UAB
Universitat Autònoma de Barcelona

Dipòsit digital de documents de la UAB

This is the **published version** of the bachelor thesis:

Cullell Homs, Júlia. *Disseny i Implementació d'una Arquitectura de Dades Cloud-Native per a la Monitorització Industrial i Analítica Avançada en Fàbriques d'Impressió 3D*. Treball de Final de Grau (Universitat Autònoma de Barcelona), 2026 (Grau en Enginyeria de Dades)

This version is available at https://ddd.uab.cat/record/337346

under the terms of the  license.

Disseny i Implementació d'una Arquitectura de Dades Cloud-Native per a la Monitorització Industrial i Analítica Avançada en Fàbriques d'Impressió 3D

Júlia Cullell I Homs

Resum– Aquest treball presenta el disseny i la implementació d'una plataforma digital avançada per a la monitorització de plantes de fabricació 3D industrial. La problemàtica central rau en la manca de control sobre la pròpia informació i en les limitacions operatives que sovint pateixen aquestes instal·lacions, on les dades de producció no són accessibles de forma immediata ni integrada. Aquesta dificultat per obtenir un flux de dades constant impossibilita tant la supervisió proactiva de la planta en temps real com l'accés a un registre històric detallat que permeti auditar la qualitat del producte un cop finalitzada la impressió. Per resoldre aquests reptes, es proposa una arquitectura orientada a esdeveniments mitjançant el protocol MQTT, que garanteix un flux de telemetria constant, fluid i amb un consum mínim de recursos. Tota aquesta informació es gestiona al núvol mitjançant un model Data Lakehouse amb processament en temps real i persistència a Google BigQuery sota una estructura evolutiva i resilient. Aquesta solució s'ha validat en un cas d'estudi real a la startup Something Added, integrant un parc de 18 impressores i 6 estacions de processament. Els resultats demostren que la nova infraestructura elimina els punts cecs d'informació, assegura una traçabilitat al llarg de tot el cicle de vida de fabricació i permet optimitzar la presa de decisions a partir de dades detallades abans inaccessibles.

Paraules clau– Indústria 4.0, Fabricació Additiva Industrial, MQTT, Sparkplug B, Internet of Things (IoT), Google Cloud Platform (GCP), Data Lakehouse, BigQuery, Cloud Run Functions.

Abstract– This project presents the design and implementation of an advanced digital platform for monitoring industrial 3D printing plants. The core problem lies in the lack of control over one's own information and the operational limitations that these facilities often face, where production data is neither immediately nor integrally accessible. This difficulty in obtaining a constant data flow prevents both proactive real-time monitoring of the plant and access to a detailed historical record that allows auditing the quality of the product once printing is completed. To address these challenges, an event-driven architecture using the MQTT protocol is proposed, guaranteeing a constant, fluid telemetry flow with minimal resource consumption. All this information is managed in the cloud through a Data Lakehouse model with real-time processing and persistence in Google BigQuery under an evolutionary and resilient structure. This solution has been validated in a real case study at the startup Something Added, integrating a fleet of 18 printers and 6 processing stations. The results demonstrate that the new infrastructure eliminates information blind spots, ensures traceability throughout the entire manufacturing lifecycle, and enables the optimization of decision-making based on detailed data previously inaccessible.

Keywords– Industry 4.0, Industrial Additive Manufacturing, MQTT, Sparkplug B, Internet of Things (IoT), Google Cloud Platform (GCP), Data Lakehouse, BigQuery, Cloud Run Functions.

1 INTRODUCCIÓ

En sectors d'alta responsabilitat com el mèdic o l'aeroespacial, on es fabriquen des d'implants ossis fins a components de motors, la impressió 3D industrial ha evolucionat de la simple prototipació cap a la producció de peces finals subjectes a normatives estrictes de seguretat i qualitat. En aquests entorns regulats, qualsevol mínima variació tèrmica o de pressió durant el procés pot comprometre la integritat del producte i, en conseqüència, la seguretat de l'usuari final. Garantir-ne la traçabilitat ja no és una opció, sinó una obligació. Un dels reptes inherents a l'enginyeria de dades industrials és adaptar l'arquitectura d'ingesta de dades a l'evolució constant de la maquinària. A mesura que les impressores incorporen actualitzacions de firmware (el programari intern que controla el funcionament del dispositiu) i nous sensors, l'estructura dels missat-

• E-mail de contacte: juliacullellhoms@gmail.com
• Treball tutoritzat per: Enric Martí Godia
• Curs 2025/26

ges emesos varia. Si no existeix una arquitectura prou flexible darrere, qualsevol d'aquestes alteracions en l'origen pot interrompre la captura de dades i trencar la continuïtat de la traçabilitat, generant costos de manteniment elevats i buits irrecuperables en l'històric de producció. El problema crític dels sistemes de monitorització tradicionals, basats en consultes síncrones (*polling*), és la seva rigidesa estructural: generen buits informatius que fan invisibles les anomalies ocorregudes entre intervals de consulta i acumulen registres redundants que inflen els costos d'emmagatzematge sense aportar valor analític. Superar aquestes limitacions és la motivació central d'aquest treball, que busca substituir aquest model fràgil i intermitent per una solució que doni el control de la informació a l'empresa operadora.

Per tot això, l'objectiu principal d'aquest projecte és dissenyar i implementar un sistema de tractament de dades industrial, escalable i independent del fabricant, capaç de garantir la interoperabilitat futura davant l'evolució dels actius de la planta, tant impressores com estacions de processament. Aquestes últimes engloben la maquinària de postprocessat encarregada de l'extracció de la peça de la cubeta i del seu desembalatge final. Això es materialitza en dues dimensions operatives clau:

- **Seguiment en viu del flux de dades**, eliminant els punts cecs informatius per permetre una monitorització en temps real.
- **Traçabilitat absoluta de cada peça impresa**, consolidant un registre històric que permeti revisar la qualitat del producte final.

Per assolir-ho, es defineixen quatre objectius específics:

1. **Model de comunicació òptim.** Analitzar i justificar el model de comunicació òptim entre els actius de planta i la infraestructura cloud, avaluant el protocol MQTT com a alternativa superior al model de consulta cíclica (polling) en termes d'eficiència de xarxa, fiabilitat i capacitat de monitorització en temps real.
2. **Model de dades agnòstic i evolutiu.** Dissenyar una arquitectura basada en Google Pub/Sub i Cloud Run Functions que permeti una entrada de dades dinàmica a BigQuery sense necessitat de reestructurar les bases de dades davant canvis de firmware.
3. **Ecosistema de gestió de dades al núvol.** Materialitzar el disseny en una plataforma operativa Data Lakehouse, validant la captació i consolidació de dades reals en un entorn de fabricació additiva industrial.
4. **Validació del salt tecnològic.** Demostrar que la nova arquitectura orientada a esdeveniments garanteix una traçabilitat completa i una visibilitat contínua de la planta, en comparació amb el sistema de consultes periòdiques anterior.

El desenvolupament d'aquest projecte ha seguit quatre fases: una anàlisi teòrica per justificar el model MQTT i revisar els estàndards de mercat existents; el disseny previ de l'arquitectura de dades (capes Bronze i Gold, Star Schema); una implementació incremental sobre el cas real de Something Added; i una validació empírica final, contrastant les hipòtesis de disseny amb dades reals de BigQuery (volum de missatges, costos i temps de resposta).

Per dur a terme aquestes fases, s'ha emprat el conjunt d'eines i serveis resumits a la Taula 1.

Eina / Servei	Ús i referències
MQTT	Protocol de comunicació entre impressores i broker [1].
MQTT Explorer	Inspecció de la jerarquia de tòpics durant la integració [2].
Telegraf	Enrutament xifrat de tòpics locals cap a GCP (Google Cloud Platform) [3].
Google Pub/Sub	Cua de missatgeria que organitza la telemetria entrant.
Cloud Run Functions	Filtratge, transformació i persistència de cada missatge.
Cloud Storage	Emmagatzematge de la dada bruta en format JSONL.
BigQuery	Emmagatzematge analític (capes Bronze/Gold) i consulta SQL.
Metabase	Disseny dels dashboards operatius [4].

Taula 1: Eines i serveis emprats en l'arquitectura dissenyada.

Pel que fa a l'estructura del document, el Marc Teòric (secció 2) introdueix els fonaments del protocol MQTT i els estàndards de semàntica industrial; el Desenvolupament (secció 3) detalla el disseny i la implementació real al cas Something Added, incloent una primera validació operativa del seu funcionament en producció; els Resultats (secció 4) contrasten quantitativament i qualitativament la nova arquitectura amb el sistema Legacy a partir de tres escenaris crítics; i les Conclusions i Millores (secció 5) recullen els resultats respecte als objectius i les línies de treball futur.

