



Hybrid clouds: the best of both worlds

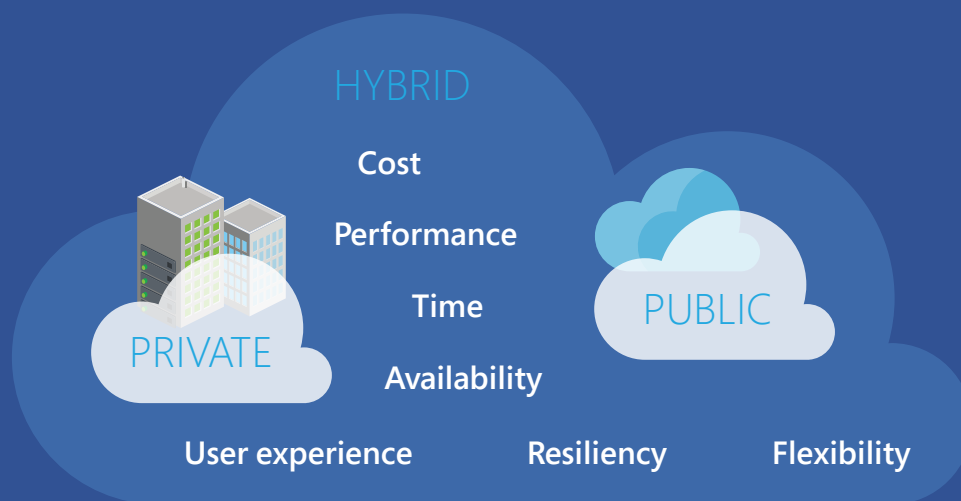
by Joe Weinman

Contents

Hybrid clouds: the best of both worlds	1
Private, public, or hybrid?	2
Transition or endstate?	3
Rationale for hybrids	4
Hybrids in space	5
Hybrids in time	7
Complex hybrids	9
Hybrid essentials	11
Summary	12
Author bio	13

Hybrid clouds: the best of both worlds

There are several compelling reasons that companies are increasingly viewing hybrid cloud solutions—constructs that couple multiple components such as private resources and public cloud services—as their target end state, rather than merely an interim step towards running all applications in the public cloud. Although there are a number of variations of hybrid architectures, they offer numerous advantages over a forced choice of one or the other. These advantages can include lower total cost, enhanced performance, greater availability and resiliency, reduced time to market and time to volume, enhanced user experience, and increased flexibility.



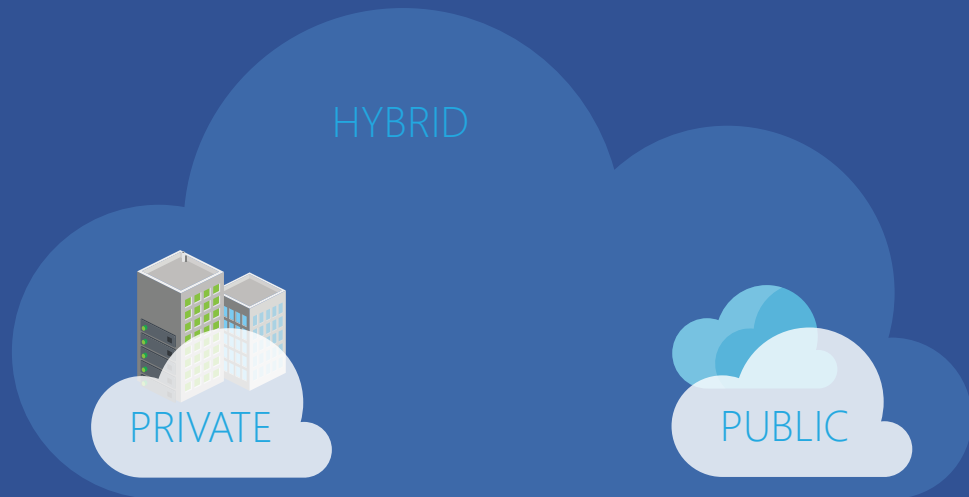
Private, public, or hybrid?

