



Proposta de Piloto

Grupo de Trabalho – Segunda Fase

GT-PID: Plataforma IaaS Distribuída

Luís Henrique Maciel Kosmowski Costa - UFRJ

04/09/2014

1. Concepção

1.1. Resumo

O GT-PID possui como objetivo o compartilhamento de recursos computacionais por universidades e centros de pesquisa brasileiros, utilizando um modelo em nuvem de Infraestrutura como Serviço (IaaS) geodistribuída. Isso permite o aumento da capacidade global do sistema, através da agregação de recursos de cada instituição

participante. Dessa forma, a plataforma proposta pelo GT-PID torna mais eficiente o uso da infraestrutura de universidades e centros de pesquisa e, conseqüentemente, dos recursos financeiros destinados à pesquisa. O usuário final do GT-PID utiliza máquinas virtuais hospedadas na infraestrutura, que executam sistemas operacionais completos, possuindo flexibilidade na utilização dessas máquinas. A plataforma do GT-PID é baseada no OpenStack, que é uma ferramenta de gerenciamento de nuvem em crescente utilização. Na Fase 1, foram realizadas modificações do OpenStack para o cenário geodistribuído. Foram modificados os processos de escalonamento e a hierarquia de usuários do OpenStack para permitir que as instituições possam gerenciar seus próprios sítios. Além disso, a interface web de gerenciamento do OpenStack foi modificada para refletir tais mudanças. Na Fase 2, pretende-se instalar um piloto da plataforma desenvolvida em três universidades, através da definição de procedimentos de instalação e refinamento de diversas características do protótipo da Fase 1, como a inclusão de ferramentas de monitoramento, configuração de rede entre máquinas virtuais de diferentes sítios, além de definição de procedimentos de atualização com as versões mais recentes do OpenStack.

1.2. Abstract

The goal of GT-PID is to share computational resources among Brazilian universities and research centers, using the Infrastructure as a Service (IaaS) model in a geodistributed fashion. This yields a global capacity increase, via resource aggregation of each participant institution. Hence, the platform proposed by GT-PID makes more efficient the utilization of the infrastructure from universities and from research centers and, consequently, of the financial investments destined to research. The final user of GT-PID uses virtual machines hosted on the available infrastructure. These virtual machines execute complete operating systems, providing users total utilization flexibility. The GT-PID platform is based on OpenStack, which is a tool for cloud management with an ascending popularity. In Phase 1, modifications to OpenStack were conducted to enable a geo-distributed scenario. Scheduling tasks and the users' hierarchical organization in OpenStack were changed to allow institutions to manage their own sites. These changes were reflected onto the OpenStack web interface for management. In Phase 2, we plan to install a pilot platform developed in three universities, which requires the definition of installation procedures and refinement of several characteristics of the prototype from Phase 1. Among these characteristics, we can cite the inclusion of monitoring tools, the configuration of the network among virtual machines located in different sites, and the definition of procedures for updates to newer versions of OpenStack.

1.3. Descrição do produto/serviço

O objetivo do projeto GT-PID é promover o compartilhamento de recursos computacionais entre universidades e centros de pesquisa a partir do compartilhamento em nuvem de todos os recursos existentes. Por um lado, provê-se uma capacidade global do sistema superior à oferecida localmente. Por outro lado, diminui-se a

ociosidade dos recursos computacionais, aumentando a eficiência e o retorno do investimento financeiro aportado na pesquisa. Para isso, utiliza-se o conceito de máquinas virtuais (*Virtual Machines* - VMs) que são oferecidas a cada solicitação de utilização da infraestrutura. As VMs são disponibilizadas em um conjunto responsável por executar as aplicações desejadas. Assim, o usuário possuirá acesso em nível de administrador a essas VMs e poderá instalar suas aplicações. O usuário, então, possuirá total flexibilidade na escolha dessas aplicações, caracterizando um serviço em nuvem IaaS (*Infrastructure as a Service*). Por exemplo, o usuário poderá utilizar as VMs para realizar suas simulações, instalando ferramentas específicas para sua pesquisa.

A infraestrutura IaaS do GT-PID permite que usuários criem e utilizem VMs, que são hospedadas nos Servidores de VMs inicialmente distribuídos pelas universidades participantes. Os Servidores de VMs consistem em PCs comuns executando uma plataforma de virtualização (hipervisor) que, no contexto atual do GTPID, é o KVM [1]. Para a gerência da infraestrutura há dois tipos de administradores no GT-PID: global e local. O administrador global é responsável pela gerência da infraestrutura IaaS como um todo, enquanto os administradores locais são responsáveis por operações de gerenciamento dentro de um único sítio. Um sítio é definido neste projeto como o conjunto de Servidores de VMs e de Discos disponibilizado por uma universidade ou laboratório de pesquisa, ou seja, uma unidade administrativa autônoma. Essa definição de administradores globais e locais foi uma proposta do GT-PID para permitir que recursos em sítios distribuídos pudessem ser gerenciados na nuvem de forma organizada, atribuindo certa autonomia às instituições participantes. Dentre as operações características do administrador local e global, estão as operações de gerenciamento que incluem, respectivamente, a migração local e migração global de VMs.

Assim, o GT-PID possui impacto potencial para a comunidade acadêmica e científica brasileira através do aumento global da capacidade de recursos computacionais disponíveis, acarretando em uma melhora da qualidade da pesquisa e dos serviços das universidades e centros de pesquisa. Além do aumento da capacidade, o serviço IaaS geodistribuído permitirá a implantação de serviços tolerantes a falhas, devido à eliminação de ponto únicos de falha na infraestrutura. Finalmente, porém não menos importante, a implantação do serviço IaaS do GT-PID poderá tornar as comunidades e centros de pesquisas brasileiros autônomos no que concerne a infraestrutura de TI, evitando a utilização de nuvens comerciais (p. ex. Amazon e Rackspace), que são sujeitas à legislação do país hospedeiro e a alterações dos termos de uso.

