Mésocentre de calcul Université de Franche-Comté

Laurent PHILIPPE

8 décembre 2009

Présentation Utilisateurs

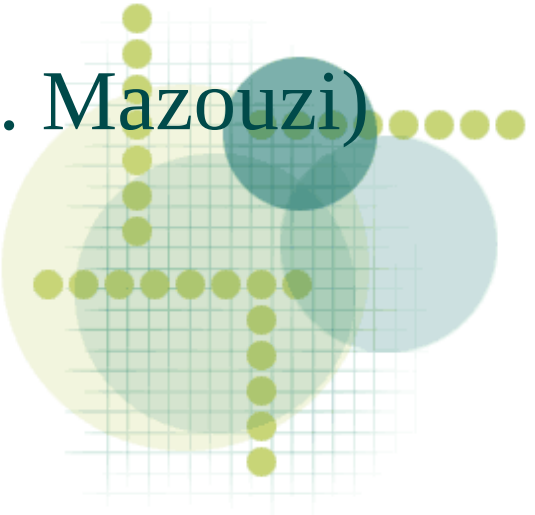


UNIVERSITÉ DE FRANCHE-COMTÉ

Présentation du mésocentre

Planning :

- 9h30 – 10h30 : le mésocentre de calcul, ressources et utilisation (L. Philippe)
- 10h30 – 11h : discussion
- 11h – 11h30 : visite cluster
- 13h30 – 12h30 : utilisation avancée (K. Mazouzi)



Plan

- Présentation mésocentre de calcul
- Ressources
- Accès
- Exécution
- Développement
- Informations générales



Plan

- **Présentation mésocentre de calcul**
- Ressources
- Accès
- Compilation
- Exécution
- Informations générales



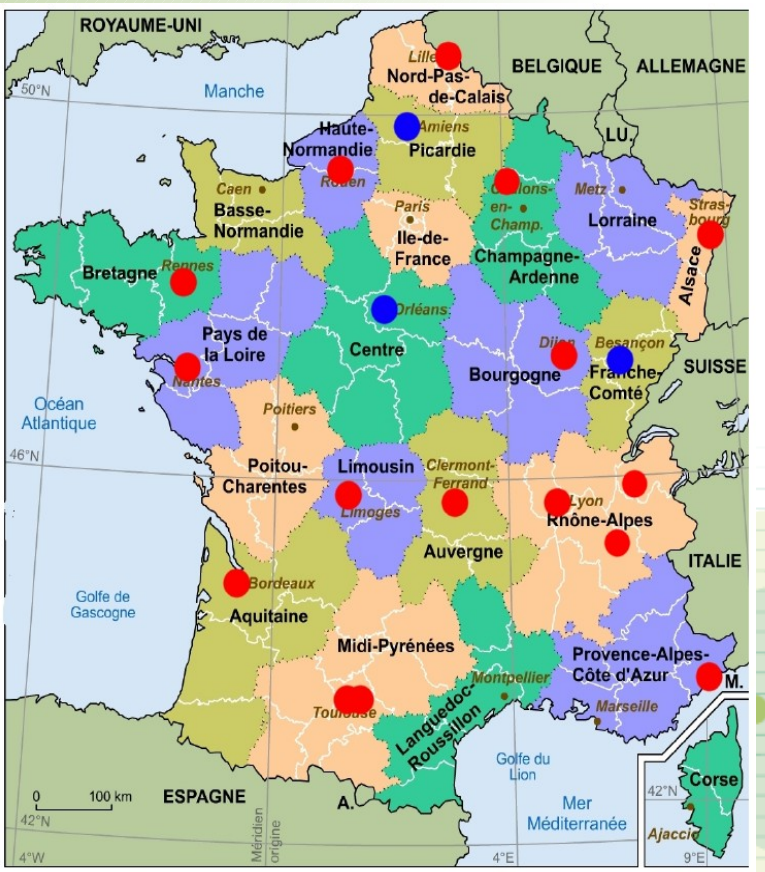
Pour quoi faire ?

- Augmenter la précision ou la capacité de traitement :
  - Modélisation
  - Simulation
  - Traitement de données
- Besoins HPC :
  - Ressources
  - Compétences



Les centres de calcul

- Europe : PRACE
- France : GENCI
  - IDRIS
  - C INES
- Régions



Objectifs

- Donner des moyens de calcul et/ou stockage
- Développer les compétences en HPC
- Lieu d'échange
- Proximité
- Mutualiser les ressources :
    - Gestion mésocentre
    - Accès prioritaire

UFC
UNIVERSITÉ DE FRANCHE-COMTÉ

utbm
université de technologie
Belfort-Montbéliard

ensmm

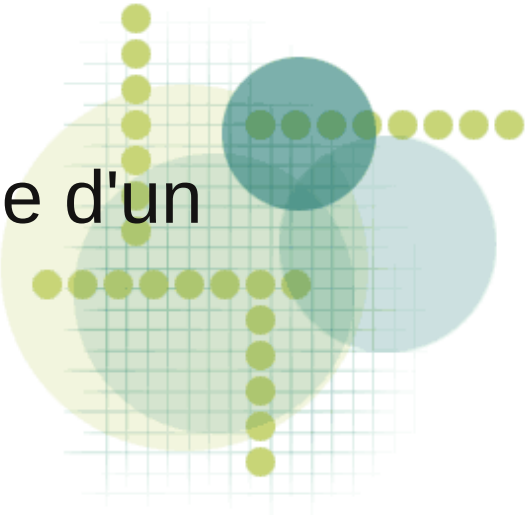
Franche-Comté
Conseil régional

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
MINISTÈRE
DE L'ENSEIGNEMENT SUPÉRIEUR
ET DE LA RECHERCHE

Pour qui ?

- TOUT LE MONDE:
  - Initiés
  - Débutants
- Chercheurs
- Enseignants : projets, TP
- Industriels

=> Besoin de puissance supérieure à celle d'un ordinateur personnel



Utilisations possibles

- Calcul :
  - Exécution à distance
  - Désynchronisée
  - Plusieurs instances
  - Exécution parallèle
  - Programme maison / Logiciel
- Stockage de données
- Exemples :
  - Calculs sur une série d'images
  - Animation d'un modèle physique



Plan

- Présentation mésocentre de calcul
- **Ressources**
- Accès
- Exécution
- Développement
- Informations générales



Ressources

- Cluster:
  - Noeuds de calcul classiques
  - Noeud GPU
  - Espace de stockage
- Personnel:
  - 1 ASI : administrateur / sites
  - 1 IGE : administrateur
  - 1 IGR : support



