z/VM – SSI & LGR Usage Scenarios

Bill Bitner
z/VM Client Focus and Care
bitnerb@us.ibm.com

For the most current version of this presentation, please see
https://www.vm.ibm.com/library/presentations/

Abstract

z/VM Single System Image (SSI) clusters and Live Guest Relocation (LGR) became available in z/VM 6.2 as a priced feature, and later was included in the base of z/VM 7.1 for free. Most people appreciate the use case of SSI and LGR for addressing planned outages such as apply service for z/VM control program, but there are other scenarios and configurations where they can be used. This presentation takes a look at ten total cases where SSI and LGR can be used.

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

BladeCenter*	GDPS*	IBM z13*	PR/SM	System z9*	zSecure
DB2*	HiperSockets	IBM z14	RACF*	System z10*	z/VM*
DS6000*	HyperSwap	OMEGAMON*	Storwize*	Tivoli*	z Systems*
DS8000*	IBM LinuxONE Emperor	Performance Toolkit for VM	System Storage*	zEnterprise*	
ECKD	IBM LinuxONE Rockhopper	Power*	System x*	z/OS*	
FICON*	IBM Z*	PowerVM	System z*		

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency which is now part of the Office of Government Commerce.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

Java and all Java based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

OpenStack is a trademark of OpenStack LLC. The OpenStack trademark policy is available on the OpenStack website.

TEALEAF is a registered trademark of Tealeaf, an IBM Company.

Windows Server and the Windows logo are trademarks of the Microsoft group of countries.

Worklight is a trademark or registered trademark of Worklight, an IBM Company.

UNIX is a registered trademark of The Open Group in the United States and other countries.

* Other product and service names might be trademarks of IBM or other companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the 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.
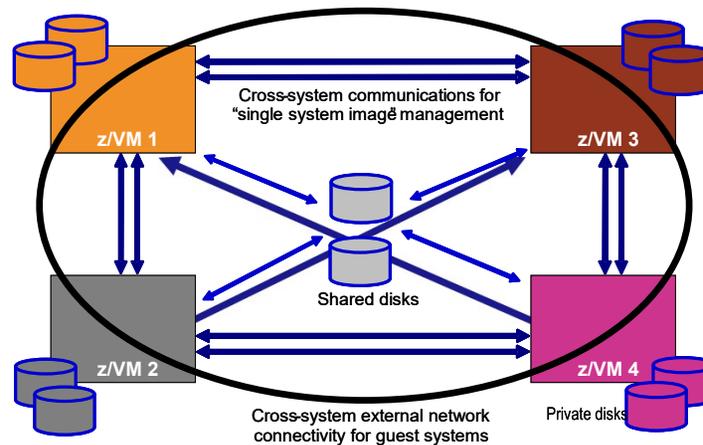
Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

This information provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g., zIIPs, zAAPs, and IFLs) ("SEs"). IBM authorizes customers to use IBM SE only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the "Authorized Use Table for IBM Machines" provided at www.ibm.com/systems/support/machine_warranties/machine_code/aut.html ("AUT"). No other workload processing is authorized for execution on an SE. IBM offers SE at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.

Single System Image (SSI) Feature
Clustered Hypervisor with Live Guest Relocation

- Included in base of z/VM 7.1 (Previously an optional priced feature)
- Connect up to four z/VM systems as members of a Single System Image cluster
- Cluster members can be run on the same or different IBM Z or LinuxONE servers
- Simplifies management of a multi-z/VM environment
  - Single user directory
  - Cluster management from any member
    - Apply maintenance to all members in the cluster from one location
    - Issue commands from one member to operate on another
  - Built-in cross-member capabilities
  - Resource coordination and protection of network and disks
- Allows Live Guest Relocation of running Linux guests

z/VM 1

Cross-system communications for "single system image" management

z/VM 3

Shared disks

z/VM 2

Cross-system external network connectivity for guest systems

z/VM 4

Private disks

What Can You Do with SSI Clusters and LGR?

1. Flexibility for Planned Outages

2. Methodically Testing at Current Levels

3. Increased Control Over Server Sprawl

4. Production with Protection

5. Managing Resource Distribution

6. Consistent Test Bed for Stress Tests

7. One From the Customers – Utility Migration LPAR

8. Local Disaster Recover (Business Continuity)

9. Come Closer for Performance

10. Migrate to New Processor

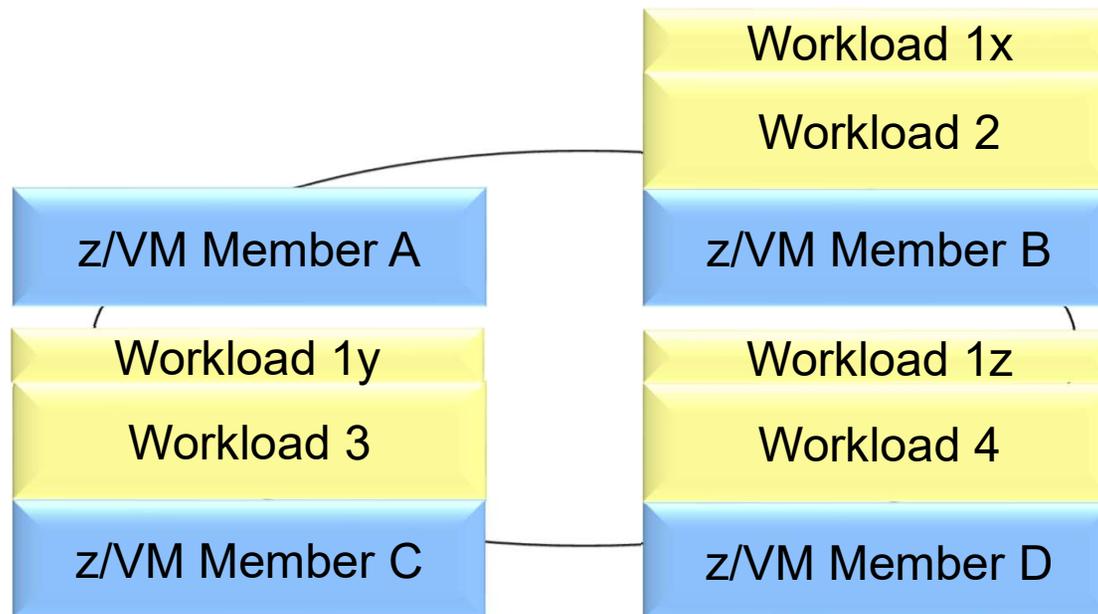# Flexibility for Planned Outages

Flexibility for Planned Outages

- The good news is workload running on z/VM is becoming more and more critical; the bad news is that brings greater availability challenges.
    - Maintenance windows for down time get smaller

- SSI and LGR allow moving work and rolling out service...

Workload 1	Workload 2
z/VM Member A	z/VM Member B
Workload 3	Workload 4
z/VM Member C	z/VM Member D
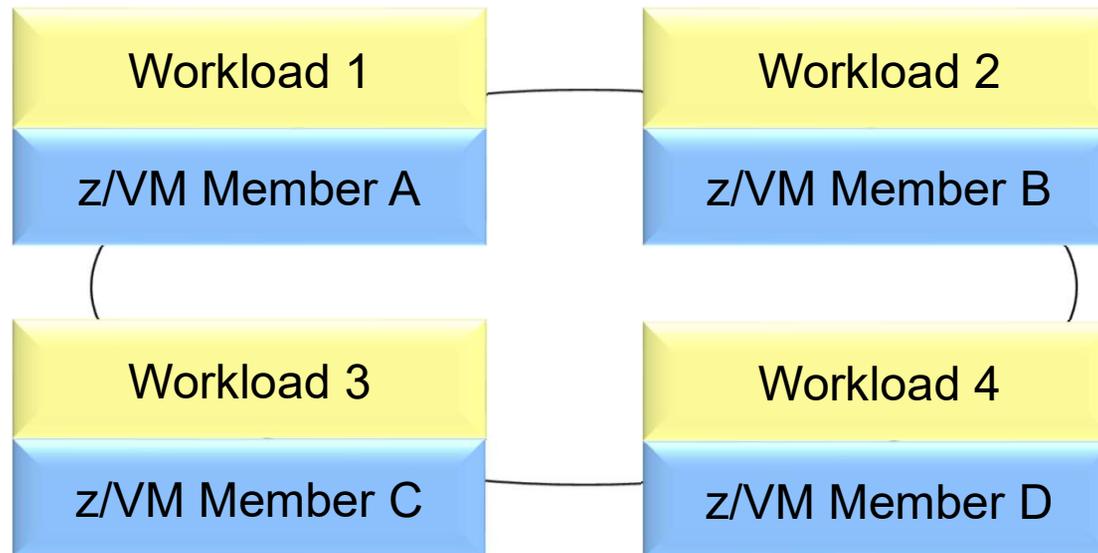
Flexibility for Planned Outages

1.Apply maintenance to Member A, having new CP load module ready for IPL.

2.Move critical work from Member A to the other 3 members in the cluster.

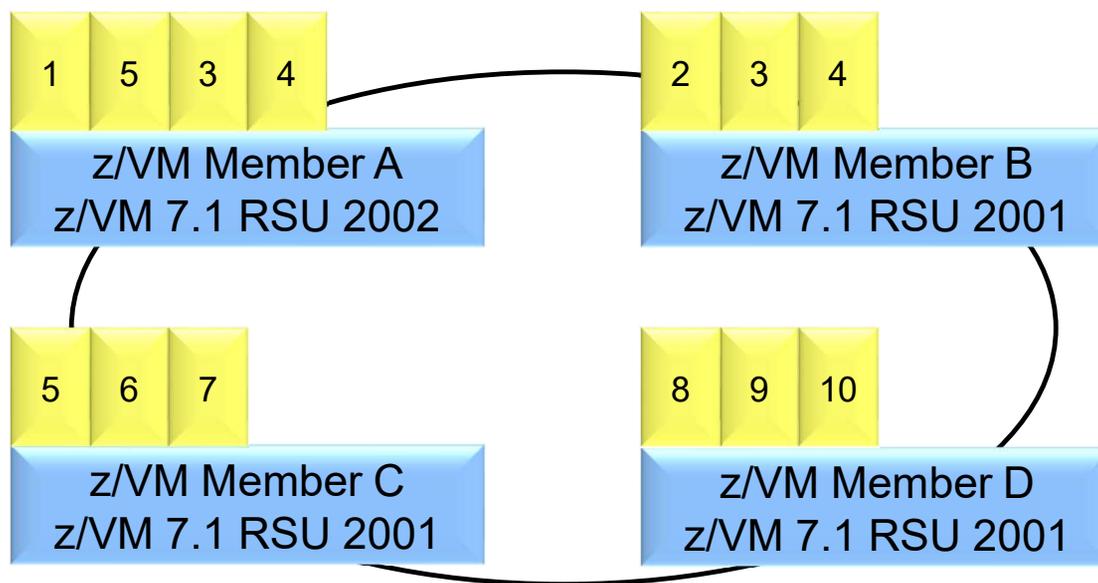3.Shutdown Member A and bring back up with new CP load module.

	Workload 1x
	Workload 2
z/VM Member A	z/VM Member B
Workload 1y	Workload 1z
Workload 3	Workload 4
z/VM Member C	z/VM Member D

Flexibility for Planned Outages

1. Move workloads back to member A

2. Repeat as appropriate for the other members.

Workload 1	Workload 2
z/VM Member A	z/VM Member B

Workload 3	Workload 4
z/VM Member C	z/VM Member D
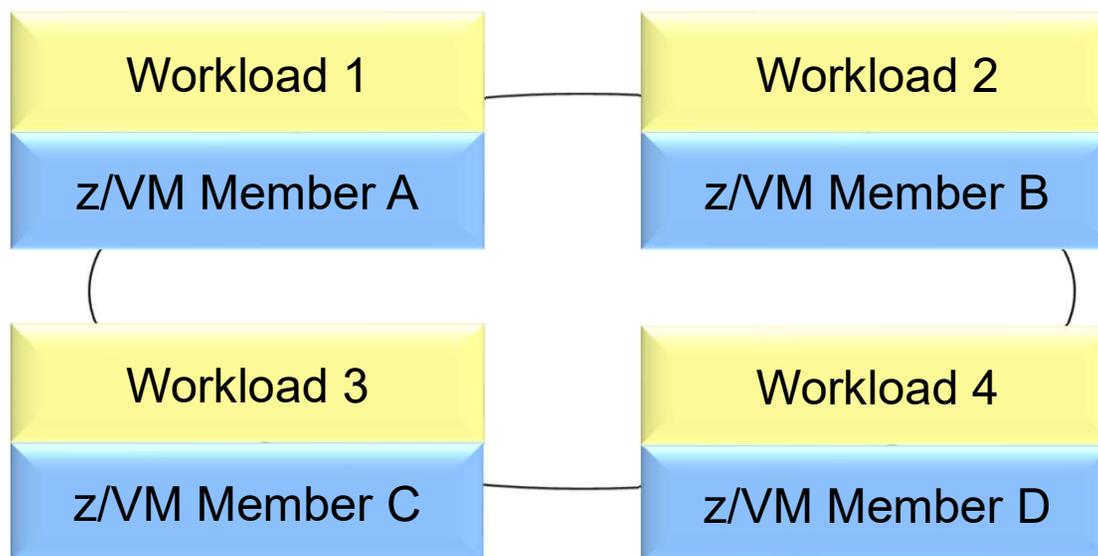
Methodically Testing at Current Levels

Methodically Testing at Current Levels

- Testing for new levels of z/VM in the past often required use of second level systems and trade-offs between matching production environment.

- z/VM SSI clusters can be used to help test and migrate throughout the members.

- Perhaps start with System A at new service level and slowly move work there to test.

1	5	3	4

z/VM Member A
z/VM 7.1 RSU 2002

2	3	4

z/VM Member B
z/VM 7.1 RSU 2001

5	6	7

z/VM Member C
z/VM 7.1 RSU 2001

8	9	10

z/VM Member D
z/VM 7.1 RSU 2001

Increased Control Over Server Sprawl

Increased Control Over Server Sprawl

- Server sprawl and the success of virtualization have led to virtual server sprawl, z/VM SSI Clusters improve the management characteristics for these environments.

- Consider customer with a single LPAR for production is sufficient today, but they are growing at a significant rate.

