Globally Unique Flight Identifier (GUFI) Format and Content

A collaborative effort to establish the Flight Information Exchange Model (FIXM) has been underway for several years. A key component of the flight model is a Globally Unique Flight Identifier (GUFI) that is included on every Air Traffic Management (ATM) flight data transaction to unambiguously identify the flight to which the data applies. The purpose of the GUFI is to eliminate problems that have occurred in the past when systems try to accurately correlate data that is received from many other systems. This paper discusses the format and content of the GUFI.

June 23, 2014

Version: 2.1

Table of Contents

Document History

Version	Version Type	Description of Changes
1		Document originally released as ATMRPP-WG23-WP/548 for the International Civil Aviation Organization (ICAO) Air Traffic Management Requirements and Performance Panel (ATMRPP) 23rd Working Group Meeting in Brussels on March 4 to 8, 2013.
2A	DRAFT	Minor changes to be consistent with Version 2 of the GUFI Requirements document.
2B	DRAFT	Added a definition of different categories of GUFIs, and a new proposal for format.
2C	DRAFT	Included discussion of URN and UUID. Separated common requirements from approach specific design considerations.
2D	DRAFT	A number of changes from the GUFI TIM held on 5/14/14. The main change is to clearly recommend the UUID approach. Also, many clarifications were added to the descriptions of the two approaches.
2E	DRAFT	A few additional editorial corrections.
2.0	Final	2E with a few minor edits. Ready to distribute.
2.1	Final	2.0 with several small editorial corrections.

1 Introduction

Developing the Flight Information Exchange Model (FIXM) is an ongoing collaborative effort across international stakeholders. A key element of the flight model is a Globally Unique Flight Identifier (GUFI). The concept of the GUFI and a proposed set of requirements have been presented in reference 1. The purpose of this paper is to discuss a format and content for the GUFI.

2 Background/Discussion

2.1 Purpose

The general purpose of the GUFI is described in detail in reference 1. In summary, the purpose is to define a globally unique identifier that can be attached to the data about a flight to support Air Traffic Management (ATM) flight data exchange (that is, flight data exchange in support of ATM functions). The GUFI will be attached to every ATM flight data transaction for a flight, allowing systems to easily and accurately correlate the data in the transactions. Using the GUFI will improve the flight data accuracy and consistency of all systems that exchange ATM flight data, and will lead to more efficient planning and execution of the operations of those flights.

There are three main reasons for having a standardized GUFI format. First, it will help assure that the GUFI requirements, as defined in reference 1, can and will be met. Second, there are many issues related to the design and implementation of the GUFI that can more easily be met with a standard format. Third, a standard format allows validity checking and provides a degree of confidence when processing received flight information. The purpose of this paper is to discuss approaches to the GUFI format and content that both meets the GUFI requirements and facilitates the GUFI implementation.

2.2 General Concept

The general concept of the GUFI, as documented in reference 1, is as follows:

- Every unique flight, defined as a single movement of an aircraft from takeoff to touchdown, must have a GUFI.
- The GUFI for a flight must be unique; that is, a flight can only ever have one GUFI, and a GUFI can only ever be allocated to one flight.
- The first time any ATM data for a flight is shared between two stakeholders (for example, the flight operator and an ATM system), a GUFI must be assigned to the flight. An example of such a transaction is a flight operator filing a flight plan with an ATM service provider. (Note: The "sharing" of data might be through a data publication or a web service, such as a service for creating a flight plan.)
- Every subsequent ATM flight data transaction for this flight must include that GUFI.
- Every system that processes data for a flight uses the GUFI to associate the data in a transaction with its internal representation of the flight's data.

This concept ensures that every system applies data to the same flight entity as every other system. All systems will have the same expectation as to what flights are operating, and the same data defining those flights.

2.3 General Approaches

Work carried out previously on unique flight identifiers falls into three categories.

- Natural identifier: in which information about the flight is encoded in the flight identifier, hence it is referred to as 'natural'. This approach is adopted by the Aviation Information Data Exchange (AIDX) Unique Flight Identifier (UFI), and presently used in a business context by members of the International Air Transport Association (IATA) (reference 2).

