Deploying Spectrum LSF on OCP: Tips and Tricks

F204240

Bill.McMillan@uk.ibm.com
Principal Product Manager

Michael Spriggs
STSM

Mark Black
Principal Architect

Tony Tang
Senior Developer

Peng Hu
Program Director

Peter Kang
Developer

George Gao
Senior Architect

IBM Spectrum Computing
IBM Cloud & Cognitive Software

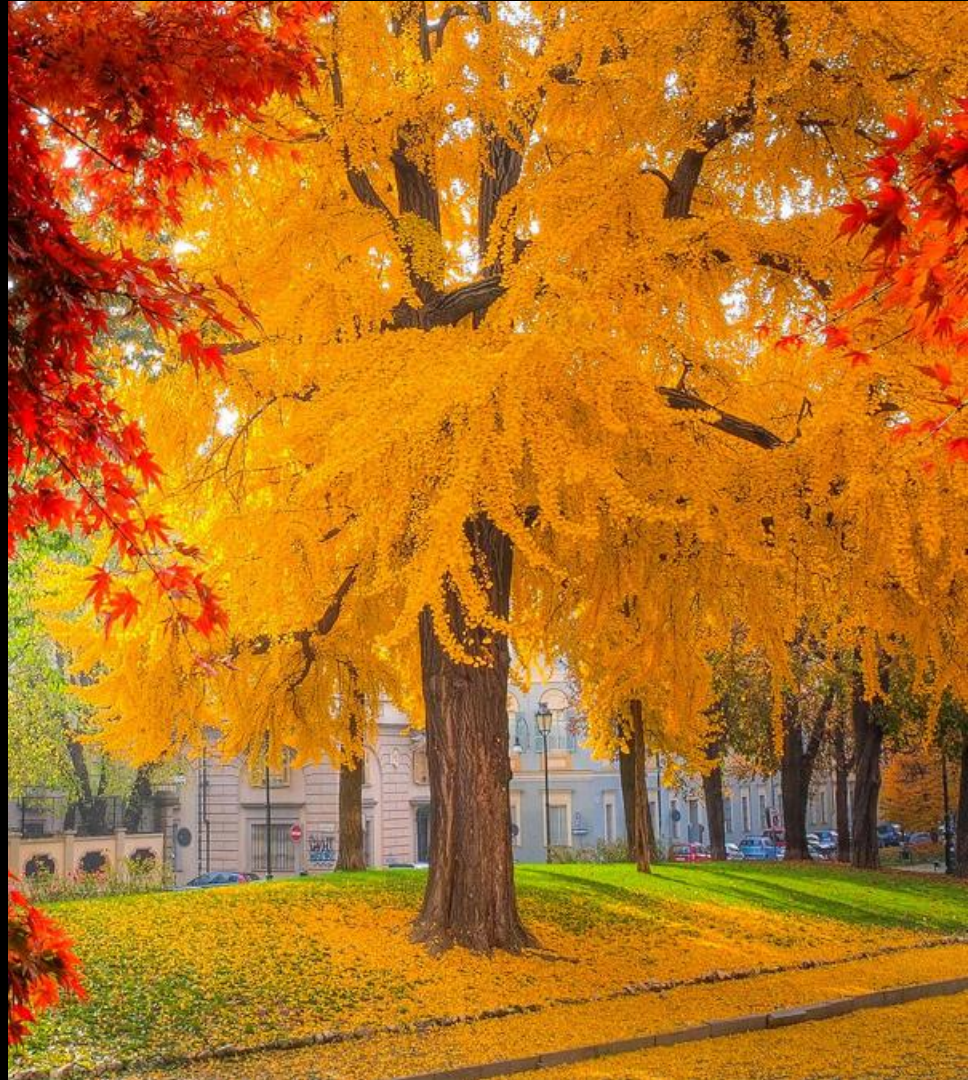
Agenda

Overview of IBM Spectrum LSF

Why use LSF with Kubernetes / OpenShift Container Platform?

Use Cases

Demo



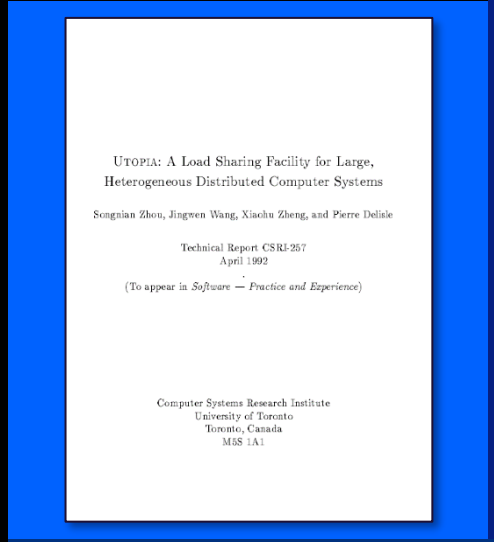
IBM Spectrum LSF

Comprehensive workload management for demanding and scalable HPC environments

Over the last 30 years LSF has evolved from a single scheduler to a full family of products.

Proven track record in delivering customer value

- 10 Major Releases
- 30+ Minor Releases
- 1000+ customer driven features
- On X86, Power, ARM, Windows (and lots of other platforms that have been consigned to history)



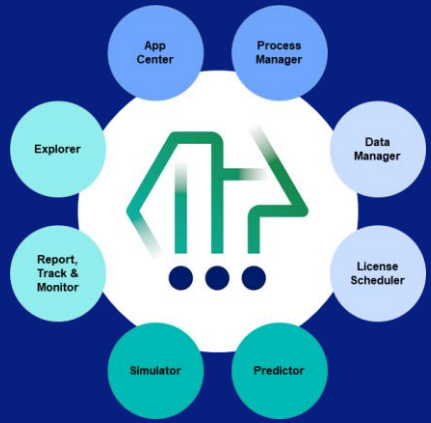
UTOPIA: A Load Sharing Facility for Large, Heterogeneous Distributed Computer Systems

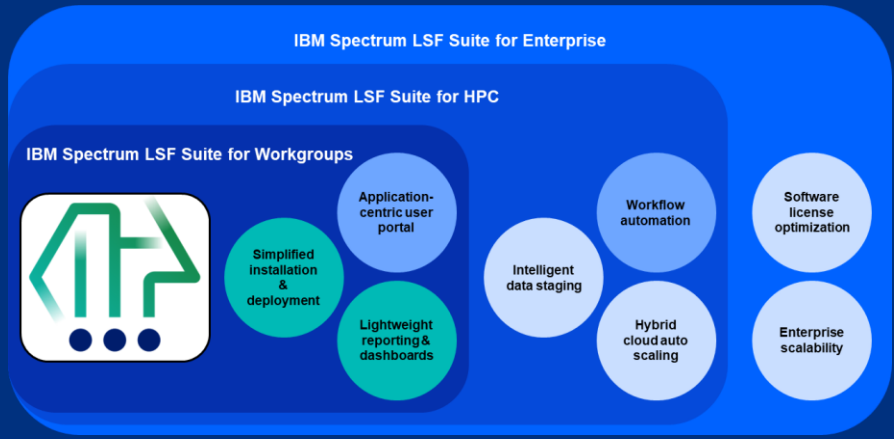
Songnian Zhou, Jingwen Wang, Xiaohu Zheng, and Pierre Delisle

