UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ESCOLA TÉCNICA SUPERIOR DE ENXEÑARÍA

Título do Traballo de Fin de Grao
Subtítulo do Traballo de Fin de Grao

Autor:
Nome do autor

Directores:
Nome do director
Nome do codirector

Grao en Enxeñaría Informática

Febreiro 2011

Traballo de Fin de Grao presentado na Escola Técnica Superior de Enxeñaría da Universidade de Santiago de Compostela para a obtención do Grao en Enxeñaría Informática

D. (Nome do Director), Profesor do Departamento de Electrónica e Computación da Universidade de Santiago de Compostela, e **D. (Nome do Codirector)**, Profesor do Departamento de Electrónica e Computación da Universidade de Santiago de Compostela,

INFORMAN:

Que a presente memoria, titulada *(Título do traballo)*, presentada por **D. (Nome do autor do traballo)** para superar os créditos correspondentes ao Traballo de Fin de Grao da titulación de Grao en Enxeñaría Informática, realizouse baixo nosa dirección no Departamento de Electrónica e Computación da Universidade de Santiago de Compostela.

E para que así conste aos efectos oportunos, expiden o presente informe en Santiago de Compostela, a (Data):

O director, O codirector, O alumno,

(Nome do director) (Nome do Codirector) (Nome do Alumno)

Agradecementos

Se se quere pór algún agradecemento, este vai aquí.

Resumo

Se se quere pór resumo, este vai aquí.

Índice general

Índice de figuras

Índice de cuadros

Capítulo 1

Introdución

Introdución: composta por Obxectivos Xerais, Relación da Documentación que conforma a Memoria, Descrición do Sistema, Información Adicional de Interese (métodos, técnicas ou arquitecturas utilizadas, xustificación da súa elección, etc.).

Capítulo 2

Gestión del proyecto

2.1. Alcance

2.2. Análisis de requisitos

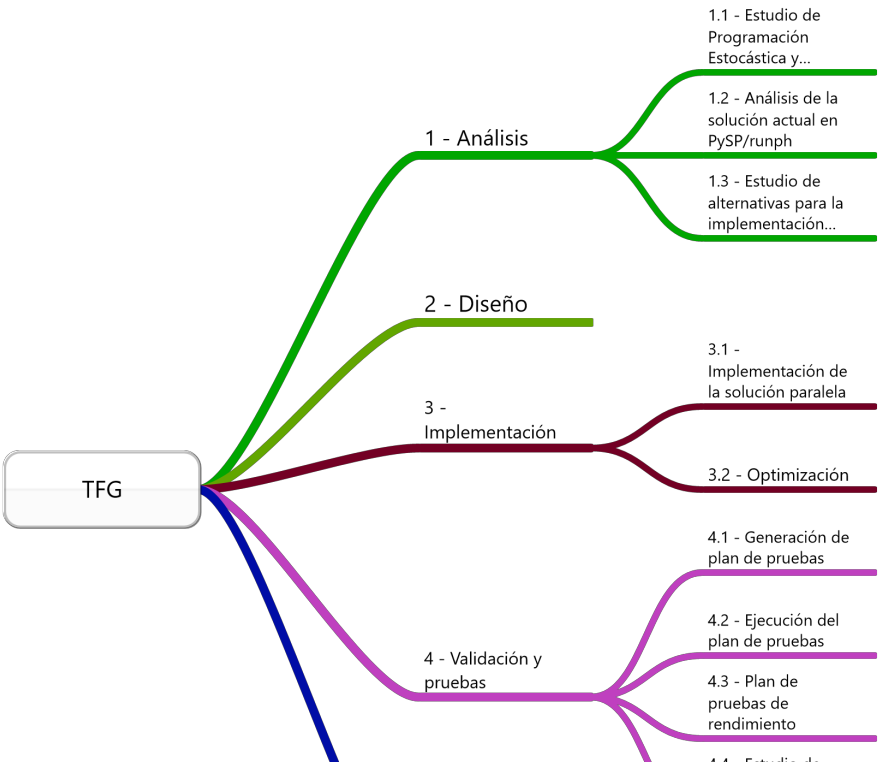
2.3. Análisis de riesgos

2.4. Análisis de costes

2.5. Gestión de la configuración

2.6. Planificación temporal

2.6.1. EDT



Id	1.1
Tarea	Estudio de Programación Estocástica y Progressive Hedging
Duración	7 días
Descripción	Se investigará el funcionamiento del algoritmo *Progressive Hedging* haciendo uso principalmente de [2] como referencia.
Salida	N/A

