

Observabilidad de infraestructuras y aplicaciones con *AWS*

UPC - Cloud Computing Architecture

\$ whoami



Marc Catrisse

- Ingeniero informático y Máster en Innovación y Investigación en la FIB
- Profesor asociado en la FIB
- Responsable técnico de proyectos y DevOPS en inLab FIB
 - Especializado en cloud (AWS)

<https://www.linkedin.com/in/marc-catrisse-99b065128/>

<https://inlab.fib.upc.edu/persones/marc-catrisse/>

Índice - ¿Qué servicios vamos a ver?

- CloudWatch
 - Logs
 - Metrics
 - Alarms
- EventBridge
- CloudTrail
 - Athena
- AWS Config
- X-Ray
- Cloudwatch Synthetics
- Servicios externos

Monitorización vs Observabilidad

- **Monitorización**

- Supervisión de métricas predefinidas con el objetivo de detectar errores en servicios específicos
- Debemos conocer las métricas clave de antemano
- Reactiva

- **Observabilidad**

- Indica la capacidad que tenemos de sacar conclusiones sobre el estado de un sistema a partir de los datos que devuelve
- Nos ayuda a comprender mejor cómo funciona nuestro sistema
- Nos permite anticiparnos a problemas desconocidos
- Proactiva
- Significado más global de sistema
 - Sistemas complejos
 - Arquitectura de microservicios (+ Servicios)

Datos a monitorizar

- Métricas

- Medidas periódicas de KPIs (Key Performance Indicators)
 - Uso de CPU
- Información agregable

- Logs

- Registros de eventos + Timestamp
- Ejemplos
 - Access Log de un servidor web

- Trazas

- Orientado a arquitecturas de microservicios
- Comportamiento de las peticiones
- Ejemplos
 - Latencias específicas

¿Qué métricas son importantes?

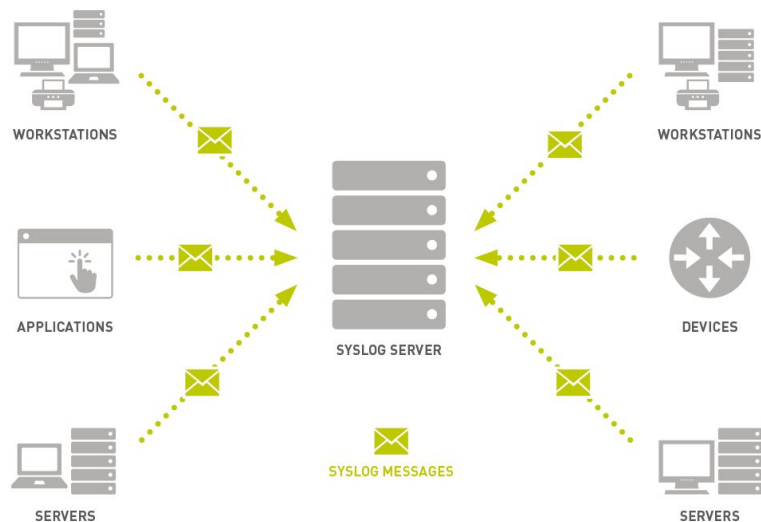
- Servicios centrados a usuarios (user-centric)
 - **Four Golden Signals**
 - **Latencia:** Tiempo que tarda en procesar una petición
 - **Tráfico:** ¿Cuanta demanda tiene el servicio? (req / s)
 - **Errores:** Ratio de peticiones que fallan (req / s)
 - **Saturación:** Carga del sistema (memoria en uso)
- Infraestructura
 - **USE**
 - **Utilización:** % CPU
 - **Saturación:** Trabajos (jobs) en cola
 - **Errores:** Número de eventos de error

Logging básico en Linux

- Linux dispone de un directorio especial para almacenar logs (/var/log)
 - **/var/log/syslog**
 - Actividad global del sistema incluyendo mensajes de arranque
 - **/var/log/auth.log**
 - Eventos de seguridad (logins)
 - **/var/log/kern.log**
 - Eventos del Kernel, errores y warnings
 - **/var/log/cron**
 - Información referente a las tareas programadas vía cron
 - **Aplicaciones externas**
 - **/var/log/apache2**
 - **/var/log/mysql**

Logging básico en Linux

- Syslog
 - Estándar para la generación de logs y transmisión de los mismos
 - Se define en el RFC 5424
 - Cómo transmitir estos datos (puerto 514 y 6514)
 - Formato de los logs
 - rsyslog / syslog-ng
 - Implementaciones de syslog
 - Servicios de procesamiento de logs



