office Protocol (POP3)	24
2.4.2 Internet Message Access Protocol (IMAP)	25
2.4.3 Simple Mail Transfer Protocol (SMTP)	25
2.4.4 Local Mail Transfer Protocol (LMTP)	26
2.5 Soluções de Software Livre	26
2.5.1 MUA	27
2.5.2 MTA	27
2.5.3 LDA	27
3 DOVECOT	29
3.1 Servidor POP3/IMAP	29
3.2 Função LDA	31
3.2.1 Atualização dos Arquivos de Índice	31
3.2.2 Controle de Quota	31
3.2.3 Interpretador Sieve	32
3.3 Implementações LDA do Dovecot	32
3.3.1 Dovecot LDA	32
3.3.2 LMTP Server	33

3.4	Formato dbox	33
3.4.1	Single-dbox (sdbox)	34
3.4.2	Multi-dbox (mdbox)	35
3.5	Ferramentas	37
4	PROPOSTA	39
4.1	Arquitetura Atual	39
4.2	Arquitetura Proposta	42
4.3	Considerações Finais	43
5	AMBIENTE DE TESTES	45
5.1	Máquinas Físicas	46
5.2	Máquinas Virtuais	46
6	AVALIAÇÃO	49
6.1	Critérios de Avaliação	49
6.2	Metodologia	50
6.3	Cenários de Testes	50
6.4	Resultados	52
6.4.1	Tempos de Recebimento e Depósito	52
6.4.2	Uso de Memória	55
6.4.3	<i>Load Average</i>	59
7	CONCLUSÃO	63
	REFERÊNCIAS	65

RESUMO

Mundialmente, o número de contas e usuários de serviço de correio eletrônico tem aumentado a cada ano. Esse crescimento é reflexo da importância do serviço de e-mail em todo o mundo. Como uma instituição de ensino, o Instituto de Informática da Universidade Federal do Rio Grande do Sul oferece aos seus alunos e funcionários um serviço de e-mail para que sejam realizadas as comunicações de cunho acadêmico. Porém, o atual serviço apresenta uma série de limitações em sua arquitetura, colaborando para que os usuários utilizem serviços de terceiros para essa finalidade. Este trabalho propõe uma mudança na arquitetura do serviço de e-mails do instituto, substituindo o atual *Local Delivery Agent*, o Procmail, por uma solução alternativa, o Dovecot *LMTP Server*. Essa modificação tem por objetivo reduzir o consumo de recursos computacionais do serviço oferecido, permitir que novas funcionalidades sejam empregadas e evitar que os usuários optem por soluções de correio eletrônico alternativas.

Palavras-chave: Dovecot, LDA, LMTP, MDA, Procmail, serviço de e-mails, SMTP.

Active Email Service Optimization

ABSTRACT

Worldwide, the number of accounts and users of email service has increased every year. This growth reflects the importance of email service around the world. As an educational institution, the Institute of Informatics of Federal University of Rio Grande do Sul offers to yours students and staff an email service for academics nature communications. However, the current service has a number of limitations in its architecture, collaborating with the users migration to third party services. This work proposes a shift in architecture from institutes email service, replacing the current Local Delivery Agent, Procmal, for an alternative solution, Dovecot LMTP Server. This modification aims to reduce the consumption of computational resources of the service offered, allowing new functionality to be employed and prevent users opt for an alternative email service.

Keywords: Dovecot, email service, LDA, LMTP, MDA, Procmal, SMTP.

LISTA DE FIGURAS

2.1	Arquitetura de um Serviço de E-mail.	21
3.1	Estrutura dos Arquivos no Formato dbox	34
3.2	Relação Entre os Arquivos de Índice no Formato mbox.	36
4.1	Arquitetura do Serviço de e-mails do INF.	39
4.2	Proposta de Arquitetura para o Serviço de e-mails do INF.	42
5.1	Arquitetura do Ambiente de Testes com Procmail.	45
5.2	Arquitetura do Ambiente de Testes com LMTP Server.	46
6.1	Tempos de Envio e Depósito para Mensagens de 1024B.	53
6.2	Tempos de Envio e Depósito para Mensagens de 10240B.	53
6.3	Tempos de Envio e Depósito para Mensagens de 102400B.	54
6.4	Tempos de Envio e Depósito para Mensagens de 1024000B.	54
6.5	Uso de Memória da Máquina Infiniti com Mensagens de 1024B.	55
6.6	Uso de Memória da Máquina Infiniti com Mensagens de 10240B.	55
6.7	Uso de Memória da Máquina Infiniti com Mensagens de 102400B.	56
6.8	Uso de Memória da Máquina Infiniti com Mensagens de 1024000B.	56
6.9	Uso de Memória da Máquina Porsche nos Cenários com 1000 Usuários.	57
6.10	Uso de Memória da Máquina Porsche nos Cenários com 100 Usuários.	57
6.11	Uso de Memória da Máquina Porsche nos Cenários com 10 Usuários.	58
6.12	Uso de Memória da Máquina Porsche nos Cenários com 1 Usuário.	58
6.13	Load Average da Máquina Infiniti com Mensagens de 1024B.	59
6.14	Load Average da Máquina Infiniti com Mensagens de 10240B.	59
6.15	Load Average da Máquina Infiniti com Mensagens de 102400B.	60
6.16	Load Average da Máquina Infiniti com Mensagens de 1024000B.	60
6.17	Load Average da Máquina Porsche nos Cenários com 1000 Usuários.	61
6.18	Load Average da Máquina Porsche nos Cenários com 100 Usuários.	62
6.19	Load Average da Máquina Porsche nos Cenários com 10 Usuários.	62
6.20	Load Average da Máquina Porsche nos Cenários com 1 Usuário.	62

LISTA DE TABELAS

5.1	Máquinas Reais.	47
5.2	Máquinas Virtuais.	47
6.1	Cenários de Teste.	51

LISTA DE ABREVIATURAS E SIGLAS

ADMREDE	Equipe de Administração de Redes
AIP	Arquivo de Índice Principal
DNS	Domain Name Server
HTTP	HyperText Transfer Protocol
IMAP	Internet Message Access Protocol
INF	Instituto de Informática
LDA	Local Delivery Agent
LMTP	Local Mail Transfer Protocol
MDA	Mail Delivery Agent
MTA	Mail Transfer Agent
MUA	Mail User Agent
NSA	National Security Agency
NFS	Network File System
POP3	Post Office Protocol
SMTP	Simple Mail Transfer Protocol
TCP	Transfer Control Protocol
UFRGS	Universidade Federal do Rio Grande do Sul

1 INTRODUÇÃO

Apesar da crescente utilização de programas de mensagens instâneas e redes sociais para comunicação na Internet, mundialmente, o número de contas de e-mails está crescendo. A tendência é que o número de contas ultrapasse a marca 4.1 bilhões em 2014 e, em 2018 atinja a marca de 5.2 bilhões. O número de usuários, tanto corporativos como individuais, devem atingir a marca de 2.5 bilhões em 2014 e, atingir a marca de 2.8 bilhões de usuários em 2018 (THE RADICATI GROUP, 2014). Esses números mostram a importância que o correio eletrônico tem nas comunicações realizadas pela Internet hoje e em um futuro próximo.

Atualmente, o Instituto de Informática da UFRGS (INF) possui 3421 contas de correio eletrônico - entre professores, alunos e técnicos administrativos. Pelo menos 160 novas caixas postais são criadas anualmente com a chegada dos novos alunos de graduação, e ainda mais são criadas com o eventual ingresso de novos professores, técnicos administrativos, alunos de intercâmbio e de pós-graduação. As caixas postais dos alunos egressos são preservadas mesmo após serem desvinculados da universidade e o serviço de e-mail é mantido para esses usuários.

No momento da realização deste trabalho, julho de 2014, 1005 contas de e-mail do INF estão redirecionando suas mensagens diretamente para servidores privados tais como Google, Yahoo e Microsoft e é possível que outras contas estejam transferindo o conteúdo de suas caixas postais para tais empresas por outros métodos - como POP3 - dificultando a identificação de que as mensagens, na prática, estão sendo armazenadas nesses servidores. Alguns dos motivos pelo qual os usuários realizam esses redirecionamentos, são a falta ou a ineficiência de algumas funcionalidades oferecidas pelo serviço de e-mails do Instituto, que muitas vezes é inferior aos oferecidos por essas outras empresas.

A atual arquitetura do servidor de e-mails do INF possui uma série de deficiências que causam um impacto significativo no uso de diversos recursos computacionais, tais como: processamento, memória, disco e enlace de rede. Tais limitações afetam diretamente o desempenho do serviço de correio eletrônico e inviabilizam a implantação de funcionalidades que tornem o serviço de e-mail mais atrativo aos usuários, colaborando assim, com o aumento no número de redirecionamentos das mensagens para serviços de terceiros. Assim, é essencial que uma nova arquitetura seja planejada, testada e colocada em produção para atender a atual e futura demanda das caixas postais, reduzindo o consumo dos recursos computacionais, permitindo que seja oferecida uma entrega de mensagens sem atrasos e incentivando os usuários a não redirecionar suas caixas postais para outras empresas ou instituições.

1.1 Objetivo

Este trabalho tem por objetivo propor melhorias para a atual arquitetura do serviço de e-mails do INF, levando em consideração a atual política adotada pela Equipe de Administração de Redes do INF (ADMREDE), de que cada serviço deve ser configurado em uma máquina virtual individual, que os programas devem ser baseados em soluções de software livre, e, que os serviços que atualmente estão em equipamentos legados devem ser movidos para novos equipamentos com hardware e software mais atualizados.

Assim, neste trabalho, é apresentada uma análise comparativa do consumo de recursos como, processamento e memória, o tempo de recebimento de e-mails e o tempo para a realização da entrega de múltiplas mensagens recebidas simultaneamente, entre a atual arquitetura e a arquitetura proposta, com intuito de justificar a mudança de uma arquitetura para a outra. Essa nova arquitetura visa atender a demanda de todas as caixas postais dos atuais e futuros usuários do INF e possibilitar que, mais tarde, sejam agregados novos recursos ao serviço de e-mails (como, por exemplo, filtragem anti-spam mais eficiente, filtros - de mensagens - personalizados e suporte a auto-configuração de programas cliente de e-mail) que possam incentivar os usuários a utilizar o serviço oferecido pelo Instituto para o tráfego de mensagens de cunho acadêmico.

1.2 Organização do Trabalho

Este trabalho está dividido em sete capítulos, contando com esta introdução. No próximo capítulo é apresentada a arquitetura típica de um serviço de e-mails e são explicadas as funções de cada componente e os principais protocolos que a compõe. A seguir, é introduzido o Dovecot, suas funcionalidades e, são listadas as vantagens que ele oferece em relação a outras ferramentas similares. A atual arquitetura do serviço de e-mails do INF, suas deficiências, o escopo que este trabalho visa atender e a arquitetura proposta são discutidas no quarto capítulo. No quinto capítulo, é descrito o ambiente utilizado para a realização dos testes e quais os critérios utilizados para a avaliação das arquiteturas. No capítulo seguinte são apresentados e discutidos os resultados da avaliação comparativa. Por último, no capítulo sete, são expostas as conclusões e os resultados obtidos a partir da avaliação realizada.

2 SERVIÇO DE CORREIO ELETRÔNICO

Para disponibilizar um serviço de correio eletrônico (e-mail) a usuários, é necessário implantar uma série de componentes que fazem parte da arquitetura desse serviço. Cada um desses componentes é responsável por uma tarefa que irá permitir que as pessoas possam encaminhar mensagens umas às outras e garantir, também, que as mensagens sejam recebidas em suas caixas postais. Um exemplo típico de uma arquitetura de um serviço de e-mails pode ser visto na Figura 2.1.

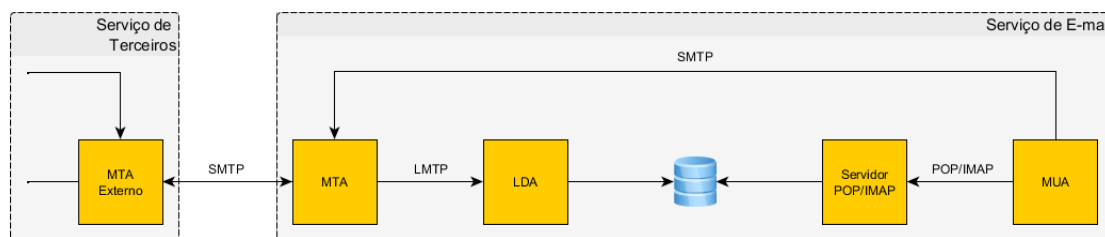


Figura 2.1: Arquitetura de um Serviço de E-mail.

Os componentes exibidos na Figura 2.1 podem ser executados em uma mesma máquina servidora ou podem estar em equipamentos distintos. Essa escolha é realizada pelo administrador do serviço de e-mails, que leva em consideração sua infraestrutura e decide como estruturar todos esses diversos componentes. É possível considerar, também, métricas qualitativas como desempenho, disponibilidade e espaço de armazenamento ocupado.

Diversos outros componentes podem ser agregados à arquitetura do serviço de correio eletrônico, oferecendo funcionalidades como verificação anti-vírus, anti-spam, controle de listas de usuários e controle de duplicatas, porém, esses serviços também consomem recursos computacionais. É necessário que o administrador dessa infraestrutura leve em consideração as características de cada componente e o tempo de processamento para cada tarefa de modo a evitar que gargalos sejam criados, afetando, assim, a qualidade do serviço oferecido.

2.1 Mail User Agent(MUA)

Do ponto de vista de um usuário, um MUA é a interface que ele utiliza para verificar as mensagens recebidas e enviar mensagens eletrônicas (e-mail) a outras caixas postais. Essas interfaces podem estar em programas instalados em um computador pessoal, tablet, smartphone, ou então, podem ser acessadas em um servidor web através de um navegador, permitindo que o usuário acesse sua caixa postal.

Do ponto de vista do serviço de e-mails, um MUA é um agente que se comunica com outros dois componentes da arquitetura: o *Mail Transfer Agent* (MTA) e o processo que possibilita o acesso às mensagens na caixa postal do usuário. Assim, quando um usuário envia uma mensagem de e-mail para alguma caixa postal, o MUA envia a mensagem para o MTA para que ele possa decidir para onde encaminhá-la. Quando, em seu dispositivo, o usuário clica na sua “Caixa de Entrada”, o MUA utiliza protocolos de comunicação como, por exemplo, o IMAP ou o POP3, para acessar as mensagens na caixa postal do usuário, armazenada no servidor de e-mails.

2.2 Mail Transfer Agent (MTA)

O MTA tem por objetivo encaminhar as mensagens que são entregues a ele para o destinatário final das mensagens. Nesse processo de entrega, a mensagem pode passar por vários MTAs intermediários. Os MTAs interagem com três componentes distintos: os MUAs, outros MTAs e, também, o *Local Delivery Agent* (LDA), não sendo visto, nem contactado diretamente pelos usuários finais.

Sempre que o MTA recebe uma mensagem de e-mail, ele avalia se o destinatário da mensagem está sob sua responsabilidade, ou seja, se a caixa postal de destino pertence a algum de seus usuários, ou se a mensagem deve ser encaminhada para outro MTA (representado pelo MTA Externo, na Figura 2.1). Caso o destinatário seja uma das caixas postais sob sua competência, o MTA encaminhará a mensagem ao LDA que, por sua vez, se responsabilizará por depositar a mensagem na caixa postal do usuário. Caso a mensagem não esteja destinada a uma das caixas postais dos usuários aos quais o MTA é responsável, ele realiza uma consulta a um servidor *Domain Name Server* (DNS) para descobrir qual o endereço do MTA responsável por aquela conta de e-mail e encaminha a mensagem a ele, utilizando o protocolo de transferência de mensagens (SMTP).

2.3 Local Delivery Agent (LDA)

Assim como o MTA, o LDA não interage diretamente com o usuário. Um LDA - também conhecido como *Mail Delivery Agent* (MDA) - recebe do MTA a mensagem de e-mail e a indicação de qual caixa postal a mensagem deve ser depositada. Dessa maneira, a atribuição do LDA é depositar essa mensagem em disco, no local destinado ao armazenamento das mensagens do destinatário do e-mail. As mensagens podem ser depositadas diretamente no sistema de arquivos do servidor de e-mails ou, dependendo da implementação do LDA que está sendo utilizado, podem ser armazenadas em um banco de dados. Dois dos formatos de armazenamento de mensagens mais utilizados pelos LDAs utilizam o sistema de arquivos do servidor de e-mails para realizar os depósitos: o formato **mbox** (QMAIL, 1996) e o formato **maildir** (BERNSTEIN, 1995), discutidos a seguir.

2.3.1 Formato Mbox

No formato *mbox*, para cada pasta utilizada no sistema de e-mail (caixa de entrada, ítems enviados, etc.), as mensagens de um usuário são armazenadas, em formato texto, em um único arquivo. Sempre que uma nova mensagem é recebida pelo LDA, ele concatena essa nova mensagem no fim do arquivo.

Uma das vantagens desse formato de arquivo, é que realizar a concatenação de uma nova mensagem ao arquivo da caixa postal é uma tarefa bastante rápida, não exigindo muitos recursos do servidor de e-mails. Outra vantagem, é que por ser um único arquivo

texto, realizar uma busca por um conteúdo específico em todas as mensagens de e-mail se torna uma tarefa bastante simples, não sendo necessário abrir múltiplos arquivos para encontrar a mensagem procurada. Realizar o *backup* dos e-mails também se torna uma tarefa simples, já que toda a caixa postal do usuário está em um arquivo único.

