Referential Integrity (RI)

Von: Arkan Abdel
Leonard Faix

Agande

- What is Referential Integrity (RI) in a Database?
- What means RI in a Data Warehouse?
- Should one have RI in a DWH or not?
- pro and cons of Referential Integrity (RI)

Referential Integrity (RI) in a Database?

- the relational data in database tables has to be universally configurable
- keys that reference elements of other tables need to be connected to those other fields
- not separately
- prevents errors

What means RI in a Data Warehouse

- Referential Integrity in the data warehouse is a form of data integrity
- Relational databases break the storage of data down into elements
- data would get dropped (If it is not implemented properly)

Integrity Constraints :

- Impose restrictions on allowable data,beyond those
- Imposed by structure and types

Referential integrity

. **Integrity of references**

. **No dangling pointers**

Student

sID	sName	GPA	HS
123	Mary	----	-----

Apply

sID	cName	Major	dec
123	Stanford	Cs	Y

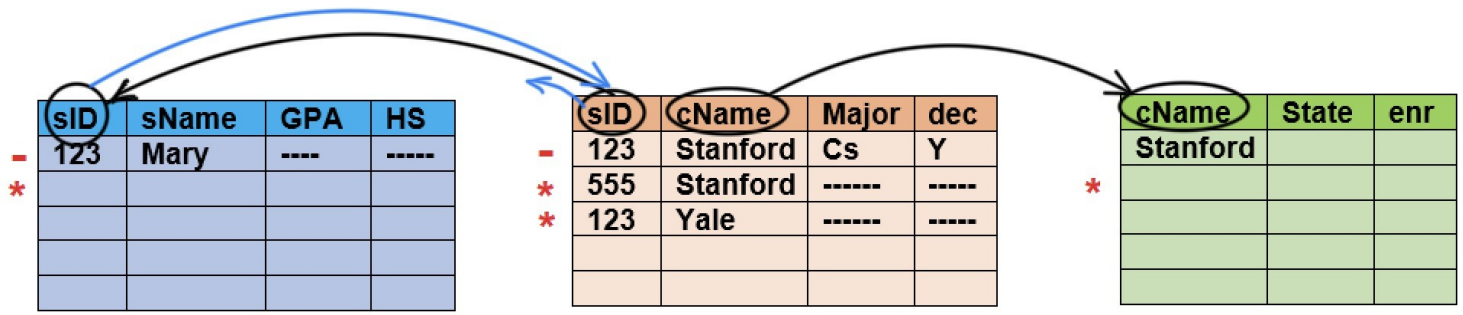
College

cName	State	enr
Stanford		

	sID	sName	GPA	HS
-	123	Mary	----	-----
*				

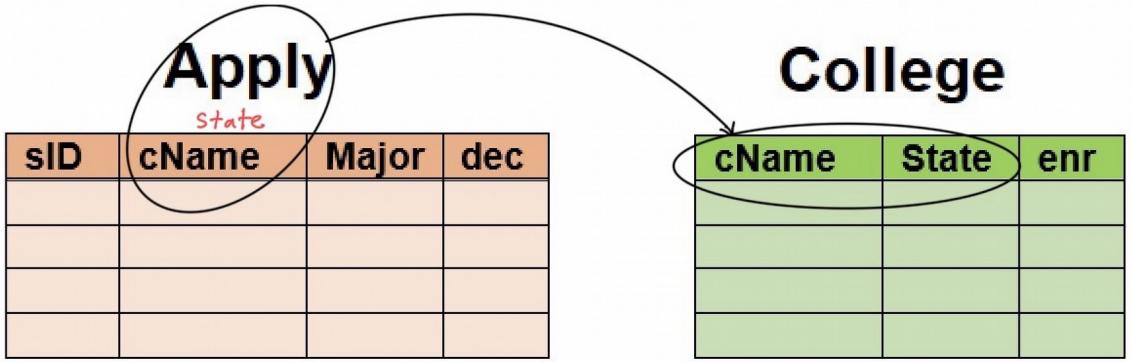
	sID	cName	Major	dec
-	123	Stanford	Cs	Y
*	555	Stanford	------	-----
*	123	Yale	------	-----

	cName	State	enr
	Stanford		
*			



Referential integrity from R.A to S.B
Each value in column A of table R must appear in column B of table S

