Trabajo Final de Grado

Grado de Tecnologías de Telecomunicación

IT-RAMS: Modelización y Simulación de la Fiabilidad, Disponibilidad y Mantenibilidad de Sistemas TIC

Autor: **Sergio Gonzalo San José**

Fecha: **Enero 2016**

Tabla de Contenidos
Índice

Formulación del Problema

Tres Preguntas

1. ¿Es posible y viable **estimar y predecir la disponibilidad y fiabilidad de los sistemas** a través sus características inherentes y los datos operacionales de los diferentes sistemas individuales?.

2. ¿Es posible **fijar requerimientos de disponibilidad y fiabilidad a los sistemas y aplicaciones** bajo desarrollo?.

3. ¿Qué criterios ha de contemplar un responsable de un servicio/sistema a la hora de decidir la **idoneidad de las inversiones en la mejora de la disponibilidad de los sistemas**?

Formulación del Problema

- La formulación del problema está enraizada en el **concepto RAMS**
  - Criticidad e impacto creciente de servicios y sistemas TIC.
  - El problema es **más extenso** → Nuevos desafíos (terrorismo, armas de destrucción masiva, …) confieren a la **Seguridad e Infraestructuras** un carácter cada vez más complejo y crítico.
  - Caracterización y modelado de los sistemas y procesos con el fin de medir y mejorar su fiabilidad y disponibilidad → **IT-RAMS (RAMS for TIC)**

- Los **retos** de este escenario son:
  - Derivación de disponibilidad y fiabilidad en base a datos de más bajo nivel.
  - Aplicación e imposición de requisitos de disponibilidad y fiabilidad a diseñadores y arquitectos del sistema (**validez en diseño y producción**)
  - Modelos de toma de decisiones enfocados al diseño e inversión.

- **El problema puede formularse por tanto del siguiente modo:**

  "*Establecer un marco de simulación matemática y estadística aplicable a infraestructuras TIC para el análisis de modos de fallo, detección de puntos críticos, inspección de propagación de errores, mitigar y reducir los efectos, y prevenir los riesgos detectados, generando una herramienta estratégica y táctica de estudio y análisis de fallos así como de orientación eficiente de inversiones.*"

Objetivos Principales

Investigación y Conocimiento

- Estudio y análisis de la tipología de **análisis y estudios RAMS**.
- **Modelos matemáticos y estadísticos** aplicables a **procesos y procedimientos de sistemas** (mto, operación, corrección, …)
- **Modelos matemáticos** aplicables al **comportamiento de los sistemas** (Hw, Sw base, Sw aplicación).
- Modelos y tipología de **arquitecturas de clustering**.

Desarrollo e Innovación (IT-RAMS)

- Construcción de **modelo de componentes y espacio de simulación estadística** (MonteCarlo) → Lenguaje Java.
- Estudios de **Análisis de Sensibilidad** y su aplicabilidad como **herramienta estratégica de diseño e inversión eficiente**.
- Simulación y análisis de sensibilidad de un **clúster de aplicación activo-pasivo**.
- **Diseño óptimo de sistemas TIC** críticos con **validación de objetivos y orientación a resultados**.

RAMS y Aproximación al Sector TIC

¿Qué es RAMS?

- **RAMS (R**eliability, **A**vailability, **M**aintainability, **S**afety)

- Conjunto de **técnicas y prácticas usadas en el sector aeronáutico, aeroespacial y defensa para el diseño de servicios críticos** cuya indisponibilidad implicaría:
  - Pérdida de vidas humanas
  - Altos costes materiales

- Fundamentada en **modelos matemáticos estadísticos y probabilísticos**.

- No sólo contempla los fallos sino también los **procesos de prevención, integración y recuperación**.

- ¿Porqué no se ha aplicado a otros sectores hasta ahora?
  - Time-to-market + servicios no críticos → inversiones no justificadas.

RAMS y Aproximación al Sector TIC

Aproximación RAMS al Sector TIC

- **Aplicación de RAMS en sector TIC**
  - Soporte al diseño óptimo de servicios de disponibilidad crítica.
  - Caracterización del estado de disponibilidad de servicios ya existentes.

