

PONTIFICIA UNIVERSIDAD JAVERIANA FACULTAD DE INGENIERIA DEPARTAMENTO INGENIERIA DE SISTEMAS

Proyecto Introducción a Sistemas Distribuidos

Período Académico 2024-10

Sistema

Distribuido: Cloud, Fog, Edge

Objetivos del Proyecto

- Desarrollar una solución a un problema de estructura distribuida
- Utilizar patrones de comunicación síncronos y asíncronos.
- Resolver problemas presentes en sistemas distribuidos, tales como fallas en los componentes y persistencia de datos.
- Reconocer atributos de calidad (e.g. desempeño, resiliencia) asociados a la implementación de un sistema distribuido



Contexto

Debido al cambio climático y en mayor medida, a las malas prácticas (intencionales y no intencionales) por parte de algunos ciudadanos, se han incrementado los incendios forestales en nuestro país. Los efectos son devastadores ya que acaban con la vida silvestre y la vegetación, los suelos pueden tardar siglos en recuperarse, se agotan las fuentes de agua, se aumenta la contaminación, etc. Por lo tanto, es importante monitorear de forma continua el estado de los suelos en zonas de bosque con el fin de detectar aumentos de

temperatura y cambios en las condiciones del ambiente para controlar a tiempo un posible incendio. Algunos de los factores que se deben monitorear son: temperatura del aire (oscila entre 11° y 29,4°), presencia de humo (Valor booleano) y humedad relativa (promedio 70%, una humedad inferior es indicativo de alguna anomalía).

En este proyecto se propone realizar un sistema distribuido utilizando los conceptos de Edge, Fog y Cloud computing ¹, en el que se simula la tolerancia a fallas y la captura y análisis de los tres parámetros previamente mencionados, de una zona boscosa de los cerros orientales de Bogotá. La arquitectura sugerida es la siguiente:

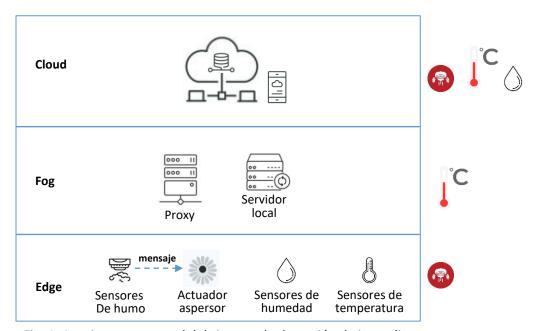


Fig. 1. Arquitectura general del sistema de detección de incendios

Descripción del Sistema a Desarrollar

A continuación, se explica cada uno de los componentes de la arquitectura (figura 1):

Capa Edge Computing

Sensores de humo: Se debe implementar un sensor de humo y levantar 10 instancias de sensores de humo, cada uno tomando muestras cada 3 segundos, las cuales se deben enviar a la capa de Fog Computing. El sensor de humo es un componente reactivo dentro de la arquitectura, es decir, tan pronto detecta una señal de humo el sensor realizará las siguientes acciones: a) Envía un mensaje directo a un sensor actuador que activará un aspersor. b) Envía el mensaje de alerta

¹ Cloud, Fog, and Edge Computing: A Software Engineering Perspective. doi: 10.1109/COMAPP.2018.8460443. url: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8460443&isnumber=8460181

a la nube, a través del proxy de la capa *Fog Computing* (en el cloud se almacena toda la información sobre las alertas en las distintas capas para fines estadísticos) c) Genera al Sistema de Calidad que se encuentra en esta capa.

- **Actuador aspersor**: Es suficiente con la implementación de un único actuador. Si recibe un mensaje de un sensor de humo, debe activarse inmediatamente (representado en un mensaje de texto).
- Sensores de temperatura: Se debe implementar un sensor de temperatura y activar 10 instancias de este sensor. Cada sensor estará tomando muestras cada 6 segundos; estas mediciones se envían al Proxy correspondiente en la siguiente capa (Fog).
- **Sensores de humedad**: Como antes, se debe implementar un sensor de humedad y activar 10 instancias del mismo, con una toma de muestras cada 5 segundos. Estos datos se envían el Proxy de la siguiente capa.

En la tabla 1 se observan valores válidos para cada variable a medir.

