egenera®

White Paper

Cloud Infrastructure Foundation

A Flexible, Responsive Cloud Is Built from the Ground Up

Executive Summary

At its heart, cloud computing is a new operational and business model for IT application hosting, rather than a new technology. It promises to deliver highly reliable, highly elastic services that respond quickly and smoothly to changing user demand. However, no matter what type of cloud environment you're considering—a large public cloud, a customizable hosted cloud or a private cloud within your own enterprise—success depends on having the right technology foundation in place.

While much of the market focus has been on the service and application cloud layers, the success of any cloud implementation is built from the ground up. The infrastructure layers must be architected and implemented properly in order to support service-level expectations. If this layer is built on technology that is itself not inherently flexible or tolerant of failure, the cloud deployment will not achieve expectations.

This paper discusses:

- The requirements and properties of a robust, highly available Infrastructure-as-a-Service layer for cloud computing
- How physical and virtual infrastructure layers need to be orchestrated automatically for a successful cloud deployment
- Why traditional approaches and tools are largely unsuitable for this task

The recommended approach to constructing cloud infrastructure is to apply PAN Manager® Software by Egenera®, which inherently provides IaaS management facilities—regardless of whether the cloud service is delivered via virtualization, physical hardware, or a mixed environment of both physical and virtual servers. Using PAN Manager massively simplifies managing both physical and virtual infrastructures, eliminates the expensive and siloed approach to provisioning new services, and integrates well with cloud management, governance processes, and self-service portals.

Defining the Cloud Computing Stack

At its heart, cloud computing is a new operational and business model for IT application hosting, rather than a new technology. Forrester Research's description of cloud computing is a good place to begin[1]:

> **A form of standardized IT-based capability—such as Internet-based services, software, or IT infrastructure—offered by a service provider that:**
>
> - Is accessible via Internet protocols from any computer
> - Is always available
> - Scales automatically to meet demand
> - Is pay-per-use
> - Offers web or programmatic control-interfaces
> - Enables full customer self-service

In essence, this is a hosted-services model, a derivative of what has been referred to as Utility Computing, except delivered universally over the internet or intranet.

Cloud Computing Doesn't Equal Virtualization

Cloud computing doesn't equal server virtualization. Indeed, cloud computing infrastructures include entirely virtualized infrastructure, entirely physical infrastructure, or a mixed infrastructure of physical and virtual servers.

For example, Google's cloud environment deploys applications that are massively horizontally deployed, so there is no need to partition servers through server virtualization technology. Even so, all of the Forrester criteria are met by Google's cloud deployment. An external hosting provider, on the other hand, typically deploys a mixed environment, with most client applications deployed on virtualized servers but with some compute-intensive applications (such as Oracle or SAP) deployed directly on physical servers. This model also meets all criteria.

The critical point is that infrastructure capacity constantly expands and/or contracts to meet user demand, and that failures are automatically bypassed so end users are protected from interruptions. Server virtualization alone is not capable of maintaining this level of service under all cloud computing deployment permutations.

Cloud Stack Components

The cloud stack contains several conceptual architectural layers (see Figure 1, at right). While many of the stack descriptions are still being fine-tuned in the market, the following definitions largely follow the NIST's *Working Definition of Cloud Computing*[2]:

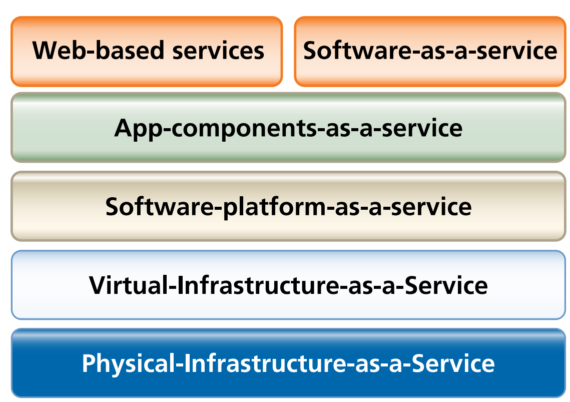

Figure 1. Cloud Stack Components

1 "The Path To Success With Server Virtualization And Cloud Computing" Frank E. Gillett, Forrester Research, November 2008
2 National Institute of Standards and Technology: Working Definition of Cloud Computing v15. http://csrc.nist.gov/

