# Red Hat OpenShift on IBM zSystems: Demystifying the steady state load

June 2023

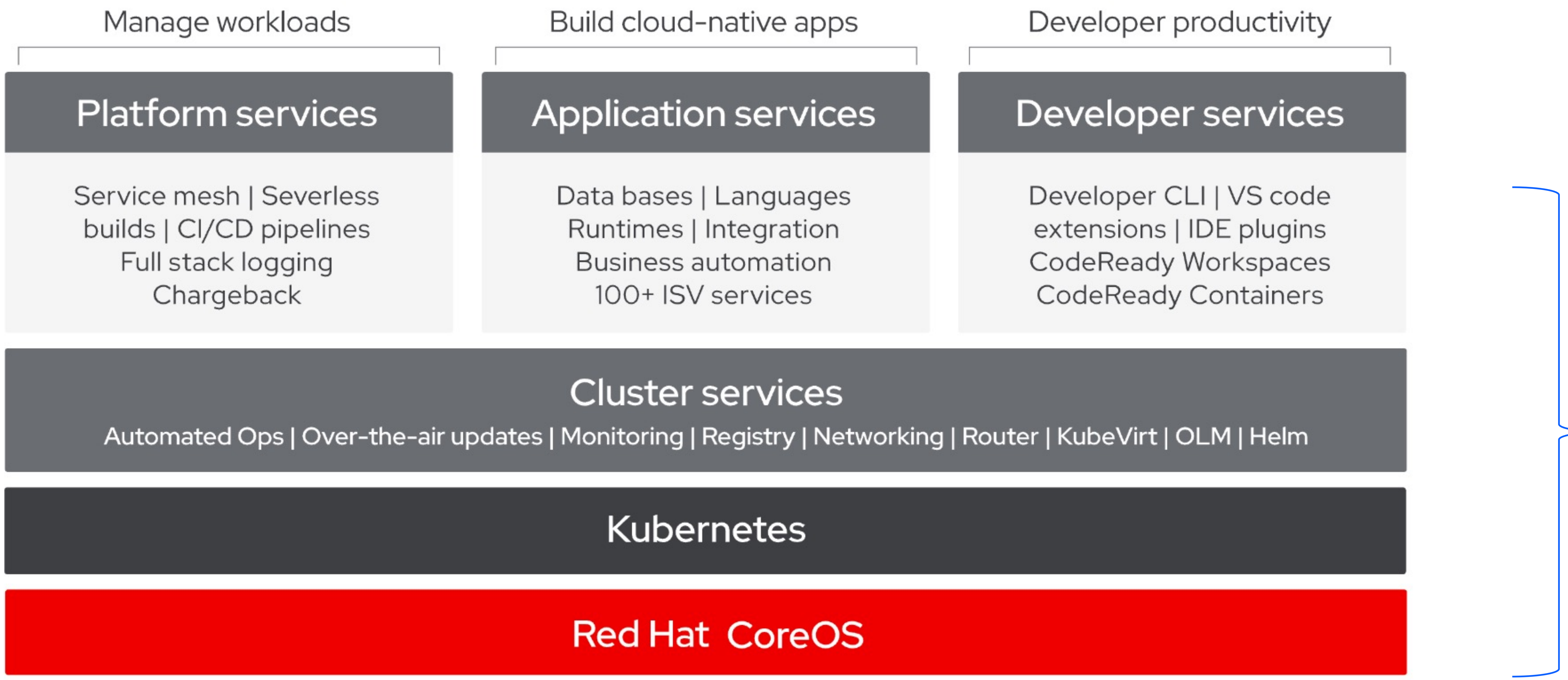
Dr.-Ing Axel Busch – Technical Lead Performance OCP on IBM zSystems and LinuxONE

# Definition of the Red Hat OpenShift steady state load

OpenShift **steady state load** is the **physical core capacity** OCP is using for **keeping the platform and the workloads up and running**.

- Supervise availability and correct function of nodes, services/operators and workloads
- Continuous collection of monitoring data
- Maintaining HA synchronization
- All other **platform services**

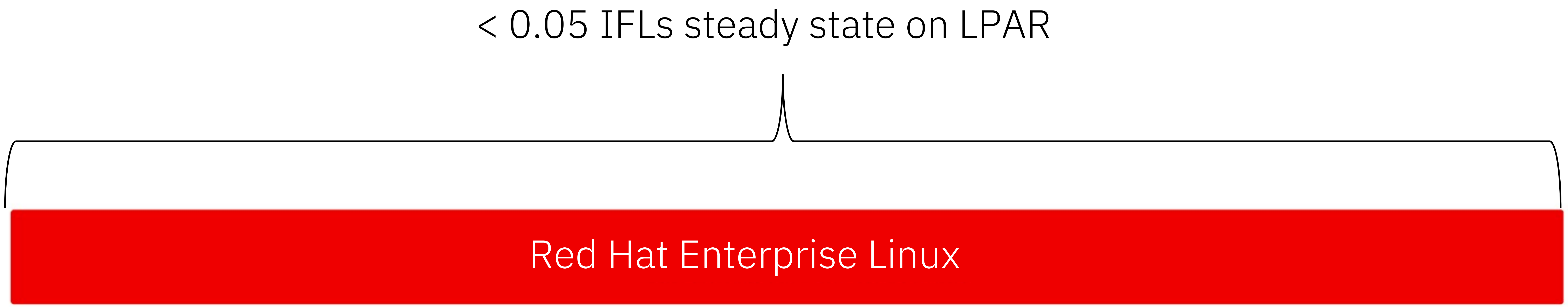
Steady state load **does not include** consumption of client workload and additionally installed operators




Approx. 1.6 IFLs (OCP 4.12, z16)

# Demystifying Red Hat OpenShift steady state load

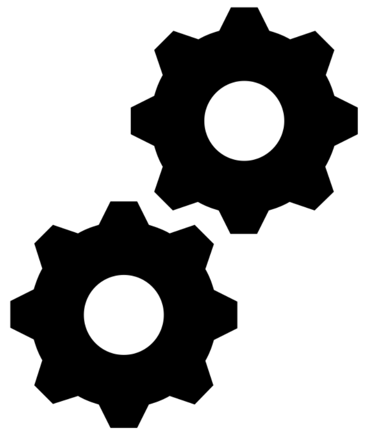
- Red Hat OpenShift steady state load is the amount of CPU utilization of LPAR when Red Hat OpenShift is up and running - without any deployment
- **But**: Red Hat OpenShift is a **platform** not an operating system!
- Nodes of Red Hat OpenShift each **have their own operating systems...**
- ... and **hundreds of processes** that implement the **Red Hat OpenShift operators**



# Business value of Red Hat OpenShift

**Cluster services**

Automated Ops | Over-the-air updates | Monitoring | Registry | Networking | Router | KubeVirt | OLM | Helm

