

# Propuesta de Lineamientos Técnicos para Proporcionar Alta Disponibilidad de Servicio en el contexto de una NGSDP

Angélica María  
Burbano Cifuentes  
Universidad del Cauca  
Grupo de Ingeniería  
Telemática  
Popayán, Colombia  
(57 2) 8209800 ext. 2127

Francisco Javier  
Calero Valenzuela  
Universidad del Cauca  
Grupo de Ingeniería  
Telemática  
Popayán, Colombia  
(57 2) 8209800 ext. 2127

Francisco Orlando  
Martínez Pabón  
Universidad del Cauca  
Grupo de Ingeniería  
Telemática  
Popayán, Colombia  
(57 2) 8209800 ext. 2127

Oscar Mauricio  
Caicedo Rendón  
Universidad del Cauca  
Grupo de Ingeniería  
Telemática  
Popayán, Colombia  
(57 2) 8209800 ext. 2127

amburbano@unicauca.edu.co fcalero@unicauca.edu.co fomarti@unicauca.edu.co omcaicedo@unicauca.edu.co

## ABSTRACT

Con el fin de integrar a terceros y alcanzar la convergencia entre el mundo de las Telecomunicaciones y el mundo Web, fueron implantadas en el modelo de negocio de los operadores las SDP (Service Delivery Platform); éstas permiten la eficiente creación, despliegue, ejecución, orquestación y gestión de servicios, generando un puente entre el mundo Telco y el mundo de las Tecnologías de la Información (TI). Actualmente, las NGSDP constituyen la tercera generación de SDP, incorporando principios SOA (Service Oriented Architecture) relacionados con la orquestación y gestión de servicios, lo cual le permite a los operadores y terceros, desarrollar servicios y desplegarlos sobre diferentes arquitecturas. Sin embargo, uno de los grandes retos en este contexto, es mantener el porcentaje de disponibilidad de los servicios tradicionales de telecomunicaciones en los nuevos servicios convergentes. El presente artículo presenta una propuesta para mantener niveles de alta disponibilidad de servicio en el contexto de una NGSDP, a través del planteamiento de una serie de criterios y lineamientos técnicos relacionados con este contexto.

## Categories and Subject Descriptors

C.4 [Performance of systems]: Performance attributes, reliability, availability, and serviceability.

## General Terms

Documentation, Performance, Experimentation.

## Keywords

NGSDP, SDP, SOA, Alta disponibilidad.

## 1. INTRODUCCION

Con el fin de alcanzar la convergencia entre el mundo de las telecomunicaciones y el mundo Web, los operadores de telecomunicaciones han empezado a implantar las SDP[1] [2], las cuales facilitan la integración de terceros y permiten la interconexión entre las diferentes redes y tecnologías, así como también el rápido desarrollo, despliegue, orquestación y ejecución de servicios. En su recorrido evolutivo, la SDP ha tenido tres generaciones; en la primera generación este término fue utilizado para describir una arquitectura de servicio común que permitiera el desarrollo de contenido móvil y de servicios de mensajería. En la segunda, la SDP evolucionó para soportar servicios de voz, multimedia, localización, presencia, entre otros [3]. La SDP actual o de tercera generación es conocida como NGSDP o SDP 2.0 y

está construida bajo principios SOA, lo cual permite una eficiente integración, orquestación y gestión del ciclo de vida de los servicios, además de facilitar la migración de servicios desde las redes tradicionales hacia las redes All-IP y los estándares de la arquitectura IMS [4]. No obstante, el gran número de usuarios con los que cuentan los operadores de telecomunicaciones y las diferencias en los porcentajes de disponibilidad de los servicios de las telecomunicaciones y la Web, plantean como reto mantener los porcentajes de disponibilidad de los servicios tradicionales de telecomunicaciones en los nuevos servicios convergentes [5]. En este sentido, la definición de lineamientos técnicos que orienten a los operadores al respecto es fundamental. Aunque se han adelantado varios estudios relacionados con el contexto de las NGSDP, éstos no se han enfocado en la definición de directrices para garantizar alta disponibilidad de servicio [6]. El presente artículo presenta una propuesta para mantener niveles de alta disponibilidad de servicio en el contexto de una NGSDP, a través del planteamiento de una serie de criterios y lineamientos técnicos relacionados con este objetivo.

