

# IDUG® 2004 – North America

## Tuning DB2 in the New Galaxy

Philip K. Gunning  
Gunning Technology  
Solutions, LLC

Session ID: E3  
May 10, 2004

Enabling Your  
On Demand DB2 World

## DISCLAIMER

- This material references numerous hardware and software products by their trade names. References to other companies and their trademarks are for information purposes only, and all trademarks are the properties of their respective companies
- DB2 is a registered trademark of IBM Corporation
- DB2 “Stinger” is a registered trademark of IBM Corp.
- AIX is a registered trademark of IBM Corporation
- UNIX is a registered trademark of The Open Group
- Linux is a registered trademark of Linus Torvalds

## Overview

- Current DB2 Galaxy
- Current DB2 Galaxy tuning points
- Transitioning to the New DB2 Galaxy
- The New DB2 Galaxy
- Key Elements
- Summary



# The Current Galaxy

Direct Attached Storage

INODE Locks

Onboard Disk

SAN

RAID-1

32-bit

SCSi-3

RAID-5

DB2

Static Parameters

Single Clustering Index

PCI

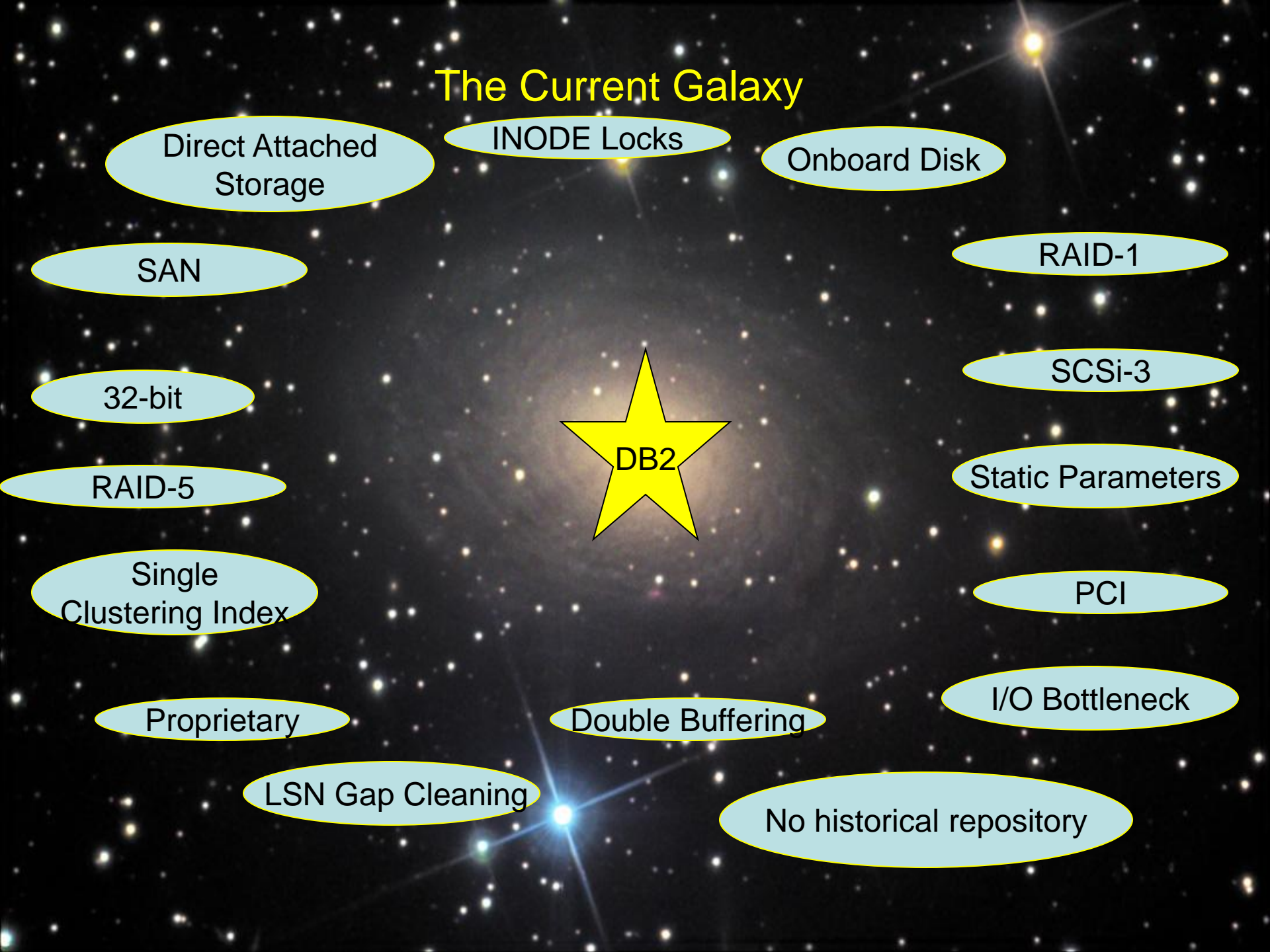
Proprietary

Double Buffering

I/O Bottleneck

LSN Gap Cleaning

No historical repository





## ONBOARD DISK

- Enterprise Database servers with onboard disk
- Limited capacity
- Limited Reliability
- Limited Availability



Server

## RAID-1

- Disk mirroring
- Writes data in two places
- With dual adapters, high availability
  - Either disk can fail, data is still accessible
  - With duplexing, disk controller can also fail
- High Performance
- *Twice the usual number of disks*



## RAID-5

- Collection of disks in an array (typically 6-7) with parity striping
- Parity data is striped across all disks in the array
- High-end storage solutions use hardware RAID-5 and large cache which reduces write penalty unless cache overrun
  - If cache overruns, performance degradation
- Provides high availability unless 2 disks fail at exact same time
  - Low probability but has happened



## RAID-10

- Implemented as a striped array whose segments are RAID-1 arrays
  - Same fault tolerance as RAID-1
  - High I/O rates achieved by striping RAID-1 segments
  - Has same overhead for fault tolerance as with mirroring alone