Por outro lado, o formato *mbox* apresenta uma quantidade significativa de desvantagens. Uma delas, é que o espaço ocupado pelo arquivo não pode crescer indefinidamente, sendo limitado pelo sistema de arquivos. Outra desvantagem é que, para excluir uma mensagem, é necessário percorrer todo o arquivo, apagar o conteúdo da mensagem e redimensionar o arquivo, tarefa que exige bastante processamento. Outro inconveniente, é que, em um arquivo grande, aumentam as chances de que o arquivo seja corrompido durante as gravações, danificando toda a caixa postal.

Sob o ponto de vista de desempenho, o formato *mbox* apresenta a desvantagem de empregar *locks* para garantir acesso exclusivo aos arquivos das caixas postais, tornando-se um gargalo. A necessidade da utilização desse recurso vem do fato de que múltiplos componentes do serviço de e-mails podem tentar acessar a caixa postal simultaneamente (LDA e MUA, por exemplo). É possível que um novo e-mail chegue ao LDA enquanto o MUA está exibindo para o usuário as mensagens em sua caixa postal. Essa situação pode gerar um comportamento estranho ao exibir os e-mails para o usuário. O problema é ainda pior caso o usuário esteja excluindo uma mensagem, pois, dessa forma, tanto o MUA quanto o LDA podem tentar editar o arquivo ao mesmo tempo.

Para evitar essas situações, sempre que um componente acessa um arquivo no formato *mbox*, ele bloqueia o arquivo (realiza o *lock*) para garantir que apenas ele terá acesso à caixa postal. Essa garantia traz custos: enquanto o usuário estiver acessando seus e-mails, nenhuma nova mensagem pode ser depositada na caixa postal e, múltiplos e-mails destinados a uma mesma caixa postal devem ser serializados (apenas um e-mail pode ser concatenado por vez em um arquivo no formato *mbox*).

Por fim, o formato *mbox* é uma família de diversos tipos de caixa postal incompatíveis entre si (POLLARD, 2004). Isso acontece porque o *mbox* é apenas um conceito e, atualmente, não existe uma padronização na implementação desse tipo de caixa postal. A RFC 4155 (HALL, 2005) é uma tentativa de criar um padrão mínimo para compatibilizar o armazenamento do tipo *mbox* entre suas diversas implementações, porém, o formato em si não possui uma formatação bem definida.

2.3.2 Formato Maildir

O formato *maildir*, assim como o *mbox*, também é um conceito, sem uma definição formal. Esse formato visa resolver algumas das deficiências do formato *mbox*, porém, acaba oferecendo, também, algumas desvantagens.

Nesse formato, cada mensagem é armazenada em um arquivo separado, com nome único (nenhum outro arquivo possui o mesmo nome) e cada pasta (caixa de entrada, ítems enviados, etc) fica em um diretório distinto. Por padrão, uma caixa postal no formato *maildir* apresenta três diretórios: *tmp*, *new* e *cur*. Quando uma mensagem é entregue pelo LDA, inicialmente, ele deposita essa mensagem no diretório *tmp*. Após realizar o depósito, o arquivo é redirecionado para a pasta *new*, onde ficará visível para o MUA. Quando o MUA acessa o novo e-mail, ele move o arquivo para a pasta *cur*, onde os e-mails ficarão armazenados até serem movidos ou excluídos.

Como cada mensagem é armazenada em um arquivo individual e, quando o MUA acessa a caixa *new*, o e-mail já foi depositado, não existe a necessidade de realização de *locks*. Assim, o formato *maildir* oferece vantagens sobre o formato *mbox*. Uma delas é

que múltiplos processos podem modificar a caixa postal simultaneamente, dessa forma, e-mails podem ser entregues paralelamente por múltiplas instâncias do LDA, mesmo quando o usuário estiver acessando sua caixa postal através de um MUA. Outra vantagem, é que localizar, acessar e excluir uma mensagem é bastante simples, uma vez que só é necessário atuar diretamente no arquivo referente à mensagem. Mais uma vantagem é que quando um arquivo é corrompido, apenas a respectiva mensagem fica comprometida. Por fim, o tamanho da caixa postal não é limitado pelo sistema de arquivos (em teoria, uma única mensagem no formato *maildir* pode ter o mesmo tamanho de uma caixa postal em formato *mbox*).

Apesar de todas as vantagens oferecidas pelo formato *maildir*, trabalhar com múltiplos arquivos também apresenta algumas desvantagens. Uma delas, é que pesquisar em todas as mensagens por um conteúdo específico exige que diversos arquivos sejam abertos, processados e fechados, tendo um tempo de resposta consideravelmente maior que o formato *mbox* que abre apenas um. Outra desvantagem é que a soma do tamanho de um grande número de arquivos pequenos impacta em uma ocupação maior de espaço em disco do que um único arquivo com as mesmas mensagens.

Na prática, o formato *maildir* costuma oferecer um desempenho superior ao *mbox* em caixas postais ativas (mensagens constantemente acessadas e modificadas). Porém, para realizar *backups* ou arquivamento de mensagens (e-mails que não serão lidos com tanta frequência) o formato *mbox* tende a se adequar melhor (FUNK, 2004).

2.4 Protocolos de Acesso e Envio de Mensagens

Para que um usuário possa acessar sua caixa postal, ele precisa de um *Mail User Agent*. Em alguns casos, o usuário acessa suas mensagens através de um navegador web. Nessas situações, o provedor do serviço de e-mails disponibiliza uma aplicação, chamada *webmail*, que será acessada utilizando o protocolo *Hypertext Transfer Protocol* (HTTP), o mesmo utilizado para acessar qualquer página na Internet.

Entretanto, existem situações onde o usuário prefere configurar um MUA em seu computador para acessar suas mensagens de e-mail sem a necessidade de acessar via *webmail*. Em alguns casos, tais aplicações oferecem, em uma mesma interface, recursos como agenda de contatos, calendário e organização de tarefas, além da visualização dos e-mails. Para que esses programas possam acessar a caixa postal do usuário é necessário utilizar algum protocolo de acesso, como o POP3 ou o IMAP, para realizar a comunicação com o servidor de e-mails. Da mesma forma, os programas que implementam os MUAs empregam protocolos específicos para enviar mensagens, por exemplo, o SMTP. Esses protocolos são brevemente apresentados a seguir.

2.4.1 Post Office Protocol (POP3)

Definido pela RFC 1939 (MYERS; ROSE, 1996), o protocolo POP3 tem como função acessar um servidor de e-mails e recuperar as mensagens de um dado usuário. Ele foi definido com base na necessidade dos usuários de ter acesso às suas caixas postais e de que nem sempre eles teriam disponível uma conexão à Internet para acessar suas mensagens no servidor de e-mails, situação bastante comum quando se utilizava conexão discada (*dial up*).

O protocolo POP3 estabelece uma conexão TCP com o servidor de e-mails, confirma as credenciais da conta do usuário (nome de usuário e senha) e realiza o *download* de todas as mensagens que estão no servidor para o computador do usuário.

Por padrão, após realizar o *download* dos e-mails, o protocolo POP3 exclui todas as mensagens da máquina servidora, liberando espaço para que novas mensagens sejam recebidas. Porém, esse comportamento não permite que o usuário acesse as mensagens - já baixadas - de outro computador ou até mesmo pelo webmail. Assim, foi criada a possibilidade de “deixar uma cópia no servidor”, possibilitando que as mensagens sejam acessadas tanto localmente quanto remotamente. Entretanto, essa opção faz com que essas mensagens ocupem espaço em disco tanto na máquina servidora quanto nas máquinas que realizaram o *download* dos e-mails.

2.4.2 Internet Message Access Protocol (IMAP)

Definido pela RFC 1730 (CRISPIN, 1994), assim como o protocolo POP3, a função do protocolo IMAP é acessar o servidor de e-mails e possibilitar ao usuário ter acesso às mensagens depositadas em sua caixa postal utilizando um MUA.

O IMAP possui diversas funcionalidades que o diferencia do POP3. O protocolo IMAP, nativamente, apenas acessa as mensagens do usuário no servidor de e-mails, não realizando o *download* para sua estação de trabalho. Assim, quando utilizado, o protocolo IMAP permite que o usuário possa acessar suas mensagens que estão no servidor de qualquer computador com conexão à Internet. Além disso, inicialmente, o protocolo IMAP não acessa todo o conteúdo das mensagens na caixa postal do usuário, mas sim, apenas informações básicas (como remetente, assunto, data de envio, etc) que ficam armazenadas no início das mensagens, área conhecida como **cabeçalho** das mensagens. Apenas quando o usuário solicita a abertura de uma mensagem em particular, então, o protocolo IMAP acessa o conteúdo completo da mensagem.

Uma das vantagens oferecidas pelo protocolo IMAP, é que ele permite que as mensagens do usuário sejam organizadas em pastas de maneira que, independentemente de que computador o usuário esteja acessando sua caixa postal, a hierarquia de pastas seja mantida. Além disso, é possível que regras de filtragem sejam criadas apenas uma vez e que sejam aplicadas sem levar em consideração qual MUA o usuário está utilizando.

Após o surgimento da Internet banda larga, as estações dos usuários permanecem *online* por muito mais tempo (CHOUDRIE; DWIVEDI, 2007) e, com a popularização da Internet móvel, é comum que um único usuário acesse sua caixa postal de múltiplos dispositivos como smartphones, tablets e notebooks, fazendo com que o protocolo IMAP seja uma excelente opção para se ter acesso aos e-mails em todos os dispositivos, sem a necessidade de ter que configurar todas suas regras de filtragem e organizar suas pastas em cada um deles.

2.4.3 Simple Mail Transfer Protocol (SMTP)

Definido, em sua atual versão, na RFC 5321 (J. KLENSIN, 2001), o protocolo SMTP, hoje, é o protocolo padrão para transferência de e-mails na Internet. Ele tem como função possibilitar a comunicação entre os MTAs e a transferência das mensagens de e-mails entre eles.

O protocolo SMTP é um protocolo baseado em texto, onde um emissor pode comunicar com o receptor fornecendo uma sequência de comandos escritos, com as informações necessárias para que uma sessão seja estabelecida. Durante essa conexão entre os envolvidos, zero ou mais mensagens podem ser transmitidas, de acordo com os parâmetros utilizados na criação dessa sessão.

AS ferramentas MUA costumam oferecer configurações que possibilitam que um usuário estabeleça uma comunicação SMTP com um serviço de e-mails para enviar suas

mensagens. Nessas situações, o programa de e-mails do usuário realiza uma comunicação SMTP com um MTA, de acordo com a definição realizada na RFC 5321, e transfere as mensagens para que ele (o MTA) decida para onde encaminhar o e-mail recebido.

Um MTA, por sua vez, pode realizar uma comunicação SMTP com outros MTAs para repassar os e-mails recebidos. Esses repasses podem ser realizados múltiplas vezes, fazendo com que as mensagens passem por quantos MTAs sejam necessários até que os e-mails cheguem em seu destino.

O protocolo SMTP funciona de maneira “*store-and-forward*” e tenta garantir que a mensagem sempre seja entregue ao destino correto. Assim, um serviço que implementa o protocolo SMTP precisa implementar um sistema de fila de mensagens para que elas (as mensagens) sejam armazenadas durante as tentativas de envio (ou depósito) dos e-mails. Se uma mensagem não chegar ao seu destino, o serviço (que assumiu a responsabilidade da entrega da mensagem) deve responder ao remetente o motivo da falha na entrega da mensagem.

2.4.4 Local Mail Transfer Protocol (LMTP)

O protocolo LMTP, definido na RFC 2033 (MYERS, 1996), foi baseado no SMTP e possui um comportamento muito similar. O LMTP surgiu para atender a necessidade de se realizar a transferência de mensagens entre processos dentro de uma mesma infraestrutura, mesmo que esses processos estejam executando em computadores diferentes.

Na prática, para realizar essa transferência, seria possível utilizar o protocolo SMTP. Porém, da maneira como o protocolo SMTP foi planejado, ou o processo (que está recebendo as mensagens) aceita a responsabilidade de entregar a mensagem para **todos** os destinatários, ou então, rejeita a mensagem completamente.

Não é responsabilidade de um LDA oferecer ao remetente garantia de entrega de mensagens, e adicionar gerência de filas em sua arquitetura só iria aumentar sua complexidade. Dessa forma, quando uma mensagem é transferida utilizando o protocolo LMTP, um serviço LMTP pode aceitar ou rejeitar o recebimento de uma mensagem, individualmente, para cada destinatário. Como um MTA naturalmente já implementa o sistema de filas para utilizar o protocolo SMTP, essa complexidade pode ser mantida no MTA, que se responsabiliza por re-tentar a transferência via LMTP ou, então, responder ao remetente que a entrega (para um destinatário em específico) não foi realizada com sucesso.

2.5 Soluções de Software Livre

Para cada um dos componentes que compõem um serviço de e-mails, diversas soluções foram desenvolvidas. Muitas dessas soluções são desenvolvidas por empresas privadas, outras são desenvolvidas por comunidades formadas por diversas pessoas espalhadas pelo mundo.

Algumas dessas soluções, privadas ou não, são proprietárias e, muitas vezes, utilizam protocolos próprios que podem não ser compatíveis com outras soluções, como, por exemplo, o Microsoft Exchange e o Lotus Notes. Normalmente, tais soluções não tem o seu código fonte divulgado, limitando suas funcionalidades àquelas implementadas pela entidade responsável por seu desenvolvimento e - em muitos casos - para utilizá-la, é necessário adquirir uma licença de uso.

As soluções de *software* livre, por definição, possibilitam que o programa possa ser executado para qualquer propósito (comercial ou não), que seja adaptado conforme a necessidade e redistribuído. Essas características incentivam que grupos de desenvolve-

dores, de empresas privadas ou independentes, em todo o mundo possam colaborar no desenvolvimento de uma ferramenta, agilizando o desenvolvimento de novas funcionalidades. O suporte é dado pelos próprios usuários que integram a comunidade de um determinado programa e não é necessário pagar por uma licença de uso.

A implantação e manutenção de soluções proprietárias, normalmente, implicam em custo financeiro (aquisição da licença e suporte técnico) e, como o desenvolvimento de novas funcionalidades fica restrito à equipe de desenvolvimento, é comum que algumas pessoas, empresas e instituições optem pela utilização de *software* livre.

2.5.1 MUA

Existem diversas soluções de *software* livre que implementam as funcionalidades de um MUA. Algumas dessas soluções se preocupam em atender usuários de um sistema operacional em particular, outras, em oferecer diversas funcionalidades extras que muitas vezes excedem o propósito de ser a interface de acesso aos e-mails. Há ainda soluções que se preocupam em consumir o mínimo de recursos da máquina onde estão sendo executadas e ainda existem as implementações elaboradas para executar em um servidor web. Alguns exemplos desse tipo de aplicação são o Mozilla Thunderbird (MOZILLA, 2003), o Alpine (WASHINGTON, 2007) e o RoundCube (TEAM, 2008).

2.5.2 MTA

Assim como nos MUAs, existem diversas soluções de *software* livre que implementam as funcionalidades de um MTA. Devido a importância desse componente, existem inúmeras implementações que agregam as funcionalidades de MTA, LDA e de servidor POP3/IMAP, oferecendo uma solução de serviço de e-mails completa, porém, essas soluções tendem a ser proprietárias. Apesar de existirem muitas soluções proprietárias, existem algumas implementações de *software* livre com uma grande aceitação entre administradores em todo o mundo, como o Courier Mail Server (VARSHAVCHIK, 2000), Sendmail (INC., 1993) e o Postfix (VENEMA, 1998).

2.5.3 LDA

Muitos MTAs agregam a função de LDA. As três implementações citadas na seção 2.5.2 são exemplos de MTA's que realizam também essa tarefa. Alguns motivos que podem fazer um administrador optar por utilizar outra aplicação para assumir a responsabilidade de um LDA são: distribuir o processamento das mensagens, formato de armazenamento das mensagens, compatibilidade com sistema de arquivos, compatibilidade com banco de dados, processamento de mensagens duplicadas, entre outros. Como exemplos de soluções de *software* livre que realizam a entrega das mensagens nas caixas postais temos o Procmal (BERG, 1990), o Maildrop (VARSHAVCHIK, 1999) e o Dovecot (SIRAINEN, 2002), este último, será apresentado com mais detalhes no próximo capítulo.

3 DOVECOT

O Dovecot é uma solução, de *software* livre, que oferece um serviço - com suporte a conexões POP3/IMAP - que realiza acesso às caixas postais armazenadas em um servidor de e-mails, permitindo que os usuários acessem suas caixas postais através de um MUA compatível com algum desses protocolos de acesso. O Dovecot oferece também as funcionalidades de LDA, privilegiando sua função de servidor POP3/IMAP.

Tendo sua primeira versão disponibilizada em julho de 2002, o Dovecot foi desenvolvido por Timo Sirainen e surgiu com a proposta de ser um *software* livre voltado aos sistemas GNU/Linux e Unix, com baixo consumo de recursos computacionais, com foco principalmente em segurança e compatível com os formatos de caixas postais *mbox* e *maildir*.