Logging básico en Linux

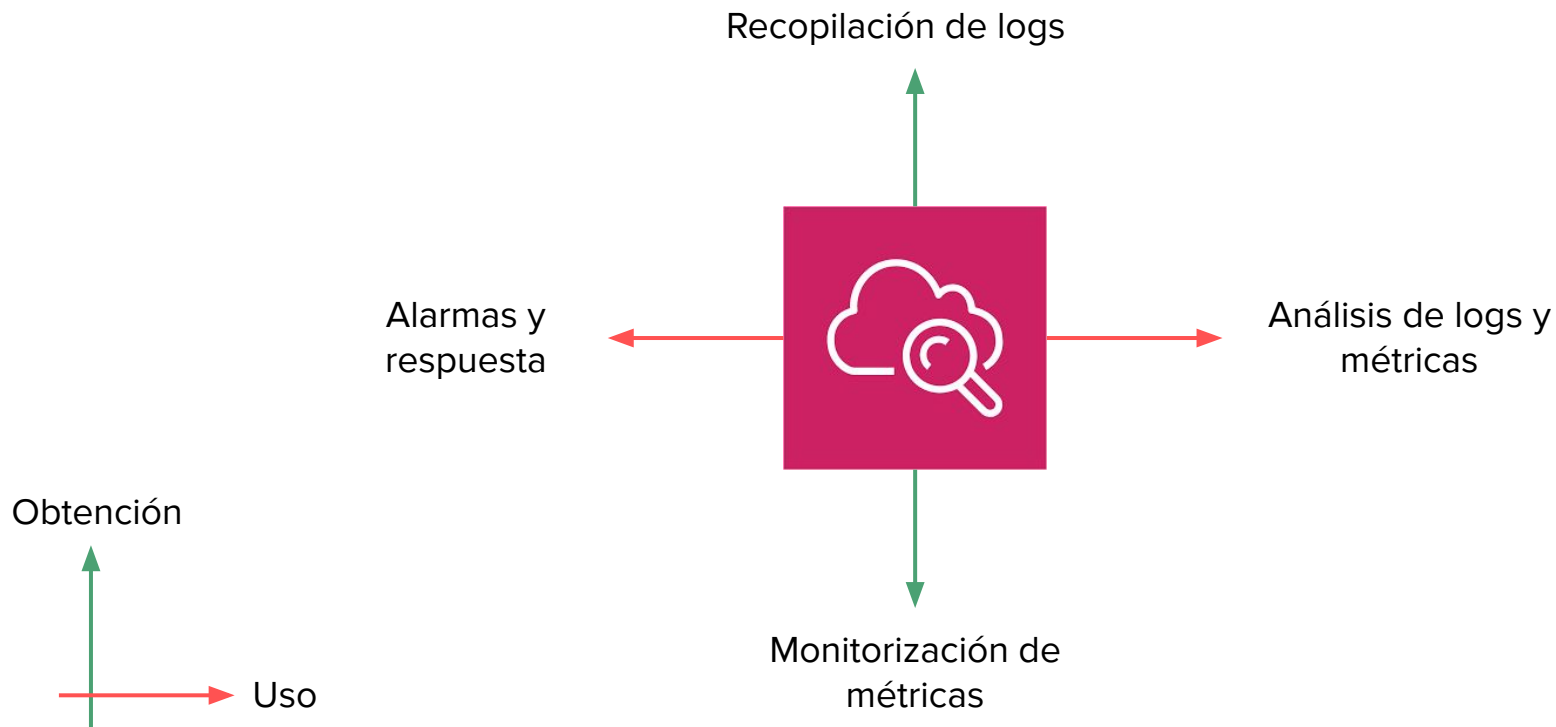
PRI	TIMESTAMP	HOSTNAME	MSG (TAG)	MSG (CONTENT)
<34>	Oct 11 22:14:15	mymachine	su:	'su root' failed for lonvick on /dev/pts/8

- PRI
 - Indica la prioridad del evento
- TIMESTAMP
 - fecha de generación
- HOSTNAME
 - origen
- MSG (TAG)
 - Indica el proceso origen o PID
- MSG (Content)
 - Contenido

CloudWatch

- Nos permite monitorizar recursos y aplicaciones en tiempo real
- Los servicios de *AWS* disponen de métricas propias
 - EC2 uso de CPU
 - RDS espacio libre de almacenamiento / IOPS
- Casos de uso:
 - Controlar recursos y rendimiento de aplicaciones (métricas)
 - Agrupar y monitorizar ficheros de logs
 - Crear y disparar alarmas
 - Basadas en métricas y logs

