White Paper: Under the Hood of Xosphere

At Xosphere, we do ***IT*** *better* by enabling your DevOps team to run your applications more efficiently. No matter what kind of applications you run, our intelligent orchestration technology finds the optimal balance between availability and cost. We dynamically move your applications from expensive On-Demand or Reserved instances to cost-effective Spot instances and back as the Spot market fluctuates, ensuring your applications are always up and running while minimizing cost.

Xosphere Installation and Configuration

Xosphere Instance Orchestrator is a cloud-native, self-hosted subscription software application. It installs into your Amazon Web Services (AWS) account using either a CloudFormation stack or a Terraform module and runs using Lambda functions.

Instance Orchestrator uses an opt-in design; it only executes on Auto-Scaling groups or individual instances that have explicitly been enabled via an AWS tag. Tags can be applied using any method or tool that is used within the organization to manage tags (for example, AWS Console, AWS CLI, AWS APIs, infrastructure-as-code platforms such as CloudFormation or Terraform, cloud management platforms, etc.). Once this enabling tag has been applied, Instance Orchestrator will automatically perform its management duties on an ongoing basis.

Additional tags can be applied to control various settings, such as:

- Specifying exactly what instance types are allowed in an Auto-Scaling group.
- Setting the bid multiplier for bidding on Spot instances.
- Enabling On-Demand instances to be replaced be equivalent burstable Spot instances.
- Indicating whether or not CloudWatch metrics will be published.
- Setting a baseline number or percentage of On-Demand instances to be running in an Auto-Scaling group.
- Indicating whether SNS notifications will be sent when Spot instances are launched in an Auto-Scaling group.
- Specifying whether or not unused Reserved instances should be taken into account when deciding whether or not to replace an instance with a Spot instance.

Xosphere Architecture and Procedure

Here's how Xosphere Instance Orchestrator runs on AWS with regard to Auto-Scaling groups:

1. Every couple of minutes, a CloudWatch event triggers the core Instance Orchestrator Lambda function to look for Xosphere-enabled Auto-Scaling groups (via tags). Within those groups, it looks at the dynamics

Xosphere Architecture and Procedure cont...

1. cont... of the Spot instance marketplace (price, availability, etc.), factors in the status of any Reserved instances already purchased (unless otherwise configured, these will never be replaced) and reviews the latest configuration and settings. It also will determine what instances would be applicable to the group based on necessary and available RAM, disk space and more, what instances are available in particular zones and more.

   If it's determined that a Spot instance would be more efficient and a replacement should be made, it proceeds to the next step. If not, the process stops until it automatically starts over again in a couple of minutes.

2. Instance Orchestrator launches a Spot instance to replace an On-Demand instance. This is done using the same AMI and user data scripts as AWS would use in a scale-out operation. All Spot instances are bid using the available On-Demand instance rate, unless otherwise configured, ensuring the lowest-available charge to the user.

3. Instance Orchestrator waits for the replacement Spot instance to be fully bootstrapped and passing EC2 and application-level health checks. Once that's done, Instance Orchestrator:

   - Attaches the new Spot instance to the Auto-Scaling group, so it looks like a scale-out operation just happened.
   - Detaches the On-Demand instance from the Auto-Scaling group.
   - Drains any load balancer connections.
   - Initiates the graceful shutdown of the operating system.
   - Terminates the On-Demand instance.

   Instance Orchestrator will only act on one instance per Auto-Scaling group at a time.

4. When a Spot termination notification is received via a CloudWatch event, the Terminator Lambda function triggers a scale-out operation on the Auto-Scaling group and begins to drain any connections on the terminating Spot instance.

5. Execution logs from Instance Orchestrator and terminator are written to CloudWatch logs.

6. SNS notifications are optionally sent when instances are launched or terminated.

7. Each Lambda function leverages its own IAM role with the least minimum privilege necessary to execute.

Sometimes, AWS will recall and terminate Spot instances. When this occurs, Instance Orchestrator will perform the opposite procedure, creating a new On-Demand instance and swapping over to it:

- When Instance Orchestrator gets an instance termination notice from AWS through a CloudWatch event, it triggers a scale-out operation on the Auto-Scaling group, so it creates an On-Demand instance itself.
- The Spot instance will be detached from the Auto-Scaling group, load balancers will be drained, and a graceful shutdown and termination will take place.

Diversification

Xosphere will use as many Spot instance pools (the different combinations of instance types and availability zones) as possible. This ensures that every Auto-Scaling group has a diverse set of instances being used – and that data is as safe as possible. By default, Instance Orchestrator will not allocate more than 20 percent of a single Auto-Scaling group on a single Spot instance pool.

