Scalability: Architecting for Growth

Glen Sheffield
shef@ca.ibm.com

Person to Person, Lisbon, November 2003

© Copyright IBM Corporation 2002

Scalability – Definition

- **How well a solution to some problem will work when the size of the problem increases.**
  - **Source : www.dictionary.com**
- **In our case – The size of the problem increases when users and/or data is added.**
  - **Need to manage growth in a predictable manner**

© Copyright IBM Corporation 2002

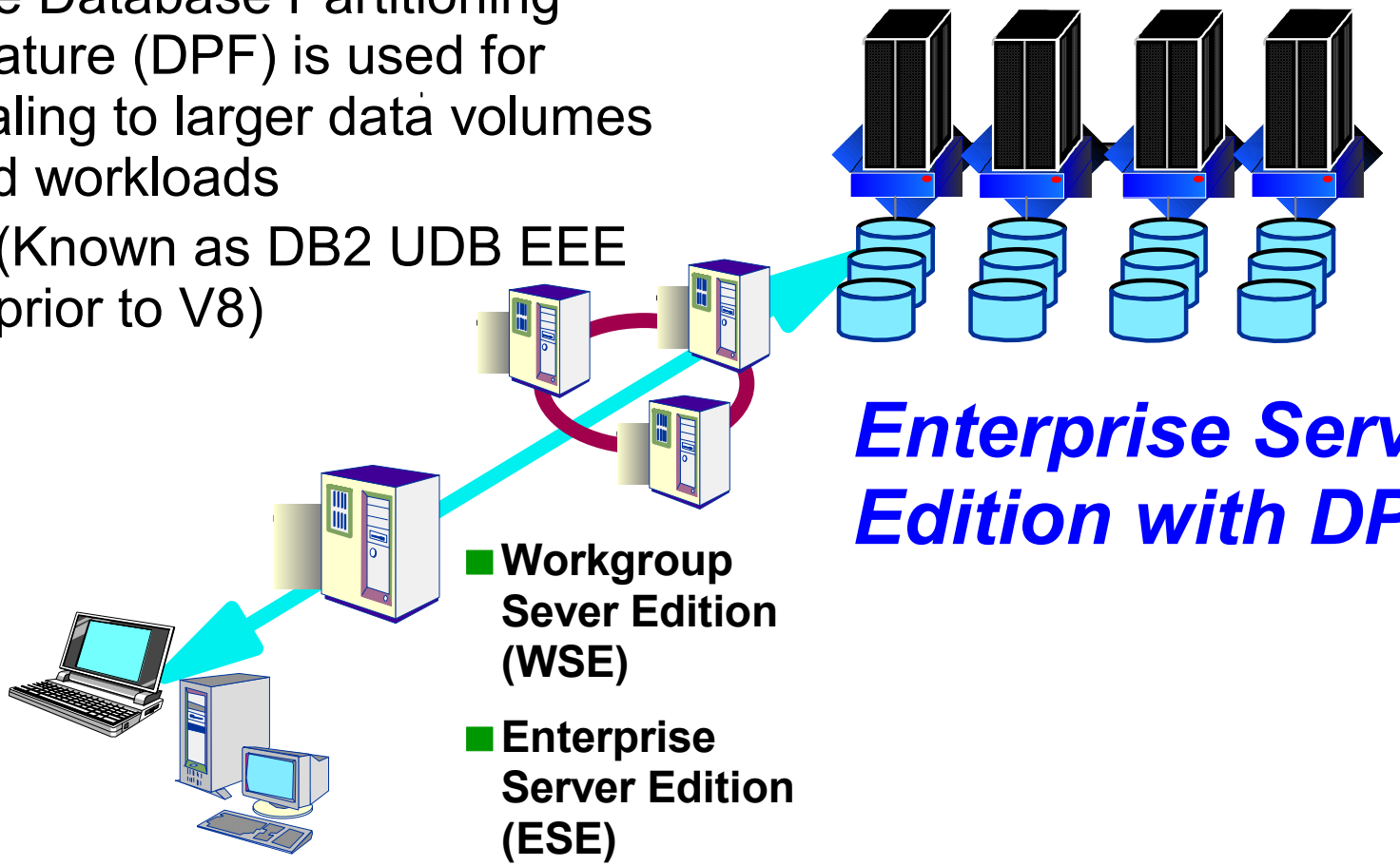
Scalability Principles

- **Scalable Database and Platform**
  - DB2 UDB and IBM Servers
- **"Basic Configuration Unit"**
  - System components must be added in proportional manner
  - CPU, Memory, I/O, interconnect
- **Tools to enhance and manage Scalability**
  - Query Patroller
  - Cube Views

