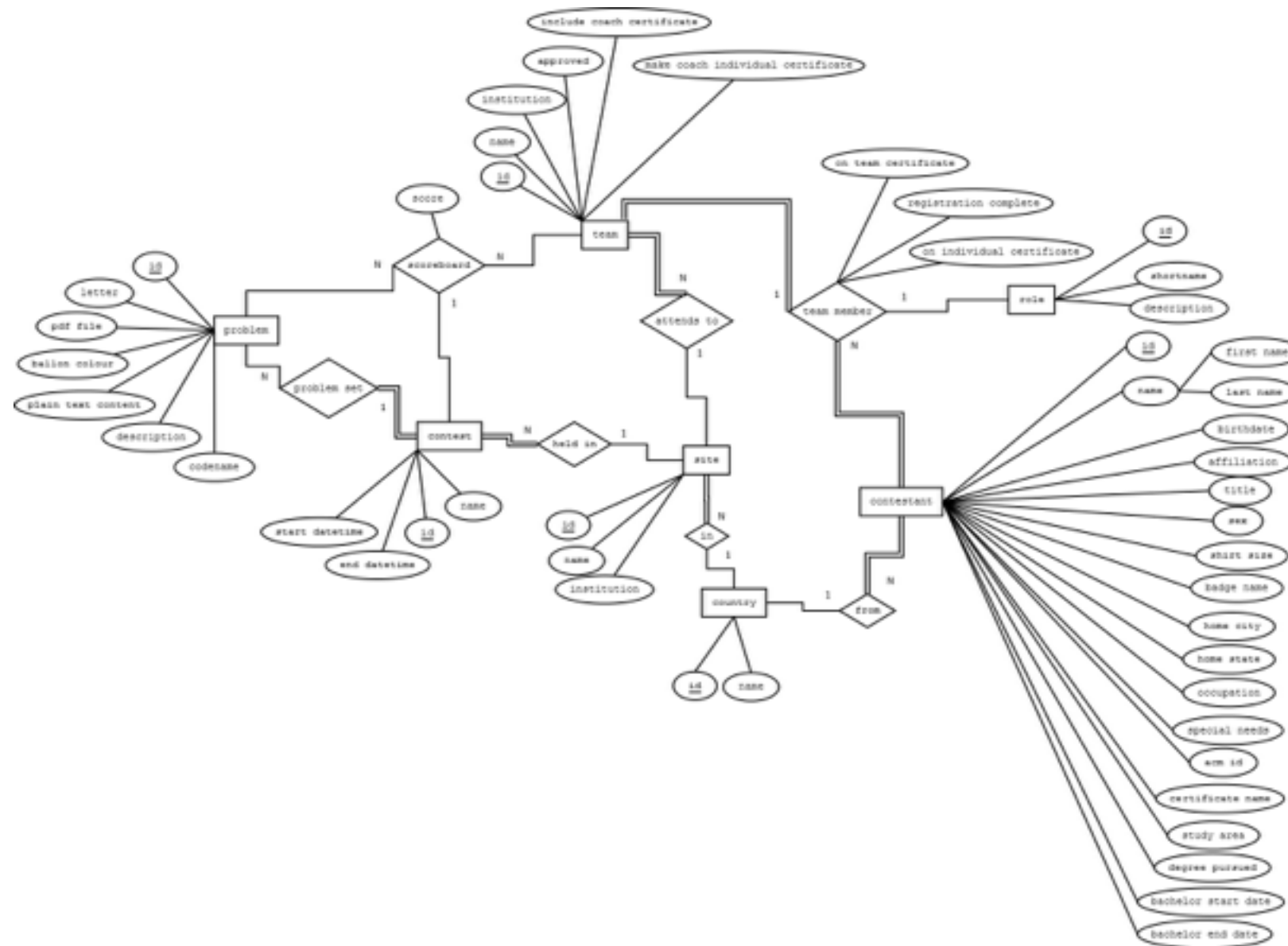


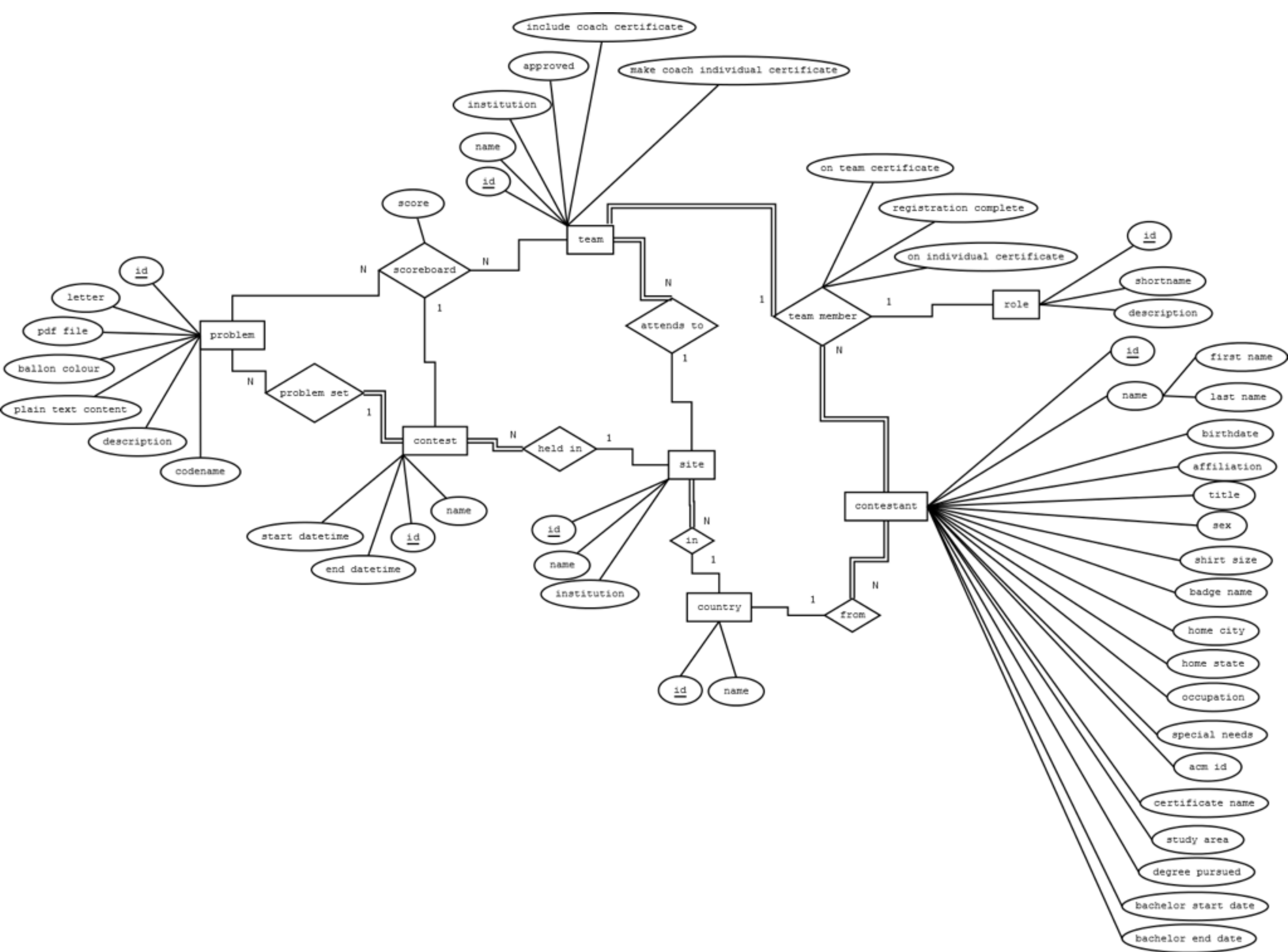
Gestión de Bases de Datos

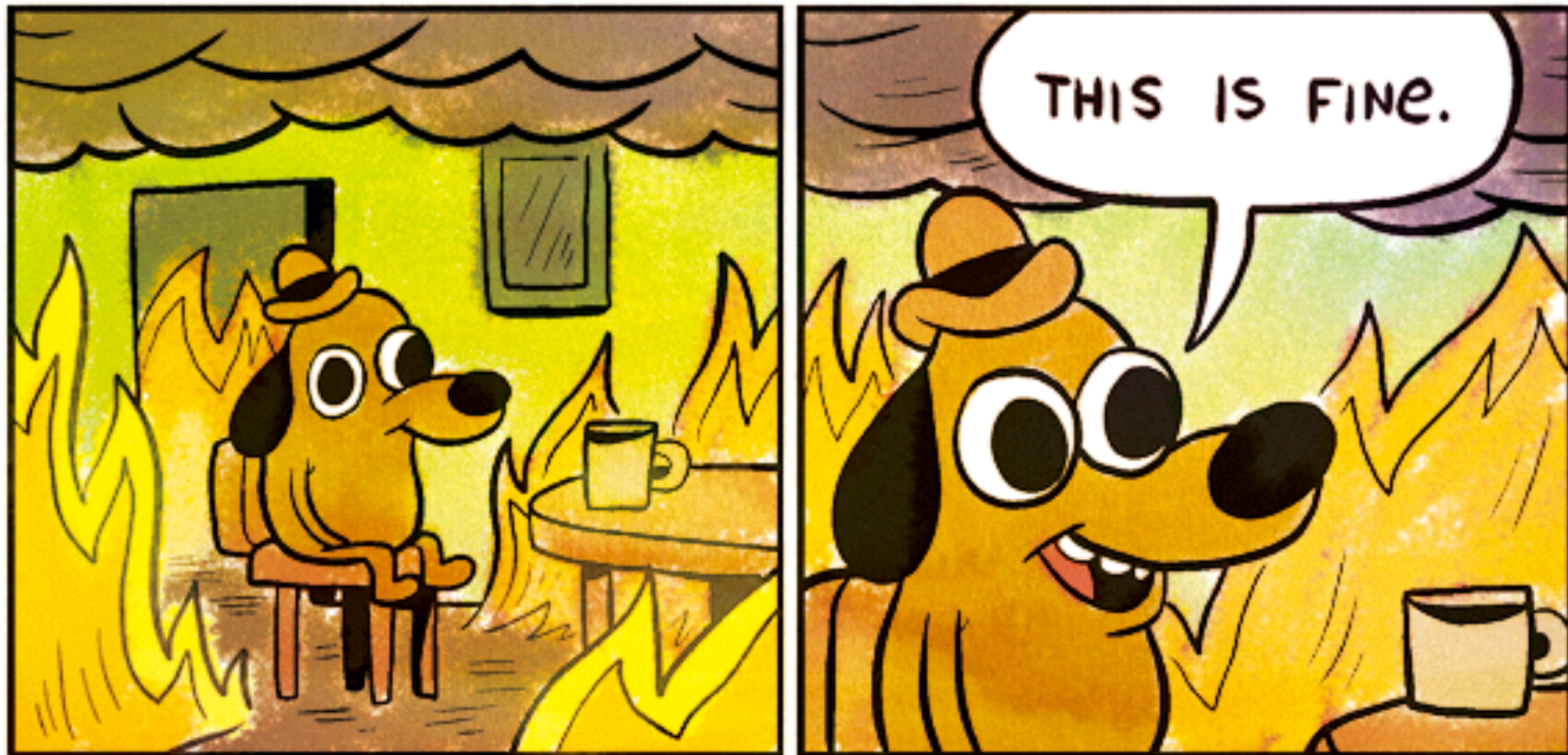
Erik Regla, Fabian Olivares



Etapas 1

Resumen





Etapas 2

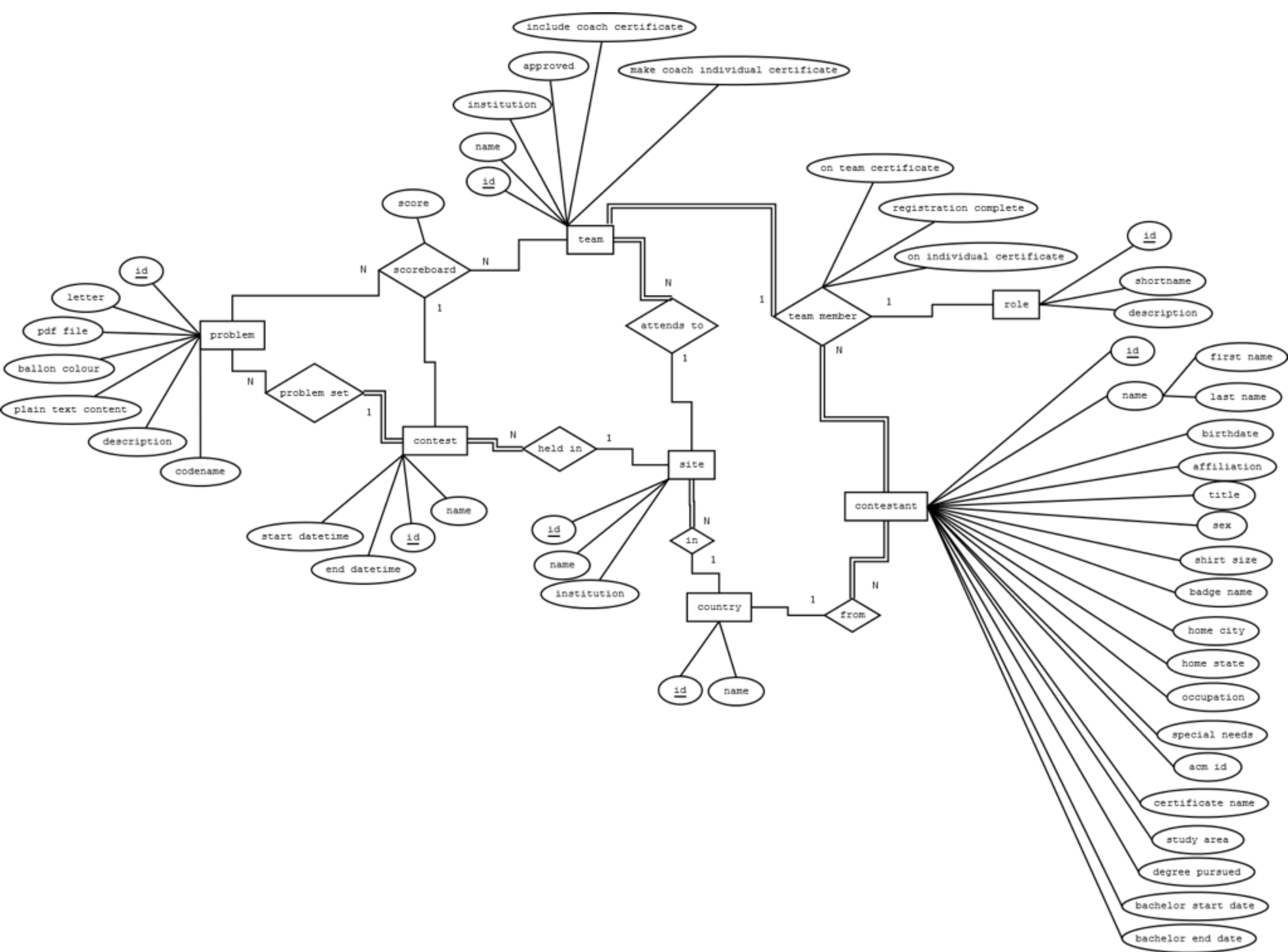
Aplicando índices

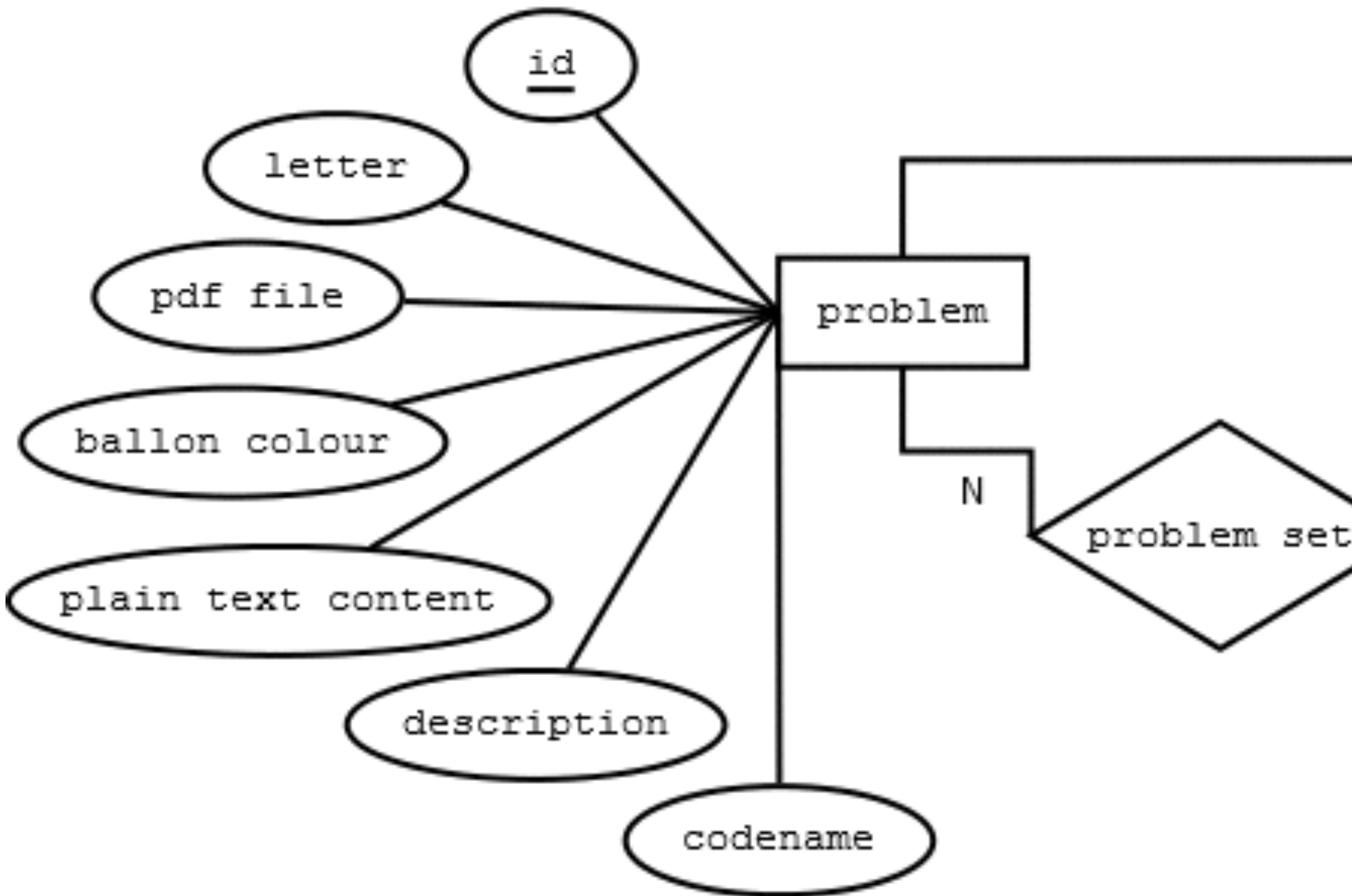
Tabla 1: Tiempos de ejecución para cada script en milisegundos

script	non_indexed	btree	hash	both
01_table_generation.sql	9	11	4	9
02_roles.sql	69760	69299	69586	69760
03_problems.sql	1964690	1958507	1982939	1964690
04_countries.sql	65882	65776	65772	65882
05_sites.sql	672127	668997	671179	672127
06_contests.sql	918502	934399	917848	918502
07_contestants.sql	2021650	2069131	2077798	2021650
08_problem_sets.sql	11099786	11066399	11095732	11099786
09_contest_sites.sql	650900	650706	651507	650900
10_teams.sql	1846802	1837683	1877973	1846802
11_scoreboards.sql	3015905	3004977	3009461	3015905
12_team_members.sql	3350787	3336618	3403136	3350787
13_queries.sql	3350787	1455	498	1558

No se puede apreciar una diferencia notable entre los tiempos de carga para bases de datos con diferentes índices, siendo la diferencia más notoria con la tablas indexadas utilizando hash (ver Figura 1.1.1). Sin embargo, el tiempo total de ejecución para las instancias es claramente diferente (ver `13_queries.sql` en Tabla 1) .

De todas las tablas, la que más demoró en cargarse fue `problems` la cual dada su estructura es bastante más pesada debido a la columna `plain_text_content`, la cual en repetidas instancias contiene una cantidad importante de texto plano (ver Figuras 1, 1.1.1, 1.1.1).





