

**eric** ENTREPÔTS, REPRÉSENTATION  
de l'INGÉNIÉRIE des CONNAISSANCES

**dis**  
UNIVERSITÉ LYON 2  
Institut de la Communication

**Décisionnel dans le Nuage**  
Sécurité et Performance

**Master 2 ECD**  
2015-2016

**Jérôme Darmont**  
eric.univ-lyon2.fr/~jdarmont/

UNIVERSITÉ LYON 1 UNIVERSITÉ LYON 2 UNIVERSITÉ DE LYON **ish**

1

**eric** ENTREPÔTS, REPRÉSENTATION  
de l'INGÉNIÉRIE des CONNAISSANCES

**Section 1**

**Décisionnel dans le Nuage**

UNIVERSITÉ LYON 1 UNIVERSITÉ LYON 2 UNIVERSITÉ DE LYON **ish**

2

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

**Le meilleur des mondes**

**Cloud Computing Benefits**

- Service storage & management
- Location-independent
- Scalability & Sustainability
- 24/7 Support
- Agile Deployment
- Virtualized & Dynamic
- Pay-as-you-use
- Utility-based, Time-sharing models
- Low Total Cost Ownership
- High-level computing

cloudcomputingadvice.com

**eric** J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance

3

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

**Du décisionnel à papa...**

Operational System, ERP, CRM, File Files → **ETL** (Extraction, Transformation, Loading) → **Data Warehouse** (Metadata, Summary Data, Raw data) → **Utilisateurs bien identifiés** (Olap Analysis, Reporting, Data Mining)

Données propriétaires

Stockage « in house »

Dimensionnement *a priori* ⇒ investissement initial important

**eric** J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance

4

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

**...à la Cloud BI !**

ITEC, RF, OAS, Local Campus

Campus ERP, Campus ERP

danbrint.wordpress.com

**eric** J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance

5

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

**Vers le décisionnel à la demande**

Données propriétaires, Données situationnelles → **Entrepôt dans le nuage** → **Analyses [collaboratives]**

[Production collaborative]

Élasticité et paiement à la demande ⇒ investissement graduel

**eric** J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance

6

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Nouvelles données, nouveaux usages

Self-service BI

Données complexes

Données ouvertes

Personal BI

Crowd sourcing

Données situationnelles

Mégadonnées

Cloud intelligence

e<sup>2</sup>ic J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 7

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Scénario écolo

Association qui surveille les activités de pêche

Règlements

Statistiques

Géolocalisation

Observations de terrain

e<sup>2</sup>ic J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 8

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Nouveau X problème X!

Sécurité

Conception - Analyse

Confidentialité

Centrée utilisateur

Disponibilité

Bottom up

Elasticité des sources

Stockage

Nuage

Technologie

Modèle économique

Performance

Accepter une perte de contrôle

e<sup>2</sup>ic J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 9

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Déjà de quoi jouer !

Google Fusion tables

Coffee worldwide 2011

Country	Population	2008-10-04	2008-10-05	2008-10-06	
Austria	8,355,260	239.43	242.46	23	
Estonia	1,340,415	240.01	241.54	23	
Belgium	10,754,528	238.14	239.14	23	
Germany	81,882,342	1928-10-04	237.75	242.53	23
Bulgaria	7,606,551	1928-10-05	240.00	243.08	23
Cyprus	793,963				
Czech Republic	10,478,543				

e<sup>2</sup>ic J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 10

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Pour voir plus loin...

A. Abello, J. Darmont, L. Etcheverry, M. Golfarelli, J.N. Mazon, F. Naumann, T.B. Pedersen, S. Rizzi, J. Trujillo, P. Vassiliadis, G. Vossen, "Fusion Cubes: Towards Self-Service Business Intelligence", *International Journal of Data Warehousing and Mining*, Vol. 9, No. 2, 2013, 66-88

- Schéma extensible
- Instances extensibles
- Drill beyond (données situationnelles)
- Quasi temps réel

e<sup>2</sup>ic J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 11

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Architecture des services

e<sup>2</sup>ic J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 12

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Architecture fonctionnelle

is it good enough?