© Copyright IBM Corporation 2002

DB2 UDB for Linux, Unix, and Windows

- The Database Partitioning Feature (DPF) is used for scaling to larger data volumes and workloads
  - (Known as DB2 UDB EEE prior to V8)



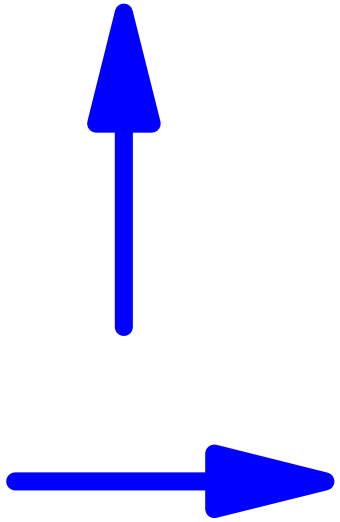
Enterprise Server Edition with DPF

- **Workgroup Sever Edition (WSE)**
- **Enterprise Server Edition (ESE)**

© Copyright IBM Corporation 2002

Hardware Scalability Models

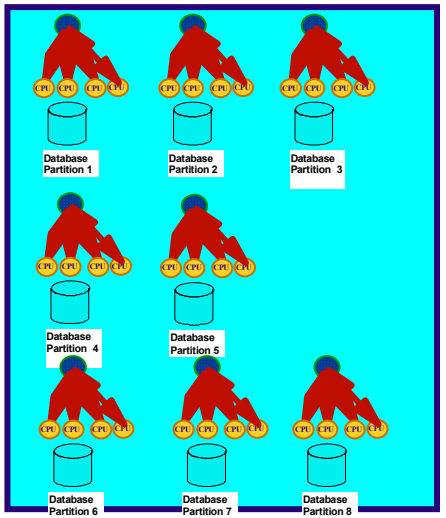
- SMP - Scaleup
  - Within a Single System OS Image
    - SMP Servers
    - e.g. IBM p690 "Regatta"
- Clustering - Scale-out
  - Build larger systems from building blocks
    - Clustered, or MPP systems
    - Building blocks can be large or small
    - Requires high performance / low latency interconnect
      - e.g. IBM High Speed Switch, gigabit ethernet, InfiniBand



© Copyright IBM Corporation 2002

SMP – Scale Up

- **Enhanced Scale up with DB2 Partitioning Feature**
  - Break the SMP into logical DB2 partitions
    - "MPP within the SMP"
    - Message passing between partitions via shared memory
  - Reduced contention for resources, e.g.
    - Dedicated memory per partition
    - Private data ownership
  - Result
    - Improved efficiency of adding CPUs
    - Better scaling

