

Calame Software S.A.S

Gathering Tools

Abaques de dimensionnement



Version 2020

Versions

Date	Description
05/03/2020	Portage



Table des matières

1. Présentation générale.....	4
2. Eléments de base pour le dimensionnement.....	4
2.1. Taille du document.....	4
2.2. Entités, documents envoyés et réponses correctrices.....	5
2.3. Restitutions.....	5
2.4. Nombre de connexions simultanées à GTServer	6
2.5. Dimensionnement d'un projet de taille moyenne	6
3. Base GT	7
3.1. Logs de polling et d'exécution de la base GT	8
4. Base client.....	10
5. Serveur GT	11
5.1. Précisions sur le fonctionnement d'un service GTServer vis-à-vis de ses clients	11
5.2. Système 64 bits	12
5.3. Processeurs/cœurs spécifiquement dédiés à un service d'une instance GT	12
5.4. Mémoire utilisée par GT.....	13
5.4.1. Tableau simplifié pour un projet de dimension « moyenne »	15
5.5. Espace disque spécifiquement dédié à GT	15
5.6. Bande passante utilisée par GT	19
5.6.1. Tableau des débits.....	20

1. Présentation générale

Les informations présentées dans ce document ne sont fournies qu'à titre indicatif afin de pouvoir réaliser des estimations en début de projet, alors même que toutes les contraintes techniques et fonctionnelles du projet n'ont pas encore été envisagées ou exprimées.

2. Eléments de base pour le dimensionnement

Les informations suivantes seront utilisées de manière récurrente pour le dimensionnement :

- Nombre de modèles de documents
- Taille de document embarquant les données pré remplies ou synchronisées ainsi que les données de réponses des correspondants.
- Nombre de campagnes à conserver dans GT (campagnes ouvertes ou fermées, non supprimées)
- Nombre d'entités adressées
- Nombre estimé de documents publiés pour chaque entité
- Nombre estimé de réponses fournies pour chaque entité (réponses correctrices)
- Nombre de gestionnaires / valideurs
- Utilisation ou non des fonctionnalités GTWeb : transfert de la réponse, synchronisation, publication d'url de document
- Nombre de connexions et demandes d'informations simultanées à GTServer depuis GTAnswer ou depuis le module d'automatisation.

Les paragraphes suivants détaillent certains indicateurs de la liste précédente.

2.1. Taille du document

La taille du document est un élément central pour la détermination des autres éléments de dimensionnement.

Idéalement, une taille moyenne et une taille maximale de document devraient être déterminées.

Si des documents GT existent déjà embarquant une volumétrie moyenne et une volumétrie maximale (données de pré-remplissage + données de réponse), il est préférable d'utiliser cette information pour chacune des campagnes.

Dans le cas où un document embarquant des données n'existe pas, on peut estimer la taille de celui-ci de deux façons différentes à partir d'un classeur Excel embarquant une

volumétrie représentative de données de pré-remplissage et de données de réponses (volumétries moyenne et maximale) :

- Construire un document en important tous les onglets Excel utilisés. Ce document ne sera pas opérationnel (i.e. ne permettra pas d'envoyer de réponse réelle) mais peut fournir une estimation assez juste de la taille du document avec les données.
- La taille du classeur Excel peut également être utilisée pour donner une idée plus approximative de la taille du document (généralement le document a une taille un peu moins élevée).

Le classeur Excel contenant les valeurs calculées aura généralement une taille plus grande que le document et sera une majoration des données au format texte.

Dans les deux cas, **le classeur Excel** utilisé pour la création rapide d'un document dans Design ou directement pour estimer la taille, devra avoir la **forme et la structure du document final** : formules utilisées dans le document final, présentation dans chaque onglet en respectant la forme des structures finales des compartiments (multi-onglets et motif, séparation des données de référence et des données de saisie avec formules dans les données de saisie pointant vers les références, etc.) La taille du document sera modifiée par l'organisation et l'articulation des compartiments entre eux.

2.2. Entités, documents envoyés et réponses correctrices

Le nombre d'entités doit être un décompte des entités pour lesquelles un document est publié mais également, le cas échéant, des « sous-entités » correspondant à des réponses reçues (axes de validation supplémentaires et réponses portant sur des domaines différents pour une même entité initiale).

