Global Harmonization Through Collaboration

FIXM Engineering Considerations

Presented By: *Ken Howard*

Date: *August 29, 2012*

EUROCONTROL

Federal Aviation Administration

AIR TRANSPORTATION INFORMATION EXCHANGE CONFERENCE - (FEATURING AIXM, WXXM AND FIXM)

August 28, 2012 - August 31, 2012
NOAA Auditorium and Science Center
Silver Spring, Maryland

Topics for This Talk

- What is the engineering analysis effort?
- How are we going about it?
- Some examples of our analysis.

Engineering the Flight Object

- One definition: “The Flight Object is an extensible and dynamic collection of data elements that describes an individual flight, and is the single common regional reference for all system information about that flight.”
- How does one build such a thing?
- The FIXM data model is a critical component of a real, living flight object, but is not the entire solution.
- The engineering analysis is looking at other parts of the solution.

What is the Problem?

- Different systems can have different ideas of what flights are planning to operate.
- Data received from different sources can be conflicting and ambiguous.
- It can be hard to correlate data from different sources.
- Systems exchange flight data using different formats, protocols, and communications paths.

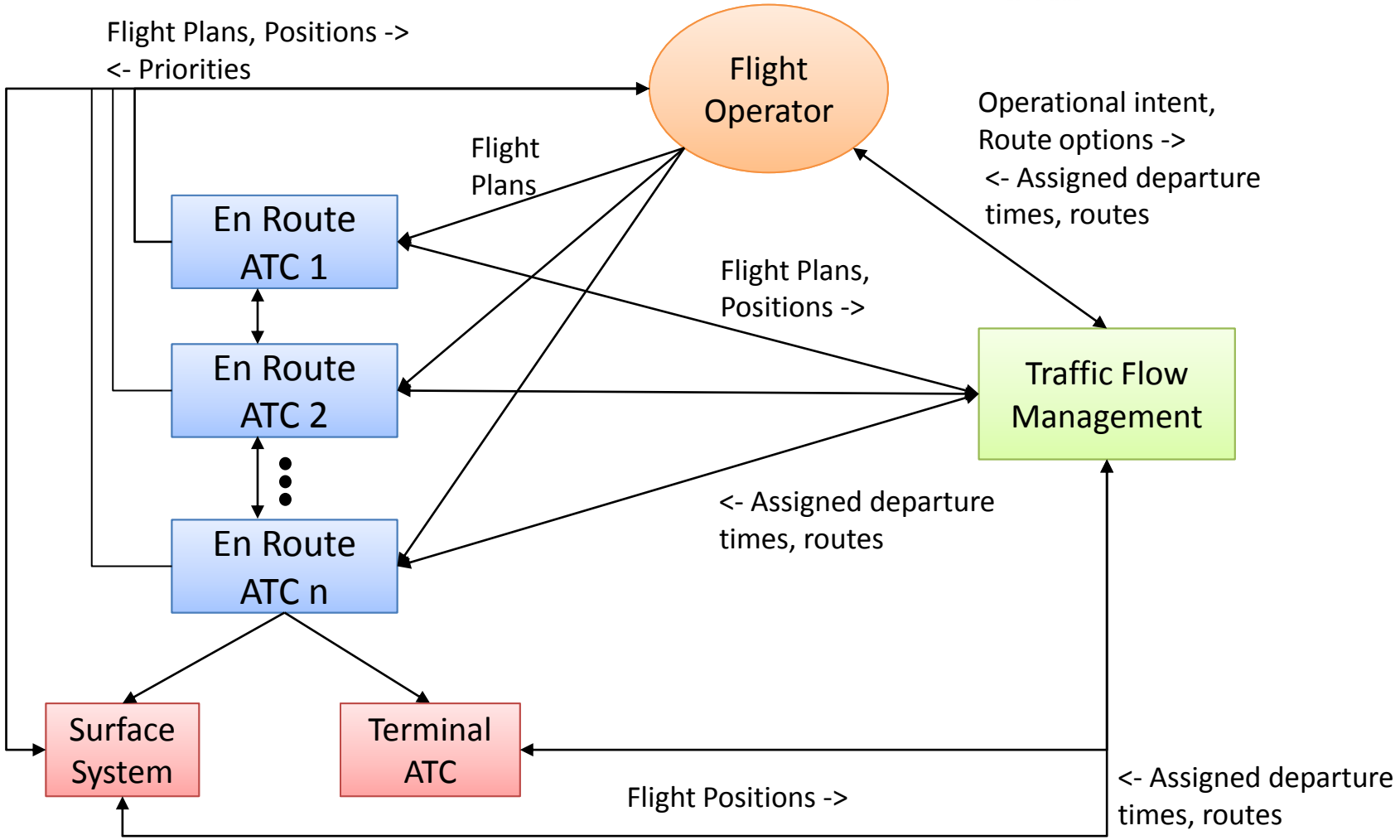
What is the Problem?

