# 格点 QCD 计算软件新框架计划

宫明

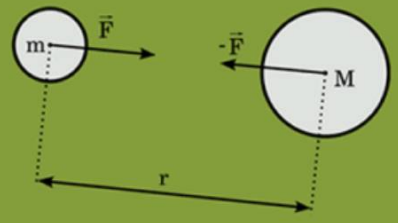
中国科学院高能物理研究所


2023 开放科学计算联盟学术年会

2023.6.28

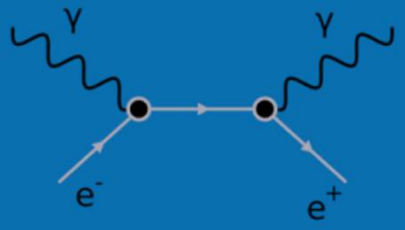
# 理论物理学对世界的理解

$$G_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + g_{\mu\nu}\Lambda = \frac{8\pi G}{c^4}T_{\mu\nu}$$



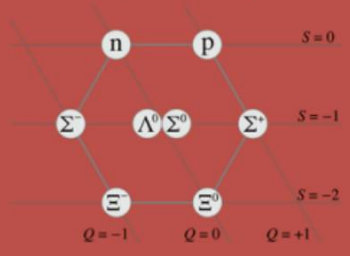

引力

$$\mathcal{L} = \bar{\psi}\left(i\gamma^\mu D_\mu - m\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$



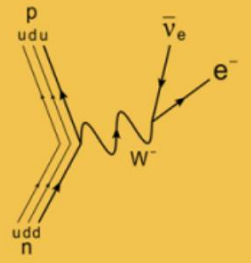

电磁力

$$\mathcal{L} = \bar{\psi}_i\left(i\gamma^\mu\left(D_\mu\right)_{ij} - m\delta_{ij}\right)\psi_j - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a$$




强核力

$$\mathcal{L} = g\left(\bar{\nu}_{eL}, \bar{e}\right)\gamma^\mu\left\{\begin{pmatrix} -\sqrt{1+\xi^2}Z_\mu & 0 \\ 0 & \frac{\xi A_\mu}{\sqrt{1+\xi^2}} - \frac{\xi^2}{\sqrt{1+\xi^2}}Z_\mu \end{pmatrix} + \frac{1-\gamma^5}{4}\begin{pmatrix} -\sqrt{1+\xi^2}Z_\mu & -\sqrt{2}W^+_\mu \\ -\sqrt{2}W_\mu & \sqrt{1+\xi^2}Z_\mu \end{pmatrix}\right\}\begin{pmatrix}\nu_{eL} \\ e\end{pmatrix}$$




弱核力

# 格点量子色动力学（Lattice QCD）

## 唯一靠谱的研究强相互作用的方案

- 从第一原理出发（没有手放进去的额外假设）
- 非微扰方法（大多数情况下微扰方法对 QCD 无效）
- 误差可控（系统误差与统计误差）

## 完全依赖计算机资源

- 基本方法是蒙特卡洛数值积分
  - 统计误差与计算资源的 1/2 次方成反比
  - 系统误差与计算资源的依赖更明显
- 主要计算核心是巨型稀疏矩阵乘法与线性方程组求解
  - 代码原理简洁，实际应用规模巨大
  - 对浮点性能、通讯带宽、通讯延迟都有很高的要求。

# 计算资源需求

## 美国高能物理学在 E 级计算机上的需求预算

| Computational Task | Current Usage | 2025 Usage | Current Storage (Disk) | 2025 Storage (Disk) | 2025 Network Requirements (WAN) |
|---|---|---|---|---|---|
| Accelerator Modeling | $\sim$ 10M – 100M core-hrs/yr | $\sim$ 10G – 100G core-hrs/yr | | | |
| Computational Cosmology | $\sim$ 100M – 1G core-hrs/yr | $\sim$ 100G – 1000G core-hrs/yr | ~10PB | >100PB | 300Gb/s (burst) |
| Lattice QCD | ~1G core-hrs/yr | $\sim$ 100G – 1000G core-hrs/yr | ~1PB | >10PB | |
| Theory | $\sim$ 1M – 10M core-hrs/yr | $\sim$ 100M – 1G core-hrs/yr | | | |
| Cosmic Frontier Experiments | $\sim$ 10M – 100M core-hrs/yr | $\sim$ 1G – 10G core-hrs/yr | ~1PB | 10 – 100PB | |
| Energy Frontier Experiments | $\sim$ 100M core-hrs/yr | $\sim$ 10G – 100G core-hrs/yr | ~1PB | >100PB | 300Gb/s |
| Intensity Frontier Experiments | $\sim$ 10M core-hrs/yr | $\sim$ 100M – 1G core-hrs/yr | ~1PB | 10 – 100PB | 300Gb/s |