LE CLUSTER BULLX **MESOCOMTE**

Architect of an Open World™

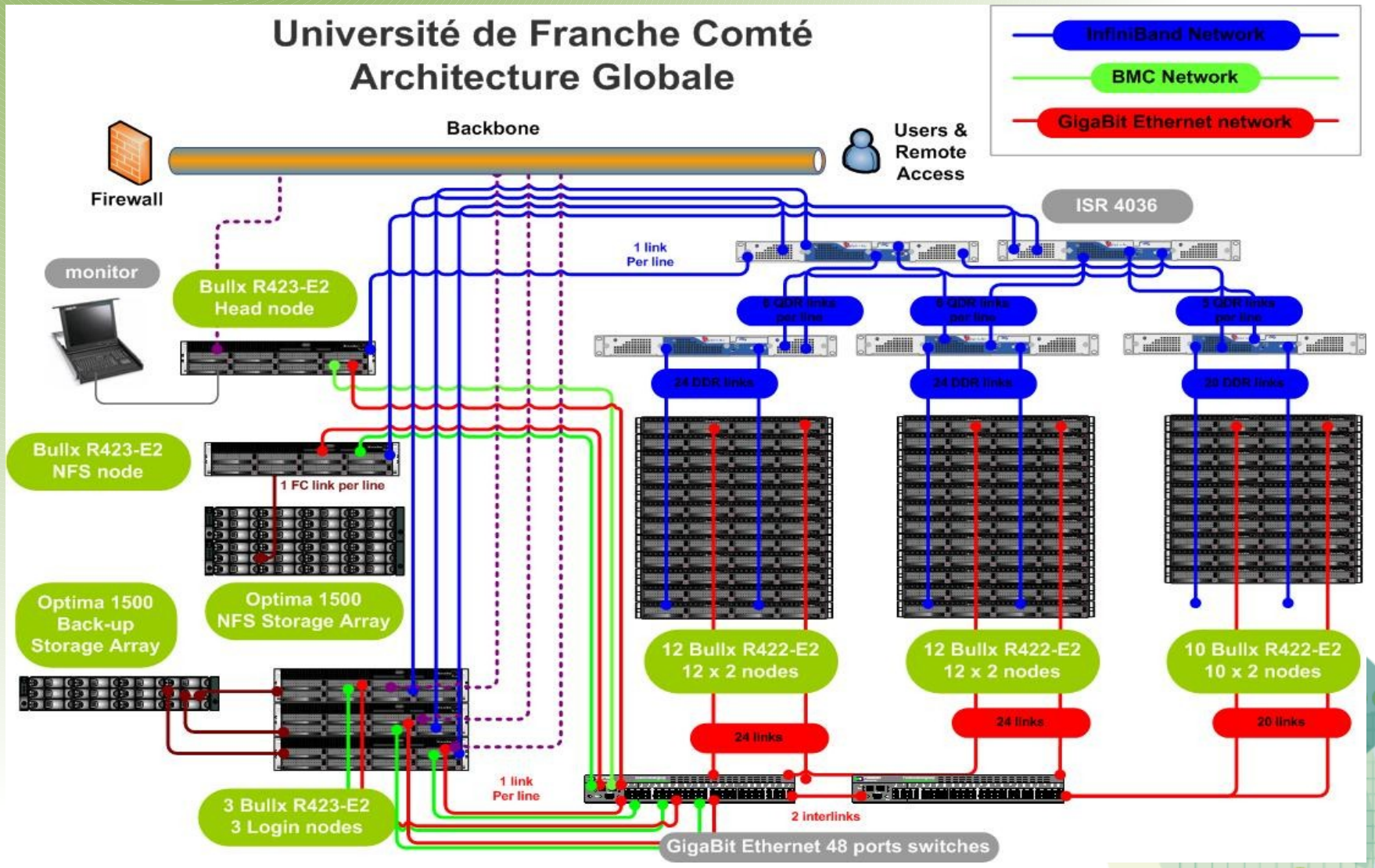
Caractéristiques

Puissance Crête de 5 Tflops

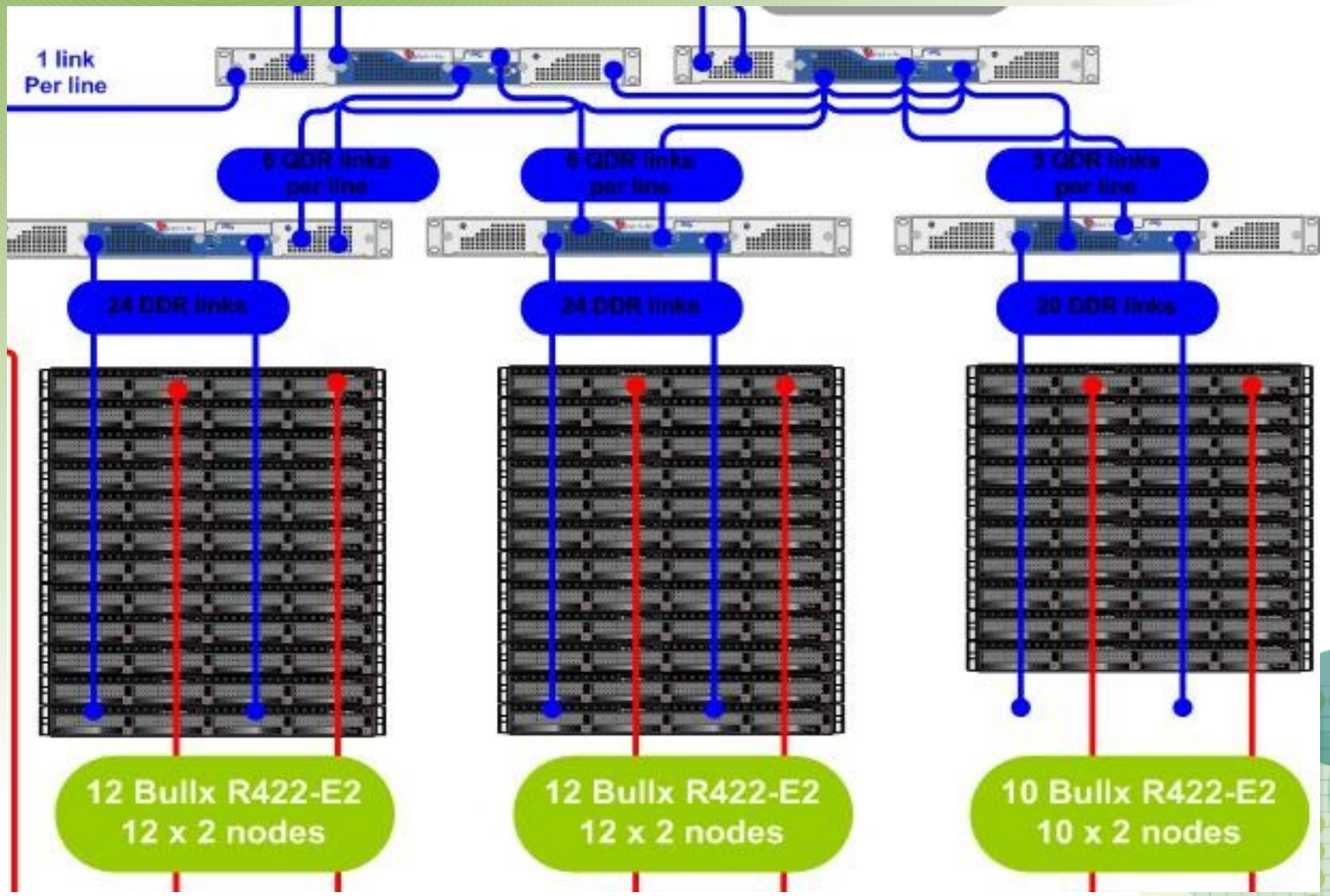
- 68 nœuds de calcul, bi-processeurs, quad cœurs
  - 544 cœurs
  - Intel Xeon Processor E5520 QC (2.26GHz-8MB-80W)
- Un noeud de calcul GP-GPU Telsa s1070
- Réseau de calcul haute performance InfiniBand DDR
- Système de fichiers partagés NFS 20TB
- Système de backup 10TB
- OS Bullx Cluster Suite (Base RHEL 5)
- Batch SGE



