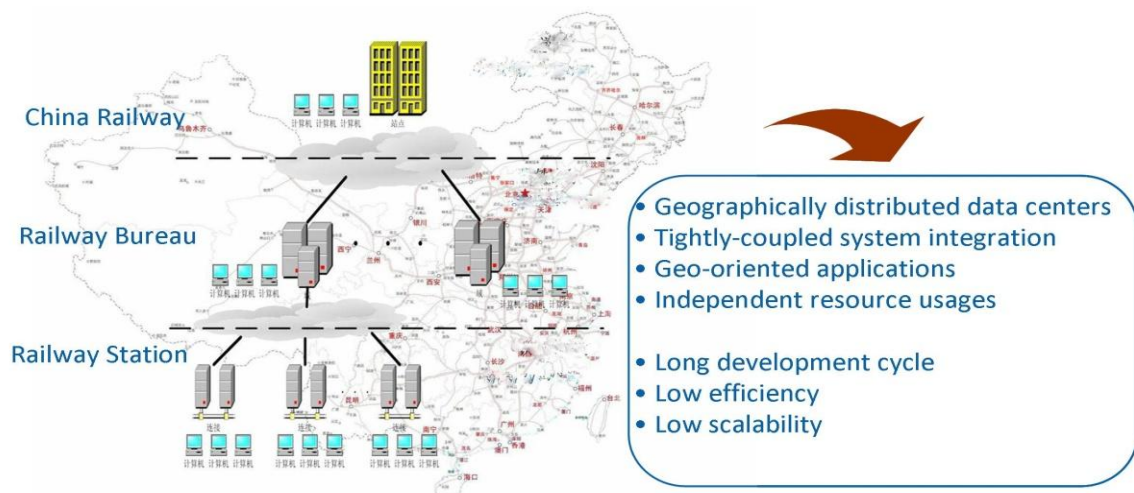




A Cloud Test Bed for China Railway Enterprise Data Center

BACKGROUND

China Railway consists of eighteen regional bureaus, geographically distributed across China, with each regional bureau having their own regional Data Center. Physically distributed and managed at a regional level, the information system of China Railway is a highly complex distributed mission critical system, integrated in networks but largely isolated in term of system resources with application silos, which is a key barrier for China Railway to drive its digital railway vision that is based on fully shared compute infrastructure and applications. The following diagram illustrates the 3 levels in the information system of China Railway.



Overall, the current information system of China Railway has the following key characteristics:

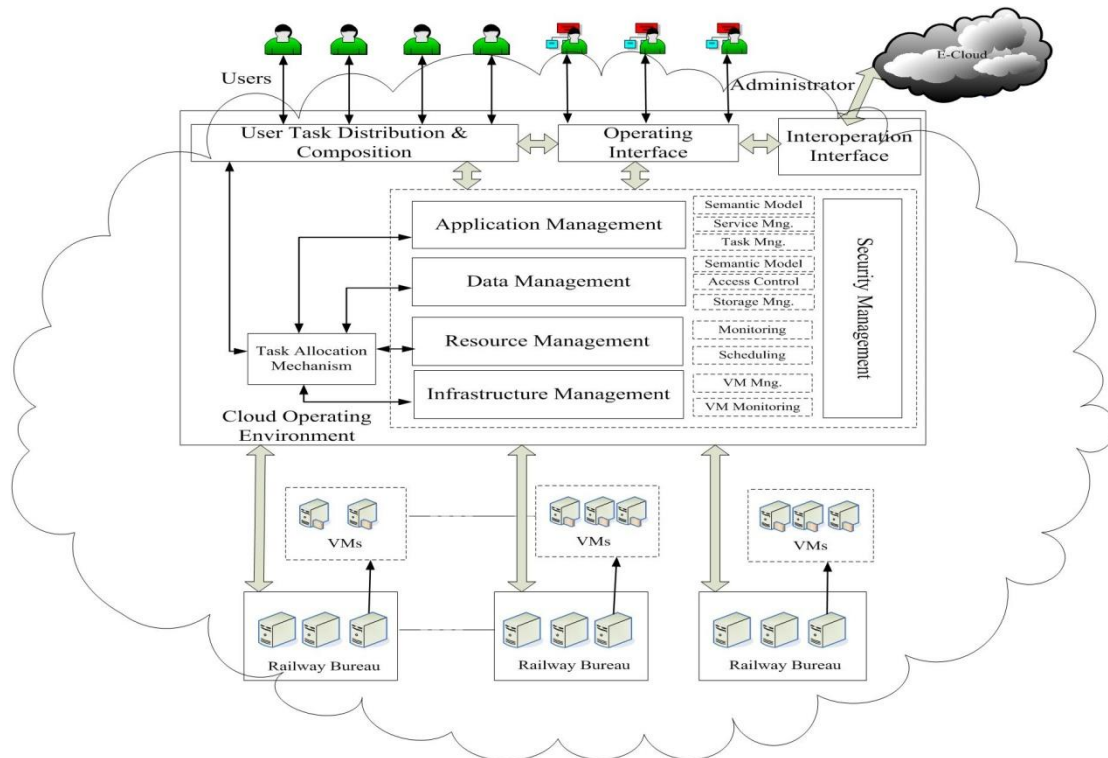
- ◆ Geographically distributed data centers
- ◆ Tightly-coupled system integration
- ◆ Geo-oriented applications
- ◆ Independent resource usages
- ◆ Multi-level usages and different user types
- ◆ Multiple application subsystems.

As an example, the railway freight transportation system manages the freight ticket information in China Railway. It consists of freight ticket sub-system, confirmation and reporting sub-system, and train dispatching sub-system. A freight ticket contains information of freight ID, source and destination information, and detailed status information of freight once it is registered for transportation. There are 120,000 freight tickets generated every day in China Railway, each is 4KB in size. The tracking period is normally 45 days, and the ticket needs to be stored online for one year. Therefore, 160GB data is generated per day and the amount of data in ready-to-be-searched state is over 1 Terabytes. In addition, 4-6 replicas of the data are required, must be considered.



CLOUD TEST BED FOR CHINA RAILWAY

In this case study, we have investigated how cloud technology could be used at Enterprise Datacenter in the business environment of China Railway by building a Cloud test bed that simulates the information system of China Railway and running the Railway applications and data in this test bed. The diagram below shows the architecture of our Cloud test bed for China Railway.



This test bed takes consideration of the following the business requirements of China Railway.

User Types and Administration Level: Railway users fall into two different categories: a customer or an administrator. A customer is a general user using a particular service provided by China Railway. An administrator manages the applications, data, resources, and the compute infrastructure according to different levels of authorization and administration inside China Railway.

User Task composition: User can utilize the application semantic model and cloud services (provided by cloud computing environment) to set up a user task. Moreover, the data must conform to the semantic constraint with each user task for improving efficiency and application consistency.

Interoperability Interface: Seamless interoperability between I-cloud and E-cloud.

Infrastructure Management: Virtual machine management and OS/application image management support dynamic resource provisioning and manage physical and virtual resources, such as creation, deployment, termination of virtual machines, instance creation and deployment of a computing service image.



Resource Management: Resources include the physical resources (computing, storage, and network etc) and virtual resources (user-oriented resource utility management). Resource Management performs monitoring, dynamic provisioning, and load balancing and usage optimization of cloud computing resources.

Data Management: Manage data storage, data access and data reliability. Data storage includes data semantic descriptions and distributed data storage allocation. Data access includes access scheduling, access control. Multi-level MOR users have different data access rights. The goal of access scheduling is to achieve data access efficiency at different locations.

Application Management: Manages execution environment, including application semantic relation, execution of cloud computing services, data access operations, and so on.

