# **Chapitre 2 : le modèle publish/subscribe**

1. **Introduction** :

Aujourd’hui de nombreuses applications utilisent des modèles fortement orientés contenu et plus complexes que le modèle Web (client-serveur) pour répondre à des exigences de communication souhaité. En effet, les applications nécessitant ces modèles sont de plus en plus utilisées dans l’Internet d’aujourd’hui et sont par exemple celles de l’Internet des Objets (IoT) telles que les systèmes de gestion des bâtiments. Ou encore des nouvelles applications de management des réseaux informatiques d’entreprises. Un des modèles utilisant des communications orientées contenu est le publication/abonnement (publish/subscribe, ou Pub/Sub). Il est déjà largement utilisé dans les réseaux et applications d'entreprise, principalement en raison de son évolutivité et de la prise en charge de la topologie de réseau dynamique.

Cependant, le modèle de communication publication/abonnement offre une topologie de réseau dynamique, une communication orienté contenu et prennent en compte les contraintes (fréquences de connexions, nombres de système destinateur, criticités des ressources) ce qui le permet d’être utilisables par les applications telles que Internet des Objets (IoT), etc.

Dans ce chapitre nous décrirons ce modèle Pub/Sub, son importance pour les systèmes et applications distribués et quelques protocoles qui utilisent.

**Définition et description** :

1. **Définition** :

Pub/Sub est un modèle de communication asynchrone de service à service utilisée dans les architectures sans serveur central [11]. Avec ce modèle on sous-entend un éditeur qui envoie un message et un abonné qui reçoit le message via un courtier de message.

1. **Description** :

La communication Pub/Sub n’a pas besoin de connaître l’adresse source ou de destination. Les sources et les destinateurs sont appelés respectivement les producteurs(éditeurs) et les consommateurs (abonnés) de messages.

Les producteurs communiquent avec les abonnés de manière asynchrone à travers un courtier en diffusant des catégories de messages sans tenir compte de la façon dont ces messages seront traités. Pub/Sub diffuse ensuite des événements à tous les services qui y sont intéressés. Cette manière d’intégration augmente la flexibilité et la stabilité du système dans son ensemble car on n’a pas besoin des appels de procédures ni attendre que les abonnés reçoivent les données [12]. C’est un modèle dynamique et découplé (les expéditeurs de messages et les destinataires peuvent agir à des intervalles de temps différents) d’où son adaptation aux applications distribués à grande échelle et peut être implémenté comme pattern de mémoire partagée [13]. Ce paradigme est fonctionnel et flexible en application distribué avec la communication centrée sur le contenu.

Subscriber

Publisher

**Figure 2.1** : présentation de Pub/Sub

La communication Pub/Sub met en relation trois entités dont le producteur (*Producer*), le consommateur (Consumer) et le courtier (*broker*).

1. **Les producteurs** :

Récoltent les informations et les publient au courtier, au lieu de destiner les messages à un consommateur il les associe à une catégorie de message.

1. **Les consommateurs** :

Un consommateur envoie une suscription événementielle à une catégorie de message et toutes informations qui répondent aux contraintes de cet abonnement sont transmises par le courtier et plusieurs consommateurs peuvent s’intéresser aux mêmes informations.

1. **Le courtier** :

Il est l’entité qui assure la livraison des messages aux consommateurs au moment opportun. Il peut posséder une petite mémoire de stockage temporaire des messages. Un exemple de courtier utilisant le protocole Pub/Sub est AMQP est RabbitMQ [14].

1. **Les catégories de Pub/Sub** :

Pub/Sub est généralement divisé en deux catégories : à base de sujet et à base de contenu. Il existe une troisième catégorie basée sur le type d’événement, il est considéré comme étant une propriété dynamique des événements avec des variantes sécurités [15].

