

# PROCESSO DE KDD PARA AUXÍLIO À RECONFIGURAÇÃO DE AMBIENTES VIRTUALIZADOS

Ana T. Winck  
Duncan D. Ruiz

Pontifícia Universidade Católica do Rio Grande do Sul -  
PUCRS

## Roteiro

2

- Descrição do Cenário
  - Realocação de Recursos
  - Processo de KDD proposto
- Fonte de Dados
  - Benchmarks
  - Estratégias de Execução
- Data Warehouse
  - Modelo Analítico
- Mineração de Dados
  - Preparação
  - Modelos Preditivos
- Apresentação e Efetivação da Reconfiguração
  - Modelos preditivos enriquecidos
- Conclusões
  - Trabalhos Futuros
- Referências

## Cenário

3

## Paravirtualizador Xen

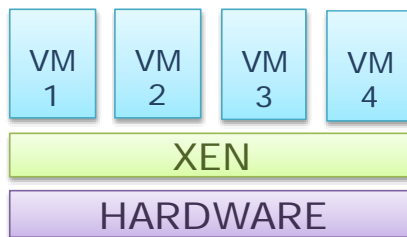


Figura 1 – Arquitetura do Xen. Adaptado de [BAR03]

4

# Realocação de Recursos

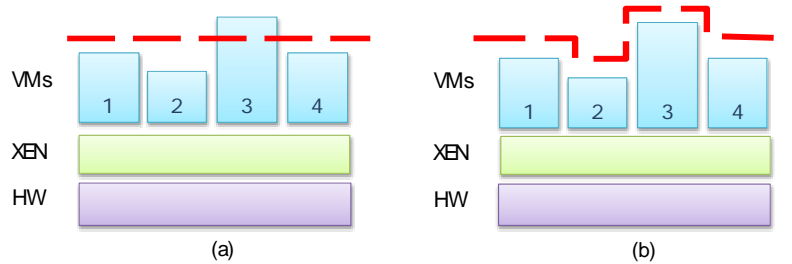
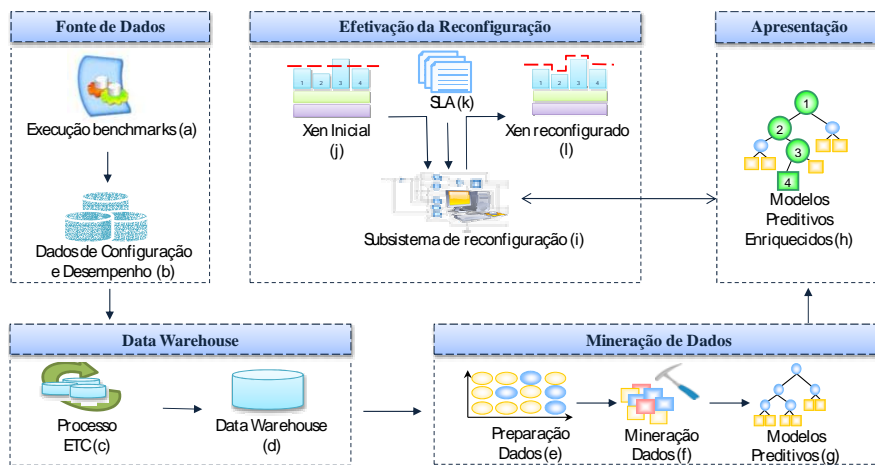


