SOLUTION BRIEF

ENABLING A HYBRID CLOUD ENVIRONMENT WITH DATA INTEGRATION

HVR SOFTWARE
REAL-TIME DATA REPLICATION

THE MOVE TO THE CLOUD

Cloud computing is fundamentally transforming the way enterprises think about IT. Information, services, and applications can now be accessed on demand, anywhere, from any device, by any user, at little or no cost. Universal access to high-speed network connectivity effectively frees users from reliance on the enterprise infrastructure. This new found freedom means that businesses can streamline the traditional obstacles and challenges faced by the CFO, CIO, and IT department. It is therefore not surprising that according to the Computerworld 2015 Forecast Study, cloud computing is projected to increase by 42%.

The hybrid cloud, according to IDC research, is set to explode in 2015. IDC reports that the combination of sales of public, private, and hybrid cloud services will top $118 billion in 2015 – and Technology Business Research predicts that the hybrid cloud alone will experience 50% higher growth in 2015 than 2014.

Gene De Libero,
Hybrid Cloud Forum

Cloud services are generally classified into three categories.

- Software as a Service (SaaS) includes applications such as Salesforce and NetSuite, email and office productivity suites such as Gmail and Office365, collaboration suites such as Basecamp and GoToMeeting, and many other related applications.
- Platform as a Service (PaaS) is the cloud rendition of application infrastructure, represented by products such as Force.com, Windows Azure, and Pivotal Cloud Foundry.
- Infrastructure as a Service (IaaS) offerings encompass hardware computing resources as well as storage and networking capabilities. Examples of common IaaS offerings include Amazon EC2, Rackspace, and Google Cloud Storage.

What defines a cloud application? The industry recognizes five essential characteristics of cloud computing. These include on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service. Of these, the notion of elasticity is of particular relevance with regard to integration. Data integration capacity and performance must be able to scale as demand for cloud services expands.

HYBRID CLOUD COMPUTING

Traditional on-premises systems are rapidly being migrated to the cloud or replaced with modernized cloud services. At the same time, entirely new applications that leverage big data and mobile technologies are being deployed exclusively as cloud services. For most organizations, this entails maintaining a hybrid cloud computing environment where on-premises and cloud-based systems co-exist and interoperate, at least temporarily. Hybrid cloud environments can be manifested in many ways, as illustrated by these two examples:

The Dutch post office, PostNL, provides the largest network for letters, parcels and e-commerce in the Benelux countries. The company is moving all its applications to the cloud. Though SaaS applications are preferred, PostNL also has numerous existing on-premises applications and the migration to the cloud will take years. During that period extensive integration between existing on-premises services to the new cloud services is necessary.

Global IT services and solution provider, CACI, delivers their OSIRIS higher education package as a cloud service. As well as wanting the economic benefits of cloud storage and processing, customers also want their data available locally for reporting and use in internal systems. This requires CACI to provide a means to deliver this data to their customers, from the cloud to their on-premises environments, in real-time.

HVR SOFTWARE
REAL-TIME DATA REPLICATION

"Hybrid cloud, and by extension hybrid IT, is here to stay. Few companies will only do public or only do private cloud computing, and no company should miss the opportunity to leverage both."

Daryl C. Plummer and David Mitchell Smith, Gartner

Hybrid cloud computing supplements the essential characteristic of elasticity with the need for interoperability between on-premises systems and cloud services, as well as between cloud services.

It therefore implies significant integration or coordination between the internal and external environments at the data, process, management and security layers.

A hybrid solution is intended to combine the benefits of both on-premises and cloud-based platforms. This approach provides enterprises with the flexibility to store and process targeted data assets in the cloud, and integrates these with on-premises data assets as required to provide an end-to-end solution. This architecture demands orchestration of services across both on-premises and cloud environments.

The benefit to the hybrid approach is that enterprises are not forced to choose between cloud or on-premises platforms, and have the flexibility to implement best-of-breed solutions that run in either environment.

CHALLENGES

Although cloud based services can deliver substantial cost savings, flexibility and accessibility, hybrid cloud computing introduces some unique challenges. The most significant of these is the added data movement operations between various cloud providers and on-premises data centers. This can be even more problematic when uncompressed data volumes are large or arrive too fast and need to be integrated quickly.

The cloud environment is also sensitive to communications issues which can cause performance and latency problems. The multi-tenant nature of a public cloud infrastructure can also cause issues — "noisy neighbors" are a potential problem. The scalability of cloud environments should address platform-related issues. However, the network could remain a problem, especially in areas with poor network reliability.

The lack of data encryption is another common challenge associated with moving data between and within cloud environments. Therefore, the need for secure, reliable and efficient integration is critical when using multiple cloud services interchangeably.

"Hybrid cloud shouldn't add more management complexity, it should simplify IT operations."