3.1 Servidor POP3/IMAP

Apesar de dar suporte ao protocolo POP3, o Dovecot foi planejado para ser um servidor IMAP, fazendo com que o seu funcionamento não seja otimizado para o primeiro protocolo. Para atender todas as especificações do protocolo POP3 (MYERS; ROSE, 1996), o Dovecot faz uma série de processamento tanto em caixas postais no formato *mbox* quanto *maildir* para garantir a comunicação com qualquer MUA. É possível configurar o *software* para não realizar esse processamento extra, porém, isso faz ele violar as especificações do protocolo POP3, podendo gerar comportamentos inesperados nos MUAs (SIRAINEN, 2012a).

Por outro lado, o Dovecot oferece um suporte otimizado ao protocolo IMAP, reduzindo o consumo de recursos no servidor de e-mails, agilizando as operações solicitadas pelo usuário e atendendo completamente (SIRAINEN, 2012b) as exigências da especificação do protocolo, descritas na RFC 1730 (CRISPIN, 1994). Esse comportamento otimizado é obtido através da gerência que o Dovecot realiza de seus arquivos de índice.

Os arquivos de índice do Dovecot (individuais para cada caixa postal e armazenados em disco) visam agilizar o acesso a algumas informações que, comumente, as implementações de servidores IMAP precisam processar toda vez que um MUA realiza um acesso. Esses arquivos permitem que os clientes IMAP acessem os dados nas caixas postais de maneira mais rápida, pois não é necessário realizar, novamente, todo o processamento para carregar os dados em memória. A atualização dos índices é realizada periodicamente (o tempo é configurável), procedimento que é chamado de **sincronização** e os dados são retirados dos cabeçalhos armazenados junto às mensagens. Existem três tipos de arquivos de índice: principal, cache e *log*.

O arquivo de índice principal (AIP) possui algumas informações individuais de cada mensagem: um identificador único, *flags*, palavras-chave, ponteiro para o arquivo de ca-

che e, caso o formato da caixa postal seja *mbox*, apresenta o deslocamento necessário, dentro do arquivo da caixa postal, para recuperar a mensagem. Existe ainda a possibilidade de novas informações serem adicionadas, caso seja utilizado algum *plugin* para estender as funcionalidades do *software*.

Além dessas informações individuais, esse arquivo apresenta informações da caixa postal como um todo, como, por exemplo, quantas mensagens existem, quantas não foram lidas e quantas foram marcadas para exclusão. Existem também informações para que seja identificada a necessidade de realizar uma nova sincronização, como o *timestamp* das pastas */cur* e */new* em caixas postais do tipo *maildir*, ou então, o *timestamp* e o tamanho do arquivo da caixa postal, no caso de caixas do tipo *mbox*. Essas informações agilizam a resposta de requisições realizadas pelo MUA, uma vez que não é necessário realizar todo o processamento na caixa postal novamente.

O AIP é o único que possui *locks* tanto para leitura quanto para escrita. Porém, o bloqueio realizado pela leitura pode durar no máximo dois segundos e é possível configurar para que o bloqueio não seja realizado. A operação de escrita bloqueia apenas durante a sincronização da caixa postal, quando o arquivo de índice é atualizado (SIRAINEN, 2009a).

O arquivo de cache é responsável por armazenar algumas informações de cabeçalho de mensagens para oferecer uma resposta mais rápida ao MUA. Essas informações variam de acordo com o perfil de utilização de cada usuário, evitando consumo desnecessário em disco. Além disso, são armazenadas informações como data de envio ou recebimento, espaço ocupado em disco, parte do conteúdo da mensagem, e a estrutura do corpo do e-mail (CRISPIN, 2003). O processamento desse arquivo é realizado, normalmente, na primeira vez que o cliente tenta acessar suas mensagens.

Clientes IMAP costumam atuar de duas maneiras:

- *Online*: realizam a mesma requisição múltiplas vezes (como no caso de webmails);
- *Offline*: realizam a requisição uma única vez, solicitando os dados de cabeçalhos das mensagens e fazendo cache local (outros tipos de cliente IMAP costumam trabalhar dessa maneira).

Para o primeiro caso, os arquivos de cache oferecem benefício pois, após a primeira requisição, não haverá a necessidade de realizar todo o processamento nas mensagens para se obter novamente as mesmas informações. Para o segundo caso, o arquivo de cache é conveniente quando o usuário costuma acessar seus e-mails de múltiplos dispositivos.

Os dados são armazenados no arquivo de cache somente quando o MUA faz a requisição de tais dados e eles são armazenados por uma semana. Após esse período, os dados vão sendo removidos. No caso de mensagens frequentemente acessadas, ou mensagens recebidas há mais de uma semana e que foram acessadas novamente, os dados são armazenados no arquivo de cache por dois meses (SIRAINEN, 2009b).

Não é realizado nenhum tipo de *lock* para o arquivo de cache durante as operações de leitura e, durante a gravação, o *lock* é realizado apenas tempo o suficiente para expandir o arquivo o suficiente para armazenar as novas informações (a gravação dos dados é realizada sem nenhum bloqueio).

Antes de realizar qualquer alteração no AIP, toda transação realizada pelo Dovecot é armazenada no arquivo de *log*. Essa prática oferece benefícios em situações em que o usuário acessa seus e-mails de múltiplos dispositivos simultaneamente, pois, ao invés de ter que comparar o atual AIP com o que foi consultado anteriormente, o MUA pode apenas

verificar as entradas mais recentes no arquivo de *log*. Outra vantagem é que as alterações no arquivo principal não precisam ser realizadas imediatamente a cada transação, mas sim quando existe um número maior de transações realizadas, reduzindo o número de escritas em disco.

Na prática, existem dois arquivos de *log*, onde, o primeiro armazena apenas as entradas mais recentes e o segundo o histórico de transações anterior. Sempre que o primeiro arquivo se torna muito grande, ele é substituído por um novo arquivo de *log* e passa a ser o arquivo que contém o histórico de transações.

Os *locks* no arquivo de *log* ocorrem apenas durante a escrita de novas transações e durante a sincronização do Dovecot.

3.2 Função LDA

O Dovecot oferece também, além de sua função de servidor IMAP/POP3, a possibilidade de atuar como um LDA. A tarefa de depositar as mensagens nas caixas postais dos usuários pode ser realizadas diretamente pelo MTA ou por outros LDAs, porém, a utilização do Dovecot como um LDA traz alguns benefícios, oferecidos, principalmente, devido a três funcionalidades: atualização dos arquivos de índice, controle de quota e suporte à linguagem Sieve (P. GUENTHER; SENDMAIL; T. SHOWALTER, 2008).

3.2.1 Atualização dos Arquivos de Índice

Realizar a atualização dos três arquivos de índice (AIP, cache e *log*), evidentemente, traz benefícios apenas caso seja utilizado o Dovecot como servidor IMAP/POP3, uma vez que esses arquivos são utilizados apenas pelo próprio Dovecot. Quando esse é o caso, a realização dessa atualização agiliza o tempo de resposta no momento em que um MUA tenta recuperar os dados da caixa postal, uma vez que não será necessário realizar o processamento de todas as mensagens e a atualização dos índices no momento em que o usuário acessar sua caixa postal.

Assim, a atualização dos índices pode ocorrer em dois momentos: no momento da entrega ou no momento do acesso. Decidir o momento da realização dessa atualização é uma decisão do administrador do serviço de e-mails e essa decisão pode influenciar diretamente na carga no servidor e no tempo de resposta ao usuário.

Realizar a atualização dos índices, no momento da entrega das mensagens, agiliza a visualização por parte do usuário. Porém, pode causar um excesso, talvez desnecessário, de processamento em horários de alto volume de mensagens, uma vez que talvez algumas caixas postais não sejam acessadas por tempo indeterminado. Realizar a atualização apenas durante o acesso à caixa postal evita o processamento desnecessário, porém, faz com que o acesso, por parte do usuário, seja mais lento, principalmente se houverem muitas atualizações pendentes (um funcionário voltando de férias, por exemplo).

3.2.2 Controle de Quota

Especificar quota para caixas postais é uma maneira bastante comum de realizar um controle de quanto espaço em disco uma determinada caixa postal pode ocupar. Existem diversas maneiras de se realizar essa tarefa como, por exemplo, deixar que o sistema de arquivos fique responsável por gerenciar o espaço ocupado por uma caixa postal, deixar essa função para um banco de dados, ou realizar esse controle durante a entrega das mensagens através do LDA.

A função de LDA do Dovecot oferece uma funcionalidade para realizar o controle de

quota dos usuários, permitindo que os usuários utilizem apenas o espaço em disco que foi disponibilizado para suas caixas postais, evitando excessos de consumo. É possível ainda informar, para o MUA, quanto da quota já foi utilizado para que o usuário possa gerenciar o seu espaço da maneira como for mais conveniente.

3.2.3 Interpretador Sieve

A terceira funcionalidade oferecida pela função LDA, é um *plugin* chamado Pigeonhole Sieve Interpreter (SIRAINEN, 2010a). Esse *plugin* oferece ao Dovecot suporte a criação de *scripts* utilizando a linguagem Sieve, chamados de filtros.

Os filtros Sieve realizam um processamento nas mensagens antes que elas sejam depositadas na caixa postal. Diversas ações podem ser tomadas, como, por exemplo, depositar uma mensagem em outra pasta, encaminhar, excluir, rejeitar ou descartar e-mails, realizar filtragem anti-spam, verificação anti-vírus, controle de duplicatas, resposta automática, entre outras. Esse processamento leva em consideração o conteúdo da mensagem e, com base nas regras definidas no filtro, a ação é realizada.

É possível ainda que, caso não exista um filtro com uma função em particular, seja desenvolvido um *script* personalizado, realizando exatamente o processamento desejado e agregá-lo ao *plugin*.

3.3 Implementações LDA do Dovecot

Para atuar como LDA, duas implementações são oferecidas pelo Dovecot: o Dovecot LDA e o LMTP Server, ambas com suporte às mesmas funcionalidades.

3.3.1 Dovecot LDA

O Dovecot LDA é um processo de curta duração chamado pelo MTA para depositar um único e-mail para algum destinatário. Assim, para cada novo e-mail: um novo processo é carregado em memória, é realizado o depósito da mensagem, a atualização dos índices e depois o processo é encerrado.

Para utilizar esse tipo de entrega, é necessário que o Dovecot LDA seja executado na mesma máquina que o MTA está instalado, uma vez que o MTA executará um processo local, passando como parâmetros de entrada a caixa-postal destino e a mensagem a ser depositada.

A vantagem de se utilizar esse método, é que o tempo que o MTA leva para transferir o e-mail para o LDA é bastante pequeno, dependendo apenas da velocidade (processamento, barramento, memória) do computador onde é executado. Para servidores de e-mail com baixo volume de dados, ou que possuem todos os componentes (MTA, LDA, servidor POP/IMAP e caixas postais) na mesma máquina, pode ser uma boa solução.

Existem diversas desvantagens no uso desse método, que tendem a se agravar em servidores de e-mail de grande porte - onde o volume de e-mails trafegados é maior, como no caso de universidades e grandes empresas. Uma dessas desvantagens, é o *overhead* de processamento causado pela necessidade de instanciar o *software* e depois encerrá-lo para cada mensagem recebida. Além desse *overhead*, cada instância do LDA consome o mesmo espaço em memória, uma vez que são processos distintos. Esse comportamento faz com que o consumo de memória cresça indefinidamente (quanto mais processos simultâneos, mais memória é ocupada) podendo afetar o desempenho da máquina onde está sendo executado.

Outra desvantagem, é que, em servidores de grande porte, é comum que os diversos

componentes de um servidor de e-mails estejam em computadores distintos, trabalhando de maneira distribuída. Trabalhar dessa maneira exige uma série de protocolos de comunicação entre os diversos componentes. Como o Dovecot LDA realiza a atualização dos arquivos de índice do Dovecot, caso essas informações estejam em outra máquina, o Dovecot LDA precisará de protocolos como o NFS (SUN MICROSYSTEMS, 1989) para acessar os arquivos de índice. Esse tipo de protocolo pode utilizar uma série de recursos de *lock* de arquivos, prejudicando operações de paralelismo na entrega das mensagens. Para atender essa demanda mais agressiva de servidores de grande porte, o *LMTP Server* tende a ser uma melhor solução.

3.3.2 LMTP Server

Todas as funcionalidades oferecidas pelo Dovecot LDA são também implementadas pelo *LMTP Server*. A diferença entre esse método e o outro, é que, ao invés de ser um processo de curta duração, o *LMTP Server* é um processo de longa duração inicializado pelo processo principal do Dovecot. Assim, apenas uma única instância do *LMTP Server* é executada, atendendo todas as mensagens enviadas pelo MTA com essa única instância.

Por receber as mensagens do MTA através do protocolo LMTP, o *LMTP Server* pode executar tanto na mesma máquina que o MTA quanto em uma máquina distinta. O *LMTP Server* atende a todas as exigências do protocolo LMTP, definidas na RFC 2033 (MYERS, 1996). Essa implementação oferece muitas vantagens para servidores de e-mail de grande porte.

Nativamente, o LMTP permite que o servidor de e-mails trabalhe de forma distribuída, evitando o problema com os *locks* durante a atualização dos arquivos de índice. Outra vantagem, é que o consumo de memória é muito menor que o Dovecot LDA, uma vez que apenas um processo é executado, independente do volume de mensagens recebidas. É possível configurar o *LMTP Server* para suportar múltiplas conexões, possibilitando que múltiplas mensagens sejam transferidas paralelamente pelo MTA.

3.4 Formato mbox

Além de oferecer suporte aos formatos *mbox* e *maildir*, o Dovecot conta com um padrão próprio de mailbox: o *dbox*, um formato de alto desempenho. O que justifica esse alto desempenho, é que as *flags* e *keywords* (informações que nos formatos *mbox* e *maildir* são armazenadas como meta dados junto aos cabeçalhos das mensagens) ficam armazenadas apenas nos arquivos de índice, não sendo necessário realizar nenhuma sincronização no momento em que o usuário realiza um acesso à sua caixa postal, nem inserir dados adicionais aos arquivos onde os e-mails estão armazenados. O acesso aos arquivos de índice e sua atualização é muito mais rápida do que realizar o mesmo processamento no arquivo da mensagem, uma vez que seu conteúdo tende a ser muito menor. Porém, isso aumenta significativamente a importância dos arquivos de índice e, caso sejam corrompidos, não é possível recuperá-los sem perda de informações.

Por ser um formato próprio do Dovecot, nem todas implementações de MTA e LDA depositam mensagens nesse formato. Também, nem todo servidor IMAP/POP3 consegue acessar informações em arquivos nesse formato. Contudo, sendo o Dovecot a solução de LDA e servidor IMAP/POP3 utilizada, é possível se beneficiar desse formato de duas maneiras: *single-dbox* e *multi-dbox*.

3.4.1 Single-dbox (sdbox)

O formato *sdbox* é bastante similar ao formato *maildir*: armazena uma mensagem de e-mail por arquivo e cada pasta é armazenada em um diretório diferente. A principal diferença entre os dois formatos é o local onde os cabeçalhos das mensagens são armazenados. Outra diferença é que o formato *sdbox* utiliza uma maneira própria de atribuir nomes únicos aos arquivos onde os arquivos de e-mail são armazenados. Por padrão, os arquivos e pastas no formato *sdbox* seguem a estrutura exibida na figura 3.1a.

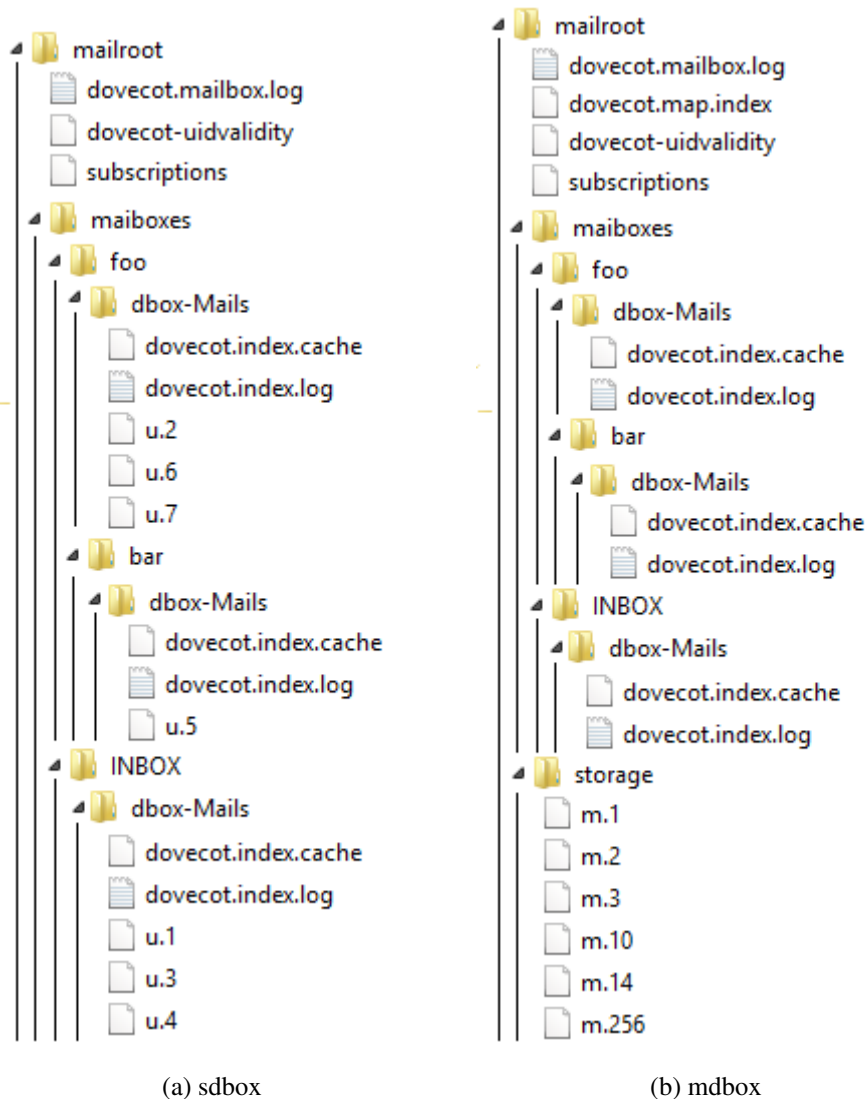


Figura 3.1: Estrutura dos Arquivos no Formato dbox

Todos os arquivos (de dados) no formato *sdbox* **armazenam uma única mensagem de e-mail** e utilizam em sua nomeação o seguinte padrão: “u.*” onde “*” é um número inteiro, crescente e único (ex. u.1, u.2, ... , u.10950). Esse nome é imutável, assim, mesmo que uma mensagem seja movida para outra pasta, ela mantém o nome original.