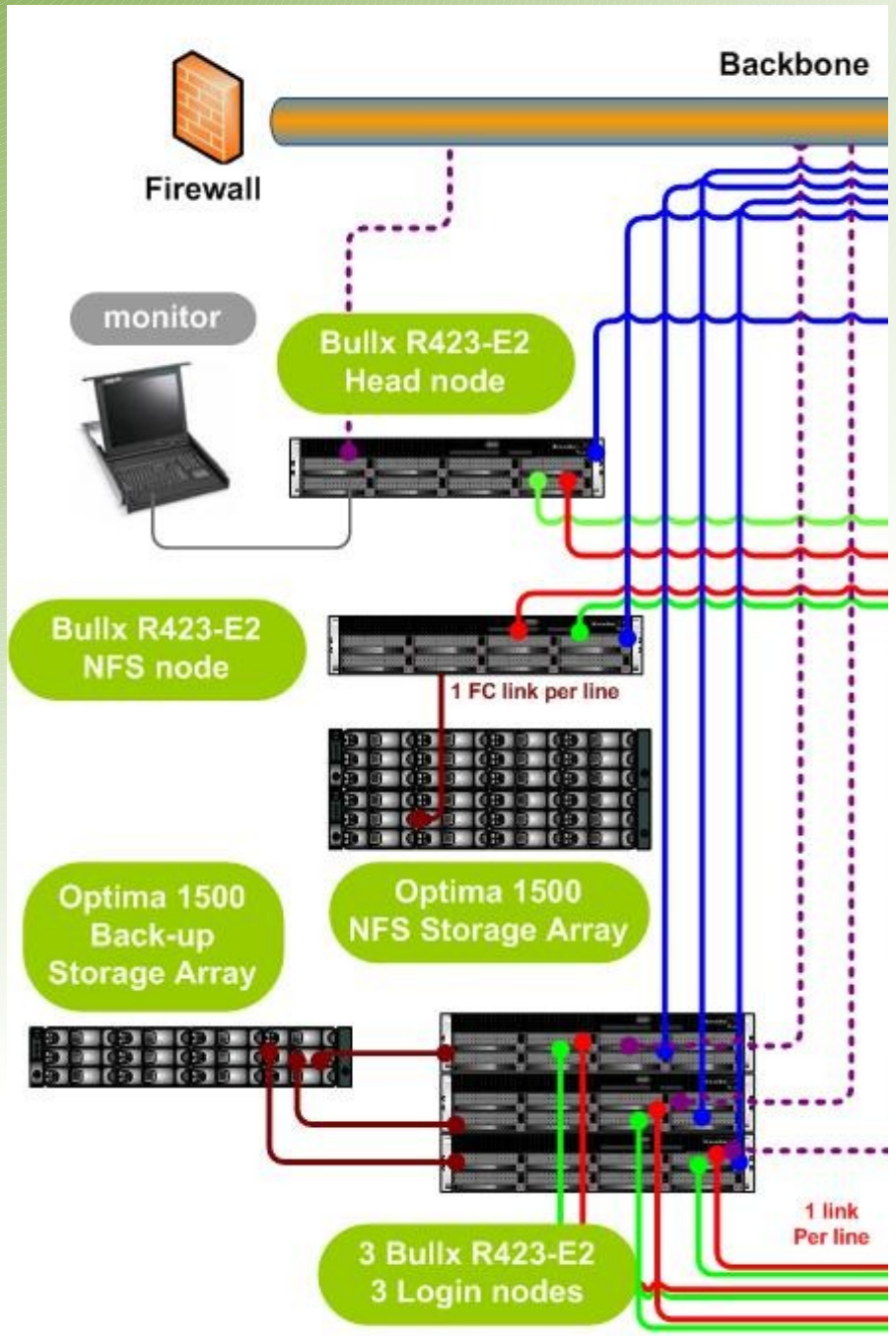
Architecture



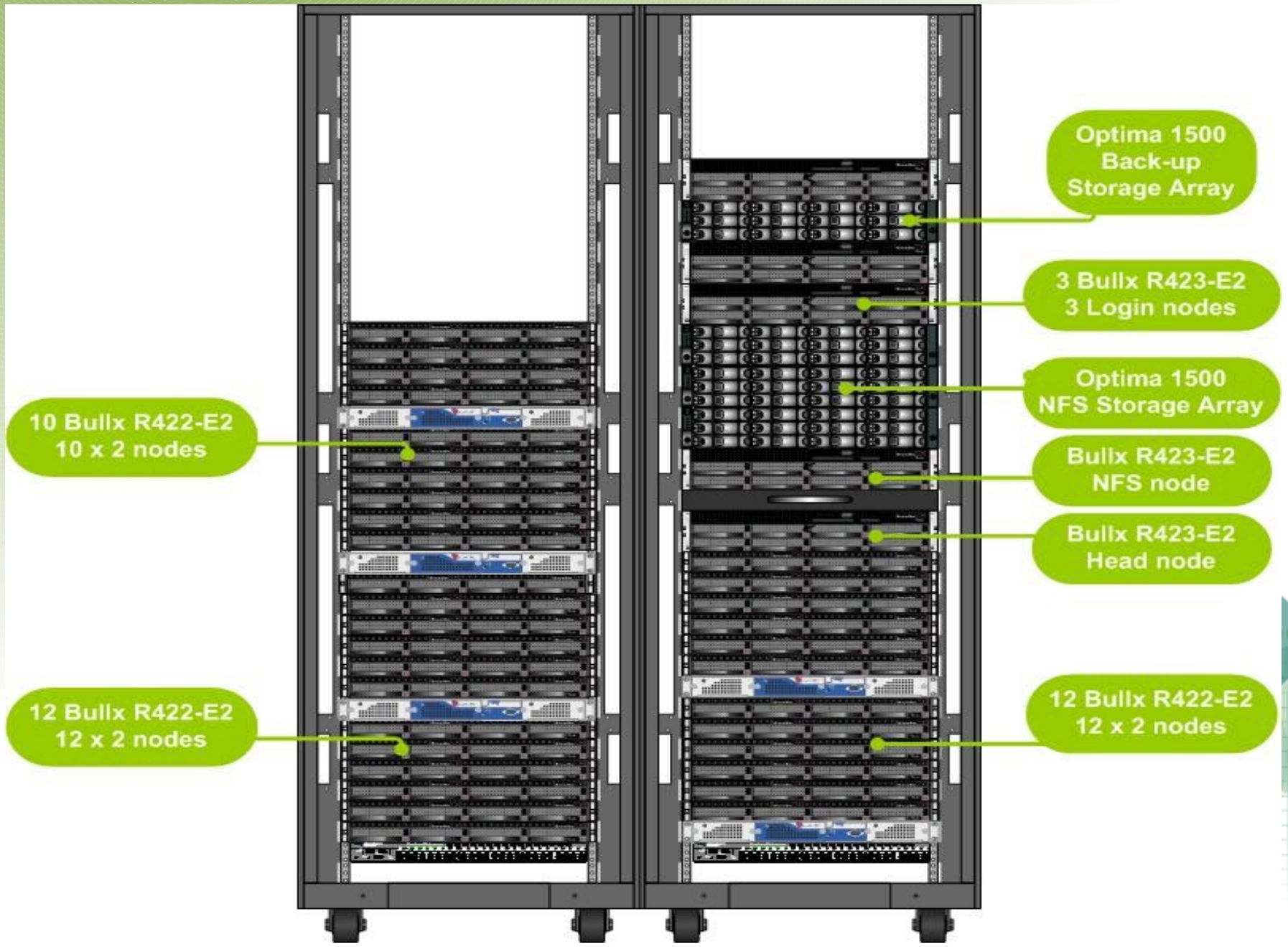
Architecture détaillée



Architecture détaillée



Installation physique



Plan

- Présentation mésocentre de calcul
- Ressources
- **Accès**
- Exécution
- Développement
- Informations générales



Accès



connexion

mesocomte0

Travail interactif

Files d'attente



mesocomte1

mesocomte2

mesocomte10-79

Devenir utilisateur

- Informations : http://meso.univ-fcomte.fr
- Charte
- Demande d'accès
- Compte : login / passwd
- Répertoire personnel : /Data/Users/login
- Accès au cluster :
  - Ligne de commande
  - Interactif



Charte

- Engagements :
  - Utilisation correcte des ressources
  - Respect de la propriété
  - Sécurité réseau
  - Logiciels sous licence
  - Citation dans les publications
- Signée par (premier accès) :
  - Utilisateur
  - Correspondant informatique
  - Directeur de thèse (doctorant)



Demande d'accès

- Connaître les personnes qui accèdent
- Identifier les besoins
- Renouvellée tous les ans
- Création/renouvellement :
    - Login
    - Répertoire personnel
- Mailing liste :

    meso-utilisateurs@univ-fcomte.fr



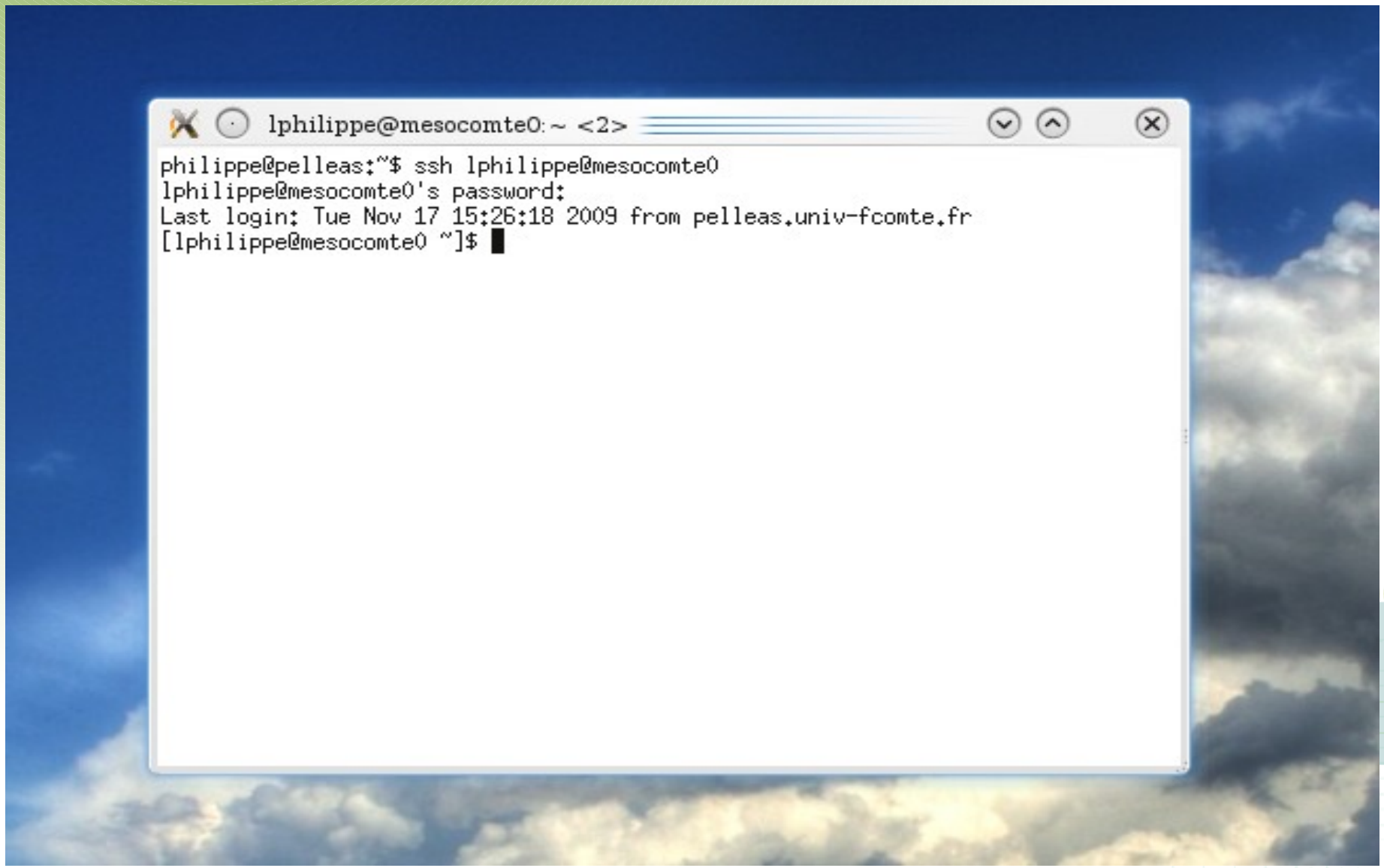
Utilisation

- Simple :
  - Connexion distante : ssh / sftp (putty, winscp)
  - Recopie données
  - Installation programme / Compilation
  - Exécution
  - Obtenir les résultats
- Avancée :
  - Parallélisme d'exécution
  - Parallélisme mémoire partagée
  - Parallélisme mémoire distribuée



Accès Linux




lphilippe@mesocomte0:~ <2>
philippe@pelleas:~$ ssh lphilippe@mesocomte0
lphilippe@mesocomte0's password:
Last login: Tue Nov 17 15:26:18 2009 from pelleas.univ-fcomte.fr
[lphilippe@mesocomte0 ~]$

Copie fichiers Linux



- PC → mesocomte0 : envoi données / programme

```
$ sftp nom_user@mesocomte0
sftp> cd repertoire_principal/sous_repertoire
sftp> put nom_fichier
sftp> bye
```

- mesocomte0 → PC : récupération données / prgm

```
$ sftp nom_user@mesocomte0
sftp> cd repertoire_principal/sous_repertoire
sftp> get nom_fichier
sftp> bye
```