There are several main options available for running applications: legacy architectures such as mainframe environments; special-purpose technologies such as high-performance computing and quantum computers; modern architectures running on dedicated, siloed resources in an enterprise data center or colocation facility, on private clouds in an enterprise data center or in a colocation facility; managed services; bare metal; virtual private clouds, and public clouds.

Although some take issue with the term “private cloud,” it is not the term that matters as much as the notion of a computing environment that can accelerate provisioning and dynamically allocate resources among workloads that vary in size, thereby achieving higher utilization and reduced costs.

From an economic perspective, the salient differences between computing options are illustrated by one extreme of owned, dedicated resources that incur a fixed cost regardless of whether they are used, and shared pay-per-use resources that incur a variable cost as they are used. Capital expenditures vs. operating expenses are often mentioned, but fixed resources can be leased and incur operating expenses, and some reserved instances have reportedly been capitalized, so that distinction is imprecise.

Any mix of these two pure approaches to meet the needs of an application or suite of applications can be viewed as a hybrid architecture.



Owned or leased	↔	Pay per use
Dedicated	↔	Shared
Fixed capacity	↔	Variable capacity
Fixed cost	↔	Variable cost

Transition or endstate?

Some view the public cloud as the end state for all computing, using the analogy of the electric utility. However, virtually every need in a variety of contexts is met by a hybrid mix of both dedicated and on-demand, pay-per-use resources. For example, transportation needs are met both by private enterprise fleets and by rental car services. Floorspace requirements are met both by owned space, such as a home, and by pay-per-use space, such as an apartment or hotel. Capital requirements are met both by equity (owned capital) and by debt (rented capital). The increasing prevalence of solar cells on buildings can be viewed as a private electric cloud servicing multiple “applications”—for example, lighting, air conditioning, cooking—that complements public electric utilities.

From looking at other industries as well as the underlying economics, one can come to several conclusions. First, although the degree of penetration varies by industry, the public cloud model—short-term, on-demand, pay-per-use “rental” of resources—almost always exists and is not only viable, but a highly successful business model offering compelling customer value. Second, the ownership model also almost always exists. Third, seamless interoperability between the two is critical. Imagine if you couldn’t pay off your mortgage, or some light bulbs in your house only ran on solar, or you couldn’t check in to or check out of a hotel.

TRANSITION OR ENDSTATE?



PRIVATE



PUBLIC



Personal car



Taxi



Equity



Debt



Home



Hotel



Generator



Electric utility

Rationale for hybrids

The underlying reasons that hybrids are prevalent across domains include economics, option value, migration and switching costs, performance tuning, and human behavior.

Generally, in the presence of demand varying off of a baseline, hybrid resources represent an optimal mix of a lower cost for dedicated resources that run the portion of the demand that is relatively consistent as well as a lower cost for on-demand, pay-per-use resources for the portion of the demand that is relatively variable. Having a car at home for the relatively consistent demand of driving to work every weekday makes sense, as does renting a car on a family vacation or taking a taxi on a business trip. Should the relative costs shift, a hybrid provides the opportunity to shift an increasing portion of demand accordingly.

There is value (“option value”) in flexibility; being locked in can destroy value. For example, being locked in to long-term leases and owned equipment can be economically unwise if public cloud providers drop prices, as has been the case for the last few years. On the other hand, being locked in to a public cloud provider could be an issue if that provider were cease operations or have degraded performance.

Moving applications from a private environment to the public cloud can incur “migration costs,” and even if the public cloud offers costs savings, the return on investment—i.e., the return in terms of cost savings or performance improvements based on the investment in application migration—will therefore vary. Ideally, applications should be developed for an environment which is fully consistent across private and public clouds.

A subtle difference between private and public clouds that is often overlooked is the performance difference between the two. While public clouds offer substantial options and numerous standard configurations, the ultimate control over physical architecture lies with owned resources that can be configured with infrastructure elements such as memory, cores, processor interconnect bandwidth and topology choices to exactly match the needs of the application. An architecture with more expensive elements that are better tuned to application requirements can offer better overall price-performance, but with the potential loss of flexibility. Such differences, if any, depend on the application. Because application performance often impacts labor productivity and revenue, the complete benefits associated with performance must be considered.

