### 수집기

#### 개요

수집기는 데이터 원천에서 데이터를 수집하는데 사용됩니다. 수집기가 데이터를 수집하는 방식은 수집 모델에 정의된 유형에 따라 다를 수 있습니다. 데이터를 어떻게 수집할 것인지 결정하는 것은 수집 모델이고, 수집기는 수집 모델에 정의된 방식에 따라 데이터를 수집하는 역할을 수행합니다.



데이터 원천

데이터 원천의 대부분은 로그프레소 소나 외부에 존재하는 네트워크 호스트입니다. 데이터 원천은 웹과 같은 서비스를 제공하는 애플리케이션 서버이거나, 파일 서버, 데이터베이스일 수도 있고, 방화벽과 같은 네트워크 보안기기일 수도 있습니다. REST API를 통해 접근 가능한 서비스일 수도 있습니다. PCAP 장치도 데이터 원천이 될 수 있습니다.

로그프레소 소나 서버 자신이 데이터 원천이 되기도 합니다. 이벤트 연관 분석을 위해 메모리에 저장된 이벤트 컨텍스트 데이터, 로그프레소 소나 또는 센트리의 로그, 외부 데이터 원천으로부터 수집된 데이터에서 파생된 또 다른 데이터 등을 수집해야 하는 경우가 있을 수 있습니다.

수집 방식에 따른 구분

수집기는 데이터 원천에서 데이터를 수집하는 방식에 따라 패시브, 액티브로 구분됩니다.

**패시브 방식**

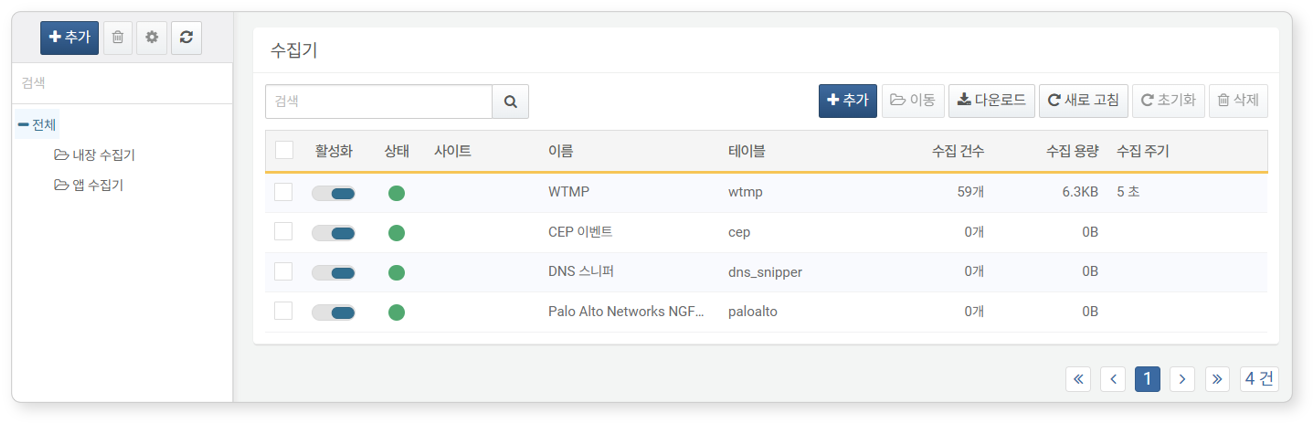
로그프레소 소나가 데이터 수신 포트를 열고, 데이터 원천이 전송하는 데이터를 수집기가 받아서 처리합니다. 위에 있는 그림에서 Syslog, SNMP Trap, NetFlow, sFlow와 같은 프로토콜로 데이터를 수집하는 수집기들이 패시브 방식입니다.

**액티브 방식**

일정한 간격 또는 지정된 일정에 따라 데이터 원천에 로그프레소 소나가 직접 접근해 데이터를 수집합니다. 위에 있는 그림에서 FTP, SFTP, SSH, HTTP, JDBC와 같은 프로토콜로 데이터 원천에 접속하는 수집기들이 액티브 방식입니다. 서비스가 제공하는 REST API를 통해 데이터를 수집하는 수집기도 액티브 방식입니다.

#### 수집기 목록 조회/검색

**수집 > 수집기**에서 수집기 목록을 조회하거나 검색할 수 있습니다. 로그프레소 소나의 초기 상태는 아무런 수집기도 제공하지 않으므로 빈 목록만 보여줍니다. 다음 그림은 **전체** 그룹의 수집기 목록 예시입니다. (기본 그룹인 **전체**는 모든 수집기 그룹에 속한 수집기와 어느 그룹에도 속하지 않은 수집기를 모두 보여줍니다.)



* **활성화**: 수집기의 활성화 토글 버튼
* **상태**: 수집기의 실행 상태(녹색: 정상, 회색: 비활성화됨, 주황: 수집 실패). 수집기를 실행하는 노드나 센트리에 아직 설정이 동기화되지 않았거나, 디스크 고갈 등의 이유로 수집기가 동작하지 않을 수 있습니다.
* **사이트**: 데이터 원천에 대한 사이트 정보
* **이름**: 수집기 식별에 사용하는 이름
* **테이블**: 수집된 데이터를 저장하는 테이블 이름
* **수집 건수**: 수집된 레코드의 개수
* **수집 용량**: 수집된 레코드의 용량(기본 단위: 바이트). 이 값은 테이블이 보여주는 디스크 사용량과 다를 수 있습니다.
* **수집 주기**: 액티브 수집기의 데이터 수집 주기 또는 일정. 패시브 수집기는 수집 주기가 없습니다.

수집기 목록에서 특정한 수집기를 찾으려면 도구 모음에 있는 검색 도구를 사용하세요. 검색 도구는 **이름**에 입력한 단어가 포함된 수집기를 찾아서 보여줍니다. 수집기 검색 도구는 대소문자를 구분하지 않습니다.

#### 수집기 그룹

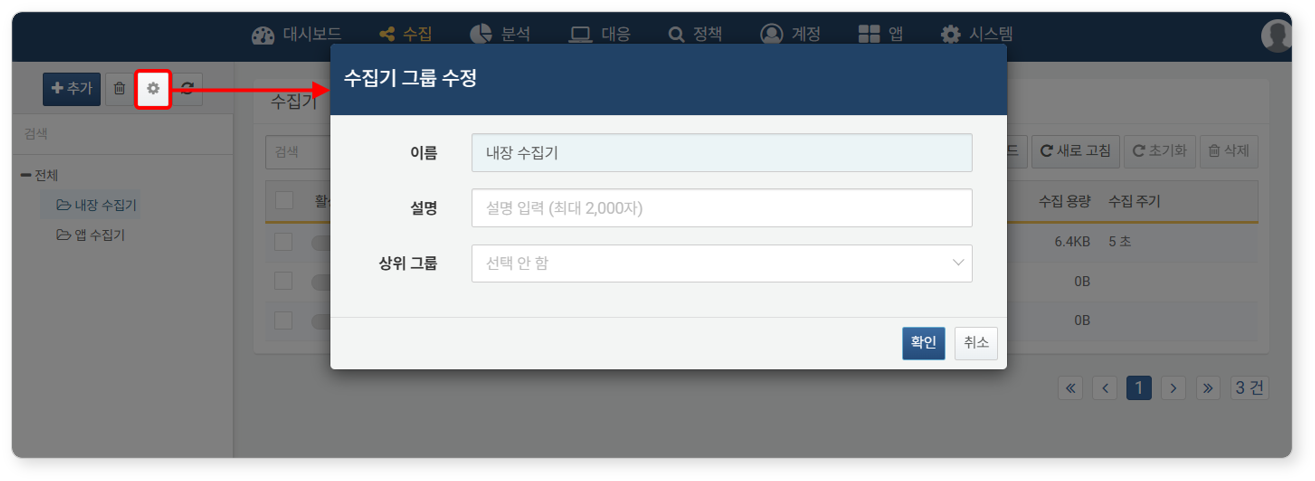
수집기를 그룹화하여 관리하려면 수집기 그룹을 사용하세요.

수집기 그룹을 추가하려면, **수집 > 수집기**에서 수집기 그룹 목록 위에 있는 **추가**를 클릭한 다음, 수집기 그룹의 **이름**, **설명**을 입력하고 **상위 그룹**을 선택한 다음 확인을 클릭하세요.

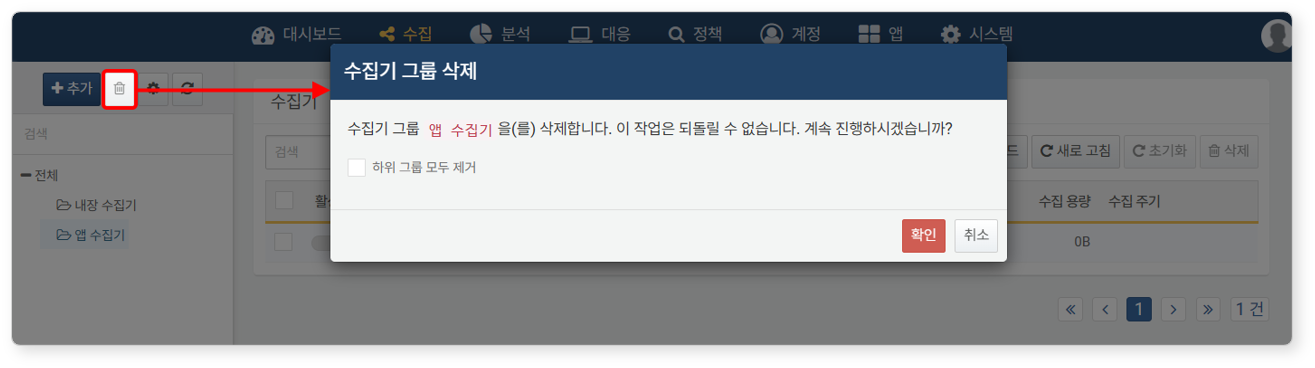


* **이름**: 수집기 그룹 식별에 사용할 이름
* **설명**: 수집기 그룹에 대한 설명
* **상위 그룹**: 추가할 수집기 그룹의 상위 그룹(기본값: 선택 안 함)

수집기 그룹을 수정하려면, 수집기 그룹 목록에서 수정할 수집기 그룹을 선택한 다음, 톱니바퀴 버튼을 클릭하세요.



수집기 그룹을 삭제하려면, 수집기 그룹 목록에서 삭제할 수집기 그룹을 선택한 다음, 휴지통 버튼을 클릭하세요. 수집기 그룹 아래에 있는 모든 그룹을 삭제하려면 **하위 그룹 모두 제거**를 선택하세요.



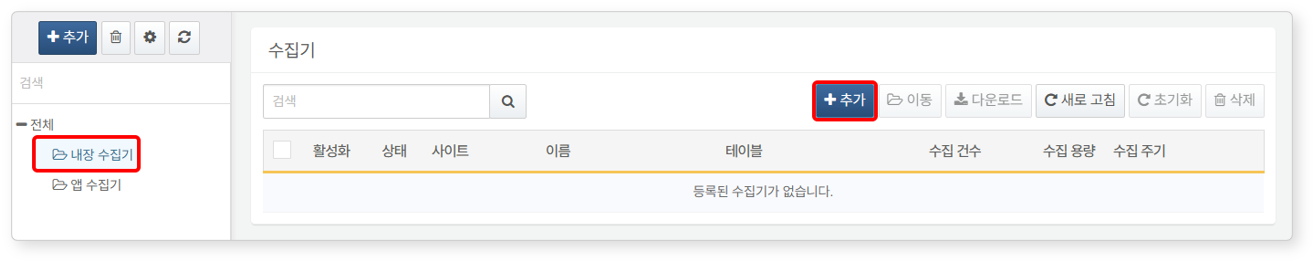
삭제할 수집기 그룹에 수집기가 포함되어 있으면 소속 그룹만 제거되고, 수집기는 소속 그룹이 없는 상태가 됩니다.

#### 수집기 추가

수집기는 파서나 로그 스키마, 수집 모델과 달리 수집할 데이터 원천에 맞게 관리자가 직접 생성해야 합니다. 수집기를 추가하려면,

**수집 > 수집기**에서 수집기를 추가할 수집기 그룹을 선택하세요. 수집기 그룹을 선택하지 않으면 추가할 수집기는 소속 그룹 없이 생성됩니다.

도구 모음에서 **추가**를 클릭하세요.



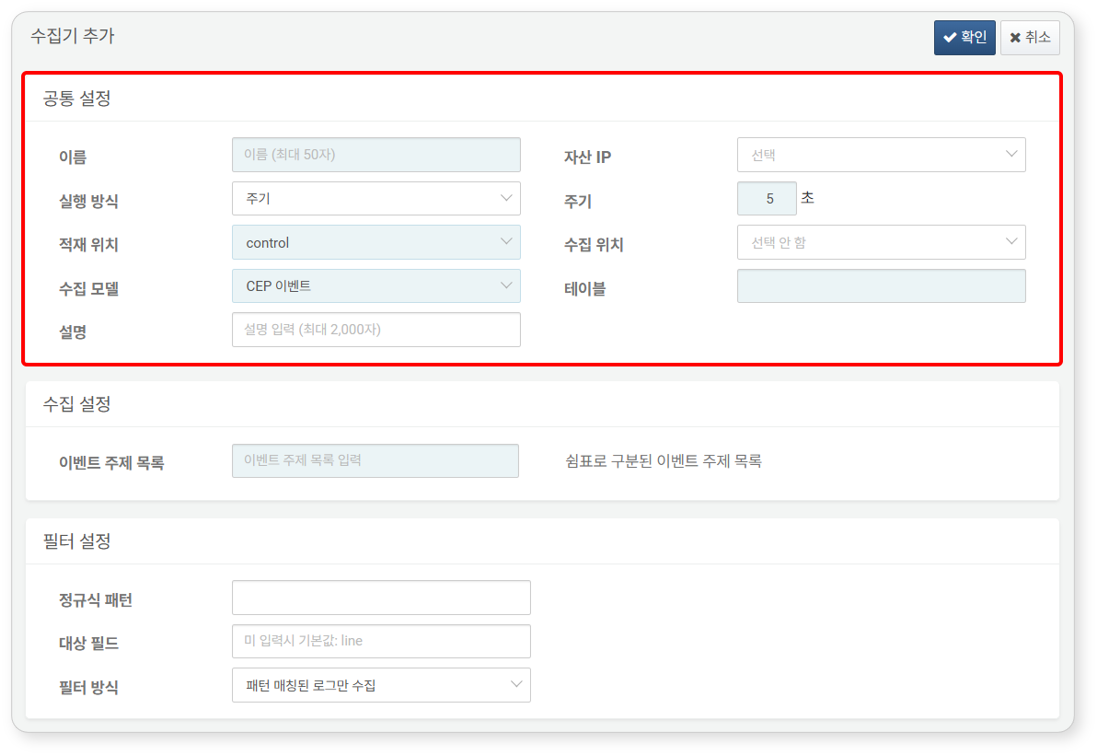
**수집기 추가** 화면에서 [공통 설정](https://docs.logpresso.com/ko/sonar/4.0/ui/section-loggers), [수집 설정](https://docs.logpresso.com/ko/sonar/4.0/ui/section-loggers), [필터 설정](https://docs.logpresso.com/ko/sonar/4.0/ui/section-loggers)에 필요한 값을 입력하거나 선택한 다음, **확인**을 클릭하세요. 설정 항목 중 배경색이 하늘색인 것은 필수 입력 항목입니다.

