



***Diseño de una arquitectura de aviónica tolerante a fallas basada
en componentes COTS para vehículos satelitales de nueva
generación***

Por ***Arias Emmanuel***

Presentado ante la Universidad Nacional de La Matanza y la Unidad de Formación Superior de la CONAE
como parte de los requerimientos para la obtención del grado de

MAGISTER EN DESARROLLOS INFORMATICOS DE APLICACION ESPACIAL

UNIVERSIDAD NACIONAL DE LA MATANZA

Buenos Aires, 2018

©UFS-CONAE 2018

©UNLAM 2018

DIRECTOR

Gustavo Wiman

INVAP, Bariloche, Provincia de Rio Negro

Parte de los resultados obtenidos en el presente trabajo fueron presentados en el *Congreso Argentino de Tecnologías Espaciales 2017 (Córdoba, Argentina)*, bajo el título: *Estudio de la confiabilidad de arquitecturas tolerantes a fallas basada en componentes COTS para aviónicas de vehículos espaciales*

*A mi madre,
gracias a tus palabras
mirá a dónde llegué. Y no me detendré.*

Abstract

The space missions development involves costs of major magnitude, the most important costs being the development, and the materials used for the manufacture of satellites. The high cost of these materials is because they are manufactured exclusively for their application to space activity, that mean that they are "qualified to fly." There are other components called COTS (Commercial Off-The-Shelf). These COTS components are considerably more economical than the classic ones, thus helping to significantly reduce costs. This is important to the Argentinian space industry. To achieve a correct application of these components, it is necessary develop techniques, strategies and architectures that ensure that the probability of catastrophic failure and degradation are compatible with the mission. In this work, we firstly studied and proposed models of different topologies that will be used to develop a fault tolerant architecture based on COTS components. Then, a communication protocol called CANae is developed. CANae is based on the CAN (Controller Area Network) protocol, and CANae is oriented to work in distributed networks. Finally, in this thesis be proposed an architecture based on both COTS components and distributed networks, that uses the protocol CANae for the communication with its components. In this way, an architectural proposal is developed to ensures that the system is reliable, even though the components that make it up are of low reliability.

Keywords: COTS Components, Fault Tolerant Architecture, Avionics, Fault Tolerance, Satellites

Resumen

El desarrollo de proyectos satelitales conlleva costos de importante magnitud, y dependen de cada misión. Estos costos pueden ser clasificados, en forma general, en 5 grupos:

- Desarrollo
- Materiales
- Ensamblado, integración, y tests
- Lanzamiento
- Operaciones

Una parte importante de los costos está conformado por el desarrollo, y sobre todo, por los materiales que se utilizan para su fabricación. Esto es debido a que se utilizan componentes que son exclusivos para el ámbito espacial, en otras palabras que se encuentran “calificados para volar”. Estos componentes son fabricados especialmente para soportar el ambiente hostil del espacio.

No se puede mencionar a ciencia cierta cuál es el costo “verdadero” de desarrollar un satélite. Este depende exclusivamente del tipo de satélite y de la misión. Lo que sí se debe tener en claro es que las tareas de desarrollo representan una parte muy importante del costo total del proyecto. Por tal motivo, este trabajo de tesis se centrará principalmente en el desarrollo (proceso de planificación, análisis, diseño e implementación.), y en los materiales utilizados en la fabricación de vehículos satelitales.

El desafío de este trabajo de tesis es analizar y estudiar arquitecturas que sean tolerantes a fallas, que permitan una correcta comunicación entre los diferentes subsistemas de un vehículo espacial de nueva generación, y que tenga como característica principal un cierto grado de confiabilidad, de modo tal que pueda ser aplicado con componentes Commercial Off-The-Shelf (COTS). Desarrollar un vehículo espacial con componente COTS, en un principio podría representar costos adicionales, ya que se le deben realizar tareas de calificación adicional, debido a que no están “preparados” para resistir las condiciones hostiles del espacio.

Estos componentes COTS tiene varias ventajas sobre los “clásicos”. Uno de sus puntos positivos, es que a la hora de desarrollar varios satélites en base a la misma ingeniería, se puede ahorrar en gran medida en los materiales que se utilizan. Los componentes COTS suelen tener un costo de compra hasta 1000 veces menores que aquellos que están calificados para volar. **Esto ayudaría a ahorrar millones de dólares de los proyectos satelitales.**

Otra de las ventajas de utilizar componentes COTS, es que la mayoría cuentan con una tecnología más avanzada que aquellos que son calificados para volar. Esta tecnología permite:

-
- Aumentar prestaciones, mediante el incremento de las capacidades de procesamiento, memoria, velocidades de procesamiento, etc.
 - Implementar funciones que son imposibles de aplicar en tecnologías viejas.
 - Reducir tiempos de desarrollo.
 - Reducir volumen, masa y consumo

Uno de los puntos en contra de la utilización de componentes COTS es que al no ser calificados para volar, es necesario llevar a cabo tareas y estrategias inteligentes, con el fin de hacer frente a esa “deficiencia”. Por ello, se exige realizar una investigación y análisis de diferentes arquitecturas de aviónica, que puedan ser utilizadas para lograr que el sistema sea tolerante a fallas, y así, cumplir con los requerimientos de una misión satelital.

En esta tesis se desarrolla un protocolo de comunicación basada en CAN, que se lo llamó CANae. Este, es el protocolo utilizado en la arquitectura tolerante a fallas que se propone en este trabajo, el cual está basado en redes distribuidas. Todos los diseños están orientados a modelos, y se utiliza el lenguaje de modelado SysML para su desarrollo y documentación.

Agradecimientos

En primer lugar, me gustaría agradecer a mi madre, que gracias a sus palabras de aliento, que me han acompañado en la distancia y a lo largo de estos dos años, hoy puedo estar concretando esta meta. Sinceramente, si no fuese por ella, hoy estaría persiguiendo otros objetivos. Desde chiquito me ha inculcado a que nunca debo bajar los brazos, aún cuando todo parezca que está perdido. Siempre hay que luchar hasta el final. Es por eso que hoy estoy aquí. Gracias mamá.

En segundo lugar, quiero agradecer a mi novia la Prof. Dra. Bioq. Brandán Yamila por su comprensión, su compañía, su cariño, su amor y sus consejos. Además, quisiera destacar (enormemente) que me guió y ayudó en el desarrollo de esta tesis. Llevó a cabo, de excelente manera, el papel de revisora y asesora de esta tesis. Gracias amor.

Por último lugar, quisiera agradecer a mis compañeros de maestría, Alfonso, Cecilia, Eduardo, Elbio, Ezequiel, Javier, José, Pablo E., Pablo S. y Ricardo, por su amistad y por compartir sus conocimientos. Gracias.

Organizacion de la tesis

Este trabajo de tesis está organizado de manera tal que, el lector logre comprender cuál es la problemática que se plantea el autor, conozca el sustento teórico sobre los cuales la tesis se apoya, y con ello, y de manera natural, se logre entender el significado y la importancia de la propuesta de la arquitectura diseñada en esta tesis.

El Capítulo 1 (página 2) se centra en que el lector entienda la motivación (Sección 1.1 página 3) que impulsan este trabajo de investigación y desarrollo, como así también las hipótesis (Sección 1.2 página 4), objetivos (Sección 1.3.1 página 5), objetivos específicos (Sección 1.3.2 página 5) y las preguntas de investigación (Sección 1.3.3 página 5) que el autor se propone resolver.

En el Capítulo 2 (página 7) se introduce al lector a la teoría en la cual se basa este trabajo de tesis. Entre las secciones más importantes que se verán se puede mencionar, que en la Sección 2.1 (página 8) se explica brevemente la terminología que se utiliza a lo largo de este trabajo. Es importante leer esta sección ya que algunos términos pueden considerarse sinónimos en el habla común, pero en esta tesis, tienen significados bien diferenciados. En la Sección 2.2 (página 10) explica qué es la fiabilidad aplicada en sistemas (de software) y cómo se puede clasificar la fiabilidad. En Sección 2.3 (página 11) se detalla el concepto de los impedimentos de la fiabilidad, cuáles son los orígenes de las fallas, los modos comunes de fallas y las fallas en el software. La tolerancia a fallas, el núcleo de este trabajo, se explica en la Sección 2.6 (página 15). En esta sección se detalla el significado de la tolerancia fallas, importante para entender este trabajo de tesis. En la Sección 2.12 (página 29) se brinda un resumen teórico de diferentes protocolos de comunicación existentes en la actualidad. En la Sección 2.16 (página 32) se detalla el protocolo CAN ya que la arquitectura propuesta en este trabajo de tesis utiliza un protocolo basado en CAN. En esta sección se describe la capa física (página 34), capa de enlace (página 35), formato del mensaje (página 35) y se comenta sobre un protocolo basado en CAN que es utilizado en aplicaciones de aeronáutica llamado CANAerospace (página 38)

En el Capítulo 3 (página 41) se presenta el estado del arte del área de trabajo, sobre la cual se desplaza esta tesis. Esto se realiza para que el lector conozca los avances que se llevaron a cabo en la temática.

En el Capítulo 4 (página 53) se lleva a cabo el estudio de la confiabilidad de tres tipos de topologías tolerantes a fallas (Sección 4.3 página 55) candidatas a ser utilizada en el diseño de la arquitectura propuesta en este trabajo de tesis. Los resultados de este estudio se presentaron en el *Congreso Argentino de Tecnología Espacial 2017*. En la Sección 4.6 (página 60) se describe resumidamente el protocolo CANae 0.1 Alpha, que está basado en CAN y que fue desarrollado en este trabajo de tesis.

En el Capítulo 5 (página 64) se presenta la arquitectura propuesta. Aquí se analizan los requerimientos (Sección 5.1 página 64), Casos de Uso (Sección 5.2 página 67), el diseño estructural (Sección 5.3 página 69) y el diseño dinámico (Sección 5.4 página 73).

En el Capítulo 6 (página 83) se presentan los resultados y conclusiones del trabajo.

En el apéndice A (página 91) se describe la especificación técnica del protocolo de comunicación CANae 0.1 Alpha, el cual es un producto de esta tesis.

En el apéndice B (página 124) se especifican los Casos de Uso desarrollados en el Capítulo 5 Sección 5.2.

En el apéndice C (página 132) tiene como propósito que el lector conozca el crecimiento en las investigaciones llevadas a cabo en el área de estudio de la presente tesis. En este apartado, se exponen gráficos para visualizar la cantidad de artículos científicos que se han publicado desde 1970. Para este fin, se utilizó la plataforma *Scopus*¹

¹www.scopus.com

Tabla de Contenidos

1. Introducción	2
1.1. Motivación	3
1.2. Hipótesis	4
1.3. Objetivo del trabajo y preguntas de investigación	5
1.3.1. Objetivo	5
1.3.2. Objetivos Específicos	5
1.3.3. Preguntas de investigación	5
2. Marco Teórico	7
2.1. Terminología	8
2.2. La fiabilidad en el software	10
2.3. Impedimentos de la confiabilidad	11
2.3.1. Orígenes de la falla	11
2.3.2. Modos comunes de fallas	11
2.3.3. Falla de causa común	12
2.3.4. Fallas en el Software	12
2.4. Medios de fiabilidad	12
2.4.1. Evitación de Fallas	13
2.4.2. Tolerancia a Fallas	13
2.4.3. Eliminación de Fallas	13

TABLA DE CONTENIDOS

2.4.4. Predicción de Fallas	13
2.5. Atributos de la fiabilidad	14
2.5.1. Confiabilidad	14
2.5.2. Disponibilidad	14
2.5.3. Seguridad	15
2.6. Tolerancia a falla	15
2.7. Redundancia en el software	17
2.7.1. Técnicas single version	17
2.7.2. Técnicas multi-version	18
2.7.3. Técnicas de detección de fallas	18
2.7.4. Técnicas de recuperación de fallas	19
2.8. Técnica de evaluación de fiabilidad	22
2.9. Medidas comunes de fiabilidad	22
2.9.1. Failure rate	22
2.9.2. Tiempo medio de falla	24
2.9.3. Tiempo medio de reparación	24
2.9.4. Tiempo medio entre fallas	25
2.9.5. Cobertura de fallas	25
2.10. Métodos de cálculos de fiabilidad	25
2.10.1. Diagramas de bloques de confiabilidad	25
2.10.2. Utilización de procesos de Markov	26
2.11. Modelos de falla	26
2.11.1. Tiempo hasta la falla	27
2.11.2. Función de confiabilidad	27
2.11.3. Tasa de falla	27
2.11.4. Tiempo medio hasta la falla	28
2.11.5. Vida restante media	28
2.11.6. Distribución binomial	28
2.11.7. Distribución exponencial	29

TABLA DE CONTENIDOS

2.12. Protocolos de comunicación de tiempo real	29
2.12.1. Sistemas de tiempo real	29
2.13. Estrategia de acceso al medio	30
2.13.1. CSMA	30
2.13.2. TDMA	30
2.13.3. Minislotting	31
2.14. Revisión de protocolos	31
2.14.1. CAN	31
2.14.2. byteflight	31
2.14.3. ARINC 659 o SAFEbus	31
2.14.4. TTP/C	31
2.15. Posibles fallas en una red	32
2.16. Protocolo CAN	32
2.16.1. Consideraciones previas	32
2.16.2. Introducción	32
2.16.3. Elementos necesarios	33
2.16.4. Capa física	34
2.16.5. Capa de Enlace	35
2.16.6. Formato del mensaje	35
2.16.7. CANAerospace	38
3. Estado del arte	41
3.1. Árboles binarios	41
3.1.1. Esquema de árbol binario con backups	42
3.2. Sistemas Hypercube	44
3.3. Redes distribuidas	44
3.3.1. Algoritmo de ruteo	45
3.4. Redes Ethernet en aviónica	45
3.4.1. Experiencia de vuelo	47

TABLA DE CONTENIDOS

3.5. Arquitectura de red basada en BUS	47
3.5.1. Evaluación de la confiabilidad de arquitecturas basadas en BUS	48
3.6. Métrica y modelado de la confiabilidad de sistemas	48
3.7. CAN en la actividad espacial	50
4. Análisis y desarrollo de arquitectura tolerante a fallas	53
4.1. Requerimientos para el análisis	53
4.2. Nomenclatura	54
4.3. Estudio de topologías de arquitecturas	55
4.3.1. Árbol binario	55
4.3.2. Red distribuida	56
4.3.3. Red hypercube	58
4.4. Topología utilizada en la arquitectura a diseñar	59
4.5. Topología propuesta	60
4.6. Protocolo de comunicación	60
5. Arquitectura propuesta	64
5.1. Árbol de requerimientos	64
5.2. Casos de Uso	67
5.3. Diseño Estructural	69
5.3.1. Bridge tolerante a fallas	69
5.3.2. Bloques Internos	72
5.3.3. Nodo Monitor	72
5.4. Modelo dinámico	73
5.4.1. Máquina de estado	73
5.4.2. Diagrama de secuencia	73
5.4.3. Diagrama de actividades	74
6. Conclusión	83
6.1. Conclusiones finales	85

TABLA DE CONTENIDOS

6.2. Perspectivas a futuros	86
Bibliografía	90
A. Protocolo de comunicación: CANae 0.1 Alpha Version	91
A.1. Introducción	91
A.1.1. Revisiones	91
A.2. Capa de aplicación	92
A.3. Tipos de servicios de la capa de aplicación	92
A.4. CANae Application Layer	93
A.5. CMS	94
A.5.1. Prioridades de los objetos	94
A.5.2. Objeto CMS	94
A.5.3. Servicios de CMS	94
A.5.4. Eventos	96
A.6. NMT	97
A.6.1. Objetos NMT	98
A.6.2. Servicios NMT	98
A.6.3. Protocolos NMT	100
A.7. DBT	103
A.7.1. Objetos y servicios de DBT	104
A.7.2. Descripción de los servicios DBT	104
A.7.3. Objetos DBT	105
A.7.4. Servicios DBT	105
A.7.5. Protocolos DBT	105
A.8. Gestor de mensajes	108
A.8.1. Receiver Manager	108
A.8.2. Prepared Message	109
A.8.3. Sorter Message	110
A.8.4. Buffer Messages	110

TABLA DE CONTENIDOS

A.8.5. Buffer Message Receive y Buffer Message To Send	111
A.9. Gestor de Nodo	111
A.9.1. Tabla primaria y secundaria de ruteo	112
A.9.2. Servicios	113
A.10.Formato de mensajes	114
A.10.1. Start of Frame	115
A.10.2. Priority	115
A.10.3. Node-ID	115
A.10.4. Remote transmission request (RTR)	115
A.10.5. Type of Message (TOM)	115
A.10.6. Bit reservado	115
A.10.7. Data length	115
A.10.8. Data field	116
A.10.9. CRC	116
A.10.10.CRC delimiter	116
A.10.11.ACK	116
A.10.12.ACK delimiter	116
A.10.13.End-of-Frame (EOF)	116
A.11.High Application Layer CANae	117
A.11.1. Network Management	117
A.11.2. Task Management	118
A.11.3. List of Tasks	120
A.11.4. Task	121
A.12.Arquitectura completa del CANae	121
A.13.Estados de nodos y red	121
B. Especificación de Casos de Uso	124
B.1. Enviar Mensaje	124
B.2. Recibir Mensaje	125

TABLA DE CONTENIDOS

B.3. Procesar Mensaje	126
B.4. Rutear Mensaje	126
B.5. Gestionar Red	127
B.6. Gestionar Tareas	128
B.7. Dividir Tareas	129
B.8. FDIR	129
B.9. Reconfigurar Arquitectura	130
C. Antecedentes	132

Índice de figuras

2.1. Fiabilidad (Hitt y Mulcare, 2001)	10
2.2. Representación de checkpoint y restart	20
2.3. Representación del proceso pares	21
2.4. Failure rate de HW vs tiempo	23
2.5. Failure rate SW vs tiempo	23
2.6. Confiabilidad vs tiempo	24
2.7. Arquitectura estándar propuesta por ISO11898	33
2.8. Tráfico en el BUS CAN	34
2.9. Frame de mensaje del CAN estándar	37
2.10. Frame del mensaje del CAN extendido	37
2.11. Formato del mensaje de CANAerospace	39
3.1. Árbol binario de 4 niveles	42
3.2. Confiabilidad con respecto al tiempo de una arquitectura de árbol binario de 4 niveles	43
3.3. Arquitectura básica TTEthernet	46
3.4. Frame de mensaje TTEthernet	46
3.5. Arquitecturas stack-trees	47
3.6. Arquitecturas que no responden al modelo stack-trees	47
3.7. Esquema CST_D	48
3.8. Modelo de sistema satelital (Hoque <i>et al.</i> , 2015)	50

ÍNDICE DE FIGURAS

4.1. Confiabilidad con respecto al tiempo de árbol binario de 4 niveles	55
4.2. Red distribuida	56
4.3. Confiabilidad de red distribuida	57
4.4. Confiabilidad de red distribuida con 4 nodos fallando	57
4.5. Red Hypercube	58
4.6. Confiabilidad de red hypercube	58
4.7. Comparación de confiabilidad	59
4.8. Arquitectura propuesta utilizando topología de red distribuida	61
4.9. Conexión entre la red y el subsistema	61
4.10. Estructura de la capa de aplicación de CANae en alto nivel	62
5.1. Diagrama de requerimientos de la arquitectura propuesta	66
5.2. Diagrama de Casos de Uso General de la arquitectura Propuesta	67
5.3. Diagrama de Casos de Uso de arquitectura propuesta	68
5.4. Diagrama de bloques de la arquitectura completa	70
5.5. Posibilidad de interconexión de nodos. Configuración 1.	71
5.6. Posibilidad de interconexión de nodos. Configuración 2.	71
5.7. Conexión entre la red y el subsistema	72
5.8. Diagrama de bloques interno del nodo.	76
5.9. Máquina de estado de la arquitectura completa	77
5.10. Máquina de estado de los nodos.	77
5.11. Diagrama de secuencia del inicio de la arquitectura.	78
5.12. Diagrama de secuencia de CA	78
5.13. Crear Objeto Red, Objeto Nodo y Objeto Nodo Remoto	79
5.14. Preparar nodo para la conexión a la red CAN	79
5.15. Preparar nodo para la conexión a la red CAN (Start Node)	79
5.16. Conectar el nodo a la red CAN	80
5.17. Diagrama de actividades del inicio de nodos	80
5.18. Diagrama de actividades del Nodo monitor	81

5.19. Diagrama de actividades del envío de mensajes	82
6.1. Comparación de confiabilidad	84
A.1. Estructura de la capa de aplicación de CANae en alto nivel	92
A.2. CANae Application Layer	93
A.3. Definición del Objeto CMS	95
A.4. Definición de bloques internos de la entidad NMT	97
A.5. Definición de la entidad NMT	98
A.6. Protocolo NMT	101
A.7. Crear Red	101
A.8. Crear Objeto Red, Objeto Nodo y Objeto Nodo Remoto	102
A.9. Preparar nodo para la conexión a la red CAN	102
A.10. Conectar el nodo a la red CAN	102
A.11. Comenzar la comunicación de nodos	103
A.12. Arquitectura de la entidad DBT	104
A.13. Protocolo create_node_table_configuration	106
A.14. Protocolo enable_distribution	106
A.15. Protocolo disable_distribution	107
A.16. Protocolo create_node_definition	107
A.17. Protocolo delete_node_definition	107
A.18. Protocolo get_checksum	108
A.19. Arquitectura de <i>Message Management</i>	109
A.20. Diagrama interno del <i>Message Management</i>	109
A.21. Arquitectura del <i>Node Management</i>	111
A.22. Frame de datos CANae	114
A.23. Diagrama de bloques del High Application Layer de CANae	117
A.24. Diagrama de bloques internos del High Application Layer de CANae	118
A.25. Diagrama de interacción del comportamiento del <i>Network Management</i>	119
A.26. Diagrama de interacción del comportamiento del <i>Network Management en caso de errores</i>	119

ÍNDICE DE FIGURAS

A.27. Diagrama de secuencia de la ejecución de tareas del Task Management	120
A.28. Diagrama de secuencia para detener procesos por parte del Task Management	120
A.29. Arquitectura completa del CANae	122
A.30. Definición de la entidad NMT	123
A.31. Definición de la entidad NMT	123
C.1. Cantidad de investigaciones con palabra clave “COTS satellite”	132
C.2. Cantidad de investigaciones con palabra clave “COTS satellite” por países	133
C.3. Cantidad de investigaciones con palabra clave “Fault Tolerance Satellite”	134
C.4. Cantidad de investigaciones con palabra clave “COTS Fault Tolerance”	134
C.5. Cantidad de investigaciones con palabra clave “COTS Fault Tolerance” por países	135
C.6. Cantidad de investigaciones con palabra clave “Fault Tolerance Satellite Architecture”	135
C.7. Cantidad de investigaciones con palabra clave “Fault Tolerance Satellite Architecture” por países	136

Índice de tablas

2.1. Disponibilidad en relación con su baja de servicio por año. Tabla modificada de Dubrova (2013)	15
4.1. Comparación de confiabilidad de topologías	60
5.1. Tabla de Requerimientos	65
A.1. Revisiones de CANae	91
A.2. Prioridad eventos	97
A.3. Prioridades de mensajes	115

Lista de acrónimos

SW	Software
HW	Hardware
FT	Tolerancia a Fallas
FA	Evitación de Fallas
FR	Eliminación de Fallas
FF	Predicción de Fallas
MTBF	Tiempo Medio Entre Fallas
MTTF	Tiempo Medio de Fallas
MTTR	Tiempo Medio de Reparación
CMF	Modo Común de fallas
FDIR	Detección, Aislación y Recuperación de Fallas
CONAE	Comisión Nacional de Actividades Espaciales
UNLAM	Universidad Nacional de La Matanza
INVAP	Investigación Aplicada
NASA	National Aeronautics and Space Administration
COTS	Commercial Off-The-Shelf
TT	Time-Trigged
CST	Complete Stack-Tree
CAN	Controller Area Network
CPU	Central Processing Unit

Introducción

Nunca me rendiría, tendría que estar muerto o incapacitado.

Elon Musk

Este capítulo se centra en que el lector logre entender los propósitos y los fundamentos de este trabajo. Se brinda el detalle de la motivación que impulsaron este trabajo de tesis (Sección 1.1 página 3). Luego, se presenta las hipótesis de este trabajo (Sección 1.2 página 4), el objetivo de este trabajo (Sección 1.3.1 página 5), los objetivos específicos (Sección 1.3.2 página 5) y las preguntas de investigación (1.3.3 página 5) que deben responderse luego de finalizado este trabajo.

En el marco del Plan Espacial Nacional, desarrollado por la Comisión Nacional de Actividades Espaciales (CONAE) de Argentina, y con el propósito de llevar a cabo actividades de investigación y aplicación, provenientes de la Universidad Nacional de La Matanza (UNLAM) se presenta esta tesis con el fin de ampliar los conocimientos y la participación de la CONAE y UNLAM, en el campo del Desarrollo Informático y Ciencias de la Computación.

Las actividades desarrolladas para este trabajo de tesis fueron realizadas, en su mayor proporción, en la Unidad de Desarrollo Investigación Aplicada (INVAP), ubicada en San Carlos de Bariloche, Provincia de Río Negro. Este trabajo se encuentra orientado a brindar un nuevo conocimiento, que ayude en cierta medida, en el desarrollo de los diferentes proyectos con los que cuenta actualmente esta empresa, agregando un grado de innovación en el resultado que se obtenga.

INVAP tiene como visión ser un referente en proyectos tecnológicos a nivel internacional/global (INVAP Sociedad del Estado, 2016), por lo tanto, debe asegurarse que cada uno de los productos que se lleven a cabo

dentro de la empresa sean competitivos. Para lograr cumplir con esto, es necesario que tales proyectos se encuentren a la vanguardia tecnológica y científica.

El desarrollo de proyectos satelitales conlleva costos de importante magnitud, que dependen de cada misión. Una parte importante de los costos está conformado por el desarrollo¹ y sobre todo los materiales que se utilizan para su fabricación. Esto es debido a que emplean componentes que son exclusivos para el ámbito espacial, en otras palabras que se encuentran “calificados para volar” que son fabricados especialmente para soportar el ambiente hostil del espacio.

Si se considera al ámbito espacial como una industria, algo que ha sido demostrado en los últimos años; y si se tiene en cuenta las intenciones de crecimiento y competitividad de la empresa INVAP y de permitir el ingreso de nuestro país en el mercado satelital (INVAP Sociedad del Estado, 2016), resulta de gran importancia lograr reducir los costos en fabricación y desarrollo de vehículos satelitales.

La National Aeronautics and Space Administration (NASA) tiene un enfoque de desarrollo bajo el lema “faster, cheaper, better” Forsberg y Harold (1999), lo cual busca desarrollar sus proyectos y misiones de forma rápida, barata y mejor. Bajo este enfoque se han realizado diversos estudios e investigaciones dando resultados sumamente positivos ((Tai *et al.*, 1999); (Chau *et al.*, 1999); (Schneidewind y Nikora, 1998); (Forsberg y Harold, 1999)). En estos trabajos se utilizan componentes que no se encuentran “calificados para volar”, los cuales también son llamados componentes COTS, o de estantería. Debe mencionarse, que entre los resultados de estas misiones también hubo fracasos en su aplicación.

A simple vista, la utilización de estos componentes ayudaría a reducir costos. Sin embargo, esto no es tan directo. Los componentes COTS al no estar calificados, requieren de tareas adicionales de “calificación”. Además deben ser aplicados a un ambiente, que asegure que no fallarán durante la misión; o si fallan, no será motivo de pérdida de la misma.

Los componentes COTS suelen tener un costo hasta 1000 veces menores que los componentes tradicionales. Por lo que el aumento en la utilización de estos componentes, aplicados al desarrollo de diferentes tipos de satélite, **permitiría reducir los costos y ahorrar millones de dólares del proyecto satelital**. Esto facilitaría el ingreso de Argentina en un mercado altamente competitivo y costoso.

El desafío de este trabajo de tesis es analizar y estudiar arquitecturas que sean tolerantes a fallas, que permitan una correcta comunicación entre los diferentes subsistemas de un vehículo espacial de nueva generación, y que tenga como característica principal un cierto grado de confiabilidad, de modo tal que pueda ser aplicado con componentes COTS.

1.1. Motivación

Los costos de un proyecto satelital se pueden clasificar, a grandes rasgos, en 5 grupos:

- Desarrollo
- Materiales
- Ensamblado, integración, y tests
- Lanzamiento
- Operaciones

¹Nota: entiéndase por desarrollo al proceso de planificación, análisis, diseño e implementación.

Este trabajo de tesis se centrará principalmente en el desarrollo (proceso de planificación, análisis, diseño e implementación.), y en los materiales utilizados en la fabricación de vehículos satelitales.

No se puede mencionar a ciencia cierta cuál es el costo “verdadero” de desarrollar un satélite. Este depende exclusivamente del tipo de satélite y de la misión. Lo que sí se debe tener en claro es que las tareas de desarrollo representan una parte muy importante del costo total del proyecto.

Desarrollar un vehículo espacial con componente COTS, en un principio podría representar costos adicionales, ya que se le deben realizar tareas de calificación adicional, debido a que no están “preparados” para resistir las condiciones hostiles del espacio.

Uno de los puntos positivos, y que motivan la aplicación de componentes COTS, es que a la hora de desarrollar varios satélites en base a la misma ingeniería, se puede ahorrar en gran medida en los materiales que se utilizan. Los componentes COTS suelen tener un costo de compra hasta 1000 veces menores que aquellos que están calificados para volar. **Esto ayudaría a ahorrar millones de dólares de los proyectos satelitales.**

Otra de las ventajas de utilizar componentes COTS, es que la mayoría cuentan con una tecnología más avanzada que aquellos que son calificados para volar. Esta tecnología permite:

- Aumentar prestaciones, mediante el incremento de las capacidades de procesamiento, memoria, velocidades de procesamiento, etc.
- Implementar funciones que son imposibles de aplicar en tecnologías viejas.
- Reducir tiempos de desarrollo.
- Reducir volumen, masa y consumo

El último punto mencionado anteriormente es de especial interés, ya que al reducir volumen y masa, permite reducir costos adicionales como el de lanzamiento.

Esta reducción de costos tienen ventajas directas a la hora de introducir a Argentina en un mercado altamente competitivo, donde la mínima reducción de estos, representa ganancias económicas importantes.

Uno de los puntos en contra de la utilización de componentes COTS es que al no ser calificados para volar, es necesario llevar a cabo tareas y estrategias inteligentes, con el fin de hacer frente a esa “deficiencia”. Por ello, se exige realizar una investigación y análisis de diferentes arquitecturas de aviónica, que puedan ser utilizadas para lograr que el sistema sea tolerante a fallas, y así, cumplir con los requerimientos de una misión satelital.

El estudio de arquitecturas tolerantes a fallas, no solamente tiene aplicación en el ámbito espacial, si no que también puede ser extendido a cualquier sistema crítico, los cuales necesitan ser robustos y tolerantes a fallas, como es el caso de aviones comerciales, plantas nucleares, automóviles, etc.

1.2. Hipótesis

En la presente tesis se propone la siguiente hipótesis: “Una arquitectura de aviónica basadas en componentes COTS, robusta y tolerante a fallas, es totalmente aplicable y utilizable en vehículos espaciales, con un alto nivel de confiabilidad, lo cual permite disminuir la complejidad de los sistemas actuales de aviónica”.

1.3. Objetivo del trabajo y preguntas de investigación

En esta sección se mencionan el objetivo principal y específicos; como así también las preguntas de investigación, que sirvieron de guía para la elaboración de esta tesis.

1.3.1. Objetivo

El objetivo de este trabajo es investigar y analizar arquitecturas de comunicación de los subsistemas de aviónica tolerante a fallas basada en componentes COTS para vehículos satelitales de nueva generación.

1.3.2. Objetivos Específicos

1. Realizar un estudio del estado de la cuestión sobre arquitecturas tolerantes a fallas para sistemas críticos.
2. Investigar y analizar arquitecturas tolerantes a fallas que aseguren la confiabilidad del sistema y que sean aplicables en la industria satelital.
3. Investigar y analizar protocolos de comunicación, para las capas superiores del modelo de OSI (modelo de interconexión de sistemas abiertos - ISO/IEC 7498-1), orientados a la tolerancia a fallas y confiabilidad de los sistemas. Realizar un estudio comparativo de los diferentes protocolos estudiados.
4. Investigar una metodología para lograr una medición de la tolerancia a fallas en arquitecturas de aviónica.
5. Desarrollar un estudio comparativo de arquitecturas tolerantes a fallas con el fin de obtener ventajas y desventajas de cada una de ellas.
6. Diseñar modelos alternativos de arquitecturas tolerantes a fallas, que tenga un grado de confiabilidad tal que permita la aplicación de componentes COTS.
7. Evaluar la confiabilidad de los modelos de arquitecturas (mediante métrica desarrollada en este trabajo o siguiendo otras estrategias).
8. Proponer el diseño de una nueva arquitectura tolerante a fallas, con un grado de confiabilidad suficiente para la aplicación de componentes COTS en aviónicas de vehículos satelitales.
9. Simular la arquitectura planteada para medir su grado de tolerancia a fallas y performance.

1.3.3. Preguntas de investigación

Para este trabajo de tesis se plantearon las siguientes preguntas de investigación, que se fueron respondiendo a lo largo de este trabajo

- ¿Es posible la realización de un método de medición del grado de tolerancia a fallas de una arquitectura de aviónica?
- ¿Cuál es la estrategia más indicada de tolerancia a fallas que permita brindar un alto grado de confiabilidad en la utilización de componentes COTS en sistemas críticos?
- ¿Cuál es la arquitectura más indicada que permita desarrollar tolerancia a fallas en sistemas críticos basados en componentes COTS?

- ¿Es factible la utilización de componentes COTS en sistemas espaciales?

Resumen

Este trabajo de tesis surge de la motivación de la Unidad de Desarrollo INVAP a introducirse en el mercado internacional de la industria espacial. Para ello, INVAP debe minimizar los costos y a la vez maximizar la confiabilidad de sus vehículos espaciales. Por tal motivo, centra su atención en los componentes COTS que les permitirá reducir considerablemente los costos de desarrollo, por otro lado, aumentar la performance de sus satélites. Para lograr introducir estos componentes en la fabricación de satélites, es necesario llevar a cabo nuevas técnicas que aseguren que el sistema tenga la confiabilidad requerida en el área espacial, aún cuando sus componentes sean de baja confiabilidad.

Marco Teórico

Solo sé que no sé nada

Platón

En este capítulo se introduce al lector a la teoría en la cual se basa este trabajo de tesis. En la Sección 2.1 (página 8) se explica brevemente la terminología que se utiliza a lo largo de este trabajo. Es importante leer esta sección ya que algunos términos pueden considerarse sinónimos en el habla común, pero en esta tesis, tienen significados bien diferenciados.

En la Sección 2.2 (página 10) explica qué es la fiabilidad aplicada en sistemas (de software) y cómo se puede clasificar la fiabilidad. En Sección 2.3 (página 11) se detalla el concepto de los impedimentos de la fiabilidad, cuáles son los orígenes de las fallas, los modos comunes de fallas y las fallas en el software.

En la Sección 2.4 (página 12) se explican los medios de la fiabilidad, es decir, los medios para lograr la fiabilidad de sistemas. Aquí se ven los conceptos de confiabilidad, disponibilidad y seguridad.

La tolerancia a fallas, el núcleo de este trabajo, se explica en la Sección 2.6 (página 15). En esta sección se detalla el significado de la tolerancia fallas, importante para entender este trabajo de tesis.

En la Sección 2.7 (página 17) se explican diferentes técnicas de redundancias en el software, las cuales son importantes durante el desarrollo de sistemas críticos.

Las técnicas de evaluación de fiabilidad (Sección 2.8 22) junto con las medidas de fiabilidad (Sección 2.9 22) exponen los fundamentos matemáticos en los cuales se basan los cálculos de tolerancia de fallas. Los métodos de cálculos de fiabilidad se presentan en la Sección 2.10 (página 25)

En la Sección 2.11 (página 26) se explican con más detalle conceptos importantes tales como, la función de confiabilidad, tasa de falla, tiempo medio hasta la falla, vida residual media.

En la Sección 2.12 (página 29) se brinda un resumen teórico de diferentes protocolos de comunicación existentes en la actualidad.

En la Sección 2.16 (página 32) se detalla el protocolo CAN que es utilizado como principal protocolo de comunicación. La arquitectura propuesta en este trabajo de tesis utiliza un protocolo de comunicación basado en CAN. En esta sección se describe capa física (página 34), capa de enlace (página 35), formato del mensaje

(página 35) y se comenta sobre un protocolo basado en CAN que es utilizado en aplicaciones de aeronáutica llamado CANAerospace (página 38)

2.1. Terminología

Existe una importante diferencia entre los significados de las palabras falla, error y avería¹, que es importante destacar. Teniendo en cuenta el Diccionario de Cambridge (Cambridge University, 2017)² el significado de fault es, “1. something that is wrong with something; 2. an imperfection, something wrong; 3. a mistake, especially something for which you are to blame”, que se adapta de mejor manera con la palabra española “falla”, que según la Real Academia Española (Real Academia Española, 2017)³ es, “1. defecto o falta; 2. incumplimiento de una obligación”. Por otro lado, Failure según Cambridge University (2017) es, “1. the fact of someone or something not succeeding; 2. the fact of not doing something that you must do or are expected to do”, que nuevamente se adapta mejor con la palabra española “avería”, la cual significa: “1. Daño que impide el funcionamiento de un aparato, instalación, vehículo, etc.” Por esto, en esta obra se utilizan las palabras: falla, error y avería, para las traducciones de fault, error y failure.

Una **avería** de sistema ocurre cuando el servicio prestado por el sistema ya no coincide con las especificaciones del mismo (Hanmer, 2007). Esto quiere decir que existe un problema que tiene una consecuencia negativa en el sistema completo, logrando que este ya no logre cumplir con sus especificaciones. Cuando el sistema no se comporta de la manera que es especificada, este ha fracasado. Esto significa que lo que se espera de un sistema se encuentra descripto, comúnmente en especificaciones o requerimientos (Pullum, 2001).

Para la IEEE (1990) avería es “la inhabilitación de una sistema o componente a llevar a cabo las funciones requeridas en los requerimientos específicos de performance del mismo”.

Hanmer (2007) ejemplifica averías de sistemas cuando: el sistema se bloquea y se detiene cuando no debería hacerlo, el sistema calcula un resultado incorrecto, el sistema no está disponible, o el sistema es incapaz de responder a la interacción con el usuario. Cuando el sistema no hace lo que debe hacer, el sistema ha fracasado. Las averías son detectados por los usuarios mientras usan el sistema.

Las averías son causados por los errores. Un **error** es una parte del estado del sistema que es susceptible de provocar un avería en el sistema. Un error que afecta al servicio, es una indicación de que una avería se ha producido (Hanmer, 2007). Un error se puede propagar, es decir dar a lugar otros errores (Pullum, 2001).

IEEE (1990) define error como “la diferencia entre un valor computado, observado o medido, con el valor verdadero, especificado o el teóricamente verdadero”.

Según Rani (2011) un error es la parte de estado del sistema que puede conducir a una avería, es una etapa intermedia entre fallas y averías.

Los errores se pueden clasificar en dos tipos: errores de tiempo y valores (Hanmer, 2007). Los errores de valores son aquellos que se manifiestan como valores discretos incorrectos o estados del sistema incorrecto. En cambio, los errores de tiempo pueden incluir aquellos que no cumplen con el total de las tareas.

¹En inglés: fault, error y failure.

²<https://dictionary.cambridge.org/>

³<http://www.rae.es/>

Hanmer (2007) especifica los siguientes casos más comunes de errores:

- Timing: existe una falta de sincronización en la comunicación de los procesos.
- Bucles infinitos: ejecución de un bucle sin detenerse, esto consume memoria, y la avería del sistema.
- Error de protocolo: errores en el flujo de comunicación ya que no coinciden los protocolos. Mensajes enviados en formato diferente, en tiempos diferentes, a lugares de sistemas incorrectos.
- Inconsistencia de datos: los errores son diferentes en diferentes lugares.
- Sobrecarga de sistema: el sistema es incapaz de hacer frente a la sobrecarga de actividades a la que es expuesta.

La causa adjudicada o la hipótesis de un error es una **falla**, también llamado “bugs”. Una **falla activa** es aquella que produce un error (Pullum, 2001). Una falla es un defecto que está presente en el sistema y que puede causar un error (Hanmer, 2007). Es la desviación actual de lo correcto (Hanmer, 2007) (Rani, 2011).

Según IEEE (1990) una falla es “un defecto en un dispositivo de hardware o componente; como por ejemplo un corto circuito o un cable cortado”. También realiza una segunda definición diciendo que falla es “un paso incorrecto, proceso, o definición de dato en un programa de computadora” (IEEE, 1990). Esta última afirmación es la que se usa en el ámbito de este trabajo.

Se puede indicar que un software contiene falla, si dado un conjunto de entradas, se obtienen resultados que son incorrectos. Una falla es una parte del software, y puede ser eliminada mediante la corrección de dicha parte. (Xie, 1991)

Se debe realizar la distinción entre fallas activas y pasivas. Las primeras son aquellas que han sido descubiertas por algún método, es decir, mediante testing o durante el funcionamiento del sistema. Si el sistema no es capaz de identificar la falla, aislar y recuperarse de esta, podría llegar a convertirse en una avería. Por otro lado, las fallas pasivas son aquellas que se encuentran dentro del sistema, pero todavía no ha sido detectada, es decir, que no causa ningún efecto observable eternamente (a nivel de sistema). Debe aclararse que una falla se manifieste o no, depende de lo que esté sucediendo en el sistema y su entorno.

Algunas fallas introducidas en el Software (SW) se detallan en Hanmer (2007), lo cual señala que pueden incluir:

- Especificaciones incorrecta de requerimientos
- Diseño incorrecto
- Errores de programación

Resumiendo, la principal diferencia entre falla (fault) y avería (failure) es que la primera, es la causa de la segunda. Debido a la existencia de una falla, las entradas del módulo (o sistema) software que contiene dicha falla, provocará como salida valores y/o resultados incorrectos, es decir, que no corresponden con los valores teóricos. Por otro lado una avería, es la consecuencia de una falla, y se refiere a cuando el sistema deja de responder o comportarse como debería hacerlo, desviándose de sus requerimientos.

Entonces, como lo indica Pullum (2001) con la tolerancia a fallas, lo que se busca es prevenir la avería mediante la “tolerancia” de fallas, las cuales son detectables cuando un error aparece. Las fallas son el motivo de errores y los errores son motivos de avería (Dubrova, 2013).

También se suele utilizar el término anomalía en las operaciones de vehículos espaciales para referirse a comportamientos anómalos o no esperados del sistema (Harlan y Lorenz, 2005)

En Dubrova (2013) se describe un ejemplo para diferenciar correctamente estos conceptos. Se considera el SW de una planta nuclear, en la cual existe una computadora que es responsable de controlar la temperatura, la presión y demás variables de interés para la seguridad del sistema. Se da el caso de que uno de los sensores detecta que la turbina principal se encuentra girando a una velocidad menor a la correcta. Esta falla hace que el sistema envíe una señal para aumentar su velocidad (error). Esto produce un exceso de velocidad en la turbina, lo cual tiene como consecuencia que la seguridad mecánica apague la turbina. En esta situación el sistema no está generando energía. Esto se considera un avería, porque el sistema no está entregando el servicio según lo establecido por los requerimientos. Pero es un avería salvable.

Otro concepto es el de **mantenibilidad**, esta es la capacidad de un sistema, bajo condiciones normales, de ser restaurado a un estado en el cual puede realizar sus funciones requeridas, cuando se realiza el mantenimiento (Rausand y Hoyland, 2004).

En secciones posteriores se ven los conceptos de confiabilidad, disponibilidad y seguridad (Sección 2.5.1, 2.5.2, 2.5.3, respectivamente).

2.2. La fiabilidad en el software

El objetivo final de la Tolerancia a Fallas (FT), es el desarrollo de un sistema fiable (Dubrova, 2013). Teniendo en cuenta que el SW que se encuentra dentro de las naves espaciales, como satélites, lanzadores, y sobre todo vehículos tripulados son críticos, ya que de ellos dependen el éxito o fracaso de una misión o la vida de seres humanos, y por lo tanto, se debe llevar a cabo un sistema fiable. La fiabilidad de un sistema es la capacidad del mismo de entregar a los usuarios un nivel deseado de servicio (Dubrova, 2013).

La fiabilidad también se la puede considerar como una propiedad global que permite justificar la confianza de los servicios de un sistema (Hitt y Mulcare, 2001). Por lo tanto, como lo indica Hitt y Mulcare (2001) la fiabilidad es un término amplio y cualitativo que está relacionado con atributos no funcionales (o “-ilities”), que buscan generar un sistema “ideal”, especialmente cuando su funcionamiento es crítico.

Como se muestra en la Figura 2.1 la consecuencia de la fiabilidad es la relación entre la evitación de fallos y la reducción de fallos, así como también la FT.

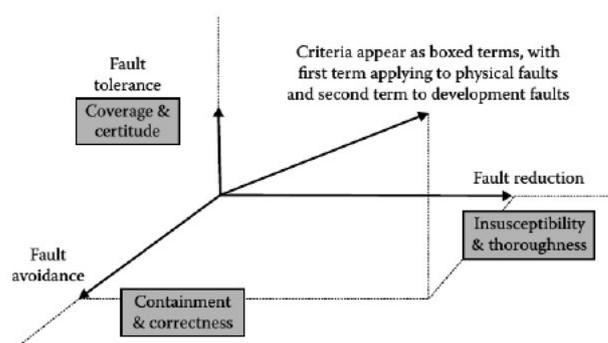


Figura 2.1: Fiabilidad (Hitt y Mulcare, 2001)

Según Pullum (2001) la fiabilidad puede ser clasificada en:

- Impedimentos: son aquellas cosas que se interponen en el camino de la fiabilidad. Son las fallas, errores y averías.
- Medios: los medios para lograr la fiabilidad, según el autor, se pueden dividir en dos grupos:

1. Aquellos que son utilizados durante la construcción del SW (Evitación de Fallas (FA)⁴ y FT).
 2. Aquellos que contribuyen con la validación del SW una vez desarrollado (Eliminación de Fallas (FR)⁵ y Predicción de Fallas (FF)⁶).
- Atributos: describen las propiedades de la fiabilidad y proporcionan una forma de evaluar el logro de esas propiedades.

2.3. Impedimentos de la confiabilidad

El impedimento de la confiabilidad o deterioro de la confiabilidad es definido en términos de fallas, errores, y avería (Dubrova, 2013). Los mismos fueron desarrollados en la sección 2.1. Lo que tienen en común estos tres conceptos, es que avisan, o dan alerta cuando algo está mal (Dubrova, 2013). La diferencia radica, en que las fallas son a nivel físico; los errores se dan a nivel computacional; mientras que las averías se dan a nivel de sistema (Dubrova, 2013).

2.3.1. Orígenes de la falla

Existen diversos orígenes de fallas. Estas pueden provenir desde terceros, en el caso de productos comprados, pueden deberse a una falta del conocimiento del problema, falta de tiempo, etc. Dubrova (2013) clasifica el origen de las fallas en cuatro grupos: *especificación incorrecta*, *implementación incorrecta*, *defectos de fabricación* y *factores externos*

Las *especificaciones incorrectas* son aquellas que surgen debidas a una incorrecta especificación de requerimientos o un mal diseño de una arquitectura o de un algoritmo (Dubrova, 2013). Estos orígenes de fallas son bastante comunes en el desarrollo de sistemas. Un ejemplo típico citado por Dubrova (2013), es el caso de requerimientos que ignoran aspectos del medio ambiente en el que opera el sistema. Una mala redacción de un requerimiento o el olvido de uno de ellos, puede traer graves problemas, atrasos y pérdida de dinero, en el diseño y producción de un sistema espacial.

Las *implementaciones incorrectas*, se refieren a las *fallas de diseño*, surgen cuando el sistema implementado no cumple con los requerimientos (Dubrova, 2013).

Otro origen de falla son los *defectos de los componentes* (Dubrova, 2013). Estos pueden incluir defectos de fabricación, defectos aleatorios dados en los componentes, etc.

Y por último se tienen las fallas que son causadas por *factores externos*, los cuales provienen del medio ambiente, usuarios u operadores (Dubrova, 2013). Ejemplos de estos factores externos pueden ser, vibraciones, cargas electrostáticas, temperatura, radiación electromagnética, envío incorrecto de comandos, etc.

2.3.2. Modos comunes de fallas

Un *Modo Común de fallas (CMF)*⁷ es una falla que ocurre simultáneamente en dos o más componentes redundantes (Dubrova, 2013).

Gangloff (1975) define los CMF como múltiples unidades de fracaso debido a una sola causa.

⁴En inglés, Fault Avoidance

⁵En inglés, Fault Removal

⁶En inglés, Fault Forecasting

⁷En inglés, common-mode faults

CMF son causados por fenómenos que crean dependencias entre unidades redundadas, lo que causa la falla de estas unidades simultáneamente (Dubrova, 2013).

Según como lo indica Dubrova (2013) el único enfoque para combatir los CMF, es mediante el diseño en diversidad. Diseño en diversidad es la implementación de más de una variante de la función en cuestión (Dubrova, 2013). Esto se puede lograr variando los algoritmos que se utilizan, diferentes equipos de trabajo realicen las mismas partes del sistema, de manera tal de tener redundancia en código, etc.

2.3.3. Falla de causa común

Una falla de causa común (CCF) se define como cualquier instancia donde múltiples unidades o elementos fallan debido a una causa común. Estos tipos de fallas pueden ocurrir debido al diseño de los equipos, catástrofes externas, fuentes de energías compartidas, la utilización de equipamientos provenientes de un mismo fabricante, etc. (Rani, 2011)

2.3.4. Fallas en el Software

El SW difiere en gran medida con el Hardware (HW). En primer lugar el SW no envejece, no se deforma, tampoco se puede quebrar ni ser afectado por el medio ambiente. Debido a que esta obra tiene como área de estudio el software de vuelo de satélites, se asume que el SW es determinístico, esto quiere decir que tendrá el mismo comportamiento, en las mismas circunstancias, al menos que exista un problema en el HW (provocado por el ambiente espacial, por ejemplo).