In any event, the ability to monitor, manage and control both types of environments together is good IT practice and a basic need.

Hybrids in space

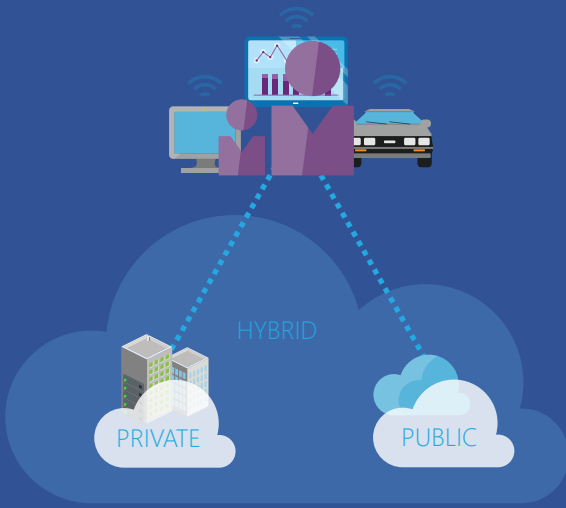
A hybrid cloud computing environment typically includes users and/or connected things, private cloud resources, and one or more public cloud services. There are multiple possible such configurations.

In the case of stateless elasticity, such as serving static web pages, load balancers can route users to private or public resources.

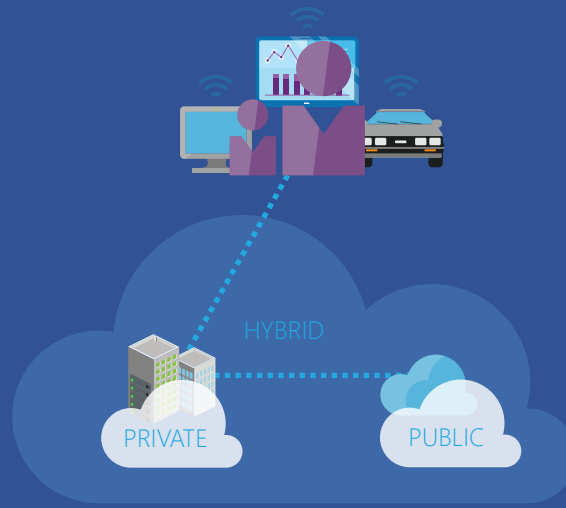
In another case, the cloud acts as the front end, and the enterprise data center is the back end. For example, the cloud may offer content delivery, say, for a web-based catalog, while financial transactions are executed at the enterprise data center when a customer orders.

In the reverse, the enterprise is the front-end while the public cloud is the back end. For example, commercial orders and transactions of record may both occur within the private cloud, and snapshots of the general ledger database, say, may be kept in the public cloud.

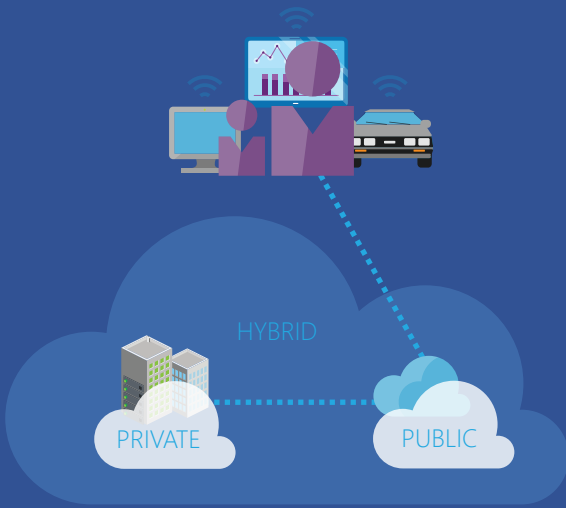
In an increasingly viable scenario—often referred to as “cloudbursting”—the same application architecture is run in both private and public clouds, and public cloud resources are increasingly brought to bear as demand exceeds the capacity of the dedicated enterprise resources.



