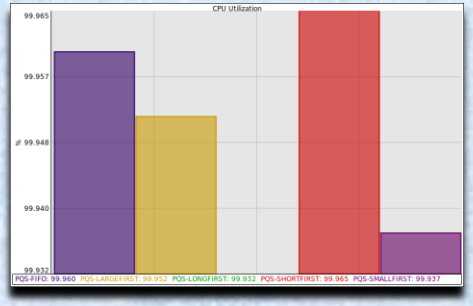
Trabajo Final de Grado (TFG)

Schedlyzer

Estudio de políticas de job scheduling mediante el simulador SST-Simulator



Autor:
José Antonio Domínguez Medina
jdominguezmed@uoc.edu

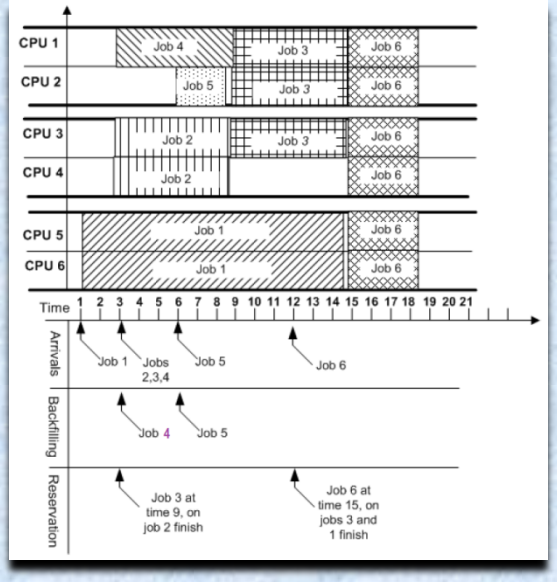
UOC-GEI
(Segundo semestre del 2015)

Consultor:
Francesc Guim Bernat

UOC
Universitat Oberta de Catalunya
www.uoc.edu

Trabajo Final de Grado (TFG)

Índice



1. Presentación.
2. Objetivos planteados.
3. Tareas involucradas.
4. Resultado obtenido.
5. Conclusiones.

José Antonio Domínguez Medina - GEI

UOC
Universitat Oberta de Catalunya
www.uoc.edu

1. **Presentación.**
2. Objetivos planteados.
3. Tareas involucradas.
4. Resultado obtenido.
5. Conclusiones.

Presentación

El presente documento pretende exponer, de forma muy sintetizada, el desarrollo realizado en el marco del Trabajo de Final de Grado en el área de Arquitectura de Computadores y Sistemas Operativos. Para ello, se realizará una presentación de los objetivos que se plantearon al inicio del mismo, se expondrán las tareas involucradas y se comentarán los resultados obtenidos que forman parte de los entregables de este TFG.

Para la realización de este TFG se ha requerido poner en práctica varias disciplinas aprendidas a lo largo de esta titulación, como pueden ser la investigación y análisis, la estructuración/esquematización de la información, la gestión de proyectos o la programación.

1. Presentación.
2. **Objetivos planteados.**
3. Tareas involucradas.
4. Resultado obtenido.
5. Conclusiones.

Objetivos planteados

(1/2)

Los objetivos principales planteados en este TFG son los siguientes:

