Basse consommation dans les systèmes embarqués

Daniel CHILLET

ENSSAT - Université de Rennes 1
INRIA/IRISA - Projet R2D2
6 Rue de Kérampont - F-22300 LANNION

chillet@enssat.fr

UNIVERSITE DE RENNES 1

INRIA

IRISA

Plan

IRISA

- Evolution technologique
- Pourquoi y-a-t il consommation ?
- Puissance statique & dynamique
- Qu'est ce qui consomme ?
- Solutions :
  - au niveau technologique
  - au niveau logique
  - au niveau architectural
  - au niveau logiciel (non traitées ici)
  - au niveau système (non traitées ici)
- Conclusions

Evolutions technologiques

IRISA

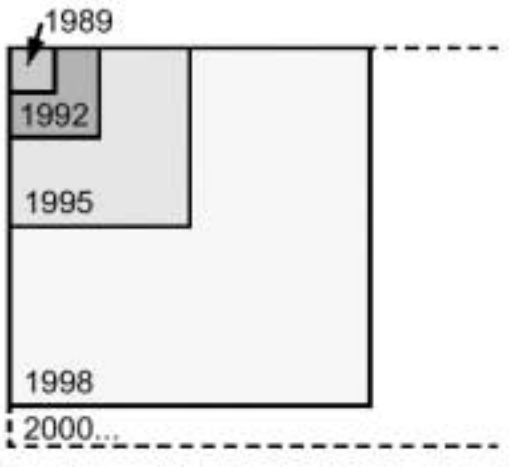
Année d'introduction	1997	1999	2001	2003	2006	2009	2012
Technologie (µm)	0,25	0,18	0,15	0,13	0,1	0,07	0,05
Fréquence (MHz)	750	1250	1500	2100	3500	6000	10000
Mémoire		256M	1G	4G	16G	64G	256G
Nombre de transistors	11M	21M	38M	77M	202M	520M	1350M
Coût par M Tr (Cent.)	500	290	166	97	42	18	8

Evolutions technologiques

IRISA

- *Scaling* technologique à chaque génération
  - Fréquence augmente de 43%
  - Capacité totale et tension d'alimentation sont réduites de 30%
  - Énergie réduite de 65%
    - $E = C*Vdd^2 = C'*0.7 * (Vdd'*0.7)^2 = 0.35*C'*Vdd'^2 = 35\% E'$
  - Puissance réduite de 50%
    - $P = f*C*Vdd^2 = 1.43*f * 0.35*C'*Vdd'^2 = 50\% P'$
    - On considère que l'activité du circuit est constante

Mais à nombre de transistors constant!



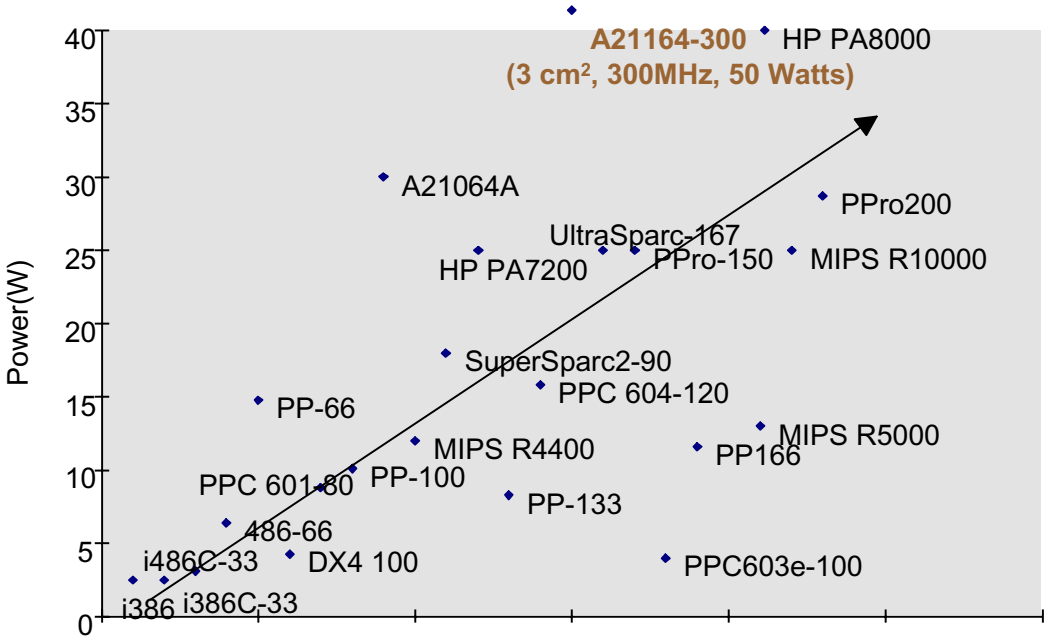
  - Densité de transistor double à chaque génération
  - Surface des puces augmente de 25%

  - Densité de puissance augmente avec un facteur 2
  - Courant d'alimentation augmente de façon importante

Effet sur la consommation

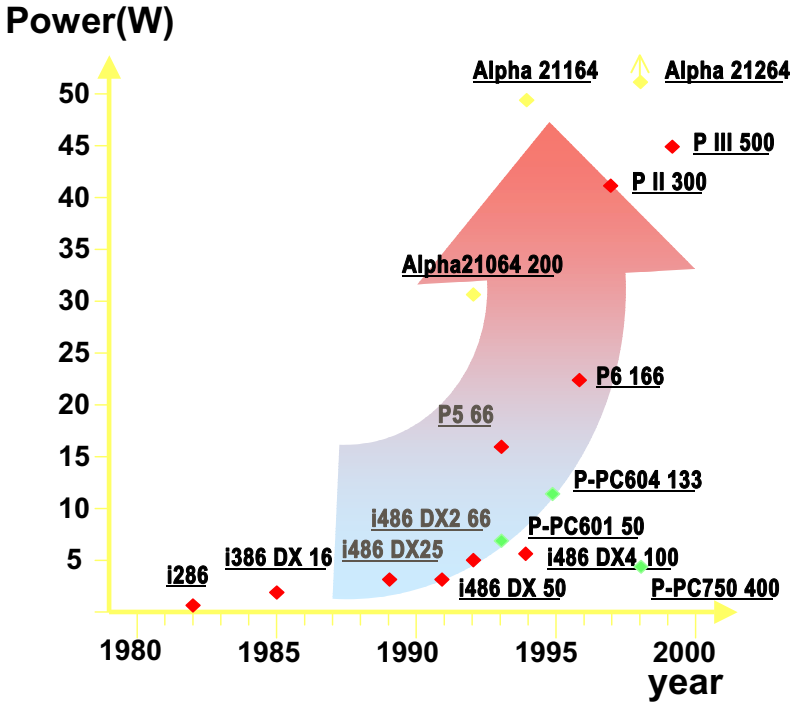
IRISA

- x4 tous les 3 ans



Evolution de la consommation

IRISA



Evolution de la consommation

IRISA

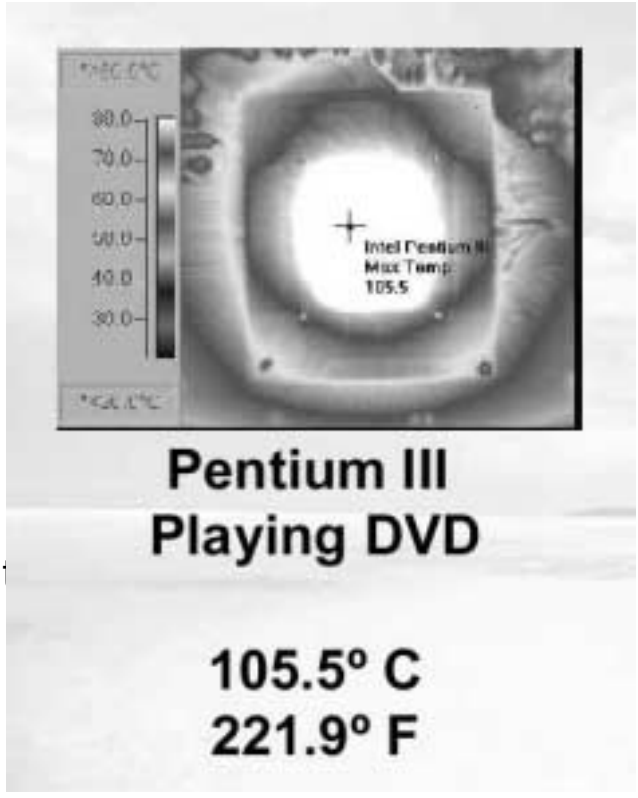
e.g. Famille DEC/Compaq [Herrick99]

- EV4
  - 200 MHz @ 3.3V
  - 30W @ 200 MHz
  - 1,7 Million Transistors
  - 233 mm²
- EV5 (21164)
  - 350 MHz @ 3.3V
  - 60W @ 350 MHz
  - 9,3 Million Transistors
  - 298 mm²
- EV6 (21264)
  - 575 MHz @ 2.2V
  - 90W @ 575 MHz
  - 15,2 Million Transistors
  - 314 mm²
- EV7 (21364)
  - > 1000 MHz @ 1.5V
  - 100W @ 1GHz
  - ~100 Million Transistors
  - ~350 mm²
- EV8
  - > 1-2 GHz (0.125 micron)
  - <150W
  - ~250 Million Transistors

Effets sur les circuits

IRISA

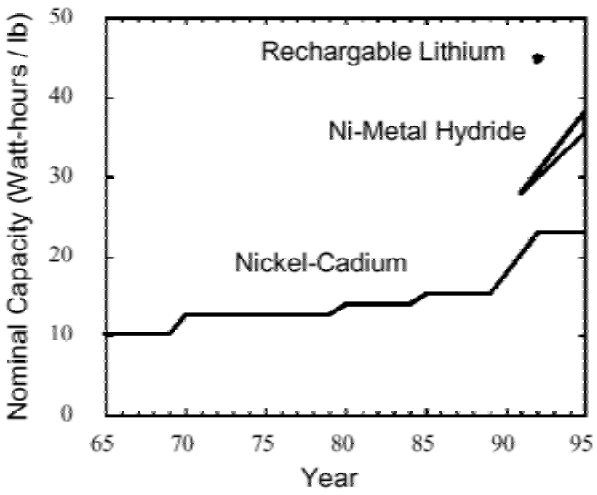
- Dissipation chaleur :
  - diminution de la fiabilité
    MTBF/2 pour une augmentation de 10°C
- Augmentation du coût de la mise en œuvre
  - 1€/W si >40W
- Augmentation volume
  - radiateur, ventilateur (volume sonore)
- Diminution du temps de fonctionnement ou augmentation du poids des batteries

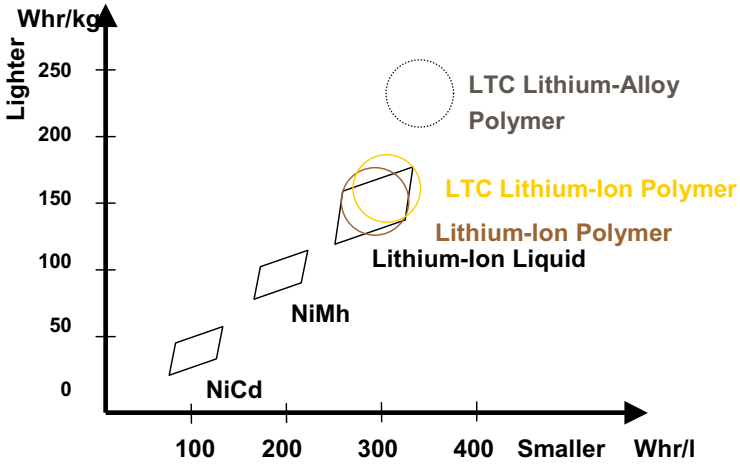


Evolution de la technologie des batteries

IRISA

- Evolution techno des batteries





[P. Senn 2000]

Evolution de la technologie des batteries

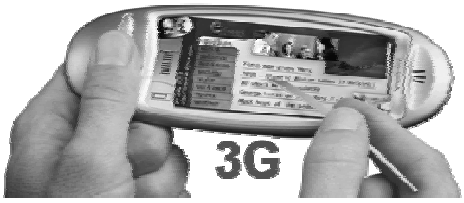
IRISA

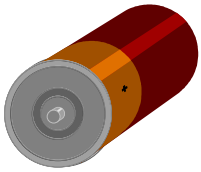
Techno	NiCd Nickel Cadium	NiMh Nickel Metal Hybride	Li-ion Lithium Ion	Li-poly
Apparition	1956	1990	1992	1996
Tension délivrée (V)	1,2	1,2	3,6	3,7
Epaisseur (mm)	> 6	>6	>6	3
Capacité (Whr/Kg)	30-50	60-90	70-140	115-140
Durée de vie (cycles)	~1000	~ 1000	~ 500	~ 500
Extralopation pour 1kg à 20 W	22 – 27 h	40 – 45 h	50 – 55 h	65 – 70 h

Evolution des applications

IRISA

- Systèmes :
  - capacité de traitement élevée :
    - multimédia (visioconf, Mpeg, …)
    - interface conviviale (reco vocale, stylo, synthèse vocale, …)
    - sécurité des info (cryptographie, authentification, …)
  - portables
  - faible coût
  - peu volumineux
  - peu lourd
  - peu bruyant
  - longue durée d'utilisation sans rechargement des batteries





Pourquoi il y a consommation ?

IRISA

- Puissance = puissance statique + puissance dynamique

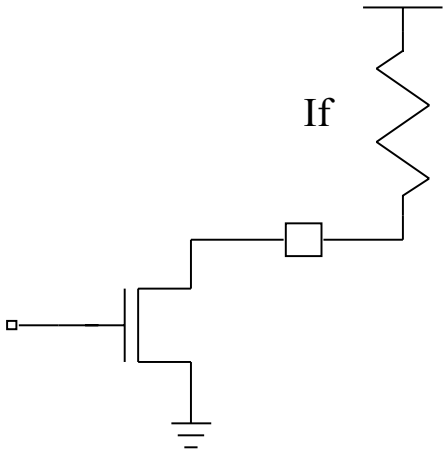
- Puissance statique :
  - même bloqué, les transistors ne sont pas complètement fermé

- Puissance dynamique :
  - lors de la commutation des transistors, il va se créer des court-circuits entre Vdd et la masse
  - charge et décharge de capacités

Puissance statique

IRISA

- Courants de fuite (Subthreshold Leakage Current)
  - Même lorsque Vgs < Vt les transistors MOS ne sont pas équivalent à un interrupteur fermé
  - Si on diminue Vdd (#Vt) alors ces courants peuvent devenir significatifs
- Courants inverses de diode Source/Drain - Substrat

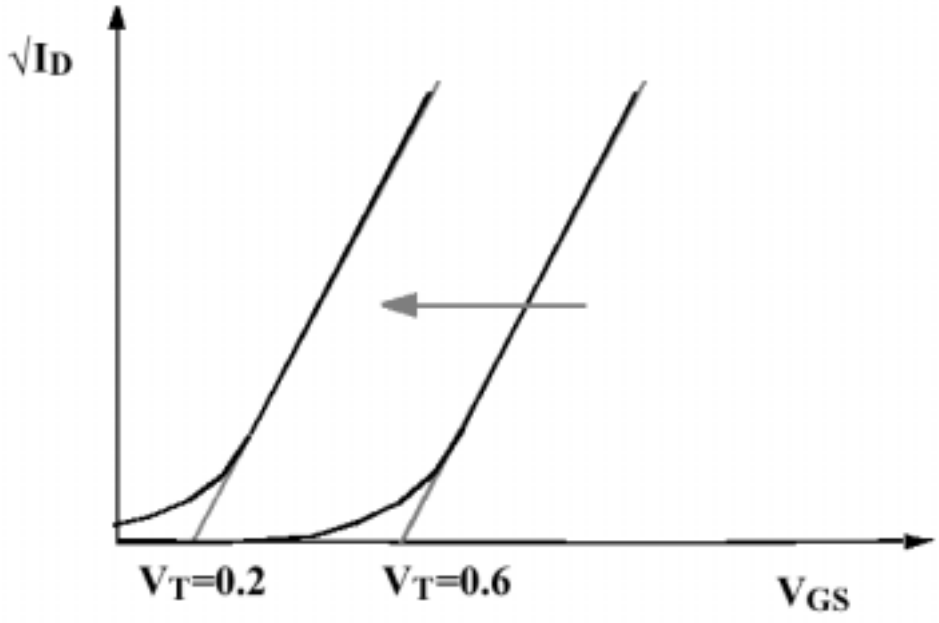


⇨ Négligeables tant que Vdd > Vt

Puissance statique

IRISA

- Impact de la tension de seuil :
  - diminution tension de seuil ➔ augmentation du courant de fuite



  - donc, pour un design low power, il vaut mieux une tension de seuil élevé

Puissance statique

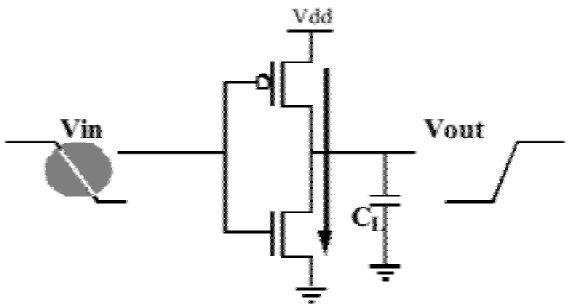
IRISA

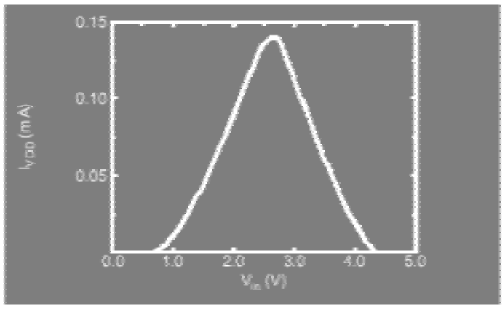
- Négligeable ou pas ?
  - souvent négligée pour les blocs logiques ayant beaucoup d'activité
  - mémoires de grandes tailles, peu d'activité (relativement au nombre de transistor) ➔ puissance non négligeable
  - techno sub-micronique profonde :
    - augmentation de l'influence des interconnexions :
      - charge supplémentaire pour les portes logiques

