

Hub Expansion vs. Operational Reality

What separates an airport that scales
from one that expands?



Table of Contents

The Capacity Paradox 3

Where Investment
Stops Paying Off 4

Aircraft Operations
+ Deep Dive 5

Passenger Operations
+ Deep Dive 7

Baggage Operations
+ Deep Dive 9

Operational Enablers 11



The Capacity Paradox

When Capacity Expansion Meets Operational Limits



Over the past decade, major hub airports have invested heavily in physical expansion. Terminals have been extended, additional gates commissioned, runway systems optimized, and passenger processing areas enlarged to accommodate growing demand. Over the last decade, the global aviation industry invested more than \$600 billion in airport infrastructure.

Yet in many cases, operational pressure has not eased; it has intensified. Despite increased physical capacity, peak-hour congestion persists. On-time performance fluctuates. Passenger processing bottlenecks reappear. Baggage systems approach throughput thresholds. The expectation that expansion alone would stabilize operations often proves overly optimistic.

The reason is structural. Capacity in hub environments is not defined solely by physical assets. It is determined by the synchronization of aircraft movements, passenger processing, and baggage handling. As scale increases, interdependencies multiply. Margins narrow. Variability becomes harder to absorb. The challenge is not insufficient infrastructure. It is operational density.

Case Study

Gate Density and OTP Decline

At a North American hub, terminal expansion and additional gate capacity enabled increased aircraft movements within existing wave structures. Despite higher capacity, on-time performance declined during peak periods. The limiting factor was not stand availability but compressed operational buffers.

-> Expansion increased throughput potential - but reduced the recovery window.

Case Study

When Passenger Throughput Becomes Unstable

At a European hub, expanded passenger processing areas and gate infrastructure were introduced to support growing transfer volumes. Yet congestion persisted during peak banks. Arrival variability triggered queue surges and delayed boarding processes, despite sufficient theoretical throughput. The system reached its coordination threshold before reaching its physical capacity limit.

-> Capacity existed - but synchronization did not.

Where Investment Stops Paying Off

Marginal Capacity Growth and the Investment Efficiency Curve

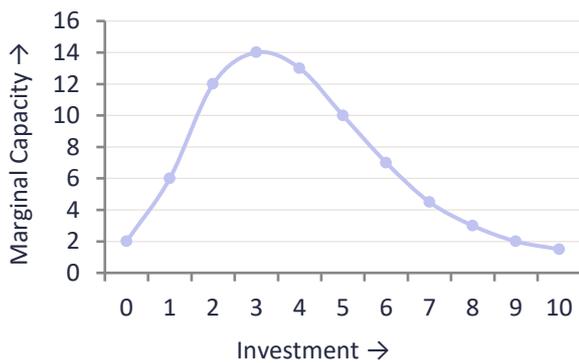
Infrastructure expansion is often assessed in terms of total capacity added. Yet a more revealing question is how efficiently additional investment translates into operational performance.

- Does increased investment automatically generate proportionally higher capacity gains?
- Or does the incremental benefit decline as the system becomes denser and more complex?

This relationship can be described through **marginal capacity growth**, the additional capacity created per additional unit of investment. In early phases of expansion, investments typically unlock substantial gains. Over time, however, as traffic intensity increases and interdependencies multiply, each additional investment may generate progressively smaller improvements in effective capacity.

Understanding this dynamic is critical when evaluating long-term expansion strategies. The following curves illustrate how marginal capacity growth behaves under different structural conditions.

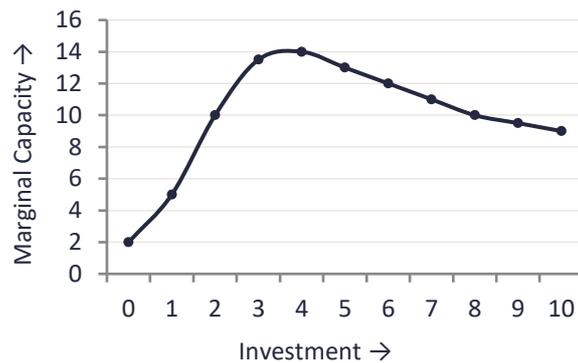
Curve 1: Infrastructure Expansion Without Operational Optimization



This curve illustrates a scenario in which physical capacity increases while operational structures remain largely unchanged. Total capacity grows, but the marginal gain flattens as density increases. Operational constraints, such as compressed buffers, gate sensitivity, and coordination limits, begin to dominate performance outcomes.

