



IBM Software Group

DB2 for z/OS V8 Unicode: preparing today for tomorrow

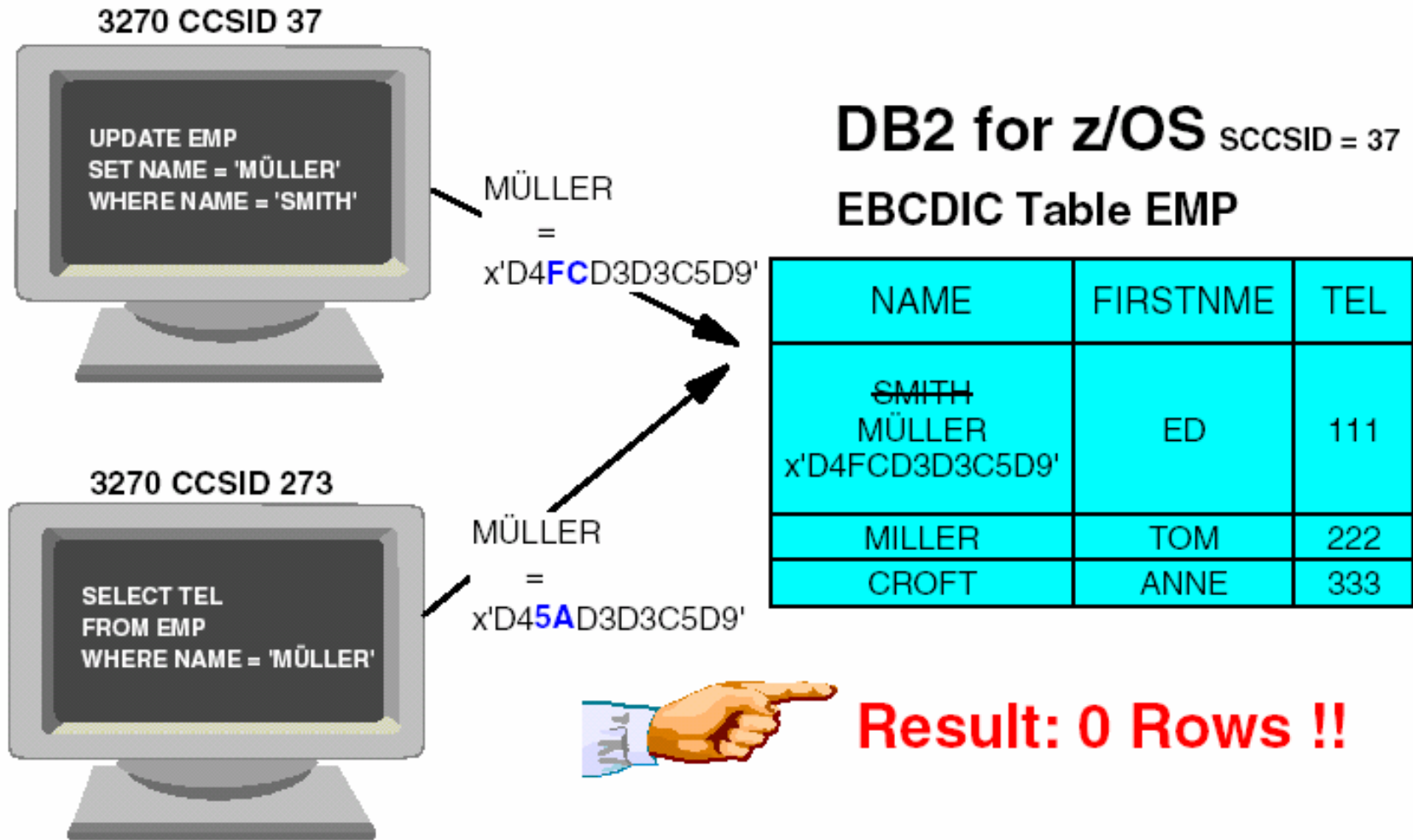
Cécile Benhamou
cecile_benhamou@fr.ibm.com

