Intel® Cluster Checker 3.0 webinar
June 3, 2015

Christopher Heller
Technical Consulting Engineer
Q2, 2015
Introduction

Intel Cluster Checker 3.0 is a systems tool for Linux high performance compute clusters

• Detects issues
• Provides diagnoses
• Suggests remedies

The third generation of Intel® Cluster Checker adds significant capabilities over previous versions and will be available as part of Intel® Parallel Studio XE 2016 Cluster Edition for Linux*
New distributed tool architecture provides:

- On-demand and background monitoring modes for distributed cluster tests
- Built-in rule based expert system technology to analyze multifaceted issues
- Built-in knowledge base facilitates remedies for common issues
- Built-in database with threshold data for major range of components
- Automated checking throughout cluster life cycle
- API to integrate in other software

Version 3.0 supports:

- Intel® Xeon® processors and Intel® Xeon® Phi™ coprocessors
- Ethernet*, Intel® True Scale Fabric, or Mellanox InfiniBand* interconnects
- Installs with Intel® Parallel Studio XE 2016 Cluster Edition for Linux*
- Also available in a stand alone package available via ICR channels

*Other names and brands may be claimed as the property of others.
Intel® Cluster Ready

One Cluster Architecture. More Opportunities. Lower Cost

• Increases “out-of-the-box” interoperability between cluster solutions and applications
• Advanced cluster quality management using Intel® Cluster Checker diagnostics
• Reduces expertise barriers for HPC and technical computing clusters

Compliant solutions platform for “volume” technical computing
Intel® Cluster Ready – The Community

www.intel.com/go/cluster
The Cluster Challenge

As a developer targeting a cluster, I want to write code that runs and performs its tasks with the best performance I can achieve - but the complexities and possible issues of clusters challenge both me, as a developer, and my users.

Intel® Cluster Checker

→ cluster systems expertise packaged into a utility
Intel® Cluster Checker 3.0 – Overview

Provides Assistance

- Cluster Health Checks (on-demand, background)
- Diagnoses and remedies for common issues
- Compliance with Intel® Cluster Ready spec

Simplifies Cluster Computing Platforms

- Reduces need for specialized expertise
- Enables cluster health checks by applications
- Extensible and customizable, API

Data Collectors

Cluster Database

Diagnostic Data

Analysis Checking for Issues

Results

Expert System

Suggesting Remedies
Intel® Cluster Checker 3.0 – Concept

Expert system concept

- Symptoms are subjective indications of health
- Signs are objective indications of health detected by direct observation
- Diagnoses are the identification of the root cause of an issue
- Remedies are methods to resolve an issue

<table>
<thead>
<tr>
<th>Concept</th>
<th>Human</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptom</td>
<td>I am nauseous and fatigued</td>
<td>My job is running slow</td>
</tr>
<tr>
<td>Signs</td>
<td>Dehydrated, Fever, Nauseous</td>
<td>DGEMM performance on nodeX is 25% of peak</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Zombie process on nodeX is using 100% cpu</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>Flu</td>
<td>Zombie process is stealing cycles</td>
</tr>
<tr>
<td>Remedy</td>
<td>Drink plenty of clear fluids, take 2 aspirin, and bed rest</td>
<td>Kill the zombie process</td>
</tr>
</tbody>
</table>
## Intel® Cluster Checker 3.0 – Features

Cluster systems expertise packaged into a utility
Increase productivity assuring the cluster works.

### Functionality
- ✓ Confidence tool for non-experts
- ✓ Checks common points of failure
- ✓ Verification of application environment and system interfaces
- ✓ Checks cluster functionality, uniformity, and performance
- ✓ Extendable, rule based expert system
- ✓ On-demand or background mode

### Features
- ✓ Support for Intel® Xeon® and Xeon® Phi™
- ✓ Ethernet*, InfiniBand*, Intel® True Scale Fabrics
- ✓ Standard performance tests (DGEMM, IMB, HPL, STREAM, ...)
- ✓ Command line and API use
- ✓ Recording data for remote support
- ✓ *Installs with Intel® Parallel Studio XE 2016 Cluster Edition for Linux*
Intel® Cluster Checker 3.0 – Installation

- Easy to install – install.sh (script), or install_GUI.sh (graphical interface)
- Standalone, or part of ‘Intel® Parallel Studio XE 2016 Cluster Edition for Linux*’
## Intel® Cluster Checker 3.0 – Operation

### Getting Started – Two ways to collect data

<table>
<thead>
<tr>
<th>Step 1 – Input</th>
<th>On-demand</th>
<th>Background (optional)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create a node file</td>
<td></td>
<td>Configure and start service *</td>
</tr>
</tbody>
</table>

| Step 2 – Measure     | Run one sweep of tests (command line tool) | Tests are run periodically in background by daemons |

| Step 3 – Result/Activity | Analyze pass/fail results with diagnostic information |