1. **Infrastructure as a Service (IaaS):** The capability provided to the user to provision processing, storage, networks and other fundamental computing resources upon which the user is able to deploy and run software, which can include operating systems and applications. IaaS may consist of two versions:

2. **Physical Infrastructure-as-a-Service (P-IaaS):** The foundational physical infrastructure elements such as physical compute, I/O, networking, load balancing, storage and availability (High Availability [HA]/Disaster Recovery [DR]) facilities.

3. **Virtual Infrastructure-as-a-Service (V-Iaas):** The virtual server management, virtual networking, and other virtualization-based components. As mentioned above, cloud architectures may not necessarily use V-IaaS .

4. **Software Platform-as-a-Service (PaaS):** The capability provided to the user to deploy onto the cloud infrastructure user-created or acquired applications created using programming languages and tools supported by the provider. The user does not manage or control the underlying cloud infrastructure including network, servers, operating systems or storage, but has control over the deployed applications and, possibly, application-hosting environment configurations. For example, this may consist of generic software platforms such as databases, application servers, etc. which run either natively (on the P-IaaS) or virtually (on the V-IaaS).

5. **Application Components-as-a-Service (A-CaaS):** The ‘personalized’ PaaS components which could then be integrated by the user into higher-level applications in the SaaS layer.

6. **Software-as-a-Service / Web-based Services: (SaaS):** The capability provided to the user to access the provider’s applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client-interface such as a web browser (e.g., web-based email). The user does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

IaaS is Fundamental to Cloud Computing Success

Because of the dynamic nature of cloud computing, the entire stack is constantly being re-purposed, re-sized and re-allocated. This requires a highly automated, reliable and elastic infrastructure to support the ever-changing user demands at the application and service layers.

The traditional approach to infrastructure, where servers, their I/O, networks and workloads are statically defined is inappropriate for this model since it inherently limits the flexibility and availability of the higher-level services (e.g., SaaS). See the Sidebar "*Why Traditional Infrastructure Approaches Can't Support Cloud Computing*" for more details.

This section discusses P-IaaS and V-IaaS properties and how they need to be closely orchestrated—constantly re-allocating physical resources, virtual resources, software payloads, networks and I/O connections—to meet user demands.

Physical Infrastructure-as-a-Service Properties

The P-IaaS layer allocates physical compute, I/O, networking, load balancing and storage components. The most essential property of the P-IaaS is to provide highly available physical CPUs (see Figure 2). This means that every CPU is failover-ready such that its workload and personality can be assumed by any other CPU in order to maintain service level agreements (SLA). These recovery servers are actual physical servers, not virtualized servers. Because everything is happening so dynamically—and as has been noted, not necessarily in a virtualized environment—this highly available server must delivered by the infrastructure itself, not by clustering, mirroring or virtualization software.

