Optimisation des comptages de Dimension

Description

Avant de décrire les méthodes possibles permettant l'accélération d'ouverture des Marques contenant des Comptages de Dimension (DimCounts), il est nécessaire d'en savoir plus sur le comptage des valeurs uniques.

Les comptages de Dimension seront relativement lents, comparés aux champs somme. C'est à cause de la façon dont travaillent les comptages de Dimensions.
Imaginez deux jeux d'enregistrements (A et B) et vous savez que dans le jeu A, il y a 23 valeurs uniques d'une Dimension. Dans l'ensemble B, il y a 15 valeurs uniques. Maintenant, vous voulez connaître le nombre de valeurs uniques de la réunion des jeux A et B.
Le nombre peut varier de 23 à 38. Vous devez garder une trace de toutes les valeurs uniques des deux ensembles et les combiner. Il s'agit d'une méthode exigeant beaucoup de volume et de temps.

Maintenant, au contraire, imaginez que les 2 jeux ont une colonne de champ somme et que le total de l'ensemble A est de 100 et le total de l'ensemble B est de 50. La seule chose que nous devons faire pour les combiner est d'additionner les valeurs : cela est rapide et facile. En fait, accumuler des nombres est l'une des instructions du microprocesseur les plus élémentaires (cela s'effectue dans le matériel).

Alors, qu'est-ce que cela signifie? Cela signifie que chaque fois que Diver doit grouper des jeux d'enregistrement, cela sera relativement lent et surtout avec grands ensembles de données.
Nous reviendrons sur ce point ultérieurement.

Comptages de Dimension précalculés

Les développeurs de Dimensional Insight savaient que les DimCounts seraient importants et qu'il fallait être plus rapide que la procédure habituelle, ils inventèrent donc les comptages de Dimension précalculés.
Cette option peut être activée pour chaque Dimension pour laquelle vous voulez calculer les valeurs uniques et de manière rapide.
En fait, cette option va effectuer des combinaisons de Dimensions, précalculer les comptages de Dimension et enregistrer le tout dans le Modèle : ainsi les comptages de Dimensions sont aussi rapides que les champs Somme !

Avec l'option de précalcul des comptages de Dimensions et l'utilisation de Marques simples, les temps de réponse sont grandement améliorés, mais ce n'est pas le cas avec des Marques complexes. La leçon à retenir sur les comptages de Dimensions est la suivante:

Chaque fois que Diver regroupe plusieurs Dimensions, il ne peut pas utiliser les valeurs précalculées DimCount!

Cela résulte de la façon dont sont calculés les DimCounts, cela a déjà été evoqué ci-dessus. Dimension A a un comptage de Dimension précalculé de 20, et la dimension B a un comptage de Dimension précalculé de 10.
En regroupant A et B, Diver ne peut pas utiliser les dimcounts précalculés. Il faudrait partir de l'ensemble des valeurs de Dimension pour les deux ensembles et faire un comptage unique.
Il n'est pas certain que Diver enregistre cette information dans le Modèle : il pourrait éventuellement prendre le jeu d'enregistrements en tant que point de départ.

Alors, que pouvez-vous faire pour vous assurer que votre Marque n'utilisera pas les DimCounts précalculés?

© Bimaxis - Tous droits réservés

Le fait d'être dans au moins l'un de ces trois cas :
1. Le Comptage de valeurs uniques de Dimension provient de plusieurs Modèles
2. Des groupes nommés combinent plusieurs valeurs de Dimension
3. Des champs infos sont promus en Dimensions dynamiques et ces Dimensions sont utilisées soit comme un filtre (par exemple, DimCount [column_a, MyDynDim = "DoNotDoThis"]), soit en groupe nommé, ou bien en tant que plongée ou bien comme QuickView.

Mais alors comment obtenir une Marque MultiTableau Croisé avec 12 QuickViews et 2 groupes nommés qui s'affiche rapidement?