Technical Report CSRI-257
April 1992

(To appear in *Software — Practice and Experience*)

Computer Systems Research Institute
University of Toronto
Toronto, Canada
M5S 1A1





LSF is primarily used where Time to Results matters

Electronics Design Automation (EDA)

Chip design & optimization, Circuit simulation & verification, Manufacturing optimization

- Some of the largest clusters with over 8K servers, 250K cores, millions of jobs per day
- Bursting to the cloud at scale: 50K cores

Automotive & Aerospace

CFD-aerodynamic modeling, FEA-impact/structural strength analysis, CAD/CAM

- Large scale: 5K+ servers, 50-150K cores, millions of jobs per week
- Bursting to the cloud at large scale: 5K cores

Life Sciences

Genome processing & sequencing, Drug design, Molecular modeling & biology simulation, Protein docking.

Assemble & map large genomes in hours instead of days & weeks

- Largest sites are ~50K cores
- Extensive use of cloud

Oil & Gas

Seismic Data Processing, Reservoir Simulation & Modeling, Geospatial analytics, Terrain/topology mapping, CFD-aerodynamic modeling, Wind simulation

- Large scale, job volume varies
- Relatively limited use of cloud

Banking, Financial Markets & Insurance

Monte Carlo simulations, Risk analysis, Fraud detection.

Calculate this **now!** 100,000s scenarios, 100s instruments & time steps

Retail

Inventory analysis, logistics & supply chain optimization, sentiment analysis, marketing offers

- Some use of LSF along side Spectrum Symphony at investment banks
- Large number of LSF and LSF Process Manager as part of the SAS Grid OEM (across FSS, Retail/Distribution, Healthcare)

Government & Defense

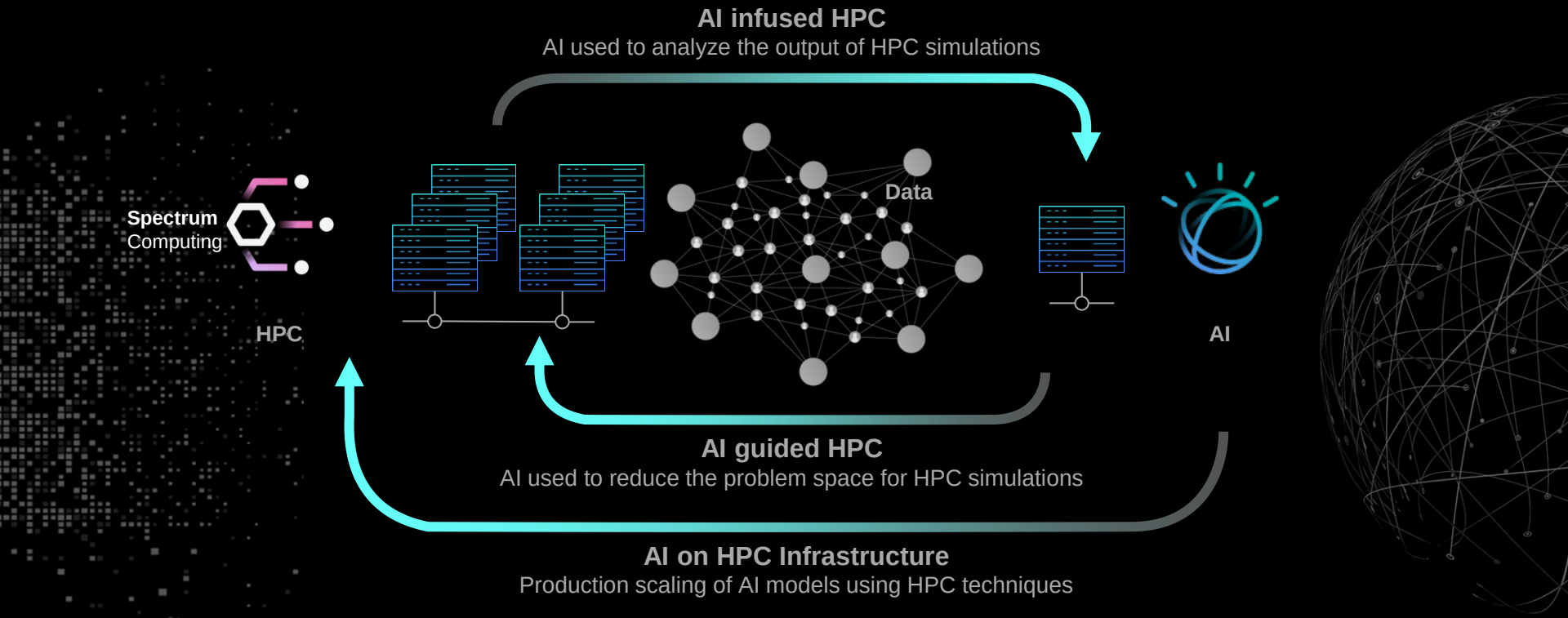
Intelligence agency, fraud analysis, climate modeling, weather forecasting, energy, nuclear stewardship, exploration

- Weather
- National Labs

Research & Academia

Why Use LSF with Kubernetes / OpenShift Container Platform?

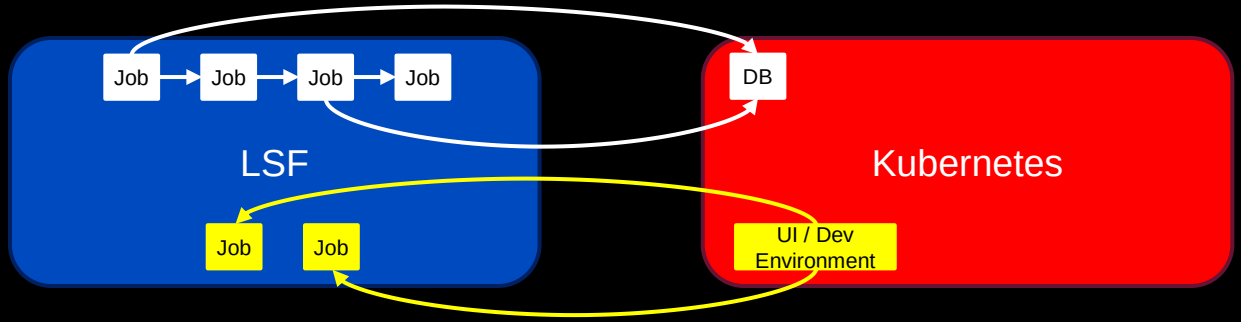
Convergence of AI and HPC



1. Side by Side

Independent environments where the applications running within them can launch work, or uses services in the other as required.

e.g. a computational workflow running in LSF could access a database hosted in K8S, or even create an instance on demand if required.

