



**Natural Attraction Forces of Data Bodies
Within a Data Grid**

To Describe Efficient Data Distribution Patterns

Proposed by:

Michael Di Stefano

mdistefano@integrasoft.com

Integrasoft, L.L.C.

September 14, 2004



Natural Attraction Forces of Data Bodies Within a Data Grid To Describe Efficient Data Distribution Patterns

Michael Di Stefano, September 14, 2004

Introduction

This paper puts forth the premise that a practical and optimal data distribution pattern of data bodies within a Data Grid can be described through forces of attraction between data bodies and the force of friction as put upon them by the physical properties of the Data Grid.

As bodies in space, the planets, stars, satellites, etc. follow natural laws of attraction and movement within the fabric of space, Single Data Bodies will follow their own set of natural laws of attraction within the physical fabric of the Data Grid. Similar to the views of Newtonian, Einsteinian, and Quantum Physics, I will apply both the micro and macro views of Data Distribution within a Data Grid. Similar to how quantum physics describes orbits of electrons around the nucleus guided by discrete quantum energy levels, the micro view for the data distribution concerns itself to the manner in which data distribution is present within a Single Data Body.

On the other hand, the macro view of data distribution concerns itself with the physical data distribution patterns of *multiple* data bodies in relation to each other within the space of the Data Grid fabric. Newtonian laws accurately describes a physical body in terms of mass and force, such as apples falling to earth in contrast to Quantum Physics that describes the internal atomic structure of the apple. I will follow the same approach, being less concerned with the internal Data Atoms and the distribution pattern that form a Single Data Body and more interested in the fact that the Single Data Body exists, has definite and measurable quantities. These Single Data Bodies exhibit forces of attraction on one another that will result in an optimal distribution pattern that best meets the requirements of the larger system thus resulting in minimal data movement of the overall system.

Einstein's Theory of Relativity describes the bending or warping of the fabric of space and time by mass. Separately, Newton's Laws and the Theory of Relativity do not describe the complete picture; a unified view has been put forth but not yet proven, states that the curvature of space and time causes mass. The forces of attraction of Single Data Bodies cannot be fully described by the forces of attraction they exert on each other. Similarly the unified view in physics, the physical characteristics of the Data Grid Fabric (DGF) will have influence on a Single Data Body and the distribution patterns among the Single Data Bodies in the Data Grid Space (DGS).

Natural Attraction Forces of Data Bodies Within a Data Grid To Describe Efficient Data Distribution Patterns

Michael Di Stefano, September 14, 2004

Observation

Most systems are composed of many different data types. For example, a portfolio management system requires market data, holdings portfolio, and risk exposure data in order to determine the best course of action for when and what to buy or sell. When this system resides in a Grid topology, the data types of market data, portfolio, and risk form their own respective Single Data Bodies in the Data Grid.

Initially, the system running in the Grid may not be optimized since the optimal data distribution patterns have not been discovered yet. One may assume some mathematical model of data distribution for initial deployment followed by continued analysis of natural data movement patterns over time. The resulting data movement patterns will suggest manual adjustments of the optimal data distribution pattern of the Single Data Bodies in the Data Grid in order to minimize data movement within the DGF during normal operation. This process will ultimately result in the formation of Single Data Bodies of definite size and shape in relation to each other to minimize system wide data movement.

While this process will ultimately result in a steady state of Single Data Body size, shape and relation to other Single Data Bodies of the system; it is one of trial and error that can span long periods of time. During this discovery phase, the Data Grid will not be operating at optimal performance levels. This is an un-acceptable behavior, as Grid Computing is starting to gain wide acceptance in both the public and private sectors with new systems migrating in mass from client server onto the Compute Grid.

Therefore, a data distribution model that accurately describes a system wide distribution pattern is mandatory so as to start the system closer to the optimal distribution pattern, which will minimize the settlement time required to achieve a steady state. This initial optimal behavior can be achieved by defining the Single Data Bodies in quantifiable terms; the laws of attraction between the Data Bodies, and the effect of the “warping” of the DGF has on them.

For the Complete Paper

Contact

Michael Di Stefano

mdistefano@integrasoftware.com