

Le calcul de la disponibilité des services (SLA)



10 décembre 2013



Directeur de division Martin Bech, martin.bech@deic.dk

La question : quel taux de disponibilité pouvons-nous garantir?



- Le temps des services offerts “au mieux” (best effort) est révolu.
- Aujourd’hui, nos utilisateurs demandent des niveaux garantis de disponibilité.
- Souvent, nos services sont eux-mêmes composés de différents éléments et services venant de fournisseurs et dotés de leurs propres niveaux garantis de disponibilité.
- Dans ces conditions, quel niveau de disponibilité peut-on garantir à nos utilisateurs?



A7 : Site 3

A B C D E F

Formulaire de demande d'un multipoint à multipoint sur RENATER

Notez que la solution technique pour la réalisation du circuit sera décidée par le GIP RENATER en fonction des besoins recensés dans ce formulaire. Veuillez apporter un soin tout particulier pour remplir ce formulaire pour accélérer son traitement. Veuillez envoyer ce formulaire complété à support-reseau@renater.fr

Seules les demande complétées par les responsables techniques des établissements seront prises en compte

Extrémités du circuit

	Nom de l'établissement	Agrément	Contact	Email
4				
5	Site 1			
6	Site 2			

Dates de disponibilité du service

Date de début	Date de fin
24	

Niveau de service requis (SLA)

26	Bande passante	
27	Type d'interfaces envisagées	
28	IPPM (gigue, délais, pertes de paquets...)	

Services transitants sur le circuit

30	IPv4 unicast	
31	IPv4 multicast	
32	IPv6 unicast	
33	IPv6 multicast	
34	MTU nécessaire	
35	Autres (détailler)	

Plan d'adressage

Décrire le type d'adressage

À RENATER, par exemple, il y a aussi une conscience du niveau de disponibilité et d'autres mesures de performance

- **La métrologie active :**
 - Gestion et surveillance proactive du réseau
 - Identification des goulots d'étranglements
 - Fourniture au client de statistiques en liens avec ces besoins
 - Quantification de l'impact de la différenciation de service et optimisation
 - Mise en place et surveillance de Service Level Agreement
 - Tarification en fonction des paramètres du SLA
- **La métrologie passive :**
 - Tarification (quantitative)
 - Sécurité
 - Diagnostic à posteriori
 - Visualisation de l'évolution du réseau sur le long terme

- Un avertissement à propos de cet objectif de niveau de service, oblige TELUS à régler le problème de manière appropriée.

Mesure	Indicateur	Standard	Objectif de niveau de service
Disponibilité du service*	Période pendant laquelle le Service peut être utilisé :	24 heures sur 24, 7 jours sur 7	Moyenne de 99,7 % sur une période de 30 jours.
Mesure du rendement du service*	Réaction du serveur Web ou FTP à une requête HTTP ou FTP de type « GET »	Temps de réponse de moins d'une seconde à compter de l'envoi-test	Moyenne de 99,7 % sur une période de 30 jours.

Le taux de disponibilité – le paramètre de performance le plus important

...calculé comme le ratio entre temps de panne et temps total



3. Service Assurance

3.1 Service Availability



Service Availability is guaranteed at 99.9%. Spin will provide a service fee rebate to customers with unavailability of greater than 40 minutes in a given month.

“Service Availability” is defined as the percentage of time the service is available, via the primary connectivity medium, during the course of a month. The Service Availability is calculated in accordance with the following formula:

$$\text{Service Availability} = \frac{\text{Total minutes for the period minus Unavailable minutes}^*}{\text{Total minutes for the period}} \times 100$$

* Unavailable minutes is the total number of minutes that the service is unavailable due to issues with the Spin network or our carrier's except for programmed outages.

3.2 Fee Rebates Due To Service Unavailability

Le taux de disponibilité est souvent lié à une clause de pénalité



The following table highlights at-a glance the main WAN and Internet services and their corresponding Service level parameters:

SERVICES	SERVICE LEVEL PARAMETER	NETWORK AVAILABILITY
Ethernet Fibre WAN (E10, E100, FastE, etc)	7 x 24 - SLA	99.7%

2.1 99.7% Service Availability - Subject to Section 5 below, if the availability of the Services is less than 99.7%, Red Fox Hosting will issue a credit to customer in accordance with the following schedule, with the credit being calculated on the basis of the monthly service charge for the affected Services:

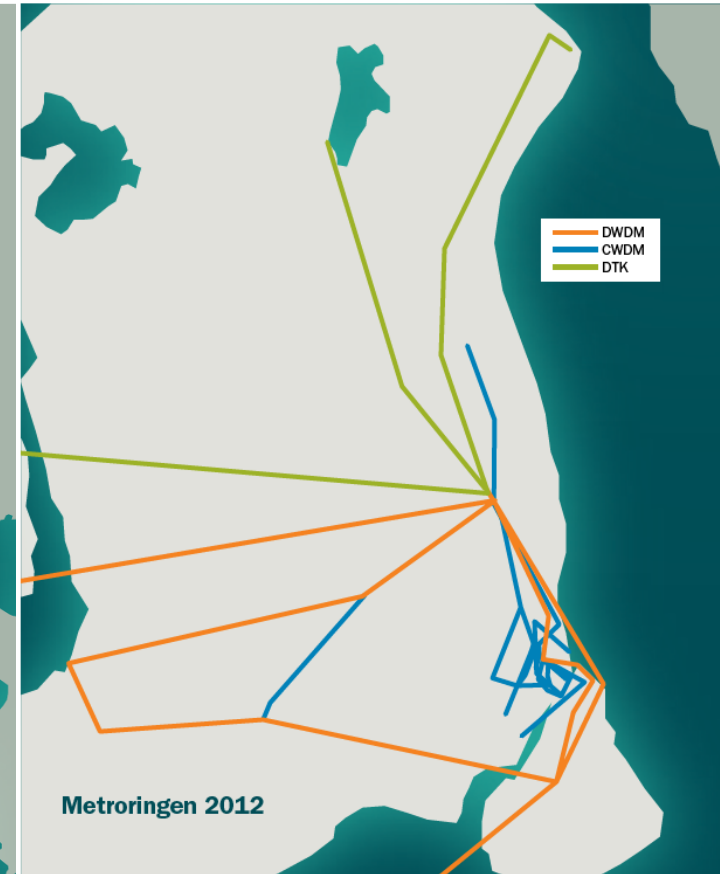
Service Availability	Credit Percentage
99.7 to 100%	0%
98% to 99.7%	10%
95% to 97.9%	20%
90% to 94.9%	30%
85% to 89.9%	40%

99,7% - qu'est-ce que ça signifie exactement?



- Un taux de disponibilité garantie de 99,7% - cela veut-il dire que le service est en panne 0,3% du temps?
- **Non** – on peut espérer qu'il est beaucoup moins en panne que cela!
- Selon nos observations pour les connexions WAN simples, elles sont en panne seulement 0,11% du temps (l'explication suit...)
- Cependant les fournisseurs ont choisi 0,3% pour payer moins souvent des pénalités
- De fait, le service ne respecte pas le niveau garanti que 6,2% de tous les trimestres, soit environ 1 fois tous les 4 ans

Det landsdækkende net 2012



Le réseau de la
recherche au
Danemark

DeiC

Relation entre disponibilité garantie et disponibilité réelle



Pour se faire une idée de cette relation, prenons les connexions WAN du NREN Danois, qui ont les propriétés suivantes :

- Ce sont des connexions WAN simples sans protection ni redondance
- À la fois fibre noire et capacité de transmission
- Nous disposons de nos propres mesures de surveillance
- Ce que nous mesurons est aussi proche de la performance du fournisseur que possible
- Ces connexions ont une disponibilité garantie de 99,7% calculée sur un trimestre (ce qui correspond à une panne de 6,5 heures/trimestre)

Les lignes WAN intéressants pour notre étude sont choisis. (liens simples de 99,7% sur base d'un trimestre)

Host list for Access 8:19:21 Friday November 11 2011.										
Samsu-Iluna System Maintenance->Show Full Log Performance										
Plot Info	Poll time /test result	Hostname /Report	System Log/ Backup Status	Action /Admin	Service	IP /Contract	Name(/ReverseIP) /External page link	Meta Group Access	Dept.	Last 90 days/Report All=Major+Minor
	600 s	fwa-lqb.dmi			RND: UDMC	130.225.245.110	/	Access		0 0 0
100% packet loss(repeat 32282):Status red alert when the second test failed at Mon Mar 28 11:31:02 2011										
	600 s	ge-kvit.or1			RND: UDMC	130.226.249.130	/	Access		2 1 1
100% packet loss(repeat 42810):Status red alert when the second test failed at Thu Jan 13 16:30:23 2011										
	600 s	3com.stat-mus.core			RND: UDMC	130.226.236.1	/	Access		2 2 0
	600 s	Bygning356.core			RND: UDMC	130.225.243.62	/	Access		0 0 0
	600 s	Ge-KVL_Carlsberg			RND: UDMC	130.226.249.14	/	Access		2 2 0
	600 s	NorduNetLy4			AMD LINE	109.105.102.37		Access	RND	1 0 1
	600 s	NorduNetLy6			AMD LINE	2001:948:0:F052::1		Access	RND	0 0 0
	600 s	NorduNetOr4			AMD LINE	109.105.102.41		Access	RND	0 0 0
	600 s	NorduNetOr6			AMD LINE	2001:948:0:F040::1		Access	RND	0 0 0
	600 s	akftest.akf			RND: UDMC	130.225.245.133	/	Access		48 10 38
	600 s	cbs			RND: UDMC	130.226.255.15	/	Access		0 0 0
	600 s	dk-bib.al.core			RND: UDMC	130.226.189.3	/	Access		0 0 0
	600 s	dmi			RND: OSPF	130.226.255.108	/	Access		2 2 0
	600 s	fe-Maskinmesterskole.ly3.core			RND: UDMC	130.225.244.21	/	Access		0 0 0
	600 s	fe-SJFI.KVL			RND: UDMC	130.226.178.1	/	Access		5 4 1
	600 s	fe-cbs.cbs			RND: UDMC	130.226.213.126	/	Access		0 0 0
	600 s	fe-fw.dmi			RND: UDMC	130.226.71.217	/	Access		0 0 0
	600 s	fe-rpc.ly8			RND: UDMC	130.226.161.17	/	Access		0 0 0
	600 s	fe-symphogen.ly3.core			RND: UDMC	130.225.247.65	/	Access		0 0 0
	600 s	fe-telestyrelsen.ly3.core			RND: UDMC	130.226.157.49	/	Access		0 0 0
	600 s	fw.djh.core			RND: UDMC	130.225.231.2	/	Access		0 0 0
	600 s	ge-AAU.UniAR.core			RND: UDMC	130.225.244.70	/	Access		0 0 0
	600 s	ge-Carl.kvlq.core			RND: UDMC	130.226.249.9	/	Access		3 2 1
	600 s	ge-NyEmdrupborg.dpu			RND: UDMC	130.226.161.121	/	Access		0 0 0
	600 s	ge-akf.ly3			RND: UDMC	130.225.245.129	/	Access		0 0 0
	600 s	ge-ar2.iha.core			RND: UDMC	130.226.249.174	/	Access		0 0 0
	600 s	ge-cbs.cbs			RND: UDMC	130.226.249.126	/	Access		2 2 0

