Ecce Version 2.1 Release Notes - February 11, 2002

Version 2.1 Patch #1 - May 10, 2002

The intent of this page is to provide information specific to version 2.1 of Ecce. Except as mentioned herein release notes from previous versions of Ecce still apply so please do not consider this as a standalone reference.

Version 2.1 contains major enhancements to the code registration and job monitoring components of Ecce, including the integration of a new electronic structure code, while addressing a number of other problems that have arisen since v2.0 was released. Much of the groundwork for the Linux port, to be released in Ecce version 2.2, was done in version 2.1. Other assorted new features have been added for version 2.1 at the request of users.

The May 10 patch of version 2.1 contains primarily job monitoring reliability improvements including an automatic job monitoring restart capability. Other functionality as described below has been added at the request of users at external sites. Only the application software has been changed for the patch. The data server component as it existed before the patch is compatible with the new application software.

Note: Ecce version 2.1 requires NWChem 4.0 or newer due to changes in the NWChem input and output file formats.

Release Notes for Recent Previous Versions

- Version 2.0 Release Notes - July 17, 2001
- Version 1.5 Release Notes - December 1, 1999

What's New? What's Fixed? What's Broken?

What's New?

1. (May 10 Patch) Job monitoring reliability enhancements
2. (May 10 Patch) eccejobstore core read/process code redesigned
3. (May 10 Patch) Messages from Ecce GUI applications now handled by eccejobstore
4. (May 10 Patch) LSF queue management system
5. (May 10 Patch) Site configurable remote xterm commands
6. (May 10 Patch) Ssh X11 forwarding used for remote xterm commands
7. (May 10 Patch) Exit button added to Gateway
8. (May 10 Patch) Delete Machine Configuration button
9. (May 10 Patch) User preferences directory renamed
10. (May 10 Patch) Customizable remote shell $LD_LIBRARY_PATH
11. (May 10 Patch) Additional launch parameters available for input file post-processing
12. Amica electronic structure code now available

13. Improved error handling in job monitoring
14. Modification date added to Calculation Manager property list
15. Search added to Calculation Manager
16. Support for reorganizing project folders in Calculation Manager
17. Calculation Manager Options menu
18. New features in MD trajectory viewer
19. Basis set library updated
20. Basis Set Tool contraction editor
21. Basis Set Tool advanced mode
22. Calculation run directory path simplified
23. Environment variable for setting location of job monitoring temporary files
24. Remote shell window width determined automatically
25. Code registration has been streamlined
26. Installation process and documentation improved
27. Improved compute server registration documentation
28. Online help bundled with Ecce data server
29. Accurate data server error messages on Ecce startup
30. Remote communication logging for diagnosing problems
31. Support for ssh version 2 and different types of authentication
32. Globus v2.0 beta toolkit integrated
33. EDSIServers data server configuration file renamed DataServers
34. Support for multiple simultaneous users running Ecce from the same UNIX account
35. User interface widget set replaced.

What's Fixed?

1. (May 10 Patch) Molecule Builder periodic table "Square Pyramidal" configuration
2. (May 10 Patch) Potential NWChem molecular orbital parsing bug
3. (May 10 Patch) Builder MD fragment file format
4. (May 10 Patch) Ecce background invocation
5. (May 10 Patch) User temporary directory space renamed
6. (May 10 Patch) Number of open files during job monitoring
7. (May 10 Patch) Remote command buffer size increased
8. Structures added from the Builder Structure Library can be the wrong display style
9. Prepending residue from Structure Library destroys residue structure of system
10. Charge and symmetry values lost on Molecule Builder drag and drop
11. Basis Set Tool ECP orbital basis sets
12. Beryllium misspelling in basis set export scripts
13. Basis Set Tool revision log
14. Basis Set Tool table resizing
15. Job launching using rsh/rcp
16. Basis Set Tool "Find Symmetry" routine
17. Automatic Ecce client machine registration
18. Calculation Manager job launch information for imported jobs

19. Job Launcher memory value lost
20. Machine configuration updated for EMSL MSCF resources
21. Gaussian-98™ Fermi Contact Coupling parsing
22. NWChem *.aoints.* files
23. Opening an enabled code or theory/runtype property causes the Calculation Viewer to crash
24. Drag and drop of moleclue while viewing a normal mode
25. Password field text entry fields
26. $HOME environment variable usage causes Molecule Builder to crash opening MD Toolkit
27. X Window System 75dpi font path on Sun
28. Fatal Ecce errors hang during exit when producing core files.

What's Broken?

1. Calculation Editor details dialogs incompatible with Linux KDE window manager
2. Several online help links in the web browser left-hand frame are invalid
3. Viewing molecular orbitals in the Calculation Viewer while a calculation is running results in coordinates that do not match the orbitals
4. Refresh the Molecule Builder geometry table in order to see the new coordinates that result from using the bond or atom manipulators
5. Molecule Builder sphere radius selection does not work
6. Rings are not displayed in the Molecule Builder and Calculation Viewer
7. Cannot add measures in the Calculation Viewer to a property that is animating
8. Color and Pixmap allocation can fail on SGIs leading to warnings or crashes
9. access_log file for SGI web server does not show proper IP addresses for connections.

What's New?