Id	1.2
Tarea	Análisis de la solución actual en PySP/runph
Duración	8 días
Descripción	Una vez conocido el funcionamiento teórico del algoritmo, estudiar su implementación sobre el proyecto Pyomo.
Salida	Diagrama de funcionamiento PH [3].

Id	1.3
Tarea	Estudio de alternativas para la implementación paralela
Duración	3 días
Descripción	Se barajarán tecnologías Big Data (Spark) o modelos tradicionales (MPI).
Salida	TODO

Id	2
Tarea	Diseño
Duración	20 días
Descripción	Generar un diseño de la implementación a realizar con la tecnología escogida. Será importante la integración con la implementación actual.
Salida	Diseño de la implementación.

Id	3.1
Tarea	Implementación de la solución paralela
Duración	15 días
Descripción	Escribir los nuevos módulos de código e integrarlos en el proyecto.
Salida	Nuevos ficheros de código y modificaciones a archivos existentes.

Id	3.2
Tarea	Optimización
Duración	5 días
Descripción	Una vez la integración es correcta y la implementación funciona, se realizarán optimizaciones de rendimiento intentando aprovechar las carácterísticas de la tecnología escogida para la nueva implementación paralela.
Salida	Modificaciones a la implementación anterior.

Id	4.1
Tarea	Generación de plan de pruebas
Duración	4 días
Descripción	Idear un plan de pruebas para el nuevo módulo.
Salida	Documento de pruebas [**?**].

Id	4.2
Tarea	Ejecución del plan de pruebas
Duración	1 días
Descripción	Implementar y ejecutar las pruebas establecidas para establecer confianza sobre el correcto funcionamiento de la implementación.
Salida	Informe de ejecución de pruebas.

Id	4.3
Tarea	Plan de pruebas de rendimiento
Duración	3 días
Descripción	Idear un plan de pruebas con problemas que permitan estudiar el rendimiento del programa.
Salida	Documento de pruebas de rendimiento [**?**].

Id	4.4
Tarea	Estudio de rendimiento
Duración	2 días
Descripción	Ejecutar el plan de pruebas anterior y compararlo con las versiones originales tanto secuencial como con Pyro.
Salida	Informe de rendimiento [**?**].

Id	5
Tarea	Documentación
Duración	10 días
Descripción	Generar los documentos asociados al desarrollo del proyecto.
Salida	Memoria del proyecto y documentos asociados.

Tareas no planificadas

Tras las modificaciones realizadas sobre el cronograma y explicadas en Subsección 2.6.3, se generan nuevas tareas para el proyecto:

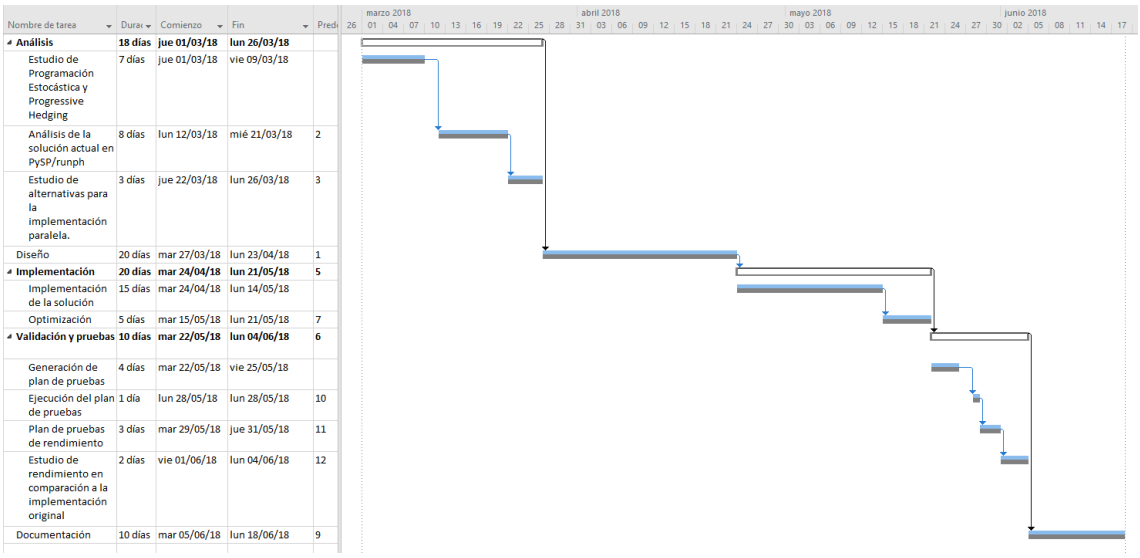
Id	3.*
Tarea	Prototipo aislado
Duración	10 días
Descripción	Crear una aplicación en Python que interactúe con Spark de forma similar a como lo hará la implementación en Pyomo.
Salida	Proyecto de pruebas en Python [4].