Puissance dynamique

IRISA

- Apparition lors de la commutation des transistors :
  - Courants de court-circuit : Icc
    - Chemin de court circuit pendant la commutation des structures logiques
    - Pour les circuits *conçus correctement* ≈ 15%

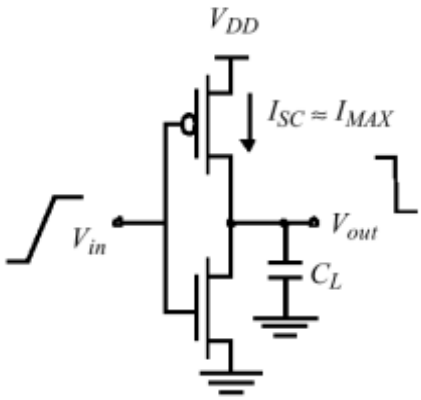




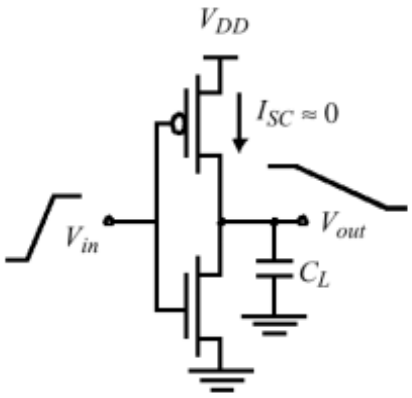
Puissance dynamique

IRISA

- Courants de court-circuit : impact des fronts lents :
  - temps de court circuit plus long
  - dépend de la charge de la porte



Capacité faible

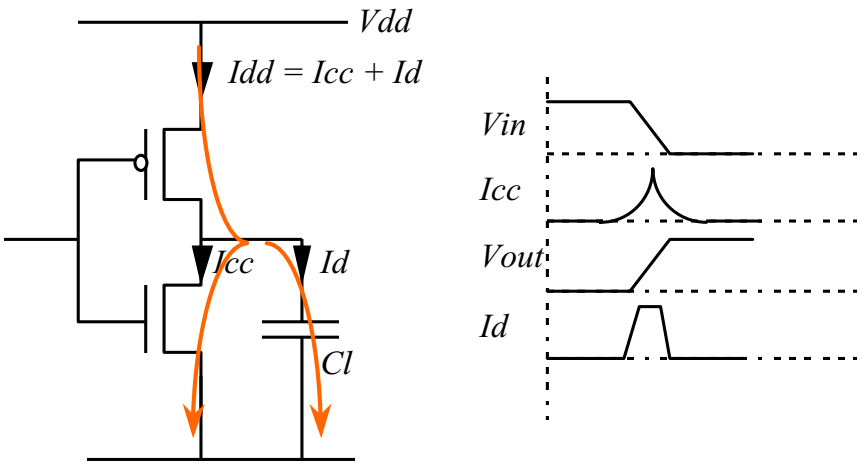


Capacité élevée

Puissance dynamique

IRISA

- Courants de charge/décharge des capacités : Id



Puissance dynamique

IRISA

- Energie par transition = $C_l \bullet Vdd^2$
- Puissance = Energie par transition • Vitesse de transition
  = $C_l \bullet Vdd^2 \bullet f$
- La puissance n'est pas fonction de la taille des transistors de la cellule considérée

$$\begin{aligned} P_{dyn} &= C_l \bullet Vdd^2 \bullet f_{0\to1} \\ &= C_l \bullet Vdd^2 \bullet f \bullet P_{0\to1} \\ &= \alpha \bullet C_l \bullet Vdd^2 \bullet f \\ &= C_{EFF} \bullet Vdd^2 \bullet f \end{aligned}$$

- C_{EFF} **est la capacité effective =** $C_l \bullet P_{0\to1}$

Puissance → ***Dépend de l'activité*** → ***Dépend des données***

Puissance dynamique

IRISA

- e.g. porte inverseuse :

A	OUT
0	1
1	0

Fonction NOT

P(0 ->1) = P(OUT = 0) • P(OUT = 1)
= 1/2 • 1/2 = 1/4

αC = Ceff = 1/4. C

Puissance dynamique

IRISA

- e.g. porte NOR statique à 2 entrées

A	B	OUT
0	0	1
0	1	0
1	0	0
1	1	0

Fonction NOR

Hypothèse : P(A=1) = 1/2 et P(B=1) = 1/2

P(OUT=1) = 1/4
P(0 ->1) = P(OUT = 0) • P(OUT = 1)
= 3/4 • 1/4 = 3/16

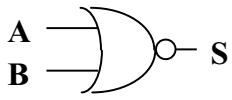
$$\alpha C = \text{Ceff} = 3/16.\ Cl$$

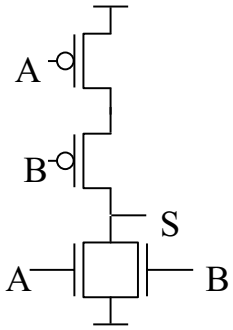
Puissance dynamique

IRISA

- Probabilités de transition d'une porte NOR

A	B	S
0	0	1
0	1	0
1	0	0
1	1	0



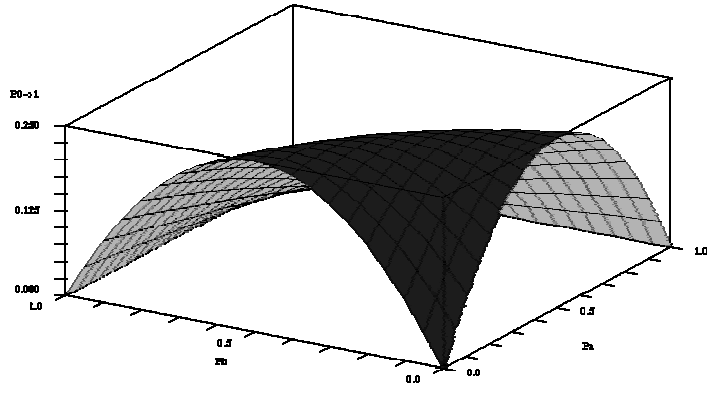


$P_S = (1-P_A)(1-P_B)$ avec $P_A=P(A=1)$ et $P_B=P(B=1)$

$P_1=P(S=1)=(1-P_A)(1-P_B)$

$P_{0\to1} = P_0.P_1 = P(S=0) \cdot P(S=1) = (1-P(S=1)) \cdot P(S=1)$

$= (1-(1-P_A)(1-P_B))\,(1-P_A)(1-P_B)$



- α est une fonction dépendant fortement de la statistique des signaux

Puissance dynamique

IRISA

- Probabilités de transition d'une porte XOR

A	B	S
0	0	0
0	1	1
1	0	1
1	1	0

$P_S = P_A + P_B - 2.P_A P_B$ avec $P_A=P(A=1)$ et $P_B=P(B=1)$

$P_1=P(S=1)=(1-P_A)(1-P_B)$

$P_{0\to1} = P_0.P_1 = P(S=0) \cdot P(S=1) = (1-P(S=1)) \cdot P(S=1)$

$= (1- (P_A + P_B - 2 \cdot P_A \cdot P_B)). (P_A + P_B - 2 \cdot P_A . P_B)$

$= (1 -(½ + ½ - 2. ½ . ½))(½ + ½ - 2. ½ . ½)$

$= (1 -(1 - ½))(1 - ½) = ½ . ½ = ¼$

Puissance dynamique

IRISA

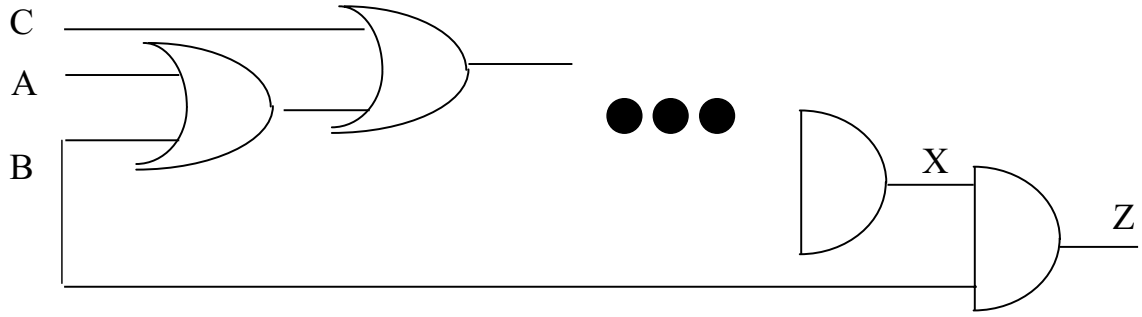
- Fonction de la statistique des entrées

	$P_{0\to1}$	Ceff
AND	$(1-P_AP_B)\ P_AP_B$	3/16
OR	$(1-P_A)(1-P_B)\ (1-(1-P_A)(1-P_B))$	3/16
XOR	$(1-(P_A+P_B-2\ P_AP_B))\ (P_A+P_B-2\ P_AP_B)$	1/4
NOR	$(1-(1-P_A)(1-P_B))\ (1-P_A)(1-P_B)$	3/16

Puissance dynamique

IRISA

- Problème de la reconvergence
  - devient très rapidement complexe



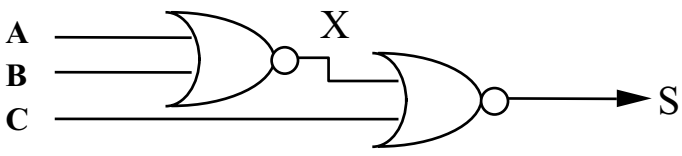
$$P(Z=1) = P(B=1) \bullet P(X=1 \mid B=1)$$

Puissance dynamique

IRISA

- Problème des glitchs
  - non pris en compte dans les probabilités

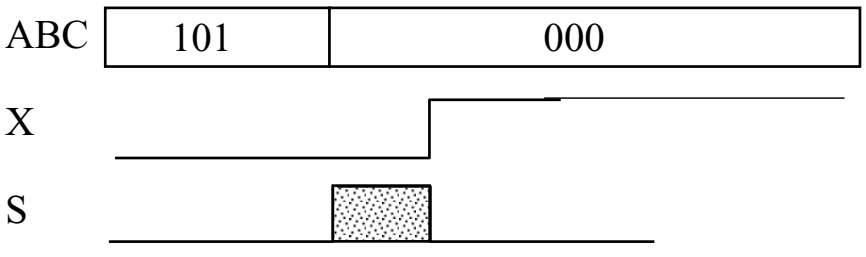
P(A) = ½
P(B) = ½
P(C) = ½



P(X =1) = 1/4
P(S = 1) = 1/2 * 3/4 = 3/8

αX = P(X=0) . P(X=1)
= (1-P(X=1)) . P(X=1)
= (1 – 1/4) . 1/4
= 3/4

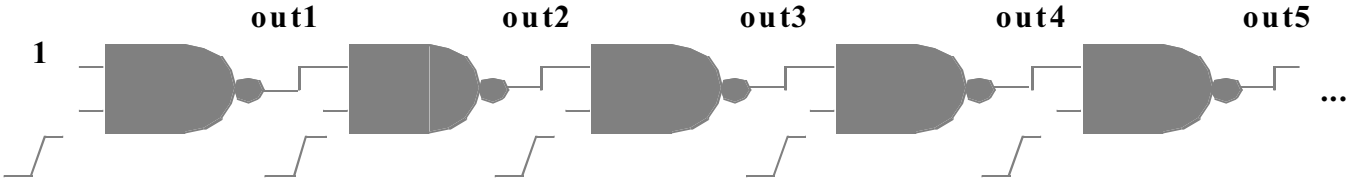
αS = P(S=0) . P(S=1)
= (1 – P(S=1)) . P(S=1)
= (1 – 3/8) . 3/8 = 5/8 .3/8
= 15/64

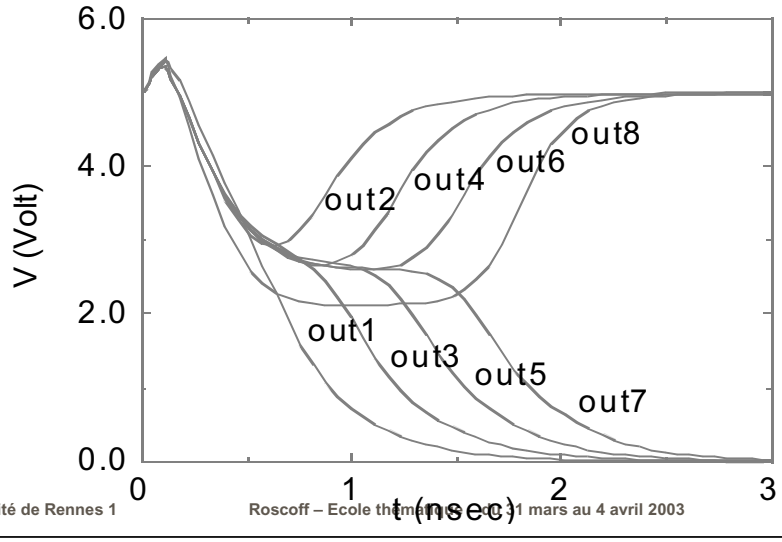


Puissance dynamique

IRISA

- Influence de la propagation (glitch) :
  - si les sorties sont à 1 et que les entrées basculent de 0 à 1

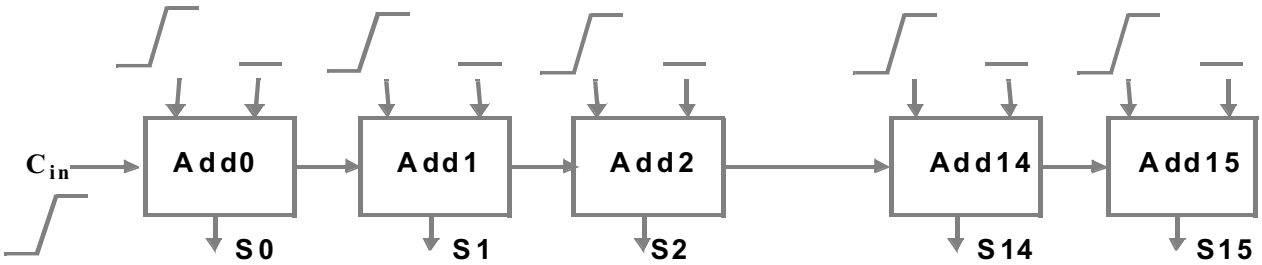


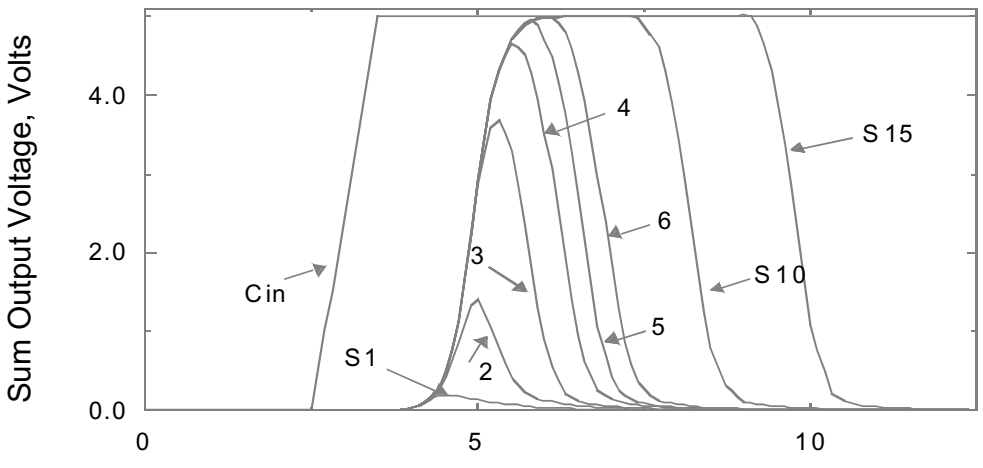


Puissance dynamique

IRISA

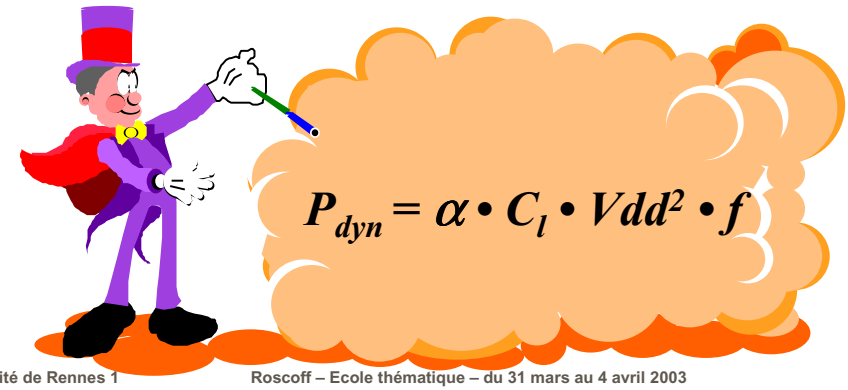
- Influence de la propagation (suite) :
  - les entrées de l'additionneur passent à 1 ainsi que la retenue





Bilan

- Puissance de charge : Pc
  - Charge et décharge des capacités du circuit
- Courants de court-circuit : Pcc
  - Chemin de court circuit pendant la commutation des structures logiques statiques
- Courants de fuite : Pf
  - Jonctions, fonctionnement sous le seuil



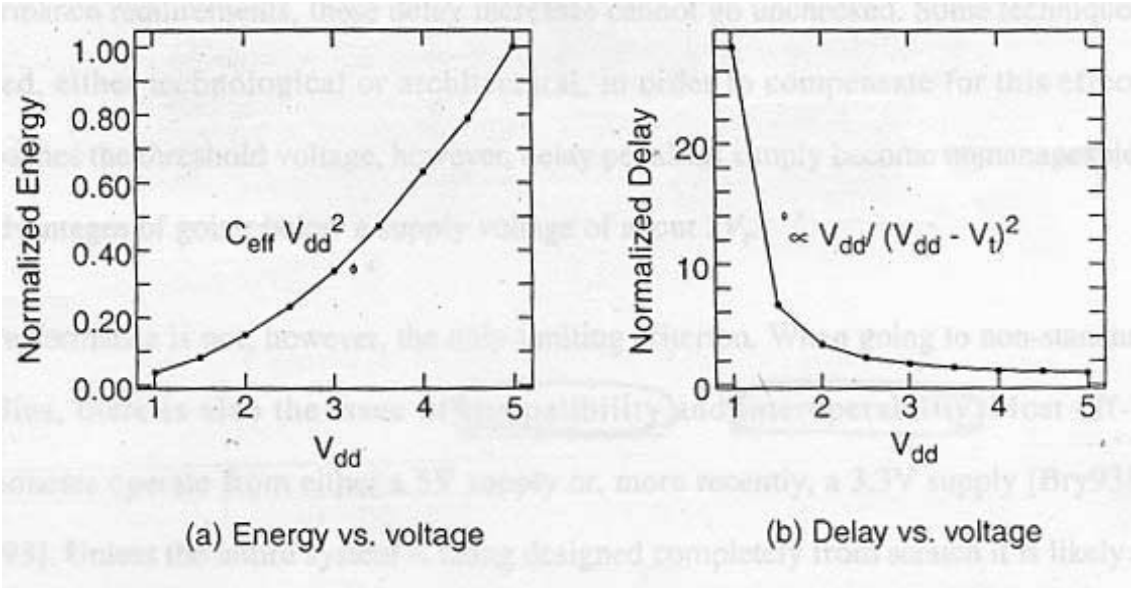
Degré de liberté pour diminuer la conso

- La puissance dépend de 4 paramètres :
  - fréquence
  - activité
  - capacité effective
  - tension d'alimentation (influence quadratique !!!)
- On peut jouer sur les 4 paramètres :
  - attention, ils ne sont pas orthogonaux
  - attention, ils ne peuvent être optimisés indépendamment les uns des autres

Dépendances entre les paramètres

IRISA

- Exemple :
  - tension versus délai



Dépendances entre les paramètres

IRISA

- Exemple :
  - capacité effective versus activité

$$\alpha C = \mathbf{Ceff}$$

Qu'est ce qui consomme ?

