Session ID: SPC202
Conversion of SAP Systems to Unicode
Introduction
Conversion Preparation
Database Export, Conversion and Import
The Unicode System
Customer Experiences
Additional Information and Contacts
What is Unicode?

“[Unicode is] the encoding standard [which] provides the basis for processing, storage and interchange of text data in any language in all modern software and information technology protocols.”

see: [http://www.unicode.org](http://www.unicode.org)

Each character has a unique number („Unicode scalar value“).

- Notation U+nnnnnnnn (where nnnnnnnn are hexadecimal digits)
Unicode...

...allows text data from different languages to be stored in one repository.

...enables a single set of source code to be written to process data in virtually all languages.

...simplifies the addition of new language support to an e-business application since character processing and storage remains unchanged.

✓ Lower costs of implementation!

✓ Faster speed to market!

✓ Better customer satisfaction!
Who needs Unicode?

Acting in global business requires support of a Global Character Set!

Companies running global business processes like Global HR Systems or Global Master Data Management

Companies offering Web Services to their customers: Global Master Data containing multiple local language characters!

Companies using Open Standards: J2EE and .NET integration (JAVA speaks Unicode!)

Collaborative Business: Integration of Third Party Products that run on different code pages
Application Impact: Global Master Data

Example: CRM Business Partner
Internationalized Software with Unicode

The **Unicode** Standard was adopted by industry leaders e.g. Apple, HP, IBM JustSystem, Microsoft®, Oracle, SAP, Sun™, Sybase, Unisys and many others.

**Unicode** is required by modern standards such as XML, Java™, ECMAScript (JavaScript™), LDAP, CORBA 3.0, and WML.

**Unicode** is the official way to implement ISO/IEC 10646 and is supported in many operating systems and all modern browsers.
**SAP Pre-Unicode Solutions**

**Single Code Page System**
One standard system code page → only a limited set of languages is supported in these systems.

**Blended Code Page System (R/3 3.0D – R/3 4.6D)**
Multibyte blended code pages which contain characters from several standard code pages. 
Blended code pages are not standard code pages, but SAP-customized code pages that were devised to support an increased number of possible language combinations in a single code page. 

a) Ambiguous Blended Code Page System:
   Two characters can share the same code point. 

b) Unambiguous Blended Code Page System:
   Each code point refers exactly to one character.

**MDMP System Configuration (R/3 3.1I – NetWeaver 2004)**
„Multiple Display/Multiple Processing“: MDMP Systems deploy more than one system code page on the application server. This allows languages to be used together in one system although the characters of those languages are not covered by the same code page.
Why convert to Unicode?

**Unicode defines the character set for...**
...efficient text processing in any language
...maintaining text data integrity

**Unicode systems...**
...need only one locale (Unicode locale = platform-independent)
...integrate seamless in existing system landscape (SAP and non-SAP systems)
...provide all ISO 639-2 language keys and 86 additional country-specific language keys
→ a total of 560 technically supported language keys!

**Users can...**
...enter and display any character from any script no matter which logon language they use
...print text data in multiple languages
Why convert to Unicode?

One character set in system landscape

Easy handling for endusers

Lower operation costs
SAP NetWeaver™ the integration platform?

Evolution of mySAP Technology

Unifies and aligns people, information and business processes

- Integrates across technologies and organizational boundaries
- A safe choice with full .NET and J2EE interoperability

The business foundation for SAP and partners

- Powers business-ready solutions that reduce custom integration
- Its Enterprise Services Architecture increases business process flexibility
SAP NetWeaver™ with non-Unicode ABAP stack

Evolution of mySAP Technology

Unifies and aligns people, information and business processes

- Integrates across technologies and organizational boundaries
- A safe choice with full .NET and J2EE interoperability

The business foundation for SAP and partners

- Powers business-ready solutions that reduce custom integration
- Its Enterprise Services Architecture increases business process flexibility
Only solution for full integration: Unicode

Evolution of mySAP Technology

Unifies and aligns people, information and business processes

- Integrates across technologies and organizational boundaries
- A safe choice with full .NET and J2EE interoperability

The business foundation for SAP and partners

- Powers business-ready solutions that reduce custom integration
- Its Enterprise Services Architecture increases business process flexibility
Unicode Conversion Project

Pre-Conversion

Set up the Unicode Conversion Project
Check Prerequisites
Plan database downtime during conversion
Enable your Customer Developments

Highly automated process
Data Analysis – special MDMP treatment
Unload /reload process for small databases
Minimum downtime tool for large databases

Post-Conversion

Unicode system is up and running
Verification of Data Consistency
Integration Testing focused on language handling
Agenda

1. Introduction

2. Conversion Preparation
   Demo

3. Database Export, Conversion and Import

4. The Unicode System
   Demo

5. Customer Experiences

6. Additional Information and Contacts
Prerequisites

Upgrade to Web AS 6.20 non-Unicode
For more information, read the appropriate Upgrade Guide.

Unicode-based mySAP components
For current status, see SAP Note 79991.

