

DisTInTo: Uma Ferramenta para Apoiar Experimentos Distribuídos em Larga Escala

José Alexandre Macedo, Guilherme Tebaldi Meira, Rafael Fávoro, Guilherme Artem, Roberta Lima Gomes, Magnos Martinello
Centro Tecnológico - UFES
Vitória, Espírito Santo, Brazil
{jamacedo, gtmeira, rfavaro, gasantos, rgomes, magnos}@inf.ufes.br

ABSTRACT

The scientific method dictates that experiments must be reproduced before they are considered valid. In computer science, reproduction requires open access to all code, scripts, and data used to produce the results. In particular, the validation of large scale distributed systems is a huge challenge to be performed in a real Internet-based testbed. In this work, we propose DiSTInTo as a tool designed to allow users to control their experiments at large-scale distributed testbeds. The tool supports the reproducibility of scientific experiments, since the setup instrumentation of machines until the crawling of the data. It provides features such as file transfer, experiments execution, real time monitoring and visualization of distributed machines. The time for experiments deployment, execution and crawling has been reduced in our P2P large scale networks tests.

Keywords

reprodutibilidade, *testbed*, sistemas distribuídos, P2P

1. INTRODUÇÃO

Sistemas distribuídos buscam compartilhar a execução de tarefas através de redes de computadores [11]. Hoje eles fazem parte da Internet e são usados para armazenar e distribuir arquivos, transmitir vídeo, realizar processamento distribuído, entre muitas outras atividades. A criação desses sistemas apresenta uma etapa fundamental no processo de desenvolvimento de qualquer *software*, a realização de testes para validação. Tais testes dependem do controle de máquinas distribuídas que possam ser utilizadas para verificar o funcionamento do sistema. Em alguns casos, dezenas de milhares de máquinas são necessárias para tornar os testes realísticos. Isso acontece, por exemplo, com aplicações *peer-to-peer* (P2P). Conforme o número de máquinas para validação de um sistema aumenta, a complexidade de manipulá-las a fim de emular comportamentos próximos dos reais também aumenta. Essa demanda se torna ainda mais crítica quando são considerados sistemas multimídia distribuídos, em que

aspectos de tempo real e interatividade também devem ser controlados.

No meio acadêmico um princípio fundamental por trás da metodologia científica é dado pela reprodutibilidade dos experimentos [4][7]. No caso de pesquisas com sistemas distribuídos, o desafio está na dificuldade de reprodução dos experimentos em uma escala global de Internet. Para isso *testbeds*, tais como o PlanetLab [1], permitem usar máquinas espalhadas pelo planeta a fim de avaliar sistemas expondo-os a todas as adversidades da Internet. Esses *testbeds* precisam ser controlados para que os experimentos possam ser configurados e posteriormente reproduzidos por outros pesquisadores.

A tarefa de manipular as máquinas distribuídas não é simples porque a comunicação pode ser prejudicada devido a falhas na rede e congestionamentos. A própria natureza dos protocolos de rede gera atrasos (latência) entre a requisição e a entrega dos dados. Nem mesmo os atrasos são constantes devido às suas variações sobre pacotes recebidos (*jitter*). Todas essas características da rede impactam tanto na comunicação entre as máquinas (durante os testes), quanto no controle delas.

A ferramenta *Distributed Testbed Instrumentation Tool* (DisTInTo) é apresentada neste contexto a fim de se tornar uma facilitadora para a manipulação de testes e experimentos relacionados a validação de sistemas distribuídos, reduzindo a complexidade envolvida nesta tarefa. O apoio envolve: a comunicação simultânea com todas as máquinas; a transferência de arquivos e *logs* por meio de armazenamento em nuvem; a realização do experimento em si; e o agrupamento de máquinas com um estado específico (ex. um erro no experimento).

O presente artigo detalha a DisTInTo da seguinte forma: Na Seção 2 são apresentados os objetivos da ferramenta. A arquitetura do sistema é descrita na Seção 3. Na Seção 4 as principais funcionalidades são mostradas. Os trabalhos relacionados são discutidos na Seção 5. Por fim, na Seção 7 as conclusões são apresentadas.

2. OBJETIVO DA FERRAMENTA

A ferramenta DisTInTo tem o objetivo de apoiar todo o processo de experimentação e testes em ambiente distribuído desde a comunicação e *setup* das máquinas até a obtenção de *logs* para análise. Essa não é uma tarefa simples porque

experimentos em ambiente distribuído podem envolver inúmeras máquinas com capacidade, comunicação e localização diferentes.

