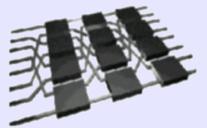
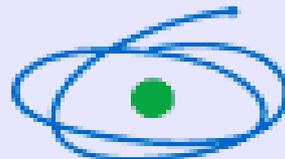

A Centralized and On-line Scheduling Solution to Dynamic MPI Programs

Márcia Cristina Cera
Nicolas Maillard
Philippe O. A. Navaux

WSPPD 2007 – 10 de agosto 2007



GPPD
Grupo de Processamento
Paralelo e Distribuído



C A P E S



Motivação

- MPI (Message Passing Interface)
 - MPI-1 (1994)
 - MPI-2 (1998)
 - Criação dinâmica de processos (MPI_Comm_spawn)
 - Acesso Remoto a Memória (RMA)
 - Entrada e saída de dados paralela
- Distribuições MPI que implementam a norma MPI-2
 - LAM/MPI, MPICH-2, OpenMPI, HP-MPI,...
- Modelo de programa: SPMD -> MPMD
 - Pode-se implementar aplicações dinâmicas
 - Uso de recursos dinâmicos
- Como tirar proveito dessas novas funcionalidades?

Contexto do Trabalho

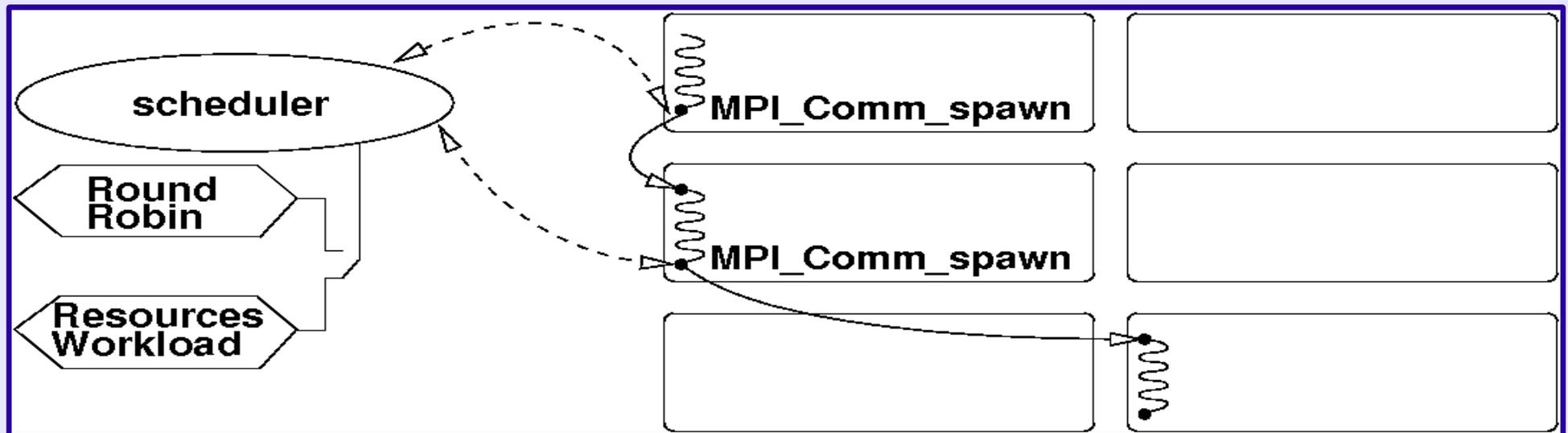
- Escalonar aplicações dinâmicas
 - **Distribuição eficiente** dos processos leva a melhora no desempenho
- Uso de **recursos dinâmicos**
 - Tarefas de aplicações dinâmicas podem ser lançadas em recursos que acabam de tornar-se disponíveis
- Objetivo:
 - Apresentar uma solução para que se possa **distribuir processos criados dinamicamente**
 - Usar a criação dinâmica de processos para destinar trabalho a **recursos dinâmicos**

Sumário

- Motivação e contexto
- Escalonador de processos dinâmicos
- Resultados experimentais
 - Sobrecusto da criação dinâmica de processos
 - Uso de recursos dinâmicos
 - Impacto das políticas de escalonamento disponibilizadas
- Conclusões e Perspectivas Futuras

Escalonador de processos dinâmicos

- Centralizado – decisões tomadas *on-line*
- Duas políticas:
 - Round_Robin
 - Baseado na carga
- Integrado na LAM/MPI 7.1.2

