### streamjoin

입력으로 받는 스트림 데이터의 필드와 서브 쿼리 결과 필드를 비교해 조인(join)합니다.

#### 문법

streamjoin [OPTIONS] KEY\_FIELD, ... [ SUBQUERY ]

#### 필수 매개변수

**KEY\_FIELD, ...**

조인의 기준이 키 필드. 필드 구분자로 쉼표(,)를 사용합니다.

**[ SUBQUERY ]**

입력 데이터와 조인할 데이터를 출력하는 서브쿼리문을 대괄호 쌍([ ])으로 감싸서 입력

선택 매개변수

**timeout=INT{s}**

서브쿼리가 완료될 때까지 대기할 시간(기본값: 무한정 대기)

**type=inner|left|leftonly**

조인 유형(기본값: inner).

1. inner: 키가 일치하는 레코드만 결합해서 출력합니다. 키를 포함하지 않는 레코드는 출력하지 않습니다. 데이터의 교집합과 유사합니다. 일반적으로 조인이라 하면 "inner 조인"을 의미합니다.
2. left: 키가 일치하는 레코드는 결합해서 출력하고, 키가 일치하지 않는 레코드는 입력 데이터의 레코드만 출력합니다.
3. leftonly: 입력 레코드 집합 중에서 서브쿼리 결과 집합과 키가 일치하지 않는 레코드만 출력합니다. 키가 일치하는 레코드는 출력하지 않습니다.

#### 설명

streamjoin 명령어는 서브쿼리의 결과를 오프힙 메모리에 적재하고 해시 조인을 수행하므로 [join](https://docs.logpresso.com/ko/query/join-command) 명령어보다 속도가 빠르며, 스트림 쿼리에서도 사용할 수 있습니다. 하지만, inner와 left, leftonly 조인만 할 수 있고, 처리할 수 있는 데이터의 크기는 메모리 풀 용량에 따라 제한됩니다. 서브쿼리가 실패하면 **\_streamjoin\_fail** 필드에 예외 메시지를 추가합니다.

로그프레소를 실행할 때 아래 옵션을 지정하여 메모리 풀 크기를 조절할 수 있습니다. 기본값은 '500M'입니다. 다음 예시와 같이 메모리 풀 크기를 지정할 수 있습니다:-Dlogpresso.streamjoin.max\_buffer\_size=1G

메모리 사용 현황을 아래의 쿼리로 조회할 수 있습니다:

1. 메모리 풀 사용 현황: system memory pools
2. 쿼리별 메모리 사용 현황: system memory objects

#### 사용 예

데이터베이스에서 가져온 데이터와 **code** 필드를 키로 하여 조인

json "[ {'code':1}, {'code':2}, {'code':3} ]" | streamjoin code [ dbquery ora select code, description from tbl\_codes ]

데이터베이스에서 가져온 데이터와 **code** 필드를 키로 하여 조인. 단, SQL 쿼리를 10초로 제한

json "[ {'code':1}, {'code':2}, {'code':3} ]" | streamjoin timeout=10s code [ dbquery ora select code, description from tbl\_codes ]

#### 호환성

type 중에서 leftonly는 ENT-3.0.2003.0 버전부터 사용 가능합니다.