Workshop “Data Bahia”
UFBA, Fiocruz, Sorbonne Universités
Salvador de Bahia, Brazil
April 25-26, 2016

UFBA FIOCRUZ SORBONNE UNIVERSITÉS

Supporting big data in Health and Bioinformatics through hybrid parallel architectures and distributed execution engines

Marcos Barreto
Distributed Systems Laboratory (LaSiD)
Computer Science Department (DCC)
Federal University of Bahia (UFBA)

marcosb@ufba.br
http://www.dcc.ufba.br/~marcoseb

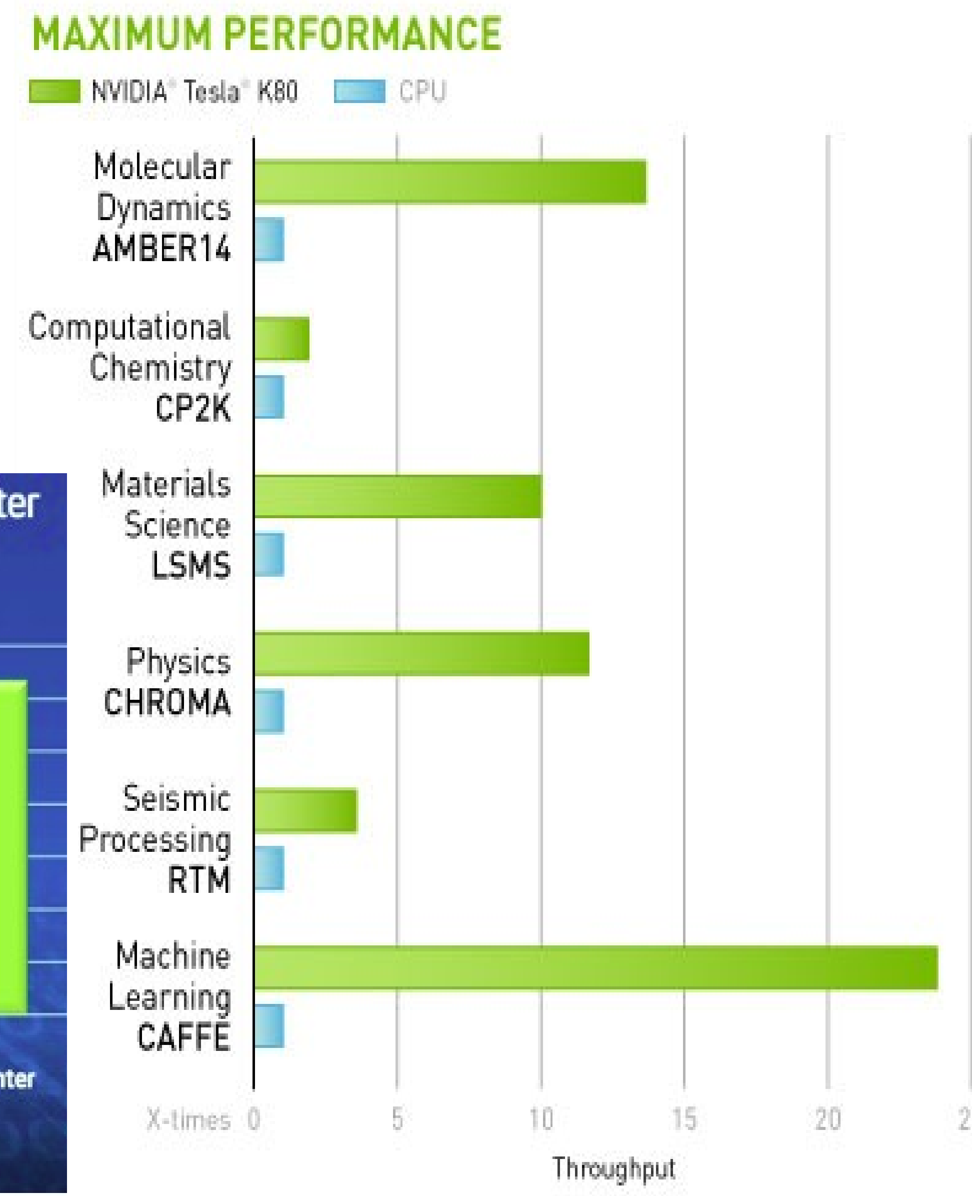
Outline

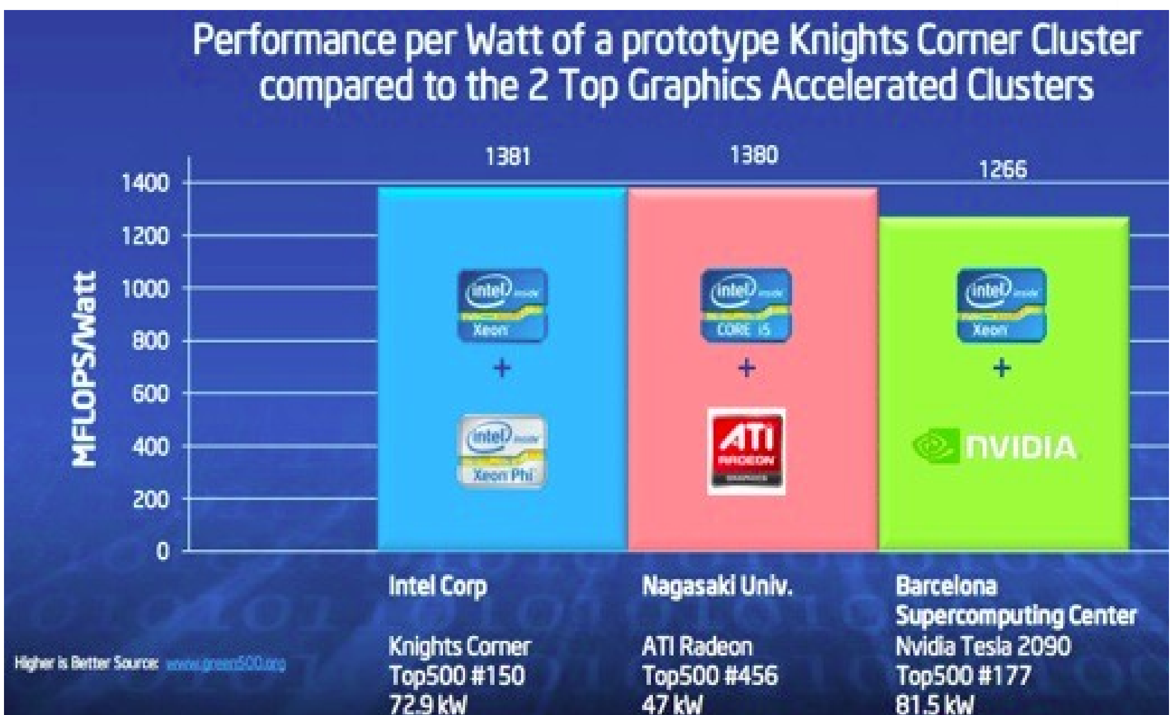
- Part I:
  - Bioinformatics workflows on hybrid parallel architectures.
- Part II
  - Social and healthcare data integration supported by distributed execution engines.

Hybrid parallel architectures

- Multicore processors + multi-GPUs + multi-MIC systems.