El SW es determinístico, siempre responde de la misma manera en el mismo ambiente, a menos que falle.

Por otro lado, el SW se lo puede actualizar varias veces a lo largo del su ciclo de vida.

En tercer lugar, arreglar bugs de SW **no significa que el mismo sea más confiable**, al contrario pueden ocurrir nuevos errores (Dubrova, 2013).

Por último el SW es mucho más complejo y menos regular que el HW. Tests tradicionales y métodos de debug pueden ser inadecuados para los sistemas de SW

2.4. Medios de fiabilidad

Los medios de confiabilidad son métodos y técnicas que permiten el desarrollo de un sistema confiable (Dubrova, 2013). Los medios se pueden dividir en dos grandes grupos (Pullum, 2001):

1. Aquellos que son empleados durante el proceso de construcción del SW,
2. y aquellos que ayudan en la validación del SW después que fue desarrollado.

Dentro del primer grupo se tiene:

- Evitación de Fallas
- Tolerancia a Fallas

Por otro lado, en el segundo grupo se puede mencionar los siguientes:

- Eliminación de Fallas
- Predicción de Fallas

2.4.1. Evitación de Fallas

FA son técnicas de mejoramiento de la fiabilidad utilizadas durante el desarrollo de SW para reducir el número de fallas introducidas durante la etapa mencionada (Pullum, 2001). Estas técnicas pueden estar presentes en las especificaciones y requerimientos del sistema, métodos de diseño de SW (Pullum, 2001).

Dubrova (2013) la denomina *prevención de fallas*⁸, y coincide con el autor anterior, definiendo FA como técnicas de control de calidad durante la especificación y fabricación de los procesos de diseño.

2.4.2. Tolerancia a Fallas

En la sección 2.6 (página 15) se discute con mayor detalle la FT.

2.4.3. Eliminación de Fallas

La FR hace referencia a las técnicas utilizadas para mejorar la fiabilidad empleadas durante la validación y verificación del SW (Pullum, 2001). Estas técnicas mejoran la fiabilidad del SW mediante la detección de fallas, usando métodos de verificación y la validación, y eliminando las fallas que se van detectando (Pullum, 2001).

Por otro lado Dubrova (2013) indica que el FR se lleva a cabo durante fases de desarrollo de SW tanto como durante el ciclo de vida de un sistema. Durante la fase de desarrollo, FR consiste en tres pasos: *verificación, diagnóstico y corrección* (Dubrova, 2013). FR durante la vida operacional de un sistema, consiste en el mantenimiento preventivo y correctivo del mismo (Dubrova, 2013).

2.4.4. Predicción de Fallas

La FF se realiza mediante la realización de una evaluación del comportamiento del sistema con respecto a la ocurrencia o la activación de una falla. Según Dubrova (2013) esta evaluación puede ser:

- Cualitativa: que tiene como objetivo clasificar los modos de fallas o combinaciones de eventos que llevan a la avería del sistema,
- Cuantitativa: que tiene como objetivo evaluar en término de probabilidad, el grado en el cual los atributos de fiabilidad son satisfechos.

FF incluye técnicas para aumentar la fiabilidad del sistema que son usados durante la validación del SW, con el objetivo de estimar la presencia de fallas y la ocurrencia o consecuencia de fracasos (Pullum, 2001)

⁸En inglés, Fault prevention

2.5. Atributos de la fiabilidad

Como se mencionó anteriormente, la fiabilidad de un sistema es la capacidad del mismo de entregar a los usuarios un nivel deseado de servicio (Dubrova, 2013).

La fiabilidad es una medida de calidad que abarca los conceptos de confiabilidad, disponibilidad y seguridad (Torres-Pomales, 2000). Estos se denominan atributos de la fiabilidad, los cuales son mencionados a continuación.

2.5.1. Confiabilidad

La confiabilidad es la probabilidad de que un sistema continúe operando correctamente durante un intervalo de tiempo dado (Torres-Pomales, 2000).

Dubrova (2013) coincide que la confiabilidad $R(t)$ ⁹ de un sistema es la probabilidad de que el sistema opere sin averías en el intervalo de tiempo $[0, t]$.

La confiabilidad es una medida de la entrega correcta del servicio que brinda un sistema (Dubrova, 2013).

En sistemas críticos como el SW de vehículo espacial, es sumamente necesario que tenga una alta tasa de confiabilidad, ya que por ejemplo, perder el contacto con la nave, podría representar la pérdida de la misión, o una gran cantidad de datos. Otro ejemplo que se puede mencionar es de un satélite geostacionario de comunicación, la pérdida de este servicio debe ser baja, casi nula (idealmente).

Coincidiendo, Pressman (2001) define la confiabilidad como la “probabilidad de tener operaciones libre de fallas de un programa de computadora, en un ambiente específico para un tiempo específico”.

La IEEE (1990) define confiabilidad como “La capacidad del sistema o componente de realizar sus funciones requeridas bajo las condiciones establecidas durante un período de tiempo especificado”.

2.5.2. Disponibilidad

Es la probabilidad de que el sistema esté operando correctamente en un determinado instante de tiempo (Torres-Pomales, 2000). La disponibilidad $A(t)$ de un sistema en el instante de tiempo t es la probabilidad que el sistema esté funcionando correctamente en el instante t (Dubrova, 2013).

Dubrova (2013) realiza una definición matemática de $A(t)$, llamándola también como *punto de disponibilidad* o *disponibilidad instantánea*. Y la define como:

$$A(T) = \frac{1}{T} \int_0^T A(t) dx$$

Para Hanmer (2007) la disponibilidad del sistema es el porcentaje de tiempo en el que es capaz de llevar a cabo una función determinada.

Para el caso de los sistemas que no pueden ser reparados el punto de disponibilidad es igual a la confiabilidad del sistema (Dubrova, 2013).

Los estados de disponibilidad pueden ser representados en términos de fuera de servicio por año. En la Tabla 2.1 que expone Dubrova (2013) se puede observar esta relación.

⁹En inglés, Reliability

Disponibilidad	Fuera de servicio
90 %	36.5 días/año
99 %	3.65 días/año
99.9 %	8.76 horas/año
99.99 %	52 minutos/año
99.999 %	5 minutos/año
99.9999 %	31 segundos/año

Tabla 2.1: Disponibilidad en relación con su baja de servicio por año. Tabla modificada de Dubrova (2013)

La IEEE (1990) define la disponibilidad como “El grado en el cual un sistema o componente se encuentra operativo y accesible cuando se requiere su uso. También es expresado en términos de probabilidad”

En satélites de órbita baja (LEO¹⁰), es necesario que el satélite se encuentre disponible al momento de su pasada por las estaciones terrenas, con el fin de descargar los datos que se fueron almacenando. Del mismo modo, el satélite debe estar disponible para llevar a cabo las funciones necesarias, y con esto cumplir con su misión (como por ejemplo la registración de imágenes en una determinada zona terrestre).

Diferente es el caso para los satélites geoestacionarios, ya que estos deberían estar disponible la mayor parte del tiempo, ya que en la mayoría de los casos son de comunicación.

Resumiendo, la disponibilidad es la probabilidad de que el sistema esté operando correctamente en el instante de tiempo en el que se requiera su servicios, es decir, en el momento que se lo opera.

2.5.3. Seguridad

La seguridad se considera como una extensión de la confiabilidad (Dubrova, 2013). Seguridad $S(t)$ es definida como la probabilidad que el sistema sea capaz de realizar sus funciones correctamente o discontinuar sus funciones en una manera a prueba de fallas (Dubrova, 2013).

Según Torres-Pomales (2000) la seguridad es la probabilidad de que el sistema llevará a cabo sus tareas de una manera no peligrosa. Un peligro se lo puede definir como “un estado o condición de un sistema, que juntos con otras condiciones ambientales del sistema, conducirá inevitablemente a un accidente” (Torres-Pomales, 2000).

La seguridad es requerida para aquellas aplicaciones de seguridad crítica donde una avería puede resultar en lesiones humanas, pérdidas de vida o desastres ambientales (Dubrova, 2013).

En el caso de los satélites, es importante asegurar altos niveles de seguridad, ya que el fracaso de una misión representa grandes pérdidas de dinero.

2.6. Tolerancia a falla

En sistemas críticos, como el de una planta nuclear, sistemas médicos, el sistema de vuelo de un avión, o el de un satélite, tanto SW (como el hardware) no deben fallar, ya que esto daría como resultado la pérdida de vidas humanas, como así también perdidas de dinero. Para el caso particular, del vehículo espacial (satélite, transbordador, lanzador), la falla del SW podría tener como consecuencia el fracaso de una misión, y/o una gran cantidad de dinero, y hasta vidas humanas en algunos caso (como pueden ser en vuelos tripulados). La principal

¹⁰Del inglés, Low Earth Orbit

diferencia entre el SW de una misión satelital, con la de un avión o una planta nuclear, o un sistema médico, es que ante alguna falla o error, se torna complicado llegar hasta el satélite para realizar una actualización o cargar un parche de SW.

La IEEE (1990) define como SW crítico a “aquel cuyo fracaso puede tener un impacto en la seguridad, o puede causar grandes pérdidas financieras o sociales”. El SW de estos sistemas críticos deben tener la capacidad de seguir funcionando, aún en la presencia de fallas, o errores. Imagínese el caso, de un avión comercial, con pasajeros a bordo, y de repente ocurre un problema debido al mal diseño del SW (por ejemplo un overflow de memoria). En esta situación es impensable que el SW se congele y que el piloto reinicie el sistema, esperar que se reestablezca al estado en el cual se encontraba antes del problema, para seguir funcionando. Lo mismo ocurre con el SW de naves espaciales, hay situaciones en la que no se puede esperar y es preferible que el sistema siga funcionando aún en la presencia de fallas.

Tal como lo indica Pressman (2001) las fallas de SW implican problemas cualitativos que son descubiertos después de que el SW es llevado a los usuarios y probados por ellos. Una gran cantidad de estudios indican que en las actividades de diseño se introducen entre un 50 y 65 por ciento de errores del total de errores que se dan durante el proceso del SW (Pressman, 2001). Esto no debe ocurrir en el ámbito espacial, ya que una vez que el sistema es utilizado, es muy difícil corregir los errores que surgen

Cabe aclarar que el SW al no ser un componente físico, no puede ser tratado de la misma manera que un componente hardware. Como ejemplifica Torres-Pomales (2000), las fallas que surgen a nivel de bit, como por ejemplo en un disco duro, son fallas del dispositivo de almacenamiento y pueden ser mitigadas con la aplicación de técnicas de redundancias. Esto no es así para el SW. Por lo tanto evitar los errores a nivel de SW no es tan trivial como en el hardware.

A nivel de SW las fallas son llamadas “bugs” (tal como se indica en la sección 2.1 en la página 8), y existe un solo tipo de fallas que es introducido durante el desarrollo del SW (Torres-Pomales, 2000). Las fallas en el SW son el principal motivo de que todo un sistema fracase.

La FT, puede ser utilizada como una capa más de protección (Torres-Pomales, 2000). Esta aplicada al SW se refiere al uso de técnicas que permiten seguir brindando el servicio en un nivel aceptable de performance y seguridad después que una falla de diseño ocurra.

Debe hacerse una diferencia entre FT y calidad. Hanmer (2007) lo define de la siguiente manera: “FT es la capacidad del sistema a ejecutarse apropiadamente a pesar de la presencia de fallas. FT ocurre en tiempo de ejecución”. Cuando se habla que un sistema es tolerante a fallas, significa que fue diseñado de tal manera, que puede seguir funcionando correctamente aún en la presencia de errores de sistemas (Hanmer, 2007).

En cambio calidad, tal como lo define Hanmer (2007), “se refiere a cuán libre de fallas está el sistema. Técnicas de calidad que indican cómo el SW es creado. Si el sistema fue testeado.”

Un sistema de alta calidad tendrá menor número de fallas, que esto representa menor número de fallas en tiempo de ejecución. La reducción del número de fallas no implica que los resultados de los defectos son menos severos (Hanmer, 2007). El sistema debe tomar medidas para reducir el impacto de los errores y fallas, y es allí donde surge la FT.

Un sistema tolerante a fallas provee una continua y segura operación, aún durante la presencia de fallas. Un sistema tolerante a fallas, es un elemento crítico para una arquitectura de vuelo, lo cual incluye hardware, SW, timing, sensores y sus interfaces, actuadores, elementos y datos de comunicación con los diferentes elementos (Hitt y Mulcare, 2001).

Este tipo de sistemas debería detectar los errores causados por fallas, evaluar los daños producidos por la falla, aislar a la misma y por último recuperarse, en ese caso se habla de arquitectura o sistemas FDIR¹¹.

¹¹FDIR, del inglés: Failure detect, isolate and recover

FT es la capacidad de un sistema a continuar funcionando a pesar de la ocurrencia de fallas (Dubrova, 2013). Un sistema tolerante a fallas debe ser capaz de manejar fallas tanto de hardware como de SW. La FT es necesaria debido a que es imposible construir un sistema perfecto.

El objetivo de la FT es el desarrollo de sistemas que tengan la capacidad de funcionar correctamente en presencia de fallas (Dubrova, 2013). La FT es alcanzada mediante la utilización de redundancias (Dubrova, 2013). *Redundancia* es la provisión de capacidades funcionales que sería innecesario para entornos libres de fallas (Dubrova, 2013). Esto significa tener hardware adicionales, check bits en una cadena de datos, o algunas líneas de código que verifique el resultado correcto del SW. La redundancia permite enmascarar una falla, o detectarla, para luego localizarla, contenerla y recuperarse de esta (Dubrova, 2013). Las técnicas de tolerancia de fallas se emplean durante la adquisición, o desarrollo del SW. Permite al SW tolerar fallas después que este haya sido desarrollado (Pullum, 2001). Cuando una falla se da, las técnicas de FT proveen mecanismos al sistema de SW para prevenir el fracaso del sistema (Pullum, 2001).

2.7. Redundancia en el software

A pesar de lo comentado anteriormente, sobre la importancia del SW, todavía existe una creencia, de que el SW aparece por arte de magia, y que los programadores no son nunca, lo suficientemente capaces, de hacer un SW libre de errores. Salvo aquellas empresas u organizaciones que tienen un proceso maduro de desarrollo de SW, el resto suele encasillarse en el pensamiento mencionado anteriormente.

Dubrova (2013) explica que la FT aplicado en el SW no está tan entendido, ni maduro, como es en el caso de la FT aplicada en hardware. Si una falla existiera en el SW, esta se haría “visible”, solo cuando las condiciones relevantes ocurran (Dubrova, 2013). Y muchas veces por tiempo o costo, no se realizan los tests cubriendo todos los posibles ambientes reales, lo cual tiene consecuencias desastrosas, tal como se expone en la sección 1.1 (página 3).

Para sistemas complejos o grandes, donde existe una gran cantidad de estados, implica que solo una pequeña porción del SW puede ser verificada correctamente (Dubrova, 2013). Los tests tradicionales y métodos de depuración actuales no alcanzan para grandes sistemas (Dubrova, 2013). La utilización de métodos formales para describir las característica requeridas por el comportamiento del SW exigen gran complejidad computacional, y solo son aplicables en ciertas situaciones (Dubrova, 2013).

Las técnicas de FT se pueden dividir en dos grupos:

- técnicas de una sola versión (single version), se utilizan cuando existe una sola versión del SW en el sistema.
- técnicas multi-versión (multi-version), se utilizan cuando se desarrollan varias versiones de una misma función.

Estas se explican en las siguientes secciones.

2.7.1. Técnicas single version

Estas técnicas son utilizadas para tolerar parcialmente las fallas del diseño de SW (Pullum, 2001). Técnicas single-version de FT se basa en el uso de redundancia aplicada a una única versión de una pieza de SW para detectar y recuperarse de fallas (Torres-Pomales, 2000).

Estas técnicas le brindan a los software que cuentan con una sola versión, un número de capacidades funcionales que no serían necesarias dentro de un ambiente libre de fallas (Dubrova, 2013).

2.7.1.1. Estructuras de software

En Torres-Pomales (2000) se mencionan dos técnicas de estructuración del SW que presentan características favorables al momento de mantener FT en el SW.

La definición de una arquitectura en el software es de suma importancia ya que proveen las bases para la implementación de FT (Torres-Pomales, 2000). Una de las técnicas utilizadas en el desarrollo del software es la modularización. Esta consiste en descomponer el problema en componentes manejables. Esto tiene como resultado que sea más eficiente la aplicación de la FT en el diseño de un sistema (Torres-Pomales, 2000).

El particionado es otra técnica mencionada en Torres-Pomales (2000), lo cual provee aislación entre módulos independientes del sistema. Esta técnica permite descomponer al problema en partes separadas (Pressman, 2001). El particionado puede ser horizontal u vertical. En el primero se descompone el problema moviéndose en forma horizontal en la jerarquía, mientras que el segundo se parte de lo más general hasta llegar a lo detallado, moviéndose verticalmente en la jerarquía (Pressman, 2001).

Sistema de cierre es un principio de FT, en el cual ninguna acción es permitible sin una autorización expresa (Torres-Pomales, 2000). Siguiendo este principio ninguna de las funciones que componen al sistema deberían tener más capacidad de la necesaria (Torres-Pomales, 2000). Las ventajas de desarrollar un sistema bajo este principio, es que es sencillo el manejo de errores, y evitar la propagación de fallas si ocurriesen.

2.7.2. Técnicas multi-version

Las técnicas de multi-version utilizan dos o más versiones diferentes del mismo módulo de SW ((Dubrova, 2013); (Torres-Pomales, 2000)), lo cual satisface el requerimiento de diversidad.

El objetivo de utilizar diferentes versiones de SW es que es construido de diferentes maneras, por lo tanto fallarían de diferente maneras (Torres-Pomales, 2000).

2.7.3. Técnicas de detección de fallas

Para los SW tolerantes a fallas de una sola versión, se suelen utilizar varios tests de “aceptación” para detectar fallas (Dubrova, 2013). Es necesario que estos SW cuenten con dos propiedades: auto protección¹² y auto check¹³ (Torres-Pomales, 2000). La auto protección significa que los componentes del sistema tienen la capacidad de protegerse así mismo mediante la detección de errores (Torres-Pomales, 2000). La propiedad de auto check significa que los componente son capaces de detectar fallas internas y tomar las acciones necesarias para evitar la propagación del error.

El resultado del sistema depende del resultado de los tests. Si el resultado pasa exitosamente el test, este es el correcto, caso contrario significa la presencia de fallas (Dubrova, 2013). Un test es más efectivo si se puede calcular de una manera simple (Dubrova, 2013).

Las técnicas utilizadas son las siguientes:

¹²En inglés, self-protection

¹³En inglés, self-checking

- *Timing checks*: se agrega a los sistemas una restricción de tiempo. Basado en esa restricción se puede deducir si el comportamiento del sistema se desvió (Dubrova, 2013). Los más utilizado es el *watchdog timer*, este es un contador que se actualiza con un *timer* que detecta si un módulo de SW se bloqueó o congeló, entonces se reinicia ese módulo o el sistema.
- *Coding checks*: se utiliza en los sistemas donde los datos se codifican usando técnicas de redundancia de datos (Dubrova, 2013).
- *Reversal checks*: son aquellos donde se toma los valores de salida, y con ellos se busca encontrar cuáles fueron los datos de entrada. Si los datos de entrada reales coinciden con los calculados (para una misma salida), este se encuentra libre de fallas (Dubrova, 2013).
- *Reasonableness checks*: usa propiedades semánticas en los datos para detectar fallas (Dubrova, 2013).
- *Structural checks*: se basa en el conocimiento de las propiedades de la estructura de datos (Dubrova, 2013).
- *Replication checks*: se basa en la comparación de resultados de varios componentes (Torres-Pomales, 2000).

Se suelen utilizar árboles de fallas, como una técnica auxiliar en el desarrollo de sistemas para la detección de fallas (Torres-Pomales, 2000). El árbol de falla permite obtener un enfoque top-down de las diferentes fallas que se pueden dar. El árbol no cubre todas las fallas que puedan darse, pero si ayudan en un alto grado en el desarrollo de SW tolerante a fallas (Torres-Pomales, 2000).

2.7.4. Técnicas de recuperación de fallas

Una vez que la falla es detectada, el sistema debe proceder a recuperarse de aquella, y volver a un estado operacional normal (Dubrova, 2013). Si los mecanismos de detección y contención de fallas fueron desarrollados correctamente, esta es contenida dentro de un set de módulos en el momento de la detección (Dubrova, 2013).

2.7.4.1. Manejo de excepciones

En muchos SW y lenguajes de programación, los errores se pueden recuperar mediante el manejo de excepciones. El manejo de excepciones es la interrupción del funcionamiento normal para responder a un funcionamiento anormal del sistema (Torres-Pomales, 2000). Los posibles eventos que pueden lanzar una excepción son:

1. Excepciones de interfaces, son lanzadas por un módulo cuando se da una solicitud inválida de algún servicio (Dubrova, 2013).
2. Excepciones locales, son lanzadas por algún módulo cuando sus propios mecanismos de detección de fallas encuentran un problema interno (Dubrova, 2013).
3. Excepciones de fracaso, son lanzadas cuando un mecanismos de detección encuentra una falla, pero es imposible recuperarse de esa falta (Dubrova, 2013).

2.7.4.2. Checkpoint y Restart

Para los software de una sola versión existen pocos mecanismos de recuperación. Checkpoint y restart es uno de ellos. También es conocido como *backward error recovery* (Dubrova, 2013). La mayoría de las fallas que se dan en los SW son debido a fallas que provienen del diseño, tal como se mencionó anteriormente. Estas fallas son activadas por entradas al sistema (Dubrova, 2013).

Este mecanismo cuenta con el módulo principal que se encuentra en ejecución combinado con un bloque que realiza tests de aceptación. Si se detecta una falla, en el bloque de testeo, se envía una señal de “reinicio”, para que el módulo principal vuelva al estado anterior, es decir, antes de producirse el error. Este estado anterior se encuentra almacenado en una memoria checkpoint (Dubrova, 2013). En la figura 2.2¹⁴, se muestra la representación de este mecanismo.

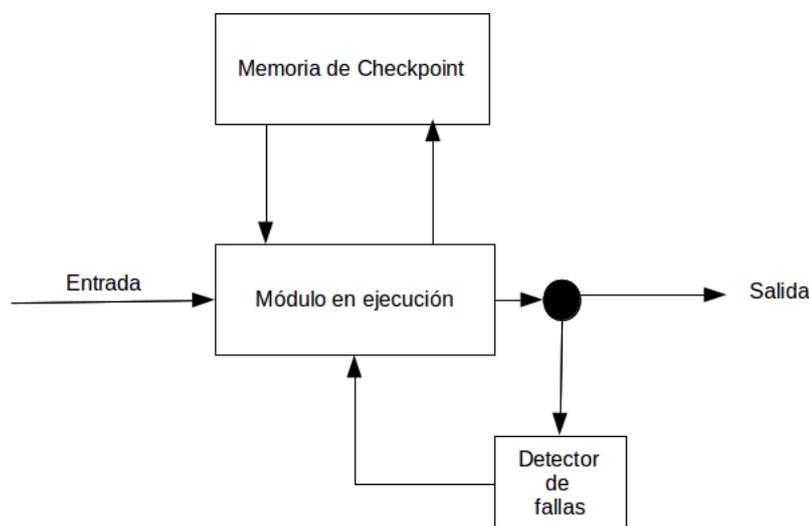


Figura 2.2: Representación de checkpoint y restart

Existe dos tipos de checkpoints, estáticos y dinámicos. Los checkpoints estáticos toman una “fotografía” del estado del sistema antes de comenzar la ejecución del SW y lo guarda en la memoria (Dubrova, 2013). Si se detecta una falla, el sistema regresa a ese estado y comienza de nuevo su ejecución (Dubrova, 2013). Los checkpoints estáticos se basan en regresar el módulo a un estado predeterminado (Torres-Pomales, 2000). Se puede regresar a un estado inicial o a un set de estados predeterminados (Torres-Pomales, 2000).

Por otro lado se encuentran los checkpoints dinámicos. Estos usan checkpoints creados dinámicamente. Estas son imágenes del estado del sistema en varios puntos durante la ejecución (Torres-Pomales, 2000).

Hay tres formas de crear los checkpoints dinámicamente:

1. Equidistantes, en el cual los intervalos que se crean los checkpoints son siempre iguales, los intervalos se elijen teniendo en cuenta el rate de falla (Dubrova, 2013).
2. Modular, en el cual los checkpoints se crean al principio o al final de la ejecución de un módulo.
3. Random, los checkpoints se crean aleatoriamente en el tiempo.

¹⁴Basado en Dubrova (2013) y Torres-Pomales (2000)

2.7.4.3. Procesos pares

Los procesos a pares utilizan dos versiones idénticas de un proceso de SW que corre en procesadores separados ((Dubrova, 2013); (Torres-Pomales, 2000)). El mecanismo de recuperación que se utiliza es el de checkpoint y restart (Torres-Pomales, 2000).

Como se puede observar en la figura 2.3¹⁵ el primer procesador se encuentra activo. Este envía un checkpoint al segundo procesador. Si una falla se detecta, el primer procesador se apaga y se cambia al segundo procesador. El segundo procesador carga el checkpoint y continúa con la operación. Toma el rol del primer procesador (Torres-Pomales, 2000). Luego el primer procesador realiza un auto test para verificar si el problema continúa. Si se encuentra que este procesador sigue teniendo problema, se continúa trabajando con el segundo procesador (Dubrova, 2013).

La principal ventaja que brinda este mecanismo según Dubrova (2013) es que permite entregar el servicio ininterrumpidamente.

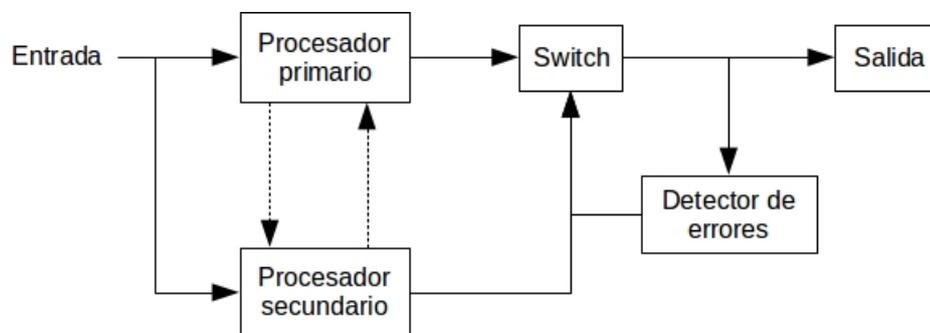


Figura 2.3: Representación del proceso pares

2.7.4.4. Diversidad de datos

La diversidad es una técnica utilizada para mejorar la eficiencia en los checkpoint y restart, usando diferentes entradas por cada reinicio. Esto se basa en que las fallas en el SW son dependientes de las entradas. Es poco probable que la misma falla se de con la misma secuencia de entrada (Dubrova, 2013).

2.7.4.5. Bloques de recuperación

Esta técnica combina las bases de la técnica de checkpoints y restart enfocada con múltiples versiones de un componente de SW en el sentido de que una versión de SW diferente es lanzada cada vez que se encuentra una falla. Los checkpoints son creados antes de que una versión de SW se ejecuta. La ejecución de las múltiples versiones pueden ser secuencial o paralelas dependiendo de la disponibilidad de la capacidad de procesamiento y performance requerida (Torres-Pomales, 2000).

Las versiones son diferentes implementaciones de un mismo programa. Solo una de estas versiones provee la salida del sistema. Si un error es detectado por el test de aceptación, se vuelve hacia atrás, se retoma el último checkpoint, y se vuelve a ejecutar el módulo de SW pero con una versión diferente a la que se ejecutó anteriormente (Dubrova, 2013).

Los checks del test de aceptación deben mantenerse simples para mantener la velocidad de la ejecución (Dubrova, 2013).

¹⁵Basada en Dubrova (2013) y Torres-Pomales (2000)

2.8. Técnica de evaluación de fiabilidad

La evaluación de la fiabilidad es de suma importancia para el desarrollo de sistemas críticos, ya que permite identificar que aspectos del comportamiento del sistema juega un papel importante (Dubrova, 2013).

1. Modelado de un sistema en la fase de diseño.
2. Aseguramiento del sistema en la fases finales de desarrollo (testing).

El análisis confiabilidad tiene tres enfoques importantes:

- Confiabilidad de HW
- Confiabilidad de SW
- Confiabilidad humana

En este trabajo de tesis se pondrá énfasis en el estudio en la confiabilidad del SW a nivel de sistema.

La evaluación de la fiabilidad tiene dos aspectos. En primer lugar se tiene una *evaluación cualitativa* que permite identificar, clasificar y medir modos de fallas, o eventos combinatoriales que puedan provocar una falla. El otro aspecto es la *evaluación cuantitativa*, la cual permite evaluar en términos de probabilidad los atributos de la fiabilidad, disponibilidad, seguridad (Sección 2.5).

El análisis de confiabilidad es de gran importancia ya que provee información que es la base de la toma de decisiones. Esto es aplicado a diferentes áreas, tales como análisis de riesgos, protección ambiental, calidad, optimización de mantenimientos y operaciones y diseño de ingeniería (Rausand y Hoyland, 2004).

2.9. Medidas comunes de fiabilidad

Las medidas de fiabilidad más comunes son las siguientes: failure rate, tiempo medio a la falla, tiempo medio de reparación y tiempo medio entre fallas.

2.9.1. Failure rate

Failure rate λ es el número esperado de fallas por unidad de tiempo (Dubrova, 2013). Es usual utilizar la dimensión *fallas/horas*.

Generalmente, λ se encuentra a nivel de componente. Para conocer el failure rate del sistema completo, se puede realizar (*grosso modo*) una sumatoria de los λ de los componentes que integran el sistema.

$$\lambda = \sum_{i=1}^n \lambda_i$$

La evolución de λ a través del tiempo, no tiene el mismo comportamiento tanto para HW como para SW. Si se divide el ciclo de vida de un sistema en las siguientes fases: mortalidad prematura (I), vida útil (II), desgaste (III) (Dubrova, 2013) se aprecia, para el caso del HW, lo que se denomina *curva de la bañera* la cual puede observarse en la Figura 2.4. En una primera fase, λ decrece, ya que a través de los procesos de testing se van

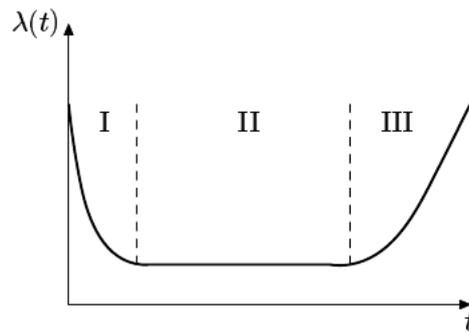


Figura 2.4: Failure rate de HW vs tiempo

descubriendo y resolviendo los errores. Luego se da un periodo de estabilización. Y al final, el HW sufre el paso del tiempo, y se desgasta, aumentando la tasa de fallas.

Para el SW es totalmente diferente. En primer lugar cuando se realiza una actualización, se aumenta la complejidad, como así también la probabilidad de fallas, con ello el failure rate. Otra diferencia sustancial con el HW es que el SW no se desgasta con el tiempo. En la Figura 2.5 se puede apreciar λ a través de tiempo. Esta curva suele llamarse *curva serrucho*. El failure rate del SW decrece en función del tiempo. En estos tipos de sistemas la tasa de falla depende de varios factores como pueden ser el proceso utilizado en el diseño y codificación, complejidad del SW, tamaño del SW, etc (Dubrova, 2013).

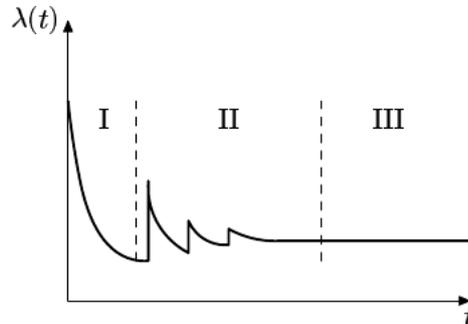


Figura 2.5: Failure rate SW vs tiempo

A lo largo de la vida del sistema se supone que el failure rate λ es constante.

La confiabilidad del sistema varía exponencialmente con respecto al tiempo (Dubrova, 2013):

$$R(t) = e^{-\lambda t}$$

Esto se conoce como *ley de la falla exponencial* (Dubrova, 2013). El gráfico de confiabilidad $R(t)$ vs tiempo se muestra en la Figura 2.6. Esta ley es utilizada para analizar la confiabilidad de componentes y sistemas de hardware. En este trabajo de tesis se asume su extensión al software. Esta ley solo puede ser utilizada en casos donde, se asume que la tasa de falla es constante. (Dubrova, 2013)

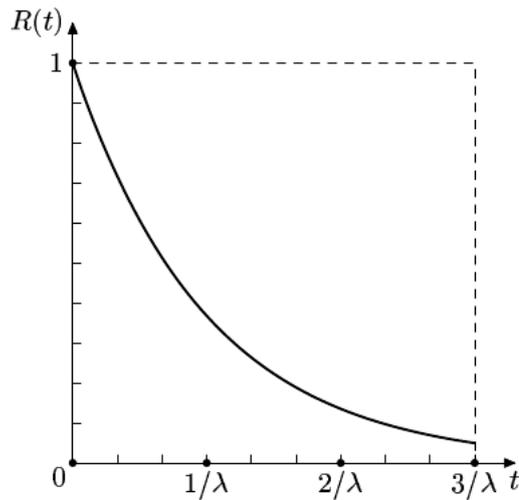


Figura 2.6: Confiabilidad vs tiempo

2.9.2. Tiempo medio de falla

El tiempo medio de falla (MTTF¹⁶) de un sistema es el tiempo esperado que transcurra hasta la primera falla que se detecte en el sistema. En terminos de confiabilidad, MTTF se define de la siguiente manera ((Dubrova, 2013); (Rausand y Hoyland, 2004))

$$\int_0^{\infty} R(t)dt$$

2.9.3. Tiempo medio de reparación

El tiempo medio de reparación (MTTR¹⁷) de un sistema, es el promedio de tiempo que se requiere para reparar al sistema. MTTR se especifica en términos de la tasa de reparación μ ((Dubrova, 2013); (Rausand y Hoyland, 2004)), el cual es el número esperado de reparaciones por unidad de tiempo:

$$MTTR = \frac{1}{\mu}$$

Esto será siempre y cuando la mantenibilidad sea exponencial. Esto se da si la probabilidad de que el sistema sea reparado en un tiempo determinado, dado que falló al principio de ese intervalo sea proporcional a ese intervalo de tiempo.

El MTTR depende de los mecanismos de recuperación ante fallas que se utilicen en el sistema, localización del sistema, scheduler de mantenimiento (Dubrova, 2013). Con esto se puede definir la disponibilidad como sigue:

$$A(\infty) = \frac{MTTF}{MTTF + MTTR}$$

Muchas veces se utiliza MDT¹⁸ en vez de MTTR, para denotar más claro que es el tiempo medio que el sistema se encuentra fuera de servicio.

¹⁶Del inglés, Mean Time To Failure

¹⁷Del inglés, Mean Time To Repair

¹⁸Del inglés, Mean Downtime

2.9.4. Tiempo medio entre fallas

El tiempo medio entre fallas (MTBF¹⁹) de un sistema es el tiempo promedio entre dos fallas del sistema.

$$MTBF = MTTF + MTTR$$

2.9.5. Cobertura de fallas

La cobertura de fallas es la probabilidad de que el sistema no interrumpirá su actividad cuando una falla se presente. En términos matemáticos la cobertura de fallas es la probabilidad condicional $P(A|B)$. Existen diferentes coberturas de fallas, dependiendo de si se está tratando con detección de fallas, localización de fallas, contención de fallas o recuperación de fallas (Dubrova, 2013). Siendo A detección, localización, contención o recuperación de fallas, y B la existencia de fallas.

2.10. Métodos de cálculos de fiabilidad

Para evaluar la fiabilidad de sistemas se pueden utilizar diagramas de bloque de confiabilidad y procesos de Markov (Dubrova, 2013)

2.10.1. Diagramas de bloques de confiabilidad

2.10.1.1. Cálculo de confiabilidad

Para medir la confiabilidad de un sistema mediante diagrama de bloques, se debe dividir el sistema objetivo en partes paralelas y en serie. Se computa la confiabilidad de las partes. La confiabilidad del sistema estará compuesta por la confiabilidad de ambas partes (Dubrova, 2013). Entonces

$$R(t) = \begin{cases} \prod_{i=1}^n R_i(t) & \text{para estructuras en serie} \\ 1 - \prod_{i=1}^n (1 - R_i(t)) & \text{para estructuras en paralelo} \end{cases}$$

Esto nos indica que un sistema paralelo, es más confiable que uno en serie, aún así si sus componentes son menos confiables. Tal como ejemplifica Dubrova (2013), si se diseña un sistema en serie de 100 componentes, con una confiabilidad de 0.999, el sistema completo tendrá una confiabilidad de: $0,999^{100} = 0,905$. Mientras que, para un sistema paralelo, con solo cuatro componentes, con una confiabilidad menor (0.95) la confiabilidad del sistema será $1 - (0,95)^4 = 0,9999375$. El punto en contra de los sistemas paralelos, es que representan un costo mayor que las estructuras en serie (Dubrova, 2013).

2.10.1.2. Cálculo de disponibilidad

Si se asume que el tiempo de falla y de recuperación son independientes, entonces se puede utilizar diagramas de bloques para calcular la disponibilidad del sistema (Dubrova, 2013). Se puede observar que el cálculo es similar al cálculo de confiabilidad.

$$A(t) = \begin{cases} \prod_{i=1}^n A_i(t) & \text{para estructuras en serie} \\ 1 - \prod_{i=1}^n (1 - A_i(t)) & \text{para estructuras en paralelo} \end{cases}$$

¹⁹Del inglés, Mean Time Between Failure

2.10.2. Utilización de procesos de Markov

El principal objetivo del análisis de los procesos de Markov es calcular $P_i(t)$ la probabilidad de que el sistema se encuentre en el estado i en el tiempo t . Con esto se puede calcular fácilmente la confiabilidad, disponibilidad o seguridad del sistema (Dubrova, 2013).

Para determinar $P_i(t)$ se debe derivar una serie de ecuaciones diferenciales, una por cada estado del sistema. Estas ecuaciones se denominan ecuaciones de estado de transición. Las ecuaciones de los estados de transición se representan en una matriz M denominada *matriz de transición*. Cada elemento m_{ij} de la matriz M es un rate de transición entre los estados i y j

$$M = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{k1} \\ m_{12} & m_{22} & \dots & m_{k2} \\ & & \dots & \\ m_{1k} & m_{2k} & \dots & m_{kk} \end{bmatrix}$$

Haciendo $P(t)$ un vector en el cual el i -ésimo elemento es la probabilidad $P_i(t)$ de que el sistema se encuentra en el estado i en el tiempo t . Con ello se tiene:

$$\frac{d}{dt}P(t) = MP(t)$$

(Dubrova, 2013).

Para calcular confiabilidad, disponibilidad y seguridad solo es necesario reemplazar en la matriz M , los rate correspondientes.

2.11. Modelos de falla

En esta sección se explica con un poco más de detalle algunos conceptos que se presentaron en secciones anteriores. Los conceptos que se detallarán son los siguientes:

- Función de confiabilidad $R(t)$
- Función de tasa de falla λ
- Tiempo medio hasta la falla (MTTF)
- Vida residual media (MRL)

También existen varias distribuciones de probabilidad que son utilizados para modelar el tiempo de vida de componentes de sistemas. Algunos de estos modelos de distribución del tiempo de vida son los siguientes:

- Distribución binomial
- Distribución exponencial
- Distribución gamma
- Distribución Weibull
- Distribución normal
- Distribución Birnbaum-Saunders
- Distribución inversa de Gauss

2.11.1. Tiempo hasta la falla

El tiempo hasta la falla (T) de un componente del sistema, es el lapso de tiempo desde que el componente empieza a operar hasta la primera falla que se produzca (Rausand y Hoyland, 2004).

Nota: Se asume que el componente que falla, no puede ser recuperado. Por otro lado, si un sistema falla, puede recuperarse (a través de técnicas de reconfiguración, componentes back-up).

T no significa que es solo una medida de tiempo. T puede significar (Rausand y Hoyland, 2004):

- Cantidad de kilómetros de un automóvil
- Número de rotación de un motor
- Número de ciclos de trabajo

Debe destacarse que esta no solo es una variable discreta. Una variable discreta puede ser aproximada a una variable continua. Suponiendo que T es una distribución continua con densidad de probabilidad $f(t)$ y una función de distribución:

$$F(t) = P(T) = \int_0^t f(u)du \text{ para } t > 0$$

$F(t)$ demuestra la probabilidad de que un componente falle dentro de un intervalo de tiempo $(0, t]$ (Rausand y Hoyland, 2004).

La función de densidad de probabilidad $f(t)$ se define como sigue (Rausand y Hoyland, 2004):

$$f(t) = \frac{d}{dt}F(t) = \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t)}{\Delta t}$$

Esto indica que para un Δt pequeño,

$$P(t < T \leq t + \Delta t) \approx f(t)\Delta t$$

2.11.2. Función de confiabilidad

La función de confiabilidad se define como sigue:

$$R(t) = 1 - F(t) = P(T > t)$$

para $t > 0$.

2.11.3. Tasa de falla

La probabilidad de que un componente falle en un intervalo de tiempo $(t, t + \Delta t]$ dado que el componente ha funcionado hasta t , se tiene:

$$P(t < T \leq t + \Delta t | T > t) = \frac{P(t < T \leq t + \Delta t)}{P(T > t)} = \frac{F(t + \Delta t) - F(t)}{R(t)}$$

Si se divide esta probabilidad por un intervalo de tiempo Δt , y haciendo $\Delta t \rightarrow 0$, se obtiene la función de tasa de falla.

2.11.4. Tiempo medio hasta la falla

El Tiempo Medio de Fallas (MTTF) de un componente está definido por ((Dubrova, 2013); (Rausand y Hoyland, 2004)):

$$MTTF = E(T) = \int_0^{\infty} t f(t) dt$$

Se demuestra en Rausand y Hoyland (2004) que

$$MTTF = \int_0^{\infty} R(t) dt$$

2.11.5. Vida restante media

Considerando un componente con tiempo hasta la falla T , que es colocado en operación en el tiempo $t = 0$ y continúa funcionando en el instante t . La probabilidad de que el componente de edad t sobreviva en un intervalo x , entonces se tiene (Rausand y Hoyland, 2004):

$$R(x|t) = P(T > x+t | T > t) = \frac{P(T > x+t)}{P(T > t)} = \frac{R(x+t)}{R(t)}$$

Entonces:

$$MRL(t) = \frac{1}{R(t)} \int_t^{\infty} R(x) dx$$

Nótese que cuando $t = 0$, el componente es nuevo, y entonces $\mu(0) = \mu = MTTF$.

2.11.6. Distribución binomial

La distribución binomial es uno de lo más utilizados (Rausand y Hoyland, 2004). Esta distribución es utilizada en los siguientes casos:

- Cuando se tienen n ensayos independientes
- Cada ensayo tiene dos posibles resultados
- La $P(A) = p$ es la misma en todos los ensayos

Entonces

$$P(X = x) = \binom{n}{x} \cdot p^x (1-p)^{n-x} \quad \text{para } x = 0, 1, \dots, n$$

Donde $\binom{n}{x}$ es el coeficiente binomial que se define como:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

y X es el número de n ensayos para alcanzar un resultado A (Rausand y Hoyland, 2004).

2.11.7. Distribución exponencial

Considerando un componente que entra en operación en el instante $t = 0$. El tiempo hasta la falla T del componente tiene una función de densidad de probabilidad de la siguiente forma (Rausand y Hoyland (2004)):

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{para } t > 0, \lambda > 0 \\ 0 & \text{para cualquier otro caso} \end{cases}$$

Esta distribución es la que se denomina distribución exponencial con parámetro λ .

La confiabilidad de esta función es:

$$R(t) = P(T > t) = \int_t^{\infty} f(u) du = e^{-\lambda t} \text{ para } t > 0$$

El MTTF es:

$$MTTF = \int_0^{\infty} R(t) dt = \int_0^{\infty} e^{-\lambda t} dt = \frac{1}{\lambda}$$

La tasa de falla de esta distribución es λ .

La distribución exponencial es lo que se utiliza más comúnmente en los análisis de confiabilidad.

Si hacemos:

$$R(MTTF) = R\left(\frac{1}{\lambda}\right) = e^{-1} \approx 0,3679$$

se puede calcular la probabilidad de que un componente sobreviva durante el tiempo medio hasta la falla (Rausand y Hoyland, 2004).

Mientras que la tasa de falla es:

$$z(t) = \frac{f(t)}{R(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$$

NOTA IMPORTANTE: Como se mencionó anteriormente, existen numerosas distribuciones de probabilidad para medir el tiempo de vida de los componentes. En este trabajo de tesis, se trabaja con la distribución exponencial debido a que es sencilla su aplicación.

2.12. Protocolos de comunicación de tiempo real

2.12.1. Sistemas de tiempo real

Un sistema de tiempo real es un sistema de computadora que depende del correcto comportamiento funcional con respecto al dominio del tiempo, esto significa que el sistema debe llevar a cabo sus funciones y obtener como resultados (correctos) dentro de las restricciones de tiempo (Lisner, 2007).

Los sistemas de tiempo real se dividen en *sistemas de tiempo real "soft"* y *sistemas de tiempo real "hard"*. En los soft real-time systems es importante cumplir con los tiempos del planificador, pero el no cumplimiento no tiene un impacto en la seguridad del sistema. Por otro lado, en los hard real-time systems, el no cumplimiento de las restricciones del tiempo, tiene como consecuencia un impacto catastrófico en el sistema.

Existen numerosos protocolos de comunicación, la mayoría abocados en la implementación de la capa 1 y 2 del modelo de ISO/OSI. Estos protocolos especifican las restricciones de hardware, la topología de red,

la arquitectura del nodo, acceso al medio, los mecanismos de detección de error, etc. (Lisner, 2007). En los sistemas críticos como en el caso de los vehículos satelitales, es necesario la aplicación de sistemas de tiempo real, que garantice una mínima latencia y un comportamiento determinístico.

La tolerancia a fallas aplicadas en protocolos de comunicación de tiempo real, permiten que la red continúe funcionando aún cuando algunos de sus componentes han fallado. Los sistemas tolerantes a fallas son diseñados a partir de un modelo de falla dado (Lisner, 2007). El modelo de falla describe la estructura del sistema y los tipos de fallas que pueden ocurrir (Lisner, 2007).

Otro aspecto importante de los protocolos de comunicación tolerantes a fallas de tiempo real, es conocer el momento en el que un nodo está disponible para enviar mensajes. Esto es llevado a cabo por las estrategias de acceso al medio. Los métodos de acceso al medio pueden ser clasificados como event-triggered y time-triggered. En el primero, un mensaje es enviado si algún nodo requiere ese mensaje. Este no es una buena opción para los sistemas de tiempo real, ya que no asegura los límites de tiempo (Lisner, 2007).

Los métodos time-triggered es hecho en ciclos. Los nodos tiene acceso al medio a través de intervalos periódicos de tiempo (Lisner, 2007). Una de las principales ventajas de este enfoque, es que es predecible, los intervalos son pre configurados. Como desventaja se puede mencionar que se hace un uso deficiente del medio, ya que cuando algún nodo no tiene nada para enviar en su ventana de tiempo, el canal se encuentra ocioso.

2.13. Estrategia de acceso al medio

2.13.1. CSMA

El esquema CSMA (Acceso Múltiple por Detección de Portadora²⁰) proviene de ideas del Ethernet alámbrico. En esta técnica los nodos esperan durante un intervalo corto de tiempo y aleatorio antes de transmitir (Tanenbaum, 2003). Este método es utilizado comúnmente en redes inalámbricas. Al protocolo CSMA se la puede dividir en:

- CSMA con Detección de Portadora
- CSMA con Detección de Colisiones

En la primera, el nodo intenta enviar un mensaje cuando el canal está inactivo. Si otro nodo lo está usando espera hasta que se desocupa y luego transmite. En el caso de una colisión de mensajes, espera un tiempo aleatorio antes de enviar el mensaje nuevamente. Esta tiene algunas desventajas, que no la hacen apropiadas para su aplicación en sistemas críticos. Para hacer frente a las desventajas del CSMA con detección de portadora, se utiliza CSMA/CD (CSMA con Detección Detección de Colisiones).

Por otro lado, también existe otro método denominado CSMA/CA (CSMA con Evitación de Colisiones), utilizado en el 802.11. Este protocolo que es similar al CSMA/CD de Ethernet, con detección de canal antes de enviar y retroceso exponencial después de las colisiones (Tanenbaum, 2003). Esta estrategia además de ser utilizada en el Wireless, se utiliza en CAN.

2.13.2. TDMA

TDMA (time division multiple access) es una técnica de multiplexación del tiempo que es utilizada en redes inalámbricas, comunicación de satélites y en diferentes protocolos de tiempo real (Lisner, 2007).

²⁰Del inglés, Carrier Sense Multiple Access

Los accesos al medio se realiza mediante ciclos, en el cual se subdivide dentro de slots de tiempo de ancho estático. Solo un nodo está permitido enviar en un solo slot.

2.13.3. Minislottting

El uso de slots de tiempo con un ancho fijo y estático del tiempo, como es el que se utiliza en TDMA, puede convertirse en un problema. En el TDMA los nodos tiene que esperar todo un ciclo para poder enviar un mensaje. Además, el nodo debe enviar mensajes vacíos (*null*) durante su slots, cuando no tiene nada que enviar. Solución de este problema es el minislottting, ya que permite la utilización de slots con ancho variable. Esto logra recortar los tiempos de espera cuando no se está utilizando el medio (Lisner, 2007). Este método permite un uso más eficiente de los ciclos. La transmisión puede tener diferentes anchos, y no hay necesidad de enviar mensajes nulos. Este método se basa en la prioridad y para evitar la monopolización del medio, algunos protocolos como el ARINC 629, utiliza timeouts para denegar el acceso al medio después de un determinado tiempo (Lisner, 2007).

