

IMPLEMENTACIÓN DE QoS EN REDES aerDOCSIS

WHITE PAPER

En este documento se explicará en detalle un término que está totalmente instaurado en el ámbito de las redes de datos, pero que en muchas ocasiones no se conoce en profundidad: la Calidad de Servicio, o QoS. Se describirá qué significa exactamente, qué formas de implementación existen y cuales son más adecuadas.

Se profundizará en la forma en la que el estándar IEEE 802.16-2012 realiza la implementación de estos mecanismos: cómo cada Estación Base aerDOCSIS realiza una asignación de recursos determinista a nivel 2 que le permite ofrecer servicios diferenciados a cada uno de sus usuarios.

INTRODUCCIÓN

La QoS (calidad de servicio) es una condición indispensable en la mayoría de las tecnologías. Hoy día, los operadores se enfrentan con clientes que exigen una conexión de banda ancha de alta capacidad, servicio de telefonía de calidad sin cortes y servicio de IPTV con contenidos multimedia.

Los servicios sobre IP pueden ser sensibles a retardos. Por ello, es importante contar con mecanismos que permitan ofrecer garantías reales. Aquí es donde nace la necesidad de mecanismos de QoS.



DEFINICIÓN E IMPORTANCIA DE LA QoS

En el ámbito de las comunicaciones, se puede decir que la QoS es el cumplimiento del **Service Level Agreement (SLA)**, también llamado Acuerdo de Nivel de Servicio. Un SLA es un **contrato que se establece entre operador y cliente en donde se acuerdan unos parámetros mínimos** de servicio, disponibilidad, rendimiento o tarificación, entre otros. Estos parámetros se pueden medir, y ya que el cliente paga por ellos, el proveedor debe poder garantizárselos.

Por lo tanto, cuando se habla de que un sistema, un estándar o un protocolo garantizan la QoS, es porque tecnológicamente tienen los mecanismos que permiten, de forma determinista y sin ninguna duda, garantizar los niveles de servicio que se establezcan en el “contrato” entre operador y cliente final.

Las tecnologías con QoS presentan numerosas ventajas para aquellos que quieran ofrecer servicios de telecomunicaciones, ya que permiten:

- **Garantías de cumplimiento del SLA.** Los operadores pueden garantizar a sus clientes unos niveles mínimos de servicio contratado, con la certeza de que la tecnología va a ser capaz de soportarlo.
- **Experiencia de usuario satisfactoria.**
- **Diferenciación de usuarios, servicios y políticas de tarificación.** Con QoS es posible proporcionar diferentes prioridades a distintos usuarios y aplicaciones, estableciendo diferentes tarifas en función de las prestaciones ofrecidas.

- **Transporte de aplicaciones exigentes.** En cualquier red de comunicaciones, los paquetes de datos pueden sufrir retardos, jitter, errores o pérdidas. Para determinadas aplicaciones esto no resulta un problema crítico ya que puede ser solucionado por protocolos de niveles superiores. Sin embargo, en otros casos no es suficiente con conseguir la información “si la red lo permite” (Best Effort) y resulta crítico garantizar unos niveles mínimos de calidad. Algunos ejemplos de este tipo de aplicaciones son:

- **Vídeo sobre IP** o cualquier tipo de streaming multimedia en tiempo real (IPTV, multicast de vídeo, etc). Se necesita sobre todo un throughput mínimo sostenido y una tasa mínima de pérdida de paquetes para no recibir vídeo con errores (pixelado, con cortes, etc).
- **Voz sobre IP y videoconferencia (VTC).** No requieren generalmente demasiado throughput, pero sí niveles bajos de jitter para que la voz sea correctamente reproducible en el destino.
- **Juegos en red** que necesitan latencias bajas y constantes.

ÁMBITO DE APLICACIÓN

Al hablar de **QoS en una red clásica de paquetes**, nos referimos a los parámetros a garantizar entre usuarios finales (QoS extremo a extremo). Sin embargo, este documento hace referencia a la QoS aplicada al medio inalámbrico (interfaz aire). No es una QoS extremo a extremo, sino en un interfaz en concreto. Para que QoS tenga sentido en un interfaz se tiene que dar alguna de las siguientes circunstancias:

- Que la capacidad neta máxima de la interfaz sea menor que la capacidad de conexión a la red.
- Que sobre la interfaz se sirvan múltiples usuarios (interfaz punto a multipunto, PtmP).
- Que sobre la interfaz se transporten múltiples servicios con distintas necesidades.