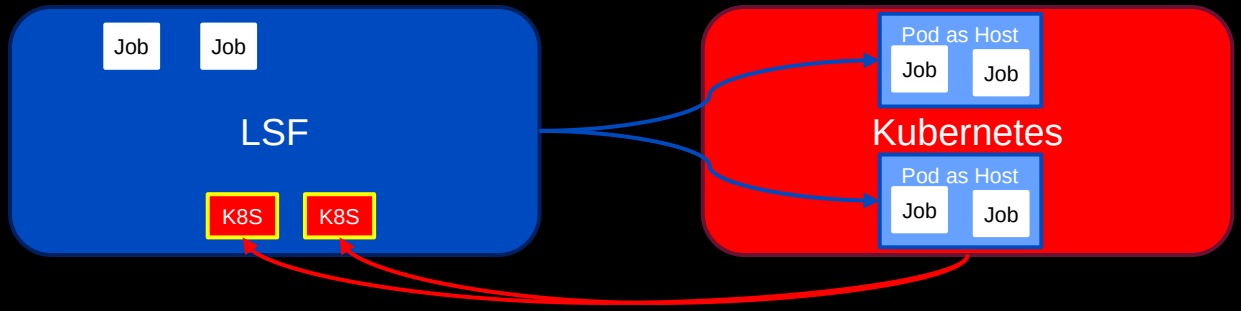


e.g. an application front end or development environment could be running on K8S and it launches work into the existing LSF environment

2. Burst Capacity “Pod as a host”

With multiple pools of resources, there will be times when there is demand in one, and idle capacity in the other.

LSF’s “Resource Connector” supports Kubernetes/OCP as an end point, and can create “pods as hosts” to extend the LSF cluster into the K8S environment, leveraging spare capacity.



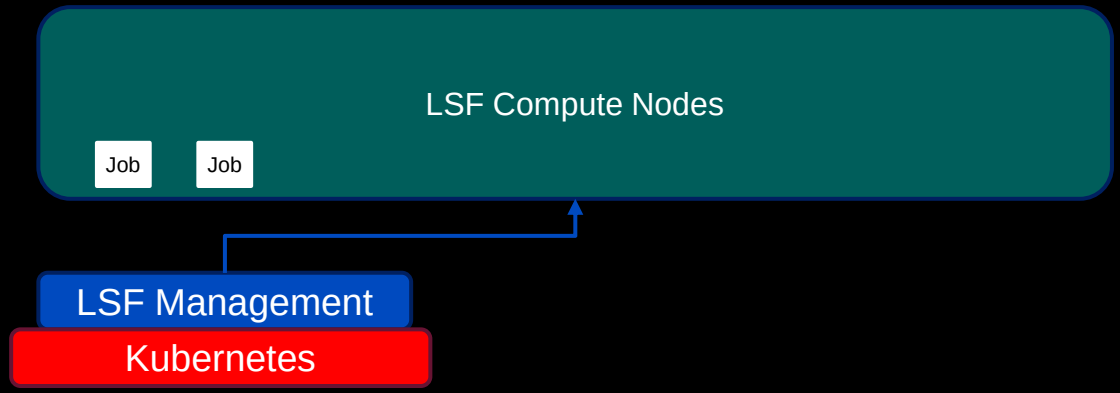
e.g. The opposite is also possible. Kubernetes services could be started on demand as LSF jobs, allowing the Kubernetes cluster to stretch into the LSF environment.

3. LSF on Kubernetes/OpenShift Container Platform

LSF itself can be run as a service/virtualized cluster on top of K8S. It can still leverage the resource connector in use case #2, to shrink and grow the LSF environment. Each LSF environment is deployed using an Operator.

4. LSF Management on Kubernetes

To address this, we can just run the LSF management services (Scheduler, UI, DB's etc) on Kubernetes and have all the compute nodes on bare metal (or VM's) allowing the site to use whatever container technology they wish.

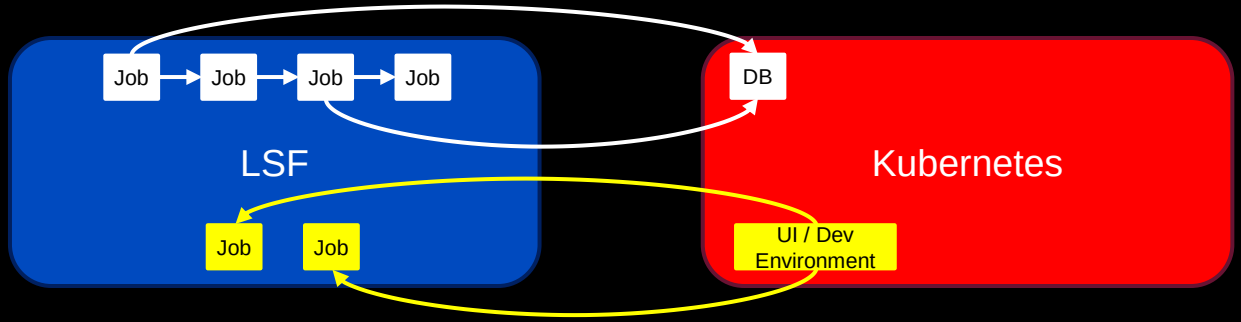


Detailed Use Cases

1. Side by Side

Independent environments where the applications running within them can launch work, or uses services in the other as required.