ParámetroValor mínimoValor máximoHumo (Variable booleana)FalsoVerdaderoTemperatura11°29,4°Humedad70%100%

Tabla 1: Rangos de los parámetros de calidad

Los sensores se simulan con hilos y al activar cada sensor, éstos deben recibir los siguientes argumentos: **Tipo de sensor y un archivo de configuración para cada uno.** En el archivo de configuración se encontrarán tres probabilidades: probabilidad de producir valores dentro del rango, probabilidad de producir valores fuera del rango y probabilidad de producir valores erróneos (para el caso del proyecto serán valores negativos). Los valores fuera del rango deben generar alertas. Por ejemplo, si el archivo de configuración para los sensores de temperatura contiene las siguientes probabilidades:

- 0,6 valores correctos
- 0,3 valores fuera del rango
- 0,1 errores

De 10 medidas producidas por el sensor de temperatura, 6 (60%) de ellas serán correctas, 3 (30%) estarán fuera del rango y 1 (10%) será invalida, por ejemplo, para el sensor de temperatura:

25 15 -1 35 29 20 27 10 17 40

En el caso del sensor de humo, solo se producirán los valores verdadero y falso y valores erróneos.

Cada sensor simulará las medidas produciendo números aleatorios que atiendan las probabilidades descritas en el archivo de configuración. Dichas medidas se enviarán al Proxy utilizando los sockets (PUSH-PULL) cada **t** unidades de tiempo, según se explicó para cada tipo de sensor. Cada medición debe ir acompañada del tiempo (u hora) en la cual se produjo.

Capa Fog Computing

Proxy: se encarga de la comunicación hacia la nube, es decir, todas las mediciones directas, cálculos y alertas generados en las capas *Edge* o *Fog* se envían a la nube a través del componente proxy. Al recibir los datos de los sensores, el proxy valida que no contengan errores

En el Fog computing, la temperatura se calcula con un promedio simple de los 10 sensores; este promedio se debe publicar por pantalla junto con la fecha de los valores tomados. En caso de que la temperatura supere el valor máximo del rango establecido para tal indicador (ver tabla), se debe activar una alerta enviando o un mensaje al sistema de calidad de la capa Fog. También debe enviar un SMS, correo, etc. a la nube para fines de almacenamiento.

El cálculo diario de la humedad se hace de acuerdo con la siguiente fórmula:

 HRd_{ijh} Es la humedad relativa del aire del día h, registrada en el sensor i, durante el mes j. Este indicador se calcula cada 5 segundos promediando el valor de los 10 sensores de humedad

El proxy reenvía las medidas calculadas de la humedad a la capa cloud donde se calculará la humedad relativa mensual.

Capa Cloud Computing

Todas las alarmas detectadas por los sensores deben ser comunicadas y almacenadas en el *cloud* con propósitos estadísticos. En el caso especial de las medidas de humedad, estas se reciben del proxy y cada 20 segundos se calcularán las muestras mensuales. La humedad relativa se calcula con la siguiente fórmula:

$$HRm_{ij} = \frac{\sum_{h=1}^{n} HRd_{ijh}}{n}$$

• HRm_{ij} Es la humedad relativa media por mes (j) del aire, registrada en el sensor (i)

- HRd_{ijh} Es la humedad relativa del aire del día h, registrada en el sensor i, durante el mes.
 Valor que viene de la capa Fog.
- n: Es el número de días registrados en el mes

Cada 20 segundos se calcula el resultado de los 4 resultados diarias para obtener la suma mensual

Si el promedio calculado de la humedad mensual está por debajo del límite, se genera una alerta que va al sistema de calidad de esta capa.

Sistema de Calidad (SC)

La única función de este proceso es esperar las posibles alarmas que se generan en cada una de las capas e imprimir las alertas por pantalla. Debe existir un SC en cada nivel (*Edge, fog, cloud*)

Detalles de Implementación

Comunicación: la comunicación entre los sensores y proxies se debe hacer con el modelo Pipeline de la librería ZeroMQ (siguiendo el patrón muchos a uno de los workers al sink). La comunicación entre los componentes de cada capa y su sistema de calidad se realizará utilizando el patrón **request-reply**. El sistema de calidad debe responder al emisor del mensaje, únicamente después de haber publicado la alerta por el canal correspondiente. La comunicación entre el proxy y la capa cloud se hace siguiendo un patrón request-reply.

