기술 백서

PAIL

물리 AI 행동을 위한 권한·증거 계층

Glowfly Identity Systems Co., Ltd.

버전 1.1 · 2026년 11월

초록

로봇이 수행하는 동작 가운데 사람이 검토하지 않은 소프트웨어가 계획한 것이 늘고 있다. 대규모 언어모델 기반 플래너, 자율 에이전트, 플릿 스케줄러, 원격 콘솔이 그 예다. 일반적인 로봇의 명령 인터페이스는 도착한 명령을 그대로 실행하며, 사고가 난 뒤에 특정 동작을 누가 어떤 한도로 허가했는지를 제3자가 확인할 수 있는 기록은 대개 남아 있지 않다.

PAIL(Physical Action Identity Layer)은 이 공백을 두 가지 기능으로 메운다. 권한 계층으로서 PAIL은 모든 물리적 행동이 서명되고 범위가 좁으며 유효 기간이 정해진 **행동 토큰**으로 뒷받침될 것을 요구하고, 로봇의 명령 경로에 놓인 게이트에서 이를 집행한다. 증거 계층으로서 PAIL은 각 인가 결정을 그 결정이 효력을 갖기 전에 변조를 탐지할 수 있는 저널에 기록하고, 공개된 규격과 도구로 제3자가 이를 검증할 수 있게 한다.

본 백서는 PAIL이 다루는 문제, 설계 원칙, 구조, 인가된 행동 하나의 일생, 장애 시 동작, 증거 모델, 그리고 PAIL이 지키는 것과 통합자의 책임으로 남는 것의 경계를 기술한다. 지금까지 측정된 결과와 그 한계도 함께 밝힌다. PAIL은 안전 시스템이 아니며 안전 시스템을 대체하지 않는다. PAIL은 요청된 행동이 인가되었는지를 판정하고 그 판정의 증명을 보존하는 계층이다.

목차

1 서론
2 문제 정의
3 설계 원칙
4 구조
5 행동 토큰
6 행동 하나의 일생
7 장애 시 동작
8 증거
9 신뢰 경계
10 통합
11 운영
12 관측
13 측정 결과
14 제공 형태
15 범위와 한계
16 결론

1 서론

이동 로봇과 매니퓰레이터의 소프트웨어에는 인지, 계획, 제어, 기능 안전을 담당하는 계층이 잘 자리 잡고 있다. 그러나 대부분의 배치에는 더 좁은 질문에 답하는 계층이 없다. **이 특정한 행동은 누가, 어떤 한도로 허가했는가.**

오랫동안 이 질문에는 별도의 계층이 필요하지 않았다. 동작은 운영자가 요청하거나 엔지니어가 작성하고 시험한 프로그램이 요청했고, 책임은 조직의 체계를 따랐다. 이 전제는 더 이상 성립하지 않는다. 언어모델 기반 플래너는 어떤 엔지니어도 예상하지 못한 동작을 제안할 수 있다. 스케줄링 에이전트는 여러 제조사의 로봇에 작업을 배정한다. 원격 콘솔은 다른 사업장에서 조작된다. 어느 경우든 결정하는 주체와 움직이는 기구 사이에는 네트워크나 조직의 경계, 또는 출력을 세부까지 예측할 수 없는 모델이 놓여 있다.

여기서 두 가지 요구가 생긴다. 첫째, 로봇은 명령의 출처와 무관하게, 인가의 증명이 없는 명령을 거부할 수단을 가져야 한다. 둘째, 조직은 사고 조사, 보험 청구, 규제 당국의 질의처럼 조직 자신이 검증 대상이 되는 상황에서도 신뢰를 유지하는 인가 기록을 가져야 한다.

PAIL은 이 두 가지를 제공하도록 설계되었다. PAIL은 물리적 행동에 대한 명시적인 권한 단위를 도입하고, 명령이 구동부에 도달하는 지점에서 이를 집행하며, 운영자나 공급자를 신뢰하지 않고도 확인할 수 있는 증거를 만든다.

2 문제 정의

2.1 출처 없는 명령

ROS 2 토픽의 속도 명령, 제조사 API의 동작 요청, 제어 소켓의 한 줄은 파라미터만을 담는다. 누가 요청했는지, 어느 작업에 속하는지, 언제까지 유효한지, 어떤 한도가 적용되는지는 담지 않는다. 제어기는 명령이 도착했기 때문에 실행한다. 의도하지 않은 노드가 명령 토픽에 발행하거나 플래너가 작업 범위를 크게 벗어난 동작을 내보내도, 로봇에는 그 명령을 정당한 명령과 구별할 근거가 없다.

사후에 남는 것은 대개 문제가 된 바로 그 시스템이 작성한 애플리케이션 로그다. 이러한 로그는 불완전할 수 있고 수정될 수 있으며, 어떤 항목이 주장하는 시점에 실제로 존재했음을 로그만으로는 보일 수 없다.

2.2 기존 계층이 제공하는 것

인접한 문제를 다루는 기존 수단이 여럿 있다. PAIL은 이들을 대체하지 않고 함께 동작하도록 설계되었다.

계층	답하는 질문	남는 공백
전송 보안 (예: SROS2가 사용하는 DDS Security)	어느 참여자가 어느 토픽에 발행·구독할 수 있는가, 메시지가 전송 구간에서 진본인가	특정 명령 하나가 특정 작업에 허가된 범위 안에 있는가
기능 안전 (안전 제어기, 보호 정지, 속도·이격 감시)	기계의 상태가 물리적으로 안전한가	그 동작을 누가 허가했으며 승인된 작업에 속하는가
플릿 관리와 접근 제어	어느 사용자나 시스템이 로그인하여 작업을 배정할 수 있는가	배정에서 구동부까지 사이에 무슨 일이 일어나며 사후에 무엇을 증명할 수 있는가
애플리케이션 로그	애플리케이션이 기록하기로 한 것	기록의 무결성, 완전성, 독립성

공백은 "이 참여자는 이 채널에서 통신할 수 있다"와 "이 기계는 안전한 상태에 있다" 사이에 있다. 개별 행동에 대한 인가가 작업과 한도에 결속되고, 실행 지점에서 집행되며, 타인이 검증할 수 있는 형태로 기록되는 것이다.

