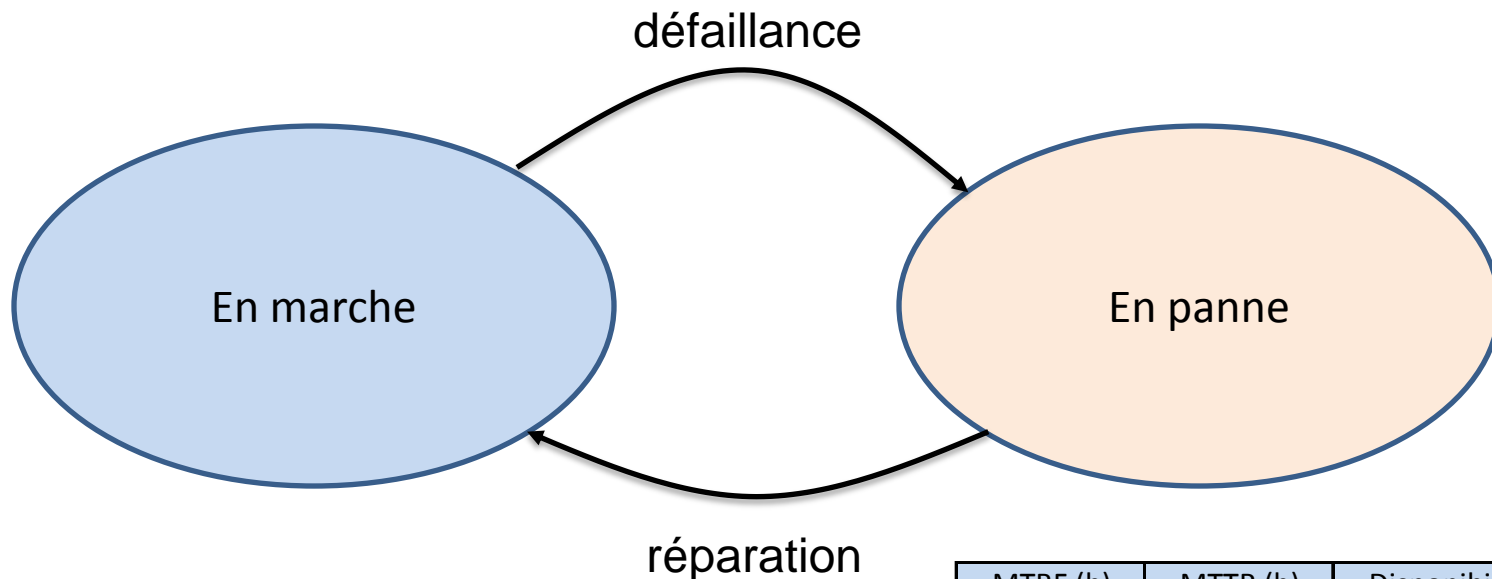




Cloud et disponibilité : est-ce compatible ?