Fallas: debe implementar una arquitectura tolerante a fallas, considerando una posible falla del componente proxy. Los estudiantes deberán contar con un respaldo que realice las mismas funciones, el cual estará en ubicado una máquina diferente (Dirección IP diferente). Debe existir también otro proceso encargado de chequear que el proxy está o no en funcionamiento (health check). Cuando se detecte que deja de funcionar, el proceso réplica debe asumir la carga. La reconexión del nuevo proceso y distribución de cargas debe hacerse de forma automática, es decir, será transparente para el cliente; las operaciones deben continuar en el punto en que se habían dejado y no iniciar desde el principio. En el informe final debe indicar que patrón de tolerancia a fallas utilizó.

Necesario para la evaluación: el día de la sustentación es importante que se pueda observar:

- Operaciones que van realizando cada uno de los procesos y resultado de la operación, si aplica.
- Información guardada en la BD (o en un archivo).
- El funcionamiento del proyecto distribuyendo los componentes en al menos tres computadoras (o máquinas virtuales): sensores en un computador, procesos proxy principal y proxy de respaldo en computadores diferentes. Los componentes de la capa cloud en otro computador.

Rendimiento del Sistema

Variables para medir el rendimiento del sistema

- 1. Cantidad de mensajes enviados (o tamaño en bytes enviado) en todo el sistema durante su funcionamiento. Tanto el total de mensajes, como el total enviado de una capa a otra.
- 2. Tiempos de comunicación entre el proxy y el nivel cloud. En este caso los estudiantes deberán mostrar:
 - a. Tiempos promedio de comunicación entre las dos capas. Desviación estándar asociada al promedio.
- 3. Cantidad total de alertas generadas por cada indicador.

Para la primera entrega los equipos deberán definir el procedimiento y las herramientas que utilizarán para valorar las métricas de desempeño citadas.

Elementos que afectan el rendimiento del sistema

En la segunda y última entrega del proyecto los equipos medirán los dos tipos de métricas citados anteriormente (a y b), en los siguientes 3 escenarios:

Escenario 1	Escenario 2	Escenario 3
Las probabilidades generadas por los sensores de humedad y temperatura se encuentran en los siguientes rangos: 0,5 valores correctos 0,4 valores fuera del rango 0,1 Errores	Las medidas de temperatura, desde los sensores, se envían cada 4 segundos.	Se mantienen las periodicidades del escenario 2 y probabilidades del escenario 1, pero se produce una falla en el proxy del nivel fog.

Cada grupo tomará medidas con los tres escenarios y colocará el valor de las variables en tablas. Adicionalmente, realizará gráficos que permitan observar cómo es el cambio de las métricas en los tres escenarios planteados. A la luz de los resultados obtenidos, cada grupo realizará un análisis y concluirá sobre la necesidad (o no) de realizar mejoras en el diseño del sistema para una detección de alarmas más eficiente y para un mejor uso de recursos (en este caso el ancho de banda de la red).

A continuación, <u>sugerimos</u> algunas herramientas de monitoreo y generación de carga:

Para generar Carga:

Apache Jmeter, Grafana k6

Para obtener el valor de las métricas de rendimiento

- Comandos del sistema operativo
- Scalene (https://github.com/plasma-umass/scalene)
- Visualvm (https://visualvm.github.io/)
- UBER JVM PROFILER (https://www.uber.com/en-CO/blog/jvm-profiler/)
- Pyroscope (https://pyroscope.io/)
- Jensor (https://jensor.sourceforge.net/)

Entregas y Rubrica de Evaluación

PRIMFRA FNTRFGA

La primera entrega se realizará en la semana 10 a través de la plataforma. Las sustentaciones serán el día correspondiente a la asignatura de acuerdo con el grupo.

La primera entrega consta de un informe donde se debe especificar:

- Modelos del sistema (Arquitectónico, interacción, fallos y seguridad)
- Diseño del sistema: Diagrama de despliegue, Diagrama de componentes, Diagrama de clases y diagrama de secuencia. Este diseño debe incluir el o los componentes para enmascarar las fallas del sistema.
- El protocolo de pruebas que utilizará para la entrega final (considere **todos** los tipos de prueba que deben realizarse a un sistema)
- Describa cómo va a obtener las métricas de desempeño: paquetes enviados y tiempos de publicación de las alarmas.
- Implementación inicial: deben tener implementada al menos la generación de mediciones desde los sensores y su recepción por parte de los procesos proxy. Esto debe estar funcionando al menos en dos computadoras.