Based on a presentation by John J. Campbell and Florence Dubois

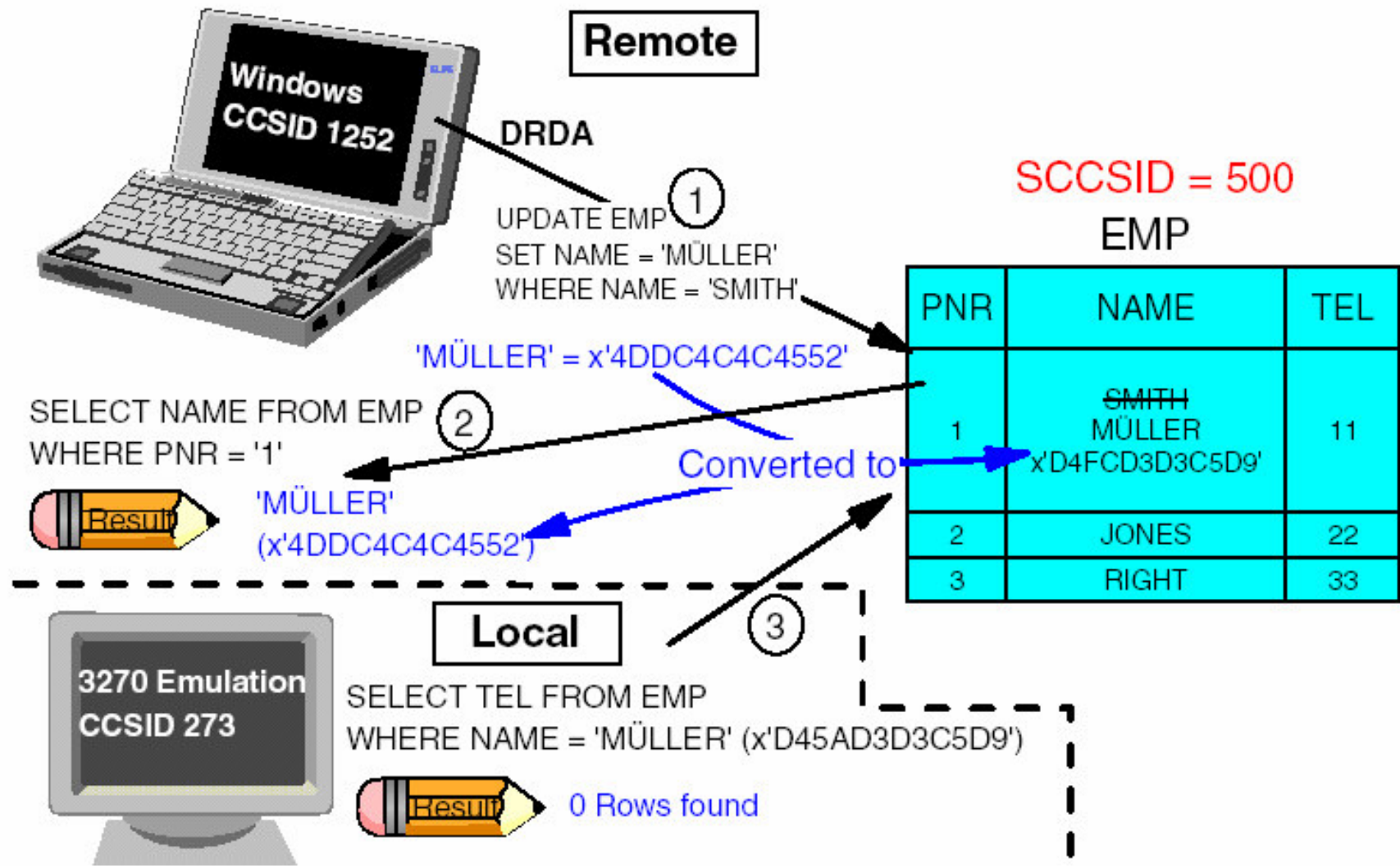
DB2 Information Management Software



Storing Characters by Local Applications



Remote (DRDA) Applications



Conversion methods

- Native DB2 - SYSIBM.SYSSTRINGS
- z/OS support for Unicode
 - Must configure Unicode Conversion Services for z/OS

```
COMMAND INPUT ==> /d uni,all
```

```
RESPONSE=SC63
```

```
CUN3000I 02.37.19 UNI DISPLAY 752
```

```
ENVIRONMENT: CREATED      04/17/2004 AT 10.08.19
```

```
MODIFIED      04/17/2004 AT 10.08.20
```

```
IMAGE CREATED 12/08/2003 AT 17.02.33
```

```
SERVICE: CHARACTER      CASE      NORMALIZATION  COLLATION
```

```
STORAGE: ACTIVE        424 PAGES
```

```
LIMIT          51200 PAGES
```

```
CASECONV: NORMAL
```

```
NORMALIZE: DISABLED
```

```
COLLATE: DISABLED
```

```
CONVERSION: 00850-01047-ER      01047-00850-ER
              00037-01200 (13488)-ER      01200 (13488)-00037-ER
              00037-01208-ER      01208-00037-ER
              00437-01208-ER      01208-00437-ER
              00037-00367-ER      01252-00037-ER
              00037-01252-ER      00367-00037-ER
              00500-01200 (13488)-ER      01200 (13488)-00500-ER
              01047-01200 (13488)-ER      01200 (13488)-01047-ER
              01047-01208-ER      01208-01047-ER
              01208-01200-ER      01200-01208-ER
              01383-01200-ER      01200-01383-ER
              00932-01200-ER      01200-00932-ER
```

```
.....
```

