

TerraServer and the Russia Adventure...

Tom Barclay

Microsoft

835 Market Street, Suite 700

San Francisco, CA 94103

(01) 415-972-6451

tbarclay@microsoft.com

Abstract

Jim Gray made many friends and few enemies in his career. Many of the projects he worked on were inspired by one or more of his friends and collaborators in industry and academia.

The original idea for an “image server” happened in the early 1990s when Jim was at Digital. Jim had met Jeff Dozier while working on the National Research Council’s “Computing the Future” committee [NRC 1994]. Jeff became Dean of the Bren School of Environmental Sciences and chair of UC Santa Barbara’s Alexandria Digital Library advisory board. Jeff asked Jim to serve on the board.

During an Alexandria board meeting, Jeff and Jim discussed the idea for a “geospatial image server”. All they needed was a lot of accessible imagery, someone that wanted it done, and someone foolish enough to load a few tera-bytes of data. The cold war ended enabling the high resolution imagery marketplace. Microsoft SQL Server wanted a scalability internet demonstration. And I walked into Jim Gray’s lab looking for a reason to move back to San Francisco. And that’s where my 10 year project with Jim Gray and TerraServer begins...

1. Why TerraServer?

Jim joined in Microsoft Research in the fall of 1995. The lab was small and focused on the interests of its founders Jim Gray and Gordon Bell – Scalable Servers and Telepresence. We slowly added staff – Joe Barrera, Don Slutz, and Jim Gemmell – and began to do small projects and prototypes. In late spring of 1996, Paul Flessner, the General Manager of the SQL Server team asked our lab to build a database application that would test and demonstrate the scalability of the next release of SQL Server code named “Sphinx”. Paul was beginning to grow the SQL Server team from several dozen engineers to several hundred and turning SQL Server into a major player in the database market.

At first Jim was reticent to take on a scalability project – where is the research in this? But Jim was keenly aware of how difficult it is to transfer ideas from research to product. He frequently told stories of how lucky they were to have the ideas in System R make to products that customers could buy. A frequent quote of Jim’s was “*ideas don’t transfer, people transfer...*” So we took on the project as a “Scale-Up” research prototype because the SQL engineering team asked us. We expected it would pay dividends later when we had other prototypes and ideas that we thought they should consider or inclusion into SQL Server.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Jim Gray Tribute, May 31, 2008, Berkeley, CA, USA.

2. Project Definition

One of Jim’s greatest abilities was to clearly define and articulate the problem. The SQL team gave us two goals:

1. Test SQL’s ability to scale up to support a database of one tera-byte or larger.
2. An internet application where SQL marketing could demonstrate Windows and SQL Server’s scalability.

In the early weeks of the project, we decided TerraServer had the following requirements to meet both goals:

- **BIG** —1 TB of data including catalog, temporary space, etc.
- **PUBLIC** — available on the world wide web
- **INTERESTING** — to a wide audience
- **ACCESSIBLE** — using standard browsers (IE, Netscape)
- **REAL** — a LOB application (users can buy imagery)
- **FREE** —cannot require NDA or money to a user to access
- **FAST** — usable on low-speed (56kbps) and high speeds(T-1+)
- **EASY** — we do not want a large group to develop, deploy, or maintain the application

An unwritten requirement was “cheap” for two reasons—(1) because TerraServer was only a prototype, test, and free demonstration; and (2) Jim Gray was a very frugal person!

3. The Search for an Interesting *and* Cheap Tera-Byte

It is a challenge to find both an interesting tera-byte *and* a cheap tera-byte. First we considered doing a genealogy database. We scoured the internet and talked to a number of experts. We decided that casual users wouldn’t know enough details about their own or their parents’ birth records in order to begin the search. We also wouldn’t necessarily have enough data for all users to be able to use the site for very long. So our search for cheap *and* interesting moved on.

It was time for Jim to call Jeff Dozier and see if there was enough high resolution imagery available to create an interesting application for consumers. Our idea was to generate small JPEG images, i.e. “tiles”, out of large, high resolution imagery captured by satellite or aircraft. We visited the Alexandria Digital Library team [ADL 1999] in the summer of 1996 in search for an interesting *and* cheap tera-byte of imagery...

The UC Santa Barbara project team was working on the Alexandria Digital Library. They had access to a number of high resolution imagery assets. Two sources were particularly intriguing – the US Geological Survey (USGS) Digital Orthophotography Quarter Quadrangle (DOQQ) program, and satellite imagery of Spot Image. There were only two problems –

1. The USGS DOQQ program was just getting underway and would only produce imagery for the conterminous United States. Over 50% of Microsoft's customers are outside of the U.S. and we wanted to reach them as well.
2. Spot Image is a French satellite firm with a large library of imagery from around the world. Unfortunately, their library was expensive and the resolution too low for consumer's to recognize their home or neighborhood.

