DOI: 10.1017/S1431927603440427
Microsc Microanal 9(Suppl 2), 2003

COMPUTATIONALLY MEDIATED EXPERIMENTAL SCIENCE

NESTOR J. ZALUZEC
Argonne National Laboratory, Electron Microscopy Center, Materials Science Div., Argonne, Il, USA

The problems being addressed by today's microanalysts are becoming increasingly complex and data intensive. Thirty five years ago, many of us were challenged by the mere process of recording a spectral profile and then using the limited resources we had at hand to analyze the data to obtain quantitative results. Today, we have within our laptop computers the processing power of supercomputers, however, computing power alone will not be sufficient to solve the next generation of problems. To truly create a new paradigm of how we, as experimentalists, work we have to consider what are all the limiting factors to employing our resources to their greatest utility, then come up with new ways of combining these resources to change the how we perform these tasks.

Taking as a fundamental the premise that we are interested in extending the range and diversity of problems that we will be dealing with in the future and not just simply improving the resolution at which we do any individual measurement , then we will be challenged to consider experiments which here-to-for were considered beyond the realm of achievability. The obvious questions thus become:

1.) In what way are our current experiments limited by the way we work, rather than our instruments?
2.) How can we push the envelope of technology to permit our solving new types of problems?
3.) Where will the breakthroughs in new ways of working be realized?

Given the ever growing tendency to add computational resources to our instruments it is clear that the next advance will be directly related to how well we can effectively merge the realms of computational science and experimental science together. In the past we have used computers to simply speed up our experiments, in the last decade we have expanded this role to permit various degrees of telepresence operation. In the coming decade the key to changing how we work will be to realize that once an effective interface of instrumentation and computational tools has been developed, then we must change the way in which we design and conduct our experiments. This means not only re-examining how we do experiments so that measurements are done not just efficiently and with a modicum of speed, but more importantly to redesign these experiments, in such a way so as to maximize the information measured from the specimen. In this way the data acquired can be “mined” for content after the fact, using tools which may not reside within the instrument room, but possibly at remote locations and not just by the instrument operator but a colleague from any location.

As example of this new type of experiment consider the technique of Position Resolved Diffraction[2]. Here a focused electron probe is sequentially positioned in a two dimensional pattern on a thin TEM specimen and at each point a complete electron diffraction pattern (EDP) is acquired, stored and ultimately analyzed (figure 1) . In the past this type of experiment would have been simply dismissed as impractical as it simply puts too great an onus on the operator conduct the experiment. Consider Table 1 which documents the data set sizes of these experiments. A minimal measurement might involve a measurement on a 64 x 64 pixel spatial array (X,Y) grid at each point measuring a diffraction pattern at 1K x 1K pixels (α x β) yielding ~ 4Gbytes of data per measurement, in comparison a study of a 1K x 1K spatial array of points produces ~ 1Tbyte of data. Both the former and later of these scenarios is well beyond the current processing capabilities of any humanly directed process, as well as for most existing desktop DAQ systems present on instruments today. This even neglects the fact that there are fundamental limitations of popular operating systems, for example, in W2K there is a limit an individual file of 2GB set by the file system. Fortunately, not all operating systems suffer this problem.

To tackle these types of problems the clear challenge to the next generation of experimentalists becomes both demanding as well as integrating new advances in information technology, networking, and processing with our methodology in such a way that we can realistically tackle the next generation of data intensive experiments. To this end, at ANL we are working with computational scientists to developing a set of Grid[2] enabled tools to facilitate network coordination of computational resources with the aim of changing the way experiments are done. The intention is for these new tools to integrate network aware resources linking: storage, communication, control together with computational power to facilitate not only data acquisition, but also data mining and remote collaboration to a degree which is unprecedented today.

References:

1.) N.J. Zaluzec, Microscopy & Microanalysis 2002, Vol. 8, Sup. 2, pg 376 Quebec City,
2.) K. Czajkowski,etal. Proc. of the 10th IEEE Int. Symp. on HPDC IEEE Press, August 2001.“Grid Information Services for Distributed Resource Sharing”

3.) This work was supported in part by the U.S. DoE under BES-MS W-31-109-Eng-38 at ANL.

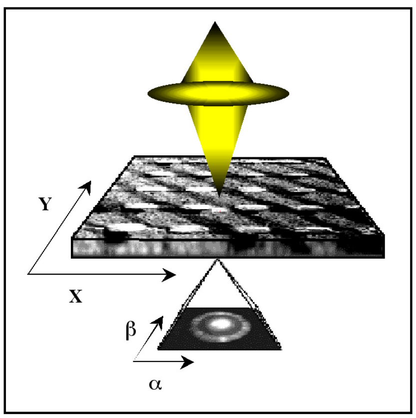

Fig.1. Position Resolved Diffraction (PRD) experimental arrangement.

αβ / XY	256^2	512^2	1024^2
64^2	262M	1G	4.2G
128^2	1G	4.2G	16.7G
256^2	4.2G	16.7G	67G
512^2	16.7G	67G	275G
1024^2	67G	275G	1T

Table 1. Data Set Size for PRD Experiments

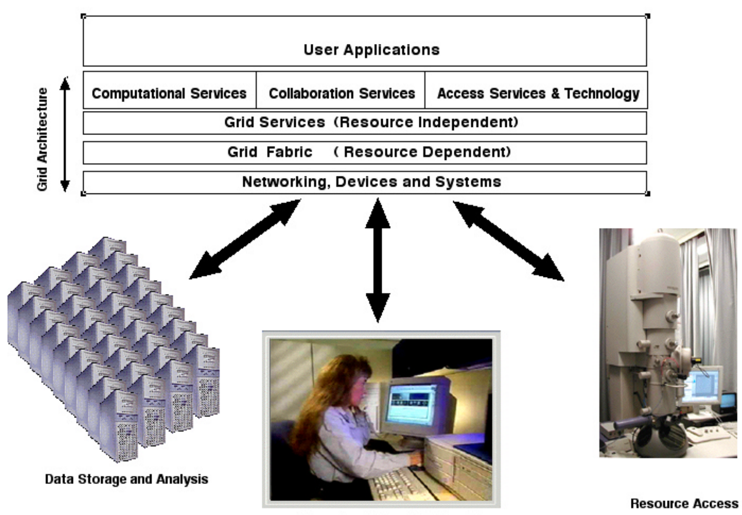

Fig. 2. Linking Grid technologies with user applications for state-of-the-art Microscopy and Microanalysis.

https://doi.org/10.1017/S1431927603440427 Published online by Cambridge University Press