

1 - Introdução

1.1 - Sistemas de Controlo Automático

Quando se inicia o trabalho com **sistemas de controlo automático**, numa indústria por exemplo, uma questão nos surge de imediato: “Qual o melhor **sistema de controlo**?”. Não é objectivo aqui nestas notas optar por um ou outro tipo específico em particular, mas sim, fornecer ao leitor um conjunto de informações que lhe permitam optar por si, ou identificar as vantagens de uns perante os outros.

A opção por um **sistema de controlo automático** em particular é muito semelhante à aquisição de um automóvel. Antes de mais existe a necessidade de determinar a sua utilização, caso contrário, poderá ser um mau investimento. “Será para conduzir de casa para o trabalho?”, “Deverá transportar a família toda?”, etc. A sua utilização irá condicionar o tipo de veículo a adquirir. Só depois se deve escolher o fabricante que mais nos convém, seleccionado entre critérios tais como: qual produz a versão mais duradoura; a mais barata; o que possui o melhor apoio ao cliente; etc.

O que foi dito anteriormente também se aplica à escolha de um **sistema de controlo automático**. Espera-se que a sua introdução venha melhorar a produtividade de um processo, a eficiência de funcionamento, bem como diminuir os custos associados à produção.

A necessidade de utilizar um **sistema de controlo** pode ser resumida no conceito “melhorar a produtividade”. São diversos os factores que podem contribuir para degradar a produtividade como, por exemplo, os representados na Fig. 1.1.



Fig. 1.1 – Factores que degradam a produtividade.

Com a utilização, cada vez maior, de microprocessadores, um conjunto de novas potencialidades e capacidades ficam disponíveis para todo o universo de necessidades. Estamos então em condições de melhorar os níveis produção, bem como a qualidade final do produto.

Produção

O **controlo** realizado com base em sistemas microprocessados permite otimizar a organização da produção bem como a utilização de equipamentos, alcançando-se um produto final com utilização óptima dos recursos disponíveis.

O controlador pode, por exemplo, memorizar o número de horas de serviço e avisar o operador do momento exacto de proceder a operações de manutenção, tomando a iniciativa de deslocar o processo produtivo para equipamentos alternativos durante o período de paragem. A manutenção assim programada aparece como uma necessidade ao contrário de um evento arbitrário.

Um **sistema de controlo** baseado num microprocessador pode requerer a intervenção do operador sempre que a produção viole limites estabelecidos, permitindo-lhe estar no topo da cadeia produtiva e executar procedimentos de afinação do processo. De uma forma alternativa, estes procedimentos podem ficar a cargo do processo automático que os realiza coerentemente.

Consistência

Os **sistemas de controlo** baseados em sistemas microprocessados apresentam um desempenho óptimo na execução de tarefas repetitivas.

A capacidade de repetir tarefas, utilizando sempre a mesma parametrização, ou, procedendo às alterações necessárias de forma a adaptar-se a novas exigências, permite obter um produto final sempre dentro das especificações. Factor importante para o valor final desse mesmo produto.

Os sistemas microprocessados permitem também a monitorização, quer dos equipamentos, quer dos operadores, no sentido de avaliar a coerência do seu funcionamento.

Sempre que um sensor não estiver devidamente calibrado, uma válvula apresente uma fuga, ou um operador execute uma acção errada, o processo pode tomar as medidas necessárias para que esses acontecimentos não se propaguem para o produto final.

Eficiência

O **controlo** baseado no microprocessamento não só permite a gestão simultânea de mais processos, como, também, permite a utilização de estratégias de **controlo** mais flexíveis.

Muitas vezes, a estratégia de **controlo** não é nova, mas apenas demasiado complexa para as limitações do equipamento utilizado.

Graças à versatilidade permitida pelos microprocessadores, é possível elaborar estratégias de **controlo** que interligam diversos níveis de comando.

Parâmetros como:

- ☐ temperatura,
- ☐ pressão,
- ☐ fluxo, e
- ☐ nível

podem agora ser conhecidos com maior exactidão, levando a uma evidente diminuição dos desperdícios.

O **controlo** permite também reduzir a quantidade de energia gasta na produção. Com as recentes evoluções nas tecnologias sensoriais, um sinal pode ser continuamente realimentado no controlador, gerindo o processo de uma forma nunca antes permitida.

São exemplos deste tipo de operações:

- ☐ a medida da composição dos gases libertados num processo de combustão para correctamente determinar a mistura ar/fuel, ou,
- ☐ a medida e monitorização da humidade relativa junto de um equipamento sensível aos efeitos da corrosão.

Muitas outras técnicas e estratégias de **controlo** fornecem uma optimização da eficiência global da produção.



Fig. 1.2 – Indústria com **sistemas de controlo automático**.

Segurança

A questão segurança pode apresentar três vertentes diferentes:

- a segurança da instalação,
- a segurança do produto, e
- a segurança do pessoal envolvido no processo.

Se uma válvula abre em demasia e um excesso de material é usado, se o nível de um líquido não sobe quando uma válvula é aberta, ou se a temperatura aumenta em mais que devia, etc., o **sistema de controlo** pode iniciar procedimentos automáticos de recuperação da falha.

Para além disso, o sistema pode recolher informação como operações realizadas, acções de **controlo** tomadas, e a resposta do processo a essas acções, para que na eminência de uma falha esta possa ser detectada a tempo de minimizar os seus efeitos.

Noutras circunstâncias, se o operador erra na introdução de dados, o **sistema de controlo** pode analisar essa entrada e pedir ao operador para a confirmar.

Claro que o sistema deve possuir mecanismos de monitorização e análise das suas próprias falhas, e, para além disso, possuir capacidade de ele próprio recuperar de um estado de falha.

Estas funcionalidades são suportadas por sistemas redundantes que são muito mais fáceis de implementar nos dias de hoje do que no passado.

Custos

Especialmente quando a instalação existente vai sofrer uma actualização, uma justificação dever ser fornecida a favor do novo **sistema de controlo**, por exemplo, que se trata de facto de uma melhoria das capacidades existentes.

Muitas vezes, um sistema microprocessado é consideravelmente menos dispendioso do que o sistema de instrumentação convencional equivalente que substitui.

O custo global de um sistema de comando não pode ser analisado em termos de custos de uma forma imediata.

Há que considerar os custos equivalentes de realizar as funções adicionais, que podem ser alcançadas sem custos acrescidos com os microprocessadores, tais como as referidas de seguida:

- ❑ capacidade de controlo sequencial automático;
- ❑ gestão da informação;
- ❑ recolha de informação para melhor compreender o processo;
- ❑ capacidade de gerar, processar e guardar informação;
- ❑ possibilidade de comparar permanentemente os resultados do processo.

1.2 - A evolução dos sistemas de controlo

Para melhor entender o conceito de **controlo distribuído** vamos dar uma breve introdução histórica relativa às salas de comando e regulação.

Claro que se trata de uma comparação algo subjectiva, dada a diversidade de sistemas encontrados actualmente no mercado. No entanto, e no intuito de introduzir alguns conceitos gerais, somos levados a generalizar.

O início do controlo de processos

O **controlo de processos** com fluidos, como por exemplo o ar comprimido ou o óleo sobre pressão, era feito tradicionalmente por uma grande quantidade de operadores, que circulavam continuamente pela instalação, observando a instrumentação de medida local e accionando as válvulas de forma a manter todo o processo sob *controlo*.

Estes procedimentos podiam ser mais complexos e exigir o registo e cruzamento de dados. A informação assim obtida era utilizada no ajuste de válvulas, amortecedores e motores.

Esta situação requeria da parte de cada operador uma sensibilidade própria que se tornou quase numa arte. Um dos objectivos a alcançar passava pela coordenação de um conjunto de operadores que conduzisse o processo da melhor forma possível.

Dado que esta sensibilidade era própria de cada operador em particular, o desempenho da instalação poderia depender de factores como o estado emocional de cada operador, ou a forma como o conjunto cooperava.

Todos estes aspectos poderiam contribuir para degradar a produtividade.

O aparecimento das salas de comando e regulação

Com os avanços tecnológicos, a capacidade de transmitir sinais pneumáticos tornou-se uma realidade. Desta forma surgiram as salas de comando e regulação.

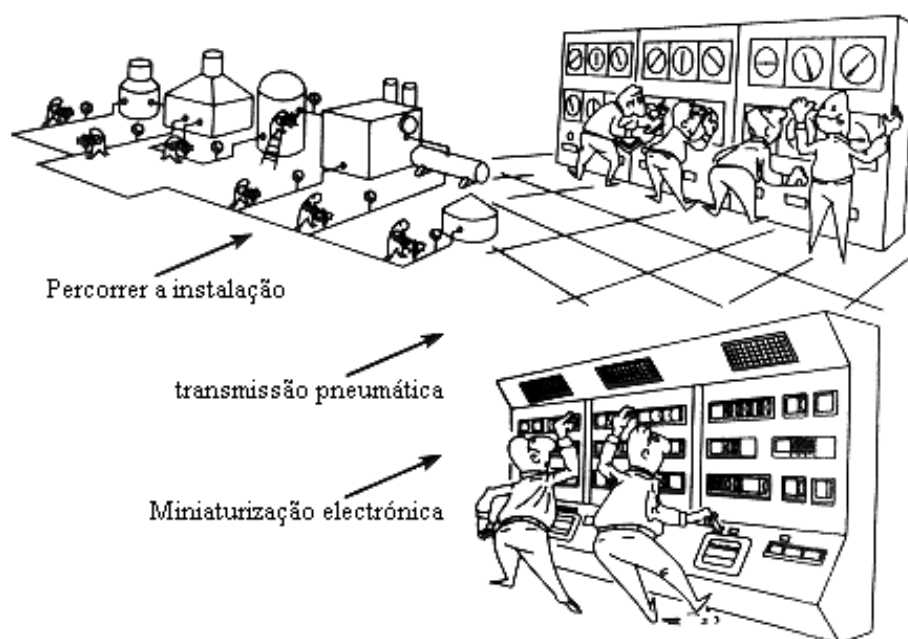


Fig. 1.3 – Evolução dos **sistemas de controlo**.