[1]

[1] ASCR/HEP Exascale requirement review report, arxiv:1603.09303

# 格点 QCD 的软件配套

## USQCD 二十多年的软件积累

| | | | | |
|---|---|---|---|---|
| **Chroma** | CPS | FUEL | MILC | QLua |
| Inverters | MDWF | QOPQDP | **QUDA** | **QPhiX** |
| QDP++ | QDP/C | **QDP-JIT** | **Grid** | |
| QLA | **QMP** | QMT | QIO | |

## CLQCD 近年来面向 E 级超算的软件发展

- 向曙光平台移植了 QUDA 和 Chroma，并自研了 pyQuda 等软件
- 在申威平台上自研了 SWLQCD 的初步版本，通过 QSunway 接口集成到 Chroma
- 在天河三原型机上移植并优化了 Chroma 和 Grid
- 为天河三异构平台研发了DDQ的支持接口

# 脱钩的危险？弯道超车的机会？

## 当代的挑战

- 新硬件平台与老软件框架之间的矛盾
  - 裱糊不可取，离开屎山、不破不立！
- 物理学的思维方式与软件工程之间的矛盾
  - 需要用高度抽象的封装隔离物理学公式和代码实现！
- 巨大的工作量与捉襟见肘的人力资源之间的矛盾
  - 让合适的人做合适的事，让计算机做不适合人类的事！

# 脱钩的危险？弯道超车的机会？

## 当代的挑战

- 新硬件平台与老软件框架之间的矛盾
  - 裱糊不可取，离开屎山、不破不立！
- 物理学的思维方式与软件工程之间的矛盾
  - 需要用高度抽象的封装隔离物理学公式和代码实现！
- 巨大的工作量与捉襟见肘的人力资源之间的矛盾
  - 让合适的人做合适的事，让计算机做不适合人类的事！

## 老笑话一则

数学家：你方程里的这几个量分别是什么？
物理学家：都是算符：D 是微分算符，U 是幺正算符，S 是旋量算符。
数学家：这几个符号表示什么？
物理学家：都是函数：f(x) 是标量函数，$\delta(x)$ 是德欧塔函数，$\Phi(x)$ 是分布函数，$\phi(x)$ 是波函数。
数学家：$A_\mu(x)$ 也是函数？
物理学家：不，$A_\mu(x)$ 是个场。

把各种物理量都分别实现为对应的程序对象？让程序具有物理意义？
不！这不是软件工程的思维方式。
PHAME (Principles of Hierarchy, Abstraction, Modularisation, and Encapsulation)