CloudWatch



CloudWatch

- **Conceptos clave:**

- Métricas
 - Estándar
 - Personalizadas
- Alarmas
 - Notificaciones

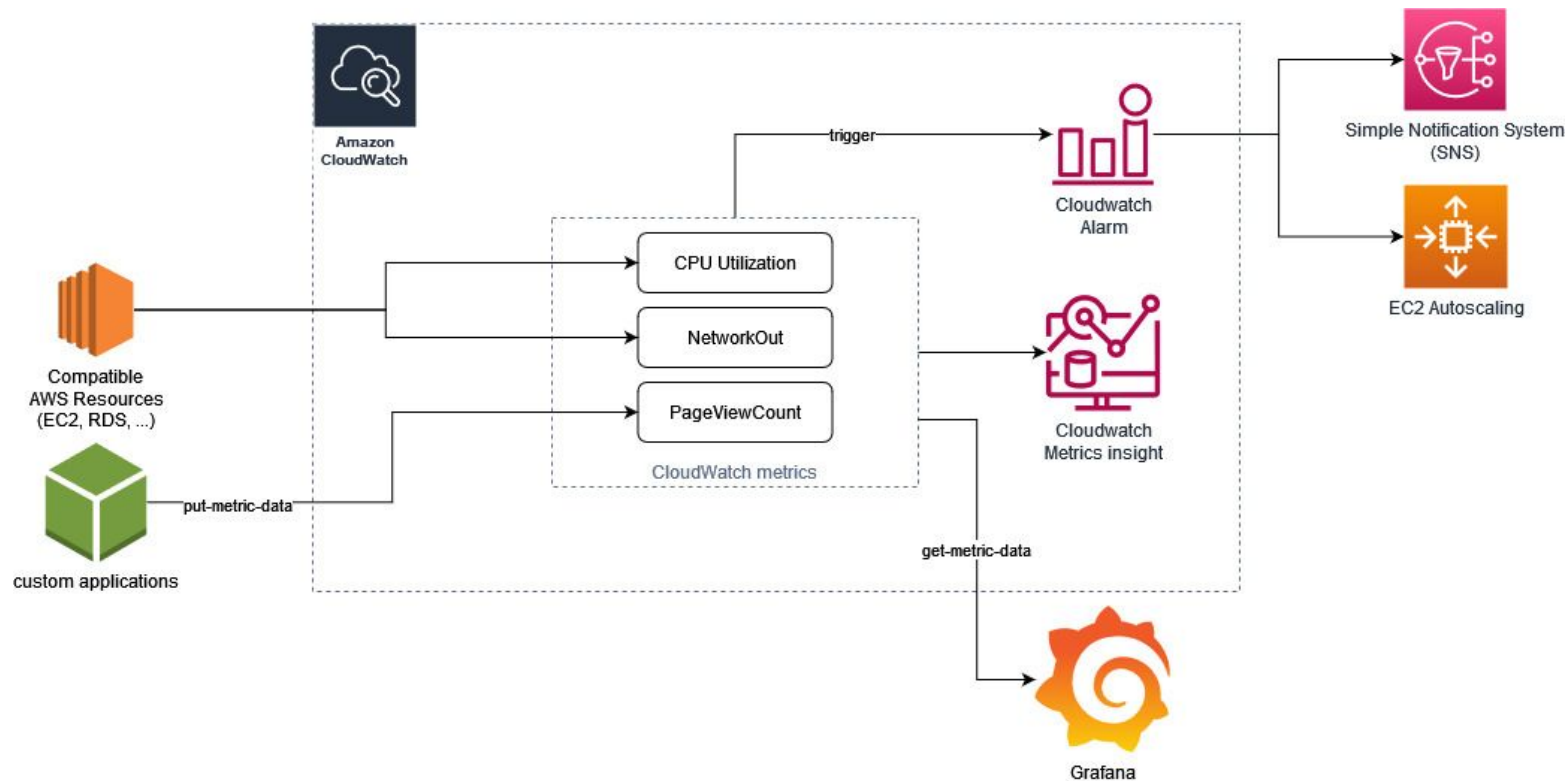
- **Limitaciones**

- CloudWatch por sí solo, no monitoriza a nivel de SO
 - Solo recolecta métricas de servicios (EC2, RDS, S3...)

- **CloudWatch agent:** recopila métricas a nivel de SO

- Espacio de disco disponible
- Métricas custom
- etc...

CloudWatch - Ejemplo - Caso de uso



CloudWatch - Métricas

Métrica: Clave + Valor + Timestamp

Namespace: Grupo de métricas relacionadas

Dimensión: Par clave valor que categoriza la métrica

MetricName + Dimension = Nueva métrica!

```
1  {
2    "Metrics": [
3      {
4        "Namespace": "AWS/S3",
5        "MetricName": "BucketSizeBytes",
6        "Dimensions": [
7          {
8            "Name": "StorageType",
9            "Value": "StandardStorage"
10         },
11         {
12           "Name": "BucketName",
13           "Value": "mySuperBucket"
14         }
15       ]
16     }
17   ]
18 }
```

\$ aws cloudwatch list-metrics --namespace "AWS/S3"

CloudWatch - Metrics - Resolución


- **Métricas de AWS**
 - Difieren según el servicio que estemos utilizando
 - **EC2**
 - **Estándar:** 5 minutos
 - **Detallada:** 1 minuto
- **Métricas personalizadas (custom)**
 - **Estándar:** 1 minuto
 - **Alta resolución:** < 1 minuto

CloudWatch - Metrics - Retención de datos

A medida que pasa el tiempo, las métricas van perdiendo resolución, agregando los datos para ser almacenados:

- Periodo de < **60 segundos** (métricas custom de alta resolución)
 - Durante **3 horas**
- Periodo de **60 segundos**
 - Durante **15 días**
- Periodo de **300 segundos (5 min)**
 - Durante **63 días (2 meses)**
- Periodo de **3600 segundos (1 hora)**
 - Durante **455 días (15 meses)**

Importante! Las métricas que no han tenido ninguna actualización no aparecen en la consola



get-metric-data / get-metric-statistics

CloudWatch - Metrics - Precio

- **Capa gratuita:**

- Todas las Métricas estándar enviadas por AWS (frecuencia de 5 minutos)
- 10 métricas de monitoreo detallado y custom (frecuencia de 1 minuto)
- 1 millón de solicitudes API (no aplicable a GetMetricData, GetMetricWidgetImage, GetInsightRuleReport)

Capas	Costo (métrica/mes)
Primeras 10 000 métricas	0,30 USD
Siguientes 240 000 métricas	0,10 USD
Siguientes 750 000 métricas	0,05 USD
Más de 1 000 000 de métricas	0,02 USD

API

GetMetricData, GetInsightRuleReport	0,01 USD por 1000 métricas solicitadas
GetMetricWidgetImage	0,02 USD por 1000 métricas solicitadas
Solicitudes de GetMetricStatistics, ListMetrics, PutMetricData, GetDashboard, ListDashboards, PutDashboard y DeleteDashboards	0,01 USD por cada 1000 solicitudes

CloudWatch - Metrics - Precio

- **Ejemplo 1 - Monitorización de EC2 Detallada:**

- Suponemos 10 VM EC2
- Cada VM EC2 dispone de 7 métricas
- Métricas totales = $7 * 10 = 70$
- $0,3 \text{ USD} * 70 = \mathbf{21 \text{ USD al mes!}}$

CloudWatch - Metrics - Precio

- **Ejemplo 2 - Métricas personalizadas:**

- Suponemos 100 VM EC2
- Publican 5 métricas personalizadas con el Agente CloudWatch cada minuto
- **Métricas**
 - $5 * 100 = 500$ métricas
 - $500 * 0,3 \text{ USD} = \mathbf{150\text{USD}}$
- **Peticiones**
 - $100 * (43200 \text{ minutos}) = \mathbf{4.320.000}$ peticiones al mes
 - 1.000.000 entran dentro de la capa gratuita
 - $3.320.000 / 1000 * 0.01 = \mathbf{33 \text{ USD}}$
- **Total = 33 + 150 USD = 183 USD**



