

Agenda

• Visão geral

• Programação Multicore

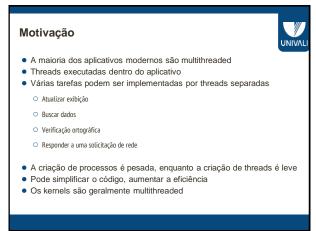
• Modelos Multithreading

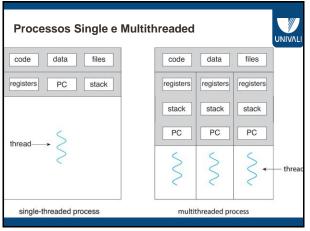
• Bibliotecas de threads

• Threading - automatização

• Problemas de encadeamento

• Exemplos de sistemas operacionais





Benefícios



- Responsividade pode permitir a execução contínua se parte do processo estiver bloqueada, especialmente importante para interfaces de usuário
- Compartilhamento de recursos threads compartilham recursos de processo, mais fácil do que a memória compartilhada ou a passagem de mensagens
- Economia mais barato do que a criação de processos, a alternância de threads é mais baixa do que a alternância de contexto
- Escalabilidade o processo pode tirar proveito de arquiteturas multicore

Programação Multicore



- Sistemas Multicore ou Multiprocessador colocando pressão sobre os programadores com desafios que incluem:
 - Dividindo atividades, balanceamento de carga, divisão de dado, dependência de dados e testar e depurar
- Paralelismo implica que um sistema pode executar mais de uma tarefa simultaneamente
 - O Dados e instrução são diferentes

6

- Concorrência oferece suporte a mais de uma tarefa que está ainda executando
 - O Processador de um único núcleo sofre concorrência por meio do escalanodor

5

Concorrência vs. paralelismo Execução simultânea em sistema single-core: T₁ T₁ T_4 T_3 T_4 ingle core T₃ T_2 Paralelismo em um sistema multi-core: T₁ T₁ core 1 T_3 T_2 T_2

time

Programação Multicore data • Tipos de paralelismo data O Paralelismo de dados: distribui core 3 subconjuntos dos mesmos dados em vários data núcleos, mesma operação em cada um task O Paralelismo de parallelism tarefas: distribuindo threads entre núcleos, cada thread executando uma operação exclusiva

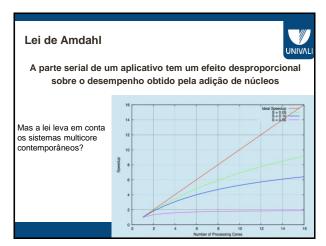
Lei de Amdahl

UNIVALI

- Identifica ganhos de desempenho com a adição de núcleos adicionais a uma aplicação que possui componentes seriais e paralelos
 - O S é a porção serial
 - O N é número de núcleos de processamento

 $speedup \le \frac{1}{S + \frac{(1 - S)}{N}}$

- Ou seja, se a aplicação é 75% paralela / 25% serial, passar de 1 para 2 núcleos resulta em speedup de 1,6 vezes
- À medida que N se aproxima do infinito, a aceleração se aproxima de 1 / S

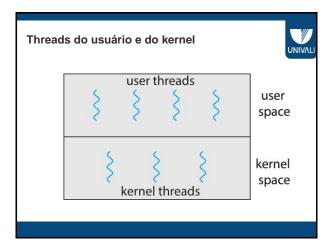


9 10

Threads de usuário e threads de kernel



- Threads de usuário gerenciamento feito pela biblioteca de threads em nível de usuário
- Três bibliotecas de threads principais:
 - O POSIX Pthreads iremos enteder essa
 - $\, \bigcirc \,$ Windows threads pesquisar no material ou internet
 - O Java threads pesquisar no material ou internet (se ficar lento, já sabe)
- Threads do kernel suportado pelo Kernel
- Exemplos praticamente todos os sistemas operacionais de uso geral utilizam:
 - O Windows, Linux, Mac OS X, iOS e Android



Bibliotecas de threads



- Biblioteca de threads fornece ao programador uma API para criar e gerenciar threads
- Duas formas principais de implementar
 - O Biblioteca inteiramente no espaço do usuário
 - O Biblioteca no nível do kernel suportada pelo SO

Pthreads



- Pode ser fornecido como nível de usuário ou nível de kernel
- Uma API padrão POSIX (IEEE 1003.1c) para criação e sincronização de threads
- Especificação, não implementação
- API especifica o comportamento da biblioteca de threads, a implementação depende do desenvolvimento da biblioteca
- Comum em sistemas operacionais UNIX (Linux e Mac OS X)
 - O Aceita no Windows via IDE CodeBLocks, CLion, Qt

13 14

#include #include #include #include #include #include <stdlib.h> #include <stdlib.h> #include <stdlib.h> #int sum; /* this data is shared by the thread(s) */ #int main(int argc, char *argv[]) ##int main(int argc, c