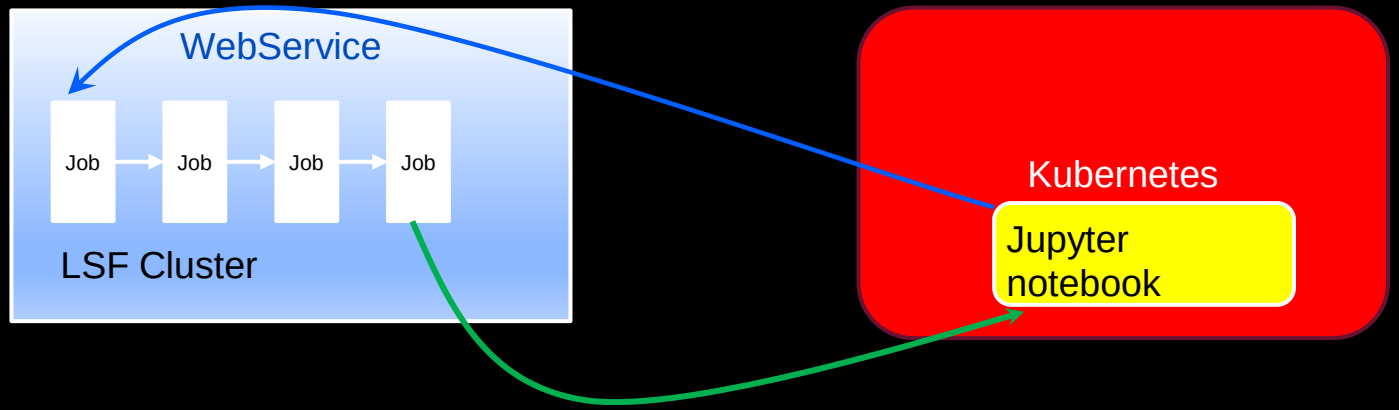
e.g. a computational workflow running in LSF could access a database hosted in K8S, or even create an instance on demand if required.



e.g. an application front end or development environment could be running on K8S and it launches work into the existing LSF environment

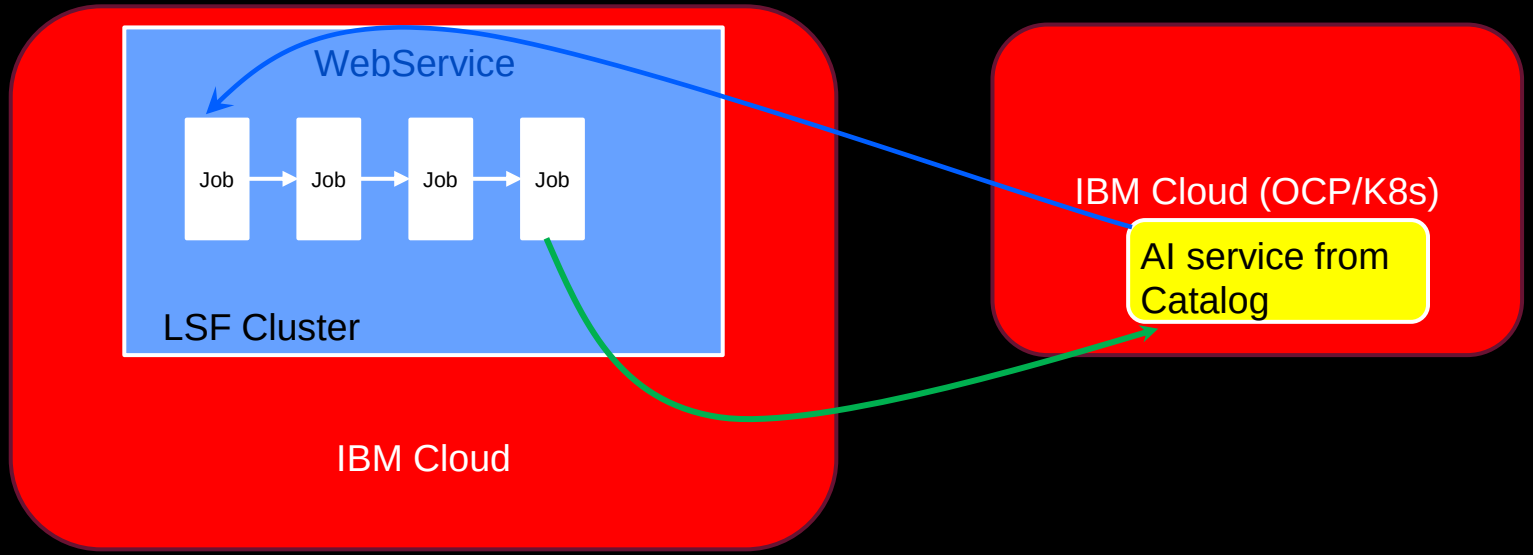
1. Side by Side

A Jupyter notebook needing to perform some machine learning calculation submits those calculations into the backend local LSF cluster.



1. Side by Side

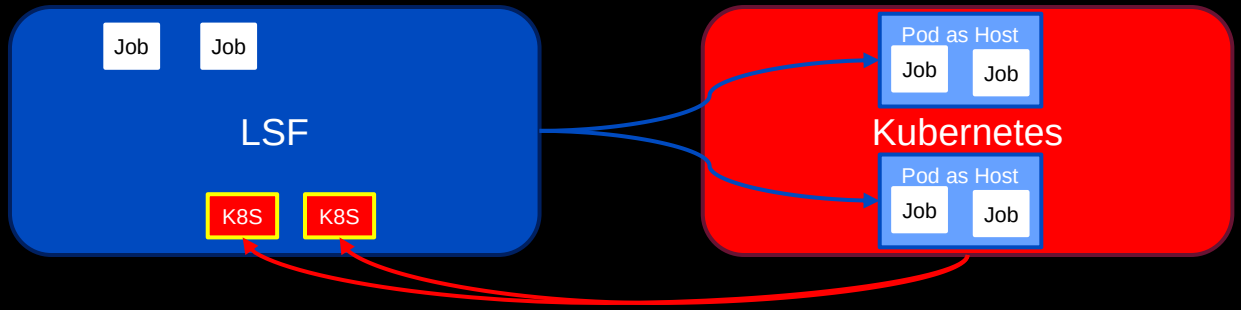
A front end GUI application/service running on Cloud submits a job to a backend LSF cluster which is part of the service.



2. Burst Capacity “Pod as a host”

With multiple pools of resources, there will be times when there is demand in one, and idle capacity in the other.

LSF’s “Resource Connector” feature supports Kubernetes/OCP as an end point, and can create “pods as hosts” to extend the LSF cluster into the K8S environment, leveraging spare capacity.



Hybrid Cloud Architecture

Workload is forwarded to the appropriate cloud based on site defined policies.

Portal, command line and restful API for submission and monitoring

IBM Spectrum LSF

IBM Spectrum LSF

Cloud resources can be autoscaled based upon workload demands and policies.

IBM Cloud

aws

openstack.

Microsoft Azure

Google Compute Engine

kubernetes OPENSHIFT

NVIDIA

AMD RADEON GRAPHICS

aspera an IBM company

IBM Spectrum Scale

Data is staged from source to the cloud before hosts are provisioned. It also caches files to avoid repeatedly moving the same files, and it returns results to on prem out of band.

Applications can be pre-installed in the cloud images, or installed on-demand by container

docker

kubernetes OPENSHIFT

podman

x86

Power

arm

On Premise

Off Premise/Cloud

LSF Resource Connector for OpenShift

https://www.ibm.com/docs/en/spectrum-lsf/10.1.0?topic=providers-configuring-openshift-lsf-resource-connector

Enabling the RC for OpenShift

1. Run the helper script
   # ./enable_lsf_rc.sh -n <LSF Operator's namespace> -s <A new service account for RC>

2. As a result, a new service account is created with a sufficient RBAC (role-based access control) permissions. openshiftprov_config.json, openshiftprov_templates.json are modified accordingly with the namespace and the service account, the service account token, and the container image name for an RC example template.

