

Center for Computational Sciences, University of Tsukuba www.ccs.tsukuba.ac.jp

High-Performance Computing Research





- **Problems:** In RPC model,
- A master and workers communicate one-by-one
- Master's bandwidth becomes a bottleneck
- Approach:
- Decouple data transfer from RPC layer

P2P overlay network for TCP programming with UDP hole punching



Our system provides a direct communication method through NATs with UDP hole punching.

- High scalability by distributed super nodes
- High connectivity with UDP hole punching
- Reliability on UDP/IP with a dedicated library
- Larger throughput than relay method
- Easy programming as like as TCP/IP



Make use of various data management layers: Broadcast-tree, BitTorrent, Gfarm etc. in order to optimize communication

Master Program	Worker's IDL
int main(void){	Define MyProcedure(int IN parameter){
 OmstPutData("MyData", data, OMST_INT * DATALEN); for(i = 0; i < n; i++){ req[i] = OmniRpcCallAsync("MyProcedure", i);	 OmstGetData("MyData", data, OMST_INT * DATALEN); /* Program in C */
} OmniRpcWaitAll(n, req);	}
}	



High Performance Data Analysis for Particle Physics using the Gfarm File System

Challenge

- 28 TB Gfarm file system
- HPC Storage Challenge Finalist 1100 compute nodes connected with GbE
 - No special hardware (No SAN)
- Scalable I/O bandwidth up to 1100 nodes
 - 55 GB/sec (50 MB/sec x # nodes)
- Search for the Direct CP asymmetry in b -> s γ decays
 - from 28 TB of real data taken at the KEKB B factory



A High-Performance FFT Library

FFTE is a Fortran subroutine library for computing the Fast Fourier Transform (FFT) in one or more dimensions. The API of FFTE is similar to sequential SGI SCSL or Intel MKL routines.

Features

FFTE:

- **CHPC Challenge benchmark** High speed. (Supports SSE2/SSE3)
- Complex and mixed-radix transforms
- Parallel transforms: Shared / Distributed memory parallel computers

Gfarm file system

Massively parallel particle physics data analysis using 1100 compute nodes



Distributed file system that federates local disks of 1100 nodes

New B Factory Computer System

TOF Cou

Belle Detector

Scalable I/O bandwidth by exploiting local I/O



(OpenMP, MPI and OpenMP+MPI)

- High portability: Fortran77 + OpenMP + MPI
- FFTE's 1-D parallel FFT routine has been incorporated into the HPC Challenge (HPCC) benchmark

