

Segurança da Informação: Proposta de arquitetura de alta disponibilidade para aplicações web

Marco Antoni¹, Gláucio R. Vivian¹, Evandro Preuss¹

¹Colégio Agrícola de Frederico Westphalen (CAFW)

Universidade Federal de Santa Maria (UFSM)

98.400-000 – Frederico Westphalen – RS – Brasil

marquinho9.10@gmail.com, {glaucio, evandro}@cafw.ufsm.br

Abstract. *This short paper discusses the context of high availability within the information security in web applications presenting a proposal for development of an infrastructure cluster with load balancing. For development of same personal computers will be used, Linux operating system and software for high availability, server will be used to provide access to web pages, databases and distributed file system, and all free tools.*

Resumo. *Este artigo aborda o contexto da alta disponibilidade dentro da segurança da informação em aplicações para web apresentando uma proposta de desenvolvimento de uma infraestrutura de cluster com balanceamento de carga. Para desenvolvimento da mesma serão utilizados computadores pessoais, sistema operacional Linux e softwares de alta disponibilidade, servidor para prover ao acesso a páginas da web, banco de dados e sistema de arquivos distribuídos, sendo todas as ferramentas livres.*

1. Introdução

Devido ao constante crescimento no uso da *Internet* motivado pelo barateamento de equipamentos como computadores, *notebooks*, *smartphones*, etc e a conexão com a *Internet* tornando-se cada vez mais barata e rápida, as organizações estão cada vez mais utilizando a *web* como uma solução eficaz de comunicação, *marketing* e comércio eletrônico. Uma empresa que trabalha com comércio eletrônico pode ter inúmeros prejuízos se o seus serviços *web* ficarem indisponíveis por qualquer período de tempo, não importando a origem do problema, seja ele por ataques de negação de serviço, congestionamento na rede, falhas na infraestrutura do *datacenter* como *hardware* defeituoso, erros de *software* ou paradas de manutenção programada. Esses são problemas que além de prejudicar a imagem da empresa irão comprometer sua confiabilidade.

Segundo [LAUREANO 2012], a segurança da informação é um ponto chave em qualquer sistema, para o mesmo ser dito seguro ele deve prover três aspectos que são a integridade, confiabilidade e disponibilidade.

O objetivo desse trabalho é propor a criação de um ambiente totalmente tolerante a falhas como acontece em *datacenters* de grande porte, utilizando apenas computadores pessoais e *softwares open source* visando garantir a disponibilidade contínua de 99,999% do sistema, ou seja ter a garantia que o sistema poderá ficar inoperante por um período máximo de aproximadamente cinco minutos durante um ano.

2. Segurança da Informação

A tecnologia da informação vem se tornando cada vez mais utilizada pelo ser humano que busca sempre utilizá-la para agilizar as tarefas do cotidiano fazendo-as de maneira rápida e eficiente produzindo melhores resultados. A *web* é a tecnologia mais usada permitindo que haja comunicação entre todos os elementos permitindo hospedar *websites*, sistemas de gestão empresarial, etc. Esses recursos disponibilizados representam a era da informação.

No que tange a segurança da informação, segundo [LAUREANO 2012], temos três aspectos que são essenciais para o bom funcionamento das organizações que são: a) Integridade: Deve garantir que a informação esteja sempre em estado consistente não podendo ter valores perdidos ou alterados por qualquer entidade não autorizada pelo dono da mesma. b) Confiabilidade: Deve garantir o acesso à informação apenas ao dono e entidades previamente autorizadas. c) Disponibilidade: Deve garantir que a informação permaneça acessível por todos que tenham autorização para acessá-la independente de qualquer problema de nível técnico ou desastres naturais.

2.1. Ferramentas *Open Source* para Alta Disponibilidade

Algumas ferramentas *open source* disponíveis para implementar alta disponibilidade são: a) DRBD: É um *software* licenciado sob a licença GPL, atualmente é mantido pela empresa austríaca LINBIT HA-Solutions GmbH [HA-Solutions 2014]. Ele é usado para a construção de *clusters* de alta disponibilidade de armazenamento através da replicação de blocos de mais baixo nível do HD. b) Heartbeat: O [Heartbeat 2014] é um *daemon* disponível para Linux que fornece a infraestrutura de *cluster*, fazendo o gerenciamento do mesmo, permitindo fazer a verificação dos serviços do *cluster*. c) MySQL com Replicação: O MySQL tem suporte a replicação permitindo que dois ou mais servidores tenham conteúdos idênticos. A arquitetura de replicação pode ser *master/master* ou *master/slave*.

3. Trabalhos Relacionados

Nesta seção serão analisados alguns trabalhos relacionados com a proposta. No trabalho de [Caciato 2012] propôs-se uma arquitetura de duas camadas, composta pela camada de virtualização e camada de aplicação. Essa arquitetura consiste basicamente em um servidor principal e outro reserva. Quando ocorrer um erro no servidor físico, a camada de virtualização realizará a migração da máquina virtual para outro servidor. Falhas nos serviços do servidor principal farão a migração do serviço para o servidor reserva. A maior desvantagem dessa estratégia é a necessidade de duplicação do número de servidores e a ociosidade do servidor reserva.

Nos trabalhos de [Chi et al. 2012] e [Li et al. 2009] propõem-se arquiteturas para alta disponibilidade com distribuição de carga e utilização de *cache*. O aplicativo utilizado para tal foi o *software* [Nginx 2014] trabalhando como *proxy* reverso. Os resultados apresentados comprovaram o aumento de disponibilidade e o número de acessos simultâneos e evitam a ociosidade de servidores reserva.

A detecção de falhas de um serviço é implementada no trabalho de [Li et al. 2009] através do uso do *software heartbeat*.