2.3 요구사항

PAIL의 설계는 다음 여섯 가지 요구사항에서 출발한다.

1. **행동 단위의 권한.** 인가는 행동 단위로 표현되어야 하며 속도, 지속 시간, 대상과 같은 명시적 한도와 명시적인 유효 기간을 가져야 한다.
2. **명령 경로에서의 집행.** 검사는 명령이 구동부에 도달하는 지점에 놓여야 하며, 인가가 없으면 동작이 없어야 한다.
3. **제3자가 검증할 수 있는 증거.** 인가 결정의 기록은 변조를 탐지할 수 있어야 하고, 운영자도 공급자도 신뢰하지 않는 당사자가 검증할 수 있어야 한다.
4. **통신 두절 시의 유한한 동작.** 인가 서비스와 연결이 끊긴 로봇은 무한정 계속 동작해서도 안 되고, 제어 주기마다 연결에 의존해서도 안 된다.
5. **로봇 스택을 교체하지 않는 통합.** 기존 명령 인터페이스에 부착되어야 하며 계획, 제어, 안전 기능은 그대로 두어야 한다.
6. **관측과 인가의 분리.** 모니터링 데이터는 인가 결정의 입력이 되어서는 안 된다.

---

3 설계 원칙

권한은 명시적이고, 범위가 정해져 있으며, 유효 기간이 한정된다. 행동은 로봇, 소속 계획과 단계, 행동 종류, 적용 한도, 유효 기간을 명시한 토큰으로 인가된다. 토큰 하나는 단계 하나를 한 번 인가한다.

검증자는 서명된 바이트를 그대로 검사한다. 게이트는 이슈어가 서명한 바이트열 그대로에 대해 서명을 검증하며, 검사 전에 메시지를 다시 직렬화하지 않는다. 권한을 전달만 하는 구성요소는 토큰의 내용을 바꾸거나 새로 만들 수 없다.

기록이 효력에 앞선다. 인가는 발급 서비스가 결과를 반환하기 전에 저널에 추가된다. 기록할 수 없으면 인가는 거부된다. 형식 검사를 통과한 요청은 거부되어도 기록된다.

기본값은 거부다. 권한이 없거나, 만료되었거나, 이미 소비되었거나, 형식이 잘못되었거나, 범위를 벗어나면 사유를 명시한 거부가 된다. 정책을 넘는 요청은 범위에 맞게 조용히 줄여지지 않고 거부된다.

로봇은 발급 키를 갖지 않는다. 서명 키는 발급 서비스에 남는다. 로봇 위의 구성요소는 권한을 검증할 뿐 만들어 낼 수 없다.

관측은 인가하지 않는다. 텔레메트리, 대시보드, 알림은 인가 기록과 로봇의 보고에서 파생된다. 관측 경로의 어떤 요소도 결정으로 되돌아가지 않는다.

증거는 독립적으로 확인할 수 있다. 체크포인트, 앵커 기록, 컨텍스트 해시의 형식은 바이트 수준에서 규정되어 고정된 시험 벡터와 함께 공개되며, 공개된 도구가 이를 검증한다. 그 가운데 앵커 증명과 컨텍스트 해시를 확인하는 도구는 기록을 만드는 구성요소와 코드를 공유하지 않는다.

---

4 구조

PAIL은 발급 서비스, 로봇 위의 구성요소들, 증거 체계, 관측 경로로 이루어진다. 그림 1은 구성요소와 이들을 잇는 두 경로를 보인다.

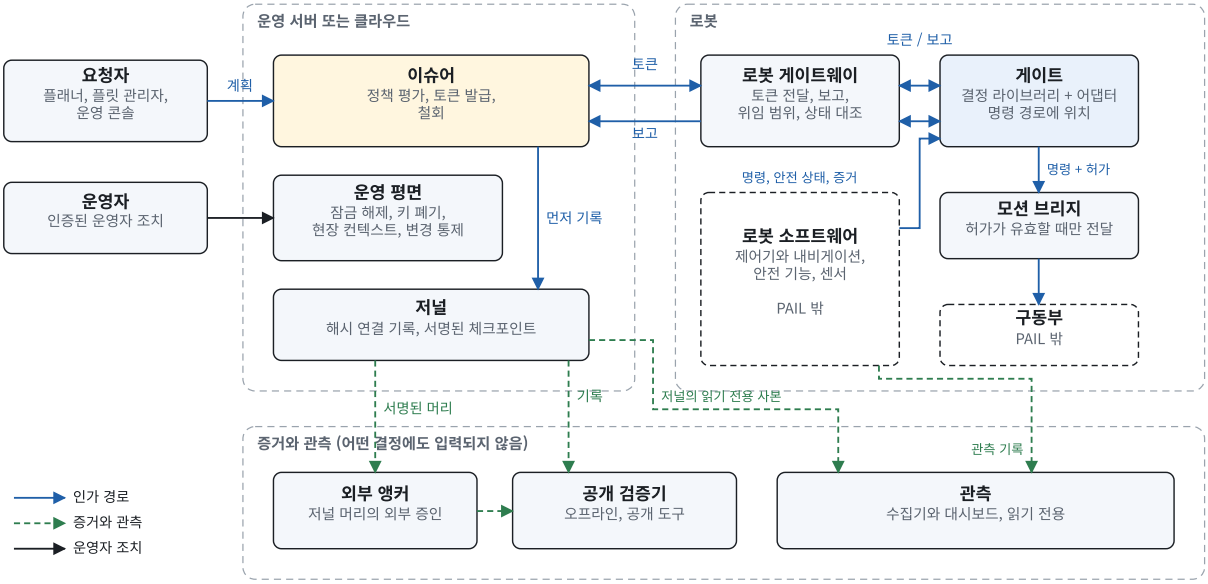

그림 1. PAIL의 구조. 인가 경로(실선)는 이슈어에서 명령 경로의 게이트로 권한을 전달한다. 관측 경로(점선)는 기록과 보고에서 파생되며 어떤 결정에도 입력되지 않는다.