Id	3.*
Tarea	Prototipo de integración inicial
Duración	5 días
Descripción	Comenzar la implementación sobre Pyomo comprobando que la integración del nuevo módulo con el código existente es correcta y el flujo de ejecución es correcto con respecto al funcionamiento anterior.
Salida	Código añadido a Pyomo.

Id	3.*
Tarea	Prototipo funcional
Duración	10 días
Descripción	Modificar el prototipo anterior añadiendo las funcionalidades esperadas del programa.
Salida	Código añadido a Pyomo.

2.6.2. Cronograma inicial

Para la realización del trabajo se plantea un ciclo de vida en cascada. Este ciclo de vida nos permitirá realizar un seguimiento del progreso del proyecto en función del tiempo disponible.



Nombre de tarea	Duración	Comienzo	Fin	Predecesoras
Análisis	**18 días**	**jue 01/03/18**	**lun 26/03/18**	
Estudio de Programación Estocástica y Progressive Hedging	7 días	jue 01/03/18	vie 09/03/18	
Análisis de la solución actual en PySP/runph	8 días	lun 12/03/18	mié 21/03/18	2
Estudio de alternativas para la implementación paralela.	3 días	jue 22/03/18	lun 26/03/18	3
Diseño	20 días	mar 27/03/18	lun 23/04/18	1
Implementación	**20 días**	**mar 24/04/18**	**lun 21/05/18**	**5**
Implementación de la solución	15 días	mar 24/04/18	lun 14/05/18	
Optimización	5 días	mar 15/05/18	lun 21/05/18	7
Validación y pruebas	**10 días**	**mar 22/05/18**	**lun 04/06/18**	**6**
Generación de plan de pruebas	4 días	mar 22/05/18	vie 25/05/18	
Ejecución del plan de pruebas	1 día	lun 28/05/18	lun 28/05/18	10
Plan de pruebas de rendimiento	3 días	mar 29/05/18	jue 31/05/18	11
Estudio de rendimiento en comparación a la implementación original	2 días	vie 01/06/18	lun 04/06/18	12
Documentación	10 días	mar 05/06/18	lun 18/06/18	9

Figura 2.2: Línea base

2.6.3. Modificaciones al cronograma inicial

Se ha realizado una estimación temporal inicial poco precisa por no utilizar ningún tipo de método probado ni datos concretos.