Investment continues, but elasticity declines.

Curve 2: Infrastructure Expansion with Operational Optimization



This curve represents expansion supported by integrated operational improvements. By addressing synchronization, flow variability, and cross-functional coordination, the system preserves elasticity under higher density. As a result, marginal capacity growth remains more stable across expansion phases.

Investment continues to translate into meaningful performance gains.



Infrastructure expansion alone does not guarantee proportional performance gains. As operational density increases, marginal capacity growth declines unless systemic constraints are addressed. Sustainable scalability therefore depends not only on investment volume, but on operational optimization. Capacity must be unlocked through integration, coordination, and flow stabilization.

Aircraft Operations

Managing Flow, Timing, and Recovery at Scale

Expansion increases aircraft movements, peak traffic, and rotation sensitivity. Physical infrastructure sets the boundary of what is possible; how precisely rotations, sequencing, and buffers are managed determines what is actually achieved within it. As volumes rise, tolerance shrinks, making coordination the decisive factor.

The three pillars of aircraft operational performance are interconnected in a specific sequence: punctuality creates the conditions for utilization, and utilization determines the real availability of capacity. Understanding that sequence is what separates reactive management from structural optimization.



On Time Performance

The Challenge

As traffic intensity increases, OTP becomes increasingly sensitive to small deviations. Punctuality under density requires more than scheduling, it requires active management. Industry benchmarks place healthy OTP above 85%; below 75%, schedule instability tends to propagate and compound downstream.¹

Operational Optimization

The following optimizations target the points where punctuality is won or lost:

- Integrated airside-terminal coordination to reduce handoff friction
- Predictive delay impact modeling across rotation chains
- Real-time disruption response frameworks with pre-defined escalation paths
- Cross-functional decision ownership across airline, handling, and airport operations

Stability under high traffic load requires orchestration, not isolated improvement.



Gate Utilization

When aircraft arrive on time, gates become available as planned. When they do not, utilization pressure concentrates, rotations compress, buffers erode, and recovery capacity shrinks. Gate utilization is therefore the direct expression of how well punctuality and planning are aligned.



What happens when utilization approaches its ceiling?

-> Deep Dive on the following page.

Gate Capacity

The Challenge

Adding gates increases nominal capacity, but peak gate demand can exceed availability even when overall gate numbers appear sufficient. Capacity is not a static number, it is a dynamic outcome shaped by how well utilization is managed. Planning benchmarks for hubs commonly use 4–5 turns per gate per day, yet demand surges often push utilization toward 80–90% of declared limits, triggering delays.²

Operational Optimization

The following optimizations close that gap between nominal capacity and effective capacity:

- Dynamic gate allocation instead of static pre-assignment
- Predictive buffer management between rotations
- Real-time visibility of inbound variability feeding gate planning
- Flexible stand assignment mechanisms during irregular operations

Effective capacity increases when gate elasticity increases.

Deep Dive

Gate Utilization Under High Load Compression

The Bridge Between Punctuality and Capacity

Gate utilization sits at the center of hub operational performance. It receives the consequences of on-time performance and directly determines the effective capacity available to the operation. When it is well managed, the system flows. When it approaches saturation without buffer, a single deviation is enough to trigger a sequence of disruptions that the operation cannot absorb.

At hub scale, the goal is not maximum utilization but optimal utilization. The highest level the operation can sustain while preserving the recovery flexibility that high traffic intensity demands.