- Various reasons to expand past a single LPAR:
  – Out growing single LPAR capacity
  – Risk management: avoiding all eggs in one basket and diversification.
  – Flexibility for software licensing

- Keep you individual systems, but prepare them to run as multi-member SSI in the future.
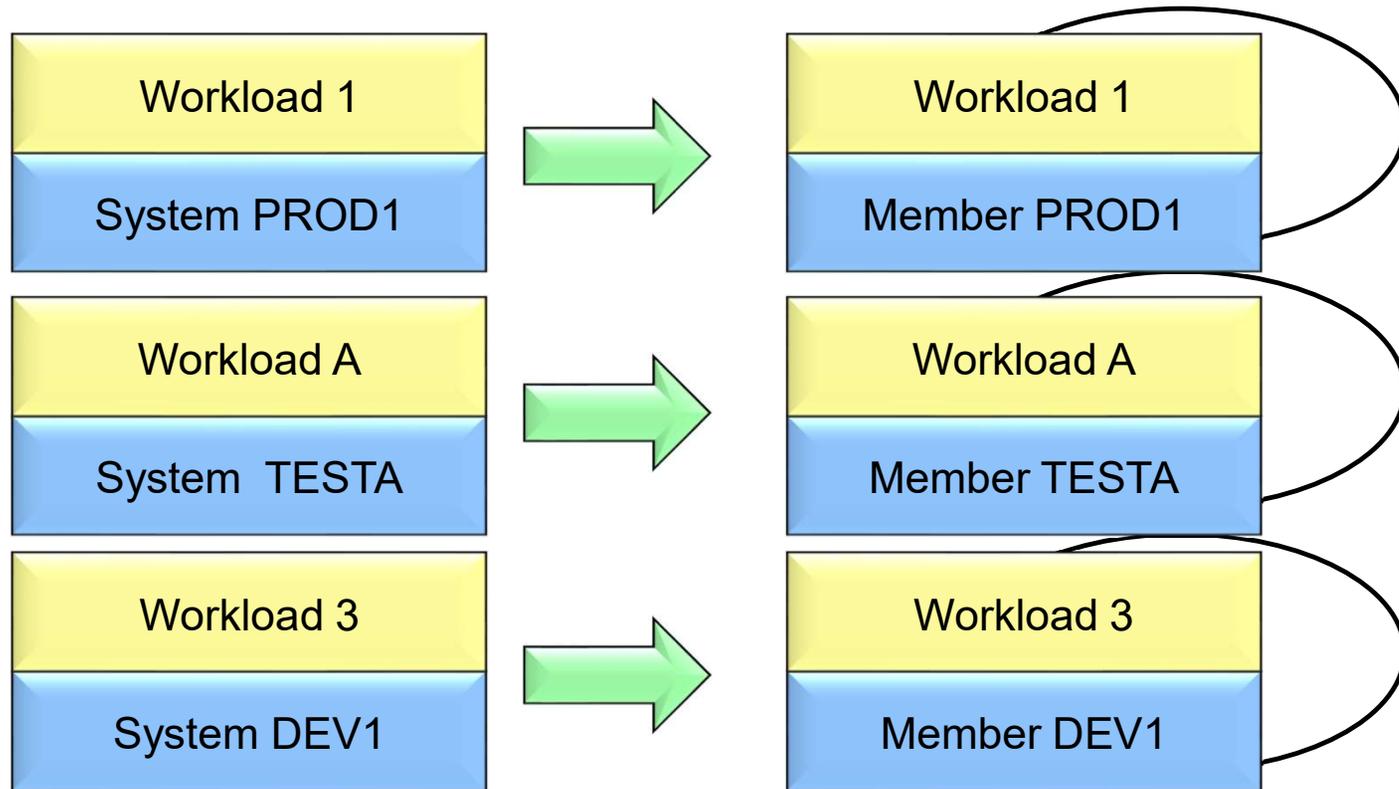 – Bring in another LPAR and bring up an additional SSI member.

Increased Control Over Server Sprawl

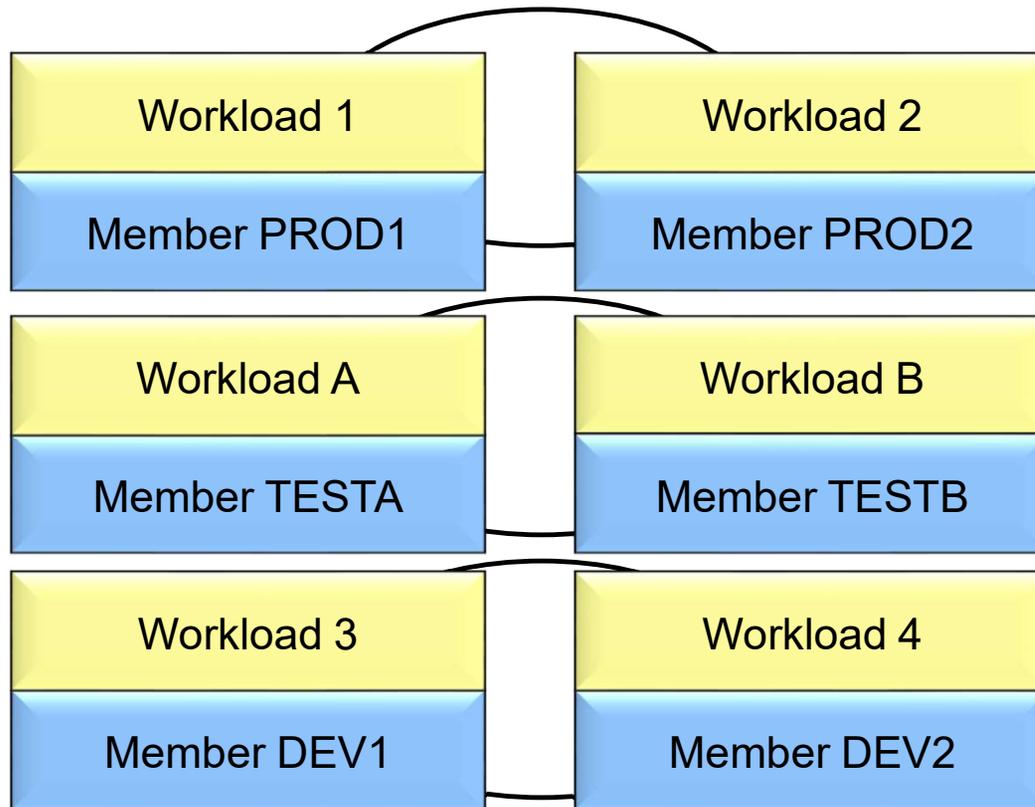Today, you may have 3 separate systems, but may not have compelling reason to combine them into a cluster.

Workload 1
System PROD1

Workload A
System TESTA

Workload 3
System DEV1
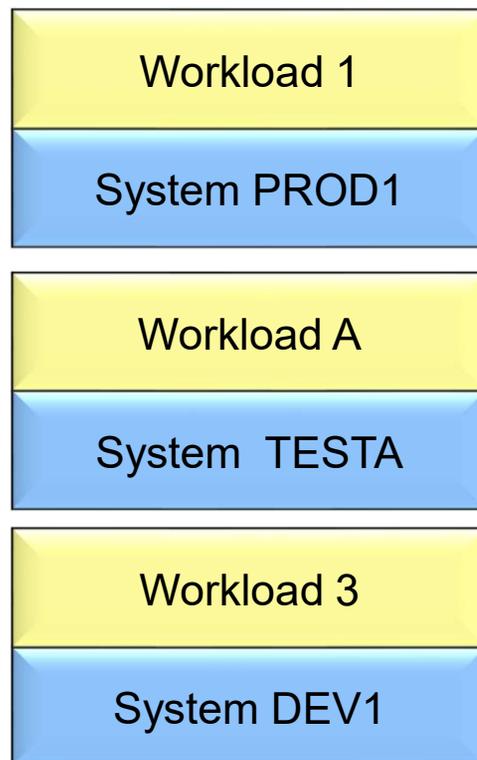
Increased Control Over Server Sprawl

Create z/VM SSI Clusters, but with a single member for now.

Workload 1	→	Workload 1
System PROD1		Member PROD1
Workload A	→	Workload A
System TESTA		Member TESTA
Workload 3	→	Workload 3
System DEV1		Member DEV1

Increased Control Over Server Sprawl

As workloads increase, create additional members in each cluster.

Workload 1	Workload 2
Member PROD1	Member PROD2
Workload A	Workload B
Member TESTA	Member TESTB
Workload 3	Workload 4
Member DEV1	Member DEV2

Production with Protection

Production with Protection

- When adding a new application or upgrading an application in production, what is your confidence that you know how it will
  - Perform?
  - Impact other production workload?
  - Meet expectations?

- Single System Image provides a way to allow workload to be part of the production environment, and yet be isolated
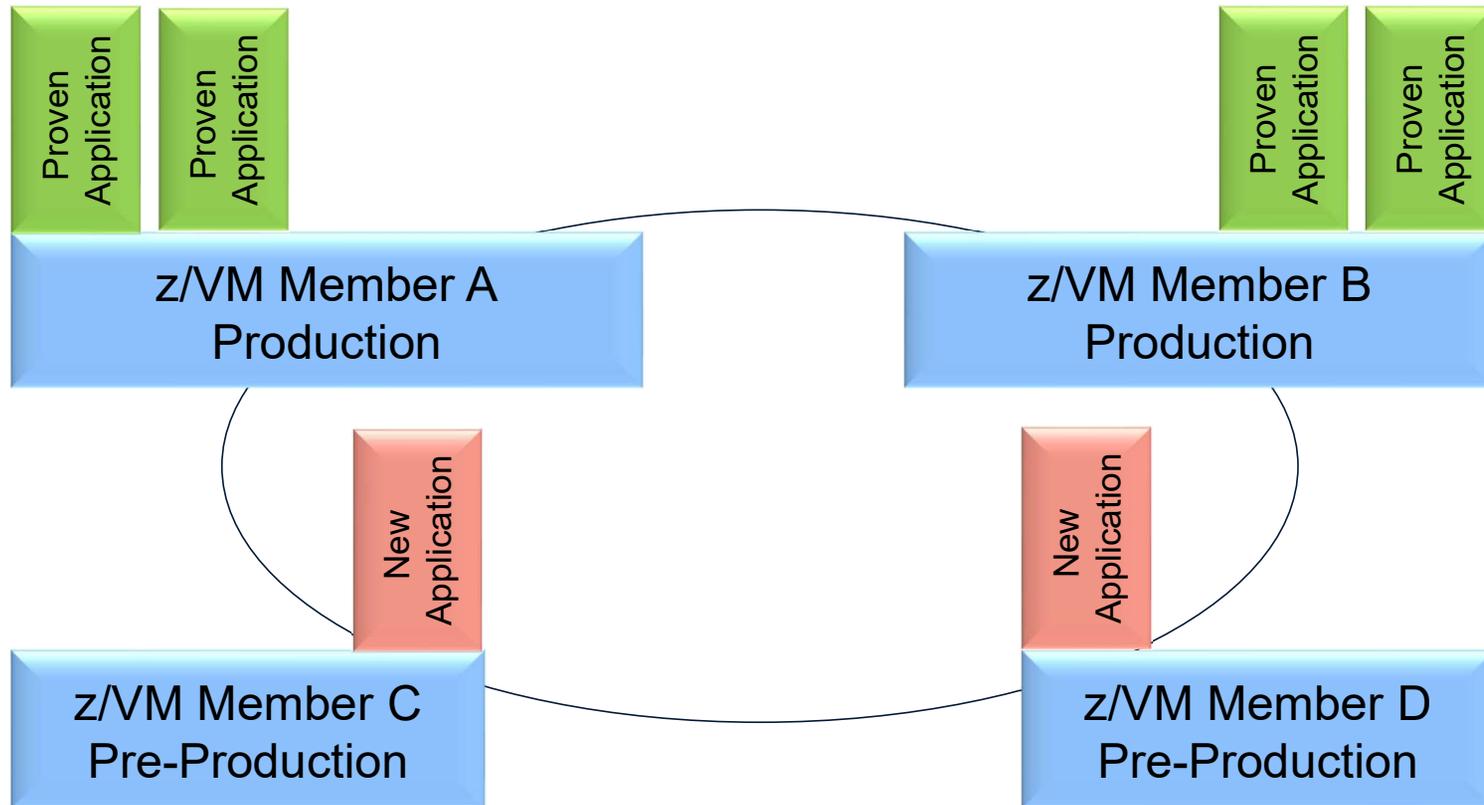
Production with Protection