1.4. Identificação do público alvo

O GT-PID originalmente foi idealizado para prover infraestrutura computacional a pesquisadores de universidades brasileiras desde que tenham acesso com um mínimo de qualidade à Internet. Tal acesso poderia ser obtido pelas universidades interconectadas pela RNP. Assim, cada grupo de pesquisa conecta um conjunto de seus servidores à infraestrutura de nuvem, e os torna disponíveis a todos os usuários da plataforma do GT-PID. Em contrapartida, o grupo de pesquisa pode usufruir da infraestrutura global do GT-PID. É importante notar que, como cada vez mais a computação é utilizada na pesquisa e os laboratórios são equipados com recursos computacionais, o GT-PID não se limita para uso apenas de grupos de computação e engenharia. A infraestrutura do GT-PID pode ser utilizada para pesquisas em qualquer área do conhecimento na qual se necessite de poder computacional. Pode-se vislumbrar em um primeiro momento pesquisas realizadas por Físicos, Biólogos, Matemáticos, Cientistas Sociais, etc. Em um segundo momento, devido à flexibilidade oferecida pelo serviço IaaS, pode-se vislumbrar aplicações não científicas que também necessitam de infraestrutura computacional, como servidores de e-mail, sites web, etc. Em caráter mais assistencial, o GT-PID ainda pode ser estendido para instituições que não possuam infraestrutura computacional passível de ser integrada à plataforma. Por exemplo, escolas, hospitais, ou mesmo laboratórios de pesquisas emergentes com recursos escassos desde que essa aplicação não esgote o poder computacional do sistema. Por exemplo, os recursos destinados a usuários dessas instituições poderiam ser limitados em uma quantidade baixa. As instituições menos privilegiadas em termos de recursos podem utilizar aplicativos do tipo desktop remoto e ter acesso a máquinas virtuais com mais recursos computacionais, de forma inclusiva. Assim, o modelo de serviço e a plataforma desenvolvida no GT-PID podem fortalecer a popularização da computação por todo o Brasil, com tecnologia e desenvolvimento nacionais.

2. Definição do piloto

2.1. Arquitetura do piloto

2.1.1. Arquitetura Básica

A arquitetura básica adotada pelo GT-PID baseia-se na plataforma OpenStack [3], que consiste em uma solução completa para gerenciamento de nuvem. Para tal, o GT-PID emprega diversos Servidores de Máquinas Virtuais (VMs) que hospedam as VMs dos usuários, como mostrado na Figura 1. Esses Servidores de VMs são organizados em sítios, que correspondem a universidades ou laboratórios de pesquisa.

Para orquestrar os recursos da nuvem, um Controlador é responsável pelo gerenciamento dos sítios. O Controlador é capaz, por exemplo, de alocar recursos destinados às VMs nos diversos Servidores de VMs. Os Servidores de VMs de cada sítio se comunicam com o Controlador através de túneis VPN (*Virtual Private Network*) estabelecidos pela Internet. A arquitetura também possui um Servidor de VMs e Discos em cada sítio. Os usuários, por sua vez, acessam o Controlador através de uma interface web que permite o gerenciamento do ciclo de vida de suas máquinas virtuais como, por exemplo, a criação e a destruição de VMs. Além disso, a interface web permite ao usuário a visualização do console das VMs.

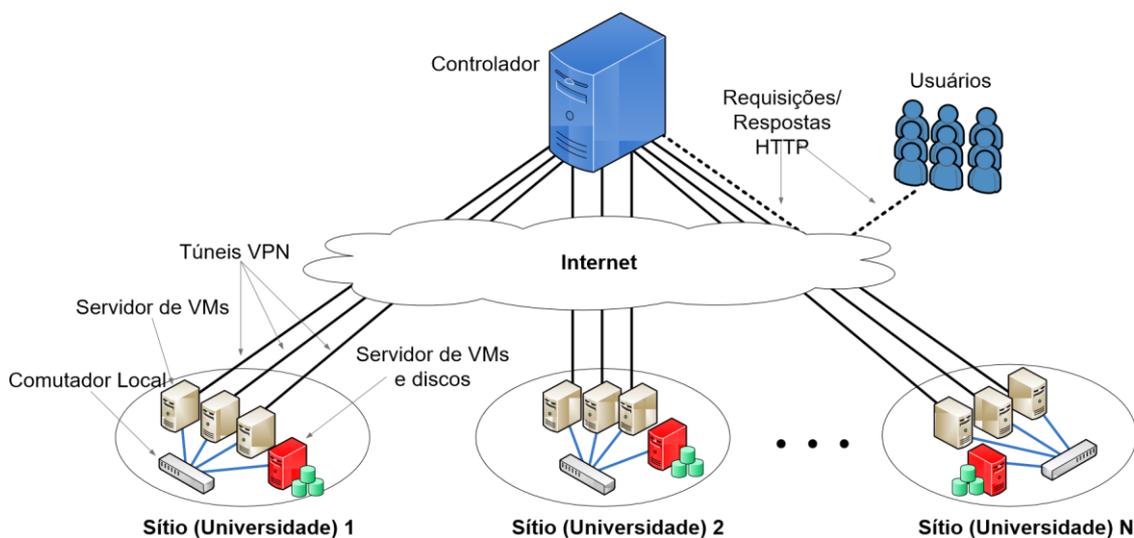


Figura 1 - Visão Geral da Arquitetura do GT-PID.

Além dos Servidores de VMs, cada sítio possui um Servidor de VMs e Discos, que é responsável por hospedar os discos virtuais das VMs, além de realizar todas as funções de um Servidor de VMs. O Servidor de VMs e Discos é importante para agilizar a migração ao vivo dentro de um sítio. A migração ao vivo consiste na transferência de VMs, em tempo de execução, entre dois Servidores de VMs, sem que haja suspensão de seus serviços. Como os discos virtuais estão centralizados no Servidor de VM e Discos, não é necessário realizar cópia de discos entre os servidores envolvidos, acelerando o processo de migração. A cópia de discos, porém, é necessária na migração entre sítios. Por fim, cada sítio possui um comutador local que interliga todas as suas máquinas, possibilitando as comunicações de VMs hospedadas em diferentes servidores e também operações de disco através de NFS (*Network File System*).