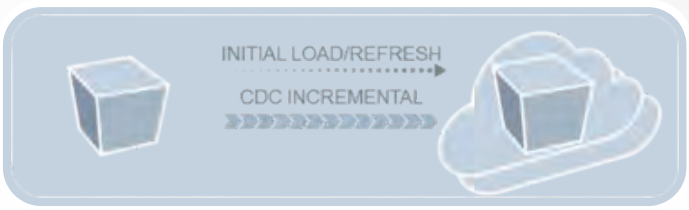
Dave Bartoletti, Forrester Research

In addition to moving data between various cloud environments, each cloud environment operates internally like a data center running multiple applications, moving data within each environment. This is especially true when on-premises applications are moved to the cloud "as is." If not re-designed for cloud deployment, they need additional integration middleware. SaaS offerings typically only provide application-level integration and not data-level integration. By adding data integration capabilities within a cloud platform, existing applications can be reused in the cloud as-is, with integration as an external service.

HVR SOFTWARE
REAL-TIME DATA REPLICATION

CLOUD-BASED DATA INTEGRATION USE CASES

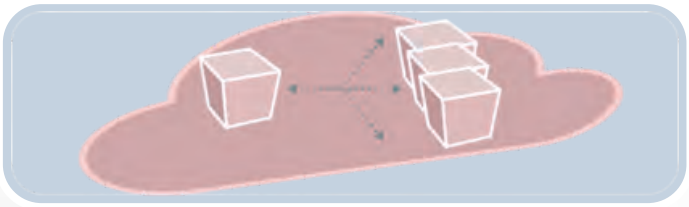
1. Adding a cloud platform to an existing IT landscape



New cloud services that are added to an existing IT landscape usually require access to corporate data. This means moving on-premises data to the cloud to perform the initial load and on an ongoing basis as changes occur.

In a later stage, feedback of data to the internal systems (for further processing, reporting or consolidation) may become a necessity as well.

2. "Lift and Shift" existing applications to the cloud



When multiple applications are moved to the cloud, integration between them becomes necessary. If the applications are not redesigned, and are merely shifted to the cloud, exchanging data between their databases is often the only way of tying them together.

2. Securing cloud data to on-premises systems

When commissioning a cloud service independent of on-premises systems, there is sometimes a need to access that data independent of the cloud service. Several use cases can coexist, such as securing data from cloud service disruptions and further use of the created data in local systems.

4. Inter-cloud data integration

Many organizations adhere to a “best-of-breed” cloud services approach. This means that different services will be sourced from different vendors. When those services operate on the same or related data, the integration challenges are significantly higher than those associated with traditional on-premises applications. Source and target interfaces are often pre-built incompatible with each other. With the addition of each new service, the likelihood of vendor-supported inter-cloud integration and interoperability is greatly diminished. Moreover, the path between the services is by definition public with all its limitations and pitfalls.

As there is no obvious role for an on-premises environment as central integrator, integration has to come from the cloud services themselves or at least be able to run from there. Hybrid cloud computing requires data integration capabilities that support any combination of the above scenarios.

ALTERNATIVE APPROACHES TO CLOUD-BASED DATA INTEGRATION

Front end integration / virtualization

This integration approach does not actually copy any data but provides data access when needed in the cloud itself. It therefore does not answer the challenges of independence of the cloud service and network.

Application / service integration

Most cloud services provide application-level interfaces based on web services. Though very flexible, this integration style is not very scalable as it is targeted to individual transactions only. Lookups are easily accomplished, but accessing large volumes of data requires some form of data integration and application awareness. When merely shifting existing applications to the cloud, adding a service layer is not desirable.

Batch data integration

This level of integration entails exporting data on request and transferring it to the desired location (often using some form of file transfer). When an update is required, all data is transferred again. Although access to all data is provided independently of the cloud service, it is best used once only. Repeating this process leads to redundant data movement, requires data to be merged, and is nowhere near real-time. Higher update frequencies lead to exploding costs.

Data replication with Change Data Capture

Change Data Capture (CDC) technology delivers a continuous flow of data to or from the cloud

service in near real-time. It provides the required independence as well as the most efficient way to move data (as it is moved only once). CDC can come with additional technologies like log-based capture and parallel integrating, minimizing the load on sources and targets.

Data integration solutions offering both batch modes and replication in the same product is clearly the preferred way to address most of these challenges.

CLOUD-SPECIFIC REQUIREMENTS

The above described hybrid cloud scenarios introduce additional requirements to a data integration environment.

In addition to supporting the full data integration lifecycle (from initial / batch loading to change data capture), it must also be possible to change the direction of data interchange when focus shifts from on-premises to the cloud. In a true hybrid situation, it must be possible to support bidirectional interchange of data to and from the cloud, and between various clouds.

The dislocation aspect of cloud computing adds demand for high efficiency and speed. Robustness is also an important factor. Data integration services need to be able to handle cloud outages, maintain integrity and be able to resume with no manual interaction.

The multi-tenancy aspect of cloud computing brings added security requirements. Protection of data needs to be an integral part of any data integration solution. Both encryption to protect eavesdropping and authentication to protect misuse of services needs to be handled.