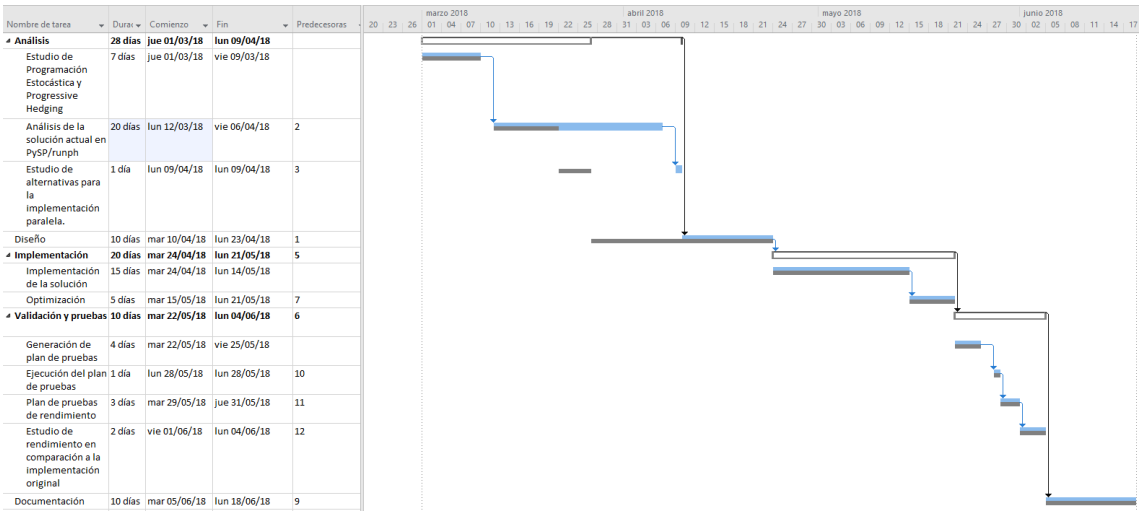
Esta es la razón principal para los fallos de cumplimiento de la planificación, al menos durante las fases de análisis y diseño.

El primer retraso se produce en la fase de análisis de la implementación actual. En esta fase se debe estudiar el funcionamiento del proyecto Pyomo y, en concreto, el módulo de resolución de problemas mediante Progressive Hedging.

A pesar de conocer el funcionamiento teórico del algoritmo mediante [**?**], Pyomo es un proyecto complejo, con multitud de funcionalidades para resolver otros tipos de problemas, soporte para plugins, etc. Todo esto hace que la complejidad del código aumente y sea necesario estudiar múltiples capas de abstracción para entender correctamente el funcionamiento del programa.

Otra complicación añadida es el personal desconocimiento del lenguaje Python previo a la realización de este trabajo.

Tras este primer retraso se intenta ajustar la planificación reduciendo el tiempo de diseño a la mitad:



Nombre de tarea	Dura	Comienzo	Fin	Predecesoras
Análisis	**28 días**	**jue 01/03/18**	**lun 09/04/18**	
Estudio de Programación Estocástica y Progressive Hedging	7 días	jue 01/03/18	vie 09/03/18	
Análisis de la solución actual en PySP/runph	20 días	lun 12/03/18	vie 06/04/18	2
Estudio de alternativas para la implementación paralela.	1 día	lun 09/04/18	lun 09/04/18	3
Diseño	10 días	mar 10/04/18	lun 23/04/18	1
Implementación	**20 días**	**mar 24/04/18**	**lun 21/05/18**	**5**
Implementación de la solución	15 días	mar 24/04/18	lun 14/05/18	
Optimización	5 días	mar 15/05/18	lun 21/05/18	7
Validación y pruebas	**10 días**	**mar 22/05/18**	**lun 04/06/18**	**6**
Generación de plan de pruebas	4 días	mar 22/05/18	vie 25/05/18	
Ejecución del plan de pruebas	1 día	lun 28/05/18	lun 28/05/18	10
Plan de pruebas de rendimiento	3 días	mar 29/05/18	jue 31/05/18	11
Estudio de rendimiento en comparación a la implementación original	2 días	vie 01/06/18	lun 04/06/18	12
Documentación	10 días	mar 05/06/18	lun 18/06/18	9

Figura 2.3: Primer retraso

Teniendo en cuenta que los primeros retrasos fueron principalmente causados por el desconocimiento de la tecnología a usar así como de una mala estimación, es muy probable que en las fases siguientes se produzcan otros retrasos. En este punto se considera entonces abandonar el ciclo de vida en cascada. Su principal atractivo era poder ajustarnos a una planificación temporal que nos permita acabar el proyecto dentro de tiempo, pero esta ventaja no se está cumpliendo en la práctica. Buscando reducir el tiempo de implementación con tecnologías que serán usadas por primera vez, se decide adaptar la planificación a un ciclo de vida por prototipos.

