Workshop on Innovative Delivery Systems in Particle Therapy

23-25 February 2017 *Molecular Biotechnology Center*

Europe/Rome timezone

Online dose calculation using GPUs

Simona Giordanengo

Istituto Nazionale di Fisica Nucleare

Torino, February 25, 2017

Online dose calculation using GPUs integrated in the CNAO DDS

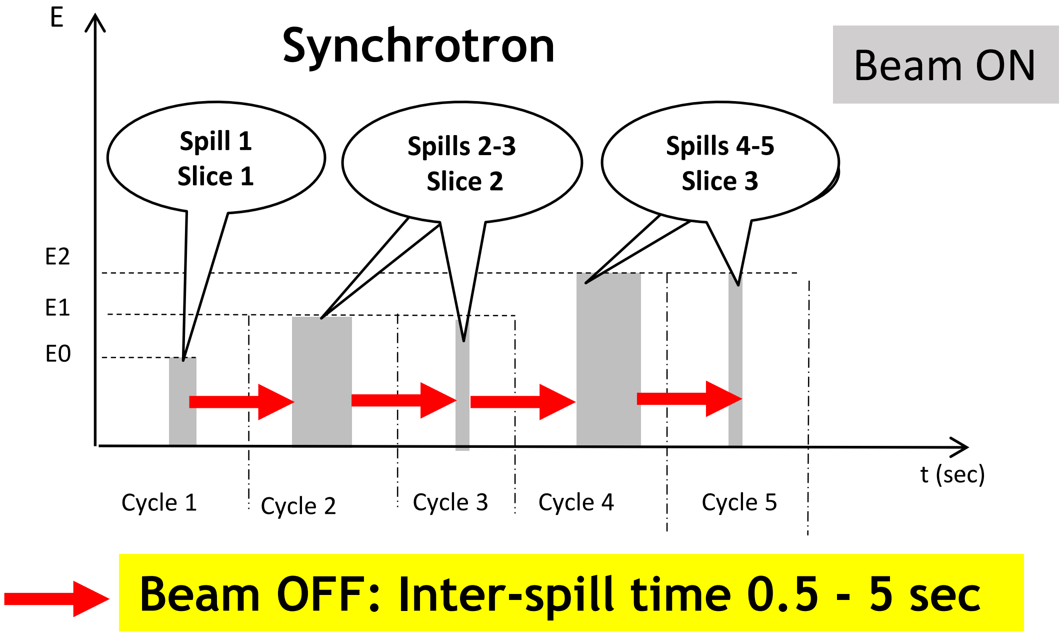
Simona Giordanengo
Istituto Nazionale di Fisica Nucleare
Torino, February 25, 2017

ONLINE → DURING treatment so integrated with the Dose Delivery Systems (DDSs)

When can be computed the dose?

ONLINE → DURING treatment so integrated with the Dose Delivery Systems (DDSs)

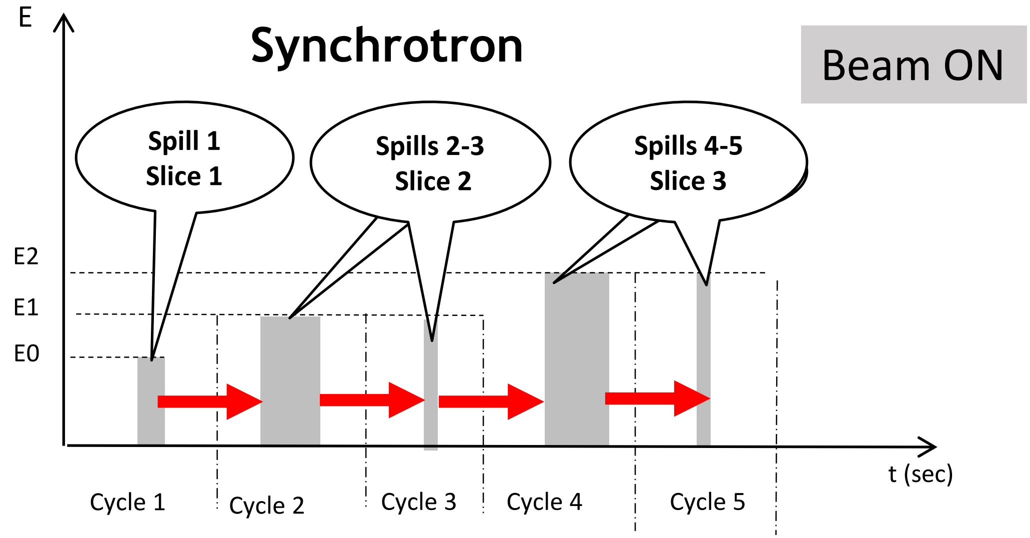
When can be computed the dose?

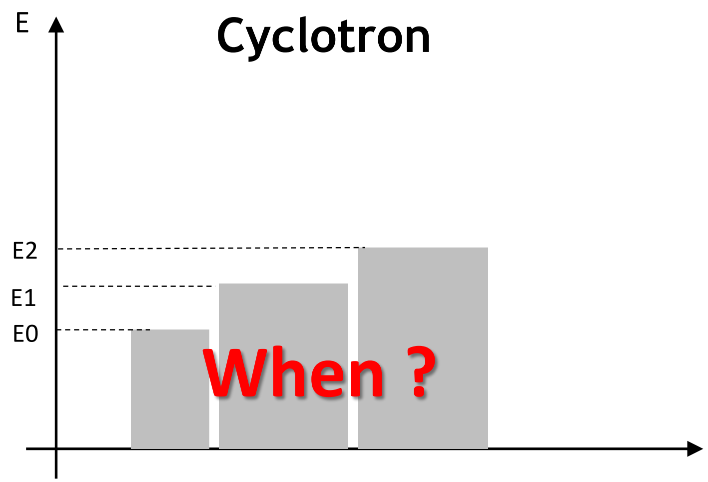


For synchrotron online means during the inter-spill time
→ Synchronous with spills

ONLINE → DURING treatment so integrated with the Dose Delivery Systems (DDSs)

When can be computed the dose?





→ **Beam OFF: Inter-spill time 0.5 - 5 sec**

Beam OFF: tens of ms

For synchrotron online means during the inter-spill time
→ Synchronous with spills

For cyclotron ...?
Could be in parallel with the delivery but not synchronous

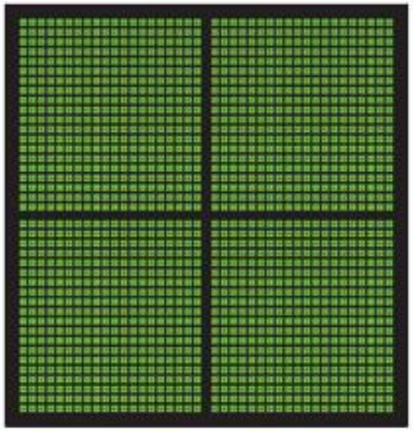
Is this feasible? Yes, using GPUs

GPU-based

- **pencil beam algorithms** (*Da Silva 2015*) and
- **Proton MC packages** (*Yepes 2009, Jia 2012, Renaud 2015, Wan Chan Tseung 2015,*) have recently been developed.

… in place of CPU

CPU
MULTIPLE CORES



GPU
THOUSANDS OF CORES

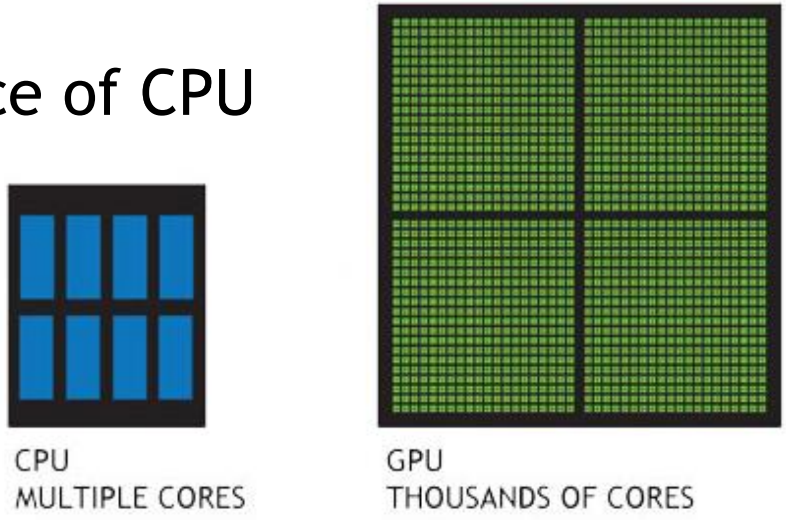
The computation time was reduced (up to a few minutes for MC-based dose simulation) via both **hardware accelerations** and the use of **simplified physics models**

Is this feasible? Yes, using GPUs

GPU-based

… in place of CPU

- **pencil beam algorithms** (*Da Silva 2015*) and
- **Proton MC packages** (*Yepes 2009, Jia 2012, Renaud 2015, Wan Chan Tseung 2015,*) have recently been developed.



PRECISION vs computation TIME is the relevant TRADE-OFF

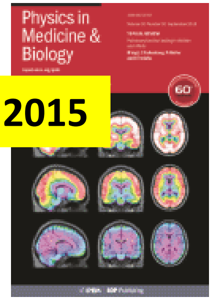
Fast MC dose calculations shows great promise for accurate **inverse plan optimization** as well as for **pre** and **post treatment independent dose calculation for dose verification in place of measurements**

BUT …. are NOT yet SUITABLE for ONLINE applications.

Sub-second forward dose calculation: Is it feasible? Yes for Pencil-Beam algorithms

Sub-second pencil beam dose calculation on GPU for adaptive proton therapy

Joakim Da Silva et al.
University of Cambridge, UK
Phys. Med. Biol. 60 (2015) *4777*



PMB 2015

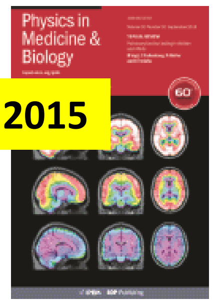
- Tesla K40 GPU (2880 cores @ 845 MHz)
- Skull base test case
  - 2 oblique beam directions
  - 38 + 45 energy layers, total 6776 spots
  - 2x2x2 mm^3 dose resolution, 1x1x1 mm^3 ray resolution
- **Total calculation time 0.22 seconds**
- **Individual energy layers: 2.2-6.4 ms**

Sub-second forward dose calculation: Is it feasible? Yes for Pencil-Beam algorithms

Sub-second pencil beam dose calculation on GPU for adaptive proton therapy

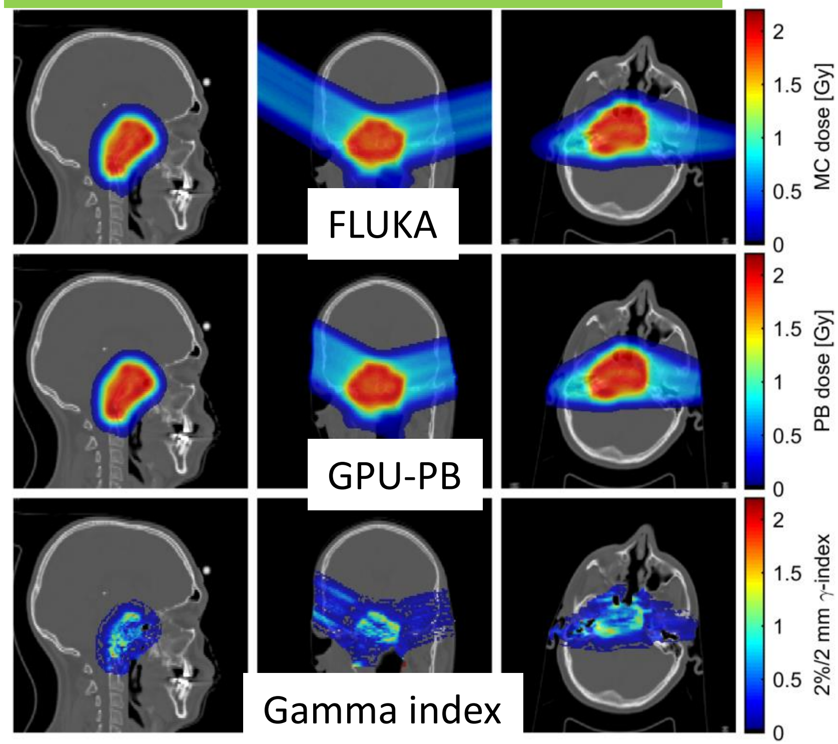
Joakim Da Silva et al.
University of Cambridge, UK
***Phys. Med. Biol. 60 (2015)** 4777*