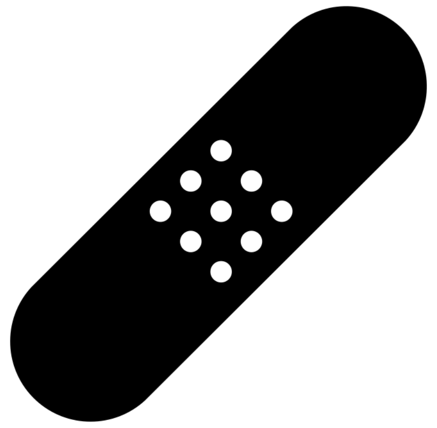


- Replaces manual recurring tasks
- No admin must setup/maintain/configure/patch the framework
- Better efficiency

# Business value of Red Hat OpenShift

**Cluster services**

Automated Ops | Over-the-air updates | Monitoring | Registry | Networking | Router | KubeVirt | OLM | Helm



- Replaces manual updating and maintenance of a complex cloud stack (comprised of lots of components)
- No review of release notes, integration tests after update,...
- Just click on update and Red Hat OpenShift updates the overall cloud stack

# Business value of Red Hat OpenShift

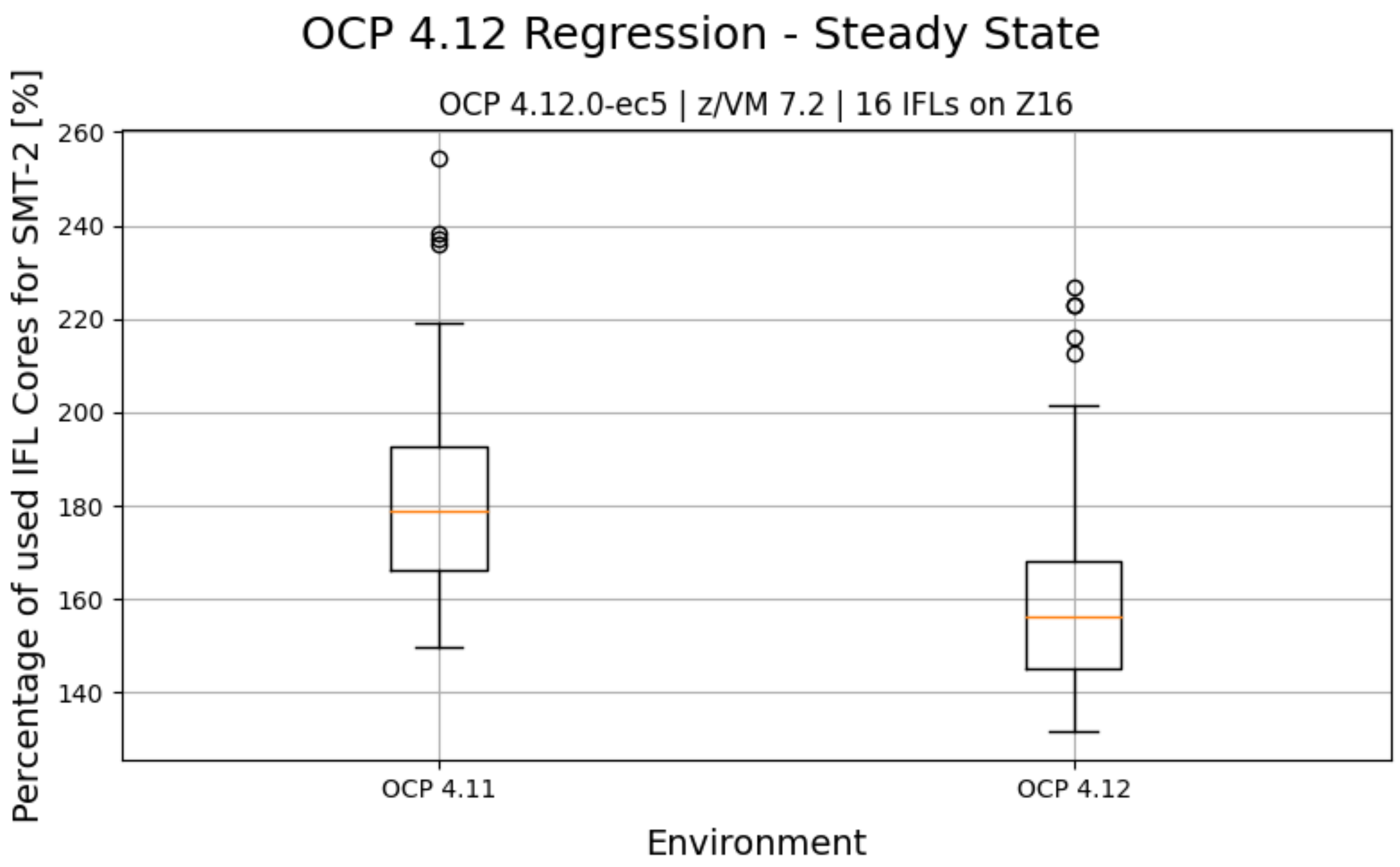
**Cluster services**

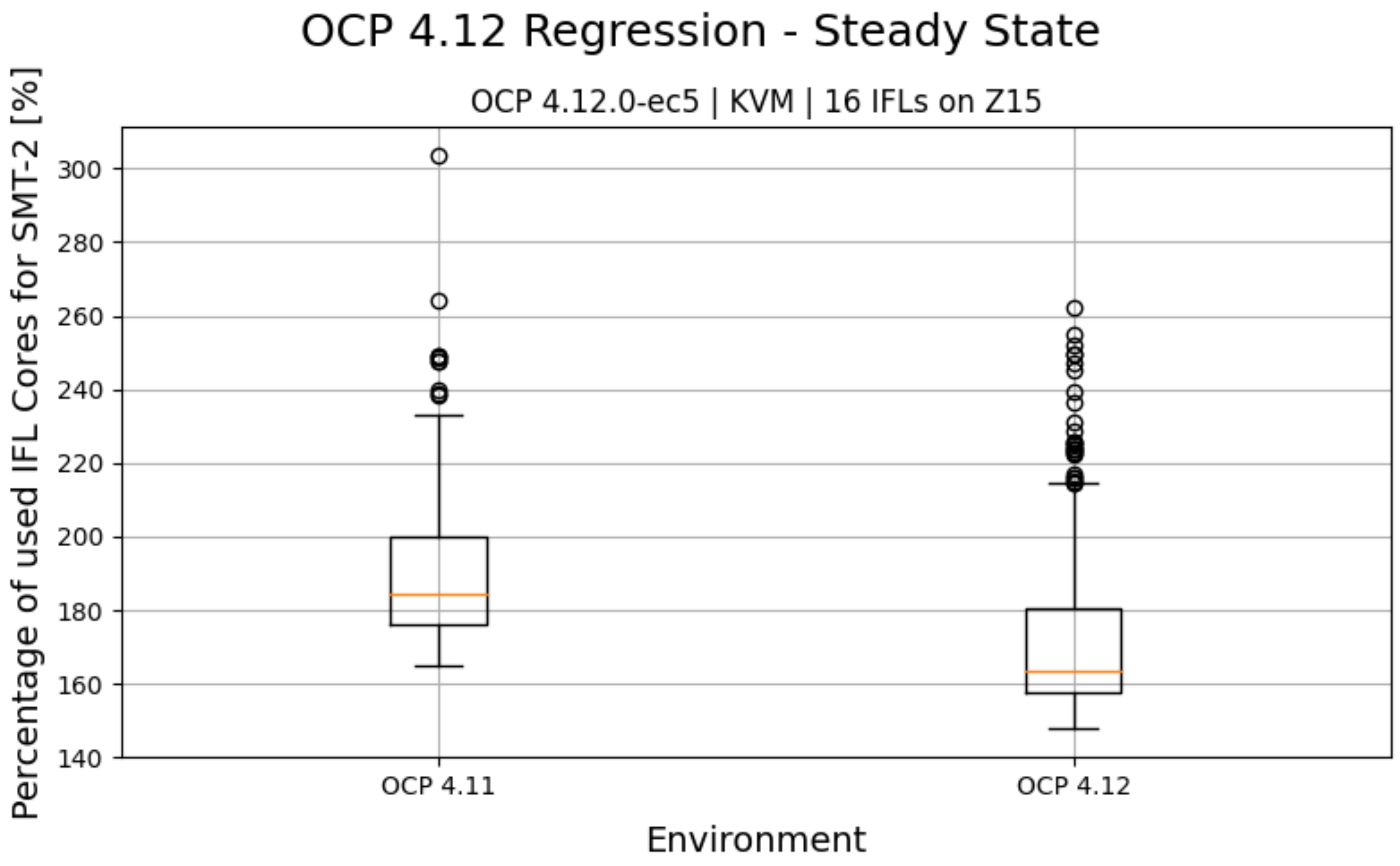
Automated Ops | Over-the-air updates | Monitoring | Registry | Networking | Router | KubeVirt | OLM | Helm