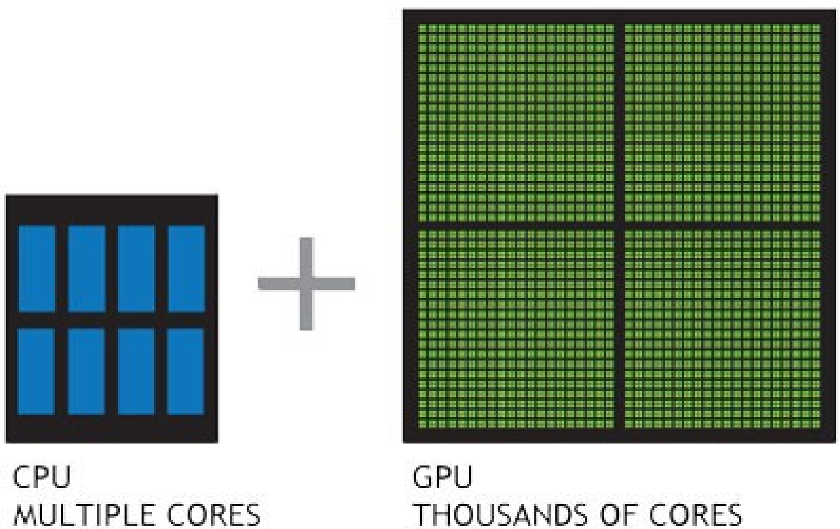
- What technology is dominant?



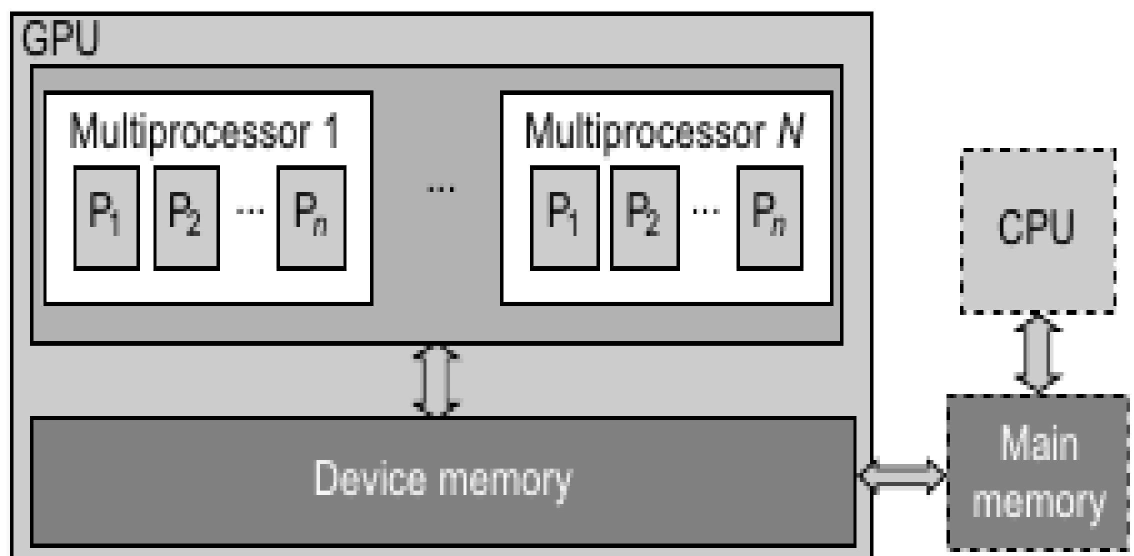


Hybrid parallel architectures

- Why they are hard to program?
  - Parallel programming => performance-oriented programming.
    - Users must explicitly deal with scalability, load (tasks + data) balancing, synchronization, and communication issues.



2 to 32 (AMD Opteron Zen)



Our approach

- <u>Performance model</u>, based on *algorithm parameters* (AP) and *system parameters* (SP).

Input size

$$Execution\,time(t) = f(s, AP, SP)$$

- # of processors, # of processes
- # and size of data blocks

- Cost of one arithmetic operation
- Communication (start-up + word transfer) times

- <u>Auto-tuning methodology</u>: estimate parameters during setup.

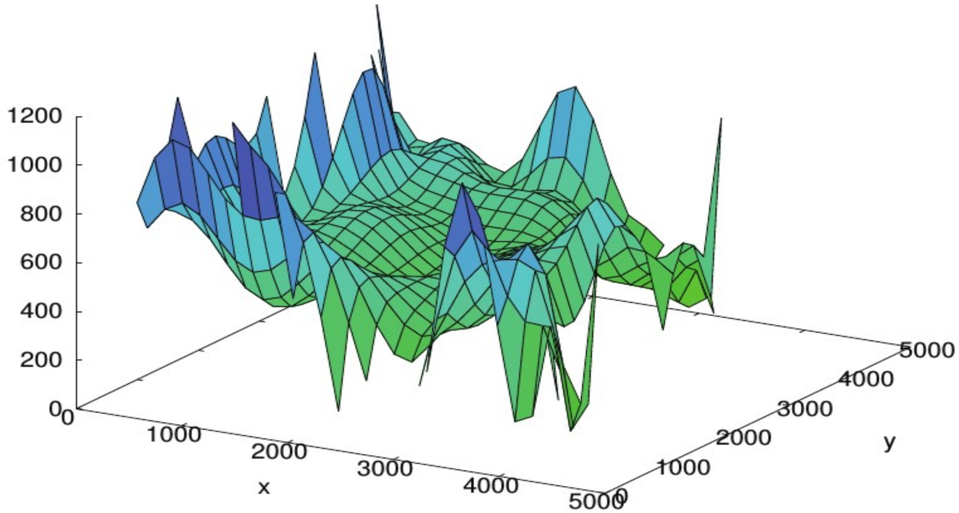
Symbol	Description
k	Computation parameter for each execution system
N	Order of the polynomial. In the experiments it ranges from 2 to 40
n	The length of the sum. In the experiments it ranges from 1.3 to 25.4 million terms
c	Number of CPU cores
w	GPU workload
g_w	Number of GPUs
t_c	Cost of initialization of a thread in CPU
t_{g_w}	Cost of initialization of a kernel in GPU
$S_{\frac{g_w}{c}}$	Relative speedup of a GPU with respect to a core in the CPU

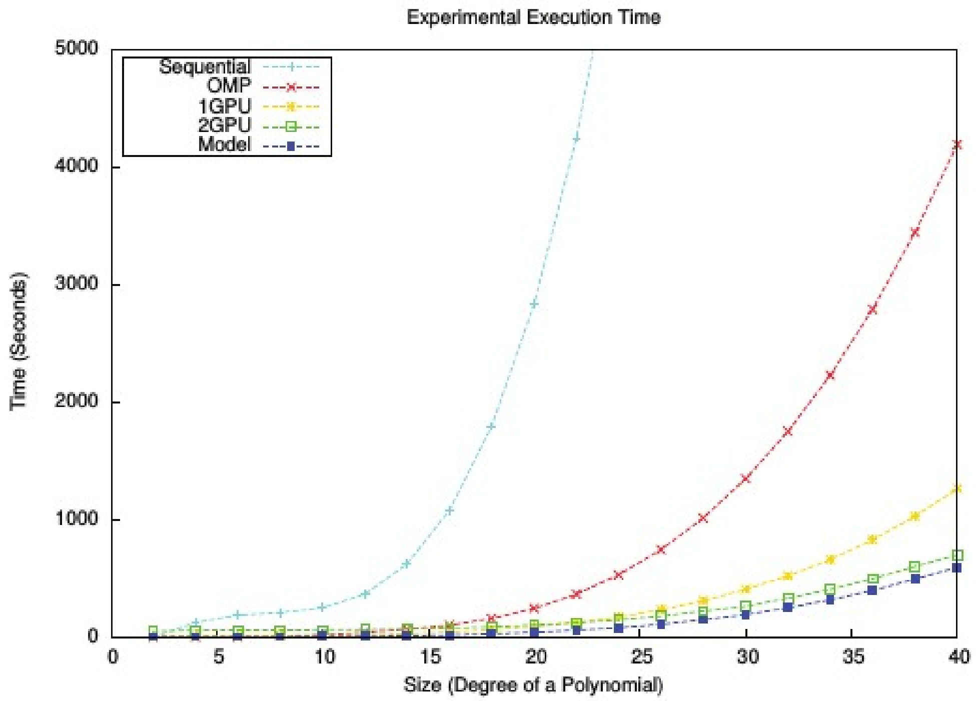
$$\boxed{t(N, n, c, w) = \frac{k \cdot N^2 \cdot n}{c + g_w \cdot S_{\frac{g_w}{c}}} + t_c \cdot c + t_{g_w(w)} \cdot g_w.}$$

