과학기술정보통신부

nipa 정보통신산업진흥원
National IT Industry Promotion Agency

2021
공개SW 페스티벌
ONLINE FESTIVAL

2021
공개SW 페스티벌
ONLINE FESTIVAL

개발자와 함께 성장하는 오픈소스

오픈소스 데이터베이스 큐브리드 개발

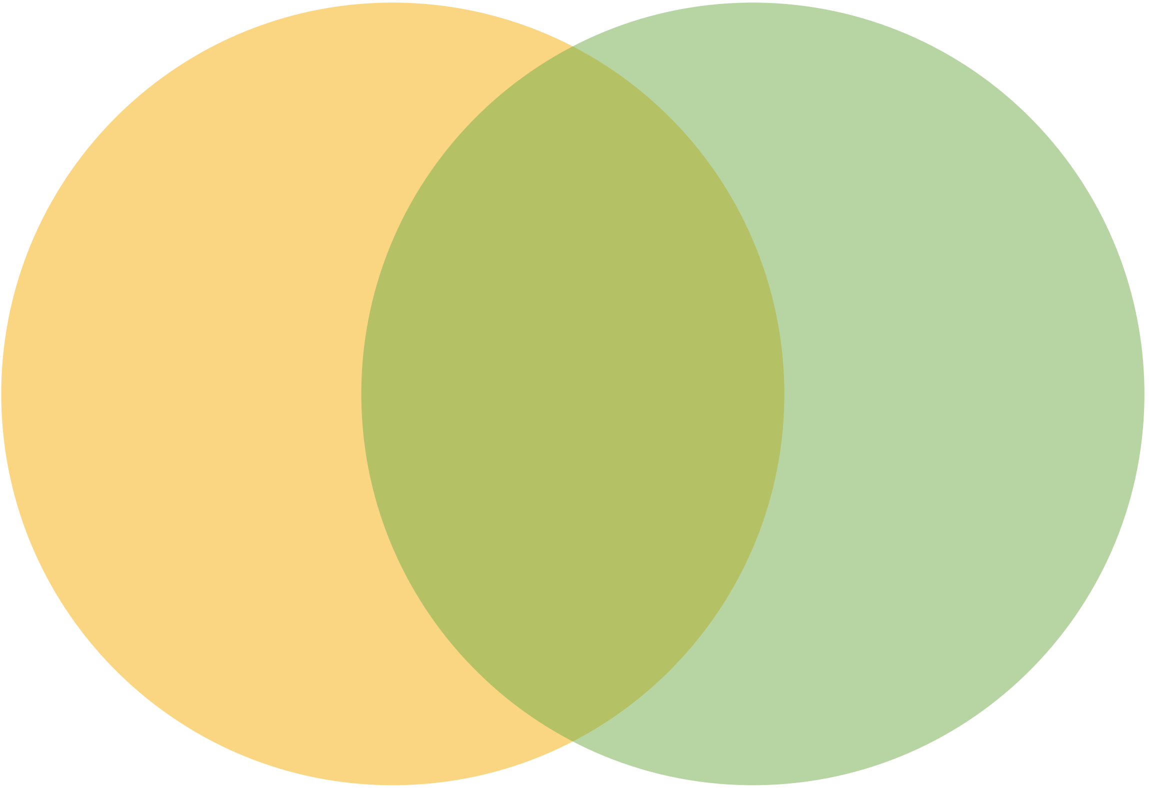
소속 : 큐브리드 (CUBRID)
이름 : 유형규

과학기술정보통신부 nipa 정보통신산업진흥원 National IT Industry Promotion Agency

개발자와 함께 성장하는 오픈소스

- 개발자는 오픈소스에 참여해 어떤 성장을 할 수 있을까?
- 오픈소스는 개발자의 참여로 어떤 성장을 할 수 있을까?

개발자가 얻고 싶은 것



오픈소스 프로젝트가 얻고 싶은 것

CUBRID 프로젝트가 하고 싶은 것

OPEN COMMUNITY!

CUBRID

오픈소스 프로젝트와 함께 성장해보기 튜토리얼

오픈소스 프로젝트에 참여하려면 어떻게 시작하고, 어떤 과정으로 기여하게 될까요?

주의)

- 개인적으로 공부했던 경험을 기반해 작성했습니다
- 더 좋은 방법이 있을 수 있습니다.
- ~~생각보다 쉽지 않을 수 있습니다...~~

신규 기여자가 좀 더 쉽게
프로젝트에 기여를 성공적으로 참여하고
성장할 수 있도록 하자

참여 할 프로젝트 찾기

- 어떤 오픈소스 프로젝트에 참여해보고 싶나요?
- 혹시 DBMS 개발에 관심이 있나요?
- 여러 소프트웨어 모듈로 이루어진 큰 프로젝트에 관심이 있나요?

그럼 CUBRID 프로젝트에 참여해볼까요?

