

# Optimisation énergétique Complexe de calcul du CEA

**Jean-Marc DUCOS, CEA DAM**

[Jean-marc.ducos@cea.fr](mailto:Jean-marc.ducos@cea.fr)

# Complexe de calcul du CEA

## CEA/DIF (Bruyères-Le-Châtel)



# Evolutions de puissances des centres de calcul du CEA

**350 Tflops**



**CCRT**

**2006 : CCRTB**  
 Capacité électrique du bâtiment : env. **1,5 MW**  
 Refroidissement par air : jusqu'à 10 kW / baie

**X 2** ||| **2009 : Extension du CCRTB**  
 Capacité électrique du bâtiment : env. **3 MW**  
 Refroidissement par eau : 25 à 35 kW / baie

**1,25 Pflops**



**TERA**

**2006 : TERA10**  
 Capacité électrique du bâtiment : **3 MW**  
 Refroidissement par air : de 5 à 8 kW / baie

**X 3** ||| **2010 : TERA100**  
 Capacité électrique du bâtiment : **10 MW**  
 Refroidissement par eau : 30 à 40 kW / baie



**Très Grand Centre de Calcul – TGCC**

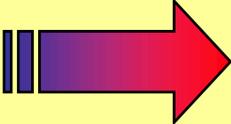
Capacité électrique du bâtiment 2010-2012 : **15 MW**  
 Refroidissement : 30 à 40 kW / baie

**> 1 Pflops en 2011**

**Il devient indispensable de maîtriser et d'optimiser la consommation électrique des centres de calcul**

## MAITRISE ET OPTIMISATION ENERGETIQUE

### EXEMPLE RETOUR D'EXPERIENCE

**TERA10**  **TERA100**

- **Configuration des infrastructures de TERA10**
  - Electricité
  - Production frigorifique
  - Traitement d'air
  
- **PUE de TERA10**
  - Mesures
  - Répartition
  
- **Optimisations énergétiques pour TERA100**
  - Généralités
  - Optimisation des équipements informatiques
  - Optimisation de l'alimentation électrique
  - Optimisation du refroidissement
  
- **PUE de TERA100**
  - Premières mesures et gain estimé
  - Comparaison avec TERA10
  
- **Prise en compte des améliorations de TERA100 pour le TGCC**

## Configuration alimentation électrique TERA10

- **4 Transformateurs de 2 MVA (15 kV / 400 V)**

- 1 transformateur secouru par groupe électrogène + 1 en redondance
- 1 transformateur alimenté en énergie normale + 1 en redondance



- **6 Onduleurs avec autonomie de 10 minutes**

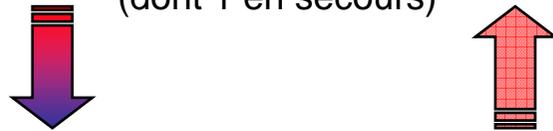
- 5 onduleurs 800 kVA pour la totalité de la configuration informatique (calculateur, stockage, réseau)
- 1 onduleur 400 kVA pour permettre de refroidir les équipements informatiques lors d'une coupure électrique (pompes de distribution)
- Pas de redondance sur les onduleurs (by-pass en cas de problème)



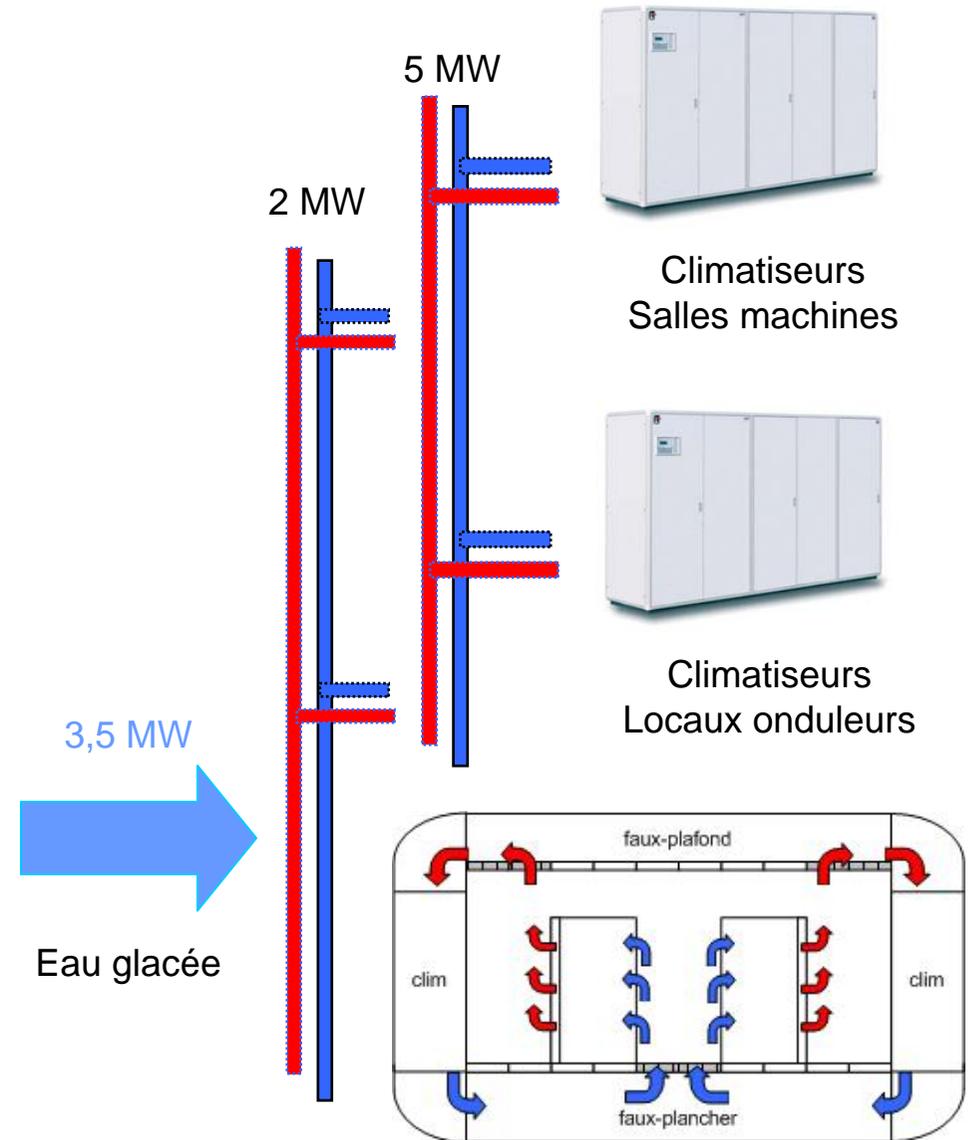
# Configuration refroidissement TERA10 (1/2)