1 Recherche par mots clés  
Contraintes de qualité

2 Query Formulation

3 Source Discovery & Selection  
Évaluation de la qualité

4 Data Acquisition  
Récriture des requêtes  
Amélioration itérative de précision et rappel

5 Data Integration  
Partage et réutilisation des mappings

6 Cube Présentation  
Mise en valeur des éléments importants  
Filtrage en fonction de la qualité

Reuses & Sharing

Requêtes  
Sources  
Mappings  
Cubes

eric J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 13

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Le décisionnel en poche

Utilisateur non expert

- Analyses à la demande
- Requêtage façon moteur de recherche
- Drill-beyond
- Quasi temps réel
- Collaboratif

Technos

- Entrepôts de données / OLAP
- Web sémantique
- Gestion de documents
- Fouille de données
- Ingénierie sociale

Problèmes ouverts

Modèle de données avancé, algèbre de requêtes  
Equilibre entre partage et confidentialité des données  
Formalisation de l'intelligence collaborative, interfaces  
Modèle économique

eric J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 14

eric ENTREPÔTS, REPRÉSENTATION  
de l'INGÉNIEURIE des CONNAISSANCES

## Section 2

### Sécurité des Données dans le Nuage

Collaboration avec V. Attasena, N. Harbi et G. Gavin

Lyon 1 UNIVERSITÉ LUMIÈRE LYON 2 UNIVERSITÉ DE LYON is

eric J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 15

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Jusqu'ici, tout va bien

Transfert réseau

Fournisseur d'accès

ED

Transfert réseau

eric J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 16

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Plus dure sera la chute

Cloud computing

Cloud service provider policies

Control & modification policies

Grid technology

Virtual machine technology

Virtual network technology

Established on network

Intruders

Inside-Intruders

Outside-Intruders

Service provider staffs & other customers

Accidental plan

Intentional plans

Service, network, electrical failure...

Delete unmodify & unaccess data

Data loss & damage

Service down

Data loss & damage

Data transfer bottlenecks

Service down

Data alteration & damage

Data pilfering

Data alteration & damage

Data pilfering

Data availability

Data integrity

Data privacy

Cloud security issues

Transparent: V. Attasena

eric J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 17

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Solution

TRUST NO ONE

eric J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 18

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Comment faire ?

Synchronous: Data replication, Data verification  
 Asynchronous: Data anonymization, Data encryption  
 Semi-synchronous: Secret sharing

k-anonymization, l-diversity, Inner code verifying, Outer code verifying  
 Homomorphic encryption (HE): Partially HE, Fully HE  
 Secret sharing: Secret sharing (SS), Multi-secret sharing (MSS), Verify secret sharing (VSS)

A Data Availability, I Data Integrity, P Data Privacy, A Data Analysis  
 Transparent: V. Attasena

e<sup>2</sup>ic J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 19

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Partager pour mieux cacher

Adil Shamir, "How to share a secret", Communications of the ACM, 22(11), 1979

- Partage :  $t$  points  $\Rightarrow$  polynôme  $f$  de degré  $t-1$
- Secret = terme constant du polynôme
- $n$  partages  $f(i)$

Figure V. Attasena  $n = 6, t = 4$

- Reconstruction : interpolation de Lagrange
- Un seul polynôme  $p(i)$  tel que  $\text{degré}(p(i)) < t$  et  $p(i) = f(i)$
- Secret =  $p(0)$

e<sup>2</sup>ic J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 20

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Repartons dans le nuage

Utilisateur ED, Fournisseur 1 ED<sub>1</sub>, Fournisseur 2 ED<sub>2</sub>, Fournisseur n ED<sub>n</sub>

Volume de données

e<sup>2</sup>ic J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 21

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### 1<sup>re</sup> approche : bpVSS

Shares & Outer signatures, CSP<sub>1</sub>, CSP<sub>n</sub>, Shared table 1, Shared table n, Original table, Secret sharing, Inner signature, Outer signatures, Linear equations, Shares, Inner signature, Homomorphic function, Digits of base-p Integer, Decimal Integer  $\Rightarrow$  Base-p Integer, Integer D

