An evaluation of the impact of end-to-end query optimization strategies on energy consumption

Eros Cedeño[1] Ana Aguilera-Faraco[2] Denisse Muñante[3] Jorge Correia[1] Leonel Guerrero[1] and Carlos Sivira[1] Yudith Cardinale[4]

[1]Universidad Simón Bolívar, Caracas, Venezuela

[2]Universidad de Valparaíso, Chile

[3]ENSIIE & SAMOVAR, Évry, France

[4]Universidad Internacional de Valencia, Spain

The 19th Int. Conf. On Evaluation of Novel Approaches to Software Engineering (ENASE'24)

Outline

Outline

Context (1/2)

Green Software

Green BY Software: The software helps a system to be more energy friendly. For instance a Smart Home.

Green IN Software: The software itself is able to be energy efficient. It could be integrated to all steps of the software development.

Figure: Green BY software *vs* Green IN Software

- **Energy awareness:** the level in which stakeholders are aware about the energy consumption that entails the building and usage of software.
- **Energy efficiency:** the level in which a software application is able to spend less energy without affecting its functionalities.

Context (2/2)

Evaluation of the energy consumed by Relational Database Systems

- Software applications intensively receive and transfer data by using a Relational Database Management Systems (*RDBMS*).
- RDBMS are software components that could require a lot of hardware resources querying larges databases.
- Resources are limited and expensive so they must be used wisely. However, developers are not aware of the resource impact they cause when building software.

Figure: Energy is expensive

Figure: Applications require more energy every time

Problem and Research objectives

Problem

Studies are traditionally interested in reducing execution time by implementing databases optimizations strategies such as index or compression. **However, there are no studies dedicated to analyse the end-to-end (server and client sides) energy impact of these strategies.**

Research objectives

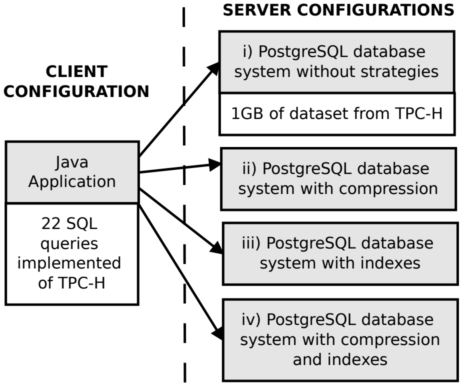
Our goal is to increase awareness about the **end-to-end energy impact** of databases optimizations strategies. It can serve as guidelines for supporting developers to choose the strategies that are more energy friendly.

- RQ1: Does the implementation of database optimization strategies affect the energy consumption of client server applications?
- RQ2: Does the addition of database optimization strategies affect the energy consumption of client server applications?

Outline

Evaluating the impact of end-to-end query optimization strategies on energy consumption

Design of the Study: Setup



Evaluating the impact of end-to-end query optimization strategies on energy consumption

Design of the Study: Algorithm

```
Require: lstOpt; // base, index, compression, index-compression
    numIterations = 30; // number of iterations
    lstQueries; // list of queries (22 queries en total)
Ensure: collect metrics for the analysis of energy consumption of a
    client-server application
for optimization in lstOpt do
    delete the database if it already exists
    create the database according to the configuration of optimization
    for query in lstQueries do
        for i in numIterations do
            restart the service of database server
            start monitoring tools
            execute query in the database server with the config.
            collect metrics (execution time from client, energy
            consumption from client and server)
            stop monitoring tools
        end for
    end for
end for
```

Evaluating the impact of end-to-end query optimization strategies on energy consumption

Juliet: Minimalist and Energy friendly tool, developed in C++

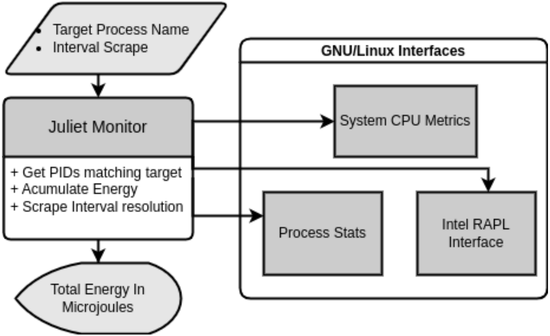


Figure: Juliet Architecture Design

Outline

Results of Study

RQ1: Does the implementation of database optimization strategies affect the energy consumption of client server applications?
RQ2: Does the addition of database optimization strategies affect the energy consumption of client server applications?