2.14. Revisión de protocolos

2.14.1. CAN

CAN es un protocolo de comunicación desarrollado por Bosch para ser utilizado en automóviles. Es protocolo ha ganado peso, a tal medida que se realizó un estándar de ISO basado CAN. Este es un protocolo del tipo event-triggered y utiliza CSMA/CA.

CAN utiliza diferentes mecanismos de detección de errores. Utiliza un CRC de 15 bits. Cada nodo puede enviar un mensaje de diagnóstico de errores, además de que lleva un contador de errores propios. Si este número de errores es grande, el mismo nodo entra en modo *error activo* (envía flags de “error activo”), *error pasivo* (envía flags de “error pasivo” y espera 8 bit times antes de repetir el envío del mensaje) y *bus off* (el nodo se apaga) (Lisner, 2007)

2.14.2. byteflight

Este se basa en el protocolo Minislottting y utilizado por BMW en automóviles. Utiliza un clock master de alta precisión para la sincronización.

Tanto CAN como byteflight, no poseen un mecanismo o técnica para proteger el canal de alguna falla de los nodos (Lisner, 2007).

2.14.3. ARINC 659 o SAFEbus

SAFEbus es una implementación del estándar ARINC 659 para su utilización en aviones comerciales. También se suele utilizar en vehículos espaciales. SAFEbus es una arquitectura que es altamente redundante.

2.14.4. TTP/C

TTP/C proviene de la familia TTP. Este protocolo utiliza el esquema TDMA en una arquitectura de doble canal. La configuración se encuentra pre definida. Esto permite al sistema ser predecible, lo cual facilita la

tolerancia a fallas. Su alto grado de tolerancia a fallas, permite desarrollar aplicaciones más confiables que otros protocolos (Lisner, 2007).

2.15. Posibles fallas en una red

Las principales fallas que se pueden producir en una red son debidas al canal de comunicación y los nodos/controladores, esto es así, ya que son los componentes más importantes de toda red (Lisner, 2007).

Las fallas en el nodo pueden ser de múltiples tipos. El nodo puede codificar erróneamente un mensaje; se pueden dar errores en el timing, es decir actúan en tiempo equivocado; el nodo puede tener un comportamiento inusual y se bloquea el servicio que está brindando.

Asimismo, las fallas en el canal de comunicación pueden ser de diferentes tipos. El canal puede generar ruido, este ruido puede modificar un mensaje (hasta el momento un mensaje correcto); el mensaje no llega a los destinatarios.

También se pueden observar que tanto errores en el nodo como en el canal se produzcan simultáneamente (Lisner, 2007).

2.16. Protocolo CAN

2.16.1. Consideraciones previas

En este trabajo de tesis se decidió continuar el desarrollo de la arquitectura con BUS-CAN ya que este es de interés para los proyectos desarrollados en INVAP. Cabe aclarar que el objetivo principal de esta tesis no es llevar a cabo una comparación exhaustiva de protocolos de comunicación tolerantes a fallas, sino elegir la más indicada y que se adapte a las necesidades de INVAP. Por tal motivo, se expone, en esta sección, las características importantes de CAN.

2.16.2. Introducción

El bus CAN (Controller Area Network) fue desarrollado por Bosh. Comenzó su desarrollo en 1983 y tuvo su primer release en 1986. Fue estandarizada por la Organización Internacional de Estandarización (ISO) bajo el nombre de ISO 11898. El Bus CAN surgió como respuesta al rápido crecimiento de la electrónica en la industria automotriz (ESD Electronics, 2017). Este protocolo se definió con el objetivo de proveer comunicación determinística de sistemas distribuidos, y que necesiten un alto grado de confiabilidad. CAN permite la conectividad vía bus serial. El bus está compuesto por 2 cables que pueden estar blindados o no (ESD Electronics, 2017).

Las características de CAN que la convierten en una tentadora opción para la aplicación en el sector espacial, son:

- Bajo costo
- Operabilidad en ambientes eléctricos complicados
- Capacidades de tiempo real.
- Facilidad en el uso

La especificaciones de CAN, originalmente desarrollados por Bosch cubre solamente las capas físicas y de enlaces de datos del modelo de referencia de ISO/OSI. Luego de la estandarización ISO provee detalles adicionales de la capa física de CAN.

A diferencia de otros protocolos como USB o Ethernet, CAN no envía grandes bloques de datos de un punto a otro. CAN envía mensajes cortos como temperatura, o RPM (Revoluciones por minutos), y son enviados como broadcast a todo el sistema (Corrigan, 2002).

El protocolo de comunicación CAN, en su estándar ISO 11898, explica cómo la información viaja entre los dispositivos conectados a una arquitectura siguiendo el modelo de OSI. La arquitectura planteada por la ISO 11898 se puede observar en la Figura 2.7 (Corrigan, 2002).

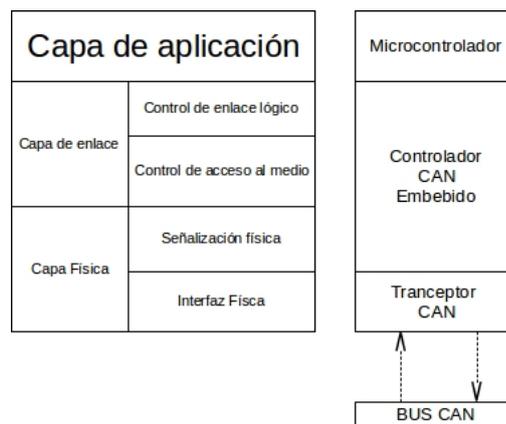


Figura 2.7: Arquitectura estándar propuesta por ISO11898

En resumen, la capa física de CAN describe la definición del bit timing, bit encoding y la sincronización, los niveles aceptables de la señal (corriente y voltaje), los conectores y características físicas de los cables (Corrigan, 2002). Por otro lado la capa de enlaces de datos, provee todos los servicios necesarios para la transmisión de un stream de bits desde un nodo a otro.

El bus CAN es un bus de tipo broadcast, es decir que todos los nodos escuchan todas las transmisiones. No hay manera de enviar un mensaje a un determinado nodo, por lo tanto no es necesario el direccionamiento de nodos. (Kvaser, 2017).

Actualmente la industria aeronáutica la utiliza. También el área agrícola en sus maquinarias, control de tráfico, sistemas de control industrial, sistemas domóticos y además, en equipamientos médicos.

2.16.3. Elementos necesarios

Como se puede observar en la Figura 2.7 cada nodo conectado a la red necesita de:

- Central Processing Unit (CPU), o microprocesador, que decida qué mensajes debe recibir y/o enviar
- Controlador CAN, muchas veces integrado en el controlador, el cual recibe la serie de bits del bus. También es el encargado de transmitir mensajes (cuando se requiera), esto significa que al mensaje lo convierte en una serie de bits.
- Tranceiver, definido en el ISO 11898.

2.16.4. Capa física

El estándar de CAN define el bit encoding, el timing, y la sincronización. La capa física es la encargada de convertir 1 y 0 en pulsos eléctricos para enviar mensajes por el canal, y también el proceso inverso cuando recibe mensajes. La capa física es implementada enteramente en HW (Corrigan, 2008).

2.16.4.1. Bus de comunicación

Para iniciar una transmisión de mensaje es necesario como mínimo dos nodos conectados al bus CAN. Esto se debe a que el dispositivo que envía un mensaje está también recibiendo su propio mensajes, con esto puede chequear cada bit que ha enviado. De esta manera, un segundo nodo responde con un ACK (acknowledgement), mientras el bit todavía está siendo transmitido por el primer nodo. Esto demuestra por qué se necesitan de dos nodos para completar la transmisión de mensajes (Corrigan, 2008).

En la Figura 2.8 se observa un ejemplo extraído de Corrigan (2008). En esta se muestra un nodo A que envía un mensaje. Luego se ve que que B y C contestan con un ACK indicando que el mensaje fue recibido sin errores. Luego B y C comienzan a transmitir hasta que C gana el bus debido a que tiene un bit dominante, y termina este completando su mensaje.

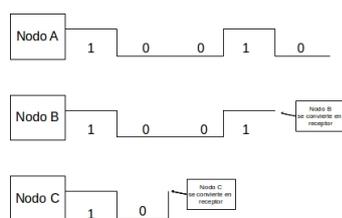


Figura 2.8: Tráfico en el BUS CAN

El medio físico es una línea de bus con dos cables, terminada por una resistencia. Los cables pueden ser paralelos, trenzados y/o blindados. Los segmentos de cable para la conexión de los nodos deben ser lo más cortos posibles.

Para mayores detalle técnico referentes al bus será necesario estudiar el estándar ISO 11898 o en Corrigan (2008).

El largo del bus va a determinar el bitrate de la comunicación. Para alcanzar un bitrate de 1Mbps, el máximo del canal de comunicación es de 40m. Si se tiene un bus de 1km el bitrate es de 0.05Mbps. Con esto se puede observar que el bitrate decae cuando la distancia se incrementa. Para CAN, el bitrate también está determinado por el total del delay del sistema (Corrigan, 2008). Esto es la suma de los delays del nodo, y el delay propio del cable.

Debe destacarse que existen diferentes capas físicas para el protocolo CAN:

- El tipo más común es el establecido en el estándar ISO 11898-2. Este protocolo cuenta con dos cables, cada uno identificado como CAN_H y CAN_L. También es llamado *CAN de alta velocidad* (1 Mbit/s). Este requiere que el largo del cable sea como máximo 40m.
- También en el estándar de ISO se establece el ISO 11898-3, el cual define otro esquema de doble cable balanceado pero para bajas velocidades. Este es tolerante a fallas si alguno de los cables entra en corto circuito. También es llamado también como *CAN de baja velocidad* (125 kbit/s)
- Se define el SAE J2411 que utiliza un solo cable para la capa física. Este es utilizado en autos (velocidad

por encima de los 50 kbit/s). No tiene demasiados requerimientos en cuanto al bitrate ni el largo del canal de comunicación. El estándar define 32 nodos por red.

- Existen modificaciones del estándar RS485 para su funcionamiento con CAN

2.16.4.2. Bit timing

El tipo de señal es codificado con Non Return Zero (NRZ). Los bits son codificados en dos estados llamados “recesivo” y “dominante”(el bit 0 es asociado al bit dominante). El protocolo permite acceso al bus multi-master con una resolución determinística ante colisiones. Para lograr la sincronización, el protocolo sincroniza durante las transiciones. Por tal motivo, deben evitarse el envío de largas cadenas de bits en un mismo estado. CAN utiliza una técnica que se denomina “bit stuffing” o “bit padding” en la cual luego de enviar 5 bits con el mismo estado, se insertan bits de relleno.

2.16.4.3. Asignación eficiente del bus

La asignación del bus depende de su aplicación. Generalmente existen 2 tipos de clases de asignación.

- **Asignación en un tiempo fijo:** La asignación se hace secuencialmente a cada participante sin importar si este necesita el bus en ese momento. Esta técnica, asigna tiempo del bus a cada nodo, y si no tiene nada que enviar, el bus se encuentra ocioso.
- **Asignación en base a sus necesidades:** el bus se asigna según las necesidades del nodo (utilizando CSMA o CSMA/CD). En CAN la asignación del bus es negociado entre los mensajes que esperan ser transmitido. CAN utiliza esta técnica.

2.16.5. Capa de Enlace

CAN utiliza el control de acceso al medio tipo *CSMA/CS+CR* (Acceso múltiple con detección de portadora, detección de colisión más resolución de colisión). CAN resuelve el problema de la colisión con la supervivencia de una de las tramas que chocan en el bus. La trama “ganadora” es aquella que tiene mayor prioridad. Por lo tanto se puede indicar que CAN por naturaleza tiene en cuenta la prioridad.

Como ya se mencionó anteriormente el bit *dominante* es el 0 y el bit *recesivo* es el 1, la resolución se realiza con una operación lógica AND de todos los bits transmitidos simultáneamente. Cada transmisor se encuentra continuamente observando y comprobando que el bit recibido se corresponda con el bit que envía. Cuando no coincide, el controlador retira el mensaje del bus y se convierte automáticamente en receptor. Como puede observarse la capa de enlace se comparte de manera similar a la capa física.

La única diferencia que presenta la capa de enlace de CAN es que todos los errores a nivel de un solo bits son detectados. Los errores de múltiples bits son detectados con una alta probabilidad (CAN-CIA, 2017).

2.16.6. Formato del mensaje

CAN utiliza un formato de mensajes cortos (94 bits). En el mensaje no está explícito ninguna dirección, por este motivo el mensaje puede ser escuchado por todos los nodos de la red (Kvaser, 2017).

Los tipos de mensajes son los siguientes:

- Frame de datos (Data Frame)
- Frame remoto (Remote Frame)
- Frame de error (Error Frame)
- Frame de sobrecarga (Overload Frame)

2.16.6.1. Data Frame

Este es el frame más común. Las partes más importantes son:

- Campo de arbitraje, el cual determina la prioridad del mensaje
- El campo de datos, que contiene desde cero hasta ocho bytes de datos
- El CRC, que está conformado por 15 bits utilizados para calcular el checksum del mensaje
- Un campo de ACK. Cualquier controlador que haya recibido el mensaje envía un bit de acuse de recibo al final de cada mensaje. El transmisor comprueba la presencia del bit ACK. En caso de no detectar este bit reenvía el mensaje. Al no poder conocer la dirección de los nodos, no se sabe si el mensaje fue recibido correctamente por el nodo receptor, solo se sabe que el mensaje fue recibido por uno o más nodos.

2.16.6.2. Remote Frame

El frame remoto es un Data Frame con dos diferencias, es marcado como Remote Frame, esto es el bit RTR es recesivo; y por otro lado no hay un campo de datos. Este frame es utilizado para pedir la transmisión de un determinado frame de datos. Por ejemplo si el nodo A transmite un Remote Frame con el campo de arbitraje en 234, entonces el nodo B, responderá con un frame de datos con el campo de arbitraje seteado a 234 (Kvaser, 2017). Este frame es poco utilizado.

2.16.6.3. Error Frame

Este frame se envía cuando un nodo detecta alguna falla en el mensaje. El envío de este frame provoca la retransmisión inmediata del mensaje. El Error Frame consiste en una bandera, el cual está compuesto por 6 bits del mismo valor, y un Error Delimiter que está compuesto por 8 bits recesivos.

2.16.6.4. Overload Frame

Este frame es similar al frame de error con el mismo formato. Es enviado cuando el nodo está ocupado. Este Frame es muy poco usado. El único controlador que generaba Overload Frame está obsoleto (Kvaser, 2017)

2.16.6.5. CAN estándar y CAN extendido

El protocolo de comunicación CAN, es un protocolo de multiple acceso con detección de colisión y arbitraje según la prioridad de los mensajes (CSMA/CD+AMP) (Corrigan, 2002). CSMA ya fue estudiado en 2.13.1. CD+AMP significa que las colisiones son resueltas mediante arbitración por corrimiento de bits. El identificador con mayor prioridad es el que siempre gana el bus.

El CAN estándar (Figura 2.9) tiene los siguientes campos:

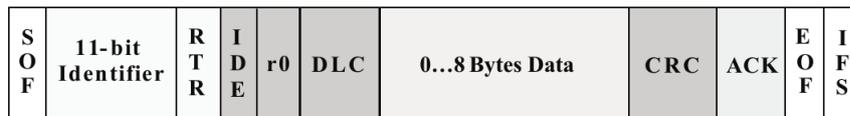


Figura 2.9: Frame de mensaje del CAN estándar

- SOF: es el bit dominante de inicio de frame, este marca el comienzo de un mensaje.
- Identificador: este es un identificador de 11 bits, en el cual el valor binario más bajo es el de mayor prioridad.
- RTR: El bit denominado Remote Transmission Request (RTR) es dominante cuando la información es requerida desde otro nodo. Todos los nodos reciben el pedido, pero el identificador determina el nodo específico.
- IDE: Es una extensión del bit identificador que significa que se está por transmitir un identificador estándar CAN sin extensión.
- r0: Bit reservado.
- DLC: 4 bits que contiene el número de bytes que van a transmitirse.
- Data: Hasta 64 bits, es el mensaje que se transmite.
- CRC: Son 16 bits que conforman el Cyclic Redundancy Check (CRC) para chequear la existencia de algún error.
- ACK: Todos los nodos que reciben el mensaje correctamente sobrescribe este bit recesivo, indicando que han recibido el mensaje sin errores. Si un nodo detecta un error y detecta la persistencia de este bit recesivo, sabe que ocurrió un error y se deberán tomar las medidas necesarias para hacer frente a ese error. Con esto, se descarta el mensaje enviado y el nodo transmisor vuelve a enviar el mensaje.
- IFS: Son 7 bits que representan el espacio entre frames. Es el tiempo que requiere el controlador para mover el mensaje hasta el buffer de mensajes recibidos.

El CAN extendido (Figura 2.10) es idéntico al CAN estándar, salvo por algunos detalles:

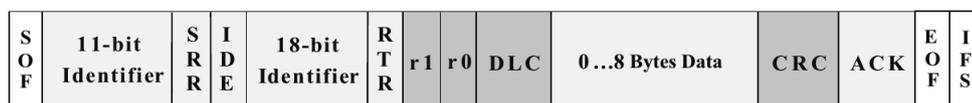


Figura 2.10: Frame del mensaje del CAN extendido

- SRR: Es el bit de petición remota substituta que reemplaza al bit RTR del mensaje estándar.
- IDE: Un bit recesivo en el identificador de extensión (IDE) indica que siguen más bits de identificación. A este le siguen 18 bits.
- r1: Este bit viene después del RTR, es un bit adicional reservado.

2.16.6.6. Arbitraje

Cualquier controlador CAN, cuando detecta que el bus está desocupado puede comenzar a transmitir. Esto puede ocasionar que dos o más controladores comiencen a transmitir al mismo tiempo. Este conflicto se resuelve de una manera ingeniosa. El nodo transmisor a medida que transmite los bits, se encuentra monitoreando el bus. Si por ejemplo, el nodo detecta un bit dominante cuando este envió un bit recesivo, el nodo inmediatamente deja de transmitir y se convierte en un receptor. Cuando el bus se encuentra desocupado, recién vuelve a transmitir (Kvaser, 2017).

Una condición importante que se debe cumplir es que dos o más nodos no pueden transmitir el mismo campo de arbitraje.

2.16.7. CANAerospace

CANAerospace es una definición de protocolos y datos que fue diseñada para ser utilizada en sistemas de comunicación altamente confiables, sobre todo aeroespaciales, que estén basadas en el Controller Area Network (CAN) (Stock Flight Systems - CANAerospace, 2017). Este, fue desarrollado por Stock Flight Systems, creado en 1997 y tiene su primer Release en 1998, y fue utilizado en varias aeronaves, como por ejemplo SOFIA Boeing 747SP²¹. También es utilizado en las interfaces de simuladores de vuelo. CANAerospace es un proyecto open source, y se continúa desarrollando. Fue publicado por NASA con el nombre de Advanced General Aviation Transport Experiments Databus Standard (AGATE).

Este protocolo nace con el propósito de crear un estándar para la creación de aplicaciones que requieran un constante monitoreo del flujo de datos y una sincronización de tiempo en sistemas con redundancias (Stock Flight Systems - CANAerospace, 2017).

El protocolo CAN, por sí solo, no cubre problemas como la representación de datos, direccionamiento, protocolo orientados a la conexión. Además, la utilización de CAN en aplicaciones espaciales requieren de la definición de un estándar que esté orientando a los requerimientos específicos de la misión. CANAerospace es creado para enfrentar estos problemas. Este protocolo es una pequeña capa de software que permite un manejo fácil del bus de datos que cumple con los requerimientos específicos de sistemas de aviónicas. CANAerospace fue instalado en una gran cantidad de aeronaves desde 1998 y ha demostrado una excelente confiabilidad en ambientes hostiles.

Algunas de las características más destacable de CANAerospace es que, no existen esclavos y maestros, por ello se dice que es una red democrática. Los mensajes se encuentran bien identificados. Existen mecanismos para informar eventos de emergencias. Es sencillo de aplicar. Y es un proyecto open source, por lo tanto no tiene costo alguno, y existen tutoriales libres y gratuitos.

2.16.7.1. Formato del mensaje CANAerospace

En la Figura 2.11 se puede observar el formato básico del mensaje de CANAerospace. Este está compuesto principalmente por cuatro bytes, los cuales se describen a continuación:

- *NODE-ID*: Este byte es utilizado cuando se produce un error en algún dispositivo y comienza a funcionar su redundancia. De esta manera, para aquellas arquitecturas que lo permiten, el protocolo puede identificar estos cambios.

²¹Ver: <https://www.sofia.usra.edu/public/about-sofia/sofia-aircraft>

- *Data-Type*: CANAerospace permite múltiples tipos de datos para cada mensaje. Este byte permite identificar cada tipo de datos de cada mensaje.
- *Service Code*: Para datos que se utilizan en las operaciones normales. Este dato debe reflejar el estado de los datos, esto ayuda a la implementación de un monitoreo integral de datos.
- *Message Code*: El número de mensajes permite detectar si el mensaje se perdió y verificar si la unidad de transmisión trabaja correctamente.

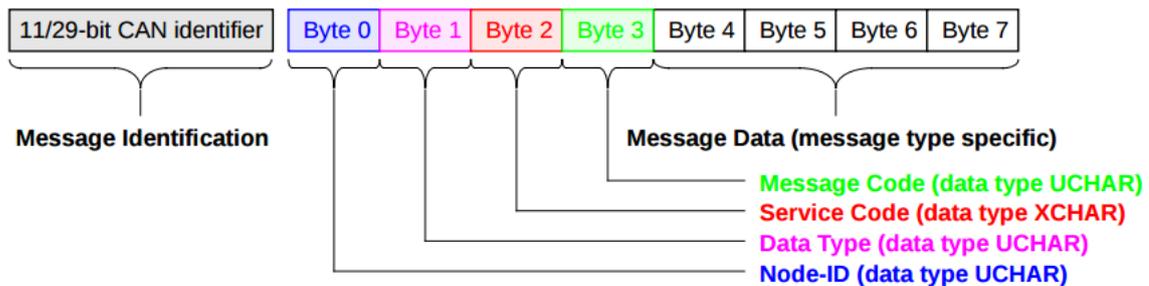


Figura 2.11: Formato del mensaje de CANAerospace

Resumen

En este capítulo se desarrollaron diferentes conceptos que son importantes ya que sustentan este trabajo de tesis.

Los principales conceptos que se necesitan conocer son:

- **Avería** de sistema ocurre cuando el servicio prestado por el sistema ya no coincide con las especificaciones del mismo. Esto quiere decir que existe un problema que tiene una consecuencia negativa en el sistema completo, logrando que este ya no logre cumplir con sus especificaciones
- Un **error** es una parte del estado del sistema que es susceptible de provocar un avería en el sistema. Un error que afecta al servicio, es una indicación de que un avería se ha producido.
- La causa adjudicada o la hipótesis de un error es una **falla**, también llamado “bugs”. Una falla es un defecto que está presente en el sistema y que puede causar un error. Es la desviación actual de lo correcto Hanmer (2007).

Por otro lado, la fiabilidad también se la puede considerar como una propiedad global que permite justificar la confianza de los servicios de un sistema. Esta puede ser clasificada en:

- Impedimentos, que son todas las cosas que se interponen en la fiabilidad de un sistema.
- Medios, son los medios para lograr la confiabilidad.

Los medios de la fiabilidad son los siguientes:

- Evitación de fallas, que son técnicas de mejoramiento de la fiabilidad utilizada durante el desarrollo de software para reducir el número de fallas introducidas durante esta etapa.

- Tolerancia de fallas, es la capacidad de un sistema a continuar funcionando a pesar de la ocurrencia de fallas.
- Eliminación de fallas, son técnicas utilizadas para mejorar la fiabilidad, que se emplean durante la validación y verificación del sistema.
- Predicción de fallas, esto se lleva a cabo mediante una evaluación del comportamiento del sistema con respecto a la ocurrencia o la activación de una falla.

Los atributos de la fiabilidad son los siguientes:

- Confiabilidad, es la probabilidad de que un sistema continúe operando correctamente durante un intervalo de tiempo dado.
- Disponibilidad, la disponibilidad $A(t)$ de un sistema en el instante de tiempo t es la probabilidad que el sistema esté funcionando correctamente en el instante t .
- Seguridad, es definida como la probabilidad que el sistema sea capaz de realizar sus funciones correctamente o discontinuar sus funciones en una manera a prueba de fallas.

En este trabajo de tesis se trabaja con el protocolo CAN. Este fue desarrollado por Bosh en 1983, para ser aplicado en la industria automotriz, como respuesta al rápido crecimiento de la electrónica en el fabricación de automóviles. Fue estandarizada por la Organización Internacional de Estandarización (ISO) bajo el nombre de ISO 11898. Este protocolo se definió con el objetivo de proveer comunicación determinística de sistemas distribuidos, y que necesiten un alto grado de confiabilidad. CAN permite la conectividad vía bus serial.

Las características de CAN que la convierten en un tentadora opción para la aplicación en el sector espacial, son:

- Bajo costo
- Operabilidad en ambientes eléctricos complicados
- Capacidades de tiempo real.
- Facilidad en el uso

Estado del arte

Locura es hacer lo mismo una y otra vez
esperando obtener resultados diferentes

Albert Einstein

En este capítulo se presenta el estado del arte del presente trabajo. Se pretende que el lector conozca los avances que se llevaron a cabo en el estudio de topologías tolerantes a fallas, tales como árboles binarios (Sección 3.1 página 41), redes hypercube (Sección 3.2 página 44), redes distribuidas (Sección 3.3 página 44), redes Ethernet (Sección 3.4 página 45), arquitecturas basadas en bus (Sección 3.5 página 47).

En la Sección 3.6 (página 48) las métrica y modelado de la confiabilidad de sistemas.

Para finalizar, en la Sección 3.7 (página 50) se comenta sobre CAN en la actividad espacial.

3.1. Árboles binarios

El concepto de arquitecturas de árboles binarios es aplicable en el desarrollo de sistemas de computadoras jerárquicas, y sobre todo en computadoras de alta performance (Raghavendra *et al.*, 1984). Existen dos diferentes mecanismos de tolerancia a fallas (Raghavendra *et al.*, 1984):

1. Esquemas con back up.
2. Esquemas con degradación de performance.

Teniendo en cuenta que estas arquitecturas son aplicadas principalmente en la construcción de circuitos VLSI¹ (Singh y Youn, 1991), se asume su aplicabilidad a arquitecturas de aviónica.

¹Del inglés, Very Large Scale Integration

La FT y la performance de los sistemas dependen de las capacidades de las redes que se utilizan para la comunicación entre unidades de procesamiento (Raghavendra *et al.*, 1984).

Un árbol binario está compuesto por nodos y enlaces (links). Existe un nodo central dónde se desprenden dos nodos hijos, estos se encuentran enlazados al nodo padre. Así recursivamente, se van generando dos nuevos hijos, por cada uno de los nodos. Los árboles binarios están divididos en niveles, que representa cada una de las generaciones de nodos.

Este tipo de topología tiene algunos problemas que la FT debe hacer frente. En las arquitecturas basadas en árboles binario, existe una cierta probabilidad de que un nodo o un link falle (Raghavendra *et al.*, 1984). Las arquitecturas de árbol binario son en general físicamente estáticas. Por lo tanto, cualquier falla en uno de sus nodos (o links) demandaría una avería a nivel sistema, lo cual daría lugar a una pérdida de misión. Para ello se debe dotar a la arquitectura de un mecanismo de reconfiguración.

La tolerancia a fallas en arquitecturas binarias ya fueron estudiadas en profundidad en Hayes (1976), Raghavendra *et al.* (1984), Singh y Youn (1991). Para lograr FT en estas arquitecturas se las deben diseñar con un número mínimo de nodos de backup y links redundantes, de modo tal de hacer frente cualquier punto de falla simple en la arquitectura.

3.1.1. Esquema de árbol binario con backups

El esquema planteado por Raghavendra *et al.* (1984) es similar al que se muestra en al Figura 3.1. En este esquema se agregan nodos y links redundantes como técnica de FT. Existe un nodo de backup por cada nivel del árbol.

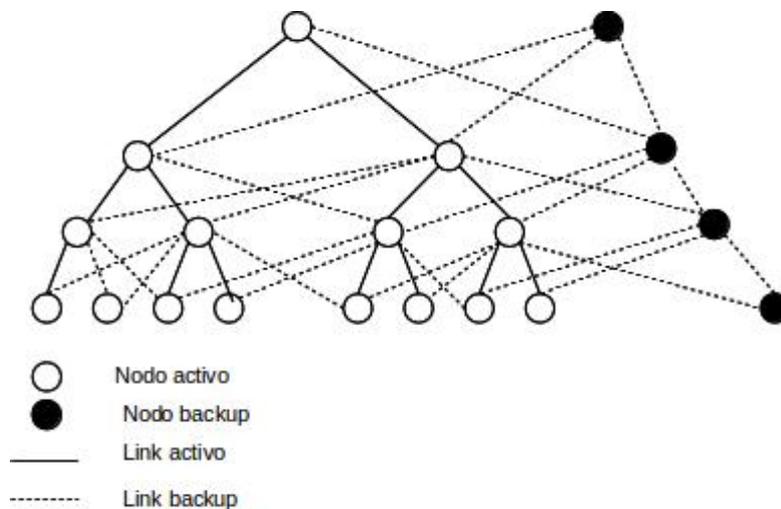


Figura 3.1: Árbol binario de 4 niveles

Esta arquitectura cuenta con una restricción, la cual indica que solo se puede tolerar una falla singular, por cada nivel de la arquitectura. Arquitecturas de este tipo, podrían tolerar más de una falla, sólo si se dan en diferentes niveles del árbol (Raghavendra *et al.*, 1984). Para agregar más tolerancia, se deberían agregar más redundancias.

Notese en la Figura 3.1 que cuando una unidad de procesamiento (nodo) falla, todos los links se deben reajustar hacia el nodo de la derecha. En este punto es importante mencionar, que ante esta situación se requiere una reconfiguración a nivel de HW, además de una reconfiguración a nivel de SW. Es necesario algoritmos de ruteo dinámicos, de modo tal de conocer los nuevos caminos que intercomunican nodos, para mantener la

operabilidad del sistema.

3.1.1.1. Estimación de la confiabilidad de un árbol binario con backup

Se asume que la probabilidad de fallas de los links es muy baja en comparación con la de los nodos. Teniendo que el rate de falla es de λ , la confiabilidad de un nodo es $R = e^{-\lambda t}$. También se sabe que un árbol binario con n niveles, tiene $2^n - 1$ nodos en total. Entonces la confiabilidad de todo el sistema es:

$$R_{nr} = R^{2^n - 1}$$

Raghavendra *et al.* (1984) incluye en sus cálculos un factor de cobertura c , el cual es la probabilidad condicional de que se lleve a cabo una recuperación exitosa, luego de que una falla se haya detectado. Entonces la confiabilidad del sistema para una arquitectura de árbol binario con redundancias (un nodo backup por nivel) es el siguiente:

$$R_{sys} = \prod_{k=0}^{n-1} [R^{2^{k+1}} + R^{2^k}(1 - R) + 2^k c R^{2^k}(1 - R)]$$

Simplificando:

$$R_{sys} = R^{2^{n+1}} \prod_{k=0}^{n-1} [(2^k c + 1) - 2^k c R]$$

En la Figura 3.2, se observa la confiabilidad de una arquitectura de árbol binario de 4 niveles, con una cantidad de $2^4 - 1 = 15$ nodos. En color negro se grafica una arquitectura sin redundancia, mientras que en color rojo, azul y cyan, se muestra arquitecturas redundadas con diferentes c (0.98, 0.99, 1, respectivamente).

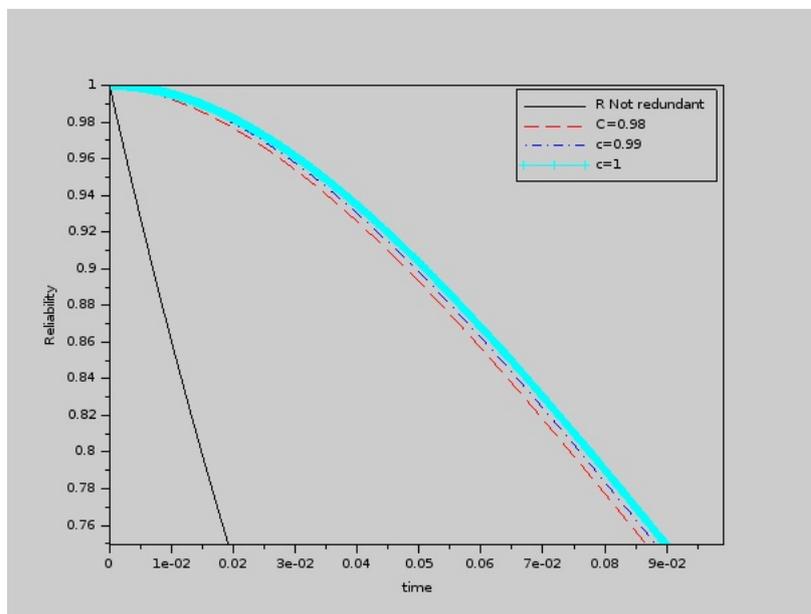


Figura 3.2: Confiabilidad con respecto al tiempo de una arquitectura de árbol binario de 4 niveles

Con esto podemos indicar, como es de esperarse, una arquitectura de árbol binario redundada permite mantener un alto nivel deseable de confiabilidad, durante un mayor lapso de tiempo, a diferencia de un sistema no redundando. También se puede concluir que con un factor de cobertura c más próximo a uno, maximiza los niveles de confiabilidad con respecto al tiempo.

3.2. Sistemas Hypercube

Hypercube es una técnica utilizada para conectar múltiples procesadores. Esta topología tiene propiedades de tolerancia a fallas de un componente, ya que si esto se produce, no se transmite a todo el sistema (Leu y Kuo, 1996). Otra ventaja que presenta esta topología es la disponibilidad de enlaces entre cualquier par de nodos (Abd-El-Barr y Gebali, 2014). Un hypercube n-dimensional tiene:

$$2^n$$

nodos y

$$n2^{n-1}$$

enlaces. Los nodos están conectados a n nodos vecinos a través de n enlaces (Leu y Kuo, 1996). Estas relaciones pueden ser representadas en forma de matrices de proximidad (Abd-El-Barr y Gebali, 2014). En Abd-El-Barr y Gebali (2014) se muestra una manera sencilla de calcular la confiabilidad de sistemas de este tipo. La probabilidad de falla de la red (p_n) depende de la probabilidad de falla de cada uno de los nodos (p_n), suponiendo que todos los nodos tienen la misma probabilidad de falla la confiabilidad del sistema sería:

$$R(N) = 1 - (NP_v)$$

A esta fórmula se le realizan leves modificaciones para obtener el siguiente cálculo de confiabilidad del sistema:

$$R_{sys} = 1 - [N(1 - e^{-\lambda t})]$$

3.3. Redes distribuidas

Una red de computadoras hace referencia a un conjunto de computadoras autónomas que se encuentran interconectadas, esto quiere decir que pueden intercambiar información (Tanenbaum y Wetherall, 2012). Una red distribuida tiene como principal característica que no existe un nodo central que “gestione” toda la red. Todas las cargas de las tareas y/o actividades son distribuidas entre los nodos que forman parte de la red.

Una red distribuida es tolerante a fallas si los nodos pueden formar subredes (Stivaros, 1992). Es decir, la red se debe mantener activa y conectada, con diferentes topologías e interconexiones, que permitan tolerar posibles fallas producidas en algunos nodos y permitir mantener la performance (Stivaros, 1992). Debido a que el procesamiento de la red se encuentra distribuida en todo el sistema, esto brinda una ventaja por encima de los sistemas centralizados desde el punto de vista de la confiabilidad (Pradhan y Reddy, 1982). Un componente importante de las redes distribuidas tolerantes a fallas, es la topología del sistema (Pradhan y Reddy, 1982).

Siguiendo la notación de Pradhan y Reddy (1982) para describir la topología del sistema se utiliza, un gráfico sin direccionamiento $G = \langle V, E \rangle$, donde V representa un set de nodos y E representa un set de relaciones. Stivaros (1992) agrega que $V(G) = \{v_1, v_2, v_3, \dots, v_n\}$ representa un vector de nodos, los cuales tienen probabilidades de operación $P = (p_1, p_2, p_3, p_n)$; y define una función de asignación π , la cual es una función que asigna V con una probabilidad $P_{\pi(v)}$.

La FT de la red G, dado el vector de probabilidades y una función de asignación π , tal como se viene discutiendo anteriormente, es la probabilidad de que la red continúe funcionando, es decir continúe conectada, aún en la falla (aleatoria) de alguno de sus nodos, esto se denota como $FT(G; \vec{P}, \pi)$.

Se dice que un subconjunto de nodos S , es un estado tolerante a fallas del sistema G , cuando estos nodos se mantienen conectados y funcionales. Se utiliza θ para indicar el conjunto de todos los S posibles. Un estado tolerante S contribuye $\prod_{v \in S} P_{\pi(v)} \prod_{v \notin S} (1 - P_{\pi(v)})$ a la probabilidad de FT.

Para calcular el total de la FT del sistema, incluyendo todo los S , se hace:

$$FT(G; \vec{P}; \pi) = \sum_{S \in \theta} Pr(S) = \sum_{S \in \theta} \prod_{v \in S} P_{\pi(v)} \prod_{v \notin S} (1 - P_{\pi(v)})$$

Una relación entre nodos es representado como ij , lo cual representa un enlace bidireccional entre nodos. El grado del nodo i representa el número de relaciones que inciden en ese nodo, el cual se escribe d_i . Así, d_i está limitado por el número de puertos de entrada y salida disponibles por cada nodo (Pradhan y Reddy, 1982). d_{ij} representa el número mínimo de *hop* (*hop* representa la transmisión a través de un link de datos) (Pradhan y Reddy, 1982)

Stivaros (1992) y Pradhan y Reddy (1982) mencionan la existencia de varias topologías que pueden ser aplicadas en una red distribuida tolerante a fallas. Una topología estrella posee una baja distancia entre nodos, pero una pobre tolerancia a fallas ((Pradhan y Reddy, 1982); (Stivaros, 1992)). La topología anillo, permite un simple ruteo, pero existen grandes distancias internodo. Un sistema completamente interconectado, presenta buenas características tolerantes a fallas, pero tiene un alto costo (Pradhan y Reddy, 1982). Pradhan y Reddy (1982) propone una topología para una arquitectura distribuida de comunicación. Esta es una topología robusta, y puede llegar a ser compleja a medida que aumentan los nodos.

3.3.1. Algoritmo de ruteo

Los algoritmos de ruteos son necesarios en el desarrollo de una arquitectura tolerante a fallas y reconfigurable. Los algoritmos de ruteo se los pueden dividir en dos: *algoritmo primario* y *algoritmo alternativo*. El primero es utilizado cuando no hay fallas de nodos. Se deben mantener los caminos para llegar, correctamente, al nodo destino. El segundo se utiliza en la presencia de alguna falla, este requiere que sea capaz de detectar la presencia de fallas, para luego reconfigurar el sistema. Estos algoritmos deben ser simples y requerir una mínima cantidad de HW y SW.

Agregado a lo mencionado en el párrafo anterior, deben existir algoritmos de diagnóstico distribuido de fallas, basados en los algoritmos de ruteo (Pradhan y Reddy, 1982).

3.4. Redes Ethernet en aviónica

TTEthernet es una tecnología de red de computadoras comercializada por TTTech Computertechnik AG para el desarrollo de aplicaciones seguras. SAE International² estandarizó esta red como SAE AS6802. TTEthernet se basa en el Ethernet clásico, en el cual se pone énfasis en las características principales que deben respetarse en sistemas críticos, tales como latencias de mensajes determinísticos, precisión de tiempo real, tolerancia a fallas (Loveless, 2015). Tiene la capacidad de transmitir datos 100 veces más rápido que la que lo hacen las tecnologías tradicionales tales como el MIL-STD-1553.

El Ethernet clásico presenta ventajas, tales como su alta velocidad de transmisión de datos, flexibilidad, y su disponibilidad y bajo costo (ya que se trata de un componente COTS) (Loveless, 2015), hacen deseable su aplicación en el área espacial. Fue utilizado en diferentes proyectos aeroespaciales y en misiones importantes tales como el Space Shuttle y la Estación Espacial Internacional (ISS) (Loveless, 2015). A pesar de esto, el

²<http://www.sae.org>

Ethernet no cumple con el determinismo requerido por las aplicaciones de tiempo real de un vehículo espacial. Por tal motivo, se desarrolla el sistema TTEthernet, el cual introduce un reloj de sincronización descentralizado, permitiendo la transmisión de mensajes Time-Triggered (TT). En este tipo de red, existe una herramienta de planning que asigna a cada dispositivo un intervalo de tiempo, en el cual puede utilizar para transmitir frames. Estos sistemas utilizan Links Virtuales (VL) para permitir el envío de mensajes time-triggered. Cada VL es asociado con un time-triggered frame a través de un identificador de tráfico crítico (CTID), este reemplaza el control de acceso al medio (MAC) (Loveless, 2015). Una red simple se muestra en la Figura 3.3.

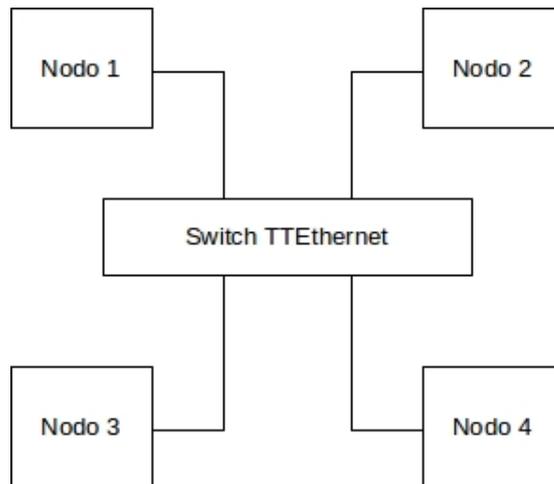


Figura 3.3: Arquitectura básica TTEthernet

TTEthernet puede actuar en dos clases de tráfico, con el objetivo de soportar diferentes niveles de criticidad de mensajes. Estas clases de tráfico son las siguientes (Loveless, 2015) (Steiner, Willfried, 2013):

- Time-Triggered, permite enviar mensajes de acuerdo a una planificación predefinida (scheduling),
- Rate-Constrained (RC), en el cual se llevan algunas restricciones de tamaño y rate de transmisión de frames,
- Best-Effort (BE), el cual se comporta de manera similar que el Ethernet

El paquete TTEthernet tiene una gran similitud con el frame del estándar IEEE 802.3 (Ethernet). Se denota en la Figura 3.4 que en lugar de la MAC del 802.3 frame se divide en el CT Marker y el CTID. El primero es un identificador estático utilizado para distinguir paquetes TT de otros tipos de tráfico.



Figura 3.4: Frame de mensaje TTEthernet

El estándar SAE AS6802 define el protocolo de sincronización del TTEthernet. El protocolo TTP fue diseñado para reducir la complejidad de las arquitecturas distribuidas tolerantes a fallas (TTTech, 2017). Existe un grupo de dispositivos de relojes locales que permite la sincronización requerida por la comunicación TT, a esto se lo denomina dominio de sincronización (Loveless, 2015). Cada dominio de sincronización es asignado a uno de los siguientes roles:

- Compression Master (CM)

- Synchronization Master (SM)
- Synchronization Cliente (SC)

3.4.1. Experiencia de vuelo

Loveless (2015) demuestra la aplicación de esta tecnología para computadoras de vuelo redundantes en misiones de naves simuladas. Esta tecnología toma tanto interés que el Sistema de Exploración Avanzada (AES)³ de la NASA lleva a cabo un proyecto denominado Avionics and Software (AS). TTEthernet también fue utilizada en una arquitectura tolerante a fallas en el Integrated Power, Avionics and Software (IPAS) del Johnson Space Center (JSC) y en el Core Flight Software (CFS) del Asteroid Redirect Mission (ARM) simulado⁴ (Loveless, 2015). También fue utilizado exitosamente durante la misión Orion (TTTech) . TTEthernet simplifica el diseño de sistemas espaciales que deben contar con tolerancia a falla y una alta disponibilidad. La seguridad y la redundancia se mantiene sin ningún tipo de aplicación extra.

3.5. Arquitectura de red basada en BUS

En Tai *et al.* (1999) se presenta una arquitectura tolerante a fallas basada en BUS. Esta topología forma parte de un programa denominado X2000, el cual tiene como objetivo desarrollar una arquitectura tolerante a fallas y basada en componentes COTS, perteneciente a NASA. Este programa se desarrolla bajo una filosofía de misiones espaciales la cual dice: “Rápido, mejor, barato”⁵.

Esta arquitectura utiliza una topología novedosa denominada *stack-tree topology* ((Chau *et al.*, 1999); (Tai *et al.*, 1999)). Una *stack-tree* es un árbol donde cada nodo rama se encuentra conectado como mucho a tres otros nodos de los cuales, como máximo, dos son nodos ramas (Tai *et al.*, 1999). Una Complete Stack-Tree (CST) es aquella en donde cada nodo rama está conectado al menos 1 nodo rama (Tai *et al.*, 1999).

En las Figura 3.5 se pueden observar ejemplos de *stack-tree*. Mientras que en la Figura 3.6 no es una *stack-tree*. Las imágenes fueron extraídas de Tai *et al.* (1999)

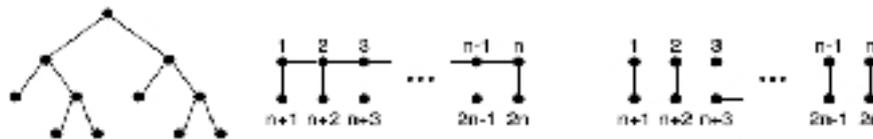


Figura 3.5: Arquitecturas stack-trees

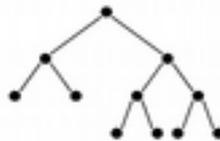


Figura 3.6: Arquitecturas que no responden al modelo stack-trees

Como el objetivo principal es desarrollar una arquitectura tolerante a fallas, esto es, que aún cuando un nodo o un link entre nodos falle, el sistema completo debe continuar funcionando, sin ningún tipo de degradación en

³Del ingles, Advanced Exploration Systems

⁴Los mencionados anteriormente son ejemplos de la utilización de TTEthernet

⁵En ingles, faster, better, cheaper

su servicio. Para cumplir con este objetivo Tai *et al.* (1999) trabajan con un esquema de denominado CST de esquema dual (CST_D)⁶. Esta puede observarse en la Figura 3.7 extraída de Tai *et al.* (1999).

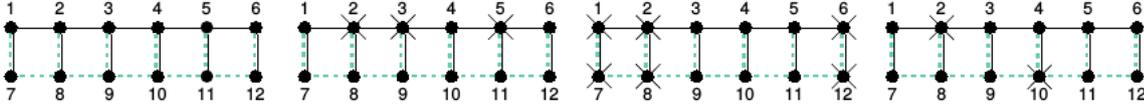


Figura 3.7: Esquema CST_D

Para aumentar la confiabilidad de la arquitectura ante la ocurrencia de fallas, se agrega links de backups que unen los nodos iniciales con el final, dando un efecto 3D de anillo (Tai *et al.*, 1999).

3.5.1. Evaluación de la confiabilidad de arquitecturas basadas en BUS

La confiabilidad de esta arquitectura (como ya se viene mencionando) es la probabilidad de que, a lo largo del tiempo de vida de la misión t , la arquitectura se encuentra funcional, en la cual todos los nodos (aquellos que no hayan fallado) se encuentren conectados (Tai *et al.*, 1999).

Tai *et al.* (1999) y Chau *et al.* (1999) asumen que la probabilidad de que un nodo falle es mucho mayor a que un enlace (el bus físico) falle.

Tai *et al.* (1999) indican que la confiabilidad de la red depende del tamaño k , siendo k el número de nodos ramas. La confiabilidad de una red basada en CST simplex es la probabilidad $U(k)$ de que los nodos no fallen, o que ante una falla, se detecte y se lleve a cabo correctamente la reconfiguración del sistema.

$$U(k) = (1 - q) \sum_{j=0}^k \binom{k}{j} (1 - q)^{k-j} (cq)^j$$

donde $q = 1 - e^{-\lambda t}$ es la probabilidad de que un nodo falle durante el tiempo de vida de la misión f . Entonces

$$R_s^{CST} = \sum_{k=1}^n (n - k + 1) U(k) (cq)^{2(n-k)}$$

donde $(cq)^{2(n-k)}$ es la probabilidad que los $2(n - k)$ nodos que conforman un cluster fallen, y esta falla es detectada y se lleva a cabo una correcta reconfiguración.

Para un CST dual ocurre de manera similar. Se define $V(k)$ de la siguiente manera:

$$V(k) = 2(1 - q)^k \sum_{j=1}^k \binom{k}{j} (1 - q)^{k-j} (cq)^j + (1 - q)^{2k}$$

Entonces: $R_D^{CST} = \sum_{k=1}^n (n - k + 1) V(k) (cq)^{2(n-k)}$

3.6. Métrica y modelado de la confiabilidad de sistemas

Es de suma importancia llevar a cabo el análisis de la confiabilidad, disponibilidad y mantenibilidad (RAM⁷) de sistemas satelitales, durante la fase de diseño (Hoque *et al.*, 2015), a fin de lograr la mínima cantidad

⁶En inglés, CST dual scheme

⁷Del inglés, Realibility, Availability, Maintainability

de fallas o incrementar el Tiempo Medio Entre Fallas (MTBF) (Peng *et al.*, 2013). Llevar esto a cabo es fundamental, ya que permite el desarrollo de estrategias que facilitan altos grados de confiabilidad, disponibilidad y mantenibilidad (Hoque *et al.*, 2015).

Existen dos categorías de medición de la confiabilidad y predicción, estas son utilizadas para asegurar la seguridad del software de sistemas críticos (Schneidewind, 1997), las cuales son:

- medición y predicción que están asociadas con las fallas y errores residuales.
- medición y predicción que están asociadas con la disponibilidad del sistema a sobrevivir durante la misión sin experimentar fallas (o pérdidas) en el sistema.

