**Identity Infrastructure as a Service**

According to Nicholas Carr’s analysis of IT’s evolution into a utility[4], infrastructure-like power generation began with a dedicated model: only organizations with large amounts of capital could afford to build and operate their own waterwheels and, later, electrical generators. This infrastructure gradually became centralized at a utility - a power plant or telephone company, for instance - due to economic forces and democratization. Finally, the estab­lishment of distribution standards and a delivery network led to wide adoption of a utility service, resulting in econo­mies of scale, cost reductions, and new capabilities.

Identity management for the cloud must also evolve to the point of being standardized and accessible by multiple applications and users. Why must it be built differently than conventional single-tenant, behind-the-firewall identity management software? Nobody has success­fully retrofitted cloud constructs and models onto legacy software, which was created with a different set of assump­tions. To be successful in the cloud era, organizations and vendors must fundamentally rethink how they manage, deliver, and consume identity.

The concept of *identity infrastructure as a service*fol­lows two major trends: the evolution of IT from capital infrastructure to a service, and the consumerization of IT.

**Identity as a service**

Rather than investing heavily in identity when developing or using an app, it makes more sense for an organization to utilize a service for its identity needs. Iden­tity as an infrastructure presents an on-demand model that delivers the right amount of capability at the right time. Companies need to think less about identity technology and focus instead on service-level agreements and ser­vice management, not infrastructure. This means moving from a company-owned to a service-provider-owned and -operated identity management approach.

An identity management solution that requires making a capital investment to use a cloud service does not reflect this shift. If each person had to buy a cell phone tower just to use a mobile phone, how many people would have mobile phones? Mobile phones are ubiquitous because users do not have to buy any infrastructure to use the net­work. They simply pay a subscription for the service and get access to the entire shared cellular infrastructure.

Consider the Internet’s Domain Name System (DNS) backbone. Major DNS name services provide these services transparently and reliably. How the provider actually runs the name servers is irrelevant. All that matters is that name services are readily available and perform as expected. Identity services must be as transparent and reliable as DNS services.

**Consumerization**

Each of us uses sophisticated, consumer-based Web apps every day in our personal lives. For example, millions of users access Amazon to search for and purchase items; hundreds of millions of people connect via Facebook; and iTunes offers millions of songs and thousands of TV shows, movies, podcasts, and audiobooks for downloading. Our experience with such sites leads us to expect enterprise systems to be just as simple, effortless, and reliable. More­over, we all prefer a “freemium” model that lets us try out a service without a commitment or investment.

We also like to do things ourselves. We generally want to use apps without IT’s help and with a minimum of administration, and we buy technology like smartphones and laptops at the grassroots level, not centrally from a company. Anyone can set up a rented development envi­ronment with a credit card and the same Amazon account they use to buy books and other goods. The same goes for identity—the line between professional and personal per­sonas has blurred. When you use Twitter or Facebook for work, for example, are you an employee or a consumer? It is often very difficult to distinguish between the two.

Cloud service providers will deliver better identity solu­tions to their customers by utilizing a specialist’s expertise rather than building it themselves. They must build scal­ability into every design decision from the beginning. Too often, scalability is an afterthought or a reaction to an overloaded system. Developers set performance ex­pectations when they build an app, and if it meets those expectations, it is deemed scalable. But problems arise when expectations change, as is the case with an unex­pectedly successful viral adoption or when porting an app server to the cloud. Using a hypervisor to abstract compute power from the app is a simple way to scale up, but most apps are not architected to readily leverage the *N*+1 scale that clouds offer.

By consolidating identity infrastructure, service pro­viders can achieve economies of scale. As with solid-state drives, the number of moving parts is reduced, resulting in greater uptime and reliability. Each identity integration point becomes a stress point, and each credential creates a broader attack surface and potential help desk expense.

**Looking Ahead**

Many large-scale identity ecosystems have the disrup­tive potential to accelerate change in cloud identity. Google Apps, which launched in March 2010, acquired more than 27 million users in its first month[5]; obviously, that number is much higher today. Other examples of such systems in­clude Twitter, which has 106 million registered users and is adding more than 300,000 per day[6], and Google Gmail, which boasts more than 170 million monthly users[7].

Facebook has exploded in popularity, with more than 550 million users—more identities than the combined populations of the US, Canada, the UK, and Italy[8]. This massive number of identities can be integrated with the identity fabric en masse due to the walled-garden aspect of Facebook’s platform and user community. When Face­book introduced support for identity sharing via OpenID in 2009, hundreds of millions of people suddenly had OpenID credentials.

When Google decided to support OpenID a year earlier, it brought more than 100 million users into the market overnight. The implication for cloud identity is that con­sumer authentication models that have been proven to scale to hundreds of millions will be part of the identity fabric and give it critical mass and scale. These examples clearly indicate that an identity access fabric linking enter­prises to the cloud is not only relevant but also necessary.

Cloud growth can be compared to a city that builds homes at an explosive pace without building the infrastructure of roadways needed to support the influx of traffic. In the case of identity management, the cloud has not kept pace with the enormous volume of user identities that network administrators must manage and secure.

To realize the cloud’s benefits, enterprises must have an identity infrastructure in place that can overcome the limitations of precloud identity architectures. This means using an identity fabric that links many apps to a single identity. The proliferation of identities demands a better identity management solution, not only to ease the burden on IT administrators’ time to manage them but also to address security and privacy concerns.

An identity fabric that provides a secure linkage be­tween the enterprise and the cloud, while reducing the number of identities, is the clear answer to enabling full-scale cloud adoption. Cloud-based identity management delivered as an infrastructure service with on-demand dial-tone quality benefits users, network administrators, application vendors, and service providers in dramatic ways. The cloud, with all its ubiquitous technology, must also make the identity management fabric ubiquitous, as its growth and acceptance depend on it.

### Auditing and compliance

A key challenge in auditing in the cloud is overcoming the lack of visibility into user access of SaaS apps. Using the public Internet rather than connecting to a company network puts users beyond the scope of network monitor­ing tools.

Unlike most enterprise networks, the cloud is globally accessible. However, regulatory compliance require­ments vary by jurisdiction and are complex and often contradictory. The industry needs a framework to meet global jurisdictional challenges. Identity is central to such a framework because many regulations center on user privacy and access.

### Cloud platform architectural requirements

The cloud has brought with it new architectures and platforms that service providers have yet to make identity-aware. Specifically, many cloud service providers offer storage- or database-as-a-service via hosted hypervi­sors like KVM or those from VMware and Xen, but such IaaSs currently lack identity and access management as a service.

With their high utilization rates, virtualized platforms cannot handle the overhead associated with precloud Web access management (WAM) use of webserver plug-ins and agents. The tight coupling of WAM with plug-ins has proven to be brittle, and the “burstable,” elastic nature of virtualized cloud platforms makes the plug-in model infeasible. The industry requires a proxy-based approach that does not place load on the virtualized Web and ap­plication servers.

In the case of SaaS apps, the identity integration challenge of enforcing access control and supporting audits stems from multitenancy and the fact that the SaaS provider owns and operates the underlying infrastructure, making it impossible to install dedicated agents or plug-ins for each application instance. Also, with most SaaS apps, collecting audit logs is problematic because they are often comingled with other tenant data. In some cases, the audit details are insufficient for answering key forensic ques­tions. What is needed is a loosely coupled, noninvasive identity management platform that can enforce policy upstream from the SaaS apps themselves.

http://www.infoq.com/cn/articles/architecting-cloud-scale-identity