PMB 2015



- Tesla K40 GPU (2880 cores @ 845 MHz)
- Skull base test case
  - 2 oblique beam directions
  - 38 + 45 energy layers, total 6776 spots
  - 2x2x2 mm^3 dose resolution, 1x1x1 mm^3 ray resolution
- **Total calculation time 0.22 seconds**
- **Individual energy layers: 2.2-6.4 ms**

VALIDATED vs Syngo and Fluka



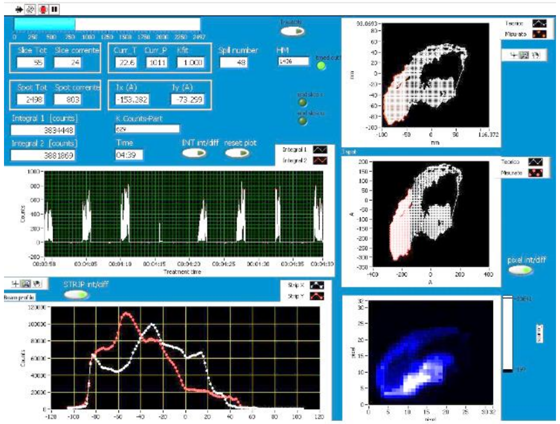
- 2%/2 mm γ-index compared to Fluka
- Passing rate: 96.7% (voxels receiving >10% of max dose)
- Corresponding value for Syngo: 96.8%

The INFN and University of Torino experience on DDS and TPS

It started in 1993 developing beam monitors and MC dose simulations for hadrontherapy

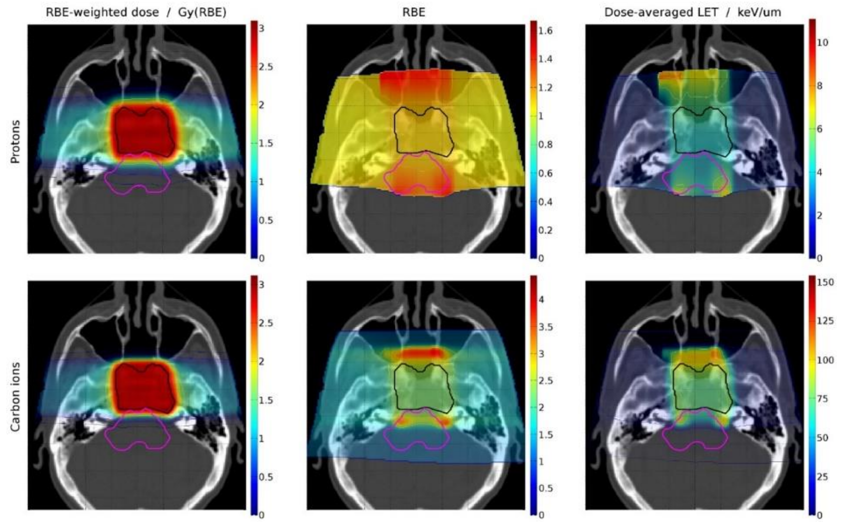
Leading to two important products

The CNAO DDS in collaboration with CNAO Foundation, which is now a CE Medical Device



Giordanengo et al. Med. Phys. 42 (2015)

The PlanKIT TPS in collaboration with IBA



Russo et al. Med. Phys. 61 (2016)

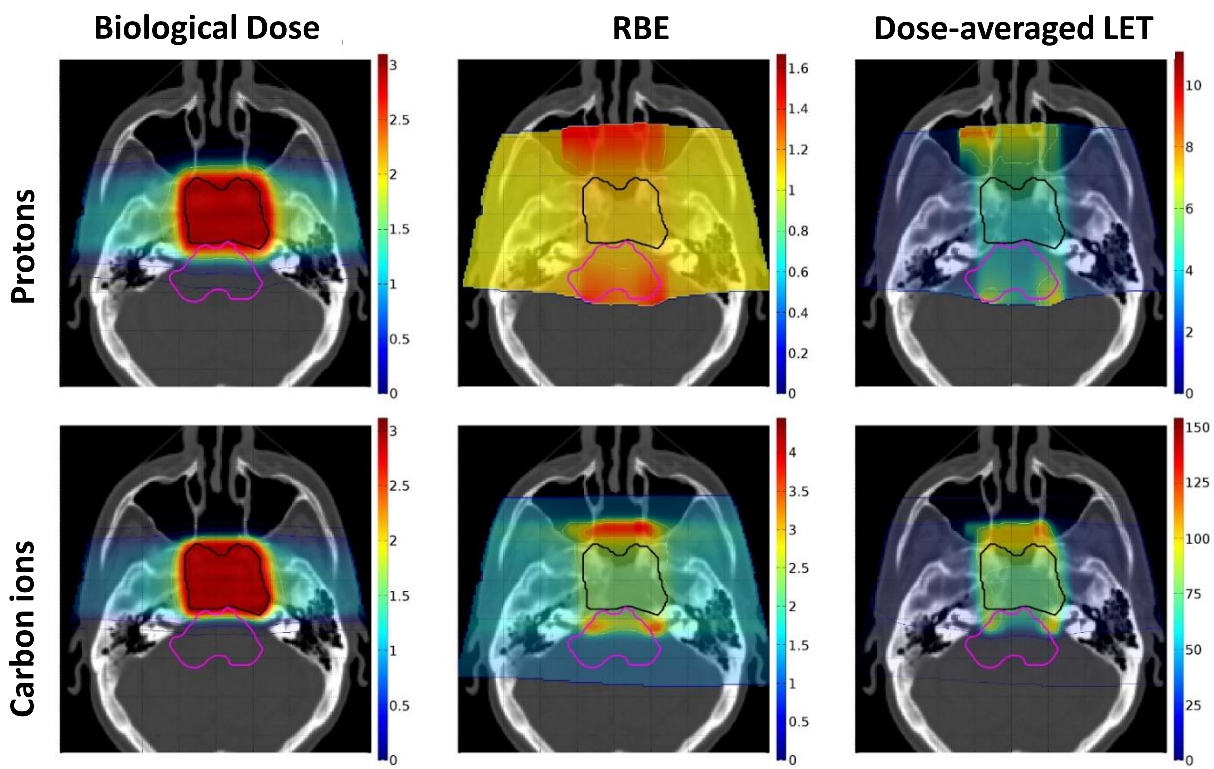
The Dose Engine Kernel (DEK) TPS

The INFN-IBA Beam Model allows computing the 3D effect of **an ion field** incident on a **water-like material.** This model is implemented in a Dose Engine Kernel (DEK) to estimate the outcome of a therapeutic ion irradiation delivered with the Pencil Beam Scanning (PBS) technique.

PlanKIT OUTCOMES:

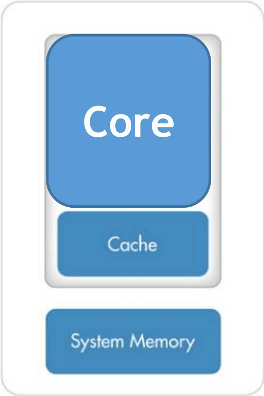
- Physical Dose
- RBE
- Biological Dose
- LET
- Dose-averaged LET
- α
- B
- Survival

For different ions!



In the following DEK doses will refer to CPU-based doses by DEK TPS

GPU-based DEK Forward Planning



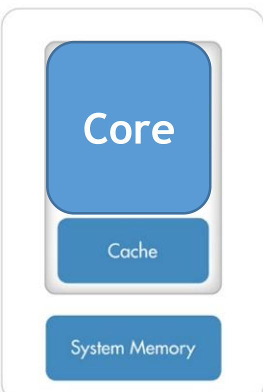
C++ DEK for single core CPU

CPU single core

Not optimize to be fast but to be versatile in order to easily be adapted to different ions, facilities and computational requirements

GPU-based DEK Forward Planning

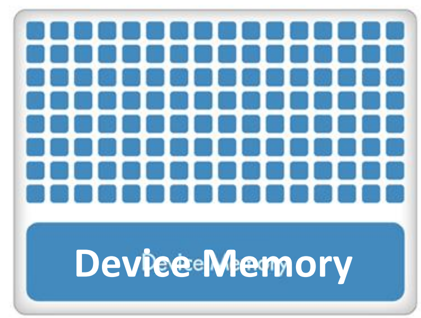
In the Framework of the INFN RIDOS project (2014-2015)
A Fast Forward Planning implementation has been done porting on CUDA the INFN-IBA TPS libraries (SW)



CPU single core

C++ DEK for single core CPU → **CUDA DEK**

GPU-DEK has been developed with CUDA for NVIDIA GPUs



GPU hundreds of cores

Nvidia
Tesla K40

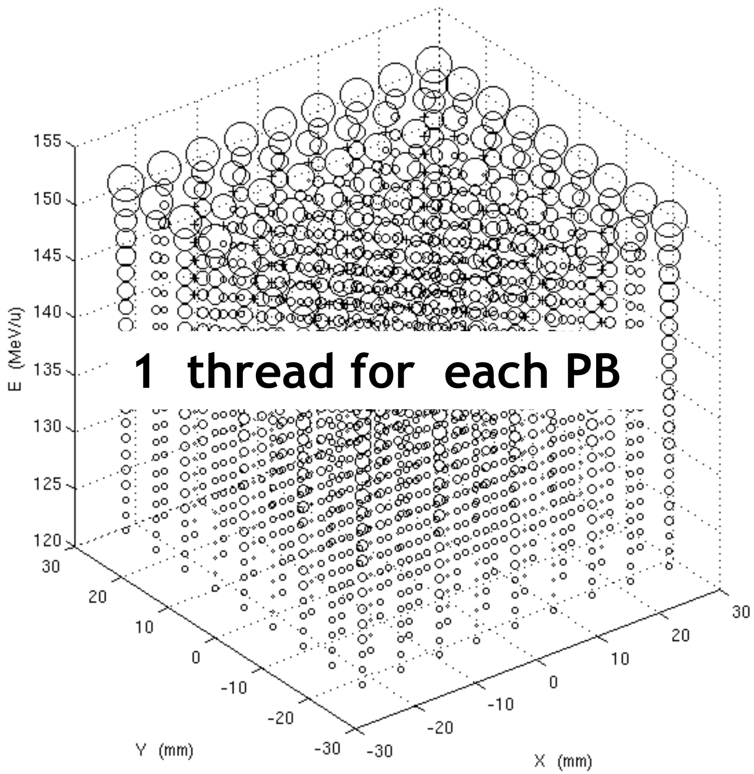
Not optimize to be fast but to be versatile in order to easily be adapted to different ions, facilities and computational requirements