  Requirement 7 of reference 1 states, "A GUFI shall be used for flight identification only". A natural identifier allows the GUFI to be used for business purposes, though the requirement is strictly speaking imposed on the use of the GUFI rather than the format and content of the GUFI. Also, because it contains natural data, which can change, there is the risk that a system will synthesize a GUFI that will not match. For these reasons, the natural identifier approach is not considered further for ATM use.

- Meta-identifier: in which information about the GUFI itself is encoded in the flight identifier, but no information about the flight is encoded. See section 4.1.
- Uninterpreted identifier: in this approach the identifier has no semantic content; it is simply a unique sequence of characters. See section 4.2.

The remainder of this document discusses standardization of the format and content of the GUFI for the meta-identifier and the uninterpreted identifier categories.

2.4 Uniform Resource Name (URN)

Requirement 6 of reference 1 states, "The GUFI definition shall be based on accepted global standards". In order to satisfy this requirement, it is proposed that, regardless of the approach taken, the GUFI is represented as a Uniform Resource Name (URN). A URN is a Uniform Resource Identifier (URI) scheme that is used to uniquely name global resources. In the context of the GUFI, the resource is a flight. Therefore, a GUFI is a URN (reference 6) with the general format:

urn:<nid>:<nss>

where <nid> is the namespace identifier and <nss> is the namespace specific string. URN is assumed in subsequent discussions.

3 Design Constraints

This section captures the primary design constraints and goals for the GUFI format, regardless of whether the meta-identifier or uninterpreted identifier approach is adopted. These are either derived directly from the GUFI requirements (reference 1) or are intended to assure that the GUFI can be easily implemented.

1) **The GUFI format must allow one system to generate a new GUFI independently of other systems and still guarantee its uniqueness.**

   This is the single, most important goal of the format. When a system generates a GUFI for a new flight, it must be able to do so with full confidence that the same GUFI will not be generated by some other system for a different flight. In addition, to meet requirement 4 from Reference 1 (a GUFI shall be unique for at least ten years), the system must ensure this uniqueness over a long time period. A system's ability to meet these requirements depends heavily on the format of the GUFI.

2) **The GUFI format must make it easy to assure that a given GUFI be assigned to only one flight within a ten-year period.**

   This is a derived from requirement 4 in reference 1. The format cannot place an unusual burden on the systems generating the GUFI to meet this requirement. As an extreme example, it would be unreasonable to expect that a system save the GUFIs it has generated in the last 10 years and check that a new GUFI doesn't match any of them.

3) **The GUFI format must not constrain the number of GUFIs that a system can generate during any given time period.**

   This is a minor but necessary consideration to ensure that the selected format does not hamper us from future growth. This is a lesson learned from some past identifiers that did not include sufficient digits in fixed size formats. For example, an individual aircraft operator might be able to create a unique string by combining a date-time with a one-digit sequence number. However, this might not be sufficient for an airline or an Air Navigation Service Provider (ANSP) during a data recovery where many flights are created in a short time. A more flexible format would be a date plus a variable length sequence number, which can grow to as large as it needs to be.

4) **The GUFI format must conform to international standards for format and content wherever possible.**

   As dictated by requirement 7 in reference 1. These could be ICAO, IATA, ISO, or other standards.

4 Alternative Approaches

This section describes and evaluates the meta-identifier and uninterpreted identifier approaches.

4.1 Meta-Identifier

The meta-identifier approach encodes data about the generation of the GUFI that serves to ensure that the GUFI design constraints from Section 3 can be easily met. There are two main thoughts behind this approach:

- If the GUFI includes a unique identifier for the system generating the GUFI, we can be confident that another system will not generate the same GUFI (meets constraint 1).
- If the GUFI includes date-time information, we can be sure that the system will not generate the same GUFI for a different flight on a different day; specifically, within the required 10-year period (meets constraint 2).