Las dos categorías mencionadas anteriormente son explicadas en Schneidewind (1997).

Según Liu *et al.* (2014) las severidades de las fallas son clasificadas como críticas, peligrosas o triviales, teniendo en cuenta la contribución de esa falla a la pérdida de la misión. Es importante, además, conocer el riesgo de una falla. El riesgo, se define como la posibilidad de que una falla produzca una lesión (por ejemplo, un astronauta en vuelos tripulados), algún daño material (por ejemplo, la destrucción del satélite), o una pérdida (por ejemplo, la pérdida de la misión).

Dependiendo de la misión, un criterio para definir si un sistema es seguro o no, es reduciendo las fallas que pueden provocar pérdidas de vida, de la misión o la obligación de abortar una misión (Schneidewind, 1997). Schneidewind (1997) define dos criterios que deben satisfacerse:

- $r(t_t) < r_c$,
- $T_F(t_t) > t_m$

dónde t_t es el Tiempo total de testing (observado o predicho); $r(t_t)$ son las fallas restantes hasta t_t ; r_c es un valor crítico de fallas restantes; $T_F(t_t)$ es la métrica para medir el riesgo; y t_m es la duración de la misión.

Lo anterior significa que un sistema crítico será seguro si: las fallas restantes en el tiempo de prueba son menores a un valor crítico de cantidad de fallas, o la duración de una misión es menor al tiempo que se de la siguiente falla.

En la literatura se utilizan modelos matemáticos para modelar los sistemas críticos y calcular así su confiabilidad. La mayoría de ellos asumen, que todas las fallas tienen igual tasa de detección de fallas, como así también la misma severidad, lo cual no es correcto (Liu *et al.*, 2014). Las técnicas de verificación formal tienen un fuerte enfoque para la verificación tanto de sistemas complejos, como de las propiedades deseadas de los sistemas (Peng *et al.*, 2013). Es importante llevar a cabo una verificación de modelo⁸, esta incluye técnicas de verificación automática para sistemas de estados concurrentes finitos. En estos modelos, las especificaciones de los sistemas son escritos en una lógica temporal y proposicional, y el procedimiento de verificación es una búsqueda del espacio de estado del diseño (Hoque *et al.*, 2015).

En Hoque *et al.* (2015) se utiliza la *verificación probabilística de modelos*, la cual es utilizada para verificar sistemas, de los cuales los comportamientos son estocástico por naturaleza. Este se basa en la construcción y análisis de modelos probabilísticos como cadenas de Markov. En Hoque *et al.* (2015) se indica que estas técnicas fueron aplicadas en misiones de NASA. Los modelos de Markov son muy utilizados para los análisis de confiabilidad y disponibilidad de sistemas complejos (Hoque *et al.*, 2015).

En Hoque *et al.* (2015) se lleva a cabo un modelo simplificado de un sistema satelital el cual se muestra en la Figura 3.8. La descripción del modelo se llevó a cabo usando el lenguaje PRISM. El modelo que se describe

⁸Model checking

en Hoque *et al.* (2015) y Peng *et al.* (2013) se refiere a un modelo a nivel de sistema/misión, muy diferente a lo que se pretende realizar en este trabajo de tesis, pero es factible basarse en los conceptos que describe la literatura, ya que lograron aumentar la confiabilidad del modelo propuesto, en contraposición de modelos “clásicos”.

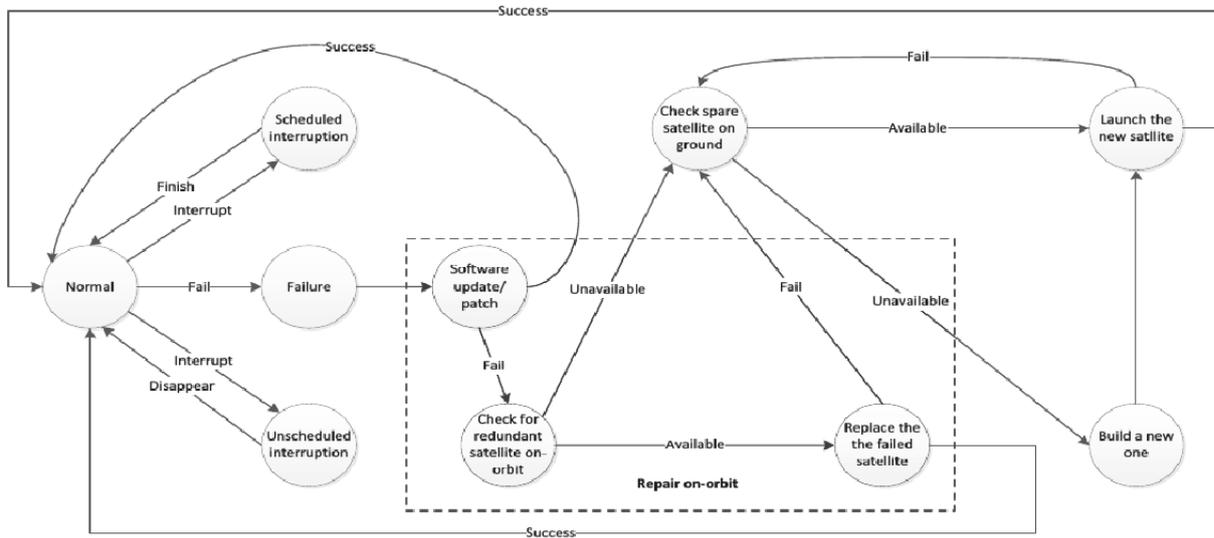


Figura 3.8: Modelo de sistema satelital (Hoque *et al.*, 2015)

Hoque *et al.* (2015) y Peng *et al.* (2013) relacionan tasa de falla λ , confiabilidad R_e y MTBF con la siguiente ecuación

$$\lambda = \frac{-\ln R_e}{MTBF}$$

Esteve *et al.* (2012) muestra el caso de estudio del desarrollo de una plataforma satelital a un alto nivel conceptual. Demuestran la correcta utilización de métodos formales y herramientas para la industria satelital. Se utiliza la herramienta desarrollada por el consorcio COMPASS⁹ para el modelado formal y análisis, que es utilizado por la industria espacial de Europa. Llegan a la conclusión de que mediante la utilización de modelos formales, se logra desarrollar un modelo completo, que cumple con todos los aspectos necesarios. Esto asegura la consistencia de los análisis (Esteve *et al.*, 2012).

El estado de la cuestión sobre la temática tratada en esta sección, indica que debería utilizarse modelos formales para el análisis de los sistemas. Se llega a la conclusión de que no existe un método (o modelo) completo y exclusivo para lo que se pretende desarrollar en este trabajo de tesis. Por ello se deberá utilizar y moldear las metodologías vistas en esta sección, para lograr cumplir con los objetivos del presente trabajo.

3.7. CAN en la actividad espacial

CAN ya es utilizado en misiones satelitales, pero en la mayoría de los casos es utilizado como bus secundario de comunicación, dejando el bus primario para otros medios de comunicación más clásicos como el MIL-STD-1553B.

⁹Correctness, Modeling and Performance of Aerospace Systems

La Agencia Espacial Europea (ESA) es la principal organización que se esfuerza en desarrollar hardware, firmware y software que implementa CAN como sistema de comunicación y control a bordo de vehículos espaciales. Esta organización definió el estándar ECSS-E-ST-15C con el objetivo de estandarizar el protocolo de comunicación CAN. La ESA observa que existe una tendencia de producir un cambio de paradigma centralizado, a funciones autónomas distribuidas, además, tiene en cuenta las ventajas que brinda CAN para el desarrollo de satélites. Por ello, lleva a cabo anualmente el *CAN in Space Workshop*¹⁰.

Debe destacarse que CAN tiene vasta experiencia de vuelo en aviónica de vehículos aéreos (helicopteros y aviones). Lo cual alienta su aplicación en la industria espacial.

Resumen

Un **árbol binario** está compuesto por nodos y enlaces (links). Existe un nodo central dónde se desprenden dos nodos hijos, estos se encuentran enlazados al nodo padre. Así recursivamente, se van generando dos nuevos hijos, por cada uno de los nodos. Los árboles binarios están divididos en niveles, que representa cada una de las generaciones de nodos. Para lograr FT en estas arquitecturas se las deben diseñar con un número mínimo de nodos de backup y links redundantes, de modo tal de hacer frente cualquier punto de falla simple en la arquitectura.

Hypercube es una técnica utilizada para conectar múltiples procesadores. Esta topología tiene propiedades de tolerancia a fallas de un componente, ya que si esto se produce, no se transmite a todo el sistema. Otra ventaja que presenta esta topología es la disponibilidad de enlaces entre cualquier par de nodos.

Una **red distribuida** tiene como principal característica que no existe un nodo central que “gestione” toda la red. Todas las cargas de las tareas y/o actividades son distribuidas entre los nodos que forman parte de la red. Una red distribuida es tolerante a fallas si los nodos pueden formar subredes. Es decir, la red se debe mantener activa y conectada, con diferentes topologías e interconexiones, que permitan tolerar posibles fallas producidas en algunos nodos y permitir mantener la performance.

TTEthernet es una tecnología de red de computadoras comercializada por TTTech Computertechnik AG para el desarrollo de aplicaciones seguras. SAE International estandarizó esta red como SAE AS6802. TTEthernet se basa en el Ethernet clásico, en el cual se pone énfasis en las características principales que deben respetarse en sistemas críticos, tales como latencias de mensajes determinísticos, precisión de tiempo real, tolerancia a fallas. Tiene la capacidad de transmitir datos 100 veces más rápido que la que lo hacen las tecnologías tradicionales tales como el MIL-STD-1553.

Para medir la confiabilidad y, que es utilizada para asegurar la seguridad del software de sistemas críticos, existen dos categorías

- medición y predicción que están asociadas con las fallas y errores residuales.
- medición y predicción que están asociadas con la disponibilidad del sistema a sobrevivir durante la misión sin experimentar fallas (o pérdidas) en el sistema.

Por último en este capítulo se comenta que el protocolo de comunicación CAN ya es utilizado en misiones satelitales, pero en la mayoría de los casos es utilizado como bus secundario de comunicación, dejando el bus primario para otros medios de comunicación más clásicos como el MIL-STD-1553B.

¹⁰<https://indico.esa.int/indico/event/162/>

La ESA observa que existe una tendencia de producir un cambio de paradigma centralizado, a funciones autónomas distribuidas, además, tiene en cuenta las ventajas que brinda CAN para el desarrollo de satélites. Por ello, lleva a cabo anualmente el *CAN in Space Workshop*

Análisis y desarrollo de arquitectura tolerante a fallas

La raza humana necesita un desafío intelectual.
Debe ser aburrido ser Dios y no tener nada que descubrir

Stephen Hawking

En este capítulo se lleva a cabo un estudio de la confiabilidad de tres tipos de topologías tolerantes a fallas (Sección 4.3 página 55), candidatas a ser utilizada en el diseño de la arquitectura propuesta en este trabajo de tesis.

Para ello, en la Sección 4.1 (página 53) se realiza una descripción de algunos requerimientos (no estrictos), que guiarán el desarrollo de una arquitectura para un misión satelital ficticia basada en componentes COTS. En base a estos “requerimientos”, se presentan los modelos para la medición de confiabilidad de las topologías mencionadas en el párrafo anterior.

En la Sección 4.6 (página 60) se describe resumidamente el protocolo CANae 0.1 Alpha, que está basado en CAN y que fue desarrollado en este trabajo de tesis. En el apéndice A se describe este protocolo con mayor detalle

4.1. Requerimientos para el análisis

En esta sección no se pretende realizar una lista de requerimientos formales para el desarrollo de una arquitectura. El objetivo de esta sección, es llevar a cabo una guía sencilla de las partes principales de un sistema satelital. Con esto en mente se podrá desarrollar diferentes topologías de comunicación, y luego analizar su tolerancia a falla.

Se supondrá una misión de 15 años (sin carga útil para simplificar el análisis) cuyo sistema estará compuesto por los siguientes subsistemas (en inglés para mantener correspondencia con la literatura) basado en Fortescue *et al.* (2003):

- Power Subsystem
- Attitude and Orbit control Subsystem
- Telemetry and Subsystem
- Thermal Subsystem
- Propulsion Subsystem
- Data Handling Subsystem

El sistema, entonces, está compuesto por 6 nodos. Los nodos se suponen computadoras con capacidad de procesamiento suficiente para cada subsistema. Cada una de estas computadoras/nodos es un componente COTS con un cierto grado de confiabilidad (se suponen lo suficientemente bajo como para no ser utilizado en forma directa en el desarrollo de satélites). A nivel de HW estos nodos cuentan con tolerancia a fallas, lo que aumenta su confiabilidad. El subsistema de Data Handling tiene que tener comunicación con todos los subsistemas, ya que será la encargada de controlar el correcto funcionamiento del sistema, enviar comandos, y empaquetar telemetría de los sensores.

4.2. Nomenclatura

Durante el análisis se utilizará la siguiente nomenclatura:

T_m	Tiempo de la misión
$TTNF$	Tiempo hasta la siguiente falla
λ	Tasa de falla
$MTBF$	Tiempo medio entre fallas
$MTTR$	Tiempo medio de reparación
$A(t)$	Disponibilidad
$R(t)$	Confiabilidad

Como se definió anteriormente el T_m es de 15 años. Para simplificar los trabajos de cálculos y ejecución del presente trabajo, se supone que la arquitectura puede fallar solo una vez durante la misión satelital. Por lo tanto, el $TTNF$ será de 15 años. La tasa de falla se define como:

$$\lambda = \frac{1}{15}$$

Suponiendo que la arquitectura no debería fallar durante los 15 años de misión se da la siguiente situación:

$$MTTBF = MTTR = 15$$

Por otro lado, La disponibilidad es $A(t) = 99\%$. La confiabilidad, como se estudió en secciones anteriores, es $R(t) = e^{-\lambda t}$

4.3. Estudio de topologías de arquitecturas

Luego de un estudio exhaustivo del estado de la cuestión (Capítulo 3 Página 41) se llegó a la conclusión de que las topologías más estudiadas, por ende más maduras y sencillas de aplicar a las actividades que se pretenden realizar en la presente tesis son:

- Árbol binario
- Red distribuida
- Arquitectura hypercube

En esta sección, se definirá modelos para medir la confiabilidad de cada una de las topologías de arquitectura que aseguren una mayor tolerancia a fallas a nivel de sistema, de modo tal que si se llega a producir una falla en cualquiera de los nodos (sistemas de procesamiento) de la arquitectura, esta puede reconfigurarse, permitiendo que esta continúe funcionando, sin sufrir ningún tipo de degradación, aún en la presencia de fallas.

4.3.1. Árbol binario

En primer lugar se planteó un árbol binario de cuatro niveles con backup, basandose en el diseño de Raghavendra *et al.* (1984) 3.2. Este diseño cuenta con $2^n - 1 = 15$ nodos. De lo estudiado en Sección 3.1 la confiabilidad puede ser calculada de la siguiente manera:

$$R_{sys} = R^{2n+1} \prod_{k=0}^{n-1} [(2^k c + 1) - 2^k c R]$$

Con esto se puede observar la confiabilidad de la red con respecto al tiempo en la Figura 4.1.

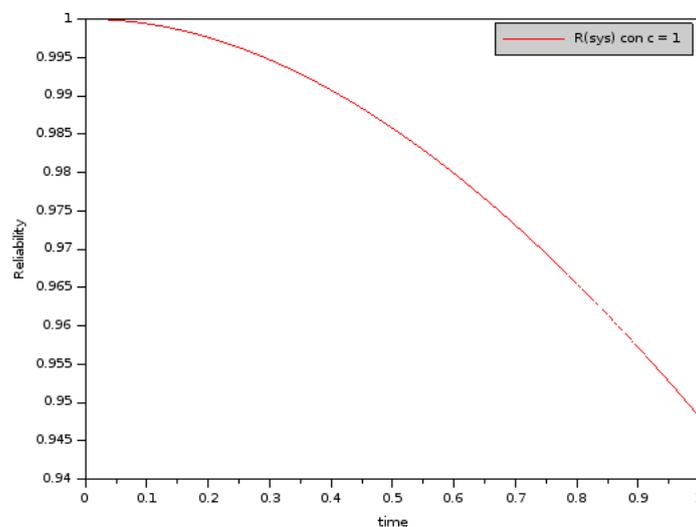


Figura 4.1: Confiabilidad con respecto al tiempo de árbol binario de 4 niveles

La cantidad de niveles que se eligió para esta topología depende del número de subsistemas que se requieren.

Los nodos backup se mantienen inactivos, es decir no participan en el procesamiento durante la vida normal del sistema. En caso de producirse una falla, se supone que estos nodos de redundancia, comienzan a funcionar automáticamente, sin ningún tipo de retraso. Esto, que no corresponde con la realidad, permite simplificar los cálculos para este trabajo de tesis.

4.3.2. Red distribuida

El estudio de una topología de red distribuida no es tan sencilla como la que se plantea para un árbol binario. Para el desarrollo de esta red, además de los 6 nodos que representa cada uno de los subsistemas requeridos, se agregan 2 nodos de redundancia (este este caso se opta por nodos de redundancias activos). Por lo tanto se tiene una red de 8 nodos. Siguiendo la metodología de desarrollo presentado por Pradhan y Reddy (1982), se llevó a cabo la red que se presenta en la Figura 4.2.

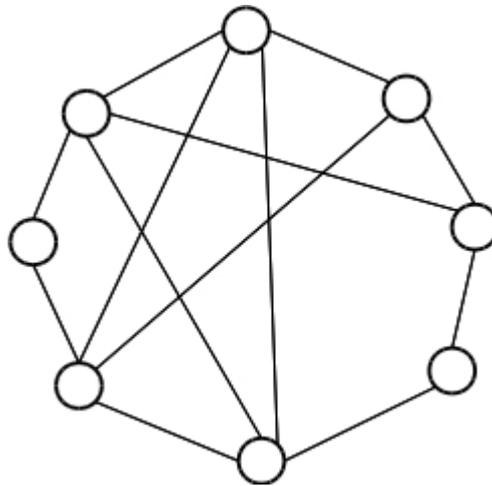


Figura 4.2: Red distribuida

Esta cuenta con 4 nodos de grado 4, 2 nodos de grado 3, y 2 nodos de grado 2. Ante cualquier falla de alguno de los nodos de la red, esta topología tiene la capacidad de formar subredes, que mantienen todos los nodos conectados (a excepción del fallado), asegurándose la funcionalidad y reconfiguración del sistema completo. Estratégicamente y para lograr la condición mencionada anteriormente, se crea la **condición** de que pueden fallar hasta 4 nodos en simultaneo. Esto aseguraría de que la red continuará funcionando aún en la presencia de fallas (definición de tolerancia a fallas).

Se modificó la fórmula desarrollada por Stivaros (1992), para lograr una coherencia entre los modelos que se plantean en el presente trabajo. Teniendo en cuenta que la confiabilidad del sistema completo es:

$$R(t) = \prod_{v \in S} e^{-\lambda t}$$

donde v representa el nodo y S es el subsistema funcional. Es decir, que este modelo recorre todos los nodos funcionales. Cuando existen nodos con fallas, y que dejan de ser funcionales, el modelo es el siguiente:

$$R_{sys} = \sum_{i=0}^k ((\prod_{v \in S} R(t)) - (\prod_{v \notin S} (1 - R(t)c)))$$

donde c se definió como una *constante de degradación del sistema* para modelar una degradación del sistema (c es definida en la subsección 3.1.1.1).

4.3 ESTUDIO DE TOPOLOGÍAS DE ARQUITECTURAS

En la Figura 4.3 se puede observar la *curva de confiabilidad* del sistema, para el caso en el que todos los nodos se encuentren funcionales.

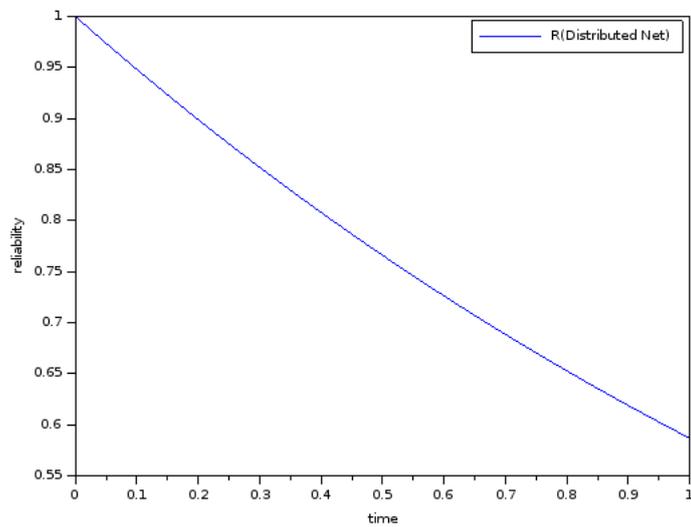


Figura 4.3: Confiabilidad de red distribuida

Para el caso de la falla de todos los nodos la *curva de confiabilidad* (Figura 4.4) muestra correctamente la degradación esperada de la confiabilidad, con respecto al sistema funcionando correctamente sin ninguna falla.

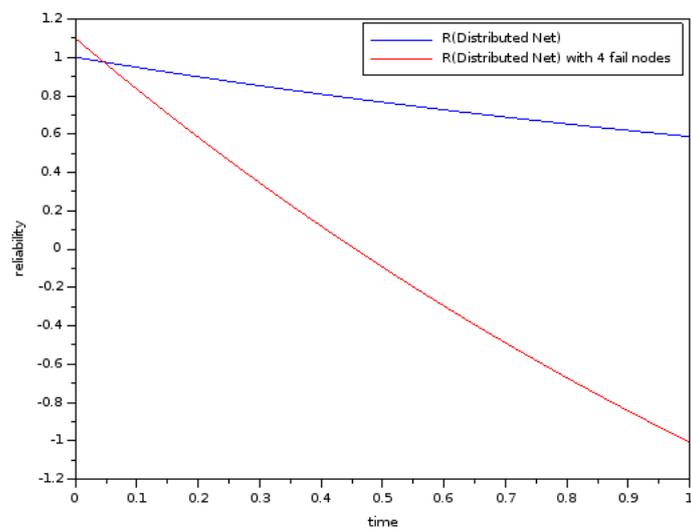


Figura 4.4: Confiabilidad de red distribuida con 4 nodos fallando

4.3.3. Red hypercube

Para el caso de la red hypercube se llevaron a cabo modificaciones al modelo planteado por Abd-El-Barr y Gebali (2014), con el propósito de mantener coherencia en los modelos y cálculos que se realizan en este trabajo. Se diseñó una red 3-dimensional, con 8 nodos, de los cuales 6 nodos corresponden a los diferentes subsistemas y 2 nodos son utilizados como redundancias activas (Figura 4.6).

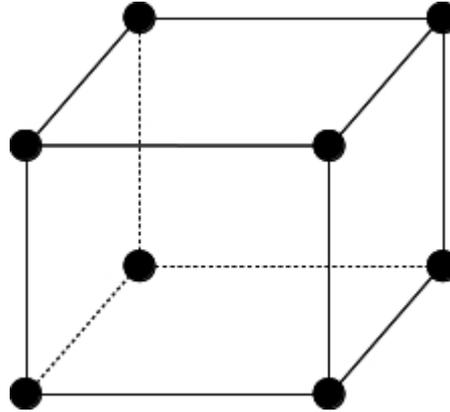


Figura 4.5: Red Hypercube

En este trabajo de tesis, la confiabilidad se calculó por medio del siguiente modelo:

$$R_{sys} = 1 - [N(1 - e^{-\lambda t})]$$

Como se puede observar el modelo es similar al modelo de árboles binario. Como se pudo estudiar en la bibliografía, árboles binarios y redes hypercube presentan varias similitudes, incluso se emebebe árboles binarios en este tipo de red. La *curva de confiabilidad* de esta red se observa en la Figura 4.6

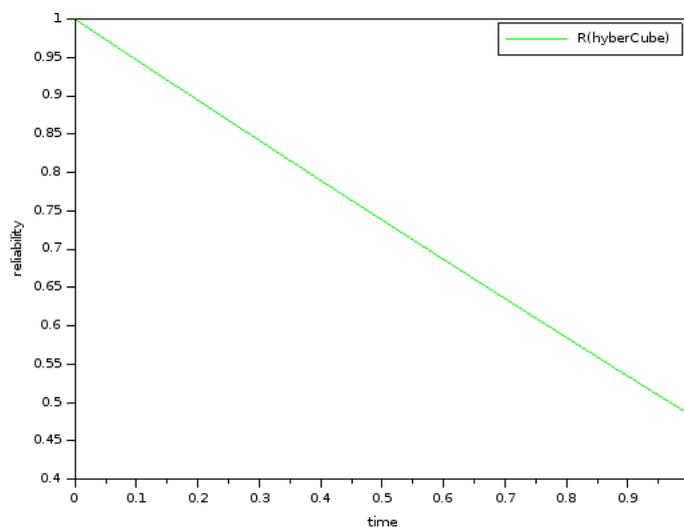


Figura 4.6: Confiabilidad de red hypercube

4.4. Topología utilizada en la arquitectura a diseñar

Teniendo en cuenta los modelos presentados anteriormente, se procede a realizar una comparación de las diferentes curvas. El resultado de esto, permitirá conocer de manera analítica qué topología presenta una mayor tolerancia a fallas a través del tiempo. Además, teniendo en cuenta lo estudiado en el estado de la cuestión, se puede realizar una comparación conceptual de las tres topologías.

La aplicación de árboles binario podría resultar una buena opción, ya que representa un desarrollo sencillo. En contraposición se puede indicar que existe un alto grado riesgo de que se produzca una falla en el nodo raíz y de su redundancia, lo cual pondría en peligro la misión. Así mismo, presenta otro punto negativo que se puede mencionar y es la gran cantidad de enlaces que esta necesita para mantener a todos los nodos de la red conectados.

Por otra parte, la red distribuida cuenta con la capacidad de distribuir el trabajo en todos sus nodos. Esto quiere decir, que si se produce una falla irreparable en cualquiera de sus nodos, la arquitectura podría continuar funcionando sin verse afectada por la ausencia de dicho nodo. Esto demanda un procesamiento computacional extra, y la necesidad de desarrollar algoritmos de ruteo especiales. Además, como punto negativo se puede mencionar que también, al igual que los árboles binarios, necesitan una gran cantidad de enlaces.

Por último, las topologías hypercube tienen un excelente respaldo teórico, exigen menor cantidad de enlaces, y pueden tolerar la falla de una gran cantidad de nodos (hasta el 50% de los nodos). Su complejidad aumenta en gran medida, cuando se desarrollan arquitecturas de más dimensiones, lo cual también incrementa su confiabilidad.

Como se mencionó en el primer párrafo, se realizó una comparación de modelos de confiabilidad de las tres topologías estudiadas. Se asumió que la distribución de la confiabilidad es exponencial, con una tasa de falla fija de $1/15$, y se estudió su evolución en un rango $[0, 1]t$. El resultado de esta comparación se la puede observar en la Figura 4.7 y en la Tabla 4.1.

Se observa que tanto las redes distribuidas como la hypercube presentan una mayor confiabilidad a través del tiempo que los árboles binarios. Si bien, las redes distribuidas y la hypercube tienen una curva similar, la primera presenta una mayor confiabilidad sostenida en el tiempo.

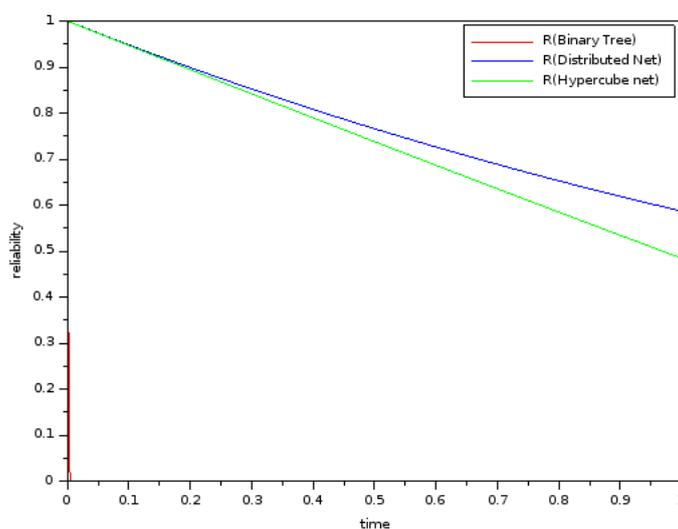


Figura 4.7: Comparación de confiabilidad

Tabla 4.1: Comparación de confiabilidad de topologías

T	Topologías de red		
	Tree Net	Distr Net	Hyper Net
0	1	1	1
0.001	0.803766	0.999947	0.999947
0.002	0.64604	0.999893	0.999893
0.003	0.519265	0.99984	0.99984
0.004	0.417368	0.999787	0.999787
0.005	0.335466	0.999733	0.999733
...
0.995	0	0.948317	0.947109
0.996	0	0.948266	0.947056
0.997	0	0.948216	0.947003
0.998	0	0.948165	0.94695
0.999	0	0.948115	0.946897

Con esto se puede concluir que la topología que presenta un mayor grado de confiabilidad es la que responde a una filosofía distribuida (bajo las condiciones en las que fueron estudiadas). Por lo tanto, la arquitectura satelital, tolerante a fallas y basada en componentes COTS que se desarrolla en la presente tesis se basa en una **topología distribuida** para interconectar los diferentes subsistemas.

4.5. Topología propuesta

Sobre la base de los resultados presentados en (Arias y Wiman, 2017) se puede establecer que la topología propuesta (red distribuida) es la más adecuada para el desarrollo de una arquitectura tolerante a fallas como se demuestra en la Figura 4.8. En esta se puede observar que cada subsistema (thermal, power, telemetry, etc.) tiene su propia CPU controladora. Estas CPU se conectan a los nodos COTS de la red.

El modelo exige como requerimiento que cada nodo debe estar compuesto por una computadora (componente COTS) que es la encargada de realizar el procesamiento de las tareas. También, debe existir un puente de comunicación entre la red y la CPU del subsistema. De este modo se hace frente a posibles fallas en la computadora del nodo. Esta conexión se observa en la Figura 5.7

4.6. Protocolo de comunicación

Como se estudió en el marco teórico, en una arquitectura de comunicación, el protocolo CAN trabaja en las capas inferiores del modelo de OSI. CAN propone cómo debe ser el medio físico, y la manera de transmisión de las señales. Además, prevé cuál es la estrategia a utilizar para sincronizar la transmisión de datos sin ningún tipo de latencia, ni tampoco colisiones de mensajes. Así, este protocolo permite que estos sean enviados teniendo en cuenta sus prioridades, eliminando las colisiones, y por lo tanto también, el tiempo de espera en que los nodos se quedan relegados cuando han tratado de enviar un mensaje al mismo tiempo.

Este protocolo no establece nada sobre rutas de mensajes, tipos de mensajes y tipos de datos, los cuales se deben tener en cuenta a la hora de desarrollar una arquitectura de aviónica. Es necesario, que exista una capa de servicios que facilite la comunicación entre aplicaciones de usuarios de diferentes nodos, y la comunicación entre las aplicaciones de usuario y el protocolo CAN (de bajo nivel) en sí. Además, teniendo en cuenta los

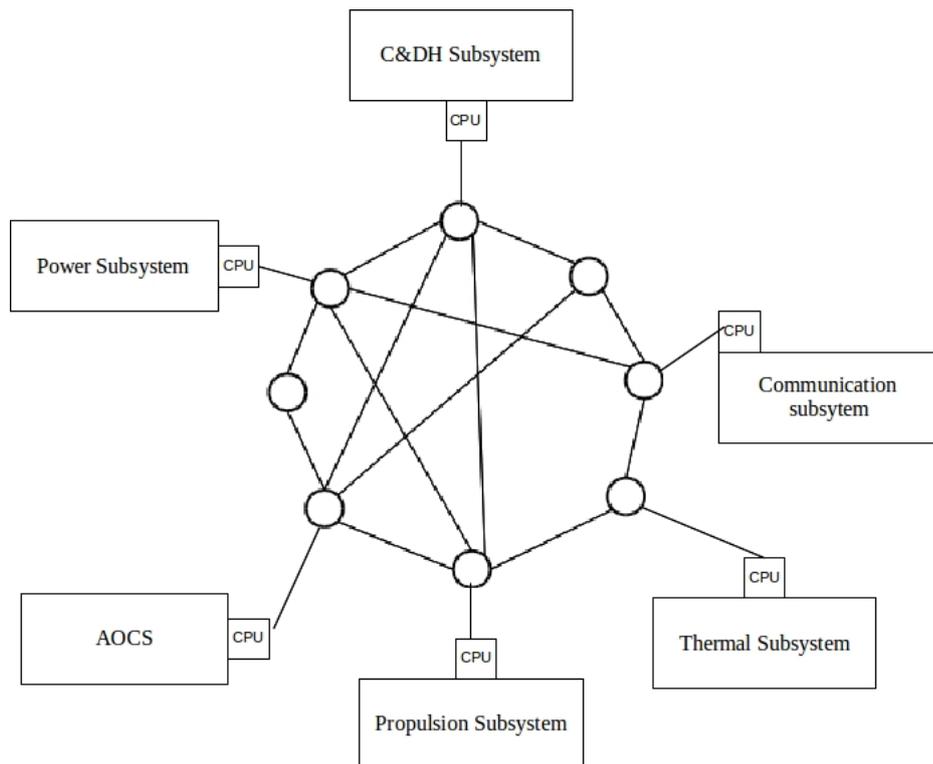


Figura 4.8: Arquitectura propuesta utilizando topología de red distribuida

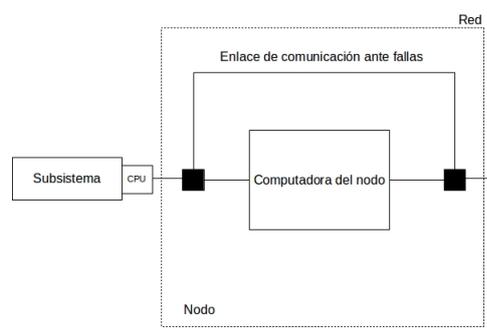


Figura 4.9: Conexión entre la red y el subsistema

resultados obtenidos en las secciones anteriores, se necesita, diseñar un protocolo de comunicación basado en redes distribuidas, y que a su vez, esté basado en CAN.

Por tales motivos, surge la necesidad de diseñar un protocolo de comunicación basado en CAN, que se “monte” sobre las capas superiores del modelo de referencia de OSI; y que permita la distribución de las tareas y el procesamiento llevado a cabo por los nodos.

El protocolo CANae se encuentra en su primera versión 0.1 Alpha. Esto hace referencia a que en esta instancia de trabajo se trata de comprender el problema y llevar a cabo un diseño preliminar del protocolo. De esta manera se podrá extraer, tanto puntos positivos, como puntos negativos; o más bien fortalezas y debilidades del stack de servicios que brinda CANae.

CANae actúa en la capa de aplicación del modelo de referencia de OSI. CANae divide esta capa en dos subcapas:

- CANae Application Layer
- CANae High Application layer

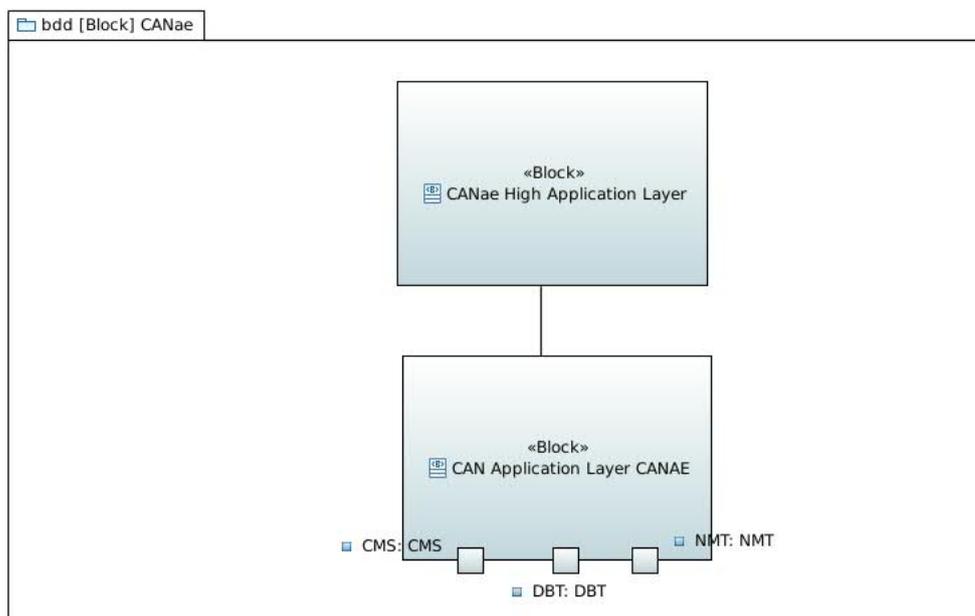


Figura 4.10: Estructura de la capa de aplicación de CANae en alto nivel

Esto puede observarse en la Figura 4.10. En el gráfico se observa que la CANae application Layer, cuenta con 3 interfaces:

- **CMS:** ofrece un ambiente orientado a objeto para diseñar aplicaciones de usuario. Esta entidad ofrece variables y eventos, y específica como un módulo puede acceder a las interfaces de CAN.
- **NMT:** ofrece un ambiente orientado a objetos para permitir que un módulo (el NMT Master) se ocupe de la inicialización y posibles fallas de otros módulos (NMT Slaves).
- **DBT:** ofrece el servicio para distribuir dinámicamente el identificador para los diferentes nodos.

Debe destacarse que, el servicio CMS ofrece un ambiente orientado a objeto para diseñar aplicaciones de usuario, es decir, que ofrece la posibilidad de modelar el comportamiento del sistema en forma de objeto. Este servicio también permite la posibilidad de crear variables y eventos, según las necesidades de los usuarios, que son utilizadas para diseñar y especificar como la funcionalidad de un módulo puede acceder a las interfaces de bajo nivel de CAN. Esto propone un grado de innovación al tratar al protocolo, a los datos, mensajes y eventos como objetos, permitiendo así un modelado orientado a objetos. Por tal motivo, este protocolo fue modelado con SysML (System Modeling Language). Para mayor detalle de estas interfaces se debe consultar en el apéndice A.

Agregando, dentro de esta capa existen dos entidades que juegan un papel importante (consultar apéndice A) las cuales son:

- **Gestor de mensajes:** este se encarga de gestionar los mensajes que son enviados y recibidos desde la red. Esta entidad debe ser capaz de determinar si el mensajes contiene datos o eventos. Trabaja en conjunto con el CMS.

- **Gestor de nodo:** esta entidad se encarga de llevar el control de los nodos existentes en la red. En este nodo se encuentra la tabla de ruteo primario y secundario necesarios para la correcta comunicación.

La subcapa denominada CAN High Application Layer tiene una función más del lado de la gestión de nodos y tareas. En esta capa se desarrollan los algoritmos necesarios para realizar el ruteo y la reconfiguración de la red ante fallas, por lo tanto en la High Application Layer se debe implementar el sistema de FDIR. El protocolo no define ningún algoritmo de ruteo, por lo que queda para el usuario la definición de los algoritmos. El protocolo recomienda desarrollar dos algoritmos, el primario y el secundario. Para aumentar la tolerancia a fallas se pueden desarrollar más algoritmos secundarios, y el switch entre algoritmos, puede depender de medidas de performance.

Así, queda conformada el protocolo de comunicación CANae como uno de los principales productos de este trabajo de tesis.

Resumen

En este capítulo se llevó a cabo un estudio de tres tipos de topologías tolerantes a fallas que fueron candidatas a ser utilizadas como base en la arquitectura que se diseñó en este trabajo de tesis.

Teniendo en cuenta los modelos que se han desarrollado en este capítulo se compararon las diferentes curvas de confiabilidad. El resultado de esto, permitió conocer de manera analítica qué topología presenta una mayor tolerancia a fallas a través del tiempo.

Se asumió que la distribución de la confiabilidad es exponencial, con una tasa de falla fija de $1/15$, y se estudió su evolución en un rango $[0, 1]t$. El resultado de esta comparación se la puede observar en la Figura 4.7 y en la Tabla 4.1.

Se observa que tanto las redes distribuidas como la hypercube presentan una mayor confiabilidad a través del tiempo que los árboles binarios. Si bien, las redes distribuidas y la hypercube tienen una curva similar, la primera presenta una mayor confiabilidad sostenida en el tiempo.

Así en este capítulo se llegó a la conclusión de que la topología que presenta un mayor grado de confiabilidad es la que responde a una filosofía distribuida. Por lo tanto, la arquitectura satelital, tolerante a fallas y basada en componentes COTS que se desarrolla en la presente tesis se basa en una **topología distribuida** para interconectar los diferentes subsistemas.

Arquitectura propuesta

Dadme un punto de apoyo y moveré el mundo

Arquímedes

La arquitectura propuesta se centra en la comunicación de la aviónica de un vehículo espacial, con la característica de que sus componentes (nodos) son de baja confiabilidad (se piensa en la utilización de componentes COTS) por lo tanto, es necesario que esta arquitectura sea tolerante a fallas. En este trabajo el protocolo de comunicación (tolerante a fallas) propuesto, se encuentra en las capas superiores del modelo de referencia OSI¹

Este capítulo está dividido en las siguientes secciones:

- *Arbol de requerimientos, Sección 5.1 (página 64).*
- *Casos de uso, Sección 5.2 (página 67).*
- *Diseño estructural (Diagramas de bloques, bloques internos), Sección 5.3 (página 69).*
- *Diseño dinámico (Diagramas de secuencia, interacción, máquina de estados), Sección 5.4 (página 73).*

5.1. Árbol de requerimientos

A continuación se detallan los requerimientos que van a guiar el diseño y desarrollo de la propuesta de arquitectura de aviónica tolerantes a fallas basada en componentes COTS para satélites. En la Figura 5.1 y en la Tabla 5.1 se pueden observar los requerimientos.

¹ISO/IEC 7498-1 - Open System Interconnection

Tabla 5.1: Tabla de Requerimientos

	ID	Name	Detail	Kind of Requirements
0	L1_001	L1_001	The architecture components shall be Commercial Off-The-Shelf category	Constraint
1	L2_000	L2_000	The architecture shall be reconfigure when a node fail.	Funcional
2	L2_001	L2_001	Each subsystem shall represented for a Components Off-The-Shelf	Constraint
3	L1_000	L1_000	Shall be develop an avionics architecture for spacecraft using Components Off-The-Shelf	Constraint
4	L1_002	L1_002	The architecture shall have at least 6 subsystems	Constraint
5	L1_003	L1_003	The architecture shall have fault tolerance techniques to assurance the mission life	No funcional
6	L1_004	L1_004	The main bus to interconnect components shall be the Controller Area Network Bus developed for Bosch	Constraint
7	L1_005	L1_005	The architecture shall be sufficient to make a master degree thesis	No funcional
8	L2_002	L2_002	The architecture shall implements a based Controller Area Network protocol to intercommunicate the components	Funcional
9	L3_000	L3_000	The intercommunication between nodes shall used a Controller Area Network protocol based developed in the master degree thesis.	Constraint
10	L1_006	L1_006	The architecture shall assurance at least 10 years mission life	No funcional
11	L3_001	L3_001	The architecture shall use the distributed network philosophy	Constraint
12	L1_007	L1_007	The thesis shall be finished in 1 year	Constraint
13	L2_003	L2_003	The nodes shall send and receive message from any nodes connected to the architecture.	Funcional

5.1 ÁRBOL DE REQUERIMIENTOS

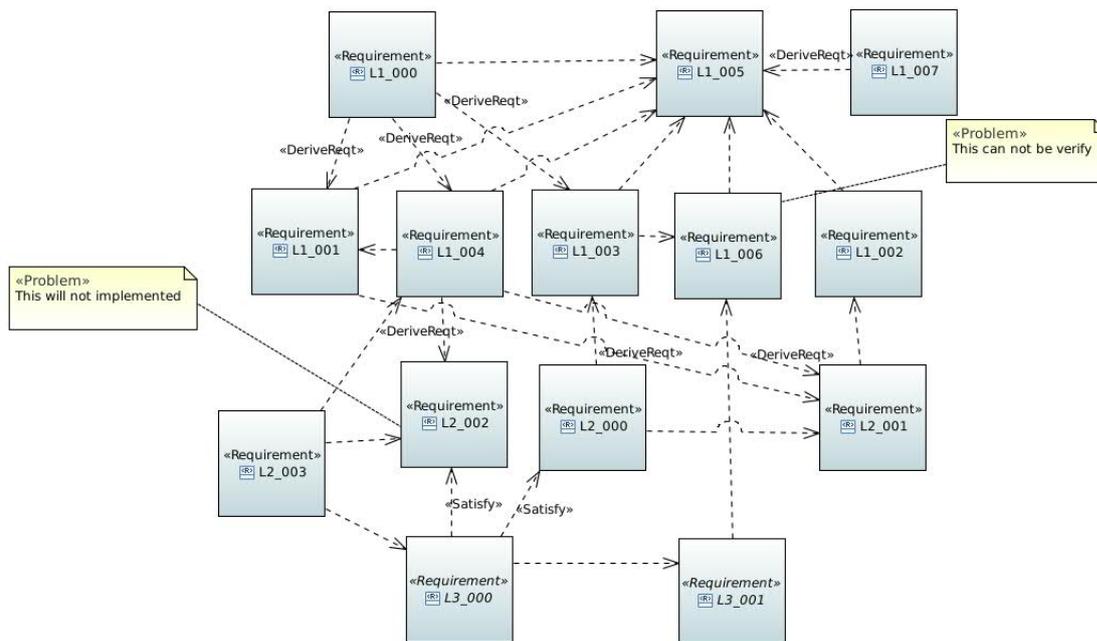


Figura 5.1: Diagrama de requerimientos de la arquitectura propuesta

A continuación se explica cada uno de los requerimientos:

- **L1_000:** Este es el objetivo principal de la presente tesis: desarrollar una arquitectura de aviónica tolerante a fallas basada en componentes COTS para vehículos espaciales.
- **L1_001:** El principal requerimiento de este trabajo de tesis, es desarrollar una arquitectura basada en componentes COTS, lo cual agrega un grado importante de innovación tecnológica y complejidad, siendo de especial interés para INVAP trabajar con estos tipos de componentes.
- **L1_002:** Este requerimiento hace referencia a que se debe asegurar el correcto funcionamiento de la arquitectura con una cantidad de al menos 6 subsistemas.
- **L1_003:** Este requerimiento exige que la arquitectura sea tolerante a fallas para así poder satisfacer el requerimiento **L1_006**.
- **L1_004:** Este requerimiento (constraint) indica que se debe pensar la arquitectura para ser utilizada con el bus CAN de Bosch. Este requerimiento es de interés para INVAP.
- **L1_005:** Este requerimiento indica que la arquitectura que se desarrolle debe ser suficiente para lograr terminar una tesis de maestría. Por el diseño y desarrollo debe ser solo el suficiente y necesario para cumplir con el presente trabajo.
- **L1_006:** Este requerimiento indica que debe asegurarse que la arquitectura sea lo suficientemente robusta como para asegurar el tiempo de vida de la misión de 10 años como mínimo. Existe un **problema** con este requerimiento, y es que para los alcances de esta tesis, será imposible verificar si se cumple con este requerimiento.
- **L1_007:** Este requerimiento indica que el trabajo de tesis tiene que ser finalizado en menos de un año. Esto tiene relación con el requerimiento **L1_005** debido a que no se alcanzará el detalle necesario para lograr la correcta implementación, verificación y validación de la arquitectura.

- **L2_000:** Este requerimiento exige que la arquitectura pueda reconfigurarse cuando un nodo falle.
- **L2_001:** Este requerimiento indica que para esta instancia de trabajo cada subsistema será tratado como un nodo dentro de la arquitectura.
- **L2_002:** Este requerimiento indica que se debe utilizar dentro de la arquitectura un protocolo de comunicación basado en el Bus CAN de Bosch.
- **L2_003:** Este requerimiento asegura que los nodos deben poder enviar y recibir mensajes de cualquier otro nodo conectado a la red CANae.
- **L3_000:** Este requerimiento indica que el protocolo de comunicación utilizado en esta arquitectura debe estar basado en el el protocolo CAN. Este protocolo fue desarrollado en el presente trabajo de tesis bajo el nombre de CANae 0.1 Alpha (Vease: Apéndice A).
- **L3_001:** Este requerimiento exige la utilización de una filosofía de red distribuída para su diseño y desarrollo.

5.2. Casos de Uso

A continuación se muestra el diagrama de Casos de Uso para la arquitectura propuesta. El diagrama de Casos de Uso tiene como propósito explicar de manera general el comportamiento a nivel de sistema de la arquitectura propuesta. Los actores identificados para esta arquitectura son los diferentes subsistemas que se conectan a la red a través de los nodos. Los subsistemas interactúan con el nodo a través de la aplicación de usuario, es decir, se encuentran asociados. Se puede observar en la Figura 5.2 que de forma general existen tres casos de usos que comprenden el comportamiento general de la arquitectura. Se tiene *Enviar mensaje* y *Recibir Mensaje* los cuales son básicos para cualquier arquitectura de comunicación. A esto se le agrega el caso de uso de *FDIR* que es el agregado que se le hace en este trabajo de tesis a través del protocolo de comunicación desarrollado denominado CANae (Ver apéndice A). Debe aclararse que en la Figura 5.2 aparece un resumen de los Actores interactuando con el sistema.