E.Besluau

Disponibilité : deux facteurs

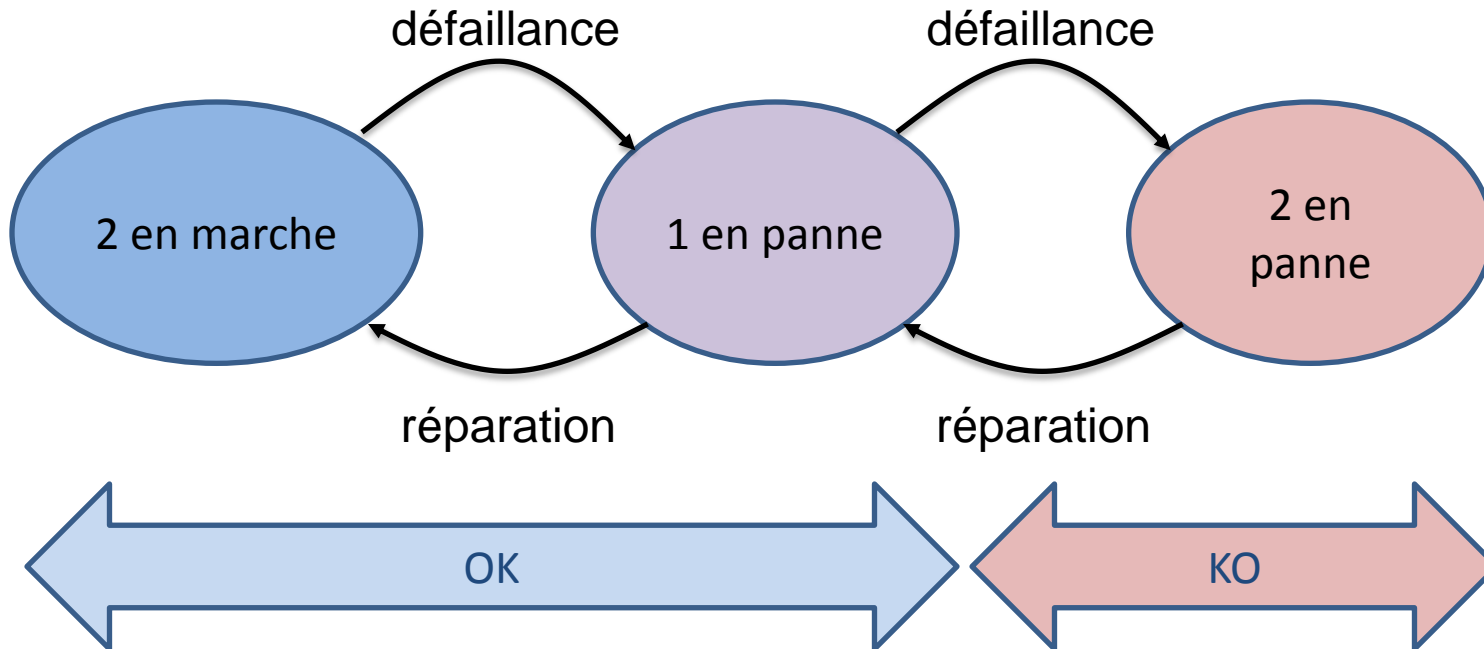


$$\text{Disponibilité} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}$$

MTBF (h)	MTTR (h)	Disponibilité
10 000	1	99,99%
50 000	5	99,99%
100 000	10	99,99%
200 000	20	99,99%
500 000	50	99,99%
500 000	5	99,999%

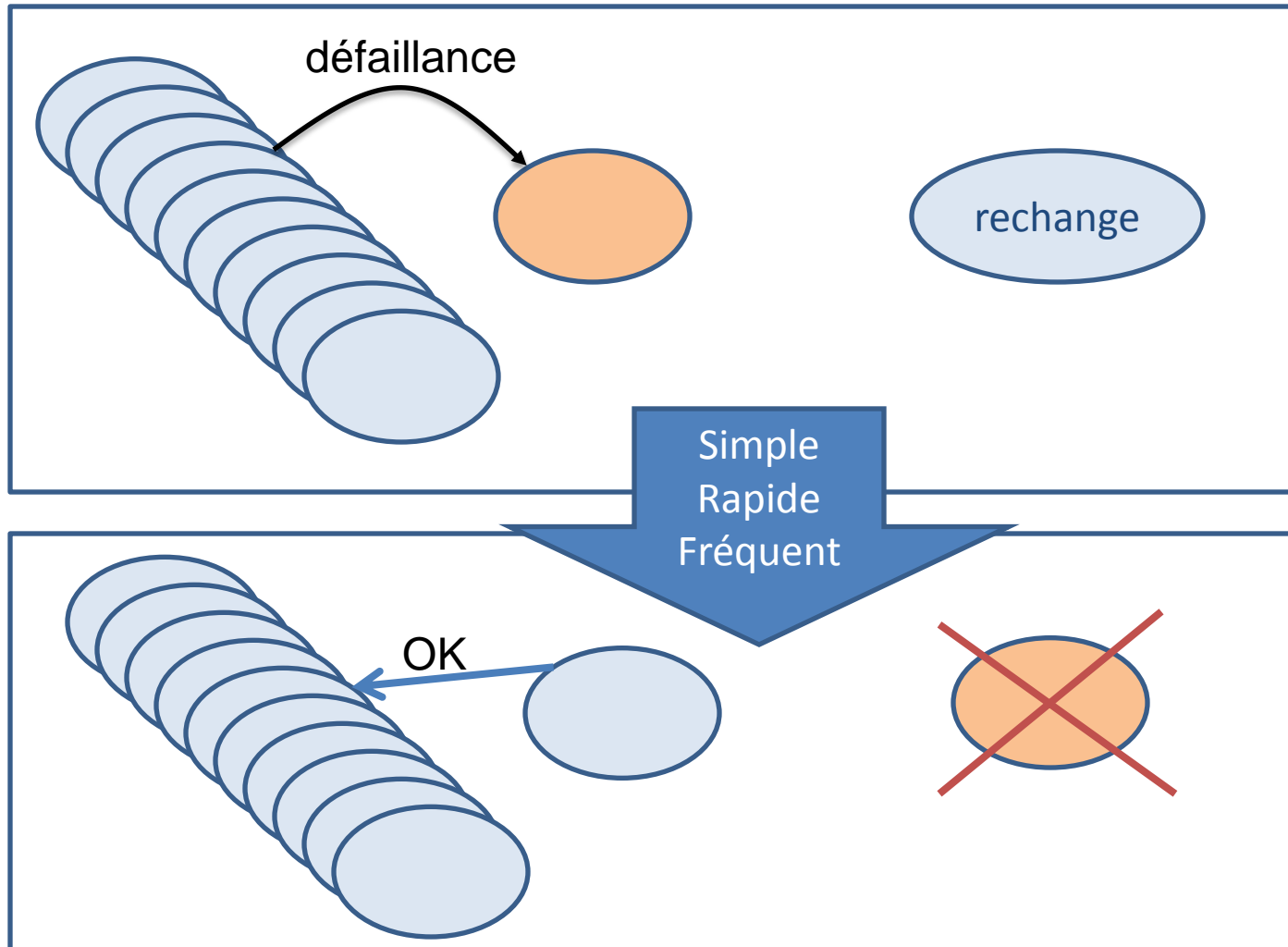
Avoir un système fiable ou facile à réparer ?

