Inside MySpace.com
January 16, 2007

By David F. Carr

Booming traffic demands put a constant stress on the social network's computing infrastructure. Yet, MySpace developers have repeatedly redesigned the Web site software, database and storage systems in an attempt to keep pace with exploding growth - the site now handles almost 40 billion page views a month. Most corporate Web sites will never have to bear more than a small fraction of the traffic MySpace handles, but anyone seeking to reach the mass market online can learn from its experience.

Story Guide:

●

A Member Rants

● **The Journey Begins**

Membership Milestones:

● **500,000 Users: A Simple Architecture Stumbles**
● **1 Million Users:Vertical Partitioning Solves Scalability Woes**
● **3 Million Users: Scale-Out Wins Over Scale-Up**
● **9 Million Users: Site Migrates to ASP.NET, Adds Virtual Storage**
● **26 Million Users: MySpace Embraces 64-Bit Technology**
● **What's Behind Those "Unexpected Error" Screens?**

Also in This Feature:

● **The Company's Top Players and Alumni**
● **Technologies To Handle Mushrooming Demand**
● **Web Design Experts Grade MySpace**
● **User Customization: Too Much of a Good Thing?**

Reader Question: Is MySpace the future of corporate communications? Write to: baseline@ziffdavis.com

Next page: A Member Rants

A Member Rants

On his MySpace profile page, Drew, a 17-year-old from Dallas, is bare-chested, in a photo that looks like he might have taken it of himself, with the camera held at arm's length. His "friends list" is weighted toward pretty girls and fast cars, and you can read that he runs on the school track team, plays guitar and drives a blue Ford Mustang.

But when he turns up in the forum where users vent their frustrations, he's annoyed. "FIX THE GOD DAMN INBOX!" he writes, "shouting" in all caps. Drew is upset because the private messaging system for MySpace members will let him send notes and see new ones coming in, but when he tries to open a message, the Web site displays what he calls "the typical sorry … blah blah blah [error] message."

For MySpace, the good news is that Drew cares so much about access to this online meeting place, as do the owners of 140 million other MySpace accounts. That's what has made MySpace one of the world's most trafficked Web sites.

In November, MySpace, for the first time, surpassed even Yahoo in the number of Web pages visited by U.S. Internet users, according to comScore Media Metrix, which recorded 38.7 billion page views for MySpace as opposed to 38.05 billion for Yahoo.

The bad news is that MySpace reached this point so fast, just three years after its official launch in November 2003, that it has been forced to address problems of extreme scalability that only a few other organizations have had to tackle.

The result has been periodic overloads on MySpace's Web servers and database, with MySpace users frequently seeing a Web page headlined "Unexpected Error" and other pages that apologize for various functions of the Web site being offline for maintenance. And that's why Drew and other MySpace members who can't send or view messages, update their profiles or perform other routine tasks pepper MySpace forums with complaints.

These days, MySpace seems to be perpetually overloaded, according to Shawn White, director of outside operations for the Keynote Systems performance monitoring service. "It's not uncommon, on any particular day, to see 20% errors logging into the MySpace site, and we've seen it as high as 30% or even 40% from some locations," he says. "Compare that to what you would expect from Yahoo or Salesforce.com, or other sites that are used for commercial purposes, and it would be unacceptable." On an average day, he sees something more like a 1% error rate from other major Web sites.

In addition, MySpace suffered a 12-hour outage, starting the night of July 24, 2006, during which the only live Web page was an apology about problems at the main data center in Los Angeles, accompanied by a Flash-based Pac-Man game for users to play while they waited for service to be restored. (Interestingly, during the outage, traffic to the MySpace Web site went up, not down, says Bill Tancer, general manager of research for Web site tracking service Hitwise: "That's a measure of how addicted people are—that all these people were banging on the domain, trying to get in.")

Jakob Nielsen, the former Sun Microsystems engineer who has become famous for his Web site critiques as a principal of the Nielsen Norman Group consultancy, says it's clear that MySpace wasn't created with the kind of systematic approach to computer engineering that went into Yahoo, eBay or Google. Like many other observers, he believes MySpace was surprised by its own growth. "I don't think that they have to reinvent all of computer science to do what they're doing, but it is a large-scale computer science problem," he says.

MySpace developers have repeatedly redesigned the Web site's software, database and storage systems to try to keep pace with exploding growth, but the job is never done. "It's kind of like painting the Golden Gate Bridge, where every time you finish, it's time to start over again," says Jim Benedetto, MySpace's vice president of technology.

So, why study MySpace's technology? Because it has, in fact, overcome multiple systems scalability challenges just to get to this point.

Benedetto says there were many lessons his team had to learn, and is still learning, the hard way. Improvements they are currently working on include a more flexible data caching system and a geographically distributed architecture that will protect against the kind of outage MySpace experienced in July.

