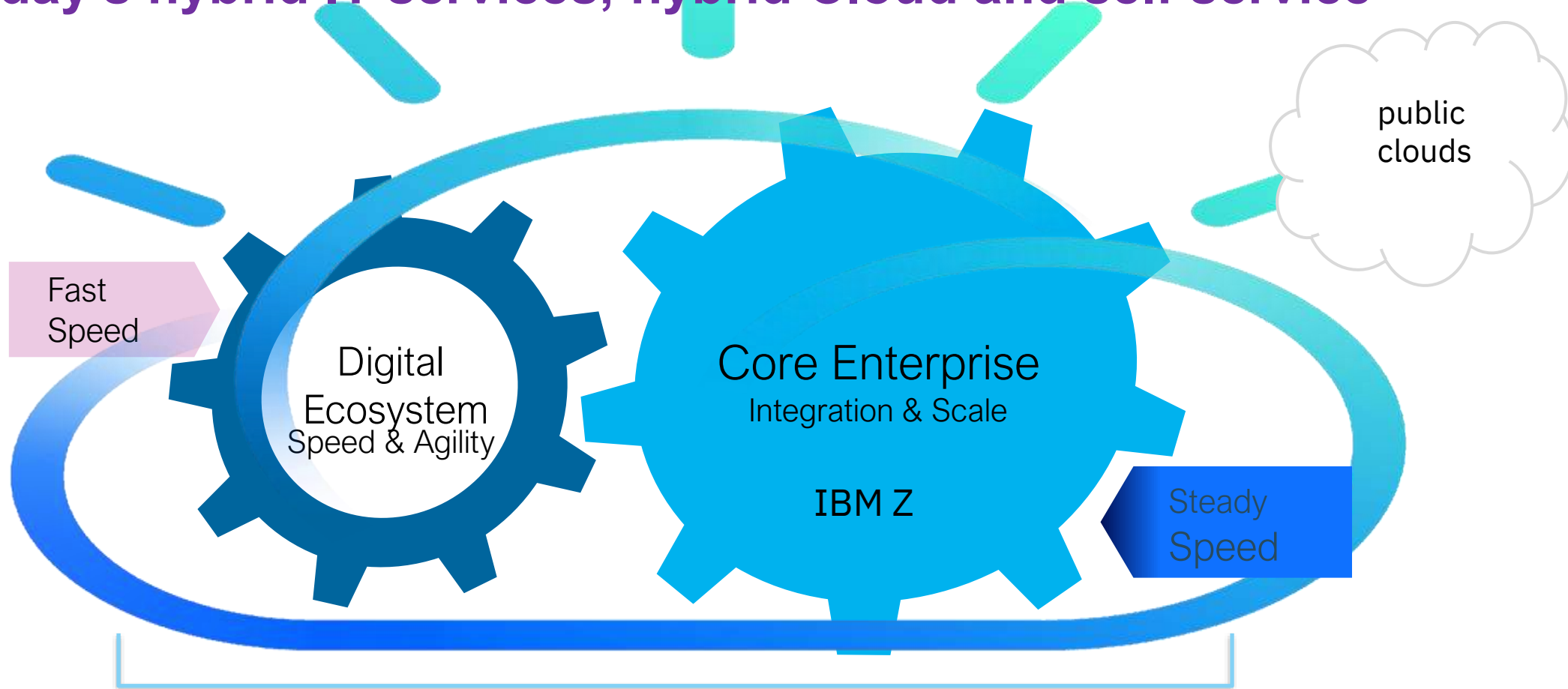


Network topologies for Linux on IBM Z and z/OS co-location and Hybrid Cloud solutions

session 201

Wilhelm Mild
IBM Executive IT Architect
IBM R & D Lab Germany
wilhelm.mild@de.ibm.com

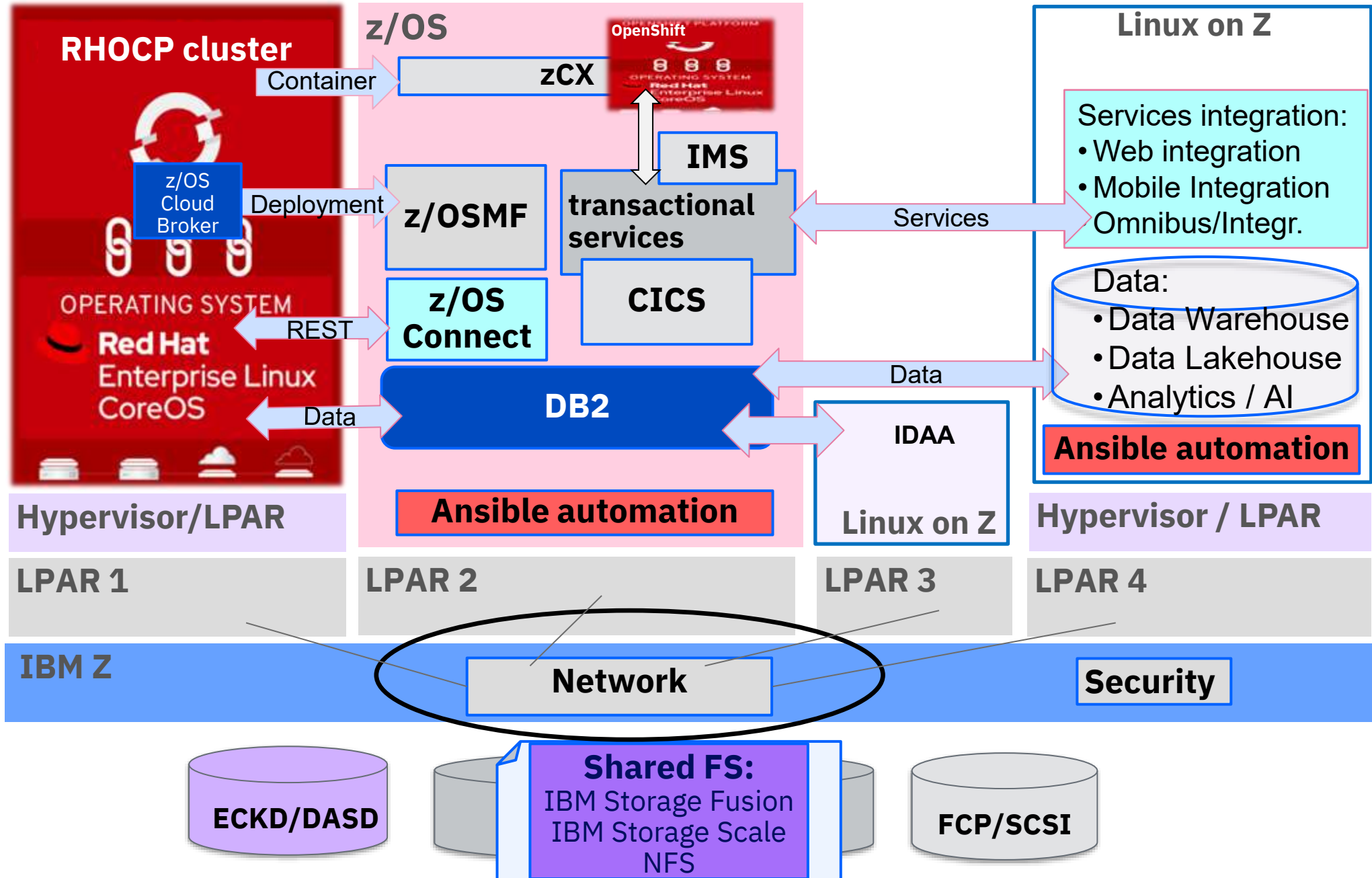
Today's hybrid IT services, hybrid Cloud and self service



Hybrid service orchestration and traditional transactional & data services orchestrator

The cloud service model with end-to-end orchestration capabilities

Overview: z/OS co-location with Linux and RH OpenShift on IBM Z



Network Adapters and internal network options

- **OSA** – Technology that has matured and is replaced with the new **Network Express** card in z17
 - is used for ETH networks
- **RoCE** – RDMA (Remote Direct Memory Access) over Converged ETH has **Network Express** as follow-up card
 - used for ETH or SMC networks
- **Network Express** – new Network card with **IBM z17 & IBM LinuxONE 5**
 - replaces as converged PCI card the OSA and RoCE cards for all protocols

Note: Network Cards work asynchronous and offload network handling, and use their own Clock speed Chips

