

Suscripciones óptimas de las APIs externas

Introducción

En este documento se recogen los análisis justificativos de la suscripciones óptimas de cada API externa usada en el proyecto. Se elaborará un estudio para cada API externa, dividiendo el documento en secciones. Cada sección está destinada a un microservicio, y cada microservicio tiene el análisis de sus APIs externas usadas. En caso de que no haya ninguna, se indicará.

Análisis justificativos de las suscripciones óptimas

User Auth

Resend

La API de **Resend** se utiliza dentro del microservicio **User Auth** para enviar correos electrónicos de confirmación de registro, verificación de email y restablecimiento de contraseña.

Planes de suscripción disponibles

Resend ofrece cuatro planes de **Transactional Emails** adaptados a distintos niveles de uso:

1. Plan Free (Gratuito)

- **Precio:** \$0/mes
- **Emails incluidos:** 3.000 emails/mes
- **Límite diario:** 100 emails/día
- **Dominios:** 1 dominio personalizado
- **Rate Limit:** 2 peticiones/segundo
- **Retención de datos:** 1 día
- **Soporte:** Ticket
- **IPs dedicadas:** No disponible

2. Plan Pro

- **Precio:** \$20/mes
- **Emails incluidos:** 50.000 emails/mes
- **Emails adicionales:** \$0.90 por cada 1.000 emails extra
- **Dominios:** Hasta 10 dominios
- **Retención de datos:** 3 días
- **Soporte:** Ticket
- **IPs dedicadas:** No disponible
- **Sin límite diario**

3. Plan Scale

- **Precio:** \$90/mes
- **Emails incluidos:** 100.000 emails/mes
- **Emails adicionales:** \$0.90 por cada 1.000 emails extra
- **Dominios:** Hasta 1.000 dominios
- **Retención de datos:** 7 días
- **Soporte:** Slack + Ticket
- **IPs dedicadas:** Disponible como add-on
- **Sin límite diario**

4. Plan Enterprise

- **Precio:** Personalizado
- **Características:**
 - Volumen de emails adaptado a necesidades específicas
 - Dominios flexibles
 - Retención de datos flexible
 - Soporte prioritario
 - IPs dedicadas con add-on
 - Sin límite diario

Configuración de dominio personalizado

Una de las ventajas clave de Resend es la posibilidad de **añadir un dominio propio de forma gratuita** incluso en el plan Free. En nuestro caso, hemos configurado el dominio **socialbeats.es** (adquirido en IONOS) como remitente verificado.

Beneficios de usar dominio propio:

- **Mayor confianza:** Los emails se envían desde direcciones como `noreply@socialbeats.es`, lo que aumenta la credibilidad.
- **Mejor entregabilidad:** Al configurar los registros DNS (SPF, DKIM, DMARC), los correos evitan ser marcados como spam.
- **Imagen profesional:** Refuerza la identidad de marca de la aplicación.
- **Sin coste adicional:** Resend permite esta configuración en todos los planes, incluido el gratuito.

Estimación de uso del proyecto

Funcionalidades que generan emails

Evento	Emails estimados/día
Registros nuevos	10-30
Verificaciones de email	10-30
Recuperación de contraseña	5-15
Cambios de contraseña	2-15

Evento	Emails estimados/día
Total diario estimado	27-90 emails

Proyección mensual

- **Escenario conservador:** ~800 emails/mes
- **Escenario de crecimiento:** ~4.000 emails/mes

Análisis de costes

Escenario	Emails/mes	Plan recomendado	Coste
Desarrollo/Beta	<1.000	Free	\$0
Lanzamiento inicial	1.000-3.000	Free	\$0
Crecimiento moderado	3.000-50.000	Pro	\$20/mes
Escala	50.000-100.000	Scale	\$90/mes

Suscripción óptima: Plan Free

Justificación

1. **Volumen suficiente para la fase actual** Con 3.000 emails/mes y hasta 100 emails/día, el plan gratuito cubre holgadamente las necesidades del proyecto en su fase de desarrollo y lanzamiento beta.
2. **Dominio personalizado incluido** Podemos enviar emails desde [@socialbeats.es](#) sin coste adicional, lo que mejora la entregabilidad y la imagen profesional.
3. **Rate limit adecuado** 2 peticiones por segundo es suficiente para el volumen actual. Además, hemos implementado un rate limiter interno con [Bottleneck](#) para respetar este límite.
4. **Circuit breaker implementado** El servicio de emails incluye un circuit breaker que protege contra fallos del proveedor, haciendo el sistema resiliente independientemente del plan.
5. **Migración sencilla** Si el volumen crece, la migración al plan Pro (\$20/mes) es inmediata y sin cambios de código.

