

CCOA: Arquitetura Aberta de Computação em Nuvem

Liang-Jie Zhang e Qun Zhou IBM

TJ Watson Research Center, Nova York, EUA E-mail:

{zhanglj, qzhou}@us.ibm.com

Abstrato

A computação em nuvem está evoluindo como uma plataforma de computação chave para compartilhar recursos que incluem infraestruturas, software, aplicativos e processos de negócios. A virtualização é uma tecnologia central para permitir o compartilhamento de recursos em nuvem. No entanto, a maioria das plataformas de computação em nuvem existentes não adotaram formalmente a arquitetura orientada a serviços (SOA) que os tornaria mais flexíveis, extensíveis e reutilizáveis. Ao unir o poder de SOA e virtualização no contexto do ecossistema de computação em nuvem, este artigo apresenta sete princípios de arquitetura e deriva dez módulos de arquitetura interconectados para formar uma arquitetura aberta de computação em nuvem (CCOA) reutilizável e personalizável. Dois estudos de caso sobre Infraestrutura e Business Cloud são usados para fornecer valor comercial e prático de infraestrutura e serviços de provisionamento de processos de negócios pela Internet. Apresentamos também alguns potenciais serviços de valor agregado do CCOA proposto para orientar o planejamento estratégico e outras práticas de consultoria de Cloud Computing.

1. Introdução

Como uma plataforma de entrega de serviços chave no campo da computação de serviços [12], a computação em nuvem fornece ambientes para permitir o compartilhamento de recursos em termos de infraestruturas escaláveis, plataformas de desenvolvimento de aplicativos e middleware e aplicativos de negócios de valor agregado. Os modelos de operação podem incluir modelos de utilidade pay as go, serviços de infraestrutura gratuitos com serviços de plataforma de valor agregado, serviços de infraestrutura baseados em taxas com serviços de aplicativos de valor agregado ou serviços gratuitos para fornecedores, mas compartilhamento de receitas geradas por consumidores.

Conforme resumido em [1], normalmente existem quatro tipos de recursos que podem ser provisionados e consumidos pela Internet. Eles podem ser compartilhados entre os usuários, aproveitando a economia de escala. O provisionamento é uma maneira de compartilhar recursos com solicitantes pela rede. Um dos principais objetivos da Cloud Computing é alavancar a Internet ou Intranet para provisionar recursos aos usuários.

O primeiro tipo de recursos são recursos de infraestrutura, que incluem poder de computação, armazenamento e provisionamento. Por exemplo, o Amazon EC2 fornece interface de serviço da web para solicitar e configurar facilmente a capacidade online [2]. O serviço Xdrive Box fornece armazenamento online para

usuários [3]. Microsoft SkyDrive oferece armazenamento gratuito serviço, com um modelo integrado offline e online que mantém arquivos relacionados à privacidade em discos rígidos e permite que as pessoas acessem esses arquivos remotamente [4]. Na área de compartilhamento de poder de computação, a iniciativa de computação em grade tomou como seu foco principal o uso de tecnologias de clustering e computação paralela para compartilhar poder de computação com outros, com base no agendamento de tarefas quando os computadores estão ociosos.

O segundo tipo de recursos em Cloud Computing são recursos de software, incluindo middleware e recursos de desenvolvimento. O middleware consiste em sistemas operacionais centrados em nuvem, servidores de aplicativos, bancos de dados e outros. Os recursos de desenvolvimento compreendem plataformas de design, ferramentas de desenvolvimento, ferramentas de teste, ferramentas de implantação e projetos de referência baseados em código aberto.

O terceiro tipo de recursos em Cloud Computing são os recursos de aplicativos. As principais empresas da indústria da informação estão gradualmente movendo aplicativos e dados relacionados para a Internet. Os aplicativos de software são fornecidos por meio do modelo Software As A Service (SaaS) ou mashups de aplicativos de valor agregado. Por exemplo, o Google usou a plataforma Cloud Computing para oferecer aplicativos Web para comunicação e colaboração [19]. Os documentos do Google movem os aplicativos de produtividade para a Web, substituindo gradualmente os aplicativos de desktop pesados. Portanto, desenvolver uma Arquitetura Aberta de Computação em Nuvem reutilizável e personalizável para habilitar o ambiente de desenvolvimento de aplicativos é a chave para o sucesso do compartilhamento de aplicativos pela Internet.

O quarto tipo de recursos em Cloud Computing são os processos de negócios. Alguns aplicativos podem ser expostos como utilitários, ou seja, subprocessos ou tarefas fracamente acoplados dentro dos processos de negócios dos clientes. O compartilhamento de processos de negócios é a terceirização de aplicativos orientada aos negócios que oferece suporte à reutilização, composição

Este artigo está organizado da seguinte forma. A primeira parte da Seção 2 discute as duas principais tecnologias de habilitação, Virtualização e SOA, para articular o valor da criação de uma arquitetura aberta para Cloud Computing. O restante da Seção 2 apresenta uma Arquitetura Aberta de Computação em Nuvem (CCOA) e seus sete princípios para construir sistemas de Computação em Nuvem extensíveis e flexíveis e formulários. A Seção 3 descreve dois estudos de caso sobre Infrastructure Cloud e Business Cloud, para ilustrar o provisionamento de serviços de infraestrutura e o processo de negócios como serviço. Algumas ofertas de valor agregado baseadas no CCOA proposto também são descritas. A seção 5 discute

trabalho relacionado e direções futuras. As conclusões são tiradas no final deste artigo.

2. Arquitetura Aberta de Computação em Nuvem

Atualmente, não há definição ou especificação padrão para Cloud Computing. Pode levar algum tempo para definir as principais características da computação em nuvem com base nas práticas de campo. A computação em nuvem envolve um conjunto de tecnologias-chave para abordar o compartilhamento de recursos com base nos requisitos de negócios. Com base em nossas práticas nas áreas de provisionamento de serviços e design de soluções, acreditamos que as duas tecnologias de habilitação a seguir podem desempenhar papéis muito importantes nessa fase revolucionária: tecnologia de virtualização e arquitetura orientada a serviços (SOA).

A tecnologia de virtualização lida com como as imagens dos sistemas operacionais, middleware e aplicativos são procriadas e alocadas para as máquinas físicas corretas ou uma fatia de uma pilha de servidores. As imagens podem ser movidas e colocadas em ambiente de produção sob demanda. Por outro lado, a tecnologia de virtualização também pode ajudar a reutilizar licenças de sistemas operacionais, middleware ou aplicativos de software, uma vez que um assinante libera seu serviço da plataforma Cloud Computing.

A SOA é a evolução de um sistema ou arquitetura de software para endereçar componentização, extensibilidade de reutilização e flexibilidade. Para construir plataformas de computação em nuvem escaláveis, precisamos aproveitar SOA para construir componentes reutilizáveis, interfaces baseadas em padrões e arquiteturas de soluções extensíveis.

