

VMware Cloud Services - VMware Cloud on AWS

by Christopher Lewis

I recently had the privilege of attending my second consecutive VMworld Europe in Barcelona. My key focus for this event was to learn as much as I could about the new VMware® Cloud Services. In this article, I focus on VMware Cloud on AWS (VMConAWS), providing an overview of what it is (and, more importantly, what it isn't), minimum sizing, pricing models and why/how VMConAWS could be the next big thing for your business along with what I see could be the main barriers to adoption.

Full Disclosure – I haven't had the opportunity to get "hands on" with VMConAWS so this review is based purely on the information that is publicly available. There is a recently released VMware Hands-On Lab (HOL) available for those who want to take a look at VMConAWS, HOL-1887-01-EMT.

Background

At VMworld 2016, VMware CEO Pat Gelsinger introduced a game changing announcement in VMware's business and cloud strategy, a partnership between VMware and Amazon Web Services (AWS). This was an important move for VMware, because at the time, AWS were (and continue to be) the market leader within the Cloud Infrastructure as a Service (World-Wide) Gartner Magic Quadrant.

Fast forward 12 months to VMworld 2017 and as part of his Day 1 Key Note, Pat Gelsinger announced that VMware Cloud on AWS (VMConAWS) was ready for initial release within the AWS US West Region.

"The partnership of VMware and AWS is a game changing move"

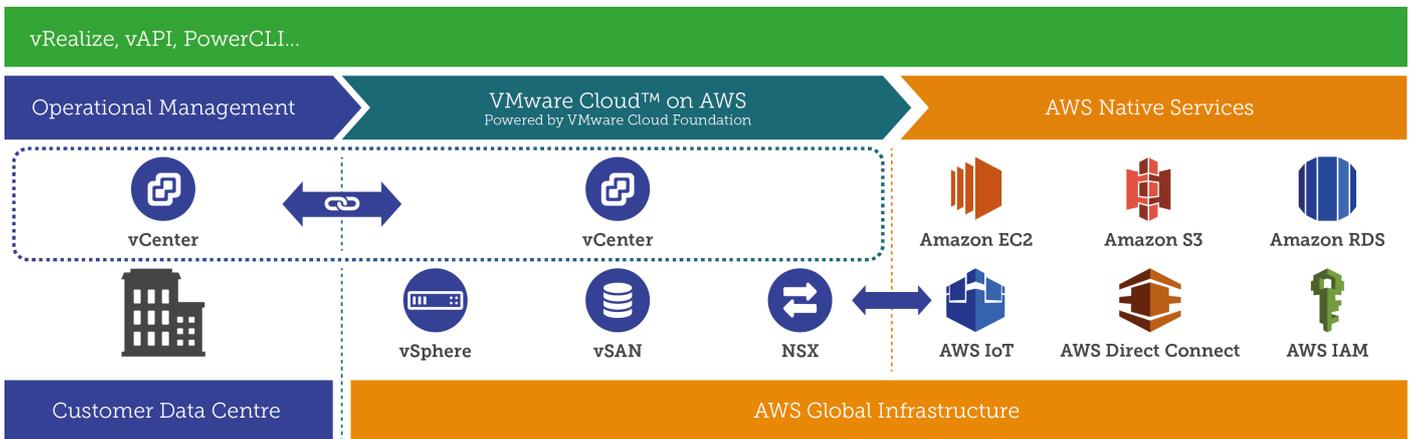
What is VMConAWS?

VMConAWS is the VMware vSphere® hypervisor deployed directly on to the AWS bare-metal hardware service within the AWS Global Infrastructure. VMConAWS leverages VMware's core technologies including VMware NSX® and VSAN™ and is managed by a dedicated VMware vCenter® Server to provide a full VMware SDDC implementation within a dedicated single tenant AWS account, using a single **AWS Virtual Private Cloud (VPC)**. VMConAWS is not nested virtualisation where the VMware ESXi™ hypervisor is installed within an AWS Elastic Cloud Compute (EC2) instance.

What is an Amazon Virtual Private Cloud (VPC)?

