

IBM eServer BladeCenter as a Dependable Web Infrastructure Platform

Steven W Hunter

IBM Corporation
1/27/2005

*IFIP Working Group 10.4
Winter Meeting 2005
University of Puerto Rico
Mayaguez, Puerto Rico*

BladeCenter Overview

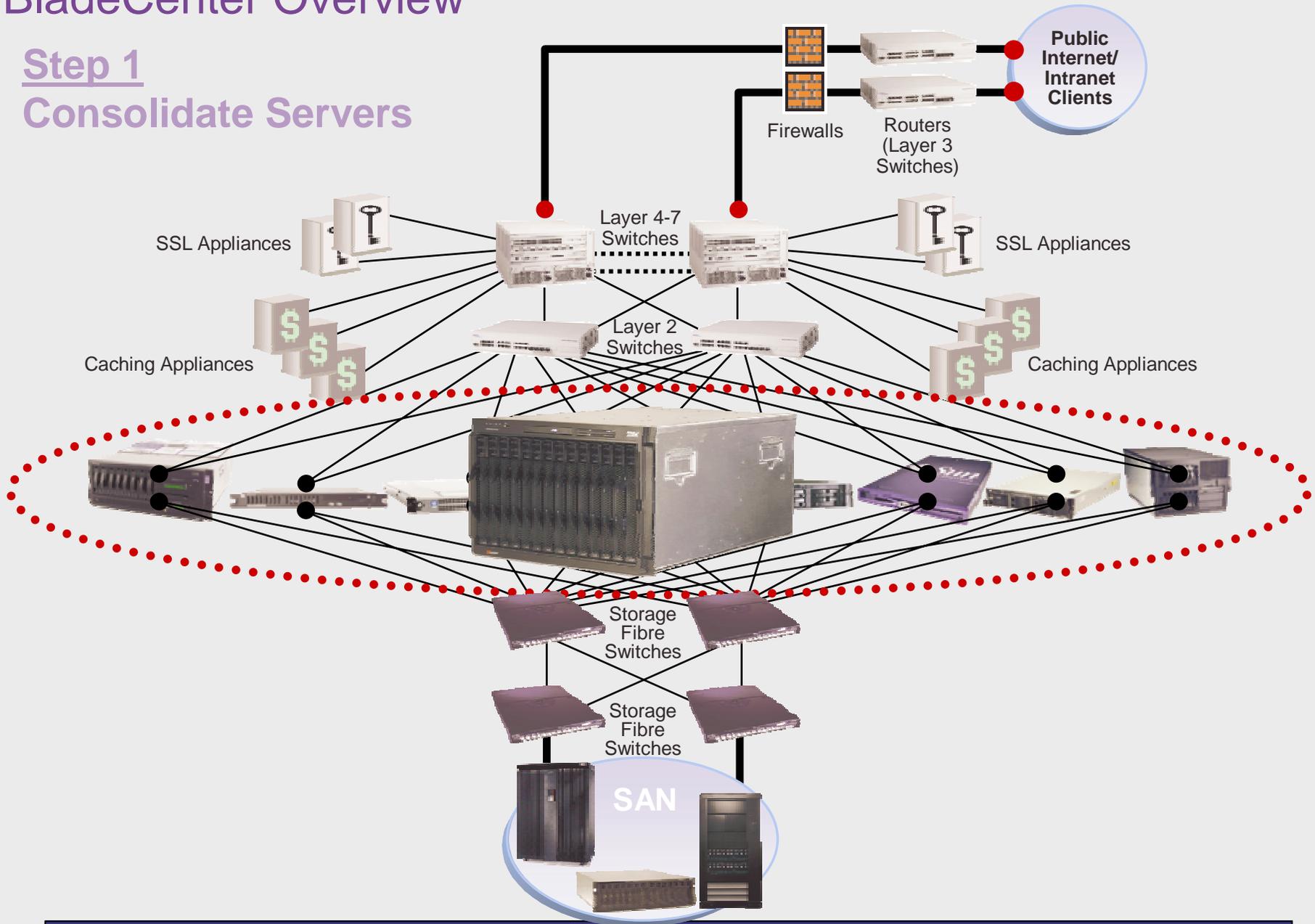


- **Modular, Scalable**
 - 1 – 14 Processor Blades
- **Density with Performance**
 - 7U Mechanical Chassis
- **Integrated Network Infrastructure**
 - Switching with point-to-point blade connections
- **Affordable Availability**
 - Redundant, Hot-swappable blades and modules
- **Advanced Systems Management**
 - Integrated service processor