<https://www.ibm.com/docs/en/linux-on-systems?topic=tips-networking>

<https://www.ibm.com/docs/en/linux-on-systems?topic=hypervisors-networking>

Networking Adapter Convergence in IBM z17 & LinuxONE 5

Up to IBM z16 &
LinuxONE 4

Channel Device

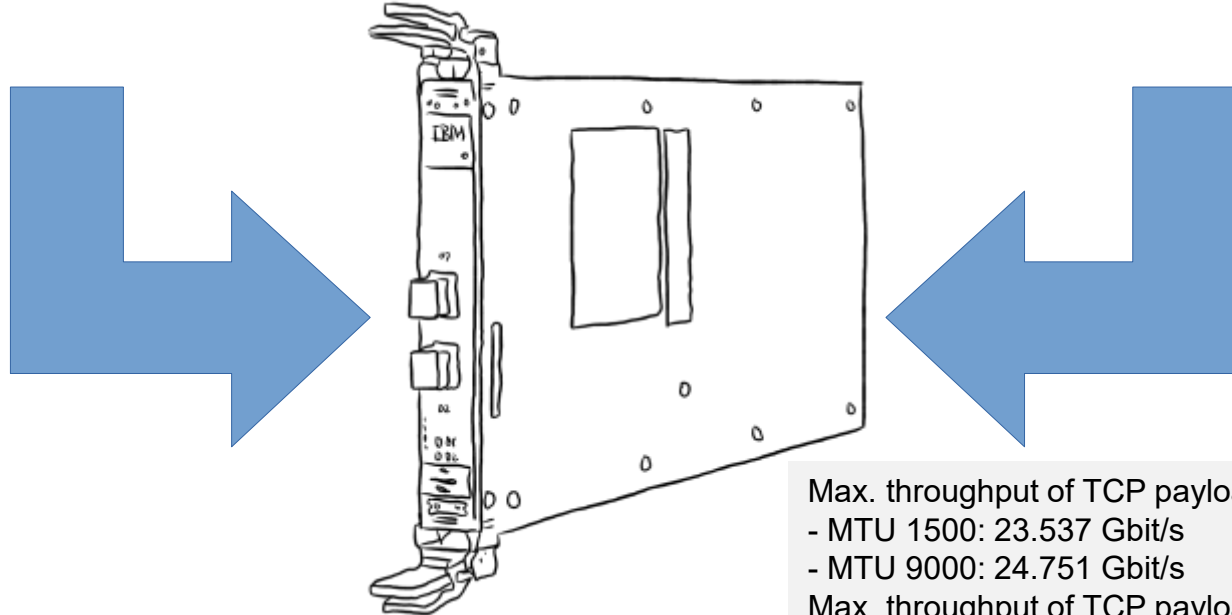


OSA-Express7S

PCI Device



RoCE Express3

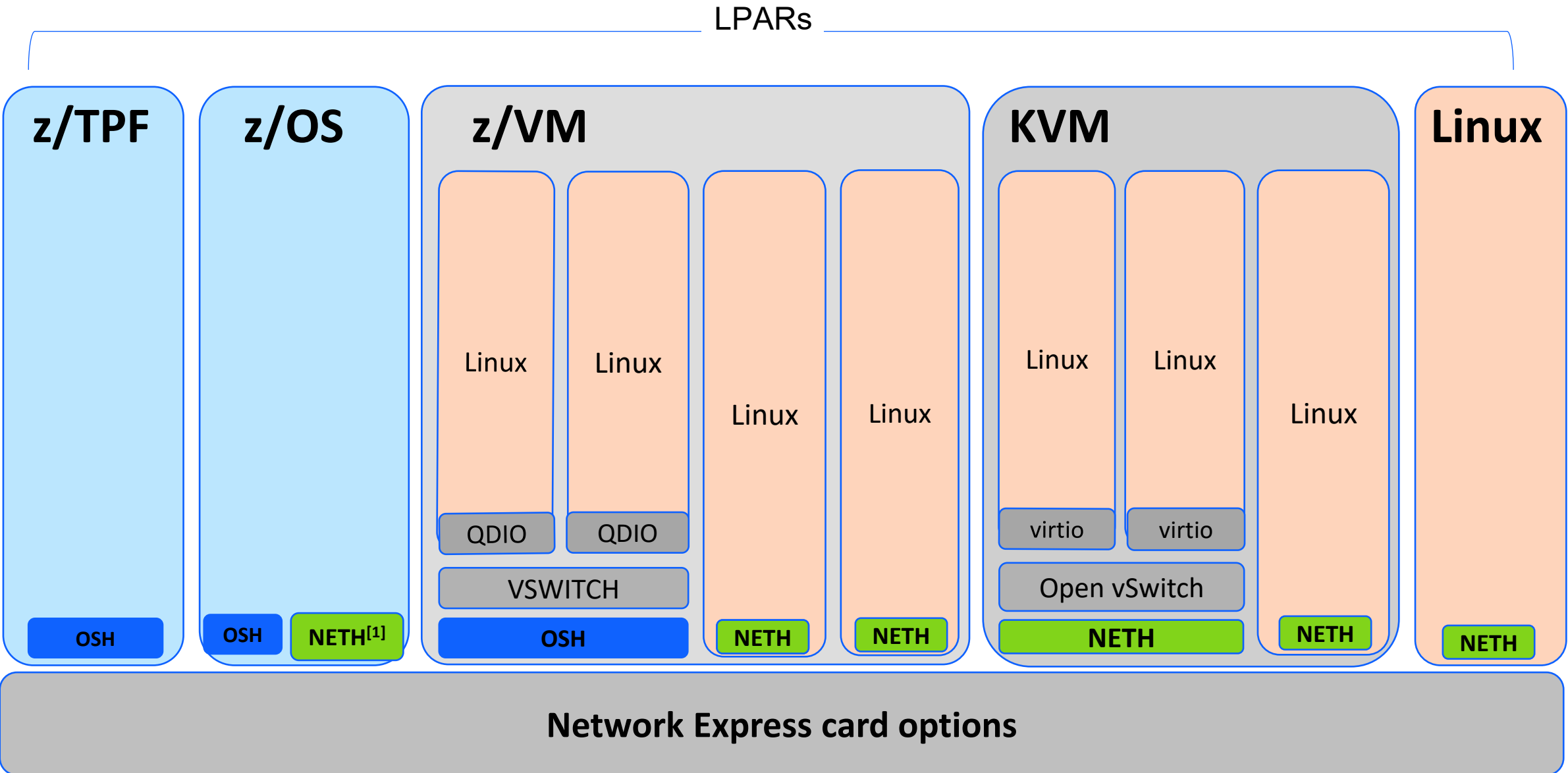


Network Express

Max. throughput of TCP payload on a 25 GbE line:
- MTU 1500: 23.537 Gbit/s
- MTU 9000: 24.751 Gbit/s
Max. throughput of TCP payload on a 10 GbE line:
- MTU 1500: 9.415 Gbit/s
- MTU 9000: 9.900 Gbit/s

- The new **Network Express** adapter can be configured as both at the same time:
 1. **OSH** channel device for z/OS and z/VM VSWITCH using Enhanced QDIO (EQDIO)
 2. **NETH PFID** (PCI device) for Linux
- Support for 10 GbE and 25 GbE

Host Operating System Converged Network Adapter Usage



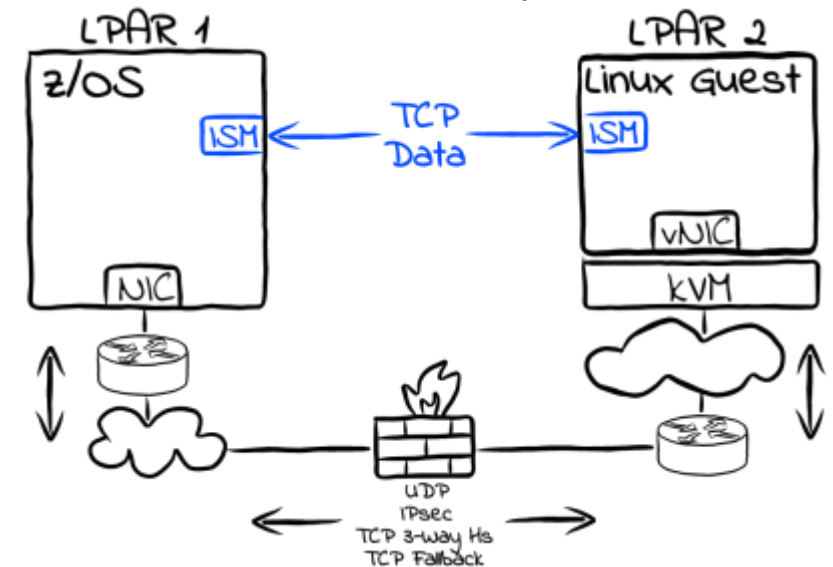
[1] SMC-R only

Network Adapters and internal network options

- **Hipersockets** – can have different / more MTU sizes for best performance
 - it is using CPU cycles at CPU Speed – if capped capacity – network can be impacted
 - up to 32 isolated networks per system possible
 - no HA for Hipersockets itself – but using HSCI – failover to another network
- **SMC** – Shared Memory Communication is an internal Network protocol for IP traffic using Internal Shared Memory (ISM) devices in IBM Z & LinuxONE via **RoCE cards**
 - **SMC-D** for IP traffic (not UDP) between LPARS / **SMC-R** for IP traffic between machines/ CECs
- **Shared ports** of OSA/RoCE cards – performance varies

Note:

Internal networks work synchronous (internal buffer copy)
(Hipersockets can use async. Completion Queue if needed)
but they use CPU cycles – which can impact network speed
with capped or limited CPU capacity

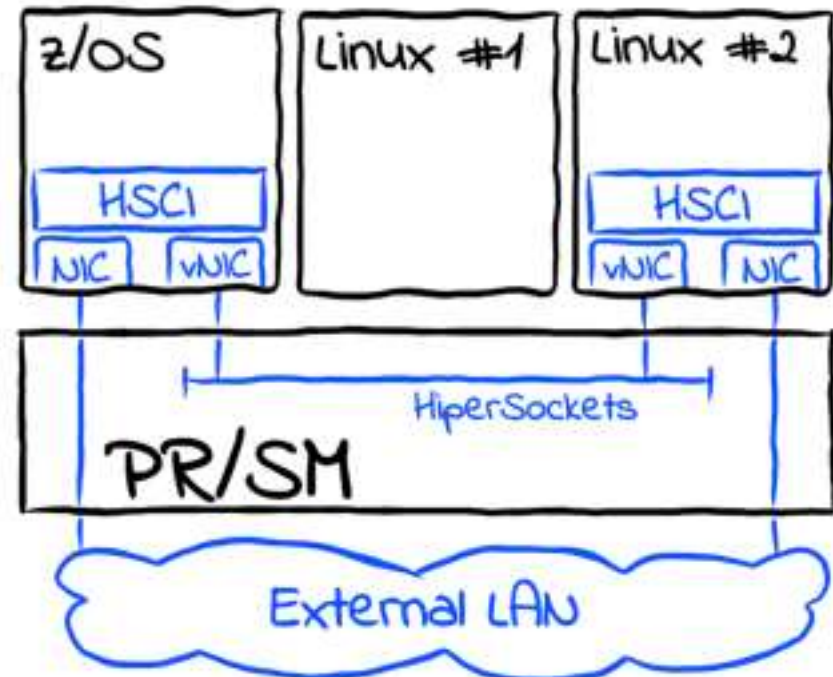


Network general best practices

Combine Networks for higher throughput and High Availability:

- **Link aggregation with 2 OSA (LACP)** for higher throughput and HA
- **Linux Channel Bonding** with the OSA ports / cards for higher throughput and HA
- **Use HSCI**, Hipersockets Converged Interface for simplification of multiple networks in the same server - supported in Linux / z/VM / KVM and z/OS

https://www.ibm.com/docs/en/linuxonibm/pdf/z_performance_of_network_adapters_Linux_on_IBM_Z.pdf



Network topologies for environments with Linux on IBM Z and z/OS Co-location:

1) Transactional interaction

2) Multi Data aggregation

3) Hybrid Cloud solutions

Network topologies for use cases with Linux on IBM Z and z/OS Co-location:

1) Transactional interaction characteristics

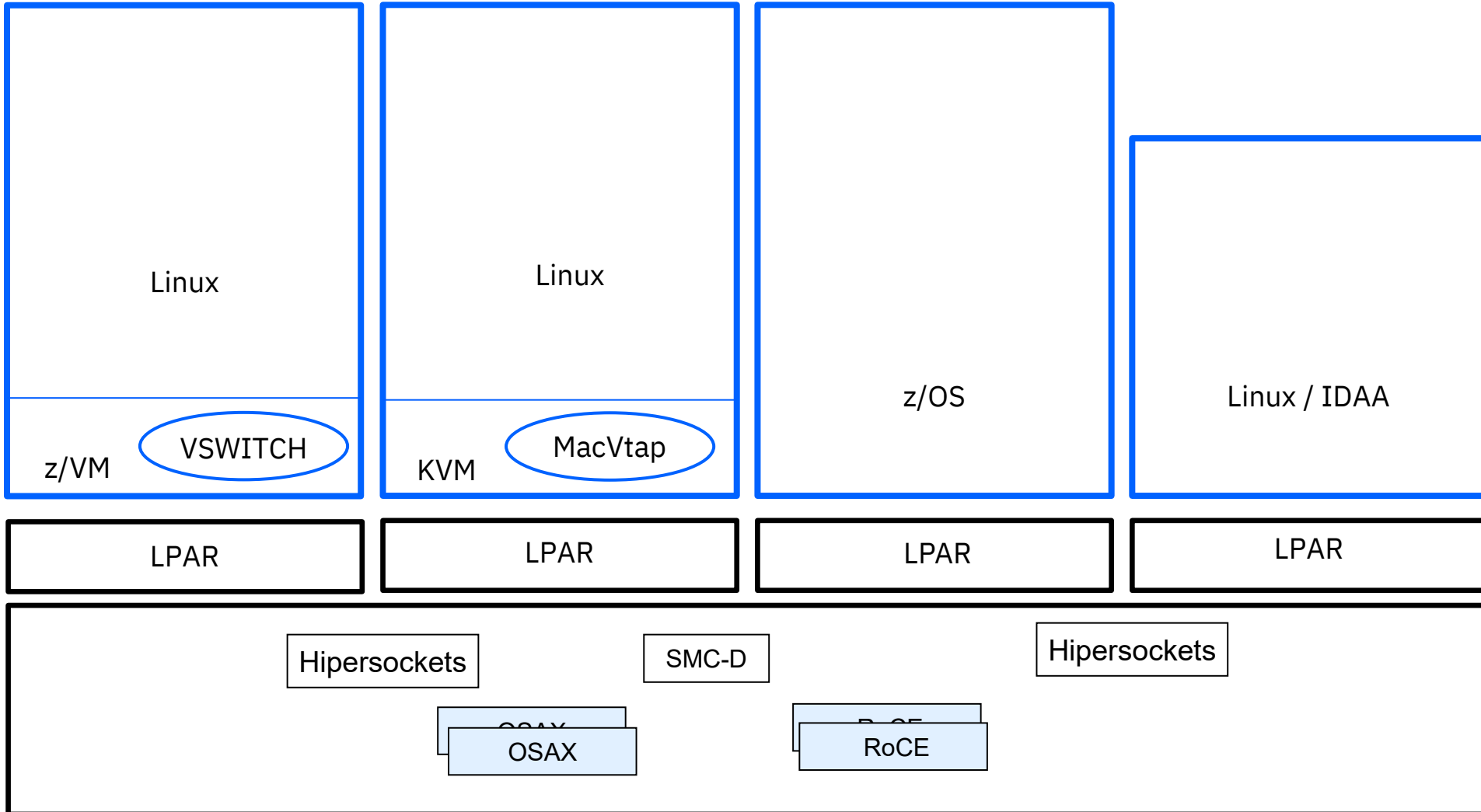
- High number of concurrent connections
- Short lived connections
- Fast turn-around response time
- Point to point connection

1) Transactional interaction characteristics

Best Practices

- Incoming/Outgoing traffic: OSA/RoCE
- Internal traffic between LPARs and Linux guests:
 - make use of Linux Channel BONDing for higher bandwidth
 - shared ports with RoCE instead of OSA
 - Hipersockets if CPU capacity available only – up to 32 individual networks
 - SMC-D if larger data movement involved over IP – it also requires an additional ETH interface (for handshaking)
- Hypervisor dependencies
 - make use of Link aggregation for higher bandwidth
 - use multiple z/VM VSWITCH for isolation and throughput
 - KVM & OpenShift Virt. (OCPV) needs a bridge to Hipersockets
 - [Improving Transactional Workloads in KVM Guests](#)
- HA/DR considerations
 - Hipersockets can be bridged across systems, not possible with SMC-D

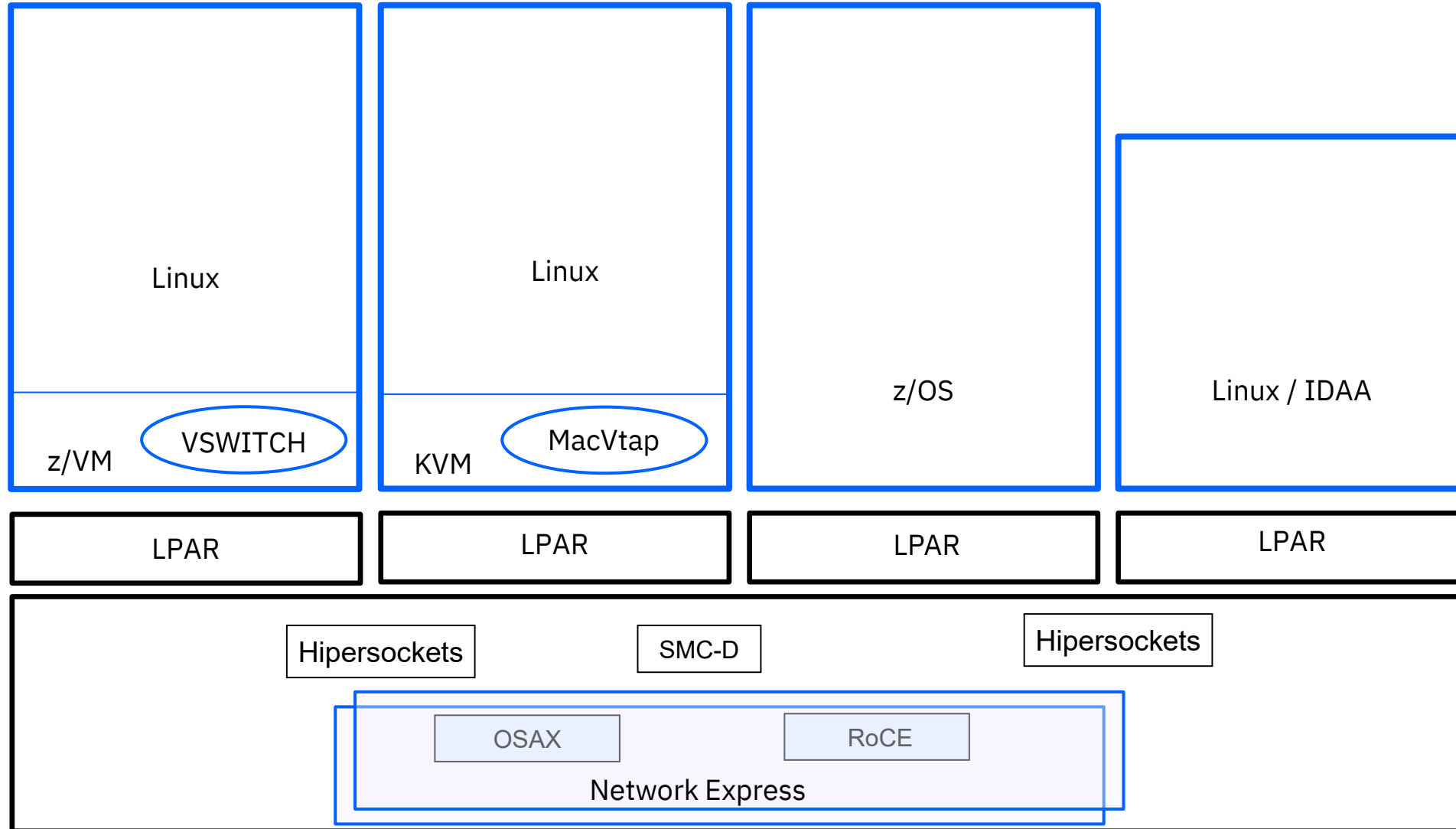
1) Transactional interaction –network topology



Best Practices

- MTU size - small
- OSA for z/VM
- OSA for KVM
- ? RoCE SMC-D
- Hipersockets –fast
- Consider HA/DR