# 脱钩的危险？弯道超车的机会？

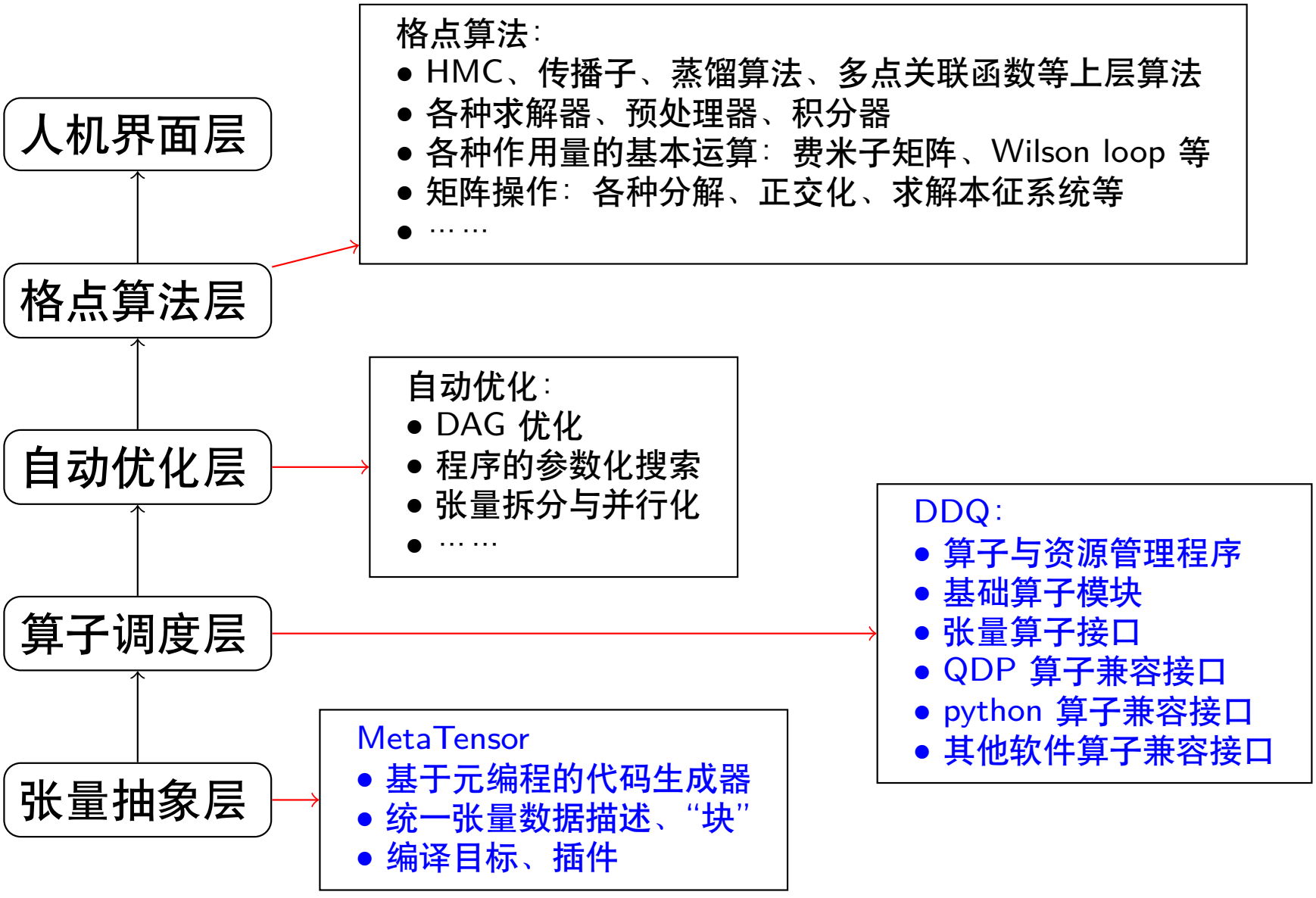
## 当代的挑战

- 新硬件平台与老软件框架之间的矛盾
  - 裱糊不可取，离开屎山、不破不立！
- 物理学的思维方式与软件工程之间的矛盾
  - 需要用高度抽象的封装隔离物理学公式和代码实现！
- 巨大的工作量与捉襟见肘的人力资源之间的矛盾
  - 让合适的人做合适的事，让计算机做不适合人类的事！

