
# Hortonworks Data Platform Reference Architecture

A PSSC Labs Reference Architecture Guide
December 2014

20432 North Sea Circle, Lake Forest, CA 92630 * www.pssclabs.com * tel: 949-380-7288 * fax: 949-380-9788

PSSC LABS

# Introduction

PSSC Labs continues to bring innovative compute server and cluster platforms to market. Focusing on specific applications where performance and reliability are critical, highlights PSSC Labs strengths. The Apace Hadoop data framework requires substantial compute and storage capabilities coupled tightly together. Hortonworks Data Platform adds layers of unification, security and management.

PSSC Labs is the first manufacturer to design a server platform specifically for Hadoop. Introducing the World's highest density, lowest power consuming, Enterprise Ready Big Data server platform designed specifically for Hadoop workloads. The Big Data D12000 offers the absolute highest possible compute and storage density combined with high performance Data IO throughput. PSSC Labs already delivers the Big Data D12000 for small POC to large production clusters spanning hundreds of nodes.

## Key Features:

- Reduce Data Center Footprint By 50%
- Reduce Power Consumption By 40%
- Nearly 50% Greater Data IO Rates
- Patent Pending Tool Free Maintenance

## Technical Specifications:

- Up to 12 x 3.5" SATAIII or 14 x SSDs in 1U Rack Space
  - 72 TB using twelve x 6 TB Hard Drives
- Supports UP to 2 x Intel® Xeon® E5 Series Processors
- Supports Up to 256 GB ECC Enterprise Memory
- 10 GigE, 40 GigE and Infiniband Network Support
- Red Hat®, CentOS, Ubuntu, MS Windows® Compatible

"We welcome PSSC Labs to the Hortonworks Technology Partner Program. Their unique server platforms are already being deployed in Apache Hadoop environments and we look forward to deepening our relationship to help joint customers achieve even greater performance from their big data deployments."

John Kreisa
Hortonworks®, Vice President

Hortonworks CERTIFIED TECHNOLOGY PARTNER

20432 North Sea Circle, Lake Forest, CA 92630 * www.pssclabs.com * tel: 949-380-7288 * fax: 949-380-9788

# Big Data D12000 Sample Configurations

Every organization or use case requires different computing needs. The Big Data D12000 offers the greatest flexibility possible. Below are three different proposed architectures for different workloads: high density storage, high computational requirements and a balanced configuration.

**Big Data D12000 High Density**

- 48 TB Total Storage
- 12 Xeon E5 Processor Cores
- 64 GB ECC Memory
- 2 x 10GigE Network Ports
- 2 x GigE Network Ports
- Remote IPMI Management
- CentOS Linux OS
- Power Draw Estimate
  215 Watt Idle / 275 Watt Full Load

**Big Data D12000 High Compute**

- 24 TB Total Storage
- 20 Xeon E5 Processor Cores
- 256 GB ECC Memory
- 2 x 10GigE Network Ports
- 2 x GigE Network Ports
- Remote IPMI Management
- CentOS Linux OS
- Power Draw Estimate
  265 Watt Idle / 345 Watt Full Load

**Big Data D12000 Balanced**

- 36 TB Total Storage
- 16 Xeon E5 Processor Cores
- 128 GB ECC Memory
- 2 x 10GigE Network Ports
- 2 x GigE Network Ports
- Remote IPMI Management
- CentOS Linux OS
- Power Draw Estimate
  230 Watt Idle / 315 Watt Full Load

## A Sample of Organizations Currently Using PSSC Labs Big Data D 12000 Servers

20432 North Sea Circle, Lake Forest, CA 92630 * www.pssclabs.com * tel: 949-380-7288 * fax: 949-380-9788

PSSC LABS

# HDP Rax: Hortonworks Data Platform Validated Turn Key Cluster Solutions

PSSC Labs offers a complete, turn-key cluster that is ready to run on delivery. PSSC Labs understands everything that is necessary for a successful deployment. All necessary hardware including servers, network equipment, power and infrastructure are included. PSSC Labs Cluster Engineers preconfigure network, storage, operating system and BIOS settings to the end user's specifications. Hortonworks Data Platform is installed at PSSC Labs factory. The final step is the running of sample data sets to ensure proper functionality and performance.

Below is an overview of each different server platform PSSC Labs offers for Hortonworks Data Platform turn-key deployments. Depending on the complexity of the environment, some software resources can be installed on different server platforms.

| **Big Data D12000 DATA NODE** | | |
|---|---|---|
| **Tech Specs** | **Key Features** | **Software Resource** |
| ○ 1U High Density Form Factor<br>○ 2 x Intel® Xeon® E5 Processors<br>○ 12 x SATAIII or SAS Hard Drives or 14 x SSDs<br>○ 12 TB to 72 TB Storage Capacity<br>○ 32 GB to 256 GB ECC Memory<br>○ 2 x GigE Network Adapters<br>○ Optional 10 GigE, 40 GigE, Infiniband Support<br>○ Dedicated IPMI / iKVM | ○ Enterprise Platform<br>○ Redundant Power Supply<br>○ Improved Data IO Throughput<br>○ 40% Reduction in Power Consumption<br>○ 2 x the Density of Standard Server<br>○ Flexible Configuration Options<br>○ 3 Year Warranty Included (24 x 7 x 365 NBD Available) | ○ DataNode Daemon<br>○ Ganglia Monitor<br>○ Region Server<br>○ Node Manager<br>○ Supervisor |

