

GLite - Installation d'un site typique

Nicolas Spalinger
HealthGrid

Emmanuel Medernach
CNRS - LPC



Credits: ressources et présentations NA3

- Pas un monologue !
- Pas tout cuit, un défi à relever ensemble
- Ressources variées, docs de référence mais changements fréquents, évolutions à suivre:
 - Twiki du CERN (liens)
 - Grid-workshop par HealthGrid et autres
 - GridPP wiki
 - Gocwiki
 - etc
- License libre type BSD, marques déposées

- Environnement de formation:
 - Machine virtuelles (vmware)
 - Schéma de répartition et comptes d'accès
 - Ressources externes du testbed
- Version visées du middleware: gLite 3.2
- OS spécialisé : SL 5.4 64bits
- Mises à jour et miroirs locaux/distants (miroir du CC IN2P3)

- Installation
- Dépôts spécialisés
- Métopaquets
 - Yum
- Configuration
 - Yaim
- Service d'arrière plan
 - Ntpd
 - openssh
 - Voir pstree, htop
 - Service inutiles désactivés: avahi, cups, etc

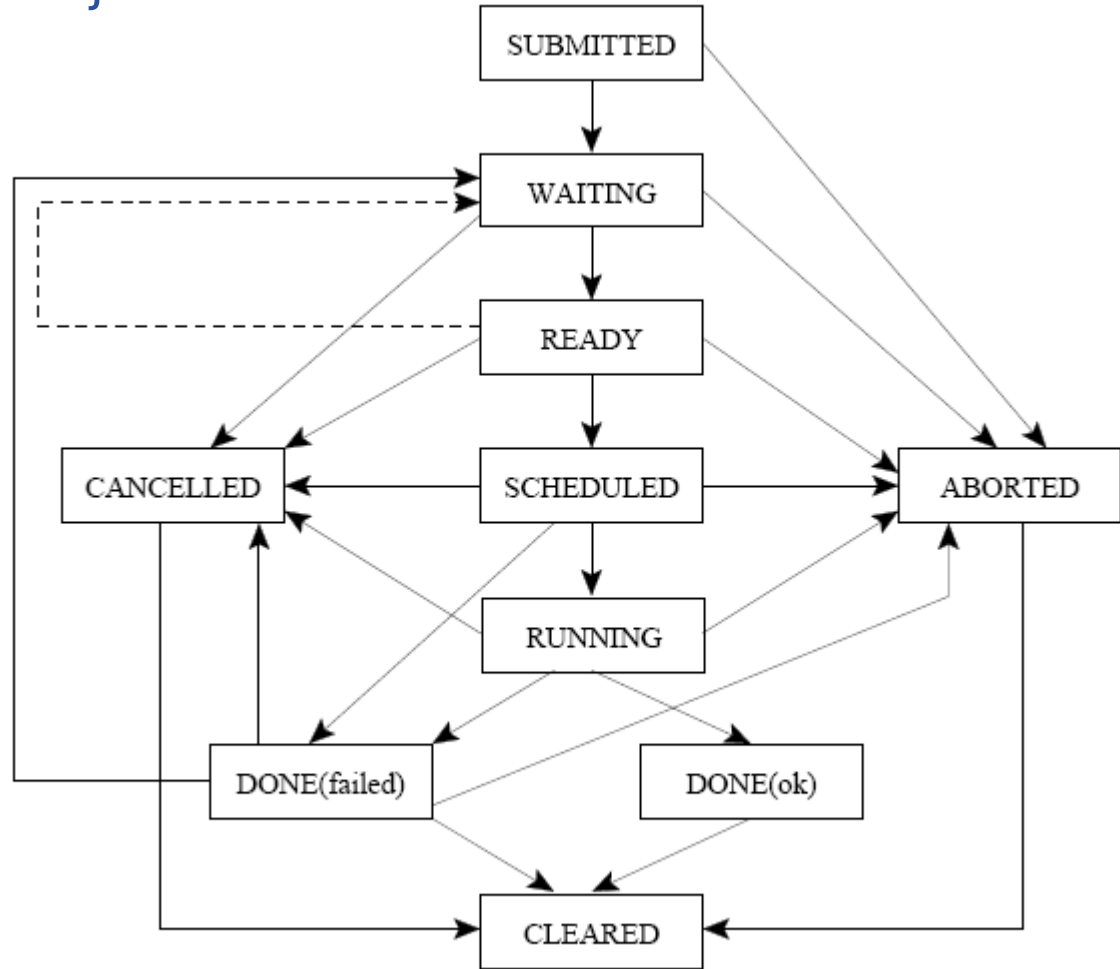
- Systeme de base pré-configuré
 - Yum et les dépôts de base à jour
 - Ntp installé et configuré
 - JDK installé et configuré (attention à la version et aux dépendances)
 - Dépôts de base définis dans les .repo

