

High Availability & Disaster Recovery: Oracle 10g and 9i

ADVANCED FEATURES

Versus The Competition

Version 1.02: Visit
www.maop.org (Fall 2005)
for upcoming revisions.

Jeffrey Bertman

Chief Engineer
DataBase Intelligence Group (DBIG)
(jefflit@dbigusa.com, 703-405-5432)
WDC / VA / MD

Ran Pan, PhD

Lead Technologist, PCAOB
(panr@pcaobus.org, 703-547-6420)
WDC / VA / MD

- IT Consulting • Strategic Planning • On-Call Support •
- High Volume (VLDB) and High Transaction Specialists •

AOTC Conference 12/8/2005

How Important are High Availability and Disaster Prevention/Recovery?

Industry	Operation	Approx Avg Downtime Cost/Hour
Financial	Brokerage	\$6,450,000
Financial	Credit Card	\$2,600,000
Media	TV Pay-per-View	\$150,000
Retail	TV Home Shopping	\$113,000
Telecom	B2B CLEC Sales & Trouble Ticket Admin (Call processing/provisioning figures generally <i>higher</i>)	\$100,000 SLA + Lost Opportunity
Retail	Catalog Sales	\$90,000
Transportation	Airline Reservations	\$89,000
Media	Teleticket Sales	69,000
Transportation	Package Shipping	\$28,000
Finance	ATM Fees	\$14,000

Figures for all above *EXCEPT* Telecom complements of HP/Compaq Users Group presentation by John Zimmerman, Digitasks Consultants, Inc.
(<http://www.decus.gr.jp/decus99/sessioncd/NOTES/UN107.PDF>)

Objectives

- **Establish**
BUSINESS
CONTINUITY

Framework:

- High Availability Components — HA (up to 24x7)
- Disaster Recovery Components — DR
- *MEASURABLE GOALS*

- Discuss Pros & Cons of Various Major Technologies — often *COMBINING* two or more:

Backups
(not enough)

RAID
(not enough)

Active-Passive
Clustering

Remote
/ Geographic
Disk Mirroring
(Snap & Block-stream Technology)

- Unidirectional Replication

- Bidirectional and
Multi-Master Replication

- Unidirectional Replication

Oracle + Third Party Tools (Examples)

→ Fault Tolerance 101 and 201

Backups (not enough):

- RMAN
- Conventional Hot Backup
- VERITAS NetBackup
- IBM Tivoli
- ETC

RAID (not enough):

- 1+0
- 5

Active-Passive Local Clustering:

- VERITAS Cluster Services (VCS)
- IBM HA/CMP
- Sun Clusters
- HP MC/ServiceGuard
- HP True64 Clusters
- Linux and Windows Clusters

Advanced: Third Party Tools (Examples)

→ Remote Disk Mirroring

RDM Tools (aka Geo-Mirroring)

Technology

- | | |
|--|-----------------------|
| - *EMC SRDF/Timefinder | Block-stream BCV |
| - *Hitachi TrueCopy/ShadowImage | Snap BCV+SPIT |
| - *Network Appliance SnapMirror | Snap SPIT |
| - VERITAS Volume Replicator | Snap SPIT |
| - *Hitachi Copy-On-Write Snapshot
(formerly known as QuickShadow) | Block-stream BCV+APIT |
| - Mendocino RealTime/Appliance | Block-stream APIT |

Legend: * = Dependent on Proprietary Disk Storage Hardware

Advanced: Oracle + Third Party Tools (Examples)

→ Database Transaction Replication

Unidirectional-Only:

- Oracle Data Guard

Uni+Bidirectional/Multi-Master:

- Oracle Streams
- Oracle Advanced Replication
- Quest SharePlex
- DataMirror iReflect
- DataMirror Transformation Server
(Heterogeneous Replication)

Advanced: Oracle + Third Party Tools (Examples)

→ Active-Active Process Clustering

Hybrid Environments:

- **Oracle Real Application Clusters (RAC)**
and associated Clusterware:
 - > **Oracle Cluster Ready Services (CRS)**
 - > **VERITAS DBE/AC**
(Bundle: VCS, DataBase Edition, VxFS)
 - > **IBM HA/CMP**
 - > **Sun Clusters**
 - > **HP MC/ServiceGuard**
 - > **HP True64 Clusters**
 - > **Linux and Windows Clusters**

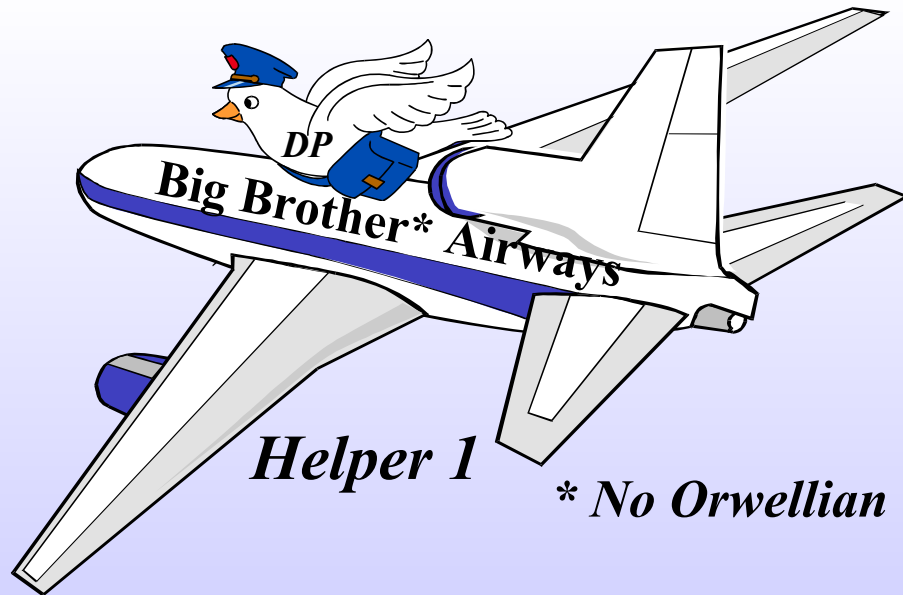
TODAY's Scope (only 2 hours)

- Technical Framework
- Specific Pros and Cons
- Your most challenging problems or nightmares are WELCOME, but SAVE any **Specific Configuration Questions** for:
 - **Q&A at End of Session**
 - If no time today, **ASAP Follow-up**