Task Assignment Mechanism: Cloud computing application is a distributed service task. The task assignment must implement a comprehensive optimization, considering semantic and performance requirements.

Security Management: Security is critical to a cloud environment, especially for mission critical railway operations. Different requirements exist at several levels in the cloud.

FEATURES OF THIS CLOUD TEST BED

1) Resource Management

Cloud computing has to be scalable to large system application. Virtualization technology is used for this cloud test bed. We used Tashi from the Intel Research Lab Intel at Carnegie Mellon University as the cluster management system for cloud computing. . Users of Tashi are able to create collections of virtual machines that run on the cluster's physical resources. These virtual machine collections form "virtual clusters" that operate on the Big Data and host services consumed by clients

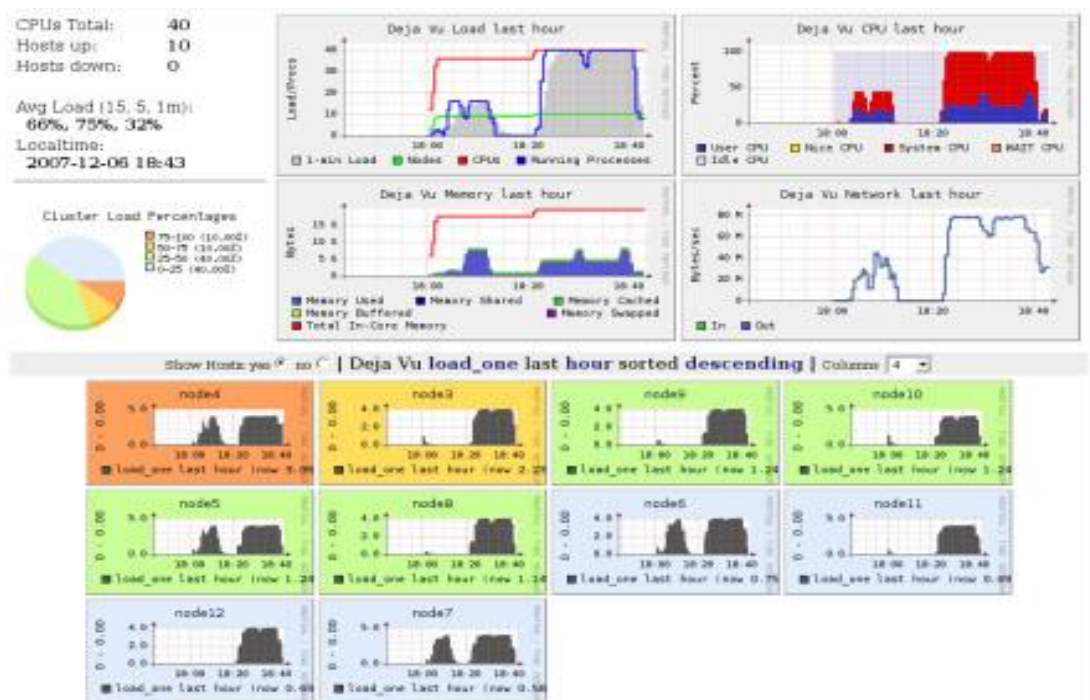
2) Dynamic Cloud Service Configuration

Cloud service customization depends on application data, cloud resources, service configuration and application semantic requirements. According to the railway data characters and application semantic model, cloud dynamic allocate computing resource, and implement relative data management operation. According to administration policy, cloud setup service configuration, and provide railway application service for different users.



3) Cloud Resource Monitoring

Cloud resource monitoring manage all kinds of physical resources (CPU, Memory, Network, etc.), cluster configuration status, service execution status, etc. The status information provides the basis for task dispatching, resource allocation.