Utilisation Linux shell



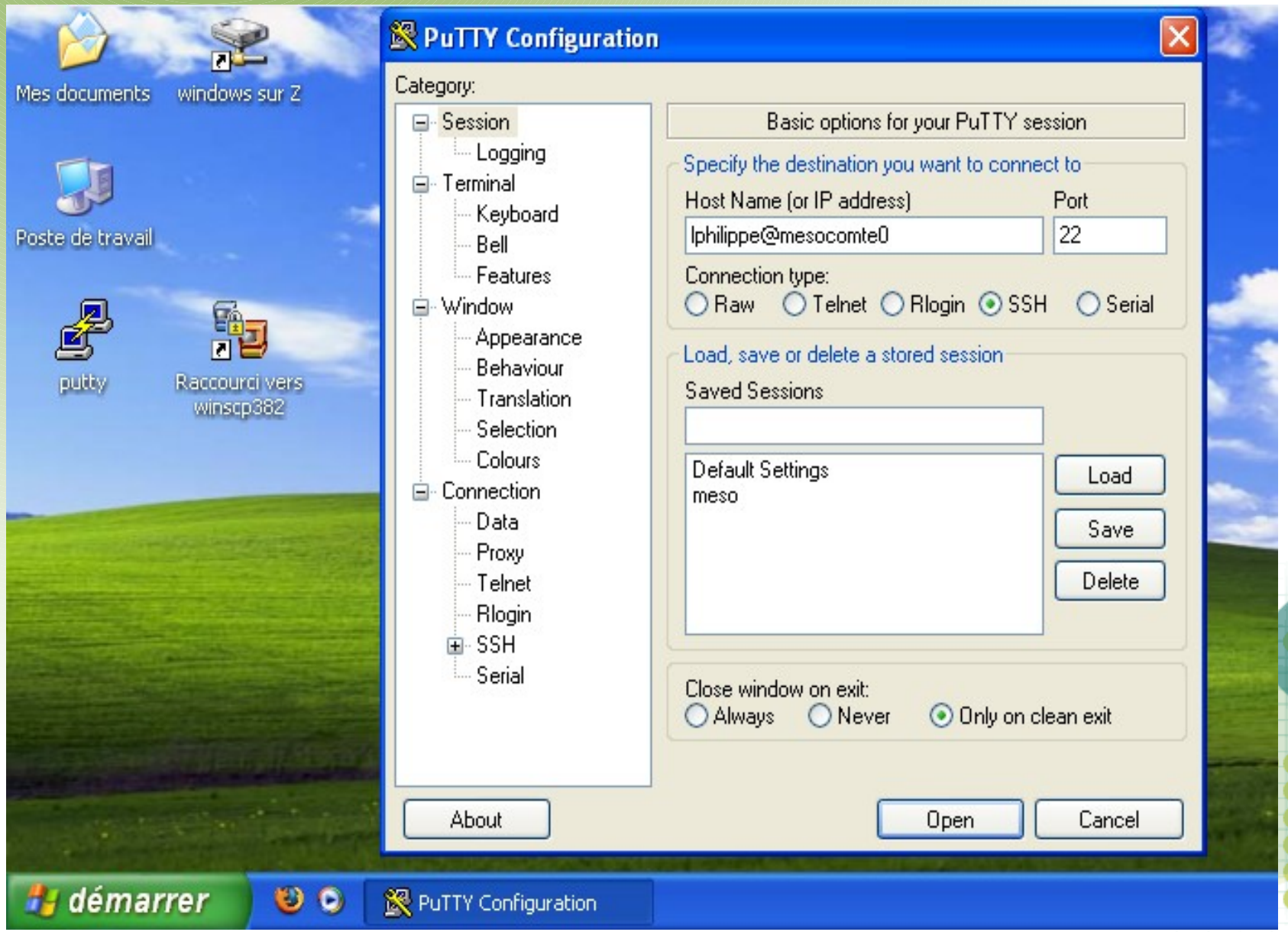
philippe@pelleas: ~/devel/meso/samples/c

```
philippe@pelleas:~/devel/meso/samples/c$ sftp lphilippe@mesocomte0
Connecting to mesocomte0...
lphilippe@mesocomte0's password:
sftp> cd samples/c
sftp> put hello.c
Uploading hello.c to /Data/Users/lphilippe/samples/c/hello.c
hello.c                                       100%   80     0.1KB/s   00:00
sftp> bye
philippe@pelleas:~/devel/meso/samples/c$
```

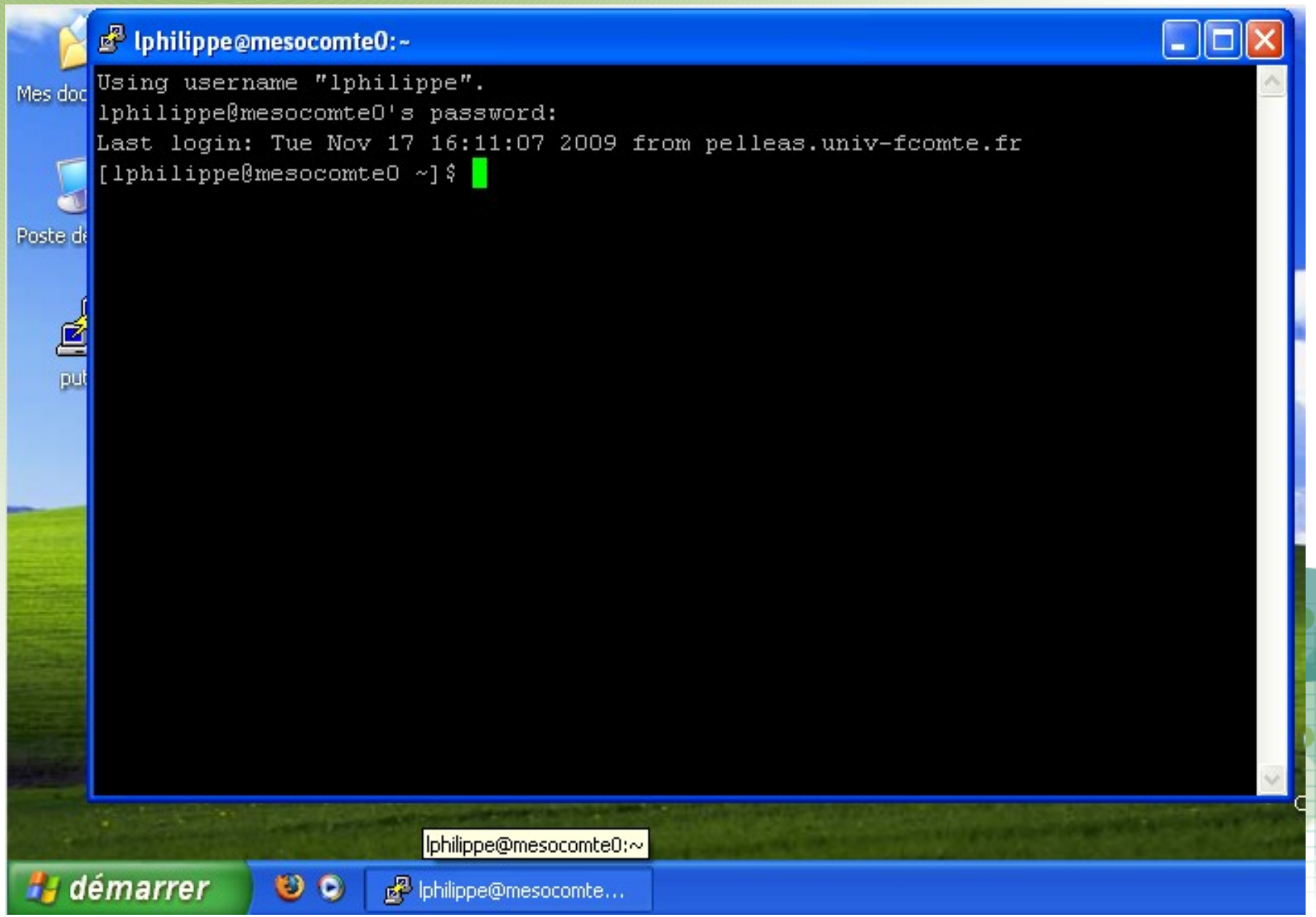
lphilippe@mesocomte0:~/samples/c

```
philippe@pelleas:~$ ssh lphilippe@mesocomte0
lphilippe@mesocomte0's password:
Last login: Tue Dec  1 13:08:53 2009 from pelleas.univ-fcomte.fr
[lphilippe@mesocomte0 ~]$ cd samples/c
[lphilippe@mesocomte0 c]$ ls
hello.c
[lphilippe@mesocomte0 c]$  gcc -o hello hello.c
[lphilippe@mesocomte0 c]$ ls
hello  hello.c
[lphilippe@mesocomte0 c]$ ./hello
Hello from mesocomte0
[lphilippe@mesocomte0 c]$ inter ./hello
Hello from mesocomte1
[lphilippe@mesocomte0 c]$ inter ./hello
Hello from mesocomte1
[lphilippe@mesocomte0 c]$ inter ./hello
Hello from mesocomte1
[lphilippe@mesocomte0 c]$ inter ./hello
Hello from mesocomte2
[lphilippe@mesocomte0 c]$
```