Most corporate Web sites will never have to bear more than a small fraction of the traffic MySpace handles, but anyone seeking to reach the mass market online can learn from its example.

Next page: The Journey Begins

The Journey Begins

MySpace may be struggling with scalability issues today, but its leaders started out with a keen appreciation for the importance of Web site performance.

The Web site was launched a little more than three years ago by an Internet marketing company called Intermix Media (also known, in an earlier incarnation, as eUniverse), which ran an assortment of e-mail marketing and Web businesses. MySpace founders Chris DeWolfe and Tom Anderson had previously founded an e-mail marketing company called ResponseBase that they sold to Intermix in 2002. The ResponseBase team received $2 million plus a profit-sharing deal, according to a Web site operated by former Intermix CEO Brad Greenspan. (Intermix was an aggressive Internet marketer—maybe too aggressive. In 2005, then New York Attorney General Eliot Spitzer—now the state's governor—won a $7.9 million settlement in a lawsuit charging Intermix with using adware. The company admitted no wrongdoing.)

In 2003, Congress passed the CAN-SPAM Act to control the use of unsolicited e-mail marketing. Intermix's leaders, including DeWolfe and Anderson, saw that the new laws would make the e-mail marketing business more difficult and "were looking to get into a new line of business," says Duc Chau, a software developer who was hired by Intermix to rewrite the firm's e-mail marketing software.

At the time, Anderson and DeWolfe were also members of Friendster, an earlier entrant in the category MySpace now dominates, and they decided to create their own social networking site. Their version omitted many of the restrictions Friendster placed on how users could express themselves, and they also put a bigger emphasis on music and allowing bands to promote themselves online. Chau developed the initial version of the MySpace Web site in Perl, running on the Apache Web server, with a MySQL database back end. That didn't make it past the test phase, however, because other Intermix developers had more experience with ColdFusion, the Web application environment originally developed by Allaire and now owned by Adobe. So, the production Web site went live on ColdFusion, running on Windows, and Microsoft SQL Server as the database.

Chau left the company about then, leaving further Web development to others, including Aber Whitcomb, an Intermix technologist who is now MySpace's chief technology officer, and Benedetto, who joined about a month after MySpace went live.

MySpace was launched in 2003, just as Friendster started having trouble keeping pace with its own runaway growth. In a recent interview with Fortune magazine, Friendster president Kent Lindstrom admitted his service stumbled at just the wrong time, taking 20 to 30 seconds to deliver a page when MySpace was doing it in 2 or 3 seconds.

As a result, Friendster users began to defect to MySpace, which they saw as more dependable.

Today, MySpace is the clear "social networking" king. Social networking refers to Web sites organized to help users stay connected with each other and meet new people, either through introductions or searches based on common interests or school affiliations. Other prominent sites in this category include Facebook, which originally targeted university students; and LinkedIn, a professional networking site, as well as Friendster. MySpace prefers to call itself a "next generation portal," emphasizing a breadth of content that includes music, comedy and videos. It operates like a virtual nightclub, with a juice bar for under-age visitors off to the side, a meat-market dating scene front and center, and marketers in search of the youth sector increasingly crashing the party.

Users register by providing basic information about themselves, typically including age and hometown, their sexual preference and their marital status. Some of these options are disabled for minors, although MySpace continues to struggle with a reputation as a stomping ground for sexual predators.

MySpace profile pages offer many avenues for self-expression, ranging from the text in the About Me section of the page to the song choices loaded into the MySpace music player, video choices, and the ranking assigned to favorite friends. MySpace also gained fame for allowing users a great deal of freedom to customize their pages with Cascading Style Sheets (CSS), a Web standard formatting language that makes it possible to change the fonts, colors and background images associated with any element of the page. The results can be hideous—pages so wild and discolored that they are impossible to read or navigate—or they can be stunning, sometimes employing professionally designed templates (see "Too Much of a Good Thing?" p. 48).

The "network effect," in which the mass of users inviting other users to join MySpace led to exponential growth, began about eight months after the launch "and never really stopped," Chau says.

News Corp., the media empire that includes the Fox television networks and 20th Century Fox movie studio, saw this rapid growth as a way to multiply its share of the audience of Internet users, and bought MySpace in 2005 for $580 million. Now, News Corp. chairman Rupert Murdoch apparently thinks MySpace should be valued like a major Web portal, recently telling a group of investors he could get $6 billion—more than 10 times the price he paid in 2005—if he turned around and sold it today. That's a bold claim, considering the Web site's total revenue was an estimated $200 million in the fiscal year ended June 2006. News Corp. says it expects Fox Interactive as a whole to have revenue of $500 million in 2007, with about $400 million coming from MySpace.