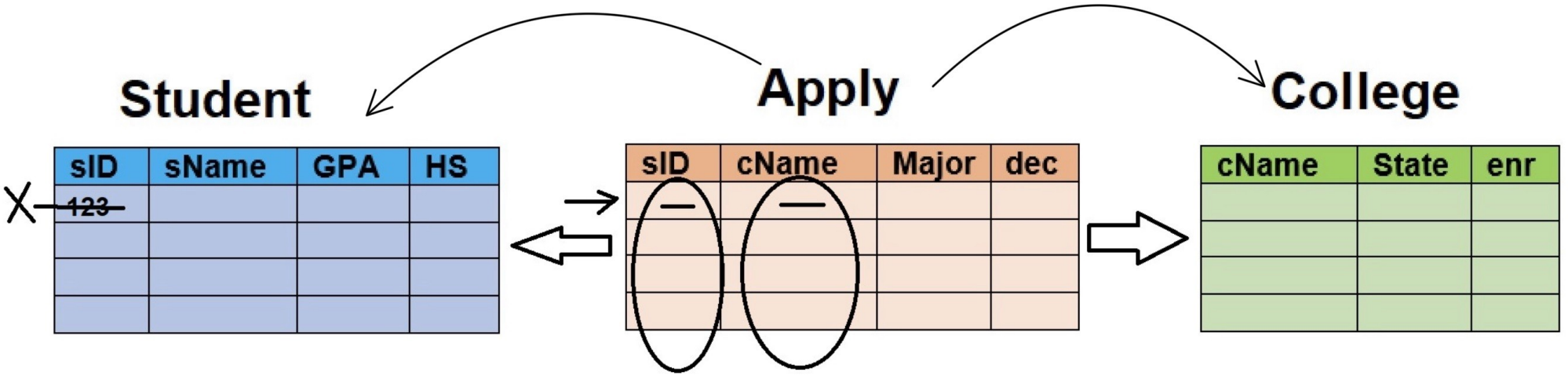
- A is called the „foreign Key“ (foreign key constraints)
- B is usually required to be the primary key for table S or at least unique
- Multi-attribute foreign keys are allowed



Referential Integrity Enforcement (R.A to S.B)

- Potentially violating modifications :
- Insert into R
- Delete from S
- Update R.A
- Update S.B

If violation → error



Referential Integrity Enforcement (R.A to S.B)

Special actions:

- Delete from S → Error

  Restrict (default), set Null, Cascade

- Update S.B

1

Student

sID	sName	GPA	HS

Apply

sID	cName	Major	dec
123	-----	------	

College

cName	State	enr

2

Student

sID	sName	GPA	HS
123	-----	-----	

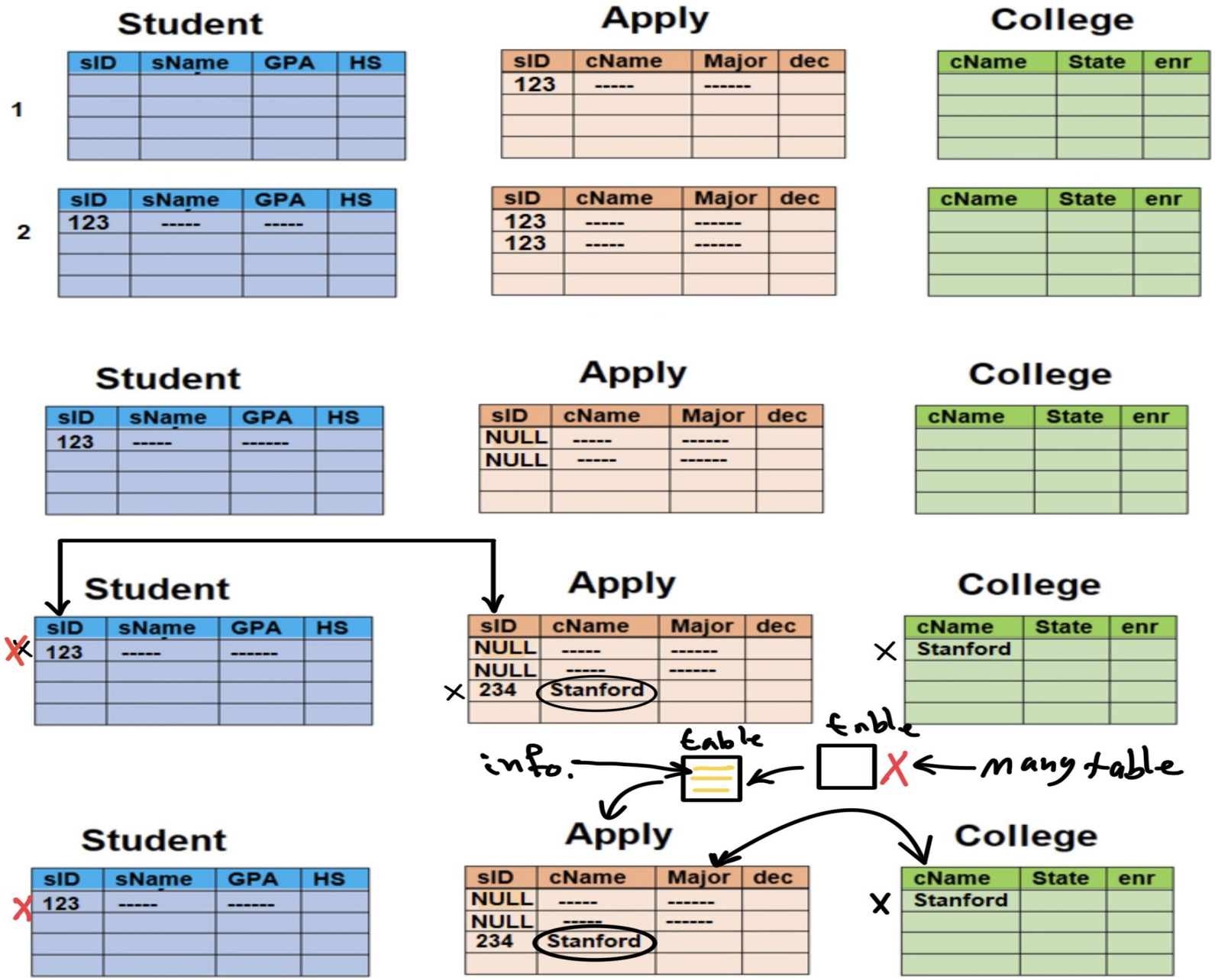
Apply

sID	cName	Major	dec
123	-----	------	
123	-----	------	

College

cName	State	enr

Referential Integrity Enforcement (R.A to S.B)



Referential Integrity Enforcement (R.A to S.B)