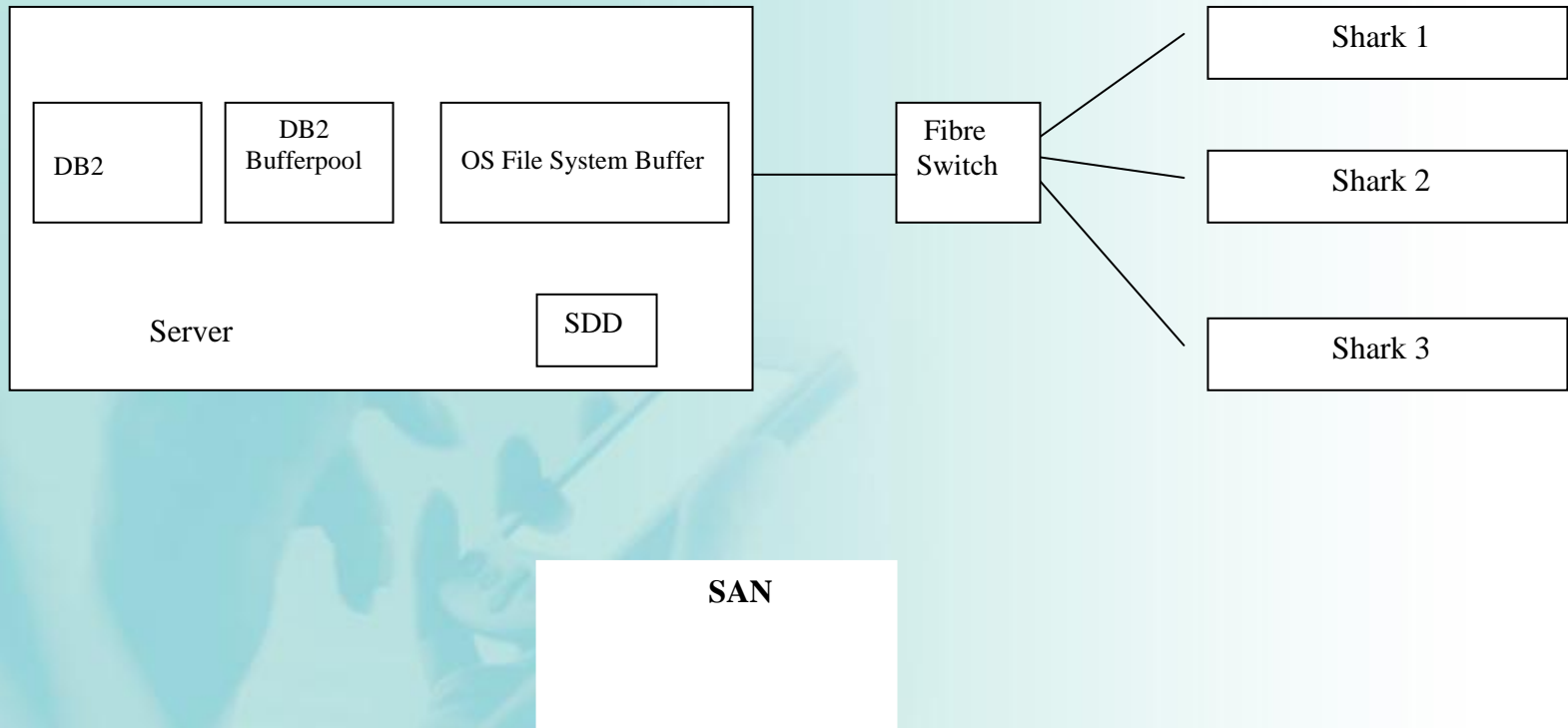


## I/O Bottleneck

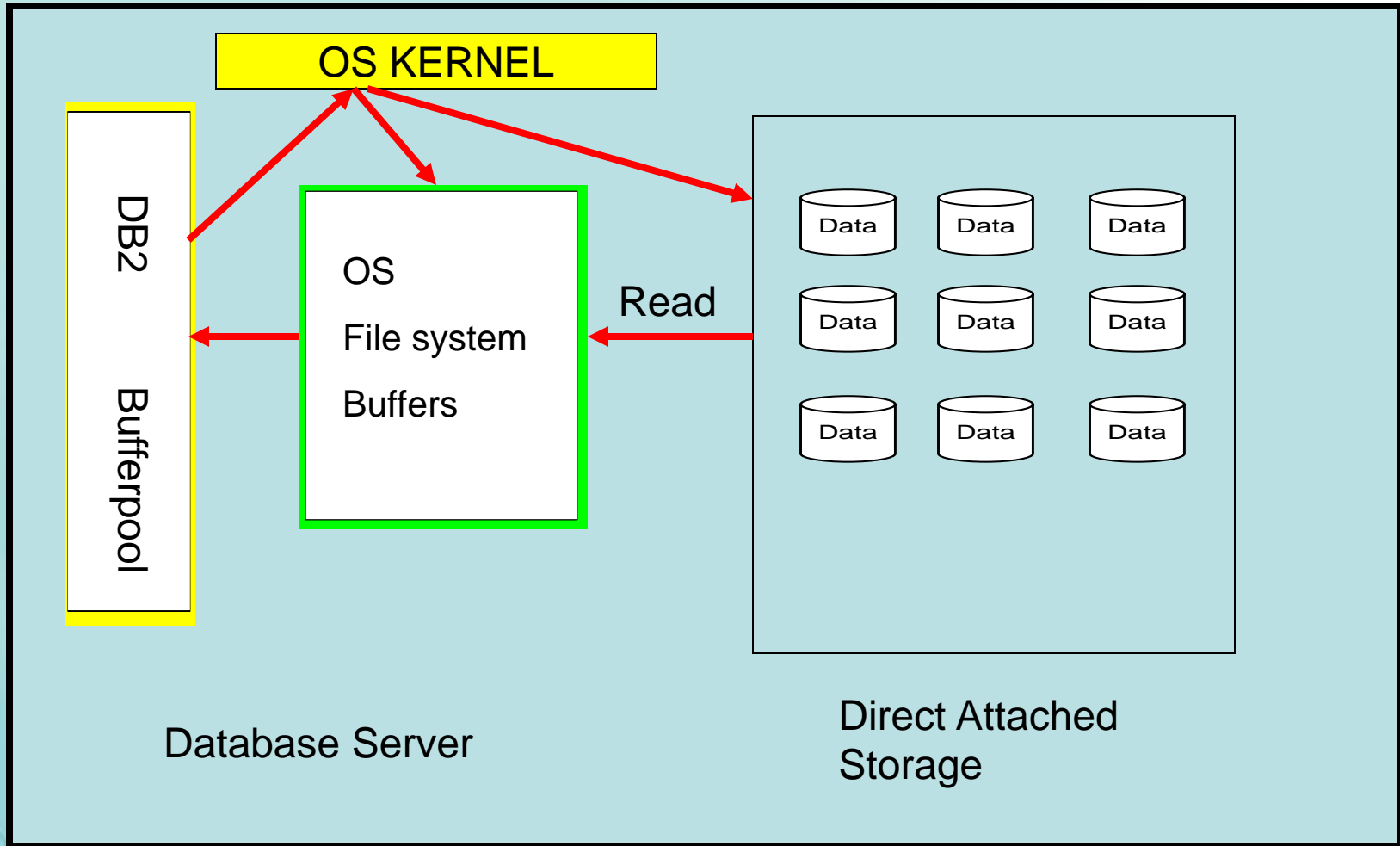
- Disks getting bigger but slower
- PCI bus saturated
- SCSI
  - Direct attached
  - Ongoing efforts
- PCI-Express
- PCI-X
- iSCSI
- TCP Offload Engines