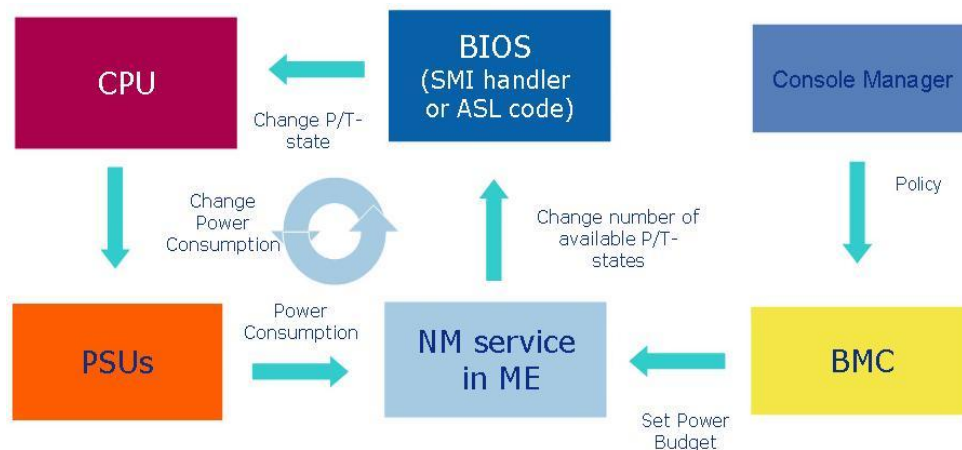


4) Power Management

Intel® Intelligent Power Node Manager (Node Manager) is an out-of-band (OOB) power management policy engine embedded in Intel Xeon 5500 server chipsets. Processors can adjust and performance throughput their power consumption through the manipulation of the P- and T-states. Node Manager works with the BIOS and OS power management (OSPM) to dynamically adjust platform power to achieve optimal performance and power balance on a single node. Node Manager has the following features:

- ◆ **Dynamic Power Monitoring:** Measures actual power consumption of a server platform within acceptable error margin of +/- 10%.
- ◆ **Platform Power Capping:** Sets platform power to a targeted power budget while maintaining maximum performance for the given power level. Node Manager receives power policy from an external management console through IPMI interface and maintains power at targeted level.
- ◆ **Power Threshold Alerting:** Node Manager monitors platform power against targeted power budget. When the target power budget cannot be maintained, Node Manager sends out alerts to the management console

Node Manager works as a closed control loop system, the following figure gives out the general idea of the power control behavior for a single server.

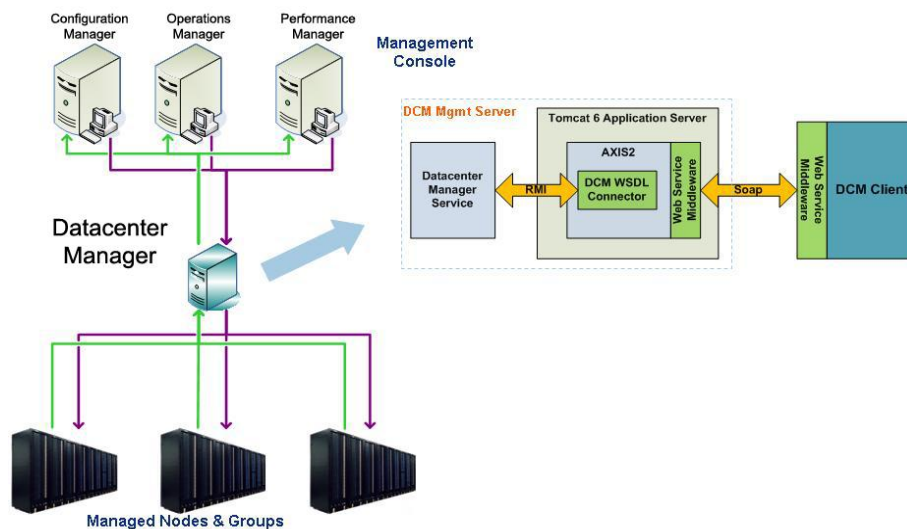


Intel® Data Center Manager is software technology that provides power and thermal monitoring and management for servers, racks and groups of servers in datacenters. It builds on Intel® Intelligent Power Node Manager and customers existing management consoles to bring platform power efficiency to End Users. Data Center Manager implements group level policies that aggregate node data across the entire rack or data center to track metrics, historical data and provide alerts to IT managers. This allows IT managers to establish group level power policies to limit consumption while dynamically