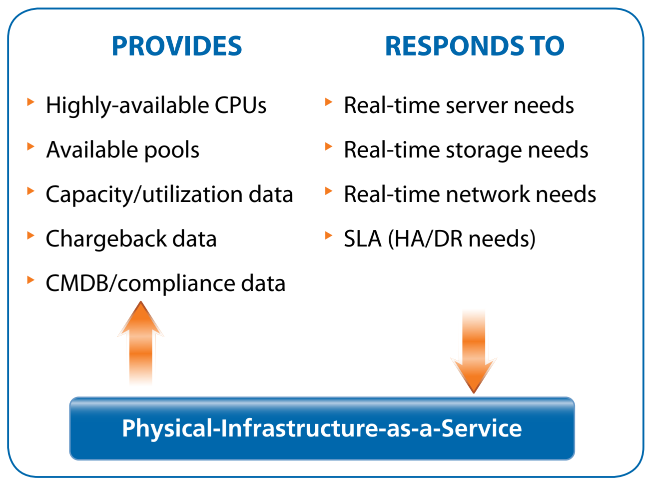

Figure 2. P-IaaS Properties

P-IaaS also should maintain one or more un-dedicated CPU pools so that raw capacity can be added or removed from duty as needed. Unlike traditional approaches, this needs to be done on the fly, defining appropriate I/O, networking and load-balancing automatically.

P-IaaS should also provide data regarding:

- Available versus committed capacity
- In-use capacity, from which chargeback data can is computed
- Raw resource dedication, from which configuration management database (CMDB), compliance and other IT service management record keeping are derived

P-IaaS also responds and adapts to real-time demand pressures. The ideal system senses SLA and/or utilization breaches, as well as changes in storage needs and/or networking needs, and automatically allocates or releases the appropriate resources with no human intervention. The P-IaaS should also adapt when SLAs are updated, allocating additional networks, ensuring HA and even ensuring entire environment DR services are available—all regardless of workload.

Why traditional infrastructure approaches can't support cloud computing

It all starts with server hardware—motherboards to be exact. When the computer industry was getting started, motherboards harbored a CPU and remedial I/O. But as processors became more sophisticated, they were integrated with complex I/O (e.g. Network Interface Cards (NICs)) as well as with storage connectivity (e.g. Host Bus Adaptors (HBAs)). Of course, there was usually a local disk. These components all added up to the motherboard's state.

This state meant that the server retained static data, specifically things like I/O addressing, storage connectivity naming and local disk data. Usually the local network had state, too, ensuring that the IP and MAC address of the motherboard were attached to switches and LANs in a particular way. For critical applications, all of these components (plus naming/addressing) were frequently duplicated for redundancy.

To replace (or clone) this server meant re-configuring all of these addresses, names, storage connections and networks—and sometimes in duplicate. As this complexity rose, so did administrative staff, operational costs and errors.

In response, vendors developed special-purpose clustering and failover software—necessarily closely coupled to specific software and hardware—to provide the re-assignment of state to the new hardware and networking. This software often required hand-crafted integration and frequent testing to ensure that all of the addressing, I/O, and connectivity operations worked properly. Many of these special-purpose systems are still in use today.

Similarly, there are equally complicated software packages for scale-out and grid computing that perform similar operations for cloning hardware to scale out systems for parallel computing, databases, etc. These systems are equally complex and usually application-specific, due again to the need to replicate stateful computing resources.

So the industry, in an effort to add "smarts" and sophistication to the server—to enable it to fail over or to scale—has instead created complexity and inflexibility.

Virtual Infrastructure-as-a-Service Properties

Much of the V-IaaS stack layer is analagous to the P-IaaS layer insofar as virtual machines also need to be dynamically created, allocated and de-allocated as needed. They also need to have their own I/O, networking and storage connections established. Each virtual machine needs to be provisioned with its O/S and application payload; and each needs to be monitored and tracked in a similar IT service management catalog, chargeback tools, etc.

One distinction, however, is that HA and DR services are not needed in the V-IaaS layer since they are already part of the P-IaaS layer. So, should hardware fail, the P-IaaS layer replicates the failed virtual hosts, networks, etc. followed by the virtual machines and guest O/Ss. To the higher-level layers in the cloud stack, a hardware failure merely manifests as a number of virtual server re-boots. Thus the need for "virtual" HA & DR is eliminated by the P-IaaS layer.

Interplay between P-IaaS and V-IaaS

There is more or less interplay between the P-IaaS and the V-IaaS layers depending on whether the cloud environment is physical, virtual or a mixed physical/virtual deployment.