Se deben implementar los requerimientos de acuerdo con el enunciado y la rúbrica.

El día de la sustentación, cada equipo tendrá 10 minutos para mostrar sus resultados y responder las preguntas que tengan los profesores.

SFGUNDA FNTRFGA

La entrega se realizará en la semana 17 a través de la plataforma. Las sustentaciones se harán los días que correspondan para cada curso. El día de la sustentación los integrantes del equipo deben **mostrar la funcionalidad del proyecto**. Deben estar presentes todos los integrantes del grupo.

La entrega se compone de:

- En un archivo .zip código fuente de los programas que conforman el sistema y un archivo Readme donde indique cómo ejecutarlo.
- Se debe complementar la documentación de la primera entrega
- Un video de máximo 10 minutos donde muestre:
 - La arquitectura implementada, distribución de los componentes entre máquinas y patrones de comunicación.
 - Tecnologías utilizadas para implementar el sistema y realizar las mediciones.

- o Cómo implementó la arquitectura confiable (tolerante a fallas)
- Un informe de máximo 4 páginas donde explican: experimentos realizados (especificaciones de hw y sw donde se realizaron las medidas, variables e instrumentos de medición), resultados obtenidos (acompañados por gráficos y tablas), conclusiones.
- Adicionalmente se le puede pedir que traiga el día de la sustentación las diapositivas que usó en su video.

Calificación I Entrega

Valor 15%, evaluado sobre 5pts

Indicador	Valoración en puntos /5pts	Excelente	Competente	Deficiente
Informe	0.5	0.5	0.25	<0.25
Diseño del	1.5	[1.5,1]	(1, 0.25]	< 0.25
Proyecto				
Protocolo de	0.75	[0.75,0.5]	(0.5, 0.25]	< 0.25
pruebas				
Modelos del	0.5	0.5	0.25	<0.25
sistema				
Obtención de las	0.25	0.25	0.15	<0.15
métricas de				
rendimiento				
Implementación	1.5	[1.5,1]	(1, 0.5]	< 0.5
Inicial				
total	5ptos			

Indicador	Excelente	Competente	Deficiente
Informe	La presentación del informe es impecable, sin problemas de ortografía o redacción	El informe tiene fallas menores en la presentación, ortografía y/o redacción	El informe tiene fallas importantes en la presentación, ortografía y/o redacción
Diseño del Proyecto	Se presentan todos los artefactos de diseño exigidos en el enunciado: diagrama de componentes,	Los diagramas incompletos o algunos diagramas están incorrectos.	No se realizan los diagramas exigidos en el enunciado o los diagramas presentados están incorrectos. No se considera TODO el

	diagrama de clases, diagrama de secuencia y diagrama de despliegue. Se incluyen todos los componentes del sistema final, incluyendo tolerancia a fallas y persistencia. Los diagramas están correctos.		sistema que se va a implementar en el diseño.
Protocolo de pruebas	En el protocolo presentado por los estudiantes, las pruebas son suficientes para evaluar la funcionalidad de la aplicación con y sin la presencia de fallas.	En el protocolo presentado no se contemplan todas las pruebas importantes. Está incompleto.	Los estudiantes no presentan protocolo de pruebas funcionales.
Modelos de sistema	Se hace una descripción de los modelos fundamentales adaptados al proyecto.	Se describen parcialmente los modelos fundamentales adaptados al problema a resolver.	No se presenta ni se describen los modelos fundamentales o la descripción no está relacionada con el proyecto a resolver.
Obtención de las métricas de rendimiento	Se describe de forma clara y completa las herramientas y metodología para obtener el valor de las métricas.	En el informe falta alguno de los siguientes puntos o están mencionados parcialmente: - Herramientas de medición que van a usar - Metodología para obtener el valor de las métricas.	No se mencionan en el informe ni las herramientas ni la metodología para obtener el valor de las métricas de rendimiento.
Implementación Inicial	Se implementa la generación de los tres tipos de mediciones desde los sensores. Se puede evidenciar que cada proxy	Se implementa la generación de las medidas desde los sensores, pero de forma parcial, es decir, se observa alguna de las siguientes deficiencias (máximo 2 de 4):	La implementación inicial es bien deficiente, se observan falencias en 3 ó 4 de los puntos

	recibe la nformación a la	a)	El sistema solo funciona en un computador	mencionados en la celda anterior (a-d)
L	que está suscrito. Los procesos se nvocan de la	b)	Los procesos no se levantan con los parámetros solicitados.	
	forma como fue especificado.	c)	No está correctamente implementado el patrón publicador/suscriptor.	
		d)	No se implementa correctamente la generación de medidas desde ellos sensores (atendiendo las	
			probabilidades sugeridas)	

