The Teraplex Experience:
Pushing the limits of Business Intelligence

For IBM Corporation, in *"Data Management Review" magazine*
Ghost-written by Robert Fripp, September '98
Released here with my client's permission

Pushing the limits, now as then

The present experimental stage of exploiting information technology (IT) in pursuit of business intelligence (BI) recalls an earlier age of commercial exploration. Magellan, Hudson and Columbus set out to discover new worlds for commerce without a critical navigational aid: knowledge of longitude. The stars told navigators how far north or south they were. But, without accurate clocks, position with respect to east or west was a matter of guesswork.

It took courage for early navigators to sail east or west after sunset. They were gambling. Did they have a night of clear sailing ahead, or would they break up in the dark on an unseen shore? I open with that analogy between sixteenth century voyages of exploration and IBM's ultra-modern Teraplex Integration Centers because our Centers were built to help and enhance that same spirit of commercial enterprise. Shifting the paradigm to serve a new millennium, our Teraplex Centers sustain commercial, competitive voyages into IT and BI unknowns. Briefly, IBM spent $47 million to set up sites where our BI customers and business partners can fine-tune their proofs of concept while scaling up to terabyte sizes, using all sorts of database and software combinations on S/390, RS/6000 and AS/400 frames. The S/390 and RS/6000 Teraplex Centers are in Poughkeepsie, New York; AS/400 customers go to Rochester, Minnesota.

Pushing the limits New Age-style

The Teraplex Centers' mandate is specific. They exist to explore the outer limits of BI in an age when massive computing power meets massive commercial data warehouse information demands.

Take different combinations of IBM hardware; software and databases from IBM and many independent vendors; add a variety of warehouse architectures; then scale any given configuration up from gigabytes to terabytes. At some point any [ITALICS] integrated system must hit the rocks. Teraplex users scale up until something snaps; then we help

them to fix, patch, rewrite code and go again. As we scale into unknown new worlds of terabyte demands, our Teraplex Centers reveal those critical break points which allow us all to chart our own route maps into the next stage and the next age of commercial penetration.

Drive it till it drops

The best analogy I can give is that of a test track for high performance vehicles, where you drive a product past its limits. Valves blow, bearings seize, brakes fail or the radiator bursts. Then you strip it down, analyze the failure, correct it, put it back on the track and kill it all over again. That is what the Teraplex Centers do for integrated BI systems. We put a premium on achieving system failures, because that's the only way to work through a problem before pushing the envelope again.

Let me illustrate the Teraplex Centers' mission by describing what can happen without them. I can describe this scenario because it's the sort of thing IBM set up the Centers to eliminate. Let's say a chief information officer is under pressure to scale up a data warehouse from a couple of hundred gigabytes to something over half a terabyte. The CIO does his homework and selects software to match his business function. But already there's a problem. Neither the customer not the software vendor has the resources to test an integrated system that is supposed to scale seamlessly into the stratosphere. Scaling up a data warehouse trips multiple stress points. How many concurrent users will the projected system handle? What happens to query response time set against a huge amount of data? What if the system crashes and you can't recover it or back it up?

Exit Big Bang. Enter trial and trial again

In the pre-Teraplex Dark Ages our hypothetical CIO might build his system, turn the key and -- crash! That's what I call the Big Bang theory. Meanwhile he's spent a lot of money, managers are screaming for results on commercial deadlines and his staff has never handled anything that size before.

That's the nightmare we work to eliminate. Teraplex staffs are dedicated to serve our customers and business partners; Teraplex users have three terabytes at their disposal; there are no commercial

pressures; and the only reason customers and software vendors come is to scale their BI systems and test their proofs of concept until they get them right.

Major scaling up requires a tight rein on the whole integrated infrastructure -- hence emphasis on Teraplex Integration Centers. It's back to the test track analogy: you're looking for the next stress point, the weak link in the chain.

Scale to any level: food for thought

A lot of IBM customers, business partners and software vendors have had the Teraplex experience. The stories they go home with are positive, and instructive.

For example, a team from Parts America spent thirty days testing their proof of concept for a data warehouse designed to manage inventory at 640 auto parts stores. Parts America had plans to scale up JDA's Retail IDEAS software on an AS/400 from 275 gb to half a terabyte. We thought of the AS/400 as a midrange machine until Parts America scored a first by scaling it up to 510 gb. Adam Brown, Parts America's Systems Director for Merchandising, had plans to scale up to 750 gb. But "as it went up we discovered so much we had not anticipated that 510 gb was as far as we needed to go."