- Four Members
  - True Production – two for redundancy
    - Full amount of resources.
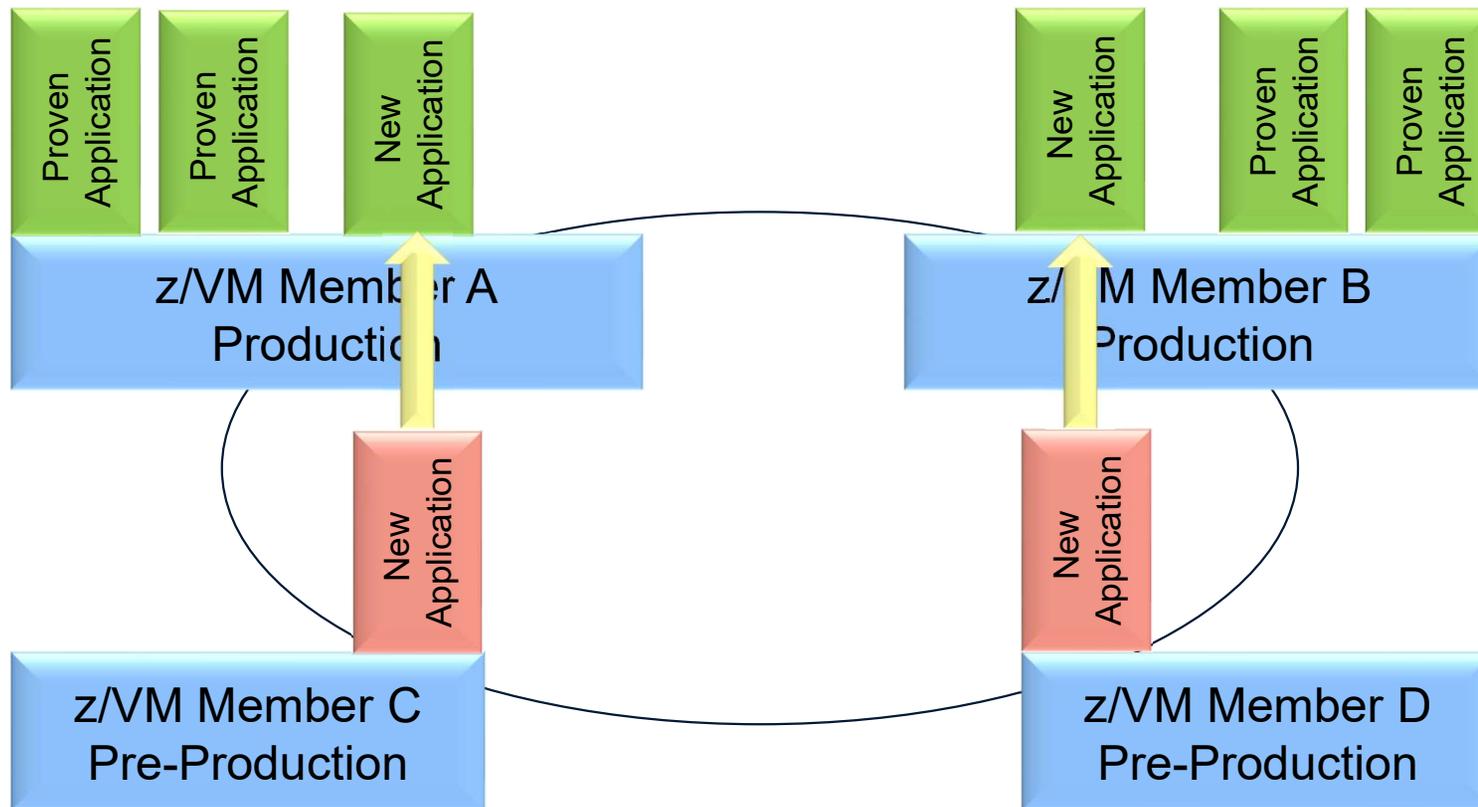 - Pre-Production: proving grounds
    - Limited resources.

```
┌────────────────────┐          ┌────────────────────┐
│  z/VM Member A     │          │  z/VM Member B     │
│  Production        │          │  Production        │
└────────────────────┘          └────────────────────┘

┌────────────────────┐          ┌────────────────────┐
│  z/VM Member C     │          │  z/VM Member D     │
│  Pre-Production    │          │  Pre-Production    │
└────────────────────┘          └────────────────────┘
```
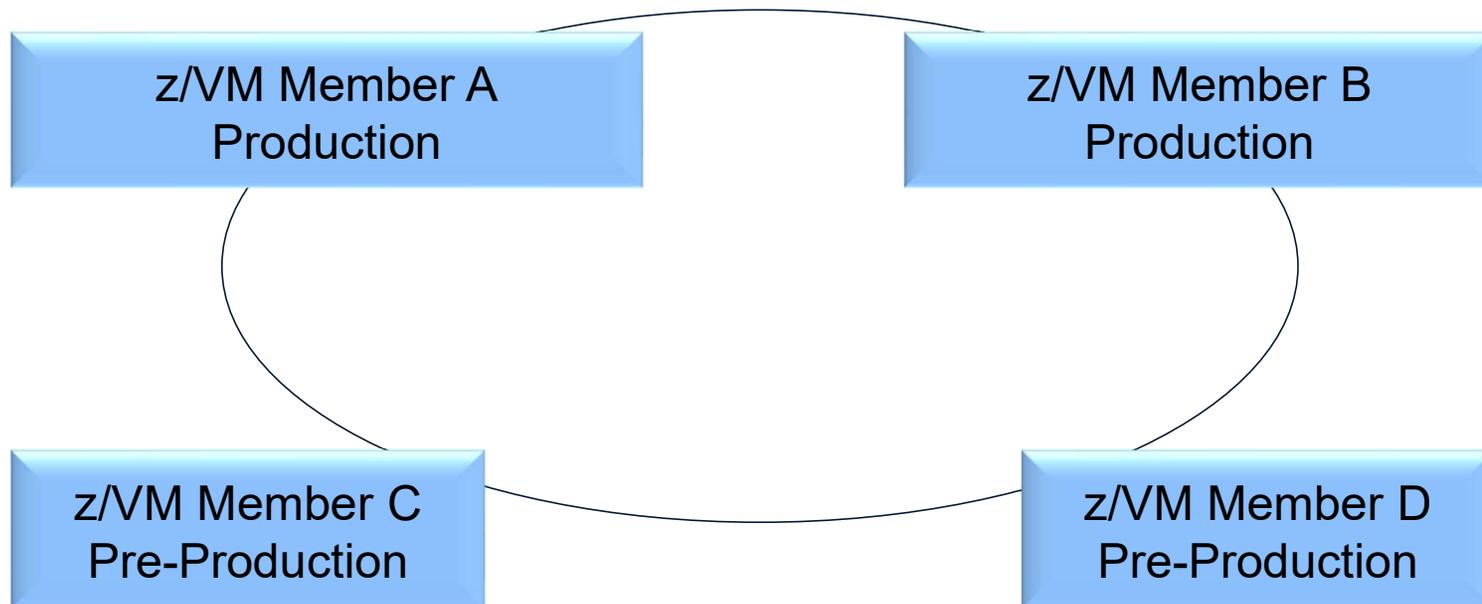
Production with Protection

- Allow new application to run in pre-production LPARs

Production with Protection

- If all goes well, move into true production

Production with Protection

- If all goes well, move into true production
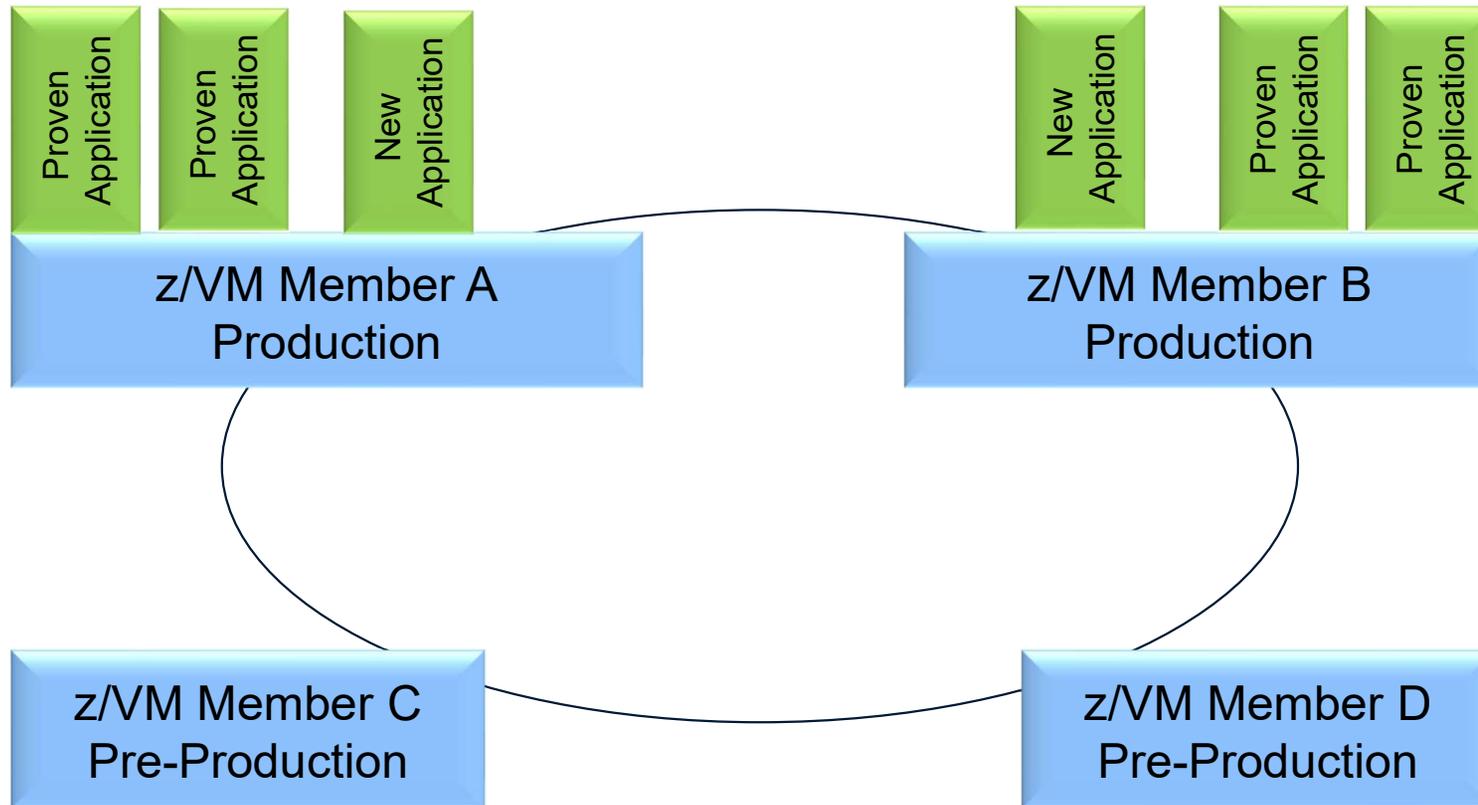
Managing Resource Distribution

Managing Resource Distribution

- Some customers have or are in processing of exceeding the capacity of a single z/VM system and split work across LPARs

- Determining how to divide the workloads across LPARs is a challenge, particularly in a dynamic world

z/VM System A   z/VM System B
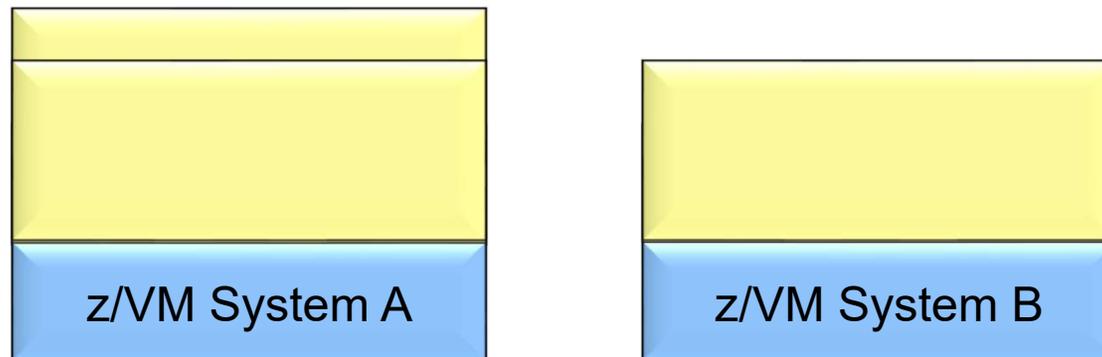
- With individual z/VM systems, one would need to define new virtual machines on B and remove the definitions on A

- Responsibility of ensuring integrity during process is on shoulders of system programmer.

- With an SSI cluster, one can more easily redistribute the load through logoff/logon or with LGR.
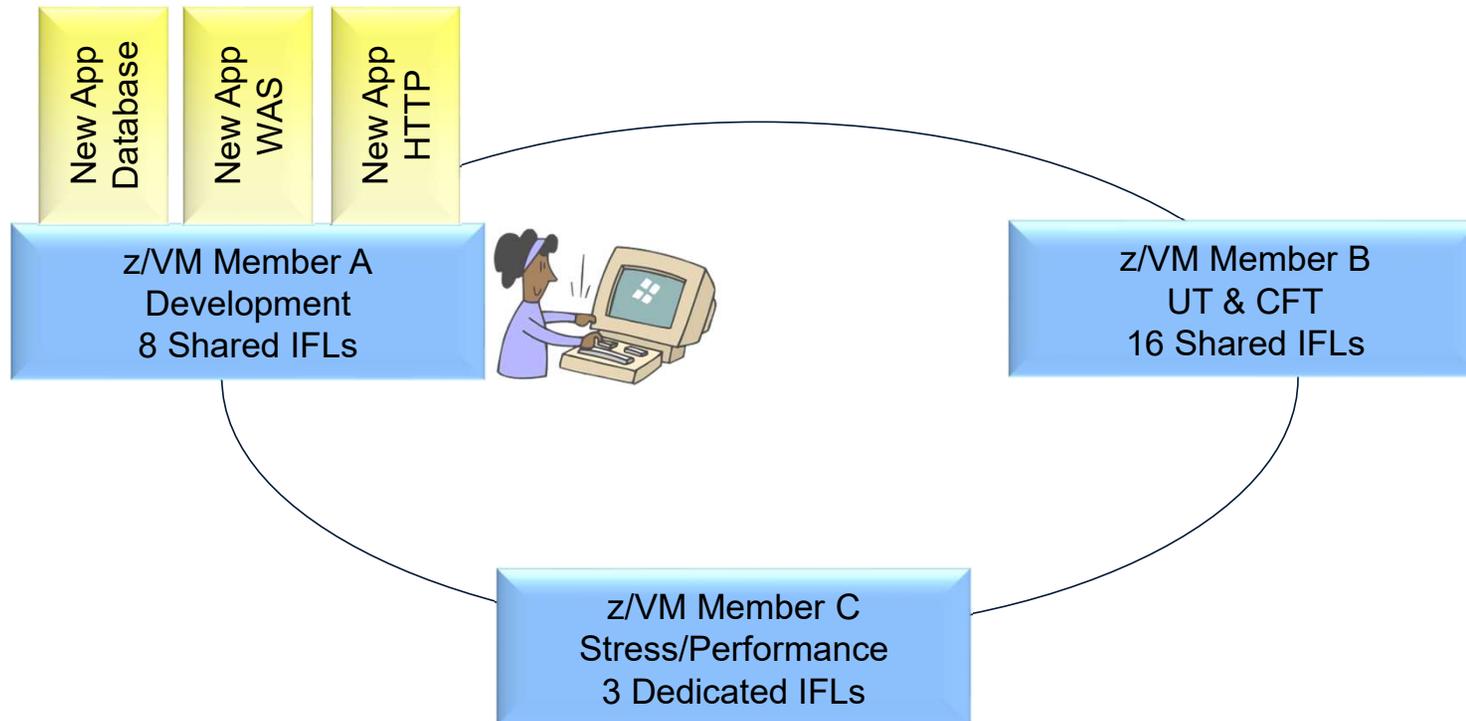
Consistent Test Bed for Stress Tests