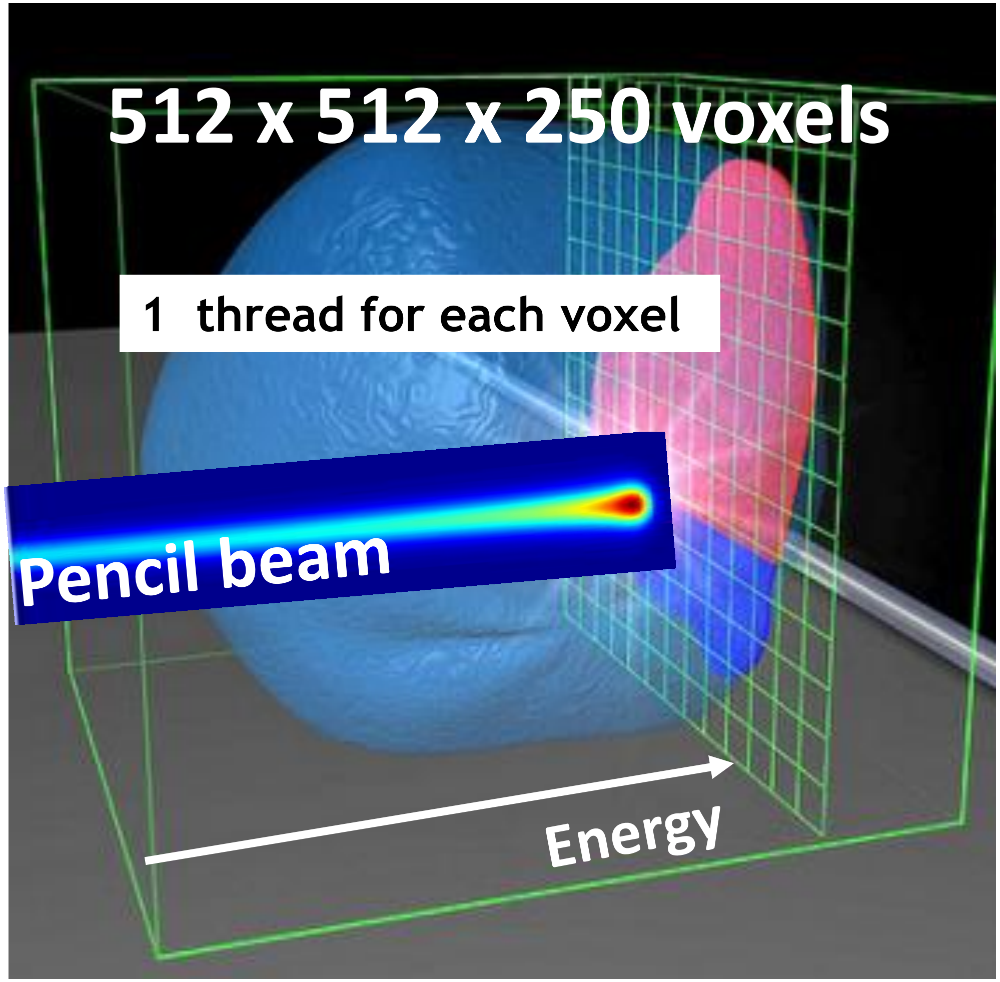
Tesla K40 cost:
11/2013 → 5 k$
Amazon today → 3 k$

Why GPUs? (I)

Thousands narrow beams over thousands voxels

1000 – 50000
pencil beams for field

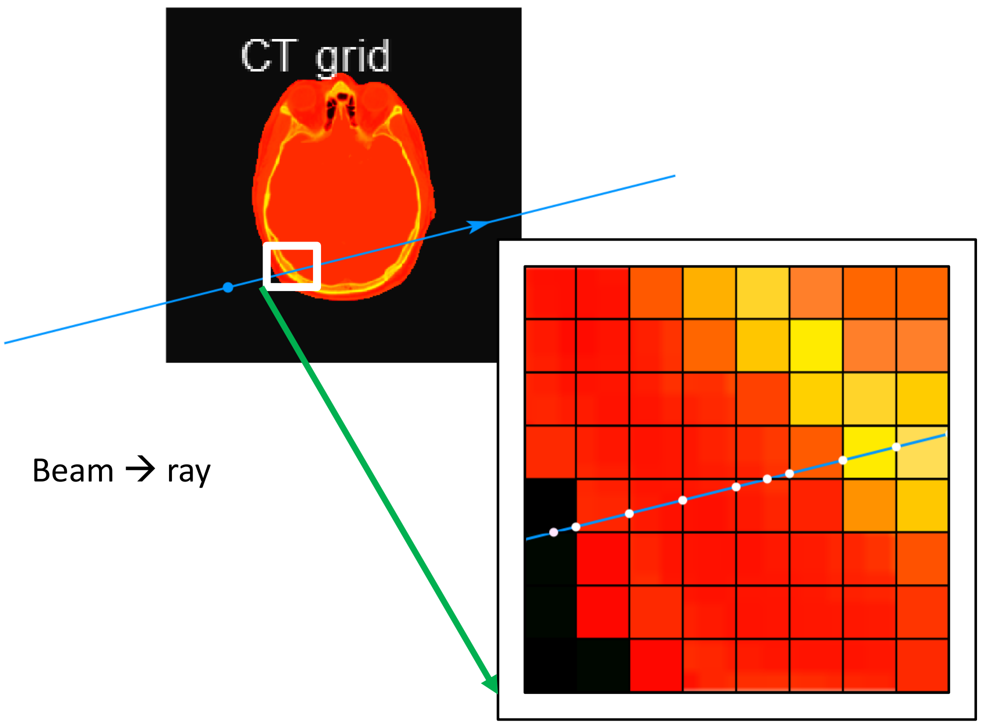


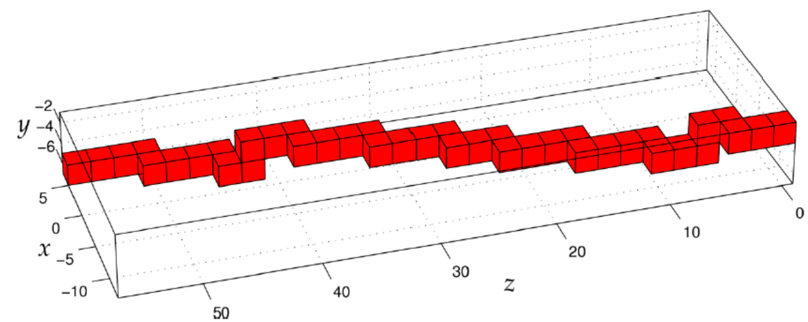


Thread = Sequence of Single Instruction on Multiple Data (SIMD) operations

Algorithms for PBS dose: (I) Raytracing

Raytracing → Find intercepts + sorting intercepts + path length evaluation

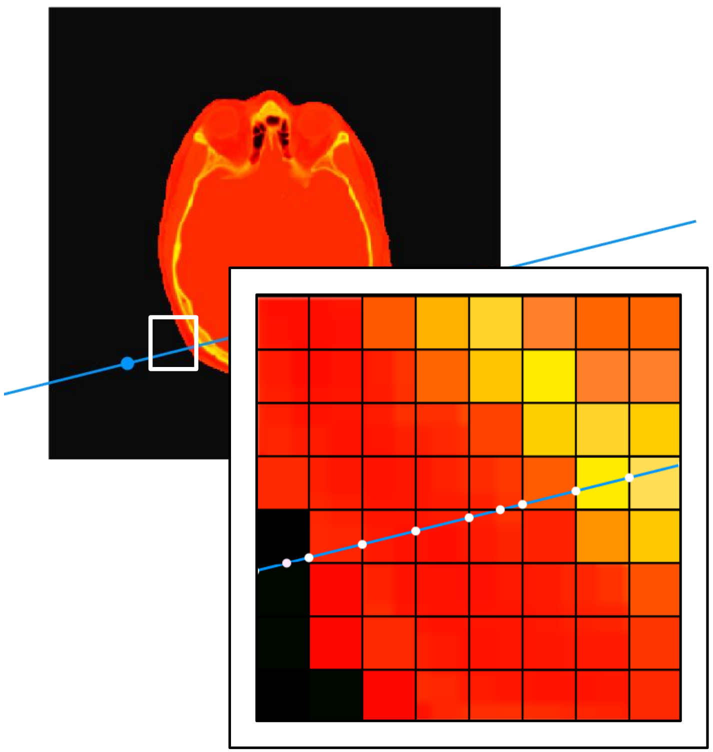




Voxels hit by a ray (pencil beam direction)

Raytracing : Find intercepts with CT voxels

GPU code: check the intercept between the ray and each plan along the 3 voxels direction (stream compaction algorithm)

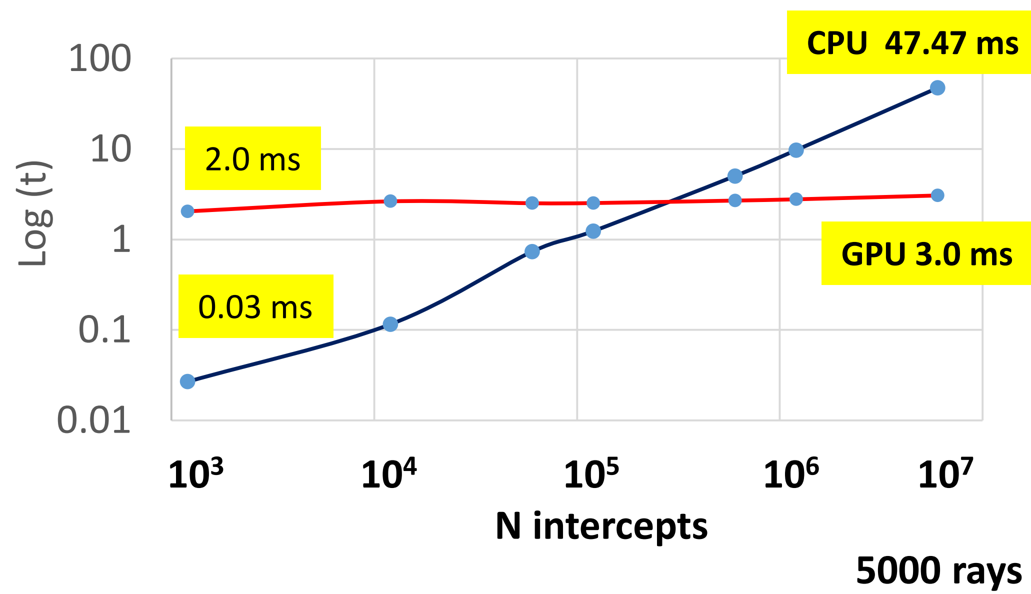


Max N intercepts: $(N_x+N_y+N_z) \times N_{ray}$

N = CT voxels ~ (512+512+256)

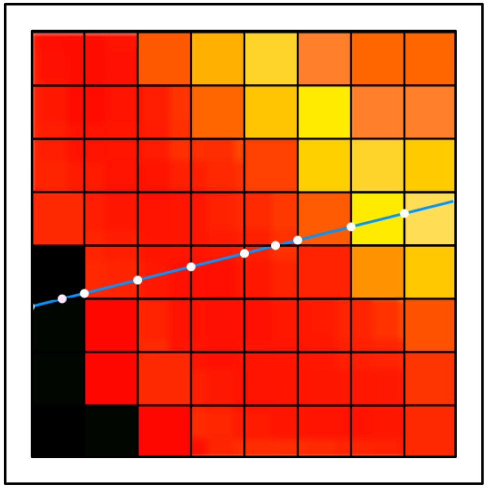
N_{ray} (Min ÷ Max) = 1 ÷ 5000000

N intercepts range : $10^3 \div 5 \times 10^9$

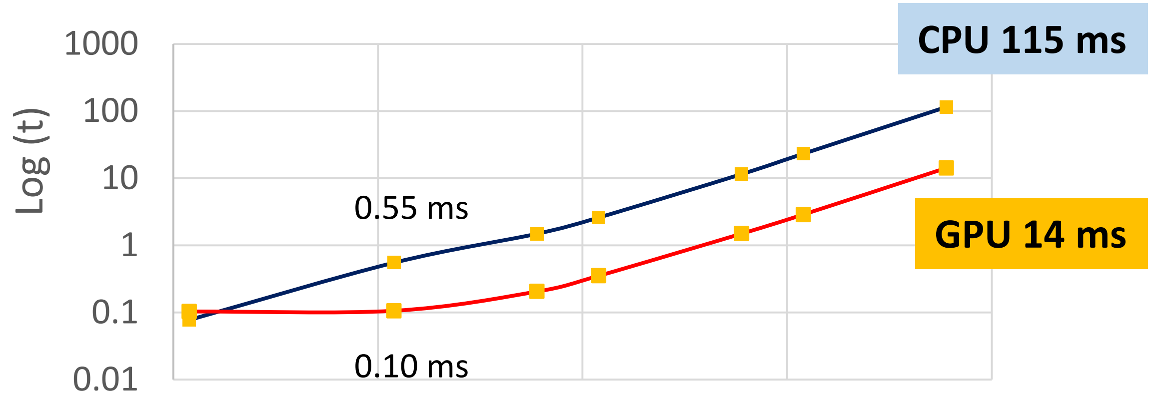


CPU: Xeon E5-1620v2-4C-16GB
GPU: Nvidia TESLA-K20c.