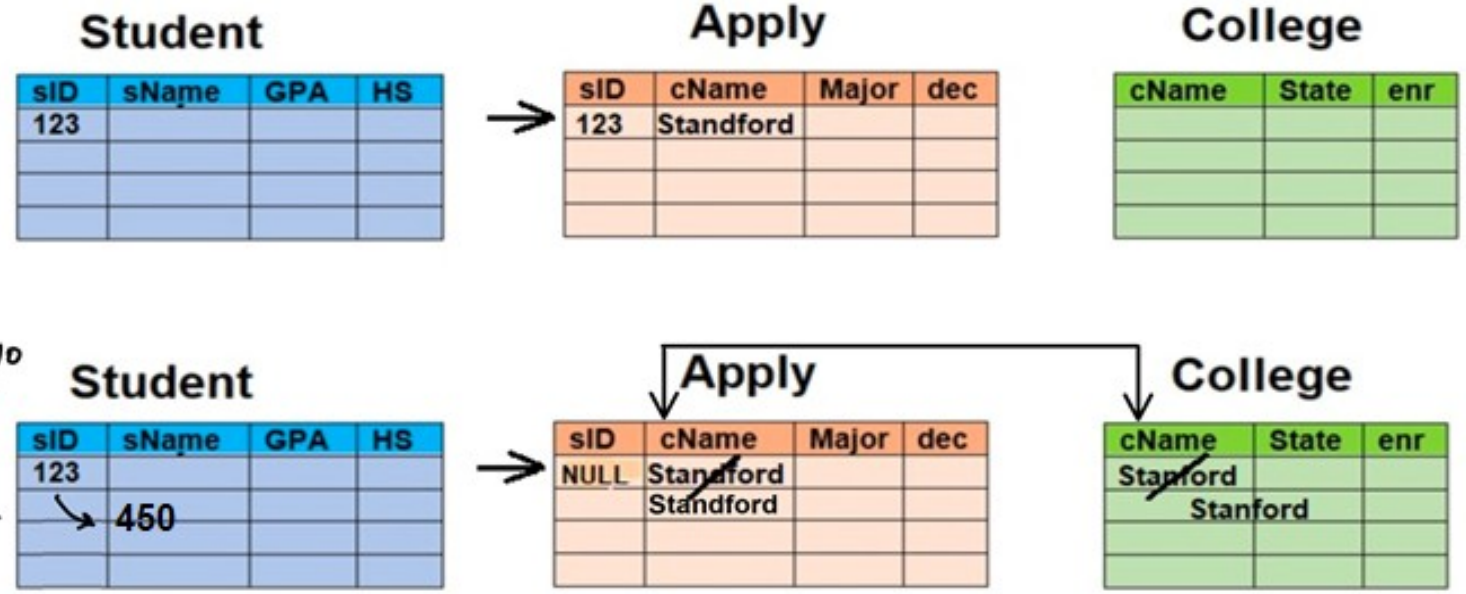
Special actions:

- Delete from S → Error

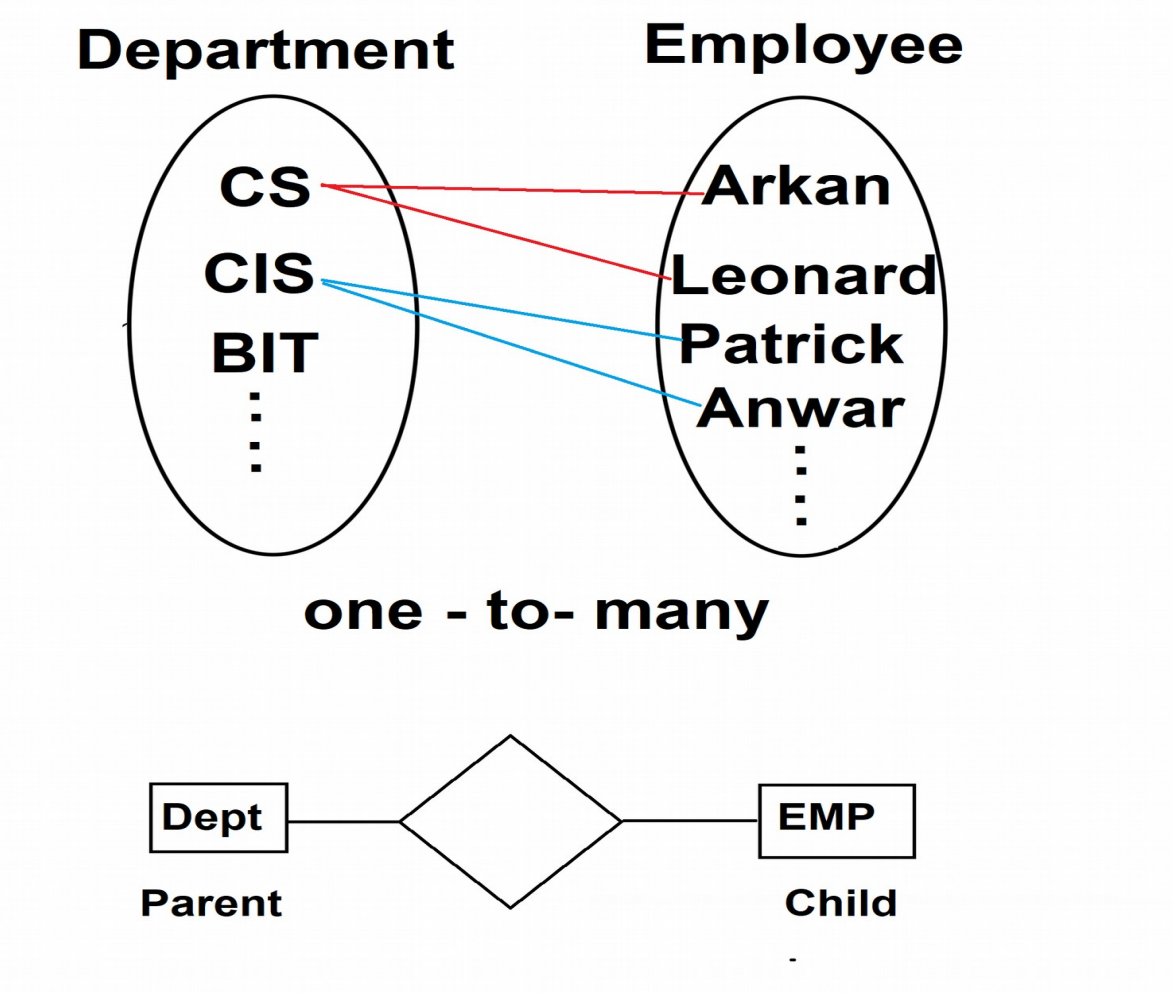
  Restrict (default), set Null, Cascade

- Update S.B

  Restrict (default), set Null, Cascade



Beispiel :


Department
Employee
CS
CIS
BIT
Arkan
Leonard
Patrick
Anwar
one - to- many
Dept
EMP
Parent
Child

EMP

EID	First	Last	Salary	Dno
10	Arkan	Abdel	42000	1
5	Leonard	Faix	60000	2
4	Patrick	Foucks	65000	1
7	Anwar	Adial	7000	2
8	jon	mark	55000	3

DEP.

DID	Dname	Location
1	CS	floor 1
2	CIS	floor 2
3	BIT	floor 3

Die Problemlösung :

Create table Dept (DID char(1) Primary Key, Dname varchar(20) , Location varchar(20) ,

on delete Cascade

on update Cascade);

+ We can create our own Data Type :

Create type **TypeName as varchar(30)** ;

create table Emp (ID……

Last **TypeName**…..

Why should I enforce RI?

- Ensuring that relationships between rows of data exist and are used as they are defined.
- User can trust data and rely on relationships

Referential Integrity in Data Warehouse

- Referential integrity is a decision, not a standard practice. It depends on the data
- ETL can ensure RI -> need strong control over ETL
- Constraints can ensure RI
  - Foreign Key...

RI by Constraints

- Constraints can increase load time and write time
- Constraints can make read queries faster
- Updates are done in the database environment, not in the warehouse envionment
- Many tables + many references -> too much development overhead

Conclusion, Considerations

- How will Referential Integrity impact the performance?
- RI can save dev and support time
- RI can cost more time, maintaining constraints
- Is the DW a read-only copy of transactional databases?
  ->Maintaining RI probably isn't worth it.
- Can the ETL maintain integrity?

Thank you for your attention

Sources

- https://datawarehouseinfo.com/implementing-referential-integrity-in-a-data-warehouse-a-controversial-decision-with-a-lasting-impact/
- https://esj.com/articles/1998/08/13/referential-integrity-for-the-data-warehouse-environment.aspx?m=1
- https://en.wikipedia.org/wiki/Referential_integrity
- sql - When is referential integrity not appropriate? - Stack Overflow
- Implement Referential Integrity Constraints for Consistency & Error Control (datawarehouseinfo.com)
- Referential integrity and its role in data warehousing | Auckland, Wellington, Christchurch, NZ (theta.co.nz)
- Referential integrity and its role in data warehousing: part two | Auckland, Wellington, Christchurch, NZ (theta.co.nz)