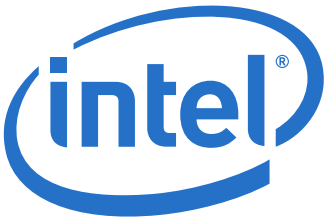

True Scale Fabric Suite Software

Release Notes

January 2014

Order Number: H31512001US

Contents

Tables

1.0 Overview of the Release

1.1 Introduction

These Release Notes provide a brief overview of the changes introduced into the Intel® True Scale Fabric Suite Software (IFS) software by this release. This release notes document includes only the IFS software and must be used in conjunction with the Intel® *OFED+ Host Software Release Notes* for a complete package. References to more detailed information are provided where necessary. The information contained in this document is intended for supplemental use only; it should be used in conjunction with the documentation provided for each component.

These Release Notes list the new features of the release, as well as the system issues that were closed in the development of Release 7.2.2.0.8.

1.2 Audience

The information provided in this document is intended for installers, software support engineers, and service personnel.

1.3 If You Need Help

If you need assistance while working with the True Scale Fabric Suite Software, contact your Intel® approved reseller or Intel® True Scale Technical Support:

- By E-mail:
  ibsupport@intel.com
- On the Support tab at web site:
  http://www.intel.com/infiniband

For OEM-specific server platforms supported by this release, contact your OEM.

1.4 New Features and Enhancements

This section list the new features and enhancements added since Release 7.2.0.0.42 and the two previous major/minor releases for the IFS.

1.4.1 Release 7.2.2.0.8 Enhancements

- Added support for
  - — RedHat EL 5.10 and 6.5
  - — CentOS 5.10 and 6.5
  - — Scientific Linux 5.10 and 6.5

1.4.2 Release 7.2.1.1.22 Enhancements

- Added support for RedHat EL 6.4, SLES 11 SP3, CentOS 6.4 and Scientific Linux 6.4

1.4.3 Release 7.2 Features

1.4.3.1 All

- All features for this release shown in Intel® *OFED+ Host Software Release Notes*

- Branding for Intel has been incorporated in this release. The following packages are included in the branding effort:
  - Intel® True Scale Fabric Suite Fabric Manager
  - Intel® True Scale Fabric Suite Fabric Viewer
  - Intel® True Scale Fabric Suite FastFabric

1.4.3.2 FastFabric

- None.

1.4.3.3 Fabric Manager

- None

1.4.4 Release 7.1.1 Features

1.4.4.1 FastFabric

- None

1.4.4.2 Fabric Manager

- None

1.4.5 Release 7.1 Features

1.4.5.1 FastFabric

- The following file names have been changed:
  - `iba_verifynodes` is renamed `iba_verifyhosts`.
  - `nodeverify.sh` is renamed `hostverify.sh`
  - `nodeverify.res` is renamed `hostverify.res`
  - `FF_NODEVERIFY` is renamed `FF_HOSTVERIFY`
- `captureall` now processes the `-D` option for all operations (host, switch and chassis). When specified, the `-D` option causes a local host capture which includes the specified level of fabric detail.
- `iba_verifyhosts` has been adjusted to use the FF_HOSTVERIFY fastfabric.conf configuration parameter to control the location where `hostverify.sh` is copied and run. The new default is `/root/hostverify.sh`. Previously the default was hard-coded as `/opt/iba/ib_tools/nodeverify.sh`.
- `mpi_apps` and `shmem_apps` now allow MPI_TASKSET to be exported or set in the `.params` file being used for the job. This variable is used by all the `run*` scripts and can optionally provide an argument to `/bin/taskset` which will be used to select the CPUs to be used for each rank in the job. This can be helpful when benchmarking to ensure consistency of runs, such as in systems with multiple CPU sockets or to avoid OS overhead which may be unique to `CPU 0`.
- `hostverify.sh` (which is used by `iba_verifyhosts`) now supports runtime creation of the `HPL.dat` file using `config_hpl2`. The HPL_CONFIG variable can be edited in the script to select an appropriate configuration file and optionally a problem size. If set to "", the tool will not create a configuration file and will expect the file to already exist, such that a single iteration of HPL will be run with an appropriate problem size. See `/root/hostverify.sh` for more information.