Le nombre de documents publié pour chaque entité peut être largement supérieur à 1 : il correspond en effet à l'ensemble des cas de nouvelles publications de documents suite à des corrections ou des mises à jours des données, avec ou sans utilisation de la synchronisation

Le nombre de réponses et de réponses « correctrices » pour une entité est le nombre de documents de réponses reçues pour une entité. Ce nombre sera généralement supérieur strictement à 1 pour chaque entité et va dépendre du « dialogue de données » pendant la campagne (validation des données que cela soit via GTAnswer ou intégré au document, demande de réponse avec des corrections, etc...)

2.3. Restitutions

Il est nécessaire de détailler le nombre de modèles différents de restitutions, la taille de chacun des classeurs Excel produits, par un modèle de document, (avec les données qu'ils devraient embarquer), ainsi que le nombre d'entité pour lesquelles cette restitution doit être publiée.

2.4. Nombre de connexions simultanées à GTServer

Le nombre de connexions simultanées à GTServer comprend les connexions des deux sources suivantes :

- GTAnswer
- Automatisation

Pour simplification, une connexion via l'automatisation sera comptée comme une connexion GTAnswer : les opérations réalisées par cette connexion sont les mêmes qu'une connexion GTAnswer. Néanmoins, suivant la fréquence d'exécution des opérations d'automatisation, il sera peut-être opportun de compter pour l'automatisation un client (de type GTAnswer) supplémentaire connecté en permanence à GTserver.

Ne devraient être dénombrées que les connexions de ces deux sources **effectuant simultanément une demande à GTServer**. Seules comptent réellement les demandes de liste de réponses, synchronisation, récupération d'un document via GTAnswer, transfert de réponse http., soit toutes les opérations clientes.

On notera NbAnswerCSimult le nombre estimé d'utilisateurs GTAnswer connectés simultanément.

Etant donné que la plupart des requêtes sont l'objet d'une connexion effective de durée très limitée, les connexions simultanées effectives sont en nombre très limité (à condition que la communication de GTServer avec la base de données GT soit rapide : taux de transfert rapide et bonne disponibilité du SGBD).

On estime généralement que moins de 5% des répondants GTAnswer sont connectés en même temps (avec une borne inférieure à 1 utilisateur Answer connecté) et que moins de 50% des utilisateurs/gestionnaires de GTAnswer sont connectés en même temps pour un même projet (avec une borne inférieure à 1 utilisateur GTAnswer).

Cependant ce nombre d'utilisateurs connectés en même temps dépend de la répartition des entités entre les gestionnaires de suivi (si, par exemple, chacun suit les entités de sa région), des calendriers entre les projets (deux projets distincts avec des gestionnaires analysant leurs projets respectifs en même temps) et au sein des projets sur la même instance GT et des pics d'affluence (par exemple, si tous les répondants GTAnswer cherchent à répondre au même moment, souvent le dernier). Cette estimation doit être réévaluée en fonction des projets sur une même instance GTServer.

2.5. Dimensionnement d'un projet de taille moyenne

Le dimensionnement peut être difficile à ajuster pour quelques caractéristiques.



Gathering Tools est une marque commerciale de Calame Software.
Pour ce manuel et toute la documentation afférente © 2003 – 2020 Calame Software S.A.S

Nous proposons dans ce document des dimensionnements exemples pour des projets de taille « moyenne », soit :

- Moins de 500 destinataires
- Des documents de moins de 300Ko
- Des transpositions ne faisant pas intervenir plus de 25 colonnes transposées pour les motifs les plus volumineux.
- Pas plus de 3 réponses correctrices pour chaque entité (soit 3 réponses correctrices + 1 réponse finale)

3. Base GT

La base GT contient l'ensemble des objets GT : modèles de documents et de restitutions, actions, campagnes lancées, modèles de mails et journaux d'exécution et de polling.

La taille des données accueillies par la base GT est essentiellement déterminée par la taille des documents (stockés sous forme de BLOB dans la base GT) auxquels s'ajoutent les logs d'exécution et de polling stockés en base. Le chiffre d'estimation fourni correspond à la taille des BLOB stockés dans la base GT et non à la taille de la base de données, la taille globale de la base pourra varier si les BLOB sont stockés de manière spécifique (cryptage, etc..).

Les documents envoyés ou les réponses reçues seront généralement les éléments les plus volumineux stockés dans la base GT (en dehors des logs de polling et d'exécution).

