

# Figure 『Man vs Machine』の深層

PRイベントの裏に隠されたヒューマノイド物流の実力と、  
次世代ベンチマークの真の課題

DATE: 2026.05.20

AUTHOR: INDEPENDENT AUDIT COMMITTEE



## 01. 華々しい成功（表層）

### 強力なPRイベントとしての 結実

- 人間 12,924個 vs ロボット 12,732個。
- 10時間の壁時計時間において、ヒューマノイドが人間のシフト平均スループットに肉薄したという事実を世界に強烈に印象付けた。



## 02. 検証の空白（限界）

### データの非対称性と 意図的な欠落

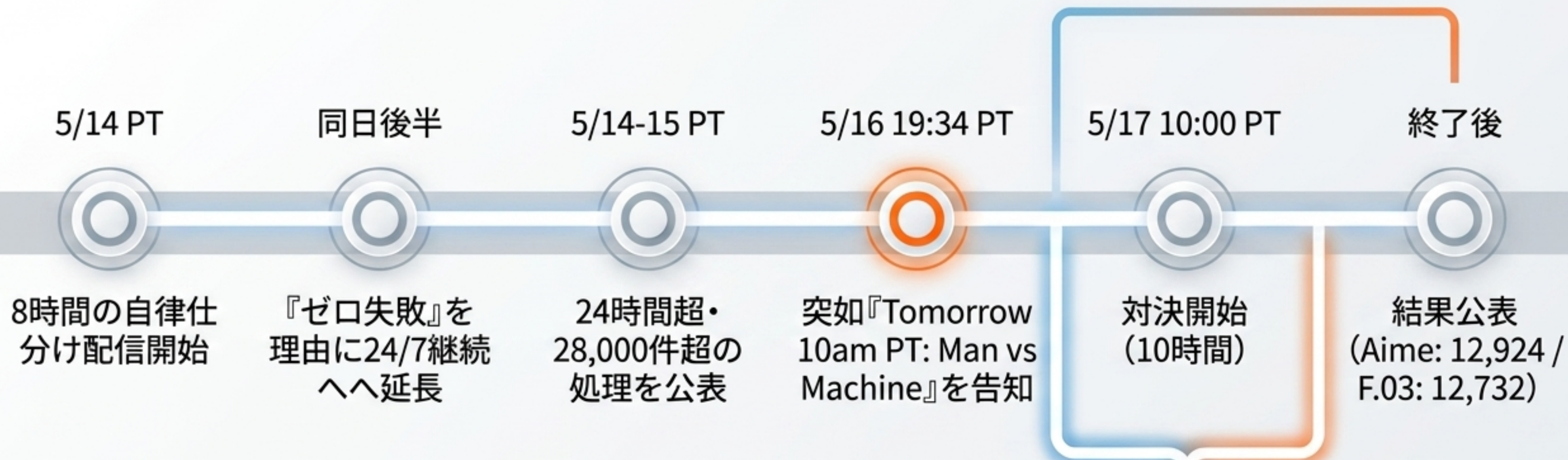
- 休憩条件の差異、複数機体の「群」運用の可能性、そして実効性に直結するエラー率（誤向き・落下・介入）が非開示。
- 厳密な実物流ベンチマークとしては成立していない。



## 03. 技術の真価（深層）

### 「個」から「群」への パラダイムシフト

- 真の技術的ブレイクスルーは、単体の瞬間速度ではない。
- 自己修復や交代ループを含む、システムとしての連続稼働能力（Fleet Reliability）の証明にある。



**監査インサイト：独立した対決イベントではなく、数日間の公開耐久デモの中に突如差し込まれた『特別企画』である。**

HUMAN (Aime)