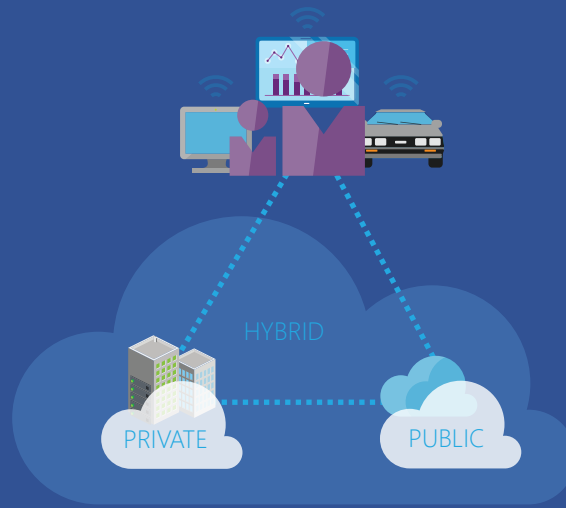
Stateless elasticity



BC/DR



Content/service delivery



Cloudbursting

Hybrids in time

Although less traditionally viewed as a hybrid, we can also think of four generic hybrids in time.

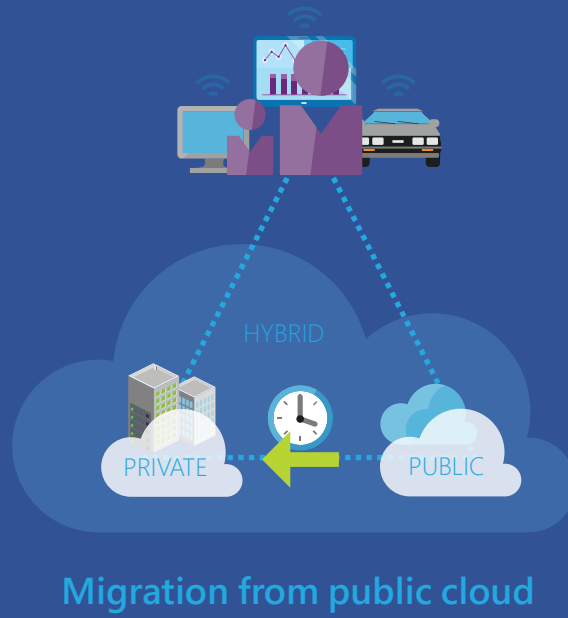
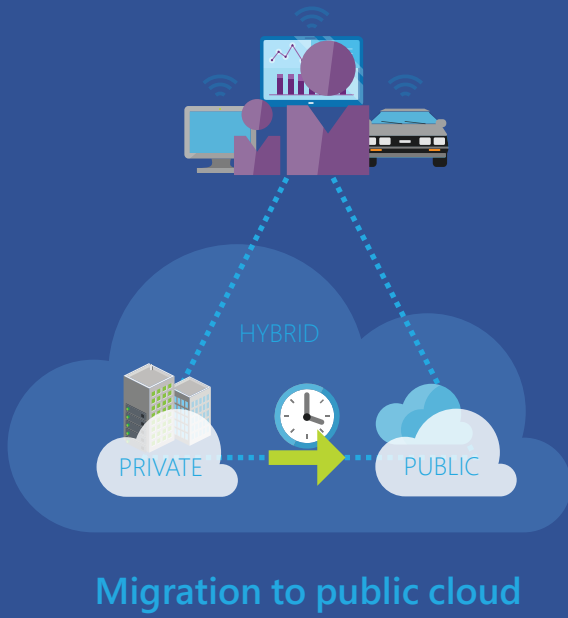
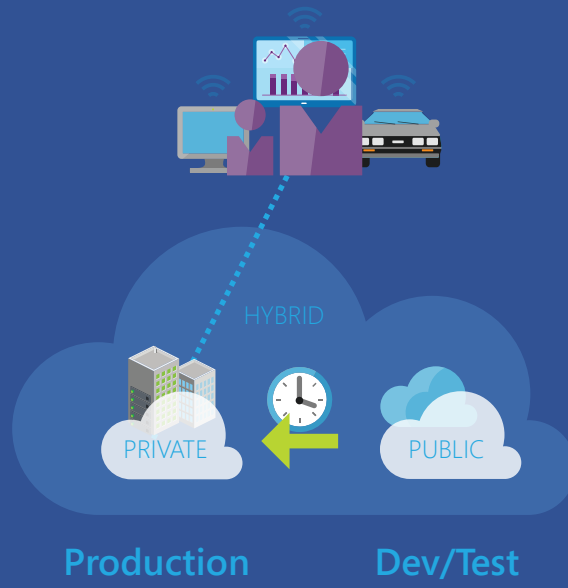
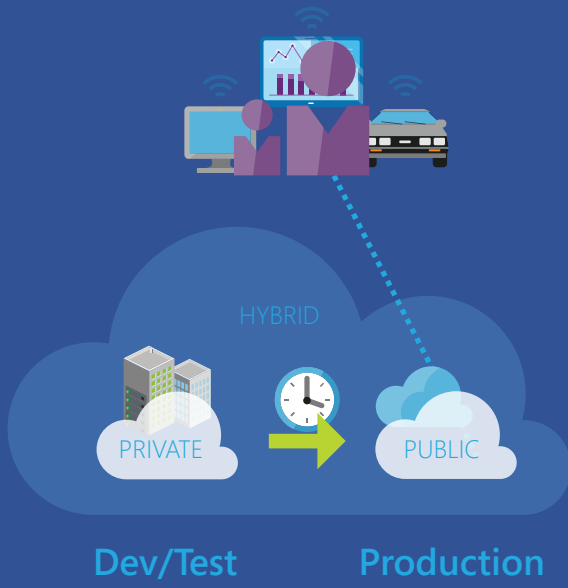
In one, development and/or testing is run in the public cloud and when it is time for production, the application is run in the enterprise.