Consistent Test Bed for Stress Tests

- Testing Challenges:
  - Controlling test environments, testing in **consistent** manner
  - Functional and QA testing of various test programs
  - Stress testing in a controlled environment

- Having an SSI cluster environment allows:
  - Virtual server with same resources, run in different members of cluster based on needs
  - Load in development probably not as heavy, run that in a smaller shared environment
  - Various testing in UT & CFT could create a heavier load for various testing
  - An isolated LPAR (member) for stress testing or establishing performance characteristics of workload.
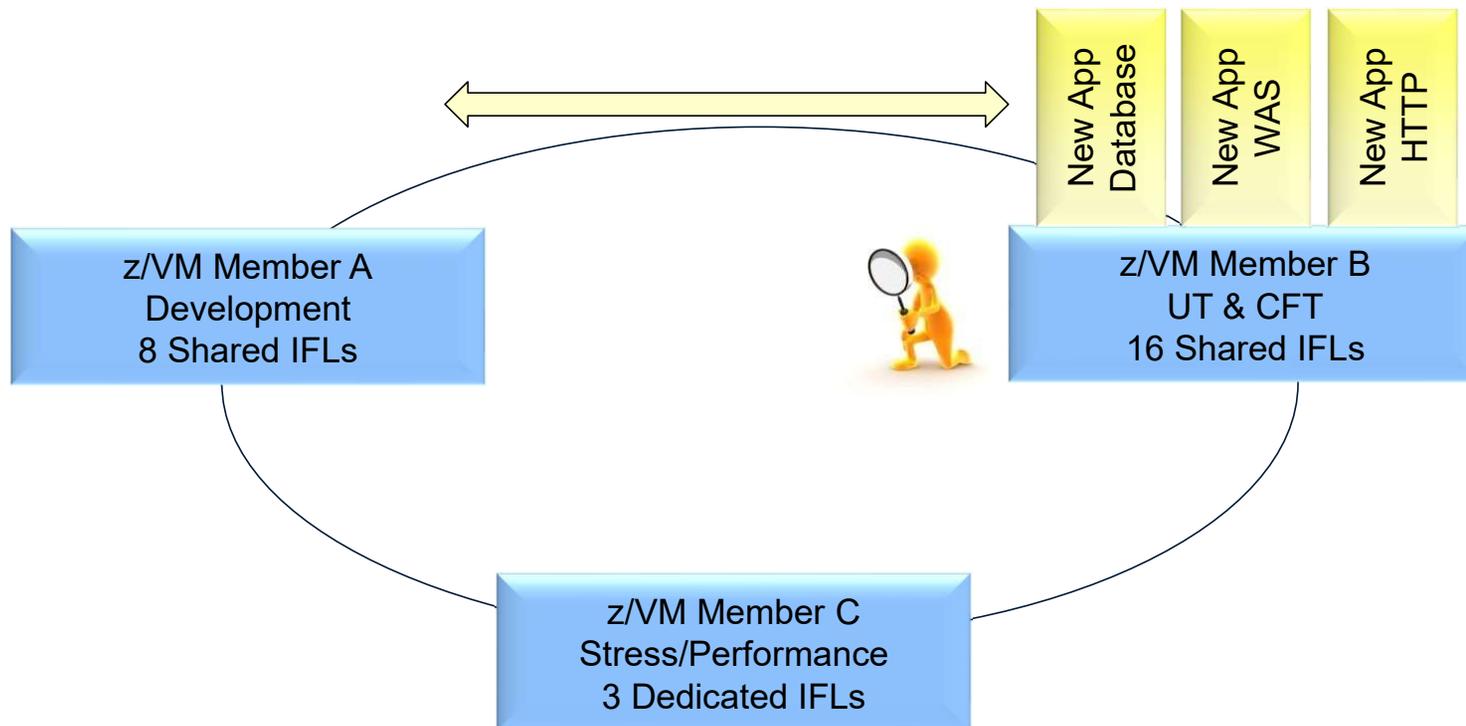
Consistent Test Bed for Stress Tests

- Consider this example with development, unit test, component function test, performance test, and stress tests.
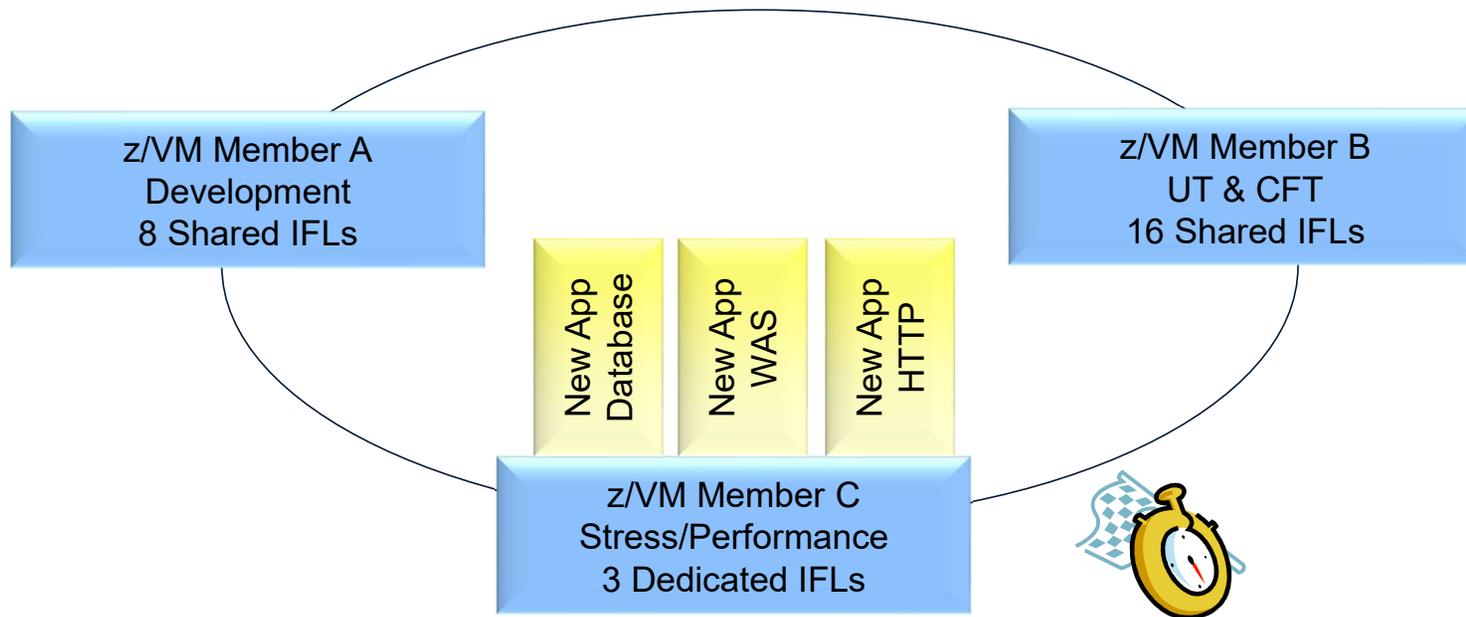- Build it all in the development member.

Consistent Test Bed for Stress Tests

- Development and Test could share the virtual machines involved, passing them back and forth between the systems as needed.

New App Database

New App WAS

New App HTTP

z/VM Member A
Development
8 Shared IFLs

z/VM Member B
UT & CFT
16 Shared IFLs

z/VM Member C
Stress/Performance
3 Dedicated IFLs

Consistent Test Bed for Stress Tests

- When ready for performance or stress test, move to Member C with Dedicated resources
- More control over what has changed

z/VM Member A
Development
8 Shared IFLs

z/VM Member B
UT & CFT
16 Shared IFLs

New App
Database

New App
WAS

New App
HTTP

z/VM Member C
Stress/Performance
3 Dedicated IFLs

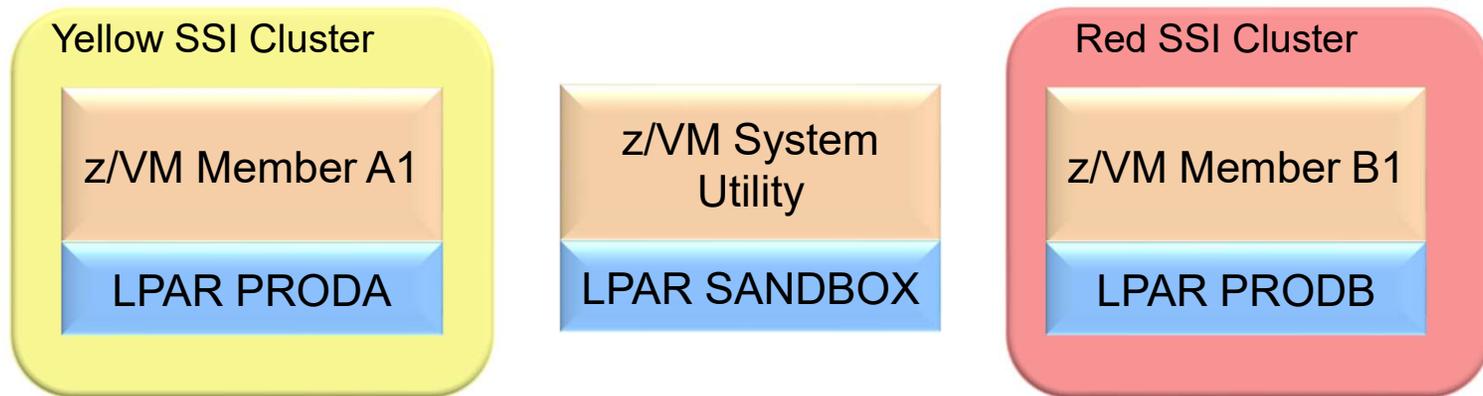# One From the Customers – Utility Migration LPAR

One From the Customers – Utility Migration LPAR

- Customer with three discrete LPARS
  - Two Production z/VM Systems
  - One System Programmer Sandbox or Utility system

- Wants benefit of updating z/VM control program without an application outage.

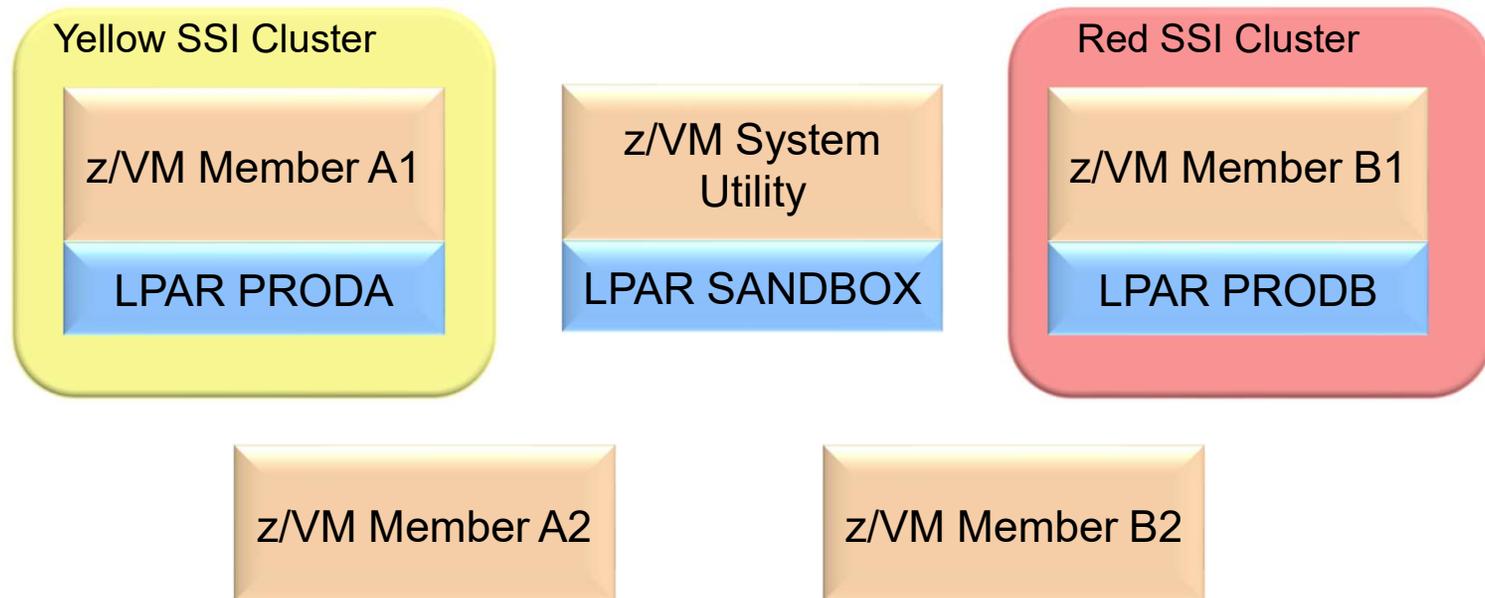| z/VM System A	z/VM System Utility	z/VM System B
LPAR PRODA	LPAR SANDBOX	LPAR PRODB
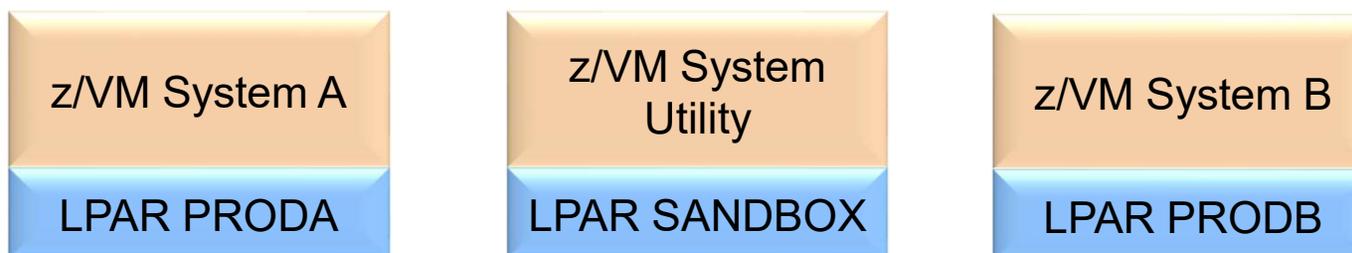
One From the Customers – Utility Migration LPAR