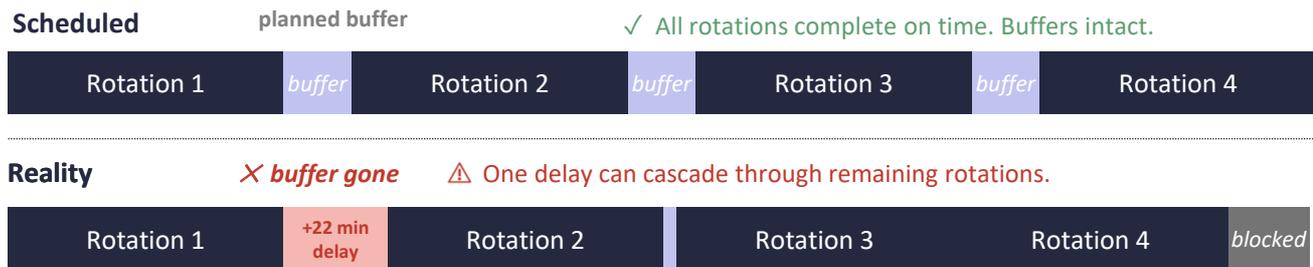
The Cascade Mechanism

The risk of saturated gate utilization is not proportional. One delayed inbound does not simply affect one rotation. It compresses the turnaround window, increases late-push probability, propagates into airside sequencing, and reaches downstream gates receiving connecting flows. In a tightly coupled wave structure, a single event can touch four or five subsequent movements before the system stabilizes, if it does at all.

This is why utilization management cannot be treated as a gate-level problem. It is a network-level problem that surfaces at individual gates. The intervention point may be local. The impact, and therefore the thinking, must be systemic.

Industry benchmarks show typical gate utilization at 65–80%, yet congestion arises when rotations lack buffers, pushing effective throughput above 80–90% of the maximum utilization a gate can handle.¹

Planned Rotation vs. Compressed Reality



A single delayed rotation erases buffer, compresses downstream movements, and cascades across upcoming rotations, without any additional disruption entering the system.

Operational Optimization

Solution 1

Connect the Planning Layer to the Ground

Gate allocation systems work on schedule data of what should be happening, while AI-based live ground detection sees what is actually happening. Integrating both closes the gap between plan and reality in real time, enabling targeted reallocation before pressure builds and transforming utilization management from reactive adjustment to continuous operational intelligence.

Solution 2

Profile Volatility, Place Buffer with Precision

Not all rotations carry equal risk. By profiling turnaround volatility across aircraft type, route origin, and time of day, buffer can be positioned exactly where the system is most exposed rather than distributed uniformly or left to chance. Buffer stops being a scheduling inefficiency and becomes a structural recovery asset.

Solution 3

Pre-Wire the Disruption Response

When utilization is high, the window between a localized disruption and a cascade is narrow. Pre-defined recovery protocols with clear trigger thresholds, pre-assigned decision authority, and cross-terminal flexibility eliminate the lag that allows manageable pressure to become systemic instability.



Passenger Operations

From Curbside Access to the Gate - Where Pressure Accumulates

Terminal space and lane capacity may expand, but performance depends on flow synchronization rather than physical footprint alone. Effective throughput is a coordination challenge, not merely a space constraint. Expansion does not always address the right constraints. Additional gates increase airside capacity while curbside and landside infrastructure remain unchanged, shifting the bottleneck rather than removing it.

Hub airports serve two fundamentally different passenger types simultaneously, origin and destination passengers entering from the landside, and transfer passengers connecting directly from the airside. Both move through different entry points, face different processing layers, and carry different timing constraints.

Check-In Capacity - The Landside Entry Point

The Challenge

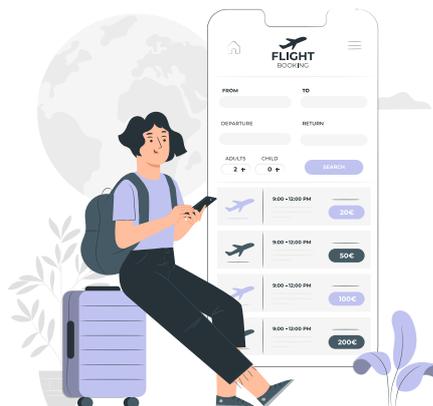
For origin and destination passengers, check-in is the first point of contact with the operation. Allocation is based on static airline demand rather than dynamic passenger arrival patterns. During peak windows, localized congestion forms despite available infrastructure elsewhere. Not because capacity is insufficient, but because it is not positioned where demand concentrates. The imbalance reduces effective capacity. IATA's optimum LoS KPIs for check-in set maximum waiting times at 20 minutes for Check-In queues and 2 minutes at Self-Service Kiosk.¹

