

***Bolt-On Strategies for Improving
Data Warehouse Performance***

**Joyce Norris-Montanari
President DBTech Solutions, Inc.**

DBTech
Solutions

Introduction

Sooner or later, every IT Organization with a successful data warehouse will eventually find itself considering how to improve query performance. Fact:: if the data warehouse is used, it will grow as the users evolve and ask for more insightful information! With this in mind, every successful data warehouse will eventually approach its scalable limits. Until now, the typical solutions were:

- Upgrade hardware and memory, and hope that 'things' run faster.
- Migrate to a new platform (database and computer) and pray the actual conversion activities don't cost more than the new platform itself.
- Redesign the warehouse using the lessons learned, in the first iterations of the data warehouse, to direct the future architecture.

However, a new, less disruptive and less costly alternative is to employ a bolt-on capable hardware and software solution such as a high-performance data warehouse appliance that solves the performance problem, behind the scenes, while leaving the application environment and end user experience intact. Sounds good, right? To help you decide if a bolt-on strategy is good for you, this article will explore some of the more common bolt-on opportunities that exist today.

What does it Mean to Bolt-On?

A bolt-on technology is built to integrate easily and seamlessly into your existing architecture. The intent is to be virtually invisible to the user. As a simple analogy, if you replaced the engine in your car with a high performance engine that had a turbo charger, your car would feel the same from the driver's perspective, all the gauges, radio, seating, etc. would be the same, but the car would be able to go faster (*Real Fast!*). Seamlessness is a key capability of bolt-on technologies.

Sometimes confused with bolt-on is the term plug-and-play. Plug-n-play refers to something that is up and running quickly with minimal setup to perform the intended function. The advent of the technology appliance is closely linked to this concept. There are many kinds of appliances popping up. For example, a data warehouse appliance is intended to be plug-n-play for the task of data warehousing. A technology that offers both bolt-on *and* plug-and-play capability, like DATAlegro for instance, is an important consideration when seeking to greatly improve data warehouse performance with minimal disruption and effort.

Bolt This On For Size!

Most of us have created an operational system or data warehouse application with an end in mind. For instance, performance of the transactional data in the operational system may have been of prime importance to a fast moving company with a lot of transactions. However, the data warehouse may have been built in increments for managed reporting and, as new phases were added,

query performance may have deteriorated. Figure 1 represents an existing DW built for reporting but not able to accommodate desired ad hoc query capability perhaps due to performance issues caused by the introduction of large volumes of data or a design not conducive to growth. The remainder of this article will focus on the use of a bolt-on data warehouse appliance to improve specific performance issues.



Figure 1: Existing DW with Performance Issues

Ad hoc Query Bolt-On

If you have not been able to offer significant ad hoc query capability to your users because of the fear of disrupting the existing data warehouse performance, consider a bolt-on appliance (Figure 2). It can provide high performance for the complexity of ad hoc queries against large volumes of data, while preserving the current performance level for the managed reporting users. This opens up the world of information to those requirements that we could not meet with managed reporting. The need of 'I want access to everything you have!' means users could be looking for one specific item or analysis of a large amount of data with a product that looks for correlations between data items which could cause unpredictable effects on the managed reporting service levels.

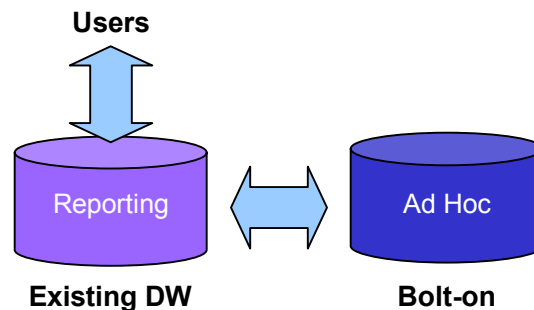


Figure 2: Bolt-on Offloads Ad Hoc Workload

Data Warehouse Reporting Bolt-On

If you have a huge number of daily reports to deliver from the data warehouse, and a shrinking batch window, consider offloading the heavy reporting infrastructure to a bolt-on appliance that can scan and aggregate huge volumes of detail data very quickly. Figure 3 shows the reporting environment is offloaded to the bolt-on appliance preserving the data warehouse as the collector of data.