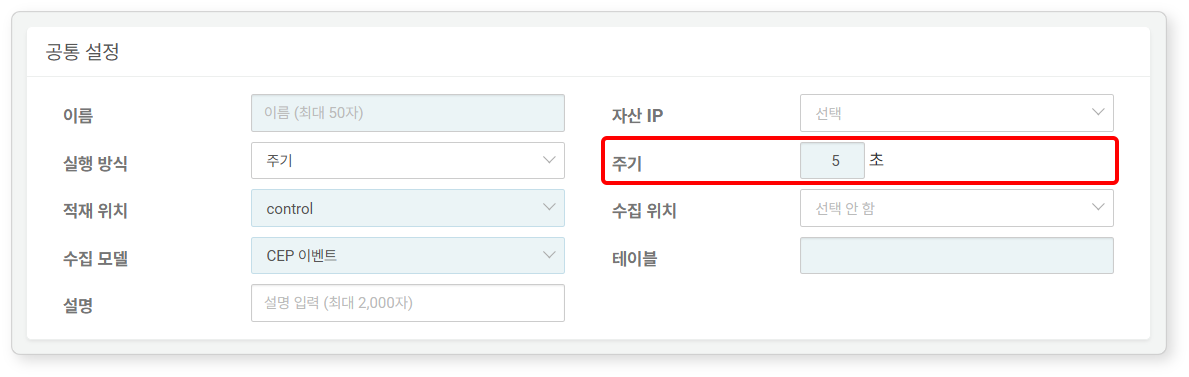
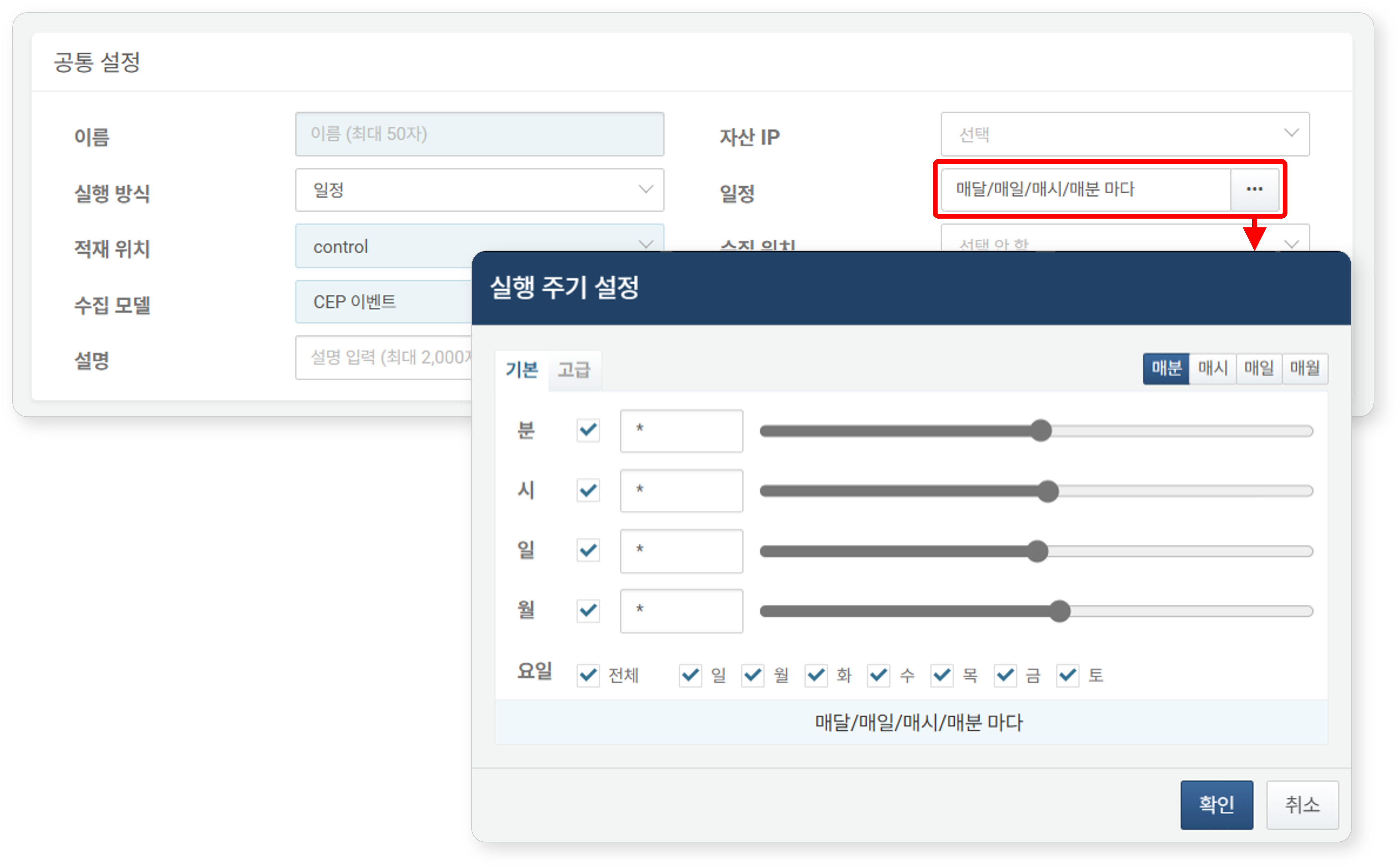
* **공통 설정**: 수집기를 구성하는데 필요한 일반 설정 항목
* **수집 설정**: 공통 설정에서 선택한 **수집 모델**에 따라 지정해야할 수집 방식 설정
* **필터 설정**: 수집된 원본 데이터에서 수집하거나 버릴 로그를 필터링할 때 적용할 정규표현식 패턴

추가된 수집기의 기본 상태는 비활성화입니다. 수집기를 활성화하려면 추가한 수집기의 활성화 버튼을 클릭해 활성화하세요.

공통 설정

공통 설정은 수집기의 일반 설정 항목들로, 수집기의 데이터 수집 주기/일정, 수집기를 실행할 노드, 수집기가 수집한 데이터를 보관할 노드와 테이블, 수집한 데이터를 정규화하는 과정에 적용할 수집 모델로 구성됩니다. 그 외에 보안정책 운영에 필요한 사이트 및 자산 IP 주소 정보와 연결하는 기능을 제공합니다.



* **이름**: 수집기를 식별하는데 사용할 이름
* **자산 IP**: 데이터 원천에 대한 사이트 및 자산 IP 주소 정보. 이 항목은 검색을 지원합니다.
* **실행 방식**: 액티브 수집기가 데이터를 수집할 **주기** 또는 **일정**(기본값: 주기)
* **주기**: 초 단위로 지정되는 수집기 실행 주기(기본값: 5초)
* 
* **일정**: CRON 문법 형식으로 지정되는 수집기의 실행 일정(기본값: 매달/매일/매시/매분 마다)
* 
* 실행 방식은 액티브 수집기에 적용됩니다. 패시브 수집기의 동작에는 영향을 미치지 않습니다.
* **적재 위치**: 수집된 데이터를 적재할 노드
* 단독실행형 로그프레소 소나인 경우, 사용자가 접속한 노드의 이름이 표시됩니다.
* [클러스터](https://docs.logpresso.com/ko/sonar/4.0/ui/section-cluster-nodes) 구성인 경우, 수집 노드에 해당하는 노드를 선택하세요. 수집 노드가 이중화되어 있으면 두 노드에 동시에 적재됩니다.
* **수집 위치**: 수집기를 실행할 센트리 또는 노드(기본값: 선택 안 함). 센트리는 로그프레소 소나 서버에 접속되어 있는 경우에만 목록에 표시됩니다. 이 항목은 검색을 지원합니다.
* **수집 모델**: 수집된 데이터 처리에 적용할 수집 모델
* [수집 모델](https://docs.logpresso.com/ko/sonar/4.0/ui/section-logger-models)에서 조회할 수 있는 수집 모델 중에서, **수집 위치**에서 선택한 전달 노드 또는 센트리에 적용할 수 있는 수집 모델만 목록에 나타납니다.
* 여기서 지정한 수집 모델에 따라 [수집 설정](https://docs.logpresso.com/ko/sonar/4.0/ui/section-loggers) 항목이 결정됩니다.
* 이 항목은 검색을 지원합니다.
* **테이블**: 파싱과 정규화를 거친 데이터가 저장될 테이블(기본값: 없음). [테이블](https://docs.logpresso.com/ko/sonar/4.0/ui/section-table-management)에서 추가한 테이블 이름을 입력하세요. 입력한 테이블이 없으면 지정된 이름으로 새 테이블을 추가해 사용합니다.
* **설명**: 수집기에 대한 설명

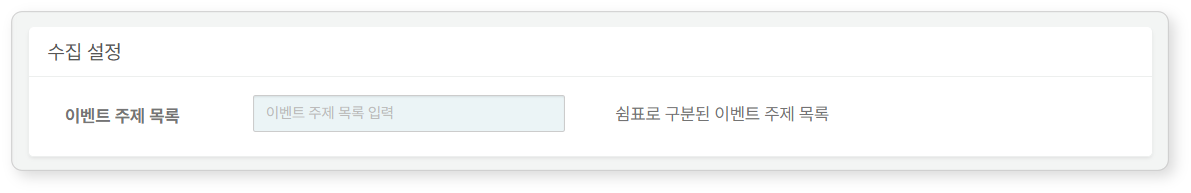
수집 설정

수집 설정은 [공통 설정](https://docs.logpresso.com/ko/sonar/4.0/ui/section-loggers)에서 지정한 **수집 모델**의 [유형](https://docs.logpresso.com/ko/sonar/4.0/ui/section-logger-models)에 따라 정의됩니다.

여기서 설명하는 내용은 모두 로그프레소 소나의 기본 수집기들을 기준으로 합니다. 그러나 앱을 설치하면 수집에 필요한 수집 모델이 함께 제공되므로, 앱 기반 수집기를 설정하는 경우가 대부분입니다. 앱 수집기들의 수집 설정은 기본 수집기의 수집 설정과 대부분 유사하므로 앱 수집기를 설정할 때 참고하세요.

CEP 이벤트

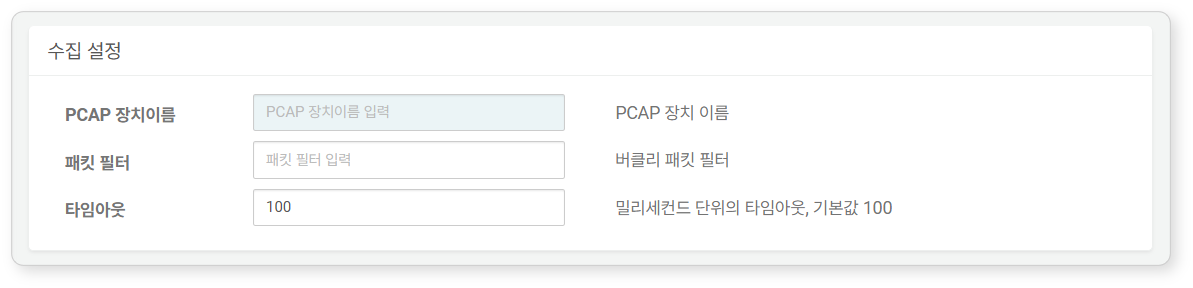
CEP(Complex Event Processing) 이벤트 데이터를 수집합니다.



* **이벤트 주제 목록**: 쉼표(,)를 구분자로 하는 이벤트 주제(topic) 목록. 주제는 인-메모리 데이터베이스의 테이블 이름과 같은 역할을 합니다.

DNS 스니퍼

PCAP 장치를 통해 미러링된 통신 패킷 중에서 DNS 통신 패킷을 수집합니다.



* **PCAP 장치이름**: 패킷 데이터를 수집할 PCAP 장치. [system pcapdevices](https://docs.logpresso.com/ko/sonar/4.0/ui/ko/query/system-pcapdevices-command) 명령으로 PCAP 장치 목록의 name 필드를 확인하고 입력하세요.
* **패킷 필터**: DNS 패킷 캡처에 적용할 필터 규칙. [PCAP 필터 문법](https://docs.logpresso.com/ko/sonar/4.0/ui/(https:/www.tcpdump.org/manpages/pcap-filter.7.html))을 참고해서 작성하세요.
* **타임아웃**: 타임아웃 시간(기본값: 100 밀리초)

FTP 디렉터리 와처

네트워크 호스트에 FTP 통신으로 접속하여 지정된 디렉터리에서 패턴과 이름이 일치하는 파일을 수집합니다.

