

GESTÃO DE ATIVOS

A organização nas mãos da TI



TENHA ACESSO ILIMITADO A TODO CONTEÚDO DO SITE DEVMEDIA

SEJA UM ASSINANTE **MVP!**

Tenha acesso garantido a qualquer tema
que você possa necessitar!



VIDEOS -AULAS

O assinante MVP acessa todas as vídeo-aulas do site DevMedia. Nenhum outro produto dá essa possibilidade. São mais de **7.000 vídeo-aulas**.



CURSOS

Nosso site possui atualmente **+ de 250 cursos online** com a qualidade que você já conhece e confia. Com o MVP você terá acesso a todos eles.



REVISTAS

Assinante MVP acessa todas as nossas revistas na versão digital, incluindo todas as edições anteriores. Quando pintar aquela dúvida você terá mais de **3.500 artigos para consultar!**



A assinatura MVP te dá a oportunidade de aumentar seu conhecimento em diversos assuntos, com qualidade e economia.

Invista na sua carreira se torne um profissional ainda mais completo!



DevMedia

POR APENAS
R\$59,90 POR MÊS

Acesse: www.devmedia.com.br/mvp

Sumário

Conteúdo sobre Novidades

06 – OpenFlow: uma rede definida por software

[André Koide da Silva]

15 – Gestão de Ativos – A organização nas mãos da TI

[Roberto Henrique]

Artigo no estilo Curso

26 – Migração de dados: metodologias e abordagens host-based – Parte 1

[Telmo Eduardo Nóbrega Reis]

Artigo no estilo Curso

37 – Monitorando em tempo real os elementos da TI – Parte 2

[Wagner Bianchi]

FÓRUM DEVMEDIA

O lugar perfeito para você ficar por dentro de tudo o que acontece nas tecnologias do mercado atual



No fórum da DevMedia você irá encontrar uma equipe disponível e altamente qualificada com consultores e colaboradores prontos para te ajudar a qualquer hora e sobre qualquer assunto. Temos as salas de Java, .NET, Delphi, Banco de Dados, Engenharia de Software, PHP, Java Script, Web Design, Automação comercial, Ruby on Rails e muito mais!

ACESSE AGORA
www.devmedia.com.br/forum



Edição

Editor

Eduardo Spínola (eduspinola@gmail.com)

Sub Editores

Marco Antônio Pereira Araújo (maraujo@devmedia.com.br)

Rodrigo Oliveira Spínola (rodrigo.devmedia@gmail.com)

Produção

Jornalista Responsável Kaline Dolabella - JP24185

Atendimento ao leitor

A DevMedia possui uma Central de Atendimento on-line, onde você pode tirar suas dúvidas sobre serviços, enviar críticas e sugestões e falar com um de nossos atendentes. Através da nossa central também é possível alterar dados cadastrais, consultar o status de assinaturas e conferir a data de envio de suas revistas. Acesse www.devmedia.com.br/central, ou se preferir entre em contato conosco através do telefone 21 3382-5038.

Publicidade

publicidade@devmedia.com.br – 21 3382-5038

Anúncios – Anunciando nas publicações e nos sites do Grupo DevMedia, você divulga sua marca ou produto para mais de 100 mil desenvolvedores de todo o Brasil, em mais de 200 cidades. Solicite nossos Media Kits, com detalhes sobre preços e formatos de anúncios.

Fale com o Editor!

É muito importante para a equipe saber o que você está achando da revista: que tipo de artigo você gostaria de ler, que artigo você mais gostou e qual artigo você menos gostou. Fique a vontade para entrar em contato com os editores e dar a sua sugestão!
Se você estiver interessado em publicar

um artigo na revista ou no site Easy Java Magazine, entre em contato com o editor, informando o título e mini-resumo do tema que você gostaria de publicar:

Eduardo Spínola - Editor da Revista
eduspinola@gmail.com

Editorial

Acada dia a gestão de ativos das empresas se torna uma tarefa mais atrelada às responsabilidades da Tecnologia da Informação, e ao invés de nos queixar, devemos nos preparar para acolher bem e aproveitar as novas oportunidades originadas a partir deste “novo” cenário. Pensando nisso, preparamos o artigo de capa desta edição, que além deste, destacará outros temas explicitados a seguir.

Em **Gestão de Ativos – A organização nas mãos da TI**, apresentaremos aos leitores todas as implicações que a falta de um sistema de gestão pode causar em uma organização e como a TI pode e deve participar diretamente neste tipo de trabalho, uma vez que é fundamental para redução de custos operacionais, mitigação de uma série de riscos à segurança e aplicação das melhores práticas de mercado.

OpenFlow: uma rede definida por software descreve um novo conceito utilizado para o controle da infraestrutura de dados: as redes definidas por software, também conhecidas como redes programáveis. Ao desvincular-se o desenvolvimento do hardware e do software dos equipamentos, há a possibilidade de conferir mais flexibilidade e agilidade na criação de novas arquiteturas e protocolos de comunicação. Este novo paradigma permite a utilização das inúmeras e conhecidas vantagens do software em uma área tradicionalmente dominada por hardwares proprietários.

No artigo **Migração de dados: metodologias e abordagens host-based – Parte 1** serão analisadas algumas estratégias e metodologias utilizadas para migração de dados corporativos de ambientes Enterprise Linux adotadas pelo mercado corporativo de TI, mais especificamente o segmento Datacenter, abordando desde o processo de classificação dos dados a serem migrados, até as estratégias que podem ser empregadas para movimentar os dados da forma mais segura e confortável, respeitando as necessidades e as diretrizes do seu negócio.

Para encerrar, em **Splunk: Monitorando em tempo real os elementos da TI – Parte 2**, colocaremos a mão na massa com um tutorial prático onde o leitor poderá explorar todo o potencial do Splunk para monitoramento da TI de uma empresa com base em dados de máquina. Extrair valor de logs e fazer com que a infraestrutura de uma empresa tenha menos problemas com downtime, por exemplo, são alguns dos pontos que fazem do Splunk uma plataforma única no mundo.

Eduardo Oliveira Spínola
eduspinola@gmail.com
twitter.com/Java_Magazine



Seu conhecimento vale muito na DevMedia

Publique artigos e vídeos gratuitamente no site e participe do programa **PayContent**.

Sua participação pode ser premiada: quanto mais você postar, mais chances de ganhar.



Para saber mais acesse:
www.devmedia.com.br/paycontent



OpenFlow: uma rede definida por software

Conheça esta tecnologia revolucionária de controle e inovação para a infraestrutura tradicional de redes de computadores

As atuais redes de computadores apresentam diversas demandas que precisam ser debatidas e tratadas, tais como: o crescimento das tabelas de roteamento, a complexidade de operação de muitos protocolos, o suporte à mobilidade dos usuários, a implementação de recursos de segurança, entre outras. A falta de ambientes que permitam a validação de novas arquiteturas e protocolos é uma das grandes dificuldades para a adoção de propostas que solucionam estas questões, pois não é possível reproduzir as mesmas características e volumes do tráfego real, além do comportamento de todos os equipamentos legados em operação.

Flexibilidade e agilidade, também são características que se fazem necessárias há algum tempo em equipamentos comuns, como *switches* e roteadores. Estes são formados por *hardwares* proprietários e engessados, os quais não possibilitam a personalização de seus recursos e funcionalidades.

As redes definidas por *software* (em inglês, *Software Defined Networking* – SDN) constituem um novo paradigma que propõe a desagregação entre o plano de dados (implementado em *hardware* especializado para suportar o desempenho requerido pelas redes atuais) e o plano de controle (executado em um ou mais servidores, os quais são responsáveis pela programação das ações realizadas pelo *hardware*).

Este artigo detalhará as necessidades que demandaram a concepção deste novo paradigma, além de apresentar os desafios atuais enfrentados pelas diferentes empresas do setor de telecomunicações para manutenção e operação de suas infraestruturas. A seguir, serão descritos os três principais componentes da arquitetura OpenFlow: as tabelas de fluxos, o canal seguro de comunicação entre

Resumo DevMan

OpenFlow: uma rede definida por software:

Este artigo descreve um novo conceito utilizado para o controle da infraestrutura de dados: as redes definidas por software, também conhecidas como redes programáveis. Ao desvincular-se o desenvolvimento do hardware e do software dos equipamentos, há a possibilidade de conferir mais flexibilidade e agilidade na criação de novas arquiteturas e protocolos de comunicação. Este novo paradigma permite a utilização das inúmeras e conhecidas vantagens do software em uma área tradicionalmente dominada por hardwares proprietários.

Em que situação o tema é útil:

O controle das redes de computadores por meio de softwares é ainda um tema incipiente e pouco difundido. Tornou-se mais conhecido após a publicação da tecnologia OpenFlow em 2008. Desde então, um número crescente de artigos e pesquisas foram desenvolvidos pelos fabricantes de hardware e pelos pesquisadores do meio acadêmico. Neste contexto, observou-se a importância da disseminação deste novo assunto entre os profissionais da área de tecnologia da informação, permitindo a contenda e o aprofundamento técnico do tema.

o plano de dados e o de controle, e o protocolo OpenFlow. Também serão abordadas as diferentes informações disponíveis para caracterização de um fluxo. A importância das redes definidas por *software* será avaliada mediante a análise de um cenário relativamente simples, mas que impõe restrições que dificultam sua resolução por intermédio dos dispositivos e protocolos padrões (não proprietários) atualmente disponíveis. Por fim, serão apresentadas as mudanças e a evolução do OpenFlow especificadas em sua nova versão, além de sua adoção pelos fabricantes de *hardware*, provedores de serviços e operadoras de telecomunicações.

Por que as redes definidas por software são necessárias?

As redes se originaram a partir da necessidade da troca de dados entre as pessoas e os computadores, possibilitando a integração entre os sistemas instalados em diferentes localidades. Atualmente, as tecnologias que disponibilizam esta infraestrutura de conectividade estão inseridas em praticamente todas as atividades diárias da sociedade contemporânea. Estão presentes nas páginas *web* que viabilizam o comércio eletrônico, nos sistemas de *internet banking* ofertados pelas instituições financeiras, nas plataformas de *e-mail*, na educação à distância, nos aplicativos de navegação e roteirização baseados em GPS (*Global Positioning System*), e, mais recentemente, em dispositivos domésticos, como *smartphones* e televisores. Diversas iniciativas de inclusão digital são desenvolvidas por órgãos governamentais e privados com o objetivo de expandir ainda mais seu alcance para toda a população. Segundo o artigo *Hobbes' Internet Timeline*, o número de dispositivos conectados à Internet, rede mundial de computadores, brevemente será superior a um bilhão de equipamentos.

As pessoas, as corporações e os serviços públicos estão totalmente dependentes desta infraestrutura de comunicação. Você já imaginou como seria um dia sem acesso à Internet? Esta importância gera um problema para o desenvolvimento de novas tecnologias e protocolos de comunicação, pois as pesquisas não podem ser realizadas diretamente no ambiente de produção (aquele disponível e utilizado pelos usuários) devido ao risco de interrupção dos serviços essenciais. Além disso, a extensa adoção das tecnologias atuais e sua economia de escala inviabilizam a criação de novos recursos e funcionalidades que carecem de *hardwares* revolucionários. Em *Redes definidas por software: uma abordagem sistêmica para o desenvolvimento de pesquisas em redes de computadores*, os autores citam que diversos pesquisadores consideram que a arquitetura das redes de computadores atingiu um nível de amadurecimento tão alto que a tornou pouco flexível, ou seja, a Internet está calcificada (em inglês, *ossified*), uma analogia ao processo que substitui as cartilagens (mais elásticas) por ossos conforme o envelhecimento dos seres vivos.

O *hardware* atualmente empregado na infraestrutura das redes é basicamente formado por dispositivos proprietários e com alto custo de aquisição. Estes possuem sua arquitetura (plano de dados) concebida a partir de ASICs (*Application-Specific Integrated Circuits*) – circuitos integrados dedicados que asseguram o alto desempenho requerido por *switches*, roteadores, *firewalls*, balanceadores de carga, entre outros. Suportam ainda uma camada complementar de *software* (plano de controle) que programa os diferentes protocolos de redes. A comunicação entre o plano de controle e o de dados ocorre por meio de APIs (*Application Programming Interfaces*) proprietárias (Figura 1).

Desta forma, a implementação de novos protocolos, a otimização nas lógicas de controle e o desenvolvimento de novos recursos, entre outras melhorias no plano de controle, estão restritas ao fabricante do *hardware*, resultando em um processo moroso e com altos custos. Este aspecto se antepõe ao crescimento exponencial do tráfego observado na Internet, bem como às tendências de

virtualização dos servidores. As arquiteturas clássicas das redes de computadores sugerem restrições às aplicações e à inovação necessária aos serviços demandados pelos negócios e usuários da era da informação. Os desafios para o setor de telecomunicações são diversos:

- As operadoras de telefonia móvel precisam lidar com o congestionamento do espectro eletromagnético (propagador dos sinais de comunicação), a transição para as tecnologias IP (Internet Protocol) e o célere crescimento de seus usuários;
- O tráfego IP global atingiu um volume sem precedentes que necessita ser suportado pelos operadores de telecomunicações;
- O número de servidores e máquinas virtuais hospedados em Data Centers também está incrementando, exigindo infraestruturas internas com maior capacidade de comutação;
- Os pesquisadores da área de tecnologia da informação tratam, armazenam e transportam volumes imensos de dados.

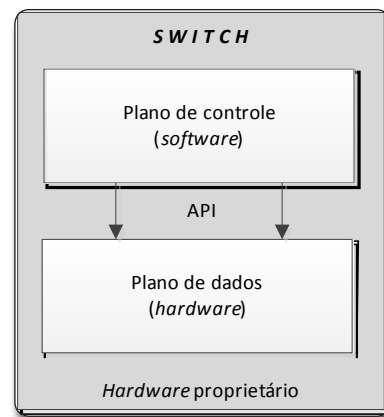


Figura 1. Hardware de um switch típico (proprietário)

Para atender a estes desafios, foi criado um novo paradigma: as redes definidas por *software* (também denominadas redes programáveis). Por meio desta arquitetura, a eficiência, a flexibilidade, a agilidade e a escalabilidade dos dispositivos de conectividade serão ampliados. As vantagens no uso dos *softwares*, já demonstradas nas diversas tecnologias existentes, foram estendidas para as redes de computadores. Segundo Marc Andreessen, co-fundador da Netscape, “*Software is eating the world*”, ou seja, as aplicações baseadas no modelo de operação da Internet serão amplamente utilizadas para transformar todos os setores da indústria.

Redes definidas por software

Estas constituem um novo padrão que propõe a desagregação do plano de controle e de dados habilitando a personalização em massa da infraestrutura para suportar serviços com diferentes características. Segundo o artigo *OpenFlow e redes definidas por software: um novo paradigma de controle e inovação em redes de pacotes*, os avanços na padronização de APIs independentes de fabricantes permitiram que a lógica de controle fosse movida para controladores externos, que podem ser disponibilizados por meio de servidores comuns e difundidos comercialmente. Esta separação

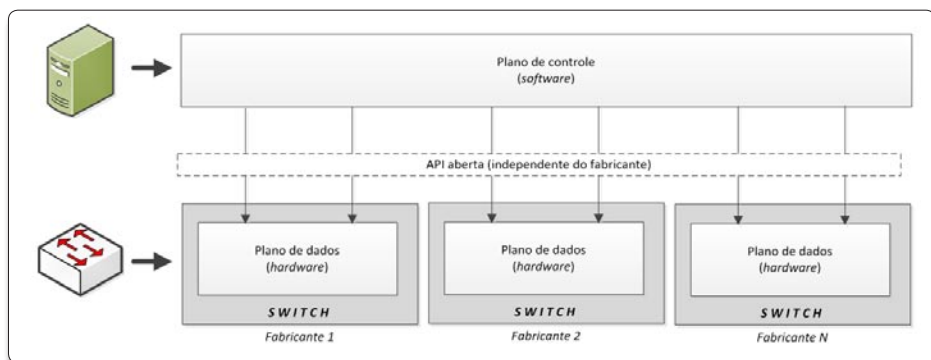


Figura 2. Redes definidas por software

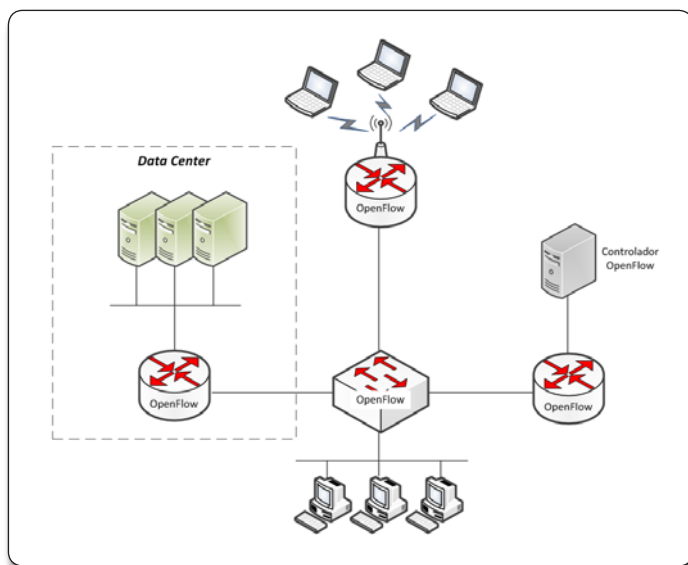


Figura 3. Exemplo de dispositivos comerciais controlados pelo OpenFlow

possibilita que o comportamento da rede definida por *software* seja determinado por seus próprios usuários e fornecedores.

O plano de controle, instalado em servidores dedicados e com alta capacidade de processamento, é responsável pela programação remota dos dispositivos de conectividade por meio de uma API aberta (não proprietária). Assim, as principais vantagens do *hardware* e do *software* são extraídas, respectivamente: desempenho para o processamento e o encaminhamento dos pacotes, e flexibilidade para inserção, remoção e especialização de suas funcionalidades (Figura 2). Neste contexto surgiu o OpenFlow, tecnologia que será abordada na próxima seção.

OpenFlow

Publicado em 2008, o OpenFlow visa suportar novas propostas de protocolos e

arquiteturas de rede sobre equipamentos comercialmente disponíveis. Descreve uma tecnologia para controlar as possíveis opções de encaminhamento de pacotes em *switches* e roteadores, entre outros dispositivos. A inteligência para a tomada destas decisões é responsabilidade de um elemento externo, chamado controlador OpenFlow, o qual pode ser instalado em servidores comuns (Figura 3).

Em *OpenFlow: Enabling Innovation in Campus Networks*, os autores afirmam que o conceito básico deste paradigma baseia-se no fato de que a maioria dos *switches* e roteadores ethernet atuais empregam tabelas de fluxos para implementar as restrições de segurança (*Access Control List – ACL*), as traduções de endereços de rede (*Network Address Translation – NAT*), as regras de qualidade de serviço (*Quality of Service – QoS*) e para coletar

estatísticas. Embora existam diferenças entre as dados armazenados (dependendo do fabricante), é possível identificar diversos campos comuns. Esta característica possibilita a especificação de um protocolo genérico para programar as tabelas de fluxos dos dispositivos de distintos fabricantes. Os fluxos são definidos pelos campos de cabeçalho do pacote processado, podendo incluir informações das camadas de enlace, rede e transporte conforme o modelo TCP/IP (*Transmission Control Protocol/Internet Protocol* – conjunto de protocolos que determinam o padrão de comunicação na Internet). A Figura 4 ilustra os campos de cabeçalho suportados pela versão inicial do OpenFlow (versão 1.0.0):

- **Porta de entrada (*ingress port*):** porta pela qual os pacotes foram recebidos;
- **Identificação da VLAN (*virtual local area network identification*):** número que possibilita a associação do pacote a sua rede local virtual;
- **Endereço MAC de origem (*source media access control address*):** também conhecido como endereço físico, indica o endereço do adaptador de rede do originador dos dados;
- **Endereço MAC de destino (*destination media access control address*):** identificador do adaptador de rede do destinatário;
- **Tipo ethernet (*ethernet type*):** especifica o tipo de pacote da camada de rede que é transportado;
- **Endereço IP de origem (*internet protocol source address*):** endereço lógico do originador dos dados;
- **Endereço IP de destino (*internet protocol destination address*):** define o endereço IP do receptor;
- **Protocolo (*protocol*):** identifica o protocolo da camada imediatamente superior (transporte);
- **Porta TCP/UDP de origem (*source transmission control protocol/user data-gram protocol port*):** indica a porta TCP/UDP do transmissor do pacote;
- **Porta TCP/UDP de destino (*destination transmission control protocol/user datagram protocol port*):** especifica a porta TCP/UDP do destinatário.

Porta de entrada	Identificação da VLAN	Ethernet			Internet Protocol			TCP/UDP	
		MAC origem	MAC destino	Tipo ethernet	IP origem	IP destino	Protocolo	Porta de origem	Porta de destino

Figura 4. Campos suportados pela versão inicial do OpenFlow (versão 1.0.0)

Em linhas gerais, a arquitetura OpenFlow é composta por três componentes (Figura 5):

- **Tabela de fluxos:** quando um pacote é recebido pelo dispositivo de rede, este é comparado com cada um dos fluxos registrados na tabela; caso encontre uma correspondência, considerar-se-á que o pacote pertence ao fluxo e aplicar-se-á as ações relacionadas. A seguir, serão atualizadas as estatísticas que controlam o número de pacotes, a quantidade de bytes e a duração daquele fluxo (Figura 6). Se um fluxo relacionado não for encontrado, o pacote será encaminhado para processamento no controlador OpenFlow;

- **Canal seguro:** assegura, por intermédio do protocolo SSL (*Secure Socket Layer*), que a troca de informações entre os equipamentos de rede e o comutador OpenFlow não tenha sua integridade comprometida. Para tanto, o SSL possibilita que todos os dados sejam criptografados;

- **Protocolo OpenFlow (OpenFlow protocol – OPF):** disponibilizado para estabelecer a comunicação entre os switches e roteadores, e o controlador. Por meio deste, são programadas as decisões tomadas nas tabelas de fluxos.

Dependendo do tipo de suporte ao OpenFlow, os dispositivos de rede podem ser agrupados em duas categorias:

- **Switches OpenFlow dedicados:** equipamentos que não implementam o processamento dos protocolos clássicos da camada de enlace e de rede;
- **Switches com suporte OpenFlow:** switches e roteadores ethernet comercialmente disponíveis que disponibilizam esta nova funcionalidade.

Os switches OpenFlow dedicados são comutadores que encaminham os pacotes para as portas somente com a especificação de um plano de controle remoto. Neste contexto, os fluxos podem ser definidos

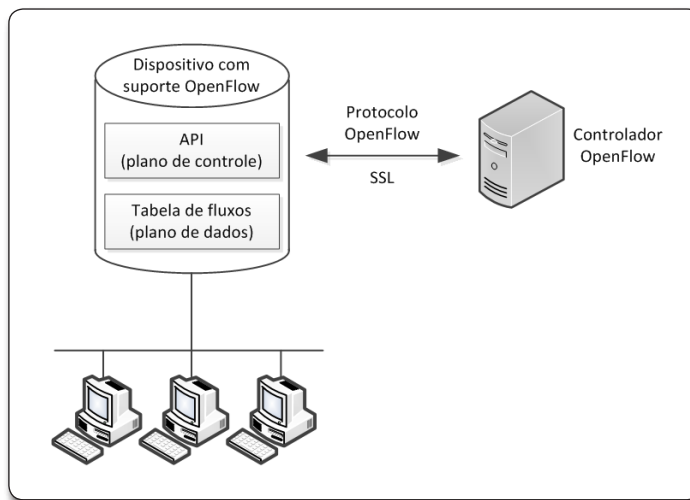


Figura 5. Componentes da arquitetura OpenFlow

Fluxos	Ações	Estatísticas
--------	-------	--------------

Figura 6. Estrutura da tabela de fluxos

pelo endereço MAC de origem, pelos endereços IP de origem e destino, pela porta TCP de destino, entre outras combinações possíveis para os campos de cabeçalho suportados (Figura 7).

