

CASO DE USO DEL INSTITUTO DE RADIOASTRONOMÍA MILIMÉTRICA

Sesión Cloud

Pablo Mellado

QUIÉNES SOMOS

- Instituto de investigación con dos sedes y dos observatorios
- En España, estamos en la estación de esquí de Sierra Nevada
- Observamos 24 horas al día, pero sólo recibimos.
- <https://youtu.be/VY3Apf8TIH0>



USO DE LA NUBE

- Disponibilidad de la conexión a internet en mitad de la montaña es un condicionante. Posible implementación de servicios poco críticos.
- Un ejemplo: formularios web para inscripción a reuniones y congresos.
- Nueva conexión por fibra durante este pasado mes cambia la situación completamente.
- Proyecto piloto: Almacenamiento de backups y archivo de observaciones en AWS

VOLUMEN DE DATOS

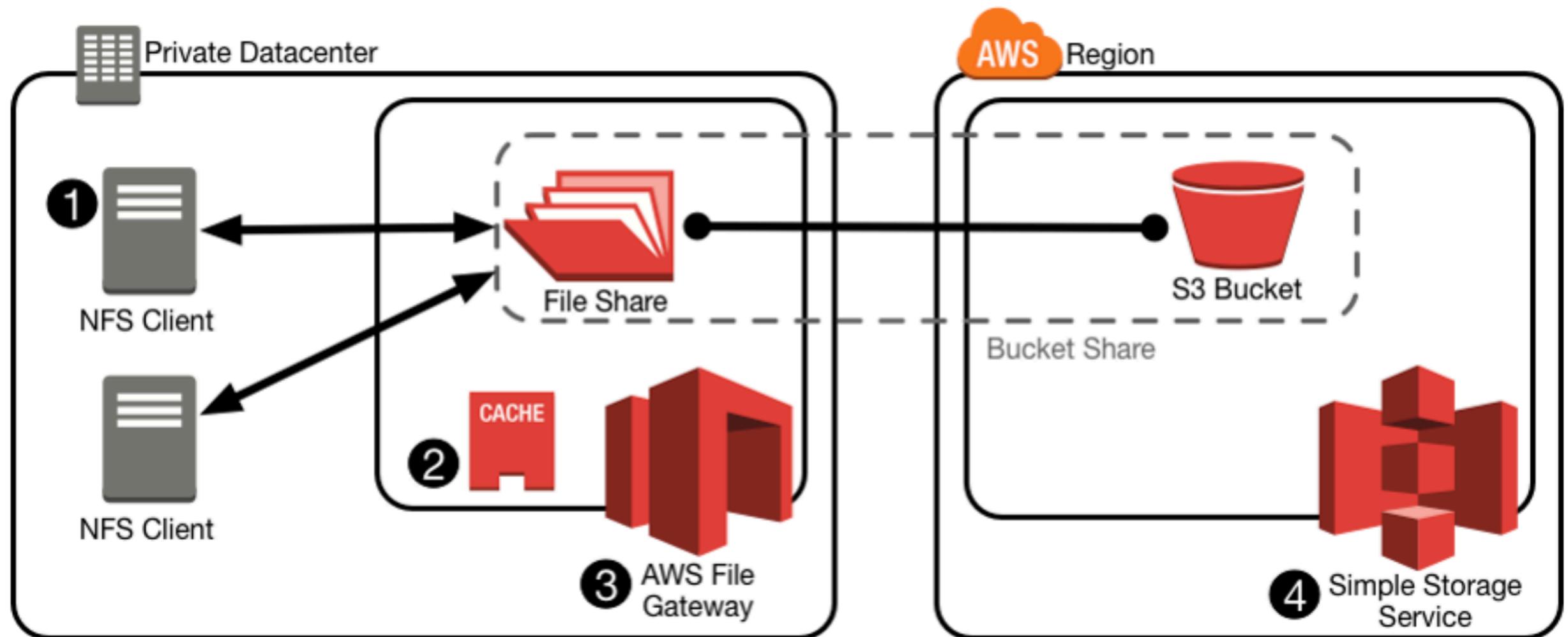
- Este año hemos generado 40TB de datos sin calibrar.
- Desde diciembre de 2013 llevamos generados 216TB (85 cintas LTO6)
- A eso hay que añadir los datos calibrados y procesados que se almacenan en un archivo:
 - 440TB