In the reverse scenario, dev/test is run in the private cloud, or even a desktop or workstation, and then the application is moved to the public cloud for production.

In another scenario, everything—development, test, and production—is run in the private cloud, but at some point, due to strategy or economics, it is migrated to the public cloud.

In the reverse scenario, a company, perhaps a startup, bases its IT initially on a public cloud strategy, and once it reaches a level of stability in both architecture and demand, migrates its applications to its own data center.

Such choices are sometimes based on economics: a social gaming company found a three-fold improvement by migrating from public to private, and a cross-platform application found a four-fold gain. Some applications are less expensive to run in the public cloud, or reduce risk through elasticity. It can also be a strategic choice: a large global streaming media company wants to close all of its own data centers to focus on its core business of creating, transcoding, delivering, and recommending content.

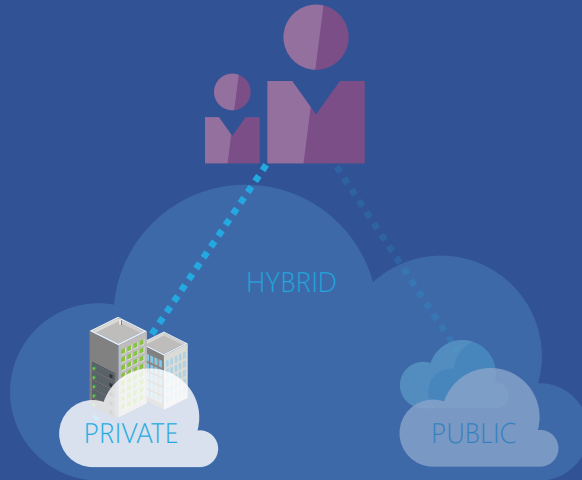


Complex hybrids

Mixes of all of these scenarios and economic and behavioral drivers can and do occur based on the requirements of the business. For example, a global entertainment company found that its web site, hosted entirely on internal resources, went down during the death of one of its most popular stars, as a global public eagerly sought information about the artist. It re-architected its application and deployment strategy so that during periods of normal load, all traffic is sent to the corporate data center. However, if demand for artist info and memorabilia spikes beyond the data center's capacity, it would be increasingly handled by a public cloud while the purchase transactions would still be handled by the company directly.