Criar uma plataforma chamada Cloud Computing é fácil desde que permita o compartilhamento de pelo menos um dos recursos. No entanto, construir uma arquitetura de computação em nuvem unificada, escalável e reutilizável para suportar o compartilhamento de todos os tipos de recursos ainda enfrenta desafios nas áreas de inovação tecnológica e melhores práticas do setor.

2.1 Modelo “OSI” para Computação em Nuvem

Identificamos os três objetivos a seguir para ajudar a enfrentar o desafio acima de definir uma boa arquitetura aberta para Cloud Computing.

O primeiro objetivo é articular uma forma reutilizável de criar uma plataforma de provisionamento escalável e configurável para Cloud Computing. Este jornal reúne as poder da Arquitetura Orientada a Serviços (SOA) e da virtualização para fornecer valor prático e de negócios para aplicativos de software emergentes, hardware e serviços de provisionamento de processos de negócios pela Internet no contexto da Computação em Nuvem.

O segundo objetivo é propor um conjunto de serviços comuns e compartilhados para a construção de plataformas de Cloud Computing, para fornecer serviços de negócios ou outras ofertas de nuvem para seus usuários consumidores corporativos em uma abordagem unificada.

O terceiro objetivo é maximizar o valor comercial potencial da computação em nuvem com base em uma infraestrutura de TI extensível e sistema de gerenciamento. Isso levará aos serviços de valor agregado da nuvem de negócios, por meio da monetização do poder combinado de SOA e Cloud Computing.

Como sabemos, os padrões OSI para Open System Interface [6] que encontrou alguns desafios para perceber seu valor no contexto de sistemas abertos genéricos. Neste artigo, tentamos limitar o escopo do sistema aberto a um domínio especializado e alavancar o pensamento orientado a serviços para ajudar a modular a arquitetura aberta para Cloud Computing. Na seção a seguir, apresentamos uma Arquitetura Aberta de Computação em Nuvem baseada em sete princípios.

2.2 Sete Princípios da Arquitetura de Computação em Nuvem

Nesta arquitetura aberta de computação em nuvem, propomos uma estrutura integrada de co-inovação e coprodução para fazer com que fornecedores de nuvem, parceiros de nuvem e clientes de nuvem trabalhem juntos com base em sete princípios, que são usados para definir dez principais módulos de arquitetura e seus relacionamentos mostrada na Figura 1. A Arquitetura Aberta de Computação em Nuvem apresentada abrange a habilitação do ecossistema de nuvem, infraestrutura de nuvem e seu gerenciamento, orientação a serviços, núcleo de nuvem em provisionamento e assinatura, ofertas de nuvem compostável, arquitetura e gerenciamento de informações de nuvem e análise de qualidade de nuvem. Esta é uma separação lógica e modularizada, que ajuda a isolar as preocupações dos detalhes de cada módulo durante o processo de projeto. Como as conexões entre os principais princípios arquitetônicos identificados para Cloud Computing são bastante complexas, as trocas de informações estão passando pela Arquitetura da Informação em Nuvem e pelo Gerenciamento do Ecossistema em Nuvem. No restante da seção, apresentaremos os detalhes de cada princípio ilustrado na Figura 1.

Princípio 1: Gerenciamento Integrado de Ecossistemas para Nuvem Uma arquitetura deve suportar o gerenciamento do

ecossistema de Computação em Nuvem. Esse ecossistema inclui todos os fornecedores de serviços e soluções, parceiros e usuários finais envolvidos para fornecer ou consumir recursos compartilhados no ambiente de computação em nuvem. Os fornecedores de nuvem expõem suas interfaces de interação de suas operações internas e capacidade de desenvolvimento de produtos para a nuvem. Os parceiros de nuvem fornecem componentes para fornecedores de nuvem ou atuam como agentes para fornecer serviços de valor agregado à nuvem clientes. Os clientes de nuvem são usuários dos serviços de nuvem que oferecem compartilhamento de recursos orientado a objetivos de negócios.

De uma perspectiva de design de arquitetura, o painel do fornecedor de nuvem fornece uma visão integrada da interação com as operações de front-end e back-end dos fornecedores. Por exemplo, no frontend, marketing e desenvolvimento de negócios atividades, portais de entrega de serviços e suporte ao cliente

pode ser ativado no painel do fornecedor da nuvem. No backend, atividades de design e desenvolvimento de soluções, ou

ambiente de hospedagem são usados para dar suporte às operações do frontend.

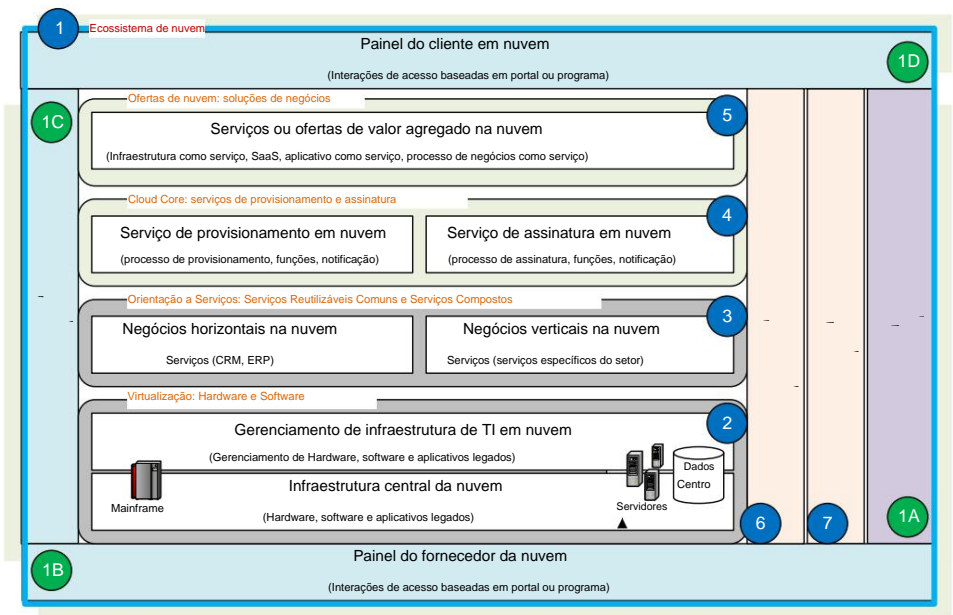


Figura 1. Diagrama de visão geral da arquitetura aberta de computação em nuvem

Como a maioria dos fornecedores de nuvem não trabalha sozinho mais, eles precisam colaborar com seus parceiros [7] na cadeia de valor do ambiente Cloud Computing. Nesse sentido, é necessário um painel de parceiros para que os parceiros participantes interajam com os fornecedores e clientes da nuvem. Por exemplo, se os parceiros de nuvem servirem como fornecedores de componentes para os fornecedores de nuvem, os blocos de construção arquitetônicos para interagir com os fornecedores e o gerente de política de colaboração são fundamentais para a integração da cadeia de valor.