3 tours aéro-réfrigérantes ouvertes  
(dont 1 en secours)



3 groupes froids (dont 1 en secours)



- **100 % de la configuration informatique refroidie par air**
  - Consigne : 22° +/- 3 ° C
  - Mise en place d'allées froides / allées chaudes
  - Soufflage faux-plancher dans les allées froides
  - Reprise en faux-plafond depuis les allées chaudes
  - Efficace jusqu'à 10-15 kW par baie pour une grosse configuration
    - ✉ Difficulté pour gérer les flux d'air
    - ✉ Problème ponctuel de point chaud
  
- **Traitement d'air**
  - Apport d'air neuf (présence ponctuelle de personnel)
  - Taux de renouvellement : 1 volume par heure
  
- **Humidification**
  - Consigne : HR = 50 % +/- 10 %
  - Fonctionne uniquement en hiver

- **Configuration des infrastructures de TERA10**

- Electricité
- Production frigorifique
- Traitement d'air

- **PUE de TERA10**

- Mesures
- Répartition

- **Optimisations énergétiques pour TERA100**

- Généralités
- Optimisation des équipements informatiques
- Optimisation de l'alimentation électrique
- Optimisation du refroidissement

- **PUE de TERA100**

- Premières mesures et gain estimé
- Comparaison avec TERA10

- **Prise en compte des améliorations de TERA100 pour le TGCC**

## L'indicateur PUE de TERA10

### ● Définition

Le PUE (Power Usage Effectivness) est un indicateur qui permet de mesurer l'efficacité énergétique d'un datacenter. Il est calculé en divisant le total de l'énergie consommée par le datacenter par le total de l'énergie utilisée par les équipements informatiques

$$\text{PUE} = \frac{\text{Consommation électrique globale du datacenter}}{\text{Consommation électrique des équipements informatiques}}$$

Actuellement le PUE est la métrique de mesure la plus utilisée dans le domaine de l'efficacité énergétique même s'il ne prend pas en compte tous les axes d'amélioration concernant la réduction de la consommation électrique globale

## PUE de TERA10 : les mesures de consommation électrique

### ● Rappel des 3 niveaux de mesure du PUE

- Niveau 1 (basique)  
Mesures mensuelles/hebdomadaires au niveau des onduleurs
- Niveau 2 (Intermédiaire)  
Mesures quotidiennes au niveau des armoires électriques de distribution secondaires
- Niveau 3 (avancé)  
Mesures en continu au niveau des équipements informatiques

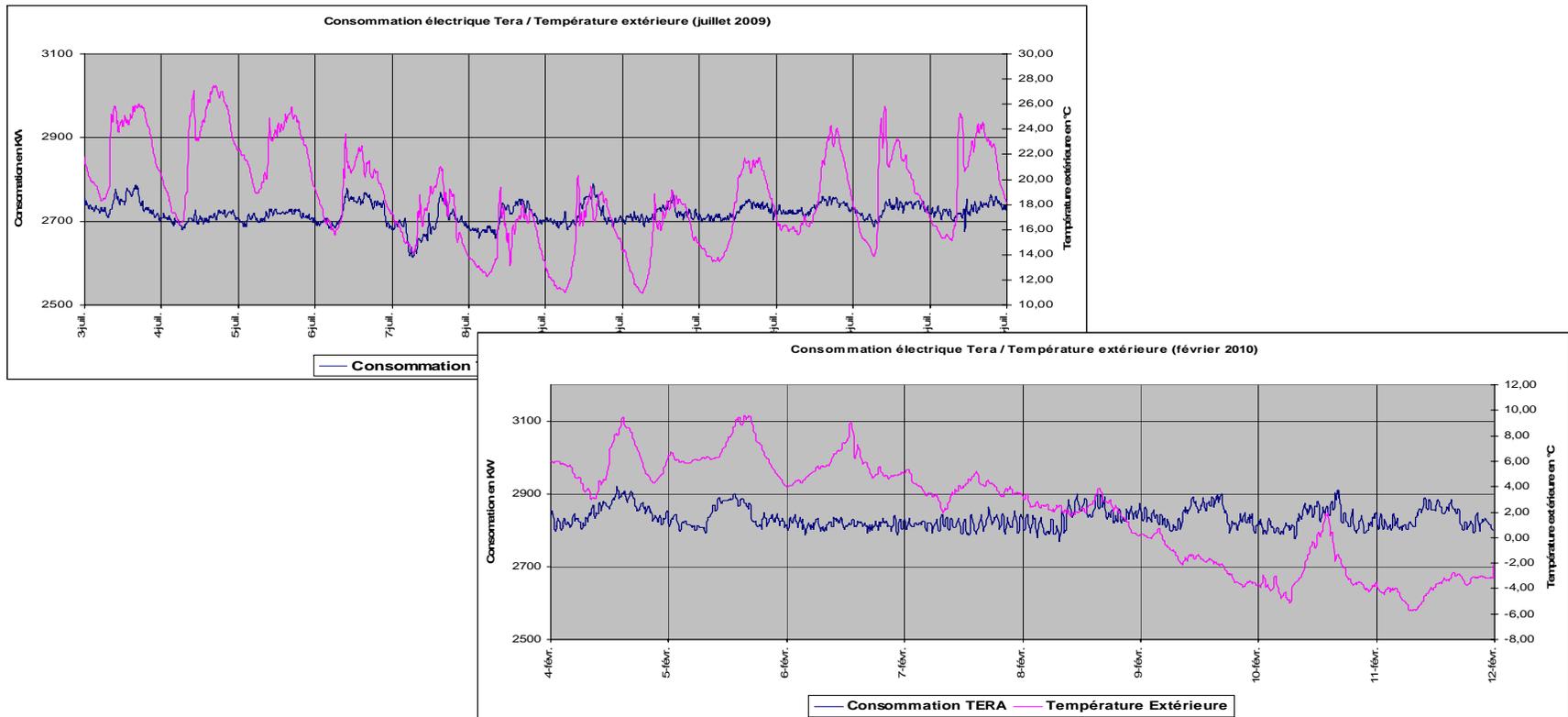
### ● Les mesures de consommation électrique sur TERA10

- Consommation des équipements informatiques
  - ☒ Mesures de niveau 2 sur une période de 5 mois
  - ☒ Relevés des consommations en continu depuis les centrales de mesures des armoires électriques en salle machines
- Consommation totale du bâtiment
  - ☒ Relevés des consommations en continu depuis les transformateurs HT