12,924

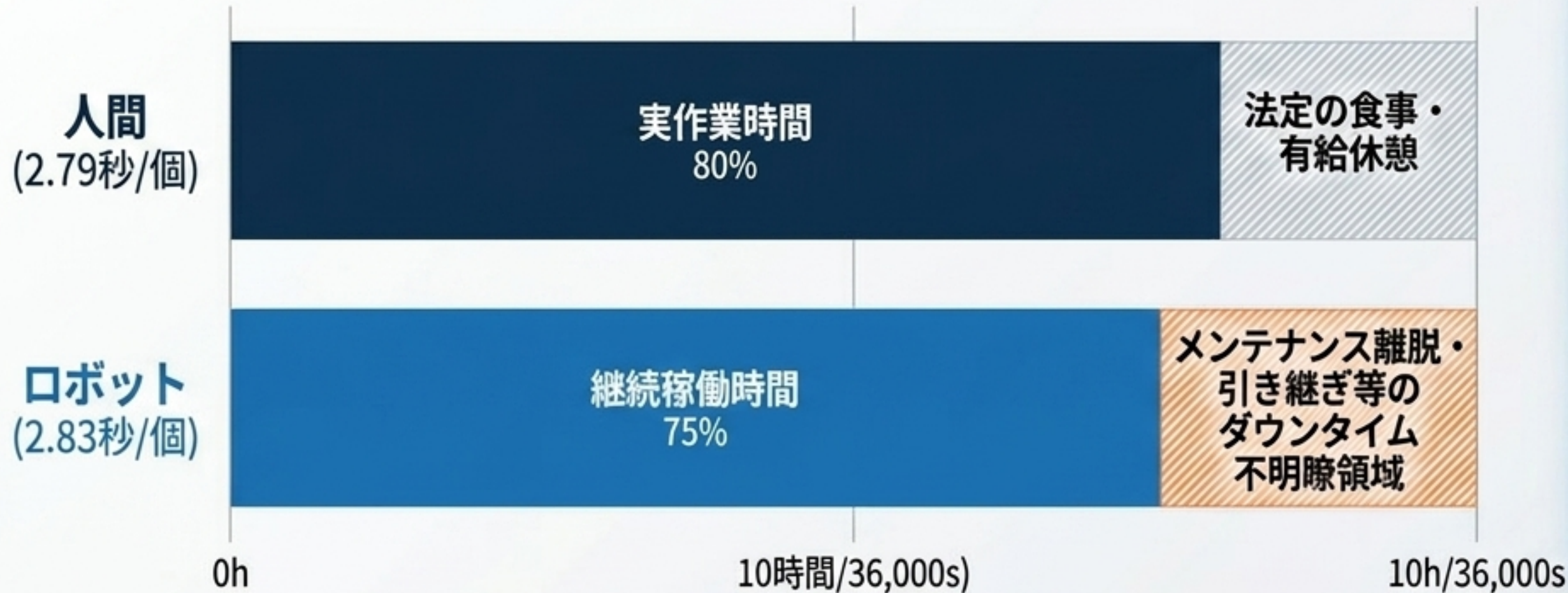
packages

差：わずか192個

MACHINE (F.03)



12,732

packages



公開された「2.8秒/個」は瞬発的なサイクルタイムではない。法定休憩などのダウンタイムを含んだ「壁時計時間(シフト全体)」の平均値である。

実作業時間あたりの真の生産性は比較不能。

評価項目	人間側 (Aime)	ロボット側 (F.03)	Audit Flag
参加者数	1名のインターン	単一機体か、チーム(群) 運用か曖昧	 非対称性の疑い
休憩・ダウンタイム	州法に基づく食事および有給休憩	休憩概念なし (別機体引き継ぎの可能性あり)	 比較不可能
作業セル環境	同一の荷物が循環するクローズドループコンベヤ	左に同じ	 実物流環境との乖離
介入・遠隔操作	不適用	No teleoperationを主張するが第三者監査なし	 証拠欠落