From Here to There

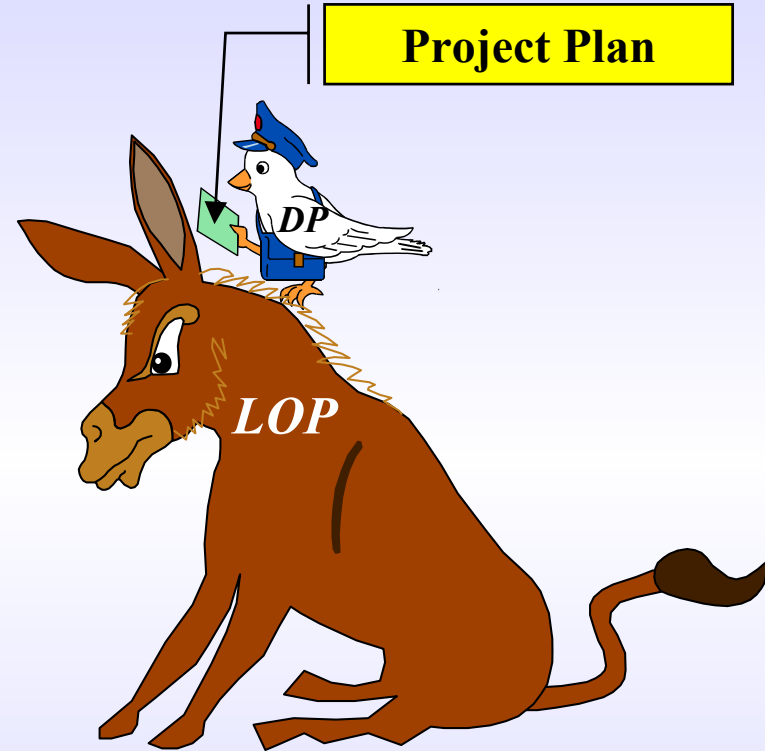


Helper 2



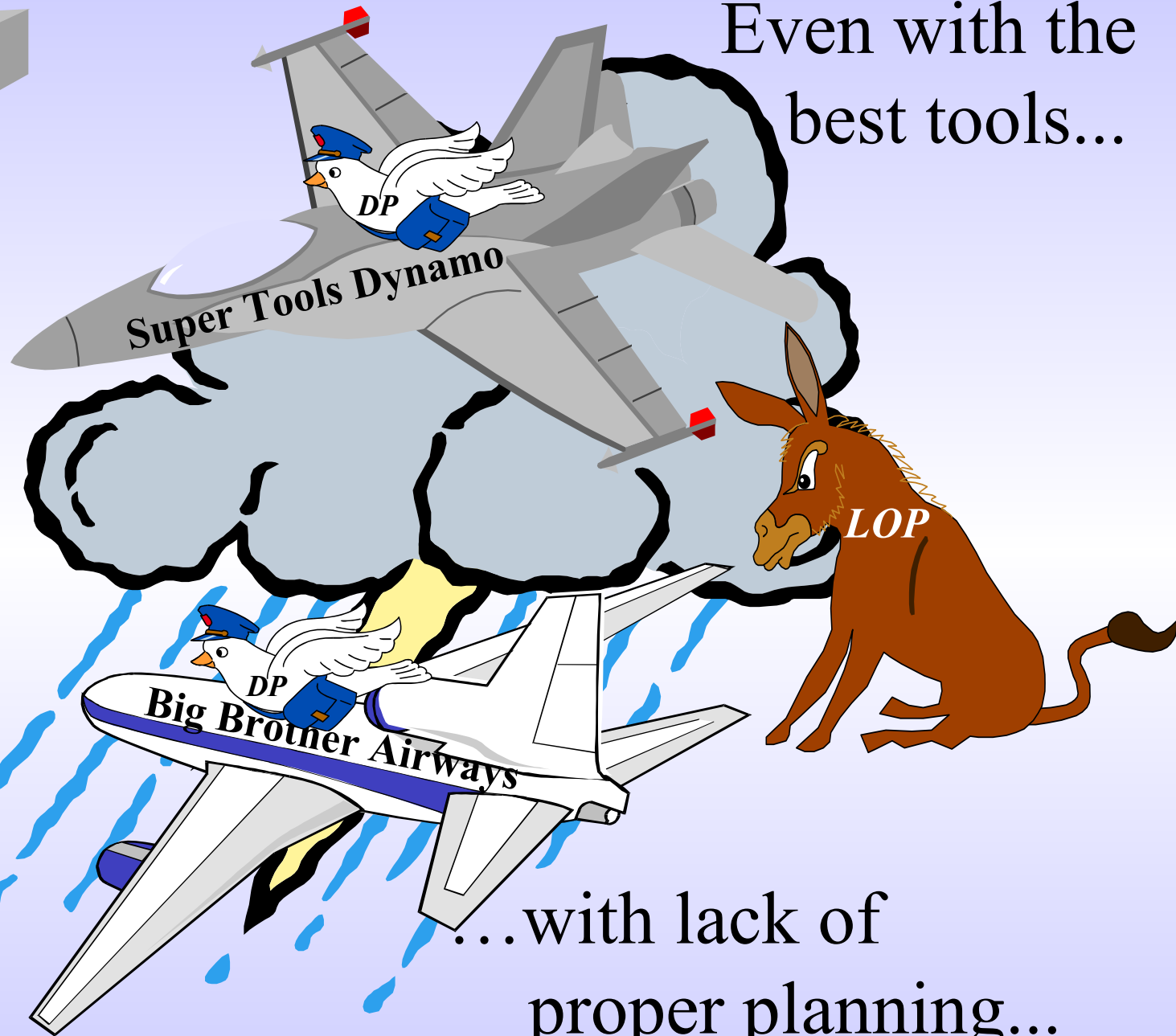
Helper 1

** No Orwellian Connotation*



*Challenge
Numero
Uno*

Even with the
best tools...



...with lack of
proper planning...

...can lead to...





Opening the door to success requires suitable technology PLUS planning and follow-through...



...to tame LOP the mule into LOP2 !

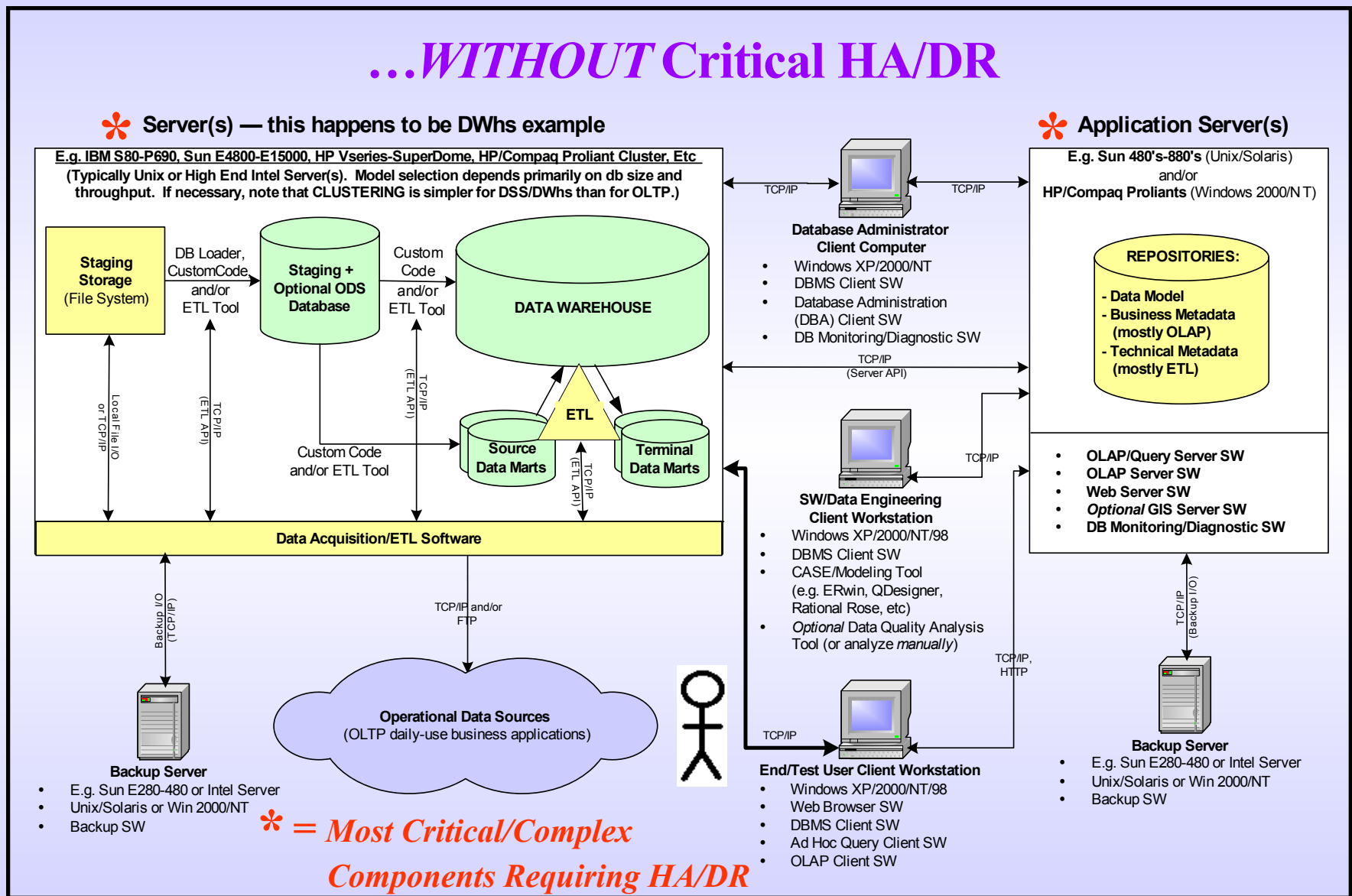
WARNING:

This Presentation is still a
Rough DRAFT.

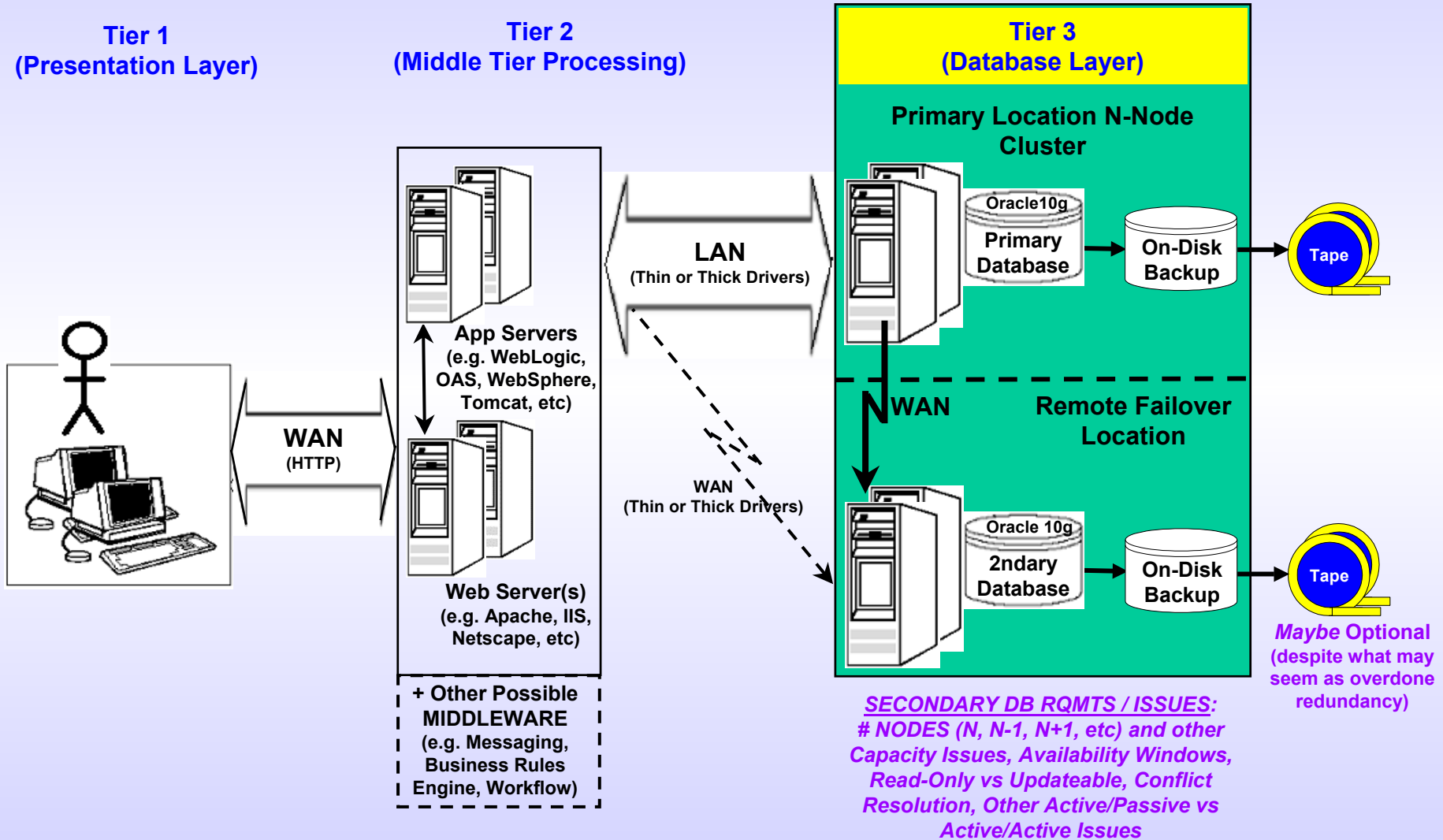
Contact Author for latest revision.
→ See Contact Info at
Beginning and End.