## PUE de TERA10 : La stabilité de la consommation électrique 1/2

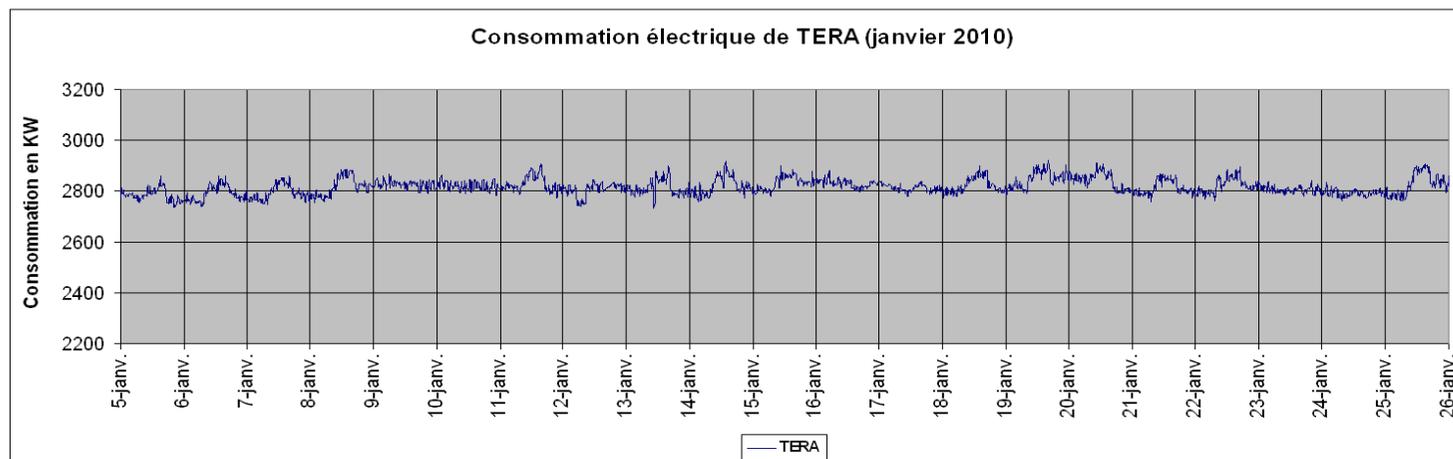
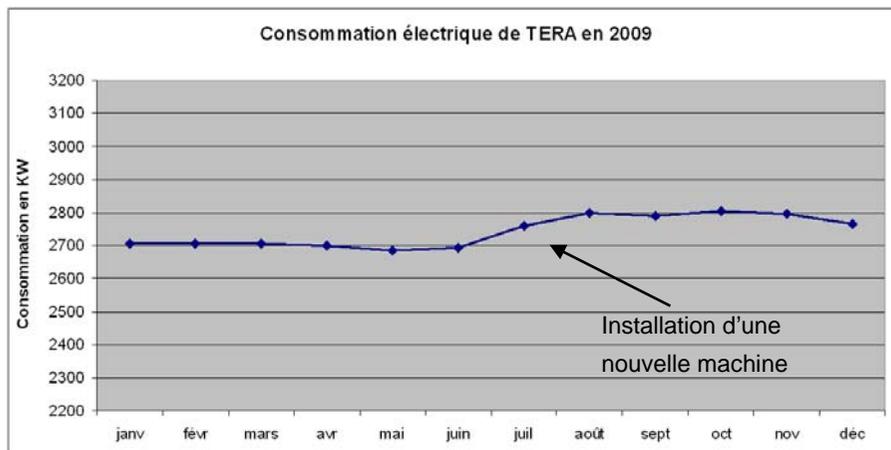
- Consommation électrique en fonction des conditions climatiques
    - Bâtiment enterré : excellente isolation thermique
    - Eté : sur-consommation liée aux tours aéro-réfrigérantes
    - Hiver : sur-consommation liée aux humidificateurs
- Très bonne stabilité de la consommation électrique



## PUE de TERA10 : La stabilité de la consommation électrique 2/2

- Très grande stabilité de la puissance électrique consommée tout au long de l'année due à une consommation constante des processeurs ITANIUM (très peu de variation en fonction de la charge de calcul)

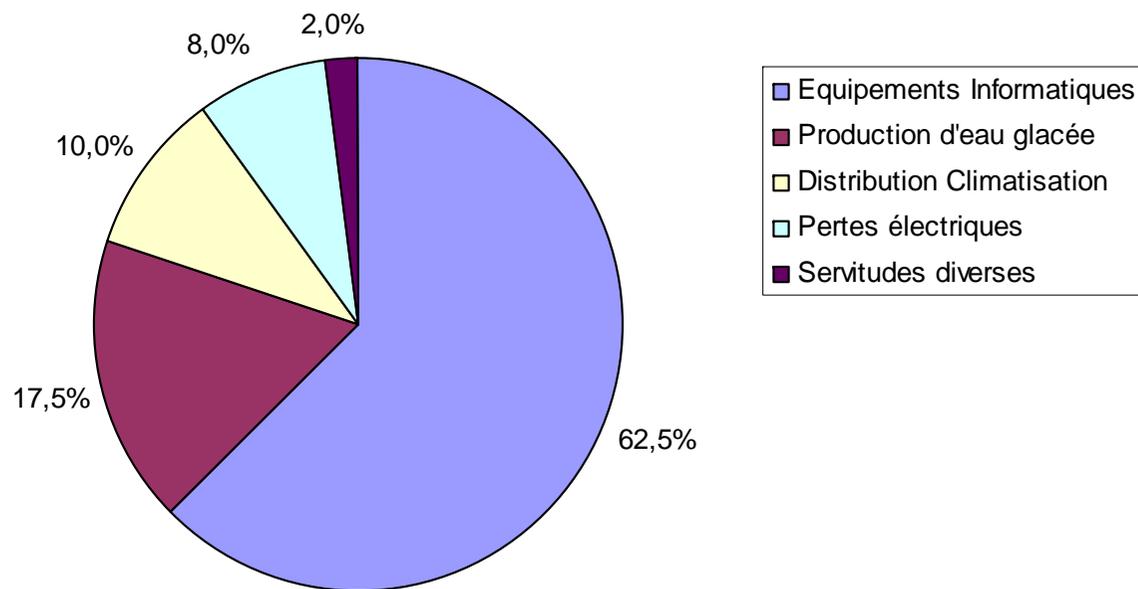
→ PUE annuel équivalent au PUE mesuré sur 5 mois



## PUE de TERA10 : répartition des consommations électriques 1/2

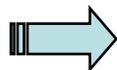
- PUE TERA10 = 1,6
  - Consommation totale du bâtiment 2,8 MW
  - Consommation de l'informatique 1,75 MW
  - Consommation des servitudes 1,05 MW

Répartition des consommations électriques



### ● Description des différents postes de répartition

- Equipements informatiques (62,5% = 1 750 kW)  
L'ensemble des équipements informatiques : Calcul, service, réseau/serveur et stockage
- Production d'eau glacée (17,5% = 490 kW)  
Les groupes froids et les tours aéro-réfrigérantes
- Distribution de la climatisation (10% = 280 kW)  
Les unités de climatisation et les pompes de distribution d'eau glacée
- Pertes électriques (8% = 225 kW)  
Pertes électriques dues au rendement des onduleurs et des transformateurs HT
- Servitudes diverses (2% = 55 kW)  
L'ensemble des postes dont la consommation unitaire est inférieure à 1 %  
Traitement d'air neuf, éclairage, pompes de relevage, systèmes de sécurité, ...