1. Based on extended testing by the Ecce team, automatic reporting of failures within EMSL, and assistance from users at external sites, several enhancements have been made to job monitoring. These significantly improve the overall reliability of job monitoring especially for jobs that may take several days to run. Most importantly, there is now an automatic restart capability built into the client-side job monitoring. Should monitoring fail for either internal (logic errors in the job monitoring related code leading to crashes or exits) or external (network glitches, resource allocation failures) reasons at any point while a job is running, the job monitoring system will attempt to restart itself. A new lightweight application named eccejobmaster starts, monitors, and restarts as necessary the primary client-side monitoring application, eccejobstore. Any failures in the server side monitoring script, eccejobmonitor, are also covered by this same mechanism as eccejobstore catches eccejobmonitor failures and a restart of eccejobstore will also result in a restart of eccejobmonitor. Besides the new eccejobmaster architecture, the second key to being able to restart job monitoring is what is known as "bookmarking". With every job output property successfully sent by eccejobmonitor and then parsed and stored by eccejobstore, a temporary file is updated with the current read offset in bytes in the job output log file. A restart operation reads this bookmark or offset and sends it to eccejobmonitor causing it to immediately skip to that location in the log file (potentially skipping many megabytes of data) and begin processing again from there. Currently up

to 25 restarts of eccejobstore will be done before eccejobmaster will conclude that a unpreventable failure has occurred and the final state of the calculation is set to monitor error. The Ecce team recognizes the critical nature of reliable job monitoring to the success and usage of our software. If you experience further problems in job monitoring please contact ecce-support@emsl.pnl.gov. Thanks to Dr. Matthew Asplund at Brigham Young University, Provo, Utah for assistance in troubleshooting job monitoring failures.

2. Ongoing concerns about the error handling and robustness of the main read/process loop in eccejobstore, which waits for messages (including properties from eccejobmonitor,) have led to a second major redesign. The new implementation uses an X Window System Intrinsics (Xt library) loop allowing the multiplexing of input from external sources including eccejobmonitor messages, inter-process communication with other Ecce GUI applications, signal handling, and inactivity timeout handling. The result has been fewer internal job monitoring failures triggering automatic restarts. The code to establish a socket based connection to the compute server has also been redesigned to handle errors and automatic changeover to a remote shell connection (ssh, telnet, etc.) more reliably.

3. The redesign of the main read/process loop in eccejobstore to be X Window System based also facilitated the integration of inter-process messaging from other Ecce GUI applications. Most importantly, eccejobstore handles renaming calculations/projects through the Calculation Manager using this capability. Previously in v2.0 and the initial v2.1 release, renaming a calculation would result in job monitoring failing, but the state of the calculation would be left as submitted or running--whatever the state was when the rename occurred. This was caused by eccejobstore not knowing the updated location of the calculation on the data server (changed by the rename) and thus losing all ability to reference/update it. Other messages pertinent to the operation of eccejobstore, such as deleting the calculation being monitored, are also handled.

4. The Load Sharing Facility, LSF, queue management system is now supported through Ecce. The machine registration GUI available to Ecce site administrators through the "ecce -admin" command, allows LSF to be selected and specific LSF queue attributes to be set when registering compute servers. The $ECCE_HOME/siteconfig/QueueManagers file lists the specific commands used by Ecce to submit and monitor LSF jobs as it does for all other queue management systems supported by Ecce. Similarly, the $ECCE_HOME/siteconfig/submit.site file lists the LSF directives needed in the job submit script. It is possible that a site will need to customize the information in either or both of these files so the Ecce integration of LSF conforms with the usage of LSF at their site. Thanks go to Dr. Jyh-Shyong Ho at the National Center for High-Performance Computing in Hsinchu, Taiwan for his work in integrating LSF.

5. Sites can now configure custom commands for invoking remote xterm shells. Remote xterms are used in various places within Ecce such as the Shell in Calculation Directory... and tail -f on Output File... menu items in the Calculation Manager, the Open Remote Shell... button on the Machine Configuration dialog, and the Open Shell menu item in the Machine Browser. A custom xterm command can be used when the standard mechanism for setting the $DISPLAY environment variable to the host machine as required by telnet and rsh (and potentially other attributes such as

$XAUTHORITY) does not allow standard Ecce remote shells running the xterm command to be created. For example, sites using the KTH (Royal Institute of Technology, Stockholm, Sweden) version of Kerberos can use the Kerberized versions of rsh and rcp to launch and monitor jobs in Ecce. However, the rxterm script in KTH Kerberos must be used to create remote xterm shells because of the customizations done in rxterm to $DISPLAY and $XAUTHORITY for Kerberos X Windows protocol forwarding. The documentation in $ECCE_HOME/siteconfig/remote_shell.site describes in detail how to use the custom remote xterm command feature. All remote xterm invocations in Ecce with the exception of those using telnet, whether a custom command or a standard remote shell such as ssh and rsh, now directly invoke the xterm command from the Ecce application client machine rather than first establishing a remote shell connection and then issuing a standard xterm command under this connection. Thanks go to Dr. Olav Vahtras of the Royal Institute of Technology in Stockholm, Sweden, for his assistance in adding support for custom remote xterm commands.

6. Invocations of remote xterms under ssh now use the X11 forwarding built into ssh. Previously Ecce required the same $DISPLAY and xhost command configuration as required for telnet and rsh and intentionally bypassed using ssh X11 forwarding. This feature simplifies what users need to do to create remote xterms especially since ssh is the most widely used remote shell with Ecce. The error messages issued by Ecce when a remote xterm command failed, explaining how to set $DISPLAY and run the xhost command often confuse users who didn't understand the standard security mechanism built into the X Window System. The situation is further exacerbated by not running Ecce GUIs locally since Ecce is able to run the xhost command on behalf of users when they are running Ecce on the workstation they are displaying to.

7. In order to support window managers not having an option to close a window, such as TWM, an explicit exit button has been added to the Ecce Gateway. By clicking anywhere on the Ecce "wave" icon or over the "exit" label below it, either an exit confirmation dialog will appear or Ecce will be immediately closed, based on the confirmation dialog preference set by the user. If the Ecce "wave" is currently animating, meaning the user is trying to start another application, then clicking on the "wave" icon or the exit label will stop the icon animation but not exit from Ecce. This feature allows users to stop a "runaway animation" if the application fails to come up for any reason. In order to add the new exit button the "windows" button previously under the Ecce "wave" has been moved to the bottom right corner of the Gateway. Other buttons in this right-side panel have been shuffled and the legend button has been removed completely. The legend displaying the symbols and colors for the different run states is now available only from within the Calculation Manager.