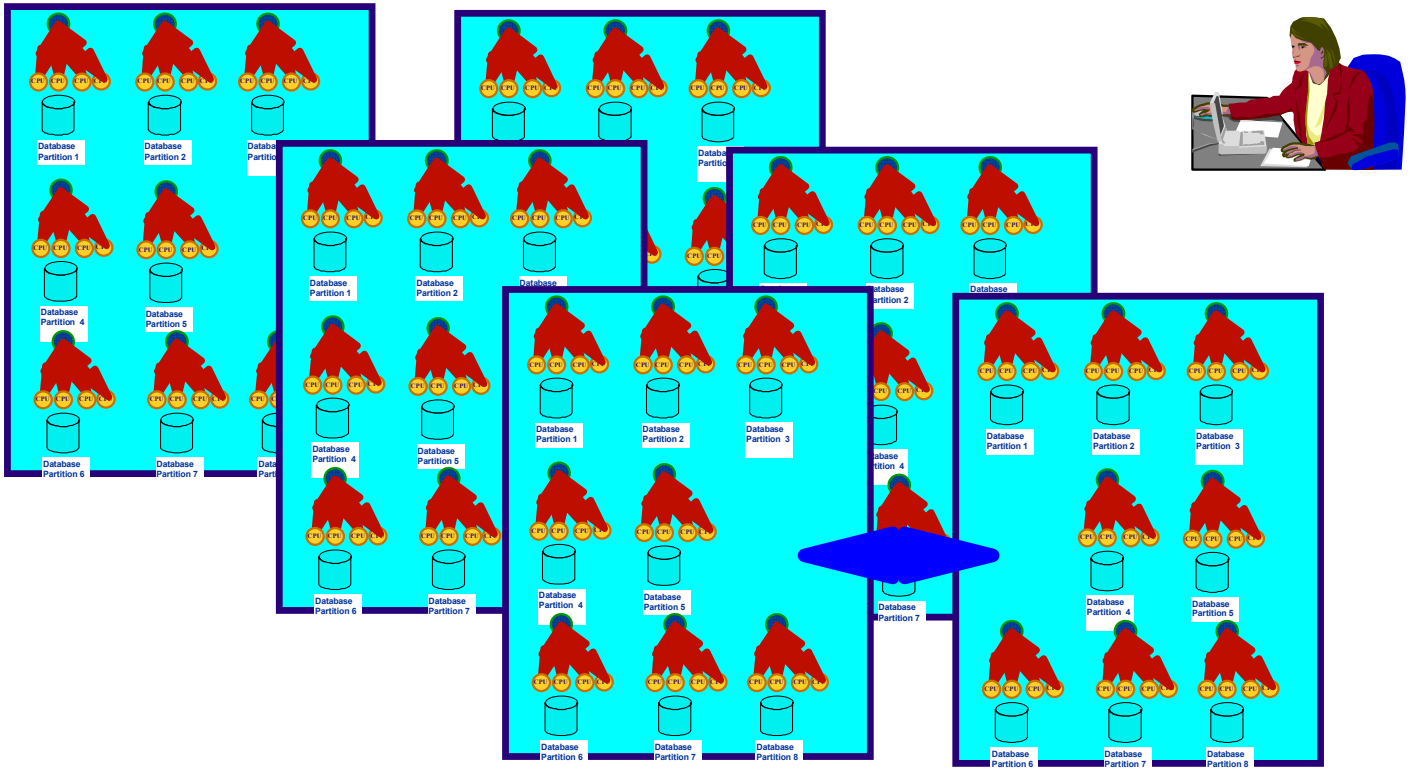


Multiple DB2 partitions leveraging CPUs and memory inside one SMP server for one query.