	Minimum	Recommandé
Taille des documents de campagnes accueillies dans la base GT	[Nb de campagnes à conserver] * [Taille du document prérempli et avec des réponses] * [Nb d'entités adressées et reçues, si axes de validation supplémentaires] * ([Nb de documents envoyés par entité] + [Nb de réponses reçues par entité])	[Nb de campagnes à conserver] * [Taille du document prérempli et avec des réponses] * [Nb d'entités adressées et reçues, si axes de validation supplémentaires] * ([Nb de documents envoyés par entité] + [Nb de réponses reçues par entité])
Taille indicative des logs de pollings et d'exécution (A RESTREINDRE PAR DES OPERATIONS SUR LA BASE GT, cf § 3.1)	512 Mo	1 Go
Taille totale des données accueillies dans la base GT	Somme des éléments ci-dessus	Somme des éléments ci-dessus

3.1. Logs de polling et d'exécution de la base GT

Les logs de polling et d'exécution inscrits dans la base ne contiennent que des informations textuelles.

Chacun des logs de polling ou d'exécution est inscrit sous la forme d'un BLOB (ou équivalent) dans la base de données GT.

Les logs de polling peuvent atteindre une taille importante surtout dans le cas de réception des réponses par mail. Cette taille dépend du nombre de pollings effectués, du nombre de réponses aux documents à intégrer par GTServer et de l'encombrement de la boîte mail dédiée à GTServer : i.e. lorsque des mails reçus dans la boîte mail de GTServer ne sont pas destinés à GTServer. Ces mails « parasites » peuvent « polluer » considérablement les logs de polling.

La transmission de réponses par mail doit être considérée comme une « connexion de secours » et doit être strictement limitée aux utilisateurs :

- **Qui ne sont pas connus d'avance / pour lesquels il n'est pas possible de créer un compte**
- **Qui, pour des raisons techniques, ne peuvent se connecter au serveur web**

Le nettoyage régulier de la boîte mail dédiée de GTServer est indispensable.

La taille des logs de polling doit être restreinte en réduisant le nombre de jours conservés dans les pollings (via GTAdmin, Onglet Messagerie).

Les logs de polling peuvent être nettoyés via l'interface GTAnswer (bouton « Historique du polling »).

Les logs de polling dans le cas de réception par Web ont des tailles nettement moins importantes (pas de mails « parasites », moins d'opérations effectuées donc loggées, etc..).

Les logs d'exécution inscrits dans la base dépendent du nombre d'exécutions d'actions (lancement, intégration, restitutions, actions SQL), du nombre de documents envoyés par action, du niveau de détail inscrit dans les journaux (si toutes les phrases SQL vers la base client sont inscrites ou pas, etc.), et enfin d'une quantité variable d'informations nécessaires pour décrire chaque exécution d'actions (par exemple, un lancement de campagne renseignera les informations de pré remplissage qui dépendront de la conception du document et de la base de données mais qui peuvent être nombreuses).

Les logs d'exécution peuvent être nettoyés directement via l'interface GTAnswer (Menu « Historique ») **ou par suppression des publications** (les publications peuvent être supprimées manuellement ou par automatisation).

Généralement, on considère qu'il n'est pas utile de dépasser 500 Mo pour l'ensemble des logs de polling et/ou d'exécution.

Dans certains cas d'utilisation comme des campagnes permanentes (réponses récurrentes à un même document pour plusieurs périodes), l'exécution de scripts d'automatisation récurrents (scripts de polling, de lancement ou d'intégration, ...), les logs d'exécution seront plus nombreux. Estimer alors à 1 Go, voire 2, la taille maximale des logs avant nettoyage sera plus prudent.

4. Base client

La base Client utilisé par GT est une base ouverte.

Le dimensionnement de cette base dépend essentiellement des données destinées à GT et des données hors GT qui vont être inscrites en base, du niveau de normalisation pour cette base, de l'indexation etc.... et le volume réel des données dépendra in fine du type de base de données choisie pour héberger la base Client (SQL Server, Oracle, Postgres, DB2).

De nombreuses données dans cette base pourront être stockées afin d'être utilisées avec d'autres outils que GT. La modélisation complète de la base client et l'évaluation de sa volumétrie dépasse le cadre d'exploitation par la seule suite GT ainsi que le cadre de ce document.

Enfin, la réduction des journaux de transactions de la base et la réduction de la base elle-même sont nécessaires pour maîtriser la taille globale de celle-ci. Cette taille pourra notamment augmenter rapidement si de nombreuses actions d'intégration sont exécutées. Les opérations réalisées sur la base en-dehors des modules GT auront également une influence sur la taille globale de la base.