- **Principales Beneficios**
  - Conocimiento óptimo de los sistemas y arquitecturas (disponibilidad global, cuellos de botella, etc.)
  - Identificación de puntos de acción (inversión) para mejorar la disponibilidad del sistema.
  - Optimización de procesos asociados a los sistemas y entornos.
  - Predicción del impacto de ciertos eventos en la disponibilidad de los servicios.

"Cualquier servicio o sistema es susceptible de ser diseñado mediante técnicas RAMS, con el objetivo de alcanzar la disponibilidad deseada a un coste óptimo".

IT-RAMS y Modelo de Componentes

Descripción del Enfoque IT-RAMS Realizado

- Basada en cuatro fases:
  1. Modelización de Componentes
  2. Construcción del Modelo de Simulación
  3. Análisis de MonteCarlo
  4. Análisis de Sensibilidad

- Aspectos de Innovación
  1. Aplicación RAMS al sector TI.
  2. Modelos estadísticos de caracterización del software.
  3. Modelos de crecimiento de fiabilidad del software.
  4. Fundamentado en múltiples estudios y análisis teóricos
  5. Aplicación de técnicas contrastadas (p.ej ESA, NASA).

Objetivos

1. Conocimiento de la disponibilidad (análisis cuantitativo).
   - Eliminación de la incertidumbre.
   - Enfoque del Sistema al Compromiso de Disponibilidad.
2. Análisis de sensibilidad (planes de mejora).
   - Estudio de diferentes escenarios orientados a la toma de decisiones de inversión en base al objetivo de disponibilidad.
   - Business Case [escenario, coste, mejora numérica de disponibilidad].
   - Estudios cuantitativos con impacto 0 en negocio.

IT-RAMS y Modelo de Componentes

Modelo de Componentes

❑ Estructura del Modelo de Componentes

- Desarrollo 100% Lenguaje Java
- Conjunto de clases encargadas de modelar el comportamiento de los diferentes elementos y servicios que componen una arquitectura TI

❑ ¿Qué Aspectos Contempla el Modelo de Componentes?.

- Hardware
- Software (COTS y desarrollado)
- Software reliabitlity growth
- Topologías HA y tolerancia: Cluster A-P (activo-pasivo)
- Corrección de fallos.
- Puestas en producción.
- Sistema de monitorización
- Equipos y modelos de reparación.
- Calendarios de servicio y reparación.
- Visión del usuario/cliente final (calendario de usuario).
- Diagnósticos del sistema

IT-RAMS y Modelo de Componentes

Modelo de Componentes

❑ Objetivos y Principios del Modelo de Componentes

- El objetivo del modelo de componentes es construir una herramienta software que permita obtener **estimaciones de la disponibilidad de un sistema IT** de acuerdo a las **características, relaciones, dependencias y comportamiento de los componentes que lo constituyen**.

- **Posibilita la definición y construcción de entidades y sistemas de alto nivel con mayor nivel de complejidad** como relaciones y combinaciones de estos componentes.

- El modelo de estudio de la disponibilidad de los componentes está basado en el **Análisis de Árboles de Fallos (AAF)**
  - Técnica deductiva para determinar las causas que han generado un fallo.
  - Permite analizar la propagación de fallos ocurridos en un componente sobre el resto de la arquitectura.
  - Esta técnica está basada en un **proceso deductivo basado en las leyes del Álgebra de Boole**, permitiendo determinar la expresión de sucesos complejos en función de fallos básicos en los elementos que intervienen.

IT-RAMS y Modelo de Componentes

Modelo de Componentes

❑ Parametrización de los Componentes

- ❑ La parametrización de la disponibilidad de los componente es dependiente de la tipología de cada entidad modelada.

❑ **COMPONENTES**

❑ **Hardware, Software (Base, Aplicación)**
- ➢ Tiempo medio entre fallos (MTBF).
- ➢ Tiempo medio de reparación (MTTR).
- ➢ Desviación típica del tiempo medio de reparación (σ).