- Note - DB2 Partitioning does not require hardware LPAR capability

© Copyright IBM Corporation 2002

Clustering – Scaling Out

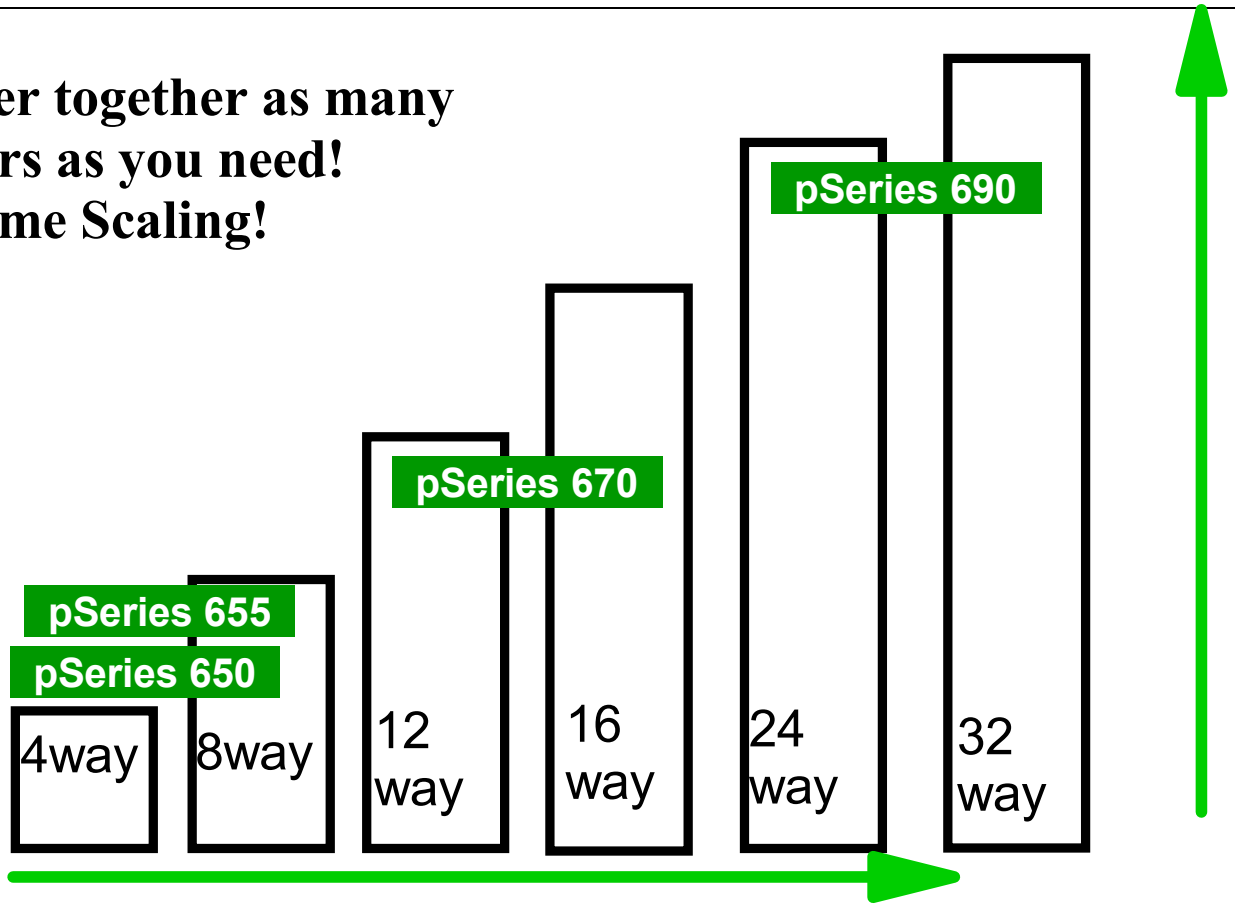


Multiple SMP servers working as a "single image" database for one query.

© Copyright IBM Corporation 2002

DB2 UDB and pSeries – Scaling Up and Out

- **Cluster together as many Servers as you need!**
- **Extreme Scaling!**



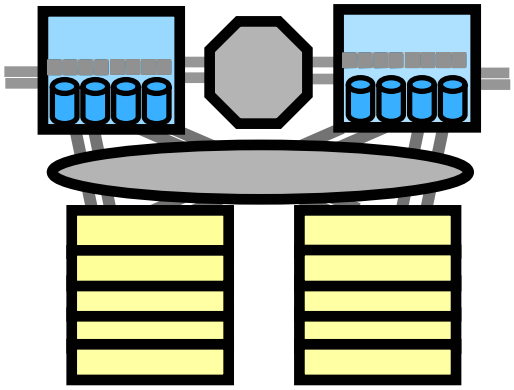
DB2 exploits all Server Sizes!

© Copyright IBM Corporation 2002

BCU – Basic Configuration Unit

- Pair of Servers
  - CPUs
  - Memory
  - I/O
- Storage
  - Balanced to Server I/O
  - SAN for ease of management
- Database Partitions
  - Balanced to Server CPUs and I/O
- Interconnect
  - Scalable, Redundant paths, non-blocking switch
- High Availability
  - Twin tailed disk resources
  - Redundant “Hot Swap” Components
  - Automated Failover and Detection software