A specific implementation of the meta-identifier approach is presented here. While there are other possible implementations, this example serves to highlight the general pros and cons of the meta-identifier approach. This implementation of the meta-identifier uses the following three fields, separated by periods:

- **Field 1**: Globally unique, predefined country, region, or organisation code. This might be composed of 2 to 10 alphanumeric characters. Some examples are: us, euro, iata.
- **Field 2**: Organization, facility and/or system code; that is, the generating entity. The system code must be unique for a given field 1 value. Whenever possible, this field should include the organization or facility code using standard names; for example, if the generator is an airline system, field 2 should use the ICAO or IATA airline code. If an organization or facility has more than one system that can generate a GUFI, field 2 must include a secondary system identifier. The format is an

organization/facility code followed by the system identifier separated by a hyphen. Total length could be 2 to 10 characters composed only of alphanumerics except for the separating hyphen. Examples: cfmu, zbw, dal, lh-1, ua-kord.

Field 3: Date-time that the identifier was created. Compliant with ISO 8601 (reference 3). Example: 2012-05-12T17:43:22Z. Specific restrictions placed by the GUFI on ISO 8601 are:

- The full date is always present;
- The full time to a granularity of seconds is always present;
- Where fractions of a second are specified, the separator ‘,’ is used;
- The time zone designator must be UTC (code ‘Z’);
- Week dates and ordinal dates are not allowed.

In addition, the label ‘gufi’ would be included as the namespace identifier.

Some example GUFIs that adhere to this format are:

- urn:gufi:euro.cfmu.2013-02-05T12:12:57,2764Z
- urn:gufi:us.ual.2013-02-05T12:12:57,4Z
- urn:gufi:iata.lh-1.2013-02-01T17:30:00Z

The idea of this format is that the combination of fields 1 and 2 form a globally unique identifier for the system generating the GUFI, therefore making it easy for a system to generate GUFIs that will be unique from other system’s GUFIs (constraint 1). The date-time in field 3 is an easy way for the generating system to create a GUFI for each flight that will be unique for at least 10 years (constraint 2). By including fractions of seconds, a system is not constrained to how many GUFIs it can create at any one time (constraint 3). By using fields 1, 2, and 3 together, it is relatively easy for a system to create a GUFI that will be globally unique for at least ten years. The only burden on the system generating the GUFI is to make sure it does not create two identical GUFIs in the same second, using fractions of a second. In addition, the information in the GUFI provides potentially useful information for troubleshooting and post-analysis. Finally, this approach uses standardized terminology and formats (constraint 4).

The pros and cons for the meta-identifier approach are as follows:

<u>Pros</u>

- Meets all of the constraints in Section 3.
- Contains potentially useful information about who generated the GUFI and when it was generated that could be used for troubleshooting and post-analysis.
- Is relatively easy to implement.

<u>Cons</u>

- Could possibly be used for business processing, which would be a violation of requirement 7 in reference 1.
- Requires significant management and governance. Some organization must be responsible for allocating field 1 values. Each field 1 country, region, or organization would have to govern the use of field 2. This could raise some issues. For example, IATA reuses airline codes over time; does this present a problem?

- While easy to implement, the meta-identifier requires each generating system to write custom code for generating and validating the GUFI.
- The new GUFI URN namespace identifier would need to be registered.

4.2 Uninterpreted Identifier

The job of creating a globally (or universally) unique identifier for a data object is not a new one. Machines on a network and software licenses are two cases where unique identifiers are created and assigned to many objects, sometimes by distributed entities. One method that has been adopted with success is the Universally Unique Identifier (UUID), standardised by the Open Software Foundation (OSF). UUID is a well-documented (see reference 5) and widely implemented concept. (A good description of UUID theory and practice can be found at reference 6). Of particular relevance to the GUFI and FIXM is the decision to use UUID as a feature identifier in the Aeronautical Information Exchange Model (AIXM); the rationale for this decision is documented in reference 7.

The UUID is a way of creating identifiers in a manner that for all intents and purposes guarantees uniqueness. As stated in reference 6, “The intent of UUIDs is to enable distributed systems to uniquely identify information without significant central coordination. … Anyone can create a UUID and use it to identify something with reasonable confidence that the identifier will never be unintentionally used by anyone for anything else.” This is close to an exact summary of the requirements for the GUFI.