For example, in a web server farm or large-scale database instance, applications are easily provisioned directly on top of the P-IaaS foundation with no need for V-IaaS. The P-IaaS continually senses SLAs, capacity and utilization status, and can take action to provision additional physical machines, load balancers, etc. to handle demand—and to release resources when appropriate.

However, with virtual environments, each CPU would first be automatically provisioned with a virtual host—presenting an elastic pool of hosts for guest virtual machines (VMs) as demand warrants. Traditionally, a virtual machine management (VMM) tool then transparently operates on top of the P-IaaS foundation. As new hosts are needed, the P-IaaS elastically provisions them onto bare-metal and signals to the VMM that new hosts are available. Conversely, when the VMM consolidates VM guests, it signals the P-IaaS management system that unused hosts can be released back into the unassigned physical resource pool for later use.

It Starts with Logically Defining the CPU

To meet the promise of cloud computing, the P-IaaS layer needs to manage an abstracted logically defined view of the CPU, including I/O, network, storage and data switching, load-balancers, etc. If this is not in place, the physical infrastructure—with its state and connections—cannot be logically defined and invoked in real-time. (See the Sidebar *"Why Don't Traditional Infrastructure Approaches Support Cloud Computing"* on how the traditional approach fails to meet cloud computing requirements.) With the physical infrastructure abstracted, it becomes irrelevant what software payload the server will handle.

For example, if additional physical capacity is needed (for a physical scale-out database or because a new virtual host is required), any physical server can be allocated, assigned the appropriate types/quantities of I/O, assigned to networks and paired with load-balancing simply and easily.

When a failure occurs, the entire state of the failed machine—including software—is replicated on a new piece of hardware, complete with original addressing, storage naming, network names, etc. The implication is clear: for most instances, this abstracted view of the physical infrastructure renders unnecessary the need for clustering and high-availability systems at the virtualization and/or software levels.

Physical IaaS Management and Control Requirements

The P-IaaS layer holistically manages the CPU, I/O, networking and storage allocations, allocating and defining infrastructure as needed. But it also needs to provide a control set that affects change on the infrastructure, as well to provide appropriate monitoring of activity and performance. Without this overall control and monitoring function, the entire infrastructure—including both physical and virtual elements—would not be able to quickly and automatically adjust to meet SLAs.

From a monitoring perspective, the P-IaaS system needs to track real-time variables in the environment, and then affect change on the environment (See Figure 3).

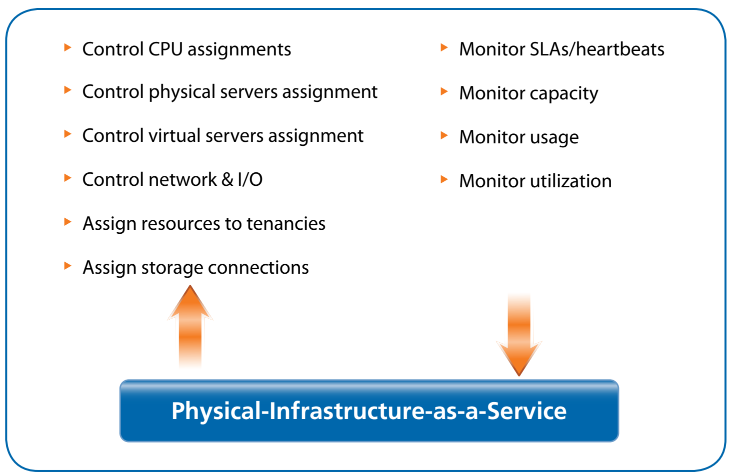

Figure 3. P-IaaS Control and Monitoring Requirements

Crucial variables to monitor include:

- Individual server SLAs & heartbeats, to detect potential failures at the hardware level
- Overall pooled compute capacity, as well as in-use capacity (which may vary from tenant to tenant). This information triggers allocation or de-allocation of resources.
- Individual physical server utilization, which indicates opportunities for de-allocation or increased consolidation of resources