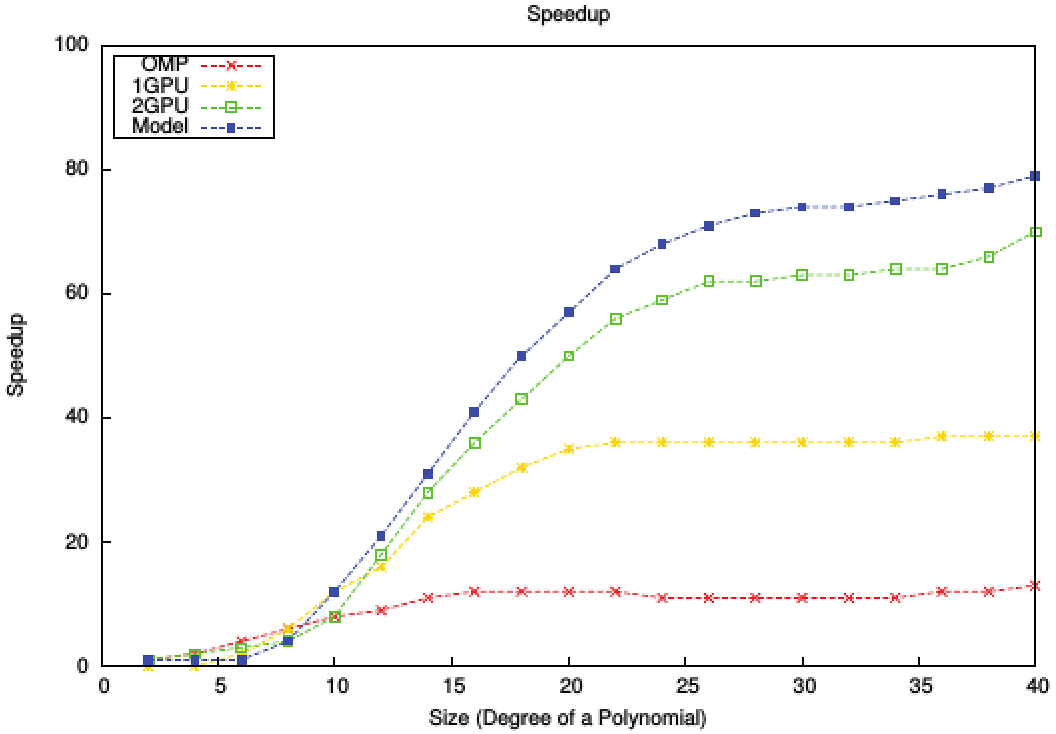
Heterogeneous Computational Model for Landform Attributes Representation on Multicore and Multi-GPU Systems

BORATTO, M.; BARRETO, M.; ALONSO, P.; RAMIRO, C. (ICCS 2012)

Degree Polynomial	Sequential	OMP	1GPU	2GPU	Model
8	84.49	12.32	12.44	14.19	13.61
12	386.17	41.85	21.36	19.39	18.04
16	1, 166.88	114.55	43.31	31.48	25.53
20	2, 842.52	268.32	90.57	57.03	49.29
24	5, 916.06	544.93	172.71	101.08	88.86
28	11, 064.96	1, 011.42	310.53	176.88	156.72
32	24, 397.66	1, 777.25	521.63	285.07	256.62
36	30, 926.82	2, 700.00	828.50	450.67	404.25
40	46, 812.70	4, 252.69	1, 261.09	666.77	600.90





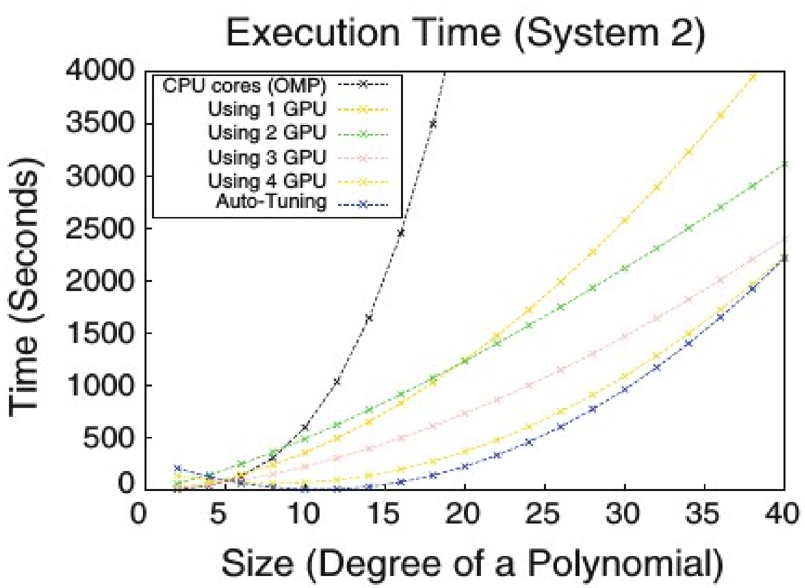


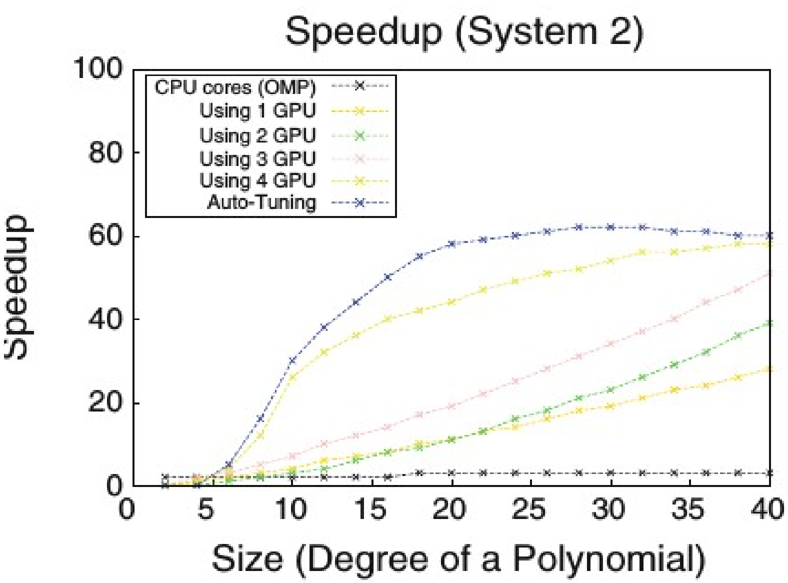
Automatic routine tuning to represent landform attributes on multicore and multi-GPU systems

BORATTO, M.; BARRETO, M.; ALONSO, P.; GIMENÉZ, D. (Journal of Supercomputing, 2014)

System 2:
- 4 NVIDIA Tesla C2070 (240 cores/GPU), 2 Intel Xeon quadcore processors (2.4 GHz, 48 GB DDR3)
- Number of CPU cores (*c*) = 16
- Workload (*w*) = (GPU, GPU, GPU, GPU, CPU) = (22%, 22%, 22%, 22%, 12%)