Transparent: V. Attasena

e<sup>2</sup>ic J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 22

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Reconstruction

Shares & Outer signatures, Verify shares, CSP<sub>1</sub>, CSP<sub>n</sub>, Select  $t$  from  $n$  shares, Shared table 1, Shared table n, Secret sharing, Original table, Linear equations, Digits of base-p Integer & Inner signature, Base-p Integer  $\Rightarrow$  Decimal Integer, Integer D

Transparent: V. Attasena

e<sup>2</sup>ic J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 23

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

Transparent: V. Attasena

### Partageons une base de données

$n=5, t=4, p=7$

Original data	ProNo	ProName	ProDescr	CategoryID	UnitPrice	CSP
124	Shirt	Red	1	142		
125	Shoe	ALL	2	142		
126	Ring	ALL	1	142		

Shares at CSP1, Shares at CSP2, Shares at CSP3, Shares at CSP4, Shares at CSP5

Shares at CSP1	ProNo	ProName	Sig-ProName	ProDescr	Sig-ProDescr	CategoryID	UnitPrice	Sig-UnitPrice
124	[31,10,21,36,16]	[4,1,1,1,1]	[2,7,26,22]	[2,1,2]	1	142	0	
125	[31,10,17,26]	[1,0,2,1,1]	NULL	NULL	2	142	0	
126	[27,21,13,6]	[2,1,3,1,1]	NULL	NULL	1	142	0	

Shares at CSP2, Shares at CSP3, Shares at CSP4, Shares at CSP5

Shares at CSP2	ProNo	ProName	Sig-ProName	ProDescr	Sig-ProDescr	CategoryID	UnitPrice	Sig-UnitPrice
124	[40,27,10,28,21]	[0,2,0,3,1]	[3,2,27,20]	[3,2,0]	1	142	1	
125	[40,27,31,27]	[0,2,1,2]	NULL	NULL	2	142	2	
126	[33,10,24,20]	[3,0,4,0]	NULL	NULL	1	142	2	

Shares at CSP3, Shares at CSP4, Shares at CSP5

Shares at CSP3	ProNo	ProName	Sig-ProName	ProDescr	Sig-ProDescr	CategoryID	UnitPrice	Sig-UnitPrice
124	[30,13,20,28,23]	[0,3,0,3,3]	[39,27,22]	[4,2,2]	1	142	0	
125	[30,13,19,17]	[0,3,4,2]	NULL	NULL	2	142	0	
126	[29,20,38,13]	[4,3,3,2]	NULL	NULL	1	142	0	

Shares at CSP4, Shares at CSP5

Shares at CSP4	ProNo	ProName	Sig-ProName	ProDescr	Sig-ProDescr	CategoryID	UnitPrice	Sig-UnitPrice
124	[44,14,21,40,22]	[4,4,1,0,2]	[39,27,22]	[4,2,2]	1	142	0	
125	[44,14,23,27]	[4,4,3,2]	NULL	NULL	2	142	0	
126	[39,21,18,9]	[4,1,3,4]	NULL	NULL	1	142	0	

Shares at CSP5

Shares at CSP5	ProNo	ProName	Sig-ProName	ProDescr	Sig-ProDescr	CategoryID	UnitPrice	Sig-UnitPrice
124	[34,16,27,42,22]	[4,4,2,2,2]	[30,32,28]	[0,3,3]	1	142	0	
125	[34,16,23,32]	[3,1,3,2]	NULL	NULL	2	142	0	
126	[30,27,19,12]	[0,2,4,2]	NULL	NULL	1	142	0	

e<sup>2</sup>ic J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 24

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Accès aux données partagées

Transparent : V. Attasena

```

SELECT P.ProdName, P.ProdNo,
       C.CategoryName, C.CategoryID,
       SUM ( P.UnitPrice * Ptax )
FROM   Product AS P
JOIN   Category AS C
ON     P.CategoryID = C.CategoryID
GROUP BY P.CategoryID
    
```

