System Architecture Design and Capacity Planning

Dave Peters

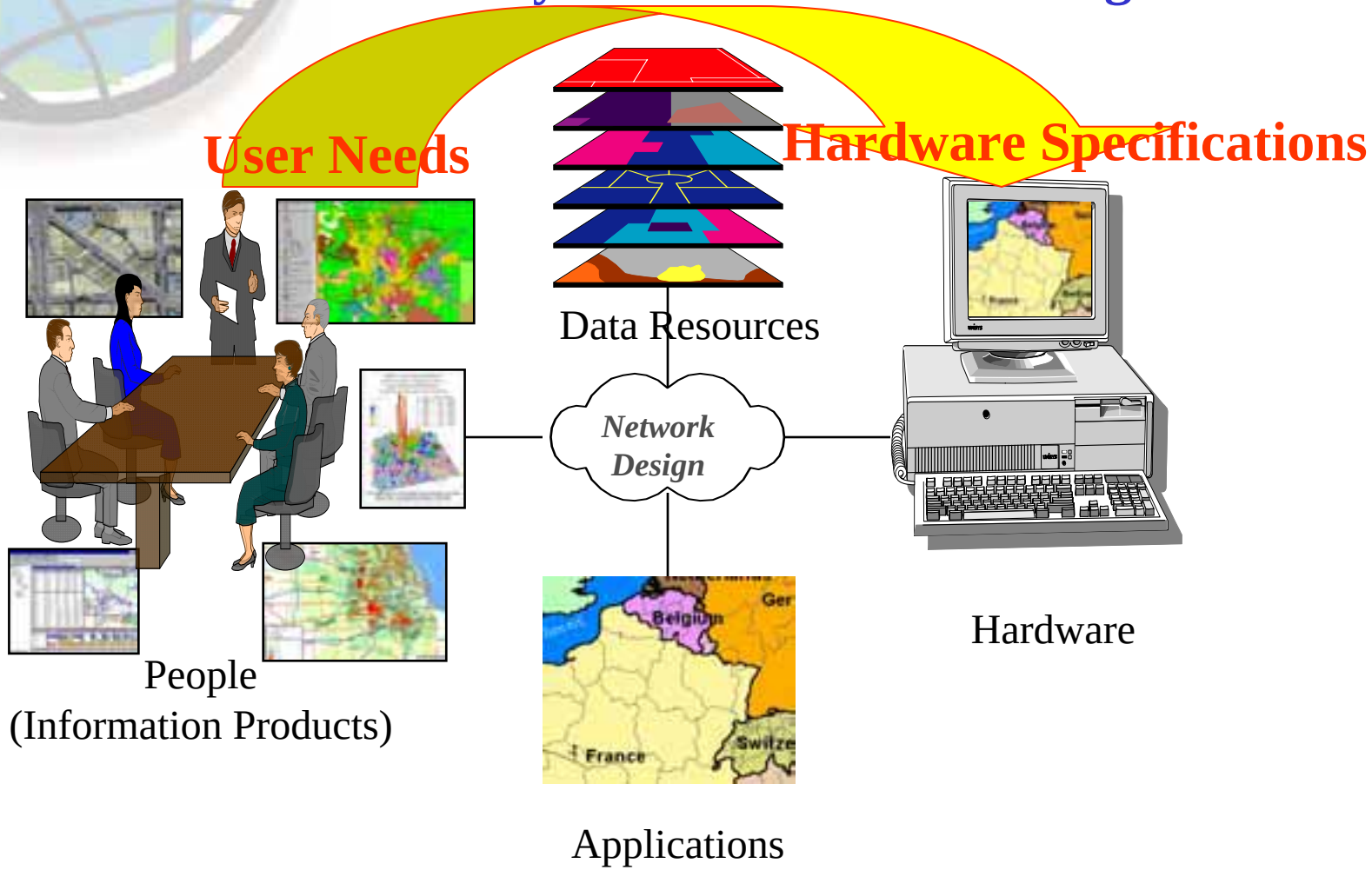
ESRI – Redlands, CA

System Architecture Design and Capacity Planning

Dave Peters
esri.com/systemdesign

System Architecture Design

What Is System Architecture Design?



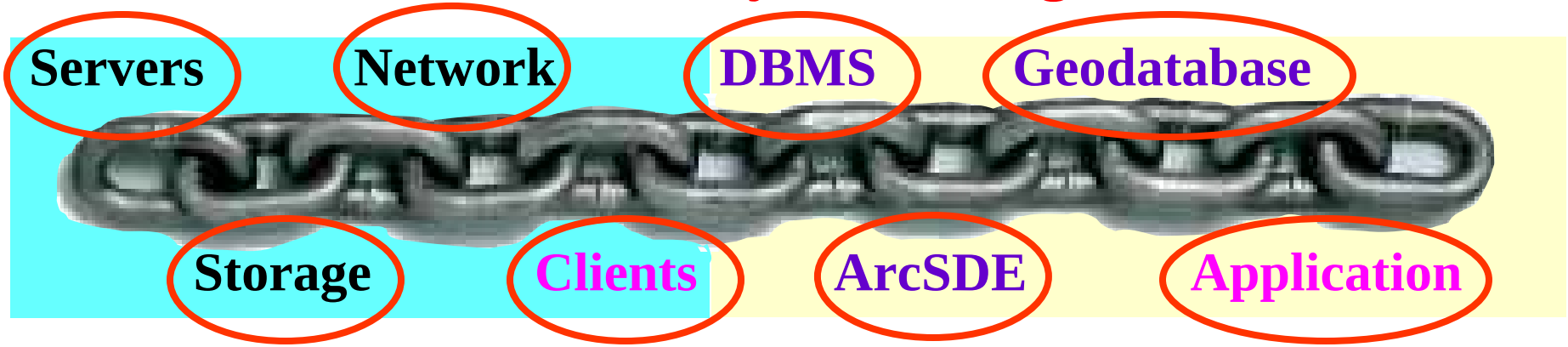
Why is System Architecture Design Important?

Performance Chain

Reduce Cost

Improve Productivity

Balanced System Design



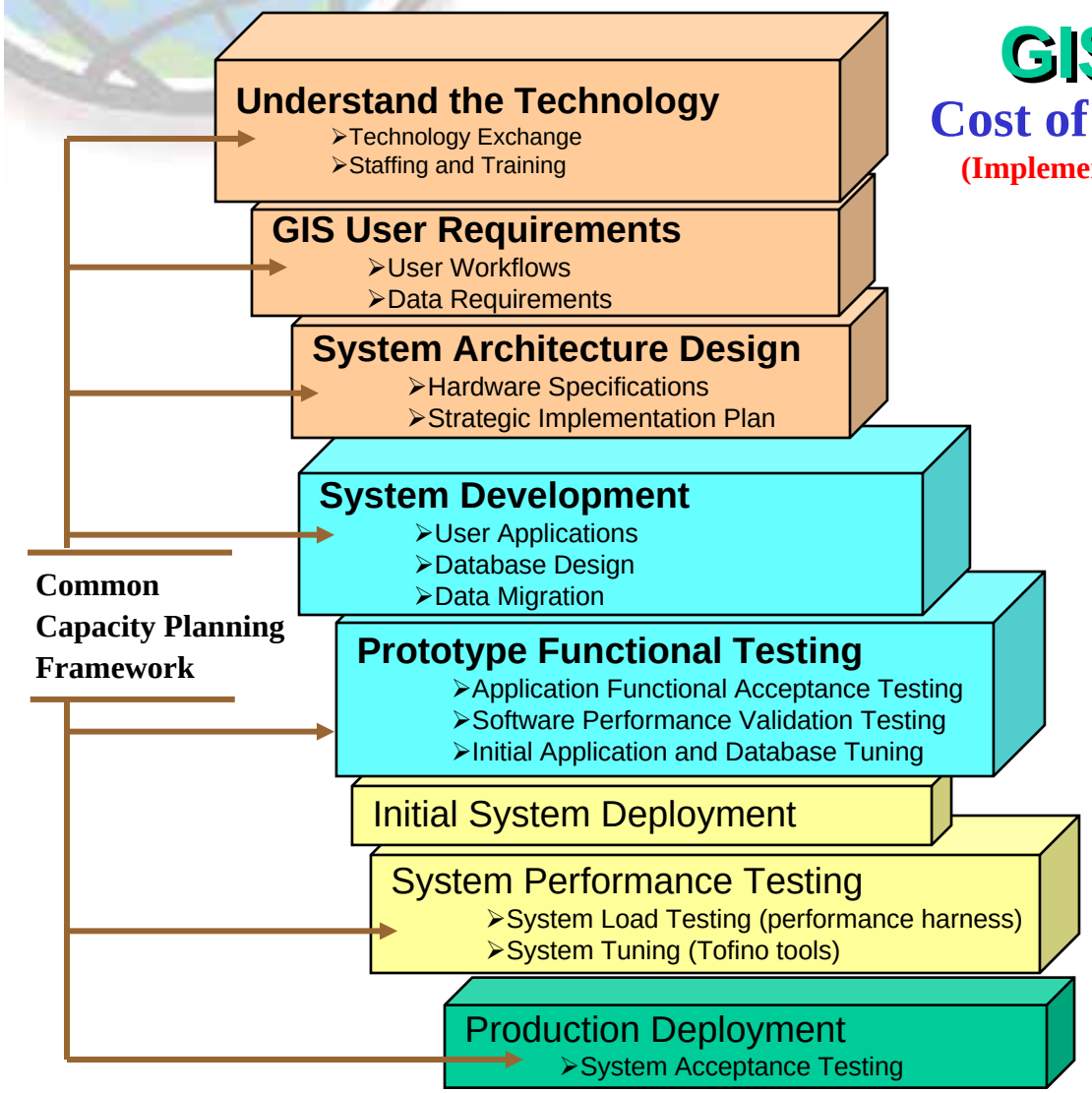
Hardware Infrastructure **Database Design** **User Workflow**

System Architecture Design

Framework for Productive Operations

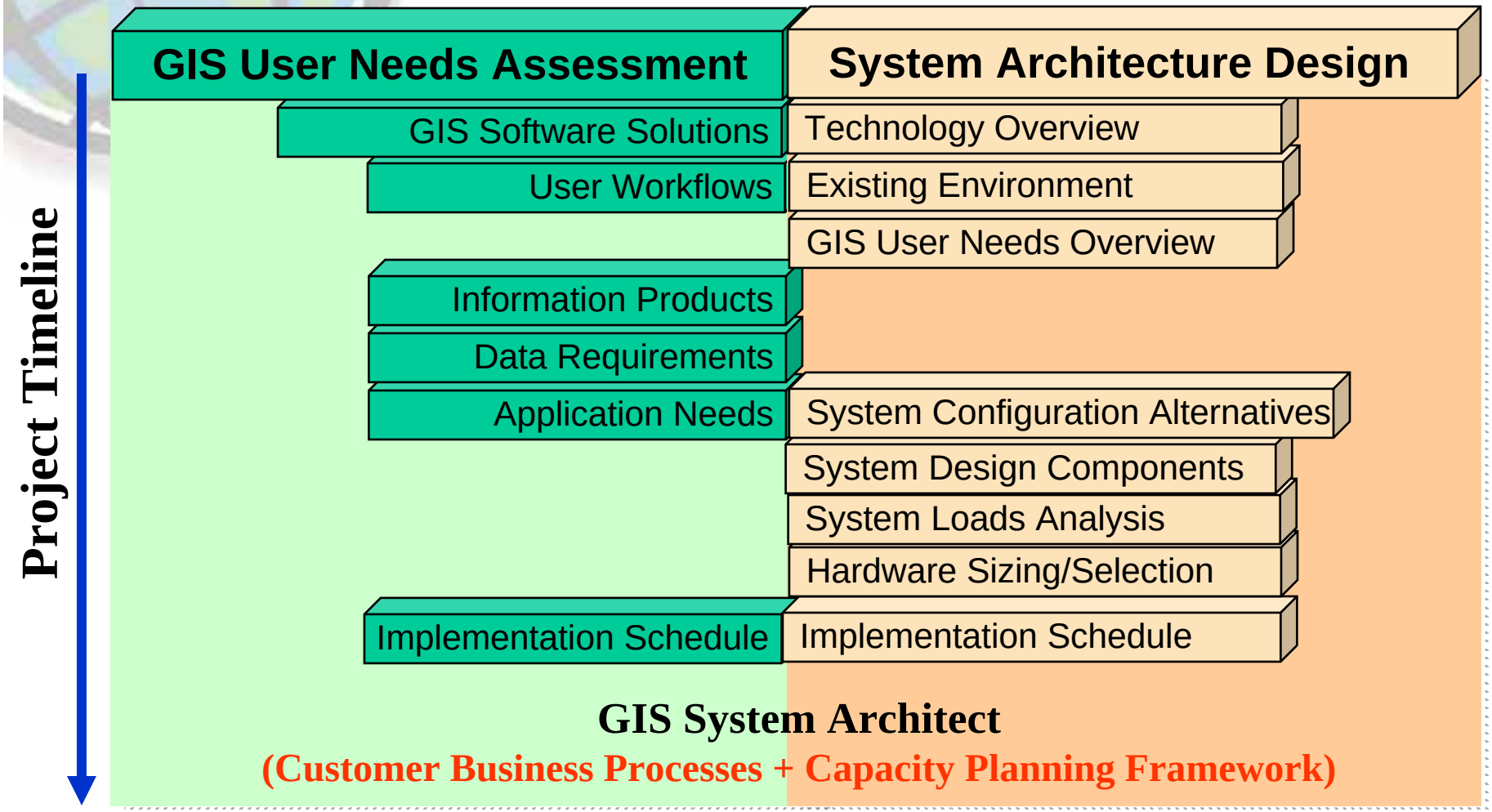
System Implementation Stages

How can we manage implementation risk?



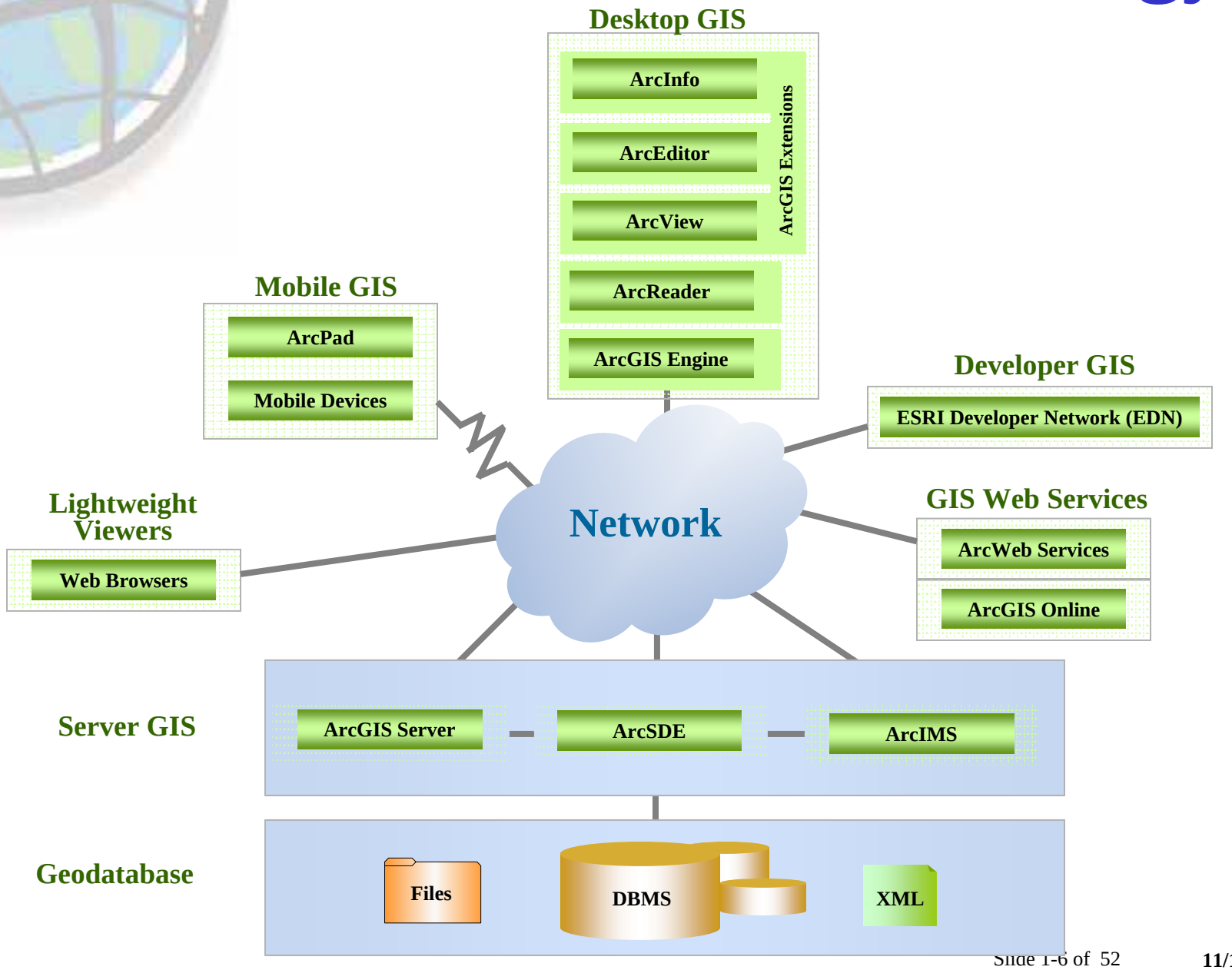
(Moody)
Database Programming & Design
October 1996, p57–64

System Design Process



Integrated Business Needs Assessment

ArcGIS Software Technology



ArcGIS Desktop

Scalable Licensing

ArcInfo is the most complete GIS available. It includes all the functionality of ArcView and ArcEditor and adds **advanced spatial analysis**, **data manipulation**, and **high-end cartography** tools. Professional GIS users use ArcInfo for all aspects of **data building**, **modeling**, **analysis**, and **map display**.

ArcEditor is a powerful GIS desktop system for **editing** and **managing** geographic data. ArcEditor is a member of the ArcGIS family of GIS products and includes all the functionality of ArcView and adds a comprehensive set of tools to **create**, **edit**, and ensure the **quality** of your data.

ArcView is full-featured GIS software for **visualizing**, **analyzing**, **creating**, and **managing** data with a geographic component. Most data has a component that can be tied to a place: an address, postal code, global positioning system location, census block, city, region, country, or other location. ArcView allows you to visualize, explore, and analyze this data, revealing patterns, relationships, and trends that are not readily apparent in databases, spreadsheets, or statistical packages.

Concurrent User Licensing
(Client Desktop and Windows Terminal Server)
- FlexLM License Manager

Single Use Licensing
(Client Desktop Only)
- Software License Key

ArcReader is a free, easy-to-use mapping application that allows users to view, explore, and print maps and globes.