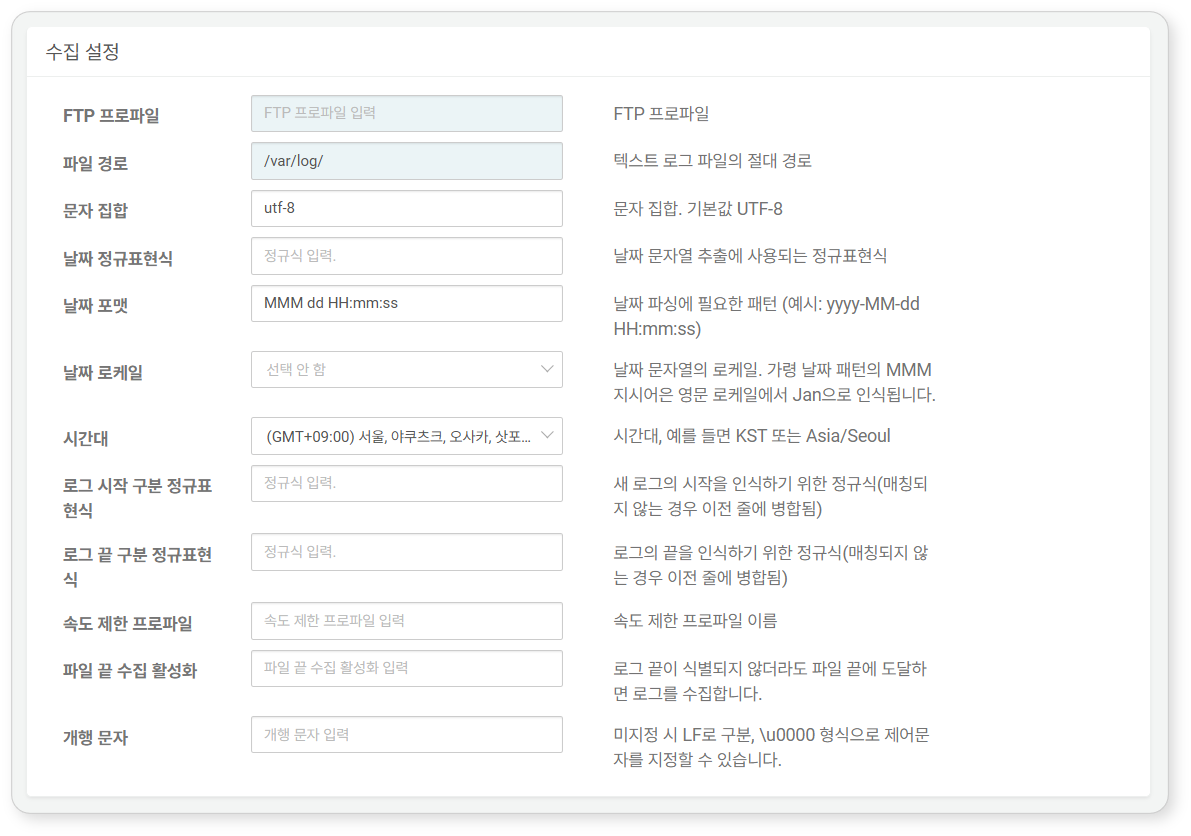


* **FTP 프로파일**: 네트워크 호스트의 FTP 연결에 필요한 [접속 프로파일](https://docs.logpresso.com/ko/sonar/4.0/ui/section-connect-profile) 이름
* **디렉터리 경로**: 파일을 수집할 대상 디렉터리의 절대 경로(예: /var/log/foo)
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식(예: bar\*.gz)
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 구분 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 구분 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **문자 집합**: 텍스트 파일에 적용할 문자 인코딩 방식(미지정 시 기본값: utf-8)
* **속도 제한 프로파일**: FTP 속도 제한 프로파일. 속도 제한 프로파일은 로그프레소 셸에서 logpresso.createFTPRateLimit 명령으로 생성할 수 있습니다.
* **파일이름 태그**: 수집한 파일 이름을 기록할 필드 이름
* **경로 태그**: 디렉터리 경로를 기록할 필드 이름
* **경로 날짜 추출 포맷**: 디렉터리 경로 문자열에서 날짜 문자열 검색에 사용할 정규표현식. 예를 들어, 디렉터리 경로에 날짜 정보(예: /var/log/foo/2022-11-02)가 있으면 디렉터리 경로 문자열에서 날짜 정보를 추출해 파일의 수집 기준 날짜로 사용합니다.
* **모니터링 대상 일수**: 파일 경로에서 날짜를 추출한 후 모니터링 대상 일수 이내에 있는 파일만 수집합니다(기본값: 없음(기간 제한 없이 파일에 있는 로그 수집)).
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.

FTP 로테이션 로그 파일

네트워크 호스트에 FTP 통신으로 접속하여 로테이션되는 로그 파일을 수집합니다.

흔히 유닉스나 리눅스 서버에서는 일정 주기로 로그 파일의 이름을 변경하여 백업하고 로그 파일을 새로 생성합니다. 그러한 파일을 로테이션 로그라 합니다.



* **FTP 프로파일**: 네트워크 호스트의 FTP 연결에 필요한 [접속 프로파일](https://docs.logpresso.com/ko/sonar/4.0/ui/section-connect-profile) 이름
* **파일 경로**: 수집할 로그 파일의 절대 경로(예: /var/log/secure)
* **문자 집합**: 파일에 적용할 문자 인코딩 방식(미지정 시 기본값: utf-8)
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 구분 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 구분 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **속도 제한 프로파일**: FTP 속도 제한 프로파일. 속도 제한 프로파일은 로그프레소 셸에서 logpresso.createFTPRateLimit 명령으로 생성할 수 있습니다.
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.

FTP 일자별 디렉터리

네트워크 호스트에 FTP 통신으로 접속하여 하루 단위로 생성되는 디렉터리를 순회하면서 파일 이름 패턴과 일치하는 모든 로그 파일을 수집합니다.



* **FTP 프로파일**: 네트워크 호스트의 FTP 연결에 필요한 [접속 프로파일](https://docs.logpresso.com/ko/sonar/4.0/ui/section-connect-profile) 이름
* **모니터링 기간**: 실시간으로 파일 변화를 모니터링할 기간(기본값: 0). 값이 0일 때, 실시간 수집을 비활성화합니다.
* **디렉터리 경로**: 파일을 수집할 대상 디렉터리의 절대 경로(예: /var/log/foo)
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식(예: bar\*.gz)
* **디렉터리 날짜 정규표현식**: 디렉터리 경로에서 날짜 문자열 검색에 사용할 정규표현식. 디렉터리 경로에 날짜 정보(예: /var/log/foo/2022-11-02)가 있으면 디렉터리 경로 문자열에서 날짜 정보를 추출해 사용합니다.
* **디렉터리 날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(미지정 시 기본값: yyyyMMdd)
* **디렉터리 날짜 사용 유무**: 디렉터리 날짜 포맷을 이용한 로그 생성 날짜 생성 여부(기본값: false)
* **과거 로그 수집 시작 일자**: yyyyMMdd 형식으로 시작 일자를 입력. 지정하지 않으면 과거 로그 수집이 비활성화됩니다.
* **과거 로그 수집 끝 일자**: yyyyMMdd 형식으로 끝 일자를 입력. 지정하지 않으면 과거 로그 수집이 비활성화됩니다.
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 구분 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 구분 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **문자 집합**: 파일에 적용할 문자 인코딩 방식(미지정 시 기본값: utf-8)
* **파일네임 태그**: 수집한 파일 이름을 기록할 필드 이름
* **속도 제한 프로파일**: FTP 속도 제한 프로파일. 속도 제한 프로파일은 로그프레소 셸에서 logpresso.createFTPRateLimit 명령으로 생성합니다.
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.

GZIP 디렉터리 와처

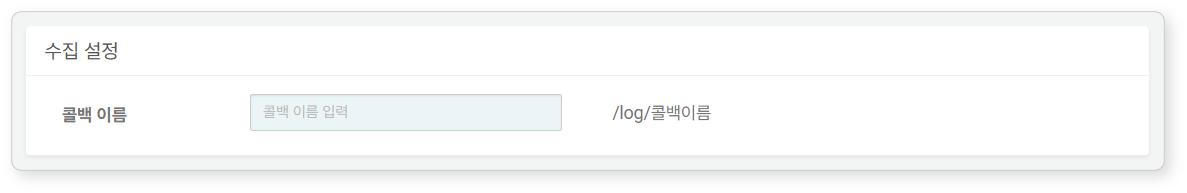
수집 위치로 지정된 노드에서 GZIP 형식으로 압축된 텍스트 로그 파일을 수집합니다.



* **디렉터리 경로**: GZIP 파일을 수집할 대상 디렉터리의 절대 경로(예: /var/log/foo)
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식(예: bar\*.gz)
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **문자 집합**: GZIP 압축을 해제한 후 텍스트 파일에 적용할 문자 인코딩 방식(미지정 시 기본값: utf-8)
* **GZIP 파일 삭제**: 수집 완료된 GZIP 파일의 삭제 여부를 나타내는 불리언 값(미지정 시 기본값: false). true를 입력하면 파일을 수집한 후 디렉터리에서 GZIP 파일을 삭제합니다.
* **파일이름 태그**: 수집한 파일 이름을 기록할 필드 이름
* **경로 태그**: 디렉터리 경로를 기록할 필드 이름
* **경로 날짜 추출 포맷**: 디렉터리 경로 문자열에서 날짜 문자열 검색에 사용할 정규표현식. 예를 들어, 디렉터리 경로에 날짜 정보(예: /var/log/foo/2022-11-02)가 있으면 디렉터리 경로 문자열에서 날짜 정보를 추출해 GZIP 파일의 수집 기준 날짜로 사용합니다.
* **모니터링 대상 일수**: 파일 경로에서 날짜를 추출한 후 모니터링 대상 일수 이내에 있는 파일만 수집합니다(기본값: 없음(기간 제한 없이 파일에 있는 로그 수집)).
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 LF(\n) 대신 CRLF(\r\n)의 사용 여부를 나타내는 불리언 값(미지정 시 기본값: false)

HTTP POST

네트워크 호스트가 로그프레소 소나에게 HTTP POST 메소드로 전송하는 로그를 수집합니다.

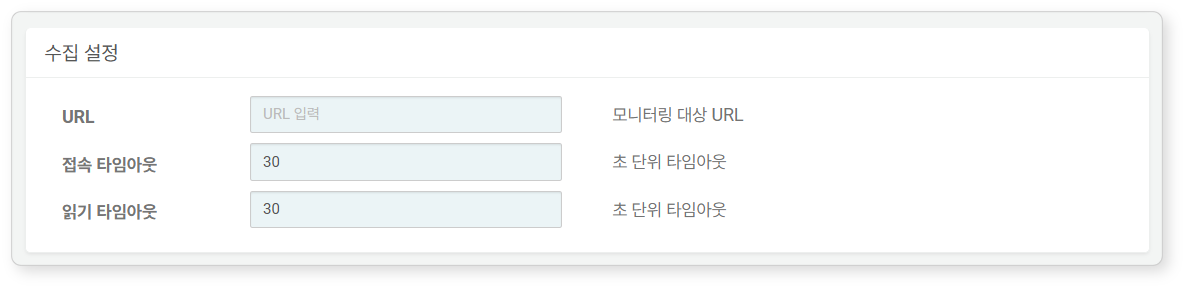


* **콜백 이름**: 네트워크 호스트가 HTTP POST 메시지를 보낼 때 사용할 콜백 주소. 콜백 이름의 형식은 /log/콜백이름입니다.

HTTP POST 수집기는 **수집 위치**에서 센트리를 지정할 수 없습니다.

HTTP 모니터

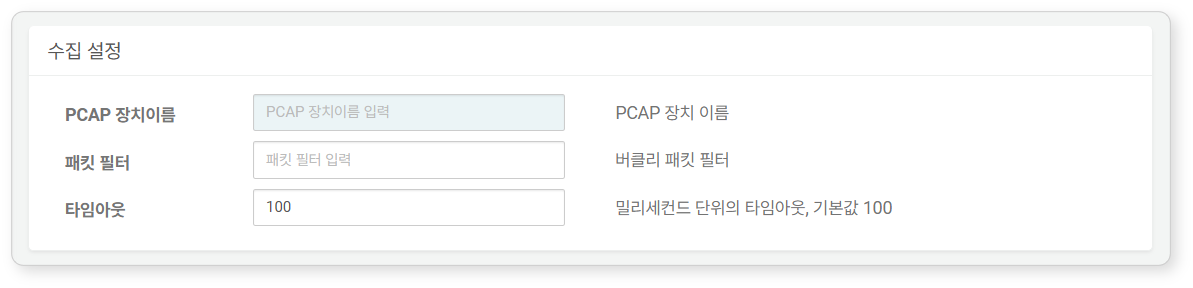
네트워크 호스트에 주기적으로 HTTP 통신으로 접속하여 웹 서비스 응답 코드와 통계 정보를 수집합니다.



* **URL**: 모니터링 대상 URL
* **접속 타임아웃**: HTTP 접속에 적용할 타임아웃(기본값: 30초)
* **읽기 타임아웃**: HTTP 접속 후 서버의 응답에 적용할 타임아웃(기본값: 30초)

HTTP 스니퍼

PCAP 장치를 통해 미러링된 통신 패킷 중에서 HTTP 통신 패킷을 수집합니다.



* **PCAP 장치이름**: 패킷 데이터를 수집할 PCAP 장치. [system pcapdevices](https://docs.logpresso.com/ko/sonar/4.0/ui/ko/query/system-pcapdevices-command) 명령으로 PCAP 장치 목록의 **name** 필드를 확인하고 입력하세요.
* **패킷 필터**: HTTP 패킷 캡처에 적용할 필터 규칙. [PCAP 필터 문법](https://docs.logpresso.com/ko/sonar/4.0/ui/(https:/www.tcpdump.org/manpages/pcap-filter.7.html))을 참고해서 작성하세요.
* **타임아웃**: 타임아웃 시간(기본값: 100 밀리초)

JDBC 수집기

SQL 쿼리를 사용하여 관계형 데이터베이스의 테이블이나 뷰에서 데이터를 수집합니다.



* **JDBC 프로파일**: 원격 호스트 연결에 사용할 JDBC 접속 프로파일
* **SQL**: 데이터 수집에 사용할 SQL문. $where 매크로와 날짜 매크로($yyyy(4자리 연도), $yy(2자리 연도), $MM(월), $dd(일), $HH(시))를 사용할 수 있습니다. 날짜 매크로는 일자에 따라 수집대상 데이터베이스 이름이나 테이블 이름이 달라지는 경우 사용하세요.
* **조건절**: $where 매크로에 삽입할 조건절. 물음표(?)를 위치 지정자 (placeholder)로 사용할 수 있습니다. 위치 지정자는 마지막 기준 컬럼 값으로 대치됩니다. 입력할 조건절은 where 문자열을 포함해야 합니다.
* **첫번째 기준 열**: 매 조회 시 마지막으로 수집했던 행 이후부터 가져올 수 있도록 검색 기준이 되는 첫 번째 컬럼 이름. 가령 시퀀스, IDENTITY, auto\_increment로 지정된 컬럼, 혹은 증가하는 타임스탬프 컬럼의 이름을 입력하세요.
* **두번째 기준 열**: 매 조회 시 마지막으로 수집했던 행 이후부터 가져올 수 있도록 검색 기준이 되는 두번째 컬럼 이름
* **첫번째 기준 열 초기 값**: 첫 번째 쿼리 시 조건절에 들어갈 초기 값. 미입력 시 첫 번째 쿼리는 **조건절**을 무시하고 수행됩니다.
* **첫번째 기준 열 초기 값 타입**: 기준 열 초기 값의 데이터 타입(string, integer, long, date 중 하나)
* **두번째 기준 열 초기 값**: 첫 번째 쿼리 시 조건절에 들어갈 초기 값. 미입력 시 첫번째 쿼리는 **조건절**을 무시하고 수행됩니다.
* **두번째 기준 열 초기 값 타입**: 기준 열 초기 값의 데이터 타입(string, integer, long, date 중 하나)
* **날짜 컬럼**: 데이터가 수집된 날짜를 가지고 있는 컬럼의 이름. 미입력 시 수집기가 수집한 날짜를 저장합니다.
* **날짜 포맷**: 날짜 컬럼이 SQL 시간 형식이 아닌 문자열에 적용할 날짜 형식(예: yyyyMMdd HH:mm:ss)
* **데이터베이스 이름**: 날짜 매크로($yyyy(연), $MM(월), $dd(일), $HH(시))를 포함한 데이터베이스 이름. 날짜 매크로는 일자에 따라 수집대상 데이터베이스 이름이나 테이블 이름이 달라지는 경우 사용합니다.
* **날짜 매크로 시작 날짜**: (날짜 매크로를 사용하는 경우) 시작 날짜
* **수집 반복 횟수 제한**: 수집 주기마다 실행할 수 있는 최대 SQL 쿼리 개수

기준 열 초기 값을 지정하지 않으면 수집기가 처음 수행할 때는 조건절을 무시하고 SQL 문을 실행합니다. 이는 기존 데이터를 모두 가져올 수 있게 하기 위한 것입니다. 모든 로그를 가져오지 않고 특정 범위의 로그만 가져오고 싶을 경우 기준 열 초기 값을 설정해야 합니다.