- `config_hpl` and `config_hpl2` (in `/opt/iba/src/mpi_apps`) now have a `-l` option. This option will cause the local `HPL.dat` file to be updated and will not update any other servers. This option can be useful when preparing for single node HPL runs.
- `iba_switch_admin getconfig` now outputs a summary of how many switches have each value for the various configuration settings. This helps to make it more obvious if all switches have the same configuration, and if not, indicates how many have each value. If some of the values are not as expected, the test.res file can be viewed to review and identify which switches have the undesirable values.
  `iba_switch_admin info` and `iba_chassis_admin getconfig` report similar summaries. `iba_chassis_admin getconfig` now includes Firmware Active and Firmware Primary in its output.
- **Build MPI Test Apps and Copy to Hosts** in the FastFabric TUI has been enhanced and renamed as **Build Test Apps and Copy to Hosts**. This selection now supports the building and copying of both MPI and SHMEM test applications, and uses the new variable, FF_SHMEM_APPS_DIR in `fastfabric.conf` file to select the location for the `shmem_apps`.
- To avoid potential confusion, all prompts in the FastFabric TUI regarding selection of a file lists hosts to act on have been changed to **Host File**. Previously this was referred to as Hosts File or Host List in various prompts and menus. All references to chassis and switch lists/files are now referred to as **Chassis File** and **Switch File**.
- A commented out sample of using taskset to select CPU cores for MPI and SHMEM sample applications is shown in the various parameter files.
- The FastFabric TUI now has a **Generate or Update Switch File** option. This can use `iba_gen_ibnodes` to generate or regenerate an ibnodes file and/or update an ibnodes file to have the proper names for each switch based on a topology `.xml` file and the analysis of the existing fabric.
- `setup_ssh` will now enable each node to ssh to itself as localhost, its hostname, or using its IPoIB name. When `setup_ssh -R` option is used, this capability is not configured.
- `iba_switch_admin` now appends to `punchlist.csv` for failing switches.
- `cmdall` now supports a `-P` option, this will cause the hostname or chassis to be prefixed to each line of the output. This can be useful when processing an output in a script or when doing simple operations such as `grep`.
- `run_nxnlatbw` has been added to `mpi_apps`. This will run the `mpi_nxnlatbw` test which was added to mpitests.
- The FastFabric TUI has been revised to improve coverage of fabric and server verification as well as take advantage of some of the newer FastFabric CLI tools. The majority of changes are in the **Host Verification** menu.
  The new capabilities include the following:
  - Provide coverage of ssh and InfiniBand* (IB) active host testing in addition to ping
  - Allow "bad" hosts to be easily skipped in subsequent tests
  - Add single node verification capabilities
  - Provide more comprehensive fabric status and verification, including topology and Bit Error rate verification
  - Allow the starting and stopping of Bit Error rate tests for HCA-SW and inter-switch links (ISLs)
  - Provide access to punchlists created by many of the updated CLI tools

- run_batch_cabletest now supports a -n option to specify number of processes to run per host.
- iba_linkanalysis is a new tool which encapsulates the capabilities for link analysis from the fastfabric TUI's **Check Status of IB Ports** option. The new tool also includes cable and fabric topology verification capabilities. This tool is built on top of iba_report and its analysis capabilities, and accepts the same syntax for input topology and snapshot files.

  In addition to being able to run assorted iba_report link analysis reports and generate the human readable output, this tool also will analyze the results and append a concise summary of issues found to the FF_RESULT_DIR/punchlist.csv file.

  See iba_linkanalysis --help for more details.
- iba_expand_file is a new low level CLI tool which can assist the user by expanding a fastfabric hosts, chassis or ibnodes file. It will expand, include, and filter out blank and comment lines. This can be useful when building other scripts which may use these files as input.
- iba_cabletest is a new CLI tool which can be used to initiate or stop Cable Bit Error Rate stress tests for HCA-SW links and/or ISLs. See the help text for more details.
- iba_findgood now automatically generates an FF_RESULT_DIR/punchlist.csv file. This file will provide a concise summary of the bad hosts found. This file can be imported into excel directly as a .csv file or can also be cut and pasted into an Excel spreadsheet and then the **Data/Text to Columns** tool bar can be used to separate the information into multiple columns at the semicolons.

  For a given run a line will be generated for each failing host. Hosts are reported exactly once for a given run. So a host which doesn't ping will NOT be listed as "can't ssh" nor "No active IB port". It should be noted that there may be cases where IB ports could be active for hosts which don't ping, especially if ethernet hostnames are being used for the ping test. However the lack of ping often implies there are other fundamental issues (such as PXE boot, inability to access DNS or DHCP to get proper hostname and IP address, etc.) which imply that reporting hosts which don't ping as also lacking active IB ports will typically be of limited value.

  Note that iba_findgood's approach to determining hosts with active IB ports is to query the SA for NodeDescriptions and therefore ports may be active for hosts which cannot be ssh'ed or pinged.
- iba_verifyhosts will now upload all of the hostverify.res files from the tested hosts. This is controlled using the -d and -u options. iba_verifyhosts will also append to the punchlist.csv file all of the failures found. If a given host has multiple failures, it will be listed in the punchlist.csv file for each failure.
- mpi_multibw has been added to mpi_apps. This is an existing benchmark based on OSU bw and multi_lat which performs multi-core pairwise bandwidth benchmarking.
- cmdall now has a -T option to set a time-limit for execution of host commands. This can be helpful such that progress will be made in large cmdall runs where some of the hosts may not be functional and might hang when executing the command or when trying to establish the ssh session. For example during initial verification of a cluster when some hosts might have hardware or OS issues.

  The -T value is not affected by FF_TIMEOUT_MULT. A value of -1 is infinite, no time-out.
- The sample parameters files for mpi_apps and shmem_apps now have commented out examples of setting the PSM Multi-rail configuration variables.