Modèle « résilient »



Risque : la « panne de mode commun »

Modèle « réparable » : le cloud



Ne touche qu'une portion du service et des usagers

Caractéristiques du cloud (1/3)

- Des moyens techniques en grosses quantités
 - 10 000 ... 100 000 serveurs ? et plus !
 - similaires en comportement
 - ‘state-less computation’ & ‘stateful storage’
- Conséquences
 - Taux de panne élevés
 - Mais : interchangeables rapidement
 - La réparation est prévue dans les procédures

Caractéristiques du cloud (2/3)

- La situation des utilisateurs est variable :
 - Profils de charge très variés
 - Taux d'utilisation changeant en cours de journée
 - Affectation sur les moyens aléatoire
- Conséquences
 - L'utilisateur peut 'échapper à la panne'
 - Les pannes affaiblissent le service sans l'arrêter
 - La panne perçue ne dure pas longtemps

Caractéristiques du cloud (3/3)

- Des outils utilisés par le prestataire :
 - Pour équilibrer les ressources
 - Pour virtualiser à plusieurs niveaux
 - Pour permettre le provisionnement par l'utilisateur
 - Qui les développe et maintient ?
- Conséquences
 - Sources de pannes de mode commun !
 - Nécessite une attention forte du prestataire
 - Retour arrière à prévoir (culture du bêta)

Le « réseau »

- Sépare l'utilisateur du prestataire
- N'est souvent pas couvert dans les SLA
- Question : Internet plus ou moins fiable qu'un « réseau privé » ?
 - Le pour : maillé, connexions multiples, points faibles résilients (serveur DNS, etc.)
 - Le contre : qui s'en occupe ? Vulnérable aux surcharges, attaques (dDoS), ruptures...
- Cet aspect est équilibré et neutre dans les choix ?

Quelques chiffres (1/2)

- Cas 1 : SLA « 99,95% sur une région »
- Interprétation et illustration :
 - $I = 5 \cdot 10^{-4}$ en utilisant 2 sites équipés
- Durée indisponible par an = 4 h 22 mn
 - Sur 1 site et son infrastructure IT : $I = 2,2 \%$
 - Indisponibilité sur un site par an : 192 h 43 mn
 - Attention : effet 24x24 / 7x7
- *Par exemple 8 arrêts ou pannes (de 24 h sur le site 1) réparées par une bascule en 30 mn vers le site 2*
- *Mais : les données doivent pouvoir suivre*

Quelques chiffres (2/2)

- Cas 2 : SLA « 99,999% d'indisponibilité »
- Interprétation et illustration:
 - Arrêt 5mn/an : très difficile sur un seul site complet !
 - Sur deux sites donc : $I = 3,2 \cdot 10^{-3}$ par site soit 28 h/an
 - *Par exemple 7 arrêts ou pannes (de 4 h sur le site 1), avec reprise en moins d'une minute sur le site 2*
 - *Cela suppose une infrastructure technique adaptée*
 - *Les applications doivent y avoir été préparées*
- Ce cas pré-suppose une ingénierie sophistiquée, technique ou financière

Bilan

- Le cloud -par construction- propose un modèle « réparable facilement »
- La disponibilité perçue va dépendre du « volume consommé » (durée, nombre de plateformes,...)
- Le prestataire maîtrise les affectations...et peut aménager sa disponibilité perçue
- Certains usages et applications vont se prêter à une bonne disponibilité (SaaS ?), d'autres pas (IaaS, PaaS ?)
- 99,95% sur deux sites suffit très souvent ! S'il faut mieux : réfléchir au choix !
- Et puis : le cloud offre aussi des solutions pour faciliter et améliorer votre disponibilité

Merci de votre attention

eb@duquesnegroup.com