But MySpace continues to grow. In December, it had 140 million member accounts, compared with 40 million in November 2005. Granted, that doesn't quite equate to the number of individual users, since one person can have multiple accounts, and a profile can also represent a band, a fictional character like Borat, or a brand icon like the Burger King.

Still, MySpace has tens of millions of people posting messages and comments or tweaking their profiles on a regular basis—some of them visiting repeatedly throughout the day. That makes the technical requirements for supporting MySpace much different than, say, for a news Web site, where most content is created by a relatively small team of editors and passively consumed by Web site visitors. In that case, the content management database can be optimized for read-only requests, since additions and updates to the database content are relatively rare. A news site might allow reader comments, but on MySpace user-contributed content is the primary content. As a result, it has a higher percentage of database interactions that are recording or updating information rather than just retrieving it.

Every profile page view on MySpace has to be created dynamically—that is, stitched together from database lookups. In fact, because each profile page includes links to those of the user's friends, the Web site software has to pull together information from multiple tables in multiple databases on multiple servers. The database workload can be mitigated somewhat by caching data in memory, but this scheme has to account for constant changes to the underlying data.

The Web site architecture went through five major revisions—each coming after MySpace had reached certain user account milestones—and dozens of smaller tweaks, Benedetto says. "We didn't just come up with it; we redesigned, and redesigned, and redesigned until we got where we are today," he points out.

Although MySpace declined formal interview requests, Benedetto answered Baseline's questions during an appearance in November at the SQL Server Connections conference in Las Vegas. Some of the technical information in this story also came from a similar "mega-sites" presentation that Benedetto and his boss, chief technology officer Whitcomb, gave at Microsoft's MIX Web developer conference in March.

As they tell it, many of the big Web architecture changes at MySpace occurred in 2004 and early 2005, as the number of member accounts skyrocketed into the hundreds of thousands and then millions.

At each milestone, the Web site would exceed the maximum capacity of some component of the underlying system, often at the database or storage level. Then, features would break, and users would scream. Each time, the technology team would have to revise its strategy for supporting the Web site's workload.

And although the systems architecture has been relatively stable since the Web site crossed the

7 million account mark in early 2005, MySpace continues to knock up against limits such as the number of simultaneous connections supported by SQL Server, Benedetto says: "We've maxed out pretty much everything."

First Milestone: 500,000 Accounts

MySpace started small, with two Web servers talking to a single database server. Originally, they were 2-processor Dell servers loaded with 4 gigabytes of memory, according to Benedetto.

Web sites are better off with such a simple architecture—if they can get away with it, Benedetto says. "If you can do this, I highly recommend it because it's very, very non-complex," he says. "It works great for small to medium-size Web sites."

The single database meant that everything was in one place, and the dual Web servers shared the workload of responding to user requests. But like several subsequent revisions to MySpace's underlying systems, that three-server arrangement eventually buckled under the weight of new users. For a while, MySpace absorbed user growth by throwing hardware at the problem—simply buying more Web servers to handle the expanding volume of user requests.

But at 500,000 accounts, which MySpace reached in early 2004, the workload became too much for a single database.

Adding databases isn't as simple as adding Web servers. When a single Web site is supported by multiple databases, its designers must decide how to subdivide the database workload while maintaining the same consistency as if all the data were stored in one place.

In the second-generation architecture, MySpace ran on three SQL Server databases—one designated as the master copy to which all new data would be posted and then replicated to the other two, which would concentrate on retrieving data to be displayed on blog and profile pages. This also worked well—for a while—with the addition of more database servers and bigger hard disks to keep up with the continued growth in member accounts and the volume of data being posted.

Second Milestone: 1-2 Million Accounts

As MySpace registration passed 1 million accounts and was closing in on 2 million, the service began knocking up against the input/output (I/O) capacity of the database servers—the speed at which they were capable of reading and writing data. This was still just a few months into the life of the service, in mid-2004. As MySpace user postings backed up, like a thousand groupies trying to squeeze into a nightclub with room for only a few hundred, the Web site began suffering from "major inconsistencies," Benedetto says, meaning that parts of the Web site were forever slightly out of date.

"A comment that someone had posted wouldn't show up for 5 minutes, so users were always complaining that the site was broken," he adds.

The next database architecture was built around the concept of vertical partitioning, with separate databases for parts of the Web site that served different functions such as the log-in screen, user profiles and blogs. Again, the Web site's scalability problems seemed to have been solved—for a while.

The vertical partitioning scheme helped divide up the workload for database reads and writes alike, and when users demanded a new feature, MySpace would put a new database online to support it. At 2 million accounts, MySpace also switched from using storage devices directly attached to its database servers to a storage area network (SAN), in which a pool of disk storage devices are tied together by a high-speed, specialized network, and the databases connect to the SAN. The change to a SAN boosted performance, uptime and reliability, Benedetto says.

