

Noyau persistant en réseaux pair-à-pair

Comment relier la taille à la durée de vie

V. Gramoli, A-M. Kermarrec, A. Mostefaoui, M. Raynal, B. Sericola



IRISA

UNE UNITÉ DE RECHERCHE À LA POINTE DES SCIENCES
ET DES TECHNOLOGIES DE L'INFORMATION
ET DE LA COMMUNICATION



// Contexte

⇒ Les systèmes dynamiques à grande échelle

- Les nœuds quittent et rejoignent le système
- Lorsqu'ils reviennent il ne possèdent pas forcément leurs données
- Les nœuds ne peuvent maintenir une information globale

⇒ Problème de persistance des données

- Pour une donnée, si tous les nœuds la détenant quittent le système, cette donnée est perdue

⇒ Constatations sur les systèmes Peer-to-Peer

- Très dynamiques
- Jamais vides



// But

⇒ Assurer la persistance des données en dépit du dynamisme du système.

⇒ Défi majeur

→ Etant donné

- La probabilité requise, p , et
- Le *va-et-vient* (« churn ») du système, ν ,

→ Il faut recopier la donnée

- Ajuster la période de recopiage, δ ,
- Ajuster la taille de la copie, q .



// Modèle du système

⇒ Système distribué à grande échelle

- n nœuds interconnectés
- Chacun avec id unique
- Sans connaissance globale

⇒ Système dynamique

- Les nœuds rejoignent/quittent le système.
- Un nœud rejoignant est nouveau.

⇒ Donnée

- Une donnée est détenue par un sous-ensembles de nœuds, le *noyau*.



// Modèle de *va-et-vient* (« churn »), ν

⇒ **Va-et-vient:**

→ Intensité du dynamisme du système.

⇒ **Il représente:**

→ Le taux de départ et d'arrivée par nœud et par unité de temps.

⇒ **On observe le système à 2 instants**

→ Soit Q le noyau de départ, et q sa taille,

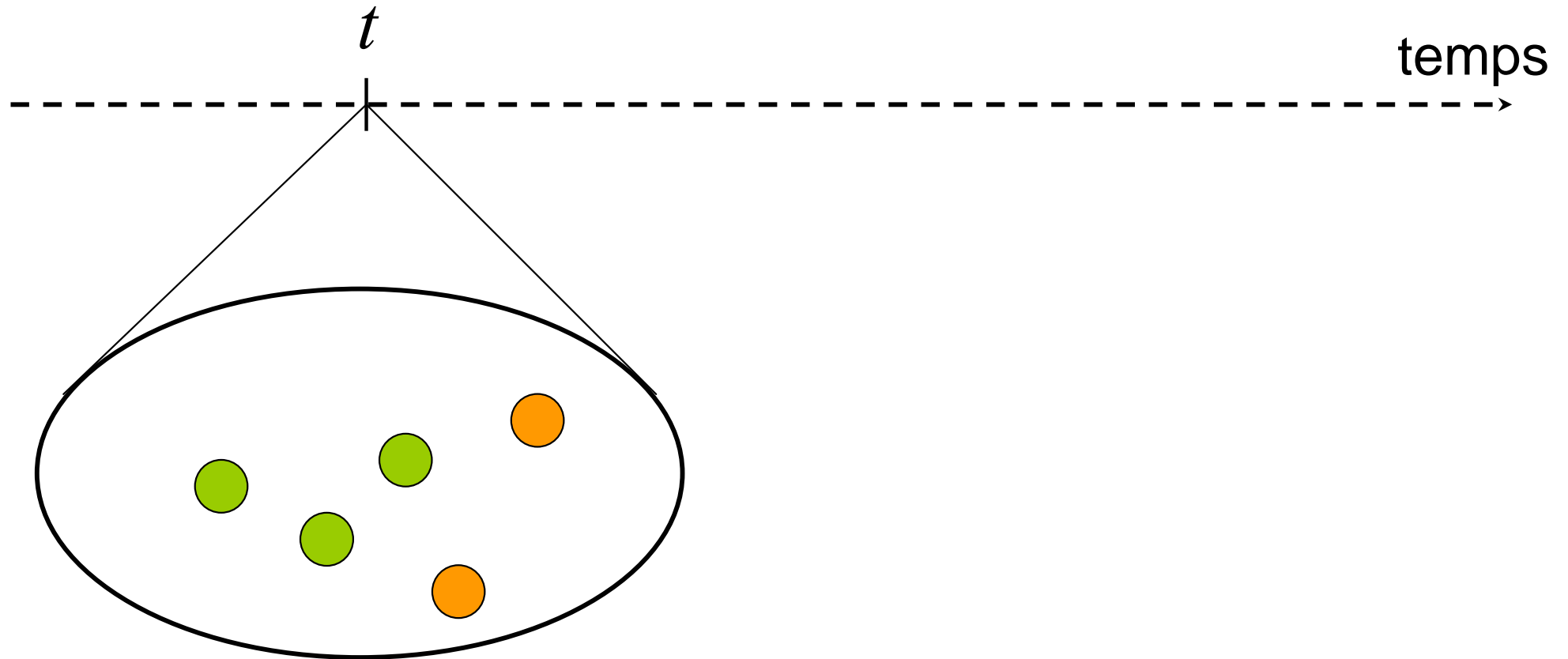
→ Soit A les nœuds remplacés, et α sa taille,