Raytracing(II-III): Sorting the intercepts and measuring the ray path length (L_i)

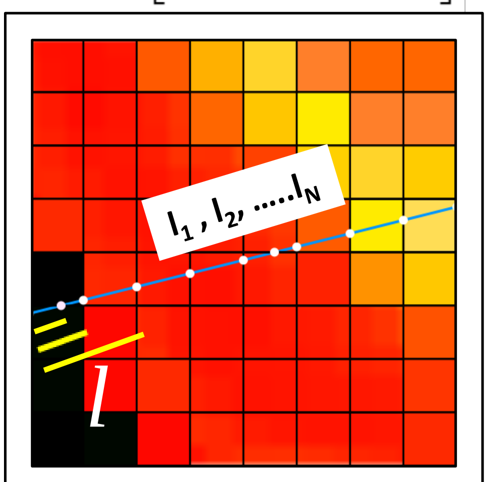


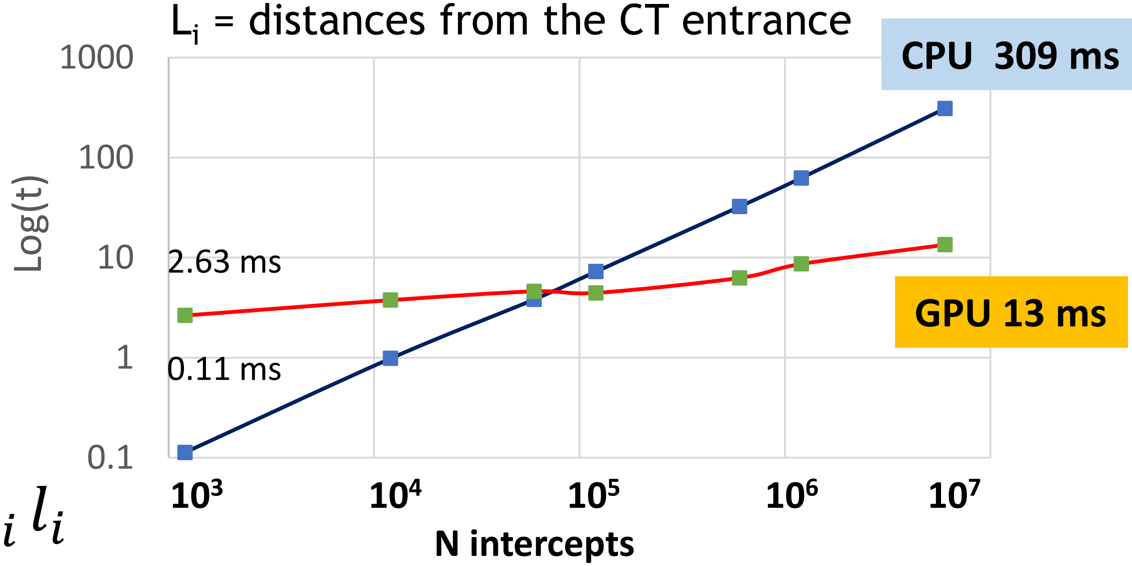
Bitonic Sort : the intercepts are sorted in function of the distance from the CT entrance



L_i = $\left[l(\vec{r}_{i+1}) - l(\vec{r}_i)\right]$

L_i = distances from the CT entrance





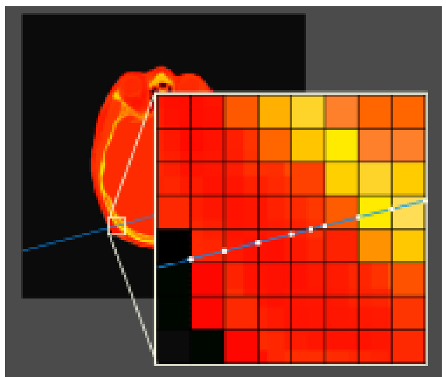
$L = \sum_i l_i$

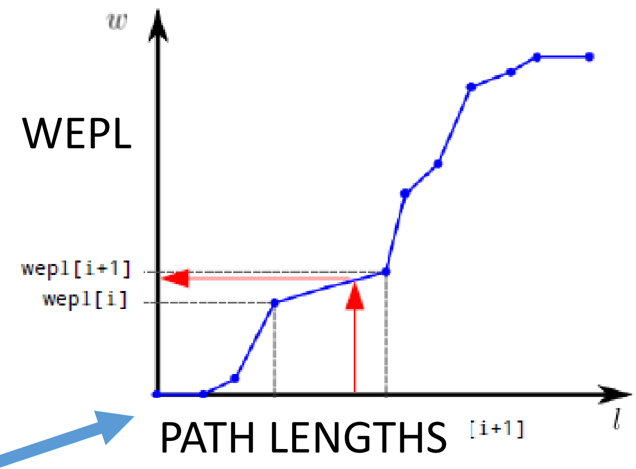
Why GPUs? (II)

Interpolations exploit the use of GPU-Texture memory

WEPL = $w(\vec{r}_{i+1}) = w(\vec{r}_i) + \rho\left(\frac{\vec{r}_{i+1} + \vec{r}_i}{2}\right)\left[l(\vec{r}_{i+1}) - l(\vec{r}_i)\right]$

Water Equivalent Path Length





RayTracing

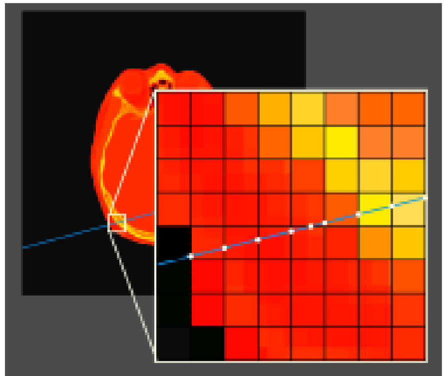
WEPL (1D Interpol)

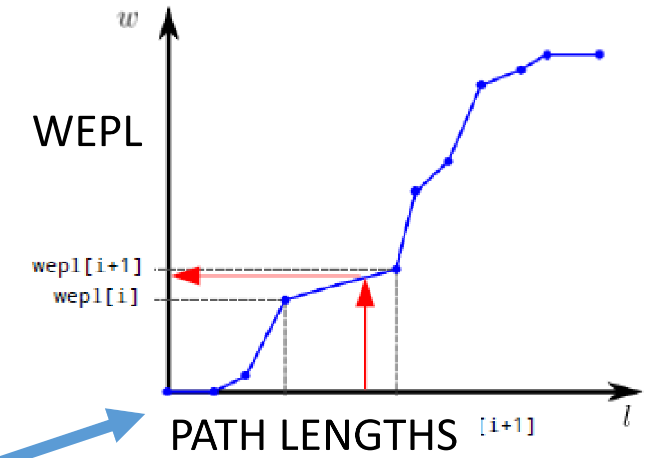
Hardware interpolation using the GPU Texture

Returns for each selected voxel the water-equivalent position in the beam reference system

Why GPUs? (II)

Interpolations exploit the use of GPU-Texture memory





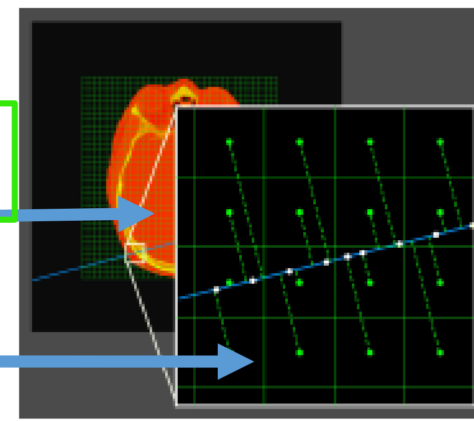
RayTracing

WEPL (1D Interpol)

Computing grid creation

Select voxels within a cut_off for each ray

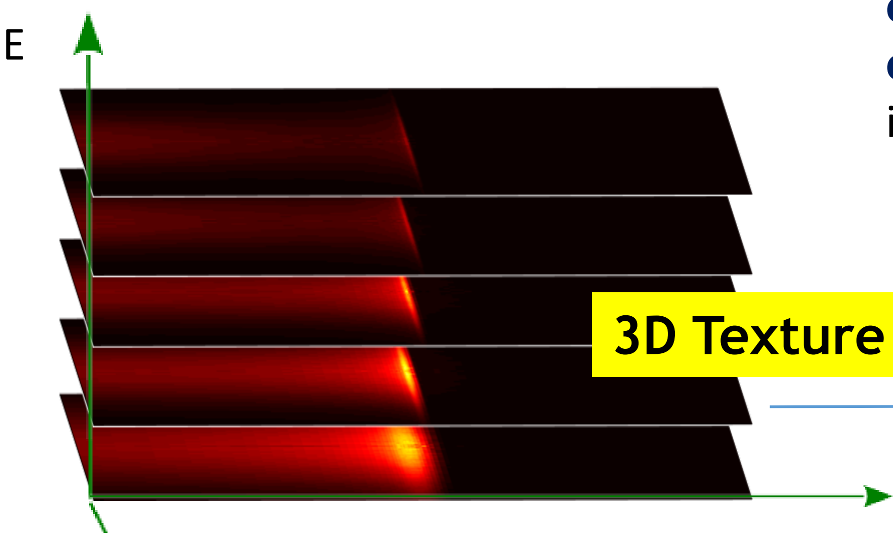
Computing grid (example 255x255x127 voxels)



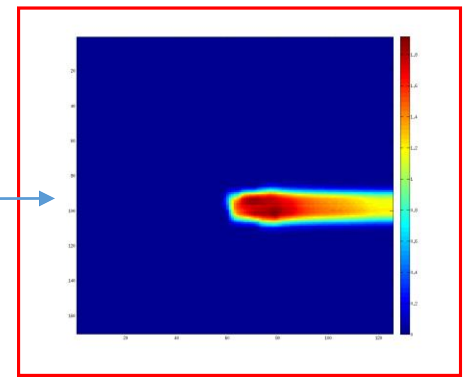
Large cut-off → beam halo is considered → better accuracy
Large cut-off → more voxels → high computation time for standard C++ codes

2nd Texture Interpolation: dose evaluation from pre-computed LUTs

The physical and biological dose contribution for each WEPL voxel **of the computing grid** is obtained through a 3D interpolation of pre-computed LUTs



3D Texture Interpolation



Dose (Gy)
Biological Dose (Gy[RBE])
LET (keV/μm)
Survival probability

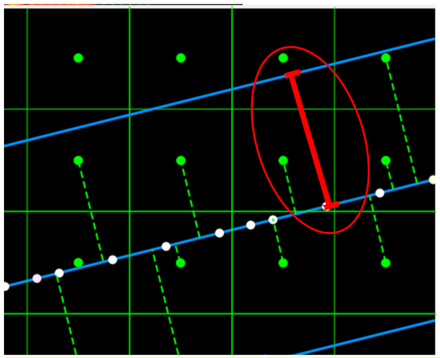
Values available in pre-computed LUTs

$$\left[D^0, \underbrace{\alpha_1 D^0, \sqrt{\beta_1} D^0}_{\text{tissue 1}}, \underbrace{\alpha_2 D^0, \sqrt{\beta_2} D^0}_{\text{tissue 2}}, \ldots\right]$$

Values used for Physical Dose

Values per particle used for Biological Dose

GPU-DEK time performances



Computing Grid = 170x170x125;
N of Rays = 1248; N of slice = 39;

CPU: XeonE5-1620v2-3.7-4C-16GB
GPU: Nvidia TESLA-K20c

Radial cut-off (mm)	Number of voxels	RIDOS F-FP time (s) Entire treatment	RIDOS F-FP time (s) Per spill	PlanKIT FP time (s)	Gain
20	44678	1.35	**0.035**	157	116
40	178702	2.33	**0.059**	636	273
50	279243	2.76	**0.071**	997	361
80	714861	4.68	**0.120**	2534	541

