



WhereScape RED
Builds data warehouses, fast.

The Dimensions of Fast

WhereScape Software

THE DIMENSIONS OF FAST

WhereScape RED enables developers to build, prototype and enhance data warehouses 10 to 100 faster than with traditional products. Since WhereScape RED enables this kind of massive productivity boost, why are some projects 100 times faster, some only 10 times faster, and some projects take the same amount of time?

Having a product that enables rapid development is critical to the increased productivity, but clearly there are other factors that influence the amount of time a project takes.

The critical factors that impact speed of delivery are:

- > Rapid development product
- > Approach
- > Data availability
- > Key user availability
- > Governance framework
- > Team size
- > Team experience
- > Technology

RAPID DEVELOPMENT PRODUCT

A rapid development product is critical to achieving orders of magnitude increases in productivity. WhereScape RED has been designed from the ground up to speed delivery. It has been designed and tested by data warehouse practitioners, and is built on real world experience not academic notions. Projects have been analyzed and dissected to see where the time is spent, and input has been sought on how to reduce elapsed time and as well as effort. Some product features were easy to identify – the ability to generate documentation at any time can shave literally weeks of the initial build, improves the ability to support the data warehouse (as the documentation is always up to date) and decreases support effort (as it can be regenerated after any changes) – and met with universal approval from developers. Some features impact the initial build, others reduce the support effort – all are focused on speeding the entire process.

Integrated meta data has a massive impact on productivity. The ability to organize and manage objects from a central location, check them out if they are being modified, and display them in a familiar manner for data warehouse professionals (load, stage, fact and dimension or model, aggregates, summaries, views, cubes etc) has a dramatic impact. Powerful impact analysis – the ability to quickly see the impact of a source system feed change or to see data lineage to identify how big an impact the change of an end user table or cube will be – supports rapid prototyping, the ability to build and the ability to enhance and maintain an in production data warehouse.

APPROACH

After the rapid development product, the two biggest impacts on speed are approach and data availability.

Using a traditional approach will take a traditional amount of time. WhereScape RED will speed up development, but you will not see orders of magnitude increases in productivity unless you are willing to do some things differently.

WhereScape RED encourages Live Prototyping – modeling done on real data with real users in small cycle times. Traditional “closed door” schema design processes – including waterfall-style schema design processes – are also fully

supported. WhereScape RED will not impact the design timelines for design tool or paper based designs, as in waterfall approaches design is done before development commences. Formal designs may be necessary in some cases, and the output from the design session can be used as a target for the development, reducing the number of iterations that are required to get the optimal model.

Prototyping has the advantage of getting to a shared understanding between users and developers of scope and complexity as rapidly as possible. If these are out of alignment then the project times will be impacted as rework will be required, or the project outcome will be put at risk when expectations are not met.

Where a model already exists (for instance the customer has a purchased a logical data model for their vertical) Live Prototyping can still be used for a shared understanding of the output, to confirm the model, or can be used for the presentation layer design using the model as a source.

WhereScape RED has been designed as a rapid development environment, and has features and functions designed to speed up delivery. It has also been designed to be extensible - at any stage you can veer from the Pragmatic Data Warehousing approach, and can modify or replace generated code with manual code. If you are an experienced data warehouse developer then this may well result in a "better" solution - more elegant, more efficient, or just more in line with the way you have always done projects. WhereScape RED generated code aims to be pragmatic - a balance between efficiency and ease of maintenance. WhereScape RED enables the developer to set the balance (for instance choosing set based code for speed or choosing cursor based code for ease of maintenance) or to set their own balance by using their own code.

DATA AVAILABILITY

The lack of good quality live data covering all scenarios (e.g. sales, returns, credits, discounts and rebates in a retail environment) has a major impact on speed. WhereScape estimates that going live with a data warehouse at the same time a new source system is being introduced takes 5-10 times as long as going live after quality data is available. With the flexibility of WhereScape RED this

reduces to 3-5 times as long – but this is still a significant impact on delivery time. In some cases there is no option, the data warehouse must go live at the same time. Any lag between source system go live and data warehouse go live will mean savings in effort as less rework will be required.