An Amazon Virtual Private Cloud (VPC) is a dedicated virtual network configured within AWS to allow the consumption of AWS services. A VPC is logically isolated from other virtual networks within the AWS cloud.

Within the context on VMConAWS, the **Customer VPC** is an existing or newly created standard Amazon VPC where AWS Native services are configured and consumed. The **Provider VPC** is the VMware owned and managed VPC that contains the VMConAWS solution for a specific customer.



The fact that VMConAWS is hosted within the AWS Global Infrastructure means that it is uniquely placed to take advantage of the enhanced capabilities of the underlying AWS hardware platform. VMware have been able to introduce technologies such as **Elastic DRS and Automated Cluster Remediation** whilst also providing tight integration into AWS Native Services (such as AWS EC2 and AWS S3). VMware are not only providing customers with the ability to manage their on-premises VMware infrastructure from within their VMware vCenter instance on VMConAWS through **Hybrid Linked Mode**, but are also supporting the ability to perform cold migration of workloads from on-premises to VMConAWS. It is capabilities like this that means VMConAWS is truly re-defining the story of hybrid cloud. In addition, as this is just another VMware vSphere environment, the VMware vRealize® Suite Cloud Management Platform can be used to monitor, manage and deploy to VMConAWS.

“VMConAWS is truly re-defining the story of hybrid cloud”

One important point to note is that whilst VMConAWS runs the full VMware SDDC stack (including vSAN and NSX), a customer's on-premises VMware infrastructure does not need to. In fact, if you wish to just consume VMConAWS as a public cloud only service, you don't need to have any on-premises infrastructure at all. Obviously, if you do have on-premises infrastructure, there are additional benefits to being fully software defined but the only parity required between the two environments is the version of vSphere and VMware vCenter (currently v6.5d).

What is the minimum footprint in the VMConAWS?

The minimum footprint of an VMConAWS solution is a four node VMware Cluster which is configured with All-Flash vSAN. Within the initial availability phase, you can scale up to 16 nodes, but the minimum is four.

Each node is configured with the following hardware:

- Dual Socket E5-2686 v4 (Custom) – 18c per socket, 2.3 GHz
- 512GB RAM
- ~10TB RAW Flash capacity – using 8 encrypted NVMe devices

What is Elastic DRS?

Elastic Dynamic Resource Scheduling (DRS) is a new type of DRS specifically developed for VMConAWS. It enables a VMware Cluster to scale elastically by dynamically adding and removing hosts based on defined resource capacity demand thresholds. Most importantly the process is transparent to the end user.

What is Automated Cluster Remediation?

Automated Cluster Remediation is the process by which VMware will swap out a failed or problematic host within an instance of VMConAWS to ensure service levels are maintained and completely transparent to the end user.

At a high level, VMware will automatically detect a failed host and then introduce a new host into the effected VMware Cluster. The failed or problematic host is put into maintenance mode and evacuated from the cluster.

What is Hybrid Linked Mode?

Hybrid Linked Mode (HLM) is the mechanism by which you can link your VMConAWS vCenter Server instance with a single on-premises vCenter Server instance (with embedded Platform Services Controller). HLM enables you to:

- Log in to your VMConAWS vCenter Instance using the on-premises credentials.
- View / Manage inventories across VMConAWS and on-premises using a single console.
- Cold migrate workloads between VMConAWS and on-premises

Unlike **Enhanced Linked Mode (ELM)**, HLM does not create a single SSO domain across the multiple sites of the hybrid cloud, it links the two different SSO domains together in a 1:1 relationship.

Which results in a VMConAWS 4 node cluster providing:

- 144 cores
- 2TB RAM
- ~40TB RAW Flash capacity

That is a significant VMware Cluster in any on-premises data centre, let alone a cloud based service!

If we look at the maximum supported cluster, which has combined resources of:

- 576 cores
- 8TB RAM
- 160TB RAW Flash capacity

I definitely cannot see too many customers needing that size for their “public cloud” based workloads.

From a storage perspective, the VMware ESXi boot volumes are provided by the **Amazon Elastic Block Store (EBS)** and the encrypted NVMe storage devices for the VMware VSAN are provided by the **Amazon Instance Store**.