O arquivo índice de *log* é o *dovecot.mailbox.log*. Os arquivos *subscriptions* e *dovecot-uidvalidity* armazenam, respectivamente, informações de quais caixas postais devem ser exibidas em um MUA utilizando IMAP e validação de identificação das mensagens, ambos conforme a especificação do protocolo IMAP (CRISPIN, 1994).

É possível especificar um novo local para armazenar os arquivos de índice, separando-os dos arquivos com as mensagens de e-mail. Porém, mesmo sendo armazenados em outro local, eles seguem sempre a mesma estrutura de nomeação e diretórios. É importante notar que, para cada pasta (INBOX, *foo*, *foo/bar*) existe um diretório com arquivos de índice (*dovecot.index.cache* e *dovecot.index.log*) próprios. Esses arquivos armazenam apenas os dados referentes aos e-mails que estão presentes em suas respectivas pastas.

3.4.2 Multi-dbox (mdbox)

O formato *mdbox* é bastante similar ao formato *mbox*: ambos **armazenam múltiplas mensagens de e-mail** em um mesmo arquivo. A principal diferença entre os dois formatos, é que, enquanto o *mbox* utiliza apenas um arquivo para cada pasta, o *mdbox* permite que múltiplos arquivos sejam utilizados (a quantidade de e-mails em cada arquivo é parametrizável). Além disso, assim como no *sdbox*, as informações dos cabeçalhos são armazenadas apenas nos arquivos de índice, com exceção de um identificador único, que são armazenados tanto nos arquivos de índice quanto junto às mensagens, e a pasta onde, originalmente, o e-mail foi depositado. Por padrão, os arquivos e pastas no formato *mdbox* seguem a estrutura exibida na figura 3.1b.

No formato *mdbox*, os arquivos que armazenam as mensagens e os arquivos de índice funcionam de maneira diferente, porém, tentando se aproximar ao máximo às definições feitas no *sdbox*. Os arquivos (de dados) no *mdbox* **armazenam uma ou mais mensagens** e sua nomeação segue o seguinte padrão: “m.*” onde o “*” é um número inteiro, crescente e único (ex. m.1, m.2, ..., m.3180). Sempre que uma nova mensagem é entregue pelo LDA, ela é concatenada ao fim do arquivo mais recente. Caso esse arquivo tenha atingido o limite máximo de mensagens que pode armazenar, um novo arquivo é criado para que a mensagem seja concatenada. É necessário destacar que, todos os e-mails, não importando a pasta (caixa de entrada, lixeira, itens enviados) em que está, é armazenado no diretório *storage*.

Os arquivos *dovecot.mailbox.log*, *subscriptions* e *dovecot-uidvalidity* seguem exatamente a mesma lógica apresentada no *sdbox*. Um novo arquivo que aparece nesse formato é o *dovecot.map.index.log*, que veremos com mais detalhes a seguir. É possível também armazenar os arquivos de índice (inclusive o *dovecot.map.index.log*) em outro diretório, separando-os dos arquivos de dados. Porém, mesmo nesse outro local, a estrutura de diretórios é mantida.

Como pode ser visto na figura 3.1b, existem dois tipos de arquivos de índice: o índice de mapeamento (*dovecot.map.index.log*) e índice de caixa postal (*dovecot.index.log*). Ambos os tipos de índices são importantes para o funcionamento do formato, porém, com funções distintas.

O **índice de mapeamento** possui um registro para cada mensagem de e-mail e, em cada registro são armazenadas as seguintes informações:

- *map_uid*: Número crescente e único para cada mensagem.
- *refcount*: Contador que é incrementado sempre que uma mensagem é copiada e decrementado sempre que uma mensagem é excluída.
- *file_id*: Número que indica em qual arquivo a mensagem está armazenada. Ex.: *file_id=10* significa que a mensagem está armazenada no arquivo m.10.
- *offset*: Deslocamento dentro do arquivo que indica onde está armazenada a mensagem.

- **size:** Tamanho ocupado pela mensagem, dentro do arquivo, incluindo os cabeçalhos.

Os **índices de caixa postal**, fazem referência aos e-mails apenas através do *map_uid*. Isso possibilita que as mensagens de e-mail sejam copiadas para qualquer uma das pastas apenas adicionando a referência ao **índice de caixa postal** e incrementando o *refcount* no **índice de mapeamento**. Na prática, múltiplas pastas podem ter a mesma mensagem de e-mail, porém, ela ocupa espaço em disco apenas uma vez. A relação entre os arquivos de índice e as mensagens pode ser observada na figura 3.2.

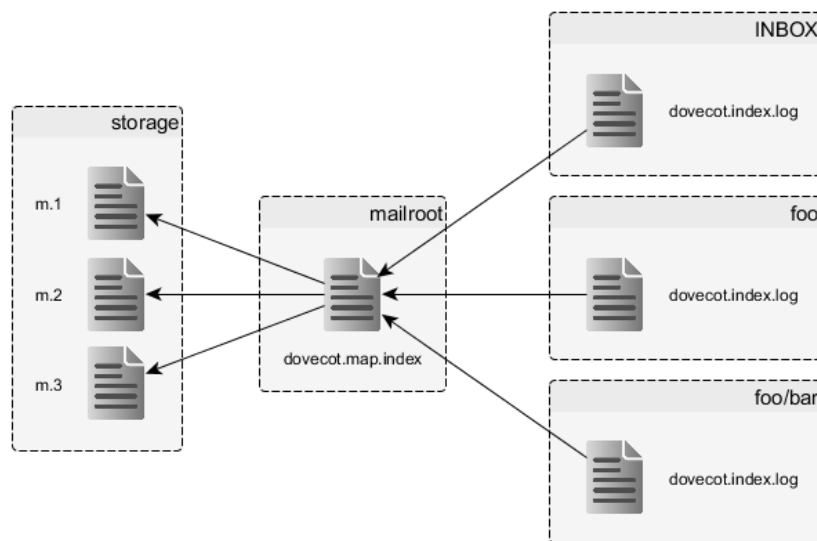


Figura 3.2: Relação Entre os Arquivos de Índice no Formato mbox.

Quando uma mensagem é excluída de alguma pasta, é removida a entrada do **índice de caixa postal** e é decrementado o *refcount* no **índice de mapeamento**. A exclusão da mensagem em si, não é realizada até o momento em que uma rotina de exclusão seja executada. Essa rotina pode (é configurável) executar durante a madrugada, quando, normalmente, existem menos acessos ao disco da máquina.

No momento que a rotina de exclusão é executada, ela identifica no **índice de mapeamento** todas as entradas que possuem o campo *refcount* igual a zero. Depois, acessa cada um dos arquivos de dados onde essas mensagens estão, e copia todas as mensagens que possuem *refcount* positivo para um novo arquivo de dados, atualiza as referências no **índice de mapeamento** e, por último, remove o arquivo de dados original.

Embora a perda dos arquivos de índice não implique em perda das mensagens de e-mail, pode não ser possível recuperar as referências realizadas pelos índices. Para minimizar as chances dessas perdas acontecerem, alguns recursos foram implementadas no Dovecot:

- Cada mensagem possui um identificador único (de 128 bits). Esse identificador é armazenado como meta dado (junto dos cabeçalhos das mensagens) nos arquivos de dados e também nos **índices de caixa postal**. Assim, caso o **índice de mapeamento** seja corrompido, é possível reconstruí-lo acessando todos os arquivos de dados e incrementar o campo *refcount* consultando os **índices de caixa postal**.

- Sempre que um arquivo de índice é atualizado, uma cópia do estado atual é armazenada com o sufixo *.backup* junto ao arquivo original. Assim, caso qualquer um dos arquivos de índice seja corrompido, é possível tentar recuperar o estado anterior.
- Caso todos os arquivos de índice (inclusive os *backups*) sejam perdidos, **não é possível recuperá-los sem perda de informação**. Essa informação é a identificação de quais mensagens cada pasta referencia. Assim, nos meta dados (junto ao cabeçalho de cada mensagem), é armazenada a pasta onde ela foi depositada originalmente. Dessa forma, ao restaurar os índices, as mensagens ficariam, inicialmente, em apenas uma pasta.

3.5 Ferramentas

Para facilitar a administração de um serviço de e-mails que utilize o Dovecot, são disponibilizadas algumas ferramentas. Dois exemplos são o *Doveadm* (SIRAINEN, 2010b) e o *Dsync* (SIRAINEN, 2010c).

O *Doveadm* é um utilitário que oferece, via linha de comando, a possibilidade de realizar, de uma maneira geral, a administração do Dovecot. Com essa ferramenta é possível fazer o Dovecot recarregar suas configurações, parar o funcionamento do serviço, revisar suas configurações, desconectar usuários, realizar testes de autenticação, gerência das mensagens nas caixas postais, entre outras várias funcionalidades. Essas funcionalidades podem ser limitadas aos administradores do serviço de e-mails ou, podem ser utilizadas para realizar processamentos solicitados por usuários, através dos filtros Sieve, por exemplo.

O *Dsync*, por sua vez, é uma ferramenta para atuar diretamente nas caixas postais. Com essa ferramenta, é possível converter as caixas postais, individualmente, para qualquer formato suportado pelo Dovecot (*mbox*, *maildir*, *mdbox* e *sdbox*). Além disso, através dessa ferramenta, é possível realizar um sincronismo das caixas postais com um outro servidor remoto, fornecendo um *backup* contínuo das mensagens dos usuários.

Realizar a manutenção das caixas postais é sempre uma tarefa delicada, pois, a perda dessas informações pode implicar em prejuízos tanto para o usuário final, que perdeu sua mensagem, como para o administrador, que terá que responder por isso. Ferramentas como o *Doveadm* e o *Dsync* oferecem, ao administrador do serviço de e-mails, uma grande ajuda para a adaptação da arquitetura do serviço às suas necessidades, tentando garantir que não haja prejuízo às mensagens.

4 PROPOSTA

Neste trabalho, é proposta a substituição do atual LDA, o Procmal, pelo Dovecot *LMTP Server* na arquitetura do serviço de correio eletrônico do INF. A proposta de nova arquitetura tem como objetivo melhorar, de forma geral, o serviço de e-mails, visando reduzir a carga sobre a máquina onde o Procmal é executado, possibilitando que ela tenha mais recursos para realizar suas outras atividades e também melhorar a qualidade do serviço oferecido aos usuários finais.

4.1 Arquitetura Atual

A atual arquitetura do serviço de e-mails do INF é apresentada na Figura 4.1. Nessa figura, é possível identificar que, nessa arquitetura, existem todos os elementos que foram ilustrados na Figura 2.1 e surgem novos elementos: o Amavis e o Mailman, que serão explicados na sequência.

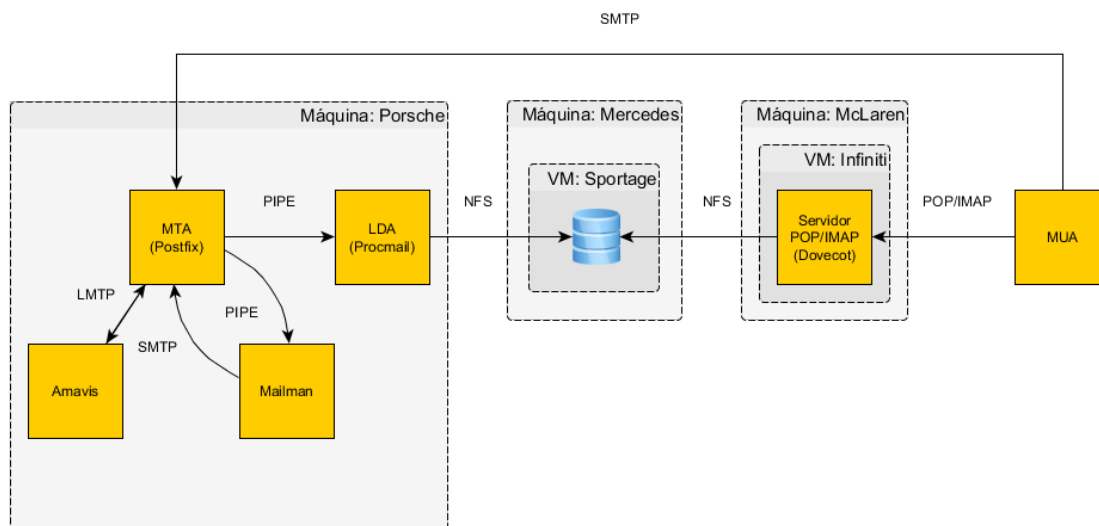


Figura 4.1: Arquitetura do Serviço de e-mails do INF.

Existem três computadores que estão envolvidos com o serviço de e-mails do INF: Porsche, McLaren e Mercedes. A máquina Porsche é responsável por executar as funções de MTA e LDA na arquitetura. As máquinas McLaren e Mercedes são servidoras de virtualização, que gerenciam múltiplas máquinas virtuais para os mais variados serviços oferecidos pela AdmRede aos seus usuários. Entre as máquinas virtuais, da servidora McLaren, existe uma chamada Infiniti, que é onde é executado o Dovecot para atender

requisições POP3 e IMAP. Entre as máquinas virtualizadas pela servidora Mercedes, há a máquina virtual Sportage, que é onde são armazenados os dados dos usuários do INF, inclusive as caixas postais, que são armazenadas no formato *maildir*.

A solução adotada para realizar a função de MTA, é o Postfix. O Procmail é o programa utilizado como LDA e o Dovecot é utilizado apenas como aplicação para possibilitar os acessos através de IMAP e POP3. O Amavis é um uma solução que realiza uma série de processamento nas mensagens para poder identificar e filtrar potenciais mensagens que sejam lixo eletrônico ou que contenham vírus. O Mailman é uma aplicação que gerencia as listas de discussão do INF (por exemplo, as listas de graduação, pós-graduação, professores e funcionários), possibilitando que uma mensagem possa ser enviada apenas para um único endereço de e-mail e que seja entregue a todas as pessoas participantes de uma determinada lista.

Entre cada um desses componentes, uma série de protocolos são utilizados. Alguns, como o POP3, IMAP, SMTP e LMTP já foram apresentados no capítulo 2 e podem ser identificados pelas setas na Figura 4.1. Além desses protocolos, outros dois tipos de comunicação são exibidos nessa imagem: o Pipe e o NFS. O Pipe é um mecanismo que possibilita que o resultado obtido (a saída) de um determinado processo possa ser utilizado como entrada (parâmetro) em um novo processo. No contexto da Figura 4.1, o Postfix gera como resultado uma mensagem de e-mail e entrega ao Procmail para que seja realizado o depósito da mensagem na caixa postal do destinatário. Já o NFS é um protocolo de compartilhamento de dados, entre máquinas, através de uma rede de computadores. No contexto do serviço de e-mails do INF, tanto a máquina Porsche quanto a máquina Infiniti, acessam através do protocolo NFS as mesmas informações que estão armazenadas apenas na máquina Sportage.

Para que seja entendido um pouco melhor a maneira como todos esses componentes interagem entre si, é descrito o seguinte cenário de execução: Uma usuária, Alice, pretende enviar uma mensagem de e-mail para Bob e ambos utilizam seus e-mails do INF. Alice abre seu programa de e-mail favorito e envia um e-mail para Bob. O MUA que ela utiliza envia a mensagem, via protocolo SMTP, ao Postfix (MTA). Ao receber essa mensagem e identificar que o e-mail de destino está sob sua responsabilidade, o Postfix encaminha a mensagem, via protocolo LMTP ao Amavis para que sejam realizadas as verificações de anti-vírus e anti-spam. Após realizado esse processamento, o Amavis devolve a mensagem ao Postfix, também através do protocolo LMTP.

Nesse momento, o Postfix verifica se essa mensagem é destinada a um usuário, ou se é destinada a uma lista de discussão. Como ela é destinada ao usuário Bob, o Postfix ignora o Mailman (o processamento nos casos de listas de discussões será discutido a seguir) e repassa, via Pipe, ao Procmail (LDA). O Procmail realiza o depósito da mensagem, via NFS, na caixa postal do usuário.

A partir desse momento, quando o usuário Bob tentar acessar sua caixa postal, o seu programa cliente de e-mail (MUA) realiza uma requisição (via POP3 ou IMAP) ao Dovecot. O Dovecot acessa, via NFS, a caixa postal de Bob, realiza um processamento para identificar se houve alguma modificação desde o último acesso, atualiza os arquivos de índice e, então, entrega as novas mensagens ao MUA de Bob, que visualiza a mensagem enviada por Alice.