1. **Pub/Sub centré sur le sujet** :

Dans cette catégorie les éditeurs sont tenus d'étiqueter (identifiés par mots-clés) chaque événement avec un nom de sujet ; les abonnés s'abonnent aux événements sous un sujet particulier. Chaque sujet est considéré comme un service événementiel propre, identifié par un nom unique et chaque événement appartient à l’un d’un ensemble fixe de sujet [15]. Après l’abonnement à un sujet, le consommateur serait notifié de tout événement des producteurs de ce sujet ou thème. L’abonnement à un sujet T peut être consulté en tant que membre d'un groupe T, et la publication d'un événement sur le sujet T se traduit donc par une diffusion d’événement parmi les membres de T [15]. C’est un modèle utilisé par Cloud Pub/Sub [16].

1. **Pub/Sub centré sur le contenu** :

Dans les systèmes basés sur le contenu, chaque abonné définit une souscription selon la structure interne des événements et chaque producteur de données attribue un nom unique et sémantiquement significatif à chaque paquet de données qu'il génère. C’est un système utilisé dans le routage statique car même si les intérêts changes dynamiquement la technique de routage reste inchangée. Les consommateurs s'abonnent à des événements sélectifs en spécifiant des filtres d’abonnement. Ces filtres définissent des contraintes, généralement sous la forme de paires nom-valeur de propriétés et d'opérateurs de comparaison de base (=, <, …), qui identifient [15]. NDN-PS (Named Data Network Pub/Sub) est un protocole qui utilise cette manière de distribution de données [17].

1. **Pourquoi Pub/Sub** :

Pub/Sub est un modèle centré courtier dans lequel il n’y a pas de saturation de plage d’adressage. Il est adapté aux systèmes distribués et aux systèmes ayant en communication plusieurs milliers d’autres appareils fixes ou mobiles comme Internet of Things (IoT). C’est une architecture pour un réseau qui ne cesse de se développer et qui accueille chaque jour de nouvel utilisateur qui échange des contenus continuellement, car elle offre à ses consommateurs la possibilité d’abonnement aux données permettant aux applications grand public (consommateurs) de recevoir efficacement des mises à jour [17]. Il est extensible, à faible consommation d’énergie et est sans serveur central ce qui assure la disponibilité des données et la tolérance aux pannes des équipements.

1. **Les protocoles qui utilisent Pub/Sub** :

Les protocoles les plus populaires utilisant le modèle de communication publication/abonnement des messages sont :

1. **Le protocole MQTT** :

MQTT (Message Queuing Telemetry Transport) [18] est un protocole léger, ouvert et simple pour transport de messagerie de publication/abonnement et Client/server. Il est conçu pour être facile à mettre en œuvre et est utilisable dans des environnements contraints tels que l’internet des objets. Il est basé sur TCP/IP, idéal pour la transmission des données utilisant une faible bande passante et est à faible consommation énergétique.

La principale caractéristique de MQTT est sa légèreté, ce protocole ne requiert que des ressources minimales et peut donc être utilisé sur de petits microcontrôleurs.

Son objectif est de faire dialoguer des équipements qui ne disposent que peu de ressources pour assurer une connexion permanente. Pour Fabien Pereira Vaz (gestionnaire technique de vente de Paessler AG), MQTT se démarque par sa souplesse et sa simplicité de mise en œuvre, en plus d'assurer une transmission de données bidirectionnelle. Le protocole prend ainsi de l'importance dans les technologies opérationnelles de l'industrie et c’est l’un des protocoles le plus utilisé dans l’internet des objets. Il utilise un courtier pour le relais des messages. Son processus se divise en quatre étapes distinctes : connexion, authentification, communication, terminaison et les accords définit pour garantir la livraison sont : plus d’une fois, moins d’une fois et exactement une fois.

1. **NDN-PS** :

NDN-PS (Named Data Network Pub/Sub) [17] est une nouvelle architecture qui fournit une communication centrée sur les données au niveau de la couche réseau. NDN implémente un modèle de communication demande-réponse asynchrone qui dissocie naturellement les producteurs et les consommateurs de données. Il définit deux types de paquets de couche réseau : intérêt et donnée. Chaque producteur de données attribue un nom unique et sémantiquement significatif à chaque paquet de données qu'il génère. Chaque consommateur émet un paquet d'intérêt avec un nom de données ou un préfixe de nom, qui est transmis en fonction du nom (préfixe). Pour chaque intérêt reçu, les producteurs NDN utilisent des stratégies de transfert pour décider où transmettre l'intérêt en tenant compte des politiques d'utilisation de la table de transfert et de la mesure des décisions de transfert précédentes.

