MaxDB – That’s Cool!

Joerg Hoffmeister, SAP AG
Germany

MySQL Users Conference, 2005
April 18-21, 2005
What’s Cool?

Jazz is cool
Some people are cool

Databases are boring – but necessary

Software is boring – and necessary too (in some cases ;-))

But single features can really be cool!
Some of MaxDB’s Cool Features

MaxDB has

… no needs for reorganization
… a complete set of backup functionality
… HA scenarios
… database snapshot functionality
… a synchronization manager
No Need for Reorganisation
No Reorg – Basics

PREREQUISITES
Space no longer needed must immediately be returned to the space management (database)
The occupation ratio of the data blocks should be held at the highest possible level
Data must be stored dense without gaps that would have to be managed.

ACCESSOIRES
Shadowing pages with our Converter
Automatic space- and I/O management
No Reorg – Basics

Handling Space
Update in place
Sort by insertion
Delete in place

Storing Data
B*Trees for almost all database items:
Tables
Indexes (secondary keys)
BLOBs
B*- Trees

K: Key Column
I: Indexed Column
L: BLOB Column

Primary Data

Index Data

BLOB data

Joerg Hoffmeister
MaxDB – That’s Cool
I/O-Concept

Concurrent asynchronous I/O

Shadow page concept
  Converter
  Managing free blocks
  Savepoints
  Backup Integration

Datawriter and server tasks

Segmentation of the data cache
Backup and Snapshot are pointers that will stay pointing to the origin if a copy is established.
SAVEPOINT

Controlled Data flush

Supported through write ahead tasks and concurrent server tasks

Storage universe

Backup

SNAPSHOT

Ex-ORIGIN

COPY ORIGIN
Converter, Free Block Mgmt, Savepoints

Converter maps current location(s) of data blocks
   Knows where block images are located in storage

Free block management provides locations to converter
   Knows about available blocks,
   when blocks can be released,
   how to combine blocks for optimized I/O

Savepoints synchronize I/O management
   Scheduled every 10 minutes (if enough volume appears),
   supported by 'write ahead' tasks (data writer)
Redundancy

Blocks are short term kept twice if modified
  Copy becomes origin at savepoint. Ex-origin is released except for backup or snapshot reasons.

Blocks are short term kept twice if modified and backed up
  Blocks are released when backup has passed them

Blocks are long term kept twice if modified and 'snapshotted'
  Blocks are released when snapshot is dropped
Backup/Recovery
Backup

Every MaxDB data backup is consistent and could be restarted even without recovering log (e.g. unsharp copies)

A savepoint is issued before the backup starts
The data belonging to this savepoint additionally contains all undo information for remaining open transactions. Thus each data backup is consistent in terms of transactions.

The savepoint checks the converter, detects those blocks relevant for the backup and marks the backup flag for those blocks in the free block manager (FBM)

A backup is executed by concurrent server tasks along informations of the free block management.
Backup Wizard

Backup Database Instance

What type of backup do you want to perform?

- Complete Data Backup
- Incremental Data Backup
- Log Backup
- Activate/Deactivate Automatic Log Backup

Last Complete Data Backup:
- Label: DAT_000000001
- Date: 24.02.2005 13:33:51
- Medium: ULF2_DB
- Volumes: 1
- Size: 960 Pages
- Log Page: 5984

Last Incremental Data Backup:
- Never

Last Log Backup:
- Never

<Back  Next>  Cancel
Recovery Wizard

Type of Recovery
Select the type of recovery that you want to perform.

- Restore last backup
- Restore a specified backup from history
- Restore a medium
- Continue restoring increment/Log
- Restore database until a specific time.

MAXDB2
State
- MAXDB2
- ADMIN

General
- Name
- Version
Backup & Recovery Variety

**Backup Types**
- DAT#  Complete
- PAG#  Incremental
- LOG#  Log backup

**Recovery path variety along multiple backup generations**

<table>
<thead>
<tr>
<th>Recovery path variety along multiple backup generations</th>
<th>DAT3</th>
<th>DAT3</th>
<th>DAT3</th>
<th>DAT2</th>
<th>DAT2</th>
<th>DAT1</th>
<th>DAT1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PAG4</td>
<td>PAG3</td>
<td>LOG5</td>
<td>PAG2</td>
<td>LOG3</td>
<td>PAG1</td>
<td>LOG1</td>
</tr>
<tr>
<td></td>
<td>LOG8</td>
<td>LOG7</td>
<td>LOG6</td>
<td>LOG5</td>
<td>LOG4</td>
<td>LOG2</td>
<td>LOG2</td>
</tr>
<tr>
<td></td>
<td>LOG9</td>
<td>LOG8</td>
<td>LOG7</td>
<td>LOG6</td>
<td>LOG5</td>
<td>LOG3</td>
<td>LOG3</td>
</tr>
<tr>
<td></td>
<td>LOG9</td>
<td>LOG8</td>
<td>LOG7</td>
<td>LOG6</td>
<td>LOG4</td>
<td>LOG4</td>
<td>LOG4</td>
</tr>
<tr>
<td></td>
<td>LOG9</td>
<td>LOG8</td>
<td>LOG7</td>
<td>LOG5</td>
<td>LOG5</td>
<td>LOG5</td>
<td>LOG5</td>
</tr>
<tr>
<td></td>
<td>LOG9</td>
<td>LOG8</td>
<td>LOG6</td>
<td>LOG6</td>
<td>LOG6</td>
<td>LOG6</td>
<td>LOG6</td>
</tr>
<tr>
<td></td>
<td>LOG9</td>
<td>LOG7</td>
<td>LOG7</td>
<td>LOG7</td>
<td>LOG9</td>
<td>LOG8</td>
<td>LOG8</td>
</tr>
<tr>
<td></td>
<td>LOG9</td>
<td>LOG9</td>
<td>LOG9</td>
<td>LOG9</td>
<td>LOG9</td>
<td>LOG9</td>
<td>LOG9</td>
</tr>
</tbody>
</table>
Backup History