J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 25

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Cloud cubes

	Time attributes			Product attributes		Aggregation attributes		Signature attributes	
	YearID	MonthID	DateID	CategoryID	ProdNo	TotalPrice	Number	Sig_TPrice	Sig-Number
#1	ALL	ALL	ALL	ALL	ALL	83231	58244	38	32
#2	ALL	ALL	ALL	1	1	26701	18254	23	54
#3	ALL	ALL	ALL	1	1	8958	7113	80	23
#4	ALL	ALL	ALL	1	2	4348	1844	82	23
	ALL	ALL	ALL	...	...	...	...	...	...
#5	1	1	ALL	ALL	ALL	44574	54542	32	23
#6	1	1	1	ALL	ALL	21158	8954	45	97
#6	1	1	1	ALL	ALL	9754	4544	23	67

Transparent : V. Attasena

J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 26

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### 2<sup>e</sup> approche : fVSS

Volume--  
Confidentialité++  
Disponibilité++  
Prix--

Transparent : V. Attasena

J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 27

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Comment ça marche ?

ProNo	ProName	ProDescr	CategoryID	UnitPrice
124	Shirt	Red	1	135
125	Shoe	NULL	2	142
126	Ring	NULL	1	142

Shares at CSP1

ProNo	ProName	ProDescr	CategoryID	UnitPrice
124	{46,2;72,1;78,3;84,4;86,9}	{44,9;68,4;67,1}	1	-110,3
125	{81,5;91,9;95,3;90,4}	NULL	2	110,7
126	{104,8;139,8;147,5;136,8}	NULL	1	196,2

Shares at CSP2

ProNo	ProName	ProDescr	CategoryID	UnitPrice
124	{94,6;121,4;122,6;134,1;136,7}	{93,3;117,5;116,3}	1	160,9
125	{59,1;80,7;87,9;77,6}	NULL	2	119,7

Shares at CSP3

ProNo	ProName	ProDescr	CategoryID	UnitPrice
125	{52,6;32,0;25,1;35,0}	NULL	2	5,3
126	{431,1;786,5;863,8;755,6}	NULL	1	-1158,7

Shares at CSP4

ProNo	ProName	ProDescr	CategoryID	UnitPrice
124	{-37,1;57,9;-58,8;-67,8;-69,7}	{36,1;54,9;-53,9}	1	-86,5

Shares at CSP5

ProNo	ProName	ProDescr	CategoryID	UnitPrice
124	{-37,1;57,9;-58,8;-67,8;-69,7}	{36,1;54,9;-53,9}	1	-86,5

Transparent : V. Attasena

J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 28

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Nous contre le reste du monde

Data privacy			Data availability	Data integrity		Ability in case CSPs fail	
Data transferring	Data Storing	CSP group cheating		Inner code verification	Outer code verification	Querying shares	Updating shares
All approaches	All approaches	Only fVSS	All Approaches	Only [21][58][60] bpVSS, fVSS	Only bpVSS, fVSS	All approaches	Only fVSS

Features	[56]	[54]	[55][57]	[59]	[60]	[58]	[21]	bpVSS & fVSS
Data Types	Positive Integers, Character, Strings,	Positive integers	Integers	Positive integers	Positive integers	Positive integers	Positive integers	Integer, Reals, Characters, Strings, Dates, Booleans
Shared data access			Yes	Yes	Yes	Yes	Yes	Yes
Updates	-	-	Yes	Yes	Yes	Yes	Yes	Yes
Exact match queries	-	Yes	Yes	Yes	Yes	-	Yes	Yes
Range queries	-	Yes	Yes	Yes	Yes	-	Yes	Yes
Aggregation queries	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes
Grouping queries	-	-	-	-	-	-	-	Yes

Transparent : V. Attasena

J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 29

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

Transparent : V. Attasena

### Expérimentons !

Approaches:

- Thompson et al.'s [58]
- Hadavi et al.'s [60]
- bpVSS
- fVSS-I (Balance share volume strategy)
- fVSS-II (Unbalance share volume strategy)

Parameter settings:

$n = 5, t = 4, |signature| = |data|/2$

Virtual machines:

Approaches	CSP1	CSP2	CSP3	CSP4	CSP5	Index server
Thompson	M3	M3	M3	M3	M3	M3
Hadavi	M3	M3	M3	M3	M3	M3
fVSS-I	M3	M3	M3	M3	M3	-
bpVSS	M3	M3	M3	M3	M3	M3
fVSS-II	M3	M3	M2	M1	M1	M3

Benchmark:

Star Schema Benchmark (SSB)  
SSB database: 757 MB

Measurements:

- Storage volume
- Data sharing time
- Data reconstruction time
- Data access time
- Data transfer volume

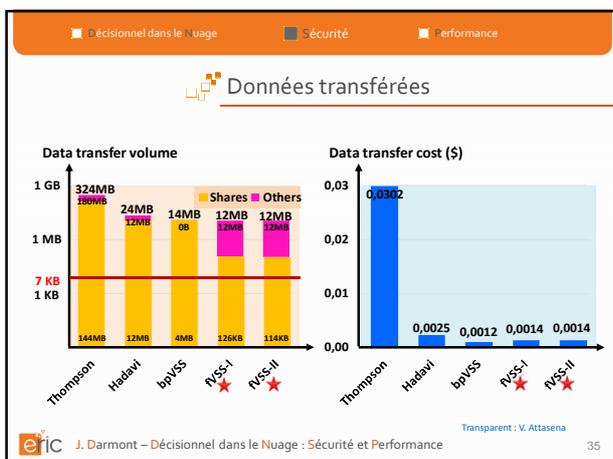
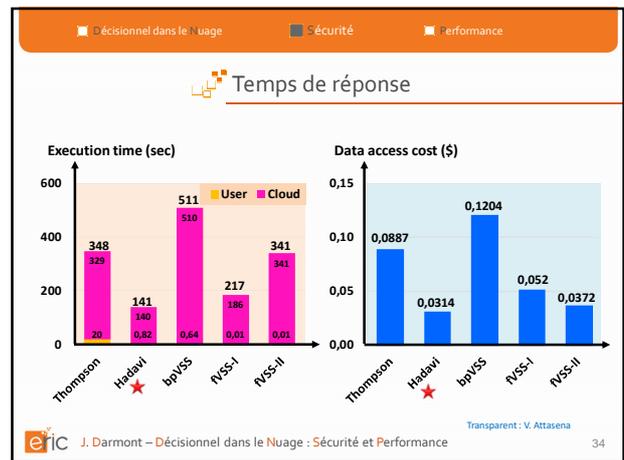
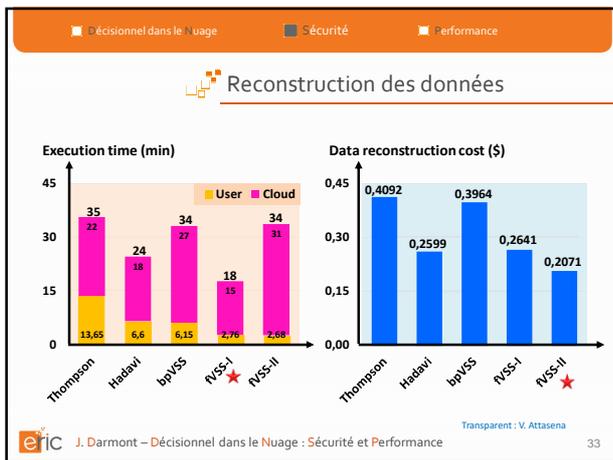
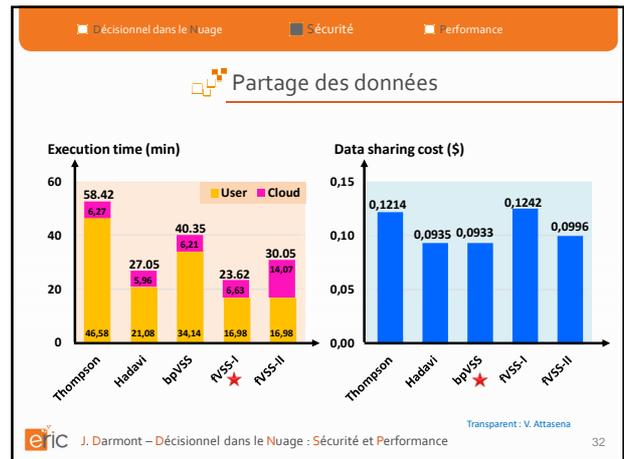
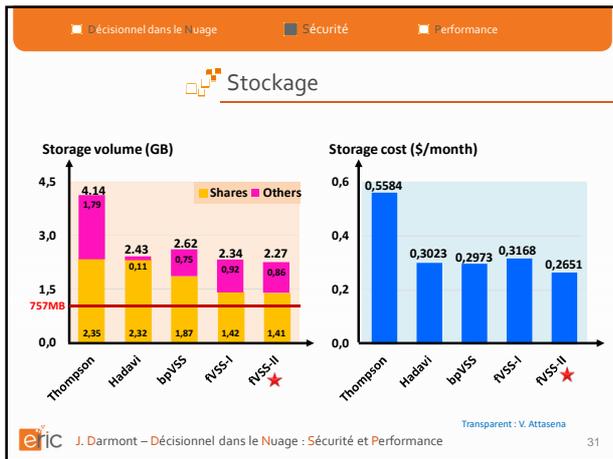
Monetary costs

