

Sumário

1	Arquitetura do middleware	p. 3
1.1	Componentes do <i>middleware</i>	p. 3
1.2	O componente <i>environment</i>	p. 5
1.2.1	Estrutura interna	p. 5
1.2.2	<i>Proxy</i>	p. 7
1.2.3	Tabela de endereços de processos	p. 7
1.3	O componente <i>process</i>	p. 9
1.3.1	Ciclo de vida de um processo	p. 11
1.3.2	Serialização de um processo	p. 13
1.4	Migração de processos	p. 14
1.4.1	Atualização da tabela de endereços dos processos	p. 15
1.5	Troca de mensagens	p. 16
1.5.1	Comunicação direta e indireta	p. 16
1.5.2	Comunicação local direta	p. 17
1.6	Serialização em grande escala	p. 19
1.7	Comunicação grupal	p. 19
1.8	Migrações em grande escala e migrações de ambientes	p. 19

1 Arquitetura do middleware

Este capítulo descreve em detalhes a arquitetura do proposto *middleware* de comunicação, tais como suas partes e suas principais funções, como troca de mensagens, migração e serialização.

1.1 Componentes do middleware

A arquitetura proposta é dividida em dois componentes principais: o ambiente (*environment*) e o processo lógico (*process*). Um ambiente representa um nó lógico do sistema de simulação distribuída, ou seja, é a representação lógica de um computador no sistema distribuído. Em uma simulação distribuída, tipicamente, cada nó físico da rede deve conter um único *environment* (excessões à esse caso serão tratados a diante).

O segundo componente da estrutura do *middleware*, o processo lógico, que é a representação lógica de um processo no sistema de simulação distribuída. Na arquitetura aqui descrita, um processo lógico somente existe dentro de um *environment*. Sendo assim, um *environment* pode ser visto como um conjunto de processos. Por sua vez, um processo somente pode estar contido por um único *environment* em um determinado momento.

Assim, sendo p um processo lógico e e um ambiente de simulação contendo n processos, então definimos:

Definição 1: Um *environment* é um conjunto de processos

$$e_k = \{p_i | i < n\} \quad (1.1)$$

Definição 2: Um processo só pode estar contido em um único ambiente em um instante bem definido

$$p_k \in e_l \rightarrow \nexists e_m, p_k \in e_m, e_l \neq e_m \quad (1.2)$$

Tanto o *environment* quanto o *process* são abstrações lógicas que representam a simulação baseado em eventos discretos. Fisicamente tanto o ambiente quanto o processo lógico são instâncias de objetos que devem ser implementadas extendendo classes bases abstratas que contém as especificações descritas por este *middleware*.

A arquitetura do *middleware* aqui proposto deve oferecer as seguintes funcionalidades:

- Comunicação entre processos por troca de mensagens.
- Serialização do conteúdo de um processo lógico.
- Migração de um processo de um *environment* para outro.
- Serialização em larga escala de um ambiente ou de toda a simulação.
- Comunicação grupal entre processos e entre ambientes.

A propriedade de comunicação por troca de mensagem é um item fundamental para a implementação de simulação distribuída. A forma de se comunicar por troca de mensagens provida pelo *middleware* baseia-se nas quatro suposições iniciais descritas por (MCQUILLAN; WALDEN, 1975) sobre os canais de comunicação inter-processos:

- O canal introduz um flutuante, porém finito, atraso nas mensagens.

- O canal possui uma flutuante, porém finita, largura de banda.
- O canal apresenta uma flutuante, porém finita, taxa de erro.
- Existe uma real possibilidade de as mensagens transmitidas da fonte para o destino cheguem ao destino em uma ordem diferente da originalmente transmitida. É assumido que tanto a fonte quanto o destino possuem, em geral, finitos tamanhos de *buffers* de armazenamento e diferentes *bandwidth*.

A capacidade de um processo lógico migrar de um *environment* para outro é um ponto fundamental para se proporcionar a capacidade de balanceamento de cargas em uma simulação distribuída. O mecanismo de migração, descrito na Seção 1.4, é a união da capacidade de serialização de um processo e da comunicação entre *environments*.