Table: Sample Query Results. **AEC**: Average of energy consumption. **RSEC**: Relative standard deviation. **DEC:** Difference of EC vs Base

Query 22						
	Client			Server		
	AEC	RSEC	%DEC	AEC	RSEC	%DEC
Base	119.39	0.06	-	4427.2	0.07	-
Index	16.22	0.07	**86.41**	1.41	0.19	**99.97**
Compression	118.93	0.06	-	4387.18	0.07	-
Index-Compression	16.26	0.05	**86.38**	1.2	0.19	**99.97**

Results of Study

RQ1: Does the implementation of database optimization strategies affect the energy consumption of client server applications?
RQ2: Does the addition of database optimization strategies affect the energy consumption of client server applications?

Table: Sample Query Results. **AEC**: Average of energy consumption. **RSEC**: Relative standard deviation. **DEC:** Difference of EC vs Base

Query 17						
	Client			Server		
	AEC	RSEC	%DEC	AEC	RSEC	%DEC
Base	43.78	0.08	-	393.82	0.03	-
Index	22.74	0.09	**48.06**	241.73	0.01	**38.62**
Compression	23.76	0.08	**45.73**	265.73	0.02	**32.53**
Index-Compression	23.39	0.10	**46.56**	238.51	0.02	**39.44**

Results of Study

RQ1 and RQ2

RQ1

- The **index** strategy reduce between [1.15, 103.17] J for the client side energy consumption, whilst for the **compression** was [1.28, 20.02] J.
- For the server side, the **index** strategy reduce between [1.66, 4425.79] J, whilst the **compression** reduced between [1.21, 128.08] J.

Conclusion: $\mathbf{I} > \mathbf{C}$

RQ2

- For the **index-compression** strategy the reduction of the energy consumption was between [0.92, 86.38] J for the client side, whilst for the server side was between [1.33, 4426] J. These values are similar to the values obtained for the **index** strategy.

Conclusion: $\mathbf{I} + \mathbf{C} \equiv \mathbf{I}$

Results of Study

Some Exceptions: ....

Query 05						
	Client			Server		
	AEC	RSEC	%DEC	AEC	RSEC	%DEC
Base	18.78	0.13	-	21.43	0.06	-
Index	18.32	0.11	-	26.50	0.07	-**23.68**
Compression	18.12	0.12	-	18.18	0.1	15.13
Index-Compression	19.14	0.11	-	25.04	0.08	-**16.84**

SELECT n_name, **sum(**l_extendedprice * (1 - l_discount)**)** **as** revenue
FROM customer, orders, lineitem, supplier, nation, region
WHERE c_custkey = o_custkey
AND l_orderkey = o_orderkey
AND l_suppkey = s_suppkey
AND c_nationkey = s_nationkey
AND s_nationkey = n_nationkey
AND n_regionkey = r_regionkey
AND r_name = 'ASIA'
AND o_orderdate $\geq$ *date* $'1993-01-01'$
AND*o_orderdate* $<$ *date* $'1993-01-01' +$ *interval* $'1'$ *year*
GROUP BY *n_name*
ORDER BY *revenue* **DESC**
LIMIT 1;

Outline

Threats to validity, Conclusion and Future work (1/2)

Threats to validity

- **Internal validity:** Juliet could introduce bias.
- **Construct validity:** Same computer used for running the client and server.
- **External validity:** Benchmarking limited to 22 specific queries, our results cannot be generalized.

Conclusion

- **Indexation tends to be more effective than data compression to reduce the energy** consumed by the execution of the majority of the queries (16/22 queries). The addition of indexation and compression does not lead to a better result comparing with indexation.
- **More experiments are needed to understand the reason of why some queries do not follow this tendency:** i) 3/22 queries increase energy consumption, and ii) 3/22 queries do not show any effect.

Threats to validity, Conclusion and Future work (2/2)

Future work

- Run new experiments in a real distributed software architecture.
- Run new experiments over other RDBMS and NoSQL datasets.
- Classify the queries according their complexity, responses time and size to run new experiments.
- Consider the trade-off of energy consumption of database optimization strategies with other resources such as CPU, I/O usage, and memory.
- Run new experiments to compare Juliet with other tools like Scaphandre or PowerAPI.

Thank you!

16-10216@usb.ve

This work has been partially supported by the **STIC AmSud project**: "GreenSE4IoT - Towards Energy-efficient Software for Distributed Systems" [STIC-AMSUD 22-STIC-04], and has been partially carried out at the **Energy4Climate Interdisciplinary Center (E4C)** of IP Paris and École des Ponts ParisTech, which is in part supported by 3rd Programme d'Investissements d'Avenir [ANR-18-EUR-0006-02].