8. A "Delete" button has been added to the Machine Configuration dialog available from the Job Launcher and Machine Browser applications. This allows the currently displayed machine configuration to be removed from Ecce if, for example, a user no longer has an account on a registered Ecce machine. Previously there was no way through the Ecce GUI to delete a machine configuration.

9. The directory where Ecce preference files are kept has been renamed from ~/.ECCE.v2 to ~/.ECCE. Starting ecce will automatically perform this rename operation if the user has an existing ~/.ECCE.v2 directory. Prior to v2.0 the preferences directory did not have the .v2 extension. However, to allow concurrent usage of v1.5 and v2.0 of Ecce, we supported both directories. Since v1.5 is no longer a factor, the preferences directory has reverted to the original name.

10. The value of the $LD_LIBRARY_PATH environment variable can now be customized for remote shell connections to a machine. This feature is analogous to customizing $PATH and was added to support executables such as queue management system job submission and query commands requiring shared libraries in non-standard locations. To specify LD_LIBRARY_PATH the $ECCE_HOME/siteconfig/CONFIG.* file must be edited for the desired machine, where "*" is the name of the machine. Machines registered by users rather than the Ecce site administrator have CONFIG.* files in the directory "~/.ECCE". A line beginning with "libPath:" should be added near the top of the file along with any existing lines specifying paths to computational codes, etc. Multiple directory entries must be separated by colons as they normally are in the $LD_LIBRARY_PATH variable. For example, to add the directories /usr/local/lib and /share/apps/lib to $LD_LIBRARY_PATH for a machine, the following line would be added to the appropriate CONFIG.* file:

    libPath: /usr/local/lib:/share/apps/lib

    This feature is not supported through the Machine Registration application because it is rarely needed and would thus confuse some site administrators and users. However, it is necessary for site administrators to use this feature should the queue management system being used on a compute server being registered in Ecce depend on shared libraries not guaranteed to be in $LD_LIBRARY_PATH for all users.

11. New parameters are available for post-processing input files based on user job launch configuration such as the total number of processors and nodes. These parameters and their usage have been documented in the existing launch post-processing scripts, $ECCE_HOME/scripts/parsers/nwchem.launchpp and $ECCE_HOME/scripts/parsers/gaussian-98.launchpp.

12. The Amica code for performing high accuracy single-point energy calculations has been registered in Ecce and is available. This code is designed to support very high levels of theory for small molecular systems. Further information on Amica can be found at http://www.hec.utah.edu/gdanitz/amica/manual.pdf. Note that unlike the Gaussian-98™ and NWChem codes, Amica will not run successfully just using the default values of the setup parameters, so it is important for prospective users to familiarize themselves with the official Amica documentation. The Amica documentation is also available through Ecce by selecting the "Help" button at the bottom of the Amica theory details dialog available from the Calculation Editor when setting up an Amica calculation. An appendix in the documentation gives further information specific to using Ecce to run Amica calculations.

13. Several improvements were added to Ecce's job monitoring facility to improve reliability and support more effective sleuthing of intermittent communications problems. Substantial effort was put into detecting errors and reporting them back to the user. There is a new job state called monitor error which indicates that an error occurred in the job monitoring. This is distinct from the incomplete state which indicates that the computational code terminated with a status other than 0. When a calculation has a state of monitor error, the View Run Log menu option in the Calculation Manager can be used to identify the problem. In addition, within EMSL all such errors will be automatically sent to the Ecce bug tracking system. Most likely you will not be contacted when this happens unless the Ecce team needs extra information in order to diagnose the cause of the failure. With v2.1 verbose logging of all data sent and processed during job monitoring has been enabled by default. This will assist the Ecce team when errors are found with real-world calculations which are often much longer and more data intensive than calculations run by the Ecce development team during testing.

    The Run Mgmt menu in the Calculation Manager now has an item Check Job Monitoring which will check if the job monitoring processes for any calculations in the currently displayed folder have failed. This detects errors that could not be caught by the monitoring process itself. Use this option if calculations seem to be hung in the submitted or running states.

    The default type of communication between the compute server and the Ecce client workstation for job monitoring has been switched from sending data over a remote shell (ssh, telnet, etc.) to creating a socket for this data. This default only applies to jobs launched within the local domain. Jobs launched outside the local domain automatically use the remote shell for sending data as was the case for all jobs prior to v2.1. The distinction between local and remote jobs was made because firewalls often prevent the use of arbitrary socket ports. Sockets have proven to be less susceptible to network related "hiccups" that can terminate Ecce job monitoring processes running over remote shells such as ssh especially with jobs running many hours or days. The environment variable ECCE_JOB_COMMS can be used by a site or user to override the default of using sockets inside and remote shells outside the local domain. The three possible values for ECCE_JOB_COMMS are "socketlocal" for automatic switching (default), "socket" for forcing the use of sockets for all jobs, and "stdio" for forcing the use of remote shells (the compute server monitoring process sends its data to standard I/O, hence the name). If an attempt to create a socket fails, such as for a firewall between the Ecce application software workstation and the compute server within the local domain, job monitoring will automatically switch over to the remote shell for sending data for that particular calculation. The View Run Log option from within he Calculation Manager logs this automatic switching. Based on the compute server configuration at a site, the Ecce site administrator may wish to switch the ECCE_JOB_COMMS variable in $ECCE_HOME/scripts/ecce_env to "stdio" if socket communications cannot normally be used. It is important to keep in mind, however, that sending data over the remote shell has proven somewhat less reliable and occasional job monitoring glitches may result from this change.