1.2 O componente environment

Essencialmente, um ambiente na arquitetura aqui proposta é uma plataforma que abriga e gerencia diversos processos lógicos. A existência desta plataforma como base para o gerenciamento dos processos é justificada quando desejamos manipular simultaneamente características de diversos processos que possuem em comum o fato de estarem no mesmo ambiente físico (como por exemplo migrar ou serializar todos os processos). Mas principalmente se justifica a existência de uma camada que seja responsável por gerenciar o ciclo de vida de processos, como criação, migração e destruição de processos.

1.2.1 Estrutura interna

Internamente, o *environment* apresenta as seguintes estruturas básicas (conforme ilustrada na Figura 1.1):

- Estrutura interna de dados do ambiente

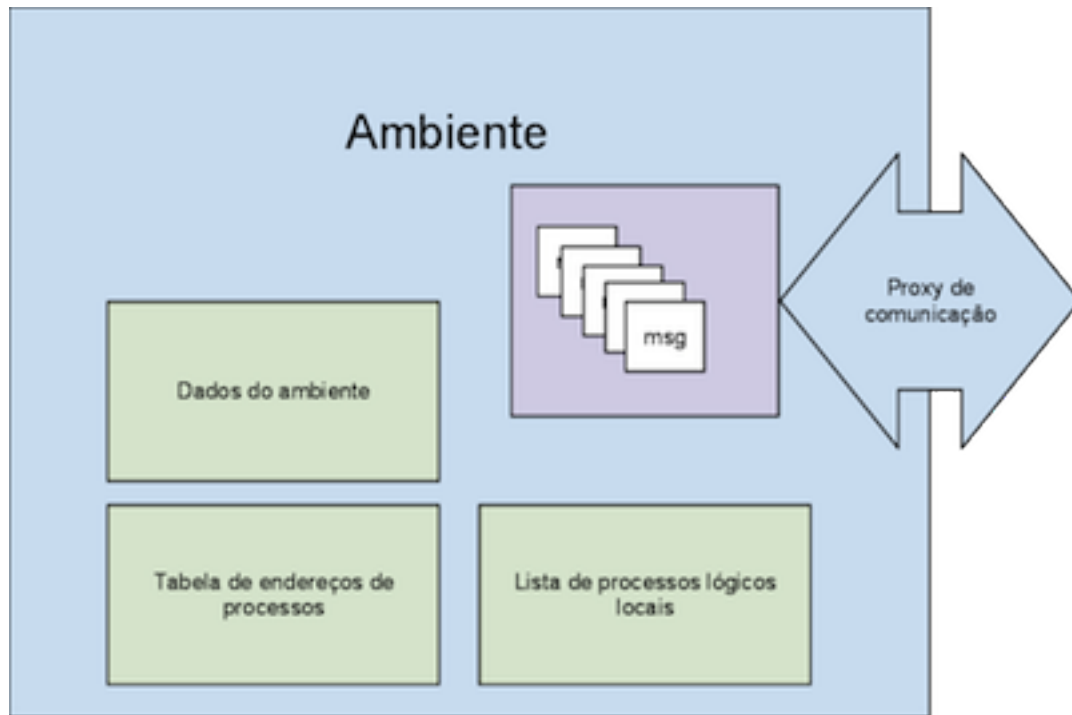


Figura 1.1: A arquitetura interna de um *environment*.

- Tabela de endereços de processos
- Lista de processos lógicos locais
- Proxy de comunicação externa

A estrutura interna de dados do ambiente é a representação de todas as variáveis pertencentes ao *environment*. Dados como o endereço físico (IP) do ambiente na rede, nome lógico do ambiente e quantidade de processos residentes no *environment* são armazenadas neste espaço.

Um *environment* possui um endereço lógico único em toda o ciclo de vida da simulação, denominado *Unique Environment Identifier* - *UEI*. Este endereço é um número natural que o representa no sistema de simulação.