3. Configure RC for OpenShift templates in openshiftprov_templates.json
   - The openshiftprov_templates.json file defines the mapping between LSF resource demand requests and OpenShift instances.
   - The template represents a set of hosts that share some attributes such as the number of CPUs, the amount of available memory, the installed software stack, operating system, and other attributes.
   - The resource connector uses the definitions in this file to map this demand into a set of allocation requests in OpenShift.

LSF Resource Connector for OpenShift (an example)

1. Enable LSF RC for OCP

```
# ./enable_lsf_rc.sh -n lsf2 -s account1 -d /opt/ibm/lsfsuite/lsf
Configuring the LSF management pod <mylsf2-master-596fbbdcbd-867sl>...
IBM Spectrum LSF resource connector for OpenShift enabled
```

2. Submit jobs to rc_example queue, LSF RC will evaluate the number of Pending jobs in “rc_example” queue, if they exceed the threshold defined in RC_DEMAND_POLICY, LSF rc pods will be provisioned from OpenShift and added to the LSF cluster to run these jobs.

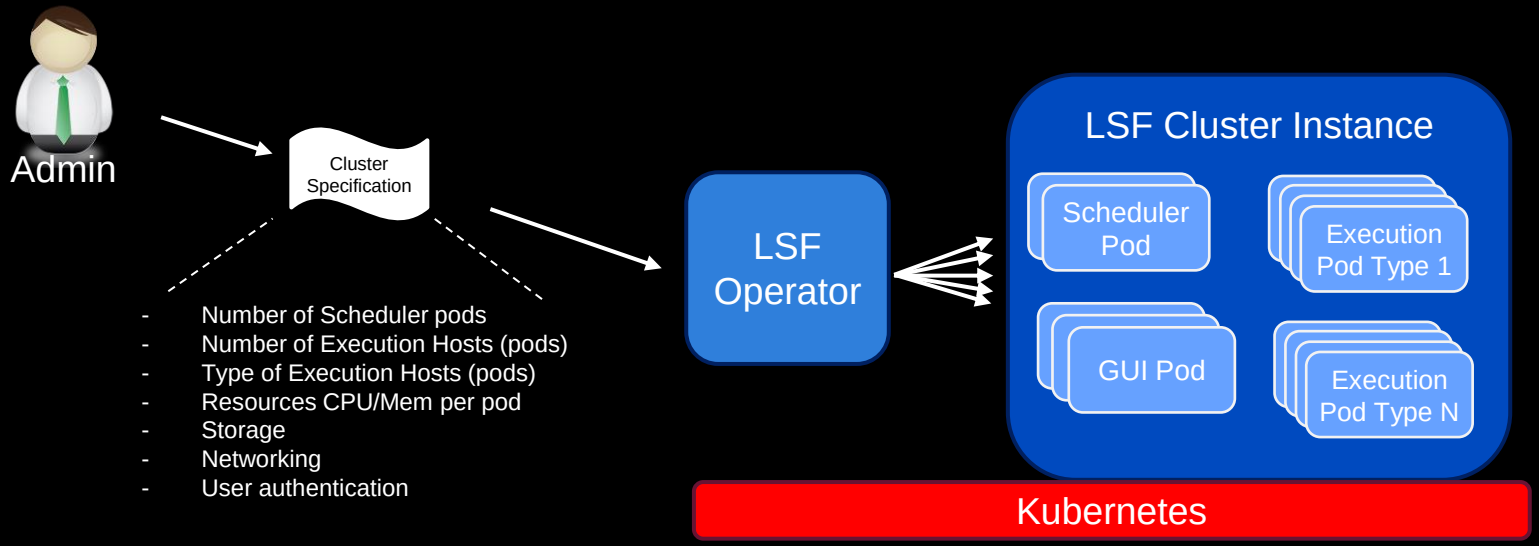
```
root@lsf1x113:/opt/LSF_Desktop_Client# bsub -R "select[openshift]" -q rc_example -J rc_job sleep 600
Job <234> is submitted to queue <rc_example>.
root@lsf1x113:/opt/LSF_Desktop_Client# bqueues rc_example
QUEUE_NAME      PRIO STATUS          MAX JL/U JL/P JL/H NJOBS  PEND   RUN  SUSP
rc_example       50  Open:Active       -    -    -    -    10     5     5     0
root@lsf1x113:/opt/LSF_Desktop_Client# bhosts
HOST_NAME          STATUS       JL/U    MAX  NJOBS    RUN  SSUSP  USUSP    RSV
gui                closed          -      0      0      0      0      0      0
lsf-rc-2f751ce0-75 closed          -      1      1      1      0      0      0
lsf-rc-6de55799-29 closed          -      1      1      1      0      0      0
lsf-rc-6e0b749e-9f closed          -      1      1      1      0      0      0
lsf-rc-a150601e-e6 closed          -      1      1      1      0      0      0
lsf-rc-d4970afb-6c closed          -      1      1      1      0      0      0
lsfmaster          closed          -      0      0      0      0      0      0
mylsf2-rhel7-7b4cf ok              -      2      0      0      0      0      0
```

3. LSF on Kubernetes/OpenShift Container Platform

LSF itself can be run as a service/virtualized cluster on top of K8S. It can still leverage the resource connector in use case #2, to shrink and grow the LSF environment. Each LSF environment is deployed using an Operator.

