Recent Activities

Rick Harper

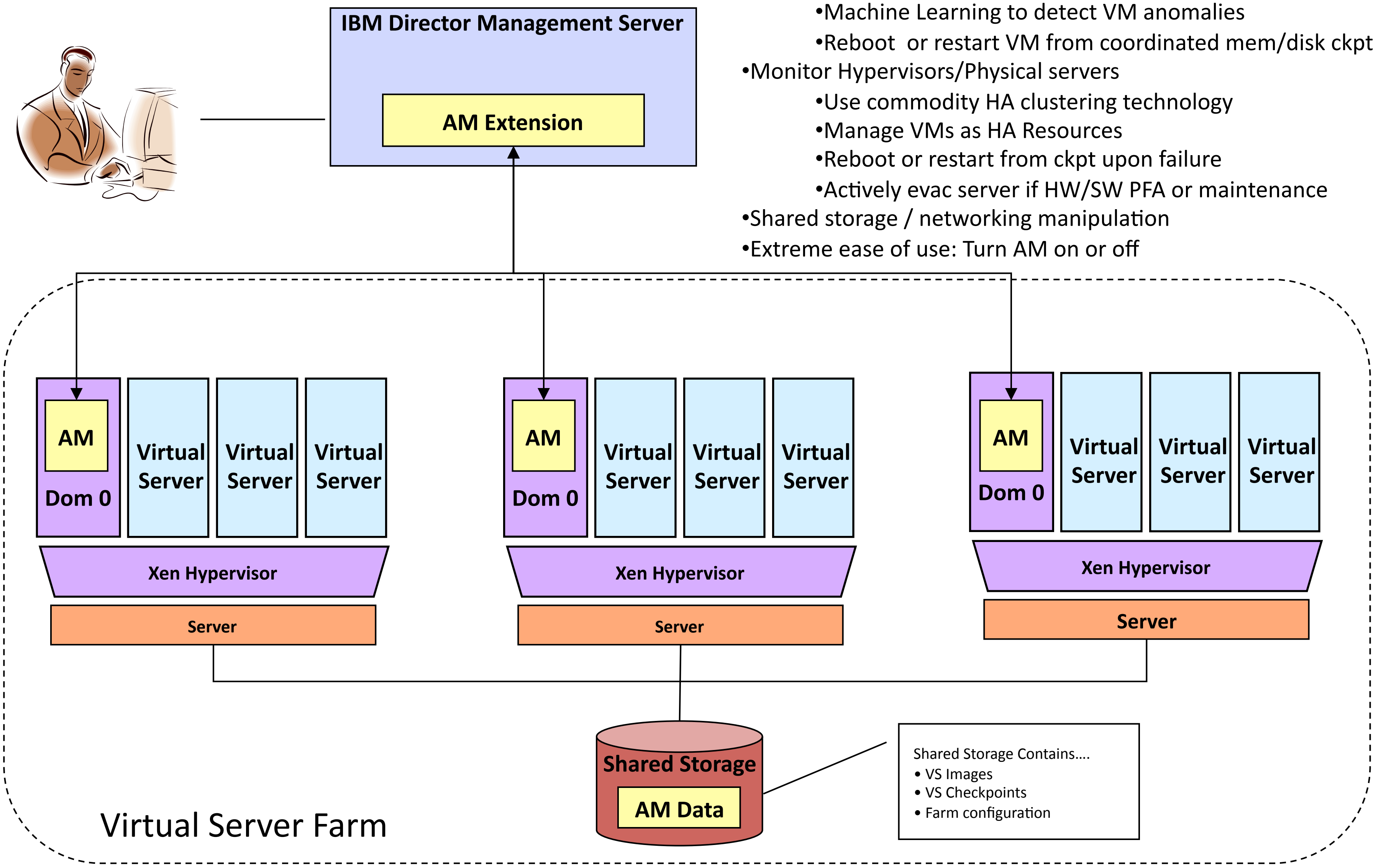
IBM Research

Overview

- Virtual Machine Availability Management
- Black Box / White Box Availability Management

Availability Manager Architecture (2007 Xen)