EC2 - Status Checks

- System
 - Comprueba el hardware (sistema de AWS) sobre el que corre nuestra instancia
 - Requiere de la intervención de AWS / Podemos optar por reiniciarla
- Instance
 - Comprueba el software y la configuración de red de la instancia
 - El problema es nuestro

The screenshot displays the AWS Management Console interface for EC2 instances. At the top, a table lists three instances: i-0c0186a12aab3741d (t2.large), i-0138edcaf722db475 (m4.large), and i-02c65b735153975ec (t3.medium). All instances are in the 'Running' state. The first instance, i-0c0186a12aab3741d, has a warning icon in the 'Status check' column, indicating a failure. Below the table, the console shows the details for instance i-0c0186a12aab3741d, with the 'Status checks' tab selected. This tab shows two categories of checks: 'System status checks' and 'Instance status checks'. The 'System status checks' section shows a green checkmark for 'System reachability check passed'. The 'Instance status checks' section shows a red warning icon for 'Instance reachability check failed', with a message indicating the failure occurred at 2020/12/16 17:30 GMT+2 (about 1 month).

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Avail	
<input checked="" type="checkbox"/>	-	i-0c0186a12aab3741d	Running	t2.large	1/2 checks ...	No alarms +	eu-w
<input type="checkbox"/>	-	i-0138edcaf722db475	Running	m4.large	2/2 checks ...	No alarms +	eu-w
<input type="checkbox"/>	-	i-02c65b735153975ec	Running	t3.medium	2/2 checks ...	No alarms +	eu-w

Instance: i-0c0186a12aab3741d

Details | Security | Networking | Storage | **Status checks** | Monitoring | Tags

Status checks [Info](#)

Status checks detect problems that may impair i-0c0186a12aab3741d from running your applications.

System status checks

- ✓ System reachability check passed

Instance status checks

- ⚠ Instance reachability check failed

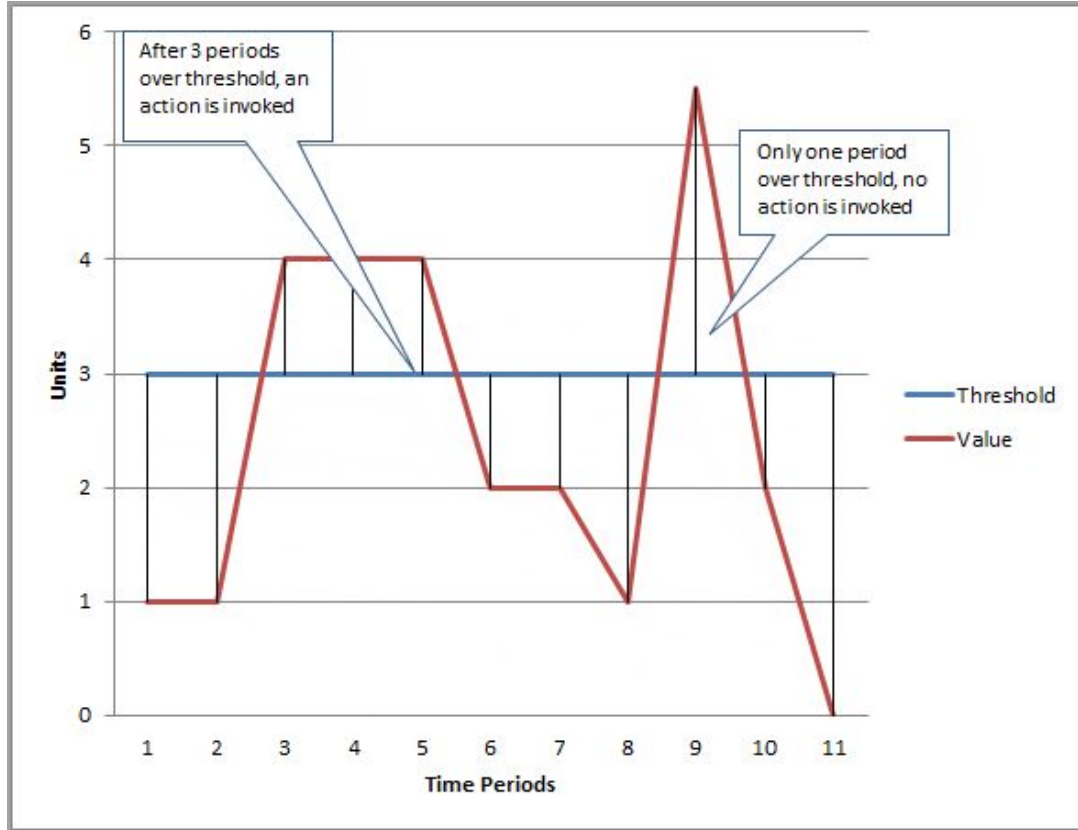
Check failure at

2020/12/16 17:30 GMT+2 (about 1 month)

CloudWatch - Alarms

- 3 estados:
 - OK
 - No se ha excedido el umbral
 - Alarma
 - Se ha excedido el umbral
 - Datos insuficientes
 - La alarma se acaba de crear, o aún no hay datos para hacer una evaluación