As VMs disponibilizadas para o usuário podem executar diferentes sistemas operacionais. O Controlador fornecerá um conjunto de imagens com sistemas

operacionais pré-instalados, e o usuário escolhe qual dessas imagens utilizará em sua VM. Além disso, o usuário pode fornecer ao Controlador sua própria imagem de VM, o que possibilita uma maior flexibilidade na escolha do sistema operacional. Uma VM básica possui um endereço IP privado (ou seja, não acessível pela Internet) e acessa a Internet através de NAT (*Network Address Translation*). Além disso, VMs que se encontram em um mesmo sítio se comunicam entre si através de seus endereços IP privados. Na Fase 2 pretende-se atribuir endereços IP públicos às VMs, de forma que seus serviços estejam acessíveis pela Internet, além de permitir que VMs de diferentes sítios possam se comunicar como se estivessem em uma rede local, através de seus IPs privados. Devido à escassez do IPv4 e o fato de prever-se uma grande quantidade de VMs instanciadas, na Fase 2 será investigado o uso do IPv6 como endereçamento de IP Público.

Ao criar uma nova VM, o usuário pode escolher duas formas possíveis de instanciação: baseada em *imagem* e baseada em *volume*. A primeira consiste na cópia de uma imagem fornecida pelo Controlador, que é destruída após o desligamento da VM. A instanciação baseada em volumes, por sua vez, copia a imagem escolhida para um volume, que é uma unidade lógica de armazenamento persistente. Assim, o usuário possui seus dados gravados mesmo após o desligamento das VMs. A quantidade de volumes disponível para cada usuário é limitada para que usuários diferentes tenham oportunidade de armazenar seus respectivos volumes. Os volumes são hospedados no Servidor de VMs e Discos.

A arquitetura do GT-PID propõe a existência de três tipos de entidade: Usuário Final, Administrador Global e Administrador Local. O Usuário Final é o cliente da nuvem IaaS, que requisita e utiliza VMs para seus fins. O Administrador Global possui total controle sobre a infraestrutura, podendo migrar VMs entre diferentes sítios, adicionar novos usuários, etc. O Administrador Local, por sua vez, é responsável pela administração de apenas um sítio, podendo realizar ações de migração das máquinas no sítio sob seu controle. Dessa forma, o Administrador Local encontra-se em um nível hierárquico intermediário entre o Administrador Global e o Usuário Final. A arquitetura do GT-PID possui, no mínimo, um Administrador Local por sítio. O papel de Administrador Local foi desenvolvido para permitir que as instituições doadoras de recursos mantenham certo nível de controle sobre sua infraestrutura.

A plataforma do GT-PID utiliza o Escalonador de Sítios, desenvolvido na Fase 1. Esse escalonador permite que o Controlador decida, com base nas informações de recursos disponíveis nos sítios, em qual sítio uma requisição de criação de VMs será atendida, caso o usuário não especifique explicitamente o sítio que deseja, o que

também é permitido. Além disso, é possível distribuir as VMs em diversos sítios de forma a aumentar a robustez de um determinado serviço, eliminando pontos únicos de falha. Após a escolha dos sítios utilizados, emprega-se no GT-PID o próprio escalonador do OpenStack, que decide quais Servidores de VMs dos sítios irão hospedar as VMs de uma requisição.

2.1.2. Módulos de Software

2.1.2.1. OpenStack

Para a construção da nuvem IaaS geodistribuída, o GT-PID modificou e utilizou diversos módulos da plataforma OpenStack. O OpenStack é uma plataforma de nuvem de código aberto, utilizado para gerenciar recursos em um centro de dados. Em linhas gerais, o OpenStack gerencia três tipos de recurso: computação (p. ex. processamento e memória), rede e armazenamento. Para tal, fornece um conjunto de serviços e APIs (*Application Programming Interfaces*) que permitem a manipulação da nuvem. Além disso, o OpenStack fornece uma interface gráfica para acessar essas APIs, disponibilizada na forma de interface Web. Os serviços do OpenStack, descritos adiante, são organizados de forma modular. Além do OpenStack, existem diversas plataformas de gerenciamento de nuvem em código aberto como, por exemplo, CloudStack [4] e Eucalyptus [5]. Escolheu-se o Openstack, pois além de atender a todos os requisitos do projeto, essa plataforma possui uma comunidade de desenvolvedores em constante crescimento, o que possibilita maior facilidade na resolução de problemas. Além disso, o OpenStack possui uma arquitetura modular, facilitando sua modificação de acordo com as necessidades do projeto.

Os serviços do OpenStack são divididos em projetos. Atualmente utilizam-se no GT-PID os projetos ilustrados na Figura 2 e descritos a seguir.

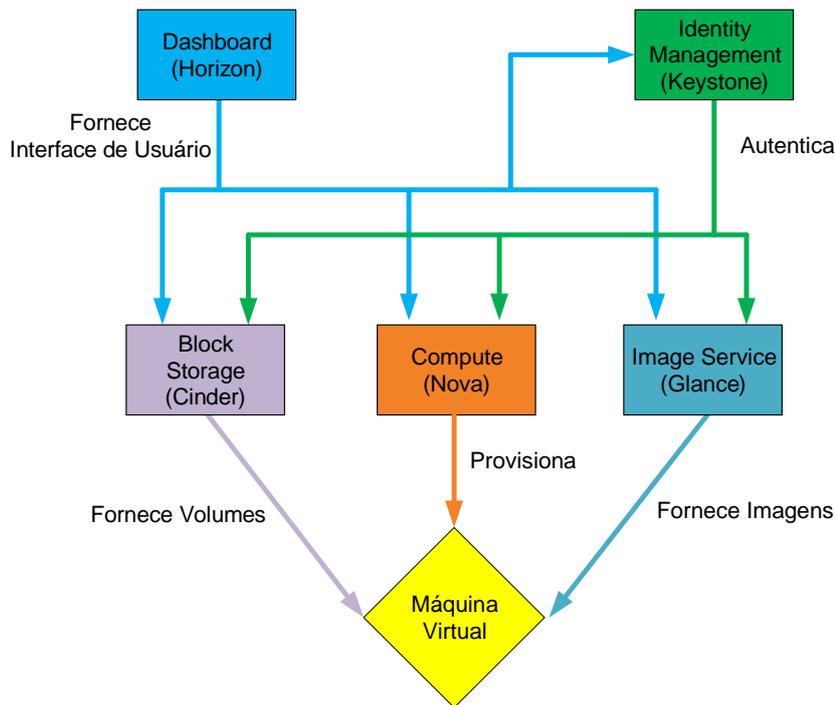


Figura 2 - Projetos do OpenStack utilizados no protótipo.