**Les besoins énergétiques étant de plus en plus importants, il devient indispensable d'optimiser les consommations électriques sur des configurations informatiques comme TERA100**

- **Configuration des infrastructures de TERA10**

- Electricité
- Production frigorifique
- Traitement d'air

- **PUE de TERA10**

- Mesures
- Répartition

- **Optimisations énergétiques pour TERA100**

- Généralités
- Optimisation des équipements informatiques
- Optimisation de l'alimentation électrique
- Optimisation du refroidissement

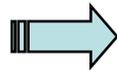
- **PUE de TERA100**

- Premières mesures et gain estimé
- Comparaison avec TERA10

- **Prise en compte des améliorations de TERA100 pour le TGCC**

- **Généralités**

Pour un ordinateur pétaflopique, les investissements en équipements techniques et leur coût de fonctionnement sur 5 ans peuvent représenter plus de 50% du coût d'acquisition du matériel informatique



**Nécessité d'optimiser le TCO (Total Cost Ownership) de l'installation en concevant le bâtiment dans l'optique de réduire les coûts de fonctionnement**

- **Axes de réduction du TCO**

- En réduisant la consommation des équipements informatiques (Cette réduction peut accroître le PUE mais réduira la facture électrique)
- En réduisant la consommation électrique des équipements techniques
  - ⊗ Réduction de l'indicateur PUE
  - ⊗ Objectif TERA100 : PUE < 1,5 au lieu de 1,6 sur TERA10
- En optimisant la consommation de fluides sur l'installation

## Amélioration de l'efficacité énergétique sur TERA100

- **Au niveau des équipements informatiques**
  - Choix de certains composants en fonction de leur efficacité énergétique : alimentations électriques, ventilateurs,...
  - Utilisation de processeurs (XEON Nehalem) de dernières générations
  - Optimisation de la gestion des jobs (scheduling, veille, ...)
  
- **Au niveau des équipements techniques**
  - Réduction du nombre d'onduleurs (uniquement pour équipements sensibles)
  - Utilisation du refroidissement par eau
  - Réduction du besoin en climatisation dans les locaux onduleurs
  
- **Solutions explorées**
  - Augmenter la température en salle machines (tests à réaliser sur TERA100)
  - Augmenter la plage d'hygrométrie en salle (tests à réaliser sur TERA100)
  - Free cooling : Non retenue sur TERA100
    - ✉ Densité par rack trop importante (env 40 kW)
    - ✉ Solution à intégrer à la conception du bâtiment
    - ✉ Investissement important

# Optimisation énergétique des composants informatiques

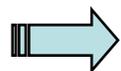


## TERA10 : 60 Tflops

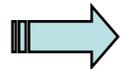
Puissance machines : 1,7 MW  
 Dissipation par rack : 5 à 8 kW par rack  
 Dissipation au m<sup>2</sup> : env. 2 kW / m<sup>2</sup>

## TERA100 : 1250 Tflops

Puissance machines : 5,5 MW  
 Dissipation par rack : env. 40 kW par rack  
 Dissipation au m<sup>2</sup> : env. 8 kW / m<sup>2</sup>



Puissance de calcul multipliée par 20



Consommation électrique des équipements informatiques multipliée par 3

**1 Tflops TERA100 consomme 7 fois moins que 1 Tflops TERA10**

### ● Réduction du nombre d'onduleurs

#### ■ Alimentation en énergie ondulée

- ✉ 20 % des équipements informatiques (disques, équipements réseaux)
- ✉ Pompes de distribution d'eau glacée
- ✉ Puissance : 1,5 MW



#### ■ Alimentation en énergie normale sans onduleur

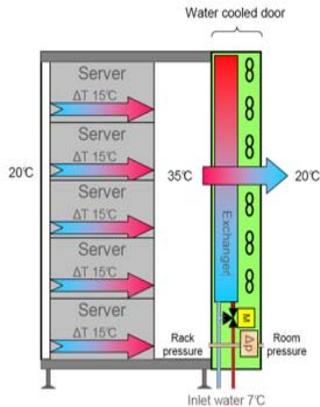
- ✉ 80 % des équipements informatiques (nœuds de calcul)
- ✉ Ces serveurs sont équipées d'ultracapacités permettant de palier aux microcoupures inférieures à 300 ms
- ✉ Puissance : 4,5 MW



#### ■ Gain énergétique : env. 450 kW par an

- ✉ Rendement onduleurs env. 90% (en fonction de la charge)
- ✉ Soit une économie d'env. 250 k€ par an