The VMConAWS deployment is based on the **VMware Cloud Foundation™** platform, therefore during initial deployment/configuration, a dedicated VMware Resource Pool is created for the **Management Plane** to ensure enough resources are available to run the platform.

When it comes to the VMware vCenter Server Instance within VMConAWS, the customer will have restricted access and will not have the ability to create any new Standard or Distributed vSwitches. All networking configurations must be completed at the **VMware NSX** layer using a combination of **Logical Switches, Edge Service Gateways (ESG)** and **Distributed Logical Routers (DLR)**.

VMConAWS “as a Service”

VMConAWS is a sold and managed “as a service” from VMware, therefore a lot of the Management and Operational responsibility resides with VMware.

Resource	Deployment	Lifecycle	Configuration
VMC Console	VMware	VMware	VMware
AWS Infrastructure (Compute, Network, Storage)	VMware	VMware	VMware
Provide VPC	VMware	VMware	VMware
Customer VPC	Customer	Customer	Customer
ESXi Hosts	VMware	VMware	VMware
vCenter	VMware	VMware	VMware
NSX	VMware	VMware	VMware
VSAN	VMware	VMware	VMware
Management Gateway	VMware	VMware	VMware
Compute Gateway	VMware	VMware	VMware
Virtual Machines	Customer	Customer	Customer
Guest OS and VM Tools	Customer	Customer	Customer

The responsibility still resides with VMware, in the event that there is an underlying hardware problem (which is ultimately an AWS responsibility) as they are providing the service. Therefore, it is VMware who will provide the single point of contact for the resolution of any issues.

What are the purchasing options?

Like with most Cloud offerings in the market place today VMConAWS comes with two pricing structures:

- **On-Demand** – This option provides the customer with the ultimate flexibility around the consumption of the service. You pay on a per host, per hour basis (billed monthly) with no upfront costs or commitment. This provides ultimate flexibility but with that flexibility comes cost as this is the most expensive way to consume this service for long periods.
- **Reservation** – This option allows customers to pay upfront for either 1 year or 3 years. This is charged on a per host basis and using this option, customers could save up to 50% on the cost of 24/7 hosting against the On-Demand pricing. Currently, Reservation pricing is paid for all up front.

Note: At initial availability, only On-Demand pricing is currently available.

In addition to the above pricing plans, VMware are also introducing a **Hybrid Loyalty Program**, where Reservation pricing customers will be able to save an additional 25% on the cost of VMConAWS, as long as a they maintain active Support and Subscription for their on-premises VMware licensing. Effectively, on a 3 years subscription, customers will save ~63% of the **On-Demand** price.

	Core Offering		
	On-Demand (hourly)	1 Year Reserved	3 Year Reserved
List Price (\$ per Host)	\$8,3681	\$51,987	\$109,366
Effective Monthly	\$6,109	\$4,332	\$3,038
Savings Over On-Demand	-	30%	50%
Effective Monthly (w. HLP of 25%)	-	\$3,249	\$2,278
Savings Over On-Demand	-	47.71%	63.34%

There are two important notes to make regarding pricing:

- The pricing mechanisms **only** include the AWS infrastructure and VMware software, however it does not include any additional AWS charges that maybe be incurred when running this service or any additional licensing that may be required. For example, you can expect to incur egress data charges in a hybrid cloud scenario where the monitoring and management toolsets are hosted on-premises and data is exchanged between the two systems.
- There is no charge for Data exchanges between a Customer VPC and the Provider VPC because it is using an internal **AWS Elastic Network Interface (ENI)**, so the data doesn't egress or ingress the Amazon ecosystem. For example, accessing data that is already in an Amazon S3 bucket from within VMConAWS will not cost anything extra.

What are some of the use cases for VMConAWS?