- Create SSI Cluster for each production System
  - Two Two-Member Clusters
  - But only include one of the production LPARs in each

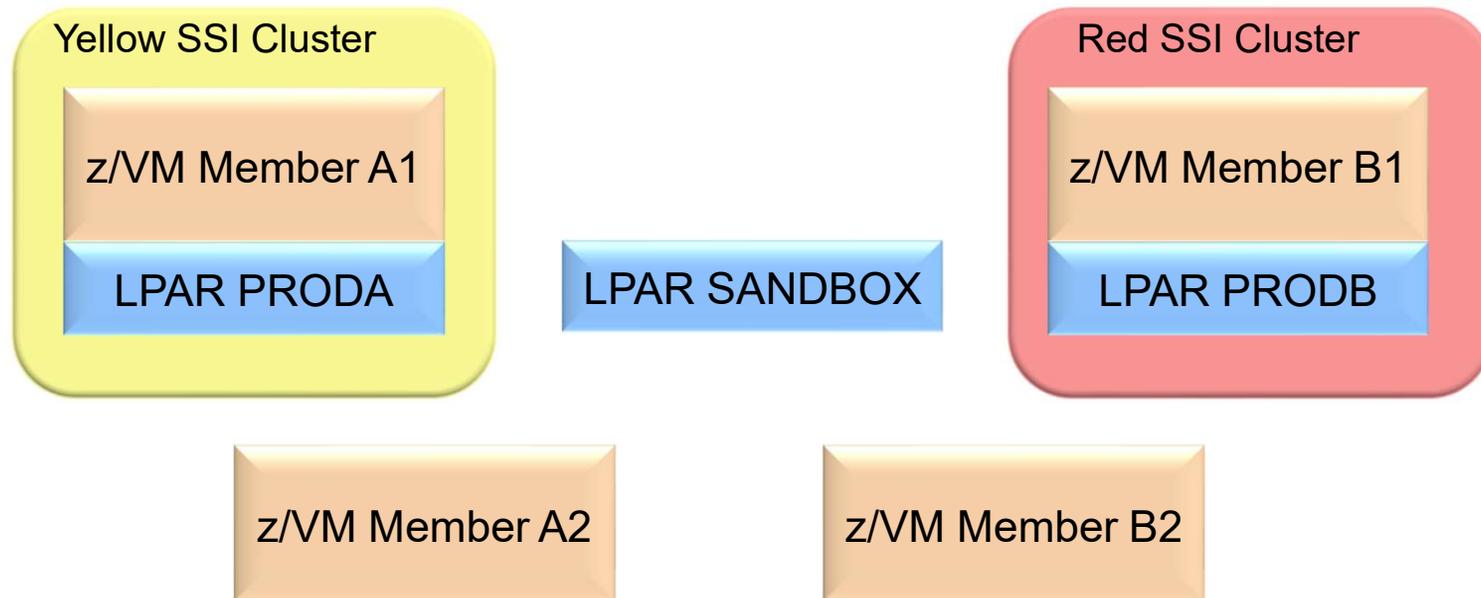- Utility System can stay a singleton or even a non-SSI system

Yellow SSI Cluster	z/VM System Utility	Red SSI Cluster
z/VM Member A1		z/VM Member B1
LPAR PRODA	LPAR SANDBOX	LPAR PRODB

One From the Customers – Utility Migration LPAR

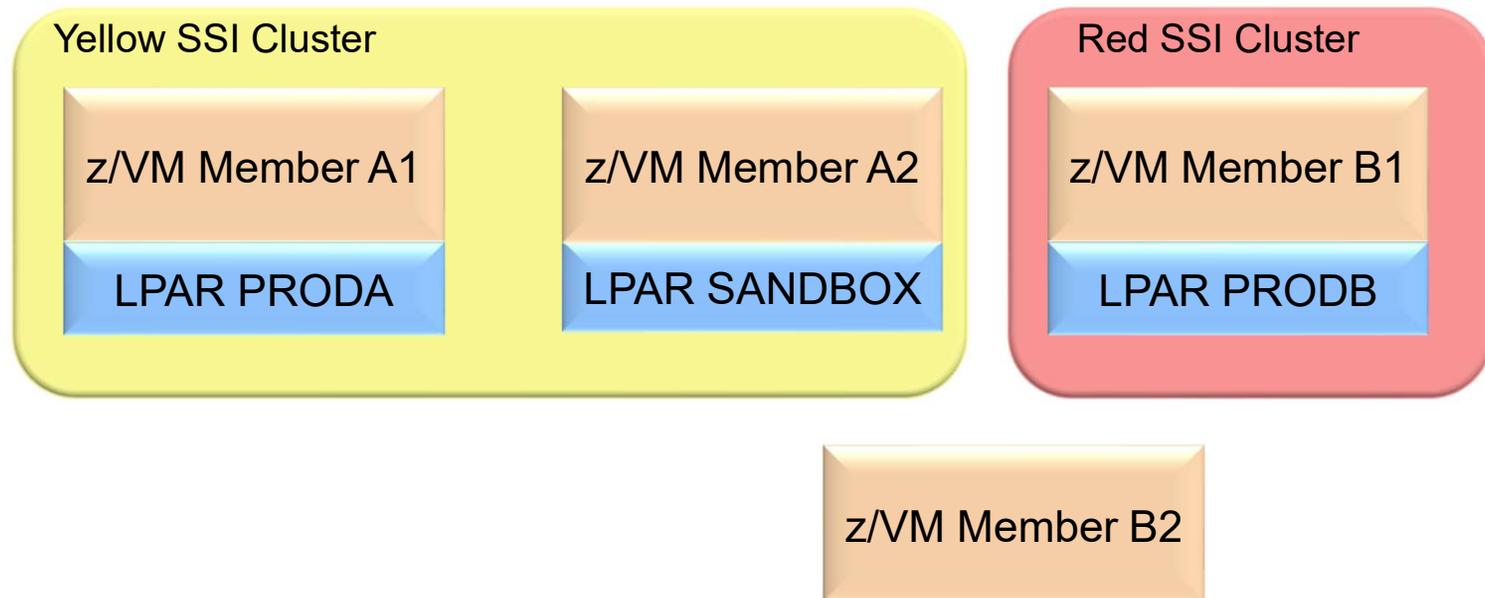▪ Clone the production members so there is a second system (member) for each of the production LPARs.

Yellow SSI Cluster	z/VM System Utility	Red SSI Cluster
z/VM Member A1		z/VM Member B1
LPAR PRODA	LPAR SANDBOX	LPAR PRODB

z/VM Member A2

z/VM Member B2

One From the Customers – Utility Migration LPAR

- To update CP on production LPAR PRODA
  Step 1: Shutdown Utility System

Yellow SSI Cluster

z/VM Member A1

LPAR PRODA

LPAR SANDBOX

Red SSI Cluster

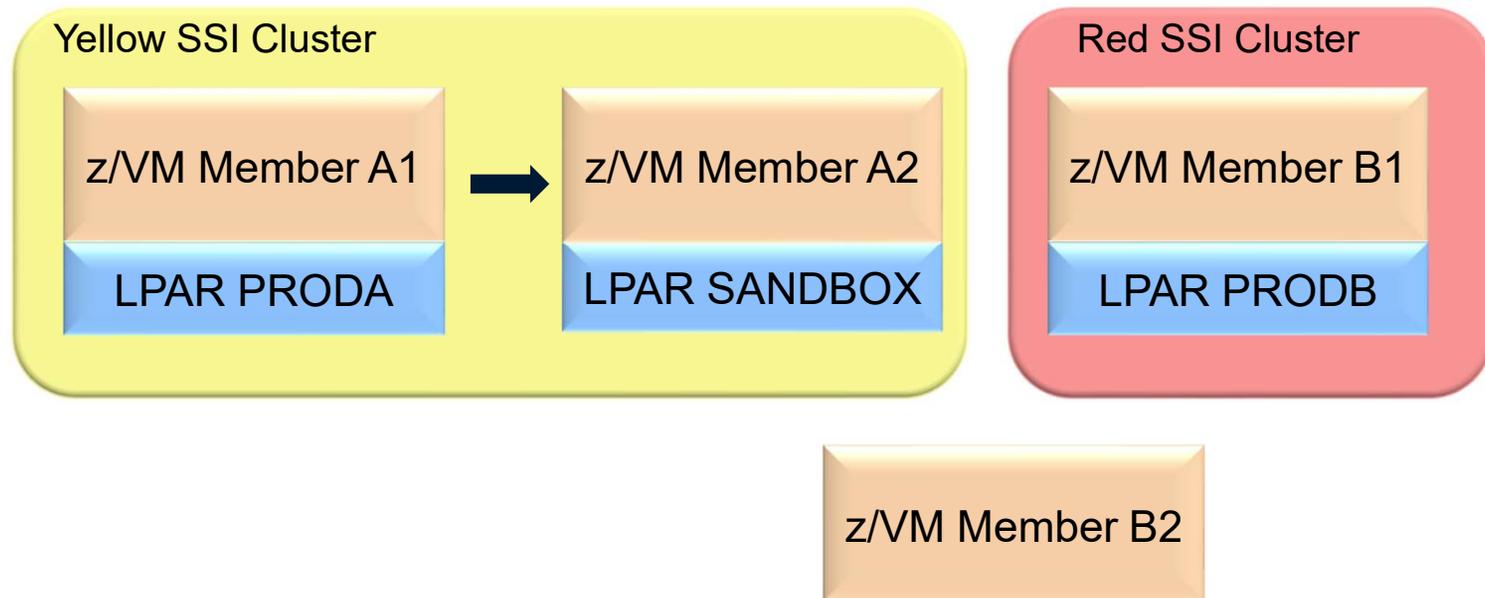z/VM Member B1

LPAR PRODB

z/VM Member A2

z/VM Member B2

One From the Customers – Utility Migration LPAR

- To update CP on production LPAR PRODA
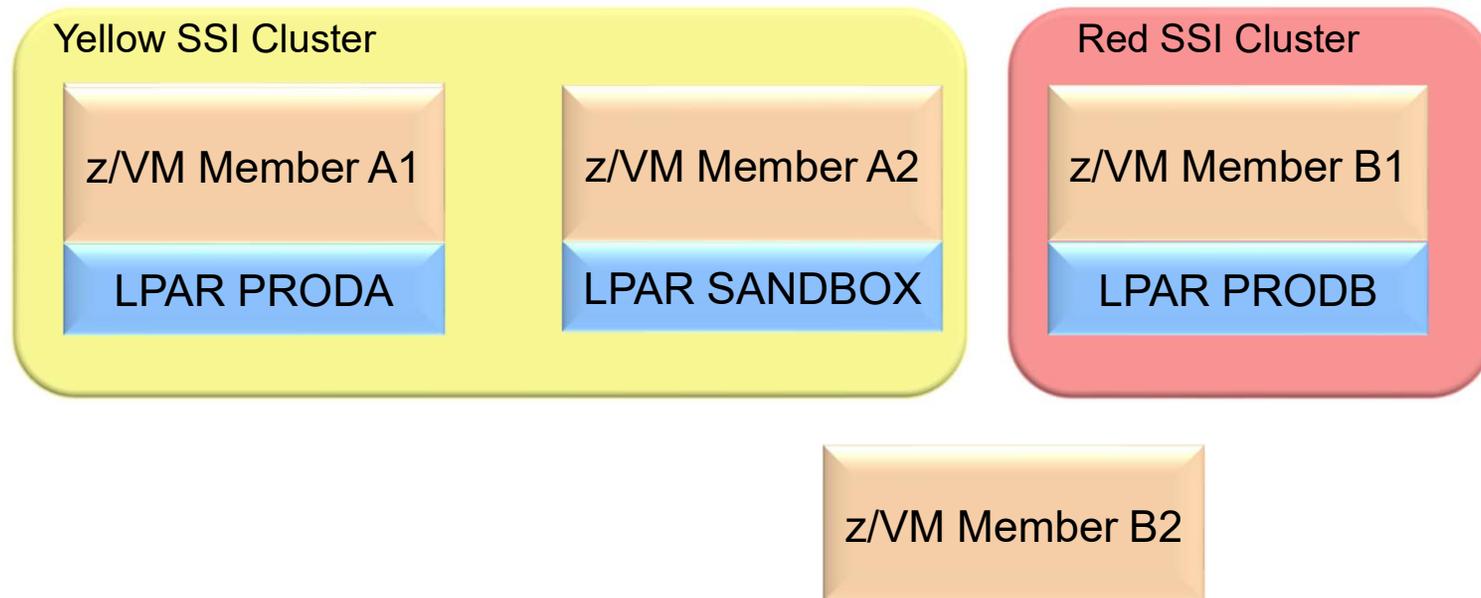 - Step 2: Bring up the other Member in SANDBOX LPAR