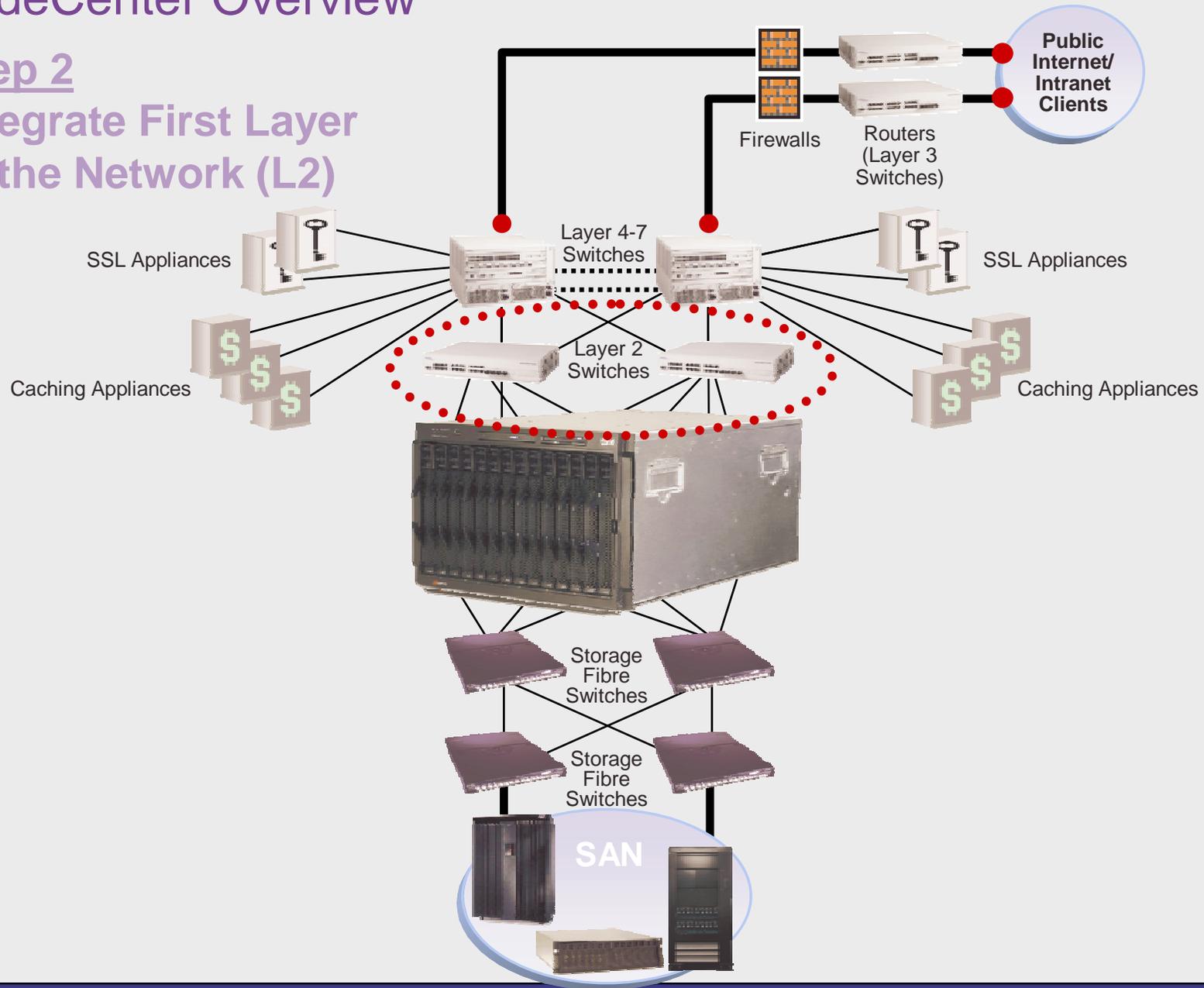
BladeCenter Overview

Step 1 Consolidate Servers



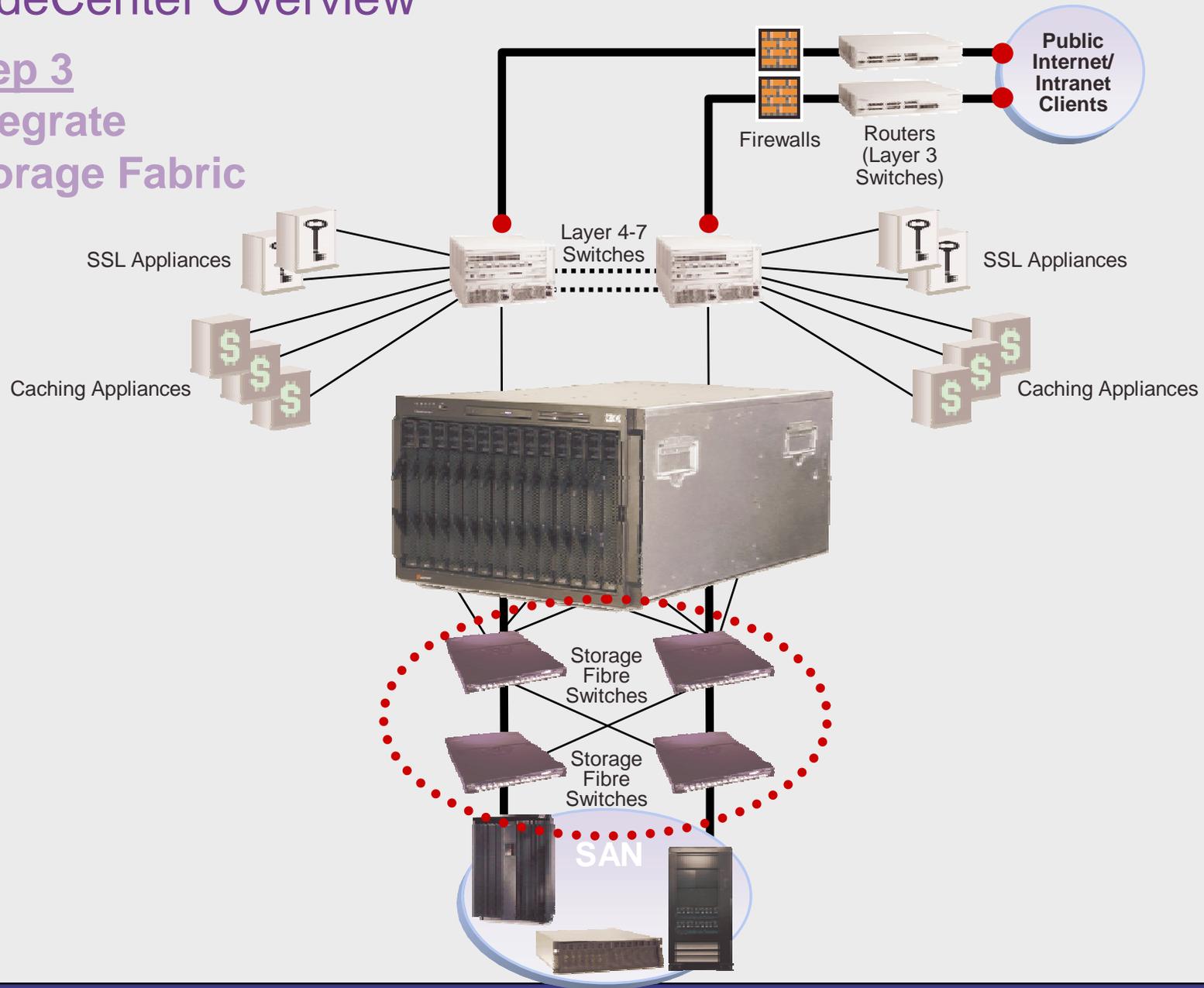
BladeCenter Overview

Step 2 Integrate First Layer of the Network (L2)