For prototyping the ability to construct small subsets of data that cover all possible scenarios is also critical to reducing cycle time.

KEY USER AVAILABILITY

Access to key users during the modeling and development process will result in better solutions with less rework. When key users are available rework decreases which reduces effort and elapsed time. WhereScape RED's code generation capabilities can mitigate the impact of rework, but it will still slow the project.

The other key benefit of key user involvement is the impact it can have on testing. If the key users have been involved in the process then they are already aware of the decisions that have been made, and testing is more confirmation than discovery.

An interesting and surprisingly common outcome of key user involvement is that often projects end up delivering less, but the stakeholders are pleased. When the development process is undertaken in isolation large amounts of time can be spent on one or two user requirements. When key users see the amount of effort that needs to be expended to meet what in some cases are actually minor benefits they will often modify requirements allowing the project to be delivered quicker or effort to be diverted to higher priority areas.

GOVERNANCE FRAMEWORK

A proper governance framework can have a significant effect on elapsed time. The governance framework needs to balance between the need to understand project status and the setting of intermediate goals that slow down overall progress. Good governance includes being aware that a project is going well, or knowing if it is going off the rails while there is still time to influence the outcomes and before all the money has been spent. Over zealous governance can result in the project team spending large periods of time preparing progress

reports and satisfying intermediate goals to the detriment of the end goal. Delivering 10 to 100 times faster means entire projects can be completed in a matter of days – over bearing governance and project structures may mean nothing is possible within such short time periods.

Governance frameworks have impacts at the technical as well as the project level. Sites that have long and complex sign off and promotion policies must add in this time to the project. If using WhereScape RED the entire project can be completed in two days rather than two weeks, yet the production promotion process still takes two weeks, then from a user perspective the development is 40% quicker, not 80%.

Another example of an intermediate goal is the imposition of external development approaches. WhereScape can be used in traditional SDLC environments, but this will generally negatively impact the development time.

TEAM SIZE

The larger the team size the harder it is to increase individual productivity. Large problems still require large teams, and then the ability to divide the work into smaller units will speed up delivery. With the productivity benefits of WhereScape RED, small focused teams can produce results very quickly. WhereScape RED is designed for rapid delivery teams of 1-10 developers. Larger teams are supported, but then WhereScape recommends multiple development environments which require a degree of process to manage, which can slow delivery. Larger teams also require a logical approach to splitting work. WhereScape RED has version control functionality, but while this can stop the overriding of changes, any bottlenecks will still impact development time.

TEAM EXPERIENCE

Experienced teams develop quicker than inexperienced teams. Developers who are familiar with data warehousing and the underlying databases will be more productive than inexperienced people. WhereScape RED will make experienced and skilled people super productive, it will make competent people orders of magnitude more productive, and while it will assist inexperienced people and teams they will not be as productive as more skilled people.

TECHNOLOGY

Undersizing hardware will impact developer productivity. Insufficient development hardware will mean that jobs take longer to run, reducing development efficiency and increasing cycle time. This can be mitigated by working on smaller data sets, but smaller data sets may not have the same range of data as larger data sets. The ability to rapidly iterate prototypes is critical to rapid prototyping. If prototype code needs optimization this negatively impacts cycle time.

If production hardware is undersized, more time will have to be spent on development. This is a compromise – more powerful hardware costs more, but if the hardware is fast you don't need to spend so much time tuning the database and optimizing jobs. Simple (as opposed to highly optimized) jobs are often easier to maintain, saving time on enhancements and maintenance as well as the initial build process.

CONCLUSION

Delivering incrementally faster can be done with incremental changes. Orders of magnitude increases in speed are possible, but they do require a change in approach. WhereScape RED provides a fast way to build data warehouses. Building 10-100 times faster requires a highly optimized approach as well as a highly optimized product.

© 2006 WhereScape Software Limited. All rights reserved.

WhereScape is a trademark or registered trademarks of WhereScape Software Limited.

Brands or product names are the trademarks or registered trademarks of their respective owners.

For more information on WhereScape RED or to schedule a web demonstration visit www.wherescape.com