4.1 구성요소

구성요소	위치	역할
이슈어	운영 서버 또는 클라우드	요청된 계획을 정책, 스킬 정의, 현재의 현장 컨텍스트에 대조하여 평가하고, 서명된 행동 토큰을 발급하며, 저널을 기록하고, 철회를 적용한다
로봇 게이트웨이	로봇	게이트를 위해 토큰을 확보하고, 연결이 없는 동안을 위한 유한한 위임 범위를 유지하며, 결정 보고를 전달하고, 장애 뒤 이슈어와 상태를 대조한다. 발급 키를 갖지 않는다
게이트	로봇의 명령 경로	C ABI를 갖는 결정 라이브러리와 로봇의 명령 인터페이스에 맞춘 어댑터. 모든 명령을 현재 토큰과 로봇의 안전·증거 입력에 대조하여 평가한다
모션 브리지	로봇, 구동부 직전	유효한 허가 신호가 있는 동안에만 명령을 전달하고, 허가가 끊기면 정지를 출력한다
저널	이슈어	인가 결정, 게이트 결정, 운영자 조치를 담는 추가 전용 해시 연결 기록. 서명된 체크포인트를 포함한다
앵커	운영자의 외부	서명된 저널 머리를 제출하는 외부 기준점. 모순 없이는 저널을 다시 쓸 수 없게 한다
검증기	어디서나, 오프라인	저널 무결성, 체크포인트 서명, 앵커 증명, 토큰의 컨텍스트 해시를 확인하는 공개 도구
관측	모니터링 망	인가 상태, 신선도, 불확실성을 표시하는 수집기와 대시보드. 읽기 전용이다
운영 평면	운영 서버	운영자 기능. 안전 정지 뒤의 잠금 해제, 키 폐기, 현장 컨텍스트 사건, 정비 잠금, 통제된 소프트웨어·정책 변경

4.2 두 개의 경로

인가 경로는 요청자에서 이슈어로, 이슈어에서 로봇 게이트웨이를 거쳐 게이트로, 게이트에서 모션 브리지를 거쳐 구동부로 이어진다. 결정 보고는 같은 경로를 거슬러 저널에 이른다. 이 경로 위의 모든 요소는 권한을 만들거나(이슈어), 검증하거나(게이트), 토큰의 내용을 바꿀 능력 없이 전달한다(게이트웨이. 게이트의 서명 검증이 토큰 내용의 변경을 막는다). 다만 토큰이 지금 활성인지와 현재의 컨텍스트 해시는 게이트웨이를 거쳐 게이트에 전달되며, 이 부분에 대한 신뢰의 상한은 토큰의 만료 시각이다.

관측 경로는 분리되어 있다. 수집기는 저널의 사본과 로봇에서 만들어진 관측 기록을 읽어 검증할 수 있는 것을 검증하고, 그 결과를 출처와 함께 제시한다. 관측자를 위해 내보내는 알림은 사건의 사본이며, 게이트와 브리지와 어댑터는 이를 읽지 않는다. 이 분리는 설계의 구조적 속성이다. 관측 구성요소가 고장 나거나 지연되거나 침해되어도 로봇에 허용된 것은 바뀌지 않는다.

4.3 참조 구현

게이트는 서로 다른 명령 경로에 내장할 수 있도록 안정된 C 인터페이스를 갖는 결정 라이브러리로 제공된다. ROS 2용 참조 통합은 베이스 제어기 앞에 놓이는 게이트 노드와 모션 브리지로 이루어진다. 두 번째 실험적 어댑터는 같은 결정 라이브러리를 ROS가 아닌 제어 소켓 앞에 두는 소프트웨어 전용 프리뷰다. 그 밖의 어댑터가 해야 할 일은 어댑터 계약과 적합성 시험 모음이 규정한다.

5 행동 토큰

5.1 내용

행동 토큰은 간결한 서명 메시지다. 필드는 여섯 묶음으로 나뉜다.

묶음	필드	목적
식별	토큰 식별자, 체인 식별자, 순번	토큰과 계획 안에서의 위치를 지칭한다
결속	로봇, 계획, 단계, 대상 객체	권한을 로봇 하나, 작업 단계 하나, 해당하는 경우 객체 하나에 묶는다
유효 기간	발급 시각, 시작 시각, 만료 시각	토큰을 정해진 시간 구간으로 제한한다. 구간의 길이는 정책이 정한다 (기본 1시간)
행동	행동 종류, 스킬과 슬롯	인가되는 행동의 종류를 명시한다
한도	최대 선속도, 최대 각속도, 최대 지속 시간, 가드레일	실행의 범위를 정한다. 모든 명령이 이에 대조된다
컨텍스트	정책 해시, 컨텍스트 스냅숏 해시, 스키마 판	권한을 정책과, 권한이 부여된 시점의 현장 조건에 묶는다

토큰은 Ed25519로 서명된다. 게이트는 원본 바이트에 대해 서명을 검증하며, 교체와 폐기를 지원하는 서명된 키 집합을 통해 서명 키를 확인한다. 키의 폐기는 게이트가 그 폐기를 담은 키 집합을 적재할 때 효력을 갖는다.

5.2 일회 사용

토큰 하나는 단계 하나를 한 번 인가한다. 게이트는 검증된 토큰 바이트에서 소비 키를 도출하고, 이미 소비한 토큰을 거부한다. 이슈어는 수신한 완료 보고가 진행 중인 단계의 토큰을 지칭하고 스킬의 완료 규칙과 일치할 때에만 계획을 진행시킨다. 따라서 한때 유효했던 토큰의 재사용은 양쪽 끝에서 거부된다.

5.3 계획과 체인

요청자는 순서가 있는 단계들로 이루어진 계획을 제출한다. 계획이 수락되면 이슈어는 **체인**을 만든다. 단계마다 토큰 하나가 있으며, 게이트에는 순서대로 하나씩 제공된다. 게이트는 현재 단계의 토큰을 보유하며, 단계가 완료되고 그 완료가 기록되면 체인이 진행한다. 체인은 완료, 요청자의 취소, 같은 로봇에 대한 새 계획의 수락에 따른 대체, 또는 철회로 끝난다.

5.4 컨텍스트와의 결속

인가는 변할 수 있는 조건에 의존한다. 사람이 구역에 들어오거나, 로봇이 정비 상태가 되거나, 구역이 통제되는 경우다. 이슈어는 현장의 모델을 유지하며, 발급 시점에 그 모델 가운데 해당 인가가 의존하는 부분의 해시를 기록한다. 게이트는 권한을 재확인할 때 이 값을 현재 값과 비교한다. 토큰이 부여된 조건이 더 이상 성립하지 않으면 권한은 회수되고 해당 체인은 철회된다. 이 해시의 계산 방법은 공개 규격에 정의되어 있으며, 공개 도구가 현장 모델과 저널로부터 이를 다시 계산한다.