AWS: STORAGE GATEWAY

- Elegimos AWS porque es el que habíamos usado alguna vez a nivel privado.
- AWS ofrece almacenamiento en la nube *híbrida* con almacenamiento en caché local (seminario).
- Tres posibilidades:
 - Volume gateway: block storage.
 - Tape gateway: using your existing tape-based processes.
 - File gateway: store files as objects in Amazon S3

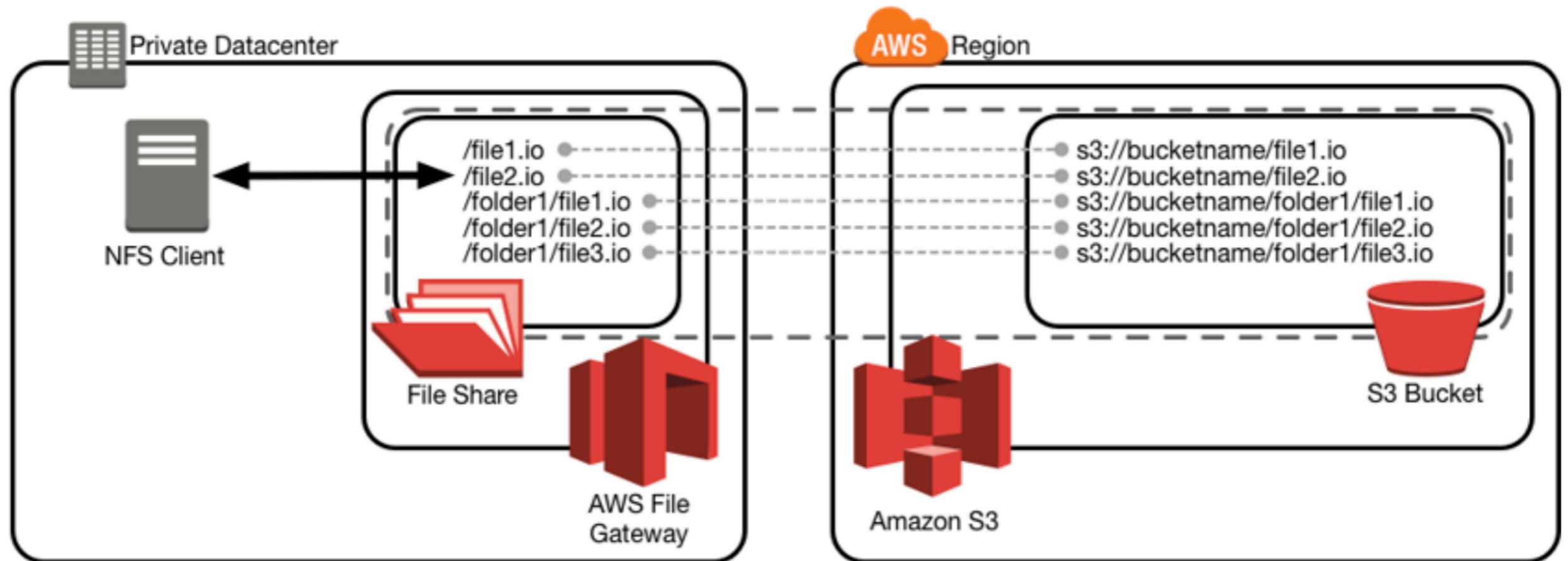
FILE STORAGE GATEWAY

- Bacula permite backups a ficheros grandes.