- Red Hat OpenShift ships a full functional monitoring stack on default setup
- No setup/maintenance/configuration needed for own monitoring solution to observe cluster core functions (i.e., kubernetes, networking,...) and own deployments
- No in-depth knowledge required about configuring monitoring correctly and efficiently

  - Allows you to do more with less people
  - Runs your cloud more efficient
  - Comes with the cost of some steady state load

# Steady state load z/VM & KVM on zSystems/IBM® LinuxONE



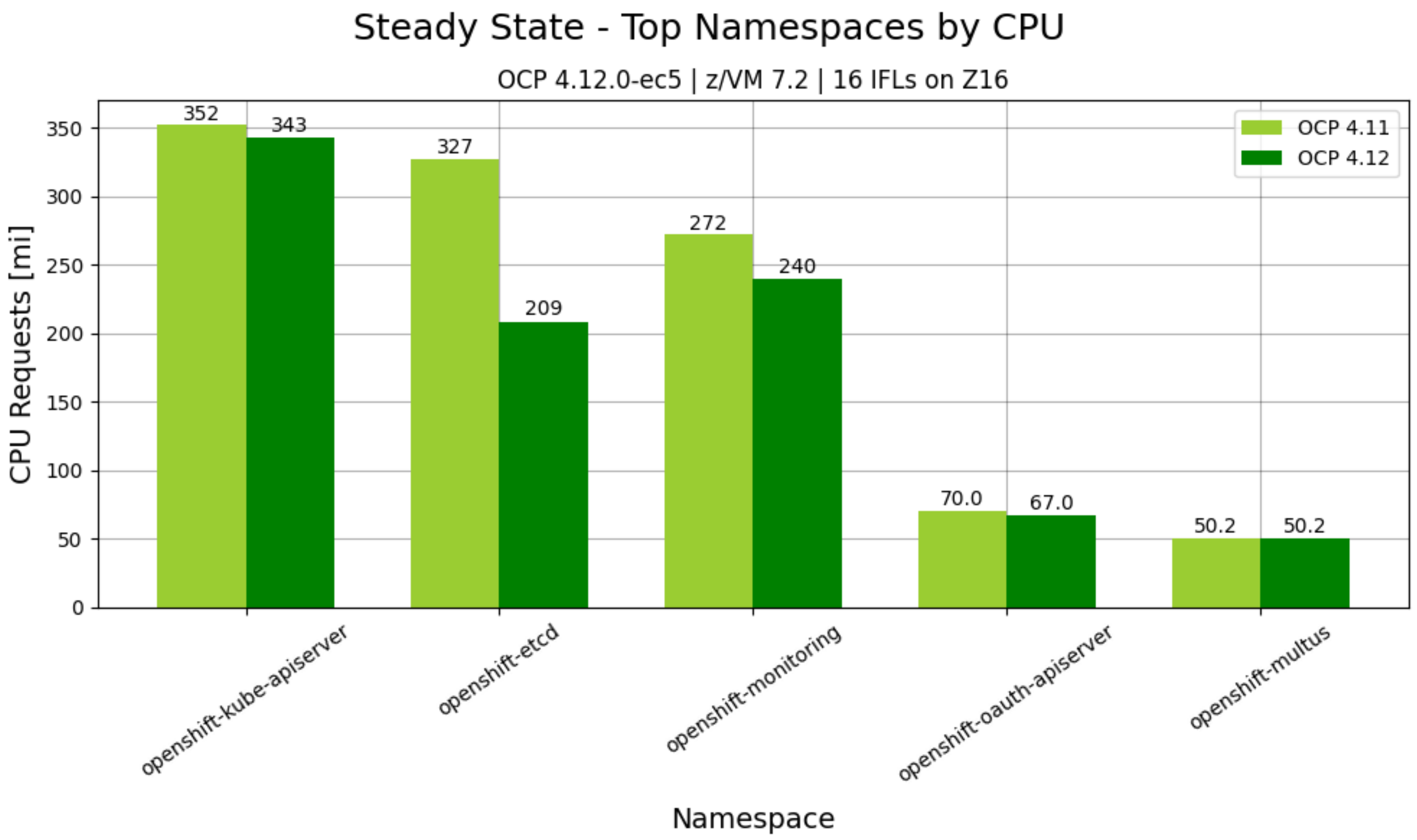




