

# JOIN – And All Its Flavors

The Standard for  
Enterprise  
Business  
Intelligence

Cesare Petrizio

Information Builders

FIRE 2008 User Conference

# JOINS – Inner and Outer

## Agenda

- JOIN Rules
- Many-to-Many (or one-to-Many) Relationship
  - Inner JOIN
  - Outer JOIN
- Many-to-One ( or One-to-One) Relationship
  - Always OUTER JOIN
- Conditional JOINS

# JOINS – Inner and Outer Rules

- All Rules are determined by the SUFFIX of the TO file
  - JOIN TO FOCUS
    - Only single target field may be specified
    - Target field must be indexed
    - Many-to-Many supported
  - JOIN TO sqltable
    - Multiple target fields may be specified
    - Indexes are not required, but preferred
    - Many-to-Many supported

# JOINS – Inner and Outer Rules

- JOIN TO Indexed Files
  - Target field/group must be primary key or alternate index
  - Multiple target fields may be specified
    - High-order elements of key or alternate index
  - Many-to-Many supported
- JOIN TO FIX
  - Multiple target fields may be specified
  - Many-to-Many not supported
  - Both files must be sorted in ascending order on the JOIN keys

# JOINS – Inner and Outer

## Syntax

JOIN

```
field1 [ AND field2 ...] [WITH fieldname] [TAG tagname]  
IN file1 TO [ALL] fielda [AND fieldb...] IN file2 [TAG tagname]  
AS joinname
```

END

```
JOIN field1 IN file1 TO [ALL] field2 IN file2 AS joinname
```

# JOINS – Inner and Outer

## Where ...

- Field1 AND field2 ...
  - Up to Four fields may be specified
- WITH fieldname
  - DEFINE-based JOIN
  - DEFINE of field specified after the JOIN
  - Fieldname specified becomes the anchor of the JOIN
- TAG tagname
  - Tagname becomes a prefix for fully qualifying fields in specified file
- joinname (default is blank)
  - Identifies JOIN for the session
  - Another JOIN with the same name will overlay
  - Specified JOIN can be CLEARed

## **Relationships – One-to-Many**

The Standard for  
Enterprise  
Business  
Intelligence

# JOINS – Inner and Outer

## EMPDATA

FILENAME=EMPDATA, SUFFIX=FOC

SEGNAME=EMPDATA, SEGTYPE=S1

FIELDNAME=PIN,	ALIAS=ID,	FORMAT=A9,	INDEX=I,\$
FIELDNAME=LASTNAME,	ALIAS=LN,	FORMAT=A15,	\$
FIELDNAME=FIRSTNAME,	ALIAS=FN,	FORMAT=A10,	\$
FIELDNAME=MIDINITIAL,	ALIAS=MI,	FORMAT=A1,	\$
FIELDNAME=DIV,	ALIAS=CDIV,	FORMAT=A4,	\$
FIELDNAME=DEPT,	ALIAS=CDEPT,	FORMAT=A20,	\$
FIELDNAME=JOBCLASS,	ALIAS=CJCLAS,	FORMAT=A8,	\$
FIELDNAME=TITLE,	ALIAS=CFUNC,	FORMAT=A20,	\$
FIELDNAME=SALARY,	ALIAS=CSAL,	FORMAT=D12.2M,	\$
FIELDNAME=HIREDATE,	ALIAS=HDAT,	FORMAT=YMD,	\$

# JOINS – Inner and Outer

## EMPDATA

PIN	LASTNAME	FIRSTNAME
---	-----	-----
000000010	VALINO	DANIEL
000000020	BELLA	MICHAEL
000000030	CASSANOVA	LOIS
000000040	ADAMS	RUTH
000000050	ADDAMS	PETER
000000060	PATEL	DORINA
000000070	SANCHEZ	EVELYN
000000080	SO	PAMELA
000000090	PULASKI	MARIANNE
000000100	ANDERSON	TIM

# JOINS – Inner and Outer Kids

FILENAME=KIDS , SUFFIX=FOC

SEGNAME=CHILDSEG, SEGTYPE=S1

FIELDNAME=EMP\_ID, ALIAS=PIN, FORMAT=A9, INDEX =I , \$

FIELDNAME=LASTNAME, ALIAS=SLN, FORMAT=A15 , \$

FIELDNAME=CHILDNAME, ALIAS=SFN, FORMAT=A10 , \$

FIELDNAME=DATE\_OF\_BIRTH, ALIAS=DOB, FORMAT=MDYY , \$

# JOINS – Inner and Outer

## Kids

EMP_ID	LASTNAME	CHILDNAME	DATE_OF_BIRTH
-----	-----	-----	-----
000000010	VALINO	ANTHONY	12/31/1980
000000010	VALINO	ANNE	11/09/1979
000000010	VALINO	ARTHUR	06/01/1982
000000010	VALINO	ASTRIC	05/03/1991
000000030	CASSANOVA	JOHN	05/07/1993
000000040	ADAMS	MARY	08/01/2000
000000060	PATEL	SAM	07/05/1998
000000070	SANCHEZ	SAMANTHA	08/04/1997