Clientes ou usuários finais de Cloud Computing podem ser agrupados em duas classes: usuários corporativos e consumidores. O painel do cliente em nuvem fornece um ponto focal para todos os tipos de usuários interagirem com serviços ou ofertas de computação em nuvem. Esse ponto focal fornece uma estrutura unificada para que os usuários consumam serviços em nuvem por meio de vários canais, como portal da Web, canal de colaboração entre empresas ou canal de representação de clientes por telefone. Há oportunidades para explorar uma arquitetura convergente de software e serviços para usuários corporativos e consumidores com base em várias estratégias de preços, habilitação de segurança e outros recursos de software e serviços. Como os usuários corporativos ou usuários consumidores são atores coexistentes no ecossistema de serviços, um usuário corporativo pode ter vários usuários consumidores. No final, eles são apenas consumidores de recursos de Cloud Computing em diferentes níveis.

Juntando todos esses painéis, a camada de gerenciamento de ecossistema de computação em nuvem (1A) fornece um processo integrado de integração e utilitários comuns para

suportar a colaboração perfeita e trocas de mensagens entre fornecedores de nuvem, parceiros e clientes.

Por exemplo, o progresso a bordo abrange o registro de entidades comerciais e usuários. As entidades de negócios incluem fornecedores de nuvem, parceiros de nuvem e clientes de nuvem corporativa. As entidades de usuário são usuários finais dentro de uma determinada entidade comercial (por exemplo, um funcionário de uma empresa ou um membro de uma comunidade registrada como uma rede social), ou usuários consumidores no espaço aberto da

Princípio 2: Virtualização para infraestrutura em nuvem

Existem duas abordagens básicas para habilitar a virtualização no ambiente de computação em nuvem. A primeira abordagem é a virtualização de hardware que consiste em gerenciar equipamentos de hardware no modo plug-and-play. Hardware equipamentos podem ser adicionados ou removidos sem afetar as operações normais de outros equipamentos no sistema. Obviamente, o desempenho ou os espaços de armazenamento podem ser alterados dinamicamente devido a essas ações de adição e remoção.

A segunda abordagem é a virtualização de software, ou seja, usar o gerenciamento de imagem de software ou tecnologia de virtualização de código de software para permitir o compartilhamento de software. Especificamente, as imagens de software podem ser criadas com base no grau de reutilização de um conjunto de sistemas de software, incluindo sistema operacional, middleware e aplicativos. A outra tecnologia de virtualização de software é a montagem e execução de código dinâmico. Nesse caso, não há imagens de software. Os elementos de código serão copiados dinamicamente dos repositórios e colados nos lugares certos com base

na lógica de negócios. Por exemplo, em um aplicativo da Internet, alguns elementos de código JavaScript podem ser recuperados dinamicamente e inseridos em um pacote Ajax correto para criar novas funções ou recursos para um cliente Web. Da mesma forma, os programas do lado do servidor podem ser montados e executados dinamicamente com base na composição de elementos de código reutilizáveis e tecnologias de compilador just-in-time. Com o desenvolvimento da tecnologia multi-core e programação paralela, essa tecnologia dinâmica de montagem e execução de código terá grandes vantagens no futuro. Especialmente, ele eliminará os requisitos de grandes espaços de armazenamento para imagens de software para middleware e ferramentas de desenvolvimento. Mas em No ambiente atual, ambas as tecnologias de virtualização de software podem coexistir e oferecer suporte uma à outra com base em cenários de uso.

Em resumo, o módulo Cloud IT Infrastructure Management abrange gerenciamento de imagens de software, virtualização de hardware e empacotamento de aplicativos legados.

Os recursos de destino gerenciados pelo módulo Cloud IT Infrastructure Management é o Cloud Core Infrastructure, que compreende todo o hardware, software e aplicativos legados de suporte para operar um ambiente de computação em nuvem. Por exemplo, o hardware pode incluir mainframe, servidores distribuídos ou clusters e armazenamentos de dados. O software pode incluir pacotes de banco de dados e middleware relacionado. Os aplicativos herdados podem envolver aplicativos desenvolvidos internamente ou aplicativos ISV que fazem parte da infraestrutura.

Note-se que este princípio de virtualização na Cloud Computing Open Architecture é uma extensão da camada de sistema operacional no SOA Solution Stack (também conhecido como Arquitetura de Referência SOA) [8] no contexto de habilitação de Cloud Computing. Por exemplo, um gerenciador de rack pode ser aproveitado para lidar com sistemas de hardware com fio. A interface de recursos de TI de montagem dinâmica permite o provisionamento de recursos de infraestrutura e pode ser usada para hospedar ou implantar aplicativos de negócios que aproveitam recursos de infraestrutura de nuvem distribuídos.

Princípio 3: Orientação de Serviço para Comum Serviços reutilizáveis

Conforme apresentado anteriormente, além da característica de virtualização, a orientação a serviços é outra força motriz para permitir que a computação em nuvem realize ainda mais o valor comercial da reutilização de ativos, aplicativos compostos e serviços de mashup. Existem dois tipos principais de serviços reutilizáveis comuns: Cloud Horizontal e Vertical Business Services.

Os Serviços de Negócios Horizontais em Nuvem consistem em vários serviços de plataforma que ocultam as complexidades de middleware, banco de dados e ferramentas. Além de oferecer middleware ou ferramentas de desenvolvimento como serviços no ambiente de Cloud Computing, alguns utilitários comuns como on-boarding, provisionamento, monitoramento, ferramentas de faturamento ou

serviços intersetoriais como Customer Relationship Management (CRM) e Enterprise Resource Planning (ERP).

Os Cloud Vertical Business Services incluem todos os serviços utilitários específicos do domínio ou do setor. Exemplos são os serviços de envio e pagamento.

Ambos os serviços reutilizáveis comuns no CCOA podem ser reutilizados para habilitar os serviços de assinatura e provisionamento do núcleo da nuvem, bem como para criar ofertas de nuvem, como infraestrutura como serviço, SaaS, aplicativo como serviço, processo de negócios como serviço, que serão descritos mais detalhadamente no Princípio 5.

Princípio 4: Provisionamento Extensível e Assinatura para nuvem

O provisionamento de serviço extensível é o recurso exclusivo de um sistema de computação em nuvem. Sem extensibilidade, a parte de provisionamento da arquitetura Cloud Computing só pode suportar um determinado tipo de compartilhamento de recursos. Isso implica que a arquitetura de provisionamento de serviços para usuários de uso gratuito e usuários pagantes é a mesma. Ambos os tipos de usuários podem ser prestadores de serviços ou consumidores de tempos em tempos. Do ponto de vista dos consumidores interessados em como acessar facilmente os serviços com base em suas próprias lógicas e objetivos de negócios. Do ponto de vista dos provedores de serviços, três níveis de fornecimento de serviços descritos na Seção 1 serão as ofertas-alvo.

O Cloud Core mostrado na Figura 1 inclui um conjunto de serviços de provisionamento e assinatura. Isso aborda uma questão-chave de como lidar com o processo de provisionamento dos provedores de serviços e o processo de assinatura dos consumidores de serviços na arquitetura de Computação em Nuvem. Classificamos o Cloud Core em Cloud Provisioning and Subscription Service. Os principais elementos de arquitetura do Cloud Provisioning Service incluem processo de provisionamento, definições de função e estrutura de notificação. Nuvem O Serviço de Assinatura envolve o processo de assinatura, definições de função e estrutura de notificação. A estrutura e os atributos de definição de função, bem como a estrutura de notificação, podem ser compartilhados para serviços de provisionamento e assinatura.