– 9% better steady state load on z/VM
– Less load peaks and lower load peaks

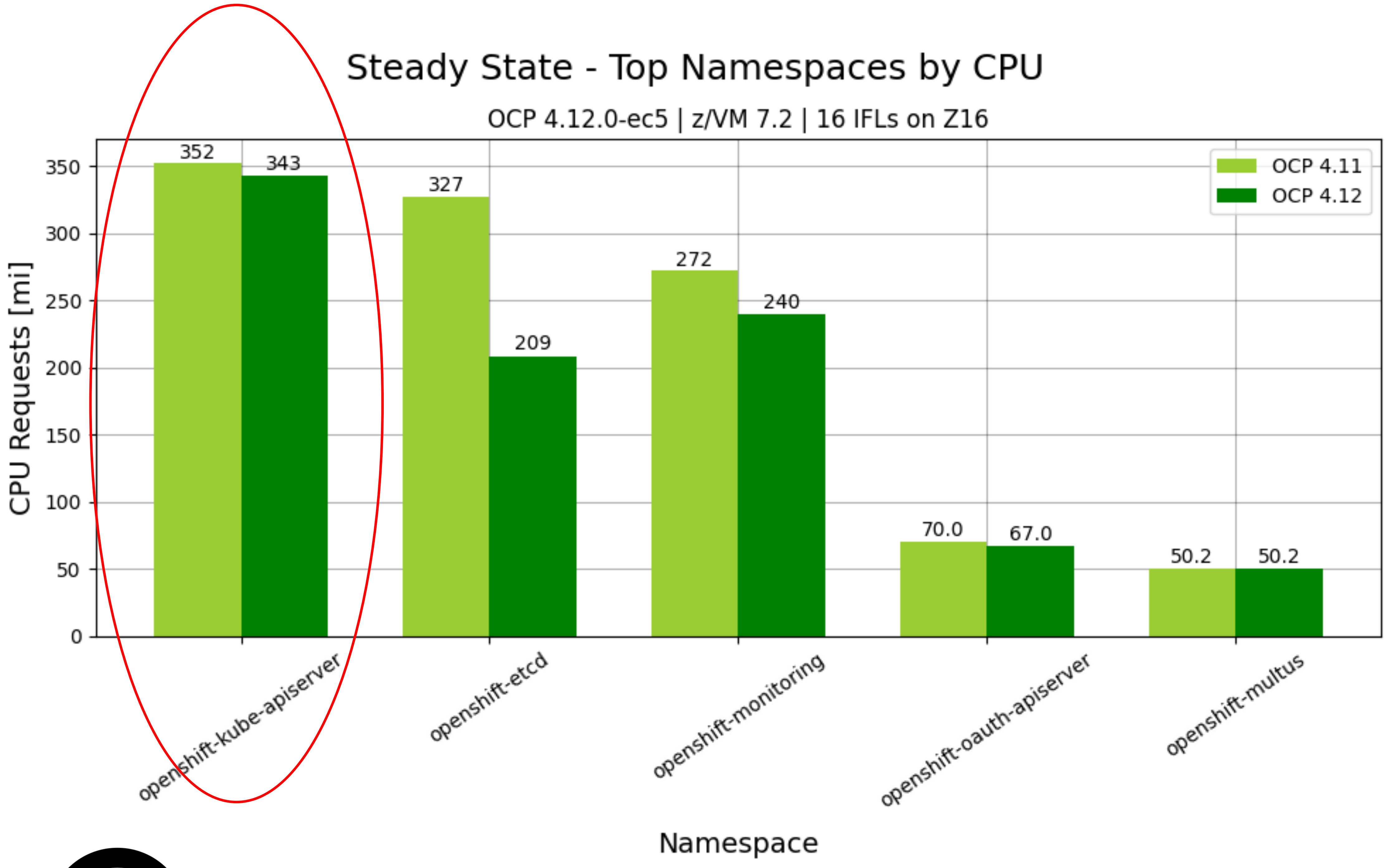
– 11% better steady state load on KVM
– Lower load peaks

- OCP 4.12 Steady state load decreased to approx. 1.6 IFLs (no additional addons)
- OCP 4.13 shows another 10% lower steady state load at approx. 1.39 IFLs (z/VM)

# Steady state: Where does it come from and why we can't get rid of it



# Steady state: Where does it come from and why we can't get rid of it



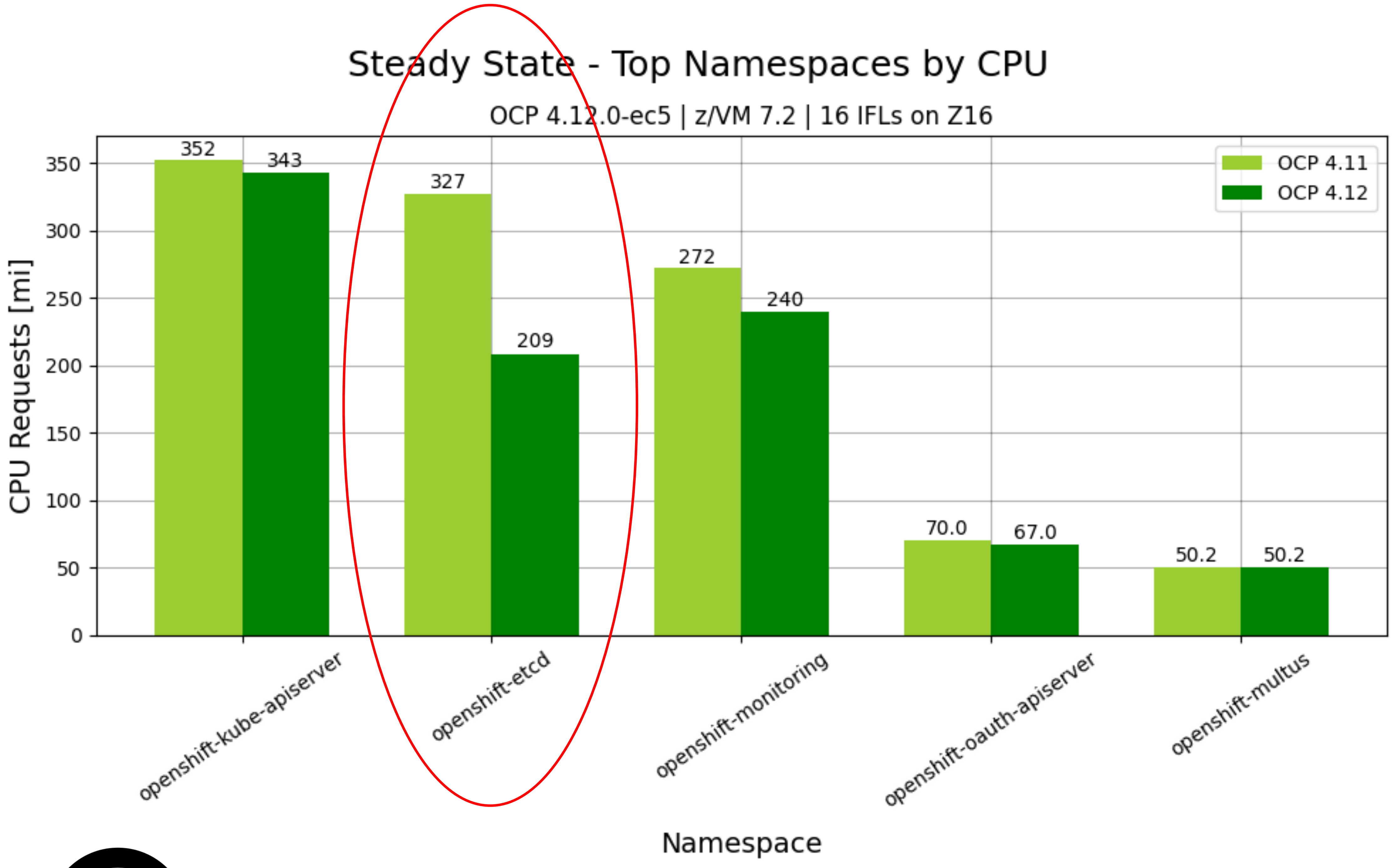

Security

kube-apiserver:

- central component of kubernetes
- handles communication of the operators
- communication broker for etcd

Enables **authentication, authorization,** and **admission controls**

# Steady state: Where does it come from and why we can't get rid of it




Security

Reliability

kube-apiserver:

- central component of kubernetes
- handles communication of the operators
- communication broker for etcd

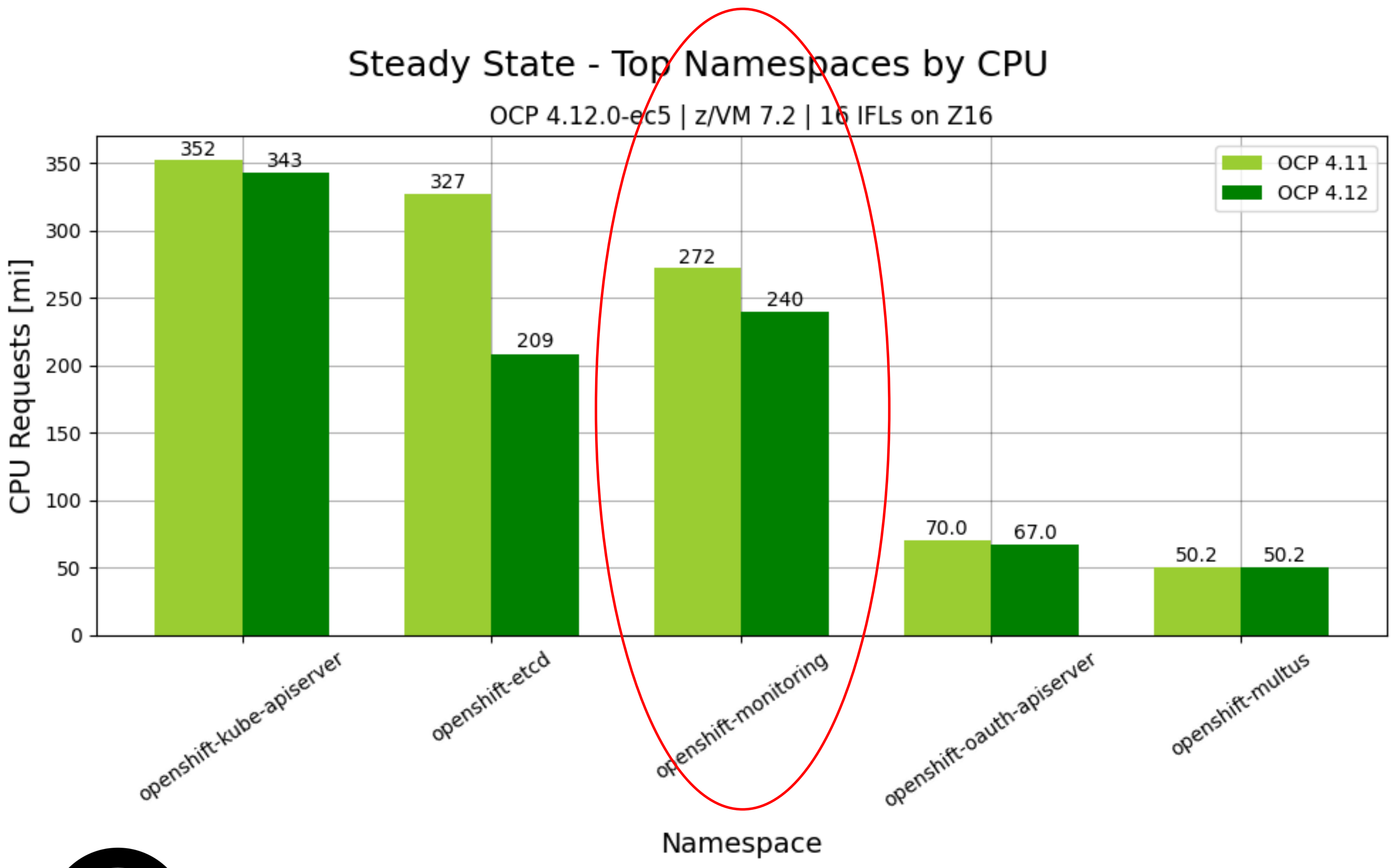
Enables **authentication, authorization**, and **admission controls**

openshift-etcd:

- central component describing the state of your cluster
- stores cluster state, networking information, and other persistent information

