

Aluno: **Lesandro Ponciano dos Santos**
Curso: **Sistemas de Informação - PUC Minas em Guanhães**

Orientador: **Prof. João Paulo D. Silva**
Curso: **Sistemas de Informação - PUC Minas em Guanhães**

Co-orientador: **Prof. Luís Fabrício W. Góes**
Curso: **Sistemas de Informação - PUC Minas em Guanhães**

CONTEXTUALIZAÇÃO

RGSA

- Aplicações que exigem alto poder de processamento;
- *Reconfigurable Gang Scheduling Algorithm* (RGSA) (Góes e Martins, 2005);
- Algoritmo reconfigurável que utiliza informações do tempo de submissão, tempo de execução e número de processos das tarefas mais executadas em uma arquitetura;
- Caracterização de Cargas de Trabalho;
- Nos testes realizados com este escalonador as tarefas são classificadas, através do Algoritmo 1, em baixo (L - *low*) e alto (H - *high*), 4 grupos possíveis (HH, HL, LL, LH).

Algoritmo 1 - Classificação Low e high

```

se job.n_processos ≤ mediana_n_processos
se job.tempo_exe ≤ mediana_tempo_exe
então job.classe = LL;
senão job.classe = LH;

senão se job.tempo_exe ≤ mediana_tempo_exe
então job.classe = HL;
senão job.classe = HH;

```

Problemas

- O algoritmo de Classificação *Low e high* é muito sensível à variações da mediana.
- No processo, não há a preocupação em gerar grupos com baixo desvio entre as tarefas que o compõe.

Proposta

- Utilização de algoritmos clássicos de Agrupamento, por similaridade, para gerar os grupos e de Classificação para identificar as regras do agrupamento e subsidiar a predição de novas tarefas.

Objetivo

- Analisar a aplicação dos algoritmos, de agrupamento, *k-means* e, de classificação, *j48* para obter modelos de similaridade de comportamento entre tarefas paralelas.