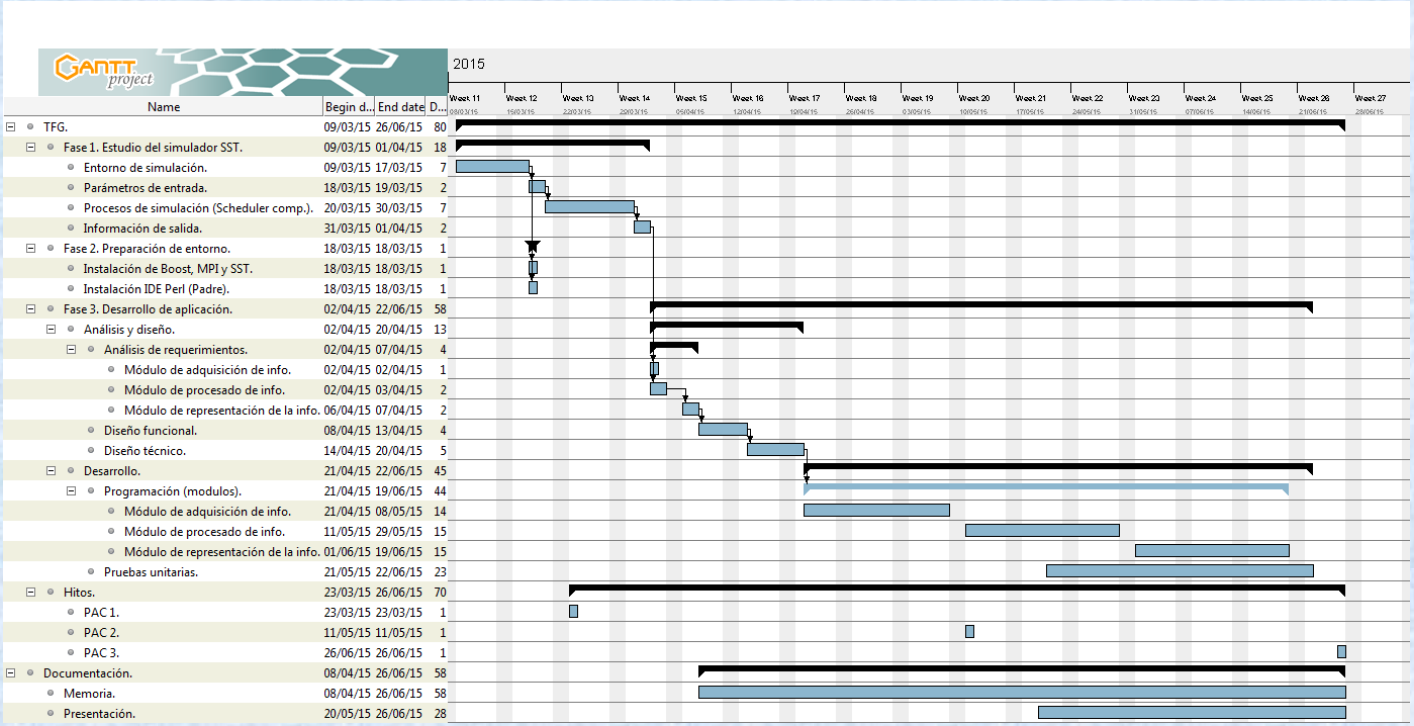
- Realizar una aplicación/tool en lenguaje de computación Perl que analice el comportamiento de las políticas de job scheduling aplicadas por un simulador de sistemas HPC, obteniendo unas métricas determinadas. Para ello la aplicación deberá obtener y procesar los resultados generados por el simulador SST-Simulator al someterle trabajos (jobs) para su planificación y ejecución.
- Realizar un análisis comparativo de las métricas obtenidas mediante los diferentes algoritmos de planificación.

Para cumplir con estos objetivo se diseña y dimensiona un plan de proyecto director que determina las tareas y duraciones de éstas para finalizarlo en los plazos que establece este TFG.

Objetivos planteados

(2/2)

A continuación se muestra el diagrama Gannt del plan de proyecto:



Name	Begin d...	End date	D...
TFG.	09/03/15	26/06/15	80
Fase 1. Estudio del simulador SST.	09/03/15	01/04/15	18
Entorno de simulación.	09/03/15	17/03/15	7
Parámetros de entrada.	18/03/15	19/03/15	2
Procesos de simulación (Scheduler comp.).	20/03/15	30/03/15	7
Información de salida.	31/03/15	01/04/15	2
Fase 2. Preparación de entorno.	18/03/15	18/03/15	1
Instalación de Boost, MPI y SST.	18/03/15	18/03/15	1
Instalación IDE Perl (Padre).	18/03/15	18/03/15	1
Fase 3. Desarrollo de aplicación.	02/04/15	22/06/15	58
Análisis y diseño.	02/04/15	20/04/15	13
Análisis de requerimientos.	02/04/15	07/04/15	4
Módulo de adquisición de info.	02/04/15	02/04/15	1
Módulo de procesado de info.	02/04/15	03/04/15	2
Módulo de representación de la info.	06/04/15	07/04/15	2
Diseño funcional.	08/04/15	13/04/15	4
Diseño técnico.	14/04/15	20/04/15	5
Desarrollo.	21/04/15	22/06/15	45
Programación (modulos).	21/04/15	19/06/15	44
Módulo de adquisición de info.	21/04/15	08/05/15	14
Módulo de procesado de info.	11/05/15	29/05/15	15
Módulo de representación de la info.	01/06/15	19/06/15	15
Pruebas unitarias.	21/05/15	22/06/15	23
Hitos.	23/03/15	26/06/15	70
PAC 1.	23/03/15	23/03/15	1
PAC 2.	11/05/15	11/05/15	1
PAC 3.	26/06/15	26/06/15	1
Documentación.	08/04/15	26/06/15	58
Memoria.	08/04/15	26/06/15	58
Presentación.	20/05/15	26/06/15	28

José Antonio Domínguez Medina - GEI

UOC
Universitat Oberta de Catalunya
www.uoc.edu

1. Presentación.
2. Objetivos planteados.
3. **Tareas involucradas.**
4. Resultado obtenido.
5. Conclusiones.

Tareas involucradas

(1/2)

Para llevar a cabo los objetivos planteados se realizan varias tareas encaminadas a finalizar con éxito la elaboración del TFG, son las siguientes:

Estudio del simulador SST.

Conjunto de tareas encaminadas a obtener unos conocimientos del simulador SST que nos permita avanzar en el desarrollo de las siguientes fases. Es necesario entender el entorno de ejecución del simulador, qué parámetros de entrada acepta, qué procesos se suceden y qué información es capaz de generar.

Preparación del entorno.

Son las tareas que se realizan para disponer de todo el entorno necesario para permitir la ejecución del simulador y que disponen de las herramientas y complementos necesarios para desarrollar la aplicación en Perl de análisis de las políticas de job scheduling.

Tareas involucradas (2/2)

Desarrollo de la aplicación/tool.

Conformada por las tareas más propias de un proyecto TIC basado en la metodología SDLC, en donde se detallan las diferentes tareas que se deben abordar de forma secuencial para elaborar y mantener un Sistema de Información durante su ciclo de vida útil. Entre estas tareas se han realizado las siguientes:

- Análisis de requerimientos/definición de funcionalidades.
- Diseño técnico.
- Ejecución de pruebas.

Análisis de resultados.

Se ha realizado un análisis/estudio comparativo de las diferentes métricas obtenidas por Schedlyzer sobre los datos proporcionados por SST-Simulator al aplicar las diferentes políticas de job scheduling.

UOC
Universitat Oberta de Catalunya
www.uoc.edu

1. Presentación.
2. Objetivos planteados.
3. Tareas involucradas.
4. **Resultado obtenido.**
5. Conclusiones.

Resultado obtenido

(1/3)

Como resultado de la consecución del proyecto se ha obtenido:

a) Documentación.

i. Memoria del TFG.

Se ha elaborado una memoria del TFG que incluye todo el estudio previo comentado referente al simulador SST, así como las tareas de adecuación del entorno necesarias para disponer de las herramientas imprescindibles para el desarrollo de la aplicación Perl de análisis de Políticas de Job Scheduling (Schedlyzer).

Esta documentación también incluye el análisis de todas aquellas tareas relacionadas con el desarrollo del ciclo de vida del proyecto antes comentadas, así como el análisis final comparativo de las diferentes métricas obtenidas por Schedlyzer sobre los datos proporcionados por SST-Simulator al aplicar las diferentes políticas de job scheduling.