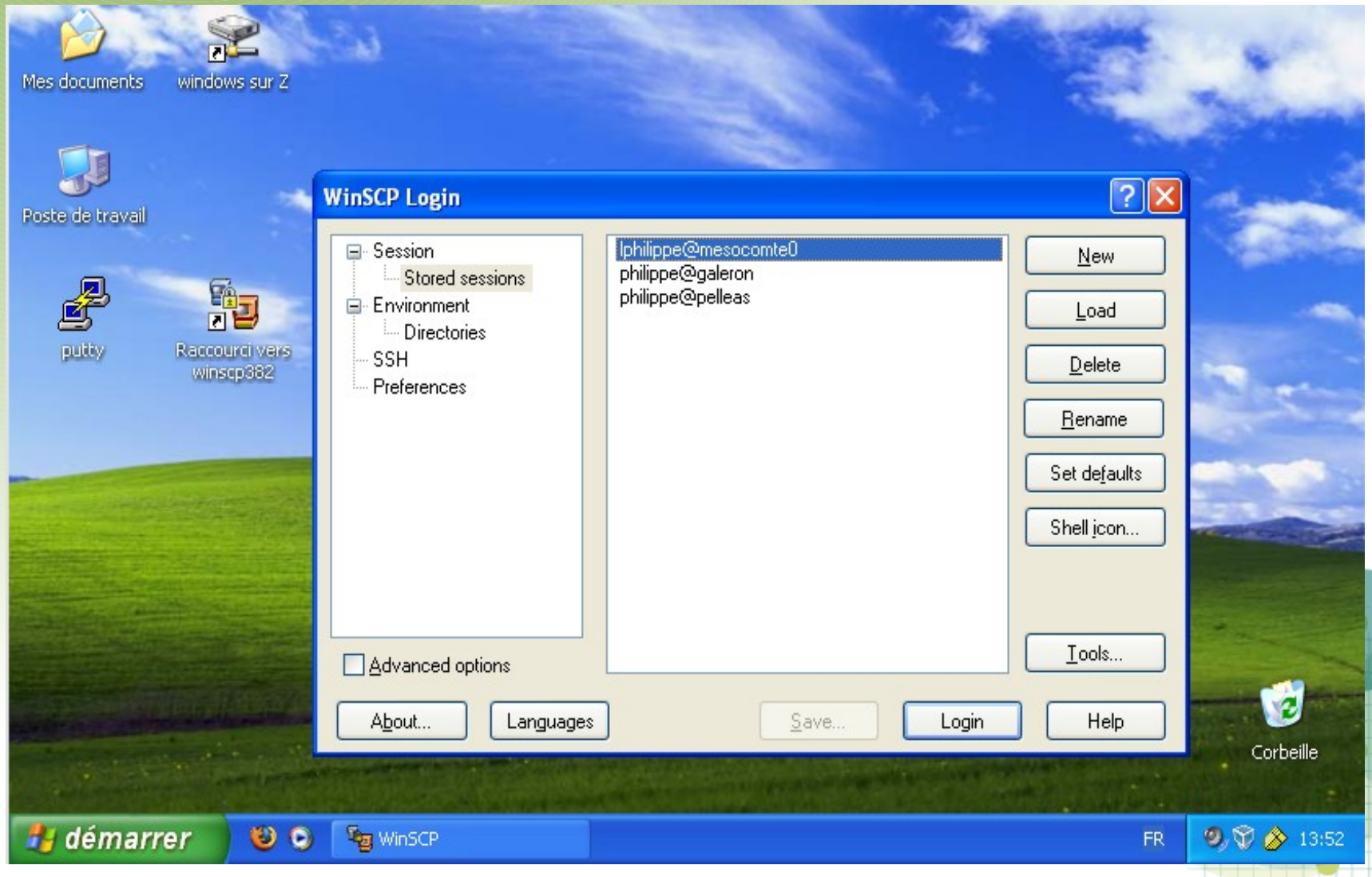
Accès Windows



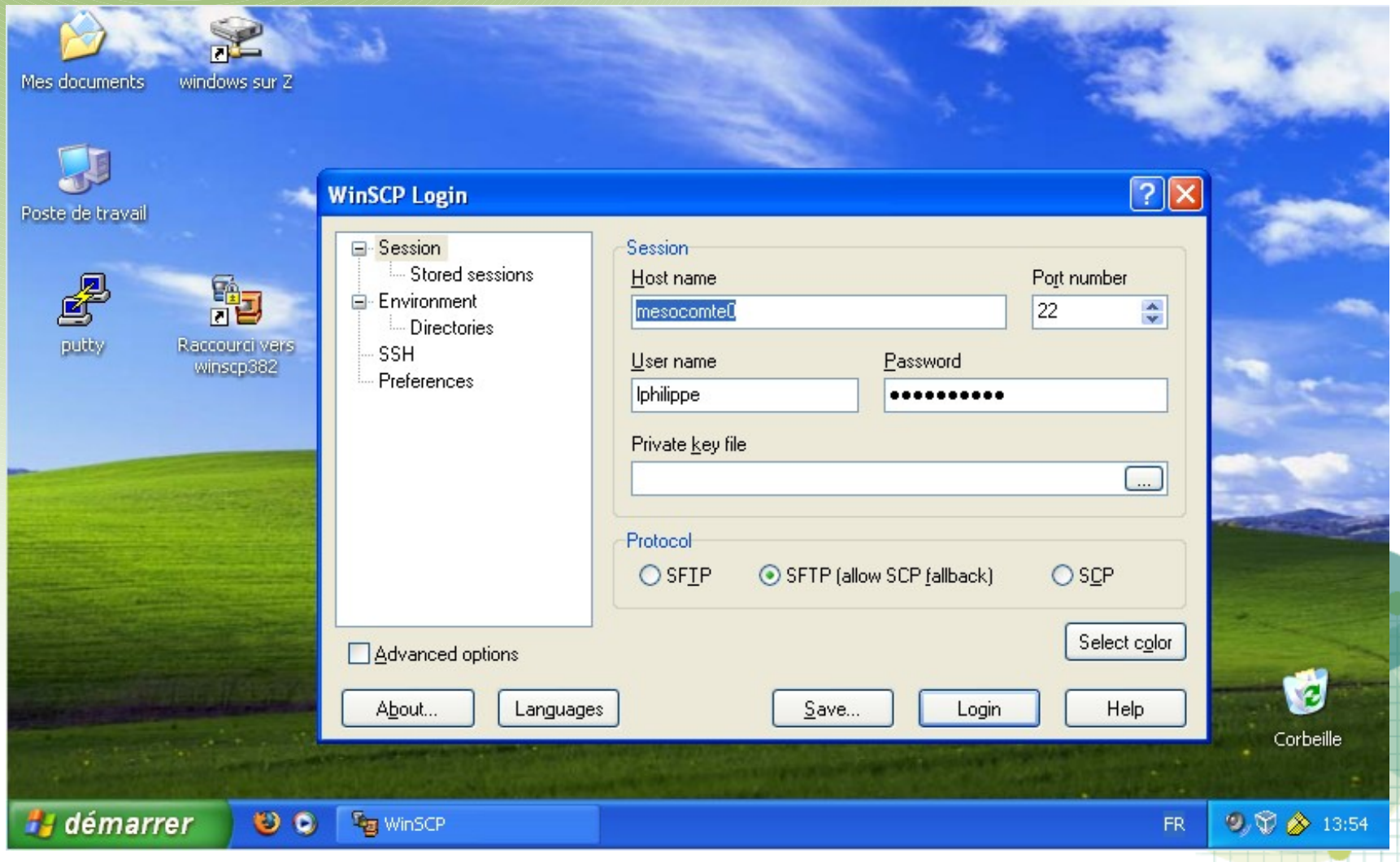
Accès Windows



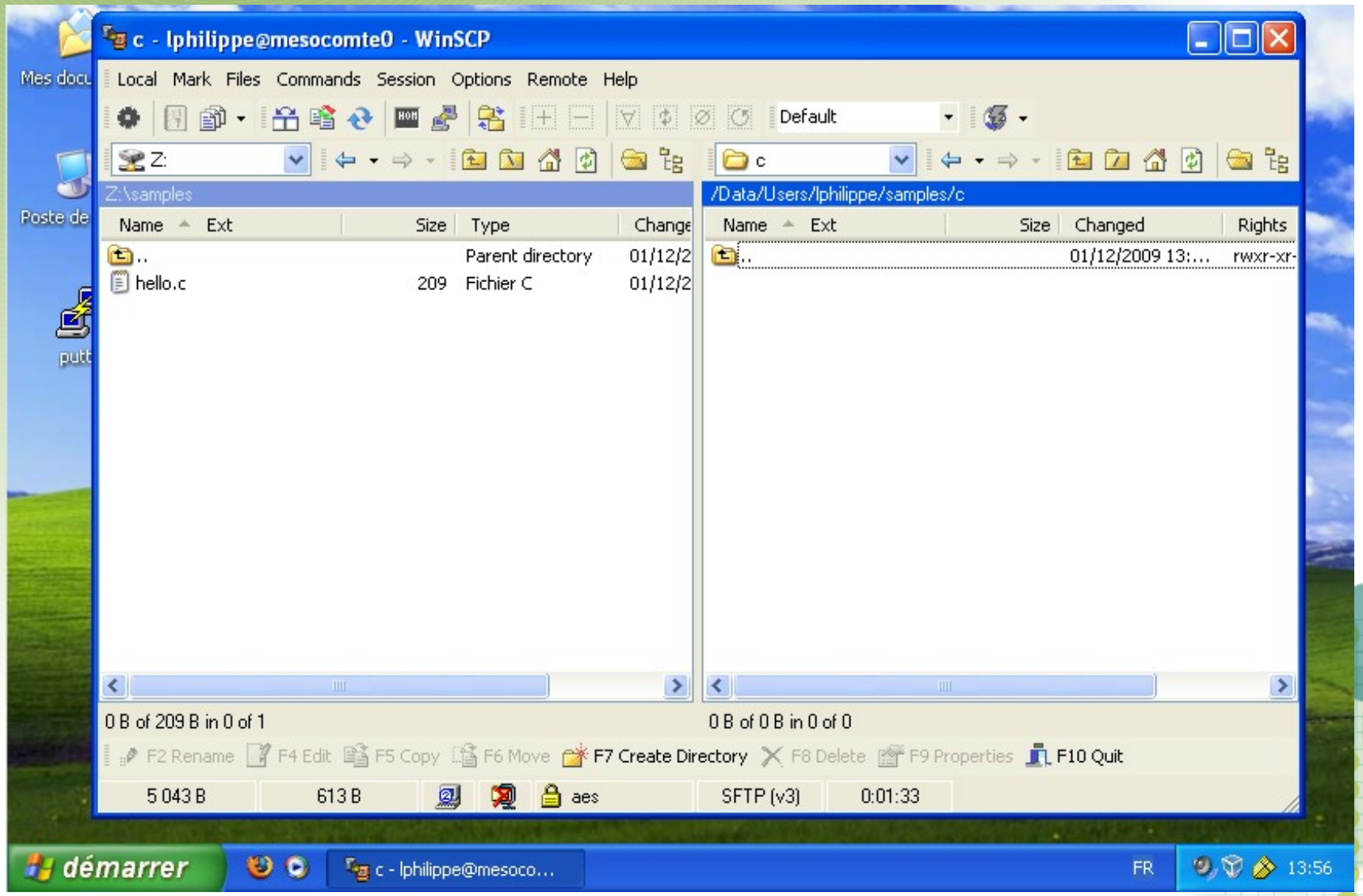
Copie fichiers Windows



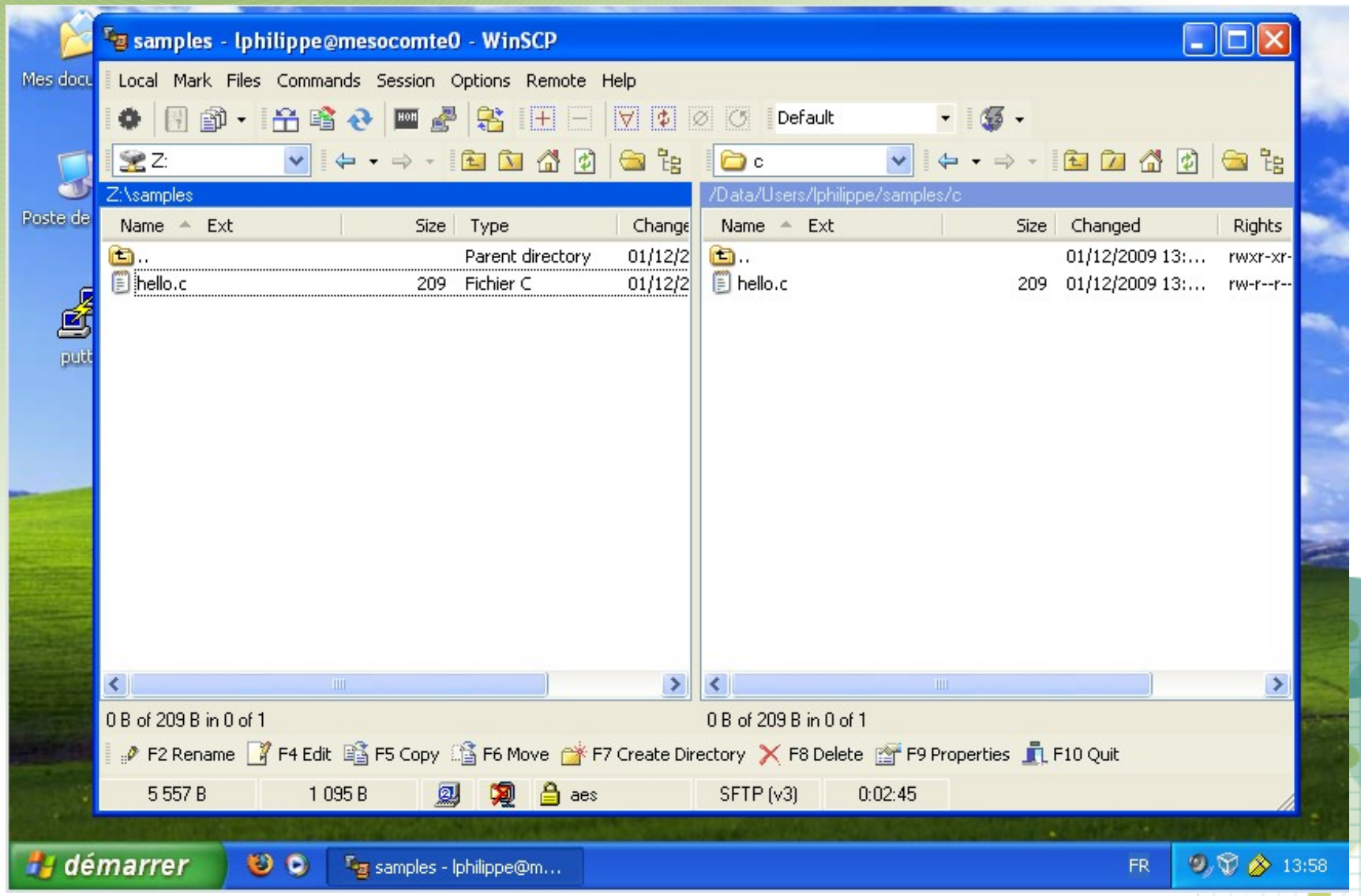
Copie fichiers Windows



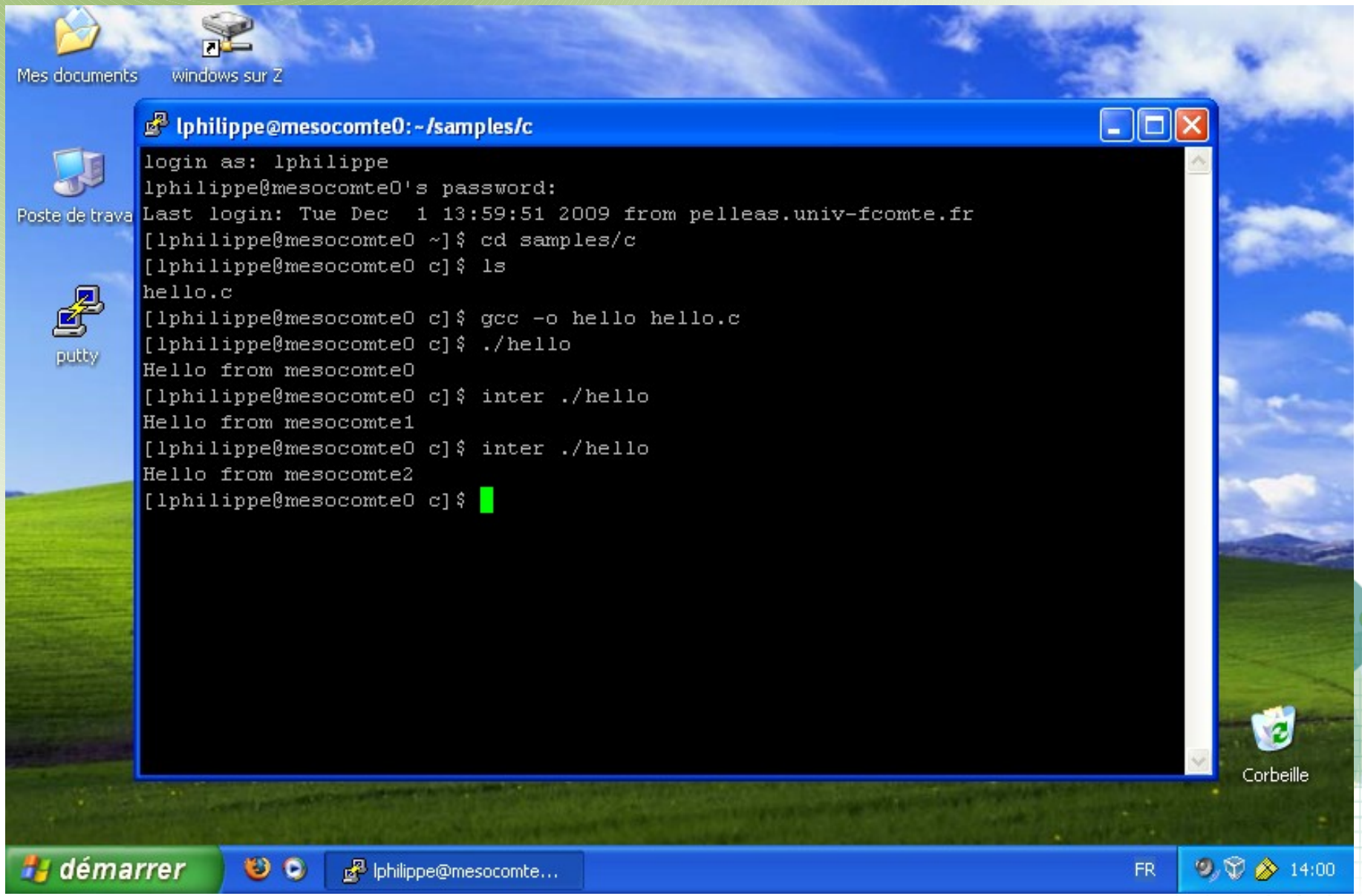
Copie fichiers Windows



Copie fichiers Windows



Utilisation Windows

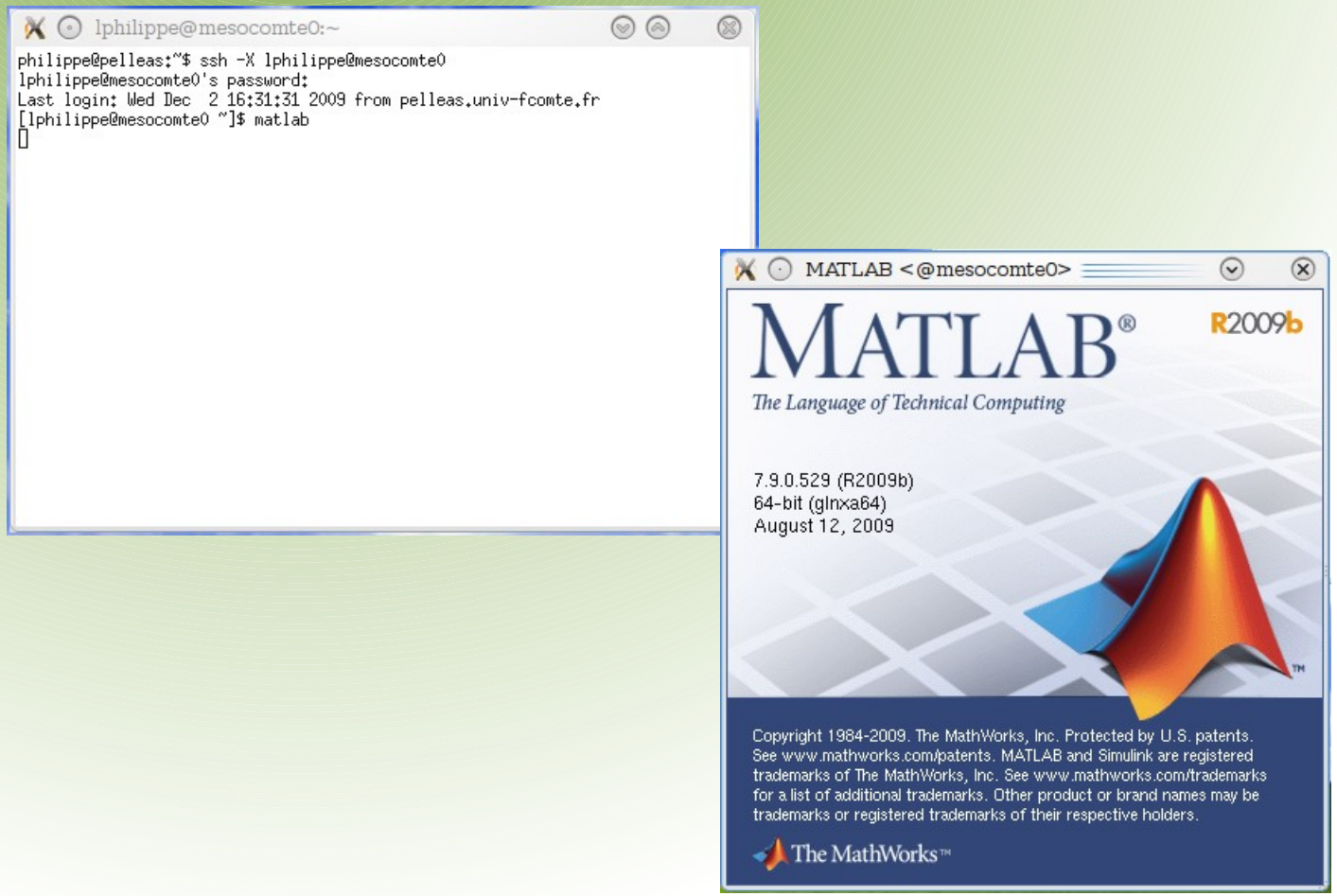


Utilisation interactive

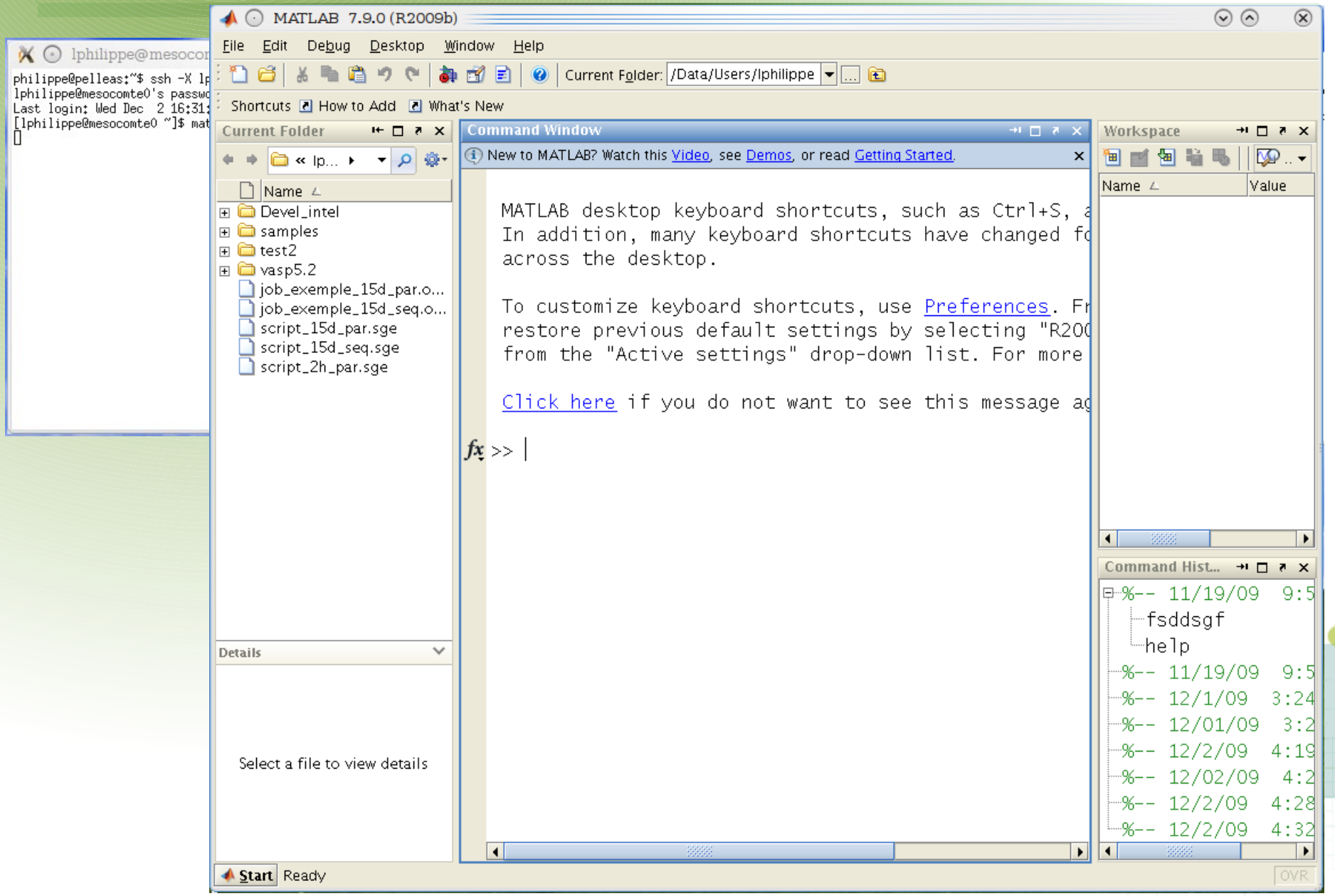
- Applications à fenêtre(s)
- Linux :
  - Disposer d'un serveur X sur la station cliente
  - ssh -X nom_user@mesocomte0
- MacOS :
  - Lancer X : Applications → Utilitaires → Lancer X11
  - ssh -X nom_user@mesocomte0
- Windows :
  - Installer Cygwin/X
  - Lancer une fenêtre Cygwin, puis Xwin dans la fenêtre
  - ssh -X nom_user@mesocomte0



Utilisation interactive Linux



Utilisation interactive Linux



Accès extérieur

- Mesocomte0 : réseau privé université