Plan de acción propuesto

Fase 1 — Desarrollo y beta

- Uso del **Plan Free**
- Dominio [socialbeats.es](#) configurado
- Monitorización del volumen de envíos

Fase 2 — Lanzamiento público

- Evaluación del volumen real
- Si se superan los 100 emails/día de forma consistente → migrar a **Plan Pro**

Fase 3 — Escalado

- Si se superan los 50.000 emails/mes → evaluar **Plan Scale**
- Considerar IPs dedicadas para máxima entregabilidad

Conclusión

El **Plan Free de Resend** es la opción óptima para el microservicio **user-auth**, ofreciendo:

- Coste cero en la fase actual
- Dominio personalizado con máxima entregabilidad
- Límites suficientes para desarrollo y beta
- Escalabilidad transparente hacia planes superiores

La configuración del dominio **socialbeats.es** con los registros DNS apropiados garantiza que los correos transaccionales lleguen a la bandeja de entrada de los usuarios y no sean filtrados como spam.

Beats Upload

AWS S3 (Amazon Simple Storage Service)

Contexto y propósito

La API de **AWS S3** se utiliza dentro del microservicio **Beats Upload** para almacenar archivos de audio (beats) e imágenes de portada. Los usuarios suben archivos mediante URLs presignadas (Presigned POST), y el contenido se distribuye a través de CloudFront (CDN) con URLs firmadas para streaming seguro.

Planes de suscripción disponibles

AWS S3 ofrece diferentes clases de almacenamiento. Para nuestro caso de uso (acceso frecuente a archivos de audio), utilizamos **S3 Standard**.

1. Free Tier (Gratuito - 12 meses)

- **Almacenamiento:** 5 GB
- **PUT, COPY, POST, LIST requests:** 2.000/mes
- **GET requests y otras:** 20.000/mes
- **Transferencia de datos:** 100 GB salida/mes (compartido con otros servicios AWS)
- **Coste:** \$0
- **Duración:** 12 meses desde la creación de la cuenta AWS

2. S3 Standard (Post Free Tier)

- **Almacenamiento:**
 - Primeros 50 TB/mes: \$0.023/GB

- Sigüientes 450 TB/mes: \$0.022/GB
- Más de 500 TB/mes: \$0.021/GB
- **PUT, COPY, POST, LIST requests:** \$0.005 por 1.000 requests
- **GET, SELECT y otras requests:** \$0.0004 por 1.000 requests
- **DELETE requests:** Gratis
- **Sin límites de almacenamiento ni requests**

Créditos promocionales

Actualmente operamos con **créditos promocionales de AWS** adicionales al Free Tier, lo que nos permite operar sin coste durante la fase de desarrollo y beta.

Estimación de uso del proyecto

Funcionalidades que generan operaciones S3

Operación	Tipo de Request
Subir beat (audio)	PUT
Subir portada (imagen)	PUT
Streaming de audio	GET (via CloudFront)
Descarga de beat	GET
Eliminar beat	DELETE (gratis)
Generar waveform	GET (se recupera el fichero de audio y se genera el waveform)

Escenario conservador (Desarrollo/Beta)

Métrica	Estimación
Usuarios activos	100-500
Beats subidos/día	5-20
Tamaño promedio beat	~5 MB
Streams/descargas/día	50-200
PUT requests/mes	~150-600
GET requests/mes	~1.500-6.000
Almacenamiento/mes	~750 MB - 3 GB

Escenario de crecimiento (Lanzamiento público)

Métrica	Estimación
---------	------------

Métrica	Estimación
Usuarios activos	500-2.000
Beats subidos/día	50-200
Tamaño promedio beat	~5 MB
Streams/descargas/día	500-2.000
PUT requests/mes	~1.500-6.000
GET requests/mes	~15.000-60.000
Almacenamiento/mes	~7.5 GB - 30 GB

Análisis de costes

Escenario	Almacenamiento	PUT requests	GET requests	Coste mensual
Conservador (Free Tier)	750 MB - 3 GB	150-600	1.500-6.000	\$0
Crecimiento (Post Free Tier)	7.5-30 GB	1.500-6.000	15.000-60.000	\$0.17 - \$0.72

Nota: Los costes del escenario de crecimiento asumen que se ha agotado el Free Tier. Mientras los créditos promocionales estén activos, el coste real es \$0.

