Load Balancing and Clustering Architecture
Statement of Confidentiality, Disclaimer and Copyright

This document contains information that is proprietary and confidential to EMTPL, which shall not be disclosed, transmitted, or duplicated, used in whole or in part for any purpose other than its intended purpose. Any use or disclosure in whole or in part of this information without the express written permission of EMTPL is prohibited.

Any other company and product names mentioned are used for identification purpose only, may be trademarks of their respective owners and are duly acknowledged.

Disclaimer

This document is intended to support administrators, technology managers or developers using and implementing Smarten. The business needs of each organization will vary and this document is expected to provide guidelines and not rules for making any decisions related to Smarten. The overall performance of Smarten depends on many factors, including but not limited to hardware configuration and network throughput.
## Contents

1. Introduction ................................................................. 4
2. Smarten Load Balancing on AWS Cloud .............................. 4
3. Smarten Integration with Spark cluster on AWS Cloud ........ 6
4. Smarten Load Balancing for On-Premises Infrastructure .......... 7
5. Smarten Integration with Spark Cluster for On-Premises Infrastructure .... 8
6. Product and Support Information ........................................ 9
1 Introduction

Load balancing and clustering is supported in Smarten for distributing workloads across multiple computing resources.

Load balancing and clustering will optimize use of resources, increase throughput, minimize response time, and avoid overload. With Managed Memory Computing, and distributed load, response time will reduce and user experience will improve. It will also provide failover mechanism and ensure high availability.

2 Smarten Load Balancing on AWS Cloud

Key components

- **AWS Elastic Load balancer (ELB)** – Requests from clients will be routed through AWS Elastic Load balancer. You can configure load balancing rules to distribute load on ELB.

- **Auto scaling** – If Auto scaling is enabled, it monitors the load on EC2 instance and based on the defined rules, it will add or remove instance under load balancer.

- **Smarten server nodes** – As per the requirement, multiple Smarten nodes can be configured in EC2 and setup with load balancer. Smarten metadata database and data files will be shared between these server nodes.
- **Smarten metadata database** – Smarten metadata database can be configured in AWS RDS (MySQL) instance. It will be shared across the all Smarten instances under load balancer.

- **Smarten Data files** – Smarten data files folder can be configured in AWS EFS and can be shared across the all Smarten instances under the load balancer.
3 Smarten Integration with Spark cluster on AWS Cloud

Key components

- **Smarten server node** – As per the requirement, Smarten node can be configured in EC2. Smarten metadata database and data files are configured with this node. AWS EMR Spark cluster is also integrated with this node and all Smarten application processing will be done on Spark cluster.

- **Smarten metadata database** – Smarten metadata database can be configured in AWS RDS (MySQL) instance. It is shared with the Smarten instance and all nodes of Spark cluster.

- **Smarten Data files** – Smarten data files folder can be configured in AWS EFS and is shared with the Smarten instance and all Spark cluster nodes.

- **Spark Cluster** – AWS EMR provides YARN based Spark cluster. Smarten can be integrated with EMR Spark cluster and it can leverage power of Spark cluster for application data processing. You can enable auto scaling feature on EMR Spark cluster to add or remove worker nodes in cluster.
4 Smarten Load Balancing for On-Premises Infrastructure

Key components

- **Load balancer** – Requests from clients will be routed through the load balancer. There can be software or hardware load balancer which is configured with load balancing rules to distribute load. For software load balancer, it can be configured using Apache web server which provides load balancing mechanism.

- **Smarten server nodes** – As per the requirement, multiple Smarten nodes can be configured and setup with the load balancer. Smarten metadata database and data files will be shared between these server nodes.

- **Smarten metadata database** – Smarten metadata database can be configured in MySQL database. It will be shared across all Smarten instances under load balancer.

- **Smarten Data files** – Smarten data files folder can be configured as a shared network drive and can be shared across all Smarten instances under the load balancer.
5 Smarten Integration with Spark Cluster for On-Premises Infrastructure

Key components

- **Smarten server node** – As per the requirement, Smarten node can be configured. Smarten metadata database and data files are configured with this node. Spark standalone cluster is also integrated with this node and all Smarten application processing will be done on Spark cluster.

- **Smarten metadata database** – Smarten metadata database can be configured in MySQL database. It is shared with the Smarten instance and all nodes of Spark cluster.

- **Smarten Data files** – Smarten data files folder can be configured as shared network drive and is shared with the Smarten instance and all Spark cluster nodes.

- **Spark Cluster** – You can setup Spark standalone cluster in your local environment. Smarten can be integrated with Spark cluster and it can leverage power of Spark cluster for application data processing.
6  Product and Support Information

Find more information about Smarten and its features at www.smarten.com

Support: support@smarten.com
Sales: sales@smarten.com
Feedback & Suggestions: support@smarten.com
Support & Knowledgebase Portal: support.smarten.com