GDE 2.0

GamBet Distributed-computing Extension

PO Box 13595, Albuquerque, NM 87192 U.S.A.
Telephone: +1-505-220-3975
E mail: techinfo@fieldp.com
Internet: http://www.fieldp.com

1 Introduction

GDE (GamBet Distributed-computing Extension) is an add-on package for **GamBet** or **Xenos** (professional versions). With **GDE**, it is possible to distribute large Monte Carlo simulations over an unlimited number of computers. The package includes three programs: `gb_sow.exe`, `gb_sow_32.exe` and `gb_reap.exe`. Figure 1 illustrates how they are used. The programs `gb_sow.exe` and `gb_sow_32.exe` have no license restrictions and may be installed on any number of worker computers. They are full-featured versions of **GamBet** 3.0 that perform the Monte Carlo calculation and consolidate all output data into a single binary file. The set of binary files is transferred to a master computer with a licensed installation of **GamBet** or **Xenos**. Here, the program `gb_reap.exe` collects information from all available data files, averages the information and generates standard **GamBet** output files with enhanced statistical accuracy.

The sole user task is to transfer files to and from the worker computers. Users have complete flexibility to configure their distributed network and to automate file transfers. Despite its simplicity, the **GDE** concept has several advantages:

- Worker calculations are independent and need not be synchronized.
- There is no overhead for parallel processing. Five quad-core computers can reduce the time to generate a required number of showers by a factor of 20.
- If one worker machine fails during the computation, data from the other computers may still be combined.
- It is easy to improve the accuracy of a calculation by adding more showers without starting over. The user simply creates more worker-computer files. They are added to the master folder and the total collection is recombined with `gb_reap.exe`.
- Communication between computers is solely through file transfers. Therefore, it is easy to carry out extended calculations at any locations via an institutional network or the Internet.

Four installation programs are included in the **GDE** package: `install_gambet_reap.exe`, `install_xenos_reap.exe`, `install_gbsow.exe` and `install_gbsow_32.exe`. Run the installer `install_gambet_reap.exe` on a computer with a licensed installation of **GamBet Pro**. The program adds `gb_reap.exe`, `gb_sow.exe`, this document and a new version of `gblaunch64.exe` to the directory `c:\fieldp\gambet`. The new **GBLaunch** has buttons for **GB_Sow** and **GB_Reap** as well as support for the new programs under the *Create task* button. The programs use the existing package activation and will run immediately. For a installation of **Xenos Pro**, run `install_xenos_reap.exe` to add components to `c:\fieldp\xenos`.

The installation programs `install_gbsow.exe` and `install_gbsow_32.exe` for worker computers may be distributed by any means (Internet download, USB drive,...). Run the appropriate installer on each worker computer. The installer creates the directory `c:\gambet_sow` with the executable program and the directory `c:\gambet_sow\reference` with Penelope data files.

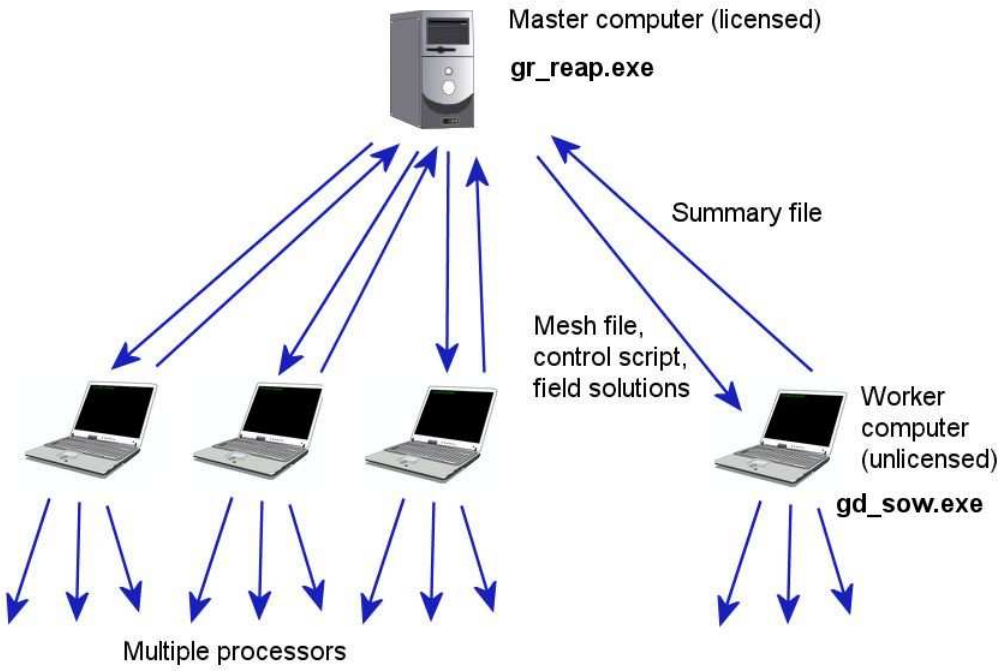

Figure 1: Organization of **GDE**.

