

## Bulk Data Implementation over public cloud with secure sharing of data file

Chetan Kumar Jain, Optum



HL7 FHIR DevDays International 2022 | Hybrid Edition, Cleveland, OH | June 6–9, 2022 | @HL7 | @FirelyTeam | #fhirdevdays | [www.devdays.com](http://www.devdays.com)

ORGANIZED BY

**firely**

**HL7<sup>®</sup>**  
International

HL7<sup>®</sup>, FHIR<sup>®</sup> and the flame Design mark are the registered trademarks of Health Level Seven International and are used with permission.

## Who am I?

- Chetan Kumar Jain
- Senior Architect, Optum
- 17 Years of architectural experience in IT industry solutioning in areas of Data Analytics, Big Data, Cloud and Interoperability implementation



# Agenda

1

- Why we need bulk data transfer?

2

- FHIR Implementation Overview

3

- Bulk File Sharing Issues

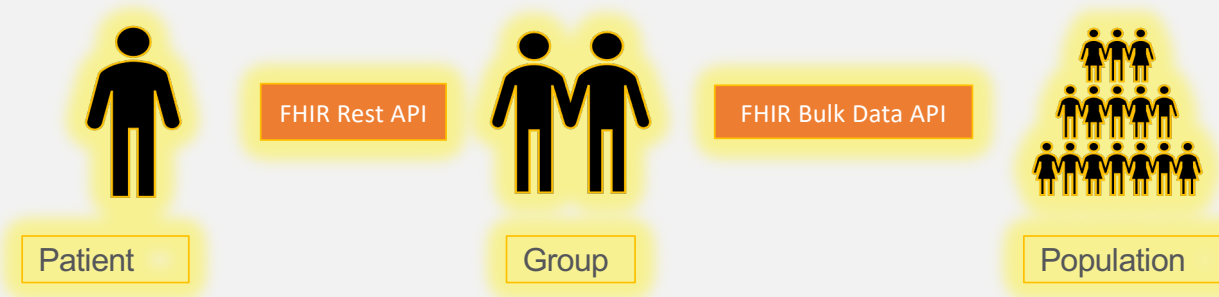
4

- Public Cloud Implementation

5

- Q&A

## Why do we need bulk transfer of data?



- **Obtaining data:** from an EHR for population-based research and training data for a machine learning algorithm or quality metrics.
- **Moving data:** in bulk between servers – perhaps when medical practices and EHR’s merge or when an organization changes their EHR. Another example would be moving directory data between servers as described in the Validated Healthcare Directory Guide (VHDir) Implementation Guide.
- **Extracting/Enriching data:** entire data sets on patient groups for analytics – patients on a clinical trial or under Case Management.

# FHIR Bulk Implementation

- **FHIR Resource** to simplify data parsing & mapping
- **FHIR Operation API** to Initiate the data extract
- **FHIR SMART Backend Services Authorization** as security model

## FHIR Operation - \$export

[FHIR Server Base]/Patient/\$export

[FHIR Server Base]/Group/[group id ]/\$export

[FHIR Server Base]/\$export

_outputFormat	The format for the generated bulk data files Currently, only <b>ndjson</b> is supported
_since	Filter results by FHIR resource modified date FHIR instant timestamp
_type	Filter results by comma delimited list of FHIR resource types

```

{
  "id": "06eb35fc-09e6-48 ... "given": ["Lucille"], "family": "Bluth"}
  "id": "cf53f382-6eb6-4f ... "given": ["George", "Oscar"], "family": "Bluth", "suffix": ["Senior"]}
  "id": "406a9c3e-50f9-4c ... "given": ["Michael"], "family": "Bluth"}
  
```

## Bulk File Sharing Issues

- **Exchange Protocol**
- **Data File Size**
- **Secure Access & Authorization**
- **Manual Process**

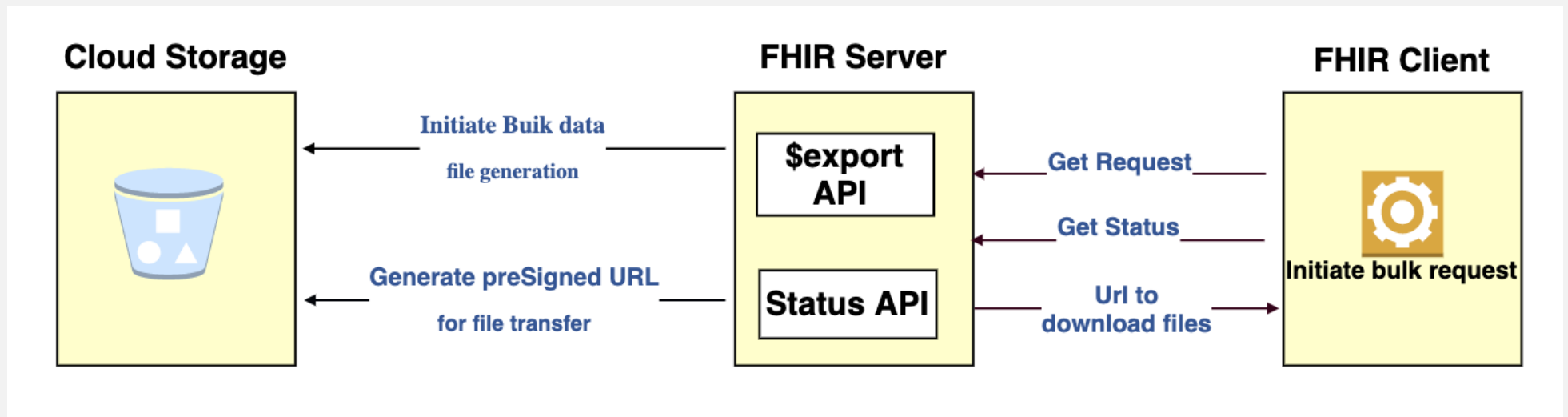
## Public cloud Storage - Self Signed URL

**AWS – Pre-Signed Url**

**Azure – Shared Access Signature**

**GCP – Signed Url**

# Public Cloud Implementation





## References

**Bulk Data Access IG** - <https://hl7.org/fhir/uv/bulkdata/>

**AWS PreSigned URL** - <https://docs.aws.amazon.com/AmazonS3/latest/userguide/using-presigned-url.html>

**Azure SAS** - <https://docs.microsoft.com/en-us/azure/storage/common/storage-sas-overview?toc=%2Fazure%2Fstorage%2Fblobs%2Ftoc.json>

**GCP Signed URL** - <https://cloud.google.com/storage/docs/access-control/signed-urls>

# Q&A

## Contact

- During DevDays, you can find / reach me here:
  - Via Whoova App – Speaker's Gallery
  - Email [chetan.jain@optum.com](mailto:chetan.jain@optum.com)

ORGANIZED BY