# JOINS – Inner and Outer Spice

FILENAME=SPICE , SUFFIX=FOC

SEGNAME=SPOUSEI , SEGTYPE=S1

FIELDNAME=PIN, ALIAS=ID, FORMAT=A9, INDEX=I,\$

FIELDNAME=LASTNAME, ALIAS=SLN, FORMAT=A15,\$

FIELDNAME=SPOUSENAME, ALIAS=SFN, FORMAT=A10,\$

FIELDNAME=SPOUSESSN , ALIAS=SSN, FORMAT=A9 , \$

# JOINS – Inner and Outer Spice

PIN	LASTNAME	SPOUSENAME	SPOUSESSN
---	-----	-----	-----
000000010	VALINO	ABIGAIL	000000011
000000030	CASSANOVA	EDWARD	000000032
000000040	ADAMS	BRIAN	000000043
000000060	PATEL	KEITH	000000064
000000070	SANCHEZ	EDWARD	000000075
000000090	PULASKI	DAVID	000000096

# JOINS – Inner and Outer Relationship

**JOIN PIN IN EMPDATA TO ALL PIN IN KIDS AS JOIN1**

**OUTER**

EMPDATA

**PIN  
LASTNAME  
FIRSTNAME  
MIDINITIAL**



**INNER**

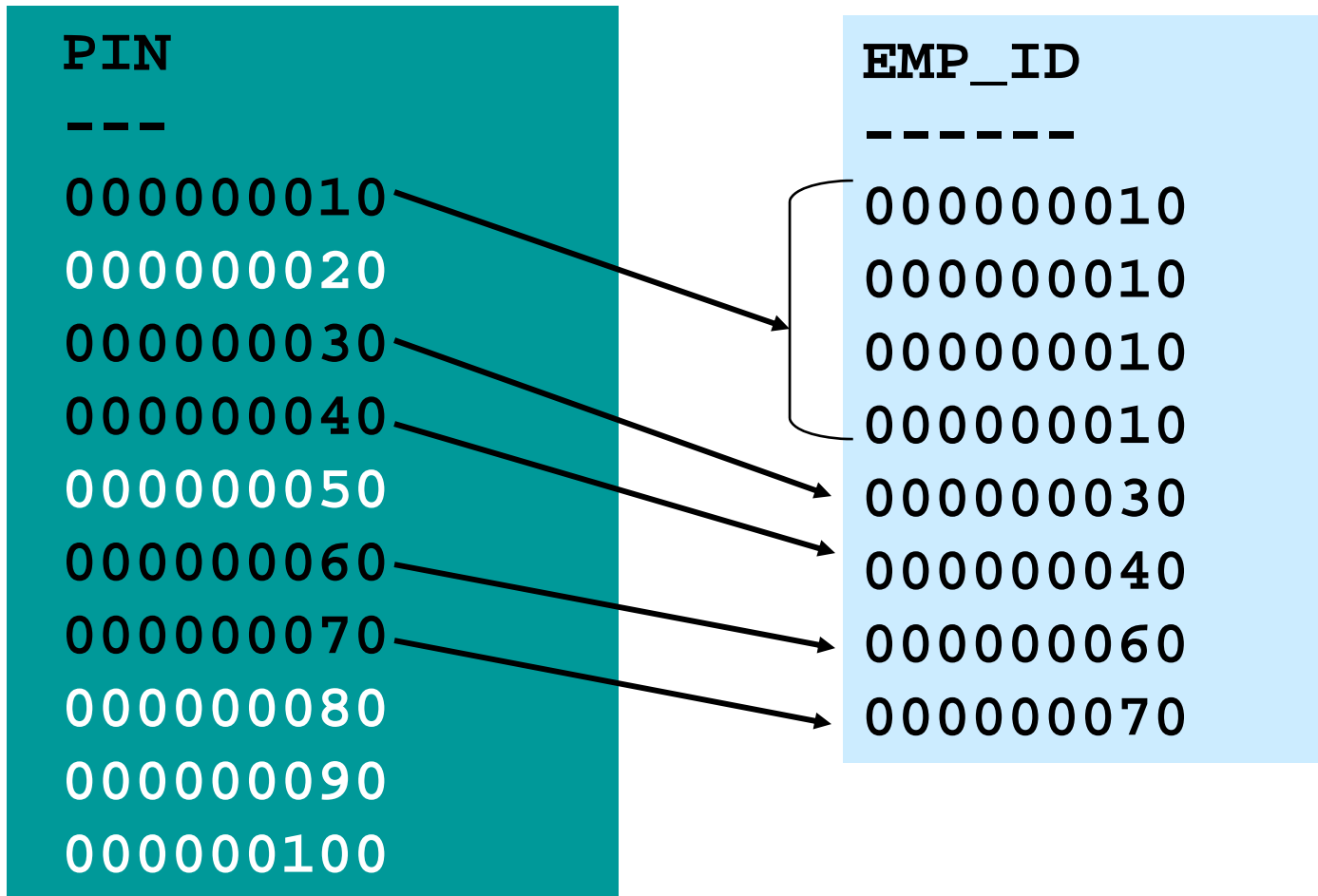
KIDS

**EMP\_ID  
LASTNAME  
CHILDNAME  
MIDINITIAL**

# JOINS – Inner and Outer

## Inner JOIN

**SET ALL = OFF**



# JOINS – Inner and Outer

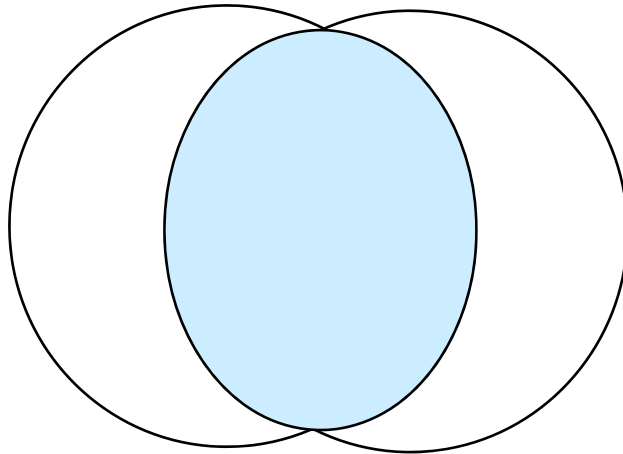
## Inner JOIN

PIN	LASTNAME	FIRSTNAME	CHILDNAME
---	-----	-----	-----
000000010	VALINO	DANIEL	ASTRIC ARTHUR ANNE ANTHONY
000000030	CASSANOVA	LOIS	JOHN
000000040	ADAMS	RUTH	MARY
000000060	PATEL	DORINA	SAM
000000070	SANCHEZ	EVELYN	SAMANTHA

# JOINS – Inner and Outer

## Outer JOIN

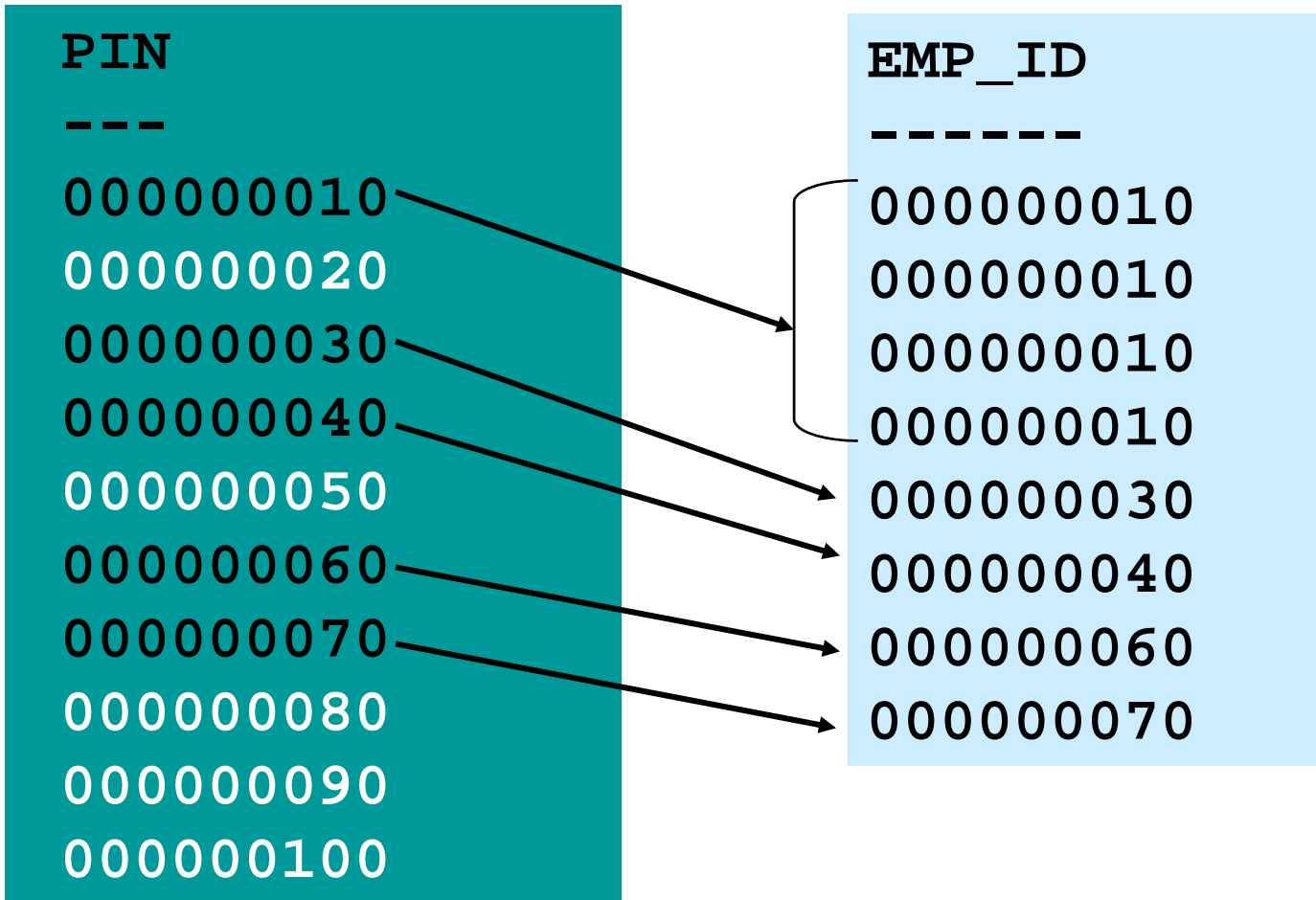
```
JOIN INNER hfld1 [AND hfld2 ...]  
      IN hostfile [TAG tag1]  
TO MULTIPLE crfield  
      [AND crfld2 ...] IN crfile [TAG tag2] [AS joinname]  
END
```