Pour rappel, les opérations de réduction de la base de données, de réduction des journaux de transactions et de sauvegardes de la base ne sont pas des opérations réalisées par les modules GT. Elles seront généralement dévolues au service administrant le serveur de base de données et/ou les bases de données.

A titre indicatif, on peut fournir une **estimation très approximative** du volume des données qui ont transité dans les documents GT. Ce premier chiffre grossier peut donner une estimation des données qui seront incluses dans la base. La modélisation des données, le niveau de normalisation de la base, les objets techniques telles que index, tables tampons, etc... ne sont pas pris en compte dans cette estimation très grossière.

L'estimation correcte de la volumétrie de **la base de données** doit être réalisée par **le concepteur du projet et** en suivant **les recommandations de dimensionnement de l'éditeur du SGBD** (SQL Server, Oracle, Postgres ou DB2).

Une estimation générale n'est pas fournie. Seule une estimation indicative est donnée pour un projet de taille « moyenne » (cf 2.5)

Volume approximatif des données pour un projet de taille « moyenne » (cf 2.5)		
	Minimum	Recommandé
Estimation indicative et très approximative des données accueillies dans la base Client	512 Mo	1,5 Go

5. Serveur GT

5.1. Précisions sur le fonctionnement d'un service GTServer vis-à-vis de ses clients

Lors d'une communication réseau avec un client (GTAnswer ou l'automatisation), deux situations peuvent se produire :

- La commande demandée par le client consiste essentiellement à transmettre une information venant de la base de données GT au client ou une information sur l'avancement de l'exécution d'une action. Le service GTServer lit dans la très grande majorité des cas des informations venant de la base GT, les communique au client réseau puis arrête la communication réseau. L'exception notable consiste en l'obtention de la liste des publications (et dans une moindre mesure de la liste des réponses) où GTServer doit effectuer un travail de réorganisation des réponses obtenues.
- La commande demandée par le client consiste à exécuter l'exécution d'une action : action de publication, action d'intégration ou action SQL. La demande d'exécution est enregistrée et placée en attente puis GTServer arrête la communication réseau. Une seule action à la fois peut être exécutée par GTServer. Les autres demandes d'exécutions sont placées en attente.

Les connexions réseau entre GTServer et un client sont donc ponctuelles dans tous les cas et, dans le cadre des opérations courantes pour un projet (i.e. hors tâches d'automatisation), deux connexions successives sont suffisantes pour obtenir toutes les informations dont le client a besoin pour une tâche élémentaire dans GTAnswer.

Deux opérations client-serveur synchrones (i.e. lorsque le client est en attente du serveur) peuvent être un peu plus lourdes que les autres :



- Une demande de synchronisation, à partir de GTAnswer, portant sur une forte volumétrie de donnée
- Une demande d'obtention de la liste des publications, à partir de GTAnswer, dans le cas où existent de très nombreuses campagnes et/ou réponses.

Dans le premier cas, l'essentiel du travail est réalisé par la base de données pour transmettre toutes ces données à GTServer puis GTServer à GTWeb et GTAnswer.

Dans le second cas, GTServer doit faire un travail de ré-ordonnancement des réponses pour produire les statistiques nécessaires.

A noter que quasiment toutes les transactions entre GTServer et un client résulteront en un accès à la base de données GT. La rapidité de communication avec le serveur hébergeant la base GT, la disponibilité et l'efficacité du SGBD hébergeant la base GT sont donc essentielles pour assurer la fluidité des transactions entre GTServer et un client (GTAnswer ou client d'automatisation)

5.2. Système 64 bits

Un système Windows 64 bits est préféré.

5.3. Processeurs/cœurs spécifiquement dédiés à un service d'une instance GT

Pour une même instance GTServer, une seule action à la fois peut être exécutée (action publication, action d'intégration, action SQL ou polling).

Idéalement, au moins 1 cœur devrait être dédié à la seule exécution des actions. Celles-ci sont placées en attente. Une seule action est exécutée à la fois pour une même instance.

On notera NbAnswerCSimult le nombre estimé d'utilisateurs répondants connectés simultanément et NbGTClisimult, le nombre d'utilisateurs gestionnaires / valideurs connectés simultanément (cf § 2.4 pour les estimations).

Réajuster les estimations pour les cas où des synchronisations GTAnswer sont transmises par GTWeb ou lorsque des périodes de pointe peuvent survenir (en fin de campagne par

exemple, avec exécution d'intégrations en plus des opérations de suivi de campagne et de réponse par des correspondants).