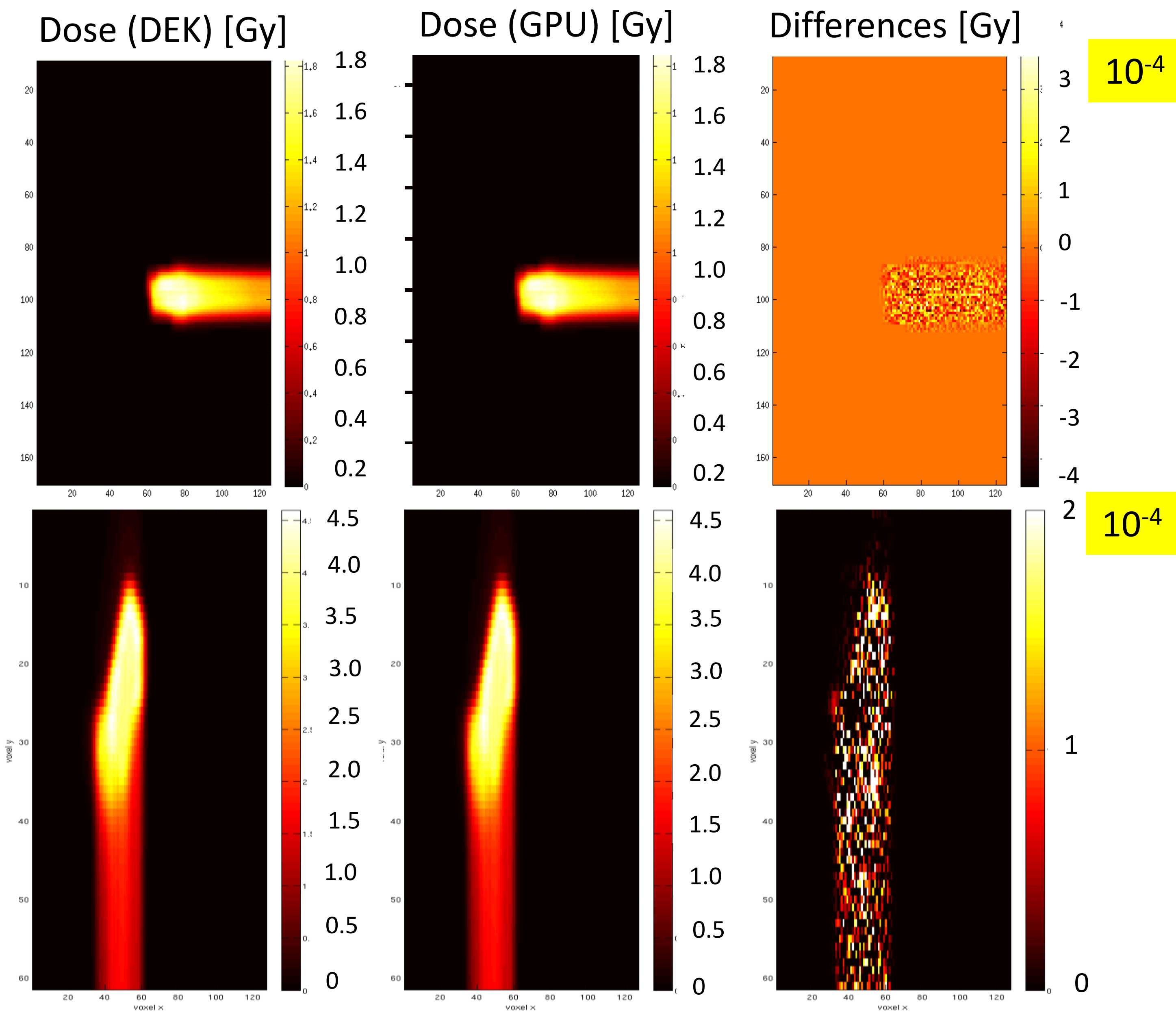
GPU-DEK vs CPU-DEK for clinical patient

Protons

Physical Dose

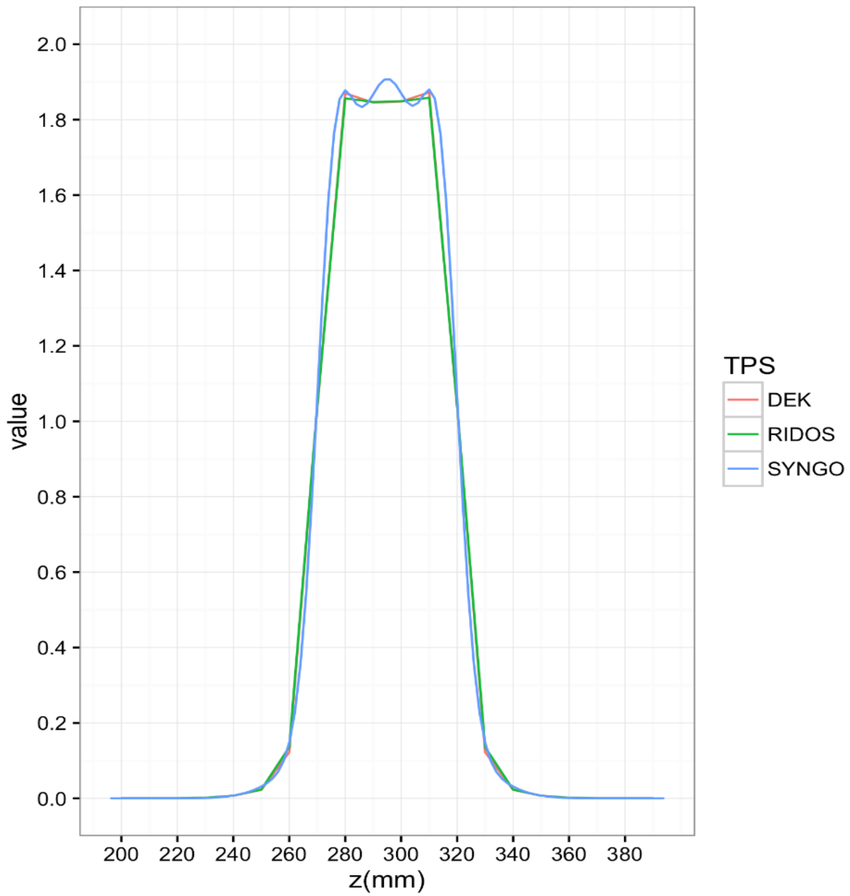
Carbon Ions

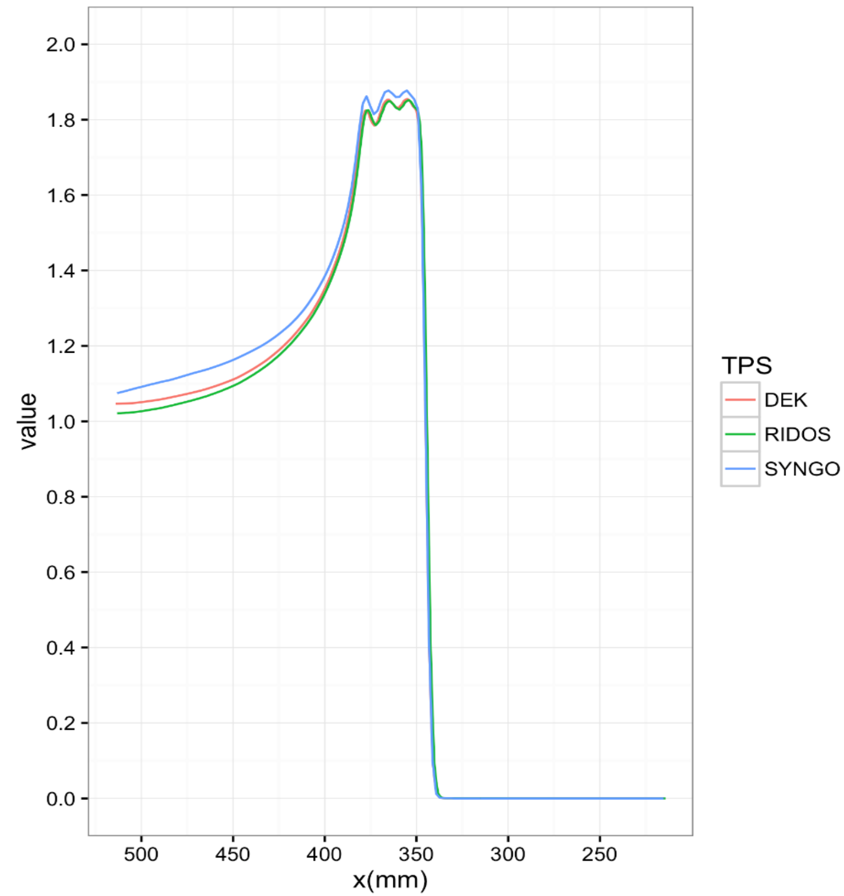
Biological Dose



Com Grid = 170x170x127;
N of Rays = 10761;
N of slice = 51;
N voxels= 7292;
radial cut-off=10 mm.