Cada registro na tabela de fluxos possui uma ação associada. Para os switches OpenFlow dedicados, foram determinadas três opções:

1. Encaminhar os dados do fluxo para uma ou mais portas, permitindo que sejam propagados pela rede;
2. Encapsular e encaminhar o pacote para o controlador por intermédio do canal seguro (SSL). Geralmente, esta ação é tomada para o primeiro pacote dos fluxos desconhecidos. Desta maneira, após o processamento pelo controlador OpenFlow, uma ação será definida, e os dados registrados em sua tabela. Em situações específicas, todos os pacotes de certos fluxos podem ser tratados individualmente pelo controlador: por exemplo, aqueles gerados por protocolos de roteamento, como OSPF (*Open Shortest Path First*) e BGP (*Border*

Gateway Protocol), ou ainda, relacionados às funções de controle, como ICMP (*Internet Control Message Protocol*) e DHCP (*Dynamic Host Configuration Protocol*), entre outros;

3. Descartar os pacotes do fluxo. Restrições de segurança, mitigação dos ataques de negação de serviços ou redução no volume de tráfego broadcast são possíveis razões para esta decisão.

Com a popularização do OpenFlow, os switches, os roteadores e os pontos de acesso wireless (*access points*) comercialmente disponíveis passaram a disponibilizar os componentes desta nova arquitetura. Estes equipamentos, denominados switches com suporte OpenFlow, implementam uma quarta ação aos fluxos:

4. Encaminhar os pacotes para o processamento padrão das camadas de rede e enlace.

Ao criar esta camada de abstração da infraestrutura física, o controlador OpenFlow possibilita o desenvolvimento de

Porta de entrada *	Identificação da VLAN *	Ethernet			Internet Protocol			TCP/UDP	
		MAC origem 00:01:1a:ca:45:ca	MAC destino *	Tipo ethernet *	IP origem *	IP destino *	Protocolo *	Porta de origem *	Porta de destino *
Porta de entrada *	Identificação da VLAN *	MAC origem *	MAC destino *	Tipo ethernet *	IP origem 189.72.2.2	IP destino 200.2.2.1	Protocolo *	Porta de origem *	Porta de destino *
Porta de entrada *	Identificação da VLAN *	MAC origem *	MAC destino *	Tipo ethernet *	IP origem *	IP destino *	Protocolo *	Porta de origem *	Porta de destino 22

Figura 7. Exemplos de registros de fluxos

serviços e aplicações para gerenciar a tabela de fluxos dos dispositivos de rede. Segundo os autores do artigo *OpenFlow e redes definidas por software: um novo paradigma de controle e inovação em redes de pacotes*, este princípio de operação é similar ao que já existe em outros sistemas de *software* que suportam recursos reutilizáveis e abstração do *hardware*. A separação do plano de controle desvincula sua evolução do plano de dados, tornando possíveis frentes paralelas e independentes de melhoria.

Aplicação do OpenFlow

Conforme exposto nas seções anteriores, a infraestrutura de redes atual é pouco flexível, pois qualquer projeto deve seguir as especificações suportadas pelos fabricantes. Os recursos disponíveis são engessados pelos protocolos padrões existentes (internet standards), tais como:

- Camada de enlace:
 - Spanning tree protocol;
 - Multiple spanning tree protocol;
 - Rapid spanning tree protocol;
 - Virtual local area network;
 - Queue-in-queue encapsulation;
 - Entre outros.
- Camada de redes:
 - Virtual router redundancy protocol;
 - Virtual private LAN service;
 - Open Shortest Path First;
 - Border Gateway Protocol;
 - Etc.

Portanto, os engenheiros de redes precisam selecionar entre as peças existentes, aquelas que se encaixam para propor as soluções para seus problemas. Por exemplo, suponha o cenário da **Figura 8**. Neste, todas as estações foram configuradas com endereçamento IP na sub-rede 1.1.1.0/24. Deseja-se separá-las em dois grupos: o primeiro, composto pelos computadores 1, 4 e 5 (azul); e o segundo, por 2, 3 e 4 (vermelho). Note que a estação 4 é a única que pode se comunicar com ambos os grupos. A infraestrutura foi configurada com os dispositivos A, B, C e D. Neste simples exemplo, qual seria

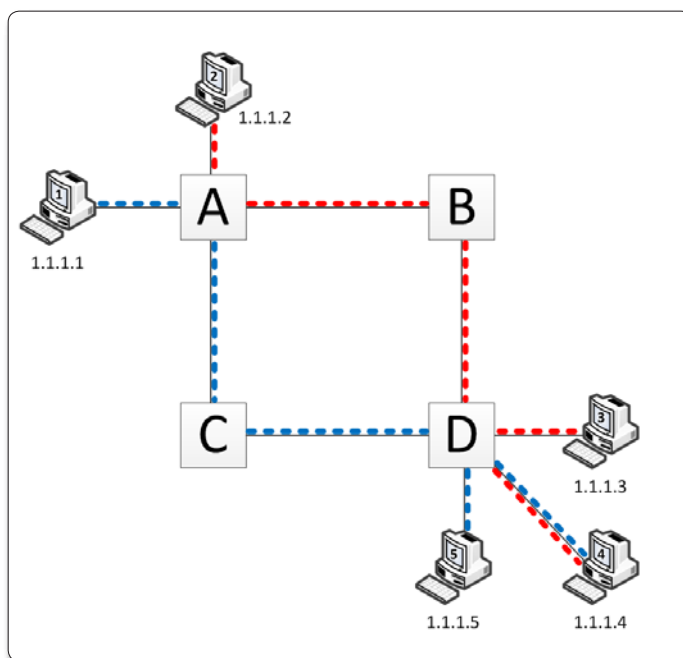


Figura 8. Exemplo de topologia com dois grupos de usuários

a topologia de rede que suportaria estes requisitos, considerando as tecnologias atualmente disponíveis?

Uma proposta inicial (**Figura 9**) poderia sugerir a utilização de *hubs* (repetidores) nas posições A, B, C e D. Entretanto, nesta topologia seria necessária a utilização do protocolo *spanning tree protocol* (ou de suas variantes) para remover logicamente os caminhos redundantes existentes (observe que existem duas possibilidades distintas entre as estações 1 e 2, e as estações 3, 4 e 5: a primeira, passando por A, B e D; e a outra, por A, C e D). Este protocolo evita a ocorrência de *loops* na comunicação de rede, desabilitando logicamente as conexões físicas duplicadas (tornando-as disponíveis somente em situações de contingenciamento). Todavia, o *spanning tree protocol* não é suportado em *hubs*, somente *switches* disponibilizam este recurso. Este cenário inicial também não atenderia ao requisito de separação em dois grupos, visto que todas as estações poderiam comunicar-se entre si.

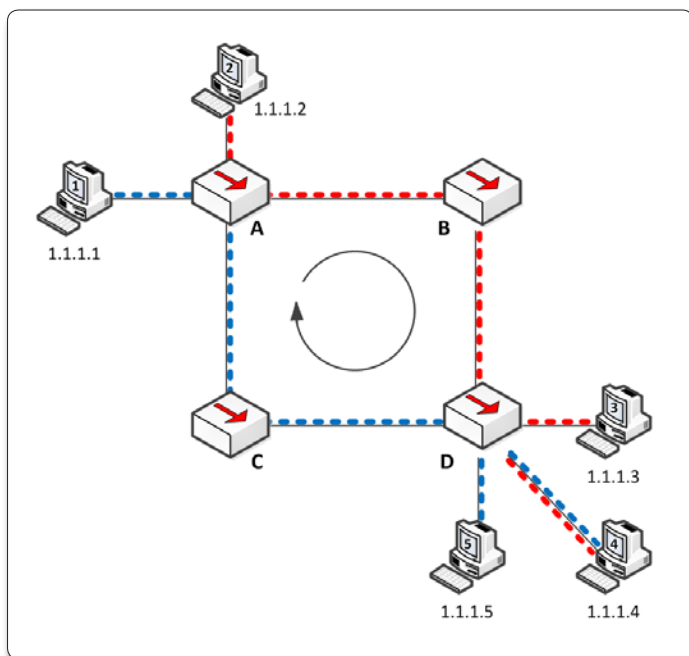


Figura 9. Proposta inicial de solução com o uso de hubs

A solução descrita na Figura 10 admite a hipótese que A, B, C e D são switches. Logo, poderiam ser criadas duas VLANs para separação das estações. O spanning tree protocol trataria a questão dos caminhos redundantes. Entretanto, como o computador 4 poderia participar de ambos os grupos? Segundo os requisitos apresentados, esta estação possui somente um adaptador de rede e um único endereço IP. A configuração de subinterfaces de redes com suporte a 802.1Q (recurso conhecido como vlan tagging, que permite a uma placa de rede ser configurada em diversas vlans) também não seria viável, visto que não é recomendada a configuração do mesmo endereço IP em ambas as subinterfaces.

A partir destes exemplos, observa-se que determinadas necessidades, por mais simples que sejam, resultam em dificuldades para os profissionais da área de infraestrutura. Neste contexto, nota-se a importância das redes definidas por software, mais especificamente, do OpenFlow. Por meio deste, os planos de dados dos switches A, B, C e D poderiam ser configurados para apresentar o comportamento descrito inicialmente, conforme as tabelas de fluxos da Figura 11.

Evolução do OpenFlow

Segundo o artigo *OpenFlow Switch Specification Version 1.1.0*, a versão 1.1.0 deste protocolo foi apresentada em Dezembro de 2010. Nesta, o switch OpenFlow suporta múltiplas tabelas de fluxos organizadas e numeradas sequencialmente a partir do número 0 (Figura 12). O processo de busca inicia-se na tabela zero. Outras tabelas podem ser consultadas dependendo das ações correspondentes aos fluxos.

Portanto, quando uma associação é encontrada, o pacote pode ser explicitamente direcionado para a tabela seguinte utilizando a instrução *go-to* (encaminha para).

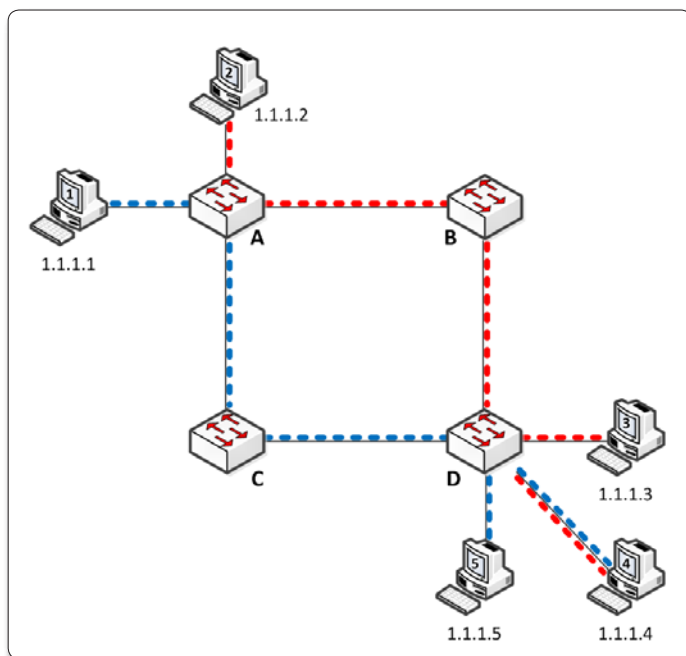


Figura 10. Proposta de topologia com switches

Não perca tempo reinventando a roda!

COBREBEMX

Componente completo para sua Cobrança por Boleto Bancário e Débito em Conta Corrente

Mais de 40 exemplos em diversas linguagens de programação

Geração e leitura de arquivos (remessa e retorno) nos padrões FEBRABAN e CNAB

Testes e Downloads gratuitos em nosso site

ACESSE E CONHEÇA O COMPONENTE EM:
WWW.COBBREBEM.COM

Switch A											
Porta de entrada	VLAN	MAC origem	MAC destino	Tipo ethernet	IP origem	IP destino	Protocolo	Porta de origem	Porta de destino	Ações	Estatísticas
Estação 1	*	*	*	*	1.1.1.1	1.1.1.4	*	*	*	Encaminhar para a porta de conexão com o switch C	Atualizar contadores
Estação 1	*	*	*	*	1.1.1.1	1.1.1.5	*	*	*	Encaminhar para a porta de conexão com o switch C	Atualizar contadores
Estação 2	*	*	*	*	1.1.1.2	1.1.1.3	*	*	*	Encaminhar para a porta de conexão com o switch B	Atualizar contadores
Estação 2	*	*	*	*	1.1.1.2	1.1.1.4	*	*	*	Encaminhar para a porta de conexão com o switch B	Atualizar contadores
*	*	*	*	*	*	*	*	*	*	Descartar	Atualizar contadores
Switch B											
Porta de entrada	VLAN	MAC origem	MAC destino	Tipo ethernet	IP origem	IP destino	Protocolo	Porta de origem	Porta de destino	Ações	Estatísticas
Conexão com o switch A	*	*	*	*	1.1.1.2	1.1.1.3	*	*	*	Encaminhar para a porta de conexão com o switch D	Atualizar contadores
Conexão com o switch A	*	*	*	*	1.1.1.2	1.1.1.4	*	*	*	Encaminhar para a porta de conexão com o switch D	Atualizar contadores
*	*	*	*	*	*	*	*	*	*	Descartar	Atualizar contadores
Switch C											
Porta de entrada	VLAN	MAC origem	MAC destino	Tipo ethernet	IP origem	IP destino	Protocolo	Porta de origem	Porta de destino	Ações	Estatísticas
Conexão com o switch A	*	*	*	*	1.1.1.1	1.1.1.4	*	*	*	Encaminhar para a porta de conexão com o switch D	Atualizar contadores
Conexão com o switch A	*	*	*	*	1.1.1.1	1.1.1.5	*	*	*	Encaminhar para a porta de conexão com o switch D	Atualizar contadores
*	*	*	*	*	*	*	*	*	*	Descartar	Atualizar contadores
Switch D											
Porta de entrada	VLAN	MAC origem	MAC destino	Tipo ethernet	IP origem	IP destino	Protocolo	Porta de origem	Porta de destino	Ações	Estatísticas
Conexão com o switch B	*	*	*	*	1.1.1.2	1.1.1.3	*	*	*	Encaminhar para a estação 3	Atualizar contadores
Conexão com o switch B	*	*	*	*	1.1.1.2	1.1.1.4	*	*	*	Encaminhar para a estação 4	Atualizar contadores
Conexão com o switch C	*	*	*	*	1.1.1.1	1.1.1.4	*	*	*	Encaminhar para a estação 4	Atualizar contadores
Conexão com o switch C	*	*	*	*	1.1.1.1	1.1.1.5	*	*	*	Encaminhar para a estação 5	Atualizar contadores
*	*	*	*	*	*	*	*	*	*	Descartar	Atualizar contadores

Figura 11. Tabelas de fluxos configuradas nos switches A, B, C e D.

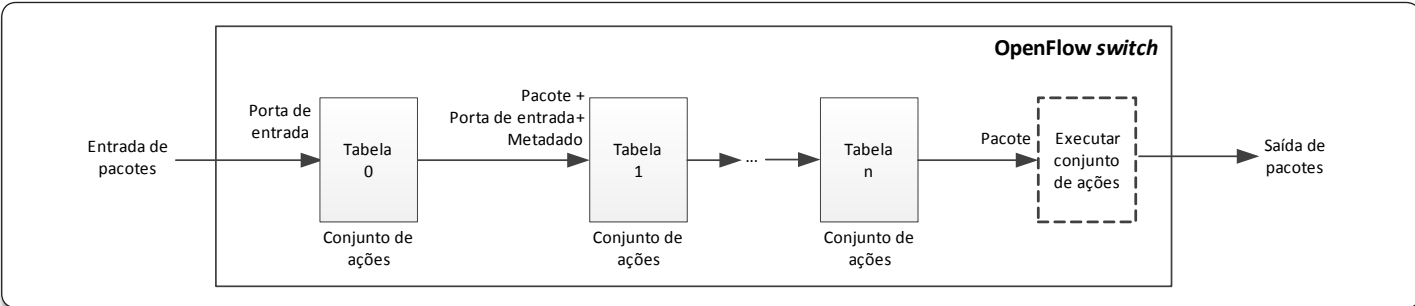


Figura 12. Múltiplas tabelas de fluxos

Porta de entrada	Metadado	Identificação da VLAN	Ethernet				MPLS		Internet Protocol				TCP/UDP	
			MAC origem	MAC destino	Tipo ethernet	Prioridade da VLAN	Etiqueta	Classe de serviço	IP origem	IP destino	Protocolo	Tipo de serviço	Porta de origem	Porta de destino

Figura 13. Campos suportados pela versão 1.1.0 do OpenFlow

Este é um procedimento que pode ser sucessivamente repetido. A configuração aplicada ao controlador OpenFlow determina diferentes comportamentos quando não se encontra a ação correspondente para um pacote nos registros das tabelas de fluxos:

1. Encaminhamento para processamento no controlador;
2. Envio para a próxima tabela de fluxos (se houver);
3. Descarte dos dados.

Nesta nova versão também foram adicionados novos campos para caracterização dos fluxos, entre eles (Figura 13):

- **Metadado (metadata):** empregado para a troca de informações entre as tabelas de fluxos no processamento sequencial;
- **Prioridade da vlan (VLAN priority):** bits PCP (*priority code point*) que indicam a prioridade do quadro (*frame*) ethernet. Geralmente utilizados para conceber classes de tráfego para voz, vídeo, dados, entre outras aplicações;
- **Etiqueta MPLS (multi-protocol label switching tag):** identificador (*tag*) configurado nas redes MPLS (infraestruturas baseadas

em pacotes rotulados, na qual cada etiqueta representa um índice na tabela de roteamento do próximo roteador);

- **Classe de serviço MPLS (MPLS traffic class):** permite a criação de níveis de prioridade para a definição de políticas de qualidade de serviço;
- **Tipo de serviço IP (IP type of service):** diferencia os pacotes transportados, classificando-os para que possam ter prioridade na transmissão.

Cada registro na tabela de fluxos está associado a um conjunto de instruções. Estas podem alterar campos dos cabeçalhos nos pacotes, realizar ações específicas e/ou continuar o processamento sequencial nas tabelas. Entre as instruções suportadas, destacam-se:

- **Aplica as ações (apply-actions):** confirma a execução das ações. Pode ser utilizada para modificar os pacotes entre duas tabelas ou para programar múltiplas ações do mesmo tipo;
- **Remove as ações (clear-actions):** remove imediatamente todas as ações;

DÊ UM SALTO EM CONHECIMENTO!



Acesse o maior portal para desenvolvedores da América Latina!

20
mil
posts

430
mil
cadastrados

10
milhões de
page-views
por mês

- **Grava as ações (*write-actions*):** grava as ações inseridas, alteradas ou removidas no conjunto existente;
- **Grava os metadados (*write-metadata*):** grava os metadados inseridos, alterados ou removidos no conjunto atual;
- **Encaminha para (*go-to*):** indica a próxima tabela no processamento sequencial, observando que sua identificação (número) deve ser maior que a anterior. Esta instrução não é válida para a última tabela de fluxos.

Por fim, a versão 1.1.0 suporta mais ações que a especificação anterior, tais como: mudanças de campos no cabeçalho, mapeamento dos pacotes para certas classes de prioridades, cópia e decremento do TTL (*time-to-live*), entre outras.

Conclusões

A infraestrutura de redes de computadores fundamentada nos protocolos clássicos das camadas de enlace e de redes está presente nos serviços disponibilizados por empresas públicas e privadas aos seus milhões de usuários. O célere crescimento da Internet, caracterizado pelo uso intensivo da navegação na *web*, do correio eletrônico, das redes sociais e dos programas de mensagens instantâneas, torna a rede mundial de computadores um serviço essencial às pessoas e corporações. Estas aplicações são suportadas, em sua maioria, por *switches* e roteadores proprietários e altamente especializados. As necessidades de desenvolvimento de novos protocolos e recursos nestes equipamentos devem ser homologadas e implementadas por seus fabricantes, conforme sua disponibilidade, interesses e procedimentos internos. Destarte, o processo de evolução geralmente é dispendioso e moroso, antepondo-se as necessidades da era da informação, marcada pela virtualização dos servidores, pela computação nas nuvens e pelo crescimento exponencial do tráfego na Internet.

Neste cenário foram concebidas as redes definidas por *software*, as quais objetivam desagregar os planos de dados (*hardware*) e controle (*software*) a fim de usufruir suas melhores características, respectivamente: desempenho e flexibilidade/agilidade. Atualmente, o OpenFlow é uma das tecnologias mais disseminadas que permite a adoção deste novo paradigma de controle e inovação. Seu potencial disruptivo já foi observado pelos principais fornecedores de *hardware*. Equipamentos de empresas como Cisco Systems, Juniper Networks, Brocade Communications, Extreme Networks, Arista, entre outras, já suportam esta tecnologia. Provedores de serviços e operadoras, como Amazon, Facebook, Google, Deutsche Telekom e Docomo, também iniciaram seus testes de homologação e possuem infraestruturas controladas por arquiteturas baseadas em OpenFlow.

Assim, observa-se uma tendência na qual as redes programáveis serão o alicerce a uma era de inovação nos dispositivos de redes,

envolvendo pesquisadores, profissionais dos setores de telecomunicações e fabricantes. Provavelmente esta evolução ocorrerá com uma velocidade nunca antes observada, possibilitando o desenvolvimento de novas arquiteturas, protocolos e serviços personalizados para as demandas atuais da era da informação.



André Koide da Silva

andre@aksilva.com

Mestre em Engenharia pela Escola Politécnica da Universidade de São Paulo, especialista em Redes de Computadores pela Universidade Estadual de Campinas e bacharel em Sistemas de Informação pela Universidade Presbiteriana Mackenzie. Atualmente é especialista em redes

de computadores pelo UOL Diveo, com ampla experiência em projetos e implementação de soluções Cisco Systems, Brocade Communications Systems e Juniper Networks. Atua também como professor na Faculdade Impacta Tecnologia e na Universidade Gama Filho. Trabalhou em empresas como Goldnet TI, EDS (Electronic Data Systems do Brasil) – atual HP Brasil (Hewlett-Packard), Alog Data Centers e Gennari & Peartree Projetos e Sistemas. Experiência de treze anos na área de redes de computadores e tecnologia da informação, além de sete anos como docente do ensino superior nestas áreas. Certificado Cisco IP Communications Express Specialist, Cisco Certified Network Associate (CCNA), Extreme Network Associate (ENA), 3Com Certified Wireless Expert, Juniper Networks Certified Internet Specialist (JNCIS-M), Project Management Professional (PMP), Information Technology Infrastructure Library (ITILv3), ITIL v3 Release, Control and Validation (RCV) e ISO/IEC 20000 - IT Service Management.



Artigo “Hobbes’ Internet Timeline 10.2”, escrito por Robert H’obbes’ Zakon.

<http://www.zakon.org/robert/internet/timeline>.

Artigo “OpenFlow: Enabling Innovation in Campus Networks” escrito por Nick McKeown, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker e Jonathan Turner.

<http://www.openflow.org/documents/openflow-wp-latest.pdf>.

Artigo “OpenFlow e redes definidas por software: um novo paradigma de controle e inovação em redes de pacotes” escrito por Christian Esteve Rothenberg, Marcelo Ribeiro Nascimento, Marcos Rogério Salvador e Maurício Ferreira Magalhães.

http://www.cpqd.com.br/cadernosdetecnologia/Vol7_N1_jul2010_jun2011/pdf/artigo6.pdf.

Artigo “OpenFlow Switch Specification Version 1.1.0”, escrito por OpenFlow Consortium.

<http://www.openflow.org/documents/openflow-spec-v1.1.0.pdf>.