SQL을 작성할 때 성능을 고려하여 다음과 같이 하는 것이 좋습니다.

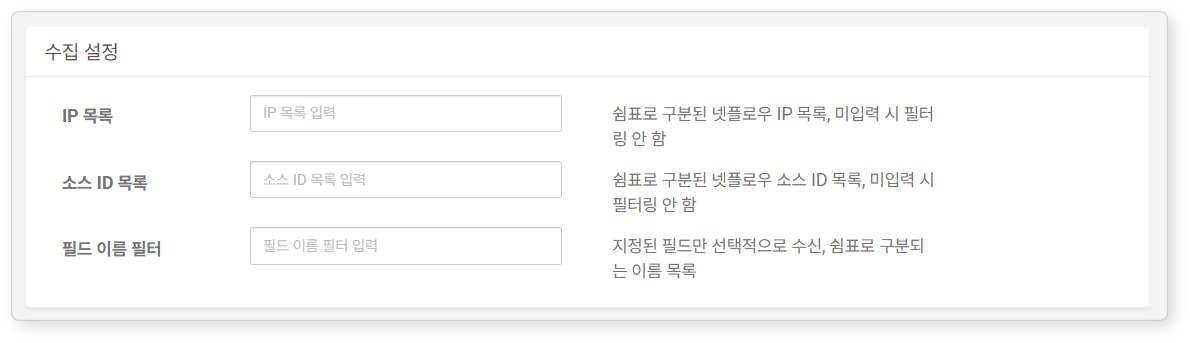
* 기준 열로 검색할 때 인덱스를 사용하는지 확인하세요. 인덱스를 사용하지 않으면 수천만 개 이상의 레코드가 있는 테이블에서 지속적으로 테이블 전체를 검색하므로 과부하를 발생시킵니다.
* 한 번에 가져올 레코드 개수를 제한하세요. 가령, 오라클의 경우 rownum을 사용하여 가져올 레코드 개수를 제한할 수 있습니다. 이를 고려하지 않으면 초기 적재 시에 너무 많은 레코드를 한 번에 가져오려고 시도하면서 문제가 발생할 수 있습니다. JDBC 수집기는 지정된 주기별로 쿼리를 수행하지만, 한 번 수집할 때 더 이상 새로운 값이 없을 때까지 쿼리를 반복 수행하므로 가져오는 레코드 개수를 제한하는 것이 좋습니다.
* 필요한 열만 SELECT 절에 명시적으로 지정하세요. JDBC 수집기는 조회되는 모든 열 값을 키/값 형태로 수집합니다. 불필요한 열을 제외하면 더 나은 성능을 기대할 수 있습니다.

SQL문에 사용할 수 있는 날짜 매크로와 관련하여,

* 일자에 따라 수집대상 데이터베이스 이름이나 테이블 이름이 달라지는 경우 날짜 매크로를 사용하세요.
* 날짜 매크로는 4자리 연도($yyyy),2자리 연도()$yy), 월($MM), 일($dd), 시($HH) 매크로를 지원합니다.
* 날짜 매크로 사용 시 수집 범위는 **날짜 매크로 시작 날짜**부터 현재 시간까지입니다. 다음 날 테이블을 미리 생성해두더라도 미리 수집되거나 오늘 일자 데이터 수집을 건너뛰지 않습니다.
* 날짜 매크로를 지정하면 최소 매크로 단위가 1씩 증가하면서 수집됩니다. 최소 매크로 단위가 $dd이면 1일씩, $HH면 한 시간씩 수집됩니다.
* 예를 들어, 데이터베이스 이름 항목에 db\_log\_$yyyy\_$MM을, SQL문 from 절에 db\_user\_$dd를, **날짜 매크로 시작 날짜** 항목에 2017-01-01 00:00:00을 입력하면 JDBC 수집기는 db\_log\_2017\_01.db\_user\_01, db\_log\_2017\_01.db\_user\_02 순으로 테이블들을 조회해서 현재 일자 테이블까지 데이터를 수집합니다.

NetFlow

네트워크 디바이스로부터 NetFlow v5/v9 네트워크 모니터링 정보를 수신합니다. NetFlow 수집기를 사용하려면 먼저 로그프레소 셸에서 lowmon.open netflow 2055 명령을 실행해 통신 포트를 개방하세요.



* **IP 목록**: 쉼표(,)를 구분자로 하는 NetFlow IP 주소 목록. 미입력 시 IP 주소를 필터링하지 않습니다.
* **소스 ID 목록**: 쉼표(,)를 구분자로 하는 NetFlow 소스 ID 목록. 미입력 시 필터링하지 않습니다.
* **필드 이름 필터**: 쉼표(,)를 구분자로 하는 필드 이름 목록. 지정된 필드만 선택적으로 수신합니다(기본값: 미지정).

PCAP 디렉터리 와처

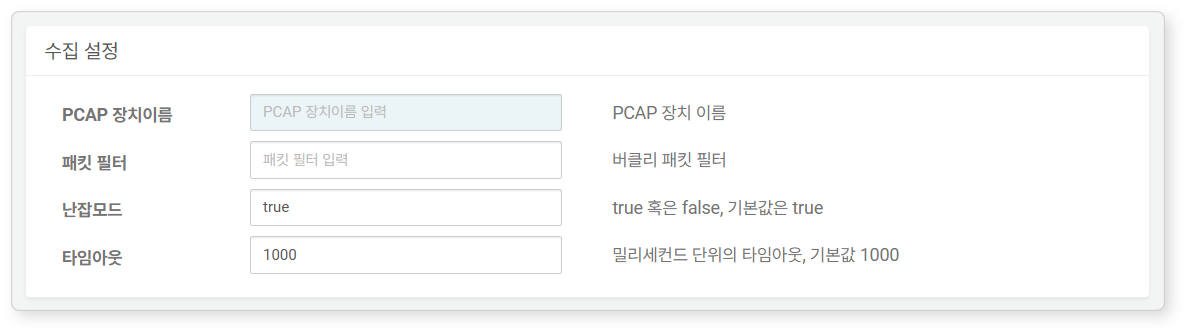
로컬 디렉터리에서 PCAP 파일을 수집합니다.



* **디렉터리 경로**: 파일을 수집할 로컬 디렉터리의 절대 경로
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식
* **로깅 모드**: 데이터 수집 모드(기본값: L4). 현재는 L4만 지원합니다.
* **파일이름 태그**: 수집한 파일 이름을 기록할 필드 이름
* **경로 태그**: 디렉터리 경로를 기록할 필드 이름
* **경로 날짜 추출 포맷**: 디렉터리 경로에서 날짜 문자열 검색에 사용할 정규표현식. 디렉터리 경로에 날짜 정보(예: /var/log/foo/2022-11-02)가 있으면 디렉터리 경로 문자열에서 날짜 정보를 추출해 사용합니다.
* **모니터링 대상 일수**: 디렉터리 경로에서 날짜를 추출한 후 모니터링 대상 일수 이내에 있는 파일만 수집합니다(기본값: 0(기간 제한 없이 파일에 있는 로그 수집))

PCAP 패킷

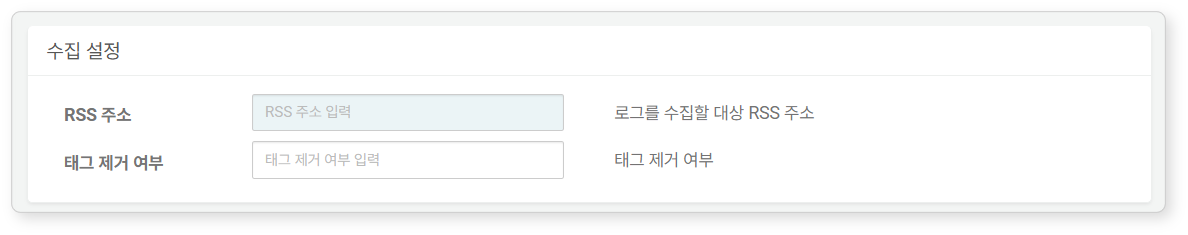
로컬 PCAP 장치로부터 패킷을 수집합니다.



* **PCAP 장치이름**: 패킷 데이터를 수집할 PCAP 장치. [system pcapdevices](https://docs.logpresso.com/ko/sonar/4.0/ui/ko/query/system-pcapdevices-command) 명령으로 PCAP 장치 목록의 name 필드를 확인하고 입력하세요.
* **패킷 필터**: DNS 패킷 캡처에 적용할 필터 규칙. [PCAP 필터 문법](https://docs.logpresso.com/ko/sonar/4.0/ui/(https:/www.tcpdump.org/manpages/pcap-filter.7.html))을 참고해서 작성하세요.
* **난잡 모드**: 난잡 모드(promiscuous mode) 사용 여부를 나타내는 불리언 값(기본값: true)
* **타임아웃**: 타임아웃 시간(기본값: 100 밀리초)

RSS 수집기

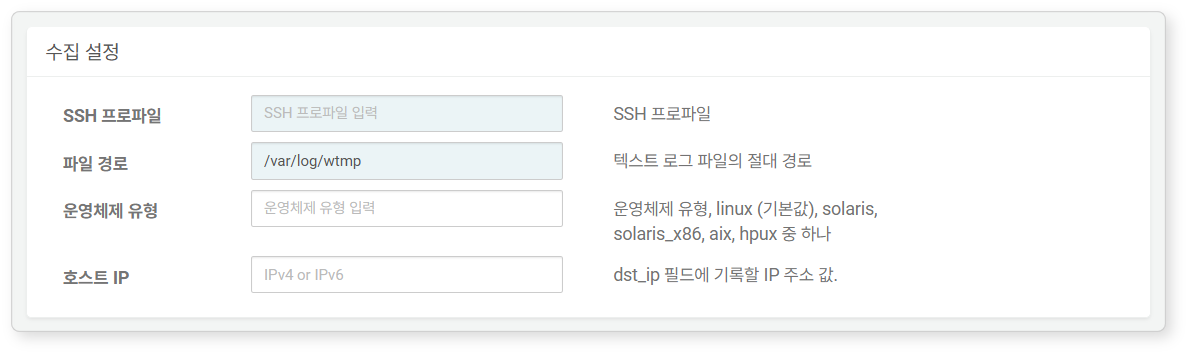
RSS 피드에서 데이터를 수집합니다.



* **RSS 주소**: 데이터를 수집할 RSS 피드 주소
* **태그 제거 여부**: 태그 제거 여부(기본값: false).

SFTP WTMP 파일

네트워크 호스트에 SFTP 통신으로 접속하여 WTMP 이벤트를 수집합니다.



* **SSH 프로파일**: 네트워크 호스트의 SFTP 연결에 필요한 [접속 프로파일](https://docs.logpresso.com/ko/sonar/4.0/ui/section-connect-profile) 이름
* **파일 경로**: wtmp 로그 파일의 절대 경로(기본값: /var/log/wtmp)
* **운영체제 유형**: 운영체제(미지정 시 기본값: **linux**). **linux**, **solaris**, **solaris\_x86**, **aix**, **hpux** 중에 선택해서 입력하세요.
* **호스트 IP**: **dst\_ip** 필드에 기록할 IP 주소. WTMP 로그는 IP 주소나 호스트 정보를 포함하지 않습니다. 하나의 테이블에 여러 WTMP 로그를 기록하는 경우, 호스트를 구분할 수 있도록 호스트 IP를 지정하세요.

SFTP 디렉터리 와처

네트워크 호스트에 SFTP 통신으로 접속하여 지정된 디렉터리에서 패턴과 이름이 일치하는 파일을 수집합니다.

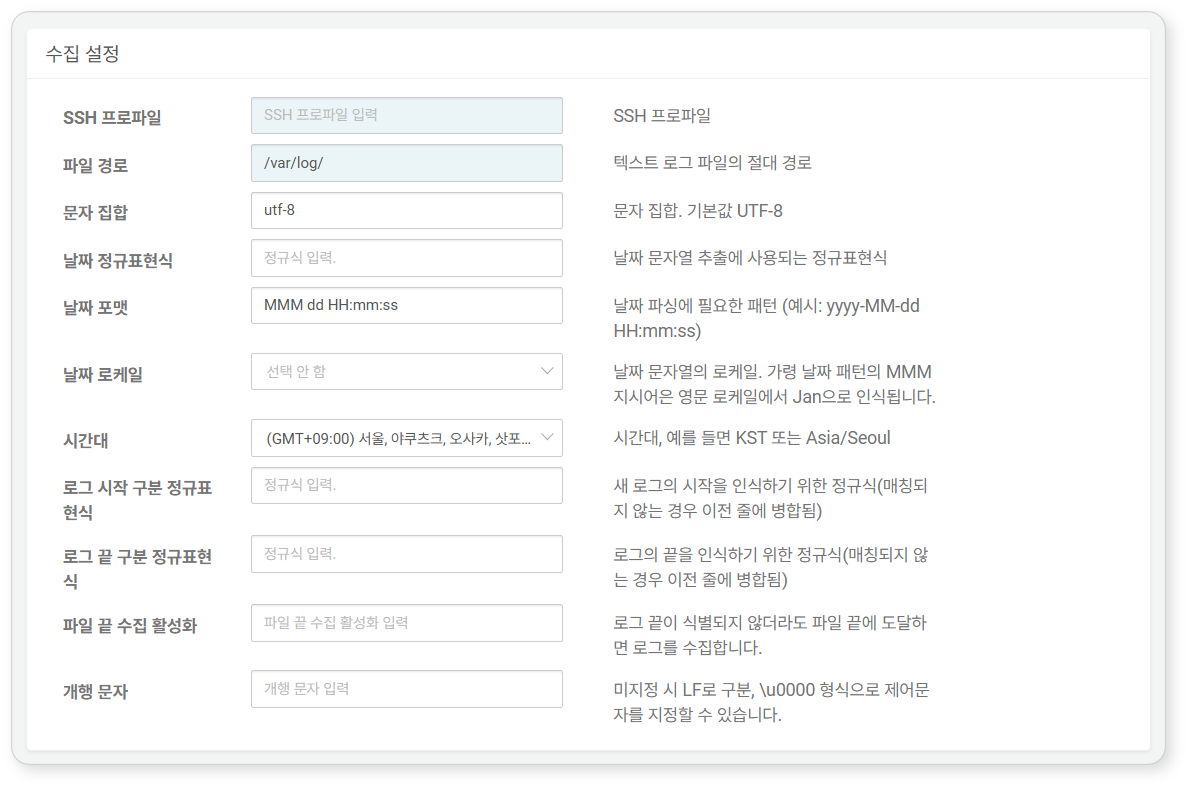


* **SSH 프로파일**: 네트워크 호스트의 SFTP 연결에 필요한 [접속 프로파일](https://docs.logpresso.com/ko/sonar/4.0/ui/section-connect-profile) 이름
* **디렉터리 경로**: 파일을 수집할 대상 디렉터리의 절대 경로(예: /var/log/foo)
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식(예: bar\*.gz)
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 구분 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 구분 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **문자 집합**: 파일에 적용할 문자 인코딩 방식(미지정 시 기본값: utf-8)
* **파일이름 태그**: 수집한 파일 이름을 기록할 필드 이름
* **경로 태그**: 디렉터리 경로를 기록할 필드 이름
* **경로 날짜 추출 포맷**: 디렉터리 경로 문자열에서 날짜 문자열 검색에 사용할 정규표현식. 예를 들어, 디렉터리 경로에 날짜 정보(예: /var/log/foo/2022-11-02)가 있으면 디렉터리 경로 문자열에서 날짜 정보를 추출해 파일의 수집 기준 날짜로 사용합니다.
* **모니터링 대상 일수**: 파일 경로에서 날짜를 추출한 후 모니터링 대상 일수 이내에 있는 파일만 수집합니다(기본값: 없음(기간 제한 없이 파일에 있는 로그 수집)).
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.