GPU-DEK vs Syngo TPS for cube in water

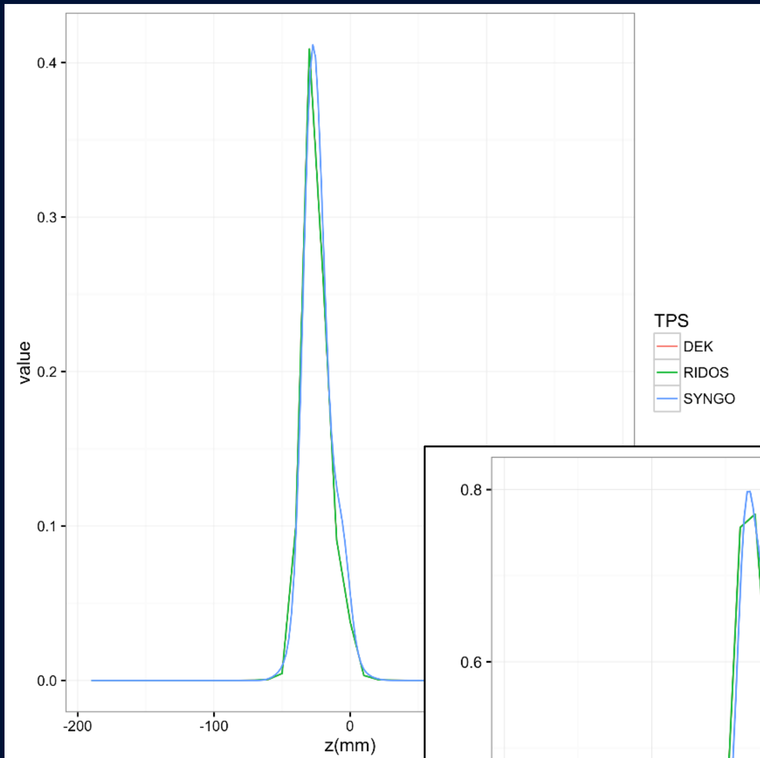


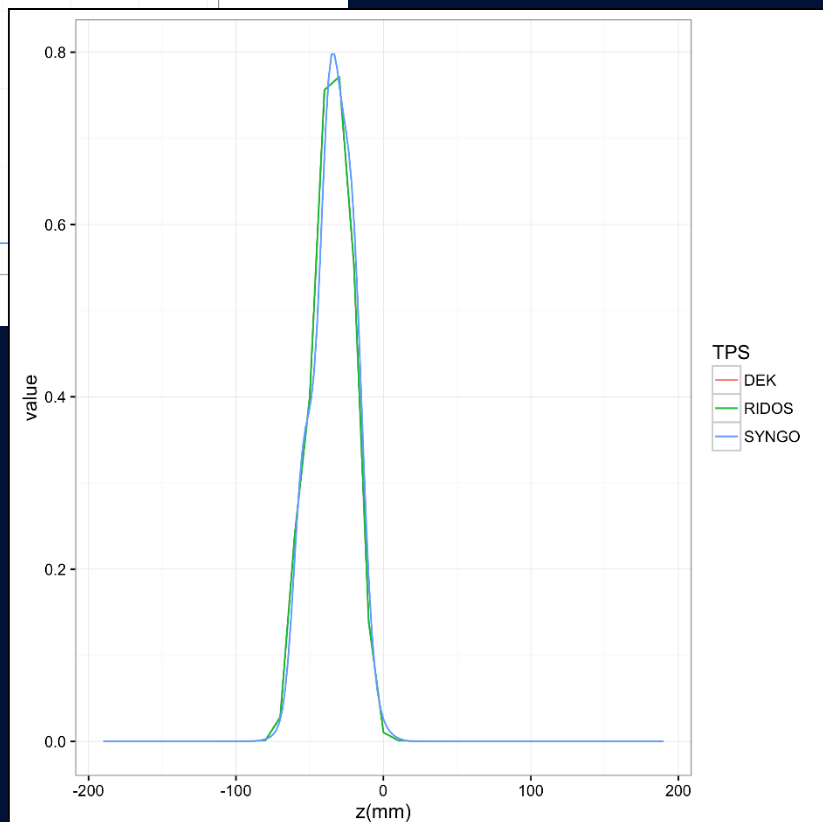


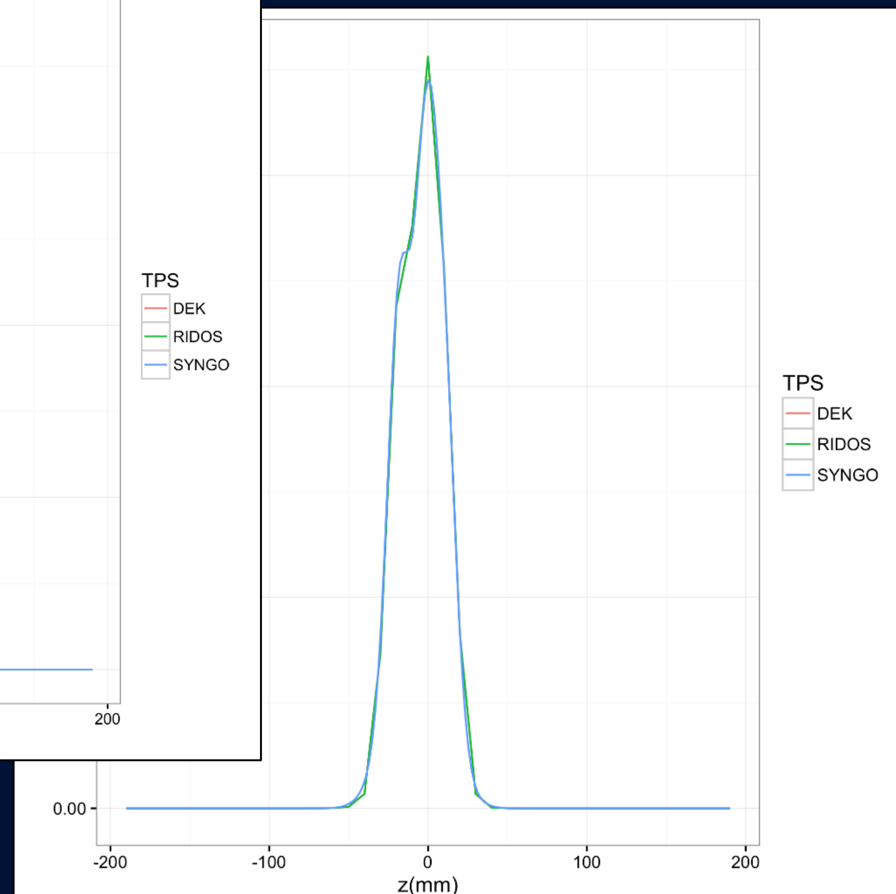
Protons – SOBP21cm Cube 3x3x3 cm^3

GPU-DEK vs SYNGO

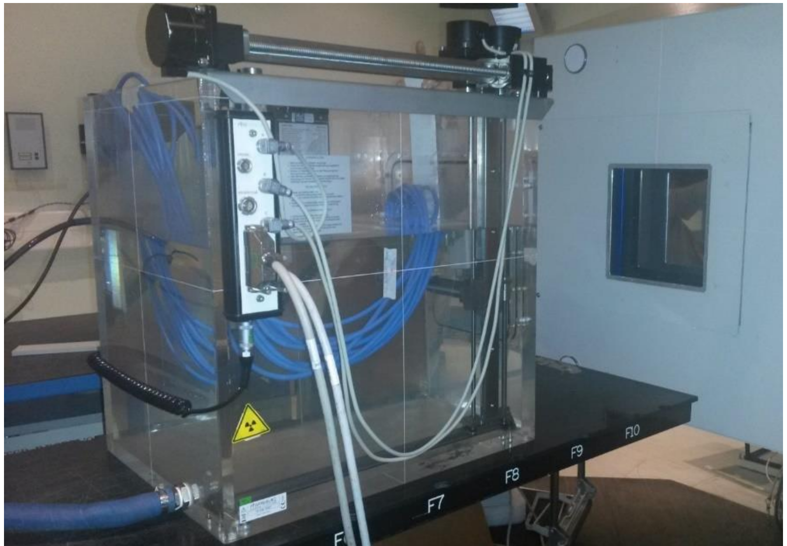
CNAO patient – dose to water proton – Small H&N tumor Lateral profiles

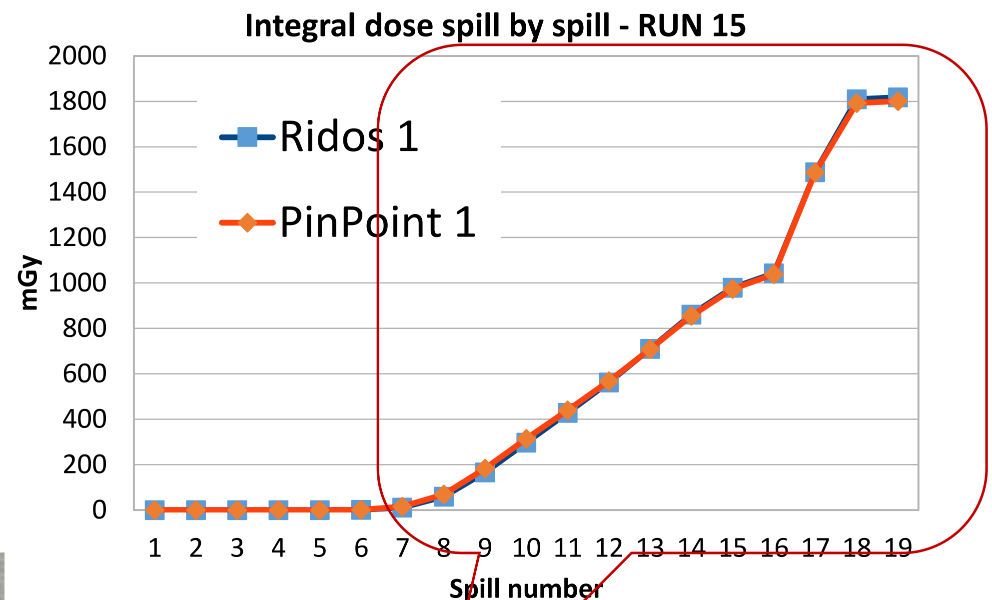


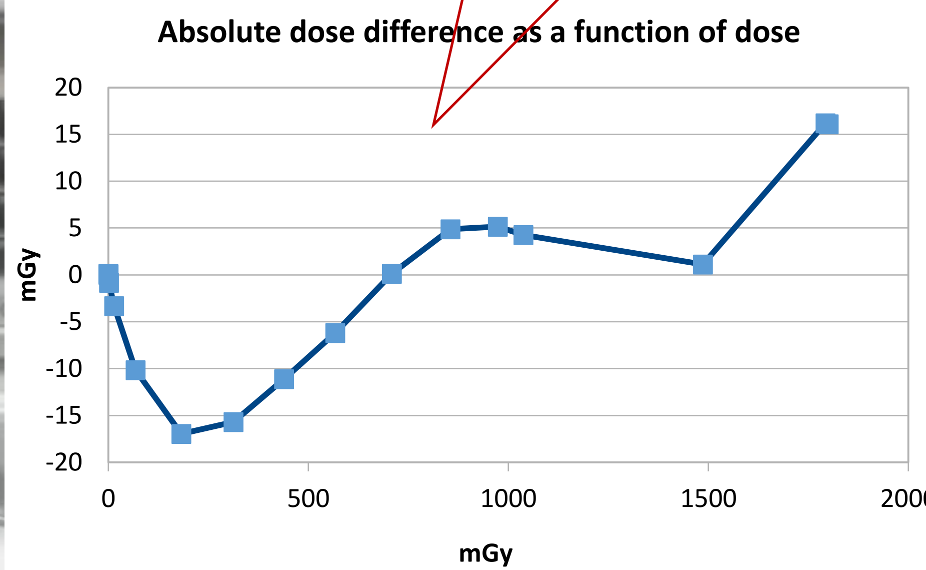




GPU-DEK vs PinPoint measurements







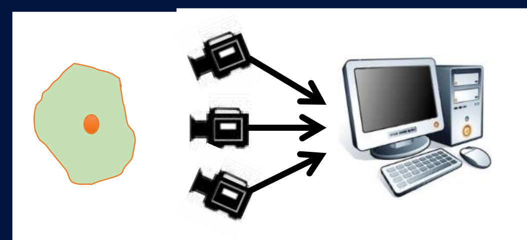
What we can do with so fast dose calculation?

Upgrade the new generation of DDS

Dose Delivery System

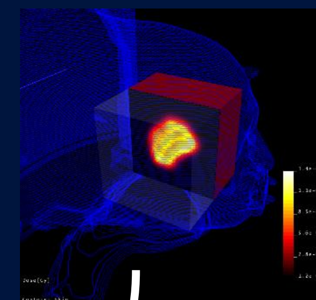
\+

Target Tracking System



\+

Treatment Planning System



New dose Planning and Delivery System

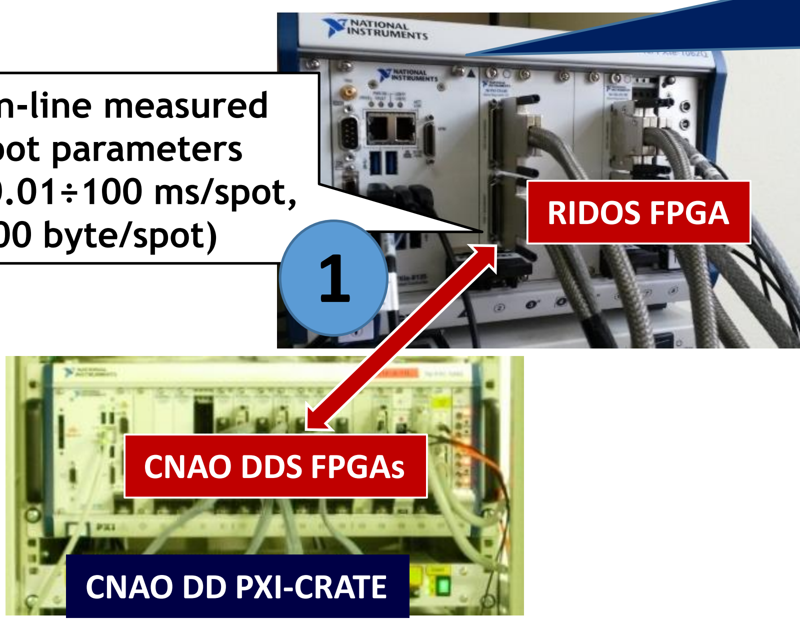
able to check ON-LINE the uncertainties in the delivered DOSE DISTRIBUTIONS due to measured patient and beam deviations from planned conditions.