- Brochure 'Support for Unicode: Using Conversion Services' (SA22-7649)



Unicode Catalog ...

- If precompile application with NEWFUN(NO)
 - Disables use of new SQL function for static SQL
 - Go back through SQL Parser during BIND processing
 - ✓ To get Unicode UTF-8 format
 - ✓ Incur extra overhead
- After entry to NFM, default for NEWFUN changes to YES
 - DBRMs switch to Unicode UTF-8 format
 - Set NEWFUN to NO
 - ✓ Get DBRMs in EBCDIC format
- At BIND time
 - Anything from DBRM in Unicode UTF-8 format goes into SYSSTMT and SYSPACKSTMT in Unicode UTF-8 format
 - Anything from DBRM in EBCDIC format goes into SYSSTMT and SYSPACKSTMT in EBCDIC format



Some Examples if using Unicode in V8

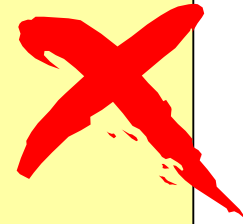
- Select rows from ET1, UT1 where C1 = 'ABC'
 - Application encoding scheme: EBCDIC

Coded SQL statement:

```
SELECT ET1.C1, UT1.C1
FROM ET1,UT1
WHERE ET1.C1 = X'C1C2C3'
AND UT1.C1 = X'414243'
```

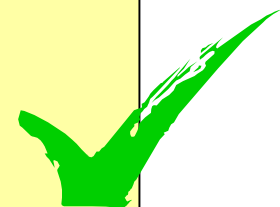
Interpreted as:

```
SELECT ET1.C1, UT1.C1
FROM ET1,UT1
WHERE ET1.C1 = 'ABC'
AND UT1.C1 = ' âä'
```



Correct encoding:

```
SELECT ET1.C1, UT1.C1
FROM ET1,UT1
WHERE ET1.C1 = X'C1C2C3'
AND UT1.C1 = X'C1C2C3'
```



Some Examples if using Unicode in V8



- Application encoding scheme: EBCDIC

Column - EBCDIC Derived value not based on a column - EBCDIC **Derived value based on a Unicode column**

```
SELECT FIRSTNME, SUBSTR(NAME, 1, 8),  
       SUBSTR('ABCD', 1, length('AB')), CURRENT DATE  
FROM UNICODETABLE, EBCDICTABLE  
WHERE NAME = :hv1  
AND FIRSTNME > 'JIM'  
AND DEPT = x'C1C2C3'
```

Host variable - EBCDIC

Special register - EBCDIC

String constant - EBCDIC



Some Examples if using Unicode in V8

■ UTILITIES

- Utility Control Statements can be specified in EBCDIC or UNICODE
 - ✓ DB2 detects which encoding scheme is being used
 - ✓ Objects names in messages will be in EBCDIC

■ EBCDIC

- //SYSIN DD *
COPY TABLESPACE A.B
/*

■ UNICODE

- //SYSIN DD *
"à|&\$"èà<àë&ää""â""""""""
/*



Conclusion

- Don't be afraid

Cyrillic

Ж	О	Й	К
046A	047A	048A	049A
Ж	О	Й	К
046B	047B	048B	049B

Mongolian

ᠠ	ᠨ
1874	1884
ᠡ	ᠯ
1875	1885
ᠢ	ᠮ
1876	1886

CJK

路	磊	綾
F937	F947	F957
露	賂	菱
F938	F948	F958
魯	雷	陵
F939	F949	F959

- But be prepared

Basic
Latin

1	A	Q
0031	0041	0051
2	B	R
0032	0042	0052
3	C	S
0033	0043	0053
4	D	T
0034	0044	0054



- Don't wait until V8 migration: Start now

