Basic Energy Sciences “Computing” (and related matters)

Dale D. Koelling

31 October 2000

http://www.sc.doe.gov/production/bes/

Organization - The Usual Charts

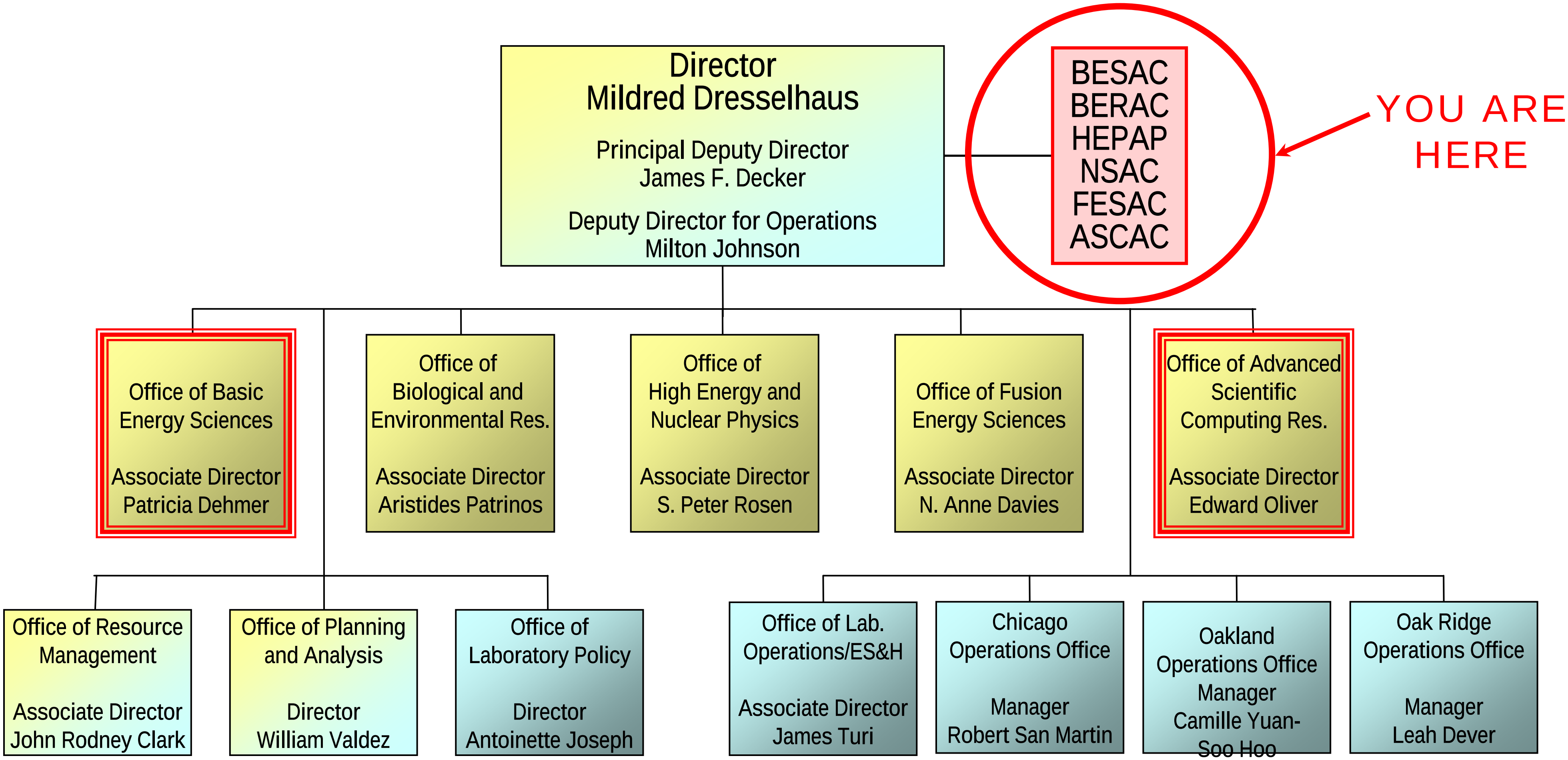
- BES and ASCR each have their advisory committees - I'm going to be telling you some of what BESAC was told at their most recent meeting. Full presentation is available at the URL:

http://www.sc.doe.gov/production/bes/BESAC/BESAC.htm

Click on "Meetings"
Under October 10-11, Click on "Presentations"
Click on "Overview from BES".

- The BES organization chart is here so you can see the people.