Détails de la surveillance



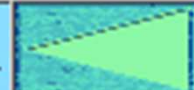
- Fait avec **ping -c 5 -s 10**
- Toutes les dix minutes
- Si la perte de paquets est de 0 et le temps de ping est sous un seuil, la mesure est "verte" (bleue en réalité)
- Si la perte de paquets est de 100%, la mesure est "rouge"
- Toute mesure entre ces deux extrêmes est "jaune"
- Toutes les mesures qui ne sont pas "vertes" sont considérées comme des pannes

Status for

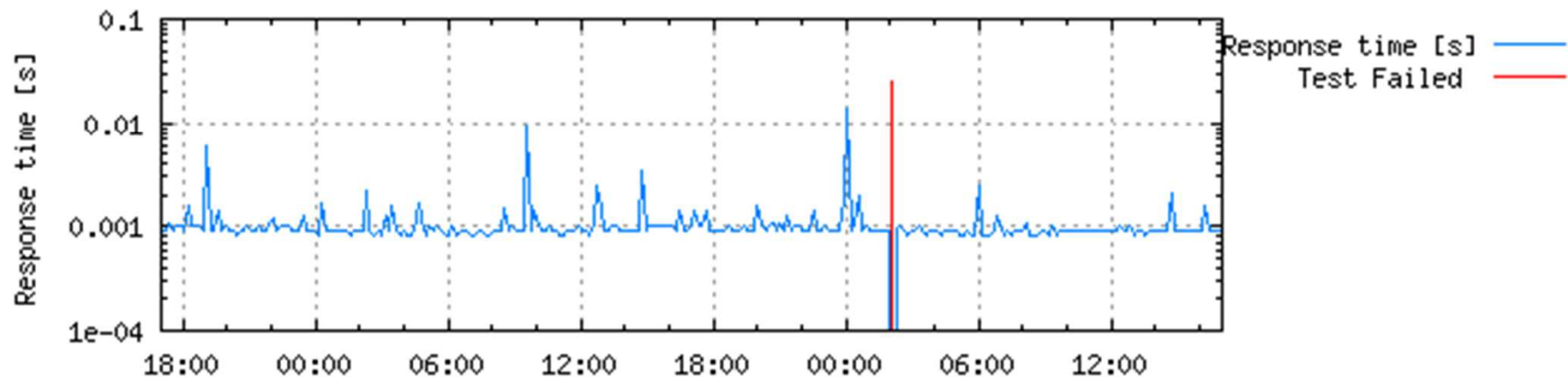
timeframe

[UDMC service.](#)

[Latest result](#)

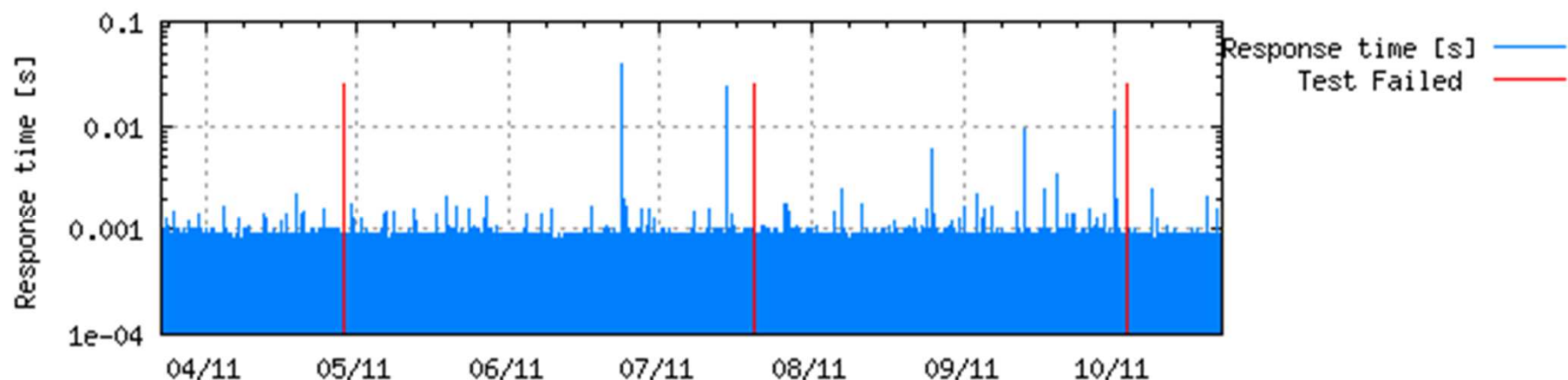


Response time timeframe:08/11/2011-17:00 - 10/11/2011-17:00



ge-NyEmdrupborg.dpu last 2 days, Yellow:Packet loss

Response time timeframe:03/11/2011-17:00 - 10/11/2011-17:00



ge-NyEmdrupborg.dpu last 7 days, Yellow:Packet loss

Response time

2 days & 7 days

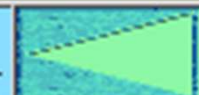
14

Status for

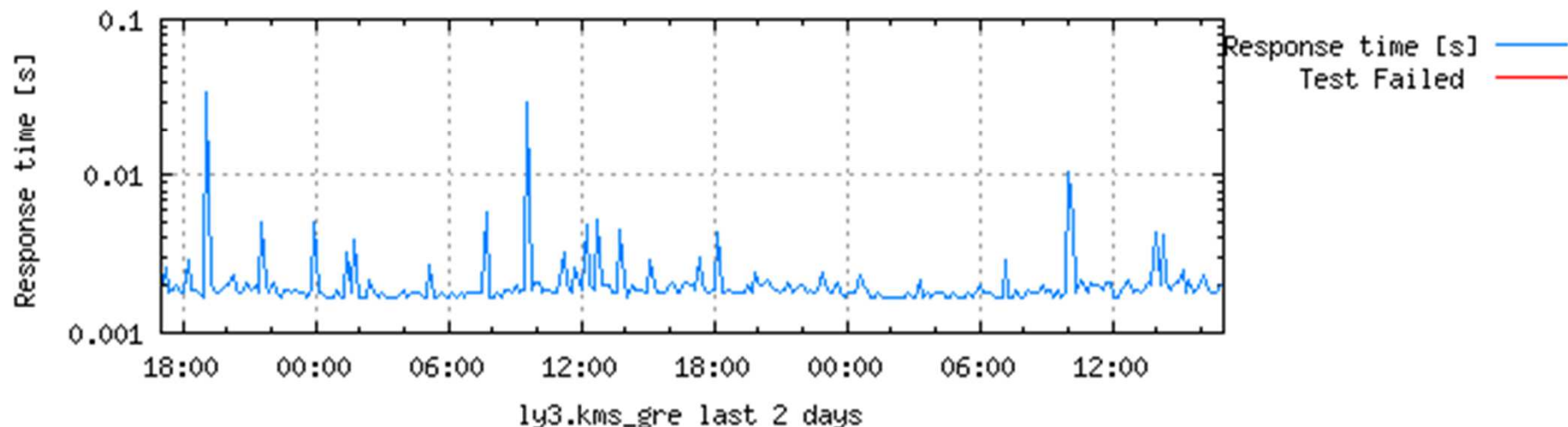
timeframe

[UDMC service.](#)

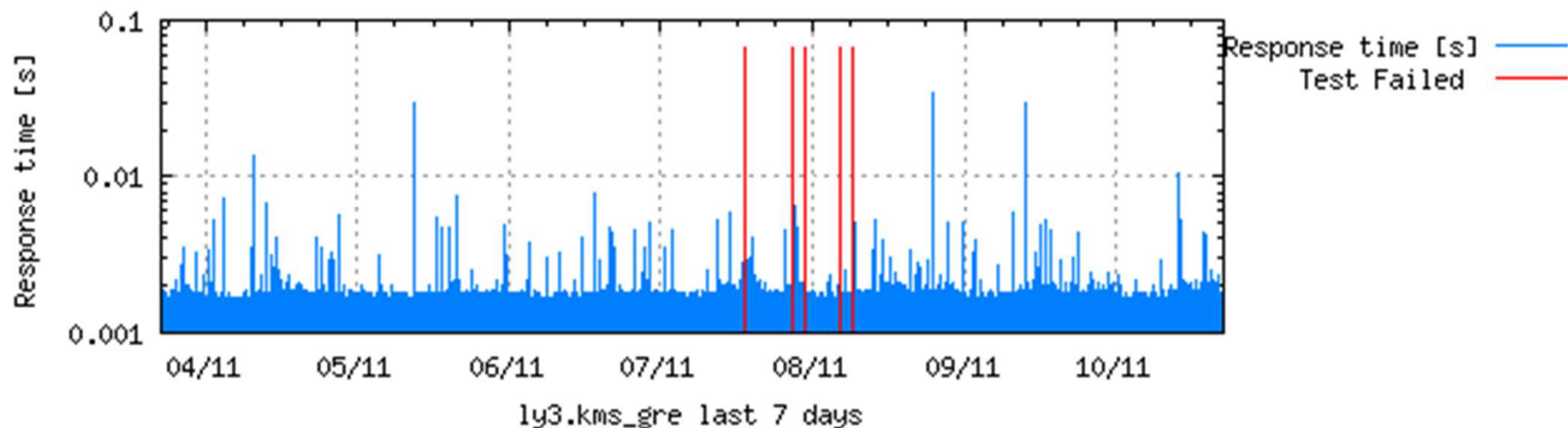
[Latest result](#)