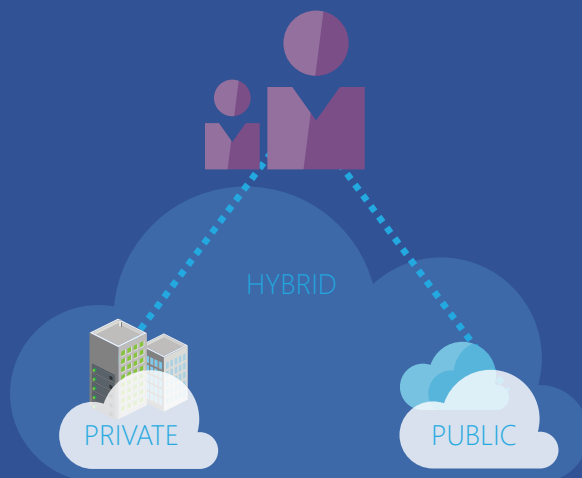
In yet another approach, which today is extremely prevalent, multiple public cloud providers complement applications or components in the enterprise data center, e.g., one SaaS provider for CRM, another for document sharing and yet another for object storage, say, backing up data from the first two providers. Over time, greater levels of application and data integration between these various components or microservices such as credit card validation will occur.

LOW DEMAND



- Purchase transactions
- Catalog
- Trailers

HIGH DEMAND

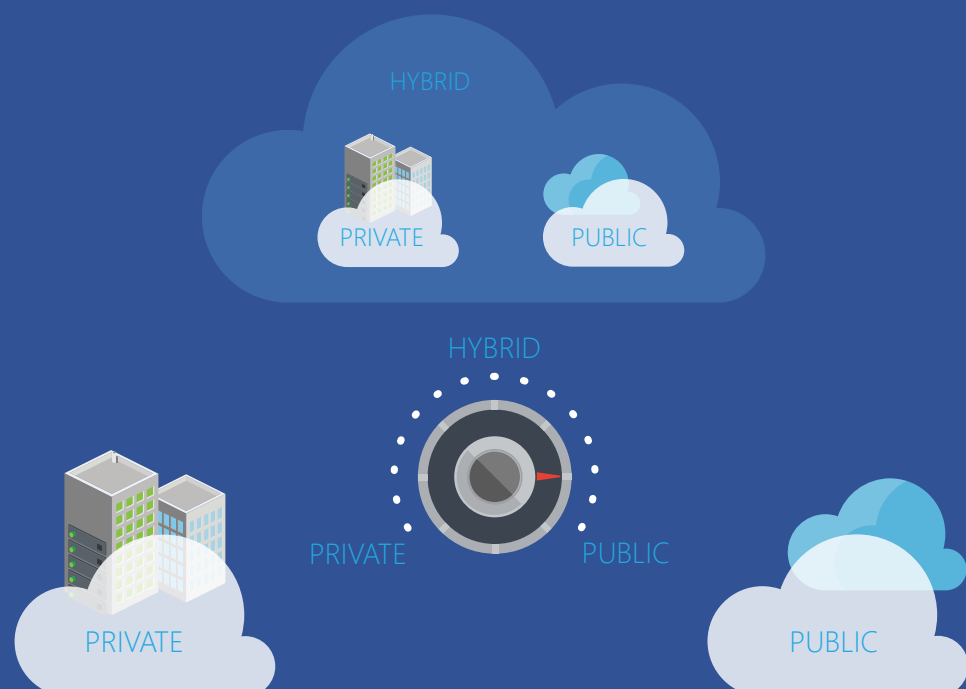


- Purchase transactions

- Catalog
- Trailers

Hybrid essentials

For these reasons, a technology approach that seamlessly links private infrastructure, platform, and/or software as a service with identical public cloud counterparts can minimize total cost, maximize business agility, and meet the constraints of real organizations run by humans and their complex decision-making processes and needs. After all, the choice of public, private or hybrid should ideally be completely separate from that of application architecture and available platform components. After all, the choice of public, private or hybrid should ideally be as simple as turning a dial and completely separate from that of application architecture and available platform components. Loosely speaking, this is something like acquiring a new car: whether to pay cash, lease, finance, or rent shouldn't impact one's selection of make and model.

