Sven Reiche :: SwissFEL :: Paul Scherrer Institute

HPC for FELs

STFC-IOP joint workshop, London, April 2017
Overview

• Codes and Interface

• Physics Problems to Solve

• Numerical Problems to Solve

• Conclusion
General Work Flow

1. Define Beam Specification at Electron Source
2. Define Beam Specification at Undulator
3. Estimate Undulator Parameters
4. Design Undulator Module & Lattice
5. Run “Ideal” Simulation
6. Run Start-end Simulation
7. Optimize Layout
8. Define Basic Lattice and Compression
9. Estimate Undulator Parameters
10. Define Wavelength Range

Domain of HPC Needs

Ming Xie Model

Madx, Lie-Track
• Electron bunch charge typically limited at 1 nC
  – $6.25 \times 10^9$ electrons
  – Each electrons has 6 variables to describe its state $(x, y, z, p_x, p_y, p_z)$
  – Each state requires a double precision (8 Byte) variable to store it

  Total required memory 380 Gbyte

• Radiation field at 1 Angstrom for a 20 micron bunch
  – 200000 longitudinal sample points
  – 2D grid with 100 grid points each = 10000 transverse modes to model radiation wave front
  – Each sample of wave front requires 2 double precision values (amplitude and phase)

  Total required memory 30 GByte

System of 500 Gbyte distributed RAM needed, fulfilled with even mid-size computer clusters
## Small Survey of Existing Codes

<table>
<thead>
<tr>
<th>Gun</th>
<th>Linac</th>
<th>BC</th>
<th>FEL</th>
<th>Type</th>
<th>Parallel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homdyn</td>
<td>X</td>
<td></td>
<td></td>
<td>Envelope</td>
<td></td>
</tr>
<tr>
<td>Madx</td>
<td></td>
<td>X</td>
<td>X</td>
<td>Envelope</td>
<td></td>
</tr>
<tr>
<td>LiTrack</td>
<td></td>
<td>X</td>
<td>X</td>
<td>Particle</td>
<td></td>
</tr>
<tr>
<td>Astra</td>
<td>X</td>
<td></td>
<td>X</td>
<td>Particle</td>
<td>Yes</td>
</tr>
<tr>
<td>GPT</td>
<td>X</td>
<td>X</td>
<td>(X)</td>
<td>Particle</td>
<td>Yes</td>
</tr>
<tr>
<td>ImpactT</td>
<td>X</td>
<td></td>
<td>X</td>
<td>Particle</td>
<td>Yes</td>
</tr>
<tr>
<td>ImpactZ</td>
<td></td>
<td>X</td>
<td>X</td>
<td>Particle</td>
<td>Yes</td>
</tr>
<tr>
<td>Elegant</td>
<td></td>
<td>X</td>
<td>X</td>
<td>Particle</td>
<td>Yes</td>
</tr>
<tr>
<td>CSRTrack</td>
<td></td>
<td></td>
<td>X</td>
<td>Particle</td>
<td>Yes</td>
</tr>
<tr>
<td>Opal</td>
<td>X</td>
<td>X</td>
<td>(X)</td>
<td>Particle</td>
<td>Yes</td>
</tr>
<tr>
<td>Puffin</td>
<td></td>
<td></td>
<td>X</td>
<td>Particle</td>
<td>Yes</td>
</tr>
<tr>
<td>Ginger</td>
<td></td>
<td>X</td>
<td></td>
<td>Particle</td>
<td>Yes</td>
</tr>
<tr>
<td>Genesis</td>
<td></td>
<td></td>
<td>X</td>
<td>Particle</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Expert Codes vs All-Inclusive Code

**All-purpose Code**
- Consistent description of beam line
- Internal format of particle representation and radiation field
- Specific problems are solved with a simpler/reduced model (e.g. 1D Green’s function of CSR)
- Algorithm not fully optimized
  - Unneeded transport of full ensemble over through simple elements (e.g. drifts)
  - Inefficient Solver (e.g. particle-particle space charge)

**Specialized Codes**
- Algorithm are fully optimized, dedicated to specific problems
- Inconsistent description of the same beam line
- Conversion between particle formats needed
- Requires up- and down-sampling of particle distribution
- Semi-automated execution.
• To ease the use of codes and open them up for more users, there is a recent trend to develop a common framework to run simulations.

• Some Examples:
  – OCELOT (DESY): Developed for easier exchange of particle distribution.
  – Virtual Accelerator (SwissFEL): Developed to have unified layout description.
Physics Problems – Space Charge in Guns

- Strongly space charge dominated
- Relativistic velocity spread in the first few centimeter
- Good knowledge of RF and magnet field maps.

[S. Bettoni et al, PR-STAB, 18 (2015) 123403]
Physics Problems – Microbunch Instability, Coherent Synchrotron Radiation and Wakefields

- Collective effect due to interaction by static and radiation field
- CSR and wakefields have transients and can be shielded.
- Problems does not scale linearly with number of macro particles
- Requires binning and filtering for numerical stability

Convergence study with IMPACT

Wakefield for LCLS II cavity

- Calculation with Dedicated Codes such as CST
- Needs ultra-fine grid to obtain single particle short range wake to use in tracking programs

[Courtesy by LBL] [Courtesy by FNAL]
- Example: SwissFEL Soft X-ray FEL
  - Distributed chicanes in undulator break section
  - Controllable transverse gradient of undulator field with APPLE-X configuration
  - Injection of tilted beam

**Basic Modes + Enhancement**

- **SASE**
- Optical Klystron
- Harmonic Lasing

**Special Modes**

- **TW Pulse***
- **Two-Color***

**Spectral Control**

- High Brightness SASE
- Large Bandwidth Mode***
- Self-Seeding

**External Synchronization**

- Mode-locked Lasing
- Slicing
- HHG Seeding

**Legend:**

- **APPLE-X Configuration**
- **Chicanes**
- **Baseline**
- **Not Baseline**
- **Tilt***
• Shot noise arises from the fluctuation in a otherwise smooth distribution if a finite number of macro particles representing the electron bunch with a much higher number of electrons.

• In frequency space the fluctuation is represented by the bunch parameter

\[ b(k) = \frac{1}{N} \sum_{j \in V} e^{ikz_j} \]

• If the positions of the electrons are uncorrelated, the bunching factor is equivalent to a 2D random walk problem

• Statistic (evaluating bunching factor for many independent shots):

\[ \langle |b| \rangle = \frac{1}{\sqrt{N_e}} \quad \langle b \rangle = 0 \]

Shot noise is not correct if number of macro particles differs from number of electrons to be simulated
• SASE FEL are starting from shot noise and thus needs a correct modelling.

• In a narrow frequency band, the correct statistic can be applied with fewer macro particles, using a quite start methods
  1. Use Halton (bit reverse) sequences and transform them to fill the transverse distribution and distribution in energy.
  2. Generate a random phase for longitudinal position, mirror that particle N times and distribute them equally within the wavelength under consideration → Quiet Start
  3. Add a random noise to the position to obtain correct statistic

[W. Fawley, PRSTAB 5 (2002) 070701]

Method can be generalized for any wavelength above a certain cut-off frequency by varying also charge and beam energy.

• However, latest configurations of FELs are breaking any reasonable application of the quiet loading methods.
• Example: EEHG

1. Quiet Loading
2. Energy Modulation + $R_{56}$
3. Repopulation From other slices
4. Harmonic Conversion e.g. 4th harmonic

Simple Solution: Resolve each individual electron in bunch (1-1 FEL simulations)
• In Linac: Combination of longitudinal space charge and compression can enhance certain frequency components in the current profile (microbunch gain curve)

Real Bunch:
• $10^9$ particles
• $3.5 \times 10^{-3}$ % initial fluctuation
• $3.5$ % final current fluctuation

Model
• $10^6$ particles
• $0.1$ % initial fluctuation
• $100$ % final current fluctuation

Requires at least $10^6$ particles to calculate numerical microbunch gain curve but full number for correct phase space at FEL entrance

*Effect is important for low energy FELs because it defines the minimum reachable energy spread and thus the required beam energy to make the FEL work.*

---

**Gain (LSC+CSR)**

- Analytical model
- Numerical simulations (Elegant)
Numerical Problems – Shot Noise (μBunching)

Simulation of Gain Curve (SwissFEL)

Measurements (LCLS)

[D. Rattner et al, PR-STAB, 18 (2015) 030704]

[Courtesy: S. Bettoni]
Numerical Problems – Compression Setup

- Non-linear forces (e.g. wakefields, CSR) the compression difficult to set-up

- Example: Maximizing the linear chirp for SwissFEL large bandwidth mode.