Los mecanismos de QoS en medios inalámbricos son especialmente importantes por:

- **Inestabilidad del medio radio:** la capacidad instantánea que ofrecen las redes inalámbricas es **variable en función de las condiciones del medio**. El tiempo requerido para transmitir un byte varía según la modulación y codificación empleadas, que se ajustan dinámicamente al medio, lo que supone una dificultad añadida para los mecanismos de QoS, que deberán **conocer la capacidad del medio en cada instante**. Esto no ocurre en redes cableadas, que generalmente tienen una alta capacidad fija y una baja probabilidad de error.
- **Escasez de espectro disponible:** el medio radio es escaso y **no es viable sobredimensionar los recursos**.
- **Control del sentido de transmisión:** en un sistema inalámbrico que soporte **ATDD (Adaptive Time Division Duplexing)** con clasificación de QoS centralizada (por ejemplo una BS que controle la QoS en ambos sentidos), el Scheduler también deberá decidir trama a trama cómo reparte el canal para UL y DL.
- **Capacidad del receptor para manejar altas tasas de error:** si una transmisión no ha podido alcanzar al equipo receptor debido a algún desvanecimiento, no es conveniente solicitar una retransmisión de datos inmediatamente ya que es muy probable que la transmisión vuelva a fallar. Además, una alta tasa de retransmisiones para un receptor aumenta la latencia media hacia ese equipo, y afecta a la QoS de otros equipos del sistema. Los mecanismos clasificadores de la QoS tendrán el difícil reto de **mantener el equilibrio del sistema, balanceando el throughput y el uso del medio radioeléctrico**.

CÓMO PROPORCIONAR QoS

El reto para conseguir QoS en redes de paquetes consiste en implementar técnicas adaptativas de QoS que puedan utilizar los recursos disponibles de la forma más eficiente posible.

El modelo más usado en la actualidad para garantizar QoS en una red o interfaz es el modelo de Servicios Diferenciados (DiffServ). Los paquetes de la red incluyen una indicación del servicio que transportan, y conociendo esta información los operadores pueden configurar sus nodos intermedios para que ofrezcan a estos paquetes un diferente “trato” a la hora de encaminarlos.

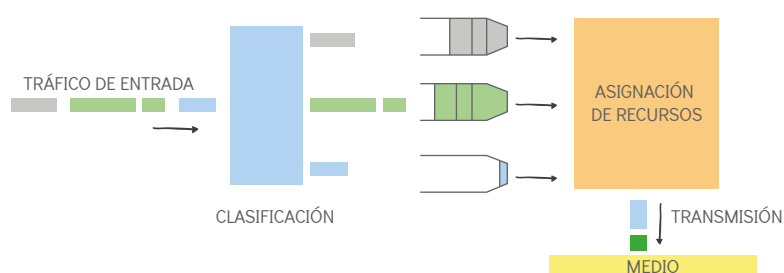


FIGURA 1 - MODELO SIMPLIFICADO DE QoS

Implementación básica

La Figura 1 representa un modelo simplificado pero útil para comprender cómo funcionan los mecanismos de QoS. Existen dos fases fundamentales que cualquier mecanismo de QoS debe implementar:

- **Clasificación**
- **Asignación de recursos**

a) **Clasificación:** Cuando a un nodo de red que implementa control de QoS le llega más tráfico del que puede sacar, debe decidir a qué paquetes va a dar mayor prioridad.

Antes de transmitir se clasificará todo el tráfico entrante. Así se podrá aplicar a cada tipo de tráfico unos criterios de prioridad concretos. Éstos habrán sido establecidos por el administrador de la red en el nodo clasificador.

Existe gran variedad de criterios de clasificación de tráfico. Pueden ser de nivel 2 (dirección MAC), de nivel 3 (dirección IP, cabecera TOS/DSCP) o de nivel 4 (puerto origen/destino).