6 행동 하나의 일생

그림 2는 계획 하나를 요청에서 기록까지 따라간다.

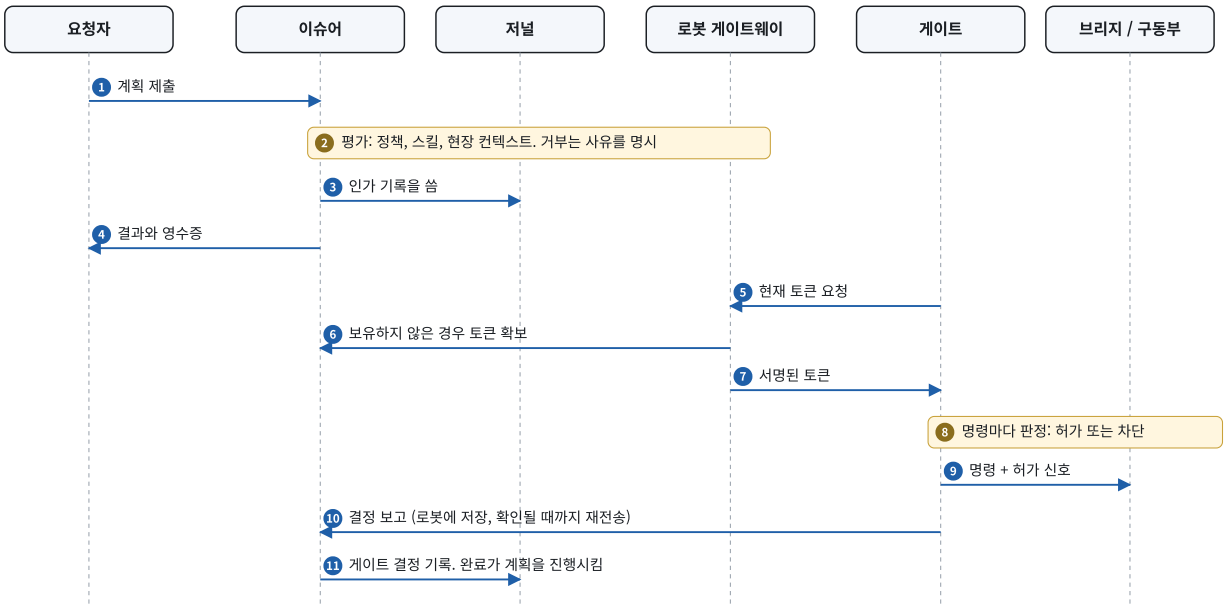

그림 2. 인가된 행동의 일생. 어느 단계에서든 거부는 사유를 명시하며 저널에 기록된다.

요청. 요청자(플래너, 플릿 관리자, 운영 콘솔)가 인증된 채널로 이슈어에 계획을 제출한다. 요청자의 신원은 요청의 내용이 아니라 채널에서 얻는다. 상태를 바꾸는 요청은 작업 식별자를 담고 영수증을 돌려받으므로, 응답이 유실된 뒤의 재시도는 두 번째 결과를 만들지 않고 원래의 판정과 체인 식별자를 돌려받는다.

평가. 이슈어는 계획을 요청자의 등록된 권한, 정책 한도, 로봇 종류에 대한 스킬 정의, 현재의 현장 컨텍스트에 대조한다. 한도를 어기는 계획은 거부되며, 거부는 어긴 한도를 명시한다. 이슈어는 요청을 수락 가능하게 줄이지 않는다. 3미터가 허용되는 곳에서 30미터를 요청한 플래너는 요청하지 않은 3미터 동작이 아니라 거부를 받는다.

기록 후 발급. 수락된 계획에 대해 이슈어는 인가 기록을 저널에 쓰고 나서야 결과를 반환한다. 저널에 쓰지 못하면 발급도 하지 못한다. 형식 검사를 통과한 뒤 거부된 계획도 기록되므로 저널에는 부여뿐 아니라 시도가 남는다.

전달. 게이트는 로봇 게이트웨이를 통해 현재 토큰을 요청한다. 게이트웨이는 이미 보유한 권한으로 답할 수 있으면 그렇게 하고, 그렇지 않으면 이슈어에 묻는다. 토큰은 로봇 쪽에서 가져간다. 이슈어가 로봇에 권한을 밀어 넣지 않는다.

판정. 게이트는 들어오는 모든 명령에 대해 정해진 순서의 검사를 수행한다. 주요 검사는 차례로 자신의 시계가 정상인지, 로봇이 보고한 비상정지 상태, 토큰의 존재, 토큰의 서명과 유효 기간, 토큰의 컨텍스트와 현재 컨텍스트의 일치, 일회 사용, 행동 종류의 일치, 토큰의 한도와 명령 파라미터의 대조, 로봇이 대상에 관해 제공하는 증거의 신선도와 일관성, 안전 신호의 신선도다. 처음 실패한 검사가 결과와 사유를 정한다. 결과는 허가, 차단, 비상정지 상태 가운데 하나다.

허가와 전달. 허가인 경우 게이트는 명령을 변경 없이 전달하고 주기적인 허가 신호를 내보낸다. 모션 브리지는 그 신호가 유효한 동안에만 명령을 구동부로 전달한다. 신호가 멈추면 브리지는 정지를 출력한다. 참조 구현에서 게이트는 모든 명령을

활성 토큰에 대조하고, 100 ms마다 게이트웨이에 현재 권한을 재확인하며, 20 Hz로 허가 신호를 내보낸다. 브리지는 100 ms의 기한을 적용한다. 이 값들은 설정 기본값이다.

보고와 완료. 게이트는 게이트웨이를 통해 이슈어에 결정을 보고한다. 단계 완료나 치명적 상태처럼 권한을 바꾸는 보고는 전송에 앞서 로봇의 영속 저장소에 기록되며, 이슈어가 확인할 때까지 재전송된다. 이슈어는 보고된 게이트 결정을 저널에 기록하며, 완료처럼 권한을 바꾸는 보고는 기록에 성공해야 처리된다. 스킬에 정의된 완료 규칙과 일치하는 완료 보고가 체인을 진행시킨다.