- `/opt/iba/src/shmem_apps` contains some `run_*` scripts for various SHMEM benchmarks which are included with SHMEM. It also includes a simple sample app, `shmem-hello.c`. The behavior of these scripts is very similar to the `mpi_apps`.
- `/root/hostverify.sh` now supports systems with multiple HCAs.The following tests are enhanced:
  - pcispeed – PCI speed and width of all Intel® HCAs are checked. All HCAs are expected to have the same speed and width
  - pcicfg – PCI Payload and Max Read Request size of all Intel® HCAs are checked. All HCAs are expected to have the same values.
  - ipath_pkt_test – The performance of each HCA is tested. All HCAs are checked against the same expected threshold

    When multiple HCAs are configured, the tests will test all HCAs. This may result in multiple FAIL messages for a single test. `ipath_pkt_test` will report a PASS or FAIL message for each HCA tested.

    To configure for multiple HCAs the variables at the start of the script must be edited.
  - HCA_COUNT – number of Intel® HCAs expected in system. If 0 then tests like `pcispeed`, `pcicfg` and `ipath_pkt_test` merely verify there are no Intel® HCAs.
  - HCA_CPU_CORE[0] – CPU core to run `ipath_pkt_test` on when testing HCA 0
  - Additional variables, such as HCA_CPU_CORE[1], must be specified when there is more than 1 HCA. It is recommended to select a CPU core (other than 0) which is in a CPU chip closest to the HCA within the server. For example Sandy-bridge systems have a PCI bus on each CPU chip and it is recommended to evaluate HCA performance by using the CPU chip whose PCI bus the HCA is connected to.
- The `/opt/iba/src/mpi_apps run_*` scripts can now optionally include the lists of host processes in a given job as part of the output and log. The following controls are available for this feature using environment variables:
  - SHOW_MPI_HOSTS – set to "y" if MPI_HOSTS contents should be output prior to starting job and set to "n" to disable defaults to "y"
  - SHOW_MPI_HOSTS_LINES – maximum lines in MPI_HOSTS to show. Default is 128

  Only lines applicable to the job will be shown. Note that the file might include some comment lines
- When using `run_batch_script` or `run_batch_cabletest`, the log files are now separated with each individual job getting its own log file with a suffix to the log filename indicating the run number within the set of batches. Such as:
  `mpi_groupstress.04Jan12165901.1`,
  `mpi_groupstress.04Jan12165901.2`

  This avoids any previous intermingling of output from multiple runs into a single log file
- Platform HPC/LSF and AC/Moab MPI job submission scripts have been enhanced to support the usage of the new PSM multi-rail feature.

  Multi-rail means a process can use multiple network interface cards to transfer messages. With modern computer servers supporting multiple network interfaces, multi-rail is the major way to improve network performance for applications.

  Before this PSM multi-rail support, PSM could use multiple cards/ports for a single application, but for a particular process in the job, only one port could be used to transfer messages. Also, all the ports had to be on the same fabric in order for the application to use them.

The new PSM multi-rail feature can be applied to a single fabric with multiple ports, or to multiple fabrics. The major goal of this feature is, to use multiple network interfaces to transfer messages to improve the message bandwidth.

- Script `iba_xlat_topology_cust` has been added, accompanied by `topology_cust.xlsx` spreadsheet. The script and spreadsheet provide a sample alternative to the standard-format topology capability to document the topology of a customer cluster (see `iba_xlat_topology`). The alternative is provided for situations in which a customer chooses not to define a fabric topology using the standard-format spreadsheet and `iba_xlat_topology`. For more information refer to the `-help` or the *FastFabric Command Line Interface Reference Guide*.
- `hostverify.sh` now supports multiple Intel® HCAs in its `pcicfg` and `pcispeed` tests. All Intel® HCAs found will be tested and expected to have the same PCI configuration and speed values for these tests to pass.
- `iba_gen_ibnodes` topology update of ibnodes data has been enhanced with support for multiple planes (fabrics). The HCA and port pair specification in `iba_gen_ibnodes` for generation of ibnodes data (PORTS and PORTS_FILE environment variables, `-p`, `-t`) can now be applied to a topology file or a snapshot file by specifying `%P` in the file name. Environment variable FF_TOPOLOGY_FILE contains a default name for a topology file.

  The following additions/changes have been made to options:

  — `-N ibnodes_file` changed to `-L ibnodes_file`; when `-L` is specified no generation of ibnodes data occurs;

  — `-s` – Specifies the update of switch names using topology XML data; formerly this was specified with `-T topology`; `-s` is required with `-L`, `-T`, and `-X`;

  — `-T topology_file` and `-X snapshot_file` continue to be used to specify topology and snapshot files respectively;

  — `-o output_file` – Write ibnodes data to `output_file` (instead of the default `stdout`); `output_file` can contain the same file name as `ibnodes_file`; `output_file` will be written as `iba_gen_ibnodes` completes successfully.

1.4.5.2 Fabric Manager

- DesiredMaxDelay in the CongestionControl section of the `ifs_fm.xml` file is now in units of nanoseconds. It was previously in units of microseconds.
- `/opt/ifs_fm/bin/fm_cmdall` command has been added. This can be used to execute a given `fm_cmd` against one or more FM instances. By default it will run the given command against all running FM instances on the local server.