Third Milestone: 3 Million Accounts

As the Web site's growth continued, hitting 3 million registered users, the vertical partitioning solution couldn't last. Even though the individual applications on sub-sections of the Web site were for the most part independent, there was also information they all had to share. In this architecture, every database had to have its own copy of the users table—the electronic roster of authorized MySpace users. That meant when a new user registered, a record for that account had to be created on nine different database servers. Occasionally, one of those transactions would fail, perhaps because one particular database server was momentarily unavailable, leaving the user with a partially created account where everything but, for example, the blog feature would work for that person.

And there was another problem. Eventually, individual applications like blogs on sub-sections of the Web site would grow too large for a single database server.

By mid-2004, MySpace had arrived at the point where it had to make what Web developers call

the "scale up" versus "scale out" decision—whether to scale up to bigger, more powerful and more expensive servers, or spread out the database workload across lots of relatively cheap servers. In general, large Web sites tend to adopt a scale-out approach that allows them to keep adding capacity by adding more servers.

But a successful scale-out architecture requires solving complicated distributed computing problems, and large Web site operators such as Google, Yahoo and Amazon.com have had to invent a lot of their own technology to make it work. For example, Google created its own distributed file system to handle distributed storage of the data it gathers and analyzes to index the Web.

In addition, a scale-out strategy would require an extensive rewrite of the Web site software to make programs designed to run on a single server run across many—which, if it failed, could easily cost the developers their jobs, Benedetto says.

So, MySpace gave serious consideration to a scale-up strategy, spending a month and a half studying the option of upgrading to 32-processor servers that would be able to manage much larger databases, according to Benedetto. "At the time, this looked like it could be the panacea for all our problems," he says, wiping away scalability issues for what appeared then to be the long term. Best of all, it would require little or no change to the Web site software.

Unfortunately, that high-end server hardware was just too expensive—many times the cost of buying the same processor power and memory spread across multiple servers, Benedetto says. Besides, the Web site's architects foresaw that even a super-sized database could ultimately be overloaded, he says: "In other words, if growth continued, we were going to have to scale out anyway."

So, MySpace moved to a distributed computing architecture in which many physically separate computer servers were made to function like one logical computer. At the database level, this meant reversing the decision to segment the Web site into multiple applications supported by separate databases, and instead treat the whole Web site as one application. Now there would only be one user table in that database schema because the data to support blogs, profiles and other core features would be stored together.

Now that all the core data was logically organized into one database, MySpace had to find another way to divide up the workload, which was still too much to be managed by a single database server running on commodity hardware. This time, instead of creating separate databases for Web site functions or applications, MySpace began splitting its user base into chunks of 1 million accounts and putting all the data keyed to those accounts in a separate instance of SQL Server. Today, MySpace actually runs two copies of SQL Server on each server computer, for a total of 2 million accounts per machine, but Benedetto notes that doing so leaves him the option of cutting the workload in half at any time with minimal disruption to the Web site architecture.

There is still a single database that contains the user name and password credentials for all users. As members log in, the Web site directs them to the database server containing the rest of the data for their account. But even though it must support a massive user table, the load on the log-in database is more manageable because it is dedicated to that function alone.

Next page: Fourth Milestone: 9 Million—17 Million Accounts

Fourth Milestone: 9 Million—17 Million Accounts

When MySpace reached 9 million accounts, in early 2005, it began deploying new Web software written in Microsoft's C# programming language and running under ASP.NET. C# is the latest in a long line of derivatives of the C programming language, including C++ and Java, and was created to dovetail with the Microsoft .NET Framework, Microsoft's model architecture for software components and distributed computing. ASP.NET, which evolved from the earlier Active Server Pages technology for Web site scripting, is Microsoft's current Web site programming environment.

Almost immediately, MySpace saw that the ASP.NET programs ran much more efficiently, consuming a smaller share of the processor power on each server to perform the same tasks as a comparable ColdFusion program. According to CTO Whitcomb, 150 servers running the new code were able to do the same work that had previously required 246. Benedetto says another reason for the performance improvement may have been that in the process of changing software platforms and rewriting code in a new language, Web site programmers reexamined every function for ways it could be streamlined.

Eventually, MySpace began a wholesale migration to ASP.NET. The remaining ColdFusion code was adapted to run on ASP.NET rather than on a Cold-Fusion server, using BlueDragon.NET, a product from New Atlanta Communications of Alpharetta, Ga., that automatically recompiles ColdFusion code for the Microsoft environment.

When MySpace hit 10 million accounts, it began to see storage bottlenecks again. Implementing a SAN had solved some early performance problems, but now the Web site's demands were starting to periodically overwhelm the SAN's I/O capacity—the speed with which it could read and write data to and from disk storage.

Part of the problem was that the 1 million-accounts-per-database division of labor only smoothed out the workload when it was spread relatively evenly across all the databases on all the servers. That was usually the case, but not always. For example, the seventh 1

million-account database MySpace brought online wound up being filled in just seven days, largely because of the efforts of one Florida band that was particularly aggressive in urging fans to sign up.