Office of Science



BESAC URL: http://www.sc.doe.gov/production/bes/BESAC/BESAC.htm
(Look under Meetings for 10-11 October meeting. Dehmer discusses both budget and research demographics in her overview presentation.)

BES Mission:

- *Foster and support fundamental research to provide the basis for new, improved, environmentally conscientious energy technologies;*
- *Plan, construct, and operate major scientific user facilities for "materials sciences and related disciplines" to serve researchers from academia, federal laboratories, and industry*

Office of Basic Energy Sciences

Patricia Dehmer, Associate Director
♌ **Iran Thomas, Deputy Associate Director**
Mary Jo Martin, Secretary

Associate Director's Office Staff Contacts
Robert Astheimer
F. Don Freeburn
Stanley Staten
Sharon Long

Materials Sciences and Engineering Division

♌ **Iran Thomas, Director**
Christie Ashton, Program Analyst
Tarra Hardeman, Secretary

Metal, Ceramic, and Engineering Sciences
Robert Gottschall
Terry Jones, Proc. Tech.

- **Structure and Composition of Materials**: Robert Gottschall; Vacant FTE; ◆Robert Hwang, SNL
- **Mechanical Behavior of Materials and Radiation Effects**: Yok Chen; ◆Robert Hwang, SNL
- **Physical Behavior of Materials**: ◆Andrew Quong, LLNL; ◆Wendy Cieslak, SNL; ◆Robert Hwang, SNL; ●Angus Rockett, U. of IL
- **Synthesis and Processing**: Alan Dragoo; Vacant FTE; Timothy Fitzsimmons; ◆Angus Rockett, U. of IL
- **Engineering Sciences**: Robert Price; ●Bassem Armaly, U. of MO; Timothy Fitzsimmons; ◆Wendy Cieslak

Condensed Matter Physics and Materials Chemistry
William Oosterhuis
Melanie Becker, Proc. Tech.

- **Experimental Condensed Matter Physics**: Jerry Smith
- **Theoretical Condensed Matter Physics**: Manfred Leiser; Vacant FTE; ◆Dale Koelling, ANL
- **Neutron and X-ray Scattering**: Helen Kerch
- **Neutron and X-ray Scattering Facilities**: Vacant FTE
- **Materials Chemistry**: Dick Kelley; Vacant FTE; Matesh Varma
- **EPSCoR**: Matesh Varma
- **Spallation Neutron Source**: Jeffrey Hoy

Chemical Sciences, Geosciences, and Biosciences Division

William Millman, Acting Director
Karen Talamini, Program Analyst
Carolyn Dorsey, Secretary

Molecular Processes and Geosciences
Paul Smith (Acting)
Diane Matthews, Proc. Tech.

- **Catalysis and Chemical Transformations**: Vacant FTE; ◆Daniel Melamed, BNL
- **Chemical Energy and Chemical Engineering**: Paul Maupin
- **Separations and Analysis**: Vacant FTE; ●Dick Gordon, Wash. State U; ◆Norman Edelstein, LBNL
- **Heavy Element Chemistry**: Paul Smith; ◆Norman Edelstein, LBNL
- **Geosciences**: Nicholas Woodward; ●Roger Turpening, MTU; ◆Henry Shaw, LLNL

Fundamental Interactions
Allan Laufer
Sharon Bowser, Proc. Tech.

- **Atomic Molecular and Optical Physics**: Eric Rohlfing
- **Chemical Physics**: William Kirchhoff; Eric Rohlfing
- **Photochemical and Radiation Sciences**: Mary Gress; Walter Stevens
- **Facility Operations**: William Millman; William Kirchhoff; Paul Smith

Energy Biosciences
Gregory Dilworth
Patricia Snyder, Proc. Tech.

- **Plant Sciences and Microbiology**: Gregory Dilworth; James Tavares
- **Biochemistry and Biophysics**: Walter Stevens; Vacant FTE; ●Sharlene Weatherwax, UCLA

Color key:
- Materials Sciences Subprogram
- Chemical Sciences Subprogram
- Engineering and Geosciences Subprogram
- Energy Biosciences Subprogram

Legend
- ♌ Dual Capacity
- ● IPA
- ◆ Detailee

21 September 2000

The Relationship: By Effort

- For this presentation, in the listing of Office of Science efforts (next slide), think of matching the red (BES) and the green (ASCR).