❑ **Software (Base, Aplicación)**
- ➢ Número de defectos software mayores (causantes de indisponibilidad)
- ➢ Transacciones por segundo (TPS).

❑ **Clúster de Aplicación**
- ➢ Tiempo medio de reparación (MTTR)
- ➢ Desviación típica del tiempo medio de reparación (σ).
- ➢ Fiabilidad porcentual del proceso failover de la aplicación al nodo pasivo.

❑ **SERVICIOS**

❑ **Servicio de Mantenimiento**
- ➢ Mantenimientos/año con indisponibilidad.
- ➢ Duración media del mantenimiento.
- ➢ Desviación de duración mantenimiento (σ).

❑ **Servicio de Monitorización**
- ➢ Probabilidad de detección de un fallo ocurrido.

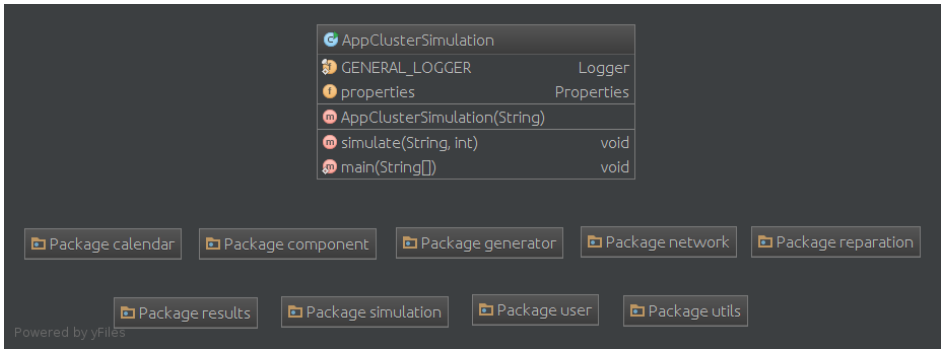
❑ **Servicio de Operación**
- ➢ Errores de operación producidos al año.

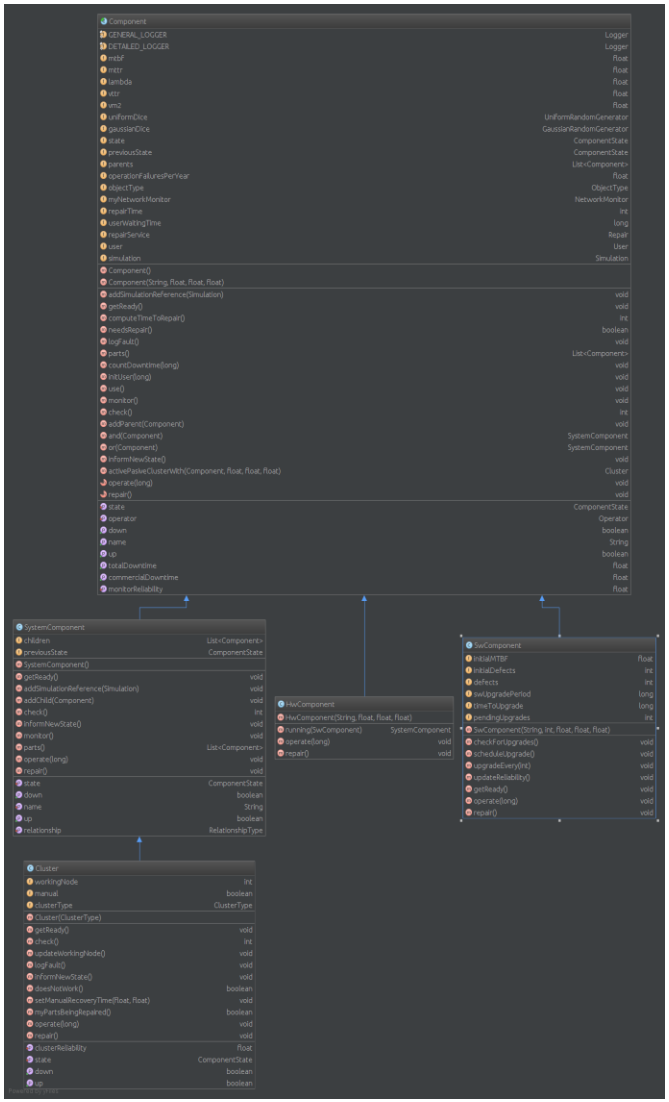
❑ **Servicio de Reparación**
- ➢ Calendario del servicio de reparación (horario de reparación, SLA).