# JOINS – Inner and Outer

## Outer JOIN

### SET ALL = ON



# JOINS – Inner and Outer

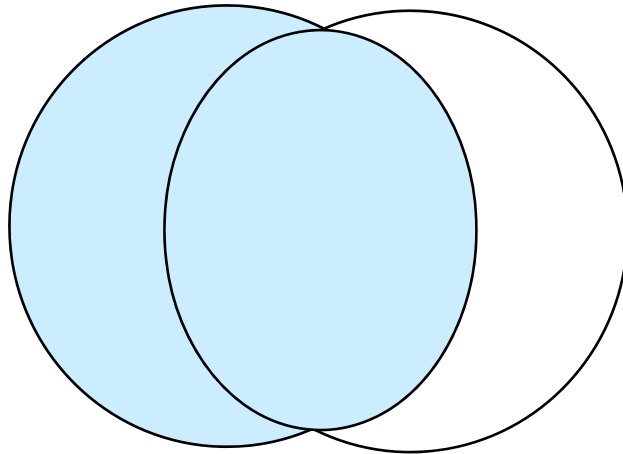
## Outer JOIN

PIN	LASTNAME	FIRSTNAME	CHILDNAME
---	-----	-----	-----
000000010	VALINO	DANIEL	ASTRIC ARTHUR ANNE ANTHONY
000000020	BELLA	MICHAEL	.
000000030	CASSANOVA	LOIS	JOHN
000000040	ADAMS	RUTH	MARY
000000050	ADDAMS	PETER	.
000000060	PATEL	DORINA	SAM
000000070	SANCHEZ	EVELYN	SAMANTHA
000000080	SO	PAMELA	.
000000090	PULASKI	MARIANNE	.
000000100	ANDERSON	TIM	.

# JOINS – Inner and Outer

## Outer JOIN

```
JOIN LEFT_OUTER hfld1 [AND hfld2 ...]  
      IN hostfile [TAG tag1]  
TO MULTIPLE crfield  
      [AND crfld2 ...] IN crfile [TAG tag2] [AS joinname]  
END
```