- Here, we consider three broad areas:
  - "Computing"
  - "Networking"
  - Applied Mathematics

Office of Science Major Research Areas

Materials Sciences

- Catalysis
- Ceramics
- Condensed Matter Physics
- Corrosion
- Electronic Properties of Materials
- Experimental Techniques & Instrument Devel.
- Intermetallic Alloys
- Magnetism and Magnetic Materials
- Materials Physics and Chemistry
- Mechanical and Physical Behavior
- Metallic Glasses
- Metallurgy, Metal Forming, Welding & Joining
- Neutron and Photon Scattering
- Nondestructive Evaluation
- Photovoltaics
- Polymer Science
- Radiation Effects
- Solid Dynamics
- Structural Characterization
- Superconductivity
- Surface Science
- Synthesis and Processing Science
- Theory, Modeling, & Computer Simulation

Geosciences

- Geochemistry of Mineral-fluid Interactions
- Geophysical Interrogation of Earth's Crust
- Rock-fluid Dynamics
- Biogeochemistry

Engineering Sciences

- Materials Engineering
- Nanotechnology and Microsystems Engineering
- Multi-component Fluid Dynamics and Heat Flow
- Nonlinear Dynamic Systems

- BES - Basic Energy Sciences
- HENP - High Energy and Nuclear Physics
- BER - Biological & Environmental Research
- ASCR - Advanced Scientific Computing Res

Chemical Sciences

- Analytical Chemistry
- Atomic, Molecular & Optical Physics
- Advanced Batteries & Fuel Cells
- Chemical Kinetics
- Chemical Physics
- Catalysis - Homogeneous and Heterogeneous Phase
- Combustion Dynamics
- Electrochemistry
- Heavy Element Chemistry
- Interfacial Chemistry
- Organometallic Chemistry
- Photochemistry
- Photosynthetic Mechanisms
- Radiation Effects
- Separations Science
- Solar Energy Conversion
- Theory, Modeling, & Computer Simulation
- Thermophysical Properties

Biosciences

- Biochemistry, Biocatalysis, Bioenergetics, Biomaterials, and Biophysics
- Extremophilic Organisms
- Fermentation Microbiology
- Photosynthetic Mechanisms
- Plant and Microbial Sciences
- Plant Genomics

Physics

- High Energy and Particle Physics
- Heavy Ion & Medium Energy Nuclear Physics
- Accelerator and Detector R&D
- Particle Astrophysics
- Physics Theory
- Plasma Physics
- Advanced Fusion Designs & Specialized Materials

Life Sciences

- Human Genome
- Structural Biology
- Microbial Genome
- Low Dose Radiation Research
- Functional Genomics
- Human Subjects in Research
- Structural Biology Facilities
- Genome Instrumentation
- Computational & Structural Biology

Medical Sciences

- Molecular Radiopharmaceutical Development
- Boron Neutron Capture Therapy
- Molecular Nuclear Medical Imaging
- Imaging Gene Expression
- Biomedical Engineering

Environmental Sciences

- Decade to Century Climate Modeling
- Atmospheric Radiation Measurement (ARM)
- Atmospheric Science & Chemistry
- Carbon Cycle Research
- Ocean Sciences
- Ecosystem Function and Response
- Information & Integration
- Integrated Assessment of Climate Change
- Bioremediation of Metals & Radionuclides
- Environmental Molecular Sciences Lab

Mathematics and Advanced Computing

- Linear Algebra Libraries
- Scientific Computing & Network Testbeds
- Advanced Computer Science
- Applied Mathematics
- Advanced Computing Facilities
- Advanced Computing Software and Collaborative Tools

BES "Computing" now

- BES traditionally utilizes just under a quarter of the resources provided by the National Energy Research Scientific Computing Center (NERSC)
  - Traditionally heavy CPU intensive.
  - Memory usage is growing very significantly.
  - Archival storage small to modest.

- BES researchers have participated significantly in the ASCR computing research centers at ANL, LANL, and ORNL as well as others such as PNNL and Sandia.

- Much work is done on workstations which now has also expanded into Beowolf clusters. [Here we motivate discussions on the relative benefit of local and global management and on "capability" versus "capacity".]

- Database sharing and archiving has been provincially organized (and may not be able to continue being so).

BES: Networking now

- BES researchers need all the traditional network services: remote login, file transfer, Email, and WEB services. This is where ESNET grew up and it has given yeoman service. Now for the world of graphical user interfaces and immersa- {desk, deck, …}.