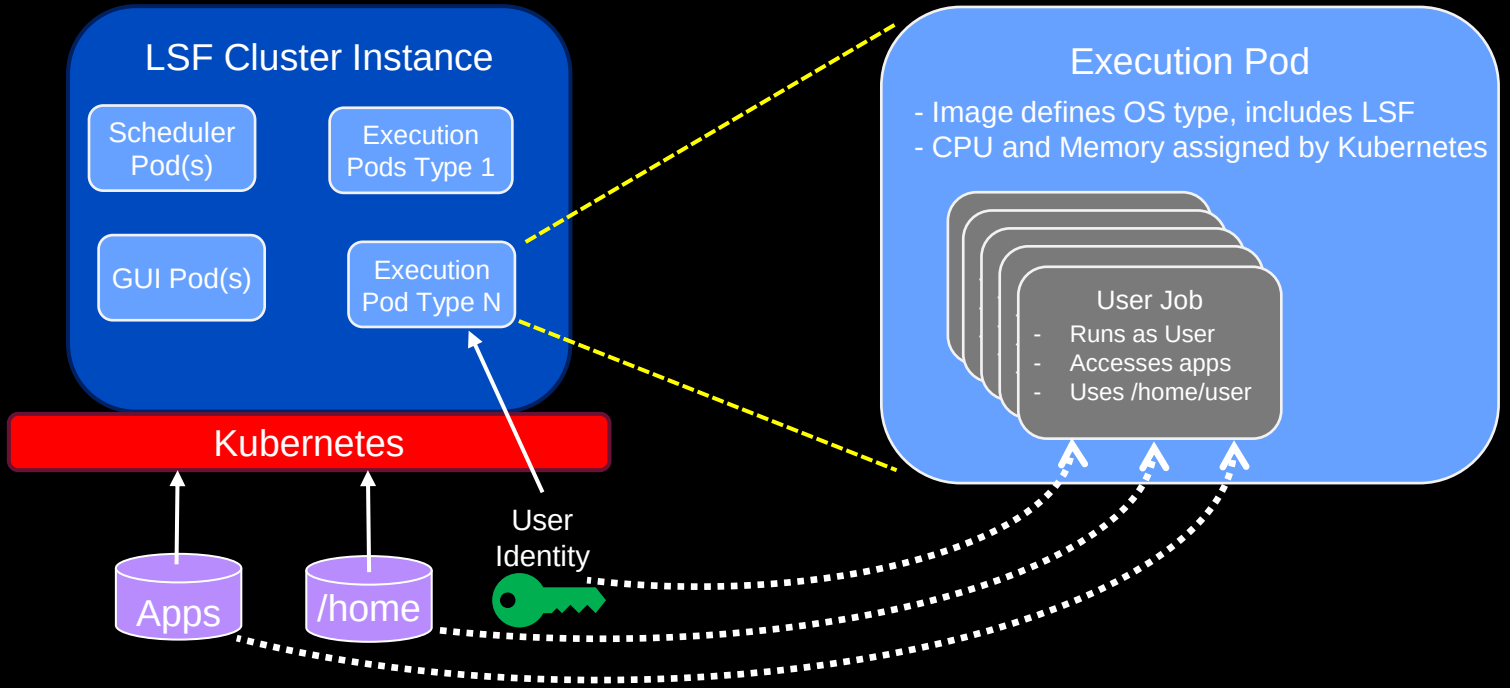
3. LSF on Kubernetes/OpenShift Container Platform

The LSF Operator deploys an LSF cluster based on the cluster specification provided by the administrator. This sets the size of the LSF Cluster, and the amount of resources to get from Kubernetes. LSF Hosts are created as Pods in Kubernetes. Jobs run **inside** pods!



3. LSF on Kubernetes/OpenShift Container Platform

Jobs run **inside** the LSF pods. LSF Operator maps Kubernetes Volumes into the pods for applications and user data. Datacentre user identity services are used inside the LSF pods, so users can access there data and applications.

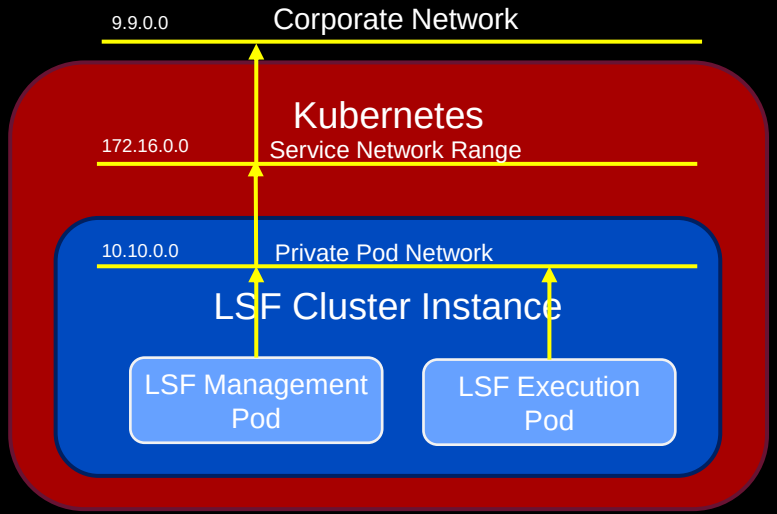


3. LSF on Kubernetes/OpenShift Container Platform

Two networking options when running in Kubernetes:

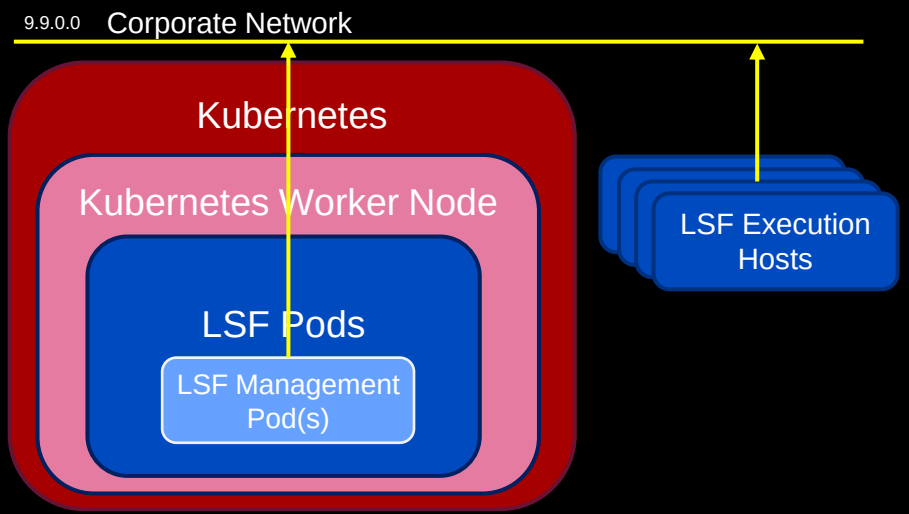
1. Kubernetes software defined networking

- Maximum Security
- LSF pods have dynamic IPs
- No access to execution hosts (pods)
- Users use PAC, Desktop client, or Login pods
- Access to management pods via “routes”



2. Physical Host Networking

- Maximum performance
- Direct access to Worker Node mounts (NFS mounts)
- LSF pods use IP stack of the Kubernetes worker nodes
- LSF execution hosts are physical machines
- 1:1 relationship between LSF management pod and physical node



Configuring LSF pods to access shared data

You can create OCP PersistentVolume (PV) to host user home directories, application executables and data, and shared with each LSF pod in the LSF on OCP.