- Monitor VMs
    - Machine Learning to detect VM anomalies
    - Reboot or restart VM from coordinated mem/disk ckpt
- Monitor Hypervisors/Physical servers
    - Use commodity HA clustering technology
    - Manage VMs as HA Resources
    - Reboot or restart from ckpt upon failure
    - Actively evac server if HW/SW PFA or maintenance
- Shared storage / networking manipulation
- Extreme ease of use: Turn AM on or off



IBM Director Management Server

AM Extension

AM

Dom 0

Virtual Server

Virtual Server

Virtual Server

Xen Hypervisor

Server

AM

Dom 0

Virtual Server

Virtual Server

Virtual Server

Xen Hypervisor

Server

AM

Dom 0

Virtual Server

Virtual Server

Virtual Server

Xen Hypervisor

Server

Shared Storage

AM Data

Shared Storage Contains….

- VS Images
- VS Checkpoints
- Farm configuration

Virtual Server Farm

Black Box /White Box AM

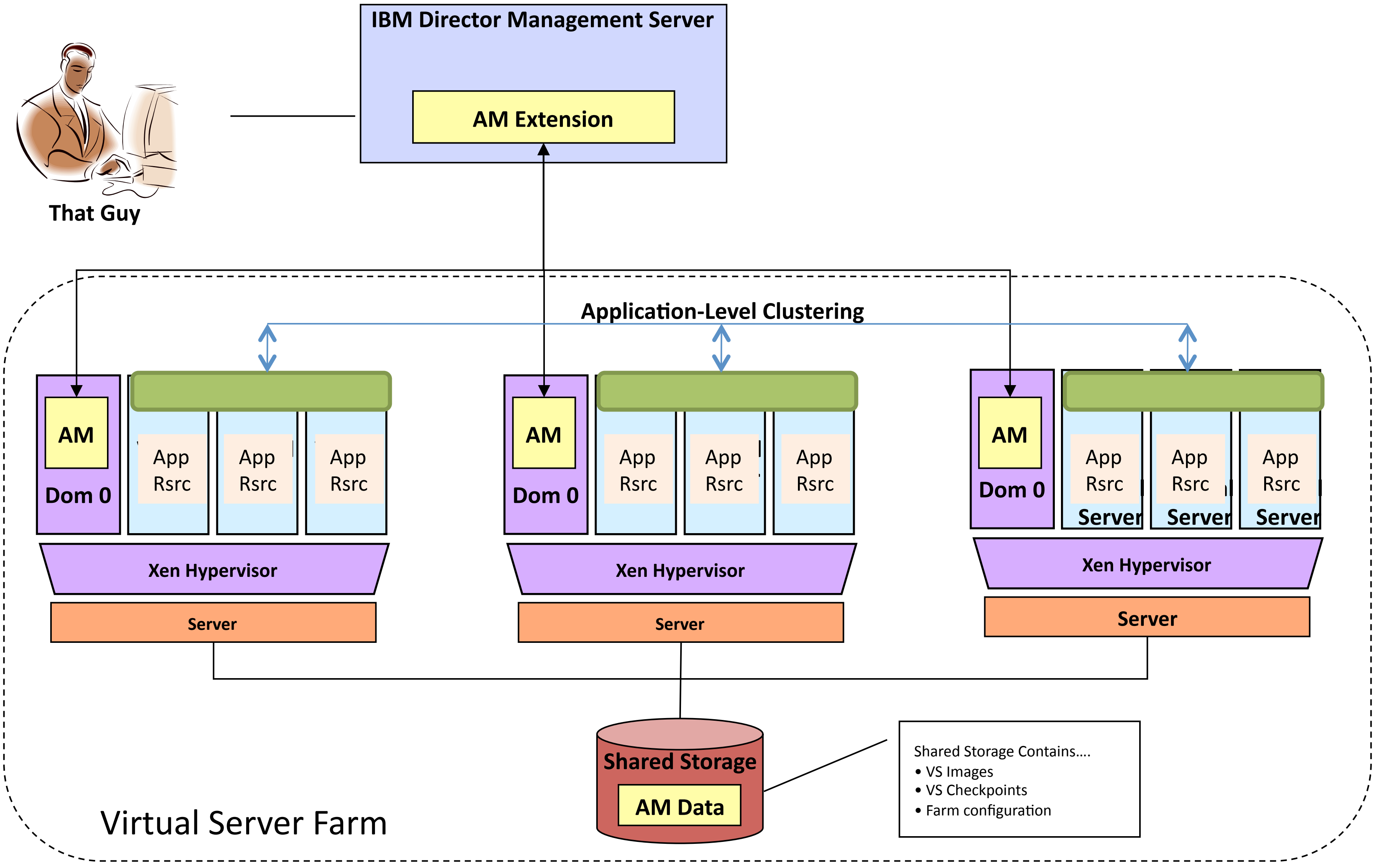
- VM AM has strengths and weaknesses
  - Payload-agnostic AM – broad market coverage
  - Can do cool things like active migration, checkpointing, active replication