A instrumentação de medida foi concentrada num único local, em conjunto com a instrumentação que controlava o funcionamento de equipamento específico.

Os operadores podiam, de um único local, registar dados e tomar as devidas acções.

Foi assim introduzido o conceito de trazer a instalação até ao operador em vez de este andar pela instalação.

Assim, os atrasos provocados pela falta de informação podiam ser reduzidos, acelerando o desempenho global do processo.

Era relativamente fácil observar simultaneamente pelo mesmo operador dois processos paralelos e tomar as medidas necessárias.

Todas estas operações eram feitas localmente por meio de ligações directas através de sinais analógicos.

Esta metodologia apresentava como lado positivo a ausência de circuitos (ou tubos, no caso pneumático), e como lado negativo o fraco **controlo**, a ausência de monitorização local do processo, e a insuficiência de avisos.

No final da Segunda Grande Guerra, a electrónica de comando e regulação ficou mais resistente e apta a ser usada em ambiente industriais.

Mais medidas eram possíveis porque também os preços dos sensores ficaram mais barato.

Passou a existir sensores para grandezas que antes não era possível medir, e passou a ser possível medir parâmetros no campo em vez de utilizar estimativas obtidas com base em testes laboratoriais.



Fig. 1.4 – Sala de comando e regulação.

A dimensão dos controladores foi reduzida, permitindo aumentar o número de controladores num painel de comando, bem como diminuir o seu preço.

Tudo isto contribuiu para o aparecimento de *salas de comando e regulação* ainda **mais eficientes**, bem como para a necessidade de circuitos dedicados mais complexos.

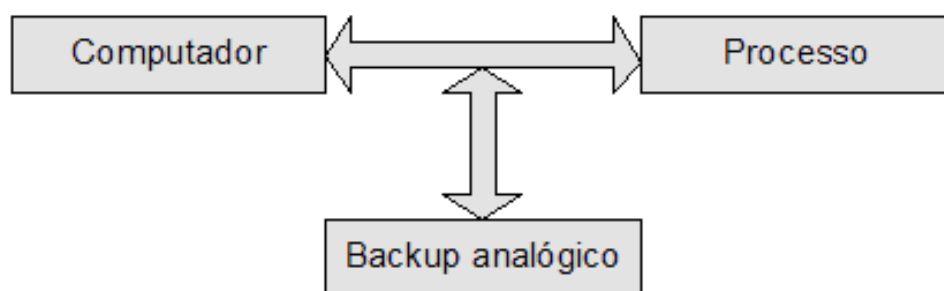
Controlo central computadorizado (Mainframe)

Com as mudanças tecnológicas a possibilitarem a redução dos preços dos computadores, a sua utilização passou a ser mais comum em instalações cada vez maiores e mais complexas, abrindo caminho ao aparecimento de *salas de comando e regulação* cada vez mais sofisticadas.

Progressivamente os computadores suportavam cada vez mais informação, um dos problemas surgidos deveu-se ao facto do desenvolvimento ser feito tendo em vista as aplicações empresariais, tais como as bases de dados ou os programas de facturação.

Durante as décadas de 60 e 70 dois tipos de computadores surgiram para o **controlo**.

- Controlo digital directo (Fig. 1.5);
- Computador supervisor central (Fig. 1.6).



Vantagens:

- Controlo sofisticado
- Controlo flexível
- Aquisição de dados

Desvantagens

- Computador ou controlador redundante
- Circuitos dedicados complexos
- Operadores qualificados
- Dispendioso

Fig. 1.5 – Controlo digital directo.

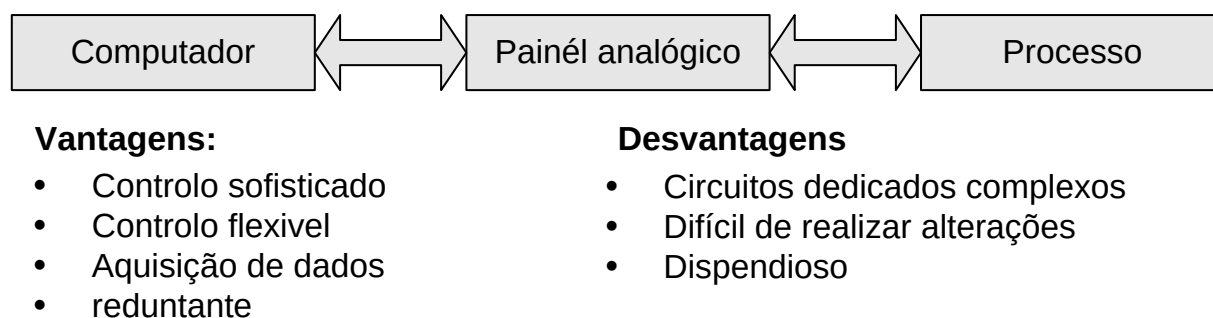


Fig. 1.6 – Computador supervisor central.

Este conceito de *sala de comando e regulação* levou até junto do operador uma imagem mais real do funcionamento de toda a instalação. No entanto, a ligação de todos os pontos remotos da instalação à *sala de comando e regulação* foi feita à custa do seguinte:

- ❑ projecto de engenharia;
- ❑ montagem e manutenção;
- ❑ circuitos e respectivos suportes;
- ❑ introdução de alterações posteriores.

Um problema sério pode também ocorrer aquando de uma falha no computador central, toda a instalação pode ficar inoperante.

Estas situações são minimizadas usando-se controladores de *backup*. Ou seja, para tornar estes sistemas mais fiáveis, é frequente ter que duplicar os **sistemas de controlo**, isto é, comprar tudo em duplicado.

Esta redundância requer por vezes a utilização de instrumentação analógica para manter a fábrica a funcionar. O pessoal qualificado para este tipo de tarefas deve ser necessariamente especializado, e, por isso, caro e de difícil recrutamento.

A utilização de computadores centralizados possui diversas vantagens e desvantagens, que podemos resumir nos pontos seguintes:

- ❑ **Vantagens**: permitir uma visão mais organizada e estruturada do processo; as estratégias de **controlo** podem ser mais flexíveis; as situações de aviso mais eficientes e fáceis de gerir; proporciona a capacidade de registar eventos e produzir históricos quase ilimitada.
- ❑ **Desvantagens**: circuitos eléctricos complexos; é um risco para a instalação; e não se pode acrescentar facilmente novas funções sem se alterar a programação. Corrigir estes tipos de desvantagens é dispendioso e muito trabalhoso.

Esta preocupação com os custos leva a que muitas vezes este tipo de sistemas só se justifique em instalações suficientemente grandes ou críticas.

Nos dias de hoje, os **sistemas de controlo computadorizados** tornaram-se mais acessíveis, porque, ao longo da sua vida útil, a empresa vai sempre necessitar de otimizar o seu funcionamento.

Controlo distribuído de processos (DCS - *Distributed Control System*)

O **controlo distribuído** foi tornado possível pela facilidade com que hoje se manipula a informação, possibilitando a tomada de decisões nela baseadas, de uma forma rápida e coerente.

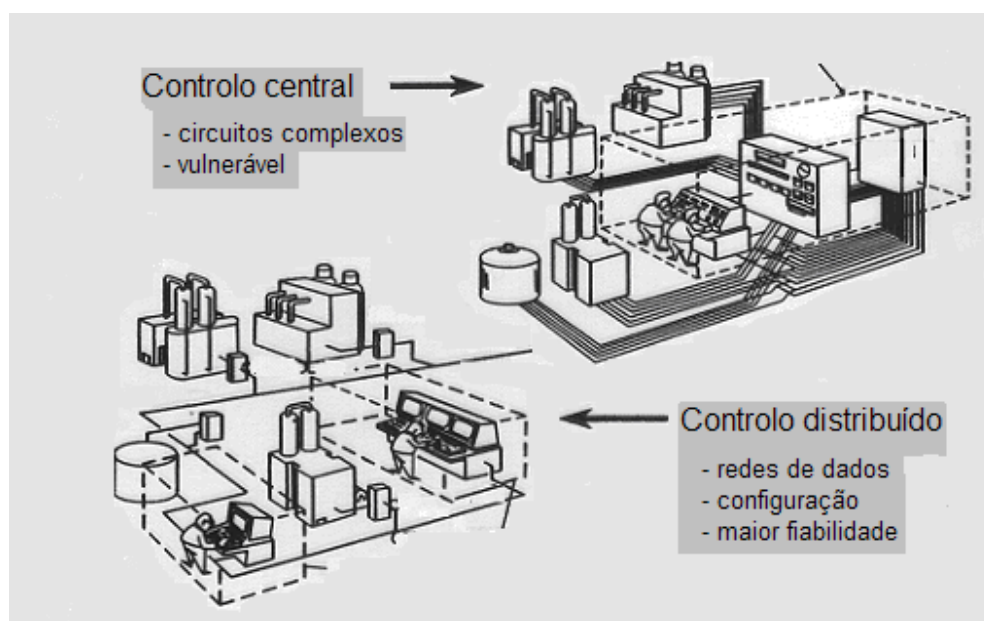


Fig. 1.7 – Evolução do **controlo centralizado** para o **controlo distribuído**.

A complexidade e a dimensão dos circuitos dedicados pode ser reduzida usando a comunicação digital, suportada por um único circuito que forma uma rede de comunicação, ligando os diversos pontos da instalação.

A arquitectura de processos distribuídos possibilita uma distribuição mais funcional das tarefas entre todos os processadores, reduzindo a probabilidade de falha ao mínimo.

Neste tipo de **sistema de controlo** o operador já não necessita de percorrer a instalação para responder a situações de alarme, podendo resolver grande parte dos problemas a partir do seu posto de comando.



Fig. 1.8 – **Posto de comando** em indústria com **controlo distribuído**.