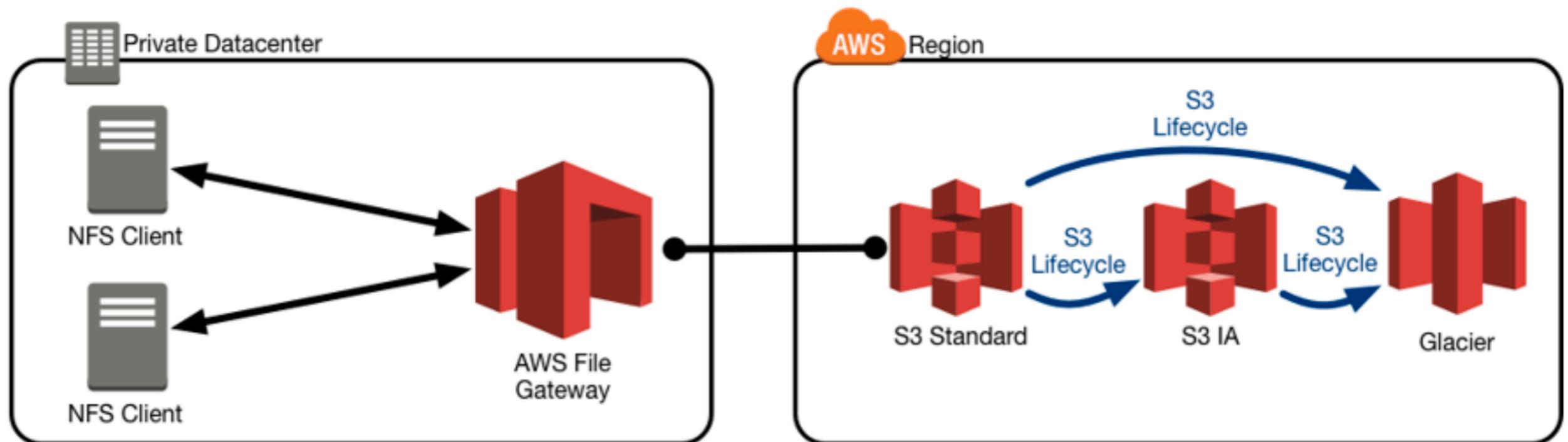
OBJETOS S3

- Correspondencia entre ficheros NFS y objetos



S3: CICLO DE VIDA

- Existen tres tipos de almacenamiento en S3.
- Cuanto más lento más barato. \$0,023 - \$0,0125 - \$0,004.
- A parte el File Gateway cobra \$0.01 por GB transferido.



FUNCIÓN LAMBDA

- Ficheros subidos a S3 son pasados a Glacier cuando ya están *finalizados* por Bacula.
- Una función AWS Lambda escrita en Python se encarga de etiquetar los objetos a *congelar*.
- Se establece en el bucket de S3 una política de ciclo de vida pasando a Glacier los objetos con la etiqueta.
- A parte se fija una alarma para que avise cuando no se trasladen los objetos a Glacier.