Artigo “Redes definidas por software: uma abordagem sistêmica para o desenvolvimento de pesquisas em redes de computadores”, escrito por Dorgival Guedes, Luiz Filipe Menezes Vieira, Marcos Menezes Vieira, Henrique Rodrigues e Rogério Vinhal Nunes.

<http://sbr2012.dcc.ufmg.br/app/pdfs/p-04/mc4.pdf>.



Gestão de Ativos – A organização nas mãos da TI

O que a TI pode fazer para assumir este novo papel nas organizações?

A gestão patrimonial de ativos é um trabalho que identifica e cataloga os bens físicos de uma organização para averiguar se determinado ativo existe na empresa, se está em sua devida localização e se está sendo utilizado pelas pessoas corretas, dentro de um prazo de vida útil adequado.

Essa tarefa sempre esteve ligada aos departamentos Fiscal e Contábil de uma organização, responsáveis por inventariar todos os ativos físicos, fixando uma plaqueta com uma numeração sequencial associada a um documento contendo suas principais características. Tais plaquetas facilitam sua identificação, localização, tempo de garantia, condição de uso e também a identificação dos responsáveis pela sua utilização, conservação e manutenção. No entanto, hoje, parte destes ativos está relacionada com algum equipamento de tecnologia, seja um computador, um dispositivo de comunicação ou mesmo algo que até há pouco tempo era encontrado em grande parte no formato físico, impresso e palpável, e hoje cada vez mais em meios virtuais: a informação!

Aos poucos, os departamentos de TI estão assumindo a responsabilidade em manter a gestão patrimonial de uma empresa, pelo menos no que tange os ativos sob sua responsabilidade (que não são poucos) e também apoiando os departamentos Fiscal e Contábil com sistemas que facilitem essa árdua tarefa. Se a revista Infra Magazine realizasse, hoje, uma pesquisa com seus leitores que trabalham na área de TI sobre quais as atuais atribuições do seu departamento, dificilmente teríamos respostas completamente semelhantes. Isso porque, ao longo dos anos visitando empresas e instituições dos mais variados setores de nossa economia, sejam elas pequenas, médias ou grandes, o que se tem notado é uma crescente absorção de tarefas que até pouco tempo atrás eram de inteira responsabilidade de outros departamentos.

Resumo DevMan

Gestão de Ativos – A organização nas mãos da TI:

Neste artigo apresentaremos aos leitores todas as implicações que a falta de um sistema de gestão pode causar em uma organização e como a TI pode e deve participar diretamente neste tipo de trabalho, uma vez que é fundamental para redução de custos operacionais, mitigação de uma série de riscos à segurança e aplicação das melhores práticas de mercado.

Em que situação o tema é útil:

A gestão de ativos não se trata apenas de uma tarefa para manter o ambiente de trabalho organizado, o que é muito importante também, mas é muito mais abrangente que isso. Fazer a gestão adequada de todos os ativos, especialmente os ativos relacionados à tecnologia, evita desperdícios com investimentos inapropriados, otimiza as atividades do negócio, permite a aderência a vários controles de normas de gestão, além de qualidade e segurança, fundamental em alguns casos para a sobrevivência da empresa em um mercado cada vez mais competitivo. Quando uma empresa não administra adequadamente seus ativos, é inevitável o aumento dos custos, que acabam parando no preço do seu produto final ou serviço.

Entre algumas tarefas adicionadas na bagagem dos profissionais de TI, certamente encontraremos serviços como manutenção e suporte de recursos de telefonia, administração de sistemas de câmeras de vigilância, inventários e gestão de ativos. Todas estas tarefas eram obrigações de outros departamentos, mas aos poucos foram absorvidas pelo departamento de TI. É sobre este assunto que trataremos neste artigo.

As recomendações da Norma ISO 27002

Ter todas as informações detalhadas e principalmente atualizadas dos equipamentos de informática e softwares utilizados numa organização não só é uma prática necessária para a localização dos ativos espalhados muitas vezes em plantas geograficamente

distantes, como também é um mecanismo que permite identificar possíveis fontes de desperdício de recursos mal ou subutilizados, ferramentas inadequadas para realização das atividades dos funcionários, depredação dos equipamentos e até fraudes ocasionadas por furtos e extravios muitas vezes cometidos pelos próprios colaboradores internos.

Também é muito importante reforçarmos que o departamento de TI não pode resumir os ativos a equipamentos de informática e softwares instalados no ambiente sob sua responsabilidade. É algo muito mais abrangente que envolve tudo o que pode ter ou fornecer um valor para uma organização, justamente porque a definição de um ativo na norma *ABNT NBR ISO/IEC 27002 - Código de Prática para a Gestão de Segurança da Informação* é bem sucinta: *Ativo: qualquer coisa que tenha valor para a organização*. Portanto, o primeiro passo é localizar e identificar todos estes ativos, que podem estar representados em diversas formas de acordo com a ISO:

- a) Ativos de informação: base de dados e arquivos, contratos e acordos, documentação de sistema, informações sobre pesquisa, manuais de usuário, material de treinamento, procedimentos de suporte ou operação, planos de continuidade do negócio, procedimentos de recuperação, trilhas de auditoria e informações armazenadas;
- b) Ativos de software: aplicativos, sistemas, ferramentas de desenvolvimento e utilitários;
- c) Ativos físicos: equipamentos computacionais, equipamentos de comunicação, mídias removíveis e outros equipamentos;
- d) Serviços: serviços de computação e comunicações, utilidades gerais, como aquecimento, iluminação, eletricidade e refrigeração;
- e) Pessoas e suas qualificações, habilidades e experiências;
- f) Intangíveis, tais como a reputação e a imagem da organização.

Como podemos notar na descrição dos tipos de ativos, fica compreensível entender os motivos que levaram a TI a assumir parcialmente a tarefa de inventariar e gerir estes itens, uma vez que boa parte deles necessita de conhecimentos técnicos para realizar a sua correta identificação, manutenção e em muitos casos a valorização estimada do ativo para a organização. Sem este apoio de especialistas da área de TI, a organização não terá condições de reduzir riscos, perdas financeiras ou ainda aplicar investimentos futuros de forma adequada.

A gestão de ativos é um elemento tão importante para o bom andamento das atividades de uma organização que há uma seção inteira destinada a este tema na principal norma para a gestão e organização da segurança das informações, *ABNT NBR ISO/IEC 27002 - Código de Prática para a Gestão de Segurança da Informação*. São dois capítulos contendo cinco controles que direcionam as organizações quanto à proteção de seus ativos, os quais estão organizados na seguinte estrutura:

7. Gestão de ativos

7.1. Responsabilidade pelos ativos

- 7.1.1. Inventário dos ativos;
- 7.1.2. Proprietário dos ativos;
- 7.1.3. Uso aceitável dos ativos.

7.2. Classificação da Informação

7.2.1. Recomendações para classificação;

7.2.2. Rótulos e tratamento da informação.

Quando observamos esta estrutura, temos a impressão de que o que está sendo solicitado é trivial, básico em qualquer processo de levantamento em um inventário. Porém, ativos não se tratam apenas de equipamentos físicos, mas de algo mais complexo que muitas vezes não parece palpável, como a informação eletrônica ou até mesmo aspectos como conhecimento, expertise e habilidade de um funcionário, ou ainda a imagem que uma organização possui perante seus clientes e o mercado corporativo.

Além disso, a norma é clara em informar que não basta apenas inventariar os ativos, mas que cada ativo deve possuir seus(s) proprietário(s) identificado(s), que se apliquem regras para o uso correto destes ativos, que estes sejam classificados e que esta classificação seja facilmente consultada por aqueles que tiverem acesso ao ativo.

Resumindo, os controles da Gestão de Ativos prevista na norma recomendam:

- **Inventário dos ativos:** Todos os ativos precisam ser identificados, localizados e que sua importância ou valor para a organização seja claramente definido para que a devida proteção possa ser aplicada sobre cada tipo de ativo. Estas informações não somente colaboram para a segurança da organização, como também para sua saúde financeira;
- **Proprietários dos ativos:** Todos os ativos relacionados com recursos de processamento da informação devem ter um proprietário definido para que este possa dar a devida classificação, especialmente para auxiliar no controle do seu acesso. Vale ressaltar que há uma nota de comentário na norma que esclarece que o termo “proprietário” é utilizado para identificar uma pessoa ou organismo que tenha responsabilidades sobre o ativo, mas não o direito de posse, pois se entende que o direito de posse é da organização que a fornece;
- **Uso aceitável dos ativos:** Depois que todos os ativos da organização forem mapeados, cabe a ela definir regras e controles para preservá-los junto aos seus funcionários, fornecedores e terceiros;
- **Recomendações para classificação:** Para dar a devida proteção ao ativo da informação, este deve ser classificado, pois com base nesta classificação é que pode ser determinada sua importância e seu valor à organização, auxiliando na definição dos perfis de acessos que serão concedidos;
- **Rótulos e tratamento da informação:** Para que essa classificação possa ser claramente entendida por todos em uma organização e válida com o passar do tempo, o ativo da informação deve ser rotulado para que as pessoas saibam facilmente quem pode ou não ter acesso.

Lendo o resumo, fica compreensível a limitação que outros departamentos possuem para realizar a gestão adequada dos ativos relacionados com tecnologia. Por isso, o envolvimento da TI tem sido cada vez maior. Seria possível escrever um novo artigo para cada um dos cinco pontos citados da norma, porém, como dito no

início, o foco será voltado ao inventário dos ativos relacionados à tecnologia, pois este é o começo de tudo. Se esta etapa for mal realizada ou não for concluída, a organização não conseguirá dar prosseguimento nas demais recomendações da ISO relacionadas à Gestão de Ativos.

Descrevendo o cotidiano real de muitas empresas brasileiras, imagine a situação apresentada no próximo tópico – ou relembre-a, caso o leitor já tenha vivenciado algo semelhante.

Uma árdua tarefa – Inventário

Vamos supor o cenário de uma empresa do ramo de autopeças que está se reorganizando, prevendo um crescimento de suas operações nos próximos anos. Fazendo parte dessa reorganização está o trabalho de inventário de todos os seus ativos, o que inclui o seu parque informático composto aproximadamente por 50 estações de trabalho, 5 servidores, 12 impressoras, entre outros periféricos distribuídos em duas plantas: uma administrativa no centro e outra fabril, numa área industrial da mesma cidade.

A Diretoria solicita ao gestor de TI a apresentação de um relatório de todos os ativos físicos sob a responsabilidade de seu departamento no prazo de 15 dias. Minimamente, o relatório deve conter a descrição do item, sua localização, seu proprietário e seu estado, se está sendo utilizado ou se está desativado, e seu número de patrimônio.

Aproveitando o ensejo, o gestor de TI repassa à sua equipe a árdua tarefa, acrescentando alguns detalhes que devem ser incluídos durante o levantamento para uso do próprio departamento de TI, que já carecia de tais informações. Além da descrição, localização, proprietário, estado do ativo e do número de patrimônio, a equipe deverá obter informações como marca, modelo, número de série, e para os computadores, informações básicas de seu hardware e software. No caso dos softwares, as informações coletadas devem incluir o CD-KEY da licença para conferir se a empresa está fazendo uso de softwares piratas ou está com o número de licenças acima do que determina o contrato. Também se deve adicionar um campo de observações para o conjunto de computadores, monitores e seus periféricos, além da identificação de sua cor predominante, pois há muita reclamação de falta de padronização de ambientes, como salas com computadores e teclados pretos usados com monitores e mouses brancos.

Depois que os técnicos retornarem, todos os dados coletados deverão ser organizados em uma planilha para gerar informações relevantes e gráficos que serão usados no relatório sumário que será encaminhado à Diretoria.

Se para uma empresa com aproximadamente 60 hosts divididos em duas plantas, este já seria um trabalho que despenderia um grande esforço por parte da equipe técnica, imaginem então uma rede com 100, 500, 1000 hosts divididos em várias plantas, em diversas localidades. Isso tudo sem esquecermos que a equipe de TI sempre está atarefada com suas outras e incessantes atividades de rotina.

Para piorar a situação, parte de todo este monumental trabalho de coleta de informações não teria validade superior a um dia,

pois quem garante que ao virar as costas os computadores continuarão com os mesmos softwares instalados ou que algum item de hardware não será alterado até a entrega do relatório?

Manter atualizado o inventário de itens tão dinâmicos como computadores que constantemente sofrem alterações de hardware e software fica completamente inviável se trabalharmos de forma manual, pois, na verdade, o que está sendo realizado não é a gestão destes ativos, mas uma fotografia do seu estado no momento da coleta de dados.

Quer outro argumento preocupante? Imagine que no meio desse processo ocorra uma denúncia anônima e a empresa seja fiscalizada por alguma associação de fabricantes de softwares contra pirataria, como a ABES (*Associação Brasileira de Empresas de Software*). Será que o gestor teria em mãos todas as licenças identificadas, incluindo o CD-KEY, contratos e notas fiscais?

E se ele tiver tais informações, será que elas estão atualizadas? Como garantir que a relação de softwares não foi modificada ao longo do tempo? Será que foram instalados softwares sem a ciência do departamento de TI, ou, pior, será que alguém na TI não fez uso de softwares piratas, na incumbência de resolver problemas rápidos, diante a estressante rotina de trabalho?

Tudo isso que apontamos aqui não é fantasia ou um mero exemplo para ilustrar o artigo. Esse cenário é muito comum no dia a dia de organizações que não possuem um planejamento para organizar a TI, pois sempre estão atarefados com as rotinas de suas atividades ou sem investimento adequado para permitir tal reestruturação.

Existe solução para a TI?

Este artigo traz duas soluções que podem auxiliar a TI na gestão de todos os seus ativos e, em alguns casos, até auxiliar a gestão de outras áreas com ferramentas integradas que mostrarão que TI não é apenas um departamento que gera custos e cuja única função é prestar suporte técnico, mas para ser um braço estratégico a toda organização. Entre várias soluções disponíveis no mercado, apresentamos uma opção de software livre, conhecida como CACIC, e outra com licenciamento comercial chamado ADOTI. O destaque de ambas as soluções reside no fato de serem Tupiniquins, com origens muito interessantes que mostram o quanto o setor de tecnologia de nosso país pode ser levado a sério quando governos e instituições acadêmicas fomentam e apoiam desenvolvedores, empreendedores, professores e universitários, permitindo assim que novas ideias e melhorias possam ser apresentadas ao mercado corporativo.

CACIC

Com um nome bem sugestivo e criativo, o CACIC foi lançado em 2005 e tem sido motivo de orgulho para o Governo Federal por ser este o primeiro e também o mais conhecido software criado e desenvolvido em uma parceria com o Ministério do Planejamento, Orçamento e Gestão (MP), da Secretaria de Logística e Tecnologia da Informação (SLTI) e da Empresa de Tecnologia e Informações da Previdência Social (Dataprev).

CACIC é uma abreviação que traduz literalmente sua funcionalidade: Configurador Automático e Coletor de Informações Computacionais, ou seja, um sistema de inventário automatizado na rede. Hoje se encontra na versão 2.6 e vem apresentando muitas melhorias ao longo de seus sete anos de vida, com o apoio da Comunidade de desenvolvedores voluntários.

Segundo informações do Portal do Software Público, o CACIC possui atualmente cerca de 30.000 usuários entre instituições do setor público, empresas do setor privado e organizações não governamentais. Além de seu reconhecimento em território nacional, o software também é utilizado pelos governos de países como Argentina, Venezuela e Paraguai.

O seu desenvolvimento e sua implantação na rede são possíveis graças a um conjunto de outras ferramentas livres, como linguagens de programação: PHP, Perl, Python e Delphi e o sistema gerenciador de banco de dados MySQL. Seu funcionamento é baseado na comunicação de um agente instalado nos computadores que mantém comunicação com o servidor CACIC. Veja a Figura 1.

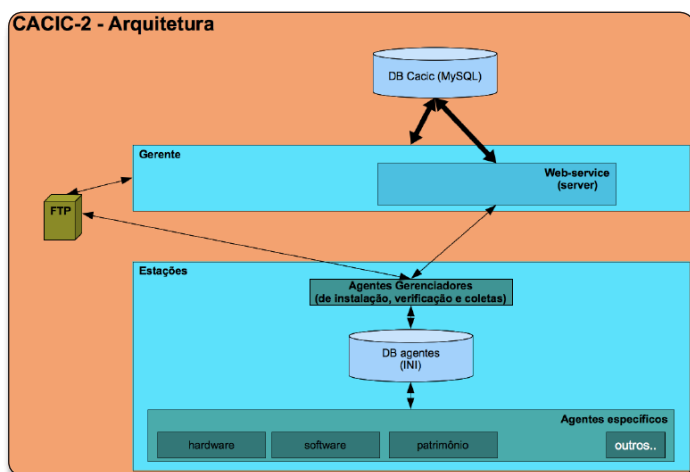


Figura 1. Arquitetura básica do CACIC série 2.x

Como já citado, o CACIC passa por melhorias a cada nova versão, com a inclusão de novas funcionalidades. Além de coletar informações do hardware e do software dos computadores na rede automaticamente, a ferramenta hoje permite o cadastro de licenças de softwares, informações adicionais dos equipamentos físicos, como cadastro de dados patrimoniais, e até suporte remoto ao usuário.

Em resumo, o CACIC permite:

- Coletar informações sobre os componentes de hardware instalados em cada computador e disponibilizá-las aos administradores de sistemas (veja a Figura 2);
- Alertar os administradores de sistemas quando forem identificadas alterações na configuração dos componentes de hardware de cada computador;
- Coletar diversas informações sobre os softwares instalados em cada computador e disponibilizá-las aos administradores de sistemas (veja as Figuras 3 e 4);

- Identificar diretórios compartilhados considerados inseguros e aplicar as restrições de segurança necessárias (veja a Figura 5);

Hardware instalado		
Placa mãe	Descrição	SH 3221
	Fabricante	Itautec Plúcio S.A.
BIOS	Descrição	F4
	Fabricante	Itautec SH 3221, LTD 4.00 PG, -512
Memória RAM	Slot 0:	1024MB (DIMM)
	Slot 1:	1024MB (DIMM)
Teclado	Descrição	Teclado padrão com 101/102 teclas ou Microsoft Natural PS/2 Keyboard
Mouse	Descrição	Mouse compatível com PS/2
Placa de som	Descrição	
Placa de vídeo	Descrição	ATI Radeon Xpress 1200
	Quantidade de cores	32
	Resolução	1280x960
	Quantidade de memória	768 MB
Unidade de Disco Óptico 1	Descrição	LITE-ON DVD C LH52C1P
CPU 1	Descrição	AMD Athlon(tm) 64 X2 Dual Core Processor 5600+
	Fabricante da CPU:	AMD
	Número de série da CPU:	0004-0F33-178B-F8FF-0000-2001
Placa de Rede 1	Descrição	Realtek RTL8169/8110 Family Gigabit Ethernet NIC - Miniporta do agendador de pacotes
	Endereço MAC da estação:	00:1D:7D:87:FE:78
	IP da estação:	10.71.0.145
	Máscara de Subrede:	255.255.255.0
	IP do gateway:	10.71.0.250
	IP do servidor DHCP:	10.71.0.210
	IP do servidor WINS primário:	10.0.134.223
	IP do servidor WINS secundário:	10.70.1.220
Placa de Rede 2	Descrição	VirtualBox Host-Only Ethernet Adapter - Miniporta do agendador de pacotes
	Endereço MAC da estação:	08:00:27:0000:0:1F
	IP da estação:	192.168.56.1
	Máscara de Subrede:	255.255.255.0
	IP do gateway:	255.255.255.255
	IP do servidor DHCP:	0.0.0.0
	IP do servidor WINS primário:	0.0.0.0
	IP do servidor WINS secundário:	0.0.0.0

Figura 2. Descrição de hardware de um computador

Versões de softwares básicos			
Sistema Operacional	Windows XP Professional	Versão do DirectX	409
Versão do internet explorer	7.0.5730.11	Versão do ODBC	3.525.1132.0 (x
Versão do Mozilla	1.9.2.6	Versão do DAO	3.6
Versão do Acrobat Reader	5.0	Versão do ADO	2.8
Versão da maquina virtual java (JVM)	1.6	Versão do BDE	5.2.0.2

Figura 3. Relação de softwares básicos de um computador

Softwares inventariados	
1 7-Zip 4.60 beta	
2 Adobe Acrobat 5.0	
3 Adobe Flash Player 10 Plugin	
4 Adobe Flash Player ActiveX	
5 Adobe Shockwave Player 11	
6 ATI Display Driver	
7 Atualizacao de Seguranca para o Windows Media Player (KB911564)	
8 AVG Free 8.0	
9 Avira NTFS4DOS 1.9	
10 Bouml 4.19.3	
11 CollabNet Subversion Client 1.5.5	
12 CutePDF Writer 2.8	
13 Desktop Delivery	
14 Diebold Printers 1.2.0.3	
15 EasyEclipse Desktop Java 1.3.1.1	
16 Ext2 IFS 1.11 for Windows XP	
17 GnuWin32: GetText version 0.14.4	
18 High Definition Audio Driver Package - KB835221	
19 High Definition Audio Driver Package - KB888111	
20 HK-Software IBEExpert Developer Studio Trial Version	
21 HyperTerminal Private Edition v7.0	
22 InterBase 6.5	
23 Java Media Framework 2.1.1e	
24 JUDE Community 5.4.1	
25 Microsoft Internationalized Domain Names Mitigation APIs	
26 Microsoft Office Enterprise 2007	
27 Pacote de Provedor de Servicos de Criptografia para o Microsoft Base Smart Card	
28 Windows Internet Explorer 7	

Figura 4. Descrição de softwares inventariados na rede

Compartilhamento de diretórios e impressoras				
Nome	Diretório	Comentário	Tipo	
print\$	c:\windows\system32\spool\drivers	Drivers de impressora	Impressora	
ShareDocs	c:\documents and settings\all users\documentos		Diretório	
Impressora2	\\http://10.71.0.1\mp01\localiponly	\\http://10.71.0.1\mp01	Impressora	

Figura 5. Compartilhamentos de diretórios e impressoras

- Coletar informações de Patrimônio (PIB, localização, etc.) de cada computador e disponibilizá-las aos administradores de sistemas (veja a Figura 6);

Figura 6. Informações de patrimônio cadastradas

- Alertar os administradores quando forem identificadas alterações na localização física do computador (veja a Figura 7);

Figura 7. Opções de envio de alerta para mudança físicas no computador

- Permitir aos administradores de sistemas o envio de mensagens administrativas aos usuários de um computador específico ou usuários de um grupo de computadores (veja a Figura 8);

Figura 8. Grupo de computadores

- Permitir a detecção do uso de dispositivos removíveis nos computadores gerenciados pelo agente CACIC (veja a Figura 9);

Fabricante	Dispositivo	Data/Hora	Evento
0951 - Kingston Technology	162a - DataTraveler 112 4GB Pen Drive	19/07/2010 às 17:44h	Remoção
0951 - Kingston Technology	162a - DataTraveler 112 4GB Pen Drive	19/07/2010 às 17:43h	Inserção
0781 - SanDisk Corp.	5567 - Cruzer Blade	19/07/2010 às 15:57h	Inserção
0781 - SanDisk Corp.	5567 - Cruzer Blade	19/07/2010 às 15:50h	Remoção
0951 - Kingston Technology	162a - DataTraveler 112 4GB Pen Drive	19/07/2010 às 14:46h	Remoção
0951 - Kingston Technology	162a - DataTraveler 112 4GB Pen Drive	19/07/2010 às 13:49h	Inserção
0781 - SanDisk Corp.	5567 - Cruzer Blade	19/07/2010 às 13:02h	Inserção
1043 - iCreate Technologies Corp.	Dispositivo Desconhecido	17/07/2010 às 19:46h	Remoção
0951 - Kingston Technology	162a - DataTraveler 112 4GB Pen Drive	16/07/2010 às 18:08h	Remoção
0781 - SanDisk Corp.	5567 - Cruzer Blade	16/07/2010 às 17:50h	Remoção
0951 - Kingston Technology	162a - DataTraveler 112 4GB Pen Drive	16/07/2010 às 17:05h	Inserção
0951 - Kingston Technology	162a - DataTraveler 112 4GB Pen Drive	16/07/2010 às 16:53h	Remoção
0951 - Kingston Technology	162a - DataTraveler 112 4GB Pen Drive	16/07/2010 às 16:33h	Remoção
0951 - Kingston Technology	162a - DataTraveler 112 4GB Pen Drive	16/07/2010 às 16:22h	Inserção

Figura 9. Relatório de mídias removíveis conectadas na rede

- Suporte Remoto Seguro, chamado de srCACIC, que permite a conexão remota ao computador cliente para realização de atendimentos de suporte técnico, desde que a conexão seja aceita pelo usuário (veja a Figura 10);

Figura 10. Telas do processo de conexão remota srCACIC

Por se tratar de um software livre com várias funcionalidades fundamentais no cumprimento das tarefas diárias das equipes de suporte, obviamente chama a atenção e cria expectativas para qualquer gestor de TI. No entanto, alguns fatores devem ser considerados antes de sua implementação, como assumir a implantação do sistema e eventuais problemas técnicos por conta da equipe interna, capacitar a equipe para desenvolvimento e manutenção da ferramenta, não ter prazo definido para resoluções de problemas técnicos e não ter garantias contratuais da continuidade do desenvolvimento do produto, por se tratar de um sistema aberto, que depende de colaboradores voluntários na comunidade. Assim, pode-se economizar em questão de licenciamento, mas acaba-se gastando com a equipe internamente, especialmente com capacitação técnica e suporte de terceiros.