O primeiro passo suportado nesse processo é a comunicação com múltiplas máquinas para a manipulação de sistemas por meio do envio de comandos. Além da comunicação, outro objetivo é tornar trivial a troca de dados entre a máquina do profissional que está orquestrando/coordenando os experimentos e as máquinas distribuídas. Isso é feito por meio de um serviço de armazenamento em nuvem que fica responsável pela transferência de arquivos. A comunicação simultânea e a troca de dados permitem realizar o *setup* junto às *software* que serão utilizados nos experimentos. Na etapa de execução de experimentos a ferramenta tem o objetivo de inicializar e parar as máquinas de acordo com *scripts* predefinidos. Adicionalmente ela permite repetir diversas vezes uma mesma configuração de um experimento. As máquinas podem ser agrupadas a qualquer momento a fim de localizar em quais delas faltam *software* para serem instalados ou onde estão ocorrendo determinados erros. O monitoramento dos experimentos também apresenta-se como um dos objetivos da ferramenta DisTInTo, permitindo o acompanhamento de pessoas a distância.

3. ARQUITETURA

A arquitetura da ferramenta DisTInTo está representada na Figura 1. Ela é composta dos módulos: Acesso SSH, Integração *Cloud*, Experimentos, Visualização das Máquinas e Monitoramento. Estes módulos correspondem a implementação base das principais funcionalidades da DisTInTo.



Figura 1: Estrutura da Ferramenta DisTInTo

O módulo **Acesso SSH** é o responsável pelo acesso remoto e simultâneo a múltiplas máquinas. Outros módulos dependem dele como o *Integração Cloud* e *Experimentos*. Sua implementação usa o *framework* Java Secure Channel (JSCh)¹ para criar um canal seguro via Secure Shell (SSH) com cada máquina.

O módulo **Visualização das Máquinas** utiliza a API Nasa World Wind² para criar uma visualização global das máquinas do experimento. Na versão atual, existe uma extensão desse módulo para visualização de máquinas do PlanetLab. A extensão foi desenvolvida com base na API do PlanetLab³.

¹<http://www.jcraft.com/jsch/> [07/2013]

²<http://worldwind.arc.nasa.gov/java/> [07/2013]

³http://www.planet-lab.org/doc/plc_api [07/2013]

O módulo de **Monitoramento** realiza o envio de mensagens para um serviço de *real-time* chamado Pusher que, por sua vez, as envia para os clientes conectados via WebSocket. O Pusher é baseado no modelo Publish/Subscribe [3], isso permite criar eventos e tratá-los de forma diferenciada na página web de acompanhamento dos experimentos. Essa página possui os códigos em JavaScript para o cliente se conectar ao Pusher e aguardar suas mensagens. Para o funcionamento ela deve ser hospedada em um servidor web (por exemplo, na pasta pública do Dropbox). A comunicação do módulo de Monitoramento com o serviço do Pusher é feita por meio da API Java Pusher⁴.

Os módulos *Integração Cloud*, *Experimentos* e *Visualização das Máquinas* possuem extensões específicas e *wrappers* que permitem novas extensões futuras. Isso garante, por exemplo, que hoje seja utilizado o *Suporte ao Dropbox* para troca de dados, mas não impede que outra solução de *cloud computing* seja adotada no futuro. O mesmo é válido para o módulo *Experimentos*, o qual suporta atualmente experimentos com redes P2P e o módulo *Visualização*, que permite visualizar máquinas do PlanetLab.

As máquinas utilizadas nos experimentos precisam estar descritas no arquivo *workstations.txt* localizado no mesmo diretório da ferramenta DisTInTo. O arquivo pode ter vários formatos que definem usuário, endereço IP e o tipo de autenticação. Esses formatos estão disponíveis no site⁵ da ferramenta. O *Suporte ao Dropbox* depende de informações de acesso que devem ser colocadas no arquivo *config.properties*.

4. PRINCIPAIS FUNCIONALIDADES

A interface principal para acesso às funcionalidades da ferramenta DisTInTo é apresentada na Figura 2. Nessa interface o profissional pode controlar experimentos distribuídos e enviar comandos para as máquinas selecionadas. Na barra superior são mostrados o número total de máquinas disponíveis e o número de máquinas selecionadas. Do lado superior direito existem as funções de conexão e desconexão com as máquinas selecionadas no painel central e também a opção de selecioná-las (apenas as conectadas, apenas as desconectadas ou todas). As funcionalidades presentes na barra inferior são descritas a seguir.