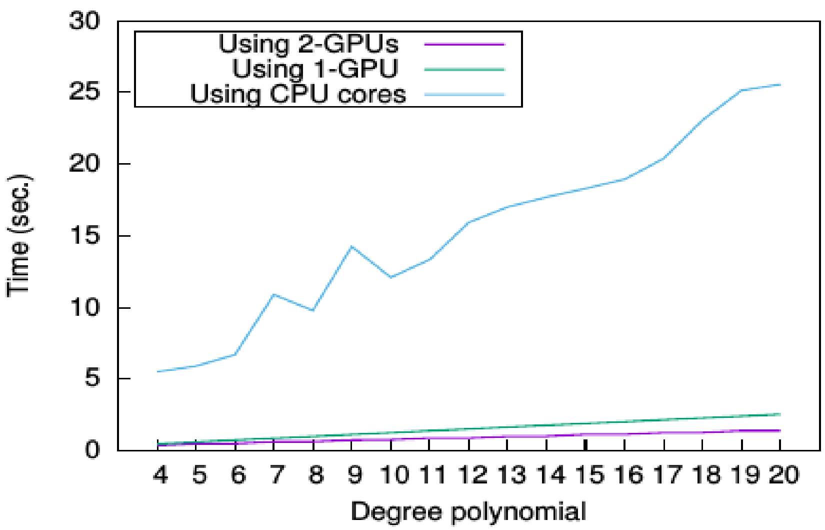
System 2	$w = 10$		$w = 15$		$w = 20$		$w = 22$	
N	c	$t(N, n, c, w)$	c	$t(N, n, c, w)$	c	$t(N, n, c, w)$	c	$t(N, n, c, w)$
10	8	59.25	10	48.50	10	50.93	10	47.53
20	8	479.88	14	1,227.78	14	940.60	14	860.79
30	10	1,623.82	14	2,113.72	14	1, 200.96	16	1,090.76
40	10	3,858.78	16	3,107.71	16	1,390.07	16	1,260.05

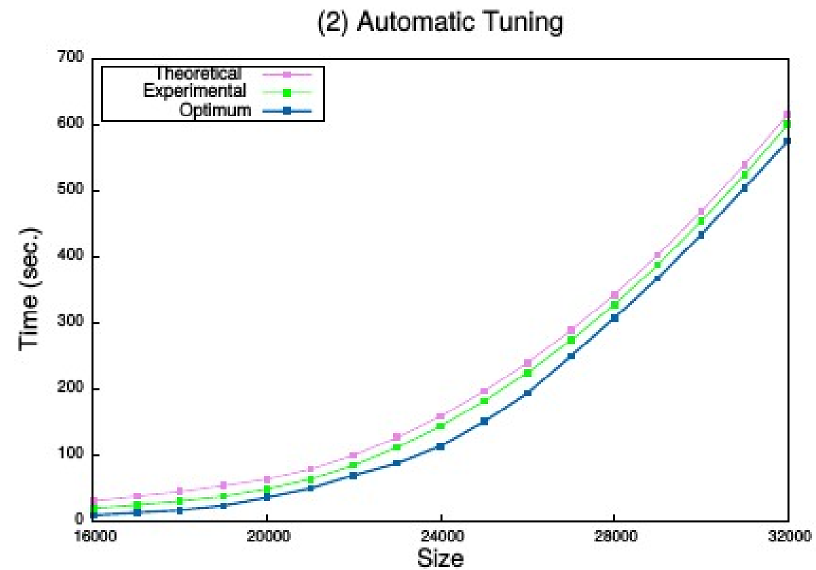




Current efforts

- Use of <u>functional performance models</u> (FPM) to data partitioning in hybrid parallel architectures.
  - Asynchronous task assignment model applied to matrix polynomials and triangular linear systems.





- Complete abstraction of the underlying architecture.

```
run(myApp, inputData, NumberofCores)
```

Current efforts - Bioinformatics

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

UNEB UNIVERSIDADE DO ESTADO DA BAHIA

- Adaptation of bioinformatics tools to hybrid architectures through performance and auto-tuning models

Biology

Phylogenetic analysis

BLAST, ClustalW, PhyML, ProtTest, Tree-Puzzle

Pharmacy

Virtual screening

Autodock, Gromacs

FIOCRUZ

Leishmania, EpiGen-Brasil

FPM + auto-tuning

MPI OpenMP TBB NVIDIA CUDA MESOS Spark

Xeon Phi, CORE i7, NVIDIA

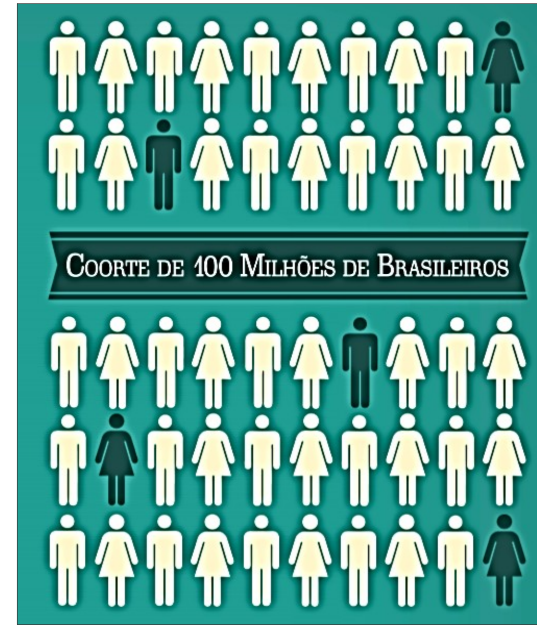
BOINC, hadoop

Support for bioinformatics applications through volunteer and scalable computing frameworks

GUTIERREZ, F.; AZEVEDO, D.; BARRETO, M.; ZUCOLOTO, R. (CLUSTER 2014)

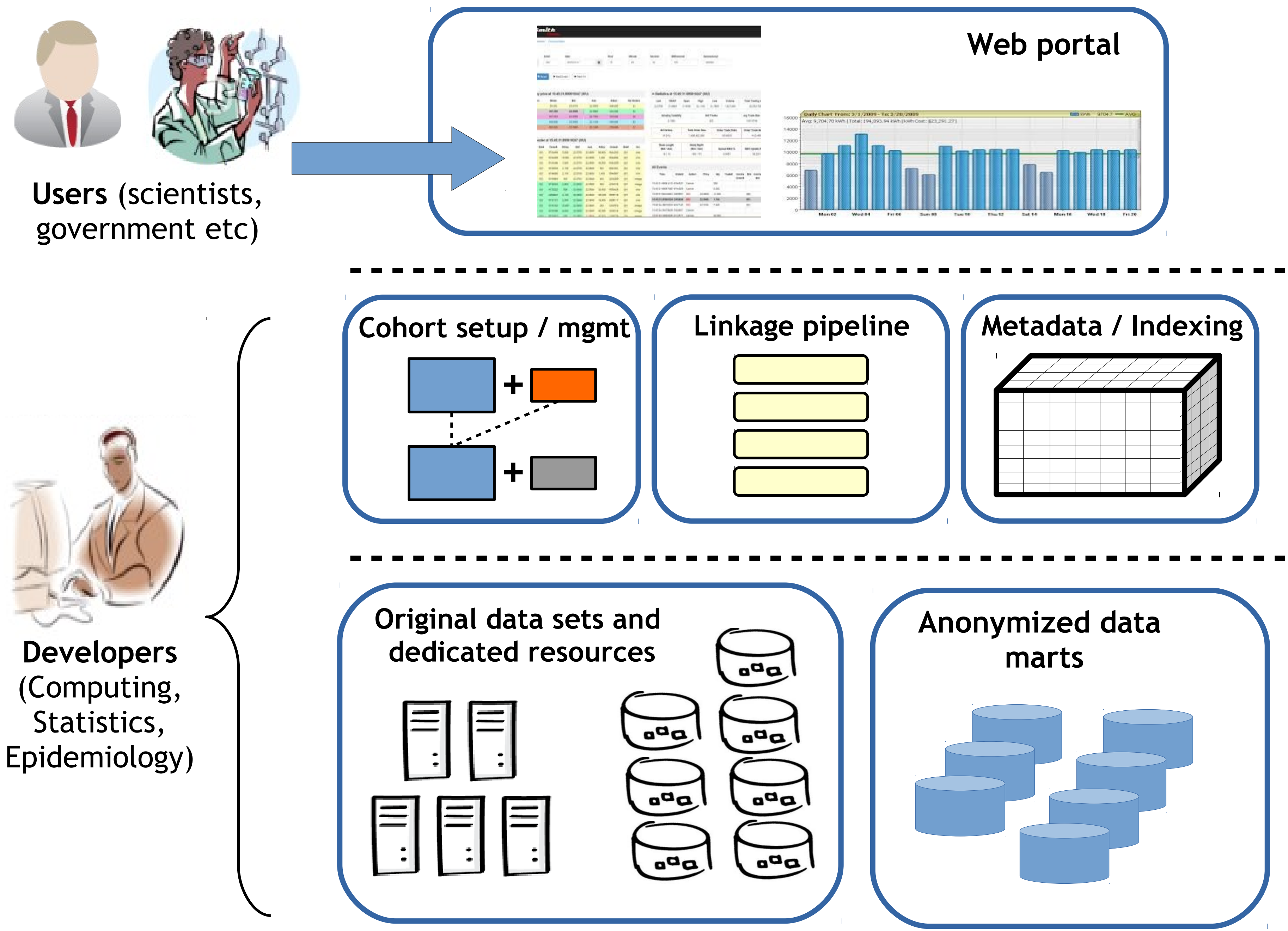
Part II – Social and healthcare data integration