Les machines physiques sont bien sûr préférées aux machines virtuelles équivalentes, pour des raisons de performance.

Même si le service GTServer peut très bien fonctionner sur une machine de type bureautique, les processeurs les plus récents sont à privilégier.

Minimum Recommandé		
	Minimum	Recommandé
Nb de cœurs disponibles pour les exécutions d'actions	1	2
Nb de cœurs disponibles pour les connexions simultanées Answer et GTAnswer (dont automatisation)	$\text{Min}([\text{NbAnswerCSimult}], [\text{NbGTcliCSimult}])$	$[\text{NbAnswerCSimult}] + [\text{NbGTcliCSimult}]$
Nb de cœurs disponibles pour un service GT	Somme des éléments précédents	Somme des éléments précédents

5.4. Mémoire utilisée par GT

La mémoire utilisée par un service GT atteindra un pic lors de l'exécution d'une action d'intégration (ou pour certains documents lors de l'exécution de l'action de publication).

Lors de l'exécution d'une action de publication, le pic de mémoire sera atteint au préremplissage d'un document (ce préremplissage est réalisé par un programme externe à GTServer). Cette taille peut être estimée par la taille mémoire maximum utilisée par

GTAnswer.exe lorsque le document correspondant à la taille maximale est chargé. Le document exemple pour estimer la taille du document (cf 2.1) peut être utilisé en le chargeant dans GTAnswer (toutes les données et formules dans leur forme définitive doivent être présentes) ; il est cependant plus précis d'utiliser l'outil d'estimation de la volumétrie de Design à partir d'un document ayant la structure et les formules du document final.

Lors de l'exécution d'une action d'intégration, la mémoire nécessaire va dépendre de la taille des données pour une campagne (taille des données embarquées dans le document, représentation en mémoire de ces données, nombre d'entités, nombre de réponses reçues pour chaque entité).

Lors des connexions clients, la mémoire nécessaire pour le service GTServer croît très faiblement avec le nombre d'utilisateurs (cf § 5.1). En comparaison avec une exécution d'action, lors de connexions avec un client, les informations transmises via les communications réseau sont beaucoup plus réduites et sont présentes en mémoire dans un temps beaucoup plus restreint que lors du pic atteint lors de l'exécution de l'action d'intégration ou lors de l'exécution de l'action de lancement.

	Minimum	Recommandé
Mémoire disponible pour une publication	Mémoire maximum utilisée par GTAnswer lorsque le document a été chargé (cf Renvoi 1) + mémoire utilisée par GTserver (cf Renvoi 2)	Mémoire maximum utilisée par GTAnswer lorsque le document a été chargé (cf Renvoi 1) + mémoire utilisée par GTServer (cf Renvoi 2)
Mémoire disponible pour une action d'intégration	Cf Renvoi 3	Cf Renvoi 3
Mémoire disponible pour un service GT et ses programmes annexes	128 Mo (mémoire de fonctionnement) + Min([NbAnswerCSimult], [NbGTClCSimult]) * 3 Mo + Maximum des tailles ci-dessus	256 Mo (mémoire de fonctionnement) + ([NbAnswerCSimult] + [NbGTClCSimult]) * 3 Mo + Maximum des tailles ci-dessus

Renvoi 1 : La mémoire maximum utilisée par GTAnswer peut être mesurée en chargeant le document dans GTAnswer et en mesurant la mémoire utilisée lors du chargement (le document peut être simulé à partir d'un classeur Excel représentatif). Une autre solution est d'utiliser l'outil « Estimation de la volumétrie à partir de GTAnswer » de Design à partir d'un document ayant les structures (motifs, multi-onglets) du document final.