In my view, different customers will have differing use cases to adopt VMConAWS. These use cases are vast and varied and can include:

- Production** – the ability to move some/all of your workloads out of a hosted data centre and re-host (without re-architecting) into a highly available public cloud service.
- Development and Test** – the ability to run the development and test workloads for your existing on-premises services, on demand in the cloud without impacting production hardware and only paying for what you use.
- Seasonal/Special Workloads** – the ability to build new or migrate application/services to meet seasonal or special workload demand that leverage the elastic scalability of the VMConAWS platform and or Native AWS services.
- Disaster Recovery** – the ability to provide a highly available short-term disaster recovery environment containing critical applications and data whilst the on-premises environment is recovered.

Is VMConAWS re-defining the term “Public Cloud”?

Gartner defines Public Cloud computing as “a style of computing where scalable and elastic IT-enabled capabilities are provided as a service to external customers using Internet technologies.” Whilst one of the key features of VMConAWS is its ability to scale elastically on demand, for me, in isolation, it is not a traditional “Public Cloud” service in the same mould as AWS, Microsoft® Azure or the Google Cloud Platform (GCP).

“A key feature of VMConAWS is its ability to scale elastically on demand”

In my view, the purpose of VMConAWS is not to compete with or replace the native Public Cloud services. After all, to fully migrate an application/service to make it available in the Public Cloud (such as AWS or Azure) would normally require the application to be re-architected using the native cloud services of the target platform. So firstly, and foremost, the purpose of VMConAWS is to give VMware customers the ability to re-host some (or all) of their VMware “Private Cloud” workloads in a Public Cloud using an “as a service” model that provides the benefit of scalability and flexibility for underlying infrastructure. Secondly, and to the benefit of AWS, the purpose is to give VMware customers the ability to leverage the plethora of native AWS services (such as the **Amazon Relational Database Service (RDS)** and **Amazon Elastic Load Balancing (ELB)**) without the need to fully re-architect the application/solution. This in turn gives the customer a more phased and structured approach to Public Cloud adoption.

Is VMConAWS heralding a new era of Hybrid Cloud?

If I'm being honest, I cannot see too many existing VMware customers migrating completely out of their on-premises infrastructure and into VMConAWS. At least not right away. A VMware Hybrid Cloud is definitely the immediate destination for most people and a step in the right direction.

Given the size and price point of VMConAWS, it is clear that VMware are targeting enterprise level customers. After all, who else could afford ~\$24K a month minimum for 24/7 hosting? However, in my experience, the majority of enterprise customers will have a large amount of technical debt in terms of legacy applications and maybe even physically

connected hardware, and these are just not supportable in a Public Cloud without the need for re-architecting/re-development. If we acknowledge the fact there are many different cultural and organisational barriers to cloud adoption and focus on the technical ones, I don't recall ever working with an enterprise customer that has successfully transitioned to become fully virtualised. Those technical barriers that exist for becoming virtualised on-premises will continue to hold true for the adoption of VMConAWS.

“Why use Any Cloud when you can use the VMware Cloud?”

Introducing a Hybrid Loyalty Program is a shrewd move from VMware, because not only are they acknowledging the fact that not everything that is virtual can or should be migrated to VMConAWS, but they are also ensuring that VMware's existing software/support and subscription revenue stream is maintained by giving customers an incentive for keeping some VMware based on-premises infrastructure.

However, it is definitely the ability to deliver truly hybrid cloud services, by combining the VMware vSphere platform and the native AWS services, that makes VMConAWS such an intriguing and viable prospect.

Whilst VMware's Cloud strategy is still very much Any Application, Any Device, Any Cloud, with the introduction of VMConAWS, I can't help but ask myself, “why use any Cloud when you can use the VMware Cloud?”.

Is VMConAWS breaking down the barriers to “Public Cloud” adoption?

One of the consistent barriers to public cloud adoption in the past has been both the lack of relevant skills and the time it takes to develop those new skills, not to mention keeping hold of those skills once developed! Whilst the support culture is slowly shifting towards hybrid cloud teams, customers will often still have different teams managing their Public Cloud presence and on-premises infrastructures. If you are using VMConAWS as purely a Private Cloud hosting platform, the skill gap is effectively eliminated because the VMware technology is the same both on-premises and in the cloud. One thing to be mindful of is that all of the customer facing networking within VMConAWS is based within **VMware NSX**.

