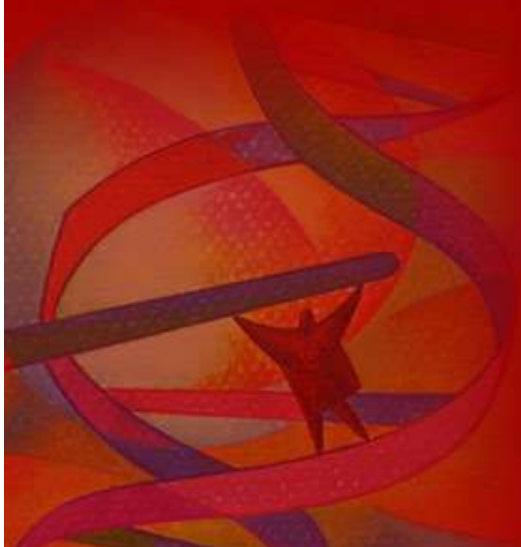




# IGS REPORT

Lauren Hyde  
Director and Lead Geneticist  
September 16, 2015

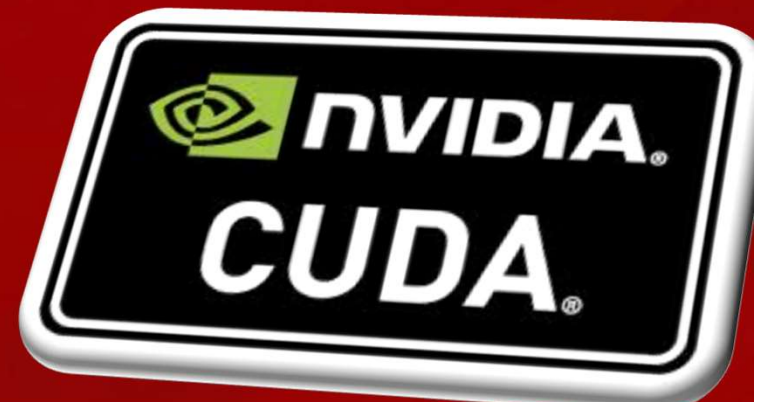
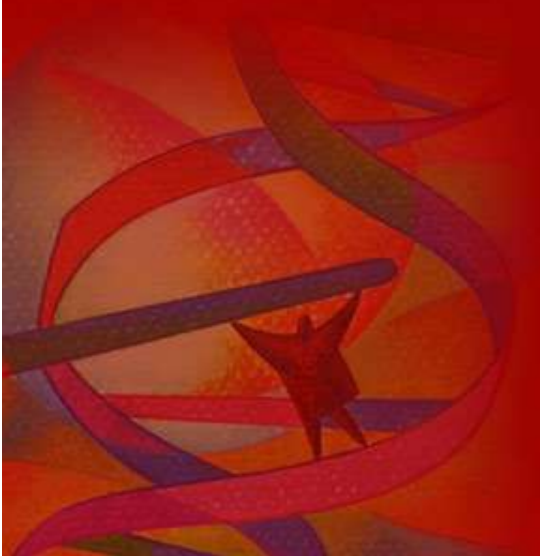


# Current IGS Database

12 breed associations

16,000,000 animals

340,000 new animals per year



# Current Software

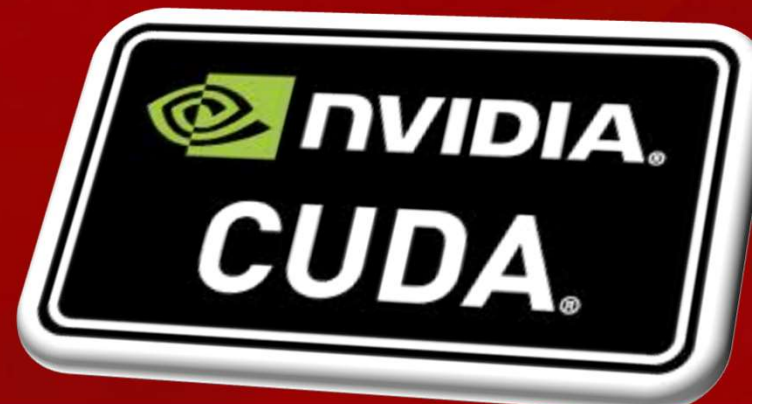
Initially developed in 1980s using Fortran 77