In turn, a control mechanism manipulates or orchestrates the interaction of the following infrastructure components:

- **CPUs/Servers:** Maintain and allocate a pool of stateless servers, i.e. servers that do not have I/O addressing, storage naming, dedicated networking, etc.
- **Virtual servers:** Allocate transparently either physical servers or virtual servers as needed
- **Server I/O:** Assign appropriate types and names of I/O ports to both physical and virtual servers. Attach to or create appropriate networks for each I/O port.

- **Storage:** Create appropriate storage connections (e.g. WW names) to each storage data port
- **Tenancies:** Maintain segregated resource pools, networks, etc. between multiple infrastructure tenants
- **SLAs:** Re-allocate automatically a physical server state to another server should a point of failure occur, yielding HA. Re-allocate automatically an entire environment to a new environment should disaster strike, yielding DR.

PAN Manager Quickly And Simply Implements a Robust IaaS

PAN Manager Software by Egenera delivers the highly available, flexible and automatic IaaS layer required to support today's cloud computing environments. It not only delivers the required properties of the P-IaaS and V-IaaS layers, it also brings simple management and control tools that enable the tight interplay between the P-Iaas and V-Iaas layers.

Since PAN Manager was designed from the ground up to meet these requirements, its implementation is simple and quick. In some situations, a production site can be up and running within days. Finally, PAN Manager has been proven in production for over eight years, with over 1,400 installations worldwide.

Software Virtualization	PAN Manager
Logically abstracts software	Logically abstracts hardware infrastructure
Allows for software portability	Allows for infrastructure portability
Permits software agility	Permits I/O, network and storage agility
Simplified logical server provisioning	Simplified logical infrastructure provisioning
Consolidates virtual servers	Consolidates I/O, networking

How does PAN Manager work? By:

- Removing state from the server through abstracting (i.e., virtualizing) all I/O, including NICs, HBAs, and KVM ports, so that they are reconfigured entirely in software.
- Abstracting the network and its components in software, essentially creating a converged network that carries both data and storage information, and then by creating reconfigurable switches and virtual networks.

The result is an elegant, simple, and easy-to-configure infrastructure that provides inherent:

- **Rapid provisioning:** Because provisioning servers isn't just software, but includes assigning infrastructure too
- **Support for any software payload:** Because infrastructure orchestration is agnostic to software payload
- **Inherent universal HA and universal DR:** Because server state and environment properties (networks, switches, load balancing, etc.) are defined logically in software

PAN Manager abstracts I/O, network, switches and storage the way that a hypervisor abstracts the O/S and applications

The software generates a P-IaaS by creating a pool of standard stateless physical x86 servers, and offers two approaches to providing a V-IaaS:

- Through an embedded VMM environment directly within the management facility, and/or
- By supporting separate third-party VMMs as a distinct layer above its P-IaaS

Thus, it is ideal both for hosting providers, who usually offer a 100% virtualized cloud-based offering, as well as for enterprise IT, who may need to support a number of mixed physical- and virtual applications in a cloud- or utility-computing-based environment.

PAN Manager Components and Functionality

The main functional components of the system include:

Stateless Environment Controls

PAN Manager eliminates the need to maintain server information on locally attached storage devices. Instead, operating system images, network names, and addresses are stored within the communication fabric and SAN, and then invoked when the server is booted. In this way, PAN Manager quickly moves applications to alternate servers without the use of clustering or mirroring software.

Converged Network Fabric

The communication fabric is a converged intra-server network that transmits storage, network, and control data internally among system components and between servers and shared external interfaces. The fabric software dynamically manages available bandwidth and provides a re-configurable, redundant and reliable communication system.