…Goal of the INFN RIDOS project

Grant for Young Researcher 2014/2015

RIDOS HW and interfaces

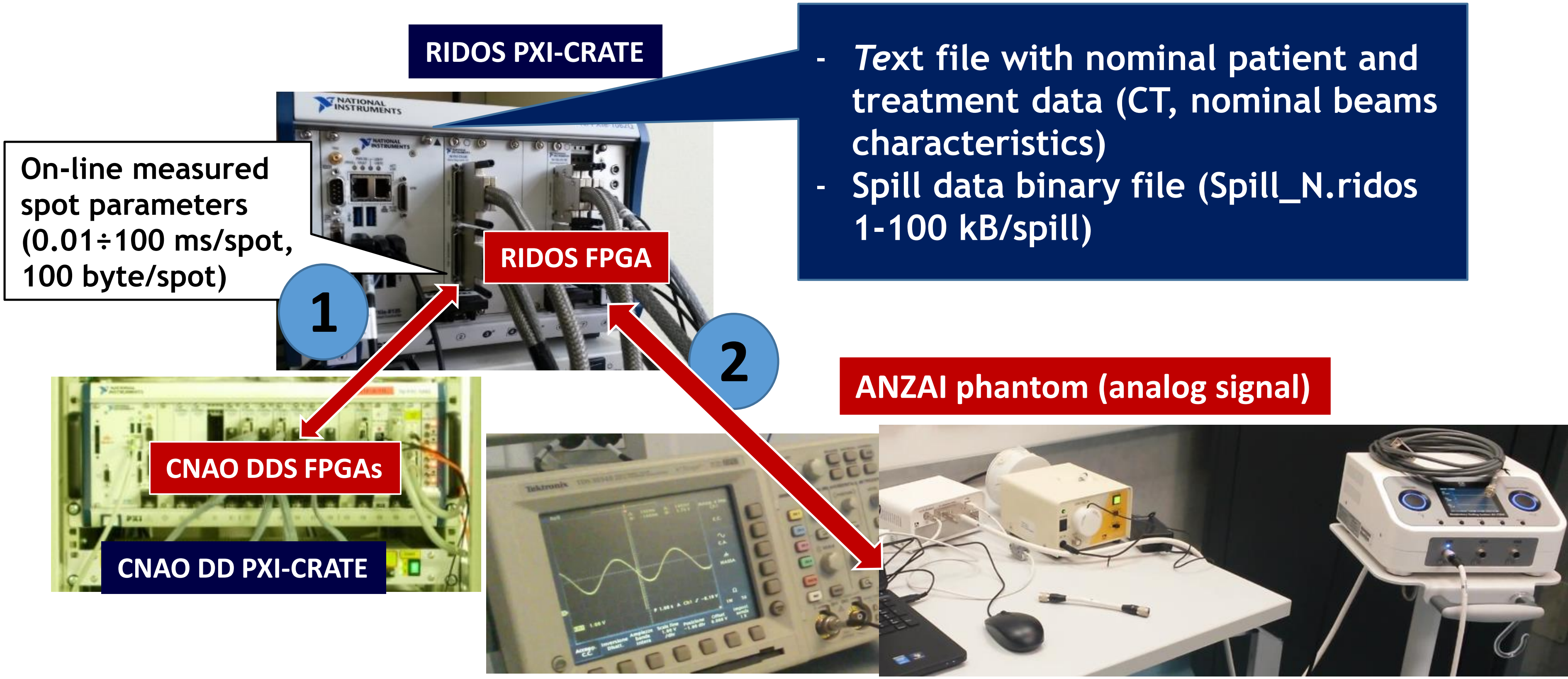
RIDOS PXI-CRATE



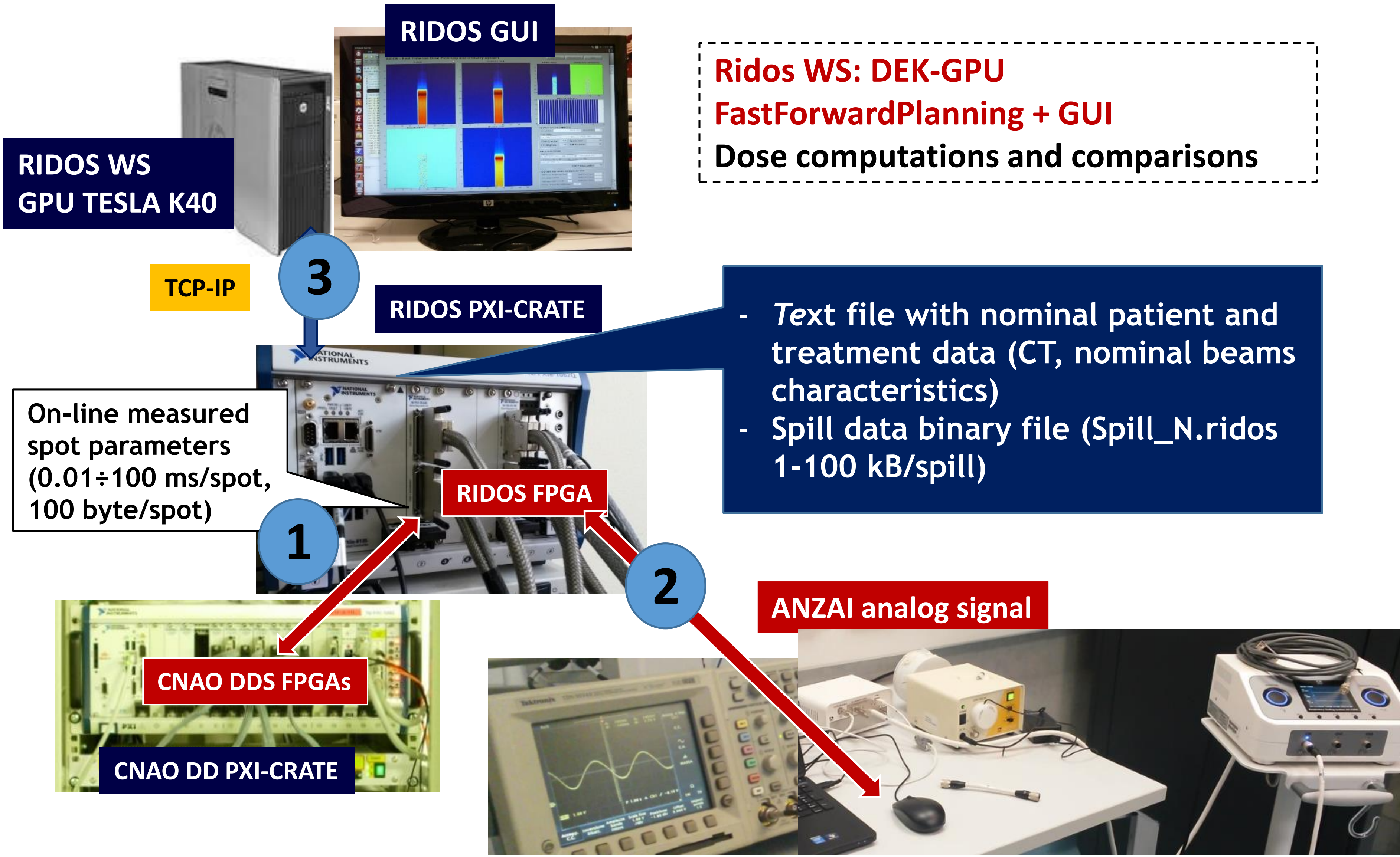
On-line measured spot parameters (0.01÷100 ms/spot, 100 byte/spot)

- *Text* file with nominal patient and treatment data (CT, nominal beams characteristics)
- Spill data binary file (Spill_N.ridos 1-100 kB/spill)