❑ **Servicio a Usuario o Cliente Final**
- ➢ Calendario del Usuario (impacto sobre uso esperado del sistema por parte del usuario).

IT-RAMS y Modelo de Componentes
Modelo de Componentes

- **Estructura de Paquetes del Modelo de Componentes**





- **Package "edu.uoc.itrams.component"**
- **Clases de modelado de componentes**
  - **Component** (clase base genérica abstracta)
  - **HwComponent**, **SwComponent** (clases hija) → Componentes básicos
  - **SystemComponent** (clase hija) → Componente Complejo = Hw+Sw
  - **Clúster** (clase hija de **SystemComponent**)

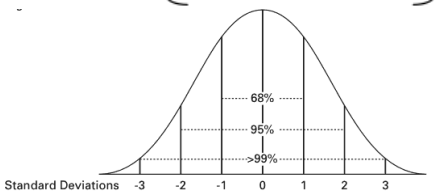
IT-RAMS y Modelo de Componentes

Modelos Estadísticos

- **Generador Distribución Uniforme de Números Aleatorios**
  - Desarrollo de generador de números aleatorios en el rango 0-2147483647
  - Basado en método propuesto por Stephen K. Park y Keith W. Miller en **Random Number Generators: Good Ones Are Hard To Find**, definido por cuatro variables:
    - La **variable "m"** o modulo (2147483647).
    - La **variable "a"** o multiplicador (48271).
    - La **variable "q"** o valor entero de la división del módulo y multiplicador (44488).
    - La **variable "r"** o resto de la división del módulo y multiplicador (3399).

- **Generador Distribución Normal a partir de Muestras Aleatorias**
  - Muestras aleatorias generadas mediante el generador uniforme
  - Basado en método propuesto por Susanna W. M Au-Yeung en **Finding Probability Distributions From Moments** definido por cuatro variables (λ1, λ2, λ3, y λ4) y la siguiente expresión:
    - Q(u) es la función de distribución inversa donde "u" toma valores entre 0 y 1 devolviendo el valor de "x" tal que F(x) = u, donde F(x) es la función de distribución acumulada.

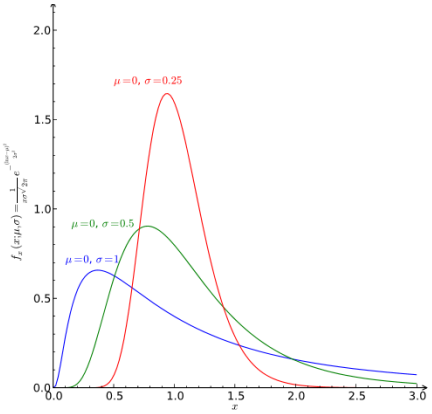
$$Q(u) = \lambda_1 + \frac{1}{\lambda_2}\left[\frac{u^{\lambda_3} - 1}{\lambda_3} - \frac{(1-u)^{\lambda_4} - 1}{\lambda_4}\right]$$



