New Technologies for HEP - The CERN openlab

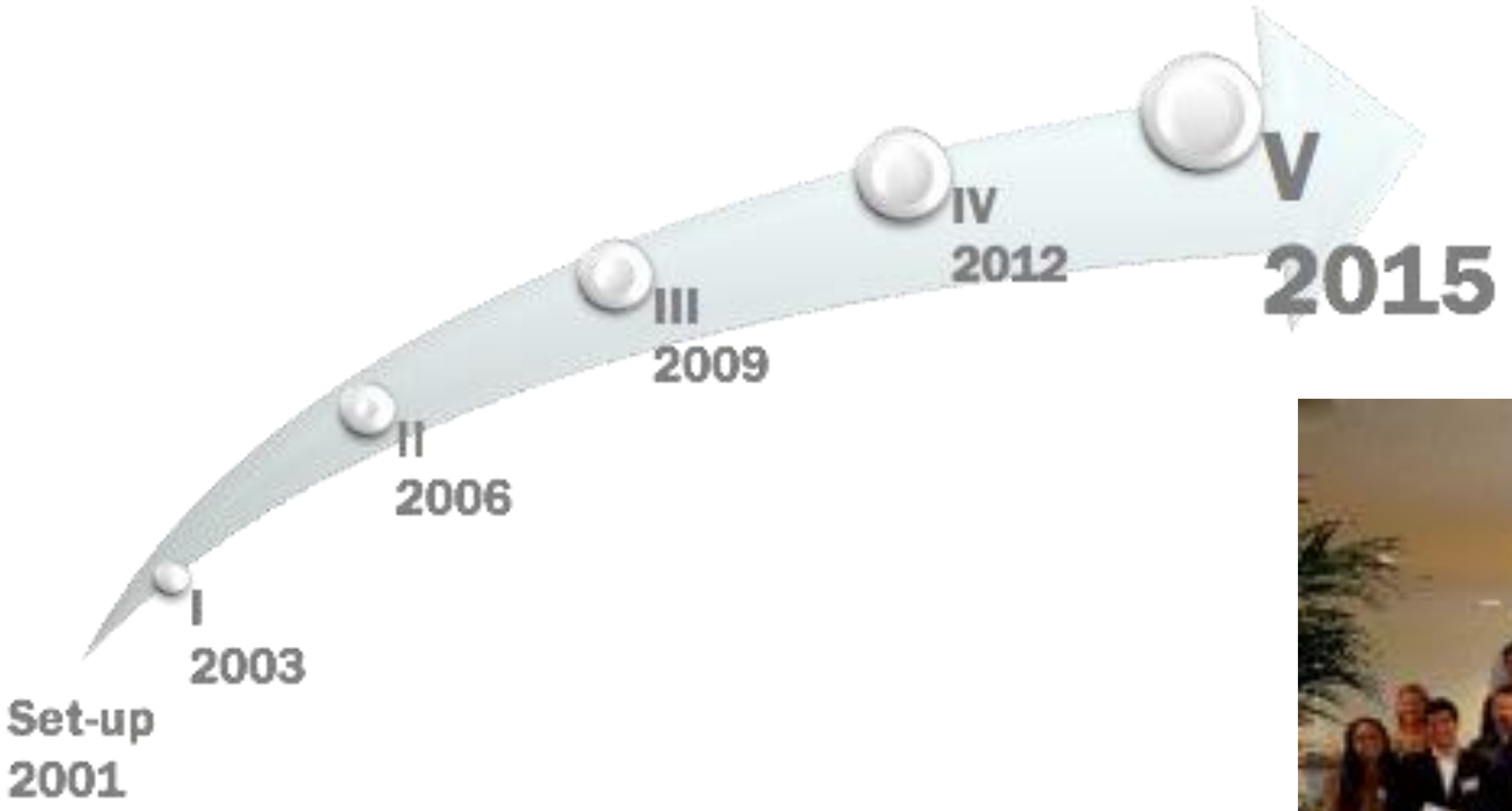
Fons Rademakers, CERN openlab Chief Research Officer
ACAT 2016, Valparaiso, 18-1-2016

CERN openlab

CERN openlab in a Nutshell

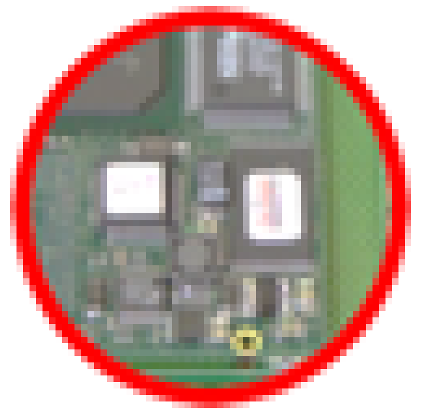
- A science – industry partnership to drive R&D and innovation with over a decade of success
- Evaluate state-of-the-art technologies in a challenging environment and improve them
- Test in a research environment today what will be used in many business sectors tomorrow
- Train next generation of engineers/employees
- Disseminate results and outreach to new audiences

The History of CERN openlab

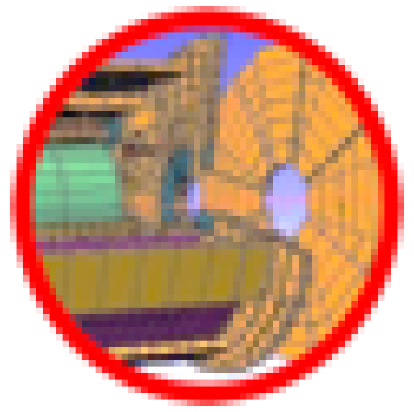


CERN openlab Board of Sponsor 2013

Information Technology Research Areas



Data acquisition and filtering



Computing platforms, data analysis, simulation

Data storage and long-term data preservation

Compute provisioning (cloud)

Networks

Data analytics

Medical applications

Who Are We Talking To

New Partners

New Educational Requirements

Multicore CPU programming, graphical processors (GPU), multithreaded software

Software & Computing Engineers

Data analysis technologies, tools, data visualization, monitoring, security, etc.