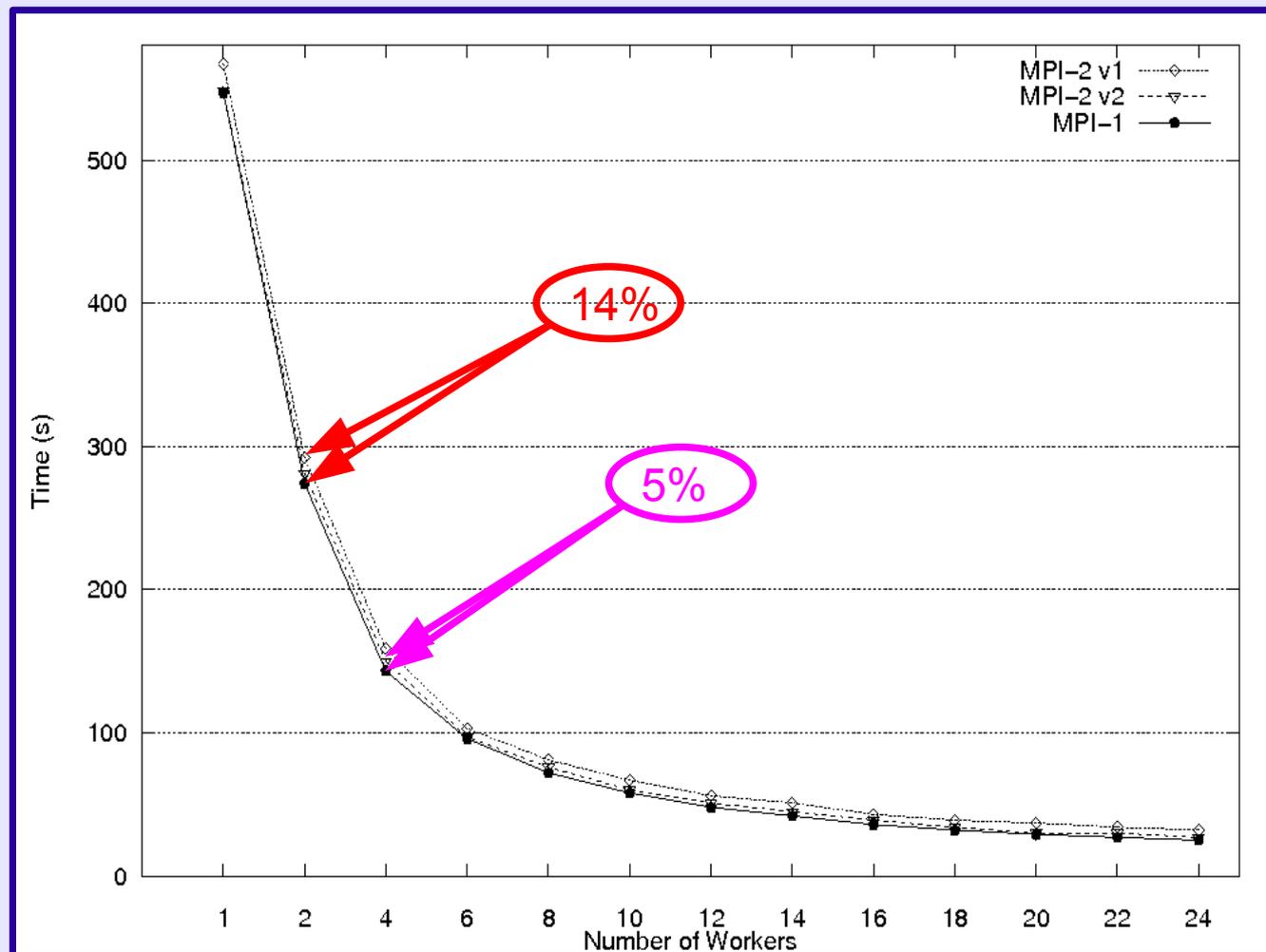


Resultados Experimentais

- Metas
 - Sobrecusto da criação dinâmica de processos
 - Comparação entre programas MPI-1 e MPI-2
 - Aplicação: Geração do fractal de Mandelbrot
 - Uso de recursos dinâmicos
 - Impacto das políticas de escalonamento disponibilizadas

Sobrecusto da criação dinâmica

- MPI-2 v1 com criação de processos a cada tarefa
- MPI-2 v2 com criação de processos por processador