Whenever a particular database was hit with a disproportionate load, for whatever reason, the cluster of disk storage devices in the SAN dedicated to that database would be overloaded. "We would have disks that could handle significantly more I/O, only they were attached to the wrong database," Benedetto says.

At first, MySpace addressed this issue by continually redistributing data across the SAN to reduce these imbalances, but it was a manual process "that became a full-time job for about two people," Benedetto says.

The longer-term solution was to move to a virtualized storage architecture where the entire SAN is treated as one big pool of storage capacity, without requiring that specific disks be dedicated to serving specific applications. MySpace now standardized on equipment from a relatively new SAN vendor, 3PARdata of Fremont, Calif., that offered a different approach to SAN architecture.

In a 3PAR system, storage can still be logically partitioned into volumes of a given capacity, but rather than being assigned to a specific disk or disk cluster, volumes can be spread or "striped" across thousands of disks. This makes it possible to spread out the workload of reading and writing data more evenly. So, when a database needs to write a chunk of data, it will be recorded to whichever disks are available to do the work at that moment rather than being locked to a disk array that might be overloaded. And since multiple copies are recorded to different disks, data can also be retrieved without overloading any one component of the SAN.

To further lighten the burden on its storage systems when it reached 17 million accounts, in the spring of 2005 MySpace added a caching tier—a layer of servers placed between the Web servers and the database servers whose sole job was to capture copies of frequently accessed data objects in memory and serve them to the Web application without the need for a database lookup. In other words, instead of querying the database 100 times when displaying a particular profile page to 100 Web site visitors, the site could query the database once and fulfill each subsequent request for that page from the cached data. Whenever a page changes, the cached data is erased from memory and a new database lookup must be performed—but until then, the database is spared that work, and the Web site performs better.

The cache is also a better place to store transitory data that doesn't need to be recorded in a database, such as temporary files created to track a particular user's session on the Web site—a lesson that Benedetto admits he had to learn the hard way. "I'm a database and storage guy, so my answer tended to be, let's put everything in the database," he says, but putting inappropriate items such as session tracking data in the database only bogged down the Web site.

The addition of the cache servers is "something we should have done from the beginning, but we were growing too fast and didn't have time to sit down and do it," Benedetto adds.

Fifth Milestone: 26 Million Accounts

In mid-2005, when the service reached 26 million accounts, MySpace switched to SQL Server 2005 while the new edition of Microsoft's database software was still in beta testing. Why the hurry? The main reason was this was the first release of SQL Server to fully exploit the newer 64-bit processors, which among other things significantly expand the amount of memory that can be accessed at one time. "It wasn't the features, although the features are great," Benedetto says. "It was that we were so bottlenecked by memory."

More memory translates into faster performance and higher capacity, which MySpace sorely needed. But as long as it was running a 32-bit version of SQL Server, each server could only take advantage of about 4 gigabytes of memory at a time. In the plumbing of a computer system, the difference between 64 bits and 32 bits is like widening the diameter of the pipe that allows information to flow in and out of memory. The effect is an exponential increase in memory access. With the upgrade to SQL Server 2005 and the 64-bit version of Windows Server 2003, MySpace could exploit 32 gigabytes of memory per server, and in 2006 it doubled its standard configuration to 64 gigabytes.

Next page: Unexpected Errors

Unexpected Errors

If it were not for this series of upgrades and changes to systems architecture, the MySpace Web site wouldn't function at all. But what about the times when it still hiccups? What's behind those "Unexpected Error" screens that are the source of so many user complaints?

One problem is that MySpace is pushing Microsoft's Web technologies into territory that only Microsoft itself has begun to explore, Benedetto says. As of November, MySpace was exceeding the number of simultaneous connections supported by SQL Server, causing the software to crash. The specific circumstances that trigger one of these crashes occur only about once every three days, but it's still frequent enough to be annoying, according to Benedetto. And anytime a database craps out, that's bad news if the data for the page you're trying to view is stored there. "Anytime that happens, and uncached data is unavailable through SQL Server, you'll see

one of those unexpected errors," he explains.

Last summer, MySpace's Windows 2003 servers shut down unexpectedly on multiple occasions. The culprit turned out to be a built-in feature of the operating system designed to prevent distributed denial of service attacks—a hacker tactic in which a Web site is subjected to so many connection requests from so many client computers that it crashes. MySpace is subject to those attacks just like many other top Web sites, but it defends against them at the network level rather than relying on this feature of Windows—which in this case was being triggered by hordes of legitimate connections from MySpace users.

"We were scratching our heads for about a month trying to figure out why our Windows 2003 servers kept shutting themselves off," Benedetto says. Finally, with help from Microsoft, his team figured out how to tell the server to "ignore distributed denial of service; this is friendly fire."