---

7 장애 시 동작

명령 경로에 놓이는 계층은 정상 동작 못지않게 무언가 실패했을 때의 동작으로 규정된다.

7.1 기본값으로서의 거부

각 구성요소는 자신의 장애가 권한을 없애는 방향으로 작용하도록 만들어졌다. 토큰이 없는 게이트는 차단한다. 토큰을 검증할 수 없는 게이트는 차단한다. 유효한 허가 신호가 없는 브리지는 정지한다. 저널에 쓸 수 없는 이슈어는 발급하지 않는다. 보안 설정이 불완전한 배치는 운영 프로파일에서 기동을 거부한다.

7.2 연결 두절

로봇은 끊길 수 있는 무선 링크 위에서 동작한다. 제어 주기마다 이슈어와의 왕복을 요구하면 링크가 단일 장애점이 되고, 연결이 끊긴 로봇이 제한 없이 계속하도록 허용하면 인가의 목적이 무너진다.

PAIL은 이 문제에 한계를 둔다. 정책이 허용하면 이슈어는 앞으로의 단계들 가운데 제한된 범위를 미리 인가할 수 있고, 게이트웨이는 이를 위임된 권한으로 보유한다. 링크가 끊긴 동안 게이트웨이는 이 범위를 로컬 상태에서 제공하며, 게이트는 각 토큰을 평소와 똑같이 검증한다. 이 범위의 깊이가 로봇이 연락 없이 진행할 수 있는 거리다. 범위가 소진되었는데 이슈어에 닿을 수 없으면 게이트는 보유한 허가의 짧은 유예가 끝난 뒤 차단한다. 위임된 권한은 그 토큰이 만료될 때, 게이트가 만료를 보고할 때, 이슈어가 철회를 알릴 때, 재연결 뒤의 대조에서 체인이 종료되었음이 드러날 때 반납된다.

7.3 재시작과 대조

게이트웨이나 게이트가 재시작한 뒤에는 재시작 전에 보유한 권한을 유효하다고 가정하지 않는다. 로봇 쪽은 자신이 보유한 것과 진행한 지점을 선언하고, 이슈어는 토큰이 다시 제공되기 전에 판정으로 답한다. 더 이상 존재하지 않는 실행 인스턴스에 위임되었던 체인은 취소된다. 이슈어와 게이트웨이는 서로 같은 판의 통신 계약을 구현하는지도 확인하며, 기동 시점의 불일치는 공유되지 않은 가정 위에서 진행하는 대신 동작을 막는다.

7.4 위험 상태

게이트는 로봇 자체의 안전 기능으로부터 안전 상태 입력을 받는다. 이 입력이 위험 상태를 알리면 게이트는 허가를 중단하고, 해당 권한이 재사용될 수 없도록 래치하며, 로봇의 안전 계층을 향해 정지 요청을 내고, 그 상태를 치명적 상태로 보고한다. 치명 보고는 확인될 때까지 재전송되며, 보고가 도착하면 이슈어는 해당 로봇의 현재 체인을 철회한다. 로봇 쪽의 래치는 이슈어의 응답에 의존하지 않는다.

7.5 계획의 변경과 컨텍스트의 변경

로봇에 대해 새 계획이 수락되면 이전 체인은 대체되고 그 권한은 게이트웨이에서 회수된다. 현장 컨텍스트가 활성 인가에 영향을 주는 방식으로 변하면 해당 체인은 철회된다. 남은 단계는 플래너로 되돌아가지 않고 변경된 컨텍스트에 대해 다시 평가된다. 정책을 더 이상 만족하지 않는 단계는 다시 인가되지 않는다.

7.6 소프트웨어 무결성

게이트는 자신이 실행 중인 소프트웨어의 측정값을 보고할 수 있다. 이를 요구하는 배치에서 이슈어는 게이트가 유효한 최신 측정값을 제시하지 않은 로봇에 대해 발급과 제공을 보류한다. 이 방식은 선택 사항이며 기본값은 꺼짐이다.

---

8 증거

8.1 저널

저널은 각 기록이 직전 기록과 암호학적으로 연결된 추가 전용 기록열이다. 기록을 삭제, 삽입, 재배열, 변경하면 연결이 깨지고 검증에서 탐지된다. 일정 간격으로, 그리고 쓰기가 멈출 때와 저널을 닫을 때 서명된 체크포인트가 삽입된다. 이슈어는 기동할 때 이 서명을 검증하며, 검증되지 않는 저널로는 기동하지 않는다. 저널에는 수락된 계획과 거부된 계획, 발급된 토큰, 완료를 포함한 게이트 결정, 철회와 그 원인, 운영자 조치, 현장 컨텍스트의 변경이 담긴다.

증거로 쓰는 데에는 두 가지 속성이 중요하다. 첫째, 기록이 결과 반환에 앞서 쓰이므로 권한은 발급되었는데 기록은 없는 상태가 존재하지 않는다. 이는 발급 프로세스 수준의 순서 보증이며, 그 영속성의 한계는 제품의 보안 문서에 함께 명시되어 있다. 둘째, 거부된 요청은 부여의 기록 없이 시도의 기록만을 남기므로 "시도되었으나 거부됨"과 "요청된 적 없음"이 구별된다.

8.2 외부 앵커

해시 연결은 내부의 일관성을 증명할 뿐, 오늘 제시된 연결이 어제 존재했던 연결임을 증명하지는 않는다. 이를 위해 저널의 머리를 서명하여 운영자의 통제 밖에 있는 앵커에 주기적으로 제출한다. 제출은 비동기이며 발급을 막지 않는다. 실패한 제출은 대기열에 남아 재시도된다.

Glowfly는 제출된 머리를 투명성 로그에 보관하고, 서명된 체크포인트를 인증서 투명성에서 정의된 것과 같은 종류의 포함 증명·일관성 증명과 함께 돌려주는 호스팅 앵커 서비스를 운영한다. 평가용 배치는 같은 클라이언트 경로를 사용하는 로컬 파일 앵커를 쓸 수 있다. 로컬 앵커는 독립된 증인이 아니며, 문서가 이를 명시한다.