Para além disso, se um processo de fabrico o requerer, podem existir diversas estações de comando ao longo da rede de comunicação. Uma estação de trabalho pode estar localizada num ponto específico, cooperando com a mesma rede, ou ligada directamente a um processo físico que deve controlar.

Como aspectos positivos podemos realçar o *controlo descentralizado* e a monitorização dos processos, traduzindo-se em circuitos mais pequenos, não dedicados, e que interligam os controladores aos processos, logo diminuindo as possibilidades de riscos e falhas ao mínimo.

O sistema assim obtido é mais estruturado, permitindo expandir a rede de comunicações sem grandes alterações de fundo e a custos reduzidos.

Sobressaem como aspectos mais negativos a necessidade de ligar os sensores e os actuadores a um concentrador, de notar ainda que a interligação de diferentes soluções, propostas por diferentes fabricantes, nem sempre é fácil.

Papel dos controladores de lógica programável (PLC) no controlo de processos

Os PLC's ("*Programmable Logic Controllers*", ou "*controladores lógicos programáveis*") foram desenhados para serem usados em funções de automatização de tarefas industriais, onde era requerida muita rapidez na execução de tarefas repetitivas.

Embora esta exigência não seja sempre um requerimento, há algumas situações onde as potencialidades dos PLC podem ser bastante atractivas e eficazes.

Actualmente os PLC podem ser muito mais eficientes no desempenho de tarefas sequenciais, de regulação e interligação, tornando-se o **controlo em tempo real** de motores e equipamentos afins muito mais prático.

Hoje os PLC já suportam a ligação a redes de comunicação, e grande parte deles utiliza redes do tipo ponto a ponto, onde um dispositivo pode dialogar com outro directamente.

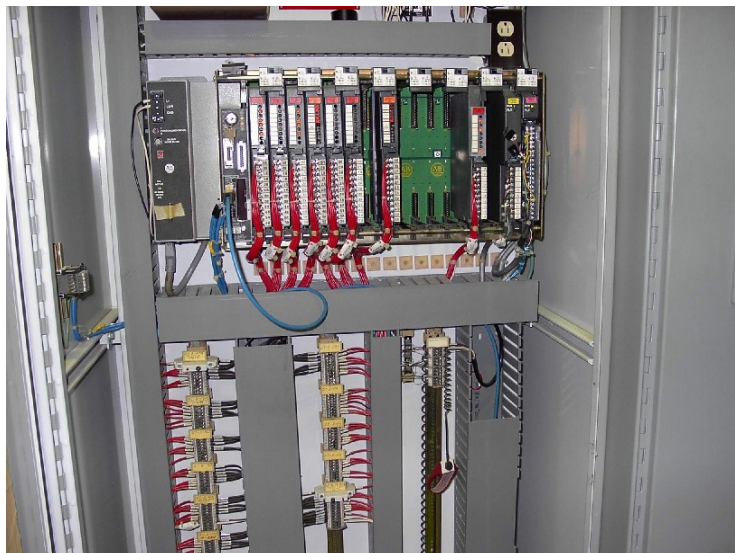


Fig. 1.9 – Controladores PLC's..

O **controlo descentralizado** está disponível em muitos dos PLC de hoje por meio de ligações de entrada/saída inteligentes.

A sua aplicação mais frequente são o

- ❑ arranque e paragem de equipamentos,
- ❑ execução sequencial de tarefas,
- ❑ manuseamento de materiais,
- ❑ etc.

Os PLC são muito frequentemente uma alternativa, em termos de custo, aos DCS, onde não há necessidade de implementar **sistemas de controlo** complexos.

As vantagens dos PLC são as seguintes:

- ❑ boas capacidades de manipulação lógica;
- ❑ muito rápidos, com tempos de resposta na ordem dos milissegundos;
- ❑ muito atraentes em termos de custos;
- ❑ capacidade para suportar condições ambientais adversas;
- ❑ altamente fiáveis e de fácil manutenção;
- ❑ compactos.

As desvantagens dos PLC são as seguintes:

- ❑ não determinísticos; não existe possibilidade de prever o tempo de resposta, o que para algumas estratégias de **controlo** (caso do **controlo** PID) é catastrófico;
- ❑ limitados na complexidade estrutural suportada, como sejam controladores em cascata;
- ❑ necessidade de um controlador pessoal para melhorar o nível de interface;
- ❑ obriga necessariamente a configurar cada PLC em particular;
- ❑ necessita operadores altamente qualificados para o dispositivo em questão.

Comparação entre os diferentes tipos de controlo

	Vantagens	Desvantagens
DCS	Risco distribuído Resposta em tempo real Suporta estratégias de controlo complexas Interface de alto nível Baixo custo de integração	Custo inicial elevado Rede proprietária Sistema operativo proprietário Interligação complexa
PLC	Suporta ambientes agressivos Sistema actual Programado com base na lógica binária Funcionamento sequencial	Interface com o operador Custos de integração Software de aplicação
Computador central	Aquisição de dados Suporta estratégias de controlo complexas Armazenamento e gestão da informação Elaboração de relatórios e históricos	Programação complexa Limitações ambientais Custos de redundância Resposta lenta

Comparação entre DCS e PLC em termos de configuração

Os sistemas DCS eram inicialmente configurados como blocos de funções representados por vários instrumentos e controladores feitos pelo mesmo fabricante. Esses blocos representavam muitas vezes as mesmas vantagens e desvantagens desse mesmo vendedor.

Era difícil encontrar uma configuração *standard* nesta área, mas muitos desses sistemas eram facilmente configuráveis. Em parte devido ao facto de ser também o vendedor a desenhar a consola de trabalho e a respectiva instrumentação. Isto garantia ao operador o funcionamento em tempo real e uma baixa probabilidade de alterar inadvertidamente as configurações do sistema.

O PLC, por outro lado, substitui regularmente os relés e são programados por lógica de relé. Quando a estratégia de **controlo** assim o exige, podem ser programados em linguagens de alto nível, tais como Pascal, Basic, ou outras semelhantes, e possuem uma interface de comunicação simples mas eficiente.

Cada PLC deve ser configurado em separado; é necessário ter um cuidado especial para evitar a duplicação de tarefas; uma tarefa complexa regra geral é confinada a um PLC individual.

Um PC deve poder ser configurado para comunicar com cada PLC em particular afim de aceder a entradas especiais, efectuar visualizações, realizar históricos, etc.

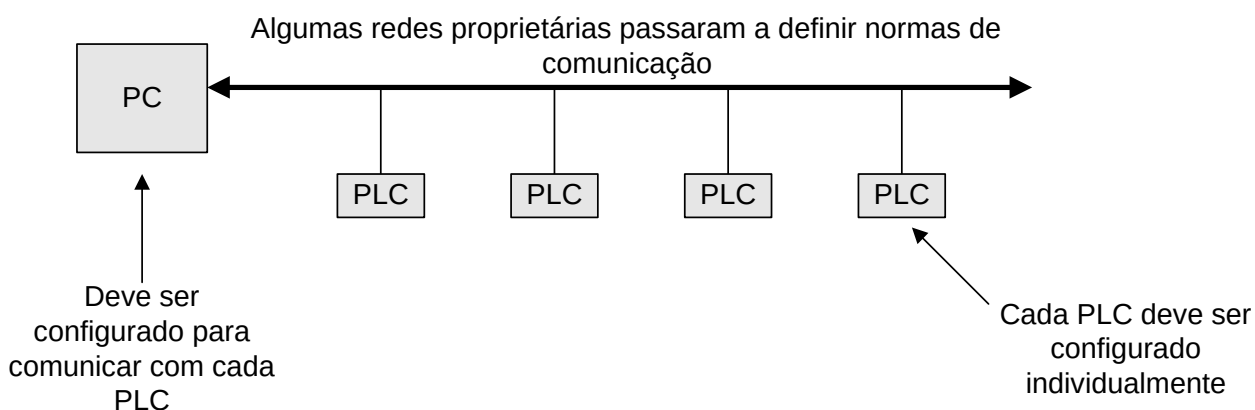


Fig. 1.10 – Sistema típico PLC.

Esta situação obriga à existência de múltiplas bases de dados que devem ser configuradas e mantidas actualizadas.

A configuração é feita, em regra geral, por uma estação de trabalho desenhada especialmente para o efeito.

Cada controlador possuirá uma base de dados que possibilita a ligação ponto a ponto em estratégias de **controlo** complexas.

As bases de dados podem residir quer na estação de trabalho quer nos controladores. Um PC, normalmente pode ser acrescentado e deve ser configurado tal como um PLC.

Há controladores com as capacidades do DCS e dos PLC, ao preço de um PLC, actualmente no mercado.

Estes sistemas podem ser configurados como um único sistema, mas a base de dados é distribuída para que cada controlador suporte a informação que lhe diz respeito.

O protocolo aberto permite a utilização de uma interface homem máquina à escolha do utilizador, desde que possua um *driver* compatível, que responda às necessidades do material já instalado.

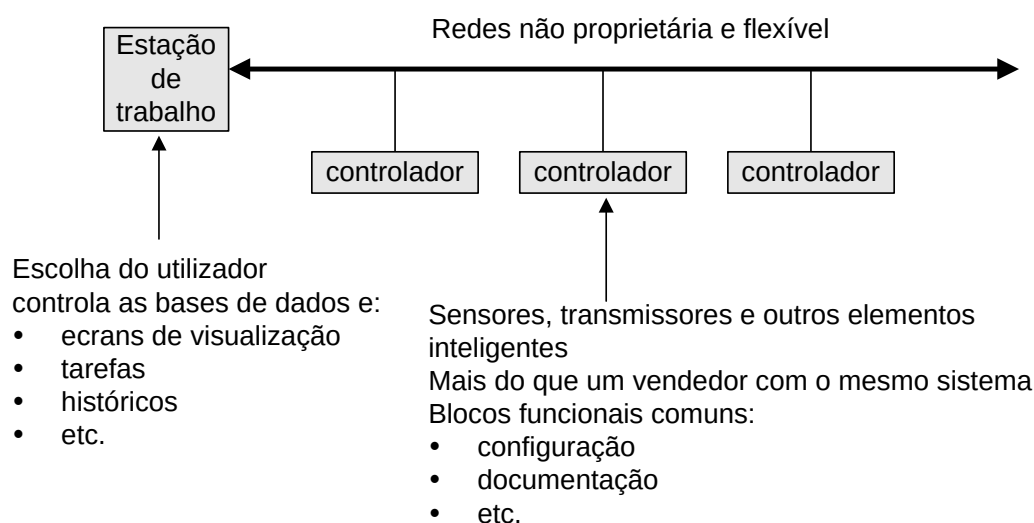


Fig. 1.11 – Emergente, mas por enquanto ainda no domínio hipotético.