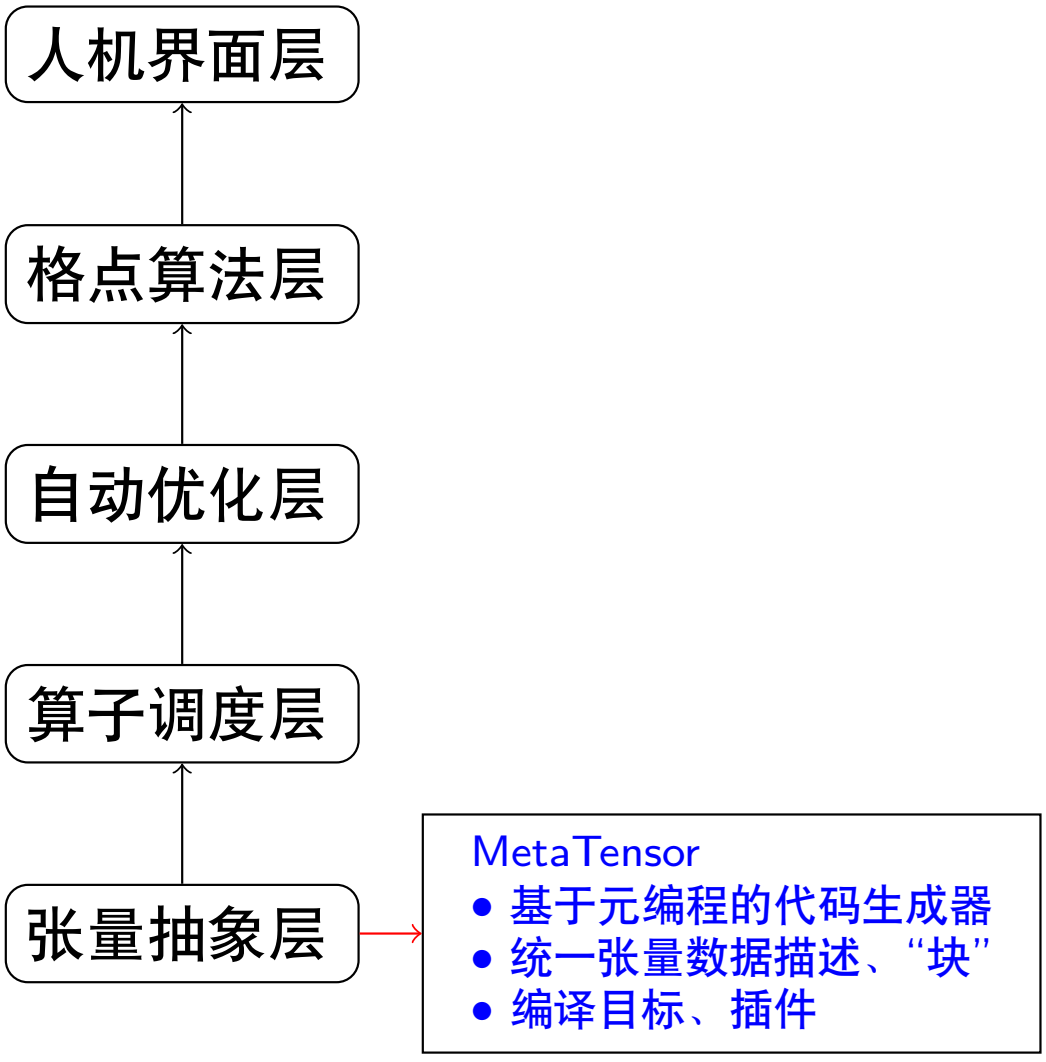
## 新软件框架的设计原则

- 要对各种硬件平台提出极简的抽象接口，并用其对代码分层
- 要对各种物理计算需求提出极简的抽象接口，并用其对代码分层
- 要保留一些适合自动优化的透明接口，对人工智能虚位以待