2 Marc Teòric

Aquesta secció presenta els fonaments conceptuals i tecnològics sobre els quals es recolza el sistema dissenyat, contextualitzant les decisions tècniques adoptades al llarg del projecte. Consta dels principals referents del sector i de la comunicació i estandarditzación pròpies del sistema industrial.

2.1 Referents

La digitalització de les plantes de fabricació additiva ha impulsat l'adopció del protocol MQTT [1] com a estàndard per a la comunicació entre impressores industrials i sistemes informàtics. A diferència del mètode de consulta cíclica (*polling*), MQTT es basa en un model de publicació-subscripció orientat a esdeveniments, on la informació només es transmet quan succeeix un canvi rellevant. Aquest paradigma estalvia recursos de xarxa, evita el col·lapse dels sistemes i habilita una monitorització en temps real, motiu pel qual s'ha consolidat com l'estàndard en l'Internet de les Coses Industrial (IIoT) [5].

Els principals referents del sector, com Markforged [6] i la nord-americana Stratasys [7], operen amb solucions que utilitzen MQTT de forma interna però reclouen la telemetria en ecosistemes tancats. Markforged ofereix l'entorn Eiger Dashboard [6], que permet monitoritzar en temps real l'estat de la flota de màquines, l'ús de material i el progrés dels treballs d'impressió, mentre que Stratasys centralitza el flux de treball a través de la plataforma GrabCAD

Print/Shop [7]. Malgrat la potència d'aquestes solucions, cap d'elles permet la interoperabilitat entre fabricants ni retorna el control de la informació a l'empresa operadora, que queda dependent del proveïdor per accedir a les seves pròpies dades de producció i no pot auditar-les de forma independent.

A més, cada fabricant utilitza nomenclatures, estructures de missatges i paràmetres de telemetria propis, generant un escenari de fragmentació semàntica que dificulta qualsevol anàlisi comparativa. Aquest context justifica la necessitat d'una arquitectura de dades agnòstica, capaç d'interpretar dades de qualsevol marca d'impressora, traduir-les a un llenguatge comú i retornar la sobirania de la informació a l'empresa operadora.

2.2 Comunicació i Estandardització Industrial

Aquest apartat s'estructura en tres blocs: el protocol de transport MQTT i el seu model de comunicació, l'organització interna dels missatges, i les especificacions semàntiques que busquen estandarditzar-ne el contingut entre fabricants.

2.2.1 El protocol MQTT i el model de publicació-subscripció

El protocol MQTT (*Message Queuing Telemetry Transport*), estandarditzat per l'organisme OASIS Open [8], defineix un model de comunicació asíncron basat en la publicació-subscripció: la impressora actua com un client que publica dades cap a un broker central únicament quan succeeix un fet rellevant, en lloc d'esperar que un sistema extern li pregunti periòdicament. Com s'observa a la figura 1, el broker actua com a node central que rep els missatges dels sensors i els distribueix de forma instantània als clients subscrits, sense que emissor i receptor hagin d'estar sincronitzats ni conèixer-se directament.

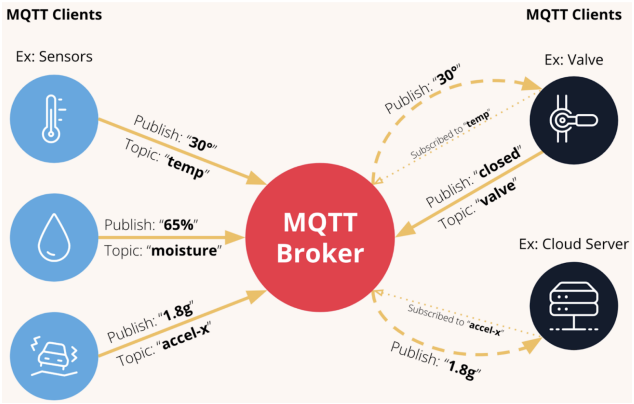


Figura 1: Diagrama de l'arquitectura de publicació-subscripció d'MQTT. [1]

A més, el protocol incorpora mecanismes de persistència de connexió, garantint un canal de comunicació fiable fins i tot en entorns amb connectivitat inestable.

2.2.2 Estructura dels missatges: tòpics i payloads

Per tal que la informació viatgi de manera ordenada a través del broker, MQTT utilitza un sistema d'adreçament jeràrquic basat en tòpics. Un tòpic és la ruta que identifica cada canal d'informació, des del node arrel fins a la mètrica concreta (per exemple: *planta/sectorA/impressora_01/temperatura*), funcionant de manera similar a l'estructura de carpetes d'un sistema de fitxers. Això permet que cada client es subscrigui únicament als canals que li són rellevants, optimitzant el flux d'informació.

Pel que fa al contingut, els missatges s'estructuren en format JSON dividint la informació en dues parts diferenciades: un bloc de **tags**, que encapsula les metadades de context (origen, *timestamp*, identificador del dispositiu), i un bloc de **fields**, que conté el payload dinàmic amb els valors en temps real dels sensors. Aquesta separació nativa entre context i telemetria és la que permet dissenyar una càrrega de dades robusta i agnòstica al núvol, independent de l'estructura interna de cada fabricant.

2.2.3 Estandardització semàntica: Sparkplug B, MTConnect i SenML

Per evitar la fragmentació en la nomenclatura de les variables entre fabricants, han aparegut diverses especificacions que busquen estandarditzar la semàntica dels missatges MQTT. L'especificació **Sparkplug B** [9] és la més rellevant en l'àmbit de la fabricació additiva industrial: com es pot apreciar a la figura 2, no només defineix com s'han d'organitzar els tòpics en quatre nivells jeràrquics (*Encoding*, *Group ID*, *Edge Node ID* i *Device ID*), sinó que també estableix una estructura rígida per al contingut dels missatges, permetent identificar instantàniament qualsevol mètrica individual dins d'entorns complexos amb centenars de nodes actius i garantint la interoperabilitat entre diferents fabricants.

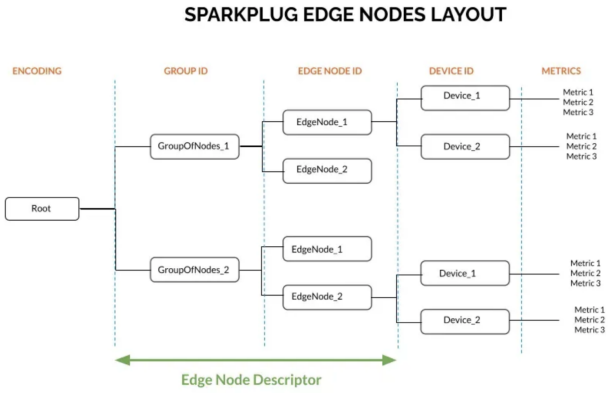


Figura 2: Estructura jeràrquica de nodes sota l'especificació Sparkplug B. [10]

En una línia similar, **MTConnect** [11] és un estàndard que defineix un vocabulari comú per a les màquines industrials. En lloc que cada fabricant hagi d'inventar els seus propis noms de variables, MTConnect estableix noms estandarditzats per als estats operatius, els paràmetres de procés i les propietats dels materials, facilitant que sistemes de diferents marques puguin comunicar-se sense necessitat de traduir el significat de cada variable. **SenML** [12], en canvi, simplifica al màxim l'escriptura per a dispositius amb recursos limitats, utilitzant etiquetes d'una sola lletra per definir el nom, la unitat i el valor de cada mètrica, garantint que la informació no perdi el seu context original sense consumir recursos de xarxa excessius.

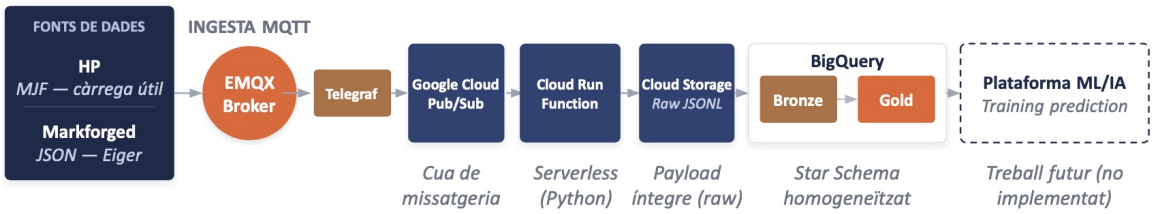


Figura 3: Esquema de l'arquitectura de dades per a la ingesta universal en una planta de fabricació additiva.

L'existència simultània d'aquestes tres especificacions (cadascuna prioritzant un objectiu diferent, ja sigui la riquesa de metadades, la compatibilitat amb maquinari divers o l'eficiència de transmissió) evidencia que el sector de la impressió 3D industrial encara no ha convergit cap a un únic estàndard semàntic, mantenint l'escenari de fragmentació que aquest treball pretén resoldre mitjançant una capa de transferència de dades agnòstica.

