

# DBSA Project 2 – Extending Databases

[mahmoud.sakr@ulb.be](mailto:mahmoud.sakr@ulb.be), [mariana.machado.garcez.duarte@ulb.be](mailto:mariana.machado.garcez.duarte@ulb.be)

Oct 30, 2022

## Introduction

Create a PostgreSQL (Postgres) extension that provides a Uniform Resource Locator URL data type. A URL represents a pointer to a "resource" on the World Wide Web. A resource can be something as simple as a file or a directory, or it can be a reference to a more complicated object, such as a query to a database or to a search engine. A URL has a structure, which is standardized by W3C. Examples are:

  
http://tools.ietf.org/html/rfc3986  
URI Scheme    Authority/Host    Path

  
http://www.google.be/search?q=ULB&start=10#1  
Query    Fragment

The goal of this project is to create an equivalent type to java.net.URL class in Postgres. More information about the java class can be found here:

<https://docs.oracle.com/javase/8/docs/api/java/net/URL.html>

## Assignment

In this assignment, you are asked to implement a Postgres extension that implements one database type URL. The new type needs to be supported by the following functions (see the link above for the details of the functions):

### Constructors

URL(varchar spec) Creates a URL object from the String representation.
URL(varchar protocol, varchar host, int port, varchar file) Creates a URL object from the specified protocol, host, port number, and file.
URL(varchar protocol, varchar host, varchar file) Creates a URL from the specified protocol name, host name, and file name.
URL(URL context, varchar spec) Creates a URL by parsing the given spec within a specified context.

### Methods and Predicates

Return Type	Method and Description
boolean	equals(URL url1, URL url2) Compares two URLs for equality. This operation must be index-supported.
varchar	getAuthority() Gets the authority part of this URL.
int	getDefaultPort() Gets the default port number of the protocol associated with this URL.
varchar	getFile() Gets the file name of this URL.

varchar	getHost() Gets the host name of this URL, if applicable.
varchar	getPath() Gets the path part of this URL.
int	getPort() Gets the port number of this URL.
varchar	getProtocol() Gets the protocol name of this URL.
varchar	getQuery() Gets the query part of this URL.
String	getRef() Gets the anchor (also known as the "reference") of this URL.
String	getUserInfo() Gets the userInfo part of this URL.
boolean	sameFile(URL url1, URL url2) Compares two URLs, excluding the fragment component. This operation must be index-supported.
boolean	sameHost(URL url1, URL url2) Compares the hosts of two URLs. This operation must be index-supported.
varchar	toString() Constructs a string representation of this URL.

### Indexing

Extend the btree index to support your URL type. The index should support the three predicates: sameHost, sameFile, and equals.

### Submission

- This project contributes 4/20 to the overall grade.
- Implement the above extension using C. It will be accepted to implement the extension using PL/pgSQL, but one mark (1/20) will be deducted.
- A link on UV will be created for submitting your extension. You are asked to submit, per group, a zip file containing: (1) your extension, (2) an SQL file for testing it, (3) if needed attach a document with any important notes or assumptions that you made.
- This assignment is to be made in groups of four. Only one submission is needed per group.

Good luck !