Em casos que o e-mail é destinado a uma lista de discussão, o Postfix envia, via Pipe, a mensagem ao Mailman. O servidor de listas, por sua vez, identifica todos os endereços de e-mail que participam dessa lista e, dividindo em grupos de cinco em cinco destinatários, re-envia, por SMTP, as mensagens ao Postfix, que encaminhará ao Amavis

e repetirá todo o processo descrito anteriormente. Isso significa que, em uma lista com quinze participantes, para um único e-mail, o Mailman enviará três mensagens de volta ao Postfix.

Embora a atual arquitetura do serviço de e-mails do INF funcione, ela apresenta uma série de limitações que prejudicam todas essas operações e, em um cenário em que o volume de e-mails é maior, tende a se agravar ainda mais. Algumas dessas limitações são descritas a seguir:

- **Múltiplos processos do Procmil:** Quando uma mesma mensagem é enviada a múltiplos usuários, para cada usuário, é criada uma nova instância do Procmil para realizar o depósito das mensagens. Esse comportamento pode gerar uma grande concorrência, pelo processador, entre todos os processos e, também, um grande consumo de memória, prejudicando o desempenho da máquina como um todo.
- **Grupos de cinco destinatários do Mailman:** Essa configuração do Mailman poderia ser alterada para um valor maior, reduzindo a carga gerada pelo Amavis. Porém, aumentando esse valor, irá implicar em mais processos simultâneos quando as mensagens forem entregues ao Procmil.
- **Reinserção de mensagens no Amavis:** Quando uma mensagem é enviada para uma lista de discussão, ela é verificada pelo Amavis. Após o Mailman reinserir as múltiplas mensagens no Postfix, cada uma dessas mensagens, individualmente, é verificada pelo Amavis novamente. Essa verificação redundante é desnecessária e apenas onera ainda mais a máquina servidora onde esses processos são executados (Porsche).
- **Disputa pelo compartilhamento NFS:** Como tanto a máquina Porsche quanto a máquina Infiniti fazem acesso ao disco da máquina Sportage, há uma disputa por esse recurso. O uso de *locks* por parte do protocolo NFS pode prejudicar o tempo de acesso aos dados, prejudicando assim o desempenho dos serviços oferecidos pelas duas máquinas.
- **Atualização dos arquivos de índices no momento do acesso às caixas postais:** O Procmil, ao depositar as mensagens nas caixas postais dos usuários, não realiza nenhuma atualização nos arquivos de índice, tampouco notifica o Dovecot de que novas mensagens foram entregues. Dessa forma, o Dovecot, enquanto o usuário acessa sua caixa postal, precisa consultar todos os arquivos na caixa postal do usuário para identificar possíveis modificações e realizar o sincronismo dos arquivos de índice. Esse procedimento é inconveniente por duas razões: em uma sessão de acesso muito longa, o Dovecot precisa ficar acessando constantemente o compartilhamento NFS para identificar se foram feitas novas modificações, onerando ainda mais o compartilhamento e, a outra razão, é que o usuário final terá que aguardar esse processamento antes de ter acesso às novas mensagens (que ainda não foram atualizadas nos arquivos de índice).
- **Manutenção da Máquina Porsche:** A máquina Porsche é o equipamento mais antigo, entre todos os servidores do INF, tanto em seu *hardware* quanto nos *softwares* utilizados. Seus recursos limitados também influenciam na qualidade do serviço de e-mails oferecido pelo Instituto.

4.2 Arquitetura Proposta

A Figura 4.2 apresenta a disposição dos componentes da arquitetura proposta. A principal alteração nessa arquitetura, em relação a original, é a mudança do LDA. O Postfix não realizará mais a entrega das mensagens ao Procmail, mas sim, ao Dovecot. O método de entrega de mensagens proposto é o *LMTP Server*, de modo que as mensagens serão transferidas da máquina Porsche para a máquina Infiniti, via protocolo LMTP, e só então as mensagens serão depositadas, via protocolo NFS, nas caixas postais dos usuários. Essa alteração pode apresentar uma possível solução para os problemas, da atual arquitetura, discutidos na seção 4.1. A maneira como cada um desses inconvenientes é endereçada é explicada a seguir:

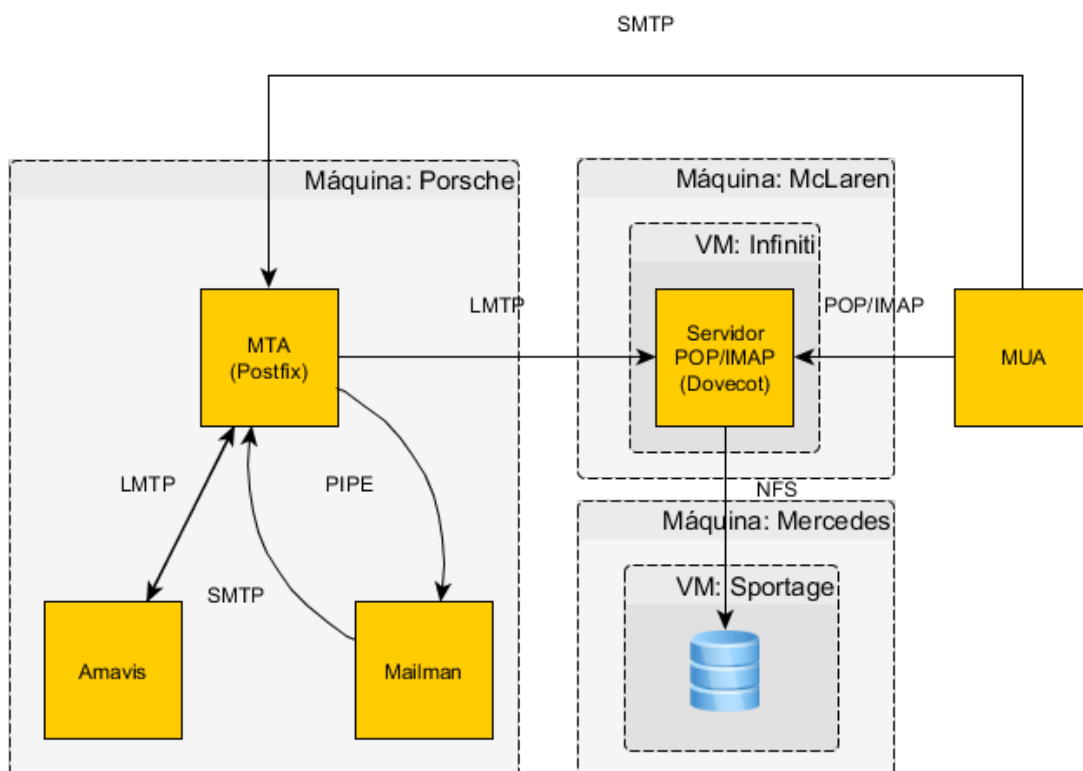


Figura 4.2: Proposta de Arquitetura para o Serviço de e-mails do INF.

- Múltiplos processos do Procmail:** Nessa nova solução, a responsabilidade pelo depósito das mensagens não ficará mais dentro da máquina Porsche, mas sim na máquina Infiniti. Isso distribui a carga do processamento das mensagens entre as diversas máquinas que já participam da arquitetura do serviço. Além disso, o *LMTP Server*, diferentemente do Procmail (que utiliza processos), faz uso de *threads* para realizar o depósito de mensagens nas caixas postais. Isso pode contribuir para um consumo menor de memória e processamento, se comparado ao método anterior.
- Grupos de cinco destinatários do Mailman:** Como não existe mais o aumento no número de processos instanciados pelo Procmail, é possível que seja aumentado o número de destinatários por mensagens nos casos de e-mails enviados às listas de discussão. Isso implica que menos mensagens serão reinseridas no Postfix e,

consequentemente, menos mensagens serão processadas pelo Amavis, reduzindo o consumo de processamento na máquina Porsche.

- **Reinserção de mensagens no Amavis:** Esse problema já seria reduzido, conforme discutido no item anterior. Esse problema seria completamente endereçado, caso a versão (instalada) do Mailman permita entregar as mensagens por LMTP. Se esse for o caso, as mensagens poderiam ser entregues diretamente ao Dovecot, não sendo necessário passar novamente pelo Postfix.
- **Disputa pelo compartilhamento NFS:** Com o abandono do Procmail, não há mais necessidades da máquina Porsche realizar acessos ao compartilhamento NFS oferecido pela máquina Sportage. Dessa forma, no que compete ao serviço de e-mails, apenas a máquina Infiniti realizará acessos a esse compartilhamento, sem o problema de concorrência pelo acesso aos arquivos.
- **Atualização dos arquivos de índices no momento do acesso às caixas postais:** Utilizando o Dovecot como LDA, os arquivos de índices são atualizados no momento do depósito das mensagens. O Dovecot confia no estado atual desses índices e não tem necessidade de ficar realizando acessos ao compartilhamento NFS, para atualização dos índices, durante a sessão do usuário. Além disso, sempre que uma nova mensagem for depositada na caixa postal de um usuário, ele terá acesso imediato a ela, sem a necessidade de aguardar pelo sincronismo dos arquivos de índice.
- **Manutenção da Máquina Porsche:** Remover a função de LDA da máquina Porsche é um primeiro passo para que toda a arquitetura do serviço de e-mails seja retirada desse equipamento e migrada para uma máquina virtual com melhor capacidade de processamento.

4.3 Considerações Finais

Realizar a substituição do Procmail, LDA da atual arquitetura do serviço de correio eletrônico do INF, pelo Dovecot *LMTP Server*, pode apresentar uma melhoria imediata no serviço de e-mail do Instituto. Essas melhorias surgem como distribuição da carga de processamento das mensagens entre as máquinas envolvidas, redução de concorrência nos compartilhamentos NFS, utilização de equipamentos com maior poder de processamento, possibilidade de otimizar os parâmetros de configuração dos diversos componentes da arquitetura e acesso mais rápido às caixas postais. Baseado nessas expectativas de melhorias, é necessária a realização de testes para mensurar e comprovar se a arquitetura proposta realmente vai melhorar o atual serviço de correio eletrônico.

5 AMBIENTE DE TESTES

Para realizar uma comparação entre as duas arquiteturas propostas, foi utilizada uma infraestrutura paralela ao servidor em produção, para não houvesse impacto negativo em nenhum dos serviços oferecidos pelo INF. Por limitações no número de máquinas disponíveis para a realização do experimento, foram utilizadas máquinas virtuais para simular as máquinas Porsche, Infiniti e Sportage. As três máquinas virtuais são executadas na máquina Ederson e, uma outra máquina física - AdmRede - é a responsável por atuar como MUA.

As duas arquiteturas criadas para realização dos testes estão representadas nas figuras 5.1 e 5.2. Nota-se que, nessas duas arquiteturas, foram retirados os componentes Amavis e Mailman. Esses componentes foram removidos pois, na ordem de participação no recebimento de mensagens, são processos que ficam entre o MTA e o LDA. Na atual arquitetura, ambos são gargalos que limitam o fluxo de mensagens dentro do serviço de e-mails. O Amavis é um gargalo porque o tempo de processamento das mensagens nesse serviço é maior que o tempo de depósito das mensagens. O Mailman é um gargalo porque ele agrava ainda mais a limitação do Amavis, fazendo com que ele tenha que processar múltiplas vezes a mesma mensagem. No ambiente de testes, esses dois limitantes foram removidos permitindo que possamos realizar uma comparação direta entre os dois métodos de entrega, Procmail e *LMTP Server*, em uma condição de estresse, sem a influência de processos externos.

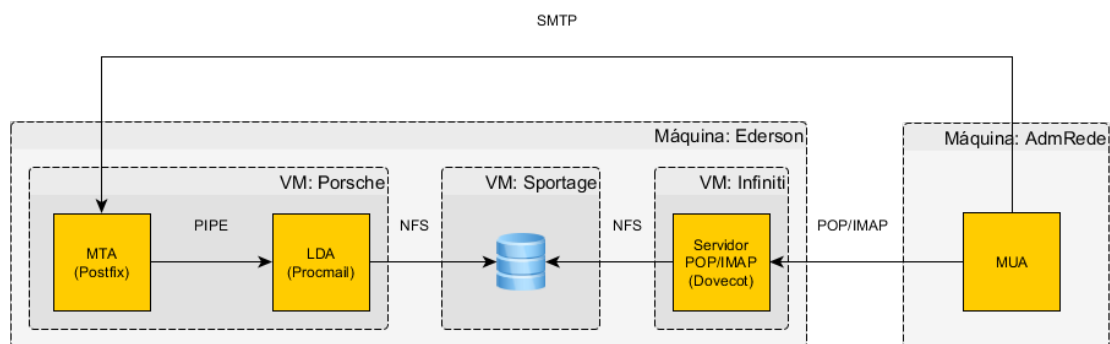


Figura 5.1: Arquitetura do Ambiente de Testes com Procmail.

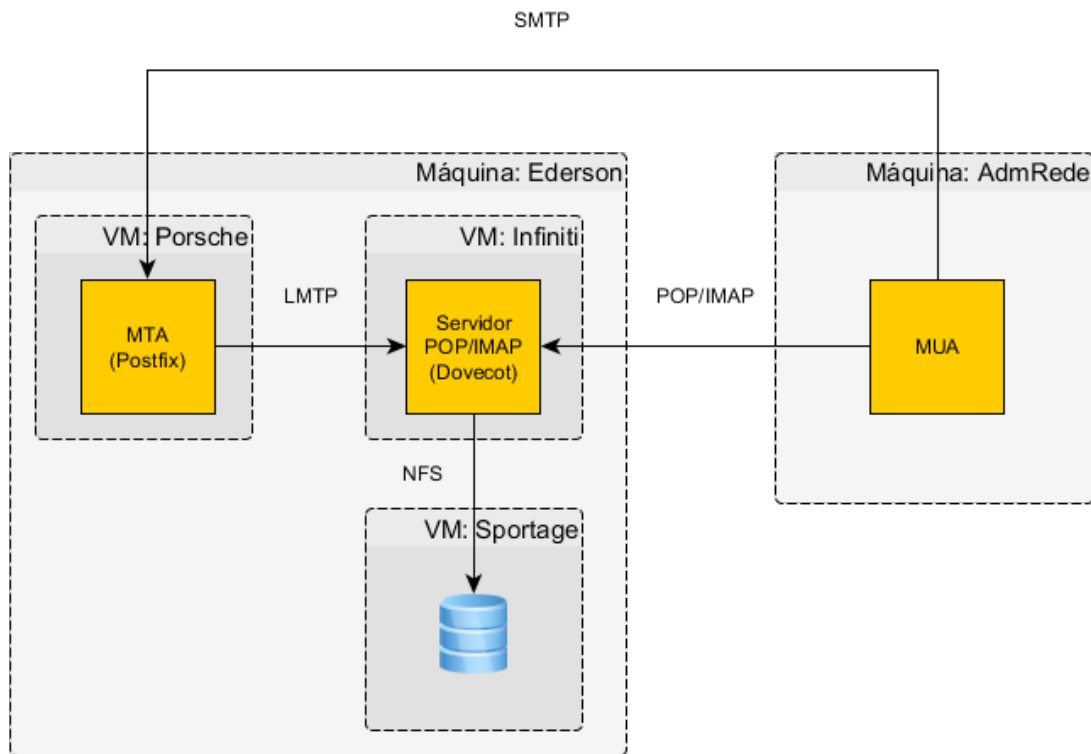


Figura 5.2: Arquitetura do Ambiente de Testes com LMTP Server.

5.1 Máquinas Físicas

Dois computadores foram utilizados para realizar os testes: a máquina Ederson e a máquina AdmRede. A máquina Ederson foi utilizada como servidor de virtualização, possibilitando que fossem instanciadas máquinas virtuais para simular as máquinas Porsche, Sportage e Infiniti, utilizadas na atual arquitetura. Para realizar sua função de máquina de virtualização, foi utilizado o programa Virtualbox. A máquina AdmRede foi utilizada para enviar mensagens destinadas a usuários em nosso serviço de e-mails. Para realizar esse envio, foi utilizado o programa *smtp-source*, que acompanha o pacote de instalação do Postfix. A especificação técnica dos equipamentos é apresentada na tabela 5.1.

5.2 Máquinas Virtuais

Todas as máquina virtuais possuem a mesma configuração de *hardware*, a exceção da máquina Sportage que possui maior espaço em disco. Foi optado por criar máquinas virtuais com as mesmas configurações para que ambos LDAs fossem executados em mesmas condições.

Na máquina Porsche foi realizada a instalação do Postfix. Ao instalar o Postfix, atualmente, é instalado também o Procmail como parte integrante do pacote da ferramenta. A máquina Infiniti é a responsável por executar o Dovecot. A máquina Sportage armazena os dados de todas as caixas postais dos usuários e realiza o compartilhamento desses dados com as outras máquinas virtuais das arquiteturas. As caixas postais estão no formato *maildir*. A especificação técnica das máquinas virtuais é apresentada na tabela 5.2.

