DB2 Data Management Software
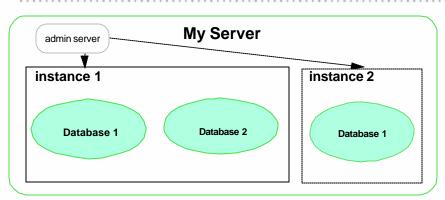
IBM

e business software

DB2 Version 8 Basic Architecture

DB2 Quickstart Education
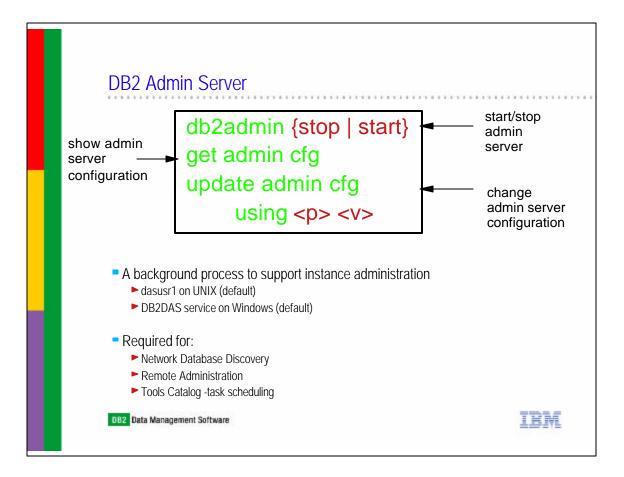Maintained by Paul Yip (ypaul@ca.ibm.com)

February 2003

IBM Software Group

The DB2 Server Picture

My Server

admin server

instance 1

Database 1

Database 2

instance 2

Database 1

- One physical machine has:
  - One Admin Server
  - One or multiple instances
- Each instance (database manager)
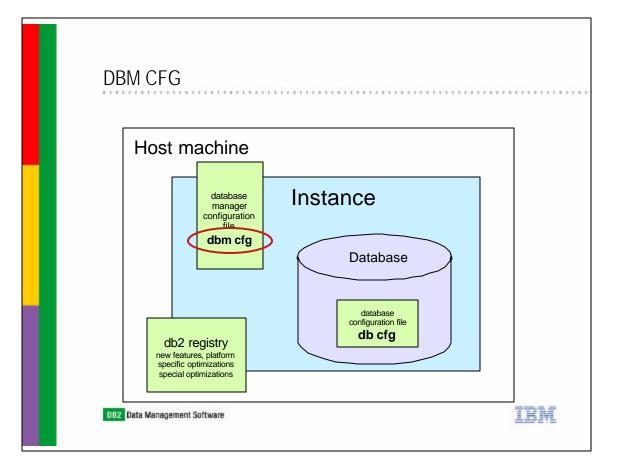 - is a logical server context
  - can have one or many databases

DB2 Data Management Software

IBM

DB2 Admin Server

db2admin {stop | start} ← start/stop admin server

show admin server configuration → **get admin cfg**

update admin cfg using <p> <v> ← change admin server configuration

- A background process to support instance administration
  - ► dasusr1 on UNIX (default)
  - ► DB2DAS service on Windows (default)

- Required for:
  - ► Network Database Discovery
  - ► Remote Administration
  - ► Tools Catalog -task scheduling
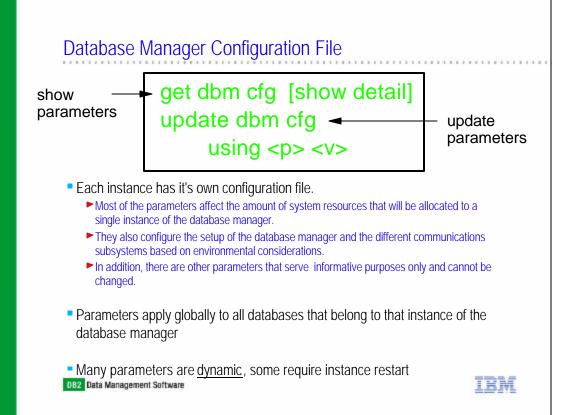
DB2 Data Management Software

IBM

Instance Concepts

- An instance, also known as Database Manager
  - ► Manages the system resources assigned to it
  - ► Can be independently stopped or started
  - ► has its own set of processes
  - ► Has its own configuration file, DBM CFG.
    - − security plan
    - −performance
    - −communications

- example: One instance for development, another for QA testing. Fully Independent.