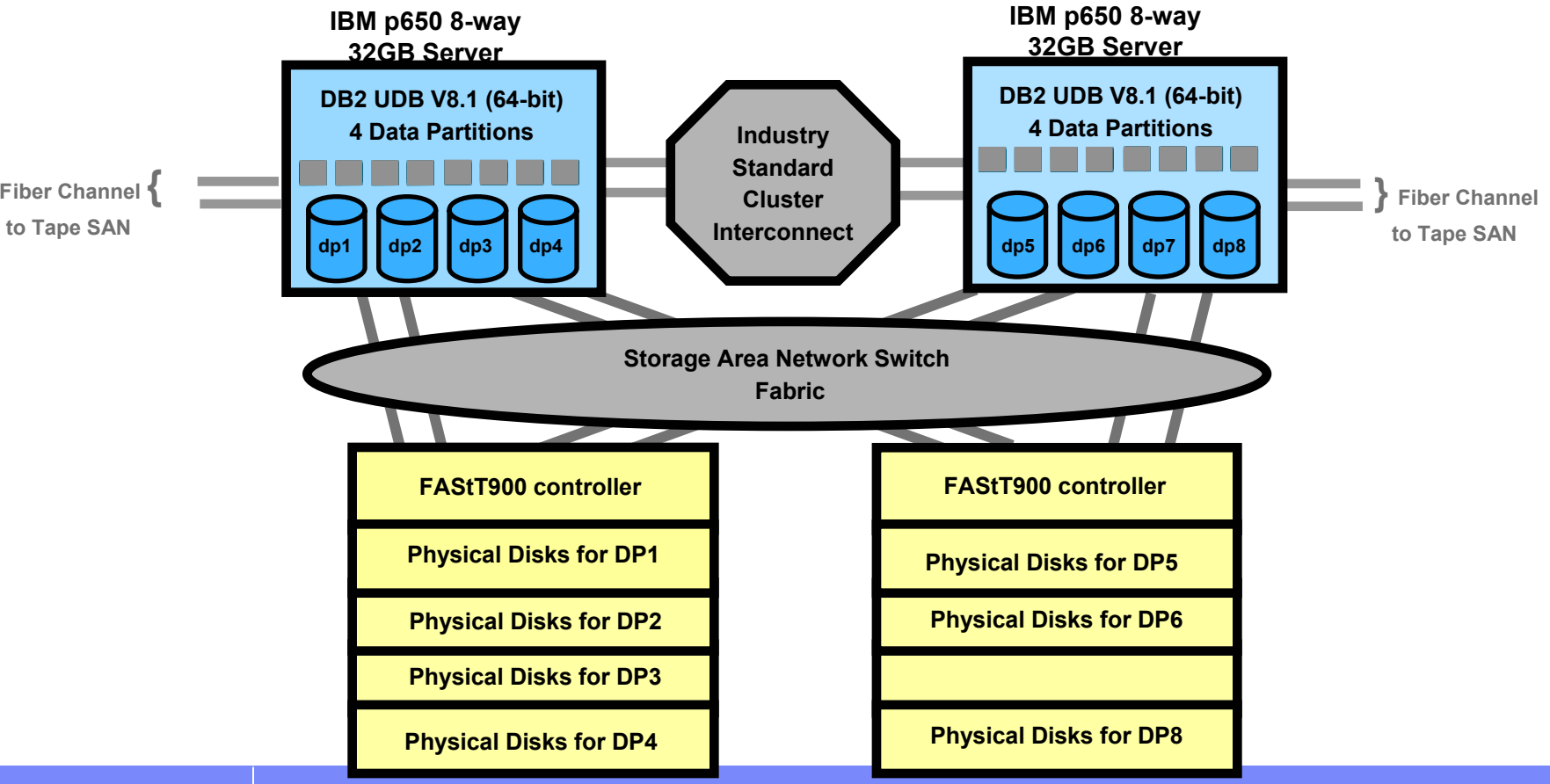
Defines Standard Building Blocks for Scalable System



© Copyright IBM Corporation 2002

BCU – Basic Configuration Unit

A "system architecture" built with scalable and highly available building blocks that provide predictable performance with growth



© Copyright IBM Corporation 2002

Example – BCU and Scalability

- Customer project, Spring 2003

- Scalability Objectives:
  - Show that the POC query throughput scales linearly as IBM hardware and DB2 partitions are added.
- Query Set 1
  - Total of 2351 queries
  - Set 1a: Start queries by timestamp
  - Set 1b: Start queries to maintain concurrency of 86
- Query Set 2
  - Total of 2042 queries (maintain concurrency of 86)