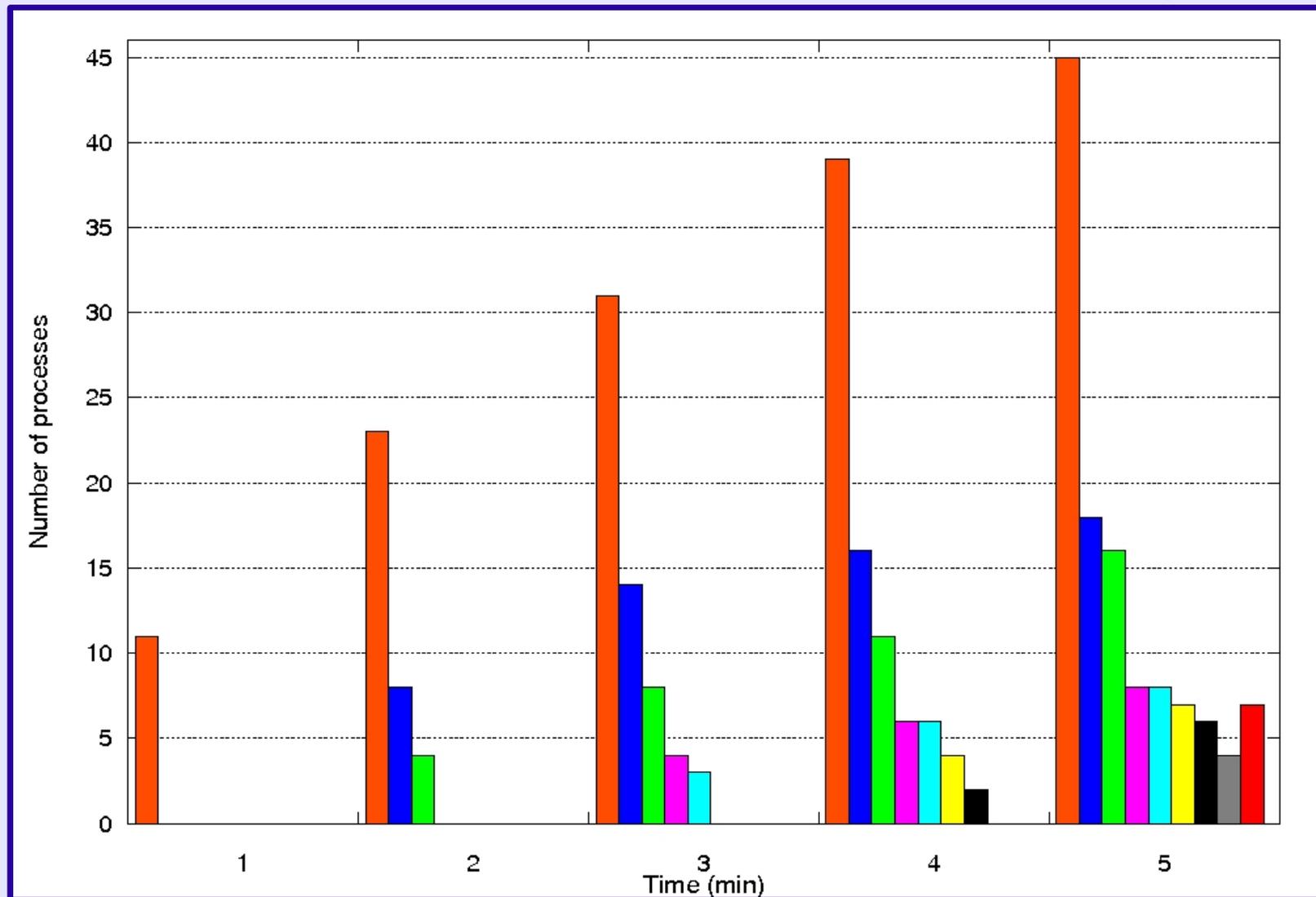


Resultados Experimentais

- **Metas**
 - Sobrecusto da criação dinâmica de processos
 - Comparação entre programas MPI-1 e MPI-2
 - Aplicação: Geração do fractal de Mandelbrot
 - **Uso de recursos dinâmicos**
 - Através da criação dinâmica de processos
 - LAM/MPI: oferece `lamgrow` para adicionar recursos
 - Aplicação: *N*-Queens -> posicionar *N* rainhas em um tabuleiro *N*×*N*
 - Impacto das políticas de escalonamento disponibilizadas