And then there was that Sunday night last July when a power outage in Los Angeles, where MySpace is headquartered, knocked the entire service offline for about 12 hours. The outage stood out partly because most other large Web sites use geographically distributed data centers to protect themselves against localized service disruptions. In fact, MySpace had two other data centers in operation at the time of this incident, but the Web servers housed there were still dependent on the SAN infrastructure in Los Angeles. Without that, they couldn't serve up anything more than a plea for patience.

According to Benedetto, the main data center was designed to guarantee reliable service through connections to two different power grids, backed up by battery power and a generator with a 30-day supply of fuel. But in this case, both power grids failed, and in the process of switching to backup power, operators blew the main power circuit.

MySpace is now working to replicate the SAN to two other backup sites by mid-2007. That will also help divvy up the Web site's workload, because in the normal course of business, each SAN location will be able to support one-third of the storage needs. But in an emergency, any one of the three sites would be able to sustain the Web site independently, Benedetto says.

While MySpace still battles scalability problems, many users give it enough credit for what it does right that they are willing to forgive the occasional error page.

"As a developer, I hate bugs, so sure it's irritating," says Dan Tanner, a 31-year-old software developer from Round Rock, Texas, who has used MySpace to reconnect with high school and college friends. "The thing is, it provides so much of a benefit to people that the errors and glitches we find are forgivable." If the site is down or malfunctioning one day, he simply comes back the next and picks up where he left off, Tanner says.

That attitude is why most of the user forum responses to Drew's rant were telling him to calm down and that the problem would probably fix itself if he waited a few minutes. Not to be appeased, Drew wrote, "ive already emailed myspace twice, and its BS cause an hour ago it was working, now its not … its complete BS." To which another user replied, "and it's free."

Benedetto candidly admits that 100% reliability is not necessarily his top priority. "That's one of the benefits of not being a bank, of being a free service," he says.

In other words, on MySpace the occasional glitch might mean the Web site loses track of someone's latest profile update, but it doesn't mean the site has lost track of that person's money. "That's one of the keys to the Web site's performance, knowing that we can accept some loss of data," Benedetto says. So, MySpace has configured SQL Server to extend the time between the "checkpoints" operations it uses to permanently record updates to disk storage—even at the risk of losing anywhere between 2 minutes and 2 hours of data—because this tweak makes the database run faster.

Similarly, Benedetto's developers still often go through the whole process of idea, coding, testing and deployment in a matter of hours, he says. That raises the risk of introducing software bugs, but it allows them to introduce new features quickly. And because it's virtually impossible to do realistic load testing on this scale, the testing that they do perform is typically targeted at a subset of live users on the Web site who become unwitting guinea pigs for a new feature or tweak to the software, he explains.

"We made a lot of mistakes," Benedetto says. "But in the end, I think we ended up doing more right than we did wrong."

Next page: MySpace Base Case

MySpace Base Case
Headquarters: Fox Interactive Media (parent company), 407 N. Maple Drive, Beverly Hills, CA 90210
Phone: (310) 969-7200
Business: MySpace is a "next generation portal" built around a social networking Web site that allows members to meet, and stay connected with, other members, as well as their favorite bands and celebrities.
Chief Technology Officer: Aber Whitcomb
Financials in 2006: Estimated revenue of $200 million.

BASELINE GOALS:

- Double MySpace.com advertising rates, which in 2006 were typically a little more than 10 cents per

1,000 impressions.

- Generate revenue of at least $400 million from MySpace—out of $500 million expected from News Corp.'s Fox Interactive Media unit—in this fiscal year.
- Secure revenue of $900 million over the next three years from a search advertising deal with Google.

User Customization: Too Much of a Good Thing?
January 16, 2007

By David F. Carr

One of the features members love about MySpace is that it gives people who open up an account a great deal of freedom to customize their pages with Cascading Style Sheets (CSS), a Web format that allows users to change the fonts, colors and background images associated with any element of the page.

That feature was really "kind of a mistake," says Duc Chau, one of the social networking site's original developers. In other words, he neglected to write a routine that would strip Web coding tags from user postings—a standard feature on most Web sites that allow user contributions.

The Web site's managers belatedly debated whether to continue allowing users to post code "because it was making the page load slow, making some pages look ugly, and exposing security holes," recalls Jason Feffer, former MySpace vice president of operations. "Ultimately we said, users come first, and this is what they want. We decided to allow the users to do what they wanted to do, and we would deal with the headaches."

In addition to CSS, JavaScript, a type of programming code that runs in the user's browser, was originally allowed. But MySpace eventually decided to filter it out because it was exploited to hack the accounts of members who visited a particular profile page. MySpace, however, still experiences periodic security problems, such as the infected QuickTime video that turned up in December, automatically replicating itself from profile page to profile page. QuickTime's creator, Apple Computer, responded with a software patch for MySpace to distribute. Similar problems have cropped up in the past with other Web software, such as the Flash viewer.