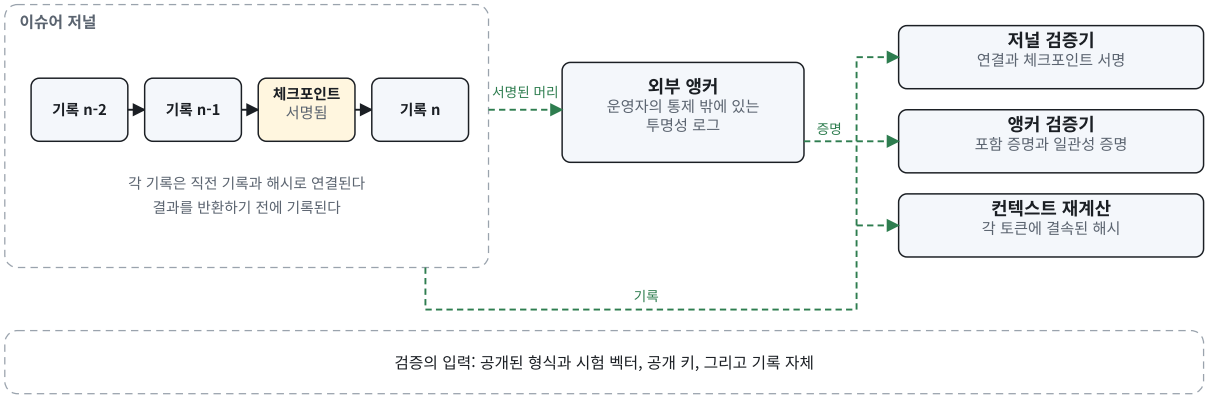

그림 3. 증거 체계. 검증에는 공개된 형식, 공개 키, 그리고 기록 자체만 필요하다.

8.3 독립 검증

체크포인트, 앵커 기록, 머클 증명, 컨텍스트 해시의 정규 바이트 인코딩과 해시 구성은 판 이름과 고정된 시험 벡터 집합을 갖는 규격으로 공개된다. 검증 도구는 셋이다. 하나는 저널의 연결과 체크포인트 서명을 확인하며, 이슈어의 공개 키를 받아 모든 기록이 서명된 체크포인트 아래에 있을 때에만 통과를 낸다. 다른 하나는 앵커 체크포인트와 그 포함·일관성 증명을 확인한다. 세 번째는 토큰에 결속된 컨텍스트 해시를 현장 모델과 저널로부터 다시 계산한다. 뒤의 두 도구는 기록을 쓰는 구성요소와 코드를 공유하지 않으며, 두 번째 구현을 만들려는 당사자는 벡터를 기준으로 이를 만들 수 있다.

8.4 증거가 보이는 것과 보이지 않는 것

저널은 누가 행동을 요청했는지, 무엇이 어떤 한도로 인가되었는지, 게이트가 각 명령에 대해 무엇을 결정했는지, 이 사건들이 어떤 순서로 일어났는지를 탐지 없이 바꿀 수 없는 형태로 보인다. 저널은 물리적 정지가 일어났음을, 로봇의 센서가 옳았음을, 인가된 행동이 취해야 할 옳은 행동이었음을 보이지 않는다. PAIL은 권한과 결정을 기록한다. 물리적 결과는 로봇 자체의 시스템이 확정한다.

---

9 신뢰 경계

PAIL은 자신이 지키는 것을 반박 가능한 주장으로 명시하고, 위임하는 것을 따로 명시한다.

9.1 네 가지 주장

주장	내용	다음이 성립하면 거짓이다
원본 바이트 서명 검증	게이트는 이슈어가 서명한 바이트 그대로에 대해 재직렬화 없이 서명을 검증한다	어떤 필드를 바꾸어도 토큰이 검증을 통과한다
이슈어 선기록	인가 기록은 발급 응답에 앞서 저널에 추가된다	저널 쓰기를 실패시킨 상태에서 토큰이 발급된다
일회 사용	토큰 하나는 단계 하나를 한 번 인가한다	소비된 토큰이 다시 수락된다
거부 기본값의 전파	판정의 기본값은 거부다. 검증되지 않은 입력은 허가로 평가될 수 없으며, 게이트는 명령마다 활성 토큰을 다시 검증한다	검증되지 않은 입력이 허가로 판정된다

9.2 위임된 책임

인가 계층이 제공할 수 없는 속성이 있으며, 이들은 통합자 또는 운영자에게 명시적으로 배정된다. 주요한 것은 다음과 같다.

- **로봇 내부의 전송 진본성.** 게이트와 브리지 사이 허가 신호의 진본성은 로봇 미들웨어의 전송 보안에 의존한다. ROS 2 배치에서는 강제 모드의 SROS2를 뜻하며, 이것이 없으면 허가 메시지가 위조될 수 있다.
- **우회 경로의 차단.** 게이트는 자신이 놓인 명령 경로를 지킨다. 구동부로 가는 다른 경로는 통합자가 막아야 한다.
- **물리적 안전.** 보호 정지, 안전 등급의 감시, 최종적인 물리적 정지는 로봇의 안전 시스템에 속한다.
- **키의 보관.** 서명 키의 물리적·절차적 보호는 배치하는 조직에 속한다. PAIL은 서명 키를 외부 키 저장소에 두는 구성을 지원한다.
- **시간.** 유효 기간은 합리적으로 동기화된 시계에 의존한다. 게이트는 설정된 허용 범위를 넘는 시계 이상을 감지하면 동작을 거부한다.

이 밖의 위임 항목과 위협 모델은 공개 문서에 정리되어 있다.

---

10 통합

10.1 어댑터 계약

PAIL을 로봇에 통합한다는 것은 결정 라이브러리를 둘러싼 **어댑터**를 작성하거나 설정하는 일이다. 어댑터의 책임은 네 가지다. 모든 명령을 게이트로 통과시키고, 허가된 명령을 변경하지 않은 채 게이트의 결정대로 행동하며, 로봇의 안전 상태와 증거 입력을 게이트에 공급하고, 허가가 끊기면 전달을 멈춘다. 이 계약은 규격으로 정해져 있으며, 적합성 시험 모음이 정의된 사례들로 어댑터를 시험한다. 토큰 없는 명령의 거부, 한도를 벗어난 명령의 거부, 게이트웨이에 닿지 못할 때의 동작, 어댑터 자체가 실패할 때의 동작이 그 사례에 포함된다.