Calificación II Entrega

Valor 25%

Informe de rendimiento 10% (Evaluado sobre 5ptos)

Resto de la entrega (corrida, sustentación, etc) 15%. Evaluado sobre 5ptos

Rúbricas

Informe de Rendimiento

	Valoración en puntos	Excelente	Competente	Deficiente
Informe: presentación. Ortografía y redacción	1	[1,0.75]	(0.75, 0.5]	< 0.5
Presentación de los datos: datos presentados en tablas y gráficos. Se grafican las variables solicitadas, los gráficos están construidos correctamente.	2	[2,1]	(1, 0.5]	< 0.5
Análisis de los resultados: el	2	[2,1]	(1, 0.5]	< 0.5

grupo comenta los			
resultados de			
interés y deriva			
conclusiones			
cónsonas con			
estos resultados			
	5		

Descripción

	Valoración en puntos	Excelente	Competente	Deficiente
Informe: presentación. Ortografía y redacción	1	La presentación del informe es impecable sin problemas de ortografía o redacción	El informe tiene fallas menores en la presentación, ortografía y/o redacción	El informe tiene fallas importantes en la presentación, ortografía y/o redacción
Presentación de los datos: datos presentados en tablas y gráficos.	2	Se grafican todas las variables solicitadas, los gráficos están construidos correctamente (títulos, unidades, etc)	Faltan algunos gráficos o, están completos, pero faltan elementos que contribuyen a su comprensión: qué se está graficando, cuáles son las unidades de medida, etc.	No se realizan gráficos o se realizan muy pocos gráficos. La presentación de los gráficos es bastante deficiente.
Análisis de los resultados: el grupo comenta los resultados de interés y deriva conclusiones cónsonas con estos resultados	2	Los estudiantes comentan los resultados obtenidos y presentan conclusiones sobre la arquitectura del sistema y los patrones de comunicación.	Los estudiantes comentan los resultados obtenidos pero no presentan conclusiones sobre la arquitectura del sistema y los patrones de comunicación.	Los estudiantes no comentan los resultados obtenidos o los comentarios son muy deficientes.

Indicador	Valoración en puntos /30	Excelente	Competente	Deficiente
Capa Edge:	0.5	0.5	0.25	< 0.25
Implementación				
de los sensores,				
actuador, sistema				
de calidad				
Capa Fodge:	0.75	[0.75, 0.5]	(0.5, 0.25]	< 0.25
funcionamiento				
del proxy y				
sistema de calidad				
Capa cloud: alertas	0.75	[0.75, 0.5]	(0.5, 0.25]	< 0.25
de humedad,				
sistema de calidad,				
persistencia				
Corrida en 3	0.75	[0.75, 0.5]	(0.5, 0.25]	< 0.25
máquinas				
Tratamiento de	0.5	0.5	0.25	< 0.25
Fallas del Proxy				
Código	0.25	0.25	(0.25, 0.15]	< 0.15
Protocolo de	0.5	0.5	0.25	< 0.25
pruebas				
funcionales				
Sustentación,	1	[1,0.75]	(0.75, 0.5]	< 0.5
video				
Total	5pts			

Descripción

Capa Edge:	Los sensores y el	Los sensores están	Los sensores están
Implementación	actuador están	implementados	implementados, pero
de los Sensores,	implementados	correctamente y, siguen	hay diferencias
actuador,	correctamente y según	la mayoría de las	importantes entre las
Sistema de	las especificaciones	especificaciones del	especificaciones del
calidad	del enunciado. Se	enunciado. En la	enunciado y lo que
	puede evidenciar la	ejecución se evidenció	implementaron los
	corrida de los 10	alguna de las siguientes	estudiantes. La
	sensores de cada tipo.	fallas (máximo 2):	información no llega a
	Se disparan las	 No corrieron los 	los proxies de la forma
	alarmas	10 sensores de	correcta.
	correctamente.	cada tipo.	
		- No funcionó el	
		actuador	
		correctamente.	