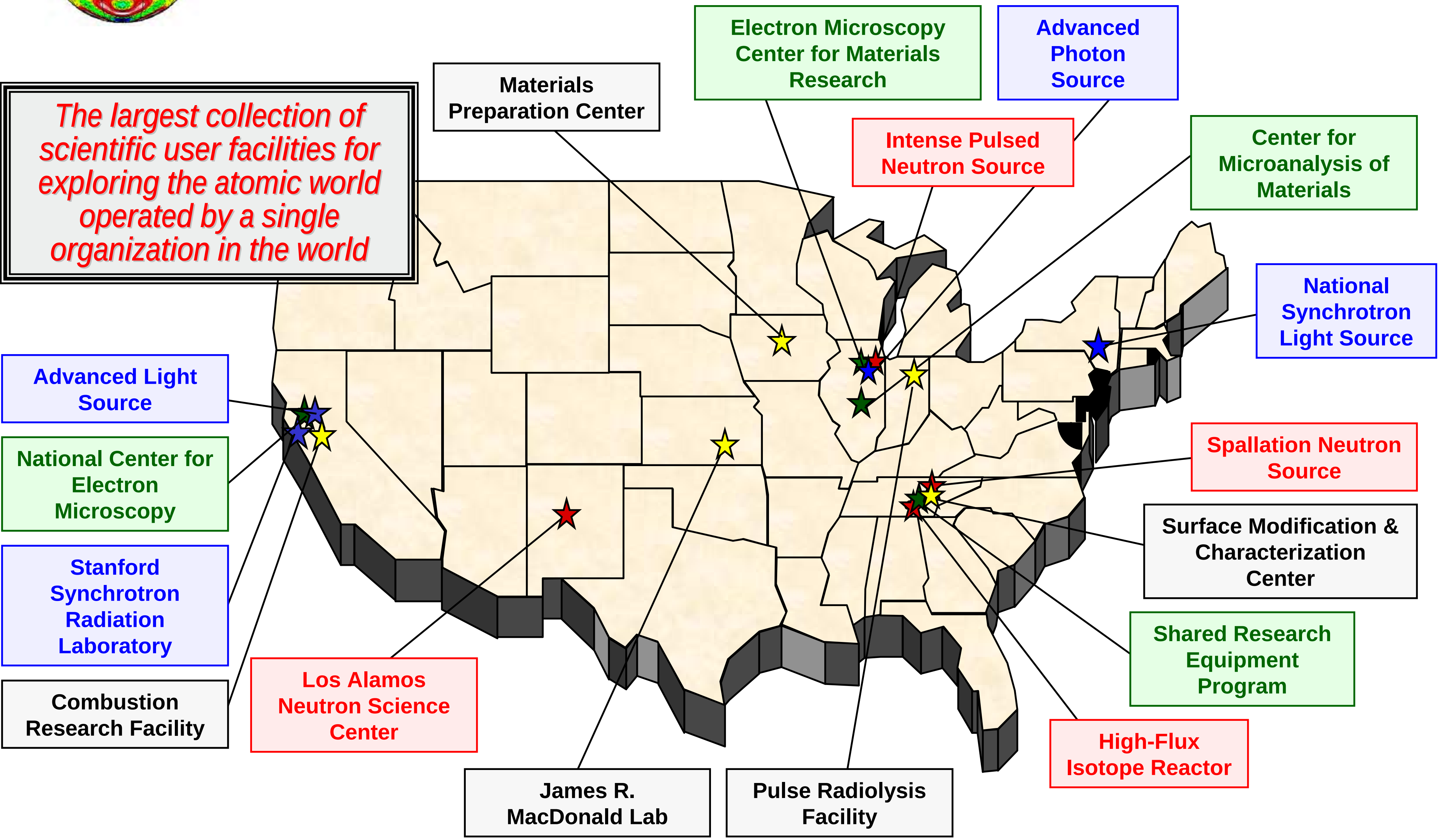
- BES operates national facilities (and other data sources) which have special communication challenges.
  - Data storage and transfer (data is getting more massive)
  - Off-line data analysis (or near-line, which is trickier)
  - Collaborator efforts (remote presence is a high bandwidth problem)
  - Remote operation (more detailed remote presence *and* big "quality of service" issue)

Two most touted examples are B. Tonner's telepresence at the ALS and the microscopy collaboratorium. See

http://tmp.amc.anl.gov

BES Facilities & Collaborative Research Centers

The largest collection of scientific user facilities for exploring the atomic world operated by a single organization in the world