Typical Web Environment (This e.g. happens to be Data Warehouse)...

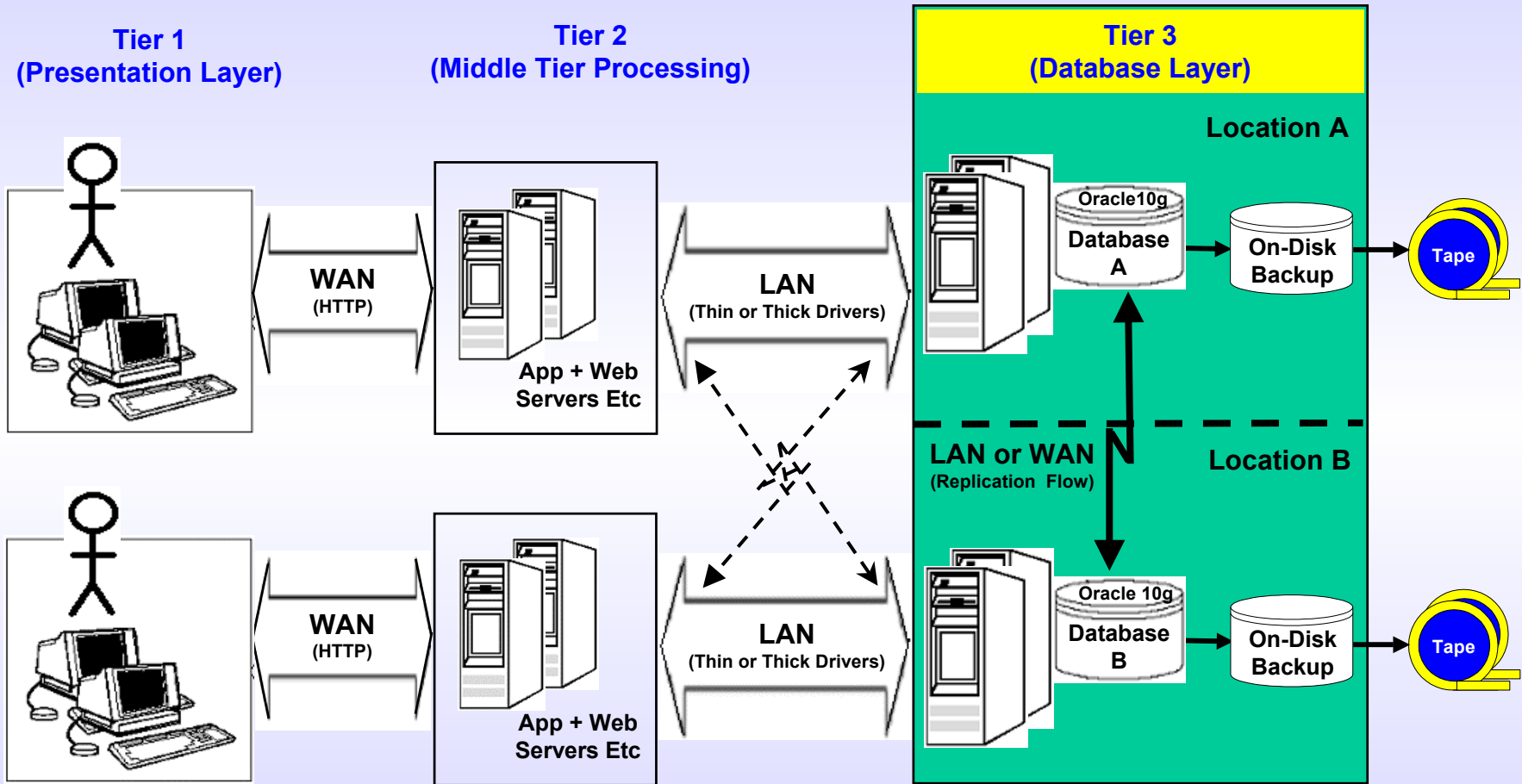
...WITHOUT Critical HA/DR



What are We Building? HA/DR Example for E-Business



What are We Building? MMR Example for E-Business

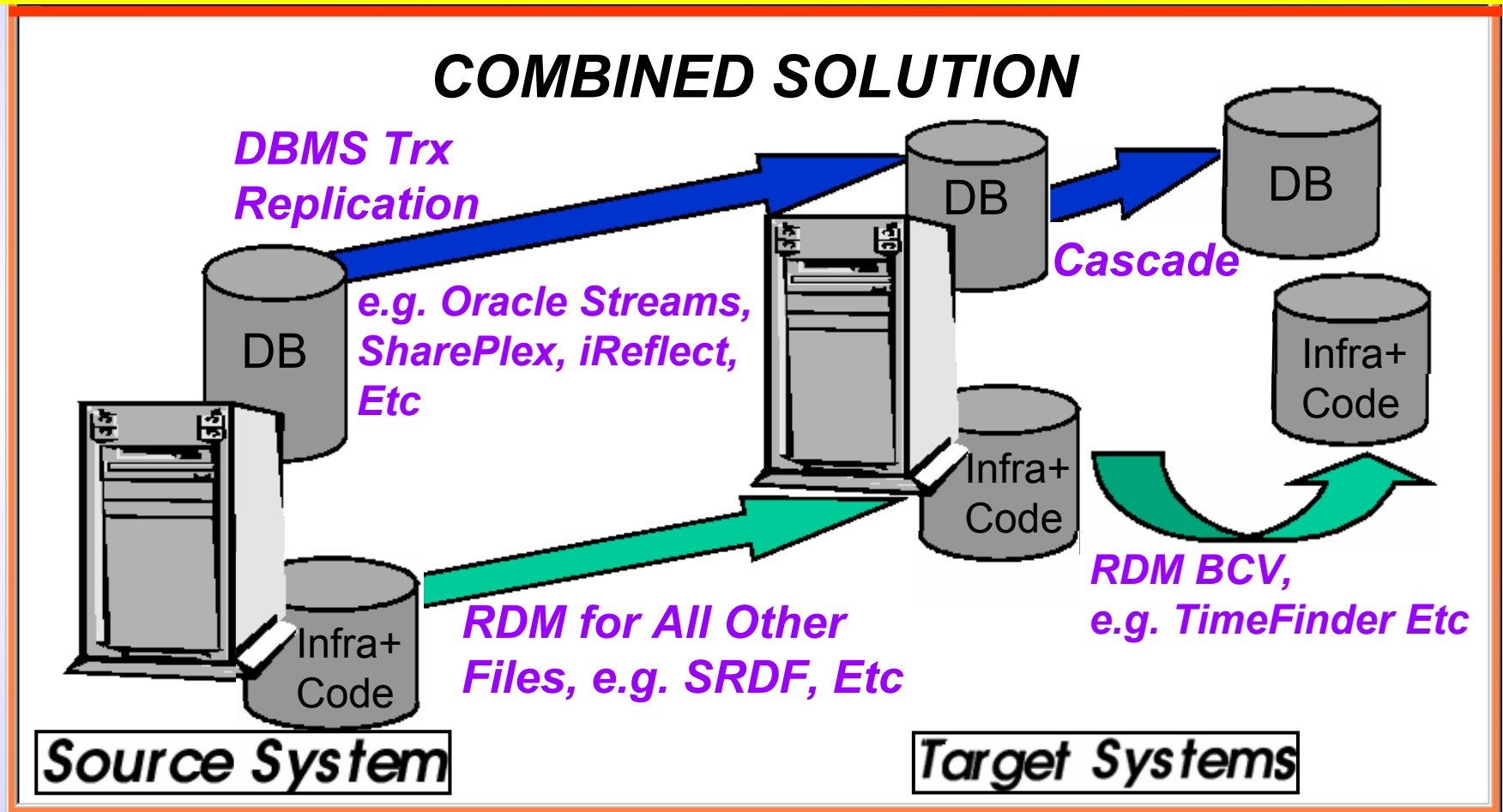


MAJOR MMR ISSUES:

*ONGOING Conflict Management
(Conflict Avoidance/Detection/Resolution)*

PERIODIC Re-Synchronization

DR Example of Using MULTIPLE Technologies to Complement One Another



Basic HA/DR

Important but generally Not Enough for Mission/Business Critical Environments

- UPS and RAID — Don't Take them for Granted ☺
Examples: RAID Level 5, Level 1+0
- Hot Backup and Exports
 - PROS:
 - Can Multiplex your Archive Logs to Remote/Failover Site
 - Fine Granularity
(Can Recover Individual Data Files or, in the case of Export, Tables)
 - CONS:
 - Too Much Downtime — Too Long to Recover
 - Possibly Significant Data Loss
(Online Redo Logs)
 - Exports generally just at Night / Low Peak
(Consistent View takes too long and too much overhead during day)

MISSION CRITICAL HA/DR

- **Document Your Requirements & Assumptions:**

EXAMPLES (examples below focus more on DR than HA)...

- Failover Database is FULL or PARTIAL Copy of Primary Database
- Failover and Failback may/may not involve manual intervention subject to specific METRICS (next slide)
- Transparent IP/Messaging Switchover Upon Failover
- End User Reconnection Requirements, e.g. Transparent App Failover (TAF) or Okay to Reconnect (using *normal* login procedure)
- Should Failover DB allow some Structure Differences to help Optimize Reporting Performance (e.g. different indexes, partitioning, tunable parms)
 - Capability for these differences to be Ignored or Reversed upon Failover AND Failback.
 - May need Manual Procedures to make this happen.

MISSION CRITICAL HA/DR

- **Document Your Required HA METRICS:**

EXAMPLES...