Implemented multi-breed international cattle evaluation (MB-ICE) in 1997

Last major modification in 2012 with multi-breed calving ease EPDs

Continually being updated using Fortran 95

Very limited



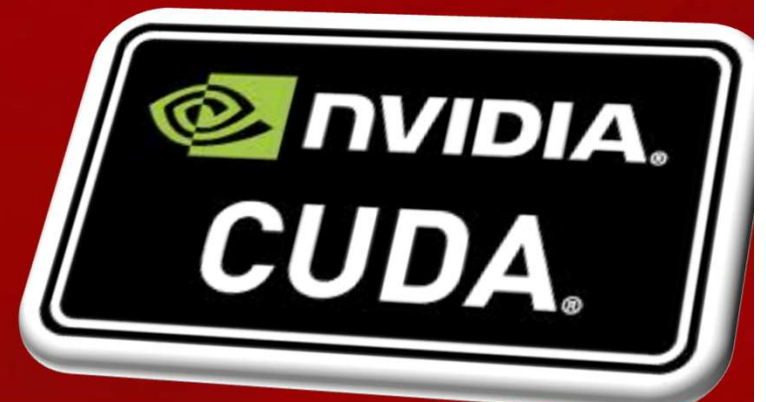
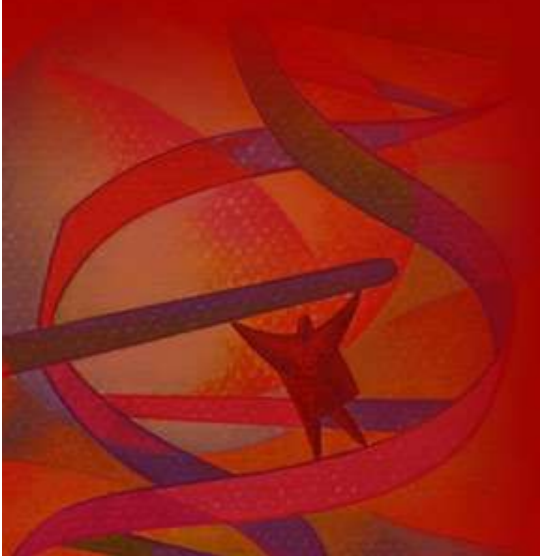
# Current Needs