Horizon

Fornece a interface web do GT-PID se comunicando através de APIs com todos os outros serviços utilizados.

Nova

Gerencia o ciclo de vida completo das máquinas virtuais, como a instanciação e destruição de máquinas virtuais. Além disso, esse serviço escalona as máquinas virtuais na infraestrutura, escolhendo qual sítio e qual servidor físico irá hospedar cada máquina virtual, etc. Outra função importante do serviço Nova é a migração de máquinas virtuais.

Cinder

Fornece armazenamento persistente às máquinas virtuais. Esse armazenamento persistente, chamado de volume na terminologia do OpenStack, pode ser visto como um disco rígido virtual.

Keystone

Realiza o gerenciamento de identidades, sendo responsável, por exemplo, pela autenticação de usuários da infraestrutura. Na arquitetura do piloto, pretende-se integrar o Keystone com o sistema CAFe (Comunidade Acadêmica Federada) da RNP, através de testes com a CAFe Expresso. Essa integração é viável, pois já foi realizada por outro

grupo de trabalho que evoluiu para serviço experimental, o SECNC (Computação em Nuvem para Ciência: Armazenamento de dados).

Glance

Fornecer imagens de máquinas virtuais. A imagem de uma máquina virtual, da mesma forma que um volume, pode ser vista como um disco rígido virtual. Entretanto, a instanciação de máquinas virtuais através de imagem não permite armazenamento persistente. Ou seja, todo conteúdo gravado no disco da máquina virtual será apagado após seu desligamento. Assim, caso o usuário necessite de persistência, o OpenStack copia a imagem para um volume e inicia a máquina virtual a partir desse volume, que é oferecido pelo serviço Cinder. Desta forma as imagens são úteis para fornecer aos usuários sistemas pré-instalados. Além disso, o próprio usuário pode fornecer novas imagens para o repositório gerenciado pelo Glance.

Os projetos descritos acima são divididos em diversos módulos que foram modificados e utilizados para atender os requisitos do GT-PID, como a existência de entidades do tipo Administrador Local e a geodistribuição em diversos sítios. A Figura 3 mostra a distribuição dos módulos e projetos do OpenStack utilizados no GT-PID. Note que os módulos de alguns projetos estão espalhados em diversas máquinas na infraestrutura, como o Nova e o Cinder.

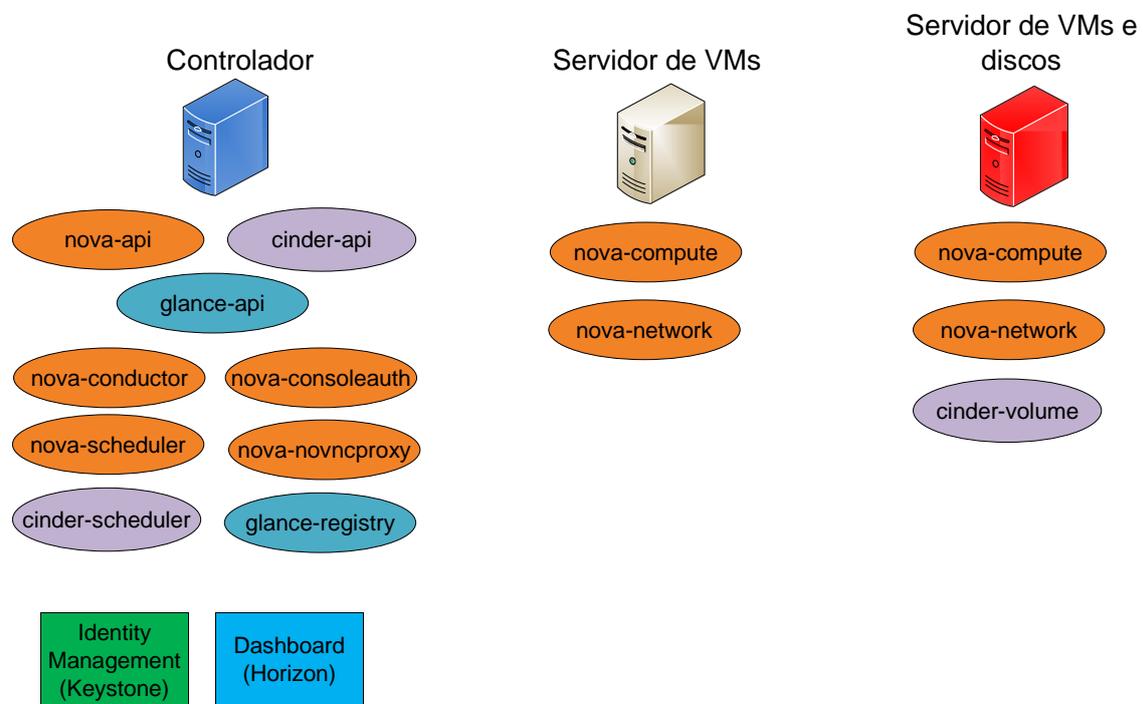


Figura 3 - Utilização dos módulos do OpenStack no GT-PID.

