



The Microsoft Training Solution  
Specialists in SharePoint 2007

Combined**knowledge**



# Scaling out a SharePoint Farm and Configuring Network Load Balancing on the Web Servers

Steve Smith  
Combined Knowledge  
MVP SharePoint Server

 Microsoft®  
**Office SharePoint®  
Server 2007**

# Scaling out a SharePoint Farm and Configuring Network Load Balancing on the Web Servers



Steve Smith – MVP SharePoint Server  
Owner  
Combined Knowledge



# Introduction

- Co Owner of Combined Knowledge Ltd and Combined Knowledge Asia Pacific
  - MVP for SharePoint Server
  - ISPA Lead for the EMEA
  - Co organiser UK SharePoint User Group
    - 5000+ members 😊 <http://suguk.org>
  - Co Author of the SharePoint Administrators Companion
  - Author of many Whitepapers on SharePoint 2007
- [www.combined-knowledge.com.au](http://www.combined-knowledge.com.au)



# Agenda

- What is scaling out in SharePoint 2007
- Why Network Load Balance
- Demo – configuring Microsoft Network Load Balancing
- Demo – Installing new SharePoint Servers into the farm and scaling out the services.
- Q & A ( time allowing )

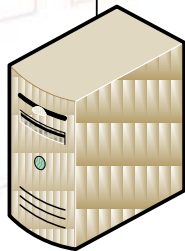
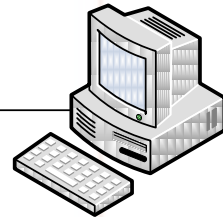
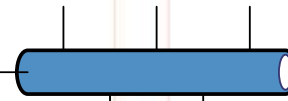
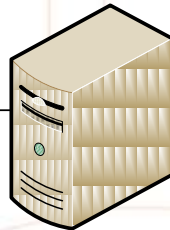
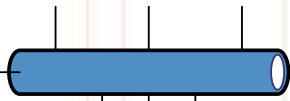
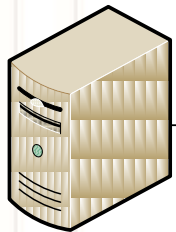
# What is scaling out ?

- Adding additional SharePoint servers into your farm
- Resilience
- Performance
- Isolation
- Architecture change – 32 to 64 bit – You know it's happening so time to start planning for it !!
- £££ \$\$\$ !! ☹