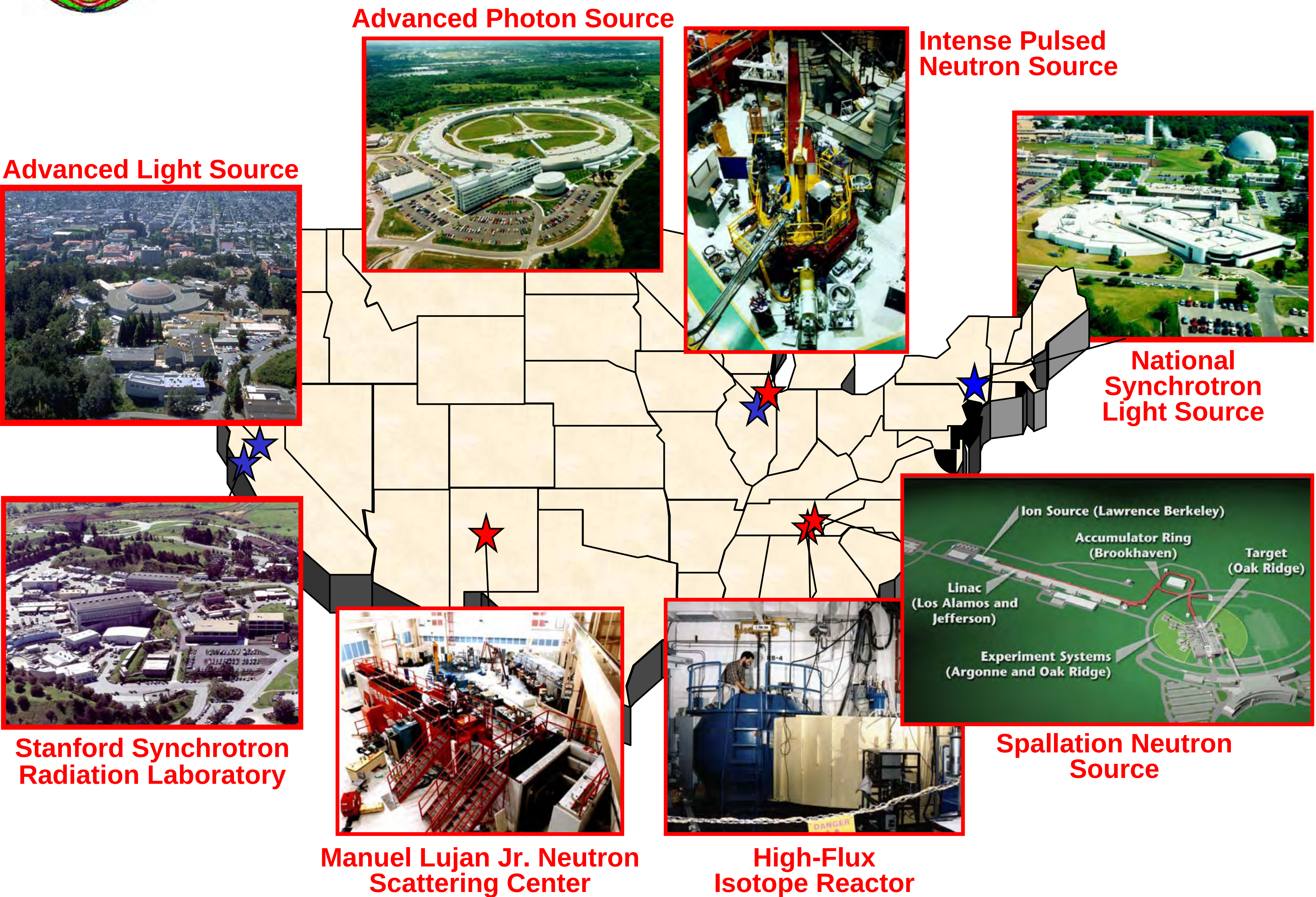
Major User Facilities:

- ***4 Synchrotron Radiation Light Sources***
- ***4 High-Flux Neutron Sources***

Collaborative Research Centers:

- ***4 Electron Beam Microcharacterization Centers***
- ***5 Special Purpose Centers***

BES X-ray and Neutron Scattering Facilities



BES: Applied Mathematics now

- BES researchers understand and utilize libraries (I personally grew up on EISPACK, LINPACK, LAPACK, …)
- Joint projects do happen at all levels --- between investigators, local groups, and ”grand challenges”. The biggest roadblock is actually agreeing on the problem!
- BUT, opportunities do go knocking. Consider the history of a couple of electronic structure thrusts
  - Car-Parrinello and friends actually determine electronic structure to derive atomic forces for Molecular Dynamics.
    - Selective use of precision.
    - Pseudopotentials.
    - Reuse of data in an iterative fashion: Lanczos
    - Non-linear optimization.
  - Order N methods seek to reduce electronic structure effort from increasing as N^3 to only increasing as N --- the number of electrons.
    - Truncation
    - Iterative solutions
    - Non-linear optimization

The Times (Problems) change!