- Accès extérieur:
  - VPN
  - Certificat :
    - Correspondant informatique
    - CRI
    - Mésocentre
  - Installation : http://vpn.univ-fcomte.fr/



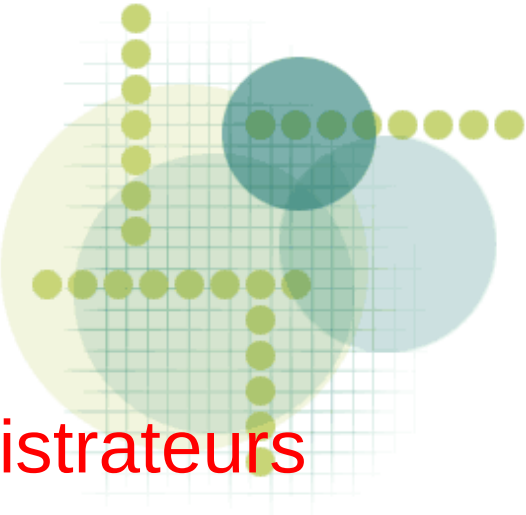
Plan

- Présentation mésocentre de calcul
- Ressources
- Accès
- **Exécution**
- Développement
- Informations générales



Exécution

- Soumission en file d'attente = préparer
- Valider l'environnement :
  - Variables d'environnement: PATH, LD_LIBRARY_PATH, etc.
  - Exécution locale
- Commande / Application :
  - Paramètres
  - Fichiers entrée / sortie :
    - Chemin d'accès
    - Positionnement



En cas de problème : demander aux administrateurs

Utilisation files d'attente

- Procédure: SGE = Sun Grid Engine
  - Soumettre un job
  - Attente
  - Fichiers de sortie
- Script :
  - File d'attente
  - Temps d'exécution
  - Fichiers de sortie
  - Exécutable
- Soumission : `qsub script`



Utilisation files d'attente

lphilippe@mesocomte0:~/samples/c