Unicode-enabled ABAP and C/C++ programs
For more information, please visit the following sessions:

Related Workshops/Lectures at SAP TechEd 2005
SPC250, Making Programs Unicode Enabled, Hands-on
SPC251, Unicode Interfaces - Data Exchange between Unicode and Non-Unicode Systems, Hands-on
Upgrade Paths to Unicode (SAP R/3 Enterprise)

Source System

- R/3 3.1i
- R/3 4.0b
- R/3 4.5b
- R/3 4.6b
- R/3 4.6c

Target System

SAP R/3 Enterprise Ext. Set 2.00
non-Unicode

Conversion

SAP R/3 Enterprise Ext. Set 2.00
Unicode

First upgrade, then conversion to Unicode!
Conversion Preparation: Concept

Before the system conversion to Unicode is executed, all text data stored in the database require language information.

CODE PAGE ASSIGNMENT

Single Code Page Systems/
Unambiguous Blended Code Page Systems
(ca. 90% of all customer installations)

MDMP Systems/
Ambiguous Blended Code Page Systems
(ca. 10% of all customer installations)

WHY?
Old solution for multiple languages: MDMP*

West European View

Japanese View

Korean View

* Check your system type with report RSCPINST → current configuration
Old solution for multiple languages: MDMP

West European View

Japanese View

Korean View

End of support with NetWeaver '04

(see notes 838402 and 79991)

As of release NetWeaver 04s MDMP will no longer be supported!

© SAP AG 2005, SAP TechEd '05 / SPC202 / 22
Content in Unicode System

... we need to know the language/code page of every character ...
MDMP -> Unicode Database Conversion

MDMP System
(logon language DE)

<table>
<thead>
<tr>
<th>MDMP System</th>
<th>Unicode System</th>
</tr>
</thead>
<tbody>
<tr>
<td>áØÝØÜ</td>
<td>синий</td>
</tr>
<tr>
<td>æÆ¶ðøö</td>
<td>파란색</td>
</tr>
<tr>
<td>blau</td>
<td>blau</td>
</tr>
<tr>
<td>blue</td>
<td>blue</td>
</tr>
<tr>
<td>□Å</td>
<td>緑</td>
</tr>
<tr>
<td>×ÖÜĐÝêU</td>
<td>緑</td>
</tr>
<tr>
<td>ÄÈ·îøø</td>
<td>緑</td>
</tr>
<tr>
<td>grün</td>
<td>grün</td>
</tr>
<tr>
<td>green</td>
<td>green</td>
</tr>
<tr>
<td>—î</td>
<td>緑</td>
</tr>
<tr>
<td>ÚàĐáÝêU</td>
<td>緑</td>
</tr>
<tr>
<td>»¡³ç»õ</td>
<td>緑</td>
</tr>
<tr>
<td>rot</td>
<td>rot</td>
</tr>
<tr>
<td>red</td>
<td>red</td>
</tr>
<tr>
<td>□Ô</td>
<td>赤</td>
</tr>
<tr>
<td>ØÖÜæè</td>
<td>赤</td>
</tr>
<tr>
<td>øë¶ðøö</td>
<td>赤</td>
</tr>
<tr>
<td>gelb</td>
<td>gelb</td>
</tr>
<tr>
<td>yellow</td>
<td>yellow</td>
</tr>
<tr>
<td>%œŒ</td>
<td>赤</td>
</tr>
</tbody>
</table>

Code Page Switch

- DE: 1100 → 4103
- RU: 1500 → 4103
- JA: 8000 → 4103
- KO: 8500 → 4103
Conversion Preparation: Concept

Problem:
Code page assignment in MDMP/Ambig. Blended Code Page Systems is complex, because...

ALL table rows in the database must be assigned a code page before the conversion!

Solution:
You can...
1. use existing Language Keys
2. create and maintain a Vocabulary (collect words* → assign language)

*‘Word’: entry in Vocabulary, default length 3-30 bytes, and contains characters outside range of common character set
Problem Description

MDMP system is a “Mixed code page solution“

Database contains textual data from several code pages.

Conversion of MDMP systems like a Single Code Page system leads to data destruction

Example:
MDMP System with Latin-1 (SAP 1100) and Latin-2 (SAP 1401)

Hex-Code on database: „0x50657472F9“ could be
Petrů (1401)
or
Petrù (1100)

To Do:
Clarify the original code page of all entries with special characters before the conversion.
Example: Russian text data in MDMP system

The Hex-code entry 0xE2D5DAE1E2 is found in a database table which has no language key.

MDMP system configuration

ISO-8859-1 (Western Europe)
ISO-8859-2 (Eastern Europe)
ISO-8859-5 (Cyrillic)

The entry could be either:

- English: âÕÚáâ
- Polish: âÕÚáâ
- Russian: тект

<table>
<thead>
<tr>
<th></th>
<th>E2</th>
<th>D5</th>
<th>DA</th>
<th>E1</th>
<th>E2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO-8859-1</td>
<td>â</td>
<td>Ő</td>
<td>Ú</td>
<td>á</td>
<td>â</td>
</tr>
<tr>
<td>ISO-8859-2</td>
<td>â</td>
<td>Ő</td>
<td>Ú</td>
<td>á</td>
<td>â</td>
</tr>
<tr>
<td>ISO-8859-5</td>
<td>т</td>
<td>е</td>
<td>к</td>
<td>с</td>
<td>т</td>
</tr>
</tbody>
</table>

Worst case: Russian speakers have to identify the data manually.
Table Categories

Code page dependent without Language Key (II)

The table contains text data with non-ASCII characters. The data is therefore code page dependent. The Language Key has to be identified.