Operational Optimization

The following optimizations convert static allocation into adaptive capacity:

- Shared-use counter flexibility across airlines and alliances
- Predictive passenger arrival modeling to anticipate demand shifts
- Demand-based counter reallocation in real time
- Integration of self-service and automated processing to reduce counter dependency

Of the three layers, check-in carries the most inherent flexibility.



Border Control

The Challenge

For transfer passengers, border processing is where arrival variability first becomes a structural problem. Aircraft arrival clusters create sudden demand spikes that exceed smooth processing capability. Unlike check-in, border control involves external agencies with limited operational flexibility, making it the layer most resistant to real-time adaptation.

Operational Optimization

The following optimizations target where arrival variability enters the system:

- Arrival clustering analysis to anticipate demand spikes before they materialize
- Cross-Terminal coordination planning aligned to inbound wave structures
- Predictive queue monitoring with early warning thresholds
- Long-term pax forecast and coordination with agency to support headcount planning

Border stability requires planning between aircraft scheduling and processing authorities. The earlier the coordination, the smaller the cascade.

Security Flow

When both flows pass through, security absorbs them simultaneously. Security is where the cascade becomes visible to passengers, and where the operation either contains the pressure or loses control of it.



Security flow is the critical bridge between arrival & terminal stability.
-> Deep Dive on the following page.

Deep Dive

Security Flow Under Peak Compression

The Cascade Receiver

According to IATA, the optimum security screening wait times should be 5–10 minutes¹, a benchmark that becomes difficult to maintain when passenger flows surge beyond planned capacity.

Security screening sits at the most exposed point in the passenger processing chain. It receives what upstream processes release, whether that is border control, a transfer connection, or a direct landside entry. It operates under its own physical constraints and must deliver passengers to gates within connection windows that do not move.

When the upstream flow is clean and predictable, security performs well. When an overload of arrivals clusters tighter than planned, or border processing runs long security absorbs the full impact. This exposure is structural. Security cannot negotiate its timing with airlines. It cannot defer a queue. The pressure arrives, and the system either handles it, or it doesn't.

Why Queue Instability Is Non-Linear

The critical operational characteristic of security queues under heavy traffic conditions is that they do not behave proportionally. Depending on utilization during peak times, it can multiply the time required to clear. For security operating near capacity, small increases in arrivals create disproportionate effects, destabilizing queues.

This means that managing security flow is not about handling average demand. It is about managing the peaks within the peaks. The moments when transfer surges, late-arriving, and irregular inbounds collide in the same window. Those moments cannot be managed reactively. By the time the queue is visible, the intervention window has already closed.

From Planning to Live Response



Operational Optimization

Solution 1

Connecting Passenger Data to Lane Activation

Integrating real-time inputs from check-in completion, biometric feeds, and transfer manifests into a predictive activation model lets lane configuration respond to what is actually coming. Combined with AI-based occupancy detection across the terminal, the system identifies congestion forming upstream and triggers lane activation before the queue builds. The operation leads demand rather than chasing it. Prediction without response architecture is incomplete.

Solution 2

Pre-Defined Response Thresholds

Prediction without response architecture is incomplete. Solution 2 closes that gap, with the usage of pre-agreed actions, that activates automatically: an extra lane opens, a staff member redeploys, a supervisor is notified. No deliberation, no escalation chain. The decision is made in advance; the threshold just executes it. For this staff communication must exist, to have flexible cross-terminal deployment, governed by standard operating procedures and realistic walking times, is what makes the response viable. SOPs that don't account for a 10-minute walk between terminals is not an SOP that works under pressure.