MODELAGEM E PREDIÇÃO

Classificação Low e High

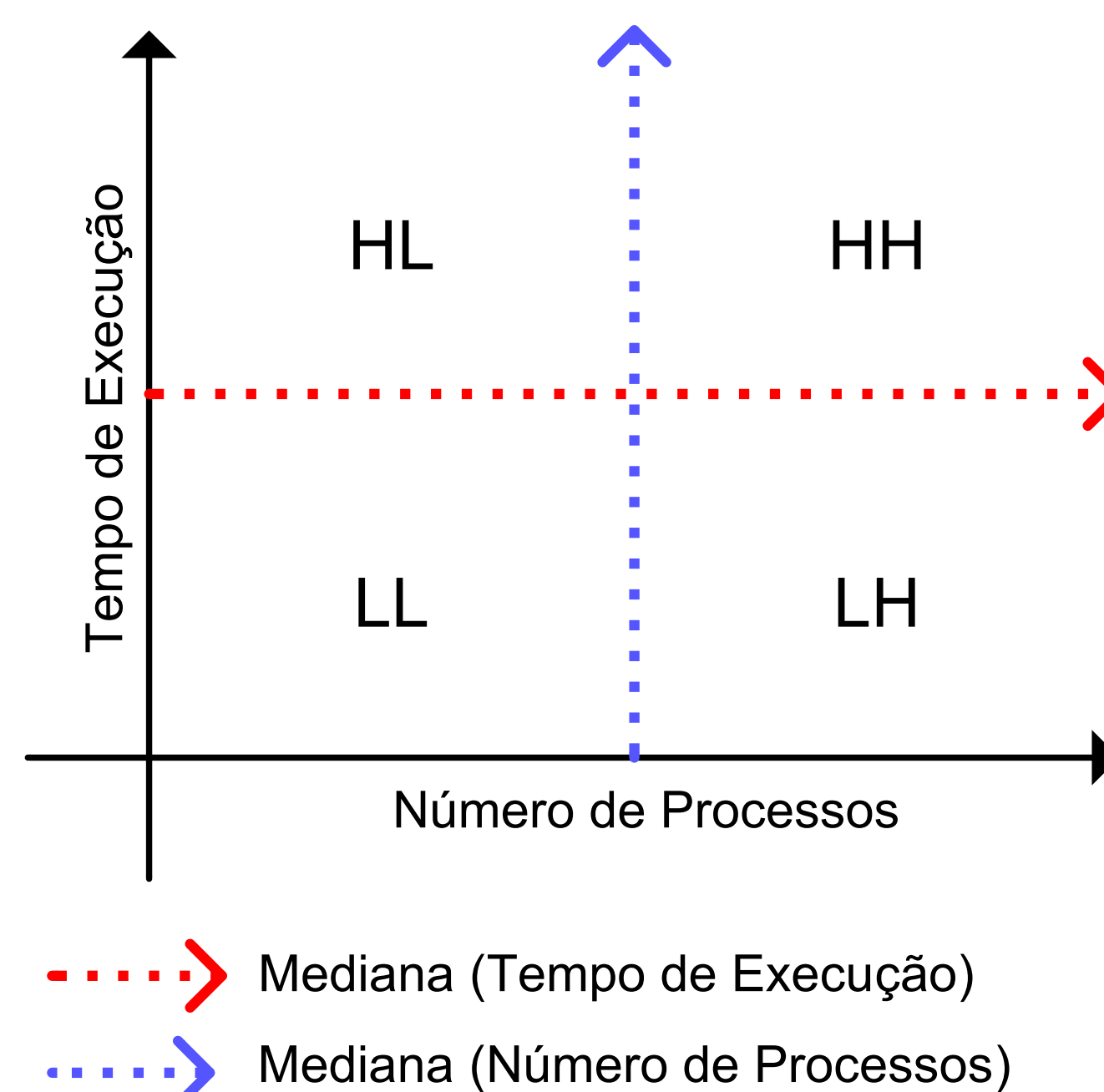


Figura 1. Classificação low e High

Agrupamento K-means e Classificação J48

- Dado um rastro, com dados de tarefas paralelas, executa-se o algoritmo *k-means* com $k=4$ e identificam-se quatro grupos (G0, G1, G2 e G3), gerados pela similaridade entre as tarefas.
- Executa-se o algoritmo *j48* para identificar as regras que definem os agrupamentos.
- As regras obtidas são utilizadas na classificação e predição do comportamento de tarefas futuras.

Algoritmo 2 - Exemplo de Regras de classificação (LANCM5 outubro/1995)

```

se job.n_processos ≤ 64
então job.classe = G1;
senão
se job.n_processos ≤ 256
então job.classe = G0;
senão
se job.tempo_exe ≤ 8357
então job.classe = G2;
senão job.classe = G3;