4.1 Acesso a Múltiplas Máquinas

A funcionalidade de acesso a múltiplos computadores permite a comunicação por meio de um canal seguro criado com o protocolo Secure Shell (SSH). A ferramenta possibilita a autenticação em cada máquina tanto via senha quanto por chave privada. Computadores em redes internas também podem ser utilizados em experimentos pela DisTInTo. O sub acesso cria um encaminhamento fazendo uso de uma máquina intermediária com IP público para alcançar uma rede interna.

As ferramentas dedicadas ao acesso múltiplo não são novidade, algumas conhecidas são: *pssh*⁶, *DSH*⁷, *sshsudo*⁸, *ten-*

⁴<https://github.com/capotej/java-libpusher> [07/2013]

⁵ <http://distinto.lprm.inf.ufes.br> [07/2013]

⁶<http://code.google.com/p/parallel-ssh/> [07/2013]

⁷<http://bit.ly/15l0UJT> [07/2013]

⁸<http://code.google.com/p/sshsudo/> [07/2013]

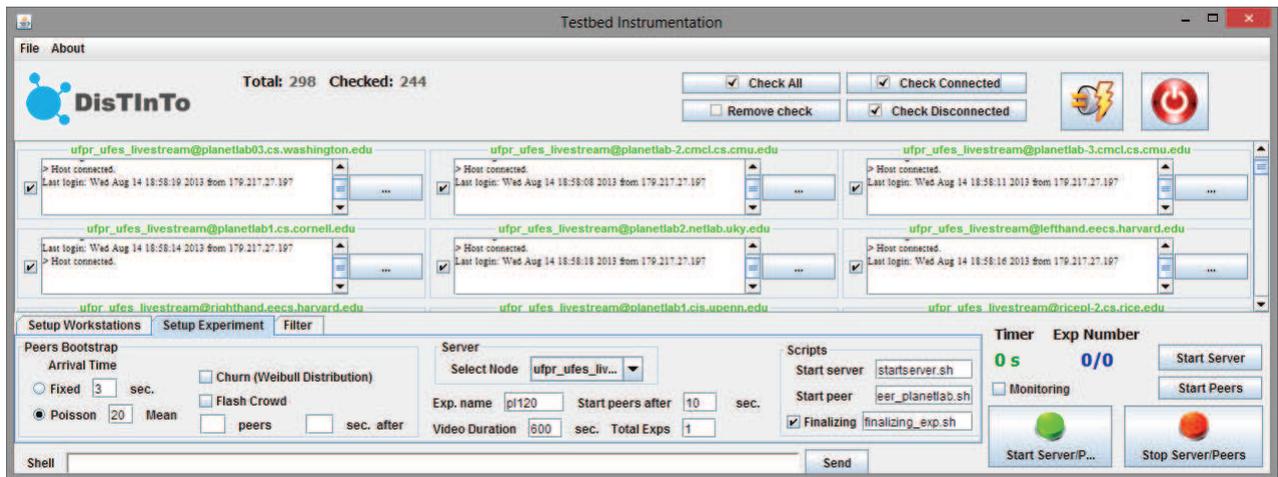


Figura 2: Interface Principal da Ferramenta

takel⁹, Cluster SSH¹⁰ e vxargs¹¹. No entanto, esse recurso foi implementado na DisTInTo por ser uma funcionalidade base para outras mais avançadas discutidas a seguir.

4.2 Integração com Armazenamento em Nuvem

Uma tarefa frequente na realização de experimentos distribuídos é a transferência de arquivos. Os sistemas sob avaliação devem ser transferidos para as máquinas do experimento e os *logs* de resultados precisam percorrer o caminho contrário para a análise. DisTInTo resolve esse problema oferecendo integração com serviços de armazenamento de arquivos em nuvem. A primeira integração disponibilizada utiliza o Dropbox, um serviço gratuito, difundido e com uma API para acesso programático ao sistema.

A DisTInTo facilita a instalação do *daemon* do Dropbox em todas as máquinas. Ela também automatiza o processo de autenticação das máquinas em uma conta criada para os experimentos, evitando a necessidade de abrir um web *browser* para autenticar cada máquina. Feito isso, uma pasta precisa ser criada e compartilhada entre a conta do experimento e a conta do instrumentador (pessoa que controla e realiza os testes). Ao final do experimento, os *logs* podem ser copiados para o Dropbox, que se encarrega de sincronizá-los com o instrumentador. A Figura 3 mostra as abas com as funcionalidades, sendo o Item 1 relativo à Integração com o Dropbox. Além da instalação, a ferramenta apoia a remoção, inicialização, parada e verificação do status de sincronização.