→ Soit Q' le noyau après remplacement.



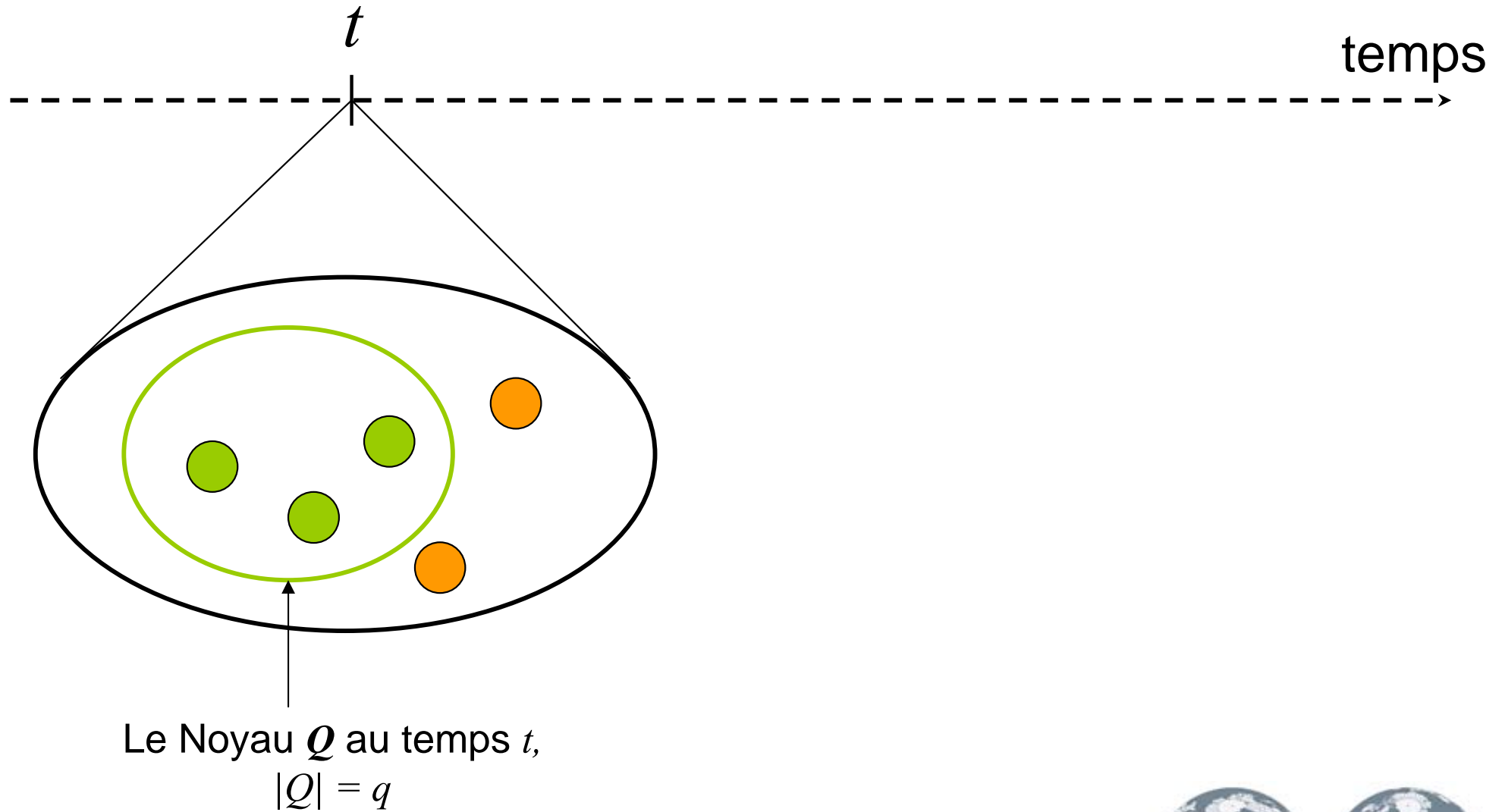
// Modèle de *va-et-vient*

- Nœuds avec la donnée.
- Nœuds sans la donnée.



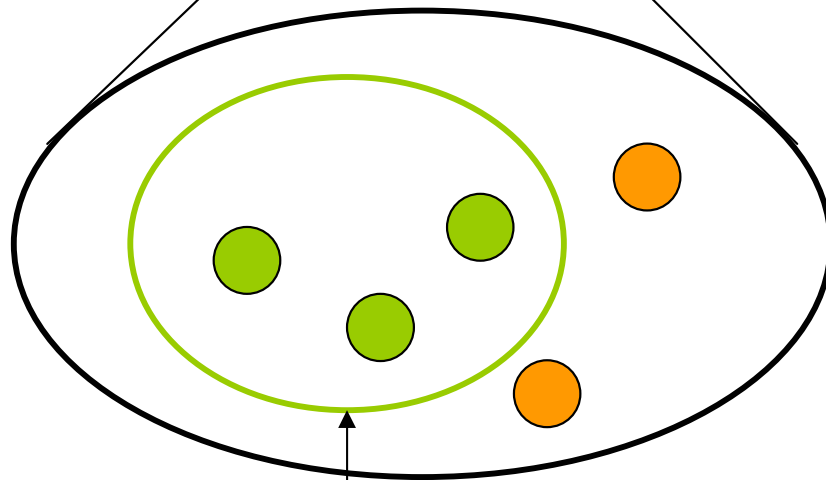
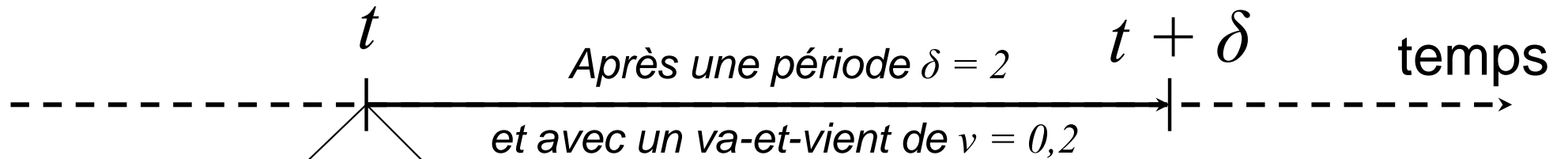
// Modèle de *va-et-vient*

- Nœuds avec la donnée.
- Nœuds sans la donnée.



// Modèle de *va-et-vient*

- Nœuds avec la donnée.
- Nœuds sans la donnée.