La creación de sucesivos prototipos permite ir acostumbrándose a las tecnologías desconocidas, en este caso Python y Spark, así como ir probando el rendimiento y la integración a medida que se desarrolla.

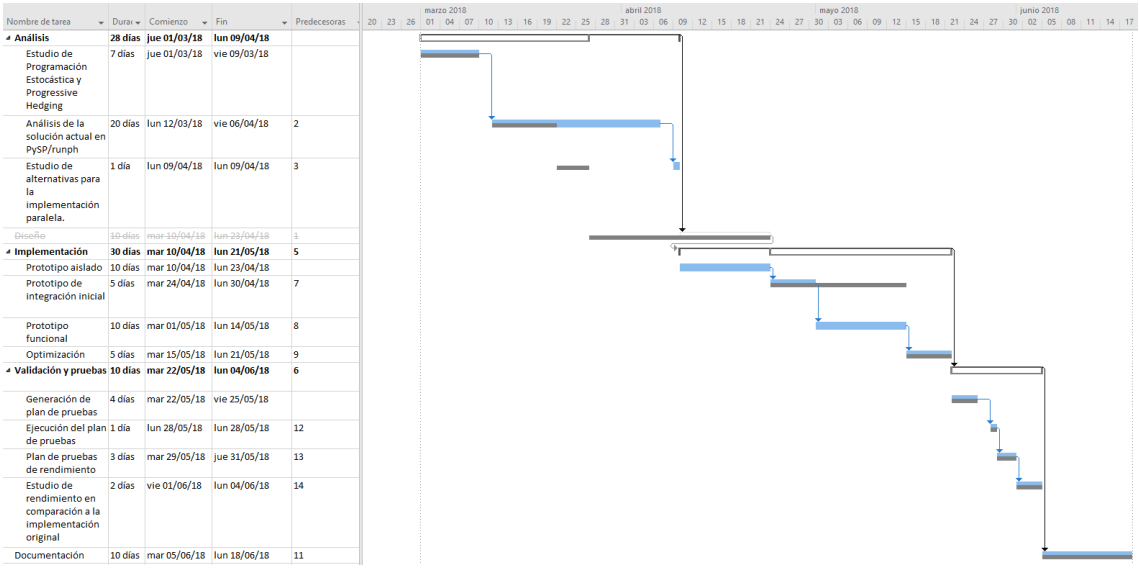
En primer lugar se crea un prototipo aislado para comprobar la implementación de Spark con una arquitectura similar a la que se implementará en Pyomo. Este primer prototipo sirve como aprendizaje de la instalación de Spark y el despliegue de una aplicación en el mismo, así como la implementación en python que interactuará con Spark. Es deseable utilizar una arquitectura de objetos python similar a la que se usará en Pyomo para concretar el uso de Spark y descubrir posibles problemas con la implementación elegida.

Posteriormente se realizarán prototipos sucesivos sobre Pyomo para integrar el nuevo módulo e ir solucionando posibles problemas de rendimiento o funcionamiento que vayan surgiendo.

Dado que la implementación partirá de un protipo inicial de baja calidad será importante realizar una fase de optimización y refactorización al final de la implementación para asegurarse un código final de calidad. Definir la calidad del código no es algo trivial y en este caso calificaremos el código como "de calidad"si cumple:

- Funciona correctamente y es resistente a errores. Para esto nos apoyaremos en un plan de pruebas funcional.
- Funciona eficientemente y otorga un buen rendimiento, en comparación a las implementaciones existentes. En este caso nos apoyaremos en el plan de pruebas de rendimiento.
- Se integra adecuadamente al proyecto actual. Debe seguir una filosofía de diseño análoga al resto del código así como funcionar correctamente de forma paralela a todo lo implementado previamente.

Tras esta modificación en la planificación, se genera una nueva planificación que podemos ver en la figura y se guardará como una nueva línea base.