Ao contrário do que será descrito na Seção 1.3 para o componente *process*, a comunicação com um *environment* se dá somente via seu endereço físico na rede.

Tal inflexibilidade é justificada quando assumimos que um *environment* tem todo o seu ciclo de vida atrelado a uma mesma máquina física (salvo exceções descritas na Seção 1.8)

1.2.2 *Proxy*

O *proxy* é a camada do *environment* que é responsável por toda troca de mensagem entre os processos. O *proxy* atua de maneira distinta em troca de mensagens entre processos que convivem no mesmo ambiente e entre trocas de mensagens entre processos que se situam em ambientes distintos. As distinções entre trocas de mensagens internas e externar serão tratadas na seção 1.5

O *proxy* também é o único canal por onde os processos recebem mensagens providas de fora do ambiente. Toda mensagem recebida pelo proxy é identificada pelo número lógico do processo destinatário. Cabe ao proxy converter o nome lógico do processo em seu endereço físico e encaminhar a mensagem ao processo de destino.

1.2.3 Tabela de endereços de processos

A tabela de endereços dos processos é uma lista associativa que possui informações básicas de endereços e status dos processos existentes. A tabela de processos lógicos contém informações não apenas dos processos residentes no *environment* em questão, mas sim dados de todos os processos existentes na simulação (mesmo que desatualizados, dependendo do instante). Dados referentes aos processos residentes no mesmo *environment* da tabela de endereço de processos estarão, invariavelmente, atualizados.

Conforme previamente mencionado, para realizar a comunicação entre processos é utilizado como identificador do processo não o seu endereço físico, mas sim seu endereço lógico. O motivo de se utilizar o endereço lógico é que este, ao contrário do endereço físico, este único em todo o ciclo de vida do processo. Isto

garante que após uma migração de um *environment* para outro, um processo possa continuar a ser encontrado pelo mesmo endereço lógico (o que não aconteceria caso o endereço físico fosse usado, uma vez que após uma migração o endereço de IP do *environment*, e a porta a qual o processo está vinculado, sofreriam variações). Esta característica de ser identificado sempre pelo mesmo endereço lógico é fundamental para evitar a constante alteração de dados referentes aos processos em inúmeras partes do sistema.

Além de armazenar dados referente aos endereços lógicos e físicos dos processos, a tabela de endereços armazena também uma variável de *status* de cada processo. Os estados de um processo podem ser:

- Ativo: Indica que o processo se encontra neste *environment* e está ativo.
- Ausente: Indica que o processo em questão se encontra em outro ambiente.
- Trânsito: Indica que o processo em questão estava em um momento anterior neste ambiente e foi migrado para um ambiente diferente, porém ainda não atualizou a tabela de endereços com seu endereço atual.
- Inativo: Indica que o processo se encontra neste ambiente, mas não está ativo.

Os estados ativo e inativo são os estados mais comuns de um processo em um sistema típico. Eles indicam que o processo em questão está, ou não, em execução naquele ambiente. Um processo em estado inativo significa que este está residente no *environment* em questão, mas não está executando no momento por algum motivo não identificado. Toda mensagem recebida para ser entregue a um processo inativo será armazenada no buffer de mensagens do *proxy* e deverá ser retirada posteriormente pelo processo.

O estado de trânsito indica que o processo esteve naquele *environment* em algum instante do passado e que sofreu uma migração, mas seu novo endereço

não foi ainda atualizado. Mais informações de como funciona a atualização de endereços durante a migração pode ser encontrado na Seção 1.4.1.

1.3 O componente process

Um processo é a menor unidade de processamento em um sistema de simulação de eventos discretos. É ele o responsável por retirar cada evento a ser executado da fila de eventos futuros e executá-lo. A representação de um processo lógico na arquitetura aqui descrita se dá pelo componente *process*. Este deve ser capaz de proporcionar mecanismos para a implementação da simulação propriamente dita (uma área onde o usuário do *middleware* implementará a sua simulação), mecanismos de troca de mensagens, serialização e locomoção (migração) de processos..