Code page dependent with Language Key (I)

The table contains text data with non-ASCII characters:
- Single code page dependent
- Multiple code page dependent

ASCII (III)

The table contains only text data with ASCII characters. The data is not code page dependent. It is irrelevant which code page is used for the conversion.

After the conversion preparation each database table is assigned to one table category!
Conversion Preparation: Transaction SPUMG

Pre-Conversion
- Set up the Conversion Project
- Check Prerequisites
- Plan database downtime during conversion
- Enable Customer Developments

Unicode Conversion

Post-Conversion
- Unicode system is up and running
- Verification of Data Consistency
- Integration Testing focused on language handling

Conversion Preparation in non-Unicode system
SPUMG

Database analysis tool for collecting words without language/code page information. SPUMG creates control information for the database conversion.

**SPUMG consists of several Scan Levels:**

- Consistency Check
- Tables without Language Information
- Tables with Ambiguous Language Information
- Tables with Language Information
- Reprocess
- INDX Table Analysis
- INDX Table Repair
SPUMG Language Lists

Languages installed in your system:

a) active
b) inactive

Active flag must be set for all languages which shall be used for code page assignment!
Maintain the SPUMG settings and then initialize the worklist for the Consistency Check!

```
<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Fallback Codepage</td>
<td>1100</td>
</tr>
<tr>
<td>Active</td>
<td></td>
</tr>
<tr>
<td>Frequency (Number of Rows)</td>
<td>1</td>
</tr>
<tr>
<td>Minimum Word Length</td>
<td>2</td>
</tr>
<tr>
<td>Word separator</td>
<td>x'50' to x'7F'</td>
</tr>
<tr>
<td>Common Character Set</td>
<td>7 Bit ASCII</td>
</tr>
<tr>
<td>Collision Resolution</td>
<td>Codepage based (coarse)</td>
</tr>
<tr>
<td>Scan previously repaired INDX records</td>
<td></td>
</tr>
<tr>
<td>R3trans Language Codepage Settings</td>
<td>R3trans uses TCPDB, TCP0D, T</td>
</tr>
</tbody>
</table>
Consistency Check

Classifies tables into tables with or without language information (Table Category 1 or 2)

Checks table consistency (existence, access)

Writes control information to SPUMG control tables (Table Category, Language field)

Single Code Page and MDMP!

Consistency Check

Without Lang. Info

With Amb. Lang. Info

With Lang. Info

Reprocess

Export Control Table

Vocab.
Tables without Language Information

This scan adds words to the Vocabulary!

All tables without language information (Table Category 2) are scanned.
Tables with Ambiguous Language Information

All tables with language information (Table Category 1) are scanned. Words with an ambiguous language are added to the Vocabulary.

Only active if ambiguous language list has been maintained!

Consistency Check
Without Lang. Info
With Amb. Lang. Info
With Lang. Info
Reprocess

Control Info
Vocab.
The Vocabulary must be maintained before the database export is executed.

You can:

- Insert language keys automatically on Scan Level *Tables with Language Information*.
- Create *Vocabulary Hints* to assign a language based on other table fields.
- Manually assign language to word in the Vocabulary.
- Reuse existing language code page assignments imported from other systems or delivered by SAP.
- Automatic language assignment with Character Statistics.
- Use of language patterns.
Vocabulary

1. MDMP only!

2. Tables with Language Information

3. Number of words in current selection: 10

<table>
<thead>
<tr>
<th>Word</th>
<th>Language</th>
<th>Collision Filled by</th>
<th>Name</th>
<th>Current Date</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>grün</td>
<td>DE</td>
<td>3</td>
<td>SEEM</td>
<td>13.09.2004</td>
<td>11:06:40</td>
</tr>
<tr>
<td>Peňošo</td>
<td>KO</td>
<td>3</td>
<td>SEEM</td>
<td>13.09.2004</td>
<td>11:06:40</td>
</tr>
<tr>
<td>ñöñôñ</td>
<td>KO</td>
<td>3</td>
<td>SEEM</td>
<td>13.09.2004</td>
<td>11:06:40</td>
</tr>
<tr>
<td>Æ-ñôñ</td>
<td>KO</td>
<td>3</td>
<td>SEEM</td>
<td>13.09.2004</td>
<td>11:06:40</td>
</tr>
<tr>
<td>AÜÅÜ</td>
<td>DE</td>
<td>3</td>
<td>SEEM</td>
<td>13.09.2004</td>
<td>11:06:40</td>
</tr>
<tr>
<td>ÆÅñôñå</td>
<td>KO</td>
<td>3</td>
<td>SEEM</td>
<td>13.09.2004</td>
<td>11:06:40</td>
</tr>
<tr>
<td>ÖÜÜBEÅ</td>
<td>RU</td>
<td>M</td>
<td>SEEM</td>
<td>13.09.2004</td>
<td>11:07:35</td>
</tr>
<tr>