Data Scientists

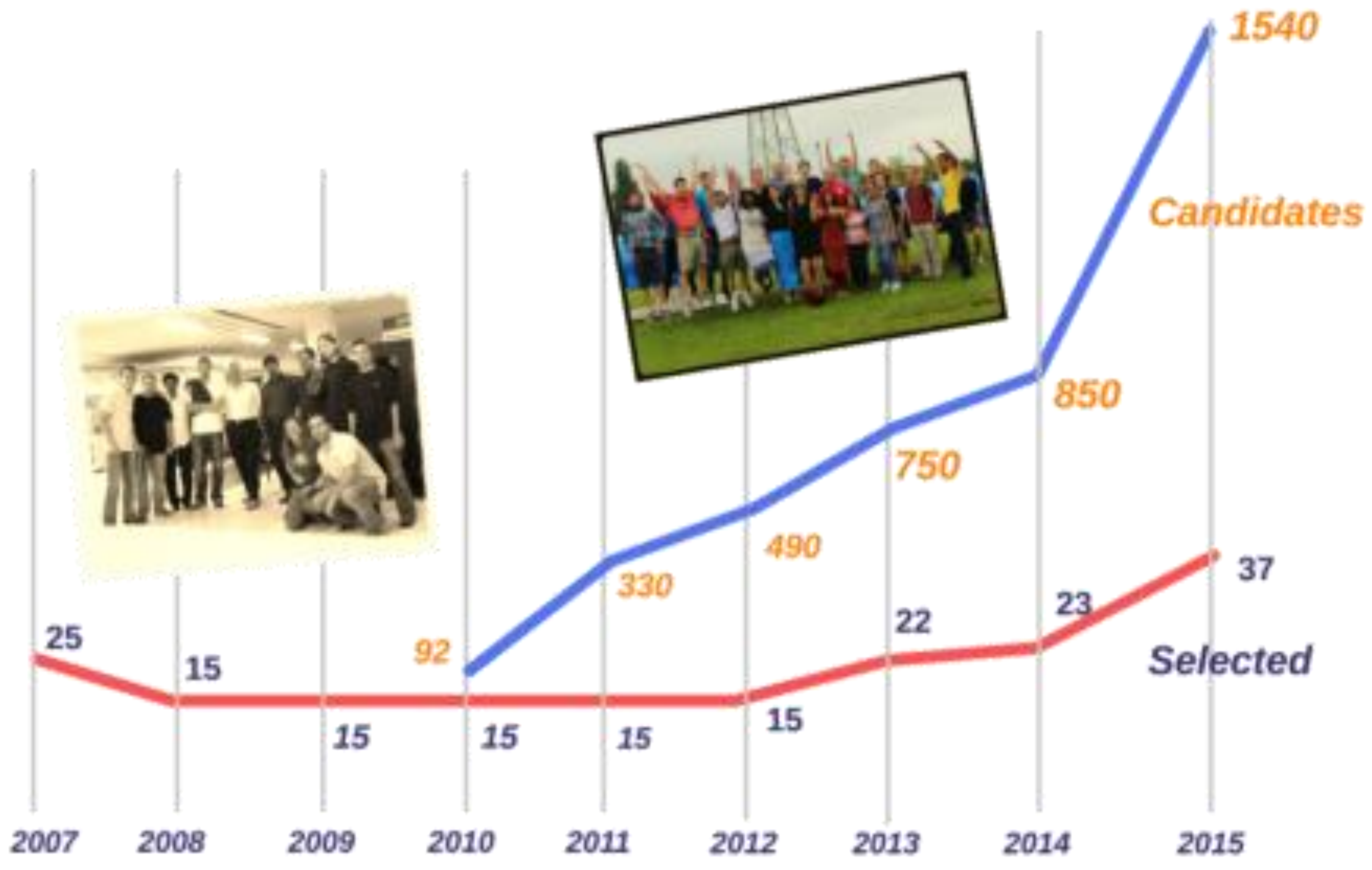
Applications of physics to medical research (hadron therapy, etc.), simulation software

Multidisciplinary applications

The Educational Program

- Most of the dedicated personnel in CERN openlab are young, talented Fellows receiving hands-on experience on new technologies
- A comprehensive offer of general and specific workshops, training events and initiatives
- Experts from industry and laboratories give lectures at events inside and outside CERN

Summer Student Program



- In 2015
  - 1540+ applicants
  - 40 selected students
  - 14 lectures
  - Visits to external labs and companies
  - Lightning talks sessions
  - Technical reports

CERN openlab Members and Projects

Intel

- High throughput computing project
  - Xeon + FPGA + omnipath, LHCb TDAQ
- Code modernization project
  - Geant V, FairRoot, Cx3D brain development simulation
- Rackscale project
  - Software defined racks
- Training, consultancy

Oracle

- Cloud and OpenStack
  - OVM integration with CERN OpenStack
- Data Analytics
  - Analytics as a Service (Endeca, Oracle R, etc.)
- Database and Systems Management
- Java Platform
- Replication using GoldenGate

Siemens

- Improve functionality, efficiency, and predictability of CERN control systems
  - Data Analytics
  - High performance archiving
  - Visualization
  - Development environment

Huawei

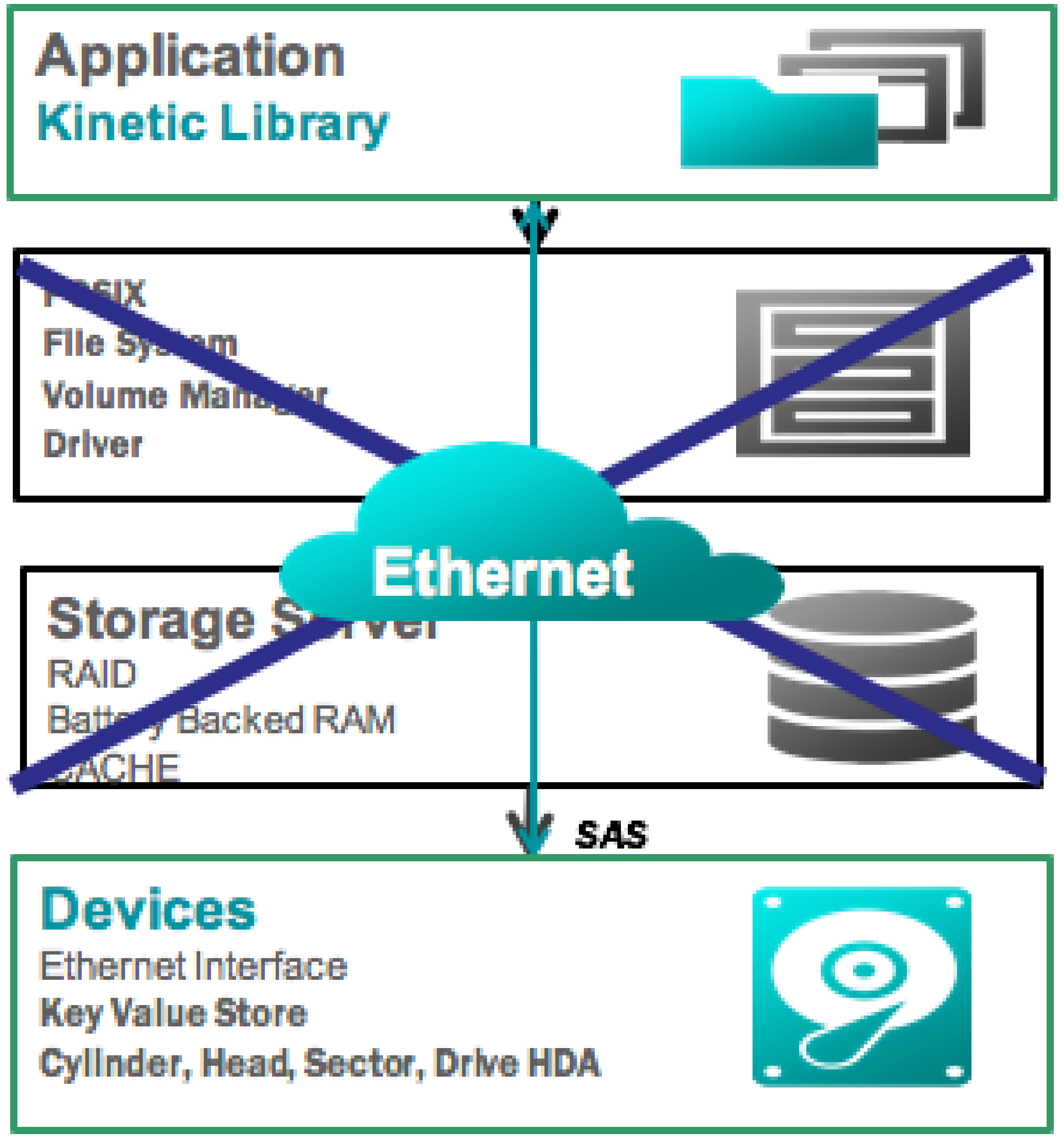
- Storage server projects
    - Test S3 compatibility
    - Test performance
    - Project finished
- ARM64 server evaluation, testing and benchmarking

Rackspace

- Cloud Federations
  - Create full orchestration capability
  - Manage virtual machines in remote clouds with a single identity
  - Done within the OpenStack development process

Seagate

- Current architectures built on layers of traditional technology
  - Translation overhead
  - Tiers of storage servers
- Kinetics cuts through these layers
  - Applications communicate directly
- Drive at higher abstraction level
  - More efficient than objects in a files system
  - Enables feature agility



KINETIC
Open Storage Project

- Started as a Seagate project, protocol & libraries now managed by the Linux Foundation