- **HA Mode Performance** (e.g. during Cluster Node or RAID Failure):

- FULL-Peak Degradation ≤ 30%
→ Can be drastically different for different envmts
- MID-Peak Degradation ≤ 20%

- **HA Uptime Percent** — aka “Number of 9’s”

Based on Mean Time Between Failure (MTBF)
and Mean Time to Recover (MTTR) :

$$\underline{\text{Target HA Percent} = \text{MTBF} / (\text{MTBF} + \text{MTTR}) * 100}$$

99.0%, TWO Nines:	Approx 88 hr (87.6)	per yr	or	7.33 hr	per month
99.90%, THREE Nines:	Approx 9 hr (8.76 hrs)	per yr	or	0.73 hr	per month
99.99%, FOUR Nines:	Approx 1 hr (53 min)	per yr	or	4 MIN	per month
99.999%, FIVE Nines:	Approx 5 minutes	per yr	or	0.4 MIN	per month

MISSION CRITICAL HA/DR

• DR METRICS EXAMPLES, Continued...

- **Geography: Failover Site must reside > X miles from Primary Site**
+ List Specific Location (just to be safe)
- **Utilization: Failover DB available for READ-ONLY / RPTG**
during specific *time windows (specific hours/schedule per day/week)*
 - **Option 1:** Unlimited Read-Only/Reporting Access
 - **Option 2:** MAYBE consider Write Access (2 or multi-way replication)
- **Return on Investment (ROI): Like ENHANCED Insurance**
 - COST of all materials, products, licensing & *maintenance* fees, labor (incl Technical Architecture/Planning thru Implementation)
 - BENEFIT of using DR site to offload or perhaps even load-balance work
 - BENEFIT of \$\$\$/hr saved by prevented downtime, which causes
 - Lost Opportunity, e.g. Lost Sales/Revenue
 - » Direct loss due to downtime
 - » Indirect loss due to loss of confidence
 - Lost Work (Data Loss *and* Business/Operational Work)
 - Backlog which may require Extra/After Hours Work
 - Many Environments: Fees (e.g. SLA, FCC Fines for Telecom Providers)

MISSION CRITICAL HA/DR

• DR METRICS EXAMPLES, Continued...

– Downtime

- Recovery Time Objective (RTO) <= 30 min
- FailOVER 1st Available Time (MTTR-Avail) <= 30 min
→ Use with CAUTION: Beware of Data Integrity Issues
- FailOVER Total Sync Time (MTTR-Synched) <= 60 min
- FailBACK Total Sync <= 60 min
- FailBACK Preparation Time (e.g. for Rebuild) <= @TBD
- Scheduled Downtime for OS/Database Maint <= 60 min

– Data Loss

- Recovery Point Objective (RPO) <= @TBD **MB** or **Trxs**
- Failover PLATFORM Latency <= 10 min
→ Get the info over there, but not necessarily applied to DB
- Failover DATABASE/Intentional Latency <= 90 min

MISSION CRITICAL HA/DR

- **DR METRICS EXAMPLES, Continued...**

(Hefty list, but still Not comprehensive — see Paper for more)

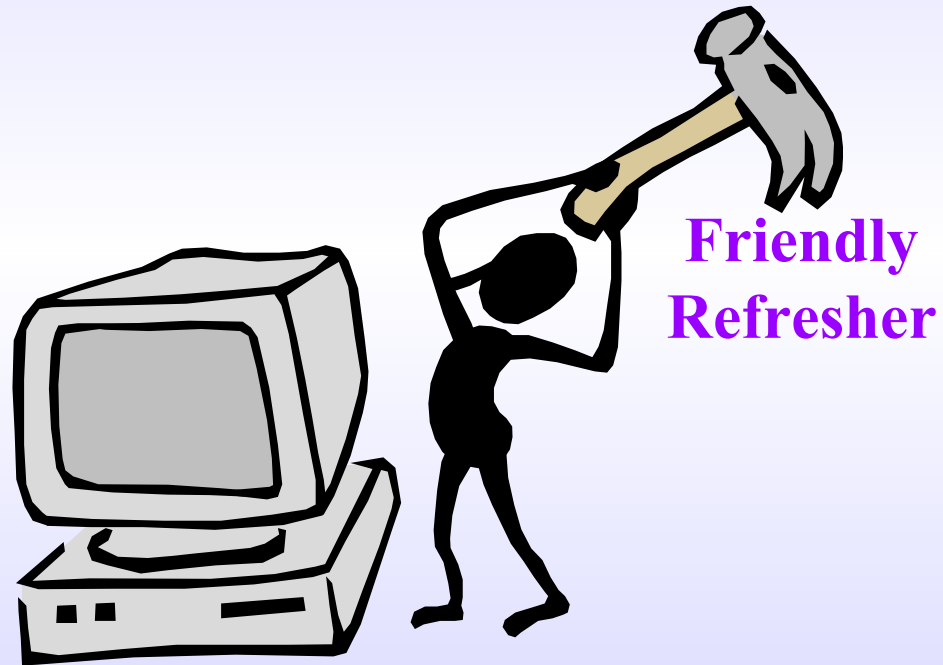
- **DR Mode Performance**

- FULL-Peak Degraded Level <= 30%
 - Can be drastically different for different envmts
- MID-Peak Degraded Level <= 20%

- **DR Mode Duration**

- Soonest #Hrs before FailBACK (*be practical*) = 24 hr
- Latest #Hrs before FailBACK = *X to Indefinite*

WAKE UP !!!



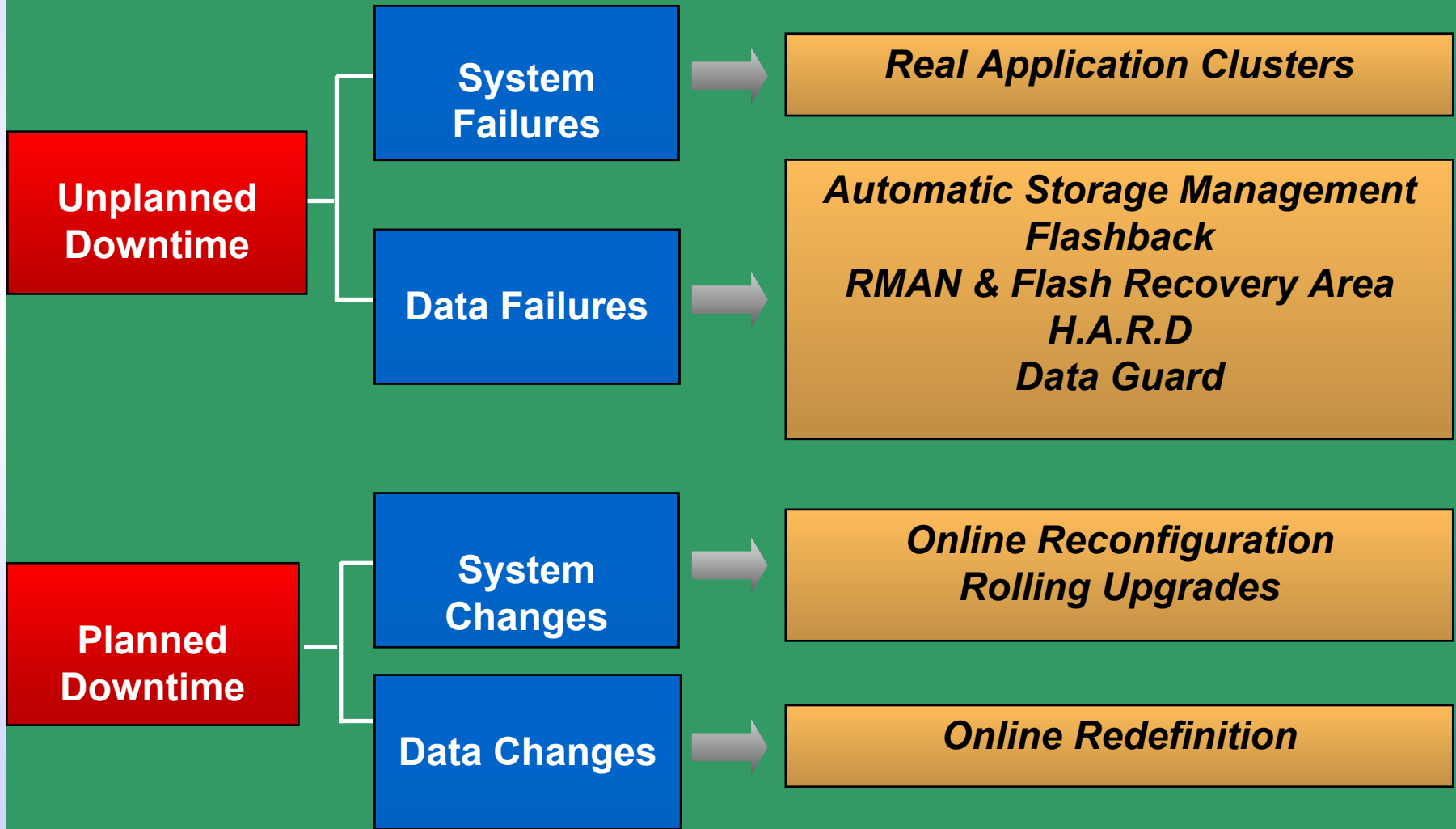
Critical DR: Remote Disk Mirroring (RDM) — PROS

- **SIMPLICITY of Replicating EVERYTHING with 1 Tool**
- **Not Prone to Human Error Corruption**
(E.g. from mismanagement of objects that can/cannot replicate)
- **Application Aware Technologies Emerging**
(Despite 1st Glance Separateness from Transactions)
- **Recovery Granularity:**
 - Block-stream Technology = Any Point in Time (APIT)
 - + Optionally Specific Point in Time (SPIT)
 - Snap Technology = Snap Intervals
 - + Optionally Specific Point in Time (SPIT)