ADOTI

O mercado corporativo também possui diversas soluções de excelente qualidade para auxiliar na gestão de ativos, porém algumas são consideradas no meio técnico como demasiadamente complexas, que se adaptam somente a organizações de grande porte, ou que podem destinar uma equipe apenas para assumir o seu gerenciamento. Algumas destas soluções, além de complexas, possuem elevados custos de licenciamento, implantação e operação, o que nos leva a entender as dificuldades de empresas do setor das PMEs (*Pequenas e Médias Empresas*) e instituições governamentais que acabam abrindo mão deste tipo de ferramenta. Como muitas destas organizações não possuem mão de obra disponível para investir tempo e dinheiro para a capacitação de suas equipes, acabam por deixar de fazer uso de uma importante solução.

Para atender a essa grande demanda, surgem soluções que focam neste nicho do mercado, inclusive soluções nacionais, como é o caso do ADOTI. Trata-se de um sistema de gestão de ativos que, de forma similar ao CACIC, realiza a coleta de informações de hardware e software dos computadores na rede por meio de um agente instalado, mas com alguns diferenciais que permitem uma gestão muito mais abrangente dos ativos de TI, como coletar informações de dispositivos como switches, impressoras, thin-clients, roteadores e qualquer outro equipamento que possua o protocolo de comunicação SNMP (*Simple Network Management Protocol*). Também permite o cadastro de qualquer outro ativo físico que seja importante para a área de TI, tais como celulares, no-breaks, aparelhos de ar-condicionado, contratos e – por que não? – até pessoas.

Desenvolvido pela eSysTech, empresa nascida do trabalho conjunto de alunos e professores na incubadora da Universidade Federal Tecnológica do Paraná (UTFPR), a solução foi criada em 2002 diante uma demanda da Positivo Informática em atender os requisitos de licitações do governo que exigiam um sistema de inventário embarcado nos computadores. Na verdade, o nome ADOTI só veio a surgir em meados de 2009, pois até então o nome da solução era PNM (*Positivo Network Manager*). Foi assim até a sua abertura para o mercado, permitindo que organizações de todos os tamanhos pudessem ter em mãos uma solução CMDB (*Configuration Management Database*).

Entre suas diversas funcionalidades, destacamos as principais ferramentas que auxiliam as atividades das equipes de TI, sem maiores dificuldades na sua implantação e operação. Confira as descrições destas ferramentas a seguir:

- Inventário completo de hardware e software automático. Com base nas informações coletadas pelo agente instalado nos computadores na rede, o administrador possui informações detalhadas do hardware, ao ponto de detectar até o número de série de componentes como memórias e discos rígidos. Em relação ao software, realiza a coleta de cada programa instalado, permitindo a coleta de todas as chaves seriais dos produtos Microsoft. Qualquer alteração no computador é registrada pelo agente e enviada ao servidor online (veja as Figuras 11 e 12);

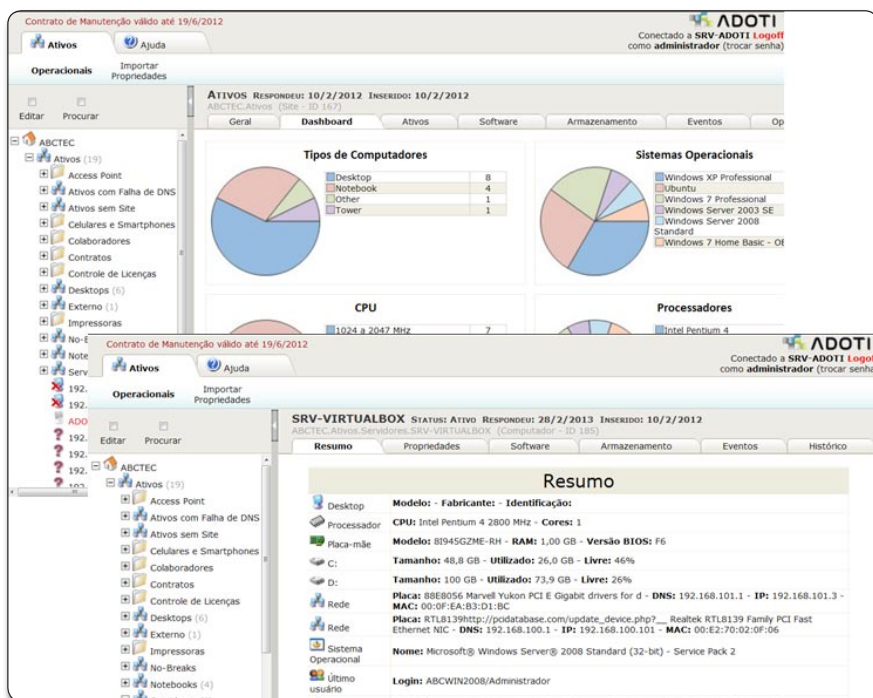


Figura 11. Informações detalhadas de hardware dos computadores na rede

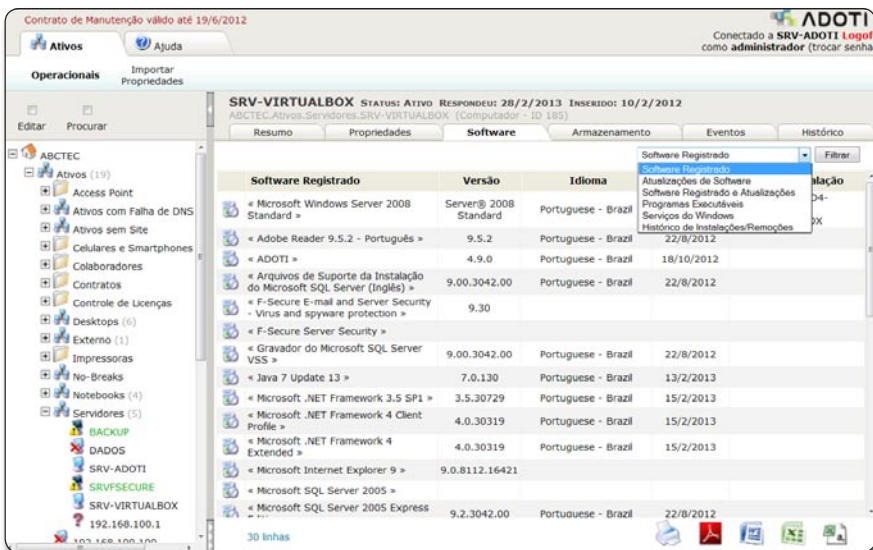


Figura 12. Informações detalhadas do software, incluindo as chaves dos produtos Microsoft

- Cadastro de qualquer ativo para complementar o inventário, permitindo que qualquer equipamento que esteja sob a responsabilidade do departamento de TI possa ser cadastrado de forma personalizada. Cada empresa define quais equipamentos serão gerenciados e quais informações serão cadastradas (veja a **Figura 13**);

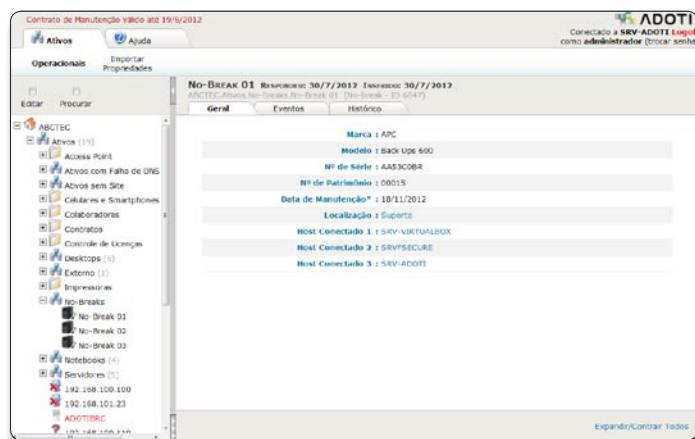


Figura 13. Exemplo de cadastro de no-breaks para controlar a data de manutenção

- Relacionamento dos ativos aos contratos ou notas fiscais, permitindo que a equipe de TI possa ter facilmente em mãos, informações complementares dos equipamentos ou softwares cadastrados no ADOTI, evitando a burocracia de solicitar uma cópia da nota ao departamento Fiscal da organização, isso quando não se perdia a própria nota original (veja a **Figura 14**). O exemplo aqui é do cadastro de uma nota fiscal, mas pode-se associar qualquer documento a um ativo, como contratos, termos de responsabilidade, fotos, planilhas, etc. E o arquivo pode ser em qualquer formato, como *.pdf, *.doc, *.xls, *.jpg, *.gif;

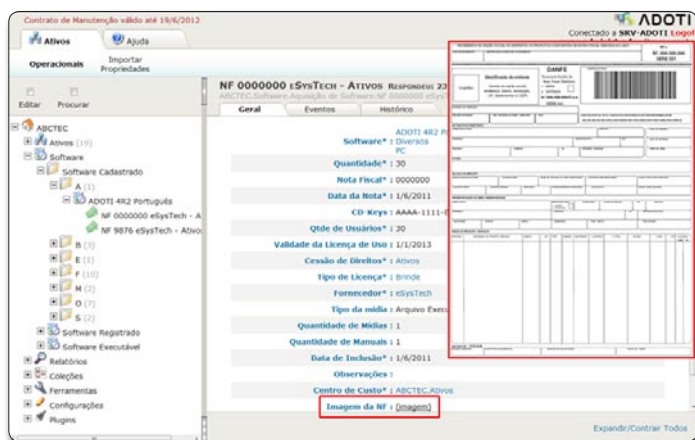


Figura 14. Imagem digitalizada da Nota Fiscal inserida no cadastro de um software

- Medição de desempenho dos computadores permite que os analistas possam verificar problemas de hardware que estejam afetando o desempenho dos computadores avaliados, sem a necessidade de retirá-lo para o laboratório (veja a **Figura 15**);

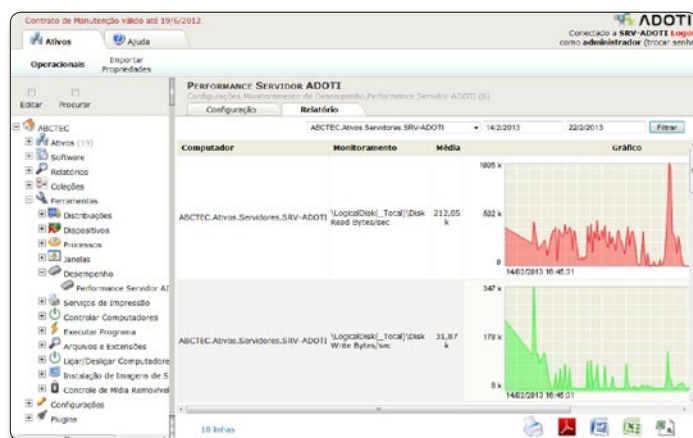


Figura 15. Checagem de performance de disco, rede, processamento e memória

- Identificação de todos os compartilhamentos na rede (veja a **Figura 16**). Este tipo de auditoria permite ao administrador identificar todos os compartilhamentos na rede, mesmos ocultos, informando os usuários e os privilégios de acesso. Ferramenta importantíssima para evitar disseminação de malwares e principalmente o vazamento de informações;

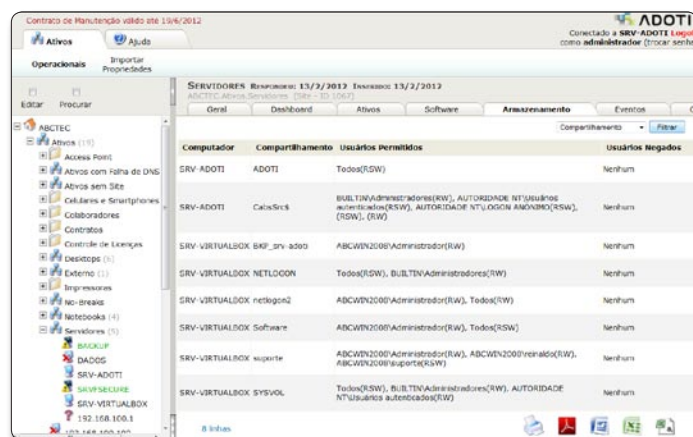


Figura 16. Diretórios compartilhados na rede

- Auditoria dos arquivos armazenados nos computadores é outra ferramenta imprescindível no controle das informações que são armazenadas nos computadores da organização. Como exemplo, tomemos o caso de arquivos multimídia (*áudio e vídeo*), que são facilmente encontrados em qualquer computador e em muitos casos até em diretórios escondidos nos servidores, mesmo que estes estejam ocultos. Além de ocuparem espaço no armazenamento da rede, muitas vezes tais arquivos expõem a organização a fiscalizações por direitos autorais e até investigações policiais, pois têm se tornando comum o armazenamento de material impróprio com conteúdo pedófilo ou discriminatório (veja a **Figura 17**);
- Monitoramento das impressões registra todas as impressões realizadas pelos computadores, com a informação da data, hora,

usuário, computador, impressora, nome do documento e quantidade de páginas. Mesmo que o usuário imprima fora da rede da empresa, o registro será enviado ao servidor ADOTI (veja a Figura 18);

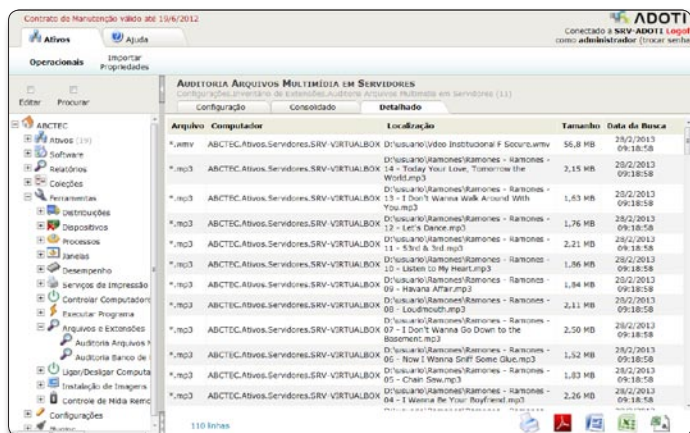


Figura 17. Auditoria de arquivos multimídia nos computadores na rede

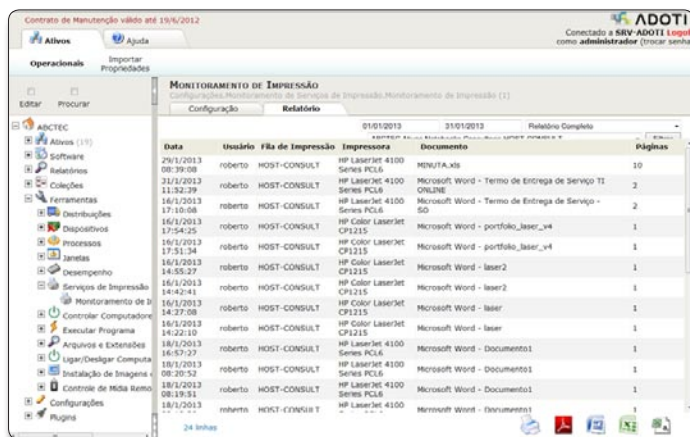


Figura 18. Monitoramento de impressões

- Monitoramento e bloqueio de janelas e aplicações permite a auditoria ou a restrição de acesso a aplicativos que não sejam autorizados pela organização. Este tipo de ferramenta é muito eficaz, pois mesmo com uma rede protegida com firewall e proxy limitando os acessos à Internet, é comum encontrar usuários utilizando aplicativos localmente de forma improdutiva. O monitoramento ou bloqueio pode ser realizado pelo nome do processo ou título de sua janela (veja as Figuras 19 e 20);
- Monitoramento e finalização de processos. Todos os processos ativos no computador podem ser monitorados e finalizados pelo analista de suporte, como processos que estejam consumindo 100% do processador, algo muito comum na plataforma Windows (veja a Figura 21);
- Bloqueio de dispositivos (USB, DVD, modem 3G, som, entre outros), de forma simples, com a possibilidade de criar exceções por usuários (veja a Figura 22);



Figura 19. Bloqueio na tentativa de abrir o Windows Media Player

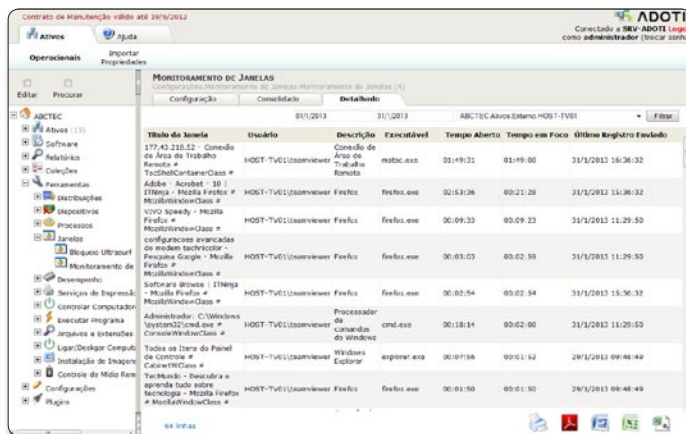


Figura 20. Monitoramento de aplicativos, diferenciando o tempo aberto e o tempo em foco

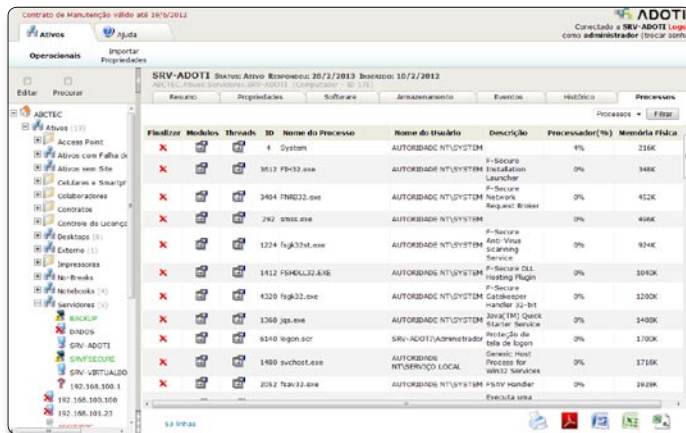
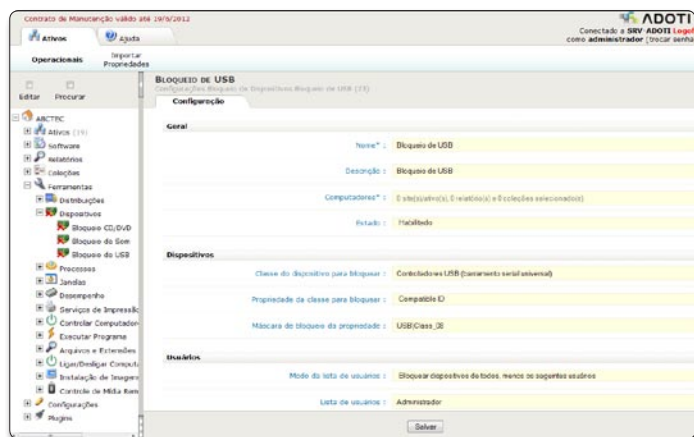


Figura 21. Monitoramento de aplicativos, diferenciando o tempo aberto e o tempo em foco

- Controle de mídias removíveis é um complemento do Bloqueio de Dispositivos, pois, além de permitir o bloqueio de leitura ou escrita em unidades de discos removíveis, permite cadastrar

- Outra funcionalidade muito interessante desta ferramenta é a possibilidade de auditar o nome e o conteúdo dos arquivos transferidos do computador para qualquer unidade de disco removível, auxiliando na redução de riscos de vazamento de informações;



Controle de Manutenção v0.01.2012

Atores

Operacionais

Ativos

Importar impressoras

ADOIT

Controlado a Nível ADOIT Local
como administrador (selecione senha)

Atividade de Manutenção

Configurar o Controle de Manutenção, Atividade de Manutenção

Configuração **Relatório**

	01/12/2012	31/12/2012			
					+ Filtros
Atividade	Data	Computador	Usuário	Arquivo Acessado	Excluído?
Ativos	6/9/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	H0ST-TLMT02\Administrador E:\Montanhas azul.jpg		0
Ativos	6/9/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	H0ST-TLMT02\Administrador E:\Web6a.jpg		0
Computers	6/9/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	H0ST-TLMT02\Administrador E:\ViewServer_Setup_2012.exe		0
Peripherals	6/9/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	H0ST-TLMT02\Administrador E:\usb.txt		0
Distributions	6/9/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Dispositivos	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Processos	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Janelas	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Desenvolvimento	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Serviços de rede	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Controlar Computação	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Executar Programa	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Arquivos e Pastas	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Liga/Desliga Computador	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Instalação de Impressora	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Controle de Mídia Removível	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Auxiliar de Mídia Removível	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Configurações	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Ativos	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0
Ativos	6/11/2012	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09	ABCTEC-Ativos.Desktops.Comercial-HOST-11:01:09		0

Figura 23. Auditoria de todos os arquivos transferidos para unidades de disco removíveis

- [illegible]

Figura 24. Clicando com o botão direito, o administrador requisita ao usuário o acesso via VNC



Figura 25. Requisição apresentada na tela do usuário

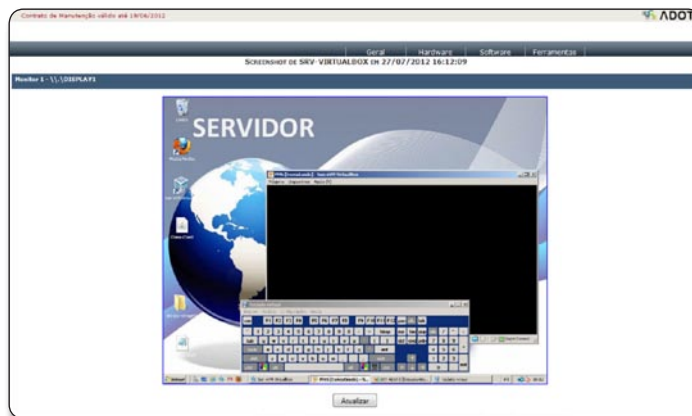


Figura 26. Imagem instantânea da tela do computador acessado

- Ligar/Desligar computadores remotamente. Ferramenta que pode ser usada de forma programada para desligar ou ligar computadores específicos em dias e horários definidos pela organização (veja a **Figura 27**);
- A distribuição de software permite a instalação remota e agendada de qualquer software e até de atualizações de BIOS nos

computadores com agente ADOTI instalado. A distribuição dos pacotes pode ser realizada via HTTP ou por meio de um diretório compartilhado na rede, ao passo que o agendamento pode ser realizado dentro de um intervalo definido de dias e horas. As instalações podem ser silenciosas ou com a intervenção dos usuários (veja a **Figura 28**);