adapting to changing server loads. The wealth of data and control that Data Center Manager provides allows data centers to increase rack density, manage power peaks, and right size the power and cooling infrastructure. It is a software development kit (SDK) designed to plug-in to software management console products. It also has a reference user interface which was used in this POC as proxy for a management software product. Key Intel® Datacenter Manager features are:

- ◆ Group (server, rack, row, PDU and logical group) level monitoring and aggregation of power and thermals
- ◆ Log and query for trend data for upto one year
- ◆ Policy driven intelligent group power capping
- ◆ User defined group level power alerts and notifications
- ◆ Support of distributed architectures (across multiple racks)
- ◆ The following figure shows the architecture of the Node Manager enabled IPDC servers which are managed by Intel DCM. Here the Datacenter Manager monitors the actual power consumptions on each server and aggregated total power consumption at the rack level. It works with Node Manager to set appropriate policies which ensures servers in a rack are delivering the best performance within the rack-level power budget. Administrator can connect to DCM server via DCM client to monitor the power consumption, query the event log, set the power policies and etc.



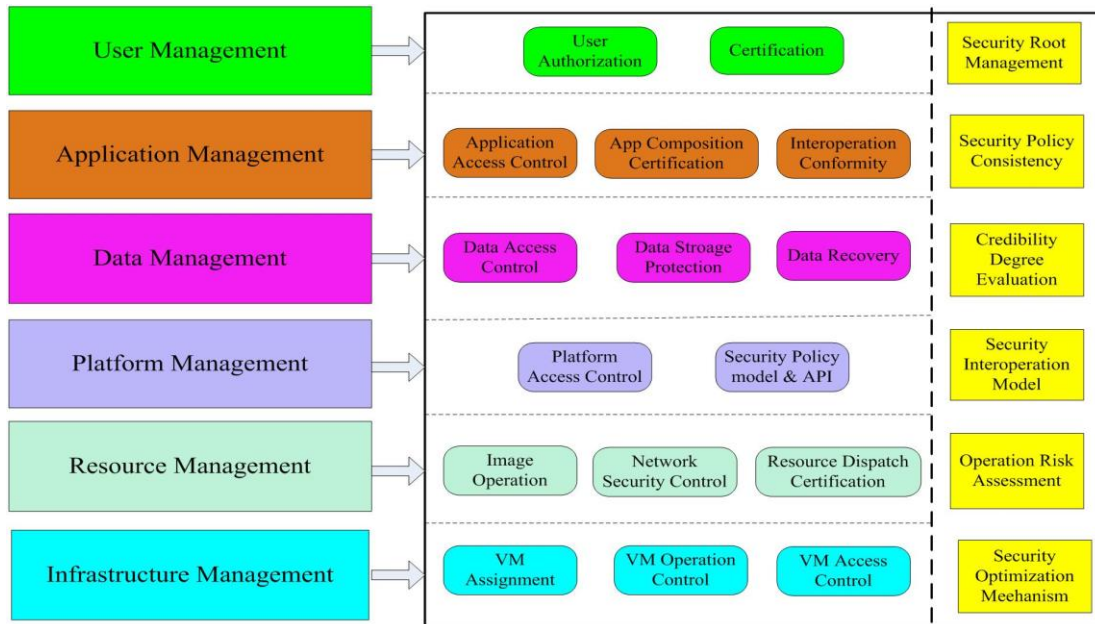
5) Security Framework and Application Protection

Security Framework

Cloud security is hierarchical and complex. It includes special security problems at different levels:



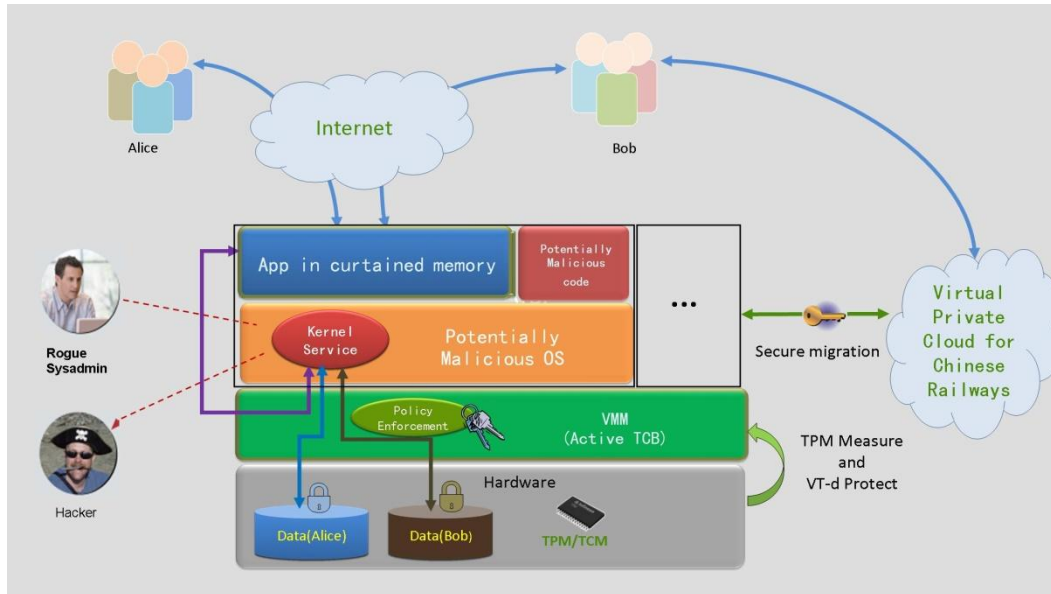
user, application, data, platform, resource level and infrastructure. According to cloud environment requirements, the integrated security includes security root management, security policy consistency, trust attestation, interoperability model, operation risk assessment, and security optimization mechanism to implement the cross-level security policy for system security and scalability security.



- ◆ Securing multi-tenant off-premises computing
- ◆ Protecting privacy and integrity of user's code/data
- ◆ Cloud resource security operation
- ◆ Infrastructure for security management
- ◆ Cross-level security policy and technology

System Implementation

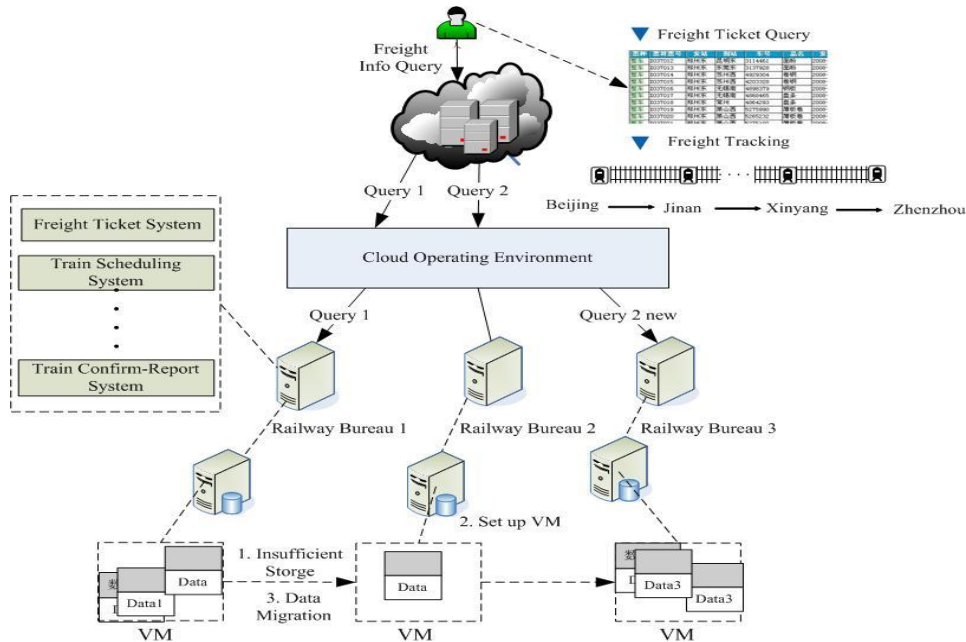
In our cloud test bed for China Railway, we utilize the TPM security root management technology to guarantee computing environment security, which is base of application level execution security, such as service semantic, service I/O operation and resource allocation.