## **Relationships – One-to-One**

The Standard for  
Enterprise  
Business  
Intelligence



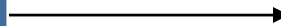
# JOINS – Inner and Outer Relationship

JOIN PIN IN EMPDATA TO PIN IN SPICE AS JOIN1

**OUTER**

EMPDATA

**PIN  
LASTNAME  
FIRSTNAME  
MIDINITIAL**



**INNER**

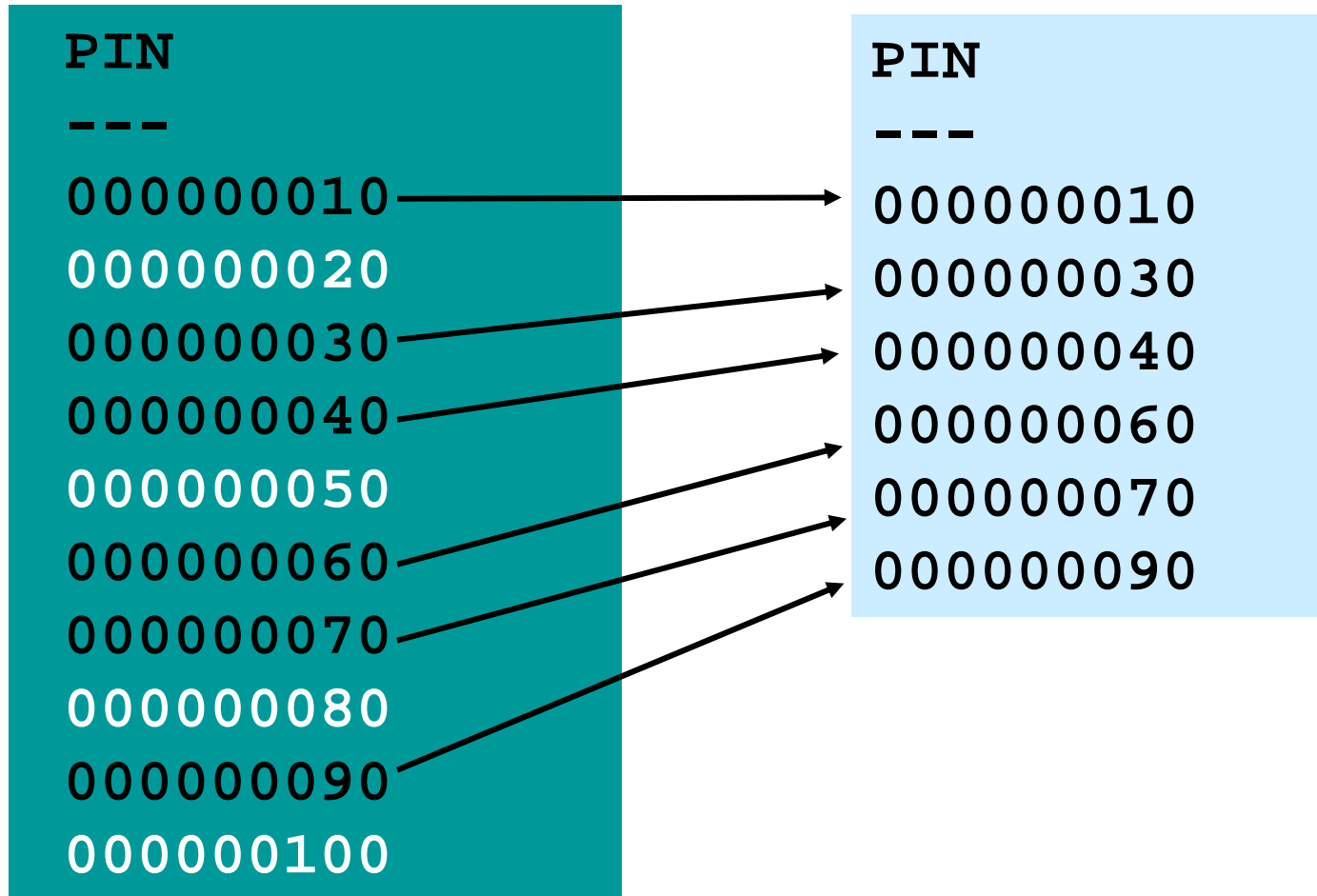
SPICE

**PIN  
LASTNAME  
SPOUSENAME  
SSN**

# JOINS – Inner and Outer

## Unique Outer JOIN

**SET ALL = OFF or SET ALL = ON**



# JOINS – Inner and Outer

## Unique Outer JOIN

PIN	LASTNAME	SPOUSENAME
---	-----	-----
000000010	VALINO	ABIGAIL
000000020	BELLA	
000000030	CASSANOVA	EDWARD
000000040	ADAMS	BRIAN
000000050	ADDAMS	
000000060	PATEL	KEITH
000000070	SANCHEZ	EDWARD
000000080	SO	
000000090	PULASKI	DAVID
000000100	ANDERSON	