**監査判定：本イベントは「一般物流における普遍的勝敗」の証明ではなく、「Figureが設計した特定環境下での長時間稼働デモ」として評価すべきである。**

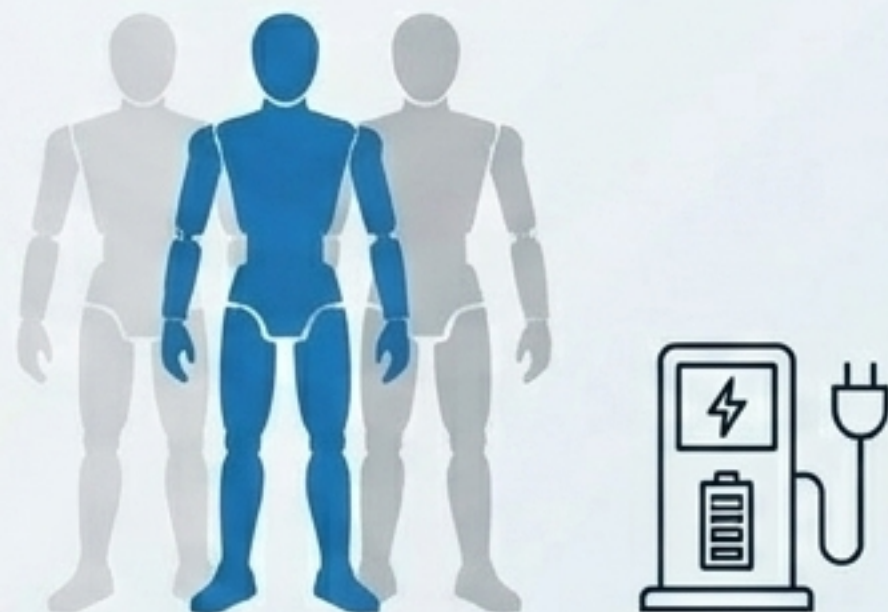
“ Figure公式ルールの記事：『the team that sorts the most packages wins』 ”