Microsoft management was willing to fund the project if we found an international imagery supplier that could supply data with the same quality as the USGS DOQQ program.

In 1996, the search engine of choice was DEC's Alta Vista. A search for "satellite" found 100k+ results and led us eventually to Aerial Images, a small firm in North Carolina. Aerial Images had just signed a contract with Sovinformspnutnik to be the exclusive U.S. reseller of de-classified Russian military imagery. Through Aerial Images, Sovinformspnutnik was distributing 1.5 meter per pixel grayscale imagery. In August 1996, international regulations on the distribution of high resolution satellite imagery had been relaxed from 5 meters per pixel (as distributed by Spot Image) to 1 meter per pixel. Sovinformspnutnik had an instant corner on the market because they were the only commercial organization that had a substantial inventory of high resolution satellite imagery. Aerial Images was looking for new markets and channels for the Sovinformspnutnik imagery library when we sent them an e-mail...

4. The Trip to Moscow

By winter of 1996, we had come to an agreement to host the Sovinformspnutnik imagery on the internet, and build Aerial Images a consumer focused e-commerce site for acquiring digital and photographic prints of Sovinformspnutnik imagery. There was only one catch – Sovinformspnutnik management wanted to meet Jim and me in Moscow before they agreed to deliver the imagery.

Off we went to Moscow in February 1997, two Californians in our brand new boots and coats along with representatives from Aerial Images, and Digital Equipment, our hardware partner on TerraServer, for a week-long series of meetings. In exchange for an eighteen month, royalty free access to the entire Sovinformspnutnik library, we had to:

1. Convince Sovinformspnutnik management that we could protect their data from theft.
2. Be able to turn off access in seconds if required by international military authorities.
3. Build an e-commerce application for Aerial Images to sell and distribute digital images and large-format prints.
4. Participate in a televised press announcement by the Russian Space Agency announcing the Sovinformspnutnik-Aerial Images-Microsoft deal...

During our first day, we were greeted at the Sovinformspnutnik office building by a soldier holding an AK47. After having our papers examined, we were escorted down a dimly hallway to a freshly painted red door where we were greeted by Mr. Fomtchenko, the General Manager of Sovinformspnutnik. We gave a detailed presentation over a span of two days. Jim was on the panel during the press conference at the Russian Space Agency. That night, Mr. Fomtchenko hosted a dinner at the Danilov Monastery built in the 13th century. We dined to a nine-

course meal and participated in 27 vodka toasts... We didn't sober up until we arrived back in the US two days later.

One of our greatest joys was hosting the TerraServer pre-launch party



Televised press announcement of TerraServer.

after

"Scalability Day", the U.S. announcement of TerraServer in New York. Attending were executives from the USGS, Sovinformspnutnik, Aerial Images, Digital Equipment other TerraServer partners. We watched our USGS partners stand and toast their Sovinformspnutnik partners; and our Sovinformspnutnik partners stand and return a toast to their USGS partners.

5. TerraServer's Legacy

TerraServer launched on June 22, 1998 running on a single DEC AlphaServer 8400 with a beta copy of SQL 7.0. The site and SQL Server was immediately overwhelmed with visitors. In the first two weeks we found and corrected over a dozen major performance bugs in the SQL beta version.

The USGS presented Jim and me with the "John Wesley Powell Award" for Industry Achievement given to those who have made a significant contribution to the mission of the USGS.

In 2005, Jim loaned me to the Virtual Earth team for two years. TerraServer imagery was extracted from the SQL databases and became the base imagery for Microsoft Virtual Earth – proving Jim right again – "*ideas transfer with people*".

The site was originally projected to be live for 18 months. This June, TerraServer will celebrate its 10th consecutive year of service. Two of us continue to operate and maintain TerraServer. Its usage has varied over the years, but today continues to enjoy 60k unique visitors per day even with sites like Virtual Earth, Google Earth, Yahoo Maps, and others.

6. Acknowledgement

Space does not permit me to fairly describe what Jim meant to me professionally and personally. Jim took a chance on me when we first met in 1991 giving me the opportunity to change careers from a field software engineer and to an entirely different area of research. In return, I taught him a little bit about Russian toasts and vodka shots...

7. References

[ADL 1999] <http://www.sdc.ucsb.edu>

[Gray] <http://research.microsoft.com/~Gray/>

[NRC 1994] Computing the Future, National Research Council, 1994 <http://www.nap.edu/readingroom/books/ctf/>

[TerraServer] Published papers about TerraServer <http://teraserver-usa.com/About.aspx?n=AboutPapers>