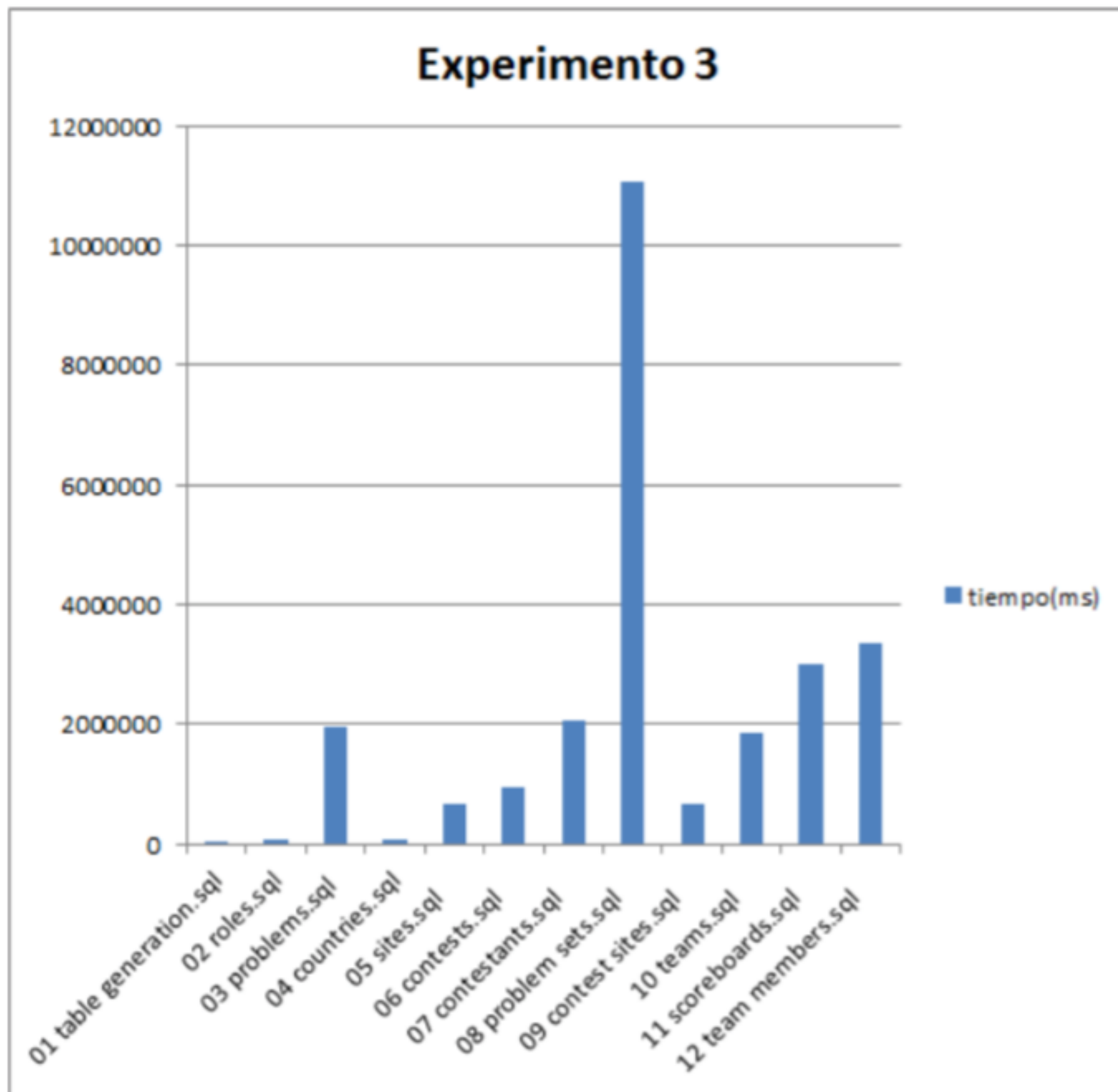


Figura 2: Tiempo de carga de experimento 3.

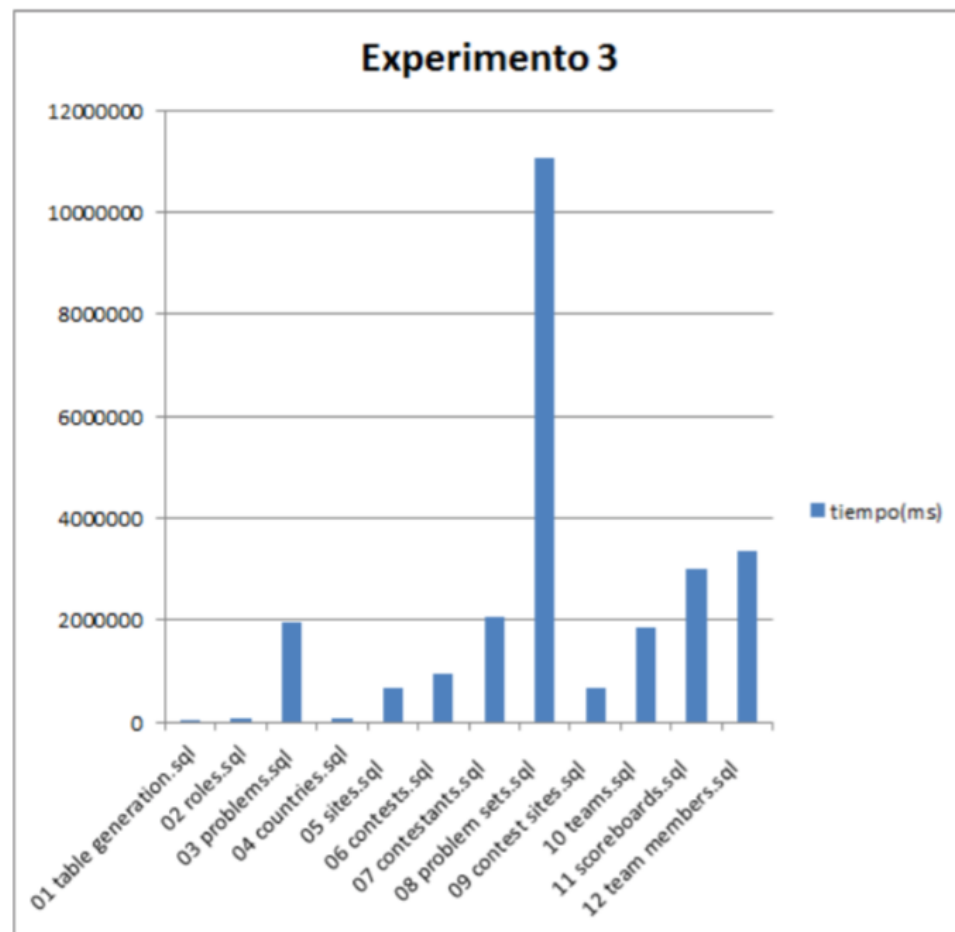


Figura 2: Tiempo de carga de experimento 3.

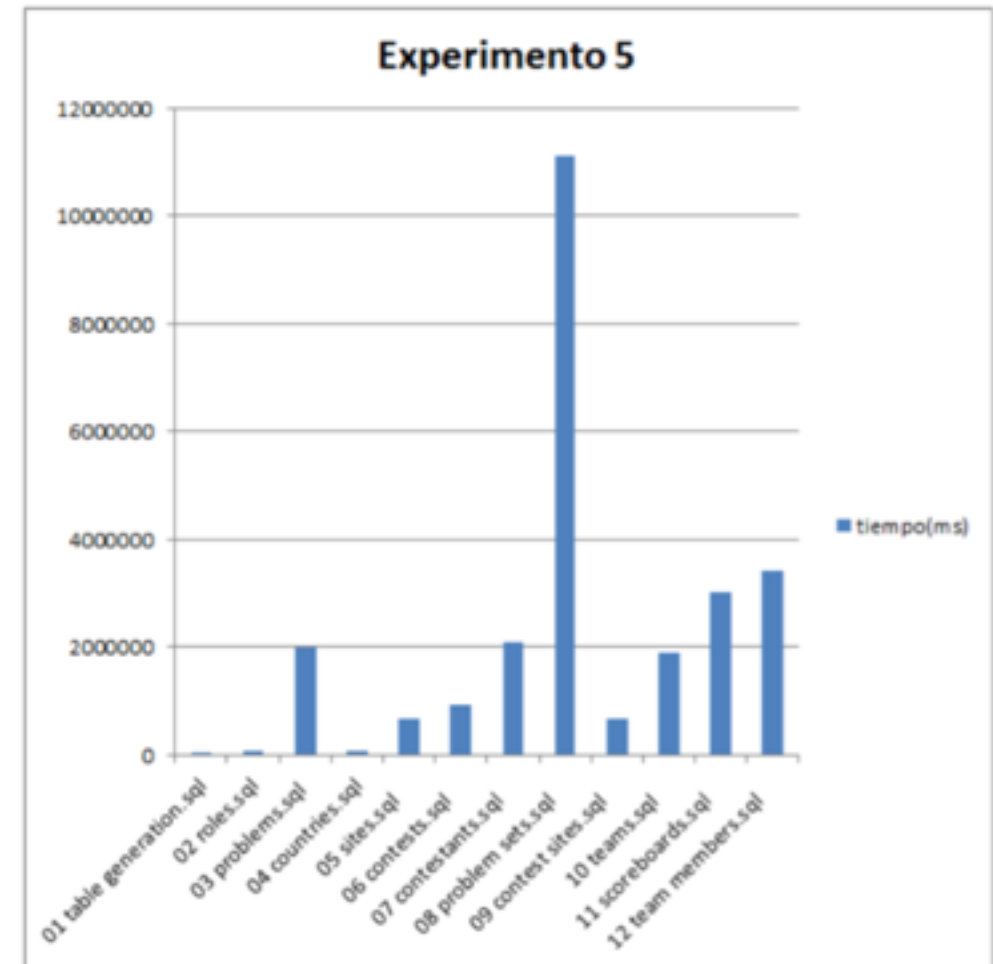


Figura 3: Tiempo de carga de experimento 5.