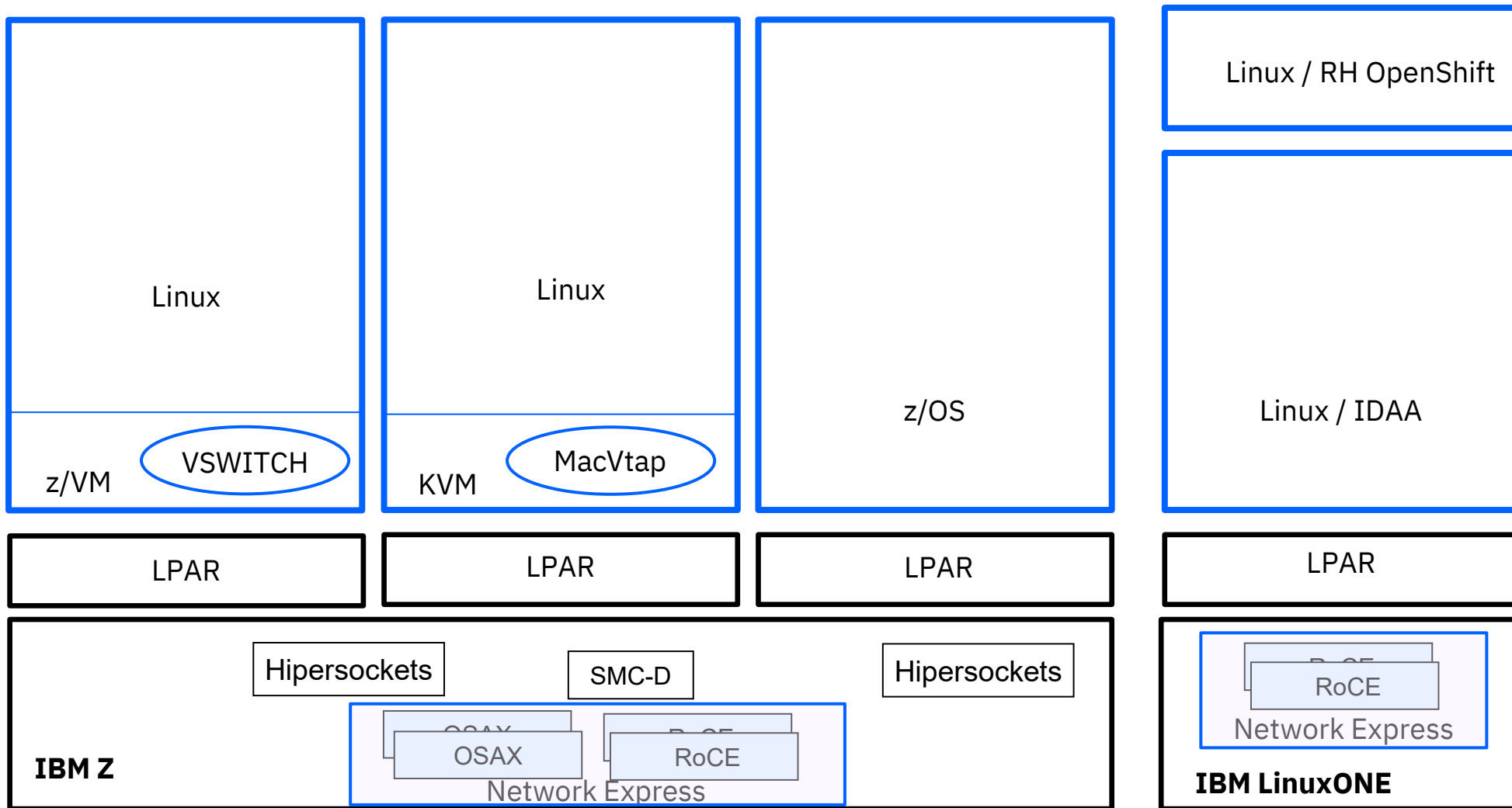
1) Transactional interaction –network topology z17



Best Practices

- MTU size - small
- OSH for z/VM
- NETH for KVM
- NETH for SMC-D/R
- Hipersockets –fast
- Consider HA/DR

1) Transactional interaction – network /w LinuxONE



Best Practices

- **RoCE /SMC**
- **OSH for z/VM**
- **NETH for KVM**
- **NETH for SMC-D/R**
- **Consider HA/DR**

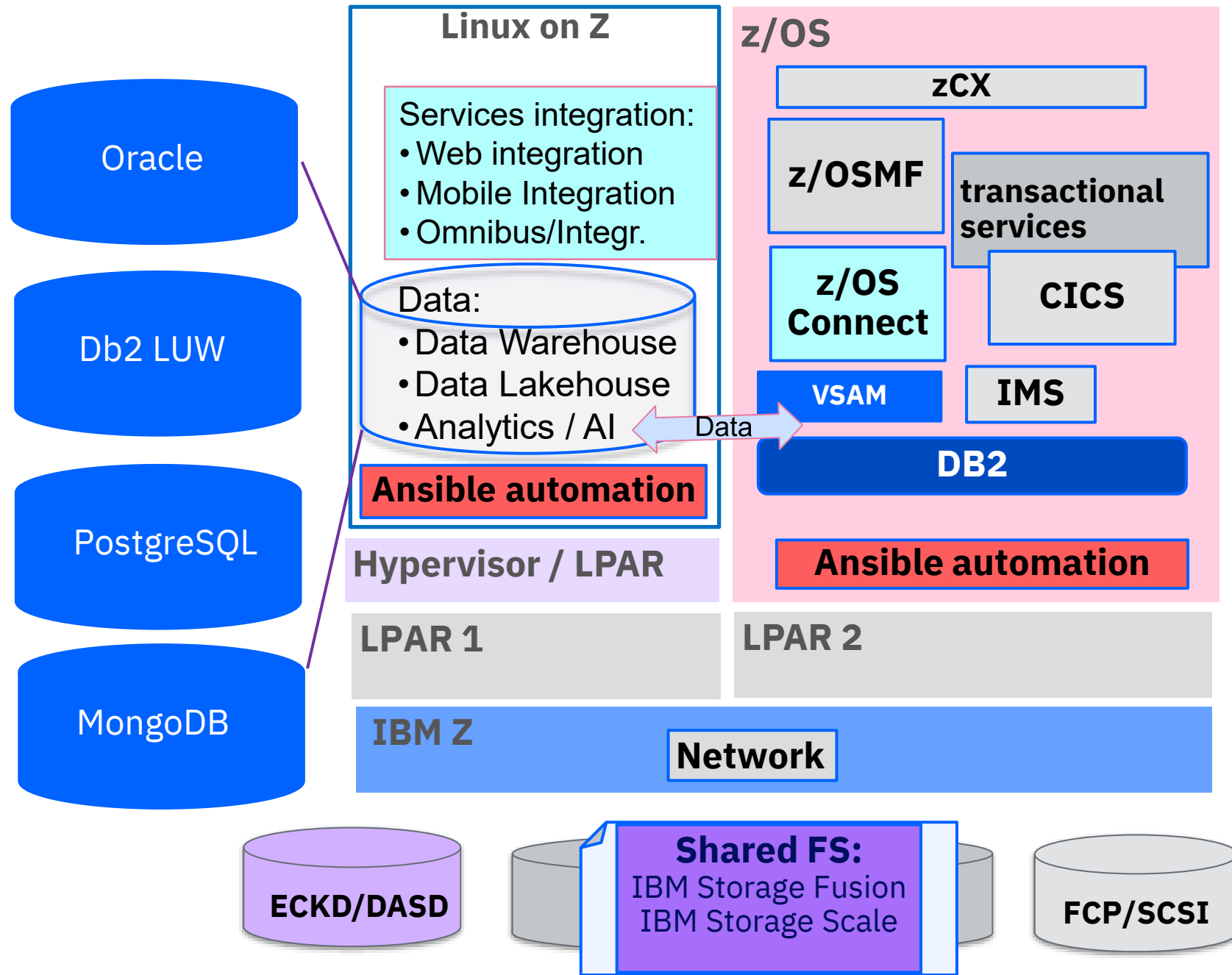
Network topologies for use cases with Linux on IBM Z and z/OS Co-location:

1) Transactional interaction

2) Multi Data aggregation

3) Hybrid Cloud solutions

Overview: Build a highly scalable Data Hub on Linux on Z co-located with z/OS



Network topologies for use cases with Linux on IBM Z and z/OS Co-location:

2) Multi Data aggregation characteristics

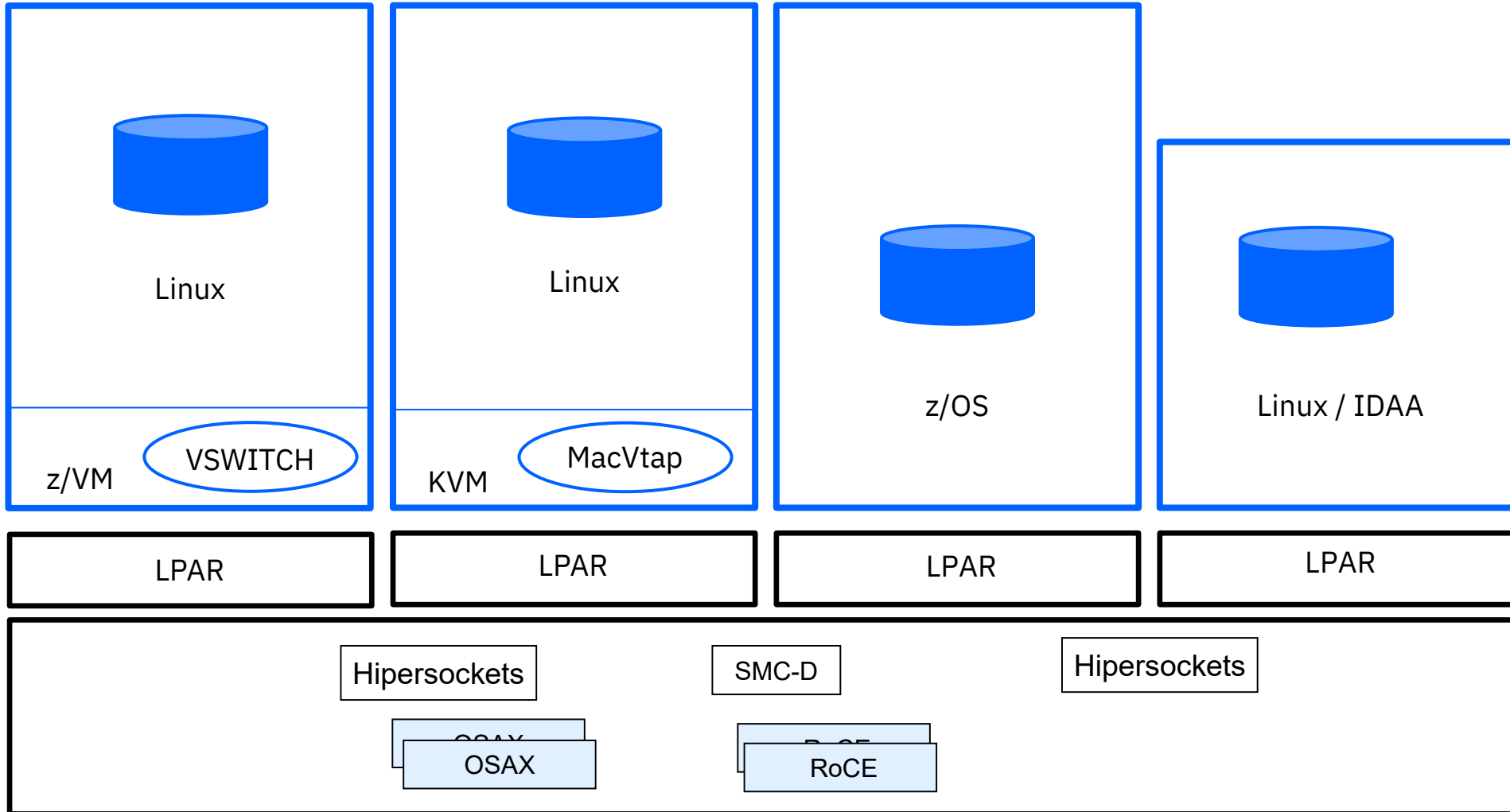
- high throughput
- reusable concurrent connections
- two phase commit series of interactions
- point to point connection
- often larger data packets than transactional

