Cloud computing explained

—by Eric Wolf

Green Mountain Falls, Colorado, is a small town of 969 people in the Rocky Mountains just west of Colorado Springs. When Trustee Marshall Worthey decided the town needed a more substantial, spatially enabled web presence, he was faced with a problem: The town's official IT infrastructure consisted of four desktop computers and a cable modem, and the software budget definitely excluded ArcServer as even a consideration. So Marshall decided to bring this 19th century mountain town into the 21st century. He decided to look to cloud computing.

Brief history

Cloud computing is far from revolutionary. In fact, it is a fairly natural evolutionary change in IT infrastructure. In the early 1980s, most people used stand-alone computers. Information and the applications used to work with these computers were marooned on the computer's desktop. By the end of that decade, more and more offices were networking their computers, allowing information to be shared easily through servers from companies like Novell. Client-server technology moved some of the applications off the desktop and to the server.

In the 1990s, the Internet made its transition from the domain of research and academia to the general public. Sharing information among offices and businesses became fairly easy. By the end of the 1990s, the Internet also began to spawn Application Service Providers who freed the IT department to maintain servers. ASPs also provided knowledge fine tuned to the applications they hosted. This was a real boon to anyone who tried to use web mapping platforms.

Three models of cloud computing

At the turn of the 21st century, Salesforce.com, instead of hosting applications created specifically for Application Service Providers began charging for access to the applications running on their servers. This model provided two significant advantages. First, the low monthly cost was easier for small companies to handle compared to having to commit to buying expensive servers and licenses to run the applications they needed. Second, the applications allowed a spread-out sales force to collaborate and share information without complex Virtual Private Networks. What Salesforce.com's applications lacked in features, they made up for with low cost of entry and ease of collaboration.

Around the same time, Amazon chose to restructure their IT infrastructure. While many people equate Amazon with books, they really are a distribution company with a comprehensive IT infrastructure for managing sales and inventory.

In the wake of the dot-com bubble, Amazon expanded their network and data center capacities. But their IT people were fraught by the fact that their infrastructure normally operated at a fraction of capacity in order to handle spikes in demand.

Amazon utilized virtualization technology that allows a computer to emulate one or more computers in software. By allowing these emulated computers to move around from physical computer to physical computer, Amazon was able to better economize their server use while increasing reliability.

Like Amazon, Google also leveraged a positive cash flow after the dot-com bubble to buy up data centers and network capacity throughout the country. Google's grand mission of indexing all of the world's information required computation power unheard of in terms of storage, processing, and communications.

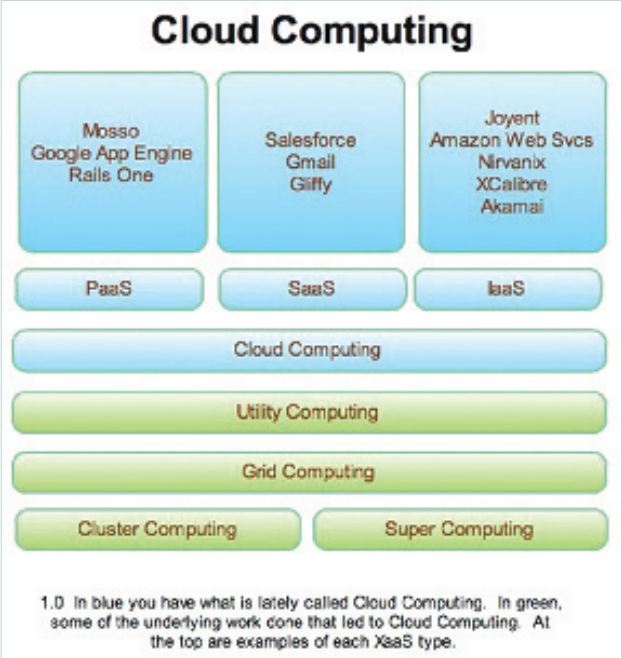
It's not hard to guess that if you have hundreds of thousands of computers, you are likely to have many of those computers failing at any one time. Google needed to figure out a way of developing applications that could leverage the power of many computers. To this end, Google created the MapReduce programming model which is based on concepts from functional programming and cluster computing. While MapReduce does leverage the power of massively distributed systems, it required a complete rewrite and redesign of software.