프로젝트 간략히 알아보기

- 참여하려는 프로젝트가 어떤 것인지 먼저 알아 보아야겠죠
- 한번 사용해보면 좋아요
- 프로젝트가 어떻게 운영되고 있는지 살펴볼 필요가 있어요

CUBRID™

데이터를 담는 공간

CUBE

+

데이터를 연결

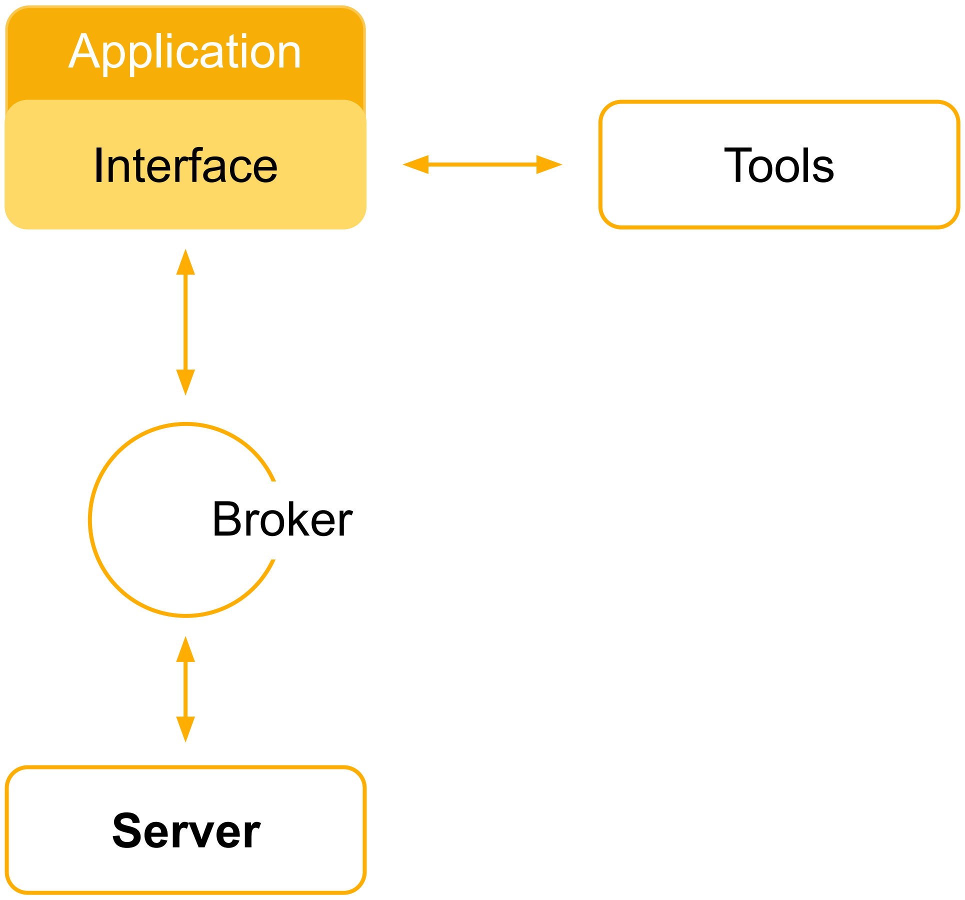
BRIDGE

데이터베이스 관리 시스템 (DBMS)

다수의 사용자들이 데이터베이스 내의 데이터를 접근할 수 있도록 해주는 소프트웨어 도구의 집합

from 위키백과

CUBRID 시스템 구조



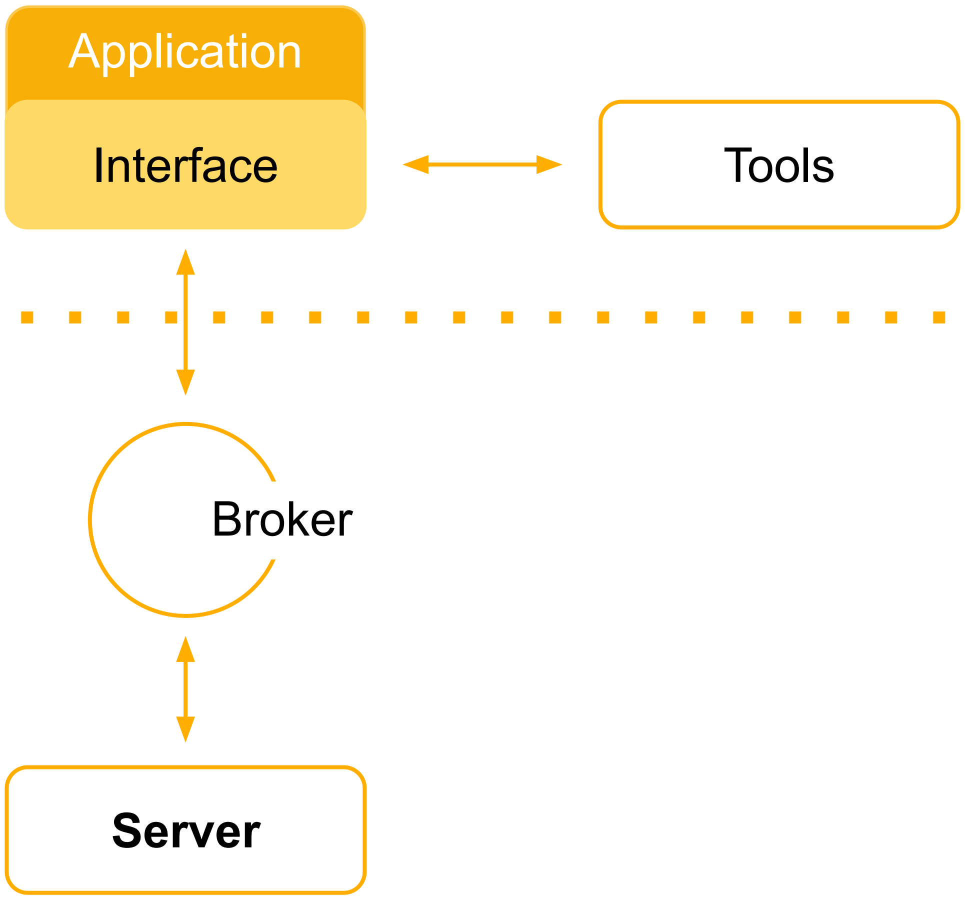
CUBRID 인터페이스/지원 도구

- 드라이버 (JDBC, ODBC, Python, GO…)
- 큐브리드 매니저(DB 모니터링, 질의 에디터)
- 마이그레이션 도구

CUBRID 엔진

- 브로커
- 데이터베이스 서버

CUBRID 오픈소스 라이선스



BSD 라이선스: 자유롭게 수정/배포 가능

CUBRID 11.0 미만: GPL v2 or later
CUBRID 11.0 이상: Apache v2.0

CUBRID 한번 사용해보기

- 프로젝트에 매뉴얼이 있다면 시작하기, 또는 Getting Started를 참고해서 사용해보세요

CUBRID
Search docs
매뉴얼 소개
CUBRID 소개
설치 및 업그레이드
시작하기
CUBRID 서비스 시작
질의 도구
관리 도구
드라이버
CSQL 인터프리터
CUBRID SQL
CUBRID 운영
CUBRID HA
CUBRID 보안
API 레퍼런스
릴리스 노트

» 시작하기

시작하기

CUBRID를 처음 사용하는데 참고할 수 있는 간략한 사용법을 설명한다. CUBRID 서비스 시작하기, CSQL 행 방법을 찾아볼 수 있다.

CUBRID의 다양한 도구 및 드라이버는 https://www.cubrid.org/downloads에서 다운로드할 수 있다.

CUBRID가 제공하는 다음 도구를 사용하면 GUI를 통해 편리하게 CUBRID의 기능을 이용할 수 있다.

- CUBRID 매니저
- CUBRID 마이그레이션 툴킷

CUBRID는 JDBC, CCI, PHP, PDO, ODBC, OLE DB, ADO.NET, Perl, Python, Ruby 등 다양한 드라이버를 API 레퍼런스를 참고한다.