Renvoi 2 : la mémoire maximum utilisée par GTServer lors d'une publication sera généralement plus réduite que la mémoire utilisée par GTAnswer sauf dans le cas où des transpositions sont utilisées. GTServer charge les données issues de la base donc non transposées, les écrit sur le disque puis les transpose, ce qui réduit leur taille sur disque (d'autant plus que le nombre de colonnes transposées est grand) et réduit le chargement par Answer. Une approximation très grossière de la RAM nécessaire pour le module de transposition (majoration du besoin mémoire de GTServer) est la suivante : pour chaque motif ayant des données transposées, on considère les données directement issues de la base (donc non transposées), la mémoire nécessaire sera la mémoire maximum de l'ensemble des estimations. Pour un motif avec des données transposées, une estimation de la taille mémoire nécessaire est la suivante : $[Taille\ Mémoire\ Transpo] \approx (13/3 * [Nombre\ de\ lignes\ dans\ BD,\ pour\ une\ entité] * [Taille\ représentation\ textuelle\ d'une\ ligne\ au\ format\ CSV,\ soit\ [NbChamps] + \Sigma(taille\ des\ champs\ texte\ réellement\ utilisés) + \Sigma(taille\ des\ champs\ numériques\ représentés\ textuellement) + \Sigma(taille\ des\ champs\ dates\ représentés\ textuellement) + \Sigma(taille\ des\ champs\ dates\ représentés\ textuellement)])$

Renvoi 3 : La mémoire utilisée par GTServer ne peut être déterminée sans entrer dans les détails de typage de chacun des champs et ce, pour chacun des motifs. De plus, le framework .NET utilisé par GTServer peut mettre en œuvre l'internalisation des chaînes (donc consommer moins de mémoire). L'expression suivante a été établie sur des statistiques d'intégration de documents-types, cette expression est approximative : $TailleMemoireGTServerIntegrationEtFonctionnement \approx (256\ Mo + 0,8 * [Nb\ Réponses,\ toutes\ les\ réponses\ y\ compris\ réponses\ correctrices] * [Taille\ Document\ moyen\ en\ Mo])$.

5.4.1. Tableau simplifié pour un projet de dimension « moyenne »

Dans le cas d'un projet « usuel » (cf 2.5).

	Minimum	Recommandé
Mémoire disponible pour GTServer et les programmes affiliés	512 Mo	1 Go

5.5. Espace disque spécifiquement dédié à GT

L'espace disque nécessaire atteindra un pic lors des publications et lors de l'intégration.

Lors de l'exécution de ces actions, des fichiers temporaires sont créés pour générer les documents.

Si les documents embarquent des pièces jointes, la volumétrie de ces pièces jointes doit être décomptée, comme éléments de pré-remplissage au lancement du document ou comme éléments de réponse par les correspondants GTAnswer. Dans le cas d'intégrations multiples sur une même campagne, il existe autant de copies du même document qu'il y a eu d'exécutions de l'action d'intégration. Ces versions des documents correspondants à d'anciennes exécutions d'intégration doivent être supprimées par le(s) concepteur(s) du projet. Les actions de restitutions dans Calame peuvent également écrire des fichiers sur disque. Ces éléments doivent être décomptés dans la volumétrie finale. Néanmoins, l'emplacement de stockage pourra se situer en-dehors d'un disque local du serveur GTServer.

Au vu de la volumétrie présentée dans les classeurs Excel exemple, prévoir 10 à 20 Go d'espace disque pour les données temporaires d'un service GTServer devrait être suffisant.

	Minimum	Recommandé
Espace disque de fonctionnement	5 Go	10 Go
Espace disque disponible temporairement pour les exécutions d'actions	Cf Espace disque pour les exécutions d'action	Cf Espace disque pour les exécutions d'action
Espace disque permanent pour les pièces jointes extraites des documents	[Taille totale des pièces jointes] = Σ (pour les entités de réponse, toutes campagnes) de [Taille totale des pièces jointes pour une réponse] * [Nombre exécutions action intégration]	[Taille totale des pièces jointes] = Σ (pour les entités de réponse, toutes campagnes) de [Taille totale des pièces jointes pour une réponse] * [Nombre exécutions action intégration]
Espace disque permanent pour les enregistrements sur disque de restitution	[Nombre de fichiers de restitutions] * [Taille des fichiers des restitutions]	[Nombre de fichiers de restitutions] * [Taille des fichiers des restitutions]
Espace disque total pour l'instance GTServer	Somme des éléments ci-dessus	Somme des éléments ci-dessus

La taille du log de polling fichier doit être maîtrisée par la limitation de la conservation de l'historique de polling (via GTAdmin).

La journalisation fichier général de GTServer est un système de rotation de fichiers définis via GTAdmin par leur nombre et leur taille : la taille des logs fichiers généraux ne pourra pas dépasser la taille définie dans GTAdmin (nombre de fichiers * taille maximale d'un fichier).