3 Desenvolupament

Aquest apartat es divideix en tres blocs: el disseny de l'arquitectura de dades proposada, la seva implementació al cas pràctic de Something Added i els resultats obtinguts a través dels dashboards de monitorització.

3.1 Disseny de l'Arquitectura

Per a materialitzar els objectius plantejats, s'ha dissenyat una arquitectura de dades basada en la separació estricta entre la captació i l'emmagatzematge, prioritzant la preservació de la dada original. El núvol escollit per a aquesta arquitectura és Google Cloud Platform (GCP), per la seva capacitat nativa de gestionar fluxos de dades massius sota un paradigma serverless, on la infraestructura s'autogestiona i escala dinàmicament segons la demanda.

Com s'observa a la figura 3, el flux s'inicia a les fonts de dades industrials (esquerra), que emeten telemetria via MQTT cap al broker EMQX del Command Center. L'agent Telegraf [3] intercepta aquests missatges i els enruta de forma segura i xifrada cap a *Google Cloud Pub/Sub*, que actua com a cua de missatgeria asíncrona organitzant els paquets segons el seu tòpic d'origen. Cada vegada que arriba un missatge, un trigger activa automàticament la *Cloud Run Function* vinculada al tòpic d'origen, que persisteix la dada bruta al *Cloud Storage Bucket* en format JSONL i la insereix a les taules analítiques de BigQuery. Finalment, aquesta base de dades centralitzada constitueix els fonaments per a una futura Plataforma ML/IA d'entrenament predictiu, orientada a anticipar fallades de qualitat i optimitzar la planificació de la producció.

L'emmagatzematge a BigQuery s'organitza en dues capes estratègiques. La **capa Bronze** utilitza una estructura híbrida que combina un esquelet de metadades fixes (timestamp, identificador del command center i tòpic MQTT) amb una columna de tipus JSON que emmagatzema el payload original íntegre tal com arriba de la impressora. La clau d'aquest disseny rau precisament en aquest camp dinàmic: si un fabricant actualitza el firmware i canvia el nom o l'estructura de les variables emeses, el sistema no col·lapsa ni perd dades, ja que el payload es guarda íntegre en el seu format natiu sense cap transformació prèvia. Això elimina la rigidesa estructural dels models relacionals convencionals, on qualsevol canvi en l'origen obliga a reescriure l'esquema de la base de dades i genera interrupcions irrecuperables en la traçabilitat.

La **capa Gold** transforma aquesta informació heterogènia en un model optimitzat per a l'anàlisi de negoci. S'ha escollit una arquitectura OLAP (*Online Analytical Processing*) en lloc d'un model transaccional perquè l'objectiu no és gestionar operacions individuals sinó realitzar consultes analítiques complexes sobre grans volums de dades històriques. Concretament, s'ha adoptat un *Star Schema* [14], l'estàndard en intel·ligència de negoci per a aquest tipus d'entorns, perquè minimitza el nombre de *joins* necessaris i optimitza la velocitat de resposta de les consultes. Per comprendre l'estructura d'aquest model, cal entendre primer el flux operatiu d'una planta de fabricació additiva: el procés s'inicia amb el carro de subministrament (*trolley*), que conté el material polimèric en pols i s'insereix a la impressora per executar el treball. Un cop finalitzada la impressió, el material es transfereix a una unitat de refredament natural (NCU). El temps de refredament varia en funció del nombre de capes del treball i del volum de la peça, i cal esperar que finalitzi completament abans de procedir al desembalatge per obtenir les peces finals. Aquest cicle de vida és el que defineix l'estructura de taules necessària.

Com es detalla a l'Annex A.1, el model s'organitza al voltant d'una taula de fets central (*JobSummary*) que registra cada treball de fabricació, envoltada de taules de dimensions descriptives com *Dimension_Printer*, *Dimension_Material*, *Dimension_NCU* o *Dimension_Trolley*, i taules de fets addicionals com *Fact_Telemetry*, *Fact_Alerts*, *Fact_Postprocessing*, *Fact_Maintenance* o *Fact_Parts_Consumption*. Aquest model constitueix el mínim viable necessari per garantir la traçabilitat completa de la planta d'acord amb els objectius del projecte. La seva naturalesa modular permet, no obstant, incorporar noves taules o columnes en el futur sense comprometre l'estabilitat de l'arquitectura existent.

Per optimitzar el rendiment de les consultes, s'apliquen dues tècniques complementàries: el *Clustering* per *PrinterSerialNumber* i *JobID* agrupa físicament els registres de la mateixa impressora en blocs contigus, de manera que quan es filtra per una màquina concreta, BigQuery accedeix directament als blocs rellevants sense escanejar tota la taula. El *Partition Pruning* per data complementa aquesta optimització permetent que el motor descarti automàticament totes les particions fora de l'interval temporal sol·licitat. Combinades, aquestes dues tècniques eviten que BigQuery hagi

d'escanejar la totalitat de l'historial de telemetria cada vegada que un enginyer vulgui consultar, per exemple, l'estat d'una impressora concreta durant un dia específic: el sistema accedeix únicament a la partició d'aquell dia i al bloc corresponent a aquella màquina, transformant consultes que podrien trigar minuts sobre volums massius de dades en respostes de pocs segons, reduint dràsticament els costos operatius i garantint la viabilitat econòmica del sistema a llarg termini.

Pel que fa a la **seguretat i governança** de la dada, la comunicació es xifra mitjançant el protocol TLS i l'autenticació es gestiona amb credencials pròpies del projecte de *Google Cloud Platform*. A nivell d'accés, s'aplica el sistema IAM (*Identity and Access Management*) per definir rols diferenciats segons el component (separant permisos d'escriptura per a les *Cloud Run Functions* dels de només lectura per a la visualització), una mesura rellevant en un projecte centrat en la sobirania de la dada. Les dades brutes es retenen indefinidament al *Cloud Storage Bucket* per garantir l'auditoria futura.

3.2 Cas Pràctic: Something Added

Something Added és una startup industrial especialitzada en la fabricació additiva de soles de bambes d'alt rendiment per a running. Aquesta startup representa un cas d'estudi ideal per validar l'arquitectura proposada: una planta en creixement actiu, amb un parc de 18 impressores HP i 6 estacions de processament, i una problemàtica real i documentada de monitorització.

Com s'il·lustra a la figura 4, el diagrama compara dues línies temporals paral·leles: la del model de consulta cíclica i la del model MQTT orientat a esdeveniments. En l'exemple representat, una alerta crítica apareix a les 10:21h i es resol per si mateixa a les 10:23h –un interval de només dos minuts–, quedant completament emmarcada entre dues consultes consecutives del sistema Legacy (10:20h i 10:25h). Com a conseqüència, aquesta incidència mai arriba a quedar registrada: el sistema de polling només comprova l'estat de la impressora en instants discrets i, si l'alerta apareix i es resol entre dos d'aquests instants, la seva existència passa completament desapercebuda.

L'arquitectura original de la planta es basava en aquest mateix model de consulta cíclica, mitjançant l'API proporcionada per HP, amb peticions HTTP programades cada 2, 5 o 10 minuts al Command Center –el servidor connectat a les impressores. A més, com que no es feia cap validació prèvia del contingut rebut, es guardaven contínuament milers de registres idèntics de manera redundant, generant sobrecostos d'emmagatzematge innecessaris. Davant d'aquesta situació, s'ha aprofitat la capacitat nativa de les impressores HP per integrar-se amb un broker MQTT local a la fàbrica, implementant el sistema dissenyat en aquest treball i establint les bases per a un control de qualitat rigorós i una analítica de dades avançada.
La implementació es descriu en quatre etapes: la connexió al broker i la inspecció inicial de la telemetria, la selecció i el filtratge dels tòpics importants, la intermediació de Telegraf cap al núvol, i el processament final mitjançant les *Cloud Run Functions*.



Figura 4: Model API Polling (Legacy) enfront de Model MQTT Event-Driven. L'alerta crítica de les 10:21h queda invisible per al sistema Legacy.

3.2.1 Connexió al Broker i Inspecció de la Telemetria

Per establir la comunicació amb el broker MQTT de la fàbrica, s'han configurat els paràmetres d'accés proporcionats per HP: protocol, adreça del host, port de comunicació i credencials d'autenticació. Aquests paràmetres garanteixen un canal de comunicació segur i estable que permet capturar cada esdeveniment generat a la fàbrica en temps real.

Per validar aquesta connexió i inspeccionar el flux de missatges abans del seu processament, s'ha utilitzat MQTT Explorer [2]. Aquest programari actua com un client d'escriptori que permet visualitzar els tòpics de forma jeràrquica i inspeccionar els paquets de dades JSON enviats per les impressores, organitzats en una estructura d'arbre que permet navegar des de l'identificador general de la fàbrica fins als canals específics de cada impressora (vegeu aquesta estructura d'arbre a l'Annex A.2).