# You are here – Small Farm

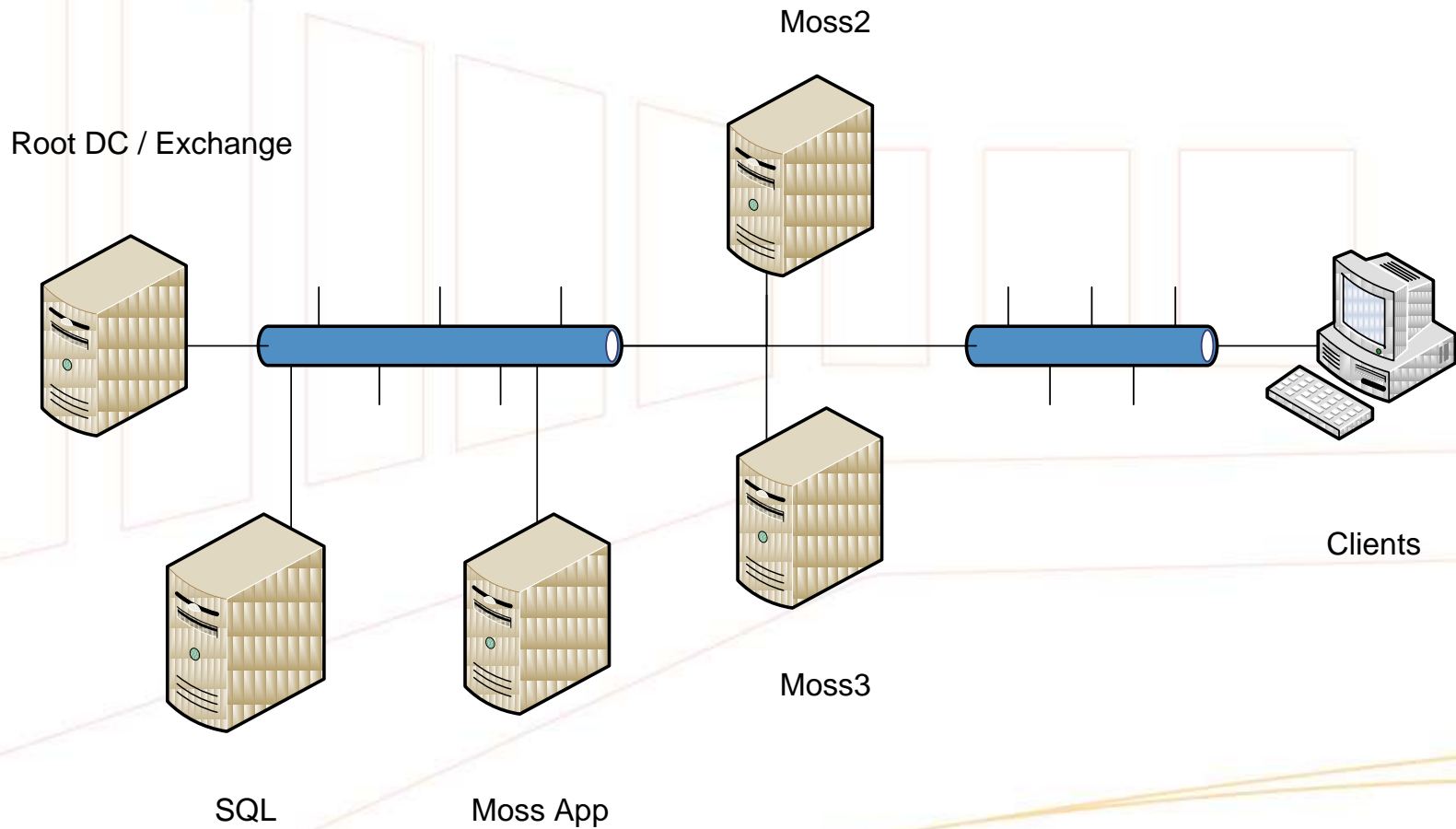
Root DC / Exchange

Moss App



SQL

# Where I want to be



# Why Network Load Balance

- More concurrent users / RPS
- On 32 bit the only way to get past the 2GB limit of available resources. On 64 bit it can mean less front servers. I.E. 2x64 bit front ends against 4x32bit boxes for the same traffic.
- Better page load time
- Caching benefits
- Resilience for the users
- Software or hardware load balancing depending on budget and needs

Demo

# Configure Microsoft Network Load Balancing



 Microsoft®  
**Office SharePoint®**  
Server 2007

Demo

# Adding two new SharePoint Servers to the farm and redistribute the services



# Summary

- It's not just about performance or resilience
- Remember the private Lan between servers
- 64 bit is the best way to keep server numbers down
- Indexing benefits from a dedicated web front ends for crawling
- Use Test tools for what if scenario's and proof of concept

# Questions?

Meet me after the Session

Contact details – [steve@combined-knowledge.com](mailto:steve@combined-knowledge.com)

whitepaper –

<http://www.combined-knowledge.com/Downloads%202007.htm>

# Agenda

- Availability (throughput, scale and latency)
- Suggested limits
- Hardware
- Topology
- Troubleshooting and testing

# Session Goals

- Provide framework for capacity planning
- Highlight top watch-outs and warnings
- Identify and demonstrate available tools for validating capacity decisions in your environment

# Questions Often Heard

- How much hardware do I need?
- Do I need a server farm?
- Do I need SQL Server?
- How much data can I store?
- How many people can I support?
- How many sites can I run on my servers?
- How do I validate it?

# Users and Throughput

- User load: Typical versus Peak
  - Typical = Average requests over standard unit of time (work day)
  - Peak = concurrency (versus user type); plan for peak
- Usage profile: user's behavior
  - Distribution of requests across content
  - Mine your IIS Logs
- Usage Rule of Thumb; assuming 10% concurrency

- 1 RPS =

Profile	Expected Rate (RPH)	Concurrent users	Total users
Light	20	180	1,800
<b>Typical</b>	<b>36</b>	<b>100</b>	<b>1,000</b>
Heavy	60	60	600
Extreme	120	30	300



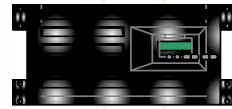
\* Note: These are consistent with WSSv2 and SPS 2003