Response time timeframe:08/11/2011-17:00 - 10/11/2011-17:00



Response time timeframe:03/11/2011-17:00 - 10/11/2011-17:00

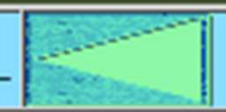


Status for

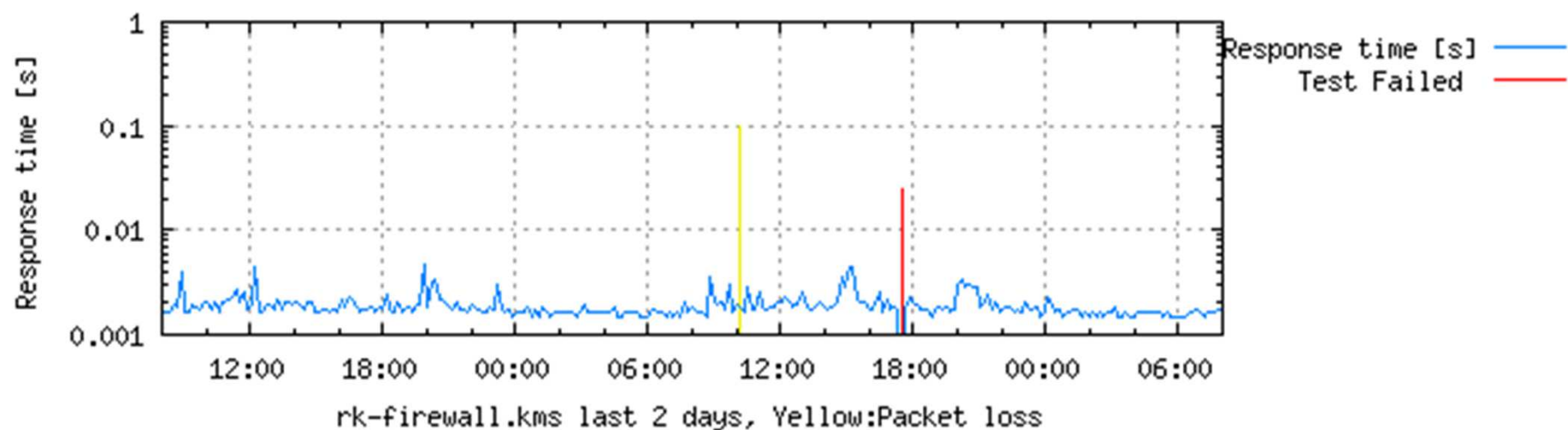
timeframe

[UDMC](#) service.

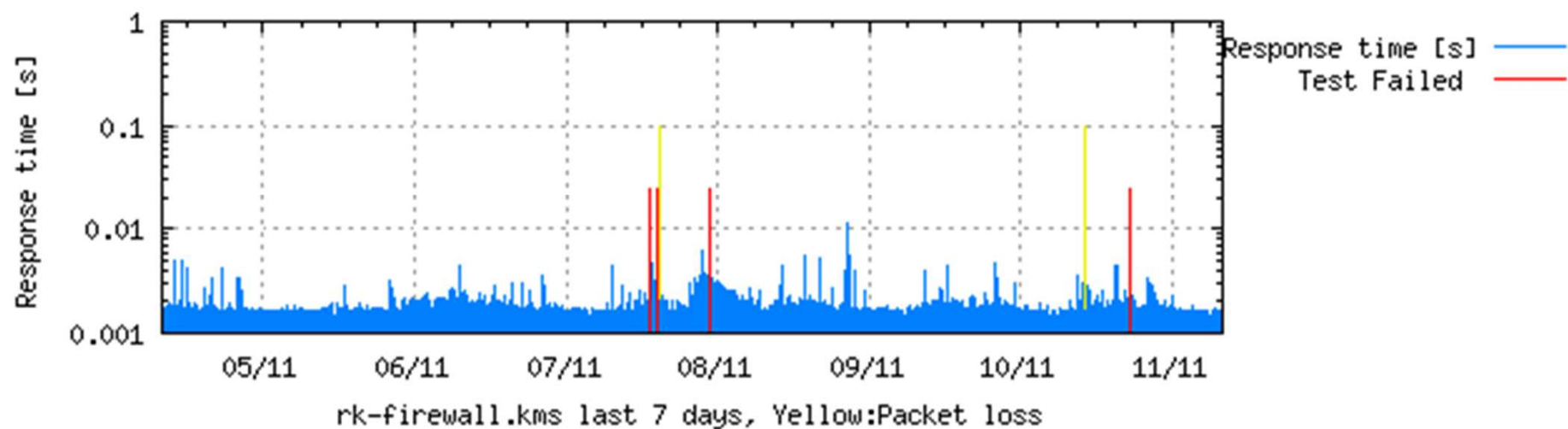
[Latest result](#)



Response time timeframe:09/11/2011-08:00 - 11/11/2011-08:00



Response time timeframe:04/11/2011-08:00 - 11/11/2011-08:00

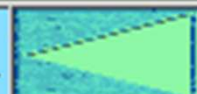


Status for

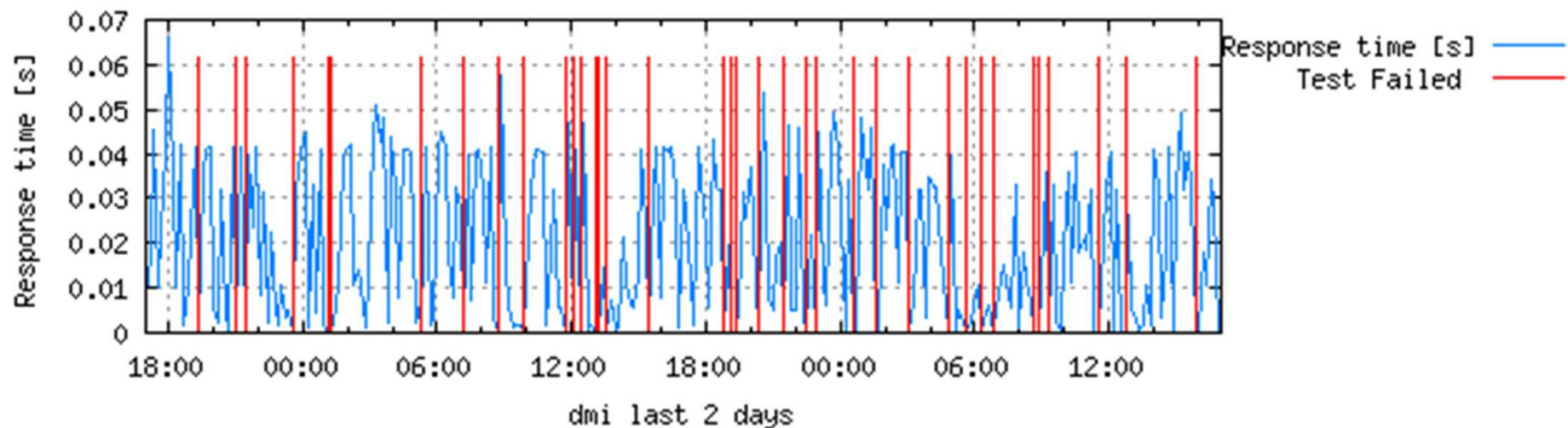
timeframe

[OSPF service.](#)

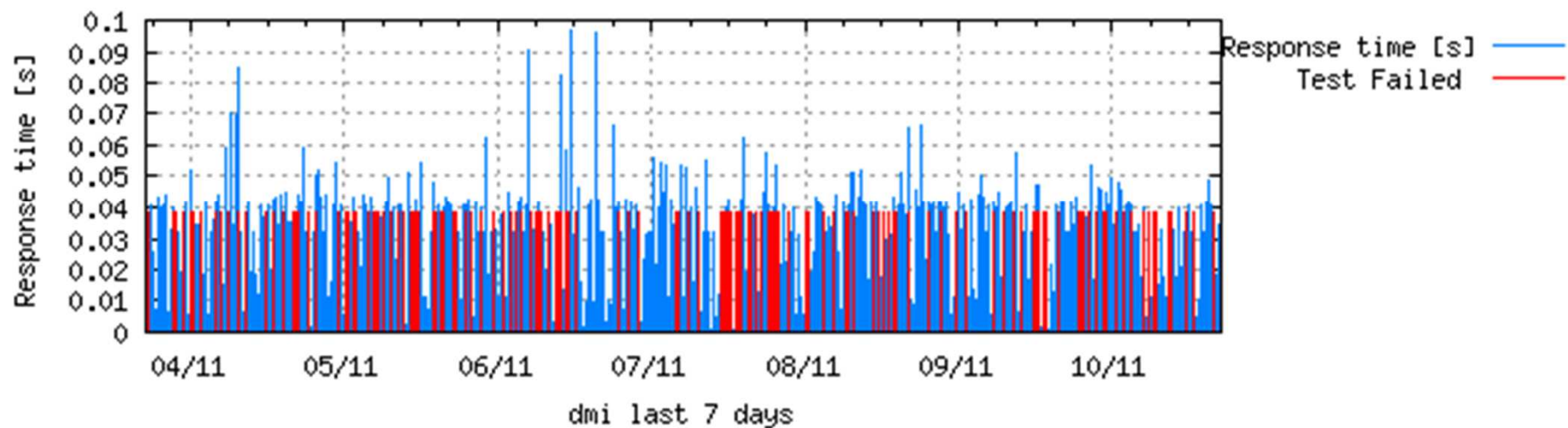
[Latest result](#)