    The compute server side job monitoring script, eccejobmonitor, has had several fixes. All jobs submitted to queue management systems now have a timeout of 3 minutes for recognizing that the job is in the queue system (in any state--idle, running, etc.). Prior to v2.1 the timeout (at which time the

job would go to the incomplete state) was 90 seconds for all but NQE jobs. The timeout has also been bumped up from 15 to 60 seconds waiting for the file .ecce.status to exist after a job is recognized as completed by the queue management system. This file contains the exit status value from the chemistry code and is used to set the state to completed or incomplete. Both of these changes to timeout values were made due to problems reported by external sites.

A feature in v2.0 and prior releases of Ecce which attempted to flush the output file by closing and reopening it has been removed for v2.1. It was found that the function to reposition a file pointer to a specific offset in the Linux version of perl does not work. Thus when running jobs to Linux compute servers there would often be duplicate sets of properties parsed for all but very short jobs (where the file reopen was never performed). The original reason this feature was added was to force output file flushes, and thus pick up new properties more quickly, when jobs were run to file systems slow to perform flushes without intervention. The loss of this benefit is not nearly as serious as the problem it introduced (due to the underlying perl deficiency) and in the past few years the underlying problem seems to have been mitigated by better file system behavior on the more finicky massively parallel machines. Finally, the SIGPIPE signal (write on UNIX pipe failed) in eccejobmonitor has been changed from a fatal job monitoring error to a warning since it appears to occur on certain machines for events that have no direct bearing on job monitoring.

The Ecce application client-side job monitoring process, eccejobstore, has also been improved. Most importantly, the primary loop which waits for data from eccejobmonitor has been reworked for reliability. Specific conditions that would have been detected as job monitoring failures in v2.0 and prior releases are now handled as non-fatal conditions. Additionally, conditions properly recognized as errors are now more tolerant to sporadic glitches in network communication through the use of a retry mechanism where a failure must occur a number of times before being treated as a fatal job monitoring error. This retry mechanism will not allow actual loss or corruption of calculation data to occur but just improves low-level communication protocol management.

The inter-process messaging done by eccejobstore to notify the Calculation Viewer of new property data now allows data to be passed both in the message itself and indirectly by passing a URL referencing the data on the Ecce DAV server. For large properties such as molecular orbitals, which can be megabytes of data, the URL reference is a much more reliable mechanism for passing data. Currently any property that is greater than 2048 bytes of data is not passed directly in the inter-process message but via a URL reference.

14. The Calculation Manager Preferences dialog now includes the modification date as a property that can be viewed in the table. The modification date does not reflect changes to meta data such as job state but is otherwise accurate. Sorting on the modification date rather than the creation date works around the problem introduced with the EMSL Ecce data server updgrade in January, 2002 where the creation date based order of calculations was lost. The creation date is an internal file system related property that cannot be changed or preserved by user commands or utilities while modification dates can be preserved during copy and move commands.

15. The Calculation Manager now includes a prototype search capability which can be accessed from the toolbar (the binoculars icon) or Calculation menu. The search will find any projects/calculations (URLs) that contain the specified search string. Searches can be interrupted by hitting the stop icon. This is a preliminary version of the search and can be slow if performed over many projects and calculations.

16. Initial support for reorganizing project folders has been added to the Calculation Manager. Folders within the tree hierarchy can be moved through the standard Motif drag and drop facilities (mouse button 2). Moving folders across authentication realms (between two different user folders) is not currently supported. The work around is a two step process: first move the desired object to the top level share directory, then move from the share area to the desired destination. Reliable support for moving calculations (from table to tree) as opposed to project folders is not yet available.

17. The Calculation Manager now includes an Options menu. The menu includes items for controlling how the tree hierarchy is sorted along with some options previously found within the Preferences dialog.

18. The MD trajectory viewer has been updated so that it will now read standard XYZ format trajectory files. A new toggle has also been added that allows the user to specify whether bonds are to be recomputed at each timestep. This feature was added primarily for viewing QM/MD simulations where bonds may be formed and broken during the course of the simulation. Note that recomputing bonds at every timestep will significantly slow down the display if the system contains a large number of atoms.

19. The Ecce basis set library has been updated to reflect the latest basis sets from the EMSL Gaussian Basis Set Web Order Form, as of February 2, 2002. For a list of the latest changes, click here.

20. The Contraction Editor, a new feature of the Basis Set Tool, was added in support of the new Amica code. It may be accessed by pressing the Edit Basis Set...button located just above the configuration table. This editor allows you to uncontract or delete specific contractions for each element/basis set/angular momentum combination.

    The Contraction Editor also allows you to add any number of tight or diffuse functions for each element/angular momentum combination. The tight or diffuse functions will be extrapolated automatically using the following algorithms:

    Tight Functions
    The largest two exponents (Emax, Emax-1) for the given element/angular momentum combination are identified, where Emax is the largest and Emax-1 is the second largest. Their ratio is then computed (ratio = Emax/Emax-1). If n is the number of tight functions to add, and i is the current tight function, then for i = 1 to i = n, Ei = ratio^i * Emax. Tight Function i is then added to the configuration with exponent Ei and coefficient 1.0. [Note: prior to adding a tight function, if only one exponent is available, the tight function will not be calculated.]

Diffuse Functions

Appended diffuse functions are calculated similarly to tight functions except for two differences. Instead of picking the largest two exponents to compute the ratio, the smallest two exponents are selected (ratio = Emin/Emin-1), where Emin is the smallest exponent, and Emin-1 is the second smallest exponent.

If only one exponent is present for the given angular momentum (i.e., shell), then the extrapolated diffuse exponent is calculated from the extrapolated diffuse exponents of the two previous shells. For example, if trying to add a diffuse function for the D shell, but only one exponent was available, then extrapolated exponents would first be determined for the P and S shells, using the same algorithm described above. The extrapolated diffuse exponent for the D shell would be then calculated as follows:

extrapolated_D_exponent = extrapolated_P_exponent ^ 2 / extrapolated_S_exponent.