Conforme mostrado na Figura 3, o nó Controlador é responsável por hospedar todos os módulos de APIs e os módulos escalonadores. As APIs são utilizadas para acessar os serviços de cada projeto, enquanto os escalonadores escolhem qual Servidor será responsável por atender uma requisição. Por exemplo, para criar uma VM, realiza-se uma requisição ao módulo `nova-api`. Esse módulo, por sua vez, consulta o `nova-scheduler` para indicar qual Servidor de VMs será utilizado no tratamento da requisição. Após a escolha do Servidor de VMs, o Controlador notifica o módulo `nova-compute` da máquina escolhida. O `nova-compute` utiliza a interação com o Glance e o Cinder, respectivamente, para instalar uma imagem na VM e criar seus volumes. Um processo semelhante ocorre nas interações entre os módulos `cinder-api`, `cinder-scheduler` e `cinder-volume`.

Além de utilizar os módulos citados acima, na Fase 2 do GT-PID pretende-se investigar a integração da plataforma do GT-PID com dois outros projetos do OpenStack: o Neutron e o Ceilometer. O Neutron oferece opções avançadas de redes, como a implementação de redes virtuais entre VMs de sítios diferentes, e suporte ao paradigma *Software-Defined Networking* (SDN). O Ceilometer fornece serviços de medição e monitoramento dos componentes da nuvem. Esse serviço se comunica com os demais serviços do OpenStack para obter as medidas desejadas como, por exemplo, o consumo de processamento por um usuário. Além disso, ele fornece alarmes que são enviados após a utilização de um determinado recurso atingir algum limiar. A partir do alarme, ações podem ser tomadas como, por exemplo, migrar as máquinas do usuário para servidores com maior capacidade de processamento. O Ceilometer fornecerá aos administradores um maior controle da utilização da infraestrutura.

2.1.2.2. DevStack

O protótipo desenvolvido na Fase 1 utiliza a plataforma Devstack para sua instalação. Essa plataforma consiste em um conjunto de *scripts* que instalam um protótipo de nuvem OpenStack de forma simples, através da configuração automática de diversos componentes auxiliares como o banco de dados MySQL e diversos pacotes do sistema operacional Linux. O DevStack também permite observar o funcionamento da plataforma em tempo real, como a troca de mensagens entre os módulos. Dessa forma, o DevStack é ideal para desenvolvimento no OpenStack. Entretanto, o DevStack não é recomendado pela comunidade OpenStack para uso em produção. Dessa forma, o DevStack será mantido para o protótipo de desenvolvimento, mas o piloto será instalado sem o DevStack. O protótipo de desenvolvimento será utilizado para desenvolver novas versões do código do GT-PID e realizar testes de novos módulos ou tecnologias a serem

inseridos no piloto. Após testados, as novas versões do código ou módulos poderão integrar a infraestrutura do piloto.

2.1.2.3. VPN

Os túneis VPN estabelecidos entre o Controlador e os Servidores de VMs são implementados pelo OpenVPN [6]. Para tal, instala-se um servidor OpenVPN no Controlador e os Servidores de VMs assumem o papel de clientes OpenVPN. Assim, cada Servidor de VMs possui um IP privado e pode acessar o Controlador pelo seu endereço IP privado (p. ex. 10.8.0.1). A utilização de VPN torna mais segura a comunicação entre Servidores de VMs e o Controlador, pois as mensagens entre essas máquinas são criptografadas, além de eliminar a necessidade de liberação de cada porta necessária pelo OpenStack nos *firewalls* das universidades. Assim, com o túnel VPN, apenas uma porta necessita ser liberada e todas as outras portas são multiplexadas pelo túnel.

Como explicitado anteriormente, a configuração utilizada do OpenVPN permite uma VPN entre cada Servidor de VMs e o Controlador. Entretanto, essa configuração não permite a comunicação na VPN entre Servidores de VMs, pois o OpenVPN é baseado no modelo cliente e servidor, no qual o Controlador é o servidor. Assim, não é possível a comunicação entre dois clientes (isto é, Servidores de VMs) OpenVPN. A existência de VPNs entre Servidores de VMs é importante, pois no GT-PID existe comunicação entre Servidores de VMs, por exemplo, na migração de VMs. No protótipo desenvolvido na Fase 1 essa comunicação é realizada por um canal não seguro e exige a liberação de uma porta específica no sítio. Para contornar tal problema, na Fase 2 serão investigadas formas escaláveis de implementar VPNs entre todos os Servidores de VMs do piloto.

2.1.2.4 Sistema Operacional Ubuntu

Todas as máquinas físicas do GT-PID executam o sistema operacional Ubuntu Server. Atualmente, a versão utilizada do Ubuntu é a 12.04. Pretende-se na continuidade do projeto atualizar para versões mais recentes.

2.1.2.5 Hipervisor KVM

Para realizar a hospedagem de VMs, os Servidores de VMs utilizam o hipervisor KVM [1]. O hipervisor é uma camada de software situada entre o *hardware* da máquina física e as VMs. O hipervisor controla o acesso ao *hardware* de um Servidor de VMs e realiza alocação dos recursos destinados a cada VM. Na Fase 2 será investigada a

possibilidade de tornar o GT-PID multiplataforma, utilizando diferentes hipervisores em conjunto, como o Xen [2] e o KVM.

2.1.2.6 NFS (*Network File System*)

Os servidores e clientes NFS utilizados são fornecidos por padrão pelo repositório de arquivos do Ubuntu.

2.1.2.7 GIT

Para realizar o desenvolvimento no GT-PID utiliza-se a ferramenta de controle de versão GIT. Para tal, mantêm-se um repositório GIT na UFRJ com o código desenvolvido. Pretende-se na Fase 2 padronizar uma metodologia para sincronizar o repositório do GT-PID com o repositório oficial do OpenStack, de forma a incorporar no projeto melhorias e correções adicionados pela comunidade OpenStack e vice-versa, caso a comunidade julgue conveniente.

2.1.3. Equipamentos necessários

A lista abaixo especifica os equipamentos necessários para a instalação do piloto em cada local. À exceção do Servidor de VMs e Discos na UFF, todas as máquinas relacionadas já estão disponíveis nas instituições participantes. Dessa forma, o Servidor de VMs e Discos da UFF será adquirido com os recursos do projeto de Fase 2.