1.4.6 Release 7.0.1 Features

1.4.6.1 FastFabric

- `iba_verifynodes` is a new tool to help perform single node verification. The actual verification is performed using `/opt/iba/ib_tools/nodeverify.sh`. A sample of `nodeverify.sh` is provided in `/opt/iba/samples/nodeverify.sh` and should be reviewed and edited to set appropriate configuration and performance expectations and select which tests to run by default.

  See `/opt/iba/samples/nodeverify.sh` and `/sbin/iba_verifynodes.sh` for more information.

- Assorted `mpi_apps` run scripts now permit use of `all` as their `number_processes` argument. When all is specified, one rank will be started for every entry in the specified `mpi_hosts` file (eg. number of ranks will be `wc -l mpi_hosts`). `all` may not function as desired if `mpi_hosts` uses more advanced syntaxes as supported by some MPIs, such as `hostname:count`. The following scripts have this capability

```
run_allhcalatency, run_alltoall3, run_app, run_bcast2, run_bcast3, run_deviation,
run_hpl, run_hpl2, run_imb, run_mbw_mr3, run_mpi_stress, run_mpicheck,
run_multi_lat3, run_pmb
```

- `gen_mpi_hosts` now has a `-s` option which will output hosts in order (host1 host1 host1 host2 host2 host2...) while retaining the ordering given in the input file. This can be helpful if a special ordering was created in the input file or host names are such that `iba_sorthosts` or a simple sort will not yield the desired order.
- `iba_report` can now obtain and display 64-bit data movement counters from the PM/PA in the Intel® True Scale Fabric Suite Fabric Manager or directly from the fabric (`-M`).

  Due to this enhancement, snapshot's generated by this new version of `iba_report` in conjunction with the `-s` option may report value out of range errors when used as input to older versions of `iba_report`. However, the thresholds specified in `iba_mon.conf` and other input configuration files continue to only support 32-bit values for data movement counter thresholds.
- `setup_ssh` now supports the `-P` option. This option will skip the ping which is normally done to verify access to the host prior to attempting ssh. This option is necessary when setting up ssh to hosts which have ping fire walled, such as for hosts being ssh'ed across the open internet.
- A new script, `run_batch_cabletest`, in `/opt/iba/src/mpi_apps`, is a specialized stress test for large fabrics. It is similar to `run_cabletest`, but is easier to set up. The hosts whose links need to be tested is set up in the `/opt/iba/src/mpi_apps/mpi_hosts` file. By default, the test is broken down into jobs of 18 hosts each. By using many small jobs, the impact of any individual host issues (host crash, hang, etc) during the test is limited to one batch of hosts. Environment variables can be used to customize the test.
- A new script, `run_batch_script`, in `/opt/iba/src/mpi_apps` makes it easier to run other `run_*` scripts as many smaller jobs. This script runs separate jobs for each `BATCH_SIZE` hosts. By using many small jobs the impact of any individual host issues (host crash, hang, etc) during the test is limited to one batch of hosts.
- The location of the `/opt/iba/src/mpi_apps` for use by FastFabric can now be configured via the `FF_MPI_APPS_DIR` parameter in `fastfabric.conf` file. This allows the directory to be copied to another location (such as a global file system or a per-MPI built copy) for use by FastFabric menus, `iba_host_admin mpiperf` and `mpiperfdeviation`. When copying, it is important not to remove or move the original files in `/opt/iba/src/mpi_apps` the presence of these tools are a pre-requisite for some FastFabric operations.
- `iba_findgood` has been added. This command can check for hosts which are pingable, ssh'able and active on IB and produce a list of good hosts meeting all criteria. The resulting "good" file can then be used in as input to create `mpi_hosts` files for use running `mpi_apps` and/or the Host Channel Adapter (HCA)-switch cabletest. Typical usage would be to quickly identify good hosts which will undergo further testing and benchmarking during initial cluster staging and bring-up. This command assumes the IB Node description for each host will be based on the `hostname -s` output in conjunction with an optional `HCA-#` suffix. Such names are the default when using OFED. Note that when using an `/etc/sysconfig/iba/hosts` file which lists IPoIB hostnames, this assumption may not be correct. The files created (good, alive, running, active, bad) are in `iba_sorthosts` order with any duplicates removed.

- The iba_sorthosts command will sort its stdin in a typical host name order. Hosts are stored alphabetically by any alpha numeric prefix and then sorted numerically by any numeric suffix. Leading zeros in the numeric suffix are optional. This command does not remove duplicates any duplicates are listed in adjacent lines.

  This command can be useful to build mpi_hosts input files for applications or cabletest which places hosts in order by name.

- The run_imb sample script has been added in /opt/iba/src/mpi_apps. This script can be used to run the Intel MPI Benchmark suite (IMB) application which is included in the mpitests rpm. IMB is a newer version of the Pallas benchmark (PMB) which is also included in FastFabric and can be run using the existing run_pmb script.