Um processo lógico possui três identificações distintas no sistema: seu endereço físico em memória, seu endereço físico na rede e seu endereço lógico no sistema de simulação. Na arquitetura do *middleware* aqui descrito, o processo é referenciado para troca de mensagens ou por seu endereço físico na rede, ou por seu endereço lógico no sistema de simulação. A distinção se a comunicação é feita através do endereço físico ou lógico se dá a partir de qual nível da arquitetura se dá a comunicação.

Como uma das premissas do *middleware* de comunicação é tratar de maneira completamente transparente para o usuário a comunicação entre os processos, o mecanismo de troca de mensagem entre dois objetos do tipo *Process* deve ser encapsulado e resolvido pelo próprio objeto, não transferindo assim ao usuário ...

Desta forma, internamente um processo lógico deve ser capaz de se comunicar com o processo alvo da maneira mais eficiente possível. Assim sendo, é natural que processos que residam em um mesmo *environment* possuam um canal diferenciado de comunicação, ao ponto que processo que estão em ambientes distintos compartilhem canais comuns de comunicação.

Em contrapartida, é justificável que processos que não se comunicam com grande frequência (ou que nunca se comunicam) não tenham conhecimentos uns dos outros. A abstração provida pelo *environment* garante o isolamento entre o endereço físico de um processo e o seu nome lógico no sistema. Este desacoplamento entre endereço lógico e físico de um processo garante também a continuidade da comunicação entre processos lógicos após sua migração. Assim sendo, uma vez que um determinado processo tenha que enviar uma mensagem para um segundo processo, cabe ao *environment* identificar o destinatário da mensagem e a despachar.

Toda comunicação que é iniciada a partir de um processo, ou seja, é uma comunicação que foi iniciada pelo usuário do *middleware*, é feita através do endereço lógico do processo. Isto garante que o usuário do *middleware* não precise se adentrar em estruturas físicas do sistema, como saber qual o endereço físico (IP:PORT) está o processo com o qual ele deseja se comunicar. Cabe ao *environment* que abriga o remetente da mensagem encaminhá-la para o endereço físico correspondente.

Por consequência, quando uma comunicação não partiu diretamente do usuário do *middleware* (como é o caso descrito anteriormente. Após a mensagem chegar no *environment* este agora vai encaminhá-la ao destinatário. Este encaminhamento é feito automaticamente pelo sistema, não necessitando de demais intervenções do usuário do *middleware*.), neste caso o endereço usado para se comunicar com o processo é o seu endereço físico na rede.

O endereço lógico de um processo é denominado *Unique Process Identifier - UPI*. Para garantir que cada processo tenha um endereço lógico único no sistema sem a necessidade de sincronização de todo o sistema em cada criação de um novo processo lógico, a construção deste endereço lógico se dá por uma tupla, onde o primeiro elemento é o *UEI* do *environment* onde este processo foi criado, e o segundo elemento da tupla é um número sequencialmente incrementado pelo ambiente a cada novo processo criado. Isto garante que cada processo possua um par de números único em todo o sistema compondo a sua tupla de identificação.

Vale ressaltar que o *UPI* é único em todo o ciclo de vida de um *process* e, portanto, não se altera em caso de migração. Mesmo após migrar de um ambiente para outro, o primeiro elemento da tupla de identificação continua representando o ambiente onde o processo foi criado pela primeira vez.

1.3.1 Ciclo de vida de um processo

Um processo lógico tal como descrito pelo *middleware* possui um ciclo de vida com fases bem definidas. Isso significa a classe base a qual o usuário do *middleware* estende para criar seu processo lógico já prevê, na forma de métodos abstratos, *slots* onde serão implementados trechos importantes para fases específicas da vida do processo. Essas fases estão ilustradas no diagrama da Figura 1.2. A representação dos *slots* onde serão inseridos os códigos, na forma de métodos abstratos é ilustrada na representação UML da Figura 1.3.