# JOINS – Inner and Outer Unique Relationship

JOIN PIN IN EMPDATA TO ID IN KIDS AS JOINU

**OUTER**

EMPDATA

**PIN  
LASTNAME  
FIRSTNAME  
MIDINITIAL**



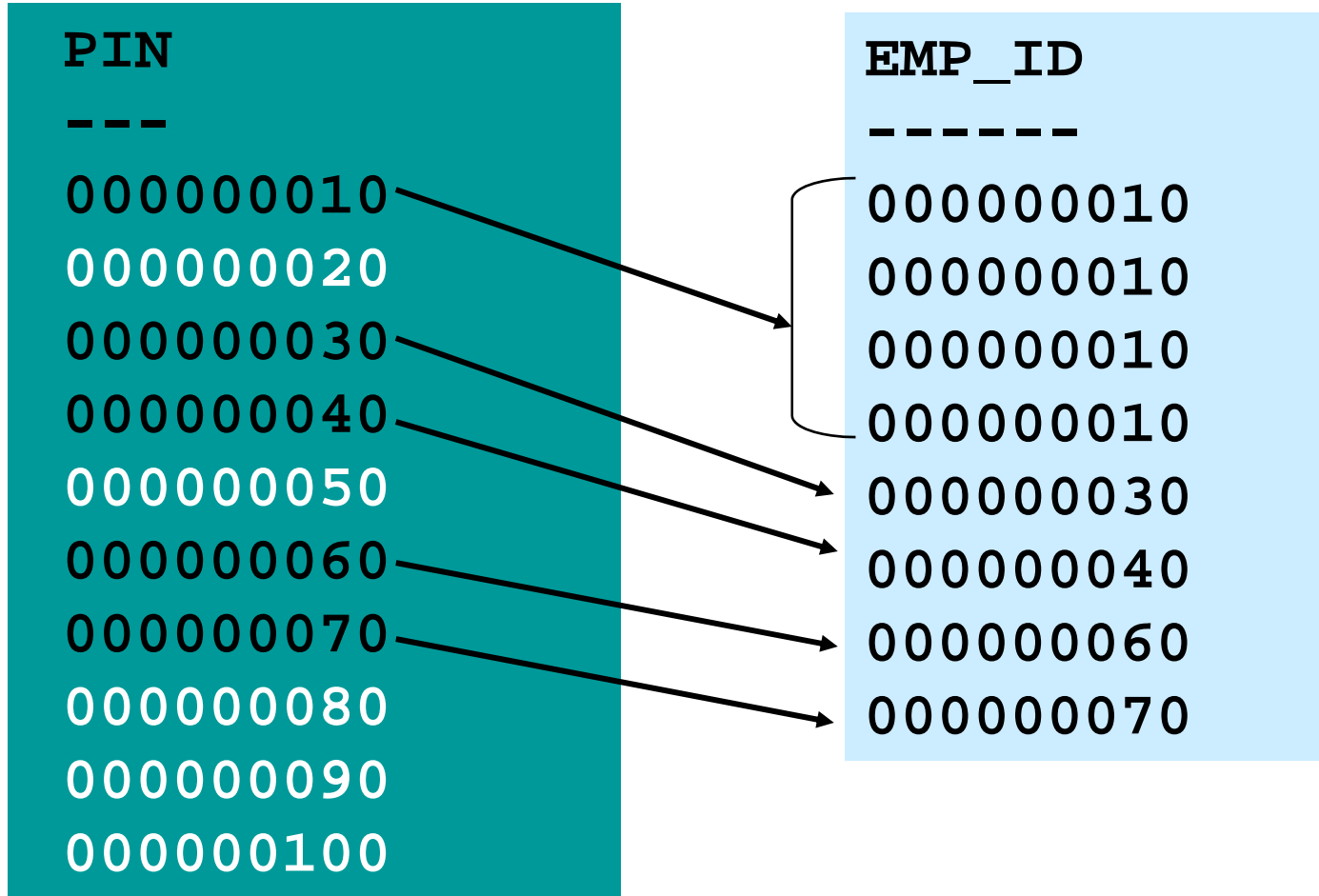
**INNER**

KIDS

**PIN  
LASTNAME  
CHILDNAME  
MIDINITIAL**

# JOINS – Inner and Outer

## Unique JOIN



# JOINS – Inner and Outer

## Unique JOIN

PIN	LASTNAME	FIRSTNAME
---	-----	-----
000000010	VALINO	ARTHUR
000000020	BELLA	
000000030	CASSANOVA	JOHN
000000040	ADAMS	MARY
000000050	ADDAMS	
000000060	PATEL	SAM
000000070	SANCHEZ	SAMANTHA
000000080	SO	
000000090	PULASKI	
000000100	ANDERSON	

## JOINS – Inner and Outer Relationship

**JOIN PIN IN EMPDATA TO PIN IN KIDS AS JOIN1**

TABLE FILE EMPDATA

PRINT CHILDSEG.CHILDNAME NOPRINT

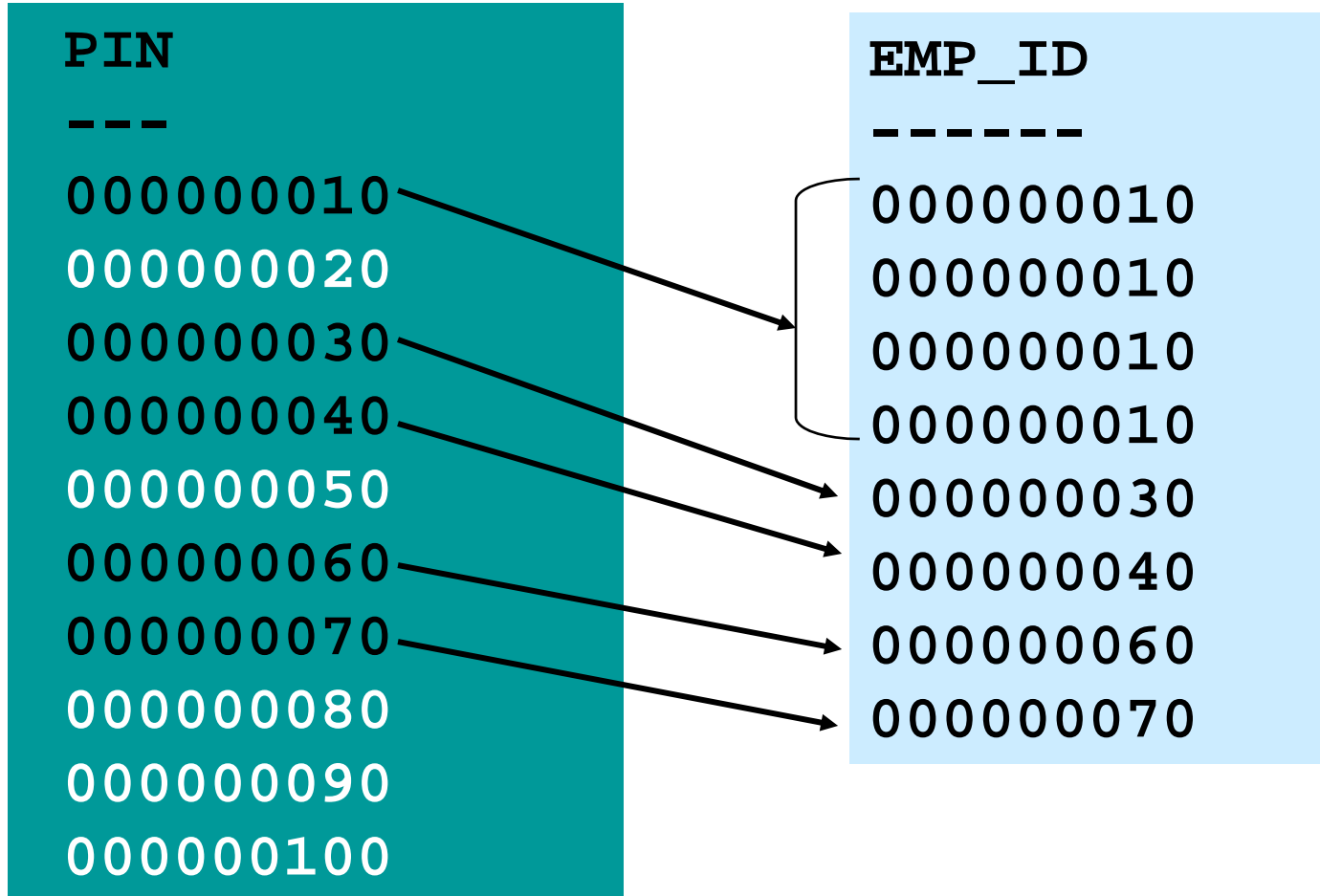
BY PIN BY LASTNAME BY CHILDNAME

WHERE PIN NE EMP\_ID

END

# JOINS – Inner and Outer

## Unique JOIN



# JOINS – Inner and Outer

## Inner JOIN

<b>PIN</b>	<b>LASTNAME</b>	<b>FIRSTNAME</b>
---	-----	-----
000000020	BELLA	MICHAEL
000000050	ADDAMS	PETER
000000080	SO	PAMELA
000000090	PULASKI	MARIANNE
000000100	ANDERSON	TIM