- The iba_xlat_topology script, accompanied by topology.xlsx, linksum_180.csv and linksum_360.csv provide the capability to document the topology of a customer cluster, and generate a topology XML file based on that topology (translate the spread sheet to a topology file). The topology file can be used to bring up and verify the cluster. The exact steps to be used in the verification process are beyond the scope of this description.

  topology.xlsx provides a standard format for representing each external link in a cluster. Each link contains source, destination and cable fields with one link per line (row) of the spread sheet. Link fields must not contain commas. Source and destination fields each are: Rack Group (rack row), Rack, Name (primary name), Name-2 (secondary name), Port Number and Port Type. Cable fields are: Label, Length and Details.

  Group and Rack names are individually optional. If either is completely empty for the entire spread sheet it will be ignored. If either is empty on a particular row, the script will default the value on that row to the closest previous value (to default to a non-empty value, at least the first row must have a value). Name and Name-2 provide the name of the node which is output as NodeDesc:

```
NodeType     Name         Name-2

Host         hostname     hostdetails

Edge Switch  switchname

Core Leaf    corename     Lnnn

Core Spine   corename     Snnn   (used only in internal core switch links)
```

  For hosts, Name-2 is optional and is output as NodeDetails in the topology XML file; also HCA-1 is appended to Name. For core leaves (and spines) Name and Name-2 are concatenated.

  Port contains the port number. If port is empty on a host node, the script will default to 1. Type contains the node type. If Type is empty on a particular row, the script will default the value on that row to the closest previous value (at least the first row must have a value). Type values are:

```
NodeType     Type

Host         CA

Edge Switch  SW

Core Leaf    CL

Core Spine   CS    (used only in internal core switch links)
```

  Cable values are optional and have no special syntax. If present they appear in the topology XML file as CableLabel, CableLength, and CableDetails respectively.

`iba_xlat_topology` takes as inputs the comma-separated-values (CSV) form of the spread sheet cluster tab, and CSV files for (internal) core switch links (12200-180 and/or 12200-360). The CSV spread sheet cluster tab is named `topology.csv`; the core switch links files are `linksum_180.csv` and `linksum_360.csv` respectively. `topology.csv` is created from the spread sheet by saving/exporting the tab as CSV to `topology.csv` (`topology.csv` should be inspected to ensure that output rows contain the correct number of link fields; extraneous entries on the spread sheet can cause excel to output extra fields). The script produces as output one or more topology files `topology.0:0.xml`. Output at the top level as well as (optionally) Group, Rack, and Switch level can be produced. Input files must be present in the directory from which the script operates.

- Tools that display byte or packet data rates have been changed to use terminology consistent with the International System of Units (SI) and the International Electrotechnical Commission (IEC). The terms applicable to the tools distinguish between data sizes in 10^6 (1,000,000) versus 2^{20} (1,048,576) bytes, and 10^3 (1000) versus 2^{10} (1024) bytes. MegaBytes and KiloPackets (various abbreviations) have been the terms used by the tools.

  The SI and IEC terms are shown in Table 1.

Table 1. SI and IEC terms

Term	Size	Abbrev	Meaning:
megabyte	10^6	M	megabyte (SI)
mebibyte	2^{20}	Mi	megabinary byte (IEC)
kilobyte	10^3	k	kilobyte (SI)
kibibyte	2^{10}	Ki	kilobinary byte (IEC)

  The tools affected are `iba_top`, `iba_report`, `iba_paquery`, `iba_pmaquery`, `iba_fequery`, `iba_porttool`, `iba_portstats`, `iba_mon` and `s20tune`. In all cases the data affected was in binary (2^n) units, so headings of MB and KP (or variations) were changed to (appropriate variations of) MiB (mebibyte) and KiP (kibipacket).

  Note that `iba_report` outputs data rate information in text and XML formats. Text data (usually displayed only to the screen) headers were changed as described above. XML data tags (for example, `<XmitDataMB>`, `<RcvDataMB>`, `<XmitDataExtMB>`, `<RcvDataExtMB>`) were not changed so that customer scripts that might parse this data would function without change.

- `iba_gen_ibnodes` has been enhanced with the capability to generate or update ibnodes data using a topology XML file representing the desired state (links and node names) of the fabric. Using the topology data, `NodeDesc` fields in the ibnodes data that may still be at their initial value will be updated to their desired value as contained in the topology data. New options are used to invoke the enhanced capability; without these options the script functions as before.

1.4.6.2 Fabric Manager

There are no new features for this release.

1.5 Operating Environments Supported

The Release 7.2.2.0.8 version of IFS allows for the Operating Systems listed in Table 2.