Nombre de tarea	Durac	Comienzo	Fin	Predecesoras
Análisis	**28 días**	**jue 01/03/18**	**lun 09/04/18**	
Estudio de Programación Estocástica y Progressive Hedging	7 días	jue 01/03/18	vie 09/03/18	
Análisis de la solución actual en PySP/runph	20 días	lun 12/03/18	vie 06/04/18	2
Estudio de alternativas para la implementación paralela.	1 día	lun 09/04/18	lun 09/04/18	3
~~Diseño~~	~~10 días~~	~~mar 10/04/18~~	~~lun 23/04/18~~	~~1~~
Implementación	**30 días**	**mar 10/04/18**	**lun 21/05/18**	**5**
Prototipo aislado	10 días	mar 10/04/18	lun 23/04/18	
Prototipo de integración inicial	5 días	mar 24/04/18	lun 30/04/18	7
Prototipo funcional	10 días	mar 01/05/18	lun 14/05/18	8
Optimización	5 días	mar 15/05/18	lun 21/05/18	9
Validación y pruebas	**10 días**	**mar 22/05/18**	**lun 04/06/18**	**6**
Generación de plan de pruebas	4 días	mar 22/05/18	vie 25/05/18	
Ejecución del plan de pruebas	1 día	lun 28/05/18	lun 28/05/18	12
Plan de pruebas de rendimiento	3 días	mar 29/05/18	jue 31/05/18	13
Estudio de rendimiento en comparación a la implementación original	2 días	vie 01/06/18	lun 04/06/18	14
Documentación	10 días	mar 05/06/18	lun 18/06/18	11

Figura 2.4: Línea base prototipos

En este punto hemos eliminado la fase de Diseño para poder aumentar el tiempo asignado a Análisis e Implementación. En caso de sufrir más retrasos en la fase de implementación podremos reducir el tiempo asignado a pruebas si el retraso no es grave. En caso de ser un retraso mayor, no cumpliremos la fecha de finalización establecida.

Capítulo 3

Planificación e presupostos

Planificación e presupostos: debe incluír a estimación do costo (presuposto) e dos recursos necesarios para efectuar a implantación do Traballo, xunto coa planificación temporal do mesmo e a división en fases e tarefas. Recoméndase diferenciar os costos relativos a persoal dos relativos a outros gastos como instalacións e equipos.

Capítulo 4

Especificación de requisitos

Especificación de requisitos: debe indicarse, polo miúdo, a especificación do Sistema, xunto coa información que este debe almacenar e as interfaces con outros Sistemas, sexan hardware ou software, e outros requisitos (rendemento, seguridade, etc).

Capítulo 5

Deseño

Deseño: cómo se realiza o Sistema, a división deste en diferentes compoñentes e a comunicación entre eles. Así mesmo, determinarase o equipamento hardware e software necesario, xustificando a súa elección no caso de que non fora un requisito previo. Debe achegarse a un nivel suficiente de detalle que permita comprender a totalidade da estrutura do produto desenvolvido, utilizando no posible representacións gráficas.

Capítulo 6

Exemplos

6.1. Un exemplo de sección

Esta é *letra cursiva*, esta é **letra negrilla**, esta é letra subrallada, e esta é `letra curier`. Letra tiny, scriptsize, small, large, Large, LARGE e moitas más. Exemplo de fórmula: $a = \int_o^\infty f(t)dt$. E agora unha ecuación aparte:

$$S = \sum_{i=0}^{N-1} a_i^2. \tag{6.1}$$

As ecuaciones se poden referenciar: ecuación (6.1).

6.1.1. Un exemplo de subsección

O texto vai aquí.

6.1.2. Otro exemplo de subsección

O texto vai aquí.

Un exemplo de subsubsección

O texto vai aquí.

Un exemplo de subsubsección

O texto vai aquí.

Un exemplo de subsubsección

O texto vai aquí.