“The real benefit of VMConAWS is the ability to integrate with and make use of native AWS cloud services”

In my opinion, whilst moving to VMConAWS just to move some or all of the workloads out of a physical hosted data centre is a valid use case, I think customers doing it only for that reason are missing the bigger picture. The real benefit of VMConAWS is the ability to integrate with and make use of the native AWS cloud services, which means, unfortunately, the skills gap (specifically in AWS) still remains and continues to get larger as new AWS services are released.

In short, if you haven't already or are not already planning to, now is the time to invest in upskilling your support teams in both **VMware NSX** and **Amazon Web Services**.

Are there any major barriers to the adoption of VMConAWS?

VMware wants VMConAWS to be a part of a Hybrid Cloud solution by enabling their customers to either transition workloads as-is to the cloud, or by allowing customers to transform their applications to become cloud ready. As with any cloud service, there are going to be barriers to adoption, and the majority are not unique to VMConAWS, but as this article is about VMConAWS, here are my top 5:

1. Price

Always one of the biggest barriers to adoption for any cloud service is price and, unfortunately, VMConAWS is no exception. That is not me saying it isn't good value for money for the resources and services being provided, but with currently only On-Demand pricing available, I can't really see many people wanting to or being able to afford to run services 24/7 in VMConAWS for a monthly fee of ~\$24.4K for a minimum size deployment.

If someone want to run services 24/7, I can definitely see the benefit for buying into the Reservation pricing when it is released, but that is still ~\$204K for 1 year and ~\$440K for 3 years for a minimum deployment that (currently) ALL has to be paid upfront.

2. Node / Cluster Size

The size of each node (in terms of CPU, RAM and HDD resources) is ultimately driving the cost of the solution and I believe VMware are excluding the SMB and even some of their Enterprise customers with the decision to have such high specification nodes without a lower specification option. The question is can you ever have too much resource capacity? Yes you can, especially when you pay by the hour and I definitely don't see too many SMB customers with the sort of resource requirements that is being offered in VMConAWS.

Looking at it from a cluster perspective, is a single cluster of up to 16 nodes going to provide enough resources for those enterprise customers that do adopt? I would definitely hope so! Would I recommend anyone deploy a full 16 node cluster? Not unless absolutely necessary because that defeats the whole point of having an elastically scalable infrastructure.

3. Requirements for VMware Hybrid Cloud

To be able to run VMConAWS and an on-premises VMware infrastructure within a Hybrid Cloud, enabling you to take advantage of Hybrid Linked Mode (HLM), the on-premises infrastructure (well at least the vCenter Server) will need to be upgraded to VMware vCenter Server 6.5d. Whilst I'm sure it is on the roadmap for a lot of customers, it potentially adds complexity and cost to the adoption. This is on a 1:1 basis therefore only one on-premises vCenter instance can link with one VMConAWS vCenter Server instance and there is only one vCenter Instance per VPC and one VPC per VMConAWS account. Customers may also be required to keep the version of on-premises VMware vCenter Server in sync with VMConAWS to maintain support for HLM.

Currently the only supported topology for HLM is VMware vCenter Server with embedded Platform Services Controller (PSC) and, therefore, you can only have one Enhanced Linked Mode (ELM) domain. This currently contradicts VMware's own recommendations for ELM within an on-premises deployment, where ELM with embedded PSC is a deprecated topology. Whilst VMware is keen to point out these are two different technologies with similar names, it doesn't help customers who want to adopt VMConAWS who already have external PSCs deployed to their on-premises infrastructure. I would expect to see HLM support for VMware vCenter Servers external PSC as coming sooner rather than later.

4. Connectivity

At initial availability, the only supported connectivity to allow Hybrid Cloud will be via a Layer 3 SSL VPN. Whilst SSL VPN is a viable connectivity option for a lot of customers because it is a secure and tested mechanism for connectivity, it is not a dedicated link. I can see a lot of enterprise customers deferring adoption of VMConAWS until the (hopefully) inevitable support for the dedicated, low latency and high bandwidth Amazon Direct Connect solution.