- CUBRID 서비스 시작
  - 셸 명령어
  - CUBRIDService 또는 CUBRID Service Tray
  - 데이터베이스 생성
  - 데이터베이스 시작
- 질의 도구

프로젝트 기여 포인트!

- 그대로 따라 했는데 실행이 안되거나
- 무언가 과정이 빠져있거나
- 설명이 부족할 때

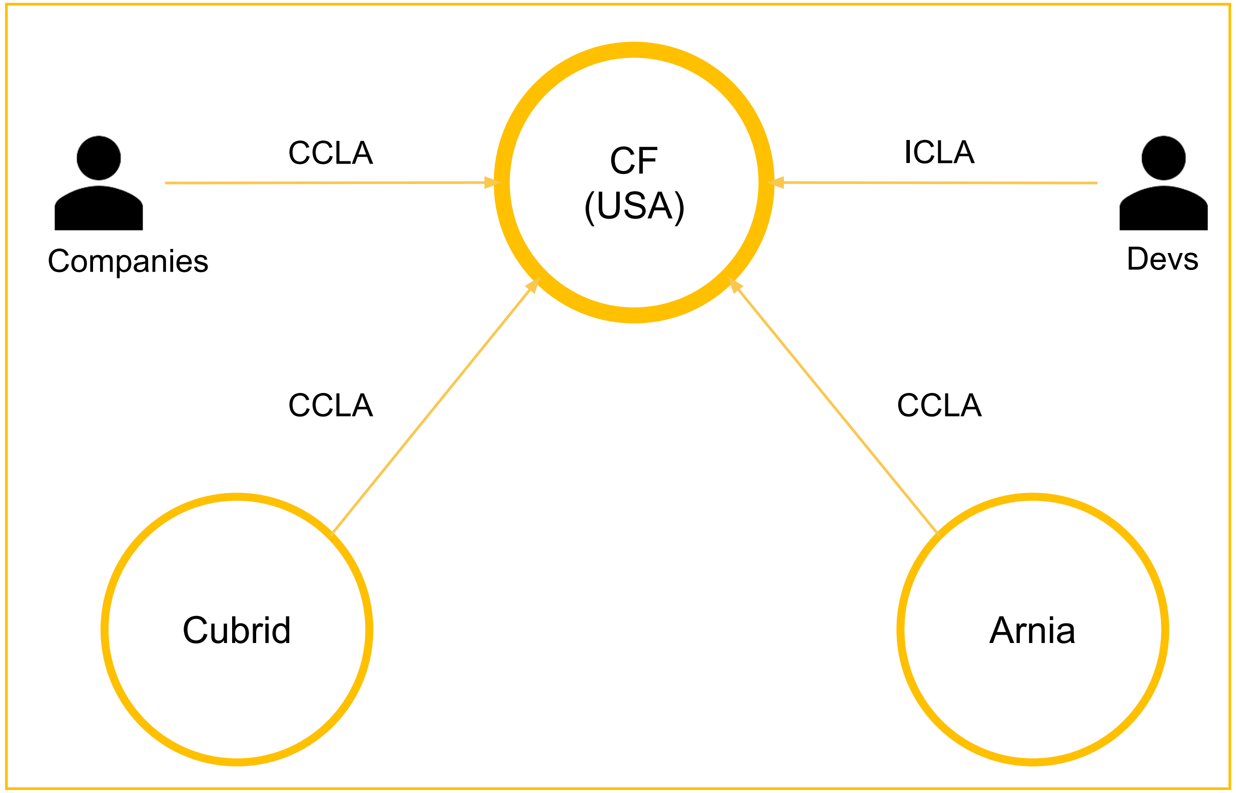
cubrid.org/manual/ko/11.0/start.html

오픈소스 프로젝트가 어떻게 운영되는지 파악해보기

- 개발 작업 (코드리뷰 포함)은 투명하게 볼 수 있나요?
- 소스 반영은 어디에서 이루어지고 있나요?
- 주로 어떤 사람들이 활동하고 있나요?

소스 코드만 공개하고 신규 참여자의 기여에 배타적이라면 ?

CUBRID 오픈소스 프로젝트의 거버넌스



CUBRID Foundation

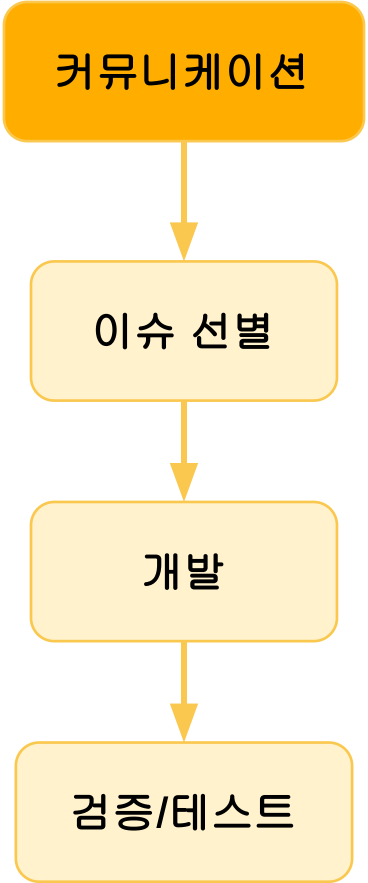
- 글로벌 오픈소스 생태계 구축
- CUBRID 차기 프로젝트 조율/런칭

- 설립: 2020년 2월
- 위치: 미국 워싱턴주 시애틀
- 대표: 김평철
- 대표 관계사
  - CUBRID Corp. / Korea
  - Arnia Corp. / Romania