1. Create PersistentVolume (PV) to connect to file system with shared data

```
# oc get pv
NAME          CAPACITY   ACCESS MODES   RECLAIM POLICY   STATUS   CLAIM                       AGE
myhomevol2    10Gi       RWX            Retain           Bound    lsf2/mylsf2-home            20h
myappsvol2    10Gi       ROX            Retain           Bound    lsf2/mylsf2-applications    20h
mydatavol2    10Gi       ROX            Retain           Bound    lsf2/mylsf2-data            20h
```

2. Configuring mounting point in lsf pod

```
# vi example-lsf.yaml

 volumes:
 - name: "Applications”
        mount: "/apps”
        selectorLabel: "apps”
        selectorValue: "apps”
        accessModes: "ReadOnlyMany”
        size: "10G”
```

```
# vi openshiftprov_templates.json

 ”templates”: [
     "mountPaths": {
         "mylsf2-pvc": "/opt/ibm/lsfsuite/lsfadmin",
         "mylsf2-home": "/home",
         "mylsf2-applications": "/apps",
         "mylsf2-data": "/data"
      },
```

Running Singularity jobs on LSF on OCP

In order to run singularity workload inside a pod, the pod needs to be in privileged mode.

All pods created by LSF Operator and LSF resource connector are non-privileged. If you would like to run singularity job in LSF cluster on OCP, you need to specify "privileged": true in LSF resource connector provisioning template.

Since privileged pods expose security risks for container user to access host properties, it is strongly suggested that you build a dedicated LSF cluster for singularity jobs on OCP, specify CONTAINER with admin-provided singularity images in rc_example queue, and only allow end-users to use this queue to submit LSF jobs.

Steps:

1. Create custom lsf-comp image with singularity installed
2. Configure provisioning template so lsf-rc pods are created in privileged mode
3. Configure CONTAINER with admin-provided singularity images in rc_example queue
4. Submit LSF jobs to execute singularity images specified in rc_example queue

Deploying LSF on OCP (an example)

1. Create a Namespace
2. Set up PV or Dynamic Storage Class to save LSF conf/work files
3. Load LSF Operator and LSF images to OCP registry
4. Prepare deployment env (lsfcluster crd, scc, service account, roles etc)
5. Create LSF operator
6. Deploy LSF cluster on OCP via LSF operator

```
# oc project
Using project "lsf2" on server "https://api.glueing.cp.fyre.ibm.com:6443".
# oc create -f operator.yaml
# oc create -f example-lsf.yaml
# oc get lsfcluster
NAME                  AGE
example-lsfcluster2   17h

# oc get pod
NAME                                  READY   STATUS    RESTARTS   AGE
ibm-lsf2-operator-7fffbb84cf-rtn7w    2/2     Running   0          18h
mylsf2-gui-95b759f59-tcvsh            2/2     Running   0          17h
mylsf2-master-596fbbdcbd-867sl        1/1     Running   0          17h
mylsf2-rhel7-7b4cfdbd98-9pgtv         1/1     Running   0          17h
```

Using LSF on OCP (via web console)

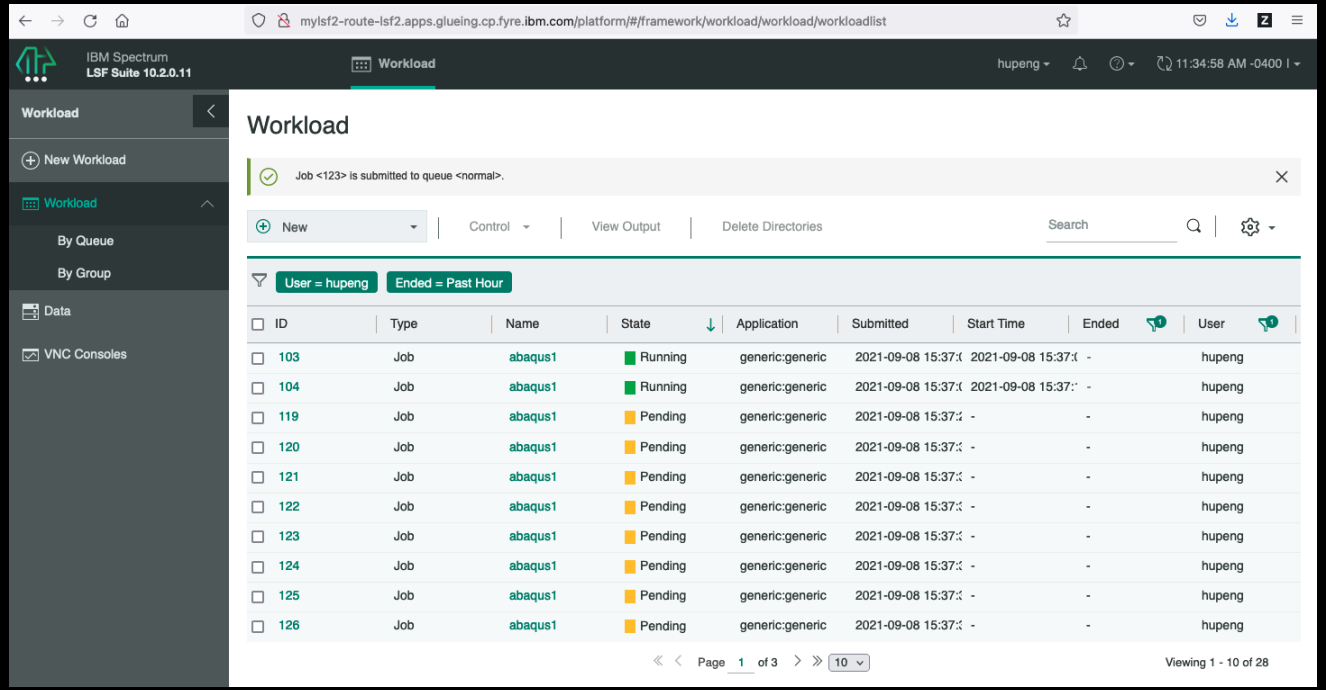
1. Obtain OCP route for LSF on OCP

oc get route

```
NAME            HOST/PORT
SERVICES
mylsf2-route    mylsf2-route-
lsf2.apps.glueing.cp.fyre.ibm.
com    mylsf2-service
```

2. Open web browser, type in mylsf2-route-lsf2.apps.glueing.cp.fyre.ibm.com, log in with your LDAP credentials

3. You can submit and manage your LSF jobs in the web console now



Using LSF on OCP (via CLI)

1. Obtain OCP route for LSF on OCP

```
# oc get route
NAME           HOST/PORT                                        SERVICES
mylsf2-route   mylsf2-route-lsf2.apps.glueing.cp.fyre.ibm.com   mylsf2-service
```

2. On your desktop, install LSF Application Center Desktop Client package, setenv LSF_DESKTOP_CLIENT=yes
3. Run paclogon with your LDAP credentials and the route

```
root@lsf1x113:/opt/LSF_Desktop_Client# paclogon
Log on to IBM Spectrum LSF Application Center
User account: hupeng
Enter password:
Specify the URL to connect to IBM Spectrum LSF Application Center. Format:
http://host_name:port_number/platform or https://host_name:port_number/platform
URL: http://mylsf2-route-lsf2.apps.glueing.cp.fyre.ibm.com:80/platform
You have successfully logged on to IBM Spectrum LSF Application Center.
```