IT-RAMS y Modelo de Componentes

Modelos Estadísticos

- **Generador Distribución Log-Normal**
  - Empleada ampliamente para el modelado de los tiempos de reparación de los sistemas.
    - Refleja una **corta duración de los tiempos** de reparación
    - Un **gran número de observaciones** están **en torno al valor modal**.
    - **Ciertas observaciones** presentan **grandes tiempos de reparación**.
  - Generada a partir de valores de una distribución normal:



$$X = e^{\mu+\sigma Z}$$

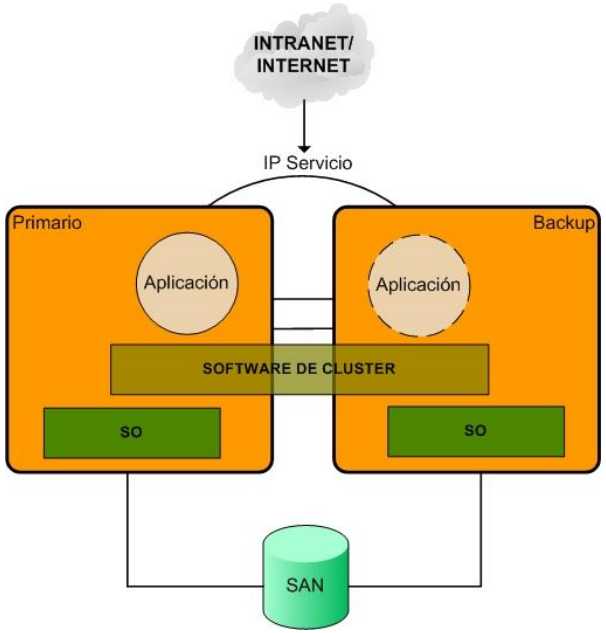
donde:

- Z = valor aleatorio de la distribución normal.
- μ = media de la distribución log-normal $\left[\mu = ln\left(\frac{m}{\sqrt{1+\frac{v}{m^2}}}\right)\right]$
- σ = desviación típica de la distribución log-normal $\left[\sigma = \sqrt{ln\left(1+\frac{v}{m^2}\right)}\right]$
- m = media de la distribución normal
- v = desviación típica de la distribución normal

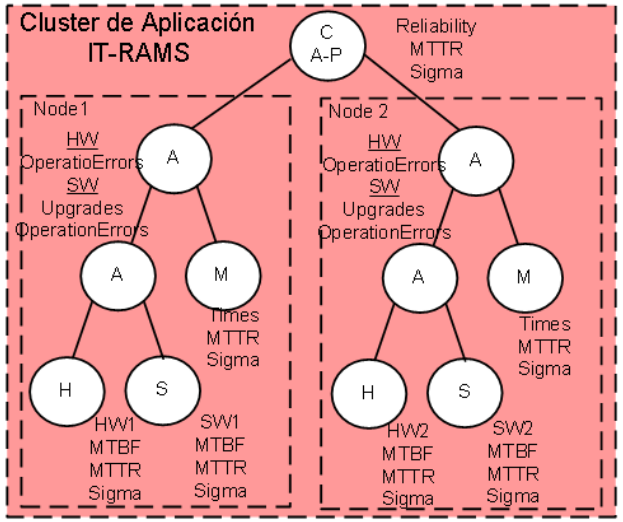
Clúster de Aplicación A-P

Descripción Funcional y Árbol de Fallos

- **Sistemas**
  - primario activo y operativo.
  - backup de respaldo.
- Fallo en sistema **primario (activo)**:
  - Proceso de **Failover**: La aplicación es trasferida por el sw de clúster al nodo de respaldo
- Fallo en sistema **backup (pasivo)**:
  - La aplicación continúa operativa en nodo primario.
  - Pérdida de tolerancia a fallos.



- **Disponibilidad es función de:**
  - Disponibilidades de componentes inferiores
  - Características y dependencias de los componentes.
  - Relación de componentes (algebra de Boole)



- **H** = Hardware
- **S** = Software
- **M** = Labor de Mantenimiento (causante de indisponibilidad).
- **C A-P** = Clúster Activo-Pasivo
- **A** = Relación booleana de tipo AND

Simulación y Resultados

Metodología de la Simulación

- **Fundamentada en 2 conceptos:**
  - **Generador uniforme** de números aleatorios
    - ✓ Caracterización definida por variable aleatoria e independiente para cada componente.
  - **Muestreo sistemático** de las variables aleatorias (simulación MonteCarlo)
    - ✓ La aleatoriedad en el modelo permite la repetitividad de las simulaciones.
    - ✓ Permite revelar patrones de comportamiento y la evaluación de función objetivo.