Figura 2 – Realocação de recursos no Xen

5

# Processo de KDD proposto



6

# Fonte de Dados

7

```
BYTE UNIX Benchmarks (Version 3.11)
System -- Linux vmlinuz 2.6.16.33-xen #7 SMP Thu Apr 19 11:34:18 AMT 2007
1686 GNU/Linux
Start Benchmark Run: Wed May 16 17:30:03 UTC 2007
1 interactive users.
Dhrystone 2 without register variables 334078.9 lps (10 secs, 6 samples)
Dhrystone 2 using register variables 343319.2 lps (10 secs, 6 samples)
Arithmetic Test (type = arithoh) 1271179.8 lps (10 secs, 6 samples)
Arithmetic Test (type = register) 56022.7 lps (10 secs, 6 samples)
Arithmetic Test (type = short) 46857.5 lps (10 secs, 6 samples)
Arithmetic Test (type = int) 55877.3 lps (10 secs, 6 samples)
Arithmetic Test (type = long) 55910.7 lps (10 secs, 6 samples)
Arithmetic Test (type = float) 54794.4 lps (10 secs, 6 samples)
Arithmetic Test (type = double) 55142.4 lps (10 secs, 6 samples)
System Call Overhead Test 35427.3 lps (10 secs, 6 samples)
Pipe Throughput Test 49001.1 lps (10 secs, 6 samples)
Pipe-based Context Switching Test 7945.3 lps (10 secs, 6 samples)
Process Creation Test 350.4 lps (10 secs, 6 samples)
Exec1 Throughput Test 150.8 lps (9 secs, 6 samples)
File Read (10 seconds) 146223.0 KBps (10 secs, 6 samples)
File Write (10 seconds) 39763.0 KBps (10 secs, 6 samples)
File Copy (10 seconds) 8379.0 KBps (10 secs, 6 samples)
File Read (30 seconds) 146950.0 KBps (30 secs, 6 samples)
File Write (30 seconds) 39265.0 KBps (30 secs, 6 samples)
File Copy (30 seconds) 6321.0 KBps (30 secs, 6 samples)
C Compiler Test 82.0 lpm (60 secs, 3 samples)
Shell scripts (1 concurrent) 172.1 lpm (60 secs, 3 samples)
Shell scripts (2 concurrent) 94.1 lpm (60 secs, 3 samples)
Shell scripts (4 concurrent) 50.9 lpm (60 secs, 3 samples)
Shell scripts (8 concurrent) 25.2 lpm (60 secs, 3 samples)
Dc: sqrt(2) to 99 decimal places 17499.9 lpm (60 secs, 6 samples)
Recursion Test--Tower of Hanoi 5555.6 lps (10 secs, 6 samples)

INDEX VALUES
TEST BASELINE RESULT INDEX
Arithmetic Test (type = double) 2541.7 55142.4 21.7
Dhrystone 2 without register variables 22366.3 334078.9 14.9
Exec1 Throughput Test 16.5 150.8 9.1
File Copy (30 seconds) 179.0 6321.0 35.3
Pipe-based Context Switching Test 1318.5 7945.3 6.0
Shell scripts (8 concurrent) 4.0 25.2 6.3
=====
SUM of 6 items 93.4
AVERAGE 15.6
```

8

Figura 4 – Exemplo Resultado Unixbench

## Estratégia de Execuções

- Distintas configurações (m diferentes situações)

- ▣ Configurações de Ambiente

- ▣ Total de VMs

- ▣ Limites de consumo de CPU

- 70%, 85%, 90% e 100%

- ▣ CAP

- Fatia do total de CPU para cada VM

- ▣ Memória

- Diferentes alocações para cada Combinação de CAP

Comb.	VM1	VM2	VM3	VM4
1	10	10	10	40
2	10	10	15	35
	VM1	VM2	VM3	VM4
Memória (MB)	70	70	70	70
	80	70	70	60
	90	70	70	50
	100	70	70	40
	110	70	60	40
	120	70	50	40
	130	70	40	40