FAST, FLEXIBLE CLOUD COMPUTING WITH HVR

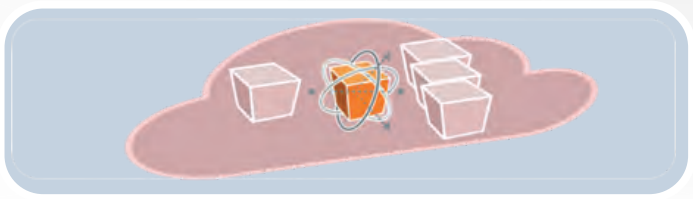
HVR's real-time data replication software mitigates the common challenges and risks associated with cloud migrations. Organizations that are migrating existing applications to the cloud or are implementing new cloud applications rely on HVR to move data quickly, efficiently and securely.

Flexible deployment options - HVR supports all prevailing cloud data integration scenarios within a single setup from a centrally managed hub and easy to use graphical user interface. HVR services are multidirectional and can operate in a mixed environment where updates happen on both sides with collision resolution handled via built-in collision detection. Common deployment scenarios include:

1. To the cloud

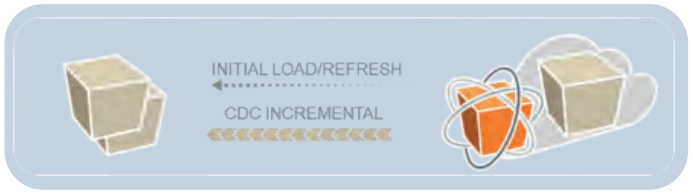
In this use case, the HVR hub installation typically is with the on-premises location as the master. HVR can be setup easily by adding a remote connection to a cloud location.

2. In the cloud



For adding data integration capabilities to a cloud environment (for example between existing applications that have just been moved there), the HVR hub can be installed in the cloud.

3. From the cloud



With the HVR hub installed in the cloud, cloud-based services can also communicate with on-premises installations.

4. Between clouds

In this use case, the HVR hub is installed in one of the clouds to facilitate data integration between cloud environments operated by multiple vendors.

Efficient and secure – HVR's log-based CDC technology only replicates incremental changes, minimizing the load on sources and targets in a non-invasive manner. Network bandwidth is further minimized through advanced data compression based on proprietary algorithms and specific knowledge of data structures. Compression ratios typically exceed 90%. Data is also encrypted on the wire to protect it from unauthorized access, even on public infrastructure. Support for 2-way SSL compression ensures uncompromised delivery of data and secures the authenticity of source and target data.

Scalable and robust – HVR's unique hub and spoke architecture delivers real-time replication with no limits on the number of sources and targets or on data volumes. It maintains the state of every replication channel, and routes data from one or more sources to one or more targets in unidirectional, bidirectional or multidirectional setups. DDL is automatically generated to set up a

In January 2015, RightScale conducted its fourth annual State of the Cloud Survey of the latest cloud computing trends, with a focus on infrastructure-as-a-service. One of the key findings of the survey was that, in 2015, 82 percent of enterprises have a hybrid cloud strategy, up from 74 percent in 2014.

Kim Weins, RightScale

target database schema based on the data source taking into consideration any required transformations. Online data load and compare functionality ensures that heterogeneous source and target systems continuously stay in sync.

Installation, configuration, diagnostics, and change management are all performed from a centralized, GUI-driven management console with context sensitive online help that minimizes ownership costs associated with system administration. Administrators can literally configure maintenance tasks for lights-out management.

Comprehensive cloud and on-premises environment support – Most hybrid cloud deployments are characterized by heterogeneous on-premises systems and diverse cloud environments. HVR supports the industry's broadest array of databases, file systems, applications, platforms, and cloud services (IaaS, PaaS, SaaS) from leading vendors such as Actian, Amazon, Cloudera, Hortonworks, IBM, MapR, Microsoft, Oracle, Pivotal, Salesforce, Teradata, and many more.

GETTING STARTED

If you're an IT professional interested in learning more about how HVR can help you address the challenges of data integration in a hybrid cloud situation, you can visit our website at http://www.hvr-software.com and watch videos to learn how the software works, how it's used in organizations like yours, and how it can potentially be used to address your issues. Through the website you can also request a free trial to test the software in your environment. Maximize your leverage of the cloud by having its pitfalls eliminated by HVR.

ABOUT HVR

At HVR, we believe it should be easy to deliver large volumes of data efficiently, reliably and at the right time into your data store of choice. Our software, the HVR High Volume Replicator, does exactly this using real-time data capture between data sources including SQL databases, Hadoop, data warehousing and business intelligence data stores as well as the most commonly used file systems.

For those organizations where real-time data replication is a mission critical process, HVR has been proven to be a reliable, secure and scalable solution by some of the largest global companies and leading government and defense organizations. HVR Software is a privately held company with offices in North America, Europe and Asia Pacific.

HVR Software US
135 Main Street, Suite 850
San Francisco CA 94105
USA

+1 (415) 655-6361

HVR Software EMEA
Daalwijkdreef 47
1103 AD Amsterdam
The Netherlands

+31 20 820 1570

HVR Software APAC
23 Adams Street
French Forest NSW 2086
Australia

+61 4040 81962

HVR Software China
2/F ShanghaiMart
2299 Yan An Road West,
Changning District Shanghai
Shanghai, 20036
People's Republic of China

+ 86 150 0066 1981

For more information, please contact us at info@hvr-software.com