© Copyright IBM Corporation 2002

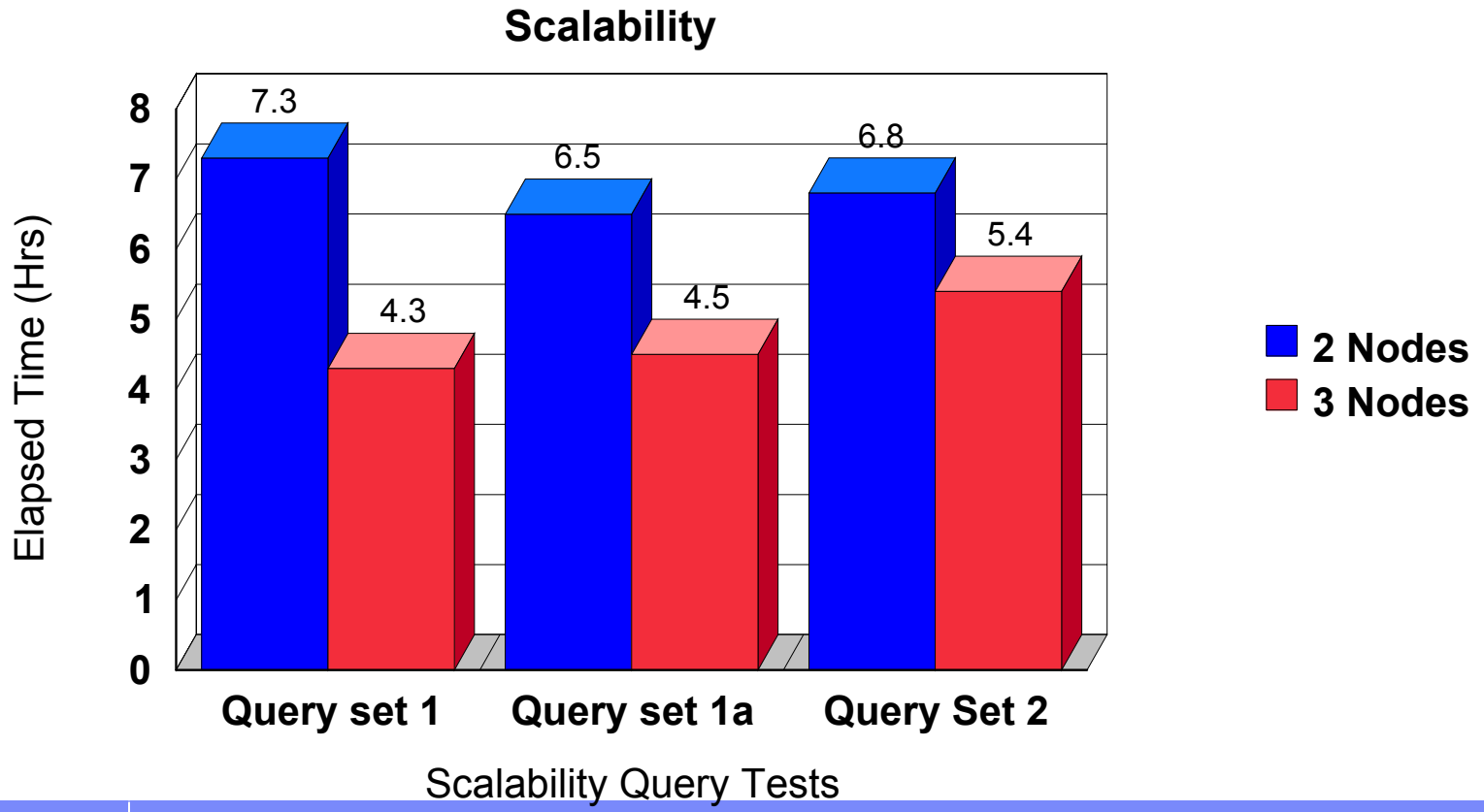
Example – BCU and Scalability

- Two hardware configurations
  - 2 Node system (2 8way P650s) – One BCU
  - 3 Node system (3 8way P650s) - One and one half BCU
- Test on 2 Node system
- Test on 3 Node system
- Execute same tests exactly on each system
  - Measure Scalability

© Copyright IBM Corporation 2002

Example – BCU and Scalability

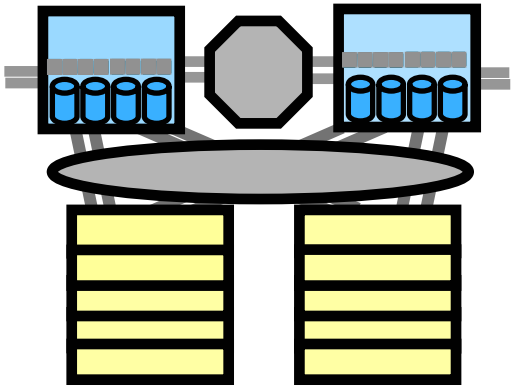
- Query Set 1a Throughput scaling is 112.1 %
- Query Set 1b Throughput scaling is 96.7 %
- Query Set 2 Throughput scaling is 84.5%