El criterio de filtrado es una decisión que toma el administrador de red. Lo importante es poder dividir un flujo común de datos en distintos flujos que podrán tener definidos unos determinados parámetros de QoS, con colas de espera diferentes. El nodo clasificador, que conoce los requisitos de QoS, podrá así realizar la asignación de recursos dando preferencia a los flujos prioritarios.

b) **Asignación de recursos:** Una vez que el tráfico esté clasificado y se sepa qué parámetros de QoS debe cumplir, se asignan los recursos en la interfaz. Para ello es necesario que los paquetes se transmitan al medio en un orden adecuado.

A la parte encargada de la asignación se le conoce como el Scheduler; un componente fundamental en cualquier arquitectura de QoS. Éste realiza la reserva de ancho de banda para los diferentes flujos de servicio activos.

En un sistema sin soporte de QoS se trata todo el tráfico de las diferentes aplicaciones a modo de cola FIFO (First- In, First-Out).

En un sistema que implementa QoS, el Scheduler asegura que el tráfico es servido en base a los requisitos de QoS y no según el orden de llegada, permitiendo que un servicio con la máxima prioridad no se vea afectado por un aumento de la carga de tráfico total de la red.

La fase de clasificación es común a todos los tipos de tecnologías que necesitan garantizar la QoS, pero la principal diferencia es la fase de asignación de recursos. Existen dos mecanismos lo bastante generales como para darles un nombre: “QoS a nivel 3 (L3QoS o IPQoS)” y “QoS a nivel 2 (L2QoS o MACQoS)”. Las diferencias entre ambas se explican a continuación.

QoS A NIVEL IP Y A NIVEL MAC

L3QoS: QoS a nivel IP - Las técnicas que se usan en este tipo de mecanismos de QoS son las que utilizan los tradicionales **conformadores de tráfico o Traffic Shapers (TS)**. Un TS realiza la fase de clasificación, la diferencia viene en la fase de Asignación de Recursos.

Una vez que el tráfico está clasificado, el TS asigna de forma estadística los recursos de transmisión al medio. Por ejemplo, si la cola de un servicio de baja latencia está muy llena, intentará vaciarla rápidamente, o si la cola de un servicio con tasa mínima garantizada tiene paquetes, intentará mantener en promedio esa tasa a la salida.

Estas técnicas de **QoS a nivel 3**, también llamadas a **nivel IP**, son las clásicas basadas en **colas de prioridades asociadas al DSCP o al TOS de las cabeceras IP**, por ejemplo.

El problema que presentan las técnicas L3QoS es que no se conoce con exactitud la capacidad y la disponibilidad del medio sobre el que se transmiten. En el medio radio, el tráfico bruto hacia un cliente es algo muy variable que depende del estado de ese radioenlace. Usar técnicas de L3QoS en estos casos, cuando se desconoce la capacidad real hacia cada usuario en cada instante, lleva a una ineficiencia insalvable: No se puede garantizar una QoS en términos absolutos, sólo relativos.

Esto quiere decir que si por ejemplo se han definido un servicio de 1Mbps y otro de 2Mbps, la única garantía que puede hacer un sistema de L3QoS es que el tráfico del primero sea la mitad que el del segundo, pero no puede garantizar cuál va a ser en realidad ese mínimo, ya que desconoce el estado y disponibilidad del medio.

Este problema aún se agrava mucho más en el caso en el que el acceso al medio esté gestionado en contienda (WiFi, Ethernet, etc.); en estos casos el propio uso del medio es estadístico, ni siquiera el nivel 2 puede saber si podrá transmitir en un momento dado. Es más, en el caso de que la red empiece a cursar mucho tráfico, es posible que un paquete jamás sea transmitido debido a las continuas colisiones.

L3QoS

- La Capa IP (nivel 3) no conoce el estado de la capa física (nivel 1) en cada instante.
- Se hace una transmisión de paquetes en base a prioridades relativas, no absolutas.
- Asignación de recursos estadística.

L2QoS

- La Capa MAC (nivel 2) conoce en todo momento el estado de su capa inferior (capa física, nivel 1).
- Asignación de recursos determinista.

L2QoS: QoS a nivel MAC - Cuando la asignación de recursos se hace a nivel 2, el sistema asigna los slots de transmisión y decide cuántos paquetes por segundo encamina, con qué tasa binaria y hacia que usuario, además, conoce en todo momento la capacidad bruta del medio, con lo que también sabe cuánto tráfico neto es capaz de transmitir hacia cada usuario.

Esto hace posible implementar algoritmos que permitan garantizar de forma absoluta la asignación de tráfico.