Cloudwatch - Alarms - Ejemplo



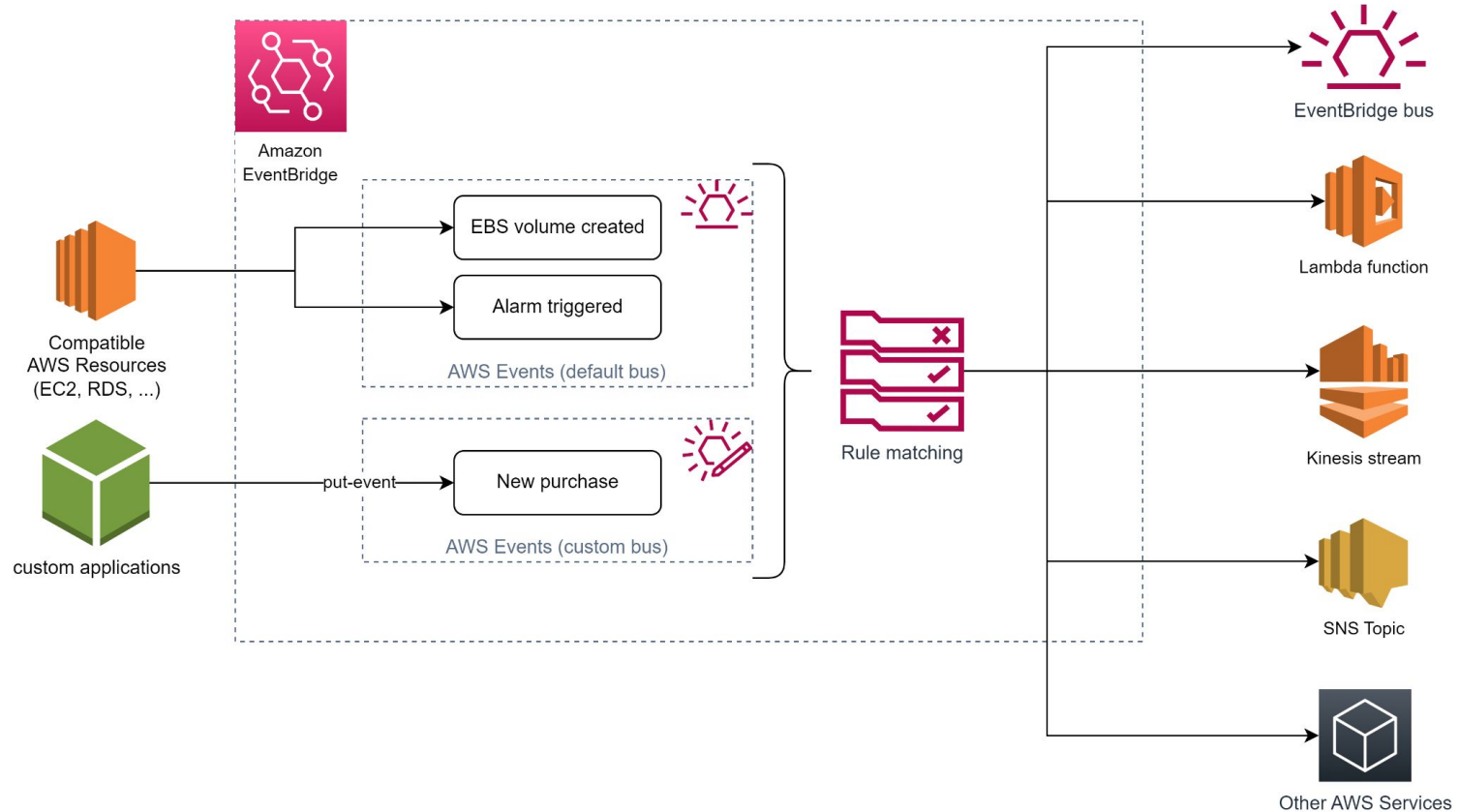
- **Period**
 - Ejemplo: 1 minuto
- **Evaluation period**
(Periodo de evaluación)
 - Ejemplo: 3 periodos
- **Datapoints to Alarm**
(Puntos de datos para la alarma)
 - Ejemplo: 3 unidades

Events (EventBridge)

