mpiexec

Article • 02/09/2022 • 6 contributors

Feedback

In this article

Starts one or more Message Passing Interface (MPI) applications on an HPC cluster.

For examples of how to use this command, see Examples.

Syntax

```
mpiexec [/genv <env_var_name> <env_var_value>]
[/gpath <path1>[;<path2>...]] [/gwdir <working_dir>] [/ghost
[/gmachinefile <file_name>] [/genvlist <env_var1>[,<env_var2
[/cores <num>] [/affinity] [/affinity_layout <algorithm>[:<t
[/port <port>] [/timeout <seconds>] [/job <string>] [/lines]
[/t [(<filter1>,<filter2>...)]] [/tracefile <file_name>]
[/tracemax <max_size>] [/debug [{0|1|2|3}]] [/logfile <log_f
[/machinefile <file_name>] [/host <host_name>]
[/hosts <num> <host1> [num1] <host2> [num2] ... <hostn> [num
[/wdir <working_dir>] [/env <env_var_name> <env_var_value>]
[/path <path1>[;<path2>...]] <application1> [<application1_p
[: [<section2_parameters>] <application2> [<application2_par
[: [<section3_ parameters>] <application3> [<application3_pa

mpiexec /configfile <file_name>

mpiexec {/? | /help | /?? | /help2 | /??? | /help3}
```

Parameters

Expand table

Parameter	Description
/genv <env_var_name> <env_var_value>	Sets an environment variable to the specified va sections of the **mpiexec** command. If you specif variable for a specific section by using the **/env** using the **/genv** parameter, the setting that the the section overrides the setting that the **/genv** To specify the values of multiple environment va parameter multiple times, once for each environ For a list of built-in environment variables that parameter, see MPICH environment variables.
/gpath <path1>[;<path2>...]	Specifies one or more paths to search on the ta in all sections of the **mpiexec** command. To spec paths with a semicolon. If the paths include spa quotation marks ("). The value you specify for th

	and is not appended to the PATH environment v If you specify the search path for a specific secti parameter and globally by using the **/gpath** par **/path** parameter specifies for the section overric parameter specifies.
/gwdir <working_dir>	Specifies the working directory to use for the ap **mpiexec** command. This directory can have a lo also specify this parameter as **/gdir**. If you specify the working directory for a specific parameter and globally by using the **/gwdir** par **/wdir** parameter specifies for the section overric parameter specifies.
/ghost <host_name>	Starts the applications in all sections of the **mpie** node. If you also specify the **/np** * parameter or parameter, the application uses one core on tha If the task that runs the **mpiexec** command is n the **/ghost** parameter specifies, the command fa nodes for the task such that you ensure that th If you specify the **/ghosts**, **/gmachinefile**, or **/np** value, you cannot also specify the **/ghost** param If you specify the host for a specific section of th the **/host** parameter and globally by using the **/** that the **/host** parameter specifies for the sectio **/ghost** parameter specifies.
/gmachinefile <file_name>	Reads from the specified file a list of nodes on w all sections of the **mpiexec** command. The forma per line, optionally followed by the number of co allowed to run on that node. You can designate lines or parts of lines in the fi number sign (#) at the start of the comment. Th end of the line. Empty lines in the file are ignore The file must exist on each node that is allocated **mpiexec** command.

	If the task that runs the **mpiexec** command is n nodes listed in the file that the **/gmachinefile** pa command fails. Specify the resources and nodes ensure that the task is allocated to all of the noc If you also specify the **/np** * option or do not sp application uses the sum of the number of core nodes in the file. If you specify a value for the **/r** sum of the number of cores that are specified fc the application runs only the total number of cc specifies. Cores on the nodes are used in the or the nodes. If you specify the **/ghost** parameter, you cannot parameter. If you specify the file that lists the nodes on whi specific section of the **mpiexec** command by usi and globally for the applications in all sections o using the **/gmachinefile** parameter, the setting parameter specifies for the section overrides the **/gmachinefile** parameter specifies.
/configfile <file_name>	Reads the remaining parameters for the **mpiexe** specified file. The lines in the file are command-l format: [<section_parameters>] <application> [<applica You can include a command in the file that span line of the command except the last line with a b lines or parts of lines in the file as comments by the start of the comment. The comment then ex Empty lines in the file are ignored. If you run the **mpiexec** command as a task in a node that is allocated to the task. If you run the the **clusrun** command, the file must exist on eac **mpiexec**.
/np {<num>|*}	Specifies the number of processes on which to r specified in the same section of the **mpiexec** cor If you specify an asterisk (*), the MPI application more than one section is specified, in which case