![Flowchart showing the process of compression setup]

[A. Saa-Hernandez et al, PRAB 19 (2016) 090702]
Example 1 – Injector Optimization

- Genetic Optimizer
- Code: Opal
- 1000 Generation
- 40 Individual selected

- Optimization:
  - Minimum core emittance
  - Minimum core energy spread

[Y. Ineichen, Ph.D Thesis, ETH-7041-02]

Choice for figure-of-merit requires some deep knowledge of beam dynamic

Better results with simple LMDIFF

[S. Bettoni et al, PR-STAB, 18 (2015) 123403]
Example 2 – FEL Performance

- Bayesian Optimizer
- Real Machine: LCLS
- Ocelot Framework

- Optimization:
  – Maximum FEL pulse energy

Choice for figure-of-merit requires some deep knowledge of beam dynamic

[M. McIntire et al, IPAC 2016, p. 2972-2975]
Example 3 – FEL Tapering

- Optimization:
  - About 20 undulator gaps
  - About 20 phase shifter
  - About 20 quadrupoles

Parameter Set too large for general optimization algorithm

Requires an self-optimization strategy:
- Kroll-Morton-Rosenbluth (constant trapping)
- Modified KRM (changing trapping phase)
- Simple undulator profile

[A. Mak et al, IPAC 2014, p. 2909-2911]
HPC is an essential tool to design a FEL facility (and even more to benchmark the measured performance)

Various codes exist, which are dedicated to specific problems, running in a parallel computer configuration

Self-interaction requires a high number of particles to use to achieve convergence

Resolving each electron is feasible, which simplifies many problems, both physically and numerically

Optimizing the system requires a lot of computational resources and is often better done on the machine.
Laser and particle driven plasma wakefield acceleration simulations using the modern particle-in-cell codes: EPOCH, SMILEI and OSIRIS

Planning your simulation campaign

Dr James Holloway
The University of Oxford, Oxfordshire, UK. OX1 3PU
Talk Outline

• Role of simulations in novel acceleration experiments
• Simulations in aid of Plasma Wakefield Acceleration (PWA)
  • Designing your simulation campaign
• Codes: The Particle in Cell (PIC) method
• Introducing EPOCH, OSIRIS and SMILEI
• Physics captured by codes
• Balancing resolution with CPU hour cost
• Computational saves with LDWFA in EPOCH
Simulations supporting experiments – Astra 2015

Design phase
• Code was written to simulate the plasma wakefield amplitude driven in a plasma by a train of laser pulses
• Code was written to model the laser chain to simulate the laser trains that would be produced
• Result fed into design of diagnostics (spectrometer gratings etc)
• Simulated laser train fed into EPOCH PIC code

Corroborating experimental results
• Wakefields of $\Delta n_e / n_e = 1$ were measured experimentally
• EPOCH used to verify data
• Paper has been submitted for publication
Simulations supporting experiments – AWAKE

• Design phase
  • Muhammad Kasim proposed oblique-angle photon acceleration as a plasma wakefield diagnostic
  • Tight requirements on angle between probing laser pulse and the proton-driven plasma wakefield
  • Experimental design extremely tight at this point!
  • Angle of probing was constrained by length and width of plasma cell and surrounding equipment.
  • Muhammad simulated available probing angles and calculated the efficacy of the diagnostic.
  • Conversed with CERN staff to find a radiation cool spot to place camera

Quantitative single shot and spatially resolved plasma wakefield diagnostics

Courtesy of E. Feldbaumer and S. Cipiccia
Simulations supporting experiments – SLAC

Corroborating experimental results
• $E = 84$ GeV electrons generated via electron-driven wakefield acceleration
• $E_{\text{avg}} = 52$ GVm$^{-1}$
• They achieved energy gain of the 3-km-long SLAC accelerator in less than a metre. Although the luminosity was greatly reduced.

Stanford Linear Accelerator Center (SLAC).

<table>
<thead>
<tr>
<th>Slac Beam Parameters</th>
<th>Plasma Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>• $E = 42$ GeV</td>
<td>• $n_e = 2.7e23$ cm$^{-3}$</td>
</tr>
<tr>
<td>• $\sigma_z = 15$ um*</td>
<td>• $\lambda_p = 100$ um</td>
</tr>
<tr>
<td>• $\sigma_r = 10$ um</td>
<td>• Element = Lithium</td>
</tr>
<tr>
<td>• $Q = 3$ nC</td>
<td>• Cell = Oven</td>
</tr>
</tbody>
</table>

The Slac beam was compressed from $\sigma_z = 6$ mm to $\sigma_z = 15$ um using a $\sim 10$m magnetic chicane, making it suitable wakefield driver.

---

Describing a plasma: PP

- The complete description (Particle-Particle approach):
  - Specify x(t) and v(t) of every particle
  - Discretise time
  - Calculate force of each particle on every other within Debye sphere
  - Number of calculations scales as $N(N-1)/2$

- Number of calculations can be reduced to $O(N \log(N))$ using the Barnes-Hut tree algorithm [1]
- Works for strongly coupled systems (small $N$ per Debye sphere)

**Used In:**
- Molecular dynamics method for condensed matter and biomolecular studies [2]
- Gravitational interaction in cosmological studies and galaxy formation

Describing a plasma: PIC

- The Vlasov description (Particle-in-Cell)
- Specify $x(t)$ and $v(t)$ of every particle
- Discretise time
- Calculate force of each particle onto local grid points
- Number of calculations scales as $\sim N$
- Number of calculations can be reduced to $O(N \log(N))$ using the Barnes-Hut tree algorithm [BH86]
- Works for strongly and weakly coupled systems

Used in:
- Plasma wakefield acceleration!
  - For example: A plasma of $T_e = 1 \times 10^5$ K, $n_e = 1 \times 10^{22}$ m$^{-3}$ has $N\Lambda_d = 106$
  - <Other weakly coupled examples>

\[ \vec{F} = q\vec{v} \times \vec{B} \]

\[ \begin{align*}
\vec{E} &= \frac{\rho}{\varepsilon_0} \\
\vec{B} &= 0 \\
\n\n\n\n\end{align*} \]
“EPOCH is a plasma physics simulation code which uses the Particle in Cell (PIC) method. In this method, collections of physical particles are represented using a smaller number of pseudoparticles, and the fields generated by the motion of these pseudoparticles are calculated using a finite difference time domain technique on an underlying grid of fixed spatial resolution. The forces on the pseudoparticles due to the calculated fields are then used to update the pseudoparticle velocities, and these velocities are then used to update the pseudoparticle positions. This leads to a scheme which can reproduce the full range of classical micro-scale behaviour of a collection of charged particles....”

“Based on the highly nonlinear and kinetic processes that occur during high-intensity particle and laser beam-plasma interactions, we use PIC codes [1,2], which are a subset of the particle-mesh techniques, for the modeling of these physical problems. In these codes the full set of Maxwell’s equations are solved on a grid using currents and charge densities calculated by weighting discrete particles onto the grid. Each particle is pushed to a new position and momentum via self-consistently calculated fields....”

“To face the diverse needs of the teams involved in its development, Smilei is developed in C++, based on an object-oriented architecture. Its modularity allows to run simulations in various dimensions, geometries, to chose between various Maxwell solvers, particle pushers, interpolators, projectors, etc. Smilei works now in 3D cartesian geometry. Maxwell’s equations are solved on the so-called Yee-mesh with centered electric and magnetic fields using the finite-difference time-domain (FDTD) method or related methods [Nuter2014]....”