2) Data aggregation characteristics

Best Practices

- Enable Linux Channel bonding and OSA Link Aggregation
- Incoming/Outgoing data: OSA/RoCE
- Internal traffic between LPARs and Linux guests:
 1. Hipersockets if CPU capacity available
 2. SMC-D/R with RoCE for IP data movement (UDP not affected)
 3. shared ports with RoCE/OSA for capped capacity systems
- Hypervisor dependencies
 - z/VM VSWITCH not compatible with RoCE
 - KVM / OCPV needs a bridge to Hipersockets
- HA/DR considerations
 - Hipersockets can be bridged across machines
 - SMC-D can not be bridged
 - best to use RoCE/OSA/Network Express for failover to DR

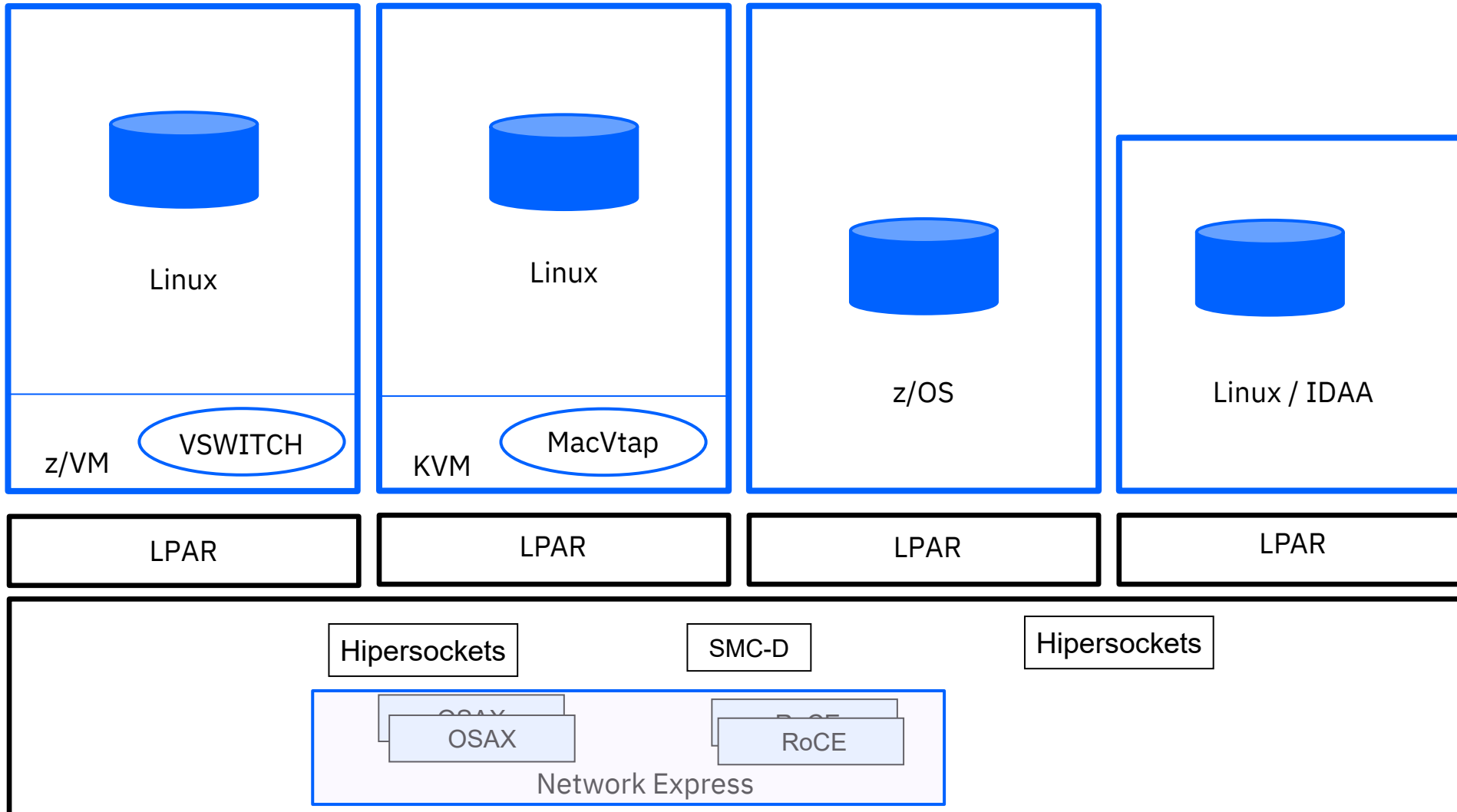
1) Data transfer –network topology



Best Practices

- MTU size - high
- Hipersockets - fast
- OSA for z/VM (VS)
- OSA for KVM
 - or SMC-D
- RoCE SMC-D
 - if possible

1) Data transfer –network topology IBM z17



Best Practices

- **MTU size - high**
- **Hipersockets - fast**
- **OSH for z/VM (VS)**
- **NETH for KVM**
 - or **SMC-D**
- **NETH for SMC-D**
 - if possible
- **Consider HA/DR failover**

Network topologies for use cases with Linux on IBM Z and z/OS Co-location:

1) Transactional interaction

2) Secure Data aggregation

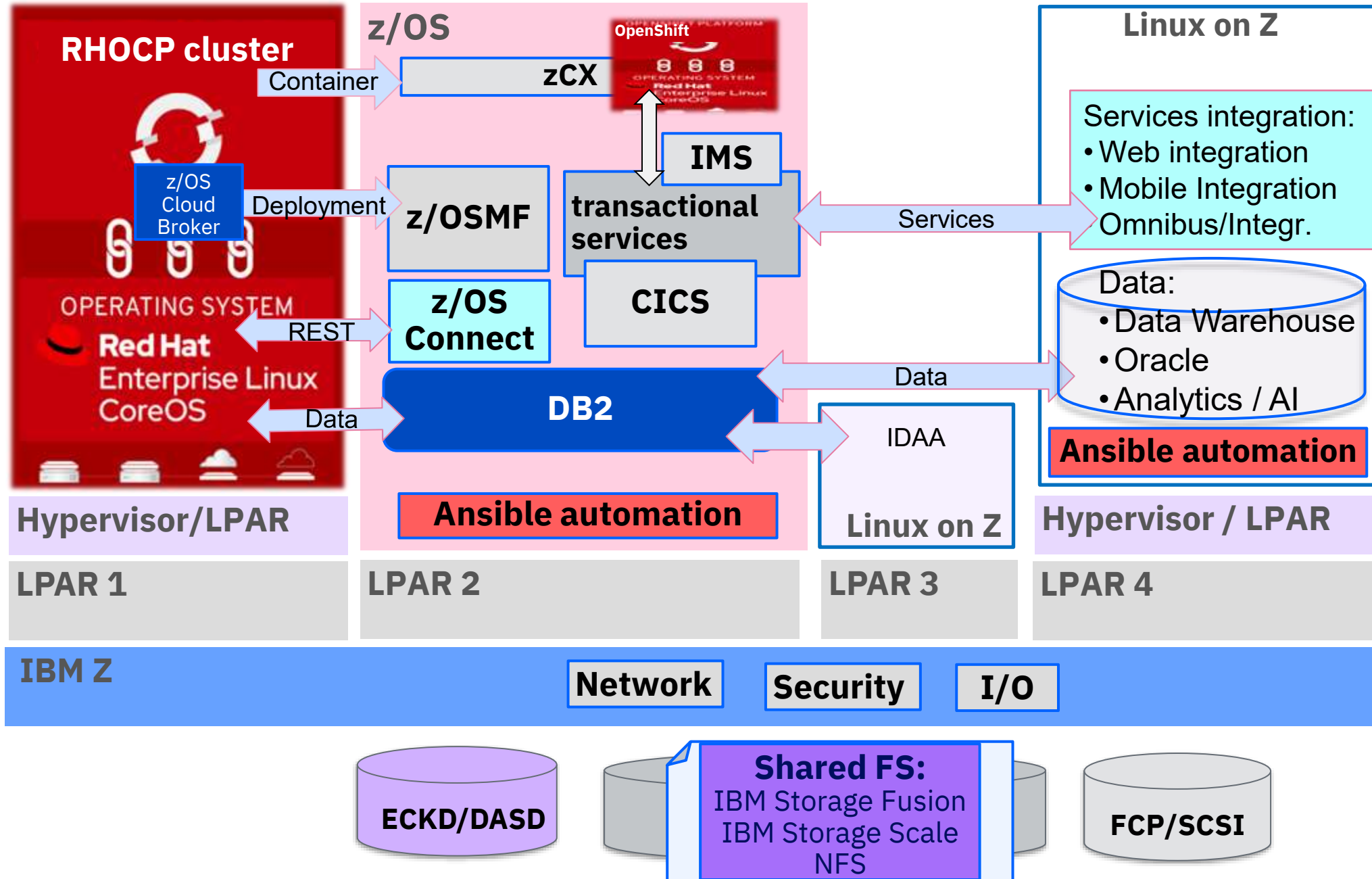
3) Hybrid Cloud solutions

Network topologies for use cases with Linux on IBM Z and z/OS Co-location:

3) Hybrid Cloud solutions characteristics

- multi segment network traffic
- multi network activities (OCP control vs. data traffic)
- broadcast connections generate more traffic
- mixed data packets control / data