La clé est de s'assurer que Diver utilise les valeurs précalculées.
Ainsi, d'une part, toutes les QuickViews et plongées doivent être basées sur des Dimensions fondamentales.
Les groupes nommés ne peuvent pas être utilisés, mais un nouveau groupe de Dimension fondamentale peut être créé.
Mais l'ajout de tout ce dont vous avez besoin en tant que Dimension fondamentale ne fonctionnera pas. Le Modèle sera trop grand: la configuration n'est pas extensible.

La meilleure solution est cependant de diviser les données en plusieurs Modèles et de s'assurer que tous les dimcounts sont dirigés vers 1 seul Modèle.
Plus vous pourrez mettre de Dimensions dans un Modèle Lookup, plus sera important le nombre de Dimensions que vous pourrez utiliser comme Dimension fondamentale sans faire les Modèles extrêmement volumineux.
Il y a plusieurs façons d'atteindre cet objectif.

Tout d'abord, il est possible de diviser le Modèle sur une QuickView. Par exemple, s'il y a une QuickView appelée «Année», il est logique de créer des Modèles par année.
Notez que cela rend également la configuration évolutive, à la fois au niveau de l'espace disque et de la performance : au fur et à mesure que les années passent, de nouveaux Modèles (relativement petits) sont créés. Les données pour chaque année sont dans un Modèle distinct et est rapide.
Vous pouvez diviser les modèles sur plusieurs QuickViews, mais vous avez juste à vous assurer que les colonnes dimcount se réfèrent toujours à 1 seul Modèle.
L'inconvénient de ceci est que vous ne pouvez pas utiliser « Toutes les valeurs », car cela n'est pas pré-calculé et Diver devra faire des regroupements sur tous les Modèles et calculer la valeur à la volée.
Une façon de tricher est de faire un Modèle de l'Année « Toutes les valeurs» et ajouter ce modèle au MultiModèle. Lors de la séparation sur 2 QuickViews et si vous voulez à la fois avoir « Toutes les valeurs », vous aurez besoin de créer le Modèle « Toutes les valeurs » pour chaque valeur de dimension de l'autre QuickView ...

Deuxièmement, la division sur les valeurs de filtre / groupes nommés. Imaginez une Marque qui possède 3 colonnes dimcount, 1 pour les jours ouvrables, 1 pour les samedis et 1 pour les dimanches.
Les définitions de colonnes sont : DimCount [Saleperson, Jour de la semaine = "journée de travail"], etc
Vous pouvez créer 3 Modèles : model_Workday.mdl, model_Saturday.mdl, model_Sunday.mdl. Les trois colonnes se référèrent à un seul Modèle et les dimcounts précalculées peuvent être utilisés.
La combinaison de plusieurs filtres est également possible : imaginez que vous voulez aussi avoir une Dimension «Holiday» avec les valeurs «vacances» et «non-vacances».
DimCount [Saleperson, Jour de la semaine = "journée de travail", Holiday = "vacances"] se réfère alors au modèle: model_Workday_holiday.mdl
Cependant, DimCount [Saleperson, Jour de la semaine = "journée de travail"] n'est plus rapide maintenant car il est nécessaire de grouper des dimcounts à partir de plusieurs Modèles.
Lorsque vous utilisez un Multitableau Croisé ou un Tableau Croisé, vous pouvez également splitter sur ces colonnes pour améliorer les temps de réponse.

© Bimaxis - Tous droits réservés

Pour avoir chaque colonne en tant que Dimension, vous pouvez utiliser les Modèles Lookup mais soyez prudent : il est facile de plonger dans une Dimension et d'obtenir une réponse lente, parce que vous avez oublié qu'il est maintenant nécessaire de regrouper des valeurs de plusieurs Modèles.

Cette méthode permet vraiment de créer une Marque DimCount spécifique qui sera incroyablement rapide plutôt que de laisser l'utilisateur faire une plongée libre dans le Modèle.

Tags

1. comptage Dimension
2. Diver
3. optimisation

© Bimaxis - Tous droits réservés