**University of Warwick**

EPOCH

http://www.ccpp.ac.uk/epoch/obtaining_epoch.pdf


http://epp.ist.utl.pt/wp/osiris/


http://www.maisondelasimulation.fr/smilei/

<table>
<thead>
<tr>
<th>Feature</th>
<th>EPOCH</th>
<th>OSIRIS</th>
<th>SMILEI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D, 2D, 3D Cartesian</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Cylindrical 2D</td>
<td>✘</td>
<td>✔</td>
<td>✘</td>
</tr>
<tr>
<td>Ionisation</td>
<td>✔</td>
<td>✘</td>
<td>✔</td>
</tr>
<tr>
<td>Binary collisions</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Collisional ionisation</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Lorentz boosted frame</td>
<td>✘</td>
<td>✔</td>
<td>✘</td>
</tr>
<tr>
<td>Laser envelope model</td>
<td>✘</td>
<td>✔</td>
<td>✘</td>
</tr>
<tr>
<td>Normalized units</td>
<td>✘</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>S.I. units</td>
<td>✔</td>
<td>✘</td>
<td>✘</td>
</tr>
<tr>
<td>QED module</td>
<td>✔</td>
<td>✔</td>
<td>✘</td>
</tr>
<tr>
<td>Dynamic load balancing</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Language</td>
<td>Fortran 90</td>
<td>Fortran 90</td>
<td>C++</td>
</tr>
<tr>
<td>Open source</td>
<td>✔</td>
<td>✘</td>
<td>✔</td>
</tr>
</tbody>
</table>
Plasma Wakefield Acceleration

Novel particle acceleration technique

Can also drive wakefields with
- Electrons
- Photons
- Positrons
- Muons (in principle)

OSIRIS simulation
Optimizing for CPU-hour cost

- CPU-hours have a cost:
  - ARC: £0.01 per cpu-hour 1st August 2016
- Aim to perform simulations that capture physics using minimal CPU-hours
- Reducing grid size (resolution) lowers cost. Too low can ‘switch off’ physics
- Reducing number of macro-particles lowers cost. Too low and the simulation has high noise. This can preferentially seed instabilities that grow from noise (hosing)
- Need to find goldilocks zone.
- Can find imperially by running test simulations

A basic electron-driven plasma wakefield simulation:

\[
E = 300 \text{ MeV}, \quad Q = 16 \text{ pC}, \quad \sigma_r = \sigma_z = 1/k_p
\]

- \(n_e = 1.12 \times 10^{17} \text{ cm}^{-3}\)
- \(\lambda_p = 100 \mu\text{m}\)
- Ion = Hydrogen
Optimizing for CPU-hour cost

Resolution parameter scan

- Amplitude of wakefield
- Chosen Resolution
- Stable wakefield amplitude

Number of grid points per plasma wavelength

Amplitude of wakefield (V/m·¹)

$10^8$
Optimizing for CPU-hour cost
Benchmark against theory and another code

Ideal driver length for Gaussian beam:

$$\sigma_{\text{ideal}} = \frac{\lambda_p}{2\pi}$$
Planning Simulation Campaign

Take this example mock campaign:

- Optimise for self modulation instability of a long electron beam
- Create stable micro-bunch train
- Drive a high amplitude wakefield via resonant excitation of the wakefield with micro-bunch train

<table>
<thead>
<tr>
<th>Simulations</th>
<th>PPC</th>
<th>Grid size</th>
<th>Simulation time (ps)</th>
<th>CPU-units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial 2D test x 10</td>
<td>8</td>
<td>300 x 60</td>
<td>200 ps</td>
<td>1</td>
</tr>
<tr>
<td>Long run test</td>
<td>8</td>
<td>300 x 60</td>
<td>2000 ps</td>
<td>10</td>
</tr>
<tr>
<td>Density scan x 20</td>
<td>8</td>
<td>300 x 60</td>
<td>2000 ps</td>
<td>200</td>
</tr>
<tr>
<td>Beam width x 10</td>
<td>8</td>
<td>300 x 60</td>
<td>2000 ps</td>
<td>100</td>
</tr>
<tr>
<td>Full final 3D</td>
<td>8</td>
<td>300 x 60 x 60</td>
<td>2000 ps</td>
<td>600</td>
</tr>
</tbody>
</table>

Total CPU-time of simulation campaign = 911 CPU-units

Note:
- Do not have to resolve Debye length as we are in moving window. No time for the plasma to numerically heat
- Smallest scale to resolve then becomes plasma wavelength
Aiming to achieve stable micro-bunch train
Hosing instability disrupts self-modulation
Increasing forward direction resolution

30 grid points per plasma wavelength

120 grid points per plasma wavelength

480 grid points per plasma wavelength
Increasing particles per cell

8 PPC

80 PPC

800 PPC
Planning Simulation Campaign

Take this example mock campaign:

- Optimise for self modulation instability of a long electron beam
- Create stable micro-bunch train
- Drive a high amplitude wakefield via resonant excitation of the wakefield with micro-bunch train

<table>
<thead>
<tr>
<th>Simulations</th>
<th>PPC</th>
<th>Grid size</th>
<th>Simulation time (ps)</th>
<th>CPU-units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial 2D test x 10</td>
<td>800+</td>
<td>300 x 60</td>
<td>200 ps</td>
<td>1 x 100</td>
</tr>
<tr>
<td>Long run test</td>
<td>800+</td>
<td>300 x 60</td>
<td>2000 ps</td>
<td>10 x 100</td>
</tr>
<tr>
<td>Density scan x 20</td>
<td>800+</td>
<td>300 x 60</td>
<td>2000 ps</td>
<td>200 x 100</td>
</tr>
<tr>
<td>Beam width x 10</td>
<td>800+</td>
<td>300 x 60</td>
<td>2000 ps</td>
<td>100 x 100</td>
</tr>
<tr>
<td>Full final 3D</td>
<td>800+</td>
<td>300 x 60 x 60</td>
<td>2000 ps</td>
<td>600 x 100</td>
</tr>
</tbody>
</table>

Total CPU-time of simulation campaign = 911 CPU-units

Total CPU-time of simulation campaign = 91,100 CPU-units

Note:

- Do not have to resolve Debye length as we are in moving window. No time for the plasma to numerically heat
- Smallest scale to resolve then becomes plasma wavelength
- **HOSING INSTABILITY GROWS FROM NOISE - SEEDED BY LOW PPC!**

“No Battle Plan Survives Contact With the Enemy”
Examples of simulation campaigns – Laser train

• A simulation campaign was recently undertaken to study the effect of laser pulse spacing in resonantly driving a wakefield
• The campaign has fed into experimental design (compressor detuning, target design etc) and the experimental proposal
• The nature of the experimental proposal was such that the simulations were complex and would take significant time to commission.
Examples of simulation campaigns – Laser train
Examples of simulation campaigns – Laser train

Commissioning:
- Bigbox
- Lambda0
- Picosecond
- Resolution
- LowHigh
- Sigmaz+
- Bandwidth
- Exponential
- Linear

Commissioning:
- PlasmaWide
- PlasmaImmobile
- PlasmaMobile
- SpikeError
- PlasmaWide
- PlasmaFinal

Production run:
- Scan1 – 40 sims
- Scan2 – 40 sims
- Scan3 – 40 sims
- Scan4 – 40 sims

Production run:
- Scan5 – 10 sims
- Scan6 – 10 sims
- Scan7 – 10 sims
- Scan8 – 10 sims

Production run:
- Scan9 – Scan20
Examples of simulation campaigns – Laser train

Commissioning:
- Bigbox
- Lambda0
- Picosecond
- Resolution
- LowHigh
- Sigmaz+
- Bandwidth
- Exponential
- Linear

Production run:
- Scan1 – 40 sims
- Scan2 – 40 sims
- Scan3 – 40 sims
- Scan4 – 40 sims

Commissioning:
- PlasmaWide
- PlasmaImmobile
- PlasmaMobile
- SpikeError

Production run:
- Scan5 – 10 sims
- Scan6 – 10 sims
- Scan7 – 10 sims
- Scan8 – 10 sims

Production run:
- Scan9 – Scan20

Production run:
- Scan1 – 40 sims
- Scan2 – 40 sims
- Scan3 – 40 sims
- Scan4 – 40 sims

Examples of simulation campaigns – Laser train

Commissioning:
- Padding
- FromFile
- Interpolate
- JustBeam

Commissioning:
- PlasmaWide
- PlasmalImmobile
- PlasmaMobile
- SpikeError

Production run:
- Scan5 – 10 sims
- Scan6 – 10 sims
- Scan7 – 10 sims
- Scan8 – 10 sims

Production run:
- Scan5 – 10 sims
- Scan6 – 10 sims
- Scan7 – 10 sims
- Scan8 – 10 sims

Examples of simulation campaigns – Laser train

Commissioning:
- Bigbox
- Lambda0
- Picosecond
- Resolution
- LowHigh
- Sigmaz+
- Bandwidth
- Exponential
- Linear