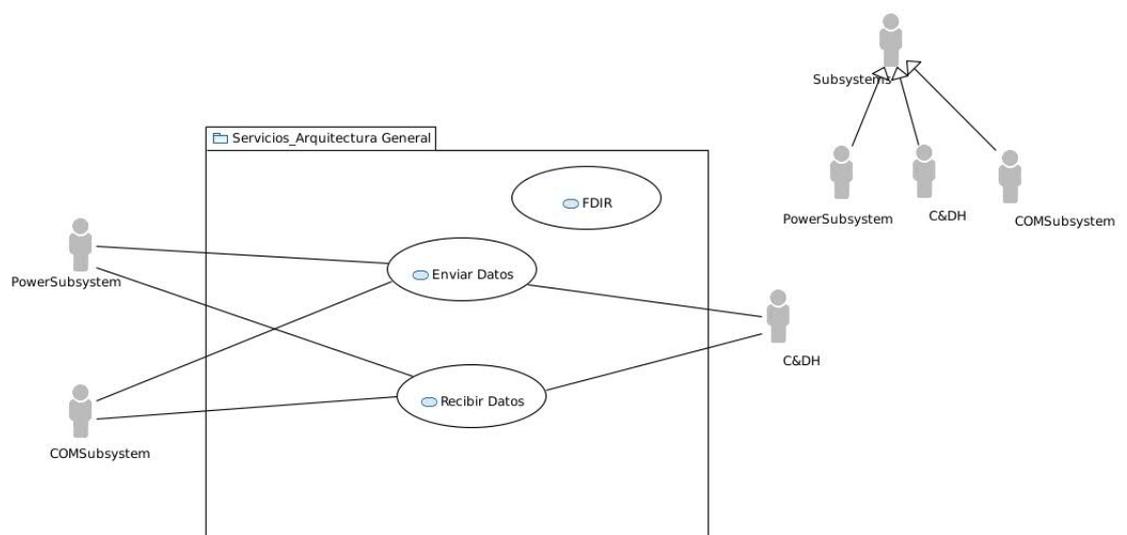


Figura 5.2: Diagrama de Casos de Uso General de la arquitectura Propuesta

De manera más detallada en la Figura 5.3 se observa un diagrama de casos de uso explayado que muestra su comportamiento.

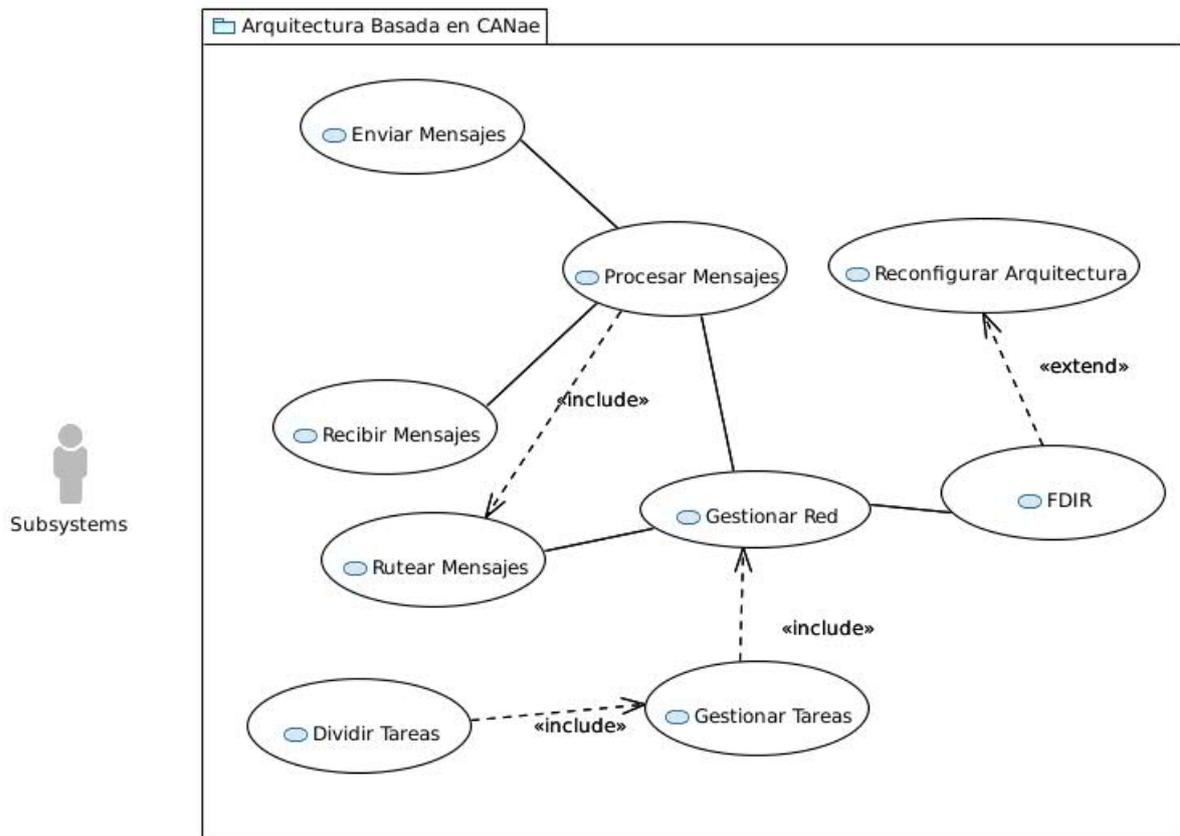


Figura 5.3: Diagrama de Casos de Uso de arquitectura propuesta

Le especificación de los casos de usos se describen en el apéndice B. Como se verifica en la Figura 5.3 el sistema tiene 9 casos de uso.

Los casos de uso de *Enviar Mensaje* y *Recibir Mensaje* denota la capacidad de la arquitectura de enviar y recibir mensajes (Data Frames de CANae) desde y hacia las aplicaciones de usuario. Cuando una aplicación necesita enviar un mensaje, estos deben, primero pasar por el caso de uso *Procesar Mensaje*, debido a que este es el encargado de empaquetar los mensajes. También en este punto se encuentra el caso de uso *Rutear Mensaje*.

El caso de uso *Rutear Mensaje* se encarga de aplicar los algoritmos necesarios para lograr entregar el mensaje de una manera eficiente, siguiendo la tabla de ruteo del CANae. Para esto se basa del caso de uso *Gestionar Red*.

El caso de uso *Gestionar Tareas* y *Dividir Tareas* se encarga de de conocer (periódicamente) el estado de la red y de sus nodos. Con esto, se puede decidir cómo se reparten las tareas de una manera eficiente, siguiendo la filosofía de arquitectura distribuída.

Por último tenemos el caso de uso *FDIR* que hace referencia a la Detección, Aislación, y Recuperación

del sistema ante fallas. Debido a la complejidad de esto, no se desarrolla este caso de uso en la presente tesis. El cual queda para **trabajos futuros** 6.2. En este caso de uso se encuentra el *Reconfigurar Arquitectura*. Este es un algoritmo de Detección, Aislación y Recuperación de Fallas (FDIR) que cuando se produce una falla en algunos de los nodos, la red se tiene que reconfigurar, para lograr continuar funcionando, aún cuando haya nodos inactivos. Esto le brindaría a la arquitectura un grado más de tolerancia a fallas.

5.3. Diseño Estructural

Este trabajo de tesis plantea una arquitectura a nivel de sistema. Esta arquitectura se encuentra basada en la red CAN debido a un interés especial de INVAP. Además se sigue una filosofía de red distribuida por las razones estudiadas en la sección 4.3. Esto le da la característica de innovador a este proyecto, ya que no se encontró en estado de la cuestión la utilización de esta filosofía en el desarrollo de vehículos espaciales, mucho menos en Argentina, que se sigue una filosofía más tradicional y conservadora a la hora de desarrollar satélites.

La arquitectura propuesta, como se mencionó anteriormente tiene como base el bus CAN. A la red se pueden conectar un número N de nodos ($N < 128$), siendo estos componentes COTS de baja confiabilidad. Esta arquitectura trabaja en las capas superiores de los protocolos de comunicación, por lo tanto, en esta instancia de trabajo, el nodo se conecta a la red mediante un microcontrolador CAN, el cual a su vez está compuesto por el CAN transceiver.

La red a bajo nivel, es decir desde el cable CAN, el CAN Transceiver y el microcontrolador CAN, deben trabajar bajo las normas preestablecidas por algún protocolo existente, que cuya selección de este, cae bajo la responsabilidad del diseñador y/o ingeniero de sistemas que vaya a implementar la arquitectura. Para esta arquitectura se aconseja el uso del protocolo CANOpen (CAN-CIA, 2017). Esto se puede observar en la Figura 5.4 que el Node (nodo de la red) tiene como *inout flowport* el microcontrolador CAN y el CAN Transceiver. El CAN Transceiver simplemente emite y recibe señales eléctricas. Las especificaciones (por ejemplo resistencias y voltaje, se muestran en el diagrama solo a modo ilustrativo) se pueden encontrar en los estándares (CAN-CIA, 2017).

Dentro de cada nodo se encuentra la aplicación de usuario que será de responsabilidad del desarrollador implementar esta aplicación haciendo uso del protocolo CAN desarrollado en este trabajo.

Para agregar tolerancia a fallas a la arquitectura se plantea la necesidad de llevar a cabo un Bridge tolerante a fallas. Este Bridge permite que la comunicación de la red con los subsistemas no se vean afectadas ante la falla de un nodo. La utilización de esta técnica se explica en la siguiente sección.

De esa manera queda definida la estructura a nivel de bloques de la arquitectura planteada.

5.3.1. Bridge tolerante a fallas

En esta sección se explica la necesidad de la utilización de un Bridge Tolerante a Fallas. No se pretende llevar a cabo una especificación técnica del Bridge, ya que escapan a los propósitos de este trabajo de tesis. En primer lugar, como la arquitectura persigue la filosofía de red distribuida, y la posibilidad de distribuir las tareas entre todos los nodos, surge la dificultad de que muchas veces no existe compatibilidad entre las computadoras de los diferentes subsistemas. Es decir, una computadora del subsistema de control térmico, por ejemplo, no puede prender y/o pagar componentes del subsistema de propulsión. Tal como se muestra en la Figura 5.5 la ejecución de tareas de un subsistema en otro está prohibido.

Esto lleva a una segunda propuesta en la cual el nodo conectado a la red distribuida, debe estar separado de la computadora que se comunica directamente con los componentes del subsistema. Con esto surge una

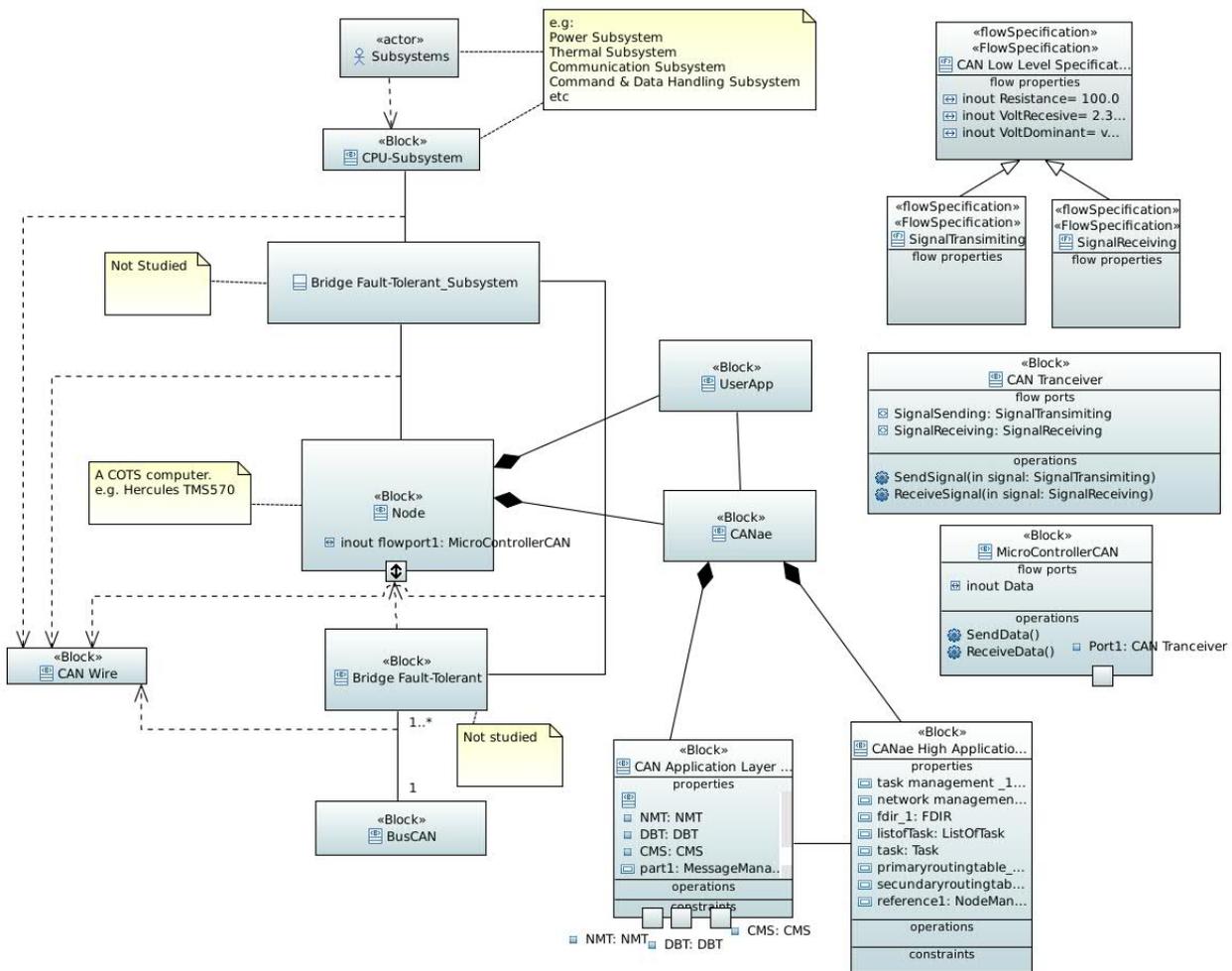


Figura 5.4: Diagrama de bloques de la arquitectura completa

segunda propuesta de desarrollo de la estructura de arquitectura. En esta, el subsistema se encuentra separado del nodo, con su propia CPU, y se conecta a la red mediante el nodo. Suponiendo que un nodo falle, debido a su baja confiabilidad (es un componente COTS), ese subsistema queda totalmente aislado de la red. En la Figura 5.6 se observa que, cuando falla el Nodo del *Subsistema 1* cualquier mensaje que envíe el *Subsistema 2* al primero, nunca se entregará. Esto provoca que la confiabilidad del sistema completo, disminuya linealmente con la confiabilidad del componente COTS.

La tercera opción y la ideal es la existencia de un Bridge, que se lo llama: *Bridge Tolerantes a Fallas*. Esto crea un puente entre la red y la computadora del subsistema permitiendo aumentar la tolerancia a fallas de la arquitectura. En caso de fallas en el nodo, el Bridge se activa ignorando al *nodo fallado* y conectando a la computadora del subsistema directamente en la red. Esto se puede observar en la Figura 5.7, siendo el Bridge el cuadrado negro.

Los detalles técnicos, electrónicos y de implementación no se desarrollan en este trabajo de tesis.

Como puede observarse en al Figura 5.4 el protocolo CANae se encuentra dentro del nodo. El desarrollador se debe encargar de implementar los protocolos, como un stack de servicios más del sistema operativo (o sistema embebido) que se encuentre en el nodo.

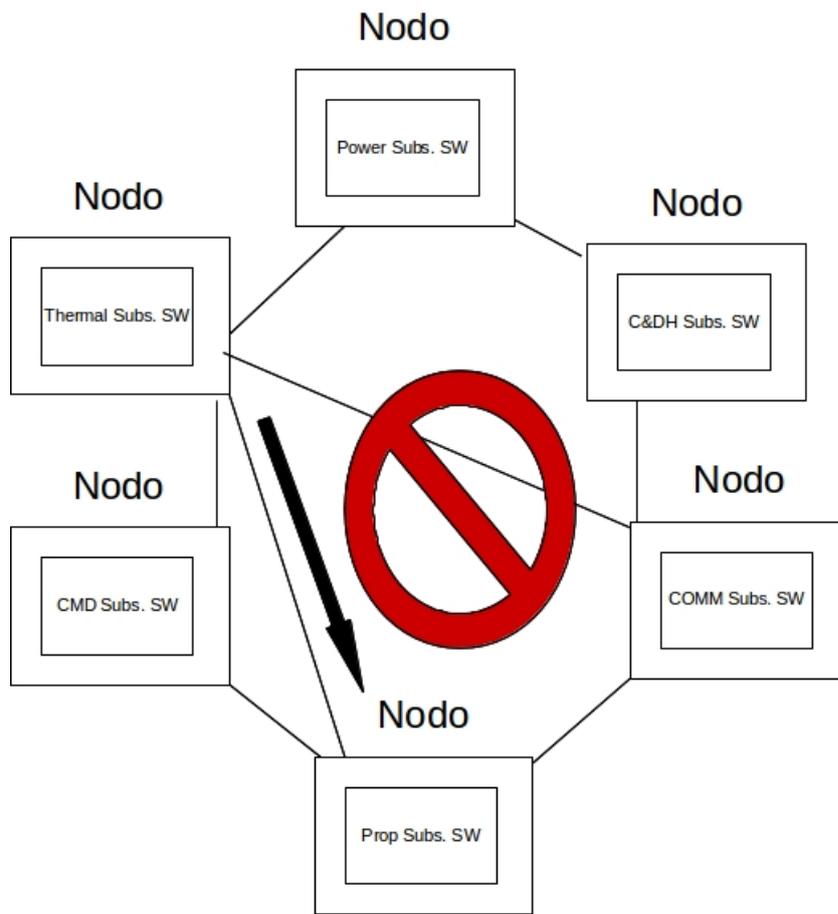


Figura 5.5: Posibilidad de interconexión de nodos. Configuración 1.

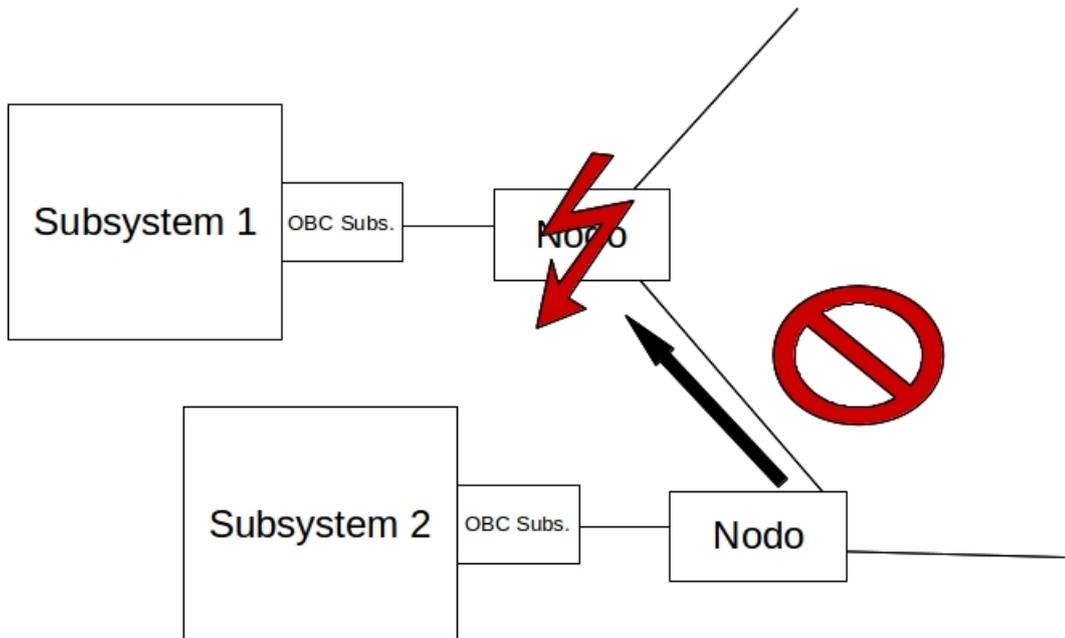


Figura 5.6: Posibilidad de interconexión de nodos. Configuración 2.

5.4. Modelo dinámico

En esta sección se modelará el comportamiento de la arquitectura. A diferencia de los modelos que se estudió en la sección anterior (diagramas de bloques, diagrama de bloques internos) que hacen referencia al *qué* del sistema, en el modelado del comportamiento se hace referencia al *cómo* del sistema. Este *cómo* es descrito por el modelado del comportamiento dentro del sistema (Holt y Perry, 2008).

En esta sección se verán 3 diferentes niveles de abstracción de los modelos de comportamiento del sistema. En primer lugar, el nivel más alto de abstracción es la máquina de estado, que considera la interacción entre bloques. El siguiente diagrama que se verá es el diagrama de secuencia, el cual pone en manifiesto los diferentes mensajes que son enviados entre las entidades. Y por último, se desarrolla el diagrama de actividades, siendo este el nivel más bajo de abstracción del modelado del comportamiento del sistema.

5.4.1. Máquina de estado

En esta subsección se describe el ciclo de vida de los bloques que conforman la arquitectura. En resumen, lo que se pretende demostrar, es el comportamiento de los objetos de la arquitectura.

En la Figura 5.9 se presenta la máquina de estado de la arquitectura completa. En esta se puede observar todos los estados por los cuales la arquitectura debe recorrer.

A continuación se procede a explicar la máquina de estado. Todo inicia en el momento en el que la red es alimentada (eléctricamente) siguiendo los protocolos de CAN (CAN-CIA, 2017). En primer lugar, la red pasa a un estado de inicialización de los nodos, en los cuales, estos son alimentados, se encienden, y comienzan con todos los procedimientos normales de cualquier sistema embebido. Los cuales son Booteo, chequeo del funcionamiento y del estado de salud del propio HW y los HW conectados a él. Esto incluye la iniciación, del protocolo CANae como un stack más del servicio del sistema operativo o sistema embebido que gobierna a cada uno de los nodos. Esto se observa en la Figura 5.10.

Una vez que todos los nodos se encuentran inicializados, y la red se encuentra estable², la arquitectura pasa a un estado de Stand-By. En este estado, la arquitectura espera por la orden del nodo monitor para comenzar a configurar la red. Aquí se sigue lo indicado en el protocolo CANae (Véase apéndice A). En primer lugar, se crean las instancias de red localmente. Luego, el nodo monitor coordina la creación de las instancias de red remota. Luego se crean los objetos nodos tanto locales como remotos. Estas actividades son necesarias para lograr establecer una correcta comunicación entre nodos utilizando el protocolo CANae (Versión 0.1 Alpha).

Una vez finalizada la creación de la red, la arquitectura completa pasa a un estado Ready, donde el nodo monitor ya puede ser retirado. Luego, la arquitectura ignora el nodo monitor (si todavía se encuentra conectado) y ya puede comenzar a trabajar normalmente.

5.4.2. Diagrama de secuencia

A continuación se muestra los diagramas de secuencias, para entender cuál es el comportamiento de la arquitectura, en un grado de abstracción menor.

En la Figura 5.11 se puede observar como en el momento de alimentar (eléctricamente) a los nodos, este comienza, en primer lugar, a ejecutar el bootstrap. Este bootea al sistema operativo (o sistema embebido según sea

²Defínase como estable, la situación en la que el protocolo CANae se inicializó correctamente y está listo para pasar a la segunda etapa.

la tecnología que se utilice). Una vez que el sistema operativo es cargado se deben ejecutar (preferentemente) dos tipos diferentes de chequeos:

- Hardware Check
- Software Check

Estos son dos piezas de código. El primero tiene como tarea principal chequear el estado de salud de los equipos de hardware, tanto del propio nodo, como así también del subsistema que tiene más próximo. Como se expuso en la subsección 5.3.1 cada nodo tiene próximo a él (físicamente hablando) un subsistema (con su computadora de vuelo). Por lo tanto el nodo se debe encargar de chequear el estado de salud de los diferentes componentes de ese subsistema. El segundo chequeo, se ocupa de verificar el estado de salud del software tanto del nodo, como del subsistema más próximo.

Luego de que los chequeos se hayan resuelto exitosamente se crea una instancia del protocolo CANae y da inicio a su servicio NMT. Este servicio se puede observar en la Figura 5.12. En primer lugar se debe esperar que el nodo monitor esté inicializado, y este le envía un mensaje de *create_remote_network()* a todos los nodos conectados a la arquitectura. Este mensaje llega al NMT del protocolo de CANae de todos los nodos.

Luego, siguiendo el protocolo de CANae, cada nodo crea su instancia de nodo. Esto puede observarse en la Figura 5.13. El objeto de *CANAppLayerController* envía el mensaje *create_node()* al *NodeManagement* de *create_newtork_object()* con el objetivo de crear una instancia local de la red.

Luego el nodo monitor envía el mensaje *prepare_remote_node()*. El protocolo CANae prevé los pasos que se muestran en la Figura 5.14

Por último, se conectan todo los nodos a la red, para lo cual deben actualizar el *Node Table Configuration*, y finalmente ya pueden comenzar a comunicarse. Esto se observan en las Figuras 5.15 y 5.16

5.4.3. Diagrama de actividades

En esta subsección se verán algunos diagramas de actividad para modelar el comportamiento de la arquitectura propuesta. Los diagramas de actividades comúnmente son utilizados, generalmente, para describir el comportamiento a muy bajo nivel (a nivel de componentes), en otras palabras para modelar procesos, mientras que los diagramas de secuencia muestran el comportamiento entre elementos, y las máquinas de estados muestran el comportamiento dentro de los elementos (Holt y Perry, 2008).

En la Figura 5.17 se puede observar el proceso, en detalle, del inicio de los nodos. Todo comienza cuando se alimenta (eléctricamente) el nodo. En primer lugar, se ejecuta (como ocurre normalmente) el bootstrap, este bootea al sistema operativo o al sistema embebido (dependiendo de como se implemente). Los siguientes pasos son los esperados para cualquier sistema espacial. Se chequea el HW propio del nodo, se chequea memoria y componentes. Luego, se chequea el estado de salud del subsistema. Para ello, se puede pedir el chequeo a la CPU del subsistema, y este a vez, chequea los componentes propios del subsistema.

Con estos pasos se puede obtener el Houskeeping (HK) tanto de HW propio del nodo como del subsistema. Este HK es recibido por el sistema operativo (o sistema embebido), y se comienza a hacer un chequeo del lado del software. Esto significa que se pueden realizar cualquier tipo de actividad para chequear que las tareas (o procesos) se ejecuten correctamente.

Para finalizar, el sistema operativo crea una instancia del stack de servicio de comunicación, dónde se encuentra el protocolo de comunicación CANae.

Debe observarse que el chequeo de HW tanto propio como el del subsistema se encuentra dentro de lo que se denomina *zona interrumpible*, esto significa que si se da alguna *Falla* (Failure en la Figura), debe activarse el sistema FDIR del nodo. Este, lo que busca en primer lugar, es detectar la falla, dónde se produjo y por qué. Luego, el sistema trata de aislar la falla para que no afecte a ningún otro componente. Por último, el sistema hará el intento de recuperarse de la falla, en el caso de que se recupere correctamente, se regresa la tarea de chequeo, para llevar a cabo estas actividades nuevamente; en caso contrario se termina la actividad.

Otra parte importante a modelar de la arquitectura propuesta, es el momento de la creación de la red CANae. Como se mencionó anteriormente el protocolo CANae (Véase apéndice A) requiere de la existencia de un nodo monitor para coordinar la creación de la red CANae. Este modelo se puede observar en la Figura 5.18.

El nodo monitor inicia la creación y configuración del nodo enviando la señal *create_remote_network()* a todos los nodos conectados a la red. Cada nodo que recibe la señal, lleva a cabo todas las actividades prevista por el protocolo CANae para la creación y configuración de la red. Crea las instancias de la red remota, los objetos red, crea la tabla de configuración de nodos y crea por último las instancias de los nodos.

Como segundo paso, el nodo monitor envía la señal de *Prepare Remote Node*, la cual es recibida por todos los nodos y se preparan internamente para comenzar a comunicarse.

Por último, el nodo monitor envía la señal para que los nodos se conecten a la red. Cada nodo se conecta, y actualiza su tabla de configuración del nodo. A partir de este momento los nodos ya pueden comenzar a enviar mensajes y el nodo monitor ya puede ser extraído de la red.

Además, se modela el “cómo” se comporta la arquitectura al momento de tratar de enviar un mensaje de un nodo a otro nodo. Esto se puede observar en la Figura 5.19. En primer lugar el nodo emisor, debe “preparar” el mensaje. Esto significa que debe empaquetarlo para, cumplir tanto con los protocolos CANae, como así también con el protocolo CAN.

Como segundo paso se debe consultar a la tabla de configuración del nodo para conocer cuál es el camino óptimo para llegar a entregar el mensaje. Para ello se le aplica algún algoritmo de ruteo.

Por último, se envía el mensaje a través del microcontrolador CAN. Este es recibido por el nodo destino y luego es procesado. Esto puede observarse, con mayor detalle en el apéndice A.

Resumen

En este capítulo se presentó la arquitectura propuesta, la cual sigue una filosofía de red distribuida que, como se estudio en capítulos anteriores, brinda una mayor tolerancia a fallas.

La innovación de este trabajo puede ser observada en el diseño de la arquitectura basado en modelos y orientado a objetos. Tanto la arquitectura, como su protocolo de comunicación (CANae), están basados en modelos, y utilizan el lenguaje de modelado SysML para su modelado y documentación.

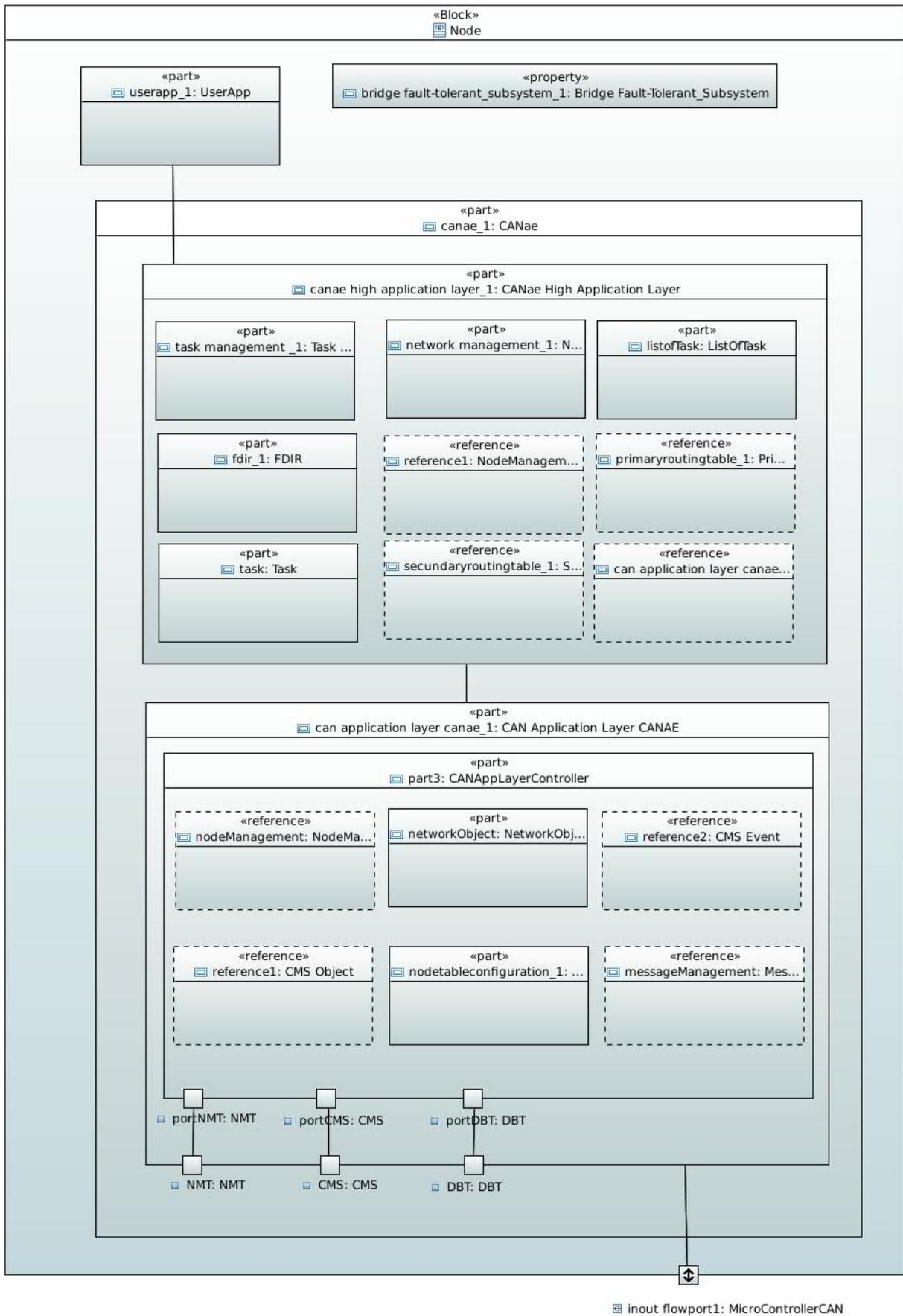


Figura 5.8: Diagrama de bloques interno del nodo.

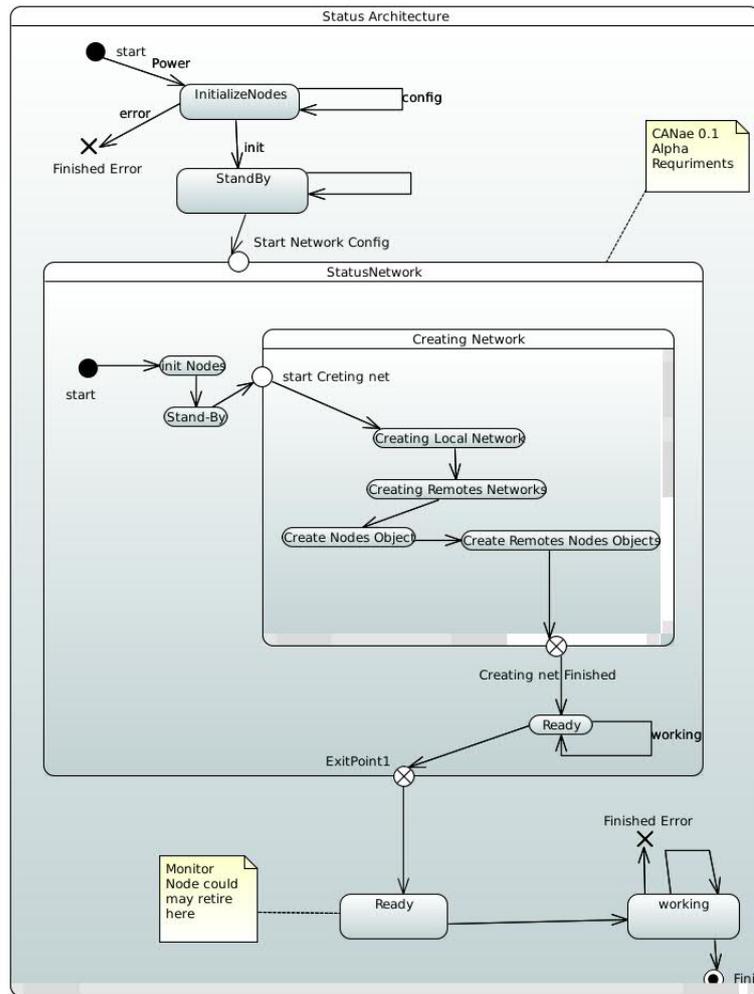


Figura 5.9: Máquina de estado de la arquitectura completa

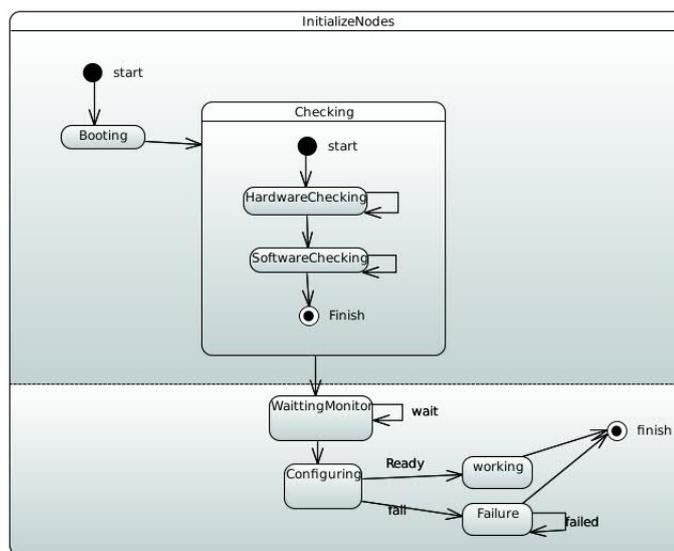


Figura 5.10: Máquina de estado de los nodos.

5.4 MODELO DINÁMICO

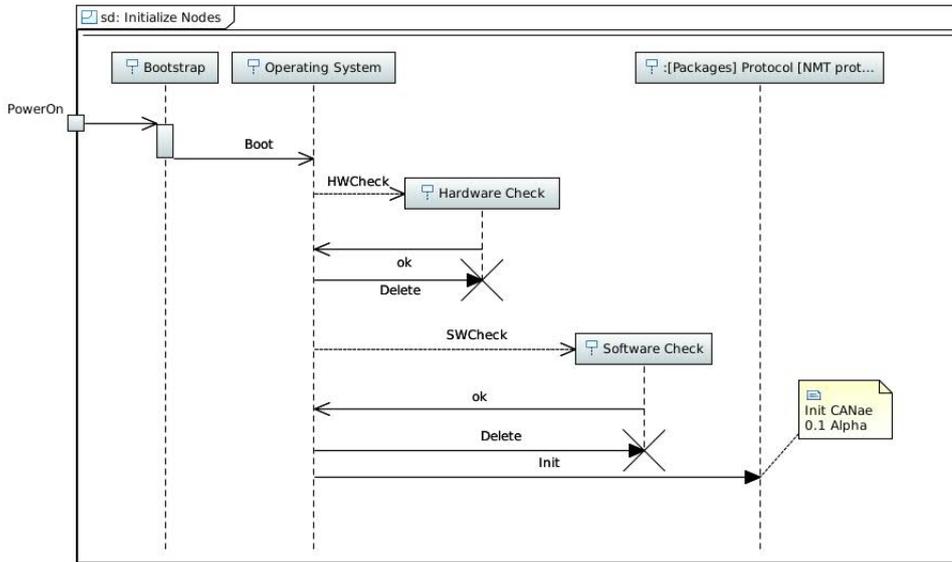


Figura 5.11: Diagrama de secuencia del inicio de la arquitectura.

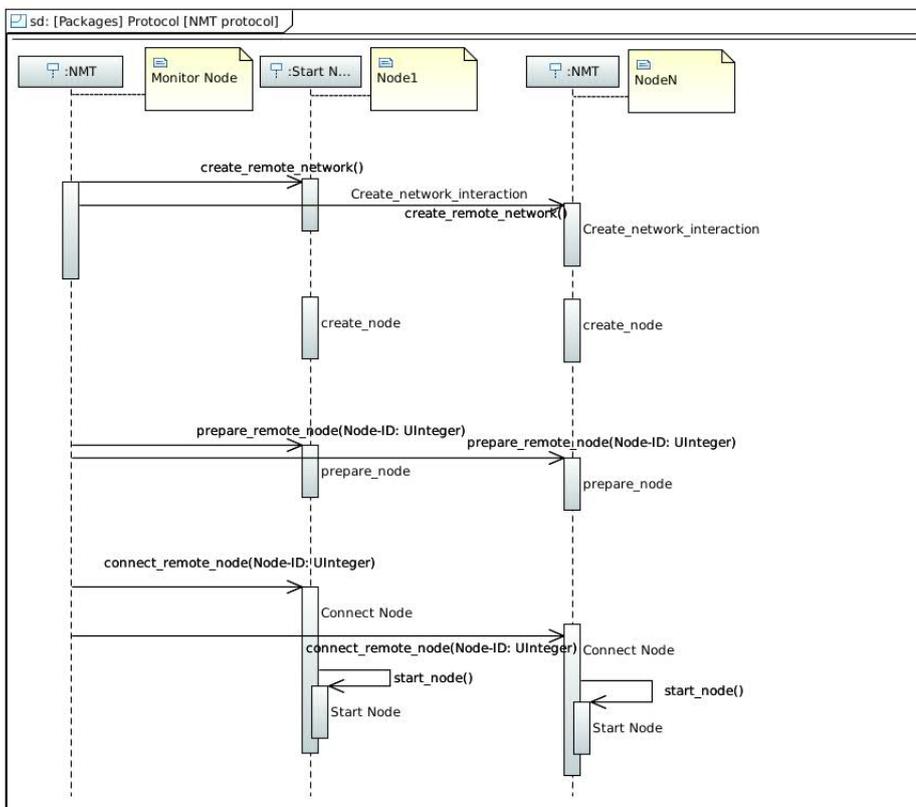


Figura 5.12: Diagrama de secuencia de CA

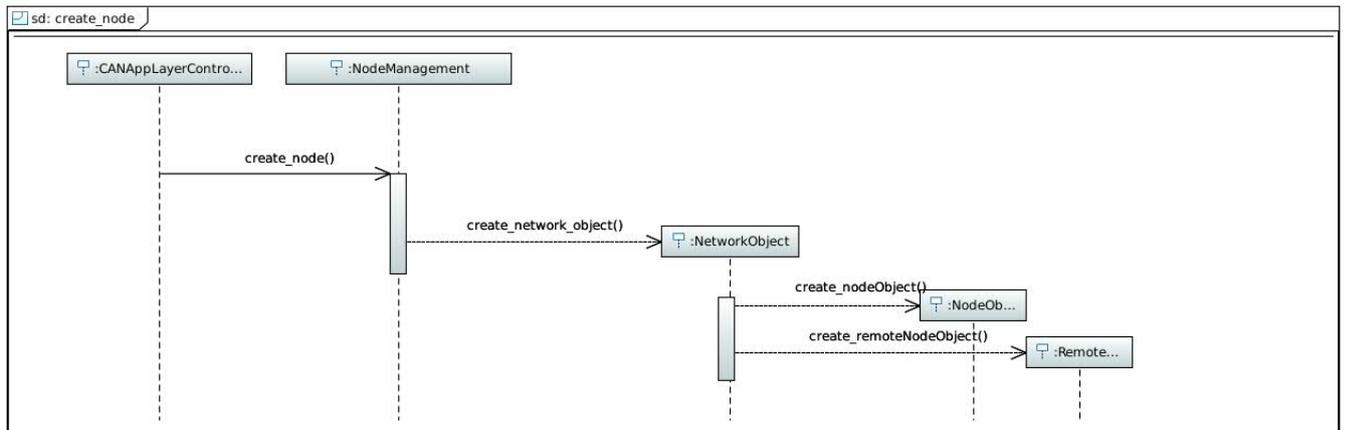


Figura 5.13: Crear Objeto Red, Objeto Nodo y Objeto Nodo Remoto

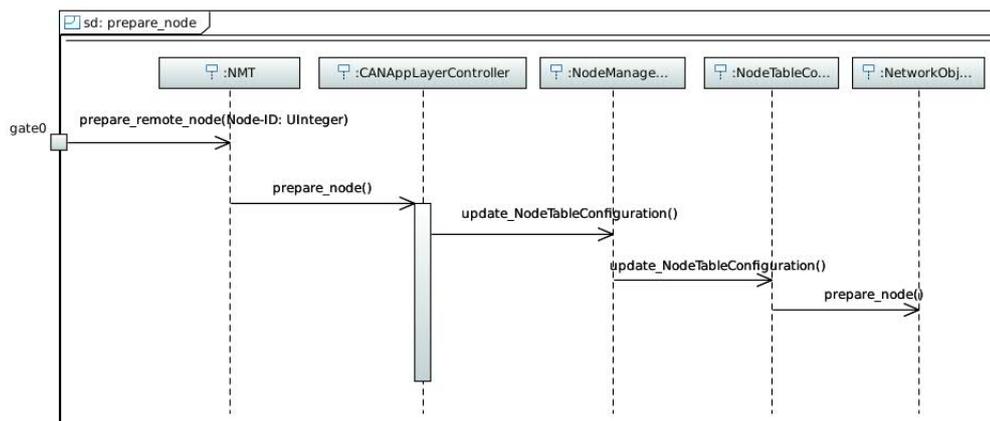


Figura 5.14: Preparar nodo para la conexión a la red CAN

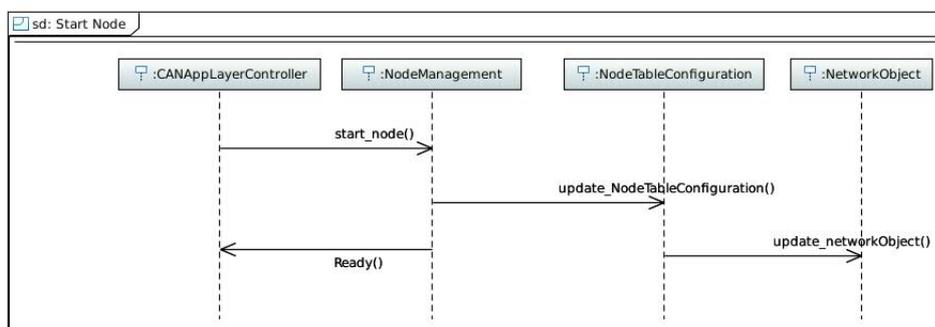


Figura 5.15: Preparar nodo para la conexión a la red CAN (Start Node)

5.4 MODELO DINÁMICO

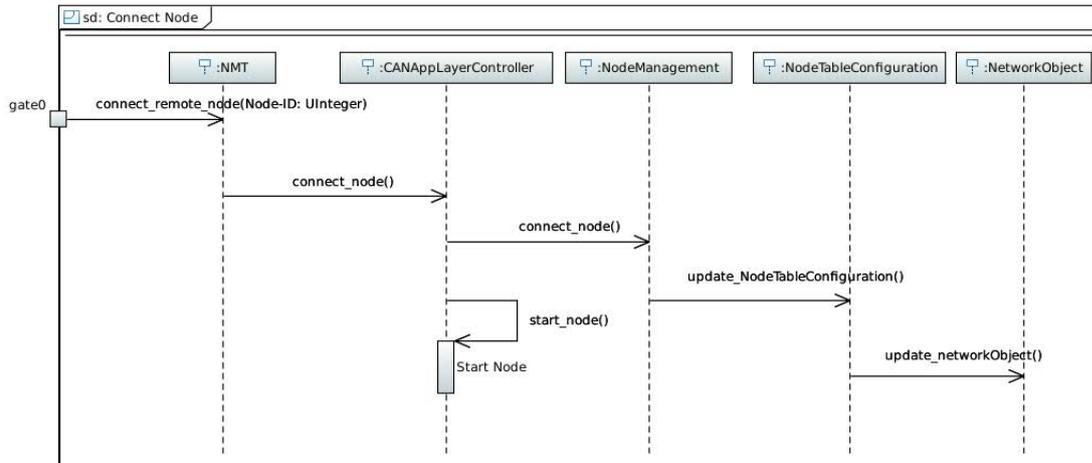


Figura 5.16: Conectar el nodo a la red CAN

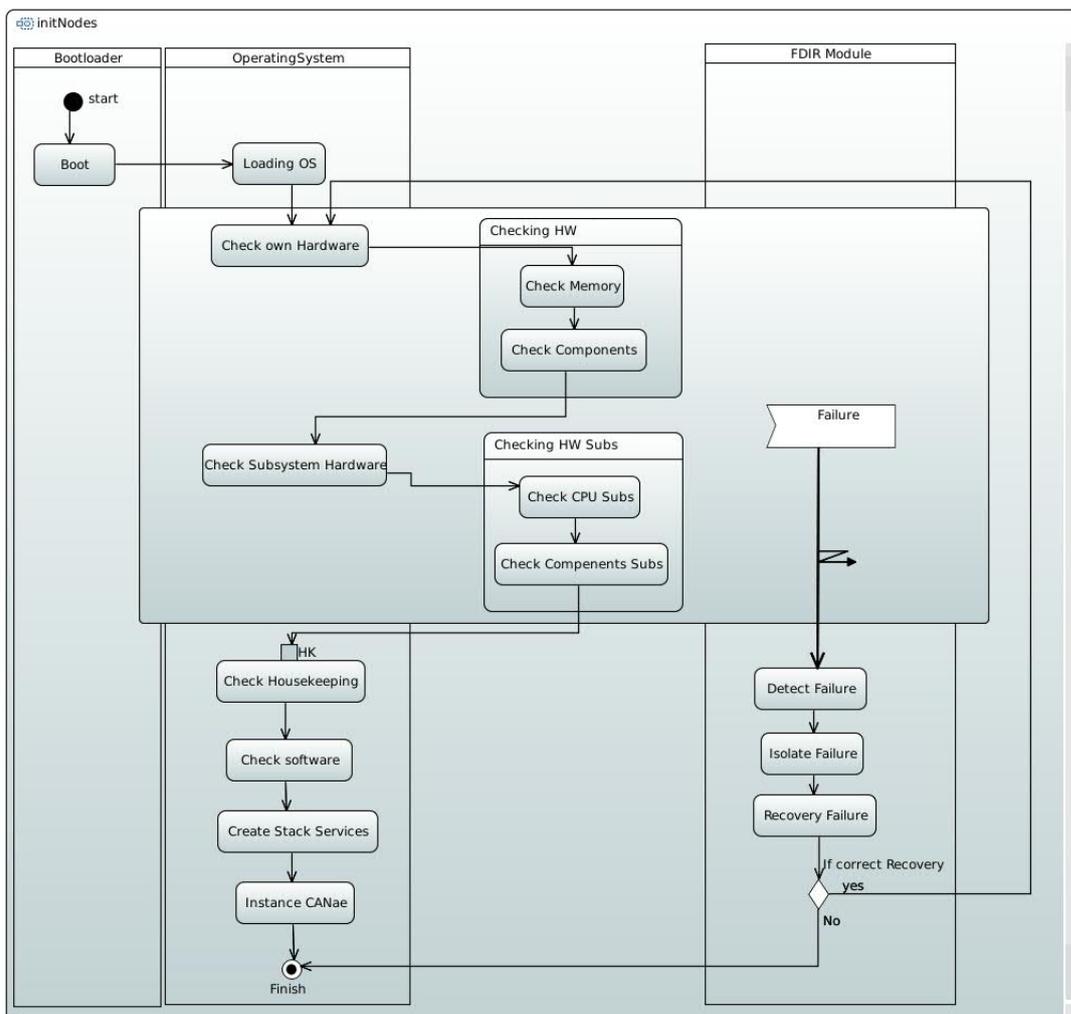


Figura 5.17: Diagrama de actividades del inicio de nodos

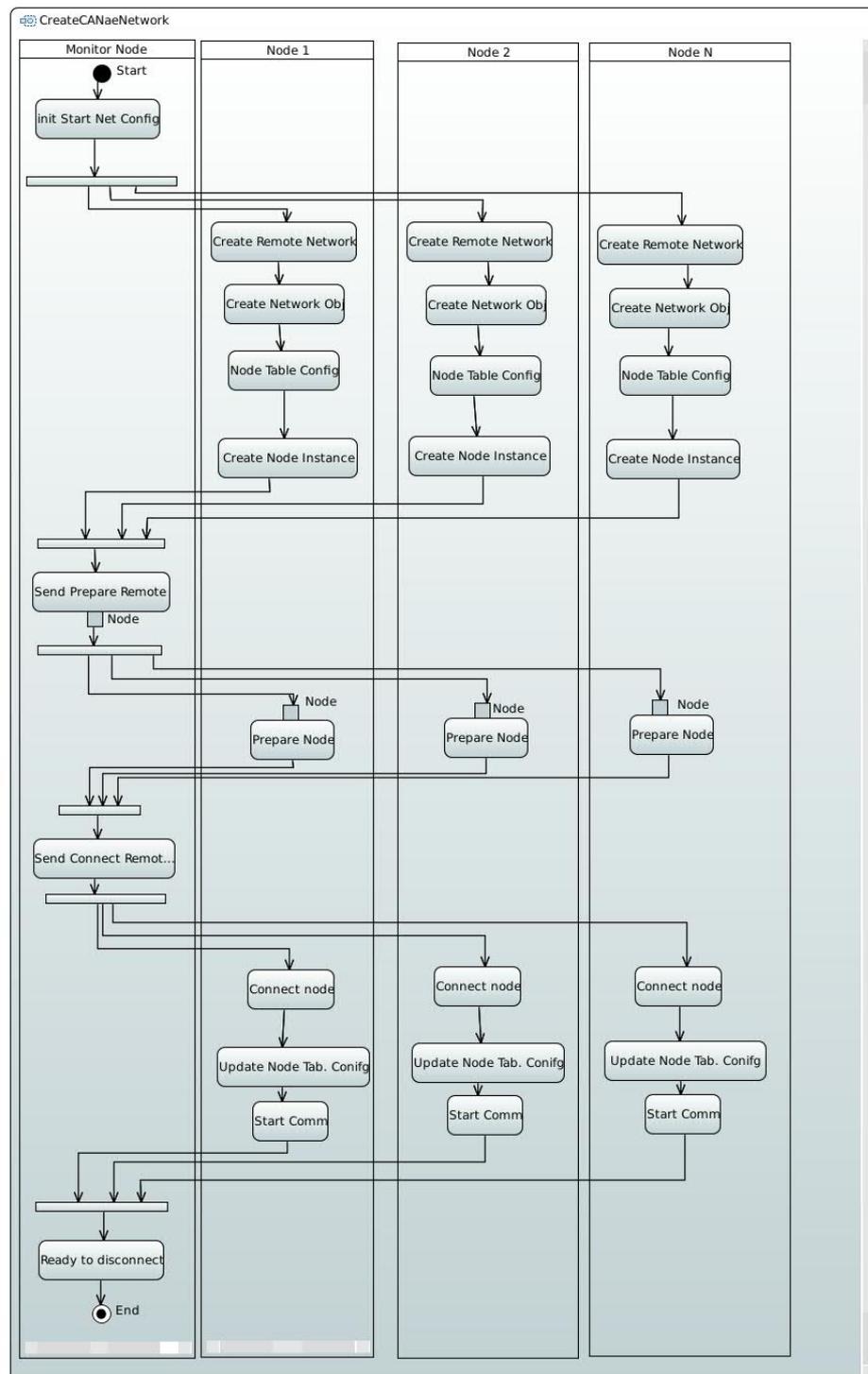


Figura 5.18: Diagrama de actividades del Nodo monitor

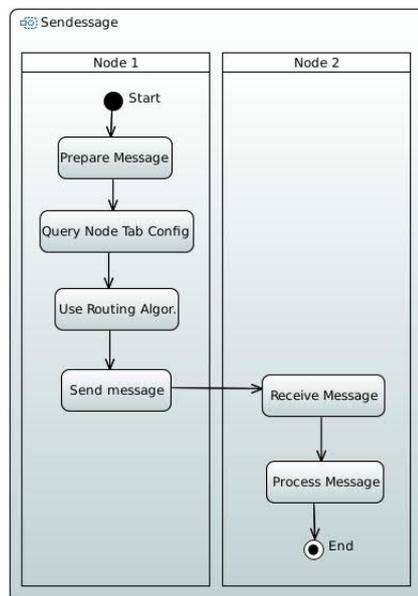


Figura 5.19: Diagrama de actividades del envío de mensajes

Conclusión

Las personas que están lo suficientemente locos como para pensar que pueden cambiar el mundo, son los que lo hacen.