Table 2. Operating Environments Supported

Operating System	Update/ SP	Version
Red Hat Enterprise Linux (RHEL) 5 X86_64 (AMD Opteron and Intel EM64T)	Update 8	2.6.18-308.el5.x86_64
	Update 9	2.6.18-348.el5.x86_64
	Update 10	2.6.18-371.el5.x86_64
RHEL 6 X86_64 (AMD Opteron and Intel EM64T)	Update 2	2.6.32-220.el6.x86_64
	Update 3	2.6.32-279.el6.x86_64
	Update 4	2.6.32-358.el6.x86_64
	Update 5	2.6.32-431.el6.x86_64
SLES 11 X86_64 (AMD Opteron and Intel EM64T)	SP2	3.0.13-0.27-default
	SP3	3.0.76-0.11-default
Community Enterprise Operating System (CentOS) X86_64 (AMD Opteron and Intel EM64T)	Update 5.8	2.6.18-308.el5.x86_64
	Update 5.9	2.6.18-348.el5.x86_64
	Update 5.10	2.6.18-371.el5.x86_64
Community Enterprise Operating System (CentOS) X86_64 (AMD Opteron and Intel EM64T)	Update 6.2	2.6.32-220.el6.x86_64
	Update 6.3	2.6.32-279.el6.x86_64
	Update 6.4	2.6.32-358.el6.x86_64
	Update 6.5	2.6.32-431.el6.x86_64
Scientific Linux X86_64 (5.x)	Update 5.8	2.6.18-308.1.1.el5.x86_64
	Update 5.9	2.6.18-348.el5.x86_64
	Update 5.10	2.6.18-371.el5.x86_64
Scientific Linux X86_64 (6.x)	Update 6.2	2.6.32-220.el6.x86_64
	Update 6.3	2.6.32-279.el6.x86_64
	Update 6.4	2.6.32-358.el6.x86_64
	Update 6.5	2.6.32-431.el6.x86_64
StackIQ Cluster Manager (Rocks+) HPC 6.1	RHEL 6.3	2.6.32-279.el6.x86_64
	RHEL 6.4	2.6.32-358.el6.x86_64
	CentOS 6.3	2.6.32-279.el6.x86_64
	CentOS 6.4	2.6.32-358.el6.x86_64
Platform HPC-3.2	RHEL 6.2	2.6.32-220.el6.x86_64
Platform HPC-4.1.1	RHEL 6.4	2.6.32-358.el6.x86_64

CPU model of Linux kernel can be identified by `uname -m` and `/proc/cpuinfo` shown in Table 3

Table 3. CPU Model of Linux Kernel

Model	uname	/proc/cpuinfo
EM64T	x86_64	Intel CPUs
Opteron*	x86_64	AMD CPUs

Note: Other combinations (such as i586 uname) are not currently supported.

1.6 Qualified Parallel File Systems

Lustre and IBM General Parallel File System (GPFS) listed below have been tested for use with this release of the Intel® OFED+ host software using the operating systems listed below:

- Lustre 2.3
  - RHEL 6.3
- Lustre 2.4.1
  - RHEL 6.4
- IBM GPFS 3.5.0.14
  - RHEL 6.4

Refer to the *Intel® OFED+ Host Software User Guide* for the latest configuration recommendations for optimizing Lustre and GPFS performance with Intel® True Scale Fabric.

1.7 Intel Interface for NVIDIA GPUs

NVIDIA's CUDA parallel computing platform and programing models have been tested for use with this release of the Intel® OFED+ host software using the operating systems listed in Table 4:

Table 4. NVIDIA's CUDA Tested with OFED+

Distributions	CUDA 5.5
RHEL 5.10	X
RHEL 6.5	X
SLES 11 SP3	X

1.8 Hardware Supported

Table 5 list the hardware supported in this release.

Table 5. Hardware Supported

HCAs
QLE7340
QLE7342
QME7342
QME7362
QMH7342
MHQH29-*
MHQH19-*
MHQH19B-XTR
MHQH29B-XTR
MHQH29B-XSR
MCX354A-QCAT
MCX353A-QCAT
NC543i (HP SL390 G7 in-built InfiniBand Host Channel Adapter)
CX-3 LOM down QDR
46M2199
46M2203

1.9 Software Supported

1.9.1 Remote Node Software Versions Supported in this Release

The Intel®True Scale Fabric Suite FastFabric management node can manage nodes with the following software:

- Host with FastFabric for OFED Enablement Tools 4.2 or later

Note: While the Intel®True Scale Fabric Suite FastFabric Management Node requires Intel® OFED+ Host Software 1.5.3 or later to run Intel®True Scale Fabric Suite FastFabric 7.2, Intel® FastFabric can manage cluster nodes running Intel® OFED+ Host Software 1.2.5 or Intel® OFED+ Host Software 1.3, OFED 1.4 or OFED 1.5, and Intel® IB Tools 4.2 or later.

- Intel® Internally Managed 9000 series Switches with 4.1 or later firmware
- Intel® Externally Managed 9024FC Switches with 4.1 or later firmware
- Intel® Internally Managed 12000 series Switches with 5.0 or later firmware
- Intel® Externally Managed 12200 Switches with 5.0 or later firmware
- Intel® 12100 Switches with 5.0 or later firmware

The Intel®True Scale Fabric Suite Fabric Manager can manage nodes with the following software:

- Host with Intel® OFED+ Host Software 1.2 or later
- Intel® Internally Managed 9000 series Switches with 4.1 or later firmware
- Intel® Externally Managed 9024FC Switches with 4.1 or later firmware
- Intel® Internally Managed 12000 series Switches with 5.0 or later firmware
- Intel® Externally Managed 12200 Switches with 5.0 or later firmware
- Intel® 12100 Switches with 5.0 or later firmware

1.9.2 Remote Node Software Versions with Reduced Capability

The Intel®True Scale Fabric Suite FastFabric can manage nodes with the following software:

- Nodes running third-party IB Stacks
- OFED nodes without the Intel® IB Tools installed
- Third Party IB Switches

The Intel®True Scale Fabric Suite Fabric Manager can manage nodes with the following software:

- Nodes running third-party IB Stacks
- OFED Nodes with Intel® OFED+ 1.2 or earlier
- Third Party IB Switches

1.10 Installation Requirements

The following list any special or release specific installation requirements for this release.