Pour l'estimation de l'espace disque pour les exécutions d'actions de lancement de campagne, de restitutions ou d'intégration, on notera :

TMaxQstComplet : Taille max d'un document avec ses données pré remplies et ses réponses (+ images et pièces jointes)

TMaxQstHorsImagesPJ : Taille max d'un document avec ses données pré remplies et ses réponses mais hors images et pièces jointes

TMaxXlsRestitComplet : Taille max d'un classeur de restitution avec toutes ses données et tous ses objets (images, objets ole, etc...)

TMaxXlsRestitHorsImagesObjets : Taille max d'un classeur de restitution avec toutes ses données mais sans les objets OLE, images, etc...

NbMaxEntitesEnvoi : Nombre max d'entités envoyées conjointement

Espace disque pour les exécutions d'action		
	Minimum	Recommandé
Espace disque de fonctionnement pour l'exécution d'action	1 Go	5 Go
Espace disque disponible pour l'exécution d'une publicatioj	[NbMaxEntitesEnvoi] * [TMaxQstComplet] + 10 * [NbMaxEntitesEnvoi] * [TMaxQstHorsImagesPJ]	[NbMaxEntitesEnvoi] * [TMaxQstComplet] + 20 * [NbMaxEntitesEnvoi] * [TMaxQstHorsImagesPJ]
Espace disque disponible pour l'exécution d'une action d'intégration	[NbMaxEntitesEnvoi] * [TMaxQstComplet] + 2 * 10 * [NbMaxEntitesEnvoi] * [TMaxQstHorsImagesPJ]	[NbMaxEntitesEnvoi] * [TMaxQstComplet] + 2 * 20 * [NbMaxEntitesEnvoi] * [TMaxQstHorsImagesPJ]
Espace disque pour les exécutions d'action	Espace disque de fonctionnement pour l'exécution d'action + Max des autres éléments ci-dessus	Espace disque de fonctionnement pour l'exécution d'action + Max des autres éléments ci-dessus

Commenté [OC1]: Pour la transposition 2 * les fichiers mais il faudrait tenir compte du nombre de colonnes transposées ou non TODO

Commenté [OC2]: Pour la transposition

Le nombre 10 ou 20 est utilisé pour donner une approximation très grossière, à partir de la taille du document, du volume des données embarquées dans le document, lorsque les données sont représentées sous la forme de fichiers texte (format utilisé par GTServer pour les données temporaires).

5.6. Bande passante utilisée par GT

Le volume des données transitant par le réseau atteindra un pic dans les situations suivantes :

- consultation des réponses à partir de GTAnswer (ou documents envoyés) par les gestionnaires pendant le suivi de campagne.
- consultation à partir de GTAnswer des fichiers temporaires de documents ou de restitutions lors de l'envoi de campagne ou lors de l'envoi de restitution.
- transmission des réponses par GTAnswer vers GtServer, via GTWeb
- synchronisation de données de GTServer vers GTAnswer, via GTWeb.

Les communications réseau avec GTServer se produisent de manière ponctuelle et pour échanger généralement des informations venant de la base de données GT et dans une moindre mesure de la base client.

Les deux goulets d'étranglement principaux sont la connexion entre le service GTServer et la base de données GT et, dans une moindre mesure, sauf au moment de l'exécution des actions de lancement/intégrations, la connexion du service GT avec la base Client.

Des débits élevés et des latences réduites sont à privilégier pour les connexions du service GT avec la base GT et avec la base Client (dans les deux sens de communication).

Les bandes passantes du tableau suivant présentent des préconisations de taux réels de transmission et dédiées au seul usage par les applications/modules Gathering Tools lors des pics de transmission.

Les opérations qui requièrent le plus fort volume de données transitant par le réseau sont marquées d'un astérisque dans le tableau qui suit.

On notera NbAnswerCSimult le nombre estimé d'utilisateurs Answer connectés simultanément et NbGTcliCSimult, le nombre d'utilisateurs GTAnswer connectés simultanément (cf § 2.4)

La latence devrait être inférieure à 100ms dans tous les cas.

5.6.1. Tableau des débits

Abréviations utilisées dans le tableau qui suit :

- [TdT] : temps souhaité de transmission de l'information en secondes. Ce temps peut varier suivant chaque opération.
- [TQstMax_kb] : Taille maximale de document (avec toutes les données pré-remplies ou synchronisées ainsi que les réponses) en kb (compter par précaution 10 fois la taille en Koctets).
- [TXLSRestMax_kb] : Taille maximale de restitution classeur Excel (avec toutes les données) en kb (compter par précaution 10 fois la taille en Koctets).