The actual ratio would then be determined as:

ratio = extrapolated_D_exponent / Emin.

Note: for P or S shells, this extrapolation method would fail because there aren't two lower shells with which to calculate the ratio. In these cases, a diffuse function will not be appended.

If you do not wish to use these calculated values, the extrapolated diffuse or tight exponents may be changed by modifying the input file via the Final Edit button in the Calculation Editor. In the future we hope to add a more convenient way to edit these values.

21. In support of the new Amica code, an Advanced Mode option was added to the Basis Set Tool. This option enables the inclusion of multiple orbital basis sets. Advanced Mode may be activated by selecting the Combine multiple basis sets toggle under the Advanced menu. Note that by default this option is off, and once activated, it will apply to all codes unless explicitly deactivated by the user.

22. The directory path to the calculation run directory on the compute server automatically created by Ecce when launching a calculation has been considerably shortened. It no longer contains the first three directories "/Ecce/users/$USER" which are superfluous. Calculation run directories now start with the top level project name created by the user. If a calculation is run outside the user's own area on the web server, such as /Ecce/share, the run directory will not be shortened to this extent, insuring a unique path on the compute server.

    In a related change, calculation run directories may now be specified using tilde, "~", notation. Previously this character was treated as a literal tilde character instead of interpreted as the user's home directory.

23. To alleviate problems arising from lack of disk space under /tmp, a new Ecce environment variable has been defined, $ECCE_TMPDIR. When set, the value of this variable is used as the directory path under which to create the temporary files needed for job monitoring. These files include the input and output property parse script files and thus can contain megabytes of data when dealing with large properties such as molecular orbitals. If this variable is not set (the default), then these files are created in a uniquely named directory created for each calculation under /tmp/ecce.

24. The width of the remote xterm shell created by doing a "shell in Calculation Run Directory..." or "tail -f on Output File..." operation in the Calculation Manager is now based on line length of the file being viewed. If all lines are 80 characters or less then the shell will have a width of 80. Otherwise the width will be 132 characters. Certain files, such as the Amica code output file, are formatted to be viewed in a 132 column text editor so this feature properly displays those files.

25. The code registration process has been substantially streamlined. The many document files required to register a code on the calculation setup side have been consolidated to a single XML format file. The output parsing remains largely the same. Documentation on the registration process is available from the main Ecce web page and here. Much of this work was motivated by the integration of Amica as the first code primarily registered outside EMSL. Most new features though will benefit the integration of all future chemistry codes.

    The Ecce application software distribution now contains all the components necessary to create user interfaces using the public domain Amulet toolkit as first introduced in v2.0 and described in the release notes here. The distribution includes examples of details dialog user interface code as well as the Amulet include files and libraries needed to compile interfaces. This code is found in the new $ECCE_HOME/codereg directory of the application software distribution.

    New objects in the Ecce Amulet library of interface objects support entry of free format text named "text_input" and an object that displays other objects in a scrolling window named "scroll_group". The "scroll_group" object also allows entire details dialog windows to be resized in height with small "+" and "-" buttons in the top right corner of the dialog. This feature is currently used for the Amica theory details dialog. Traditional window resize done through the window manager is not supported by Amulet so this was the most efficient solution to achieve the same result. All type-in fields (text_input, integer_input, float_input, and exponent_input) now have an Amulet slot named "WIDTH" which allows the number of characters displayed in the field to be changed from the default.

    The Calculation Editor supports a toggle button named "Use As Irreducible Fragment In Input". Whether this toggle is displayed (shown under the "Charge" field) is controlled by the code registration XML file. Currently it is applicable only to Amica. When the toggle is set by the user then the chemical system shown is taken to represent an irreducible fragment meaning the full chemical system (based on the symmetry specified) is used to determine the number of atoms and electrons in the calculation, but only the irreducible fragment is printed in the input file. If the toggle is not set then the object shown is treated as the complete chemical system to use for the calculation regardless of any specified symmetry. Because this can be a difficult concept, the small chemical

system viewer in the Calculation Editor updates based on the toggle setting to show what would currently be used by the computational code as the chemical system for the calculation.

The Calculation Editor now allows the summary fields on the main window below the "Theory Details..." and "Runtype Details..." buttons to be registered along with the code. Previous versions of Ecce had hardwired summary fields built into the Calculation Editor application code. These summary fields are specified in the code registration XML file using the unique keys for objects first given in the Amulet user interface code. Another v2.1 feature allows individual summary fields to be hidden when the user has not changed the value from the default and shown only when they are changed from the default. This feature is used extensively in Amica where nearly all theory details dialog objects are specified as summary fields but only the ones explicitly changed by the user are displayed on the main Calculation Editor window. The NWChem and Gaussian-98™ Calculation Editors still use "static" summary fields which are always displayed regardless of whether the user has changed anything. Another feature for summary fields displays only the labels for fields that are toggles where the label is only shown when the toggle is set (on). This feature is not used by default so normally the label is shown along with the current value of "true" or "false".

The list of theories shown in the "Theory" menu in the Calculation Editor is now condensed to those theories that are supported by the current code. With previous versions of Ecce all theories supported by any Ecce registered code were displayed which created a very long list in addition to being confusing to users. Additionally, it is now possible to specify only theory names rather than requiring theory category and name pairs. In the case of theory names only, the Calculation Editor "Theory" menu behaves like a regular Motif option menu instead of a two-level hierarchical menu. Finally, it is no longer mandatory to have both a theory and runtype details dialog at all. Any combination of the two dialogs, including neither, is allowed. When there is no details dialog the corresponding button in the Calculation Editor is hidden. All of the information characterizing theories and runtypes for a code is now specified in the XML registration file.