- **Metodología de Simulación:**
  - **Espacio temporal analizado de 1 año.**
  - **Nivel de granularidad de 1 minuto (525.600 eventos por simulación).**
    - ✓ Cada minuto se determina el estado operativo o fallido de cada componente.
    - ✓ La operatividad o fallo vendrá dada por:
      - ✓ Características inherentes del componente (muestreo de su variable aleatoria asociada a su tasa de fallos).
      - ✓ Errores de operación (muestreo de la variable aleatoria de su tasa de fallos)
      - ✓ Relaciones con el resto de componentes del sistema de acuerdo al álgebra de Boole (árbol de fallos).
  - **La disponibilidad del sistema** vendrá dada como consecuencia de las **relaciones y disponibilidades de los componentes de más bajo nivel**.

Simulación y Resultados

Caracterización de la Simulación

- **La caracterización de los componentes es clave para la evaluación de la viabilidad del modelo**
  - Los valores **definen el comportamiento real de los componentes** a evaluar.
  - La **opción preferente** siempre es:
    - ✓ **Histórico de comportamiento** en mismo entorno/instalación/cliente
    - ✓ Información proporcionada por el **fabricante** (hardware)
  - **En ausencia de la opción preferente**:
    - ✓ Experiencia de **comportamiento en entornos similares**.
    - ✓ **Objetivos que se desean cumplir** (válido para diseño de servicios).

- **Hardware**
  - MTBF = 5000 hrs
  - MTTR = 5.5 hrs
  - Sigma = 1.9 hrs
  - Monitorización = 0.7 (70% fallos detectados)
  - Errores Operación = 2/año
- **Usuario**
  - Tiempo reacción del usuario tras indisponibilidad del servicio = 30 min

- **Software**
  - #defectos causantes de indisponibilidad = 15
  - MTBF = 500 hrs
  - MTTR = 2.2 hrs
  - Sigma = 1.5 hrs
  - Monitorización = 0.5 (50% fallos detectados)
  - Errores Operación = 2/año
  - Periodicidad de upgrades Sw = 60 días

- **Clúster**
  - Fiabilidad failover = 0.7
  - MTTR failover manual= 1.5 hrs
  - Sigma MTTR = 0.8 hrs
- **Calendarios**
  - Reparación = 7 x 24 x 365
  - Usuario = L-V 09:00-18:00