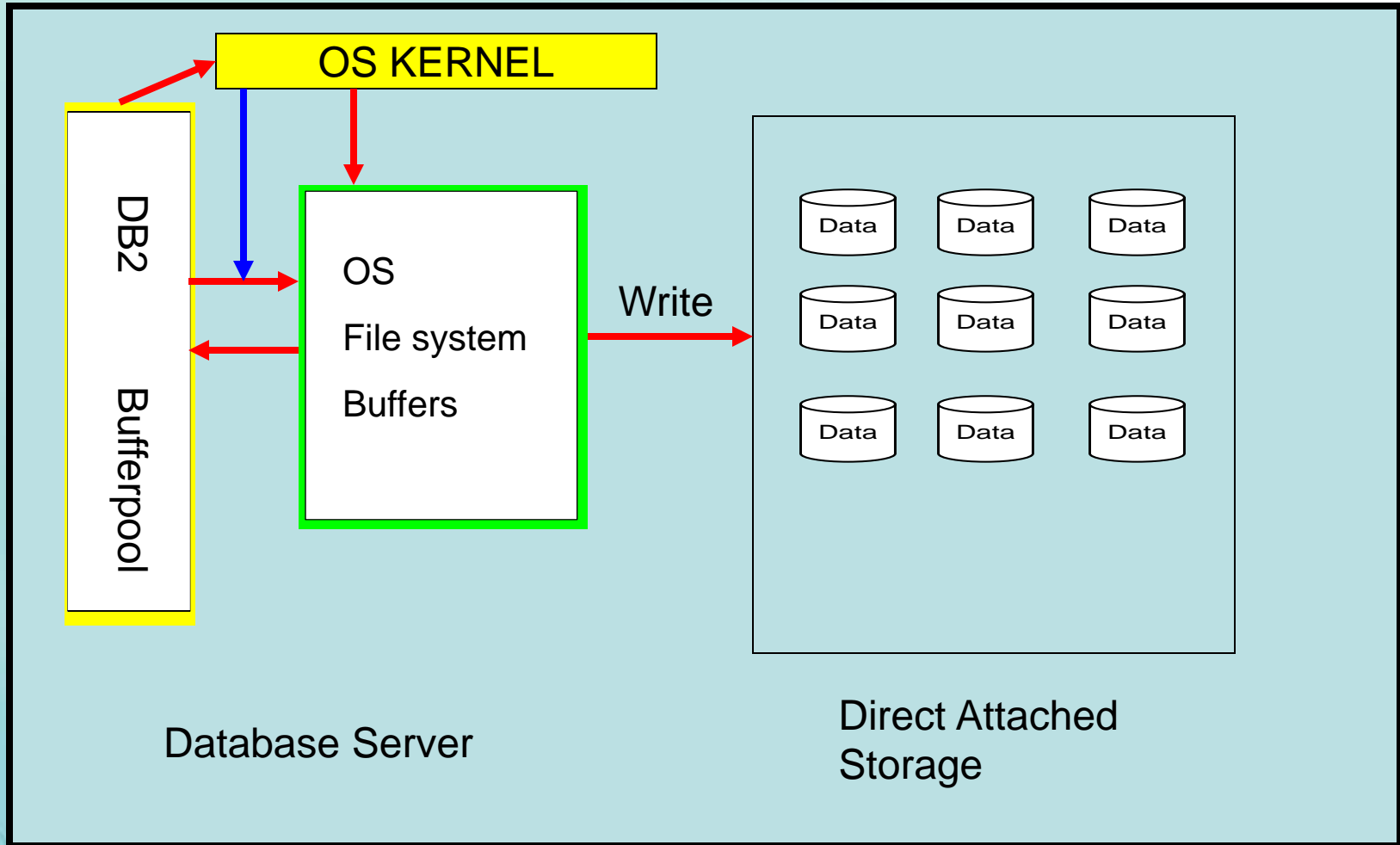
## Storage Architecture



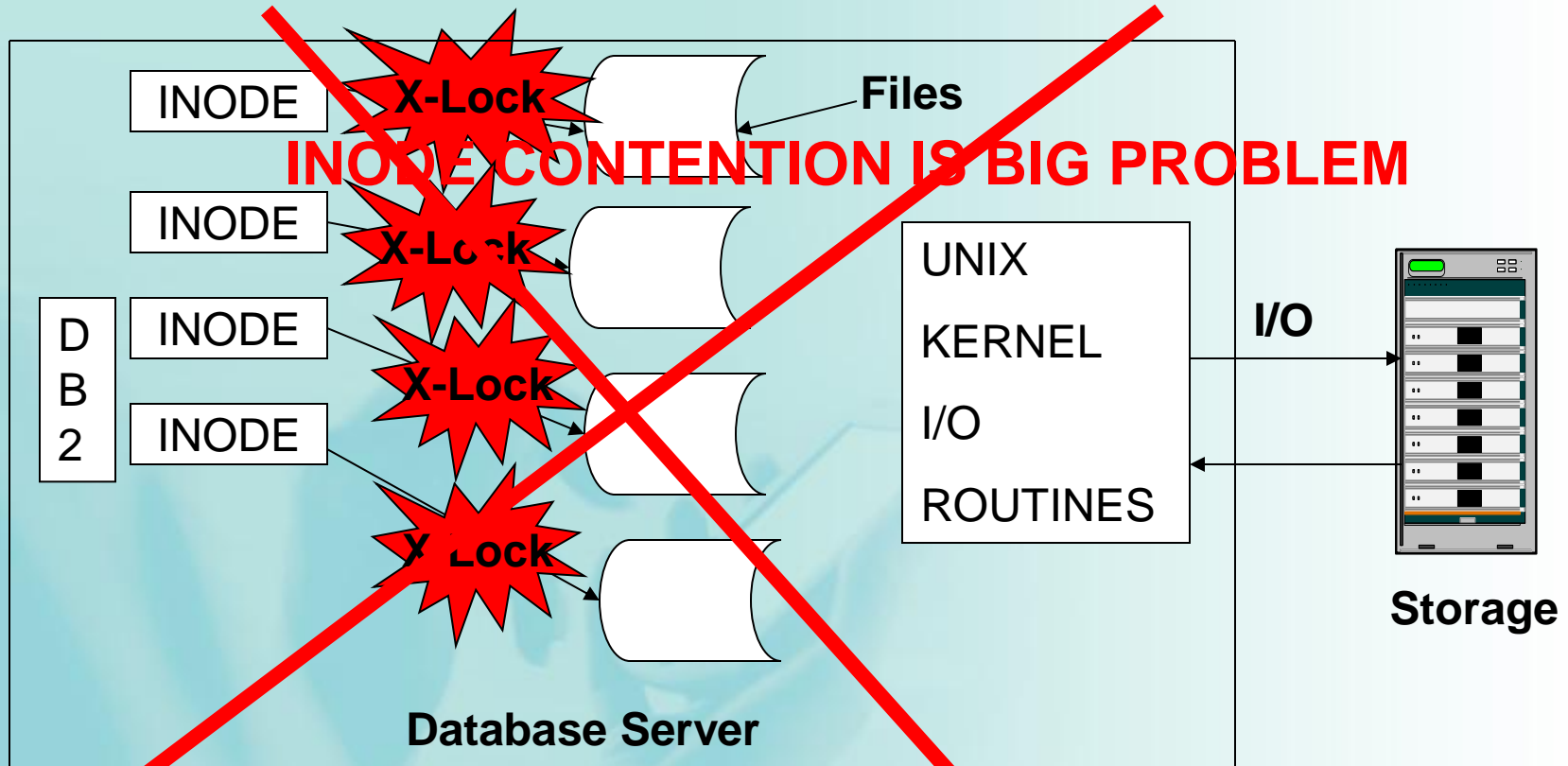
## Double Buffering – READ Operation



## Double Buffering – WRITE Operation

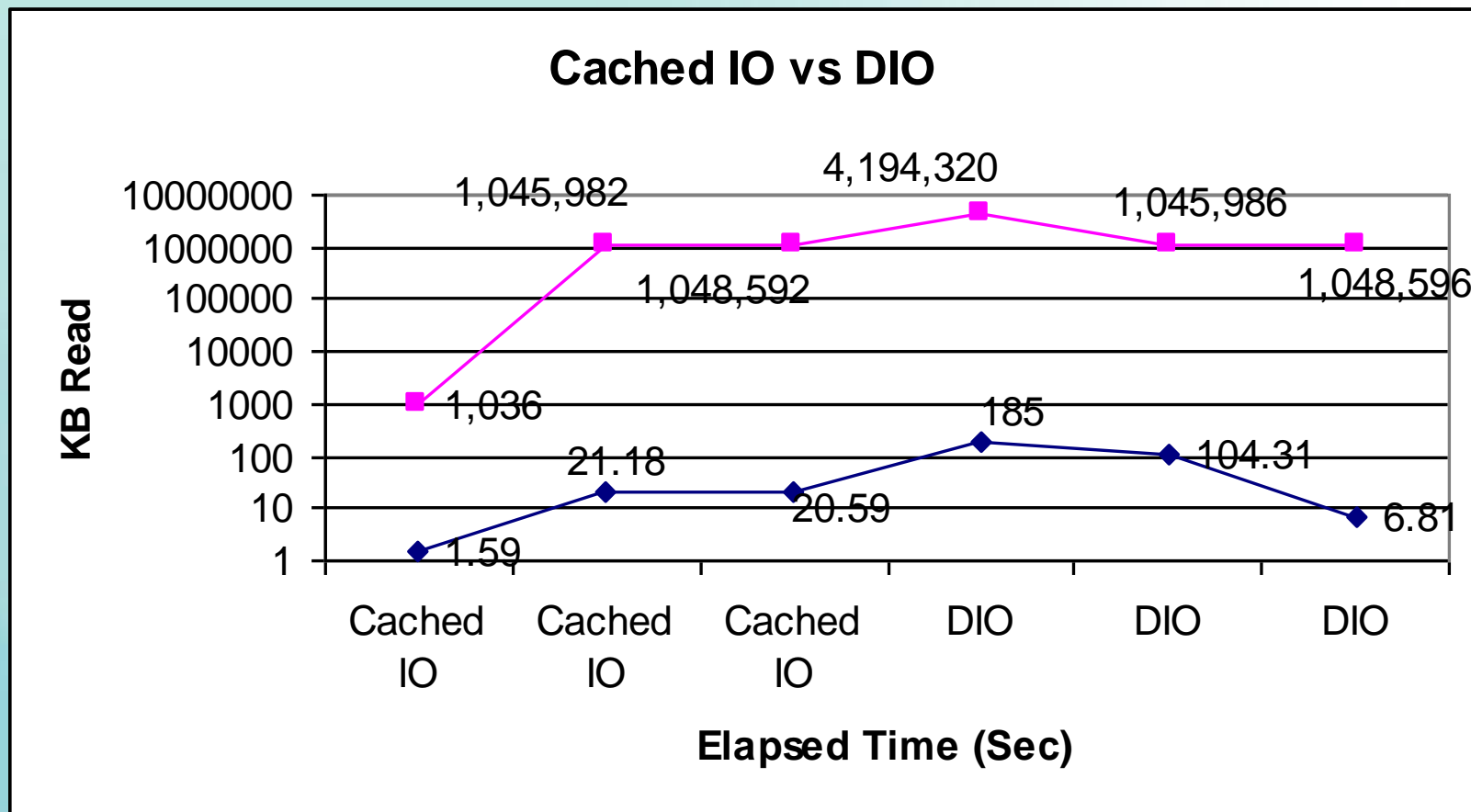


## Current Situation: INODE Lock Contention





## CACHED I/O vs DIRECT I/O



## LACK of HISTORICAL PERFORMANCE CAPABILITY

- Prior to V8.1, historical performance data only available via event monitors or snapshot to file
- Customized scripts required to format and insert into DB2 tables
- Third party vendor tools required
- Difficult to answer
  - What happened 3 hours ago?
  - What happened yesterday?
  - Has our database activity changed over the last day, week, or month?



# The New Galaxy

- At this point in time, we are kind of in between the current and new DB2 Galaxy
- Some have transitioned, many are in process and many haven't started yet
- So, just what is the new DB2 Galaxy?
- I'll attempt to define that for you in the next several slides

# The New Galaxy

Direct I/O

Concurrent I/O

Connection  
Concentrator

Balanced I/O

iSCSI

Write-to-table

DB2

RAID

64-bit

Dynamic

Information Integration  
(II)

Linux

Storage Provisioning

SQL Snapshot F(x)