	Ederson	AdmRede
Processador	Intel(R) Core(TM) i3-2100 CPU @ 3.10GHz	Intel(R) Core(TM)2 Duo CPU P8800 @ 2.66GHz
Memória	2x Kingston DDR3 1333MHz de 4096MB	2x DDR2 2048MB
Rede	Intel(R) 82579V Gigabit Network Device	Gigabit Ethernet
Disco Rígido	Western Digital(R) WD5000 - SATA 1.0 - 500GB	250GB HDD - SATA 1.0
Sistema Operacional	Ubuntu 13.10 - Saucy Salamander - 64bits	Ubuntu 12.04.4 LTS - Precise Pangolin
Programas	Virtualbox 4.2.16-dfsg-3	Postfix 2.9.6-1 12.04.1

Tabela 5.1: Máquinas Reais.

	Porsche	Infiniti	Sportage
Processador	100% - 1 core	100% - 1 core	100% - 1 core
Memória	1024MB	1024MB	1024MB
Rede	<i>Bridged Adapter</i>	<i>Bridged Adapter</i>	<i>Bridged Adapter</i>
Disco Rígido	8GB fixo	8GB fixo	15GB fixo
Sistema Operacional	Debian GNU/Linux 7.4 - Wheezy	Debian GNU/Linux 7.4 - Wheezy	Debian GNU/Linux 7.4 - Wheezy
Programas	Postfix 2.9.6-2 amd64	dovecot-core, dovecot-imapd, dovecot-lmtpd, dovecot- managesieved, dovecot-sieve 1:2.1.7-7 amd64	nfs-common, nfs-kernel-server 1:1.2.6-4 amd64

Tabela 5.2: Máquinas Virtuais.

6 AVALIAÇÃO

A avaliação será realizada baseando-se no comportamento das duas formas de entrega, Procmail e *LMTP Server*, e tem como objetivo apresentar, na prática, o impacto que cada um dos métodos utilizados causa nos tempos de entrega das mensagens e consumo dos recursos computacionais das máquinas envolvidas no processo.

6.1 Critérios de Avaliação

Durante a execução dos testes, serão coletadas algumas informações que permitirão realizar uma comparação direta entre os dois métodos de entrega de mensagens. Essa comparação permitirá avaliar que método oferece vantagens sobre o outro e em que condições. Para realizar essa avaliação, serão levados em consideração os seguintes critérios:

- **Tempo de Recebimento:** Esse dado exibirá o tempo que o MTA destinatário leva para receber todas as mensagens que estão sendo enviadas pelo MTA remetente. Se o MTA destinatário levar menos tempo para receber as mensagens, é um indício de que mais recursos (processador ou memória) ele teve a disposição para processá-las. Caso os recursos da máquina estejam sendo disputados por outros processos, a tendência é que esse recebimento seja mais lento.
- **Tempo de Depósito:** Essa informação exibirá o tempo que cada LDA leva para realizar o depósito de todas as mensagens nas caixas postais do usuários, levando em consideração o processamento do controle de duplicatas e, quando for o caso, da atualização dos arquivos de indexação. Com esse dado será possível comparar o atraso que um método pode ter em relação ao outro para realização da entrega das mensagens.
- **Uso de Memória:** Essa métrica apresentará o consumo de memória durante o recebimento das mensagens. Esse dado é importante pois permite identificar se há diferença, entre os métodos, na utilização desse recurso. Um consumo menor de memória possibilita que o equipamento utilize este recurso com outros processos que, possivelmente, estejam sendo executados em paralelo, como verificação anti-vírus, por exemplo. Caso algum dos métodos de entrega esgote esse recurso, é possível que o equipamento utilize o disco-rígido como memória, reduzindo consideravelmente o desempenho das tarefas sendo executadas.
- **Load Average:** Essa medição exibirá quantos processos ficaram aguardando - em média, no último minuto - pelo processador, em cada um dos métodos de entrega de mensagens. Assim como o uso da memória, a tendência é que o desempenho das

tarefas executadas fique prejudicado, caso o número de processos aguardando seja muito alto, uma vez que os processos passarão a competir pelo recurso. Foi optado por utilizar o *Load Average* ao invés do percentual de uso do processador, pois, a variação desse percentual é bastante brusca e, a captura desse dado em um determinado instante pode não corresponder ao uso real durante o intervalo de tempo entre as medições. Por outro lado, o *Load Average*, por levar em consideração a variação no último minuto, consegue exibir um comportamento mais constante, privilegiando a avaliação. Outro motivo, é que o uso do processador tende a ser de 100% em ambos os métodos de entrega, dificultando a identificação de, qual método, de fato, necessita de maior poder de processamento.

6.2 Metodologia

A partir da máquina AdmRede, será realizado o envio de diversas mensagens, utilizando o programa *smtp-source*, destinadas a usuários válidos de nosso serviço de e-mails. Essas mensagens devem ser processadas pelo MTA, entregues ao LDA e depositadas nas caixas postais dos usuários.

A obtenção do **tempo de recebimento** será realizada na máquina AdmRede, através do comando *time*. Esse comando inicia uma contagem imediatamente antes de começar o envio das mensagens e finaliza a contagem imediatamente após o envio da última. O valor obtido utilizando essa ferramenta exibe o tempo com precisão na ordem de milissegundos, que será arredondada para ordem de segundos.

O **tempo de depósito** será obtido na máquina que realizará o depósito diretamente nas caixas postais (máquinas Porsche ou Infiniti). Essa métrica será obtida através das mensagens de log armazenadas no arquivo */var/log/mail.log*. Para cada uma das entradas nesse arquivo de *log*, há o horário de entrega das mensagens com precisão na ordem de segundos. Assim, realizando a diferença do momento da última entrega com a primeira, é possível obter o tempo total da entrega das mensagens nas caixas postais dos usuários.

A cada dez segundos, o **uso de memória** será obtido, tanto na máquina Porsche quanto na máquina Infiniti, utilizando o comando *free*. É possível obter, através desse comando, o real consumo de memória realizado pelos processos, desconsiderando os *buffers* e *cache* que são utilizados pelo sistema operacional.

A partir do arquivo */proc/loadavg*, será extraída a carga média, no último minuto, do processador - o **Load Average**. Nesse arquivo, também são exibidas as médias dos últimos cinco e quinze minutos, porém, esses dados apresentam ruídos devido ao curto espaço de tempo entre cada teste executado, assim, a única informação que será utilizada é a média do último minuto. Essa informação será coletada a cada dez segundos.

6.3 Cenários de Testes

O volume de e-mails diários recebidos pelo servidor em produção do INF é de, em média, dezoito mil mensagens diárias. Além disso, o tamanho médio das mensagens armazenadas nas caixas postais dos usuários é de, aproximadamente, 100KB. Com base nessas informações, foram elaborados dezesseis cenários de teste para cada método de entrega de mensagens. Em cada um dos cenários, um total de dez mil e-mails serão enviados às caixas postais dos usuários. Serão variados os números de mensagens enviadas juntamente com o número de destinatários por mensagem, de modo a manter o total de dez mil mensagens a serem depositadas. Além disso, será variado, também, o tamanho

das mensagens enviadas. Os tamanhos de mensagens escolhidos para a realização dos testes foram de 1024B, 10240B, 102400B e 1024000B. Todos os dezesseis cenários estão resumidos na Tabela 6.1.

Tamanho das Mensagens(B)	Número de E-mails	Número de Destinatários
1024	10000	1
1024	1000	10
1024	100	100
1024	10	1000
10240	10000	1
10240	1000	10
10240	100	100
10240	10	1000
102400	10000	1
102400	1000	10
102400	100	100
102400	10	1000
1024000	10000	1
1024000	1000	10
1024000	100	100
1024000	10	1000

Tabela 6.1: Cenários de Teste.

Com esses cenários, é possível explorar as situações em que há múltiplas mensagens para um único usuário, mais usuários do que o limite de entrega simultânea permite e, também, casos intermediários, com múltiplos usuários e mensagens, explorando o paralelismo dos dois LDAs. Além disso, é possível identificar qual o impacto do tamanho das mensagens no consumo de recursos e nos tempos de recebimento e depósito das mensagens.

6.4 Resultados

Em sua configuração padrão, o Postfix possui uma característica de que, ao utilizar o Procmail como LDA, no máximo duas mensagens podem ser entregues simultaneamente para um mesmo usuário e que múltiplos processos (limitados a cem) podem ser instanciados simultaneamente para entregar mensagens a usuários distintos. Com o *LMTP Server*, o comportamento é diferente. O Postfix, em sua configuração padrão, realiza no máximo vinte conexões simultâneas com um mesmo servidor LMTP e, para cada conexão, é respeitado um limite de cinquenta destinatários. Esses comportamentos apresentam um grande impacto na análise dos resultados que são apresentados e discutidos a seguir.

6.4.1 Tempos de Recebimento e Depósito

Analisando os gráficos apresentados nas figuras 6.1, 6.2, 6.3 e 6.4, é possível identificar que o tempo de recebimento foi menor nos testes que utilizaram o *LMTP Server* como agente de entrega (LDA). Esse resultado é um indício de que, utilizando esse método, a máquina Porsche pode atender mais frequentemente as demandas do Postfix, agilizando o processamento das mensagens que estão sendo recebidas. A exceção desse comportamento foi apenas nos casos em que foram enviadas dez mil mensagens para um único usuário, o que mostra que, nessas situações, o Procmail obteve melhores resultados para esse critério de avaliação. Esse último comportamento é devido ao fato de que, em ambos os métodos de entrega, o Postfix recebe dez mil mensagens distintas. Após o recebimento, são realizadas dez mil retransmissões para o *LMTP Server* utilizando sempre dez threads simultâneas, enquanto, através do Procmail, apenas duas instâncias (por vez) do processo são criadas para realizar o depósito das mesmas dez mil mensagens.

Por outro lado, em todos os testes, o tempo de depósito das mensagens, utilizando o Procmail como LDA, foi mais rápido que os depósitos realizados utilizando o *LMTP Server*. Esse resultado se deve ao fato de o Procmail realizar o depósito das mensagens, utilizando o NFS, diretamente nas caixas postais dos usuários, a partir da máquina Porsche. Já, utilizando o *LMTP Server*, as mensagens inicialmente são transferidas pelo Postfix (através do protocolo LMTP) da máquina Porsche para a máquina Infiniti e, só então, é realizado o depósito das mensagens, também por NFS, pelo Dovecot, existindo, assim, um caminho maior percorrido pelas mensagens para chegar ao mesmo destino - as caixas postais. Além disso, existe mais o tempo do processamento dos arquivos de índices, por parte do *LMTP Server*, que também influencia no atraso no tempo de depósito. A atualização dos arquivos de índice, quando é utilizado o Procmail, é realizado apenas quando um usuário acessa sua caixa postal, processamento que não foi contabilizado durante a execução dos testes.

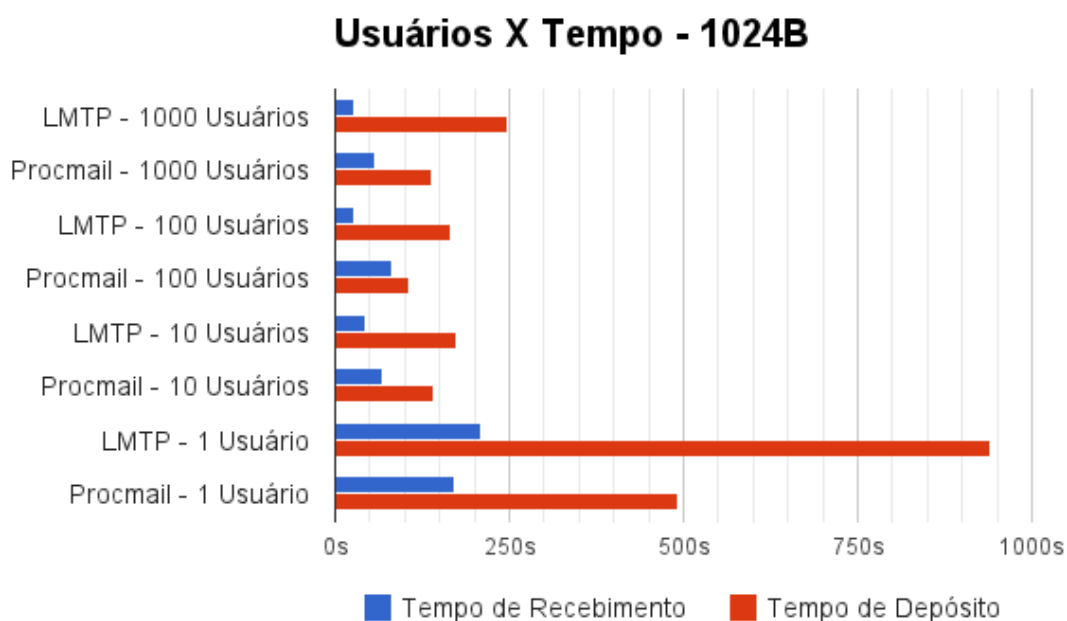


Figura 6.1: Tempos de Envio e Depósito para Mensagens de 1024B.

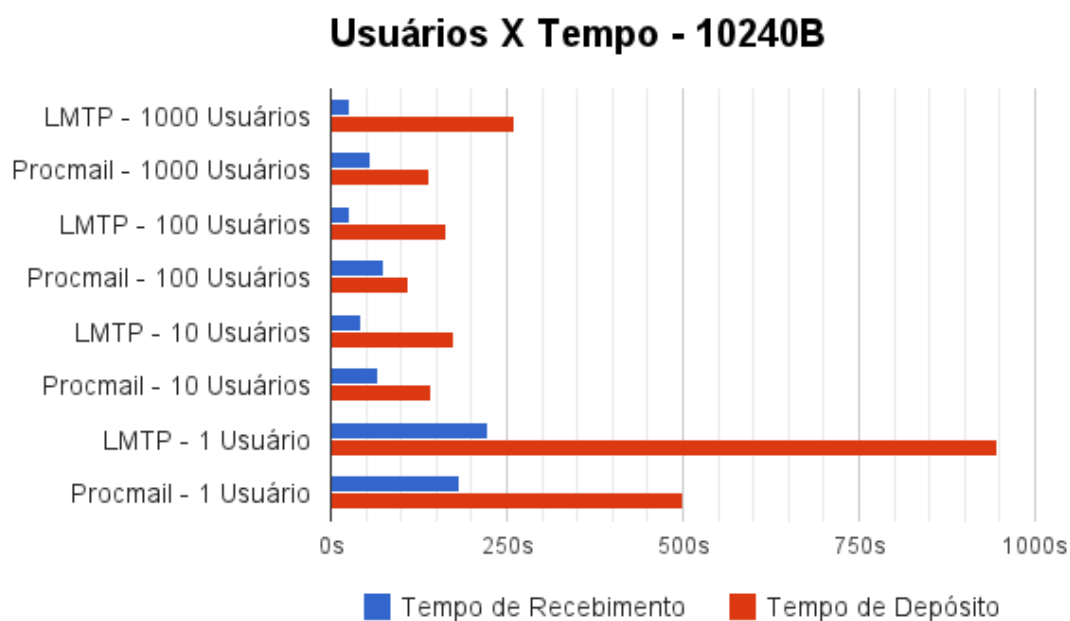


Figura 6.2: Tempos de Envio e Depósito para Mensagens de 10240B.

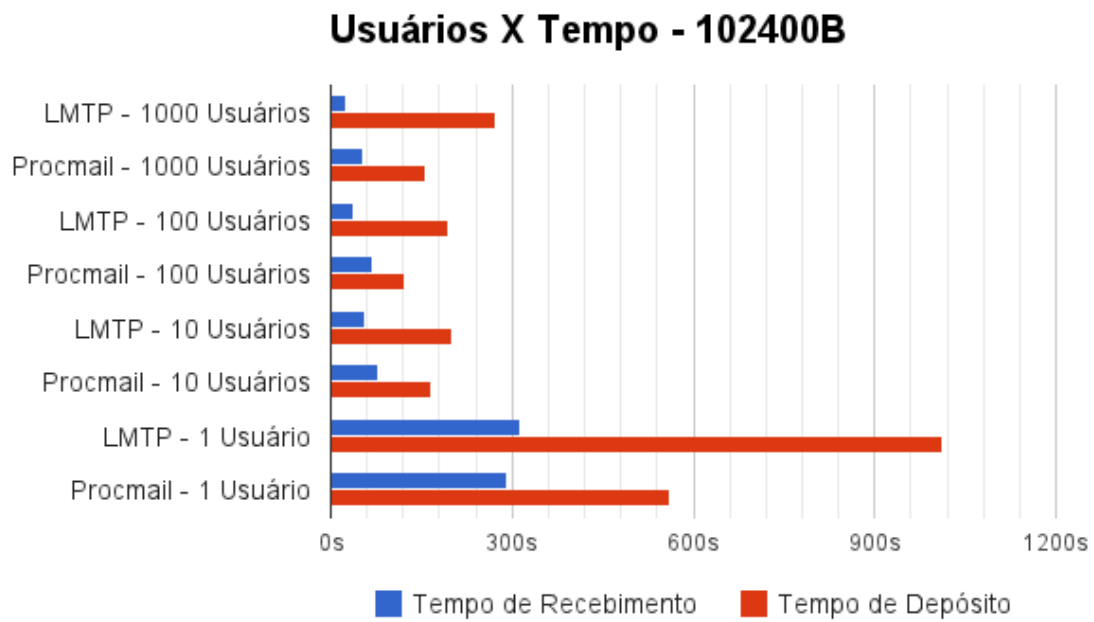


Figura 6.3: Tempos de Envio e Depósito para Mensagens de 102400B.