5. Industry Compliance/Regulation

Currently there is no support for, or the ability to deliver solutions that are required to meet some of the industry compliance checks such as PCI-DSS and HIPAA. Using PCI-DSS as an example, AWS itself is PCI-DSS Level 1 Certified Service Provider, but VMware will need to go through the accreditation process as they are essentially the Service Provider. This means, that currently, VMConAWS is not a suitable platform for deploying or migrating applications that store, transmit or process card holder data. To be clear, just because AWS is a PCI-DSS "compliant" Service Provider, doesn't mean applications deployed to it are automatically compliant. Ensuring overall compliance is still down to the application owner or business to achieve. The feeling coming out of VMworld is that this is being worked on so won't be a barrier for long.

Key Points on VMware Cloud on AWS

- You do not need an existing AWS account to use this service
- VMware will create and manage a VMConAWS account on behalf of the customer
- VMConAWS is not nested virtualisation, it is VMware vSphere on AWS bare-metal hardware
- Management of On-Premises infrastructure via Hybrid Linked Mode requires vCenter Server 6.5d
- You do not actually need an on-premises infrastructure to use VMConAWS – you can just use it as a standalone Public Cloud Service
- If you do have VMware infrastructure on-premises, it does not need to use VMware NSX or VSAN to use VMConAWS
- The minimum number of nodes is 4 (maximum is 16)
- Only available in AWS US West Region Availability Zone (currently) – rollout to all AWS AZ
- Customer networking with VMConAWS is via VMware NSX only
- Currently available with On-Demand Pricing only
- Data egress charges will apply
- Traffic between the Customer VPC and Provider VPC is free!
- VMware have a rapid development cycle for new features to be released

Final thoughts

There is no doubt we are in the midst of a cloud revolution that spans all industry sectors. Now more than ever before, businesses are moving workloads and applications into Public Cloud services as a way to meet the business requirements to reduce cost, add flexibility, increase productivity and expand the global footprint. However, I still feel wholesale adoption of "traditional" Public Cloud services (such those provided by AWS, Azure and GCP) are being stifled by the fact that legacy applications often need to be re-architected or re-developed to fit within the Public Cloud model. This is where VMConAWS is an interesting proposition, as it provides the ability for businesses to either start from scratch in VMConAWS, or migrate their existing workloads "to the cloud" and instantly reap the benefits from the scalability and flexibility of the AWS infrastructure, without the immediate need to re-architect from the ground up.

“We are in the midst of a cloud revolution”

Once “in the cloud”, by providing application owners the ability to access and consume native AWS Cloud Services (as an example Amazon RDS), businesses are empowered to make a choice about the future of their applications. Do I keep my legacy application running in a cloud hosted workload or do I choose to transform (and enhance) the way my application works using native cloud services.

As with any new endeavour, there is always going to be a number of technical, cultural and organisational barriers to adoption and I have highlighted a few. However, we need to remember that VMConAWS is in initial release and although the features/capabilities that are available today will be suitable for a lot of customers, we can expect the joint partnership of VMware and AWS to continue to deliver not only feature enhancements, but new functionality and availability of services across all AWS availability zones.

In summary, I believe that the partnership of VMware and AWS is a game changing move that may just pave the way for IT to start delivering real benefits back to the business. The first step is always about being able to articulate the business benefits and, with VMware Cloud on AWS, the journey “to the cloud” has just become so much simpler. It certainly is going to be interesting to watch how cloud adoption changes in the next 12 months.

If you are interested in VMware cloud products and services and want to know how they can benefit your organisation, [contact us](#) and we'll be happy to use our wealth of knowledge and experience to assist you.

Xtravirt is a leading, independent provider of enterprise virtualisation solutions. We deliver data centre, workspace and cloud transformational solutions to clients across public and private sectors, both in the UK and internationally.

xtravirt
Dedicated to smarter business

© Copyright 2017 Xtravirt Ltd. All right reserved. Information contained herein is subject to change without notice. Xtravirt Ltd shall not be liable for technical or editorial errors or conclusions contained herein. Xtravirt and the Xtravirt logo are registered trademarks of Xtravirt Ltd. The names of actual companies and products mentioned herein may be the trademarks or registered trademarks of their respective owners.