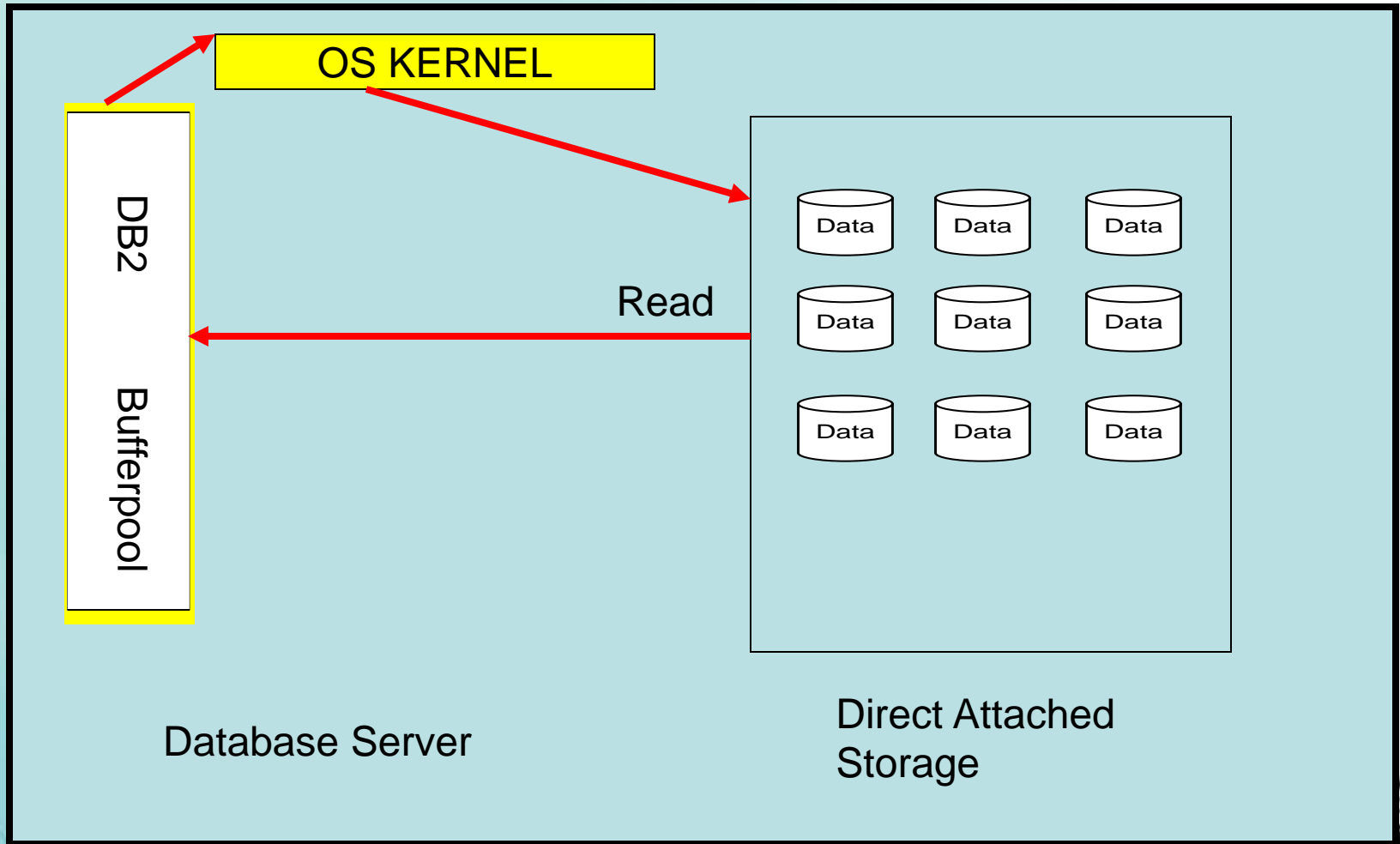


## DIRECT I/O

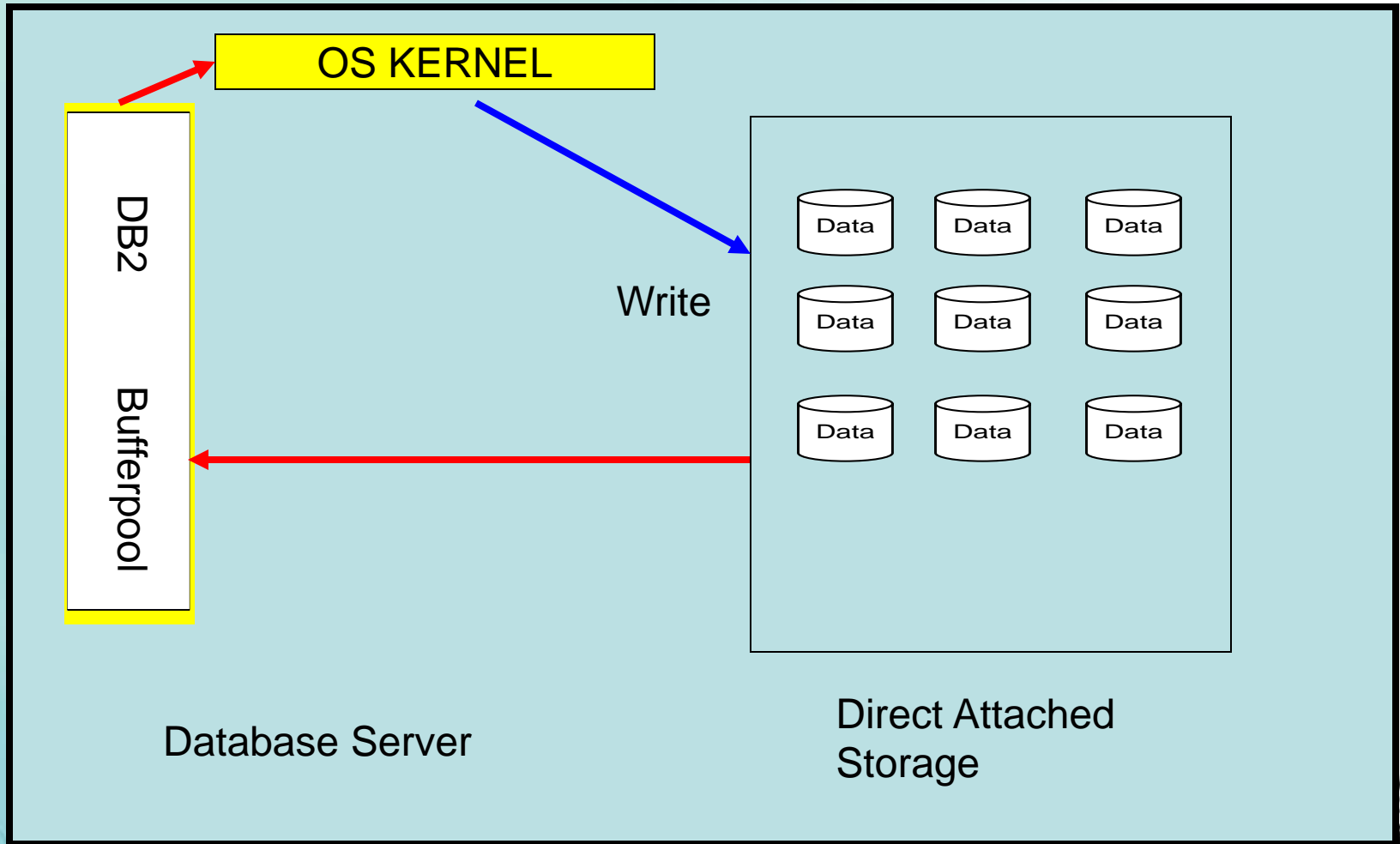
- On AIX, limited Direct I/O support added in DB2 V8.1.4 for all SMS tablespaces except the following:
  - temporary tablespaces
  - Long fields
  - LOBs
- Not yet available for DMS, possibly future release or Fixpak
- Previously available on Windows™ for SMS and DMS using the DB2NTNOCACHE registry variable
- Eliminates double-buffering
- Reduced path-length
- INODE contention still a problem (AIX)



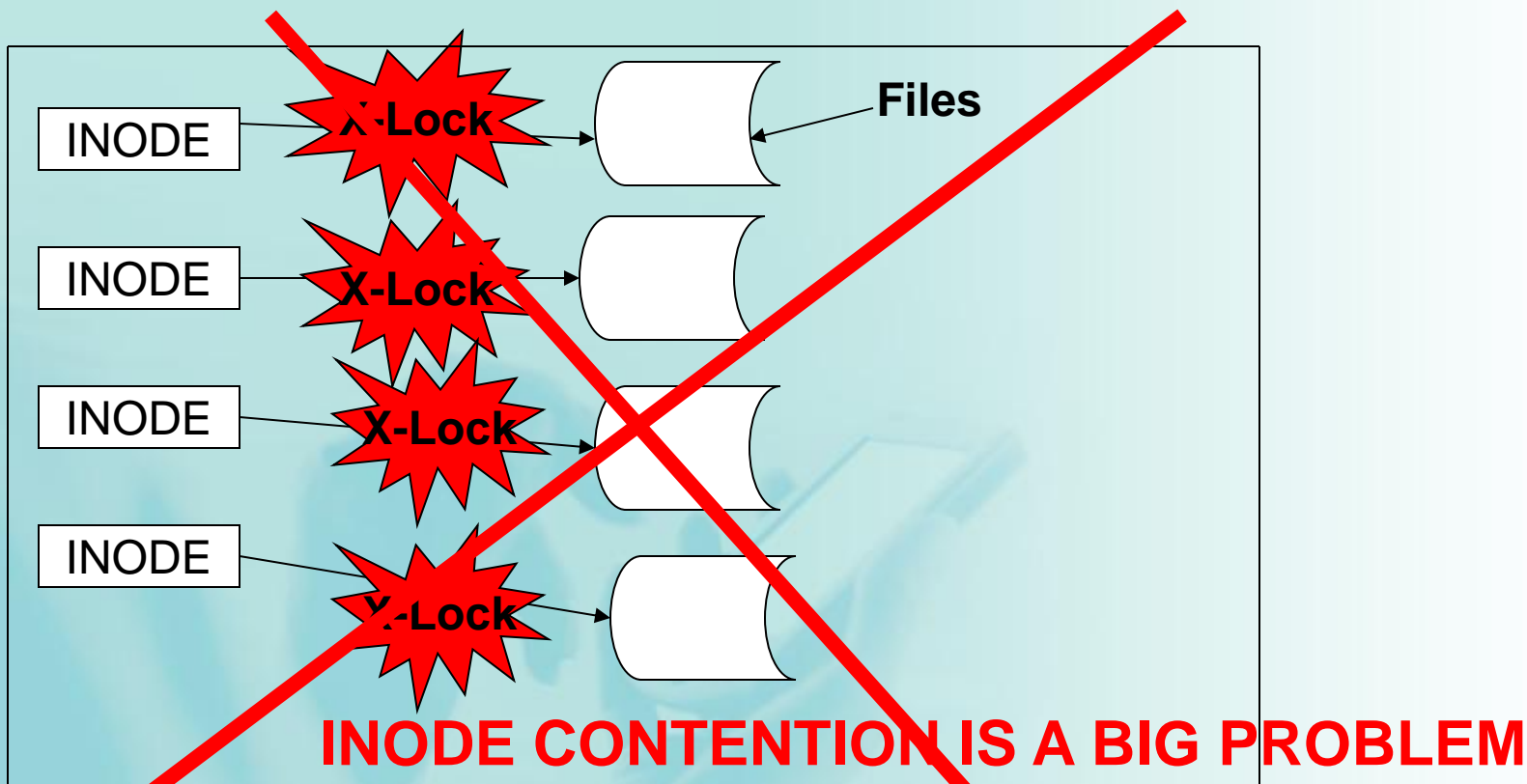
## DIRECT I/O – READ Operation



## DIRECT I/O – WRITE Operation



## DIRECT I/O : INODE Lock Contention



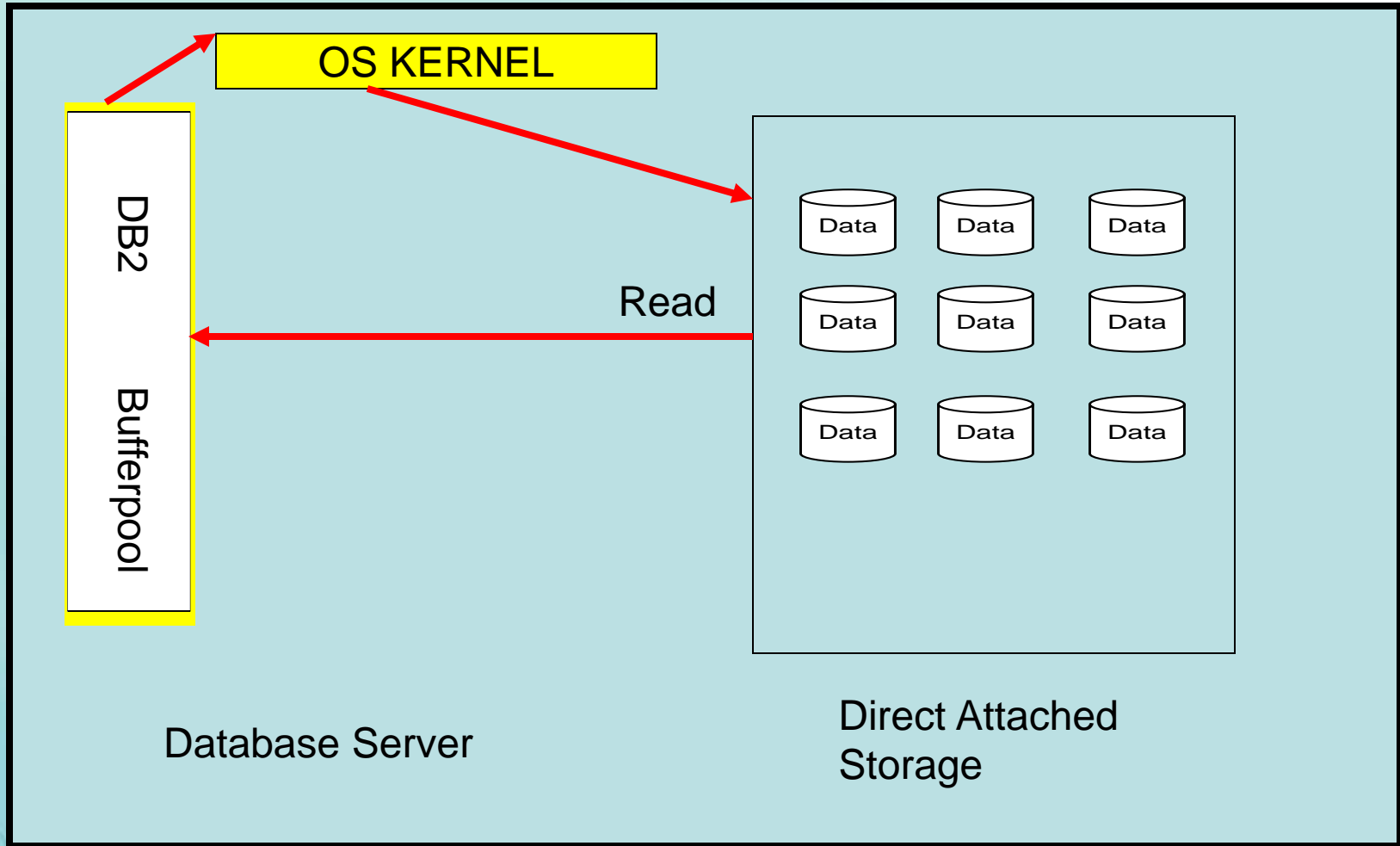
x-lock = read-shared, write-exclusive inode lock

## Concurrent I/O (CIO)

- Concurrent I/O not yet available for DB2 UDB as of V8.1.4 on AIX
- Is needed to reduce I/O bottleneck
- Uses Direct I/O implicitly, and is enabled using similar commands
- Reduces INODE lock problems by taking a read-shared lock, essentially relying on application locking semantics
  - Except when file is truncated or extended

