ESG Lab First Look

Velostrata: Decoupling Compute from Storage in Real-time Hybrid Cloud Environments

Date: **March 2016** Author: **Kerry Dolan**, **Senior Lab Analyst**

Hybrid Cloud Challenges: [1]

#1 Where **Performance** ranked among top 10 concerns about moving applications & data from on-premises to the cloud.

#3 Where **inefficiency of workload portability** ranked among top 10 concerns about moving applications & data from on-prem to cloud.

While hybrid cloud computing sounds promising, it's not so easy to accomplish. Many organizations want the option to run workloads in the cloud to handle spikes in demand, but to do that means renting compute and storage resources—where you pay for the storage even when it's not in use—and it may take weeks or months to move applications and data. It means extra cost, security concerns, and application changes, plus it's difficult to bring those workloads back on-premises.

Velostrata Architecture Overview

Velostrata

Velostrata software streams production workloads to the public cloud in minutes. It decouples compute from storage so adminstrators can move VMs to the cloud, leave the storage on-premises, and still maintain performance. Administrators make no changes to applications, images, networks, or storage. With the vSphere plug-in, they simply right-click on a VM, select "run in the cloud," and Velostrata does the rest. Velostrata snapshots the VM, shuts it down, makes required operating system changes, handles network configuration, and streams the VM to the cloud in minutes. Using Velostrata caching, dedupe, and prefetch algorithms, the VM boots natively in the public cloud, and pre-fetches necessary data to run the workload—similar to a streaming video that starts playing before the entire file has been downloaded. Authoritative data stays on-premises, where it is protected. This unique approach enables support for data center extension and cloud migration use cases. After installation and a one-time network set up, administrators manage and monitor VMs from vCenter. Velostrata currently works with VMware vSphere in the data center and Amazon Web Services (AWS) for the public cloud, with additional hypervisor and cloud support on the horizon. A RESTful API enables integration with third-party cloud orchestration products.

The Velostrata architecture includes a Data Center Edge virtual appliance on-premises, and a dual-node Cloud Edge virtual appliance in a customer-owned AWS virtual private cloud (VPC), spanning AWS Availablity Zones. The data center and VPC are connected via VPN or DirectConnect, with encrypted traffic between them. Each Cloud Edge supports up to 50 concurrent VMs, and can scale out by adding Cloud Edge virtual appliances. Velostrata technology creates a multi-tiered, read-write cache; for reads, the cache includes data that workloads frequently access or are likely to access in the near future, enabling SAN-grade performance. Writes are acknowledged locally for low latency, and updates may be sent back on-premises asynchronously or configured to persist in the public cloud. Writes are mirrored across Availablity Zones, and a write journal is saved in an object store to ensure data resiliency. Bi-directional deduplication and compression reduce the amount of data moving across the WAN.

---
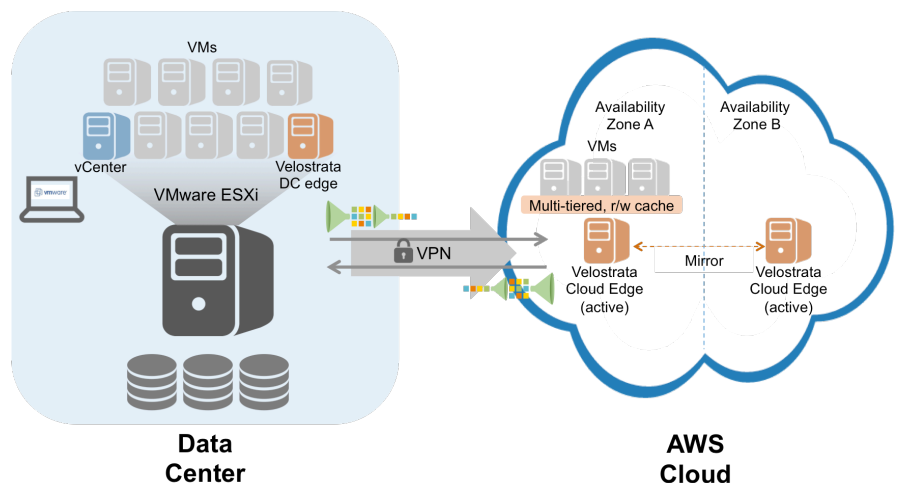
[1] Source: ESG Research Report, *The Cloud Computing Spectrum, from Private to Hybrid*, to be published.

ESG Lab Demo Highlights

ESG Lab performed hands-on testing of Velostrata, moving a VM to the cloud while monitoring performance.