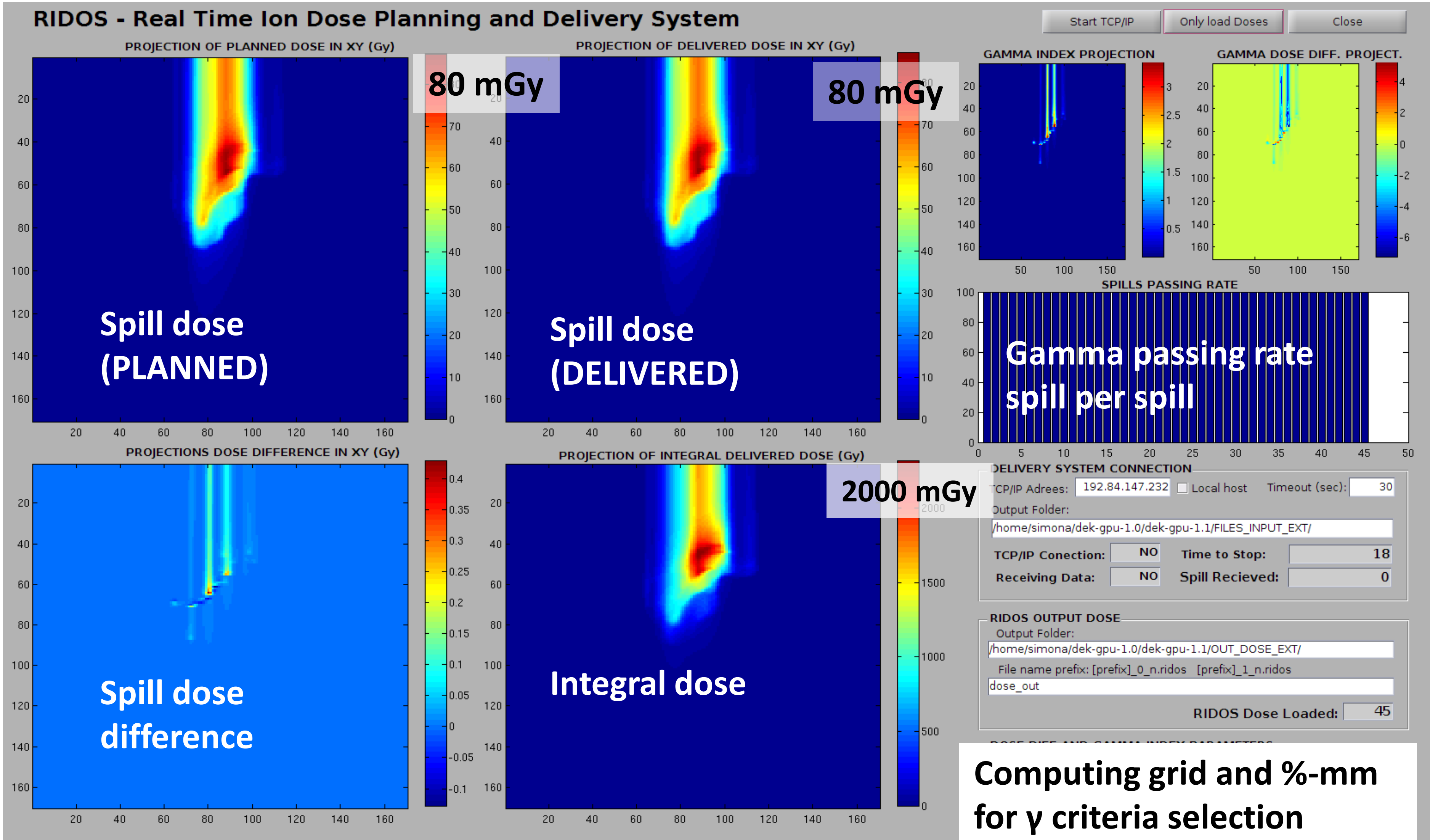
RIDOS HW and interfaces

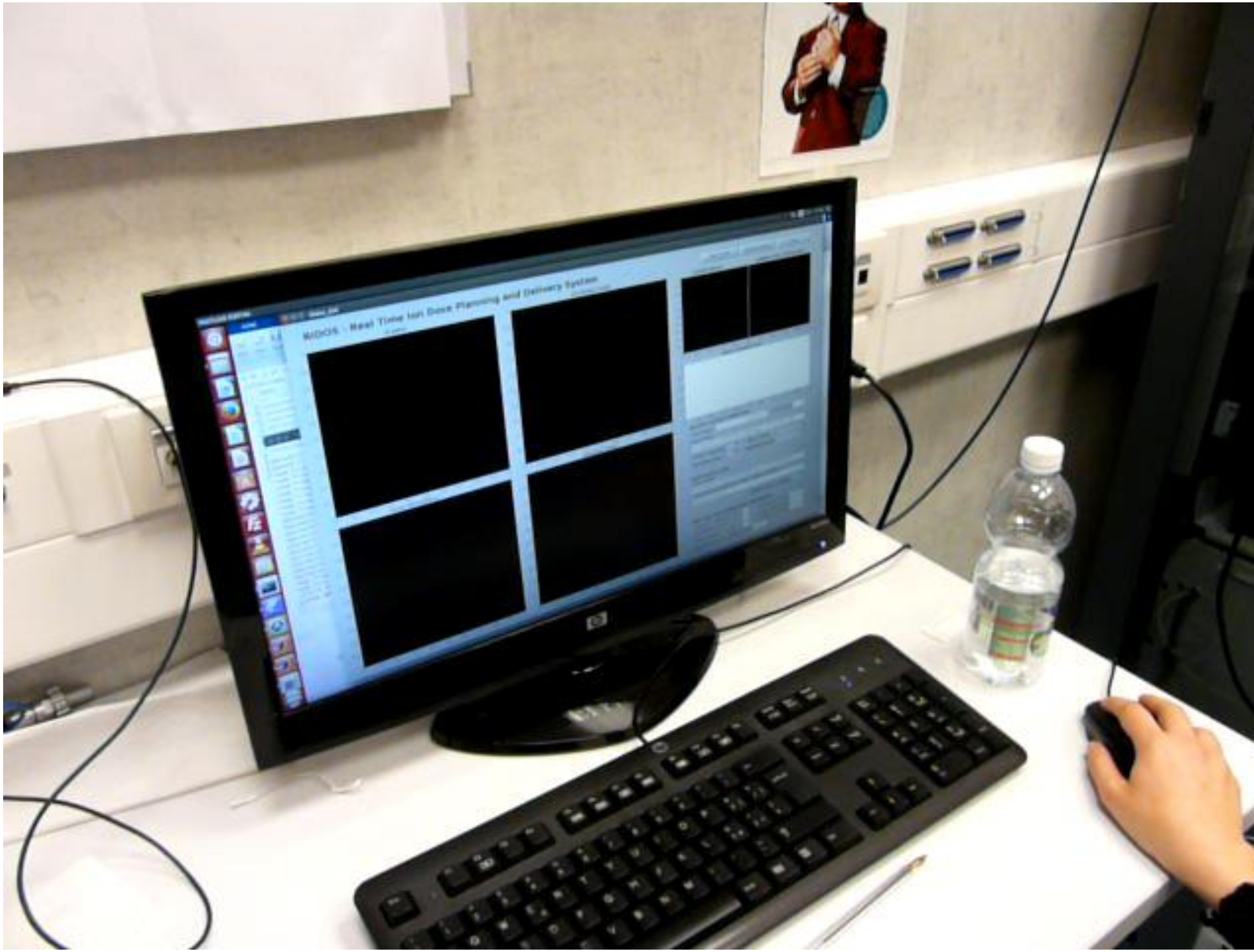


RIDOS HW and interfaces

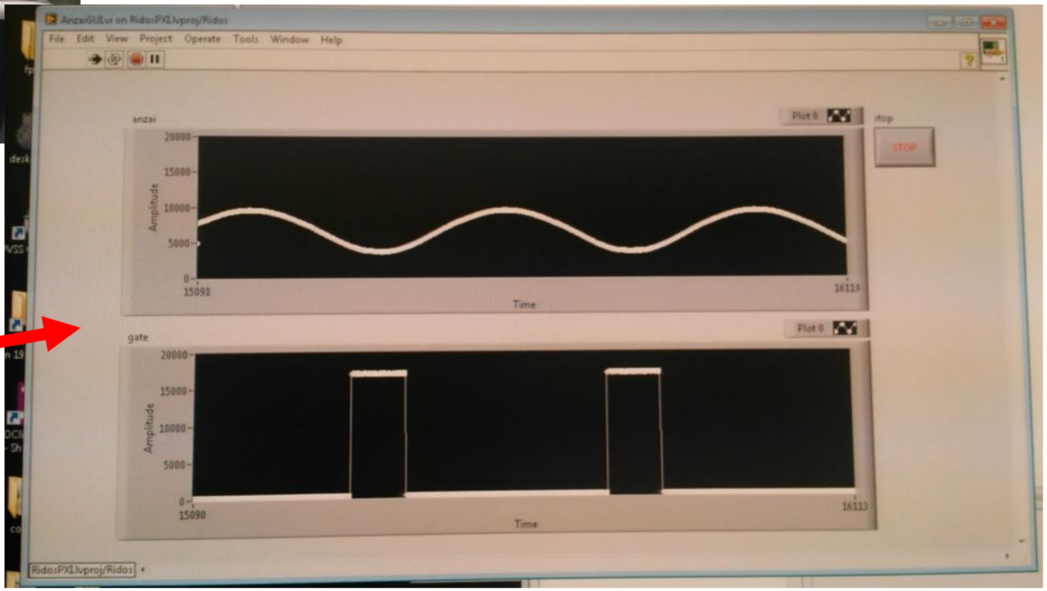


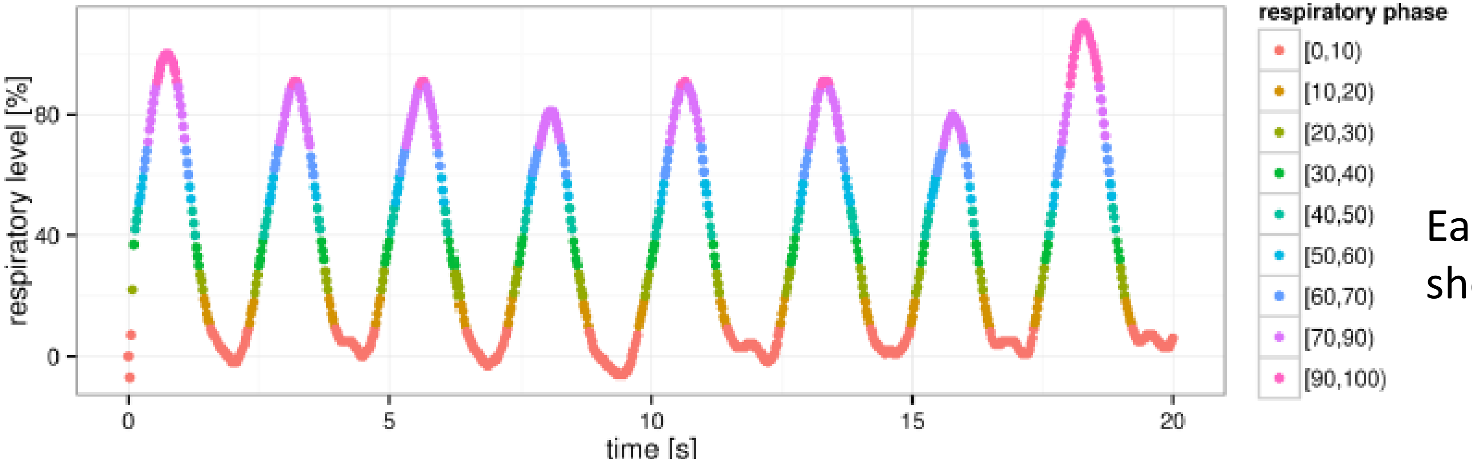
RIDOS GUI in Matlab




RIDOS - Real Time Ion Dose Planning and Delivery System

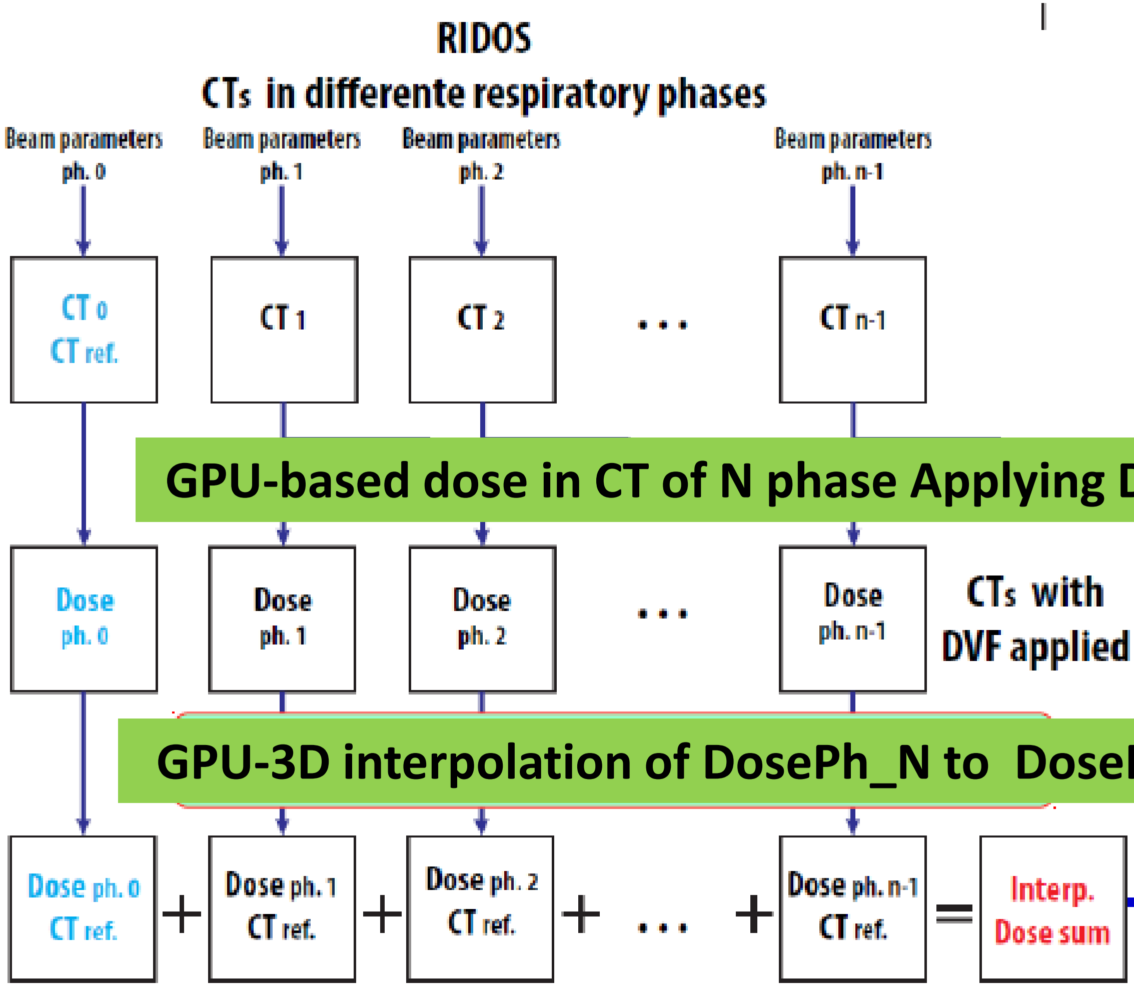
RIDOS "4D"





Each spot colour shows one phase

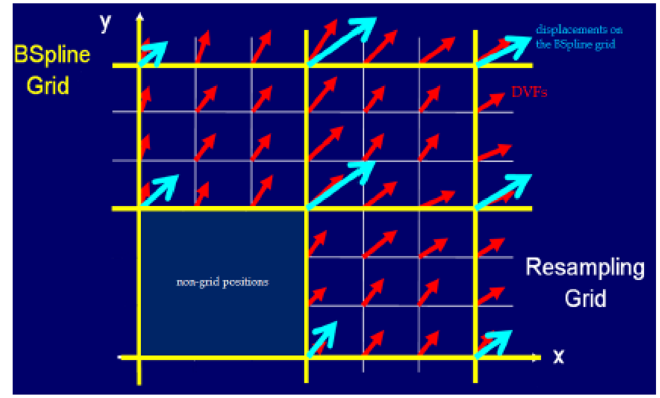
Dose accumulation with Deformable Image Registration (DIR)



GPU-based dose in CT of N phase Applying DVF

GPU-3D interpolation of DosePh_N to DoseRef

DVF = Deformation Vector Field



Coordinates translation Voxel by voxel on GPU

Map back to the reference CT

Doses accumulated spill by spill in different respiratory phases

Preliminary results of a GPU-based Dose accumulation in different respiratory phase

NVIDIA Quadro K2000
CUDA Cores: 384
Memoria Globale: 2 GB

NVIDIA Tesla K20
Cuda Cores: 2496
Memoria Globale: 5 GB

Tempi interpolazione	CPU	GPU K2000	GPU K20	Gain
DVF interp.	0,15 s	0,011 s	0,006 s	~14 / ~25
Dose interp.	5,1 s	0,18 s	0.08 s	~30 / ~63

Future and perspectives: Adaptive radiotherapy with PBS (?)

Adaptive radiotherapy consists in performing dose feedbacks

INTER-fraction dose feedback will be soon available for proton beams with commercial TPS and also with Fast MC

INTRA-fraction dose feedback is possible …if you have access to DDS HW and FIRMWARE and you have a sub-seconds dose calculation

BUT: *on-line patient imaging* is mandatory for significant clinical improvement

Conclusions

GPUs allow online accurate PB dose calculation

A FAST GPU-based Dose Engine Kernel for different ions has been integrated in the CNAO DDS and is under test.

The online doses can be calculated on different CTs exploiting the use of 4D-CT to improve treatments of moving targets.

Vincenzo Monaco
Anna Vignati
Andrea Attili
Felix Mas Milian
Marco Donetti
Mario Ciocca
Silvia Molinelli
Roberto Cirio
Flavio Marchetto
Germano Russo
Roberto Sacchi
Lorenzo Manganaro
Giorgio Magro
Mohammad Varasteh Anvar

REFERENCES

Jia X, Schumann J, Paganetti H and Jiang S B 2012 GPU-based fast Monte Carlo dose calculation for proton therapy *Phys. Med. Biol.* **57** 7783–97

Yepes P, Randeniya S, Taddei P J and Newhauser W D 2009 Monte Carlo fast dose calculator for proton radiotherapy: application to a voxelized geometry representing a patient with prostate cancer *Phys. Med. Biol.* **54** N21–8

Wan Chan Tseung H, Ma J and Beltran C 2015 A fast GPU-based Monte Carlo simulation of proton transport with detailed modeling of nonelastic interactions *Med. Phys.* **42** 2967

Ma J, Beltran C, Seum Wan Chan Tseung H and Herman M G 2014 A GPU-accelerated and Monte Carlo-based intensity modulated proton therapy optimization system *Med. Phys.* **41** 121707