- Certificats serveurs (/certificats)
- Certificats utilisateurs → distribution
- Récupération et sécurisation de son certificat personnel: mécanisme habituel
- Utilisation dans le navigateur et client mail
- Mise en place et usage du certificat sur l'UI
 - Export
 - Installation dans ~/.globus
 - Utilisation une fois l'UI installée (ou via portail par la suite)

UI: User Interface



- Etats possibles d'un job :



- Installation d'une UI: User Interface
 - Dépôts
 - Metapaquet installé via yum
 - site-info.def
 - Attention aux VOs de type DNS
 - Yaim pour la configuration
 - Wiki grid-workshop pour méthodologie complète
 - Mise en place du certificat personnel
 - Test de requêtage de ressources:
 - lcg-info(sites)

- Test de soumission
 - voms-proxy-init
 - glite-wms-job-delegate-proxy
 - glite-wms-job-submit
 - glite-wms-job-status
 - glite-wms-job-output

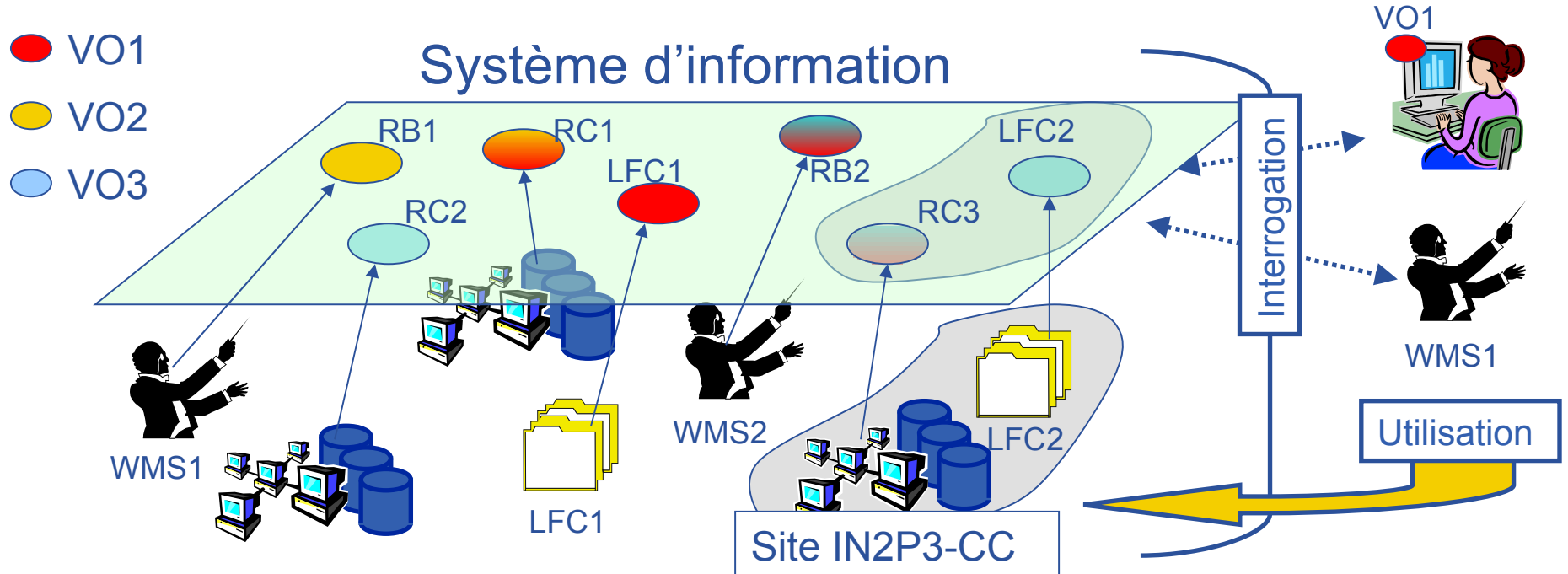


Site BDII

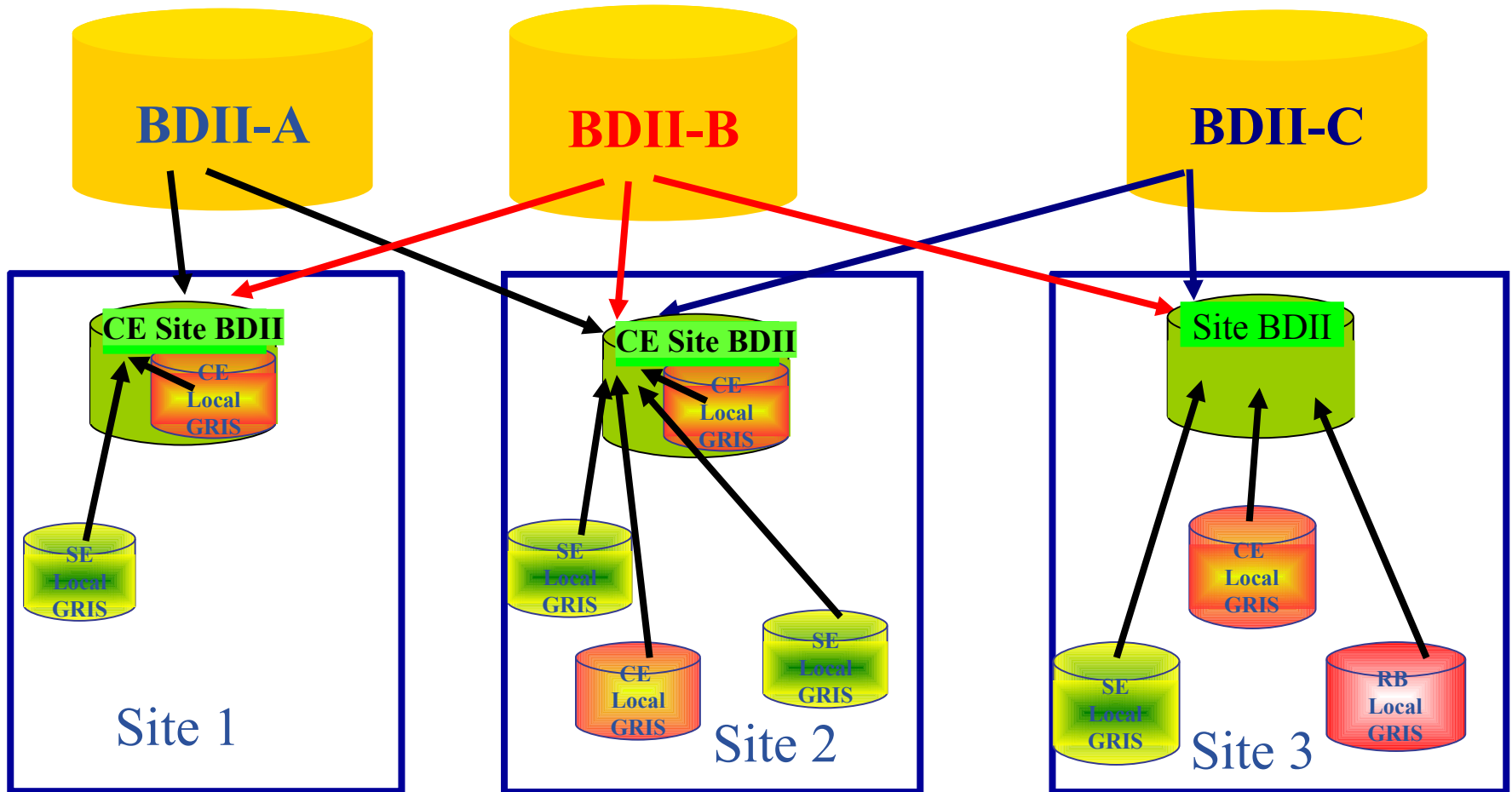


Principes

- Chaque site public
 - Une description des ressources/services qu'il fournit par VO
 - L'état actuel de ses ressources (CPUs libres, Espace de stockage, etc.)
- Chaque VO publie (anecdotique)
 - Ce qu'elle a installé sur chaque site ("Tags" des Software Managers)
- Les acteurs interrogent le SI pour savoir comment utiliser les services/ressources de la grille



- **LCG-2/GLITE : Agrégations des systèmes d'information des sites**
 - BDII (Berkeley Database Information Index): même fonctionnalité que le GIIIS, avec une implémentation différente (meilleure gestion de cache)
 - 2 niveaux d'agrégation: de GRIS (site BDII) et de site BDII (top BDII)



- Annuaire des services locaux du site
- Technologie LDAP: openldap
- Port dédié
- Service de fusion de données ldap
- Hiérarchie entre les différents BDII: BDII, site BDII, topBDII

- Installation
- Depuis VM de base
 - ntp
 - dépôts spécifiques
 - Remplir le site-info.def et autres fichiers
 - Yaim
 - Tests
- Wiki grid-workshop pour méthodologie complète