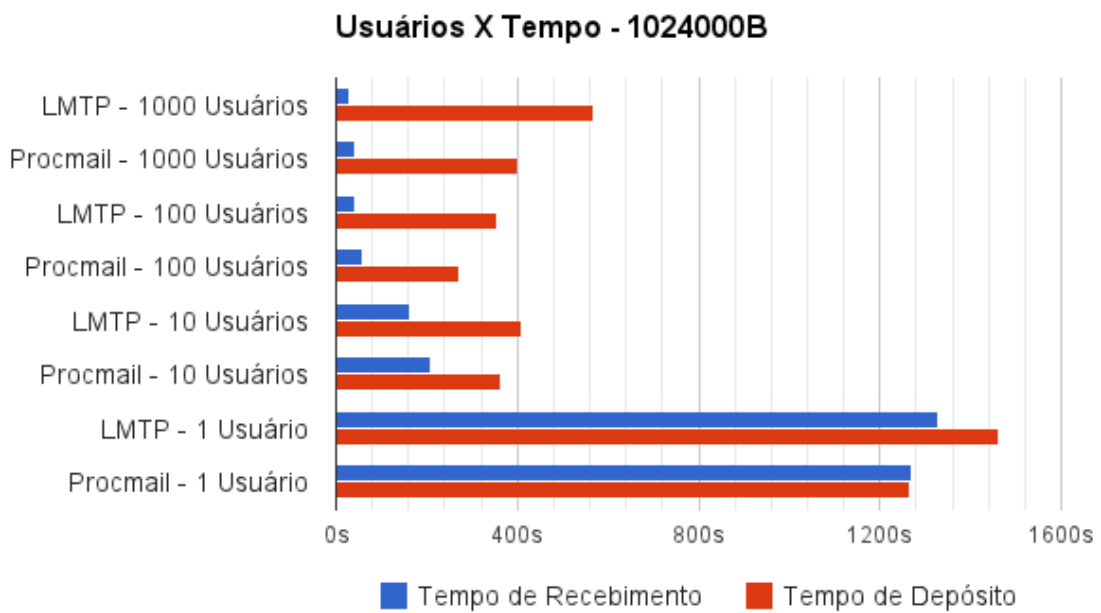


Figura 6.4: Tempos de Envio e Depósito para Mensagens de 1024000B.

6.4.2 Uso de Memória

Para todos os casos de entrega de mensagens através do Procmal, a máquina Infiniti não apresenta nenhuma variação no consumo de memória, se mantendo sempre com um uso médio de 70MB. Esse comportamento já era esperado, uma vez que, utilizando esse método de entrega de mensagens, a máquina Infiniti não tem nenhuma participação no depósito das mensagens, uma vez que o Procmal é executado apenas na máquina Porsche. O processamento dos arquivos de índice são realizados apenas durante o acesso dos usuários às caixas postais e esse processamento não foi contabilizado durante os testes.

Por outro lado, nas Figuras 6.5, 6.6, 6.7 e 6.8 é possível identificar que, durante a entrega utilizando o *LMTP Server*, o consumo de memória - da máquina Infiniti - tende a aumentar junto com o aumento do número de destinatários. Isso ocorre, pois, com o aumento do número de destinatários, aumenta o número de arquivos de índices atualizados e as informações de cache que são mantidas em memória.

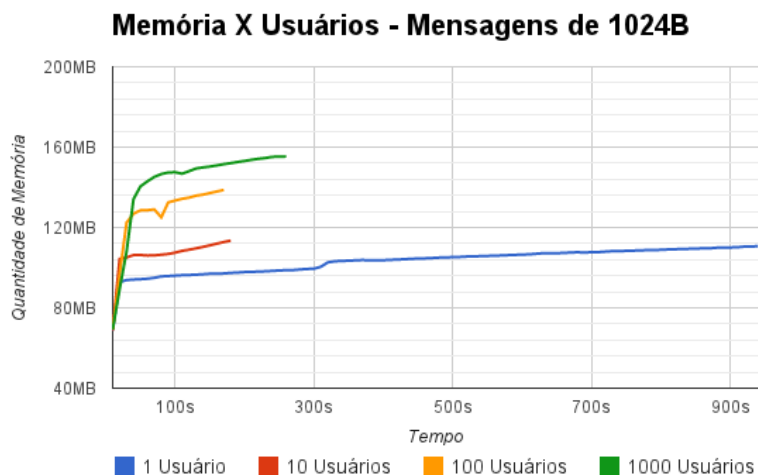


Figura 6.5: Uso de Memória da Máquina Infiniti com Mensagens de 1024B.

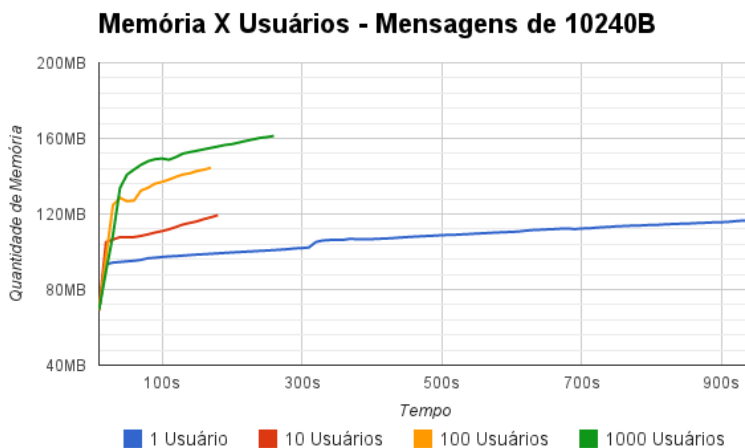


Figura 6.6: Uso de Memória da Máquina Infiniti com Mensagens de 10240B.

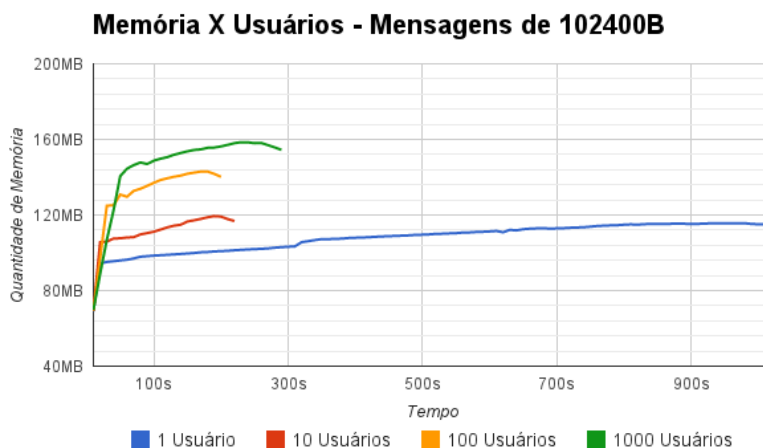


Figura 6.7: Uso de Memória da Máquina Infiniti com Mensagens de 102400B.

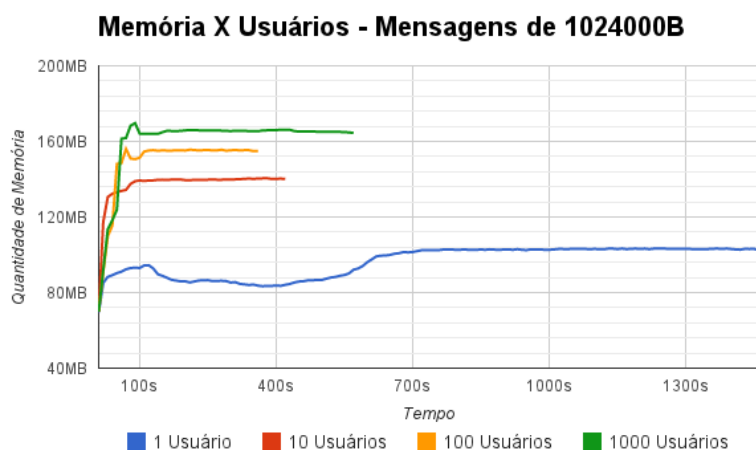


Figura 6.8: Uso de Memória da Máquina Infiniti com Mensagens de 1024000B.

As Figuras 6.9, 6.10 e 6.11 possibilitam realizar uma comparação entre o consumo de memória, na máquina Porsche, durante os testes utilizando as duas arquiteturas. É possível identificar que, para os mesmos testes, o consumo de memória utilizando o *LMTP Server* é inferior ao consumo realizado pelo Procmail. A exceção são os casos de dez mil mensagens enviadas para um único usuário, apresentado na Figura 6.12, pois, nessa situação, o Procmail instancia, no máximo, dois processos para realizar a entrega das mensagens, enquanto o outro método instancia vinte *threads* simultâneas para a transmissão das mensagens.

Na prática, com a adoção do *LMTP Server*, toda a memória que antes era consumida na máquina Porsche pelos processos do Procmail passou a ser utilizada na máquina Infiniti para atender as *threads* do *LMTP Server*. A vantagem de ser realizada a distribuição do consumo desse recurso nos servidores do INF, é que a máquina Porsche passa a ter mais memória para destinar aos seus demais processos, como o Amavis, o Mailman e o Postfix.

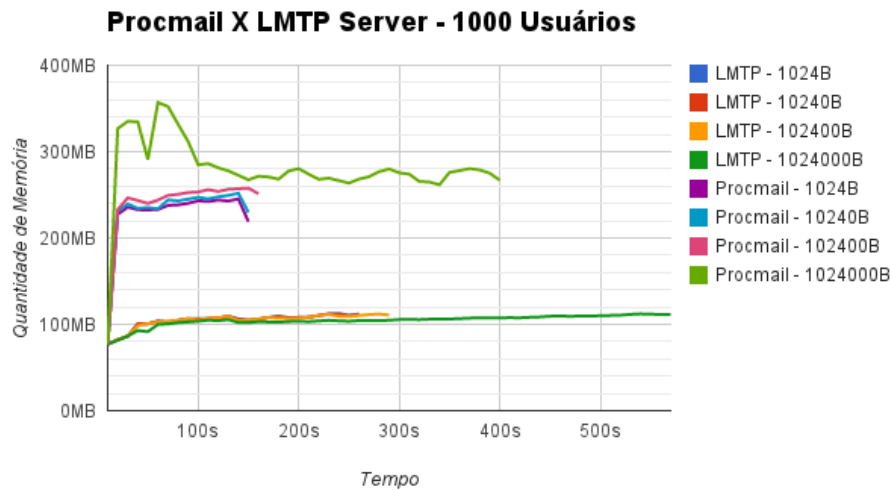


Figura 6.9: Uso de Memória da Máquina Porsche nos Cenários com 1000 Usuários.

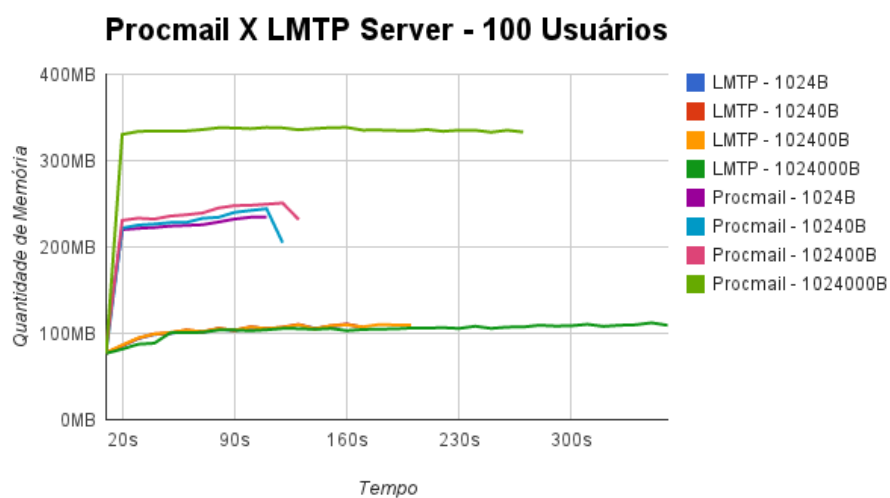


Figura 6.10: Uso de Memória da Máquina Porsche nos Cenários com 100 Usuários.

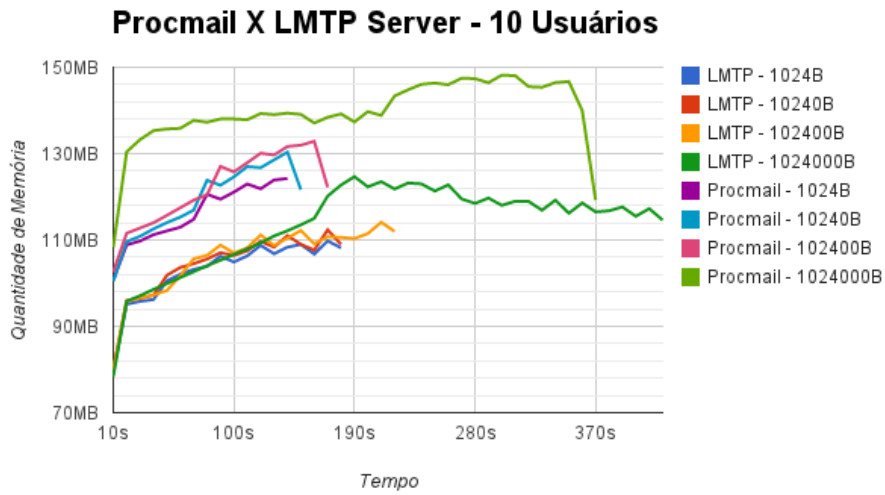


Figura 6.11: Uso de Memória da Máquina Porsche nos Cenários com 10 Usuários.

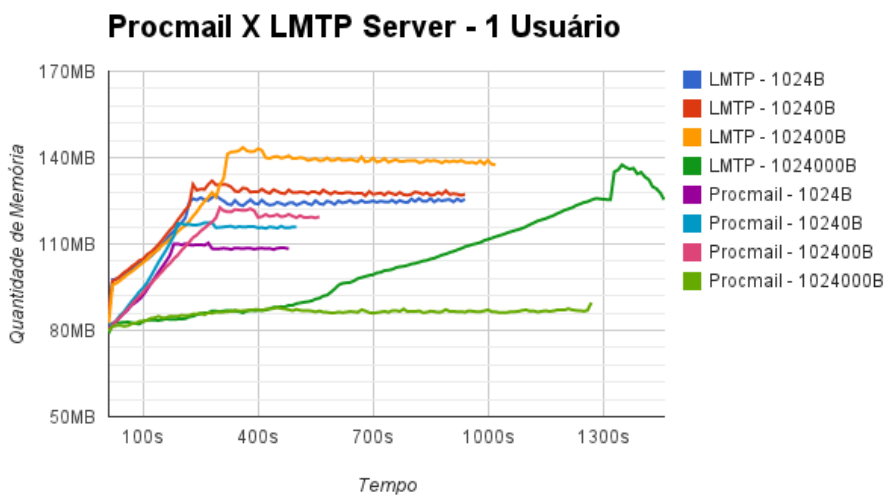


Figura 6.12: Uso de Memória da Máquina Porsche nos Cenários com 1 Usuário.

6.4.3 Load Average

Assim como no uso de memória, o *load average* não apresenta nenhuma variação na máquina Infiniti, durante a execução dos testes utilizando o Procmail como LDA, pois essa máquina não participa do depósito das mensagens nesse método.

Utilizando o *LMTP Server*, o *load average*, na máquina Infiniti, tende a aumentar com o aumento do número de usuários por mensagem, como pode ser visto nas Figuras 6.13, 6.14, 6.15 e 6.16. Como o Postfix realiza no máximo vinte comunicações simultâneas utilizando o protocolo LMTP, o *load average* chega no máximo a vinte - situação alcançada nos testes com mil e cem destinatários por mensagem. Para os testes com dez destinatários por mensagem o *load average* chega a dez e para o caso de um destinatário por mensagem, a métrica tende a se manter em dois.

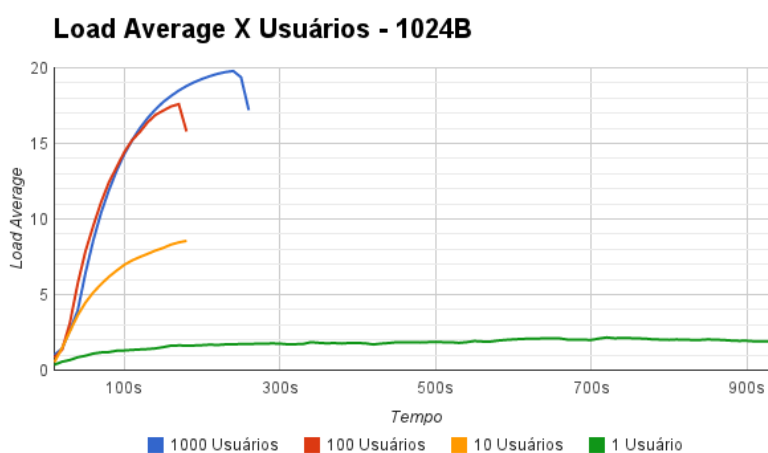


Figura 6.13: Load Average da Máquina Infiniti com Mensagens de 1024B.