10.2 SROS2와의 관계

PAIL과 SROS2는 서로 다른 계층에서 동작하며 함께 사용된다. SROS2는 어느 참여자가 어느 토픽에서 통신할 수 있는지를 정하고 전송 중인 메시지를 보호한다. PAIL은 주어진 명령이 주어진 단계에 인가된 범위 안에 있는지를 정하고 그 판정을

기록한다. PAIL은 로봇 내부 망의 메시지 진본성을 SROS2에 의존하며 이를 직접 제공하려 하지 않는다.

10.3 요청자

동작을 요청하는 시스템은 공개된 프로토콜 정의에 따른 gRPC 인터페이스로 통합한다. 계획을 제출하고 취소하며, 사유와 호출자가 취해야 할 처분과 재시도가 의미를 갖는 가장 이른 시점을 담은 구조화된 거부를 받는다. 요청자는 로봇의 내부 인터페이스를 알 필요가 없다.

---

11 운영

가동 중인 시설에서의 인가에는 운영자 기능이 필요하며, 이 기능들도 발급과 같은 규율을 따른다. 모든 조치는 인증되고 저널에 기록되며, 잠금 해제와 키 폐기는 운영자 키의 서명을, 보류 해제는 승인자의 서명된 승인을 요구한다.

- **안전 정지와 잠금 해제.** 위조되거나 형식이 잘못된 권한을 가리키는 보고, 또는 이슈어 자신의 저널 장애는 이슈어를 새 계획을 받지 않는 안전 정지 상태로 만든다. 이 상태를 벗어나려면 운영자 키로 서명되고 짧은 시간 동안만 유효하며 한 번만 쓸 수 있는 잠금 해제가 필요하다.
- **키 폐기.** 발급 키를 폐기하면 이슈어 전체가 영구 잠금 상태가 되며, 다시 시작해도 잠금이 유지된다. 게이트는 그 키의 폐기를 담은 새 키 집합을 적재한 뒤부터 그 키를 거부하므로, 폐기 뒤에는 새 키 집합을 모든 게이트에 적용한다.
- **현장 컨텍스트 사건.** 운영자와 현장 시스템은 사람이 구역에 들어오거나 로봇이 정비에 들어가는 것과 같은 조건을 이슈어에 알린다. 이 사건들은 인가가 의존하는 현장 모델을 갱신한다.
- **정비 잠금.** 로봇은 한 명 이상의 보유자에 의해 정비 잠금될 수 있다. 모든 잠금이 해제되어야 그 로봇에 대한 발급이 재개된다.
- **통제된 변경.** 로봇의 소프트웨어와 설정 변경은 보류 상태에서 수행할 수 있다. 보류 중에는 해당 로봇에 대한 발급이 중지되고, 실행기가 제출한 정지 증거가 기록에 묶이며(물리적 정지의 증명은 아니다), 해제에는 승인이 필요하다. 정책과 현장 모델의 변경은 서로 다른 두 당사자의 승인을 받도록 할 수 있으며, 이슈어가 기동 시 이를 검증한다.

이 기능들의 계약은 규격의 일부다. 일부는 선택 사항이며 설정으로 켠다.

---

12 관측

관측 경로는 운영자에게 인가 상태를 보여 주되 그 일부가 되지 않는다. 수집기는 이슈어가 저널을 계속 쓰는 동안 그 사본을 검증하고, 로봇의 관측 기록을 서명이 검증된 것, 인증된 연결로 올라온 것, 자기 보고로 구별하여 받아들이며, 결과를 읽기 전용 인터페이스와 대시보드로 제공한다.

이 설계는 불확실성을 특히 주의하여 다룬다. 표시되는 각 값에는 그 값을 최신으로 보는 근거가 따라붙는다. 수신 시각이나 로봇이 응답한 도전값으로 확정된 한계가 그 예다. 근거가 만료된 값은 마지막으로 알려진 값이 아니라 오래됨 또는 알 수 없음으로 표시된다. 관측 기록은 자신이 물리적 정지를 증명하지 않는다고 명시한다. 수집기가 고장 나도 로봇은 보유한 권한으로 계속 동작한다. 수집기의 어떤 것도 권한을 부여하거나 회수할 수 없다.

13 측정 결과

다음 결과는 릴리스와 함께 공개된 측정 보고서에서 가져왔다. 각 결과는 명시된 구성과 명시된 빌드에 대해서만 유효하다.

플랫폼	범위	결과
소프트웨어 꾸러미 (Docker, Linux arm64·amd64)	모의 TurtleBot3 베이스, 비 ROS 어댑터 프리뷰, 다중 로봇 플릿, 게이트 단위 시험, 수명 주기, 계획 대체	0.2.0 배포물에서 TurtleBot3 모의 시험 8개 점검 통과 · 어댑터 프리뷰 15/15 · 플릿(로봇 2대) 각 15/15 · 게이트 시험 37/37 · 수명 주기 21/21 · 계획 대체 8/8
TurtleBot3 실물 (바퀴를 띄운 상태)	엔코더 기반 관측과 수락 판정	시험판 빌드에서 13개 장면 전부 통과, 3회 반복. 그 빌드에 한정된다
Jetson Orin Nano Super (CPU만 사용)	실제 게이트와 브리지를 쓴 합성 참조 구성	7개 시나리오 전부 통과. 8시간 분할 반복 시험에서 실행·취소 752회, 주입한 결함 22건, 정리 전부 통과
Isaac Sim (L40S 가상 머신 1대)	시뮬레이터 SDK 종단에 대한 우회·수명 주기 시험 행렬	릴리스 실행의 초기 상태 반복 2회에서 14개 사례 모두 기대한 결과: 13건 통과, 우회를 탐지해야 하는 양성 대조 1건 탐지. 이전 시험판의 실행 한 번에서 판정 불가 1건이 있었으며 이를 공개 문서에 밝혔다

게이트와 브리지가 속도 명령 경로에 더하는 지연은, 직접 구독자와 비교하여, 2026년 10월 16일에 SROS2 강제 모드를 켠 상태로 측정했다.