DB2 Data Management Software

IBM

DBM CFG

Host machine

Instance

database manager configuration file
dbm cfg

Database

database configuration file
db cfg

db2 registry
new features, platform specific optimizations special optimizations

DB2 Data Management Software

IBM

Database Manager Configuration File

show parameters →

get dbm cfg [show detail]
update dbm cfg
using <p> <v>

← update parameters

- Each instance has it's own configuration file.
  - ► Most of the parameters affect the amount of system resources that will be allocated to a single instance of the database manager.
  - ► They also configure the setup of the database manager and the different communications subsystems based on environmental considerations.
  - ► In addition, there are other parameters that serve informative purposes only and cannot be changed.

- Parameters apply globally to all databases that belong to that instance of the database manager

- Many parameters are dynamic, some require instance restart

DB2 Data Management Software

IBM

Common Instance (DBM CFG) Parameters (cont..)

- DIAGLEVEL
  - ▶ Diagnostic log (db2diag.log) detail level

- NOTIFY LEVEL
  - ▶ Notification log (*.nfy) detail level

- MON_HEAP_SZ
  - ▶ Database performance monitor heap size
  - ▶ Default generally too small - recommend 128 or higher

- INTRA_PARALLEL
  - ▶ Enable/disable intra-partition parallelism
  - ▶ Keep disabled for OLTP workloads

- SVCENAME
  - ▶ TCP/IP Service Name, as mapped in /etc/services file
  - ▶ Must be set to support remote communications

DB2 Data Management Software

IBM

Common Instance (DBM CFG) Parameters

- MAXAGENTS
  - ▶ Maximum number of database manager agents available at any given time to accept application requests. This value limits the memory usage of the instance, because each additional agent requires additional memory.

- NUM_POOLAGENTS
  - ▶ A guideline for how large the agent pool should grow. The pool can save the cost associated with allocating an agent.
  - ▶ Not significant in environments where connection pools are used

- NUMDB
  - ▶ Limits the maximum number of concurrently active databases.
  - ▶ Value should reflect reality as closely as possible.

- SHEAPTHRES
  - ▶ Controls the total amount of memory for sorting available across the entire instance for all sorts
  - ▶ start with roughly 10x SORTHEAP (DB CFG)

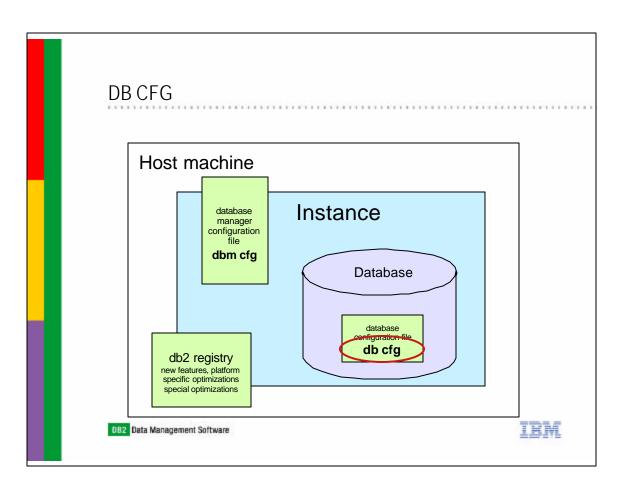**DB2** Data Management Software

IBM

Databases Concepts

- A collection of tables and related objects.