Planner: Calculating the Costs of a Web Site Makeover
January 18, 2007

By Sean Nolan

At your consumer products company, "Web 2.0" is officially in danger of becoming this year's "thinking outside of the box": buzzword fodder that's big on elocution, but short on execution.

That's why this six-month project to enliven and build out your consumer Web site will place so much emphasis on defining exactly what your company wants to accomplish with "Web 2.0"—a catchall that means different things to different businesses. For your company, it will mean adding more dimension to a flat consumer Web site and, most important, developing more direct connections to a customer base that gets harder to reach each day.

View the PDF -- Turn off pop-up blockers!

Part of your motivation here will be pure survival. Old-world media and marketing approaches are increasingly less effective in the face of rapidly fragmenting customer niches. But far from a business problem, those niches represent "an opportunity for customer engagement," says Patricia Seybold, founder and CEO of the Patricia Seybold Group and author of Outside Innovation: How Your Customers Will Co-Design Your Company's Future. "Instead of combating fragmentation, companies should be leveraging online tools and online communities to leverage customer fragmentation and to address more customers' needs, not fewer."

Delivering that online promise will mean actively and constructively involving your customers in your business processes through those online tools—blogs, surveys, contests, forums, ratings, and other one-to-one communication exchanges that will do more than just create a multilayered online presence. They'll also use a data analytics backbone to channel that customer interaction into tangible feedback on everything from smarter product development to more effective advertising to "viral" (buzzwords never die) word-of-mouth sales.

To see the details behind this Planner and fill in your own estimates, click on the "Get the Tool" icon above and download the interactive worksheet.

Web Design Experts Grade MySpace
January 15, 2007

By David F. Carr

MySpace.com's continued growth flies in the face of much of what Web experts have told us for years about how to succeed on the Internet. It's buggy, often responding to basic user requests with the dreaded "Unexpected Error" screen, and stocked with thousands of pages that violate all sorts of conventional Web design standards with their wild colors and confusing background images. And yet, it succeeds anyway.

Why?

"The hurdle is a little bit lower for something like this specifically because it's not a mission-critical site," says Jakob Nielsen, the famed Web usability expert and principal of the Nielsen Norman Group, which has its headquarters in Fremont, Calif. "If someone were trying to launch an eBay competitor and it had problems like this, it would never get off the ground." For that reason, he finds it difficult to judge MySpace by the same standards as more utilitarian Web sites, such as a shopping site where usability flaws might lead to abandoned shopping carts.

On most Web sites designed to sell or inform, the rampant self-expression Nielsen sees on MySpace would be a fatal flaw. "Usually, people don't go to a Web site to see how to express yourself," he says. "But people do go to MySpace to see how you express yourself, to see what bands you like, all that kind of stuff."

The reliability of the service also winds up being judged by different standards, according to Nielsen. If a Web user finds an e-commerce site is down, switching to a competitor's Web site is an easy decision. "But in this case, because your friends are here, you're more likely to want to come back to this site rather than go to another site," Nielsen says. "Most other Web sites could not afford that."

From a different angle, Newsvine CEO Mike Davidson says one of the things MySpace has done a great job of is allowing millions of members to sort themselves into smaller communities of a more manageable size, based on school and interest group associations. Davidson has studied MySpace to glean ideas for social networking features he is adding to his own Web site for news junkies. As a Web developer and former manager of media product development for the ESPN.com sports news Web site, he admires the way MySpace has built a loyal community of members.

"One of the things MySpace has been really great about is turning crap into treasure," Davidson says. "You look at these profile pages, and most of the comments are stuff like, 'Love your hair,' so to an outsider, it's kind of stupid. But to that person, that's their life." The "treasure" MySpace extracts from this experience is the billions of page views recorded as users click from profile to profile, socializing and gossiping online.

On the other hand, parts of the MySpace Web application are so inefficient, requiring multiple clicks and page views to perform simple tasks, that a good redesign would probably eliminate two-thirds of those page views, Davidson says. Even if that hurt MySpace's bragging rights as one of the most visited Web sites, it would ultimately lead to more satisfied users and improve ad rates by making each page view count for more, he argues.

"In a lot of ways, he's very right," says Jason Feffer, a former MySpace vice president of operations. While denying that the Web site was intentionally designed to inflate the number of page views, he says it's true that MySpace winds up with such a high inventory of page views that there is never enough advertising to sell against it. "On the other hand, when you look at the result, it's hard to argue that what we did with the interface and navigation was bad," he maintains. "And why change it, when you have success?"

Feffer, who is currently working on his own startup of an undisclosed nature called SodaHead.com, says one of the biggest reasons MySpace succeeded was that its users were always willing to cut it some slack.