Um problema encontrado em sistemas CMS é o grande uso do banco de dados para armazenamento de configurações e dos dados propriamente ditos. Nos trabalhos de [Li et al. 2009] e [Chi et al. 2012] não há preocupação com a replicação da base de

dados. Já em [Caciato 2012] utiliza-se uma instância do DRBD[HA-Solutions 2014], porem não ha detalhes sobre a implementação.

4. Proposta da Arquitetura

Com base nos trabalhos relatados, propõem-se uma arquitetura utilizando o Nginx como *Proxy* reverso para realizar a distribuição de carga entre dois servidores HTTP, replicando esse serviço em outra instancia gerenciada pelo [Heartbeat 2014]. Os outros dois servidores HTTP trabalham usando o Nginx, e usam o sistema de arquivos OCFS sobre o DRBD, permitindo que a escrita seja feito em ambos. O banco de dados Mysql[Dubois 2005] estará rodando em duas máquinas distintas, utilizando mecanismos próprios de replicação *Master/Master* para garantir a consistência das informações entre as duas base dados. O objetivo dessa configuração é garantir que a informação armazenada em uma base seja replicada para a outra base de dados. Finalmente pretende-se utilizar o *software heartbeat* para detectar falhas e redistribuir os serviços a fim de garantir a alta disponibilidade. Na figura a seguir pode-se visualizar a arquitetura proposta.

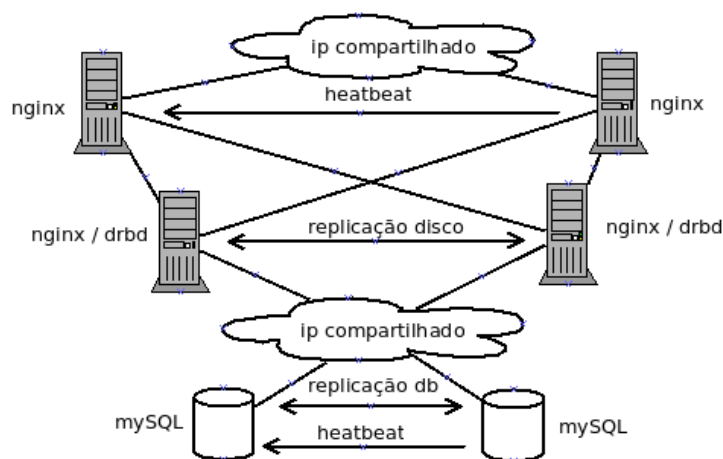


Figura 1: Proposta da arquitetura de alta disponibilidade

Esta proposta de arquitetura foi elaborada com base nos tecnologias apresentadas na secao 3. Procurou-se corrigir a ociosidade do servidor relatada no trabalho de [Caciato 2012] bem como melhorar a sua implementacao atravez do uso de outras solucoes utilizadas como o heartbeat, DRBD e a replicacao do banco de dados.

5. Resultados Esperados

Espera-se obter a comprovação da arquitetura proposta através da análise do aumento do número de acessos simultâneos bem como a obtenção de 99,999% de disponibilidade dos serviços. Para tanto, serão mensurados a disponibilidade do serviço na web por meio da ferramenta Sentinela Web[Web 2014] durante o período de um ano e comparada com um servidor simples. A análise do aumento do número de acessos simultâneos será realizada através do aplicativo *apache benchmarking tool*[Foundantion 2014], método também utilizado por [Caciato 2012] para avaliar a performance.

6. Considerações finais

Com base nos estudos das soluções já existentes e visando manter os três princípios da segurança da informação, principalmente a disponibilidade, propôs-se um modelo de arquitetura de alta disponibilidade visando corrigir as deficiências apontadas pelas soluções estudadas. Espera-se constatar a viabilidade da solução apresentada através do monitoramento da disponibilidade e testes de acesso simultâneos, comprovando a utilização do *Nginx* como *proxy* reverso, DRBD como replicador de sistema de arquivos e a replicação *master/master* do SGBD.

Referências

- [Caciato 2012] Caciato, L. E. (2012). High availability for critical services using open software and virtualization. In *Brazilian Symposium on Computing System Engineering*, pages 214–217. IEEE.
- [Chi et al. 2012] Chi, X., Liu, B., Niu, Q., and Wu, Q. (2012). Web load balance and cache optimization design based nginx under high-concurrency environment. In *Digital Manufacturing and Automation (ICDMA), 2012 Third International Conference on*, pages 1029–1032. IEEE.
- [Dubois 2005] Dubois, P. (2005). *MySQL : guia de estudo para certificação*. Ciência Moderna - Rio de Janeiro, 1ª edition.
- [Foundantion 2014] Foundantion, A. (2014). Apache http server project. Disponível em: <http://httpd.apache.org/>.
- [HA-Solutions 2014] HA-Solutions, L. (2014). Drbd. Disponível em: <http://drbd.org/>.
- [Heartbeat 2014] Heartbeat (2014). Disponível em: <http://linux-ha.org/wiki/Heartbeat/>.
- [LAUREANO 2012] LAUREANO, M. (2012). *Segurança da Informação*. Editora LT - Curitiba, 1ª edition.
- [Li et al. 2009] Li, F.-f., Xiang-zhan, Y., and Wu, G. (2009). Design and implementation of high availability distributed system based on multi-level heartbeat protocol. In *IITA International Conference on Control, Automation and System Engineering*, pages 83–87. IEEE.
- [Nginx 2014] Nginx (2014). Disponível em: <http://nginx.org/>.
- [Web 2014] Web, S. (2014). Disponível em: <http://www.sentinelaweb.com.br/>.