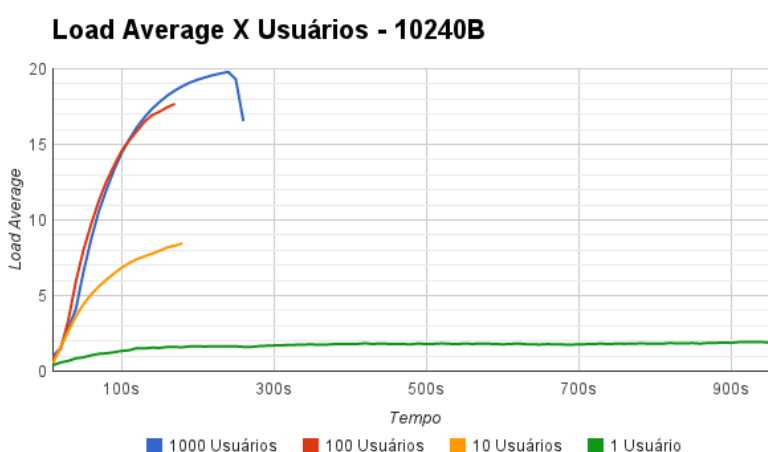


Figura 6.14: Load Average da Máquina Infiniti com Mensagens de 10240B.

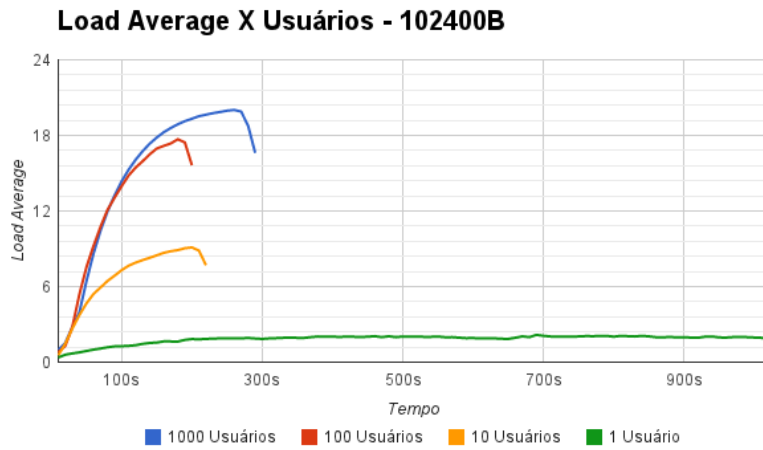


Figura 6.15: Load Average da Máquina Infiniti com Mensagens de 102400B.

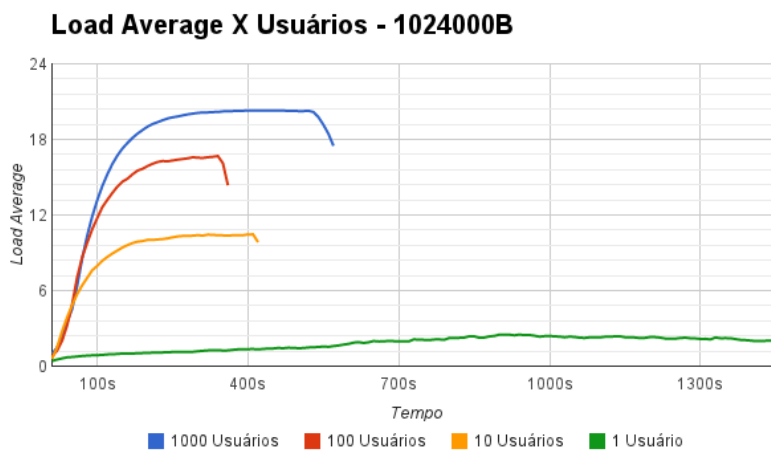


Figura 6.16: Load Average da Máquina Infiniti com Mensagens de 1024000B.

Na máquina Porsche é possível comparar os dois métodos e, em todos os casos, o *LMTP Server* apresenta uma melhora significativa nesse critério de avaliação, como pode ser visto nas Figuras 6.17, 6.18 e 6.19. Como o Procmail tende a instanciar, para cada mensagem, um processo para cada destinatário (limitado a cem), a concorrência pelo processamento é muito maior do que através do *LMTP Server*, que mantém sempre, no máximo, vinte threads simultâneas que são executadas rapidamente. No gráfico apresentado na Figura 6.20 é possível notar que, no caso em que dez mil mensagens são enviadas para um único usuário, o consumo do recurso, utilizando o Procmail, tende a se aproximar do outro método de entrega, pois, nessa situação, apenas dois processos são instanciados simultaneamente enquanto, utilizando o *LMTP Server*, se mantém as vinte *threads*.

Assim como no consumo de memória, ao se utilizar o *LMTP Server*, a carga de processamento no depósito das mensagens foi transferida da máquina Porsche para a máquina Infiniti, porém, com uma concorrência muito menor entre os processos, principalmente nos casos em que as mensagens foram enviadas para dez mil e mil usuários, onde o *Load Average* chega a cem, para o Procmail, e estabiliza em vinte, para o *LMTP Server*. Essa redução no *Load Average* certamente é um dos principais fatores que influenciaram a redução do tempo de recebimento das mensagens, uma vez que a máquina Porsche pode destinar mais processamento ao Postfix. No servidor em produção, isso pode beneficiar também outros processos, como o Amavis e o Mailman, agilizando assim a entrega dessas mensagens ao LDA.

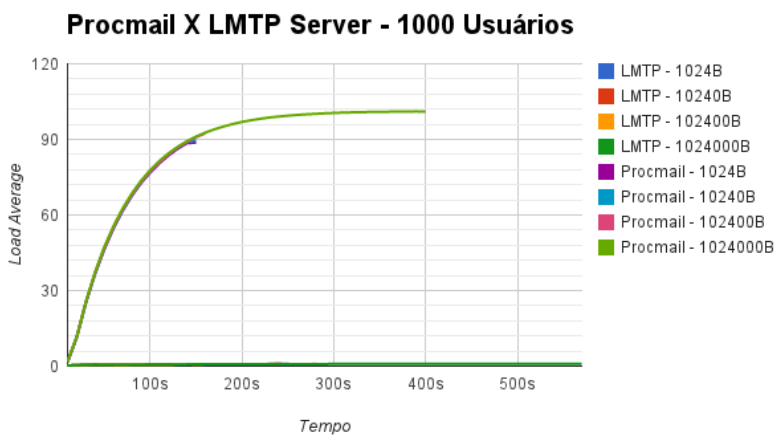


Figura 6.17: Load Average da Máquina Porsche nos Cenários com 1000 Usuários.

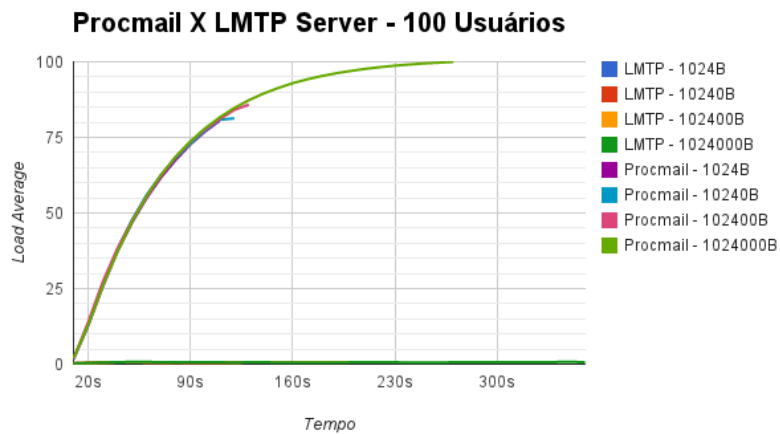


Figura 6.18: Load Average da Máquina Porsche nos Cenários com 100 Usuários.

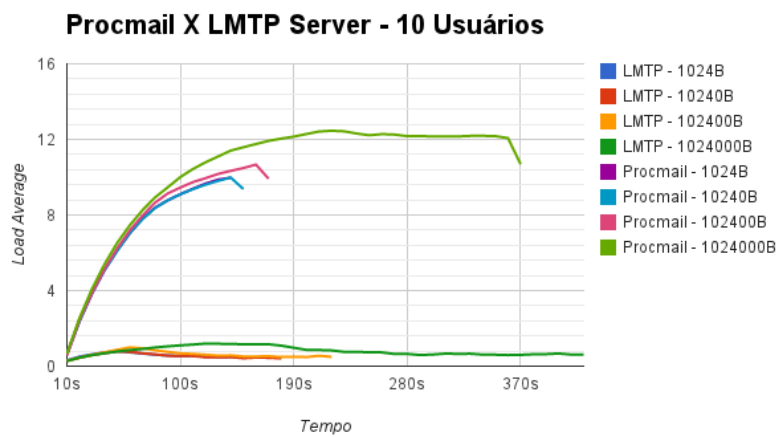


Figura 6.19: Load Average da Máquina Porsche nos Cenários com 10 Usuários.

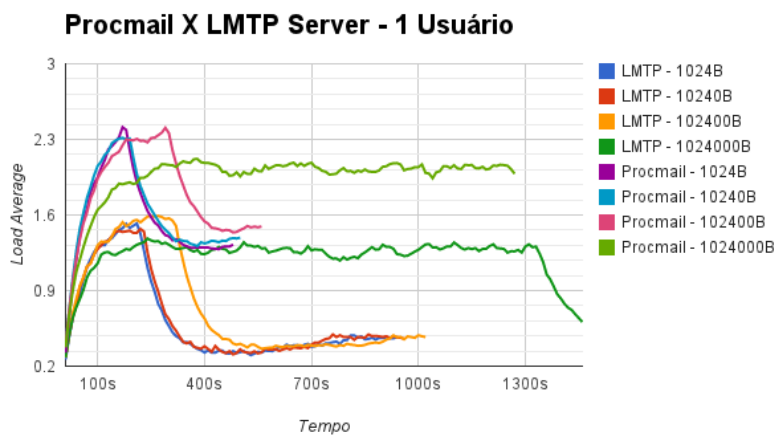


Figura 6.20: Load Average da Máquina Porsche nos Cenários com 1 Usuário.

7 CONCLUSÃO

Com base nos resultados obtidos no capítulo 6, é possível concluir que realizar a mudança do sistema de entrega de e-mails do Procmail para o Dovecot apresenta, sim, uma melhora na qualidade do serviço de e-mails do Instituto de Informática da UFRGS. Apesar do tempo de depósito das mensagens ser maior na arquitetura proposta, esse é um custo bastante pequeno, comparado a todos os benefícios oferecidos pela nova arquitetura. A distribuição do consumo de memória entre as máquinas envolvidas no serviço e a redução da carga de processamento na máquina Porsche são grandes avanços, pois, esses mesmos recursos, podem ser disponibilizados para os outros componentes do serviço de e-mails, como o Amavis, o Mailman e o Postfix e ainda permite que sejam alterados parâmetros nesses componentes de forma a remover os gargalos que eles apresentam na atual arquitetura.

Realizar os depósitos das mensagens através do Dovecot remove a concorrência pelo compartilhamento NFS e também evita que a ferramenta realize consultas, durante a sessão do usuário, para ver se foi feita alguma modificação nas caixas postais, reduzindo o tráfego de rede e os acessos NFS. Realizar o processamento dos arquivos de índice no momento da entrega das mensagens permite que os usuários tenham um acesso mais rápido às suas mensagens e abre a possibilidade para que sejam criadas regras personalizadas para a filtragem das mensagens em tempo de depósito, não sendo mais necessário realizar a mesma configuração em diferentes programas clientes de e-mail.

As melhorias oferecidas pela nova arquitetura não se limitam apenas ao que foi discutido neste trabalho. A adoção do Dovecot permite que novas análises sejam realizadas para que o serviço de e-mail seja ainda melhor, como, por exemplo: utilização dos formatos *dbx*; novas funcionalidades com os filtros Sieve, como, por exemplo, verificação anti-spam com treinamento de acordo com o retorno dos usuários; otimização de parâmetros de configuração do Postfix, como o número de conexões LMTP simultâneas e número de usuários por conexão.

Os resultados obtidos neste trabalho, em conjunto com as possibilidades de expansão do serviço de e-mails utilizando essa nova arquitetura resultou, em meados de junho de 2014, na implantação da arquitetura proposta nos servidores de e-mails em produção do Instituto de Informática da UFRGS. Essa arquitetura possibilitará a equipe de administração de redes agregar, futuramente, diversas funcionalidades que beneficiarão todos os usuários do Instituto, incentivando assim, o uso dos recursos oferecidos pela instituição de ensino para os assuntos acadêmicos.

REFERÊNCIAS

- BERG, S. R. van den. **Procmail**. Disponível em: <<http://www.procmail.org/>>. Acesso em: 08 mai. 2014.
- BERNSTEIN, D. J. **Using maildir format**. Disponível em: <<http://cr.yp.to/proto/maildir.html>>. Acesso em: 10 mai. 2014.
- CHOUDRIE, J.; DWIVEDI, Y. K. Broadband Impact on Household Consumers: online habits and time allocation patterns on daily life activities. **Int. J. Mob. Commun.**, Inderscience Publishers, Geneva, SWITZERLAND, v.5, n.2, p.225–241, Dec. 2007.
- CRISPIN, M. **Internet Message Access Protocol - Version 4**: RFC 1730. [S.l.]: Internet Engineering Task Force, Network Working Group, 1994.
- CRISPIN, M. **INTERNET MESSAGE ACCESS PROTOCOL - VERSION 4rev1**: RFC 3501. [S.l.]: Internet Engineering Task Force, Network Working Group, 2003.
- FUNK, R. **Mailbox Formats**. Disponível em: <<http://rfunk.livejournal.com/1571.html>>. Acesso em: 24 jun. 2014.
- HALL, E. **The application/mbox Media Type**: RFC 4155. [S.l.]: Internet Engineering Task Force, Network Working Group, 2005.
- INC., S. **Sendmail**. Disponível em: <<http://www.sendmail.com/>>. Acesso em: 08 mai. 2014.
- J. KLENSIN, E. **Simple Mail Transfer Protocol**: RFC 5321. [S.l.]: Internet Engineering Task Force, Network Working Group, 2001.
- MOZILLA. **Thunderbird - Programa feito para tornar emails simples**. Disponível em: <<http://www.mozilla.org/>>. Acesso em: 08 mai. 2014.
- MYERS, J. **Local Mail Transfer Protocol**: RFC 2033. [S.l.]: Internet Engineering Task Force, Network Working Group, 1996.
- MYERS, J.; ROSE, M. **Post Office Protocol - Version 3**: RFC 1939. [S.l.]: Internet Engineering Task Force, Network Working Group, 1996.
- P. GUENTHER, E.; SENDMAIL, I.; T. SHOWALTER, E. **Sieve**: an email filtering language: RFC 5228. [S.l.]: Internet Engineering Task Force, Network Working Group, 2008.

- POLLARD, J. d. B. **“mbox” is a family of several mutually incompatible mailbox formats.** Disponível em: <<http://homepage.ntlworld.com./jonathan.deboynepollard/FGA/mail-mbox-formats.html>>. Acesso em: 10 mai. 2014.
- QMAIL. **mbox - file containing mail messages.** Disponível em: <<http://qmail.org/qmail-manual-html/man5/mbox.html>>. Acesso em: 10 mai. 2014.
- SIRAINEN, T. **Dovecot.** Disponível em: <<http://www.dovecot.org/>>. Acesso em: 08 mai. 2014.
- SIRAINEN, T. **Dovecot’s index files.** Disponível em: <<http://wiki2.dovecot.org/Design/Indexes>>. Acesso em: 13 mai. 2014.
- SIRAINEN, T. **Cache file.** Disponível em: <<http://wiki2.dovecot.org/Design/Indexes/Cache>>. Acesso em: 13 mai. 2014.
- SIRAINEN, T. **Pigeonhole Sieve Interpreter.** Disponível em: <<http://wiki2.dovecot.org/Pigeonhole/Sieve>>. Acesso em: 15 mai. 2014.
- SIRAINEN, T. **Doveadm.** Disponível em: <<http://wiki2.dovecot.org/Tools/Doveadm>>. Acesso em: 06 jul. 2014.
- SIRAINEN, T. **Dsync.** Disponível em: <<http://wiki2.dovecot.org/Tools/Dsync>>. Acesso em: 06 jul. 2014.
- SIRAINEN, T. **Dovecot as a POP3 server.** Disponível em: <<http://wiki2.dovecot.org/POP3Server>>. Acesso em: 12 mai. 2014.
- SIRAINEN, T. **IMAP Server Compliancy Status.** Disponível em: <<http://imapwiki.org/ImapTest/ServerStatus>>. Acesso em: 12 mai. 2014.
- SUN MICROSYSTEMS, I. **NFS: network file system protocol specification: RFC 1094.** [S.l.]: Internet Engineering Task Force, Network Working Group, 1989.
- TEAM, T. R. **Roundcube Open Source Webmail Software.** Disponível em: <<http://roundcube.net/>>. Acesso em: 08 mai. 2014.
- THE RADICATI GROUP, I. **Email Statistics Report, 2014-2018.** Disponível em: <<http://www.radicati.com/wp/wp-content/uploads/2014/04/Email-Statistics-Report-2014-2018-Executive-Summary.pdf>>. Acesso em: 06 jul. 2014.
- VARSHAVCHIK, S. **Maildrop.** Disponível em: <<http://www.courier-mta.org/maildrop/>>. Acesso em: 08 mai. 2014.
- VARSHAVCHIK, S. **Courier Mail Server.** Disponível em: <<http://www.courier-mta.org/>>. Acesso em: 08 mai. 2014.
- VENEMA, W. **Postfix.** Disponível em: <<http://www.postfix.org/>>. Acesso em: 08 mai. 2014.
- WASHINGTON, U. of. **Alpine Messaging System.** Disponível em: <<https://www.washington.edu/alpine/>>. Acesso em: 08 mai. 2014.