

Armazenamento não volátil

- Armazenamento não volátil em sistemas
 - ROM: armazenamento de dados para leitura (read only)
 - NVRAM: armazenamento de dados para leitura/escrita (Non Volatile)
 - Memórias flash são exemplos típicos
- NVRAM são usados normalmente para armazenar:
 - Boot loader
 - Imagem do sistema operacional
 - Aplicações e bibliotecas
 - Arquivos e dados de configuração

Memórias flash

- Organizadas em regiões de 32, 64 ou 128 KB (*eraseblocks*)
- Possuem três operações básicas: *read*, *write* e *erase*
 - Para escrever é necessário apagar
- O procedimento de *erase* apresenta alguns detalhes:
 - Executado sobre um *eraseblock* inteiro (não dá para apagar bytes!!)
 - Um *eraseblock* tem um número limitado de apagamentos (vida útil)
 - *Wear leveling*: escrever uniformemente na *flash* para aumentar a vida útil
- Construídas de duas formas diferentes (NOR e NAND)

Flash NOR versus Flash NAND

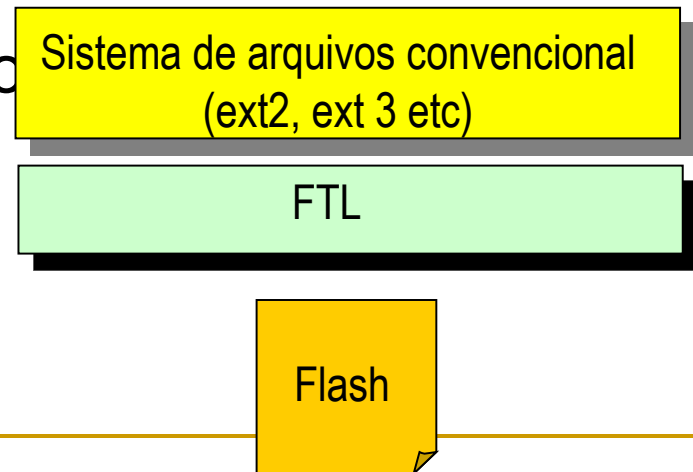
	NOR	NAND
Acesso a dados	Acesso randômico a dados	Área é dividida em blocos e estes em páginas. O acesso é a páginas
Operações	<i>Read, write e erase (eraseblocks)</i>	<i>Read e write (páginas) e erase (blocos)</i>
Interface c/ sistema	Barramento de dados e endereços	Diferentes formas. Normalmente através de um barramento de 8 bits (dados/comandos)
Código para execução	Permite execução do código diretamente armazenado	Necessário copiar código para memória
Desempenho	Leitura rápida, escrita e apagamento lentos	Leitura, escrita e apagamento rápidos
<i>Bad blocks</i>	Normalmente não tem esse problema porque foram projetadas para armazenar dados de sistema.	Dispositivos de armazenamento de baixo custo. Possui bad blocks, mas eles são marcados e não usados.
Utilização	Execução de código (<i>boot loader</i>)	Armazenamento de dados (mp3, set-top boxes, etc).

Memória *flash* e sistemas de arquivos

- Sistemas de arquivos convencionais não são adequados:
 - O apagamento é feito sobre blocos grandes ao passo que dispositivos orientados a disco os blocos são pequenos
 - Usam sistema de cache para desempenho
 - Provoca inconsistências do sistema de arquivos em caso de corte de alimentação
 - Convencionais: *sync* periódicos e saída via *shutdown* tratam disso
 - Embarcados podem ficar sem energia repentinamente (bateria)
- Vantagem dos sistemas de arquivos convencionais
 - Existem, tem um conjunto apropriado de primitivas e vasto suporte do sistema operacional

Flash Translation Layer (FTL / NFLT)

- Camada de software que emula um dispositivo de bloco para flash
 - ❑ Objetivo é usar toda a infra-estrutura já existente de sistemas de arquivos
- Desvantagem:
 - ❑ Não trata das inconveniências dos sistemas de arquivos tradicionais
 - ❑ Desempenho
 - ❑ Trata uma flash como se fosse orientado a bloco



FLT – Memórias NOR
NFLT – Memórias NAND

File System nativo para flash memories

■ JFFS2

- ❑ Não usa translation layer
- ❑ Implementa um filesystem com log estruturado (journalling)
 - Confiabilidade em falta de energia
- ❑ Utiliza wear-level (necessidade da flash)
- ❑ Utilizado em NOR flash, compact-flash e cartões com interface IDE
 - Não utilizada em NAND flash diretamente (em desenvolvimento)
- ❑ Brinde: Compressão dos dados

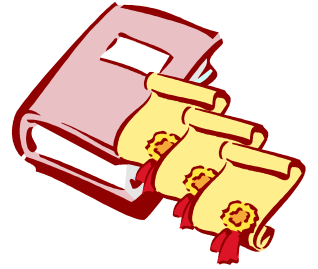
Memory Technology Device (MTD)

- Solução software livre para tratar dispositivos NVRAM (flash)
- Integrado no núcleo do linux a partir da versão 2.4
 - <http://www.linux-mtd.infradead.org>
- Filosofia: tratar dispositivos de memória como tal e não como discos
- Composto por duas partes: drives e aplicativos
- Dispositivos MTD
 - Flash (NAND, NOR non CFI e NOR CFI compliant)
 - Common Flash Interface
 - Flash disks (ATA-based e linear)
 - Armazenamento de massa

Diferenças entre discos e dispositivos MTD

Discos (<i>block devices</i>)	MTD
Composto por setores	Composto por blocos (eraseblocks)
Setores são pequenos (512, 1024 bytes)	Erase blocks são grandes (32, 64, 128 KB)
Duas operações básicas: read e write	Três operações: read eraseblock, write erasabloc e erase eraseblock
Bad-sectors são remapeados e escondidos pelo hardware (LBA disk drives)	Bad-blocks devem ser tratados pelo software
Setores de disco tem um grande ciclo de vida	Eraseblocks tem ciclo de vida limitado a 10^4 ou 10^5 ciclos de apagamento

Bibliografia



- P. Ragavahn, A. Lad, S. Neelakandan. Embedded Linux System Design and Developement. Auerbach, 2006.
- Artigo jffs2