1. **Pub/Sub sur Google Cloud** :

C’est un ensemble d’applications implémentées comme un ensemble de service connecté par la messagerie asynchrone. Le producteur crée le message et l’envoie pour un topic laquelle est appelé « ressource », le message est stocké jusqu’à l’acquittement de tous les consommateurs. Pour retirer ce message, une application consommatrice doit créer un abonnement pour ce topic, et Pub/Sub Cloud livre le message à tous ce qu’y sont abonnés. L’acquittement des messages permet la non livraison des messages plus d’une fois. Cette communication peut s’effectuée entre un producteur et plusieurs consommateurs, plusieurs producteurs et un seul consommateur et plusieurs producteurs et plusieurs consommateurs. Une application productrice peut-être n’importe quelle application capable d’envoyer une requête http aux APIs du Google. Quant aux applications consommatrices, elles doivent être non seulement capables d’envoyer une requête https, mais aussi accepter la requête POST de https [16]. C’est un modèle idéal pour un environnement où il y a beaucoup d’échanges de messages.

1. **Le protocole AMQP** :

AMQP (Advanced Message Queuing Protocol) [19] est un protocole réseau ouvert (open source) pour la messagerie. Il agit au niveau de la couche application autorisant un échange de messages entre deux parties avec utilisation d’une file d’attente (Queuing). Il est standard et peut être utilisé pour la mise en place d’une topologie d’échange de type publication/abonnement. L’objectif principale de ce protocole est de mettre en place un serveur de messagerie orienté message en utilisant la fille d’attente et l’interopérabilité pour permettre une production/consommation sans un problème de langage de programmation voire de système d’exploitation. Il est efficace dans l’interconnexion des objets hétérogènes. AMQP est plus adapté aux situations exigeantes : une faible connexion, la fiabilité de la communication et une importante sécurité [20]. Il est utilisé dans activeMQ (broker open source populaire multi-protocole dont AMQP) par la fondation Apache pour fournir son implémentation dans plusieurs langages tels que : C, C++, java... et d’autre plateforme comme IBM (MQLight) [20].

1. **Pub/Sub et Client/serveur** :

Pub/Sub est une architecture basée sur un composant central courtier (peut être un serveur). C’est un modèle dans lequel un consommateur n’a pas besoin de connaître les producteurs de donnée et un producteur n’a pas non plus besoin de connaître un consommateur de donnée.

Le modèle client/serveur est un modèle de communication d’internet à base de requête/réponse. Un client demande des informations ou des services et le serveur répond par les données ou services demandés. Le serveur doit connaître l’adresse du client pour pouvoir communiquer avec lui et inversement.

Pub/Sub

consumer

producer

requête

serveur

client

consumer

réponse

consumer

producer

**Figure 2.2** : comparaison Pub/Sub et client/serveur

**Quelques briefs description de cette comparaison** :

|  |  |
| --- | --- |
| **Pub/Sub** | **Client/server** |
| Centré sur le contenu | Centré sur l’adresse IP |
| Routage basé sur le contenu | Routage basé sur IP |
| Élimination du DNS | Impossible de fonctionner sans DNS |
| Données distribuées | Partages de données |
| Pas de serveur central pour les informations | Un serveur central est utilisé |
| QoS plusieurs vers plusieurs | QoS un client et un serveur |
| Communication asynchrone | Communication synchrone |
| Les deux entités peuvent débuter la communication. | Client débute la communication |

1. **Protocole à implémenter AMQP** :

Parmi les protocoles cités nous implémentons AMQP car c’est un protocole qui assure l’échange des messages chiffrés et assure l’interopérabilité entre les applications. Il utilise l’architecture publication/abonnement utilisé dans les systèmes de messageries ainsi que pour la gestion des périphériques IoT**.**

Une communication AMQP est un transfert de message pair-à-pair entre les nœuds d’un réseau. Elle est constituée d’un ensemble de conteneurs qui peuvent contenir plusieurs nœuds. Un conteneur peut être un broker ou une application cliente. Un nœud peut être un producteur ou un consommateur responsable d’un stockage ou d’une livraison de message.