- **Adaptation du système de refroidissement**

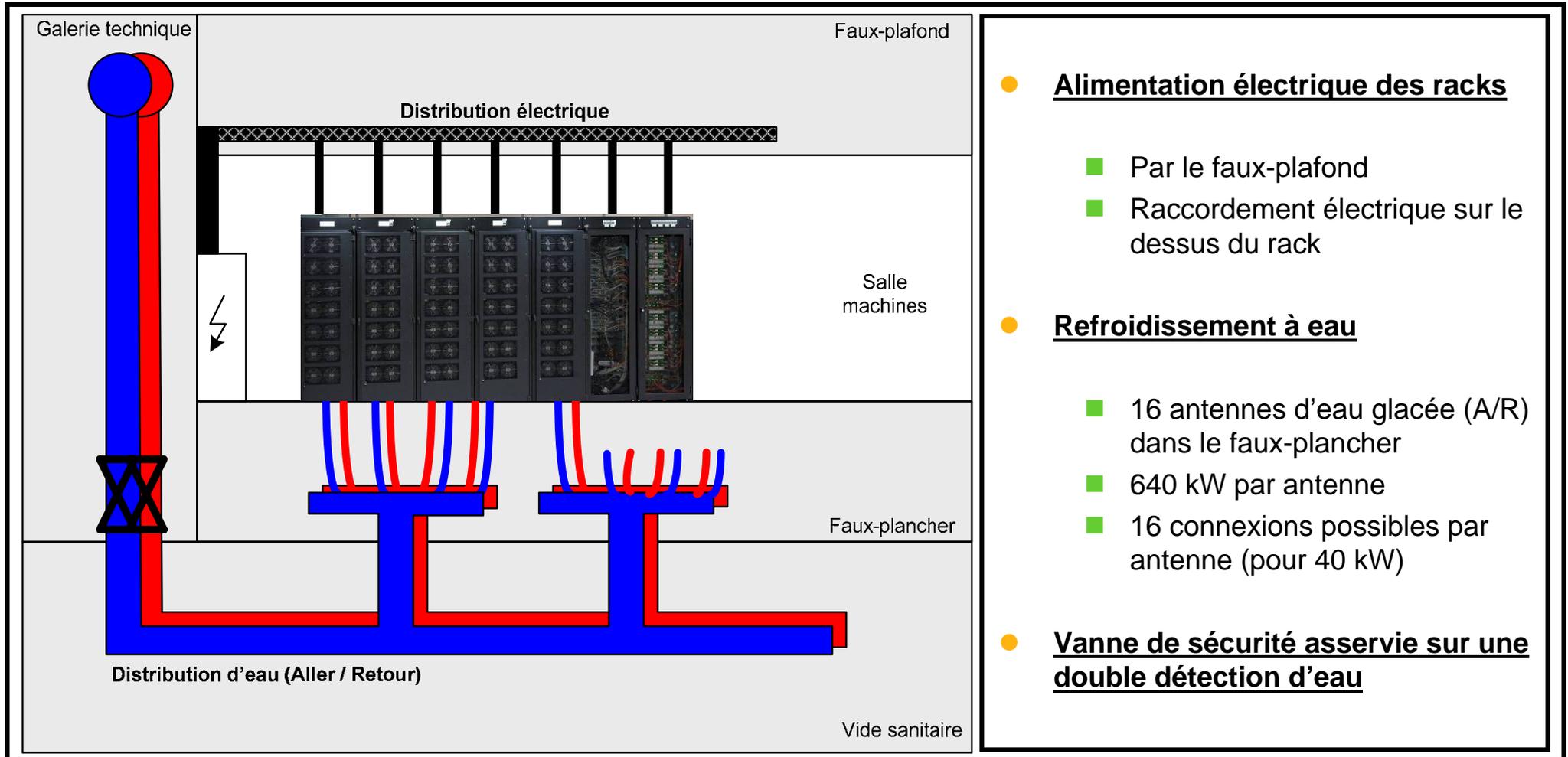


- Augmentation de puissance – Ajout de 2 groupes froids de 2 MW
  - ☒ Configuration : 4 groupes froids + 1 secours
  - ☒ Capacité : 7,5 MW
- Augmentation de la densité par rack – Utilisation du refroidissement par eau
  - ☒ Porte froide intégrée derrière le rack : solution peu encombrante
  - ☒ Permet de dissiper jusqu'à 40 kW par rack
  - ☒ Refroidissement direct des composants informatique
  - ☒ Attente d'une amélioration énergétique significative

- **Configuration du refroidissement pour TERA100**

- Refroidissement par eau avec échangeur intégré au rack (porte froide)
  - ☒ Pour 80 % de la consommation électrique des équipements informatiques (nœuds de calcul) – 170 baies
- Refroidissement par air avec climatiseurs
  - ☒ Pour les équipements informatiques moins denses (disques et réseaux) – 110 baies
  - ☒ Pour les locaux onduleurs





- **Configuration des infrastructures de TERA10**

- Electricité
- Production frigorifique
- Traitement d'air

- **PUE de TERA10**

- Mesures
- Répartition

- **Optimisations énergétiques pour TERA100**

- Généralités
- Optimisation des équipements informatiques
- Optimisation de l'alimentation électrique
- Optimisation du refroidissement

- **PUE de TERA100**

- Premières mesures et gain estimé
- Comparaison avec TERA10

- **Prise en compte des améliorations de TERA100 pour le TGCC**

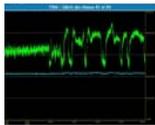
## PUE de TERA100 : premières mesures et gain estimé

- Premières mesures du PUE de TERA100
  - Mesures réalisées lors de la mise au point de TERA100 avec TERA10 toujours en production
  - Puissance IT consommée
    - ☒ 1/3 pour TERA10
    - ☒ 2/3 pour TERA100

→ PUE mesuré = **1,39**

Cette première valeur est un bon indicateur qui nous permet de cibler un **PUE de 1,4** sur l'installation finale TERA100 en production

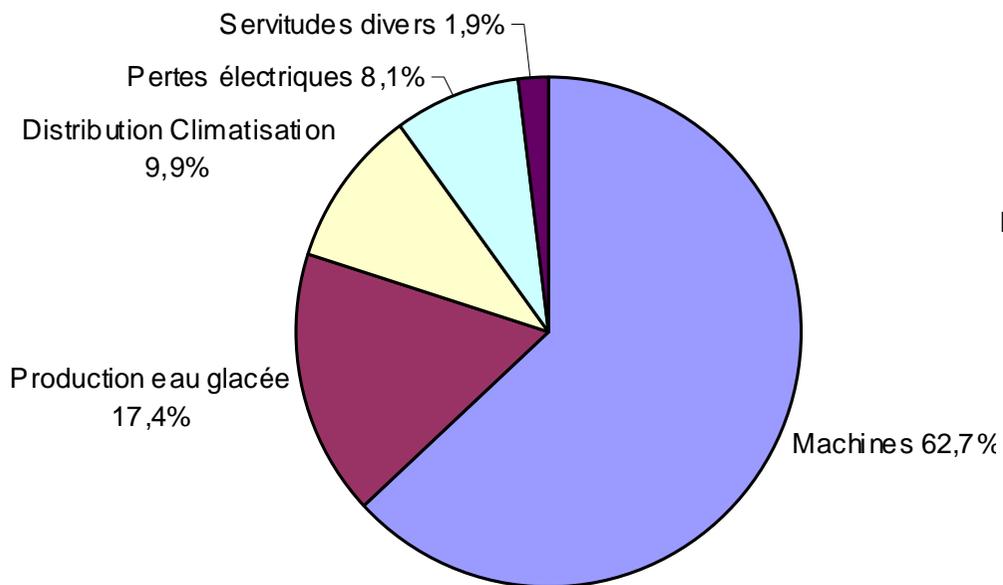
- Gains envisagés avec une puissance machine TERA100 estimée à 5,5 MW
  - Si PUE=1,6 (équivalent TERA10) → consommation totale = 8,8 MW
  - Si PUE=1,4 (estimation TERA100) → consommation totale = 7,7 MW
  - Gain d'environ **1MW** sur la puissance électrique de l'installation, soit une économie annuelle d'environ **650 000 Euros**  
(Coût moyen de l'électricité au CEA en 2010 = env 75 € / MWh)