<td>ÚÖÜýäÅ</td>
<td>JA</td>
<td>M</td>
<td>SEEM</td>
<td>13.09.2004</td>
<td>11:07:46</td>
</tr>
<tr>
<td>UabaYeÅ</td>
<td>A</td>
<td>SEEM</td>
<td>13.09.2004</td>
<td>11:04:20</td>
<td></td>
</tr>
<tr>
<td>ÆÔÝÜÅ</td>
<td>A</td>
<td>SEEM</td>
<td>13.09.2004</td>
<td>11:04:20</td>
<td></td>
</tr>
</tbody>
</table>
All tables with language information (Table Category 1) are scanned.

This scan assigns a language to words in the Vocabulary based on the values of other tables in the database.

Problem: Vocabulary Collisions

- Consistency Check
- Without Lang. Info
- With Amb. Lang. Info
- With Lang. Info
- Reprocess

MDMP only!
Reprocess

This scan simulates R3load behavior.

It checks if the code page information in the Vocabulary is sufficient for the conversion. If not, this scan creates a Reprocess Log for the table.

(row identifier + language assignment)
The Reprocess Log contains:
- table name,
- key values,
- reason why no code page could be assigned.

Users can assign a language to each entry here. This information is used in the Unicode system to repair wrongly converted data.
INDX-type Tables: Special Treatment

Why must INDX-type tables be scanned separately?

INDX-type tables consist of:

- a transparent part which is treated like the other transparent tables in the database during the conversion preparation.

- a binary part which contains the code page information used during the EXPORT TO DATABASE /IMPORT FROM DATABASE statement.

In MDMP Systems the handling of INDX-type tables is improper in a way that a wrong code page might be stored in the binary part of the table.
## INDX-type Tables: Structure

<table>
<thead>
<tr>
<th></th>
<th>MANDT</th>
<th>RELID</th>
<th>SRTFD</th>
<th>SRTF2</th>
<th>......</th>
<th>CLUSTR</th>
<th>CLUSTD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key</td>
<td>C(3)</td>
<td>C(2)</td>
<td>C(n)</td>
<td>INT4</td>
<td>INT2</td>
<td>LRAW</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Character area**

- **Transparent Part**
- **Binary Part**
These two scans treat the binary part of INDX-type tables:

1. All INDX-type tables without language information are scanned and all text hidden in the INDX-cluster part is analyzed. Words are added to the Vocabulary.
2. The Vocabulary is used for assigning correct code pages before the database export.

Consistency Check ➔ INDX Table Analysis ➔ Vocabulary ➔ INDX Table Repair ➔ Vocabulary ➔ Maintenance

Adjust Code Page Info in INDX cluster data
Final Steps in non-Unicode Systems

Newly created tables can be added to SPUMG by updating the Worklist. You can display the new tables in the Update Log.

Perform additional preparation steps as described in the documentation „Unicode Conversion Guide“.

One important additional preparation steps is Generating the Unicode nametab.

Note:
Afterwards you must not import new or modify existing DDIC objects! Otherwise the Unicode nametab will become invalid.
Demo

SPUMG
Agenda

1. Introduction
2. Conversion Preparation
   Demo
3. Database Export, Conversion and Import
4. The Unicode System
   Demo
5. Customer Experiences
6. Additional Information and Contacts
Database Export, Conversion and Import

Pre-Conversion
- Set up the Conversion Project
- Check Prerequisites
- Plan database downtime during conversion
- Enable Customer Developments

Unicode Conversion
- SAPinst (R3load)
- Database Export, Conversion & Import

Post-Conversion
- Unicode system is up and running
- Verification of Data Consistency
- Integration Testing focused on language handling

Conversion Preparation in non-Unicode system

Conversion of data from the old code system

Conversion of data to the Unicode system

Completion of the project
When you perform a Unicode Conversion you perform a System Copy!

Documentation:
“Homogeneous and Heterogeneous System Copy for SAP Systems based on”:

SAP Web AS 6.20
SAP Web AS 6.40
SAP Web AS ABAP 6.40 SR1

- Preparation tasks
- Export the database with SAPinst
- Install database on target system
- Import the database dump with SAPinst
- Post-Processing tasks

You can download the applicable System Copy Guide from SAP Service Marketplace → Quick Link / instguides.

For information about Java System Copy, see SAP Note 795267
The system setup tool SAPinst is used for the entire system copy - internally SAPinst uses the program R3load.