Começam a aparecer novas soluções e abordagens do problema. Cada vendedor fornece ao cliente um *software* de configuração, a carregar no PC, para que o dispositivo possa ser configurado com todas as funcionalidades fornecidas pelo vendedor. O *software* para que isto seja possível deve ser escrito de forma universal para todos os sistemas.

Cada vez mais esta é a realidade, e começa já a aparecer no mercado estruturas compostas por dispositivos, como sejam válvulas e transmissores inteligentes, que desempenham estas funcionalidades.

Sistemas SCADA

Nos anos mais recentes, o uso do termo SCADA (*Supervisory Control and Data Aquisition*) tem vindo a ser aplicado não só ao que é na realidade a aquisição de dados, mas também, ao **controlo**.

Usado em larga escala em sistemas de supervisão, *controlo* e aquisição de dados, na implementação de acções de **controlo** e recolha de informação a partir do exterior da instalação.

Os sistemas SCADA não têm tradicionalmente sido usados no **controlo** remoto de instalações, mas apenas, e só, no arranque/paragem de unidades industriais.

São exemplos disso a utilização no comando de postos de transformação, bombas de água ou gás, localizados em locais remotos.

Muitas vezes, a comunicação não é feita por meio de circuitos dedicados, mas sim através de transmissões via rádio, linhas telefónicas, ou mesmo ligações via satélite.

Os atrasos verificados nas comunicações, típicos neste tipo de sistema, fazem com que não se baseiem na monitorização e **controlo** de aspectos particulares do processo.

Da unidade de **controlo** e supervisão apenas se espera que desligue ou ligue determinada unidade em particular, ou, que retire uma unidade que tenha sofrido uma avaria e introduza outra, que se encontre de reserva, na cadeia de funcionamento.

Qualquer comunicação num sistema SCADA deve permitir possíveis atrasos, a que são chamados de latência, entre o requerimento de uma acção e a satisfação dessa ordem.

Um factor de grande importância é também o tratamento de eventuais falhas de comunicação.

Esta possibilidade torna qualquer acção de **controlo** difícil, porque qualquer acção tomada exige uma resposta em tempo útil. Uma solução técnica utilizada consiste em testar antes de executar (*"check-before-execute"*) qualquer acção tomada com base numa comunicação.

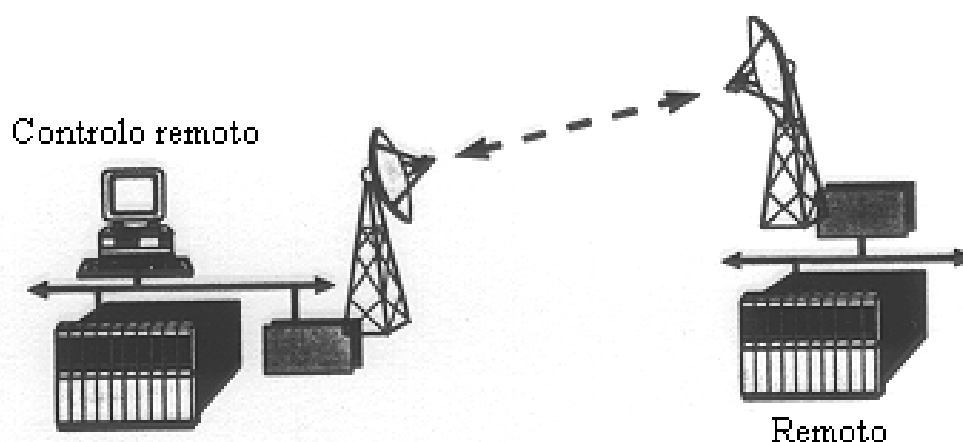


Fig. 1.12 – Sistema SCADA típico.

Importância dos PC no controlo

Quer se goste, ou não, a década de 90 tornou-se na era *Microsoft*. Cada vez mais esta empresa influencia a nossa tecnologia.

A capacidade de negócio e o volume de produtos compatíveis, fez com que muitos dos produtos desenvolvidos por esta empresa estabelecesse normas a preços reduzidos.

Esta situação abriu caminho ao aparecimento no mercado de produtos para todos os níveis de **controlo de processos**.

A configuração normalmente requerida para o PC é a estação de trabalho. O UNIX tornou-se no cavalo de batalha das capacidades de processamento e estabilidade. Ao longo destes anos evoluiu para uma plataforma de trabalho a usar em situações de processamento crítico. No entanto, está rapidamente a ser ultrapassado pelo Windows NT, ou semelhantes.

A indústria de **controlo de sistemas** nunca teve a vantagem proporcionada pela quantidade de volume de vendas.

Pelo menos quando comparada com a área de negócios, e muito menos quando comparada com o *software* de consumo.

A investigação dedicada às novas tecnologias é sempre direccionada para as mais rentáveis, ou seja, para as que podem suportar os custos envolvidos.

Aos outros utilizadores apenas resta esperar e proceder às alterações necessárias para o seu próprio uso.

O objectivo central é o processamento de tempo real

A grande batalha não é entre o UNIX e o NT, porque nenhum deles é um sistema de tempo real por natureza. Existem tantas definições de sistemas de tempo real como fabricantes. De uma forma simplista, para efeitos de **controlo de sistemas**, o tempo real significa fazer uma tarefa sempre que necessário.

Outras definições identificam sistemas de tempo real determinísticos como sistemas que cumprem sempre as funções no momento exacto, e não determinísticos como sistemas de tempo real que vão cumprindo as funções à medida das capacidades de processamento disponível na máquina.

Nos sistemas determinísticos, uma latência não é aceitável, seja qual for o motivo. A sua ocorrência pode ser catastrófica e os custos extremamente altos. Os sistemas não determinísticos, por seu lado, são caracterizados por uma evolução dos custos associada à evolução da latência, havendo um ponto até ao qual essa relação ainda é permitida.

Um sistema operativo de tempo real tem que incluir algumas capacidades, tais como a execução multi-tarefa, para que o *software* mais crítico possa sempre ser executado, mesmo às custas do sacrifício do menos crítico.

Há também, para além das exigências requeridas para tomar as acções de **controlo**, a necessidade de mostrar ao utilizar que a acção foi realizada, mesmo que a inércia do equipamento introduza algum atraso nesse caso.

1.3 - Dispositivos de processamento - Pormenores práticos

Este capítulo discute os conceitos básicos que tornam os sistemas de processamento distribuído possíveis. Embora se trate uma matéria já sobejamente conhecida para muitas pessoas, que operam regularmente com sistemas computacionais, não deixa de ser interessante rever alguns conceitos.

Circuitos integrados

O desenvolvimento tecnológico que tornou viáveis os **sistemas de controlo distribuídos** foi a integração em larga escala de circuitos electrónicos (LSI – *Large Scale Integrated*).

A arquitectura de computadores está bem estabelecida desde os anos 60, mas foi durante a última década que a integração, numa única pastilha de silício, de grandes quantidades de circuitos foi tornada possível.

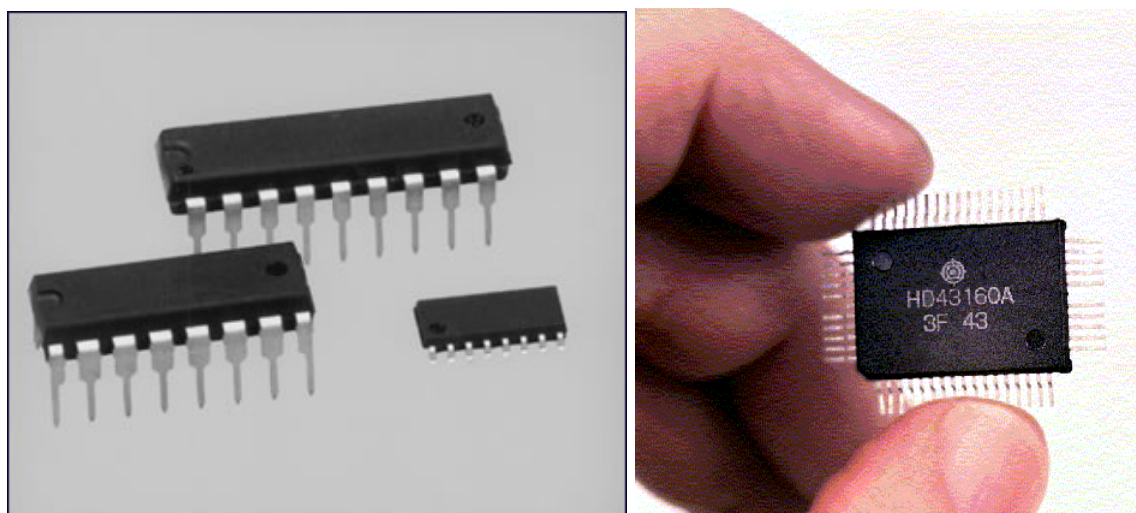


Fig. 1.13 – Circuito integrados.

Como principal consequência o consumo energético e a quantidade de calor foram reduzidos, aumentando a fiabilidade e reduzindo os custos.

Ao computador assim obtido, de reduzidas dimensões, mas com elevada capacidade de processamento, convencionou-se chamar-se micro-computador.