- December 2015 plugfest demonstrated Seagate / WD / Toshiba interoperability

http://www.openkinetic.org

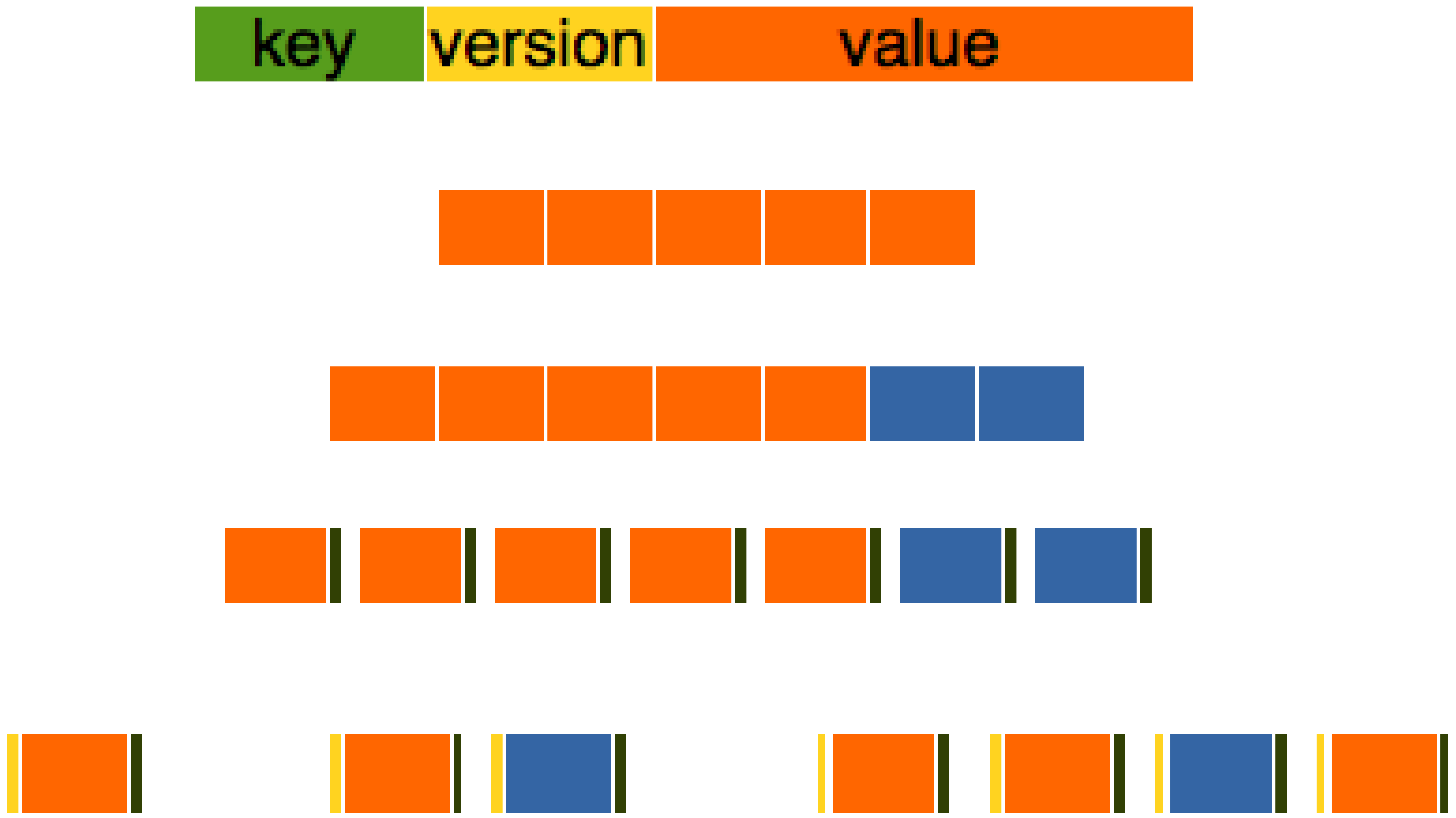
The Kinetic Key-Value Protocol

- Put/Get/Delete/… with a few extra’s

- Checksum: can be verified by the drive
  - No need to read data for scrubbing
- Version: test-and-set functionality
  - Drive-side concurrency resolution

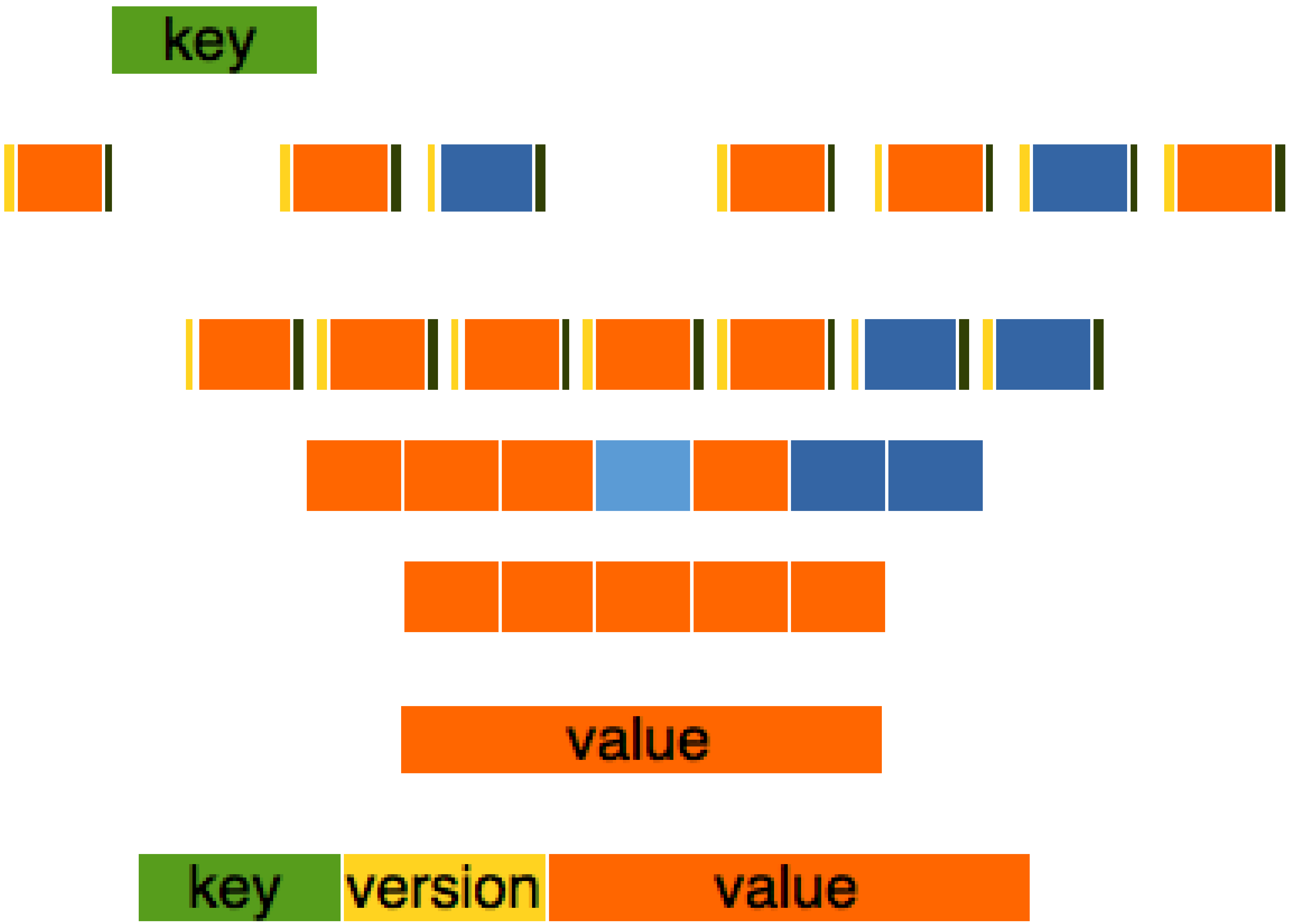
Cluster Logic - Put

- Put request
  - Chunk value
  - Erasure coding
  - Calculate crc
  - Assign drivers
  - Flush chunks