- Le PUE final de TERA100 sera impacté par une bonne optimisation de la production de froid qui varie en fonction de la charge du calculateur

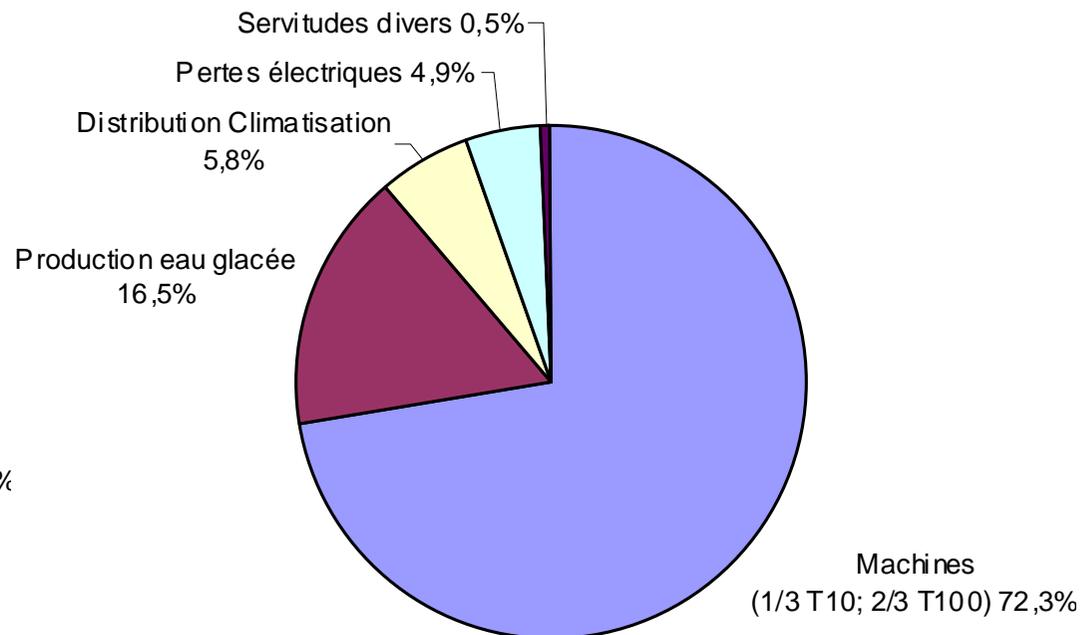
# PUE de TERA100 : Comparaison avec TERA10

**PUE TERA10 = 1.6**



**PUE TERA10/100 = 1.4**

(Mesures sur 4 mois)

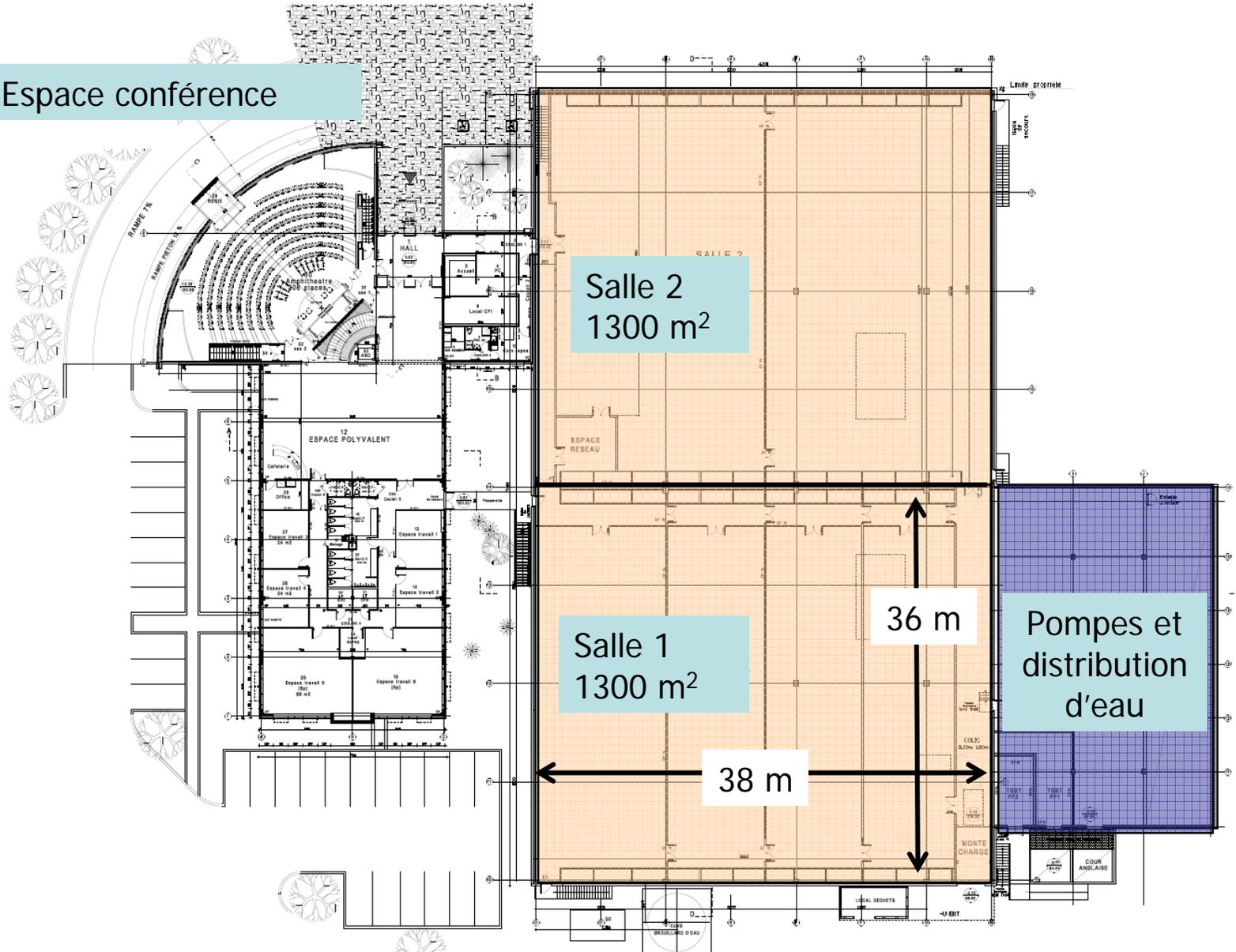


- **Configuration des infrastructures de TERA10**
  - Electricité
  - Production frigorifique
  - Traitement d'air
  