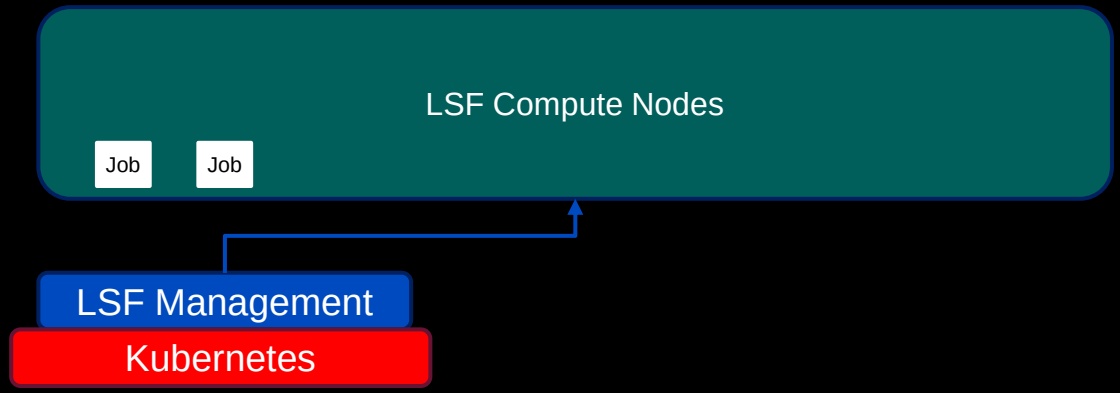
4. You can use LSF CLI to submit and manage your jobs now.

```
root@lsf1x113:/opt/LSF_Desktop_Client# bsub -J testJob -q priority sleep 60
Job <133> is submitted to queue <priority>.
root@lsf1x113:/opt/LSF_Desktop_Client# bjobs 133
JOBID   USER    STAT  QUEUE      FROM_HOST   EXEC_HOST   JOB_NAME   SUBMIT_TIME
133     hupeng  RUN   priority   gui         mylsf2-rhel testJob    Sep  8 15:49
```

Demo

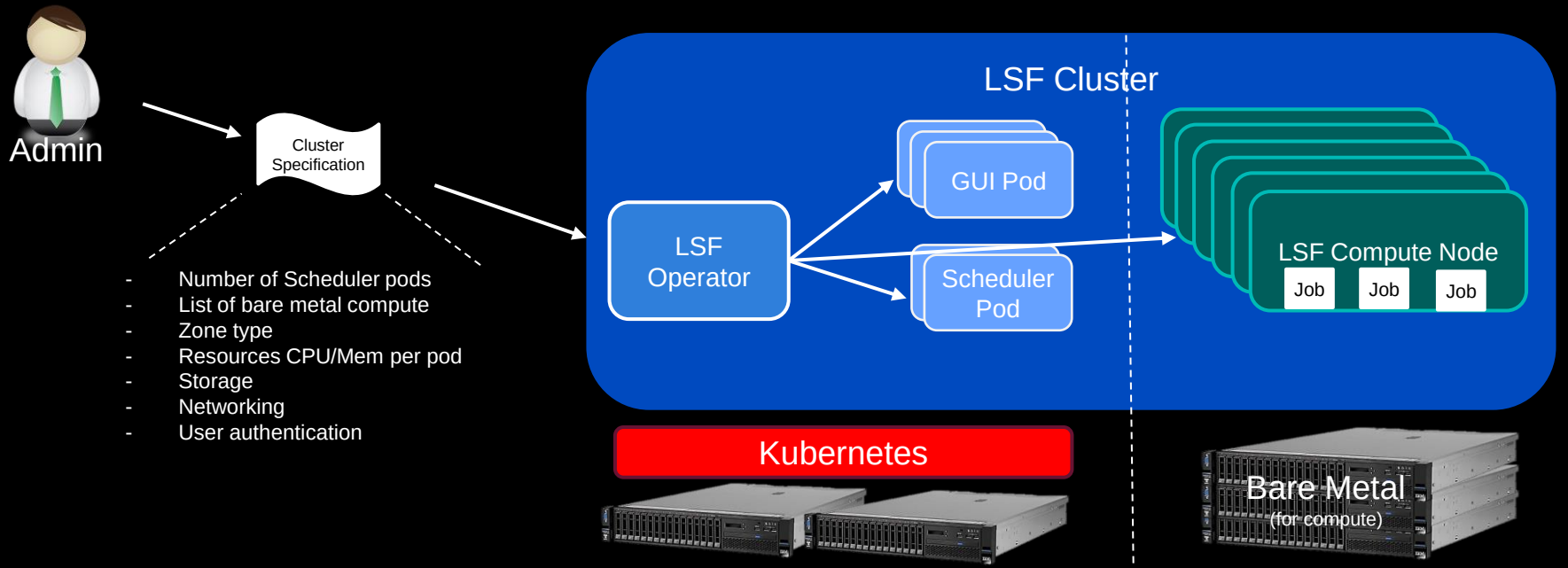
4. LSF Management on Kubernetes

To address this, we can just run the LSF management services (Scheduler, UI, DB's etc) on Kubernetes and have all the compute nodes on bare metal (or VM's) allowing the site to use whatever container technology they wish.



4. LSF Management on Kubernetes

The LSF Operator can deploy LSF management functions on Kubernetes, and Compute on bare metal machines. Pods use HostNetwork, and HostPath on Kubernetes gives pods same access as LSF Compute Nodes



In Conclusion

How you introduce K8S/OCP into an existing HPC environment will depend on the long term goals.

Compared to a traditional bare metal install, establishing an HPC environment on OCP is more involved.

However, once the Operator is configured, creating new clusters, or extending existing clusters, is straightforward.

Thank you

Bill.McMillan@uk.ibm.com
Principal Product Manager, IBM Spectrum Computing

Please complete the session evaluation!

Notices and disclaimers

U.S. Government Users Restricted Rights — use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. **This document is distributed "as is" without any warranty, either express or implied. In no event, shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity.** IBM products and services are warranted per the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts. In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply."

Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.

Performance data contained herein was generally obtained in a controlled, isolated environments. Customer examples are presented as illustrations of how those

customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer's responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer follows any law.

Notices and disclaimers

- Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products about this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. **IBM expressly disclaims all warranties, expressed or implied, including but not limited to, the implied warranties of merchantability and fitness for a purpose.**
- The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.
- IBM, the IBM logo, ibm.com and [names of other referenced IBM products and services used in the presentation] are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at: www.ibm.com/legal/copytrade.shtml


IBM