- Systems cannot get access to all the data they need.
- Rules for data access are inconsistent.
- Flight data can be lost when flights transition from one system to another.
- There is a lot of redundant data and data processing.

Goals and Objectives

- Data should be easy to access:
  - Available through a common infrastructure.
  - Loosely-coupled interfaces.
  - Delivery must meet consumer performance requirements.
  - Data providers must be able to control access.
- Data should be easy to process:
  - Consistent format and consistent meaning.
  - Easily correlated.
  - Authoritative data easily determined.
- Approach must be evolutionary!

What is the Scope?



International Relevance

- We are primarily analyzing engineering in the FAA environment.
  - This is where our experience lies.
- However, the problems faced when exchanging data between FAA systems and its data consumers are surely the same as problems faced when exchanging data between any two countries, regions or ANSPs.
- As a result, the engineering analysis should have an international relevance as well.

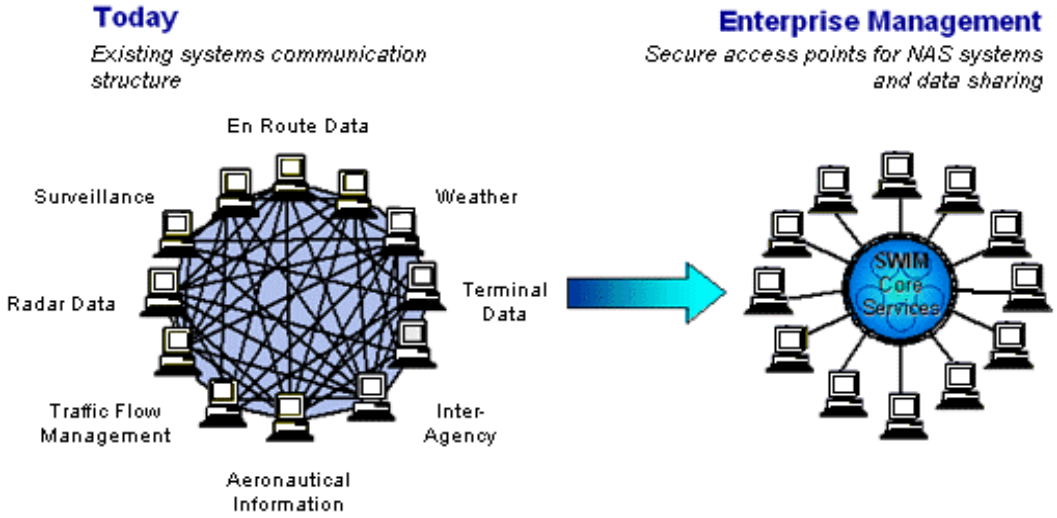
Engineering Topic Areas

- Engineering analysis is organized into a number of topics.
  - High-level Architecture*
  - Data Correlation (GUFI)*
  - Data Reconciliation*
  - Data Exchange Modes
  - Data Customization
  - Access Control
  - Compression*
  - Data Validation*
  - Data Discovery
  - Failure Modes
  - Infrastructure*
  - Archival
  - Transition

*These topics will be discussed briefly today.

High Level Architecture

- Problem: What is the overall architecture for storing and/or distributing data?
  - Put another way, what exactly does this picture mean?
  - One wire? One protocol? One format? One data source?



High Level Architecture

- Some of these questions are easier to answer than others:
  - FAA telecommunications infrastructure makes it possible to use one connection.
  - SWIM/SOA defines a loosely coupled approach based on standard protocols.
  - FIXM is defining a consistent data model and format.
- Other questions require more analysis; example: Data Reconciliation.