Overview: z/OS co-location with Linux and RH OpenShift on IBM Z



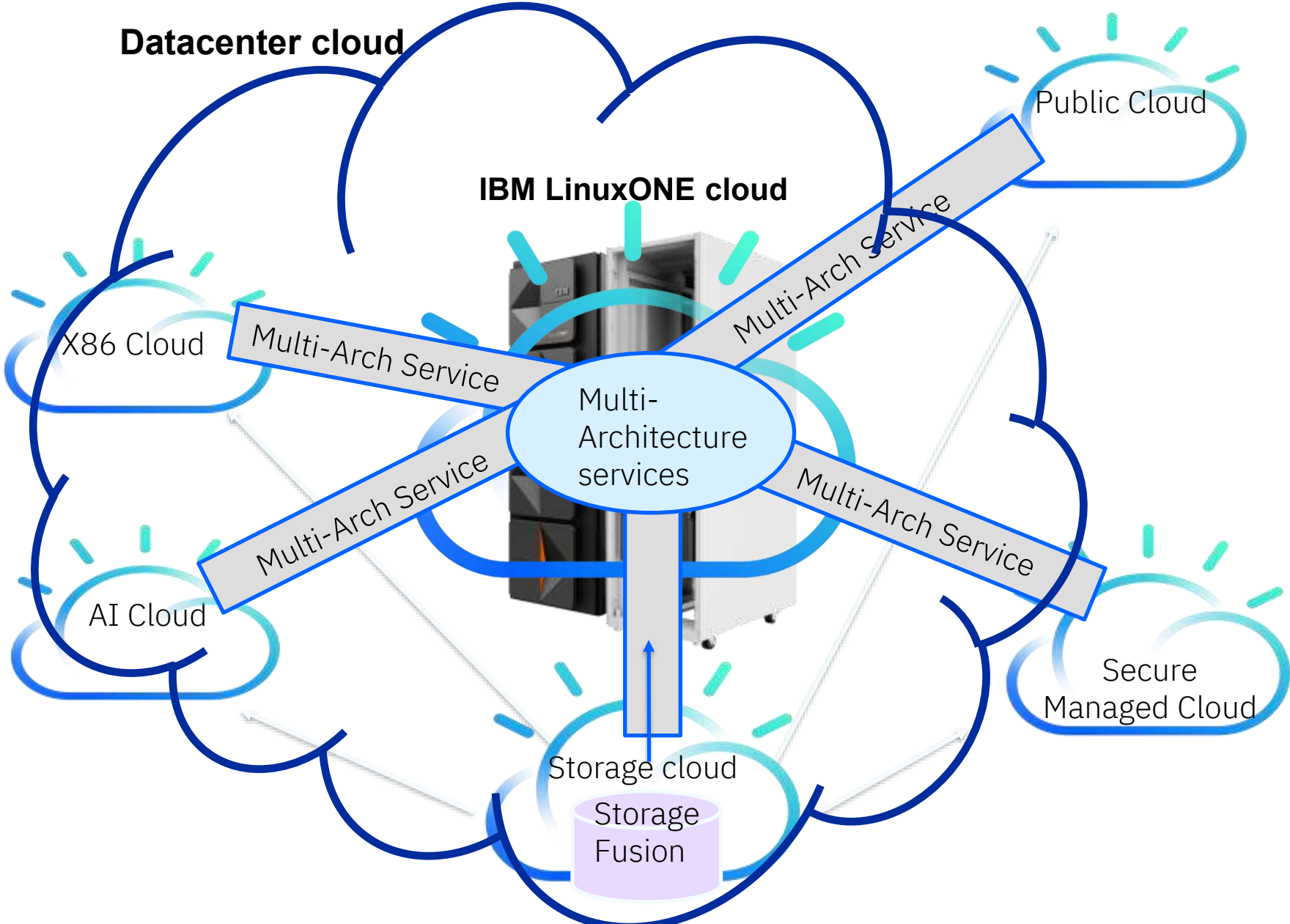
Datacenter Multi-Architecture cloud services with IBM LinuxONE

Common Cloud services characteristics

- Self Service
- Automation
- Charge back

DevSecOps

- Develop once
- Deploy multi-Arch
- Run on best fit

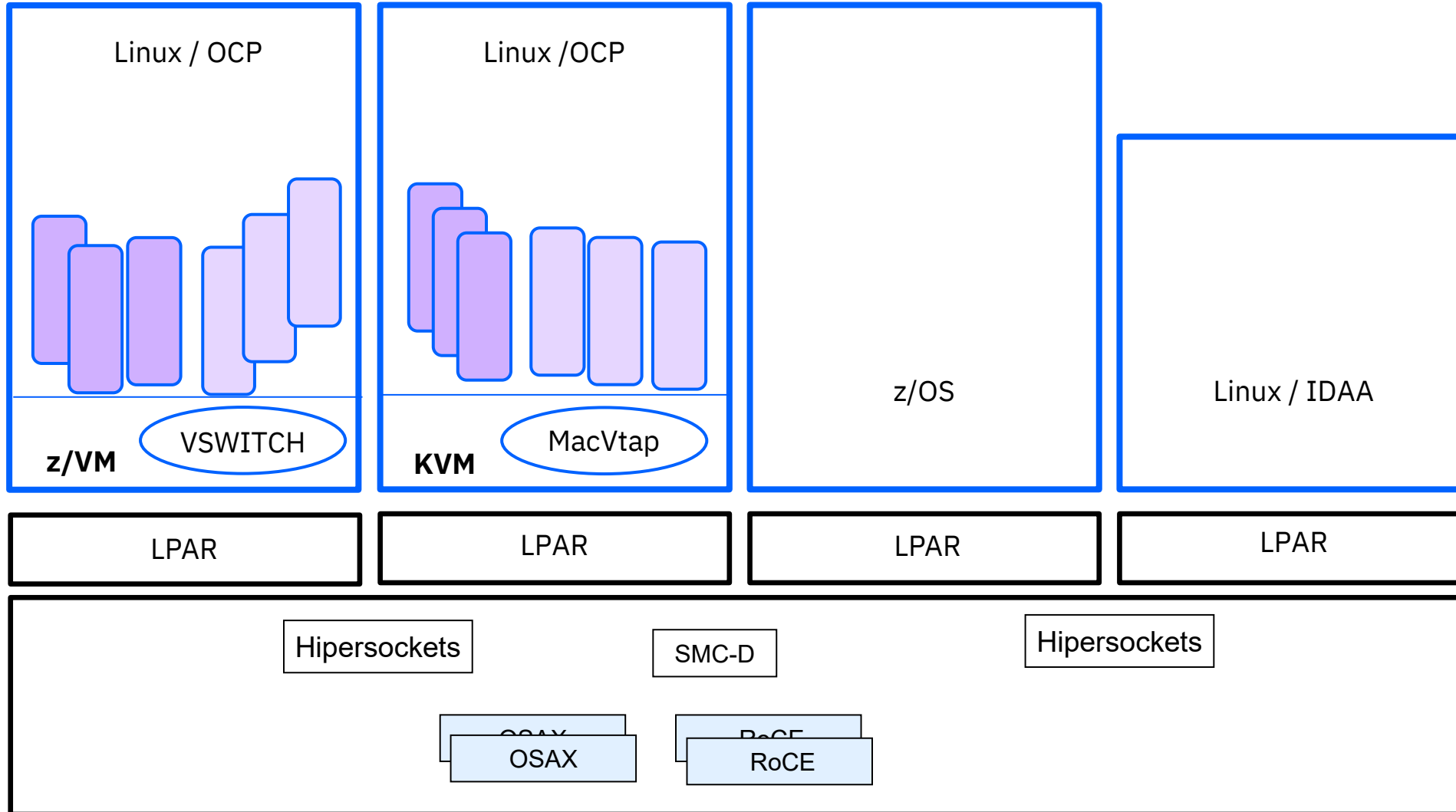


3) Hybrid Cloud characteristics

Best Practices

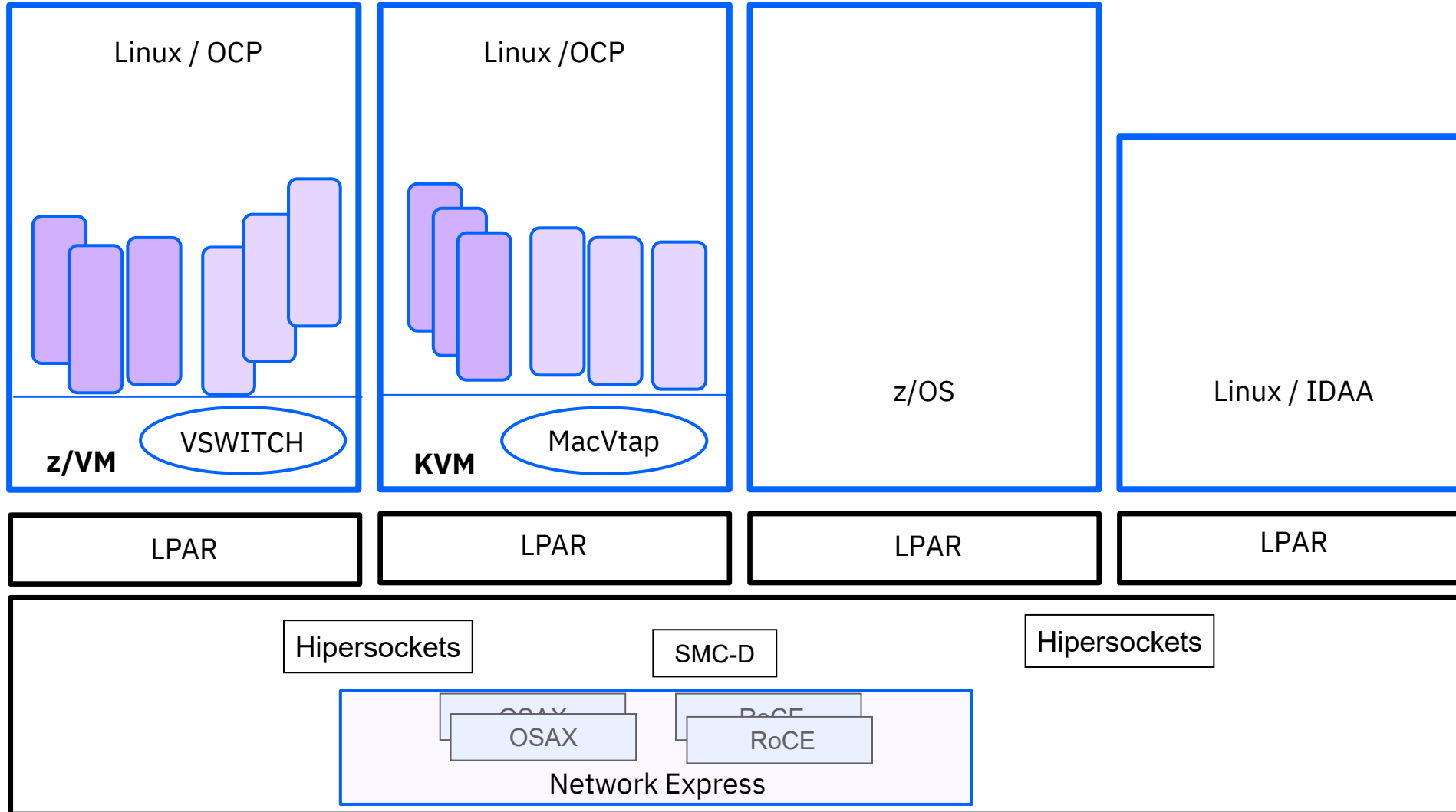
- Incoming/Outgoing data: OSA/RoCE/Network Express
- Internal traffic between LPARs and Linux/OpenShift nodes:
 1. Hipersockets if CPU capacity available
 2. shared ports with RoCE/OSA for capped capacity
 3. SMC-D not yet available for OpenShift
- Hypervisor dependencies
 - z/VM VSWITCH not compatible with RoCE
 - KVM / OCPV needs a bridge for Hipersockets
- HA/DR considerations
 - Hipersockets can be bridged across machines
 - SMC-D can not be bridged