Simulación y Resultados

Realización y Productos de la Simulación

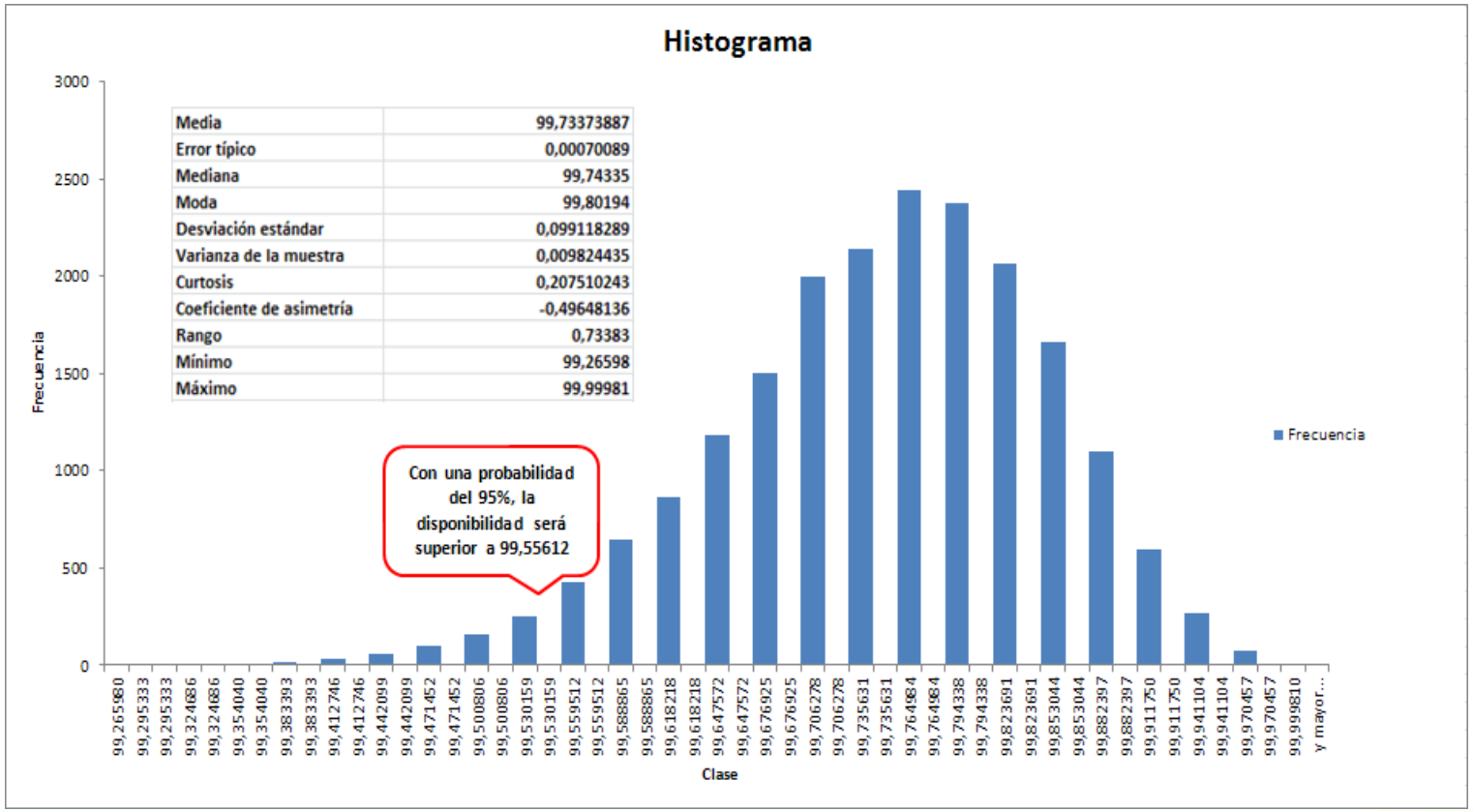
- **Realización de 20.000 simulaciones del modelo del Clúster A-P**
  - Se garantiza la **unicidad de comportamiento** en cada componente
  - En cada simulación se usa un **conjunto de semillas distinto**
  - Cada simulación registra un fichero con los eventos producidos.

```
….
Wed, Feb 21, 05:28: Sw_Component_1 has failed due to operator error!
Wed, Feb 21, 05:28: My network monitor informs that Sw_Component_1 has failed
Wed, Feb 21, 05:28: Sw_Component_1 is scheduled for repair
Wed, Feb 21, 05:28: Sw_Component_1: repair action will take 71 minutes
Wed, Feb 21, 05:28: my state has changed: Cluster_Node_1 is down
Wed, Feb 21, 05:28: cluster active node Cluster_Node_2 still operating
Wed, Feb 21, 06:39: Sw_Component_1 has been repaired.
Wed, Feb 21, 06:39: my state has changed: Cluster_Node_1 is up
Wed, Feb 21, 06:39: cluster AppCluster is fully recovered
Wed, Feb 21, 06:39: Scheduling a software upgrade for Sw_Component_1
…
```

Simulación y Resultados

Resultados de la Simulación

- **Disponibilidad del Clúster de Aplicación (medido sobre el Calendario de Usuario)**
  - **Disponibilidad media** = 99,733%
  - **Compromiso con Clientes** → Con probabilidad=95%, disponibilidad > 99,55612%



Análisis de Sensibilidad

Escenarios Contemplados

- **Permite la evaluación de distintas variaciones (escenarios) sobre el modelo con el fin de evaluar el impacto sobre los resultados de disponibilidad obtenidos.**

- **<u>Escenarios Evaluados</u>:**

  - **Escenario I: Incremento de la Formación del Personal Operador**
    - ✓ Tanto a nivel Hw como Sw, con el objetivo de reducir los errores de operación a 0.