Production run:
- Scan1 – 40 sims
- Scan2 – 40 sims
- Scan3 – 40 sims
- Scan4 – 40 sims

Examples of simulation campaigns – Laser train

Commissioning:
- Bigbox
- Lambda0
- Picosecond
- Resolution
- LowHigh
- Sigmaz+
- Bandwidth
- Exponential
- Linear

Production run:
- Scan1 – 40 sims
- Scan2 – 40 sims
- Scan3 – 40 sims
- Scan4 – 40 sims

Commissioning:
- PlasmaWide
- PlasmaImmobile
- PlasmaMobile
- SpikeError

Production run:
- Scan5 – 10 sims
- Scan6 – 10 sims
- Scan7 – 10 sims
- Scan8 – 10 sims

Production run:
- Scan9 – Scan20

Examples of simulation campaigns – Laser train

Commissioning:
- Bigbox
- Lambda0
- Picosecond
- Resolution
- LowHigh
- Sigmaz+
- Bandwidth
- Exponential
- Linear

Production run:
- Scan1 – 40 sims
- Scan2 – 40 sims
- Scan3 – 40 sims
- Scan4 – 40 sims

Commissioning:
- PlasmaWide
- PlasmaImmobile
- PlasmaMobile
- SpikeError

Production run:
- Scan5 – 10 sims
- Scan6 – 10 sims
- Scan7 – 10 sims
- Scan8 – 10 sims

Production run:
- Scan9 – Scan20
Examples of simulation campaigns – Laser train

- Commissioning:
  - Bigbox
  - Lambda0
  - Picosecond
  - Resolution
  - LowHigh
  - Sigma +/
  - Bandwidth
  - Exponential
  - Linear

- Commissioning:
  - Padding
  - FromFile
  - Interpolate
  - JustBeam

- Commissioning:
  - PlasmaWide
  - PlasmaImmobile
  - PlasmaMobile
  - SpikeError
  - PlasmaWide
  - PlasmaFinal

- Production run:
  - Scan1 – 40 sims
  - Scan2 – 40 sims
  - Scan3 – 40 sims
  - Scan4 – 40 sims

- Production run:
  - Scan5 – 10 sims
  - Scan6 – 10 sims
  - Scan7 – 10 sims
  - Scan8 – 10 sims

- Production run:
  - Scan9 – Scan20
Examples of simulation campaigns – Laser train

Commissioning:
- Bigbox
- Lambda0
- Picosecond
- Resolution
- LowHigh
- Sigmaz+
- Bandwidth
- Exponential
- Linear

Commissioning:
- Padding
- FromFile
- Interpolate
- JustBeam

Production run:
- Scan1 – 40 sims
- Scan2 – 40 sims
- Scan3 – 40 sims
- Scan4 – 40 sims

Production run:
- Scan9 – Scan20

Commissioning:
- PlasmaWide
- PlasmalImmobile
- PlasmaMobile
- SpikeError
- PlasmaWide
- PlasmaFinal

Production run:
- Scan5 – 10 sims
- Scan6 – 10 sims
- Scan7 – 10 sims
- Scan8 – 10 sims
Examples of simulation campaigns – Laser train
Examples of simulation campaigns – Laser train

Production jobs
14,130 cpu-hours

Commissioning jobs
444 cpu-hours
Examples of simulation campaigns – Laser train

Commissioning jobs
444 cpu-hours

Production jobs
14,130 cpu-hours

Cost of ~£150
(at £0.01 per cpu-hour)
Note: less than the cost of ~8 staff days

Commisioning took 3% Of total CPU-time

Commissioning jobs
444 cpu-hours
Examples of simulation campaigns – Laser train

Commissioning took 3% of total CPU-time

Production jobs: 14,130 cpu-hours

Commissioning jobs: 444 cpu-hours
Examples of simulation campaigns – Laser train
Examples of simulation campaigns – Diamond beam

- Aim to demonstrate micro-bunching of 3 GeV Diamond beam
- Diamond beam is $\sigma_z = 27$ mm
- Laser wavelength is 0.0008 mm
- Discrepancy in lengths warrants huge grid to accommodate (1,200,000 x 120)
- Problem further compounded by need for long simulation time of nano-seconds
- Won $2 \times 10^6$ cpu-hours on the DiRAC super computer to perform simulation campaign.
- This was still not enough! Needed to find a way to reduce cpu-time required

---

1) The long beam given transverse momentum by strong wakefield.
2) Micro-bunches form as half e- defocused, other half focused
3) At the moment the beam has achieved best micro-bunching*, pass into the second plasma cell.
Examples of simulation campaigns – Diamond beam

- The first plasma stage has a laser driven wakefield
- High resolution needed to capture the physics of the laser pulse
- The rest of the simulation could be supported by a grid 100 times coarser
- Solution:
  1. Run the simulation at high resolution for the first plasma stage
  2. Bin down the grid in the restart timestep by x 100 (ScaleDownSDF.m available upon request)
  3. Restart the low resolution simulation from the binned timestep

1) The long beam given transverse momentum by strong wakefield.
2) Micro-bunches form as half e- defocused, other half focused
3) At the moment the beam has achieved best micro-bunching*, pass into the second plasma cell.
Future technology for PIC codes
Summary
smilei
ionisation has built in look up tables
Micro-Bunching Via Self Modulation

- Seed instability
- Wakefield modulates beam
- Beam drives harder wakefield
- Feedback mechanism
- Takes time for instability to saturate

This is not the scheme I use

Micro-bunched beam.
• Strong transverse kick from laser wakefield
• **Propagate beam through vacuum**
• Pass micro-bunches into second plasma stage when on-axis number density maximised
The ‘Drift Space’ Design

1) The long beam given transverse momentum by strong wakefield.
2) Micro-bunches form as half e- defocused, other half focused
3) At the moment the beam has achieved best micro-bunching*, pass into the second plasma cell.

Electron beam
- $E = 300$ MeV
- $\varepsilon_p = 0$
- $\sigma_z = 0.6$ cm
- $\sigma_r = \sqrt{2} / k_p$
- $Q = 0.1$ nC
- $E_r = 1$ GVm$^{-1}$

The Plasma
- $n_e = 1.11 \times 10^{22}$ m$^{-3}$
- $\lambda_p = 300$ um

• First plasma cell analogous to lens
• Longer first cell results in a stronger focus and shorter vacuum needed
• However strong focus results in higher emittance micro-bunches
Recent Developments in Scientific Visualisation: Bivariate, Multi-variate and Analytical Techniques
Scientific Visualization

Scalar Field
\( f : \mathbb{R}^d \rightarrow \mathbb{R}^l \)

Vector Field
\( f : \mathbb{R}^d \rightarrow \mathbb{R}^4 \)

Multi Field
\( f : \mathbb{R}^d \rightarrow \mathbb{R}^l \)

Wiebel et al., 2007

Stompel & al., PG 2002
Visualizing Scalar Fields

One slice at a time

Volume Rendering
(computed X-rays)

Isosurfaces
(extracted geometry)
Trends

• Simple visualisations swamp user bandwidth
• Complex visualisations require:
  • Interfaces for parameter choice
  • Analysis of data to guide choices
  • Techniques for non-scalar, non-vector data
• Scalability
• In situ analysis
Isosurfaces
Isosurface Statistics

Cell Intersection Statistics for Monkey MRI T1 Data Set

- Duffy 11 (Eq 30)

Carr, Duffy & Denby 2006
Scheidegger, Schreiner, Duffy, Carr & Silva, 2008
Duffy, Carr & Möller, 2011
Contour Spectrum

- Bajaj, Pascucci & Schikore, Vis 1997
Flexible Isosurfaces
Topological Landscapes

Fig. 9. Methane dataset: with average volume-to-area projection error of 1.5%. (Left) Topological landscape. (Right) Volume rendering using the same color scheme.