Response time timeframe:08/11/2011-17:00 - 10/11/2011-17:00



Response time timeframe:03/11/2011-17:00 - 10/11/2011-17:00



Le calcul de la disponibilité relative est totalement traditionnel



- Pour chaque connexion WAN
- Pour chaque trimestre

Nous comptons

- Le nombre de mesures de panne, N_{panne}
- Le nombre de mesures en ordre, N_{ordre}

Et nous calculons la disponibilité relative du trimestre:

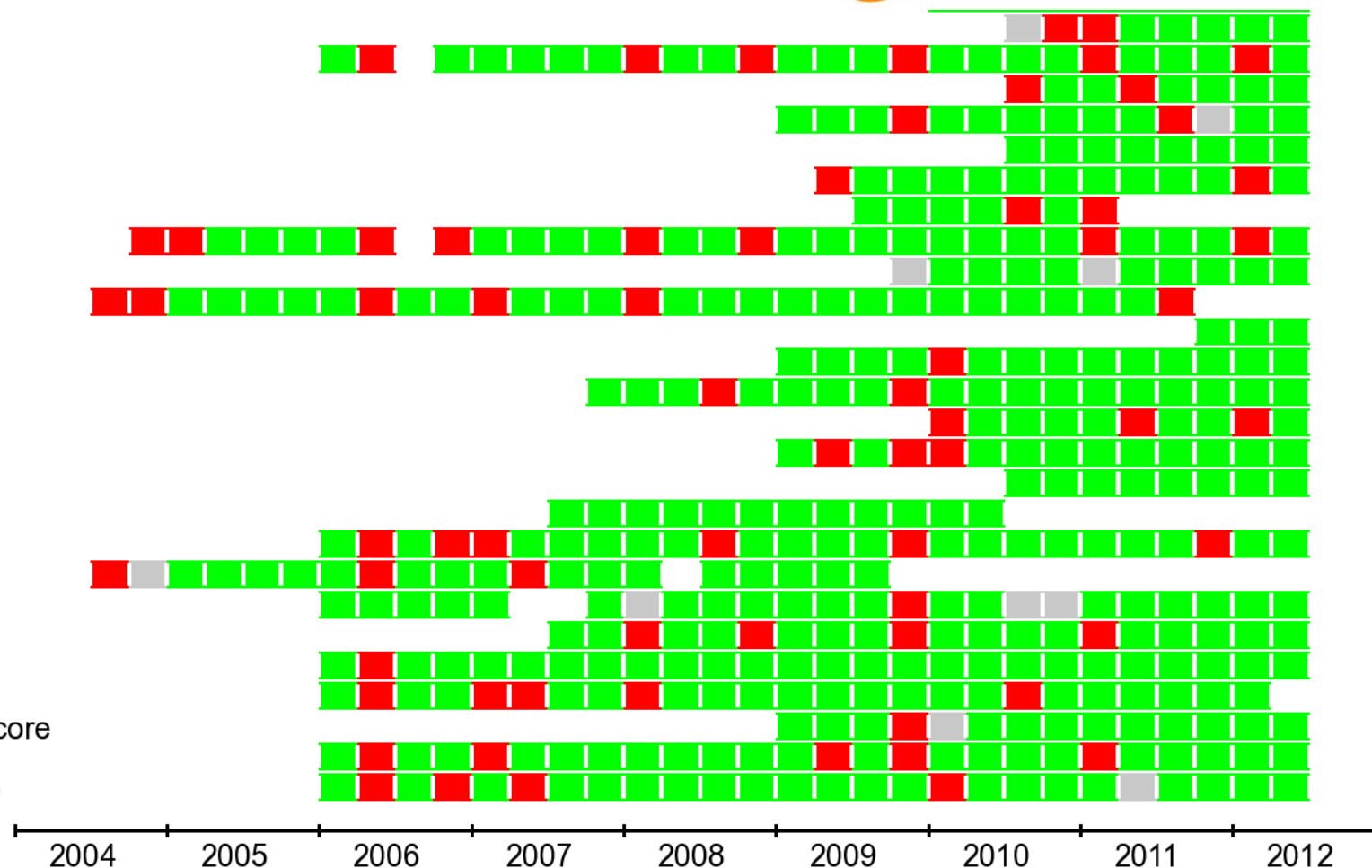
$$\frac{N_{\text{panne}}}{N_{\text{ordre}} + N_{\text{panne}}}$$



Observation de 27 connexions WAN non protégées – toutes avec une disponibilité garantie de 99,7%



- ge-hor.Riso.core
- ge-dpu.cbs
- ge-ver.kvlg.core
- ge-akf.ly3
- dk-bib.al.core
- ge-kolss.or1
- 10g-Aal.Ly0.core
- SYMBION
- ge-kvl.KbhG
- LO-SKOLEN
- fwa-dmi.ly8
- ge-ver.eva
- ge-ly3.ihk.core
- ge-ihk.ar5
- 10g-Ruc.Ly0.core
- fw.djh.core
- ge-cbs.ly3
- horsketten
- ge-ar2.ihk.core
- fe-hha.ar3
- ge-Scion.ly3.core
- Ge-KVL.Carlsberg
- ge-dpu.ly3.core
- ge-nbi.ly3.core
- ge-horsketten.ruc3.core
- fe-hhk.kvlg.core
- ge-AAU.UniAR.core

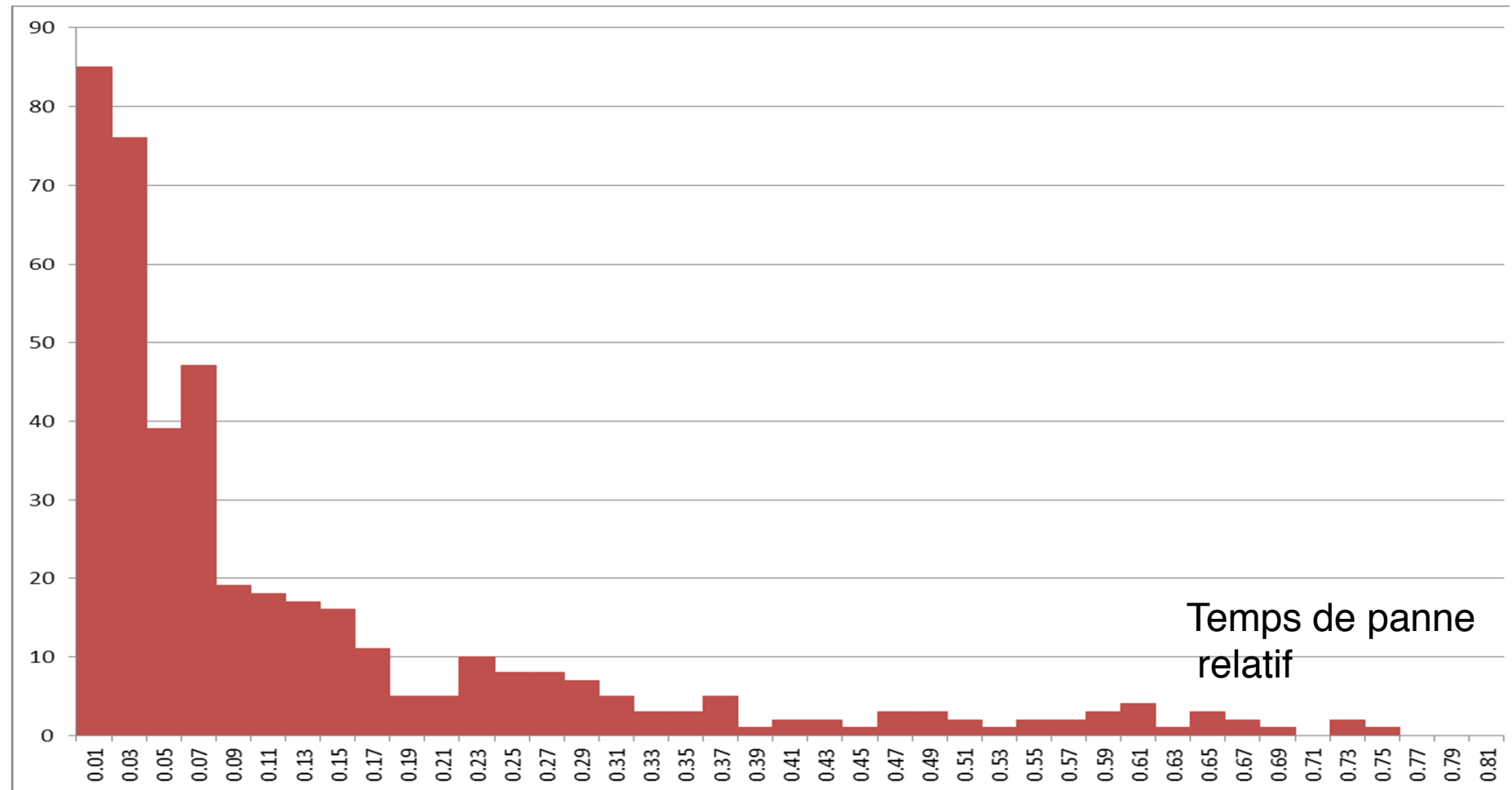


Légende : ■ ≥99.7% ■ <99.7% ■ exclues



Mesures de 450 trimestres (correspondant à 112 années)

