Serveur de traitement SIG

Province Sud

Vendredi 23 mars 2018 à 09h00

CAPS de la Province Sud - Salle 4a02

Auteur : Roman de Scorraille

Le constat

Fréquence d'utilisation d'un logiciel

- **Quotidiennement**
  L'agent utilise en permanence le logiciel pour son métier
- **De temps en temps**
  L'agent a un besoin très ponctuel d'un logiciel comme par exemple transformer une donnée dans un certain format

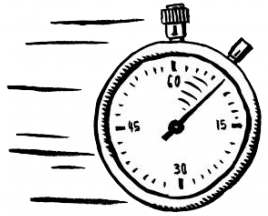
Besoin de performance

- **Quotidiennement**
  L'agent a besoin systématiquement des traitements lourds pour son métier
- **De temps en temps**
  L'agent a un besoin très ponctuel d'un traitement lourd comme par exemple transformer une donnée dans un certain format

Un serveur de traitement pour quoi faire ?

- Pouvoir accéder à une panoplie de logiciels sous licences ou pas

- Réaliser des traitements rapides grâce à une configuration adaptée



- Ne pas gêner l'ordinateur de l'agent pendant les traitements longs

- Diminuer le coût (matériel / logiciel)

La solution fonctionnelle

- Avoir un ordinateur personnel performant si c'est justifié (usage quotidien)
- Disposer d'une licence individuelle ou flottante si l'usage du logiciel est quotidien
- En cas de besoin ponctuel (logiciel , performance), l'agent a la possibilité de se connecter sur un ordinateur distant

Les contraintes techniques

- Accès distant rapide sur un réseau « lent »
- Pouvoir utiliser la 3D. Le bureau distant de Windows ne supportant pas la 3D
- Pouvoir accéder aux données sécurisés

La solution technique

Un serveur robuste avec un logiciel de streaming video (HP remote)

- Diffusion de la 3D
- Pas de RDP (License logiciel)

Une seule session utilisateur anonyme.
L'agent doit réserver la session sur le calendrier outlook (comme une ressource)

- Achat d'une seule licence par logiciel
- 100% des ressources du serveur affecté

Réalisation d'un script de montage réseau que l'agent exécutera avec son login/ mot de passe

- Accès aux données sécurisées

Caractéristiques du serveur de la province sud

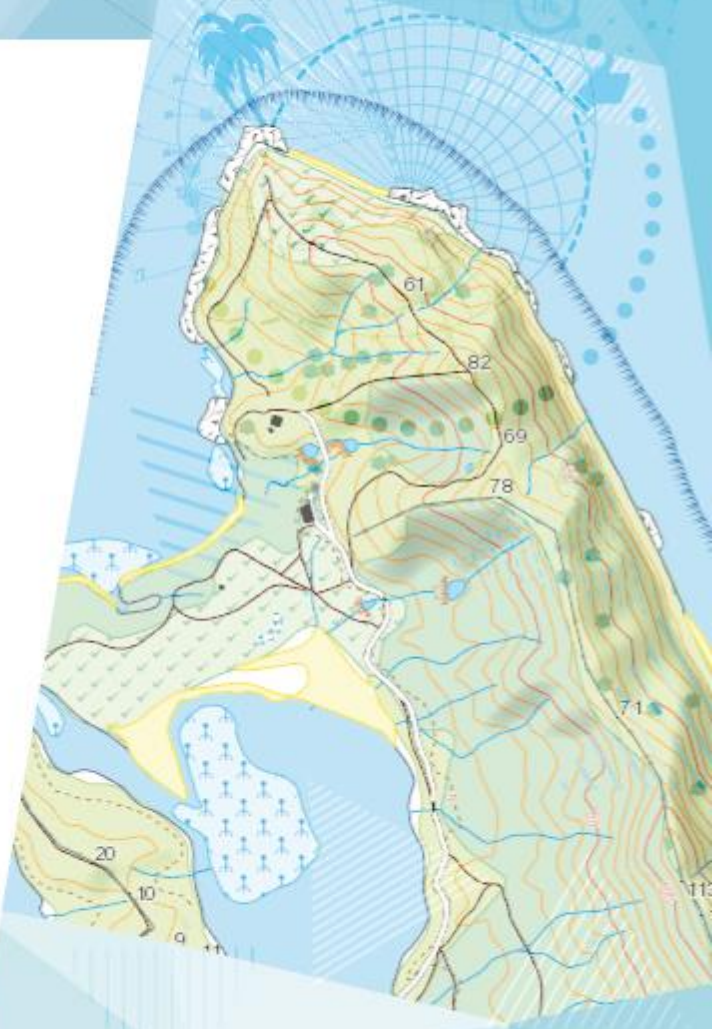
- 2 CPU XEON E5-2620 v4 @ 2.1 GHz
  - 16 cores
  - 32 threads
- 128 Go de ram
- 2 ssd
  - 512 Go moyenne performance
  - 1.4 To haute performance
    - Lecture: 2 Go/s
    - Ecriture 1 Go/s
- 1 carte graphique quaddro nvidia