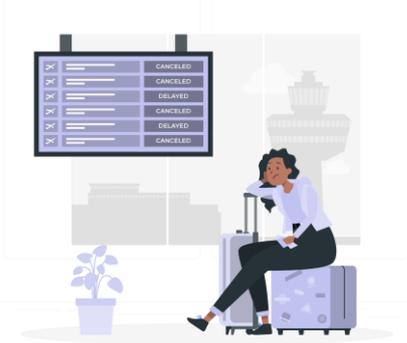
Solution 3

Transfer Flow Separation and Priority Routing

Mixing origin passengers and tight-connection transfers in the same queue creates a predictable conflict. Dedicated fast-track lanes calibrated to connection time thresholds and activated dynamically during peak times can remove transfer pressure from the main queue. It reduces peak load and keeps the primary lane structure stable.

Baggage Operations

The Invisible Backbone of Hub Capacity



As traffic intensity increases, baggage complexity grows disproportionately as well, particularly in transfer-heavy environments. While terminals and gates may expand visibly, baggage systems operate behind the scenes and often approach critical thresholds unnoticed.

Baggage flow stability directly influences aircraft departure readiness, passenger experience, and overall operational resilience. In high-density systems, small disruptions in baggage handling can rapidly turn into delays and missed connections.

Transfer Baggage Stability

The Challenge

According to the latest industry estimates, about 41% of baggage mishandling occurs during transfers, illustrating how sensitive the connection window is to disruption.¹ Short connection times increase transfer bag risk at the source and baggage misalignment becomes more likely and affects downstream flights and passenger satisfaction before the sortation system has even had a chance to respond. The problem is that the window between inbound and outbound narrows faster than the physical process can adapt.

Operational Optimization

The following optimizations target the connection window before it closes:

- Risk-based prioritization of short-connection bags from inbound gate assignment
- Integrated aircraft-baggage sequencing visibility across ramp and control
- Cross-functional coordination between inbound handling and baggage control
- Real-time exception management for bags approaching connection thresholds

Stability requires synchronization across airside and baggage domains.

Baggage Sortation Throughput

Transfer instability feeds directly into sortation. Sortation is where transfer risk is either resolved or becomes a departure delay. IATA does not specify fixed service level benchmarks for baggage delivery². In practice, industry targets typically aim for first bag delivery within 15 minutes and last bag delivery between 25–60 minutes, depending on aircraft type, parking position, and airline service level agreements.



Sortation throughput bridges airside and baggage delivery

-> Deep Dive on the following page.

Carousel & Arrival Flow Management

The Challenge

Arrival peaks can overwhelm the reclaim areas, creating congestion that affects terminal flow and passenger perception. Of the three baggage layers, this is the most visible to passengers and the most recoverable, because the pressure arrives here last and with the most lead time for intervention.

Operational Optimization

The following optimizations distribute reclaim demand before physical congestion forms:

- Predictive carousel assignment based on inbound flight clustering
- Dynamic reallocation responding to real-time arrival variability
- Integration with passenger arrival forecasting to anticipate congestion before it forms
- Carousel assignment based on real-time flight-based log load information

Reclaiming stability protects the overall terminal throughput and passenger journey's perception.



Deep Dive

Baggage Sortation, the Hidden Bottleneck

The Hidden Bottleneck

Sortation is the least visible part of the passenger journey and among the most consequential for hub performance. Passengers never see it. Airlines monitor it indirectly through departure readiness. Airport operations teams watch it through exception reports that often arrive too late to prevent the cascade they describe. According to the latest industry estimates from SITA, global baggage mishandling averages about 6.3 bags per 1,000 passengers, though transfer-heavy hub airports typically experience significantly higher exposure.¹

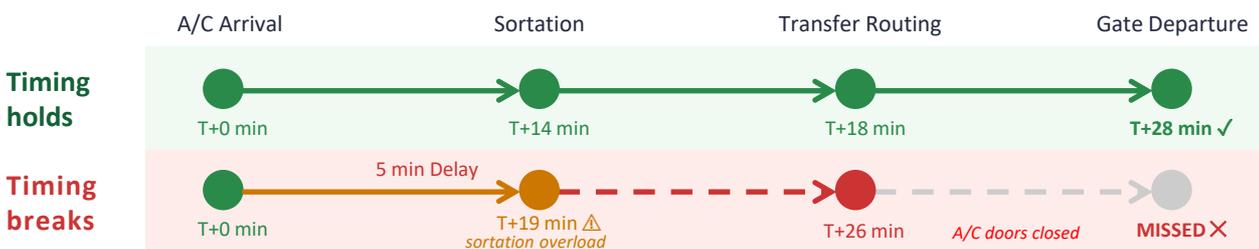
At average load, sortation systems perform well. The problem is not average load. It is the clustered transfer wave, the moment when three inbound arrivals close within a narrow window and simultaneously release hundreds of transfer bags into a system sized for sequential, not concurrent, throughput. That is when the margin disappears.