**HUMAN**



**HUMAN: 完全なる「個」**  
独立した1人のAime

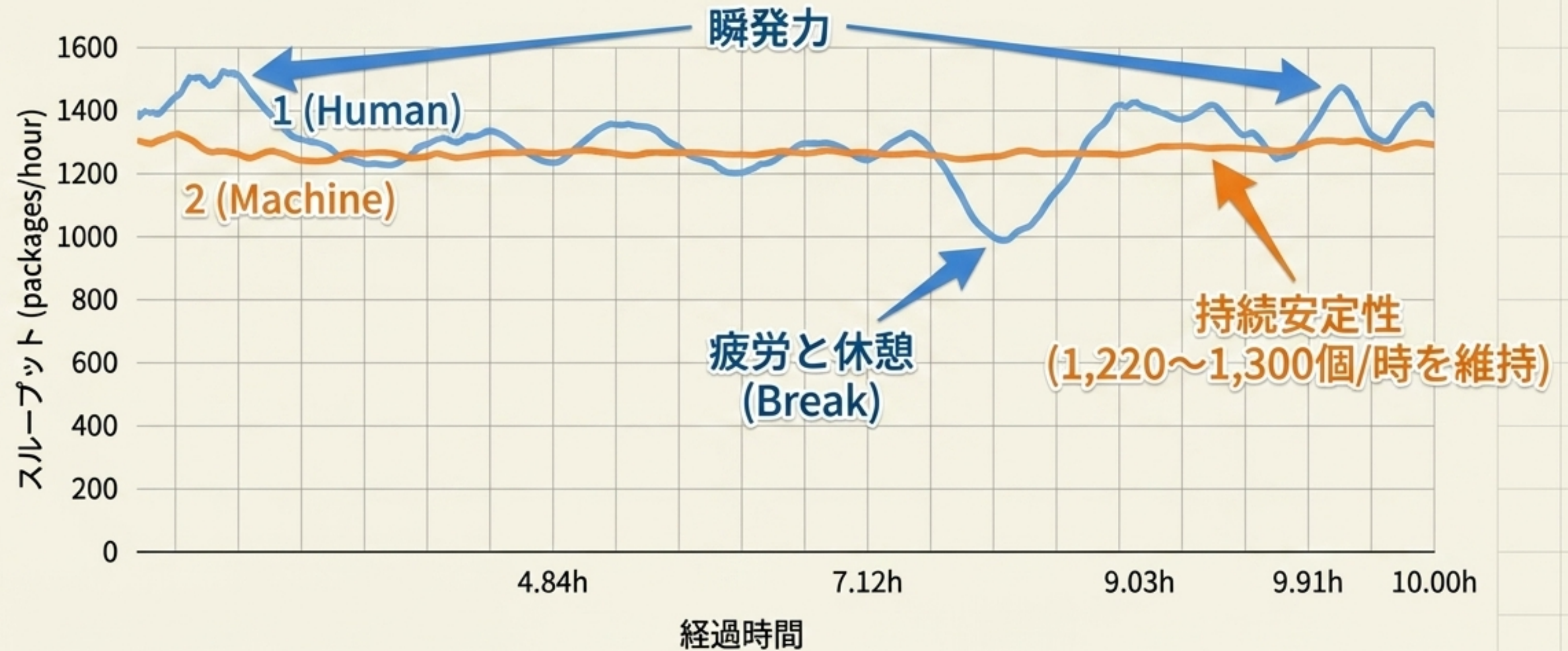
**MACHINE**



**MACHINE: 交代を前提とした「群」**  
F.03 + 待機・充電システム (Bob, Jim, Rose...)

ルール文面は「ロボット」ではなく「チーム」を主語としている。  
背景で充電待機する他機体の存在や、自動引き継ぎ機能の言及を考慮すると、  
これは「人間1人 vs ロボット運用システム全体」の対決であった可能性が高い。

## 対決中の近似区間平均スループット



Figureの狙い通り、「人間のボラティリティ (変動)」対「ロボットのサステナビリティ (持続性)」という明確なコントラストが描かれている。

## 開示された指標

総仕分け数:  
**12,732個**

シフト平均速度:  
**2.83秒/個**

全体のスピードのみが開示

⊘ REDACTED 総誤り率 (Total Error Rate)

⊘ REDACTED 誤向き率 (Misorientation Rate)

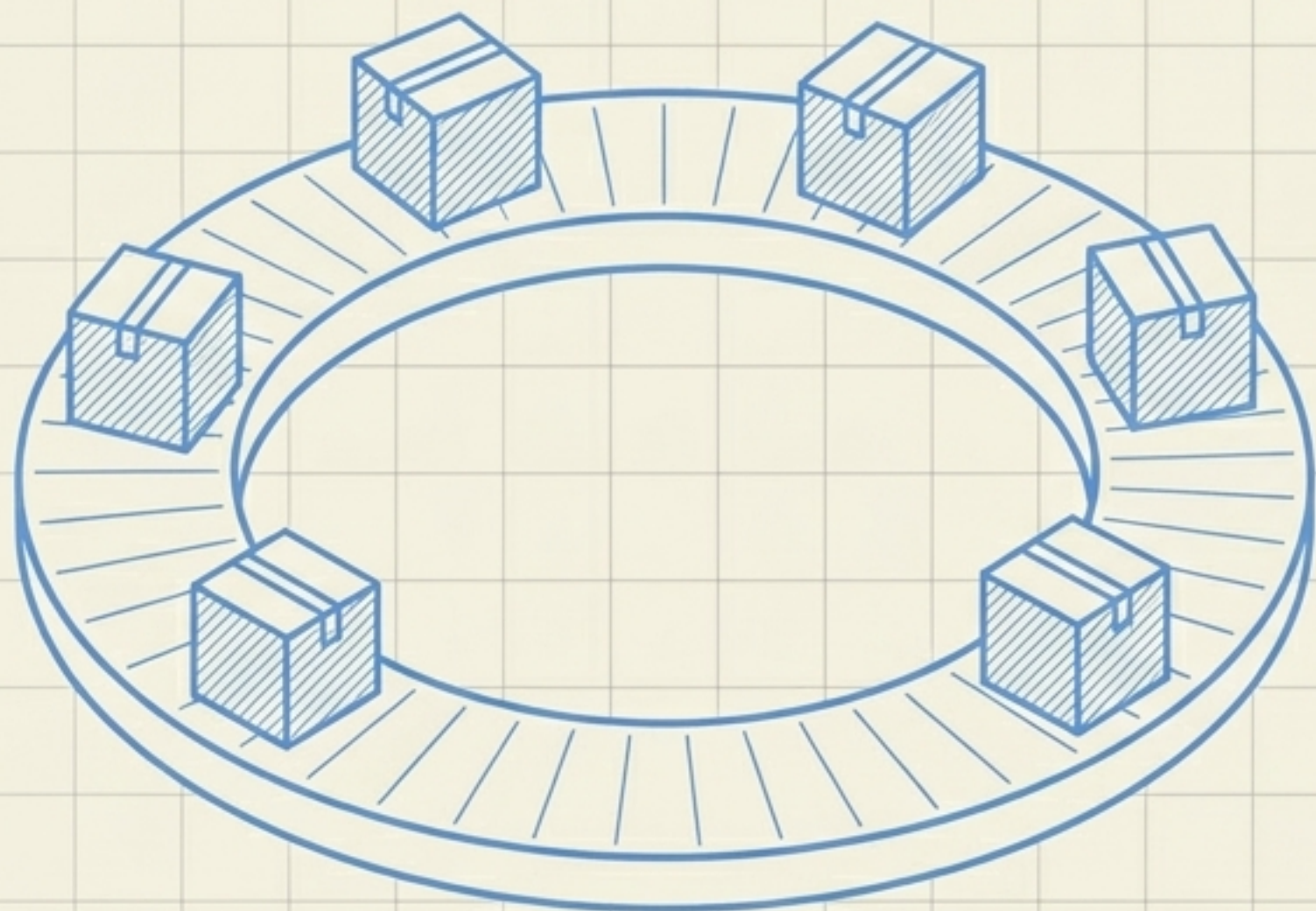
⊘ REDACTED 落下率 (Drop Rate)