Une communication entre deux conteneurs nécessite un établissement de connexion. Les messages envoyés lors d’une connexion sont fragmentés en un ou plusieurs trames. Une connexion est alors constituée de l’envoie d’une séquence de trame ordonnée car la réception de la dernière trame implique que les premières trames sont déjà reçues.

1. **Architecture AMQP** :
2. **Les entités de communication** :
3. **Nœuds** : les nœuds sont les entités responsables de production, de stockage, de livraison ou de consommation des messages. Il peut être un producteur, un consommateur ou une queue.

* **Source ou producteur** : est un nœud qui crée et envoie les messages.
* **Cible ou consommateur** : est un nœud qui récupère et traite les messages.
* **File d’attente (queue)** : est la destination ciblée pour le stockage des messages. Elle peut être plusieurs, chacune ayant un identifiant unique qui la différencie des autres et est réalisée à l’aide d’une mémoire. Elle permet de distribuer les messages aux différents consommateurs. Un message est stocké dans une file d’attente à condition que les règles de comparaison entre l’identifiant de la file d’attente et celui du message soient respecté.

Broker

Queue1

**Exchange**

Queue2

Queue3

**Figure 2.3** : broker, le gestionnaire des queues

1. **Conteneurs** : est un ensemble de nœud ou une application client avec ces producteurs et/ou ces consommateurs, ou un broker avec ces queues. Le placement des nœuds dans un conteneur peut être plat ou hiérarchique [20] :

**Topologie hiérarchique des nœuds**

**Topologie plate des nœuds**

b1

a

Node path /a

Conteneur

Conteneur

Node path a

a

b2

b

/

b

b3

c

Node path c

b4

c

Node path /b/b4

d

d

**Figure 2.4** : différentes topologies des nœuds dans un conteneurs

1. **Unité de communication AMQP** :

Une trame est l’unité de base échangée de AMQP et chaque trame est composée de trois parties distinctes :

Header

Body

Extended header

**Figure 2.5** : paquet AMQP

* Frame header (8 bits) : il est de taille fixe et contient la version protocole AMQP utilisé.
* Extended header (taille variable) : Le traitement de cette zone dépend du type de trame car il contient des données supplémentaires.
* Frame body (taille variable) : c’est le corps de la trame, il contient les données à transmettre et peut avoir différentes formes telles que : *open/close* (la trame d’établissement ou de fin de connexion), *begin/end* (établissement ou finalisation d’une session), *attach/dettach* (concerne les liens), *transfert* (pour envoyer un message), *disposition* (utilisé pour informer le récepteur des changements d’état de livraison).

1. **Connexion entre pairs** :

Tout connexion AMQP commence par les échanges TCP, ensuite chaque pair doit envoyer un paquet *open* qui indique ses limites et ses capacités pour établir les deux Chaines de connexion unidirectionnelle appelé chacun un lien AMQP. Cet échange TCP concerne les informations sur la version du protocole et de la sécurité. Ce lien peut être par exemple producteur-queue-consommateur. Ces deux liens de sens opposé constituent une connexion bidirectionnelle avec le control de flow d’où une session.

Session

conteneur

conteneur

Sessi sessionon

lien

lien

**Figure 2.6** : Session AMQP

La session est éphémère c’est-à-dire qu’après la déconnexion les pairs doivent refaire le même processus [19].

Pour la fermeture de la connexion les deux points de communication doivent envoyer un paquet *close* avec un code de fermeture de session. Pour cela chaque pair doit attendre avec un délai raisonnable pour la réception de ce paquet pour éviter toute erreur car c’est le dernier paquet envoyé.

1. **Contrôle de flow** :

Le point de terminaison de session attribue à chaque trame sortante un identifiant de transfert implicite à partir d'une séquence. Chaque point de terminaison de session maintient l'état suivant pour gérer les trames de transfert entrantes et sortantes. Le contrôle de flow détermine l’identifiant du prochain paquet entrant attendu et sortant, le nombre de paquet entrant et sortant tout en respectant la taille de la fenêtre (la taille maximale de paquet recevable par le consommateur) ce qui facilite la communication avec une faible bande passante. Cette connexion étant l’envoie d’une séquence ordonnée de trame, les deux entités vérifient le numéro de séquence car le contrôle de flow est basé sur le nombre de paquet échangé. Les deux points de communication incrémentent l’identifiant du paquet précédemment reçu comme l’identifiant du prochain paquet à envoyer et peut aussi décrémente la taille de la fenêtre.