  …but
  - No visibility into payload hence cannot respond to application faults
- Application-level AM has strengths and weaknesses
  - Provides ability to respond to application faults
  - Many important applications are handled using current technology

  …but
  - Not payload-agnostic – must be crafted for each app
  - Cannot actively migrate work (failover model, typically)

- Combining the two yields significant availability advantages
- There is a lot of potential cross-fertilization
  - Ex: Applying Virtual Resource Placement algorithms to Application Resource placement and failover planning

Black Box / White Box AM

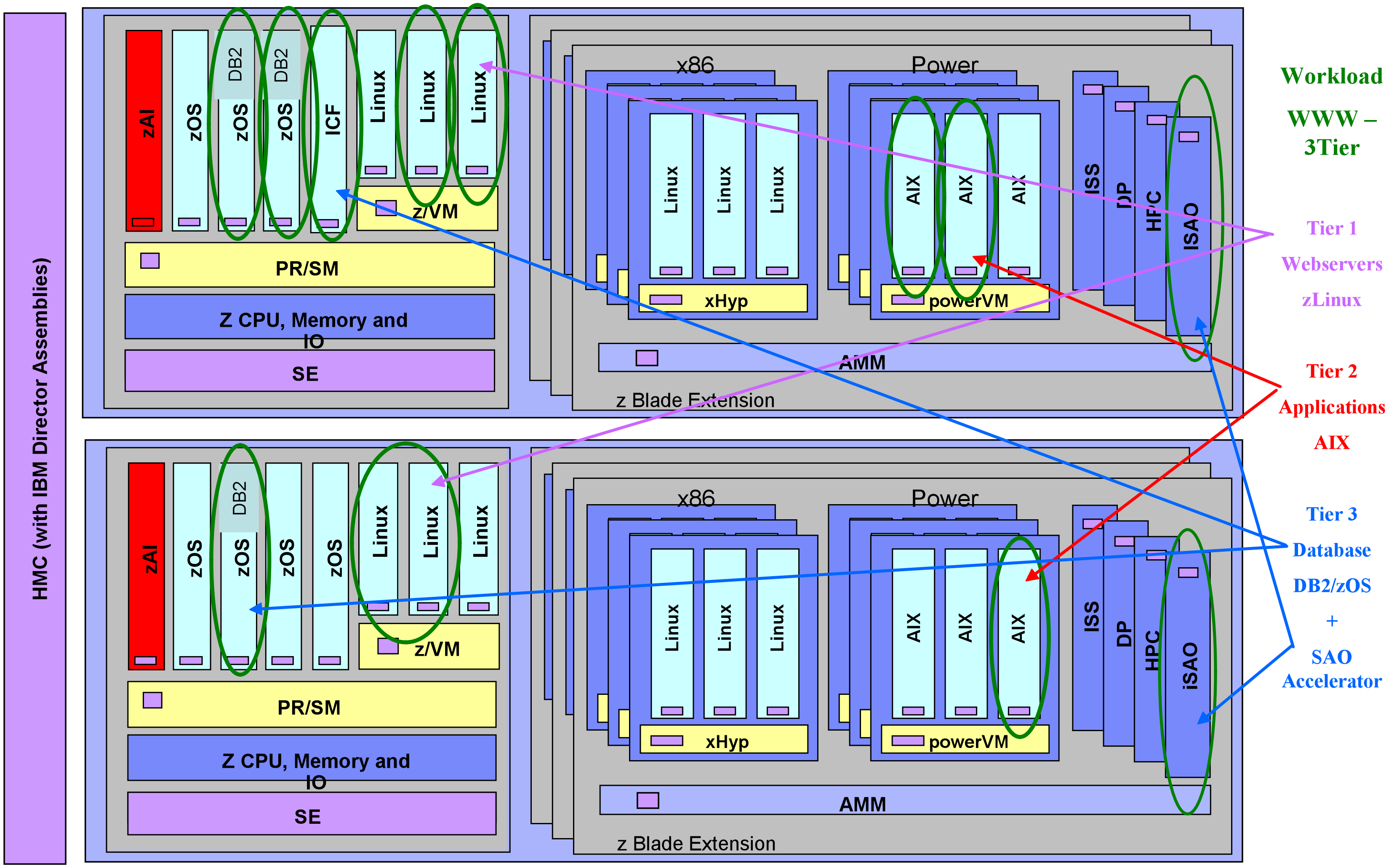


Shared Storage Contains….

- VS Images
- VS Checkpoints
- Farm configuration

Black Box / White Box AM