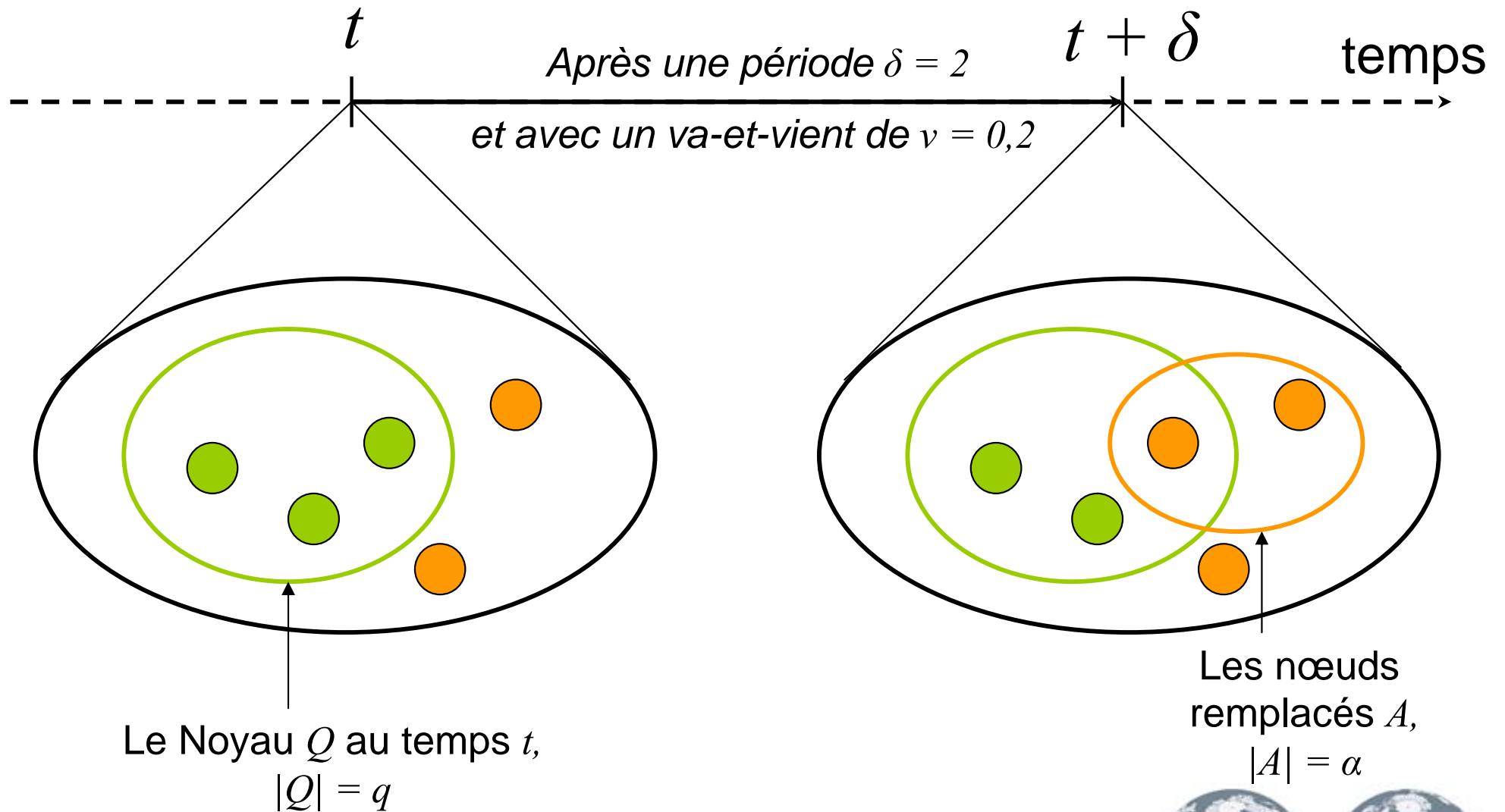


Le Noyau Q au temps t ,
 $|Q| = q$



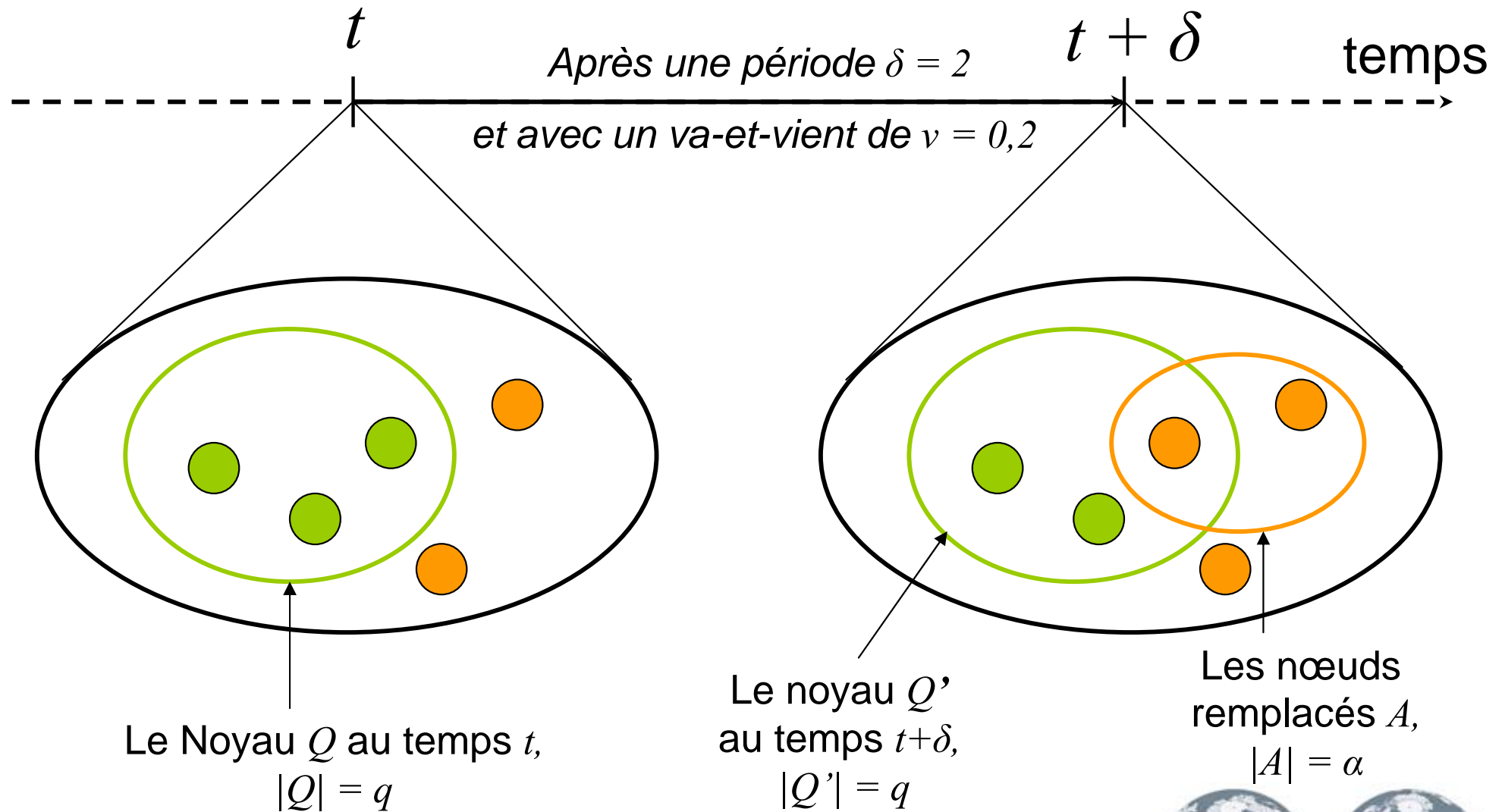
// Modèle de *va-et-vient*

- Nœuds avec la donnée.
- Nœuds sans la donnée.



// Modèle de *va-et-vient*

- Nœuds avec la donnée.
- Nœuds sans la donnée.



// Modèle de *va-et-vient*

⇒ Evolution du nombre des nœuds initialement présents

→ t_0 n nœuds initiaux

→ t_1 $n - n\nu = n(1-\nu)$ nœuds initiaux

...

→ t_i $n(1-\nu)^i$ nœuds initiaux

→ t_{i+1} $n(1-\nu)^i - n(1-\nu)^i\nu = n(1-\nu)^{i+1}$ nœuds initiaux

⇒ On choisit $\alpha = \lceil n - n(1-\nu)^\delta \rceil$ le nombre de nœuds remplacés après δ unités de temps



// Disponibilité d'une donnée

- ⇒ Initialement, q nœuds ont la donnée
- ⇒ Les nœuds remplacés sont choisis aléatoirement de façon uniforme
- ⇒ Combien de copies de la donnée restent-il de disponible après δ unités de temps dans un système avec va-et-vient ν ?