"Creating a culture where users are sympathetic is very important," Feffer says. Especially in the beginning, many users thought the Web site was "something Tom was running out of his garage," he says, referring to MySpace president and co-founder Tom Anderson, who is the public face of the service by virtue of being the first online "friend" who welcomes every new MySpace user.

That startup aura made users more tolerant of occasional bugs and outages, according to Feffer. "They would think that it was cool that during an outage, you're putting up Pac-Man for me to play with," he says. "If you're pretending to be Yahoo or Google, you're not going to get much sympathy."

MySpace is starting to be held to a higher standard, however, since being purchased by News Corp. in 2005, and the reaction was different following a 12-hour outage this past summer, Feffer says: "I don't think anyone believed it was Tom's little garage project anymore."

MySpace Insiders
January 16, 2007

By [David F. Carr](#)
Rupert Murdoch
Chairman, News Corp.
As the creator of a media empire that includes 20th Century Fox, the Fox television stations, the New York Post and many other news, broadcast and music properties, Murdoch championed the purchase of MySpace.com as a way of significantly expanding Fox Interactive Media's presence on the Web.

Chris Dewolfe
CEO, MySpace
DeWolfe, who is also a co-founder of MySpace.com, led its creation while employed by Intermix Media and continues to manage it today as a unit of News Corp.'s Fox Interactive Media. Previously, he was CEO of the e-mail marketing firm ResponseBase, which Intermix bought in 2002.

Tom Anderson
President, MySpace

A co-founder of MySpace, Anderson is best known as "Tom," the first person who appears on the "friends list" of new MySpace.com members and who acts as the public face of the Web site's support organization. He and DeWolfe met at Xdrive, the Web file storage company where both worked prior to starting ResponseBase.

Aber Whitcomb
Chief Technology Officer, MySpace
Whitcomb is a co-founder of MySpace, where he is responsible for engineering and technical operations. He speaks frequently on the issues of large-scale computing, networking and storage.

Jim Benedetto
Vice President of Technology, MySpace
Benedetto joined MySpace about a month after it launched, in late 2003. On his own MySpace profile page, he describes himself as a 27-year-old 2001 graduate of the University of Southern California whose trip to Australia last year included diving in a shark tank. Just out of school in 2001, he joined Quack.com, a voice portal startup that was acquired by America Online. Today, Benedetto says he is "working triple overtime to take MySpace international."

Jason Feffer
Former vice president of operations, MySpace
Starting with MySpace's launch in late 2003, Feffer was responsible for MySpace's advertising and support operations. He also worked with DoubleClick, the Web site advertising vendor, to ensure that its software met MySpace's scalability requirements and visitor targeting goals. Since leaving MySpace last summer, he has been working on a startup called SodaHead.com, which promises to offer a new twist on social networking when it launches later this year.

Duc Chau
Founder and CEO, Flukiest
Chau, as an employee of Intermix, led the creation of a pilot version of the MySpace Web site, which employed Perl and a MySQL database, but left Intermix shortly after the production Web site went live. He went on to work for StrongMail, a vendor of e-mail management appliances. Chau now runs Flukiest, a social networking and file-sharing Web site that is also selling its software for use within other Web sites.

MySpace Tech Roster
January 16, 2007

By David F. Carr
MySpace has managed to scale its Web site infrastructure to meet booming demand by using a mix of time-proven and leading-edge information technologies.

APPLICATION	PRODUCT	SUPPLIER
Web application technology	Microsoft Internet Information Services, .NET Framework	Microsoft
Server operating system	Windows 2003	Microsoft
Programming language and environment	Applications written in C# for ASP.NET	Microsoft
Programming language and environment	Site originally launched on Adobe's ColdFusion; remaining ColdFusion code runs under New Atlanta's BlueDragon.NET product.	Adobe, New Atlanta
Database	SQL Server 2005	Microsoft
Storage area network	3PAR Utility Storage	3PARdata
Internet application acceleration	NetScaler	Citrix Systems
Server hardware	Standardized on HP 585 (see below)	Hewlett-Packard
Ad server software	DART Enterprise	DoubleClick
Search and keyword advertising	Google search	Google

Standard database server configuration consists of Hewlett-Packard HP 585 servers with 4 AMD Opteron dual-core, 64-bit processors with 64 gigabytes of memory (recently upgraded from 32). The operating system is Windows 2003, Service Pack 1; the database software is Microsoft SQL Server 2005, Service Pack 1. There's a 10-gigabit-per-second Ethernet network card, plus two host bus adapters for storage area network communications. The infrastructure for the core user profiles application includes 65 of these database servers with a total capacity of more than 2 terabytes of memory, 520 processors and 130 gigabytes of network throughput. *Source: MySpace.com user conference presentations*