

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

Nome André Sancho Cabral Guilhoto do Vale Duarte

Doutoramento em Engenharia Física Tecnológica

Orientador Prof. Bernado Brotas de Carvalho

Resumo

O sistema de controlo e aquisição de dados é uma parte importante de qualquer experiência moderna de física. À medida de as fronteiras do conhecimento avançam assim também crescem as especificações para estes sistemas e isto tem de ser tomado em conta ao desenhá-los. Além disso, visto que estas experiências podem durar vários anos ou mesmo décadas, é necessário considerar a futuras expansões e atualizações. Em investigação de energia de fusão é comum ver os subsistemas em torno do reator como fazendo 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 de 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.

Esta tese é baseada no trabalho realizado no CODAC do Tokamak COMPASS, uma máquina de tamanho médio para investigação em fusão, que esteve instalada em Culham, Reino Unido e foi transferida para Praga, na República Checa. Este CODAC é baseado no sistema FireSignal, que está a ser utilizado no tokamak ISTTOK (Lisboa, Portugal) e foi usado pelo CASTOR, que antecedeu o COMPASS em Praga. A operação do COMPASS é baseada em placas de controlo e aquisição de dados, desenvolvidas pelo IPFN. Ao desenvolver *software* e *drivers* para integrar no CODAC diagnósticos e atuadores foi necessário não só ter em conta a física por detrás destes bem como também as capacidades do hardware. Durante este trabalho, um sistema de controlo em tempo real baseado na plataforma MARTe foi integrado com sucesso, tendo partes da operação sido automatizadas e permitindo que os dados fossem transferidos para a base de dados central. 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. Foi também necessário desenvolver ferramentas de acesso à base de dados que tivessem em conta as características do COMPASS.

Esta tese apresenta uma análise geral de desenho de CODAC para experiências de física. 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.

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

Title A software architecture for data acquisition, process and storage

Abstract

The control and data acquisition system is an important part of every modern physics experiment. As the boundaries of knowledge move forward, so do the requirements for these systems and this must be taken into account when designing them. Also, since experiments can span the course of several years or even decades, expansion and upgrade must also be considered. In fusion energy research it is common to view the sub-systems around the reactor as part of a whole and having the different tasks of configuration, control, data acquisition, data access and machine security as part of a unified system. Such a system is called a CODAC, standing for Control, Data Access and Communications. Typically they follow a hierarchical, tree-like structure and use proprietary technology.

This thesis is based on the work developed on the CODAC of COMPASS Tokamak, a medium-sized fusion research machine that was installed in Culham, UK, and was transferred to Prague, in the Czech Republic. The CODAC is based on the FireSignal system that is being used at the ISTTOK tokamak (Lisbon, Portugal) and was used in CASTOR, the predecessor of COMPASS in Prague. The operation of COMPASS is based on the ATCA control and acquisition boards developed on IPFN. When developing software and drivers for integrating diagnostics and actuators on the CODAC, it was necessary not only to take into account the physics behind them, but also the hardware capabilities. During this work, a real-time control system based on MARTE was successfully integrated, with parts of the operation being automated and data being transferred to the main database. Diagnostics such as reflectometry were also successfully integrated on the system, allowing operators to configure them on a common interface and automating the setup phase as much as possible. It was also necessary to develop database access tools that took into account the characteristics of COMPASS.

This thesis makes a broad analysis of CODAC design for physics experiments in general. Common challenges such as: increasing data volume, integration of faster electronics and new technologies; are taken into consideration to propose a structure that is more flexible and with expansion in mind.

Keywords: CODAC, real-time, data storage, signal query, scalability.

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 tratar os vários diagnósticos e atuadores em torno do reator como subsistemas que fazem parte de um todo e assim agregar 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 desenvolvendo uma série de módulos de interface para atingir algum nível de integração de novos subsistemas. Além disso, no caso dum *upgrade* completo do sistema, o que é de habitual numa experiência que dure mais de uma década, 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 do IPP-CZ em Praga que antecedeu o COMPASS. 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: dispositivos 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 comprimir o tamanho dos dados sem perda significativa da informação no 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 ainda 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 uma 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 interagem com as várias máquinas como se fossem uma única entidade utilizando um nível adicional de abstração para que não tenham de se preocupar com os detalhes do seu funcionamento respectivo. Os dados adquiridos dos diagnostics experimentais devem ser armazenados num formato *standard* que seja transparente para os utilizadores sendo aconselhável comprimir dados antigos, que habitualmente deixam de ser acedidos regularmente. Compressão de dados em tempo real também pode ser aplicado a dados enviados em modo *streaming*, de modo a evitar atrasos imprevisíveis na rede, no entanto deve-se sempre ter em conta o tempo de compressão. Nesta secção (qual ?) é discutida a evolução de paradigmas de pesquisa de dados em investigação de fusão. Partindo do esquema inicial, em que os dados eram indexados por pulso e tempo absoluto ou relativo a um dado instante inicial (“ T-zero”), passando por uma nova indexação por eventos, até às propostas de indexação e pesquisa por padrões nos sinais. Considerando-se todos estes aspectos conclui-se que fará sentido ter a indexação por pulso em paralelo com outros métodos, mesmo em plasmas longos, sobretudo por ser um método intuitivo. Ao desenhar interfaces para operadores e utilizadores dos sistemas CODAC, o programador deverá ter em conta os seus hábitos e preferências, pois pessoas de diferentes origens e áreas têm diferentes maneiras de abordar o mesmo problema. Não se deve concentrar todas a funcionalidades necessárias para a operação da experiência numa única aplicação monolítica. Embora inicialmente tentadora a ideia de ter muitas funcionalidades integradas no mesmo produto, isso irá tornar a aplicação pesada no uso de memória e demasiado confusa para os utilizadores. É aconselhável desenvolver 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.