Pel que fa a l'estructura dels missatges, les impressores HP segueixen l'estàndard Sparkplug B [9] descrit a la secció 2.2.3. Com s'observa a la Figura 5, el camp **tags** conté el *topic* complet d'origen –identificador únic del dispositiu i la mètrica– juntament amb el *host* i el *target_topic* assignat per Telegraf, mentre que el camp **fields** conserva el payload original íntegre dins de *value*, sense cap processament previ.

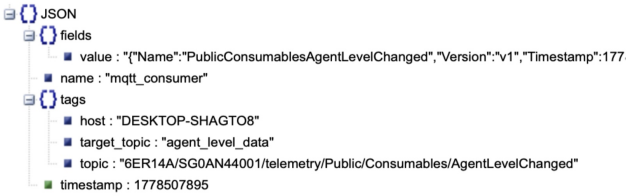


Figura 5: Estructura en arbre del missatge JSON rebut a *Google Cloud Pub/Sub*, amb les metadades de context (*tags*) i el payload dinàmic de la telemetria (*fields*) diferenciats.

3.2.2 Selecció i Filtratge de Tòpics Crítics

Un cop explicada la jerarquia de tòpics disponibles –les rutes jeràrquiques que identifiquen cada canal d'informació, tal com s'ha descrit a la secció 2.2.2, s'ha realitzat un procés rigorós de selecció i filtratge, amb un doble objectiu: per una banda, capturar l'estat de la fàbrica en temps real; per altra banda, emmagatzemar de manera immutable la informació detallada de cada peça un cop finalitzat el procés de fabricació. Aquest criteri de selecció optimitza els recursos al núvol, evitant la captació de telemetria redundant o sorollosa.

Després d'una exhaustiva fase d'avaluació, s'han seleccionat deu tòpics crítics distribuïts en tres mòduls funcionals que cobreixen les necessitats de traçabilitat de la planta (es referencien pel nom de l'esdeveniment dins del tòpic, omitint el path jeràrquic complet per simplicitat; per exemple, *BuildStart* correspon en realitat al tòpic complet *[ID]/[Sèrie]/telemetry/Public/BuildManager/BuildStart*):

- **1. Mòdul de Control d'Estat i Sistemes (3 tòpics):** *Heartbeat* (estat de connexió de la màquina en temps real), *SystemErrorRaised* (notificació i obertura d'incidències tècniques) i *SystemErrorCleared* (resolució d'incidències tècniques). Aquest mòdul garanteix la visibilitat absoluta de l'estat operatiu de cada impressora.
- **2. Mòdul de Producció (4 tòpics):** *BuildStart* (inici del treball de fabricació), *BuildEnd* (finalització del treball a la impressora), *BuildReport* (resum complet de mètriques i paràmetres del build acabat, incloent dades capa per capa) i *BuildUnit/StatusChanged* (canvis d'estat a la unitat de construcció). Aquest mòdul permet reconstruir el cicle de vida complet de cada peça fabricada.
- **3. Mòdul de Logística i Manteniment (3 tòpics):** *AgentLevelChanged* (nivell d'agents i consumibles de fusió químics dels quatre tancs), *WiperStatusChanged* (estat del mecanisme de neteja) i *FilterStatusChanged* (estat i saturació dels filtres d'extracció). Aquest mòdul anticipa les necessitats de manteniment preventiu de la planta.

A l'Annex A.3 es detallen els deu tòpics d'informació seleccionats. Amb aquesta selecció es disposa del volum de dades necessari i suficient per materialitzar el model dimensional en estrella definit a la nostra arquitectura.

3.2.3 Telegraf: Intermediació cap al Núvol

Per interconnectar el broker MQTT local amb Google Cloud Platform (GCP), s'ha implementat l'agent Telegraf al Command Center de la fàbrica. La seva funció principal és assignar a cada tòpic industrial seleccionat una ruta específica cap a *Google Cloud Pub/Sub*, garantint que cap missatge es perdi en el trànsit local. L'autenticació es realitza mitjançant un fitxer de credencials en format JSON que permet autenticar-se de manera xifrada contra el projecte de Google, assegurant un canal de comunicació segur entre la planta i el núvol.

Els missatges arriben de manera segura als tòpics de *Google Cloud Pub/Sub*, que els reben de forma aïllada segons la seva tipologia –*system-alert-raised*, *agent-level-data*, *heartbeat-data*–, els organitzen i els mantenen a l'espera de ser processats per les *Cloud Run Functions*. Aquest disseny modular impedeix l'encreuament de telemetries de diferent naturalesa tècnica i facilita l'escalabilitat futura de la planta de fabricació.

3.2.4 Cloud Run Functions: Processament Actiu i Lògica d'Optimització

El darrer pas del sistema és la transformació i l'emmagatzematge intel·ligent de les dades. Cada vegada que un tòpic de Pub/Sub rep un missatge, s'activa automàticament la *Cloud Run Function* específica, programada en Python, amb tres objectius: persistir la dada bruta al *Cloud Storage* en format JSONL, inserir-la a la taula Bronze de BigQuery mantenint el format natiu, i processar-la per inserir-la a la taula Gold optimitzada per a consulta ràpida.

La intel·ligència d'aquestes funcions rau en el filtratge intel·ligent segons la naturalesa de cada tòpic. El tòpic *Heartbeat*, per exemple, s'emet de manera asíncrona cada vegada que la impressora finalitza una capa –aproximadament cada quinze segons–, però això no implica un canvi en l'estat general de la màquina, que sol mantenir-se constant durant centenars de capes. Per això, la funció descarta el missatge si l'estat és idèntic a l'últim registrat, consolidant només els canvis reals –com el pas a *CoolingHot* o l'aparició d'un error– i el progrés de capes a intervals regulars (cada 100 capes), suficient per al seguiment del percentatge d'avenç sense comprometre la captura íntegra dels canvis d'estat crítics. Aquest mecanisme evita milers de files redundants sense renunciar a la fiabilitat informativa que el distingeix del model de polling Legacy.

De manera similar, el nivell dels quatre tancs d'agents químics es reporta nativament cada vegada que disminueix un 1%, però la funció només persisteix el registre si la variació respecte a l'últim valor és com a mínim del 5%. L'anàlisi d'una mostra real de 482 esdeveniments d'una impressora durant un mes confirma que aquest filtre redueix els registres persistits en un 51%, malgrat que el consum real –amb oscil·lacions per recàrregues parcials– no segueix un patró perfectament lineal.

Pel que fa a la persistència, la informació de Something Added respecta l'estructura híbrida del disseny teòric: cada missatge s'insereix amb les metadades fixes i el payload original íntegre, sense parsing previ. A la capa Gold, les funcions apliquen una lògica de dues fases: el *BuildStart* genera un registre amb la planificació esperada (material, volum, capes estimades), i el *BuildEnd* l'actualitza –via *JobID*– amb els detalls reals de tancament (temps d'execució, consum de material, estat final), permetent una traçabilitat completa del cicle de vida de cada impressió.

Finalment, pel que fa a la gestió d'errors, les *Cloud Run Functions* incorporen reintents automàtics (fins a 3 intents) davant de fallades transitòries, complementats amb un software extern de monitorització. Durant el període de validació no s'ha registrat cap error de processament, gràcies a l'estructura consistent dels missatges Sparkplug B emesos per les impressores HP.

3.3 Dashboards de Monitorització

Per a la capa de visualització, s'ha optat per Metabase [4] enfront d'altres alternatives del mercat com Looker Studio [15], Grafana [16] o Power BI [17]. La decisió respon principalment a la seva naturalesa open-source –que evita costos de llicència recurrents en una startup en fase de creixement–, a la seva integració nativa amb BigQuery mitjançant consultes SQL directes –sense necessitat de connectors o transformacions intermèdies–, i a la simplicitat de la seva interfície de creació de dashboards.