Data Reconciliation

- Problem: How do you combine conflicting data from different sources to provide consistent, authoritative data?
- Examples of conflicting or ambiguous data:
  - Position updates from adjacent or overlapping facilities.
  - Route of flight as filed by flight operator, adapted by departure facility, adapted by arrival facility.
  - Predicted departure time as:
    - Planned by flight operator
    - Filed on a flight plan
    - Assigned by a national TFM ground delay program.

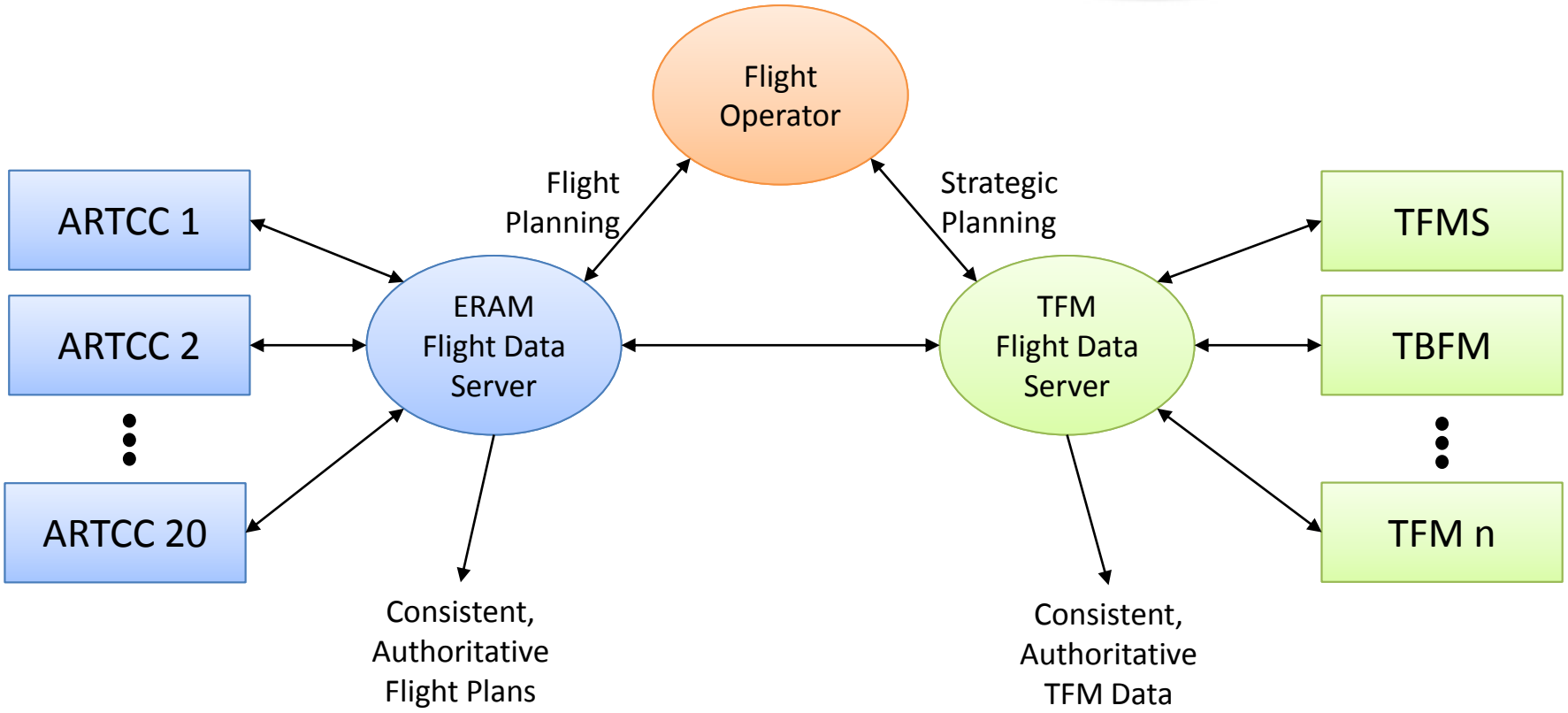
Data Reconciliation

- Alternative approaches using distributed architecture.
  - Allow only one authoritative source to publish data at a time; negotiate who is the publisher.
    - Works well for coordinating discrete data, such as position updates, between peer-level systems.
    - Doesn't handle multiple versions of routes, for example.
  - Make one system the "manager" of the data but allow other systems to contribute data.
    - Works well for merging conflicting data (e.g., routes) from peer-level systems.
    - Do you want each system to maintain and publish data for every other systems? (e.g., ATC and TFM)

Data Reconciliation

- Alternative approaches using centralized architecture.
  - Implement a big centralized database.
    - Good for achieving consistency and makes it easy for a consumer to get all the data it wants from one place.
    - This would be a complex process and could have performance problems.
- Hybrid of the above:
    - Each “data domain” (for example, En Route ATC or TFMS) publishes its own data.
    - Within each domain, use either a centralized or “distributed manager” approach to produce consistent data for that domain.
    - Negotiate between domains for data that overlaps.