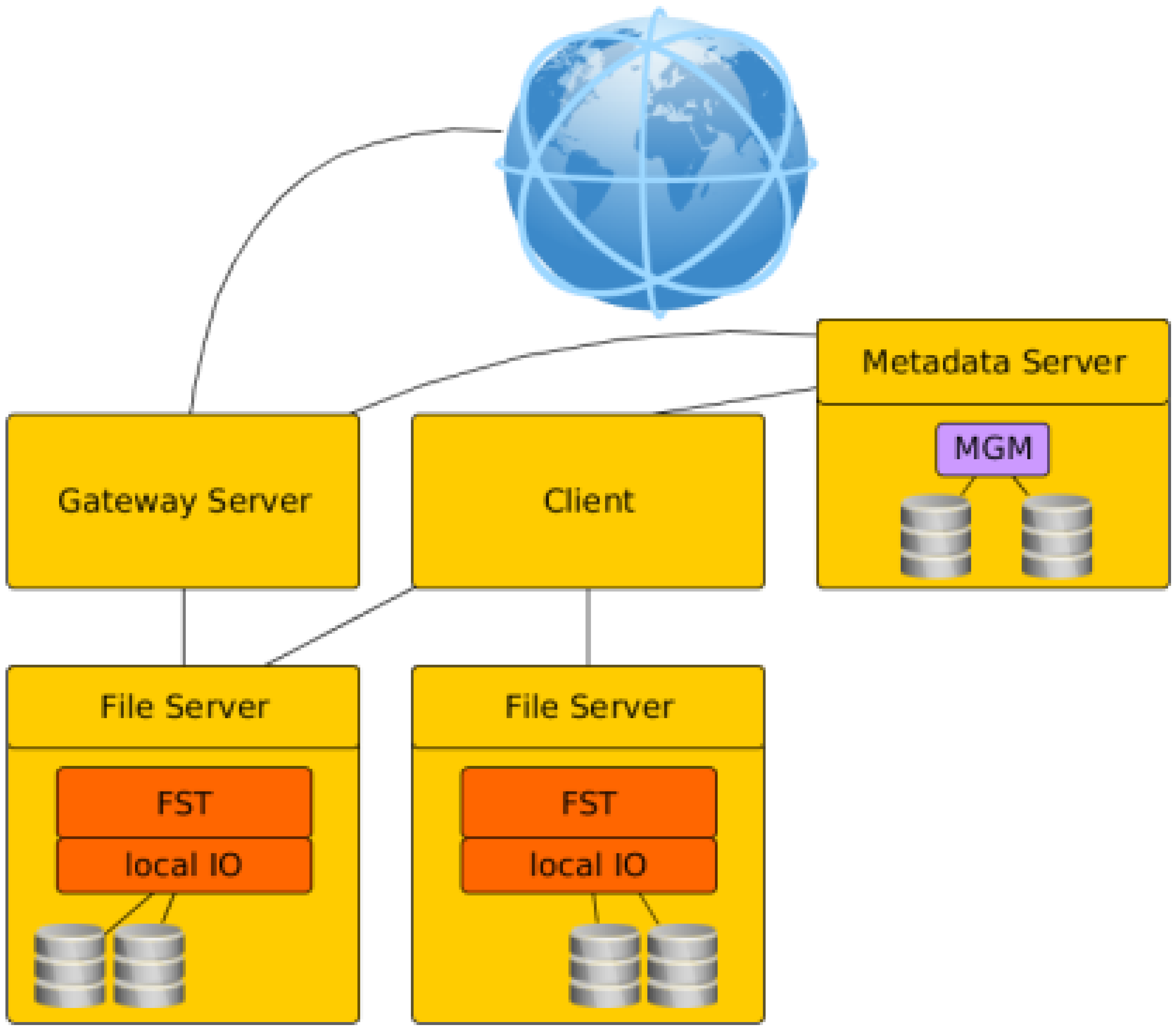


Cluster Logic - Get

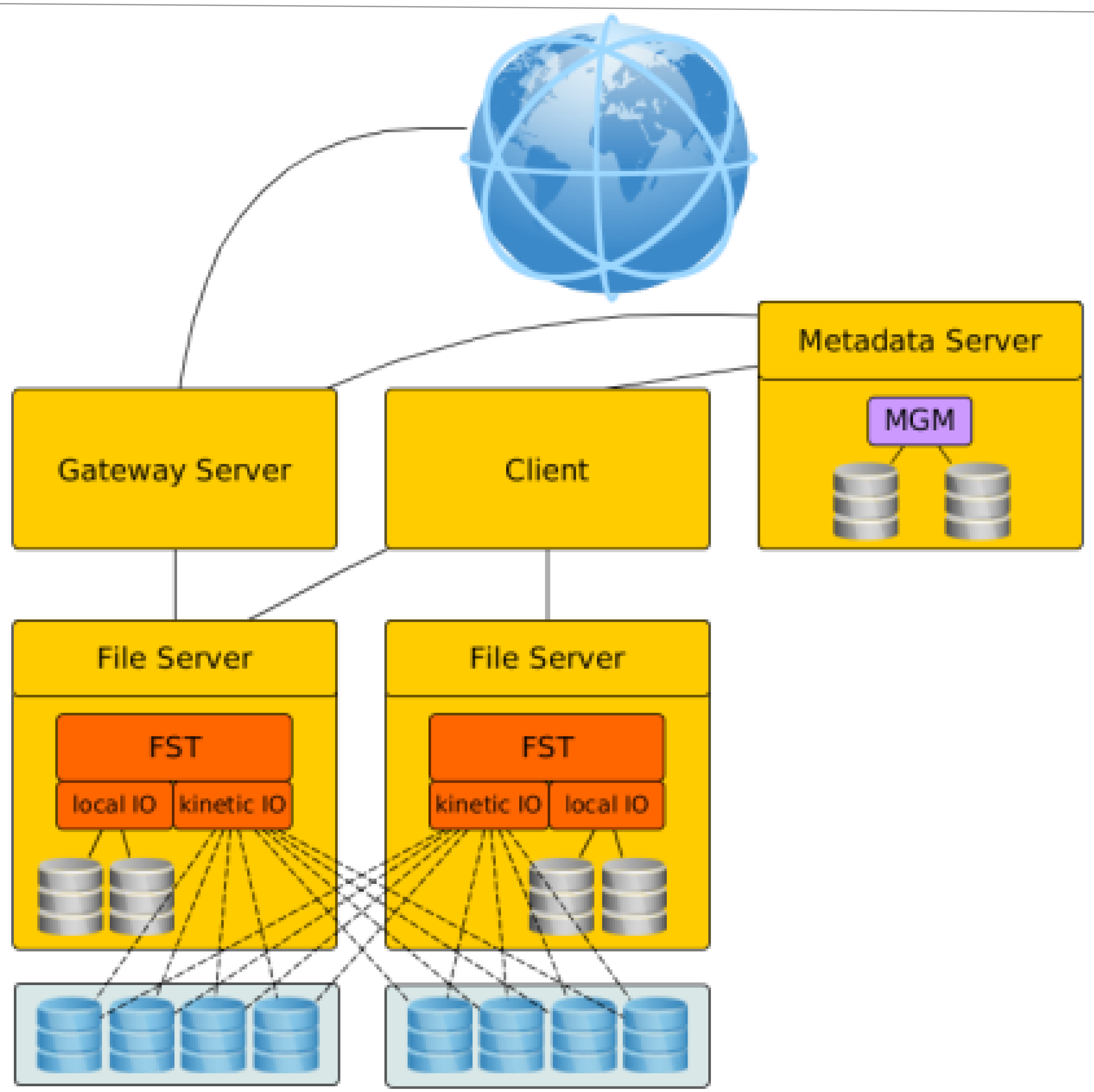
- Get request
  - Identify drives
  - Read chunks
  - Verify crc and versions
  - Erasure decode
  - Concatenate value