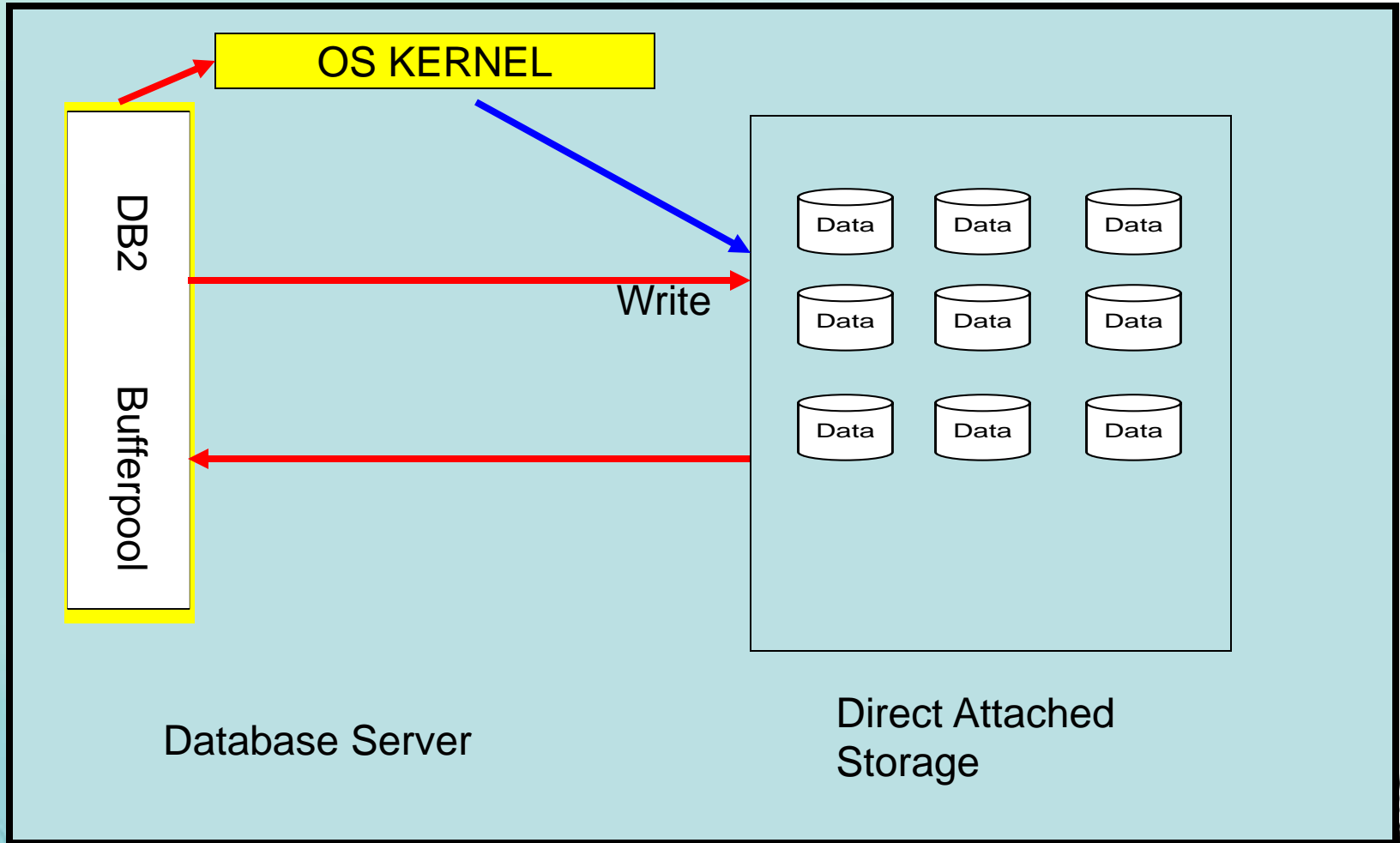


## CIO USES DIRECT I/O IMPLICITLY

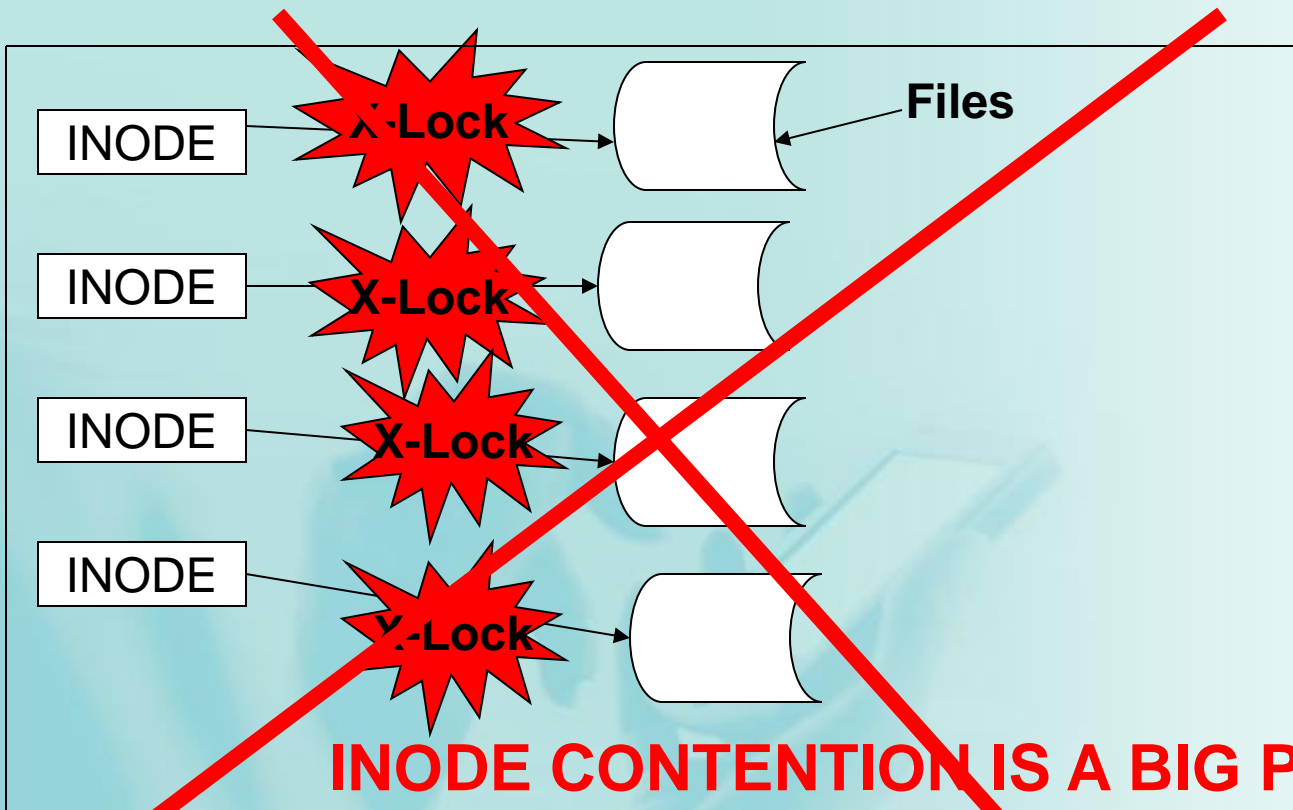




## CIO USES DIRECT I/O IMPLICITLY

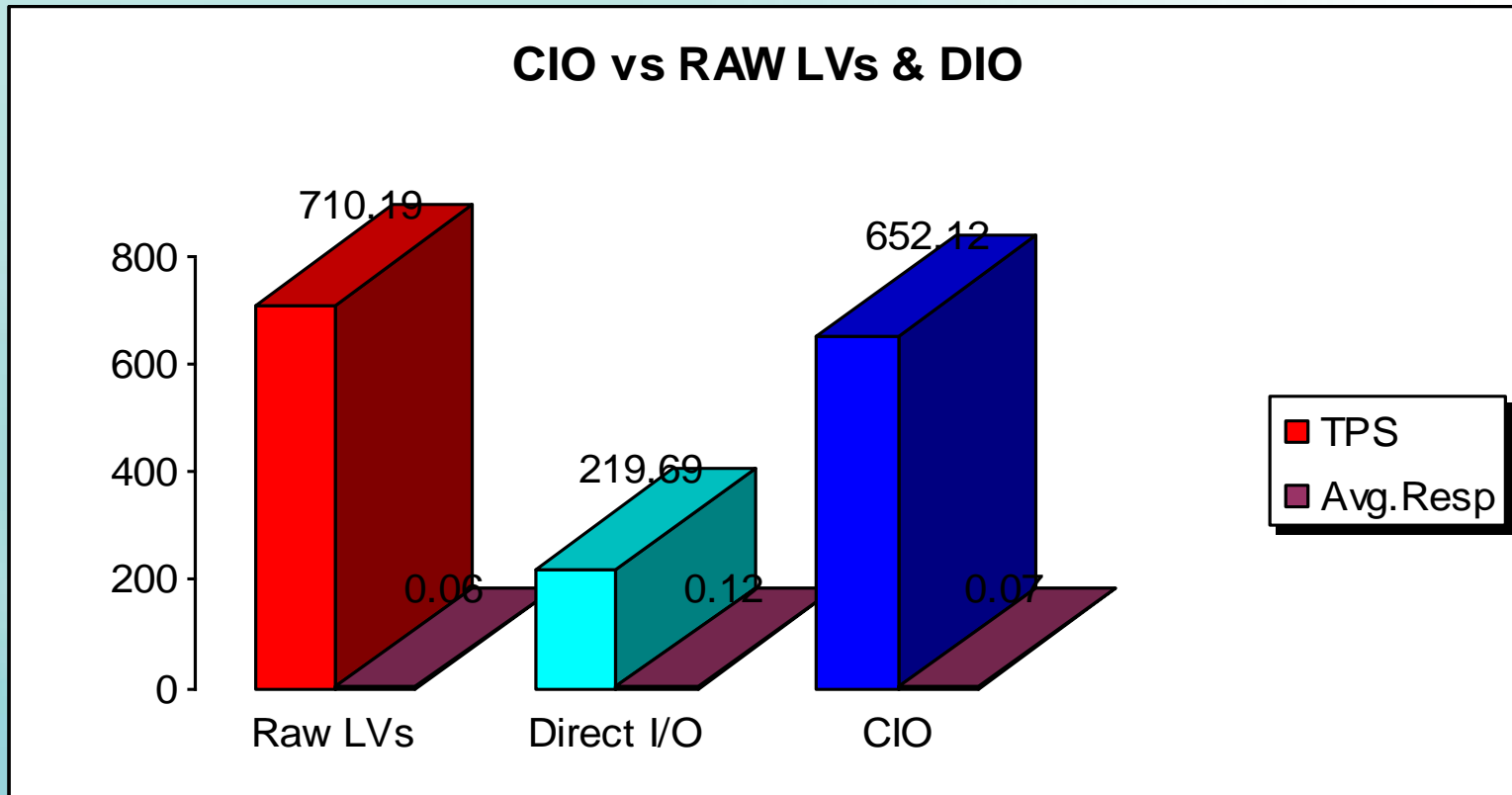


## Elimination of INODE Contention with CIO!



x-lock = read-shared, write-exclusive inode lock

## CIO Performance



## CIO I/O Performance Improvements

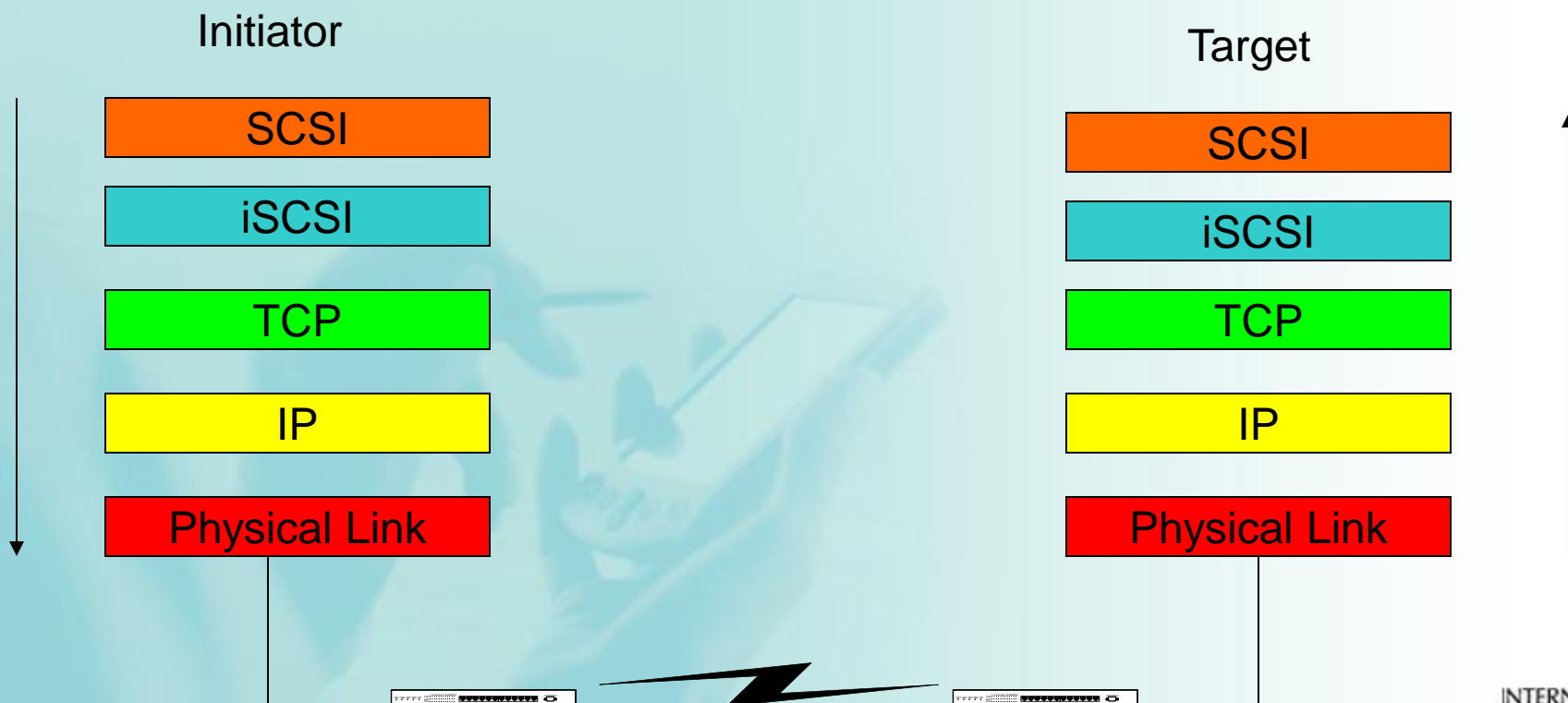
- CIO achieves approximately 92% of performance of RAW Logical Volumes without the complexity and management of RAW Logical Volumes
- 70% better than Direct I/O
- **This is a serious performance multiplier that you should keep on your radar screen!**



## iSCSI

- What is iSCSI?
- iSCSI and DB2
  - Enables companies that do not have SANs to implement SANs with iSCSI
  - Enterprise Storage consolidation
  - Recoup allocated but unused storage
- How does iSCSI affect DB2?
- The Future

## TCP/IP Encapsulates iSCSI Packet



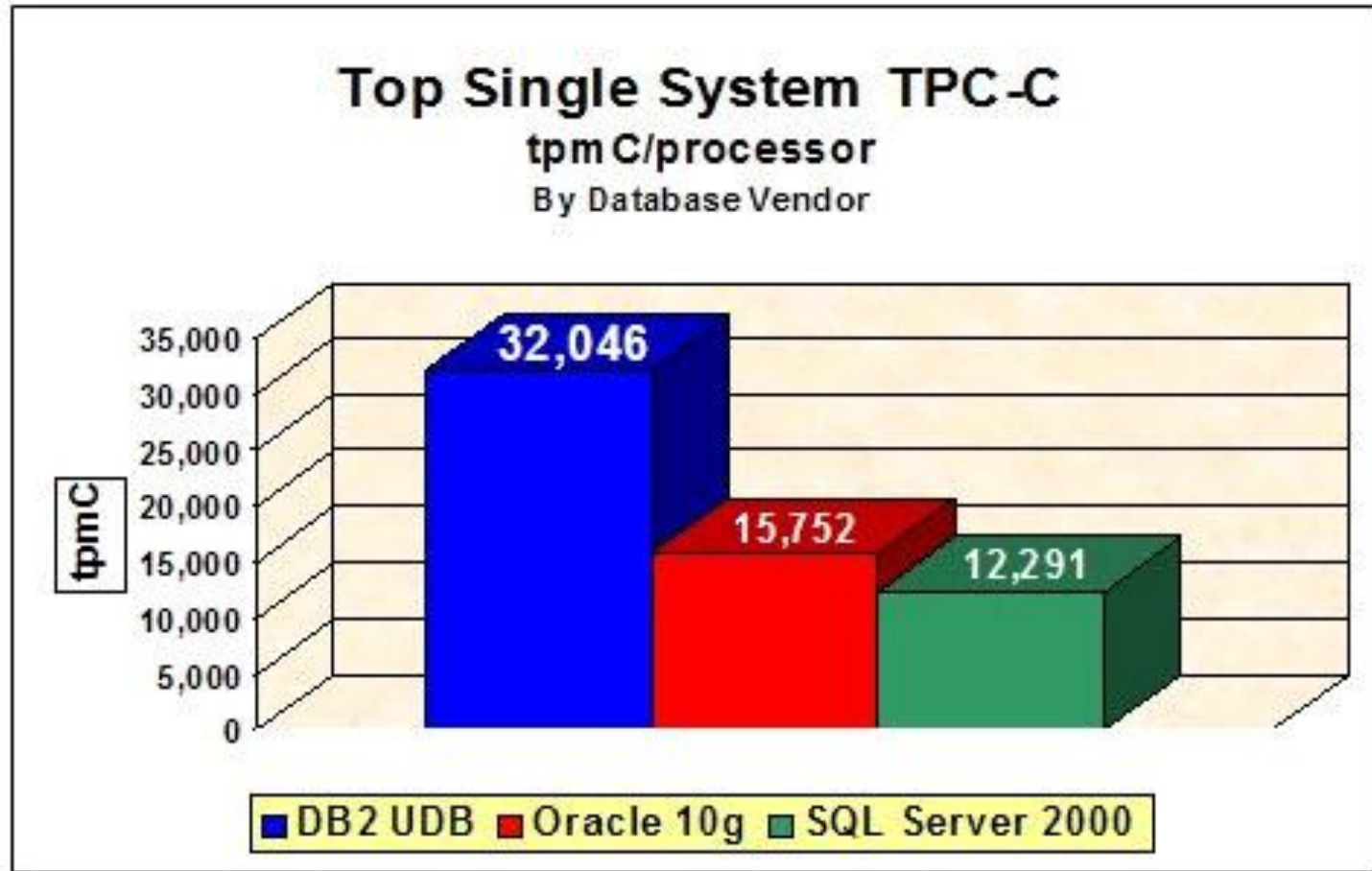


## Page Cleaning Improvements DB2 V8.1.4

- New page cleaning algorithm
- New DB2 Registry variable
  - DB2\_USE\_ALTERNATE\_PAGE\_CLEANING
- Use DB2SET command
  - DB2SET DB2\_USE\_ALTERNATE\_PAGE\_CLEANING=ON
- New algorithm reduces potential for flooding RUN queue



## DB2 Stinger TPC Record



Source: [www.tpc.org](http://www.tpc.org)