Internal Logical View across 2 z196 CECs with zBX



Workload

WWW – 3Tier

Tier 1

Webservers

zLinux

Tier 2

Applications

AIX

Tier 3

Database

DB2/zOS

+

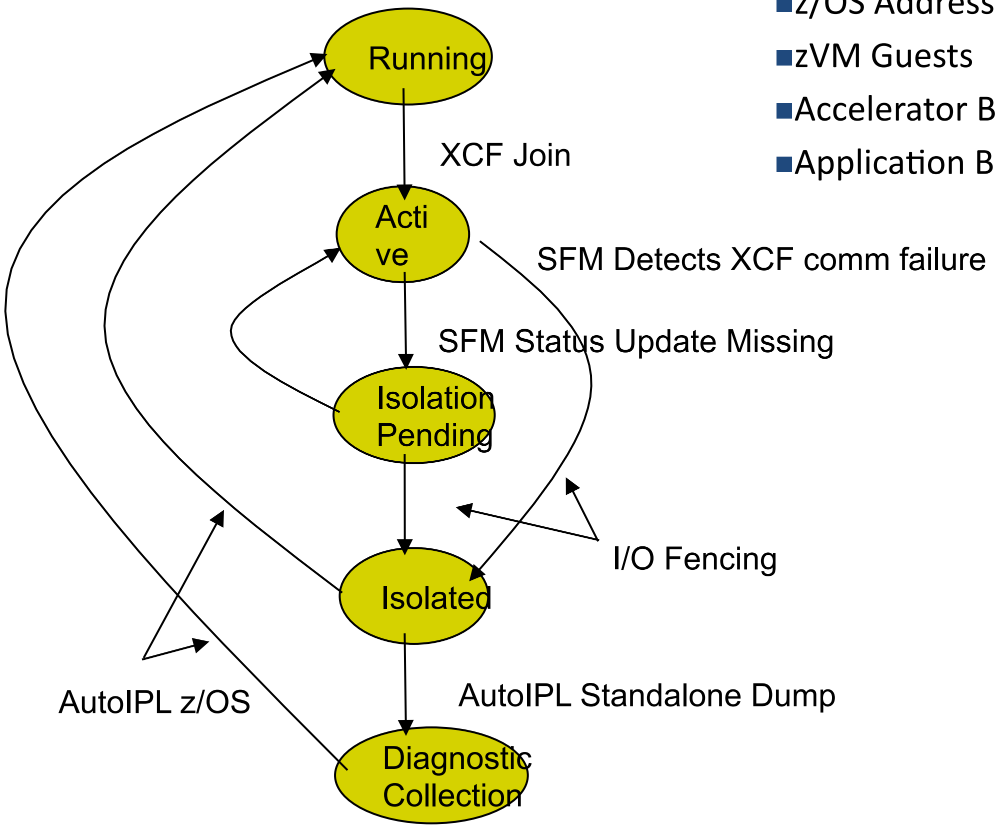
SAO Accelerator

Availability and Performance Managers on z

- Existing System z Managers
  - XCF Topology/Comm. Services
  - XCF Status Monitoring
  - Sysplex Failure Manager (SFM)
  - AutoIPL
  - CF Resource Manager (CFRM)
  - Sysplex Distributor
  - Geographically Distributed Parallel Sysplex(GDPS)
  - Automatic Restart Manager (ARM)
  - Workload Manager (WLM)
  - Intelligent Resource Director (IRD)
  - eWLM
  - Resource Recovery Services (RRS)
  - Predictive Failure Analysis
  - Health Checker
  - Tivoli System Automation for z/OS
  - Runtime Diagnostics (RTD)