	remaining available cores. The absence of the **/n** **/np ***. You can also specify this parameter as **/n**. If you specify the **/hosts** parameter, you cannot parameter.
/machinefile <file_name>	Reads from the specified file a list of nodes on w that is specified in the same section of the **mpie** **/machinefile** parameter. The format for the file i optionally followed by the number of cores on v to run on that node. You can designate lines or parts of lines in the fi number sign (#) at the start of the comment. Th end of the line. Empty lines in the file are ignore The file must exist on each node that is allocatec **mpiexec** command. If the task that runs the **mpiexec** command is n nodes listed in the file that the **/machinefile** par fails. Specify the resources and nodes for the tas the task is allocated to all of the nodes in the file If you also specify the **/np *** option or do not sp application uses the sum of the number of cores nodes in the file. If you specify a value for the **/r** sum of the number of cores that are specified fc the application runs only the total number of cc specifies. Cores on the nodes are used in the orc the nodes. If you specify the **/host** or **/hosts** parameter, yo **/machinefile** parameter. If you specify the file the lists the nodes on whic specific section of the **mpiexec** command by usi and globally for the applications in all sections o using the **/gmachinefile** parameter, the setting parameter specifies for the section overrides the **/gmachinefile** parameter specifies.

/host <host_name>	Starts the application that is specified in the sam command as the **/host** parameter on the specifi * parameter or do not specify the **/np** paramete on that node. If the task that runs the **mpiexec** command is n the **/host** parameter specifies, the command fail nodes for the task such that you ensure that th If you specify the **/hosts**, **/machinefile**, or **/np** p you cannot also specify the **/host** parameter. If you specify the host for a specific section of th the **/host** parameter and globally by using the **/** that the **/host** parameter specifies for the sectio **/ghost** parameter specifies.
/hosts <num> <host1> [num1] <host2> [num2] ... <hostn> [numn]	Specifies the number of nodes on which you wa specified in the same section of the **mpiexec** cor parameter, then lists the names of each node ar of processes to run on that node. If the task that runs the **mpiexec** command is n nodes that the **/hosts** parameter specifies, the c resources and nodes for the task such that you allocated to all of the nodes in the list. If you specify the **/host** or **/machinefile** parame **/hosts** parameter.
/cores <num>	Specifies the numbers of cores to use on each n number of cores that are specified for each nod the file that is specified by the **/machinefile** para all sections of the **mpiexec** command. You can also specify this parameter as **/c**.
/affinity	Sets the affinity mask for each of the processes starts to a single core. This setting should only be used if a job is using sharing a node (not running exclusively), this aff assign the same affinity mask to all jobs on the

	nodes, you must either set affinity explicitly, or a manage affinity. For more information, see job s For more information about affinity, see Perform Cluster for Parallel Applications .
/affinity_layout	Specifies the algorithm used to distribute rank p Specifying a different algorithm will change the runs on. Changing the target allows larger units rank process. Target is typically used only for multithreaded ap Following are the values for the algorithm: Disabled = 0 - Does not assign affinity to any p Spread = 1 - Distributes the processes as far as Sequential = 2 - Distributes the processes per c Balanced = 3 - Distribute the processes over the Following are the values for the target: L - Assigns each process to a logical core. P - Assigns each process to a physical core. N - Assigns each process to a NUMA node. This parameter was introduced in HPC Pack 201 previous versions.
/wdir <working_dir>	Specifies the working directory to use for the ap section of the **mpiexec** command as the **/wdir** p have a local or a remote path. You can also spec If you specify the working directory for a specific parameter and globally by using the **/gwdir** par **/wdir** parameter specifies for the section overric parameter specifies.