Weber et al., 2007
Multi-Fields

• Simplest approach:
  • Combine / compare scalars
  • Statistical
  • Sequential
  • Topological
  • Bivariate
  • Fully multi-variate (not yet)
Combustion Analysis

Bremer et al., Vis 2009
Vortex Detection

(a) $\lambda_2$ Contour Tree
(b) Similarity Browser
(c) Pressure Contour Tree
(d) Main Render View

Schneider et al., Vis 2008
Ensemble Analysis

- Regions where extrema occur in every run

Günther & al., 2014
Bivariate Visualisation

- Need to generalise:
  - Histograms
  - Isosurfaces
  - Interfaces
  - Topological Analysis
(Continuous) Scatterplots

Fig. 8. The left side shows the discrete scatterplot of the “blunt-fin” data set, whereas the continuous version is shown on the right side. Both types of scatterplots visualize the scalar data value along the horizontal axis and the magnitude of the gradient along the vertical axis. Choosing these data dimensions, material and boundary identification is possible by finding arc-like structures.
An Example: Ethanediol

• Data distribution in range (Bachthaler & Weiskopf 2008)

• Assumes $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$
• Level sets of a point in range

\[ f^{-1}(h) = \{ x \in \mathbb{R}^3 : f(x) = h : h \in \mathbb{R} \} \]

• Intersection of isosurfaces of two functions

\[ f^{-1}(h) = f_1^{-1}(h_1) \bigcap f_2^{-1}(h_2) : h = (h_1, h_2) \in \mathbb{R}^2 \]
Fiber Surfaces

- Inverse images of *lines, curves, polygons &c.*
- Each point on the line generates a fiber
- Their union gives a surface
• Mental model:
  • draw a loop around the desired data
  • It shows it immediately as geometry
Example: Thermal Plume
Open Surfaces
• Notice the overlapping features
• This is where the topology kicks in
• Quotient space under continuous contraction of the fiber components
Separable Sheets (Regions)
Turbulent Flow

- Regions of shared behaviour
- Peeling away shows small regions
- In this case, vortex street
Flow vs Curl
Scalability

- Continuous Scatterplots - excellent
- Fiber Surfaces - excellent
- Topological Analysis - Work in Progress
- Large Memory Footprint
- Stubbornly Serial
- At present: 10GB data is practical limit
Implementations

• Prior to 2016: roll your own
• 2017:
  • Cont. Scatterplots & Fiber Surfaces in vtk
  • Parallel Contour Trees through vtk-m
  • AND (April 2017)
    • Topology Toolkit (https://topology-tool-kit.github.io)
  • Usable through ParaView
RCUK Cloud Working Group: supporting the research community in the application of cloud computing technologies

Networkshop45
Nottingham, Wednesday 12th April 2017

Philip Kershaw
Technical Manager, Centre for Environmental Data Analysis, RAL Space, STFC; Chair, RCUK Cloud Working Group
Overview

- [What is Cloud?]
- Origins of WG
- November Workshop
- Legal, policy, regulatory issues
- Technical Integration
- Next steps
Cloud 101: need to understand in order to exploit

“Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction.” – NIST SP800-145

5 essential characteristics
- On-demand self-service
- Broad network access
- Resource pooling
- Rapid elasticity
- Measured service

3 service models
- IaaS (Infrastructure as a Service)
- PaaS (Platform as a Service)
- SaaS (Software as a Service)

4 deployment models
- Private cloud
- Community cloud
- Public cloud
- Hybrid cloud
RCUK Cloud Working Group Origins

- **Cloud Computing for Research and Innovation** - Report Aug 2015, with input from members of research community and contributions from industry

- Where are we with adoption of cloud?
  - identifies the major technical and policy issues that are seen to be preventing widespread take up of cloud services

- What needs to be done?
  - Four high level recommendations

- How do we get there – 5 year roadmap
  - to investigate these issues and provide closer integration of public and private sector resources to improve the capability of the UK research community
1. **Community building**
   - *Cloud Computing Working Group* (in place since 2015) to provide a clear community focus for cloud computing

2. **Technical integration**
   - of the UK National e-Infrastructure resources to promote workload mobility and to reduce technical barriers to entry.

3. **Training and Support**
   - Equip the research community with the right skills and support to fully exploit UK National e-Infrastructure cloud resources.

4. **Legal, Policy and Regulatory Issues**
   - Policy changes needed within RCUK to grow the adoption of cloud computing
   - Policy actions that RCUK can initiate externally on behalf of the UK cloud computing community.

RCUK Cloud Working Group

National e-Infrastructure Project Directors Group

Liaise with

Research Councils UK National e-Infrastructure Group

Reports to

European and wider international initiatives

Track and collaborate with

Research Community, cloud providers

Initiated by WG and SIG

Challenges and Ideas

Cloud Special Interest Group

Build relationships with

Task Force

Reports to

RCUK Cloud Working Group

Initiated by WG and SIG

Challenges and Ideas

RCUK Cloud special Interest Group

Build relationships with

Task Force
RCUK Cloud WG Membership

- David Colling, Imperial College
- Tim Cutts, Sanger Institute
- David Fergusson, University of Edinburgh
- Martin Hamilton, Jisc (Chair of SIG)
- Adam Huffman, Francis Crick Institute
- Philip Kershaw, CEDA, STFC (Chair)
- Steven Newhouse, EMBL-EBI & ELIXIR
- David Salmon, Jisc
- Simon Thompson, Birmingham University
- Jeremy Yates, UCL (Secretary)
Community building: workshops

- Imperial College, December 2015
  - ~40 attendees
  - Invited presentations

- Crick Institute in November 2016
  - ~120 attendees
  - Combination of invited speakers and talks solicited from the community
    - Talks from the big 3 hyper-scale providers
    - Community and private cloud, OpenStack
    - Legal, policy and regulatory issues
    - Lightning talks from the community
    - Breakout/Interactive session

- Outcomes
  - Desire for collaboration around task forces e.g. HTC and HPC compute (need to co-ordinate with HPC-SIG)
  - Cost and performance issues on-prem. compared with public cloud

Photos courtesy of Martin Hamilton, Jisc
Community building: website

- [https://cloud.ac.uk/](https://cloud.ac.uk/)

- Minutes from working group public

- Reports on and presentations from workshops:
  - [https://cloud.ac.uk/2017/03/20/cloud-workshop/](https://cloud.ac.uk/2017/03/20/cloud-workshop/)
  - [https://cloud.ac.uk/workshops/nov2016/](https://cloud.ac.uk/workshops/nov2016/)

- A resource to report back on our findings
  - Technical suitability of workloads for cloud e.g. HPC and parallel file systems
  - Legal, regulatory, policy issues, costs
Legal, Policy and Regulatory Issues

- The goal is to bottom out real and perceived issues around the use of public cloud and to provide guidance to the research community.
  - Build on work that has already done
  - Recognise that this is a changing and evolving area

- Strawman prepared January 2016

- Questionnaire to obtain feedback from the research community (Martin Hamilton)

- Session at November 2016 Workshop at the Crick, input from:
  - EMBL-EBI – experience from recent public cloud procurement
  - QMUL Cloud Legal Project

- Briefing note from the WG to provide guidelines currently in preparation

**Was legal advice sought?**

- No: 26%
- Yes: 44%
- Not sure: 30%

**Response** | **Organisations**
--- | ---
Review is still ongoing | 10
OK to use public cloud | 5
Not OK to use public cloud | 1
Occasional OK but restrictions around data, security and budget | 1
We have many use cases, hard to find general answer | 1
N/A – did not need permission for this particular use case | 1
Inconclusive result | 1
Technical Integration

• What are the opportunities, challenges, barriers?
  – Managing and tracking costs
  – Matching research workloads (e.g. HPC) with public cloud architectures
  – Matching more traditional data access with cloud native e.g POSIX, parallel file systems and object stores
  – Hybrid public/private – ability to move data and compute easily between providers

• How can we inform ourselves?
  – Interaction and communication within the community to track developments e.g. OpenStack Scientific WG, contact with representatives from public cloud providers
  – Dedicated pilots or ‘task forces’ organised through the WG to target particular areas of interest . . .
## Workflow Types and Key Bottlenecks

<table>
<thead>
<tr>
<th>Code/Workflow Requirement</th>
<th>Type</th>
<th>Interconnect</th>
<th>On chip memory Cache</th>
<th>OK on Cloud infrastructure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loosely Parallel: Independent Jobs run on a single core or node. No, or very little, communication between cores</td>
<td>HTC</td>
<td>1-10Gbs, 100μsec</td>
<td>1-2 MB</td>
<td>Yes</td>
</tr>
<tr>
<td>Weakly Parallel: Single instance of a code, run on many cores. Sporadic communication between cores at a high rate</td>
<td>HPC</td>
<td>50Gbs, 2-5μsec</td>
<td>1-2MB</td>
<td>No</td>
</tr>
<tr>
<td>Strongly Parallel: Single instance of a code, run on many cores. High and constant communication between cores</td>
<td>HPC</td>
<td>100-200Gbs, 0.2-0.1μsec</td>
<td>&gt;1GB</td>
<td>No</td>
</tr>
<tr>
<td>Data Intensive: Multiple distinct activities, coupled via intensive file system operations</td>
<td>HPC</td>
<td>50Gbs, 2-5μsec</td>
<td>&gt;1GB</td>
<td>No</td>
</tr>
</tbody>
</table>
A Hierarchy of systems – latencies (signal travel time) and bandwidth (Signalling Rate)