- **PUE de TERA10**
  - Mesures
  - Répartition
  
- **Optimisations énergétiques pour TERA100**
  - Généralités
  - Optimisation des équipements informatiques
  - Optimisation de l'alimentation électrique
  - Optimisation du refroidissement
  
- **PUE de TERA100**
  - Premières mesures et gain estimé
  - Comparaison avec TERA10

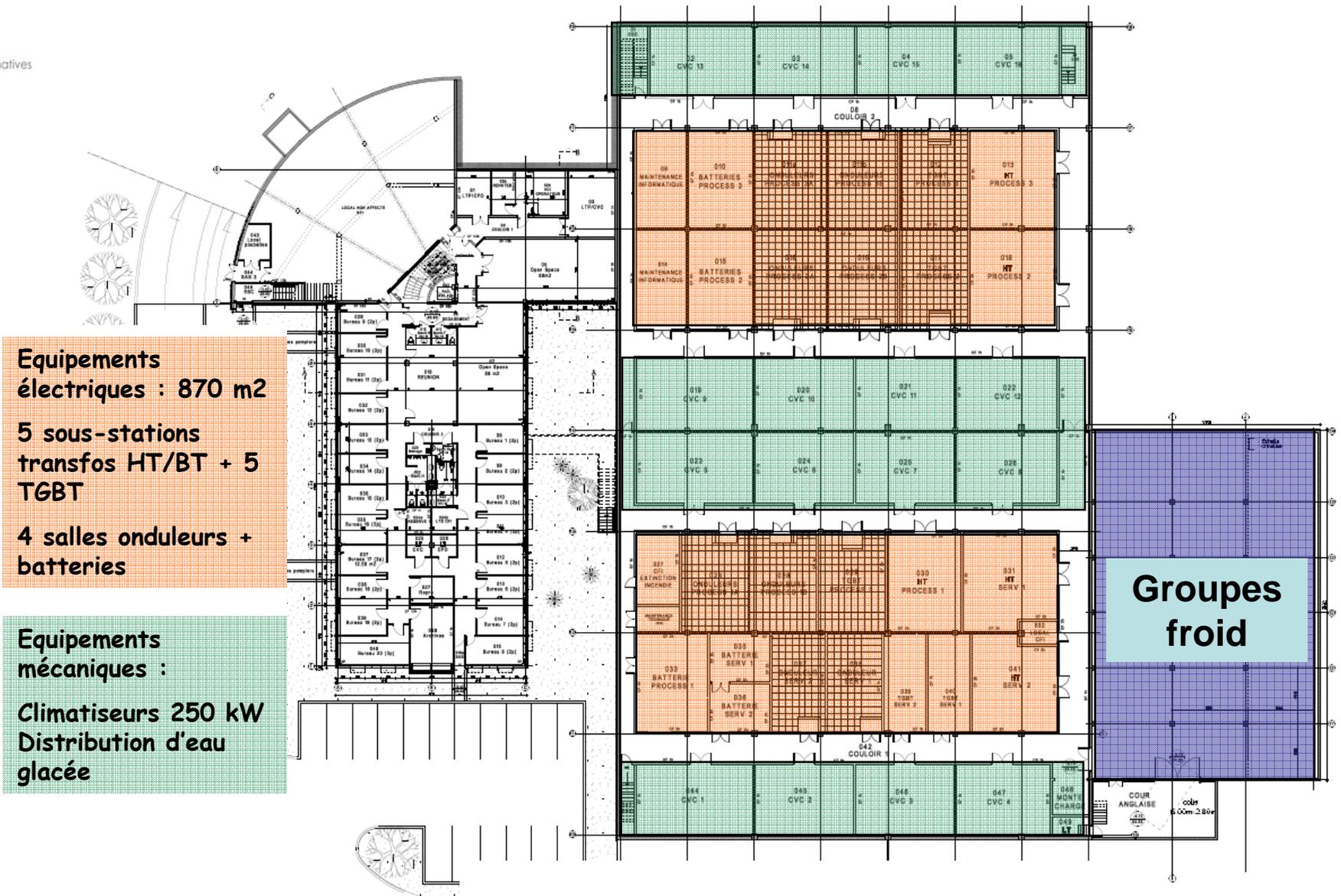
- **Prise en compte des améliorations de TERA100 pour le TGCC**

# RDC du TGCC – Salles machines

Espace conférence

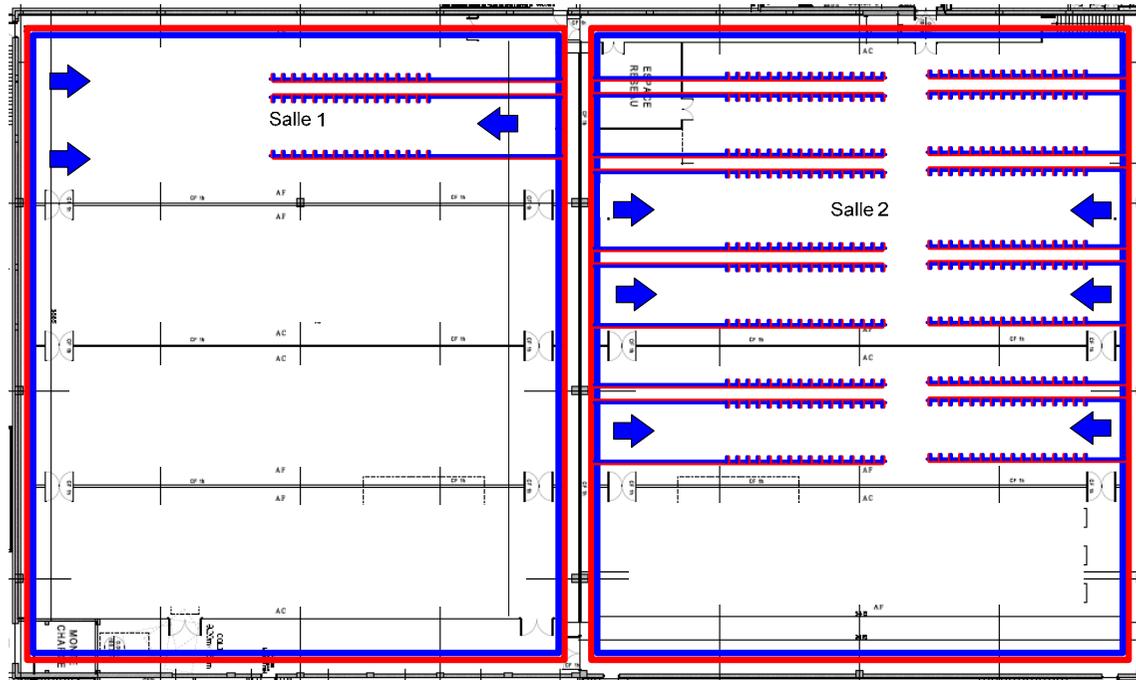


# S/S du TGCC – Equipements électriques et mécaniques



**Equipements électriques : 870 m<sup>2</sup>**  
**5 sous-stations transfos HT/BT + 5 TGBT**  
**4 salles onduleurs + batteries**