Among others etcd is required to restore your cluster after downtime

# Steady state: Where does it come from and why we can't get rid of it



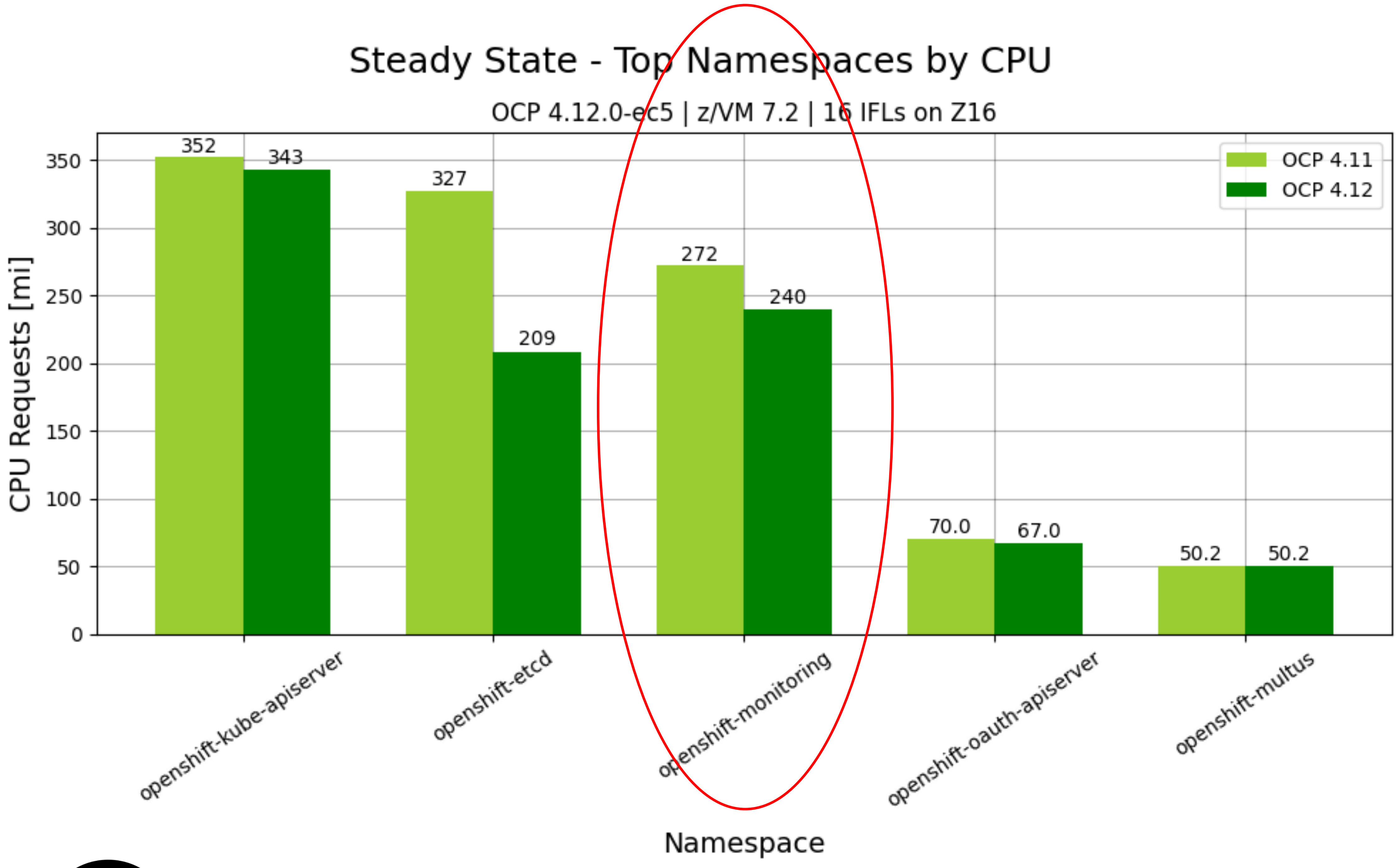

Security Reliability Observability

openshift-monitoring:

- manages and updates the Prometheus-based cluster monitoring stack
- Includes several critical components: Prometheus, Alertmanager, kube-state-metrics, node_exporter,...

**Continuous and real-time observation and monitoring** of your cluster, workloads, networks, hardware,...

# Steady state: Where does it come from and why we can't get rid of it





openshift-monitoring:

- manages and updates the Prometheus-based cluster monitoring stack
- Includes several critical components: Prometheus, Alertmanager, kube-state-metrics, node_exporter,...

✓ **Continuous and real-time observation and monitoring** of your cluster, workloads, networks, hardware,...

Many others ...

✓ kube-scheduler: Schedules pods on nodes

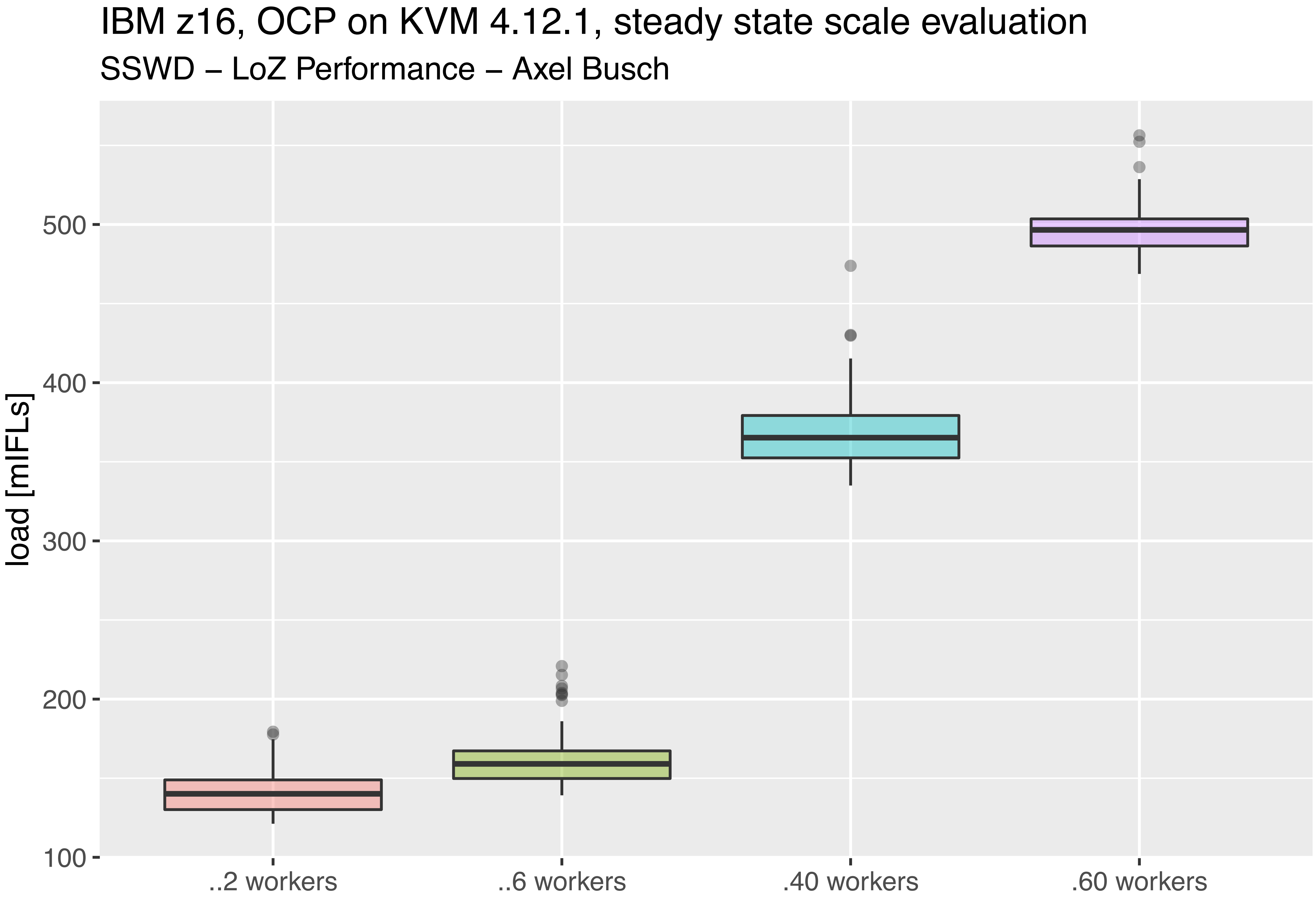
✓ kube-controller manager: Determines and adjusts cluster state

✓ OpenShift-SDN/OVN-kube/kube-proxy: Manage and make network connectivity available

# Update: Steady State IFL estimation model for KVM on z16

*Can we calculate IFL requirements for given cluster sizes before installation time?*

