PJAS Selection - Autumn 2016

Group/Section	Project name	Project description	Required skills
CM-RPS	Resource management automation	CERN runs an OpenStack production cloud with over 190,000 cores across two data centers on 7,500 hypervisors to provide the compute resources for the Large Hadron Collider and other experiments at CERN. The continued growth of the cloud infrastructure in terms of use cases and cores requires automation of manpower intensive tasks. The project will design and implement workflows for the management of the resources such as migrations of virtual machines, workflows for hardware repair and expiration of Personal virtual machines to improve utilization	Linux administration. Python or Puppet would be desirable
ST-FDO	CERNBOX TESTING	CERNBOX is a service running at CERN to provide at the same time synchronisation services (based on the OwnCloud software) and high-performance data access and sharing (based on EOS, the CERN disk storage system for large-scale physics data analysis). Within CERNBOX we have developed a testing Python toolkit named Smashbox. Smashbox is maintained by CERN on GitHub with several external contributions. This will be the platform for investigating the behavior and the stability of CERNBOX and to evaluate the interoperability of similar cloud services. We need to extend the present system to cover the following use cases/scenarios: Replicate the Linux/Mac testing on Windows (running natively test clients on Windows boxes) Replicate the SAMBA mount testing on Windows (running natively test clients on Windows boxes) Fully integrate the Windows client distribution within the IT infrastructure (CMF)	This project requires proven experience in developing in Python on the Windows platform.
ST-FDO	LARGE-SCALE ERASURE-CODING DISK FARMS	EOS is the disk solution for LHC. It has been developed at CERN to handle the analysis of LHC data crossed the 100-PB mark of used disk space. This corresponds to about 800 M files being made available to CERN researchers and engineers. EOS software continues to be enhanced at CERN where it delivers a scalable storage system built on heterogeneous storage units. Data are served off large sets of disks (currently 50,000 disks in 1600 servers) and are efficiently accessed thanks to the cooperation of a fast name space (capable of sustaining requests well above 100 kHz) and a wide selection of transfer protocols. Data durability is achieved by maintaining multiple replicas or erasure-code fragments across the EOS disk farm. The project is to investigate new deployment scheme actively using the erasure-code capabilities to achieve: Reduced space overhead compared to equivalent multiple-replica schemes mprove single stream performance in reading from multiple disks Validate operational modes not requiring disk replacements by tuning the redundancy level against the disk lifetime.	This project requires proven experience in dealing with Linux system (Foundation of system administration and Python scripting)
DB-IMS	Enable Docker deployment for Weblogic applications for faster provisioning and easier operation	WebLogic Server is the Oracle application server for building and deploying critical enterprise Java applications. At CERN, it is mainly used as bases for administrative and engineering applications. The project will take advantage of latest Openstack and container technologies (Magnum, Kubernetes). It'll provide highly availably, scalable environment for Java applications developers. The aim is to look at the overall process of Weblogic deployment and adapt it for container like environment. It requires to investigate all aspects of development and deployment process: testing, application and infrastructure updates, single-sign on integration, security, access control, scaling monitoring as well as operations.	Programming languages: Java, shell scripting, python, Middleware: Java EE application servers, specially Oracle Weblogic. Apache HTTP Server and related backend servers communication protocols (HTTP(S), TCP, Web Sockets, proxy protocol), LDAP System administration: Linux, Docker Configuration Management: Puppet, RPM