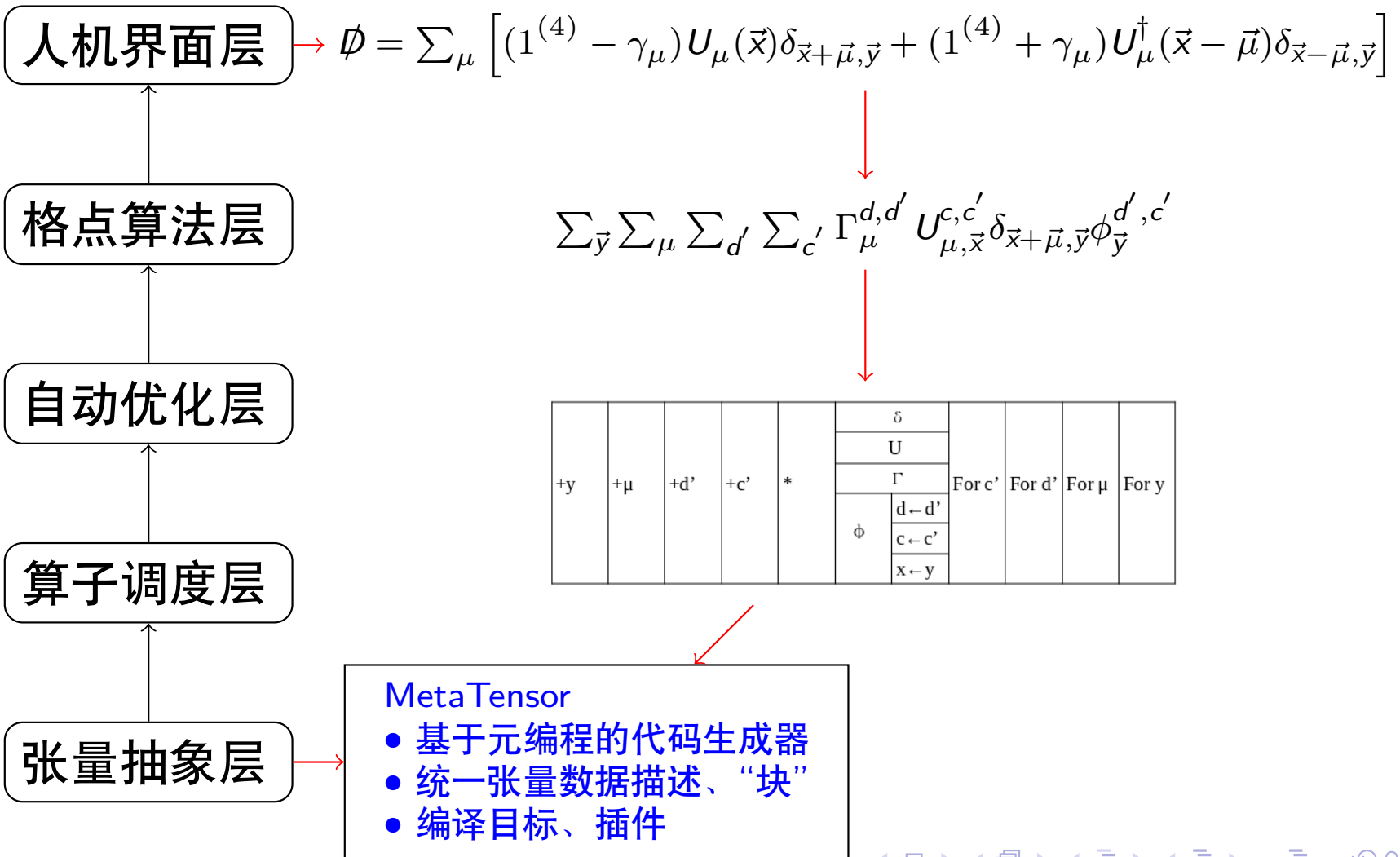
# 格点 QCD 新软件框架设计



# 格点 QCD 新软件框架设计

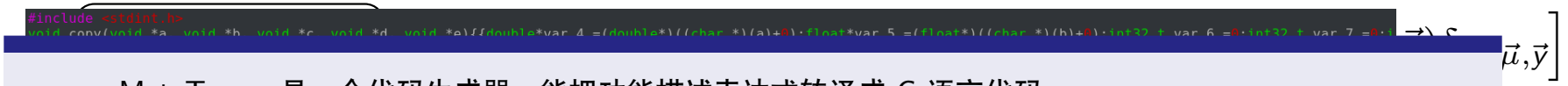


# 格点 QCD 新软件框架设计



# 格点 QCD 新软件框架设计

# 格点 QCD 新软件框架设计

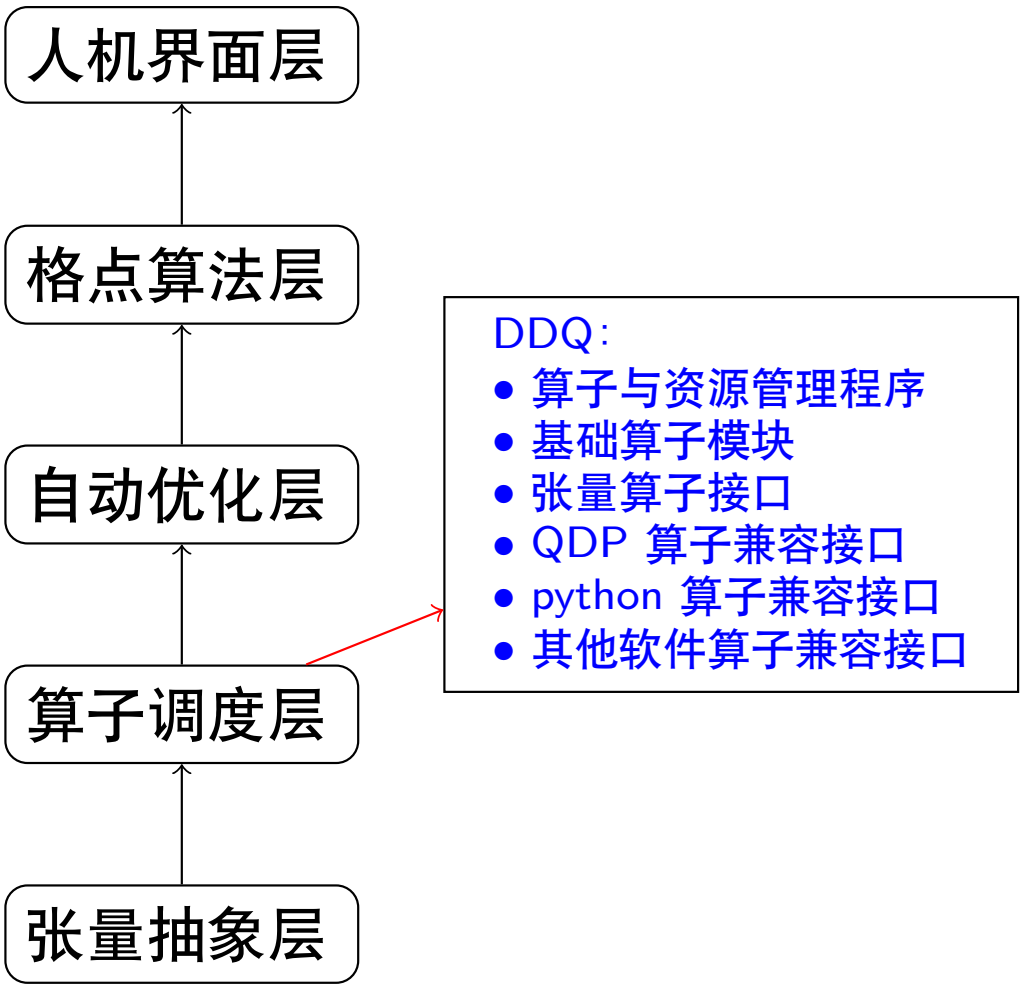



- MetaTensor 是一个代码生成器，能把功能描述表达式转译成 C 语言代码
- 它包含 4 个插口，用于功能的不断扩充：
  - 1、datacopy ：
    - （本地/远程/集合/同步）数据拷贝
    - 数据布局转置/归并
    - 数据类型/大小端转换
    - 数据归约
    - 数据广播
  - 2、block2data：
    - 所有张量计算
    - 迭代计算
  - 3、optimize：
    - 程序等价推演
    - 编译期代码优化
  - 4、compile：
    - 孤立代码的编译入口
- 目前成功运行了“Hello World”程序：第一个 datacopy 插件