En este contexto, el contenido del artículo se ha organizado de la siguiente forma: en la sección II se muestran algunos trabajos relacionados, señalando sus aportes y diferencias. En la sección III se explica brevemente el nuevo concepto de NGSDP y su arquitectura. En la sección IV se analizan los conceptos de disponibilidad y alta disponibilidad en el entorno de las telecomunicaciones y se caracterizan algunos aspectos que un sistema altamente disponible debe cumplir. En la sección V se definen los criterios técnicos asociados a la alta disponibilidad y se proponen algunos lineamientos en el marco de dichos criterios, esto debido a que el trabajo se encuentra en construcción. En la sección VI se presentan algunas conclusiones y el trabajo a desarrollar en la presente investigación.

## 2. TRABAJOS RELACIONADOS

A continuación, se exponen tres de los trabajos más relevantes para el contexto de la presente investigación.

En [7], se propone una solución de alta disponibilidad para los servicios críticos de un centro de datos; esta solución se basa en una serie de fases y lineamientos, con las cuales se alcanza dicho objetivo. Sin embargo el trabajo es desarrollado para el contexto específico del centro de datos de la Universidad del Cauca. El trabajo aporta conceptos relacionados con la alta disponibilidad y un modelo para la construcción de lineamientos.

En [6], se proporcionan criterios técnicos para el suministro de servicios de valor agregado. No obstante dichos criterios son propuestos para el caso específico de una NGN (Next Generation Network) en el contexto Colombiano. El trabajo aporta: i) una base de conocimiento alrededor de la SDP y su función dentro de las redes NGN, ii) un modelo para la construcción del prototipo con el cual se verificara el presente trabajo.

En [3], se analizan las nuevas características y el modelo conceptual de la NGSDP. Sin embargo no se muestran las características con las que esta plataforma debe cumplir a fin de brindar servicios altamente disponibles.

### 3. NGSDP: PLATAFORMA PARA LA ENTREGA DE SERVICIOS EN NGN

La NGSDP, es una SDP construida bajo principios SOA e integrada con IMS y que proporciona un puente entre las redes de telecomunicaciones y la Web 2.0; su principal innovación es la capa de orquestación y gestión de servicio. El modelo conceptual de la NGSDP propuesto por el grupo Moriana (Figura 1) [8] tiene una estructura en capas, como se describe a continuación:



Figura 1. Arquitectura NGSDP.

**Capa de abstracción de red:** conformada por servicios habilitadores de bajo nivel que proveen acceso a las capacidades básicas de la red de telecomunicaciones y a los servicios subyacentes, como: control de llamada, mensajería móvil, gestión multimedia, localización, presencia, gestión de sesión, entre otros.

**Capa de creación y ejecución de servicio:** se compone de uno o más entornos SEE (Service Execution Environment), entre los cuales se pueden nombrar: SIP Servlets, Parlay, JAIN, etc. Esta capa también incluye la gestión de interacción del servicio y funcionalidades de service broker<sup>1</sup>.

**Capa de servicios de telecomunicación y habilitadores de servicio:** contiene servicios críticos del core como: VPN (Virtual Private Network), Pago por adelantado, IP Centrex o telefonía VoIP clase 5. Igualmente, permite la implementación de habilitadores de servicio de alto nivel como: administración de identidad, de contenido y de medios, distribución de llamadas, conferencia, etc.

**Capa de orquestación y gestión de servicio:** incorpora los principios de SOA y los servicios Web a la plataforma SDP. Esto incluye los ESB (Enterprise Service Bus) para la integración con los sistemas OSS/BSS (Operational Support System / Business Support Systems), así como también la integración al interior de la SDP. Adicionalmente, incluye características comunes de la

SDP como repositorios de servicios, de perfiles y gestión de la identidad compartida entre todos los servicios desplegados.