ArcGIS Server

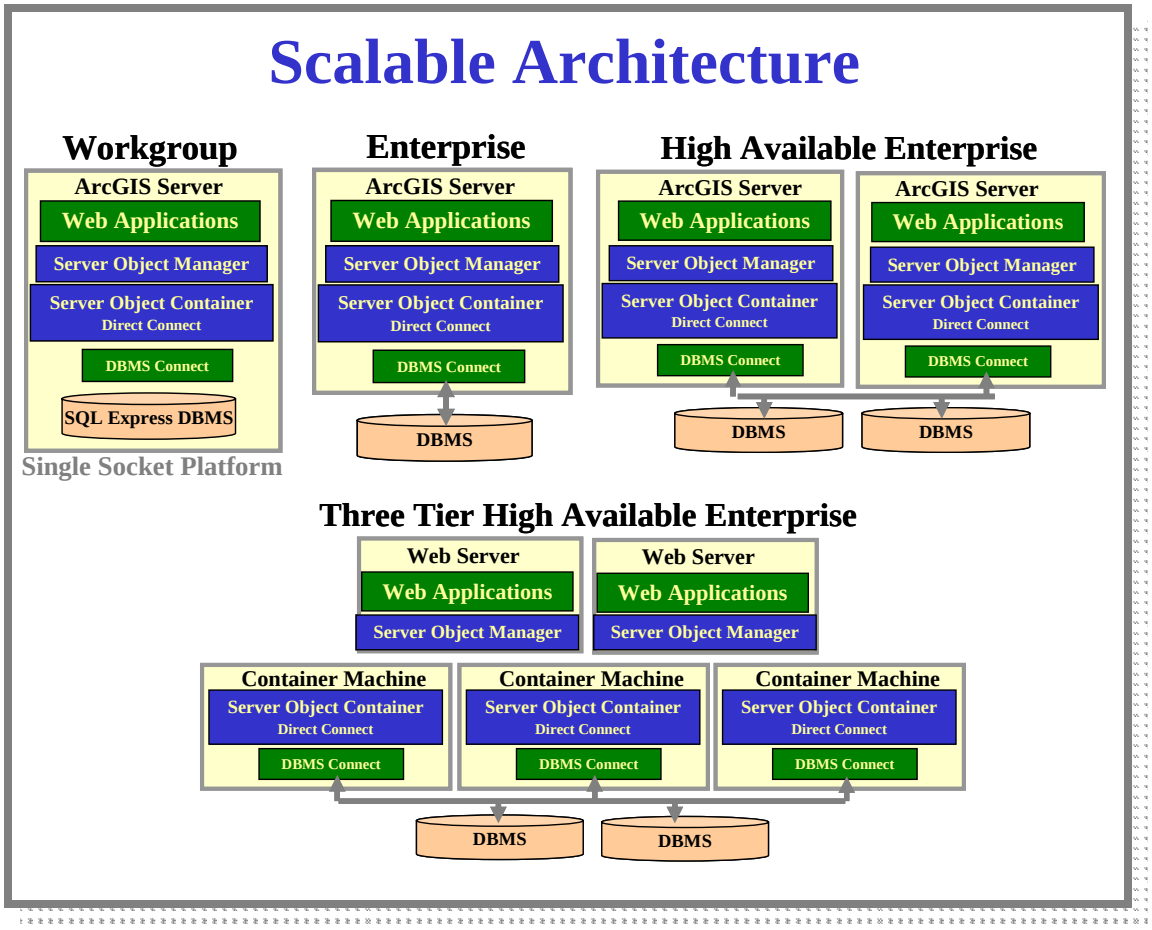
Scalable Licensing

Advanced
- Web Editing
- Mobile ADF Clients
- Advanced Geoprocessing
- Advanced Extensions

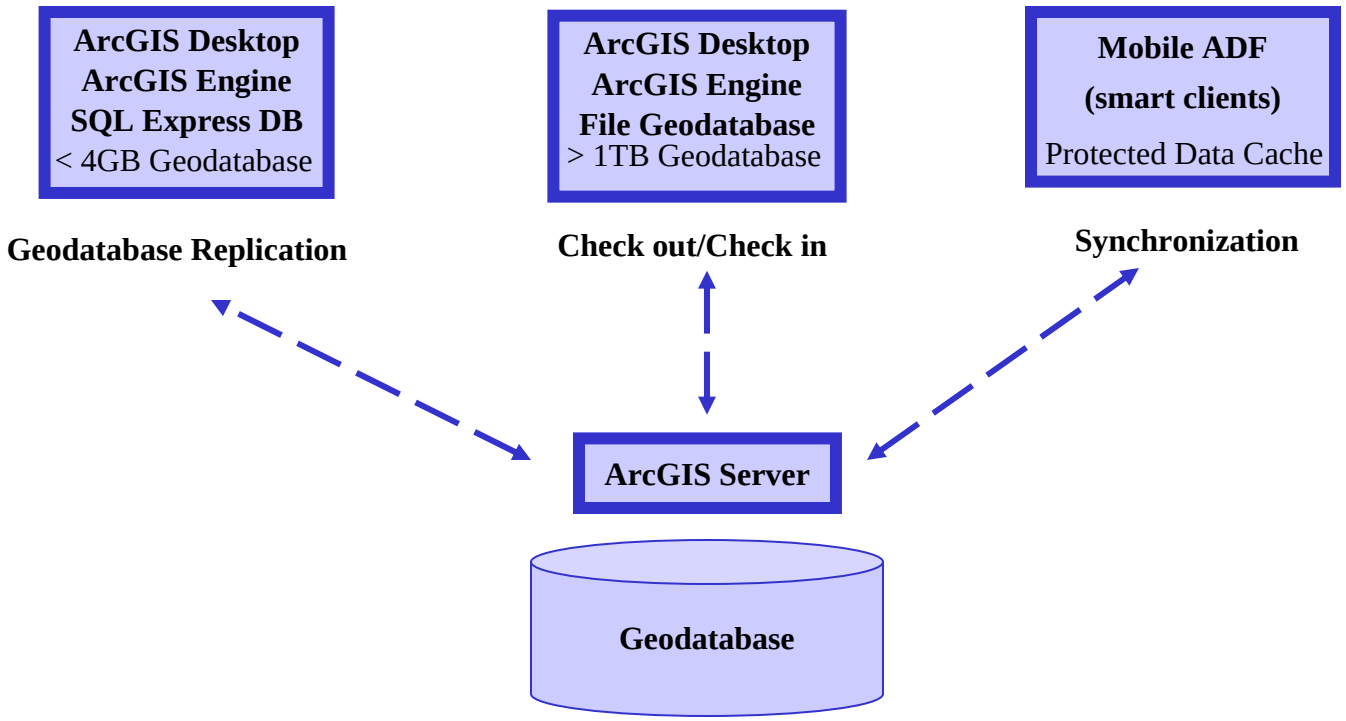
Standard
- Standard Map Publishing
- Globe Server (ArcGIS Explorer)
- Standard Geoprocessing
- Standard Extensions

Basic
- Geodatabase Management (SDE)
- Check in/Check out
- Geodatabase Replication

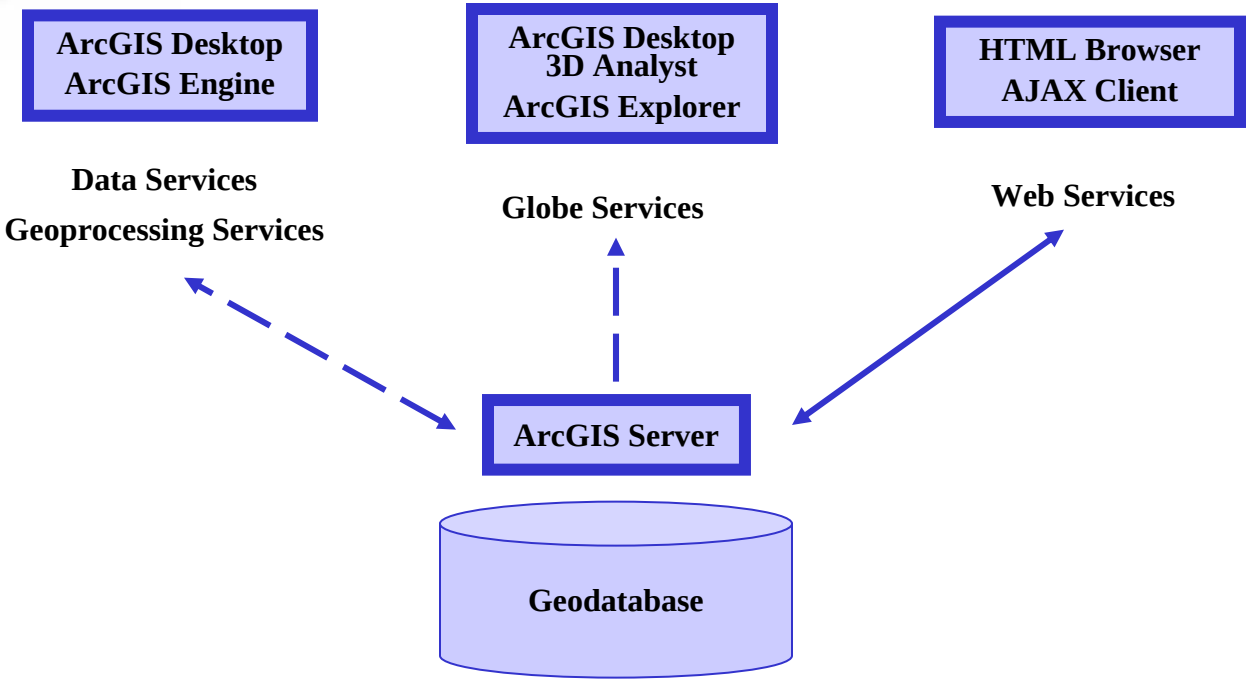
Scalable Architecture



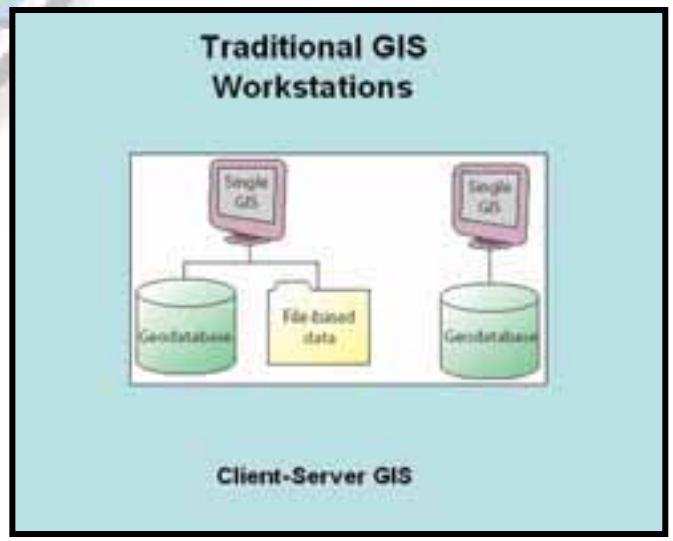
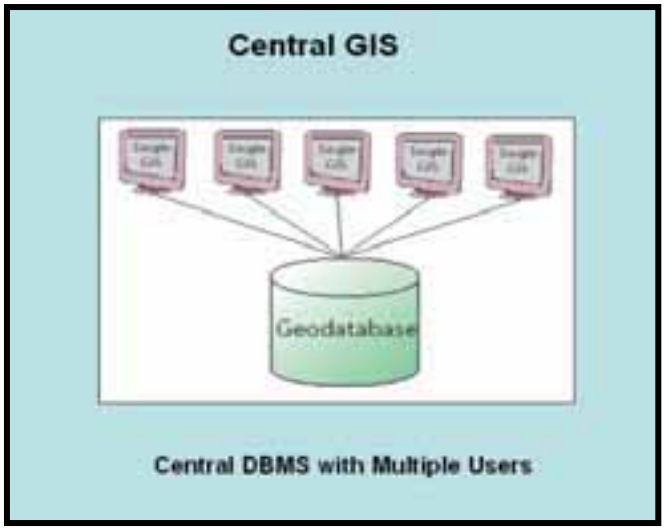
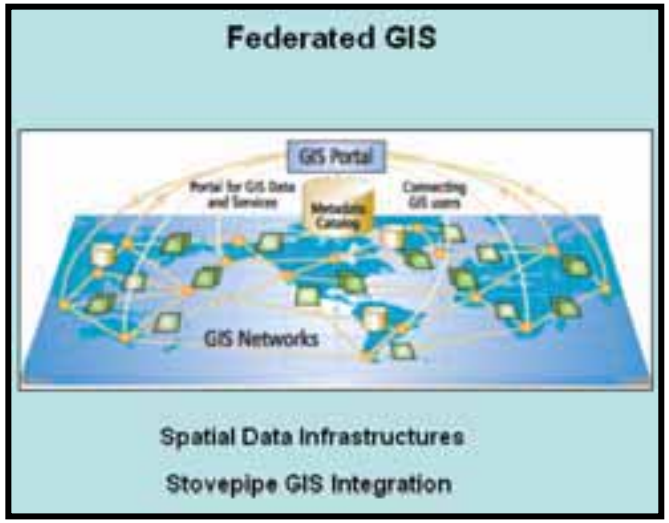
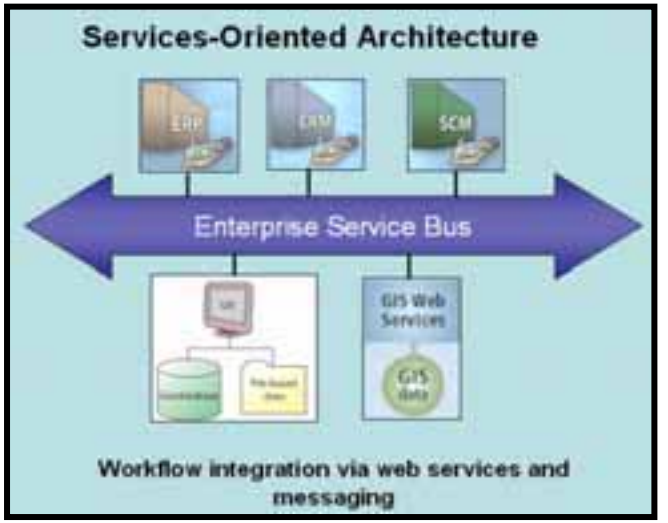
Mobile Operations



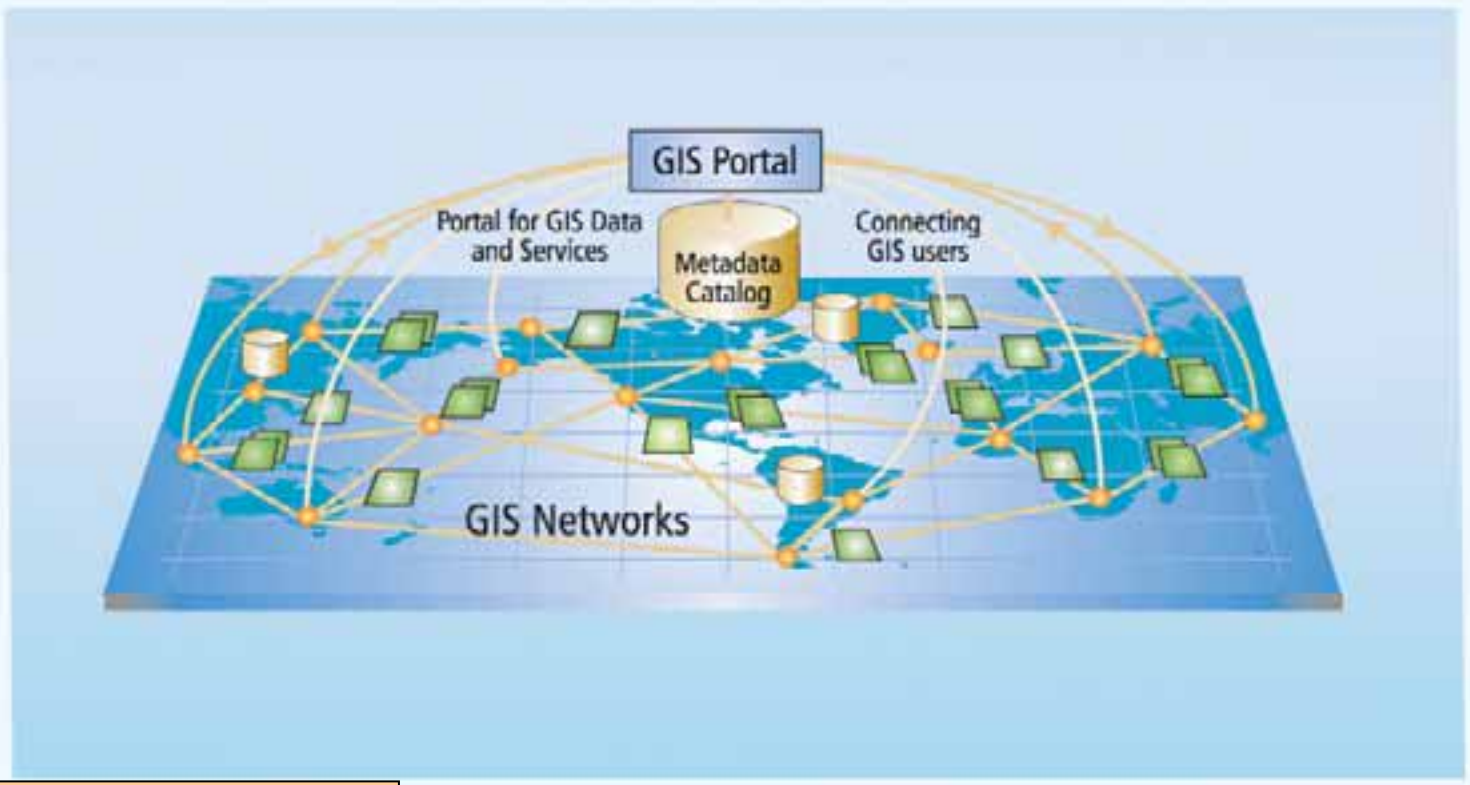
Web Operations



GIS Is Deployed In Many Ways

	Departmental	Enterprise
Traditional		
Emerging		

Federated Architecture



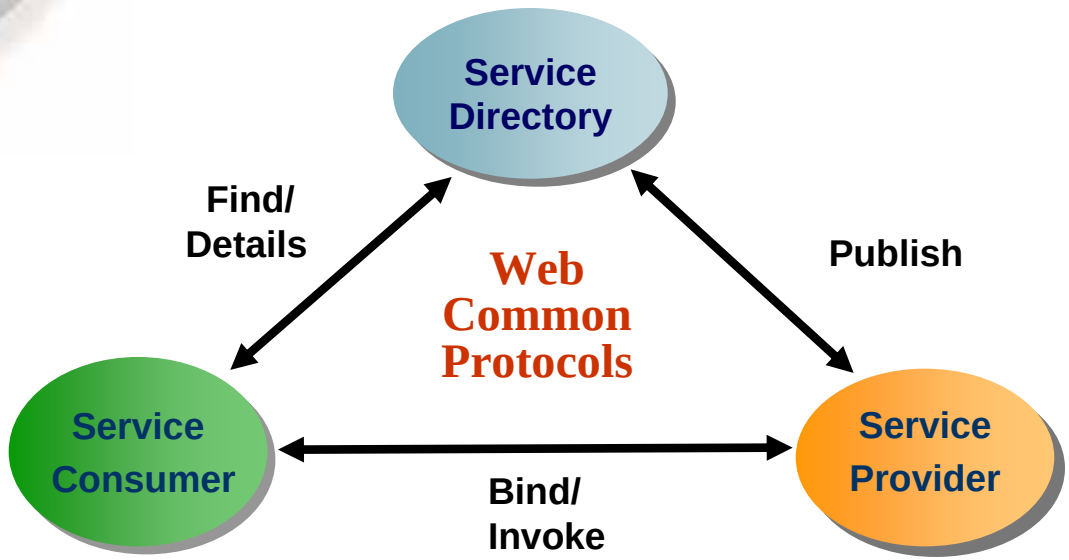
Homeland Security
- **ETL Data Flows**
- **Geodatabase Replication**

Spatial Data Infrastructure

Heterogeneous GIS Integration

Portal Toolkit
- **Map Viewer**
- **Metadata Catalog**

Service-Oriented Architecture



An approach for building distributed computing systems based on encapsulating business functions as services that can be easily accessed in a loosely coupled fashion

Why SOA?