/env <env_var_name> <env_var_value>	Sets an environment variable for the application section of the **mpiexec** command as the **/env** pa value of an environment variable for a specific se parameter and globally by using the **/genv** para **/env** parameter specifies for the section overrid parameter specifies. To specify the values of multiple environment va parameter multiple times, once for each environ For a list of built-in environment variables that y parameter, see MPICH environment variables.
/genvlist <env_var1>[, <env_var2>...]	Passes the values of the specified environment v **mpiexec** starts. The list is a comma-separated lis names. For a list of built-in environment variables that y parameter, see MPICH environment variables.
/exitcodes	Prints the exit codes of the processes that **mpie** output at the end of the run.
/priority {0|1|2|3|4}	Sets the priority for the processes that **mpiexec** Following are the possible values: 0 - Idle 1 - Below normal 2 - Normal 3 - Above normal 4 - High
/port <port>	Specifies the port on which the smpd.exe proces this parameter as **/p**.
/path <path1>[;<path2>...]	Specifies one or more paths to search on the tar that is specified in the same section of the **mpie** parameter. To specify multiple paths, separate p paths include spaces, enclose the paths in quota you specify for this parameter does not replace

	PATH environment variable. If you specify the search path for a specific secti parameter and globally by using the **/gpath** par **/path** parameter specifies for the section overric parameter specifies.
/timeout <seconds>	Sets the amount of time that job that runs the **r** before timing out, in seconds.
/job <string>	Associates an MPI job with a job that is created Scheduler Service. The string is passed to **mpiex** Service.
/lines	Prefixes each line in the output of the **mpiexec** c process that generated the line. You can also sp
/trace [(<filter1>,<filter2>...)]	Traces Microsoft-Message Passing Interface (MS application. If you do not specify any filters, the MS-MPI events for the application. Optionally, y enable only trace events of interest. List the eve equivalent hexadecimal values, by using a comm parentheses. **Important:** While this type of tracir simple test application, in most practical applicat such a large trace file that it is not very useful. T manageable trace files, use an Event Tracing for Xperf or Logman to turn on tracing only for the By default, trace logs are written to the directory node. Use the **/tracefile** parameter to specify ar Following are the event filter names and equival you can include in the list of filters: all (0xffffffff) - All API and communication events api (0x00007fff) - All API events pt2pt (0x00000001) - Point-to-point APIs poll (0x00000002) - Point-to-point polling APIs, MPI_TestXXX coll (0x00000004) - Collective APIs

	rma (0x00000008) - One-sided APIs comm (0x00000010) - Communication APIs errh (0x00000020) - Error handler APIs group (0x00000040) - Group APIs attr (0x00000080) - Attribute APIs dtype (0x00000100) - Data type APIs io (0x00000200) - Input/output APIs topo (0x00000400) - Topology APIs spawn (0x00000800) - Dynamic process APIs init (0x00001000) - Initialization APIs info (0x00002000) - Information APIs misc (0x00004000) - Miscellaneous APIs interconn (0x000f8000) - All interconnectivity co icsock (0x00008000) - Socket interconnectivity icshm (0x00010000) - Shared memory intercon icnd (0x00020000) - NetworkDirect interconnec You can also specify this parameter as **/t**. This parameter is deprecated as of HPC Pack 20 (SP3).
/tracefile <file_name>	Specifies the name of file to use for the trace log default file is %USERPROFILE%\mpi_trace_*job_identifier.task_* You can also specify this parameter as **/tf**.