프로젝트 좀 더 자세히 알아보기

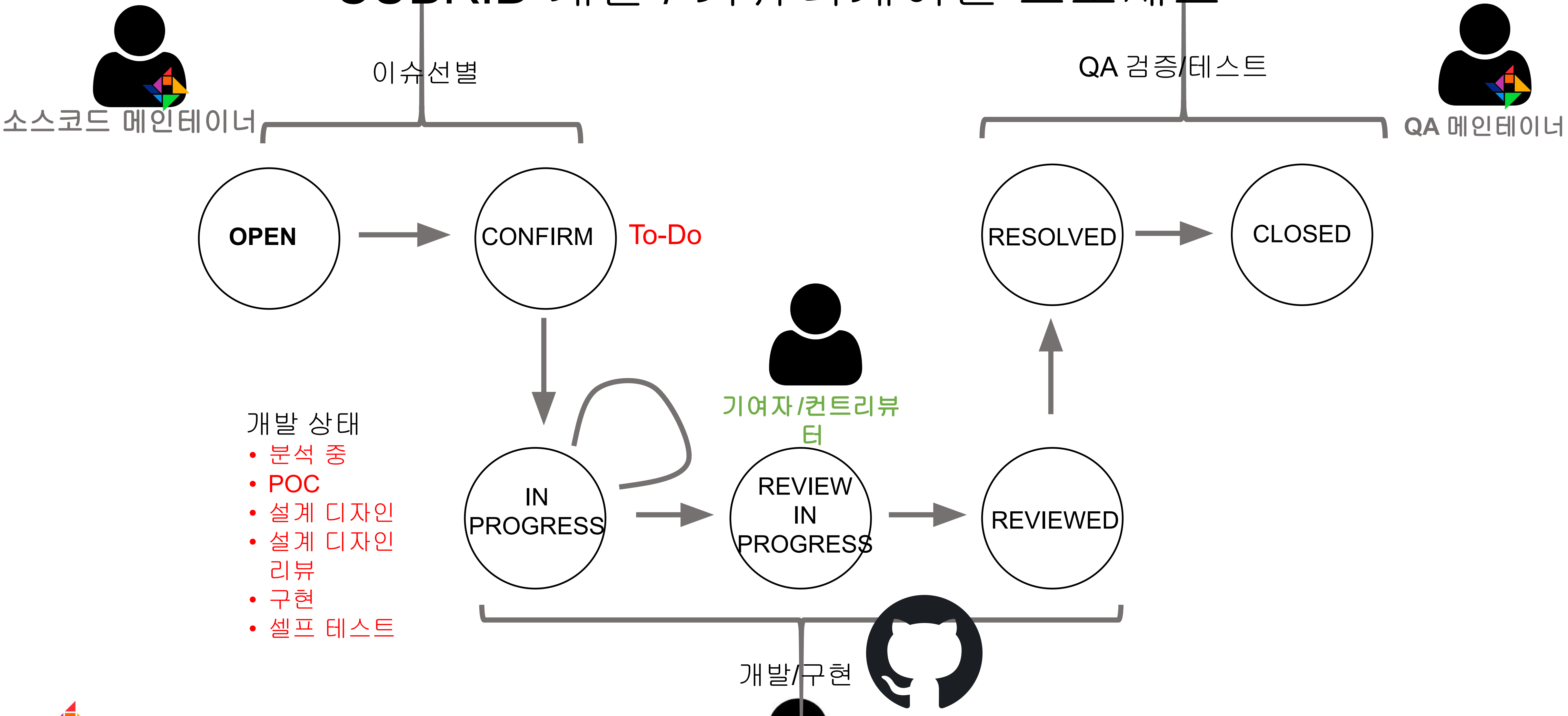
- 프로젝트의 개발 프로세스 파악하기
- 주요 릴리즈 토픽/최근 진행하고 있는 개발 주제를 확인해봅니다
- 소스로부터 직접 빌드해보기
- 프로젝트 구조와 아키텍쳐를 파악해보면 큰 그림을 이해하기 좋아요

CUBRID 의 개발 프로세스



- 개발 참여를 위한 프로젝트, 기능 추가, 버그 수정에 대한 제안과 토론
- 질문/답변, 정보 제공을 포함한 모든 오픈소스 커뮤니티 활동
- 개발 토픽/버그 리포트 : jira.cubrid.org
  - JIRA에 모든 의사결정과정/진행사항이 공유
  - 모든 프로젝트, 기능 추가, 버그 수정은 JIRA의 이슈 생성으로 시작
- JIRA에서 이슈를 생성하면, 이슈의 상태로 개발 작업을 관리

CUBRID 개발 / 커뮤니케이션 프로세스



주요 릴리즈 된 토픽 확인

- CUBRID는 Github에서 프로젝트를 릴리즈
  - Releases · CUBRID/cubrid · GitHub

CUBRID 11.0

CUBRID 11.0 is the latest stable version that includes new features, significant changes and enhancements.

CUBRID 11.0

- is a version with improved security.
- is more stable, faster, and more convenient for administrators.
- fixes a large number of critical bugs..
- includes useful SQL extensions: Supporting RVC (Row Value Constructor) and various REGEXP functions.
- includes code refactoring and modernization.

CUBRID 11.0 **improves security** by providing data encryption and packet encryption. This version prevents abnormal data loss by supporting table-based TDE (Transparent Data Encryption) and packet encryption between the driver and server.

CUBRID 11.0 is **faster**. This version supports hash scan and improves the performances by up to 10 times in join query that could not perform index scans. By supporting the cache of search query results through hints, data change is minimal, and the performance of the workload with complex queries is maximized.

CUBRID 11.0 **improves administrator convenience** by providing new functions for administrators. This version supports statement-based replication through hints on the HA environment, improving the replication time when deleting and updating a large amount of data. By separating the Java SP server from the DB server, the influence of the DB server is minimized from the start/stop of the Java SP server. In addition, the DDL audit function is provided so that DDL change can be tracked.

주요 릴리즈 된 토픽 확인

- Change logs를 보면 중요한 주제 키워드를 얻어갈 수 있어요

Change logs:

Enhancements

Security

0cf60dd [CBRD-23608] Implement "Transparent Data Encryption" (#2497, #2511, #2514, #2516, #2519, #2526, #2546, #2551)
9123b9b [CBRD-23637] Implementation of secure CUBRID jdbc driver (#2346, #2360, #2393)
9c93886 [CBRD-23687] Apply SSL Broker/CAS to damson (#2348, #2350, #2387, #2474)
35c5e35 [CBRD-23688] Implementation of secure CUBRID CCI driver (#2362, #2472, #2480)
f9abbb1 [CBRD-23726] Add DDL Audit log (#2481, #2527)

SQL extension

6d597de [CBRD-22856] Support Row Value Constructor (#1564)
3df36a1 [CBRD-22783] Add support for REGEXP_REPLACE function (#1509, #1647, #1671, #2181, #2186, #2189, #2194, #2198, #2210, #2212)
7d24258 [CBRD-23071] REGEXP_SUBSTR function (#1709)
34e53ff [CBRD-23073] REGEXP_INSTR function (#1722)
20c0159 [CBRD-23084] REGEXP_COUNT function (#1741)
0583db9 [CBRD-23085] REGEXP_LIKE function (#2203)

예시)

- Transparent Data Encryption
- Secure JDBC Drvier
- Row Value Constructor
- REGEX_XXX Function

…

주요 릴리즈 된 토픽 확인

- JIRA의 설명과 토론 과정, 그리고 코드 반영과 테스트 코드 반영을 확인

CUBRID / CBRD-23608

Implement "Transparent Data Encryption"

Description

Transparent Data Encryption

Transparent Data Encryption (TDE) encrypts data stored in persistent storage (at rest) transparently. "Transparent" here means that users are not aware that encryption is occurring. In other words, if you use the TDE feature on your database, it works the same as before, which means you don't need to make any changes to existing schemas or queries. All encryption and decryption occur when data moves up and down between memory and disk automatically behind your back.