⊘ REDACTED 介入率 (Intervention Rate)

### なぜこれが致命的か？

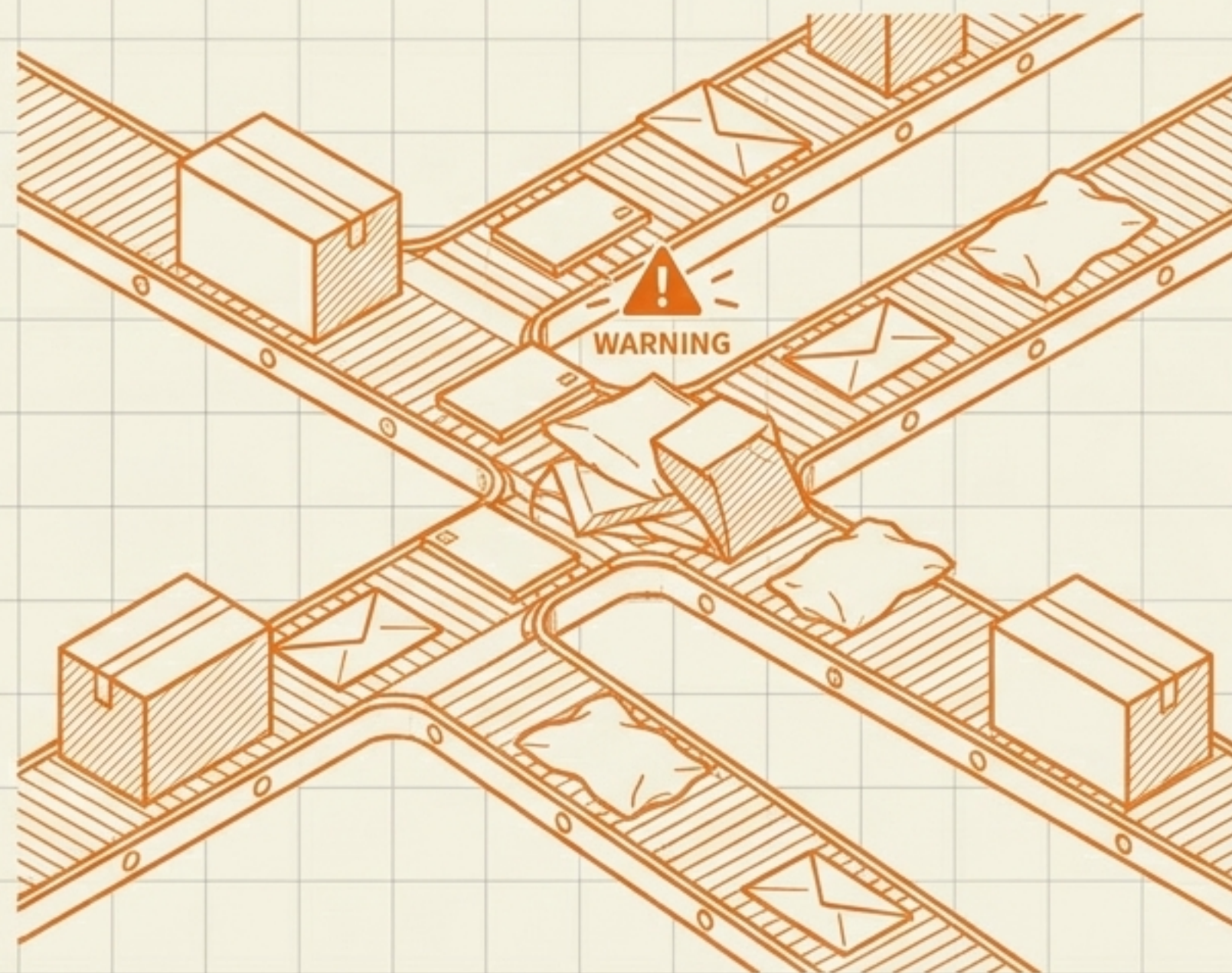
わずか1-2%の「誤向き配置」があるだけで、スキャナ待ちや手直しが発生し、実効スループットは劇的に低下する。速度が僅差であっても、エラー回復コストを含めれば実際の生産性は逆転しうる。