Share volume assignment:

Number of shared records (% of original)

Approaches	CSP1	CSP2	CSP3	CSP4	CSP5
Thompson	100%	100%	100%	100%	100%
Hadavi	100%	100%	100%	100%	100%
bpVSS	60%	60%	60%	60%	60%
fVSS-I	81%	81%	60%	39%	39%
fVSS-II	81%	81%	60%	39%	39%

J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 30



**eric** ENTREPÔTS, REPRÉSENTATION  
de l'INGÉNIERIE des CONNAISSANCES

# Section 3

## Performance de l'Accès aux Données stockées dans le Nuage

Collaboration avec B. Bachelet, S. Bimonte et L. d'Orazio

UNIVERSITÉ LYON 1    UNIVERSITÉ LYON 2    UNIVERSITÉ DE LYON    **is**

37

■ Décisionnel dans le Nuage    ■ Sécurité    ■ Performance

## Payer plus pour gagner plus... de perfs

L'infonuagique d'un point de vue économique

Élasticité      Paiement à la demande

**eric** J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance      38

■ Décisionnel dans le Nuage    ■ Sécurité    ■ Performance

## Soyons terre à terre !

Fournisseur d'accès

- Index
- Caches
- Vues matérialisées
- Partitionnement
- ...

$Coût_{global} = Coût_{transfert} + Coût_{calcul} + Coût_{stockage}$

**eric** J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance      39

■ Décisionnel dans le Nuage    ■ Sécurité    ■ Performance

## Problématique et contributions

Sélection de vues à matérialiser      Optimisation multicritère

- Modèles de tarification flexibles
- Modèles de coût pour la matérialisation des vues
- Modèle détaillé du processus d'optimisation

**eric** J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance      40

■ Décisionnel dans le Nuage    ■ Sécurité    ■ Performance

## Coût de transfert

$$C_t(D, Q, A) = C_t^-(D, Q) + C_t^+(A)$$

ascendant      descendant

$$\approx C_t^+(A)$$

EC2

Volume	Coût
0 Go – 1 Go	0
1 Go – 10 To	0,12 \$ / Go
10 To – 40 To	0,09 \$ / Go

- D : Ensemble de données
- Q : Charge de requêtes
- A : Réponse aux requêtes

**eric** J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance      41

■ Décisionnel dans le Nuage    ■ Sécurité    ■ Performance

## Coût de calcul

$$C_c(Q, IC) = \sum_{i=1}^{n_Q} \sum_{j=1}^{n_{IC}} t(Q_i, IC_j) \times c_c(IC_j)$$