Durante décadas, o desenho dos semicondutores usados com maior sucesso baseavam-se na tecnologia CMOS (*Complementary Symetry Metal-Oxide Semiconductor*).

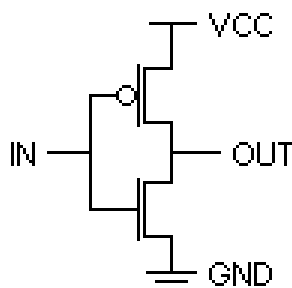


Fig. 1.14 – Um circuito lógico CMOS (inversor).

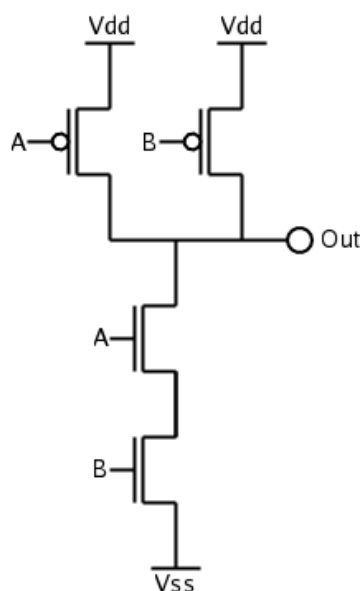


Fig. 1.15 – Um circuito lógico CMOS (porta NAND).

Embora os componentes miniaturizados não dissipem muito calor, existem limitações para o número de circuitos a incluir num único integrado. Mas mesmo assim o seu número tem aumentado rapidamente.

A tecnologia CMOS permite a redução de calor gerado num circuito, porque a corrente eléctrica apenas existe quando se dá a transição de estado.

Num regime estacionário, o que ocorre durante grande parte do tempo de funcionamento, o consumo energético é reduzido e logo a quantidade de calor minimizada.

Embora exista uma grande variedade de componentes eléctricos, os mais comuns são:

- o a resistência que origina uma oposição à passagem da corrente;
- o o condensador que armazena carga eléctrica;
- o os díodos que só permitem a passagem de corrente num sentido; e
- o o transístor que funciona principalmente em dois modos diferentes: num, o fluxo de corrente pode ser controlado numa situação *on/off*, e no outro, funciona como um amplificador de sinal.

Um circuito integrado (CI) não mais é do que um conjunto destes componentes ligados entre si partilhando o mesmo substrato e respectivo encapsulamento.

Os semicondutores são criados introduzindo deliberadamente impurezas (*outros átomos*) na estrutura cristalina formada pelos átomos de silício, a que se dá a designação de "dopping" ou dopagem.

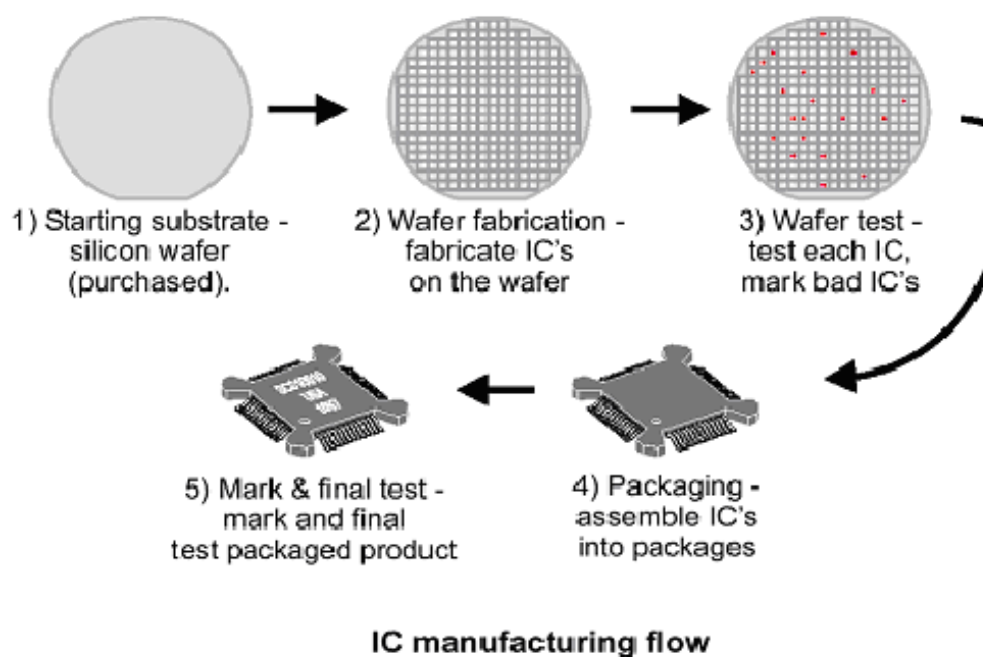


Fig. 1.16 – Procedimento de fabrico de um circuito integrado.

O silício usado no fabrico de CI é extremamente puro, apresentando uma forma cristalina, usado na forma de discos (*wafer*) com uma poucos milímetros de espessura e qualquer coisa como 100mm a 300mm de diâmetro.

Os discos de silício são planos, apresentam um aspecto altamente polido e estão livres de partículas no início do processo de fabrico, que é formado por cinco etapas principais (Fig. 1.16).

Objecto	Tamanho aproximado
Grão de areia	100 a 1000 μ m
Diâmetro de cabelo humano	30 a 200 μ m
Bactéria	1 a 30 μ m
Fumo do tabaco	10nm a 1 μ m
Vírus	3 a 40nm
Átomos	0.2 a 0.3nm

Tab. 1.1 – Comparação entre a dimensão de diversos objectos.

O processo de fabrico actualmente em produção utiliza dimensões entre os 130nm e os 70nm.

Na Tab. 1.1 é possível ter uma ideia comparativa das dimensões com objectos comuns.

É de se observar que 70nm é ligeiramente maior do que um vírus grande, menor do que o fumo do tabaco, e muito menor do que uma bactéria.

O processo de fabrico de CI consiste em selectivamente alterar as propriedades do silício para assim construir componentes eléctricas e respectivas ligações.

A produção decorre em salas extremamente limpas (*Cleanroom*), sendo normal uma concentração de partículas inferior a 100/pt².

Das diferentes etapas de fabrico há a destacar o processo de fotolitografia, em que um material fotossensível é depositado sobre o disco de silício, depois de seco uma máscara é aplicada sobre o material quando exposto à luz ultravioleta.

Este padrão é depois revelado e atacado com produtos específicos, para assim construir as camadas necessárias. O processo é resumido pela Fig. 1.17.

As características do silício podem ser alteradas por implantação iónica.

A introdução controlada de impurezas permite formar os componentes eléctricos desejados e respectivas ligações.

A título de exemplo é ilustrado na Fig. 1.18 o processador Pentium IV.

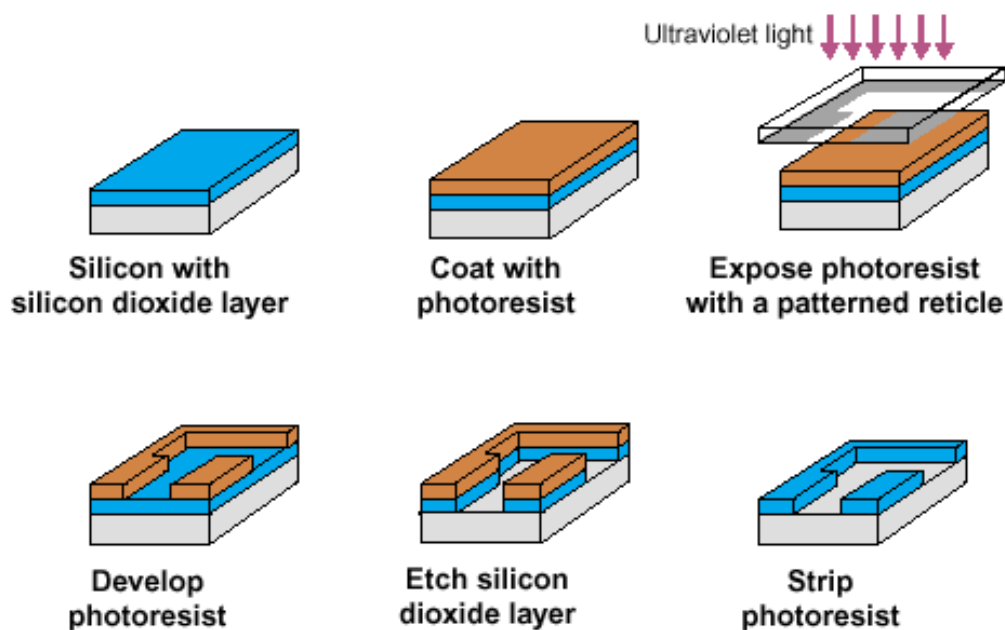


Fig. 1.17 – Procedimento de fabrico de um circuito integrado.



Fig. 1.18 – Processador Pentium IV.

No fim do processo de construção os CI de um mesmo disco de silício são testados, marcando os que não estão correctos, e cortados.

A Fig. 1.19 representa esse processo que termina com o encapsulado do CI num formato específico, afim de obter a robustez mecânica necessária à sua utilização.

Os tipos de encapsulamento mais usados actualmente podem ser vistos, por percentagem de utilização, na Fig. 1.20.