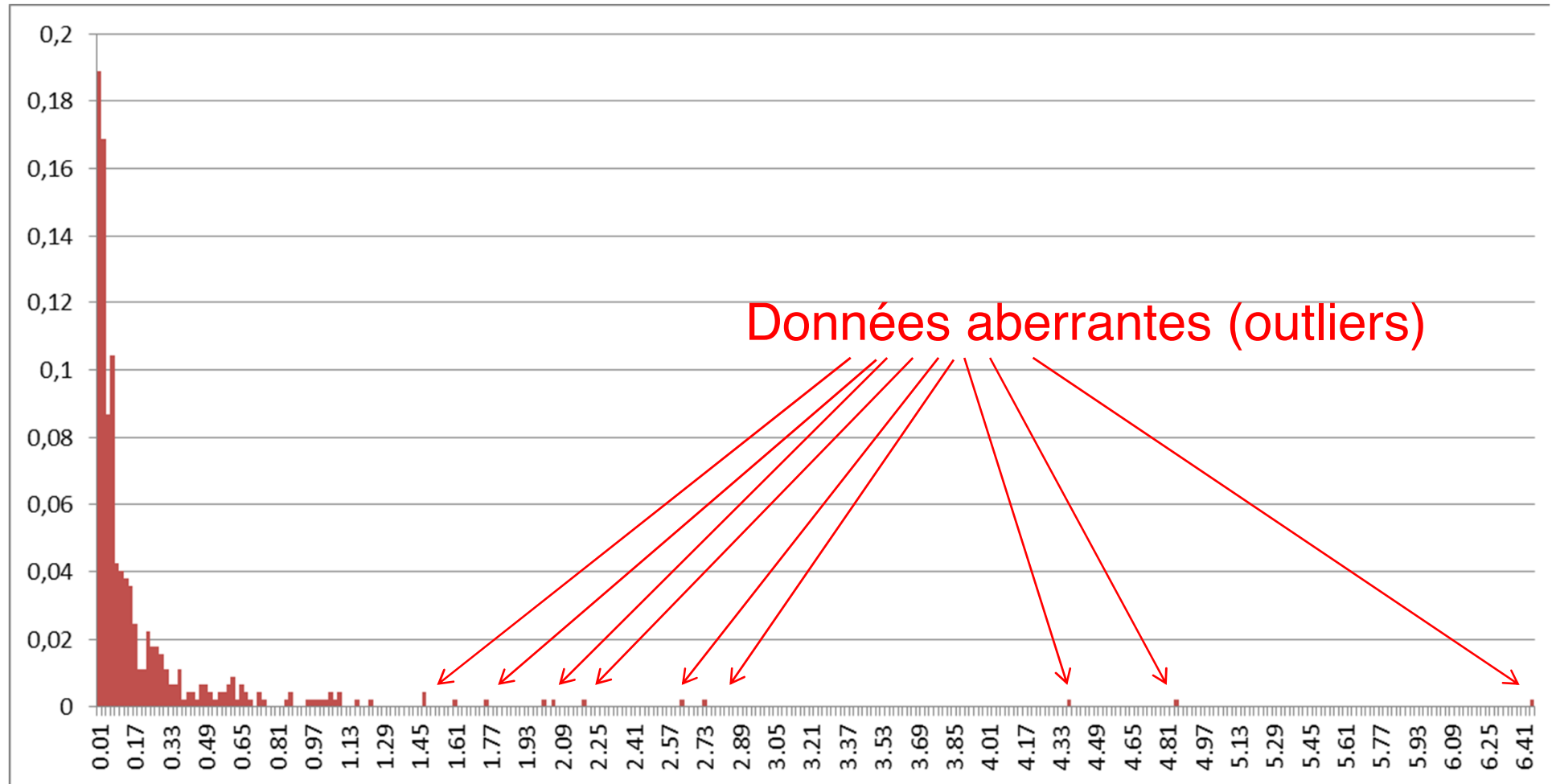
Nbre
de trimestres



Toutes les données ne sont pas représentées car certains trimestres connaissent un temps de panne plus élevé

La longue queue

(ici comme une distribution de probabilités pour l'ensemble des données)



Nous excluons les valeurs élevées (données aberrantes)



Nous ignorons :

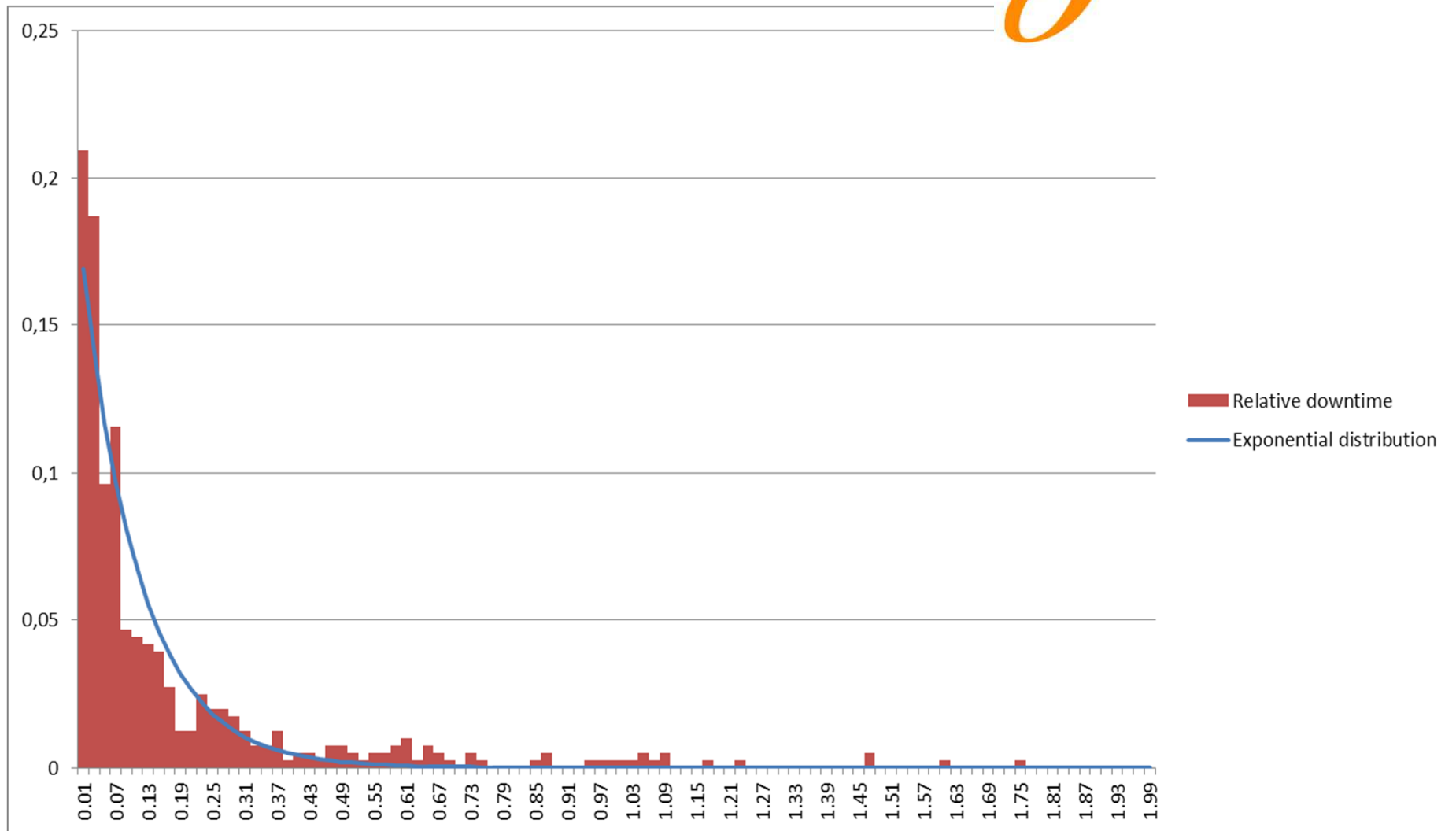
- Certaines pannes mesurées ne sont pas imputables à une faute du fournisseur mais renvoient à la pratique consistant à arrêter la surveillance pendant une période après l'arrêt d'une connexion
- Certaines sont liées à une maintenance qui n'est pas enregistrée dans le système de surveillance
- Un petit nombre est dû à des pannes accidentelles tragiques dont on peut présumer qu'elles n'avaient pas été prévues par le fournisseur et qui n'ont donc pas été prises en compte lors du calcul des prix, des garanties et des pénalités

Mesures erronées

Absence
d'enregistrement
des exceptions

Force Majeure

Cela ressemble à une distribution exponentielle



La distribution exponentielle de la probabilité



- Fonction de répartition $F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$
- Fonction de densité $f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$
- L'espérance ou la moyenne est $E(X) = \frac{1}{\lambda}$ (qui peut aussi être utilisée pour estimer λ)
- La distribution exponentielle est « sans mémoire », ce qui veut dire que pour une variable aléatoire X , on a

$$P(X > t + v | X > v) = P(X > t) \Leftrightarrow P(X > t + v) = P(X > t) \cdot P(X > v)$$

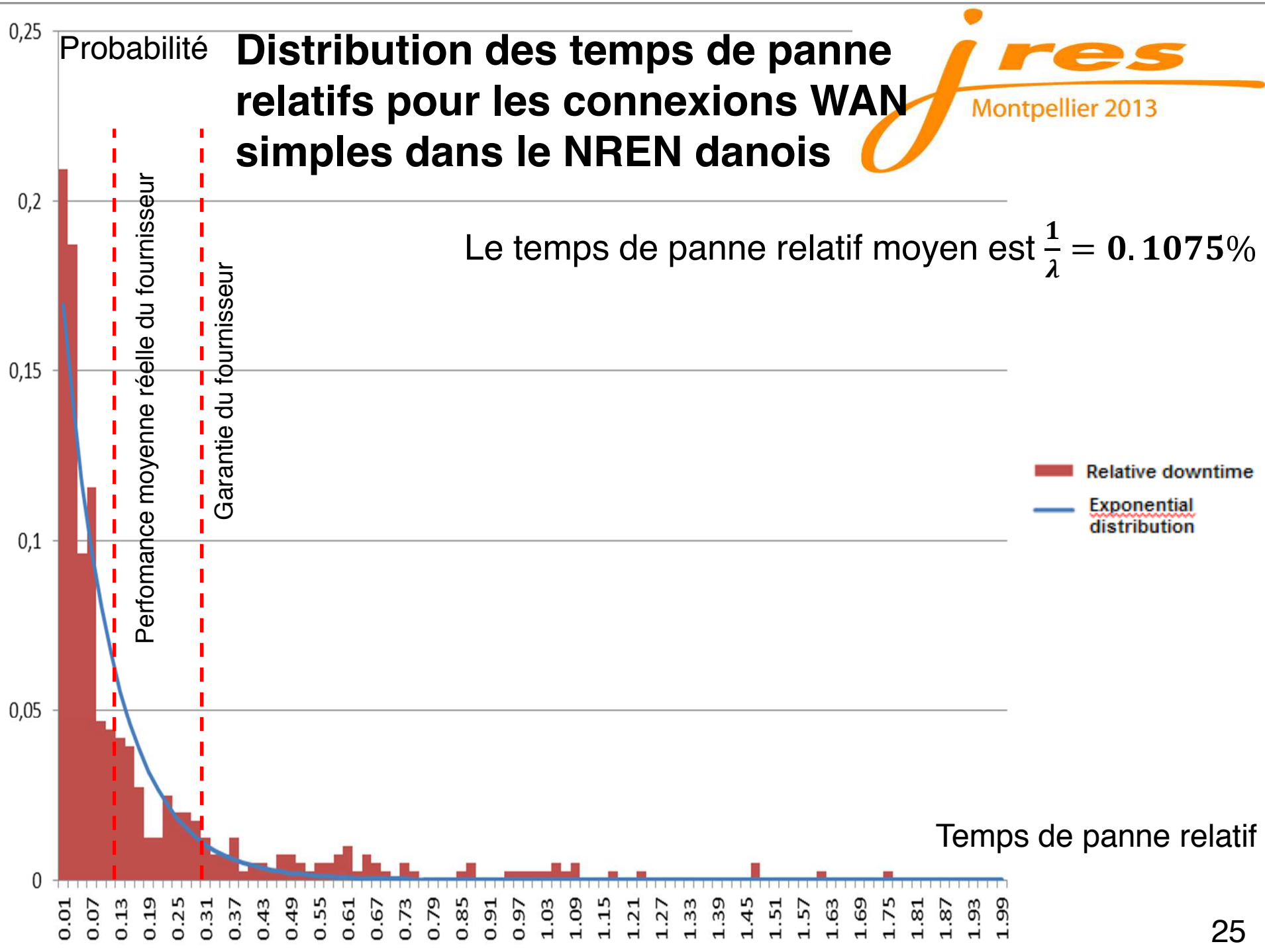
ou:

”la probabilité d’une heure supplémentaire de panne ne dépend pas du fait qu’il y ait eu des pannes auparavant”

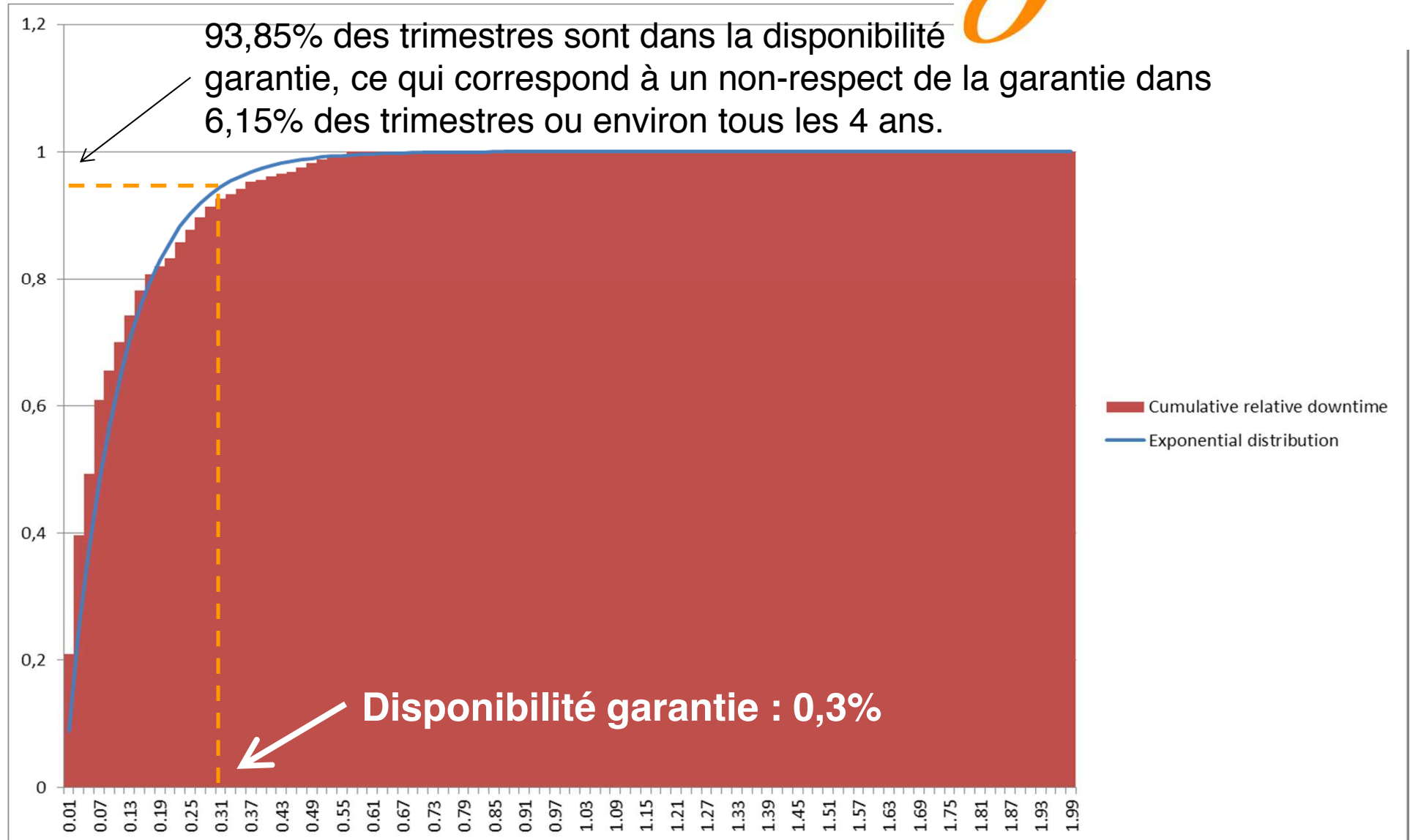
- La distribution exponentielle est la seule distribution statistique continue qui possède cette propriété.

Probabilité Distribution des temps de panne relatifs pour les connexions WAN simples dans le NREN danois

Le temps de panne relatif moyen est $\frac{1}{\lambda} = 0.1075\%$



Distribution cumulée



Quelques suppositions audacieuses et on peut calculer notre SLA



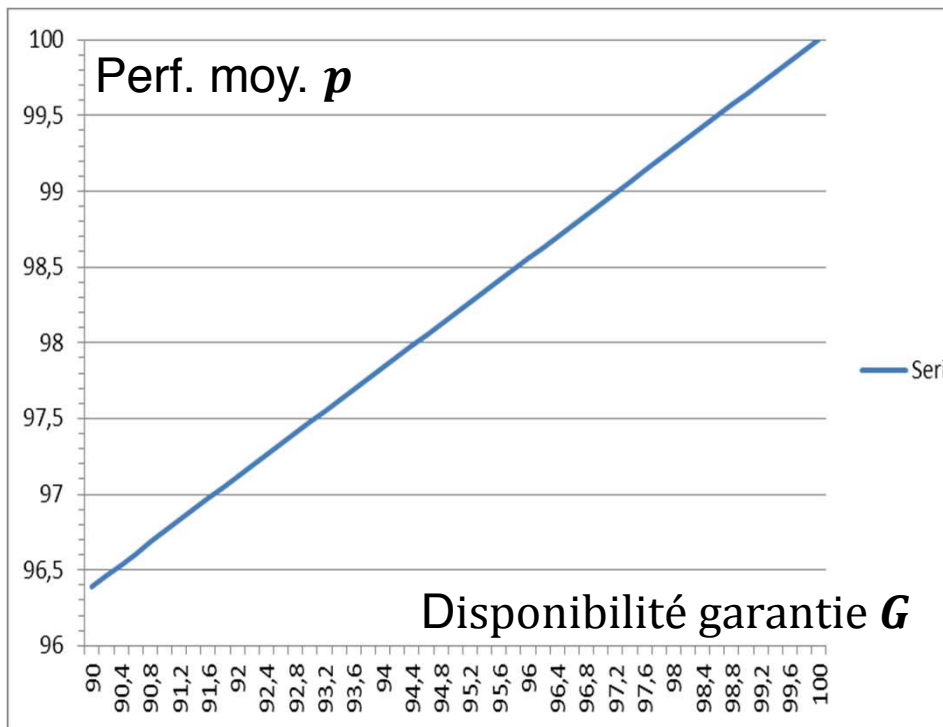
- Le temps de panne est distribué de manière exponentielle
- Dans cette classe de services (connexions WAN simples), les fournisseurs (et nous) avons: MTBF = θ = 4 ans = 16 trimestres
- Si la disponibilité garantie est G , on peut trouver la performance moyenne, p , de la manière suivante

$$1 - \frac{1}{\theta} = F(1 - G) = 1 - e^{-\lambda(1-G)}$$
$$\Leftrightarrow p = 1 - \frac{1}{\lambda} = 1 - \frac{1 - G}{\ln \theta}$$

Conversion en performance moyenne



Dispo. garantie <i>G</i>	Perf. moy., <i>p</i>
95%	98.20%
96%	98.56%
97%	98.92%
98%	99.28%
99%	99.64%
99.5%	99.82%
99.7%	99.892%
99.8%	99.928%
99.9%	99.964%
99.95%	99.982%
99.99%	99.9964%
100%	100%



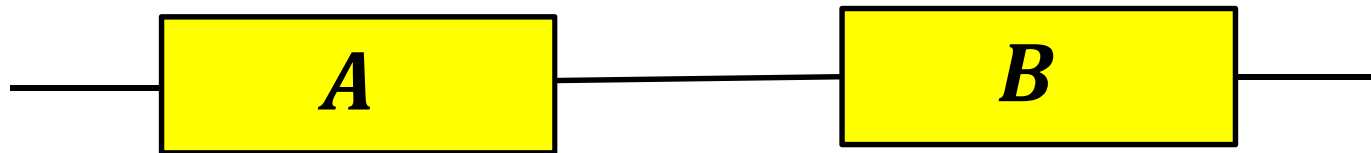
$$G = 1 - (1 - p) \ln \theta$$



Connexion en série : quand les deux parties doivent être disponibles

La performance combinée de deux services *A* et *B* qui doivent être tous les deux disponibles pour que le service soit disponible est:

$$p_{A \wedge B} = p_A \cdot p_B$$



Si l'on a deux services ayant des disponibilités garanties respectives de 99,7% et de 99,5%, les performances moyennes seront: $p_A = 99,892\%$ et $p_B = 99,82\%$