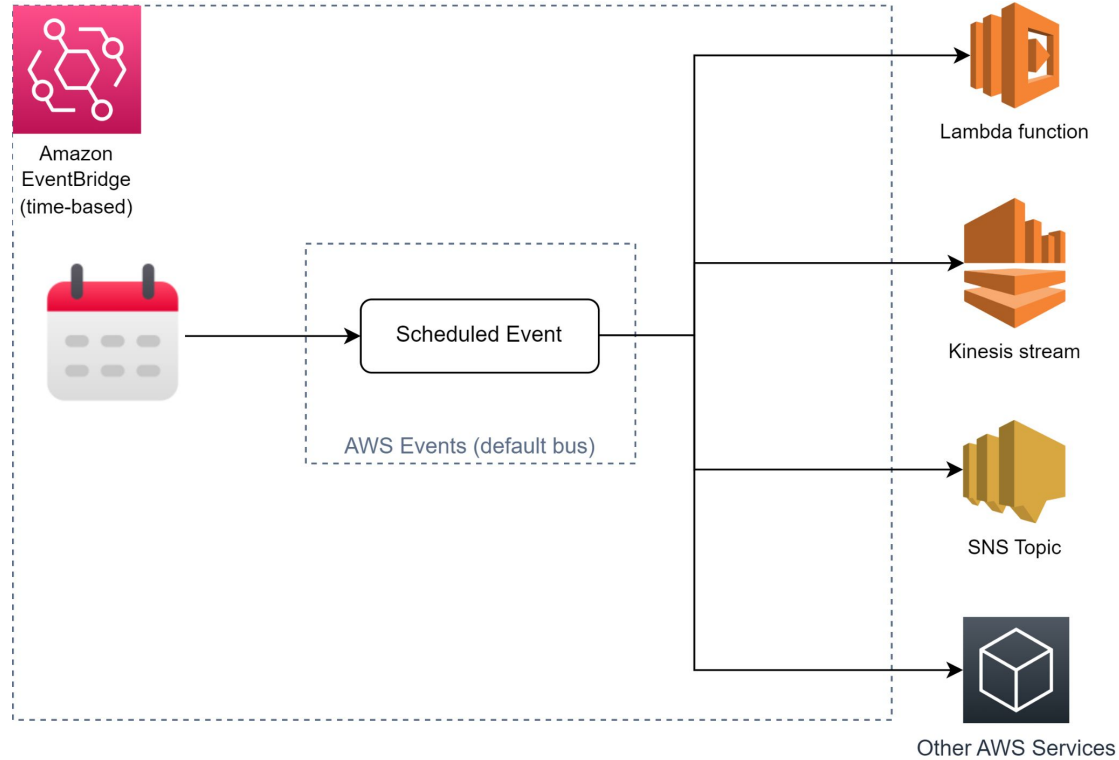


- Captura eventos generados en nuestra cuenta *AWS* y actúa acorde:
 - Creación de un volumen EBS
 - Activación de una alarma
- **EventBridge** es un bus de eventos sin servidor que facilita la creación de aplicaciones basadas en eventos

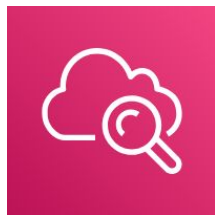
Events (EventBridge) - Ejemplo



Events (EventBridge) - Ejemplo CRON



CloudWatch Logs



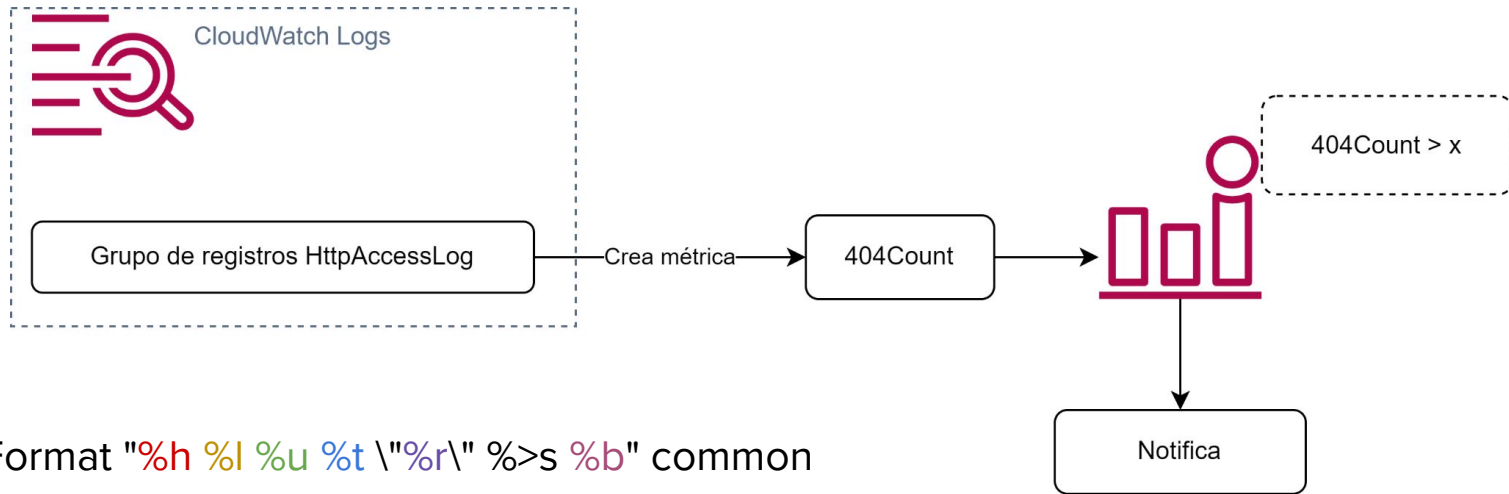
- CloudWatch logs nos permite monitorizar y almacenar archivos de registro de:
 - Instancias EC2 (CloudWatch Agent)
 - Servicios de AWS
 - AWS Route 53 / AWS Lambda / AWS CloudTrail
 - Otros
- Funcionalidades
 - Recopilación automática de registros
 - Agregación de datos en **grupos de registros**
 - Capacidad de configurar **filtros de métricas**
 - Búsqueda de patrones (Errores en el Access Log de Apache)
 - Crear métricas en base a registros
 - Consulta de registros y creación de visualizaciones con **CloudWatch Logs Insights**
 - Consulta los registros en tiempo real con **CloudWatch Live Tail**

Cloudwatch - Ejemplo - CloudWatch Logs Insights

```
1 fields @timestamp, @message
2 | filter Severity="ERROR"
```

#	@timestamp	@message																								
▼ 1	2021-11-09T06:54:13.38...	{"Severity": "ERROR", "message": "This is where the message detail would go", "IP Address..."}																								
<table><tr><th>Field</th><th>Value</th></tr><tr><td>@ingestionTime</td><td>1636458856910</td></tr><tr><td>@log</td><td>231551015239:application.log</td></tr><tr><td>@logStream</td><td>i-00985c85e1895c55a</td></tr><tr><td>@message</td><td>{"Severity": "ERROR", "message": "This is where the message detail would go", "IP Address": "10.5.23.124", "1"</td></tr><tr><td>@timestamp</td><td>1636458853386</td></tr><tr><td>Browser</td><td>Firefox v11</td></tr><tr><td>IP Address</td><td>10.5.23.124</td></tr><tr><td>message</td><td>This is where the message detail would go</td></tr><tr><td>Severity</td><td>ERROR</td></tr><tr><td>Timestamp</td><td>2021-11-09T11:54:13.291Z</td></tr><tr><td>User ID</td><td>joe.user@gmail.com</td></tr></table>			Field	Value	@ingestionTime	1636458856910	@log	231551015239:application.log	@logStream	i-00985c85e1895c55a	@message	{"Severity": "ERROR", "message": "This is where the message detail would go", "IP Address": "10.5.23.124", "1"	@timestamp	1636458853386	Browser	Firefox v11	IP Address	10.5.23.124	message	This is where the message detail would go	Severity	ERROR	Timestamp	2021-11-09T11:54:13.291Z	User ID	joe.user@gmail.com
Field	Value																									
@ingestionTime	1636458856910																									
@log	231551015239:application.log																									
@logStream	i-00985c85e1895c55a																									
@message	{"Severity": "ERROR", "message": "This is where the message detail would go", "IP Address": "10.5.23.124", "1"																									
@timestamp	1636458853386																									
Browser	Firefox v11																									
IP Address	10.5.23.124																									
message	This is where the message detail would go																									
Severity	ERROR																									
Timestamp	2021-11-09T11:54:13.291Z																									
User ID	joe.user@gmail.com																									
▶ 2	2021-11-09T06:54:13.38...	{"Severity": "ERROR", "message": "This is where the message detail would go", "IP Address..."}																								
▶ 3	2021-11-09T06:54:13.38...	{"Severity": "ERROR", "message": "This is where the message detail would go", "IP Address..."}																								
▶ 4	2021-11-09T06:54:13.38...	{"Severity": "ERROR", "message": "This is where the message detail would go", "IP Address..."}																								