Resultado obtenido

(2/3)

ii. Presentación del TFG.

Documento en formato Power Point que sintetiza el trabajo realizado y que permite obtener una visión general y completa de las tareas realizadas y del producto final obtenido. Ese documento es éste mismo que se está presentando.

b) Aplicación.

Se ha desarrollado una aplicación en lenguaje de computación Perl que analiza los resultados de los procesos de las políticas de job scheduling del simulador SST-Simulator (ficheros .time), proporcionando resultados -métricas- que permiten evaluar la idoneidad de éstas en un contexto de ejecución determinado.

UOC
Universitat Oberta de Catalunya
www.uoc.edu

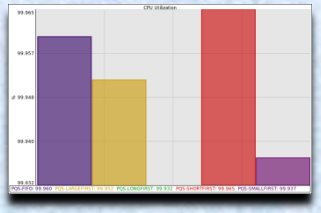
Resultado obtenido

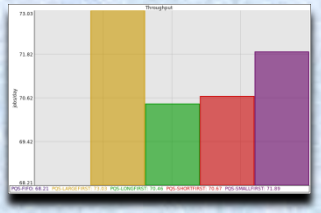
(3/3)

Las métricas obtenidas por la aplicación son las siguientes:

- Uso del procesador (CPU utilization).
- Productividad (throughput).
- Tiempo de retorno (Turnaround Time)
- Tiempo de espera (Waiting Time)

Por cada una de las métricas obtenidas, la aplicación genera una gráfica con todas las políticas de planificación consideradas en los ficheros .time de entrada.





José Antonio Domínguez Medina - GEI

UOC
Universitat Oberta de Catalunya
www.uoc.edu

1. Presentación.
2. Objetivos planteados.
3. Tareas involucradas.
4. Resultado obtenido.
5. **Conclusiones.**

Conclusiones

(1/2)

El estudio de las políticas de job scheduling en sistemas multiprocesador, en donde los trabajos sometidos son ejecutados concurrentemente, es una tarea muy compleja y está sometida a multitud de variables que deben tenerse en cuenta para poder determinar la idoneidad de una determinada política de planificación; empezando por obtener una muestra representativa de los datos a analizar. Precisamente por no disponer de todas estas variables, en la elaboración del análisis de las métricas obtenidas me he visto obligado a introducir conceptos teóricos, como el de fragmentación de recursos o grado de paralelismo, para poder explicar/desarrollar algunos resultados.

En general, podernos convenir en que la implementación del algoritmo EASY, utilizando la técnica de backfilling, obtiene mejores resultados que el algoritmo Prioritario para todas la métricas obtenidas en el sistema simulado y, por tanto, un mejor rendimiento en su conjunto.

Conclusiones

(2/2)

Para obtener las métricas, utilizadas como base del estudio comparativo, se ha realizado una aplicación en Perl para calcularlas, a partir de los ficheros .time generados por SST-Simulator al procesar el workload de SDSC y, para cada uno de los algoritmos/comparadores considerados en este estudio. Una vez obtenidas las métricas, éstas han sido representadas gráficamente mediante diagrama de barras. Esto ha requerido no sólo introducirme en un nuevo lenguaje de computación, como ha sido el Perl, sino también adentrarme en conceptos avanzados como es la modulación de código en paquetes y la generación gráfica en Linux.

Finalizar comentando que la realización de este TFG me ha proporcionado un mayor conocimiento sobre las políticas de planificación de procesos y, que ha sido muy gratificante aprender el lenguaje de computación Perl y adentrarme en el estudio de una parte tan fundamental del desarrollo de los Sistemas Operativos como son las job scheduling policies.

Trabajo Final de Grado (TFG)

Muchas gracias

José Antonio Domínguez Medina - GEI

UOC
Universitat Oberta de Catalunya
www.uoc.edu