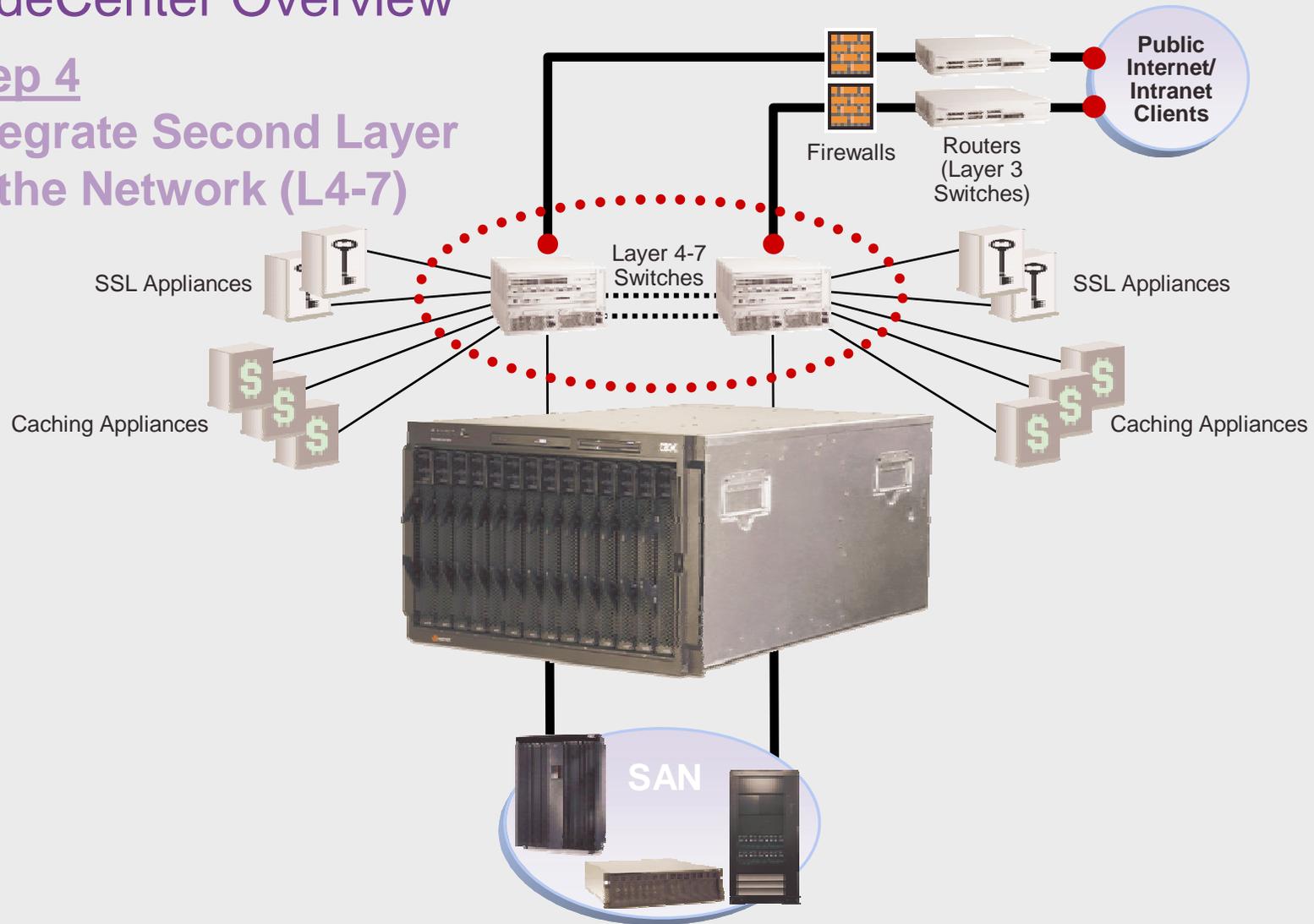
BladeCenter Overview

Step 3 Integrate Storage Fabric