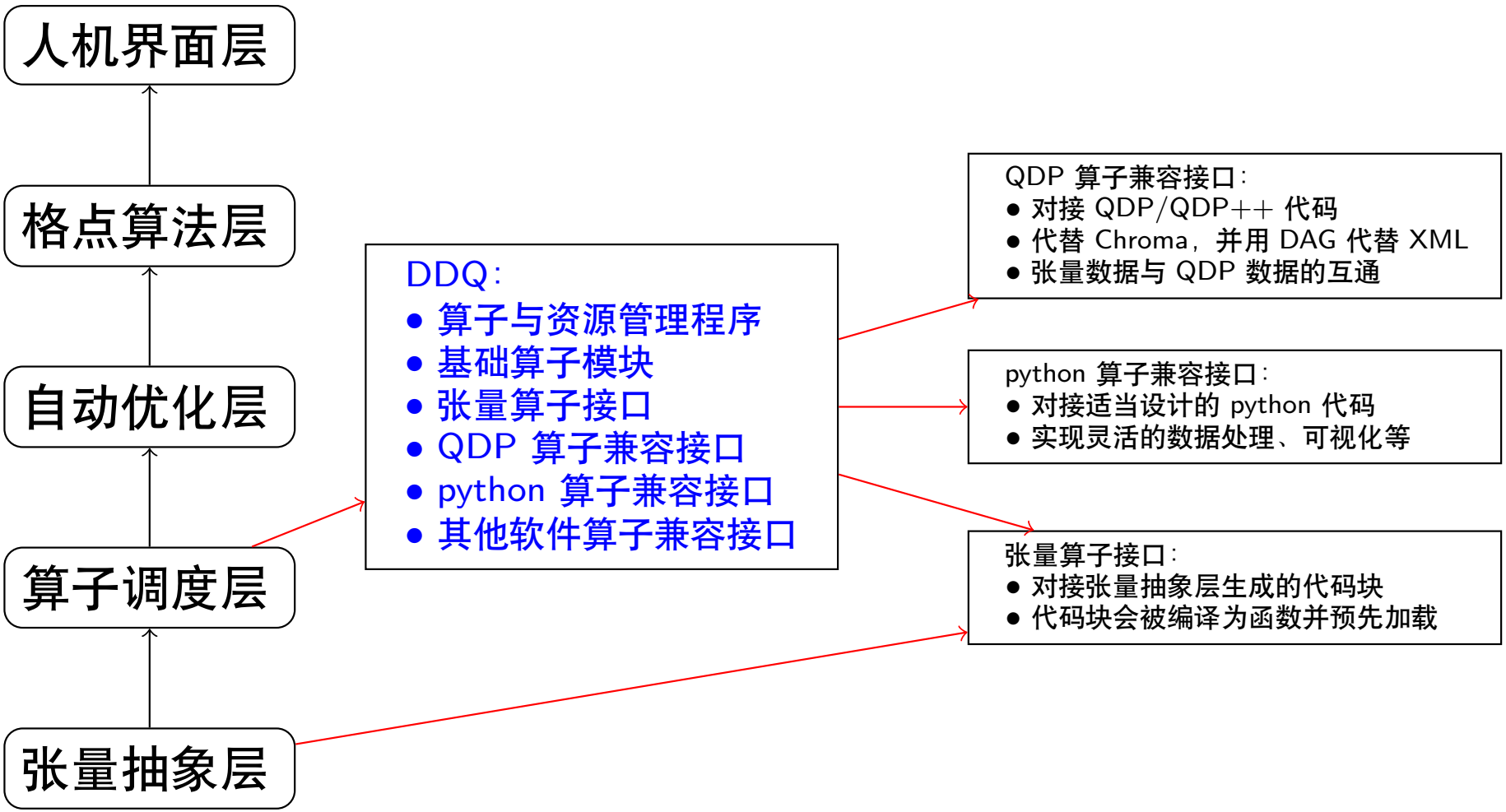
张量抽象层

MetaTensor
- 基于元编程的代码生成器
- 统一张量数据描述、“块”
- 编译目标、插件

# 格点 QCD 新软件框架设计



# 格点 QCD 新软件框架设计



# 格点 QCD 新软件框架设计

人机界面层

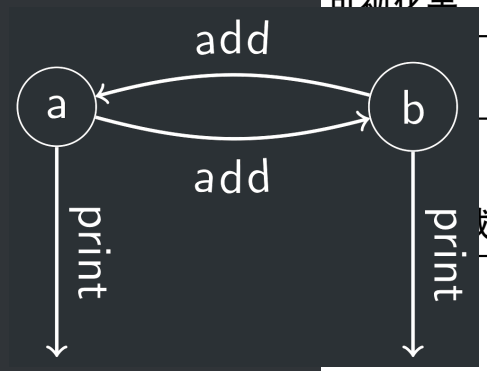
格点算法层

自动优化层

算子调度层

张量抽象层

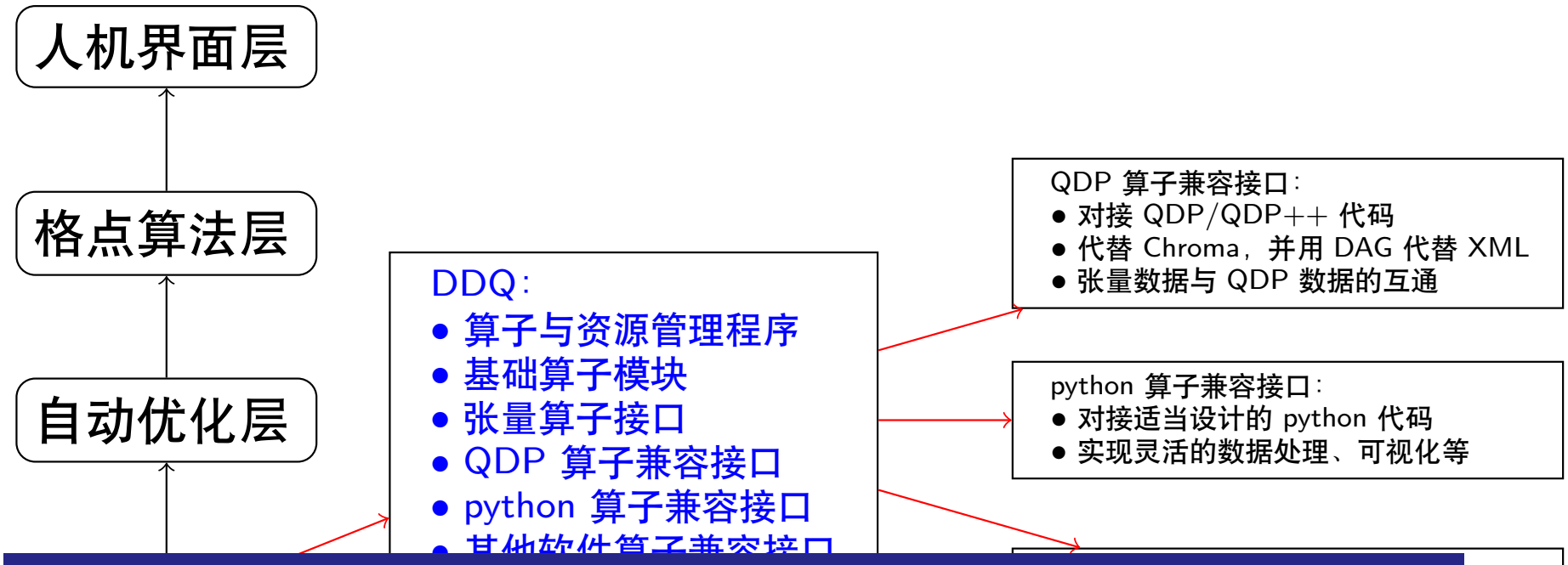
```
task_ret        f_print(void **inputs, void **outputs);
task_ret        f_add(void **inputs, void **outputs);

int     main()
{
        obj     f_a, f_p, a, b;
        ddq_op  ab, ba, pa, pb;

        unsigned long   a0, b0;
        a0 = b0 = 1;

        f_a = obj_import(f_add, NULL, obj_prop_ready);
        f_p = obj_import(f_print, NULL, obj_prop_ready);

        a = obj_import(&a0, NULL, obj_prop_consumable | obj_prop_ready);
        b = obj_import(&b0, NULL, obj_prop_consumable);

        ab = ddq_spawn(processor_pthread, 1, 1);
        ab->f = f_a;
        ab->inputs[0] = a;
        ab->outputs[0] = b;