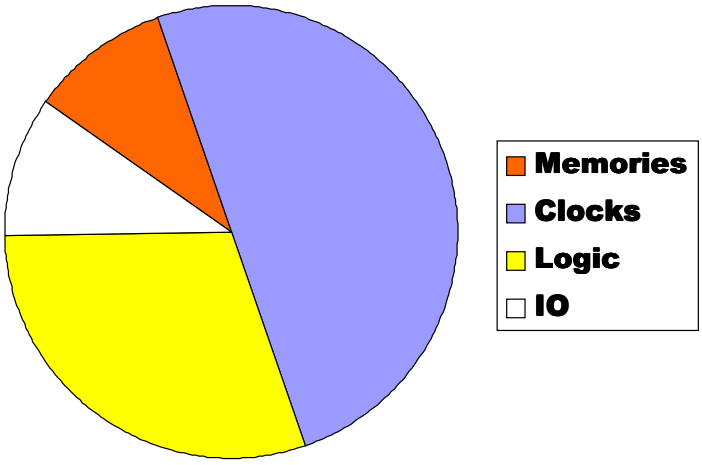
IRISA

- La logique
- Les opérateurs
- La mémoire
- Les interconnexions
- L'horloge
- L'écran
- Les disques
- …
- …

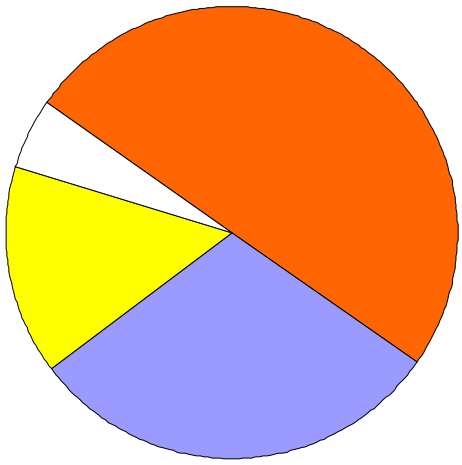
?

Qu'est ce qui consomme ?

IRISA



DEC Alpha uP

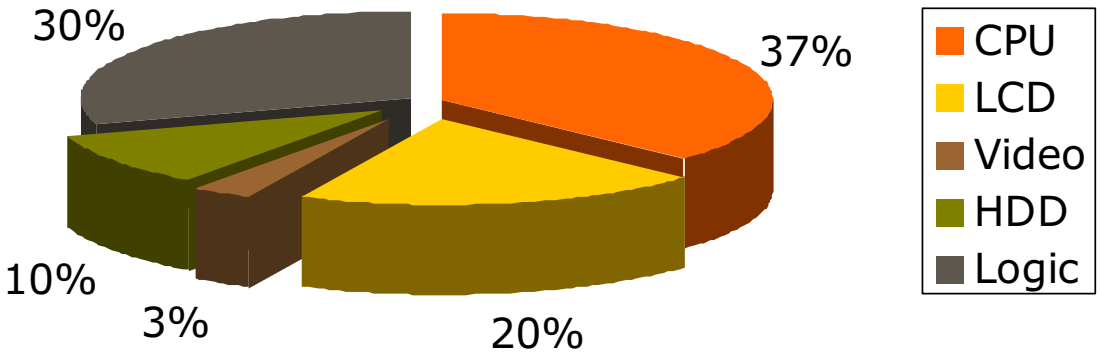


Telecom chip

Qu'est ce qui consomme ?

IRISA

- PC Portable
  - Puissance totale (application Word) : 19.1 W



[Source Hitachi]

Qu'est ce qui consomme ?

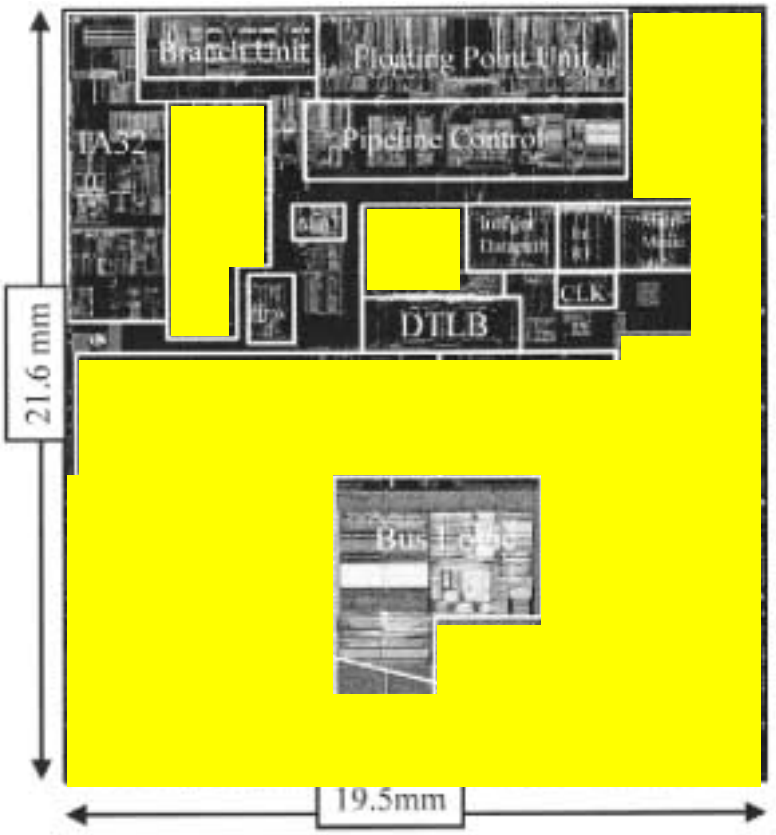
IRISA

- Le CPU :
  - les unités fonctionnelles évidemment
  - le contrôleur :
    - contrôle de l'exécution des instructions :
      - lecture, décodage, exécution, rangement de résultat
  - les caches qui sont de plus en plus intégrés dans le chip

Qu'est ce qui consomme ?

IRISA

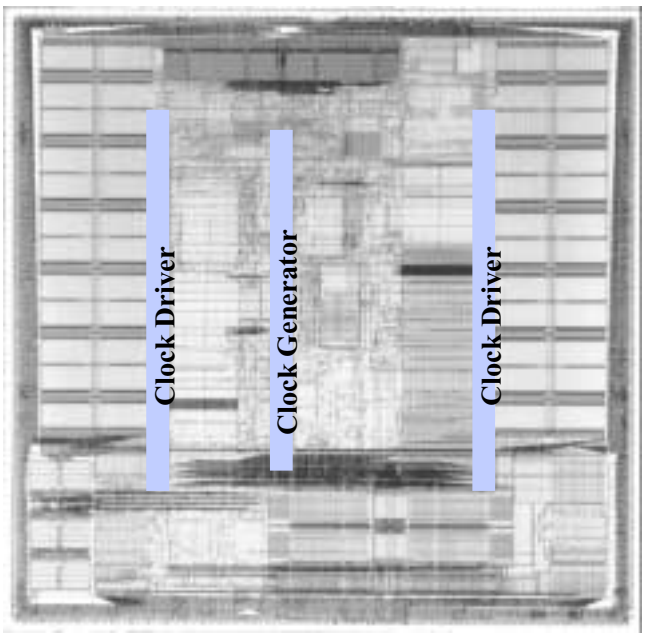
- La mémoire
  - interne (cache, scratch memory)
  - externe (passage par les entrées sorties du chip)
  - exemple de l'Itanium 2
    - environ 60 à 70 % de la surface consacrée à la mémoire cache



Qu'est ce qui consomme ?

IRISA

- L'horloge :
  - 40-50% de la puissance dissipée par l'arbre d'horloge

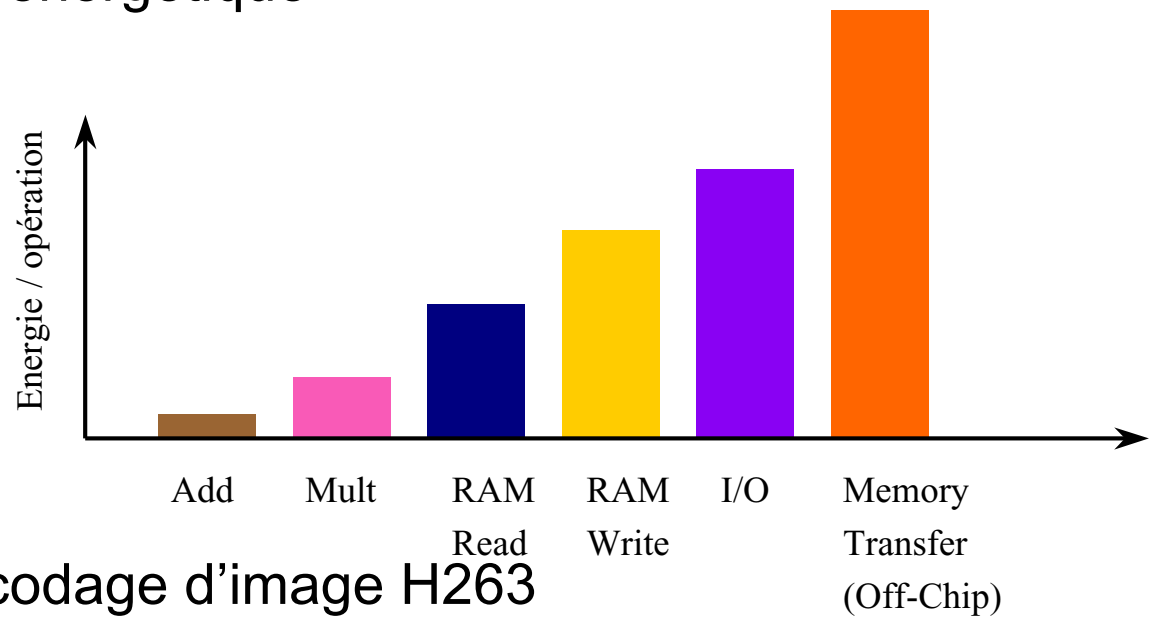


DIGITAL Corp.
Processeur alpha 21164
1995
2.5M Portes
9.3M Transistors
298 mm^2
300 MHz
64/128 Bits
0.5µ
60W @ 3.3V

Qu'est ce qui consomme ?

IRISA

- Bilan énergétique



- E.g. codage d'image H263

  1500K Opérations et 500K Transferts mémoire

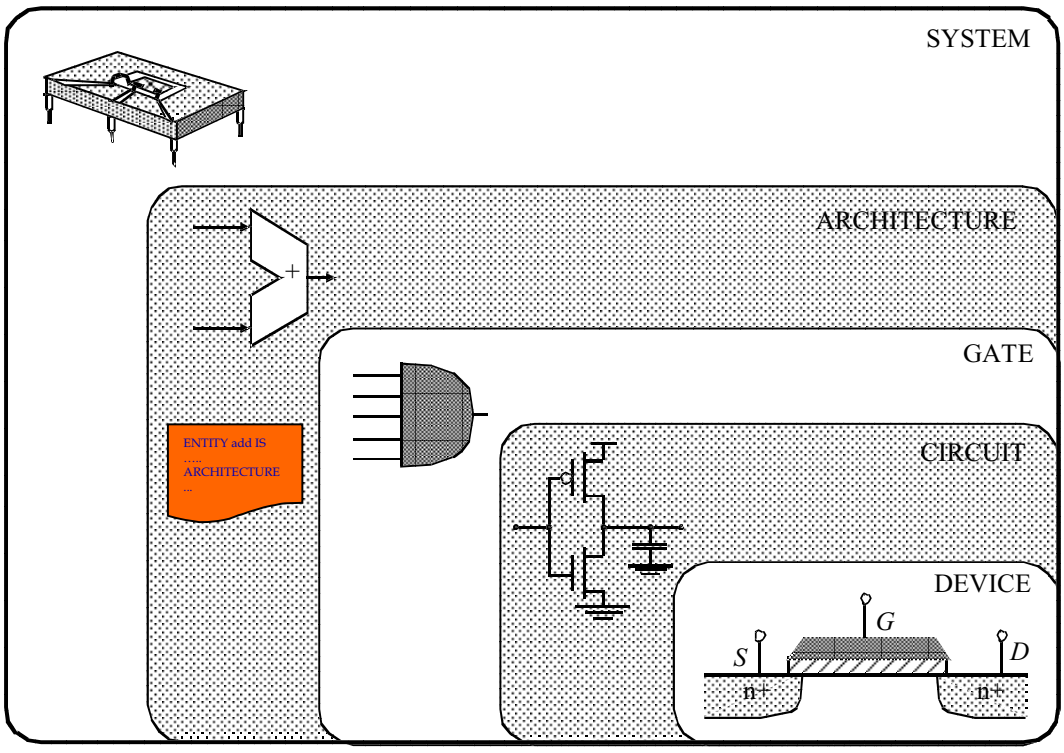
  Energie transfert = 33 . Energie Opération

  -> Puissance mémoire # 10 . Puissance opérations

Où intervenir ?