  - **Escenario II: Sustitución de las Plataformas HW (MTBF 50% Superior) Y Nuevo Acuerdo SLA con Fabricante (MTTR 50% Inferior)**
    - ✓ Este escenario contempla el uso de nuevas máquinas con un MTBF superior (7500 hrs).
    - ✓ El establecimiento de un acuerdo SLA más restrictivo con el fabricante para reducir los tiempos de recuperación a la mitad (pasar de un MTTR=5,5 hrs a MTTR=2.75 hrs).

  - **Escenario III: Incremento de la Fase de Pruebas SW hasta MTBF 50% superior y reducir el número de defectos mayores un 66%**
    - ✓ Este escenario contempla incrementar la duración de las pruebas hasta obtener un MTBF del software de 750 hrs (incremento de 250 hrs del MTBF).
    - ✓ El numero de defectos causantes de indisponibilidad se reducen de 15 a 5.

  - **Escenario IV: Sustitución del Sistema de Monitorización por otro de Mayor Sensibilidad (detección del 95%)**
    - ✓ En este escenario, la granularidad de detección de fallos se incrementa de un 70% a un 95%.

Análisis de Sensibilidad

Análisis Comparativo de Resultados y Valoraciones

- **El Escenario IV es aquel que proporciona un mayor porcentaje de mejora de disponibilidad.**
  - **5,88 horas adicionales al año** de disponibilidad de la aplicación para el usuario.
  - **Business Case**
    - Porcentaje de mejora
    - Cuantificación económica de la implementación de cada escenario.
    - Impacto en negocio derivada de la indisponibilidad (p.ej. Pérdida por minuto de indisponibilidad en App venta online).
  - **Herramienta de Planificación de Negocio, Orientación de las Inversiones, y Soporte a la Toma de Decisiones**

Escenario	Disponibilidad Mínima (con probabilidad 95%)	Mejora Disponibilidad (%)	MEJORA (Disponibilidad para Usuario)		
			En minutos	En horas	En dias
Modelo Original	99,55612%	-	-	-	-
Escenario I	99,62671%	0,07059%	99,10836	1,651806	0,068825
Escenario II	99,59988%	0,04376%	61,43904	1,023984	0,042666
Escenario III	99,66153%	0,10541%	147,99564	2,466594	0,102775
Escenario IV	99,80765%	0,25153%	353,14812	5,885802	0,245242

Conclusiones y Líneas Futuras
Conclusiones

❑ Investigación y Conocimiento

- Análisis y estudio de la disponibilidad y fiabilidad de los sistemas
- Técnicas RAMS y aplicación de actividades a sistemas TIC
- Modelización de la fiabilidad software
- Estudio de modelos estadísticos y probabilísticos
- Habilitación de nuevas líneas de estudio, evolución, trabajo futuro, y aplicabilidad.

❑ Innovación

- Modelización estadística de componentes reutilizables de simulación (lenguaje Java)
- Herramienta de estudio y análisis sistemático basada en MonteCarlo
- Análisis cuantitativo de la disponibilidad de sistemas TIC

❑ Análisis de Sensibilidad

- Viabilidad como solución de diseño y optimización de sistemas sin impacto en negocio
- Herramienta de soporte para la toma de decisiones y orientación de las inversiones

Conclusiones y Líneas Futuras
Líneas Futuras

- Estudio de la Aplicabilidad de Nuevos Modelos Estadísticos
- Modelización de Conceptos de Indisponibilidad por Incapacidad y Calidad de Servicio (QoS)
- Incorporación de Métricas de Tráfico como Elemento del Modelo (transacciones de servicio)
- Ampliación del Modelo de Componentes
- Corrección Imperfecta de Defectos

Muchas Gracias

Sergio Gonzalo San José
Email: sgonzalos@uoc.edu
www.uoc.es