- The **'100 million cohort project'** challenge
  - Introductory conference: Mauricio Barreto - *“Evaluating the impact of social protection policies on health: the 100 million Brazilian cohort”*



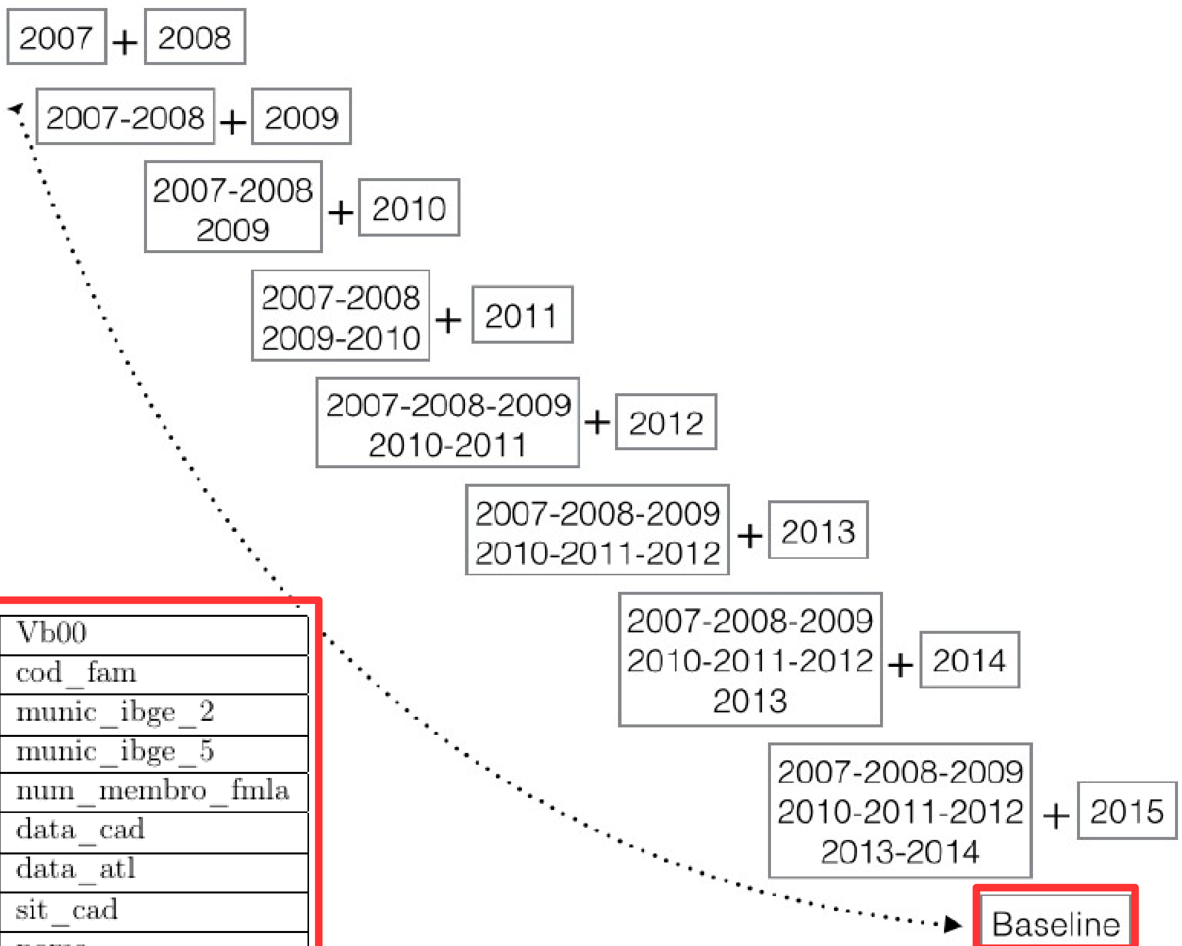
Ministério da Saúde
FIOCRUZ
Fundação Oswaldo Cruz
UFBA
UnB
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE
SENAI CIMATEC
SISTEMA FIEB
Federação das Indústrias do Estado da Bahia
SECRETARIA DE CIÊNCIA, TECNOLOGIA E INOVAÇÃO
BAHIA GOVERNO DO ESTADO
RIO+ WORLD CENTRE FOR SUSTAINABLE DEVELOPMENT
UNDP
BILL & MELINDA GATES foundation
SUS
Ministério da Saúde
Ministério do Desenvolvimento Social e Combate à Fome
Ministério da Ciência e Tecnologia
GOVERNO FEDERAL BRASIL PÁTRIA EDUCADORA