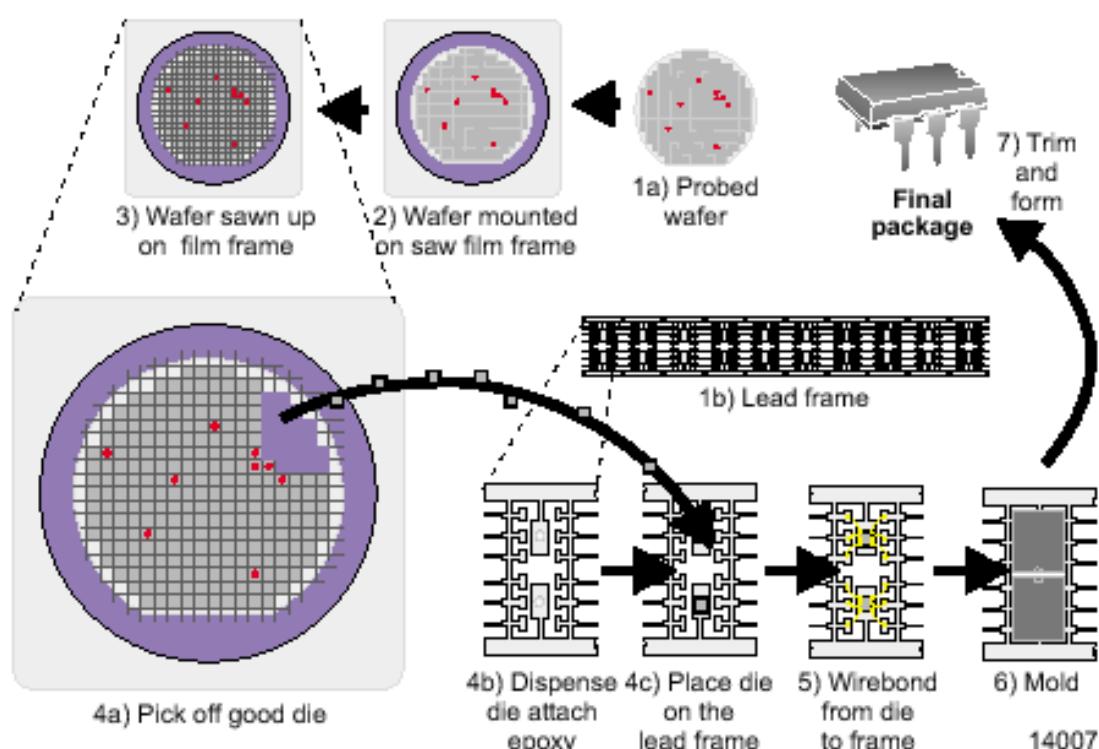


Fig. 1.19 – Processo de encapsulamento.

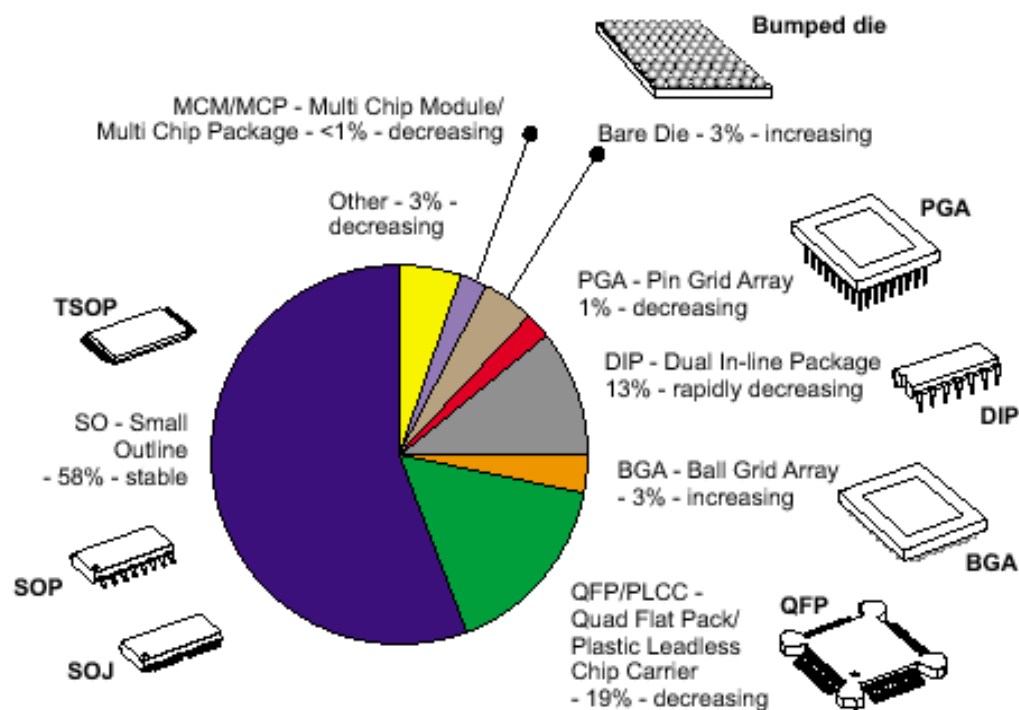


Fig. 1.20 – Tipos de encapsulamento.

Variantes tecnológicas de sistemas de memória

A memória do sistema é o local onde o computador mantém os programas e a informação que está a usar. O próprio tema memória é de certa forma ambíguo, já que pode corresponder a diferentes dispositivos presentes no computador. Mas de uma forma generalizada podemos afirmar que diz respeito ao armazenamento das instruções a serem executadas pelo computador, e dos dados que elas operam.

A memória representa um papel muito importante nos seguintes aspectos:

- **Desempenho:** A quantidade e o tipo de memória em uso são um importante contributo para o desempenho geral. Em muitos aspectos, pode tornar-se tão importante como o próprio processador. Memória insuficiente pode fazer com que um processador trabalhe a menos de 50% das suas capacidades normais. Este ponto, apesar de importante, é muitas vezes desprezado.
- **Suporte de software:** Novas aplicações de *software* requerem cada vez mais memória. Mais capacidade de memória possibilita o acesso a aplicações informáticas cada vez mais eficazes.
- **Fiabilidade e estabilidade:** Más memórias são muitas vezes a fonte dos misteriosos problemas com que se depara a informática. O desempenho do computador em fiabilidade e estabilidade pode ser optimizado utilizando memória de boa qualidade, para além de adequadas às necessidades.
- **Actualização:** Existem muitos tipos diferentes de memórias no mercado, algumas mais universais que outras. Ao fazer uma escolha adequada estamos a viabilizar a hipótese de futuras evoluções com um mínimo de custos.

Esta secção descreve resumidamente as diferentes tecnologias mais utilizadas normalmente.

Memórias apenas de leitura (*Read-Only Memory*)

Esta secção discute os principais tipos de memórias usadas nos computadores, bem como as suas principais diferenças. O objectivo é explicar as principais tecnologias usadas no desempenho de diferentes aplicações. Um tipo de memória usado nos processadores é chamado de memória apenas de leitura (*Read-Only Memory*), ou de uma forma abreviada ROM.

A ROM é um tipo de memória que apenas pode ser lida, em oposição com a RAM que tanto pode ser lida como escrita. Existem duas razões principais que justificam o seu uso:

- **Performance:** A informação é armazenada de forma permanente numa ROM, independentemente de estar, ou não, a ser alimentada. Permitindo a sua extracção do sistema e armazenamento para posterior utilização. Por este motivo é também chamada de memória não volátil.
- **Segurança:** O facto de as ROM não poderem ser modificadas facilmente fornece uma medida de segurança contra alterações acidentais do seu conteúdo. É difícil encontrar vírus que infectem memórias ROM, apesar de não ser tecnicamente impossível nas memórias do tipo EPROM.

As memórias apenas de leitura são normalmente usadas ao nível do sistema para armazenamento de programas ou dados que desejamos sempre disponíveis. A utilização mais comum é o exemplo do programa de BIOS do computador.

Estando permanentemente gravada nesta memória, o computador pode executar os procedimentos de arranque após ter sido ligado.

Embora a principal característica de uma memória ROM seja o facto de o seu conteúdo não poder ser alterado, existem situações em que mudar o conteúdo pode revelar-se útil. Existem diferentes variantes de memórias ROM que podem ser alteradas mediante determinadas circunstâncias.

Em seguida é feita referência das diversas variantes, com uma descrição sumária da sua alterabilidade.

- **ROM:** Uma memória ROM, construída através de lógica na própria pastilha de silício, deve desempenhar uma função fixa e não pode ser alterada. Esta característica é muito inflexível, por isso este tipo de dispositivo apenas é usado em situações em que o seu conteúdo não mais necessita de ser modificado, ou em produções em massa de um dado produto.
- **ROM programável (PROM – *Programmable ROM*):** Esta memória ROM pode ser programada usando equipamento especial, mas apenas pode ser escrita uma vez. Este dispositivo é útil em situações em que é necessário possuir diferentes versões do conteúdo.
- **ROM reprogramável (EPROM – *Erasable Programmable ROM*):** Uma EPROM é uma ROM em que o seu conteúdo pode ser apagado e reprogramado. Uma pequena janela de vidro, localizada no topo do circuito integrado, torna visível o integrado. A luz ultravioleta, com uma frequência específica pode ser utilizada para apagar o conteúdo da memória. Para isso, o integrado é sujeito, durante um período de tempo predefinido, à incidência dessa luz. Como é óbvio, uma EPROM é muito mais útil do que uma ROM.

- **ROM electricamente reprogramável (EEPROM – *Electrically Erasable Programmable ROM*):** O próximo nível de programação é a memória EEPROM, que pode ser programada por *software*. Este é o tipo de ROM mais flexível, sendo actualmente usada para conter os programas de BIOS dos computadores. Sempre que nos referimos a BIOS do tipo *flash*, será sinónimo de que a BIOS pode ser reprogramada através de *software* específico para o efeito. É claro que este dispositivo está preparado para apenas ser programado esporadicamente.

Devemos ter em atenção o facto de que, tal como as memórias do tipo RAM, também as memórias do tipo ROM não necessitam de ser lidas de forma sequencial. Outro factor a ter em atenção é o facto de as memórias ROM serem muito mais lentas do que as memórias RAM, tipicamente com tempos de acesso duplos.

Memórias de acesso aleatório (RAM – *Random Access Memory*)

O tipo de memória que contém os programas e a informação que estão a ser executados denomina-se por RAM.

Este tipo de memória difere das memórias ROM no sentido em que tanto pode ter acessos de leitura como de escrita.

Também é conhecida por memória volátil, em contraponto com as ROM, dado que a informação que contém é perdida sempre que a sua alimentação é retirada. Estas memórias são também chamadas de memórias de leitura escrita (RWM – *Read-Write Memory*).