O primeiro método invocado automaticamente pelo processo na sua criação é o método `on_create`. Este método é chamado apenas uma única vez em todo o ciclo de vida do processo lógico. Imediatamente após executar o código contido no método `on_create`, o próximo método invocado automaticamente pelo sistema é o método `run`.

O conteúdo do método `run` é onde o corpo da simulação deve estar implementado. É neste ponto do código onde a simulação deverá retirar o próximo evento da lista de eventos futuros e executá-lo. Vale ressaltar que como descreve (ALVES; WALBON; TAKAHASHI, 2009), o processo de se retirar eventos da fila de eventos futuros para execução não pode se dar na forma de um laço de repetição, pois tornaria a execução do método `run` não-preemptiva. A solução apontada por (ALVES; WALBON; TAKAHASHI, 2009) é a de, após realizar a execução de um item da fila de eventos futuros, enviar uma mensagem para si mesmo agendando a execução do evento seguinte. Isto garante a preempção da simulação para tratamento de mensagens externas, migrações, etc.

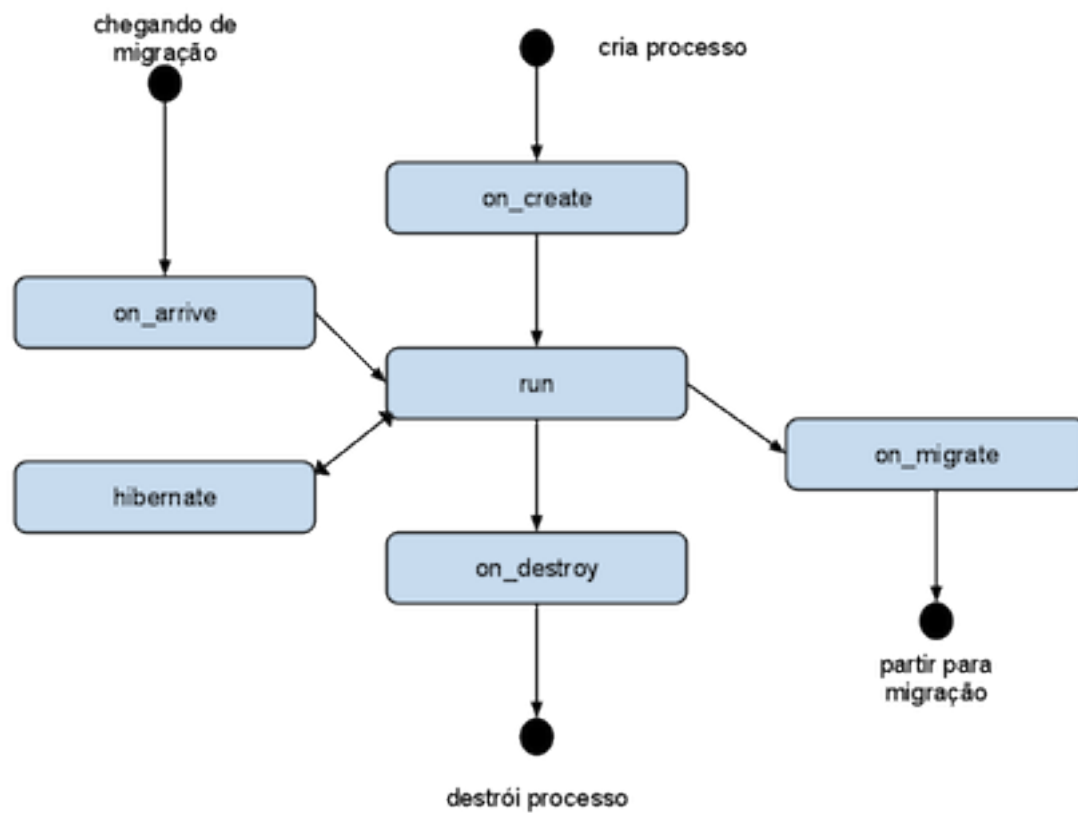


Figura 1.2: Ciclo de vida de um processo lógico.

