Journal of the Arkansas Academy of Science

Volume 48 Article 49

1994

Fluid Dynamics Model of Data Acquisition and Date Analysis for High-Energy Physics

Charles M. Byrd
University of Arkansas at Little Rock

Christine A. Byrd
University of Arkansas at Little Rock

Wilson H. Howe
University of Arkansas at Little Rock

W. J. Braithwaite
University of Arkansas at Little Rock

A Fluid Dynamics Model of Data Acquisition and Data Analysis for High-Energy Physics

Charles M. Byrd, Christine A. Byrd, Wilson H. Howe, and W.J. Braithwaite
Department of Physics and Astronomy
University of Arkansas at Little Rock
Little Rock, AR 72204

The current paradigm for accelerator-based high-energy physics experiments involves the design of two distinct detector subsystems for managing data, namely data acquisition and data analysis. In order to establish the design parameters for detector components, the choice of technology for these data management subsystems must be made years in advance. By the time detectors come on-line, hardware will often be obsolete, because the technology choice was made years earlier. The majority of the data analysis will occur off-line, making expensive data storage necessary. Furthermore, the present gap between data acquisition and data analysis would be reduced by having a homogenous software environment throughout the experiment.

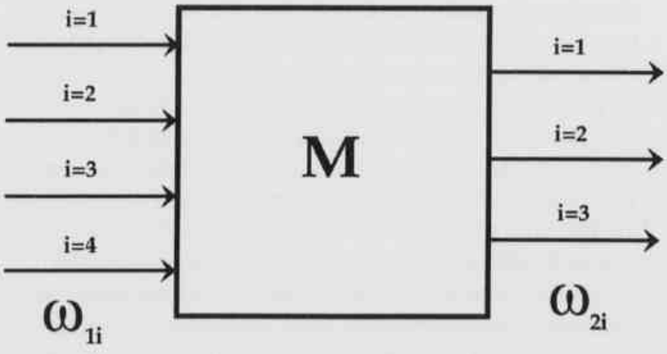
The Relativistic Heavy Ion Research Group at the University of Arkansas at Little Rock has developed an alternative archetype for data processing in high-energy physics which makes use of the transputer model to achieve a balance between data communication and data processing. By providing an integrated data acquisition and data analysis environment, the use of transputers allow computing resources to be maximized while minimizing data storage requirements for high-energy physics experiments. By taking advantage of high performance communication hardware developed by GE Corporate Research and Development, a flexible data management system is being designed to provide necessary parameters for detector component development without forcing premature technology choices. The hardware independence of this model insures the experiment can always be operated with state of the art technology. Using a homogenous software environment, this transputer computing model allows a significant portion of the analysis to occur on-line, reducing the volume of data passed to storage systems.

The data streams generated at the Relativistic Heavy Ion Collider (RHIC), Brookhaven and the European Center for Nuclear Research (CERN), Geneva are so enormous that a fundamentally new way of thinking about data acquisition and processing must be developed. The volume of data from the final state particles is so large that it becomes reasonable to use a macroscopic model for what is usually thought of as microscopic. Fluid-dynamics is a near-perfect macroscopic model for data acquisition because it provides a way to quickly obtain consequences to proposed changes. Therefore, to aid in the future development of the algorithms for massively parallel processors, an analogy was drawn between the data flow and fluid flow. In this paper, the fundamental principles of fluid dynamics (Shapiro, 1954) are applied to problems in data acquisition for high-energy physics.

A very simple fluid dynamics model is one which has a control volume M with a mass flow rate in ω_1 and a mass flow rate out ω_2. An Equation for the overall mass balance follows (where $d\theta$ is change in time):

$$\omega_1 - \omega_2 = dM/d\theta,$$
$$\text{or, } \omega_2 - \omega_1 + dM/d\theta = 0.$$

These equations say, simply, that the difference in mass flow rate in and mass flow rate out is equal to a rate of accumulation of mass in the control volume. Usually, one must deal with multiple inputs and outputs in a system such as the one below.