// Disponibilité d'une donnée

⇒ Observation préliminaire

→ Le nombre $\beta = |Q' \cap A|$ de nœuds qui avaient la donnée et quittent le système est borné:

$$\underbrace{\max(0, a + q - n)}_{a} \leq \beta \leq \underbrace{\min(a, q)}_{b}$$



// Disponibilité d'une donnée

⇒ Probabilité que $\beta = k$ copies de la donnée aient été remplacées ?

$$\Pr[\beta = k] = \frac{\binom{q}{k} \binom{n - q}{\alpha - k}}{\binom{n}{\alpha}}.$$



// Chercher une donnée

- ⇒ Initialement, q nœuds ont la donnée.
- ⇒ δ unités de temps plus tard, on tire aléatoirement et de façon uniforme q nœuds du système.
- ⇒ Quelle est la probabilité qu'on trouve la donnée critique après ces δ unités de temps dans un système avec va-et-vient v ?



// Chercher une donnée

⇒ Probabilité de ne pas trouver la donnée

→ Tirage aléatoire, uniforme et sans remise de q nœuds.

Soit $E = Q' \setminus A$.

$$\Pr[x \notin E] = \Pr[\{x \notin Q'\} \cup \{x \in Q' \cap A\}],$$

$$\Pr[x \notin E] = \Pr[x \notin Q'] + \Pr[x \in Q' \cap A]. \quad (\text{évts disjoints})$$

$$\begin{aligned} \Pr[x_1 \notin E, \dots, x_q \notin E / \beta = k] &= \sum_{k=a}^b \prod_{i=1}^q \left(1 - \frac{q-k}{n-i+1}\right), \\ &= \sum_{k=a}^b \frac{\binom{n-q+k}{q}}{\binom{n}{q}}. \end{aligned}$$