9

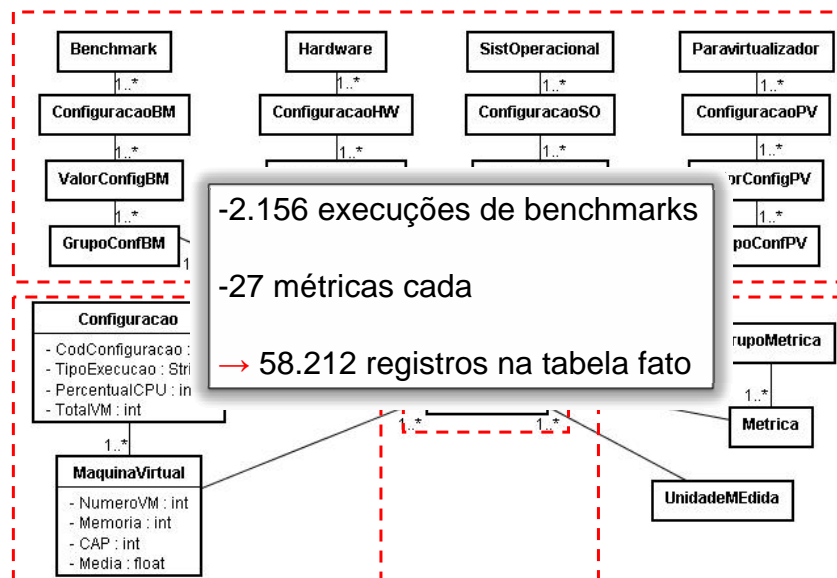
## Planejamento de Execuções

Ambiente	Escalonador: Credit / Xen: 3.0.4 / Máquina: Xeon / Kernel: 2.6.16.33 Memória Dispo.: 280 MB / Benchmark: Unixbench 3.11										
CPU	85%				CPU	85%					
VMs	VM1	VM2	VM3	VM4	VMs	VM1	VM2	VM3	VM4		
CAP	10	10	20	45	CAP	10	15	25	35		
	Memória (MB)					Memória (MB)					
Configuração	239	70	70	70	70	Configuração	281	70	70	70	70
	240	80	70	70	60		282	80	70	70	60
	241	90	70	70	50		283	90	70	70	50
	242	100	70	70	40		284	100	70	70	40
	243	110	70	60	40		285	110	70	60	40
	244	120	70	50	40		286	120	70	50	40
	245	130	70	40	40		187	130	70	40	40

10

# Data Warehouse

11



12

## Mineração de Dados

13

## Mineração de Dados

- Indica parâmetros para reconfiguração
- Tarefas preditivas
  - ▣ Classificação
    - Árvores de decisão (C4.5 [QUI96])
    - Weka
- Preparação dos dados para a mineração
  - ▣ [TAN06]

14

# Preparação Atributo Preditivo

- Mapeamento entre as configurações
  - ▣ Valores atributo média
  - ▣ Somatório das médias das VMs
  - ▣ Configuração inicial e configuração alvo
  - ▣ Critério de benefício
    - Diferença entre o somatório das configurações
  - ▣ Matriz Quadrada

15

# Preparação Atributos Explanatórios

16

CPU		85%			
VMs		VM1	VM2	VM3	VM4
CAP		10	10	20	45
		Memória (MB)			
Configuração	239	70	70	70	70
	240	80	70	70	60
	241	90	70	70	50
	242	100	70	70	40
	243	110	70	60	40
	244	120	70	50	40
	245	130	70	40	40

- ~~CPU~~
  - % CPU de origem
  - CAP
    - ▣ Primeira Configuração
  - Memória
    - ▣ Primeira Memória
- ~~Valores Ordenáveis~~

- Combinação de valores

```

@relation Arff_85

@attribute CAP_INICIAL {218, 225, 232, 239, 246, 253, 260, 267, 274, 281, 288,
295, 302, 309, 316, 323, 330, 337}
@attribute MEM_INICIAL {70, 80, 90, 100, 110, 120, 130}
@attribute CAP_ALVO {218, 225, 232, 239, 246, 253, 260, 267, 274, 281, 288, 295,
302, 309, 316, 323, 330, 337}
@attribute MEM_ALVO {70, 80, 90, 100, 110, 120, 130}
@attribute class {0, 1}

@data

281, 70, 239, 70, 0
281, 70, 239, 90, 0
281, 70, 239, 110, 1
281, 70, 239, 130, 0

281, 90, 239, 70, 0
281, 90, 239, 90, 0
281, 90, 239, 110, 1
281, 90, 239, 130, 0

281, 110, 239, 70, 0
281, 110, 239, 90, 0
281, 110, 239, 110, 1
281, 110, 239, 130, 0

281, 130, 239, 70, 0
281, 130, 239, 90, 0
281, 130, 239, 110, 1
281, 130, 239, 130, 0