Princípio 5: habilitação configurável para nuvem Ofertas

As ofertas de nuvem são os produtos ou serviços finais que são provisionados pela plataforma de computação em nuvem. Como todas as ofertas de nuvem devem atender a determinados objetivos de negócios, as ofertas de nuvem também são conhecidas como soluções de negócios em nuvem. Em alinhamento com a categorização de compartilhamento de recursos descrita na Seção 1, o CCOA define seus aspectos de oferta nos quatro níveis a seguir: infraestrutura como serviço, software como serviço (SaaS), aplicativo como serviço e processo de negócios como

Aproveitando a extensibilidade e a configurabilidade da arquitetura de referência SOA, o CCOA proposto busca a habilitação configurável de plataformas e serviços de Cloud Computing. O gerenciamento modularizado do ecossistema de nuvem, virtualização, orientação a serviços e núcleo de nuvem formaram uma base sólida para garantir que uma plataforma de computação seja configurável, compostável e gerenciável.

Na categoria de infraestrutura como serviço, o CCOA aproveita o núcleo da nuvem para oferecer suporte ao provisionamento e assinatura dos recursos de infraestrutura de TI virtualizados, explorando alguns serviços reutilizáveis comuns e recursos de gerenciamento de ecossistema de nuvem. Algumas ofertas de nuvem de exemplo são nuvem de armazenamento e nuvem de infraestrutura de teste.

A maioria das ofertas de nuvem são entregues ou acessadas por meio de navegadores da Web. Por exemplo, podemos usar o navegador da Web para fazer upload de fotos, arquivos de áudio, arquivos de vídeo e outros documentos para uma nuvem de armazenamento. Em todas as outras ofertas de nuvem, as interfaces da Web têm sido comprovadamente canais eficazes para permitir que clientes de nuvem interajam com parceiros e fornecedores de nuvem no ciclo de vida da entrega de serviços.

Na área de software como serviço, muitas histórias de sucesso demonstram serviços reutilizáveis comuns, como CRM como serviço (por exemplo, Salesforce.com), pagamento como serviço (por exemplo, o PayPal do eBay) e envio como um serviço, aproveitando o núcleo da nuvem para gerenciar o provisionamento e a assinatura. Enfatiza-se que muitos serviços de valor agregado podem ser construídos em cima dos recursos reutilizáveis comuns. serviços no CCOA antes de serem provisionados para clientes de nuvem como ofertas de nuvem. Portanto, aplicativos compostos ou tecnologias de composição de serviços em SOA podem desempenhar um papel importante na construção de serviços de nuvem de valor agregado no contexto do CCOA.

Na área de aplicativo como serviço, muitos aplicativos autônomos ou aplicativos baseados em rede podem ser provisionados como serviços em nuvem para compartilhamento na cadeia de valor de Computação em Nuvem. Por exemplo, as ferramentas de desenvolvimento baseadas na Web podem ser boas ofertas de nuvem para a comunidade de desenvolvimento de aplicativos, para projetar e desenvolver aplicativos para nuvem ou aplicativos independentes, aproveitando os recursos de desenvolvimento da Computação em Nuvem, como software de modelagem empresarial, software de modelagem de processos de negócios, modelagem de arquitetura software, software de desenvolvimento de aplicativos e repositórios de ativos relacionados. Especificamente, no espaço de design de soluções orientadas a serviços, oferecer o serviço SOMA-ME [13] como um design no ambiente de computação em nuvem é uma maneira natural de capacitar os membros da comunidade de massa de design e desenvolvimento de aplicativos. Não há limite absoluto entre software como serviço e aplicativo como serviço. No CCOA, um aplicativo é a integração de um conjunto de pacotes de software com base em determinadas lógicas e objetivos de negócios.

No ambiente de Cloud Computing, o processo de negócio como serviço é um novo modelo para compartilhar as melhores

práticas e processos de negócios entre clientes de nuvem e parceiros na cadeia de valor. Um exemplo de processo de negócio como oferta de serviço no ambiente de Cloud Computing é o teste de software, que é um processo de negócio muito importante no ciclo de vida do desenvolvimento de software. Como os sistemas ou aplicativos de software precisam ser implantados e executados em diferentes sistemas operacionais, ambientes de middleware ou diferentes versões ou configurações deles, é muito difícil para desenvolvedores individuais ou empresas configurarem uma “fábrica” de testes para abordar essas questões de teste em na área de serviços de aplicação.

A atualização ou manutenção do ambiente de teste envolve a migração de hardware, software e repositórios de conhecimento de teste. São as duas maiores despesas, além dos engenheiros de teste que manipulam o ambiente para realizar os testes.

Princípio 6: Representação Unificada de Informações e Exchange Framework

A representação de informações e troca de mensagens dos recursos de Cloud Computing são muito importantes para possibilitar os recursos colaborativos e eficazes da Cloud Computing. No CCOA, os recursos de computação em nuvem incluem todas as entidades de negócios (por exemplo, clientes, parceiros e fornecedores de nuvem) e os recursos de suporte, como módulos relacionados à virtualização, módulos relacionados à orientação a serviços, núcleo de nuvem e ofertas de nuvem. Por exemplo, no gerenciamento de ecossistema de nuvem, país, local e organização estão associados às entidades de negócios na Computação em Nuvem. Os role players são usados para definir suas funções na dinâmica do ecossistema Cloud Computing. Projeto, tarefa, documentos, transações, processos de negócios, links de referência, anotações e eventos são recursos potenciais para apoiar a colaboração entre várias entidades de negócios no ambiente de Computação em Nuvem.

No CCOA, o módulo de arquitetura de informações em nuvem permite a representação dessas entidades de nuvem (entidades de negócios e recursos de suporte) em uma estrutura unificada de descrição de entidade de computação em nuvem. Por exemplo, tecnologias como Resource Description Framework (RDF), Web Services Resource Framework (WSRF) e XML são candidatos para implementar essa estrutura unificada. As mensagens trocadas entre entidades de nuvem formam padrões de troca de mensagens. O formato da mensagem e os padrões de troca de mensagens podem ser reutilizados para dar suporte a vários cenários de negócios. Os protocolos de roteamento e troca de mensagens, bem como a capacidade de transformação de mensagens, formam uma base para a arquitetura de informações em nuvem.

Assim como o sangue nos corpos humanos, a arquitetura de informações em nuvem usa seu “sangue” de informações para formar “fluxo de sangue” para fazer com que todos os vários módulos se comuniquem entre si de maneira eficaz no CCOA.

Princípio 7: Qualidade e Governança da Nuvem

O último e mais importante módulo do CCOA é o Cloud Quality and Governance mostrado na Figura 1. Este módulo é responsável pela identificação e definição de indicadores de qualidade para o ambiente Cloud Computing e um conjunto de orientações normativas para reger o projeto, implantação, operação, e gerenciamento das ofertas de nuvem.