Suscripción óptima: Free Tier + Créditos Promocionales

Justificación

- Cobertura total en fase actual** El Free Tier (5 GB storage, 2.000 PUT, 20.000 GET) cubre holgadamente el escenario conservador.
- Créditos promocionales como buffer** Los créditos de AWS proporcionan un margen de seguridad adicional para picos de uso o crecimiento inesperado.
- Costes mínimos post Free Tier** Incluso en el escenario de crecimiento, el coste es inferior a \$1/mes, haciendo S3 extremadamente económico.
- Sin compromisos ni contratos** S3 es pay-as-you-go puro, sin costes fijos ni mínimos.
- Controles de estabilidad implementados** El código incluye Bottleneck (5 operaciones concurrentes) y toobusy-js para proteger contra sobrecarga.

Plan de acción propuesto

Fase 1 — Desarrollo y beta

- Uso del **Free Tier + Créditos Promocionales**
- Monitorización de uso en AWS Console

Fase 2 — Lanzamiento público

- Evaluación del consumo real vs estimaciones
- Si se supera el Free Tier → transición automática a S3 Standard (pay-as-you-go)
- Coste esperado: < \$1/mes

Fase 3 — Escalado

- Si almacenamiento > 50 GB → considerar S3 Intelligent-Tiering para optimizar costes
- Si tráfico > 100.000 requests/mes → evaluar Reserved Capacity

Conclusión

El **Free Tier de AWS S3 + Créditos Promocionales** es la opción óptima para el microservicio, ofreciendo:

- Coste cero en la fase actual
- Límites suficientes para desarrollo y beta
- Costes mínimos en escalado (< \$1/mes)
- Escalabilidad transparente sin cambios de código
- Controles de estabilidad implementados (Bottleneck + toobusy-js)

Beats interaction

OpenRouter API

Contexto y propósito

La API de **OpenRouter** se utiliza dentro del microservicio **Beats-interactions** para habilitar funcionalidades de inteligencia artificial orientadas al análisis, clasificación y moderación de contenido (comentarios, descripciones, metadatos, etc.).

OpenRouter actúa como un *gateway* unificado que da acceso a **más de 400 modelos** (número que se actualiza constantemente), permitiendo seleccionar dinámicamente el modelo más adecuado según coste, latencia y calidad.

Planes de suscripción disponibles

OpenRouter ofrece tres planes principales adaptados a distintos niveles de uso y madurez del proyecto.

1. Plan Gratuito (Free Tier)

- Acceso a modelos gratuitos (variantes con sufijo **:free**)
- **Límite de 50 solicitudes por día** a modelos gratuitos
- **Límite de 20 solicitudes por minuto (RPM)** a modelos gratuitos
- Sin necesidad de tarjeta de crédito para empezar
- Ideal para:
 - Prototipado
 - Pruebas técnicas

- Validación inicial del flujo de IA
- **Coste:** \$0
- **Nota:** El uso gratuito de modelos populares puede estar sujeto a limitaciones adicionales por parte del proveedor durante horas pico

2. Plan Pay-as-you-go (Pago por uso)

- Acceso a **más de 400 modelos de IA** (catálogo completo)
- **Sin límites en modelos de pago** (requiere saldo mínimo de \$10 en créditos)
- **Modelos gratuitos (:free) limitados a:**
 - **1.000 solicitudes por día** (con al menos \$10 en créditos)
 - **50 solicitudes por día** (con menos de \$10 en créditos)
 - **20 RPM** (solicitudes por minuto) en ambos casos
- Facturación transparente basada en:
 - Tokens de entrada y salida
 - Coste específico de cada modelo
- Tarifa de plataforma:
 - ~5.5% sobre el coste del modelo
 - Mínimo de \$0.80 (para pagos no cripto)
- Sistema de créditos prepagados con recarga automática o manual
- Sin compromisos ni contratos a largo plazo
- Protección DDoS mediante Cloudflare
- Ideal para:
 - Productos en producción
 - Escenarios con demanda variable o crecimiento progresivo

3. Plan Enterprise

- Precios personalizados
- Descuentos por volumen
- Créditos prepagados con compromiso anual
- Facturación empresarial (PO, invoices, etc.)
- Enrutamiento regional y *failover* automático
- Políticas avanzadas de retención de datos (incluido **Zero Data Retention**)
- Soporte técnico dedicado
- Ideal para:
 - Empresas con alto volumen
 - Requisitos legales, de privacidad o rendimiento específicos