26. The installation scripts and on-line installation and administration documentation for both the application software and the data server have been significantly improved for v2.1. Both the application software and data server installation scripts support merging previous versions of Ecce. For the application software this allows the machine registration information and data server configuration from a previous release to be preserved (those files in $ECCE_HOME/siteconfig). For the data server this allows all existing user calculation data and structures created in the Molecule Builder Structure Library to be preserved along with data server user names and passwords. This merge feature will also be used to update the format of data stored with the application software or on the data server when they are changed to add new functionality. Thus, it is a "schema upgrade" feature in addition to saving the Ecce site administrator from manually copying selected data from a previous version of Ecce into the new installation.

    The on-line documentation is more consistent end-to-end and some explanations have been simplified and/or clarified based on repeated reviews and installation walkthroughs. The Ecce Installation and Administration documentation is available from the main Ecce web page and here.

27. The process for registering compute servers as well as new queuing systems has been simplified and the documentation updated. All configuration files associated with the registration process now reside in the $ECCE_HOME/siteconfig directory. The documentation can be accessed from the main Ecce web page and here.

    The Ecce site administrator can now create siteconfig directory configuration files (files named CONFIG.*) based on the hardware vendor and/or hardware model of machines. This allows specifying configuration items such as paths to chemistry codes based on the type of hardware platform instead of requiring a CONFIG.* file for every single machine. Machines matching this vendor or model can all share a single CONFIG.* file. For example, within EMSL there are CONFIG.SUN, CONFIG.SGI, and CONFIG.SGI.O2 files since paths to codes are the same for all Sun workstations, SGI workstations other than the O2 model, and the SGI O2s, respectively. There are no CONFIG.* files for the individual Sun and SGI workstations as would be required with Ecce v2.0. These CONFIG.* files must be created by hand rather than with the "ecce -admin" machine registration application.

28. The online Help within Ecce is now bundled with the Ecce data server distribution rather than being served from the EMSL public web server. This significantly shortens access time for those sites who are physically quite some distance from EMSL. Additionally, it allows the help to be viewed by those sites without external internet access due to firewall restrictions or lack of a network. By upgrading to the v2.1 data server, references to Ecce help web pages in the v2.1 application software will automatically point to the Ecce Apache/mod_dav data server.

29. Error messages related to problems with the data server detected when Ecce is first started have been improved. Ecce now accurately indicates when:

    - The data server httpd daemon is not running
    - The data server has not been properly configured as a DAV server
    - Specific libraries and documents on the data server cannot be found such as the Gaussian basis set library
    - The application software has not been correctly installed defining a data server that should be used

30. Remote shell communication can now be logged to the shell window where Ecce has been started for diagnosing problems. Issue the command "setenv ECCE_RCOM_LOGMODE true" before starting Ecce to enable this logging. An "unsetenv ECCE_RCOM_LOGMODE" will turn off remote communication logging. Any passwords sent to remote shell and copy commands are elided from the output. Remote communication logging can produce a significant amount of output. However, it is an extremely valuable technique for tracking down problems where the error messages issued by the applications for remote operations don't provide enough information about what went wrong.

31. Version 2 of ssh can now be used as a remote communications shell. Only openssh has been tested. The same "ssh" remote shell selection in the Machine Configuration dialog supports both version 1 and version 2 of ssh. Ecce v2.1 does not try to guarantee that password based authentication will be used as was the case with Ecce v2.0, but relies on the sshd server daemon and the ssh client to

negotiate the type of authentication. It is possible that Ecce will not recognize the protocol used by a particular type of authentication that has not been encountered within EMSL. Please contact ecce-support@emsl.pnl.gov if this is the case (the ECCE_RCOM_LOGMODE environment variable can be used to diagnose the problem). There is a variation with how Ecce v2.1 uses ssh, named sshpass in the Machine Configuration dialog and Machine Registration application, which tries to insure that password authentication is used. This variation is equivalent to the normal ssh authentication used with Ecce v2.0 except it also supports ssh v2. Also, as documented in the release notes for v2.0, the $ECCE_HOME/siteconfig/remote_shells.site file contains documentation on configuring remote shell usage at a site.

32. Globus v2.0, from Argonne National Laboratory, has been integrated for launching jobs within Ecce. This was done with an alpha version of Globus v2.0 for demonstration at the IEEE/ACM Supercomputing 2001 conference in November 2001. It has not been tested with the currently available Globus v2.0 which is referred to as a beta release. For this reason we recommend that any sites wishing to use Globus for launching jobs within Ecce contact ecce-support@emsl.pnl.gov prior to installing Ecce so that we may provide further guidance. Use of Globus within Ecce also requires that a special patch to ssh be downloaded and built to support Globus GSI authentication. It is the responsibility of any site wishing to use Globus to install both the Globus v2.0 Beta Toolkit clients and the patched version of ssh on the local machine running Ecce. The Globus web site has the install directions and software download for the ssh patch. Additionally, the Globus server side daemons and a patched version of sshd must be available on the compute server.

33. The $ECCE_HOME/siteconfig/EDSIServers file in v2.0 has been renamed to $ECCE_HOME/siteconfig/DataServers for v2.1. This file is maintained by the Ecce site administrator and contains the URLs to the data servers to be used with an application software installation. When upgrading to the v2.1 data server following the installation documentation, this change will automatically be made.

34. A minor new feature that helps to support group Ecce workshops where all participants may use the same UNIX account has been added. Previously if the same account was used for more than one person running Ecce applications at the same time, the passphrase window would come up each time applications were started instead of only when the gateway was started. Now the passphrase will only be asked for once regardless of how many people are currently running Ecce from the same account.