SFTP 로테이션 로그 파일

네트워크 호스트에 SFTP 통신으로 접속하여 주기적으로 로테이션되는 로그 파일을 수집합니다.

흔히 유닉스나 리눅스 서버에서는 일정 주기로 로그 파일의 이름을 변경하여 백업하고 로그 파일을 새로 생성합니다. 그러한 파일을 로테이션 로그라 합니다.



* **SSH 프로파일**: 네트워크 호스트의 SFTP 연결에 필요한 [접속 프로파일](https://docs.logpresso.com/ko/sonar/4.0/ui/section-connect-profile) 이름
* **파일 경로**: 수집할 로그 파일의 절대 경로(예: /var/log/secure)
* **문자 집합**: GZIP 압축을 해제한 후 텍스트 파일에 적용할 문자 인코딩 방식(미지정 시 기본값: utf-8)
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 구분 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 구분 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.

SFTP 멀티 로테이션 로그 파일

네트워크 호스트에 SFTP 통신으로 접속하여 주기적으로 다른 경로에 백업하는 방식으로 로테이션하는 로그 파일을 수집합니다.



* **SSH 프로파일**: 네트워크 호스트의 SFTP 연결에 필요한 [접속 프로파일](https://docs.logpresso.com/ko/sonar/4.0/ui/section-connect-profile) 이름
* **디렉터리 경로**: 로그 파일을 수집할 대상 디렉터리의 절대 경로(예: /var/log/foo)
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식(예: bar\*.gz)
* **문자 집합**: 파일에 적용할 문자 인코딩(미지정 시 기본값: utf-8)
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 구분 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 구분 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **파일네임 태그**: 수집한 파일 이름을 기록할 필드 이름
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.

SFTP 일자별 디렉터리

네트워크 호스트에 SFTP 통신으로 접속하여 하루 단위로 생성되는 디렉터리를 순회하면서 파일 이름 패턴과 일치하는 모든 로그 파일을 수집합니다.



* **SSH 프로파일**: 네트워크 호스트의 SFTP 연결에 필요한 [접속 프로파일](https://docs.logpresso.com/ko/sonar/4.0/ui/section-connect-profile) 이름
* **모니터링 기간**: 실시간으로 파일 변화를 모니터링할 기간(기본값: 0). 값이 0일 때, 실시간 수집을 비활성화합니다.
* **디렉터리 경로**: 파일을 수집할 대상 디렉터리의 절대 경로(예: /var/log/foo)
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식(예: bar\*.gz)
* **디렉터리 날짜 정규표현식**: 디렉터리 경로에서 날짜 문자열 검색에 사용할 정규표현식. 디렉터리 경로에 날짜 정보(예: /var/log/foo/2022-11-02)가 있으면 디렉터리 경로 문자열에서 날짜 정보를 추출해 사용합니다.
* **디렉터리 날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(미지정 시 기본값: yyyyMMdd)
* **디렉터리 날짜 사용 유무**: 디렉터리 날짜 포맷을 이용한 로그 생성 날짜 생성 여부(기본값: false)
* **과거 로그 수집 시작 일자**: yyyyMMdd 형식으로 시작 일자를 입력. 지정하지 않으면 과거 로그 수집이 비활성화됩니다.
* **과거 로그 수집 끝 일자**: yyyyMMdd 형식으로 끝 일자를 입력. 지정하지 않으면 과거 로그 수집이 비활성화됩니다.
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **문자 집합**: 파일에 적용할 문자 인코딩 방식(미지정 시 기본값: utf-8)
* **파일이름 태그**: 수집한 파일 이름을 기록할 필드 이름
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.

SNMP GET

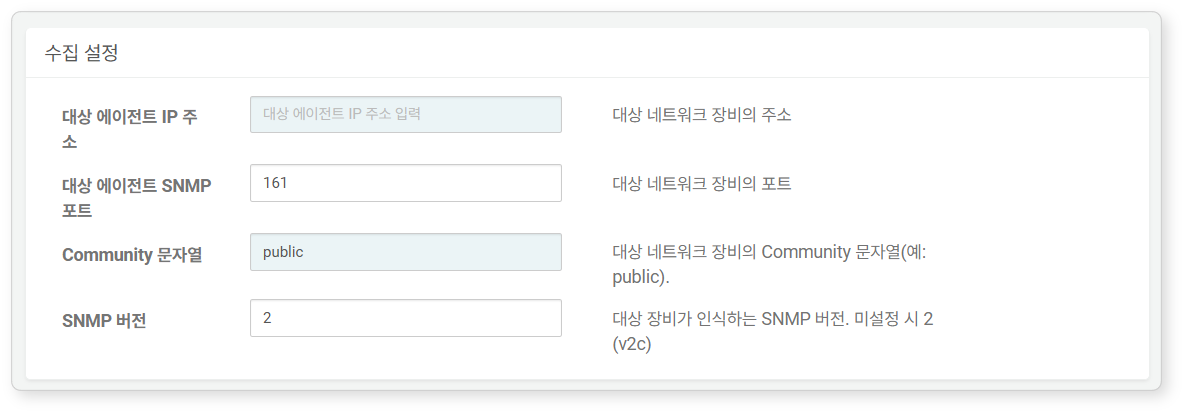
SNMP 에이전트를 대상으로 SNMP 쿼리를 수행한 결과를 수집합니다.



* **SNMP 버전**: v1, v2c 중 하나를 입력(기본값: 없음)
* **IP 주소**: SNMP 에이전트의 호스트 이름 혹은 IP 주소(기본값: 없음)
* **SNMP 포트**: SNMP 쿼리를 수행할 때 사용할 포트 번호(기본값: 161)
* **SNMP 커뮤니티**: SNMP 인증에 사용할 커뮤니티 문자열(예: public)
* **OID 매핑 목록**: 수집할 대상 Object ID와 필드 이름의 쌍을 쉼표(,)로 구분하여 입력. 로그를 수집할 때 OID가 지정한 필드 이름으로 변환되어 값과 함께 기록됩니다.
* **타임아웃 (초)**: 쿼리 응답 만료시간(단위: 초, 기본값: 5)
* **재시도 횟수**: 쿼리 실패 시 재시도 횟수(기본값: 2)

SNMP 인터페이스 통계

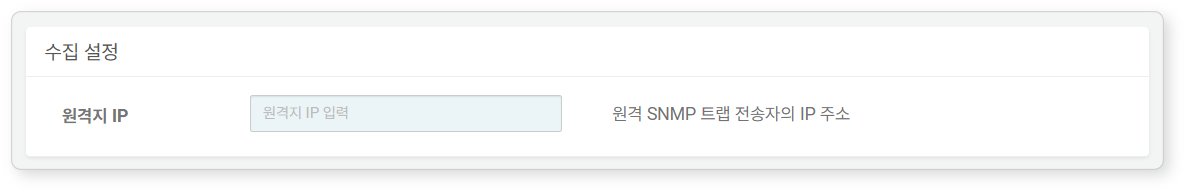
SNMP 쿼리를 통해 인터페이스별 트래픽 통계를 수집합니다.



* **대상 에이전트 IP 주소**: 대상 네트워크 장비의 주소(기본값: 없음)
* **대상 에이전트 SNMP 포트**: SNMP 쿼리를 수행할 때 사용할 포트 번호(기본값: 161)
* **Community 문자열**: 대상 네트워크 장비의 커뮤니티 문자열(예: public).
* **SNMP 버전**: 대상 장비가 인식하는 SNMP 버전 번호를 1, 2 중에서 입력(기본값: 2). 2는 v2c를 의미합니다.

SNMP 트랩 수신기

네트워크 호스트가 로그프레소 소나로 전송한 SNMP 트랩 메시지를 수집합니다.



* **원격지 IP**: SNMP 트랩 메시지 전송 호스트의 IP 주소

SNMPv3 GET

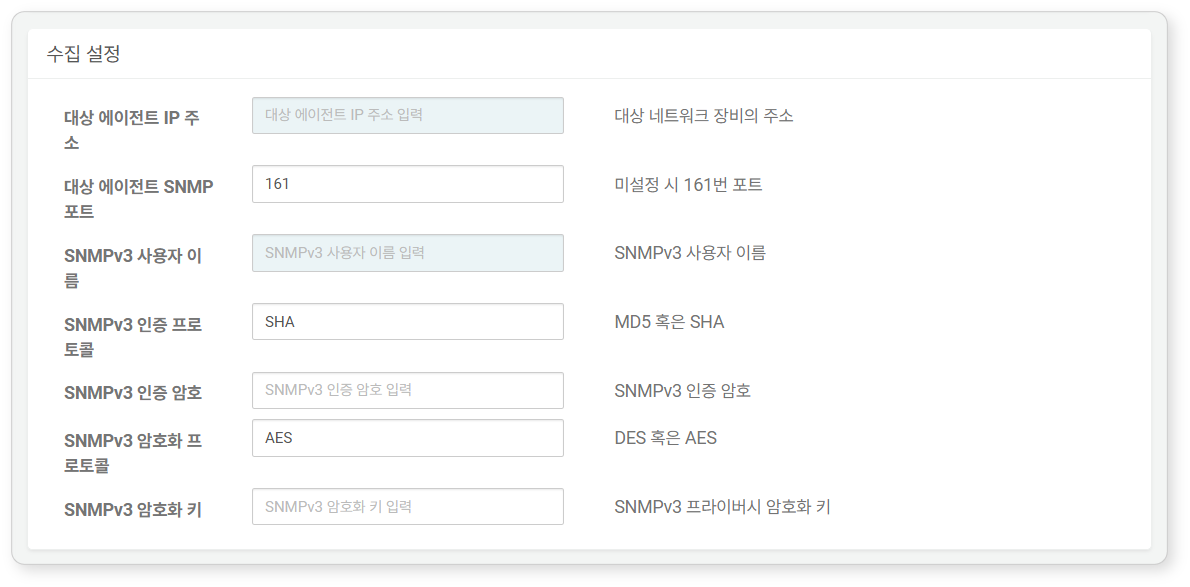
SNMPv3 에이전트를 대상으로 SNMP 쿼리를 수행한 결과를 수집합니다.



* **IP 주소**: SNMP 에이전트의 호스트 이름 혹은 IP 주소(기본값: 없음)
* **SNMP 포트**: SNMP 쿼리를 수행할 때 사용할 포트 번호(기본값: 161)
* **SNMPv3 사용자 이름**: SNMPv3 통신 인증용 사용자 이름
* **SNMPv3 인증 프로토콜**: MD5, SHA 중에서 입력(기본값: SHA).
* **SNMPv3 인증 암호**: SNMPv3 인증 암호
* **SNMPv3 암호화 프로토콜**: DES, AES 중에서 입력(기본값: AES).
* **SNMPv3 암호화 키**: SNMPv3 암호화 키
* **OID 필드 매핑 목록**: 수집할 대상 Object ID와 필드 이름의 쌍을 쉼표(,)로 구분하여 입력. 로그를 수집할 때 OID가 지정한 필드 이름으로 변환되어 값과 함께 기록됩니다.
* **타임아웃 (초)**: 쿼리 응답 만료시간(기본값: 5)
* **재시도 횟수**: 쿼리 실패 시 재시도 횟수(기본값: 2)

SNMPv3 인터페이스 통계

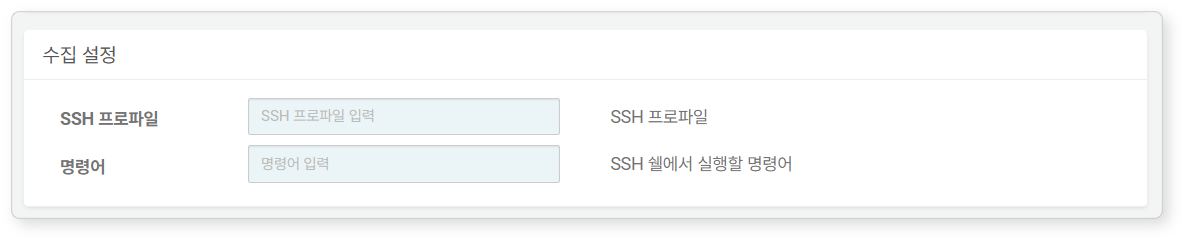
SNMPv3 쿼리를 통해 인터페이스별 트래픽 통계를 수집합니다.



* **대상 에이전트 IP 주소**: 대상 네트워크 장비의 주소(기본값: 없음)
* **대상 에이전트 SNMP 포트**: SNMP 쿼리를 수행할 때 사용할 포트 번호(기본값: 161)
* **SNMPv3 사용자 이름**: SNMPv3 통신 인증용 사용자 이름
* **SNMPv3 인증 프로토콜**: MD5, SHA 중에서 입력(기본값: SHA).
* **SNMPv3 인증 암호**: SNMPv3 인증 암호
* **SNMPv3 암호화 프로토콜**: DES, AES 중에서 입력(기본값: AES).
* **SNMPv3 암호화 키**: SNMPv3 암호화 키

SSH 실행

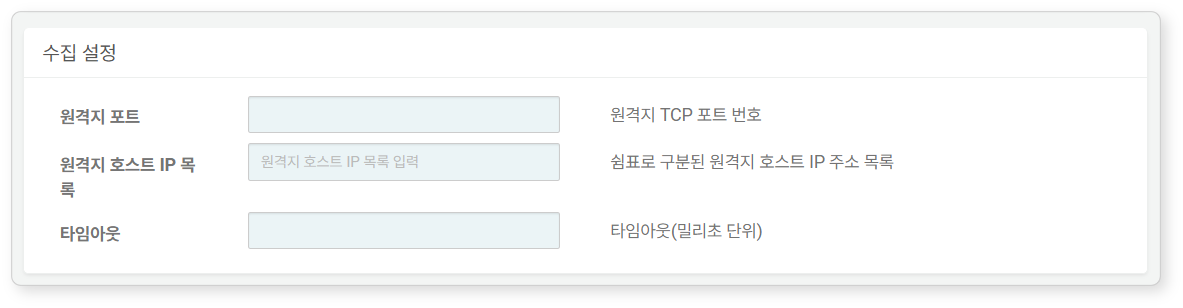
네트워크 호스트에 SSH로 접속해 명령어를 실행하고, 표준 출력을 수집합니다.