SE: Storage Element DPM head

- **installation depuis VM blank**
 - Ntp
 - site-info.def
 - Yaim
 - create-default-dirs-DPM.sh
 - mysql
 - /etc/shift.conf
 - CE_LIST dans site-info.def, le CE le plus proche
 - Debug logs
 - Wiki grid-workshop pour méthodologie complète
- **dpm-qryconf**
- **déclarer le SE dans le sBDII**



SE: Storage Element

DPM disk

- **installation depuis VM blank**
 - Ntp
 - site-info.def
 - Yaim
 - Debug logs
 - Wiki grid-workshop pour méthodologie complète
- **indiquer SE_LIST dans site-info.def, le SE le plus proche**
- **dpm-addfs**
- **Tests de lecture et transfert depuis DPM head**



CE: Computing Element



- Install depuis VM de base
 - Ntp
 - JDK
 - Dépôts
 - site-info.def et autres fichiers
 - Yaim
- Wiki grid-workshop pour méthodologie complète

- services d'arrière-plan pbserver / maui
- logs aussi dans
 - /var/spool/pbs
 - /var/spool/maui
- **Déclaration du CE dans le sBDII**



WNs: Worker Nodes



- **WN1** Install depuis VM de base
 - Configuration similaire au CE
- **authentication / mise en confiance entre CE et WN via clés ssh**
- **debuggage via logs dans**
 - /var/spool/pbs sur le WN
- Wiki grid-workshop pour méthodologie complète



Tests

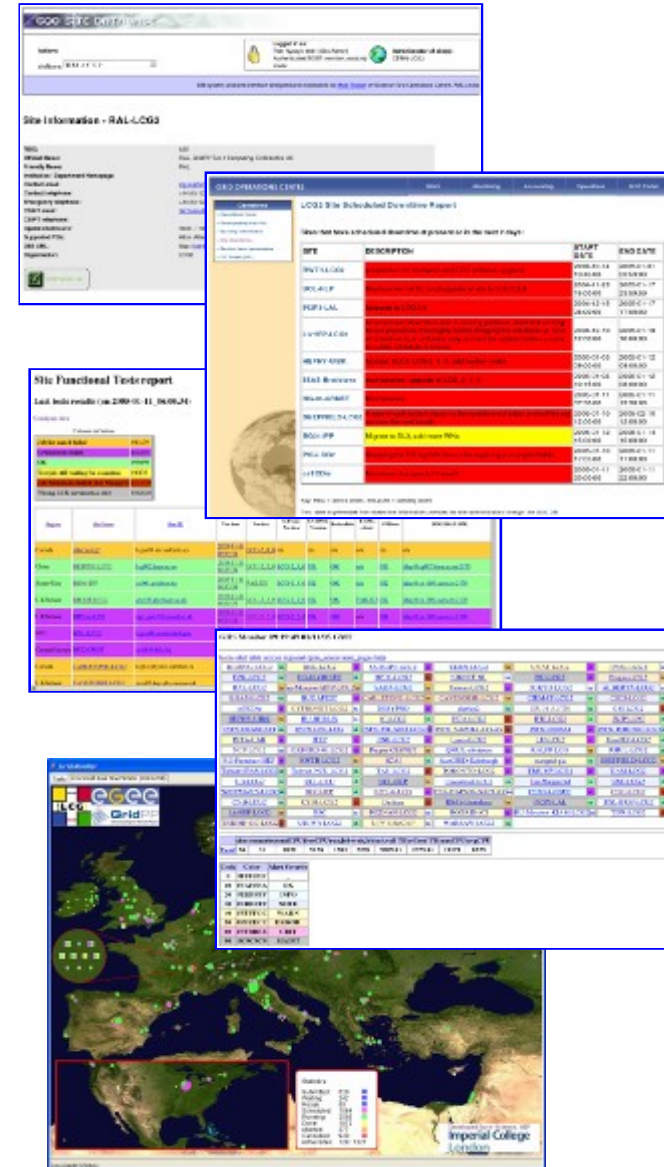


- **test des éléments installés**
- **gstat expose les commandes de requête LDAP pour vérifier la publication des services**
- ***BDII ports:**
 - **2170: sBDII**
 - **2170 topBDII**
 - **2135 service locaux: CE/SE/etc**

- **localement**
- **SAMAP portail de test distant**
 - besoin de déclaration du DN
- **où chercher pour les erreurs:**
 - /var/log/messages
 - /var/spool/
 - /opt/glite/* aussi (dimensionnement généreux des partitions puisque logs pas FHS-compliant)