35. All Ecce user interface widgets were reduced to a single 3rd party widget set. This reduces development licensing costs, increases flexibility in interfaces, and will simplify the porting of Ecce to Linux. The new widget set replaces many of the “convenience widgets” that give Ecce interfaces some of the MS Windows look on top of Motif including tabbed dialogs, iconic push buttons with selection state, and drop down combo boxes. The most intensive widget conversion was for the project folder outline view in the Calculation Manager and Molecule Builder Structure Library. In addition to porting over to a single 3rd party widget set, one widget (for evenly spacing objects such as OK/Cancel/Help buttons at the bottom of a dialog) did not have an analog in the new product so a public domain X Window widget was found and integrated.

What's Fixed?

1. The configuration or shape known as "square pyramidal" in the palette shown in the periodic table of shapes to use for building structures has been fixed allowing it to actually be used to build molecules. Previously it would not be added into the small palette of shapes shown in the Molecule Builder main window when it was selected in the periodic table window (available from the "more..." button on the Molecule Builder).

2. A bug in the perl parse script used by Ecce to parse NWChem molecular orbitals during job monitoring previously could lead to a job monitoring failure as well as a failure to parse the molecular orbitals. In rare circumstances, observed to this point with unusually large molecular orbital data, an internal FORTRAN compiler error on IBM AIX, systems if not others, could generate invalid data. A repeating data specification, e.g. 10*0.0, with an invalid count specification, e.g. -0.0*0.0, sometimes can be generated. This would cause the Ecce perl parsing script to go into an infinite loop eating up disk space until there was none left. Although the underlying problem is apparently in the FORTRAN compiler, Ecce has been fixed so that it treats any invalid repeating count specifications as a count of 0.

3. The output format used to generate NWChem molecular dynamics fragment files has been fixed to correspond with the format expected by the NWChem MD module.

4. Certain operations related to running the "which" shell command from within Ecce would fail when Ecce had been invoked in the background from the command line with "ecce &". The UNIX "which" command does not execute properly when invoked in the background so the implementation was switched to perform the same underlying functionality without using the "which" command. Ecce now runs properly when started in the background.

5. The name of the temporary directory created for each user have been changed to fix a potential directory ownership bug introduced with version 2.1. Previously the first person running Ecce on a machine would create a directory named /tmp/ecce used to hold temporary files for job monitoring. If the write permissions on this new directory, based on the umask of the creator, prevented other users from writing there, Ecce would not be usable for others without manually changing the directory permissions. Now each user creates their own unique directory under /tmp named /tmp/ecce_$USER on each machine where they start Ecce. This directory contains all the temporary files used for job monitoring and various other purposes while Ecce is running. A uniquely named subdirectory under /tmp/ecce_$USER, starting with the name of the calculation, is created for each running job. The files in this directory can be useful for diagnosing problems as well as being used during job monitoring by Ecce.

6. A long standing bug in inter-process communication code related to opening but never closing files was found and resolved. This problem only manifested itself with the eccejobstore job monitoring application due to the volume of messaging it performs compared to other Ecce GUI applications. It was possible for the number of files currently open to exceed the resource limit resulting in eccejobstore crashing for jobs running more than a few hours.

7. A mechanism was found through the public domain 3rd party library used by Ecce to wrap remote communications, libexpect, to increase the size of the internal buffer used to hold the results of remote commands performed by Ecce. The default size is 2,000 characters which easily overflows on commands like "ps" on a busy machine. The buffer size has been increased to 50,000 characters which under any normal circumstances should hold the output for any command issued by Ecce.

8. Structures added from the Molecule Builder Structure Library are inserted into the visualization workspace using the correct Molecule Builder display style set in the Edit, Display All menu option.

9. Prepending residues by attaching a residue from the Structure Library to the front end of amino acid or nucleotide chain would result in incorrect assignment of the atoms to their corresponding residue and incorrect assignment of internal data pointers. Additional editing of this structure would eventually result in a builder crash. This bug has been fixed.

10. The charge and symmetry point group for a chemical system is now preserved when using a drag and drop operation between the Molecule Builder and the Calculation Editor. Drag and drop does an implicit export of the current chemical system to the standard Ecce chemical system format, MVM, which is passed to the application where the chemical system is dropped. The charge and symmetry group attributes were not part of the MVM format until Ecce v2.1. This also means a chemical system exported as an MVM file by the user has charge and symmetry group attributes so they will be restored when importing back into the Molecule Builder.

11. Exponents and coefficients for some ECP orbital basis sets were being written incorrectly to the input file. This problem has been corrected and verified against data from the EMSL Gaussian Basis Set Web Order Form.

12. The element Beryllium had been misspelled as "Berylllium" in all the scripts that formatted basis sets for codes. Therefore any calculation done prior to v2.1 with Beryllium would be suspect for errors with the basis set specification in the input file.

13. The basis set library revision log was not accessible from the Basis Set Tool. This has been corrected.

14. The bottom half of the Basis Set Tool (which contains the configuration table) can now be enlarged. This feature was added so that configuration tables containing a large number of elements (five or greater) can be viewed without scrolling.

15. Job launching using rsh as the remote shell was failing in v2.0 because it tried to use scp instead of rcp as the corresponding remote copy command. Ecce v2.1 correctly uses rcp with rsh.

16. The “Find Symmetry” routine has incorporated some improvements that make it more robust. Previously, the routine would produce distortions of the structure if used on C60 or cubane. The “Find Symmetry” routine still returns C1 if used on C60 but does not distort the geometry and will correctly identify the symmetry group of cubane. These changes reflect changes that have been made in the automatic symmetry recognition module in NWChem.

17. The feature which automatically registers the local machine where the application software is being run, has been moved from the Job Launcher to the Calculation Manager. This eliminates a “catch-22” situation for those users who want to import an existing calculation into Ecce without ever first starting the Job Launcher. Importing a calculation must be done on a machine that has been registered due code sharing between importing and job monitoring performed on calculations launched in Ecce.

18. Prior to v2.1, when importing a calculation run outside of Ecce using the Calculation Manager, certain job launch attributes such as the number of nodes and number of processors, were left uninitialized and displayed as random integer values such as 2 billion. These are now set to zero so these random values are no longer displayed in the Calculation Manager. Since these values are not parsed from the calculation output file this seemed like the most reasonable solution.