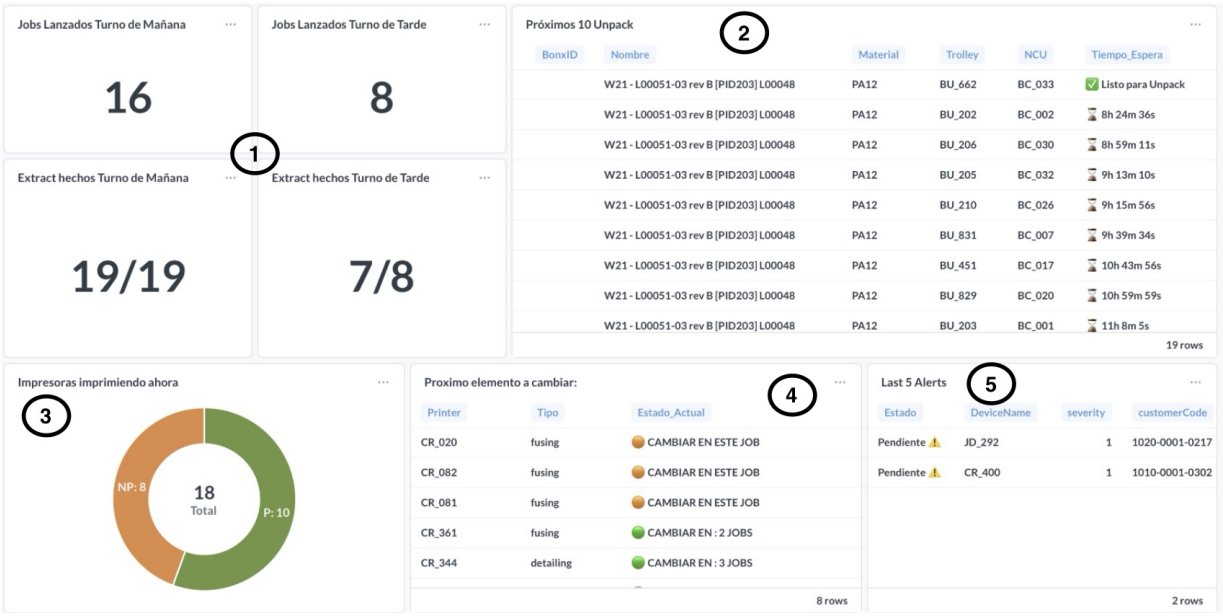


Figura 6: Interfície visual del panell de control operatiu (Factory Screen) que permet als operaris tenir una gestió integrada de la planta.

Per aprofitar tota la informació unificada arran de la transició cap a la nova base de dades, s'ha dissenyat amb Metabase un entorn visual interactiu connectat directament a les taules de fets i dimensions de la base de dades Gold de BigQuery. Aquest sistema es divideix en dos panells especialitzats: una pantalla de fàbrica global per a la coordinació operativa diària, i un panell de seguiment del registre de cada treball de fabricació. D'aquesta manera es cobreixen les necessitats de tots els perfils d'usuari de la planta, tant del director, que necessita conèixer el registre històric, com de l'operari a peu de fàbrica, que ha de poder fer un seguiment en viu.

3.3.1 Pantalla de Fàbrica Global (Factory Screen)

Aquest primer panell es projecta en una pantalla de grans dimensions al mig de la planta per coordinar el treball de l'equip a peu de fàbrica. Com s'observa a la figura 6, integra en una sola vista tota la informació operativa necessària per gestionar el dia a dia de la producció.

El panell ofereix cinc blocs d'informació clau.

El bloc **(1) control de torns** mostra en temps real quantes feines s'han llançat al torn de matí i de tarda, i quantes operacions d'extracció s'han completat o queden pendents durant aquell torn —independentment de si el treball corresponent es va llançar en un torn anterior, ja que el temps de refredament d'una peça pot allargar-se més enllà del torn en què es va iniciar la impressió. El bloc **(2) pròxims desembalatges** mostra una cua amb els següents deu treballs a desempaquetar, indicant per a cadascun si encara es troba en fase de refredament –amb el temps restant estimat– o si ja ha finalitzat aquest procés i està llest per a l'*unpack*; aquesta dada és important per a la qualitat del producte, ja que si l'operació de desembalatge es realitza abans que el refredament natural a la NCU hagi finalitzat, la peça encara calenta pot deformar-se irreversiblement. El bloc **(3) estat de les impressores** mostra mitjançant un gràfic circular quantes màquines estan imprimint (P, en verd) i quantes estan sense imprimir (NP, en taronja). El bloc **(4) manteniment de consumibles** desglossa per impressora i tipus d'agent –fusing o detailing– quantes feines falten abans que s'esgoti cada consumible, prediuint quin serà el proper element que s'haurà de canviar i permetent planificar-ne la reposició abans que es produeixi una aturada no planificada. Finalment, el bloc **(5) alertes actives** mostra les darreres incidències detectades, identificant el dispositiu afectat i el seu nivell de severitat. En conjunt, aquests cinc blocs permeten a l'operari seguir tot el cicle de vida de la producció –des de l'inici del torn fins al desembalatge final– des d'una única pantalla.

La captura real de producció presentada a la figura 6 il·lustra el funcionament del sistema: en el moment de la consulta, el gràfic **(3)** reflecteix que 18 de les impressores estaven actives (10 imprimint, 8 sense activitat), mentre el control de torns (1) mostrava 16 feines llançades i 19 extraccions completades al torn de matí —xifra superior al nombre de feines llançades aquell mateix torn, ja que inclou extraccions de treballs iniciats en torns anteriors i finalitzats durant aquest—, coincidint amb l'estat real de la planta en aquell instant. A més, el sistema havia identificat de forma autònoma cinc impressores amb consumibles pròxims a esgotar-se al requadre **(4)**, i dues alertes actives de severitat baixa al llistat **(5)**. Si bé aquesta captura puntual no permet generalitzar el comportament del sistema a llarg termini, sí que constitueix una evidència concreta que la informació mostrada no és una simulació, sinó el reflex de l'estat operatiu real de la planta capturat via MQTT.

3.3.2 Històric i Traçabilitat Avançada de Treballs (Tracking Prints)

El segon component està destinat al seguiment detallat de cada feina enviada a producció i representa la millora més significativa de la nostra arquitectura respecte al sistema anterior. Com s'observa a la figura 7, cada fila de la taula

deviceName	BU	BuildUUID	Job_name	Material	Status	LayerPct	StartTime	FinishedTime
CR_020	BU_204	d380305f-a71f-469f-a257-ff98368f0604	[PID203] W25 - L00051-03 rev B Upper mode x15	PA12	Printing	71.9%	June 18, 2026, 9:12 AM	
CR_062	BU_202	1edbc0aa-adee-48e2-8e19-58880995cdf2	[PID203] W25 - L00051-03 rev B Upper mode x15	PA12	Printing	78.2%	June 18, 2026, 8:36 AM	
CR_096	BU_004	3f62d623-4824-48a2-9b6f-7dcd97682987	[PID203] W25 - L00051-03 rev B Upper mode x15	PA12	Printing	87.8%	June 18, 2026, 7:42 AM	
CR_400	BU_754	9298ef01-c2ca-4779-92e4-ad550c45ed55	[PID203] W25 - L00051-03 rev B Upper mode x15	PA12	Printing	87.7%	June 18, 2026, 7:42 AM	
CR_082	BU_210	a19bd6e9-b42d-4995-b65a-8044f835f7f5	[PID203] W25 - L00051-03 rev B Upper mode x15	PA12	Printing	87.6%	June 18, 2026, 7:40 AM	
CR_403	BU_756	b0328f6c-42cc-4ea1-a76c-6f9644317df6	[PID203] W25 - L00051-03 rev B Upper mode x15	PA12	Printing	87.7%	June 18, 2026, 7:39 AM	
CR_202	BU_005	08e2b3d2-2278-4a99-9eed-374faf820b6f	[PID203] W25 - L00051-03 rev B Upper mode x15	PA12	Printing	89.9%	June 18, 2026, 7:28 AM	
CR_361	BU_006	5736e5f7-dae1-4b67-afd5-1d527dc0d9d2	[PID203] W25 - L00051-03 rev B Upper mode x15	PA12	Printing	89.9%	June 18, 2026, 7:24 AM	
CR_362	BU_451	fc418901-55fb-4a36-a7f4-1469c0b13bcf	[PID203] W25 - L00051-03 rev B Upper mode x15	PA12	Printing	89.9%	June 18, 2026, 7:22 AM	
CR_204	BU_051	607b7328-3e22-4785-a419-f287b3e0e23c	[PID203] W25 - L00051-03 rev B Upper mode x15	PA12	Printing	94.4%	June 18, 2026, 7:04 AM	
CR_400	BU_943	c209d6ac-19c7-440d-8281-d711e2d6e0ec	L00051 [PID203] W18	PA12	Printin...	0%	April 30, 2026, 10:06 PM	April 30, 2026, 10:21 PM
CR_361	BU_746	4a2de44c-863a-4198-aab5-2f3df8dbbd7a	L00051 [PID203] W18	PA12	Succee...	100%	April 30, 2026, 10:02 PM	May 1, 2026, 7:50 AM
CR_202	BU_207	f4c7b6ee-0163-4dc5-b45c-24214e014a5d	L00051 [PID203] W18	PA12	Succee...	100%	April 30, 2026, 10:02 PM	May 1, 2026, 7:55 AM
CR_344	BU_001	c5e57c31-b2c0-4e7b-82cb-81e801caff1d	L00051 [PID203] W18	PA12	Warning	100%	April 30, 2026, 10:01 PM	May 1, 2026, 7:53 AM
CR_082	BU_202	b98abe41-22fa-426e-8ecd-7f97117c58a8	L00051 [PID203] W18	PA12	Succee...	100%	April 30, 2026, 9:35 PM	May 1, 2026, 7:27 AM
CR_362	BU_005	252d6d52-6365-4063-bb71-a4858ca4305a	L00051 [PID203] W18	PA12	Warning	100%	April 30, 2026, 8:56 PM	May 1, 2026, 6:46 AM
CR_062	BU_205	e7fa587f-c952-497b-838a-36cfd57c7284	L00051 [PID203] W18	PA12	Succee...	100%	April 30, 2026, 7:21 PM	May 1, 2026, 5:12 AM

Figura 7: Panell de seguiment (Tracking Prints) amb capacitat de filtratge dinàmic de comandes i espai de temps.

correspon a un treball d'impressió individual, identificat pel seu nom de fitxer (*JobName*) i un identificador únic de construcció (*BuildUUID*), mentre que les columnes recullen la informació rellevant per al seu seguiment: la impressora i la unitat de refredament (*Trolley/NCU*) assignades, el material utilitzat, l'hora d'inici, l'estat actual del treball (*Status*) i el percentatge de progrés (*Percent_Complete*). Aquesta estructura permet localitzar qualsevol comanda i avaluar el seu estat de forma centralitzada.