Enabling Your On Demand DB2 World

## Information Integration

- “The integration of disparate data into DB2 or accessible from and/or managed by DB2”
  - Data is in all types of forms in many different “places”
  - Can access from DB2 via Federation
  - Optimizer improvements for accessing Federated data
  - Web services
- Data is needed “On Demand” from anywhere, from anyplace, at any time

## Health Indicators

- Database Manager or Database health indicators for an instance can be obtained by issuing the following command:
  - “db2 get alert cfg for DBM”
  - “db2 get alert cfg for database”



## Database Health Indicators

Indicator Name	= db.sort_shrmem_util
Type	= Threshold-based
Warning	= 70
Alarm	= 85
Sensitivity	= 0
Formula	= ((db.sort_shrheap_allocated/sheapthres_shr)*100);
Actions	= Disabled
Threshold or State checking	= Enabled



## Memory Tracker

- DB2 Memory Tracker is a new tool available in DB2 V8.1
- Similar usage to UNIX commands such as VMSTAT, IOSTAT
- Run from DB2 command line
- Command and options follow:  
`db2mtrk -i | -d | -p [-m | -w] [-v] [-r interval [count]] [-h]`
- Handy for tracking memory usage over time interval
- Can be used for point based monitoring and could be scripted and data stored in DB2 tables to show DB2 memory/heap activity over time