BladeCenter Overview

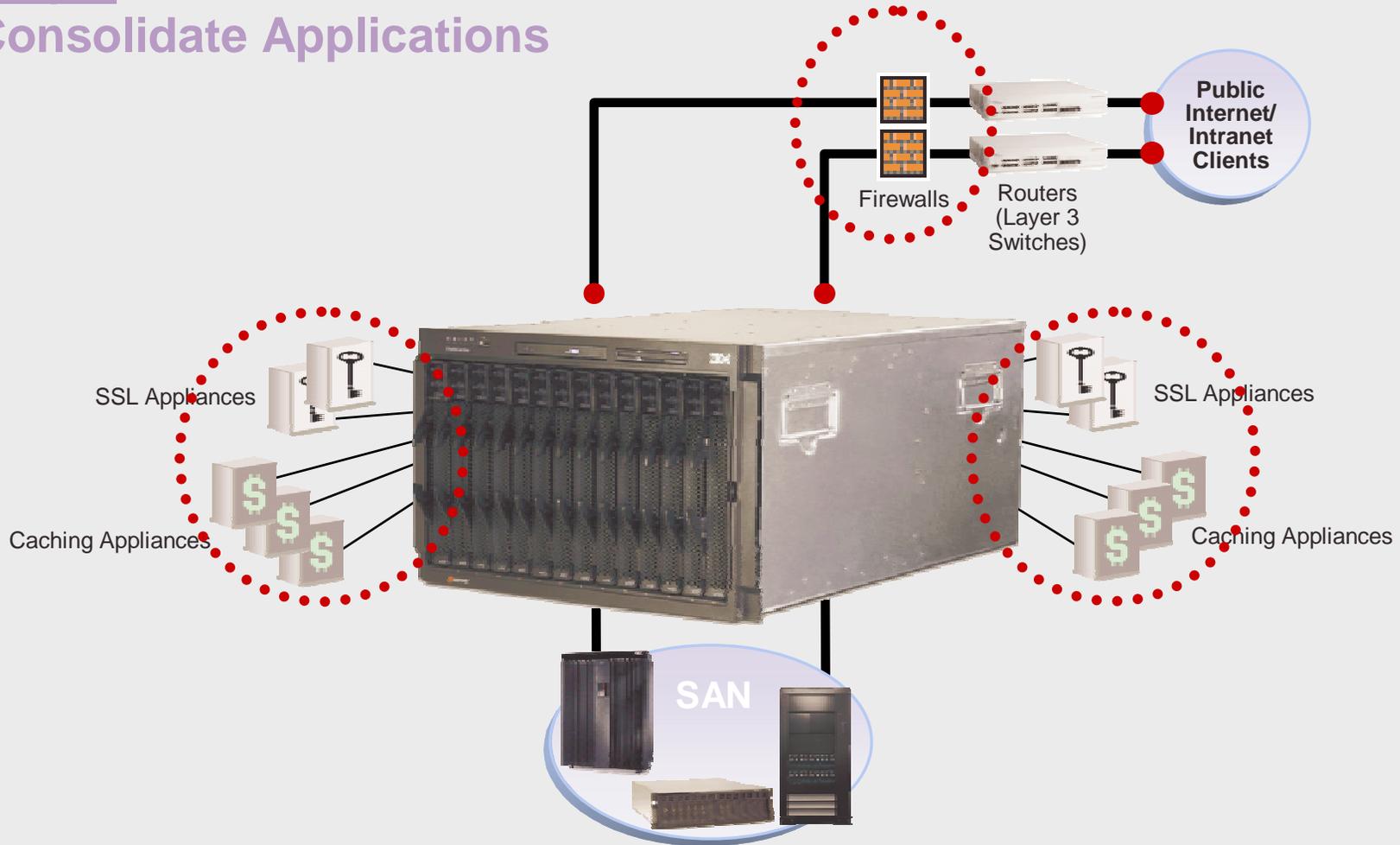
Step 4 Integrate Second Layer of the Network (L4-7)