```
Connection to mesocomte0 closed.
philippe@pelleas:~$ ssh lphilippe@mesocomte0
lphilippe@mesocomte0's password:
Last login: Thu Dec  3 11:17:57 2009 from pelleas.univ-fcomte.fr
[lphilippe@mesocomte0 ~]$ cd samples/c
[lphilippe@mesocomte0 c]$ ls
hello  hello.c  script_2h_seq.sge
[lphilippe@mesocomte0 c]$ qsub script_2h_seq.sge
Your job 32869 ("lphilippe_exemple_2h") has been submitted
[lphilippe@mesocomte0 c]$ qstat -t
job-ID  prior   name       user         state submit/start at     queue
---------------------------------------------------------------------------
  32869 0.00000 lphilippe_ lphilippe    qw    12/03/2009 11:25:03
[lphilippe@mesocomte0 c]$ ls
hello  hello.c  script_2h_seq.sge
[lphilippe@mesocomte0 c]$ qstat -t
[lphilippe@mesocomte0 c]$ ls
hello  hello.c  lphilippe_exemple_2h.o32869  script_2h_seq.sge
[lphilippe@mesocomte0 c]$ more lphilippe_exemple_2h.o32869
Hello from mesocomte11
[lphilippe@mesocomte0 c]$
```