Caractéristique d'un ordinateur individuel performant (poste du bureau SIG)

- 1 cpu i7-7700 @ 3.6 GHz
  - 4 cores
  - 8 threads
- 32 Go ram
- 1 ssd moyen de 512 Go
  1 disque magnétique de 4 To

Tests sur l'extraction du contour d'un raster (Algorithme traitement parallèle)

1. Pré-découpage du raster en tuiles
2. Calcul binaire du raster (masque noData)
3. Polygonisation du masque
4. Union et dissolution des polygones extraits
5. Suppression des trous
6. Union et dissolution des polygones des résultats
7. Génération d'un fichier shp



Tuile 1	Tuile ….	Tuile n
Phase 2,3,4,5	Phase 2,3,4,5	Phase 2,3,4,5
Phase 6,7		

Benchmark sur l'extraction du contour d'un raster (2016-Noumea-serail)

	1 cpu	4 cpu	8 cpu	16 cpu	32 cpu
E5 (windows)	2h43 X1				0h18 X8,2
I7 (windows)	1h50 X1.47	0h41 X4	0h25 X6.3		

Processeurs Intel: XEON (E) vs CORE (I)

	XEON série E	CORE série I
Avantage	• Multi-processeur • Mémoire (1.5To /cpu) • Nombre de core	• Rapidité par core • Prix
Inconvénient	• Core moins rapide • Prix	• Mono-processeur • Limité à 128Go de mémoire • Nombre de core plus faible

Processeurs Intel: XEON (E) vs CORE (i)

	E7-8894	**i9-7980XE**
Caractéristiques	3,4 GHz (turbo) 24 cores 48 threads	*4.2 GHz (turbo)* 18 cores 36 threads
Prix	10 000€	1800 €

R : idéal pour la parallélisation des tâches

```
library(doParallel)
library(foreach)


#on prépare 4 processus en parallèle
cl <- makeCluster(4)
registerDoParallel(cl)


#on prépare une liste de raster à traiter
listRaster <- list.files(path = ".", pattern = "*.tif",full.names = TRUE, recursive = TRUE)

#on itère la liste en distribuant sur les processus. Il est possble de récupérer l'index de la liste
foreach(rasterIn = listRaster,i=icount()) %dopar% {
    print(
        sprintf("traitement du raster n° %1$s (%2$s) ",i,rasterIn)
    )
    ....
}

#on arrête les processus
stopCluster(cl)
```

Conclusion

- Peu de logiciels SIG sont capables de paralléliser les traitements
- Il faut mieux regarder la performance brute d'un core
- Privilégier un processeur de type core série i qu'un xeon sauf s'il faut un grand besoin de mémoire vive
- Un ssd performant n'est pas nécessaire, les traitements des données seront toujours plus long que la capacité de lecture du ssd

Merci de votre attention