CloudWatch - Logs - Creación de métricas a partir de logs



LogFormat "%h %l %u %t \"%r\" %>s %b" common

127.0.0.1 - Marc [01/Feb/2023:13:55:36 +0100] "GET /hello_world HTTP/1.0" 200 2326

CloudWatch - Logs - Precio

- **Capa gratuita:**

- 5GB (recopilación, almacenamiento, análisis...)
- 1800 minutos de uso de Live Tail al mes (aprox 1h al día)



Recopilación (captura de datos)

0,57 USD por GB

Almacenamiento (archivado)

0,03 USD por GB

Análisis (consultas de Logs Insights)

0,0057 USD por GB de datos escaneados

Detección y máscara (Data Protection)

0,12 USD por GB de datos analizados

Analizar (Live Tail)

0,01 USD por minuto

CloudWatch - Logs - Precio

- **Ejemplo 1 - Logs HTTP:**

- Suponemos que registramos 1 GB diario durante 30 días = 30 GB al mes
- Captura de datos
 - 0 a 5GB = 0 (Capa gratuita)
 - 5 a 30 GB = $0,57 \text{ USD} * 25 = \mathbf{14,25 \text{ USD}}$
- Almacenamiento
 - 0 a 5GB = 0 (Capa gratuita)
 - 5 a 30 GB = $0,03 \text{ USD} * 25 = \mathbf{0,8 \text{ USD}}$
- Total
 - $\mathbf{14,50 + 0,8 = 15,05 \text{ USD}}$

VPC Flow Logs



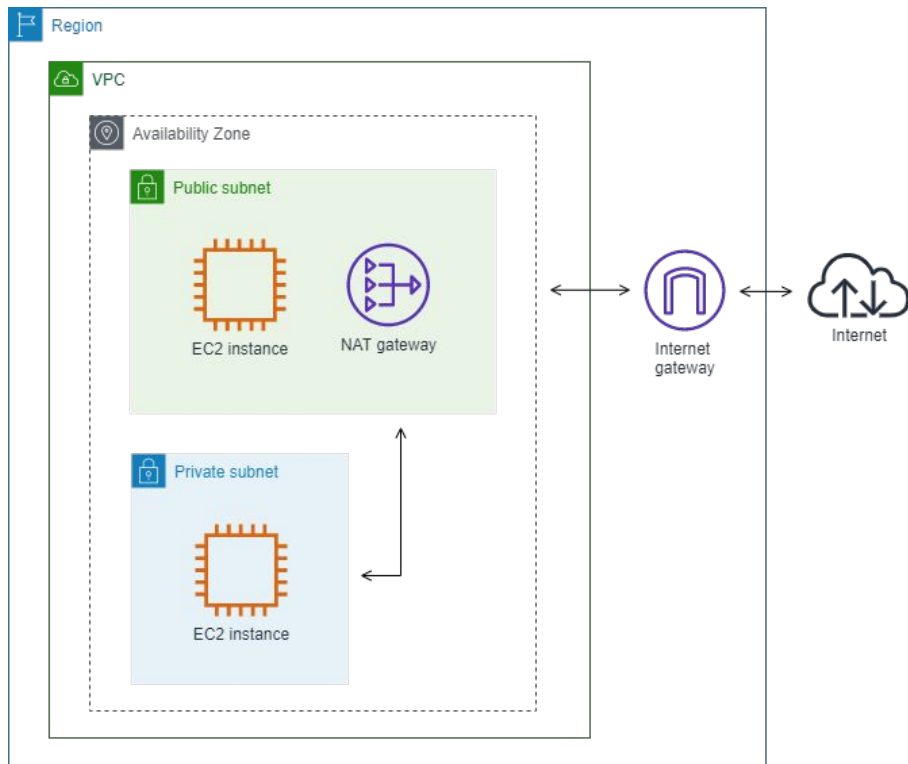
- Nos permite capturar información referente al **tráfico IP** que fluye a través de una o más interfaces de red
- Esta información puede distribuirse a
 - Bucket S3
 - Athena
 - CloudWatch Logs
 - Amazon Kinesis Data Firehose

VPC Flow Logs - Ejemplo

- Tráfico a través de un NAT Gateway

```
aws ec2 create-flow-logs \  
  --resource-type NetworkInterface \  
  --resource-ids eni-11223344556677889 \  
  --traffic-type REJECT \  
  --log-group-name my-flow-logs \  
  --deliver-logs-permission-arn arn:aws:iam::123456789101:role/publishFlowLogs
```

```
- eni-X 10.0.1.5 10.0.0.220 10.0.1.5 203.0.113.5
```



```
instance-id interface-id srcaddr dstaddr pkt-srcaddr pkt-dstaddr
```