```

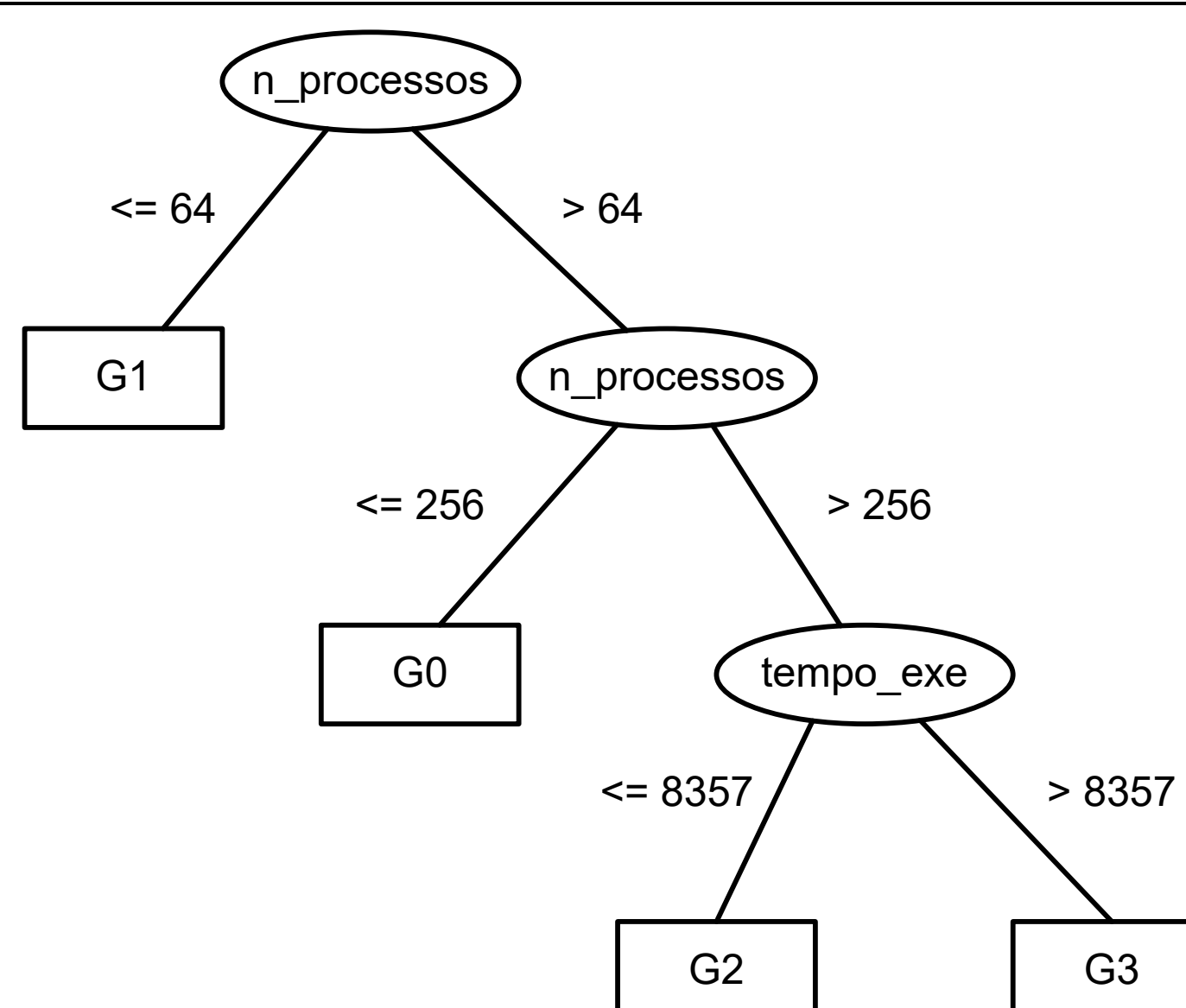


Figura 2. Árvore de decisão j48

ESTUDO DE CASO

Planejamento dos Experimentos

- São utilizadas tarefas submetidas, ao longo de sete meses, em quatro supercomputadores reais (HPC2N, SDSC, SHARCNET, LANCM5);
- Para predição de comportamentos, caracteriza-se um mês e testa-se a capacidade de predição para o mês subsequente;

Resultados Parciais

Tabela 1. Rastros, Períodos e Número de Tarefas

	Número de Tarefas						
	Abr.	Mai.	Jun.	Jul.	Ago.	Set.	Oct.
HPC2N	29970	17817	20545	24121	7241	5159	11444
SDSC	5259	8395	8358	5339	7204	8782	9147
SHARCNE.	87505	76813	52551	95118	90225	63162	158681
LANCM5	5790	6831	6216	6969	6126	6303	5271

Tabela 2. Acerto na predição Low e High

Meses	Dias da Semana						
	D	S	T	Q	Q	S	S
mai.	73	56	43	65	63	58	65
jun.	77	62	71	71	71	66	71
jul.	46	57	53	56	69	49	55
ago.	55	51	58	57	52	49	58
set.	65	61	63	57	56	66	60
out.	50	63	76	74	67	61	64

Tabela 3. Acerto na predição K-means e J48

Meses	Dias da Semana						
	D	S	T	Q	Q	S	S
mai.	76	84	70	82	76	82	82
jun.	76	86	90	92	94	94	90
jul.	66	76	76	80	80	82	74
ago.	76	84	82	66	70	70	70
set.	90	86	86	68	70	76	74
out.	78	90	82	92	90	88	86