1.10.1 Package Installation Requirements:

Intel® True Scale Fabric Suite Software (IFS) package should be installed on the head node and Intel® OFED+ Host Software package should be installed on all other nodes except the head node.

When using Intel®True Scale Fabric Suite FastFabric toolset to install other nodes, `IntelIB-Basic.`*`DISTRO`*`.`*`VERSION`*`.tgz` should be downloaded. This file is specified by default in `fastfabric.conf` through the `FF_PRODUCT` and `FF_PRODUCT_VERSION` parameters and is used to install all other nodes.

1.10.2 Software and Firmware Requirements

All IFS software on a given node must be at the same release level. The Intel® OFED+ Host Software is installed as part of the package. Prior to installing the Intel® True Scale Fabric Suite Software release, any versions of the SilverStorm IB stack (and any other vendor's IB stack) must be uninstalled

Note: When using the Intel® True Scale Fabric Suite (IFS) Software installation wrapper, the wrapper install enforces this requirement.

1.11 Changes for this Release

The following sections describe the changes that have been made to the Intel® True Scale Fabric Suite Software package between versions 7.2.0.0.42 and 7.2.2.0.8, including the following releases:

- 7.2.0.0.42
- 7.2.1.1.22

For detailed information about any of the previous releases listed, refer to the Release Notes for the specific version.

1.11.1 Changes to Hardware Support

For changes to hardware supported refer to the Intel® *OFED+ Host Software Release Notes*.

1.11.2 Changes to Operating System Support

For changes to new operating systems supported for the releases refer to the Intel® *OFED+ Host Software Release Notes*..

1.11.3 Changes to Industry Standards Compliance

Refer to the Intel® *OFED+ Host Software Release Notes* for information about the changes to industry standards compliance.

1.12 Product Constraints

The following is a list of product constrains for this release:

1.12.1 FastFabric Toolset Product Constraints

- The product supports a default HCAs configuration of Port 1 on the HCAs as the active port and Port 2 on the HCAs as the standby port. The following FastFabric operations may not work correctly with a HCA configuration of 2 active ports, or a configuration which has Port 2 of the HCAs as the active port:
  - Host Setup using FastFabric->Configure IPoIB IP Address
  - Host Admin using FastFabric->Verify Hosts ping via IPoIB
- All commands that are to be run on the chassis (Intel® and SilverStorm switches and gateways) should be invoked with the `-noprompt` option to avoid command execution time-out. This applies both to chassis commands invoked from the FastFabric TUI (Run a command on all chassis), as well as those invoked from the command line using the FastFabric `cmdall` command.

1.12.2 Fabric Manager

- Virtual Fabrics in this release leverage IBTA standard Partitioning Features. However, some OFED applications have limitations with regard to partitioning.
  - FastFabric – FastFabric tools are fully supported. Intel recommends that FastFabric be installed on an admin node which is a Member in the Default Partition (0xffff).
  - IPoIB – Intel recommends configuring Virtual Fabrics so that the first PKey on the port is the one desired for IPoIB on the host. Refer to the Configuration section of the *Intel*® True Scale Fabric Software Installation Guide for detailed information.
  - mvapich1 – To control the PKey, the VIADEV_DEFAULT_PKEY must be exported at job startup. Refer to the Configuration section of the *Intel*® True Scale Fabric Software Installation Guide for detailed information.
  - Open MPI – To control the PKey, the OMPI_MCA_btl_openib_ib_pkey_val must be exported at job startup. Refer to the Configuration section of the *Intel*® True Scale Fabric Software Installation Guide for detailed information about this feature.
  - mvapich2 – To control the PKey, the MV2_DEFAULT_PKEY must be exported at job startup. Refer to the Configuration section of the *Intel*® True Scale Fabric Software Installation Guide for detailed information.

1.13 Product Limitations

There are no product limitations for this release.

1.14 Other Information

The following is a list of need-to-know information for this release:

1.14.1 FastFabric Toolset Information

The FastFabric Toolset is automatically uninstalled if the base OFED release is uninstalled.

1.14.2 Fabric Manager Information

When there are many changes in the fabric (ISLs, switches going down) it is possible that many loops are no longer viable and the distribution of ISLs in the loops is becoming unbalanced. These changes can cause the loop test utilization to drop. Restarting of loop test will stop all traffic and compute fresh loop routes with balanced distribution of ISLs in loops.