High Level Architecture



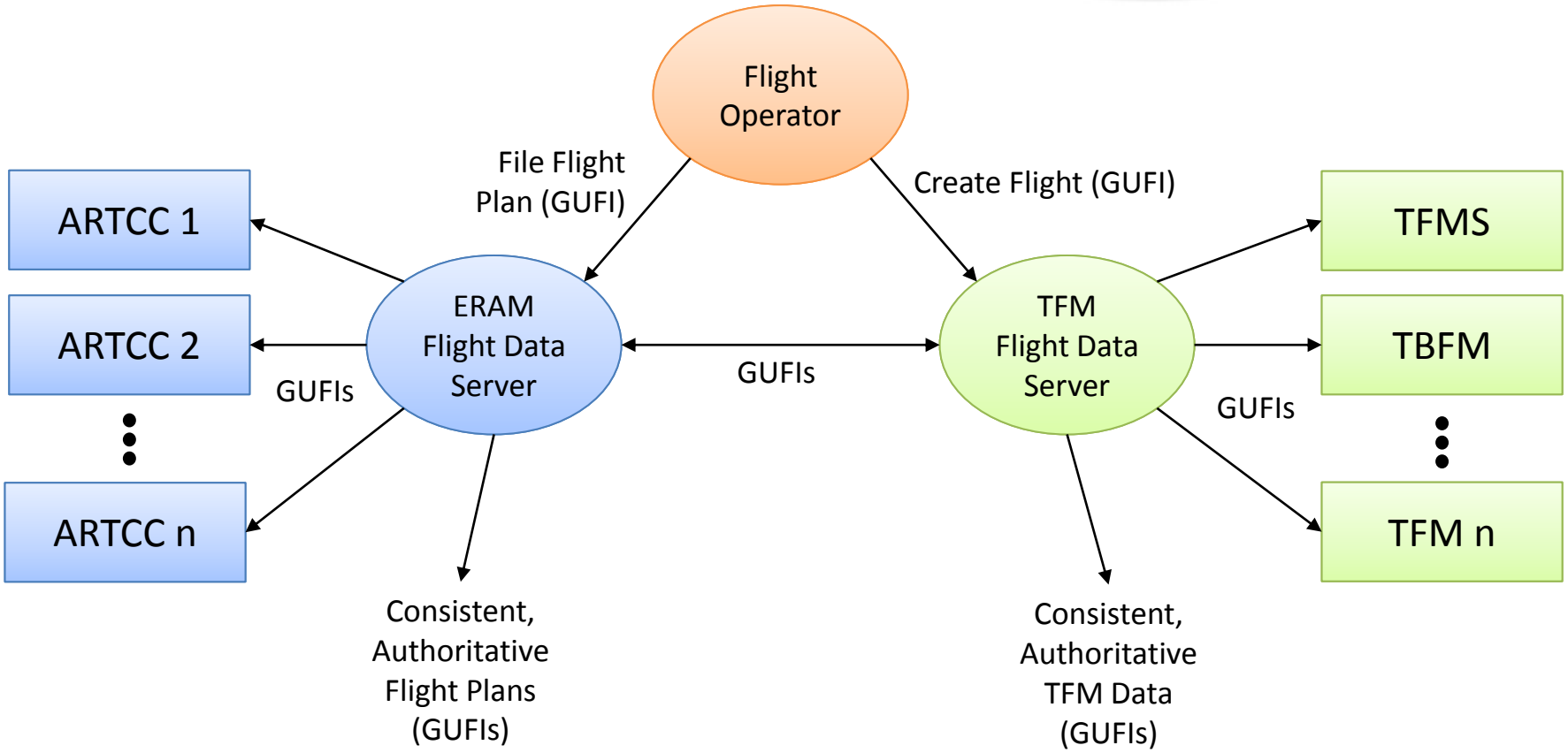
Data Correlation (GUFI)

- Problem: How does a consumer correlate data received from different sources?
- Why is this so hard?
  - Within FAA ATC/TFM domain, flights are created from several different sources:
    - Schedule data, “CDM” data, flight-operator flight plans, flight plans from adjacent ANSPs.
  - Data from each source might differ in:
    - Call sign, origin, destination, departure time etc.
- Solution: A Globally Unique Flight Identifier (GUFI).