- Each database includes
  - ► system catalog tables (database meta data)
  - ► a database configuration file (db cfg)
  - ► a recovery log

- Databases are isolated from each other

DB2 Data Management Software

IBM

DB CFG

Host machine

Instance

database manager configuration file
dbm cfg

Database

database configuration file
db cfg

db2 registry
new features, platform specific optimizations special optimizations

DB2 Data Management Software

IBM

Database Configuration File

show parameters →

get db cfg [show detail]
update db cfg for <dbname> ← update parameters
using <p> <v>

- Parameters that affect the database.
  - if there are two databases in an instance, there will be <u>two</u> database configuration files

- Parameters include:
  - location of database logs
  - defaults for query optimization
  - size of memory for lock management
  - etc..

- Many parameters are dynamic. Others may require all applications disconnect. Changes to take effect on first re-connection

DB2 Data Management Software

IBM

Common Database (DB CFG) Parameters

- SORTHEAP
  - Defines the maximum number of memory pages for sorts
  - Set this parameter big enough to avoid sort overflows, but not so big that the optimizer attempts too many sorts

- MAXAPPLS
  - Specifies the maximum number of applications (local and remote) that can be connected to the database

- AVGAPPLS
  - Specifies the expected number of concurrently executing (large query) applications
    - affects how much buffer pool the optimizer expects to have available

- NEWLOGPATH
  - change the location of the db2 logs.
  - recommendation: always move logs onto separate, dedicated disks

- MIRRORLOGPATH
  - path to mirrored logs

IBM

Common Database (DB CFG) Parameters (cont...)

- CHNGPGS_THRESH
  - Specifies the percentage of the buffer pool which must be "dirty" before page cleaners are triggered
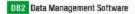 - frequent, small writes vs. infrequent, large writes

- NUM_IOCLEANERS
  - Specifies the number of asynchronous page cleaners
  - Asynchronous page cleaners can save an agent from having to write changed data to the disk when it needs to bring data into the buffer pool

- NUM_IOSERVERS
  - Specifies the number of I/O servers for a database
  - I/O servers perform prefetching which can reduce the amount of I/O that agent has to perform to execute a query

- LOGBUFSZ
  - Specify buffer size for database logging
  - defaults generally too small

DB2 Data Management Software

IBM

Common Database (DB CFG) Parameters (cont...)

- UTIL_HEAP_SZ
  - usually increased when tuning for utilities such as LOAD

- LOCKLIST
  - Memory storage size for lock list

- MAXLOCKS
  - Maximum percent of LOCKLIST that can be used by an application

- SHEAPTHRES_SHR
  - Sortheap threshold for shared sorts (for intra-partition parallelism)

- SORTHEAP
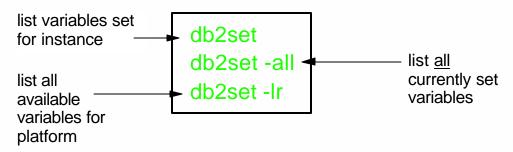 - maximum size of sort heap that a db2 agent can allocate

- LOCKTIMEOUT
  - maximum time an application will wait on locks before rolling back
  - Default is wait forever. Therefore, should be changed to reflect needs

DB2 Data Management Software

IBM

DB2 Registry

list variables set
for instance → **db2set**

db2set -all ← list <u>all</u>
currently set
variables

list all
available
variables for
platform → **db2set -lr**
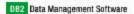
- A place to specify environment or platform specific settings
  - ► e.g. settings for RAID, optimizer paths not yet fully integrated, communication protocols, environment specific performance enhancements

- Settings can be applied to various levels
  - ► environment
  - ► instance
  - ► global (all instances)