La tecnología aerDOCSIS, por ejemplo, es un sistema de L2QoS. La Estación Base (BS) es el nodo maestro de la red que asigna la transmisión de datos tanto en la bajada hacia los usuarios (Downlink), como en la subida desde los usuarios (Uplink). El tener un nodo central permite eliminar la contienda en el Acceso al Medio, lo que garantiza que la BS puede conocer en todo momento la disponibilidad del medio radio. Además, la BS aerDOCSIS conoce la calidad del enlace de cada uno de los clientes que tiene conectados, ya que sabe el estado de las modulaciones y, por lo tanto, la capacidad física existente con lo que puede asignar de una forma totalmente determinista el tráfico, tanto en bajada como en subida.

Por supuesto, la calidad de servicio a nivel MAC no es exclusiva de aerDOCSIS. Por ejemplo, DVB-RCS, un protocolo estándar para el acceso múltiple vía satélite, es un esquema parecido: un nodo central que asigna tráfico, un conocimiento exhaustivo de la capacidad y disponibilidad del medio y, por tanto, una QoS que se puede garantizar.

QoS DE ALBENTIA SYSTEMS

Acceso al medio - En situaciones de alto tráfico o número de estaciones elevado, el acceso al medio aleatorio da lugar a un gran número de colisiones, con lo que el uso del canal es ineficiente y el throughput neto total del sistema se ve drásticamente reducido. Por el contrario, con la QoS a nivel 2 de aerDOCSIS la Estación Base actúa de "árbitro" y gestiona el acceso al medio de todas las estaciones suscriptoras (CPEs).

Es una tecnología entramada en la que la BS asigna a cada usuario registrado en la celda intervalos de transmisión (time slots) utilizando TDMA, para que puedan transmitir en sus slots sin necesidad de "competir" entre ellos. Es un Acceso al Medio completamente determinista, en el que la BS conoce en todo momento las modulaciones que tienen todos los usuarios y, por tanto, el estado del medio radio en cada instante, y, así, es capaz de reservar slots dependiendo de los criterios de QoS que estén establecidos. Por eso, se habla de una asignación de recursos a nivel 2 (L2QoS).

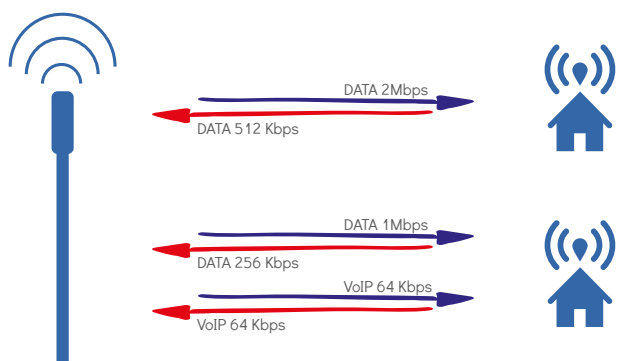


FIGURA 2 - BS con dos CPE y distintos flujos de servicio

Sistema aerDOCSIS

- Entramado: TDD y TDMA (time slots).
- Centralizado: BS ejerce de árbitro, controla todas las transmisiones.
- Capa MAC determinista: sin colisiones ni contienda.
- Conocimiento total del medio radio: la BS conoce siempre el estado del medio radio y las demandas de tráfico.

Flujos de servicio

La tecnología aerDOCSIS basa la transmisión de datos en el concepto de **Flujos de Servicio (service-flows o flows)**, que son conexiones de datos unidireccionales e individuales que se pueden establecer entre la BS y cada CPE.

Son como "canutos" o "tuberías" que se crean en el aire y que son el soporte sobre el que circulan los paquetes de datos entre BS y CPE. La Figura 2 muestra un ejemplo de BS con dos equipos CPE y distintos flujos de servicio.

Los paquetes de datos no pueden transmitirse al aire si no son introducidos previamente en uno de estos flujos. Cada cliente puede tener varios flujos de distintos tamaños y con distintos requisitos de QoS por los que transportar paquetes de datos asociados a distintas aplicaciones: datos, voz, vídeo, etc.