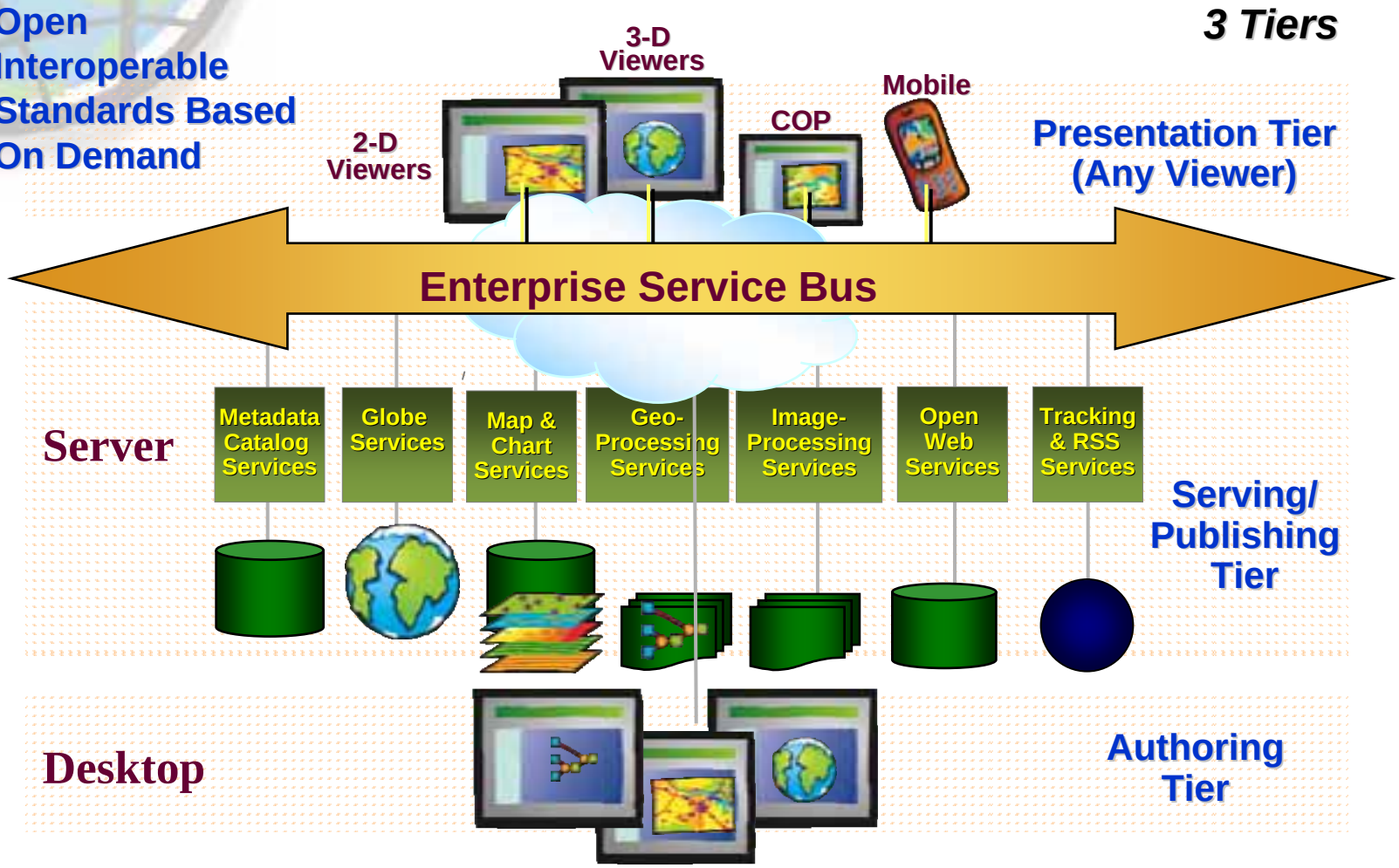
- Technology Change ➡ Component Architecture
- Business Continuance ➡ Reduce Vendor Dependence
- Leverage Investments ➡ Reusable Components
- Customer Flexibility ➡ More Vendor Choices
- Business Integration ➡ Open System Communications

SOA supports more adaptive enterprise operations

GIS in a Services Oriented Architecture

Providing A Framework For Integration

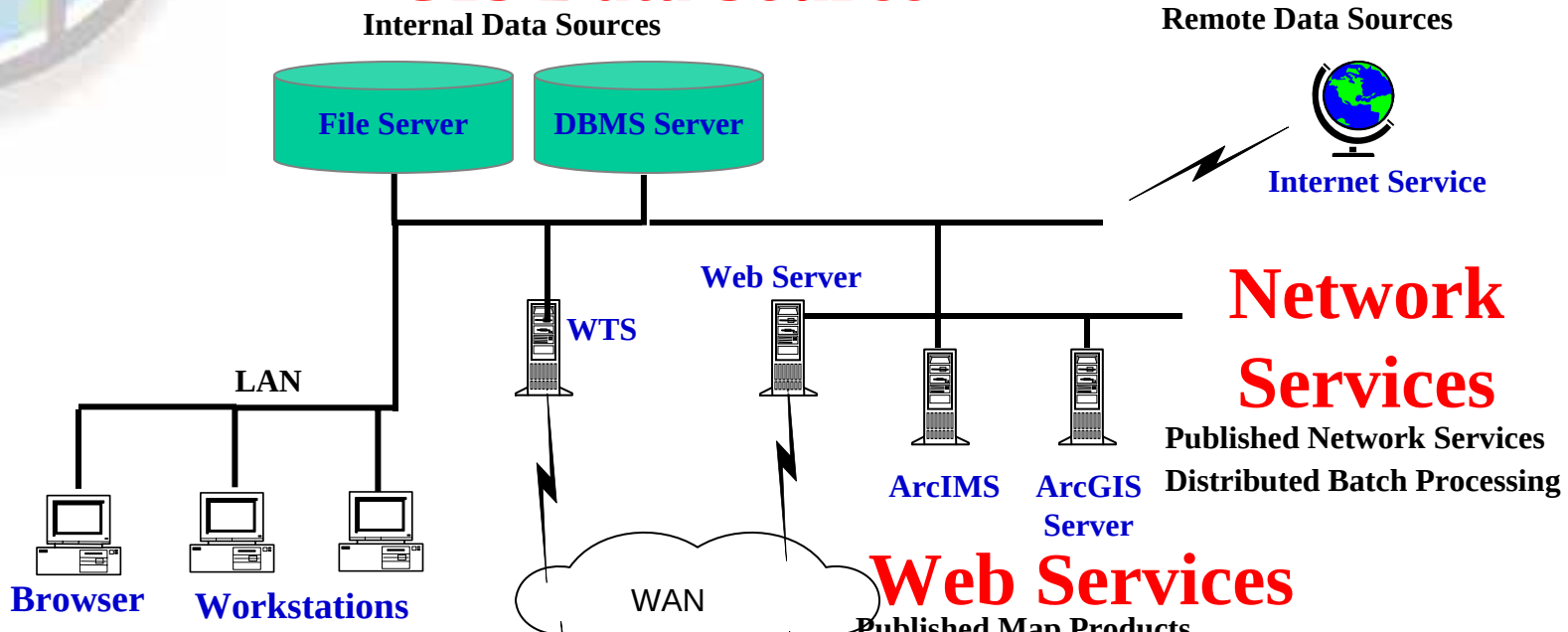
- Open
- Interoperable
- Standards Based
- On Demand



. . . Standard IT Infrastructure

Software Technology Selection

GIS Data Source



Network Services

Published Network Services
Distributed Batch Processing

Web Services

Published Map Products
Published Geoprocessing Services
Integrated Business Workflows

Desktop Applications

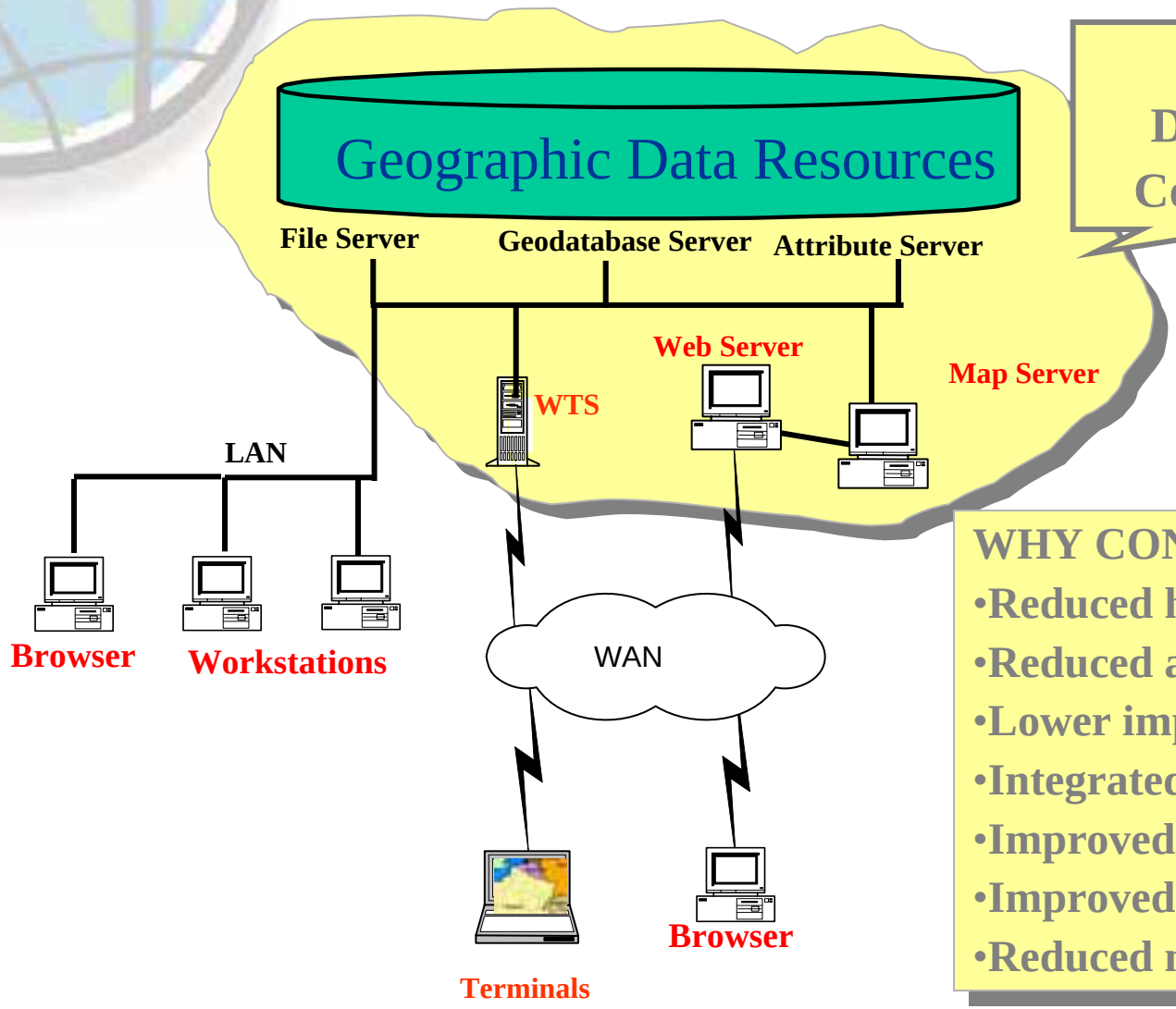
Professional GIS Users
Database Maintenance
Heavy Business Workflows
GIS Project Efforts
Data Conversion

Mobile Applications

Disconnected Operations
Loosely Connected Workflows

Database Configuration Options

Centralized Computing Environment



Central
Data Center
Consolidation

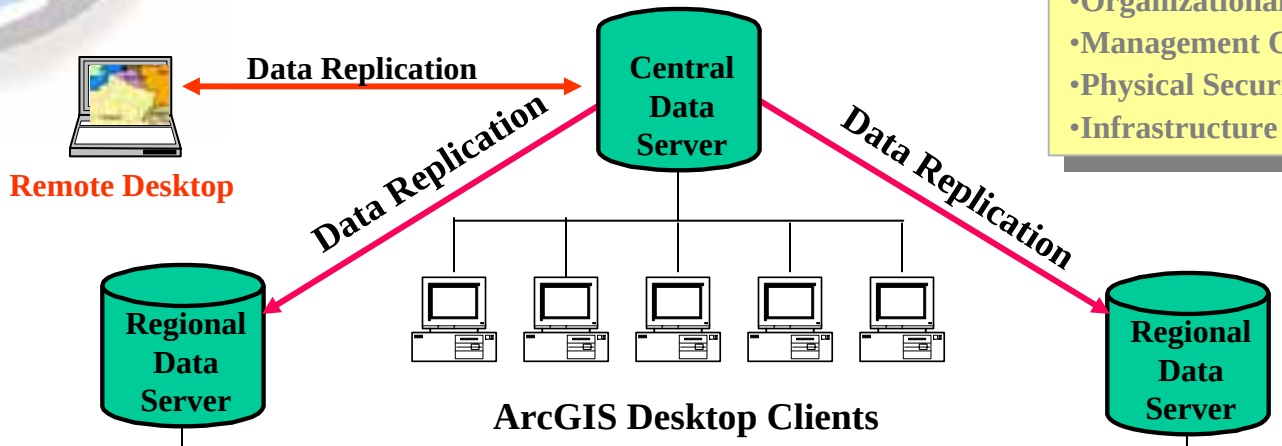
WHY CONSOLIDATE?

- Reduced hardware cost
- Reduced administration cost
- Lower implementation risk
- Integrated operations
- Improved data access
- Improved security
- Reduced network traffic

Database Configuration Options

Distributed Computing Environment

Why use distributed architecture?
- Organizational Precedence
- Management Confidence
- Physical Security
- Infrastructure Limitations

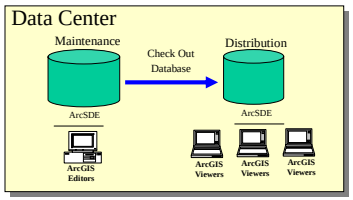


Geodatabase Single- generation Replication
(Data Checkout – ArcGIS 9.1)

ArcGIS Desktop Clients

Peer to Peer – Database Checkout

One Direction Checkout/Refresh Transaction

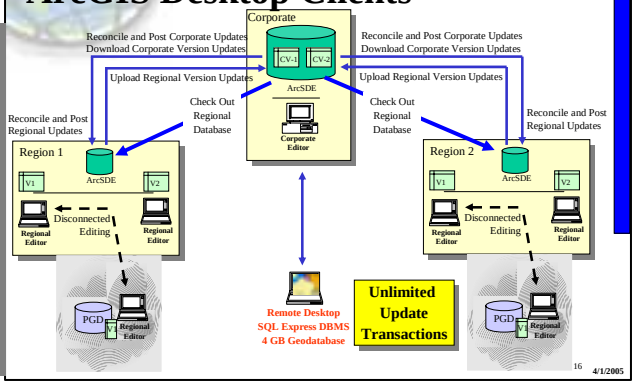


ArcGIS 9.2 → Incremental Updates

DISTRIBUTED DATA SOLUTIONS
- Increased hardware cost
- Increased administration cost
- Higher implementation risk
- Disconnected operations
- Limited data access
- Reduced security
- Increased network traffic

Geodatabase Multi-generation Replication
(ArcGIS 9.2)