El panell ofereix cercadors dinàmics per localitzar comandes de clients específics, registra automàticament la unitat de refredament natural assignada a cada cubeta i calcula els temps totals de refredament. El punt més valuós d'aquest panell és el **control de qualitat i feines fallides**: el sistema contrasta directament les peces sol·licitades en el disseny amb les realment completades de forma òptima, permetent identificar desviacions de producció abans inadvertides i decidir si cal reprogramar o repetir un lot concret.

Aquesta capacitat de seguiment en directe és possible gràcies al creuament de les dades de progrés de cada impressió, registrades pel sistema de filtratge intel·ligent descrit a la secció 3.2.4, que permet mostrar a la columna *Percent_Complete* el percentatge exacte d'avenç de cada treball sense necessitat d'esperar a la seva finalització.

La captura real demostra aquesta capacitat amb un cas concret extret de la figura 7: una de les comandes registra un estat *Succeeded* a la columna *Status*, amb totes les peces completades sense cap merma, mentre que el sistema manté el desglossament exacte de quantes unitats s'han fabricat correctament respecte a les sol·licitades inicialment per a cada lot. Aquesta capacitat d'auditoria automàtica, abans completament inexistent amb el sistema Legacy, valida que la infraestructura no només captura dades en brut sinó que les transforma en informació de negoci necessària per a la presa de decisions.

4 RESULTATS

Aquest apartat es divideix en dos blocs: la validació del salt tecnològic assolit, contrastant quantitativament i qualitativament la nova arquitectura amb el sistema Legacy, i la validació en producció sobre diferents escenaris del cas pràctic de Something Added.

4.1 Validació del Salt Tecnològic

Per a validar el nostre sistema, resulta imprescindible contrastar el funcionament de la nova arquitectura orientada a esdeveniments enfront del comportament del sistema Legacy. Donat que els dos sistemes operen sota paradigmes fonamentalment diferents –un basat en consultes periòdiques i l'altre en esdeveniments asíncrons– la validació combina una dimensió qualitativa –demostrar que la nova arquitectura és capaç de capturar informació que l'anterior, per disseny, era incapaç de registrar– amb una dimensió quantitativa, mitjançant l'anàlisi de dades reals de producció que permeten mesurar l'abast d'aquesta millora. La validació s'estructura en tres escenaris crítics de producció i infraestructura que demostren el valor real de la transició tecnològica.

4.1.1 Escenari 1: Eliminació dels Punts Cecs Informatius

Tal com s'ha justificat a la secció 2 i s'il·lustra a la figura 4, el sistema Legacy és estructuralment incapaç de registrar esdeveniments ocorreguts entre dues consultes consecutives. Aquesta capacitat de resposta immediata és especialment important en aplicacions on les propietats mecàniques de la peça són determinants per al rendiment final, com és el cas de les soles esportives d'alt rendiment fabricades per Something Added.

Com a exemple real: el dia 17 de juny de 2026 es va registrar una alerta de severitat crítica (nivell 3) a l'impressora CR_344 a les 06:04:13h, resolta de manera automàtica només 14 segons després, un esdeveniment que qualsevol

model de consulta cíclica –inclosa la freqüència més agressiva de 2 minuts disponible a l'antiga API– hauria deixat invisible a l'històric.

Per quantificar l'abast d'aquest problema, s'ha analitzat el registre complet de *Fact_Alerts* a BigQuery durant 30 dies, identificant 157 incidències. D'aquestes, un **42,0%** van tenir una durada inferior als 2 minuts, i un **65,6%** inferior als 10 minuts: més de quatre de cada deu incidències reals haurien passat desapercebudes fins i tot amb la configuració de polling més exigent disponible, confirmant que l'eliminació dels punts cecs és una millora estructural i no un benefici puntual.

4.1.2 Escenari 2: Eficiència en el Volum de Dades i Contenció de Costos

El sistema Legacy opera mitjançant una replicació cega: encara que l'estat de la màquina no hagi variat, la informació es torna a escriure a la base de dades en cada consulta, generant milers de registres redundants sense valor analític. La nova arquitectura trenca amb aquesta ineficiència mitjançant el filtratge intel·ligent per tòpic, descartant tot missatge que no representi un canvi real d'estat.

S'ha analitzat el volum real de registres a *Fact_Heartbeat* durant 14 dies d'operació normal de les 18 impressores, obtenint una mitjana de 2.060 registres diaris. Atès que aquest tòpic reporta l'estat cada quinze segons, el volum potencial sense filtratge hauria estat de 18 impressores $\times$ 4 missatges/minut $\times$ 1.440 minuts/dia $\approx$ 103.680 missatges/dia, de manera que la lògica implementada redueix el volum d'escriptures en aproximadament un **98%**.

De manera complementària, s'ha comparat el volum de dades processat per una consulta filtrada per un dia concret enfront de la mateixa consulta sense aprofitar la partició per *event_date*: 2,13 MB davant de 102,95 MB, una reducció també del **98%**, ara en dades escanejades per consulta. Tot i coincidir en xifra, es tracta de dues optimitzacions independents –escriptura i lectura– que es complementen per garantir la viabilitat econòmica del sistema.

Aquesta eficiència es tradueix en un cost operatiu contingut: l'anàlisi de la facturació real a Google Cloud Platform durant el mateix període mostra un cost mitjà de 8,10€/dia (uns 243€/mes) pels serveis de BigQuery, Cloud Run Functions i Cloud Storage. Aquesta xifra no es pot contrastar amb un cost equivalent del sistema Legacy, que no disposa de facturació centralitzada ni de registre històric de costos –una de les mancances de control que motiven aquest treball–. Tot i així, el nou model *serverless* elimina costos estructurals propis de l'arquitectura anterior, com el manteniment del servidor on-premise o la dependència d'un servei de *tunneling* extern (LocalXpose [13]).

4.1.3 Escenari 3: Arquitectura Cloud-Nativa versus Sistema Local

En l'entorn Legacy, la base de dades resideix en un ordinador físic a les oficines de la fàbrica (*On-Premise*), amb recursos limitats que impedeixen guardar històrics massius o connectar noves màquines sense col·lapsar el sistema. L'accés a les dades fora de la xarxa local només és possible mitjançant un túnel configurat amb LocalXpose [13], una solució que afegeix un punt addicional de fragilitat i dependència tècnica a un sistema ja limitat per disseny.

La nova arquitectura planteja un model híbrid: el broker MQTT es manté localment al Command Center per garantir la transferència de dades a peu de planta, mentre que tota la resta s'executa al núvol. El volum d'emmagatzematge s'adapta elàsticament al flux real de missatges, garantint que el sistema no deixi mai de transferir, transformar i persistir dades independentment del volum generat, i fent accessible la informació des de qualsevol ubicació en tot moment.

5 CONCLUSIONS I MILLORES

S'ha desenvolupat un sistema de tractament de dades industrial, escalable i independent del fabricant, que s'ha validat en l'empresa Something Added, assolint els quatre objectius específics plantejats:

- **Model de comunicació òptim**: l'anàlisi de 157 incidències reals confirma que un 42% haurien quedat invisibles sota qualsevol freqüència de polling, validant empíricament MQTT com a alternativa superior.
- **Model de dades agnòstic i evolutiu**: la capa Bronze permet incorporar nous fabricants reconfigurant Telegraf i adaptant les *Cloud Run Functions* corresponents, sense redissenyar l'esquema de dades al núvol –capacitat encara no validada amb maquinari de tercers.
- **Ecosistema de gestió de dades al núvol**: materialitzat en una plataforma Data Lakehouse en producció real, amb un cost contingut de 8,10€/dia gràcies al filtratge intel·ligent i les tècniques de Clustering i Partition Pruning, que redueixen en un 98% tant el volum d'escriptura com el de lectura.
- **Validació del salt tecnològic**: demostrada pels tres escenaris de la secció 4, confirmant una millora estructural –no puntual– respecte al sistema Legacy.