R3load performs the database export including conversion to Unicode by using the Export Control Table and the Vocabulary. R3load writes a log in case code page information is not available, and it performs the database import.

The Distribution Monitor is an additional tool for reducing the system downtime to an acceptable limit by taking advantage of additional hardware (e.g. application servers) of the customer’s system landscape (pilot project).
SAPinst Export (SAP R/3 Enterprise Ext. 2.00)

Intro

Welcome
Welcome to the System Landscape Implementation Manager (SAPinst). To start an installation, choose one of the Following:

- ABAP: SAP R/3 Enterprise 4.70 Extension Set 2.0 Server for MS SQL Server
- ABAP: SAP R/3 Enterprise 4.70 Extension Set 2.0 Dialog Instance for MS SQL Server
- [ADAP]: Export SAP System Server Database
  - Configure the MSOS Node A for MS SQL Server
  - Configure the MSOS Node B for MS SQL Server

Advanced Options

Advanced Export Options

- split STR files
- top 50 (default: split largest 50 tables in extra STR files)
- -package 1000 (default: split STR files in 100MB packages)
- -limit 300 (default: split table greater than 300MB)

General Parameters

General Export Parameters
Enter General Export Parameters.

- export packages in alphabetical order
- export packages in custom order

Data file code page: 4103
Number of parallel jobs: 9

© SAP AG 2005, SAP TechEd '05 / SPC202 / 51
R3load - Export

SAPinst is used as „Frontend-tool“
- Logs are written to the SAPinst Directory
- Export Files are written to the Export Directory

Conversion (Non-UC → UTF-16) during Export:

R3load.exe -datacodepage 4103 -e ....
Example: „für“ (0x66 FC 72 → 0x0066 00FC 0072)

Non-UC R3load

SAP System
non-Unicode

Export DIR:
SAPAPPL0.001
SAPAPPL1.001
...
...

Multiple processes possible
- Default: One process (STR file) per data class (see SE11 → Technical Settings)
- Advanced Export options: 50 largest tables get own process each

Space needed on Export Dir:
~10-30% of DB size

DB independent format (depending on platform big or little endian)
SAPinst is used as „Frontend-tool“
- Non-UC DB can be deleted (not recommended) or UC DB can be installed on different server
- Import Procedure nearly same as New Installation

UC R3load
R3load.exe -i ....

Export DIR:
SAPAPPL0.001
SAPAPPL1.001
...
.
.

Multiple processes possible.
Number of processes same as during export (default)
Database Export, Conversion and Import: Overview

SPUMG

R3load Log

Conversion

non Unicode → R3load → Data

EXPORT

IMPORT

Unicode → R3load

Unicode

Data

Control Table

Vocabulary

Re-process Log
Database Export, Conversion and Import: Overview

EXPORT

SPUMG

Re-process Log
Control Table
Vocabulary

R3load Log

non Unicode

R3load

IMPORT

Data

R3load

Unicode
## Agenda

1. **Introduction**
2. **Conversion Preparation**
   - Demo
3. **Database Export, Conversion and Import**
4. **The Unicode System**
   - Demo
5. **Customer Experiences**
6. **Additional Information and Contacts**
1. Install Unicode Kernel + other Unicode Executables
2. Start Unicode System
3. Logon to Unicode System
Database Export, Conversion and Import: Overview

- **SPUMG**
  - Re-process Log
  - Control Table
  - Vocabulary

- **R3load**
  - Export Log

- **Conversion**

- **Export**
  - non Unicode
  - R3load
  - Data

- **Import**
  - Unicode
  - R3load

- **Data**

© SAP AG 2005, SAP TechEd '05 / SPC202 / 58
Database Export, Conversion and Import: Overview

SPUMG

R3load Log

SUMG

Data

non Unicode

R3load

Unicode

EXPORT

IMPORT

Re-process Log
Control Table
Vocabulary

Re-process Log
Control Table
Vocabulary

© SAP AG 2005, SAP TechEd '05 / SPC202 / 59
Run report UMG_SCAN_STATISTICS. This report collects and analyzes SPUMG statistics, e.g. database size, number of Vocabulary entries, R3load runtime, Reprocess Log size.

For enhancement monitoring, SAP kindly requests customers to send an e-mail with the following information to internationalization@sap.com:

- UMG_STATS_<SID>.xml (UMG_SCAN_STATISTICS file)
- database size before and after the conversion
- export and import runtimes
Unicode System: Transaction SUMG

Pre-Conversion
- Set up the Conversion Project
- Check Prerequisites
- Plan database downtime during conversion
- Enable Customer Developments

Unicode Conversion

Post-Conversion
- Unicode system is up and running
- Verification of Data Consistency
- Integration Testing focused on language handling

SUMG Conversion Completion in the Unicode System