ArcGIS Desktop Clients



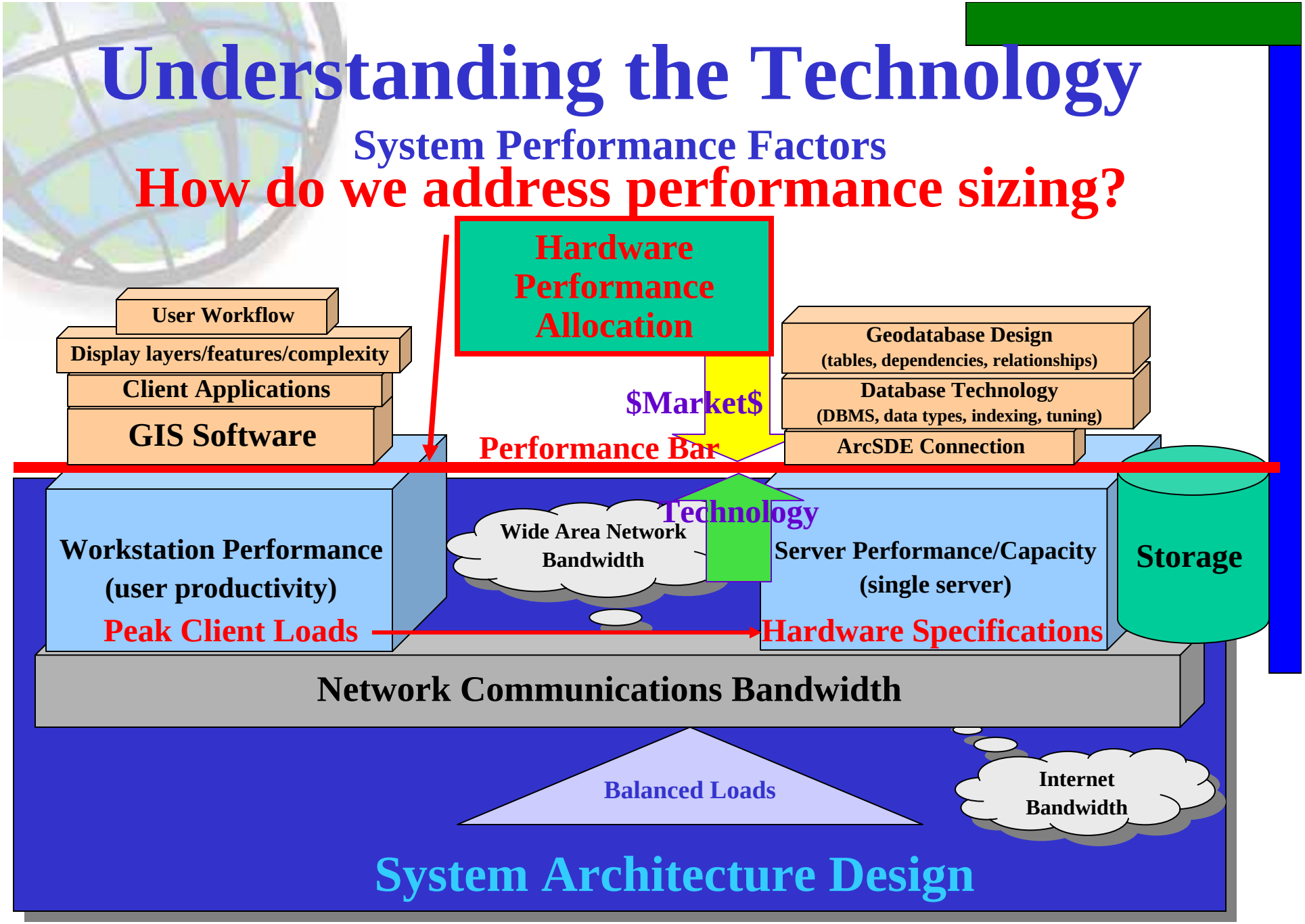
System Design Strategies

Capacity Planning

Understanding the Technology

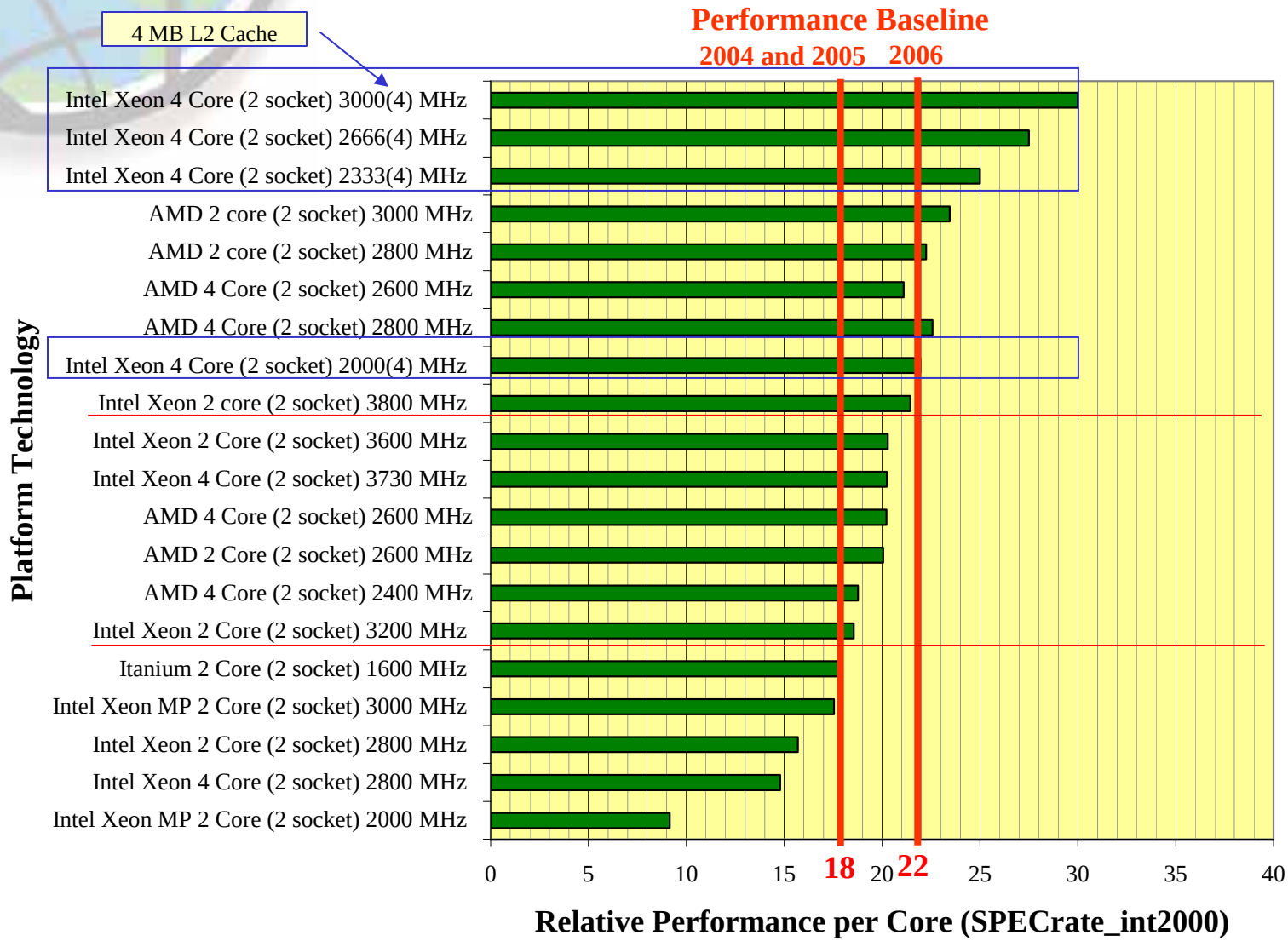
System Performance Factors

How do we address performance sizing?



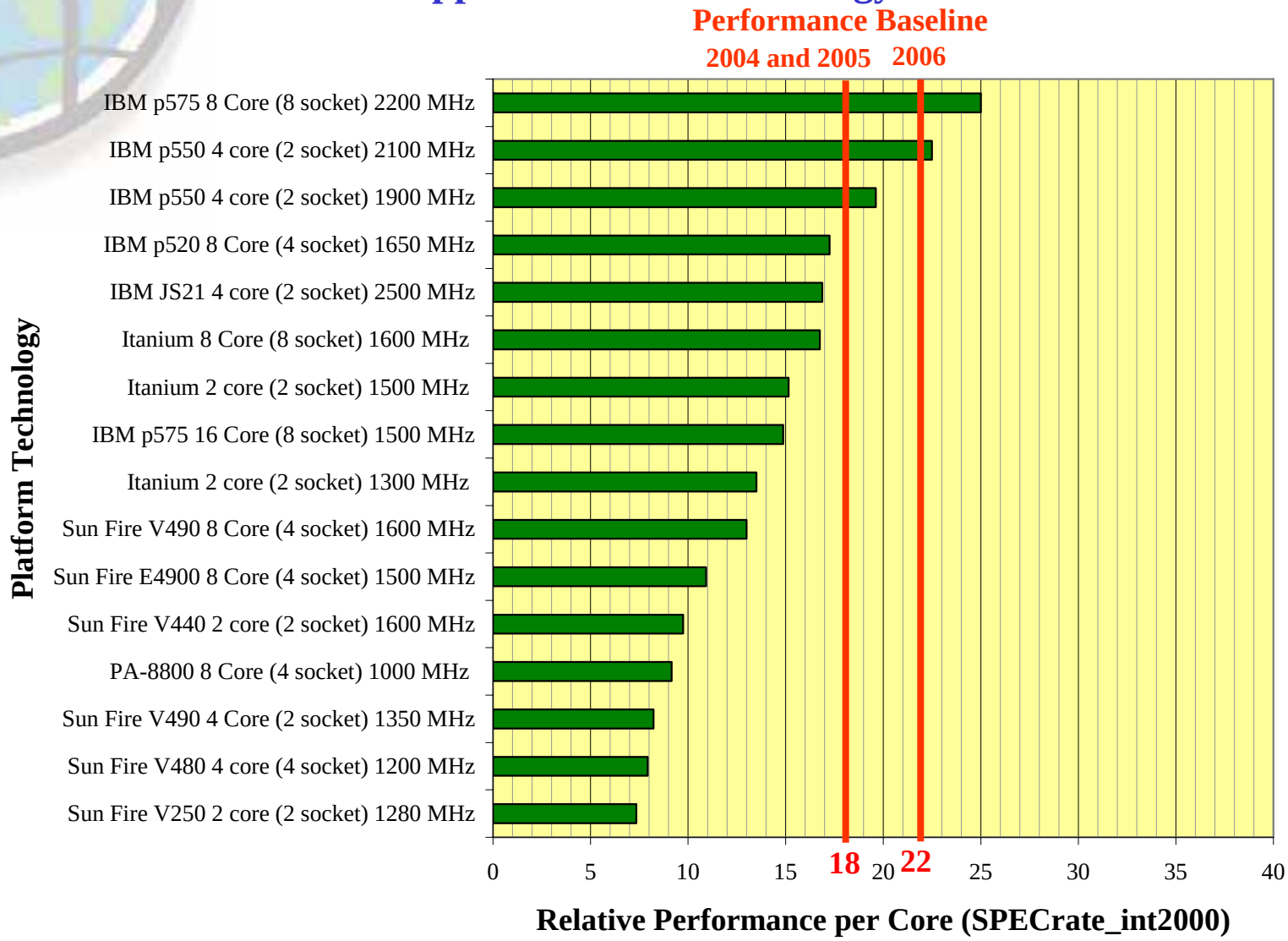
Platform Performance Makes a Difference

2006 Supported Windows Technology

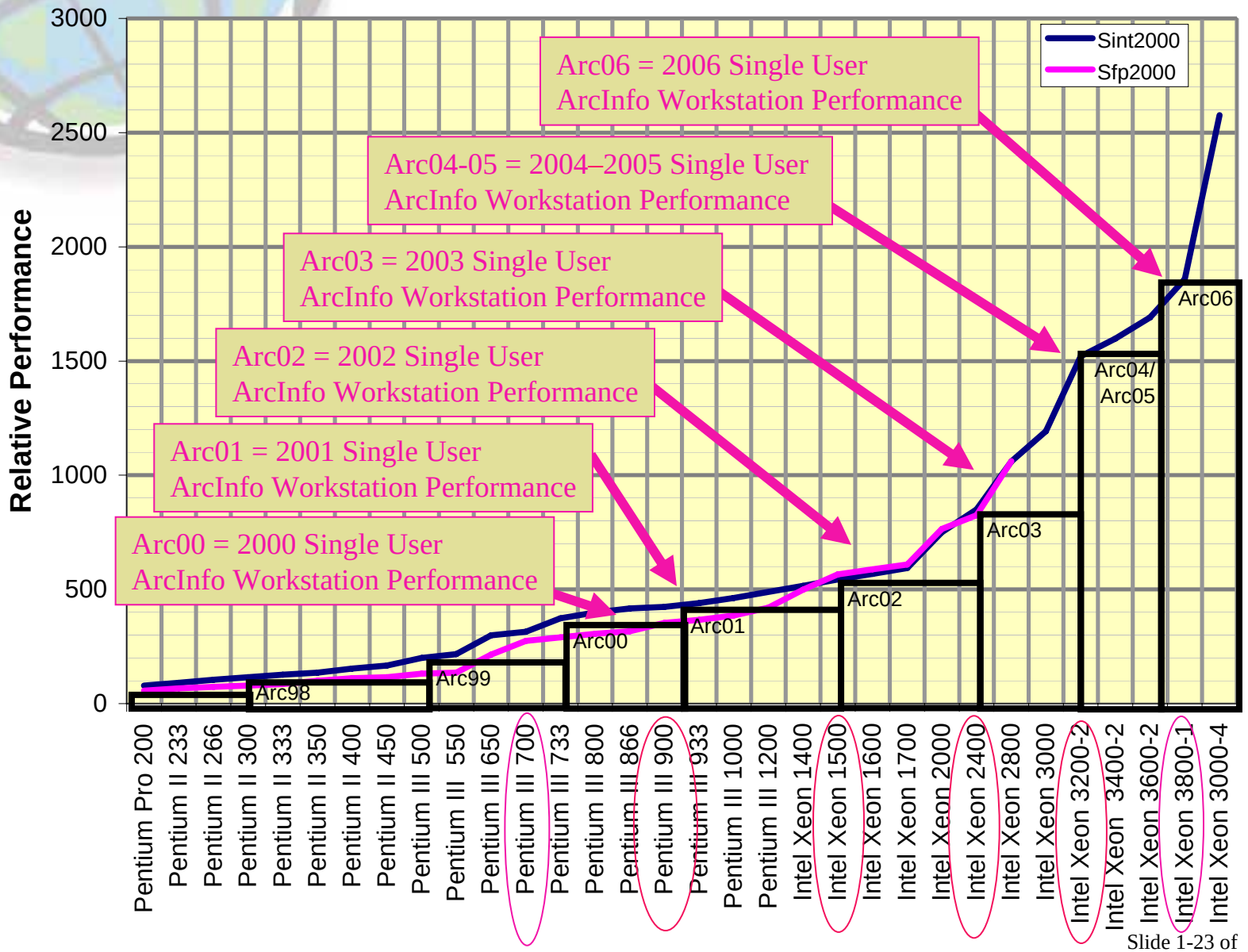


Platform Performance Makes a Difference

2006 Supported UNIX Technology



Platform Performance Change



How do we handle change?

Theory of Relative Performance

The relative performance of two servers is directly proportional to their compute capacity.

$$\frac{\text{Performance of Server A}}{\text{Performance of Server B}} = \frac{\text{Clients of Server A}}{\text{Clients of Server B}}$$

How do we measure relative performance?

Standard Performance Evaluation Corporation (SPEC Mission)

To develop technically credible and objective benchmarks so that both computer designers and purchasers can make decisions on the basis of realistic workloads

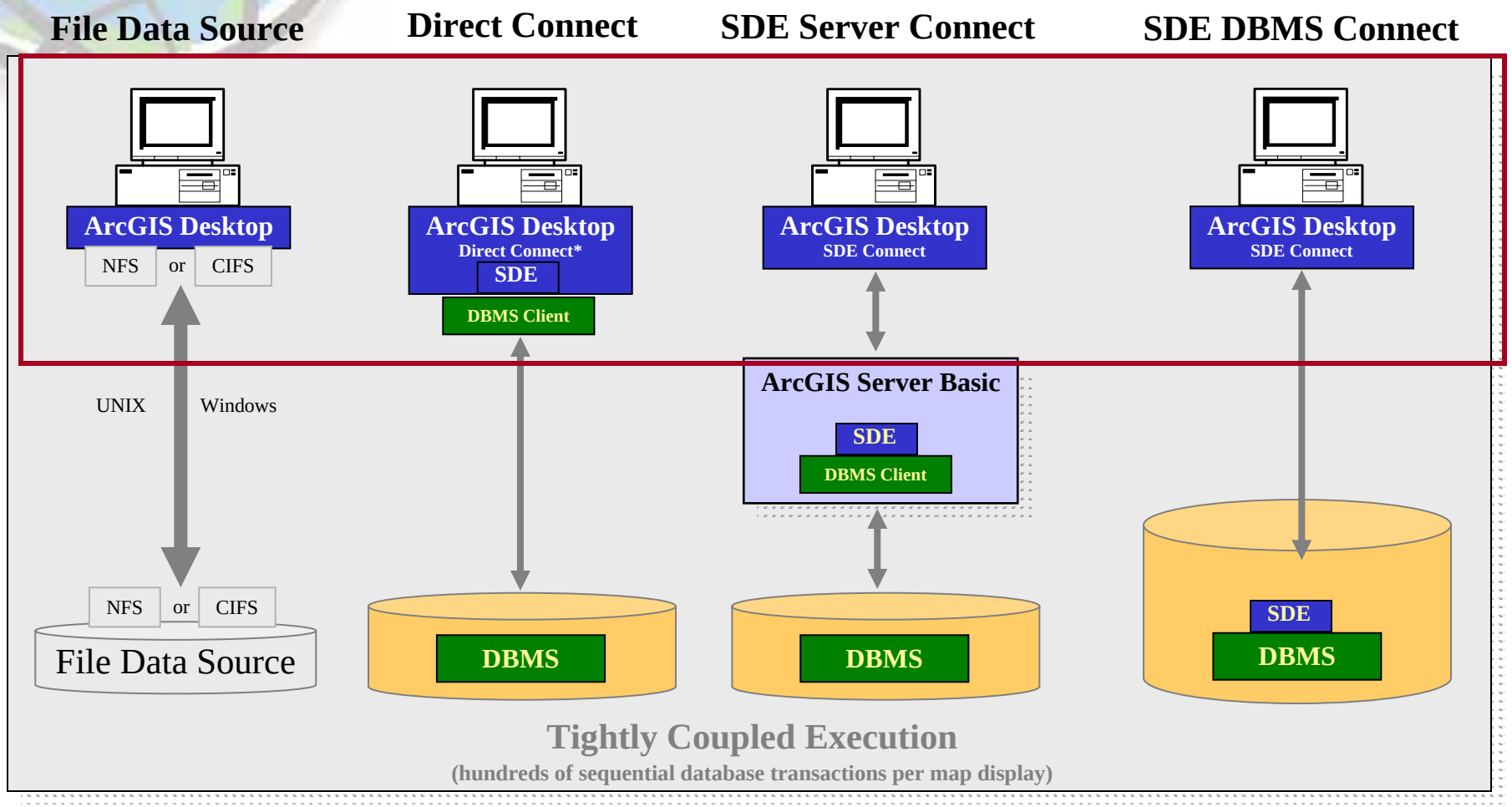
ESRI System Design History

- 1992–1996, SPEC92 benchmark results
- 1996–2000, SPEC95 benchmark results
- 2000–2006, SPEC2000 benchmark results
- New release UC2007+, SPEC2006 benchmark results