	This parameter is deprecated as of HPC Pack 20 (SP3).
/tracemax <max_size>	Specifies the maximum size of the trace log file, i at least *max_size* megabytes of free space availa specify for the trace file. The binary tracing data is written by using a circ exceeds the maximum size of the file, the tracing the beginning of the file. As a result, the tracing most recent *max_size* megabytes of tracing data Each binary record in the trace log file has a tim can display the information in chronological ord cause when the tracing data is overwritten. The default value for this parameter is **10240**. Sp a trace log file of unrestricted size. You can also specify this parameter as **/tm**. This parameter was introduced in HPC Pack 200 HPC Pack 2008 R2 with Service Pack 3 (SP3).
/debug [{0|1|2|3}]	Prints the specified level of debugging informati specify the **/debug** option without a value, **mpie** as value of **2**. Following are the possible debugg 0 - No debugging information is displayed. 1 - Information about errors is displayed. 2 - Extensive debugging information is displayed 3 - Both errors and extensive debugging inform You can also specify this parameter as **/d**.
/logfile <log_file_name>	Specifies the name of file to use for the log file, i
<section*N*_parameters>	Specifies parameters that apply only to the secti in which they occur, and includes the following p - **/env**

	- **/path** - **/wdir** - **/host** - **/hosts** - **/machinefile** - **/np** The Syntax section explicitly lists these paramete uses the <section*N*_parameters> placeholder fo information about sections in an **mpiexec** comm the <application*N*> [<application*N*_parameters
<application*N*> [<application*N*_parameters>]	Specifies the MPI applications that you want to parameters for those applications. You can speci by including multiple colon-separated sections i If you include multiple sections to specify multip need to have compatible requirements for the r need to run and the roles that the applications s example, you can use two sections to specify a r run on rank 0 and a subordinate process that d
/pwd <string>	Authenticates the user with the provided passw when MS-MPI Launch Service is being used.
/saveCreds	Notifies the Launch Service to save credentials. T /pwd is provided. After a successful invocation of /saveCreds, it is password with /pwd unless the password is cha
/unicode	Switches **mpiexec** output to unicode stream. Th **mpiexec**. Unicode path and executables are sup option.
/?	Displays Help for commonly used parameters of command prompt. You can also specify this parameter as **/help**.
/??	Displays Help descriptions for all parameters of examples of the **mpiexec** command at the comm You can also specify this parameter as **/help2**.

/???	Displays Help for the environment variables tha command at the command prompt. You can also specify this parameter as **/help3**.

Remarks

In most cases, you should run the **mpiexec** command by specifying it in a task for a job. You can run **mpiexec** directly at a command prompt if the application requires only a single node and you run it on the local computer, instead of specifying nodes with the **/host**, **/hosts**, or **/machinefile** parameters.

If you run the **mpiexec** command by using the **clusrun** command, the **mpiexec** command behaves as if you specified the command locally at the command prompt on each node on which **clusrun** runs the **mpiexec** command. In this case, you again can only run the **mpiexec** command if the application requires only a single node, and only if you do not specify other nodes by using the parameters of the **mpiexec** command.

Examples

To run four application1 processes on the local host with four cores, use either of the following commands:

```
mpiexec application1
```

```
mpiexec /np * application1
```

To run one master process and three subordinate processes on the local host with four cores, use the following command, in which the absence of **/np** from the second section is equivalent to **/np ***:

```
mpiexec /np 1 master : subordinate
```

To run one master process and 31 subordinate processes on the hosts listed in a myhosts.txt file that lists four hosts with eight cores each, use the following command:

```
mpiexec /gmachinefile myhosts.txt /np 1 master : subordinate
```

To trace the point-to-point and collective MPI calls that the two application1 processes are making and the interconnectivity communication that is generated by MPI, and limit the maximum size of the tracing file to 20,480 megabytes, use the following command:

```
mpiexec /trace (pt2pt,coll,interconn) /tracemax 20480 /np 2
```

To run the application1 process on the hosts that are listed in a myhosts.txt file and include a connectivity table in the standard output, use the following command:

```
mpiexec /machinefile myhosts.txt /env MPICH_CONNECTIVITY_TAB
```

Additional references

- mpicsync

- etl2clog
- etl2otf
- clusrun
- Command-Line Syntax Key