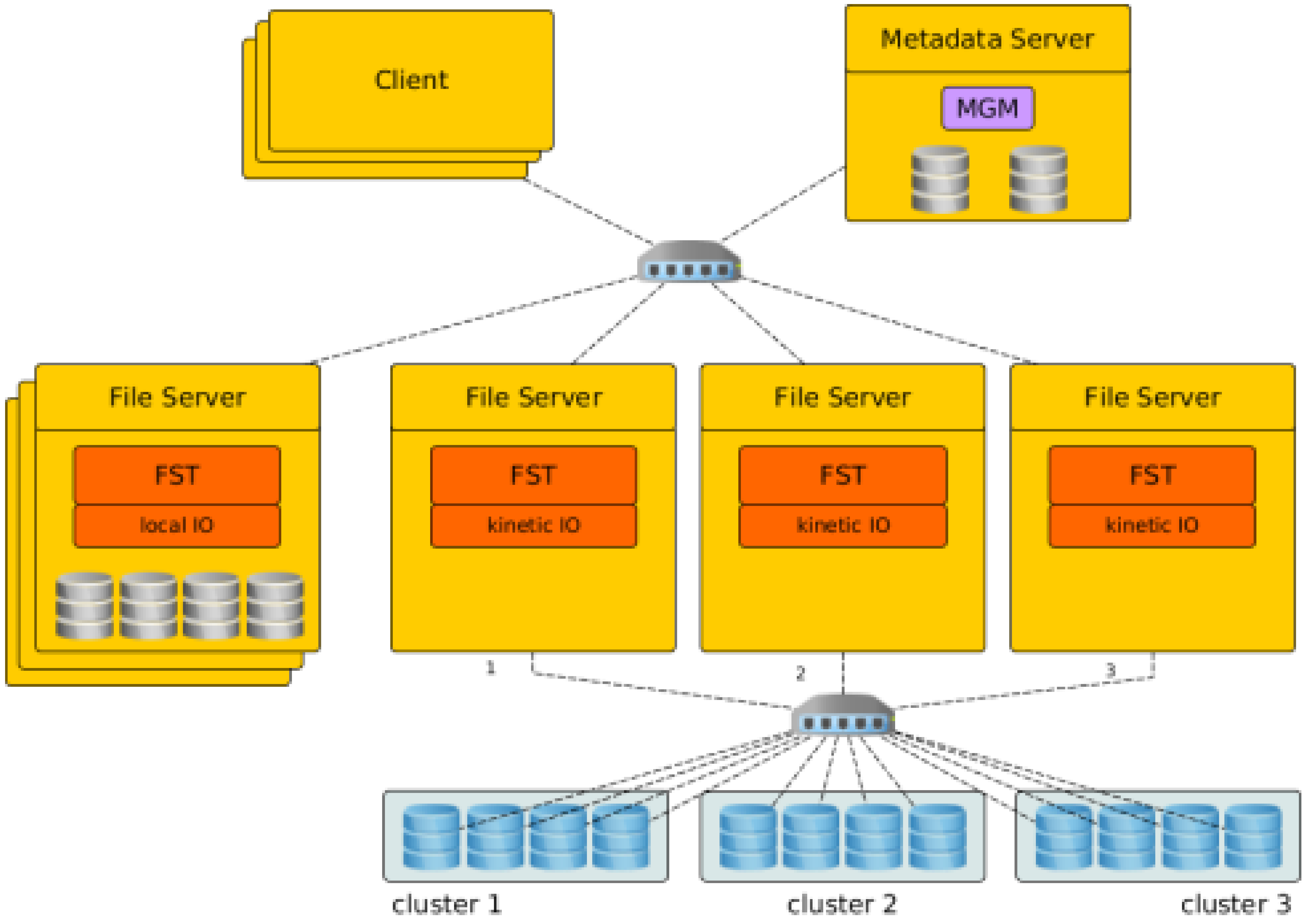
Basic EOS Architecture With I/O Plugin



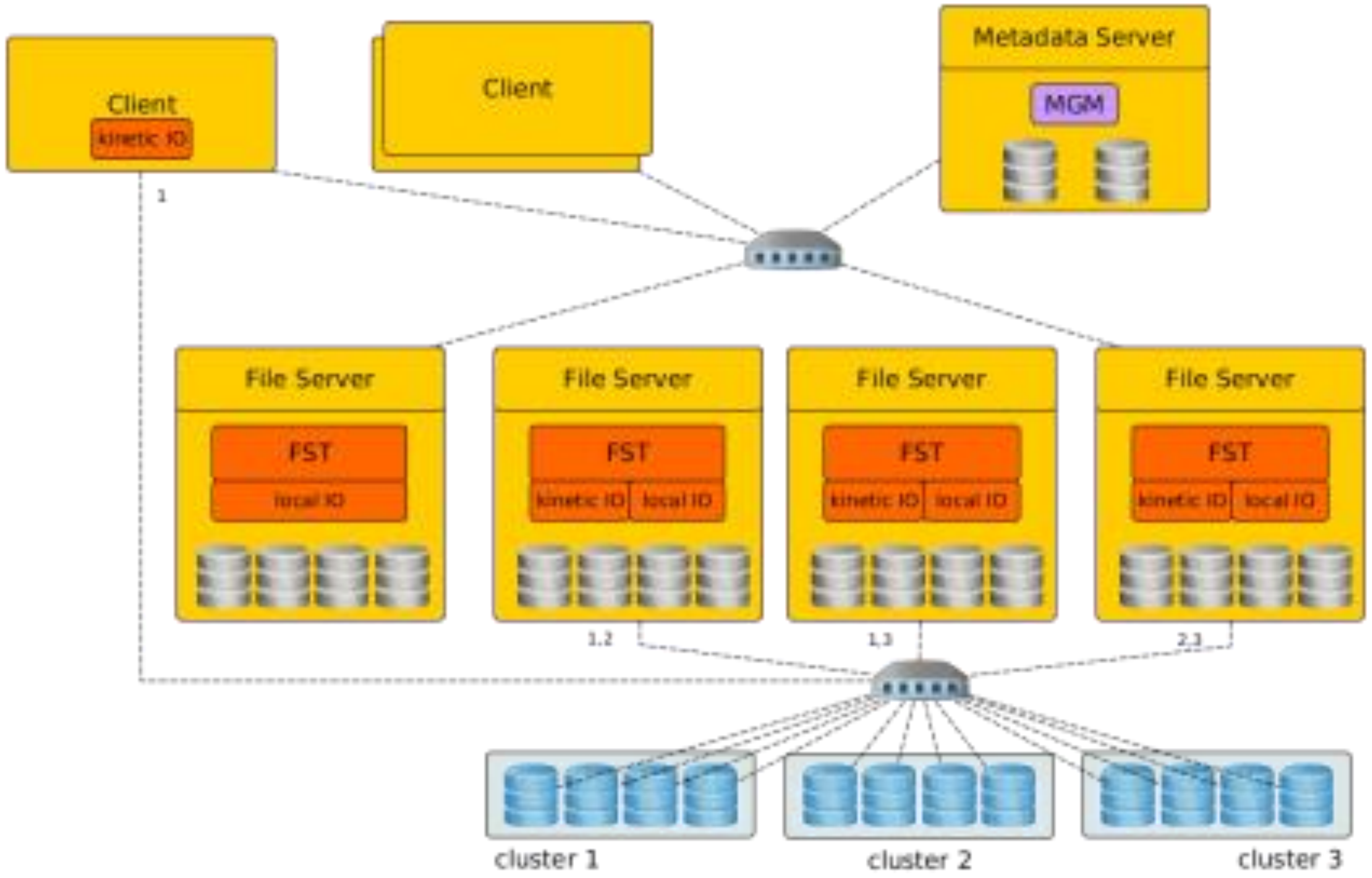
Basic EOS Architecture With I/O Plugin and Kinetic Support



Deployment Models - Dedicated



Deployment Models - Client Side Mounting



IDT

- RapidIO low-latency switch technology
  - Test and evaluate in analytics clusters
  - Test and evaluate in TDAQ environment