Estimación de uso del proyecto

Funcionalidades previstas

- Análisis semántico y lírico de beats
- Generación automática de descripciones
- Clasificación y etiquetado por género o mood

- Sistemas de recomendación personalizados
- Moderación de contenido generado por usuarios

Estimación de volumen

- Usuarios activos iniciales: **100–500**
- Interacciones por usuario/día: **5–20**
- Solicitudes diarias estimadas: **500–10.000**
- Tokens promedio por solicitud: **500–2.000** (entrada + salida)

Modelos candidatos

- **Modelos ligeros / medianos**
 - GPT-3.5
 - Claude Haiku
 - Llama 3
 - Clasificación, análisis simple y tareas repetitivas
- **Modelos avanzados**
 - GPT-4
 - Claude Sonnet
 - Recomendaciones complejas, análisis profundo y generación de contenido de mayor calidad

Análisis de costes

Escenario conservador (500 solicitudes/día)

- Uso mensual: ~15.000 solicitudes
- Tokens mensuales estimados: 15M – 30M
- Coste con modelos económicos (ej. GPT-3.5 a \$0.50 / \$1.50 por 1M tokens):
 - **\$7.50 – \$45 / mes**
- Tarifa de plataforma (5.5%):
 - **\$0.80 – \$2.50 / mes**

Total estimado:

\$8.30 – \$47.50 / mes

Escenario de crecimiento (5.000 solicitudes/día)

- Uso mensual: ~150.000 solicitudes
- Tokens mensuales estimados: 150M – 300M
- Coste estimado: \$75 – \$450 / mes
- Tarifa de plataforma: \$4.13 – \$24.75

Total estimado:

\$79.13 – \$474.75 / mes

Escenario de alto volumen (10.000+ solicitudes/día)

- Uso mensual: 300.000+ solicitudes
- Tokens mensuales: 300M+
- Coste estimado: \$150 – \$900+ / mes

En este punto, el **Plan Enterprise** suele resultar más rentable gracias a los descuentos por volumen y acuerdos personalizados.

Suscripción óptima: Plan Pay-as-you-go**Justificación****1. Flexibilidad total**

El proyecto se encuentra en una fase temprana, donde el tráfico real aún es incierto. El modelo pay-as-you-go permite escalar sin riesgo.

2. Sin costes fijos ni compromisos

No hay contratos ni pagos mínimos, lo que reduce la fricción financiera en las primeras etapas.

3. Acceso completo al ecosistema de modelos

Más de 400 modelos disponibles (y en constante crecimiento) para optimizar coste vs calidad según cada caso de uso.

4. Control total del gasto

La facturación por tokens permite monitorizar y ajustar el consumo en tiempo real.
Presupuesto recomendado inicial: **\$50–\$100/mes**.

5. Escalabilidad natural

Al superar consistentemente los \$500/mes, se puede migrar fácilmente al plan Enterprise con mejores condiciones.

6. Ventajas técnicas clave

- Compatibilidad directa con el SDK de OpenAI
- Failover automático entre modelos
- API unificada y estable
- Soporte para Zero Data Retention (cumplimiento de privacidad)

Plan de acción propuesto**Fase 1 — Desarrollo (Mes 1–3)**

- Uso del **Plan Gratuito**
- Validación técnica y pruebas funcionales
- Ajuste de prompts y flujos de IA

Fase 2 — Lanzamiento beta (Mes 4–6)

- Migración al **Plan Pay-as-you-go**
- Presupuesto inicial: \$50–\$100/mes
- Monitorización de consumo y rendimiento

Fase 3 — Escalado (Mes 7+)

- Evaluación mensual de costes
- Si el gasto supera los \$500/mes de forma estable → considerar **Plan Enterprise**

Optimizaciones de coste recomendadas

- Usar modelos más económicos para tareas simples
- Implementar caché de respuestas frecuentes
- Aplicar rate limiting por usuario
- Monitorizar métricas de uso desde el dashboard de OpenRouter
- Separar workloads críticos y no críticos por modelo

Conclusión

El **Plan Pay-as-you-go de OpenRouter** es la opción más equilibrada para el microservicio **Beats-interactions**, ofreciendo:

- Flexibilidad total
- Costes controlados
- Escalabilidad progresiva
- Acceso a modelos de última generación

Esta estrategia permite crecer de forma sostenible, minimizando riesgos financieros y manteniendo un alto nivel técnico desde el inicio.