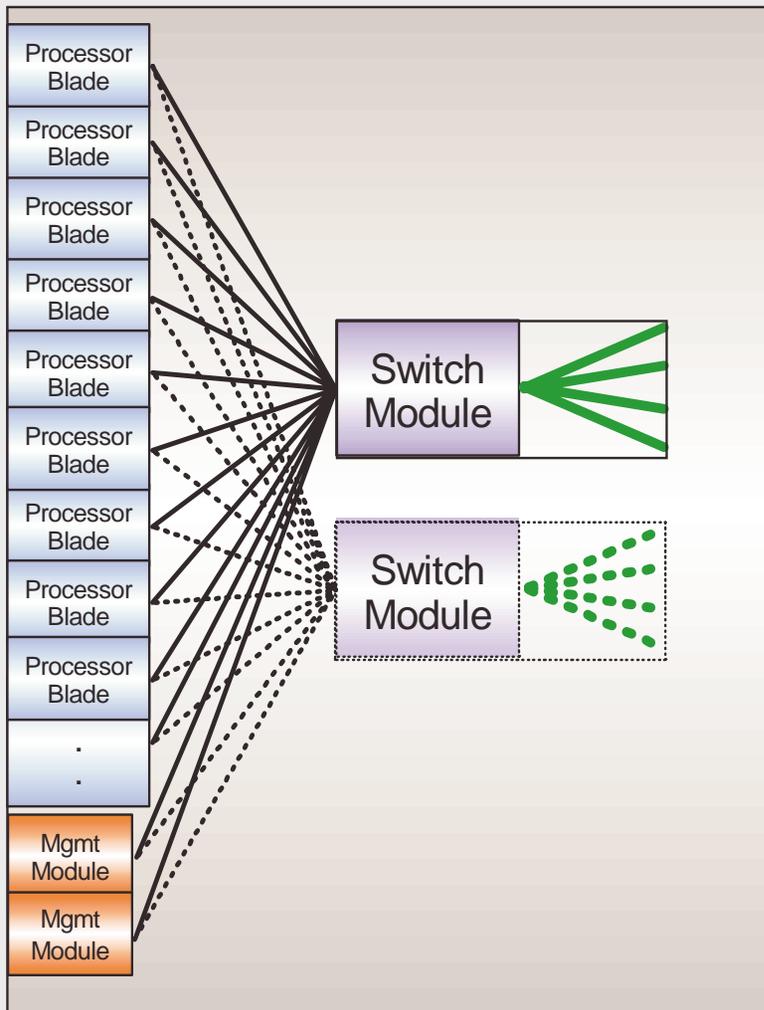
BladeCenter Overview

Step 5

Consolidate Applications



BladeCenter Overview



Switching Modules

- **Fibre Channel**
- **Ethernet**
- **Others...**

Blade I/O Card

- **I/O expansion card matches switch technology in the corresponding slot**



BladeCenter Overview

Gigabit Ethernet Switches (Layer 2)

- **Commodity level networking**
- **Link aggregation**
- **VLAN partitioning and management**

Advanced Switching (Layer 2-7)

- **Load Balancing**
- **Content-based switching**

Fibre Channel Switches

- **Lower cost via integration**
- **Full support of FC-SW-2 standards**

Power (4 x 1800W load-balancing)

- **Upgradeable as required**
- **Redundant and load balancing for HA**

Calibrated, vectored cooling™

- **Fully fault tolerant**
- **Allow maximum processor speeds**

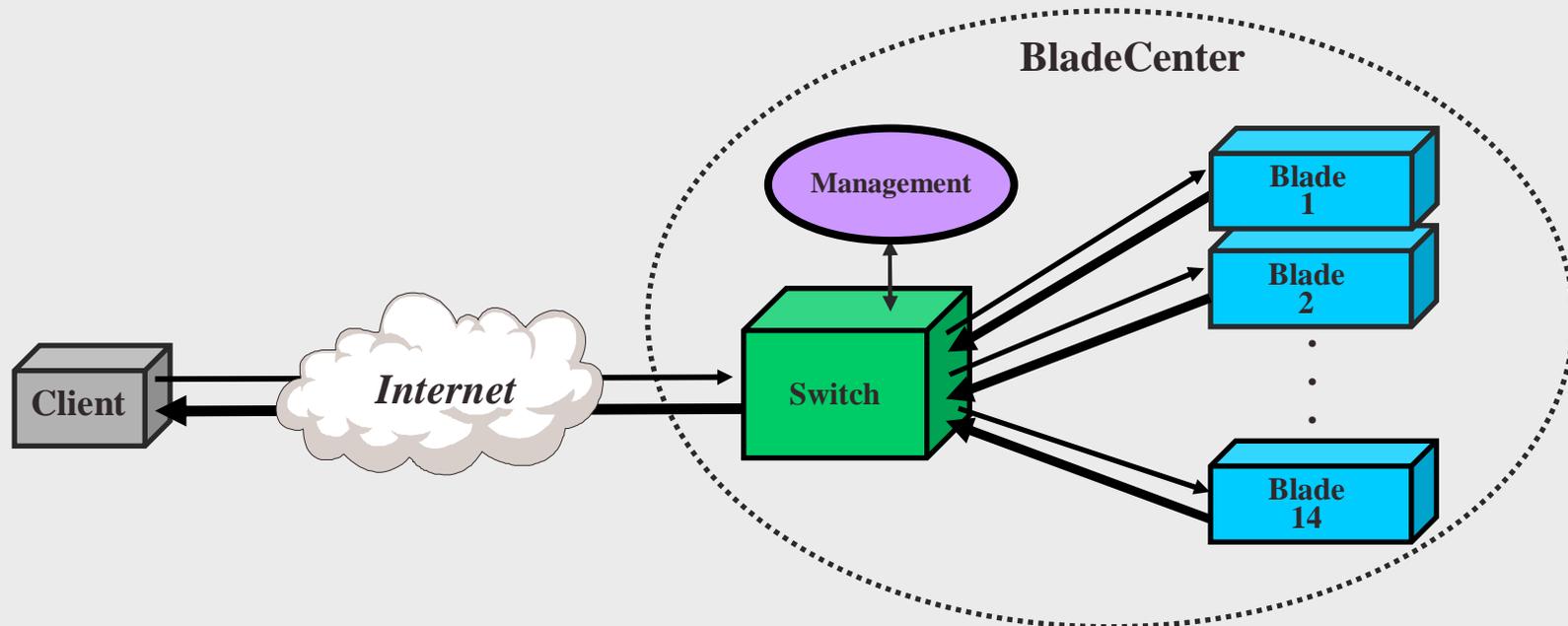
KVM Switches / Management Modules

- **Full remote video redirection**
- **Out-of-band / lights out systems management**