Da perspectiva dos indicadores de qualidade, os parâmetros de Qualidade de Serviços (QoS) podem ser usados diretamente para definir a confiabilidade, o tempo de resposta, a segurança e a integridade das entidades em nuvem. A integridade pode ser verificada por meio de habilitação de rastreabilidade e validação de conformidade. A segurança é um aspecto muito importante da qualidade da nuvem. Somente entidades comerciais ou usuários autorizados podem acessar os recursos corretos de Cloud Computing. Controle de acesso, privacidade e proteção das entidades de nuvem formam os aspectos da base de confiança e segurança no CCOA.

Do ponto de vista de governança, muitas práticas recomendadas de governança SOA podem ser emprestadas para habilitar a nuvem Ambiente de computação e ofertas de serviços. Por exemplo, para lançar uma iniciativa de nuvem, um centro de excelência ou competência pode ser formado para melhor comunicação e coordenação entre líderes de negócios

e equipes técnicas. Na fase de execução do projeto, monitoramento, medição, cobrança e tratamento de exceções devem ser habilitados e coordenados em todas as linhas de negócios que produzem, operam e gerenciam vários módulos das ofertas de nuvem.

3. Estudos de caso do CCOA

Existem vários cenários de uso do CCOA proposto e nenhuma restrição específica com o design de partes específicas do CCOA. Por exemplo, podemos usar o CCOA para fazer um planejamento estratégico de alto nível para um empresa para executar a iniciativa de transformação em nuvem. Também podemos usar o CCOA para construir uma pilha de infraestrutura para reduzir o custo operacional. Da mesma forma, podemos usar o CCOA para orientar as práticas de desenvolvimento ou implantação de um aplicativo de computação em nuvem. Também podemos criar novas aplicações de Cloud Computing de valor agregado com base nos ativos existentes de forma sistemática, seguindo os princípios do CCOA.

Para um cenário de computação em nuvem interconectado, o CCOA pode ser usado como base arquitetônica para orientar o design, desenvolvimento, implantação e gerenciamento da entrega de serviços colaborativos na cadeia de valor da nuvem. Do ponto de vista das metodologias, podemos usar a abordagem de baixo para cima para identificar recursos e provisionamento de infraestrutura. Também podemos usar a abordagem de cima para baixo para criar ofertas de nuvem e aproveitar ou criar infraestrutura de nuvem para dar suporte às ofertas.

Nesta seção, apresentamos dois exemplos de ofertas de nuvem baseadas em CCOA. O primeiro caso é sobre Infrastructure Cloud, que é um exemplo de infraestrutura como serviço. O segundo caso estuda o Business Cloud, que é um exemplo de processo de negócio como serviço em ambiente de Cloud Computing.

3.1 Infraestrutura em Nuvem

Neste estudo de caso, construímos nossa própria nuvem privada para permitir o compartilhamento de recursos de computação. O cenário de uso é oferecer aos solicitantes de serviço um modelo de pagamento para uso para fornecer aos servidores um conjunto selecionado de pacotes de software pré-instalados para reduzir o tempo de produção com base nos envios de solicitações.

Usamos os 7 princípios para instanciar o CCOA para uma Nuvem de Infraestrutura para fornecer Oferta de Provisionamento de Servidor para clientes em nuvem. Dentro do ecossistema de nuvem (Módulo 1), o módulo Cloud Ecosystem Management (1A) fornece gerenciamento de contas de usuário de serviços da Web. O Cloud Vendor Dashboard (1C) é responsável por lidar com o processo de login e fornece um portal da Web para que os fornecedores de nuvem acessem o ecossistema de nuvem. O módulo Cloud Partner Dashboard (1C) na Figura 2 não é usado neste caso. A razão é que toda a capacidade de virtualização de hardware e software é fornecida exclusivamente por fornecedores de nuvem internos. Nenhum parceiro de nuvem está envolvido nesta Nuvem de Infraestrutura. O Cloud Client Dashboard oferece suporte ao processo de entrada e a um portal baseado na Web para que os clientes em nuvem enviem solicitações de pedidos e acessem informações relacionadas nos servidores provisionados. O WebSphere Portal Server no Módulo 2 foi usado para implementar portlets relacionados à UI.

No módulo de virtualização (Módulo 2), o Cloud Core Infrastructure inclui um grande número de servidores, data centers e pacotes de software de suporte, como WebSphere Application Server (WAS), banco de dados DB2 e LDAP. Neste estudo de caso, usamos o WebSphere Process Server para suportar processos BPES e o WebSphere Portal Server para suportar portlets. O LDAP é usado para gerenciar usuários e informações de controle de acesso. O Cloud Infrastructure Management é ativado pelo Tivoli Enterprise Portal Server e Tivoli Provisioning Manager, para manipular a virtualização e o fornecimento de hardware neste Infrastructure Cloud.

No módulo Orientação a Serviços (Módulo 3), identificamos e definimos um conjunto de serviços reutilizáveis comuns com base no pensamento orientado a serviços. Nuvem Serviços de negócios horizontais contém serviço de gerenciamento de pedidos na nuvem, serviço de gerenciamento de recursos, serviço de notificação por e-mail e gerenciamento de fluxo de trabalho serviço. Aqui eles são todos implementados como serviços da Web. Os serviços de negócios verticais em nuvem do módulo 3 não são usados neste estudo de caso, uma vez que essa nuvem de oferta de provisionamento de infraestrutura não aproveita nem fornece serviços de negócios específicos do domínio ou do setor.

No módulo Cloud Core (Módulo 4), Cloud

Os serviços de provisionamento são suportados por um conjunto de processos de negócios, como o provisionamento de serviços (ServiceP), Desprovisionamento de Serviço (ServiceDeP) e Contrato Alteração incluindo Extensão e Contrato de Contrato

Terminação. Os serviços de provisionamento em nuvem são principalmente usado por fornecedores de nuvem para compartilhar seletivamente recursos de nuvem. O Serviço de Assinatura em Nuvem é habilitado por Oferecendo processos de negócios de Criação e Criação de Pedidos.