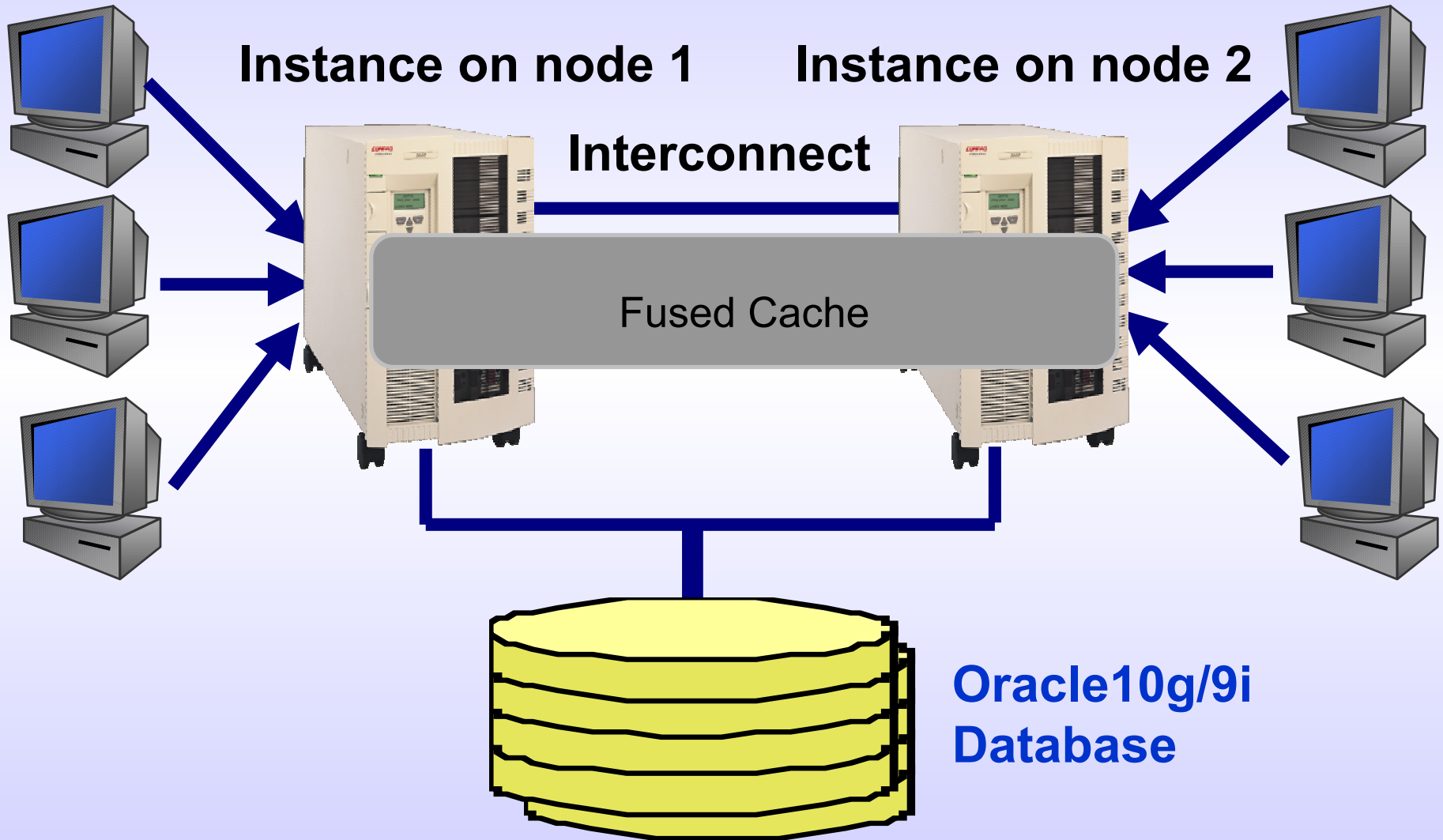
Critical DR: Remote Disk Mirroring (RDM) — CONS

- To Avoid Fractured Blocks, Some Type of Quiesce OR Checkpoint Processing must occur Frequently
- CANNOT Replicate BI-directionally
- More Prone to Block Level Media Corruption (versus Transactional Propagation)

Oracle's Integrated HA Solutions



Critical DR: Oracle 9i/10g RAC



Critical DR: Oracle RAC — PROS

- **Primary Benefits**
 - Increased SCALABILITY / Performance — *If Configured Well*
 - **DEFAULT** scalability for “most” apps using 9i/10g FULL Cache Fusion: 80% without customization according to Oracle — **HIGHER** for Read-Only DSS/Data Warehouse DBs.
 - Remember extra NETWORK/Interconnect traffic between server nodes for OLTP DBs, even with full Cache Fusion.
 - Increased Availability (focus on Server versus Storage/Media)
- **NODE Fault Tolerance** w/out Need for Complex Database Failover
 - WARNING: More HA vs DR Focused...
Cannot handle Site Destruction, Media Failure, or DB Corruption.
- **Increased App Resilience** — especially via TAF
 - For QUERIES. NOTE: Requires OCI-compliance, e.g. Oracle Forms, SQL*Plus, etc or for J2EE world need JDBC THICK (vs thin) Client.
- **Increased Scale-Out Uptime** via Shared Disk Architecture
 - Server Nodes can be added WITHOUT Redistributing Data across nodes — significant availability and complexity advantage vs Shared Nothing envmts, although we can spend hours **debating** related pros/cons, e.g. see Required Downtime in CONS section (which, BTW, actually occurs less often and takes less time vs data redistribution in shared nothing).

Critical DR: Oracle RAC — PROS

- **Increased Uptime during OS Upgrades/Patches**
 - Can shutdown one instance while others remain up, but schedule for off-peak since less nodes available for load balancing.
- **Increased Uptime during DB Upgrades/Patches**
 - “Rolling Upgrades” allow shutdown one instance while others remain up, but schedule for off-peak since less nodes available for load balancing.
 - **DEPENDS ON:**
 - **Oracle Corp spec for each upgrade/patch**
 - **Whether Data Dictionary is affected**
 - **Whether using shared Oracle Home via clustered filesystem or separate homes per node/instance**
- **Easily Complements other more DR Focused Solutions**

Critical DR: Oracle RAC — CONS

- **Extensive Interconnect Network Traffic in OLTP Envmnts**

(Unless data is segregated or partitioned *and* user/session access to data is vertically or horizontally partitioned across instances — not just via least load or round robin load balancing — “Instance Pinging” Acts like Block Level Locking)

- Oracle 8.1.6 through 9i IMPROVED this “instance pinging” problem gradually across Releases 1 through 3 (e.g. CACHE FUSION Phases 1 and 2).
- 9i/10g NO LONGER a Disk I/O problem, but still extensive NETWORK/Interconnect traffic.

- **Potential Complexity of RAW DEVICES**

(More difficult to maintain/backup vs standard/cooked filesystems, although eventually becomes routine — ALSO can lead to increase in wasted **DISK SPACE** due to raw device sizing options)

- Oracle 8i on most platforms and 9i on MORE platforms (e.g. finally on AIX 5.x) allows use of 3rd Party Clustered Filesystem to simulate regular filesystems, e.g. VERITAS.
- Oracle 9i/10g has ORACLE CLUSTERWARE on Win 2000 and Linux. For 10g, Cluster Ready Svcs (CRS) combines GSD, Oracle Cluster Services, and OCFS.

- **Potential Complexity of MULTIPLE Oracle HOMES**

- Clustered filesystem can RESOLVE this — but since Oracle OCFS / CRS currently available only on Win 2000 and Linux, consider third party, e.g. VERITAS DBE/AC with VxFS.

- **Extra Cost of Third Party (or HW Vendor) Clusterware**

(E.g. IBM HA/CMP, Sun Clusters, HP MC/ServiceGuard, HP/Compaq Proliant Clusters, VERITAS Database Edition/Advanced Cluster (DBE/AC). Remember clustered filesystem sw as well !

- Oracle 9i/10g has ORACLE CLUSTERWARE (See similar note above).

Critical DR: Oracle RAC — CONS

- **Distance Limitations** — LOCAL Failover Only (same site/room)
(Most clustering HW/SW requires or works better if servers co-reside at same site due to High Speed Interconnect between nodes)
→ **Geo-Clustering is starting to emerge, but since data replication makes sense between sites...**
- **High End Scalability Limitation** — Max# Supported Nodes
 - Clustering Software Limitation
(Number of nodes on even large platforms may be limited to 8 or 16 — best support on Alpha True64 which has become less popular for not-great reasons outside scope of this presentation)
 - Scalability Degradation for MANY Nodes Updating Same Rows/Blocks
→ **9i/10g FULL Cache Fusion improves this (minimizes disk I/O), but still involves extensive network interconnect I/O.**
- **Extra Maintenance Complexity** — e.g. During OS/DB Upgrades/Patches
→ **Oracle 9i Clustered Filesystem / 10g CRS improves this somewhat — less Oracle Home maintenance — but currently available only on Win 2000 and Linux. Also consider 3rd Party clustered filesystems, e.g. Veritas VxFS.**
- **Planned Downtime Required during DB Upgrades/Patches**
(For envmts > 2 nodes, actually less downtime withOUT shared Oracle Home via clustered clustered filesystem — unrelated to use of clustered filesystem for shared data files)

Critical DR: Automatic Storage Management (ASM provided in 10g Only) — PROS

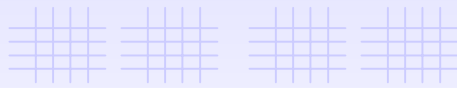
- **Improved Data Protection:**
 - Built-in Normal (2-way) and High Redundancy (3-way) as well as External Redundancy option (3rd party)
 - **WONDERFUL:**
FAILURE GROUPS support true disk separation (e.g. for redo log mirroring) *if* Oracle manages all disks
- **Centralized, Simplified, Consistent Management:**
 - Integrated filesystem and volume mgr for all ORACLE db files
 - One Vendor for Small to Mid Storage Requirements
 - Hides LUNs and Data File Nomenclature
 - POLICIES promote Consistency, e.g. for auto file names *and* mirror settings of data files vs redo/archive logs
 - IMPROVED RELIABILITY
(Centralized, Simplified, Consistent Mgt promote Maintainability)