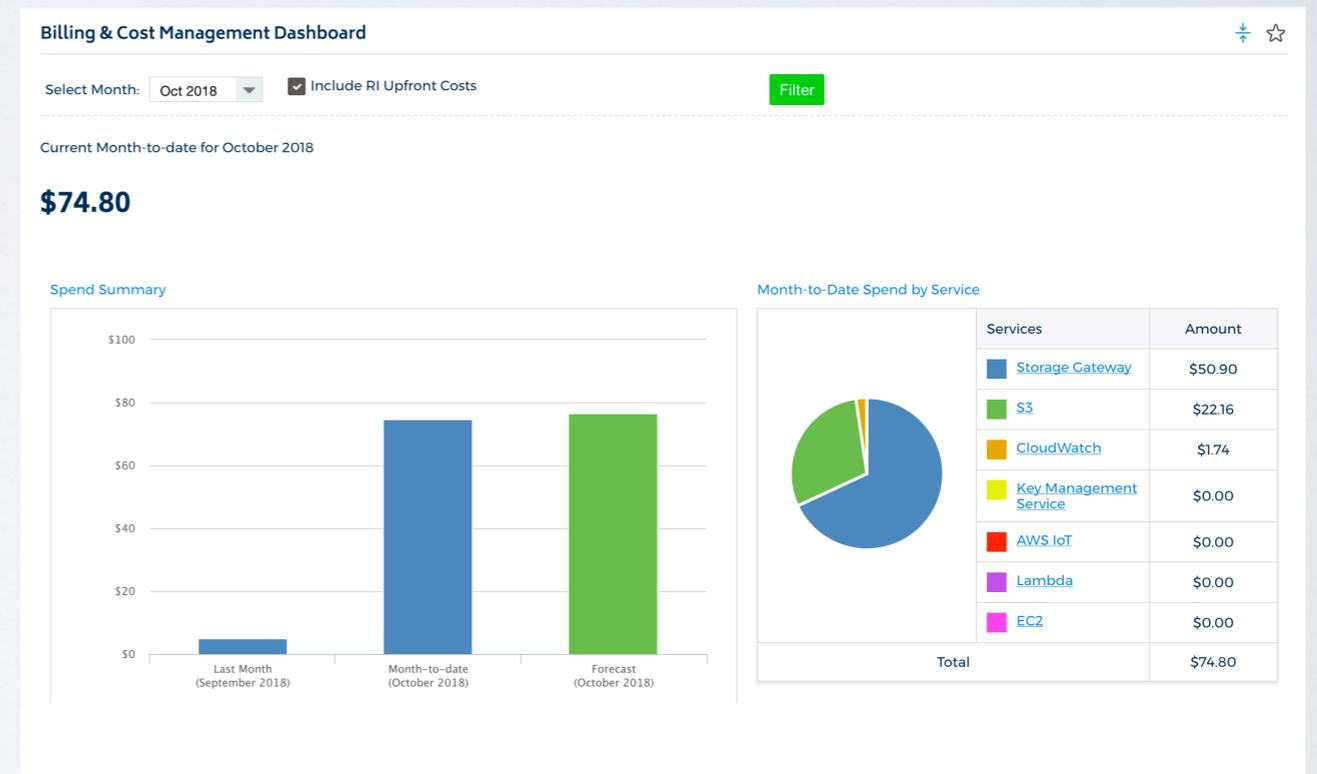
The screenshot displays the AWS Lambda console for the 'archive_backups' function. The top navigation bar shows the AWS logo, 'Services', 'Resource Groups', and user information. The breadcrumb trail is 'Lambda > Functions > archive_backups'. The function's ARN is 'arn:aws:lambda:eu-west-1:167381000869:function:archive_backups'. Below this, there are buttons for 'Throttle', 'Qualifiers', 'Actions', 'Test', and 'Save'. The 'Configuration' tab is active, showing a 'Designer' section with a list of triggers on the left (API Gateway, AWS IoT, Alexa Skills Kit, Alexa Smart Home, CloudWatch Events, CloudWatch Logs, CodeCommit) and a central workspace where 'archive_backups' is connected to 'CloudWatch Events'. On the right, there are boxes for 'Amazon CloudWatch Logs', 'Amazon Resource Group Tagging API', and 'Amazon S3'. Below the Designer, the 'Function code' section shows 'Code entry type' set to 'Edit code inline', 'Runtime' set to 'Python 3.6', and 'Handler' set to 'lambda_function.lambda_handler'. The code editor displays Python code for the lambda handler, including a function to get the bucket and a try-catch block for handling errors.

EVALUACIÓN DE COSTES

- Coste de almacenamiento de 1TB:
 - 4,1 USD al mes (491,52 USD durante 10 años).
 - 336€ más IVA en hierro local (sólo equipamiento).
- Nuestro archivo de 440TB almacenado en Glacier sale por unos 1800 USD al mes.
- Si a eso añadimos nuestros datos en crudo desde diciembre 2013, salen unos 655TB por un coste de 2685€ mensuales.