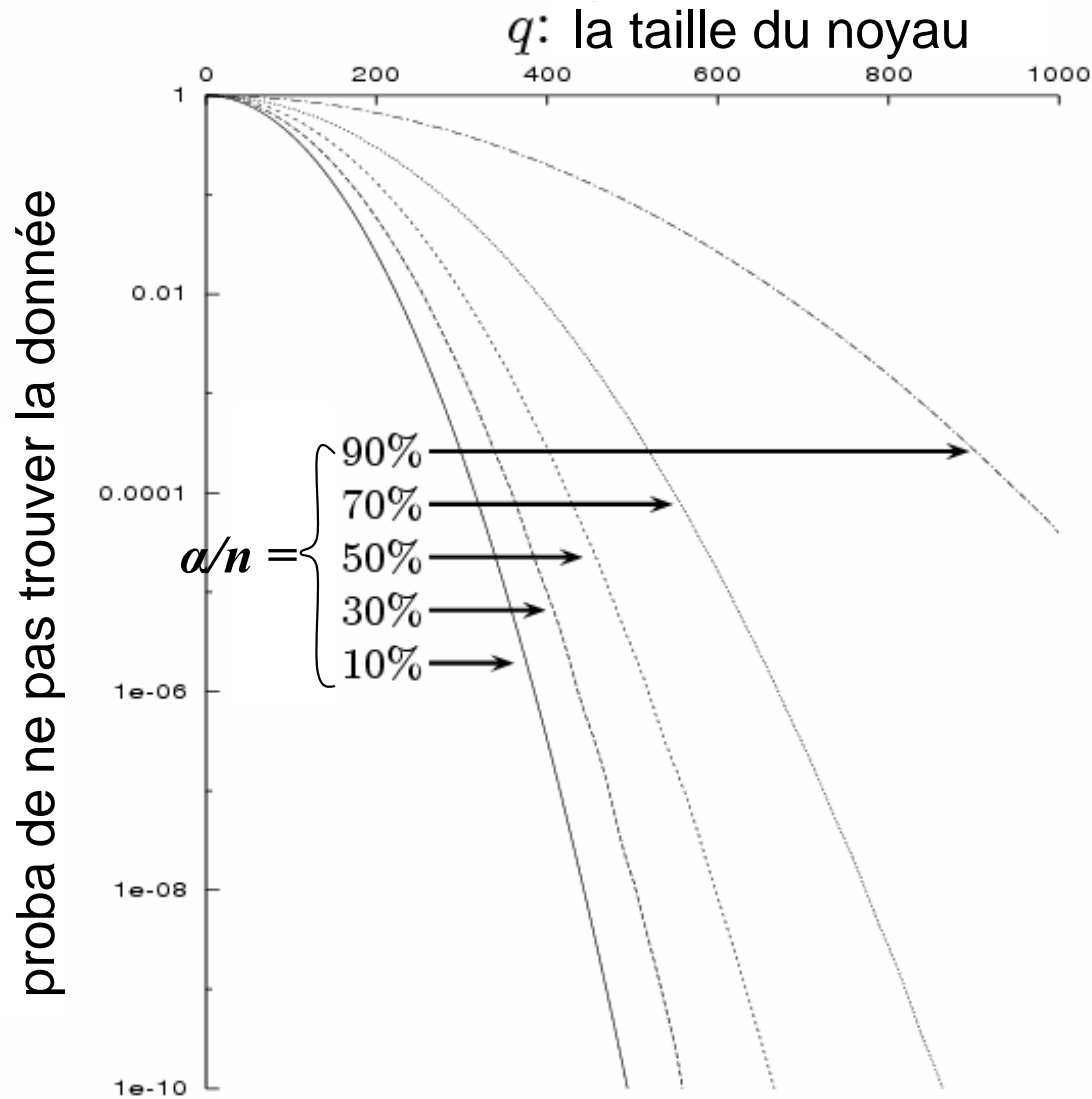
// Chercher une donnée

⇒ Probabilité de ne pas trouver la donnée

$$\begin{aligned}\Pr[x_1 \notin E, \dots, x_q \notin E] &= \sum_{k=a}^b \frac{\binom{n-q+k}{q}}{\binom{n}{q}} \Pr[\beta = k], \\ &= \frac{\sum_{k=a}^b \left[\binom{n+k-q}{q} \binom{q}{k} \binom{n-q}{\alpha-k} \right]}{\binom{n}{q} \binom{n}{\alpha}}.\end{aligned}$$



// Taille de Noyau pour $n = 10^4$



// Probabilité, dynamisme, durée de vie et noyau

⇒ Variation du va-et-vient et de la probabilité

Proba de trouver	α/n	Taille du noyau pour		
		$n = 10^3$	$n = 10^4$	$n = 10^5$
99%	static	66	213	677
	10%	70	224	714
	30%	79	255	809
	60%	105	337	1071
	80%	143	478	1516
99.9%	static	80	260	828
	10%	85	274	873
	30%	96	311	990
	60%	128	413	1311
	80%	182	584	1855



// Conclusion

➔ **Retrouver une donnée est paradoxalement facile!**

➔ **Applications de stockage**

→ Modifier les données en q nœuds

→ Accéder les données à jour en contactant q nœuds

→ Les noyaux sont des quorums probabilistes

➔ **Futures recherches**

→ Modéliser le churn en utilisant un modèle plus réaliste (Markovien continue).

→ Spécifier un protocole pour la cohérence/persistance probabiliste des données en systèmes dynamiques.



// Des références

- ⇒ **A Quorum based protocol for searching objects in P2P ntwks.**
K. Miura, T. Tagawa, and H. Kakugawa.
IEEE Trans. on Parallel and Distributed Systems, 17(1):25–37, 2006.
- ⇒ **Probabilistic quorums for dynamic systems.**
I. Abraham and D. Malkhi.
Distributed Computing, 18(2):113–124, 2005.
- ⇒ **Reconfigurable distributed storage for dynamic ntwks.**
G. Chockler, S. Gilbert, V. Gramoli, P. M. Musial, and A. A. Shvartsman.
In *Proc. of 9th Int'l Conf. on Principles of Distributed Systems*, 2005.