Steve Jobs

No es un secreto que llevar a cabo una misión satelital representa grandes costos y largos períodos de trabajo de desarrollo de la misión. La NASA también padeció estos problemas hace algunas décadas. Pero, al ser una organización que suele ir en contra de las corrientes y filosofías normales de desarrollo, contar con personal altamente capacitado y disponer de recursos económicos, pudo afrontar los obstáculos e innovar. Así, es como surge su lema *Faster, Better and Cheaper* (FBC). El cual fue introducido por el administrador de la NASA Dan Goldin a principios de la década de los 90's. Dan Goldin tenía una vasta experiencia en el desarrollo de pequeños satélites, y estaba convencido de que el desarrollo "clásico" de misiones satelitales de NASA necesitaba un cambio de enfoque, con el objetivo de lograr reducir el tiempo de desarrollo, y pasar de escala de tiempo de décadas a años y, por sobre todo, lograr reducir costos (Paxton, 2007).

Es sabido que el lema FBC produce una gran cantidad de fallas y carencias en las misiones satelitales (Paxton, 2007). Sin embargo, esto es lo que impulsa la innovación, el desarrollo y entrenamiento de nuevos managers, ingenieros y científicos, los cuales serán los líderes de las nuevas generaciones (Paxton, 2007). Por tales motivos, es importante, que empresas y organizaciones del ámbito espacial en Argentina, dejen de lado el "clasicismo", y se animen, y por supuesto, alienten a las futuras generaciones, a transformar el tradicional *caminar con los manuales de NASA bajo el brazo*. Lo cual implica volcarse al camino de la innovación y el desarrollo de nuevas tecnologías y técnicas para la realización de satélites (que en este trabajo se denominan *vehículos de nueva generación*). Esta visión fue adoptada por empresas tales como INVAP y otras tanto nacionales como extranjeras, que están creciendo a paso agigantados llevando en su interior el cambio de filosofía que se exige, para lograr reducir costos, y alcanzar la capacidad necesaria para producir a gran escala.

El presente trabajo tiene como objetivo demostrar que es posible desarrollar vehículos espaciales utilizando componentes COTS de baja confiabilidad, trayendo consigo una reducción considerable del costo y del tiempo de desarrollo. De más está decir, que esta tesis no busca desarrollar en profundidad una arquitectura que alcance los niveles de detalles que son esperables en una misión satelital, debido a que para el cumplimiento de este objetivo se requiere de la colaboración interdisciplinaria de equipos de ingenieros, managers y científicos. Si no, que el presente trabajo pretender servir como base o puntapié inicial de líneas de investigación y desarrollo en el campo de estudio. Debido a que como se demostró, es factible el desarrollo de arquitecturas "económicas",

con la aplicación de técnicas y filosofías que van más allá de lo tradicional.

Luego de llevar a cabo un análisis de los sistemas satelitales y de aviónica que se han publicado en diferentes trabajos de investigación, y como resultado de diversas reuniones con varios equipos dentro de la Unidad de Desarrollo INVAP, se llegó a la siguiente conclusión: se debe promover el trabajo sobre el desarrollo de aviónicas aún cuando sus componentes sean de baja confiabilidad, la comunicación entre componentes, y que existe una tendencia a implementar redes de componentes dentro de los sistemas.

Una vez realizado este análisis previo, se estudiaron diferentes redes que tengan la característica de ser tolerante a fallas y que sean factibles de ser aplicadas en el ambiente espacial, como se indicó en la sección de Marco Teórico (Capítulo 2 Página 7) y en el Estado del Arte (Capítulo 3 Página 41). Las cuales fueron las siguientes:

- Árboles binarios
- Redes hypercube
- Redes distribuídas

Los resultados del análisis de la confiabilidad se encuentran detallada en la sección 4.3 (Página 55) y en Arias y Wiman (2017). Resumiendo, como se puede observar en la Figura 6.1, una vez realizada una comparativa entre redes tolerantes a fallas, la que presenta un mayor grado de confiabilidad sostenible en el tiempo es la que corresponde a la red distribuida. Este resultado representa un importante aporte para el conocimiento de tipos de topologías confiables tolerantes a fallas, y el desarrollo de modelos para lograr medir la confiabilidad de esas topologías. Debido a que permitirá reemplazar la estructura tradicional de un simple bus de comunicación, centralizada en una On Board Computer (OBC), comúnmente la computadora del subsistema de CDH (Command and Data Handling), por una estructura distribuida, dónde no existen rangos de importancia entre componentes, sino que todas las computadoras tiene la capacidad suficiente, y deben ejecutar las tareas de todos los subsistemas de manera distribuida.

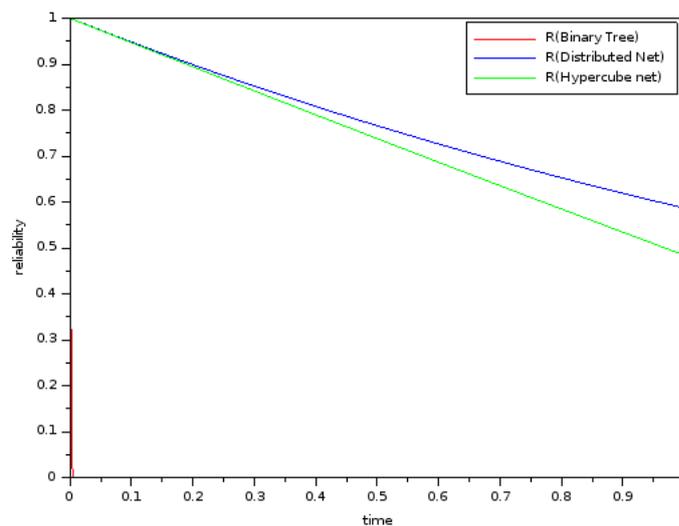


Figura 6.1: Comparación de confiabilidad

Un requerimiento que fue planteado por la Unidad de Desarrollo INVAP fue trabajar con el bus de comunicación CAN. Este es un protocolo de comunicación de bajo nivel, que demuestra ser altamente confiable por

lo que es utilizado en la industria automotriz, que fue la promotora de su creación. Esto, nos da una clara razón de que CAN es un protocolo confiable, debido a que en la fabricación de automóviles se tiene que cumplir el objetivo de que el producto final no ponga en riesgo la vida de personas. Además, por su confiabilidad, y a su excelente respuesta en sistemas de tiempo real, CAN es utilizado también en el ambiente espacial (entre otras industrias), pero relegado a ser un bus secundario.

En base a estos resultados y restricciones se plantea como un desafío a futuro implementar, crear, diseñar, una arquitectura que sea capaz de trabajar con una filosofía de red distribuida, la cual se encuentra descrita en la sección 5 (Página 64). El grado de innovación que coloca esta nueva arquitectura se ve reflejada en que no existe una computadora única (la connotación “única” hace referencia a la computadora propiamente dicha más sus redundancias, o bien, cualquier (micro)controlador más sus redundancias) por cada subsistema como se trabaja “tradicionalmente” sino que, existen computadoras híbridas capaces de llevar a cabo tareas de todos los sistemas en forma distribuida. Para ello se planteó un protocolo basado en CAN y CANOpen. Este protocolo, hasta el momento, se encuentra en su versión 0.1 Alpha. El protocolo presenta la innovación de estar orientado a objetos, y para ello se utilizó el lenguaje SysML (System Modeling Language) para llevar a cabo su modelado y la documentación del mismo. Esto facilita la interpretación del stack de servicio que debe ser implementado a la hora del desarrollo. Este protocolo se encuentra documentado en el apéndice A (Página 91).

En base a los resultados obtenidos y a la arquitectura propuesta basada en CANae (apéndice A) se concluye de que es factible desarrollar una *arquitectura de aviónica tolerantes a fallas basada en componentes COTS para vehículos satelitales* utilizando el bus CAN, como bus principal de comunicación. Si bien esto exige un cambio de paradigma importante, los datos recolectados en el presente trabajo demuestran que es posible diseñar arquitecturas fiables basadas en componentes que sean baratos y de baja confiabilidad.

6.1. Conclusiones finales

En el presente trabajo de tesis se demostró que es factible la utilización de componentes COTS en sistemas espaciales. Si bien, no se pudo responder si es posible la realización de un método de medición del grado de tolerancia a fallas de una arquitectura de aviónica, se logró desarrollar una manera de comparar confiabilidad de topologías de arquitectura, basándose en la tolerancia a fallas.

En este trabajo se demostró que la estrategia más indicada de tolerancia a fallas que permite brindar un alto grado de confiabilidad en la utilización de componentes COTS en sistemas críticos, es desarrollar una arquitectura basada en una red distribuida, debido a que de esta manera se logra distribuir las tareas entre todos las computadoras que conforman la arquitectura, y ante cualquier eventualidad (o falla), la red se puede reconfigurar sencillamente, y al no estar ligada una computadora a un subsistema (enfoque tradicional), las tareas se redistribuyen dinámicamente sin perjudicar el funcionamiento normal del sistema, en otras palabras aumenta su tolerancia a fallas.

La arquitectura basada en CANae, demostró ser la arquitectura más indicada que permita desarrollar tolerancia a fallas en sistemas críticos basados en componentes COTS, debido a que cumple con los requisitos para sentar las bases de trabajos futuros en la implementación de arquitecturas confiables basadas en componentes COTS.

La principal ventaja que brinda CANae, es su stack de servicio. Este permite su aplicación en redes distribuidas basadas en el protocolo CAN. CANae otorga las bases para la implementación de técnicas que logra la coordinación de nodos en una red distribuida, y es el mismo protocolo que contempla la posibilidad de que, al momento de la detección de una falla, este puede detectar el nodo con problema, y aislarlo; así de esta manera evitar una avería que podría tener como efecto negativo, y como consecuencia, la pérdida de la misión espacial.

Otra mejora que impone CANae es la utilización de tablas ruteos para ser utilizados por algoritmos de

ruteos de tolerancia a fallas. Esto agrega dinamismo a la red, manteniendo una tolerancia a falla aceptable para ser utilizada en misiones espaciales.

6.2. Perspectivas a futuros

A partir de análisis realizado en este trabajo de tesis, se logra ver el gran potencial que existe en el estudio de arquitecturas de aviónicas tolerantes a fallas, y de los componentes COTS introducidos en el sistema, donde el producto esperado debe ser más confiable que los componentes que la conforman.

Por estos motivos, a futuro se pretende continuar con las siguientes líneas de trabajo:

- Diseño detallado, desarrollo e implementación del protocolo CANae.
- Diseño detallado de la arquitectura propuesta.
- Estudio de nuevas técnicas de tolerancia a fallas aplicadas a los diferentes niveles de detalle de las arquitecturas de aviónica.
- Desarrollo de algoritmos de ruteo para la distribución de tareas en redes distribuidas.
- Estudio de tecnología Wireless como medio de comunicación alternativo al cableado.

Bibliografía

- Abd-El-Barr, M. y Gebali, F. (2014). Reliability analysis and fault tolerance for hypercube multi-computer networks. *Information Sciences*, 276:295 – 318.
- Alena, R., Gilstrap, R., Baldwin, J., Stone, T., y Wilson, P. (2011). Fault tolerance in ZigBee wireless sensor networks. In *Aerospace Conference, 2011 IEEE*, pages 1–15.
- Anderson, T. y Knight, J. C. (1983). A Framework for Software Fault Tolerance in Real-Time Systems. *IEEE Trans. Software Eng.*, (3):355–364.
- Arias, E. y Wiman, G. (2017). Estudio de la confiabilidad de arquitecturas tolerantes a fallas basada en componentes cots para aviónicas de vehículos espaciales. Asociación Argentina de Tecnología Espacial - Instituto Universitario Aeronautico.
- Becker, K. y Voss, S. (2016). *A Formal Model and Analysis of Feature Degradation in Fault-Tolerant Systems*. Springer International Publishing.
- Cambridge University (2017). <https://dictionary.cambridge.org/>.
- CAN-CIA (2017). <https://www.can-cia.org>.
- Chau, S. N., Alkalai, L., Tai, A. T., y Burt, J. B. (1999). Design of a fault-tolerant COTS-based bus architecture. *IEEE Transactions on Reliability*, 48(4):351–359.
- Chau, Savio N. and Smith, Joseph and Tai, Ann T. (2001). A design-diversity based fault-tolerant cots avionics bus network. In *Dependable Computing, 2001. Proceedings. 2001 Pacific Rim International Symposium on*, pages 35–42.
- Corrigan, S. (2002). Introduction to the Controller Area Network (CAN). Technical Report SLOA101A, Texas Instruments.
- Corrigan, S. (2008). Controller Area Network Physical Layer Requirements - ICP, Industrial Interface. Technical Report SLLA270, Texas Instruments.
- Crawley, E., de Weck, O., Eppinger, S., Magee, C., Moses, J., Seering, W., Schindall, J., Wallace, D., y Whitney, D. (2004). Engineering Systems Monograph, The influence of architecture in engineering systems.
- Douglas Isbell and Don Savage (1999). Mars Climate Orbiter Failure Report - NASA.
- Dubrova, E. (2013). *Fault-Tolerant Design: an introduce*. Springer-Verlag New York, New York, U.S.A., 1 edition.

- Ducard, G. J. J. (2007). *Fault-Tolerant Flight Control and Guidance Systems for a Small Unmanned Aerial Vehicle*. PhD thesis, Swiss Federal Institute of Technology Zurich.
- Edwards, C., Lobaerts, T., y Smaili, H. (2010). *Fault Tolerant Flight Control. A benchmark challenge*. Springer-Verlag.
- Eickhoff, J. (2012). *Onboard Computers, Onboard Software and Satellite Operations*. Springer-Verlag Berlin Heidelberg, Alemania.
- ESD Electronics (2017). <http://www.esd-electronics-usa.com/Controller-Area-Network-CAN-Intro>
- Esposito, S., Albanese, C., Alderighi, M., Casini, F., L., G., Esposti, M. L., Monteleone, C., y Violante, M. (2015). Cost-based high-performance computing for space applications. *IEEE Transactions on Nuclear Science*, 62(6):2687–2694.
- Esteve, M.-A., Katoen, J.-P., Nguyen, V., Postma, B., y Yuste, Y. (2012). Formal correctness, safety, dependability, and performance analysis of a satellite. pages 1022–1031.
- Eterno, J. S., Weiss, J. L., Looze, D. P., y Willsky, A. (1985). Design issues for fault tolerant-restructurable aircraft control. In *Decision and Control, 1985 24th IEEE Conference on*, pages 900–905.
- Forsberg, K. y Harold, M. (1999). 4 System Engineering for Faster, Cheaper, Better. *INCOSE International Symposium*, 9(1):924–932.
- Fortescue, P., Stark, J., y Swinerd, G. (2003). *Space Systems Engineering*. Wiley, West Sussex, England, 3 edition.
- Fortescue, P., Stark, J., y Swinerd, G. (2016). *Space Systems Engineering*. Springer International Publishing.
- Friedenthal, S., Moore, A., y Steiner, R. (2008). *A Practical Guide to SysML: Systems Modeling Language*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- Gangloff, W. C. (1975). Common mode failure analysis. *IEEE Transactions on Power Apparatus and Systems*, 94(1):27–30.
- Hamilton, Deirdre L. and Walker, Ian D. and Bennett, John K. (1999). Fault tolerance versus performance metrics for robot systems. *Reliability Engineering & System Safety*.
- Hanmer, R. S. (2007). *Patterns for Fault Tolerant Software*. John Wiley and Sons Ltd, England.
- Harlan, D. M. y Lorenz, R. D. L. (2005). *Space Systems Failures: Disasters and Rescues of Satellites, Rockets and Space Probes*. Springer Praxis Books. Praxis, 1 edition.
- Hayes, J. P. (1976). A graph model for fault-tolerant computer systems. *IEEE Transactions on Computers*.
- Hess, Ronald and Vetter, T. K. and Wells, S. R. (2005). Design and evaluation of a damage-tolerant flight control system. In *Institution of Mechanical Engineers Part G Journal of Aerospace Engineering*.
- Hitt, E. F. y Mulcare, D. (2001). Fault-Tolerant Avionics. In Spitzer, C. R., editor, *Avionics Development and Implementation Second Edition*, chapter 8. CRC Press, Williamsburg, Virginia, U.S.A.
- Holt, J. y Perry, S. (2008). *SysML for systems engineering*. The Institution of Engineering and Technology, London, United Kingdom.
- Hoque, K., Mohamed, O., y Savaria, Y. (2015). Towards an accurate reliability, availability and maintainability analysis approach for satellite systems based on probabilistic model checking. volume 2015-April, pages 1635–1640.

- IEEE (1990). IEEE Standard Glossary of Software Engineering Terminology. Std. 610.12-1990, IEEE.
- INVAP Sociedad del Estado (2016). <http://www.invap.com.ar/es/>.
- Izadi, B. A. y Özgüner, F. (2004). An augmented k-ary tree multiprocessor with real-time fault-tolerant capability. *The Journal of Supercomputing*, 27(1):5–17.
- Krodel, J. y Romanski, G. (2008). Handbook for real-time operating systems integration and component integration considerations in integrated modular avionics systems. Techn. Rep. DOT/FAA/AR-07/48, U.S. Department of Transportation.
- Kvaser (2017). <https://www.kvaser.com/can-protocol-tutorial/>.
- Leu, Y.-R. y Kuo, S.-Y. (1996). A fault-tolerant tree communication scheme for hypercube systems. *IEEE Transactions on Computers*, 45(6):641–650.
- Lisner, J. C. (2007). *A Fault-tolerant Dynamic Time-triggered Protocol*. PhD thesis, Dependability of Computing Systems - University of Duisburg-Essen.
- Liu, Y., Li, D., y Guo, C. (2014). Software reliability modeling with fault detection data when knowing fault severity. pages 558–562.
- Loveless, A. T. (2015). chapter On TTEthernet for Integrated Fault-Tolerant Spacecraft Network. AIAA SPACE Forum. American Institute of Aeronautics and Astronautics. 0.
- Lyu, M. R. (1995). *Software Fault Tolerance*. John Wiley and Sons Ltd, Chichester, New York, USA.
- M., P. (2006). Dmt and dt2: two fault-tolerant architectures developed by cnes for cots-based spacecraft supercomputers. In *12th IEEE International On-Line Testing Symposium (IOLTS'06)*.
- MARTE (2016). <http://www.omgmarte.org/>.
- Mayo, J. R., Armstrong, R. C., y Hulette, G. C. (2016). *Leveraging Abstraction to Establish Out-of-Nominal Safety Properties*. Springer International Publishing.
- Nelson, V. P. (1990). Fault-tolerant computing: fundamental concepts. *Computer*, 23(7):19–25.
- Papyrus (3 de Junio de 2016). <http://eclipse.org/papyrus/>.
- Paxton, L. J. (2007). “faster, better, and cheaper” at nasa: Lessons learned in managing and accepting risk. *Acta Astronautica*, 61(10):954 – 963.
- Peng, Z., Lu, Y., Miller, A., Johnson, C., y Zhao, T. (2013). A probabilistic model checking approach to analysing reliability, availability, and maintainability of a single satellite system. pages 611–616.
- Pradhan, D. K. y Reddy, S. M. (1982). A Fault-Tolerant Communication Architecture for Distributed Systems. *IEEE Transactions on Computers*, C-31(9):863–870.
- Pressman, R. S. (2001). *Software Engineering, A Practioner's Approach*. McGraw-Hill, fifth edition edition.
- Pullum, L. (2001). *Software fault tolerance techniques and implementation*. Artech House, England.
- Raghavendra, C. S., A., A., y Ercegovac, M. D. (1984). Fault tolerance in binary tree architectures. *IEEE Transactions on Computers*, (6):568–572.
- Rani, S. (2011). *Software and Hardware Reliability of Fault Tolerant Systems*. PhD thesis, Shobit Institute of Engineering Technology.

- Rausand, M. y Hoyland, A. (2004). *System Reliability Theory. Models, statistical methods, and applications*. John Wiley Sons, Inc., Hoboken, New Jersey, USA, 2 edition.
- Real Academia Española (2017). <http://www.rae.es/>.
- RTI (2015). *RTI Purchase Order Terms and Conditions v1.12. Spanish Translation*. RTI International, 3040 East Cornwallis Road, Research Triangle Park, USA.
- Schneidewind, N. F. (1997). Reliability modeling for safety-critical software. *IEEE Transactions on Reliability*, 46(1):88–98.
- Schneidewind, N. F. y Nikora, A. P. (1998). Issues and methods for assessing cots reliability, maintainability, and availability.
- Singh, A. D. y Youn, H. Y. (1991). A modular fault-tolerant binary tree architecture with short links. *IEEE Transactions on Computers*.
- Steiner, Willfried (2013). An Introduction to TTEthernet. TTTech Computertechnik AG.
- Stivaros, C. (1992). A measure of fault-tolerance for distributed networks. In *Computing and Information, 1992. Proceedings. ICCI '92., Fourth International Conference on*, pages 426–429.
- Stock Flight Systems - CANAerospace (2017). <http://www.stockflightsystems.com/canaerospace.html>.
- Stone, T., Alena, R., Baldwin, J., y Wilson, P. (2012). A viable cots based wireless architecture for spacecraft avionics. In *Aerospace Conference, 2012 IEEE*, pages 1–11.
- SysML (2 de Junio de 2016). <http://sysml.org/>.
- Tai, A. T., Chau, S. N., y Alkalai, L. (1999). Cots-based fault tolerance in deep space: Qualitative and quantitative analyses of a bus network architecture. In *High-Assurance Systems Engineering, 1999. Proceedings. 4th IEEE International Symposium on*, pages 97–104.
- Tanenbaum, A. (2003). *Redes de computadoras*. Editorial Alhambra S. A. (SP).
- Tanenbaum, A. S. y Wetherall, D. J. (2012). *Redes de computadora*. Pearson Educación, Mexico, quinta edition.
- Torres-Pomales, W. (2000). Software Fault Tolerance: A tutorial. Technical Report NASA/TM-2000-210616, National Aeronautics and Space Administration Langley Research Center, Hampton, Virginia.
- TTTech. Technical report, TTTech Computertechnik AG, Vienna, Austria.
- TTTech (20 de Enero del 2017). <http://tttech.com/>.
- Watkins, C. B. y Walter, R. (2007). Transitioning from federated avionics architectures to integrated modular avionics. In *2007 IEEE/AIAA 26th Digital Avionics Systems Conference*.
- Woerner, D., Deutsch, L., y Salvo, C. (2000). The X2000 Program: An Institutional Approach to Enabling Smaller Spacecraft.
- Xie, M. (1991). *Software Reliability Modeling*. World Scientific Publishing Co. Pte. Ltd, Singapore, 1 edition.
- Zhang, C., Jaimoukha, I. M., y Sevilla, F. R. S. (2016). Fault-tolerant observer design with a tolerance measure for systems with sensor failures. In *2016 American Control Conference (ACC)*, pages 7523–7528.
- Zhang, Y. y Jiang, J. (2008). Bibliographical review on reconfigurable fault-tolerant control systems. *Annual Reviews in Control*, 32(2):229–252.

Protocolo de comunicación: CANae 0.1 Alpha Version

A.1. Introducción

En este apéndice se explica el diseño de la versión Alpha del protocolo CANae. Este protocolo se basa en la capa de aplicación del protocolo CAN y CANopen. El nombre CANae proviene de la unión de CAN y CONAE.

A.1.1. Revisiones

Tabla A.1: Revisiones de CANae

Ver.	Who.	When	What
0,1 Alpha	Emmanuel Arias	25/06/2017	Initial Version

A.2. Capa de aplicación

La capa de aplicación de CANae está dividida entre la denominada *CANae Application Layer* y la *CANae High Application Layer* Figura A.1. La funcionalidad que brinda la capa de aplicación está dividida en diferentes elementos de servicios dentro de la capa de aplicación. Un elemento de servicios brinda un determinado servicio.

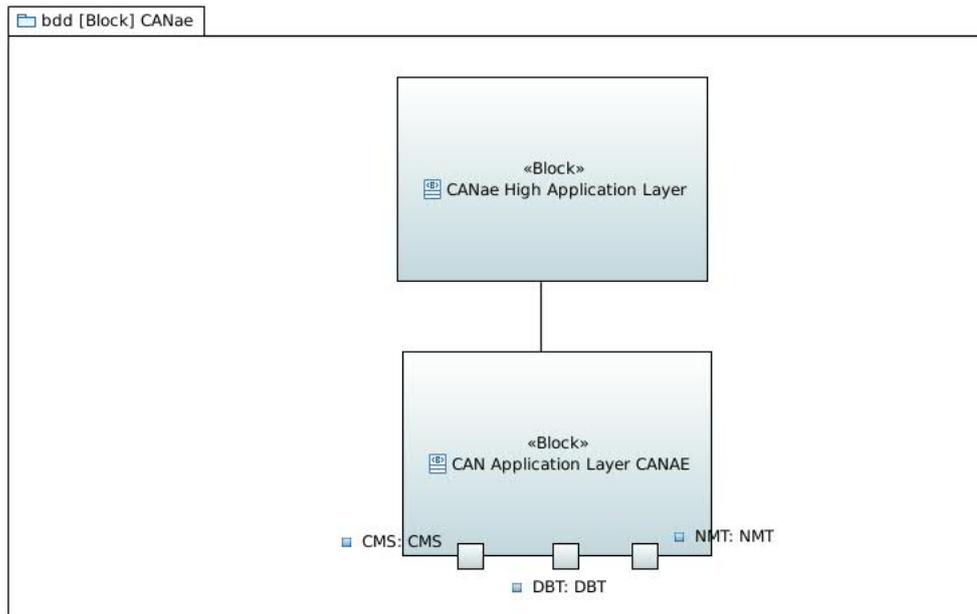


Figura A.1: Estructura de la capa de aplicación de CANae en alto nivel

Una capa de aplicación tiene las siguientes 4 funciones primitivas:

- Request: Una aplicación emite una petición a la capa de aplicación para solicitar un servicio.
- Indication: Emitido por la capa de aplicación con destino una aplicación, para reportar de algún evento interno detectado por la capa de aplicación, o indicar que un servicio fue solicitado.
- Response: Emitido por la aplicación con destino la capa de aplicación para responder a una Indication previa.
- Confirm: Emitido por la capa de aplicación para reportar el resultado de una petición previa.

A.3. Tipos de servicios de la capa de aplicación

Un tipo de servicio define las primitivas que intercambian las aplicaciones de usuario y los diferentes elementos de servicio que se encuentran dentro de la capa de aplicación. Los tipos de servicios posibles en CANae son los siguientes:

- Local service: Solo envuelve un elemento de servicio local.
- Confirmed service: Implica que uno o más elementos de servicio pares. La aplicación de usuario emite un request a su local service element. Este se transmite a sus elementos de servicios pares, la cual es

enviada a la aplicación como una indicación. La otra aplicación emite un response que se transmite a la aplicación originante para confirmar la recepción del request.

- Provide initiated service: Implica solo el elemento de servicio local. El elemento de servicio detecta un evento no solicitado por la aplicación y emite una Indication.

A.4. CANae Application Layer

En CANae divide la tradicional capa de aplicación del modelo OSI en dos capas. En esta sección se estudia la *CANae Application Layer*. Esta capa está conformada por 3 elementos de servicio:

- CAN based Message Specification (CMS): ofrece un ambiente orientado a objeto para diseñar aplicaciones de usuario. Esta entidad ofrece variables y eventos, y especifica cómo un módulo puede acceder a las interfaces de CAN.
- Network Management (NMT): ofrece un ambiente orientado a objetos para permitir que un módulo (el NMT Master) se ocupe de la inicialización y posibles fallas de otros módulos (NMT Slaves).
- Distributor: ofrece el servicio para distribuir dinámicamente el identificador para los diferentes nodos.

Estos están basados en la capa de aplicación de CAN para la industria (CAN in Automation [CiA] International Users and Manufacturers Group e.V. CAN Application Layer for Industrial Applications). Estos elementos de servicios determinan el “qué” puede hacer la capa. Estos son representados como interfaces (Figura A.2).

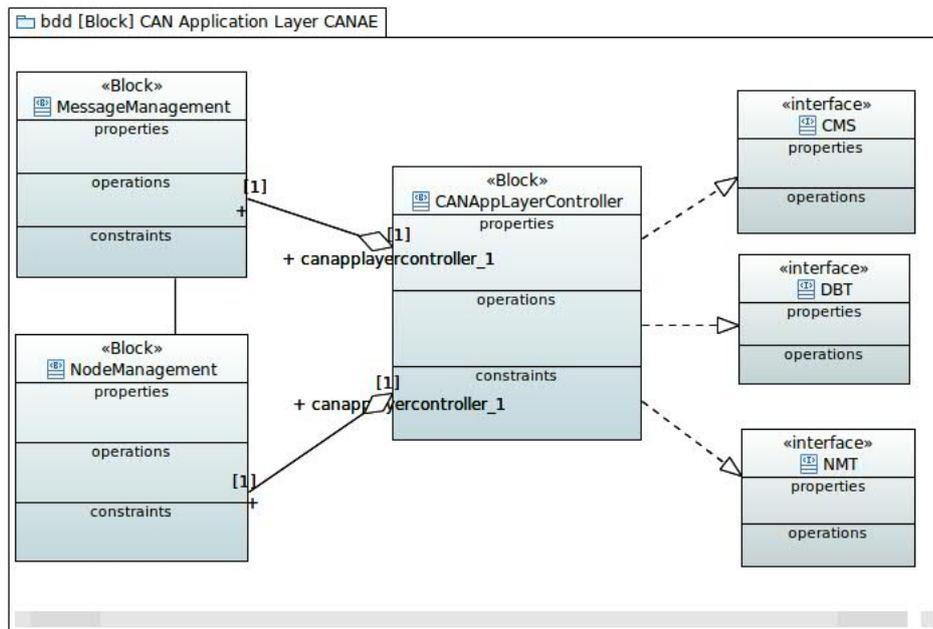


Figura A.2: CANae Application Layer

Dentro de esta capa existen dos entidades importantes:

- Gestor de mensajes: este se encarga de gestionar los mensajes que son enviados y recibidos desde la red. Esta entidad debe ser capaz de determinar si el mensajes contiene datos o eventos. Trabaja en conjunto con el CMS.

- Gestor de nodo: esta entidad se encarga de llevar el control de los nodos existentes en la red. En este nodo se encuentra la tabla de ruteo primario y secundario necesarios para la correcta comunicación.

A.5. CMS

CMS ofrece un ambiente orientado a objeto para diseñar aplicaciones de usuario. CMS ofrece a la aplicación la posibilidad de modelar su comportamiento en la forma de objeto. Este elemento de servicio ofrece variables, eventos que se utilizan para diseñar y especificar cómo la funcionalidad de un módulo puede acceder a las interfaces CAN.

El servicio asume que no hay fallas provenientes de la capa de enlace de datos y la capa física de la red CAN. Las fallas, si ocurren, son resueltas por el NMSE (Network Management Service Element).

A.5.1. Prioridades de los objetos

La arbitración del protocolo CAN se realiza teniendo en cuenta la prioridad de los identificadores de los frames que se envía. La arbitración es no destructiva, debido a la necesidad de tener un procesamiento de mensajes de tiempo real. La prioridad se determina teniendo en cuenta cada uno de los bits. El indentificador de los objetos y eventos de CMS está compuesto por 8 bits más 3 bits que los diferenciará entre objetos y eventos. Los eventos tiene mayor prioridad.

000	(ID)	Tipo de evento
111	(ID)	Prioridad objeto

La prioridad de los objetos es un UInteger de 1 byte. La prioridad más baja es 0 y la prioridad más alta de 1. Por lo tanto la prioridad va desde 00000000b (mayor prioridad) hasta 11111111b (menor prioridad).

A.5.2. Objeto CMS

Un objeto CMS define una estructura de dato que se debe respetar para poder enviar información (variables) a través de la red. El objeto está compuesto por los siguientes campos:

- Name: string de 6 caracteres (6 bytes)
- User_type: {client, server} (1 bit)
- Priority: UInteger (1 bytes)
- Datatype: UInteger identificador del tipo de datos (1 byte)
- Data: Indefinido.

A.5.3. Servicios de CMS

Con respecto a las variables de CMS existen una serie de servicios brindados por CMS, que pueden ser tanto *local services* como *remote services*. Su clasificación depende de si se está accediendo a variables almacenadas dentro del nodo, o bien se está intentando acceder (escribir o leer) una variable remota. Un evento está definido de la siguiente manera:

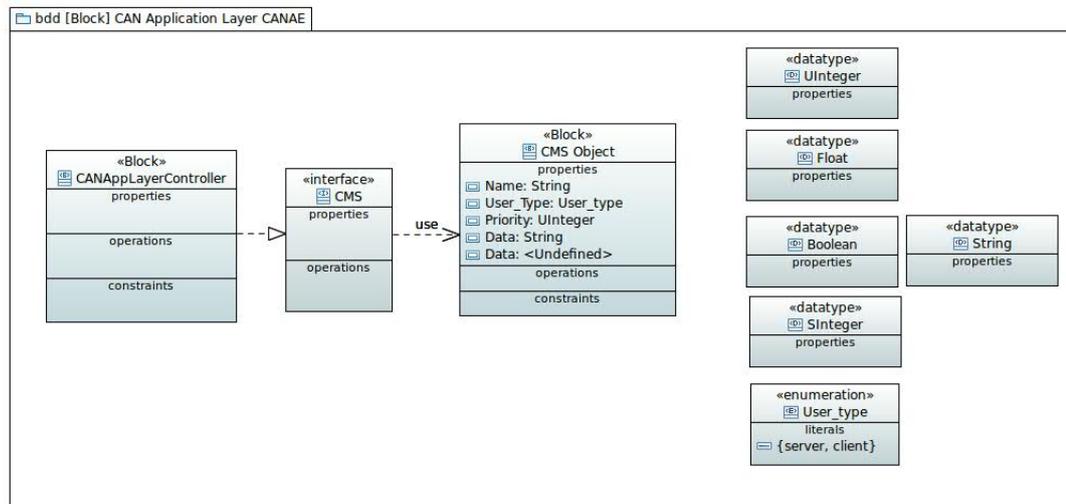


Figura A.3: Definición del Objeto CMS

- Name: string de 6 caracteres (6 bytes)
- User_type: {client, server} (1 bit)
- Priority: UInteger (1 bytes)
- Event_Type: UInteger identificador del tipo de datos (1 byte)
- Data: Indefinido.

A.5.3.1. Servicios locales

- Boolean define_variable(String name, DataType data_type)
 - **String name:** Nombre de la variable.
 - **DataType data_type:** Tipo de dato de la variable.
- Boolean set_variable(String name, UInteger destination, Data data, UInteger priority)
 - **String name:** Nombre de la variable.
 - **UInteger destination:** Dirección del destinatario. Esta dirección debe ser provista por el NMT.
 - **Data data:** Datos a enviar.
 - **UInteger priority:** Prioridad de la variable.
- Boolean update_variable(String name, data data)
 - **String name:** Nombre de la variable
 - **Data data:** Datos a enviar

A.5.3.2. Servicios remotos

- Boolean write_variable(String name, Data data)
 - **String name:** Nombre de la variable a escribir
 - **Data data:** Dato a escribir
- Data read_variable(String name)
 - **String name:** Nombre de la variable a leer

A.5.4. Eventos

Los eventos se utilizan para modelar comportamientos asíncronos, tales como temperatura excedida en un tiempo determinado. La ocurrencia de los eventos es detectada por el nodo, que en ese momento actuará como server y se notifica a todos los clientes. Los servicios al igual que las variables se dividen en servicio local o servicio remoto. Los servicios se describen a continuación.

A.5.4.1. Servicios locales

- define_event(String name, UInteger priority, UInteger event_type)
 - **String name:** Nombre del evento
 - **UInteger priority:** Prioridad del evento
 - **UInteger event_type:** Tipo de evento según tabla A.2

A.5.4.2. Servicios remotos

- Event notify_event()
- Boolean store_event(String name)
 - **String name:** Nombre de evento a almacenar
- Event read_event(String name)
 - **String name:** Nombre de la variable a leer

A.5.4.3. Prioridad de eventos

La prioridad de los eventos se define como una tabla que se define al momento de implementar el protocolo. Al desarrollar la tabla de prioridades se debe cuidar la relevancia de los eventos. La prioridad más baja es 0. El tamaño de la prioridad del evento es de 1 byte. Por lo tanto la prioridad va desde 00000000b (mayor prioridad) hasta 11111111b (menor prioridad).

Esta tabla de prioridad de eventos, debe ir acompañada de un diccionario de eventos, donde se almacena los datos relevantes al evento. En esta se hace una descripción del evento. Sirve de guía para la implementación y desarrollo de aplicaciones que identifican eventos.

El diccionario debería incluir los siguientes campos:

Tabla A.2: Prioridad eventos

00000000	Nombre evento	Mayor prioridad
00000001	Nombre evento	
00000010	Nombre evento	
...
11111111	Nombre evento	Menor prioridad

- Nombre del evento
- Descripción del evento
- Prioridad
- Umbral límite inferior (En el caso de que existiera)
- Umbral límite superior (En el caso de que existiera)

A.6. NMT

Esta es una entidad de la capa de aplicación que permite el correcto funcionamiento de la red. Esta entidad tiene los servicios necesarios para que cada nodo tenga un conocimiento de todos los nodos conectados en la red. Esto permite el armado de una tabla de nodos, que es mantenida en todo los nodos. Esta entidad exige la existencia de un nodo monitor, que es el encargado de monitorear y configurar la red. Luego, una vez en funcionamiento este nodo monitor no será vital para la red. En el NMT se deben implementar los algoritmos de ruteo. El entorno NMT se puede observar en la Figura A.5. Se puede observar también el diagrama de bloques internos en la Figura A.4

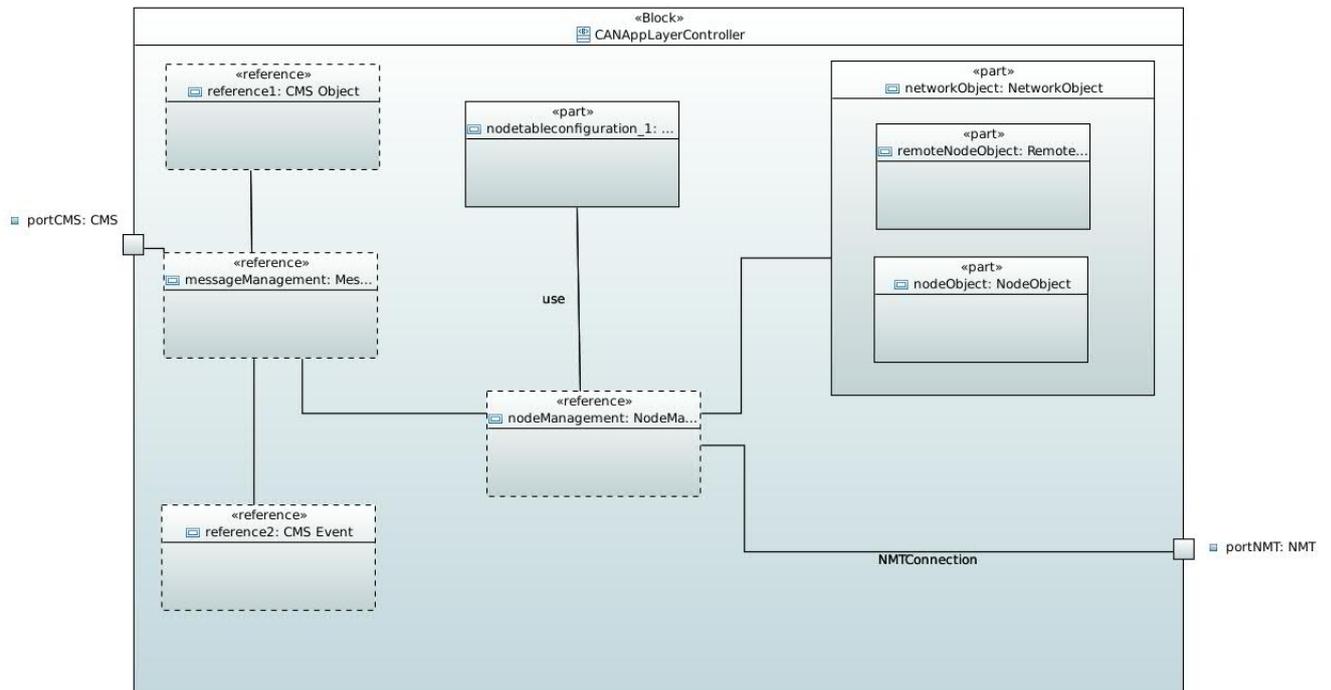


Figura A.4: Definición de bloques internos de la entidad NMT

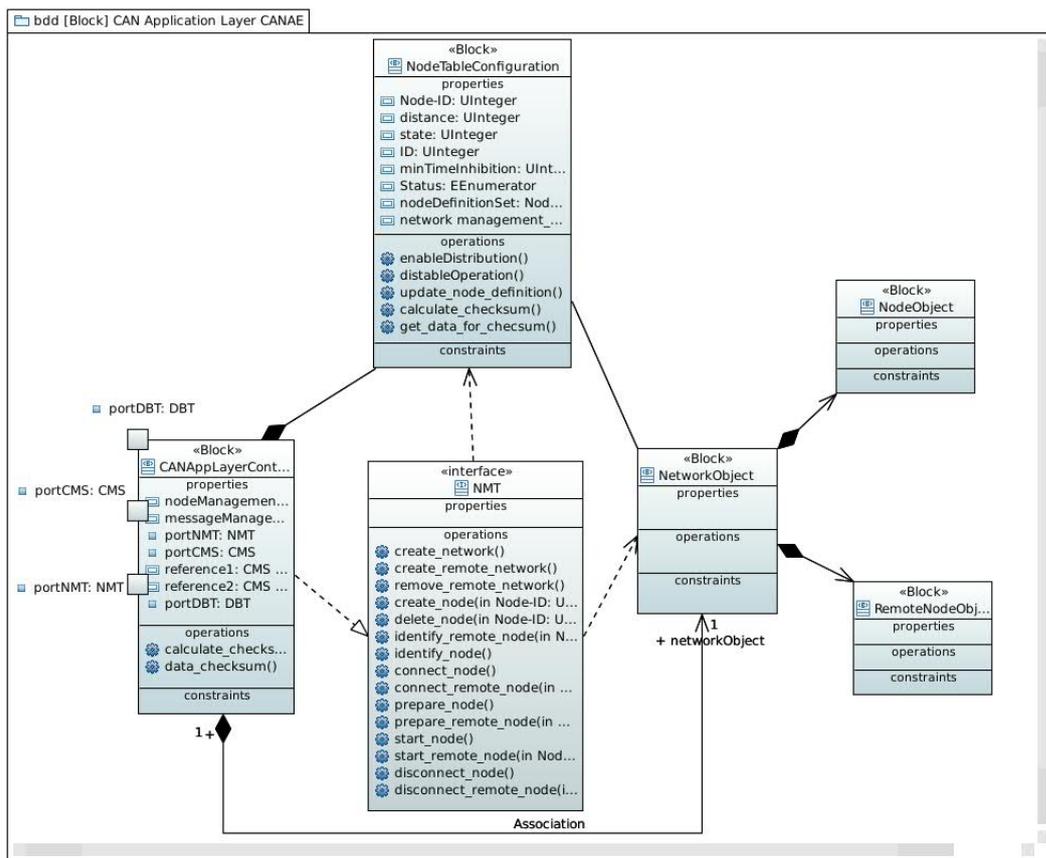


Figura A.5: Definición de la entidad NMT

Por medio de este servicio es posible conocer el estado de la red CAN. Los mensajes de NMT CAN son enviados con máxima prioridad.

A.6.1. Objetos NMT

El Network Management utiliza tres objetos diferentes para modelar una red CAN.

- **Network Object:** Representa todos los módulos conectados en la red CAN. El objeto de red existe en todos los nodos.
- **Remote Node Object:** Cada nodo conectado a la red CAN tiene representado todos los demás nodos.
- **Node Object:** Cada nodo que es gestionado por el servicio NMT está representado por un node object. Este objeto se encuentra modelado en cada uno de los nodos.

Cada nodo y su objeto (remoto y local) es unívocamente identificado en la red por su NMT Address. Esta dirección no se puede cambiar. Esto significa que por cada nodo existe un objeto nodo y un objeto nodo remoto con el mismo NMT Address, replicada en todos los nodos de la red.

A.6.2. Servicios NMT

La entidad NMT ofrece los siguientes servicios:

A.6.2.1. Module Control Service

EL NMT monitor inicializa los nodos NMT que forman parte de la red CAN distribuida, a través de este servicio se asegura que todos los nodos se encuentran configurados y funcionando correctamente.

A.6.2.2. Error Control Services

A través de este servicio, el NMT detecta fallas en la red CAN, ya sea fallas producidas en la capa de enlace de datos (fallas locales) y/o fallas producidos en otros nodos.

A.6.2.3. Configuration Control Services

A través de este servicio se realiza la configuración de la red.