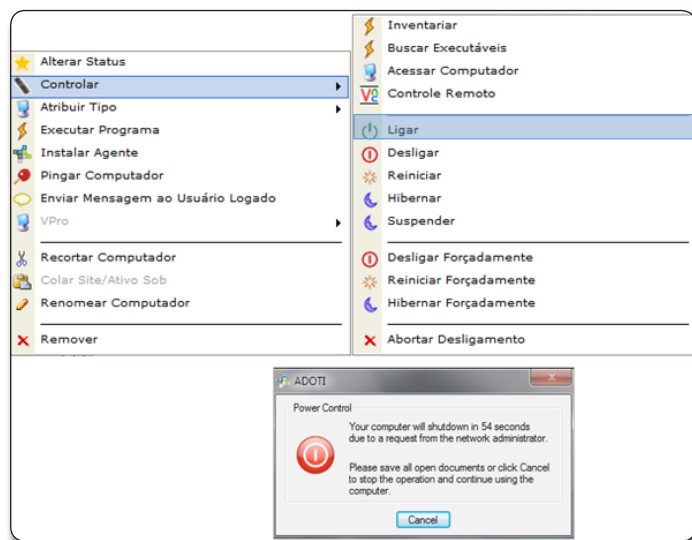


Figura 27. Menu com as opções disponíveis e um exemplo de mensagem para desligamento

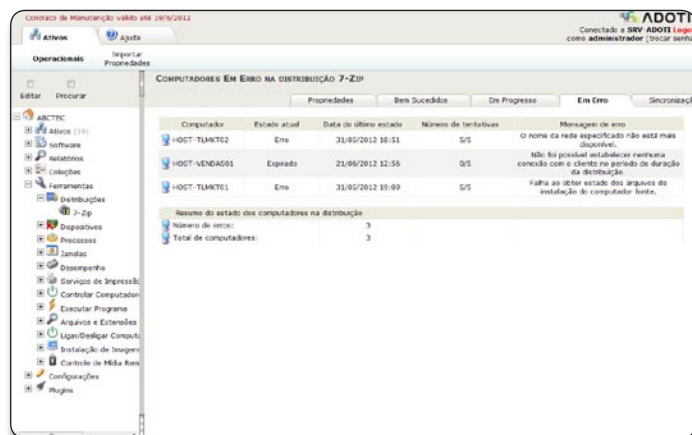


Figura 28. Exemplo de instalação do software TeamViewer realizada remotamente

- Envio de alertas por e-mail de qualquer evento que seja monitorado pelo agente do ADOTI, seja um evento de hardware, seja de software. Entre os principais tipos de eventos, os mais importantes para a gestão preventiva e proativa da TI está no acompanhamento e na identificação imediata das alterações de hardware em qualquer computador na rede, instalação ou remoção de programas, erros de discos detectados pelo HDD-Smart, falhas de dispositivos e vencimento de datas de manutenção programadas, como a de no-breaks ou ainda datas relacionadas a vencimentos de contratos ou prazos de garantia de um equipamento, por exemplo. No exemplo da **Figura 29** temos um alerta informando sobre alterações de software em um determinado computador:

- Sensores de hardware permitem que sejam monitorados itens como temperatura do processador, do disco, velocidade dos coolers de refrigeração e até a abertura de tampa do gabinete do computador, com possibilidade de receber alertas por e-mail (veja a **Figura 30**);

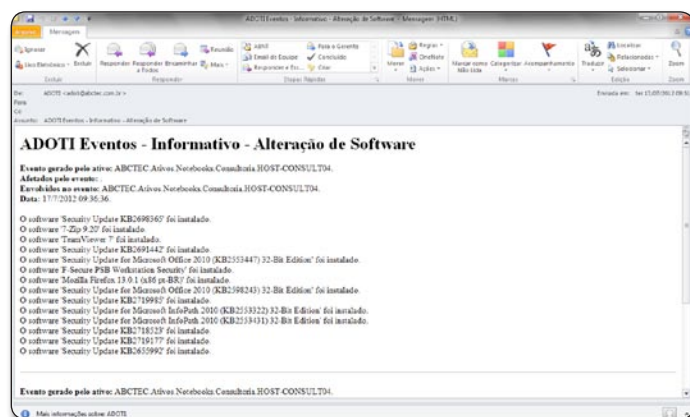


Figura 29. Alerta informando sobre alterações de software em um computador

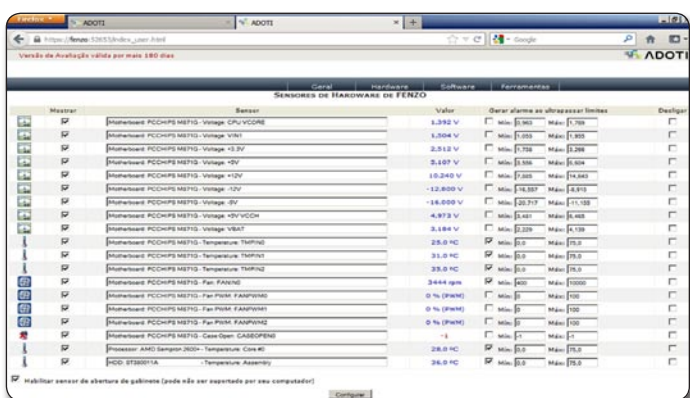


Figura 30. Sensores de hardware, com possibilidade de determinar intervalos para alarmes



Conclusão

Ainda teríamos uma série de ferramentas e possibilidades de suas aplicações a serem descritas tanto para a solução open source CACIC como para a solução licenciada ADOTI, mas o principal objetivo deste artigo foi apresentar os reais problemas que a falta de um sistema de gestão de ativos pode causar a uma empresa e principalmente ao departamento de TI, que hoje se vê responsável por trazer alguma resposta a esta questão. Dessa forma, apresentar estas soluções acessíveis a qualquer organização nos traz a satisfação de poder colaborar efetivamente para o desenvolvimento dos nossos leitores.

Mais importante do que escolher uma solução, é saber tirar dela o máximo proveito para que o investimento seja rapidamente ressarcido e principalmente reconhecido pelos demais setores da organização, de forma que possam entender que a TI está ali não apenas para consertar computadores e reinstalar sistemas, mas para prover soluções que auxiliem as atividades da organização e, por que não, contribuir com ferramentas que possam auxiliar na mitigação dos riscos à segurança da organização, garantir a disponibilidade dos recursos e também reduzir os custos operacionais. A TI precisa mostrar que esta área não é um ralo que apenas suga dinheiro na aquisição de equipamentos e softwares caríssimos, mas um apoio que pode sustentar todas as áreas na busca de melhores resultados.



Roberto Henrique

roberto@abctec.com.br

É Analista de Segurança da Informação na ABCTec, com 14 anos de experiência na área de TI (Suporte, Gestão, Consultoria), especializado em análise de vulnerabilidades e no tratamento de incidentes de segurança da informação. Formado em Análise e Desenvolvimento de Sistemas e atualmente cursando a Pós-Graduação em Investigação de Fraudes e Forense Computacional, possui as certificações ISFS ISO/IEC 27002 Certified, F-Secure Certified Technical Professional - FSCTP, D-Link DBC, Microsoft MCP/MCDST. Escreve artigos para sites e revistas especializadas em Tecnologia e Segurança da Informação. Foi membro do Núcleo de TI do CIESP - São Bernardo do Campo e atualmente membro do Comitê Técnico ABNT/CB21:CE27 sobre Segurança da Informação.



Site da ADOTI

<http://www.adoti.com.br/>

Site da CACIC no Portal do Software Público

http://www.softwarepublico.gov.br/ver-comunidade?community_id=3585

Norma ABNT NBR ISO/IEC 27002:2005

<http://www.abntcatalogo.com.br/norma.aspx?ID=1532>

Blog do autor

<http://abctec.blogspot.com>



Conhecimento faz diferença!

engenheria de software magazine

Edição 29 :: Ano 3

Automação de Testes

Definição, preocupações e custo

+ de 290 vídeos para assinantes

Faça um upgrade em sua carreira

Em um mercado cada vez mais focado em qualidade, ter conhecimentos aprofundados sobre requisitos, metodologia, análises, testes, entre outros, pode ser a diferença entre conquistar ou não uma boa posição profissional. Sabendo disso a DevMedia lançou uma publicação totalmente especializada em Engenharia de Software. Todos os meses você pode encontrar artigos sobre Metodologias Ágeis; Metodologias tradicionais (document driven); ALM (application lifecycle); SOA (aplicações orientadas a serviços); Análise de sistemas; Modelagem; Métricas; Orientação à Objetos; UML; testes e muito mais. **Assine Já!**



Faça já sua assinatura digital! | www.devmedia.com.br/es

Migração de dados: metodologias e abordagens host-based – Parte 1

Conheça as técnicas e metodologias adotadas para migração de dados corporativos em ambiente Linux



ESTE ARTIGO FAZ PARTE DE UM CURSO

Existem muitas técnicas utilizadas em migrações de dados corporativos, podendo variar desde procedimentos executados no Storage, migrações de dados via *Fabric* (*SAN – Storage Area Network*), migrações via Host/Sistema Operacional ou até mesmo via gerenciadores de volume ou aplicação. Cada uma das abordagens repercute em ações e pré-requisitos que devem ser levados em consideração para o sucesso da migração dos dados.

Independentemente da técnica ou da metodologia que será empregada no processo, existe um fator que pode inviabilizar ou impactar o procedimento, que é a circunstância em que o dado se encontra armazenado. Normalmente, muitos administradores começam a planejar ou definir alguns aspectos da migração de dados sem ao menos entender como o dado encontra-se gerenciado e hospedado logicamente.

O dado a ser migrado pode estar sob um gerenciador de volumes como o LVM (*Logical Volume Manager*), que possibilita migração online, sem impacto para a aplicação, porém o dado também pode estar armazenado em um *File System* estático (*/dev/sda1*), significando que haverá necessidade de efetuar uma parada da aplicação em algum momento para gerar a consistência dos dados, de modo que parte do processo ou até mesmo todo ele precise ser executado de forma offline. Em outros mo-

Resumo DevMan

Metodologias e abordagens host-based – Parte 1:

Neste artigo serão apresentadas algumas estratégias e metodologias utilizadas para migração de dados corporativos de ambientes Enterprise Linux adotadas pelo mercado corporativo de TI, mais especificamente o segmento Datacenter, abordando desde o processo de classificação dos dados a serem migrados, até as estratégias que podem ser empregadas para movimentar os dados da forma mais segura e confortável, respeitando as necessidades e as diretrizes do seu negócio.

Em que situação o tema é útil:

Este tema é especialmente útil em migrações de dados quando o próprio host atua como movimentador de dados em cenários diversos, nos quais segurança, tempo e conforto operacional sejam levados em consideração.

mentos o dado também pode estar armazenado em um *RAW device* e requerer algum procedimento específico na aplicação para ser migrado adequadamente.

Esses e outros pontos importantes que devem ser levados em consideração para realizar uma migração de dados bem sucedida serão abordados neste artigo, cobrindo algumas metodologias adotadas pelo mercado de TI de várias corporações.

O artigo também visa dar ênfase às situações de migração executadas via host, utilizando ferramentas do próprio sistema operacional ou do gerenciador LVM. Em futuros artigos serão abordadas estratégias para migração de dados via *Storage* ou *Fabric* (*SAN – Storage Area Network*), entre outras. Em razão de sua vasta

abrangência, o artigo foi dividido em duas partes. Na primeira parte contida nessa edição, o leitor terá contato com o ambiente proposto, sendo este composto por três servidores Linux com alguns *File System* e Volumes LVM. Cada servidor encontra-se em um cenário característico que viabilizará as estratégias de migração que serão apresentadas no artigo.

Na segunda parte deste artigo, o leitor será apresentado aos procedimentos técnico-operacionais, incluindo os comandos que serão utilizados para migrar os dados de nossos três servidores do ambiente proposto, além de cobrir aspectos de validação de migração de dados e *Clean-Up* do ambiente envolvido.

Classificação dos dados

Quando pensamos em dados pessoais, não parece uma tarefa tão difícil migrar dados de um dispositivo para outro, porém, quando pensamos em dados corporativos, várias questões entram em jogo, como a janela permitida de *downtime* do ambiente, SLA contratado, impacto a clientes externos, comunicação, gerência de mudanças, necessidade e tempo de rollback em caso de falhas ou comportamentos não esperados, entre outros fatores que também influenciam a decisão do melhor método para efetuar a migração dos dados. Em muitos desses casos, nem sempre o método inicialmente definido se mostra o mais adequado quando contextualizamos o custo-benefício da estratégia escolhida.

Independentemente da técnica a ser escolhida para migrar os seus dados, o primeiro passo de uma migração bem sucedida consiste em classificar os dados a serem migrados. É necessário compreender questões importantes, como: se os dados encontram-se armazenados em RAW Devices, se estão encapsulados sob algum gerenciador de volumes (LVM), se estão armazenados em *File Systems* dinâmicos ou estáticos, se são *hard partitions*, entre outras. Sem esse processo não é possível saber qual técnica pode ser adotada para migrar os dados da forma mais adequada às necessidades do negócio da companhia.

Usar ou não um gerenciador de volumes?

Em linhas gerais, uma das mais importantes tarefas durante a migração de dados ocorre antes mesmo de os dados serem migrados, que seria no momento inicial, quando os dados são armazenados. Algumas perguntas que visam balizar essa dúvida seriam “Qual o melhor *File System* para armazenar esse tipo de informação, dadas as características do meu dado?” e “Devo utilizar algum gerenciador de Volumes?”

Para simplificar esse entendimento, vamos imaginar que um dado servidor recebeu um novo disco/LUN (*Logical Unit Number*), sendo este contido em um Storage e conectado ao host através de uma SAN (*Storage Area Network*). Essa nova LUN pode ser utilizada de várias formas.

Num primeiro momento, podemos entrar no utilitário FDISK, criar uma nova partição do tipo “Linux”, criar um *File System* nessa partição e montar a nova área em disco da seguinte forma: `mount -t ext3 /dev/sdf1 /mountpoint`. Esse método é conhecido como *Hard Partition* e impossibilita migrações de dados online.

Outra possibilidade seria utilizar essa nova LUN sob um gerenciador de volumes, no caso do Linux, o LVM (*Logical Volume Manager*). Nesse cenário, várias opções de migração de dados online são permitidas, como espelhamento e movimentação de PV (*Physical Volumes*), que serão abordados mais adiante.

LVM

LVM é uma tecnologia que permite utilizar um disco físico ou LUN de forma mais interessante e atraente, oferecendo melhor nível de aproveitamento do espaço em disco e gestão da área de armazenamento e, consequentemente, garantindo segurança e desempenho oferecendo recursos como espelhamento dos dados (RAID 1) e operações de redimensionamento dinâmico de volumes. Além disso, a utilização de um gerenciador de volumes permite manobras dinâmicas, como migração de dados e movimentação de estruturas lógicas com a aplicação em estado online.

O LVM trabalha com uma unidade primária chamada *Physical Extents* (PE), normalmente tendo o tamanho inicial de 4 MB. Quando disponibilizamos uma nova área em disco para ser utilizada com LVM, durante o processo de inicialização essa nova área é dividida logicamente em tamanhos iguais (*Extents*). Quando um novo volume de 10 GB é criado no LVM, este precisa alocar os 10 GB de armazenamento em “x” *Extents* de 4 MB. Neste caso, esse volume seria composto por 2560 *Extents*. Em operações como criação ou redimensionamento de volumes, o administrador pode informar o tamanho desejado do volume em *Extents* (PE) ou em tamanho físico, como Megabytes ou Gigabytes. Adicionalmente, vários comandos LVM mostram as informações na unidade *Extents*. Deste modo, é interessante que o leitor crie familiaridade com esta unidade para melhor compreensão desses comandos.

Para aqueles leitores que não estão familiarizados com o LVM e precisam utilizar um disco nele, primeiramente é necessário criar um PV (*Physical Volume*), que numa analogia seria equivalente a uma partição. Esse processo cria a estrutura lógica que permite endereçar o espaço em disco em endereçamentos LVM, conforme o Estágio 1 da Figura 1.



Figura 1. Etapas para utilização de um novo disco no LVM

Após essa etapa, é necessário criar o *Volume Group* (VG), que representa uma estrutura lógica com propósitos de associar um ou mais *Physical Volumes* (PVs), conforme visto no Estágio 2 da Figura 1. Vamos imaginar que existam dois discos físicos a serem

espelhados. Nesse cenário, teremos dois PVs (um em cada disco físico), os quais irão compor um único VG (*Volume Group*).

Por fim, temos os *Logical Volumes* (LVOLs), que representam os *File Systems* que serão criados na estrutura LVM, conforme observado no Estágio 3 da **Figura 1**. Os LVOLs são criados dentro do VG, que é composto por um ou mais PVs.

Dessa forma, os passos macro para realizar a adição um novo disco no LVM seriam:

- Criar uma partição do tipo 8E (Linux LVM) via *FDISK* ou outro utilitário;
- Criar um *Physical Volume* (PV) no disco;
- Adicionar o novo *Physical Volume* (PV) ao *Volume Group* (VG);
- Criar um novo *Logical Volume* (LVOL) no *Volume Group* (VG).

Quando nenhum Gerenciador de Volumes é utilizado

Nesses casos, normalmente trata-se de um *File System* montado como */dev/sdb1* (*Hard Partition*) ou um *RAW Device*. Quando nenhum gerenciador de volumes é utilizado, não é possível executar migrações de dados entre discos/LUNs de forma online por meio de ferramentas do sistema operacional, ou seja, os dados são migrados necessitando de uma parada da aplicação ou serviço.

Existem algumas soluções e produtos comerciais pagos que permitem o processo online, porém, em linhas gerais, nessa situação é necessário gerar um ponto de consistência dos dados, normalmente o *umount* do *File System* ou parada da aplicação, traduzindo-se em um processo offline. Ademais, algumas aplicações como bancos de dados com propósitos comerciais têm recursos que permitem a migração de dados online, mesmo que os dados não se encontrem hospedados em algum gerenciador de volumes.

Em alguns casos, o dado pode estar armazenado em um *RAW device* utilizado por um banco de dados, como o Sybase, Oracle, entre outros. Nesses casos, existem ferramentas do próprio banco de dados que podem permitir a migração online, no entanto, via sistema operacional, quando os dados não estão encapsulados sobre algum gerenciador de volumes, o processo de migração de dados ocorre de forma offline.

Outro ponto interessante é que preferencialmente devem ser escolhidos *File Systems* dinâmicos para armazenar os dados. *File Systems* dinâmicos são aqueles que permitem operações de *Grow* (expansão) e *Shrink* (redução) de forma dinâmica, ou seja, com a aplicação em estado online. Em casos menos comuns, porém não menos importantes, a LUN origem é maior que a LUN destino, podendo ser necessário executar operações de *Grow* ou *Shrink* antes de migrar os dados.

Quando algum Gerenciador de Volumes é utilizado

Nesses casos, normalmente é um *Logical Volume* (LVOL) pertencente a algum *Volume Group* (VG) LVM. Esse volume pode ser utilizado como *File System* ou *RAW device*. Essa situação é um dos cenários mais interessantes de se trabalhar quando se objetiva disponibilidade e conforto operacional, pois manobras como espelhamento, migração de PV (*Physical Volumes* LVM), entre

outras, podem ser utilizadas sem necessidade de efetuar parada da aplicação. Quando optamos por utilizar algum gerenciador de volumes, a migração ocorre na camada do gerenciador, ou seja, uma operação *Block Level*. Já quando não utilizamos um gerenciador de volumes, a cópia ocorre em cenário *File Level*, ou seja, é influenciada pela quantidade de arquivos e diretórios existentes no *File System*, e a princípio é menos performática, pois mais camadas devem ser acessadas para chegar até o dado (LVM e *File System*). A **Nota do DevMan 1** traz mais detalhes sobre as operações *Block Based* e *File Based*.



Nota do DevMan 1

Operações File Based x Block Based

As operações *Block Based* se comportam como cópias de *RAW Devices* por meio do comando *DD*. O tempo de cópia não é influenciado pela quantidade de arquivos a serem migrados. A ferramenta utilizada na cópia requisita os blocos existentes nos volumes origem e os copia para os volumes destino.

Já as operações *File Based* são influenciadas pela quantidade de arquivos existentes no sistema de arquivos a ser migrado. A ferramenta utilizada na cópia requisita os arquivos existentes no sistema de arquivos origem e os copia para o sistema de arquivos destino.

Para melhor compreensão do leitor entre a diferença prática entre operações *Block Level* e *File Level*, imagine um *File System* de 10 GB com 232 milhões de arquivos muito pequenos na casa de alguns poucos *Bytes*, como é muito comum nos ambientes de Telecom (arquivos CDR).

Podemos afirmar certamente que em uma migração *Block Level* o processo levaria alguns minutos, enquanto que numa cópia *File Level* o processo certamente levaria algumas horas, pois os 232 milhões de ponteiros dos *File Systems*, que na verdade referenciam arquivos e diretórios, precisam ser traduzidos para posições físicas de disco. Nos dois casos, são apenas 10 GB a serem migrados. Não é incomum que administradores já tenham vivenciado experiências de migrações de dados em que alguns poucos *Gigabytes* de informação demoraram mais de um dia para serem migrados via operação *File Level* devido à baixa taxa de transferência (*Throughput*) de cópia, parametrização indevida de *File System*, opções de montagem impróprias para necessidades da aplicação, entre outras situações.

Nas operações *Block-Level*, não importa se o dado se encontra hospedado em um *RAW device* ou um *File System* de 232 milhões de arquivos. A única informação que importa é a volumetria a ser migrada, ou seja, os zeros (0) e uns (1) armazenados em disco. O tempo de cópia não sofre influência da quantidade de arquivos existentes.

Cabe ressaltar que também existem alguns gerenciadores de volumes comerciais voltados ao segmento *Datacenter* que oferecem outros recursos mais avançados direcionados ao ambiente Enterprise. Alguns produtos muito utilizados em ambiente Linux são o Oracle ASM (*Oracle Automatic Storage Management*), Symantec Storage Foundation, entre outras soluções interessantes.

Multipathing

Uma situação muito comum em ambientes corporativos e SAN de porte Enterprise é a utilização de algum software de Multipathing nos sistemas operacionais para acesso dos servidores ao Storage onde o dado reside. A **Nota do DevMan 2** traz mais detalhes sobre a tecnologia Multipathing.



Nota do DevMan 2

Multipathing

Multipathing é uma tecnologia utilizada quando existe mais de uma controladora ou caminho de acesso a um disco/LUN. Neste caso, um mesmo dispositivo pode ser visto e reconhecido pelo sistema operacional por dois ou mais caminhos de acesso, `/dev/sda` e `/dev/sds`, por exemplo. Porém, para que o sistema operacional tire proveito da redundância dos múltiplos caminhos de acesso de um disco, é necessário configurar algum software de multipath no sistema operacional.

Existem softwares freeware e outros proprietários distribuídos pelo fabricante do Storage. Cada um oferece vantagens específicas a cada uma das necessidades do negócio.

A utilização de Multipathing faz sentido quando o sistema operacional reconhece os discos/LUNs por mais de um caminho lógico. Em ambientes de alta disponibilidade é comum encontrar servidores com duas ou mais controladoras Fiber Channel (HBAs) e o Storage possuir mais de uma controladora também. Neste cenário, um único disco/LUN é reconhecido em nível de sistema operacional por mais de um caminho lógico. Exemplo: `/dev/sdb`, `/dev/sdc`, `/dev/sdd` e `/dev/sde`.