- **ressources et tâches du sysadmin de grille**
- **mailing-listes**
 - lcg-rollout (archives publiques)
 - Operations france-grille
 - autres
 - par VO (communauté précise)
- **“known issues” pages pour chaque update gLite**
- **Gocwiki: page des erreurs connues**
- **Twiki du CERN: troubleshooting**

- **COD operators**
 - Rotation sur les ROCs pour fournir une équipe d'opérateurs de surveillance (CIC On Duty)
- **User support : GGUS**
 - <https://gus.fzk.de/>
- **Outils de monitoring**
 - CIC portal
 - <http://cic.gridops.org/>
 - Point d'entrée sur les autres outils
 - Vue intégrée des autres outils à l'usage des COD opérateurs
 - Site Functional Tests (SFT) -> Service Availability Monitoring (SAM)
 - <https://lcg-sam.cern.ch:8443/sam/sam.py>
 - Grid Operations Centre Core Database (GOCDB)
 - <https://goc.gridops.org/>
 - GIS monitor (Gstat)
 - <http://goc.grid.sinica.edu.tw/gstat>



The collage displays several screenshots of monitoring and support tools:

- Top Left:** A screenshot of the 'GOC DBIC Data' interface, showing a search bar and a list of data entries.
- Top Right:** A screenshot of the 'GOC OPERATIONAL CENTER' interface, displaying a table with columns for 'SITE', 'DESCRIPTION', 'START DATE', and 'END DATE'. The table lists various sites and their operational status.
- Middle Left:** A screenshot of the 'Site Functional Test report' interface, showing a table with columns for 'Name', 'Status', and 'URL'. It lists various sites and their functional test results.
- Middle Right:** A screenshot of the 'GOC DBIC Data' interface, showing a table with columns for 'Name', 'Status', and 'URL'. It lists various sites and their functional test results.
- Bottom Left:** A screenshot of the 'Gstat' interface, showing a map of Europe with various sites marked and a table of data.
- Bottom Right:** A screenshot of the 'Gstat' interface, showing a map of Europe with various sites marked and a table of data.

- **test régulier depuis différents portails**
- **rapport régulier via le CIC portal**
- **Messages broadcast**
- **gocdb**
- **monitoring spécifique par VO: alice monalisa / atlas et lhcb dashboard**
- **aussi le RTM et grid-observatory**
- **nagios spécialisé grille**
- **snort**
- **monitoring de bande passante**
- **accounting plus poussé: APEL à configurer ou monbox**

- **GGUS: bugtracker dédié géré par le ROC (inscription)**
- **le ROC fait l'interface entre sysadmins, VOs, devs du middleware, coordination globale**
- **Recommandations pour les mises à jour:**
 - dégré d'urgence des mises à jour indiquées: date limite prévue

- **Pre-prod / phase de certification**
- **procédure sur le gocwiki: "How to get your site certified"**
- **Via un contact du ROC: inscription, choix de nom de site. validation par les spécialistes du ROC.**
- **important: responsabilité et indépendance du site**
 - Sinon blacklisting en 2 jours
- **création d'un nouveau site: nouveau nom**

- **SAMAP → Nagios**
- **Dashboards du ROC**
 - Dashboards spécifiques aux Vos
- **Outils locaux de vérification du réseau, de la température, etc.**

- **backup:**
 - DB du SE (difficile de sauvegarder toute les données en double)
 - CE: sauvegarder les softs et les tags
 - logs (requis juridiques un an au moins)
- **firewall/filtrage de ports de grille sur JRA1 cvs**
- **optimisation des images, services à désactiver**
- **les WNs peuvent éventuellement être NAT-és derrière le CE**
 - le CE fait donc passerelle (perfs moins bonnes mais utiles quand on n'a pas beaucoup d'adresses)

- **ACLs sur le SE par roles comme décrits dans le VOMS correspondant**
- **nécessité de mettre en place des quotas:**
 - mécanisme des tokens
- **Stop ou decommissioning:**
 - fermer les queues batch, passage en mode drain
- **Broadcast:**
 - Suivi et utilisation

- **Quattor pour déploiement automatisés (quattor.org)**
 - déploiement basé sur des règles compilées "template"
 - generation d'un profile kickstart
 - fait aussi DHCP
 - rapidité mais flexibilité très relative: trop envahissant
 - Alternatives : puppet, cfengine, etc
- **Nodetracker**
 - Voir le nodetracker sur le wiki pour l'évolution des services

- **Formulaire de retours pour formations EGEE**
- **Favoriser la dissémination de la connaissance dans votre institut, contribuer aux discussions, se documenter, participer...**
- **A vous de jouer...**