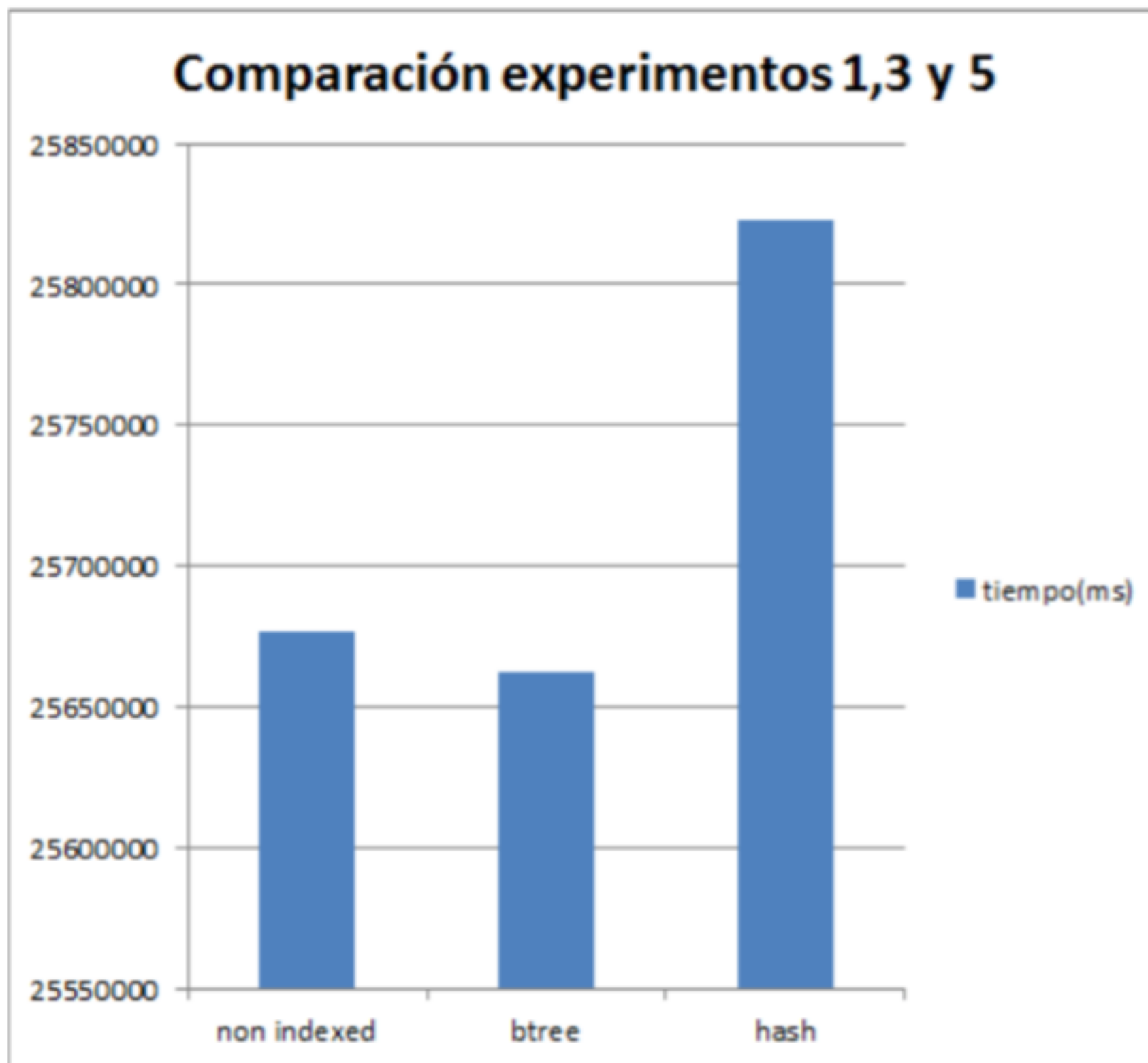


Figura 4: Comparación en tiempo de carga de experimentos 1, 3 y 5.

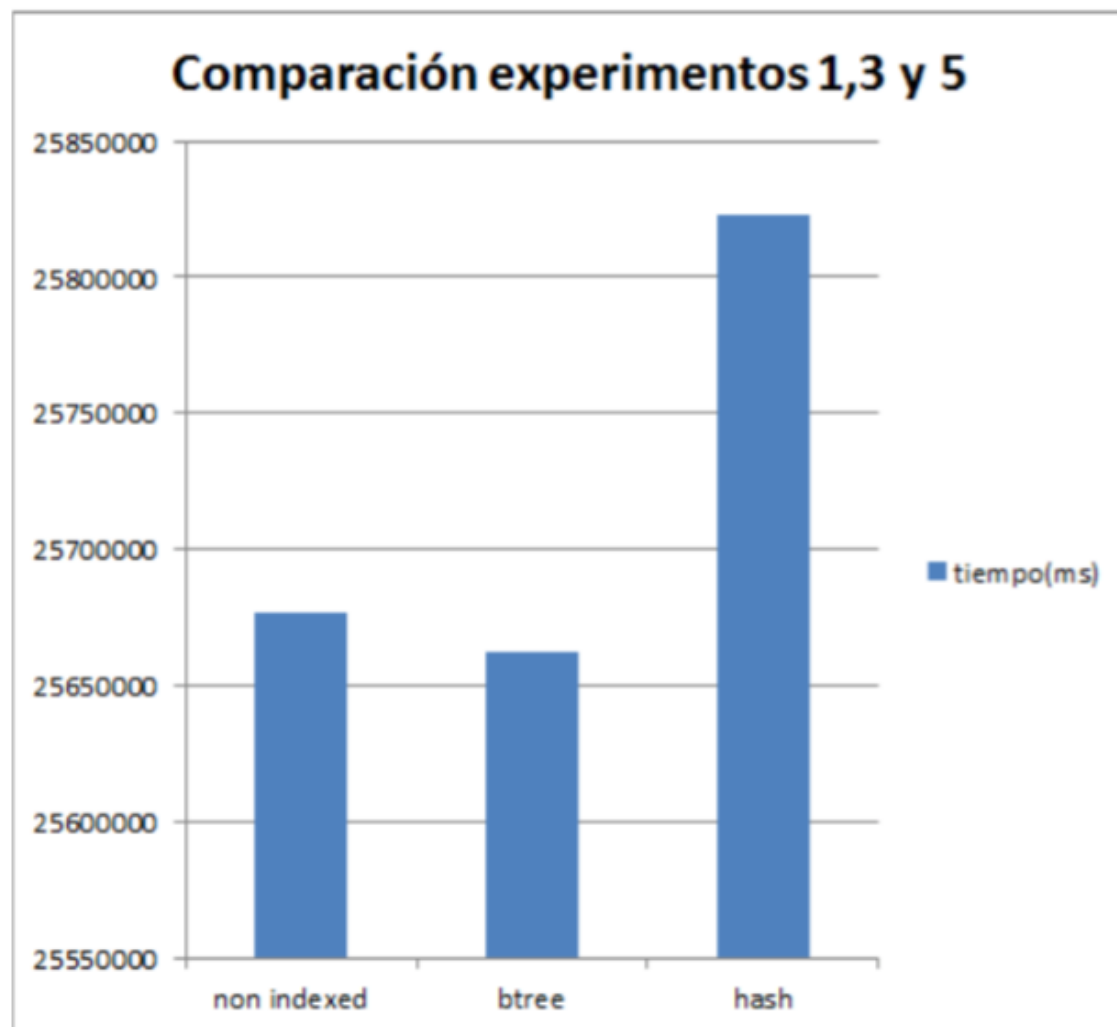


Figura 4: Comparación en tiempo de carga de experimentos 1, 3 y 5.

Table Name	Tuples inserted	Size
contest_sites	330000	14 MB
contestants	880000	157 MB
contests	440000	33 MB
countries	33000	2040 kB
problem_set	5600000	236 MB
problems	880000	534 MB
roles	33000	2032 kB
scoreboards	1500000	75 MB
sites	330000	28 MB
team_members	1600000	80 MB
teams	880000	101 MB