Il exige aussi le contrôle de flow des liens car il possède un état qui indique si le lien est disponible pour l’envoie des messages et pour le mettre à jour l’envoyeur doit le modifier pour le receveur [19].

1. **Etat de la connexion et de la réception des messages** :
2. **Etat de la connexion** :

Toute connexion AMQP commence par les échanges TCP,

Il s’agit des étapes principales indiquées par les paquets de comment la communication se passe entre les pairs :

* Open/Close : cette étape implique l’envoie des paquets d’initiation ou de fermeture de la connexion ce qui veut dire que la connexion est soit ouverte ou fermée.
* Begin/End : pour l’établissement ou la fermeture de la session les pairs s’envoi ces paquets. A ce stade la session est ouverte ou fermée et la session fait partie de la connexion TCP établie.
* Attach/Detach : indique l’attachement ou le détachement d’un lien a une session.
* Send/Receive : indique l’envoi des paquets et leurs réceptions sur les deux liens établis.

1. **Etat de la réception des messages** :

Les messages reçus peuvent être de différente forme tel que :

* *Accepted* : le message est bien reçu et est traité par le consommateur
* *Rejected* : les messages invalides que les consommateurs ne peuvent pas traiter donc ils sont rejetés.
* *Released* : les messages invalides qui doivent être re-envoyé pour qu’ils puissent être traités.
* *Modified* : les messages non traités car ils ont été modifiés.
* *Received* : cet état indique à l’éditeur de re-envoyer le message mais il indique à la cible que le message ne sera pas re-envoyer [19].

1. [**La qualité de services (QoS)**](https://www.ibm.com/docs/en/ibm-mq/9.2?topic=reliability-message-delivery-amqp#q123795___qos) :

AMPQ est efficace, portable, multiport et réputé par sa fiabilité de livraison [21]. Il fonctionne bien dans les environnements multi client et obéit à des règles strictes, garantissant l'interopérabilité des clients de différents fournisseurs. Ce protocole offre trois manières de communication producteur/consommateur grâce au principe de réglage automatique (*setled*) de ces entités de communication. Ces trois manières sont :

* ***at-most-once*** : un seul envoi avec la possibilité qu'il soit manqué ; si un producteur choisit ce réglage il supprime sa copie du message avant que les files en acquitte. De même si un consommateur choisit cette option la file n’attend pas l’acquittement de leur part avant de le jeter.

consommateur

producteur

Transfert (delivery\_tag,setled, …)

**Figure 2.7**: QoS “at-most-once”

* ***at-least-once*** : livraison garantie avec la possibilité de duplication de message, contrairement au premier le producteur et la file d’attente attend respectivement l’accusé de réception de la file d’attente et le consommateur.

consommateur

producteur

Transfert (delivery\_tagsetled , …)

ACK (setled, …)

**Figure 2.8**: QoS “at-least-once”

* ***exactly-once*** : livraison unique garantie. Car avant l’envoie du message le producteur et le consommateur s’envoient quelques messages d’acquittement [21].

1. **Fonctionnement** :

AMQP est un protocole de communication entre les applications à travers un réseau. Il permet une communication asynchrone car l’émetteur et le consommateur n’ont pas à agir au même rythme et le consommateur n’est pas non plus obligé de traiter l’information et d’en accuser la réception. Il s’appuie sur le broker pour la distribution des messages aux différentes destinations. Les messages étant stockés dans une file d’attente, le producteur peut ainsi travailler sans avoir de période d’inactivité.

Les messages sont distribués dans les différentes files d’attentes en utilisant les identifiant de ces files et le *binding* des messages.

Broker

Queue

Echange

Consommateur

Producteur

Queue

Queue

**Figure 2.9** : éléments intermédiaires de AMQP

**Exemple d’application** :

Cet exemple met en place un producteur, un consommateur et un serveur RabbitMQ avec le langage Java en localhost.