Conversion Preparation in non-Unicode system

SAPinst (R3load) Database Export, Conversion & Import

SPUMG Conversion Completion in the Unicode System
SUMG: Conversion Completion

In the Unicode system you might recognize data which have not been converted correctly. In order to ‘complete’ the conversion procedure, the data can be converted once again in transaction SUMG, using the correct language information (code page).

SUMG provides two Completion Types

**Automatical Completion:**
Using the R3load Log and the language information stored in the Reprocess Log

**Manual Repair:**
- Repair Hints
- Manual Language (CP) Assignment
Automatical Completion

1. Enter path and name of the file which contains the path(s) of the XML-file(s).

2. R3load Log
   - table name, key value, code page used for the conversion, error messages

   Reprocess Log
   - table name, key value, lang. which were manually assigned in SPUMG

   Conversion Completion
   - Status, Table Name, Repair Category, Status Text
Manual Repair: Repair Hints

1. Create Condition

2. Create Hint

3. Save Hint

4. Execute Hint

Once created you can reuse Hints in other systems.
1. Assign current lang.
2. Assign correct lang.
3. Save entry to database

Tables can be manually added. Then for each broken entry the correct language must be recognized and assigned.
Agenda

1. Introduction

2. Conversion Preparation
   Demo

3. Database Export, Conversion and Import

4. The Unicode System
   Demo

5. Customer Experiences

6. Additional Information and Contacts
Agenda

1. Introduction
2. Conversion Preparation
   - Demo
3. Database Export, Conversion and Import
4. The Unicode System
   - Demo
5. Customer Experiences
6. Additional Information and Contacts
Customer Experiences

- Customer Experiences: Hardware Requirements
- Conversion Project: Timeline
- Customer Experiences: Runtimes
- Customer Project Example
## Customer Experiences: Hardware Requirements

<table>
<thead>
<tr>
<th>Unicode Encoding Form</th>
<th>Database Size Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>UTF-8</td>
<td>up to +10%</td>
</tr>
<tr>
<td>UTF-16</td>
<td>up to +60...70%</td>
</tr>
</tbody>
</table>

The figures are based on parallel benchmarking of Unicode / non-Unicode test systems and on customer experiences.
Conversion Project: Timeline

How long does it take to convert a system?

The project timeline depends on a number of factors:
Dependencies

Single Code Page or MDMP Conversion?
- MDMP requires more preconversion tasks and postconversion handling in the Unicode system

Biggest Tables – Optimize parallelization of Export/Import processes (splitting of R3load packages)

Processing of Cluster Tables
- Sizes of cluster tables (compared to transparent tables)

Hardware
- Number and speed of CPUs
- Performance of disks
- Separate server available for Unicode System?

Time spent on optimization (problems not related to runtime)
Depending on hardware and parallelization effort, the R3load runtimes throughput is 100 – 200 GB/h.
Runtime measurements show that using additional hardware (CPU power) speeds up the overall runtime of the system copy process.
R3load processes are distributed among multiple application servers.
Distribution Process

PREPARATION PHASE

Generation of control information and Migration Monitor properties.

Preparation is performed only on one designated machine.

EXPORT AND IMPORT PHASE

Both are started on each machine. Export and Import can run in parallel.

Preparation Phase must be completed first!

SAP Note
Pilot Release
855772
Paul Hartmann AG

Conversion of MDMP system to Unicode:

3 Standard single code pages with 16 installed system languages
ISO-8859-1 (1100): DA, DE, EN, ES, FR, IT, NL, NO, PT, SV
ISO-8859-2 (1401): CS, HU, PL, SK, SL
ISO-8859-7 (1700): EL
# Customer Project Example: Schedule

<table>
<thead>
<tr>
<th>Date Range</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/23/2004</td>
<td>Kick-Off-Meeting</td>
</tr>
<tr>
<td>02/27 – 04/01/2004</td>
<td>SPUMG tests (afterwards RESET_ALL)</td>
</tr>
<tr>
<td>04/02 – 08/04/2004</td>
<td>Conversion of first sandbox system</td>
</tr>
<tr>
<td>08/10 – 10/14/2004</td>
<td>Conversion of second sandbox system</td>
</tr>
<tr>
<td>10/12 – 10/30/2004</td>
<td>Conversion of development system (DEV)</td>
</tr>
<tr>
<td>10/12 – 11/02/2004</td>
<td>Conversion of production system (PRD)</td>
</tr>
<tr>
<td>10/31/2004</td>