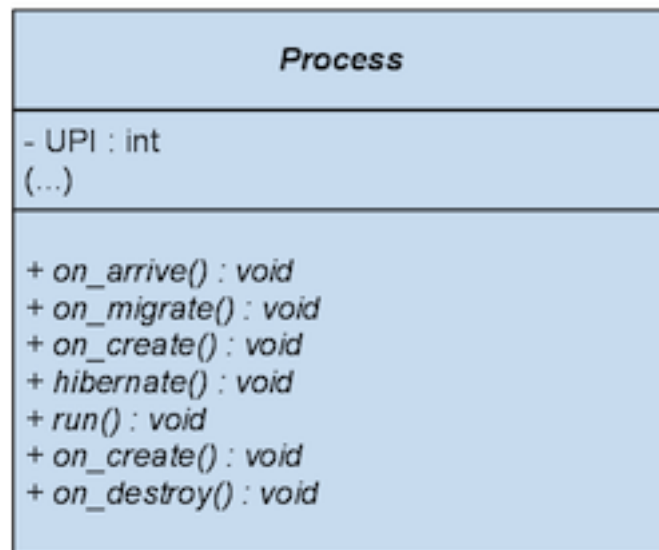


Figura 1.3: Representação em diagrama de classe do componente *process*.

Os demais métodos invocados automaticamente pelo sistema são `on_migrate`, invocado antes da migração. `on_arrive`, invocado assim que o processo chega no destino. `hibernate`, invocado quando um processo é adormecido e `on_destroy`, invocado ao se destruir um processo lógico.

1.3.2 Serialização de um processo

Um processo deve possuir a propriedade de ser serializado quando conveniente. Um processo não possui a capacidade de se serializar, ou de serializar um outro processo. O processo de serialização de um processo se dá somente a partir de chamadas internas do *middleware*. Ao usuário cabe garantir que todo o código escrito seja serializável.

1.4 Migração de processos

Um processo pode migrar de um ambiente para outro quando necessário ou conveniente para a simulação. Quando a migração ocorrer, uma série de ações ocorrem na seguinte ordem:

1. O estado do processo na tabela de endereços é modificado de Ativo para Trânsito.
2. O método `on_migrate` é invocado ainda no ambiente de origem do processo.
3. O processo a ser migrado é serializado pelo *environment* de origem e enviado para o ambiente de destino.
4. O ambiente de destino recebe o processo serializado e o desserializa, tornando-o um novo objeto em memória, mas mantendo os dados originais do processo.
5. O ambiente de destino atualiza o estado do processo em sua tabela local de Ausente para Inativo.
6. O ambiente de destino envia uma mensagem para o ambiente de origem indicando que o processo chegou, e qual o novo endereço físico do processo.
7. O *environment* de origem atualiza o estado do processo para Ausente. Atualiza também o endereço físico do processo na tabela de endereços.
8. O processo executa o método `on_arrive` no ambiente de destino.
9. O *environment* muda o status do processo de Inativo para Ativo.
10. Por fim, o processo recupera todas as mensagens do buffer de mensagens do proxy de comunicação, resgatando eventuais mensagens recebidas enquanto o seu estado era Inativo.
11. O processo, já em estado ativo, executa o método `run`.

Assim que o processo termina o ciclo de ações de migração ele está apto a continuar a simulação do ponto onde parou no ambiente antigo. Isto se dá porque o processo foi serializado e todos os dados foram mantidos tais como estavam instantes antes da migração.

Vale ressaltar que uma vez que esteja em processamento (executando o método *run*), o processo só executa a migração após o término da execução do evento em questão. Sendo assim, o método *run* deve ser implementado de maneira a possibilitar a preempção de eventos.

1.4.1 Atualização da tabela de endereços dos processos

Uma vez que um processo migra de um *environment* para outro, instantes após a migração apenas os dois ambientes envolvidos possuem os dados atualizados. Todo ambiente diferente dos envolvidos no processo de migração possui em suas tabelas de endereços de processos dados incorretos quanto a sua localização, portanto, enviariam mensagens para o ambiente antigo, ao qual o processo destinatário não mais pertence.