```

17

Figura 7 – Exemplo de conteúdo do arquivo arff para a mineração

## Modelos Preditivos

```

CAP_ALVO = 281
| CAP_INICIAL = 239
| | MEM_INICIAL = 70: 1
| | MEM_INICIAL = 80: 1
| | MEM_INICIAL = 90: 1
| | MEM_INICIAL = 100: 1
| | MEM_INICIAL = 110: 0
| | MEM_INICIAL = 120: 1
| | MEM_INICIAL = 130: 1

```

```

CAP_ALVO = 246
| CAP_INICIAL = 253
| | MEM_ALVO = 70: 0
| | MEM_ALVO = 80: 0
| | MEM_ALVO = 90: 1
| | MEM_ALVO = 100: 0
| | MEM_ALVO = 110: 1
| | MEM_ALVO = 120: 0
| | MEM_ALVO = 130: 1

```

18

Figura 8 – Exemplo de modelos preditivos

## Apresentação e Efetivação da Reconfiguração

19

## Apresentação

- Protocolo de comunicação
  - ▣ Modelos preditivos
  - X
  - ▣ Subsistema de reconfiguração
  
- Modelos Preditivos Enriquecidos
  - ▣ Instruções em alto nível
  - ▣ Evento-Condição-Ação

20

# Modelos Preditivos Enriquecidos

CONFIGURAÇÃO INICIAL					CONFIGURAÇÃO ALVO								INSTRUÇÕES ENRIQUECIDAS										
					CAP				MEM				VM1		VM2		VM3		VM4				
	VM1	VM2	VM3	VM4	VM1	VM2	VM3	VM4	VM1	VM2	VM3	VM4	CAP	MEM	CAP	MEM	CAP	MEM	CAP	MEM			
CAP	10	10	20	45	10	15	25	35	70	70	70	70		70	15		25		35	70			
									80	70	70	60		80	15		25		35	60			
									90	70	70	50			15		25		35				
MEM	90	70	70	50	10	15	25	35	100	70	70	40		100	15		25		35	40			
									110	70	60	40		110	15		25	60	35	40			
									120	70	50	40		120	15		25	50	35	40			
									130	70	40	40		130	15		25	40	35	40			
CAP	10	15	20	40	10	10	25	40	90	70	70	50		70	10		25	70		50			
									110	70	60	40		110	10		25	60					
MEM	130	70	40	40					130	70	40	40			10		25	40					

21

## Conclusões

22

## Conclusões

- Processo completo de KDD
- Modelo analítico abrangente
- Descoberta de padrões para a reconfiguração
- Uso rigoroso de técnicas de preparação
- Modelos preditivos enriquecidos

23

## Trabalhos Futuros

- Data Warehousing
  - ▣ Planejamento com diferentes ambientes
  - ▣ Mais execuções de benchmarks
  - ▣ ETC com suporte computacional
- Descoberta de padrões
  - ▣ Explorar outros critérios de benefício
  - ▣ Inserir as políticas de SLA já na mineração

24

## Referências

- [BAR03] BARHAM, P., DRAGOVIC, B., FRASER, K., HAND, S., HARRIS, T., HO, A., NEUGEBAUER, R., PRATT, I., WARFIELD, A. 2003. Xen and the art of virtualization. In: ACM SYMPOSIUM ON OPERATING SYSTEMS PRINCIPLES (SOSP'03), 19., 2003, New York. **Proceedings...** New York: ACM Press, Out. 2003, p. 164-177.
- [QUI96] QUINLAN, J. R. **C4.5: programs for machine learning**. California: Morgan Kaufman, 1996. 302 p.
- [TAN06] TAN, P-N., STEINBACH, M., KUMAR, V. **Introduction to data mining**. Boston: Addison Wesley, 2006. 769 p.

25

# PROCESSO DE KDD PARA AUXÍLIO À RECONFIGURAÇÃO DE AMBIENTES VIRTUALIZADOS

Ana T. Winck  
Duncan D. Ruiz

Pontifícia Universidade Católica do Rio Grande do Sul -  
PUCRS