Database Manager

Backup History

<table>
<thead>
<tr>
<th>Label</th>
<th>Action</th>
<th>Start</th>
<th>Result</th>
<th>Medium</th>
<th>Size (Pages)</th>
<th>Volumes</th>
<th>Next Log Page</th>
<th>From Page</th>
<th>To Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOG_000000004</td>
<td>SAVE WARM</td>
<td>14.03.2005 21:54:07</td>
<td>OK</td>
<td>MDM_LOG</td>
<td>1016</td>
<td>1</td>
<td>3702</td>
<td>4720</td>
<td></td>
</tr>
<tr>
<td>PAG_000000002</td>
<td>SAVE WARM</td>
<td>14.03.2005 21:52:40</td>
<td>OK</td>
<td>MDM_INCREMENT</td>
<td>4248</td>
<td>1</td>
<td>4602</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOG_000000003</td>
<td>SAVE WARM</td>
<td>14.03.2005 21:52:08</td>
<td>OK</td>
<td>MDM_LOG</td>
<td>1016</td>
<td>1</td>
<td>2698</td>
<td>3701</td>
<td></td>
</tr>
<tr>
<td>LOG_000000002</td>
<td>SAVE WARM</td>
<td>14.03.2005 21:52:05</td>
<td>OK</td>
<td>MDM_LOG</td>
<td>1016</td>
<td>1</td>
<td>1694</td>
<td>2697</td>
<td></td>
</tr>
<tr>
<td>LOG_000000001</td>
<td>SAVE WARM</td>
<td>14.03.2005 21:52:04</td>
<td>OK</td>
<td>MDM_LOG</td>
<td>1016</td>
<td>1</td>
<td>0</td>
<td>1833</td>
<td></td>
</tr>
<tr>
<td>DAT_000000001</td>
<td>SAVE WARM</td>
<td>14.03.2005 21:51:26</td>
<td>OK</td>
<td>MDM_COMPLETE</td>
<td>240</td>
<td>1</td>
<td>1058</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HISTLOST</td>
<td></td>
<td>14.03.2005 21:50:35</td>
<td>OK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HISTLOST</td>
<td></td>
<td>14.03.2005 21:50:35</td>
<td>OK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Guided Recovery

Select Items for Recovery
Select items to specify the use of incremental backups.

Select the incremental backup items you want to use for recovery. The use of incremental backups is much faster than the recovery of logs. We recommend that you use the specified incremental backups.

Start Recovery
Check your selection and start the recovery.

You have completed the steps required to perform a recovery. Your recovery is defined below:

- DAT_00002 CompleteBU MAXDEMO_COM_01
- DAT_00002 CompleteBU MAXDEMO_COM_02
- PAG_00003 INC-BU MAXDEMO_INC
- LOG_00043 logautomatic MAXDEMO_LOG.048
- LOG_00043 logautomatic MAXDEMO_LOG.049
- LOG_00050 logautomatic MAXDEMO_LOG.050
- LOG_00051 logautomatic MAXDEMO_LOG.051
- LOG_00052 logautomatic MAXDEMO_LOG.052

Make the specified medium available for recovery. Choose ‘Start’ to begin the recovery. If you want to restore the database until a particular point in time, specify the date and time.

Restore database until a specific time.

Database Manager
Snapshot
Snapshot

Freezing a database image

Create Snapshot (ADMIN)
Revert to Snapshot (ADMIN)
Drop Snapshot (ADMIN)

Ideas of use:
Very fast point in time resetting (e.g. during upgrades)
Restoring training-systems to a defined status
Master - Slave Support w/ Snapshots

Master

Data 01.01.2005

Complete

Slave

Data

Create Snapshot

Data 07.01.2005

Incremental

Restore Snapshot

Data 14.01.2005

Incremental

Data 07.01.2005

Incremental

Data 14.01.2005

Incremental
Snap Demo: Database after setup
Let's take a snapshot
Now some data has been loaded
Proof: Data is there
Return to admin: prepare reverting

### Database Manager

**Name** | **State** | **Data** | **Log** | **Sessions** | **Data Cache Hit** | **Auto Log**
--- | --- | --- | --- | --- | --- | ---
MAXDB1 | Offline | | | | | |
MAXDB2 | Offline | | | | | |
MAXSNAP | Admin | | | | | |
MDSM | Not connected | | | | | |
MDSLA | Not connected | | | | | |