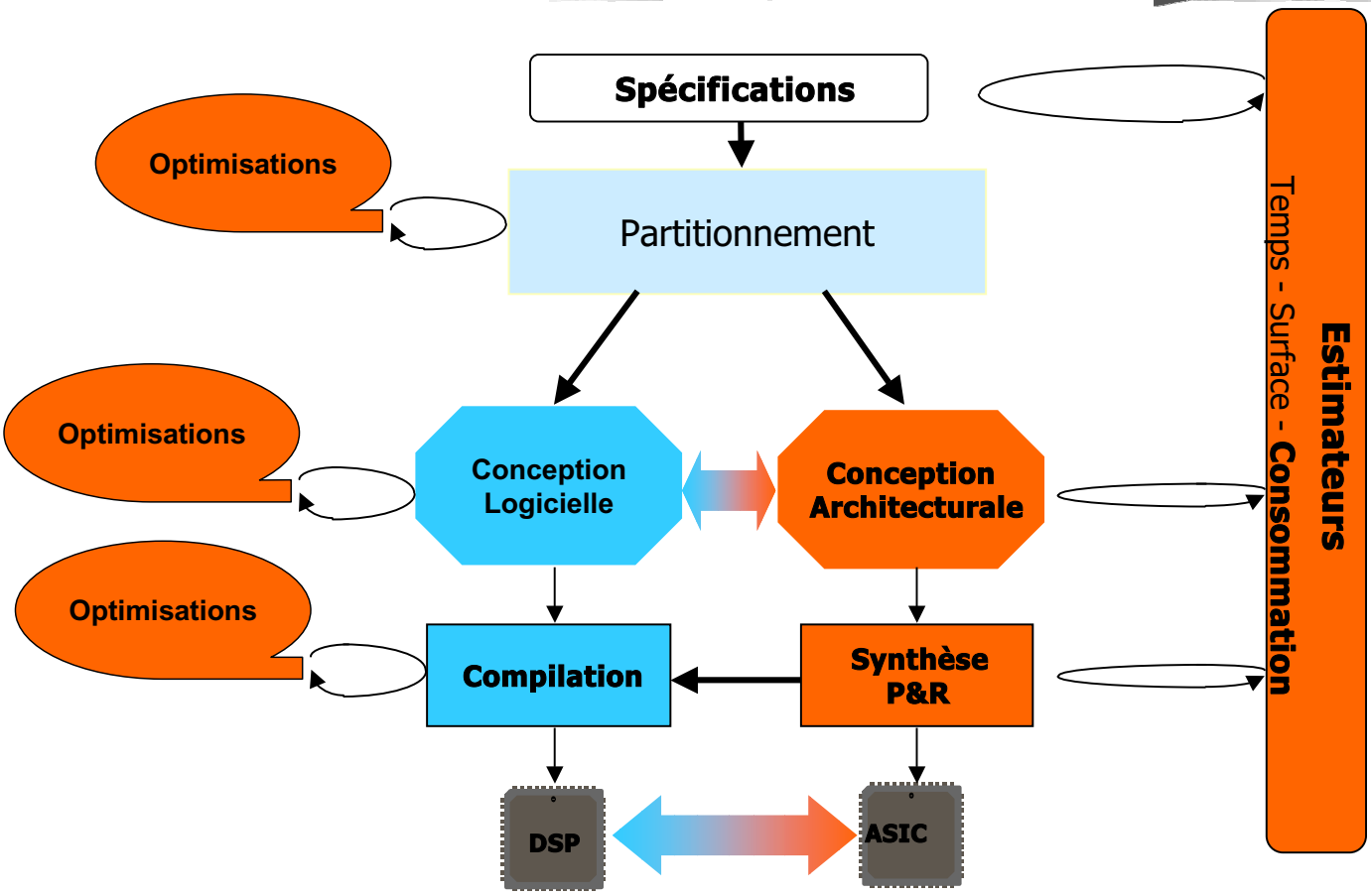
IRISA

- Interventions possibles à tous les niveaux



Optimisations & estimations

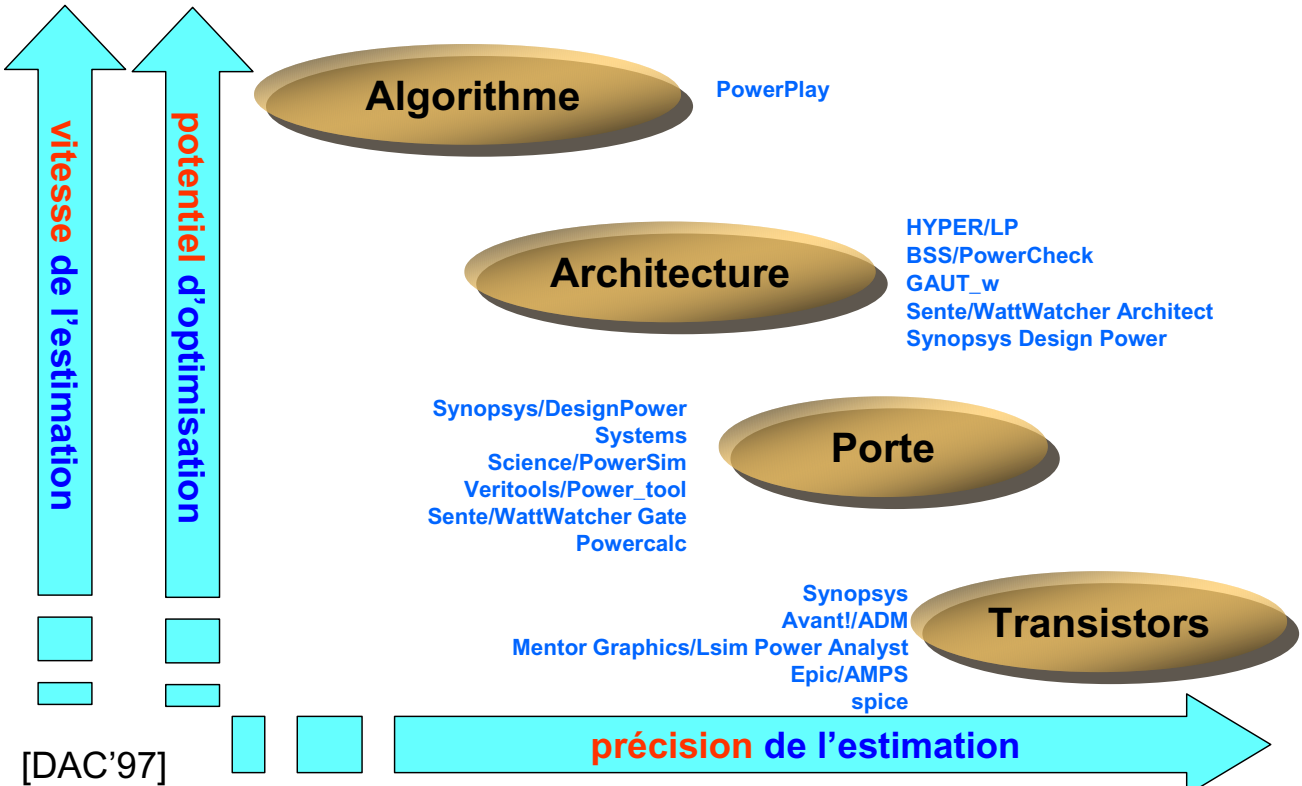
IRISA



Optimisations & estimations

IRISA

- Besoins en estimateurs



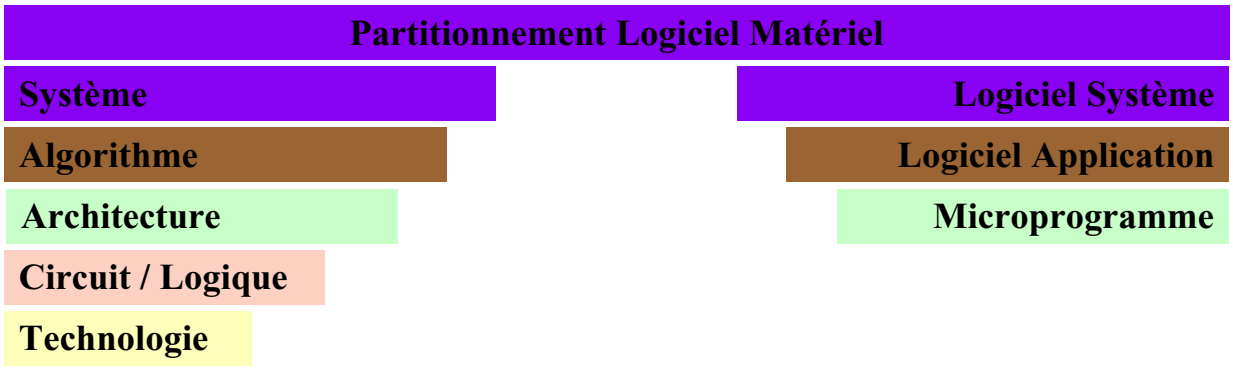
[DAC'97]

Optimisations & estimations

IRISA

- Où gagner en consommation ?

Partitionnement logiciel/matériel 10-95%



Système et Algorithme 30-95%
Architecture 10-90%
Logique 15%
Physique 20%
Technologie 30%

Logiciel 10-80% (sur le système)

Optimisations : au niveau technologique

IRISA

- Dimensionnement des transistors :

Capacité →

Transistors de petite taille

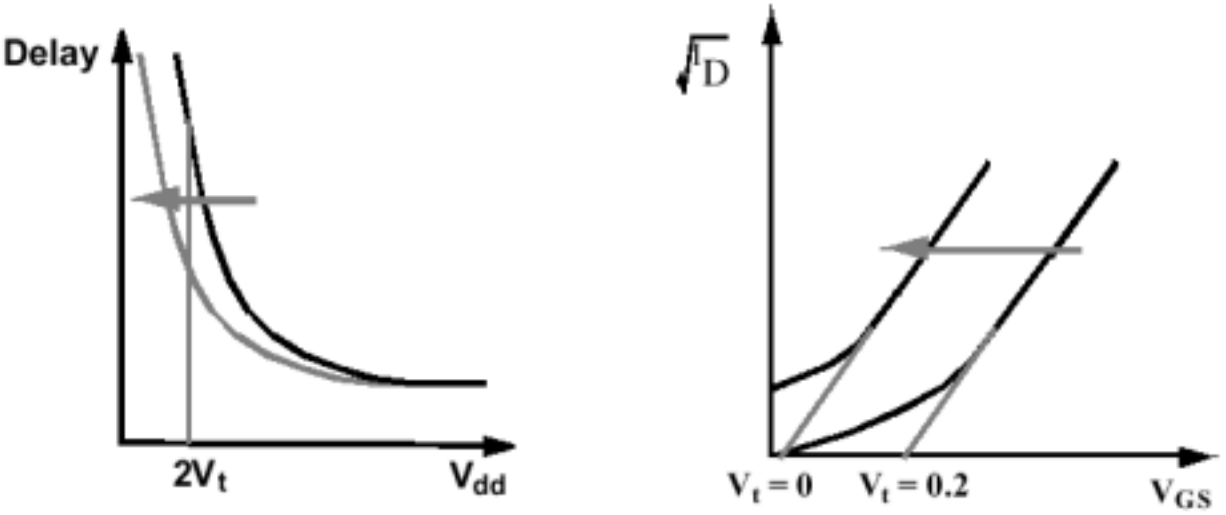
Transistors de taille importante

← Tension d'alimentation

Optimisations : au niveau technologique

IRISA

- Augmentation de la tension de seuil des transistors :
  - diminution du courant de fuite :
    - impact sur la puissance statique et dynamique



Optimisations : au niveau technologique

IRISA

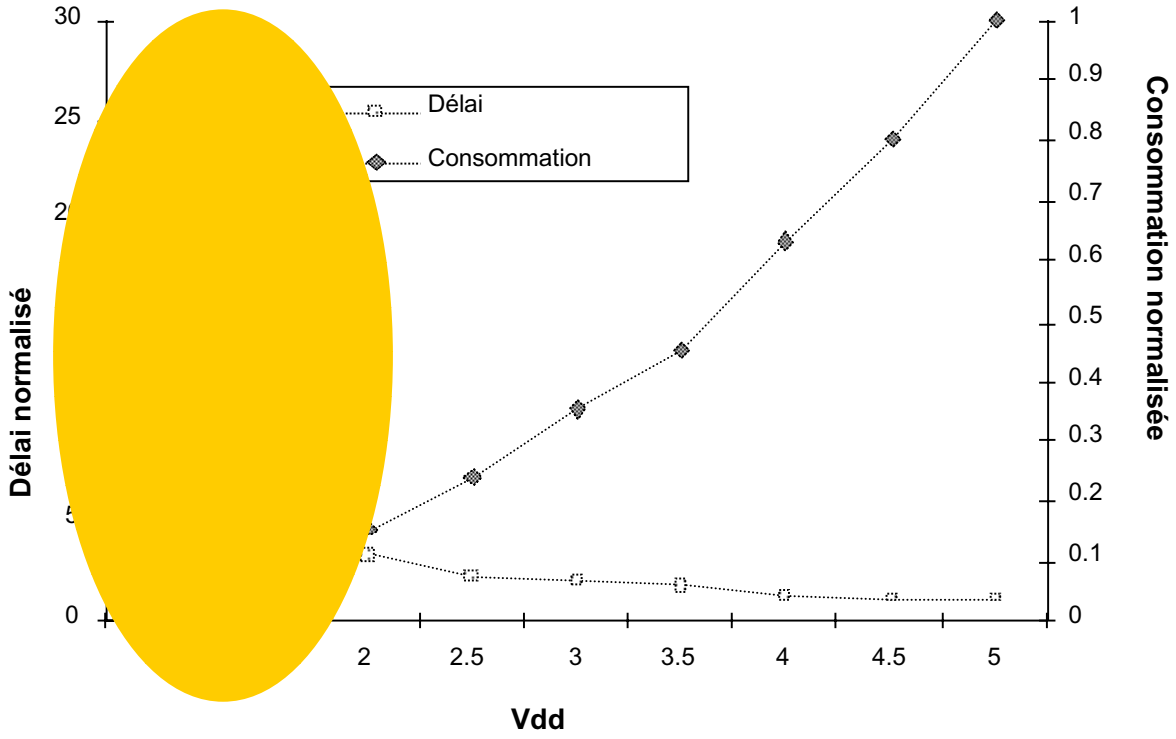
- Abaisser la tension d'alimentation :
  - intervient de façon quadratique
  - de tout le circuit lorsque celui-ci est inactif ou en sous utilisation
  - d'une partie du circuit lorsque celle-ci est inutilisée ou non critique

  - disponibilité de plusieurs tensions d'alimentation
  - gestion dynamique de la tension
  - attention aux interfaces avec le monde extérieur

Optimisations : au niveau technologique

IRISA

- Abaisser la tension d'alimentation : suite

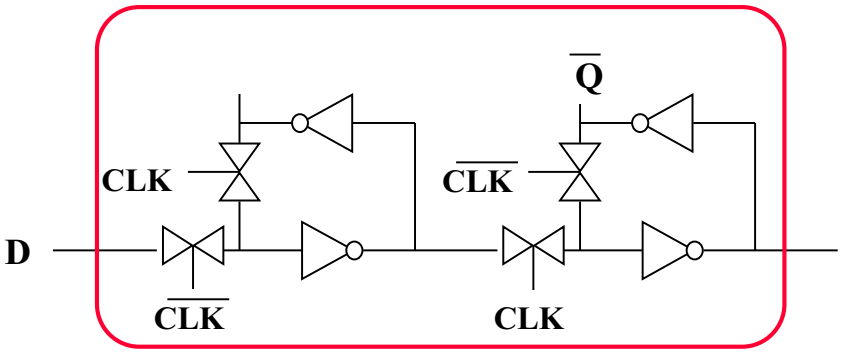


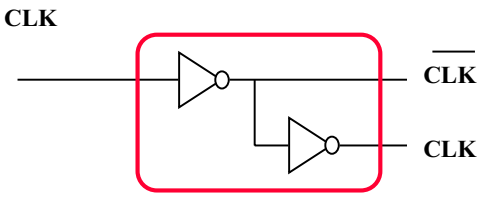
Proche de la tension de seuil

Optimisations : au niveau logique

IRISA

- Gated clock :
  - la puissance d'une bascule dépend :
    - de l'activité de la bascule : transitions de 0 à 1
    - mais même sans activité, la bascule consomme





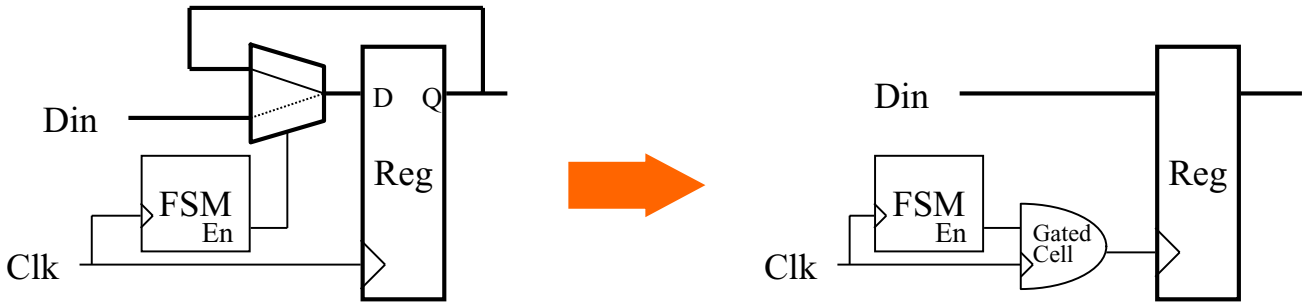
  - puissance : $P_{basc}=\alpha P_{dyn}+P_{horl}$

  - objectif du gated clock : $P_{basc}=\alpha(P_{dyn}+P_{horl})$

Optimisations : au niveau logique

IRISA

- Gated clock suite :
  - Supprimer les commutations des blocs qui n'ont pas besoin d'être rechargés

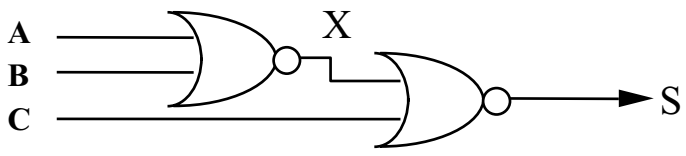


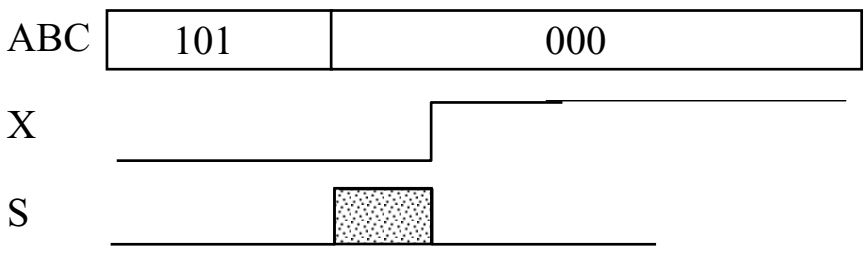
  - Modification de la machine d'états
  - Gain important
  - Mais... Circuit non totalement synchrone

Optimisations : au niveau logique

IRISA

- Glitch (hasards dynamiques)
  - **Transitions dynamiques parasites**
  - Consommation majeure (ET INUTILE) de courant

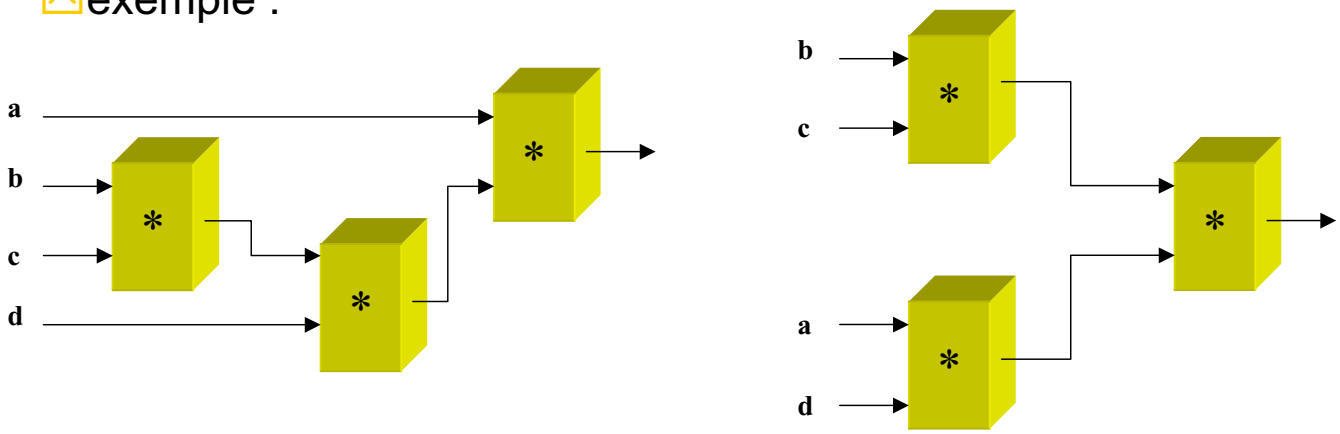




Optimisations : au niveau logique

IRISA

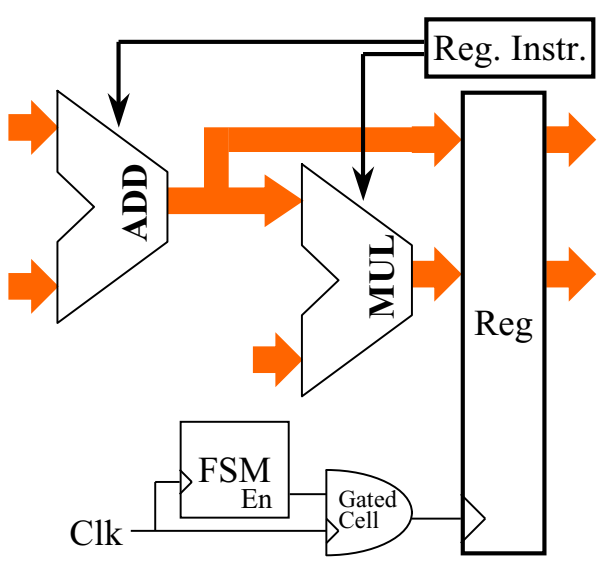
- Elimination des glitchs :
  - par réorganisation des fonctions combinatoires lorsque cela est possible :
    - limitation de la longueur du chemin :
      - stabilité de la sortie obtenue plus rapidement
      - moins de commutations
    - limitation de la capacité effective
  - exemple :