**Capa de exposición de servicio:** permite el acceso seguro y la gestión de terceros sobre los habilitadores de servicio y a las capacidades de bajo nivel de la red del operador.

### 4. ALTA DISPONIBILIDAD EN EL CONTEXTO DE UNA NGSDP

Como se mencionó en la sección I, en un futuro ambiente de convergencia entre las redes y servicios de telecomunicaciones con el mundo TI, la alta disponibilidad juega un papel de suma importancia, ya que los nuevos servicios convergentes deben aportar las experiencias del mundo Web con la disponibilidad del mundo de las telecomunicaciones, lo cual genera un importante desafío para el despliegue de servicios sobre una NGSDP. A continuación se describe el concepto de alta disponibilidad en este contexto.

#### 4.1 El concepto de Disponibilidad

La disponibilidad es un cálculo porcentual del tiempo en el cual la aplicación está realmente disponible para resolver las solicitudes de servicio en comparación con el tiempo de ejecución total disponible previsto. El cálculo formal de la disponibilidad está expresado de manera general mediante la ecuación básica (Ec.1), la cual relaciona el tiempo promedio entre fallas (MTBF-Mean Time Between Failure) y el tiempo medio para la reparación una vez ocurrida la falla (MTTR-Mean Time To Repair), donde A (Availability), representa la disponibilidad general[7].

$$A = \frac{MTBF}{(MTBF + MTTR)} \quad \text{Ec.1. Fórmula para el cálculo de disponibilidad.}$$

#### 4.2 El concepto de Alta Disponibilidad

La alta disponibilidad garantiza un grado absoluto de continua funcionalidad dentro de una ventana de tiempo. En la Tabla 1, se muestra el número de minutos o segundos estimados de tiempo de paradas del sistema (downtime) en relación con el número de minutos en 365 días al año[9].

Tabla 1. Porcentaje de disponibilidad en Sistemas de Telecomunicaciones

% Disponibilidad	Paradas en minutos	Paradas por año
90	52.560	36,5 días
99	5.256	4 días
99.9	525,6	8,8 horas
99.99	52,56	53 minutos
99.999	5,26	5,3 minutos
99.9999	0,53	32 segundos

Los tiempos de caídas y paradas planeados y no planeados casi nunca se presentan en las redes de telecomunicaciones, suministrando una disponibilidad cercana al 99.999%. Por otro lado, los periodos planeados y no planeados son muy comunes en el mundo TI e internet y éstos pueden durar varias horas, proporcionando una disponibilidad de servicio considerablemente menor a la de los servicios de telecomunicaciones. Con el objeto de mantener el porcentaje de disponibilidad de los servicios tradicionales en los nuevos servicios convergentes desplegados sobre las plataformas NGSDP, es necesario implementar las características de carrier grade en las NGN.

En este sentido, Open Cloud, define las características de carrier grade que se deben cumplir en las NGN, así [10]:

<sup>1</sup> Service Broker, resuelve interacciones de servicio y otros conflictos entre los servicios activados simultáneamente en la misma sesión. Componente clave de una SDP convergente.

- Ningún punto único de fallo.
- 99.999% de disponibilidad.
- Disponibilidad continua de servicios y de sesiones de servicio.
- Auto monitoreo y auto reparación.
- Protección de sobrecarga y gestión de la calidad de servicio.
- Baja latencia en la ejecución.
- Alto rendimiento (gran cantidad de llamadas).

## 5. LINEAMIENTOS TÉCNICOS PARA GARANTIZAR ALTA DISPONIBILIDAD DE SERVICIO EN UNA NGSDP

### 5.1 Definición de criterios técnicos asociados a la alta disponibilidad

La alta disponibilidad es caracterizada por diferentes empresas de telecomunicaciones y consultoras de tecnología mediante diferentes criterios. Estos criterios han sido definidos por cada una de las empresas de acuerdo a sus necesidades, ya que no existe un conjunto bien definido de criterios técnicos que permitan tener servicios altamente disponibles en un sistema.