DB2 Data Management Software

IBM

Database Objects

- Table
  - ► A collection of data logically arranged in columns and rows
- View
  - ► A representation of the data in tables
- Index
  - ► An ordered set of keys each pointing to a row in a table
- Schema
  - ► All DB2 UDB objects have a two part name and the schema is the first half of that name.
  - ► A schema is an identifier that helps group database objects together.
  - ► A schema can be owned by an individual who can then control access to the objects within it.
  - ► Schemas can be either implicit or explicit when accessing an object.
- Database Application Objects
  - ► Sequences
  - ► Triggers
  - ► User Defined Functions (UDFs)
  - ► Stored Procedures

DB2 Data Management Software

IBM
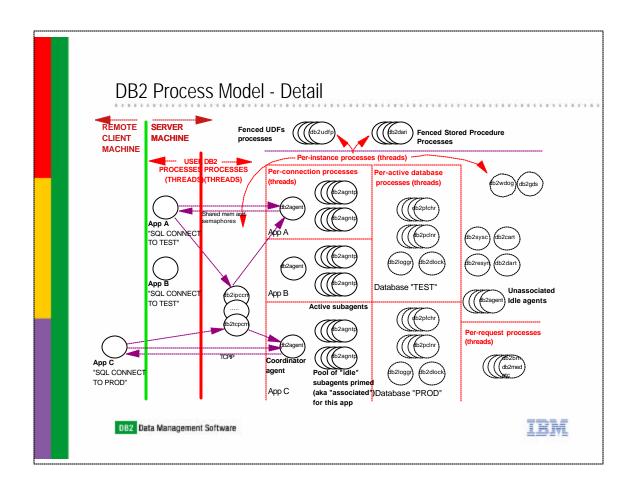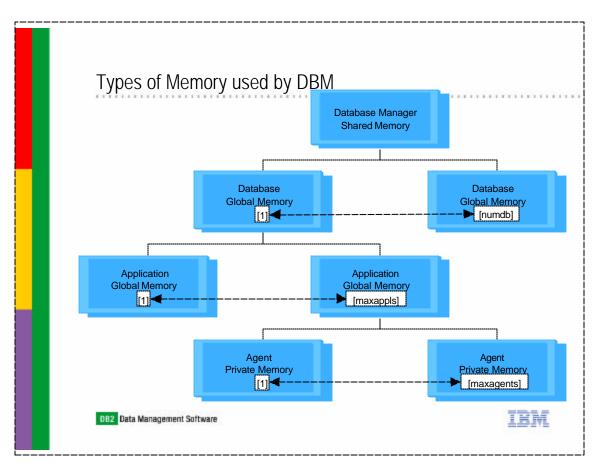
DB2 Object Heirarchy

```
                        ┌─────────────┐    ┌───────────────────────┐
                        │ Instance 1  │────│ dbm configuration file │
                        └─────────────┘    └───────────────────────┘

                                  ┌ Catalog ┐                      ┌ Catalog ┐
          Database                                Database
┌───────┐    1         ┌──────┐                     2        ┌──────┐
│Lock List│───────     │ Logs │                              │ Logs │
└───────┘             └──────┘                              └──────┘
                    ┌────────────────┐                    ┌──────────────┐
                    │ DB config file │                    │ DB config file│
                    └────────────────┘                    └──────────────┘
                                                              ┌──────────┐
                                                              │ Lock List │
                                                              └──────────┘

  ┌ Table1 ┐  ┌ Table2 ┐  ┌ Table3 ┐           ┌ Table2 ┐

  ┌ Index1 ┐                                    ┌ View1 ┐

                  ┌ View1 ┐   ┌ Index1 ┐

  ┌ BLOBs ┐    ┌ View2 ┐         ┌ Index2 ┐

                ┌ View3 ┐        Table
                                 spaces
```

DB2 Data Management Software

IBM

DB2 Process Model - Detail

REMOTE CLIENT MACHINE **SERVER MACHINE**

Fenced UDFs processes db2udfp db2dari **Fenced Stored Procedure Processes**

USER PROCESSES (THREADS) **DB2 PROCESSES (THREADS)**

Per-instance processes (threads)