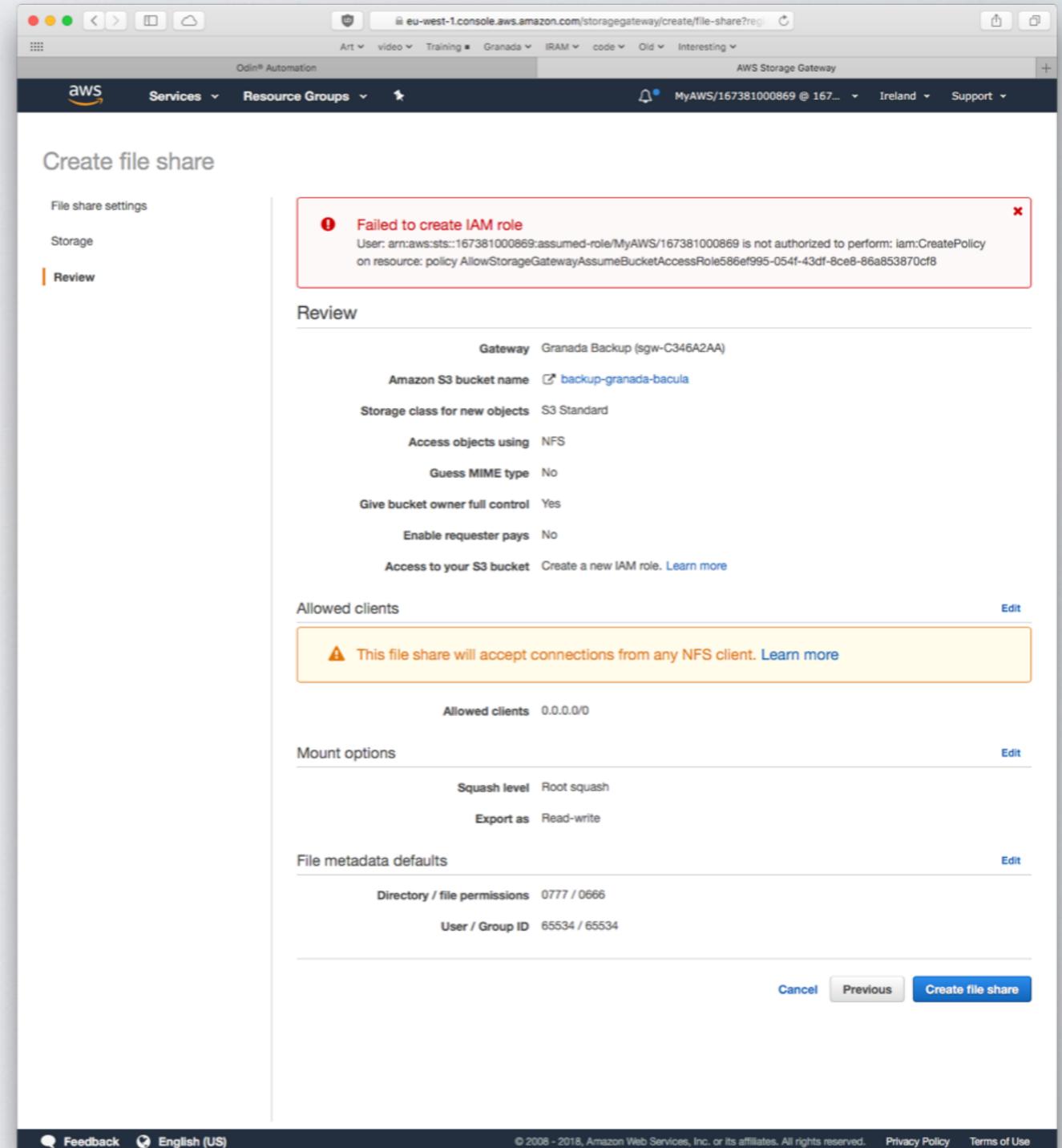
SPARKLE

- Proveedor de servicios AWS dentro del acuerdo marco de GÉANT.
- Sparkle y Comparex fueron los proveedores que contactamos y nos quedamos con el que fue más rápido en gestionarnos el alta.



INCONVENIENTES DE SPARKLE

- Tener un proveedor que intermedio facilita la facturación pero introduce otros inconvenientes:
 - No tienes permisos de acceso a los detalles de gasto que ofrece AWS.
 - La creación de nuevos usuarios tienes que hacerla a través de tu proveedor.
 - Apareció un error como el de la captura al intentar crear un rol automático. Resuelto.



RECURSOS

- Curso online: AWS Cloud Practitioner Essentials (Digital)

<https://www.aws.training/learningobject/curriculum?id=16357>

- Canal de YouTube de Miguel Arranz (VideoCursosCloud)

<http://www.youtube.com/c/MiguelArranz>

<http://videocursoscloud.com/>

CONCLUSIONES

- El almacenamiento en la nube puede salir más barato a largo plazo.
- Desplegar infraestructura en la nube puede ser una solución muy eficaz. Pago por uso y alta disponibilidad.
- Requiere de cierto aprendizaje pero puede ser muy rentable profesionalmente.
- Es crítico disponer de una buena conexión a internet.

ask no
questions
and
be told
no lies

¿ PREGUNTAS ?