플랫폼	명령 주기	중앙값	99백분위
Jetson Orin Nano Super	10, 20, 50 Hz	3.8~4.3 ms	5.4~7.0 ms
TurtleBot3 (Raspberry Pi 3B+), 바퀴를 띄운 상태	10, 20 Hz	14.3~14.7 ms	21.9~23.4 ms

이 수치는 정상 상태에서 명령이 입력 토픽에서 구동 토픽까지 지나는 시간이다. 같은 게이트와 브리지의 이전 빌드로 얻은 값이며 현재 릴리스에서 다시 측정하지 않았다. TurtleBot3의 수치는 측정 프로브가 프로세서를 함께 사용했으므로 상한이다. 드라이버, 펌웨어, 모터 제어 루프는 포함하지 않으며 PAIL은 이들에 관여하지 않는다. 이 수치는 지연에 대한 보증이 아니다.

측정하지 않은 것: 위에 열거하지 않은 로봇, 바닥 주행, 다른 시뮬레이터나 가속기, 그리고 어떤 형태의 기능 안전 인증.

14 제공 형태

공개 Core. PAIL Core는 Apache License 2.0에 따라 소스로 공개된다. 게이트 결정 라이브러리, ROS 2 게이트와 모션 브리지, 로봇 게이트웨이, 어댑터, 프로토콜 정의와 클라이언트 예제, 저널 라이브러리, 관측 수집기와 대시보드, 검증기, 규격과 시험 벡터, 적합성 시험 모음으로 이루어진다.

이슈어. 이슈어는 연구, 개발, 상호운용성 시험, 내부 지속적 통합을 위한 평가용 실행본으로 제공된다. 소스는 공개하지 않는다. 운영 환경에서의 사용은 별도로 라이선스한다.

상용 제공. Glowfly는 공개 Core와 별도로 다음을 제공한다. 운영용 이슈어, 단일 노드의 영속성 특성을 수용할 수 없는 배치를 위한 복제 이슈어, 호스팅 앵커 서비스, 그리고 11절에 기술한 변경 통제 계약을 구동하는 플릿 운영 계층.

특허. Glowfly는 본 백서에 기술된 기술과 관련하여 등록된 특허와 출원 중인 특허를 보유하고 있다. Apache 라이선스가 적용되는 자료에 대한 특허 권리는 해당 라이선스 3조가 부여하는 바에 따른다.

규제에 관한 입장. PAIL은 어떤 규정이나 기능 안전 표준에 대한 적합성도 주장하지 않는다. PAIL은 적합성 평가, 사고 조사, 감사가 사용할 수 있는 기록을 만든다.

---

15 범위와 한계

PAIL은 요청된 행동이 인가되었는지를 판정한다. 그 행동이 현명한지는 판정하지 않는다. 올바르게 인가된 행동도 잘못된 행동일 수 있으며, 이를 막는 것은 계획, 인지, 안전 기능의 몫이다.

PAIL은 게이트가 설치된 명령 경로를 지킨다. 게이트를 우회하는 경로에 대해서는 보호를 제공하지 않으며, 로봇 내부의 메시지 진본성은 로봇 미들웨어의 전송 보안에 의존한다.

이슈어 선기록은 발급 프로세스 안에서의 순서 보증이다. 전원 상실 시의 동작은 배치의 영속성 설정에 따르며, 보안 문서가 이를 기술한다.

각 판의 알려진 한계는 보안 정책 문서가 밝힌다. 예를 들어 이전의 서명된 체크포인트 지점까지 잘라 낸 저널은 그 자체로는 유효한 기록으로 보이므로, 외부 앵커가 보관한 머리와 대조해야 드러난다. 발급 키의 폐기가 실행 중인 게이트에 닿으려면 새 키 집합의 적용이 필요하다는 점도 같은 문서에 운영 절차와 함께 적혀 있다.

13절의 측정은 정해진 조건의 소수 플랫폼에 한정된다. 다른 하드웨어, 다른 환경, 다른 규모에서의 동작을 확립하지 않는다.

---

16 결론

로봇 동작의 출처가 엔지니어가 만든 프로그램에서 생성된 계획으로 옮겨 가면서, 누가 행동을 허가했는지와 그것을 어떻게 보일 수 있는지는 그 자체로 시스템 요구사항이 된다. PAIL은 이를 요구사항으로 다룬다. 권한을 명시적이고 서명되었으며 한도가 정해진 객체로 만들고, 명령이 구동부에 도달하는 지점에서 그 권한을 집행하며 조건이 바뀌거나 구성요소가 실패하면 이를 회수하고, 각 결정을 효력이 생기기 전에 조직 밖의 당사자가 공개된 도구로 검증할 수 있는 형태로 기록한다.

이 접근은 의도적으로 좁다. PAIL은 전송 보안, 기능 안전, 플릿 관리를 대체하지 않으며, 자신이 지키는 속성과 통합자에게 남기는 속성을 명시한다. 그 경계 안에서 PAIL은 이 계층들이 제공하지 않는 것을 제공한다. 물리적 행동 하나하나에 대한 권한의 증명, 그리고 기록을 보유한 쪽을 신뢰하지 않아도 되는 그 증명의 증거다.

참고문헌

1. PAIL Core 규격, 판 PAIL-CORE-2026.1, 시험 벡터 집합 `bbddff16…`. 정규 형식, 어댑터 계약, 통신 계약, 시험 벡터를 포함한다. PAIL Core 소스와 함께 공개.
2. PAIL 신뢰 경계, 위협 모델, 보안 정책. PAIL Core 소스와 함께 공개.

3. Object Management Group, **DDS Security Specification**.
4. ROS 2 문서, **SROS2: security enclaves and access control**.
5. S. Josefsson, I. Liusvaara, **Edwards-Curve Digital Signature Algorithm (EdDSA)**, RFC 8032.
6. B. Laurie, A. Langley, E. Kasper, **Certificate Transparency**, RFC 6962.

문의

Glowfly Identity Systems Co., Ltd. · www.glowfly.ai

공개 소스에 관한 질문은 프로젝트의 이슈 트래커에서 다룬다. 취약점으로 의심되는 사항은 프로젝트 보안 정책에 적힌 주소로 신고한다.

© 2026 Glowfly Identity Systems Co., Ltd. PAIL Core는 Apache License 2.0에 따라 제공된다. 본 문서는 변경 없이 재배포할 수 있다.