© Copyright IBM Corporation 2002

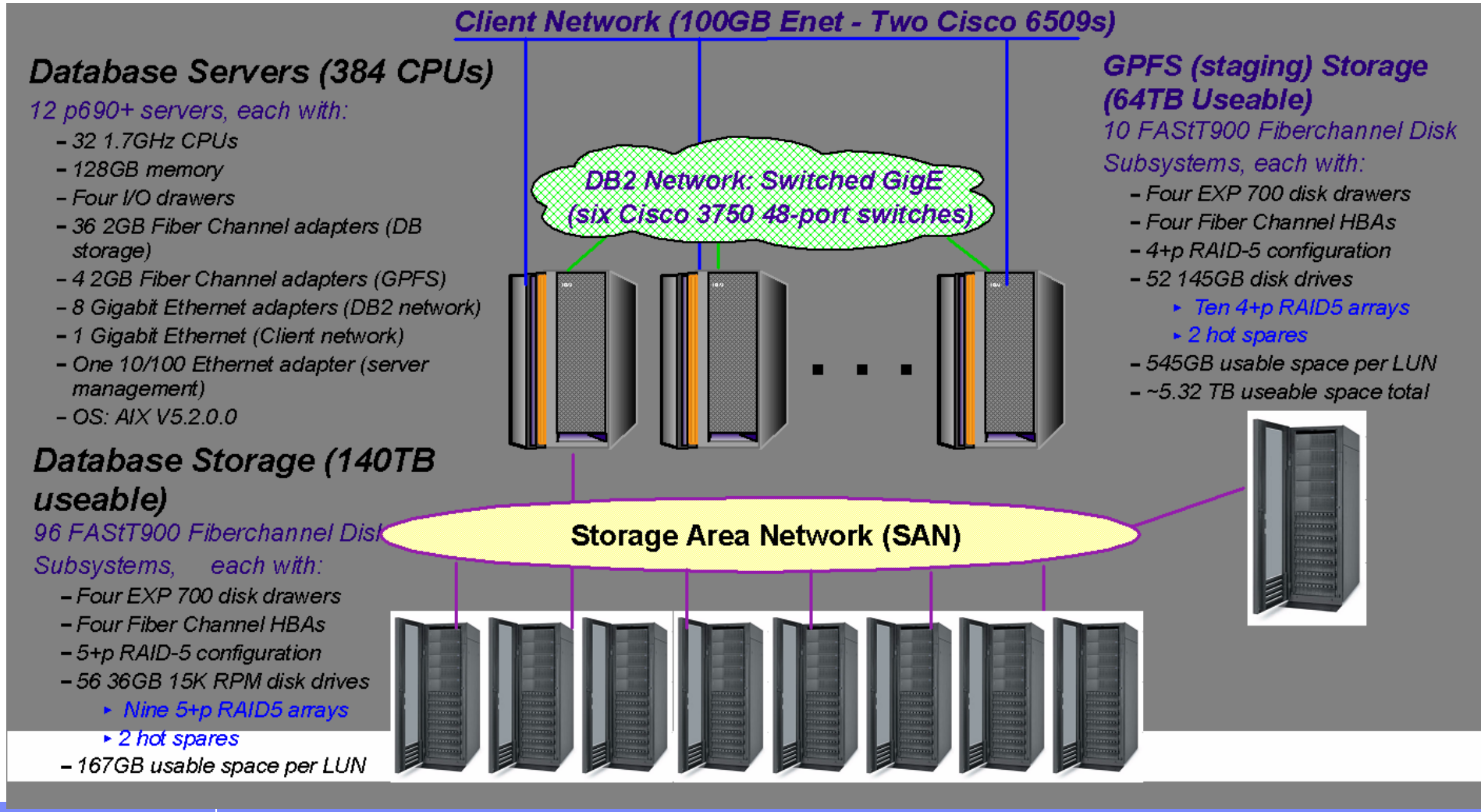
Scalability – BCU for a VERY large databasse

- **32 Terabytes "Raw" Data**
  - 97 Terabytes including index, temp, logs
- **Concurrent Query, Insert, Delete**
  - 800 queries per hour
  - Peak of 1.6m inserts and deletes per second
- **DB2 UDB V8.1**
  - Exploitation of MDC
    - Latitude, Longitude, Time
  - Binary Large Objects stored within DB2
- **12 P690 Servers**
  - 384 Processors
  - Gigabit Ethernet interconnect
- **FAStT 900 Storage**
  - 140TB usable after Raid 5