**Equipements mécaniques :**  
**Climatiseurs 250 kW**  
**Distribution d'eau glacée**



### Comme pour TERA100

- Limitation du nombre d'onduleurs
- Utilisation du refroidissement par eau

#### Electricité

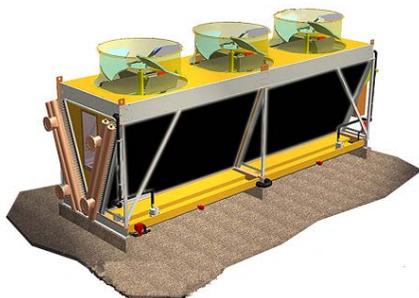
- Armoires électriques ondulées pour les équipements sensibles (20 %)
- Armoire électrique normale pour les nœuds de calcul (80 %)

#### Refroidissement

- Une boucle d'eau de 8 MW par salle machine
- Climatiseurs (+ redondance) pour 20 % des équipements informatiques
- Refroidissement par eau pour 80 % de la configuration (nœuds de calcul) :
  - 23 antennes d'eau glacée (A/R) dans les faux-plancher  
Capacité : 640 kW par antenne
  - Jusqu'à 16 piquages par antenne pour la connexion des échangeurs

#### Sécurité d'eau

- Détection d'eau dans le faux plancher, sous les baies avec échangeur
- Mise en sécurité des antennes d'eau glacée en cas de fuite



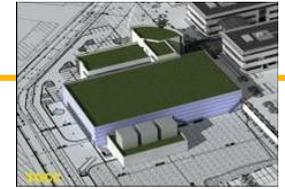
### ● Utilisation de tours de refroidissement hybrides air / eau

- Fonctionnement par air : 70 % de l'année
- Fonctionnement par eau : 30 % de l'année
- Réduction de la consommation d'eau : gain de 150 000 m<sup>3</sup> par an
- Réduction des coûts de fonctionnement : économie de 400 k€ par an  
Mais investissement initial plus élevé (+ 1,2 M€ pour 7,5 MW calculateurs)
- Moins d'émission de vapeur – réduction du risque de légionellose

### ● Récupération de chaleur pour chauffer les bâtiments



- Récupération à partir des condenseurs des groupes froids
- Nécessité d'adapter la température d'eau par rapport au système de chauffage
- Retour sur investissement sur une période relativement longue

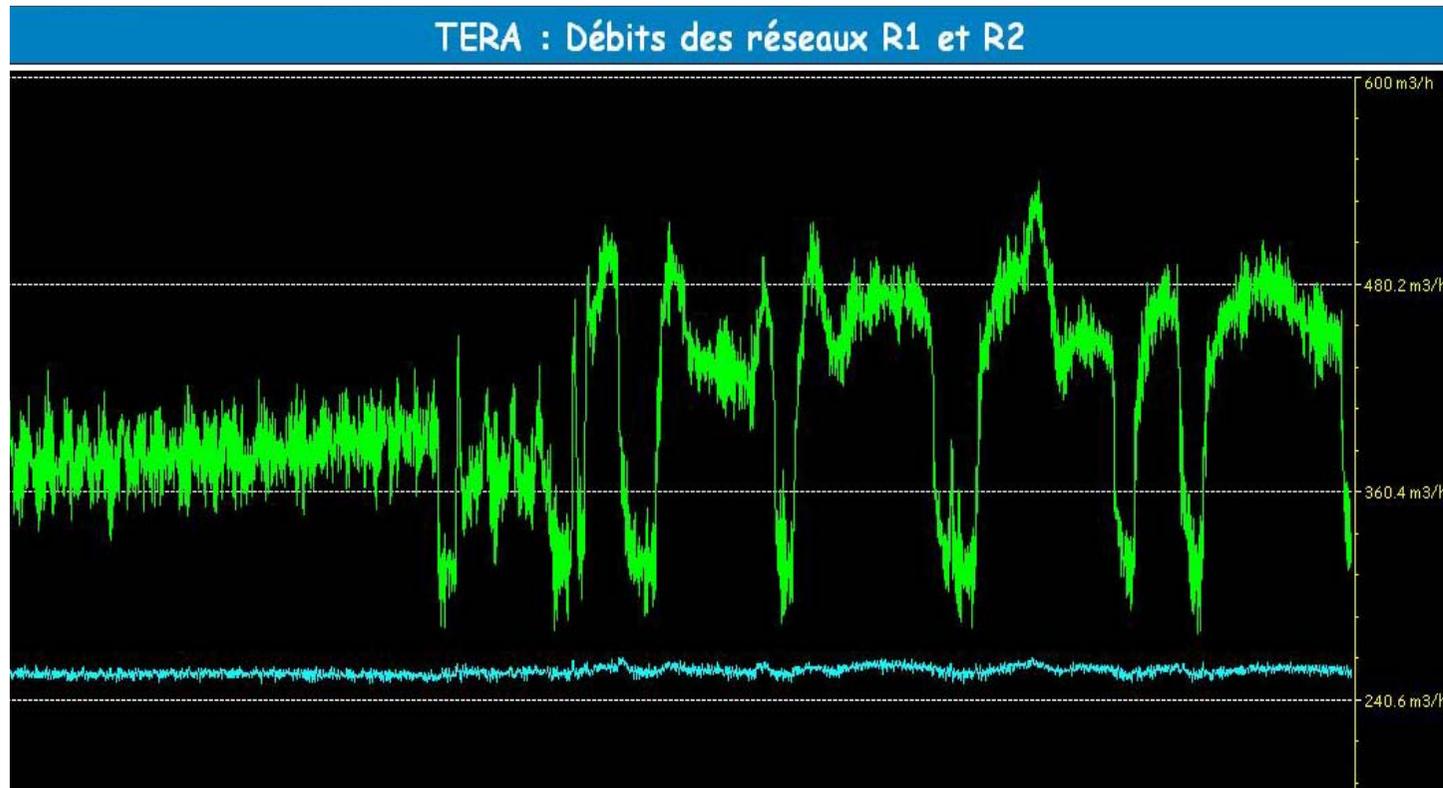


**Merci pour votre attention**

**Avez-vous des questions ?**



## Variation des débits dans les réseaux d'eau glacée (16/08/10)



- Optimisation de la production de froid délicate
  - Variation brusque de 1,5 à 2 MW sur la puissance froid (Passage de Linpack à Idle)
  - Variation d'un facteur 2 sur le débit d'eau glacée