Analytics and Dashboards

Estudio de Viabilidad Técnica y Económica: Integración de Azure Translator API en Ecosistema de Producción Musical

Esta sección analiza la viabilidad de integrar el servicio de traducción de Microsoft Azure en nuestra plataforma social de *beats*. Se evalúa el impacto del uso de una caché en Redis y la eficiencia de costes bajo distintos escenarios de consumo y planes de usuario.

Introducción y Contexto

En este ecosistema, la API de Azure se utiliza para traducir frases inspiracionales relacionadas con el mundo de la música. Debido a que el proyecto se encuentra en una fase académica, este estudio es fundamental para determinar la suscripción óptima y asegurar que la arquitectura de la aplicación sea escalable antes de enfrentarse a datos reales de tráfico.

Definición de Asunciones del Modelo

Para proyectar el consumo de recursos, he establecido las siguientes asunciones basadas en el diseño técnico del sistema:

- **Estrategia de Caché Asimétrica:** Se ha implementado una capa de persistencia con Redis. La API de origen de las frases actualiza el contenido cada 1 hora, mientras que nuestra caché de traducción en Redis tiene una persistencia de 24 horas.
- **Carga Operativa Base:** Se estima que el sistema procesará una frase nueva cada hora, resultando en 24 frases únicas al día.
- **Métrica de Caracteres:** Se estima una longitud media de 150 caracteres por frase.
- **Direccionalidad Lingüística:** El sistema traduce bidireccionalmente entre inglés y español (\$D=2\$).
- **Distribución de Usuarios:** Se contemplan tres planes (FREE, PRO, STUDIO), aunque en el modelo actual el contenido traducido es global y compartido.

Análisis de Datos y Resultados

Cálculo del Volumen Horario

Basado en la configuración de la caché de Redis, el volumen de caracteres (\$V_h\$) que la API de Azure debe procesar es independiente del número de usuarios (debido al caché global), y se calcula por hora:

$$V_h = (F_h \times C \times D)$$

Donde:

- \$F_h\$ (Frases por hora) = 1 (Nueva frase cada hora)
- \$C\$ (Caracteres promedio) = 150
- \$D\$ (Direcciones de traducción) = 2

Sustituyendo los valores: $V_h = (1 \times 150 \times 2) = 300 \text{ caracteres/hora}$

Análisis de Sensibilidad: Escenario de Aleatoriedad

En un escenario crítico donde la frase fuese aleatoria por usuario (mayor tasa de *Cache Miss*), suponiendo un pico de 1.500 frases únicas al día (~63 frases/hora):

$$V_{h(\text{estrés})} = (63 \times 150 \times 2) = 18.900 \text{ caracteres/hora}$$

Incluso bajo estrés, el consumo se mantiene lejos del límite.

Suscripción Óptima

Comparativa con el límite horario de Azure:

Plan Azure	Límite Gratuito (F0)	Consumo Estimado
Caracteres / Hora	2.000.000	300
Coste Económico	0,00 €	0,00 €
Margen de Seguridad	-	>99,9%

Planes de precios y Evolución del Modelo

En el estado actual de la aplicación, los planes FREE, PRO y STUDIO no afectan la viabilidad, ya que la traducción es un recurso compartido. Sin embargo, se plantean los siguientes escenarios futuros para ajustar la API a los planes de precios:

1. **Traducción de Usuario (PRO/STUDIO):** Permitir la traducción de perfiles o descripciones de *beats* personales, lo que vincularía el coste al crecimiento de usuarios.
2. **Personalización de Idiomas (STUDIO):** Acceso a más de dos idiomas de forma simultánea.
3. **Prioridad de Refresco:** Los usuarios STUDIO podrían forzar la actualización de la frase, saltándose la caché horaria de la API de frases.

Conclusiones Finales

El estudio concluye que la **suscripción óptima es el Plan F0 (Gratuito)**. La implementación de Redis como puente entre la API de frases y la API de traducción permite que el proyecto sea económicamente viable a coste cero. La arquitectura asimétrica diseñada no solo garantiza la gratuidad del servicio en su fase académica, sino que establece una base sólida para escalar a planes de pago (S1) solo cuando se introduzcan funcionalidades de traducción personalizada por usuario.

Social