Autonomic Web Computing with BladeCenter

- **Integrated switching enables autonomic functions through a common control point**
 - Layer 2 switching provides basic standard functionality
 - Layer 4 (load balancing) and Layer 7 (content switching) for advanced web clustering
 - Layer 4/7 enables control point for directing traffic to up to fourteen blades
 - Web clusters are a popular method of workload management
- **Examples of autonomic functions include performance, management, health, power, etc.**
 - Automated workload management supports performance optimization and failover of blades
 - VM technology applied to blades to further improve granularity
 - Software health addressed with rejuvenation techniques
 - Power management can be addressed at multiple levels



Autonomic Web Computing with BladeCenter

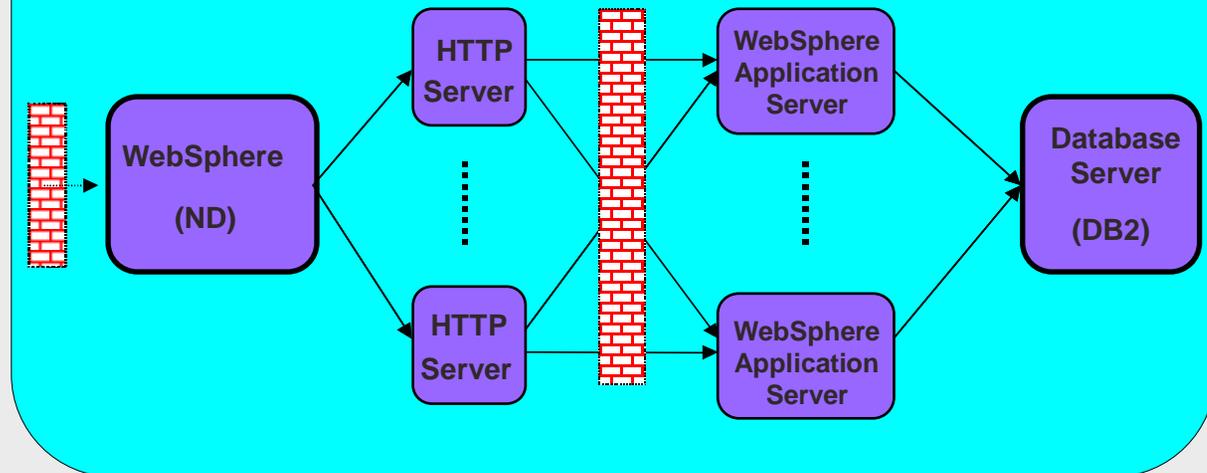
Multi-Tier Infrastructure

- Front-End Load Balancer
- Web Servers
- Application Servers
- Data Base Server

Infrastructure Automation

- Initially configures chassis & network and dynamically configures new and failover blades
- Automatically deploys and configures software stack (OS, middleware & apps) & network VLANs
- Monitors CPU load and predicts need for additional capacity (configures from free pool)

Typical 4 tier web infrastructure for e-business



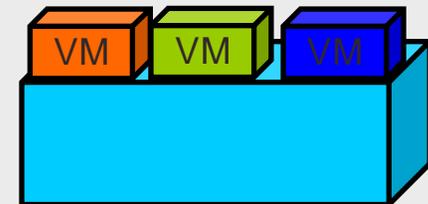
Solution Details

- Opus automatically provisions HTTP and WAS tiers
- IBM Tivoli Intelligent Orchestrator 1.1 (ITITO) policy-based analysis can determine when to schedule provisioning
- Opus utilizes IBM Director, Remote Deployment Manager for bare-metal install of Linux or Windows OS
- Opus workflows to install WebSphere Application Server/IBM HTTP Server/J2EE application, update Load balancer and HTTP Plug-in configuration files