Direct incorporation of genomics into EPDs

Improved accuracy computation

Ability to handle quickly growing large database

Efficiency and speed

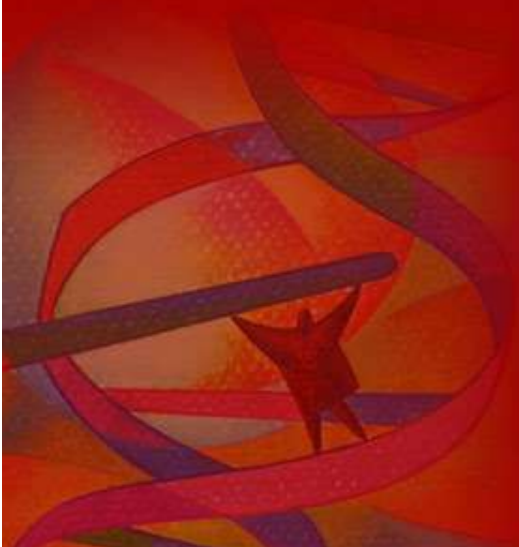




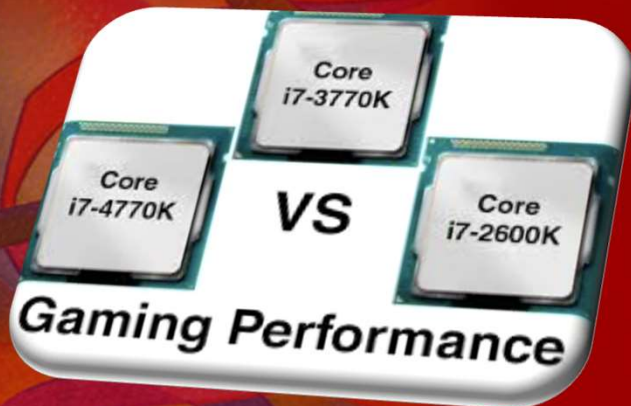
# BOLT

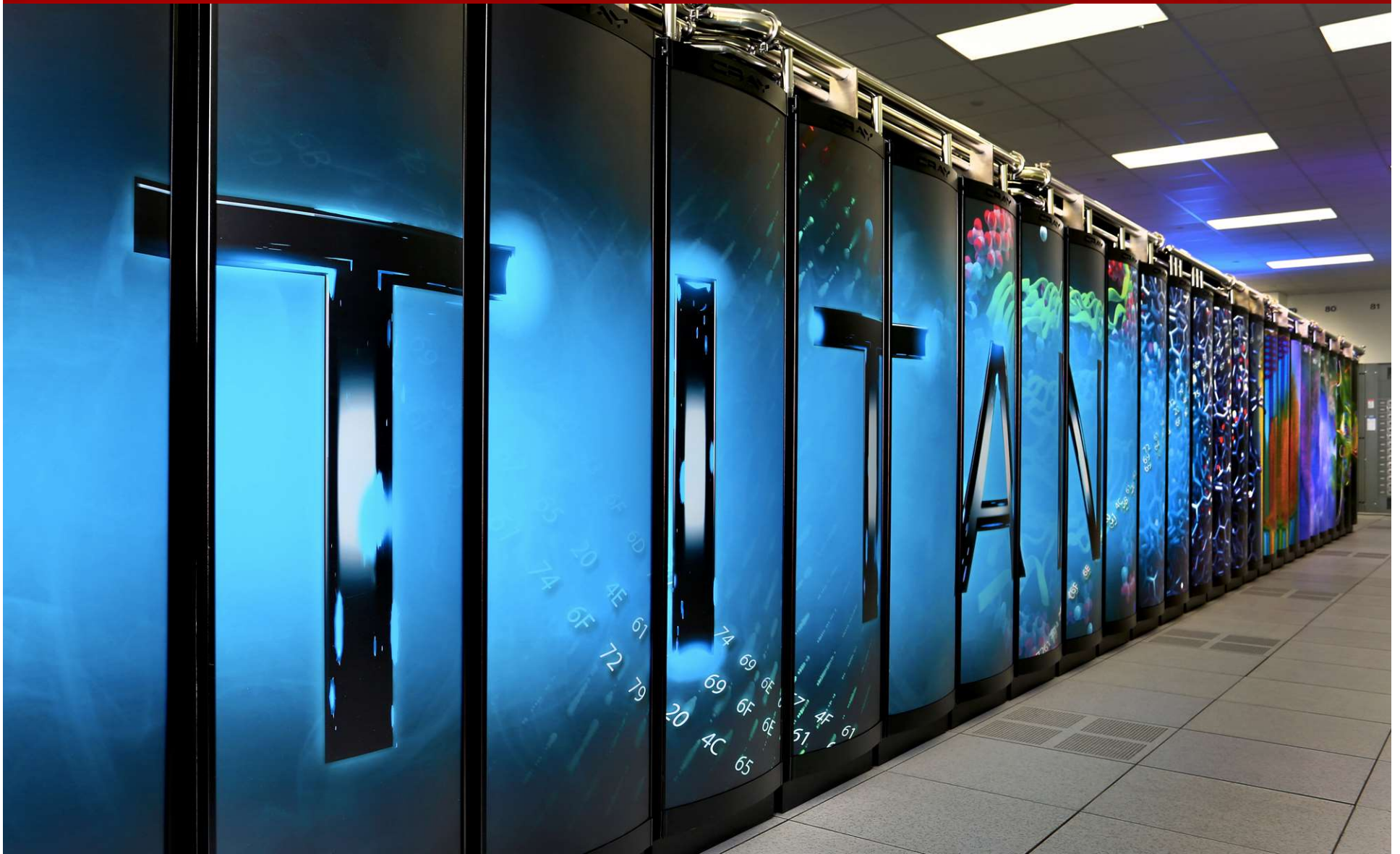
Biometry Open Language Tools

The next generation in  
genomic evaluation



# Leverage technology built for computer gaming





Intel Thermal  
Solutions  
Liquid Cooler

i7-4930K  
3.4Ghz

A5105 Sabertooth MB

2600MHz Rampage

GEFORCE GTX

Titan  
Tesla K20c  
GTX 760

Antec 1200  
HIGH CURRENT PRO  
CONTINUOUS POWER





# BOLT Major Features

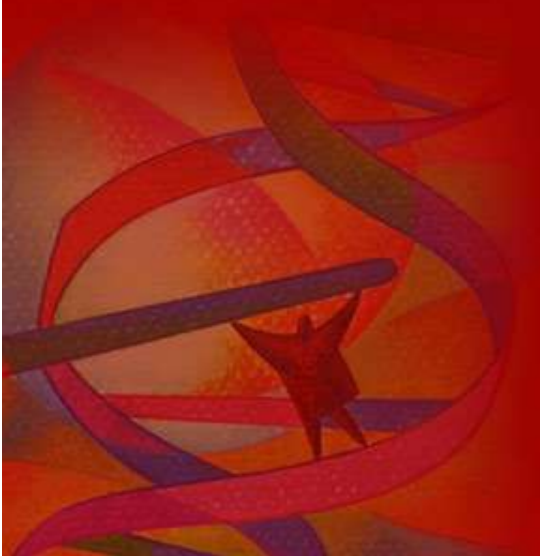
Fine grid to “embarrassingly” parallel capability