Parts America's experience is typical. They had a debate at the Teraplex as to how their warehouse design would affect its size; and whether Parts America or JDA (the Retail IDEAS folks) were correct about data configuration. Apart from proving their concept, Parts America's Teraplex experience saved a lot of grief by probing their warehouse architecture step by step, showing how each change they made affected its growth and performance.

MicroStrategy also went home happy from a Teraplex experience. MicroStrategy's DSS Server analytical processing application was new to the S/390. We helped them stress-test their product running with DB2.

A real customer environment, without the customer

Here was another first. These applications had never been tested with such a high volume of transactions (370 fairly complex reports per hour) with so many concurrent connections between DSS Server and DB2. They took the combination up to 50 concurrent connections representing several hundred users, against a 300 gb database. Ashutosh Jhaveri, MicroStrategy's Engineer for the Very Large Database Technology Group, told us they came because "We were getting strong customer demand for certification of DSS Server Version 5.1 for DB2 390. The Teraplex Center is probably the best place we knew that could recreate a real customer environment and simultaneously run large-scale tests."

In the end, Jhaveri says, "We had a good time. The experience we got there led to us being able to fully integrate and optimize Version 5.1 for DB2."

Working with Teraplex users, IBM has discovered that the experience is a two-way street. We provide a dedicated resource staffed by people who debug problems as they arise in an environment free of commercial deadlines. Our users soon discover that knowing an IBMer who has fired up terabytes on a daily basis for a couple of years is a good friend to have. Meanwhile, we are also making changes to match our users' experience. As MicroStrategy's Jhaveri puts it: "IBM made several code changes as well as we did, so it was a mutual relationship that we established."

Mass suicide: a test bed for survival

I think the ultimate test of the Teraplex ideal to date has been the work we did on Sears' proof of concept. Sears wanted to see if Makuro merchandise planning software would scale successfully to their massive, multi-terabyte data warehouse.

So Sears set out to simulate an environment as close as possible to production conditions. Sears' Director of Merchandising Systems, Joseph Lichocki, sent a team to the RS/6000 Teraplex Center with a straightforward criterion: we have 800 stores, 8,000 classes of merchandise. Let's see this thing work from the perspective of concurrent users.

Sears was tough. They pushed the system to 300 concurrent users

and then -- even Lichocki admits they weren't being fair to Makuro -- they had those 300 users do the same intense update simultaneously. The way he tells it: "We had them all commit suicide at the same time."

Sears is now working into a two-phase implementation. First come national class [of merchandise] plans. After that they hit a point where they really stretch technology: they will develop store by store class plans in a database running close to two terabytes.

To see it from Sears' point of view, you have to put yourself in the position of a major retailer. "We were really looking at start of day type pressures on the software and the hardware," says Lichocki. Could an RS/6000 supporting Makuro sustain the workload? "On a typical day we'd be heavy in the morning and then taper off, but when our financial month ends, the question becomes: can the system sustain 300 concurrent users for 48 hours?"

The Teraplex Centers: designed to take the load

Sears' team worked through their proof of concept at the Teraplex in February before implementing Makuro- based production on August 3rd, giving Lichocki a breathing space for shake-down and roll-out. The way he tells it, "I was able to jump six months ahead and say, will this work? How often do you have a chance to do that?"

The Teraplex experience helped Sears assess hardware decisions, too. "In the end," says Lichocki, "we were able to tell -- with tremendous help from IBM -- that beyond a point it really doesn't pay to get more equipment. So we ran our tests not only from a concurrent viewpoint, but from the number of nodes in the planning engine and the data engine. There were probably 2,000 combinations of environments that we ran." Look at the thing from the point of view of a CIO charged with the responsibility of setting up a system of unprecedented size. "It's got to be a crystal ball," says Lichocki. "No matter what kind of experience you and your team have had at lower levels, all kinds of things happen when you start to stress them up."

Too true! But, to return to our analogy with old seafarers' voyages of exploration: if we're all in the same ship steering in the dark, we can at least shift the major stresses. We can take the loads off information

managers and shift them to the Teraplex Centres. That's why they're here. The last word goes to Joseph Lichocki's verdict on Sears' Teraplex experience. It speaks volumes. He says, "It gave me peace of mind."