Yellow SSI Cluster		Red SSI Cluster
z/VM Member A1	z/VM Member A2	z/VM Member B1
LPAR PRODA	LPAR SANDBOX	LPAR PRODB

z/VM Member B2

One From the Customers – Utility Migration LPAR

- To update CP on production LPAR PRODA
  - Step 3: Move work from A1 to A2

Yellow SSI Cluster

z/VM Member A1 → z/VM Member A2

LPAR PRODA

LPAR SANDBOX

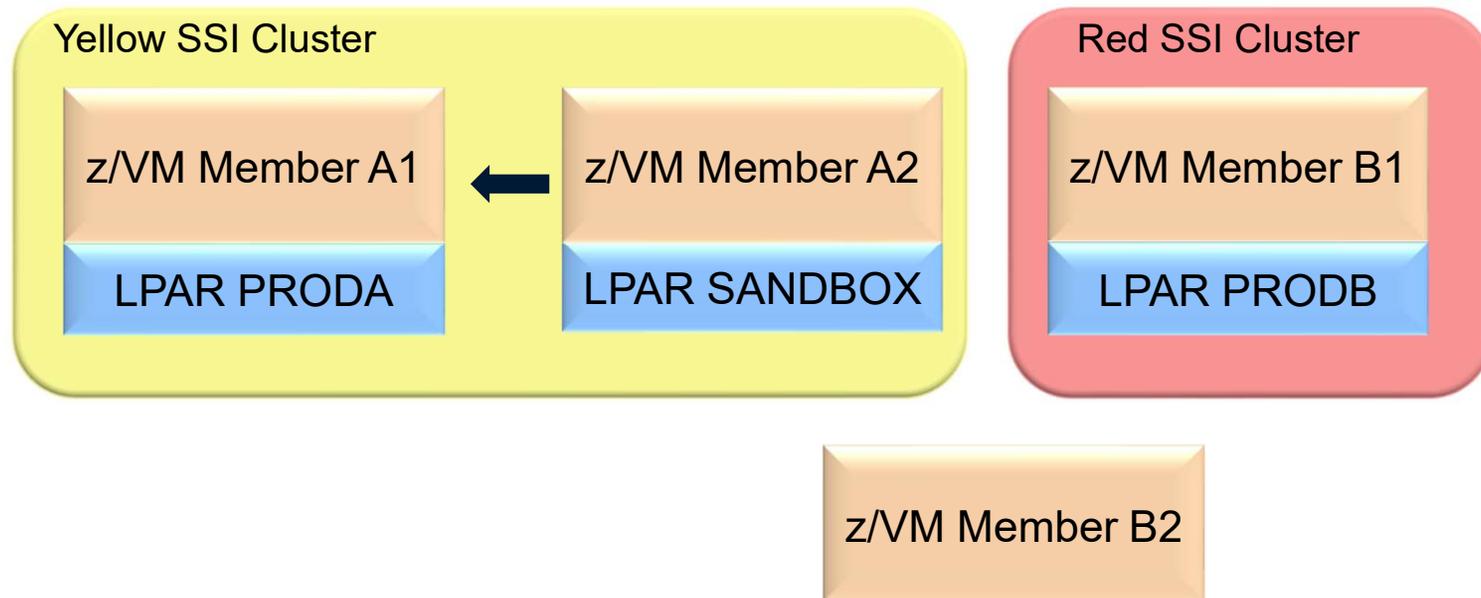Red SSI Cluster

z/VM Member B1

LPAR PRODB

z/VM Member B2

One From the Customers – Utility Migration LPAR

- To update CP on production LPAR PRODA
  Step 4: Bounce A1 to pick up service

Yellow SSI Cluster

z/VM Member A1

LPAR PRODA

z/VM Member A2

LPAR SANDBOX

Red SSI Cluster

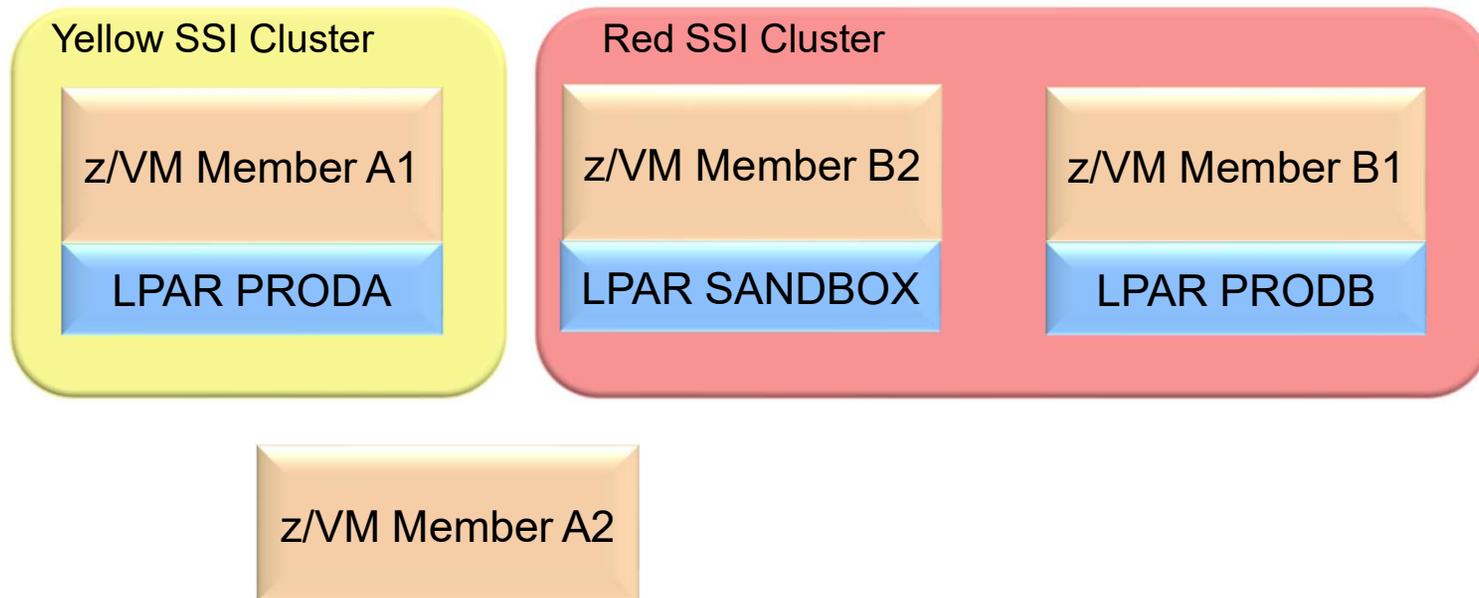z/VM Member B1

LPAR PRODB

z/VM Member B2

One From the Customers – Utility Migration LPAR

- To update CP on production LPAR PRODA
  - Step 5: Move work back to A1 from A2

# One From the Customers – Utility Migration LPAR

- Repeat on Red SSI Cluster

Yellow SSI Cluster

z/VM Member A1

LPAR PRODA

Red SSI Cluster

z/VM Member B2

LPAR SANDBOX

z/VM Member B1

LPAR PRODB

z/VM Member A2

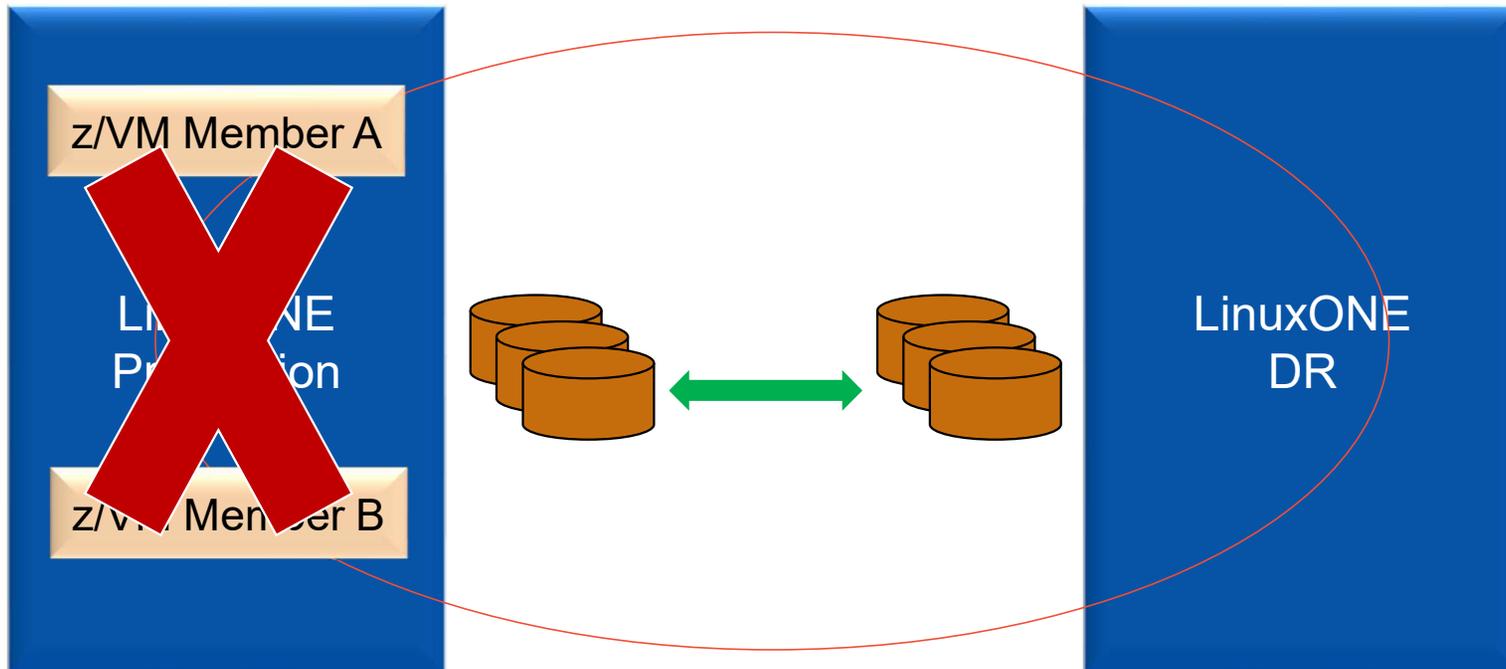# Local Disaster Recover (Business Continuity)

Local Disaster Recovery (Business Continuity)

- Four Members Defined:
  - 2 Members active in production (A & B)
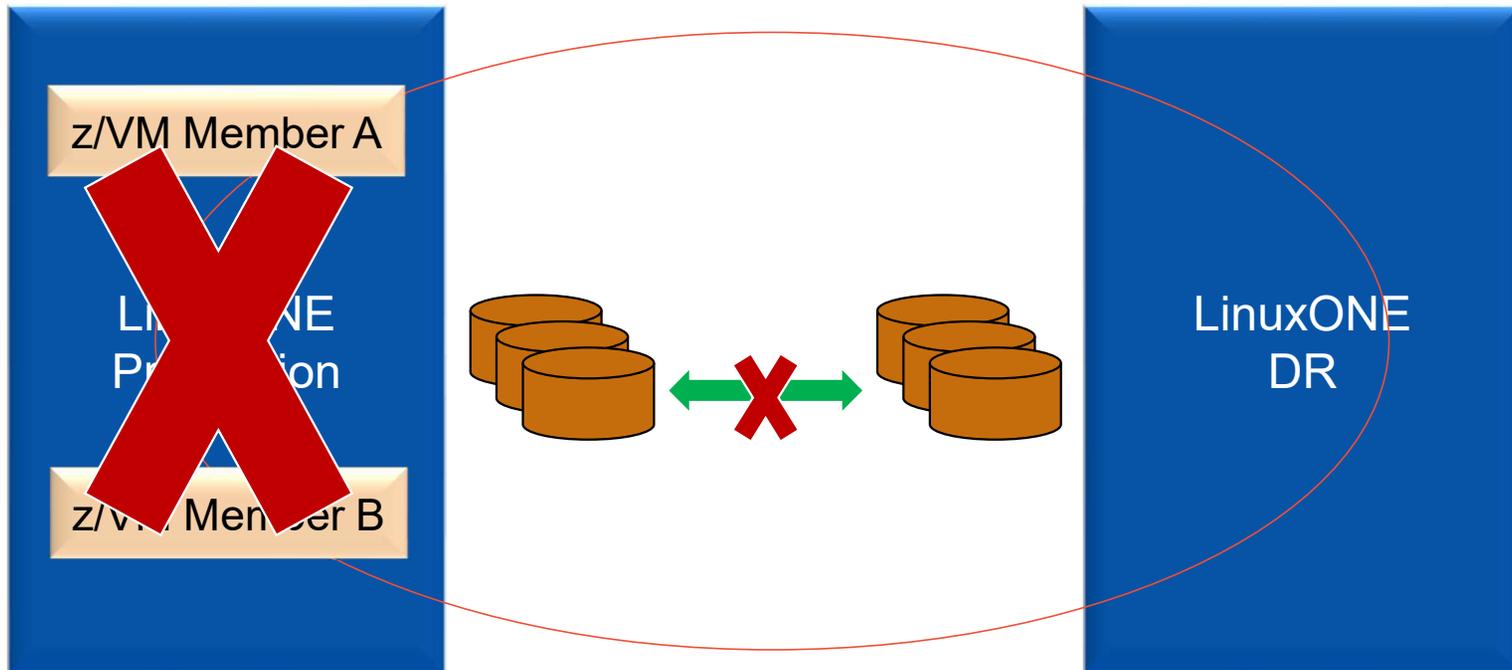 - 2 Members standby in DR (C & D)
  - Mirrored DASD
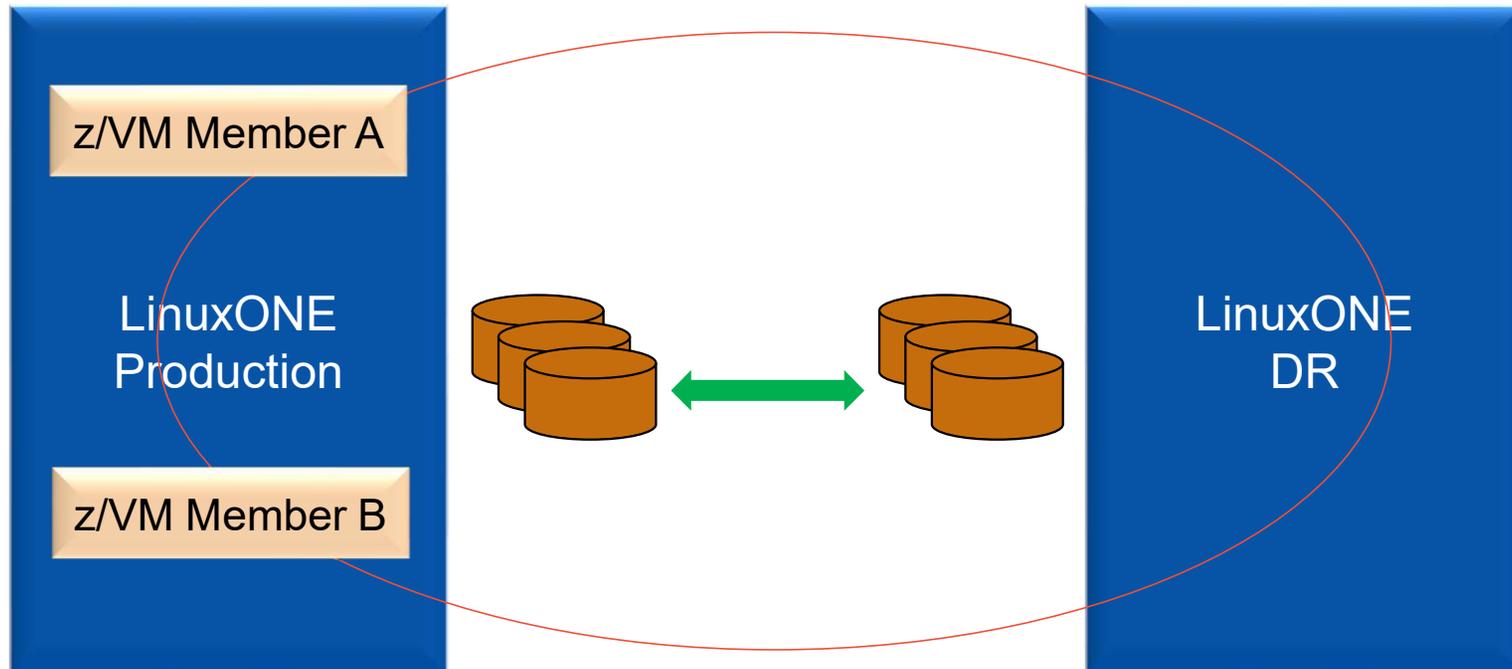
Local Disaster Recovery (Business Continuity)
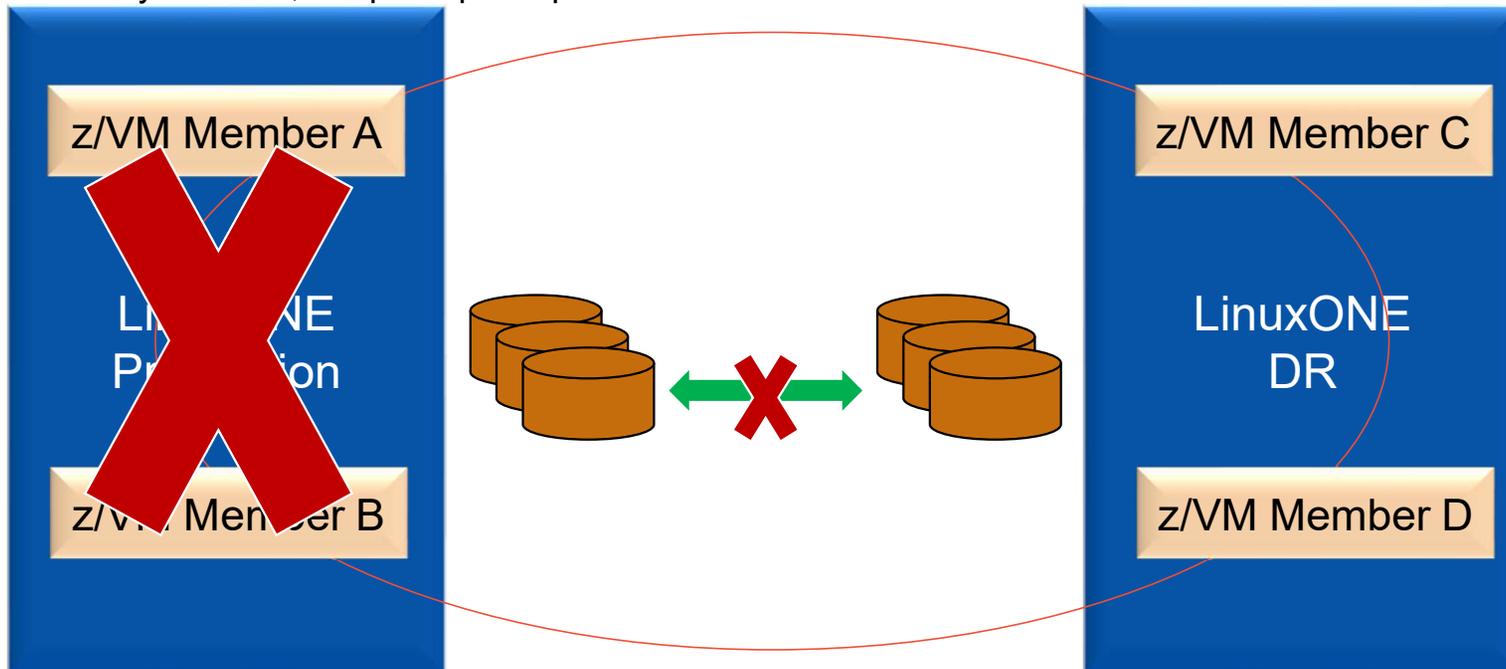
- Assume Production Side goes down

