



Universidade Federal do Ceará

Disciplina: Sistemas Distribuídos

Ano/Semestre: 2019.1

Professores: Windson Viana e Fernando Trinta

Roteiro de Estudo sobre Representação de dados de comunicação em SD

1- Contexto

O problema da representação de dados para comunicação aparece quando queremos comunicar partes de um sistema que se encontram em plataformas ou sistemas operacionais distintos. Esse problema pode ser visto de duas formas. Primeiro, como codificar os dados do meu sistema para serem transmitidos em um stream de bytes de forma que seus valores tenham a mesma semântica em duas plataformas. Segundo, como estruturar (organizar, concatenar ou agrupar) os dados para que a outra parte do sistema possa recuperá-la e utilizá-la.

A primeira parte do problema está relacionada à codificação dos dados (*encoding*). Vale lembrar, por exemplo, que uma sequência de bits representando um inteiro (e.g., #A AFF) pode ter valores distintos dependendo do sistema operacional ou da linguagem de programação (i.e., um problema conhecido como *endianess*¹). Além disso, o mapeamento dos bits para um caractere depende de uma tabela de codificação (e.g., ASCII, Unicode, UTF-8, Base-64). A conversão de dados de um formato/tipo/objeto, como uma instância de String em Java, em outro formato (e.g., um stream binário em UTF-8) é um exemplo de codificação. Geralmente, a codificação refere-se ao uso de um codec (tabela combinada entre as partes) para converter um objeto em um stream de bits. Uma pequena revisão ilustrada do problema pode ser vista em:

¹ <https://www.youtube.com/watch?v=B4t1lq3SIAY>

<https://www.youtube.com/watch?v=SbzxEt4oXGY>.

No caso específico dessa aula, nossa preocupação maior é com a segunda parte do problema, a estruturação dos dados a serem enviados. A estruturação de dados para transporte ou armazenamento requer serialização de bytes em uma ordem específica. Isto é, para que a sequência resultante de bytes seja escrita de uma maneira que possa ser lida por outro processo/máquina. Por exemplo, deseja-se transmitir as coordenadas geográficas, o nome e o id de usuário entre duas partes de um sistema (e.g., aplicação móvel e o servidor Web). Em qual sequência esses dados devem ser agrupados? Como separar essas informações de forma que o outro processo possa recuperá-las?

Em Sistemas Distribuídos, os mecanismos que se encarregam disso são chamados de processos de *marshalling* e *unmarshalling*. De certa forma, a serialização de objetos, vista em sala, é um exemplo desse processo. No decorrer do roteiro, veremos quatro outras formas de representação de dados que podem ser utilizadas para esse fim. Bibliotecas em várias linguagens de programação estão disponíveis para transformar um objeto em uma representação nesses formatos (e.g., Java Object to JSON, Python object to XML,...)

2- Pré-teste

Antes de iniciar a aula você tem 20 minutos para responder o seguinte teste abaixo: <https://forms.gle/4TPtD9jPEtMoUnQw8>

3 - Atividade a ser realizada em Sala

Formato de entrega: relatório com respostas de cada questão

Informação: Essa prática foi idealmente pensada para ser feita em laboratório, caso não seja possível, leve papel e caneta para produzir as respostas e depois envie o relatório no período indicado pelos professores

1 - Imagine, então, um aplicativo móvel que mostra Trailers de lançamentos do

cinema e suas avaliações do IMDB. Para tal, ele pagina sua interface e faz o download da descrição de apenas 10 filmes seguindo o formato listado abaixo. Mostre como seria a representação dessa lista de 10 filmes em dos dois formatos escolhidos pela sua equipe. Para o Protocol Buffer, mostre como seria o código da mensagem, já que o payload seria binário.

```
class Filme {  
    int FilmeID;  
    String titulo;  
    int ano;  
    int Avaliacao_IMDB;  
    List<Filme> tres_filmes_mais_relacionados;  
    URL linkparaTrailerNoYoutube;  
}
```

2- Nessa parte da atividade em sala, vamos mesclar as práticas de PBL (Problem Based Learning) e Aula Invertida para fazermos uma análise comparativa das 4 tecnologias de representação de dados externos. 5 casos serão apresentados pelos professores e você deve escolher uma tecnologia adequada e uma inadequada para cada um deles produzindo aqui uma justificativa para tal

3- Para essa atividade, escolha uma das estruturas de representação externa e implemente o processo de serialização e desserialização dos dados dos filmes do item 1. Você pode simular a comunicação do aplicativo móvel com o servidor usando sockets e usando o console como “interface gráfica” **(a ser feita em casa caso não tenha laboratório)**

4- Por fim, compare, de forma superficial, as quatro tecnologias estudadas para seu uso em *marshaling* e a técnica de serialização de objetos da sua linguagem escolhida. Use os seguintes critérios: tamanho dos dados gerados, facilidade de programação, interoperabilidade e desempenho.

4 - Avaliação da Aula Invertida

Contribua para a melhoria da disciplina respondendo ao questionário abaixo:

<https://forms.gle/ZGMDHns2cDTUWeqk8>