1.15 Documentation

Table 6 lists the Release 7.2 related documentation. All related documentation is available on the Intel® download site.

Documentation for Intel® Partners is available at the vendors web site.

Table 6. Related Documentation for this Release

Document Title	Document Number	Revision
Intel® Hardware Documents		
Intel® True Scale 12000 Hardware Installation Guide	G91928	001US
Intel® True Scale 12000 Users Guide	G91930	001US
Intel® True Scale 12000 CLI Reference Guide	G91931	001US
Intel® Adapter Hardware Installation Guide	G91929	001US
Intel® OFED+ Documents		
Intel® True Scale Fabric Software Installation Guide	G91921	002US
Intel® OFED+ Host Software User Guide	G91902	003US
Intel® *OFED+ Host Software Release Notes*	H31513	001US
Intel® IFS Documents		
Intel® True Scale Fabric Suite FastFabric User Guide	G91916	001US
Intel® True Scale Fabric Suite Fabric Manager User Guide	G91918	001US
Intel® True Scale Fabric Suite FastFabric Command Line Interface Reference Guide	G91904	001US
Intel® True Scale Fabric Suite Software Release Notes	H31512	001US
Intel® Fabric Viewer Documents		
Intel® *True Scale Fabric Suite Fabric Viewer Online Help*	N/A	N/A
Intel® True Scale Fabric Suite Fabric Viewer Release Notes	H31503	001US

2.0 System Issues for Release 7.2.2.0.8

2.1 Introduction

This section provides a list of the resolved issues in the True Scale Fabric Suite Software that were verified by this release. It also lists the open issues with a description and workaround for each.

2.2 Resolved Issues in this Release

Table 7 is a list of issues that are resolved in this and the previous two releases.

Table 7. Resolved Issues

Product	Release	Description
IFS/ HCA	7.2.0.0.42	IPoIB connections no longer become hung (`ib0: transmit timed out`) under certain stressful conditions.
IFS/ FastFabric	7.2.1.1.22	Release 7.2 FastFabric can now get, modify, analyze, and push a Fabric Manager configuration file to a switch (i.e. switch with ESM) running release 7.1.x or earlier firmware,
IFS/ FastFabric	7.2.2.0.8	Result of iba_verifynodes for C-states are no longer misleading on SLES 11.
IFS/ HCA	7.2.2.0.8	IFS now works properly with SLES11SP3 kernel 3.0.93-0.8.

2.3 Known Issues

The subsections below catalog the known open issues for the release as well as a description and a workaround by component.

2.3.1 Severity

This document provides a level of severity for each issue listed The levels are:

- **Critical** – Could result in a service outage
- **Major** – Could degrade system performance
- **Minor** – Could cause minimal impact to ongoing operations
- **None** – No operational impact

2.3.2 Open Issues Table

Table 8 is the list of open issues for 7.2.2.0.8. The table is sorted by product, then severity.

Table 8. Open Issues

Product/ Component	Severity	Description	Workaround
IFS/ FastFabric	Minor	While trying to rebuild mvapich-psm and mvapich-verb using FastFabric with PGI 11.7, it fails with the following error message: `/usr/bin/ld: final link failed: Nonrepresentable section on output`	Add the following line to the login scripts, in the environment such that is it available for both interactive and non-interactive logins. `export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$PGI /linux86-64/11.7/libso` After making the change to the login scripts, exit and log back into the server so its defined and then run the `do_*_build` script.

Table 8. Open Issues (Continued)

Product/ Component	Severity	Description	Workaround
IFS/ Fabric Manager	None	When the `LogFile` parameter is in use, the Fabric Manager outputs to the named file instead of `syslog`. If a high LogLevel is selected, the log file can grow quickly and consume too much disk space.	Limit use of `LogFile` to short duration debug type operations and use `syslog` for normal Fabric Manager operation.

IFS/ Rolls/Kits	Minor	In Release 7.2.1 of IFS and OFED kit, Platform HPC 4.1.1 GUI shows Version as 1.5.4.1 and release version is no longer displayed.	kit name (.bz2) file shows right version after OFED version (1.5.4.1) as 7.2.1.1.19.
			
			`kit-intel_ifs-1.5.4.1-7.2.1.1.19-rhels-6-x86_64.tar.bz2`
			
			`kit-intel_ofed-1.5.4.1-7.2.1.1.19-rhels-6-x86_64.tar.bz2`

IFS/ Rolls/Kits	Minor	If the IFS kit is already installed and running the `updatenode` command on the Installer and/or compute node (`updatenode<headnode/computenode>` command) emits errors similar to the following will be displayed.
		
		`compute000: Error: Package: opensm-devel-3.3.13-1.x86_64 (installed)`
		
		`compute000:Requires: opensm-libs = 3.3.13-1`
		
		`compute000:Removing: opensm-libs-3.3.13-1.x86_64 (installed)`
		
		`compute000: opensm-libs = 3.3.13-1`
		
		`compute000: Updated By: opensm-libs-3.3.15-1.el6.x86_64 (xCAT-rhels6.4-path0)`
			These errors may be safely ignored.

§ §