| **SURESTORE U2000 MANAGEMENT NODE (NAME NODE & SECONDARY NAME NODE)** | | |
|---|---|---|
| **Tech Specs** | **Key Features** | **Software Resources** |
| ○ 2 x Intel® Xeon® E5 Processors<br>○ 1 TB to 24 TB SATA III, SAS, SSD Hard Drives<br>○ 32 GB to 256 GB ECC Memory<br>○ 2 x GigE Network Adapters<br>○ Optional 10GigE, 40GigE, Infiniband Support<br>○ Dedicated IPMI / iKVM | ○ Enterprise Platform<br>○ Redundant Power Supply<br>○ Redundant Storage<br>○ Raid Levels 0,1,5,6,10,50<br>○ Flexible Configuration Options<br>○ 3 Year Warranty Included (24 x 7 x 365 NBD Available) | ○ App Timeline Server<br>○ DRPC Server<br>○ Ganglia Monitor<br>○ HDFS Client<br>○ NameNode / Secondary NN<br>○ Oozie Server / Client<br>○ Yarn Client<br>○ Zookeeper Server / Client<br>○ HDFS Client<br>○ MySQL Server |

| **SURESTORE U1000 EDGE NODE** | | |
|---|---|---|
| **Tech Specs** | **Key Features** | **Resources** |
| ○ 2 x Intel Xeon E5 Processors<br>○ 1 TB to 12 TB SATAIII, SAS, SSD Hard Drives<br>○ 32 GB to 256 GB ECC Memory<br>○ 2 x GigE Network Adapters<br>○ Optional 10GigE, 40GigE, Infiniband Support<br>○ Dedicated IPMI / iKVM | ○ Enterprise Platform<br>○ Redundant Power Supply<br>○ Redundant Storage<br>○ Raid Levels 0,1,5,6,10,50<br>○ Flexible Configuration Options<br>○ 3 Year Warranty Included (24 x 7 x 365 NBD Available) | ○ Hive Server / Client<br>○ Tez Client<br>○ Nimbus<br>○ Nagios Server |

PSSC LABS

# HDP Rax Turn-Key Cluster Sample Configurations

**HDP Rax 150**

- 250 TB Total Storage
- 2 Name Node
- 5 Data Nodes
- 60 Xeon E5 Processor Cores
- 320 GB ECC Memory
- 10GigE Network Backplane
- Remote IPMI Management
- CentOS Linux OS
- HDP Installation Service
- HDP Validation Service

**HDP Rax 500**

- 500 TB Total Storage
- 2 Name Nodes
- 1 Edge Node
- 10 Data Nodes
- 160 Xeon E5 Processor Cores
- 1280 GB ECC Memory
- 10GigE Network Backplane
- Remote IPMI Management
- CentOS Linux OS
- HDP Installation Service
- HDP Validation Service
- Rack & Roll Service

**HDP Rax 1500**

- 1500 TB Total Storage
- 2 Name Nodes
- 1 Edge Node
- 30 Data Nodes
- 480 Xeon E5 Processor Cores
- 3840 GB ECC Memory
- 10GigE Network Backplane
- Remote IPMI Management
- CentOS Linux OS
- HDP Installation Service
- HDP Validation Service
- Rack & Roll Service


"We believe strongly in our ability to deliver the highest performance, highest reliability server platforms to Hortonworks end users. Our experience delivering clusters ranging from several hundred TBs to several dozen PBs ensures a successful Hortonworks Data Platform deployment."

**Larry Lesser**
PSSC Labs, CTO

20432 North Sea Circle, Lake Forest, CA 92630 * www.pssclabs.com * tel: 949-380-7288 * fax: 949-380-9788

PSSC LABS

# Total Cost of Ownership Comparison

PSSC Labs goal is to offer solutions with the absolute lowest total cost of ownership. The below chart compares different server manufacturer's solution for a 1 Petabyte (raw) Hadoop environment. PSSC Labs HDP Rax 1000 requires 50% less rack space and consumes 40% less power

| | PSSC Labs HDP Rax 1000 for 1PB Total Storage Space | Dell Configuration for 1PB Total Storage | HP Configuration for 1PB Total Storage | Lenovo Configuration for 1PB Total Storage |
|---|---|---|---|---|
| Required Data Center Footprint | Single x 42U Rack | Two x 42U Rack | Two x 42U Rack | Two x 42U Rack |
| Power Consumption Estimate* | 4300 Watts Total @ Idle<br>5500 Watts Total @ Load | 5800 Watts Total @ Idle<br>8500 Watts Total @ Load | 6000 Watts Total @ Idle<br>8800 Watts Total @ Load | 5700 Watts Total @ Idle<br>8700 Watts Total @ Load |
| Required Power Circuits | Single x 30 Amp / 208V / Single Phase | Two x 30 Amp / 208V / Single Phase | Two x 30 Amp / 208V / Single Phase | Two x 30 Amp / 208V / Single Phase |
| Pre-installation and Validation of Hortonworks Data Platform at Factory | Yes. Hortonworks Data Platform preinstalled and tested. | No. Additional services and fees required. | No. Additional services and fees required. | No. Additional services and fees required. |
| Onsite Physical Installation | Yes. Cluster arrives pre-racked, cabled and labeled. | No. Additional services and fees required. | No. Additional services and fees required. | No. Additional services and fees required. |
| Cluster Management Training | Yes. | No. Additional services and fees required. | No. Additional services and fees required. | No. Additional services and fees required. |
| Dedicated Remote Monitoring Capabilities | Yes. IPMI 2.0 Network Standard | Yes. | Yes. | Yes. |
| Hardware Warranty | 3 Year NBD Service Available. | 3 Year NBD Service Available. | 3 Year NBD Service Available. | 3 Year NBD Service Available. |

*Dell, HP and Lenovo power estimates based on manufacturers website power draw estimates.