# Customization affecting performance

- Event handlers
- Workflows
- Web Parts
  - Badly coded
  - Too many SQL round trips
  - Content by Query limited to 1,000 lists by default – fails if more.
- Field Controls
- Authentication and Role Providers
  - Anonymous - Fastest
  - Kerberos
  - NTLM
  - Basic
  - Forms - Slowest
- iFilters and Protocol handlers

# Growth and Capacity Planning

## Dedicated Crawl Target Front-end



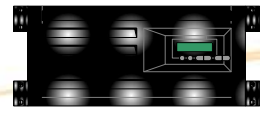
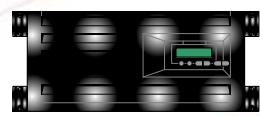
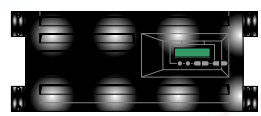
Client traffic

Indexing

NLB Cluster1: NLB enabled

NLB Cluster1: NLB Stopped

Windows SharePoint Services



# Latency

- Latency components

- Server processing
  - SQL processing, # SQL round trips, AJAX processing, security trimming
- Client processing
  - JavaScript, CSS, AJAX requests, HTML load, Client machine specs
- Wire transfer
  - Bandwidth, size of download

- Recommendations

- #1 killer of latency = custom web parts / field controls
  - Watch for: SQL round trips, unnecessary data, excessive client side script
- Re-use existing client code versus adding more
- Design code for performance – e.g. HTML table definitions
- Profile your solutions

# Data – Scale

- How many objects?
  - Infrastructure: Portal sites, team sites, personal sites, document libraries, etc.
  - Data: documents, lists, profiles, etc.
- Recommendations
  - Carefully plan your site hierarchy and deployment
    - Minimize # web applications and application pools
    - Limit # of shared service providers
    - Plan for database growth
  - Follow data and feature best practices and suggested limits

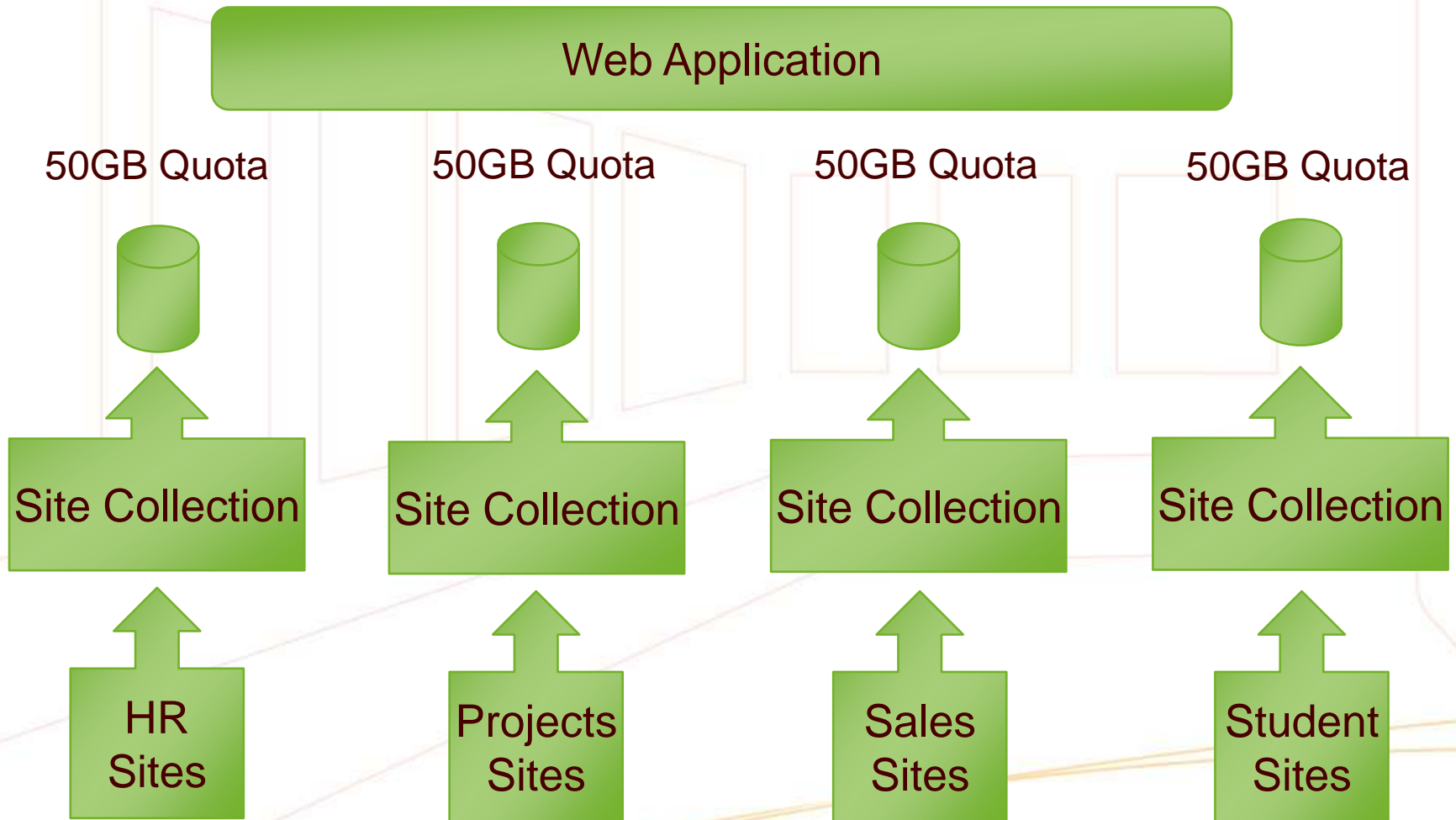
# Data – Suggested Limits ( not hard coded )