These three companies' initiatives embody the three main models of cloud computing. Salesforce.com provides an example of Software as a Service (SaaS) as does Google Docs. Amazon created what is now called Infrastructure as a Service (IaaS). Google's model is commonly referred to as Platform as a Service (PaaS).

Like cloud computing in general, none of these models is entirely new; they reflect a step in the evolution of technology. Cloud computing and the three models have the following common properties: elasticity, device/location independence, and a utility pricing model.

Elasticity refers to the ability of the cloud to provide as much (or as little) computational power and storage as your application demands. Device and location independence frees the application from the reliability issues associated with a single machine or even a single data center. Utility pricing means you only pay for what you use, with little or no initial costs other than the costs associated with moving your application to the cloud.

Cloud computing, p. 33



1.0 In blue you have what is lately called Cloud Computing. In green, some of the underlying work done that led to Cloud Computing. At the top are examples of each XaaS type.

Examples

Cloud computing is especially significant for geospatial information. Building a spatially enabled web application for the town of Green Mountain Falls, Colorado, would have been impossible given the budget constraints of the small town. By leveraging the cloud, Marshal Worthey found that the application could be created with similar development costs as a stand-alone web mapping server application running ESRI ArcServer, but with significantly lower ongoing infrastructure costs.

The US Geological Survey's Earthquake website provides near real-time maps of seismic activity. This site has seen as many as 250,000 accesses per second immediately following an earthquake. The USGS had two options to manage this traffic successfully: create a massive IT infrastructure to handle the activity—and have that infrastructure idle most of the time—or move it out to the cloud. All of the content on earthquake.usgs.gov is mirrored across thousands of servers by IaaS provider Level 3 Communications. The USGS is only billed for the actual activity on the site.

Working with about a dozen government and non-government organizations, FortiusOne created an application for real-time monitoring of elections in Afganistan. FortiusOne relied on cloud computing to collect, aggregate, and map information about tribal demographics, violence, and political support to determine where such issues would impact voter turnout. The application needed to be tested and rolled out in a matter of weeks. Not only did cloud computing provide a platform insulated from the harsh Afgan environment, it also allowed the various agencies to work together to create a successful application. Because of the low cost of operating the application, it is still used by humanitarian agency staff working on the ground in Afghanistan.

Cloud, from p. 31

The Greater New Orleans Community Data Center (GNOCDC) is a non-profit tasked with democratizing data about the Crescent City. Though the situation in post-Katrina New Orleans is extreme, the lessons learned are universal—how to be responsive to a rapidly changing environment, nimbly scale when demand for data spikes, store and process data remotely to maximize collaboration, and streamline workflows so a small staff can do big things with minimal error.

GNCOCDC is using a combination of Salesforce.com to manage data requests from their website and WeoGeo.com to handle the management of the spatial datasets for content. Being able to use the open APIs of WeoGeo to help their customers discover, customize, and download datasets in a hosted, scaleable product has freed them from managing their IT infrastructure. Instead they can now concentrate on creating data products for their users. Because WeoGeo is able to convert more than 250 different file formats, users downloading datasets can request data in native formats they are familiar with, rather than spending time trying to figure out what software they need to use to do this.

Conclusions

Cloud computing continues to be touted as the "next big thing" in IT—even though it has all the signs of being a disruptive technology. Feature sets in cloud-based applications are not as rich as desktop applications, but many are "good enough". Further, cloud computing provides a level of collaboration that far exceeds anything available on the desktop. Given its low initial and on-going costs, we are likely to see many projects leverage the cloud.

An ecosystem of smaller, faster applications is quite likely to result in software solutions that are hard to live without. Fortunately, as those applications become popular, the elasticity of the cloud platform will allow it to grow with the number of users.

In fact, this article was written using Google's docs word processor. Because a cloud computing application was used, it was very easy for many people to contribute to it, including Eric Wolf in Denver, Colorado, and Scott Haefner in Menlo Park, California, from the USGS. Sean Gorman in the Washington D.C. area contributed content, which was reviewed by his colleague Andrew Turner, who happened to be in England at the InterGEO conference. James Fee, from WeoGeo added a paragraph while attending WhereCampPDX in Portland, Oregon. Marshall Worthey from Green Mountain Falls, Colorado, and Adam Estrada from ZekiahTech in La Plata, Maryland, reviewed the facts on their cloud computing-based application.

Despite the spatial distribution of the contributors, cloud computing provided greater collaboration than would have been possible even compared to using e-mail to share word processing files.