* requires root privilege
Quickstart for on-demand execution mode

- Install Intel® Cluster Checker 3.0
- Create a node list
  - One line per node
  - Node role can be given with the “# role: “ tag (optional), e.g.
    - master # role: head
    - node00 # role: compute
- Collect the data – execute “clck-collect”:
  - $ source /opt/intel/clck/3.0.X.XXX/bin/clckvars.[c]sh
  - $ clck-collect -a -f <full_path_to_node_list>
- Execute the Analyzer:
  - $ clck-analyze
Intel® Cluster Checker 3.0 – Operation

Terminology used in diagnostic output

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undiagnosed Signs</td>
<td>Issues with no diagnosis.</td>
</tr>
<tr>
<td>Diagnosed Signs</td>
<td>Issues that contributed to diagnosis</td>
</tr>
<tr>
<td>Diagnoses</td>
<td>Potential root cause of an issue. (Rule-based expert system typically combines one or more findings to reach a diagnosis)</td>
</tr>
<tr>
<td>Confidence</td>
<td>Level of certainty that a sign / diagnosis is correct (0 to 100%)</td>
</tr>
<tr>
<td>Severity</td>
<td>Level of seriousness of a sign / diagnosis (0 to 100%)</td>
</tr>
</tbody>
</table>
Intel® Cluster Checker 3.0 – Operation

Discover a functional problem and understand its root cause

```
[1cr@yoda ~]$ click-analyze
Intel® Cluster Checker 3.0.0 Beta Update 2 (build 20150822)
Running connector extensions... done
Analyzing: 100%, rules completed/remaining: 560/0
Nodes being tested:
compute-0-0.local, compute-0-1.local, compute-0-2.local, compute-0-3.local,
compute-0-4.local, compute-0-5.local, compute-0-6.local,
yoda.clusterlab.intel.com.
1 diagnosed:
  1. The software image and/or configuration is not consistent with the rest of the cluster.
     [node(s),severity,confidence,remedy]
     [compute-0-0.local, 90, 75]
     [compute-0-1.local, 90, 75]
     [compute-0-2.local, 90, 75]
     [compute-0-3.local, 90, 75]
     [compute-0-4.local, 90, 75]
     [compute-0-5.local, 90, 75]
     [compute-0-6.local, 90, 75]
     [yoda.clusterlab.intel.com, 90, 75]
0 undiagnosed signs:
This analysis took 2.9035 seconds.
```
Intel® Cluster Checker 3.0 – Operation

Encounter a cluster performance issue and know its results

```
[icr@jargon ~]$ clock-analyze
Intel(R) Cluster Checker 3.0.0 Beta Update 2 (build 20150522)
Running connector extensions... done
Analyzing: 97%, rules completed/remaining: 208/6  Analyzing: 97%, rules
completed/remaining: 208/6  Analyzing: 97%, rules completed/remaining: 208/6
Analyzing: 100%, rules completed/remaining: 687/0
Nodes being tested:
compute-0-0-mic0.local, compute-0-0.local, compute-0-1-mic0.local,
compute-0-1.local, compute-0-2-mic0.local, compute-0-2.local,
compute-0-3-mic0.local, compute-0-3.local, jango.clusterlab.intel.com.
```

0 diagnoses:

1 undiagnosed signs:

1. The latest DGEMM result of 1007.58 GFLOPS is only 76.06 percent of the
theoretical peak of 1324.80 GFLOPS

[node(s): compute-0-2.local]
[severity = 50]
[confidence = 99]

This analysis took 5.5949 seconds.
Intel® Cluster Checker 3.0 – Operation

Everything fine - group of nodes is validated OK, ready to run applications

$ click-analyze
Intel® Cluster Checker 3.0.0 Beta Update 2 (build 20150522)
Running connector extensions... done
Analyzing: 100%, rules completed/remaining: 560/0
Nodes being tested:

0 diagnostics:
0 undiscovered signs:
This analysis took 2.8755 seconds.
Showing specific test data