Multi-GPU (CUDA)

Multi-core

Multi-threaded asynchronous execution

Shared memory for performance



# BOLT Major Features

Novel algorithms

Written in C

Integrated with Unix user utilities

Extremely flexible

Complete set of extensible tools/commands

Simple user API



# Biometry Open Language Toolkit



corr	fblockinv	mmultongpu	mtmgpu	csolve
fsolve	insbmtx	mnvar	absorbx	csolves
fsolves	invert	mprin	mmultgpu	tsolve
absorbxm	csub	genomult	invnrm	pcgmgpu
astarsetup	csubm	gpustat	pedrecode	sthmgibbs
cadd	cudacheck	grpcnt	lambayes	permsub
cgen_z	cularnd	grpmn	libbolt	rank
chcat	cvcat	grps	libboltcuda	sdate
cholesky	cvcatcsr	grps2	load2csc	shmgr
cln	diag	ident	minmax	sp2mm
cmult	include	impute	stack_ped	ssgibbs
cnewr	fbcsrmv	imputegpu	mmult	transM
cnewr2	fblockinv			

# BOLT



Emphasizes sampler-based strategies to:

- Incorporate genomics
- Maximize accuracy

$$P(\theta | \mathbf{X})$$

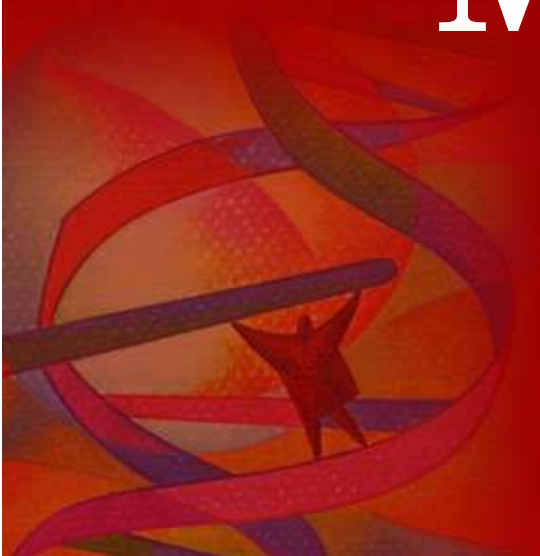
# Mixed Model Equations

$$\begin{bmatrix} X'X & X'Z_d & X'Z_m & X'Z_{pe} \\ Z_d'X & Z_d'Z_d + A^{-1}\alpha_1 & Z_d'Z_m + A^{-1}\alpha_2 & Z_d'Z_{pe} \\ Z_m'X & Z_m'Z_d + A^{-1}\alpha_2 & Z_m'Z_m + A^{-1}\alpha_3 & Z_m'Z_{pe} \\ Z_{pe}'X & Z_{pe}'Z_d & Z_{pe}'Z_m & Z_{pe}'Z_{pe} + I\alpha_4 \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u}_d \\ \hat{u}_m \\ \hat{u}_{pe} \end{bmatrix} = \begin{bmatrix} X'y \\ Z_d'y \\ Z_m'y \\ Z_{pe}'y \end{bmatrix}$$

# Genomic Information



$$\mathbf{M} = \begin{matrix} \text{Animals} \rightarrow & \begin{matrix} \text{Markers} \rightarrow \\ \mathcal{S}_{11} & \cdots & \mathcal{S}_{1m} \\ \vdots & \cdot & \vdots \\ \vdots & \cdot & \vdots \\ \mathcal{S}_{n1} & \cdots & \mathcal{S}_{nm} \end{matrix} \end{matrix}$$



# Example of Performance Computing

M'M



6,625,000,000,000,000  
(6.6 quadrillion) flop

Compute Time: 48m  
(2.29 Tflops)

# BOLT Implementation

Scheduled for early 2016

Data issues (international IDs, breed compositions)

Spring bull sales

Development and testing

Growth traits—framework built, testing well underway

Carcass traits—initial framework built, start testing

Threshold traits—start after growth complete