Controlador (instalado na UFRJ):

PC com processador Intel Core i7-3770 3.4GHz/
RAM 8GB DDR3 / HD 1TB

Sítio da UFRJ:

Servidor de VMs e Discos:

PC com processador Intel Xeon E3-1270V2 3.5GHZ/
RAM 32GB DDR3 / HD 3TB

Servidor de VMs:

PC com processador Intel i7-4930K 3.40GHz/
RAM 32GB DDR2 / HD 3TB

Rede:

Comutador 1Gbe com, no mínimo, 4 portas

Sítio da UERJ:

Servidor de VMs e Discos:

PC com processador Intel Xeon E3-1270V2 3.5GHZ/ RAM 32GB DDR3 /
HD 3TB

Servidor de VMs:

Servidor HP DL320E G8 Intel Xeon E3-1240V2 3.4GHz/ RAM
16GB / HD 500GB

Rede:

Comutador 1Gbe com, no mínimo, 4 portas

Sítio da UFF:

Servidor de VMs e Discos:

PC com processador Intel Xeon E3-1270V2 3.5GHZ/ RAM 32GB DDR3 /
HD 3TB

Servidor de VMs:

PC com processador Intel i7-4930K 3.40GHz/
RAM 32GB DDR2 / HD 3TB

Rede:

Comutador 1Gbe com, no mínimo, 4 portas

2.2. Instituições participantes

Desde a Fase 1 a equipe do GT-PID conta com participantes da UFRJ (Universidade Federal do Rio de Janeiro), UERJ (Universidade do Estado do Rio de Janeiro) e UPMC (*Université Pierre et Marie Curie* - Paris, França). Na Fase 2, pretende-se instalar o piloto em três instituições, através da entrada da UFF (Universidade Federal Fluminense) no projeto. Optou-se, na Fase 2, por escolher instituições geograficamente próximas de forma a facilitar possíveis deslocamentos para configuração do piloto. Após a definição dos procedimentos de instalação de sítios na Fase 2 (veja a Seção 2.3.1) esse requisito não será mais necessário, visto que tais procedimentos poderão ser seguidos por qualquer pessoa com experiência em Linux. Dessa forma, durante a Fase 2 serão analisadas possibilidades de parcerias com mais instituições para uma possível instalação do serviço experimental. A lista abaixo apresenta a relação de cada sítio do piloto e seus respectivos contatos:

- UFRJ – Universidade Federal do Rio de Janeiro

Contato: Prof. Miguel Elias Mitre Campista

- UERJ – Universidade do Estado do Rio de Janeiro

Contato: Prof. Marcelo Gonçalves Rubinstein

- UFF – Universidade Federal Fluminense

Contato: Prof. Igor Monteiro Moraes

Além das parcerias referentes à hospedagem do piloto, o projeto possui as seguintes parcerias:

- UFRJ – Universidade Federal do Rio de Janeiro

Prof. Pedro Braconnot Velloso

Prof. Otto Carlos Muniz Bandeira Duarte

- UPMC – Université Pierre et Marie Curie Prof. Stefano Secci

2.3. Refinamento do protótipo

Para a Fase 2 do GT-PID são necessários os refinamentos descritos nas subseções seguintes.

2.3.1. Instalação e configuração da versão de produção do GT-PID

Como mencionado na Seção 2.1, a instalação do protótipo desenvolvido na Fase 1 se baseia no DevStack, que não é adequado para ambientes de produção. Para a implantação do piloto, a plataforma do GT-PID será instalada manualmente através da definição de procedimentos de instalação do Controlador e dos sítios. Além disso, para a implantação do piloto serão realizados aprimoramentos na segurança da plataforma. Por exemplo, o protótipo da Fase 1 utiliza uma comunicação não segura entre Servidores de VMs para realizar a migração de VMs. Esse fator será corrigido através da configuração escalável de um canal seguro entre todos os Servidores de VMs.

2.3.2. Integração da plataforma à federação CAFe

O protótipo desenvolvido na Fase 1 utiliza o módulo Keystone para controle de acesso à nuvem, que acessa um banco de dados local com informações dos usuários e os papéis que esses exercem na nuvem (isto é, Administrador Global, Administrador Local e Usuário Final). Para a implantação do piloto, o módulo Keystone será modificado para utilizar informações de federação fornecidas pela CAFe. Dessa forma, a autenticação na nuvem poderá ser realizada através das credenciais fornecidas pela instituição do usuário.

2.3.3. Sincronismo com a comunidade de desenvolvimento do OpenStack

A comunidade de desenvolvimento do OpenStack possui um grande número de participantes e é bastante ativa. Dessa forma, a plataforma está em constante aprimoramento e novas versões do código são disponibilizadas periodicamente. O protótipo da Fase 1 foi desenvolvido com base na versão estável Havana do OpenStack, lançada em Outubro de 2013. Essa versão é a penúltima versão estável do OpenStack até a data de submissão deste documento. A atual versão estável é a IceHouse, lançada em Abril de 2014, e a próxima versão, Juno, está prevista para Outubro de 2014. Dado o constante aprimoramento do OpenStack, é necessário atualizar a plataforma do GT-PID a cada atualização estável promovida pela comunidade. Para tal, a equipe do GT-PID entrará na Fase 2 para a comunidade de desenvolvedores OpenStack através da inscrição em Fóruns online de colaboração e assistirá periodicamente as palestras proferidas nos *OpenStack Summit*. O *OpenStack Summit* é uma conferência realizada periodicamente com desenvolvedores e usuários OpenStack. Nessa conferência são proferidas palestras sobre casos de uso do OpenStack e novas funcionalidades adicionadas ao código. Todas as palestras do *OpenStack Summit* são disponibilizadas no canal YouTube da *OpenStack Foundation*. Para sincronizar a plataforma do GT-PID com as atualizações do OpenStack, serão definidas na Fase 2 metodologias baseadas no GIT, visto que essa é a plataforma de controle de versão utilizada pela comunidade OpenStack e também pelo GT-PID na Fase 1. Caso seja conveniente, a equipe do GT-PID poderá enviar partes das modificações realizadas no código à comunidade OpenStack, contribuindo ativamente para o desenvolvimento da plataforma.