Why Throughput Sensitivity Is Non-Linear

Sortation systems, like security queues, do not degrade proportionally. A 20% increase in concurrent transfer volume does not produce a 20% increase in processing time. Depending on how close the system is already operating to capacity, it can produce delays that compound across multiple downstream flights. Bags miss their connections not because the belt is broken, but because the sequencing window closed before the bag reached the correct sorting outlet.

This is a timing problem, not an infrastructure problem. And it is precisely the kind of problem that cannot be solved by adding belt capacity alone, because the constraint is not speed, it is synchronization.

When 5 Minutes Is the Difference



A 5 min sortation delay is sufficient to miss the Outbound flight

Operational Optimization

Solution 1

Predictive Transfer Load Modeling Connected to Live Flight Data

Integrating live flight data, inbound delays, gate reassignments, revised arrival sequences into a predictive load model allows the system to anticipate volume spikes before they hit the belt. When overlapping transfer waves are identified early, storage activation and dynamic routing can redistribute load before congestion forms. The system prepares for what is coming rather than reacting to what has arrived.

Solution 2

Risk-Based Bag Sequencing with Connection Threshold Triggers

Not all bags carry equal urgency. A thirty-minute connection and a four-hour connection occupy the same belt, but they do not carry the same risk. Pre-defined connection time thresholds automatically elevate short-connection bags in the sortation sequence, ensuring priority routing before the window closes. When triggers are system-executed rather than manually flagged, the response is consistent and fast enough to matter.

Solution 3

Cross-Domain Synchronization: Ramp, Sortation, and Gate in One View

The core vulnerability is the gap between domains. Ramp knows when the aircraft landed. Sortation knows where the bag is. Gate knows when the door closes. But without a shared real-time view, the window between them is managed by exception and communication rather than visibility. Connecting all three into a unified operational layer allows the operation to see which bags are at risk and how much time remains to intervene.

Operational Enablers

The Structural Layer Beneath Operational Performance

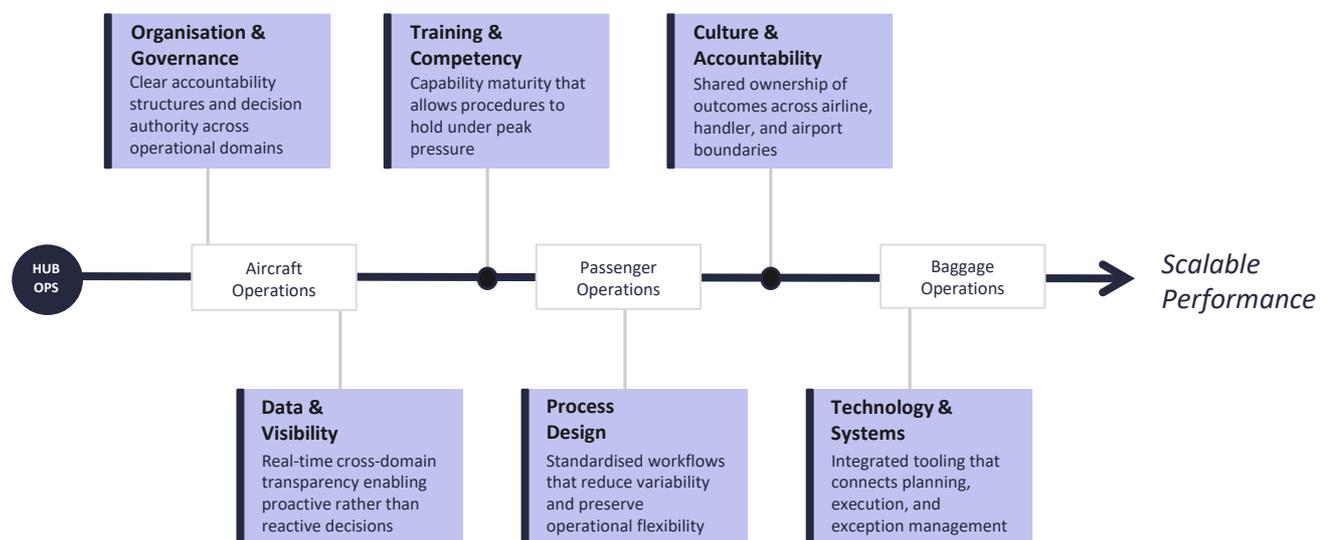
The three operational domains covered in this report: aircraft flow, passenger processing, and baggage handling share a common dependency. Their performance under high operational load is not determined by infrastructure alone. It is determined by the structural layer beneath them: how decisions are made, how information moves, and how people and systems behave when the margin for error narrows.