	1	I	
Capa Fog: Proxy Sistema de Calidad	El proxy recibe la información adecuada de los sensores, almacena los datos y genera las alarmas correspondientes. Se envía de forma correcta la información a la siguiente capa. El proxy se implementa según las especificaciones del enunciado.	- No se disparan las alarmas correctamente Las medidas no se producen con la probabilidad especificada Los sensores no se inician según las especificaciones del enunciado. El proxt recibe la información correctamente de los sensores. No obstante, alguno de los siguientes aspectos está parcial, incorrecto o incompleto: - Generación de alarmas - Argumentos de inicialización de este tipo de procesos Envío de información a la siguiente capa.	La implementación del Proxy es deficiente en más de uno de los siguientes aspectos: - Recepción de las medidas Generación de alarmas - Argumentos de inicialización de este tipo de procesos Envío de información a la siguiente capa.
Capa cloud	El sistema recibe las medidas de las capas anteriores, detecta las alarmas, realiza los cálculos correspondientes y se implementa la persistencia.	La capa se implementa, pero 1 ó 2 de las siguientes funcionalidades no está del todo correcta: - Tratamiento de alarmas - Cálculos mensuales de los diferentes indicadores - Persistencia - Manejo de alarmas.	La capa se implementa, pero hay deficiencias en tres o cuatro de las siguientes funcionalidades: - Tratamiento de alarmas - Cálculos mensuales de los diferentes indicadores - Persistencia - Manejo de Alarmas

Corrida en 3 máquinas	El sistema funciona en al menos tres	El sistema funciona solo en dos computadoras	Todos los componentes del sistema se instalan y
maqamas	computadoras (o	y/o máquinas virtuales	funcionan en una sola
	máquinas virtuales)		computadora.
	según la arquitectura		
	sugerida en el enunciado del		
	proyecto		
Tratamiento de	Ante la falla del proxy	Ante la falla del proxy se	No se implementa la
Fallas del proxy	la inicialización de un	levanta un proceso	tolerancia a fallas de
l amae act proxy	proceso homólogo	homólogo que asume	este componente.
	funciona de forma	de forma correcta sus	
	correcta. La falla se	funciones. La falla no se	
	detecta de forma	detecta de forma	
	automática usando un	automática. Los	
	patrón, como por	estudiantes no explican	
	ejemplo, el health	claramente los patrones	
	check. Los estudiantes	de resiliencia	
	explican el o los	implementados.	
	patrones de resiliencia implementados		
Código	El código está bien	El código es funcional	El código es deficiente,
Coulgo	estructurado, bien	pero no está	sin orden ni
	indentado y	debidamente	documentación
	documentado	documentado	
Protocolo de	Se hacen todas las	Faltaron algunas	Las pruebas funcionales
pruebas	pruebas necesarias	pruebas funcionales	son insuficientes.
funcionales	para evaluar la		
	funcionalidad del		
	sistema		
Sustentación,	Se siguen todas las	Hay alguna falla menor	Hay <i>fallas</i> importantes
video	reglas dadas para la	en uno o dos de los	en 3 o más <i>de los</i>
	sustentación, se	siguientes elementos:	siguientes elementos:
	presentan todos los	- Seguimiento de	- Seguimiento de
	elementos	las reglas dadas	las reglas dadas
	importantes del	para la	para la
	sistema y se	sustentación	sustentación
	responden de forma	- Presentación de	- Presentación de
	adecuada todas las	las	las
	inquietudes del	funcionalidades	funcionalidades
	evaluador. Se entrega el video en	del sistema	del sistema - Respuesta
	el que se evidencia	- Respuesta correcta a todas	correcta a todas
	claramente la	las preguntas	las preguntas
	arquitectura de la	del evaluador.	del evaluador.
	aplicación, su	- El video no	
	funcionamiento y los	ilustra de	

elementos de	manera	- El video es
tolerancia a fallas.	suficiente la	deficiente o no
	arquitectura, el	se adjunta.
	funcionamiento	
	de la aplicación	
	o la tolerancia	
	de las fallas.	