2.3.4. Inclusão de ferramentas de monitoramento

Na Fase 2 serão estudados mecanismos de monitoramento de recursos para o uso com a plataforma do GT-PID. Com tais mecanismos será possível analisar estatísticas de uso da plataforma, como consumo em tempo real de rede e de CPU. O OpenStack possui um módulo de monitoramento de recursos denominado Ceilometer, conforme mencionado na Seção 2.1. Esse módulo é um candidato para executar as funções de gerenciamento a serem implementadas na Fase 2.

2.3.5. Inclusão de opções avançadas de rede

Em relação à rede, duas questões permaneceram em aberto na Fase 1. A primeira consiste em realizar a comunicação entre VMs que estiverem em sítios diferentes. A solução direta para tal é atribuir IPs públicos às VMs de forma que sejam acessíveis pela Internet. Entretanto, o IP público é um recurso escasso e é atribuído a VMs que necessitam de acesso externo por diferentes usuários. Além disso, em alguns casos não é desejável que o acesso de rede das VMs esteja disponível publicamente, por questões

de segurança. Quando as VMs necessitam apenas de comunicação entre elas, é aconselhável construir uma rede sobrecamada de nível 2 e realizar a comunicação entre elas a partir de IPs privados. Dessa forma, apesar de estarem em diferentes sítios, as VMs se enxergarão como se estivessem na mesma rede local, não necessitando de IPs públicos para se comunicarem. Essa solução de sobrecamada de nível 2 é fornecida pelo serviço de rede do OpenStack, denominado Neutron. Atualmente o GT-PID utiliza um esquema simples de rede, fornecido pelo Nova, que provê acesso à Internet às VMs através de NAT. Na Fase 2 será investigado o módulo Neutron, fornecendo conectividade em nível 2 às VMs e permitindo novas funcionalidades no GT-PID, como comunicação segura entre VMs e aplicação do conceito de SDN.

A segunda questão em aberto consiste no gerenciamento do endereço IP público da VM quando esta é migrada para outro sítio, visto que o endereçamento IP na Internet identifica a localização da máquina e necessita ser alterado após a migração. Por exemplo, a UFRJ e UERJ possuem faixas de endereços IP diferentes. Ao realizar a migração entre a UFRJ e UERJ o endereço IP da VM necessitaria ser alterado. Entretanto, essa mudança de endereço IP tornaria a VM inacessível a aplicações que tentarem acessar seus serviços pelo endereço IP antigo. Dessa forma, é necessário considerar a relação da mobilidade das VMs e seus IPs públicos no GTPID. Na Fase 1 realizaram-se testes com o protocolo LISP (*Location ID Separation Protocol*), que permite a manutenção do endereço IP da VM após uma migração. Optou-se na Fase 1 em não concentrar esforços na integração do LISP com o OpenStack, visto que o esquema de rede utilizado pelo GT-PID (isto é, o módulo simples de rede do Nova) será possivelmente substituído pelo módulo avançado Neutron, recomendado pela comunidade OpenStack. Dessa forma, a integração do LISP, ou de outra forma de mobilidade de IP público, será analisada na Fase 2.

Por fim, pretende-se estudar a utilização do IPv6 no contexto do GT-PID. Como o espaço de endereçamento do IPv4 está se esgotando e o número de VMs fornecidas pela infraestrutura pode chegar na ordem de centenas ou até milhares, é necessário definir uma forma eficiente de distribuir IPs públicos. Como o espaço de endereçamento do IPv6 é maior, pode ser interessante sua utilização no contexto do GT-PID. O protocolo LISP suporta o IPv6 e pode apresentar uma alternativa para a transição de IPv4 para IPv6, visto que possibilita o encapsulamento de pacotes IPv6 em pacotes IPv4.

2.4. Ferramentas de suporte à operação (para propostas de serviço)

O protótipo desenvolvido na Fase 1 utiliza o módulo Horizon do OpenStack para gerenciar a nuvem, oferecendo uma interface web de gerenciamento. Esse módulo foi modificado para atender os requisitos do GT-PID, como a existência de entidades do

tipo Administrador Local e a passagem de parâmetros para o Escalonador de Sítios. As principais funções de gerenciamento permitidas pelo módulo Horizon no GT-PID são:

- criação de usuários e definição de seus papéis;
- migração entre sítios (global) e dentro de um sítio (local);
- visualização dos recursos disponíveis (p. ex., número de CPUs virtuais, quantidade de memória RAM, etc.);
- listagem das máquinas e sítios que integram o protótipo;
- criação, suspensão e destruição de máquinas virtuais.

Para a implantação do piloto serão adicionadas ferramentas de monitoramento da infraestrutura, como o Ceilometer, que exibirão suas informações na interface do Horizon. O Ceilometer permite a coleta dinâmica de estatísticas, diferente das quantidades estáticas exibidas pelo protótipo da Fase 1. Dentre os tipos de informação exibidos, podem ser citados:

- estatísticas de uso da rede;
- estatísticas de uso de CPU, RAM e disco;
- estatísticas de downloads de imagens de VMs oriundas do Controlador.

Como mencionado na Seção 2.3, a Fase 2 implantará novas funcionalidades de rede através do módulo Neutron. A utilização dessas funcionalidades e o gerenciamento de rede poderão ser realizados pela interface web. Em particular, o Neutron permite, entre outras funções, as seguintes ações de gerenciamento de rede:

- criação de redes virtuais entre as VMs;
- criação de roteadores virtuais;
- atribuição de endereços IP públicos.

Por fim, na Fase 2 serão definidos procedimentos e implementados scripts para adicionar novos sítios à infraestrutura.