Temps de traitement      Coût de location

- Q = {Q<sub>i</sub>} / i = 1..n<sub>Q</sub>: Charge de requêtes
- IC = {IC<sub>j</sub>} / j = 1..n<sub>IC</sub>: Configuration d'instances de calcul

**eric** J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance      42

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Coût de stockage

S3	
Volume	Coût
0 To - 1 To	0,140 \$ / Go
1 To - 450 To	0,125 \$ / Go

$$C_S(D) = \sum_{k=1}^{n_D} \frac{\text{Coût de stockage}}{\text{Taille des données}} \times \frac{\text{Durée de stockage}}{t(D_k)}$$

- $D = \{D_k\} / k = 1..n_D$ : Données stockées par périodes de temps

e<sup>2</sup>iC J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 43

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Coût de calcul avec vues matérialisées

Temps de traitement

$$C_c(Q, V, IC) = T(Q, V) \times c_c(IC_0) \times n_{IC}$$

Coût de location

$$T(Q, V) = T_{proc}(Q, V) + T_{mat}(V) + T_{maint}(V)$$

Exécution requêtes    Matérialisation    Maintenance

- $Q$ : Charge de requêtes
- $V$ : Ensemble de vues matérialisées
- $IC$ : Configuration d'instances de calcul

e<sup>2</sup>iC J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 44

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Coût de stockage avec VM

Coût de stockage    Durée de stockage

$$C_S(D, V) = c_s(s(D) + s(V)) \times t$$

Taille des données

- $D$ : Ensemble de données
- $V$ : Ensemble de vues matérialisées

e<sup>2</sup>iC J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 45

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Processus d'optimisation

Sélection de vues (algorithme existant)

Programme linéaire

e<sup>2</sup>iC J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 46

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Problèmes d'optimisation

Trouver un ensemble de vues matérialisées  $V \subseteq V_{cand}$

■  $MV_1$

- Minimiser  $T_{proc}$
- Contrainte:  $C \leq C_{max}$

■  $MV_2$

- Minimiser  $C$
- Contrainte:  $T_{proc} \leq T_{max}$

■  $MV_3$

- Minimiser  $\alpha \times T_{proc} + (1 - \alpha) \times C$

e<sup>2</sup>iC J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 47

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Environnement expérimental

Star Schema Benchmark

- Données : 5,5 Go
- 4 séries de requêtes
- 2 Go RAM
- 8 Go disque
- Hadoop 0.20.2
- Pig 0.9.1
- Quadri-pros 800 MGHz
- 96 Go RAM

e<sup>2</sup>iC J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 48

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Paramètres

Période d'expérimentation : 1-24 [12] (mois)

Fréquence de la charge : 1-5 [4] (par semaine)

Nombre de nœuds : 5-20 [10]

eric J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 49

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Résultats expérimentaux

Gain de performance : 110 %

Gain de coût : 30 %

eric J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 50

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Et ensuite ?

Modèle de tarification

Sélection des vues  
Optimisation des coûts

eric J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 51

eric ENTREPÔTS, REPRÉSENTATION & INGÉNIÉRIE des CONNAISSANCES

Pour aller... plus haut !

- Cloud Intelligence Workshop [eric.univ-lyon2.fr/cloud-i/](http://eric.univ-lyon2.fr/cloud-i/)
- Cloud Computing Research Group [liris.cnrs.fr/cloud/wiki](http://liris.cnrs.fr/cloud/wiki)
- Thèse de S. Sobati Moghadam
- Projet ANR **BI4people**

Lyon 1 UNIVERSITÉ LUMIÈRE LYON 2 UNIVERSITÉ DE LYON

eric J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 52

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### Confidentialité de données décisionnelles partagées dans le nuage

Travail basé sur CryptDB  
Réalisé avec S. Sobati Moghadam et G. Gavin

css.csail.mit.edu/cryptdb/

eric J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 53

■ Décisionnel dans le Nuage ■ Sécurité ■ Performance

### BI4people

eric IIRIT Institut de Recherche en Informatique de Toulouse LI Université de Lyon elico TRIMANE

eric J. Darmont – Décisionnel dans le Nuage : Sécurité et Performance 54