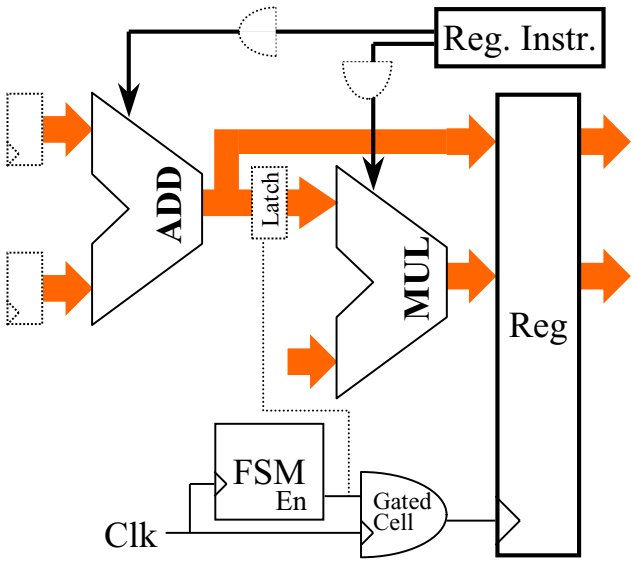
Optimisations : au niveau logique

IRISA

- Activer les UF uniquement lorsque cela est nécessaire



Le multiplieur consomme même si ses sorties sont inutilisées



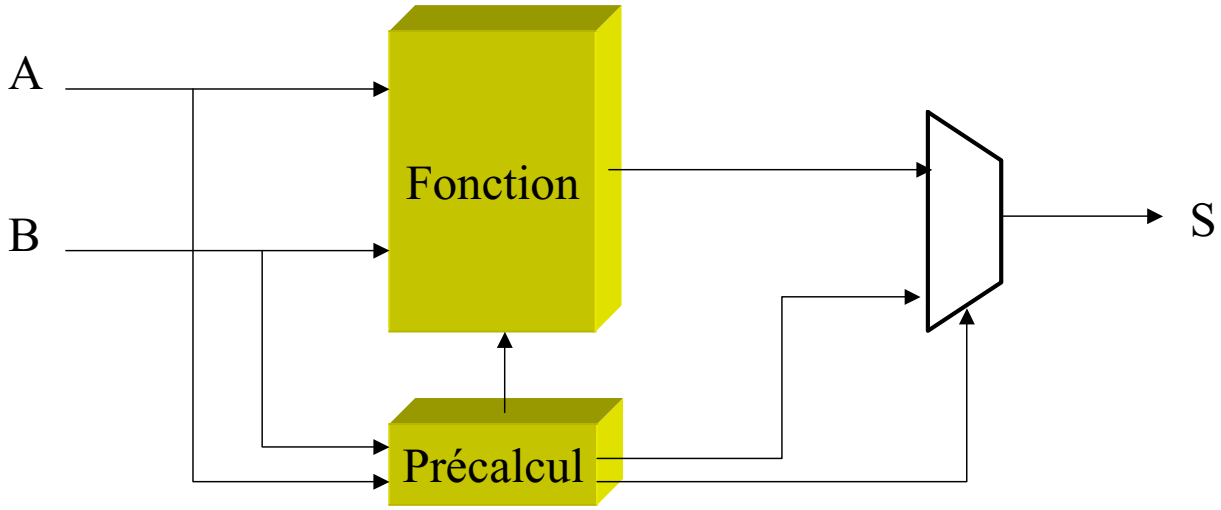
Les entrées du multiplieur sont bloquées si ses sorties sont inutilisées

Les instructions sont bloquées

Optimisations : au niveau logique

IRISA

- Pré calcul :
  - principe général : éviter le passage par une fonction coûteuse lorsque le résultat peut être pré calculé

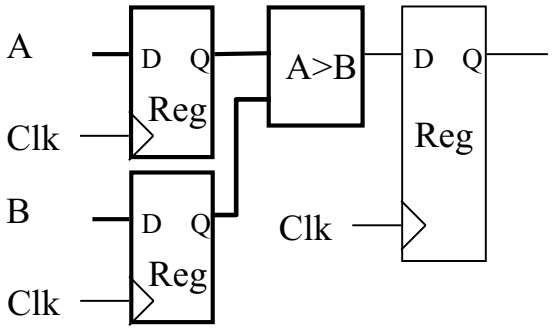


Optimisations : au niveau logique

IRISA

- Pré calcul : exemple
  - comparateur A > B
  - si les 2 MSB sont différents alors le résultat peut être déterminé sans soustraire A et B :
    - Si A[MSB] != B[MSB] && A[MSB] == 0 alors A > B est vraie (1)
    - Si A[MSB] != B[MSB] && A[MSB] == 1 alors A > B est faux (0)

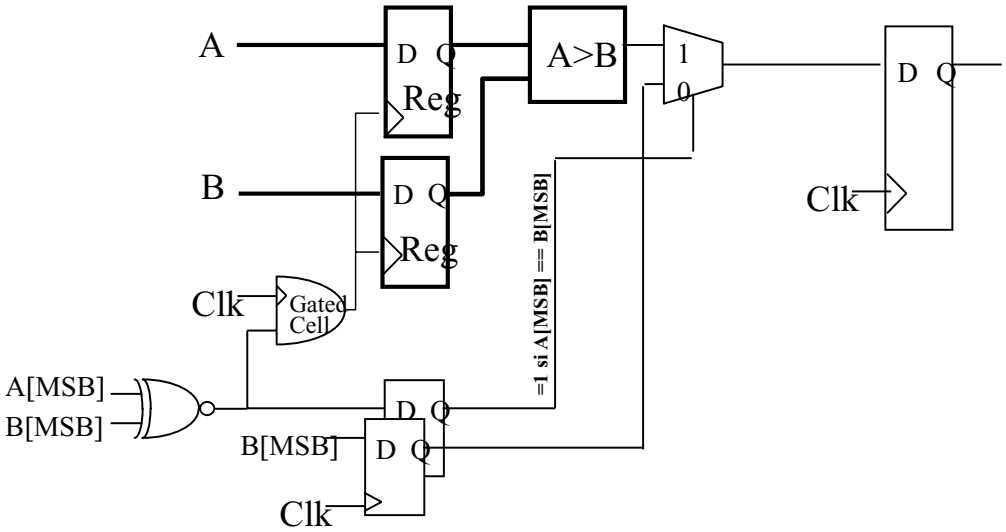
Schéma classique



Optimisations : au niveau logique

IRISA

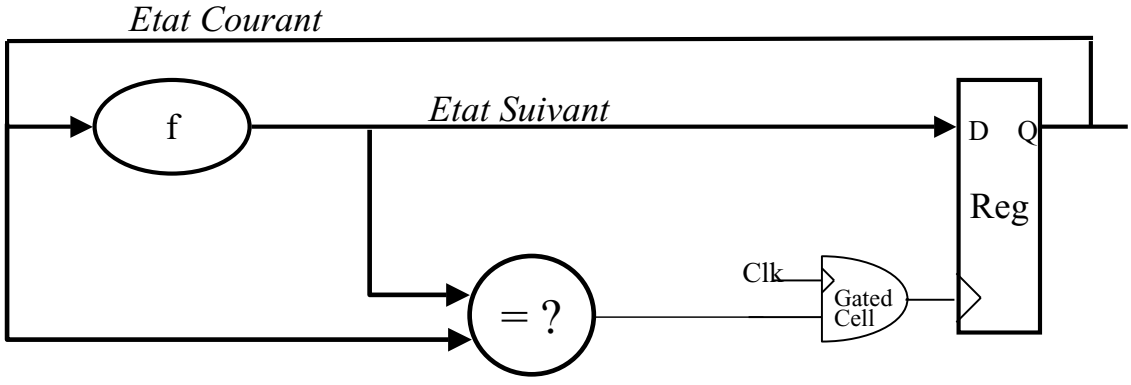
- Pré calcul : exemple
  - si les 2 MSB sont égaux alors il faut faire la soustraction
  - sinon, il suffit de prendre la valeur du MSB de B :
    - si MSB B == 1 alors A > B est vraie
    - sinon A > B est faux



Optimisations : au niveau logique

IRISA

- Post calcul :
  - principe général : bloquer un chargement si l'état suivant est le même que l'état courant



Optimisations : au niveau logique

IRISA

- Post calcul :
  - exemple

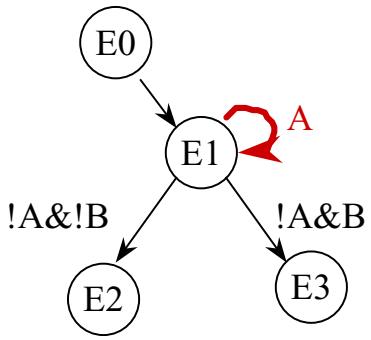


Schéma classique

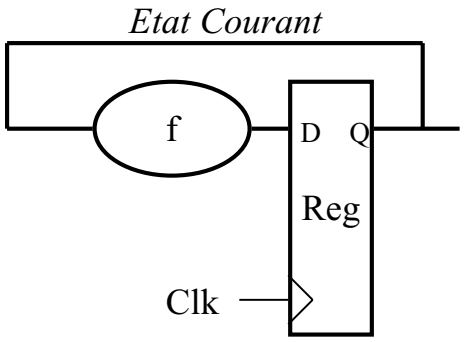
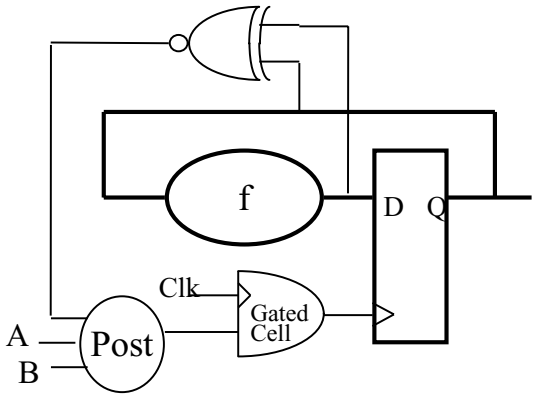


Schéma bloquant le chargement du même état

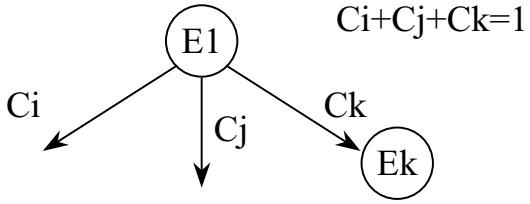


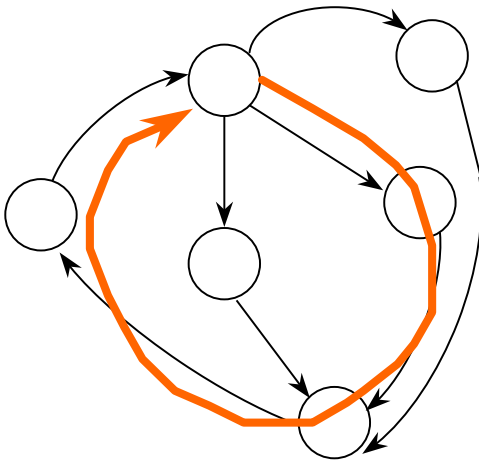
Optimisations : au niveau logique

IRISA

- Les machines d'états :
  - Code pour les états d'une machines

Gray	**Binaire**
- Activité	+ Activité





Coder le graphe avec le minimum de transitions

Si Prob(Ck) importante Alors
les codes de E1 et Ek doivent différer
avec un minimum de positions
(distance de Hamming)

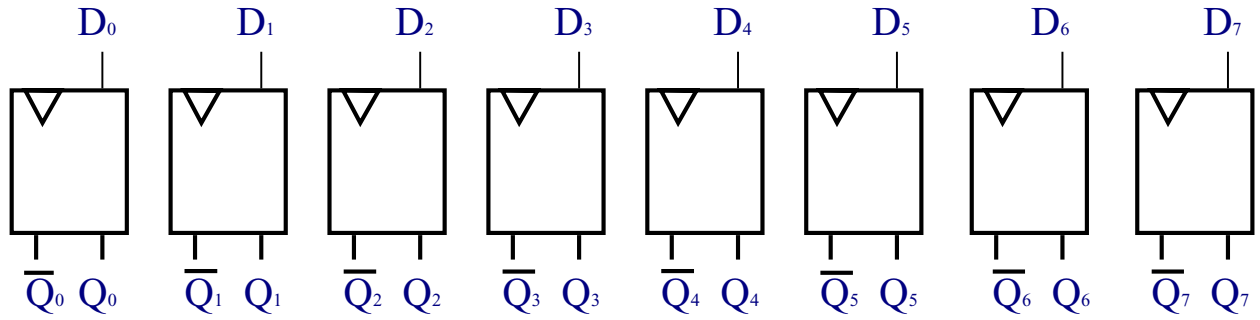
Chemin le plus probable

A coder avec un Min de transitions

Optimisations : Dépendance aux données

IRISA

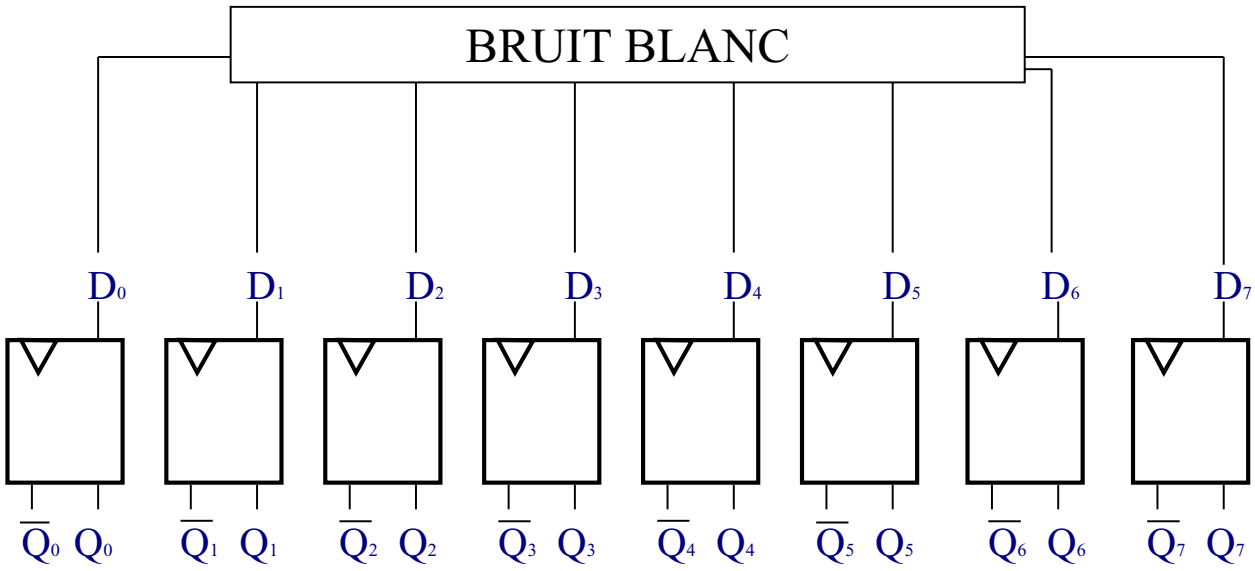
- e.g. Registre 8 bits
- Consommation d'une bascule :
  - 9.5µW/MHz par transition de 0 à 1 à la sortie



Optimisations : Dépendance aux données

IRISA

- Signal aléatoire :
  - Environ 4 bascules commutent pour un signal aléatoire en entrée :
    - deux bascules commutent de 1 à 0 et deux de 0 à 1 :
      - donc 2*9,5 = 19 µW/MHz

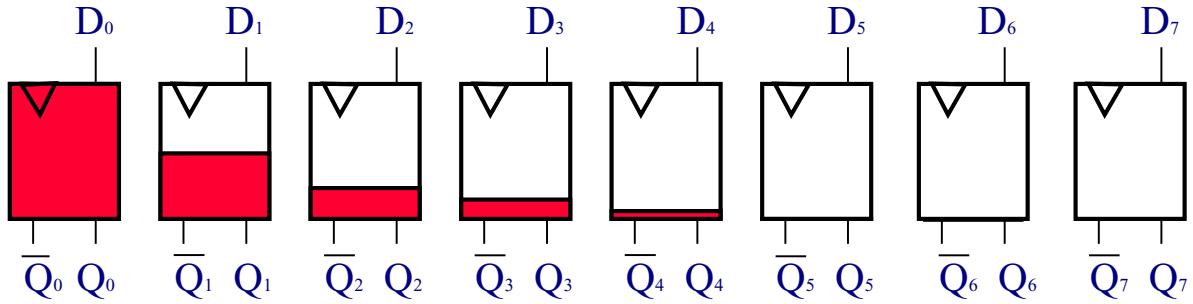


Optimisations : Dépendance aux données

IRISA

- Comptage binaire en complément à 2
  - Proba transition = PT(b0) + PT(b1) + PT(b2) + …….
  - 1 + 1/2 + 1/4 + 1/8 ... = 2
    - conso = 2 * 9,5 / 2 = 9,5 µW/MHz

0 0 0
0 0 1
0 1 0
0 1 1
1 0 0
1 0 1
1 1 0
1 1 1

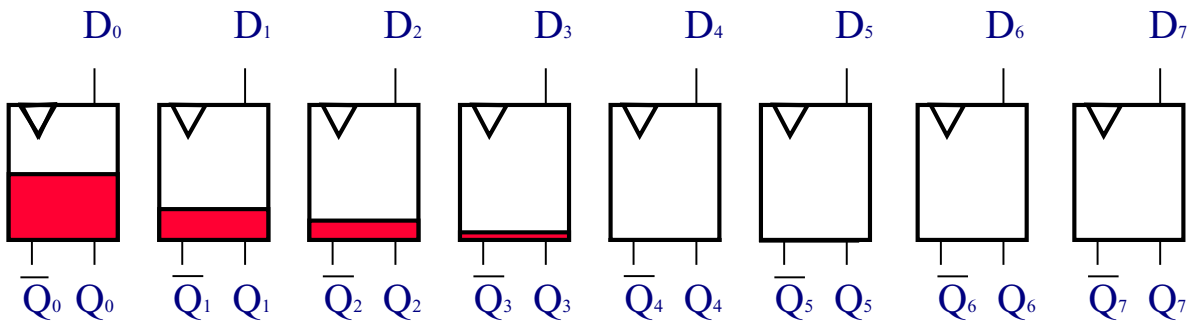


Optimisations : Dépendance aux données

IRISA

- Comptage binaire en codage gray
  - Proba transition = PT(b0) + PT(b1) + PT(b2) + …….
  - 1/2+1/4+1/8...=1
    - conso = 1 * 9,5 / 2 = 4,75 µW/MHz

0 0 0
0 0 1
0 1 1
0 1 0
1 1 0
1 1 1
1 0 1
1 0 0



Optimisations : Dépendance aux données

IRISA

⌘ Dépendance aux données ???