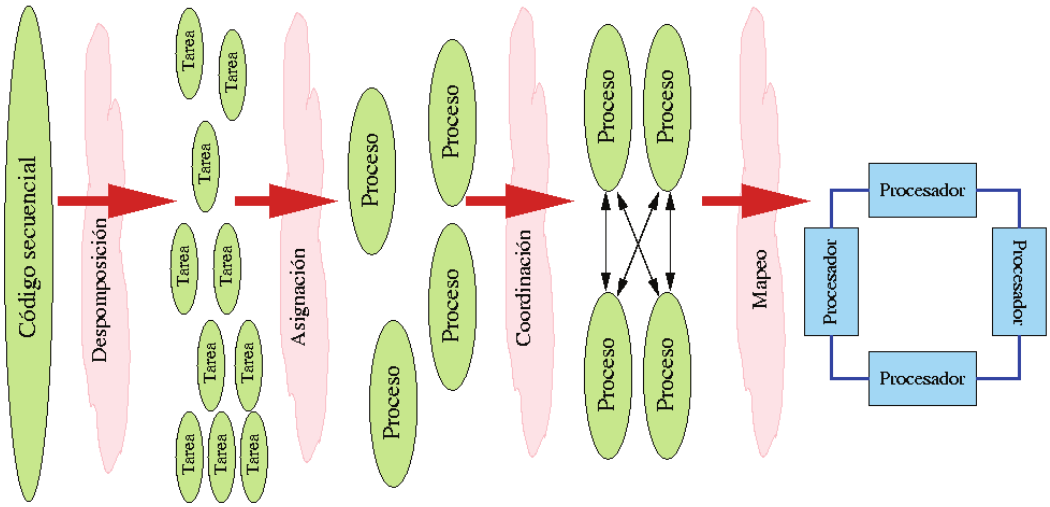


Figura 6.1: Esta é a figura de tal e cal.

Izquierda	Derecha	Centrado
ll	r	cccc
llll	rrr	c

Cuadro 6.1: Esta é a táboa de tal e cal.

6.2. Exemplos de figuras e cadros

A figura número 6.1.
O cadro (taboa) número 6.1.

6.3. Exemplos de referencias á bibliografía

Este é un exemplo de referencia a un documento descargado da web [**?**]. E este é un exemplo de referencia a unha páxina da wikipedia [**?**]. Agora un libro [**?**] e agora unha referencia a un artigo dunha revista [**?**]. Tamén se poden pór varias referencias á vez [**?**, **?**].

6.4. Exemplos de enumeracións

Con puntos:

- Un.
- Dous.

- Tres.

Con números:

1. Catro.
2. Cinco.
3. Seis.

Exemplo de texto verbatim:

```
O texto        verbatim
     se visualiza tal
            como se escribe
```

Exemplo de código C:

```c
#include <math.h>
main()
{   int i, j, a[10];
    for(i=0;i<=10;i++) a[i]=i; // comentario 1
    if(a[1]==0) j=1; /* comentario 2 */
    else j=2;
}
```

Exemplo de código Java:

```java
class HelloWorldApp {
    public static void main(String[] args) {
        System.out.println("Hello World!"); // Display the string.
    }
}
```

Capítulo 7

Conclusións e posibles ampliacións

Conclusións e posibles ampliacións

Apéndice A

Manuais técnicos

Manuais técnicos: en función do tipo de Traballo e metodoloxía empregada, o contido poderase dividir en varios documentos. En todo caso, neles incluirase toda a información precisa para aquelas persoas que se vaian a encargar do desenvolvemento e/ou modificación do Sistema (por exemplo código fonte, recursos necesarios, operacións necesarias para modificacións e probas, posibles problemas, etc.). O código fonte poderase entregar en soporte informático en formatos PDF ou postscript.

Apéndice B

Manuais de usuario

Manuais de usuario: incluirán toda a información precisa para aquelas persoas que utilicen o Sistema: instalación, utilización, configuración, mensaxes de erro, etc. A documentación do usuario debe ser autocontida, é dicir, para o seu entendemento o usuario final non debe precisar da lectura de outro manual técnico.

Apéndice C

Licenza

Se se quere pór unha licenza (GNU GPL, Creative Commons, etc), o texto da licenza vai aquí.

Bibliografía

[1] Matei Zaharia, Patrick Wendell, Andy Konwinski, Holden Karau, *Learning Spark: Lightning-Fast Big Data Analysis*, 1ª edición, O'Reilly Media, 2015.

[2] John R. Birge, François Louveaux, *Introduction to Stochastic Programming*, 2ª Edición, Springer, 2011.

[3] PONER AQUÍ LA RUTA

[4] Proyecto de pruebas en: "/DistributedTest"