Cuando un terminal (CPE) entra en una celda aerDOCSIS, se le asignan los flujos de servicio que la BS tenga especificados en su Base de Datos de provisionamiento para ese usuario. Cada uno de estos servicios llevará asociado, de forma independiente al resto de servicios, las siguientes propiedades:

- **Un contrato de QoS:** define las características de cada flujo y se guarda en la Base de Datos de la BS. Son parámetros como la tasa binaria máxima sostenible, tasa mínima garantizada, tipo de servicio, retardo máximo, etc. En resumen, parámetros que recogen las características de cada filtro.
- **Una política de filtrado:** una serie de reglas que permitirán determinar qué paquetes de datos se van a cursar por cada servicio. Se encargan de cumplir la fase de Clasificación de paquetes presente en todo proceso de QoS.

En conclusión, aerDOCSIS soporta QoS diferenciado para distintos usuarios y para distintos flujos, empleando estos flujos de servicio para llevar a cabo las dos tareas necesarias de cualquier mecanismo de QoS: Clasificación del tráfico y Asignación de recursos. Cada uno de estos flujos se destinará a cursar distintos tipos de tráfico y se le asignarán unas políticas de filtrado y unos recursos en el medio. Usar servicios para transportar datos, además de ser diferente a otras tecnologías, es fundamental para ofrecer garantías de QoS a nivel de operador.

En redes de operadores siempre se hace una sobreprovisión de servicios, es decir, se dimensiona la estación base para vender más capacidad de forma no garantizada. Gracias a la QoS esto no es un problema, ya que cada usuario tiene definidos sus flujos de servicio diferenciados con los máximos permitidos y mínimos garantizados. La clave está en que la suma de todos los mínimos garantizados no supere la capacidad neta proporcionada por la base. Una vez que la base haya garantizado los mínimos, comenzará a asignar recursos a todos los demás servicios que están demandando cursar tráfico atendiendo a las políticas de QoS y de forma proporcional, hasta completar su capacidad total. De modo, un cliente con 20Mbps contratados recibirá el doble de recursos que uno con 10Mbps contratados.

Tipos de servicio

El estándar IEEE 802.16-2012 define 5 diferentes tipos de servicios en función de la QoS que pueden ofrecer. Si bien el estándar define estos tipos sólo en el UL, en general se aplican tanto al UL como al DL. Estos 5 tipos se pueden agrupar en 3 grandes grupos:

- **BE (Best Effort):** este método ofrece un servicio eficiente para el tráfico “Best Effort” y son empleados habitualmente para servicios de datos, que no suelen requerir niveles mínimos de servicio. Estos servicios son atendidos en función de la disponibilidad de la celda en cada momento y cuando todos los servicios con una prioridad superior hayan sido ya atendidos. Están pensados para soportar tasas binarias de este tipo:

Máximo Tráfico sostenido = X
Mínimo Tráfico garantizado = 0

- **(x)RTPS:** engloba tres tipos de servicios: RTPS (Real Time Polling Service), NRTPS (Non Real Time Polling Service) y eRTPS (Extended Real Time Polling Service). Se diferencian por las latencias que entregan al usuario final y están orientados a aplicaciones con requerimientos de capacidad mínima garantizada. Tienen una tasa mínima garantizada y son ideales para productos como VoIP. Cuando el usuario necesita transmitir algo lo solicita a la BS, que le dará un trato preferencial frente a flujos BE. Están pensados para soportar tasas binarias de este tipo:

Máximo Tráfico sostenido = Y
Mín. Tráfico garantizado = X (X<Y)

- **UGS (Unsolicited Grant Service):** se utilizan para aplicaciones de tráfico constante como transmisión de vídeo ininterrumpido o tramas E1/T1. Si un CPE tiene asignado un servicio de este tipo, la BS le garantizará siempre un tiempo de transmisión en la trama, independientemente del tráfico. Ofrecen una mayor calidad de servicio, a costa de reservar continuamente tiempo en la trama, con lo que conviene dimensionar correctamente este tipo de servicios para no reducir demasiado el throughput total agregado de la celda. No se suelen utilizar en redes de operador. Siguen un patrón de tráfico de este tipo:

Máximo Tráfico sostenido = X
Mínimo Tráfico garantizado = 0

Excepto en el caso de UGS, un CPE que quiera realizar una transmisión tendrá que solicitarlo a la BS en un periodo de la trama destinado a ello mediante un mensaje denominado Bandwidth Request. Cuando la BS recibe esta petición, le asignará el ancho de banda que pueda, hasta el máximo delimitado en el parámetro “Tasa Binaria Máxima” de cada flujo, y cursando en primer lugar aquellos flujos con un tipo de QoS mayor.