- After presenting an extensive set of statistics to BESAC, Dehmer characterized what they all meant (the next foil). The simple way to say it is the old phrase: “Hang together or hang separately.”

- Terms like “Correlated Electrons”, “Complex Materials”, “Multiple Scales”, “Nanoscience”, “Microstructure”, and many more have significantly increased problem size and complexity. (That means that as the hardware gets better, we’ll be anxious to use it. But that’s not the primary issue here!)

What Did All That Mean?

- *BES work at the DOE labs -- once dominated by individual investigator/small group activities -- is now dominated by world-class scientific facilities serving the Nation, by collaborative research centers, by research associated with the themes of these facilities and centers, and by other research uniquely suited to the laboratories. This trend is supported by numerous blue-ribbon panels.*
- *Work at universities is a critical component of our portfolio. It has remained a constant fraction of the research portfolio for more than a decade, and it will so continue.*
- *Laboratory activities are increasingly linked to activities at other institutions.*
- *"Flat funding" for the physical sciences in SC <u>isn't</u>. It's much worse than flat.*

Lets hang together...

- *BES work at the DOE labs -- once dominated by individual investigator/small group activities -- is now dominated by world-class scientific facilities serving the Nation, by collaborative research centers, by research associated with the themes of these facilities and centers, and by other research uniquely suited to the laboratories. This trend is supported by numerous blue-ribbon panels.*
- *Laboratory activities are increasingly linked to activities at other institutions.*

Implying ---

- ⇦ **More networking (ESNET Steering Committee)**
- ⇦ **More collaborat{ing | iums}**
- ⇦ **Generally working together.**

- **Quantum chemists have been leaders at code sharing. However, this often requires that one become an expert at the code to benefit. An especially profitable step would be to codify the inputs. But even then, one can utilize the codes as canned entities only for problems smaller than those of interest. Code sharing is less advanced amongst the materials scientists although growing. (CVS is still virtually unknown!)**

- **One response to the increased complexity of the problems is the Computatinal Materials Sciences Network (http://cmpweb.ameslab.gov/ccms/). Cooperative Research Teams are assembled to work on large scale projects. Current projects are**
  - **Microstructural Evolution based on Fundamental Interfacial Properties.**
  - **Microstructural Effects on the Mechanics of Materials.**
  - **Polymers at Interfaces.**
  - **Magnetic Materials Bridging Basic and Applied Science.**
  - **Excited State Electronic Structure and Response Functions.**

(To the best of my knowledge; electronic notebooks, shared whiteboards, etc. are not in use.)

- **An announced objective of the users group at ORNL is to maximize remote operation of instruments at both the High Flux Isotope Reactor and at the Spallation Neutron Source. This cannot fully supplant the actual presence of the investigator but it can reduce it --- the more complete the telepresence (higher the bandwidth and cleverer the connection), the more successful it will be.**

Initiatives, etc.

◆ *"Flat funding" for the physical sciences in SC isn't. It's much worse than flat.*