Background

With the authorization system provided by databases, only authorized users can access the data. However, if a malicious user accesses data **directly** by taking a database file or disk, data leakage can occur. If data is always encrypted when it comes down (be flushed) to disk, this vulnerability can be eliminated. In particular, it is getting much more important as the cloud environment becomes more common. In the cloud environment, an administrator for a database does not directly manage the physical equipment for the database but runs the database on an external infrastructure (IasS, PasS) provided by a cloud service provider. Or you can even leave the management of the database to the service provider and use the database as a service (DBasS). In this case, there is a possibility for other than the DBA to be able to access the physical storage device, so sensitive data must be encrypted before flushing to disk.

One solution is TDE, an encryption method that is efficient because data is encrypted in the database kernel level and minimizes changes to applications. Other major databases already provide TDE feature though details like encryption granularity and target are slightly different. And the government also recommends TDE for security.

Data encrypted together

TDE encrypts sensitive user data in tables. Encrypting user data in a table means that all related data have to be encrypted together. As the table is encrypted, the data that should be encrypted together is as follows.

- Indices created on Encrypted Table
- REDO, UNDO log for encrypted data
- Temporary data including join intermediate results or query results
- Backup of encrypted data
- Data and ~~replication logs~~ in HA environment

SQL Interface

You can use SQL to create an encrypted table with the additional table option, 'ENCRYPT'. It'll look like the following, which could be changed in detail later.

```
CSQL> CREATE TABLE table_a (att1 INT, att2 VARCHAR(20)) ENCRYPT[=AES|ARIA];
```

진행하고 있는 개발 주제 확인

- JIRA에서 현재 IN PROGRESS 상태인 New Feature 이슈 확인

project = CBRD AND issuetype in ("New Feature", Improvement) AND status in ("In Progress", "REVIEW IN PROGRESS") ORDER BY summary ASC, created DESC

–10 of 10

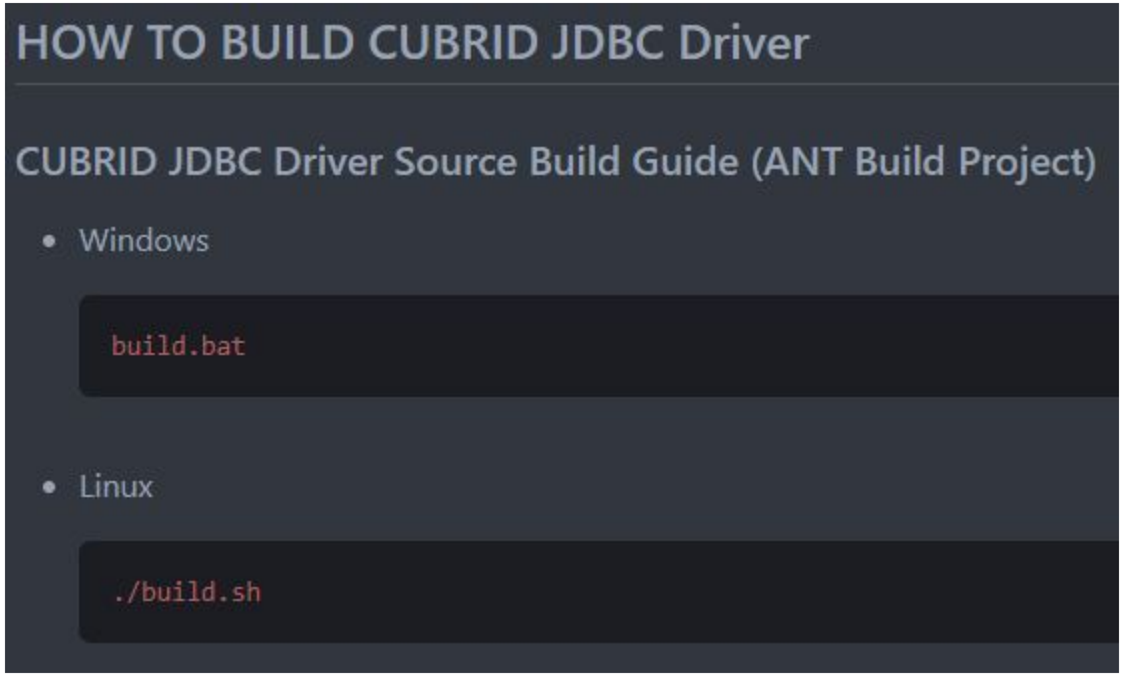
T	Key	Summary
	CBRD-23902	Clean up the unused parts related to grammer for Lex and Bison.
	CBRD-21614	Enhance logdump
	CBRD-24042	Enhance view transformation
	CBRD-24083	Even if only the column name is given in where, it evaluates to true or false.
	CBRD-20152	Implement overflow column
	CBRD-22111	Improve performance monitoring
	CBRD-23852	Performance degradation occurs due to the use of double-write-buffer when performing utilities such as createdb, addvoldb, loaddb, etc
	CBRD-24069	Prevent emergency_patch -r if the active volume is sane enough

View transformation이 뭐지?

Overflow column?

소스로부터 직접 빌드 해보기

- 프로젝트에 직접 참여해보기 위한 시작점
- 생각보다 이 지점에서 기여해볼 부분이 나오는 경우가 많아요



HOW TO BUILD CUBRID JDBC Driver

CUBRID JDBC Driver Source Build Guide (ANT Build Project)

- Windows

```
build.bat
```

- Linux

```
./build.sh
```

프로젝트 기여 포인트!