Como punto de partida para caracterizar la alta disponibilidad, se tomaron algunos criterios de este tipo de organizaciones<sup>2</sup> [11] [12], para conformar una base inicial, la cual fue sometida posteriormente a una serie de filtros (Figura 2.); el resultado final es un conjunto de criterios, que para el contexto del presente trabajo, deben ser cumplidos en su totalidad por una NGSDP.

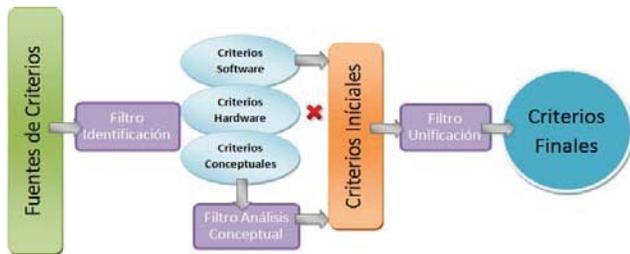


Figura 2. Filtrado de criterios

Las etapas de filtrado se describen a continuación:

**Filtro identificación:** clasifica los criterios en: hardware, software o conceptuales. Aquellos criterios que requieran capacidades hardware para su implementación, fueron descartados; esto debido a que actualmente, las fallas de los componentes hardware son más bajas que las producidas por los componentes software. Por lo tanto, es esencial diseñar sistemas software de mayor calidad.

**Filtro análisis conceptual:** del conjunto de criterios conceptuales se seleccionan los de mayor relevancia<sup>3</sup> para el modelo de negocio de los operadores de telecomunicaciones.

Como resultado de los filtros identificación y análisis conceptual se obtiene el conjunto de criterios iniciales.

**Filtro unificación:** se comparan los criterios del conjunto inicial con el fin de evitar la duplicación de los mismos, ya que puede existir el caso que uno o varios criterios sean definidos por más de

una empresa u organismo de estandarización. En caso de existir duplicación, éstos son agrupados.

Después de realizar este proceso, se obtuvieron los criterios relacionados con la alta disponibilidad en el contexto de una NGSDP. A continuación se realiza una breve descripción del alcance de cada uno de ellos:

**Balanceo de cargas:** aumenta el rendimiento de los sistemas, evitando los denominados cuellos de botella que se generan por la mala distribución de la carga de trabajo entre los equipos y redes.

**Escalabilidad:** capacidad del sistema para manejar el crecimiento continuo de trabajo de manera fluida o hacerse más grande sin perder calidad en los servicios ofrecidos.

**Diseño de red:** topologías y configuraciones de red bien definidas para asegurar que no exista ningún punto único de fallo.

**Rendimiento:** se refiere al desempeño en términos de tiempo de respuesta y el rendimiento de una red, equipo, servicio u otro.

**Contención contra fallas:** capacidad del sistema para reaccionar ante posibles fallas, de forma que el rendimiento del sistema no se vea afectado en caso de un siniestro.

**Construcción en redundancia:** duplicación de información de los componentes de un sistema a fin de incrementar su fiabilidad.

**Requisitos de tiempo de funcionamiento:** hace referencia a la cantidad total de tiempo en la cual el sistema está disponible para las aplicaciones de usuarios finales.

**Requerimientos de resiliencia:** conjunto de aplicaciones, datos y entornos de sistemas que es necesario conservar en caso de una parada del sistema de producción.

**Requerimientos de distancia:** la distancia entre los sistemas tiene ventajas, pero cuenta con varios límites físicos y prácticos. Entre más separados estén los sistemas, se añadirá más latencia.

**Cobertura de paradas:** se refiere a los tipos de paradas imprevistas que pueden existir en un sistema y cómo se pueden resolver para no afectar de forma considerable su disponibilidad.