Managed Entities

- z/OS LPAR
- z/OS Address Space
- zVM Guests
- Accelerator Blades
- Application Blades



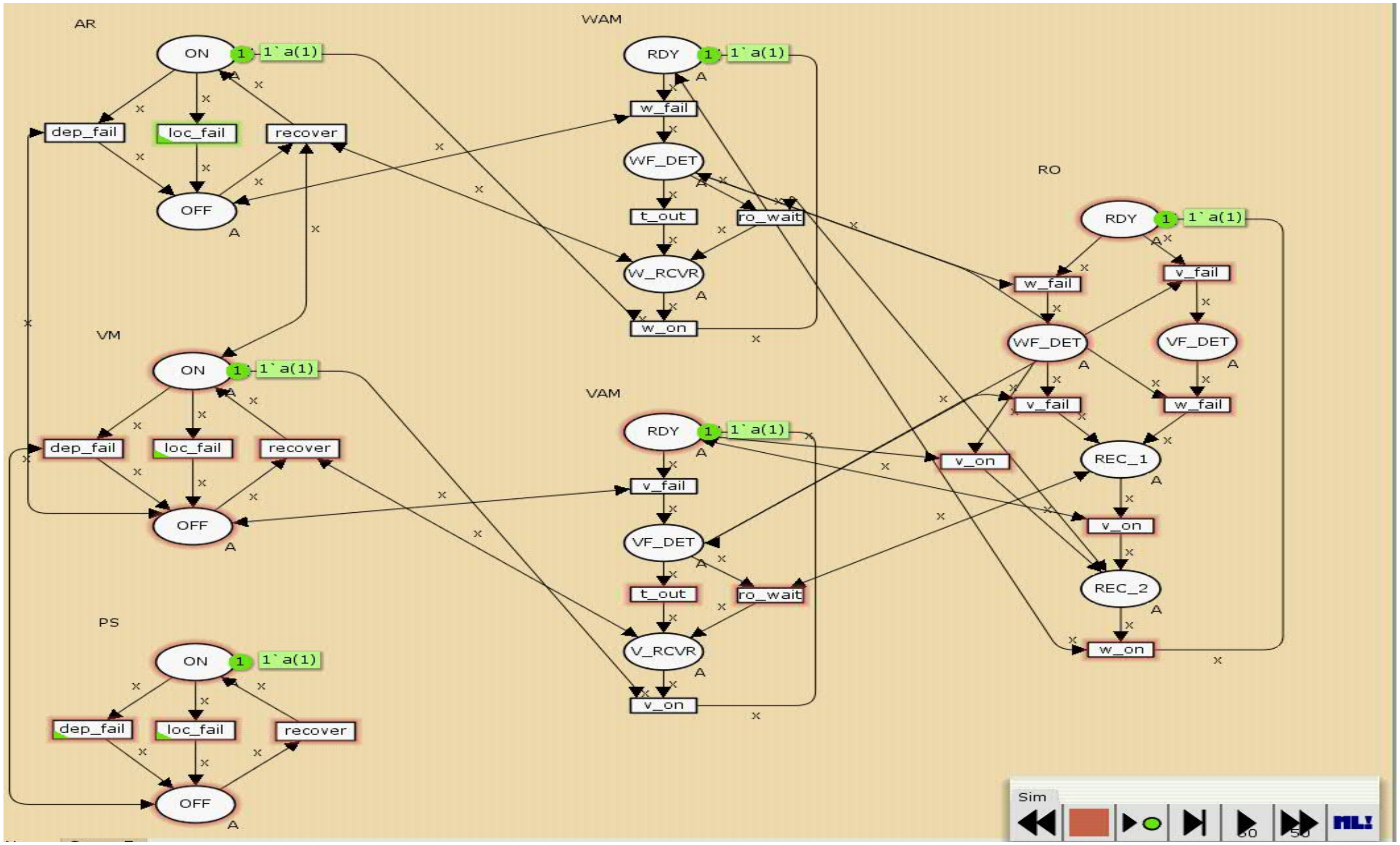
Axes of Discrimination

- This is a high dimensionality problem
- Must accommodate and abstract the following dimensionality
  - – Horizontal and Vertical Spatial Composition
    - – Horizontal Composition: Different managers manipulate different resources
    - – Vertical Composition: Different managers manipulate related / dependent resources
  - – Behavioral Composition
    - – How different managers work together to concurrently manipulate related / dependent resources