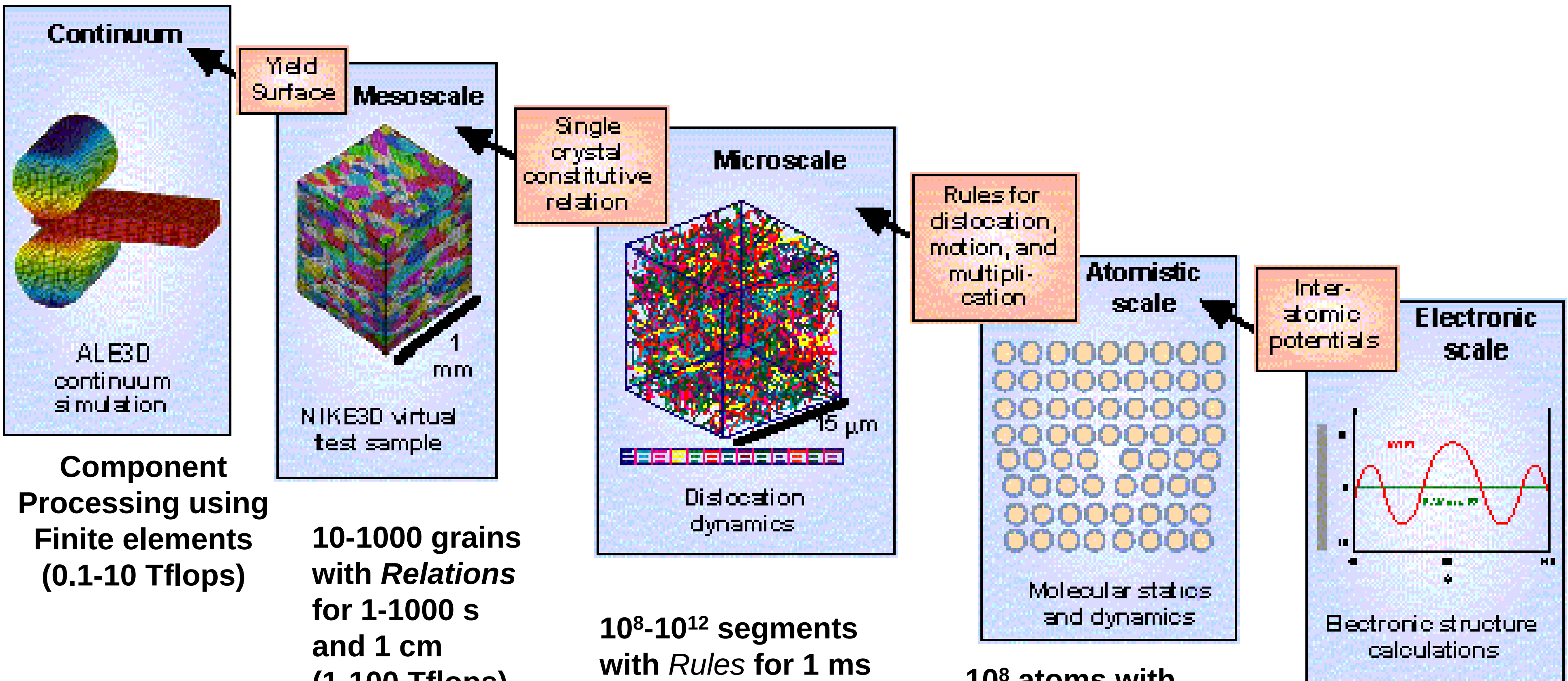
☆ **BUT, there are the initiatives.**

🕐 **While the core program diminishes.**

The Nanoscale Initiative is the most recent *interagency* initiative.

- **The only dedicated funding for computation is $2M for quantum chemistry. Demand is expected to be much higher!**
- **The possibilities are enormously exciting. They have only become realistic because fabrication techniques have become capable of creating things at this scale. We are roughly interested in assemblies consisting of between ~100 and ~10^7 atoms and in the interactions between these entities.**
- **Theorists have long been at this scale. This size is reachable for direct atomic calculations --- the usual extrapolations and other schemes are not needed. This will allow a more direct comparison of theory and experiment. It will also facilitate studies of the holy grail: bridging scales.**
- **Many fundamental and technical advantages arise.**
  - **Many major macroscopic effects cannot develop in particles of this small size: dislocations, for example.**
  - **An appreciable fraction of the atoms are at or near the surface. (I.e., Surface Stabilization can more readily act.)**
  - **The finite size is less than the mean free path and the level spacing due to quantization can be larger than thermal of driving energies. These will shut down various loss mechanisms.**
  - **The smaller size permits charging of the particle to have a much more significant effect.**

Mechanical Properties over Many Length Scales



Component Processing using Finite elements (0.1-10 Tflops)

10-1000 grains with *Relations* for 1-1000 s and 1 cm (1-100 Tflops)
<u>NEW</u>

10^8-10^{12} segments with *Rules* for 1 ms and 100 µm (1-100 Tflops)
<u>NEW</u>

10^8 atoms with *semi-empirical potentials* for 1-10 ns and 100 nm (1-100 Tflops)

~500 electrons with density functional theory (1-100 Tflops)

***<u>The</u>* issue is how to succinctly transfer information from one scale to the next!!**

BES "Computing"

- The increased complexity of the problems like "Correlated Electrons", "Complex Materials", "Microstructure", and "Nanoscience" will push the bifurcation of the community described by Ken Wilson in his appendix to the Lax report: those pushing the envelope will be very hungry for cycles and memory; those doing problems comparable to what has been done will be much happier at home. (Is NERSC like a beam line or like a sausage factory?)