As memórias RAM são muito mais rápidas do que as memórias ROM, devido à forma como armazenam a informação. Existem diferentes tipos de memórias RAM, incluindo as memórias estáticas (SRAM – *Static RAM*) e muitas variantes de memórias RAM dinâmicas (DRAM – *Dynamic RAM*).

Memórias RAM estáticas (SRAM)

As memórias estáticas armazenam a informação sem necessidade de estarem constantemente a serem refrescadas, desde que a memória esteja alimentada. Isto é contrastante com as memórias dinâmicas DRAM que devem ser refrescadas várias vezes por segundo para poderem manter o seu conteúdo.

- **Simplicidade:** As SRAM não requerem circuitos externos de refrescamento ou qualquer outro tipo de tarefa de forma a manter o conteúdo da memória intacto, que não seja manter a alimentação.

- **Velocidade:** As memórias SRAM são muito mais rápidas do que as memórias DRAM.

Em contraste as SRAM's possuem as seguintes fraquezas comparativamente às DRAM's:

- **Custo:** Uma memória SRAM, com a mesma capacidade de uma memória DRAM, é comparativamente mais cara.
- **Tamanho:** As memórias SRAM exigem mais espaço que as memórias DRAM.

Estas vantagens e desvantagens mostram que em termos de performance o desempenho das memórias SRAM é superior, quando comparado com as memórias DRAM. Infelizmente 32MB de SRAM são proibitivamente dispendiosos para além de ocuparem muito espaço.

É este o motivo porque as memórias do tipo DRAM são usadas para implementar os sistemas de memória. As memórias SRAM são usadas para implementar as memórias *cache* que necessitam de ser mais rápidas.

As memórias SRAM são fabricadas de uma forma muito semelhante à dos processadores: matrizes de transístores altamente integrados numa pastilha de silício.

Cada bit de um memória SRAM é composto por quatro a seis transístores, ocupando assim muito mais espaço que uma memória DRAM que apenas usa um transístor e um condensador.

Como uma memória SRAM é composta por uma malha de centena ou milhares de transístores, o seu fabrico é muito mais fácil comparativamente ao fabrico de um processador, que é uma estrutura não repetitiva.

Esta é uma das razões porque as memórias SRAM custam menos do que um processador.

Memórias RAM dinâmicas (DRAM)

Uma memória dinâmica é um tipo de RAM que apenas retém a informação se for continuamente refrescada por um circuito especial denominado por circuito de refrescamento.

Estes circuitos lêem o conteúdo de cada célula de memória muitos milhares de vezes por segundo, independentemente desta (a memória) estar a ser usada, ou não, pelo processador.

Dada a forma como estas células de memória são construídas, a acção de leitura faz com que o conteúdo da célula de memória seja refrescado.

Se esta tarefa não for executada periodicamente o seu conteúdo será perdido, mesmo que continue a ser alimentada. É devido a esta tarefa de refrescamento que esta memória se chama dinâmica.

Todos os computadores usam memórias DRAM para implementar a memória do sistema, em alternativa às memórias SRAM. Mesmo tendo consciência da necessidade de circuitos de refrescamento e tratando-se de memórias mais lentas.

Pode parecer estranha esta situação, mas o facto de se usar este tipo de memórias está relacionado com o preço mais baixo das DRAM's, e por elas ocuparem 1/4 de área de uma memória SRAM de capacidade equivalente.

O facto de necessitarem um circuito de refrescamento é facilmente ultrapassado.

As memórias DRAM são mais pequenas e mais baratas do que as memórias SRAM, dado usarem apenas um transístor e um condensador por cada bit.

O condensador quando armazena carga eléctrica suporta um 1 lógico, quando descarregado suporta um 0 lógico.

O transístor é usado na leitura do conteúdo do condensador. O problema do condensador é que apenas pode suportar uma carga durante um pequeno período de tempo, descarregando ao longo do tempo.

Como estes condensadores são muito pequenos o tempo de descarregamento é muito rápido.

É por este motivo que o circuito de refrescamento é necessário e tem como função: ler o conteúdo de cada célula e refrescar antes que se perca.

As memórias DRAM são fabricadas usando um processo similar ao dos computadores: num substrato de silício é criada uma malha de transístores e de condensadores que suportam cada bit.

Os dispositivos são mais baratos porque a sua estrutura repetitiva é muito simples e ocupam pouco espaço.

Por isso não é muito complexo construir num único integrado uma estrutura repetida formada por milhares de transístores e condensadores.

Existem muitas variantes tecnológicas de memória DRAM e velocidades a que se encontram disponíveis.

Constituição interna dos microprocessadores

A título de exemplificar o funcionamento interno de um processador vamos neste ponto analisar brevemente a constituição de um processador com algumas funcionalidades especiais como é o caso dos processadores digitais de sinal (DSP – *Digital Signal Processor*).

Os DSP's (*Digital Signal Processor*) programáveis são dispositivos especialistas no processamento digital de sinal em situações de cálculo numérico intensivo, é natural que uns estejam mais adaptados a uma determinada tarefa do que outros.

A escolha de um DSP em particular deve observar características, tão importantes para o produto final como para a fase de desenvolvimento, como sejam o desempenho, o custo, a integração, a facilidade de desenvolvimento, e o consumo energético.

Embora exista uma grande diversidade de dispositivos, todos eles partilham as mesmas funcionalidades básicas.

O desenvolvimento de uma aplicação compreende duas vertentes diferentes a considerar na escolha de um DSP.

Por um lado temos as ferramentas de *software*: *assemblers*; *linkers*; *debuggers*; bibliotecas de *software*; e sistemas operativos de tempo real.

Por outro lado temos ferramentas de *hardware*: sistemas de desenvolvimento; e emuladores.

Normalmente o desenvolvimento é feito em linguagem *assembly* por fornecer um código mais otimizado.

No entanto, o tempo despendido no desenvolvimento é mais logo e o resultado final mais difícil de manter e alterar, comprometendo a portabilidade do código para outras famílias de DSP.

As linguagens de alto nível, como o C ou ADA, também são utilizadas, mas geram código pouco eficiente, mais lento, e demasiado longo para determinadas aplicações.

Os processadores mais modernos proporcionam mecanismos internos de emulação e desenvolvimento, acedidos por interface série JTAG em conformidade com o *standard* IEEE 1149.1, em alternativa, o desenvolvimento pode ser feito com auxílio de um programa monitor residente na memória do DSP.

Esta última opção, apesar de eficaz, consome alguns recursos do processador dificultando o trabalho de desenvolvimento.

A performance do processador é avaliada usando diversos critérios, sendo o mais fundamental o tempo de execução de uma instrução.

O seu recíproco dividido por um milhão define o número de instruções executadas por segundo (MIPS).

O problema deste parâmetro é que a quantidade de trabalho realizada por uma instrução varia de máquina para máquina.

A forma mais eficaz de resolver o problema é definir um conjunto de funções, como por exemplo filtros FIR ou IIR, e avaliar o desempenho das máquinas em comparação.

É necessário ter alguma atenção ao avaliar parâmetros como milhões de operações por segundo (MOPS) e milhões de operações de vírgula flutuante por segundo (MFLOPS), porque diferentes fabricantes possuem diferentes opiniões sobre o assunto.

Analisar a velocidade do relógio é também complicado, já que uma instrução pode ser executada à mesma frequência ou duas ou quatro vezes mais rápido.

Tudo dependendo se o DSP possui, ou não, uma PLL (*Phase-Locked Loops*) interna que lhe permita obter múltiplos da frequência de entrada do relógio.

Uma das características mais importantes num DSP programável é a forma de representação numérica.

Muitos deles utilizam vírgula fixa, embora a representação em vírgula flutuante seja mais flexível e ofereça uma gama dinâmica mais alargada, libertando o programador de preocupações como a precisão.

Em contraste, os DSP de vírgula fixa exigem do programador uma maior atenção no escalonamento dos valores numéricos nas diferentes etapas do processamento de forma a assegurar uma precisão adequada com a gama dinâmica disponível.

Como resultado, os DSP de vírgula flutuante são mais fáceis de programar do que os de vírgula fixa, mas são também mais dispendiosos.

Contudo é sempre possível emular operações de vírgula flutuante em processadores de vírgula fixa através de rotinas especializadas, à custa de uns ciclos máquina extras.

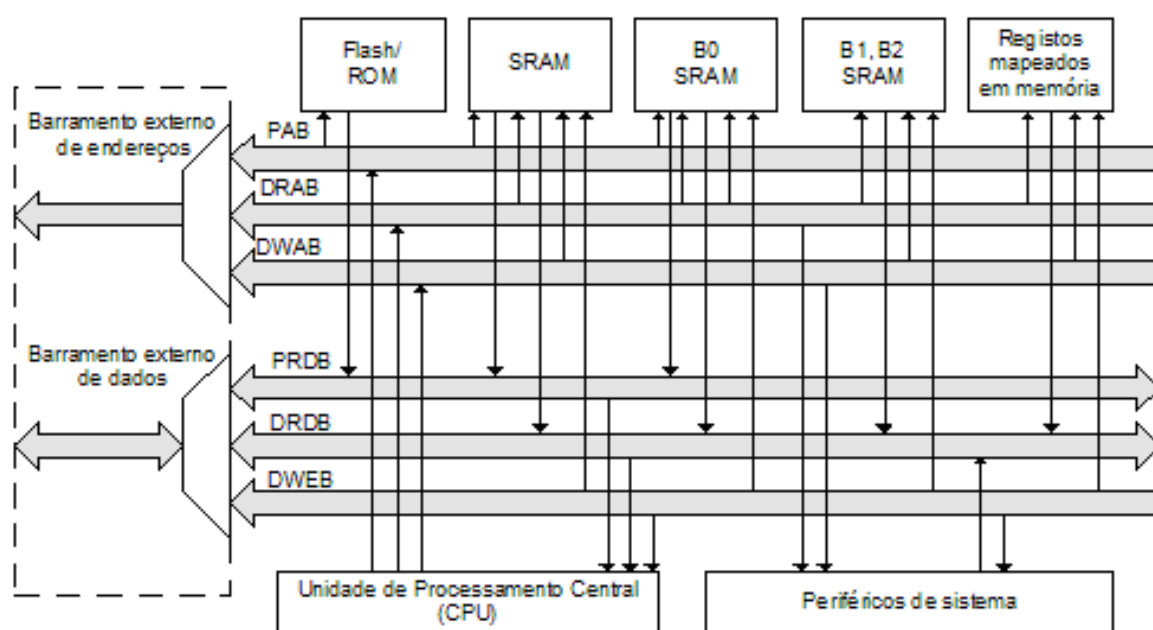


Fig. 1.21 – Barramento de informação do TMS320F240 da *Texas Instruments*.