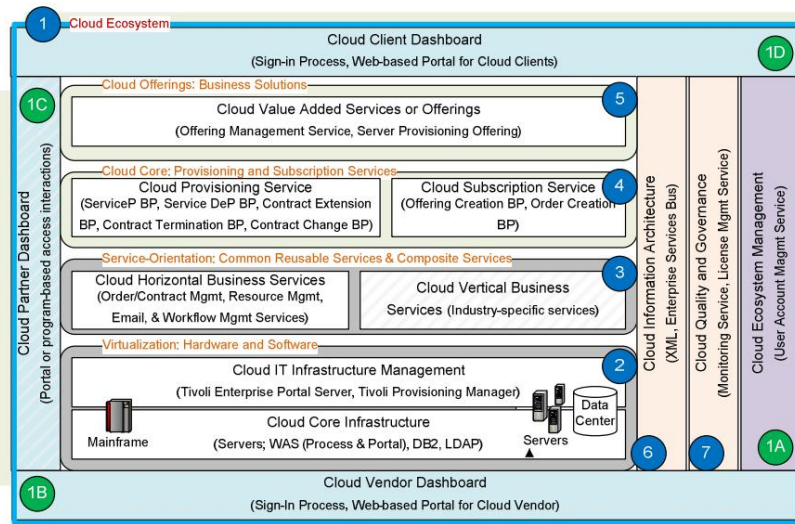


Figura 2. Arquitetura instanciada para infraestrutura em nuvem

No módulo Oferta de Nuvem (Módulo 5), a Nuvem de Infraestrutura fornece Oferta de Provisionamento de Servidor apenas para clientes em nuvem. Mais serviços de valor agregado podem ser habilitados e oferecidos nesta Infrastructure Cloud ambiente sem alterar os módulos básicos de sua arquitetura. Por exemplo, a oferta de armazenamento em nuvem pode ser criada para que o fornecedor de nuvem compartilhe seus data centers com clientes em nuvem. Todos os negócios de criação e provisionamento de serviços de valor agregado em nuvem são gerenciados pelo Serviço de Gerenciamento de Ofertas, que é realizado em um serviço Web neste estudo de caso.

No módulo Cloud Information Architecture (Módulo 6), definimos estruturas de dados para pedidos de assinatura, contratos, SLAs, informações de projetos e cenários de negócios. Essas estruturas de dados e dados são armazenados em Banco de dados DB2 no módulo Virtualização (Módulo 2). Também definimos estruturas de dados para usuários e grupos. No módulo Virtualização, o LDAP é usado para capturar as informações relacionadas ao usuário e controle de acesso. A tecnologia Enterprise Services Bus (ESB) é usada para implementar o roteamento e a transformação de mensagens no módulo de arquitetura de informações em nuvem.

No módulo Cloud Quality and Governance (Módulo 7), acordos de nível de serviço, contratos e estatísticas de recursos são definidos com base nas estruturas de dados definidas no módulo Cloud Information Architecture. Por exemplo, o horário de início e término dos serviços de provisionamento ou desprovisionamento são capturados neste módulo para garantir que os serviços sejam entregues de acordo com os contratos. No caso de exceções, também podemos usar este módulo para rastrear de onde vêm e fornecer soluções.

Da perspectiva de ativação, usamos o Tivoli Enterprise Portal Server para monitorar as operações de serviço no módulo Virtualização (Módulo 2). Enquanto isso, O Tivoli Provisioning Manager pode planejar e monitorar todas as tarefas de fornecimento. No aspecto de governança, organizações como o Centro de Excelência em Computação em Nuvem e a Equipe de Operações em Nuvem são criadas para apoiar seu planejamento estratégico geral, projeto e revisão de arquitetura, alinhamento e coordenação. A produtividade geral, a reutilização e a relação custo-benefício também são identificadas e refletidas no design da arquitetura da solução em nuvem com base nos sete princípios do CCOA. Por exemplo, algumas políticas sobre quando usar serviços da Web e onde alocar máquinas também fazem parte do processo de governança.

Como exemplo, usamos o Cloud Vendor Dashboard para monitorar os usos e estatísticas de recursos que são concedidos pelo módulo Cloud Quality and Governance.

A Figura 3 ilustra o gráfico de dados estatísticos dos recursos de Cloud Computing. Este diagrama de uso mostra quantos recursos estão disponíveis para assinatura de clientes em nuvem. Os recursos são categorizados com base nos tipos de servidor e sistemas operacionais (por exemplo, Windows e Linux). Nesse caso, existem Xen-VM, xSeries, Xen Cluster, VMware, DynamicP6 e DynamicP. Temos dois sistemas operacionais (Windows e Linux) para provisionamento.

Conforme mostrado na Figura 3, há uma máquina Xen-VM assinada com Linux e uma Xen-VM para assinatura. Existem 17 xSeries para assinatura. Existem 13 máquinas Xen Cluster assinadas com Linux e 9 máquinas para assinatura. Há um inscrito

Máquina VMware com Windows e uma máquina VMware para assinatura. Há 1 DynamicP6 subscrito com Linux e 5 máquinas para subscrição.

Existem 4 servidores DynamicP para clientes em nuvem para se inscrever.

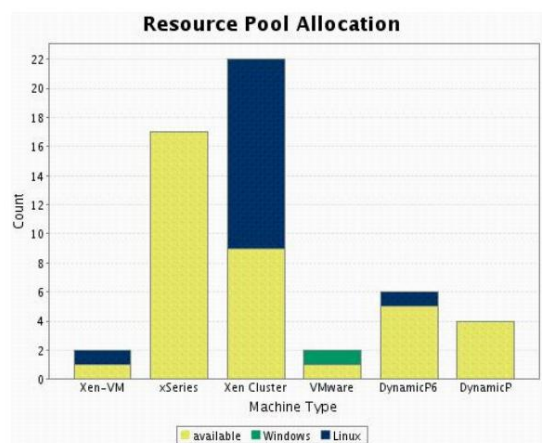


Figura 3. Estatísticas de recursos de nuvem

A Figura 4 exibe uma tela de criação do pedido no Painel do cliente em nuvem. Os clientes de nuvem podem criar pedidos depois de ler e aceitar o contrato de licença que é mantido no módulo Cloud Quality and Governance.

OS	Type	No. of CPUs	Memory(GB)	CPU Speed(MHz)	Storage(GB)	Quantity	Available	
<input checked="" type="radio"/> Windows	Xen-VM	2	2	3200	20	1	1	Add to Cart
<input type="radio"/> AIX	LPAR-PS	2	2	2100	25	1	4	Add to Cart
<input type="radio"/> Linux	Xen-VM	2	2	3200	20	1	1	Add to Cart
<input type="radio"/> LAMP	Xen-VM	2	2	3200	20	1	1	Add to Cart

Clustered offerings: select server(s)							
OS	Type	No. of CPUs	Memory(GB)	CPU Speed(MHz)	Storage(GB)	Nodes	Available
<input type="radio"/> Linux Cluster OS	Xen Cluster	2.0	2	3200	250	2	9

Figura 4. Uma tela de exemplo do painel do cliente do Cloud

3.2 Nuvem de Negócios

O Business Cloud abrange todos os cenários de processos de negócios como serviço no ambiente de Cloud Computing. O CCOA pode ser usado para dar suporte às ofertas de nuvem da nuvem de negócios. Nesta seção, empregamos uma nuvem pública para demonstrar o uso do CCOA para permitir que os negócios ofereçam ofertas de nuvem. Para tornar um processo de negócios mais inteligente, a inteligência humana pode ser introduzida nele como um subprocesso, tarefa humana ou tomada de decisão dentro de uma tarefa.