- ◆ Security Root Management
 - Intel-TXT/TCG-TPM measured and Intel-VT-d protected Trusted Computing Base (TCB)
- ◆ Virtual Machine Isolation & Access Control
 - Protected VMM playing the role of a TCB as the most privileged servicing software
 - The TCB interposing between protected apps and untrustworthy OSs or peripherals to guard against potential damage in their services:
 - ✓ Memory Curtaining
 - ✓ IO Sealing

CHINA RAILWAY APPLICATIONS RUNNING ON CLOUD TEST BED

We utilize the Tashi cloud middleware to manage the virtual machine resource, including CPU, memory and storage space of every physical machine within cluster management. We use Hadoop distributed files system (HDFS) provides data management mechanism. Based on the HDFS, MapReduce distributed programming model provides parallel search mechanism for processing a large-scale data. Resource management integrates the application semantic, infrastructure characters, and physical resource to implement resource allocation and deployment. With the help of those system supporting technologies, application-oriented services can be developed for the different users.



Freight Ticket Search

Utilizing the distributed data from each railway bureau, it can in parallel search data conforming to specific semantic characters. Data sources that a new freight search relies on can be determined according to business routine in the railway and freight transportation. The distributed programming model – MapReduce used to map a search job to multiple computers for data search, and produce search result by integrating the semantic features of the railway freight system. The railway freight system involves many railway subsystems, such as freight information sub-system, train scheduling sub-system, train confirm report sub-system, etc. Through integrating those heterogeneous dataset, we can retrieve more detailed freight transportation information.

Data Storage Expansion

When specific railway bureau’s storage has shortage, the cloud computing environment can provide services to dynamically expand storage capacity at a different location – Tashi sets up a VM with required computing, storage capacity, and image instance of Hadoop datanode. This newly created VM and data storage could be additional capability for the railway bureau to deal with a sudden increase of business demand and guarantee the freight searching at all time.

SUMMARY

In this effort, we have realized that an Enterprise Datacenter built with cloud technologies can offer key advantages including Efficiency, Scalability, Reliability, and Flexibility. The value of a cloud based enterprise data center for China Railway is in no doubt quite significant and promising for China



Railway. We have also realized, however, our work is the beginning of a much bigger effort, which includes not only overcoming technology challenges but also working towards solutions for gaps in government policy, enterprise organizational challenges, business rules and practices, standardization, etc.

CONTACT INFORMATION

More information about this project or any technical aspects of the test bed, please contact anyone below.

Beijing Jiaotong University:

Professor Feng Liu. Email: fliu@bjtu.edu.cn

Professor Baopeng Zhang. Email: bpzhang@bjtu.edu.cn

Intel Corporation:

Kai X. Miao. Email: kai.miao@intel.com

Jackson He. Email: Jackson.he@intel.com

Xinran (Ryan) Wu. Email: ryan.x.wu@intel.com

EMC Research:

Wenbo Mao. Email: mao_wenbo@emc.com