Proposed platform



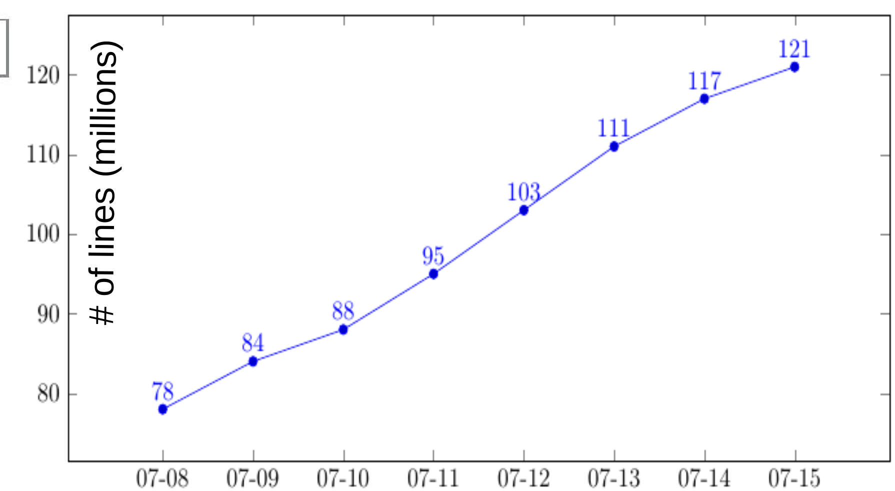
Cohort setup and management

- Longitudinal merge of CadastroÚnico (CadU) based on NIS (social ID) attribute

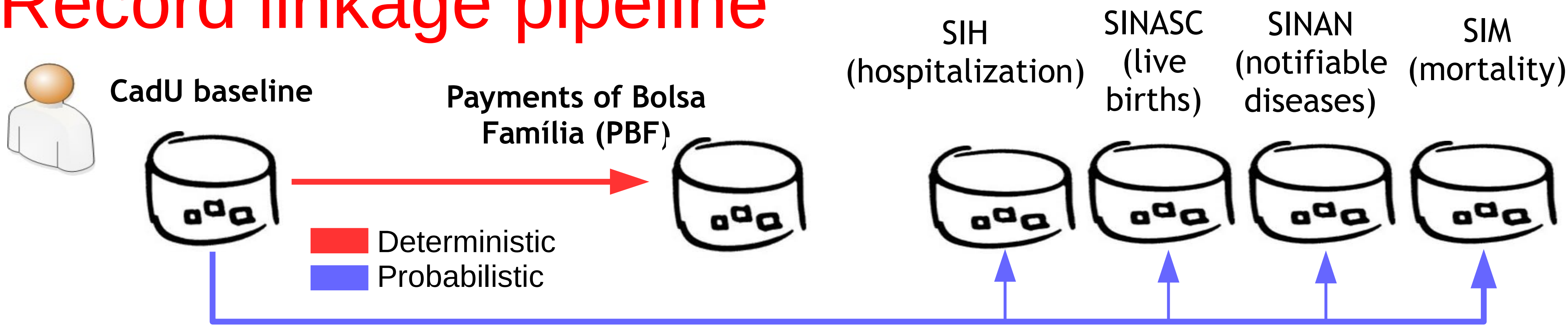


Vb00
cod_fam
munic_ibge_2
munic_ibge_5
num_membro_fmla
data_cad
data_atl
sit_cad
nome
nis_at l
cod_sex
data_nasc
cod_parent
nome_mae
munic_ibge_nasc
nis_orig

	Table	Filesize	# of lines	Version
2007	A	11,4GB	21.028.364	V6
	B	86,8GB	79.050.446	
2008	A	12,5GB	22.767.472	
	B	100,1GB	89.915.568	
2009	A	13,5GB	24.661.693	
	B	108,8GB	97.640.845	
2010	A	14,3GB	26.107.223	
	B	114,4GB	102.663.287	
2011	1	25GB	27.014.194	V7
	4	4,3GB	106.433.938	
2012	1	11GB	30.268.867	
	4	27GB	115.636.503	
2013	1	6.5GB	32.897.120	
	4	29GB	123.116.446	
2014	1	7.1GB	35.439.015	
	4	34GB	130.430.300	
2015	1	7.6GB	35.439.015	
	4	36GB	136.368.326	



Record linkage pipeline



Data quality assessment

- CadU baseline + SUS files
- Metrics for qualitative analysis
- Candidate attributes for linkage

SPSS, R, STATA

Data conditioning

- ETL-based routines (cleansing, standardization)
- Anonymization (Bloom filter)
- Blocking routines
- Comparison blocks

ATYIMO

Record linkage

- Linkage parameters
- Linkage routines (deterministc and probabilistic)
- Data marts

Spark

Accuracy assessment

- Assessment metrics (sensitivity, specificity, VPP etc)
- Controlled scenarios
- Accuracy results

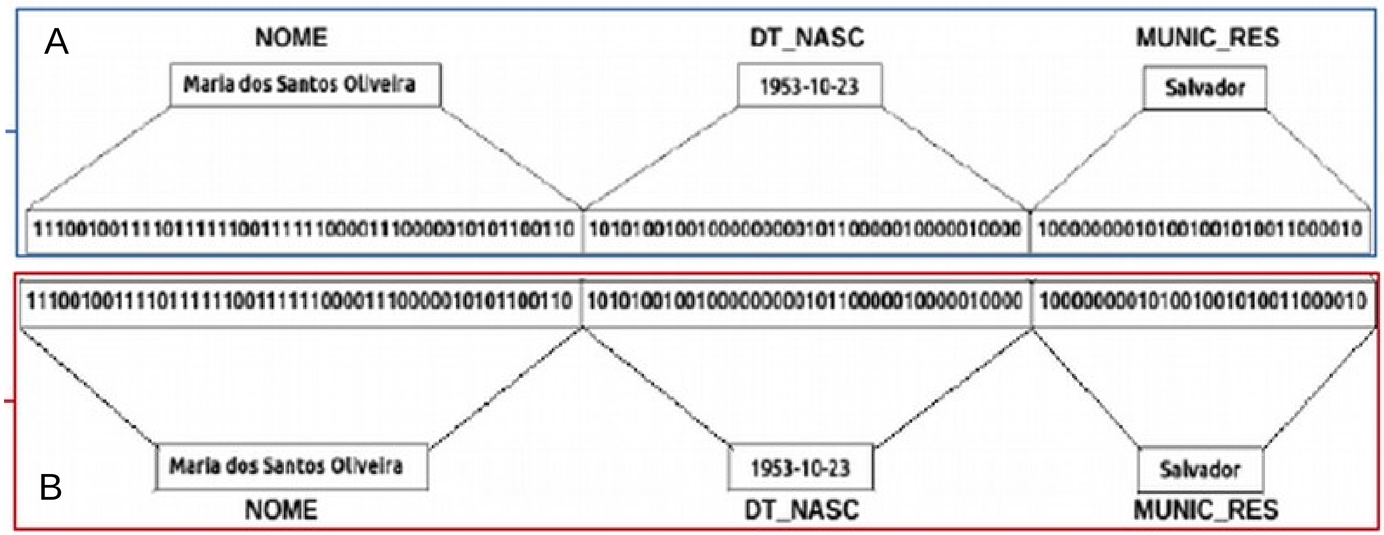
SPSS, R, STATA

A Spark-based workflow for probabilistic record linkage of healthcare data

PITA, R.; PINTO, C.; MELO, P.; SILVA, M.; BARRETO, M.; RASELLA, D. (BeyondMR - EDBT/ICDT 2015)

Record linkage pipeline - methods

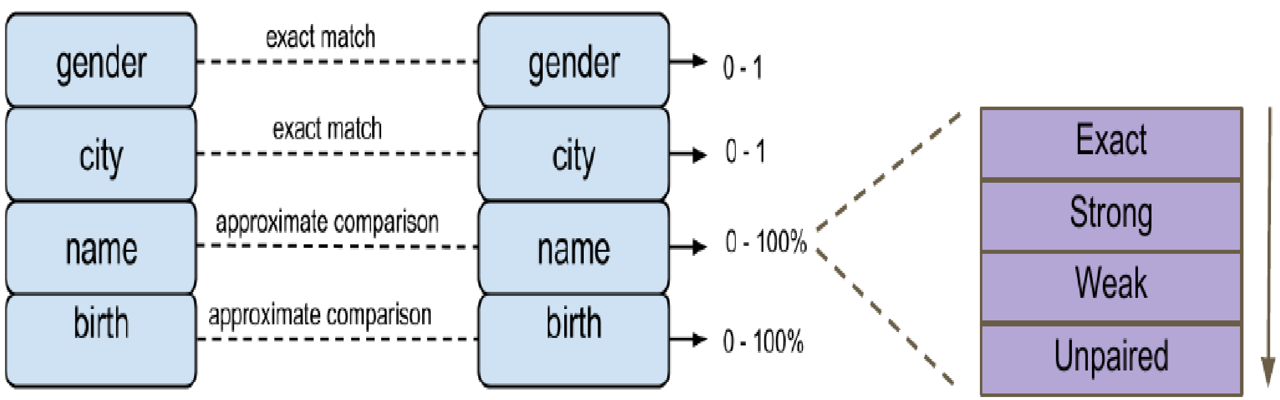
- **Full probabilistic**: Sorensen (Dice) index applied to Bloom filters.