# ITUC® 2004 - North America Memory Tracker Output

gtsaix01

Memory for database: SAMPLE

Backup/Restore/Util Heap has max size of 20660224 bytes  
Package Cache has max size of 536870912 bytes  
Catalog Cache Heap has max size of 536870912 bytes  
Buffer Pool Heap has max size of 536870912 bytes  
Buffer Pool Heap has max size of 536870912 bytes  
Buffer Pool Heap has max size of 536870912 bytes  
Buffer Pool Heap has max size of 536870912 bytes  
Buffer Pool Heap has max size of 536870912 bytes  
Lock Manager Heap has max size of 638976 bytes  
Database Heap has max size of 6127616 bytes  
Other Memory has max size of 12517376 bytes  
Total: 3798040576 bytes

Memory for agent 24254

Application Heap has max size of 1277952 bytes  
Application Control Heap has max size of 704512 bytes  
Total: 1982464 bytes

Tracking Memory on: 2004/01/12 at 02:16:36

Memory for instance

Database Monitor Heap has max size of 540672 bytes  
Other Memory has max size of 13467648 bytes  
Total: 14008320 bytes

Memory for database: SAMPLE

Backup/Restore/Util Heap has max size of 20660224 bytes  
Package Cache has max size of 536870912 bytes  
Catalog Cache Heap has max size of 536870912 bytes  
Buffer Pool Heap has max size of 536870912 bytes  
Buffer Pool Heap has max size of 536870912 bytes  
Buffer Pool Heap has max size of 536870912 bytes  
Buffer Pool Heap has max size of 536870912 bytes  
Buffer Pool Heap has max size of 536870912 bytes  
Lock Manager Heap has max size of 638976 bytes  
Database Heap has max size of 6127616 bytes  
Other Memory has max size of 12517376 bytes  
Total: 3798040576 bytes

Memory for agent 24254

Application Heap has max size of 1277952 bytes  
Application Control Heap has max size of 704512 bytes  
Total: 1982464 bytes

Start | 2 Windo... | DB2 Basics... | Adobe Re... | Microsoft ... | GTSAX01... | 2 Excee... | WS\_FTP Pro | db2mtrk - ... | Telnet 192... | 2:22 AM

Enabling Your On Demand DB2 World



Independent • Not for Profit • User Run

## Snapshot Monitoring via New SQL Functions in V8.1

- New in v8 is the ability to issue snapshot commands via SQL functions. These 15 new functions make it easy for you to write programs that issue snapshot commands and process the snapshot data. They can also be issued via the CLP. Previously this was only available via the administrative API.
- This is a real breakthrough in monitoring capability in DB2 and along with write-to-table event monitors provides a historical repository for use in reviewing performance problems, trend and historical analysis

## Performance Repository

SQL SNAPSHOT FUNCTIONS  
+ WRITE-TO-TABLE EVENT MONITORS  
= DB2 PERFORMANCE REPOSITORY

- New powerful SQL Snapshot functions in V8.1 and write-to-table event monitors give us true historical capability.



## SQL SNAPSHOT FUNCTIONS

- Sample insert statement for using an SQL snapshot function and inserting the output into a DB2 table:

```
INSERT INTO BP_SNAP
SELECT * FROM TABLE( SNAPSHOT_BP( 'sample', -1 ))
as SNAPSHOT_BP;
```



# IDUG® 2004 – North America

Control Center

Control Center Selected Edit View Tools Help

Control Center

192 - NDE2B71F (DB2inst1) - SAMPLE1 (SAMPLE) - User Defined Functions

Name	Schema	Specific name	Result	Input parameters	Comment
OID2PATH	SYSFUN	OID2PATH	VARCH... (BIGINT)		
PATH2OID	SYSFUN	PATH2OID	BIGINT (VARCHAR())		
USERS_GROUPS	SYSFUN	USERS_GROU...	VARCH... (INTEGER,VARCH...		
GROUPS	SYSFUN	GROUPS	VARCH... ()		
USERS	SYSFUN	USERS	VARCH... ()		
GROUPS_FOR_USER	SYSFUN	GROUPS_FOR...	VARCH... (VARCHAR())		
SNAPSHOT_DBM	SYSPROC	SNAPSHOT_D...	TIMEST... (INTEGER)		
SNAPSHOT_FCM	SYSPROC	SNAPSHOT_FCM	TIMEST... (INTEGER)		
SNAPSHOT_FCMNODE	SYSPROC	SNAPSHOT_FC...	TIMEST... (INTEGER)		
SNAPSHOT_SWITCHES	SYSPROC	SNAPSHOT_S...	TIMEST... (INTEGER)		
SNAPSHOT_APPL_INFO	SYSPROC	SNAPSHOT_AP...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_APPL	SYSPROC	SNAPSHOT_AP...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_STATEMENT	SYSPROC	SNAPSHOT_ST...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_LOCKWAIT	SYSPROC	SNAPSHOT_LO...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_AGENT	SYSPROC	SNAPSHOT_AG...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_SUBSECT	SYSPROC	SNAPSHOT_S...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_DATABASE	SYSPROC	SNAPSHOT_DA...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_BP	SYSPROC	SNAPSHOT_BP	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_TBS	SYSPROC	SNAPSHOT_TBS	(VARCHAR(),INTE...		
SNAPSHOT_LOCK	SYSPROC	SNAPSHOT_LO...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_TABLE	SYSPROC	SNAPSHOT_TA...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_TBS_CFG	SYSPROC	SNAPSHOT_TB...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_QUIESCERS	SYSPROC	SNAPSHOT_Q...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_CONTAINER	SYSPROC	SNAPSHOT_C...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_RANGES	SYSPROC	SNAPSHOT_RA...	TIMEST... (VARCHAR(),INTE...		
SNAPSHOT_TBREORG	SYSPROC	SNAPSHOT_TB...	TIMEST... (VARCHAR(),INTE...		
DB_PARTITIONS	SYSPROC	DB_PARTITIONS	SMALLI... ()		

160 of 160 items displayed

Default View\*

View

## WRITE-TO-TABLE Event Monitors

- Can use to support development efforts and once you have developed “canned” SQL to query the tables, it is very easy to review and asses performance and adequacy of SQL and resources
- Third party vendor query tools also can be used to speed analysis of data





## Write-to-table Event Monitors

Event type	Target table names	Available information
DEADLOCKS	CONNHEADER DEADLOCK DLCONN  CONTROL	Connection metadata Deadlock data Applications and locks involved in deadlock Event monitor metadata
DEADLOCKS WITH DETAILS	CONNHEADERE DEADLOCK DLCONN  DLLOCK CONTROL	Connection metadata Deadlock data Applications involved in deadlock Locks involved in deadlock Event monitor metadata
STATEMENTS	CONNHEADER STMT SUBSECTION  CONTROL	Connection metadata Statement data Statement data specific to subsection Event monitor metadata
TRANSACTIONS	CONNHEADER XACT CONTROL	Connection metadata Transaction data Event monitor metadata
CONNECTIONS	CONNHEADER CONN CONTROL	Connection metadata Connection data Event monitor metadata
DATABASE	DB CONTROL	Database manager data Event monitor metadata
BUFFERPOOLS	BUFFERPOOL CONTROL	Buffer pool data Event monitor metadata
TABLESPACES	TABLESPACE CONTROL	Tablespace data Event monitor metadata
TABLES	TABLE CONTROL	Table data Event monitor metadata

## WRITE-TO-TABLE Event Monitors

- General Consideration for Write-to-Table Event Monitors
- When the CREATE EVENT MONITOR statement is issued, all event monitor target tables are created. If the creation of a table fails for any reason, an error is passed back to the application program and the CREATE EVENT MONITOR statement fails.
- During CREATE EVENT MONITOR processing, if a table already exists, but is **not defined** for use by another event monitor, no table is created, and processing continues. A warning is passed back to the application program.

# IDUG® 2004 – North America

**Control Center**

Control Center Selected Edit View Tools Help

192 - NDE2B71F (DB2inst1) - SAMPLE1 (SAMPLE) - Event Monitors

Name	Status	Event Types	Auto restart	File I/O	Buffer size
PGUNN	Stopped	Connections ...	No	Blocked	4

**Create Event Monitor**

192 - NDE2B71F (DB2inst1) - SAMPLE1 (SAMPLE)

Name:

**Event Types**

- Database
- Bufferpools
- Tables
- Tablespaces
- Deadlocks
  - With details
- Connections
- Transactions
- Statements

**Activate**

All event monitors are terminated when the database is stopped

- Start now
- Restart automatically

Output options

OK Cancel Show SQL Help

1 of 1 items displayed

Default View View

# IDUG® 2004 – North America

The screenshot displays the IBM Control Center interface. On the left, a tree view shows the hierarchy: Control Center > Cataloged Systems > CPQ47403191915 > 192.168.1.101 > Instances > NDE2B71F (DB2inst1) > Databases > SAMPLE1 (SAMPLE) > Event Monitors. The main pane shows a table of Event Monitors for PGUNN.

Name	Status	Event Types	Auto restart	File I/O	Buffer size
PGUNN	Stopped	Connections ...	No	Blocked	4

An "Event Analyzer for PGUNN" dialog box is open, showing a table of monitored periods:

ACC_CURS_BLK	AG...	APPL_ID	APPL_...	APPL_...	APPL_SECTIO...	APPL_SECT...
0	60	*LOCAL_db2inst1_05B417170004			0	

The dialog box also includes a "Navigate to" dropdown menu set to "Data Elements" and "Close" and "Help" buttons.