ArcGIS Desktop Sizing Models

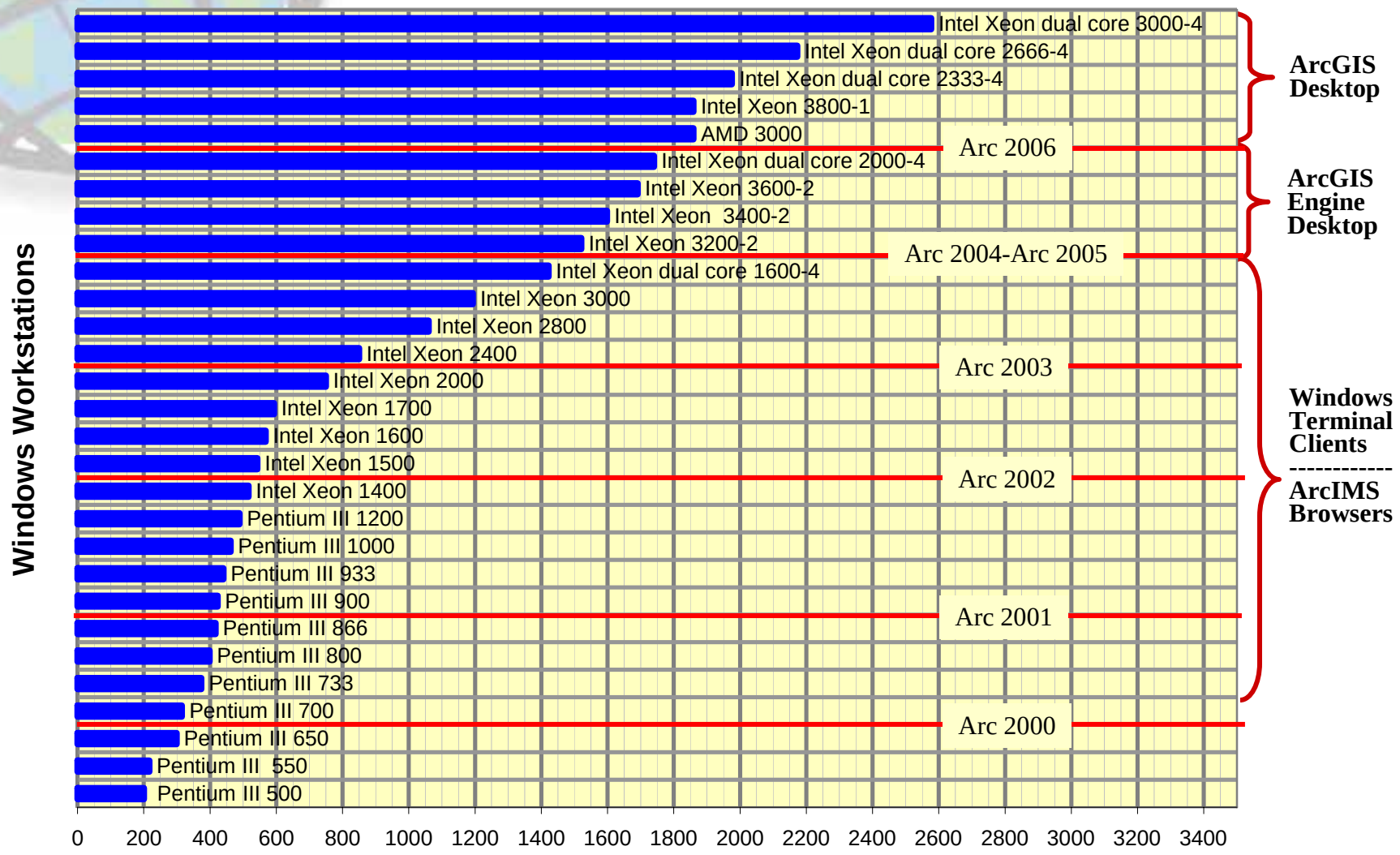
Client/Server Configurations

ArcGIS Desktop Workstations



*Note: ArcSDE executables are included in Direct Connect API

Workstation Platforms

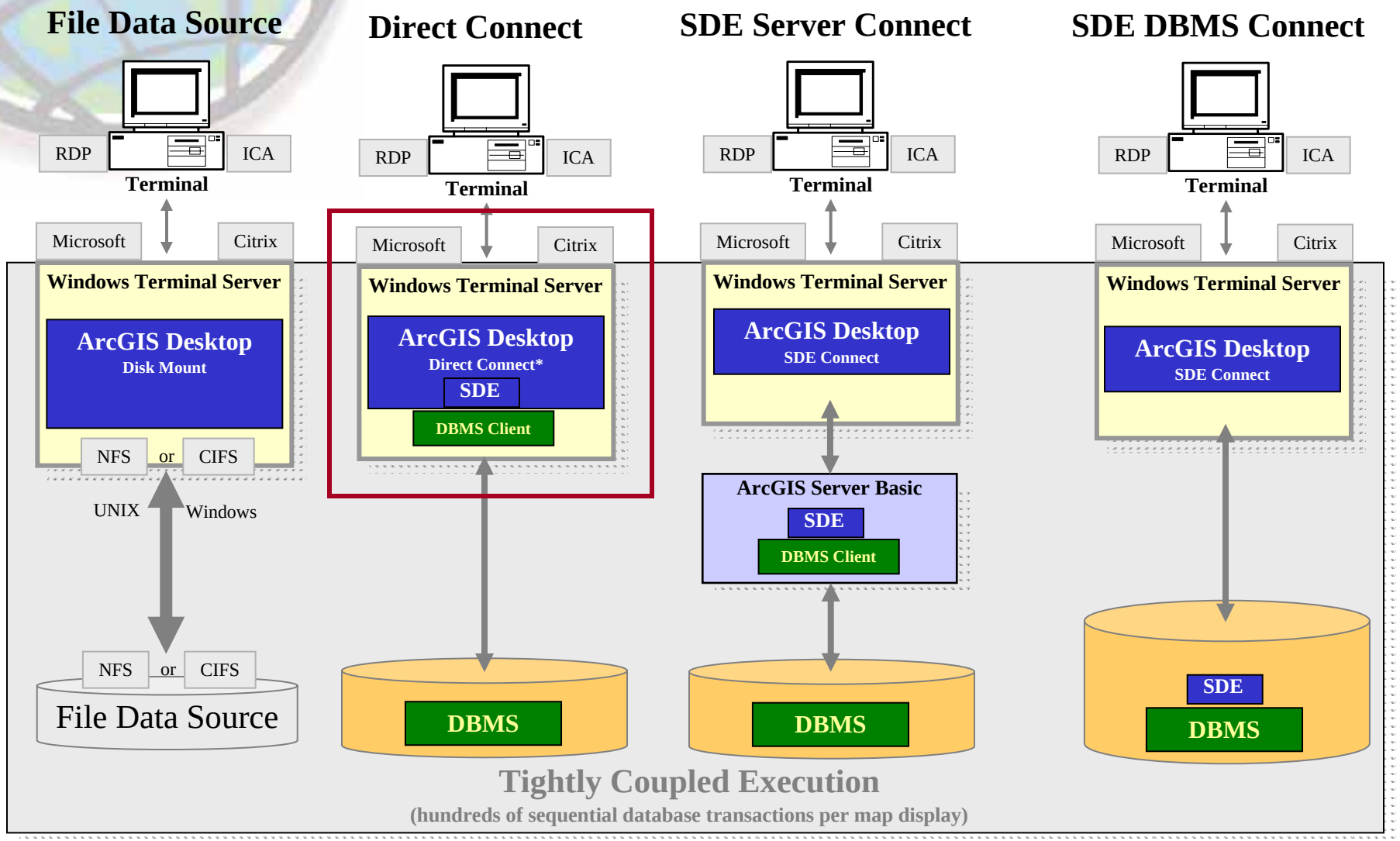


ArcGIS Memory = 1 GB

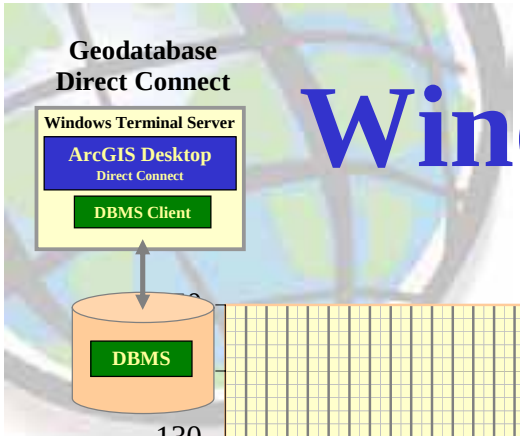
Up to 2 GB may be required to support large file-based data sources

Client/Server Configurations

ArcGIS Desktop Server

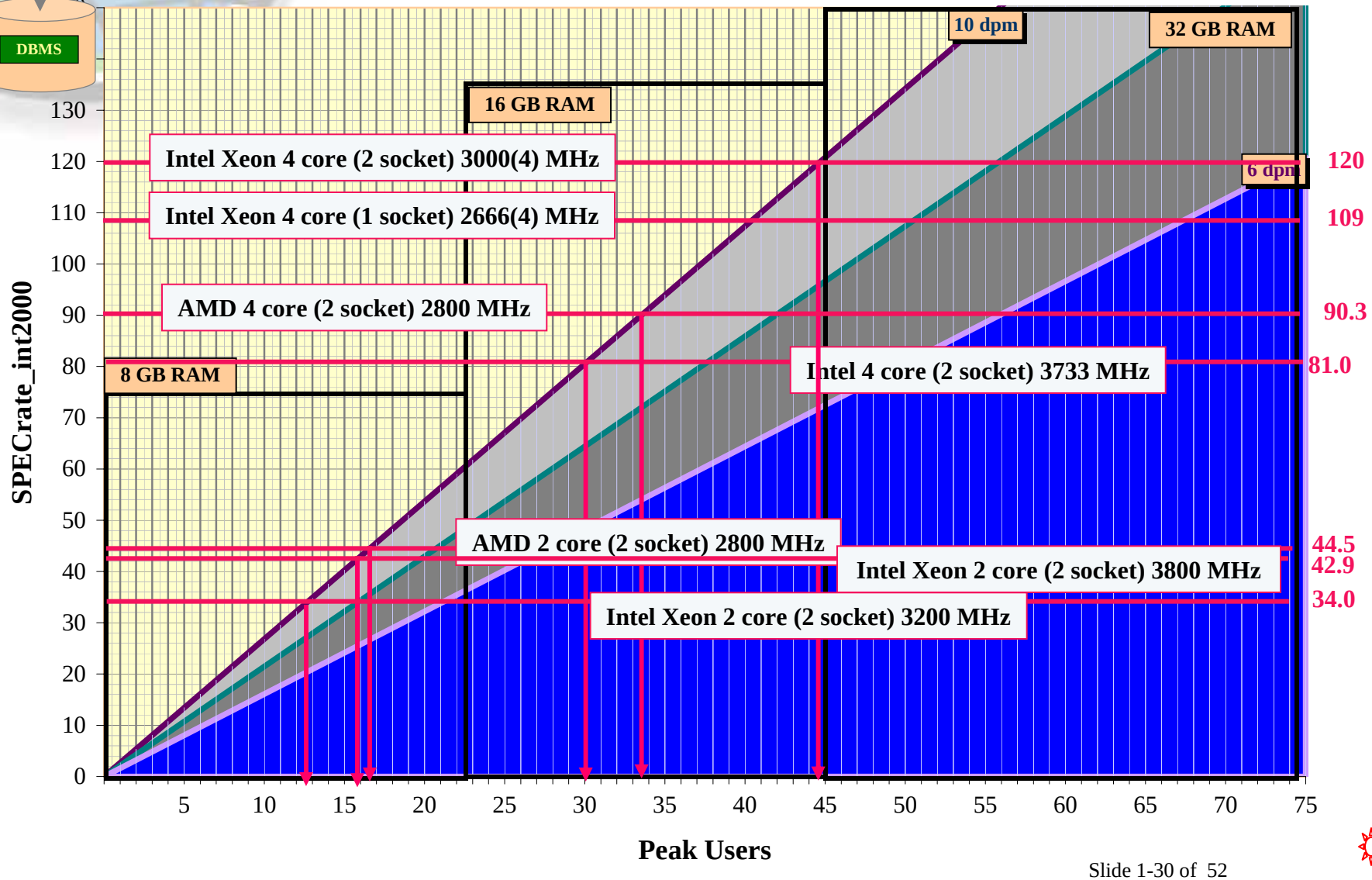


*Note: ArcSDE executables are included in Direct Connect API



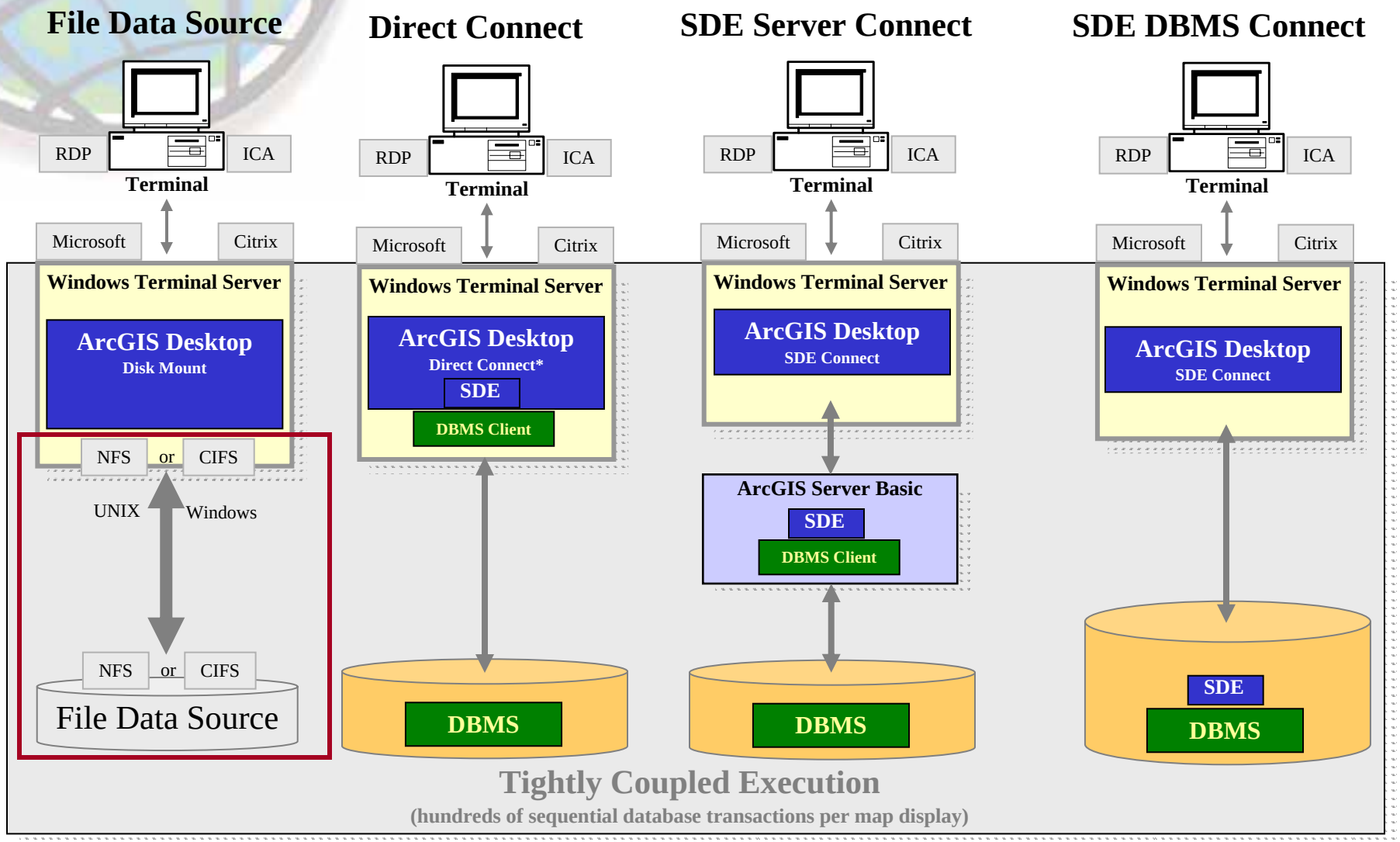
Windows Terminal Server

Geodatabase Direct Connect



Client/Server Configurations

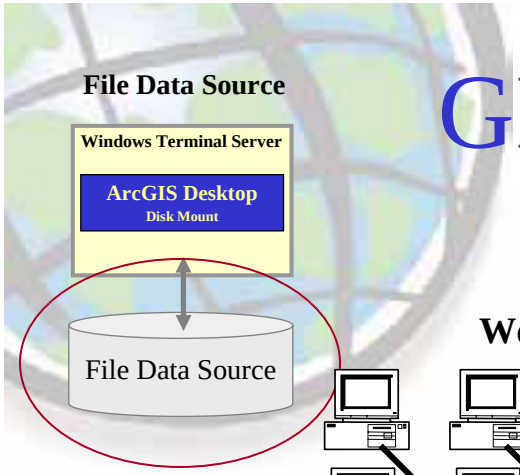
Centralized ArcGIS Desktop Client

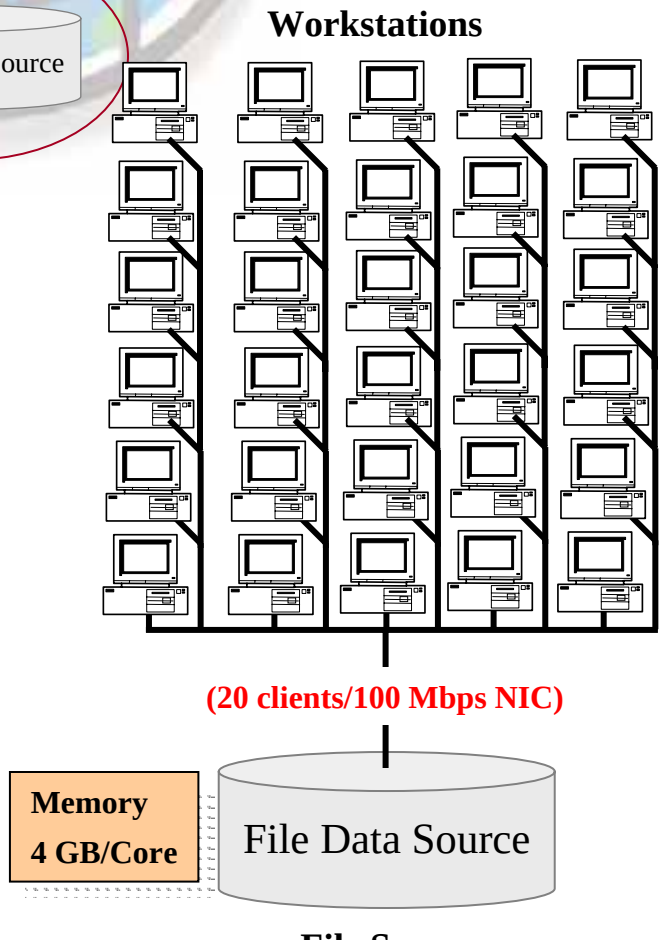


*Note: ArcSDE executables are included in Direct Connect API

GIS Data Server Sizing

GIS File Server





ArcGIS Desktop service time varies with size of the GIS data files

GIS File Servers

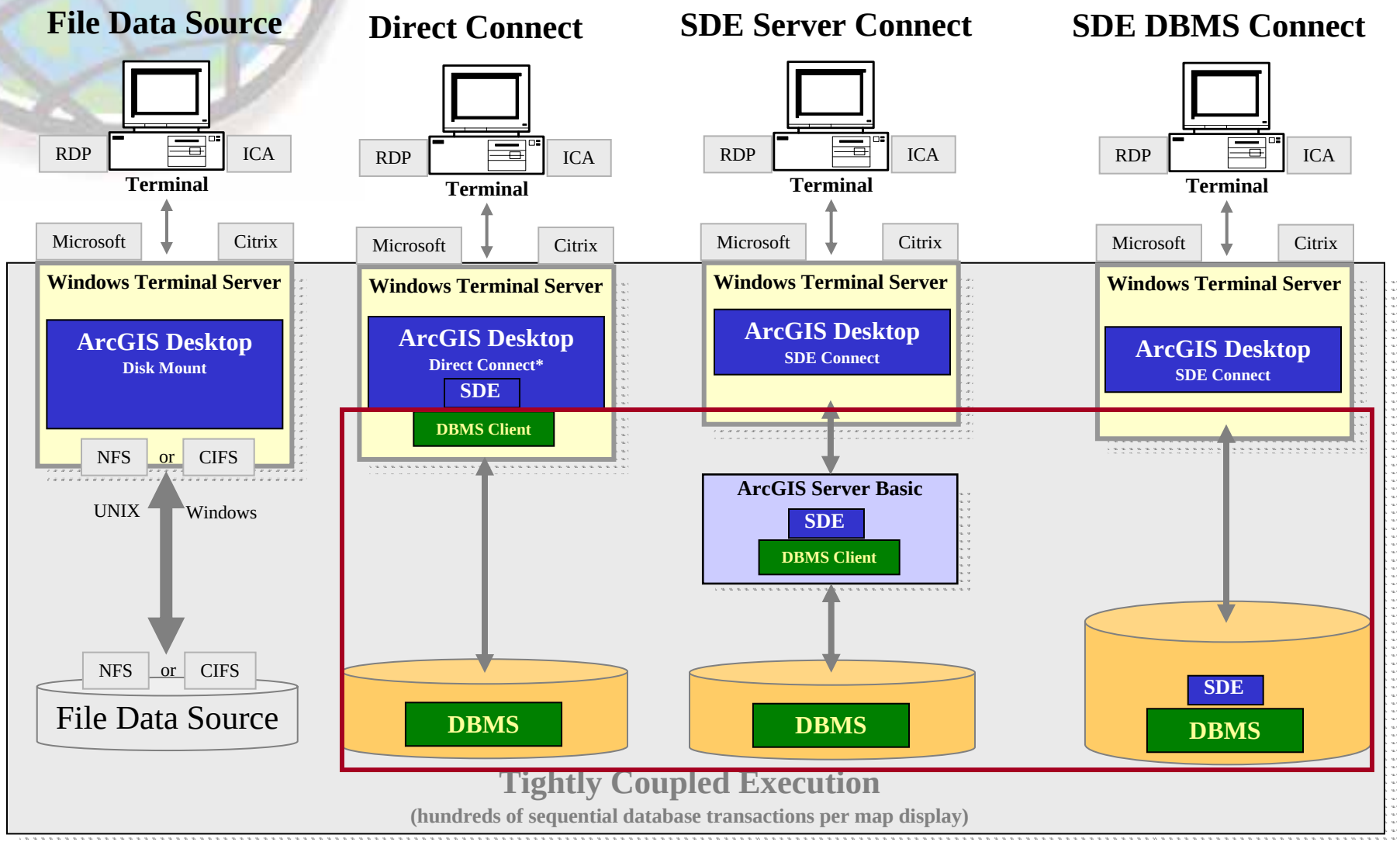
Capacity Planning Matrix

Traffic per Map Display (Mb)		File Server NIC Bandwidth (Mbps) 10	100	200	1000
10	Peak concurrent Clients	6	60	120	600
20		3	30	60	300
30		2	20	40	200
40		2	15	30	150
50		1	12	24	120
60		1	10	20	100
70		1	9	17	86
80		1	8	15	75
90		1	7	13	67
100		1	6	12	60

