Sample Project: Batch Factory

Code	IT1402
Programme	TRAIN-PTES
Department	IT
Responsible	43253 - Dr. Ulrich Schwickerath
Created by	34019 - Mr. Juan Guijarro
Updated by	23616 - Dr. Miguel Marquina
Date Created	21-MAR-12
Date updated	12-JUN-14

Title

Batch Factory

Description

Batch Factory (IT-PES project #2): One of the key computing services at CERN is the central network batch system. It is currently running on 4000 servers, physical and virtual, which conform an IBM/Platform LSF facility. As we are moving towards a cloud-computing model based on OpenStack, the number of virtual worker nodes is currently increasing with the goal of reaching several tens of thousands.

Management tools are necessary to manage a cluster of this scale in a flexible way,as well as to embrace new advantages offered by the cloud computing model. These tools would allow to create, and assign roles to, virtual machines. They would be used to drain and remove large numbers of worker nodes and they would ensure the stability of the entire cluster by retiring faulty hosts and maximizing the capacity by making an optimal usage of available resources allocated for this service.

This project is an excellent opportunity to learn about modern techniques used within both the field of Cloud Computing as well as the field of Configuration Management. The project also requires taking a deep in sight into the largest cluster at CERN and the state of the art Large Scale System Administration practices which are followed to manage it. The horizontal service management approach used within the IT Department implies doing an integration type of work which requires a broad range of skills and is likely to be defined by a breadth of knowledge rather than a depth of knowledge. These skills are likely to include system administration, software engineering, interface protocols, and general problem solving skills. It is likely that the problems to be solved have not been solved before except in the broadest sense. They are likely to include new and challenging problems with an input from a broad range of engineers where the system integration engineer, i.e. the selected candidate for this project, "pulls it all together."

A DevOps approach has been chosen for development and operations, which is the best option to the interdependence of software development and IT operations. It aims at helping us to rapidly produce computing services.

Requirements: Expertise in Python development and good knowledge of the Linux operating system. Experience of batch systems and cloud software is desirable but not required.

Skills

Databases: NoSQL. Information Technologies: Developing distributed computing systems (e.g. clusters, batch systems), Developing with virtualised infrastructure (e.g. Openstack), System administration (e.g. with GNU/Linux, Microsoft Windows, Networks), Using

software development tools (e.g. Git, Jira, Trac). Programming Languages: Python, Shell Script

Disciplines

Information Technologies

To edit this project go to https://hrapps.cern.ch/auth/f?p=131:4:::::P4_ID:1402