2 Setting up a run

This section shows how to carry out a distributed calculation manually. The following section gives suggestions for automation. To begin, create and test a calculation with a moderate number of showers using the standard **GamBet** programs. Inspect the `GLS` listing file to make sure everything is correct. Then, collect the required input files:

- The **GamBet** control script (`GIN`).
- A 2D or a 3D geometry mesh file (`MOU` or `MDF`).
- Input particle definitions if they are not included in the control script (`SRC` or `PRT`).
- Optionally, field-solution files for electron dynamics (`EOU` for **EStat**, `POU` for **PerMag**, `HOU` for **HiPhi** or `GOU` for **Magnum**).

Next, create a working directory on each worker computer for the calculation. Copy the complete set of input files to the worker computers.

Run **GB_Sow** to start a calculation on a worker computer, either by using the desktop shortcut or by double-clicking `gb_sow.exe` in your file manager. The program starts in the simple window shown in Fig. 2. Use the editing commands in the *File* menu to inspect **GamBet** control scripts and other text files. To initiate a calculation, click on *Start run*, navigate to the working directory and choose the **GamBet** control script. The program displays the progress of the calculation as shown in the figure. A run from control script `FPREFIX.GIN` creates two files. The text file `FPREFIX.GLS` is the standard **GamBet** diagnostic listing. The binary data file has a name of the form `FPREFIX_SOW.AAAAA`, where the suffix `AAAAAA` is a random string of

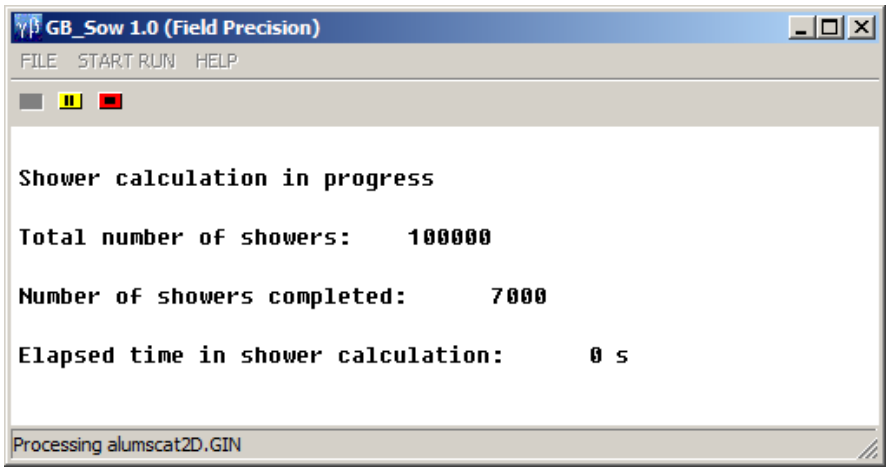

Figure 2: **GB_Sow** running in a window

six upper-case characters. The purpose of the suffix is to ensure that each data file accumulated on the master computer has a unique name. There are over 300 million suffix permutations, so the chances of a file conflict are low. Finally, note that you can launch several simultaneous calculations on a single multi-core or multi-processor worker computer. Be sure to create a data directory for each program instance to avoid file conflicts.

When all worker calculations are complete, copy the binary data files to the processing directory on the master computer. It's a good idea to save data files on the worker computers, because they may be recombined later with additional files for increased accuracy. In addition to the worker data files, the processing directory must contain the **GamBet** script `FPREFIX.GIN` and the mesh file (`MOU` or MDF) used for the calculation. To avoid invalid results, these files must be consistent with those used for the worker calculations. Set the data folder in **GBLaunch** to the processing directory and run **GB_Reap** to open the window of Fig. 3. This program also has a simple editor accessible from the *File* menu. You can open the **GamBet** manuals from the *Help* menu.

Click on the *Process* command and choose the script `FPREFIX.GIN`. **GB_Reap** first reads run parameters from `FPREFIX.GIN` and mesh properties from the `MOU` or `MDF` file. The program then collects a list of available files with names of the form `FPREFIX_SOW.AAAAA` and processes each one. The program checks that all data files have been created from identical meshes and issues a warning if the length of the **GamBet** script does not match that of the script used for the worker calculation. **GB_Reap** creates the following output files:

- `FPREFIX.GDE` (text). A listing of run information and statistical quantities, the average of calculated quantities in the worker files.
- `FPREFIXESC.SRC` (text). The sum of escape particle data from the worker files. If a single worker run creates N_e escape particles and there are N_w worker files, then the master escape file contains about $N_e \times N_w$ data lines. For runs in the *ContinuousBeam* and *PulsedBeam* modes, current and flux values are divided by N_w.
- `FPREFIX.G2D` or `FPREFIX.G3D` (binary or text). Dose distribution files that may be analyzed with **GBView2** or **GBView3**. Dose values from the worker files are averaged. The format is determined by the *Format* command in `FPREFIX.GIN`.