This would especially work well in data warehouses that are global and receive data from various sites multiple times a day, or data warehouses that have concentrated on data acquisition and not delivery of the information into the hands of users.

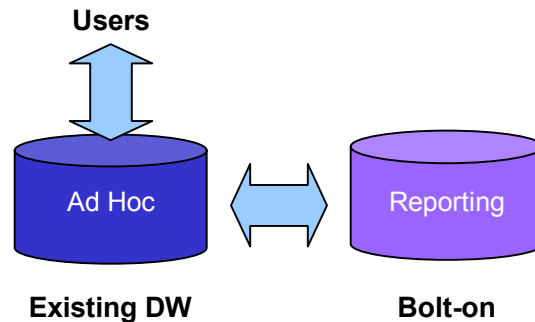


Figure 3: Bolt-on Offloads Reporting Workload

Operational Reporting Environment

Operational reporting got you down? Consider an appliance! Here is the scenario, the reporting in the operational systems are getting slower and slower. The batch and ad hoc reports are starting to affect the performance in the transaction system. We have considered replication into a redundant reporting environment, but can't justify the expense of the disk space, software, and memory required to set up the environment. Instead of using replication, you may want to consider using a bolt-on appliance for operational reporting, with loads at specified intervals. The database administrator can configure Oracle to recognize that certain tables are managed by the appliance (i.e. the tables that are used most often for reporting the operational data).

Other Scenarios

It is important to remember that bolt-on is a *strategy*, not a technology. Therefore, in addition to the previous scenarios where only a particular workload or group of users is handled by the bolt-on system, there is no reason that a bolt-on capable appliance could not also be considered for the following instances:

- New Data Warehouse – where the requirement is *"We need a faster stand-alone data warehouse that takes less effort to deliver the required performance."* In this instance, we could consider an appliance for the new data warehouse, with the knowledge that we can add another appliance when performance starts to deteriorate.
- New Data Warehouse and Data Mart – where the requirement is *"I have built the data warehouse with dependent data marts and would like an environment where everything queries faster."* In this instance, the appliance would house the data warehouse and the data marts creating a fast scaleable platform for use today and tomorrow.

- Data Marts – where the requirement is *“We built data marts that are growing bigger each day. The query performance is dwindling as we add more data.”* A data warehouse appliance is a cost effective solution for a large data mart that is continually growing.
- Pilot Project – where *“We have no business intelligence and would like to build a pilot project to show return on investment quickly. Along with the data warehouse we would like to have a lower cost of ownership.”* A data warehouse appliance, by its plug-n-play nature, is designed for quick and easy setup. We are talking days here, not months!

Summary

With a bolt-on, adding more computing power is as easy as adding another appliance. This may be more cost effective as a company grows instead of upgrading the existing computing platforms by adding more processors, disk space, memory, or bringing in a clustering technique to gain horsepower. The appliance is capable of delivering valuable information to the end users faster. In summary, the data warehouse appliance can be used to improve performance for the end users in many ways. This paper has barely scratched the surface of what we can do with such exciting technology. So if you are having performance issues in your data warehouse bolt it on and off we go.

About the Author

Joyce Norris-Montanari **Joyce Norris-Montanari** is President of DBTech Solutions. Her architectural expertise encompasses working with multiple vendors on numerous platforms and advising clients on tool selection through evaluation and analysis, especially on ETL, profiling, and metadata products. In addition to her data warehouse and operational data store experience, Joyce has a strong background in application system development in the mainframe, client/server and Internet environments. In addition to her international consulting work, Joyce speaks frequently at data warehouse conferences, is an instructor for public seminars, is a contributor to several trade publications and has co-authored with W. H. Inmon, Dan Meers and R. H. Terdeman on *Data Warehousing and E-Business*, published by Wiley & Sons. Joyce has managed and implemented data warehouses and operational data stores for pharmaceuticals, restaurants, telecommunications, government, medical organizations, financial institutions, oil and gas, insurance, research and development, and retail environments. Ms. Norris-Montanari can be reached at JMontanari@earthlink.net.



Thank you for reading this complimentary paper provided by DATAllegro.
For more information on this topic or DATAllegro products, please contact us at
(949) 330-7690, or visit www.datallegro.com.