Data Correlation (GUFI)

- Every system keeps track of “unique flights”.
  - Note: This differs from the current ERAM-GUFI approach in which a GUFI refers to a flight plan, not a flight.
- Every flight gets a GUFI at the time it is first created in an ANSP system.
- Every subsequent transaction for that flight includes the GUFI.
- Every system uses the GUFI as its primary mechanism for matching data to unique flights.

Data Correlation (GUFI)



GUFI Issues

- Who generates the GUFI: flight operator or ANSP?
  - Either approach works. The important point is *when* the GUFI is created, not by whom.
  - Solution might want to allow for both.
- Format: What does the GUFI look like?
  - Possible format:
    - <country code>.<org code>.<date>.<time>.<sequence number>
    - Examples: us.faa.20120210.0631.17; fr.f9893rl.20110930.1745.1
  - As long as a GUFI generator can be guaranteed a way to make its GUFIs unique, the format may not matter.

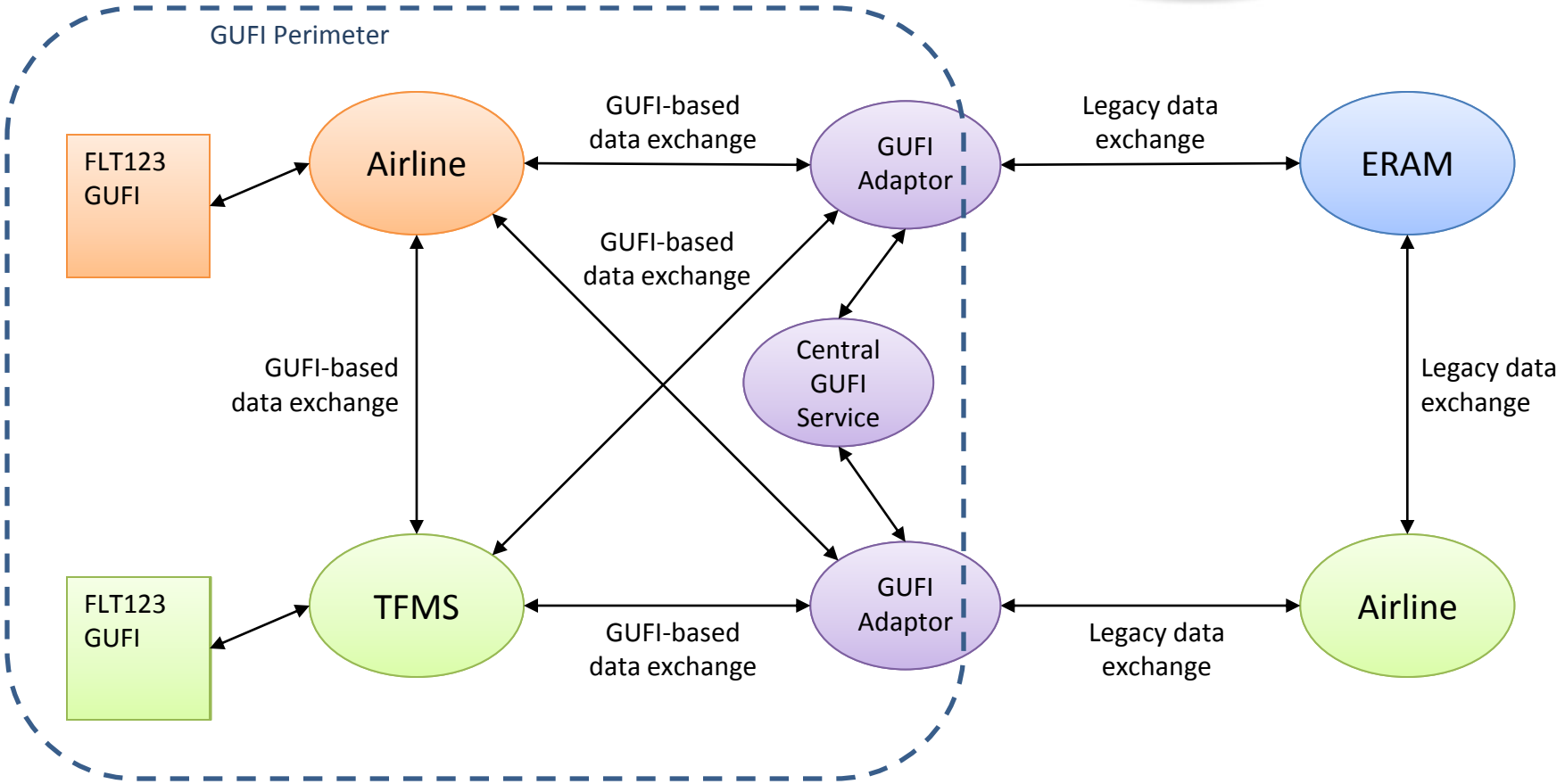
GUFI Issues

- Transition: How can the GUFI be introduced in an evolutionary manner?
  - ANSP systems are likely not all going to convert to the GUFI at the same time.
  - Flight operators will not all convert to the GUFI at the same time.
  - If there is no benefit from the GUFI during the transition, no one will convert.
  - The transition period could well last forever.
    - That is, some systems may never convert.

GUFI Issues

- Transition Approach: Create a “GUFI perimeter”.
  - Systems inside the perimeter use the GUFI in their data exchanges
  - Systems outside the perimeter do not.
  - Create “adaptors” at each perimeter crossing that convert between GUFI and non-GUFI messages.
  - Create a centralized “GUFI service” that allows the converted systems and adaptors to synchronize GUFIs.

GUFI Transition Approach



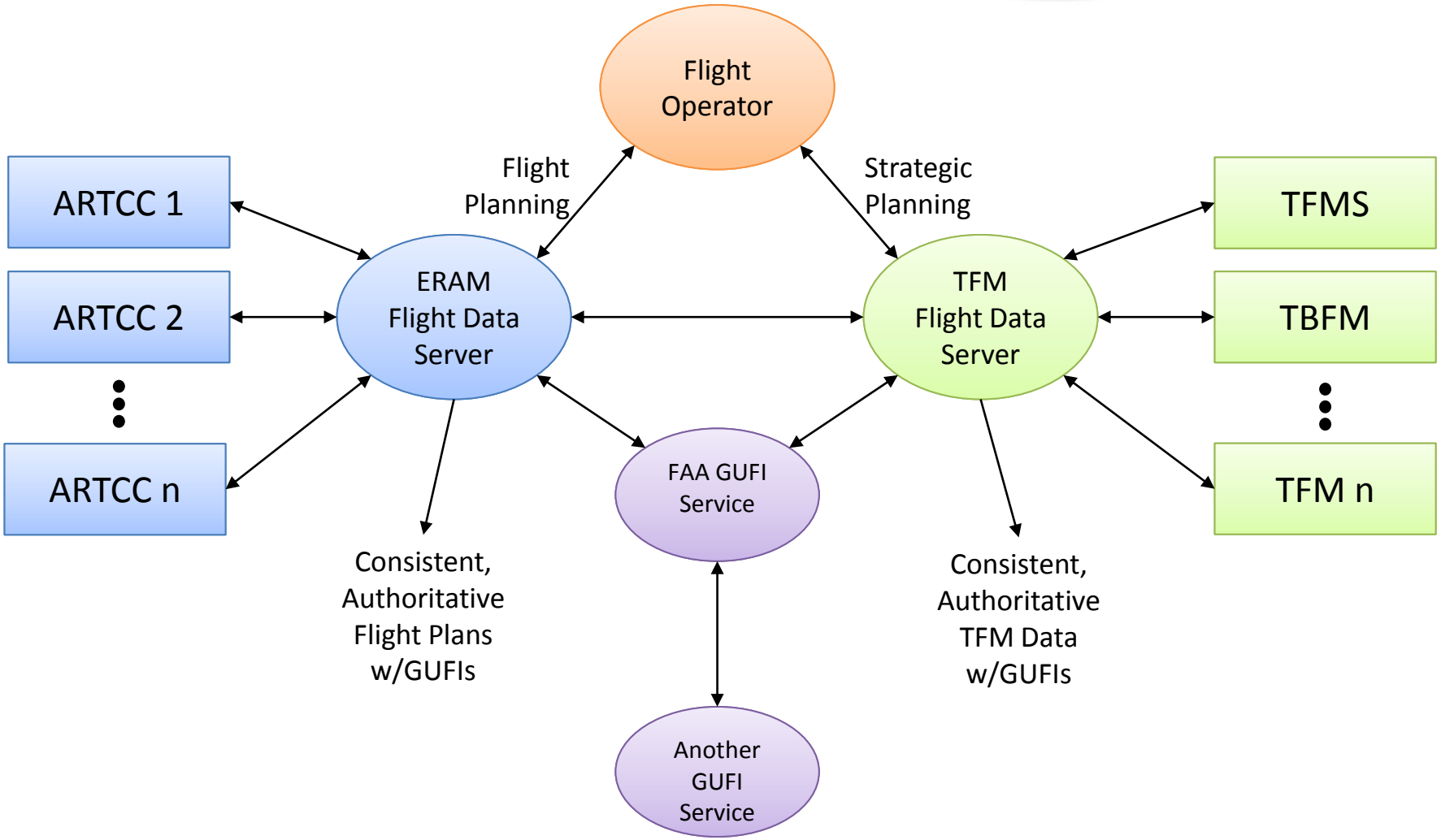
NOTES:

ANSP systems can be converted independently of each other, while the airline gets full benefit.
Adaptors can also be applied to data exchange between ANSP systems.

GUFI Issues

- International: How do we coordinate GUFIs across international domains?
  - Easiest if general philosophy and GUFI format is universal.
  - If so, the international issues are the same as the domestic issues, just more players involved.
    - For example, an international data feed is another external source that can create flights with or without a GUFI.
  - Rather than one GUFI service, maybe we have several that hand-shake with each other to ensure common GUFIs are used.

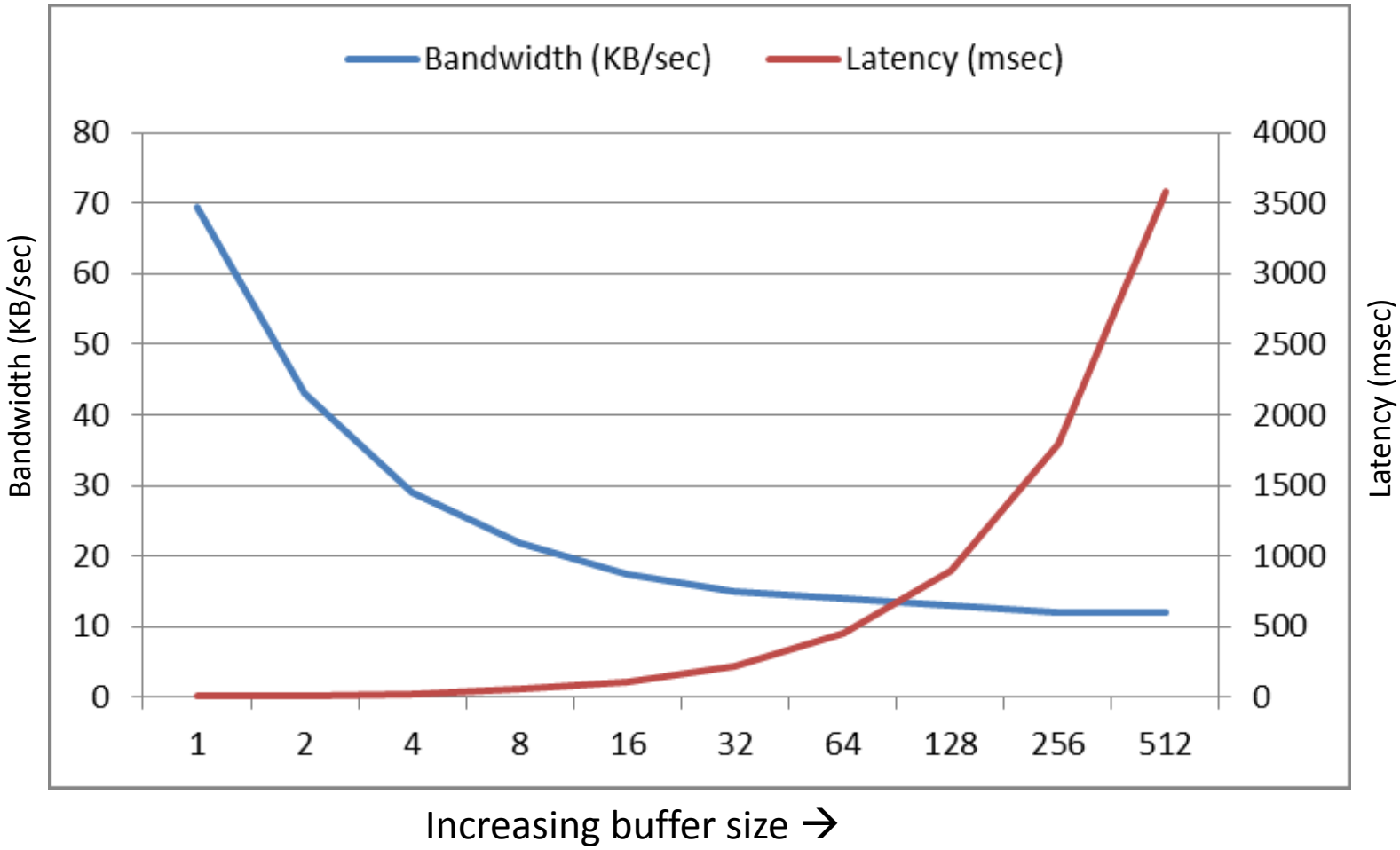
Architecture with GUFI



Compression

- Problem: How can we minimize bandwidth requirements for transmitting verbose XML data?
- Conventional analysis typically looks at only compression efficiency using large batches of data.
- Flight data consists of streams of small messages to be delivered quickly.
  - Single messages do not compress well.
  - “Batching” gives you a way to improve compression efficiency, but …
  - Need to look at impact on latency.

Compression Example



Compression

- Initial recommendations:
  - Batching is a very effective way to increase compression efficiency.
  - Each system must determine its own trade-offs between efficiency and latency.
  - For batches >> 1, GZIP works better than EXI.
  - NOTE: Alternate views on EXI compression are being presented in the WXXM session.

Data Validation

- Problem: How should data validation best be performed?
- COTS schema validation tools offer and easy way for a system to validate incoming data.
  - Can validate items such as:
    - Required fields exist; unknown fields do not.
    - Format.
    - Range.
    - Allowed values.
  - Quality of the validation depends on how restrictive the schema definition is.

Data Validation

- Sample tools have been tested and work well while introducing minimal additional latency.
- It is up to each system to determine whether they need to perform validation, but it can be done very easily with minimal impact.