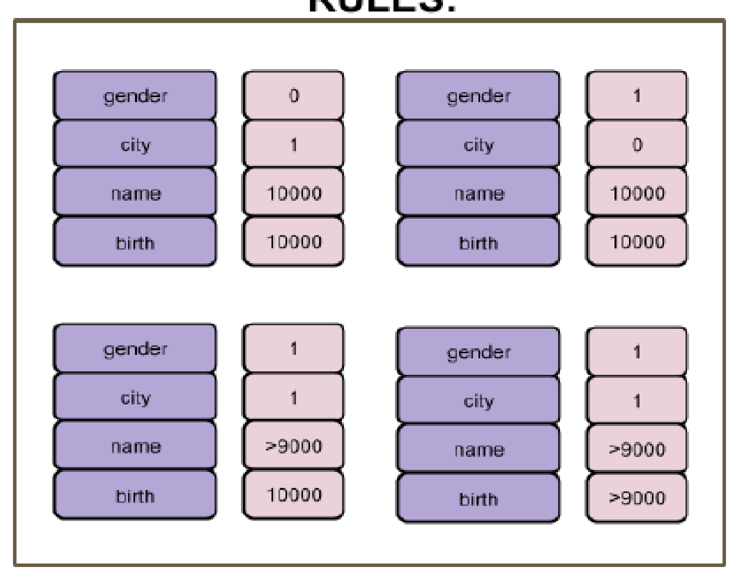
$$D_{a,b} = \frac{2h}{|a| + |b|} = [0,\ 1]$$

h = number of 1's at same position in both Bloom filters
a = number of 1's in Bloom filter A
b = number of 1's in Bloom filter B

- **Hybrid approach**: individual comparison of attributes based on different rules



RULES:



Correlação probabilística de bases de dados governamentais
PINTO, C.; PITA, R.; MELO, P.; SENA, S.; BARRETO, M. (SBBD 2015)

Record linkage pipeline - results

- **Controlled scenario**: 2 databases

Databases	Total # of records	True matches
Rotavirus (diarrhea)	686	486 (positive exams)
Other causes (children treated at outpatient clinics)	9,678	

- 4 simulated scenarios
  - different percentage of changes in records

	Cenário 1 (10,3%)	Cenário 2 (11,3%)	Cenário 3 (10,3%)	Cenário 4 (5,15%)
Tradicional (sem blocos)	482	481	479	482
Tradicional (com blocos)	444	332	466	458
Alternativo (sem blocos)	482	482	480	486
Alternativo (com blocos)	482	482	472	486

Tradicional = full prob. Alternativo = hybrid prob.

- Main metrics:
  - Sensitivity ('sensibilidade')
  - Positive predictive value (VPP)

	Com blocagem		Sem blocagem	
Dice	**Sensibilidade (%)**	**VPP (%)**	**Sensibilidade (%)**	**VPP (%)**
10000	69.3	100.0	8.8	100.0
9800	71.2	100.0	12.8	100.0
9600	75.3	100.0	59.5	100.0
9400	79.4	100.0	86.6	100.0
9200	82.3	100.0	95.3	100.0
9000	86.4	100.0	98.1	100.0
8800	91.4	100.0	98.8	100.0
8600	91.4	100.0	99.0	100.0
8400	91.4	100.0	99.2	99.8
8200	91.4	100.0	99.2	99.8
8000	91.4	100.0	99.2	99.8
7000	91.4	100.0	99.2	98.2

Record linkage pipeline - results

Databases	Linked pairs		True positives (within linked pairs)	
	Full prob.	Hybrid prob.	Full prob.	Hybrid prob.
CadU (tuberculosis) X SINAN (Sergipe)	398	311	309 (77,63%)	299 (96,14%)
CadU (tuberculosis) X SINAN (Sta. Catarina)	661	500	551 (83,35%)	462 (92,4%)
CadU (tuberculosis) X SIH (Sergipe)	40	24	23 (57,5%)	23 (95,83%)
CadU (tuberculosis) X SIH (Sta. Catarina)	140	95	83 (59,28%)	86 (90,52%)

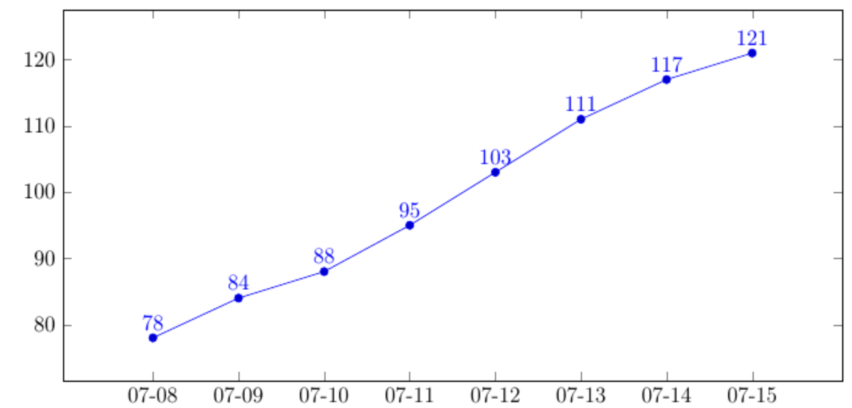
Sergipe: CadU (1,447,512), SIH (49), SINAN (624)
Sta. Catarina: CadU (1,988,599), SIH (330), SINAN (2049)

Databases	Linked pairs	True positives
SIM (Manaus) X BCG vaccination (Manaus)	2,264	2,172 (95,9%)

BCG Manaus: 156,331 SIM: 16,260

Dice	Linked pairs	True pairs	Sensitivity (%)	Specificity (%)	PPV (%)
≥ 10000	952	952	43,83	100,00	100,00
≥ 9800	1234	1234	56,81	100,00	100,00
≥ 9600	1680	1678	77,26	97,83	99,88
≥ 9400	1960	1950	89,78	89,13	99,49
≥ 9200	2151	2119	97,56	65,22	98,51
≥ 9110	2247	2169	99,86	15,22	96,53