• Point-to-point latencies via the main switch are quoted
  • 40-100 Gb/s; 10-80 microsecs Ethernet – packet can take 10-80 microsecs to be delivered
  • 100-200 Gb/s 0.6-1.5 microsecs for Infiniband
  • SMP, 56Gb/s, 0.2 microacrsec for latency

• A core is operating at 2.6GHz = 0.38 ns per cycle
  • 0.2 Microsec = 520 cycles (SMP) – needed so the cache-coherence can be maintained so cores’ RAM can appear as a single RAM
  • 1 microsec = 2600 cycles – reduces wait time for comms, but memory to memory comms is impracticable. Weakly Scaling codes and Strongly scaling codes (e.g. BG/Q 30GB/s)
  • 10 microsec = 26000 cycles – wait times mean too much time is spent waiting – Loosely parallel codes
  • 100 microsec = 260,000 cycles – embarrassingly parallel codes
Types of Computing

• High Throughput Computing (HTC) computes huge numbers of the same simple workflow
  – e.g. equipment simulations for LHC, bioinformatics, stacking images and spectra from medical scanners, telescopes, ISIS and DIAMOND.
  – This is the loosely coupled version of parallelism (embarrassingly parallel).
  – It really is about high throughput - producing large numbers of similar things.
  – **HTC tends to be coupled to experimental equipment in research e.g. the Large Hadron Collider**
  – Most HTC calculations now becoming data intensive as the measurements are so much more accurate and much larger

• High Performance Computing uses an aggregate of resources to run one instance of a code.
  – This requires high performant CPU, Interconnect and file system (Input/Output) performance >10x that needed for HTC
  – Job sizes run from 32 to 100,000 cores
  – Outputs can be TB to PB in scale
  – Special onchip memory (CACHE) configurations are required
  – The use of parallel file systems to have high IO rates to disk systems, 10-100 GB/s
  – Used to calculate weather, climate, materials properties, aerodynamics, chemical properties and kinetics, transport systems, environmental systems, structure of sub-atomic particles, planet formation
Technical Integration – current activities

• Portability between cloud platforms
  – Particle Physics Cloud Pilot
  – Workshop on use Terraform and Ansible (in planning)

• Use of parallel file systems with on-prem cloud:
  – https://cloud.ac.uk/reports/spectrumscale/
  – Workshop on use of FUSE file system in the summer (in planning)

• Bulk data movement – co-ordinating with work in the NeI Project Directors Group
Particle Physics Cloud Pilot
(an example task force)

• Goal:
  – Explore strategies for making code deployments interoperable between providers and so avoid vendor lock-in

• To investigate:
  – Ability to port given workloads between different public cloud providers with minimal changes
  – Challenges related to workload needs and cloud topology
  – Focus on compute aspects rather bulk data movement
  – Focus on functionality rather than performance

• Domain-specific use case: workloads for Particle Physics CMS and ATLAS experiments
  – Work carried out by Andrew Lahiff, Scientific Computing Dept., STFC
  – Started mid-2016
Particle Physics Cloud Pilot: technical approach

- Uses container-based solution from the ground up as a means of abstraction for portability
  - Docker + Kubernetes

- Kubernetes
  - Supports abstraction e.g. underlying storage with StorageClasses
  - Powerful for automated deployment and scaling

- On hyper-scale providers: AWS, Azure and Google Cloud Platform
  - Possible with thanks to the providers for donated free credits

- Target workloads
  - CMS Monte Carlo simulation - compute intensive
  - LHCb Monte Carlo just completed, ATLAS jobs planned
  - Next steps: focus on io intensive workload
Particle Physics Cloud Pilot: summary by provider

- **Google Compute Platform**
  - Supports Kubernetes out-of-the-box
  - Supports auto-scaling in response to demand
  - Web portal Kubernetes interface – very easy to get up and running
  - Fairly extensive exploration of functionality

- **Azure**
  - Out-of-the-box since close 2016 with web portal interface
  - Supports other container orchestration technologies e.g. DCOS
  - Current work
    - integration of Azure Kubernetes cluster with ATLAS Big PanDA workload management system
    - Use of Azure Blob storage with DynaFed storage caching system
  - Scope for further extensive testing over the coming months

- **AWS**
  - No Kubernetes out-of-the-box but 3rd party solutions like StackPoint ([https://stackpoint.io/](https://stackpoint.io/)) can be used to overlay Kubernetes cluster across a given cloud provider tenancy
  - Very limited testing in this pilot to date
Particle Physics Cloud Pilot: federation across Google and Azure

Example with 1 cluster in Azure & 4 clusters in different regions in Google Cloud Platform

- single API to deploy applications on both Azure and Google Cloud Platform
- run one command to create squids in all 5 clusters

Courtesy of Andrew Lahiff, STFC
Next steps

• Technical Integration
  – Particle Physics Cloud Pilot: in i/o intensive workloads, close and report findings
  – HTC and cloud (depending on interest from the community): David Colling (Imperial) and David Salmon (Jisc)
  – HPC and cloud: discussions underway with HPC-SIG, tracking HPC on public cloud work e.g. NERC NCAS and Azure

• Workshops – training and dissemination
  – Lustre on cloud (planned): Simon Thompson, Birmingham
  – Using Terraform and Ansible to make workloads portable (planned): Steven Newhouse

• Legal, policy and regulatory issues, costs
  – Provide briefing note to provide guidelines to community: David Salmon
  – Track cost models of public providers

• Community: engage with Boston Open Research Cloud initiative
  – https://drive.google.com/drive/folders/0B4Y7flFgUgf9dElkaFkwUrhKblU
  – Attending first meeting in May following OpenStack Summit
Further information

- Working Group website: [https://cloud.ac.uk/](https://cloud.ac.uk/)
- NIST SP800-145 Cloud definition: [http://dx.doi.org/10.6028/NIST.SP.800-145](http://dx.doi.org/10.6028/NIST.SP.800-145)
- POSIX, Parallel file systems and cloud: [https://cloud.ac.uk/reports/spectrumscale/](https://cloud.ac.uk/reports/spectrumscale/)
- Boston Open Research Cloud initiative
  - [https://drive.google.com/drive/folders/0B4Y7flFqUgf9dElkaFkwUUhKblU](https://drive.google.com/drive/folders/0B4Y7flFqUgf9dElkaFkwUUhKblU)
- [philip.kershaw@stfc.ac.uk](mailto:philip.kershaw@stfc.ac.uk), @PhilipJKershaw
STFC Facilities Computing

Brian Matthews

Group Leader, Data Science and Technology
Scientific Computing Department

brian.matthews@stfc.ac.uk
Scientific Computing Department

- ~180 between RAL and DL
- Software engineering expertise
  - Data management systems
  - Visualisation and analytics
  - Applied maths
  - Computational science
- Systems hosting and management
  - GRID-PP Tier 1
  - JASMIN
  - Facilities Computing
To/from CERN:
Up to 40Gb/s

To/from SuperJanet5:
Up to 40Gb/s
- 10,000 batch cores
- ~15PB disk storage
- ~16 PB tape storage
- ~450kW power

RAL data centre is one of 11 Tier1 sites around the world which receive CERN LHC data

Also extensive private Cloud

Higgs discovery
- 11 PB received by RAL
- 21 PB sent by RAL
- 60 PB processed by RAL
JASMIN

Urgency to provide better environmental predictions
• HPC for higher-resolution models
But…
• Massive data requirement: observational data transfer, storage, output, post-processing

- 16 PB Fast Storage (Panasas, many Tbit/s bandwidth)
- 1 PB Bulk Storage
- Elastic Tape
- 4000 cores: half deployed as hypervisors, half as the “Lotus” batch cluster.
- Some high memory nodes, a range, bottom heavy.
Rutherford Appleton Laboratory Campus