19. The Job Launcher memory field was not being saved properly with a calculation when the default was overridden. The value would always restore to 100 megabytes. It is now saved and restored properly.

20. Machine registration information such as number of processors, job time limits, and paths to codes for the EMSL MSCF parallel machines including mpp1, jupiter, colony (pier.emsl.pnl.gov), ecs1, and nwtest, has been updated.

21. For parsing output from Gaussian-98™ calculations, the Fermi Contact Coupling is now parsed properly. The format of this data in the output file changed from Gaussian-94™ to Gaussian-98™ but Ecce had not been updated. A couple bugs in the Calculation Editor when handling Gaussian-98™ semi-empirical calculations have also been fixed. One bug in particular would result in the input file always being overwritten when the Calculation Editor was brought up overwriting any “Final Edit” changes.

22. The NWChem *.aoints.* files are now deleted by Ecce when a job completes. These temporary files can eat up gigabytes of disk space and are usually left by NWChem when jobs exit with other than a zero status value. Ecce only deletes these files when no explicit “scratch directory” is specified for a calculation in the Job Launcher. When a scratch directory is given then it is assumed the user has specified a type of disk that is automatically cleaned on a regular schedule. If this is not the case then it is the user’s responsibility to manage the *.aoints.* files and Ecce does nothing special in addition to normal NWChem behavior.

23. The Calculation Viewer no longer has a “drop-down” type code or theory/runtype property, displaying the setting in the property title bar. This fixes the crash in previous versions.

24. Using the drag site icon in the Calculation Viewer normal mode property window, users can work with a molecule at any time step in the Molecule Builder.

25. Behavior when text is entered in password fields, such as the initial login passphrase window that comes up when starting Ecce and the password field in the Machine Configuration dialog, has been improved. In previous versions the value would be automatically cleared when you moved the mouse pointer out of the window and then back in again. Plus it was possible to corrupt the value in the field and eventually crash applications by selecting multiple characters or the whole value and deleting or by deleting characters from the middle of the value instead of the end. Now these fields only allow characters to be added or deleted from the end of the text, one at a time. Also the value isn't automatically cleared by leaving and reentering the window.

26. Version 2.0 of Ecce had a small number of instances in the code that used a non-standard environment variable set within EMSL, $HOME. For sites outside EMSL where this is not set, some applications would fail including crashes. Most notably, the Molecule Builder MD Toolkit would crash immediately when invoked due to this error. Ecce v2.1 sets its own environment variable, named ECCE_USERHOME, which points to the user's home directory.

27. The X Window system path to 75dpi font path has been fixed to allow the fonts to be located either in the /usr/lib/X11/fonts/75dpi or /usr/openwin/lib/X11/fonts/75dpi (standard Sun location) directory. An error in the v2.0 Ecce startup script only recognized them under the first location. Fonts used in Ecce applications such as Helvetica must be found in the 75dpi directory in order for interfaces to display properly.

28. Fatal errors caught by Ecce, referred to as "asserts", no longer produce core files. These core files would often take minutes to write thus making it appear that Ecce was in an infinite loop. Now the command line prompt where Ecce was started will return instantly after a fatal assertion is issued. Please report these fatal assertions to ecce-support@emsl.pnl.gov. There are also less severe warning assertions where Ecce will continue operating. We would appreciate if these were reported as well.

What's Broken?

1. The public domain Amulet GUI toolkit does not work with Ecce running under the Linux KDE window manager remotely logged in to a Sun or SGI. The Calculation Editor details dialogs, developed with Amulet, will not come up under KDE. The Calculation Editor will issue an error message stating that the Linux GNOME window manager should be used instead. When the Linux port of Ecce is released, v2.2, this will no longer be an issue as it is unlikely anyone will run Ecce applications remotely on a Sun or SGI from Linux when they can run locally.

2. A number of the online help links that are in the left-hand frame of the web browser when online help is invoke point to invalid URLs. The online help was not fully converted for deployment with Ecce data server and these links still reference the online help being on the EMSL public web server. These links will be fixed by the next major release. The majority of online help is functional and the bad links in the left-hand frame just provide supporting information rather than core Ecce application documentation.

3. If you are viewing molecular orbitals as a calculation is currently running, the coordinates might not match the orbitals.

4. You must manually refresh the geometry table after using the bond rotator or the atom manipulator to update the coordinates.

5. The sphere radius selection mechanism in the Molecule Builder, which is initiated by clicking and dragging on an atom, currently does not work.

6. Display of aromatic rings in the Molecule Builder and Calculation Viewer is disabled. Double and triple bond displays are still supported.

7. Do not try to select atoms to use with measures while a calculation property is animating. Press stop, select the atoms you are interested in, and then start the animation again.

8. There has been an ongoing X Windows Server resource problem on SGIs where allocating colors and/or pixmaps can fail. This is dependent upon the X server resource requirements for all applications currently being run, Ecce applications and others, and the hardware (model, memory, graphics card) configuration of the workstation. Usually, a series of X Windows allocation failure warnings will be printed to the window where Ecce was started and you may notice certain colors/pixmaps are missing from applications. However, it is also possible for applications to crash from the failure to allocate colors or pixmaps. The workaround to this is to close any unneeded applications, both Ecce applications and others--especially those which allocate many colors such as netscape. Logging out from the workstation (thus shutting down the window manager session), and then back in again, is another means to free up resources.

9 There is a known bug where the access_log file for an Apache web server running on an SGI will not show correct IP addresses for clients connecting but show 255.255.255.255 instead. This is not something the Ecce team can fix. According to reports, it is a deficiency in the IRIX libraries/ compiler so SGI will need to resolve this. If you are concerned, please consider using a data server running on Linux or a Sun.