## 今回のテスト環境（クローズド・ループ）



- 同一SKU（全く同じ箱）が循環
- 長時間の「繰り返し耐久試験」に特化

## 実際の物流現場（オープン・ストリーム）



- 新規SKU（多様な形状・重さ）の連続流入
- ラベル異常や例外処理が頻発
- 商用化の真の判断軸（汎化能力）

### 高頻度・低遅延の視覚 (Upgraded Vision)

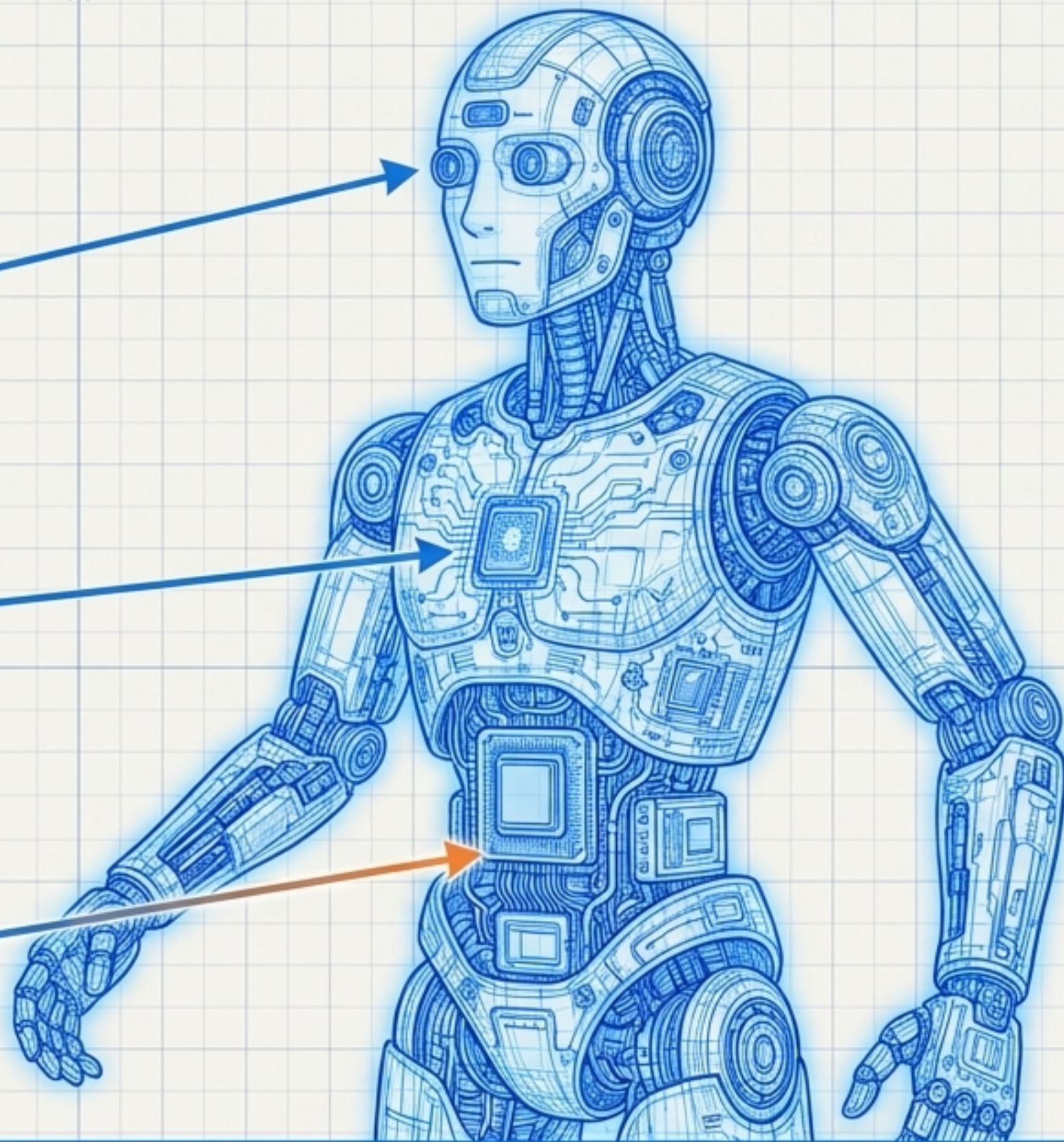
フレームレート2倍、レイテンシ1/4、視野60%拡大。動くベルト上の物体認識を強化。