A UUID is a 128-bit (16-byte) sequence that can be represented in a text string as 32 hexadecimal digits in groups of 8, 4, 4, 4, and 12, separated by hyphens. As stated earlier, it is desirable to express this using a URN. Per Request for Comments (RFC) 4122, reference 5, when represented as a URN, the namespace identifier for a UUID is ‘uuid’. The namespace specific string is the string representation of the bit sequence. An example of the resulting expression of a UUID is:

urn:uuid:f81d4fae-7dec-11d0-a765-00a0c91e6bf6

There are five accepted versions of the UUIDs (reference 6). The Eurocontrol-FAA AIXM Design Team analysed these methods and determined that Version 4, based solely on random number generation, is best for use in AIXM (reference 7). The considerations for GUFI are similar to AIXM, and so it is reasonable to assume that UUID Version 4 is best of for the GUFI as well. One of the benefits of Version 4 is that it is supported by virtually every commonly used development language and platform; in fact, a benefit of UUID in general is that no custom code needs to be written to generate or validate UUIDs.

The UUID approach does not explicitly and absolutely guarantee uniqueness, but relies on probabilities. The Version 4 UUID uses 122 random bits (6 bits are used to identify the UUID version). Generating a random number of this size virtually guarantees that one system will never generate the same GUFI as another. According to reference 6, after generating approximately 69 billion (69,000,000,000) GUFIs, the probability of the next GUFI being a duplicate would be one in 2.5×10^{15} (2,500,000,000,000,000). For perspective, if systems worldwide generate 1,000,000 GUFIs a day, it would take 189 years to generate 69 billion GUFIs.

In summary, the use of a UUID meets all the design constraints from Section 3. A system can easily generate GUFIs that are unique from other systems (constraint 1), GUFIs would be unique for at least 10 years (constraint 2), there are no constraints to the number of GUFIs generated (constraint 3), and the UUID is a standardized approach (constraint 4).

The pros and cons for the uninterpreted, UUID approach are as follows:

<u>Pros</u>

- Meets all of the constraints in Section 3.
- Is extremely easy to implement.
- Supported by development platforms, requiring no customized code to generate or validate the GUFI.
- Cannot possibly be used for business processing, thus supporting requirement 7 in reference 1.
- Requires no management or governance of unique codes.
- Already adopted by AIXM.
- Already has a registered URN.

<u>Cons</u>

- Contains no information about who generated the GUFI or when.

4.3 Recommended Approach

The uninterpreted, UUID approach has four significant advantages over the meta-identifier approach:

1. It requires no management or governance of unique codes.
2. It is supported by library functions on all major development platforms.
3. It has already been adopted by AIXM.
4. It is an accepted, worldwide standard.

While the meta-identifier approach is certainly feasible, these advantages make it clear that the UUID is the preferred approach to implementing the GUFI.

Appendix A Acronym List

Acronym	Definition
AIDX	Aviation Information Data Exchange
ANSP	Air Navigation Service Providers
ATM	Air Traffic Management
ATMRPP	Air Traffic Management Requirements and Performance Panel
FIXM	Flight Information Exchange Model
GUFI	Globally Unique Flight Identifier
IATA	International Air Transport Association
ICAO	International Civil Aviation Organization
ISO	International Organization for Standardization
OSF	Open Software Foundation
RFC	Request For Comments
UFI	Unique Flight Identifier
URI	Uniform Resource Identifier
URN	Uniform Resource Name
UUID	Universally Unique Identifier

Appendix B References

1. Flight Information Exchange Model (FIXM), "Globally Unique Flight Identifier (GUFI) Requirements", Version 2 Draft E, May 19, 2014.
2. IATA, "Recommended Practice 1797a: Aviation Information Data Exchange, Passenger Services Conference Resolutions Manual", 33rd Edition, 2013, (Found at http://www.iata.org/publications/Pages/pscrm.aspx.)
3. Date and time format – ISO 8601, International Organization for Standardization. (http://www.iso.org/iso/home/standards/iso8601.htm.)
4. URN Syntax, RFC 2141, (http://tools.ietf.org/html/rfc2141.)
5. A Universally Unique Identifier (UUID) URN Namespace, RFC 4122, (http://www.ietf.org/rfc/rfc4122.txt)
6. Universally unique identifier, Wikipedia, (http://en.wikipedia.org/wiki/Universally_unique_identifier)
7. AIXM 5, "Feature Identification and Reference – use of xlink:href and UUID", Version 1.0, 29 April 2011.