Conforme ilustrado na Figura 5, uma empresa chamada Cloud Publishing Business (CPB) gostaria de explorar a Plataforma de computação em nuvem para ajudar a estender seu processo de negócios de proteção de direitos autorais que inclui três subprocessos principais. O primeiro subprocesso é a Verificação Automática, que aciona uma ferramenta de detecção de plágio de terceiros para gerar um alerta com um relatório resumido quando um novo manuscrito é enviado. Muitos cenários de negócios podem ser

O usuário da nuvem pode escolher os recursos disponíveis e suas configurações, como sistema operacional (SO), tipo de servidor, número de CPUs, tamanho da memória (GB), velocidade da CPU (MHz), tamanho do armazenamento (GB) e quantidade mostrada na Figura 4.

O número correspondente de recursos disponíveis também é mostrado no painel com base na seleção das configurações.

Existem duas categorias de ofertas de servidor: padrão e em cluster. Para ofertas em cluster, os usuários da nuvem podem escolher o número de nós na solicitação de pedido. Feito o processo de seleção de configuração, o pedido pode ser adicionado ao carrinho de compras. As telas subsequentes no painel permitem que os usuários escolham middleware e visualizem as informações de preço.

O painel também permite que os usuários selecionem um membro da equipe para gerenciar ou usar esses recursos. Depois que um pedido é enviado, os gerentes de projeto ou administradores do Cloud Vendor Dashboard podem revisá-lo, aprová-lo, rejeitá-lo ou excluí-lo. Após a aprovação do pedido, o usuário poderá visualizar seu status. Todas as lógicas de negócios e fluxos de tela são implementados como processos e portlets BPEL. A estrutura de dados de um pedido na Figura 4 é definida no módulo Cloud Information Architecture (Módulo 6) mostrado em Figura 2.

inteligência para analisar o relatório resumido validando ou corrigindo as descobertas e gerar um relatório novamente. Por exemplo, um artigo publicado em anais de conferência pode ser aprimorado e submetido a um periódico para consideração, desde que este artigo tenha pelo menos 30% de novos conteúdos. Ou este artigo pode ser aprimorado com base em suas próprias entradas de blog no espaço de rede social. Nesses casos, os seres humanos podem realizar um trabalho mais eficaz do que os computadores em termos de validação, correção ou aprimoramento de um manuscrito. Este é o segundo subprocesso denominado Validação e Análise Inteligentes.

Neste estudo de caso de nuvem de negócios, assumimos que a CPB foi convencida a executar esse subprocesso de Validação e Análise Inteligente de maneira escalável. Então o desafio passa a ser como acessar rapidamente seres humanos qualificados para realizar um serviço de validação e análise com garantia de qualidade. Como exemplo, o CPB aproveita o mercado de força de trabalho, como o Amazon Mechanical Turk (AMT) [11] para implementar esse subprocesso. Usamos o CCOA para apoiar as ocupações de execução de negócios e

infraestrutura de apoio. Conforme mostrado na Figura 5, o mercado de força de trabalho é habilitado pelo Amazon Mechanical Turk, que implementa os principais módulos do CCOA para dar suporte aos negócios de publicação baseados em nuvem. O que o CPB precisa fazer é aproveitar as APIs Mechanical Turk em seu Cloud Client Dashboard para criar seu próprio Portal CPB, que pode coexistir com o portal do cliente da Amazon. Do ponto de vista do modelo de negócios, o CPB pode criar um Serviço de Detecção de Direitos Autorais para ser usado internamente ou como uma oferta de nuvem de valor agregado

para outros editores ou autores assinarem. O Portal CPB fornece interfaces para que os clientes CPB enviem requisitos e descrições de cargos. CPB lança processo geral de negócios de Detecção de Direitos Autorais e gera o relatório de resumo inicial após a conclusão do primeiro subprocesso. Em seguida, o CPB Portal cria as Tarefas de Inteligência Humana (TITs) correspondentes com base nas APIs ou serviços Web do Mechanical Turk da Amazon.

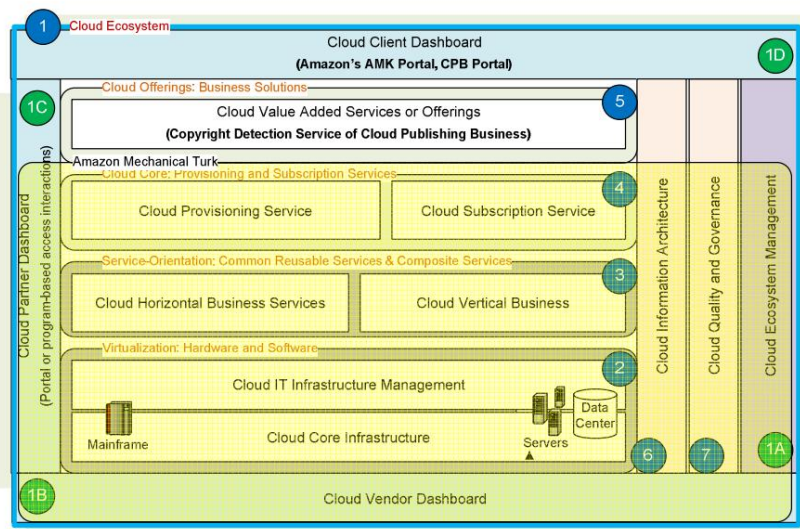


Figura 5. Arquitetura para negócios de publicação em nuvem

Teoricamente, milhões de HITs e requisitos de qualificação podem ser criados automaticamente no CPB Portal e entregues à força de trabalho Mechanical Turk da Amazon

Mercado. Somente o trabalho qualificado será aprovado e pago através da AMT. Os resultados do segundo subprocesso (Validação e Análise Inteligentes) podem ser retornados programaticamente ao Portal CPB.

No terceiro subprocesso Tomada de Decisão e Notificação, a equipe de administração de casos do CPB utiliza um modelo de tomada de decisão para resumir os resultados da investigação com uma recomendação e anexar relatórios de análise dos dois subprocessos anteriores e entregá-los à equipe de tomada de decisão. Assim que uma decisão for tomada, a equipe de administração do caso enviará notificações ao consumidor de nuvem de negócios que

solicitou este Serviço de Proteção de Direitos Autorais na nuvem de negócios do CPB.

Pode ser um bom exercício para os profissionais técnicos mapear as principais capacidades da AMT para o módulos correspondentes no CCOA. A conexão entre o AMT e o Amazon Elastic Compute Cloud (Amazon EC2) [2] pode ser habilitada na oferta de nuvem de negócios do CPB por meio de serviços web da Amazon. No entanto, não é abordado neste artigo devido a restrições de espaço.

4. Trabalhos Relacionados e Discussões

Na área de provisionamento de aplicativos ou processos de negócios, publicamos vários artigos sobre a estrutura de perfis [10][16], processos de provisionamento e assinatura, tratamento de solicitações em lote e monitoramento [9][17]. No desenvolvimento de aplicativos para o ambiente virtualizado de computação Grid, propusemos o conceito Grid Solution Sphere para liberar o valor comercial da Grid Computing [14][15][17]. As práticas acumuladas desses exercícios foram usadas para articular o provisionamento e a aplicação aspecto de composição do CCOA.