En la pràctica, això es tradueix en un **seguiment en viu** de la fàbrica i una **traçabilitat absoluta** de la peça, consolidada a BigQuery per analitzar consums, diagnosticar fallades i predir temps d'execució futurs. Com a línies de millora futures, es proposen:

- **Analítica predictiva**, amb models de Machine Learning entrenats sobre l'històric, per anticipar fallades i reduir el temps d'inactivitat.
- **Manteniment intel·ligent**, generant automàticament ordres de treball davant d'alertes crítiques.
- **Integració física**, amb etiquetes electròniques (ESL tags) que mostrin informació rellevant a peu de màquina.
- **Validació amb més fabricants**, per confirmar l'agnosticisme amb maquinari real d'altres proveïdors.

Aquest treball presenta limitacions d'abast en la validació empírica, relacionades amb la durada de les finestres temporals i amb l'ús d'un únic fabricant (vegeu l'Annex A.4).

En definitiva, el projecte demostra que la combinació d'un protocol de comunicació estàndard amb una arquitectura cloud-native permet transformar una planta opaca i reactiva en un ecosistema de dades transparent, aportant un valor de negoci mesurable.

AGRAÏMENTS

En primer lloc, vull agrair l'esforç i la dedicació de l'Enric, el meu tutor d'aquest projecte, per la seva orientació i els seus consells al llarg de tot el projecte, i per haver acceptat la meva proposta amb ganes des de l'inici.

També vull agrair a les meves companyes de la universitat, la Mireia i la Martina, que han fet especials aquests quatre anys. I a en Carlos, la meva parella, per estar al meu costat en cada moment d'aquest camí i fer-me costat sense demanar res a canvi.

Finalment, als meus germans, al meu pare, i sobretot a la meva mare. Perquè si estic entregant aquest projecte i, amb això, finalitzant la carrera, és gràcies a la constància que ha tingut amb mi, per animar-me sempre a seguir endavant i per confiar en mi quan ni jo mateixa em pensava que podia.

ÚS D'INTEL·LIGÈNCIA ARTIFICIAL

Durant la redacció d'aquest treball s'ha utilitzat l'assistent d'Intel·ligència Artificial Claude com a eina de suport per millorar la redacció del text. El disseny, la implementació i les conclusions del treball són responsabilitat exclusiva de l'autora.

REFERÈNCIES

[1] MQTT. 2026. MQTT: The Standard for IoT Messaging. [online] Available at: https://mqtt.org/ [Accessed 21 March 2026].

[2] MQTT Explorer. 2024. MQTT Explorer – A comprehensive MQTT Client. [online] Available at: https://mqtt-explorer.com/ [Accessed 19 March 2026].

[3] InfluxData. 2024. Telegraf Open Source Data Collector. [online] Available at: https://www.influxdata.com/time-series-platform/telegraf/ [Accessed 20 March 2026].

[4] Metabase. 2024. Metabase – Business Intelligence, Dashboards, and Data Visualization. [online] Available at: https://www.metabase.com/ [Accessed 10 April 2026].

[5] Industry IoT Consortium. 2026. *What Is IIoT? The Industrial Internet of Things Primer*. [online] Available at: https://www.iiconsortium.org/2019/09/what-is-iiot-the-industrial-internet-of-things-primer/ [Accessed 28 June 2026].

[6] Markforged. 2026. Eiger Software – Slicing & Print Management. [online] Available at: https://markforged.com/software/eiger [Accessed 18 June 2026].

[7] Stratasys. 2026. GrabCAD Print Software: Solution for Additive Manufacturing. [online] Available at: https://grabcad.com/print [Accessed 24 March 2026].

[8] OASIS Open. 2019. MQTT Version 5.0 Standard. [online] Available at: https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.html [Accessed 26 March 2026].

[9] Eclipse Foundation. 2022. Sparkplug 3.0.0 Specification. [online] Available at: https://sparkplug.eclipse.org/specification/version/3.0/documents/sparkplug-specification-3.0.0.pdf [Accessed 5 April 2026].

[10] HiveMQ. 2024. Understanding MQTT Topic Namespace: Sparkplug Edge Nodes Layout. [online] Available at: https://www.hivemq.com/sassets/f/243938/960x540/37dd324b85/understanding-mqtt-topic-namespace-iiot-sparkplug-edge-nodes-layout.webp [Accessed 1 April 2026].

[11] MTConnect Institute. MTConnect Standard Downloads. [online] Available at: https://www.mtconnect.org/standard-download20181 [Accessed 6 April 2026].

[12] IETF. RFC 8428: Sensor Measurement Lists (SenML). [online] Available at: https://datatracker.ietf.org/doc/html/rfc8428 [Accessed 6 April 2026].

[13] LocalXpose. 2026. LocalXpose: Secure Tunneling & Localhost Tunnel Software. [online] Available at: https://localxpose.io/ [Accessed 18 June 2026].

[14] Kimball Group. Dimensional Modeling Techniques. [online] Available at: https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/ [Accessed 18 June 2026].

[15] Google. 2026. Looker Studio. [online] Available at: https://lookerstudio.google.com/ [Accessed 18 June 2026].

[16] Grafana Labs. 2026. Grafana: The Open and Composable Observability Platform. [online] Available at: https://grafana.com/ [Accessed 18 June 2026].

[17] Microsoft. 2026. Power BI – Data Visualization. [online] Available at: https://www.microsoft.com/en-us/power-platform/products/power-bi [Accessed 18 June 2026].

A ANNEXOS

A.1 Esquema en Estrella (Star Schema) de la Capa Gold

La figura 8 mostra el disseny complet de l'esquema en estrella de la base de dades analítica Gold, amb les taules de dimensió i fets interconnectades. Al centre hi ha la taula de fets *JobSummary*, que concentra la informació de cada treball de fabricació (impressora, material, trolley, estat final). Al seu voltant, les taules de dimensió (*Dimension_Printer*, *Dimension_Material*, *Dimension_Trolley*, *Dimension_NCU*) aporten el detall descriptiu dels actius de la planta, mentre que les taules de fets addicionals (*Fact_Telemetry*, *Fact_Alerts*, *Fact_Maintenance*, *Fact_Postprocessing*, *Fact_Parts_Consumption* i *Fact_Heartbeat*) registren els esdeveniments concrets que tenen lloc al llarg del cicle de vida de cada peça.

