

Slurm Workload Manager

Samuel Praça de Paula
MAC5742 - Computação Paralela e Distribuída

Junho de 2015

SLURM: Simple Linux Utility for Resource Management

Escalonador de processos desenvolvido desde 2002.

SLURM: Simple Linux Utility for Resource Management

Escalonador de processos desenvolvido desde 2002.

Lawrence Livermore National Laboratory, Linux NetworX, HP,
Groupe Bull.

SLURM: Simple Linux Utility for Resource Management

Escalonador de processos desenvolvido desde 2002.

Lawrence Livermore National Laboratory, Linux NetworX, HP,
Groupe Bull.

Desde 2010: **SchedMD**;

SLURM: Simple Linux Utility for Resource Management

Escalonador de processos desenvolvido desde 2002.

Lawrence Livermore National Laboratory, Linux NetworX, HP,
Groupe Bull.

Desde 2010: **SchedMD**; CEA, Cray, Intel, NVIDIA...

Informações básicas

Linux, AIX, *BSD, MacOS, Solaris

Informações básicas

Linux, AIX, *BSD, MacOS, Solaris

GNU General Public License v2.0

Informações básicas

Linux, AIX, *BSD, MacOS, Solaris

GNU General Public License v2.0

500 mil linhas em linguagem C.

Informações básicas

Linux, AIX, *BSD, MacOS, Solaris

GNU General Public License v2.0

500 mil linhas em linguagem C.

Usado em 6 dos 10 primeiros computadores do TOP500.
(e.g. Tianhe-2, IBM Sequoia.)

Escalonamento de tarefas

O que é? Por que precisamos fazer?

Escalonamento de tarefas

O que é? Por que precisamos fazer?

Gerenciamento de tarefas batch.

Escalonamento de tarefas

O que é? Por que precisamos fazer?

Gerenciamento de tarefas batch.

Faz sentido principalmente quando há muitos recursos.

Slurm é também descrito como um “gerenciador de cluster”

(Obs.: MPI)

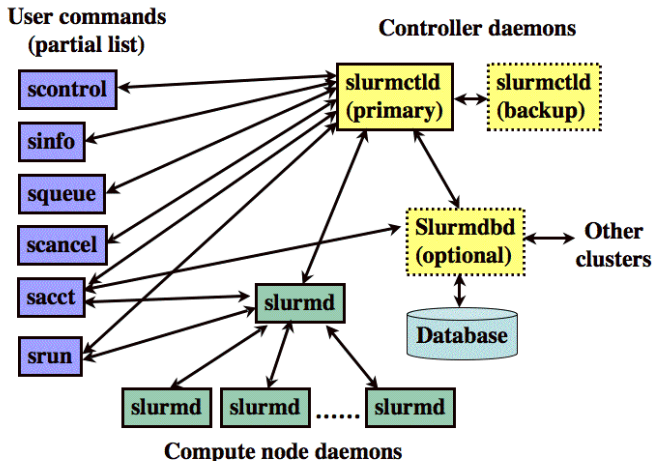
Escalonamento de tarefas

Que tipo de coisa permite fazer?

Que tipo de coisa permite fazer?

- Configurar recursos
- Estabelecer relações de precedência
- Prioridade de tarefas
- etc...

Uma ideia da arquitetura do Slurm...



Uma ideia do tipo de algoritmo usado

Alocação de processadores por localidade.

Uma ideia do tipo de algoritmo usado

Alocação de processadores por localidade.

Curvas de Hilbert para transformar em Bin Packing unidimensional.

Curva de Hilbert

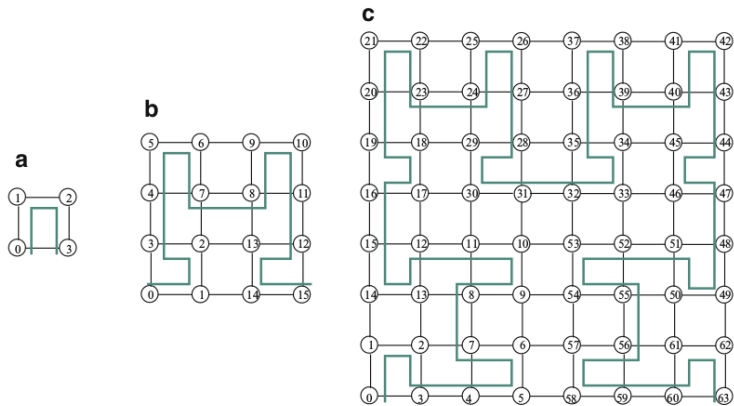


Fig. 5.26 Example of applying Hilbert space-filling curve to number the processors in a 2D-mesh. The solid line shows the fitting of the Hilbert curve. (a) First, (b) second, and (c) third iteration

Bin Packing

Dada a capacidade V que minhas “bins” têm, e cada item j tem tamanho p_j , quero minimizar uso de bins que comportam todos os itens.

Bin Packing

Dada a capacidade V que minhas “bins” têm, e cada item j tem tamanho p_j , quero minimizar uso de bins que comportam todos os itens.

NP-difícil! Mas há várias aproximações fáceis de executar.

First fit.

Bin Packing

Dada a capacidade V que minhas “bins” têm, e cada item j tem tamanho p_j , quero minimizar uso de bins que comportam todos os itens.

NP-difícil! Mas há várias aproximações fáceis de executar.

First fit. $11/9 \text{ OPT} + 1$

Quanto vale a pena gastar com o escalonamento?

HOW LONG CAN YOU WORK ON MAKING A ROUTINE TASK MORE EFFICIENT BEFORE YOU'RE SPENDING MORE TIME THAN YOU SAVE?
(ACROSS FIVE YEARS)

HOW OFTEN YOU DO THE TASK

	50/DAY	5/DAY	DAILY	WEEKLY	MONTHLY	YEARLY
1 SECOND	1 DAY	2 HOURS	30 MINUTES	4 MINUTES	1 MINUTE	5 SECONDS
5 SECONDS	5 DAYS	12 HOURS	2 HOURS	21 MINUTES	5 MINUTES	25 SECONDS
30 SECONDS	4 WEEKS	3 DAYS	12 HOURS	2 HOURS	30 MINUTES	2 MINUTES
1 MINUTE	8 WEEKS	6 DAYS	1 DAY	4 HOURS	1 HOUR	5 MINUTES
5 MINUTES	9 MONTHS	4 WEEKS	6 DAYS	21 HOURS	5 HOURS	25 MINUTES
30 MINUTES		6 MONTHS	5 WEEKS	5 DAYS	1 DAY	2 HOURS
1 HOUR		10 MONTHS	2 MONTHS	10 DAYS	2 DAYS	5 HOURS
6 HOURS				2 MONTHS	2 WEEKS	1 DAY
1 DAY					8 WEEKS	5 DAYS

HOW MUCH TIME YOU SHAVE OFF