Current efforts



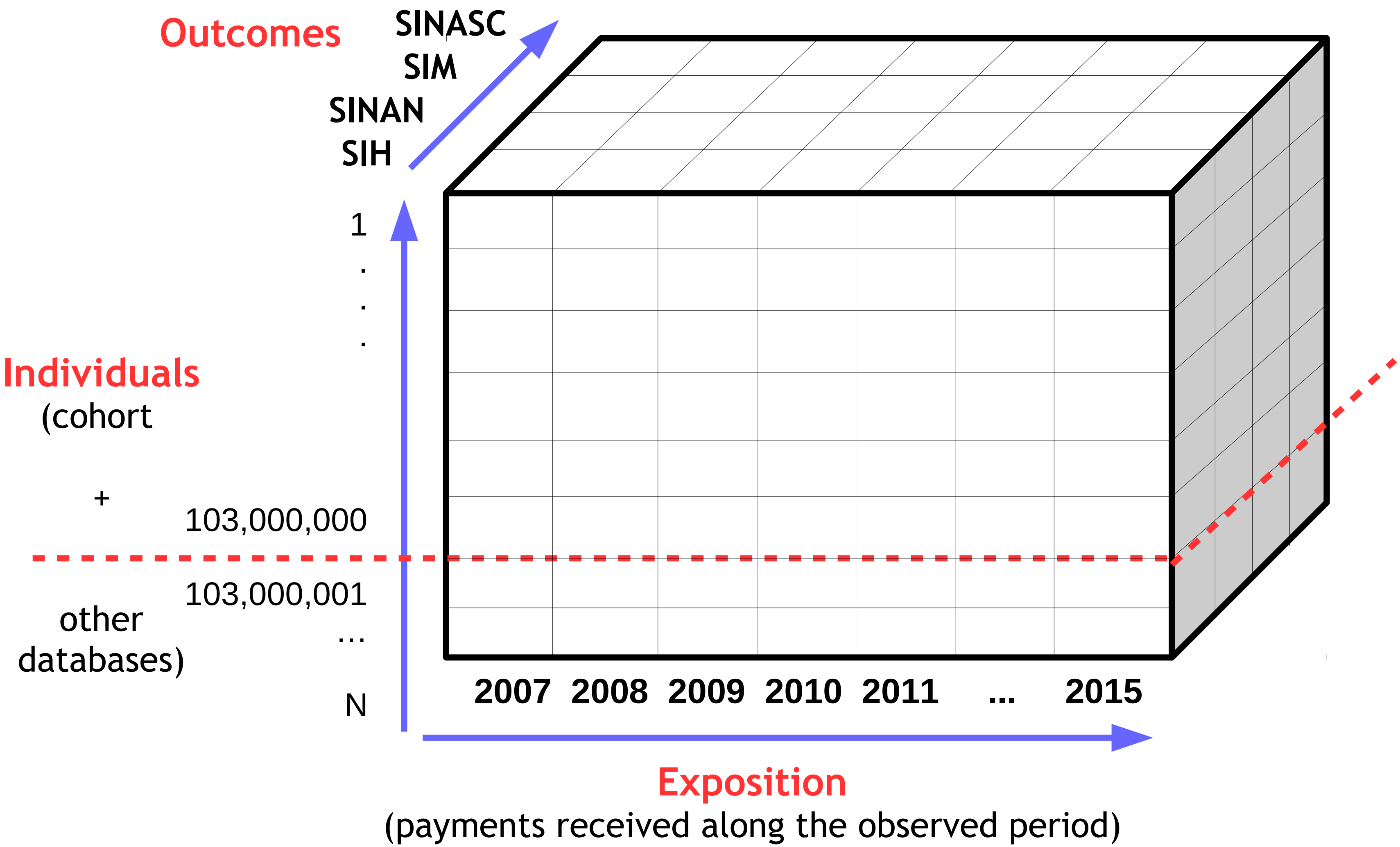
- **Cohort setup & management**
  - 121 million records found! => deduplication
  - How to deal with family dynamics in CadU?
    - People changing NIS code during relisting.
  - Construction of the cohort's profile
    - baseline + variables to be used in each desired study.
    - currently 15 sub-projects (tuberculosis, leprosy, HIV, suicides, nutritional evaluation etc).
- **Scalability tests of linkage routines**
  - Performance and scalability evaluation using the CIMATEC's Yemoja (#2 in LatAM) supercomputer

Number of nodes (20 cores / node)	Larger Database (# records)	Smaller Database (# records)	Execution Time (seconds)
30	1,000	1,000	11
30	100,000	1,000	24
30	500,000	1,000	98
40	1 million	1,000	240
100	40 million	1,000	200
60	81 million	1,000	

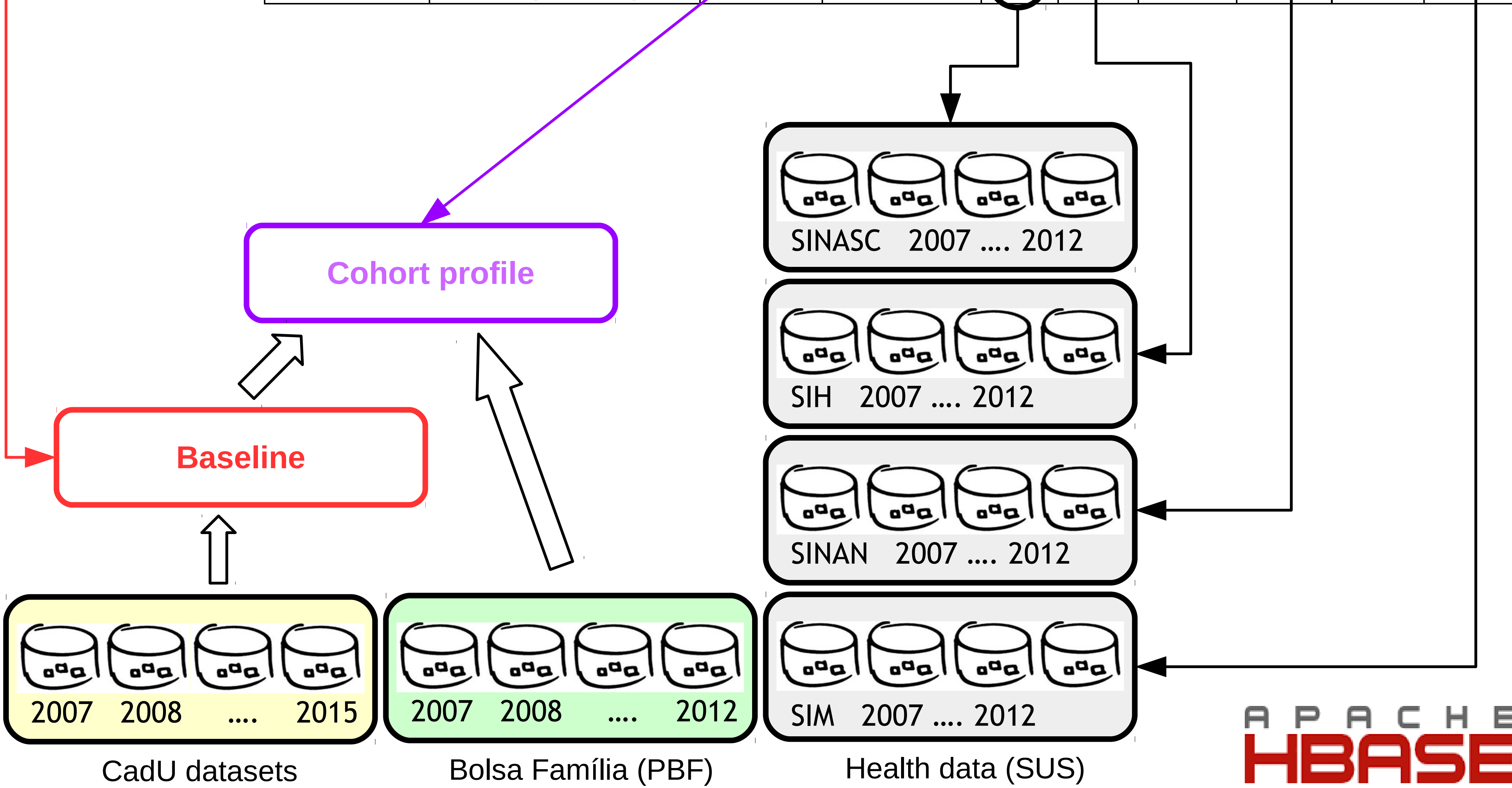
java.lang.OutOfMemoryError

Current efforts

- **Metadata & indexing**



Individuals	Datasets	Baseline	Exposition	Outcomes					
	C | B | H | N | M | A			2007	2008	2009	2010	2011	2012
123	CadastroÚnico	X							
	Bolsa Família		X						
	SIH (hospitalization)				X	X		X	X
	SINAN (notifications)					X		X	
	SIM (mortality)								X
	SINASC (live births)			X		X			



Workshop “Data Bahia”
UFBA, Fiocruz, Sorbonne Universités
Salvador de Bahia, Brazil
April 25-26, 2016

Thank you!

Merci beaucoup!

Marcos Barreto
Distributed Systems Laboratory (LaSiD)
Computer Science Department (DCC)
Federal University of Bahia (UFBA)

marcosb@ufba.br
http://www.dcc.ufba.br/~marcoseb