Local Disaster Recovery (Business Continuity)

- Assume Production Side goes down

- Sever mirroring of DASD

# Local Disaster Recovery (Business Continuity)

- Bring up Member C & D

- Logon virtual machines (shared directory)

- Not a High Availability Solution, but perhaps helpful.

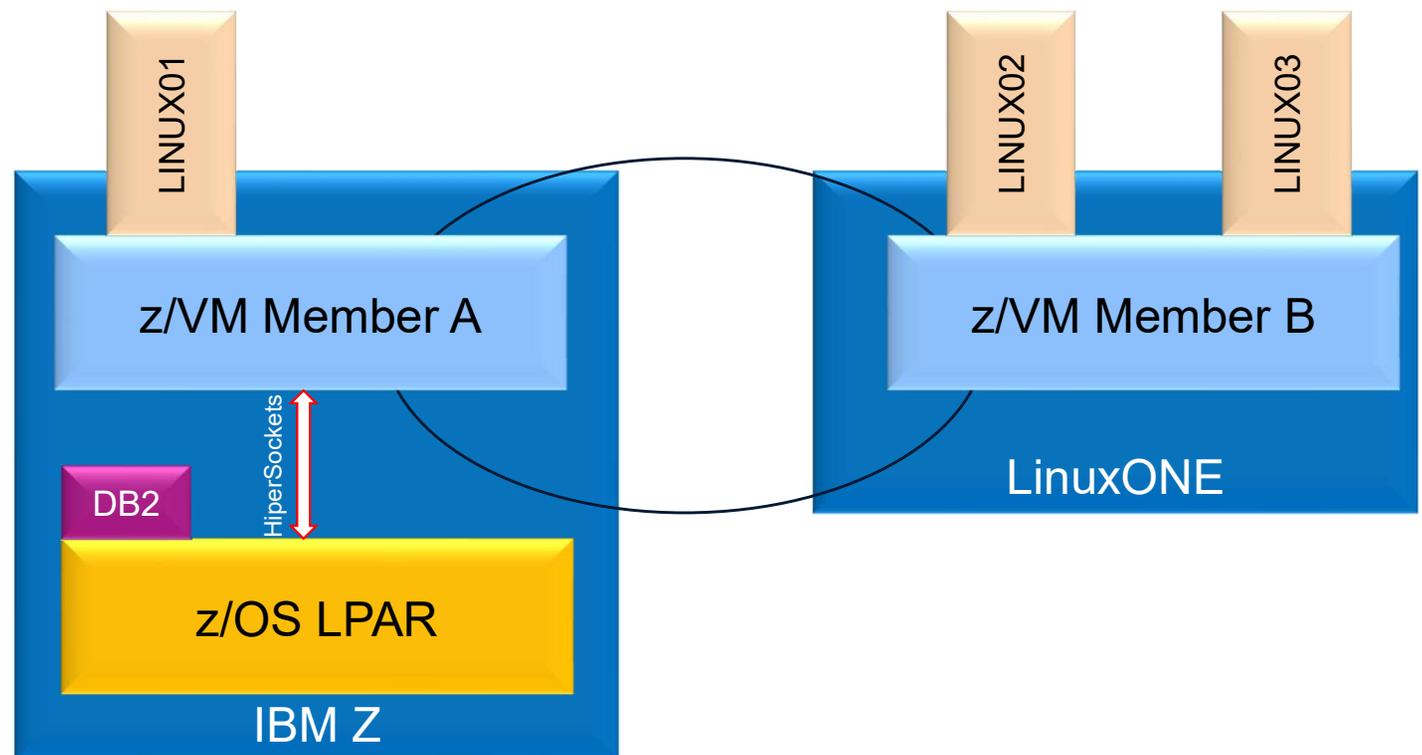# Come Closer for Performance

Come Closer for Performance – Example 1

- Various workloads on Linux guests, many of which use DB2 on z/OS.

- If on same LPAR, they can benefit from the performance of HiperSockets

- LINUX01 is important, so it is closer to z/OS.

LINUX01

LINUX02

LINUX03

z/VM Member A

z/VM Member B

LinuxONE

HiperSockets

DB2

z/OS LPAR

IBM Z
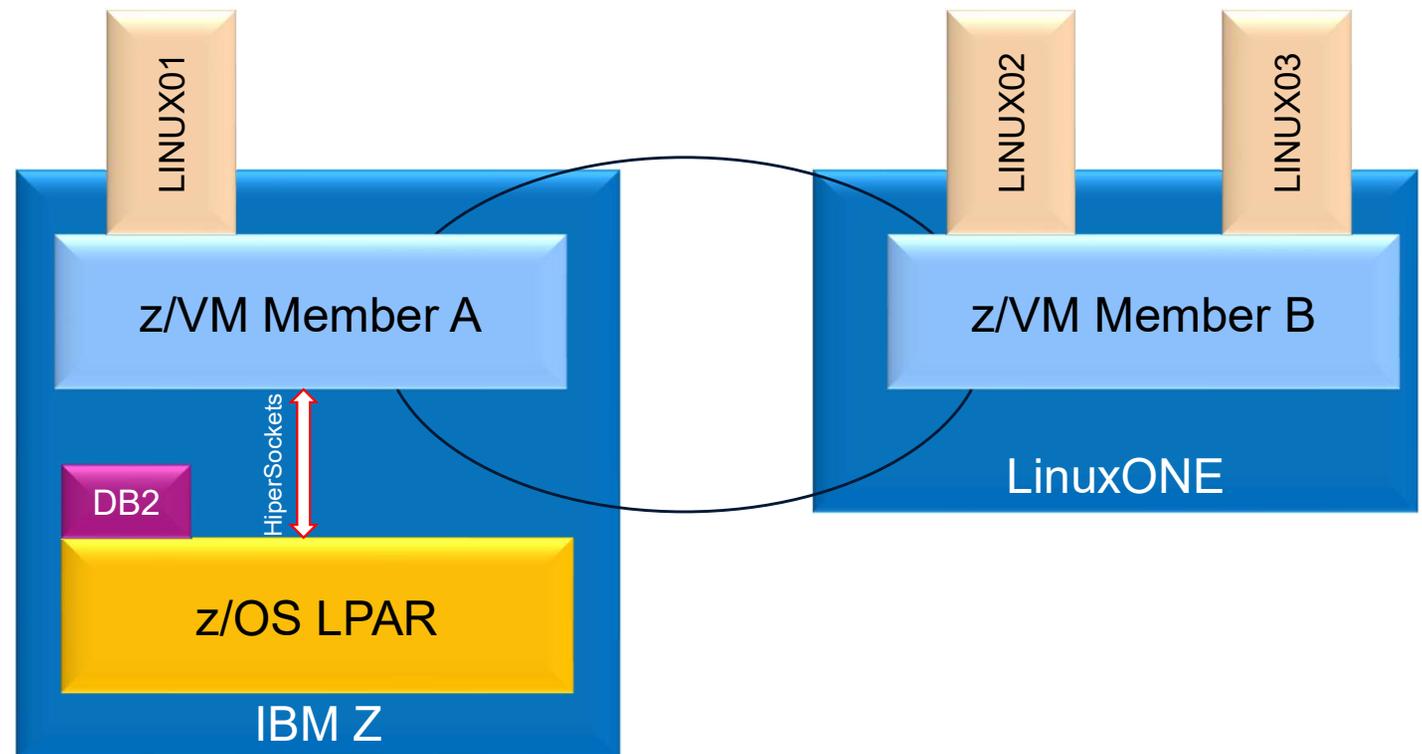
Come Closer for Performance

- A month later, LINUX02 has critical workload increase and could benefit from HiperSockets performance.

- No Problem, just move LINUX02 to Member A with LGR and perhaps LINUX01 to Member B to make room.

Come Closer for Performance – Example 2

- Spectrum Protect server running on one member

- Backing up Linux guests on Member B requires data to me moved over wired network.

- Backing up Linux guests on Member A could be faster because of in memory network.

- No problem, just move Linux guest closer to the Spectrum Protect Server
  – Really only beneficial if disk data backed up is significantly more than memory size.
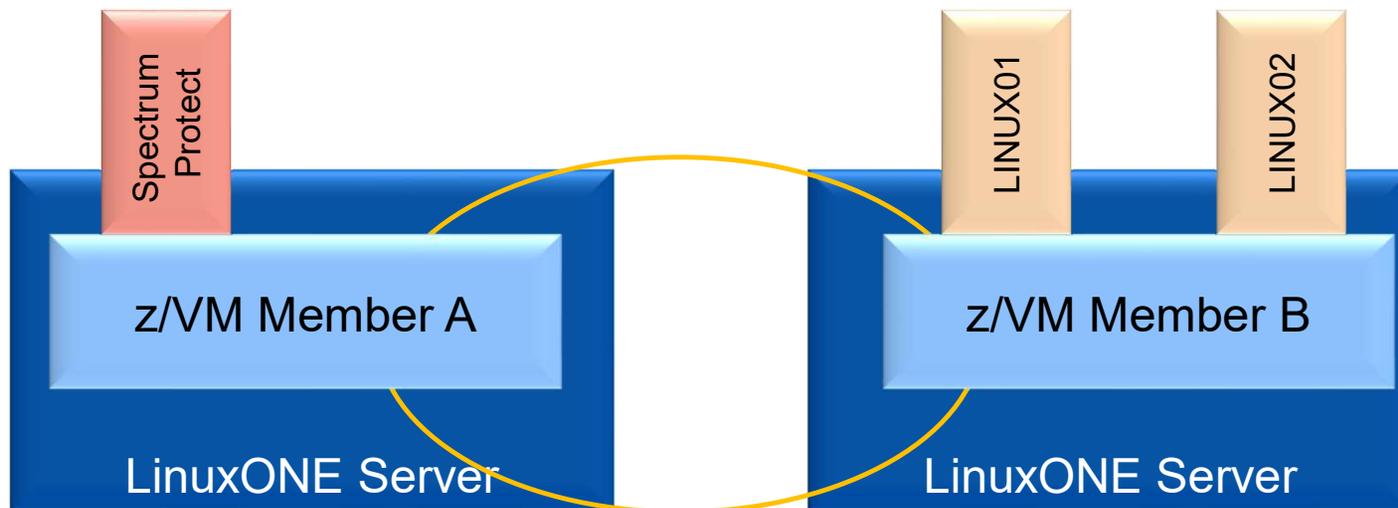
Migrate to New Processor