Infrastructure: SWIM

- Problem: What common infrastructure should we use to exchange flight data?
- The FAA's System Wide Information Management program defines a standards-based Service Oriented Architecture (SOA) for exchanging data.
  - The FuseSource tool set meets these standards and makes it relatively easy to become compliant.

Infrastructure: SWIM

- SWIM defines approaches for a number of engineering topics:
  - Data Exchange Modes:
    - Pub-sub using JMS.
    - Request-response using web services.
  - Data Access Control: security-related services.
  - Data Customization: property-based filtering.
  - Data Discovery: NAS Service Registry.

Infrastructure: SWIM

- One potential concern is the performance of the COTS-based solution.
  - We have prototyped a FuseSource-based flight object server in a closed lab environment that:
    - Processes on the order of 2.5 million messages per hour.
    - Performs schema validation, filtering, data enhancement, and compression.
    - Has average latencies of under 1 second on a single server.

Early Implementation: FDPS

- SWIM re-plan includes development of the Flight Data Publication Service, FDPS.
- FDPS will make en route ATC (ERAM) data available via SWIM.
- FDPS will use the FIXM data model/schema, will adopt other Flight Object qualities, and will provide valuable feedback to Flight Object project.
- FDPS will be a working prototype by the end of 2013; operationally deployed in 2015.

Summary

- Final outputs of this effort will include:
  - Recommendations, best practices, and lessons learned for implementing systems.
  - Trade-off analyses that will help system designers make smart choices.
  - Recommendations for Flight Object Exchange Services (FOXS) that should be developed to support implementing systems.

Questions

Contact Information

- Engineering analysis documents are available at:
  - www.FIXM.aero
    - Documents
      - Engineering
- I can be reached at:
  - Ken.Howard.CTR@dot.gov
  - Or come by the FIXM booth.