Average 50 Mb Traffic per Map Display (assume 1 display every 6 seconds)

Client/Server Configurations

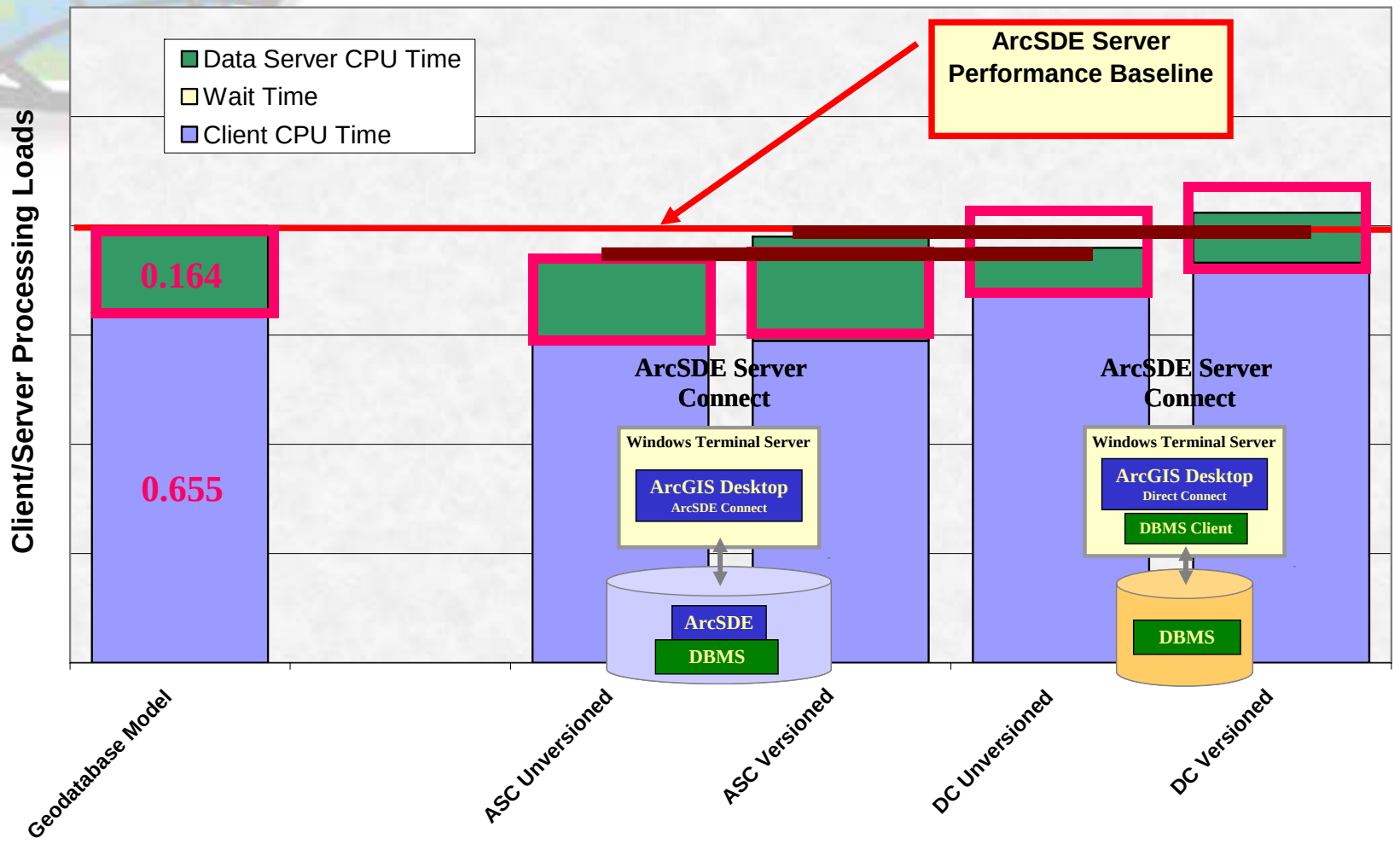
Centralized ArcGIS Desktop Client



*Note: ArcSDE executables are included in Direct Connect API

Geodatabase Performance

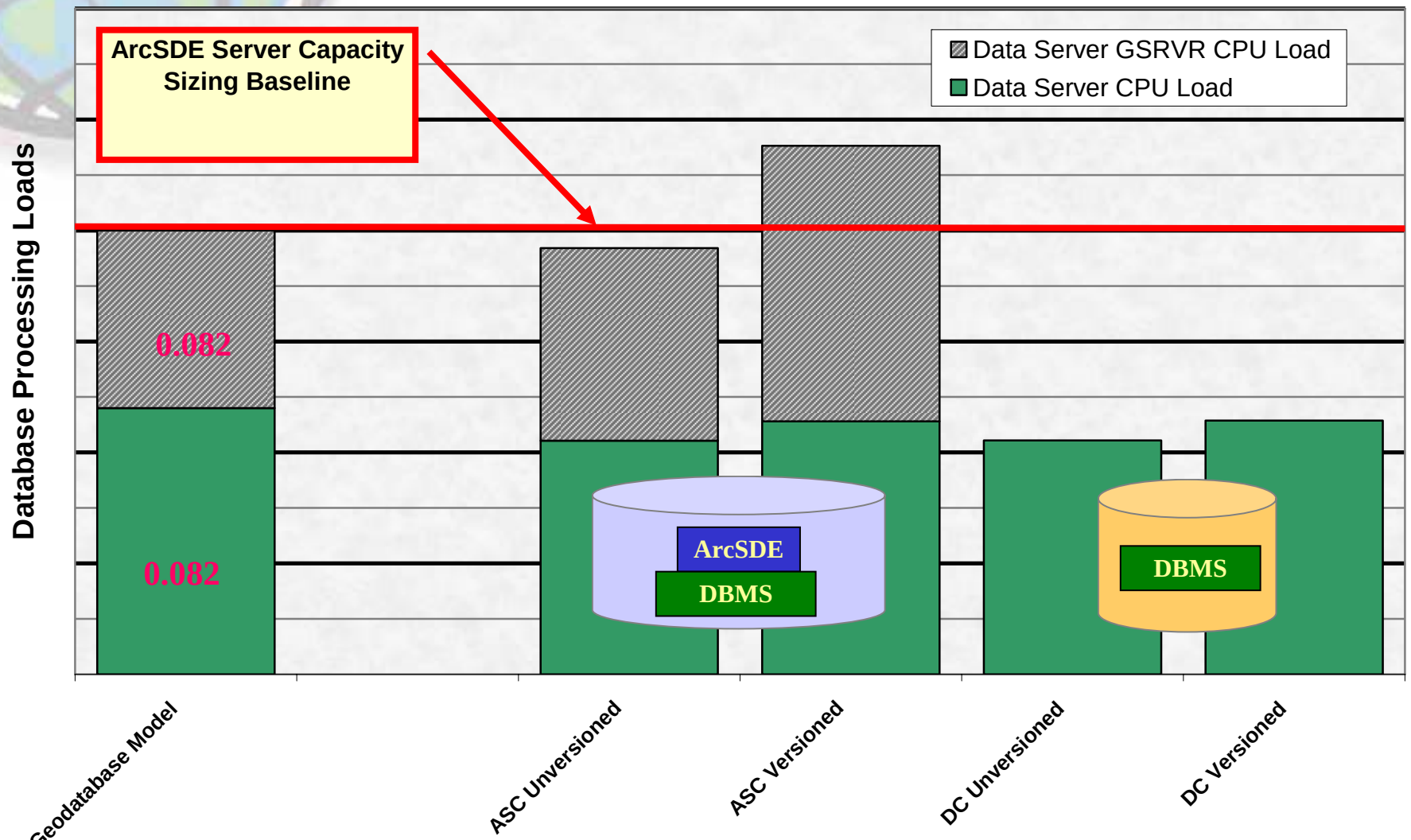
(ArcGIS 9.0 ArcMap/ArcSDE Vector and Raster Client/Server core Loads)



ASC = ArcSDE installed on database server
DC = ArcSDE direct connect communications
Model does not apply to Oracle Spatial data types (use DC configurations only)

Geodatabase Loads

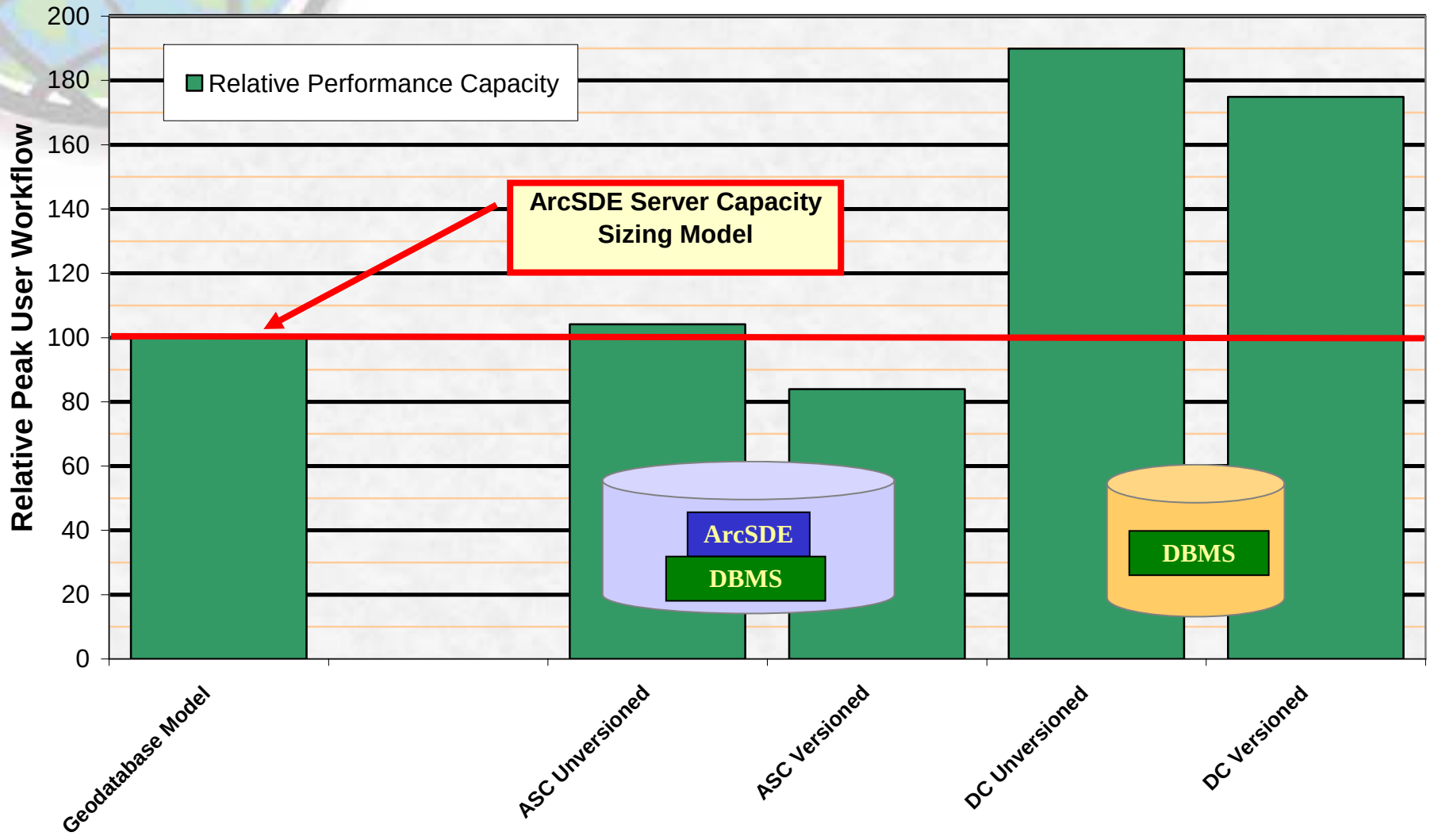
ArcGIS 9.0 Data Server Loads



ASC = ArcSDE installed on database server
DC = ArcSDE direct connect communications
Model does not apply to Oracle Spatial data types (use DC configurations only)

Relative Capacity Sizing

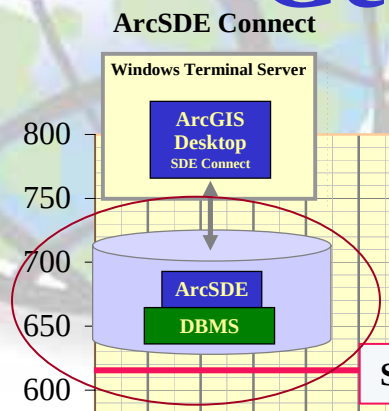
ArcGIS 9.0 Test Results



ASC = ArcSDE installed on database server
DC = ArcSDE direct connect communications
Model does not apply to Oracle Spatial data types (use DC configurations only)

Geodatabase Server Sizing

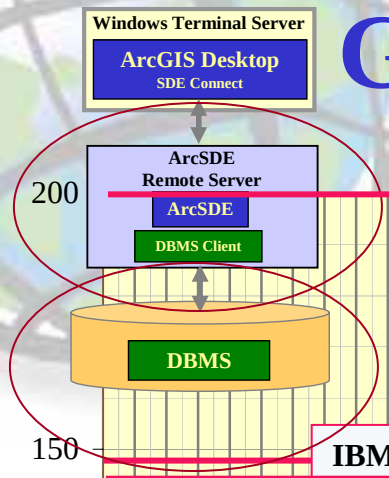
ArcSDE DBMS Server



Server memory = 4 GB/Core

Geodatabase Server Sizing

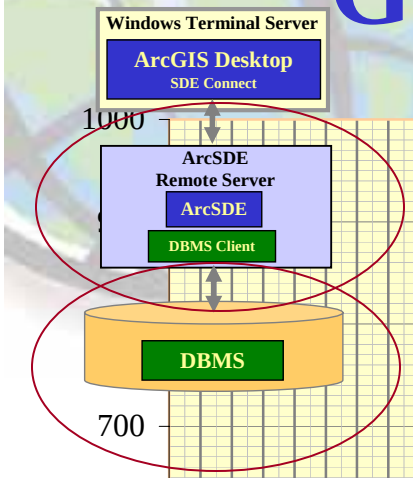
DBMS Direct Connect and ArcSDE Remote Server (up to 8 core)