Autonomic Web Computing with BladeCenter

Virtual Machines

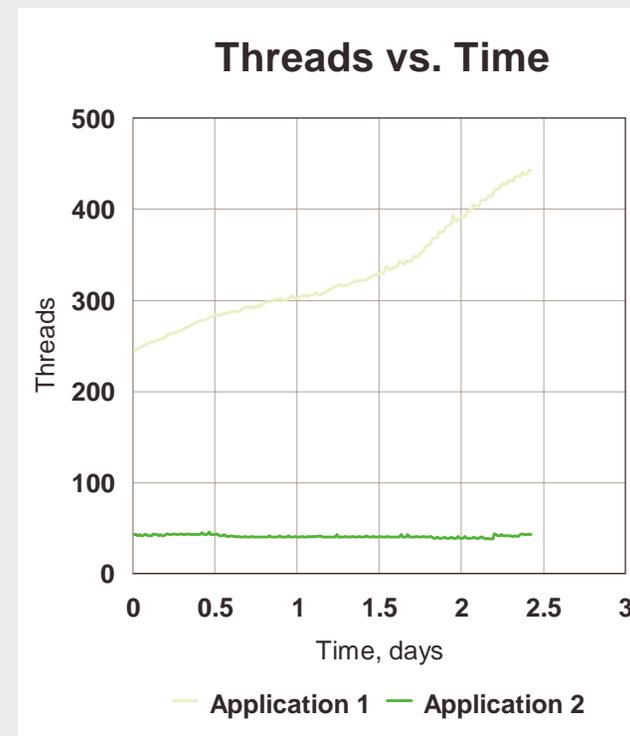
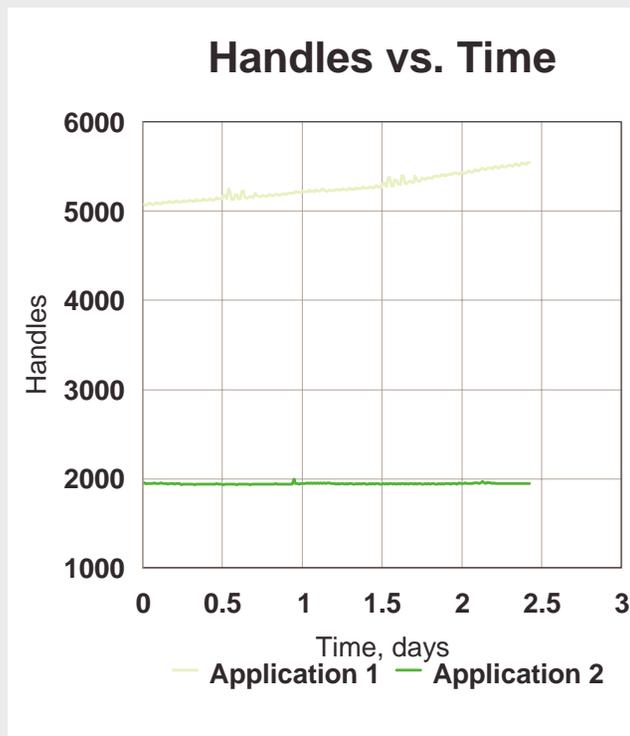
- **VM technology such as VMware applied to blades for server consolidation**
 - **Orchestration and provisioning tools also apply to virtual machines**
- **VMware's VMotion technology enhances failover by transferring the entire system and memory state of a running virtual machine from one ESX Server to another**
 - **The Systems' disk, including all of its data, software and boot partitions, must be stored on a shared storage infrastructure such as a SAN**
 - **Keeps track of on-going memory transactions in a bitmap, which is kept small**
 - **When the memory and system state has been copied to the target server. VMotion:**
 1. **Suspends the source VM**
 2. **Copies the bitmap to the target ESX Server**
 3. **Resumes the VM on the target ESX Server**
- **The process takes less than 2 seconds (i.e., "hiccup time") on a Gigabit Ethernet network and appears as no more than a temporary network loss to the app, service and/or user.**
 - **It's necessary to keep this length of time minimal, since it leverages the operation of the TCP protocol for guaranteed delivery of lost packets.**



Autonomic Web Computing with BladeCenter

Software Rejuvenation

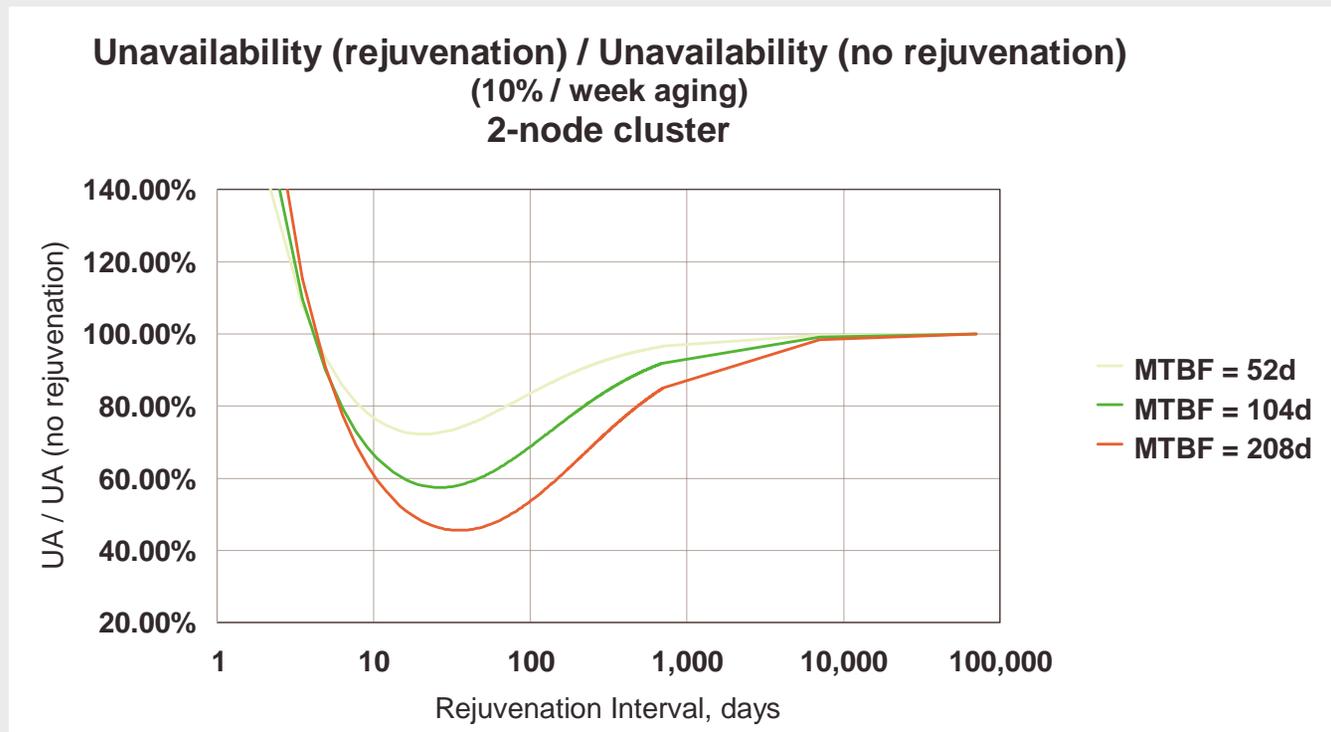
- System outages are far more likely to be a result of software errors than hardware failures
- Software (OS, middleware, applications, actually, state) ages with time...
 - memory leaks, handle leaks, nonterminated threads, unreleased file-locks, data corruption
 - ...resulting in Bad Things (outages, hangs, ...)
- Software failure prediction and state rejuvenation is a proactive technology designed to mitigate the effects of software aging