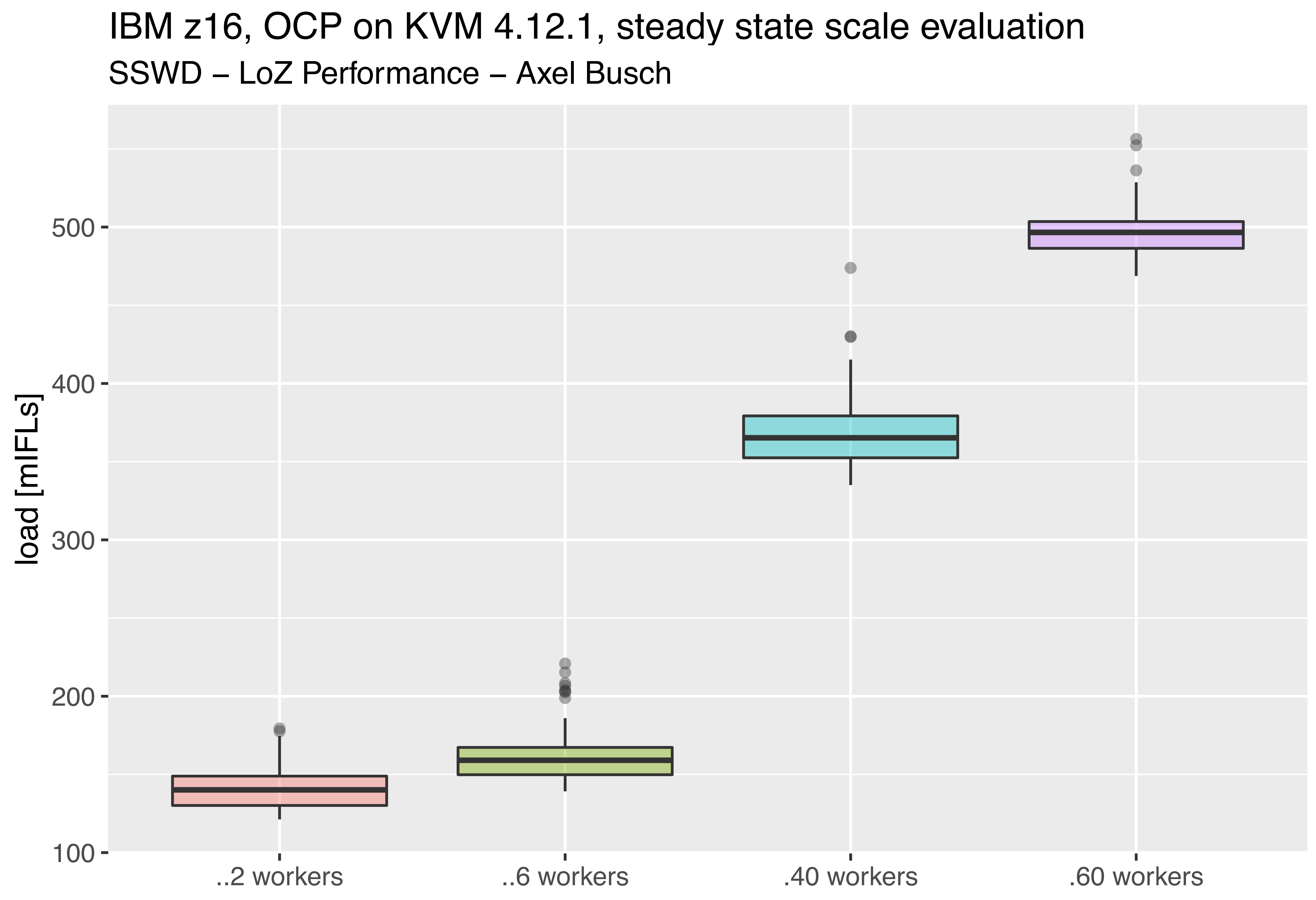
*Estimating steady state load when we install 2, 6, 40,60 compute nodes?*




Cluster setup:

- z16 KVM cluster
- LPAR with 12 IFLs + SMT-2 = 24 processors
- Red Hat OpenShift version 4.12.1 Each control/compute node with 4 vCores and 16 GiB memory
- No other processes running in the same LPAR

# Update: Steady State IFL estimation model for KVM on z16




Estimation function for OpenShift 4.12:

$$f(x) = 1.261 + x * 0.061$$

constant overhead — compute node dependant overhead

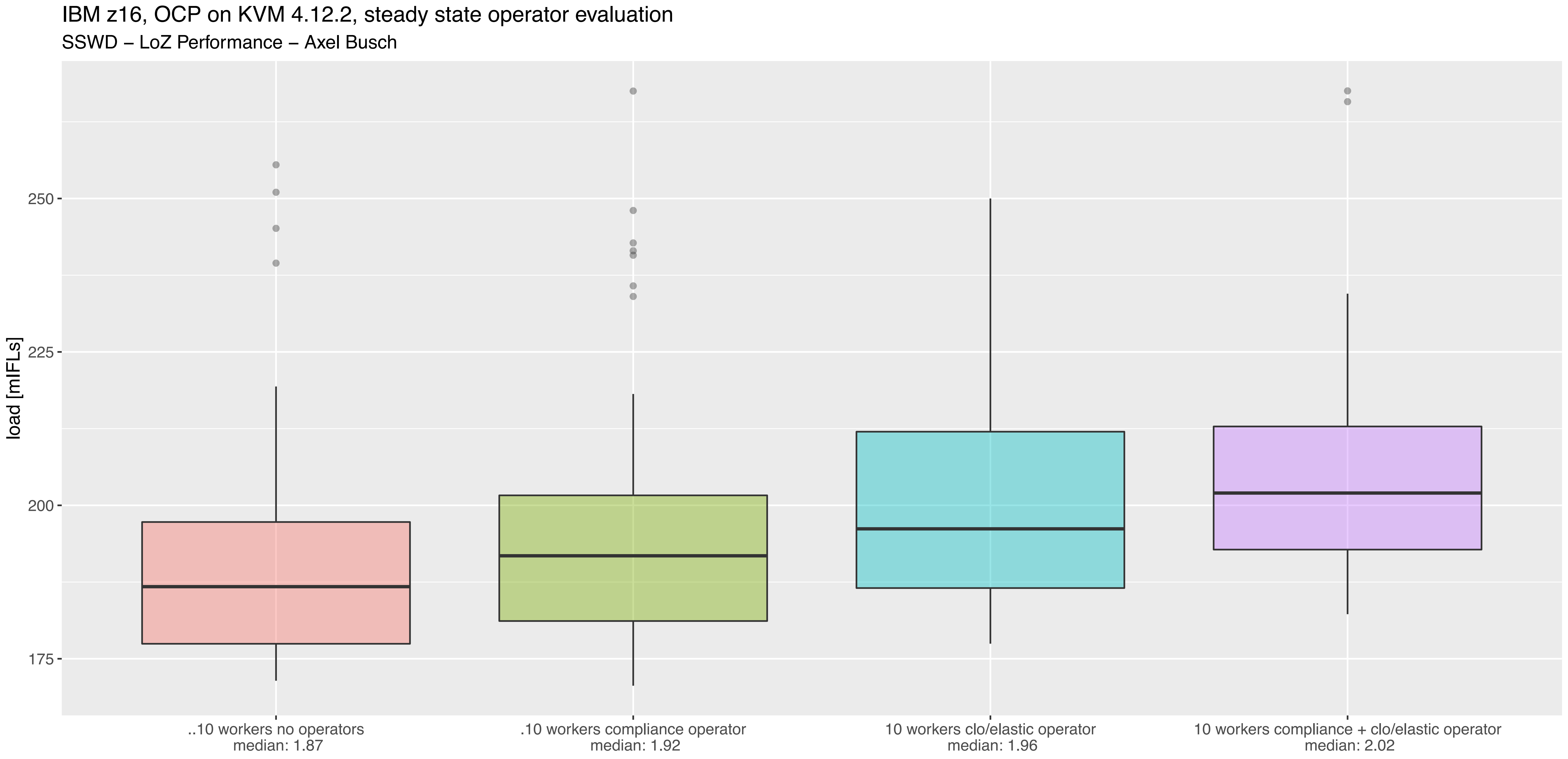
- Estimation accuracy is < 5% for 2-100 workers
- Lower constant overhead compared to 4.10
- 18% lower overhead for each node*
- No additional operators installed