O número de bits de um processador condiciona o seu preço. Normalmente os processadores de vírgula flutuante trabalham com dados de 32 bits, enquanto que os de vírgula fixa operam dados de 16 bits.

De observar, que o número de bits que compõe uma palavra de dados, não é obrigatoriamente igual, ao número de bits que compõe uma palavra de código.

É possível com máquinas de 16 bits realizar trabalho idêntico ao de máquinas com 32 bits, isto à custa de concatenar duas palavras, executando um número suplementar de ciclos de máquina para processar a informação.

O desempenho de um DSP é largamente condicionado pela estrutura hierárquica da memória.

Para otimizar o processamento dos dados, simultaneamente à leitura de uma instrução é feita a leitura dos operandos (Fig. 1.21).

Este requerimento é conseguido utilizando memórias múltiporto, memórias separadas para dados e programa, ou memória *cache* para as instruções.

Paralelamente, são proporcionados mecanismos que possibilitam o acesso a memórias mais lentas gerando ciclos de espera suplementares.

O *pipelining* das instruções consiste numa sequência de operações que ocorrem em paralelo durante a execução de uma instrução.

Um *pipeline* de quatro níveis pode, simultaneamente, realizar a operação de *leitura da instrução*, *descodificação da instrução*, *leitura de operandos*, e *execução da instrução*.

Como os quatros níveis são independentes, estas operações podem ocorrer simultaneamente.

Esta característica permite que a CPU possa ir buscar o código da instrução enquanto simultaneamente carrega os operandos e/ou guarda, em memória, o resultado de uma operação anterior.

Este tipo de operação está sujeito a algumas restrições e regras, por exemplo, só ser suportada por algumas instruções, ou um dos endereços de memória residir dentro do integrado. Esta funcionalidade é suportada pela existência de diferentes barramentos, memória de duplo-acesso, e bancos de memória independentes.

Uma funcionalidade comum a todos os DSP é a operação de multiplicação e acumulação (MAC), suportada por uma única instrução e tornada possível porque o multiplicador e o acumulador estão intercalados no mesmo caminho da informação como se pode ver na Fig. 1.22.

O acumulador é formado por um número de bits superior ao barramento de dados, para que o DSP possa suportar a execução sequencial de operações de multiplicação e acumulação, sem que se verifiquem situações de *overflow*.

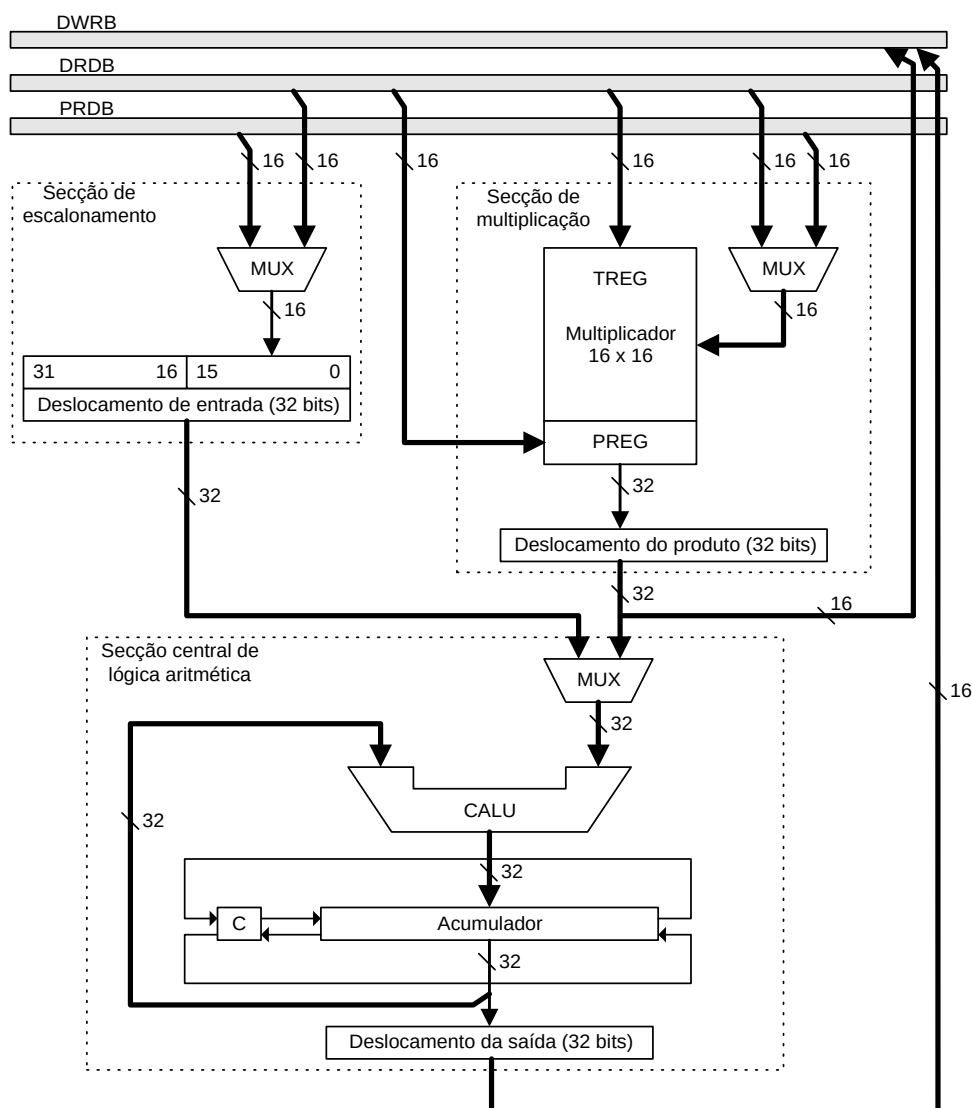


Fig. 1.22 – Unidade de processamento do TMS320F240 da Texas Instruments.

A existência de uma, ou mais, unidades de gestão de endereços (Fig. 1.23) facilita o processamento aritmético dentro do DSP, tornando-o rápido e eficiente. Uma vez gerado o endereço, toda a sua gestão é realizada separada do restante processamento.

Os modos de endereçamento ajustam-se às necessidades do processamento digital de sinal, sendo o mais comum o indirecto, suportados por registos auxiliares, com incremento ou decremento, ideal em situações onde o cálculo intensivo com dados colocados em endereços de memória consecutivos se justifica.

Outro modo de endereçamento característico, denominado por *bit-reversed*, facilita o trabalho de algoritmos como o da FFT (*Fast Fourier Transform*).

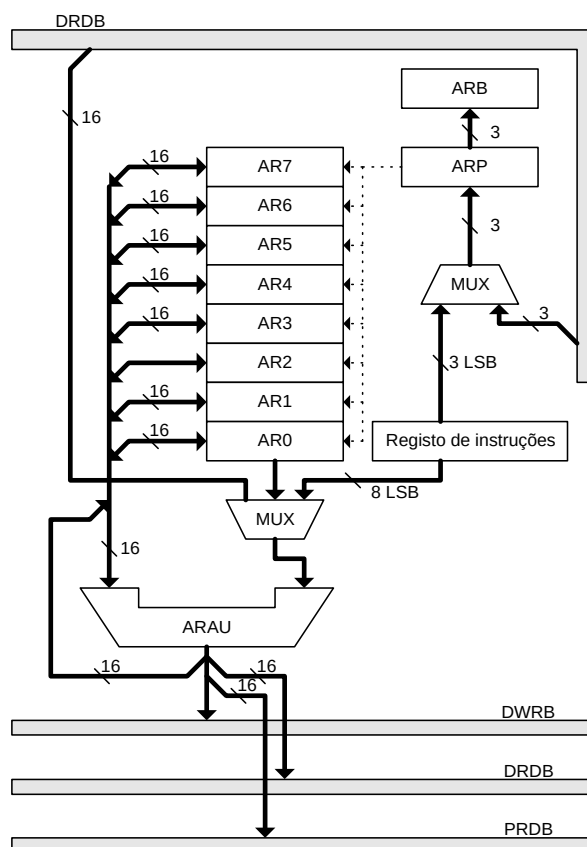


Fig. 1.23 – Unidade de registos auxiliares do TMS320F240 da *Texas Instruments*.

Muitos algoritmos de processamento de sinal são iterativos, para otimizar este aspecto, os DSP incorporam instruções que possibilitam a execução repetida de pedaços de programa em ciclo fechado, sem necessidade de ocupar ciclos de máquina adicionais a actualizar e testar o contador de iterações.

Finalmente, para permitir o acesso à máquina, paralelamente ao processamento de dados, a baixo custo e com elevada performance, é comum os DSP possuírem um, ou mais, portos série ou paralelo de I/O com mecanismos que englobam funcionalidades como sejam interrupções de alta prioridade ou o acesso directo à memória.

De especial importância é também o consumo energético do processador, para o reduzir e permitir a sua utilização em aplicações portáteis de consumo reduzido, é normal existirem dispositivos a funcionarem a tensão reduzida (3.3.V ou 3V), com modos de adormecimento, frequências de relógio programáveis, e desactivação de periféricos quando não estão a ser usados.

