**Título** Arquitetura de Software para Aquisição, Processamento e Armazenamento de Dados

**Resumo alargado**

O sistema de controlo e aquisição de dados é uma parte importante de qualquer experiência moderna de física. À medida que as fronteiras do conhecimento se alargam assim também crescem as especificações para estes sistemas, pelo que isto tem de ser tomado em conta ao desenhá-los. Além disso, visto que estas experiências podem durar vários anos ou mesmo décadas e que a sua natureza inovadora obriga a constantes melhoramentos, é necessário considerar futuras expansões e atualizações. Em investigação de energia de fusão é comum ver os vários diagnósticos e atuadores em torno do reator como subsistemas que fazem parte de um todo e assim ter as tarefas de configuração, controlo, aquisição de dados, acesso a dados e segurança da máquina como parte dum sistema unificado. Tal sistema é chamado CODAC, vindo do inglês *Control, Data Access and Communications* (controlo, acesso a dados e comunicações). Tipicamente estes seguem uma estrutura hierárquica em árvore e usam tecnologia proprietária. A este esquema, apesar da sua modularidade, falta-lhe flexibilidade e torna-se por vezes necessário improvisar para conseguir algum nível de integração de novos subsistemas. Além disso, no caso dum *upgrade* completo do sistema, o que é de esperar numa experiência que dure mais de dez anos, pode-se assistir a um processo lento e oneroso, caso o sistema não esteja preparado para tal.

Esta tese é baseada no trabalho realizado no CODAC do tokamak COMPASS, uma máquina de tamanho médio para investigação em energia de fusão, que esteve instalada em Culham, Reino Unido e foi transferida para Praga, na República Checa. Um tokamak consiste numa câmara toroidal onde um plasma é confinado por campos magnéticos gerados por bobinas externas e pela própria corrente de plasma, conseguindo-se assim as temperaturas e pressões necessárias para ocorrer fusão nuclear. O CODAC do COMPASS é baseado no sistema FireSignal, que está a ser utilizado no tokamak ISTTOK (Lisboa, Portugal) e que fora usado pelo CASTOR, o tokamak que antecedeu o COMPASS em Praga. A operação do COMPASS é baseada em placas de controlo e aquisição de dados, desenvolvidas pelo Instituto de Plasmas e Fusão Nuclear (IPFN) do Instituto Superior Técnico. Estas placas são baseadas no *standard* ATCA e possuem características tecnologicamente avançadas, tais como: FPGA para processamento de sinal e ligações série multi-Gigabit/s para partilha de ultrarrápida de dados entre canais da mesma placa.

Ao desenvolver *software* e *drivers* para integrar no CODAC diagnósticos e atuadores foi necessário não só ter em conta a física por detrás destes como também as capacidades do hardware.

Durante este trabalho, um sistema de controlo em tempo-real baseado na plataforma MARTe (*Multithreaded Application Real-Time executor*) foi integrado com sucesso, tendo partes da operação sido automatizadas e permitido que os dados fossem transferidos diretamente para a base de dados central. O MARTe é uma plataforma que consiste numa coleção de bibliotecas C/C++ que permite o desenvolvimento e lançamento de aplicações de análise e controlo tempo-real. Foi desenhado para ser compatível e correr em várias arquiteturas e sistemas operativos, como por exemplo Windows, Linux, Linux-RTAI e RSX. Para a parte de aquisição de dados, também foram usadas as mesmas placas, mas correndo um *firmware* que lhes permite adquirir 2 milhões de amostras por segundo, a 18 bits por canal, num total de 32 canais por placa. Devido às diferenças na operação do ISTTOK e do COMPASS, foi necessário adaptar o *software* e *drivers* que permitiam a ligação entre as placas de aquisição de dados e o FireSignal, adicionando novos recursos e configurações. Estes estão também agora disponíveis para o ISTTOK, onde foram feitas igualmente modificações para reduzir o tamanho dos dados sem perda significativa de qualidade de sinal. Diagnósticos, tais como a reflectometria, foram igualmente integrados com sucesso, permitindo aos operadores configurá-los usando uma interface comum, e automatizando a fase de preparação o mais possível. Neste caso específico, foi desenvolvida uma biblioteca para o FireSignal e os geradores de micro-ondas do diagnóstico, para configurar os mesmos antes de cada descarga. Após uma fase inicial em que se usou uma base de dados semelhante à usada para o ISTTOK, tornou-se necessário implementar uma nova base de dados, que tivesse mais em conta as características do COMPASS, e desenvolver novas ferramentas para lhe aceder. Foi assim desenvolvido o Compass DataBase (CDB) com o objetivo principal de permitir maior fluxo de dados que o sistema anterior e simplificar a mudança de diagnósticos entre placas de aquisição. Para manter compatibilidade com as aplicações de FireSignal já existentes, foi desenvolvida uma biblioteca para permitir a ligação entre estas e a nova base de dados.