### 全身制御・触覚統合 (Helix AI)

視覚・触覚・固有感覚をエンドツーエンドで統合し、狭い整流タスクにおける器用さを実現。

### 連続性・回復ループ (Fleet Reliability)

ソフトウェア/ハードウェアのスタック時に自動リセットし、必要に応じて自律的に別機体へ引き継ぐ。



デモの真のブレイクスルーは「単体の速さ」ではなく、「疲労をシステム設計によって『運用可能な連続性』に変換したこと」にある。

## 労働衛生論に基づく正当性

- 倉庫の反復仕分けは筋骨格系負荷が極めて高い（NIOSH基準）。
- 人間を過酷な身体負荷から解放するための自動化には、強力な正当性がある。

## 敵対的なPRメッセージ

- 『Man vs Machine』
- 『This is the last time a human will ever win』
- 支援技術ではなく「人間の置き換え・見世物化」を強調。

摩擦の発生：技術的・倫理的な正当性があるにもかかわらず、コミュニケーションを不必要に対立的にしたことで、社会的受容（Social Acceptance）のハードルを自ら上げてしまった。

# 真のベンチマーク成立に向けた6つの監査要件

- 秒単位のイベントログ**  
把持、誤向き、落下、リセット等の全タイムスタンプ。
- 機体IDと交代ログ**  
対決中の同一個体の継続性、ファームウェア情報。
- 採点ルールブック**  
既正向き荷物の扱いや、取り逃しのペナルティ定義。
- エネルギーログ**  
SOC（充電率）推移、待機機体数、交代に要した時間。
- 荷物マニフェスト**  
SKUの多様性、重量、寸法、材質の分布データ。
- 第三者監査報告**  
遠隔操作の不使用（No teleoperation）の外部署名。

これらが揃って初めて、印象的な「PRデモ」は再現可能な「業界ベンチマーク」へと昇華する。

## 1. デモとしての評価 (The Demo)

【強力な技術的マイルストーン】 特定タスク（狭い整流作業）において、人間の1シフト平均に匹敵するスループットを長時間維持できることを証明した。

## 2. ベンチマークとしての評価 (The Benchmark)

【非対称な条件下での不成立】 エラー率の非開示と、単体（人間）対システム（ロボット群）の曖昧な構図により、実物流の普遍的勝敗を示すデータとしては機能しない。

## 3. 未来へのインサイト (The Future)

【競争軸は『個の速度』から『群の連続稼働』へ】 我々が注目すべきは、「人間より速く動くロボット単体」ではない。自己修復と交代を繰り返し、24時間365日止まらない『フリート（群）としての自律システム』の完成である。これこそが物流自動化の次なる主戦場となる。