1) Hybrid Cloud –network topology



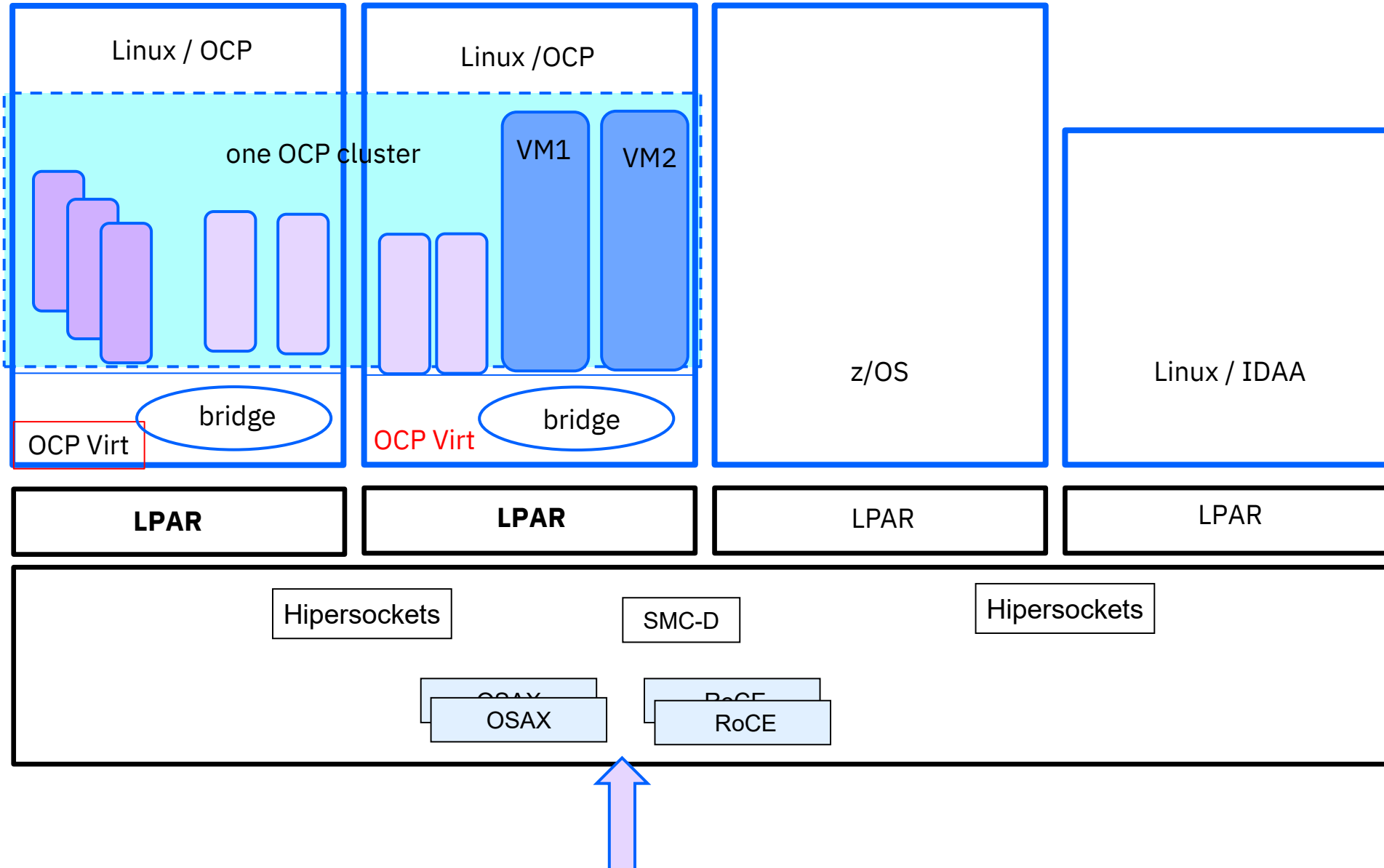
- ### Best Practices
- Multiple Networks
 - Zoned Networks
 - MTU size mixed
 - Hipersockets
 - ok if possible
 - OSA for z/VM (VS)
 - RoCE for SMC-D
 - if possible

1) Hybrid Cloud –network topology



- ### Best Practices
- Multiple Networks
 - Zoned Networks
 - MTU size mixed
 - Hipersockets
 - ok if possible
 - OSA for z/VM (VS)
 - RoCE for SMC-D
 - if possible

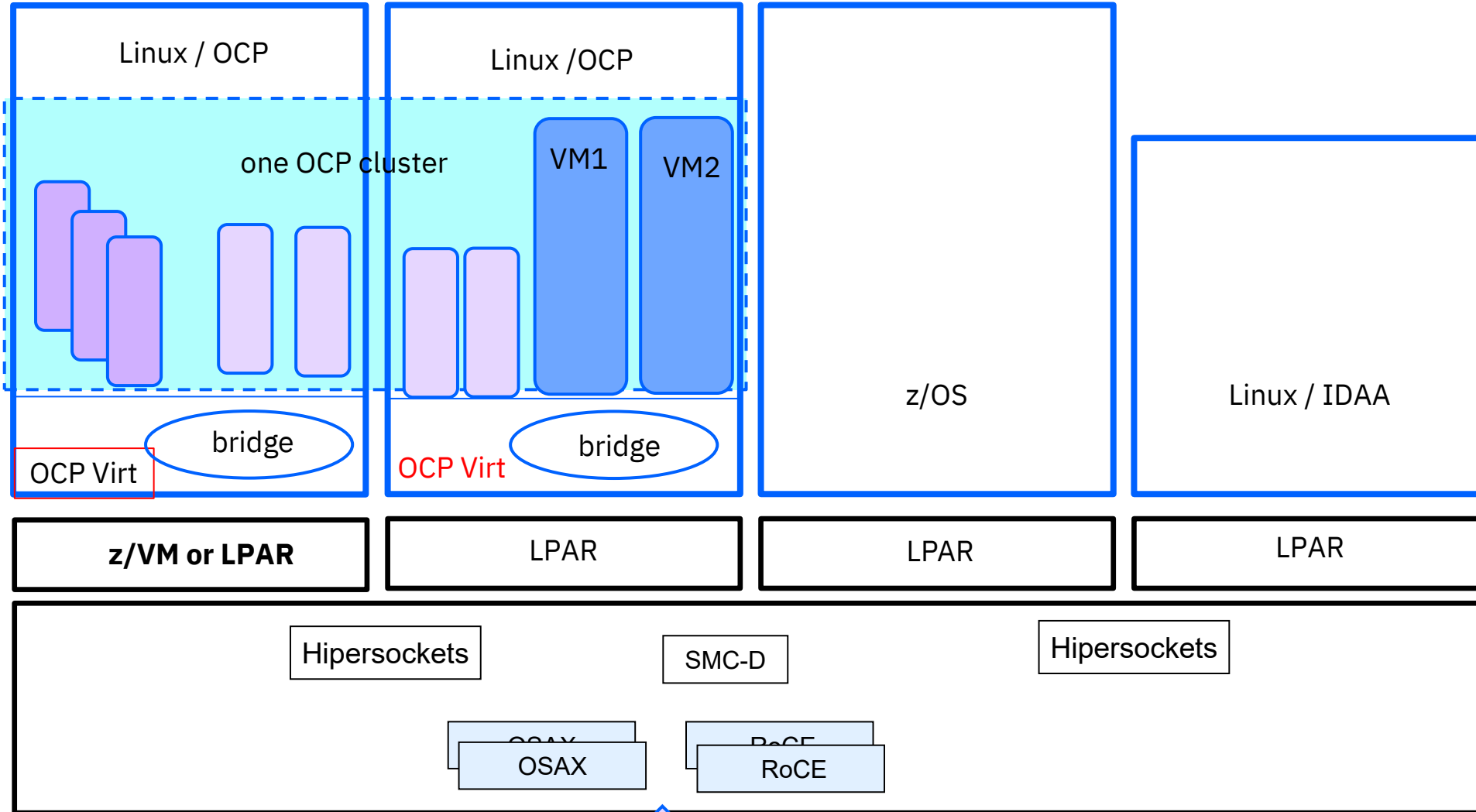
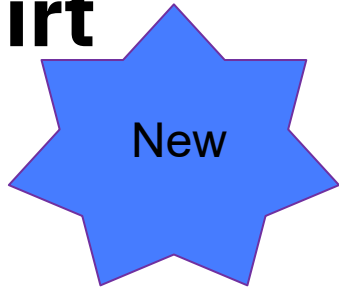
1+) Hybrid Cloud –network topology /w OpenShift Virt



Best Practices

- **OCP Virt is similar to KVM networks**
- **Multiple Networks**
- **MTU size mixed**
- **Hipersockets**
 - ok if possible
- **OSA for z/VM (VS)**
- **OSA for KVM**
 - or **SMC-D**
- **RoCE SMC-D**
 - if possible