Critical DR: Automatic Storage Management (ASM provided in 10g Only) — PROS

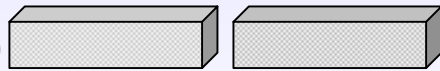
- **Increased Availability:**
 - Online changes to storage resources
(Add disks, remove disks, and migrate to new storage)
- **Automatic I/O Balancing and Resource Utilization:**
 - Distributes database files across all available disks
(at extent level)
- **It's FREE :-)**

Automatic Storage Management (ASM) versus Traditional Storage Management

Tables (Logical)



Tablespace (Logical)



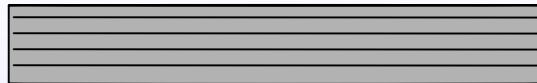
Database Files



File System



Logical Vol

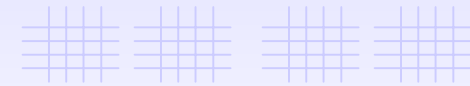


Disks

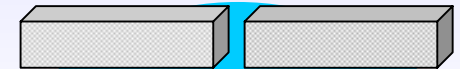


**Traditional
Storage
Management**

Tables



Tablespace

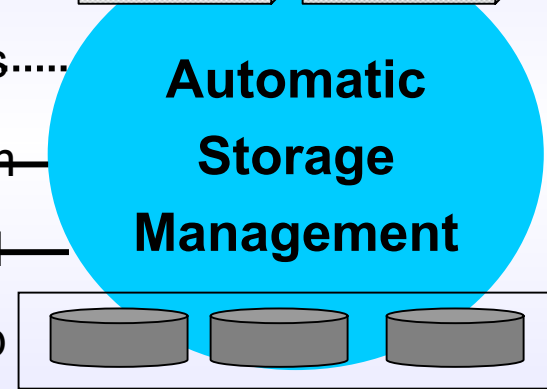


...File Names...

~~File System~~

~~Logical Vol~~

Disk Group



**Automatic
Storage
Management**

Critical DR: Automatic Storage Management (ASM provided in 10g Only) — CONS

- **Manages ONLY Oracle Database Storage**

NOT binary execs, config files, alert/error logs, source code, DDL, batch scripts, etc

- **For High Volume Mass Storage Units,
still need to manage underlying disk geometry**

E.g. logical-to-physical partition ratios, stripe width/depth, hypersplits, etc

Oracle 10g Recovery Manager (RMAN) — PROS

- **Optimized Backup Time and Disk Space**
 - Backup only used blocks
 - Incremental and differential backups — can backup only changed blocks and merge blocks to reduce number of backup pieces
 - Binary compression saves approx 50-75% disk space
 - WARNING: Can increase MTTR !!!
 - New in 10g
- **Resumeable Backup/Restore in Case of Failure**
 - Saves Time and Decreases MTTR
- **Continues Backup/Restore stream if multiple parallel/failover channels**
 - Improved Fault Tolerance
 - New in 10g
- ***Online* Block Corruption Detection and Repair/Recovery**
 - Increases Availability and Decreases or Eliminates MTTR

Oracle 10g Recovery Manager (RMAN) — **PROS**

- **Recover Tablespace CLONING now FULLY Automated**
(Auto creates clone on same server as defective db, and auto-drops when done)
 - Decreases MTTR
 - New in 10g
- **Can Control Intrusiveness via Backup *Duration* option**
 - Improves SLA Compliance ability
 - New in 10g
- **Faster Incremental Backups via intelligent delta traversal versus long file scans**
 - Improves SLA Compliance ability
 - New in 10g
- **Centralized, OS Independent Mgt of backups, recovery, and supporting infrastructure, e.g. backup scripts, configuration options, retention policies, etc**
 - Improves Maintainability thru Centralized, Simplified, Consistent Mgt

Oracle 10g Recovery Manager (RMAN) — PROS

- **Tight Integration with Flash Recovery Area**
(To automate Retention and Reclamation management
+ can configure automated Notification of low disk space and backup obsolescence)
→ More Maintainability Improvements
- **Fine-granular Data Operations vs 3rd Party Technologies**
(e.g. tablespace, data file, archive log, control file, block levels)
→ Provides variety of Recovery Scenarios to accommodate MTTR
- **Fine-granular Time Based Operations**
(e.g. recovery to present, specific point in time, or specific SCN)
- **Extensible to third party storage media managers**
(e.g. to access tape libraries and integrate with ASM)

Oracle 10g Recovery Manager (RMAN) — CONS

- Protection is limited to individual databases
- Must complement with other framework for **ADVANCED HA/DR** such as site failover

Oracle 10g Flashback Technology

- **Fast Recovery:** In just minutes
- Like a **REWIND BUTTON** (slider)
- **Easy to use** (particularly in 10g)
- **Quick** to analyze and repair
- Fixes data at **Row, Table** and **Database** Levels
- Operates on **just changed data blocks**

Oracle 10g Flashback Technology — Function Levels

9i/10g

✓ Oracle Flashback Query

Causes
NO Data
Change

✓ Oracle Flashback Version Query

✓ Oracle Flashback Transaction Query

10g

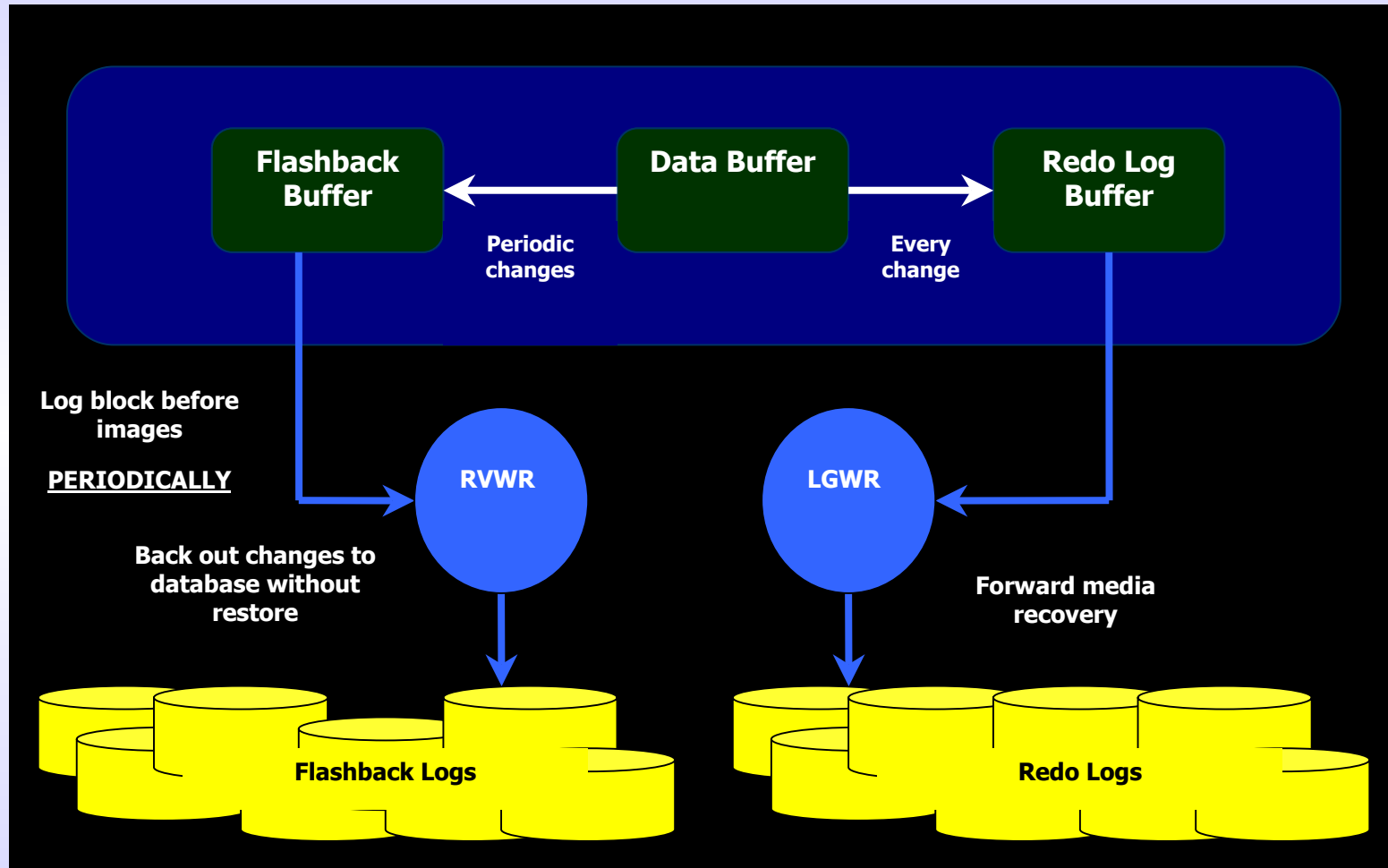
✓ Oracle Flashback Table

✓ Oracle Flashback Database

✓ Oracle Flashback Drop Table

Causes
Data
Change

Oracle 10g Flashback Database Architecture



Oracle 10g Flashback and UNDO

- **Limited by the amount of UNDO data retained with the UNDO_TABLESPACE**
- **Flashback time is bounded by the UNDO_RETENTION parameter**
- **Enough undo information to reconstruct the snapshot.**
- **Can be enabled at a session level.**
- **HOW FAR into the past depends on above UNDO info**

Oracle 10g Flashback — PROS

FAST, Easy, Reliable Recovery from Logical/Human Errors:

- **Accessible through Well Known Interfaces:**
SQL*Plus and RMAN
 - Easy to Automate (script) — Promotes Reliability
- **Flashback Database:** Effectively REWINDS entire database
 - Decreases MTTR
 - New in 10g
- **Flashback Table:** ~*Online* Objects Level Restoration
 - Increases Availability for DR/Diagnostics BUT Acquires Exclusive Locks
 - Decreases MTTR (often faster to restore just table)
 - High Automation ALSO restores SUB-objects such as indexes
constraints, triggers — can even REVERT table to original state when done
 - New in 10g
 - **WARNING** — Does NOT work on SYS or After: Truncate, Alter Table Move, Alter Table DDL such as Drop Column or Drop Partition, or Create Cluster