Cisco

- Build a rack-scale system with a modern OS including the following ideas:
  - Data plane OS for virtualized high-throughput I/O
    - Multi-kernel operating systems (Arrakis, Barrelfish)
    - Data transfer without kernel mediation (Cisco usNIC and libfabric)
  - Multicore systems
    - Decouple the CPU, kernel and the OS
  - Scaling beyond a single chassis
    - Using asynchronous message exchange

Brocade

- Build intelligent system that can optimize routing of data traffic entering and leaving an organization and drop network attacks
- The optimal routing or drop will be decided based on the information coming from network itself, from db of trusted applications and other sources

Yandex

- Data popularity project
  - Based on data usage patterns determine the data storage class
- Data verification project
  - Automatic detection of anomalies in the LHCb detector operating mode

Comtrade

- Customization and packaging of EOS

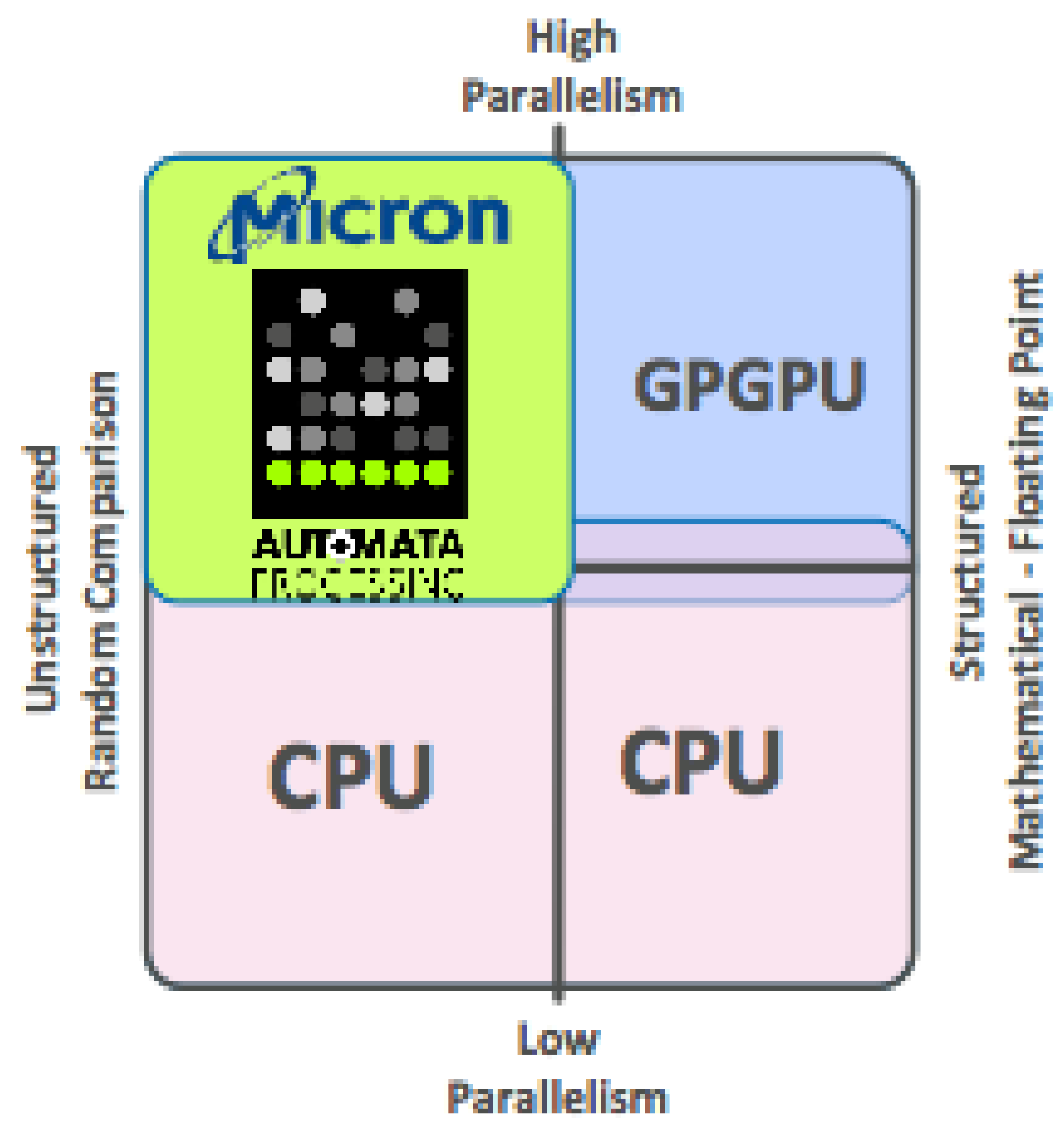
Micron (not yet, but hopefully soon a project)

- Automata processor evaluation
  - On the fly HEP pattern recognition processing
- NVRam 3DXPoint technology (developed with Intel)
  - Persistence storage with the speed of RAM, highly reduced I/O bottleneck
  - Reduced need for caches, language performance more important as the I/O waits are reduced

Automata Processor

Micron's ***Automata Processor*** is a revolutionary new class of programmable accelerator

- An industry-first hardware implementation of highly-parallel Non-deterministic Finite Automata (NFA)
- Orders of magnitude (>100x) faster than CPU's for pattern matching and graph analytics
- Rapidly reconfigurable for complex algorithms
- Simple parallel programming with familiar tools



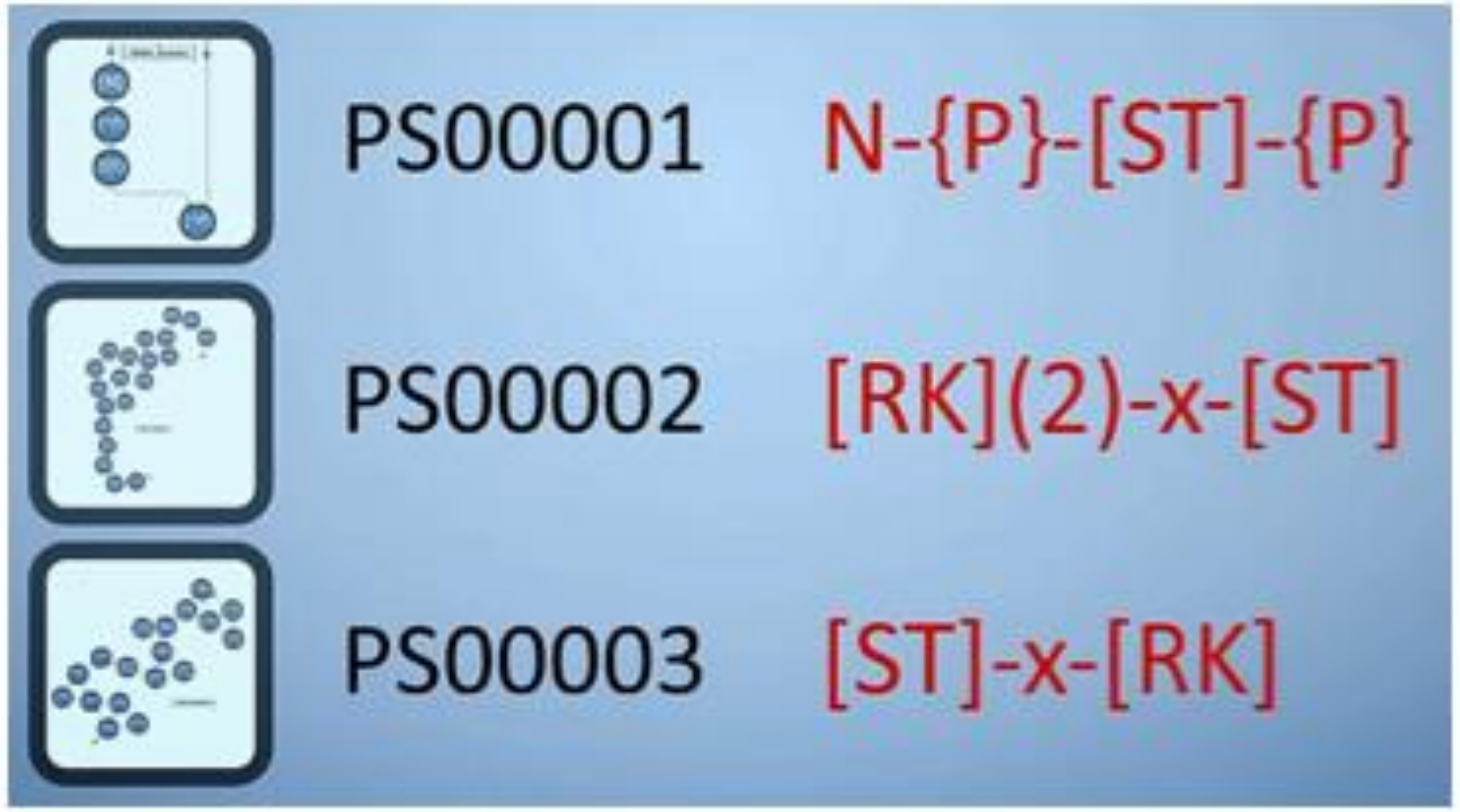
Automata is a Multiple Instruction – Single Data (MISD) processor