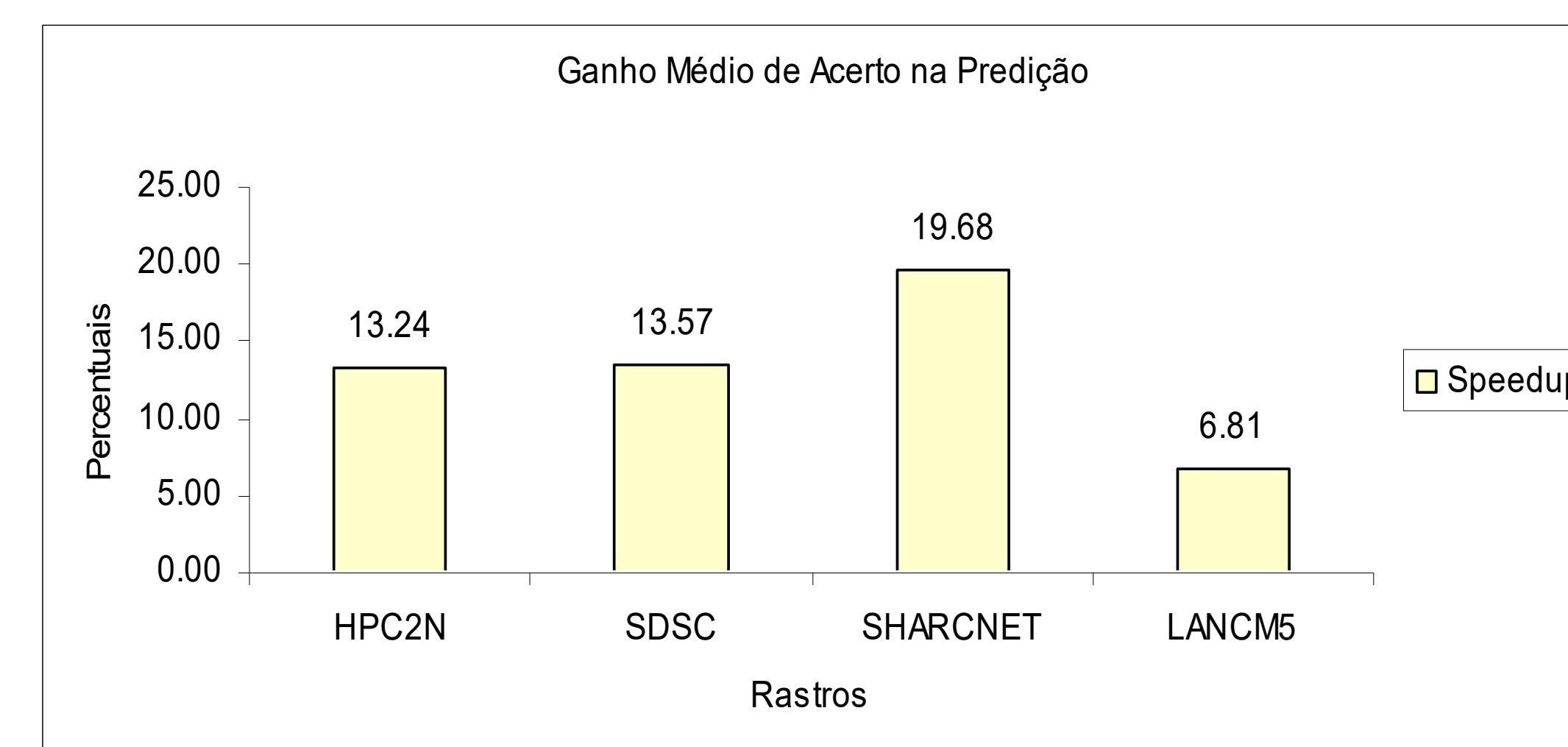


Gráfico 1. Speedup na predição (k-means j48 e low e high)

Principais conclusões

- Os algoritmos *K-means* e *j48* geram, na média, maior acerto na predição;
- A implementação do algoritmo *low e high* é mais simples;
- Os algoritmos *K-means* e *j48* extraem mais informação (regras), da base de dados, em relação ao algoritmo *low e high* (Mediana)