Comparación experimentos 2, 4, 6 y btree + hash

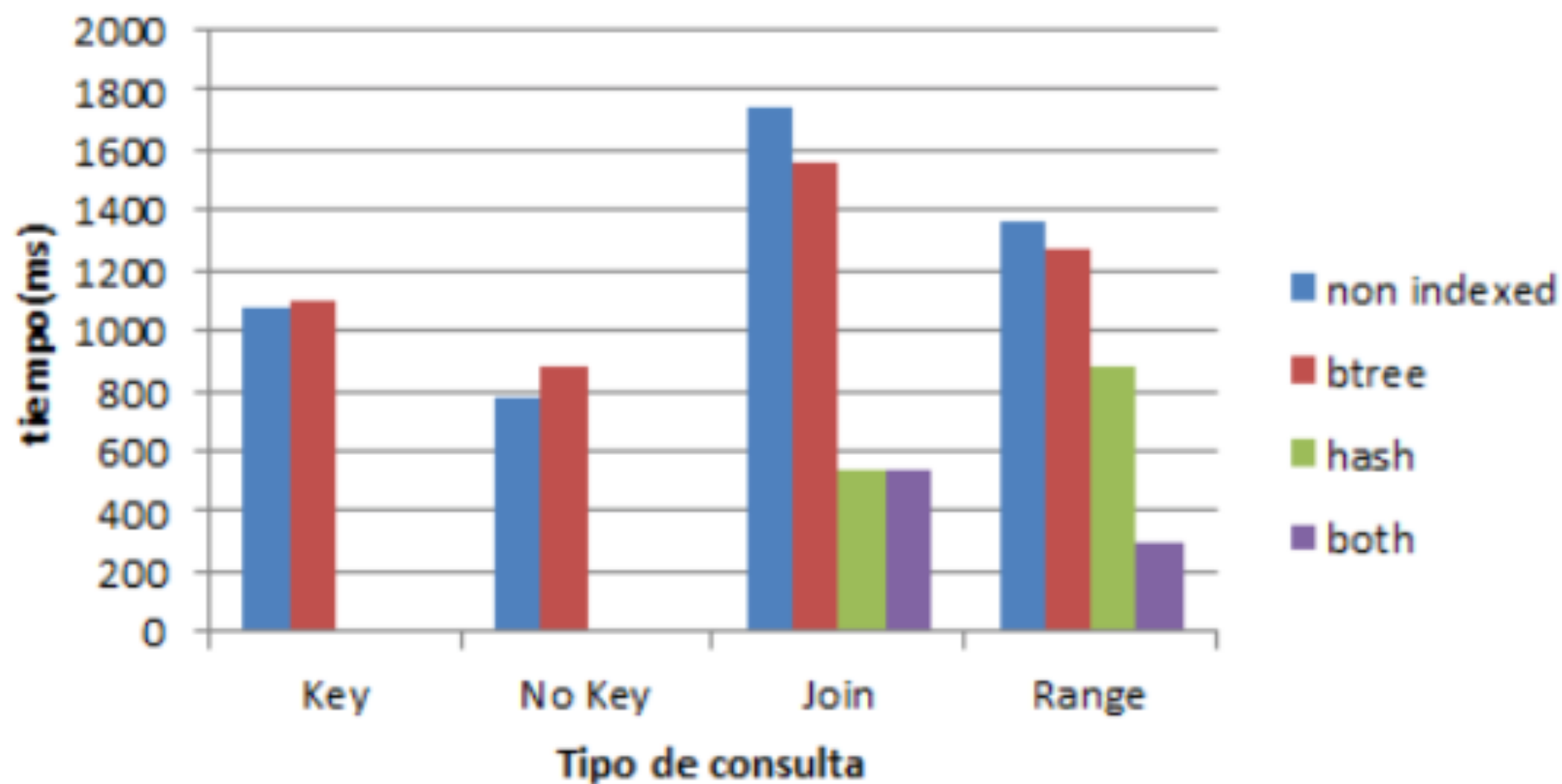


Figura 10: Comparación en tiempo de experimentos 2, 4 y 6.

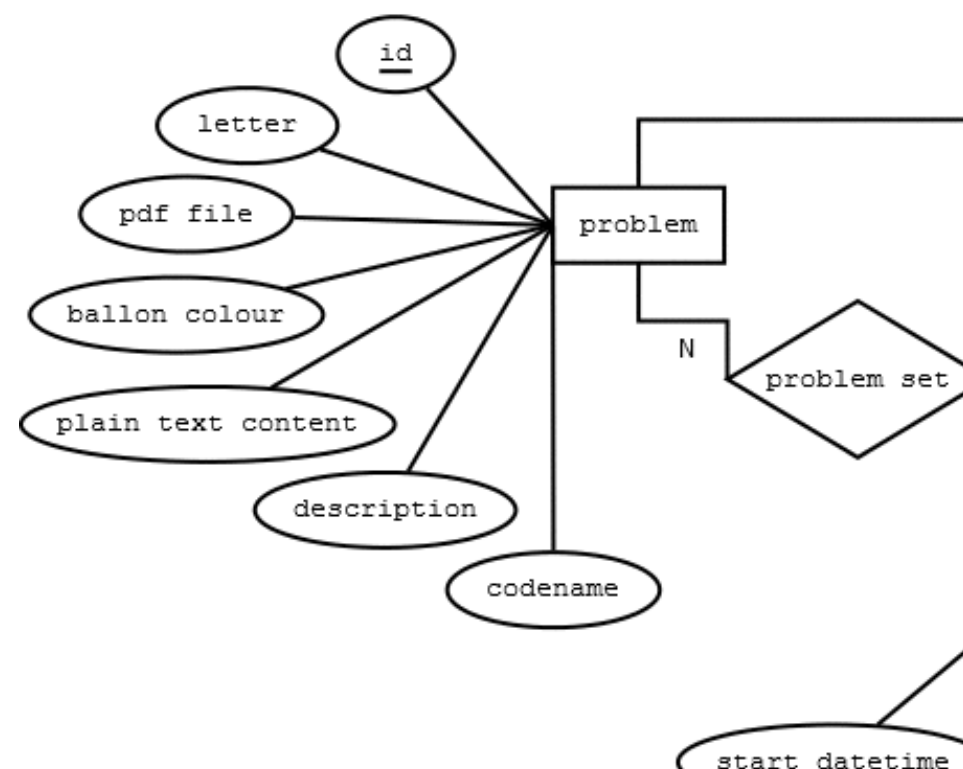


Etapas 3

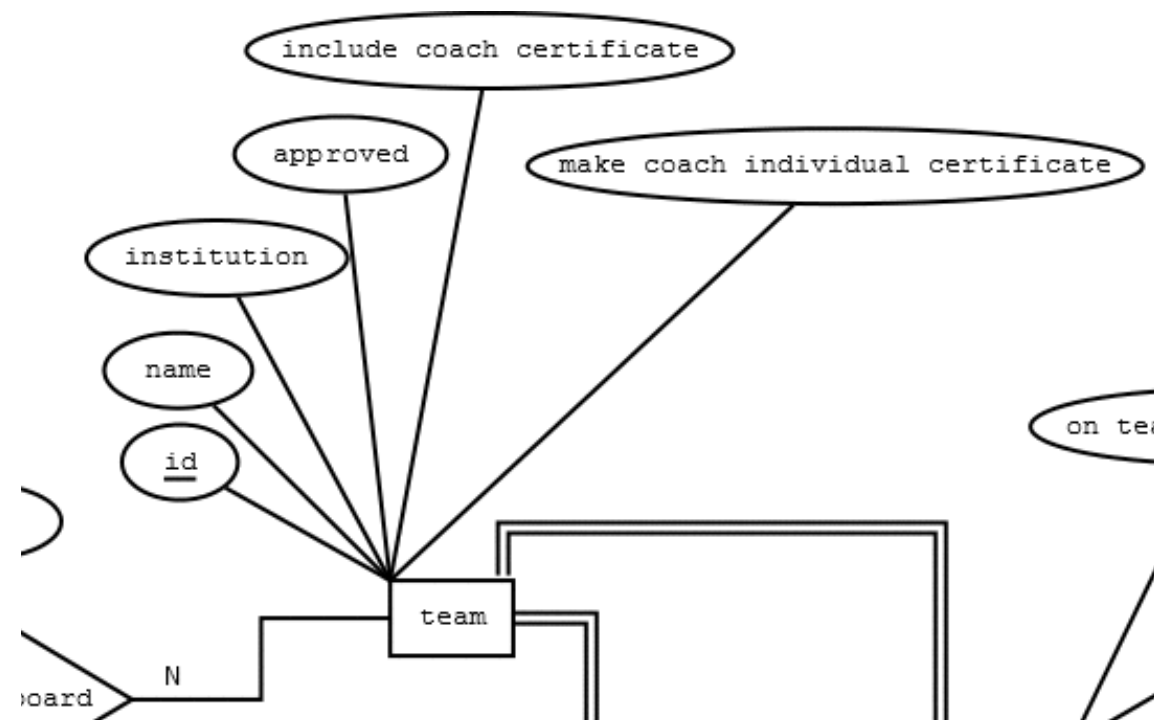
Desnormalizando las tablas

1. *Relación 1-1:* Se separa columna `plain_text_content` de `problems` porque es muy grande y no se consulta frecuentemente.
2. *Relación 1-N no clave:* Nueva columna `contest_name` en tabla `problems` dado que consultar a que competencia pertenece un problema dado es bastante frecuente.
3. *Relación 1-N clave:* Columna `country_id` añadida a tabla `teams` dado que es muy frecuente consultar el país atribuido a un equipo.
4. *Relación N-M:* Nueva columna `team_name` a tabla `team_members` ya que es frecuente consultar los equipos en los que ha participado un competidor.
5. *Tabla de búsqueda:* Se separa al malesscrito `ballon_colour` de la tabla `problems` ya que son 20 colores que se utilizan. De este modo, reducir espacio en la base de datos.

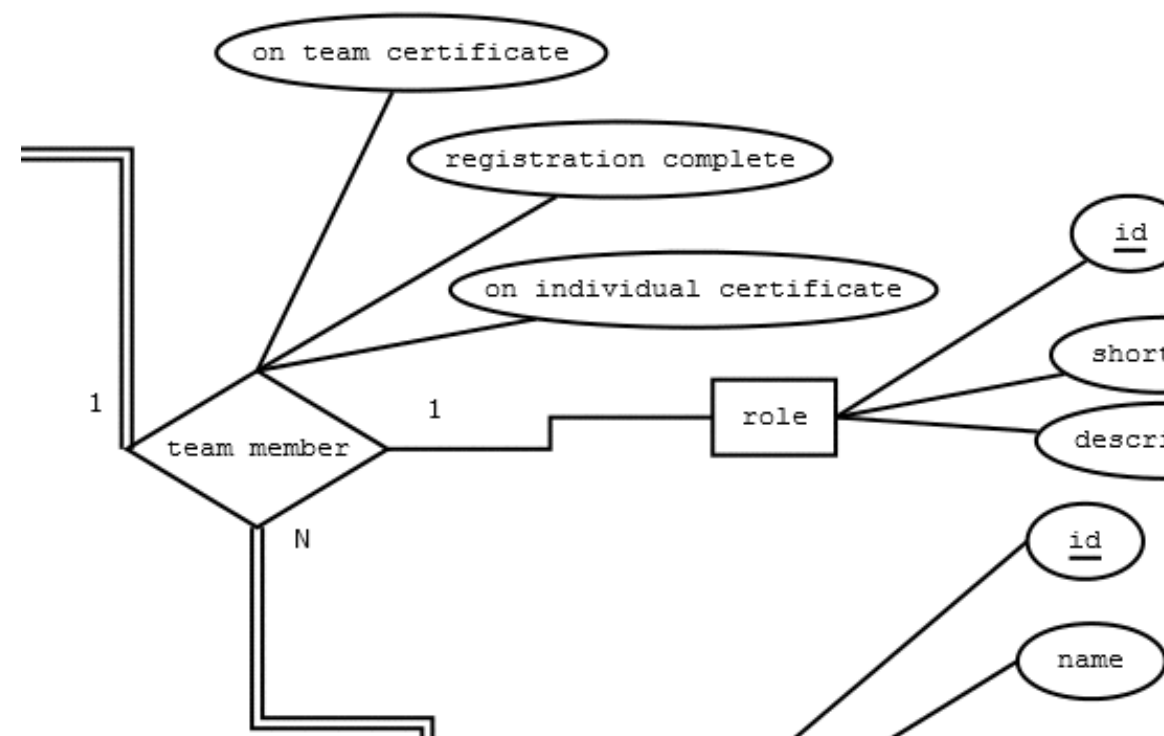
1. *Relación 1-1:* Se separa columna `plain_text_content` de `problems` porque es muy grande y no se consulta frecuentemente.
2. *Relación 1-N no clave:* Nueva columna `contest_name` en tabla `problems` dado que consultar a que competencia pertenece un problema dado es bastante frecuente.
3. *Relación 1-N clave:* Columna `country_id` añadida a tabla `teams` dado que es muy frecuente consultar el país atribuido a un equipo.
4. *Relación N-M:* Nueva columna `team_name` a tabla `team_members` ya que es frecuente consultar los equipos en los que ha participado un competidor.
5. *Tabla de búsqueda:* Se separa al mailescrito `ballon_colour` de la tabla `problems` ya que son 20 colores que se utilizan. De este modo, reducir espacio en la base de datos.