- Non-von Neumann architecture evaluates streaming data against ***all*** instructions in parallel
- Enables **deep analysis** of data streams containing **spatial** and **temporal** information
- Complexity of expressions (instructions) has ***no impact*** on execution time

Parallel Programming, Automata Style

Automata are discrete patterns (graphs) that are "placed" into the programmable fabric of the chip

- A single chip can be configured with 1000's of patterns (automata)
- Every automaton evaluates each input symbol on every clock cycle



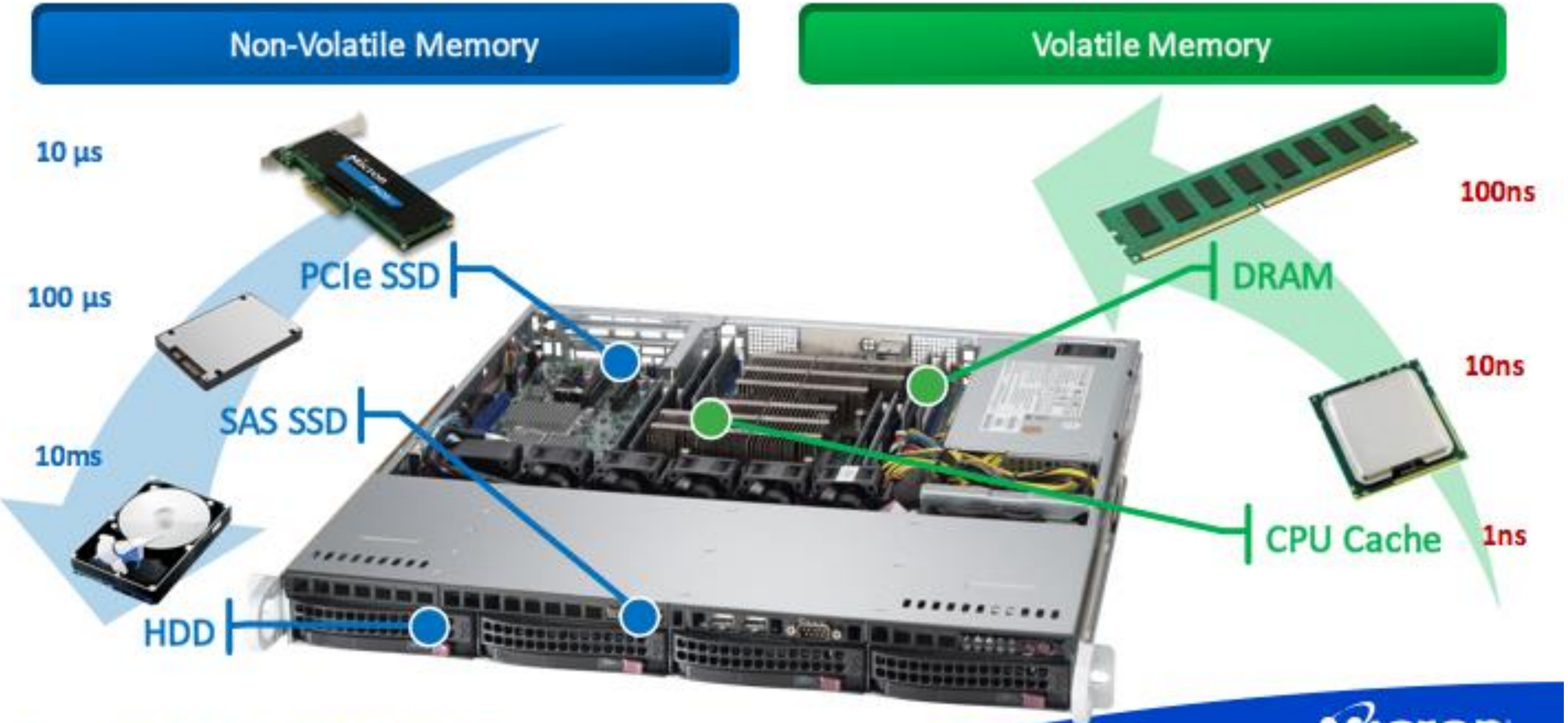
What must the programmer do in order to execute the Automata in parallel?

- *Each automaton is a discrete pattern, no manipulation of data required*
- *Each state transition is fully resolved on each clock cycle by design*
- *Correct operation is guaranteed by design*

Parallel operation is intrinsic to the design – no special skills needed to achieve high levels of parallelism!

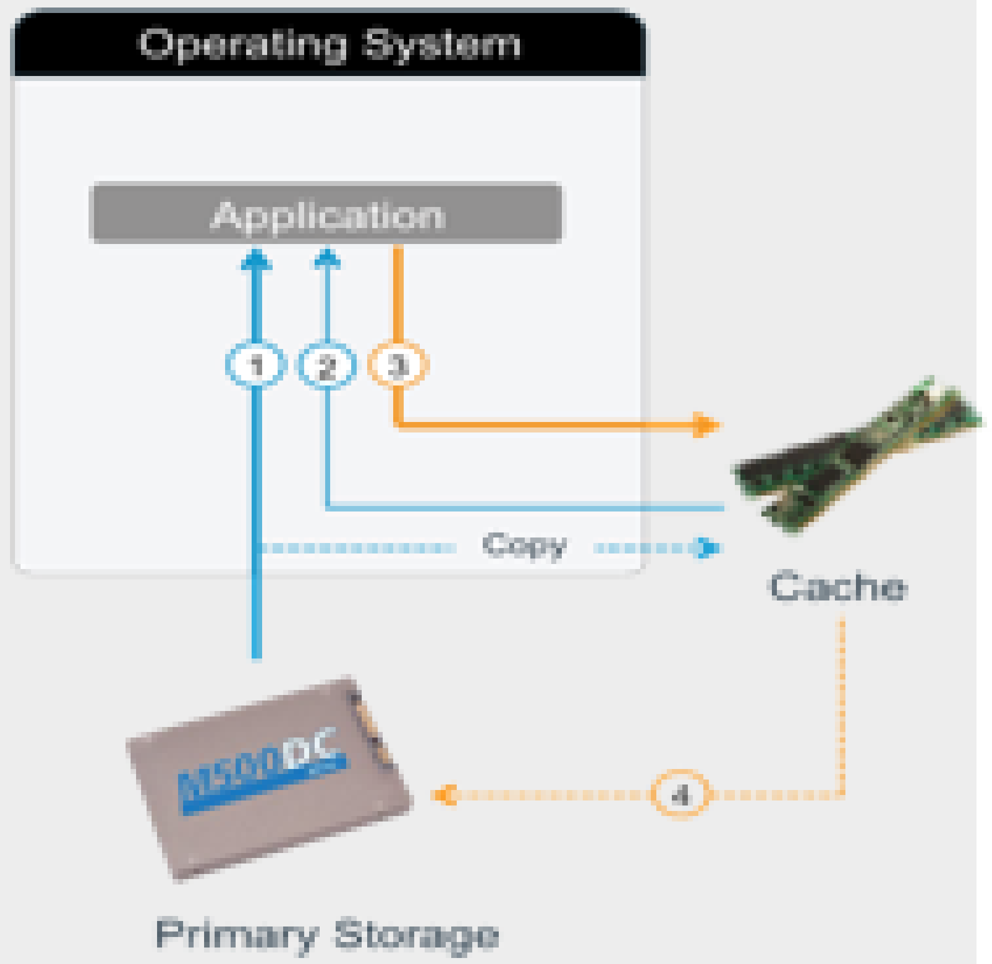
The Challenge: Nonvolatile Memory Latency