**Clustering de Base de datos:** permite a dos o más equipos acceder concurrentemente a una sola base de datos compartida, apareciendo frente a las aplicaciones como una base de datos unificada; brindando beneficios de disponibilidad y escalabilidad.

**Confiabilidad:** se refiere a la probabilidad que un dispositivo o sistema lleve a cabo sus actividades prescritas sin fallas durante un tiempo determinado.

**Presupuesto:** cada solución de alta disponibilidad tiene un costo asociado y un presupuesto. Desde el punto de vista técnico, se puede obtener una disponibilidad constante con un tiempo de inactividad igual a cero, pero su costo podría ser muy elevado.

**Buenas prácticas:** procedimientos documentados para desarrollar y mantener una infraestructura de red robusta.

**Copia de Seguridad:** es la copia total o parcial de la información que puede utilizarse para restaurar el sistema original después de una eventual pérdida de información.

### 5.2 Definición de Lineamientos técnicos

A continuación se presentan algunos de los lineamientos técnicos definidos para alcanzar alta disponibilidad en el contexto de una NGSDP, de acuerdo a las necesidades de los operadores; para cada uno de los criterios, estas necesidades son determinadas

<sup>2</sup> Cisco, IBM, Oracle, OpenCloud.

<sup>3</sup> Relevancia debido a que: i) son nombrados por otros autores, ii) son tenidos en cuenta por más de una de las empresas consultadas y iii) por el criterio de los autores.

mediante una serie de preguntas, con las cuales el operador podrá determinar si es necesaria la implementación del criterio y cuáles de los lineamientos son los más adecuados para su caso específico. A continuación se presentan algunos lineamientos para tres de los criterios presentados en la sección anterior, esto debido a que el trabajo aun se encuentra en construcción.

### 5.2.1 Balanceo de Cargas

Los operadores tienen diferentes opciones a fin de implementar esta técnica en sus sistemas, como se muestran a continuación:

**Balanceo de cargas de capa 2:** también llamado agregación de puertos, proporciona redundancia y tolerancia a fallos.

**Balanceo de cargas de capa 4:** distribuye las peticiones hacia los servidores en la capa de transporte, es orientada a la conexión.

**Balanceo de cargas de capa 7:** distribución del trabajo en la capa de aplicación, provee calidad de servicio y mejora el rendimiento debido a que las peticiones son asignadas según su contenido.

**Balanceo de cargas MPLS (Multiprotocol Label Switching):** distribuye el tráfico de la red entre los LSP (Label Switch Paths), evitando la congestión y mejorando el rendimiento.

**Balanceo de cargas DNS (Domain Name System):** se devuelve una dirección IP distinta cada vez que se solicite.

**Balanceo de cargas en base de datos:** las peticiones de acceso a la base de datos se distribuyen entre clusters de servidores.

**Balanceo de cargas SIP (Session Initiation Protocol):** mediante un algoritmo de selección se reparten las nuevas solicitudes entre los servidores disponibles.

**Balanceo de cargas computacional:** se dividen las diferentes tareas entre los nodos de un clúster, logrando mayor rendimiento.

### 5.2.2 Escalabilidad

Se consideran dos puntos de vista: escalabilidad vertical y horizontal, las cuales tienen objetivos similares pero difieren en su aplicación, como se aprecia a continuación:

**Escalabilidad vertical:** implica un cambio interno en la arquitectura software; es la acción de repartir toda la arquitectura funcional en componentes que se pueden utilizar por separado.

**Escalabilidad horizontal:** implica un cambio en la periferia de la arquitectura; es la capacidad de las arquitecturas software el establecimiento de conexiones con el software de otras arquitecturas de una manera eficiente.

### 5.2.3 Copia de Seguridad

Este proceso puede ser realizado de forma manual o automática, y se tienen diferentes opciones con el fin de implementarlo.