Uso de Recursos Dinâmicos

- *N*-Queens em tabuleiro 18x18 usando 5 máquinas

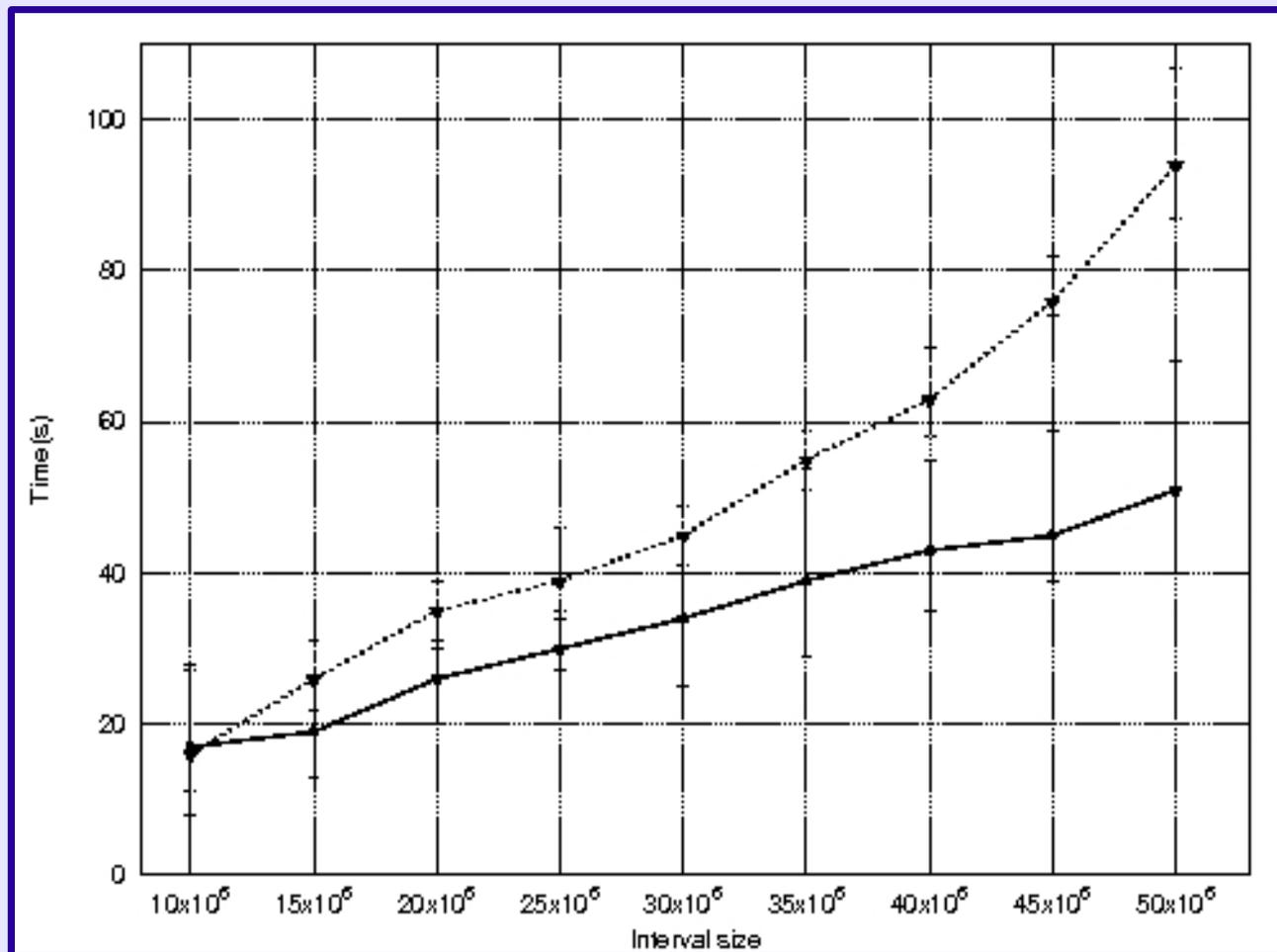


Resultados Experimentais

- Metas
 - Sobrecusto da criação dinâmica de processos
 - Comparação entre programas MPI-1 e MPI-2
 - Aplicação: Geração do fractal de Mandelbrot
 - Uso de recursos dinâmicos
 - Através da criação dinâmica de processos
 - LAM/MPI: oferece `lamgrow` para adicionar recursos
 - Aplicação: *N*-Queens -> posicionar *N* rainhas em um tabuleiro *N*×*N*
 - Impacto das políticas de escalonamento disponibilizadas
 - Round-Robin: padrão
 - Baseado na carga: eficiente em ambientes heterogêneos ou para aplicações irregulares
 - Aplicação: Computação de números primos

Impacto das políticas de escalonamento

- Round-Robin: pontilhado
- Baseado na carga: linha



Conclusão

- Criação dinâmica de processos do MPI-2 é uma ferramenta promissora desde de que bem empregada
 - Desenvolvimento de **aplicações dinâmicas**
 - Uso de **recursos dinâmicos**
- Restrição na comunicação de processos criados dinamicamente
 - **Pai <-> filhos**; não existe mais um comunicador global
- Como tratar a eliminação de nós
 - Precisa de **Tolerância a Falhas** para manter a integridade das aplicações

Perspectivas Futuras

- Modelos de programa dinâmico para MPI-2
 - Divisão e Conquista (D&C)
 - outros?
- Estratégia de **escalonamento**
 - Work Stealing
 - Hierarchical Work Stealing
 - Work Stealing Genérico?
- Escalonador de programas MPI-2 associados a escalonadores de *batch*
 - OAR?

Obrigada....