A.6.2.4. Servicios a implementar

La funciones típicas que se deben implementar en el NMT son los siguientes:

- Boolean create_network()
- Boolean create_remote_network()
- Boolean remove_remote_network()
- Boolean create_node(UInteger Node-ID, UInteger address, String description)
 - **UInteger Node-ID**: ID del nodo a crear.
 - **UInteger address**: Dirección del nodo creado. UInteger entre [0,255]
 - **String description**: Descripción del nodo.
- Boolean delete_node(UInteger Node-ID)
 - **UInteger Node-ID**: Dirección del nodo a eliminar. UInteger entre [0,255]
- UInteger identify_remote_node(UInteger Node-ID)
 - **UInteger Node-ID**: Dirección del nodo a identificar
- UInteger identify_node()
- Boolean connect_node()
- Boolean connect_remote_node(UInteger Node-ID)
 - **UInteger Node-ID**: ID del nodo remoto a conectar.
- Boolean prepare_node()
- Boolean prepare_remote_node(UInteger Node-ID)
 - **UInteger Node-ID**: ID del nodo a pasar a modo preparado

- Boolean `start_node()`
- Boolean `start_remote_node(UInteger Node-ID)`
 - **UInteger Node-ID**: ID del nodo que comenzará a participar en la red.
- Boolean `disconnect_node()`
- Boolean `disconnect_remote_node(UInteger Node-ID)`
 - **UInteger Node-ID**: ID del nodo a desconectar.

A.6.3. Protocolos NMT

A.6.3.1. Crear red CANae

Para conectar correctamente la red CANae se debe respetar el protocolo NMT para crear la red. En primer lugar el nodo monitor debe enviar el mensaje `create_remote_network()` para avisar a todos los nodos que deben crear sus propias instancias de *NetworkObject* a través de `create_network()`. Automáticamente, los nodos deben crear una propia instancia del nodo mediante `create_node()`. Luego, el nodo monitor envía el mensaje de `prepare_remote_node()` para prepara todos los nodos a conectarse a la red CANae. Paso siguiente, el nodo monitor envía la señal de `connect_remote_node(UInteger Node-ID)` con la dirección de todos los nodos que figuran en *NodeTableConfiguration* preprogramada. Así, cada nodo, correctamente preparado, se conecta a la red a través de `connect_node()`.

Este protocolo puede ser observado en la Figura A.6

A.6.3.2. Crear Red

Para lograr la correcta conexión de los nodos a la red, estos deben configurarse. Cuando se alimenta la red, los nodos deben iniciar en estado de *escucha* esperando la orden del nodo monitor de crear la red mediante `create_remote_network()`. Esta se envía en broadcast a todos los nodos conectados. Cuando el controlador de la capa de aplicación recibe el mensaje a través del servicio NMT, este manda un mensaje al gestor del nodo interno, el cual crea la *tabla de configuración del nodo* y la actualiza con la información necesaria.

A.6.3.3. Crear Objeto Red, Objeto Nodo y Objeto Nodo Remoto

Antes de comenzar a funcionar la red, cada nodo debe crear un objeto de red exigidos por el protocolo NMT. Una vez que el proceso Crear Red termina, el Gestor de la capa de Aplicación envía un mensaje `create_node()` al Gestor del Nodo. El Gestor de Nodo crea un *Network Object*. Este a su vez crea un *Node Object* y un *Remote Node Object*. Estos objetos pertenecen a la entidad NMT (Figura A.8).

A.6.3.4. Preparar nodos para conexión

Para lograr la conexión de los nodos a la red CAN, es necesario que el nodo monitor envíe la señal `prepare_remote_node(UInteger Nodo-ID)` a cada uno de los nodos conectados a la red siguiendo la propia tabla de configuración de nodos preprogramada. Cada nodo recibe la señal procesa aquella que le pertenece, y comienza a preparar el nodo para lograr la correcta conexión. Para ello debe actualizar su propia tabla de configuración de nodos, como así también actualizar su *Objeto Red* (Figura A.9).

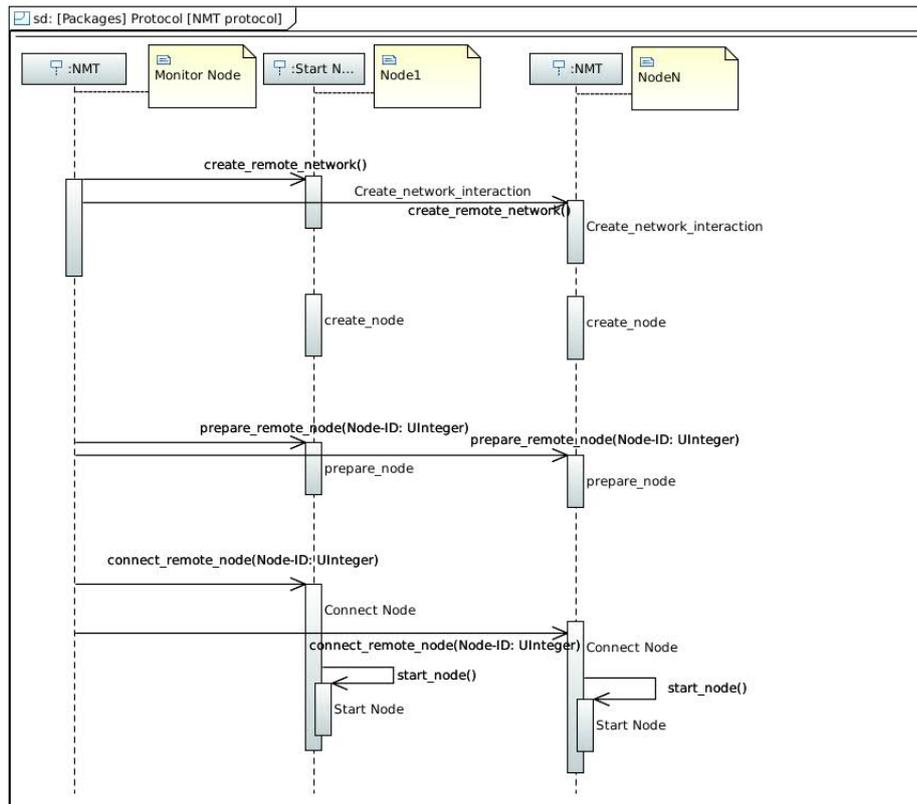


Figura A.6: Protocolo NMT

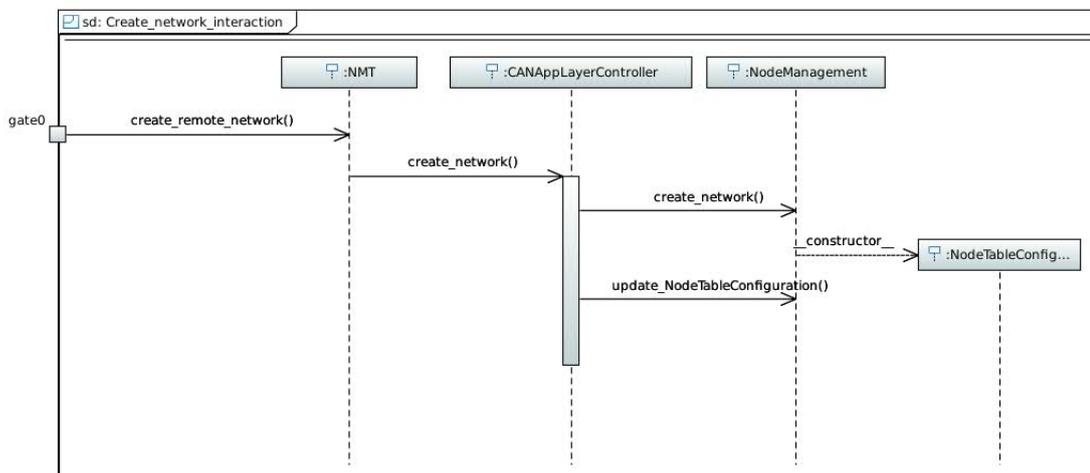


Figura A.7: Crear Red

A.6.3.5. Conectar nodos en la red

Para conectar los nodos a la red se tiene que modificar la información de la *Node Table Configuration* y en el *Objeto de Red*. Una vez realizado esto se puede comenzar con la comunicación de los nodos (Figura A.10).

A.6 NMT

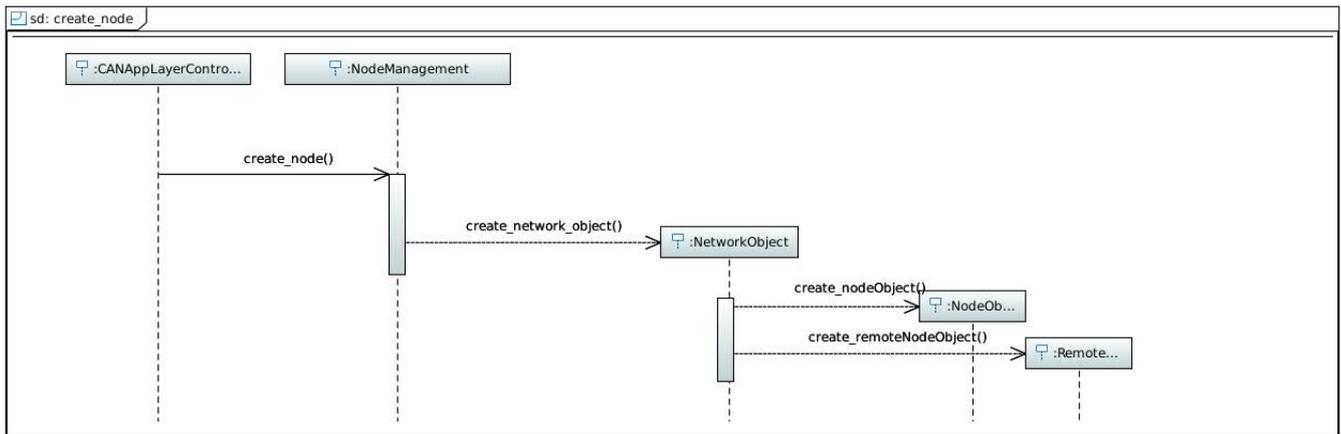


Figura A.8: Crear Objeto Red, Objeto Nodo y Objeto Nodo Remoto

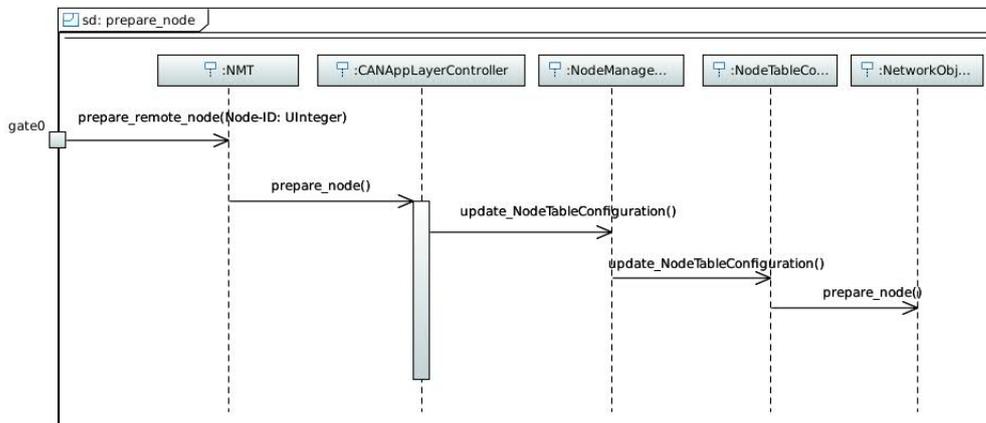


Figura A.9: Preparar nodo para la conexión a la red CAN

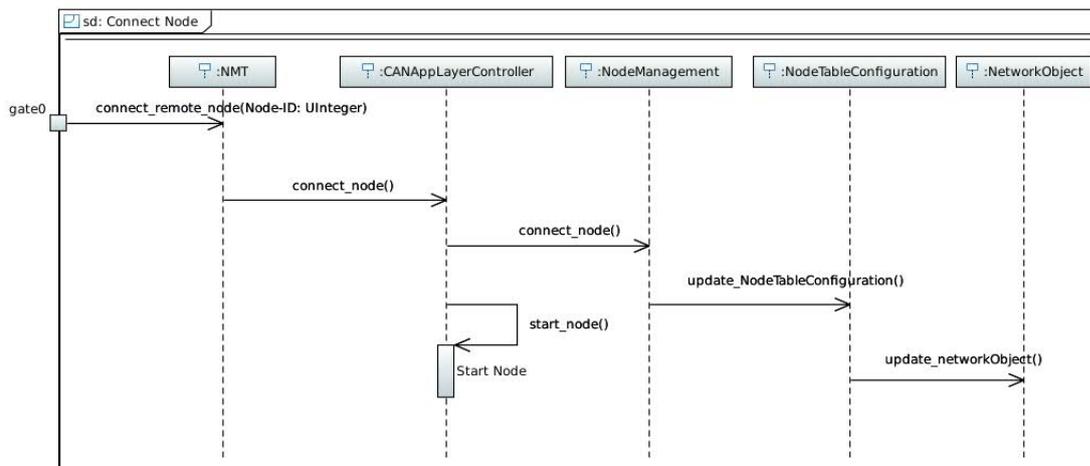


Figura A.10: Conectar el nodo a la red CAN

A.6.3.6. Comenzar la comunicación

Para que el nodo pueda comenzar a comunicar debe encontrarse en un estado *Start*. Esto representa una actualización de la *Node Table Configuration* y del *Objeto de Red* (Figura A.11).

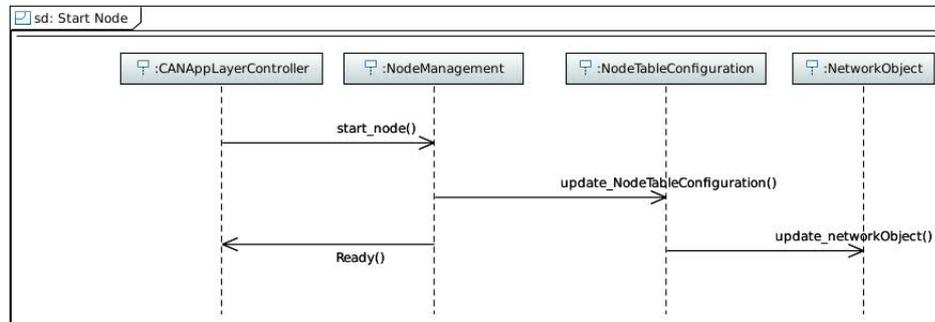


Figura A.11: Comenzar la comunicación de nodos

A.7. DBT

La principal complicación del desarrollo de una red distribuida basada en el BUS CAN, es que puedan comunicarse correctamente entre sí, cómo se deben asignar los identificadores (Node-ID) correctamente a cada uno de los nodos y cómo se asignan los tiempos de inhibición. Los Node-ID y los tiempos de inhibición se deben distribuir entre los nodos de tal manera de asegurar que:

- Se prevean los conflictos entre nodos. Por ejemplo, diferentes funciones que usan el mismo identificador.
- Se prevean la falta de coincidencia. Por ejemplo, que existan diferentes indenticadores para el mismo nodo.
- ofrecer el control integrado para un comportamiento dinámico del sistema.

El protocolo CAN (CAN Application Layer for Industrial Applications CiA/DS204-1) dispone la existencia de 3 métodos para distribuir identificadores y tiempo de inhibición a un módulo.

- **Distribución estándar:** en este método los identificadores y tiempos de inhibición son estandarizados por el módulo proveedor (nodo monitor. Una distribución estándar requiere la estandarización de todas las funciones y su correspondiente identificador.
- **Distribución estática:** En este método los identificadores y los tiempos de inhibición son fijos. Todos los identificadores y los tiempos de inhibición son fijados en el momento del desarrollo.
- **Distribución dinámica:** En este método los identificadores y los tiempos de inhibición son distribuidos vía red CAN a través del servicio estándar y un protocolo definido.

En esta instancia de trabajo, para el desarrollo del protocolo CANae, se utilizó el *método de distribución estática*. Esto exige que los nodos ya cuenten con un identificador y tiempos de inhibición por defecto. Esto se asigna en el momento del desarrollo del diseño e ingeniería de la arquitectura de aviónica. Esto reduce considerablemente la complejidad de la entidad DBT. Si bien originalmente, el DBT es un elemento de servicio

de la capa de aplicación CAN que ofrece una distribución dinámica de identificadores y tiempos de inhibición, en CANae es utilizado para asegurar la correcta comunicación y consistencia de los nodos y la red en su conjunto.

La arquitectura del DBT se puede observar en la Figura A.12

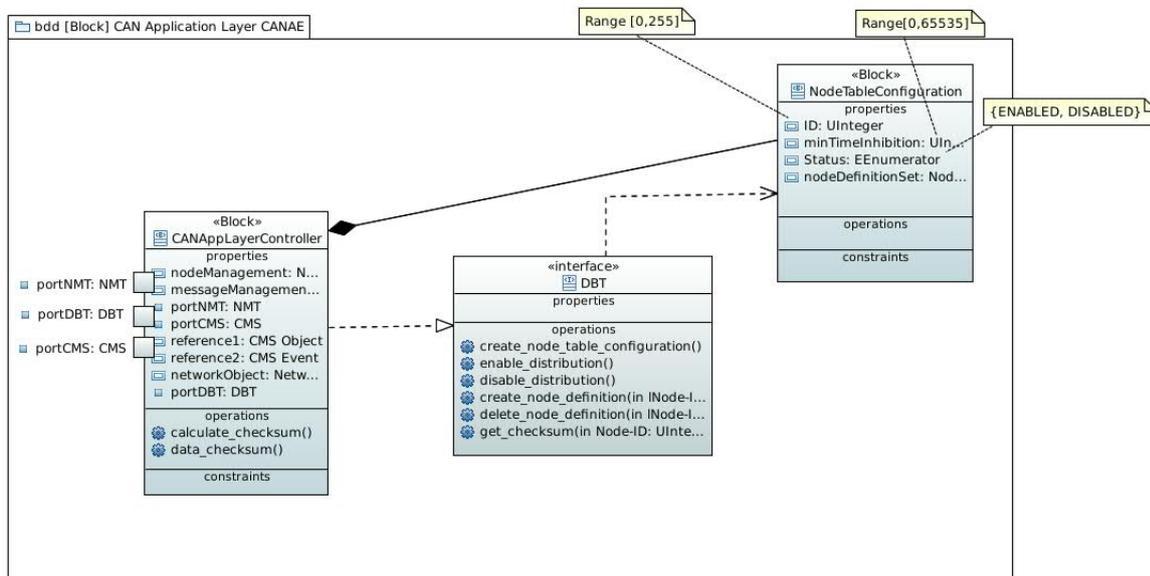


Figura A.12: Arquitectura de la entidad DBT

A.7.1. Objetos y servicios de DBT

El DBT de CANae utiliza 2 objetos para modelar su funcionamiento:

- **Node-ID Database (NodeTableConfiguration):** Esta tabla contiene la definición de todos los nodos conectados a la red. Esta tabla está compuesta por Node-ID Definitions
- **Node-ID Definition:** define todos los atributos de un Node-ID. Contiene una definición de usuario creado por el usuario.

El DBT de CANae ofrece las siguientes categorías de servicio:

- **Distribution Control Services:** esta es tarea del nodo monitor de enviar el identificador y los tiempos de inhibición a todos los nodos conectados a la red. Los datos que son enviados por el nodo monitor son corroborados por los nodos, comprobando que el ID enviado coincida con su propio Node-ID.
- **Consistency Control Services:** a través de este servicio cada nodo puede detectar inconsistencia en el *NodeTableConfiguration* e inconsistencia entre nodos.

A.7.2. Descripción de los servicios DBT

Los servicios se describen en forma de tabla que contiene los parámetros de cada función que se define para ese servicio. Los parámetros determinan el tipo de servicio. Todos los servicios asumen que no ocurrieron ningún tipo de error en la capa de física ni en la capa de enlace de datos.

A.7.3. Objetos DBT

Los objetos que utiliza la entidad DBT para modelar el servicio son los siguientes:

- **Node Table Configuration (Node-ID Database)**
 - **Atributos**
 - Estado: Uno de los valores ENABLED, DISABLED. Este atributo indica si el DBT Monitor es capaz de distribuir los Node-ID y tiempos de inhibición. En esta instancia de trabajo el valor del estado en el DBT del Nodo Monitor será DISABLED
 - Node Definition Set set de todas las definiciones.
 - **Node-ID Defintion**: Las definiciones contiene los siguientes datos:
 - ID: Valor en el rango de [0,255]
 - Mínimo tiempo de inhibición: en el rango [0,65535] indicando el valor mínimo en unidades de 100 μ sec para el tiempo de inhibición.

A.7.4. Servicios DBT

A.7.4.1. Distribution Control Services

- create_node_table_configuration()
- enable_distribution()
- disable_distribution()
- create_node_definition(UInteger lNode-ID, UInteger hNode-ID, UInteger minInhibitTime)
 - **UInteger lNode-ID**: Límite inferior del rango de IDs.
 - **UInteger hNode-ID**: Límite superior del rango de IDs.
 - **UInteger minInhibitTime**: Tiempo mínimo de inhibición de cada nodo.
- delete_node_definition(UInteger lNode-ID, UInteger hNode-ID)

A.7.4.2. Consistency Control Service

- get_checksum(UInteger Node-ID, Float checksum)
 - **UInteger Node-ID**: ID del nodo que desea controlar la consistencia.
 - **UInteger checksum**: checksum de la DBT Definition.

A.7.5. Protocolos DBT

A.7.5.1. Crear NodeTableConfiguration

En la tabla de configuración del nodo juega un papel similar al que lo haría el Node-ID DATABASE del estándar CAN. La principal diferencia, es que esta tabla se encuentra presente en todos los nodo y no en el

“Master”, ya que este protocolo está pensado para ser utilizado en redes distribuidas, donde no existe un nodo central o master.

El nodo monitor envía la señal a los nodos de crear la tabla de configuración del nodo a través de *create_node_table_configuration()*. Este es recibido a través de la interfaz DBT del nodo, y crea el objeto *NodeTableConfiguration*. Esto se observa en la Figura A.13.

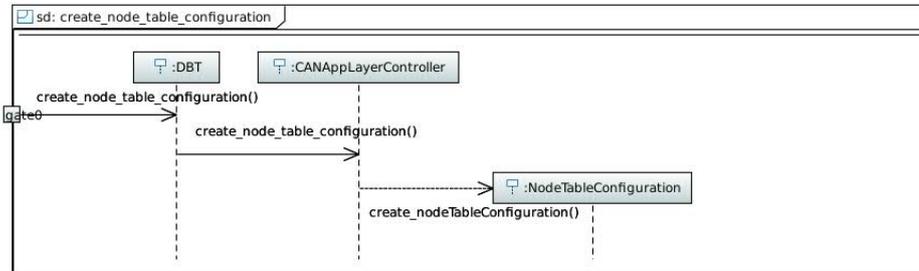


Figura A.13: Protocolo create_node_table_configuration

A.7.5.2. Habilitar distribución

La distribución de Node-ID es utilizada cuando se usa el método de asignación dinámica. En esta versión del protocolo la distribución del ID se encuentra desactivada. En la Figura A.14 se puede observar el protocolo.

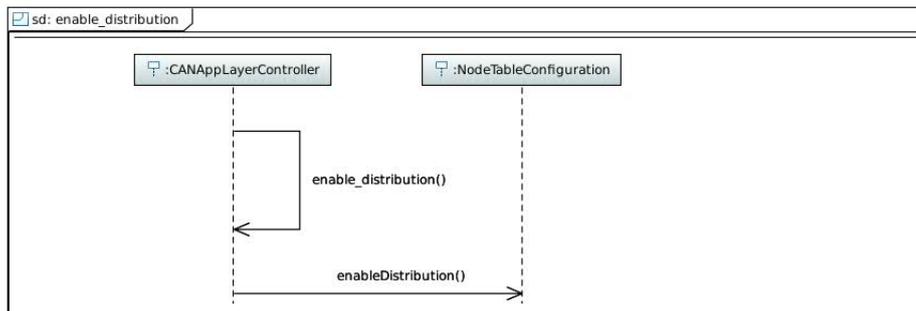


Figura A.14: Protocolo enable_distribution

A.7.5.3. Deshabilitar distribución

Esta opción se encuentra por defecto en esta versión del protocolo CANae. El ID se distribuye estáticamente. Son preestablecidos durante el desarrollo. En la Figura A.15

A.7.5.4. Crear definición del nodo

El nodo monitor envía la señal para crear las definiciones del nodo. Para ello es necesario que el nodo monitor envíe los rangos permitidos para la definición de los nodos. Esto se define a través de las variables *lNode-ID* y *hNode-ID*. En esta versión del protocolo estas variables no son utilizadas ya que los Node-ID ya se encuentran definidas estáticamente. También el nodo monitor envía el tiempo mínimo de inhibición explicado anteriormente. El nodo monitor hace este proceso con todos los nodos de la red a través del mensaje *create_node_definition*. Este se lo puede observar en la Figura A.16.

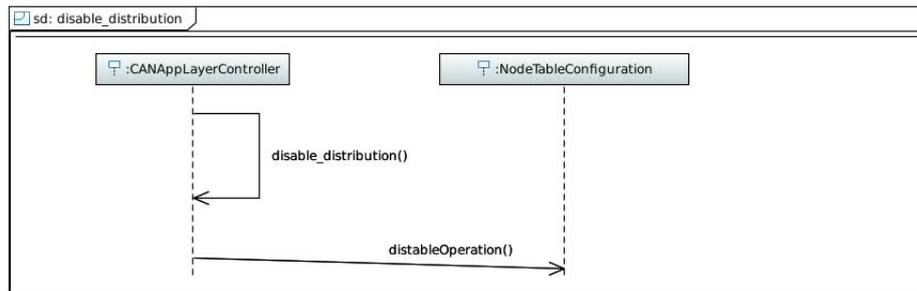


Figura A.15: Protocolo disable_distribution

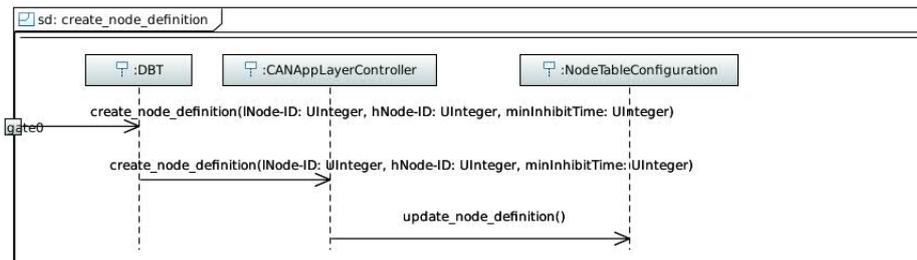


Figura A.16: Protocolo create_node_definition

A.7.5.5. Eliminar la definición del nodo

Esta función es utilizada para eliminar la definición del nodo. Se lleva a cabo para eliminar un nodo de la red. Se hace a través de la función *delete_node_definition()*. Esta se la puede observar en la Figura A.17

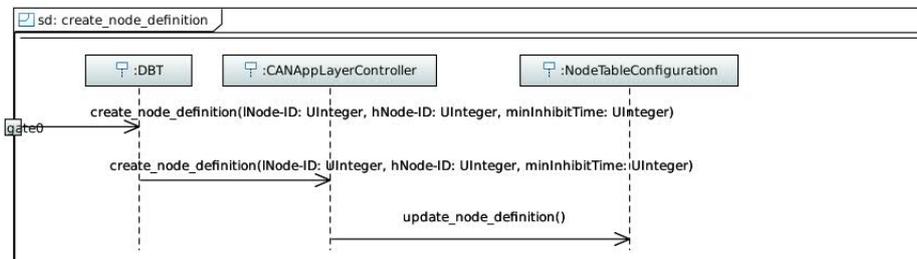


Figura A.17: Protocolo delete_node_definition

A.7.5.6. Checksum

Para poder asegurar la integridad de los nodos conectados a la red, se puede utilizar este mecanismo para comprobar de que la información, referente a la definición de los nodos, se encuentra correctamente en todos los nodos. Para ello, cada nodo lleva a cabo un checksum de su información y la envía por la red. Como esta está basada en el Bus-CAN, el mensaje llega correctamente a todos los nodos. Los nodos realizan sus comprobaciones y responden. Si la información es correcta deben responder con un mensaje que contengan bits NO dominantes. En el caso de que se produzca alguna diferencia en los checksum deberán responder con bits dominantes. Cuando esto se produzca deberán llevarse a cabo medidas de recuperación o reconfiguración de la red. Este protocolo se lo puede observar en la Figura A.18

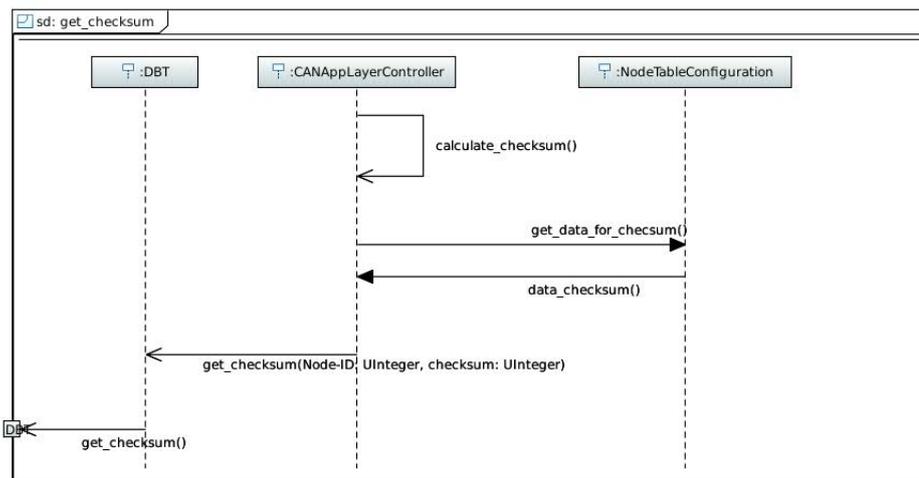


Figura A.18: Protocolo get_checksum

A.8. Gestor de mensajes

El **Message Management** se encarga de gestionar los mensajes que son enviados y recibidos desde la red CAN. Esta entidad debe ser capaz de determinar si el mensaje (enviado o recibido) contiene datos o eventos. El gestor de mensajes tiene cuatro entidades:

- *Receiver Manager*: es el encargado de recibir los mensajes de la capa inferior a través de la interfaz NMT. Esta entidad recibe y despaquetiza los mensajes.
- *Prepared Message*: es el encargado de recibir los mensajes desde las capas superiores para ser enviado a la red.
- *Sorter Message*: es el encargado de clasificar los mensajes, ya sea aquellos que envía a la red CAN como los que recibe de esta. Su principal función es verificar el tipo de mensaje que está recibiendo o enviando (datos o eventos). Los eventos tiene mayor prioridad que los datos y deben atenderse con urgencia.
- *Buffer Message*: este sirve de buffer para almacenar mensajes tanto que se envían como lo que se reciben. Este buffer es utilizado por el *Sorter Message*.

En la Figura A.19 se muestra la composición del gestor de mensajes. Por otra parte en la Figura A.20 muestra el diagrama interno del gestor de mensajes.

A.8.1. Receiver Manager

Esta entidad perteneciente al *Message Management* se encarga de recibir los mensajes de la capa inmediatamente inferior a través de la interfaz NMT (y sus protocolos). Esta entidad recibe y despaquetiza los paquetes. Se supone que los mensajes llegan sin ningún tipo de error.

Las principales funciones de esta entidad son las siguientes:

- `receive_message(Message message)`: esta es la operación que se debe llamar para recibir el mensaje proveniente de la capa inferior.
 - **Data message**: Es el mensaje que recibe de las capas inferiores de la red CAN.

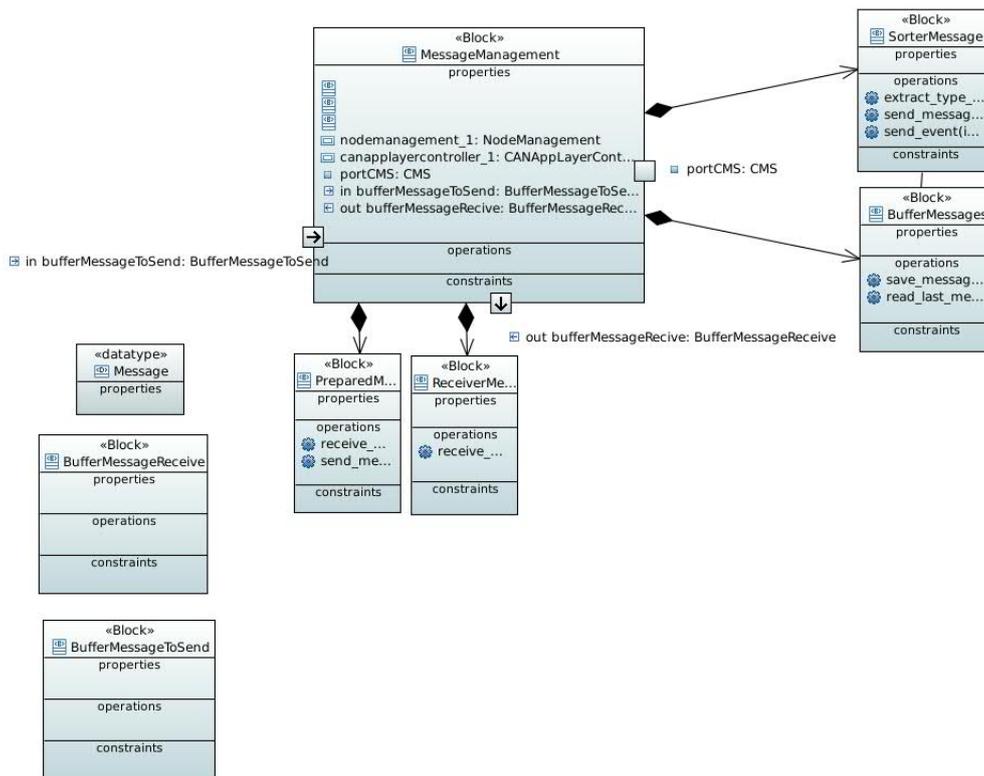


Figura A.19: Arquitectura de *Message Management*

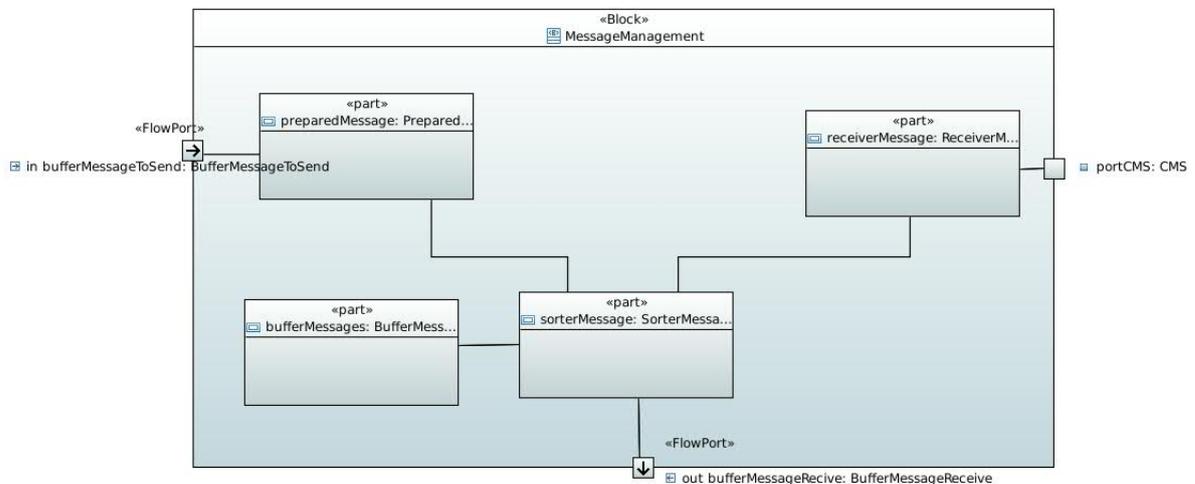


Figura A.20: Diagrama interno del *Message Management*

A.8.2. Prepared Message

Esta entidad es la que recibe los mensajes desde las capas superiores y lo envía al *Sorter Message*.

- receive_message(Message message): esta es la operación que se debe llamar para recibir el mensaje proveniente de la capa superior.
 - **Message message:** Es el mensaje que recibe de las capas inferiores de la red CAN.

- `send_message(Message message)`: esta es la operación que se debe llamar para enviar el mensaje al *Sorter Message*.
 - **Message message**: Es el mensaje que se envía al *Sorter Message*

A.8.3. Sorter Message

Esta entidad es el encargado de clasificar los mensajes que provienen tanto del *Receiver Message* y *Prepared Message*. Este clasifica los mensajes recibidos o por enviar, de modo tal de conocer aquellos que sean datos o eventos. Debe tenerse en cuenta que los eventos tiene mayor prioridad que los datos, debido a que (en su mayoría) informan de algún problema a la red.

Las principales funciones con las que cuenta esta entidad son las siguiente:

- Boolean `extract_type_message(Message message)`: esta operación permite extraer el tipo de información que contiene el mensaje: datos o evento.
 - **Message message**: es el mensaje que recibe de las capas inferiores de la red CAN.
 - **Return Boolean**: retorna 1 se se trata de un evento, 0 si se trata de un dato.
- Boolean `send_message(Message message)`: esta función envía los mensajes a las capas inferiores a través del protocolo NMT.
 - **Message message**: es el mensaje formateado listo para enviar.
 - **Return Boolean**: retorna 1 si la operación se llevó correctamente.
- Boolean `send_event(Message message)`: esta función envía el evento en forma de mensajes formateado.
 - **Message message**: es el mensaje formateado listo para enviar.
 - **Return Boolean**: retorna 1 si la operación se llevó correctamente.

A.8.4. Buffer Messages

Esta es la entidad que permite almacenar los mensajes que son recibidos y enviados. Esta entidad debe controlar la integridad de los mensajes que se encuentran almacenados. Además se encarga de escribir y leer los mensajes almacenados en el buffer. El buffer es una FIFO¹.

Las principales funciones de esta entidad son las siguientes:

- `save_message(Message message)`: esta función almacena el mensaje en el buffer de mensajes
 - **Message message**: es el mensaje formateado listo para enviar.
- `Message read_last_message()`: esta función lee el último mensaje.
 - **Return Message**: retorna el mensaje extraído del buffer.

¹First In First Out

A.8.5. Buffer Message Receive y Buffer Message To Send

La comunicación entre el Gestor de Mensajes y las capas, tanto superior como inferior, se realizan por medio de buffers. Se pueden dar las siguientes situaciones:

- Llegada de un mensaje de la capa superior para ser enviada a través de la red. En este caso la aplicación de usuario necesita enviar un mensaje, para ello debe colocar el mensaje en el *Buffer Message To Send*. El *Gestor de Mensaje* debe detectar (por medio de algún mecanismo) que se ha escrito en el buffer, para tomar el mensaje.
- Enviar un mensaje a la red. Cuando el gestor necesita enviar un mensaje lo tiene que almacenar en el *Buffer Message To Send*. Esto “alertaría” por medio de algún mecanismo a las capas inferiores, de que existe un nuevo mensaje a enviar

A.9. Gestor de Nodo

El *Node Management* es una entidad que se encuentra en el NMT, y es la encargada de llevar a cabo el control de los nodos existentes en la red. En esta entidad se encuentra almacenada la tabla de ruteo primario y secundario, utilizado para lograr la correcta comunicación y configuración de la red.

El objetivo del gestor del nodo, es conocer todos los nodos que se encuentran conectados en la red, y la distancia que existe entre el nodo actual y los demás.

Por medio de las tablas de ruteo (primaria o secundaria en caso de fallas) el nodo conoce a qué nodos puede enviar mensajes y a cuáles no. Debe aclararse que nada impide enviar un mensaje a un nodo “desconectado”, pero esta desición es ineficiente y altamente peligrosa, ya que se puede estar enviando mensajes críticos a un nodo “caído”.

El gestor de nodo de cada dispositivo conectado a al red, es el encargado de llevar a cabo el protocolo *heartbeat*, el cual le indica a todos los nodos conectados a la red que este está “vivo”

En la Figura A.21 se puede observar la arquitectura del Gestor de Nodos.

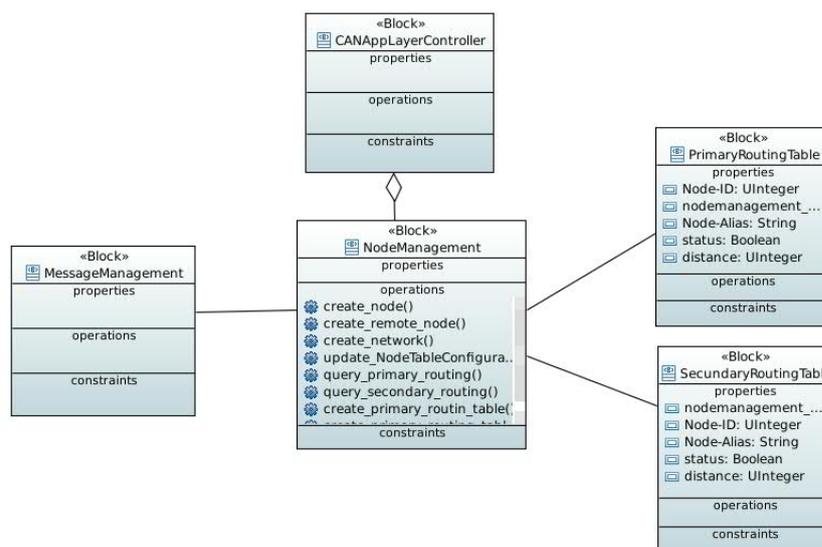


Figura A.21: Arquitectura del *Node Management*

A.9.1. Tabla primaria y secundaria de ruteo

Los principales objetos del *Node Management* son las tablas de ruteo. Estas se las puede representar como una lista, en la cual se incluye información de importancia para mantener actualizada la estructura de la red. Los nodos se basarán sobre esta lista para tener conocimiento de los nodos conectados a la red y con esto, llevar a cabo una correcta distribución de las tareas.

Los algoritmos de ruteos se deben basar en estas tablas, ya que contienen toda la información necesarias para cumplir con sus objetivos.

A continuación se detalla la tabla primaria y tabla secundaria de ruteo.

A.9.1.1. Tabla primaria

La Tabla Primaria de Ruteo es la tabla principal estática utilizada por el *Node Management* para la gestión de los nodos de la red. Esta tabla mantiene la a los n-ísimos nodos conectados a la red, desde el punto de vista del nodo “actual”. Es decir, para cada nodo existe una y sola una tabla, y esa tabla es única en la red, ya que es creada desde el punto de vista del nodo “host”.

La estructura de la lista es simple. Cada renglón de la lista es un nodo, siendo el primer renglón obligatoriamente el nodo de la tabla, llamado nodo “host”. Los siguientes renglones de la lista representan los demás nodos conectados a la red. Por convención, los nodos se listan comenzando por los vecinos de derecha a izquierda. Luego se continúa la lista tratando de seguir la convención de derecha a izquierda.

La tabla tiene la siguiente estructura:

- **UIntegerNode-ID:** Es el ID único para cada nodo.
- **String Node-Alias:** Es el alias para cada nodo. Este alias puede ser utilizado para comodidad por parte del usuario para referenciar los nodos.
- **UInteger Address:** Es la dirección del nodo.
- **Boolean Status:** Es el estado del nodo. Pude ser { CONNECTED, DISCONNECTED }.
- **UInteger Distance:** Es la distancia (cantidad de nodos) que hay entre el nodo “host” y los demás nodos de la red.

A.9.1.2. Tabla secundaria

La Tabla Secundaria de Ruteo es una tabla estática similar que la Tabla Primara de Ruteo, con la diferencia sustancial que es utilizada cuando hay fallas en el ruteo. Esta tabla se genera cuando se producen fallas en uno o más nodos, y se necesita conocer la estructura de la red, para luego tomar desiciones sobre la configuración de la misma.

La estructura de la Tabla Secundaria de Ruteo es similar a la Tabla Primara de Ruteo. La diferencia sutancial es que no se encuentran todos los nodos de la red, sino aquellos que se encuentran “funcionales”.

La estructura de la tabla se detalla a continuación:

- **UIntegerNode-ID:** Es el ID único para cada nodo.

- **String Node-Alias:** Es el alias para cada nodo. Este alias puede ser utilizado para comodidad por parte del usuario para referenciar los nodos.
- **UInteger Address:** Es la dirección del nodo.
- **Boolean Status:** Es el estado del nodo. Puede ser { CONNECTED, DISCONNECTED }.
- **UInteger Distance:** Es la distancia (cantidad de nodos) que hay entre el nodo “host” y los demás nodos de la red.

A.9.2. Servicios

Los servicios brindados por el *Node Management* son los siguientes:

- Boolean `create_node(UInteger Node-ID, UInteger address, String description)`: esta función crea el objeto nodo. Este servicio ya fue descrito en A.6.2.4. Entre otras cosas, este servicio crea la primera entrada en la Tabla Primaria de Ruteo.
 - **UInteger Node-ID:** ID del nodo a crear.
 - **UInteger address:** Dirección del nodo creado. UInteger entre [0,255]
 - **String description:** Descripción del nodo.
- Boolean `connect_remote_node(UInteger Node-ID)`: envía la señal a sus nodos vecinos para que creen una referencia de este nodo en los demás.
 - **UInteger Node-ID:** ID del nodo remoto a conectar.
- Boolean `create_network()`: esta función fue descrita en A.6.2.4. Este servicio crea una instancia de la Tabla Primaria de Ruteo.
- Boolean `update_NodeTableConfiguration()`: esta función actualiza la tabla de configuración del nodo (Vista en: A.6.2.4).
- Data query `_primary_routing()`: esta función lleva a cabo una consulta a la Tabla Primaria de Ruteo. Se la utiliza cuando se necesita extraer datos de la configuración de la red.
- Data query `_secondary_routing()`: esta función lleva a cabo una consulta a la Tabla Secundaria de Ruteo. Se la utiliza cuando se necesita extraer datos de la configuración de la red luego de fallas en los nodos.
- Boolean `add_node(UInteger Node-ID, String Node-Alias, Boolean status, UInteger distance, UInteger address)`: este servicio se lleva a cabo cuando llega el mensaje `connect_remote_node(UInteger Node-ID)`. Este servicio agrega un nodo (renglón) en la Tabla Primaria de Ruteo.
 - **UInteger Node-ID:** ID del nodo a agregar.
 - **String Node-Alias:** Alias del nodo a agregar.
 - **Boolean status:** Estado del nodo. CONNECTED=1, DISCONNECTED=0. Por defecto los nodos están en estado 1 (CONNECTED).
 - **UInteger distance:** Distancia que existe entre el nodo “host” y el nodo a agregar.
 - **UInteger address:** Dirección del nodo a agregar.

A.10. Formato de mensajes

Los mensajes enviados entre los nodos deben tener un formato compatible con las redes CAN convencionales. Por ello el mensaje tiene un tamaño que varía desde los 45 bits hasta los 109 bits (5 bytes hasta los 13 bytes), dependiendo de la cantidad de datos en el *Data Field*.

En la Figura A.22 se observa el formato de Frame de datos de los mensajes CANae.

SOF	Priority	Node-ID	RTR	Type of Message	Reserved bit	Data length	Data field	CRC	CRC delimiter	ACK	ACK Delimiter	EOF
1	4	7	1	1	1	0-8	0-64	1	1	1	1	1

Figura A.22: Frame de datos CANae

Los campos del mensaje son los siguientes:

- Start of Frame (SOF) - 1 bit: Este bit marca el comienzo de un frame. Suele ser 0.
- Priority - 4 bits: Este campo declara la prioridad del frame.
- Node-ID - 7 bits: El ID del Nodo origen.
- Remote transmission request (RTR) - 1 bit: Bit dominante (0) para frames de datos o eventos, y bit recesivo (1) para frames remotos. Un frame remoto es un mensaje que es enviado en forma automática por algún sensor. Los frames remotos se utilizan para enviar telemetría por parte de los sensores y componentes.
- Type of Message (TOM) - 1 bit: Bit dominante (0) si es un evento, y bit recesivo (1) para frames remotos.
- Bit reservado - 1 bit: Este bit está reservado para futuras aplicaciones. En esta versión debe ser un bit dominante (0).
- Data length - 6 bits: Cantidad de bytes del campo de datos.
- Data field - 0-64 bits (0-8 bytes): Datos.
- CRC - 15 bits: Cyclic redundancy check.
- CRC delimitier - 1 bit: Debe ser recesivo (1).
- ACK - 1 bit: El transmisor debe poner este bit en recesivo (1), y el receptor puede responder este bit con un bit dominante (0).
- ACK delimiter - 1 bit: Debe ser recesivo (1).
- End-of-Frame (EOF) - 7 bits: Final del frame de datos.

El *Data Field* tiene el siguiente formato: el primer byte siempre es el ID de la función, evento y/o comando. Esto significa que es posible definir un número de $2^8 = 256$ funciones, eventos y comandos diferentes.

Los siguientes bytes (del 1 al 63) son argumentos de las funciones o datos.

Debe entenderse que esto no es mandatorio, ya que se pueden definir ID de dos o más bytes reduciendo la cantidad de datos.

A.10.1. Start of Frame

Este bit marca el comienzo de un mensaje CANae. Este es simplemente un bit dominante (0). Para que un dispositivo pueda enviar mensajes, el bus debe estar en estado IDLE.

A.10.2. Priority

Este campo indica la prioridad del mensaje. La prioridad está compuesta por 4 bits, por lo tanto acepta una cantidad de 2^4 niveles de prioridad. La prioridad máxima es 0000, mientras que la mínima es 1111.

Tabla A.3: Prioridades de mensajes

Priority Field	Prioridad	
0000	0	Alta prioridad
0001	1	...
...
1111	15	Baja prioridad

A.10.3. Node-ID

Este indica el ID del nodo que envía el mensaje. Este campo está compuesto por 7 bits, lo cual posibilita la existencia de hasta $2^7 = 128$ nodos conectados a la red.

A.10.4. Remote transmission request (RTR)

Este permite diferenciar un frame que sea remoto o un mensaje “normal”. Un frame remoto es aquel que se utiliza para enviar telemetría. Comúnmente es enviado por sensores y componentes. El bit dominante (0) indica que es un frame “normal” (contiene un mensaje o un evento), un bit recesivo (1) indica que es un frame remoto.