Código Pthreads para unir 10 threads



#define NUM_THREADS 10

/* an array of threads to be joined upon */
pthread_t workers[NUM_THREADS];

for (int i = 0; i < NUM_THREADS; i++)
 pthread_join(workers[i], NULL);</pre>

15 16

Threading implícita



- Com o número de threads aumenta, a manutenção do programa é mais difícil com threads explícitas
- Criação e gerenciamento de threads feito por compiladores e bibliotecas de tempo de execução em vez de programadores
- Cinco métodos explorados
 - O Thread Pools (falaremos)
 - O Fork-Join (pesquisar)
 - OpenMP (falaremos)
 - O Grand Central Dispatch (pesquisar)
 - O Intel Threading Building Blocks (pesquisar)

Pools de threads



- Criar vários threads em um pool onde eles aguardam o trabalho
- Vantagens:
 - O Normalmente, um pouco mais rápido para atender a uma solicitação com uma thread existente do que criar um nova thread
 - O Permite que o número de threads no(s) aplicativo(s) seja vinculado ao tamanho do pool
 - O Separar a tarefa a ser executada da mecânica de criação da tarefa permite diferentes estratégias para executar a tarefa
 - ou seja, as tarefas podem ser agendadas para serem executadas periodicamente
- A API do Windows oferece suporte a pools de threads:

```
DWORD WINAPI PoolFunction(AVOID Param) {
   /\star \star this function runs as a separate thread.
```

17 18

Java Thread Pools



- Três métodos para criar pools de threads na classe Executors:
 - static ExecutorService newSingleThreadExecutor() • static ExecutorService newFixedThreadPool(int size)
 - static ExecutorService newCachedThreadPool()

import java.util.concurrent.*; public class ThreadPoolExample public static void main(String[] args) {
 int numTasks = Integer.parseInt(args[0].trim());

/* Create the thread pool */
ExecutorService pool = Executors.newCachedThreadPool(); /* Run each task using a thread in the pool */
for (int i = 0; i < numTasks; i++)
 pool.execute(new Task());</pre>

/* Shut down the pool once all threads have completed */ pool.shutdown();

OpenMP



- Conjunto de diretivas do compilador e uma API para C, C++, FORTRAN
- Fornece suporte para programação paralela em ambientes de memória compartilhada
- Identifica regiões paralelas blocos de código que podem ser executados em paralelo

#pragma omp parallel Crie tantos threads quanto houver núcleos

#include <omp.h> #include <stdio.h> int main(int argc, char *argv[])

/* sequential code */ #pragma omp parallel

printf("I am a parallel region.");

/* sequential code */

return 0;

19 20

OpenMP



• Execute o loop for em paralelo

```
#pragma omp parallel for
for (i = 0; i < N; i++) {
   c[i] = a[i] + b[i];
}</pre>
```

Problemas com Threads



- Semântica das chamadas do sistema fork() e exec()
- Manipulação de sinais
 - O Síncrono e assíncrono
- Cancelamento de thread de destino
- Armazenamento local de thread
- Ativações do Escalonodar

21 22

Cancelamento de Thread



- Encerrando um thread antes que ele tenha terminado
- O thread a ser cancelado é o thread de destino
- Duas abordagens gerais:
 - O cancelamento assincrono encerra o thread de destino imediatamente
 - O cancelamento adiado permite que o thread de destino verifique periodicamente se deve ser
- Código Pthread para criar e cancelar um thread:

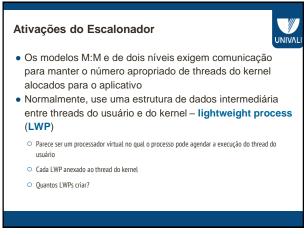
- pthread_t tid;
- /* create the thread */
 pthread_create(&tid, 0, worker, NULL);
- . . .
- /* cancel the thread */
 pthread_cancel(tid);
- /* wait for the thread to terminate */
 pthread_join(tid,NULL);

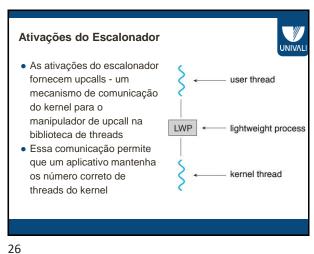
Armazenamento local de thread



- Thread-local storage (TLS) permite que cada thread tenha sua própria cópia dos dados
- Útil quando você não tem controle sobre o processo de criação de threads (ou seja, ao usar um pool de threads)
- Diferente das variáveis locais
 - O Variáveis locais visíveis apenas durante a invocação de uma única função
 - O TLS visível em invocações de função
- Semelhante a dados estáticos
 - O TLS é exclusivo para cada thread

23 24





Universidade do Vale do Italjal
Escola Politécnica
Engerharia de Computação

Threads e Concorrência