1. *Relación 1-1*: Se separa columna `plain_text_content` de `problems` porque es muy grande y no se consulta frecuentemente.
2. *Relación 1-N no clave*: Nueva columna `contest_name` en tabla `problems` dado que consultar a que competencia pertenece un problema dado es bastante frecuente.
3. *Relación 1-N clave*: Columna `country_id` añadida a tabla `teams` dado que es muy frecuente consultar el país atribuido a un equipo.
4. *Relación N-M*: Nueva columna `team_name` a tabla `team_members` ya que es frecuente consultar los equipos en los que ha participado un competidor.
5. *Tabla de búsqueda*: Se separa al maletín `ballon_colour` de la tabla `problems` ya que son 20 colores que se utilizan. De este modo, reducir espacio en la base de datos.



1. *Relación 1-1*: Se separa columna `plain_text_content` de `problems` porque es muy grande y no se consulta frecuentemente.
2. *Relación 1-N no clave*: Nueva columna `contest_name` en tabla `problems` dado que consultar a que competencia pertenece un problema dado es bastante frecuente.
3. *Relación 1-N clave*: Columna `country_id` añadida a tabla `teams` dado que es muy frecuente consultar el país atribuido a un equipo.
4. *Relación N-M*: Nueva columna `team_name` a tabla `team_members` ya que es frecuente consultar los equipos en los que ha participado un competidor.
5. *Tabla de búsqueda*: Se separa al malesscrito `ballon_colour` de la tabla `problems` ya que son 20 colores que se utilizan. De este modo, reducir espacio en la base de datos.



1. *Relación 1-1*: Se separa columna `plain_text_content` de `problems` porque es muy grande y no se consulta frecuentemente.
2. *Relación 1-N no clave*: Nueva columna `contest_name` en tabla `problems` dado que consultar a que competencia pertenece un problema dado es bastante frecuente.
3. *Relación 1-N clave*: Columna `country_id` añadida a tabla `teams` dado que es muy frecuente consultar el país atribuido a un equipo.
4. *Relación N-M*: Nueva columna `team_name` a tabla `team_members` ya que es frecuente consultar los equipos en los que ha participado un competidor.
5. *Tabla de búsqueda*: Se separa al malesscrito `ballon_colour` de la tabla `problems` ya que son 20 colores que se utilizan. De este modo, reducir espacio en la base de datos.

Tabla 3: Estadísticas de migración

Query	Aff. Rows	Time(ms)	Prev. Size (MB)	Final size(MB)
C_1, C_2	1760088	7669	problems = 534	d1_problems_content = 335 d1_problems = 91
C_3	12520000	51977	problems = 534	d1_problems = 2698MB
C_4	2090000	3054	sites = 28 teams = 101	d3_teams = 107
C_5	4080000	4195	team_members = 80 team = 101	d4_team members = 137
C_6, C_7	1760020	7070	problems = 534	d5_colours = 0.08 d5_problems = 403

Tabla 3: Estadísticas de migración

Query	Aff. Rows	Time(ms)	Prev. Size (MB)	Final size(MB)
C_1, C_2	1760088	7669	problems = 534	d1_problems_content = 335 d1_problems = 91
C_3	12520000	51977	problems = 534	d1_problems = 2698MB
C_4	2090000	3054	sites = 28 teams = 101	d3_teams = 107
C_5	4080000	4195	team_members = 80 team = 101	d4_team members = 137
C_6, C_7	1760020	7070	problems = 534	d5_colours = 0.08 d5_problems = 403

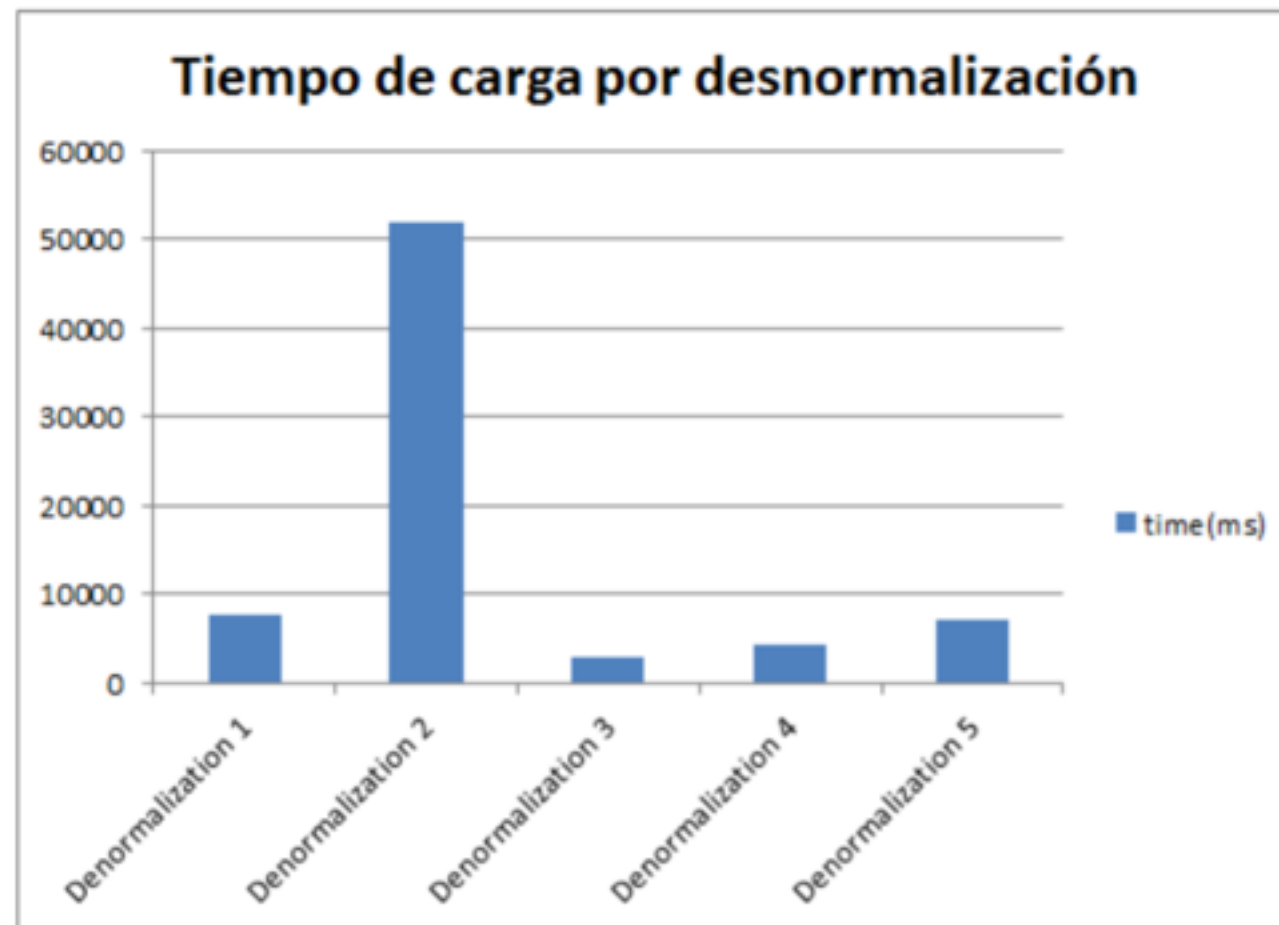


Figura 9: Tiempos de carga para BD desnormalizada.

Tabla 3: Estadísticas de migración

Query	Aff. Rows	Time(ms)	Prev. Size (MB)	Final size(MB)
C_1, C_2	1760088	7669	problems = 534	d1_problems_content = 335 d1_problems = 91
C_3	12520000	51977	problems = 534	d1_problems = 2698MB
C_4	2090000	3054	sites = 28 teams = 101	d3_teams = 107
C_5	4080000	4195	team_members = 80 team = 101	d4_team members = 137
C_6, C_7	1760020	7070	problems = 534	d5_colours = 0.08 d5_problems = 403

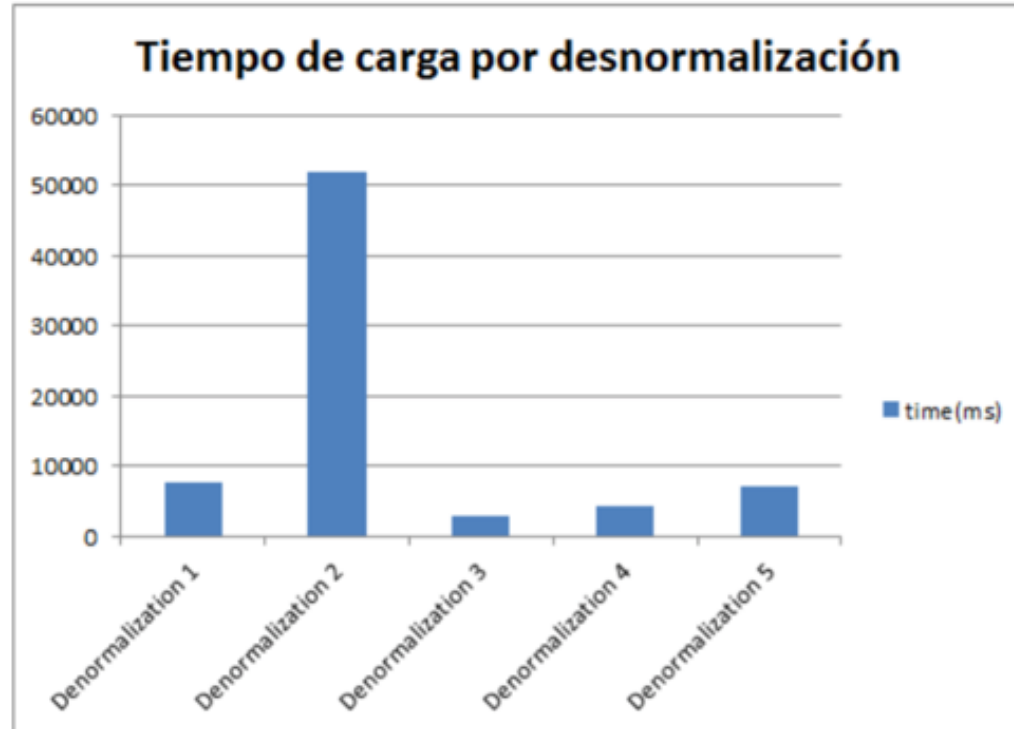


Figura 9: Tiempos de carga para BD desnormalizada.