4.3 Observação e Alerta de Acontecimentos

Experimentos com máquinas distribuídas estão suscetíveis a eventos inesperados como erros e quedas de conexão. Acompanhar manualmente o funcionamento dos sistemas em todas as máquinas se torna mais difícil à medida que a quantidade de máquinas aumenta. Para auxiliar nesse acompanhamento, DisTInTo fornece a funcionalidade de filtros.

⁹<http://bit.ly/14fXaW6> [07/2013]

¹⁰<http://sourceforge.net/projects/clusterssh/> [07/2013]

¹¹<http://vxargs.sourceforge.net/> [07/2013]

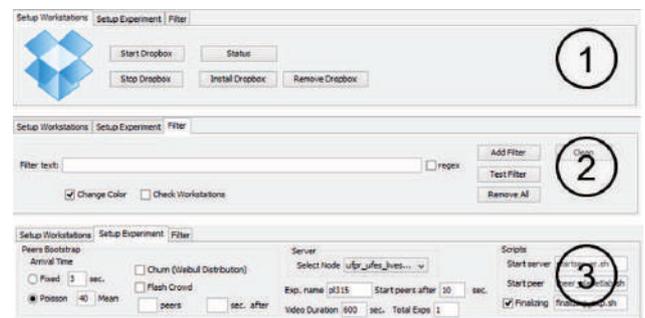


Figura 3: Funcionalidades da Ferramenta

Um filtro monitora as mensagens de saída dos terminais das máquinas em busca de uma determinada *string* ou expressão regular inseridas pelo instrumentador. Como exemplo, é possível criar um filtro para encontrar todas as máquinas que apresentam um erro do tipo *NullPointerException*. Quando a mensagem é encontrada duas ações podem ocorrer: (i) a máquina é destacada com cor diferente, indicando que algo aconteceu e merece atenção; e (ii) a máquina é selecionada e agrupada a todas com o mesmo filtro para que, posteriormente, sejam enviados comandos para elas. O Item 2 da Figura 3 apresenta a aba responsável pelo gerenciamento dos filtros.

4.4 Apoio a Experimentos com Redes P2P

A ferramenta tem como um dos objetivos apoiar a realização do experimento em si. Dessa forma, a primeira classe de experimentos suportados são os de redes P2P. As aplicações P2P comerciais possuem características de uso mapeadas em pesquisas [8][10][2]. Os experimentos com essas redes precisam simular diferentes características para obter resultados realísticos.

O Item 3 da Figura 3 mostra a aba com as funcionalidades para interação com as máquinas em experimentos de redes P2P. A chegada dos *peers* (inicialização do sistema nas máquinas) pode ser feita com um intervalo fixo entre as chegadas ou seguindo uma distribuição de Poisson. As redes reais

são dinâmicas e os clientes entram e saem com frequência, o que é chamado de *churn*. A ferramenta possibilita simular esse comportamento em máquinas reais com uma frequência baseada na distribuição de Weibull. Ela permite também a inicialização simultânea (*flash crowd*) de um grande número de *peers* na rede P2P. Um experimento pode ser programado para repetir várias vezes a fim de obter resultados confiáveis. Os comandos para inicializar e finalizar o servidor e os *peers* são definidos em *scripts* com comandos para execução em Linux.

4.5 Visualização Geográfica das Máquinas

Os experimentos distribuídos podem ser afetado pela localização das máquinas. Em *testbeds* como o PlanetLab elas estão espalhadas por todo o mundo. Conhecer as localizações permite ao instrumentador compreender o cenário com o qual se está trabalhando e obter informações que auxiliam na interpretação dos resultados. A ferramenta possui essa funcionalidade implementada para visualização das máquinas de usuários do PlanetLab. Além do planisfério 2D, DisTInTo permite a visualização por meio de um globo tridimensional e a utilização do recurso de zoom para análise de máquinas aglomeradas em locais próximos.

4.6 Monitoramento Web

Os experimentos distribuídos podem envolver outros participantes além do instrumentador. O acompanhamento desses participantes permite a tomada de decisões em tempo de execução. Como exemplo de decisão está a espera para entrada nos experimentos. DisTInTo facilita o monitoramento por meio de uma página web usada para recepção em tempo real de informações sobre experimentos.