I/O Virtualization

I/O Virtualization utilizes the fabric to create logical paths between servers, their storage and network resources. It lets the system provision servers configure, and manage HBAs and NICs for each server with software, eliminating the need to cable. I/O Virtualization enables shared access to resources while protecting data flows from unauthorized access.

Server Profiles / Server Definition Management

The server definition profile is a portable XML specification of each application service within the environment. PAN Manager uses the specifications to provision servers and move applications quickly from one server to another, whether physical or virtual. The definition includes the operating system and application images as well as bindings to the application's storage and network resources.

On-demand Provisioning

With PAN Manager, IT can rapidly deploy server definitions on any physical or virtual server. This wire-once capability enables IT staff to allocate and repurpose servers from the management console, without requiring manual configuration.

Automated Management and Monitoring

PAN Manager monitors each server's health for resource drains and component failures, and initiates recovery operations according to established user-defined policies. Recovery may take the form of a software or hardware restart. Alternatively PAN Manager can migrate the application and its resources to another server.

Chargeback Management

Chargeback Management collects utilization data that are imported into billing and analysis applications. These metrics help IT better determine how costs are distributed, how to effectively map applications to available servers, and how to plan for future resource investment.

Integrated Virtual Server Management

PAN Manager supports industry-standard server virtualization products, allowing IT to manage both physical and virtual computing resources with one set of tools and processes. In addition, PAN Manager seamlessly integrates virtual machine technology, providing integrated support for virtual servers.

Universal High Availability and Scaling

PAN Manager's N+1 availability allocates one server as backup to multiple production servers. This feature automatically moves applications to alternate physical and/or virtual servers, due to component failure or the need for more CPU and memory.

Universal Disaster Recovery

PAN Manager can migrate all or part of the data center configuration to a secondary location. PAN Manager moves each physical and virtual server configuration in its entirety, including operating system, hypervisors, network configuration and applications. With this approach, a single site can be used as backup for multiple production sites.

Simplified Control and Interaction

From a cloud management perspective, PAN Manager Software provides a number of control methods to help govern its actions. The console is a web-based GUI with a mature set of controls available to administrators. Entire application, server and environment profiles are created for future manual (or automatic) instantiation, and domain sizes and properties can be changed in real-time (e.g. to be used in multi-tenancy scenarios). The GUI is simple to use although its single management window enables users to control dozens of variables.

PAN Manager can also be used in an embedded management scenario, where the GUI is not the primary source of control. Instead, both a command-line interface and a web-services API are available for integration with other administrative and self-serve consoles (see Figure 4). Further, the system's resource configuration collector (RCC) provides step-by-step records of every environmental variable for use by chargeback/accounting, configuration management and other external monitoring systems.