Beneficios de los servicios aerDOCSIS

El método de transmisión basado en flujos de servicio permite ser **muy eficiente a la hora de asignar los recursos en un sistema compartido** como es el medio radio y, además, **proporciona control absoluto sobre el uso del ancho de banda**, evitando abusos o simplemente que el propio tráfico de datos de un usuario entrecorte sus propias llamadas de VoIP. **El resultado es una mejor experiencia de navegación** así como el uso de servicios diferenciados por parte de los usuarios y una capacidad mejorada para proveer servicios de datos de diferentes niveles y servicios de valor añadido para el operador de la red.

Para establecer flujos de servicio asociados a distintos productos comerciales, **el operador tiene que:**

- Definir los **productos que se van a ofrecer**
- Definir los **servicios aerDOCSIS que van a soportar los productos** comerciales
- **Provisionar los servicios** a los clientes en función de los productos comerciales que han contratado

Algunos de los beneficios que obtienen el operador y sus clientes con este modelo son:

- **El operador obtiene una mejora del uso de la capacidad radio**, al evitar abusos y reducir sobreprovisiones, y haciendo posible el aumento del número de usuarios por celda
- **El cliente consigue un servicio de calidad** y ajustado a lo que contrata, evitando que otros clientes hagan abuso de su ancho de banda
- **Permite ofrecer nuevos servicios:** la VoIP con calidad "Carrier Class" es ahora posible, ya que los servicios permiten establecer la máxima prioridad para las llamadas y evitar que el tráfico de la celda entrecorte la voz o imposibilite su realización
- En general, **la experiencia del cliente mejora** frente a otros medios inalámbricos e incluso supera en muchos casos la experiencia obtenida con medios cableados

En resumen, **los servicios de aerDOCSIS permiten ofrecer al operador un modelo de datos equivalente al de un acceso cableado de última milla** (QoS, separación voz de y datos, etc), pero con todas las ventajas del acceso radio: inmediatez, reducción de costes, escalabilidad, entre otras.

CONCLUSIONES

1* **Definición de QoS:** conjunto de mecanismos que garantizan unas mínimas condiciones de rendimiento en una comunicación, estableciendo un contrato entre operador y cliente que recoge el compromiso de servicio acordado y que el operador tendrá que ser capaz de cumplir siempre, utilizando para ello una tecnología que se lo permita.

2* **La necesidad de QoS:** ofrecer QoS es muy importante para los proveedores de servicios, ya que pueden garantizar a sus clientes unos niveles mínimos de calidad y crear distintos niveles de servicio con diferente tarificación. Además, hoy en día existen muchas aplicaciones sobre IP que por su propia naturaleza requieren necesariamente un estricto control de QoS: VoIP, videovigilancia, IPTV, etc.

3* **QoS en redes inalámbricas:** éstas son usadas cada vez más en redes de operadores, por lo que deben estar preparadas para ofrecer un amplio rango de servicios. Estas tecnologías presentan más dificultades que las redes cableadas: el medio de propagación, inestable por naturaleza y con mayor probabilidad de error, unido a un espectro radioeléctrico limitado, suponen un reto adicional al problema del QoS inalámbrico, especialmente en entornos NLOS.

4* **QoS en Alcentia Systems:**

- **Capa MAC:** el nivel de Acceso al Medio es la piedra angular de esta tecnología, permite implementar de forma determinista cualquier mecanismo de QoS.
- **Layer2QoS:** uno de los puntos fuertes de aerDOCSIS es que implementa mecanismos de QoS hasta nivel 2 que le permiten ofrecer servicios diferenciados de forma determinista, ya que la BS conoce en todo momento el estado del medio radio y, por lo tanto, es capaz de asignar los recursos en cada momento.
- **Flujos de servicio:** los datos que fluyen por el aire deben ser transportados por flujos de servicio. La tecnología aerDOCSIS se basa en diferenciar aplicaciones utilizando la diferenciación de flujos.
- **Tipos de servicio:** los flujos de datos en aerDOCSIS pueden ser de cinco tipos distintos en función del tipo de priorización que se quiera realizar.

WHITE PAPER