VPC Flow Logs - Precio

	Envío a CloudWatch Logs	Envío a S3	Envío a Kinesis Data Firehose	
Datos recibidos				
Primeros 10 TB al mes	0,57 USD por GB	0,285 USD por GB	0,285 USD por GB	
Próximos 20 TB al mes	0,285 USD por GB	0,171 USD por GB	0,171 USD por GB	
Próximos 20 TB al mes	0,114 USD por GB	0,086 USD por GB	0,086 USD por GB	
Más de 50 TB al mes	0,057 USD por GB	0,057 USD por GB	0,057 USD por GB	
Datos almacenados	0,03 USD por GB	A partir de 0,023 USD/GB (Standard) hasta 0,00099 USD/GB (Glacier Deep Archive)		No se aplica
Formato convertido a Apache Parquet	N/D	0,034 USD por GB*	N/D	

Cloudwatch > Pricing > Logs > Vended Logs

Quizz!

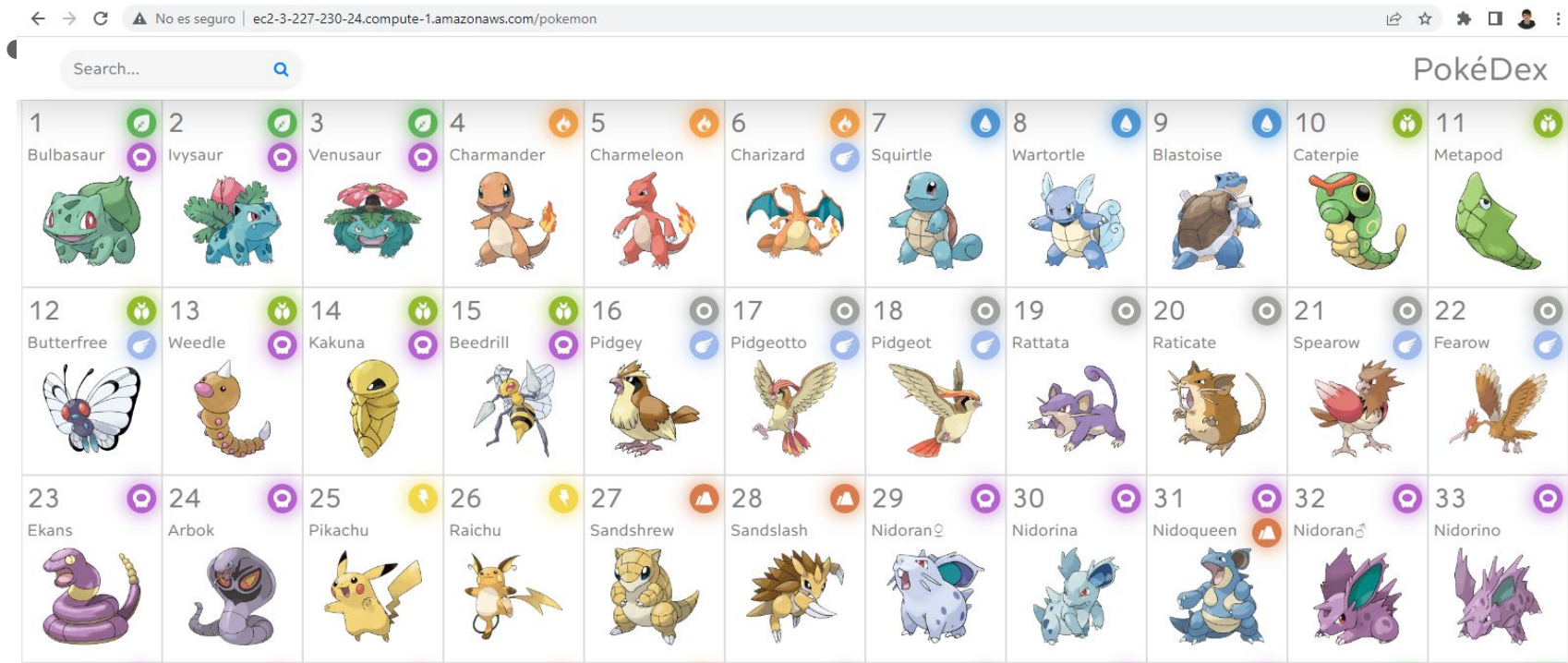


Lab 1 - Pokedex

- Pokemon Go ha vuelto a hacerse viral gracias a una nueva actualización
- El CEO ha decidido publicar una aplicación Web que ofrece el servicio de Pokedex
 - Objetivo: Conseguir aumentar la presencia de la marca en los círculos de jugadores
- La aplicación parece que está gustando aunque algunas personas se quejan de errores puntuales
- Problema! No la estamos monitorizando!
 - ¿Errores?



Lab 1 - Pokedex




Lab 1 - La Pokedex

← → ↻ ⚠ No es seguro | ec2-3-227-230-24.compute-1.amazonaws.com/pokemon/94

← PokéDex

GENGAR

Shadow Pokémon



EVOLUTION CHAIN

ID #94

Height 1.5m (4'11")

Weight 40.5kg (89.3lbs.)

Abilities CURSED-BODY

Type Ghost Poison

Forms GENGAR MEGA GENGAR GIGANTAMAX GENGAR

Base **Min** **Max**

HP	60	
Attack	65	
Defence	60	
Sp. Attack		130
Sp. Defence	75	
Speed		110
Total	500	

Lab 1 - La Pokedex

- Despliegue con **Terraform**
- Servidor NGINX
 - App Web Angular