Server memory = 4 GB/Core

Geodatabase Server Sizing

Large DBMS Direct Connect and ArcSDE Remote Server

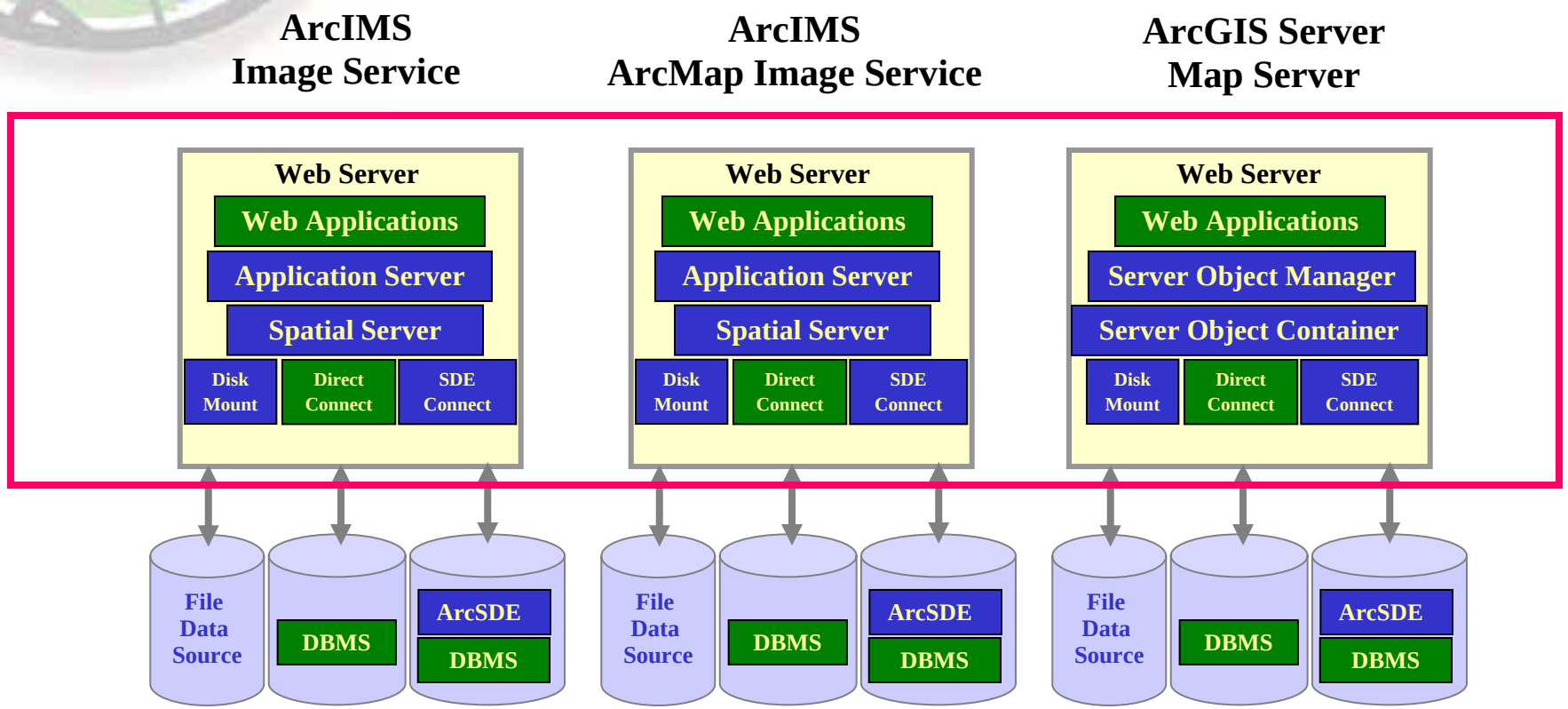


* ArcSDE Direct Connect Support Only

Server memory = 4 GB/Core

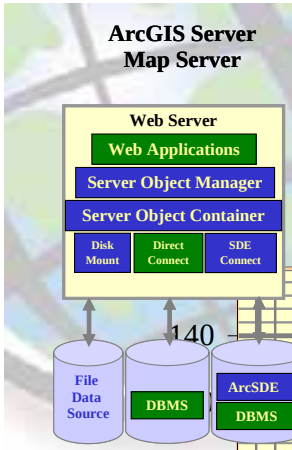
Server Performance and Scalability

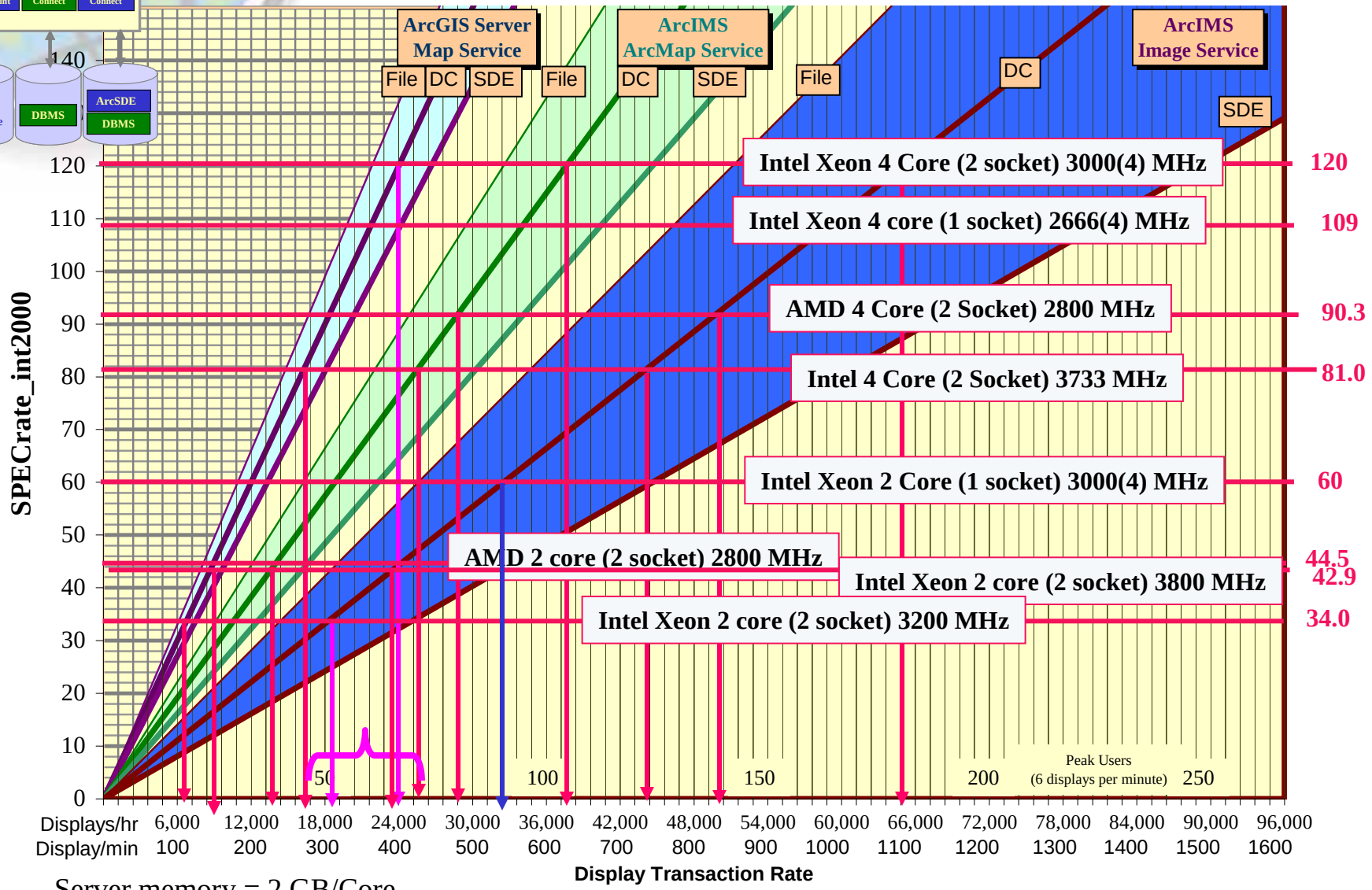
Two Tier Implementations



ArcIMS/ArcGIS Server Sizing

Web Two Tier Architecture

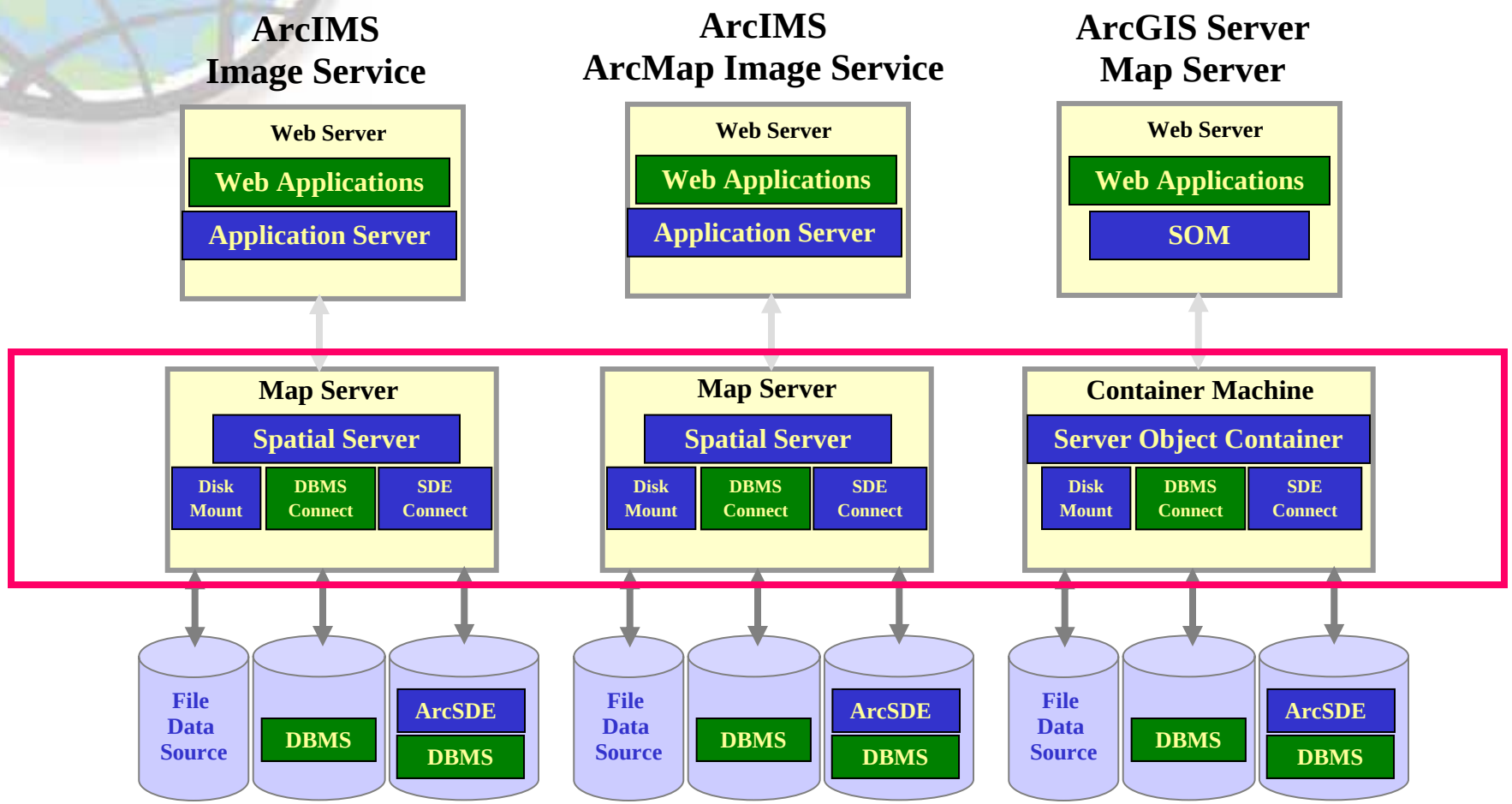




Server memory = 2 GB/Core

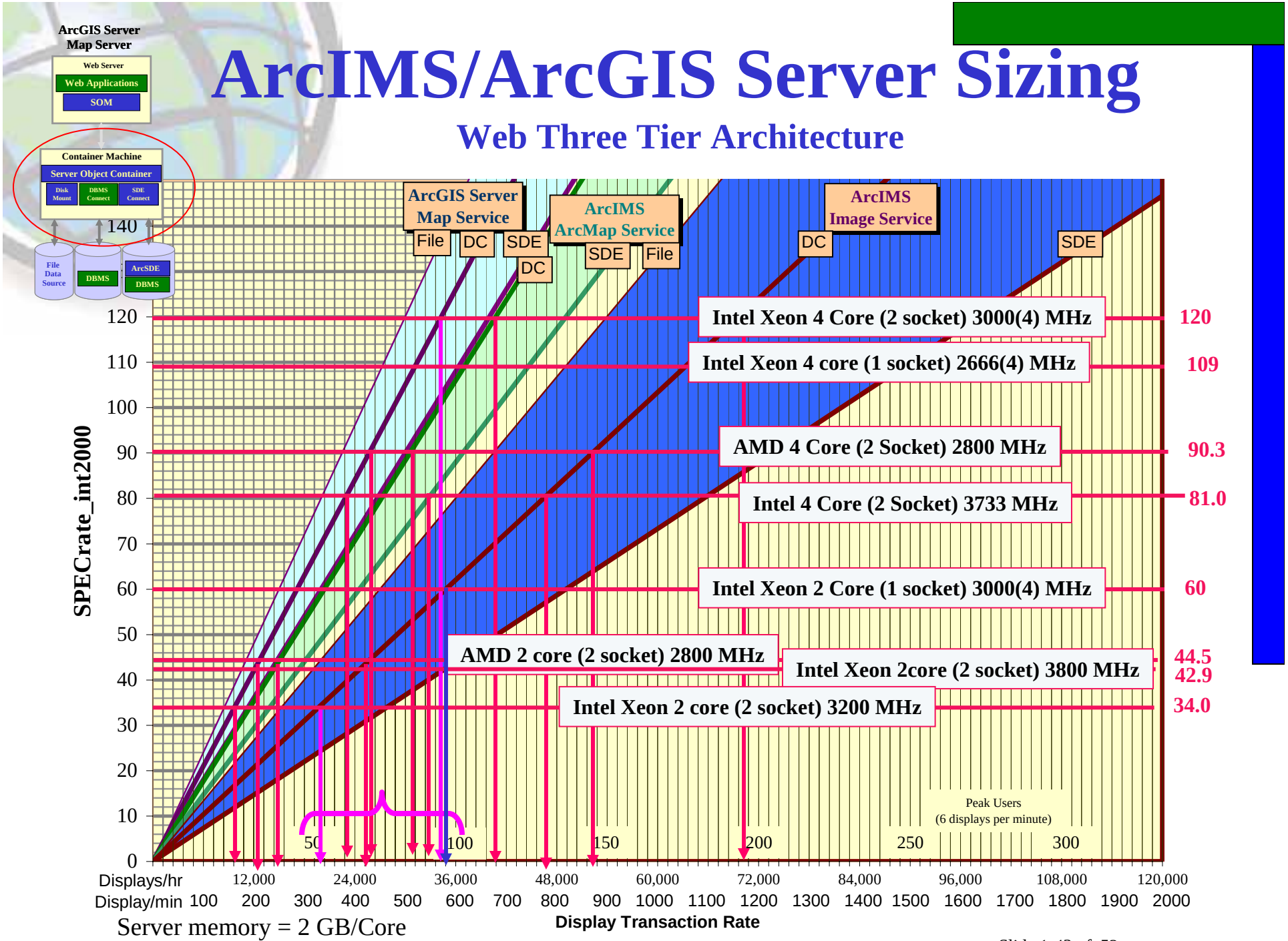
Server Performance and Scalability

Larger Three Tier Implementations



ArcIMS/ArcGIS Server Sizing

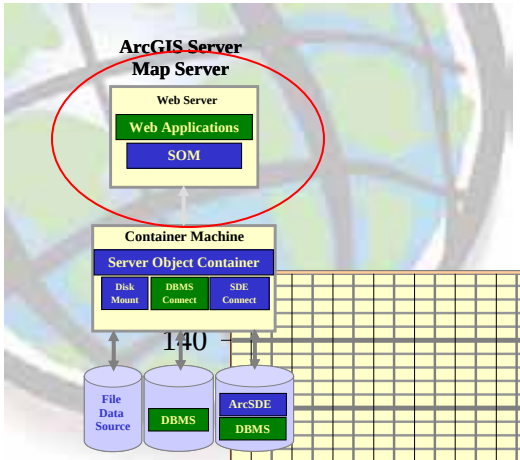
Web Three Tier Architecture

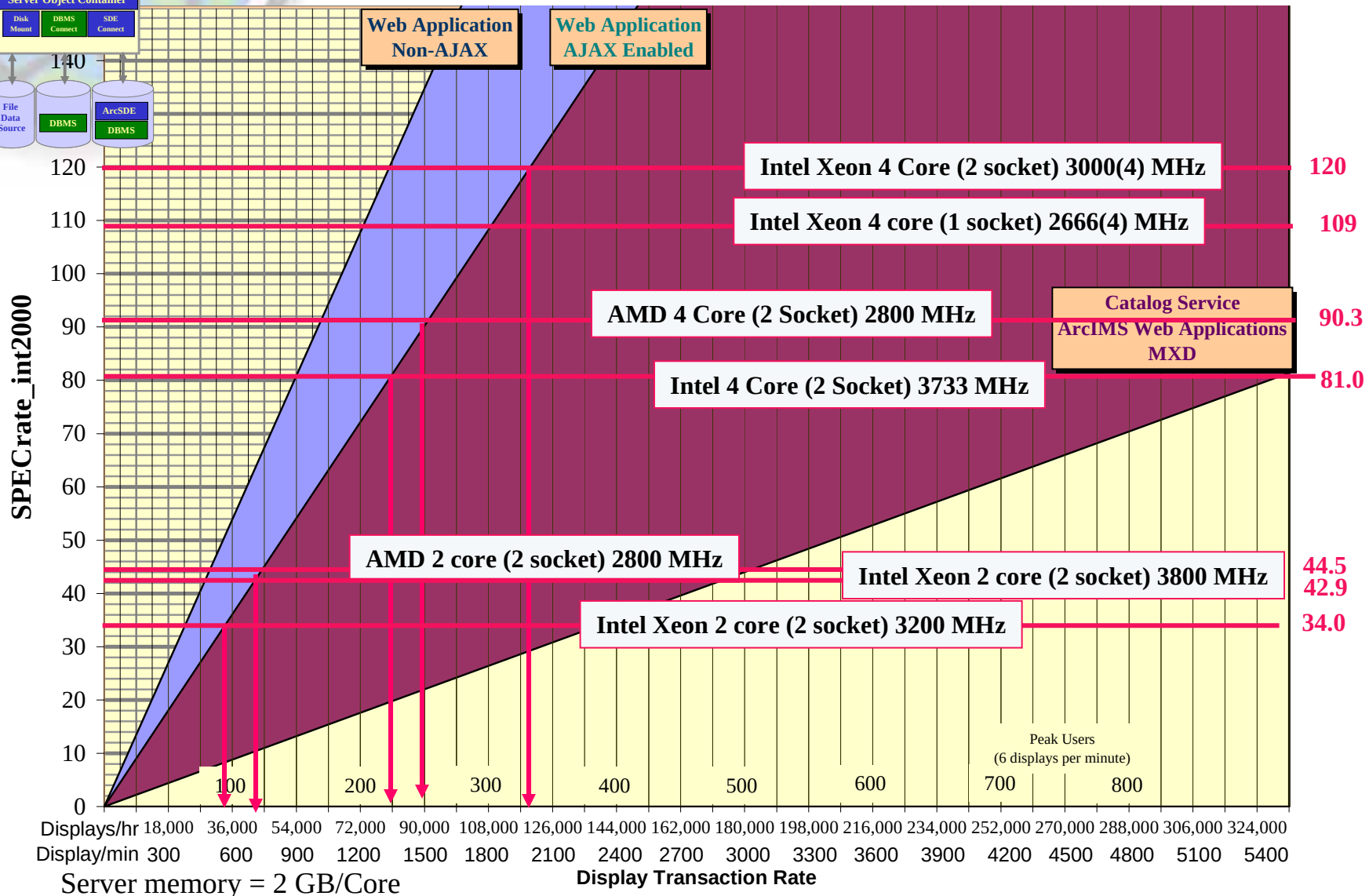


Server memory = 2 GB/Core

Web Server Sizing

Web Application Server





Server memory = 2 GB/Core

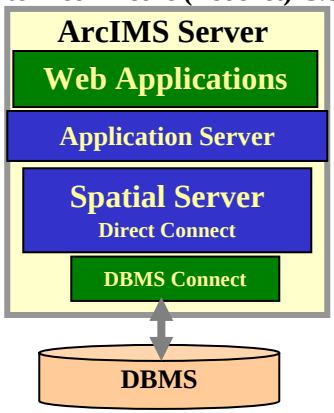
ArcIMS Image and ArcMap Service

Scalable Architecture (based on Intel Xeon single core socket 3.8 GHz Platform Baseline)

[Single Socket (1 single core) Workgroup Server/Dual Socket (2 single core) Enterprise Servers]