# IDUG® 2004 – North America

Data Element	Value
ACC_CURS_BLK	0
AGENT_ID	130
APPL_ID	COA80168.L405...
APPL_PRIORITY	0
APPL_PRIORITY_TYPE	65536
APPL_SECTION_INSERTS	0
APPL_SECTION_LOOKUPS	2
APPL_STATUS	11
AUTHORITY_LVL	18292030
BINDS_PRECOMPILES	0
CAT_CACHE_HEAP_FULL	0
CAT_CACHE_INSERTS	0
CAT_CACHE_LOOKUPS	0
CAT_CACHE_OVERFLOWS	0
CAT_CACHE_SIZE_TOP	0
COMMIT_SQL_STMTS	2
COORD_NODE	130
DDL_SQL_STMTS	0
DEADLOCKS	0
DIRECT_READ_REQS	0
DIRECT_READ_TIME	0
DIRECT_READS	0
DIRECT_WRITE_REQS	0
DIRECT_WRITE_TIME	0
DIRECT_WRITES	0
DISCONN_TIME	Jan 17, 2004 2:...
DYNAMIC_SQL_STMTS	2
ELAPSED_EXEC_TIME	123
EVMON_FLUSHES	0
FAILED_SQL_STMTS	0
HASH_JOIN_OVERFLOWS	0
HASH_JOIN_SMALL_OVERFLOWS	0
INT_AUTO_REBINDS	0
INT_COMMITS	0
INT_DEADLOCK_ROLLBACKS	0
INT_ROLLBACKS	0
INT_ROWS_DELETED	0
INT_ROWS_INSERTED	0
INT_ROWS_UPDATED	0
LOCK_ESCALS	0
LOCK_TIMEOUTS	0
LOCK_WAIT_TIME	0
LOCK_WAITS	0
PARTIAL_RECORD	0

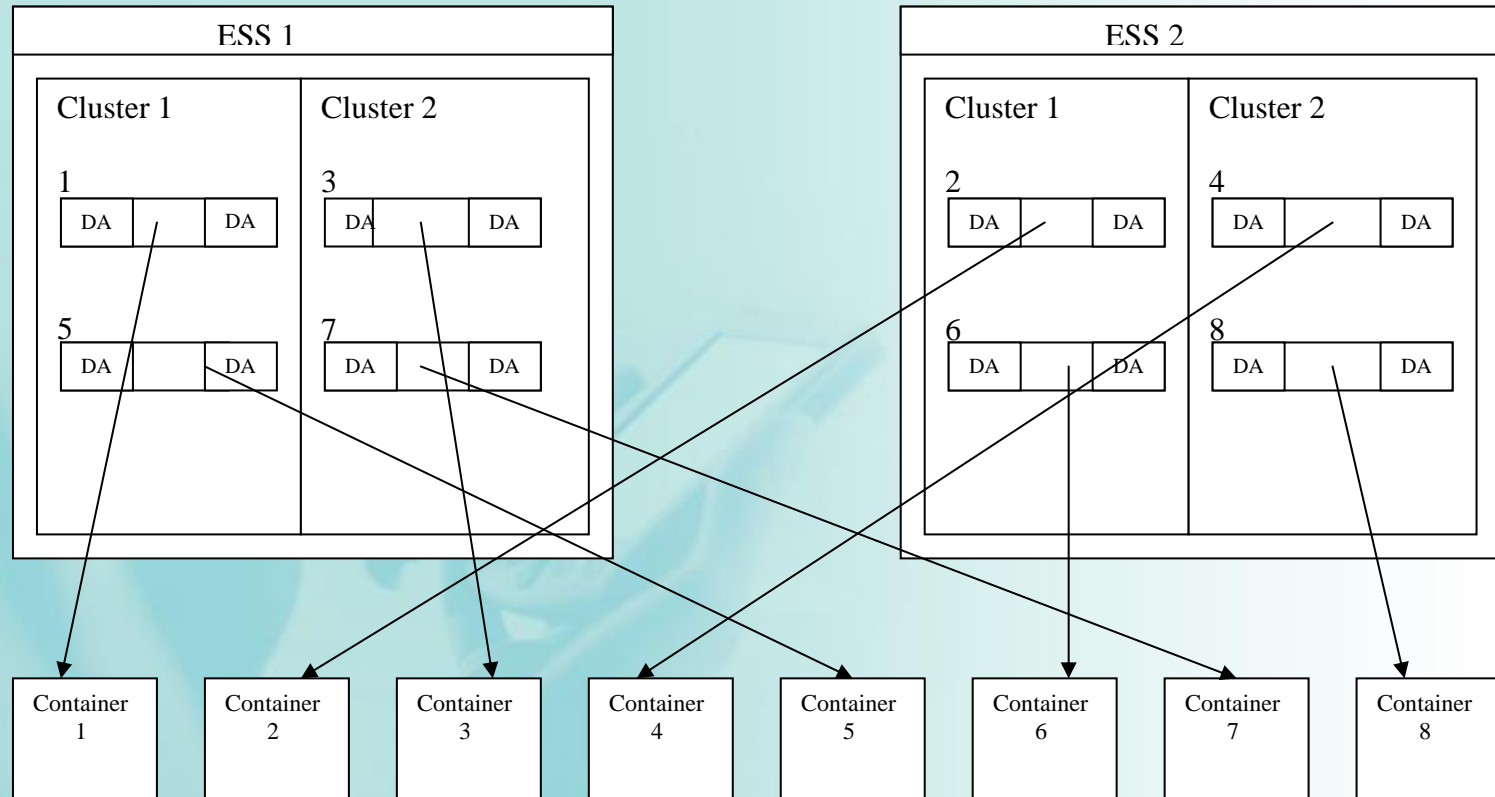
## Balanced I/O

- Key for good performance
- Balance I/O across
  - Storage devices (cabinets)
  - Clusters within a cabinet
  - Span disk adapters
  - Engage as many arrays as possible



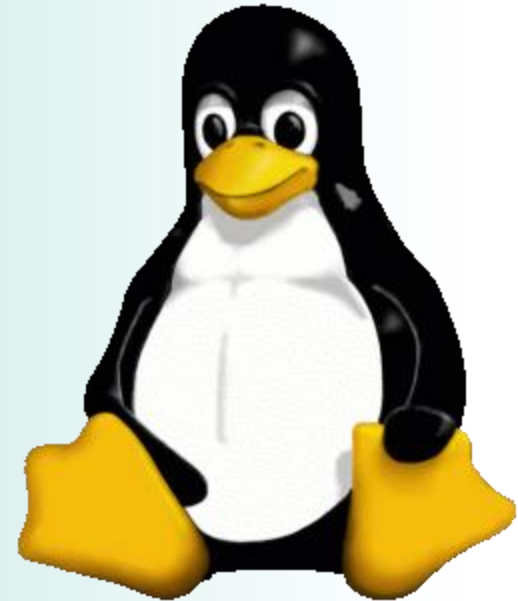


## SHARK



## DB2 on LINUX RULES!

- DB2 Linux Information on the Web:
- <http://www-306.ibm.com/software/data/db2/linux/>
- <http://www-306.ibm.com/software/data/db2/linux/validate/>
- DB2 UDB V8.1 (and above) setup information



IBM Widens Embrace of 64-Bit Linux - Microsoft Internet Explorer provided by Compaq

File Edit View Favorites Tools Help eSend

Address <http://www.eweek.com/article2/0,1759,1540390,00.asp?kc=EWNWS030104DTX1K0000599>

My Search Google AltaVista More WebFerret Norton AntiVirus

**eWEEK** ENTERPRISE NEWS & REVIEWS ZIFF DAVIS MEDIA

NEWS · REVIEWS · OPINIONS · CASE STUDIES · RESEARCH · TOOLS · DISCUSSIONS

SEARCH eWEEK

SUBSCRIBE TO eWEEK My Account | Sign In Not a member? Join now

Home > Linux & Open Source > News > **IBM Widens Embrace of 64-Bit Linux**

**Linux & Open Source**

**IBM Widens Embrace of 64-Bit Linux**

By Jeffrey Burt  
March 1, 2004

IBM has been a champion of Linux for several years—extolling the virtues of the open-source operating system running on its xSeries systems, which use industry-standard 32-bit Intel Corp. processors.

**ADVERTISEMENT**

**See how the right middleware can improve your IT infrastructure.**

- e-business On Demand Operating Environment
  - Overview of an on demand operating environment
  - Integration
    - Automation
      - Automation blueprint
      - Autonomic computing and automation
      - Automation requirements
    - Virtualization
      - On demand technologies
        - Technology summary
      - On demand operating environment offerings
        - Integration offerings
        - Automation offerings
        - Virtualization offerings

But the company lately has been ramping up a major initiative to push Linux on its iSeries and pSeries servers, which use IBM's family of Power processors, as a way to expand its reach, particularly as an entry point for 64-bit computing. A focal point of the strategy, dubbed Linux-on-Power, this year will be IBM's PowerPC 970 processors.

The Linux-on-Power group, which was established in January to promote the initiative, is responsible for its own profits and losses and has engineers dedicated to the task of improving the open-source

**TABLE OF CONTENTS**

- Introduction
- IBM's Not Alone
- Vocal Linux Supporter

**TOPIC CENTERS**

JUMP TO TOPIC CENTER:

-Select-

**LINUX & OPEN-SOURCE VIEW**

**STEVEN J. VAUGHAN-NICHOLS**

**Open Software, Secure Software**

Even experts can disagree. Linux & Open-Source Center Editor Steven Vaughan-Nichols thinks the open-source approach does lead to more secure software.

**BREAKING NEWS**

3:46 PM  
Texas State's Accuses

Done

attempt on Coproot, at the same time withdrawing its board state and extending

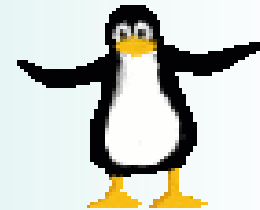
<http://eletters.eweek.com/zd1/cts?d=79-511-2-3-260015-60183-1>

Start Inbox - Outlook Express ... IBM Widens Embrace ...

Internet 4:06 PM

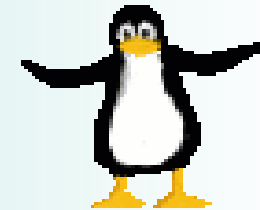
## The DB2 for Linux Tsunami

- The Linux momentum has reached critical mass and the Linux Tsunami has hit land!
- DB2 Rocks on Linux!
- IBM early embracer and supporter
- DB2 for Linux continuous improvement
- If you are not riding the Linux Tsunami you are .....



## V8.1 Linux Enhancements

- Support for the following processors:
  - AMD64
  - Intel 32-bit(x86)
  - Intel 64-bit
  - iSeries and pSeries
- Larger bufferpools
- Auto-configuration of kernel parameters
- Alternate Fixpacks
- Be on the lookout for new enhancements!



## Storage Provisioning

- Storage provisioning is concerned with easing the storage administration burden for DBAs and System Administrators
- Future plan
- Based on similar concepts used on mainframe for managing storage
  - System Managed Storage (SMS)
    - Don't confuse this with System Managed Space (SMS)!





## 64-bit Addressability

- Fully supported by DB2 V8.1
- Enables use of larger bufferpools, sortheaps, package cache, catalog cache, locklist and other memory areas
- Enables use of a large physical memory (2 power 64)
- DB2 supports IA-64 and AMD processors
- Enables substantial amount of scalability
- Especially beneficial to relational databases that use large amounts of memory

# IDUG® 2004 – North America



**Enabling Your  
On Demand DB2 World**

## **Tuning DB2 in the New Galaxy**

Questions?



**THANK YOU!**

[pgunning@gunningts.com](mailto:pgunning@gunningts.com)  
Session ID: E3