<b>Object</b>	<b>Scope</b>	<b>Guideline</b>
<b>Site collections</b>	<b>Database</b>	<b>50,000</b>
<b>Web sites</b>	<b>Site collection</b>	<b>250,000</b>
<b>(sub) Web sites</b>	<b>Web site</b>	<b>2,000</b>
<b>Lists</b>	<b>Web site</b>	<b>2,000</b>
<b>Items</b>	<b>List</b>	<b>10 M</b>
<b>Documents</b>	<b>Doc Library</b>	<b>2 M</b>
<b>Documents</b>	<b>Folder</b>	<b>2,000</b>
<b>Document size</b>	<b>File</b>	<b>2 GB</b>
<b>Indexed Documents (MOSS)</b>	<b>SSP</b>	<b>50 M</b>
<b>Search Scopes (MOSS)</b>	<b>Site Collection</b>	<b>1,000</b>
<b># Profiles (MOSS)</b>	<b>SSP</b>	<b>5 M</b>

# Physical Data Storage & Site Collections

- What about Content DBs?
  - Site Collections cannot span Content DBs
- But what if I want to keep my DBs manageable? (75 - 100GB)
  - Then you need multiple Site Collections
- But aren't there features that don't work across Site Collections?
  - Content Types
  - Shared Columns
  - Content by Query Web Part
  - Workflow
  - Information Management & Retention Policies
  - Document Link Fix Up
  - WSS v3 Search cannot span Site Collections
  - Quotas are at the Site Collection Level