Na área de orientação a serviços, SOA Solution Stack (S3) define uma arquitetura de referência SOA no contexto de design e entrega de soluções [8]. O CCOA proposto também é uma evolução natural dessas práticas e do S3.

Uma vez que um sistema de Cloud Computing é colocado em operação, há dois custos principais associados, ou seja, taxa de serviço de manutenção e custo de consumo de energia. Existe uma tendência de que a construção de novos data centers de Internet reduz drasticamente o consumo de energia e o uso de espaço físico. Garantir que as tecnologias verdes façam parte da fase de design é muito importante para operações eficazes de uma plataforma de computação em nuvem no futuro. O Global Technology Outlook 2008 da IBM Research identificou os Centros de Dados de Escala da Internet como uma resposta para abordar negócios de infraestrutura em torno da Computação em Nuvem [18].

O objetivo da abordagem proposta é reunir o poder da SOA e da virtualização para ajudar a perceber e explorar o valor comercial da computação em nuvem. Gostaríamos de incentivar a comunidade de pesquisa a trabalhar em conjunto para refinar o CCOA e focar na análise teórica do CCOA para construir bases sólidas para Cloud Computing. Quanto aos tópicos de pesquisa futura, trabalharemos nos blocos de construção arquitetônicos detalhados e nos padrões de interação para todos os dez módulos definidos no CCOA, bem como nas interfaces de serviço padrão (ou APIs) entre os blocos de construção arquitetônicos orientados à solução e sua infraestrutura de hospedagem em Cloud Computing. Por meio dessas interfaces, a infraestrutura de hospedagem e os serviços ou ofertas em nuvem de valor agregado determinam cooperativamente provisionamento adequado de recursos e ações de gerenciamento. Por exemplo, uma plataforma de computação em nuvem fornece serviços básicos (por exemplo, estilo Hadoop [5] e além) para ajudar a criar recursos de processamento de dados ultraescaláveis e resilientes em soluções de negócios. Além disso, pode-se construir rapidamente serviços novos ou compostos que alavancam ativos e aplicativos orientados a serviços existentes.

5. Conclusões

Neste artigo, propomos a Computação em Nuvem Arquitetura Aberta (CCOA) baseada em sete princípios de arquitetura e dez módulos de arquitetura, integrando o poder da arquitetura orientada a serviços (SOA) e tecnologia de virtualização de hardware e software. Também apresentamos uma nuvem de infraestrutura como um estudo de caso para ilustrar como usar o CCOA para permitir o compartilhamento de recursos em nível de infraestrutura como uma oferta de nuvem. Também estudamos uma nuvem de negócios para ilustrar como usar o CCOA para separar a preocupação do design de negócios desde a habilitação da infraestrutura até o Cloud Publishing Business. Os estudos de caso e a análise mostraram que o CCOA proposto é uma ferramenta extensível e configurável. arquitetura para fornecer orientação normativa e habilitar infraestrutura, software, aplicativo e compartilhamento de processos de negócios de forma unificada.

Referências

- [1] Liang-Jie Zhang, Carl K Chang, Ephraim Feig, Robert Grossman, Keynote Panel, Business Cloud: Bringing The Power of SOA and Cloud Computing, pp. xix, 2008 IEEE International Conference on Services Computing (SCC 2008), julho 2008
- [2] Amazon Elastic Compute Cloud (Amazon EC2), <http://aws.amazon.com/ec2/>, 2009 [3] Serviço XDriver Box, <http://www.box.net/xdrive>, 2009
- [4] Serviço Microsoft Skydrive, <http://skydrive.live.com/>, 2009
- [5] Hadoop Open Source <http://hadoop.apache.org/core/>, 2009 [6] Modelo OSI, <http://www.osi.org>, 2009 [7] John Y. Sayah, Liang-Jie Zhang, Habilitação de colaboração empresarial sob demanda com serviços da web, Decision Support System, 40 (2005), pp.107-127.
- [8] Ali Arsanjani, Liang-Jie Zhang, Michael Ellis, Abdul Allam, Kishore Channabasavaiah, "S3: A Service-Oriented Reference Architecture," IT Professional, vol. 9, não. 3, pp. 10-17, maio/junho de 2007
- [9] Liang-Jie Zhang, Henry Chang, Tian Chao, Jen-Yao Chung, Zhong Tian, Jingmin Xu, Yingnan Zuo, Shunxiang Yang, Qingyun Ao, A Manageable Web Services Hub Framework and Enabling Technologies for e-Sourcing, IEEE Conference em Sistema, Homem e Cibernética (SMC'02), 2002.
- [10] Shun Xiang Yang, Liang-Jie Zhang, Tian Chao, Jing Min Xu, Ying Nan Zuo, Zhong Tian e Henry Chang: estrutura e sistema de perfil adaptável para provisionamento de serviços em soluções de e-business. Pesquisa e aplicativos de comércio eletrônico 3(2): 139-151 (2004)
- [11] Amazon Mechanical <http://aws.amazon.com/mturk/>, 2009 [12] Liang-Jie Zhang, Jia Zhang, Hong Cai, Services Computing, Springer and Tsinghua University Press, 2007, ISBN: 978-3-540-38281-2, julho de 2007 [13] L.-J. Zhang, N. Zhou, Y.-M. Che, A. Jaladeen, K. Ponnalagu, RR Sindhgatta, A. Arsanjani e F. Bernardini, SOMA-ME: Uma plataforma para o design orientado a modelos de soluções SOA, IBM SYSTEMS JOURNAL, 47, NO 3, 2008, pp.397-413.
- [14] Liang-Jie Zhang, Jen-Yao Chung, Qun Zhou, Discover Grid Computing, IBM developerWorks Journal, fevereiro de 2003, pp.14-19.
- [15] Liang-Jie Zhang, Qun Zhou, Jen-Yao Chung, Desenvolvendo aplicativos de computação em grade, IBM developerWorks Journal, maio de 2003, pp.10-15.
- [16] Jing Min Xu, Ying Nan Zuo, Shun Xiang Yang, Zhong Tian, Henry Chang, Liang-Jie Zhang, Tian Chao: Portal de Membros e Sistema de Aprovisionamento de Serviços para uma Infraestrutura de Hubs: E Hub Gerenciado. ICEIS (4) 2003: pp. 143-150.
- [17] Liang-Jie Zhang, Haifei Li, Herman Lam, Toward a Business Process Grid for Utility Computing, IT Professional, Volume: 6 Issue: 5, set.-out. 2004, pp. 408-418 [18] IBM Outlook, [http://www03.ibm.com/procurement/proweb.nsf/objectdocswebview/file3+-+ibm+gto+overview+-+agerwala/\\$file/3+-+ibm+gtooverview+-+agerwala.pdf](http://www03.ibm.com/procurement/proweb.nsf/objectdocswebview/file3+-+ibm+gto+overview+-+agerwala/$file/3+-+ibm+gtooverview+-+agerwala.pdf) [19] Aplicativos da Web do Google para Comunicação e Colaborações. <http://www.google.com/apps>