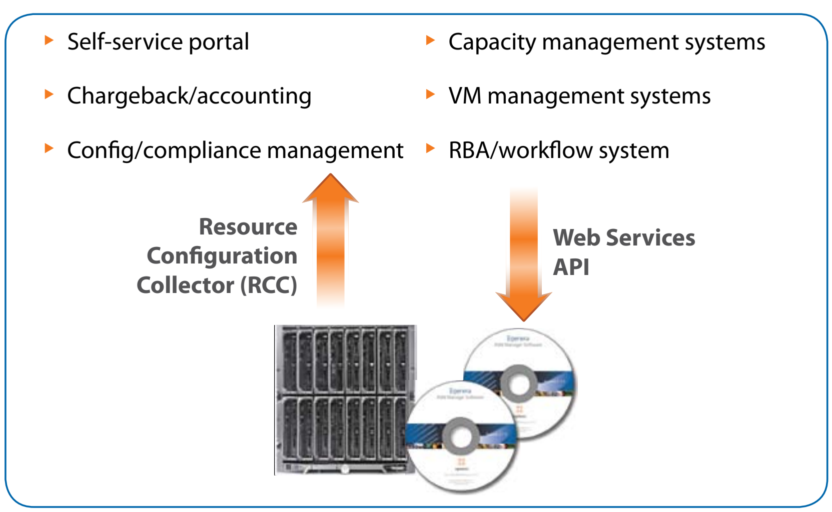

Figure 4. Cloud Management Services

PAN Manager Use Cases

Public Iaas: Building a Hosted Compute Service - SAVVIS

In 2004, SAVVIS set a goal to become the industry's first totally virtualized utility computing data center, integrating virtualized servers, storage, networks and security into an end-to-end solution. Today, the service provider houses over 1,425 virtual servers running on 70 industry standard servers, 370 terabytes of 3PAR® storage and 1,250 virtualized firewalls.

SAVVIS has built huge, scalable service platforms that can be leveraged by multiple tenants/clients with full security based on PAN Manager Software. The result is a powerful value proposition that gives SAVVIS clients:

- Easy access to a broad range of hosting services including data centers, operations and managed hosting
- Access to the latest technologies without upfront capital costs or the risk of obsolescence
- Hosted services at half the price of traditional outsourcing
- Exceptional performance guarantees
- New resources provisioned in minutes
- Services that can grow—or shrink—dynamically
- Faster application time to market
- Better agility, flexibility and responsiveness to market opportunities

As SAVVIS' CTO observed, "The beauty of Egenera is that the physical element is fungible. You don't care where an application is running."

Internal IaaS: Utility Computing - U.S. Census Bureau

To continue providing the best mix of timeliness, relevancy, quality and cost for the data it collects and services it provides, the U.S. Census Bureau launched an initiative to overhaul its data center by establishing the Census Bureau Utility Computing Environment (CBUCE).

A directive of the Census CIO, CBUCE sought to reduce the cost and complexity of IT infrastructure by recentralizing the Bureau's IT function. PAN Manager Software is being used for this set of mission-critical, compute-intensive applications in the new utility environment. The first project, Master Addresss File (MAF), is designed to be a complete and current list of all addresses and locations where people live or work, covering an estimated 115 million residences as well as 60 million businesses and other structures in the U.S. The results of this internal IaaS hosting model includes:

- Cost savings
- Better service levels, resource utilization and IT responsiveness
- Faster application time to market
- Simpler management
- Mission-critical level of reliability across all applications

CBUCE has become such a showcase facility that other government agencies are looking to come into the environment.

Next Steps

Building a reliable, flexible cloud computing infrastructure starts with a reliable, flexible Infrastructure-as-a-Service foundation. That foundation must inherently be physical/virtual agnostic; highly adaptable for automatically provisioning/releasing of pooled resources; and able to provide availability and disaster recovery services across all payloads.

PAN Manager integrates well within existing accounting, ITSM, governance, and user self-service portals—whether it is being used within an enterprise or within a hosted-services provider. It is simple to implement, enabling a production site to be online within days in some cases.

Learn more about how PAN Manager, proven in production for over eight years with over 1,400 installations worldwide, can power your cloud environement. Visit our website at www.egenera.com or call us at +1-508-858-2600.

egenera®

East Coast Headquarters
Egenera, Inc.
165 Forest Street
Marlboro, MA 01752
U.S.A.
Main: 508-858-2600
Sales: 508-858-3600
Fax: 508-481-3114
www.egenera.com

West Coast Headquarters
Egenera, Inc.
5201 Great America Parkway
Suite 446
Santa Clara, CA 95054
U.S.A.
Main: 508-858-2600

Europe/Middle East/Africa
Egenera Ltd.
Egenera Ltd.
Squires House
205A High Street
West Wickham
Kent BR4 0PH
United Kingdom
Voice: +44 208 133 8678

Asia Pacific Headquarters
Egenera, K.K.
Shinjuku NS Bldg. 6F,
2-4-1 Nishishinjuku,
Shinjuku-ku
Tokyo 163-0806 Japan
Voice: +81-3-5321-7157
Fax: +81-3-5321-7158
japan.egenera.com