Ao receber uma mensagem de um antigo hospedeiro, um *environment* (que possui o seu novo endereço), redireciona a mensagem ao proxy do ambiente que possui atualmente o processo em questão. Porém, isso inclui mais um intermediário no processo de transmissão de mensagens. Sendo assim, duas ações, em momentos distintos, são efetuadas para garantir a atualização das tabelas de endereços de processos. Primeiro, o antigo *host* do processo, ao receber a mensagem, além de repassá-la ao atual *host* também devolve uma mensagem para o remetente da mensagem, notificando-o que o endereço do ambiente que contém o processo mudou, e atualiza este endereço.

Em um momento distinto, uma segunda ação de sincronismo de tabelas é disparada. Esta ação é iniciada de forma independente por cada *environment*, enviando uma mensagem para os demais ambientes, notificando-os de quais processos lógicos

encontram-se em seu poder. Isto garante uma atualização constante das diversas tabelas de endereços de processos existente na simulação.

1.5 Troca de mensagens

No modelo de comunicação baseado em proxy, quando um processo deseja enviar uma mensagem a um segundo processo, a mensagem é primeiramente enviada ao proxy do *environment* ao qual o processo remetente reside, e o proxy é responsável por resolver a correlação entre endereço lógico e o endereço físico, e enfim enviar a mensagem ao processo destinatário. O principal motivo de se deixar o proxy responsável pelo envio da mensagem é se possibilitar que existam uma pequena quantidade de tabelas de correlação de endereços, e assim, que existam menos atualizações de tabelas em uma migração.

Em um modelo onde cada processo resolveria o envio de mensagem diretamente, a quantidade de tabelas de endereços seria proporcional à quantidade de processos (e não proporcional à quantidade de ambientes, como é na arquitetura proposta). Ao haver uma migração, existirão menos tabelas a serem atualizadas na arquitetura baseada em comunicação inter-proxies, ao passo que na comunicação direta entre processos, a quantidade de tabelas a serem atualizadas seria bem maior.

1.5.1 Comunicação direta e indireta

Quando a comunicação é feita por dois processos lógicos residentes em diferentes *environments* (e por consequência, em máquinas distintas da rede) a comunicação se passa através do *proxy* do *environment* do remetente, e esse é responsável por enviar a mensagem ao *proxy* do processo destinatário. O fluxo de comunicação é ilustrada de maneira simplificada na Figura 1.4.

Neste caso, quando ao processo P1 enviar uma mensagem ao processo P3, o

próprio processo requisita diretamente ao *environment* o envio da mensagem. A mensagem é então encaminhada de P1 para o *proxy* do seu *environment* que detecta, através da tabela de endereço de processos, o endereço físico do processo alvo.

Uma vez que a mensagem a ser enviada está em posse do *proxy* do *environment* destinatário, duas modalidades de comunicação podem ocorrer. A modalidade padrão é a comunicação indireta, onde o *proxy* em posse da mensagem irá enviá-la ao *proxy* do ambiente que contém o receptor da mensagem. Uma vez que a mensagem esteja no *proxy* do destinatário, este a encaminha para o processo (Figura 1.4).

Uma alternativa à comunicação indireta é a comunicação direta, onde o *proxy* do processo de origem, uma vez obtido o endereço físico do processo destinatário, envia a mensagem diretamente para o processo receptor (Figura 1.5). O método direto é útil por eliminar um intermediário na transmissão da mensagem, mas possui um inconveniente que pode anular o ganho da eliminação do intermediário. Caso haja uma migração, a mensagem não é automaticamente redirecionada para o novo ambiente em que o processo se localiza, mas sim é levantada uma excessão na comunicação, o que levaria a um tratamento de excessão, que por sua vez seria encarregado de localizar o processo em seu novo endereço físico e retransmitir a mensagem.