Enterprise

Intel Xeon 2 core (2 socket) 3.8 GHz



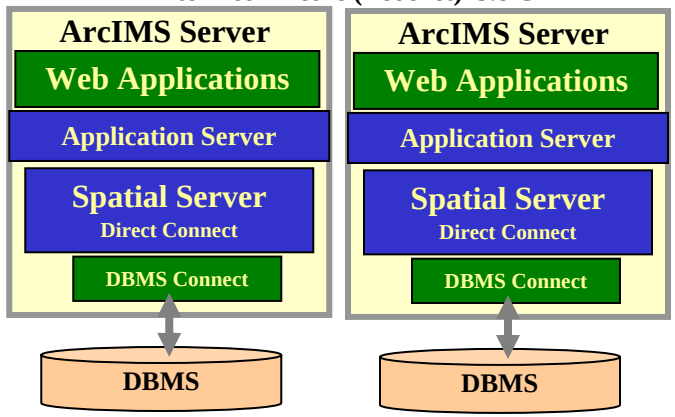
Target Capacity

Image Server = 387 dpm (25,740 dph)

ArcMap Server = 224 dpm (13,800 dph)

High Available Enterprise

2 x Intel Xeon 2 core (2 socket) 3.8 GHz



Target Capacity

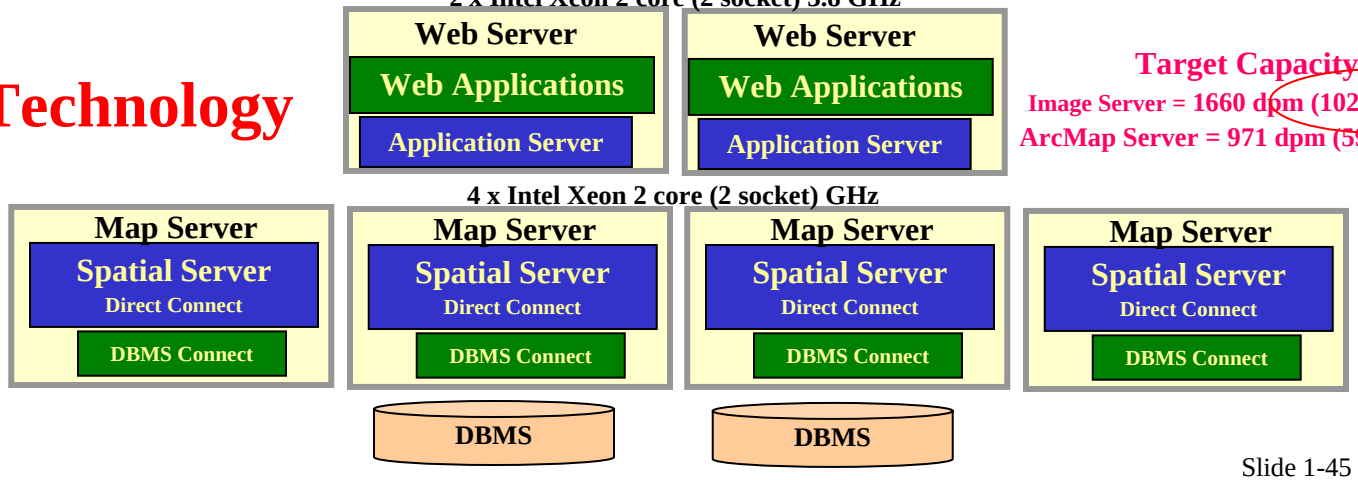
Image Server = 775 dpm (51,480 dph)

ArcMap Server = 448 dpm (27,600 dph)

Three Tier High Available Enterprise

2 x Intel Xeon 2 core (2 socket) 3.8 GHz

4 x Intel Xeon 2 core (2 socket) GHz



Target Capacity

Image Server = 1660 dpm (102,120 dph)

ArcMap Server = 971 dpm (59,760 dph)

2006 Technology

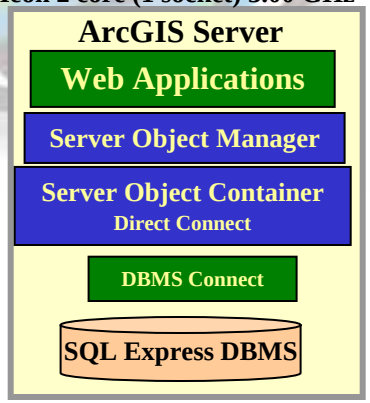
ArcGIS Server Map Service

Scalable Architecture (based on Intel Xeon new dual core socket 3.0 GHz Platforms)

[2 core (1 socket) Workgroup Server 4 core (2 socket) Enterprise Servers]

Workgroup

Intel Xeon 2 core (1 socket) 3.00 GHz



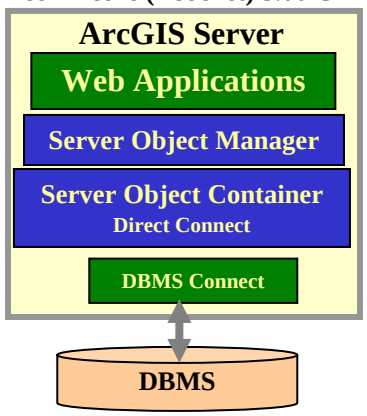
Target Capacity

Map Server = 182 dpm (10,905 dph)

Up to 30 concurrent user sessions (6 dpm)

Enterprise

Intel Xeon 4 core (2 socket) 3.00 GHz



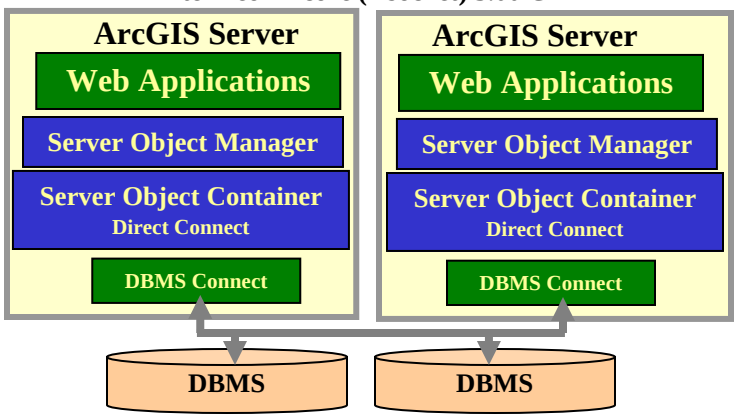
Target Capacity

Map Server = 400 dpm (23,995 dph)

Up to 67 concurrent user sessions (6 dpm)

High Available Enterprise

2 x Intel Xeon 4 core (2 socket) 3.00 GHz



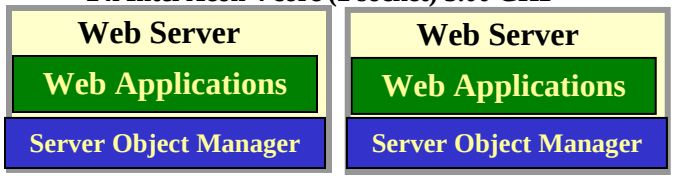
Target Capacity

Map Server = 800 dpm (47,989 dph)

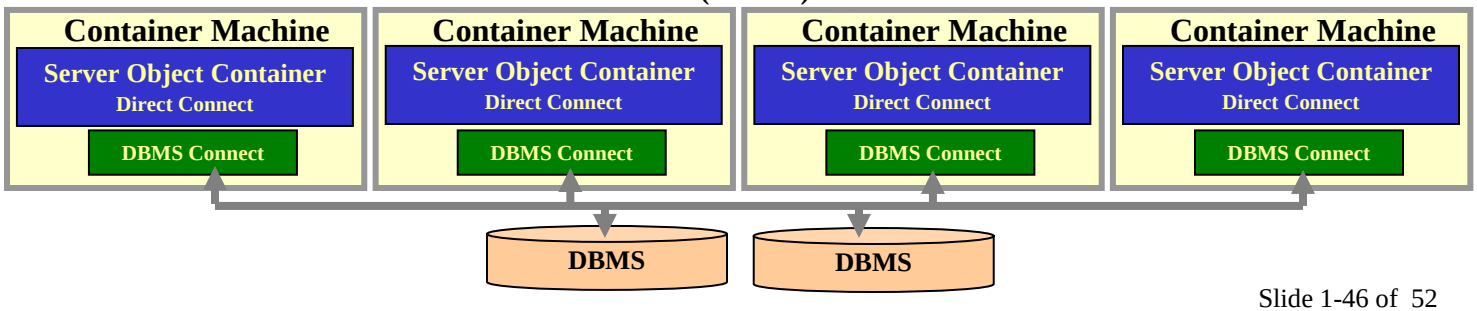
Up to 133 concurrent user sessions (6 dpm)

Three Tier High Available Enterprise

2 x Intel Xeon 4 core (2 socket) 3.00 GHz



4 x Intel Xeon 4 core (2 socket) 3.00 GHz



Target Capacity

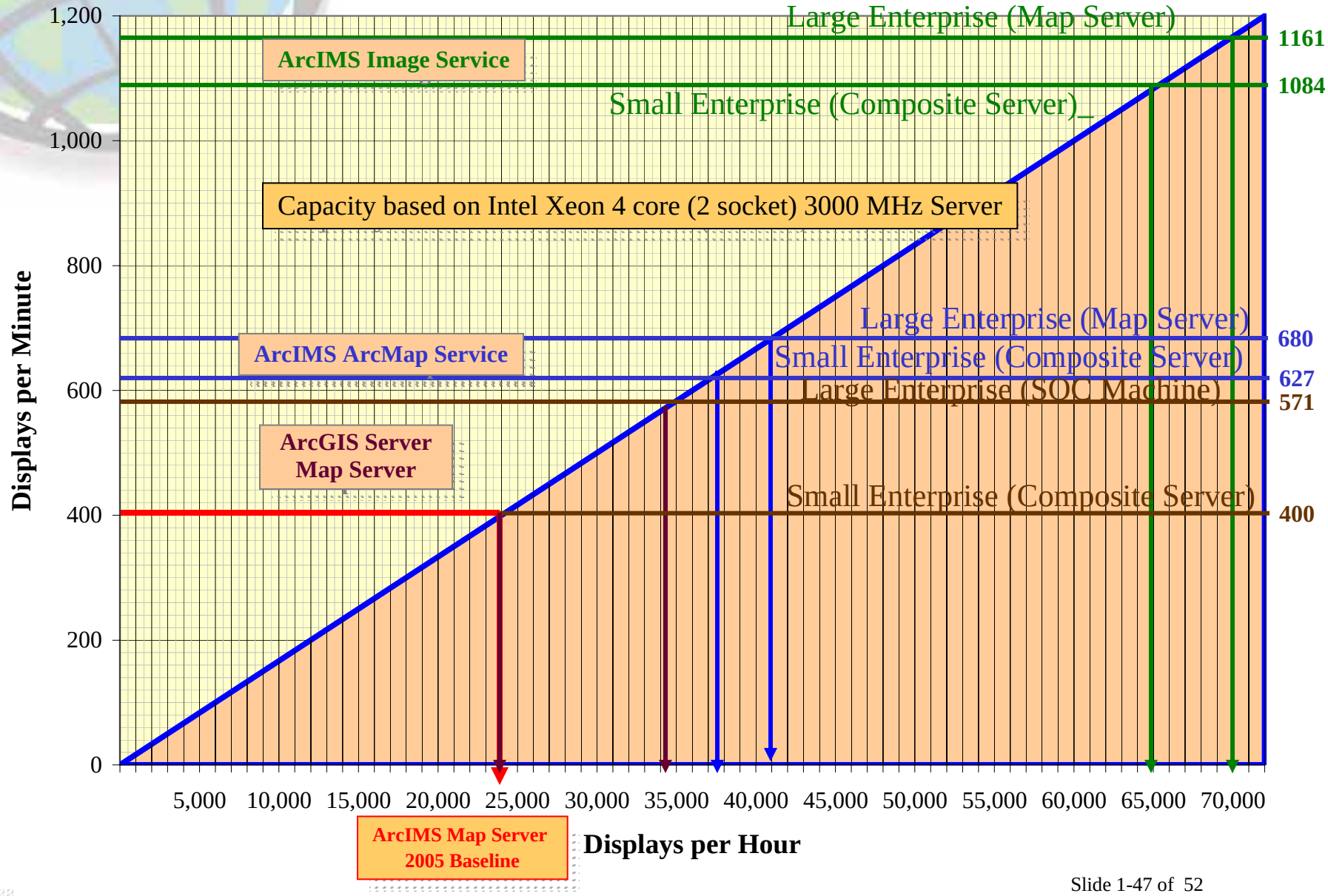
Map Server = 2,285 dpm (137,099 dph)

Up to 381 concurrent user sessions (6 dpm)

2007 Technology

Displays per Hour Conversion

(DBMS Direct Connect Data Source)



Selecting the Right Hardware

Capacity and Performance

Understand the Difference

Hardware Selection Does Make a Difference

Vendor Published Hardware Metrics

(www.specbench.org)

Hardware vendor benchmarks are published on the SPEC website

New June 2006 Processors

OLD Processors

Relative Performance

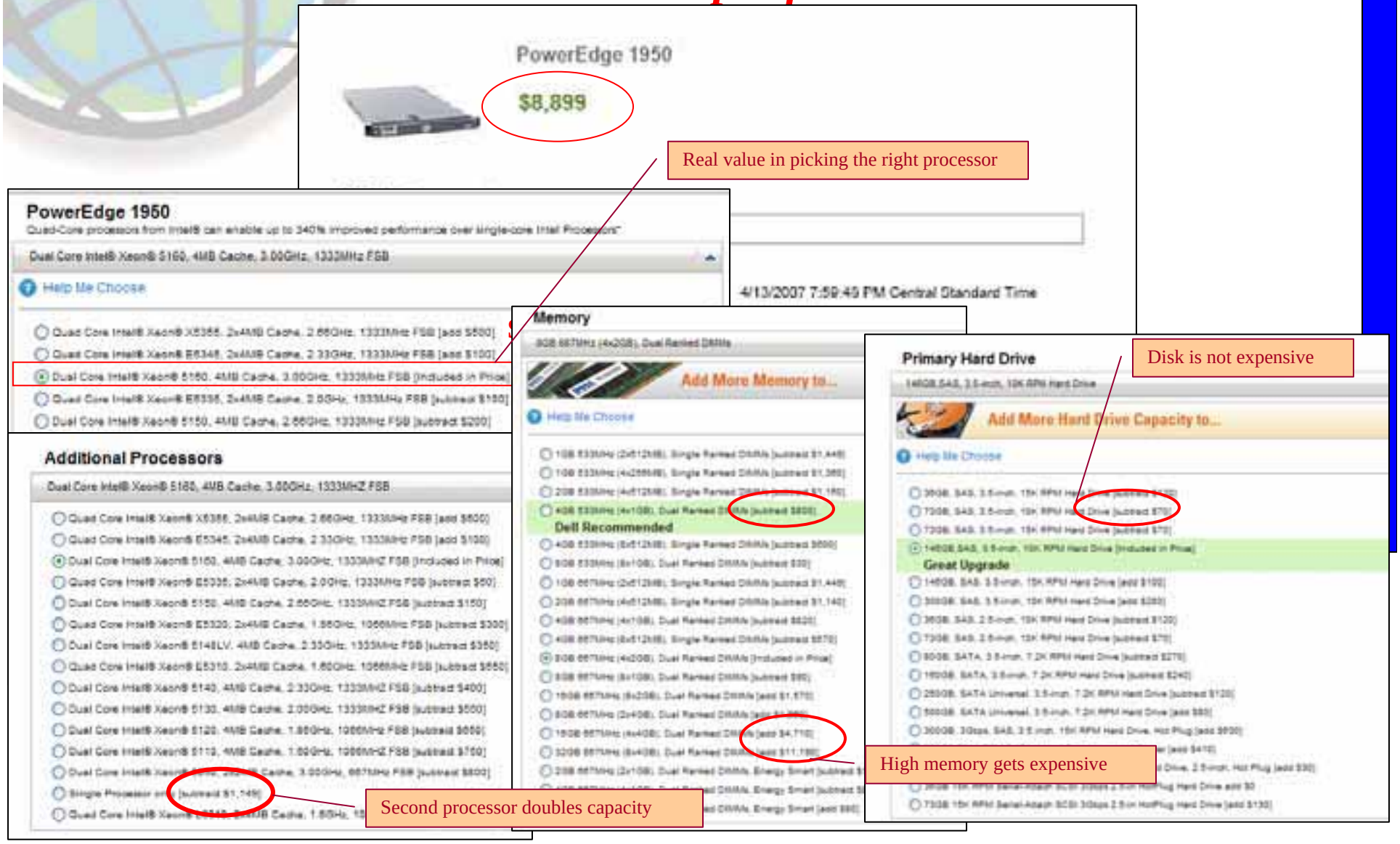
Company	System	# cores	# chips	# cores per chip	Processor MHz	Baseline	result/ core
Dell	PowerEdge 2950 (Intel® Xeon® processor 5160, 3.00GHz)	4	2	2	3000	123	30.8
Dell	PowerEdge 1950 (Intel Xeon processor 5150, 2.66GHz)	4	2	2	2666	111	27.8
Dell	PowerEdge 1950 (Intel Xeon processor 5140, 2.33GHz)	4	2	2	2333	101	25.3
Dell	PowerEdge 1950 (Intel Xeon processor 5130, 2.00GHz)	4	2	2	2000	88.6	22.2
Dell	PowerEdge 1950 (Intel Xeon processor 5080, 3.73GHz)	4	2	2	3733	82.2	20.6
Dell	PowerEdge 1950 (Intel Xeon processor 5120, 1.86GHz)	4	2	2	1866	82.1	20.5
Hewlett-Packard Company	ProLiant DL380 G5 (1.86 GHz, Intel Xeon processor 5120)	4	2	2	1860	80.7	20.2
Dell	PowerEdge 1950 (Intel Xeon processor 5060, 3.20GHz)	4	2	2	3200	73	18.3
Dell	PowerEdge 1950 (Intel Xeon processor 5110, 1.60GHz)	4	2	2	1600	72	18.0
Dell	PowerEdge 1950 (Intel Xeon processor 5050, 3.00GHz)	4	2	2	3000	67	16.8
Fujitsu Siemens Computers	PRIMERGY RX300 S3, Intel Xeon processor E5310, 1.60 GHz	8	2	4	1600	132	16.5
Dell	PowerEdge 1855 (2.8GHz Intel dual-core Xeon)	4	2	2	2800	57.3	14.3

Higher MHz means Higher Energy Bills

We use vendor published SPECrate_int2000 benchmarks for capacity sizing

Identifying the Right Platform

How do we select the platform we want?



Platform Vendor Selection (Section Overview)

- **Factors Supporting Hardware Selection**
  - Platform Performance
  - Purchase Price
  - System Supportability
  - Vendor Relationships
  - Total Life-Cycle Costs

System Design Strategies

Capacity Planning Tool

DEMO

System Design Strategies TRD
www.esri.com/systemdesign
+
System Architecture Design Strategies Training Class