

# Paradigmas de Sistemas Distribuídos

## Sistemas Distribuídos em Grande Escala

### “Agregação para dispositivos IoT”

Trabalho Prático

2021/2022

#### Informações gerais

- Cada grupo deve ser constituído por até 5 elementos, de ambas ou de uma das UCs.
- Deve ser entregue o código fonte e um relatório de até 6 páginas (A4, 11pt) no formato PDF.
- O trabalho deve ser entregue até às 23:59 do dia 3 de Junho de 2022 no *eLearning* de qualquer das UCs.
- A apresentação do trabalho ocorrerá em data a anunciar.
- O trabalho será classificado separadamente para cada uma das UCs, de acordo com os objetivos de cada uma delas.

#### Resumo

Elabore um protótipo de uma plataforma para suporte a recolha e agregação de dados enviados por dispositivos IoT (e.g., drones, veículos), em número potencialmente elevado. Esta plataforma é instanciada em vários nós espalhados geograficamente em diferentes *zonas*, para permitir escalabilidade e disponibilizar informação agregada. O protótipo é constituído por dois componentes de software em cada zona, instanciados aos pares e que comunicam entre si: o *colector* e o *agregador*.

Cada dispositivo, consoante a zona em que se encontre, pode autenticar-se no colector da zona, estabelecendo uma sessão onde pode enviar informação para ser agregada. Assume-se que os colectores têm a informação necessária para autenticar os dispositivos (identificação e uma senha).

Os colectores não comunicam entre si mas apenas com o agregador da própria zona. Para o sistema poder funcionar o mais autonomamente possível, independentemente de conectividade global, cada dispositivo deve comunicar diretamente apenas com o colector da zona onde se encontra. A informação relevante deve ser enviada do colector para o agregador respetivo.

Os agregadores das diferentes zonas devem comunicar entre si, de modo a cada um ter informação agregada, não apenas da sua zona, mas relativamente a todo o sistema. Os agregadores formam uma rede sobreposta, em que cada um apenas comunica com alguns vizinhos. Cada agregador tem também papel de publicador de informação agregada, da sua zona ou global. Esta informação, pública e em tempo real, deverá poder ser subscrita por consumidores arbitrários.

## Funcionalidade

### Dispositivos e Consumidores

Cada dispositivo tem um identificador único, uma senha e um tipo (todos *strings*), imutáveis e conhecidos pelos colectores. Cada dispositivo, depois de autenticado, enviando o seu identificador e a senha, pode enviar eventos, identificados por um tipo (*string*), como por exemplo “alarme”, “avaria”, “travagem” (sem mais dados associados ao evento em questão).

O trabalho deve incluir implementações simples de dispositivos e de consumidores que permitam demonstrar o funcionamento da plataforma. É admitida uma implementação centralizada que permita simular um número elevado de dispositivos ligados a várias zonas, para carregar o sistema e avaliar o seu desempenho.

### Colectores

O papel do colector é permitir que muitos dispositivos se liguem em simultâneo, garantindo a sua autenticidade, e possam enviar eventos, encaminhando a informação relevante para o agregador.

Serve ainda para poupar trabalho ao agregador, podendo fazer algum pré-processamento na avaliação do estado de atividade dos dispositivos: um dispositivo é considerado *ativo* se enviou pelo menos um evento no último minuto (sendo portanto considerado inativo se passou mais de um minuto sem eventos enviados).

### Agregadores

O objetivo do agregador é resumir informação proveniente dos dispositivos (via colectores) de modo a permitir responder a perguntas sobre o estado global do sistema. Por exemplo: “quantos automóveis estão *online* (autenticados) no sistema”, ou “quantos alarmes foram despoletados”. Devem poder ser respondidos em concreto os seguintes pedidos:

- o número de dispositivos de um dado tipo online no sistema;
- se um dado dispositivo (com um dado identificador) está online no sistema;
- o número de dispositivos ativos no sistema;
- o número de eventos de um dado tipo (e.g., “alarme”) ocorridos no sistema.

Assuma um número muito elevado de dispositivos e eventos por dispositivo, de modo que não é viável estar a enviar uma mensagem por cada evento que chega, de cada dispositivo, em cada zona, para outras zonas. Mas assumo que as conexões/desconexões de dispositivos são menos frequentes, podendo despoletar mensagens individuais. Em termos de conectividade, não assumo que os dispositivos se movimentam diretamente para zonas vizinhas na rede sobreposta entre agregadores.

### Notificações

Cada agregador serve ainda como publicador de eventos relativos à sua zona, que podem ser subscritos por clientes arbitrários. Uma aplicação cliente deverá permitir registar (e cancelar) o interesse em ser notificada em tempo real sobre diferentes ocorrências:

- quando deixar de haver dispositivos online de um dado tipo na zona;
- atingido record de número de dispositivos online de um dado tipo na zona (com informação do seu valor); um cliente poderá estar interessado em saber se foi atingido algum record, para algum tipo de dispositivo (e não um tipo em particular);

- a percentagem de dispositivos online na zona face ao total online subiu, tendo ficado em mais de  $X\%$  dos dispositivos online, para  $X \in \{10, 20, \dots, 90\}$ .
- a percentagem de dispositivos online na zona face ao total online desceu, tendo ficado em menos de  $X\%$  dos dispositivos online, para  $X \in \{10, 20, \dots, 90\}$ .

## Objetivos

**PSD** O trabalho deve ser efectuado utilizando Java (agregador), Erlang (colector), ZeroMQ (todos os componentes), sockets com comunicação orientada à linha (comunicação entre os dispositivos e os colectores). Deve em particular aproveitar a facilidade que Erlang dá a gerir o estado de sessão, e de diferentes padrões de mensagens oferecidos pelo ZeroMQ.

**SDGE** O trabalho deve construir uma rede sobreposta entre os agregadores e usar mecanismos eficientes para construir, representar, agregar e propagar informação, apropriados para grande escala em termos de dispositivos e eventos gerados por estes, permitindo movimentação de dispositivos que se traduza em saltos entre zonas não contíguas em termos de conectividade.