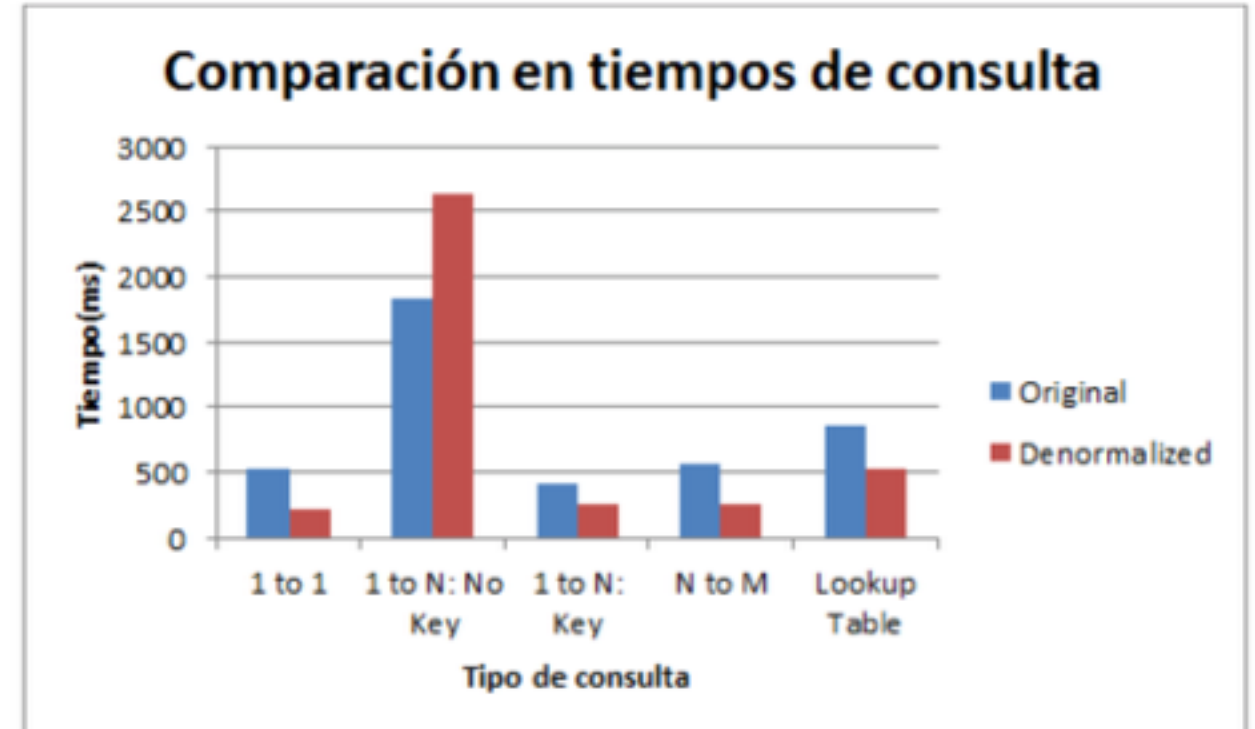


Figura 12: Comparación entre tiempos de consulta de BD original y desnormalizada.

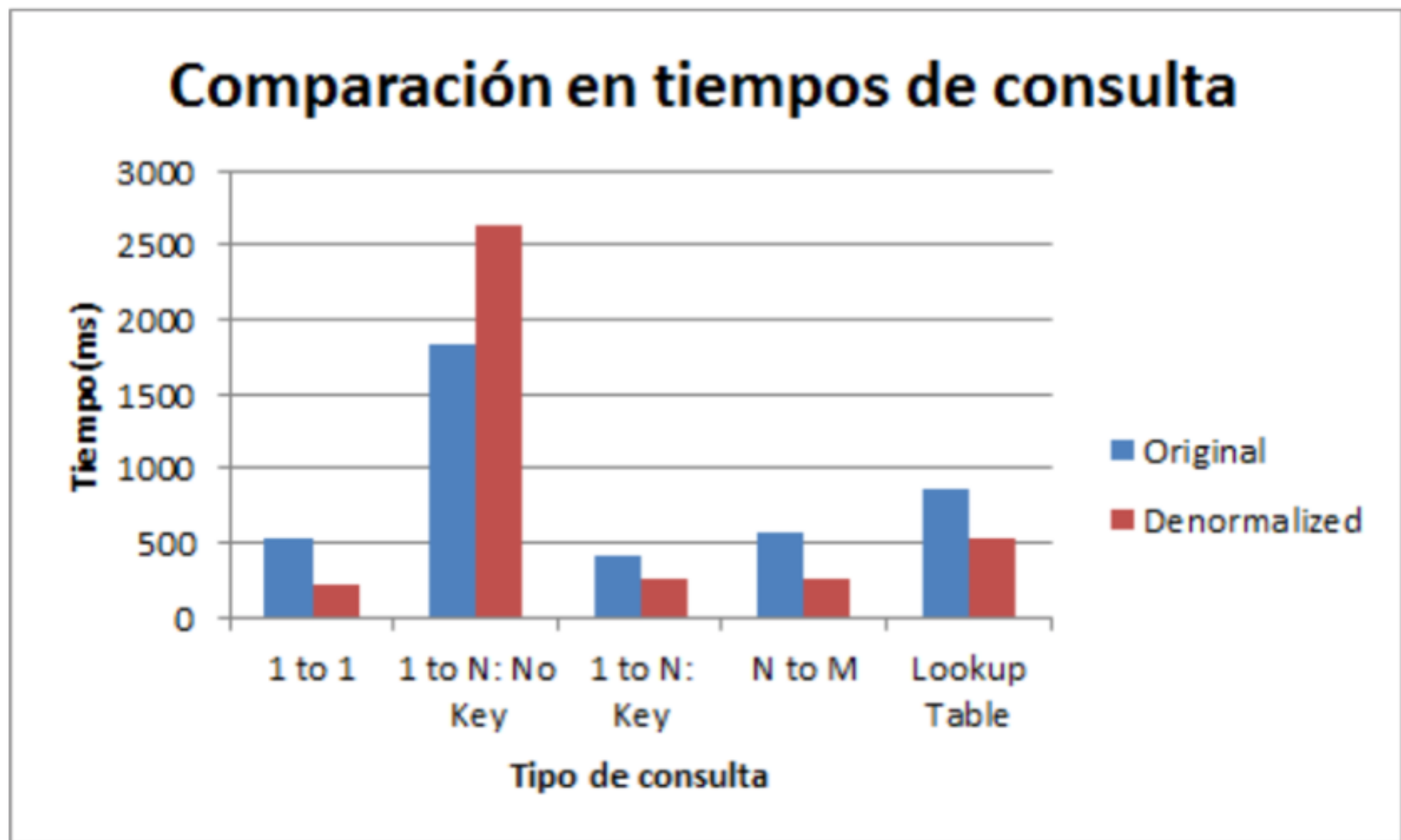


Figura 12: Comparación entre tiempos de consulta de BD original y desnormalizada.



Etapa 4

Vistas materializadas y procedimientos almacenados

Tabla 5: Tiempos de ejecución para PgAdmin3 y Conector en milisegundos

Query	PgAdmin3	Conector
FINAL_1	1600	1906
FINAL_2	1600	1821
FINAL_3	32	1
FINAL_4	1800	1722
FINAL_5	1800	1640

Tabla 5: Tiempos de ejecución para PgAdmin3 y Conector en milisegundos

Query	PgAdmin3	Conector
FINAL_1	1600	1906
FINAL_2	1600	1821
FINAL_3	32	1
FINAL_4	1800	1722
FINAL_5	1800	1640