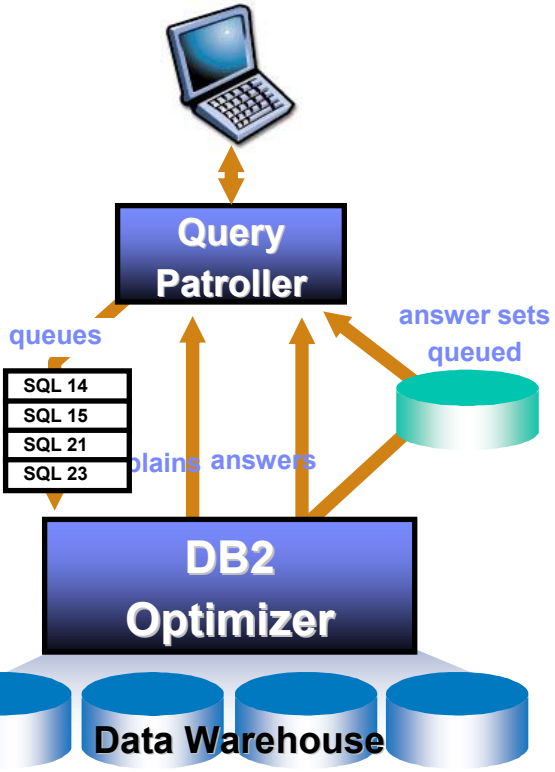
© Copyright IBM Corporation 2002

Scalability –BCU for VERY large Database



© Copyright IBM Corporation 2002

Enhance Scalability with Query Patroller



- Query Patroller will give you the following capabilities:
  - Predictive Governing
    - regulate data access **before** entering the system
  - Priority queuing
    - Assign different priorities to different users **and** build up job queues to make them run efficiently
  - Query monitoring
    - Real time view and access to queries
  - Historical analysis
    - Over a period of time, view what data is being accessed and by wh
    - Over a period of time, view what data is not being accessed.
  - Result caching
    - Optionally save query results to be used over again

© Copyright IBM Corporation 2002

Customer experience with Query Patroller

- **Real Customer workload - DB2 V8.1 on Cluster of 4 IBM P650 servers**
  - Set up 5 queues
    - Shortest queries get 70% of system
    - Start 2350 queries at once
    - Query Patroller set up to manage based on cost
- **Overview Results**
  - Shortest 600 queries completed in 10 minutes
    - Previous tests required over 1 hour
    - Short queue emptied in less than 45 minutes
    - Best overall throughput achieved for all queries
- **Lessons Learned**
  - Better throughput/less contention – **Enhanced Scalability**
  - Substantial improvement in Workload Management using Query Patroller V8.1

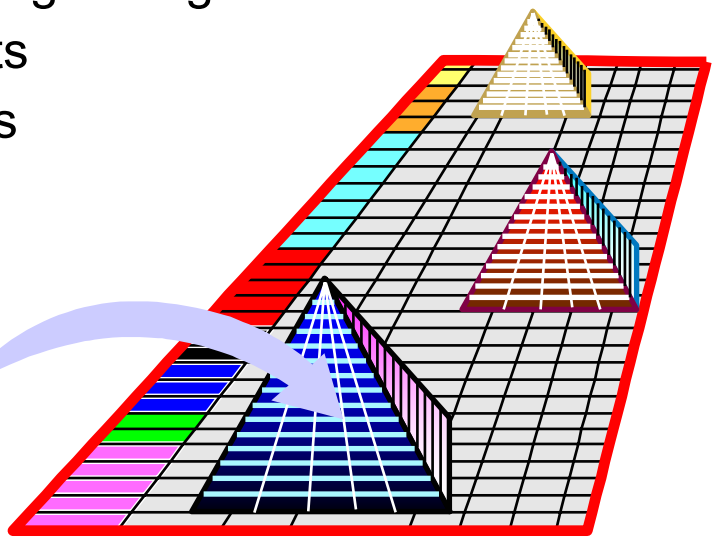
© Copyright IBM Corporation 2002

Enhance Scalability with Cube Views

- Wizard to design & generate "cube views"
  - aka multidimensional MQTs
- Multi-dimensional metadata inside DB2
  - Import/export for metadata exchange bridges
- Use Standard SQL for "rollup" reports
- DB2 optimizer query rewrite to MQTs
  - fast performance
- Included in DB2 Warehouse Edition

Cube Views

aka multidimensional MQTs



© Copyright IBM Corporation 2002

Scalability: Architecting for Growth

- **Scalable Database and Platform**
  - DB2 UDB and IBM Servers
- **"Basic Configurable Unit"**
  - System components must be added in proportional manner
  - CPU, Memory, I/O, interconnect
- **Tools to enhance and manage Scalability**
  - Query Patroller
  - Cube Views

© Copyright IBM Corporation 2002