- ***The high end is still important.*** The problems are there to be brought out when the required resources are feasible to acquire.

- ***But so is the "workstation" environment.*** It needs to be enhanced to pull some of the clutter off the high end. The workstation environment now is clustered: DQS, MPI-Lite, M-VIA (Virtual Interface Architecture), MVICH, …

- ***The third category is the dedicated engine.*** Many of the facilities can generate data faster than it can be piped away. On-the-spot processing can help: image enhancement and recognition are important here.

BES: Networking

- ***BES must need enhanced networking resources.*** **The detailed strategy should be hammered out in the ESNET Steering Committee. Help to be forward looking is definitely needed.**

- ***BES operated national facilities (and other data sources) generate communication challenges.***
  - **Data storage and transfer (data is getting more massive)**
  - **Off-line data analysis (or near-line, which is trickier)**
  - **Collaborator efforts (remote presence is a high bandwidth problem)**
  - **Remote operation (more detailed remote presence *<u>and</u>* big "quality of service" issue)**
- ***(Are we going to do wide area parallelism?)***

- ***The data has to get through!***
  - **Fire walls and overly restrictive AUP's ARE isolating us!**
  - **ESNET must interface with VBS to reach the universities.**

- *Work at universities is a critical component of our portfolio. It has remained a constant fraction of the research portfolio for more than a decade, and it will so continue.*

BES: Applied Mathematics

- BES researchers utilize libraries: *advertise and document* as well as create.

- Joint projects need to happen. The biggest roadblock is actually agreeing on the problem! A well defined problem is a very valuable possession.

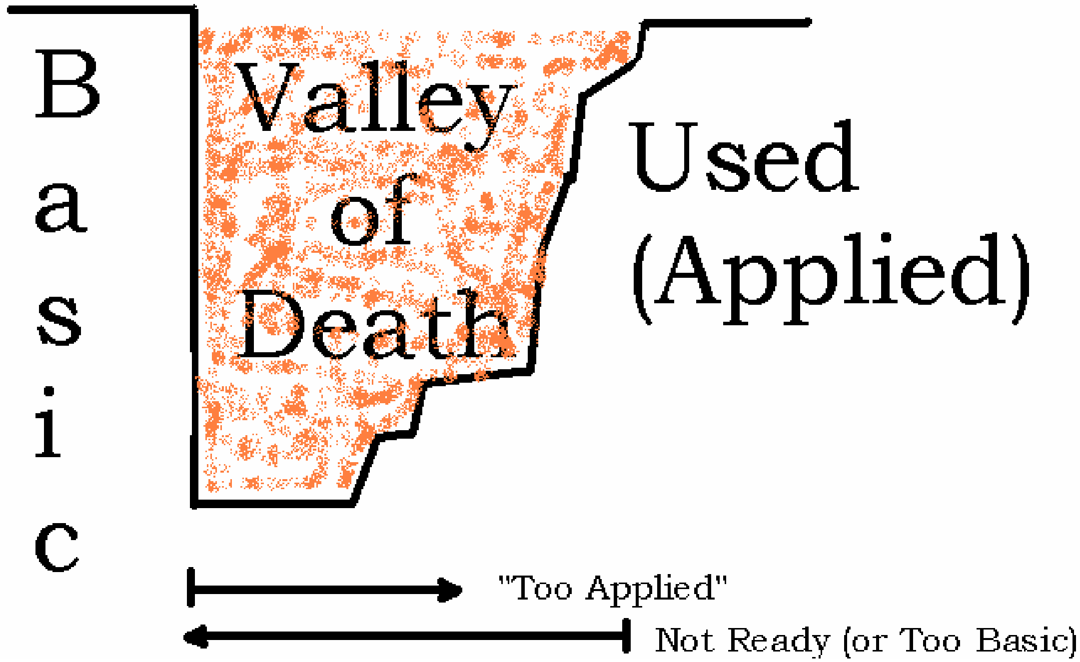
- A very major class of important projects can be understood and organized by a Halloween-consistent invitation:

Bridge the "Valley of Death"

The Valley of Death

Something is developed as a part of a basic effort.

- Idea
- Discovery
- Code
- Data Base



They represent the product of a very fine effort. But to be useful, they need to be further developed. But now they become orphans falling into the "Valley of Death". They are no longer the forefront development and so "not basic" but they are not fully developed ready to be used. Much of the required expertise is not the same as that which generates the basic product.

- Data Base Management
- Neural Nets
- Validation
- Management
- Interfaces

Clearly, bridging the Valley of Death is a highly leveraged endeavor.