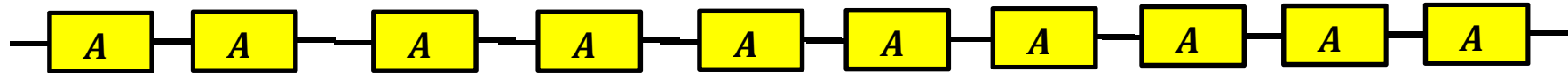
Et la performance combinée:

$$p_{A \wedge B} = p_A \cdot p_B = 99,892\% \cdot 99,82\% = 99,71\%$$

Et la disponibilité garantie correspondante (que l'on peut promettre à notre utilisateur) est:

$$G_{A \wedge B} = 1 - (1 - p_{A \wedge B}) \ln \theta = 99,2\%$$

Cas spécial de combinaison en série: beaucoup de systèmes identiques



La performance combinée est:

$$p_{10foisA} = p_A^{10}$$

Si l'on souhaite par exemple avoir une disponibilité garantie de **99,9%** pour chaque service, alors: $p_A = 99,964\%$ et la performance combinée:

$$p_{10foisA} = p_A^{10} = 0,99964^{10} = 99,004\%$$

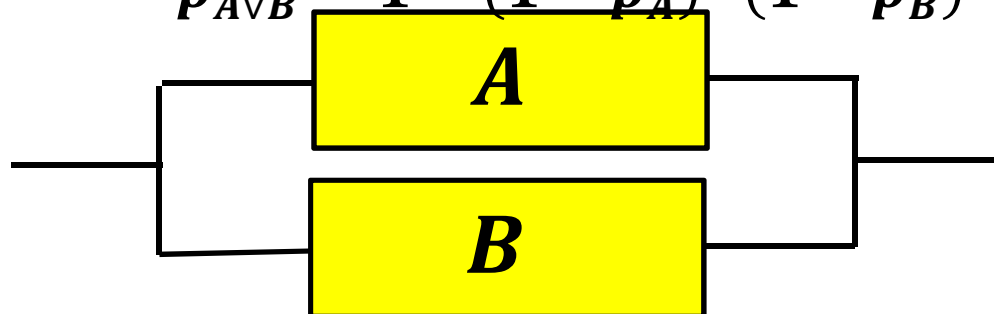
Et la disponibilité garantie correspondante (que l'on peut promettre à notre utilisateur) est:

$$G_{10foisA} = 1 - (1 - p_{10foisA}) \ln \theta = 97,2\%$$

Connexion en parallèle: quand seul un système doit être disponible

La performance combinée de deux services *A* et *B* qui se suppléent est

$$p_{A \vee B} = 1 - (1 - p_A) \cdot (1 - p_B)$$



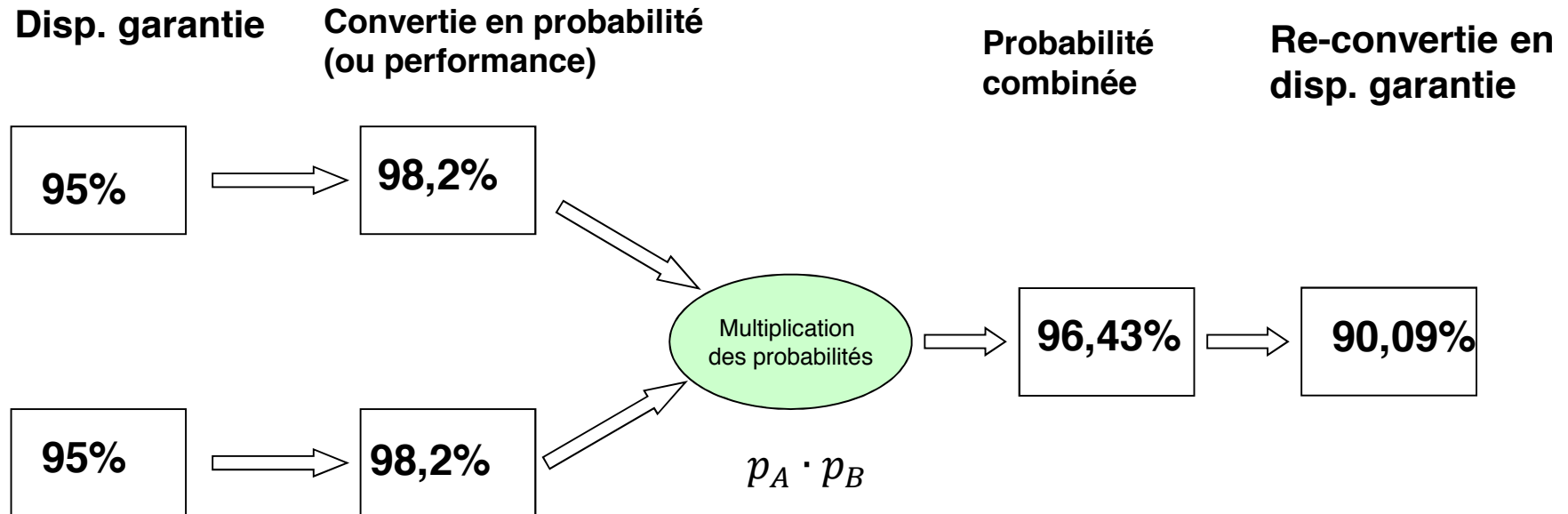
Si l'on a deux services ayant des disponibilités garanties respectives de 99,7% et de 99,5%, alors $p_A = 99,892\%$ et $p_B = 99,82\%$ et la performance combinée

$$p_{A \vee B} = 1 - (1 - p_A) \cdot (1 - p_B) = 1 - 1,08\% \cdot 1,80\% = 99,999806\%$$

Et la disponibilité garantie correspondante (que l'on peut promettre à notre utilisateur) est:

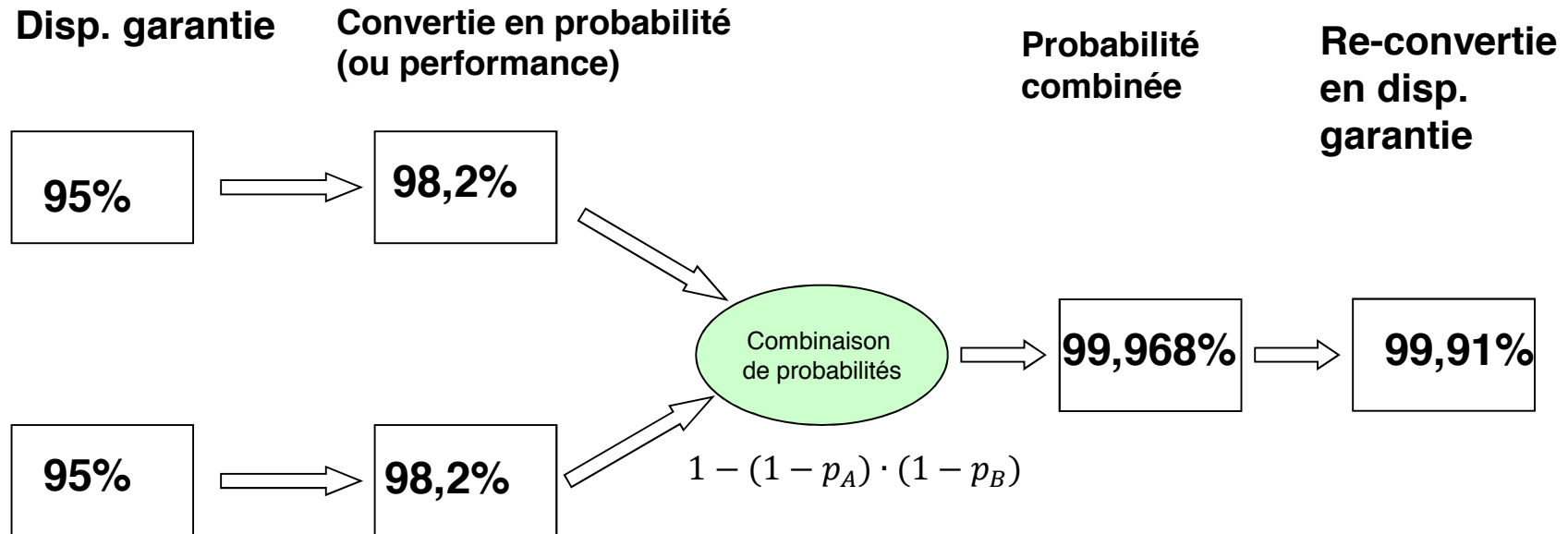
$$G_{A \vee B} = 1 - (1 - p_{A \vee B}) \ln \theta = 99,99946\%$$

Nous maîtrisons maintenant le processus avec une connexion en série de deux services à 95% comme exemple



Ce qui est (légèrement) différent du résultat plus imprécis si l'on traite les disponibilités garanties comme des probabilités
 $95\% \cdot 95\% = 90,25\%$

Autre exemple avec une connexion en *parallèle* de deux services à 95% comme exemple



Ce qui est (légèrement) différent du résultat plus imprécis si l'on traite les disponibilités garanties comme des probabilités

$$1 - (1 - 95\%) \cdot (1 - 95\%) = 99,75\%$$

Pourquoi utiliser cette méthode ?