This layer is invisible in normal conditions. It becomes decisive under pressure.

Airports that have invested heavily in physical expansion but underinvested in governance clarity, cross-functional visibility, and operational culture find that their new capacity performs below expectation. Not because the infrastructure is wrong, but because the enabling conditions for coordinated performance were never built alongside it.

The following six enablers on this page are not enhancements. They are prerequisites.

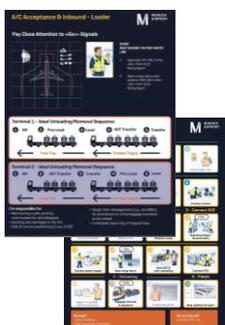
The structural layer that determines whether operational improvements hold at scale



Solution Showcase

The following solutions illustrate how targeted interventions at the enabler level can unlock operational stability without requiring physical expansion.

Standardised Arrival Execution Protocols



Operational Context
High-density arrivals created variability in ground handling execution

Intervention
Structured task cards defining standardized aircraft arrival procedures and sequencing responsibilities.

Real-Time Aircraft & Baggage Performance Visibility



Operational Context
Fragmented visibility across aircraft and baggage domains limited proactive response.

Intervention
Integrated KPI dashboard providing real-time cross-functional performance transparency.



Infrastructure creates the possibility. Coordination determines the outcome.

Closing

The Opportunity in Operational Density

The pressure hub airports face today is real. Congestion persists despite expansion. On-time performance fluctuates. The instinct is to build more gates, more lanes, more belts. And yet the evidence points consistently in a different direction. The constraint is not capacity. It is coordination, which can be improved without waiting for a new terminal to open. Implementing buffer architecture, threshold protocols, sortation sequencing. The tools exist. The data exists. What is needed is the decision to treat operational synchronization as seriously as physical infrastructure.

Behind every delayed rotation, every security queue that exceeds its threshold, every missed connection, there is a passenger. The aviation industry is built on movement. But its purpose is people. The airports that remember this will find that passenger experience and operational resilience are not competing priorities. They are the same priority, approached from different directions.

Operational density is not a crisis. It is an invitation to operate differently.

From Expansion to Performance

- 01 Physical Expansion**
Gates, terminals, and runways are the infrastructure foundation
- 02 Operational Planning**
Rotation sequencing, buffer architecture, wave structure design
- 03 Cross- Domain Visibility**
Real-time data connecting aircraft, passenger, and baggage domains
- 04 Coordinating Architecture**
Decision authority, threshold protocols, cross-functional response
- 05 Synchronized Operations**
Aircraft, passengers, and baggage aligned in real time



- 06 Operational Performance**
Resilience, throughput stability, and passenger experience at scale

 *Every operational decision made well is a satisfied passenger who got home on time.*

Authors



Krisztian von Treuenfels-Gergely

Munich Airport International

E-mail: Krisztian.treuenfels@munich-airport.de



Rania Saoud

Munich Airport International

E-mail: Rania.saoud@munich-airport.de