Até então, se nenhum software de *Multipathing* for utilizado, na prática, a redundância de acesso ao disco não existe. O efeito prático para o sistema operacional é como se houvessem quatro discos diferentes com um único caminho de acesso, em vez de um único disco com quatro caminhos de acesso.

A tecnologia de *Multipathing* permite que o sistema operacional possa “entender” que existem vários caminhos de acesso a um mesmo disco físico. Existem softwares nativos de cada sistema operacional e também softwares proprietários em caso de fabricantes que comercializam *Storages* de médio e grande porte, como HP, EMC e Hitachi. Cada um com comandos, procedimentos e requisitos muito específicos.

Cada modelo de *Storage* permite ou não características como balanceamento de carga no acesso aos dados que possibilitam trabalhar de forma ativa/ativa quando há mais de uma controladora para acesso aos discos. Também existem modelos de *Storage* que possuem duas controladoras, porém estas funcionam somente em regime ativo/passivo, ou seja, a segunda controladora só é utilizada em situações de falha da primeira controladora.

Dessa forma, o software de *Multipathing* somente pode trabalhar nas configurações suportadas pelo modelo do *Storage* em questão, variando normalmente de forma macro em ativa/ativa ou ativa/passiva, sendo possível parametrizar esse tipo de acesso ao disco normalmente nas seguintes variantes:

- **Round Robin:** Modalidade em que todas as requisições de I/O são distribuídas entre os possíveis caminhos de acesso ao dado. O acesso 1 é feito por meio da controladora 1, o acesso 2 é feito por meio da controladora 2, o acesso 3 por meio da controladora 1, o acesso 4 pela controladora 2, e assim por diante;
- **Preferred:** Nessa modalidade, todas as requisições de I/O são direcionadas a uma controladora que o administrador define como preferencial. A segunda controladora só é utilizada em caso de falha da primeira;
- **Custom:** Modalidade específica do fabricante do *Storage* em função da tecnologia utilizada. Alguns fabricantes oferecem em seus softwares de *Multipathing* recursos como verificação de latência e tempo de resposta de acesso ao disco/LUN, ou seja, o caminho que responder com menor latência (mais rápido) é o utilizado para acessar a informação. Já existe outra modalidade customizada em que a carga de consumo da controladora é verificada. A controladora do *Storage* que estiver com menor consumo atende às requisições de I/O. Existem inúmeros outros algoritmos de balanceamento de carga variando de fabricante para fabricante com os mais diferentes propósitos, como beneficiar o processo de escrita ou de leitura, bem como privilegiar o acesso sequencial ou randômico aos dados.

Para melhor compreensão didática, os aspectos de *Multipathing* não serão abordados em detalhes nesse artigo, pois, em função do fabricante ou tecnologia que o leitor dispuser, os comandos requeridos podem mudar de forma que dificultaria a compreensão técnico-didática desse artigo. Cada fabricante mantém sua própria documentação, além de existirem diversas outras fontes e artigos na Internet sobre o assunto Multipathing. A tecnologia *Multipathing* será abordada em detalhes em futuros artigos da Infra Magazine.

Vale lembrar aos leitores menos experientes que LVM (*Logical Volume Manager*) não deve ser associado a *Multipathing*. É possível utilizar *Multipathing* e não utilizar LVM, por exemplo. Fazendo uma analogia, *Multipathing* está relacionado à forma ou modo de acesso físico ao dado, ao passo que o LVM está relacionado à forma lógica como o dado é armazenado em disco.

Cenário de testes

Para elaboração desse artigo, vamos imaginar que nosso ambiente a ser migrado seja composto de três servidores Linux, com os hostnames `RJOSRVUUX01`, `RJOSRVUUX02` e `RJOSRVUUX03`, respectivamente. Cada *host* tem seus dados armazenados de uma forma peculiar que será avaliada e terá um método de migração próprio a ser definido adiante, no decorrer do artigo. A **Figura 2** ilustra de forma macro o cenário que será tratado nos próximos passos do artigo.

Para obtermos as informações necessárias que viabilizem a avaliação de como os dados encontram-se hospedados nos hosts e Sistemas Operacionais envolvidos, é necessário executar comandos específicos do gerenciador de volumes ou do sistema operacional, conforme será explicado adiante.

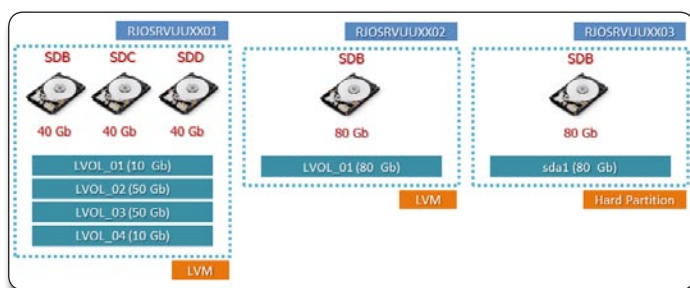


Figura 2. Dados a serem migrados

Para cada um dos casos a serem analisados, pode ser necessário executar comandos específicos mediante o resultado obtido na análise inicial. Quando os dados estiverem sob gerência do LVM, será necessário avaliar *outputs* de comandos como *pvdisplay*, *lvdisplay* ou *vgdisplay*. Já em situações em que os dados estejam hospedados em *hard partitions* (*/dev/sda1*), não é necessário avaliar os comandos LVM, apenas os comandos específicos do sistema operacional. Por isso a etapa de levantamento é muito específica e peculiar a cada ambiente.

Mapeamento do servidor RJOSRVUXX01

Na **Listagem 1**, temos todos os passos associados com o mapeamento do servidor RJOSRVUXX01, que tem seus *File Systems* a serem migrados sob gerência do LVM (*Logical Volume Manager*). Como pode ser observado, serão executados comandos específicos de LVM para efetuar o mapeamento dos dados a serem migrados.

Para facilitar a compreensão da sequência de comandos exibida na **Listagem 1**, a dividimos em várias partes e agora explicaremos cada uma separadamente:

1. Conforme observado na **Figura 2**, o host RJOSRVUXX01 tem seus *File Systems* a serem migrados sob gestão de um gerenciador de volumes. Vamos então documentar o cenário no qual iremos trabalhar. Para realizar o mapeamento dos dados é necessário associar cada *File System* que será migrado com seu respectivo LVOL (*Logical Volumes*). Utilizando o comando *df -h* ou outro de sua preferência, verifique os LVOLs e seus respectivos *Mount Points* e VGs (*Volume Groups*).

Como exemplo, a primeira linha do comando *df -h* mostra que o LVOL **lvol_01**, pertence ao Volume Group (VG) **db_vg** e está montado no diretório */srv/db/MYDB*. Tome nota dos demais LVOLs. Em negrito, encontram-se relacionados os nomes dos LVOLs que serão migrados (**lvol-01**, **lvol-02**, **lvol-03** e **lvol-04**) juntamente com os atuais *Mount Points* desses volumes (*/srv/db/MYDB*, */srv/db/data01*, */srv/db/indx01* e */srv/db/arch*);

2. Por meio do comando *pvs*, verifique quais PVs (*Physical Volumes*) fazem parte do VG **db_vg**. Em nosso caso, */dev/sdb1*, */dev/sdc1* e */dev/sdd1*, conforme as marcações em negrito. Além disso, também repare que cada um desses PVs a serem migrados possui 40 GB. Essa informação será utilizada mais adiante no artigo. Até o presente momento, foram mapeados quatro LVOLs hospedados em três PVs;

3. Conforme o comando anterior, todos os PVs já estão devidamente mapeados. Agora vamos registrar como esses PVs encontram-se do ponto de vista físico, ou seja, qual a controladora, target e LUN desses PVs. Essa informação pode ser utilizada em caso de *Rollback*, além de dar uma visão da distribuição física dos discos/LUNs entre os barramentos e os controladores existentes no servidor.

Utilizando o utilitário *lsscsi* (nem sempre instalado por padrão), verifique as LUNs que farão parte da migração e registre o caminho físico associado com cada LUN. Observe os *device paths* marcados em negrito. Nesse caso, temos o disco físico */dev/sdb* associado ao *device path* 1:0:0:0. O disco */dev/sda* não está marcado porque é a área de Boot do servidor;

4. Conforme explicado inicialmente nesse artigo na seção LVM, o LVM trabalha com o conceito de *Extents*. Portanto, é importante registrar não só o tamanho dos PVs em GB, mas também seu tamanho em *Extents* (áreas de 4 MB). Essa informação será útil na criação dos volumes em um momento futuro. Assim, por meio do comando *PVDISPLAY*, verifique a atual parametrização e quantidade de *Extents* alocados para os PVs.

Nesse caso temos os PVs */dev/sdb1*, */dev/sdc1* e */dev/sdd1*, cada um com 10239 *Extents*, o que representa 40 GB. Também podemos ver que os *Extents* (PE) são de 4 MB.

Em alguns casos muito específicos, aplicações *Enterprise*, visando obter melhor desempenho nas operações de I/O, recomendam alterar o valor *default* do tamanho dos *Extents*. Esta configuração pode ser feita no momento da criação das estruturas LVM. A princípio, o tamanho inicial de 4 MB atende muito bem qualquer aplicação. Em nosso caso, os PVs têm *Extents* de 4 MB;

5. Verifique se todos os LVOLs foram devidamente mapeados ou se faltou algum. Até então nos baseamos na saída do comando *df -h*, que não é o melhor método. Poderiam existir LVOLs não montados, ou então os LVOLs poderiam ser RAW devices que não apareceriam na saída desse comando.

O comando adequado para o mapeamento é o *lvs*, que mostra todos os *Logical Volumes* (LVOLs) existentes no VG (*Volume Groups*) **db_vg**. Ao executá-lo, podemos notar que todos os quatro *Logical Volumes* (**lvol_01**, **lvol_02**, **lvol_03** e **lvol_04**) foram devidamente mapeados.

Note também que os *Logical Volumes* têm respectivamente 10 GB, 50 GB, 50 GB e 9,98 GB de tamanho;

6. Como já vimos anteriormente, o LVM trabalha com o conceito de *Extents* (áreas de 4 MB), de forma que é muito importante sempre registrar o tamanho das estruturas LVM não em tamanho (GB ou MB), por exemplo, mas em *Extents*. Assim, em momentos de migração de dados, ao utilizar os tamanhos em *Extents*, sempre garantimos que as novas estruturas LVM serão migradas consistentemente do ponto de vista físico de volumetria.

Dando continuidade à fase de levantamento de informações, através do comando *lvdisplay*, verifique a atual parametrização e quantidade de *Extents* dos LVs. Podemos observar que o **lvol_01** tem 2560 *extents*, o **lvol_02** tem 12800 *Extents*, o **lvol_03** também possui 12800 *Extents*, e, por fim, o **lvol_04** tem 2555 *Extents*.

Essa informação pode ser útil em momentos de *Rollback* e criação de volumes, sendo importante para efeitos de planejamento e organização dos dados que serão migrados. Baseado nos dados obtidos

na execução dos procedimentos da **Listagem 1**, concluímos a fase de mapeamento do servidor RJOSRVUXX01 e consolidamos as informações coletadas, conforme observado na **Figura 3**.

Listagem 1. Sequência de comandos para mapeamento do host RJOSRVUXX01.

```
##### Trecho 1 #####
sysadmin@rjosrvuux01:~$ df -h | tail -4
/dev/mapper/db_vg-lvol_01 9.9G 4.5G 5.4G 45% /srv/db/MYDB
/dev/mapper/db_vg-lvol_02 50G 22G 28G 44% /srv/db/data01
/dev/mapper/db_vg-lvol_03 50G 22G 28G 44% /srv/db/indx01
/dev/mapper/db_vg-lvol_04 9.9G 8.1G 1.8G 82% /srv/db/arch

##### Trecho 2 #####
sysadmin@rjosrvuux01:~$ sudo pvs
PV      VG      Fmt Attr PSize PFree
/dev/sdb1 db_vg   lvm2 a-  40.00g 8.00m
/dev/sdc1 db_vg   lvm2 a-  40.00g  0
/dev/sdd1 db_vg   lvm2 a-  40.00g  0

##### Trecho 3 #####
sysadmin@rjosrvuux01:~$ lsccsi --size
[0:0:0:0] disk ATA VBOX HARDDISK 1.0 /dev/sda 8.58GB
[1:0:0:0] disk ATA VBOX HARDDISK 1.0 /dev/sdb 42.9GB
[2:0:0:0] disk ATA VBOX HARDDISK 1.0 /dev/sdc 42.9GB
[3:0:0:0] disk ATA VBOX HARDDISK 1.0 /dev/sdd 42.9GB

##### Trecho 4 #####
sysadmin@rjosrvuux01:~$ sudo pvdisplay /dev/sdb1 /dev/sdc1 /dev/sdd1
--- Physical volume ---
PV Name      /dev/sdb1
VG Name      db_vg
PV Size      40.00 GiB / not usable 3.00 MiB
Allocatable  yes
PE Size      4.00 MiB
Total PE     10239
Free PE      2
Allocated PE 10237
PV UUID      LWfb6K-5Za8-5X1n-3CD4-iTmw-N126-s4bUI3

--- Physical volume ---
PV Name      /dev/sdc1
VG Name      db_vg
PV Size      40.00 GiB / not usable 3.00 MiB
Allocatable  yes (but full)
PE Size      4.00 MiB
Total PE     10239
Free PE      0
Allocated PE 10239
PV UUID      rle6sc-zB55-90Fq-KN0f-kem0-J17a-KitlgB

--- Physical volume ---
PV Name      /dev/sdd1
VG Name      db_vg
PV Size      40.00 GiB / not usable 3.00 MiB
Allocatable  yes (but full)
PE Size      4.00 MiB
Total PE     10239
Free PE      0
Allocated PE 10239
PV UUID      XnttFc-mMLh-il7g-VelV-yVVd-Hqsy-F4X456

##### Trecho 5 #####
sysadmin@rjosrvuux01:~$ sudo lvs db_vg
LV      VG Attr LSize Origin Snap% Move Log Copy% Convert
lvol_01 db_vg -wi-ao 10.00g
lvol_02 db_vg -wi-ao 50.00g
lvol_03 db_vg -wi-ao 50.00g
lvol_04 db_vg -wi-ao 9.98g

##### Trecho 6 #####
sysadmin@rjosrvuux01:~$ sudo lvdisplay /dev/db_vg/lvol_0*
--- Logical volume ---
LV Name      /dev/db_vg/lvol_01
VG Name      db_vg
LV UUID      qRY5U9-DiYo-YyCC-meP7-DkQd-7dy9-Wndqk9
LV Write Access read/write
LV Status     available
# open        1
LV Size       10.00 GiB
Current LE    2560
Segments      1
Allocation    inherit
Read ahead sectors auto
- currently set to 256
Block device  252:2

--- Logical volume ---
LV Name      /dev/db_vg/lvol_02
VG Name      db_vg
LV UUID      rvCeKG-UdyT-MmaT-xCv3-LnQe-2r35-s5hOuG
LV Write Access read/write
LV Status     available
# open        1
LV Size       50.00 GiB
Current LE    12800
Segments      2
Allocation    inherit
Read ahead sectors auto
- currently set to 256
Block device  252:3

--- Logical volume ---
LV Name      /dev/db_vg/lvol_03
VG Name      db_vg
LV UUID      zLMppE-6Lpl-sE4V-XMvO-etZs-zzoQ-CvymTE
LV Write Access read/write
LV Status     available
# open        1
LV Size       50.00 GiB
Current LE    12800
Segments      2
Allocation    inherit
Read ahead sectors auto
- currently set to 256
Block device  252:4

--- Logical volume ---
LV Name      /dev/db_vg/lvol_04
VG Name      db_vg
LV UUID      KfMVnC-S5w4-2JQT-mlgv-VLH5-pGG1-J32CFL
LV Write Access read/write
LV Status     available
# open        1
LV Size       9.98 GiB
Current LE    2555
Segments      1
Allocation    inherit
Read ahead sectors auto
- currently set to 256
Block device  252:5
```

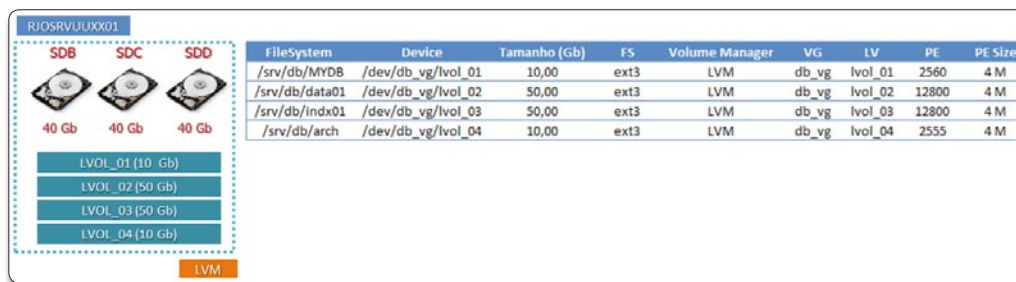


Figura 3. Dados a serem migrados do servidor RJOSRVUXX01

Mapeamento do servidor RJOSRVUXX02

Na **Listagem 2**, temos todos os passos associados com o mapeamento do servidor RJOSRVUXX02, que possui um único *File System* a ser migrado sob gerência do LVM. Como pode ser observado, serão executados comandos específicos de LVM para efetuar

o mapeamento dos dados que serão migrados.

Para facilitar a compreensão da sequência de comandos exibida na **Listagem 2**, a dividimos em várias partes e agora explicaremos cada uma separadamente:

1. Conforme observado na **Figura 2**, o host RJOSRVUXX02 também possui seus *File Systems* sob gestão de um gerenciador de volumes. Para concluir a etapa de levantamento dos dados, é necessário associar cada *File System* que será migrado com seu respectivo LVOL. Utilizando o comando `df -h` ou outro de sua preferência, verifique os LVOLs e seus respectivos *Mount Points* e VGs.

Podemos notar que existe apenas o lvol **lvol_01** associado ao *Mount Point* `/srv/app`;

2. Agora vamos documentar os PVs responsáveis por armazenar o **lvol_01** por meio do comando `pvs`. Notamos que o VG **app_vg** só possui um PV (`/dev/sdb1`), de 80 GB;

3. Nesse momento, o único PV já está devidamente mapeado. Agora vamos registrar como esse PV encontra-se do ponto de vista físico, ou seja, qual a controladora, target e LUN do mesmo. Essa informação pode ser utilizada em caso de *Rollback* e oferece uma visão clara da distribuição física dos discos/LUNs entre os barramentos e controladores existentes no servidor. Utilizando o utilitário `lsscsi`, verifique o caminho físico da LUN que será migrada. Neste caso temos o disco físico `/dev/sdb` associado ao *device path* `1:0:0:0`. O disco `/dev/sda` não está marcado em negrito, pois é a área de Boot do servidor;

4. Por meio do comando `pvdisplay`, observe o número de *Extents* (áreas de 4 MB) associado ao PV `/dev/sdb1`. Podemos notar que ele apresenta 20479 *Extents*. Para efeitos de planejamento de migração, é muito importante registrar essa informação, que além de poder ser útil em caso de *Rollback*, ajuda a ter uma visão clara sobre os PVs que estão sendo tratados durante processo de migração de dados;

5. Até o presente momento utilizamos o comando `df -h` para verificar os LVOLs que serão migrados. Porém, podemos ter mais algum LVOL não montado ou um *raw device*, por exemplo, que não seria identificado por esse comando. Dessa forma, verifique se todos os LVOLs foram mapeados ou se faltou mais algum utilizando o `lvs`.

Com isso podemos confirmar que no VG **app_vg** só existe o **lvol_01** com o tamanho de 79,90 GB;

6. Agora, por meio do comando `pvdisplay`, vamos observar a atual parametrização do LVOL **lvol_01** e sua quantidade de *Extents*.

Listagem 2. Sequência de comandos para mapeamento do host RJOSRVUXX02.

```
##### Trecho 1 #####
sysadmin@rjosrvuux02:~$ df -h | tail -1
/dev/mapper/app_vg-lvol_01 79G 54G 25G 68% /srv/app

##### Trecho 2 #####
sysadmin@rjosrvuux02:~$ sudo pvs
PV          VG      Fmt Attr   PSize PFree
/dev/sdb1   app_vg  lvm2 a-    80.00g 96.00m

##### Trecho 3 #####
sysadmin@rjosrvuux02:~$ sudo lsscsi --size
[0:0:0:0] disk ATA VBOX HARDDISK 1.0 /dev/sda 8.58GB
[1:0:0:0] disk ATA VBOX HARDDISK 1.0 /dev/sdb 85.8GB

##### Trecho 4 #####
sysadmin@rjosrvuux02:~$ sudo pvdisplay /dev/sdb1
--- Physical volume ---
PV Name           /dev/sdb1
VG Name           app_vg
PV Size           80.00 GiB / not usable 3.00 MiB
Allocatable       yes
PE Size           4.00 MiB
Total PE          20479
Free PE           24
Allocated PE      20455
PV UUID           BTmBkB-eMIG-XD8R-afM3-la9m-RnDo-o6EWQx

##### Trecho 5 #####
sysadmin@rjosrvuux02:~$ sudo lvs app_vg
LV   VG   Attr LSize  Origin Snap%  Move Log Copy%  Convert
lvol_01 app_vg -wi-ao 79.90g

##### Trecho 6 #####
sysadmin@rjosrvuux02:~$ sudo lvdisplay /dev/app_vg/*
--- Logical volume ---
LV Name           /dev/app_vg/lvol_01
VG Name           app_vg
LV UUID           sR8ilj-IUct-SUeM-d6p9-2pKI-0LxR-SsPdGa
LV Write Access   read/write
LV Status         available
# open            1
LV Size           79.90 GiB
Current LE        20455
Segments          1
Allocation         inherit
Read ahead sectors auto
- currently set to 256
Block device      252:2
```

Essa informação será utilizada mais adiante no artigo. Podemos observar que o `lvol_01` possui 20455 *Extents* (em áreas de 4 MB). A partir da coleta de dados realizada com a execução dos procedimentos da **Listagem 2**, concluímos a fase de mapeamento do servidor `RJOSRVUXX02` e consolidamos as informações coletadas, conforme observado na **Figura 4**.

Mapeamento do servidor `RJOSRVUXX03`

Na **Listagem 3**, temos todos os passos associados com o mapeamento do servidor `RJOSRVUXX03`, que possui um único *File System* a ser migrado. Diferentemente dos outros dois servidores já mapeados (`RJOSRVUXX01` e `02`), o servidor `RJOSRVUXX03` não possui seus discos/LUNs gerenciados pelo LVM. O *File System* encontra-se no estado *Hard Partition*, ou seja, não dinâmico. Sendo assim, serão executados comandos específicos do sistema operacional para efetuar o mapeamento dos dados que serão migrados.

Para facilitar a compreensão da sequência de comandos exibida na **Listagem 3**, a dividimos em várias partes e agora explicaremos cada uma separadamente:

1. Por meio do comando `df -h`, verifique os *File Systems* que serão migrados. Podemos notar que existe somente o device `/dev/sdb1` com o *Mount Point* `/srv/monitor`;
2. Nesse momento, vamos documentar como o disco/LUN `/dev/sdb` encontra-se do ponto de vista físico, ou seja, qual sua controladora, target e LUN. Essa informação pode ser utilizada em caso de *Rollback* e oferece uma visão clara da distribuição física dos discos/LUNs entre os barramentos e controladores existentes no servidor.

Utilizando o utilitário `lsscsi`, verifique o caminho físico da LUN que será migrada. Neste caso, temos o disco físico `/dev/sdb` associado ao device path `1:0:0:0`. O disco `/dev/sda` não está marcado porque é a área de Boot do servidor;

3. Como o device `/dev/sdb1` não está sob gerência do LVM, precisamos utilizar ferramentas do próprio sistema operacional para coletar dados para o planejamento do processo de migração. Assim, por meio do utilitário `FDISK`, verifique a atual parametrização do device, como tamanho e tipo de partição.