Signal	Bruit blanc	Signal de comptage binaire complémént à 2	Signal de comptage binaire gray
Consommation	19 µW/MHz	9,5 µW/MHz	4,75 µW/MHz

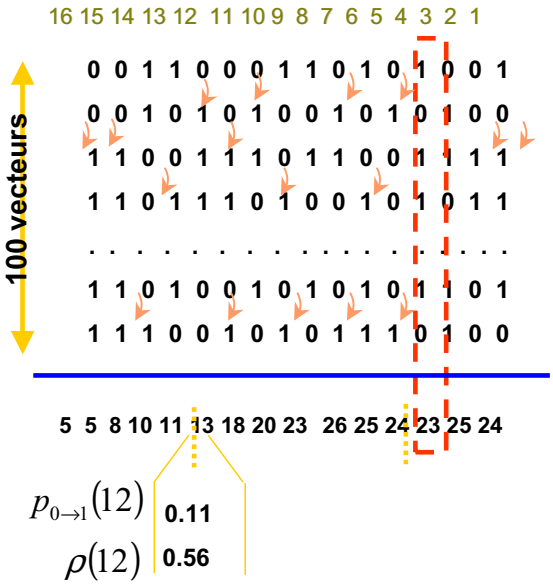
⌘ D'où l'importance d'avoir un modèle de signal d'entrée

⌘ et de pouvoir propager ce modèle dans l'architecture

Optimisations : Dépendance aux données

IRISA

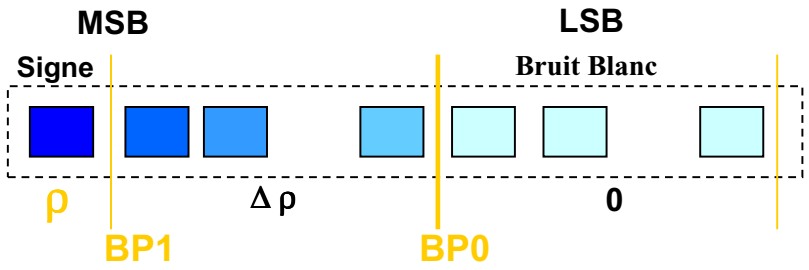
[PREECE81]



$$\rho = 1 - 4 \cdot p_{0\to1}$$

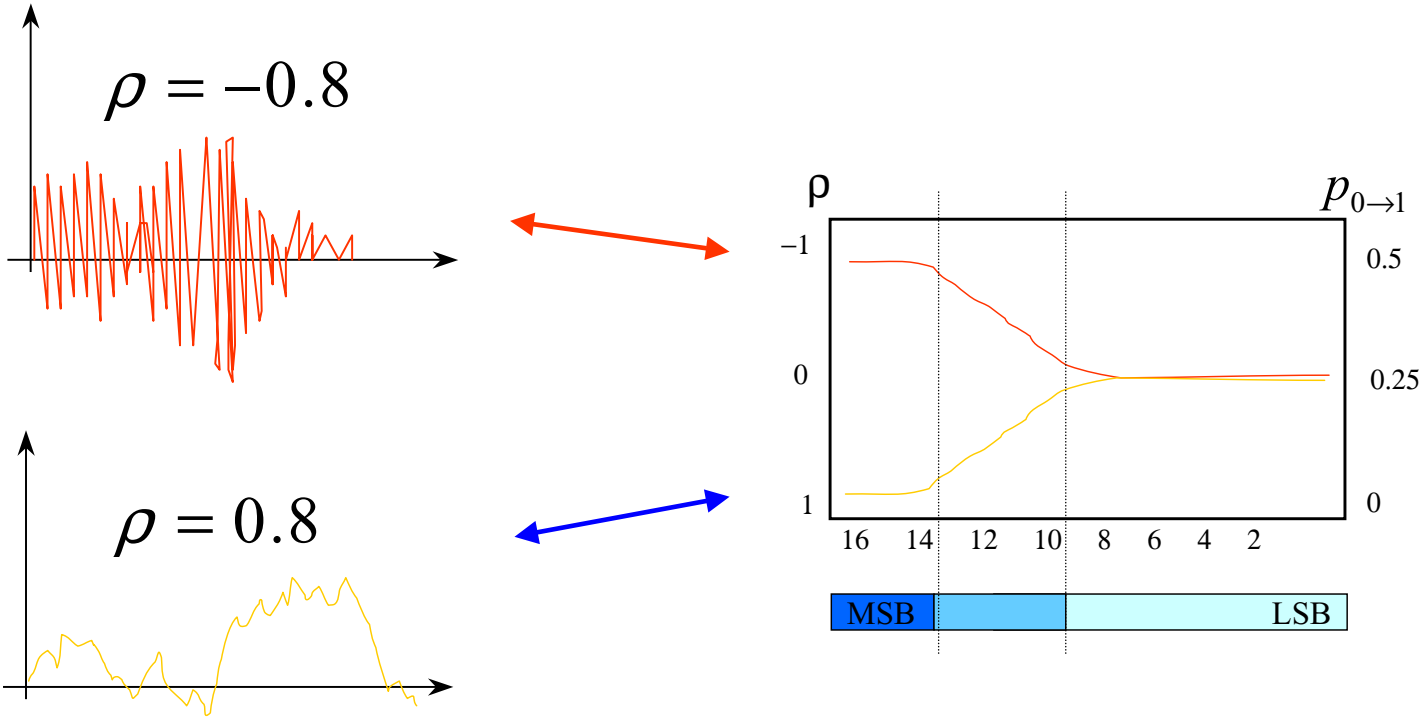
Dual Bit Type

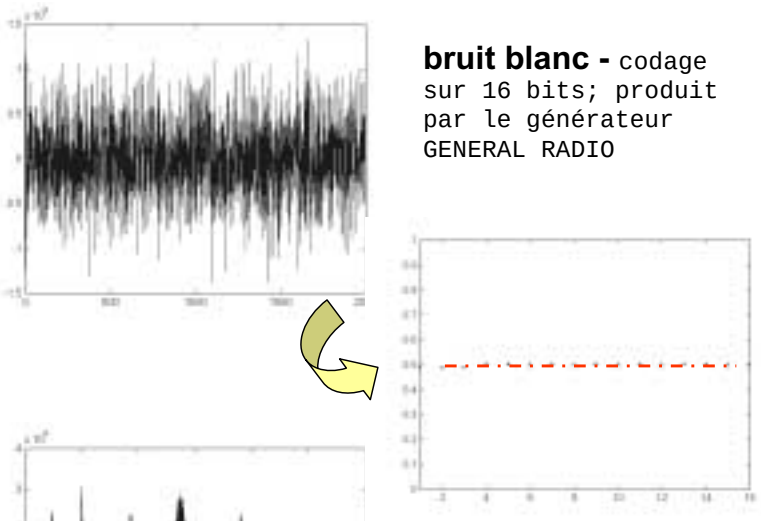
[LANDMAN96]



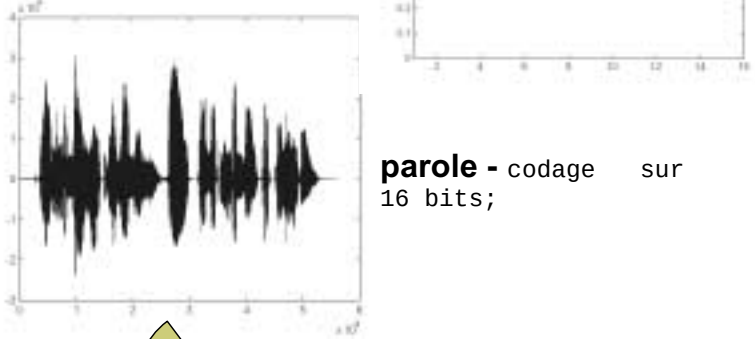
⌘ Signaux audio, vidéo, radio, ...

Optimisations : Dépendance aux données





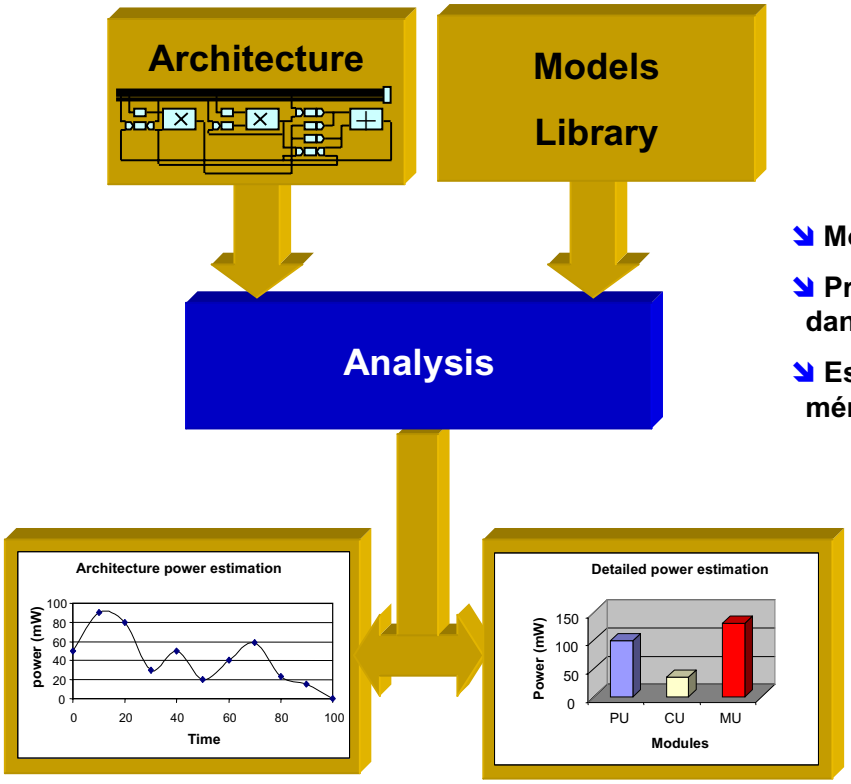
bruit blanc - codage sur 16 bits; produit par le générateur GENERAL RADIO



parole - codage sur 16 bits;

image - codage sur 8 bits;

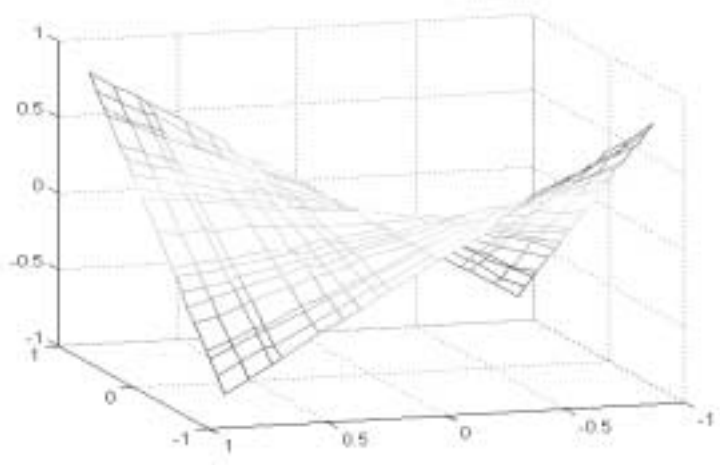
Optimisations : Dépendance aux données



- Modèle probabiliste du signal (DBT)
- Propagation des propriétés du signal dans l'archi
- Estimation conso, : traitement, mémoire, contrôle

Optimisations : Dépendance aux données

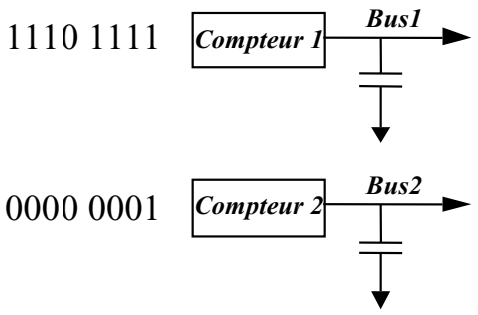
- Propagation du signal dans l'architecture



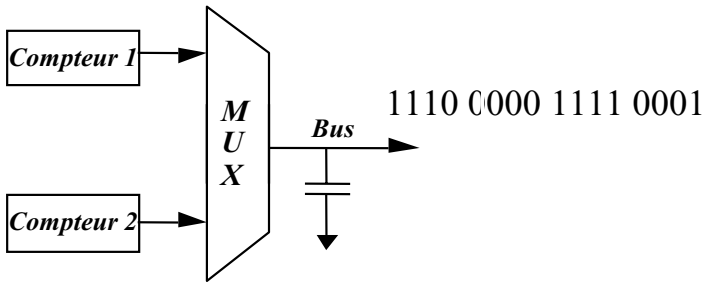
Optimisations : au niveau architectural

IRISA

- Le partage de ressources augmente l'activité
  - Destruction des corrélations entre signaux



a) Implémentation parallèle



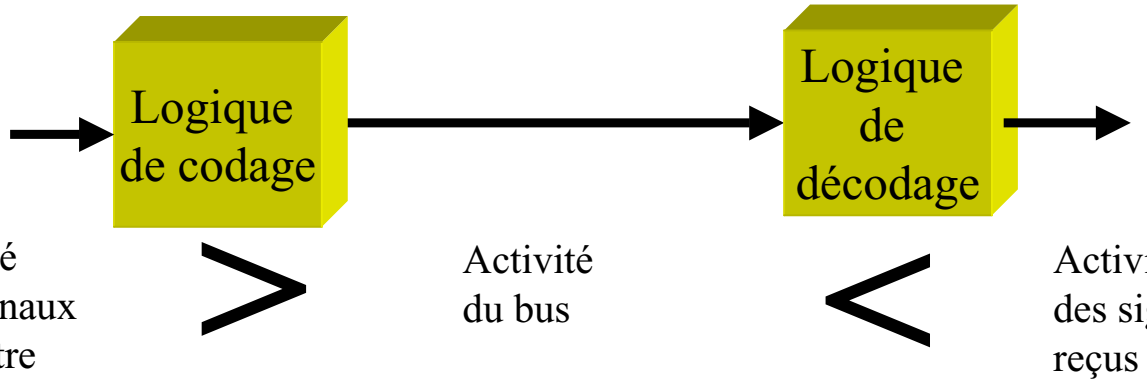
b) Multiplexage temporel

  - Effet de la réutilisation des ressources :
    - non respect des flots de données
    - augmentation de l'activité
    - augmentation de la conso

Optimisations : au niveau architectural

IRISA

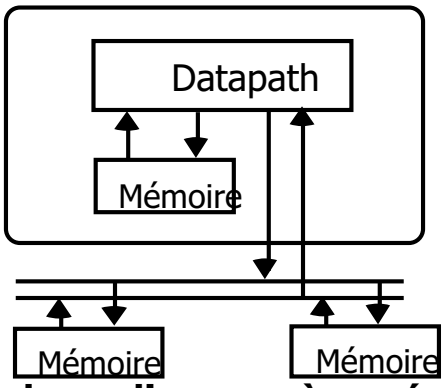
- Encodage de bus pour diminuer l'activité
  - e.g. liaison mémoire ou cache (instruction ou données) vers CPU
  - objectif : diminuer le nombre de transitions entre les valeurs véhiculées :
    - codages :
      - activité binaire > activité gray



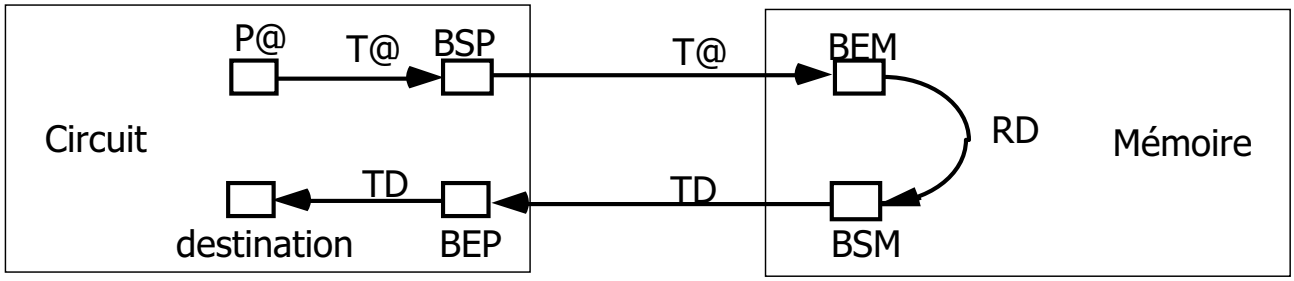
Optimisations : au niveau architectural

IRISA

- Les mémoires :
  - **Modélisation générale**



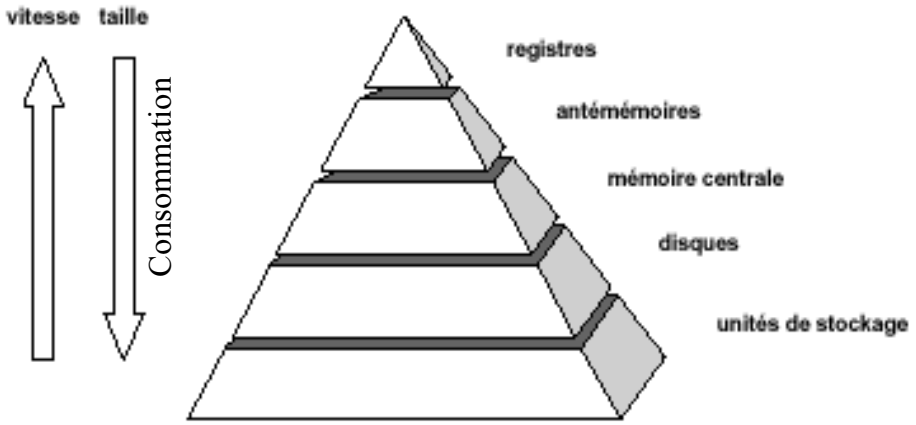
  - **modélisation de la consommation lors d'un accès mémoire**