* **SSH 프로파일**: 네트워크 호스트의 SSH 연결에 필요한 [접속 프로파일](https://docs.logpresso.com/ko/sonar/4.0/ui/section-connect-profile) 이름
* **명령어**: 로그인 셸에서 실행할 명령어

TCP 포트 상태 수집기

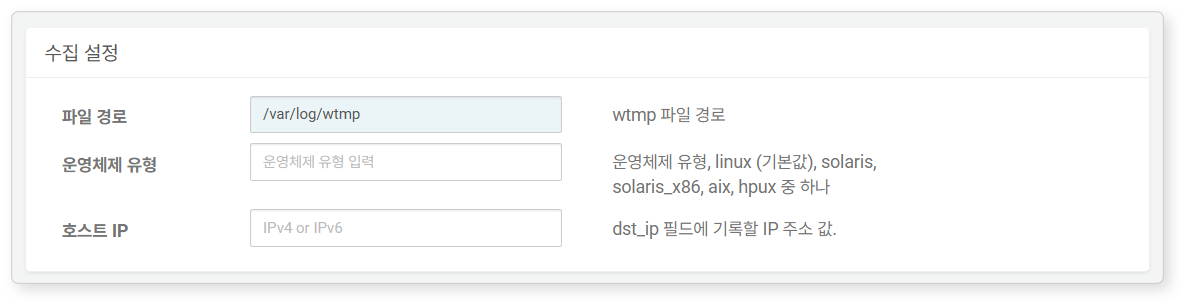
네트워크 호스트의 TCP 포트 개방 여부를 확인하고 결과를 수집합니다.



* **원격지 포트**: 원격지 TCP 포트 번호
* **원격지 호스트 IP 목록**: 쉼표(,)를 구분자로 하는 원격지 호스트 IP 주소 목록. 도메인 주소(예: localhost, example.com)를 입력해도 IP 주소로 변환하여 인식합니다.
* **타임아웃**: 연결 타임아웃 시간(단위: 밀리초, 기본값: 0)

WTMP

로컬 호스트의 WTMP 이벤트를 수집합니다.



* **파일 경로**: wtmp 로그 파일의 절대 경로(기본값: /var/log/wtmp)
* **운영체제 유형**: 운영체제(미지정 시 기본값: **linux**). **linux**, **solaris**, **solaris\_x86**, **aix**, **hpux** 중에 선택해서 입력하세요.
* **호스트 IP**: **dst\_ip** 필드에 기록할 IP 주소. WTMP 로그는 IP 주소나 호스트 정보를 포함하지 않습니다. 하나의 테이블에 여러 WTMP 로그를 기록하는 경우, 호스트를 구분할 수 있도록 호스트 IP를 지정하세요.

디렉터리 와처

로컬 디렉터리에서 패턴과 이름이 일치하는 파일을 수집합니다.



* **디렉터리 경로**: 파일을 수집할 대상 디렉터리의 절대 경로(예: /var/log/foo)
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식(예: bar\*.gz)
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **문자 집합**: 파일에 적용할 문자 인코딩 방식(미지정 시 기본값: utf-8)
* **파일이름 태그**: 수집한 파일 이름을 기록할 필드 이름
* **경로 태그**: 디렉터리 경로를 기록할 필드 이름
* **경로 날짜 추출 포맷**: 디렉터리 경로 문자열에서 날짜 문자열 검색에 사용할 정규표현식. 예를 들어, 디렉터리 경로에 날짜 정보(예: /var/log/foo/2022-11-02)가 있으면 디렉터리 경로 문자열에서 날짜 정보를 추출해 파일의 수집 기준 날짜로 사용합니다.
* **모니터링 대상 일수**: 파일 경로에서 날짜를 추출한 후 모니터링 대상 일수 이내에 있는 파일만 수집합니다(기본값: 없음(기간 제한 없이 파일에 있는 로그 수집)).
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.

로테이션 로그 파일

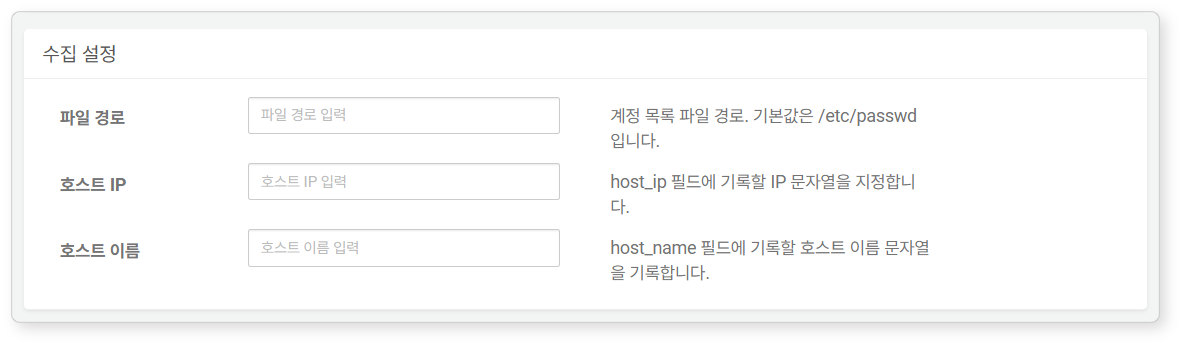
로컬 호스트에서 로테이션되는 로그 파일을 수집합니다.



* **파일 경로**: 수집할 로그 파일의 절대 경로(예: /var/log/secure)
* **문자 집합**: 파일에 적용할 문자 인코딩 방식(미지정 시 기본값: utf-8)
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.

리눅스 계정 이벤트 탐지

로컬 호스트에서 계정 목록 파일 passwd의 변경사항을 수집합니다.



* **파일 경로**: 계정 정보가 있는 파일 경로(기본값: /etc/passwd)
* **호스트 IP 목록**: **host\_ip** 필드에 기록할 IP 주소 문자열(기본값: 없음)
* **호스트 이름**: **host\_name** 필드에 기록할 호스트 이름 문자열(기본값: 없음)

리커시브 GZIP 디렉터리 와처

로컬 호스트의 지정된 디렉터리에서 파일이름 패턴과 일치하는 모든 텍스트 GZIP 로그 파일을 수집합니다.



* **디렉터리 경로**: GZIP 파일을 수집할 대상 디렉터리의 절대 경로(예: /var/log/foo)
* **디렉터리 경로 패턴**: 로그 파일이 있는 디렉터리를 선택할 때 사용할 정규표현식
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식(예: bar\*.gz)
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **문자 집합**: GZIP 압축을 해제한 후 텍스트 파일에 적용할 문자 인코딩 방식(미지정 시 기본값: utf-8)
* **하위 디렉터리 포함**: 하위 디렉터리 포함 여부를 나타내는 불리언 값(미지정 시 기본값: false)
* **파일이름 태그**: 수집한 파일 이름을 기록할 필드 이름
* **경로 태그**: 디렉터리 경로를 기록할 필드 이름
* **경로 날짜 추출 포맷**: 디렉터리 경로 문자열에서 날짜 문자열 검색에 사용할 정규표현식. 예를 들어, 디렉터리 경로에 날짜 정보(예: /var/log/foo/2022-11-02)가 있으면 디렉터리 경로 문자열에서 날짜 정보를 추출해 GZIP 파일의 수집 기준 날짜로 사용합니다.
* **모니터링 대상 일수**: 파일 경로에서 날짜를 추출한 후 모니터링 대상 일수 이내에 있는 파일만 수집합니다(기본값: 없음(기간 제한 없이 파일에 있는 로그 수집)).
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.
* **GZIP 파일 삭제**: 수집 완료된 GZIP 파일의 삭제 여부를 나타내는 불리언 값(미지정 시 기본값: false). true를 입력하면 파일을 수집한 후 디렉터리에서 GZIP 파일을 삭제합니다.

리커시브 디렉터리 와처

로컬 호스트의 지정된 디렉터리에서 파일이름 패턴과 일치하는 모든 텍스트 로그 파일을 수집합니다.



* **디렉터리 경로**: 파일을 수집할 대상 디렉터리의 절대 경로(예: /var/log/foo)
* **디렉터리 경로 패턴**: 로그 파일이 있는 디렉터리를 선택할 때 사용할 정규표현식
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식(예: bar\*.gz)
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **문자 집합**: 파일에 적용할 문자 인코딩 방식(미지정 시 기본값: utf-8)
* **하위 디렉터리 포함**: 하위 디렉터리 포함 여부를 나타내는 불리언 값(미지정 시 기본값: false)
* **파일이름 태그**: 수집한 파일 이름을 기록할 필드 이름
* **경로 태그**: 디렉터리 경로를 기록할 필드 이름
* **경로 날짜 추출 포맷**: 디렉터리 경로 문자열에서 날짜 문자열 검색에 사용할 정규표현식. 예를 들어, 디렉터리 경로에 날짜 정보(예: /var/log/foo/2022-11-02)가 있으면 디렉터리 경로 문자열에서 날짜 정보를 추출해 GZIP 파일의 수집 기준 날짜로 사용합니다.
* **모니터링 대상 일수**: 파일 경로에서 날짜를 추출한 후 모니터링 대상 일수 이내에 있는 파일만 수집합니다(기본값: 없음(기간 제한 없이 파일에 있는 로그 수집)).
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.

멀티 로테이션 로그 파일

로컬 호스트에서 주기적으로 다른 경로에 백업하는 방식으로 로테이션하는 로그 파일을 수집합니다.



* **디렉터리 경로**: 로그 파일을 수집할 대상 디렉터리의 절대 경로(예: /var/log/foo)
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식(예: bar\*.gz)
* **문자 집합**: 파일에 적용할 문자 인코딩(미지정 시 기본값: utf-8)
* **하위 디렉터리 포함**: 하위 디렉터리 포함 여부를 나타내는 불리언 값(미지정 시 기본값: false)
* **디렉터리 경로 패턴**: 로그 파일이 있는 디렉터리를 선택할 때 사용할 정규표현식
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 구분 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 구분 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **파일네임 태그**: 수집한 파일 이름을 기록할 필드 이름
* **경로 태그**: 디렉터리 경로를 기록할 필드 이름
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.

설정 파일 변경 탐지

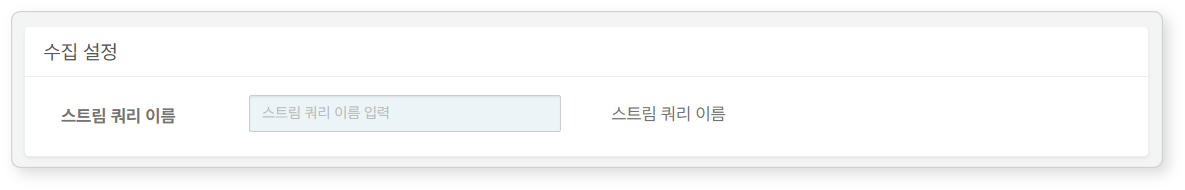
설정 파일의 변경 내역을 수집합니다.



* **디렉터리 경로**: 파일을 수집할 대상 디렉터리의 절대 경로(예: /etc)
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식(예: bar\*.gz)
* **파일이름 태그**: 수집한 파일 이름을 기록할 필드 이름
* **파일 내용 추가**: 파일 내용의 기록 여부(기본값:false). 수집한 내용은 **file\_content** 필드에 기록합니다. 변경 내역은 파일 전체에 대해 탐지하지만, 파일 내용은 128KB까지만 기록합니다.
* **변경사항 추가**: 수집 대상 설정 파일의 변경 내역 기록 여부(기본값: false) 이 옵션을 선택하면 변경전 내용과 변경 후 내용을 **file\_diff** 필드에 기록합니다. 변경 내역은 최대 128KB까지 기록합니다.
* **파일 인코딩**: 텍스트 파일에 적용할 문자 인코딩 방식을 목록에서 선택(기본값: utf-8)

스트림 쿼리 출력

스트림 쿼리의 출력을 수집합니다.



* **스트림 쿼리 이름**: 수집할 스트림 쿼리의 이름. 스트림의 이름은 [system streams](https://docs.logpresso.com/ko/sonar/4.0/ui/ko/query/system-streams-command) 명령의 출력 필드 **name**에서 확인할 수 있습니다.

시스로그 수집기

원격 호스트가 전송하는 시스로그 메시지를 수집합니다.



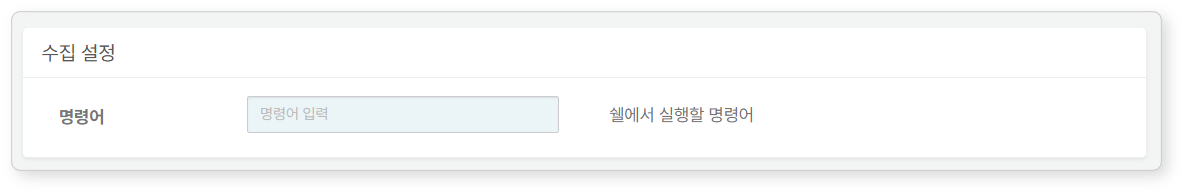
* **원격지 IP**: 시스로그 메시지를 전송하는 원격지 IP 주소 또는 주소 대역. 입력값은 단일 IP 주소(예: 192.168.0.1), 시작 주소와 끝 주소 사이에 하이픈(-)이 있는 IP 주소 범위(예: 192.168.0.1-192.168.0.254), CIDR 표기법에 따른 IP 주소 대역(예: 192.168.0.0/24) 형식으로 사용해야 합니다.
* **포트**: 시스로그 메시지를 수신할 포트 번호
* **시스로그 설비 목록**: [RFC 5424](https://tools.ietf.org/html/rfc5424)에 정의된 PRI 상수값. 여러 개의 PRI 상수를 지정하려면 구분자로 쉼표(,)를 이용하세요. PRI 상수값은 Facility 값의 8배수에 Severity 값을 더한 값입니다. 다음 표를 참조해서 계산된 값을 사용하세요.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Facility(↓)<br />Severity(→) | 0/Emer | 1/Alert | 2/Crit | 3/Error | 4/Warn | 5/Notice | 6/Info | 7/Debug |
| 0 / kern | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 / user | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 2 / mail | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| 3 / deamon | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 |
| 4 / auth | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 |
| 5 / syslog | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 |
| 6 / lpr | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 |
| 7 / news | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 |
| 8 / uucp | 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 |
| 9 / clock | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 |
| 10 / authpriv | 80 | 81 | 82 | 83 | 84 | 85 | 86 | 87 |
| 11 / ftp | 88 | 89 | 90 | 91 | 92 | 93 | 94 | 95 |
| 12 / ntp | 96 | 97 | 98 | 99 | 100 | 101 | 102 | 103 |
| 13 / audit | 104 | 105 | 106 | 107 | 108 | 109 | 110 | 111 |
| 14 / alert | 112 | 113 | 114 | 115 | 116 | 117 | 118 | 119 |
| 15 / solaris-cron | 120 | 121 | 122 | 123 | 124 | 125 | 126 | 127 |
| 16 / local0 | 128 | 129 | 130 | 131 | 132 | 133 | 134 | 135 |
| 17 / local1 | 136 | 137 | 138 | 139 | 140 | 141 | 142 | 143 |
| 18 / local2 | 144 | 145 | 146 | 147 | 148 | 149 | 150 | 151 |
| 19 / local3 | 152 | 153 | 154 | 155 | 156 | 157 | 158 | 159 |
| 20 / local4 | 160 | 161 | 162 | 163 | 164 | 165 | 166 | 167 |
| 21 / local5 | 168 | 169 | 170 | 171 | 172 | 173 | 174 | 175 |
| 22 / local6 | 176 | 177 | 178 | 179 | 180 | 181 | 182 | 183 |
| 23 / local7 | 184 | 185 | 186 | 187 | 188 | 189 | 190 | 191 |