– Design *goal*

  - Managers need no modification to work together
  - Manipulate existing functionality and interfaces to achieve interoperability

Colored Petri-Net (CPN) Model of Behavioral Composition

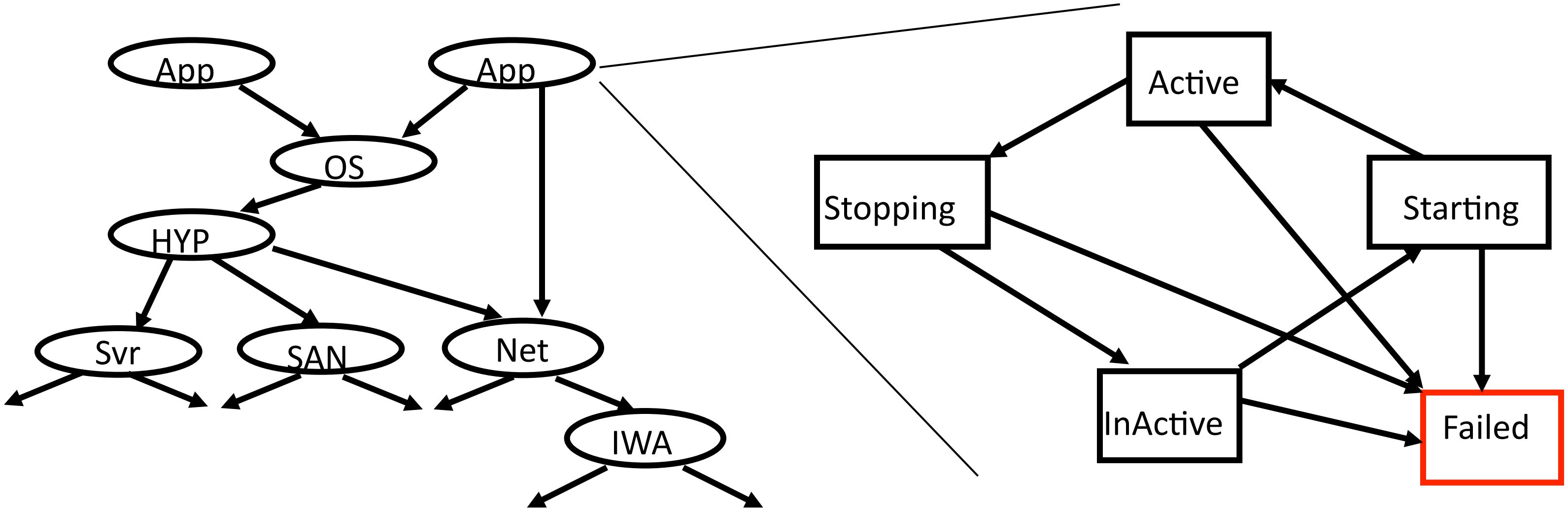


- Resources: 1 Application Resource, 1 Virtual Machine, 1 Physical Server
- Resilience Managers: White Box Availability Manager, Virtual Machine Availability Manager, Resilience Orchestrator

Resource Relationship Model to support Concurrent Multimanager Composition

Models Relationships:
“Resources” are related in complex and dynamic ways
Dependency; Start Order; Locality; Collocation etc

Models Behavior:
Resources have extremely complex internal behavior; but common mechanisms can be defined: such as Detection, Notification, Manipulation, Analysis



- How can multiple managers coexist and coordinate manipulation of shared and/or dependent resources?
- How can multiple managers quiesce/reactivate each other’s resource management?
- How can multimanager concurrency be supported?
  - RRM indicates which manager ops interact and which don’t
  - Any manager (Availability or other) wanting to manipulate a resource can lock the (actualized and possibly contingent) dependency network prior to manipulation
    - Acyclic digraph nature of RDM reduces possibility of deadlock
  - How do managers communicate with each other
    - Via manipulation of the shared resource dependency model