Move VM to Cloud in Minutes

Running On-prem:
- 2 vCPU (2.4GHz), 8 GB RAM
- ~11K pages
- 2.03 seconds average response time

Running in cloud with Velostrata:
- 16 vCPU (2.8GHz), 30 GB RAM
- ~57K pages
- 0.38 seconds average response time

- ESG Lab started with an on-premises VM with 2 vCPUs, 8 GB of RAM, a 40GB boot drive, and a pre-defined data store. We ran a simulated SugarCRM application and used a load generation tool (WAPT) to simulate concurrent users making requests to the application that was CPU-bottlenecked. Performance monitoring demonstrated about 11K pages/sec, plus a high (and variable) response time, averaging more than 2 seconds. We also created a file (*ESG.txt*) on the desktop.
- Next, in vSphere we right-clicked on the VM and selected *Velostrata Operations/Run in the cloud*. After selecting the Velostrata Cloud Extension we chose an AWS Instance Type with more resources (16 vCPUs, 30 GB of RAM), selected to write back to storage, confirmed security and network configs, and clicked *Finish*.
- A vCenter screen refresh displayed a cloud instance window showing the VM's state and status as it was shapshotted on-prem; shut down; streamed to the cloud; and rebooted. Velostrata handled changing the VM configuration from working with VMware vSphere on-premises to AWS Xen in the cloud, including driver, registry, and licensing changes.
- In about five minutes, ESG Lab logged into the VM running in the cloud. Velostrata is integrated with Active Directory and DNS, so there is no change in how you access the VM. The vSphere portlet showed the VM running in the cloud, along with the AWS instance details. We validated that it was the same VM by the presence of the *ESG.txt* file on the desktop.
- Next, we re-started the WAPT workload generator. With the CPU bottleneck alleviated, performance monitoring demonstrated higher pages/sec (57K) and a low and consistent response time, averaging 0.38 seconds.
- Several notable features: a) As workloads run in the cloud, the cache warms, improving performance over time. b) Once the VM boots (natively, not nested) in the AWS cloud, all AWS services are available to it. c) With the VM running in the cloud, the customer pays just for cache for the active data set, not storage. d) All cache data at rest in the cloud is encrypted automatically with AES-256 and all data is encrypted in transit using SSL, on top of a VPN. e) Administrators can reconfigure the VM in the cloud. f) Returning to on-premises resources is equally simple and fast.
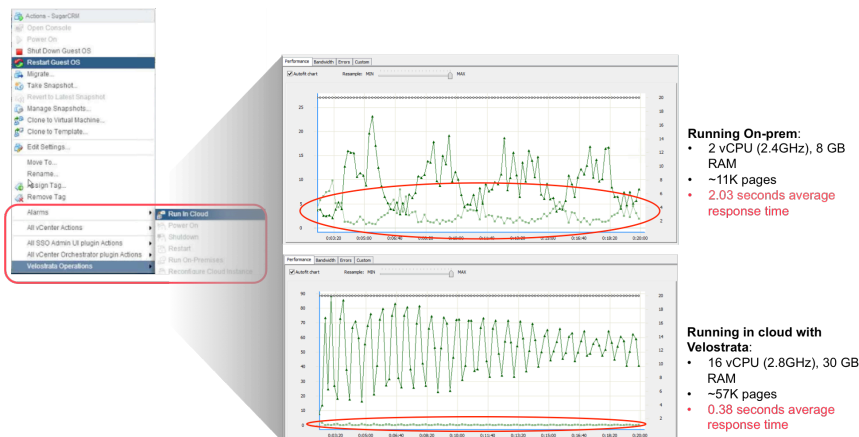
First Impressions

Many data centers are overprovisioned to accommodate peak loads for existing applications and to provide capacity for new projects with variable loads. However, due to the inherent variability of business demands, those same data centers may also experience times when they are underprovisioned, negatively impacting application response time and availability. The promise of hybrid cloud computing has been the option to offload jobs during high utilization times, but that has not been the reality due to the cost, complexity, and migration time of moving compute and storage resources.

Velostrata enables that reality by decoupling compute from storage and leaving storage on-premises. This enables real-time, simple, fast "cloud-bursting" for CPU-limited workloads, with performance equal to or better than on-premises—and without costly data center overprovisioning or expensive, high performance cloud storage. Only recently out of stealth, Velostrata's solution is just now being proven in the field, but ESG Lab's first impression is that the performance claims for production workloads are real. We moved a VM to the cloud in about five minutes, requiring very little management, and demonstrated the ability to eliminate CPU performance bottlenecks by choosing a cloud instance with additional resources to boost performance.