- As CPU technology scales, memory IO creates significant performance bottlenecks
- Huge latency gap in memory hierarchy between volatile and non-volatile technologies
- Latency gap widens with the introduction of DDR4



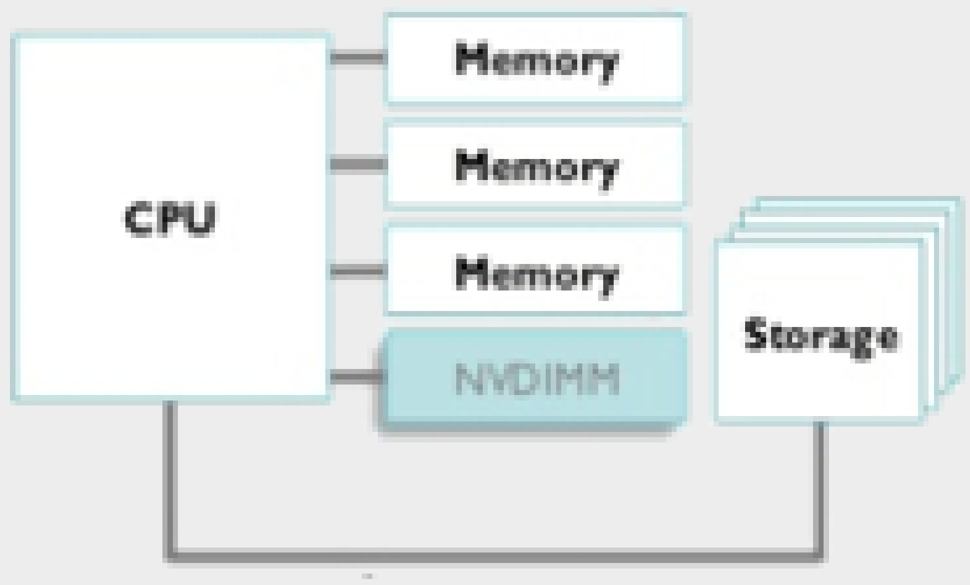
Use Cases and Persistent Variables

Case #1: Write Caching For MLC SSDs



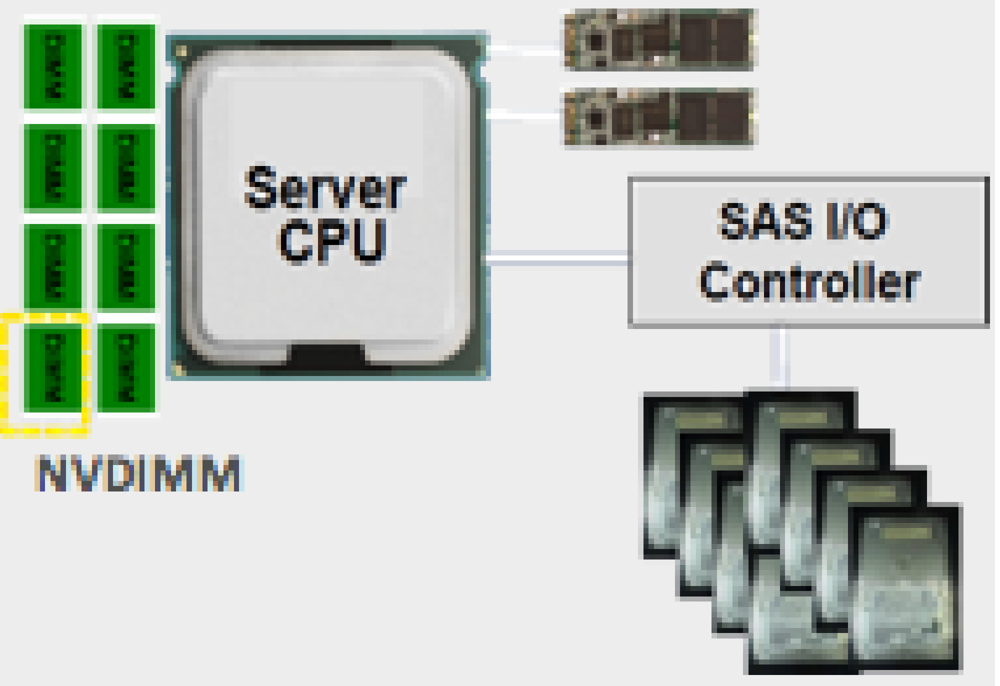
- Extends life of SSD and improves performance for write-intensive apps

Case #2: Low Write Latency Persistent Storage



- >100x lower write latency versus PCIe SSD w/ unlimited endurance

Case #3: Unified Open Software-Defined Server RAID



- Scalable unified RAID performance for SSDs and HDDs

Persistent variables: Metadata, Checkpoint State, Host Caching, RAMDisk, RAID Compute, Write Buffer, SSD Mapping, Journaling, Logging

Intel Modern Code Developer Challenge

The Challenge - Speedup Brain Development Simulation Code

- Original code is 14000 lines of Java
- Recoded in C++
- CERN openlab provided a summer student to start this task
- Intel provided tools and hardware
- A 500 line kernel from this program was used for the Challenge
- This kernel took 45 hours to run with the target set of parameters

The Prizes

- 1 Grand Prize: CERN openlab fellowship
- 3 First Prizes: visit to CERN
- 3 Second Prizes: visit to SC'16

Contestant Engagement

- 17000 students reached
- 2077 students registered for the challenge
    - 130 universities
    - 19 countries
- Over 1200 code downloads
- 1000 students accessed free training

Grand Prize Winner

Mathieu Gravey

Alès School of Engineering

France

- Original code C/C++ running on single core single thread Xeon Phi 7120A
- Final optimized code runs on Xeon Phi 7120A taking advantage of all cores and threads

From 45 hours to…..

8 minutes
24 seconds

320x Increase

Mathieu’s Optimisations

- Change from AoS to SoA to allow vectorisation and improved cache layout
- Custom memory allocator, reuse memory for many small memory allocations
- Use OpenMP for parallelisation over all Xeon-Phi cores
- Use icc Cilk+ scatter/gather intrinsics

Code Modernisation Can Payoff Big Time

Idea: Create CERN Modern Code Developer Challenge

- Find critical pieces of code in CERN programs
- Put them up for the acceleration challenge
- Keep running scores of fastest times to create competition
- Allow students to refine their submissions till end of challenge
- Thinks of some nice prizes
- Also a perfect recruitment tool ;-)

Conclusions

- CERN openlab, a science – industry partnership to drive R&D and innovation
- A number of very interesting projects underway, with a lot of potential
- Some technologies will change the way programs are written
  - New languages, memory, disc, network and CPU technologies
- Very interesting times, indeed

CERN openlab

EXECUTIVE CONTACT

Alberto Di Meglio, CERN openlab Head
alberto.di.meglio@cern.ch

TECHNICAL CONTACTS

Maria Girone, CERN openlab Chief Technology Officer
maria.girone@cern.ch

Fons Rademakers, CERN openlab Chief Research Officer
fons.rademakers@cern.ch

COMMUNICATION CONTACT

Andrew Purcell, CERN openlab Communications Officer
andrew.purcell@cern.ch

ADMIN CONTACT

Kristina Gunne, CERN openlab Administration Officer
kristina.gunne@cern.ch