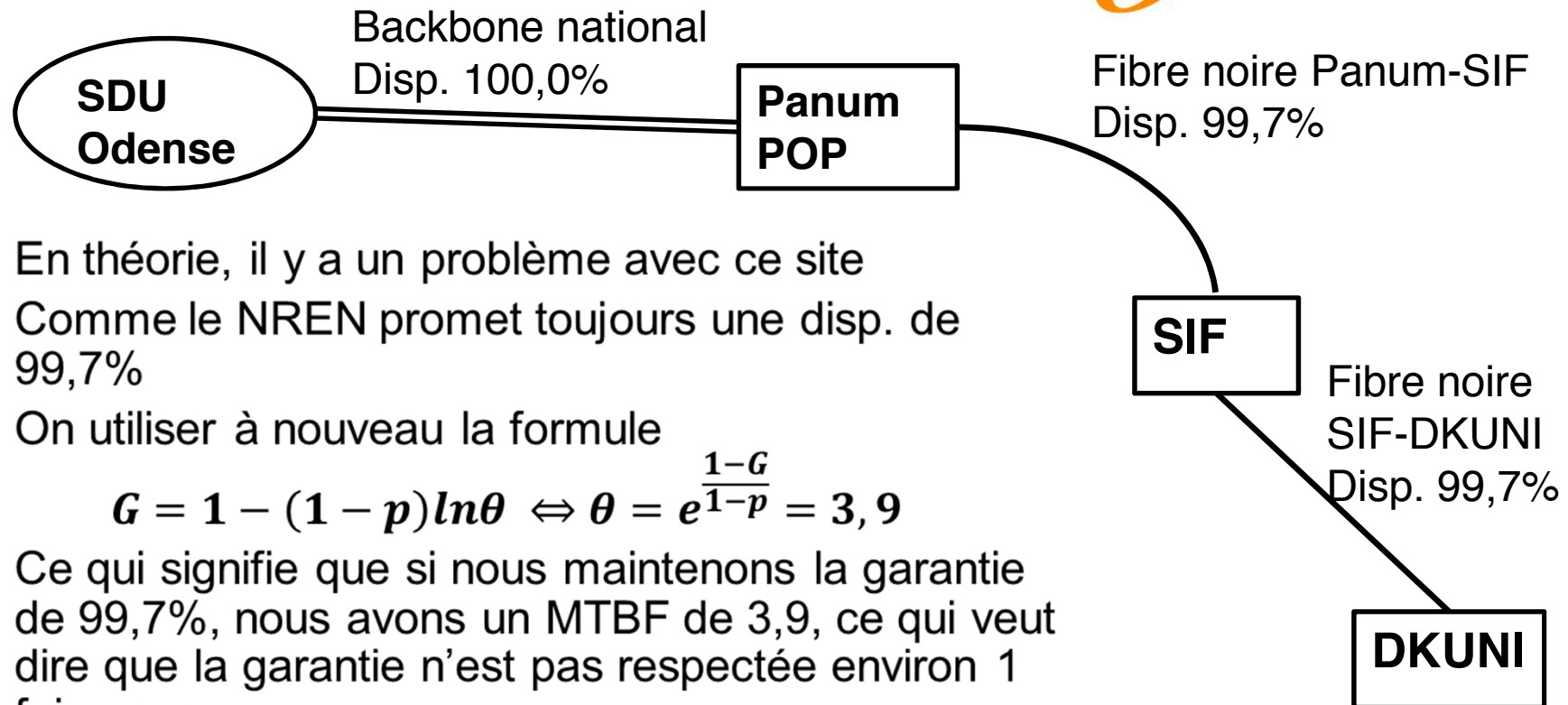
- Pour une connexion de lignes WAN avec une disponibilité garantie de 99,7%, la disponibilité garantie résultante est de 99,40032%, ce qui n'est pas significativement différent de $99,7\% \cdot 99,7\% = 99,4009\%$ (30 sec sur un trimestre)
- Cependant – comme nous l'avons vu – pour des temps de panne plus longs et des situations complexes, cette nouvelle méthode produit des résultats significativement plus précis
- Et cela permet aussi autre chose: calcul des pénalités escomptées si l'on change la disponibilité garantie – voir les exemples sur les diapositives suivantes

Le SLA du NREN danois (pour la connectivité de base)



- La disponibilité garantie est de 99,7%, mesurée chaque trimestre pour chaque connexion
- La disponibilité garantie est de 100,0% pour les connexions optiques protégées par redondance

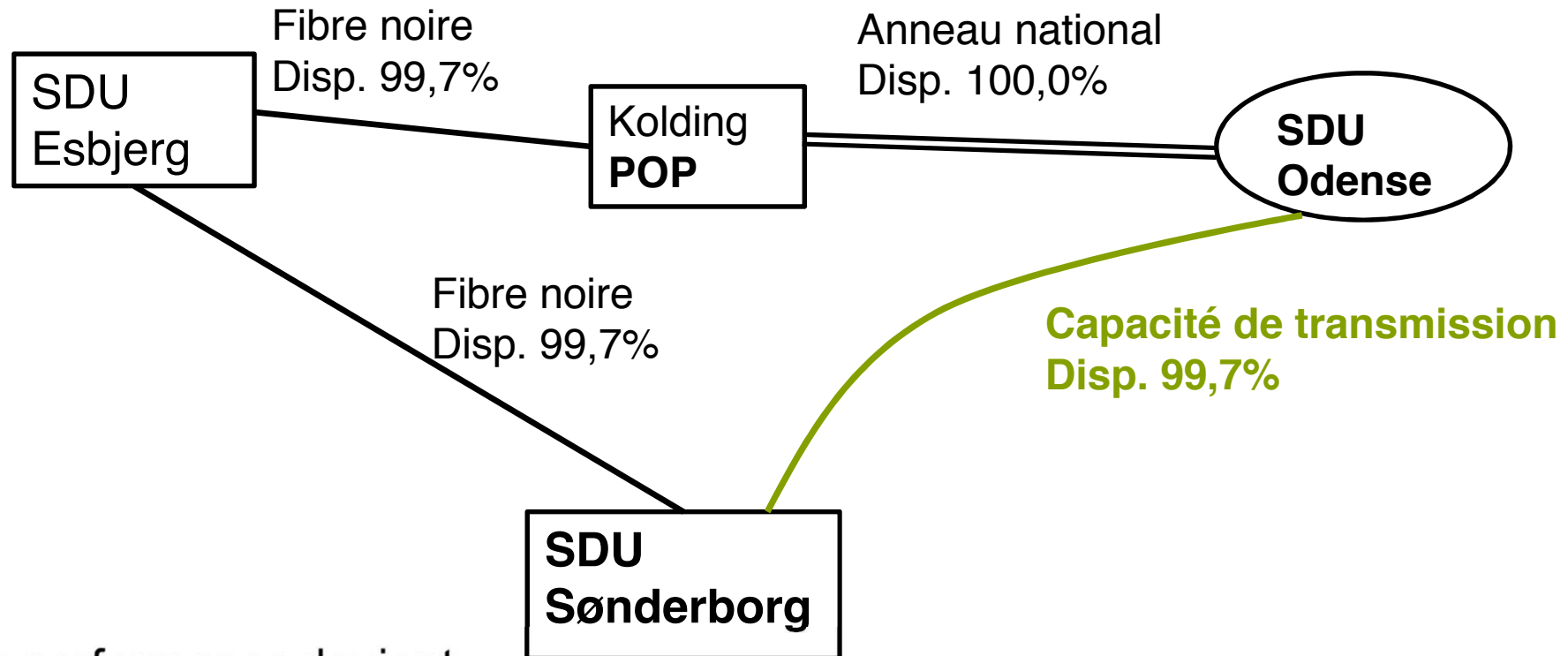
Exemple: Odense-Copenhagen P2P



- En théorie, il y a un problème avec ce site
- Comme le NREN promet toujours une disp. de 99,7%
- On utilise à nouveau la formule
$$G = 1 - (1 - p) \ln \theta \Leftrightarrow \theta = e^{\frac{1-G}{1-p}} = 3,9$$
- Ce qui signifie que si nous maintenons la garantie de 99,7%, nous avons un MTBF de 3,9, ce qui veut dire que la garantie n'est pas respectée environ 1 fois par an
- Cependant, dans ce cas, nous estimons que les longues fibres présentent plus de risques que les courtes et nous promettons donc un 99,7% confortable (et nous avons commandé des fibres redondantes)

Disp. résultante pour DKUNI: 99,4%

Exemple 2: Odense-Sønderborg P2P



La performance devient

$$1 - (1 - 99,892\% \cdot 99,892\%)(1 - 99,892\%) = 99,99976\%$$

Et la disponibilité garantie résultante est donc de **99,99935%**

Vous pouvez maintenant le faire vous-même!

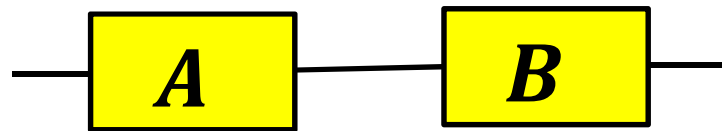


Vous pouvez maintenant calculer :

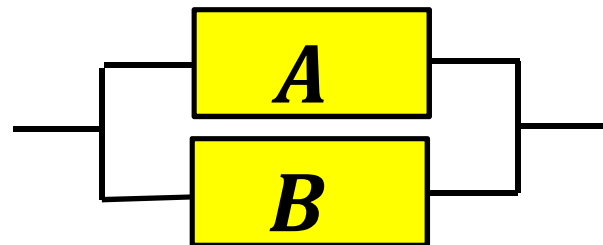
- Quel SLA vous pouvez promettre à vos utilisateurs
- Si les SLA dans les contrats de vos fournisseurs sont adéquats étant donné ce que vous avez déjà promis à vos utilisateurs
- Quelles pénalités vous pouvez vous attendre à payer ou à recevoir dans le futur
- Les SLA suffisants à préciser quand vous achèterez des services dans le futur

Bon courage...

À votre manuel-formulaire pour combiner des services avec des temps de panne distribués de manière exponentielle



$$G_{A \wedge B} = 1 - \ln \theta \left(1 - \left(1 - \frac{1 - G_A}{\ln \theta} \right) \left(1 - \frac{1 - G_B}{\ln \theta} \right) \right)$$



$$G_{A \vee B} = 1 - \frac{(1 - G_A)(1 - G_B)}{\ln \theta}$$

Où G est la disponibilité garantie ,
et θ le nombre moyen de périodes entre l'absence de respect de la disponibilité garantie.

La relation entre G et la performance moyenne p :

$$G = 1 - (1 - p)\ln \theta \Leftrightarrow p = 1 - \frac{1 - G}{\ln \theta} \Leftrightarrow \theta = e^{\frac{1 - G}{1 - p}}$$

Questions?



Merci de votre attention et
patience

Martin BECH
martin.bech@deic.dk