- How to build 문서가 없는 경우
- 필요한 라이브러리를 제대로 포함 하지 않은 경우
- 빌드 전 환경 설정이 불편한 경우

소스로부터 직접 빌드 해보기

- 최근 버그…

CUBRID APIs / APIS-883

Failed to build since setting ANT environment variable wrong

환경 변수를 잘못 설정해서 빌드가 실패 해요

(1) ANT_PATH has to be found from %ANT_FILE% variable defined at the beginning of the build.bat file.
(2) but, It tries to find the ant from invalid variable %ANT%.
(3) Also, even if ANT_HOME is set correctly, It tries to find ant at invalid ANT_FILE

```
set ANT_FILE =%ANT_HOME%\bin\ant   --- (1)

...

call :FINDEXEC ant ANT_PATH %ANT%  --- (2)
if "%ANT_PATH%" == "" (
  if "%ANT_HOME%" == "" (
    echo "[ERROR] set environment variable is required. (ANT Or ANT_HOME)"
    GOTO :EOF
  ) else (
    call :FINDEXEC ant ANT_PATH "%ANT_FILE%" --- (3)
  )
)
```

ANT_HOME이 아니라 ANT_PATH라고
오타가 난 것 같은데요?

CUBRID 프로젝트 구조

https://github.com/CUBRID

Engine

cubrid

cubrid-manager-server

Interface

cubrid-cci

cubrid-odbc

cubrid-php

cubrid-go

cubrid-python

…

Tool

cubrid-manager

cubrid-migration

QA / CI

cubrid-testcases

cubrid-testtools

cubridci

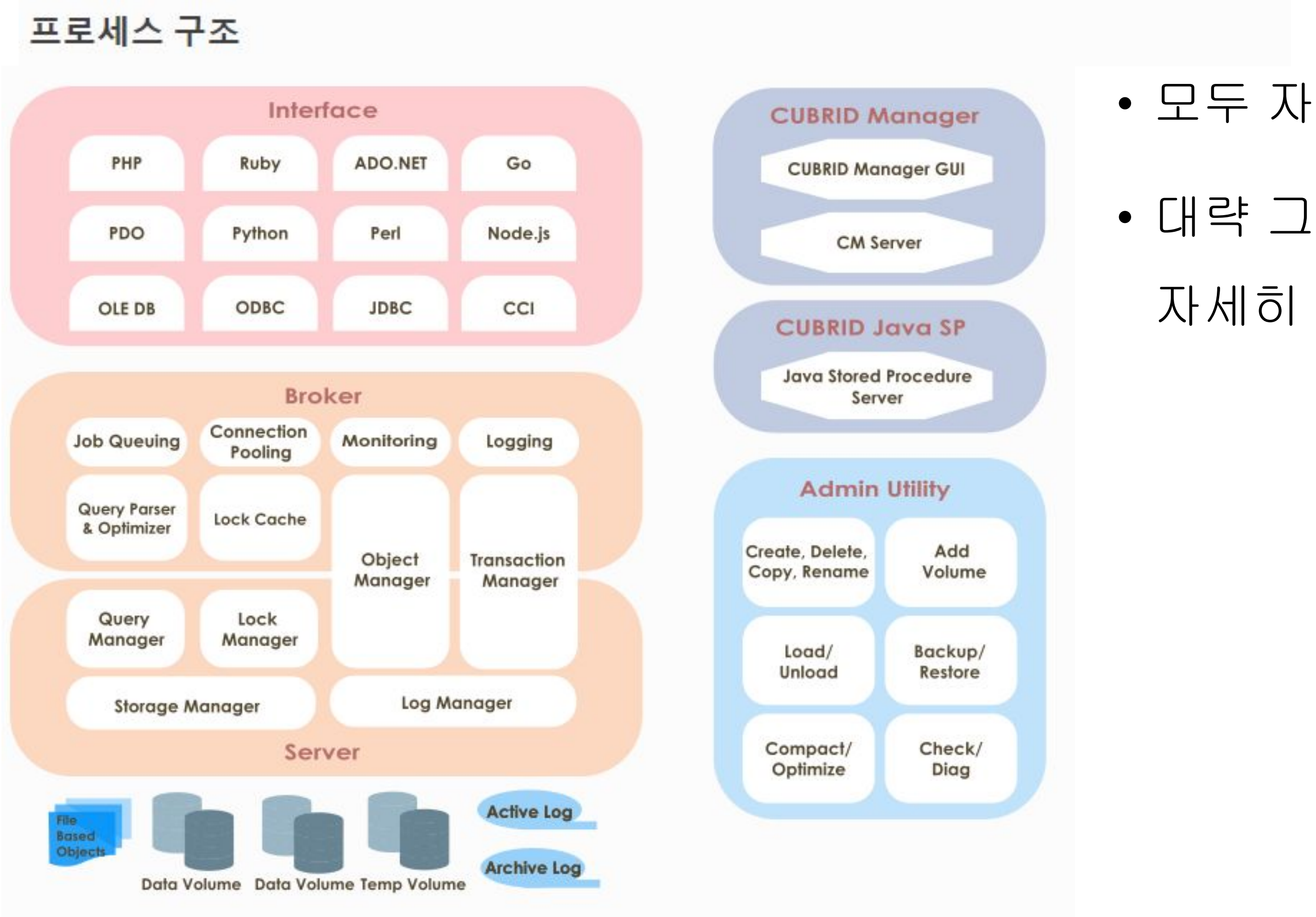
Culture

Document

cubrid-manual

CUBRID 시스템 아키텍처

cubrid.org/manual/ko/11.0/



- 모두 자세히 파악할 필요는 없어요
- 대략 그림만 파악해두고 직접 개발해볼 때 자세히 보면 되죠!

프로젝트에 참여해보기

- 커뮤니티 참여 가이드가 있다면 먼저 잘 읽어봅시다
- 작은 이슈부터 시도해보면서 전체적인 기여 프로세스를 직접 체화 해봅시다
- 이슈와 코드를 보고 기여할 수 있는 부분을 더 찾아봅니다
- 프로젝트의 테스트 프로그램을 실행해보면 프로젝트를 더 깊게 이해할 수 있어요

개발자 기여 가이드 - dev.cubrid.org

Contribution Guidelines

Contribution Guide (ko)

- Overview
- Preface

WORKFLOW WITH JIRA ISSUES

- Introduction
- JIRA 이슈 등록 <OPEN>
- 이슈 선별 (Triage) <Confirmed>
- 작업 진행 <In Progress>
- 작업 리뷰 <Review In Progress, Reviewed>
- 작업 완료 <Resolved>
- 이슈 종료 <Closed>

WORKFLOW WITH GITHUB

- Introduction
- Git Workflow
- Git 브랜치 모델

Overview

CUBRID는 오픈소스 프로젝트로, 일반적인 오픈소스 프로젝트 개발 프로세스를 기반으로 운영합니다. CUBRID의 개발자들은 오픈소스 개발 프로젝트를 위한 협업 도구를 이용해 CUBRID라는 하나의 큰 프로젝트에 함께 기여하고 있습니다.