Cabe salientar que em caso de falha na transmissão da mensagem na modalidade de comunicação direta, a excessão levantada deve ser tratada não pelo processo, mas pelo *environment*, pois a mensagem encontra-se em posse do *proxy*

1.5.2 Comunicação local direta

Uma terceira modalidade de comunicação é a comunicação direta, via endereço físico (Figura 1.6), entre processos que habitam um mesmo ambiente. Nesta modalidade um processo que se comunica constantemente com outro processo no mesmo *environment* adquire o seu endereço físico e faz a comunicação direta, sem

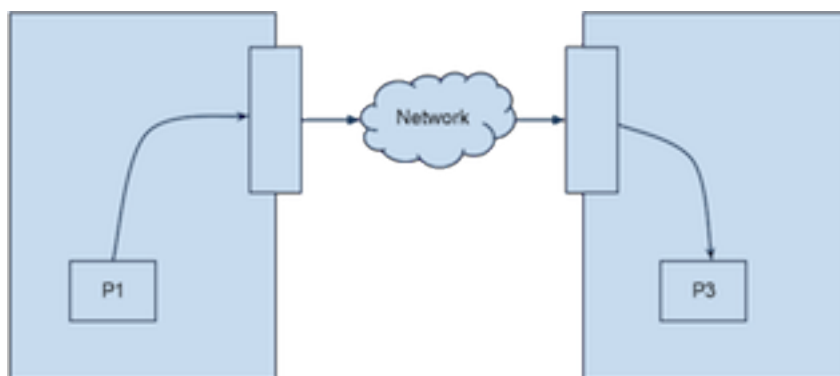


Figura 1.4: Comunicação indireta *proxy-process*.

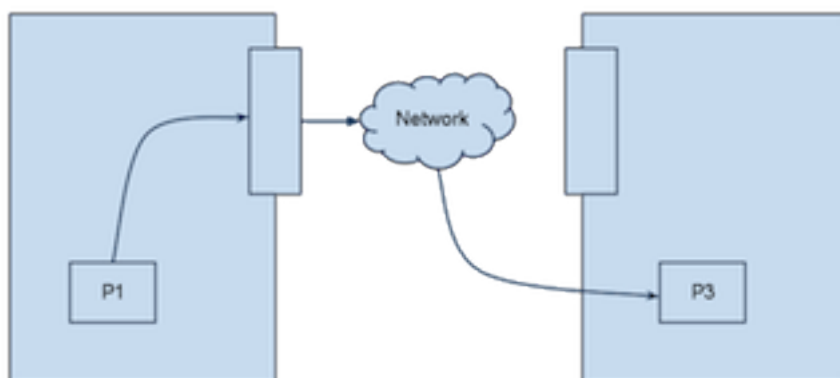


Figura 1.5: Comunicação direta *proxy-process*.

Figura 1.6: Comunicação direta *process-process*.

a necessidade de se passar por um *proxy* de comunicação.

Assim como na comunicação direta, neste caso há um ganho na transmissão da mensagem entre origem e destino por se excluir o *proxy* como intermediário da transmissão. Entretanto, no caso de uma migração, as mensagens não seriam automaticamente redirecionadas para o processo destinatário, cabendo assim ao usuário do *middleware* o tratamento de exceção caso esta aconteça.

A comunicação local direta, assim como a comunicação direta, são opções que devem ser exploradas em casos onde a migração de processos é reduzida ou nula.

1.6 Serialização em grande escala

1.7 Comunicação grupal

1.8 Migrações em grande escala e migrações de ambientes

Referências Bibliográficas

ALVES, A. R.; WALBON, G.; TAKAHASHI, W. *Agentes móveis e Simulação Distribuída*. 2009.

MCQUILLAN, J. M.; WALDEN, D. C. Some considerations for a high performance message-based interprocess communication system. In: . [S.l.]: Proceedings of the 1975 ACM SIGCOMM/SIGOPS workshop on Interprocess communications, 1975. v. 9.