<td>Go-Live</td>
</tr>
<tr>
<td>11/02/2004</td>
<td>Systemcopy PRD → quality system (QAS)</td>
</tr>
</tbody>
</table>

The schedule also includes: application and integration tests, preconversion data reduction (archiving and deletion of obsolete data), runtime measurements and optimization
### Customer Project Example: PRD System

<table>
<thead>
<tr>
<th>SAP Release</th>
<th>R/3 Enterprise 4.7 Ext. Set 1.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>System type</td>
<td>Production System (PRD)</td>
</tr>
<tr>
<td>Platform (non-Unicode)</td>
<td>AIX 5.2</td>
</tr>
<tr>
<td>Platform (Unicode)</td>
<td>AIX 5.2</td>
</tr>
<tr>
<td>Database version</td>
<td>DB2 UDB V8 FP3</td>
</tr>
<tr>
<td>Hardware (non-Unicode)</td>
<td>5x PowerPC_POWER4 (24 CPUs, 1.4-1.5 GHz, 47 GB RAM)</td>
</tr>
<tr>
<td>Hardware (Unicode)</td>
<td>5x PowerPC_POWER4 (24 CPUs, 1.4-1.5 GHz, 127 GB RAM)</td>
</tr>
<tr>
<td>Database size (non-UC, in KB)</td>
<td>551.775.684</td>
</tr>
<tr>
<td>Database size (UC, in KB)</td>
<td>523.101.248</td>
</tr>
<tr>
<td>Export size (in KB)</td>
<td>44.467.397</td>
</tr>
<tr>
<td>PSAPTEMP size (in KB)</td>
<td>133.600.000</td>
</tr>
<tr>
<td>SAP Basis 6.20 Support Package</td>
<td>SAPKB62040</td>
</tr>
</tbody>
</table>
## Runtimes of automated conversion steps

<table>
<thead>
<tr>
<th>Step</th>
<th>Runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fill TRDIR Language Keys</td>
<td>7 min</td>
</tr>
<tr>
<td>Set missing Language Flags</td>
<td>47 min</td>
</tr>
<tr>
<td>Scan: Consistency Check</td>
<td>24,5 h</td>
</tr>
<tr>
<td>Check cluster tables</td>
<td>20,7 h</td>
</tr>
<tr>
<td>Scan: Tables without Language Information</td>
<td>144,8 h</td>
</tr>
<tr>
<td>Scan: Tables with Ambiguous Language Information</td>
<td>6,3 h</td>
</tr>
<tr>
<td>Scan: INDX Table Analysis</td>
<td>1,7 h</td>
</tr>
<tr>
<td>Scan: Tables with Language Information</td>
<td>-</td>
</tr>
<tr>
<td>Import of maintained Vocabulary</td>
<td>8 min</td>
</tr>
<tr>
<td>Scan: Reprocess</td>
<td>107,2 h</td>
</tr>
<tr>
<td>Scan: INDX Table Repair</td>
<td>11 min</td>
</tr>
<tr>
<td>Generate the Unicode-Nametab</td>
<td>1,5 h</td>
</tr>
<tr>
<td>SUMG: Automatical Completion</td>
<td>7,9 h</td>
</tr>
</tbody>
</table>

"*" = Parallelization possible by scheduling several Jobs.
<table>
<thead>
<tr>
<th>Action</th>
<th>Runtime (Sandbox/DEV)</th>
<th>Runtime (PRD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unicode-enabling of customer programs</td>
<td>200 days</td>
<td>-</td>
</tr>
<tr>
<td>Deletion of obsolete data</td>
<td>3 days</td>
<td>0,5 days</td>
</tr>
<tr>
<td>Vocabulary Maintenance</td>
<td>33 days*</td>
<td>6 h</td>
</tr>
<tr>
<td>Resolve collisions</td>
<td>10 days</td>
<td>-</td>
</tr>
<tr>
<td>Reprocess Log Maintenance</td>
<td>5 days</td>
<td>5,5 days</td>
</tr>
<tr>
<td>INDEX Log Maintenance</td>
<td>0,5 h</td>
<td>0,5 h</td>
</tr>
<tr>
<td>SUMG: Manual Repair</td>
<td>12 h</td>
<td>12 h</td>
</tr>
<tr>
<td>Application tests</td>
<td>...</td>
<td>-</td>
</tr>
</tbody>
</table>

*new Vocabulary Maintenance Tools (see SAP Notes) have not been used.
## Conversion of Production System: Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00 – 23:00</td>
<td><strong>Uptime</strong> Finish SPUMG Preparation (Repairlog), Export of proposal pool, stop of transports</td>
</tr>
<tr>
<td>21:30 – 23:00</td>
<td>Offline Backup of non-UC; generation of Unicode nametab</td>
</tr>
<tr>
<td>23:00 – 00:00</td>
<td>Offline Backup of non-UC; preparation tasks</td>
</tr>
<tr>
<td>00:00 – 01:00</td>
<td>Offline Backup of non-UC; preparation tasks</td>
</tr>
<tr>
<td>01:00 – 20:00</td>
<td>Export</td>
</tr>
<tr>
<td>20:00 – 22:00</td>
<td>Drop database, delete syslog files</td>
</tr>
<tr>
<td>22:00 – 22:30</td>
<td>Preparation of Import</td>
</tr>
<tr>
<td>22:30 – 01:00</td>
<td>Create tablespaces</td>
</tr>
<tr>
<td>01:00 – 10:15</td>
<td><strong>Downtime</strong> Import (incl. create all indexes and views)</td>
</tr>
<tr>
<td>10:15 – 14:45</td>
<td>Update statistics</td>
</tr>
<tr>
<td>14:45 – 15:30</td>
<td>First steps in UC</td>
</tr>
<tr>
<td>15:30 – 17:40</td>
<td>Offline Backup of UC</td>
</tr>
<tr>
<td>17:40 – 23:00</td>
<td>SUMG; RFC destinations; key-user tests</td>
</tr>
<tr>
<td>23:00</td>
<td><strong>Uptime</strong> Release for endusers</td>
</tr>
</tbody>
</table>
Agenda

1. Introduction
2. Conversion Preparation
   Demo
3. Database Export, Conversion and Import