이 문서는 **협업 도구를 활용한 개발 워크 플로우를 설명하고 개발 문화에 대한 상호 합의를 위한 기본 원칙 수립 목적**으로 작성되었습니다. 이 문서를 기반으로 상호 간의 합의로 상황에 따라 유연하게 적용하고, 문서를 수정/발전 시켰으면 합니다.

- 공정과 도구보다 개인의 상호작용을
- 포괄적인 문서보다 작동하는 소프트웨어를
- 계획에 따르기보다 변화에 대응하기를

CUBRID 개발 프로세스는 다음 두 도구를 기반으로 운영됩니다.

- Issue Tracking: JIRA (jira.cubrid.org)
- Code Repository and Code Review: Github (github.com/CUBRID)

CUBRID의 모든 프로젝트, 기능추가, 버그수정은 JIRA 이슈 생성으로부터 시작합니다. JIRA의 Issue workflow에 따라 문제를 정의, 분석하는 단계를 지나 설계 및 구현을 시작하게 됩니다. 설계 및 구현 단계에서는 Github에서 브랜치 생성, Pull Request, 코드 리뷰를 통해 검토/검증 과정을 거친 후 CUBRID에 코드를 반영합니다. 위 과정을 계속 반복하여 CUBRID를 계속 발전시킵니다.

- 개발 프로세스 전반에 대한 설명
- 각 단계 별 기여 가이드
  - JIRA 이슈 작성과 필수 내용
  - 이슈 상태 별 절차 정의
  - 코드 리뷰
  - 자동화 도구 설명
  - QA 리포트 확인과 테스트 실패 시 절차
  - 매뉴얼 작성 가이드
  - 신규 기여자 가이드

처음엔 코드 주석 실수, 간단한 로직 문제를 고치는 것도 좋아요

```diff
   * GNU General Public License for more details.
   *
   * You should have received a copy of the GNU General Public License
-  * along with this program; if not, write to the Free Softwarehttp://jira.cubrid.org/browse/CBRD-20692
+  * along with this program; if not, write to the Free Software
   * Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston, MA 02110-1301 USA
   *
   */
```

```diff
   char tmp[PATH_MAX];
   std::size_t len, suffix_len;

-  if (er_Msglog_filename != NULL)
+  if (er_Msglog_filename == NULL)
    {
      er_Accesslog_filename = NULL;
      return;
```

처음에 시도해 볼만한 방법

이미 구현된 기능과 비슷한 이슈를 찾아보고 참고해보기

예시)
REGEXP_REPLACE (‘Open Source’, ‘a|e|i|o|u’, ‘#’); --- Op#n S#urc#
기능이 구현된 PR을 발견했는데
REGEXP_COUNT () 함수가 없다?!

프로젝트 기여 포인트!

그냥 프로젝트 메인테이너 (커미터) 에게 물어보기

- 혹시 기여할 만한 작업이 있을까요?
- 이 기능에 관심이 있어서 기여해보고 싶은데 어떤 걸 보면 좋을까요?

비교적 간단하지만 시간이 없어서 해결하지 못한 이슈가 쌓여 있는 경우가 많음

대부분 이슈 생성 전 커뮤니케이션 채널이 있어요

CUBRID의 경우

버그/문제 토론, 질문, 정보 제공
(한글) cubrid.com, 자유게시판/Q&A
(영문) 레딧 CUBRID, https://www.reddit.com/r/CUBRID/

JIRA 작성

CUBRID / CBRD-23629

Create utility for Java SP server

Description

When the java_stored_procedure system parameter is on, Java SP server is only able to start with the database server (cub_server) together. So there was no way to restart the java stored procedure server without restarting the cub_server.

This work could be a preliminary task to add a system parameter to set JVM options for the Java se

Proposal

I propose to develop a process that acts as a service utility of CUBRID. users can manage a utility b

Usage example

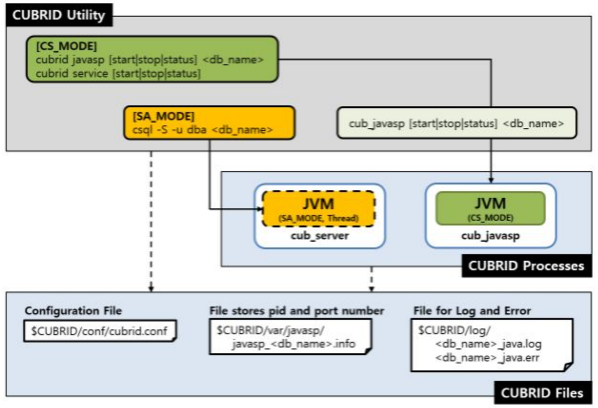
```
usage: cubrid javasp <command> [args]

Available command:
    start       [database-name]
    stop        [database-name]
    restart     [database-name]
    status      [database-name]
```

cubrid javasp status will show the following information

- The database name
- The process id
- The port number for which communicate with CAS
- Applied JVM options

Architecture Overview



The followings are added newly in the figure above:

- **cub_javasp**: It has the main routine for javasp server process. it processes commands given by cubrid utility.
- **$CUBRID/var/javasp/javasp_<db_name>.info**: a file to store pid and port number of running javasp server. When javasp server start, PID and port number are written in this file. From other commands, cubrid utility reads this file and try to connect with written information to the running javasp server and check the status of the server.
- **$CUBRID/log/<db_name>_java.err**: error message is written when error occurs in process of startup JVM.
- **Able to be managed by \`cubrid service\` command**: javasp server is a service utility so that users can register it in the service section in the conf file ($CUBRID/conf/cubrid.conf) and manage this utility by \`cubrid service\` command.

In Standalone mode (SA_MODE), it operates as a thread in cub_server as before. In CS_MODE, users can start and manage the javasp server with the cubrid utility or the cub_javasp utility.

테스트 도구 활용해보기

프로젝트에서 유닛 테스트/기능 테스트 도구가 있다면 활용해보세요

cubrid-testtools

This repository includes functional tools to test CUBRID.

CTP (CUBRID Test Program)

CTP is designed to execute all kinds of functional testings for CUBRID.

How to install CTP

Quick Start

See CTP/README.md for a quick start.

