Architecture_for_kafka_cluster_MSK_

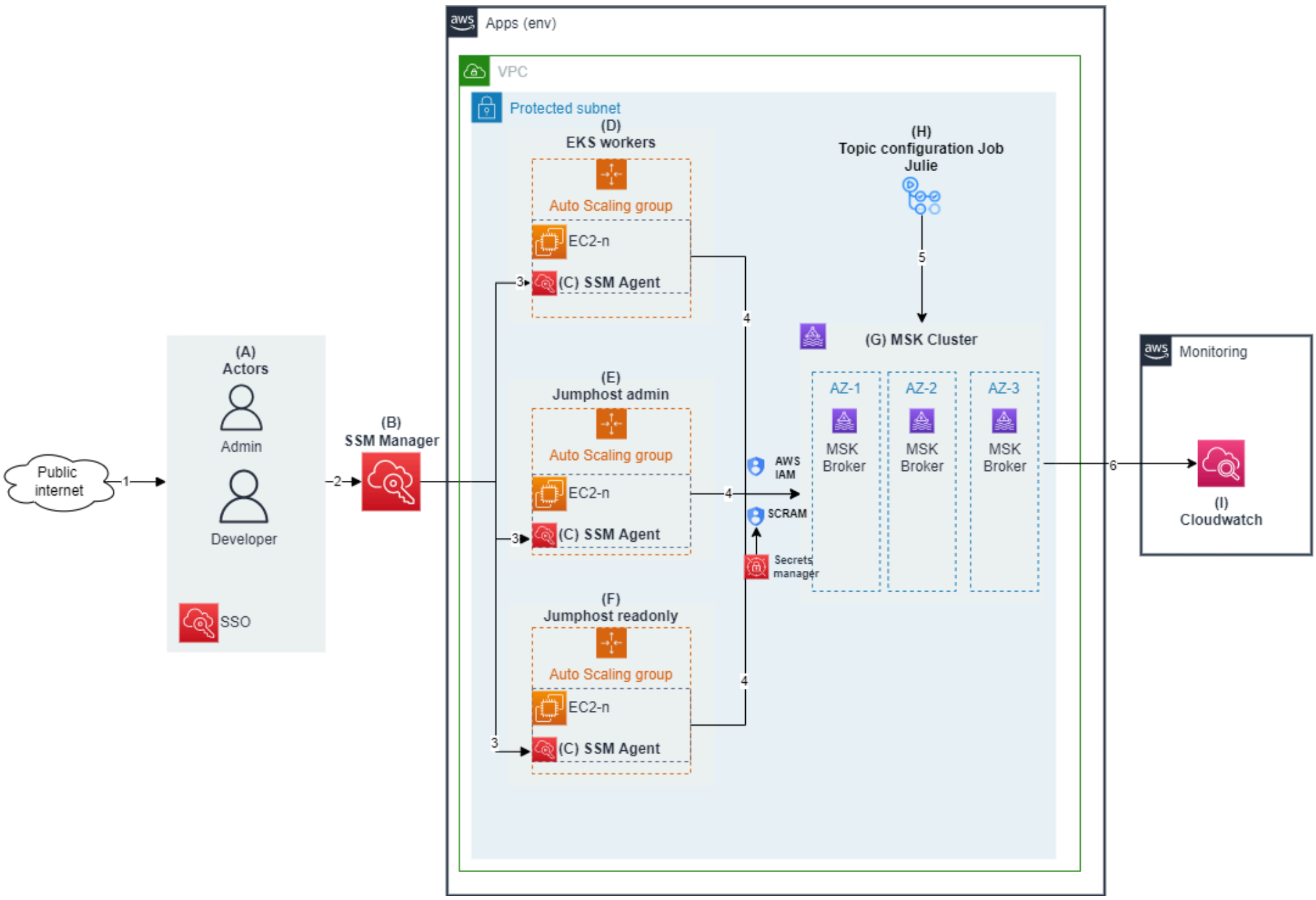
Last edited by **czerniga-gft** 8 months ago

Architecture for kafka cluster (MSK)

Important Parametres of MSK Cluster

- default.replication.factor 3 for 3-AZ clusters, 2 for 2-AZ clusters (default)
- kafka_number_of_brokers = 3 (default)
- num.partitions = 1 (default)
- auto.create.topics.enable = false (default)
- broker logs with short retention (3days) without alerts on logs in broker yet
- kafka_broker_instance_type = kafka.t3.small
- kafka_broker_volume_size = 10 (broke volume size in GiB)
- subnets: protected

Diagram of Architecture for MSK



Relevant elements of architecture

- **A** – Actors (Admin and Developer) - Accounts with different rights and accesses to resources

- **B** – Session Manager which allows to manage EC2 instances by communicating with SSM Agents **(C) **(without the need to open inbound ports, or manage SSH keys)
- **C** – Session Manager Agents - Software that runs on EC2 instances. It allows SSM to communicate with an instance
- **D** – Auto Scalling Group of Kubernetes workers
- **E** – Auto Scaling group of jumphosts with admin rights
- **F** – Auto Scaling group of jumphosts with readonly rights
- **G **– MSK Cluster containing multi-az, 3 brokers of kafka.t3.small, all az
- **H **– Topic configuration job (deployed on MSK cluster by Julie)
- **I **– Cloudwatch

- ## **Description of architecture/flow **

  - 1 – Developer or admin connecting with usage of SSO (or IAM) from Cloud CLI to specific roles **(A)** with specific rights
  - 2 – Specific users communicating with SSM Manager **(B) **
  - 3 – SSM Manager **(B) **mediating between actors and particular SSM Agent **(C)** in one of the EC2 auto scaling groups **(D) (E) (F)**
  - 4 – Achieving access** **to the MSK Cluster **(G)** with usage of Cloud IAM and SCRAM authentication
  - 5 – Topic configuration Job **(H)** is deploying topic schemas to MSK Cluster **(G)**
  - 6 – Cloudwatch **(I)** alerts are receiving metrics from MSK Cluster **(G)**

Attachments:

MSK Diagram.drawio.png (image/png)

Comments:

Hi Szczepankiewicz, Jarosław , as you know I am reviewing our TM assets in the light of usefulness for our customers.

I have few inputs for this asset. I think we should describe what problems this architecture solves e.g.

1. This diagram shows ways how an MSK cluster can be accessed by humans as well as machines securely.
2. It also "recommends" usage of SSM Manager to be able to access Jumphost over the internet, without having to open any inbound ports or managing SSH keys. We should add why Is SSM the best/better approach? What are the other ways to do it. Most customers would have something for PAM (privileged access management) for accessing jump servers, so is it possible to expand this architecture to integrate with any 3rd party PAM solution customer is using? Customers would most likely want that.
3. We should show integration of IAM with a 3rd party directory services. Any customer would want SSO these days thru their own profiles directory.
4. We should expand on what is the benefit of using "Julie" with Kafka. Is Julie a recommended approach? How does it compare to other options?
5. The logs of MSK are streamed to Cloudwatch. Shall we rather point that to "Observability architecture"

So, this diagram is more than just a Kafka Architecture because it also sheds light on various expects related to access of MSK, operations on MSK and observability.
This asset seems more like a "Kafka-Ops - how to access, operate and monitor MSK right way" paper.

Notes:

This asset is aligned APAC's second offering - Cloud Transformation.

Posted by pvga at Aug 16, 2023 10:05