Oracle 10g Flashback — PROS

...CONTINUED... (FAST, Easy, Reliable):

- **Flashback Drop:**
Can See and UNdrop a variety of objects until “PURGE” command is used or automatic reclamation event
→ New in 10g
- **Flashback Query — Reduced Limitations in 10g:**
 - OLDER than 5 Days (9i is 5 Day max)
 - Every 6 Seconds (9i is every 5 Minutes)Can cause SMON_SCN_TIME table to store older data by setting appropriate init parms (e.g. undo_retention, undo_management=auto) + UNDO tablespace setting retention_guarantee.
→ Improved in 10g

Oracle 10g Flashback — PROS

...CONTINUED... (FAST, Easy, Reliable):

- **Flashback Versions Query:**

AMAZING diagnostic/auditing feature which facilitates recovery from logical corruption or construction of audit trail information

→ New in 10g

- **Flashback Transaction Query:**

Once Flashback Versions Query has identified one or more responsible transaction ids, you can use Flashback Transaction Query to further investigate the suspect transaction(s)

→ New in 10g

- **It's FREE :-)**

Oracle 10g Flashback — CONS

- **Restricted to Same-Database Operations**
- **Rewind Limitations:**
 - FB Table: No Rewind past Truncate Table, Create Cluster, Alter Table Move, Alter Table DDL (e.g. drop column or drop partition)
 - FB Table: Not Supported for Sys/System Tables, External Tables, Temp Tables, Views
 - FB Table: Statistics are Not Restored (but CAN rewind past analyze)
 - FB DB: No Rewind past Control File Restore or Recreation
 - FB DB: No Rewind past Data File or Tablespace Drop
 - FB DB: No Rewind past Manual RESIZE of a Data File
 - Both: Acquires Exclusive DML Locks (even though “Online” operation)
- **Query Limitations:**
 - Every 5 Minutes (9i) or 6 Seconds (10g)
 - Older than 5 Days (9i) — Artificial Restriction Lifted in 10g

Oracle 10g Flashback — CONS

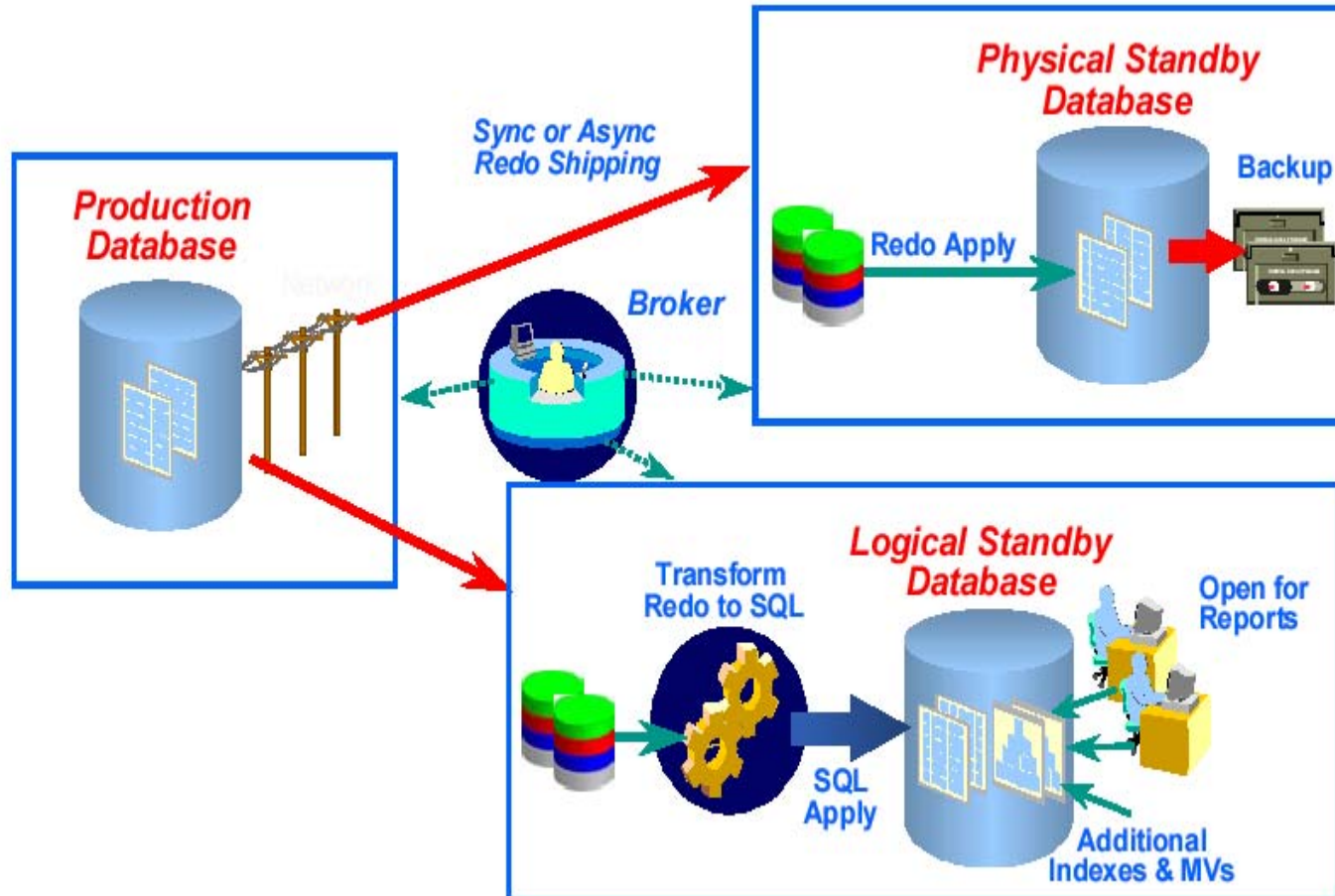
- **UNdrop Limitations:**

- Works only on NON-Sys/System, Locally Managed Tablespaces.
- Tables with Fine Grained Access policies are NOT protected by the recycle bin.
- Partitioned index-organized tables are not protected by the recycle bin.
- The recycle bin does not preserve referential integrity.

Miscellaneous Restrictions (not highly significant for HA/DR planning):

- No fixed size for the recycle bin. The time an object remains in the recycle bin can vary.
- Objects in the recycle bin are restricted to query operations only (no DDL or DML).
- Flashback query operations must reference the recycle bin name.
- Tables and all dependent objects are placed into, recovered, and purged from the recycle bin at the same time.

Oracle DataGuard / Standby — Architecture



Critical DR: Oracle DataGuard / Standby — PROS

- **Excellent for SITE Failover and Planned Switchovers**
 - New in 10g: Rolling Upgrades now officially supported
- **Use Standby Database for REPORTING + BACKUPS**
 - **WARNING for Reporting: Better in Logical vs Physical Standby**
 - **WARNING for Backups: Consider Network Communications in ALL Recovery Scenarios (including recovery to primary site).**
- **Intentional Database Latency Easy to Implement**
- **Integrated with Oracle RAC and FLASHBACK**
 - **WARNING Prior to 10g: DG not supported with RAC via DG Broker or OEM GUI due to bug reading redo threads.**
 - New in 10g
- **Improved OEM GUI Database Control Web Console**
 - (Uses DG Broker under the hood – versus direct extended SQL and PL/SQL)
 - Improves Reliability via increased Usability
 - New in 10g
- **Zero Downtime Instantiation of Standby DB**
 - (No need to Quiesce or Shutdown to install Resource Manager on Primary DB — necessary info now in Standby Control File)
 - Improves Availability
 - New in 10g

Critical DR: Oracle DataGuard / Standby — PROS

- **It's FREE for Oracle Enterprise :-)**
- **Does Not Propagate Disk Block Level Corruption**
- **Mid Level DBA Maintenance** — Easier than Some Other DR Options
(Except for NoLogging. UnRecoverable table/index operations and certain infrastructure operations, e.g. add data file)
- **Small to Zero Data Loss**
 - Oracle 9i/10g reduce to Minimal and Zero data loss without manual intervention or risk in shipping current redo after failure. Supports “Maximum Protection” and “Maximum Availability” modes (very intrusive on performance especially if communication bandwidth not abundant) as well as “Maximum Performance” modes which writes to remote redos Asynchronously.
- **Single Vendor Support — Very Standard/Proven Solution**
- **Minimal Performance Impact on Primary Server/DB**
(REALITY requires some impact since tend to use smaller REDO log size OR force frequent log switches)
 - Oracle 9i resolves this, e.g. “Real Time Apply” feature propagates REDO vs Archive Logs
- **Minimal EXTRA Impact on Network Bandwidth**
(Archive Logs should be duplexed anyway, although ZERO data loss requires extra bandwidth)

Critical DR: Oracle DataGuard / Standby — CONS

- Some Data Loss Very Likely
→ Oracle 9iR2+10g resolves this (standby redos + Real Time Apply)
- Negative Primary Server/Database Performance Impact if reduce potential data loss by reducing Redo Log size
(Possibly N/A if using small redos for regular archiving also)
→ Oracle 9iR2+10g resolves this (standby redos + Real Time Apply)
- Limited Ability for Failover Database to support Read-Only Users
(Data load/synch must be suspended while db mounted read-only — can increase MTTR)
→ Oracle 9iR2 resolves this (“Logical Standby” converts log trxs to SQL trxs versus using logs directly in Recovery Mode)
- Increased Chance of Infrastructure/Integrity Problems in Failover Database — Can Cause Failover Usage Problems
(Cumbersome to propagate non-archived data/operations — requires manual procedures for NoLogging, Truncate, etc plus infrastructure operations, e.g. add data file)
→ Oracle 9iR2 improves this (not fully resolved) via supplemental archiving mode, but more intrusive (adjust capacity plans, e.g. some extra space and performance requirements)

Critical DR: Oracle DataGuard / Standby — **CONS**

- **Can Write to Secondary DB BUT CANNOT Replicate Back**
(No support for 2/multi-way replication – see “Option 2” for Utilization Metric)
→ **Consider Streams or Advanced Multi-Master Replication — see Pros/Cons there**
- **High SwitchBACK / FailBACK Time — Cumbersome Process**
(Must recreate database from scratch/backup)
→ **Oracle 10g Resolves this (use Flashback database, then apply archive logs)**
- **Planned Downtime Required for Primary AND Secondary databases simultaneously during database upgrades/patches**
(Particularly if change in Oracle’s redo/archive log structure, although the downtime can be minimized)
→ **Oracle 10g Resolves this (Rolling Upgrade officially supported)**

Critical DR: Quest SharePlex — PROS

SAME AS FOR ORACLE DATA GUARD Except Add the Following Advantages...