# JOINS – Inner and Outer

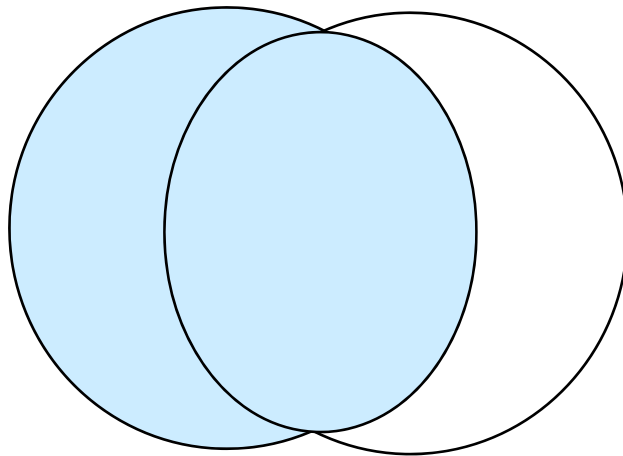
## Outer JOIN

```
JOIN LEFT_OUTER hfld1 [AND hfld2 ...]  
      IN hostfile [TAG tag1]
```

```
TO UNIQUE crfield
```

```
      [AND crfld2 ...] IN crfile [TAG tag2] [AS joinname]
```

```
END
```



# Conditional JOINS Syntax and Examples

The Standard for  
Enterprise  
Business  
Intelligence



# Conditional JOINS

## Syntax

```
JOIN [LEFT_OUTER|INNER] FILE from_file
      AT from_field [TAG tag1]
TO {ALL|ONE} FILE to_file
      AT to_field [TAG tag2]
[AS as_name]
      [WHERE expression1 ;
      WHERE expression2 ; ... ; ]
END
```

# Conditional JOINS

## Insurance Rates

TABLE FILE RATES PRINT \*

AGE	EAGE	RATE_PER_THOUSAND
---	----	-----
20	26	\$8
27	35	\$9
36	42	\$11
43	48	\$14
49	53	\$24
54	60	\$30
61	65	\$36
66	999	\$42

# Conditional JOINS

## Insurance Rates

```
JOIN FILE EMPDATA1 AT BIRTHDATE
      TO ALL FILE RATES AT AGE AS J1
      WHERE EMPDATA1.BAGE GE RATES.AGE;
      WHERE EMPDATA1.BAGE LE RATES.EAGE;

END

TABLE FILE EMPDATA1
HEADING
"TO: <FIRSTNAME <LASTNAME "
"</1 Thank you for choosing our company for your <0X insurance
needs."
"Thank you for choosing our company for your insurance needs. "
"Since your birthdate is <BIRTHDTATE ,your current rate is <0X"
  per <RATE_PER_THOUSAND"
"unit of coverage. This is your rate through age <EAGE . "
ON TABLE SET PAGE OFF
BY PIN NOPRINT PAGE-BREAK
END
```

# Conditional JOINS

## Insurance Rates and Letters

To: DANIEL VALINO

Thank you for choosing our company for your insurance needs.  
Since your birthdate is 07/20/1959 , your current rate is \$11 per  
unit of coverage. This is your rate through age 42 .

To: MICHAEL BELLA

Thank you for choosing our company for your insurance needs.  
Since your birthdate is 07/27/1952 , your current rate is \$24 per  
unit of coverage. This is your rate through age 53 .

## Conditional JOINS

### Insurance Rates – Another Approach

```
JOIN FILE EMPDATA1 AT BIRTHDATE  
TO ONE FILE RATES AT AGE AS J1  
  WHERE EMPDATA1.BAGE GE RATES.AGE;  
END  
  
-RUN  
-STEP2  
  
TABLE FILE EMPDATA1  
PRINT PIN BIRTHD BAGE  
  RATE BY AGE AS 'MINIMUM AGE'  
END
```

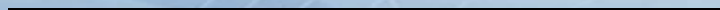
# Conditional JOINS

## Insurance Rates – Another Approach

MINIMUM AGE	PIN	BIRTHDATE	BAGE	RATE_PER_THOUSAND
20	000001020	07/24/1975	26	\$8
	000001030	04/17/1978	23	\$8
	000001040	05/07/1977	24	\$8
	000001050	03/20/1978	24	\$8
	000001060	03/06/1979	23	\$8
	000001070	03/10/1981	21	\$8
	000001090	04/08/1979	22	\$8
	000001120	04/08/1979	22	\$8
	000001140	04/24/1978	23	\$8
27	000001010	07/17/1972	29	\$9
	00000108W	02/19/1974	28	\$9
	000001100	05/21/1973	28	\$9
	000001110	05/16/1974	27	\$9
36	000000010	07/20/1959	42	\$11
	000000040	05/08/1960	41	\$11