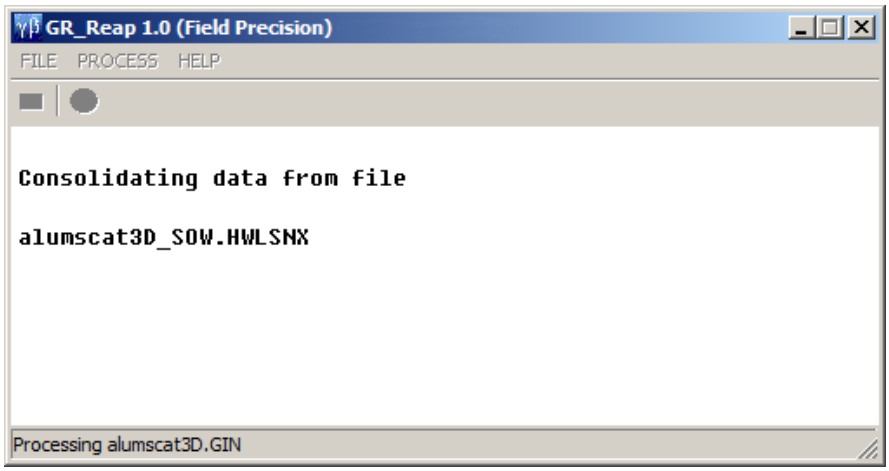

Figure 3: **GB_REAP** running in a window

To limit the size of the final escape file, you can use the *NoEscape* command in some or all of the worker scripts.

Note that some **GamBet** script commands are not supported in **GB _Sow**. The following commands are ignored if they appear in `FPREFIX.GIN`:

- `FORMAT` (`GEOMETRY` section). Output to the summary file is always in binary format.
- `RSEED` (`PROCESS` section). This command was removed to ensure statistical independence of worker runs. If the user supplied a seed, every run would generate an identical sequence of seemingly random showers. In this case, nothing would be gained by averaging the worker runs. In the default mode, **GB_Sow** sets the seed from the system clock, which will be different for each worker calculation.
- `MCA` (`PROCESS` section). The multi-channel analysis function is not supported.
- `PARALLEL` (`PROCESS` section). The parallel-processing mode of **GamBet** is not supported. Instead, launch individual instances **GB_Sow** on multi-processor or multi-core worker computers.

3 Automating runs

GB_Sow and **GB_Reap** may be run from the Windows command prompt. The feature is critical to enable automation. The programs are launched in the background with commands of the form:

```
program_directory\gb_sow64.exe data_directory\fprefix
program_directory\gb_reap64.exe data_directory\fprefix
```

Users can create sophisticated automation sequences by calling the **GamBet** programs from Perl or Python scripts. This section discusses a simpler approach, Windows batch files. To illustrate, this set of commands launches four simultaneous instances of **GB_Sow** on a worker computer.

```
start \b c:\gb_sow\gb_sow64.exe c:\gbdata\liver_study\directory01\liver_dose
start \b c:\gb_sow\gb_sow64.exe c:\gbdata\liver_study\directory02\liver_dose
start \b c:\gb_sow\gb_sow64.exe c:\gbdata\liver_study\directory03\liver_dose
start \b c:\gb_sow\gb_sow64.exe c:\gbdata\liver_study\directory04\liver_dose
```

The requirement is that the input files (including the control script `liver_dose.gin`) have been copied to the four data directories. The *Start* command with the \b option runs **GB_Sow** in the background without opening a window.

Ideally, a batch process would generate a binary data file and then automatically upload it to a server accessible to the master computer. One possible approach uses the freeware FTP utility **WinSCP**, available at

`http://winscp.net/eng/index.php.`

The program should be installed and `c:\program files\winscp` added to the `PATH` environmental variable. To illustrate the procedure, consider a calculation controlled by the **GamBet** script `brainscan.gin` in the data directory `c:\datafiles\brainscan01`. When the calculation is complete, the binary summary file should be uploaded to the password-protected directory `public_html/gambetprocess` at the URL `mycompany.com`. The user name and password for FTP transfers is `myuser` and `mypassword`. The user creates a text file `autorun01.bat` in the directory `c:\datafiles\brainscan01` with the content:

```
start \b \w c:\gb_sow\gb_sow64.exe brainscan01
winscp.exe /console /script=autorun01.txt
```

The batch file runs **GB_Sow** with directives to run in the background (\b) and to wait for completion (\w) before the next step. The second command runs **WinSCP** in the console mode and calls the script `autorun01.txt` for instructions. The second script has the contents:

```
open ftp://myuser:mypassword@mycompany.com/public_html/gambetprocess
option confirm off
put brainscan01_sow.*
exit
```

The first line opens a password-protected FTP session and transfers to the master directory. The second line ensures that the script will not pause for confirmation when replacing existing files. The third line copies the data file.