**RabbitMQ** : est un server (ici broker = gestionnaire de queues) localhost de simulation du protocole AMQP qui écoute par défaut sur le port 15672 pour les versions 3.x.

Publie Consomme

**Producteur**

**Consommateur**

**RabbitMQ**

Vérifie

**Figure 2.10** : élément intermédiaire de RabbitMQ

Apres avoir lancé le programme producteur (**Prod**), il produit un message et ce message sera stocké dans la queue définie et visible dans **RabbitMQ**. Le programme consommateur (**Cons**) va ensuite vérifier la présence d’un message dans la même queue et le consomme (affiche le message trouvé) s’il y en a.

**NB** : Une exécution continue de **Prod** produit plusieurs messages qui seront stockés dans la même queue.

1. **Transmission des messages aux consommateurs** :

Après la production des messages, chaque message est transmis au broker par les producteurs et le broker à son tour livre les messages aux consommateurs. Le broker AMQP est constitué de l’*exchange* et des queues. Le transfert de message entre *l’exchange* et les queues s’effectue par la correspondance entre la *routing\_key* du message et le *binding\_key* des files d’attente.

Le protocole AMQP comporte quatre types *d’exchange*, ces quatre types deroutages sont :

* Direct exchange ou unicast : le message est livré à une seule file d’attente si la *routing\_key* correspond au *binding\_key* donc un seul destinateur. C’est une communication one-to-one.
* Fanout exchange ou broadcast : c’est une communication one-to-many car ce type d’exchange ne fournit pas de *routing\_key* donc le message est livré à toutes les files d’attentes.
* Topic exchange ou multicast : c’est comme direct exchange sauf que la comparaison entre les *keys* ne concerne qu’une partie des keys. C’est possible d’avoir une communication *many-to-many*.
* Header exchange : utilise l’en-tête des messages à la place de la routing\_key. Les informations que contient cet entête seront comparées au *binding\_key*.

Par défaut, c’est-à-dire default exchange, le *binding\_key* d’une queue est utilisé comme le *routing\_key* ce qui stockage directement les messages dans cette queue.

Si un ou plusieurs consommateurs s’abonnent à une queue pour un type de message. Le message leurs serait livré une fois son stockage dans la file d’attente.

1. **Inconvénient :**

Ce protocole n’est pas compatible à ses versions précédentes car actuellement il n’existe que deux versions qui indépendantes aux niveaux de broker, *binding* et *exchange* [22].

1. **Sécurité**:

AMQP est un protocole d’encodage binaire qui utilise un protocole de sécurisation SSL/TLS et de l’authentification SASL. Chacun de ces deux protocoles consiste en quelque échange de frame header suivies des négociations de TLS ou de SASL. Cette négociation est un échange de certificat de sécurité pour les échanges sur réseau informatique.

SSL/TLS permet de crypter les informations sur les différentes chaînes pour assurer la confidentialité des données [19].

SASL est un mécanisme d’authentification qui vérifie les différents accès aux informations et peut être par exemple (nom d’utilisateur et mot de passe).

1. **Conclusion** :

Les services de communication de nouvelle génération doivent pouvoir coopérer pour répondre à des besoins spécifiques tout en gardant leur autonomie. Les réseaux sont mis en place notamment dans le but de transférer des données d’un système à un autre ou de fournir des ressources partagées des serveurs, des bases de données ou d’une imprimante. Ceci nécessite de maîtriser leurs architectures.

Pub/Sub est une norme qui préconise comment les équipements devraient communiquer entre eux tout en permettant aux systèmes ouverts capable de respecter son architecture d’être apte à échanger des informations avec d'autres équipements hétérogènes et qui sont issus de constructeurs différents. Les protocoles de cette architecture fourni une connexion logique qui est à la base de ces échanges entre ces systèmes informatiques indépendants afin de répondre aux contraintes des applications et systèmes actuelles.

Dans le prochain chapitre, nous présenterons une spécification pour les APIs utilisant le modèle Pub/Sub. Plus précisément le protocole AMQP pour la génération des codes clients pouvant implémenter ces APIs.