Per-connection processes (threads) **Per-active database processes (threads)**

db2agent db2agntp db2pfchr db2wdog db2gds

App A "SQL CONNECT TO TEST" Shared mem and semaphores App A db2pclnr db2sysc db2cart

App B "SQL CONNECT TO TEST" db2agent db2agntp db2loggr db2dlock db2resyn db2dart

App B Database "TEST" db2agent **Unassociated Idle agents**

db2ipccm ..... db2tcpcm **Active subagents**

App C "SQL CONNECT TO PROD" TCPIP db2agent **Coordinator agent** db2pfchr **Per-request processes (threads)**

db2agntp db2pclnr db2bm db2med etc

App C **Pool of "idle" subagents primed (aka "associated") for this app** db2loggr db2dlock Database "PROD"

DB2 Data Management Software IBM

Types of Memory used by DBM

Database Manager Shared Memory

Database Global Memory [1] ←→ Database Global Memory [numdb]

Application Global Memory [1] ←→ Application Global Memory [maxappls]

Agent Private Memory [1] ←→ Agent Private Memory [maxagents]

DB2 Data Management Software IBM

Memory Usage in DB2 UDB

	Database Manager Shared Memory (including FCM)	
Monitor heap (mon_heap_sz)		Audit buffer size (audit_buf_sz)

Database Global Memory

Utility heap (util_heap_sz)		Database heap (dbheap)
Backup buffer (backbufsz)	Bufferpools (buffpage)	Log buffer (logbufsz)
Restore buffer (restbufsz)	Extended memory cache	Catalog cache (catalogcache_sz)
Package cache (pckcachesz)	Locklist (locklist)	

Application Global Memory (app_ctl_heap_sz)

Agent Private Memory

Application heap (applheapsz)	Agent stack (agent_stack_sz)	Statistics heap (stat_heap_sz)	Sort heap (sortheap)
	DRDA heap (drda_heap_sz)	UDF memory (udf_mem_sz)	Statement heap (stmtheap)
Query heap (query_heap_sz)	Java heap (java_heap_sz)	Client I/O block (rqrioblk) (remote)	

Agent/Application Shared Memory

Application support layer heap (aslheapsz)
Client I/O block (rqrioblk) (local)

IBM

Connection Concentrator

Clients

DB2

- Connection Concentrator
  - ► Support large number of concurrently connected users
  - ► new concept of logical agents and database agents
  - ► n:m architecture

DB2 Data Management Software

IBM

Database Agents

- Communication between the database and client
  - ► Unix processes vs. Intel threads

- Connection Concentrator is enabled when:
  max_connections > max_coordagents (DBM CFG)

- Processes/threads
  - ► Logical agent
    - ▬ Represents a connected application to the database manager
    - ▬ Contains information and control blocks required by an application
    - ▬ Controled by MAX_COORDAGENTS (DBM CFG)
  - ► Worker agent
    - ▬ Carries out application requests
    - ▬ No permanent attachment to any particular application

DB2 Data Management Software

IBM

Query Parallelism

- Inter-query Parallelism
  - ▬ Multiple queries running at the same time

- Intra-query Parallelism
  - ▬ Intra-partition parallelism
  - ▬ Inter-partition parallelism

DB2 Data Management Software

IBM

Intra-partition Parallelism

SELECT...
FROM...

Data

Database Partition

DB2 Data Management Software

IBM

Inter-partition Parallelism (ESE with partitioning option)

SELECT...
FROM...

A query is divided into parts, each being executed in parallel.

| Data | Data | Data | Data |

Database Partition · **Database Partition** · **Database Partition** · **Database Partition**

DB2 Data Management Software

IBM

Inter and Intra partition parallelism !!

SELECT...
FROM...

Data   Data   Data   Data

DB2 Data Management Software

IBM