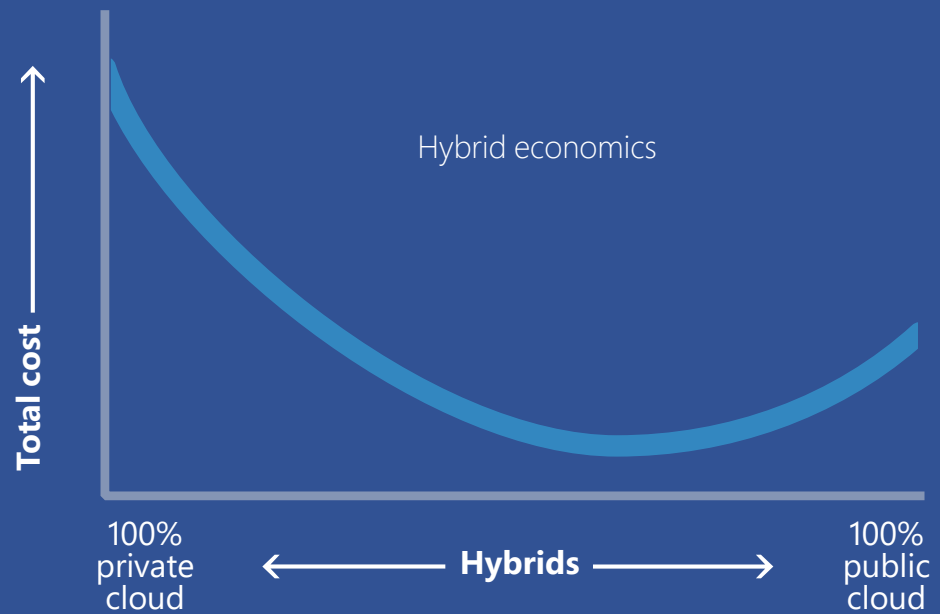


Requirements to enable this strategy are straightforward. Customers should have the option to deploy an application easily in either a private cloud, public cloud, or both without any rewriting whatsoever, and automated management tools should be available to enable various application components to autoscale and shift seamlessly back and forth between public and private and to provide key data such as application performance and cost differentials, so as to intelligently optimize between the two environments.

Beyond these high-level requirements are specific technologies for migrating virtual machines, virtual hard disks, managing configuration of virtual private networks, security, policy management, and the like.

Summary

For any given need, a single solution is rarely found—within the field of computing or outside of it. Instead, we find a mix of solution elements with various attributes: old and new, generic and custom, owned and rented. Various configurations of hybrid architectures and their corresponding business value can be enabled by choosing a strategic, foundational cloud solution that fully supports both public and private cloud environments, organizing and skilling the IT function to both run internal production and act as a broker for external cloud services, and periodically evaluating needs and optimizing for a balance of cost, time, availability, flexibility, resiliency, user experience and performance.



Author



Joe Weinman is the author of *Clouconomics: The Business Value of Cloud Computing*, and the forthcoming book *Digital Disciplines*, which explores four business strategies to exploit emerging digital technologies such as cloud computing. He has held a variety of executive management positions at organizations such as Bell Labs, AT&T, and HP. He has been awarded 21 patents in various information and communication technologies.



©2014 Microsoft Corporation. All rights reserved.

This document is provided "as-is." Information and views expressed in this document, including URL and other Internet Web site references, may change without notice. You bear the risk of using it. This document does not provide you with any legal rights to any intellectual property in any Microsoft product. You may copy and use this document for your internal, reference purposes.