Functional Testing Guide

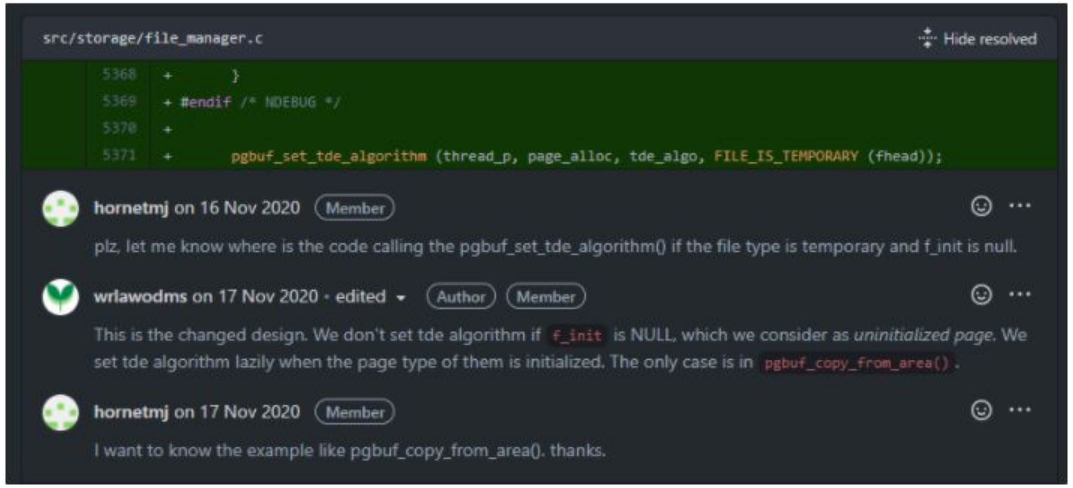
- SQL/MEDIUM/SQL_BY_CCI
- Shell for Linux
- Shell for Windows

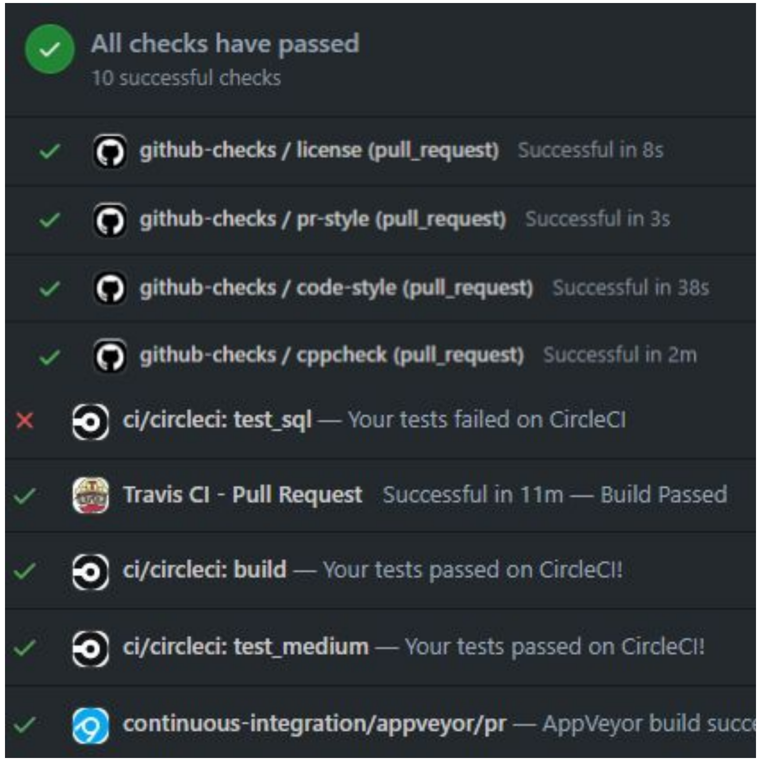
linux

Bits	Category	Total	Testing	Success	Fail Total	Fail New	Rate Test Rate	Rate Fail Rate	Elapse Time	Test Date	Unstable Case Total	Unstable Case New	Unstable Case Trend	Unstable Case Verified Rate
64bits	sql	15698	15698	15698	0	0	100%	0%	3012	2021-02-25 00:22	0	0	0	100%
	sql_debug	15698	15698	15698	0	0	100%	0%	10532	2021-02-25 00:23	0	0	0	100%
	medium	970	970	970	0	0	100%	0%	232	2021-02-25 00:21	0	0	0	100%
	medium_debug	970	970	970	0	0	100%	0%	860	2021-02-25 00:23	0	0	0	100%
	sql_by_cci	15698	15698	15697	1	1	100%	1.00%	2921	2021-02-25 01:10	1	1	+1	100%
	shell	2809	2781	2777	4	1	100%	0.14%	13777	2021-02-25 11:30	0	0	0	100%
	shell_debug	2809	2781	2772	9	6	100%	0.32%	26321	2021-02-25 07:39	6	6	+6	100%
	shell_heavy	100	100	99	1	0	100%	1.00%	24612	2021-02-26 15:10	0	0	0	100%
	shell_long	71	71	71	0	0	100%	0%	60658	2021-02-26 08:18	0	0	0	100%
	cci	321	321	321	0	0	100%	0%	15023	2021-02-25 09:28	0	0	0	100%
	cci_debug	321	321	321	0	0	100%	0%	17853	2021-02-25 05:17	0	0	0	100%
	ha_shell	323	323	320	3	1	100%	0.93%	14109	2021-02-25 04:58	1	1	+1	100%
	ha_repl	15698	15504	15503	1	1	100%	0.01%	32089	2021-02-26 01:52	1	1	+1	100%
	ha_repl_debug	15698	15504	15504	0	0	100%	0%	59437	2021-02-25 16:52	0	0	0	100%
	isolation	6770	6752	6752	0	0	100%	0%	1981	2021-02-25 05:31	0	0	0	100%
	isolation_debug	6770	6752	6752	0	0	100%	0%	2539	2021-02-25 01:02	0	0	0	100%

https://github.com/CUBRID/cubrid-testtools

걱정하지 마세요, 코드리뷰와 자동 테스팅 도구가 있으니까요





마무리

OPEN COMMUNITY!

CUBRID™

What's Next?

CUBRID 참여 시 프로젝트 이해를 더 도와줄 수 있는 내용

- 빌드 방법
- 디버깅/성능 측정 가이드
- 모듈 별 설명 자료
- 기술적인 내용
- …

감사합니다

과학기술정보통신부

nipa 정보통신산업진흥원
National IT Industry Promotion Agency