Neste caso, o disco/LUN `/dev/sdb` possui 85.9 GB e somente uma partição (1) com o Id 83 do tipo Linux. Outro ponto importante é verificar as informações de geometria física, como o tamanho em cilindros ou o tamanho em quantidade de blocos de 512 bytes, que serão utilizadas para recriar as partições no novo disco/LUN com as mesmas características físicas do disco/LUN antigo;

4. No momento da migração de dados, precisamos recriar esse *File System* no novo disco/LUN. Para isso é importante entender como o *File System* está montado e é utilizado atualmente, bem como saber: Qual o tipo de *File System* (`ext3`, `ext4`, `ReiserFS`, etc.)? Quais as opções de montagem? Está em apenas leitura ou leitura e escrita?

A partir do comando `mount -v`, verifique a atual parametrização do *File System* e se não existe nenhuma opção de montagem diferente das parametrizações default.

Em nosso caso, o *File System* `/srv/monitor` é do tipo `ext3` e possui opções default de montagem em modo Read-Write.

A partir da coleta de dados realizada com a execução dos procedimentos da **Listagem 3**, concluímos a fase de mapeamento do servidor `RJOSRVUXX03` e consolidamos as informações coletadas, conforme observado na **Figura 5**.

Planejamento

A etapa de planejamento é uma das mais importantes do processo de migração de dados. Todo tempo gasto nessa etapa se traduz em conforto operacional, segurança e tempo poupado durante a migração dos dados. O contrário também é verdadeiro, ou seja, toda migração mal planejada ou baseada em informações inexatas resultam em problemas, os quais podem causar perda de dados.

Outro ponto importante é respeitar todas as diretrizes de negócio, como a janela produtiva do ambiente em questão, *Downtime* máximo permitido para a aplicação (quando aplicável), possibilidade de *Rollback* em caso de problemas, incluir tempo de *Rollback* no planejamento da janela de migração, entre outras.

Listagem 3. Sequência de comandos para mapeamento do host `RJOSRVUXX03`.

```
##### Trecho 1 #####
sysadmin@rjosrvuux03:~$ df -h | tail -1
/dev/sdb1          79G  65G  14G  82% /srv/monitor

##### Trecho 2 #####
sysadmin@rjosrvuux03:~$ sudo lsscsi --size
[0:0:0:0] disk ATA  VBOX HARDDISK  1.0 /dev/sda  8.58GB
[1:0:0:0] disk ATA  VBOX HARDDISK  1.0 /dev/sdb  85.8GB

##### Trecho 3 #####
sysadmin@rjosrvuux03:~$ sudo fdisk -l /dev/sdb

Disk /dev/sdb: 85.9 GB, 85899345920 bytes
86 heads, 10 sectors/track, 195083 cylinders, total 167772160 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0xbc8fdd60

   Device Boot      Start         End      Blocks    Id      System
/dev/sdb1   *    2048    167772159    83885056    83          Linux

##### Trecho 4 #####
sysadmin@rjosrvuux03:~$ mount -v | grep /srv/monitor
/dev/sdb1 on /srv/monitor type ext3 (rw)
```

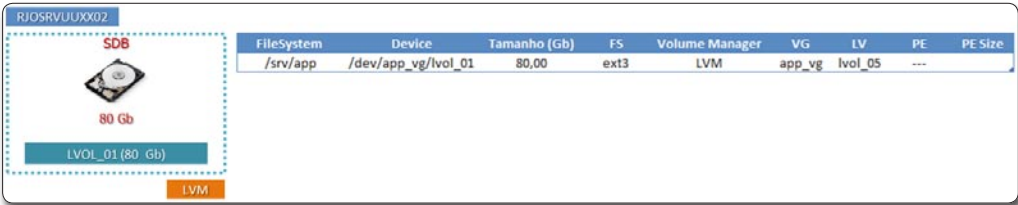


Figura 4. Dados a serem migrados do servidor `RJOSRVUXX02`

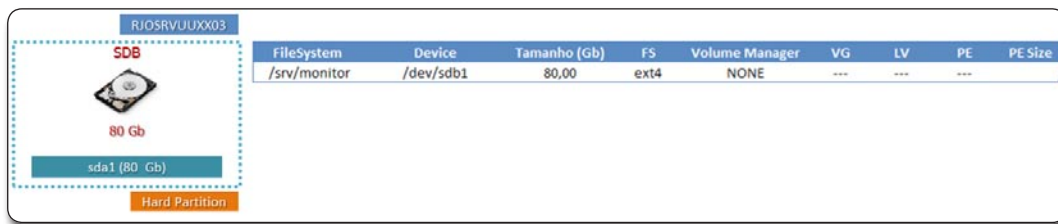


Diagrama de um servidor RJOSRVUXX03 com uma SDB de 80 Gb e uma tabela de dados.

FileSystem	Device	Tamanho (Gb)	FS	Volume Manager	VG	LV	PE	PE Size
/srv/monitor	/dev/sdb1	80,00	ext4	NONE	---	---	---	---

Figura 5. Dados a serem migrados do servidor RJOSRVUXX03

Trabalharemos com a ideia de que existem três pilares básicos a serem considerados no momento da escolha da estratégia de migração: tempo, complexidade operacional e segurança de execução. Esses pilares são grandezas diretamente proporcionais, ou seja, procedimentos com nível de complexidade baixo resultam em menor nível de segurança. Já procedimentos com maior nível de complexidade e que demandam mais tempo e atenção, costumam proporcionar maior conforto operacional e segurança durante a execução da atividade. No entanto, ainda cabe um estudo caso a caso, e essa opinião é influenciada pela cultura individual de cada administrador.

Matriz de Compatibilidade

Todo equipamento ou software utilizado no segmento *Enterprise* tem sua devida documentação nos respectivos fabricantes bem atualizada. É altamente recomendado que as matrizes de compatibilidade e conectividade sejam validadas. Normalmente, para se conectar um dado equipamento, o fabricante recomenda a atualização de *patches* do sistema operacional e a re-parametrização de drivers *scsi*, *fc* ou da controladora *Fiber Channel*. Em alguns casos, quando existe alguma solução de alta disponibilidade envolvida como *Clusters*, pode ser necessário habilitar parâmetros específicos nas portas do Storage ou até mesmo *flags* de controle nas LUNs mapeadas aos servidores.

Todas essas informações referentes a pré-requisitos de conectividade e suporte dos equipamentos envolvidos costumam estar descritas nas documentações, ou podem ser obtidas juntamente ao suporte técnico do fabricante do equipamento ou produto.

Se essas informações não forem validadas antes da migração, podem resultar em problemas e inviabilizar o sucesso da atividade. Não é incomum administradores sofrerem problemas de conectividade entre sistema operacional e *Storage* por não se adequarem às matrizes de compatibilidade mantidas pelos fabricantes dos equipamentos e softwares de infraestrutura.

Backup dos Dados

As ferramentas utilizadas nesse artigo garantem a consistência dos dados, no entanto existem fatores externos como falha humana, quedas de energia, *bugs* de sistema operacional, entre outros, que podem comprometer o sucesso da atividade. Portanto, tenha certeza de que existam backups íntegros para casos extremos de necessidade. Caso haja disponibilidade de tempo, também simule um restore parcial de alguns dados, a título de validação.

Em alguns casos, backups podem ser armazenados em dispositivos que podem sofrer desgaste e ação do tempo e do ambiente, como fitas (DDS, LTO, DLT). Deste modo, é sempre interessante verificar se a mídia está em bom estado.

Estratégias de Migração

Conforme abordado anteriormente, é possível realizar a migração dos dados utilizando diversas técnicas, como migração via *Storage*, via *Fabric* (SAN) e via aplicação. Porém, esse artigo pretende abordar técnicas de migrações que são executadas via ferramentas do próprio sistema operacional ou recursos do gerenciador de volumes LVM.

As técnicas que serão utilizadas para migrar nosso ambiente proposto são:

- **LVM – PV MOVE:** Procedimento que utiliza o próprio gerenciador de volumes para migrar os dados por meio do comando *PVMOVE*. Essa técnica consiste em migrar um PV do disco/LUN origem para o disco/LUN destino;
- **LVM – Split Mirror:** Técnica que consiste num *re-layout* do volume, transformando-o em um espelhamento (RAID 1), e depois do momento de sincronismo do espelhamento, efetua *SPLIT* dos dados. Dessa forma, o dado é migrado entre a LUN origem (um lado do *mirror*) e a LUN destino (outro lado do *mirror*);
- **Migração via ferramentas do sistema operacional:** Essa é a técnica utilizada quando os dados a serem migrados não estão sob gestão do LVM ou algum outro gerenciador. Normalmente os *File Systems* são montados por meio do caminho lógico */dev/sdNx*, ou seja, encontram-se em estado *Hard Partition*, conforme discutido anteriormente. Neste caso, realizaremos o procedimento utilizando o comando *RSYNC*, pois é uma ferramenta *freeware* disponível em praticamente todos os sabores de Unix e distribuições Linux, e oferece muitos recursos para manter a consistência dos dados, além de poder ser executada de forma incremental.

Para melhor compreensão do leitor, a **Tabela 1** procura ilustrar as estratégias de migração que abordaremos nesse artigo, juntamente com um breve sumário de suas respectivas vantagens e desvantagens do ponto de vista de migração de dados.

Baseado nas informações coletadas na fase de mapeamento dos dados a serem migrados referentes aos servidores RJOSRVUXX01, 02 e 03, as seguintes estratégias de migração serão utilizadas:

1. RJOSRVUXX01 – Estratégia LVM PV Move;
2. RJOSRVUXX02 – Estratégia LVM Split Mirror;
3. RJOSRVUXX03 – Estratégia Rsync.

Planilha De/Para

A Planilha De/Para é uma nomenclatura não oficial para um documento que relaciona os discos/LUNs, *File Systems* e Volumes

Técnica	Operação	Vantagens	Desvantagens
LVM PV Move	Block Based	• Procedimento mais simples quando comparado à técnica LVM Split Mirror; • Pode ser executado com a aplicação online, sem qualquer impacto aos serviços. • Velocidade da cópia não é influenciada pela quantidade de arquivos ou diretórios existentes nos volumes que serão migrados.	• Não permite reaproveitamento dos dados das LUNs antigas. Em caso de rollback, o dado precisa ser copiado novamente entre as LUNs destino e origem. • Todo o volume deve ser copiado. Caso o volume tenha o tamanho de 10 GB e tiver apenas 100 MB de dados, os 10 GB precisam ser tratados.
LVM Split / Mirror	Block Based	• Pode ser executado com a aplicação online, sem qualquer impacto aos serviços. • Os dados e estruturas LVM são preservadas nas LUNs antigas. • É possível ter acesso aos dados novos e antigos. • Em caso de rollback, basta efetuar UMount/MOUNT usando a estrutura LVM antiga. • Velocidade da cópia não é influenciada pela quantidade de arquivos ou diretórios existentes nos volumes que serão migrados.	• Procedimento pouco mais complexo quando comparado à técnica LVM pvmove. • Mais comandos são executados e maior nível de atenção é requerida. • Todo o volume deve ser copiado. Caso o volume tenha o tamanho de 10 GB e tiver apenas 100 MB de dados, os 10 GB precisam ser tratados.
Rsync	File Based	• Procedimento mais simples quando comparado às estratégias que envolvem o gerenciador de volumes. • Apenas os arquivos e diretórios existentes no File System serão migrados. • Caso o volume tenha o tamanho de 10 GB e tiver apenas 100 MB de dados, apenas os 100 MB de dados serão copiados.	• Velocidade da cópia é influenciada pela quantidade de arquivos e diretórios existentes nos volumes que serão migrados. • Dependendo da quantidade de arquivos e do nível de fragmentação do File System, a cópia pode demorar muito. • Para ter consistência dos dados migrados, a migração deve ser executada com a aplicação offline, trazendo maior downtime e impacto aos negócios.

Tabela 1. Estratégias de migração e suas vantagens e desvantagens

CURSOS ONLINE

A Revista .net Magazine oferece aos seus assinantes uma série de Cursos Online de alto padrão de qualidade.

.net
m a g a z i n e



CONHEÇA OS CURSOS MAIS RECENTES:

- Curso Padrões de Projeto com C#
- Curso básico de ASP .NET
- Curso de Introdução ao .NET Framework
- Curso Básico de C#
- C# 5 e suas novidades
- ASP.NET MVC – Sistema de Vestibular

Para mais informações :
www.devmedia.com.br/curso/netmagazine
 (21) 3382-5038



armazenados nos dispositivos que serão migrados com as informações dos novos dispositivos que receberão os dados oriundos do processo de migração. O objetivo desse documento é manter informações importantes sobre os dados que serão migrados e facilitar qualquer tipo de consulta de forma rápida.

A quantidade de informações nesse documento pode variar em função dos *Storages* ou gerenciadores de volumes envolvidos. Não existe um padrão exigido pelo mercado ou qualquer definição das informações que devem existir nesse documento, no entanto, conforme foi sinalizado anteriormente, quanto mais informação houver, melhor.

Quando temos acesso administrativo ao *Storage*, além do acesso aos sistemas operacionais dos servidores, é interessante do ponto de vista do mapeamento, incluir na Planilha De/Para informações de controle adicionais como nome e número da LUN, *serial number*, identificação do *Storage* utilizado, entre outros dados específicos da camada *Storage*. Em situações de verificação ou *Rollback*, o fácil acesso a esses dados se traduz em uma tomada de decisão mais rápida.

Em nosso artigo, não temos a necessidade do mapeamento dessas informações oriundas da camada *Storage*. Portanto, nosso documento De/Para será mais simples, apenas relacionando a LUN antiga à nova LUN, conforme pode ser observado na **Tabela 2**.

HOST	VG	Old Dev	New Dev	PE Size	PE	SIZE
SRVUXX01	db_vg	/dev/sdb1	---	4 M	10239	40 GB
		/dev/sdc1	---	4 M	10239	40 GB
		/dev/sdd1	---	4 M	10239	40 GB
SRVUXX02	app_vg	/dev/sdb1	---	4 M	20478	80 GB
SRVUXX03	NONE	/dev/sdb1	---	---	---	80 GB

Tabela 2. Exemplo de documento De/Para

Note que a coluna *New Dev* ainda encontra-se em branco e será preenchida na parte dois deste artigo, somente depois que os

novos discos forem mapeados nos três servidores e tenhamos o endereçamento lógico desses em nível de sistema operacional.

Conclusão

A primeira parte deste artigo procurou trazer aos leitores a complexidade e a abrangência de um projeto de migração de dados, abordando algumas estratégias e metodologias que podem ser empregadas durante o planejamento e a execução das migrações de dados corporativos ou até mesmo pessoais.

Como podemos perceber, existem muitas outras tarefas em uma migração de dados do que simplesmente movimentar os dados entre o dispositivo de origem e o dispositivo de destino. Como exemplo, temos tarefas relacionadas a processos internos da companhia, gestão de mudanças, atualização dos controles internos das equipes de infraestrutura, janelas de execução, procedimentos de *rollback*, vasto planejamento prévio, levantamento detalhado dos dados, planilha De/Para, entre outros.

Na segunda parte desta série, serão apresentados os procedimentos técnico-operacionais, incluindo todos os comandos necessários para realizar a migração dos *File Systems* e volumes LVM existentes nos servidores RJOSRVUXX01, 02 e 03. Adicionalmente, serão tratados aspectos como a validação dos dados pós-migração e o *Clean-Up* do ambiente envolvido.



Telmo Eduardo Nóbrega Reis

telmoeduardo@gmail.com

Analista de sistemas e consultor especialista em Storage, SAN e sistemas operacionais UNIX/Linux. Atualmente trabalha como arquiteto de soluções em projetos que envolvam tecnologias de alta disponibilidade, migração e replicação de dados, Capacity & Planning, análise e reestruturação de Infraestrutura LAN, WAN e SAN, Continuidade de Negócio, Disaster & Recovery e Cloud Computing. Atuou em projetos importantes em companhias como Banco Votorantim, VIVO, TIM, Nextel, Tivit, Diveo, HP, EMC, Symantec, Sun Microsystems, BM&F Bovespa, Embraer, Huawei, entre outras. Iniciou sua jornada no mercado de TI na companhia Sun Microsystems (agora Oracle) onde atuou como instrutor certificado por cinco anos lecionando sistema operacional Solaris, SunCluster e produtos relacionados a Storage.



Monitorando em tempo real os elementos da TI – Parte 2

Aprenda a analisar sua infraestrutura de TI com a ferramenta Splunk



ESTE ARTIGO FAZ PARTE DE UM CURSO

Não basta somente querer monitorar o ambiente e obter respostas em tempo real, é necessário que o departamento de TI das empresas tenha as pessoas certas e a ferramenta de coleta e análise de informação mais adequada para que o trabalho seja realizado. Muitas são as empresas que têm interesse em entrar para o mundo da análise de dados e ter suas decisões baseadas em números ou fatos, e é isso que o Splunk pode entregar, desde que a equipe de TI dê ao pessoal de negócios o apoio necessário para coletar os dados certos, disponibilizar um software para que analistas de infraestrutura possam analisar dados e extrair de lá, do Big Data, as respostas às perguntas do dia a dia. O mais interessante é que com o Splunk é possível conseguir respostas em segundos ao invés de dias, o que acontece com empresas que ainda tocam grandes Data Warehouses.

Analisar através de gráficos os pacotes que trafegam em redes, roteadores e switches, ou ainda fazer o monitoramento de servidores Windows, Unix e Linux, é parte do que o Splunk lhe dará possibilidade de fazer com poucos cliques. Deste modo, demonstraremos neste artigo algumas das funcionalidades do Splunk na prática.

Iniciando o Splunk

O primeiro passo para iniciarmos com o Splunk é acessar o site do produto e fazer o download da plataforma.

Resumo DevMan

Splunk: Monitorando em tempo real os elementos da TI – Parte 2:

Neste artigo colocaremos a mão na massa com um tutorial prático onde o leitor poderá explorar todo o potencial do Splunk para monitoramento da TI de uma empresa com base em dados de máquina. Extrair valor de logs e fazer com que a infraestrutura de uma empresa tenha menos problemas com downtime, por exemplo, são alguns dos pontos que fazem do Splunk uma plataforma única no mundo.

Em que situação o tema é útil:

Imagine que você precisa de uma ferramenta para consolidar logs de várias máquinas servidoras em uma só e, através de uma interface prática, analisar os problemas que vêm acontecendo relacionados com a compra de um produto em sua plataforma de e-commerce. Ou mesmo, imagine ter vários servidores Linux/Unix/Windows remotos, sendo necessário consolidar todos os logs dos ativos de TI para investigação de erros ou até ameaças de segurança. A possibilidade de investigar problemas e responder a eventos que impactam o funcionamento do seu ambiente fazem do Splunk uma plataforma de Big Data ideal para análise de toda a sua infraestrutura de TI.

Quando estiver na tela de download, perceba que você poderá selecionar o arquivo a ser baixado de acordo com a arquitetura da máquina, 32 ou 64 bits, e com o sistema operacional que esta roda. Mais a frente serão listados os sistemas operacionais suportados, mas, neste momento, selecione e baixe o arquivo adequado e faça a instalação do Splunk. Para sistemas operacionais Microsoft, a instalação segue o modelo NNF (*Next-Next-Finish*), enquanto que para SOs Unix-like, basta descompactar o arquivo “tar.gz” ou “.zip” no diretório /opt.

Feita a instalação, precisamos então iniciar o Splunk e trocar as credenciais que vêm configuradas por padrão (usuário *admin* e senha *changeme*). Para usuários de Windows, é mais fácil iniciar os serviços do Splunk (Splunkd e Splunkweb) a partir do painel de serviços. No Mac ou Linux, é necessário utilizar a linha de comando para iniciar, parar, reiniciar ou checar o status atual dos serviços do Splunk, o que é demonstrado nas **Listagens 1 e 2**.

Listagem 1. Iniciando o Splunk na linha de comando com MAC.

```
# Mac – iniciando o Splunk pela primeira vez
localhost:~wbianchi$ splunk start --accept-license

Splunk> See your world. Maybe wish you hadn't.

Checking prerequisites...
Checking http port [8000]: open
Checking mgmt port [8089]: open
Checking configuration...
Done.

Checking index directory...

Validated databases: _audit _blocksignature _internal _thefishbucket history main
summary

Done.

Success.

Checking conf files for typos...

All preliminary checks passed.
Starting splunk server daemon (splunkd)...
Done.
Starting splunkweb...
Done.

If you get stuck, we're here to help.
Look for answers here: http://docs.splunk.com/Documentation/Splunk
The Splunk web interface is at http://localhost:8000
```

Note que tanto no Linux quanto no Mac utilizamos a flag *--accept-license* para aceitar a licença e evitar que todo o texto da licença seja exibido no terminal. Ao final da saída do comando de start, perceba que já é informado o endereço para ser acessado via browser para então fazermos o primeiro login em nossa instalação do Splunk. Para verificar se o processo ou serviço está realmente ativo, utilize a opção *status*, como exibido na **Listagem 3**.

É interessante, em qualquer tipo de sistema operacional que for utilizado, configurar o diretório “Bin” (*splunk/bin*) da instalação do Splunk na variável PATH para que se tenha mais agilidade ao trabalhar com a plataforma na linha de comando. Vale ressaltar que muito do que faremos com a interface web (*SplunkWeb*) podemos fazer via linha de comando.

Após o primeiro login no Splunk, o usuário será convidado a trocar a senha atual do usuário *admin* para uma senha mais segura. Inicialmente a senha é “*changeme*”, como já informado, mas uma vez acessado você poderá colocar aquela que desejar. Após efetuar a troca da senha, seu browser será levado à tela de

welcome. Na parte inferior da página, o usuário já poderá clicar em “default app” para fazer as primeiras configurações, como colocar o seu nome em vez de utilizar o nome Administrador, definir o *timezone* no qual você está inserido e definir, dentre os aplicativos padrão Splunk, qual será aquele que será aberto após fazer login com o seu usuário.

Após fazer as alterações, clique em logout e efetue login novamente para verificar se o nome que você configurou já foi adotado pelo Splunk. Este aparecerá no menu superior, na parte direita da tela.

Listagem 2. Iniciando o Splunk na linha de comando com Linux-Unix.

```
# Linux – iniciando o Splunk pela primeira vez
root@localhost:~# /opt/splunk/bin/splunk start --accept-license
splunkd 3997 was not running.
Stopping splunk helpers...
Done.
Stopped helpers.
Removing stale pid file... done.

Splunk> Be an IT superhero. Go home early.

Checking prerequisites...
Checking http port [8000]: open
Checking mgmt port [8090]: open
Checking configuration...
Done.

Checking index directory...

Validated databases: _audit _blocksignature _internal _thefishbucket history main
summary

Done.

Success.

Checking conf files for typos...

All preliminary checks passed.
Starting splunk server daemon (splunkd)...
Done.
Starting splunkweb...
Done.

If you get stuck, we're here to help.
Look for answers here: http://docs.splunk.com/Documentation/Splunk

The Splunk web interface is at http://localhost:8000
```

Listagem 3. Exibindo o status atual da instância do Splunk.

```
localhost:~ wbianchi$ splunk status
splunkd is running (PID: 1132).
splunk helpers are running (PIDs: 1168 1169 1177 1178 1179 1180 1361).
splunkweb is running (PID: 1148)
```

Colocando a mão na massa

Como já abordado neste artigo, o Splunk é uma plataforma universal de dados de máquina que possibilita a coleta e a indexação dos dados, estejam eles onde estiverem. É possível coletar desde dados locais até dados localizados remotamente, como em uma mesma LAN ou WAN. Se os dados não estive-

rem na máquina local onde está instalado o Splunk, podemos instalar o *Universal Forwarder* para coletar os dados e enviá-los automaticamente para uma instalação ou instância de Splunk central. Isso ajuda a centralizar o monitoramento a partir da análise dos dados de todos os softwares que suportam os serviços oferecidos pela TI de uma empresa, independentemente de onde estiverem tais softwares.