Esta tese apresenta uma análise geral de desenho de CODAC para experiências de física, baseado no trabalho que foi desenvolvido no COMPASS e no ISTTOK. Desafios comuns como: crescimento de volume de dados, integração de eletrónica mais rápida e novas tecnologias são estudadas com o objetivo de propor um estrutura mais flexível e com expansão em mente. Conclui-se que esta deve assentar em três princípios: escalabilidade, adaptabilidade e autonomia. Isto é, capacidade de crescer em número de diagnósticos, número de controladores e volume de dados; capacidade de se adaptar a novas tecnologias, permitindo a convivência de tecnologia de ponta com tecnologia estável; capacidade de modificar *hardware* e *software* de acordo com as suas necessidades, sem demasiada dependência de empresas ou tecnologias específicas. São igualmente feitas considerações mais específicas sobre a estrutura do sistema, recolha e acesso de dados, e ferramentas para os utilizadores. A ideia dum sistema composto por módulos com funções específicas é expandido, sendo feita uma proposta a favor duma topologia em malha híbrida, por ser mais flexível que a topologia em árvore, a mais vulgarmente usada. É também proposta uma base de dados distribuída por vários computadores, para evitar congestionamentos na rede. A solução escolhida deve permitir que os subsistemas vejam as várias máquinas como se fossem uma única e que não tenham de se preocupar com os pormenores do seu funcionamento. Os dados devem ser armazenados num formato *standard* que seja transparente para os utilizadores e é aconselhável comprimir dados antigos, que não sejam acedidos regularmente. Compressão de dados em tempo real também pode ser aplicado a dados enviados em *streaming*, de modo a evitar atrasos imprevisíveis na rede, mas deve-se sempre ter em conta o tempo de compressão. Nesta secção é discutida a evolução de paradigmas de pesquisa de dados em investigação de fusão. Da maneira inicial, em que os dados eram indexados por pulso e tempo, passando pela indexação por eventos, até às propostas de pesquisa por padrões nos sinais. Apesar de tudo, ainda faz sentido ter a indexação por pulso em paralelo com outros métodos, mesmo em plasmas longos, por ser um método intuitivo. Ao se desenhar interfaces para utilizadores, o programador deve ter em conta os hábitos e preferências dos utilizadores, pois pessoas de diferentes origens e áreas têm diferentes maneiras de pensar o mesmo problema e o que parece óbvio e natural a um pode não o ser a outro. Não se deve concentrar todas a funcionalidades necessárias para a operação da experiência numa única aplicação monolítica. Pode ser tentador a ideia de ter muitas funcionalidades integradas no mesmo sítio, mas isso apenas torna a aplicação pesada no uso de memória e confusa para os utilizadores. É aconselhável usar várias aplicações para tarefas específicas (configuração, monitorização e visualização de dados), que partilhem a estética e filosofia de interface.

**Palavras-chave:** CODAC, tempo real, armazenamento de dados, pesquisa de sinais, escalabilidade.