

O Manifesto pela Reactividade

Publicado em 16 de Setembro de 2014. (v2.0)

As diversas organizações que trabalham nos mais variados domínios estão a descobrir, de forma independente, padrões de desenvolvimento de software muito semelhantes. Estes sistemas são mais robustos, mais resilientes, mais flexíveis e muito mais capazes de responder às exigências de hoje em dia.

Estas alterações acontecem devido às dramáticas mudanças dos requisitos das aplicações dos anos mais recentes. No passado mais recente, uma aplicação era apenas constituída por algumas dezenas de servidores, com tempos de resposta na ordem dos segundos, incontáveis horas de manutenção offline e dados apenas na ordem dos Gigabytes. Hoje em dia, podemos encontrar aplicações nas mais diversas áreas, desde dispositivos móveis até serviços de nuvem executados em milhares de processadores de múltiplos cores. Os utilizadores destes sistemas são muito mais exigentes e esperam tempos de resposta na ordem dos milisegundos com disponibilidade permanente de 100%. O volume de dados passou a ser medido em Petabytes. Simplesmente, as arquitecturas do passado que tão boa conta deram de si, não são capazes de dar resposta às necessidades de hoje.

Acreditamos que é necessária uma abordagem coerente às arquitecturas destes sistemas, e acreditamos que todos os aspectos necessários já foram identificados individualmente: queremos sistemas que sejam Responsivos, Resilientes, Elásticos e Conduzidos por Mensagens. São os chamados Sistemas Reactivos.

Sistemas desenvolvidos como Sistemas Reactivos são mais flexíveis, desacoplados e [escaláveis](#). Assim, são mais fáceis de manter e de mudar. São significativamente mais tolerantes a [falhas](#) e quando estas ocorrem, são tratadas com elegância. Sistemas Reactivos são altamente responsivos, dando aos [utilizadores](#) um verdadeiro feedback interactivo.

Sistemas Reactivos são:

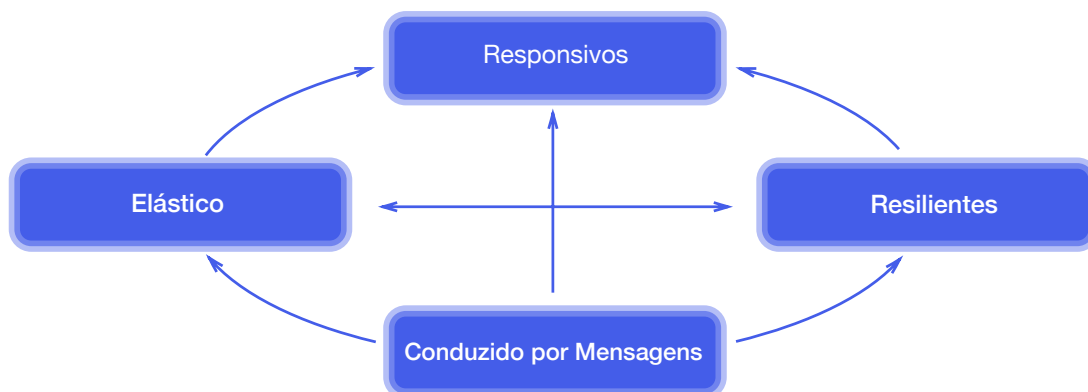
Responsivos: O [sistema](#) responde atempadamente, se possível. Responsividade é a pedra basilar da usabilidade e utilidade, mas muito mais que isso, significa que os problemas podem ser detectados mais rapidamente e tratados com eficácia. Sistemas Responsivos estão focados em oferecer tempos de resposta rápidos e consistentes, estabelecendo margens de tolerância que garantam uma qualidade de serviço regular. Este comportamento consistente, simplifica o tratamento de erros,

transmite confiança ao utilizador final e encoraja a interacção.

Resilientes: O sistema continua responsivo perante a [falha](#). Aplica-se não só apenas a sistemas críticos de alta disponibilidade - qualquer sistema que não seja resiliente não será responsivo após uma falha. Resiliência é obtida através de [replicação](#), contenção, [isolamento](#) e [delegação](#). As falhas são contidas em cada [componente](#), que por sua vez estão isoladas umas das outras e assim garantem que partes do sistema possam falhar e recuperar sem comprometer o sistema como um todo. A recuperação de cada componente é delegada para outro componente externo e a alta-disponibilidade é assegurada por réplica, quando necessário. O tratamento das falhas de um componente não é da responsabilidade dos seus clientes.

Elástico: O sistema mantém-se responsivo face a variações de carga. Sistemas Reactivos podem reagir a mudanças no ritmo de acessos, aumentando ou diminuindo os [recursos](#) alocados para servir estes mesmos acessos. Isto implica que não possa existir qualquer ponto de contenção, possibilitando a divisão ou replicação de componentes, assim como a distribuição dos acessos entre eles. Sistemas Reactivos suportam algoritmos escaláveis, assim como de previsão e de reacção porque geram métricas de desempenho em tempo real. Conseguem [elasticidade](#) de uma forma eficaz em hardware de baixo custo.

Conduzido por Mensagens: Sistemas Reactivos dependem de mensagens [assíncronas](#) para assegurar que os componentes estão desacoplados, isolados e localizados transparentemente, com meios para oferecer delegação de [erros](#) na forma de mensagens. O roteamento de mensagens permite gerir a carga e a elasticidade enquanto que o controlo de fluxo cria e monitoriza filas de mensagens no sistema, que quando necessário são roteadas por caminhos alternativos. Transparência na localização das mensagens como um meio de comunicação, torna possível a gestão de falhas de forma idêntica, quer se trate de um “cluster” ou de apenas uma máquina. Comunicação [não bloqueante](#) permite aos destinatários consumir apenas os [recursos](#) necessários quando estão activos, traduzindo-se numa carga muito menor no sistema.



Os grandes sistemas são compostos por sistemas mais pequenos, que por sua vez dependem das propriedades Reactivas dos seus componentes. Isto significa que os Sistemas Reactivos aplicam estes princípios a todos os níveis, permitindo diversas combinações entre os vários componentes. Os maiores sistemas do mundo dependem de arquitecturas baseadas nestes princípios, servindo as necessidades de biliões de pessoas diariamente. É tempo de aplicar estes princípios conscientemente desde o princípio, do que ter que os descobrir de cada vez que implementamos um novo sistema.