# Database file per site collection



# Data – Storage Requirements

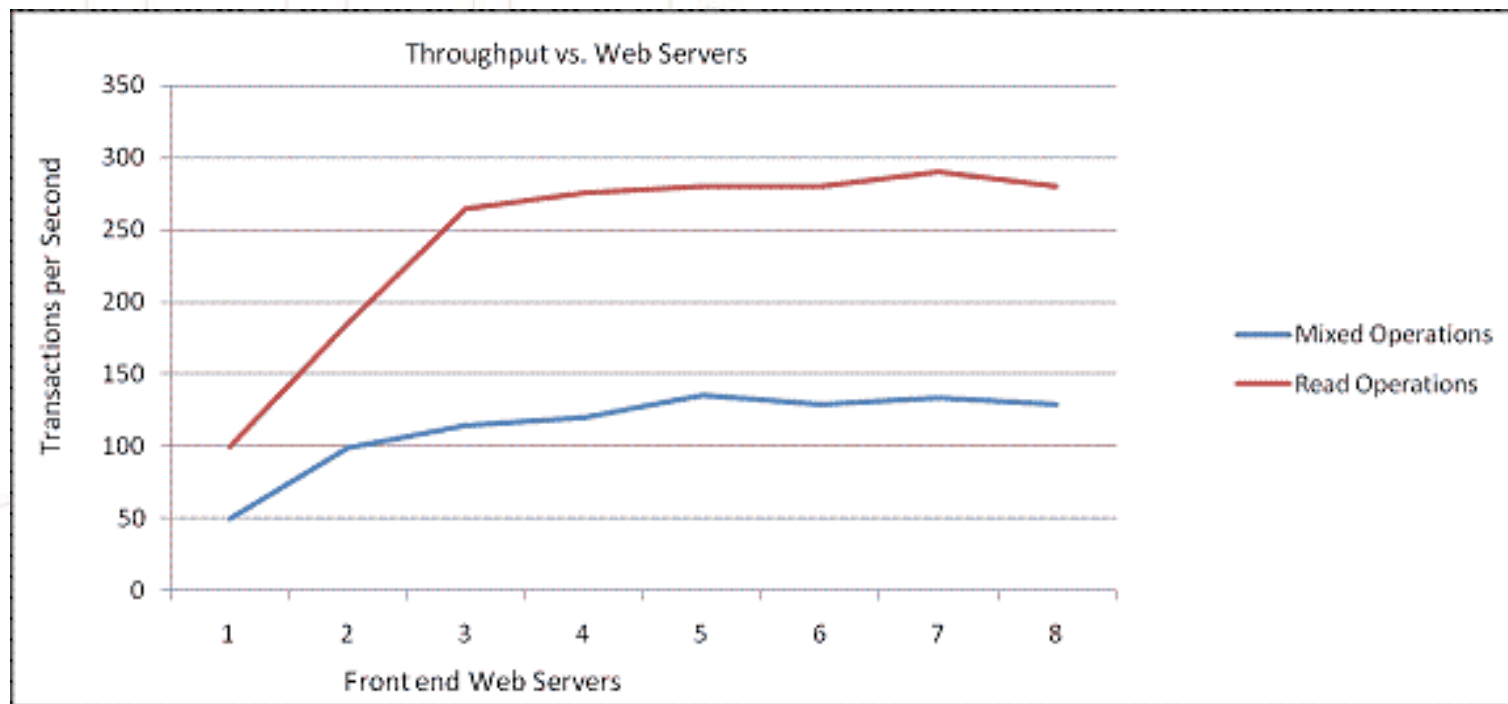
- Primary metric: Document storage
  - Plan for 1.2 – 1.5 x file system size for SQL Server  
**note:** metric is closely tied to RAID level used on SQL disks
- Secondary metric = Index
  - Index server: 10-15% of total size of all content indexed for single server
  - Query server: *same* size as Index server

# Hardware – How SharePoint Scales?

- Designed to grow with organization needs
  - Server resources: x32, x64, CPU, RAM, HDD
    - Recommend 64 bit for back end services which can leverage additional addressable memory >4GB.
    - SQL - Index – Excel – Search – Web
    - Mixed 64 / 32 bit farms supported – Do not Mix Roles. All Webs 64 Bit
    - SQL: HDD configuration critical
  - Server Farm
    - Topology restrictions removed
    - WFE, query, index, excel calc, project, SQL
  - Shared Services
    - On by default (up to 20/farm)
  - Adopted WSS adage: content only limited by HW capability\*
    - Sites: Portals are "just another site"

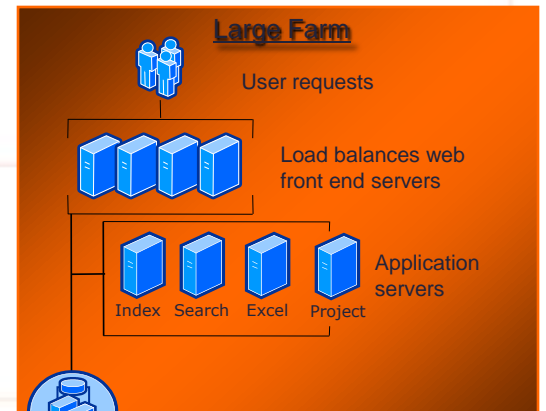
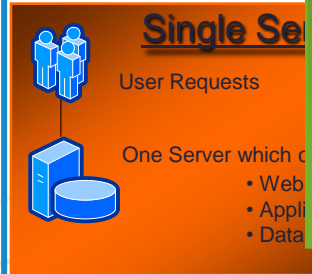
\* see limits and storing content

# Hardware vs. Throughput



# Picking Your Topology

Performance



## Medium Farm

- Strengths**
- Scalability
  - High user load capability
  - High availability & reliability
  - High performance
  - High availability & reliability
  - High performance
  - High availability & reliability
  - High performance
- Limitations**
- High cost
  - High user load
  - High performance
  - High availability & reliability
  - High performance
  - High availability & reliability
  - High performance
  - High availability & reliability
  - High performance