## **Conditional JOINS – Clearing JOINS**

The Standard for  
Enterprise  
Business  
Intelligence



# Conditional JOINS

## Lots of JOINS

```
-* 1 FOR EMPLOYEE
JOIN FILE EMPLOYEE AT CURR_SAL TO ALL
      FILE CAR      AT RETAIL_COST AS CARALL
END

-* 2 FOR EMPLOYEE
JOIN CJC IN EMPLOYEE TO JOBCODE IN JOBFIL AS BJ
-* 3 FOR EMPLOYEE
JOIN FILE EMPLOYEE AT CURR_SAL TO ALL
      FILE CAR      AT RETAIL_COST AS CAREMP
WHERE EMPLOYEE.CURR_SAL GT (5 * CAR.RETAIL_COST);
END

-* 4 FOR CAR
JOIN FILE CAR AT RETAIL_COST TO ALL FILE EMPLOYEE AT
CURR_SAL AS EMPCAR
WHERE EMPLOYEE.CURR_SAL GT (5 * CAR.RETAIL_COST);
END
```

# Conditional JOINS

## Lots of JOINS

```
-* 5 FOR EMPLOYEE
  JOIN FILE EMPLOYEE AT LAST_NAME TO ONE
      FILE RETIRED AT FOCLIST AS EMPRET
WHERE RETIRED.NAME CONTAINS EMPLOYEE.LAST_NAME ;
END
-* 6 FOR CAR
JOIN COUNTRY IN CAR TO COUNTRY IN WORLD AS AJ
END
-RUN
? JOIN
```

# Conditional JOINS

## Current JOINS

### JOINS CURRENTLY ACTIVE

HOST			CROSSREFERENCE					
FIELD	FILE	TAG	FIELD	FILE	TAG	AS	ALL	WH
-----	-----	----	-----	-----	----	--	---	---
CURR_SAL	EMPLOYEE		RETAIL_COST	CAR		CARALL	Y	Y
RETAIL_COST	CAR		CURR_SAL	EMPLOYEE		EMPCAR	Y	Y
COUNTRY	CAR		COUNTRY	WORLD		AJ	N	N
CJC	EMPLOYEE		JOBCODE	JOBFILE		BJ	N	N
CURR_SAL	EMPLOYEE		RETAIL_COST	CAR		CAREMP	Y	Y
LAST_NAME	EMPLOYEE		FOCLIST	RETIRED		EMPRET	N	Y
>								

# Conditional JOINS

## Insurance Rates – Another Approach

### JOIN CLEAR CAREMP

#### JOINS CURRENTLY ACTIVE

HOST			CROSSREFERENCE					
FIELD	FILE	TAG	FIELD	FILE	TAG	AS	ALL	WH
-----	-----	---	-----	-----	---	--	---	--
CURR_SAL	EMPLOYEE		RETAIL_COST	CAR		CARALL	Y	Y
RETAIL_COST	CAR		CURR_SAL	EMPLOYEE		EMPCAR	Y	Y
COUNTRY	CAR		COUNTRY	WORLD		AJ	N	N
CJC	EMPLOYEE		JOBCODE	JOBFILE		BJ	N	N
CURR_SAL	EMPLOYEE		RETAIL_COST	CAR		CAREMP	Y	Y
LAST_NAME	EMPLOYEE		FOCLIST	RETIRED		EMPRET	N	Y

# Conditional JOINS

## Clearing JOINS

### JOIN CLEAR CARALL

#### JOINS CURRENTLY ACTIVE

HOST			CROSSREFERENCE					
FIELD	FILE	TAG	FIELD	FILE	TAG	AS	ALL	WH
-----	----	---	-----	----	---	--	---	--
CURR_SAL	EMPLOYEE		RETAIL_COST	CAR		CARALL	Y	Y
RETAIL_COST	CAR		CURR_SAL	EMPLOYEE		EMPCAR	Y	Y
COUNTRY	CAR		COUNTRY	WORLD		AJ	N	N
CJC	EMPLOYEE		JOBCODE	JOBFILE		BJ	N	N
CURR_SAL	EMPLOYEE		RETAIL_COST	CAR		CAREMP	Y	Y
LAST_NAME	EMPLOYEE		FOCLIST	RETIRED		EMPRET	N	Y

## **Conditional JOINS – Caveats**

The Standard for  
Enterprise  
Business  
Intelligence



# Conditional JOINS

## Rules and Caveats

- The conditional JOIN is supported for
  - FOCUS
  - VSAM
  - ADABAS
  - IMS
  - All relational data sources
- Optimization of the conditional JOIN syntax differs
  - Specific data sources involved in the join
  - Complexity of the WHERE criteria

# Conditional JOINS

## Rules and Caveats

- Where possible, use EQ-JOIN
  - Index/Key always used
  - No TABLE Scan
- Conditional JOIN
  - JOIN large file to small
  - Pages may remain in memory
- EQ-JOIN
  - JOIN small file to LARGE
  - Reduced I/O for non-matches.

Thank You