Migrate to New Processors

- Four Members Defined:
  - 2 Members on each of 2 IBM Z (or LinuxONE) Servers

z/VM Member A	z14	z/VM Member B

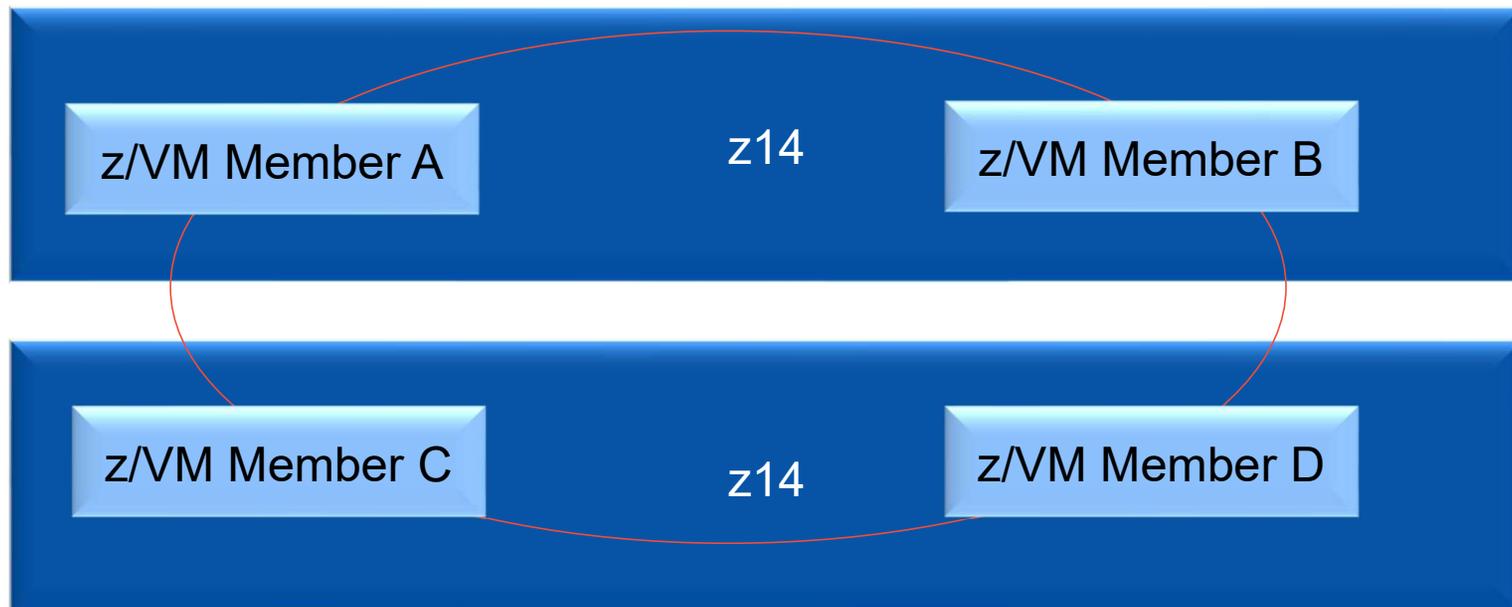| z/VM Member C	z14	z/VM Member D
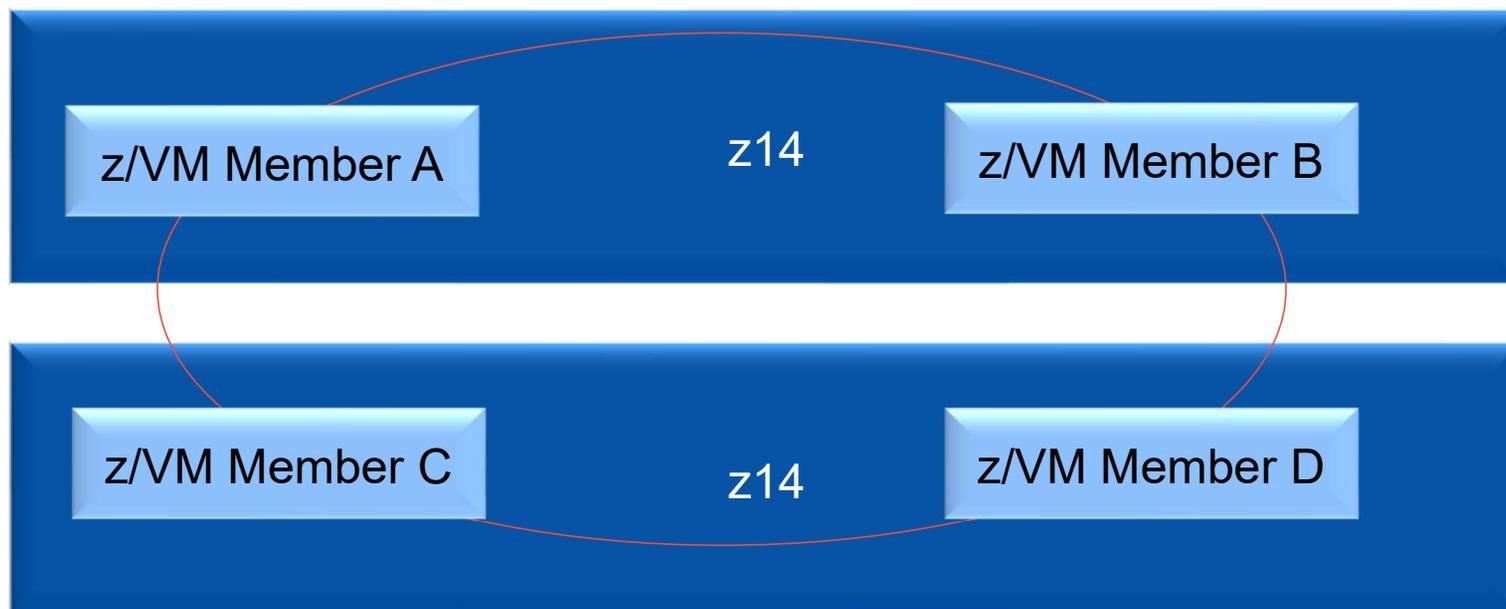
Migrate to New Processors

- Move work off of second z14 to first z14, onto just Members A & B
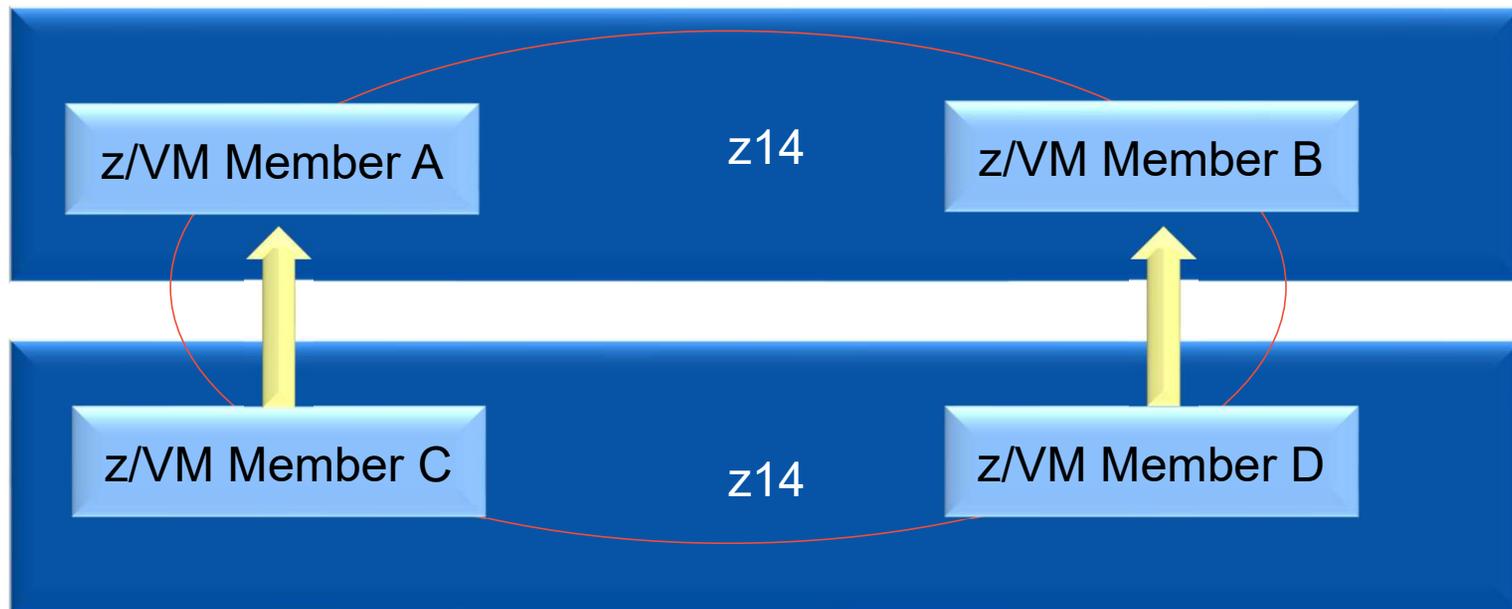
Migrate to New Processors

- Move work off of second z14 to first z14, onto just Members A & B

- Shutdown Members C & D

```
┌─────────────────────────────────────────────────────────────┐
│                              z14                              │
│  ┌──────────────────┐                  ┌──────────────────┐   │
│  │  z/VM Member A   │                  │  z/VM Member B   │   │
│  └──────────────────┘                  └──────────────────┘   │
└─────────────────────────────────────────────────────────────┘

┌─────────────────────────────────────────────────────────────┐
│                              z14                              │
│  ┌──────────────────┐                  ┌──────────────────┐   │
│  │  z/VM Member C   │                  │  z/VM Member D   │   │
│  └──────────────────┘                  └──────────────────┘   │
└─────────────────────────────────────────────────────────────┘
```

Migrate to New Processors

- Push out z14 and pull in the new z15

- Start up Members C & D on the new z15
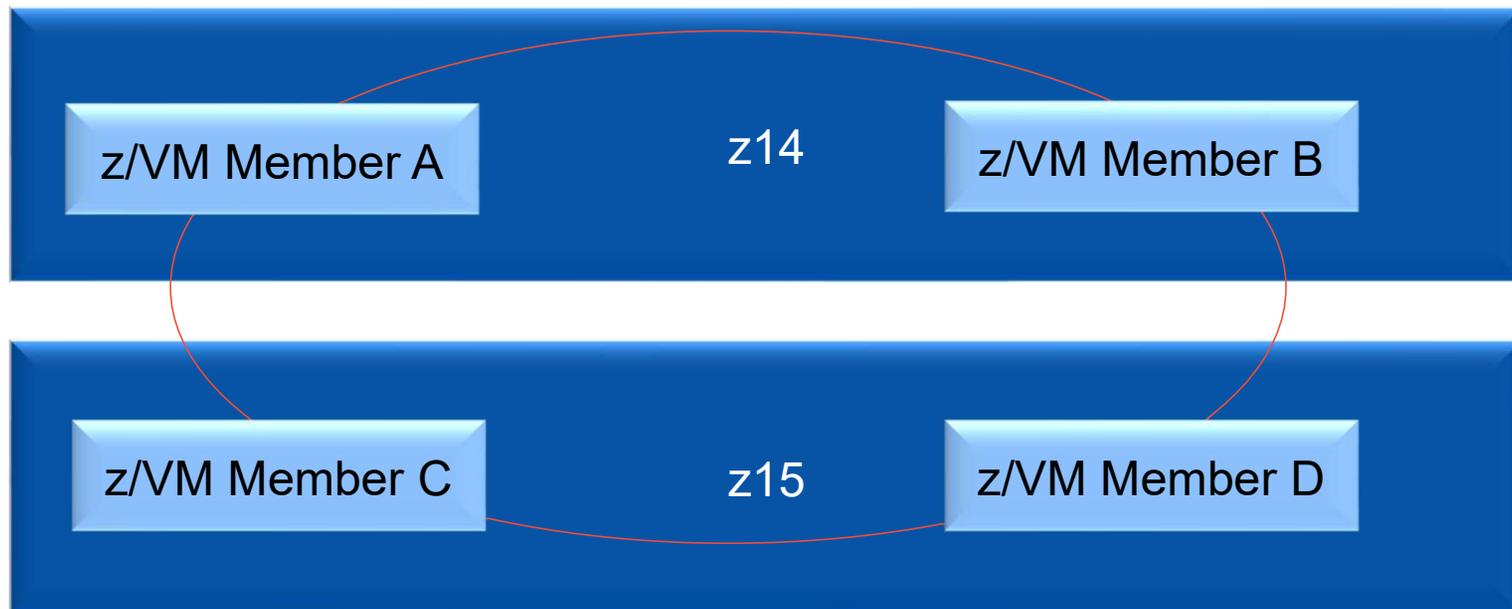
Migrate to New Processor

- Now, move Member A and B workloads to the Members C and D.

z14

z/VM Member A

z/VM Member B

z15

z/VM Member C

z/VM Member D

Migrate to New Processor

- Shutdown Members A and B

- Pull out old z14

- Push in new z15

z14

z/VM Member A   z/VM Member B

z15

z/VM Member C   z/VM Member D

Migrate to New Processor

- Bring back up Members A and B

- Move workloads back to Members A & B

Migrate to New Processor

- Running on new processors without shutting down servers!!

- Would need to re-boot Linux to pick up new z15 hardware facilities.

Summary

Summary: SSI & LGR – Another Milestone for Virtualization

Manage Resources & Workloads	For decades, IBM Z and LinuxONE have has shown the strength of moving resources to the work that needed it. SSI and LGR add more value by allowing work to move to the resources in a non-disruptive manner.
Optimize Success	The SSI clustering takes advantage of hardware and software technology to optimize success and minimize the complexity of clustering, with low overhead and without specialized hardware.
Protect the Advantage	Guest mobility in is a remarkable technology. z/VM Live Guest Relocation takes it to the next level. Exploiting LGR doesn't mean giving up the rich resource control and management features you have come to love with z/VM.