Em um cenário mais simplista, ou seja, “começando pelo começo”, nesta parte do artigo teremos uma prática pela qual buscaremos explicar a indexação de dados com o Splunk. Vamos criar um índice exclusivo onde serão indexados os dados, buscar uma fonte de dados, investigar e buscar por ocorrências nos dados, criar painéis com relatórios e gráficos e finalizar com a criação de alertas. Isso nos dará um overview geral das capacidades técnicas para se trabalhar com o Splunk, além de passarmos pelas principais features que o produto oferece. Para criação deste cenário, o leitor precisará fazer o download dos arquivos de logs de exemplo que estão comprimidos em um zip chamado *Sampledata.zip*, disponível na documentação online do Splunk (veja a seção **Links**). Após baixar o arquivo, posicione-o em um diretório de fácil acesso para trabalharmos com ele dentro do Splunk.

Conhecendo e criando um novo índice

Como citamos anteriormente, e precisamos somente recordar neste tópico, o Splunk armazena os dados de máquina coletados em buckets que fazem parte de índices. Estes índices são mantidos pelo próprio Splunk em uma estrutura denominada MapReduce, que paraleliza as consultas aos dados fazendo com que a performance seja muito superior a muitos sistemas de armazenamento de informação existentes.

Por padrão o Splunk apresentará alguns índices para armazenamento de metadados criados automaticamente no processo de instalação. Os índices padrões do Splunk para armazenamento de metadados e dados são os seguintes:

- **summary:** índice interno utilizado para armazenar consultas que manipulam grandes volumes de dados. O resultado é armazenado nesse índice para auxiliar na recuperação dos dados quando a mesma consulta for executada pela segunda vez. Se comprarmos com um banco de dados, seria este um sistema de cache de resultados vinculado às consultas, ora denominado Summary Indexing. Os arquivos do índice de sumário estão localizados em `$SPLUNK_HOME/var/lib/splunk/summary`;
- **_internal:** de longe, um dos índices mais interessantes do Splunk, pois fornece várias possibilidades, como verificar as métricas de processamento, a atual fila de consultas, violações de licenças, informações de DEBUG, etc. Os arquivos do índice `_internal` estão localizados em `$SPLUNK_HOME/var/lib/splunk/_internaldb`;
- **_audit:** este índice contém informações sobre auditoria de processos, histórico de alterações no sistema de arquivos e um histórico de consultas do Splunk. Os arquivos deste índice estão localizados em `$SPLUNK_HOME/var/lib/splunk/audit`;
- **_thefishbucket:** índice que contém informações sobre os arquivos que vêm sendo acessados pelo Splunk para escrita

de dados e leitura dos mesmos. Este índice está localizado em `$SPLUNK_HOME/var/lib/splunk/fishbucket`;

- **main:** índice padrão para onde vão todos os dados quando o administrador ou usuário Splunk não especifica um índice exclusivo para uma nova fonte de dados (especificar um índice exclusivo para cada nova fonte de dados adicionada ao Splunk é parte do manual de boas práticas). É também chamado de `defaultdb` (no sistema de arquivos) e está localizado em `$SPLUNK_HOME/var/lib/splunk/defaultdb/db`.

Os índices internos do Splunk que têm nomes que iniciam com *underline* são índices que armazenam informações de processamento do próprio Splunk, tais como auditoria de processos, histórico de consultas, histórico de alertas e violações de licenças. Já o índice *main*, que também é criado no momento da instalação, é o índice padrão do Splunk para dados de usuário.

É interessante criar um novo índice sempre que um tipo de informação diferente for adicionado ao Splunk. Os índices são como apontadores para os dados que estão armazenados em estruturas denominadas *buckets*, e estes, por sua vez, estão organizados sobre o disco da máquina onde roda o Splunk.

Para criarmos um novo índice, de maneira simplificada, podemos seguir por duas vias: pelo *Splunk Manager* ou pela linha de comando (ver **Listagem 4**).

Listagem 4. Criando um novo índice na linha de comando.

```
root@splunk01:~# splunk add index splunkbr # commando simples
Index "splunkbr" added.
```

Veja os passos para criação do índice via Splunk Manager:

1. Faça login no Splunk com o usuário `admin` e a senha que configurou;
2. Acesse o menu *Manager* no canto superior direito;
3. Acesse o link *Indexes*, no agrupamento de menus *Data*;
4. Na tela que são listados todos os índices que sua instância Splunk possui, clique em *New*;
5. No campo *Index name*, informe o nome `splunkbr`;
6. Clique em *Save*.

Após clicar em *Save*, o browser será retornado à lista de índices que a instância Splunk possui. Nesse momento podemos verificar que o nosso índice foi realmente criado, uma vez que é parte da lista exibida na tela atual. Outra forma de listar o índice criado é via linha de comando, o que pode ser verificado na **Listagem 5**.

Listagem 5. Listando o índice que acabamos de criar.

```
root@splunk01:~# $SPLUNK_HOME/bin/splunk list index | grep splunkbr
splunkbr
/opt/splunk/var/lib/splunk/splunkbr/db
/opt/splunk/var/lib/splunk/splunkbr/db
/opt/splunk/var/lib/splunk/splunkbr/db
```

No tópico anterior, foi solicitado ao leitor que fizesse o download de um arquivo contendo vários logs. Nesse arquivo existem logs de três servidores Apache e logs de uma instância MySQL. O processo de adição de dados ao nosso Splunk então se inicia quando, a partir da Home do Splunk, clicamos em *Add Data* e selecionamos a opção *Files & Directories*. A opção *Skip Preview* poderá ser selecionada, seguindo com *Upload and Index a file*, já que temos arquivos estáticos. É nessa tela que vamos selecionar o índice que será utilizado para armazenar tais dados. Os passos para adicionar dados ao Splunk estão descritos a seguir:

1. Na Home do Splunk, clique em *Add Data*;
2. Clique na opção *Files & Directories*;
3. Clique em *Next* na opção *Consume any files on this Splunk Server*;
4. Marque *Skip Preview* e *Continue*;
5. Na tela seguinte, selecione a opção *Upload and index a file*;
6. Clique no botão *Browse* para selecionar o arquivo “*Sample-data.zip*”;
7. Clique em *More settings*;
8. No campo *Set host*, selecione o valor “*regex on path*”;
9. No campo *Regular expression*, informe a seguinte expressão regular, de acordo com o sistema operacional que está trabalhando:

Linux: `Sampledata.zip:./([^\s]+)/`
Windows: `Sampledata.zip:\\.\\([^\s]+)/`

10. Em *Set the destination index*, selecione *splunkbr*, que é o índice que criamos.
11. Clique em *Save*.

A tela seguinte, logo após clicar em *Save*, já lhe dará a opção de iniciar as buscas por dados que estão, nesse momento, sendo indexados pelo Splunk. Além disso, a *Search App*, que é o aplicativo padrão do Splunk, nos dará um sumário de todos os eventos e padrões já indexados e encontrados, respectivamente, pelo Splunk. Clique no link *Start Searching* e vamos em frente.

Search App e a SPL (Search Processing Language)

Por meio da *Search App*, o administrador do Splunk já conseguirá ter um sumário do que está sendo indexado no Splunk. À primeira vista já é possível avistar um campo de busca que é reconhecidamente um Google para o lado interno da TI, no qual você pergunta por determinados termos e o Splunk lhe trará todos os eventos que tenham aquela dada string.

Esta App conta com três áreas de sumário muito importantes, que são *Source*, que se refere ao arquivo que foi ou está sendo indexado (*path* + arquivo), o *sourcetype*, que é o tipo de informação que o Splunk está indexando e foi identificada pelo próprio Splunk, e *hosts*, que é nome do host onde reside o arquivo que está sendo indexado pelo Splunk. Estes três *fields* estarão presentes em qualquer fonte de dados que for indexada no Splunk, sendo que eles também podem ser utilizados para efetuar buscas.

Um exemplo interessante para ser explorado nesse momento é justamente o de clicar sobre o link *apache1.splunk.com*, localizado

em hosts. Após clicar, o Splunk iniciará uma busca pelos eventos que tenham sido originados pelo host *apache1.splunk.com*, e exatamente 9.199 linhas devem ser retornadas.

Percebam que, agora, no campo de busca, temos uma consulta por host igual ao nome do host que clicamos no sumário *hosts*. Perceba também que do lado esquerdo da tela surgiram os *fields*, estruturas que armazenam as informações que foram identificadas no processo de indexação e também de consulta (a cada consulta pode ser que um novo *field* seja descoberto!).

Os *fields* podem ser utilizados para buscar dados. Por exemplo, digamos que, nesse contexto, em que estamos trabalhando com dados de máquina que se referem aos logs de acesso do Apache, o administrador queira saber se existem erros de status 503 acontecendo durante a operação. Se percorrermos os *fields* atualmente descobertos pelo Splunk, será fácil achá-lo. Se ele já foi descoberto pelo Splunk, podemos simplesmente escrever uma consulta conforme a **Listagem 6**.

Listagem 6. Consulta por erro de status 503 em um servidor Apache.

```
host="apache1.splunk.com" status=503  
15 matching events
```

Vale salientar que existe um AND implícito entre a primeira e a segunda cláusula da consulta da **Listagem 6**, ou seja, queremos recuperar o host igual a *apache1.splunk.com* onde também haja *status* igual a 503. Nesse ponto, podemos ainda explorar um pouco mais a Search App para criar uma consulta que nos dê uma posição da quantidade de erros de status 503 (*Service Unavailable*) que vêm acontecendo em todos os hosts nos quais o Splunk vem coletando os logs do Apache e os indexando. Para isso, precisamos modificar a consulta da **Listagem 6** e colocá-la conforme a **Listagem 7**, onde utilizamos o comando *stats* e apontamos que queremos buscar dados somente no índice que criamos, o *splunkbr* (lembra?).

Listagem 7. Consulta por erro de status 503 nos servidores Apache.

```
index=splunkbr status=503 | stats count by host  
15 matching events
```

Após rodar a consulta da **Listagem 7**, podemos perceber que estamos chamando o índice criado anteriormente com a presença do sinal pipe “|”, que funciona como no Linux/Unix, direcionando o resultado da primeira consulta para a segunda. Nos resultados da consulta, vemos que somente o Apache do servidor *apache1.splunk.com* tem problemas com *status* 503 (*Service Unavailable*), o que nos coloca em uma posição mais proativa em relação aos problemas que possam estar acontecendo com este software, naquele servidor. Neste ponto vale lembrar o que no primeiro artigo discutíamos sobre os profissionais que por uma ou várias horas estariam engajados em achar os problemas decorrentes de uma parada nos serviços de tecnologia nos arquivos de logs, ou melhor, nos arquivos de logs dos vários softwares espalhados por vários servidores.

Com o Splunk é possível centralizar toda a informação de várias máquinas e obter as respostas para perguntas como: “desde quando nosso serviço tem falhado?” ou “quantas vezes nosso serviço falhou hoje?”.

O *Time Range Picker*, localizado ao final do campo de busca, permite ao usuário da *Search App* ajustar o intervalo de tempo sobre o qual se deseja fazer as pesquisas e investigações. É possível buscar dados em vários intervalos pré-definidos, sendo que o usuário ainda poderá personalizar o intervalo no qual busca por informações, clicando em *custom* e selecionando dados de um período específico. Para listar os períodos suportados pela ferramenta, basta clicar no botão do *Time Range Picker*. Na **Listagem 8** realizamos uma busca agrupando os status pela quantidade que cada um deles ocorreu ao longo do arquivo de log de acesso do servidor Web Apache. Perceba que, como temos três servidores Apache sendo monitorados, os erros de todos eles são agrupados em um só resultado. A **Figura 1** apresenta o resultado da consulta da **Listagem 8**.

Na consulta da **Listagem 8**, como pode ser observado, utilizamos o comando *rename* para criar um apelido para as colunas do resultado.

Um livro com a referência completa da linguagem de consulta utilizada no Splunk, de nome SPL ou *Search Processing Language*, pode ser baixado gratuitamente. Veja o endereço para isso na seção **Links**.

Outro exemplo interessante de comando que pode ser utilizado em consultas a dados no Splunk, ainda quando buscamos agrupar dados, é o *top*. Este comando agrupa as ocorrências de um determinado dado (em nosso exemplo utilizaremos o campo status) de acordo com o contexto atual e nos traz uma coluna adicional que informa o percentual de ocorrência de cada status. A coluna *percent* é adicionada automaticamente ao resultado. Veja na **Listagem 9** um exemplo da utilização do comando *top*, e na **Figura 2** o seu resultado.

Todas as consultas realizadas neste tópico, exibidas nas **Listagens 6 a 9**, ainda podem ser executadas na linha de comando, chamando o programa *client splunk* e executando a consulta, como exibido na **Listagem 10**. As cláusulas *search* e *rtsearch* podem ser utilizadas para se consultar utilizando a CLI e representam respectivamente uma busca simples e uma busca por dados em tempo real.

Criando Dashboards

Este é um dos pontos máximos da entrega de valor do Splunk. Após entrar com as consultas, aquelas que praticamos nas listagens anteriores, o administrador terá a oportunidade de gerar relatórios baseados em gráficos para tornar o processo de análise dos dados ainda mais inteligível.

Tudo começa na *Search App*, onde, após escrever uma consulta e visualizar os resultados, teremos a possibilidade de visualizar dois botões, ambos localizados logo abaixo do campo de busca, sendo um deles denominado *Create*. Ao clicar neste botão serão verificadas algumas opções de criação de objetos Splunk, como *Dashboard Panel*, *Alert*, *Report*, entre outros.

Nesse momento, já passada a busca de dados, vamos nos ater aos relatórios que criaremos em forma de gráficos. O processo aqui é simples. Uma vez que a consulta é escrita e os resultados já estejam sendo exibidos na área de resultados, basta clicar no botão *Create* e, em seguida, em *Dashboard Panel*. Assim que um novo *Dashboard Panel* é criado, ele pode contar com vários gráficos, os quais o administrador do Splunk poderá separar por áreas da empresa. Estes gráficos também são conhecidos por painéis.

Listagem 8. Consulta por ocorrência de status nos logs dos servidores Apache monitorados.

```
index=splunkbr
| stats count by status
| rename status as ReqStatus
| rename count as ReqQtd
| sort -ReqQtd
```

Listagem 9. Consulta com o comando *top*.

```
index=splunkbr
| top status
| sort -status
```

Listagem 10. Consultando dados utilizando a linha de comando.

```
$ splunk search 'index=_internal | stats count'
#
Preview of: index=_internal | stats count
count
-----
562490
```

	ReqStatus ↕	ReqQtd ↕
1	200	62337
2	302	1831
3	404	301
4	304	203
5	301	105
6	503	15

Figura 1. Resultado da consulta da **Listagem 7**

	status ↕	count ↕	percent ↕
1	200	62337	96.210952
2	302	1831	2.825966
3	404	301	0.464564
4	304	203	0.313310
5	301	105	0.162057
6	503	15	0.023151

Figura 2. Resultado da consulta da **Listagem 8**

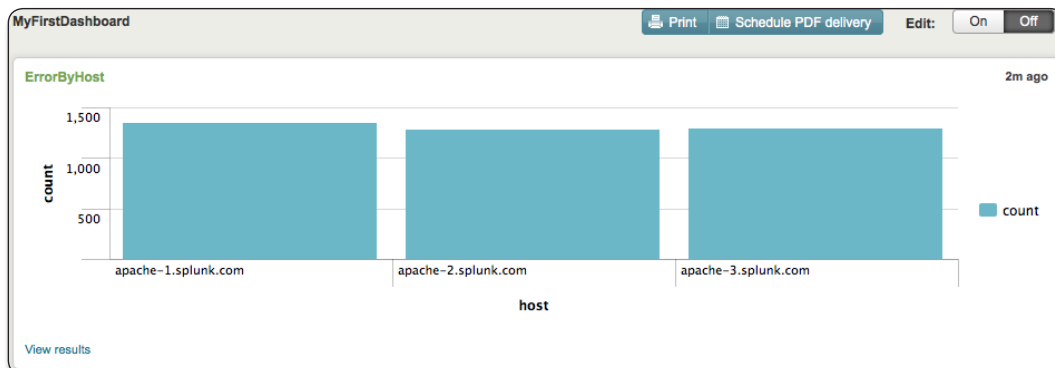


Figura 3. Painel de gráficos e o gráfico que criamos com a consulta da Listagem 8

Imaginemos então que precisamos exibir um gráfico em cima dos resultados da consulta da **Listagem 8**, aquela que exibe a quantidade de erros de status 503 dos servidores Apache que estamos monitorando. Como estamos utilizando os dados de exemplo que baixamos da documentação online do Splunk, estamos consultando os mesmos dados. Voltando ao Splunk, execute os passos a seguir para geração do primeiro *Dashboard Panel* que conterá um gráfico baseado na consulta da **Listagem 8**:

1. Entre com a consulta da **Listagem 8** no campo de busca de dados, na Search App;
2. Após executar a consulta, clique no botão *Create*, localizado no canto direito da tela, logo abaixo do final do campo de busca;
3. Nesta primeira tela, informe o nome da consulta “Error-ByHost” no campo *Name*. Feito isso, clique em *Next*;
4. Selecione a opção *New dashboard Name* e informe o nome “MyFirstDashboard”. Novamente, clique em *Next*;
5. Na terceira tela do Wizard, altere o campo *Visualization* para *Column* e clique em *Next* mais uma vez;
6. Na quarta e última tela, teremos a oportunidade de clicar sobre o link com o nome que demos ao nosso painel de gráficos no passo 4, *MyFirstDashboard*, e então visualizá-lo. Veja na **Figura 3** como ficou este painel.

CURSOS ONLINE

A Revista Java Magazine oferece aos seus assinantes uma série de Cursos Online de alto padrão de qualidade.



CONHEÇA ALGUNS DOS CURSOS:

- Curso de noSQL (Redis) com Java
- Curso Básico de JDBC
- Java Básico: Aplicações Desktop
- JSF com Primefaces
- Conhecendo o Apache Struts

Para mais informações :

www.devmedia.com.br/curso/javamagazine

(21) 3382-5038



Este processo de criação de gráficos que reportam respostas a perguntas internas ou externas de uma empresa com base nos dados de máquina pode ser realizado da forma como fizemos ou ainda salvando as consultas e adicionando-as aos painéis. Mas para que este artigo não fique tão longo e/ou cansativo, tal processo será mostrado em um próximo artigo.

Criando alertas

Para finalizar a apresentação do produto, precisamos abordar os alertas, que são, na visão de muitos dos clientes que o Splunk tem no mundo, um tópico de bastante importância, uma vez que se trata da facilidade que o Splunk apresenta de alertar usuários a partir do envio de e-mails, execução de scripts ou, ainda, na recente versão 5.0 do Splunk, o envio de relatórios personalizados no formato PDF. Este recurso é muito valioso, pois, numa situação em que seu supervisor precisa de um relatório de estoque ou da quantidade de transações realizadas em um sistema todas as manhãs e tardes, por exemplo, basta escrever uma consulta para recuperar os dados que ele precisa visualizar e, em seguida, criar um alerta sobre tal consulta, programando o envio deste alerta no formato PDF diretamente para o e-mail do seu supervisor.

Da mesma maneira como criamos um painel de gráficos ou *Dashboard Panel*, a partir do mesmo menu temos a oportunidade de criar os alertas, que seguem o mesmo padrão de criação. Após escrevermos a consulta da **Listagem 6**, a qual nos traz as ocorrências do erro de status 503 no servidor *apache1.splunk.com*, criaremos um gráfico para nos alertar sempre que um erro dessa natureza for retornado como resultado da consulta.

Ao escrever a consulta no campo de busca, atente-se às sugestões que são disponibilizadas pelo Splunk logo abaixo deste campo. Após executar a consulta e visualizar os resultados, clique no botão *Create* e selecione a opção *Alert*. Feito isto, um *Wizard* será aberto e iniciaremos a implementação de um alerta, cujo processo é apresentado a seguir:

1. Na primeira tela do *Wizard*, dê o nome "MyFirstAlert" ao seu alerta e clique em *Next*;
2. Na segunda tela, teremos a oportunidade de configurar o envio de e-mail (marcando a opção *Send email*) e rodar um script, que deverá estar armazenado em `$SPLUNK_HOME/bin/scripts/` (marcando a opção *Run a script*). Feito isto, preencha o campo com um script teste chamado "echo.sh" e clique em *Next*;
3. Na terceira tela, clique em *Finish*.

Terminados esses passos, o seu primeiro alerta sobre a análise dos dados retornados pela consulta da **Listagem 6** estará criado. Um ponto interessante é que, caso esse erro aconteça muito em

produção, você ainda poderá utilizar a opção *throttling* para adormecer os alertas por x segundos, minutos ou, ainda, horas, para que o problema possa ser resolvido e a caixa de e-mails do pessoal não vire mais um problema. Ao disparar e-mails como alertas podemos ainda enviar os resultados de uma consulta anexados nos formatos ".pdf" ou ".csv", ou ainda colocar a consulta e o resultado no corpo do e-mail (*inline*).

Conclusão

Neste artigo sobre o Splunk, foi feito um overview geral da solução Splunk para o monitoramento de TI de uma organização, independentemente do seu tamanho: pequena média ou grande. O Splunk é uma solução que oferece um framework com vários tipos de soluções e que ainda suporta o desenvolvimento de novas Apps sem necessidade de ferramenta terceira adicional. Ainda, oferece uma API chamada REST, que permite que as linguagens PHP, Python, Java e JavaScript sejam utilizadas para a criação de extensões do Splunk – seja para um tratamento adicional dos dados, seja para implementação de novos templates para deixar o Splunk com a cara da sua empresa.



Wagner Bianchi

wbianchi@splunk.com

Com MBA em Bancos de Dados Oracle pela Uniara (SP), Pós-graduação em Administração de Empresas e Gestão de Negócios pela Fundação Getúlio Vargas (SP/MG) e Nível Superior em Gerenciamento de Bancos de Dados pela Faculdade Infórium de Tecnologia (MG), Wagner vem atuando como Consultor de TI com foco em servidores de bancos de dados há mais de 10 anos, tendo passado por empresas como IBM, Oracle e Percona. Atualmente é Sr Sales Engineer na empresa norte-americana Splunk, Inc., onde mantém foco em Big Data e monitoramento de serviços, alinhando a TI aos Negócios.



Exploring Splunk

<http://www.splunk.com/goto/book>

Documentação do Splunk

<http://docs.splunk.com/Documentation>

Download do arquivo de logs (Sampledata.zip)

<http://tinyurl.com/bgmkod3>

Site oficial Splunk em português

<http://pt.splunk.com>

Splunk Processing Language Book

<http://www.splunk.com/goto/book>



Somos tão apaixonados por tecnologia que o nome da empresa diz tudo.

Porta 80 é o melhor que a Internet
pode oferecer para sua empresa.

Já completamos 8 anos e
estamos a caminho dos 80, junto
com nossos clientes.

Adoramos tecnologia.
Somos uma equipe composta
de gente que entende e
gosta do que faz,
assim como você.



Estrutura

100% NACIONAL.
Servidores de primeira
linha, links de alta
capacidade.



Suporte diferenciado

Treinamos nossa equipe
para fazer mais e melhor.
Muito além do esperado.



Serviços

Oferecemos a tecnologia
mais moderna, serviços
diferenciados e
atençados com as suas
necessidades.



1-to-1

Conhecemos nossos
clientes. Atendemos
cada necessidade de
forma única.
Conheça!



Porta 80
WEB HOSTING

Hospedagem | Cloud Computing | Dedicados | VoIP | Ecommerce |
Aplicações | Streaming | Email corporativo

porta80.com.br | comercial@porta80.com.br | twitter.com/porta80

SP 4063-8616 | RJ 4063-5092 | MG 4063-8120 | DF 4063-7486