Autonomic Web Computing with BladeCenter

Software Rejuvenation

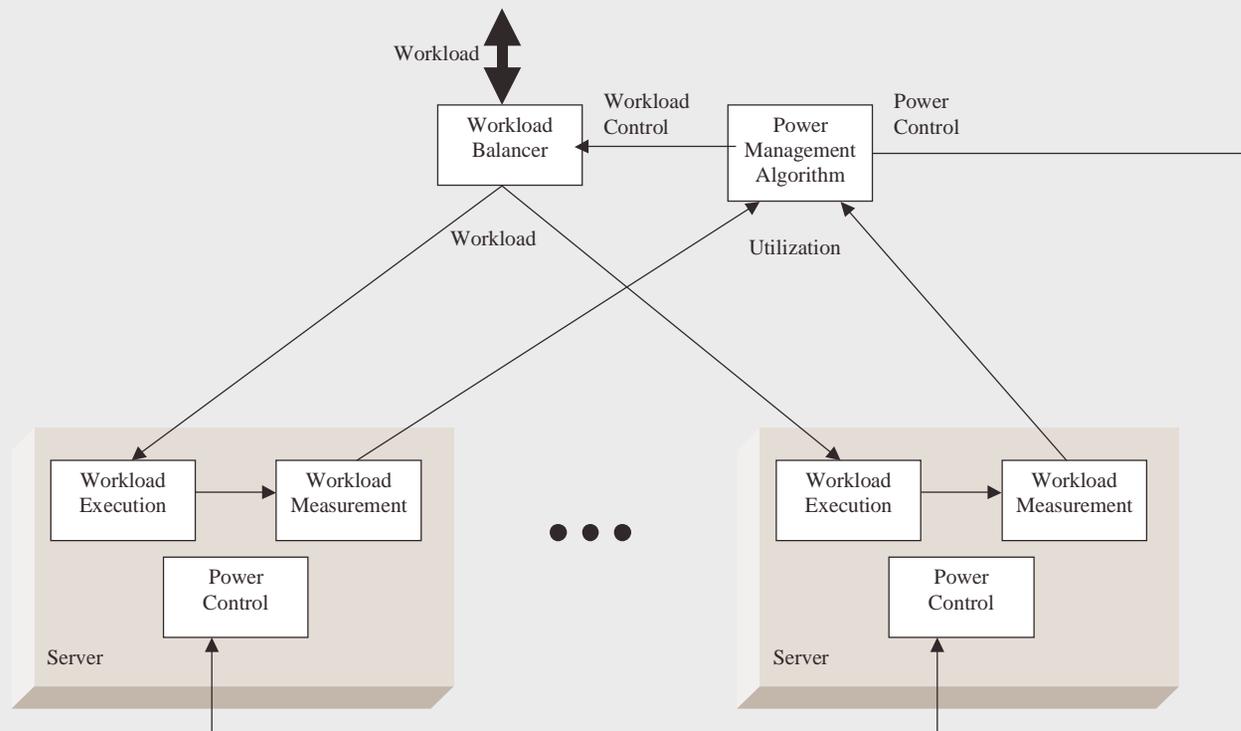
- **Develop proactive self-healing systems**
 - Reduce probability of "Bad Things" due to software aging
 - Detect and predict resource exhaustion
 - Invoke timely corrective action via Software Rejuvenation
 - Resetting of software state to initial level of resource consumption
 - Apply technology to web clustering
 - More info: <https://www.research.ibm.com/journal/rd/452/castelli.html>



Autonomic Web Computing with BladeCenter

Power Management

- Predictive algorithm that measures and predicts workload and determines when to place servers in a low power state
- Objective is to minimize energy consumption, unmet demand, and power cycles
 - Automatically adapts to short term and seasonal workload variations
 - Automatically adapts algorithm "gains" to workload dynamics
 - Energy savings of 20% or more can be achieved
 - More info: <http://www.research.ibm.com/journal/rd/475/bradley.pdf>



Questions?