- ISIS Neutron Source
- Central Laser Facility
- R89 Data Centre
- Diamond Light Source
Computing for Facilities

- SCD provides infrastructure support for STFC facilities
  - ISIS, CLF, DLS
- Computing in facilities too (especially DLS)
  - Controls, DAQ
  - Reduction and analysis packages
  - Support for user during the visits
- SCD provides
  - Central storage and archiving
  - Dedicated HPC
  - Computational software development

- There is a huge potential for future computing infrastructure support
  - As Facilities Science becomes ever more Data Intensive

Diamond

ISIS
SCARF: Cluster for STFC

- STFC local HPC cluster supporting STFC scientists, collaborators and facility users
  - Experimental simulations, computational models
  - Users from SCD, ISIS, CLF and others including RAL Space and Diamond
- 6200 cores,
  - 5000 generally available, remainder are for particular user communities (CLF Plasma Physics, ISIS IBIS)
  - with access to 5 terabytes of Fast Parallel IO memory (Panasas)
- Additional 4800 cores this year
  - 3200 common + 1600 for CLF
- 374 registered users (169 active)
Supporting Facilities Data Management

- **STFC Scientific Computing Department**
  - Support three STFC Funded facilities on the RAL campus
  - Provide data archiving and management tools

- **ISIS Neutron and Muon Source**
  - Provide tools to support ISIS’s data workflows
  - Support through the science lifecycle
  - Support for richer metadata
  - Provide data archiving
  - Support data Publication

- **DLS Synchrotron Light Source**
  - Data Archiving
  - Limited metadata in archive
  - Managing the scale of the archive

- **Central Laser Facility**
  - Real-time data management and feedback to users
  - Rich metadata on laser configuration
  - Access to data
  - Fast-data rates (0.05Hz to 10 Hz) on the horizon – will change the
DLS Per Beamline Data Rates

Tomography largest generator of data

But now EM
2 Microscopes @
  5TB/day

XFEL from 2017
**Diamond Data Volumes**

- **X-Ray Detectors**
  - 2016: Eiger 16M for MX
    - 6.75 - 13.5 GB/s
  - Beamlines with TB/day

- Data collected by DLS
  - ~7.2 PB
  - 1,500,000,000 files
  - Rate double ~7.5 months
    - cataloguing 12000 files/min

Tomography Beamlines have collected over 2PB of data.

**New Arrival Electron Microscopes:**
- 2 Microscopes @ > 5TB/day

**EU-XFEL:**
- 5000 frame/sec: ~ 50 GB/s
Data Management Systems

ICAT Data Management Suite

Integrated data management pipelines
- From data acquisition to storage to publication

**Metadata as Middleware**
- A Catalogue of Experimental Data
- Automated metadata capture
- Integrated with the User Office and data acquisition system

Providing access to the user
- TopCat web front end
- Integrated into Analysis frameworks
  - Mantid for Neutrons, DAWN for X-Rays

15 years effort to build data management systems

DLS Archive of 4.7PB, 1100 million files (April 2016)
- 4.7PB, 1100 million files (April 2016)
- 2.2PB, 620m Jan 2015

ISIS Data Archive
- ~50Tb
- Full experimental Metadata

ICAT Open Source Collaboration: [www.icatproject.org](http://www.icatproject.org)
The amount of data catalogued by Diamond has now surpassed 1.5 billion files and totals more than 7.2 petabytes.

- But that’s the easy bit!
- Users need to analyse the experimental data

A Single Tomographic Image (Diamond I12)
- 120GB raw data
  - 30 mins collection
  - Can collect many in a visit
- 80 Gb reconstructed image
  - 30 mins on 24 GPU cards (FBP algorithm)
- Iterative algorithms
  - Less raw data (12GB per image – 3 mins)
  - Much longer processing (~10 hrs)

THIS IS HARD FOR USERS
ISIS Data Analysis Challenges

• Neutrons are rare: Data sets tend to be smaller
• Lots of modelling and simulation to interpret the data (e.g. RMC)
  • Combining models, simulations and data analysis
  • Complex algorithms and tools
  • Good visualisation

Summarising:
• Neutron science is now: DATA INTENSIVE
• The users can’t handle the data + algorithms
• The science is being affected by the computing

THIS IS HARD FOR USERS
Facilities Science is Data Intensive

Before Visit (at Univ)

Modelling of sample
Simulation of experiment
Re-analysis of previous results

In Experiment (at Facility)

Rapid analysis and visualisation for steering
Quick simulations for steering

In Visit (at Facility)

Data reduction and sampling
- User leaves with “clean” data

After Visit (at Univ)

Detailed analysis – from data to information.
Incorporating results from other experiments.
Compare with models.
Develop published data and papers

ISSUES
- Models more complex
- Codes more complex
- Data volumes greater
- Access to HPC for quick turn around

DO MORE IN THE STFC DATA CENTRE
- Data archives available
- Computational clusters

MAKE COMPLEX COMPUTING TASKS ROUTINE
- Software expertise
- Give remote access to users

ISSUES
- Data movement
- Data storage
- Models more complex
- Data analysis more complex
- Access to HPC/HTC
- Expertise rarer!
Tomographic Reconstruction

- Support in-experiment and post-experiment tomographic reconstruction
  - Round-trip the data to HPC CPU/GPU clusters in experiment time
  - A tomographic image reconstruction toolbox with different algorithms
  - High throughput image reconstruction framework – time scheduled
  - Visualisation on the beamline or remote
  - An integral component of IMAT’s in-experiment data analysis capability through Mantid (ISIS) and DAWN (DLS),

- Maximise the science resulting from Data collected on facility instruments.

STFC Scientific Computing: Erica Yang, Srikanth Nagella, Martin Turner, Derek Ross
STFC ISIS: Winfried Kockelmann, Genoveva Burca, Federico Montesino Pouzols
DLS: Mark Basham
CCP4-DAAaaS

CCP4 – Macro-Crystallography suite
  • proteins, viruses and nucleic acids
  • determine macromolecular structures by X-ray crystallography
  • Used by DLS users
    • But need post-experimental access

Data Analysis as a Service
  - Remote access to data and compute via SCD Cloud
  - CCP4 s/w maintained on Cloud via VM packaging and distribution (CVMFS)
  - User Portal provides access to right data and compute and workflows.
Materials Workbench

• Vibrational motion of atoms crucial for a material’s properties
  – e.g., how well it conducts electricity or heat

• Peaks in INS spectrum correspond to specific atomic vibrations

• Peak assignment: what specific vibrational motions of atoms give rise to specific peaks?

• INS spectra can be computed for a given atomic structure

• Calculations allow us to see what specific vibrational motion of atoms occur, and at what frequency

• Materials Workbench
  – An environment for running simulations and comparing with experimental data
Data Analysis Algorithms

• Feature extraction from full-spectrum neutron images using Principal Component Analysis
  • Joe Kelleher, Genoveva Burca, ISIS

- Processing energy-discriminating image data
  • Image processing technique
  • Emphasises different features

- Requires
  • Linear algebra algorithms
  • Cluster computing

- Working with SCD.
Ada Lovelace Centre

An initiative to support the future needs for data intensive science within large-scale facilities (STFC + DLS + CCFE)

Supplement Advances in detectors and sources with

Provision of high performance, data intensive computing resources &

Advanced Algorithm Development in Facilities Science Domains &

Integrated Science data-spaces and workflows to support users remotely

Supported by Research in Data and Computational science

Providing Training in Data and Computational Science
Summary

• SCD provides computing infrastructure, software and services to STFC programmes and facilities
  – In particular ISIS, DLS and CLF
  – Data archiving and distribution
  – HPC Clusters + private cloud
  – Computational Science Software (and increasingly data science).

• Changing facilities science landscape
  – Ever more data and compute intensive
    • Much more than traditional Facilities Science
  – Computing bottleneck for the science
  – More direct support for facility users in experiment interpretation
Report from the Computer Advisory Panel
David Colling on behalf of CAP
Charlotte Jamieson
Head of Enabling Themes
STFC
What is CAP?