* z15 vs. z16 did not show significant changes in steady state load

# Steady State IFL Estimation –
How to use the approximation model?

- For each individual setup regarding number of compute nodes, you can **forecast the minimum required IFLs** for steady state before installation time
- Use the model for estimating minimum IFL requirements: If 60 compute nodes are required it makes sense to think about using at least 5 IFLs
- Cluster efficiency goes up when using less **compute nodes but instead installing more virtual CPU in each compute node.** This improves cluster steady state efficiency as well as workload efficiency

# Steady state with additional operators

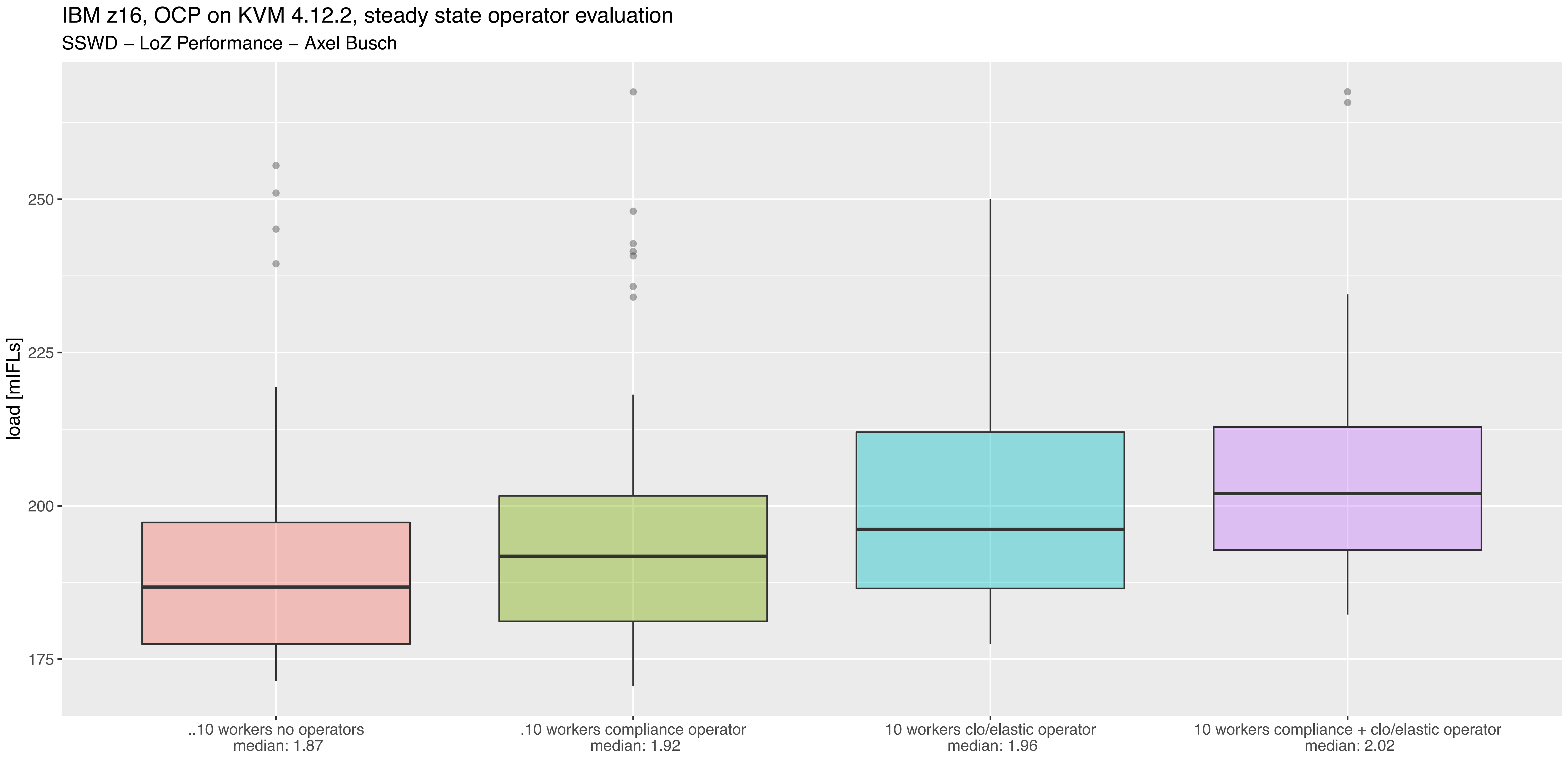



| operator | z16 IFL | ratio |
|---|---|---|
| OpenShift base | 1.87 | 1.0 |
| Compliance | 1.92 | 1.027 |
| CLO + Elastic | 1.96 | 1.048 |
| Compliance + CLO + Elastic | 2.02 | 1.08 |
| Compliance + CLO + Elastic + FIPS* | 2.40* | 1.28* |

* Extrapolated numbers from FIPS Red Hat OpenShift 4.11

→ Quite comfortable within the 3+3 IFL offering

# Steady state with additional operators




- Operators are just installed, no instances created (except ODF)
- Overhead seems to be constant: does not increase with additional compute nodes significantly

| operator | z16 IFL | ratio |
|---|---|---|
| OpenShift base | 1.87 | 1.0 |
| Compliance | 1.92 | 1.027 |
| CLO + Elastic | 1.96 | 1.048 |
| Compliance + CLO + Elastic | 2.02 | 1.08 |
| Compliance + CLO + Elastic + FIPS* | 2.40* | 1.28* |
| ODF 4.12** | 3.23** | 1.72** |
| ODF + Compliance + CLO + Elastic + FIPS*** | 3.76*** | 2.01*** |

- Extrapolated numbers from FIPS Red Hat OpenShift 4.11
- ** Due to a bug, the value was adjusted for the overhead
- *** Calculated and adjusted

# 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.