- Zero or Near-Zero Data Loss
(Feeds directly from the redo logs and now de-escalates if trx throughput causes redo wrap-around before copy is complete — then escalates back to redo logs after Oracle catches up)
- Can Propagate JUST USER-SPECIFIED TABLE/INDEX **ROWS**
- Secondary DB can be used Read-Only WITHOUT Restriction
→ Oracle 9iR2 resolves this difference (see previous slide)
- Secondary DB allows Structural Differences to Optimize for Reporting (e.g. different indexes, partitions, etc.)
→ Oracle 9iR2 resolves this difference (see previous slide)
- Secondary DB allows WRITES **and** Replicates Back
(2/multi-way replication supports robust conflict resolution *as well as* monitoring and exception/comparison analysis tools — easier and more comprehensive than Oracle advanced replication)

Critical DR: Quest SharePlex — PROS

- **Propagates LONG Data Types (in addition to LOBs)**
(Although more modern applications use LOBs instead of LONG data types, Oracle Advanced Replication does not propagate LONGs — and it is Oracle's only 2/multi-way replication vehicle)
- **Decreased Planned Maintenance Time**
(Maintenance Downtime NOT required for Primary AND secondary databases simultaneously during database upgrades/patches, even if change in Oracle's redo/archive log structure)
- **Low FailBACK Time (Simple vs Oracle DataGuard/Standby)**
(Recreate Primary database with virtually a few clicks using *Overdrive*)
 - **Oracle 10g improves this (see previous slide)**
 - WARNING: If recreate primary db from a secondary db that has structure changes, modified to improve reporting performance (e.g. indexes/partitions/etc), those structures will be recreated the same way in the primary database — probably need manual intervention in this case.

Critical DR: Quest SharePlex — CONS

- NOT FREE (compared to Oracle Data Guard)
- Increased Chance of Infrastructure/Integrity Problems in Failover Database — Can Cause Failover Usage Problems (cumbersome to propagate non-archived data/operations — requires manual procedures for NoLogging, Truncate, etc plus infrastructure operations, e.g. add data file)
 - Same as Oracle DataGuard in that manual procedures are required for NoLogging, Truncate, etc plus infrastructure operations, e.g. add data file
 - WORSE than Oracle DataGuard in that SharePlex does not propagate CREATE/DROP DDLs or ALTER TABLE DROP/DISABLE COLUMN
 - v4.5+ improves this (less restrictions on propagated data/operations)
 - Improves even MORE (not completely) by using Oracle 9i R2+ supplemental archive mode...
but more intrusive (adjust capacity plans, e.g. some extra space and performance requirements)

Critical DR: Oracle Advanced Replication — PROS

- Secondary Database can be used extensively for BOTH Read-Only plus WRITING (updates/inserts/deletes)
- HW and Operating System Independent
(Remote sites can be heterogeneous as long as NET8 or higher is supported)
- Can HELP Performance in Some Scenarios
 - Data Proximity / Geographic Data Caching
 - Site-Level Load Balancing
(although overhead may offset advantages and not incredibly scalable)
- Interoperability with Non-Oracle Databases using Optional Database Gateways
- ZERO Down Time, although with Potentially Significant Data Loss
(Typically still need users to reconnect unless supporting Oracle Transparent Application Failover/TAF which requires no coding mods for OCI compliant apps as of Oracle 8.1.6)

Critical DR: Oracle Advanced Replication — CONS

- Significant Performance Impact on Primary Database AND Network Bandwidth
(Performance impact continues to exist even after v8i improvements such as internalized triggers and data streaming vs vanilla two-phase commit propagation)
→ TIP: Using custom **PROCEDURAL REPLICATION** for some operations can substantially reduce the amount of data being propagated, substantially increasing performance — use with Caution though (typically with batch process which writes many thousands or more rows in short time).
- Significant Data Loss Potential due to Latency
- Limitations on Data Type Propagation
 - No LONGS (although LOBs are propagated)
 - Some other restrictions — but getting better
→ **Oracle 9i propagates UDTs/Objects and 8.1.7 propagates CLOB indexes.**
 - No Sequence Infrastructure/DDL (only the generated sequence values for the target table)
→ **Technically, sequence values can be staggered across databases, e.g. odd/even or every N, BUT business restrictions may preclude staggering or even max+1 approach.**
@TBD: Consider using Procedural Replication (e.g. DBMS_DEFER pkg) for sequence propagation, BUT Data LATENCY may add addtl complexity.

Critical DR: Oracle Advanced Replication — CONS

- **Complex Planning/Admin for Propagating DDL**
(using DBMS_REPCAT and REPUTIL packages, e.g. REPUTIL.from_remote)
 - This compares to “Increased Chance of Infrastructure/Integrity Problems in Failover Database” in Oracle DataGuard and Quest SharePlex CONS sections, but for Advanced Replication envmts, we don’t expect Infrastructure such as data FILES to be propagated, only the contents of those files, e.g. data/indexes/etc.
- **Increased Admin/Monitoring Requirements**
(e.g. **high water mark on replication logs**; increased **temp space** usage; **rollback space** usage especially when purging jobs; **quiescing** required when performing certain admin tasks; constant monitoring of **Deferred Transaction Queue** DEFTRAN and DEFERROR)
- **Dependent on Hard-Coded Passwords for DB Links**
(security consideration for repadmin user)
- **Conflict Resolution is NOT As Easy As Two Words!!!**
(**IF** secondary database used for WRITES, conflict resolution involves COMPLEX ONGOING planning, monitoring, administration, and possibly coding)

Critical DR: Oracle STREAMS — PROS

- 10g is Much Faster and Less Performance Degradation than Advanced Replication — Uses LOG SCRAPING/Mining
- Improved Scalability
(E.g. new STREAMS_POOL_SIZE so no longer need to over-allocate general Shared Pool)
→ New in 10g
- Improved Integration with RAC
(Auto Failover from Primary to Secondary in Active-PASSIVE implementation)
→ New in 10g
- Improved Data Types Support — including Long, Long Raw
→ New in 10g: ALSO incl NCLOB, Binary, Float, Binary Double as well as
IOTs (and Function Based Indexes + Descending Indexes ???)
- Cascading Replication — “DIRECTED”
→ New in 10g

Critical DR: Oracle STREAMS — PROS

- Improved Ability to Monitor — New Data Dict Views
 - New in 10g
- Hot Mining of Current Redo
 - Reduces Data Loss
 - New in 10g
- Write Order Easier to Maintain
 - Versus Advanced Replication DEFERROR Order By
 - No Replication Groups (has DISadvantages too, e.g. during Maintenance)
- It's FREE :-)

Critical DR: Oracle STREAMS — CONS

- Still CUMBERSOME to Instantiate and Maintain
— Especially for Many Tables
- More Potential for Data Loss e.g. versus Data Guard or RDM
 - Due to Human Error (Usability)
 - Due to Latency Management
- LATENCY ISSUE for 9i: Frequency of Data Propagation
(job_queue_interval versus UNdocumented _job_queue_interval)

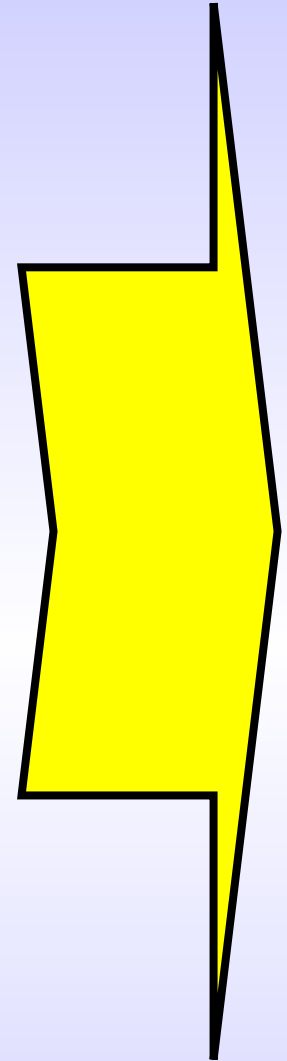
GOTO VERBAL MODE

DO NOT PASS GO.

DO NOT COLLECT \$2000.

- See upcoming revisions for additional 9i RAC info + Remote/Geographic Disk Mirroring HA/DR technologies, incl:
- Vyant RealTime
 - EMC SRDF/TimeFinder
 - SNAP Technology

For latest copy: jefflit@dbigusa.com or 703-405-5432



So Now What?

THE BIG QUESTION...



Will We SUCCEED?



Summary

- We Covered:
 - Technical Framework
 - Major Pros and Cons
- Your most challenging problems or nightmares?
- **Now is Time for:**
 - **Specific Configuration Questions, Etc.**
 - **General Q&A**
 - If no time today, **Can Follow-up Offline**

*Questions...
Comments?*

Thank You!

Jeffrey Bertman

**Chief Engineer
DataBase Intelligence Group (DBIG)
(jefflit@dbigusa.com, 703-405-5432)
WDC / VA / MD**

Ran Pan, PhD

**Lead Technologist, PCAOB
(panr@pcaobus.org, 703-547-6420)
WDC / VA / MD**

- IT Consulting • Strategic Planning • On-Call Support •
- High Volume (VLDB) and High Transaction Specialists •

AOTC Conference 12/8/2005