With multiple inputs and outputs the mass balance equations are as follows (subscripts indicate mass balance of each flowing substance):

$$\omega_{2i} - \omega_{1i} + \frac{\partial M_i(\theta)}{\partial \theta} = 0$$

$$X_{ki} = \frac{\omega_{ki}}{\omega} \quad \text{(for k = 1 or 2)}$$

$$\omega_{2i} X_{2i} - \omega_{1i} X_{1i} + \frac{\partial M_i(\theta)}{\partial \theta} = 0$$

The concentration X_{ki} (k = 1 or 2) is the ratio of the k-th mass flow rate to total flow rate.

The analogy between fluid flow and data flow follows: volume is analogous to memory, mass/sec is analogous to bits/sec, chemical reaction is analogous to data processing, ω_1 (mass in) is analogous to D_1 (data in), ω_2 (mass out) is analogous to D_2 (data out), and $dM/d\theta$ (accumulation of mass) is analogous to dP/dt (accumulation of processed data). Data flow is reduced at each step of the analysis through dP/dt. A large dP/dt results in a proportionately large reduction in ω_2.

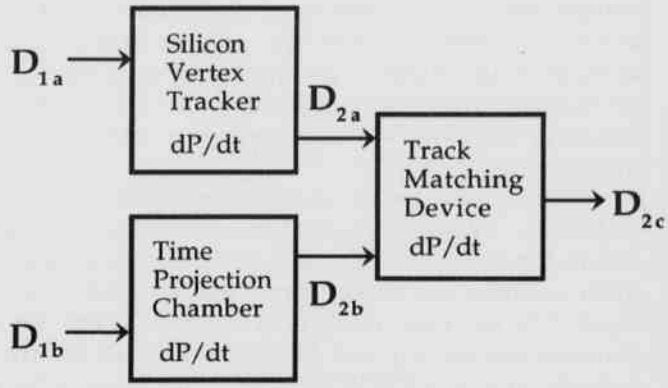
The simplest example to model using the principles of fluid dynamics is a two-detector, two-stage data acquisition system. This modeling example is a frequent occurrence in high-energy physics, as the output of two independent systems may be merged to give a unified set of data describing the particle events. This type of paired groupings of data streams may be carried out more than once in modeling the full data acquisition and analysis for the entire system.

An example of merging data streams from a two-detector, two-stage system is found in the track matching effort of the Silicon Vertex Tracker (SVT) and the Time Projection Chamber (TPC) of the STAR Instrument at RHIC. Another example is in the track matching of two vertex TPCs in the NA49 Experiment at CERN. Each example uses a track matching device following the two-detector data streams.

The diagram at the right represents a two-stage system being modeled. Note the analysis software for each detector component (represented by dP/dt) filters out the most interesting events, significantly reducing the outgoing data flow to the track matching device, which in turn filters the data further (also represented by dP/dt).

Using a tangible example of the diagram to the right analysis, particles registered by the SVT and the TPC must be matched by a track matching device. The track matching device must wait for data from both detector components before further processing the data. Since the SVT is a faster device than the TPC, data streams originating from the SVT will reach the track matching device sooner than data streams from the TPC. Thus, in the interest of efficiency, transputers processing elements within the track matching device not currently being used may be reallocated to the TPC by taking advantage of their reconfigurable backplanes. Processor idle time may be minimized under master computer control by this type of dynamic reallocation of elements (Byrd et al., 1993).

There are many advantages to using integrated data acquisition. For instance, data archival costs may be substantially reduced due to prefiltering data used by trigger software. Since interesting data may be extracted by the analysis software at each level, cost of off-line analysis may be significantly reduced, as less archival storage is needed due to prior rejection of the less interesting events.



Acknowledgements

This work was supported in part by U.S. Department of Energy and the Arkansas Science and Technology Authority. This work acknowledges the advice and support of the STAR Collaboration at RHIC and of the NA49 collaboration at CERN. The second and third authors acknowledge financial support from the UALR Donaghey Scholars Program.

Literature Cited

Byrd, C.M., C.A. Byrd, D.L. Roetzel and **W.J. Braithwaite.** 1993. Role of MPP Granularity in Optimizing Monte Carlo Programming. Proceedings of the International Conference on Monte Carlo Simulation in High-Energy and Nuclear Physics. (World Scientific Publ. Co., Singapore) 40-44.

Shapiro, A.H. 1954. The Dynamics and Thermodynamics of Compressible Fluid Flow. Ronald Press Co., New York p. 12-15.