5. TRABALHOS RELACIONADOS

Ferramentas de suporte ao processo de experimentação distribuído são pouco encontradas na literatura. Os trabalhos geralmente têm seu foco na qualidade e no monitoramento do ambiente [6][9]. Em contraste com os trabalhos anteriores, a ferramenta DisTInTo auxilia na instrumentação, na implantação e monitoramento dos experimentos de sistemas distribuídos em larga escala.

O trabalho mais semelhante ao DisTInTo é a plataforma *Testbed for Future Internet Services* (TEFIS)[5]. O TEFIS tem o objetivo de disponibilizar um único ponto de acesso para múltiplos *testbeds* (intermediando serviços como o PlanetLab). Assim como a DisTInTo, o TEFIS busca apoiar o processo de experimentação como um todo e para isso ele permite o planejamento do *workflow* do experimento. Com relação a manipulação dos resultados, o TEFIS armazena em banco de dados os *logs* dos experimentos, enquanto a DisTInTo faz uso dos serviços de armazenamento em nuvem. Outro ponto comum está em facilitar o processo de experimentação em larga escala. Porém, enquanto a TEFIS possui uma abordagem voltada para auxiliar todo ciclo de desenvolvimento de novos serviços, não apenas experimentação e testes, a DisTInTo provê mais flexibilidade ao permitir a utilização de qualquer máquina (conectável via ssh) nos experimentos, independentemente dos *testbeds* como é o caso do TEFIS.

6. AGRADECIMENTOS

Agradecemos à FAPES e FINEP (Projeto CPID-CIDIG) e à CAPES (Projeto RH-TVD #225/2008) pelo financiamento parcial deste projeto.

7. CONCLUSÃO

O artigo apresentou a ferramenta DisTInTo com o objetivo de apoiar o processo de experimentação de sistemas distribuídos. Os primeiros testes práticos, realizados com um sistema de distribuição de vídeo P2P, contemplando cerca de 300 máquinas mostram que a instrumentação se torna mais rápida e em poucos minutos as máquinas podem estar configuradas para a realização e repetição de experimentos. DisTInTo encontra-se disponível³ para download sem limites de uso.

As melhorias futuras pretendem estender a ferramenta DisTInTo de forma a apoiar o processo de experimentação como um todo. Isso significa incluir todas as etapas envolvidas até a obtenção dos gráficos com resultados finais. Pretende-se ainda criar novas integrações com outros tipos de serviço de armazenamento em nuvem para transferência de dados e novas facilidades para experimentos com outros sistemas distribuídos.

8. REFERÊNCIAS

- [1] B. Chun, D. Culler, T. Roscoe, A. Bavier, L. Peterson, M. Wawrzoniak, and M. Bowman. Planetlab: an overlay testbed for broad-coverage services. *ACM SIGCOMM Computer Communication Review*, 33(3):3–12, July 2003.
- [2] C. Dana and D. Li. Bass: Bittorrent assisted streaming system for video-on-demand. In *International Workshop on Multimedia Signal Processing*, pages 1–4, 2005.
- [3] P. T. Eugster, P. A. Felber, R. Guerraoui, and A.-M. Kermarrec. The many faces of publish/subscribe. *ACM Computing Surveys*, 35(2):114–131, 2003.
- [4] Y. Gil, E. Deelman, M. Ellisman, T. Fahringer, G. Fox, D. Gannon, C. Goble, M. Livny, L. Moreau, and J. Myers. Examining the challenges of scientific workflows. *Computer*, 40(12):24–32, 2007.
- [5] A. Gracia, F. Perrina, B. Pickering, F. Schreiner, et al. Deliverable title d3. 1.1 teagle assessment and tefis portal specifications. 2010.
- [6] K. N. Ramachandran, K. C. Almeroth, and E. M. Belding-Royer. A framework for the management of large-scale. wireless network testbeds. In *1st workshop on Wireless Network Measurements*, 2005.
- [7] W. Schütz. *The testability of distributed real-time systems*, volume 245. Springer, 1993.
- [8] T. Silverston and O. Fourmaux. Measuring p2p iptv systems. In *NOSSDAV'07*, 2007.
- [9] J. Stéger, S. Laki, and P. Mátray. A monitoring framework for federated virtualized infrastructures. *Measurement-based experimental research: methodology, experiments and tools*, 7586, 2012.
- [10] D. Stutzbach and R. Rejaie. Understanding churn in peer-to-peer networks. In *6th ACM SIGCOMM conference on Internet measurement*, IMC'06, pages 189–202, 2006.
- [11] A. S. Tanenbaum and M. Van Steen. *Distributed systems*, volume 2. Prentice Hall, 2002.