Security

Xosphere Instance Orchestrator is a self-hosted subscription software that installs into an AWS account. Since Instance Orchestrator is not interacting with your account from a SaaS platform, there is no requirement to delegate any amount of access using IAM roles. Xosphere will have no remote access to your account, and all sensitive data within the account and generated by Instance Orchestrator will remain within the AWS account.

The only data not completely contained inside your AWS account are aggregated usage statistics shared with Xosphere for billing purposes. All instance usage data is available on the Xosphere dashboard, with no identifying information displayed.

In addition, Instance Orchestrator follows recommended best practices and leverages a series of IAM roles that ensures that each Lambda function executes with the least minimum privileges required to perform its duties. This means that the CloudFormation stack or Terraform modules used to install Instance Orchestrator will require being run with a user account that has permissions to create IAM roles.

AWS Lambda runs your function code securely within a VPC by default. Except for Kubernetes integration (as described below), Instance Orchestrator does not interact directly with any underlying AWS instance, so no configuration is required to enable the Lambda functions to access resources inside your private VPC. Instead, Instance Orchestrator will leverage the permissions from its IAM role and interact with the AWS API to manage AWS resources within the account.

Stateful Applications

There are two primary differences between how Xosphere is used with stateless applications, as described above, and stateful applications.

Stateful Applications cont...

1. In use cases when Auto-Scaling groups don't exist, you can create an Xogroup instead – a cluster of like-minded instances that perform the same operations as each other. Each individual instance within the Xogroup has to be tagged with the Xogroup name for proper detection. Xogroup configuration settings can be configured in the Xosphere dashboard, or by applying configuration tags to any one of the instances within the Xogroup.

2. When instances are replaced, the process looks like a rolling restart instead of a scale-out operation.

    - Snapshots are taken of any EBS volume running in an Xogroup instance every 15 minutes. To save disk space, only the three most recent snapshots are kept.
    - Xosphere preserves key instance attributes such as IP addresses (private and Elastic IPs), load balancer relationships, EBS volumes (both root and data volumes), Elastic Network Interfaces, Elastic GPUs and more.
    - Load balancers are drained, and the instance is gracefully shut down.
    - The replacement instance is launched, the snapshot is applied, health checks are completed and operations resume as normal.

Containers

Xosphere is an ideal fit for container environments. Almost all of the above descriptions and procedures that apply to Auto-Scaling groups also apply for containers, with two additions:

- The process of draining instances supporting containers has an additional step. After load balancer connections are gracefully drained, Xosphere will automatically detect if the terminating instance is a Kubernetes Node or an ECS Container Instance. In either case, Xosphere will automatically interface with the container orchestration platform to ensure the Kubernetes Pods or ECS Tasks are drained and rescheduled to run on other Nodes or Container Instances in the cluster.
- In the case of Kubernetes, the Xosphere installation process requires two additional steps. Xosphere requires a DaemonSet to be deployed to the Kubernetes cluster. Installation of the DaemonSet is simple and can be accomplished by running a single kubectl command. In addition to the DaemonSet, one of the Xosphere Lambda functions must be placed in a VPC that has network access to the target Kubernetes Nodes. These steps are only required for Kubernetes; in the case of ECS, Xosphere leverages the AWS API, so these steps are not required for ECS.

xosphere™

Intelligent Cloud Orchestration

The Value and Benefits of Using Xosphere

Instance Orchestrator delivers a wide range of benefits and can have a transformative impact on your organization. They include:

Time Savings and Ease of Use

- Our platform requires no changes and works out-of-the-box in less than two minutes, enabling you to save immediately. In addition, we enable you to manage and configure your entire cloud infrastructure using existing processes and tools, not by toggling back and forth between multiple interfaces.

Security

- Xosphere has no visibility into any sensitive information stored within your account, and as such doesn't expose it to outside risks such as corruption, hacking, theft or worse.

Minimized Risk

- Our technology executes without replacing native AWS functionality such as Auto-Scaling groups, so your applications' uptime and availability always remain in your control. We also mitigate risk by using varied Spot pools, so if one becomes unavailable for any reason your applications will keep running.

High-Touch Support

- Every Xosphere customer service and support resource is based inside the United States, so we deliver the optimal human touch in the quickest time possible.

Cost Savings

- Our unique pricing model will save you up to 80% off On-Demand instance costs, or up to 40% off Reserved instance costs. You'll find our approach is fairer to your environment and less expensive than anyone else's.

About Xosphere

Xosphere and its groundbreaking Instance Orchestrator technology exist to optimize and improve efficiencies in your computing environment. Simply put, we do ***IT*** *better*.

Contact: galbright@cloudinsyte.com for more info or a free trial.

White Paper: Under the Hood of Xosphere