Flux concerné (émetteur -> destinataire)	Débit minimum en kb/s	Débit recommandé en kb/s
GTAnswer ou Automatisation → GTServer (pour un travail de gestionnaire)	$128 * [\text{NbGTcliCSimult}] / [\text{TdT}]$	$256 * \text{NbGTcliCSimult} / [\text{TdT}]$
GTServer → GTAnswer ou Automatisation (*)	$2 * \text{Max}([\text{TQstMax_kb}],$ [TXLSRestMax_kb]) . $[\text{NbGTcliCSimult}] / [\text{TdT}]$	$2 * \text{Max}([\text{TQstMax_kb}],$ [TXLSRestMax_kb]) . $[\text{NbGTcliCSimult}] / [\text{TdT}]$
GTWeb → GTServer (avec réponse Answer par GTWeb) (*)	$2 * [\text{TQstMax_kb}]$. $[\text{NbAnswerCSimult}] / [\text{TdT}]$	$2 * [\text{TQstMax_kb}]$. $[\text{NbAnswerCSimult}] / [\text{TdT}]$
GTServer → GTWeb (avec synchronisation des données vers GTAnswer) (*) Cf Renvoi 1	$2 * [\text{TQstMax_kb}]$. $[\text{NbAnswerCSimult}] / [\text{TdT}]$	$2 * [\text{TQstMax_kb}]$. $[\text{NbAnswerCSimult}] / [\text{TdT}]$
GTServer → GTWeb (sans synchronisation des données vers GTAnswer)	$128 * [\text{NbAnswerCSimult}] /$ [TdT]	$256 * [\text{NbAnswerCSimult}] /$ [TdT]

Flux concerné	Débit minimum en kb/s	Débit recommandé en kb/s
GTServer ↔ SGBD (base GT et base Client) (*)	20 Mb/s (download et upload)	>100 Mb/s (download et upload)
GTServer → Serveur de messagerie SMTP (*)	[nombre de restitutions ou document envoyés lors de l'exécution d'une même action] * 2 * Max([TQstMax_kb], [TXLSRestMax_kb]) / [TdT]	[nombre de restitutions ou document envoyés lors de l'exécution d'une même action] * 2 * Max([TQstMax_kb], [TXLSRestMax_kb]) / [TdT]
Serveur de messagerie SMTP → GTServer	1 Mb/s	5 Mb/s
GTServer → Serveur de messagerie POP/IMAP (si pas de réponse par GTWeb)	128 kb/s (si la boîte de mail GTServer est nettoyée des mails non destinés à GTServer)	256 kb/s (si la boîte de mail GTServer est nettoyée des mails non destinés à GTServer)
Serveur de messagerie POP/IMAP (si pas de réponse par GTWeb) → GTServer (*)	2 * [TQstMax_kb] / [TdT]	4 * [TQstMax_kb] / [TdT]

Flux concerné	Débit minimum en kb/s	Débit recommandé en kb/s
GTAnswer → GTWeb (avec réponse GTAnswer par GTWeb) (*)	$2 * [TQstMax_kb]$. $[NbAnswerCSimult] / [TdT]$	$2 * [TQstMax_kb]$. $[NbAnswerCSimult] / [TdT]$
GTAnswer → GTWeb (sans réponse GTAnswer par GTWeb)	$128 * [NbAnswerCSimult] / [TdT]$	$128 * [NbAnswerCSimult] / [TdT]$
GTWeb → GTAnswer (avec synchronisation des données vers GTAnswer) (*) Cf Renvoi 1	$2 * [TQstMax_kb]$. $[NbAnswerCSimult] / [TdT]$	$2 * [TQstMax_kb]$. $[NbAnswerCSimult] / [TdT]$
GTWeb → GTAnswer (sans synchronisation des données vers GTAnswer)	$128 * [NbAnswerCSimult] / [TdT]$	$128 * [NbAnswerCSimult] / [TdT]$

Note 1 : l'astérisque pointe les flux les plus volumineux.

Note 2 : le facteur « 2 » qui intervient dans de nombreux calculs permet d'exprimer une majoration du fait que les données sont encodées en base64.

Renvoi 1 : Concernant l'évaluation de la volumétrie des données transitant par les flux réseaux pour la synchronisation : les données subissent un processus similaire à celui des données se trouvant dans un document. De ce fait, la taille des données de synchronisation n'excédera pas la taille du document final (puisque celui-ci contient toutes les données pré-remplies et/ou synchronisées ainsi que les réponses).