Availability

# Hardware – Single server



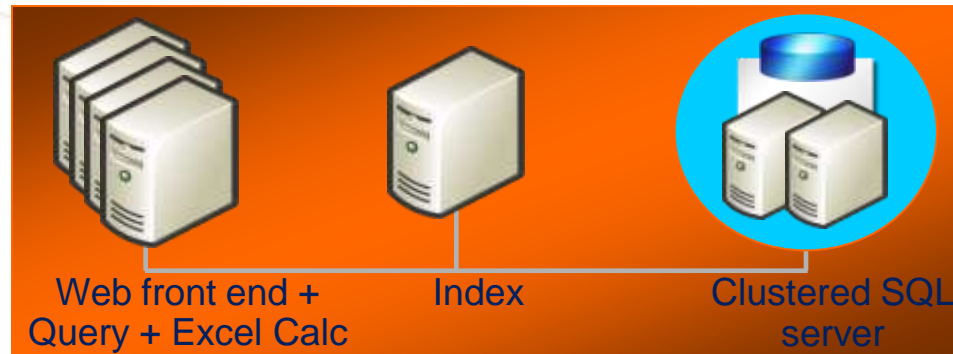
One Server containing

- Web front end
- Application
- Database

Server type	RAM	HDD	CPU
Single server	2 GB	100 GB	1 x 2.8 Ghz Pentium-4 (32bit)

- SQL Express appropriate for up to 500 (typical) users
- SQL appropriate up to 5,000 (typical) users
  - Host: 1,000 team, portal, and MySites
  - Store: 10,000 docs
  - Index: 100,000 docs (11 docs/sec)
  - 10 rps using “common” operations

# Hardware – 4x1x1 farm



Server type	RAM	HDD	CPU
Front end servers	2 GB	200 GB	2 x 2.8 Ghz AMD 64bit
Index server	4 GB	200 GB	2 x 2.8 Ghz AMD 64bit
SQL Server computer	4 GB	200 GB	4 x 2.8 Ghz, dual core, AMD 64bit

- Highly available
  - Users: 100,000s of users
  - Host: 10,000s of team, portal and MySites
  - Store: 1,000,000s of documents
  - Index: 1,000,000s of documents

# "Right" Sizing Your Installation

- Farm rules of thumb:\*

Farm Type	# Users	Comments
Single server (SQL Express)	≤ 500	Does not meet "high availability" bar
Single server (SQL)	≤ 5,000	Does not meet "high availability" bar
Medium farm (2 x 1 x 2)	≤ 100,000	No single point of failure
Large farm (4 x 3 x 2)	≤ 500,000	No single point of failure

\* # users vary depending on usage profiles, operational mix, and hardware

# Caching

Use this type of caching...	At the...	Notes
Output caching and cache profiles	Individual page level	Ideal for heavily accessed Web sites that do not need to present new content frequently.
Object caching	Individual Web Part control, field control, and content level Set at Site collection , Site and page level.	Includes cross-list query caching and navigation caching 100MB by default
Disk-based caching for Binary Large Objects (BLOBs)	Individual BLOB level and caches images, sound, movies, and code Configured per Web App	Supports .gif, .jpg, .js, .css, and other image, sound, and code files that are stored as binary large objects 1GB Space used minimum

# Testing Server Load

demo

# Troubleshooting

- Poor throughput
  - SQL – use SQL best practices for perf; especially disk perf
  - Asynchronous operations and timer job conflicts
  - Resource contention: # web apps, app pools, DBs, etc.
  - Check ANY custom code for SQL round trips and payload
  - Look for mis-configured network components (NICs, routers, etc.)
- Poor end user perceived latency
  - Page payload
    - OOB 1st page download ~ 200 KB; should not be significantly higher
    - Use page compression where possible
  - Caching strategy: Turn on BLOB and Output caching whenever possible
  - Client machine specs
  - Mis-configured network

# Call To Action

## Do

- Use the Framework
  - Use the rules to right-size your deployment
  - Use the tools to validate your environment
- Use the Documentation
  - Capacity planning - TechNet
  - High availability deployment

## Be Aware

- Capacity Planning whitepaper will be out shortly!

# Q & A