Optimisations : au niveau architectural

IRISA

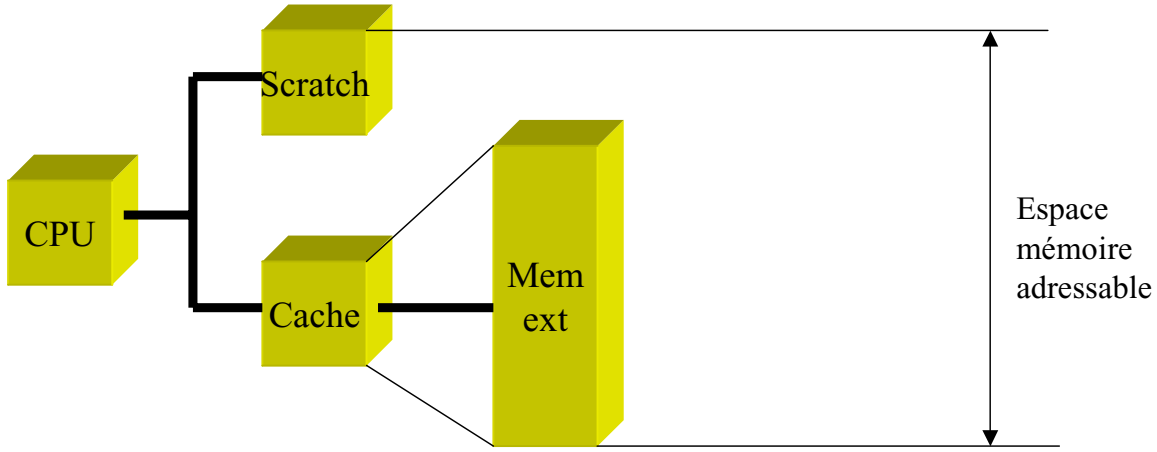
- Organisation mémoire :
  - Constats et motivations :
    - croissance importante du coût engendré par la mémorisation
    - la mémoire est très souvent le goulot d'étranglement du système
    - modèle sans hiérarchisation :
      - mal adapté, beaucoup d'accès à une grosse mémoire
      - nécessite de disposer de mémoires importantes et avec temps d'accès court



Optimisations : au niveau architectural

IRISA

- Organisation mémoire :
  - placement des données les plus souvent utilisées dans des mémoires internes (cache)
  - minimisation de la taille mémoire totale par réutilisation des points mémoires
  - mise en place de buffers ou de scratchpad pour limiter les accès aux mémoires



Optimisations : au niveau architectural

IRISA

- techniques :
  - exploiter les profils des accès aux données
    - localités spatiales et temporelles des données
  - exploiter le déterminisme de l'application :
    - nombre d'accès à chaque données
    - dates des accès
    - sens des accès
    - décision de placement (dans le scratch-pad ou en mémoire principale)

Optimisations : au niveau architectural

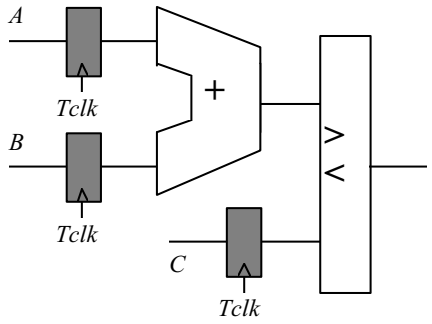
IRISA

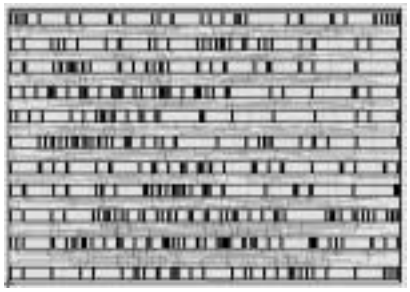
- Compenser la réduction de Vdd par des optimisations architecturales ou logiques
  - Exemple : architecture 16 bits (décodage de Viterbi)

$P_{ref} = C_{ref} \cdot V_{ref}^2 \cdot F_{ref} = 14.7$ mW

C_{ref} déterminé par simulation (Cl = 31pF)

$F_{ref} = 1 / 40$ns $V_{ref} = 5$V





Surface = 788x555 μ^2 = 0.44mm ^{2}

3823 Transistors

56 mm d'interconnexions

Optimisations : au niveau architectural

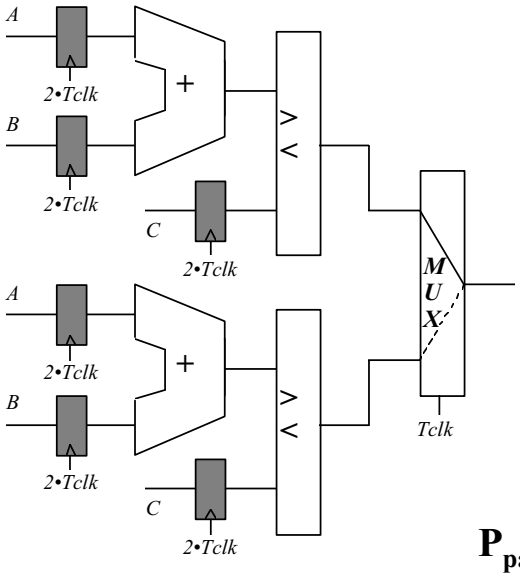
IRISA

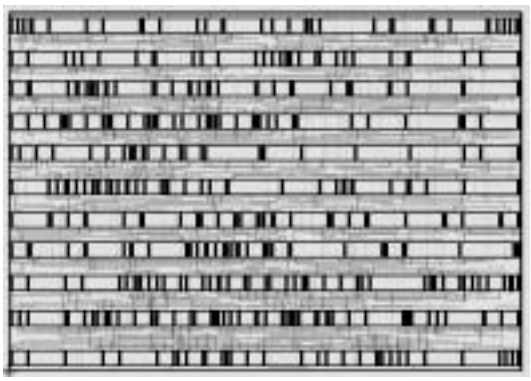
- Parallélisation de l'architecture

Surface = 1130x768μ^2 = 0.87 mm ^{2}

7638 Transistors

111 mm d'interconnexions





$P_{parallèle} = (2.15 \cdot C_{ref}) \cdot (0.58 \cdot V_{ref})^2 \cdot (0.5 \cdot F_{ref})$

$\# 0.36 \cdot P_{ref} = 5.3$ mW

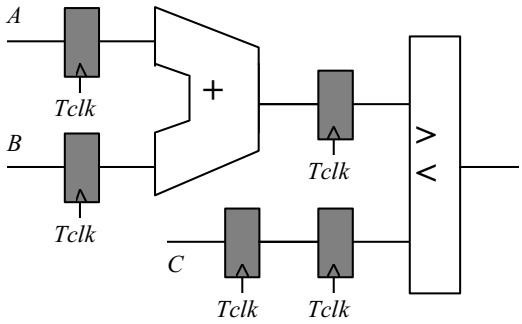
$C_{parallèle} = 2.15 \cdot C_{ref}$

$F_{parallèle} = 1 / 80$ns $V_{dd\ parallèle} = 2.9$V

Optimisations : au niveau architectural

IRISA

Pipelinage de l'architecture



$$P_{pipeline} = (1.15 \bullet C_{ref}) \bullet (0.58 \bullet V_{ref})^2 \bullet F_{ref}$$
$$\# 0.39 \bullet P_{ref} = \mathbf{5.7\ mW}$$

$$C_{pipeline} = 1.15 \bullet C_{ref}$$
$$F_{pipeline} = F_{ref} \qquad V_{dd\ pipeline} = 2.9V$$

Optimisations : au niveau architectural

IRISA

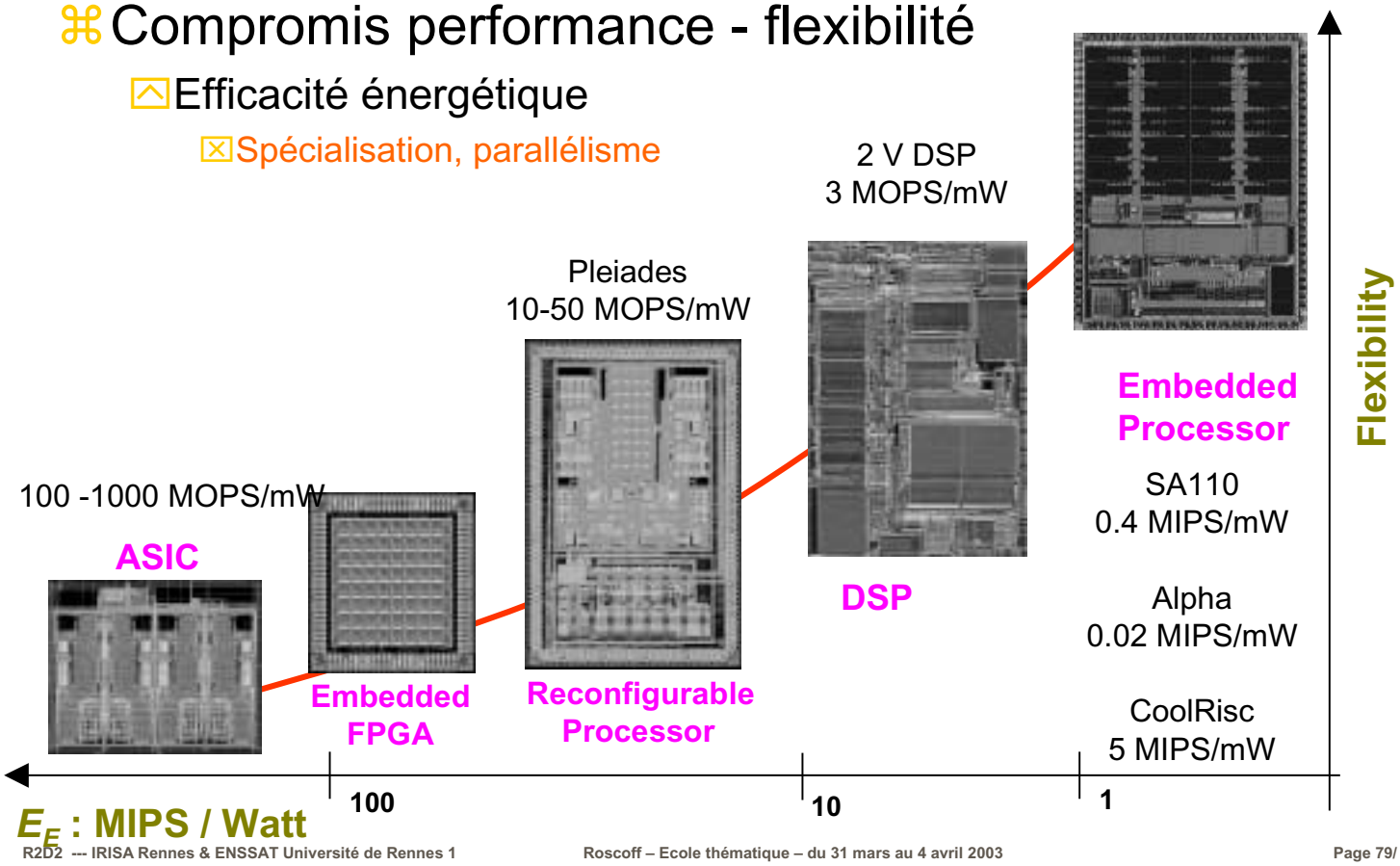
Parallélisation ou pipelinage ?

Type d'Architecture	Vdd	Surface	Puissance
Référence	5V	1	1
Parallèle	2.9V	1.98	0.36
Pipeline	2.9V	1.1	0.39
Parallèle/Pipeline	2.0V	2.2	0.2

Alternatives architecturales

IRISA

- Compromis performance - flexibilité
  - Efficacité énergétique
    - Spécialisation, parallélisme



Alternatives architecturales

IRISA

- Il ne faut plus penser uniquement aux performances
- Efficacité énergétique
  - Nb Opérations Utiles / Energie Utilisée = NbOp / J
    = MOPS/mW
    - MOPS = Millions d'Opérations Par Seconde
  - MOPS/mW = $(Fclk.Nop)/(Nop.Aop.Csw.Vdd^2.Fclk)$
    = $1/(Aop.Csw.Vdd^2)$
    - Nop : Nb d'opérateurs
    - Aop : Surface moyenne d'un opérateur = Achip/Nop
    - Csw : Capacité efficace par mm^2
  - dépend de la spécialisation du chip !!!
- Efficacité en surface
  - $MOPS/mm^2$

Dart : une archi reconf

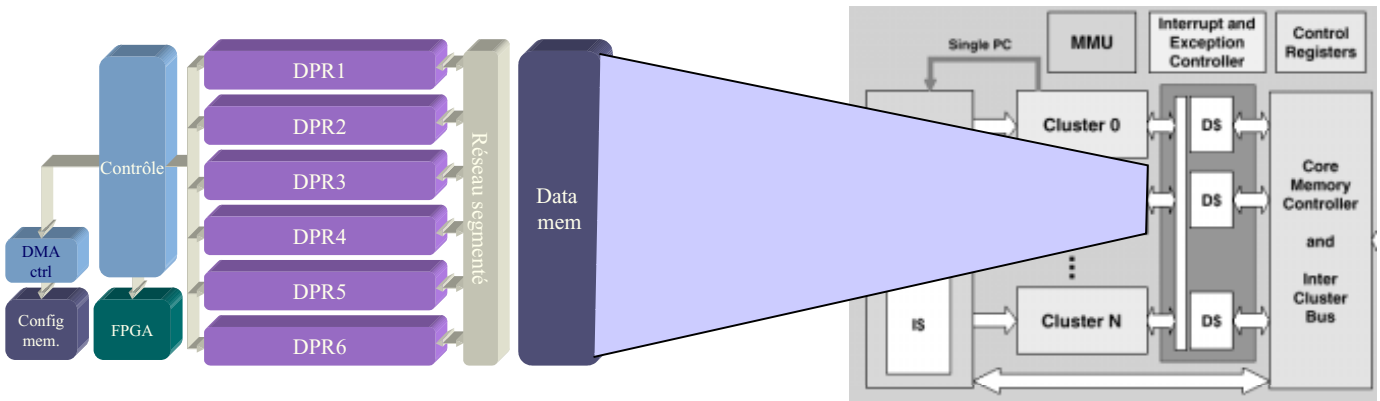
IRISA

- Dart : architecture reconfigurable
  - objectifs :
    - forte puissance de calcul
    - faible consommation
  - principes :
    - recherche d'un bon compromis performance / flexibilité :
      - reconfiguration matérielle : exécution des cœurs de boucles
      - reconfiguration logicielle : exécution d'instructions classiques
    - remplacement du contrôle d'instructions d'exécution en instructions de reconfiguration
  - collaboration ENSSAT/UBO-AS/STMicroelectronics

Dart : une archi reconf

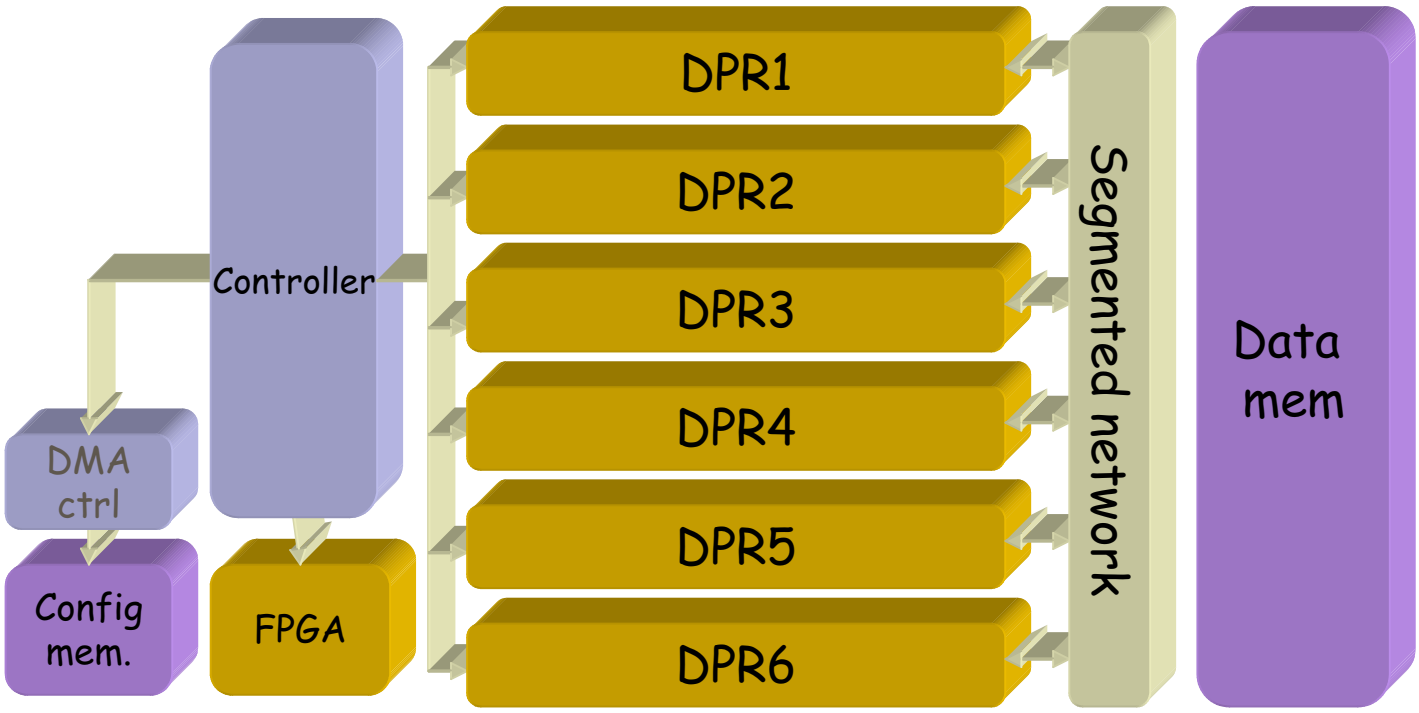
IRISA

- Architecture :
  - Reconfiguration multigrain dynamique
    - Fonctionnel (DPR), logique (FPGA)
  - Faible consommation
  - Distribution hiérarchique des ressources
    - calcul, stockage, interconnexions, contrôle