* PRI 상수값을 지정하지 않으려면 -1을 입력하세요.
* 아무 값도 입력하지 않으면 모든 facility 값을 입력한 것과 같습니다.
* **날짜 정규표현식**: 시스로그에서 날짜 및 시각을 추출하는데 사용할 정규표현식
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 날짜 형식(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다. 가령 날짜 파싱 포맷 지시자 중에서 MMM은 **Jan**(날짜 로케일이 en일 때) 혹은 **1월**(날짜 로케일이 ko일 때)로 해석됩니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **정규식 필터**: 추출할 데이터를 선별할 정규식 패턴(폐기 예정)

외부 프로그램

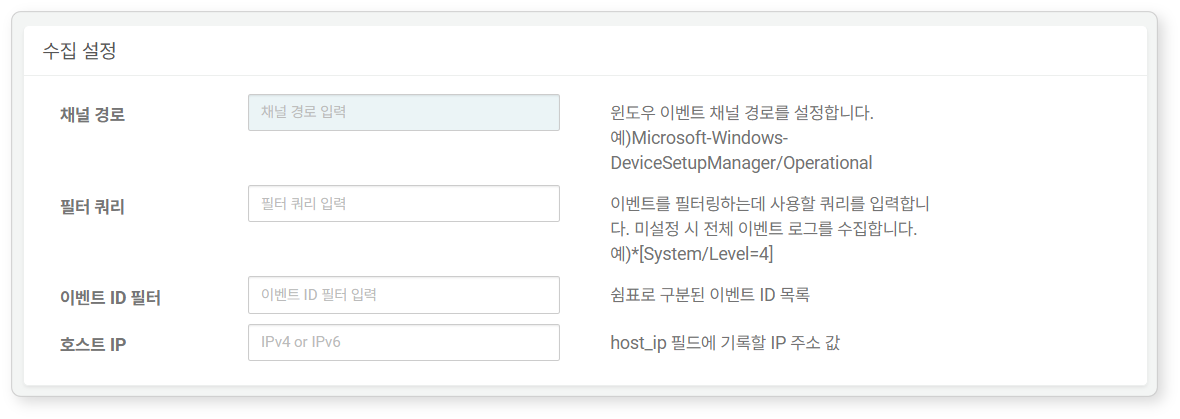
로그프레소 소나의 외부 프로그램을 실행하고 프로그램의 표준 출력을 수집합니다.



* **명령어**: 셸에서 실행할 명령줄(예: uname -a)

윈도우 XML 이벤트 로그

윈도우 센트리에서 이벤트 로그를 수집합니다. 이 수집기는 윈도우 센트리에 설정할 수 있습니다.



* **채널 경로**: 윈도우 이벤트 채널 경로(예: Microsoft-Windows-DeviceSetupManager/Operational)
* **필터 쿼리**: 이벤트를 필터링하는데 사용할 쿼리 문자열(예: \*[System/Level=4], 기본값: 모든 이벤트 로그)
* **이벤트 ID 필터**: 쉼표(,)를 구분자로 하는 이벤트 ID 목록
* **호스트 IP**: **host\_ip** 필드에 기록할 IP 주소 값

일자별 디렉터리

로컬 호스트에서 하루 단위로 생성되는 디렉터리를 순회하면서 파일 이름 패턴과 일치하는 모든 로그 파일을 수집합니다.



* **모니터링 기간**: 실시간으로 파일 변화를 모니터링할 기간(기본값: 0). 값이 0일 때, 실시간 수집을 비활성화합니다.
* **디렉터리 경로**: 파일을 수집할 대상 디렉터리의 절대 경로(예: /var/log/foo)
* **파일이름 패턴**: 수집할 로그 파일 이름을 나타내는 정규표현식(예: bar\*.gz)
* **디렉터리 날짜 정규표현식**: 디렉터리 경로에서 날짜 문자열 검색에 사용할 정규표현식. 디렉터리 경로에 날짜 정보(예: /var/log/foo/2022-11-02)가 있으면 디렉터리 경로 문자열에서 날짜 정보를 추출해 사용합니다.
* **디렉터리 날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(미지정 시 기본값: yyyyMMdd)
* **과거 로그 수집 시작 일자**: yyyyMMdd 형식으로 시작 일자를 입력. 지정하지 않으면 과거 로그 수집이 비활성화됩니다.
* **과거 로그 수집 끝 일자**: yyyyMMdd 형식으로 끝 일자를 입력. 지정하지 않으면 과거 로그 수집이 비활성화됩니다.
* **날짜 정규표현식**: 로그에서 날짜 및 시간 정보 문자열 추출에 사용할 정규표현식
* 다음은 아파치 로그 등에서 쓰이는 날짜 정보를 정규표현식으로 표현한 예시입니다.
* \[(\d{2}/\S{3,10}/\d{4}:\d{2}:\d{2}:\d{2} \+\d{4})\]
* 날짜 정규표현식에서 괄호쌍(( ))으로 감싼 문자열 그룹들을 이어붙여서 하나의 날짜 문자열을 만들어냅니다. 이렇게 생성된 날짜 문자열로부터 **날짜 포맷**, **날짜 로케일**에 따라 시간 정보 추출에 사용합니다.
* **날짜 포맷**: 날짜 문자열을 파싱하는데 사용할 포맷(예: yyyy-MM-dd HH:mm:ss). **날짜 로케일**에 따라 파싱하는 형식이 다를 수 있습니다.
* **날짜 로케일**: 날짜 문자열에 적용할 언어를 en, ja, zh, ko 중에서 선택(기본값: 선택 안 함)
* **시간대**: 시간대를 목록에서 선택. API를 이용해 수집기를 구성하는 경우 **KST** 또는 **Asia/Seoul**와 같은 형식으로 시간대를 지정할 수 있습니다.
* **로그 시작 정규표현식**: 로그 레코드의 시작을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **로그 끝 정규표현식**: 로그 레코드의 끝을 인식하는 정규표현식. 하나의 로그 레코드가 여러 줄에 걸쳐 있을 때 사용합니다(미지정 시 개행 문자를 로그 구분자로 인식).
* **문자 집합**: 파일에 적용할 문자 인코딩 방식(미지정 시 기본값: utf-8)
* **파일이름 태그**: 수집한 파일 이름을 기록할 필드 이름
* **파일 끝 수집 활성화**: 로그 끝이 식별되지 않더라도 파일 끝에 도달하면 로그를 수집(미지정 시 기본값: false)
* **개행 문자**: 개행 문자로 사용할 문자를 유니코드 이스케이프 시퀀스 형식(예: LF를 의미하는 \u000A)으로 입력(미지정 시 기본값: \u000A). 유니코드 검색은 [SYMBL](https://symbl.cc/en/unicode-table/)을 참고하세요.

웹키퍼 (구 버전)

마이크로소프트 SQL 서버를 데이터베이스로 사용하는 소만사 웹키퍼에서 데이터를 수집합니다(엘라스틱 서치를 사용하는 최신 버전의 웹키퍼는 앱 스토어에서 [웹키퍼](https://logpresso.store/ko/apps/webkeeper) 앱을 내려받아 사용하세요.)



* **데이터베이스 IP**: 접속할 마이크로소프트 SQL 서버의 IP 주소
* **데이터베이스 접속 포트**: 데이터베이스 접속 포트(기본값: 1433)
* **데이터베이스 계정**: 접속 가능한 데이터베이스 계정
* **데이터베이스 암호**: 데이터베이스 계정의 암호
* **에이전트 UID**: 네트워크 에이전트의 UID
* **파일 목록 추가**: 첨부파일 목록 수집 여부를 나타내는 불리언 값(기본값: false)
* **수집 기준 일자**: yyyy-MM-dd HH:00:00 형식으로 입력하는 수집 기준 일자. 미지정시 로거 시작 시점부터 데이터를 수집합니다.
* **수집 반복 횟수 제한**: 수집 주기마다 실행할 수 있는 최대 SQL 쿼리 개수(기본값: 0, 제한 없음)

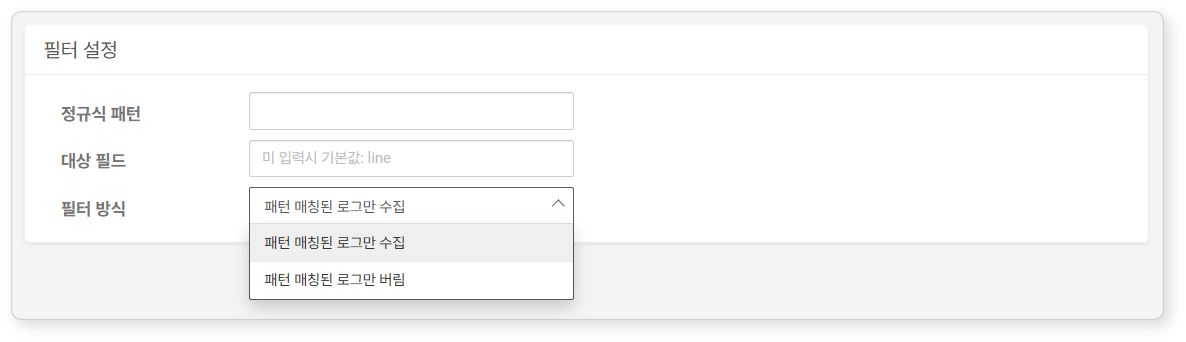
수집 설정이 불필요한 수집기들

다음 수집기들은 별도의 수집 설정 항목이 없습니다.

|  |  |
| --- | --- |
| 수집기 | 설명 |
| sFlow | 네트워크 기기의 sFlow v5 패킷을 수집 |
| 디스크 사용량 | 노드 또는 센트리 호스트의 디스크 사용현황 정보를 수집 |
| 센트리 성능 로그 | 센트리의 성능 로그를 수집 |
| 시스템 경보 | 시스템 경보를 수집 |
| 포워더 성능 로그 | 전달 노드의 성능 로그를 수집 |

필터 설정

필터 설정은 수집된 원본 레코드로부터 저장하거나 버릴 레코드를 정의하는 기능입니다.

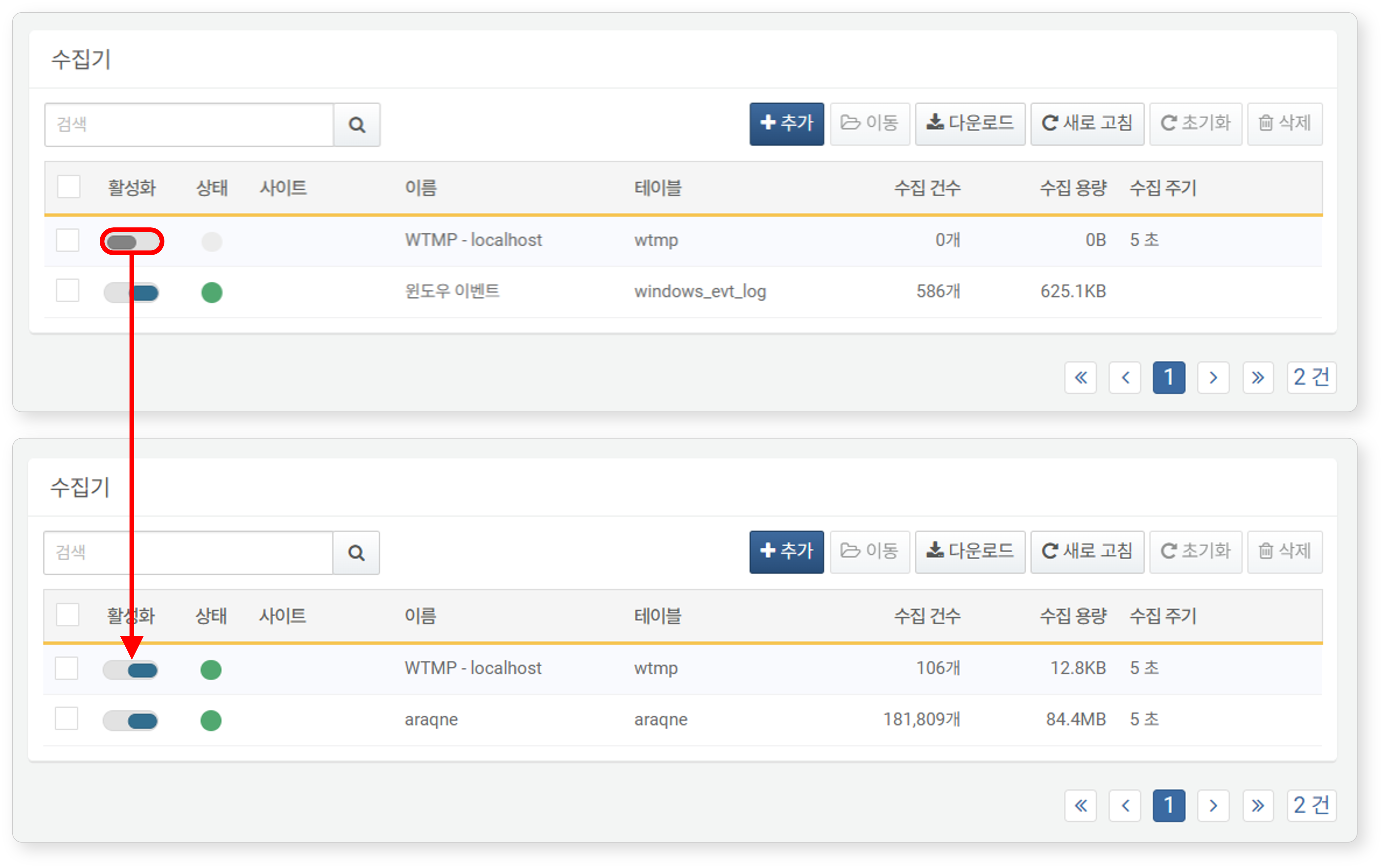


* **정규식 패턴**: 검색할 로그 레코드에 대한 정규표현식. 정규식 패턴을 지정하지 않으면 모든 원본 레코드에 **필터 방식**에서 지정한 동작을 수행합니다.
* **대상 필드**: 정규식 패턴을 이용해 검색할 입력 데이터의 필드(기본값: line)
* **필터 방식**: 정규식 패턴과 일치하는 레코드에 대한 처리 방식으로, **패턴 매칭된 로그만 수집**(기본값), **패턴 매칭된 로그만 버림** 중에서 선택

#### 수집기 제어

활성화/비활성화

추가된 수집기의 기본 상태는 비활성화입니다. 수집기를 활성화하려면 추가한 수집기의 활성화 토글 버튼을 클릭해 활성화하세요.