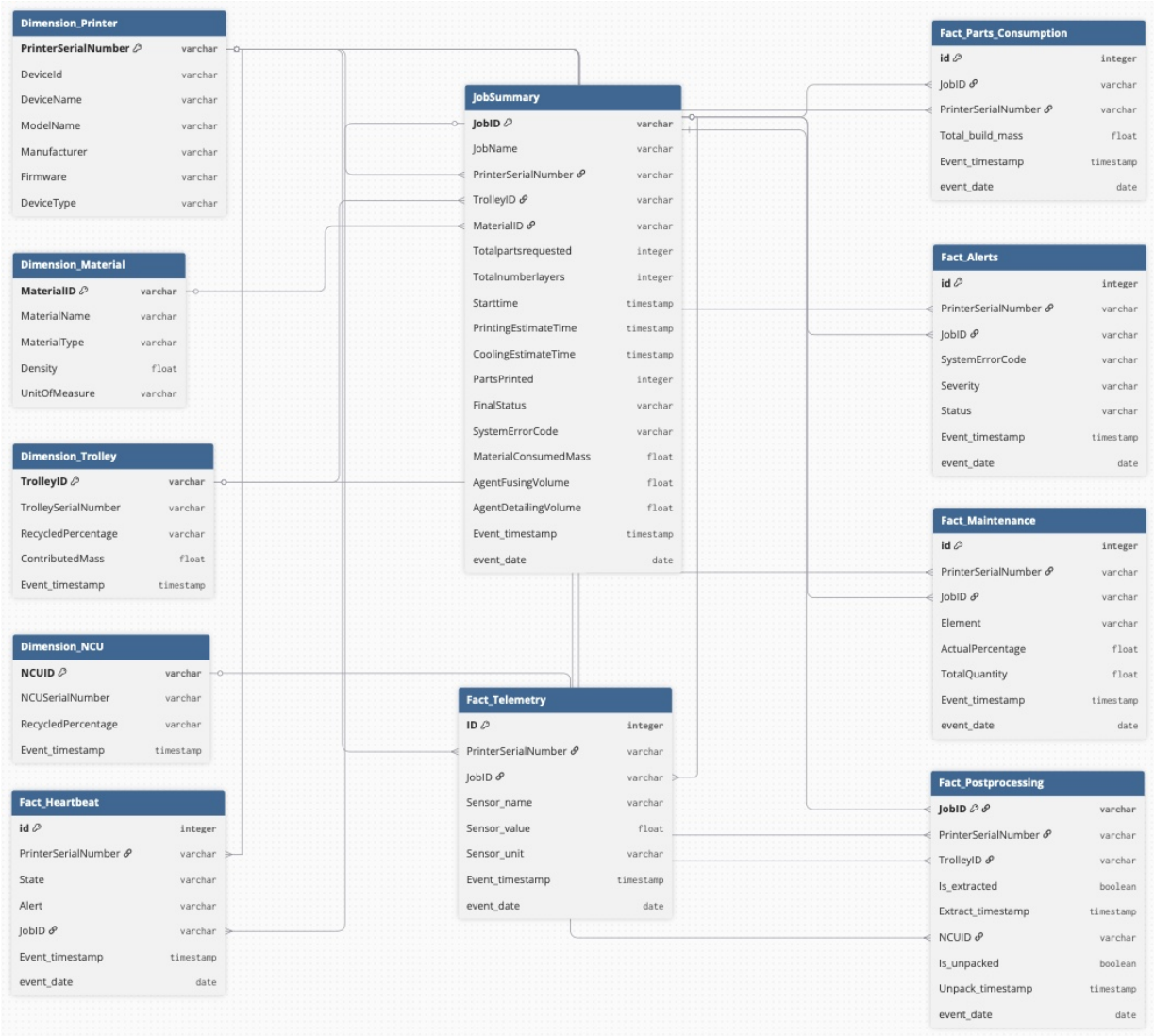


Figura 8: Disseny de l'esquema en estrella (Star Schema) de la base de dades analítica Gold, amb les taules de dimensió i fets interconnectades.

A.2 Inspecció de la Telemetria amb MQTT Explorer

La figura 9 mostra la visualització en arbre de la jerarquia de tòpics i la inspecció del payload JSON de telemetria emès per una impressora, tal com es mostra a través del client MQTT Explorer [2]. A l'esquerra es pot observar l'estructura jeràrquica completa dels tòpics, organitzada per identificador de dispositiu i submòduls funcionals (*DeviceInformation*, *BuildManager*, *Consumables*, etc.), mentre que al panell de la dreta es mostra el contingut del missatge seleccionat, amb el seu *timestamp* i un historial dels valors anteriors, que permet comparar canvis entre missatges consecutius.

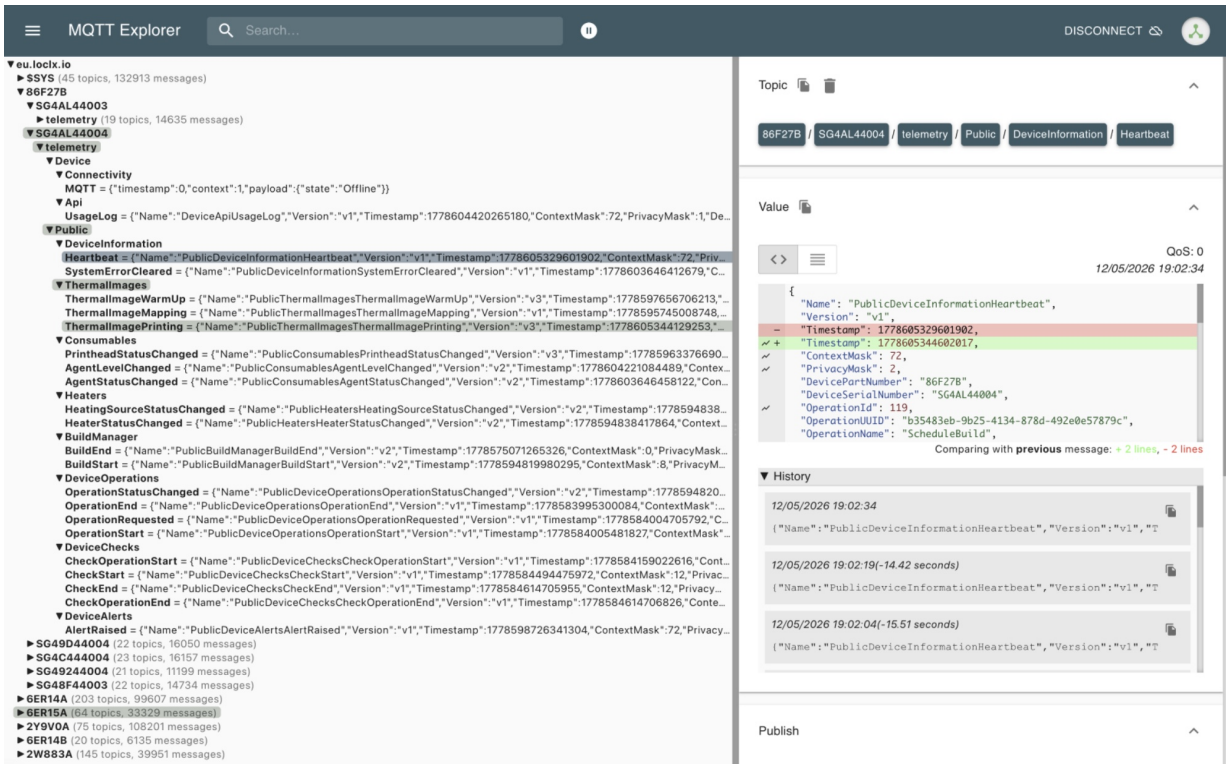


Figura 9: Visualització en arbre de la jerarquia de tòpics i inspecció del payload JSON de telemetria emès per una impressora a través de MQTT Explorer.

A.3 Topics MQTT Seleccionats: Path Jeràrquic Complet

La Taula A.3 detalla el path jeràrquic complet dels deu topics crítics seleccionats, organitzats pels tres mòduls funcionals descrits a la secció 3.2.2. Cada path segueix una estructura comuna (`+/+/telemetry/Public/...`), on el símbol `+` actua com a comodí MQTT d'un sol nivell, permetent que el mateix tòpic s'aplique de manera genèrica a qualsevol impressora de la planta sense necessitat de definir una subscripció diferent per a cada dispositiu.

Mòdul	**Esdeveniment**	**Path complet**
Control d'Estat i Sistemes	Heartbeat	`+/+/telemetry/Public/DeviceInformation/Heartbeat`
	SystemErrorRaised	`+/+/telemetry/Public/DeviceInformation/SystemErrorRaised`
	SystemErrorCleared	`+/+/telemetry/Public/DeviceInformation/SystemErrorCleared`
Producció	BuildStart	`+/+/telemetry/Public/BuildManager/BuildStart`
	BuildEnd	`+/+/telemetry/Public/BuildManager/BuildEnd`
	BuildReport	`+/+/telemetry/Public/BuildManager/BuildReport`
	BuildUnit/StatusChanged	`+/+/telemetry/Public/BuildUnit/StatusChanged`
Logística i Manteniment	AgentLevelChanged	`+/+/telemetry/Public/Consumables/AgentLevelChanged`
	WiperStatusChanged	`+/+/telemetry/Public/Consumables/WiperStatusChanged`
	FilterStatusChanged	`+/+/telemetry/Public/Consumables/FilterStatusChanged`

Taula 2: Path jeràrquic complet dels deu topics MQTT seleccionats per a la integració amb Something Added.

A.4 Limitacions de la Validació

La validació presentada en aquest treball està subjecta a certes limitacions d'abast, pròpies del temps i recursos disponibles en un projecte de tipus TFG.

En primer lloc, les xifres quantitatives reportades es basen en finestres temporals relativament curtes: 14 dies per a l'anàlisi del volum de *Fact_Heartbeat*, 30 dies per a l'estudi de durada de *Fact_Alerts*, i una mostra de 482 esdeveniments per al filtratge d'agents químics. Si bé aquestes mostres són suficients per demostrar la validesa del disseny, una validació a més llarg termini (per exemple, sobre un trimestre complet) permetria confirmar l'estabilitat d'aquestes mètriques davant de variacions estacionals de producció.

En segon lloc, el cas pràctic s'ha implementat íntegrament sobre impressores HP, que segueixen l'estàndard Sparkplug B. La capa Bronze ha estat dissenyada de manera agnòstica al fabricant per construcció –preservant el payload original sense parsing previ, independentment de la seva estructura interna– de manera que la incorporació d'un nou fabricant no requeriria redissenyar l'esquema de dades al núvol, si bé caldria reconfigurar Telegraf per subscriure's als seus tòpics propis i adaptar o crear noves Cloud Run Functions amb la lògica de filtratge corresponent a la seva telemetria. Tot i això, aquesta capacitat d'interoperabilitat no s'ha pogut validar empíricament amb maquinari real d'altres proveïdors dins l'abast d'aquest treball, i es proposa com a línia de treball futur confirmar-ho en un entorn multifabricant real.