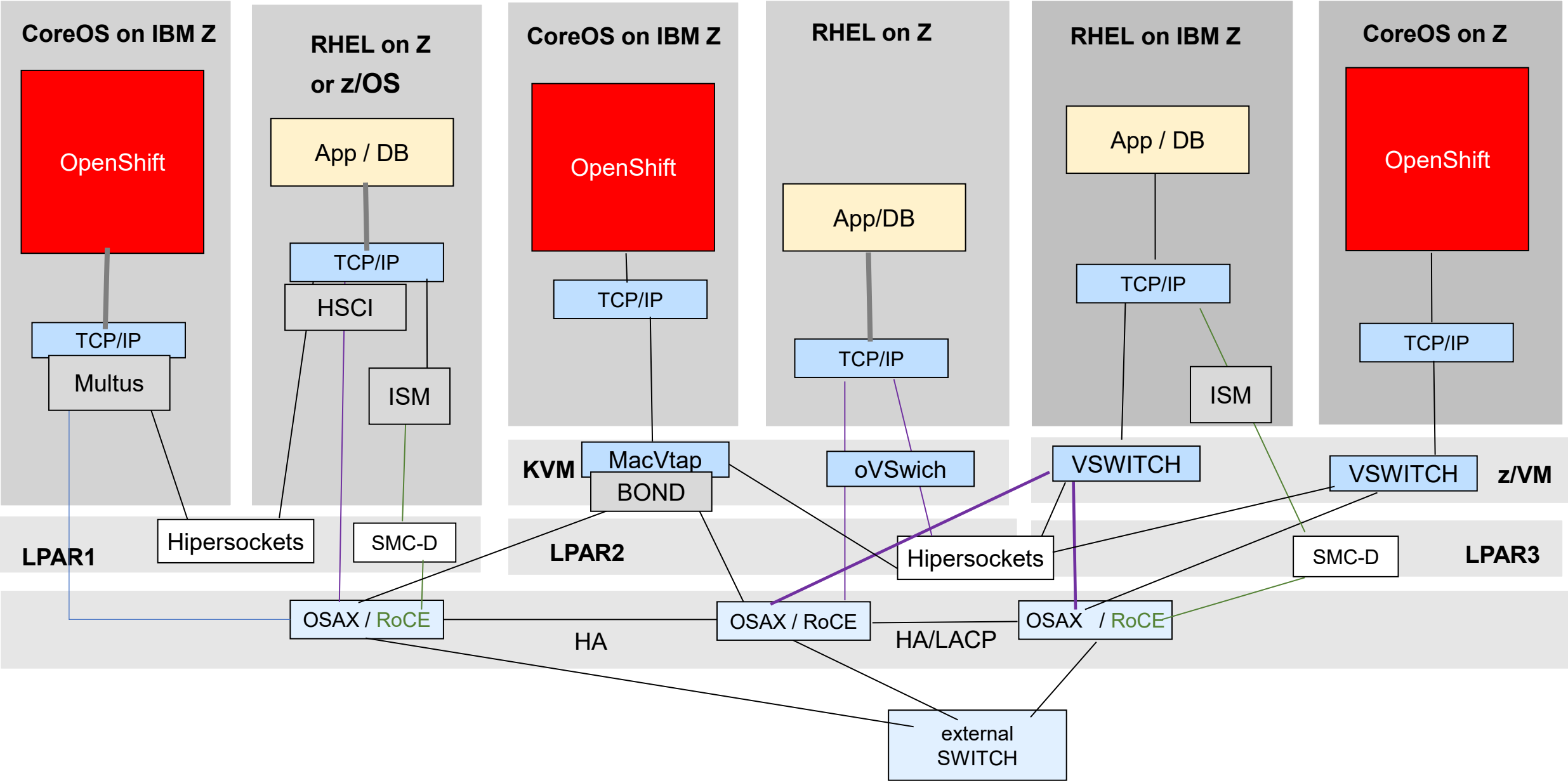
1+) Hybrid Cloud –network topology /w OpenShift Virt



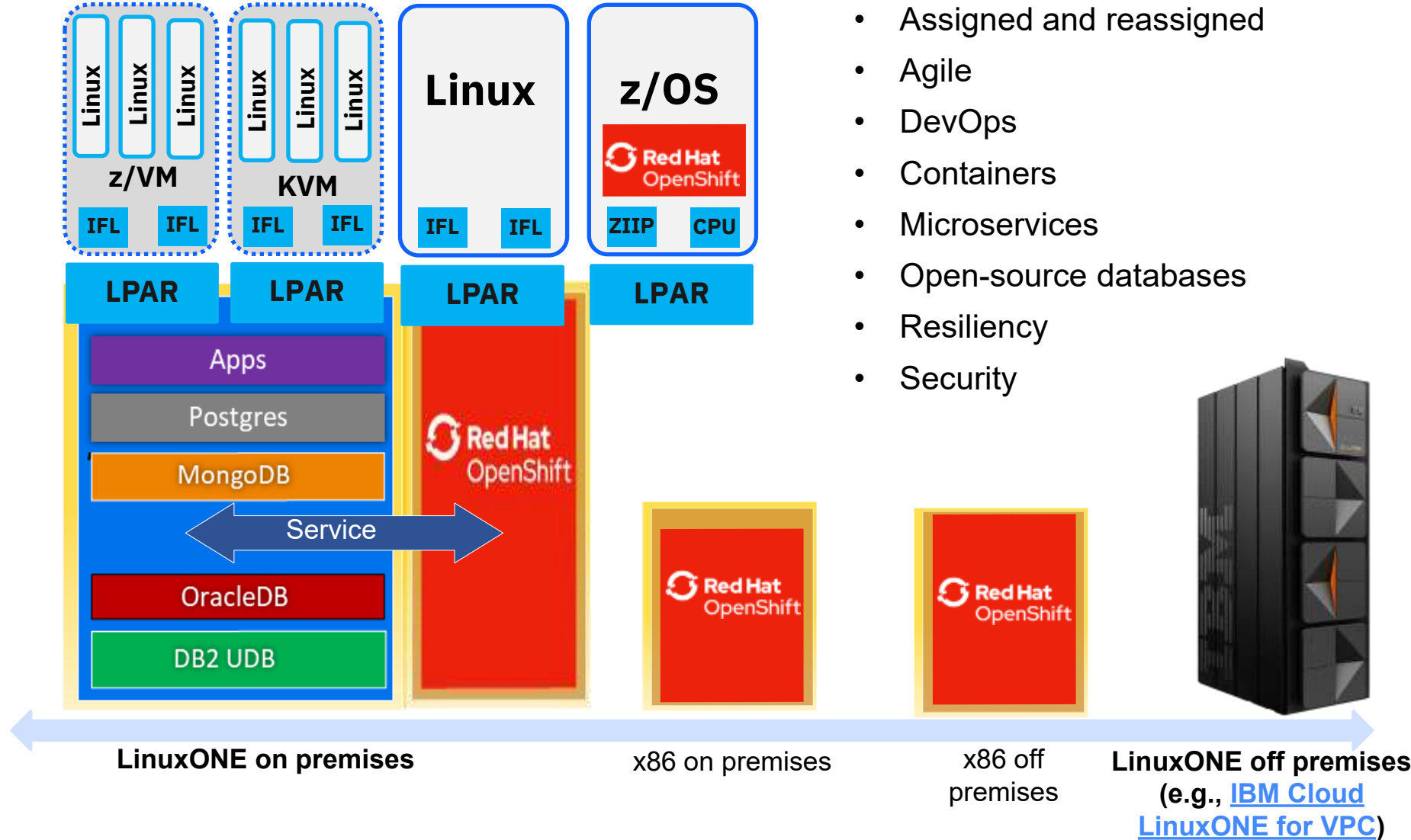
Best Practices

- OCP Virt is similar to KVM networks
- Multiple Networks
- MTU size mixed
- Hipersockets
 - ok if possible
- OSA for z/VM (VS)
- OSA for KVM
 - or SMC-D
- RoCE SMC-D
 - if possible

Network options overview on IBM Z and LinuxONE



IBM Z in hybrid experience



- On-demand
- Elasticity
- IT resources dynamically
- Assigned and reassigned
- Agile
- DevOps
- Containers
- Microservices
- Open-source databases
- Resiliency
- Security



Experience more with IBM

[Visit us at the IBM Booth #113](#)

After a full day of technical sessions, take a break with us!

Connect with our experts, snap a photo with the z17 Plexi or the latest Telum II, and get an up-close look at our Spyre Accelerator.

Come back each day for fresh topics and demos at our expert stations.



Think 2026

Join 5000+ senior business and technology leaders who are seizing the AI revolution to unlock unprecedented growth and productivity at **Think 2026**.

Find out more information using the QR code below.



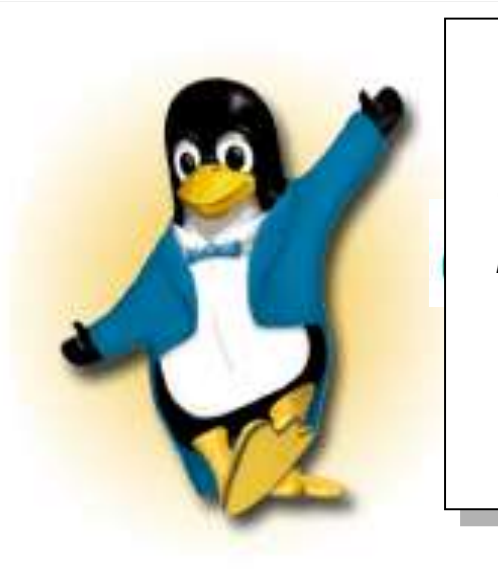
IBM Digital Asset Haven

IBM Digital Asset Haven is the operational backbone for financial institutions and regulated enterprises entering the digital asset economy.

Find out more information using the QR code below.



Questions?



Wilhelm Mild
IBM Executive IT Architect



*IBM Deutschland Research
 & Development GmbH
 IBM Campus 1
 71139 Ehningen
 Germany*

wilhelm.mild@de.ibm.com

Notices and disclaimers

- © 2019 International Business Machines Corporation. No part of this document may be reproduced or transmitted in any form without written permission from IBM.
- **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 continued

- 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