**Copia de seguridad completa:** copia total de la información; requiere mayor espacio de almacenamiento y mayor tiempo.

**Copia de seguridad incremental:** copia los datos que hayan variado desde la última copia de seguridad. Por lo cual guarda una menor cantidad de datos que una copia de seguridad completa; sus operaciones son más rápidas y ocupan menor espacio.

**Copia de seguridad diferencial:** almacena todos los datos que hayan sido modificados desde la copia de seguridad anterior. Sin embargo, cuando se vuelve a ejecutar, continúa copiando los datos que hayan cambiado. Almacena más datos que una copia de seguridad incremental, pero menos que una completa.

**Copia de seguridad completa sintética:** reconstruye la imagen de la copia de seguridad completa utilizando todas las copias de seguridad incrementales o la copia de seguridad diferencial guardada y se restaura en un medio alterno.

**Copia en espejo:** copia los datos almacenados en el disco de copia de seguridad a otro juego de discos.

**Copia de seguridad incremental inversa:** se utiliza para añadir soporte de la copia de seguridad de tipo incremental.

**Protección de datos continua (CDP):** permite un mayor número de puntos de restauración que las opciones de copia de seguridad tradicionales.

## 6. CONCLUSIONES

En este artículo se presentó una caracterización de los criterios más relevantes alrededor de la alta disponibilidad en el contexto de una NGSDP, tomando como punto de partida las contribuciones realizadas al respecto por varias empresas del sector; igualmente, fueron definidos un conjunto de lineamientos con los cuales se da cumplimiento a cada uno de los criterios propuestos, facilitando a los operadores la selección de las directrices correctas de acuerdo a su contexto específico.

Actualmente se encuentra en construcción una estructura conceptual de referencia para la organización de los criterios definidos; igualmente, se encuentra en desarrollo un prototipo sobre el cual se realizarán las pruebas correspondientes para validar la aplicación satisfactoria de los lineamientos propuestos para garantizar una alta disponibilidad de servicio en el contexto de una NGSDP.

## 7. REFERENCIAS

- [1] S. H. Maes, "Service delivery platforms as IT Realization of OMA service environment: service oriented architectures for telecommunications," 2007, pp. 2883-2888.
- [2] M. Camelo, *et al.*, "Convergencia de servicios en redes de próxima generación," *Revista SISTEMAS ACIS*, pp. 42-51, Octubre-Diciembre 2008.
- [3] H. Lu, *et al.*, "The Next Generation SDP Architecture: Based on SOA and Integrated with IMS," 2009, pp. 141-145.
- [4] Moriana, *SDP 2.0 Service Delivery Platforms in the Web 2.0 Era*, 2008.
- [5] C. Ghazel, "Dimensioning of Next Generation Networks Multiservice Gateway for Achieving a Quality of Service Target," 2008, pp. 260-267.
- [6] C. F. Estrada and J. A. Caicedo, "Criterios Técnicos para el Aprovisionamiento de VAS en una NGN dentro del Contexto Colombiano," Investigación, Grupo de Ingeniería Telemática, Universidad del Cauca, Popayán, 2010.
- [7] S. M. Pantoja and J. J. Imbachí, "Propuesta de Solución de Alta Disponibilidad de los Servicios Críticos del Centro de Datos de la Universidad del Cauca," Investigación, Universidad del Cauca, Popayán, 2010.
- [8] Moriana, "SDP 2.0-Third Generation SDP Architecture," in *SDP 2.0 Service Delivery Platforms in the Web 2.0 Era*, K. Kimbler, Ed., ed, 2008, pp. 88-95.
- [9] E. Ciurana, "Scalability & High Availability," 2009.
- [10] OpenCloud, "Rhino 2.1: Overview And Concepts," pp. 32-34, March 2009.
- [11] K. Loney and B. Bryla, "Oracle Database 11g Alta Disponibilidad Handbook," *Informe Ejecutivo de Oracle*, 2007.
- [12] Cisco, "Designing High-Availability Services," ed, pp. 173-201.