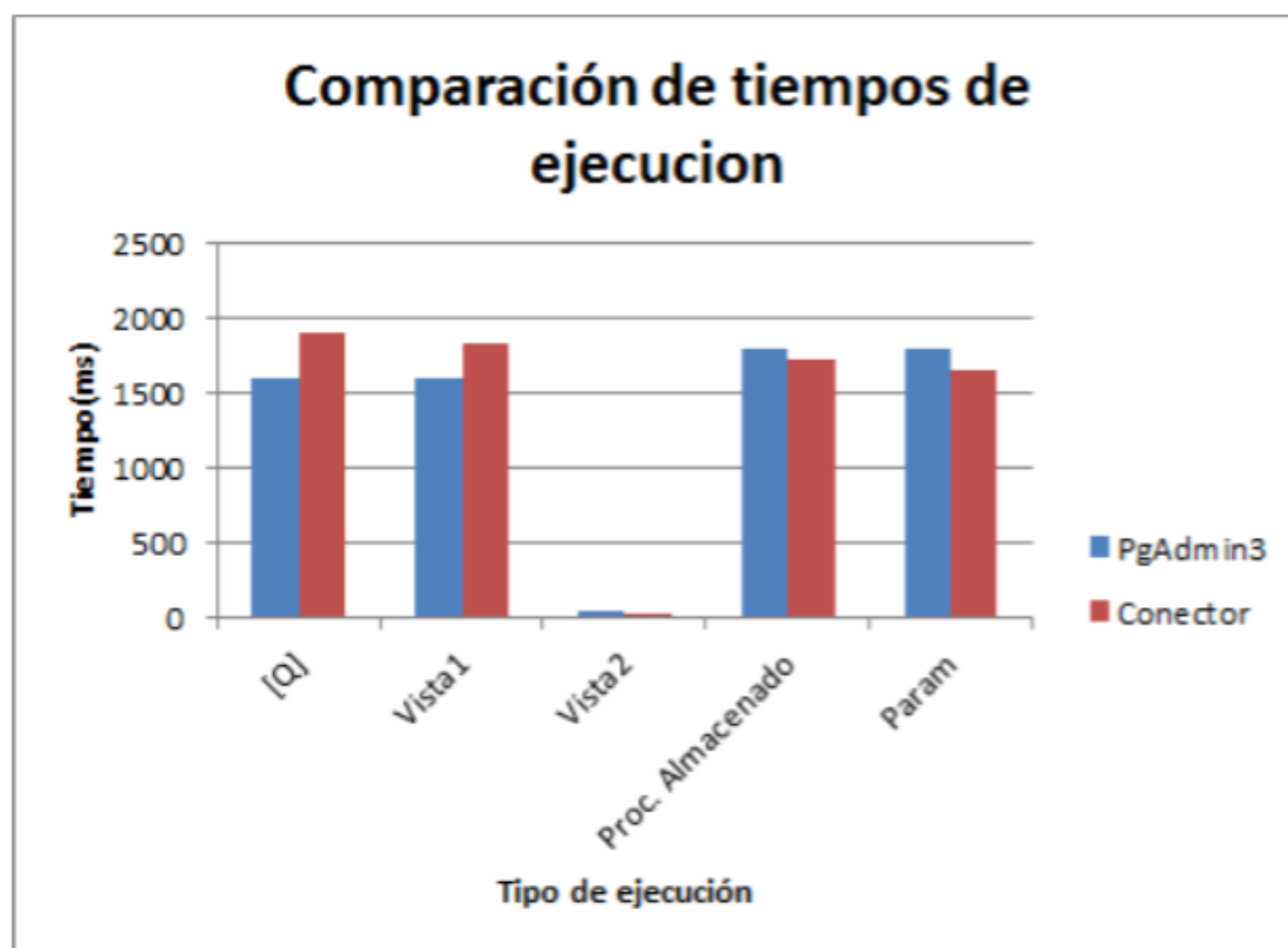


Figura 13: Comparación de tiempos de ejecución.

Comparación de tiempos de ejecución

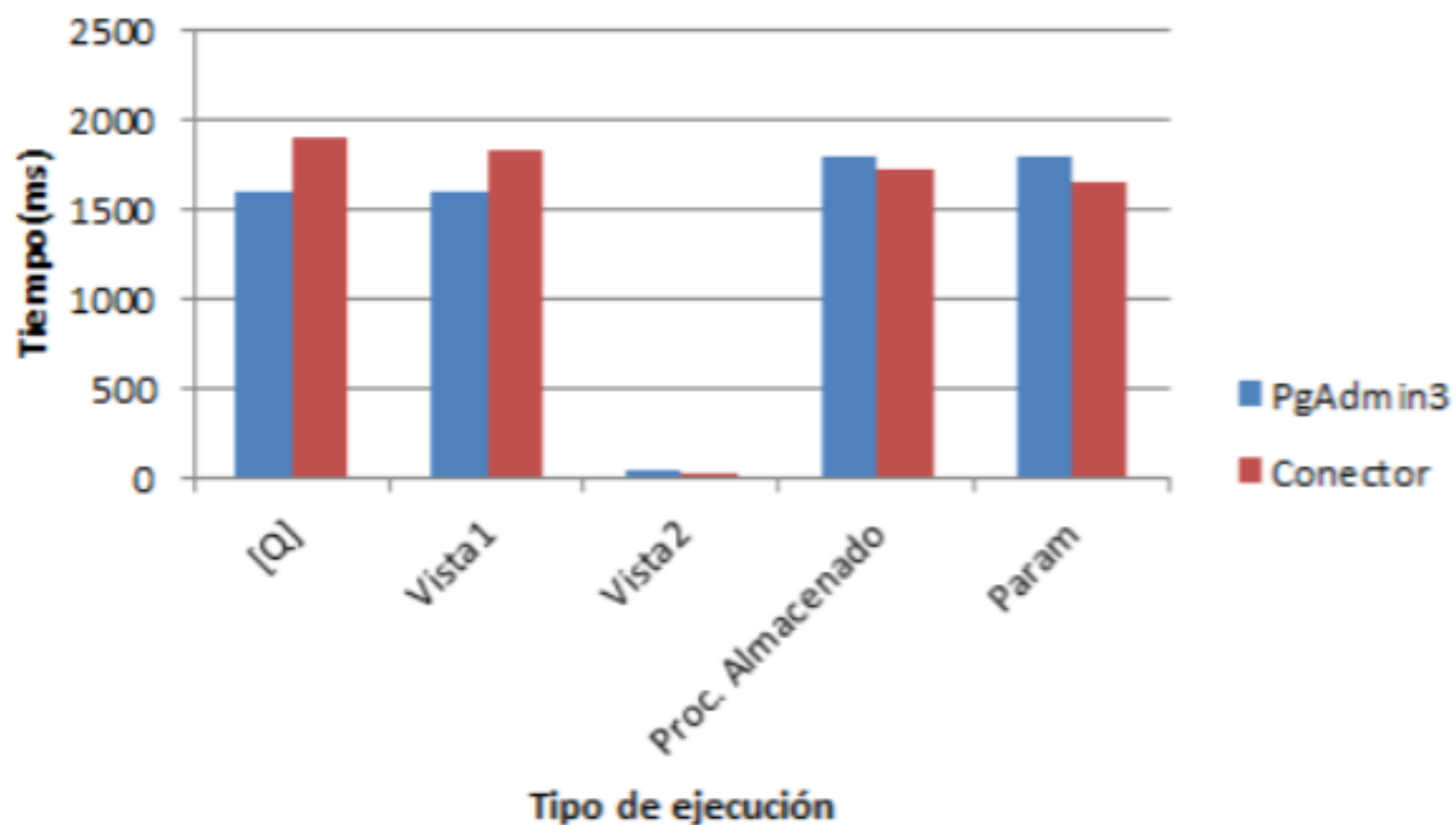


Figura 13: Comparación de tiempos de ejecución.

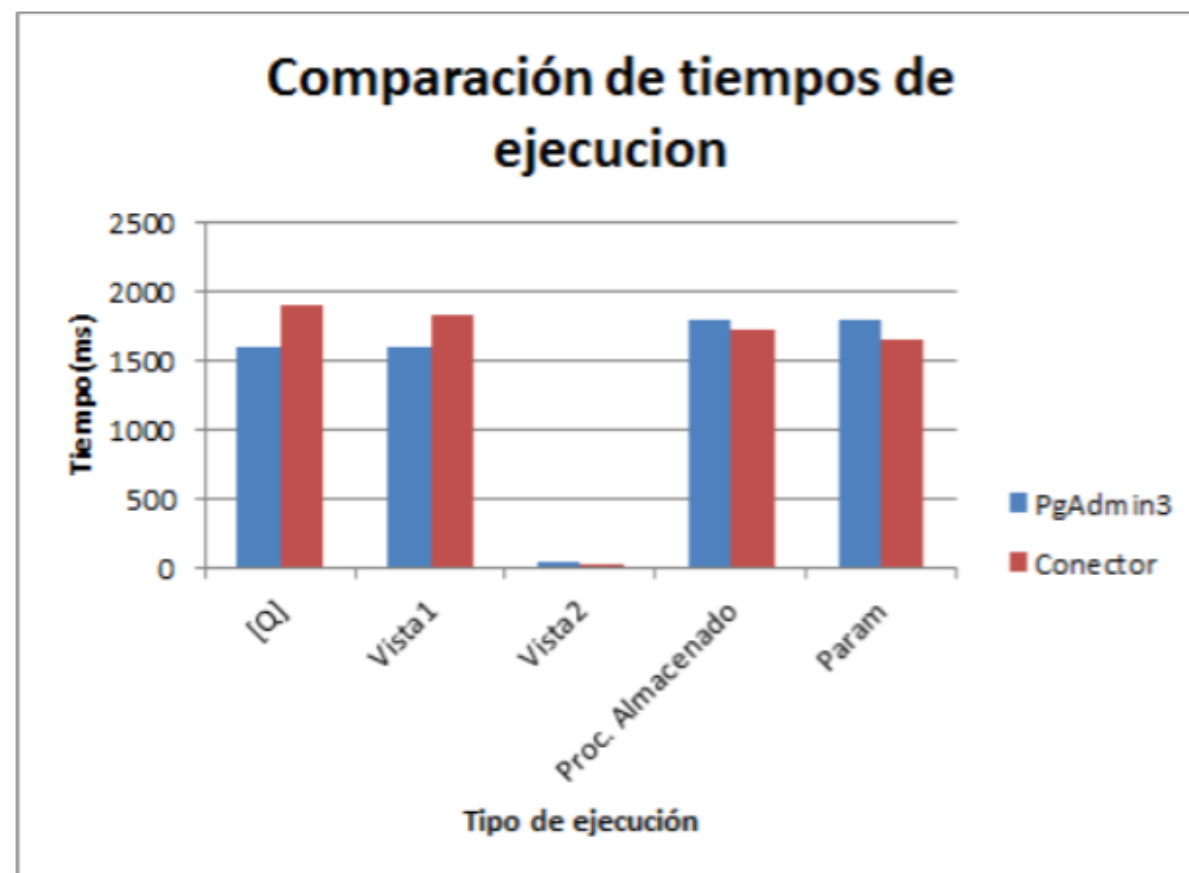


Figura 13: Comparación de tiempos de ejecución.

Table Name	Tuples inserted	Size
contest_sites	330000	14 MB
contestants	880000	157 MB
contests	440000	33 MB
countries	33000	2040 kB
d1_problems	880000	91 MB
d1_problems_content	880000	335 MB
d2_problems	5600000	2698 MB
d3_teams	880000	107 MB
d4_team_members	1600000	137 MB
d5_colours	20	16 kB
d5_problems	880000	403 MB
materialized	1598667	305 MB
problem_set	5600000	236 MB
problems	880000	534 MB
roles	33000	2032 kB
scoreboards	1500000	75 MB
sites	330000	28 MB
team_members	1600000	80 MB
teams	880000	101 MB

Gracias c:

[https://github.com/ereгла/3407B414_A-](https://github.com/ereгла/3407B414_A-databases_management)
[databases_management](https://github.com/ereгла/3407B414_A-databases_management)

Si es que alguien quiere ver las fuentes~