3. Cronograma

A Fase 2 do GT-PID está dividida em 10 atividades baseadas nos requisitos apresentados nas Seções 2.3 e 2.4, que serão executadas seguindo o cronograma da Tabela 1. A seguir, essas atividades são brevemente descritas.

- **Atividade 1 – Gerenciamento do projeto:** Esta atividade envolve a realização de reuniões semanais entre os membros da equipe do GT, atualização periódica do sítio do GT, execução dos relatórios de acompanhamento do projeto, elaboração de documentação e busca de parceiros para uma possível fase de serviço experimental.
- **Atividade 2 – Definição de requisitos e casos de uso:** Desde o início da Fase 2 até a apresentação para o comitê de avaliação, serão realizadas reuniões com membros da RNP para definição de requisitos e casos de uso da plataforma. Busca-se nesta atividade uma integração com a área de Serviços da RNP para entender as necessidades reais dos clientes da nuvem. As diferentes atividades

realizadas ao longo do projeto poderão ser alimentadas por essa atividade, dependendo dos requisitos e casos de uso levantados.

- **Atividade 3 – Instalação e configuração da versão de produção:** Nesta atividade serão definidos os procedimentos para instalação dos componentes da infraestrutura, sem a utilização da versão de protótipo DevStack, e aprimorados diversos aspectos de segurança e operacionais. Após a definição e teste dos procedimentos de instalação e configuração, será possível elaborar uma documentação com o passo-a-passo da inserção de um sítio na infraestrutura e instalação do Controlador. Nesta atividade também será analisada a possibilidade de tornar a infraestrutura multiplataforma, permitindo a coexistência de diferentes hipervisores, como o KVM e o Xen. Esta atividade será finalizada com a implantação do piloto em três instituições (UFRJ, UERJ e UFF) e entrega de seu código-fonte, com a respectiva documentação, de acordo com o cronograma de entregas definido no edital.
- **Atividade 4 – Integração com a CAFe:** Nesta atividade o módulo Keystone será integrado com a CAFe através da utilização da CAFe expresso. O desenvolvimento e os testes serão implementados na versão protótipo do GTPID, para posteriormente integrar essa solução no piloto em desenvolvimento na Atividade 3.
- **Atividade 5 – Inclusão de ferramentas de monitoramento:** Nesta atividade serão estudadas ferramentas de monitoramento, em especial, o módulo Ceilometer do OpenStack. A partir do estudo, essas ferramentas serão implantadas no piloto.
- **Atividade 6 – Planejamento e participação no WRNP:** Nesta atividade a demonstração do piloto no WRNP 2015 será planejada, através da listagem dos equipamentos utilizados, da escolha dos membros da equipe que participarão da demonstração e da confecção de recursos visuais. O piloto a ser apresentado no WRNP consiste na infraestrutura instalada na Atividade 3.
- **Atividade 7 – Inclusão de opções avançadas de rede:** Nesta atividade serão analisadas funções avançadas de rede, como a possibilidade do uso do protocolo LISP e do módulo Neutron.
- **Atividade 8 – Definição de procedimentos de atualização:** Nesta atividade serão definidos os procedimentos de atualização do código do GT-PID de acordo com o código atual do OpenStack. Para tal, serão realizados diferentes testes mesclando o código do GT-PID com as versões disponibilizadas pelo OpenStack na época da execução da atividade. Esta atividade também envolve a participação em fóruns do OpenStack e acompanhamento das palestras proferidas no *OpenStack Summit*. Além do alinhamento com a comunidade OpenStack, esta atividade prevê o alinhamento com outros projetos da RNP que utilizam o OpenStack, como o SE-CNC (Computação em Nuvem para Ciência), através de possíveis integrações e padronização da nomenclatura utilizada. No fim desta atividade o código do GT-PID estará atualizado em relação à última versão estável do OpenStack.

- **Atividade 9 – Realização de testes no piloto:** Nesta atividade serão realizados testes de desempenho no piloto, mostrando seus aspectos de escalabilidade e qualidade do serviço. Os resultados culminarão na elaboração do relatório de Avaliação dos Resultados do Piloto (RT6).
- **Atividade 10 – Análise e correções de problemas e vulnerabilidades:** Nesta atividade serão corrigidos possíveis problemas e vulnerabilidades descobertos durante a execução do projeto. Além disso, nesta atividade serão executados refinamentos sugeridos pela RNP e outros membros da comunidade no WRNP, no workshop de disseminação e na apresentação para o comitê de avaliação. As correções realizadas nesta atividade serão integradas ao código final que será entregue em 30/11/2015.

Atividades	2014		2015											
	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Ativ. 1														
Ativ. 2														
Ativ. 3														
Ativ. 4														
Ativ. 5														
Ativ. 6														
Ativ. 7														
Ativ. 8														
Ativ. 9														
Ativ. 10														

Tabela 1 - Cronograma de implantação e testes do piloto.

Atividade 1 – Gerenciamento do projeto

Atividade 2 – Definição de requisitos e casos de uso

Atividade 3 – Instalação e configuração da versão de produção

Atividade 4 – Integração com a CAFe

Atividade 5 – Inclusão de ferramentas de monitoramento

Atividade 6 – Planejamento e participação no WRNP

Atividade 7 – Inclusão de opções avançadas de rede

Atividade 8 – Definição de procedimentos de atualização

Atividade 9 – Realização de testes no piloto

Atividade 10 – Análise e correções de problemas e vulnerabilidades

4. Referências

- [1] KVM - *Kernel-based Virtual Machine*, <http://www.linux-kvm.org/> - Acessado em Agosto de 2014.

[2] Xen – The Xen Project, <http://www.xenproject.org/> - Acessado em Agosto de 2014. [3] *OpenStack - Cloud Software*, <http://www.openstack.org> - Acessado em Agosto de 2014.

[4] *Apache CloudStack - Open Source Cloud Computing*, <http://cloudstack.apache.org/> - Acessado em Agosto de 2014.

[5] *Eucalyptus - Open Source AWS Compatible Private Cloud*, <http://www.eucalyptus.com/> - Acessado em Agosto de 2014.