A.10.5. Type of Message (TOM)

Este campo es de 1 bit. Diferencia frames de mensajes con frames de eventos. Si el bit es dominante (0) significa que se trata de un frame de evento. Si el bit es recesivo (1) significa que se trata de un frame de mensaje. Debe tenerse en cuenta que si el bit RTR es recesivo (1), el bit TOM obligatoriamente debe ser recesivo (1) indicando que el frame trae consigo un mensaje con los datos de telemetría.

A.10.6. Bit reservado

Bit reservado.

A.10.7. Data length

Indica el tamaño del campo de datos (*Data Field*). Este campo esta compuesto por 6 bits.

A.10.8. Data field

En este campo se agregan los datos del mensaje. El primer byte del *Data Field* es el código de la función y/o evento. Los bytes subsiguientes pueden ser los argumentos y/o datos de la función o evento que se está enviando en el frame. Se pueden enviar solo una función a la vez, y hasta 7 datos o argumentos.

A.10.9. CRC

En este campo se almacena el CRC del frame. Para calcular el CRC se utiliza el polinomio de CAN.

$$x^{15} + x^{14} + x^{10} + x^8 + x^7 + x^4 + x^3 + 1$$

El algoritmo para resolver el CRC se muestra en A.1.

Listing A.1: Función para resolver CRC

```
uint16_t crc_c(uint8_t data , uint16_t crc){
uint8_t i;

crc ^= (uint16_t)data << 7;
for (i = 0; i < 8; i++){
    crc <<= 1;
    if(crc & 0x8000){
        crc ^= 0xc599;
    }
}
return crc & 0x7fff;
}
```

A.10.10. CRC delimiter

Este es un bit recesivo (1) para indicar que el CRC Field finalizó.

A.10.11. ACK

El transmisor debe poner este bit en recesivo (1), y el receptor puede responder el mensaje colocando este bit en dominante(0).

A.10.12. ACK delimiter

Este bit indica el final del ACK. Este es un bit recesivo (1).

A.10.13. End-of-Frame (EOF)

Indica el final del frame de datos. Debe ser recesivo, es decir, los 7 bits en 1.

A.11. High Application Layer CANae

La denominada *High Application Layer* tiene una función más del lado de la gestión de nodos y tareas. En esta capa se desarrollan los algoritmos necesarios para realizar el ruteo y la reconfiguración de la red ante fallas, por lo tanto en la *High Application Layer* se debe implementar la FDIR. El protocolo no define ningún algoritmo de ruteo, por lo que queda para el usuario la definición de los algoritmos. Se recomienda desarrollar 2 algoritmos, el primario y el secundario. Para aumentar la tolerancia a fallas se pueden desarrollar más algoritmos secundarios, y el switch entre algoritmos, puede depender de medidas de performance.

En las Figuras [A.23] y [A.24] se pueden observar los diagramas de bloques y diagramas de bloques internos para el *High Application Layer* de CANae.

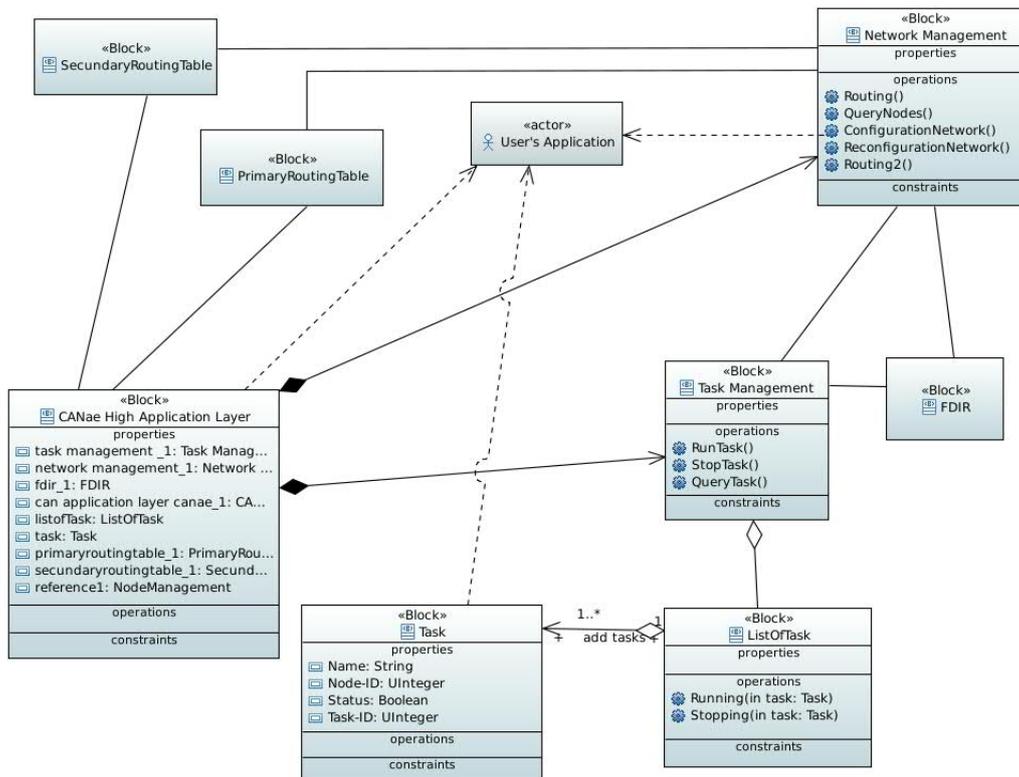


Figura A.23: Diagrama de bloques del High Application Layer de CANae

Esta capa está compuesta por varias entidades, las cuales tienen solo una instancia por entidad.

En la *High Application layer* se encuentra la entidad más importante, la cual es el FDIR. El desarrollo de esta entidad se encuentra fuera del alcance de esta versión (0.1 Alpha) del protocolo.

A.11.1. Network Management

Esta entidad se encarga de la gestión de la red desde el punto de vista del nodo. La entidad utiliza el *NodeTableConfiguration* del *CANae Application Layer*.

Los servicios brindados por el *Network Management* son los siguientes:

- **Routing()**: este servicio lleva a cabo el ruteo de los mensajes. Este servicio consulta la tabla *NodeTableConfiguration*.

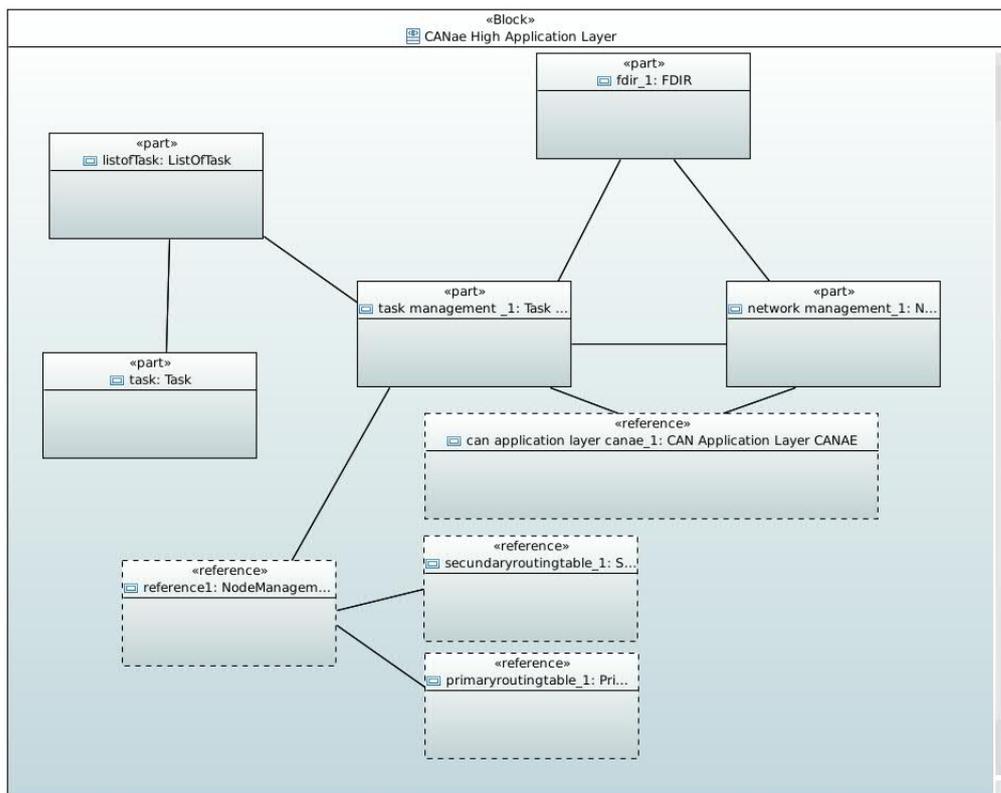


Figura A.24: Diagrama de bloques internos del High Application Layer de CANae

- **QueryNodes()**: este servicio hace consulta de los nodos conectados. Este puede ser utilizado por el servicio Routing()
- **ConfigurationNetwork()**: este servicio se lleva a cabo solo una vez, durante el encendido y configuración del nodo.
- **ReconfigurationNetwork()**: este servicio lleva a cabo la reconfiguración de la red. Este servicio es activado por el FDIR una vez que detecta algún error.
- **Routing2()**: Este servicio se lleva a cabo cuando se detecta un error en la red. Lo activa el FDIR.

El comportamiento del *Network Management* es sencillo, cuando se enciende por primera vez el nodo, de manera automática, la aplicación de usuario debe llamar el servicio *ConfigurationNetwork()* al *Network Management* esta entidad se comunica con el *Node Management* para que se genere la “Tabla Primaria de Ruteo”. Una vez que esta tabla se encuentre generada (tal como se hace referencia en A.9.1), la aplicación de usuario puede enviar señales de *QueryNodes()* y *Routing()*. Esto se puede ver visualmente en la Figura A.25.

En caso de errores, esto debe ser captado por el FDIR del sistema. El comportamiento del mismo se muestra en la Figura A.26

A.11.2. Task Management

Esta entidad se encarga de la gestión de las tareas. Debe aclararse que esta es una abstracción, y no depende del Sistema Operativo que se utilice para la codificación del sistema. El objetivo de esta entidad es mantener la distribución de las tareas entre los nodos de la red. Esto permitiría que cuando deja de funcionar (detectado por

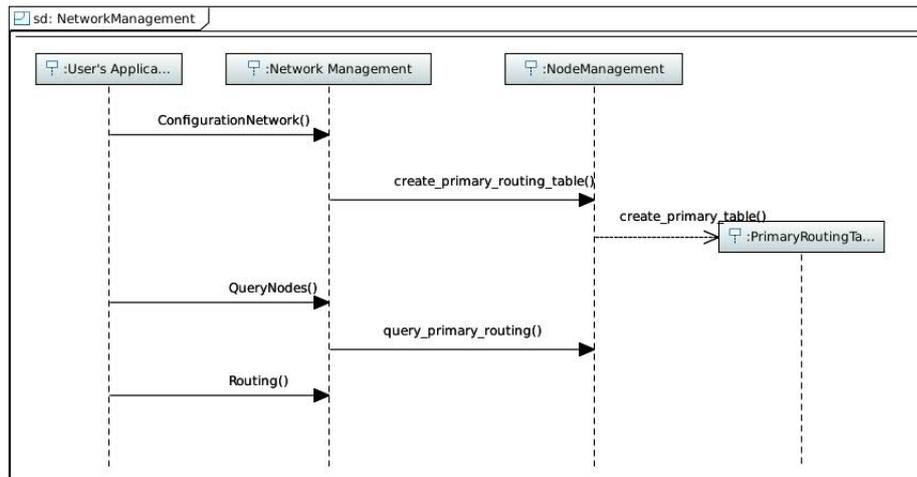


Figura A.25: Diagrama de interacción del comportamiento del *Network Management*

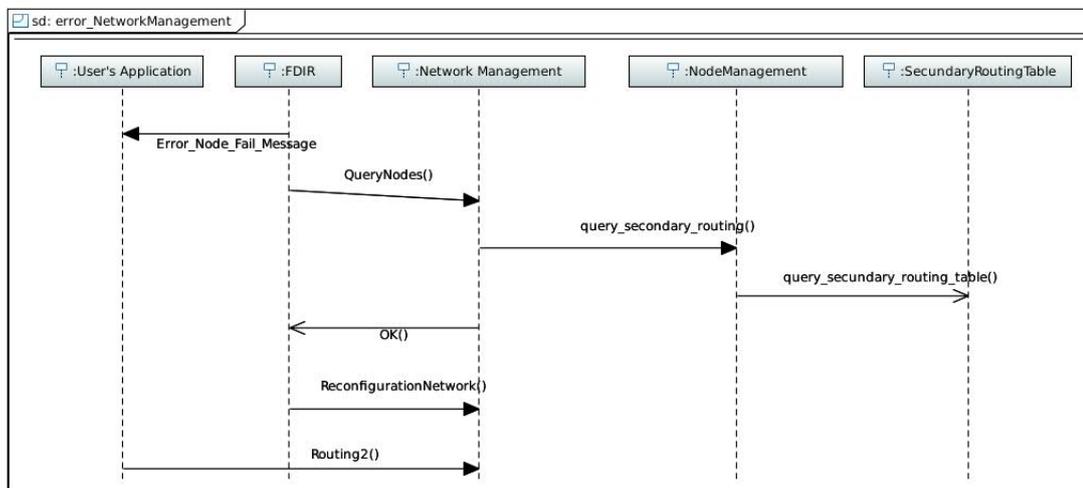


Figura A.26: Diagrama de interacción del comportamiento del *Network Management* en caso de errores

el FDIR) uno de los nodos, el *Task Management* gestiona las tareas que se desarrollan en los nodos. Para ellos se hace uso de la “Lista de Tareas” (*ListOfTask*).

Los servicios que ofrece el *Task Management* son los siguientes:

- **RunTask()**: este servicio activa las tareas para ejecutarse en el nodo.
- **StopTask()**: este servicio detiene las tareas que se ejecutan en el nodo.
- **QueryTask()**; este servicio hace consulta de las tareas que existe la “lista de Tareas”.

El comportamiento de esta entidad es sencilla. La aplicación de usuario le envía una señal de *QueryTasks()* con el motivo de conocer todas las tareas que se están ejecutando en la red. El *Task Management* lleva a cabo un *query_tasks_in_network()* al *Network Management* el cual le provee la información. La primera comprobación que debe realizar la aplicación de usuario es que la tarea no exista en otro nodo ejecutándose. Luego de esta comprobación, ya tiene permitido llevar a cabo el *RunTask()*. Por último, el *Task Management* manda el mensaje *Running(Task task)* al *ListOfTasks* (ver A.11.3). El comportamiento de la entidad se la puede observar en la Figura A.27.

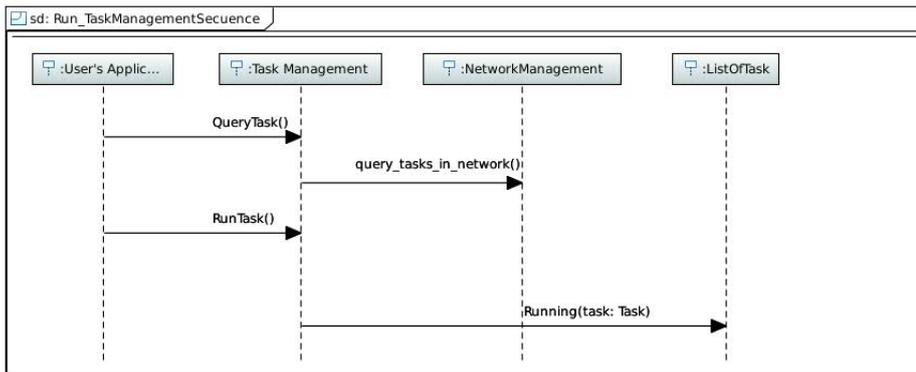


Figura A.27: Diagrama de secuencia de la ejecución de tareas del Task Management

Por otro lado cuando la aplicación de usuario necesita detener una tarea, o debido a que el FDIR detecto una falla y necesita detener la ejecución de una tarea, el proceso que debe llevar el *Task Management* es el siguiente, en primer hace un *QueryTasks()* para conocer las tareas que se están ejecutando y sus estados. Luego manda un señal de *StopTask()* al *Task Management*. Esta, envía mensajes de *query_tasks_in_network()*, y luego el *Stopping(Task task)* para detener la ejecución del proceso. En la Figura A.28

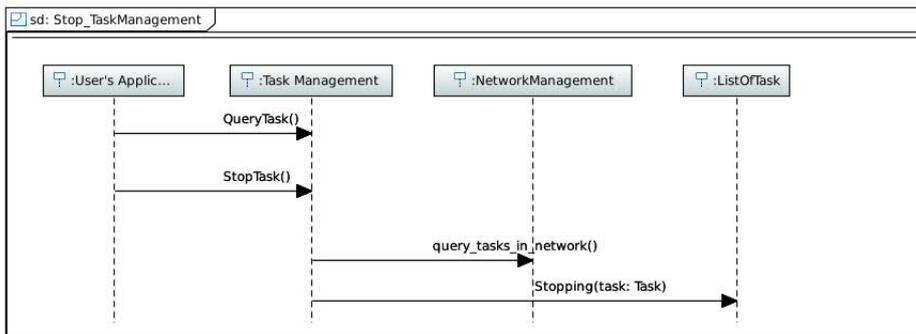


Figura A.28: Diagrama de secuencia para detener procesos por parte del Task Management

A.11.3. List of Tasks

Este es un objeto que mantiene una lista de las tareas que se están definidas en el sistema. Estas tareas pueden o no estar ejecutándose. Esta lista tiene un instancia a todas las tareas.

Los servicios que se ejecutan en esta entidad son los siguientes:

- **Running(Task task):** este servicio ejecuta una tarea.
 - **Task task:** tarea a ejecutar.
- **Stopping(Task task):** este servicio detiene una tarea.
 - **Task task:** tarea a ejecutar.

A.11.4. Task

Este es un objeto que representa la abstracción de las tareas. Las propiedades de esta entidad son las siguientes:

- **String Name:** es el nombre de la tarea.
- **UInteger Task-ID:** es el ID de la tarea.
- **UInteger Node-ID:** es el ID del nodo donde se está ejecutando la tarea.
- **Boolean Status:** es el status de la tarea {RUNNING, STOPED}

A.12. Arquitectura completa del CANae

CANae nace en el principio de la ingeniería en sistemas basada en modelo utilizando el lenguaje de modelado SysML para su diseño y desarrollo. Debido a la complejidad de los sistemas espaciales, a la necesidad de lograr un entendimiento completo por parte de los desarrolladores, y para lograr una correcta interpretación del protocolo se decidió basarse en este principio. En esta sección se pretende mostrar solo con motivo informativo la arquitectura estática completa del protocolo, para demostrar la complejidad del mismo, pero a la vez lo sencillo que resulta su interpretación, con conocimientos básicos de SysML.

En esta se puede observar como interactúan las diferentes entidades, objetos y eventos que se desarrollaron a lo largo de este documento.

En la Figura A.29 se observa la arquitectura de la capa de aplicación de CANae denominada: **CAN Application Layer CANae**. La cual como se mencionó con anterioridad se encuentra basada en la capa de aplicación de diferentes protocolos que existen para las redes CAN.

A.13. Estados de nodos y red

El protocolo de comunicación CANae está diseñado para ser sencillo en su uso, por lo tanto, los nodos como la red pasan por estados de funcionamiento que son sencillos y atómicos. En esta versión se pretende desarrollar las bases para los futuros avances sobre el mismo.

En primer lugar el ciclo de vida de un nodo comienza cuando es alimentando. El nodo empieza en un estado *Stand-by* hasta que se inicializa internamente. Luego pasa a un estado de *Waiting Monitor*, donde espera las ordenes de un *Monitor Node*. Luego cuando recibe la señales del monitor comienza a configurarse internamente (ver: A.7 y A.6). Luego de esto pasa a un estado *Working*. Aquí pueden ocurrir dos situaciones. En primer lugar, el nodo se apaga intencionalmente. La segunda situación es que el nodo falla y es detectado por el FDIR, entonces pasa a un estado *Failure*.

Por otro lado en el entorno de la red CANae, el ciclo de vida de la red comienza con la alimentación de los nodos. En primer lugar entra en un estado de *Init Nodes* donde todos los nodos de la red se encuentran desarrollando su ciclo de vida, tal como se indicó anteriormente. Luego, la red pasa a un estado *Stand-By* para que tiempo después empiece el estado de *Creating Network*. Dentro de este estado, se dan diferentes estados, se necesita que todos los nodos se encuentren coordinados. Estos se detallan en A.6. Luego de su creación toda la red se encuentra en estado *Ready* donde se lleva a cabo su normal funcionamiento.

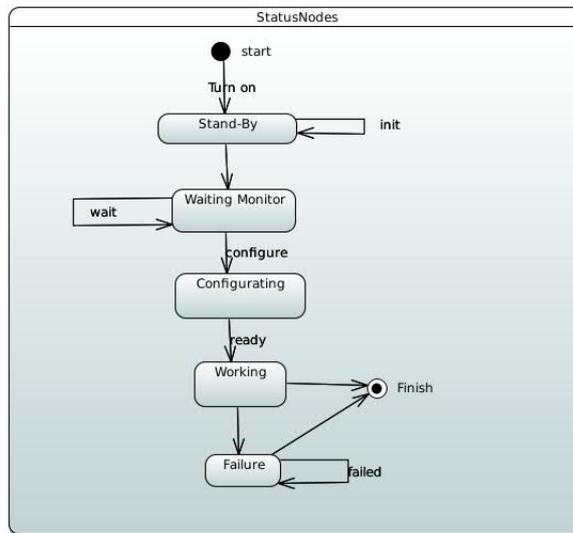


Figura A.30: Definición de la entidad NMT

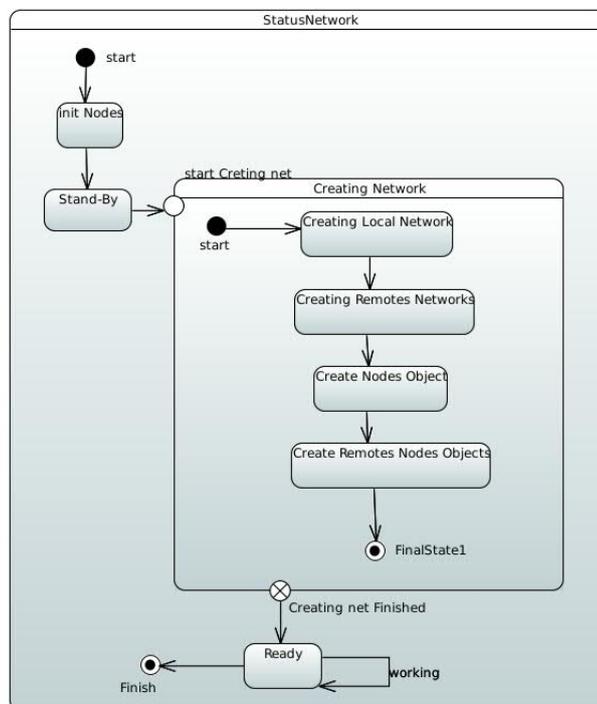


Figura A.31: Definición de la entidad NMT

Especificación de Casos de Uso

B.1. Enviar Mensaje

Use Case ID	El ID del Caso de uso es ARQ001
Historia del Caso de Uso	
Creado por	Arias Emmanuel
Fecha de Creación	12 de Junio del 2017
Fecha de la última actualización	N/A
Definición del Caso de Uso	
Actor	<ul style="list-style-type: none"> ▪ <i>UserApplication</i>. Este es la aplicación del usuario que se encuentra ejecutándose en el nodo.
Descripción	Este caso de uso recibe los mensajes que la aplicación de usuario necesita enviar por la red hacia otro nodo en el sistema.
Precondición	<ul style="list-style-type: none"> ▪ El nodo debe estar funcional y conectado a la red. ▪ El protocolo CANae debe haber finalizado su inicialización. ▪ El nodo destino tiene que estar funcional.
Postcondición	<ul style="list-style-type: none"> ▪ Debe asegurarse el correcto envío del mensaje.
Prioridad	Prioridad media.

Curso normal de eventos	El caso de uso recibe el mensaje proveniente de la aplicación del usuario. En el siguiente paso prepara el mensaje siguiendo los lineamientos del protocolo CANae. Por último lo envía a las capas inferiores para ser enviado a través del protocolo CAN.
Cursos alternativos	N/A
Excepciones	En caso de producirse un error en el caso de uso, se lleva a cabo una excepción, para avisar a la aplicación de usuario que se produjo un error.
Includes	N/A
Notas y problemas	N/A

B.2. Recibir Mensaje

Use Case ID	El ID del Caso de uso es ARQ002
Historia del Caso de Uso	
Creado por	Arias Emmanuel
Fecha de Creación	12 de Junio del 2017
Fecha de la última actualización	N/A
Definición del Caso de Uso	
Actor	<ul style="list-style-type: none"> ▪ <i>UserApplication</i>. Este la aplicación del usuario que se encuentra ejecutándose en el nodo.
Descripción	Este caso de uso recibe los mensajes proveniente de las capas inferiores de CAN y luego es enviado a la aplicación del usuario.
Precondición	<ul style="list-style-type: none"> ▪ El mensaje tiene que llegar sin errores. ▪ El nodo debe estar conectado y funcional en la red. ▪ El nodo debe estar preparado (no ocupado) para recibir el mensaje.
Postcondición	N/A
Prioridad	Prioridad media.
Curso normal de eventos	N/A
Cursos alternativos	N/A
Excepciones	N/A
Includes	N/A
Notas y problemas	N/A

B.3. Procesar Mensaje

Use Case ID	El ID del Caso de uso es ARQ003
Historia del Caso de Uso	
Creado por	Arias Emmanuel
Fecha de Creación	12 de Junio del 2017
Fecha de la última actualización	N/A
Definición del Caso de Uso	
Actor	N/A
Descripción	Este caso de uso se encarga de recibir los mensajes desde la aplicación de usuario y de las capas inferiores de CAN, para realizar un procesamiento. Este procesamiento incluye empaquetado y desempaqueado, y verificación de la ocurrencia de errores, clasificación de mensajes que contienen datos o eventos.
Precondición	<ul style="list-style-type: none"> ▪ El nodo debe estar conectado y funcional en la arquitectura.
Postcondición	N/A
Prioridad	Prioridad alta.
Curso normal de eventos	El caso de uso recibe el mensaje proveniente de las capas inferiores del CAN. Se suponen que llegan sin errores. Se desempaqueta, se clasifica (si es datos o evento) y luego se envía a la aplicación de usuario.
Cursos alternativos	N/A
Excepciones	N/A
Includes	Ver Rutear Mensajes. B.4
Notas y problemas	N/A

B.4. Rutear Mensaje

Use Case ID	El ID del Caso de uso es ARQ003.1
Historia del Caso de Uso	
Creado por	Arias Emmanuel
Fecha de Creación	12 de Junio del 2017
Fecha de la última actualización	N/A
Definición del Caso de Uso	
Actor	N/A
Descripción	Este es el encargado de rutear los mensajes. Debe indicar que nodos intermedios debe pasar el mensaje hasta llegar al nodo destino.

Precondición	<ul style="list-style-type: none"> ■ El nodo debe estar conectado y funcional en la arquitectura. ■ El mensaje debe estar clasificado y empaquetado. ■ Se debe conocer el nodo destino.
Postcondición	N/A
Prioridad	Prioridad alta.
Curso normal de eventos	Este caso de uso una vez que se clasifica y empaqueta el mensaje, debe consultar al Caso de Uso <i>Gestión de Red</i> para verificar que nodos intermedios debe pasar el mensaje hasta llegar al nodo destino.
Cursos alternativos	N/A
Excepciones	N/A
Includes	N/A
Notas y problemas	N/A

B.5. Gestionar Red

Use Case ID	El ID del Caso de uso es ARQ004
Historia del Caso de Uso	
Creado por	Arias Emmanuel
Fecha de Creación	12 de Junio del 2017
Fecha de la última actualización	N/A
Definición del Caso de Uso	
Actor	N/A
Descripción	Este caso de uso se encarga de gestionar la red. Se vale de las bondades ofrecidas por CANae para conocer cuáles son los nodos conectados a la red y la distancia que existe entre el nodo objetivo y el destino.
Precondición	<ul style="list-style-type: none"> ■ El nodo debe estar conectado y funcional en la arquitectura. ■ El protocolo CANae debe haber finalizado su inicialización. ■ El protocolo CANae tiene que haber realizar la tabla de ruteo.
Postcondición	N/A
Prioridad	Prioridad alta.

Curso normal de eventos	El caso de uso utiliza las bondades del protocolo CANae para conocer cuáles son los nodos conectados a la red y su distancia. Este caso de uso es la base para que se realice una correcta gestión de tareas a realizar en el sistema. Luego con esto se pueden dividir las tareas que se llevarán a cabo en los diferentes nodos de la red.
Cursos alternativos	N/A
Excepciones	N/A
Includes	Ver Gestionar Tareas B.6
Notas y problemas	N/A

B.6. Gestionar Tareas

Use Case ID	El ID del Caso de uso es ARQ004.1
Historia del Caso de Uso	
Creado por	Arias Emmanuel
Fecha de Creación	12 de Junio del 2017
Fecha de la última actualización	N/A
Definición del Caso de Uso	
Actor	N/A
Descripción	Este caso de uso se encarga de gestionar las tareas que se deben ejecutar en la red. Para ello debe implementarse un algoritmo que asegure la correcta comunicación con los nodos para lograr la gestión de tareas del sistema.
Precondición	<ul style="list-style-type: none"> ■ El nodo debe estar conectado y funcional en la arquitectura. ■ El protocolo CANae debe haber finalizado su inicialización. ■ El protocolo CANae tiene que haber realizado la tabla de ruteo.
Postcondición	N/A
Prioridad	Prioridad alta.
Curso normal de eventos	En primer lugar debe consultar la tabla de ruteo para conocer la posición de los nodos. Con ello y la lista de tareas del sistema, lleva a cabo un algoritmo de gestión de tareas, donde se negocia con los demás nodos qué tareas se van a llevar a cabo en los diferentes nodos.
Cursos alternativos	N/A
Excepciones	N/A

Incluye	Ver Dividir Tareas B.7
Notas y problemas	N/A

B.7. Dividir Tareas

Use Case ID	El ID del Caso de uso es ARQ004.1.1
Historia del Caso de Uso	
Creado por	Arias Emmanuel
Fecha de Creación	12 de Junio del 2017
Fecha de la última actualización	N/A
Definición del Caso de Uso	
Actor	N/A
Descripción	Este caso de uso se encarga de realizar al división de las tareas, una vez que se ha llevado a cabo la negociación.
Precondición	<ul style="list-style-type: none"> ▪ El nodo debe estar conectado y funcional en la arquitectura. ▪ El protocolo CANae debe haber finalizado su inicialización. ▪ El protocolo CANae tiene que haber realizar la tabla de ruteo.
Postcondición	<ul style="list-style-type: none"> ▪ Se actualiza la lista de tareas en los nodos.
Prioridad	Prioridad alta.
Curso normal de eventos	Luego de recibir el resultado de la negociación, actualiza la lista de tareas que se llevan a cabo en los nodos. Informa a los demás nodos.
Cursos alternativos	N/A
Excepciones	N/A
Incluye	N/A
Notas y problemas	N/A

B.8. FDIR

Use Case ID	El ID del Caso de uso es ARQ005
Historia del Caso de Uso	
Creado por	Arias Emmanuel
Fecha de Creación	12 de Junio del 2017

Fecha de la última actualización	N/A
Definición del Caso de Uso	
Actor	N/A
Descripción	Este caso de uso se encarga de la detección, aislamiento y recuperación de errores que se produzcan tanto en el nodo como en la red.
Precondición	<ul style="list-style-type: none"> ■ El nodo debe estar conectado y funcional en la arquitectura. ■ El protocolo CANae debe haber finalizado su inicialización. ■ El protocolo CANae tiene que haber realizado la tabla de ruteo.
Postcondición	<ul style="list-style-type: none"> ■ Se toma una medida para lograr aislar y recuperar la arquitectura del error que se ha producido.
Prioridad	Prioridad alta.
Curso normal de eventos	Este caso de uso debe observar como se encuentra la red en cada momento. Para ello hace uso de las bondades que le puede brindar el caso de uso <i>Gestionar Red</i> . Cuando detecta que se ha producido una falla en el nodo o en la red, lanza una excepción para manejar el problema.
Cursos alternativos	Ver <i>Reconfigurar Arquitectura B.9</i>
Excepciones	N/A
Incluye	N/A
Notas y problemas	No se llega a realizar un diseño y desarrollo de este caso de uso en esta instancia del trabajo de tesis.

B.9. Reconfigurar Arquitectura

Use Case ID	El ID del Caso de uso es ARQ006
Historia del Caso de Uso	
Creado por	Arias Emmanuel
Fecha de Creación	12 de Junio del 2017
Fecha de la última actualización	N/A
Definición del Caso de Uso	
Actor	N/A

Descripción	Este caso de uso se encarga de la detección, aislamiento y recuperación de errores que se produzcan tanto en el nodo como en la red.
Precondición	<ul style="list-style-type: none"> ■ El nodo debe estar conectado y funcional en la arquitectura. ■ El protocolo CANae debe haber finalizado su inicialización. ■ El protocolo CANae tiene que haber realizado la tabla de ruteo.
Postcondición	<ul style="list-style-type: none"> ■ Se toma una medida para lograr aislar y recuperar la arquitectura del error que se ha producido.
Prioridad	Prioridad alta.
Curso normal de eventos	Este caso de uso debe observar como se encuentra la red en cada momento. Para ello hace uso de las bondades que le puede brindar el caso de uso <i>Gestionar Red</i> . Cuando detecta que se ha producido una falla en el nodo o en la red, lanza una excepción para manejar el problema.
Cursos alternativos	Ver <i>Reconfigurar Arquitectura B.9</i>
Excepciones	N/A
Includes	N/A
Notas y problemas	No se llega a realizar un diseño y desarrollo de este caso de uso en esta instancia del trabajo de tesis.

Antecedentes

El objetivo de este apéndice es que el lector conozca el crecimiento de las investigaciones llevadas a cabo en el área de estudio de la presente tesis. Esto ayudará a comprender y respaldar las motivaciones del presente trabajo. Así, se podrá demostrar que existe una gran cantidad de investigadores (los cuales se incrementan con el paso de los años), que se enfocan en la problemática a la cuales se enfrenta esta tesis. Para ello, se utilizó la base de datos Scopus¹, la cual es una base de datos que contiene abstracts, citas, y artículos científicos.

En primer lugar se realizó la búsqueda de las palabras claves “COTS satellites”, obteniendo los resultados que se muestran en la Figura C.1. Se observa que existe una tendencia creciente en las investigaciones que contienen estas *key words*. Se encontró una cantidad total de 2047 artículos científicos.

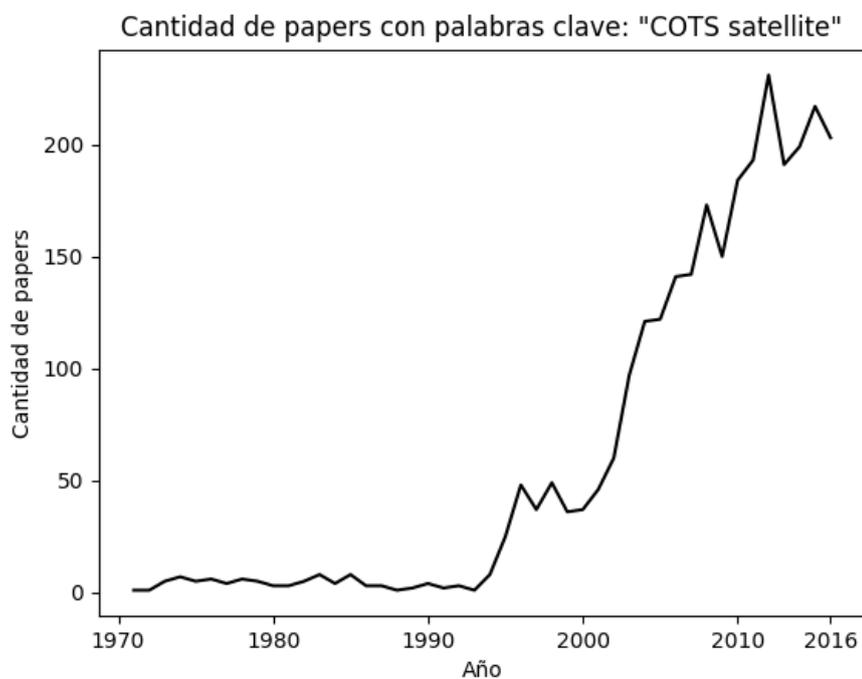


Figura C.1: Cantidad de investigaciones con palabra clave “COTS satellite”

¹www.scopus.com

Por otro lado, se llevó a cabo una búsqueda de las mismas palabras claves, pero esta vez, la consulta se realizó para conocer los principales países involucrados en estas investigaciones. De esto, se obtiene el gráfico de la Figura C.2. Como se observa el país con más investigaciones es Estados Unidos.

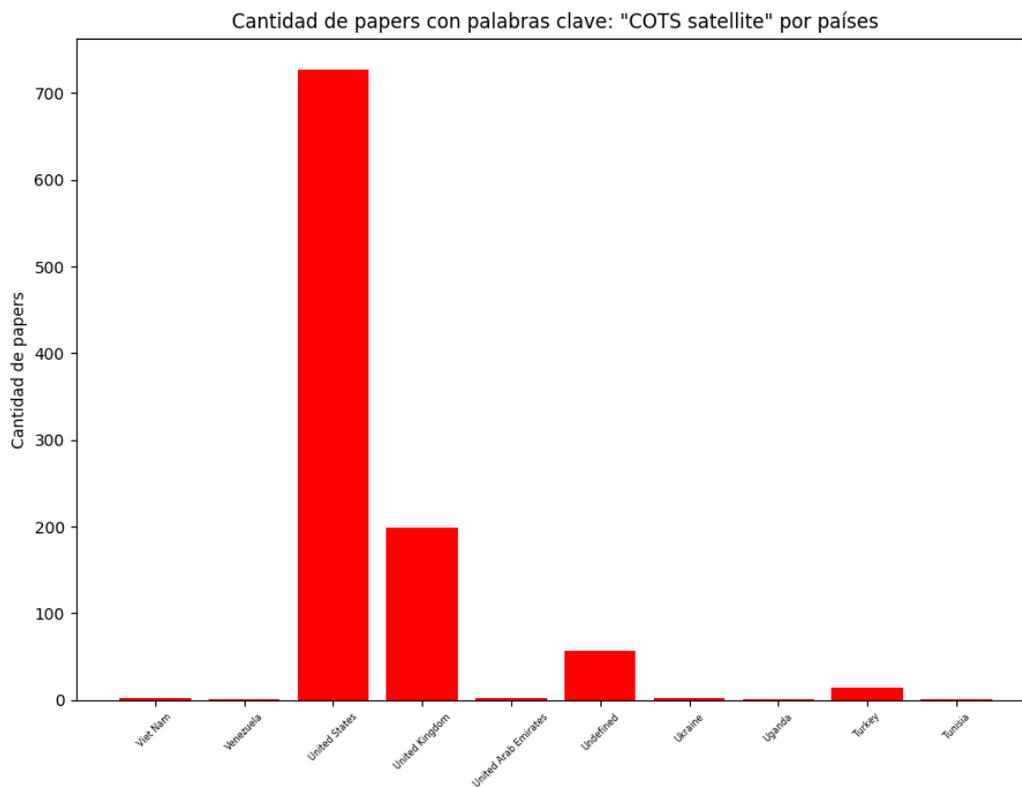


Figura C.2: Cantidad de investigaciones con palabra clave “COTS satellite” por países

De la misma manera se realizó una consulta de las palabras claves: “Fault Tolerance Satellite”, para conocer tanto la cantidad de artículos existentes (Figura C.3). Nuevamente, se puede observar un creciente interés en llevar a cabo investigaciones con respecto a las áreas de interés de este trabajo de tesis.

Luego se realizó una búsqueda con las siguientes *key words*: “COTS Fault Tolerance”. Como resultado se obtuvo que existe una lateralización en la tendencia de las investigaciones en estas áreas como se puede observar en la Figura C.4. Nuevamente, el país con más investigaciones en el área sigue siendo Estados Unidos C.5.

Por último se llevó a cabo una consulta con las palabras claves: “Fault Tolerance Satellite Architecture”. Los resultados se pueden observar en las Figuras C.6 y C.7. Como en las consultas anteriores, se observa un creciente interés por el estudio de esta área de trabajo y el país como mayor publicación, sigue siendo Estados Unidos.

De estas búsquedas sencillas, se puede llegar a la conclusión de que en general, existe una tendencia creciente en investigaciones en el área que compete este trabajo de tesis. Además, se puede obtener como conclusión de que existe una mayor cantidad de investigaciones en la dirección de “COTS” y “Satellites”. Esto se puede justificar con la existencia de una necesidad de reducir costos a la hora de la fabricación de satélites.

Por otro lado, lo que resulta interesante (o más bien preocupante) es que no hay una tendencia creciente, en las investigaciones sobre “COTS” y “Fault Tolerance” que acompañe a lo dicho en el párrafo anterior. Esto pue-



Figura C.3: Cantidad de investigaciones con palabra clave "Fault Tolerance Satellite"

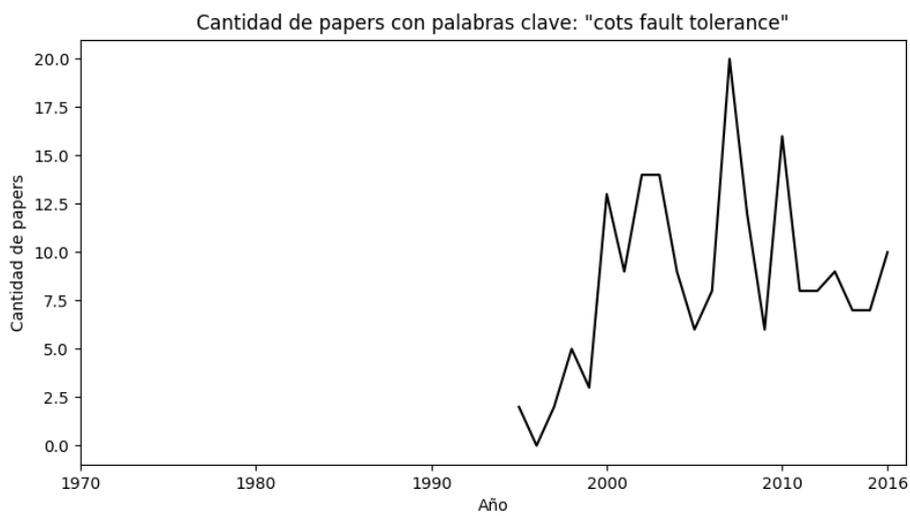


Figura C.4: Cantidad de investigaciones con palabra clave "COTS Fault Tolerance"

de traer consecuencias negativas para aquellos que traten de desarrollar vehículos satelitales con componentes COTS, y no cuenten con estrategias de tolerancia a fallas.

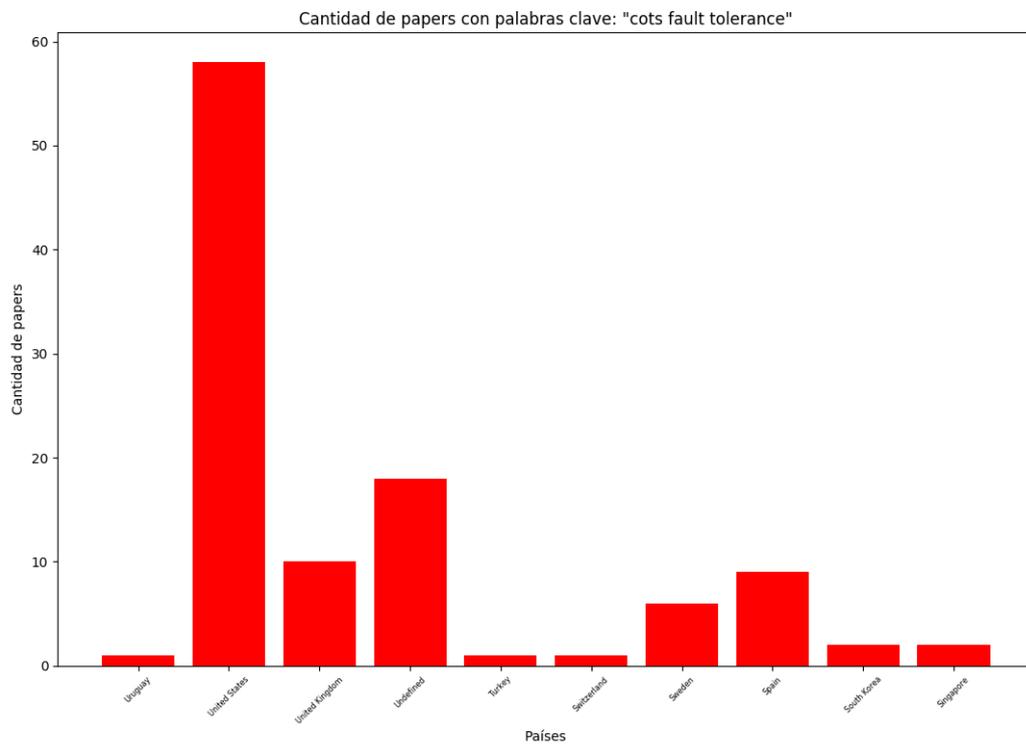


Figura C.5: Cantidad de investigaciones con palabra clave “COTS Fault Tolerance” por países

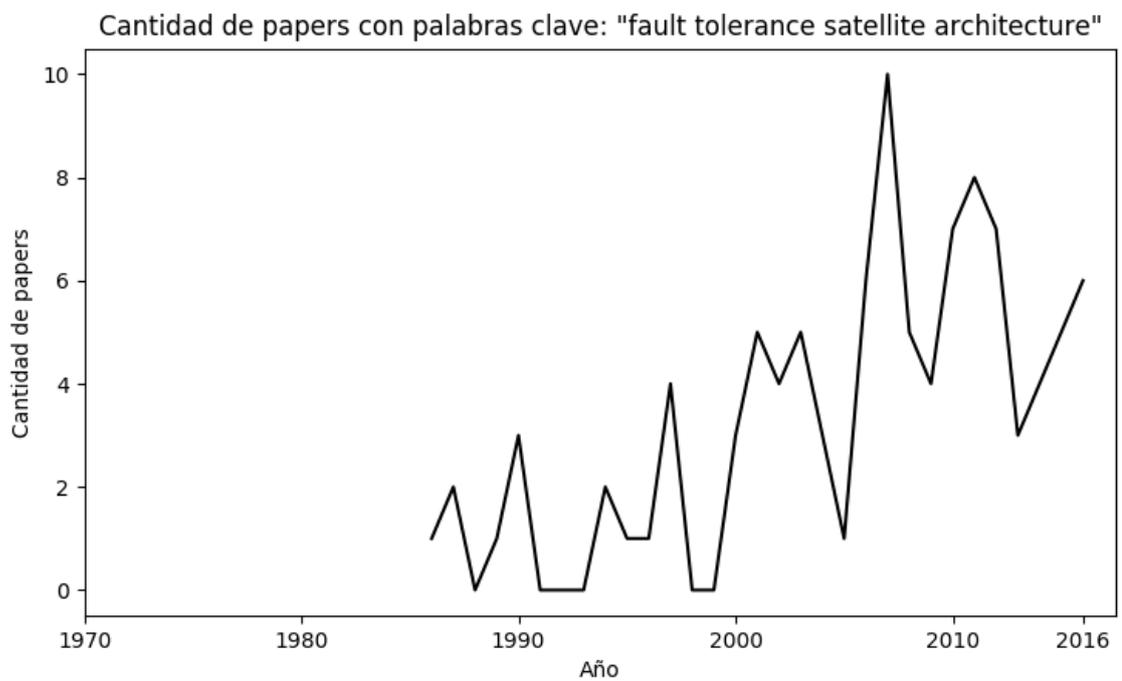


Figura C.6: Cantidad de investigaciones con palabra clave “Fault Tolerance Satellite Architecture”

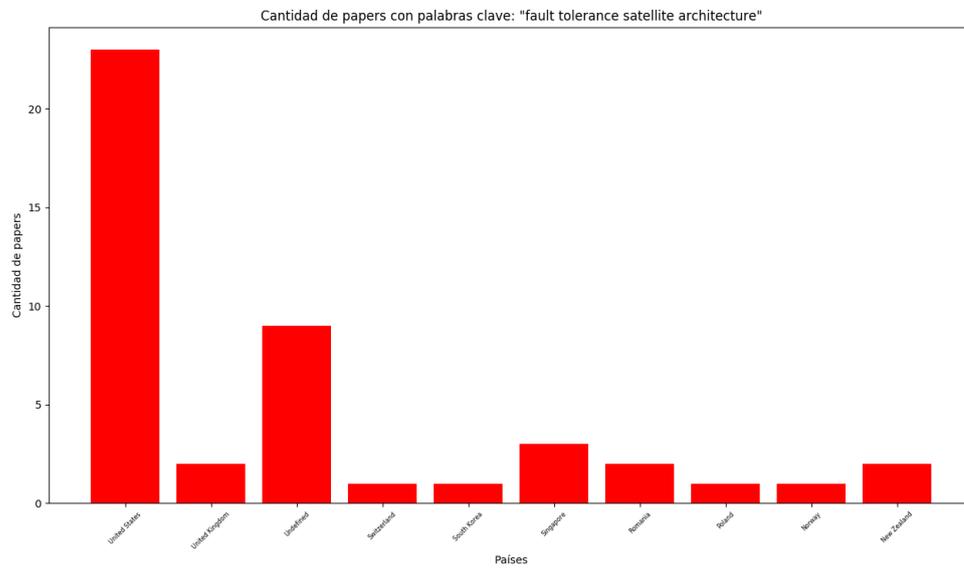


Figura C.7: Cantidad de investigaciones con palabra clave "Fault Tolerance Satellite Architecture" por países