### MAXSNAP

**MAXSNAP**

**Data:**

**Log:**

**Sessions:**

### General

<table>
<thead>
<tr>
<th><strong>Name</strong></th>
<th><strong>Version</strong></th>
<th><strong>Operating System</strong></th>
<th><strong>Run Directory</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>MAXSNAP</td>
<td>7.5.00.07</td>
<td>Windows XP (WIN32)</td>
<td>c:\documents and settings\all users\application data\db\data\work\MAXSNAP</td>
</tr>
</tbody>
</table>
Action: Revert to snapshot
After reverting to initial state
Standby database
Standby Database (w/ Log Shipping)

- Master
- Standby
- Data
- Log
- Log Backup
- Staging Directory
- Log Backup
- Log Restore
- Initial data backup
Standby Database: Demo Example

MDMaster

Log

C:\ShadowMedia

MDSlave

Log

Data

Log Backup

Log Restore

Initial data backup

C:\shadowmedia\MDMASTER_COM
Standby Database: Demo Example

![Diagram showing Standby Database with MASTER, SLAVE, and FILESYSTEM sections]
Standby Database: Slave Steps

Standby Mode
Initialize once with complete backup from Master
Redo logs as they appear

Start Slave to online mode in case of emergency
   If possible, back up last piece of log from master
   Redo all 'open' log backups (there should be none)
   Redo final log piece
   Restart slave to be the new master
Standby Database: Slave Steps live
Standby Database: Slave Steps live (2)
These steps are required if slave should be started
Standby Database online!

After a log recovery the database instance has **successfully restarted**. Choose 'Close' to end the Recovery Wizard.
... but looks smaller ... ?
Indexes
due to the time impact are not created immediately during log recovery. It is left to the administrator when this fits to the schedule.
Now: Same Size!
Hot Standby – standby within seconds

![Diagram of Hot Standby system with labels: Master, Standby, Cluster, Application, RECONNECT, IP SWITCH, After Images, Data, Archive Log, Data, Storage System, continuous RESTART.]
Synchronization Manager
MaxDB - Synchronization Architecture

- Master DB
- Synchronization Service
- Message Server
- Message DB
- Synchronization Service
- Client 1 DB
- Synchronization Service
- Client 2 DB
- Synchronization Mgr GUI
MaxDB - Sync Manager Features

Replication of tables from a master database to client databases

Synchronization Manager decouples master and client DBMS's

Initial replication of the complete DBMS state

Delta replication of transactions

Point-to-point replications (queues)

Broadcast replications (publish/subscribe)

Support for bi-directional replications with conflict resolution

Admin tool to define replication scenarios

Supports MaxDB and MinDB
Example: Master table for Synchronization
SyncMan GUI – Replication Tables

Column Group 'Country'

Column Group 'District'
Vertical Projection: Bob's Client Tab

Column Selection
Horizontal Selection: Mona's Client Tab
SyncMan: Monitor

These tables are created on unit activation
Bob adds a record – but only to his columns
The view of the master shows NULL values
Mona updates the row's missing columns
Now this is the final master view
THE END

Thank you!
Have a nice further conference

joerg.hoffmeister@sap.com
• No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP AG. The information contained herein may be changed without prior notice.

• Some software products marketed by SAP AG and its distributors contain proprietary software components of other software vendors.

• Microsoft, Windows, Outlook, and PowerPoint are registered trademarks of Microsoft Corporation.

• IBM, DB2, DB2 Universal Database, OS/2, Parallel Sysplex, MVS/ESA, AIX, S/390, AS/400, OS/390, OS/400, iSeries, pSeries, xSeries, zSeries, z/OS, AFP, Intelligent Miner, WebSphere, Netfinity, Tivoli, and Informix are trademarks or registered trademarks of IBM Corporation in the United States and/or other countries.

• Oracle is a registered trademark of Oracle Corporation.

• UNIX, X/Open, OSF/1, and Motif are registered trademarks of the Open Group.

• Citrix, ICA, Program Neighborhood, MetaFrame, WinFrame, VideoFrame, and MultiWin are trademarks or registered trademarks of Citrix Systems, Inc.

• HTML, XML, XHTML and W3C are trademarks or registered trademarks of W3C®, World Wide Web Consortium, Massachusetts Institute of Technology.

• Java is a registered trademark of Sun Microsystems, Inc.

• JavaScript is a registered trademark of Sun Microsystems, Inc., used under license for technology invented and implemented by Netscape.

• MaxDB is a trademark of MySQL AB, Sweden.

• SAP, R/3, mySAP, mySAP.com, xApps, xApp, SAP NetWeaver and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP AG in Germany and in several other countries all over the world. All other product and service names mentioned are the trademarks of their respective companies. Data contained in this document serves informational purposes only. National product specifications may vary.

• These materials are subject to change without notice. These materials are provided by SAP AG and its affiliated companies ("SAP Group") for informational purposes only, without representation or warranty of any kind, and SAP Group shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP Group products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.