        ba = ddq_spawn(processor_pthread, 1, 1);
        ba->f = f_a;
        ba->inputs[0] = b;
        ba->outputs[0] = a;

        pa = ddq_spawn(processor_pthread, 1, 0);
        pa->f = f_p;
        pa->inputs[0] = a;

        pb = ddq_spawn(processor_pthread, 1, 0);
        pb->f = f_p;
        pb->inputs[0] = b;

        ddq_loop();

        return 0;
```




代码
AG 代替 XML
的互通

n 代码
可视化等

# 格点 QCD 新软件框架设计




- DDQ 是一个运行时的算子调度器，能根据依赖关系动态安排算子的运行
- 它是跨平台的，通过插件支持各种硬件设备
  - 目前已初步支持了 CPU 单核、多核、GPU、天河三 DSP 等处理器
- 它可以自动实现异构并行化、双缓冲优化等等

# 目前状态

## 分层进展情况

- 张量抽象层的 MetaTensor 的设计环节基本完成，代码工作进展顺利
- 算子调度层的 DDQ 已初步成型，目前测试正常，正在继续完善
- 自动优化层的预研工作得到了乐观的结果，目前留出接口即可
- 格点算法层早已预研完成，等待底层工作完成后再开始研发
- 人机界面层初步设计了基于图形界面的新编程语言

## 预期进展

- 2 年内，实现最底两层的基本功能
- 4 年内，实现格点 QCD 的实用计算
- 未来保持持续更新，并积极研发新算法进行自动优化

## 资金支持

- 基金委重大项目《基于国产超算的格点量子色动力学关键科学问题研究》(2023-2027 年) 的子课题《基于国产超算的格点软件与数据》的主要任务

# 谢谢！