“The purpose of the Computing Advisory Panel (CAP) is to advise the STFC Executive on the strategy for, and management of, provision for computing resources (including data handling, data storage, software and hard provisions, skills, and developments in high performance and high throughput computing) in support of programmes either funded or delivered by STFC.”
Membership

• Professor Stephen Fairhurst – Cardiff University (Chair)
• Professor David Colling – Imperial College London
• Professor Martin Dove – Queen Mary University of London
• Dr David Fergusson – The Francis Crick Institute
• Dr John Pasley – University of York
• Professor Arttu Rajantie – Imperial College London
• Dr Debora Sijacki – University of Cambridge
• Dr Stuart Sim – Queen’s University Belfast
• Mr Ash Vadgama – AWE
• Professor Daniel Watts – University of Edinburgh

In attendance: STFC
David Corney
Anthony Davenport
Charlotte Jamieson
CAP activities

- Reports to Science Board
- Provides input through e.g. balance of programmes
- Input to the 2015 "Computing Strategic Review"
- Review and provide feedback to computing models (e.g. SKA)
- Review computing needs and provisions at STFC facilities
Computing Eco system

- There is a need for all scales of computing (from Tier 1 to Tier 3)
- Includes high performance and high throughput
- Innovation & development occur at Tiers 2 and 3, then migrate to Tier 0 and 1. (less true in HEP than many other communities)
- Expect an order of magnitude increase in required hardware, storage, network bandwidth over next 5 years.
Key issues

• Computing is essential to nearly all areas of STFC science and this is a growing dependence.
• Still sometimes treated as optional bolt on - “if we had provided for our actual computing needs then we couldn’t have built …”
• Need for computing will grow by an order of magnitude over the next ~5 years with new projects such as LSST, SKA, HL-LHC …
• Hardware is important but so are Software Engineers and Data Scientists.
Computing Costs

- Naively costs scale linearly with computing

- Can achieve saving with:
  - scale
  - sharing infrastructures (UK T0)
  - Employing better software engineers

These may require money to be spent to save money

However this never happens and even with Moore’s Law more money will needed to be spent on computing if the needed expansion is to be realised.
People

- Hardware is useless without people to support it
- Too often infrastructure is “batteries not included”
- Expert software engineers and data scientists are increasingly important
- Faculty, postdocs and PhD students don’t know how to make best use of multi-core machines, novel architectures, etc
- “If I had £1M to spend on computing, then it's better spent on £800k on hardware and £200k on people - I'm sure they can find a 25% speed-up on my codes.
- We need to find a career path for data scientists
- Rapidly increasing computing requirements
- Included in the ongoing Balance of Programmes exercise
- Experiments: GridPP (LHC), SKA, LSST, CTA
- DiRAC (Distributed Research using Advanced Computing)
- Grant-funded computing
DiRAC Services
(www.dirac.ac.uk)

Dr Jeremy Yates, Director
jyates@star.ucl.ac.uk
HPC and Data Intensive Science for Particle Physics, Astro Physics and Nuclear Physics

- Simulations
- Data Modelling
- 1 call per year
- Seedcorn projects can be applied for at any time
- Have a suite of systems that are focussed in particular problem types
- 1000 users and 80 projects with typically 300 users using the system each month and 44 active projects
- Cover the whole PPAN science range
• Solve CFD, MHD and radiative HD in unstructured grids
  • Particle in a Cell (SPH)
  • Grid methods (AMR)
• PDEs in structured grids as part of a MCMC formalism
• Solve Radiative Transfer (PDE/PiC) and Chemical kinetics (ODE)
• Require balanced systems between cores, RAM, interconnect (IB) and IO. All need to be fast and “matched”.
<table>
<thead>
<tr>
<th>Service</th>
<th>Cores or Tflops Available to DiRAC-2.5</th>
<th>Functionality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BG/Q (80% of original level)</strong> <em>(many core)</em></td>
<td>76800 <em>(1.04 Pflops)</em></td>
<td>PDE solver for structured Grid using MCMC methods</td>
</tr>
<tr>
<td><strong>Data Centric 2.5 Xeon Clusters</strong> <em>(multi-core)</em></td>
<td>14900 cores</td>
<td>Large Scale CFD, MHD, Radiative Hydro Hamiltonian</td>
</tr>
<tr>
<td><strong>Data Analytic 2.5 Xeon Clusters</strong> <em>(multi-core)</em></td>
<td>17750 cores</td>
<td>Data Intensive Research</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Small Scale CFD, MHD, Radiative Hydro Hamiltonian for data modelling</td>
</tr>
<tr>
<td><strong>Data Analytic 2.5 KNL Cluster</strong> <em>(many-core)</em></td>
<td>75 Tflops <em>(although 500 Tflops jobs can be run)</em></td>
<td>PDE solver for structured Grid using MCMC methods</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Machine Learning</td>
</tr>
<tr>
<td><strong>Data Analytic 2.5 GPU Cluster</strong> <em>(many-core)</em></td>
<td>150Tflops <em>(although 1 Pflops jobs can be run)</em></td>
<td>CFD, MHD, Radiative Hydro Hamiltonian</td>
</tr>
<tr>
<td><strong>Complexity 2.5 (multi-core)</strong></td>
<td>7000 cores</td>
<td>Large Scale Radiative Hydro with physico-chemical processes</td>
</tr>
</tbody>
</table>
## Workflow Types and Key Bottlenecks

<table>
<thead>
<tr>
<th>Code/Workflow Requirement</th>
<th>Type</th>
<th>Interconnect</th>
<th>On chip memory Cache</th>
<th>OK on Cloud infrastructure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loosely Parallel: Independent Jobs run on a single core or node. No, or very little, communication between cores</td>
<td>HTC</td>
<td>1-10Gbs, 100μsec</td>
<td>1-2 MB</td>
<td>Yes</td>
</tr>
<tr>
<td>Weakly Parallel: Single instance of a code, run on many cores. Sporadic communication between cores at a high rate</td>
<td>HPC</td>
<td>50Gbs, 2-5μsec</td>
<td>1-2MB</td>
<td>No</td>
</tr>
<tr>
<td>Strongly Parallel: Single instance of a code, run on many cores. High and constant communication between cores</td>
<td>HPC</td>
<td>100-200Gbs, 0.2-0.1μsec</td>
<td>&gt;1GB</td>
<td>No</td>
</tr>
<tr>
<td>Data Intensive: Multiple distinct activities, coupled via intensive file system operations</td>
<td>HPC</td>
<td>50Gbs, 2-5μsec</td>
<td>&gt;1GB</td>
<td>No</td>
</tr>
</tbody>
</table>
A Hierarchy of systems – latencies (signal travel time) and bandwidth (Signalling Rate)

• Point-to-point latencies via the main switch are quoted
  • 40-100 Gb/s; 10-80 microsecs Ethernet – packet can take 10-80 microsecs to be delivered
  • 100-200 Gb/s 0.6-1.5 microsecs for Infiniband
  • SMP, 56Gb/s, 0.2 microsec for latency
• A core is operating at 2.6GHz = 0.38 ns per cycle
  • 0.2 Microsec = 520 cycles (SMP) – needed so the cache-coherence can be maintained so cores’ RAM can appear as a single RAM
  • 1 microsec = 2600 cycles – reduces wait time for comms, but memory to memory comms is impracticable. Weakly Scaling codes and Strongly scaling codes (e.g. BG/Q 30GB/s)
  • 10 microsec = 26000 cycles – wait times mean too much time is spent waiting – Loosely parallel codes
  • 100 microsec = 260,000 cycles – embarrassingly parallel codes
Types of Computing

- High Throughput Computing (HTC) computes huge numbers of the same simple workflow
  - e.g. equipment simulations for LHC, bioinformatics, stacking images and spectra from medical scanners, telescopes, ISIS and DIAMOND.
  - This is the loosely coupled version of parallelism (embarrassingly parallel).
  - It really is about high throughput - producing large numbers of similar things.
  - **HTC tends to be coupled to experimental equipment in research e.g. the Large Hadron Collider**
  - Most HTC calculations now becoming data intensive as the measurements are so much more accurate and much larger

- High Performance Computing uses an aggregate of resources to run one instance of a code.
  - This requires high performant CPU, Interconnect and file system (Input/Output) performance >10x that needed for HTC
  - Job sizes run from 32 to 100,000 cores
  - Outputs can be TB to PB in scale
  - Special onchip memory (CACHE) configurations are required
  - The use of parallel file systems to have high IO rates to disk systems, 10-100 GB/s
  - Used to calculate weather, climate, materials properties, aerodynamics, chemical properties and kinetics, transport systems, environmental systems, structure of sub-atomic particles, planet formation