4. The Unicode System
   Demo
5. Customer Experiences
6. Additional Information and Contacts
Additional Information and Contacts

- Printing in Unicode Systems
- Upgrade and Conversion Paths
- Platforms and Databases
- Documentation
- References and Contacts
Printing in Unicode Systems

Unicode UTF-8 Printing Solutions
High-speed printing of documents with single-byte and multi-byte native language characters.

1. HP Unicode Printer + HPUTF8 device type + SAP Andale_* Fonts
2. Lexmark Unicode Printer + LEXUTF8 device type + SAP Andale_* Fonts

- SAPscript forms
- Smart Forms
- ABAP Lists

Cascading Fonts Printing Solution
Printing of all Unicode characters on any printer with SAP!pd + cascading fonts-enabled SAPWIN device type. No modification of forms required!

- SAPscript forms
- Smart Forms
- ABAP Lists
Upgrade and Conversion 4.6C/D -> ERP 2005

**Release**
- 4.6C or 4.6D

**Action**
- Toolimport
- Upgrade Preparation
- Unicode Conversion Preparation
- Upgrade
- Unicode Conversion

**Timeline**
- Uptime
- Downtime
SAP supports Unicode systems on the following platforms:

### Platforms and Databases

<table>
<thead>
<tr>
<th>Database system</th>
<th>Operating System</th>
</tr>
</thead>
<tbody>
<tr>
<td>SQL Server</td>
<td>W2K¹</td>
</tr>
<tr>
<td>Oracle</td>
<td>x</td>
</tr>
<tr>
<td>DB/2</td>
<td>x</td>
</tr>
<tr>
<td>SAP DB</td>
<td>x</td>
</tr>
<tr>
<td>Informix²</td>
<td>-</td>
</tr>
</tbody>
</table>

Default = 64 bit versions

¹32 bit versions still available

²There will be no support for Informix and Tru64!
The following documents are required for the conversion of non-Unicode SAP systems to Unicode:

- **System Copy Guide**: available for Web AS 6.20/6.40/6.40SR1
- **SAP Note 548016**: application-specific information.
https://service.sap.com/unicode@sap

internationalization@sap.com
References and Contacts

https://service.sap.com/globalization

Multinational Issues

Enabling Business Globalization

mySAP.com e-business solutions are purpose-built for deployment anywhere in the world, as our unparalleled customer base in over 120 countries proves.

Country Versions, Languages

Our solutions follow global best business practices, with specific functions and features to account for varying national tax laws, accounting laws, and manufacturing regulations in over 40 countries: our country versions.

You can run all country versions simultaneously in the same system. And what’s more, each user can log on in the language that suits them best, with 30 languages to choose from. The benefits speak for themselves.

Service

Our global service network will help you implement and run your mySAP.com solution, wherever you are. With development centers, offices, and partners worldwide, we couldn’t be better placed to help.

globalization@sap.com
Questions?

Q & A
Further Information

Public Web:
Please complete your session evaluation.

Be courteous — deposit your trash, and do not take the handouts for the following session.

Thank You!
Cascading Fonts: Customizing

Customizing SAP Cascading Fonts

Global Cascading Fonts Data

Cascading SAP Font Name
- ANDALE_J
- CHINESE
- ONGMO
- ONGMO
- COURIER
- COURIER
- SYL
- SYL
- SYL
- SYL
- SYL
- SYL
- SYL
- SYL
- SYL

Device type dependent data of HELVE

Script | Sub SAP Font
--- | ---
ARABIC | ANDALE_J
BASIC LATIN | HELVE
CYRILLIC | HELVE
GREEK | HELVE
HEBREW | HELVE
LATIN EXTENDED-A | HELVE
LATIN EXTENDED-B | HELVE
LATIN 1 SUPPLEMENT | HELVE
OTHERS | HELVE
THAI | THANSAN

© SAP AG 2005, SAP TechEd '05 / SPC202 / 96
Both documents were printed out with font 'HELVE'.
Common Character Set

Example: MDMP with Latin-1 and Latin-2

ISO 8859-1 (Latin-1)

A | i | φ | £ | ¥ | £ | œ | « | ¬ | ¬ | Ø | – | – | – |
B | ± | ² | ³ | µ | ¶ | · | ¹ | ² | » | ¼ | ½ | ¾ | › |
C | À | ¿ | À | À | â | ã | å | æ | ç | È | Ê | Ë | Ì |
D | Ð | Ñ | Ò | Ô | Õ | Õ | Ô | × | Ò | Ù | Ù | Ù | Ù |
E | ä | á | â | ã | å | æ | ç | ë | ë | é | é | é | ë |
F | ë | ñ | ö | ò | ö | ò | ò | ò | ò | ò | ò | ò | ò |

Common Characters

ISO 8859-2 (Latin-2)

A | Å | Ö | Å | ö | Å | ä | Å | Å | Ł | Ś | Ś | Ś | ź |
B | Å | Ï | Å | Å | Å | æ | Ç | Ç | Ç | Ç | Ç | Ç | Ć |
C | Ć | Ė | Ė | Ė | Ė | Ý | Ý | Ý | Ý | Ý | Ý | Ý | Ý |
D | Ń | Ñ | Ń | Ń | Ń | Ń | Ń | Ń | Ń | Ń | Ń | Ń | Ń |
E | Ć | Ć | Ć | Ć | Ć | Ć | Ć | Ć | Ć | Ć | Ć | Ć | Ć |
F | đ | ň | ň | ň | ň | ň | ň | ň | ň | ň | ň | ň | ň |

© SAP AG 2005, SAP TechEd '05 / SPC202 / 98