Exemple de script

```
#!/bin/bash

# Pour utiliser ce  script, faire  qsub de celui-ci
# apres renseignement des parametres
# qsub <nom du script>

# Temps max pour l'execution du processus
# Plus il est court, plus la priorite est importante
# Le job sera tue s'il depasse ce temps
#$ -l h_rt=2:00:0

# Nom du fichier ou stocker la sortie
#$ -o $JOB_NAME.o$JOB_ID

# Nom du fichier ou stocker les erreurs
#$ -e $JOB_NAME.e$JOB_ID

# Nom du job
#$ -N lphilippe_exemple_2h

# Queue pour les jobs courts < 2h
#$ -q normal2h
```

```
# SGE doit etre bavard
#$ -V

#######################################
### COMMANDES A LANCER, DEBUT ###
#######################################

DATADIR=$HOME/samples/c

# Mettre ici le nom de votre executable

./hello

#######################################
###  COMMANDE A LANCER, FIN  ####
#######################################

exit 0
```

Attention à la défintion des variables d'environnement

Files d'attente

- Files d'attentes :
  - interactive : partagée → 3 noeuds
  - normal2h < 2 h → 4 noeuds
  - normal15d < 15j → 67 noeuds
  - normal2m < 2 m → 4 noeuds
  - bigmem15d : 96 Go → 1 noeud
  - GPU
- Exemples de scripts :
  - /Data/Commun
  - Séquentiel / Parallèle



Files d'attente

- Commandes de base :
  - qdel : détruire un job (-N ou -u user)
  - qstat : état des jobs (-f : état des files)
  - qhost : état des machines (-F -h machine)



Plan

- Présentation mésocentre de calcul
- Ressources
- Accès
- Exécution
- **Développement**
- Informations générales



Développement

- Langages :
  - C, C++
  - Fortran
  - Java
  - ...
- Parallélisme :
  - Open – MP
  - MPI : Intel + Bull



Logiciels - Bibliothèques

- Installées :
  - Algèbre linéaire :
    - Blas,
    - Lapack,
    - Scalapack
  - FT : FTW
  - Parallélisme :
    - Threads : OpenMp
    - Passage de messages :
      - MPI
      - Communication classique



Logiciels – Open Source

- Installés :
  - Compilateur gnu : gcc, gfort
  - Java : jdk sun
  - Dynamique Moléculaire : NAMD
  - Modélisation électromagnétique : Meep
  - Calcul de statistiques : R
- Prévus :
  - Gamess, Ab-initio
  - Scilab



Logiciels – Commerciaux

- Compilateur : Intel Cluster Toolkit
- Matlab :
  - Simulink
  - Image Processing Toolbox
  - Optimization Toolbox
  - Signal Processing Toolbox
  - Symbolic Math Toolbox
  - Statistics Toolbox
  - Compilateur
- Gaussian
- Molpro



Développement

- Compilation:
  - ifort / icc
  - gcc / gfort
- Mise au point : gdb / idb
- Optimisation :
  - Compilation : -O[1,2,3]
  - Document Bull : optimisation boucles

http://mesoserver/wiki : → documents → présentation



Plan

- Présentation mésocentre de calcul
- Ressources
- Accès
- Exécution
- Développement
- **Informations générales**



Les noeuds du cluster

- Nœud de login : mesocomte0
- Nœuds interactifs : mesocomte0,1,2
- Nœud d'administration : mesocomte8
- Nœud fichiers : mesocomte9
- Nœuds de calcul : mesocomte[10-77]
- Nœuds Tesla : mesocomte78,79




Ganglia:: Cluster Report - Mozilla Firefox
Fichier Édition Affichage Historique Marque-pages Outils Aide
http://194.57.91.98/ganglia/
shared:logiciels_installes | ...
Gestion enseignant - Projet ...
Ganglia:: Cluster Report
Ganglia .sourceforge.net
Cluster Report for Tue, 01 Dec 2009 09:04:40 +0100
Get Fresh Data
Metric load_one
Last hour
Sorted descending
Physical View
Grid > mesocomte > --Choose a Node
Overview of mesocomte
CPUs Total: 584
Hosts up: 73
Hosts down: 0
Avg Load (15, 5, 1m): 82%, 82%, 82%
Localtime: 2009-12-01 09:04
Cluster Load Percentages
100+ (73.97%)
25-50 (1.37%)
0-25 (24.66%)
mesocomte Load last hour
Load/Procs
1-min Load Nodes CPUs Running Processes
mesocomte CPU last hour
Percent
User CPU Nice CPU System CPU WAIT CPU Idle CPU
mesocomte Memory last hour
Bytes
Memory Used Memory Shared Memory Cached
Memory Buffered Memory Swapped Total In-Core Memory
mesocomte Network last hour
Bytes/sec
In Out
Show Hosts: yes no | mesocomte load_one last hour sorted descending | Columns 4
mesocomte53
load_one last hour (now 9.20)
mesocomte16
load_one last hour (now 9.10)
mesocomte28
load_one last hour (now 9.08)
mesocomte14
load_one last hour (now 9.02)
mesocomte40
load_one last hour (now 9.01)
mesocomte48
load_one last hour (now 9.01)
mesocomte43
load_one last hour (now 9.00)
mesocomte58
load_one last hour (now 9.00)

Ganglia .sourceforge.net

Host Report for Tue, 01 Dec 2009 09:07:56 +0100

Get Fresh Data

Last hour

Node View

Grid > mesocomte > mesocomte16

mesocomte16 Overview

This host is up and running.

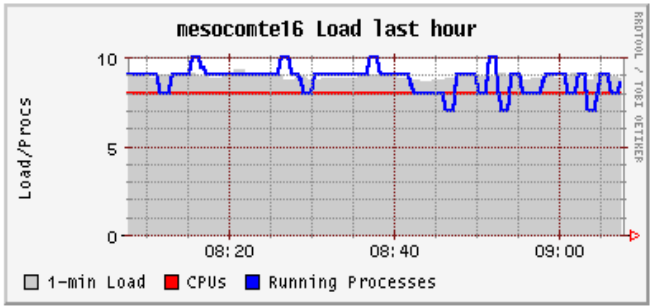
Time and String Metrics

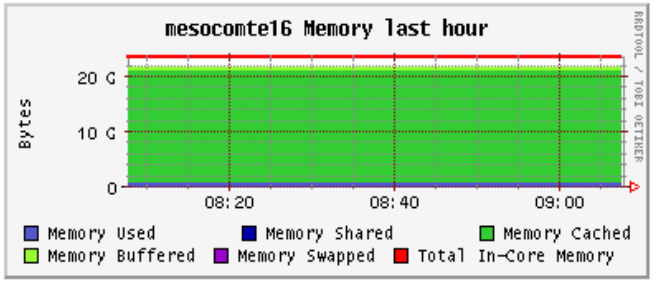
boottime	Sun, 15 Nov 2009 17:16:27 +0100
gexec	ON
gmond_started	Sun, 15 Nov 2009 17:17:25 +0100
last_reported	0 days, 0:00:04
machine_type	x86_64
os_name	Linux
os_release	2.6.18-128.el5.Bull.3
uptime	15 days, 15:51:17

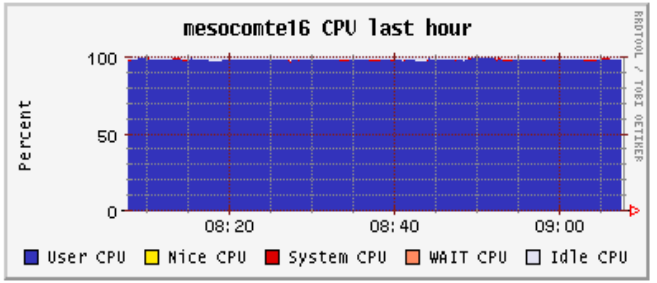
Constant Metrics

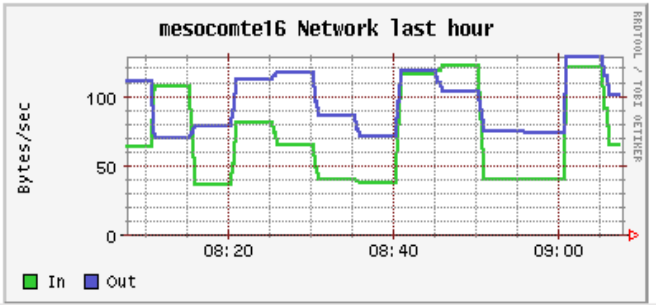
cpu_num	8 CPUs
cpu_speed	2668 MHz
mem_total	24546096 KB
swap_total	16386292 KB

Gmetrics









Limitations

- Système Linux (Redhat)
  - ? Windows :
    - Java
    - Programmes avec DLL
    - Emulation / Machines virtuelles
- Ondulation seulement sur noeuds maîtres
  - Points de reprise
  - Re-soumission
- Sauvegardes (en cours)
  - Limitées
  - Pas d'archivage
- Soumission batch principalement



Organisation

- Comité de pilotage :
    - Décisions stratégiques
- Comité scientifique :
    - Choix techniques, proposition d'investissement
- Groupe des utilisateurs :
    - Réunion annuelle
    - Formation
    - Proximité = propositions



Comité scientifique

- FEMTO-ST : Eric Duverger, Lamine Boubakar
- UTINAM : Françoise Gazelle, Jean-Marc Petit, Christophe Ramseyer
- LIFC : Raphael Couturier
- ENSMM : Michel Devel
- UTBM : Sébastien Roth
- CRI : Marc Hamelin
- ...



Prévisions

- Logiciels :
  - COMSOL
- Matériel :
  - Stockage / Base données
  - Visualisation distante
  - Noeuds



Contacts et informations

- Site WEB : http://meso.univ-fcomte.fr
- Site intranet : http://mesoserver.univ-fcomte.fr
- Liste de diffusion : meso-utilisateurs@univ-fcomte.fr
    - Adhésion nécessaire, ouverture de compte
- En cas de problème : svpmeso@univ-fcomte.fr

- Citer l'utilisation / Reporter au mésocentre