Using clckdb to show specific test/node data

```
$ clckdb --hostname compute-0-3.local --provider uname
rcwid: 421
Provider:  uname
Hostname:  compute-0-3.local
num_nodes:  1
node_names:
Exit_Status:  0
Timestamp:  1432729262 : Wed May 27 12:21:02 2015 UTC
Duration:  0
Encoding:  0
Stdout_size:  553
Kernel-name:  Linux
hostname: compute-0-3.local
kernel-release:  2.6.32-431.11.2.el6.x86_64
kernel-version:  #1 SMP Tue Mar 25 19:59:55 UTC 2014
machine:  x86_64
processor:  x86_64
hardware-platform:  x86_64
operating-system:  GNU/Linux
type:  Linux compute-0-3.local 2.6.32-431.11.2.el6.x86_64 #1 SMP Tue Mar 25 19:59:55 UTC 2014 x86_64 x86_64 x86_64 GNU/Linux
Stderr_size:  0
OptionID:  None
```
Intel® Cluster Checker 3.0

Concept of Operation
Intel® Cluster Checker 3.0

Distributed architecture

Intel® Cluster Checker – User Interface

- ‘clck’ command (user/root)
- API for custom interfaces
- ISV applications
- Cluster management system
- Resource management system
- Provisioning system

Cluster

- User Interface ‘clck’
- Database
- Rule based expert system
- API

On-demand mode (Background mode)

Test data providers (Background ‘clkdd’)

Optimization Notice

Copyright © 2015, Intel Corporation. All rights reserved. *Other names and brands may be claimed as the property of others.
## Intel® Cluster Checker 3.0

### Architecture elements

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Intel® Cluster Checker 3.0</th>
<th>Component /opt/intel/clck_latest/ (default)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front End</td>
<td>User Interface</td>
<td>bin/clck-collect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bin/clck</td>
</tr>
<tr>
<td></td>
<td>Database</td>
<td>~/.clck/3.0.0/clck.db</td>
</tr>
<tr>
<td></td>
<td>Expert system</td>
<td>kb/</td>
</tr>
<tr>
<td></td>
<td>API</td>
<td>include/</td>
</tr>
<tr>
<td>Test data providers</td>
<td>Test data providers</td>
<td>providers/</td>
</tr>
<tr>
<td>'clckd' background daemon</td>
<td>'clckd' background daemon</td>
<td>bin/clckd</td>
</tr>
</tbody>
</table>
Expert system concept

- Symptoms are subjective indications of health
- Signs are objective indications of health detected by direct observation
- Diagnoses are the identification of the root cause of an issue
- Remedies are methods to resolve an issue

<table>
<thead>
<tr>
<th>Concept</th>
<th>Human</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptom</td>
<td>I am nauseous and fatigued</td>
<td>My job is running slow</td>
</tr>
<tr>
<td>Signs</td>
<td>Dehydrated, Fever, Nauseous</td>
<td>DGEMM performance on nodeX is 25% of peak</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Zombie process on nodeX is using 100% cpu</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>Flu</td>
<td>Zombie process is stealing cycles</td>
</tr>
<tr>
<td>Remedy</td>
<td>Drink plenty of clear fluids, take 2</td>
<td>Kill the zombie process</td>
</tr>
<tr>
<td></td>
<td>aspirin, and bed rest</td>
<td></td>
</tr>
</tbody>
</table>
Intel® Cluster Checker 3.0

API allows integration into other software

- An application can check its cluster environment, health, functionality
- A monitoring system can display ‘clck’ status information
- A deployment system can configure or trigger ‘clck’ data collection/analysis
- A resource manager can control ‘clck’ background execution
- A job scheduler can validate node groups before applications run
How to program the API – e.g. C++ sample code snippets

INPUT AND CONFIGURATION

// INPUT AND CONFIGURATION

ANALYSIS

// ANALYSIS

RESULTS PROCESSING

// RESULTS PROCESSING

// set up filters: confidence, severity, nodes, types, etc.
clk::Layer::Filter filter;

// set up sorting order
std::vector<clk::Layer::Sorting> sorting;

// signs and diagnoses (filtered and sorted)
std::vector<std::shared_ptr<clk::Fault>> faults = layer.get_faults(filter, sorting);

// process signs and diagnoses
for (auto &fault : faults) {}
Where to get Intel® Cluster Checker

Software

- Install with Intel® Parallel Studio XE 2016 Cluster Edition for Linux*, free 30-day evaluation available
- Find pre-installed with cluster systems shipped with ‘Intel® Cluster Ready’ certification

Support

- http://premier.intel.com – the Intel software support portal

Further information

- http://www.intel.com/go/cluster - all details on Intel® Cluster Ready program, partners, and products
Legal Disclaimer & Optimization Notice

INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS”. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO THIS INFORMATION INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

Copyright © 2015, Intel Corporation. All rights reserved. Intel, Pentium, Xeon, Xeon Phi, Core, VTune, Cilk, and the Intel logo are trademarks of Intel Corporation in the U.S. and other countries.

Optimization Notice

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Notice revision #20110804