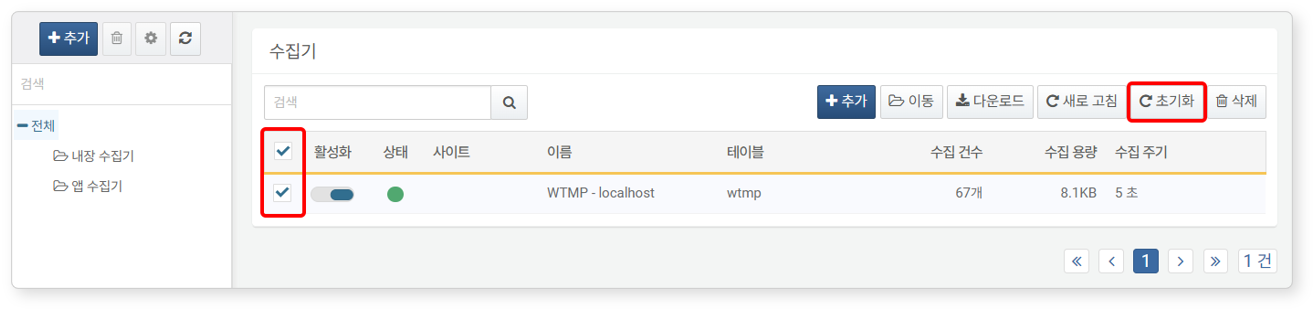


수집 상태 초기화

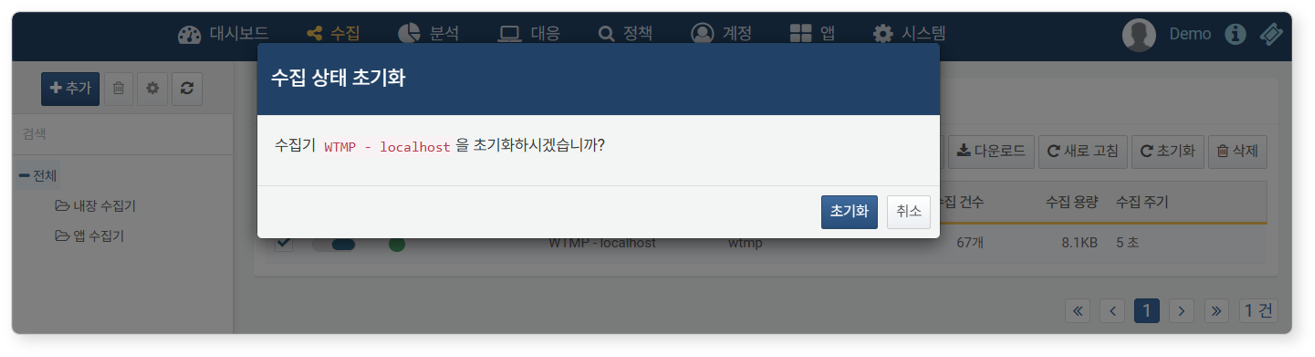
수집기의 **수집 건수**, **수집 용량**을 초기화할 수 있습니다. 수집기를 초기화하려면,

수집기 목록에서 초기화할할 수집기 정보가 있는 행의 체크박스를 선택하세요.

도구 모음에서 **초기화**를 클릭하세요.



**수집 상태 초기화** 대화상자에서 초기화할 수집기 목록을 확인하고 **초기화**를 클릭하세요. 초기화하지 않으려면 **취소**를 클릭하세요.

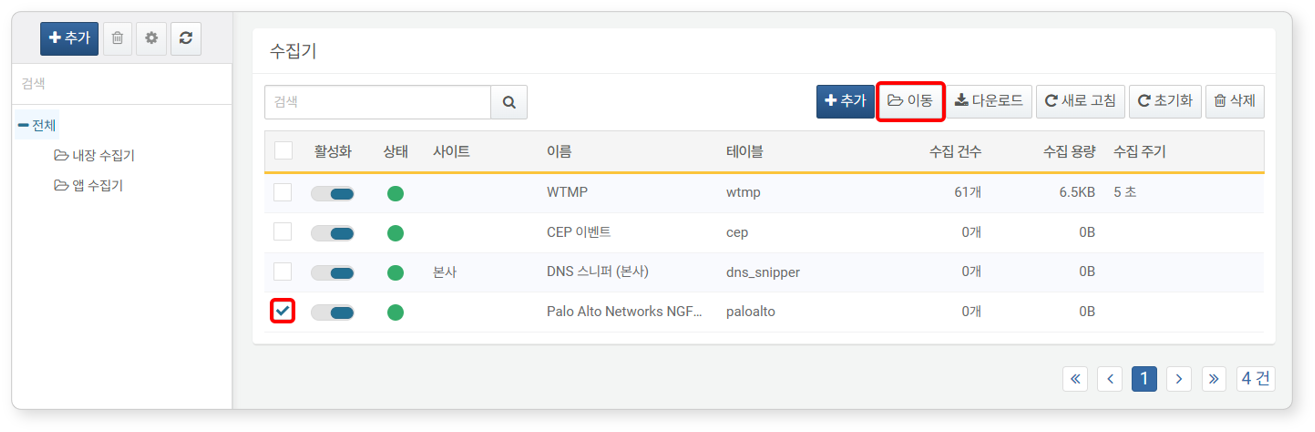


수집기를 초기화하면 이전에 수집한 데이터를 다시 수집합니다. 테이블에 중복된 데이터가 발생할 수 있으므로 운영 환경에서는 가급적이면 사용하지 마세요.

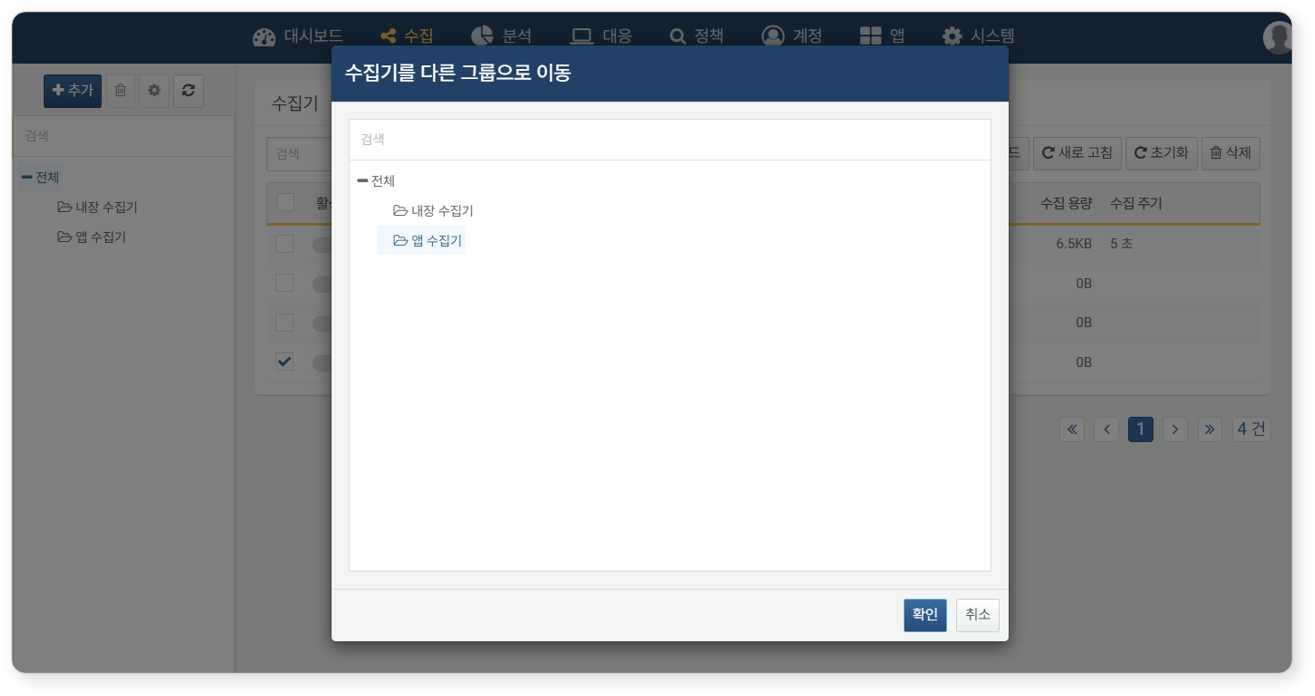
수집기 그룹 변경

수집기를 수집기 그룹에 할당하거나, 소속 수집기 그룹을 변경하려면,

수집기 목록에서 수집기 그룹을 할당/변경할 수집기 정보가 있는 행의 체크박스를 선택한 다음, 도구 모음에서 **이동**을 클릭하세요.



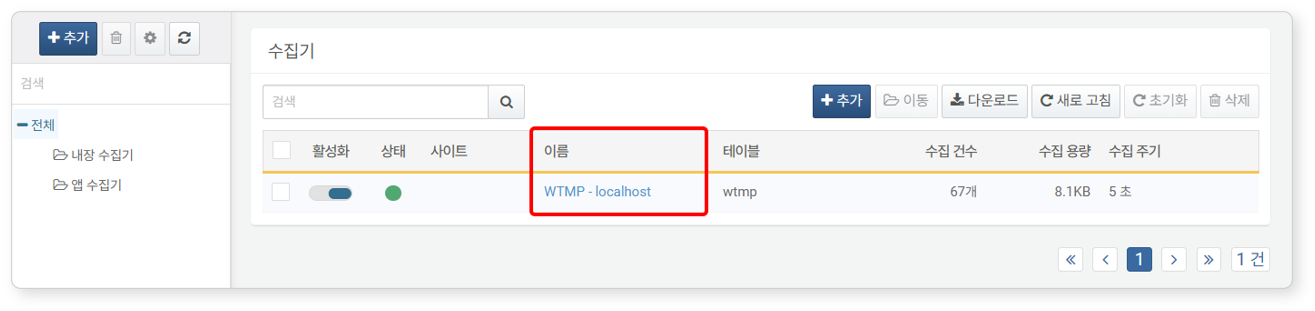
**수집기를 다른 그룹으로 이동** 대화상자에서 수집기를 할당할 그룹을 선택하고 확인을 **클릭**하세요. 수집기 그룹을 변경하지 않으려면 **취소**를 클릭하세요.



#### 수집기 수정

수집기를 수정하려면,

수집기 목록에서 수정할 수집기의 이름을 클릭하세요.



**수집기 수정** 화면에서 정보를 수정한 후 **확인**을 클릭하세요.

* 수정할 속성에 대한 설명은 [수집기 추가](https://docs.logpresso.com/ko/sonar/4.0/ui/section-loggers)를 참조하세요.
* **적재 위치**, **수집 위치**, **수집 모델**은 수정할 수 없습니다.

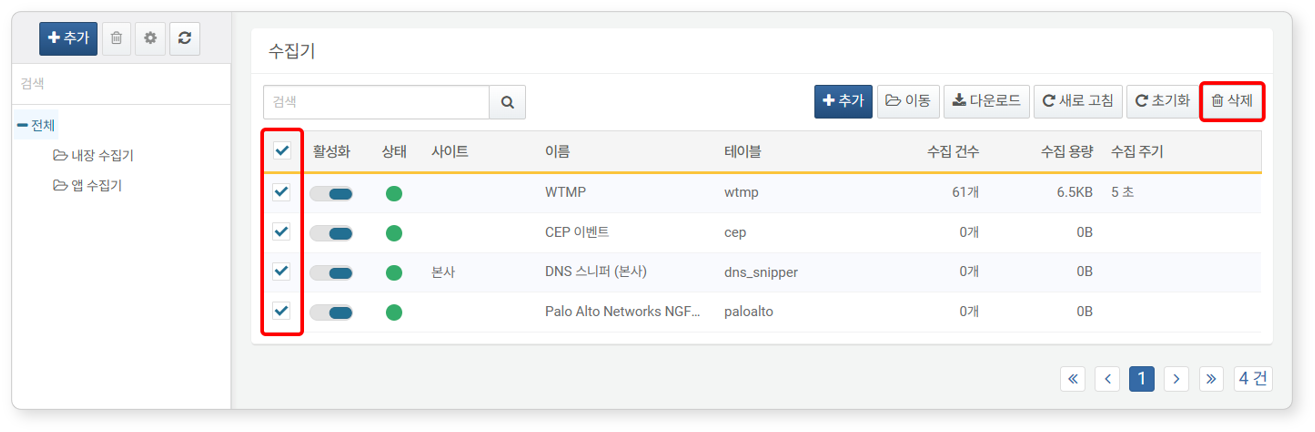
#### 수집기 삭제

실시간 탐지 시나리오에서 사용 중인 수집기를 삭제하면 시나리오는 삭제된 수집기에서 데이터를 받을 수 없게 됩니다. 수집기를 삭제하기 전에 반드시 삭제할 수집기를 사용하는 기능이 있는지 확인하세요.

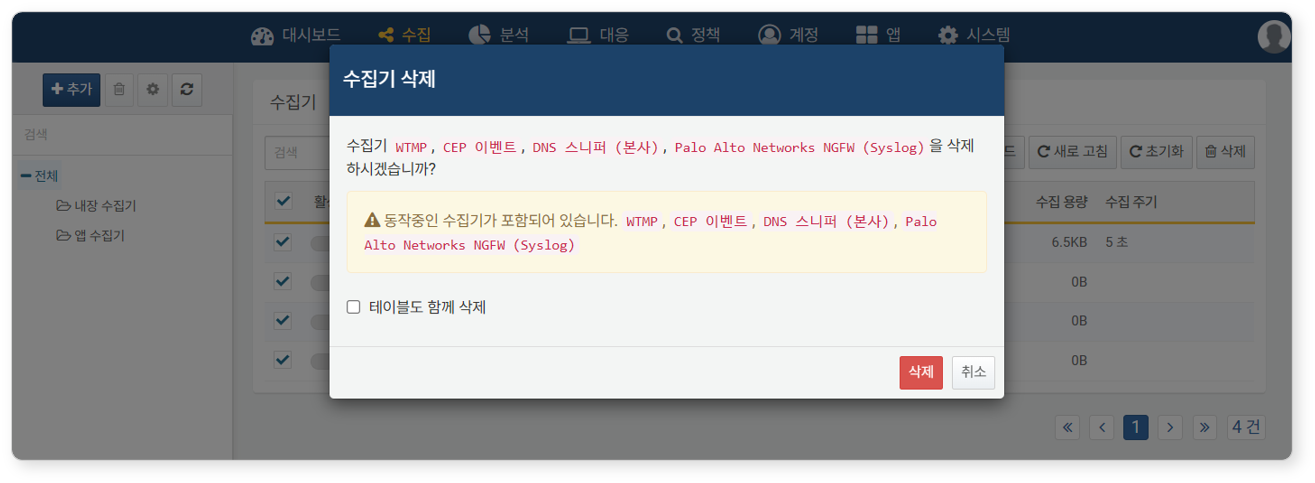
수집기를 삭제하려면,

수집기 목록에서 삭제할 수집기 정보가 있는 행의 체크박스를 선택하세요.

도구 모음에서 **삭제**를 클릭하세요.



**수집기 삭제** 대화상자에서 삭제할 수집기 목록을 확인한 후 **삭제**를 클릭하세요. 삭제하지 않으려면 **취소**를 클릭하세요.



* **테이블도 함께 삭제**를 선택하면 테이블 및 테이블에 저장된 데이터도 함께 삭제됩니다.
* 활성화된 수집기를 삭제하려 하면 그림과 같이 경고가 나타납니다.