Dart : une archi reconf

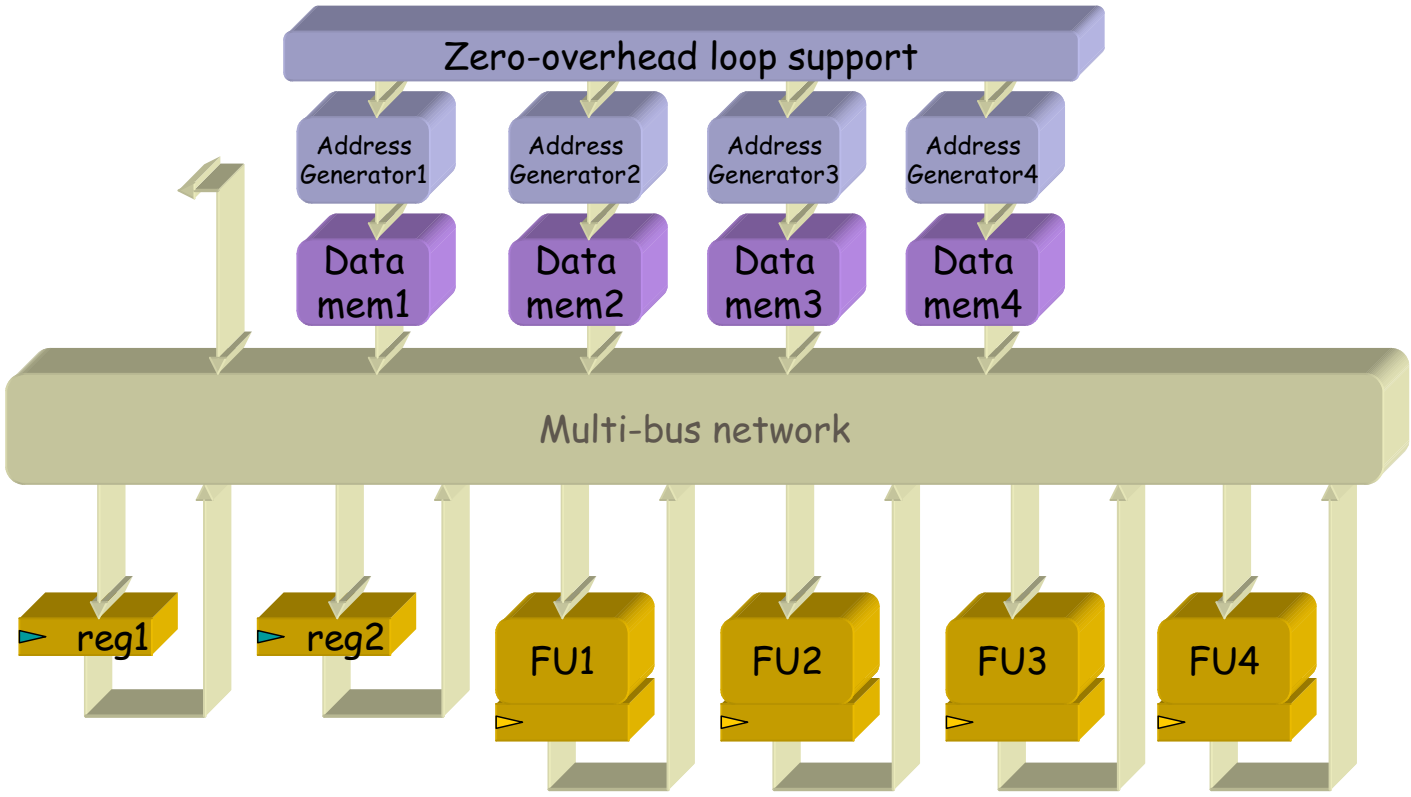
IRISA



Dart : une archi reconf

IRISA

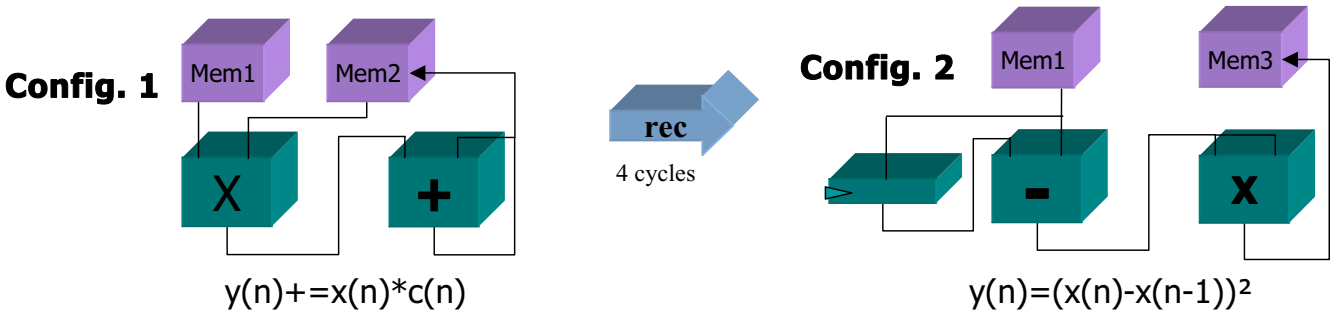
- Data Path Reconfigurable



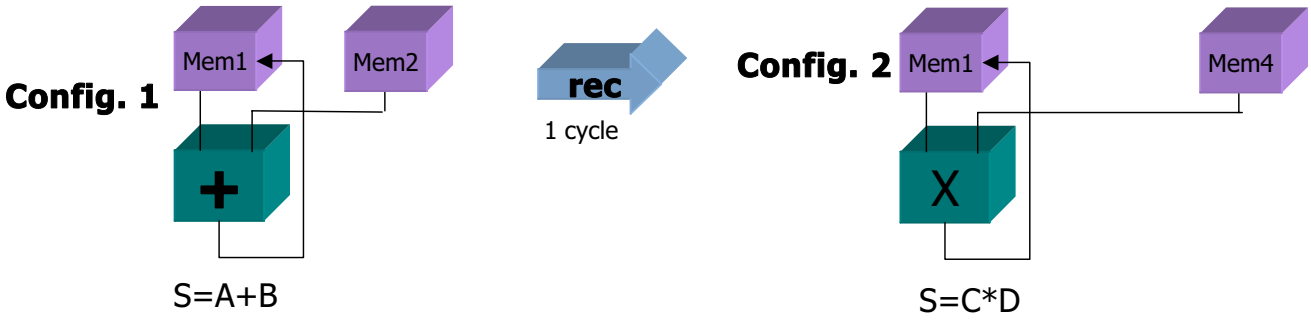
Dart : une archi reconf

IRISA

⌘ Reconfiguration HW pour optimiser le chemin de données :



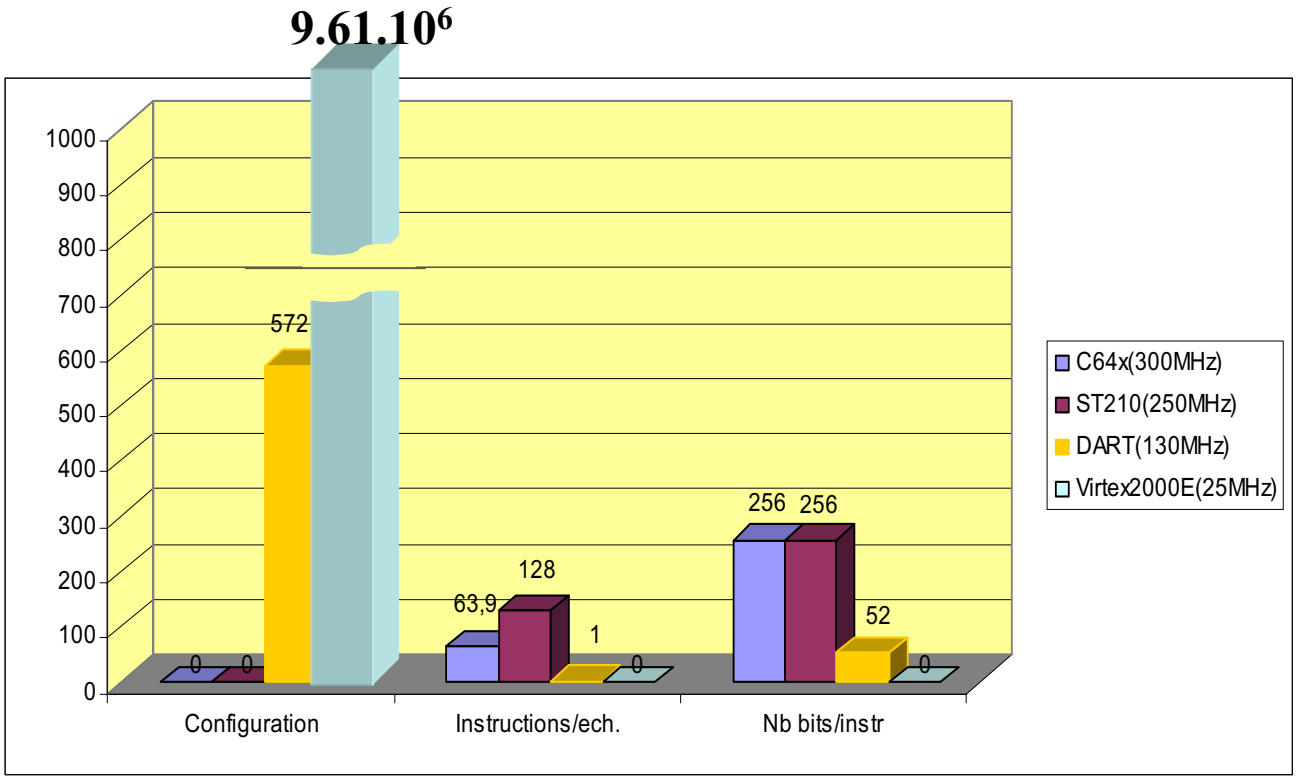
⌘ Reconfiguration SW pour reconfigurer le DPR à chaque cycle :



Dart : une archi reconf

IRISA

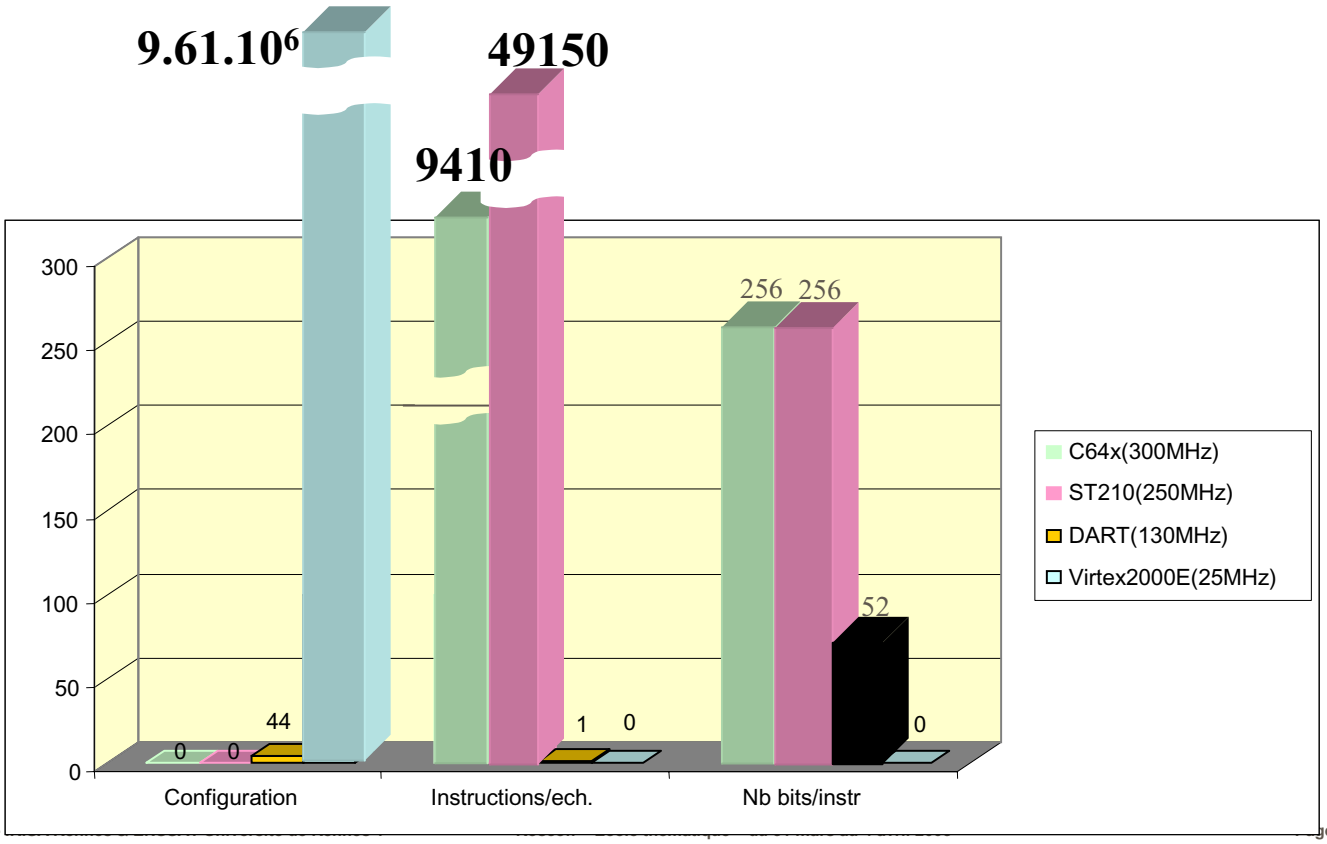
⌘ Résultats : volume des configurations pour le FIR



Dart : une archi reconf

IRISA

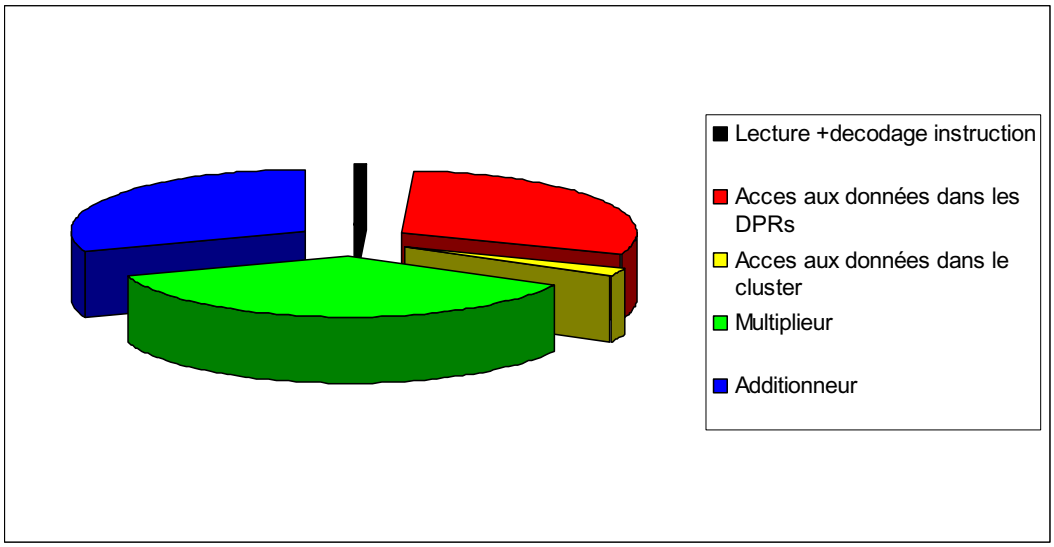
- Résultats : volume des configurations pour le Rake Receiver



Dart : une archi reconf

IRISA

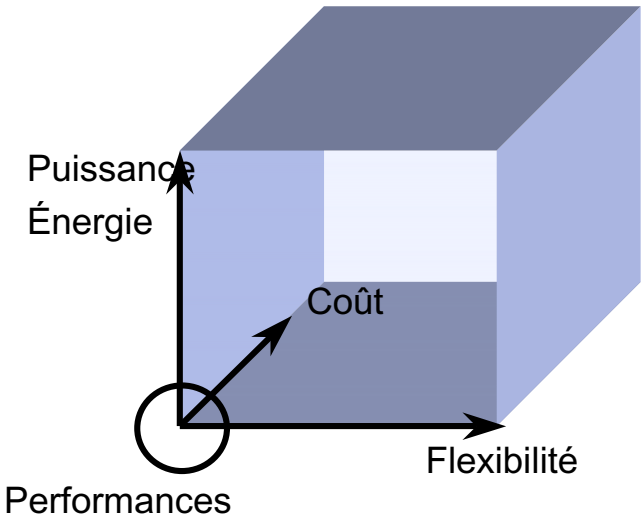
- Répartition de la consommation dans DART lors de l'exécution d'un filtrage (1966GMACs)



149mW

Conclusions

- De nouvelles métriques pour la conception
  - Puissance ou énergie devient la contrainte dominante
- La consommation doit être estimée et optimisée à tous les niveaux, en particulier au niveau système
- La diminution de la tension d'alimentation doit se faire en conservant les performances du circuit
- La diminution de l'activité de commutation des signaux internes ou externes



Organisation **En collaboration avec**

I.S.E.P ENSSAT LANNION

FTFC'2003

4ème journées d'études
Faible Tension Faible Consommation

15-16 mai 2003, Paris, France

Cercle national des armées
8, place Saint - Augustin, 75008 Paris

IEEE

Circuits and Systems Society