Data Management Challenges

Martin Merck

SAC Review 2009

Madison

Mai 20th-21th 2009

Agenda

SAC Review
Madison
Mai 20th-21th
2009

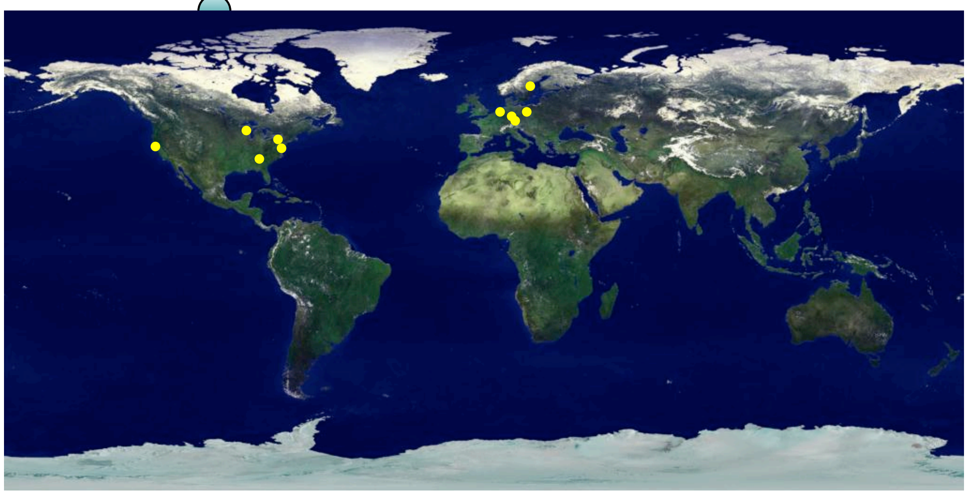
- Current status
- Projected data and computing needs
- Data Storage & Transfer
- Computing Resources
- Data production processing
- Simulation Production
- Outlook

Current status

Available computing resources

SAC Review
Madison
Mai 20th-21th
2009

- Datacenter at Madison
  - Central data repository for IceCube
  - Central compute cluster
  - Access and user accounts for the full collaboration
  - Central infrastructure services (DB, WWW, Mail, etc.)



- Collaboration contributed computing resources
  - Computing resources for MC production
  - Temporary storage for MC datasets
  - Secondary data storage for European access

Current status

Logical view of the IceCube Datacenter

SAC Review
Madison
Mai 20th-21th
2009

IceCube Computer Technology Overview

Network 222 Data Center

Data Warehouse

Experimental, Simulation, and Analysis Data

442 TB SAN
12 Servers
30 RAID Arrays

NPX 2 Cluster

256 Opteron Cores
64 Servers
PBS queue

Secondary Storage

FileTek HSM 600TB
Near-line tape based filesystem

Backups

QualStar Tape Library
400TB Onsite tape media
Offsite copy => Chamberlin Hall

Infrastructure

18 Racks — UPS
158 Hosts — KVM
34 Arrays — ILO
45 kw power (60 kw Capacity) — AC

Access

Shell login, data rsync, grid FTP
6 Hosts

General Use CPU

cobalt/guppy/cobalt64
8 Hosts

Network Services

Web, Printing, E-mail, DBs, Tracker, Wiki
15 Hosts

Infrastructure Services

LDAP, Monitoring, DHCP, Syslog, DNS, RHN
12 Hosts

File Services

User Data (47TB) Home Dirs

Networking & Security

Switches, Firewalls, VPNs, WAPs

Projected data and computing needs

Increase in data rate

SAC Review
Madison
Mai 20^{th}-21^{th}
2009

- Raw data files have increased considerably in size
  - Adding SLC readouts increases raw data by a factor of > 2
  - IC40 ~ 325 GB / day
  - IC59 ~ 660 GB / day
  - IC80 ~ 1000 GB /day
  - Compression is useful for limiting satellite transfer needs
  - SLC hits can't be compressed after feature extraction and increase processed data volume
- Need to increase storage for analysis datasets and simulation beyond original plan
  - Some increase from contingency
  - Continuous increase due to technology evolution and standard replacement cycle
- Simulation resources increased due to good GRID usage
  - More MC data to store
  - Use filtering also on MC data (only partly a reduction)

Projected data and computing needs

Increase in data rate

SAC Review
Madison
Mai 20^{th}-21^{th}
2009

	IC40	IC59	IC86
Trigger rate	1200 Hz	1800 Hz	2700 Hz
Raw data uncompressed	630 GB/day	1300 GB/day	2000 GB/day
Raw data compressed	325 GB/day	660 GB/day	1000 GB/day
Filtered data	35 GB/day	55 GB/day	82 GB/day
Level1 data	50 GB/day	70 GB/day	105 GB/day
Level2 data	85 GB/day	200 GB/day	300 GB/day
Storage 1 Year	60 TB	120 TB	180 TB
Storage 5 years Level2			550 TB
Analysis data			200 TB
Total exp data			930 TB
Simulation data			~ 1.1 TB

Projected data and computing needs

Increase in data rate

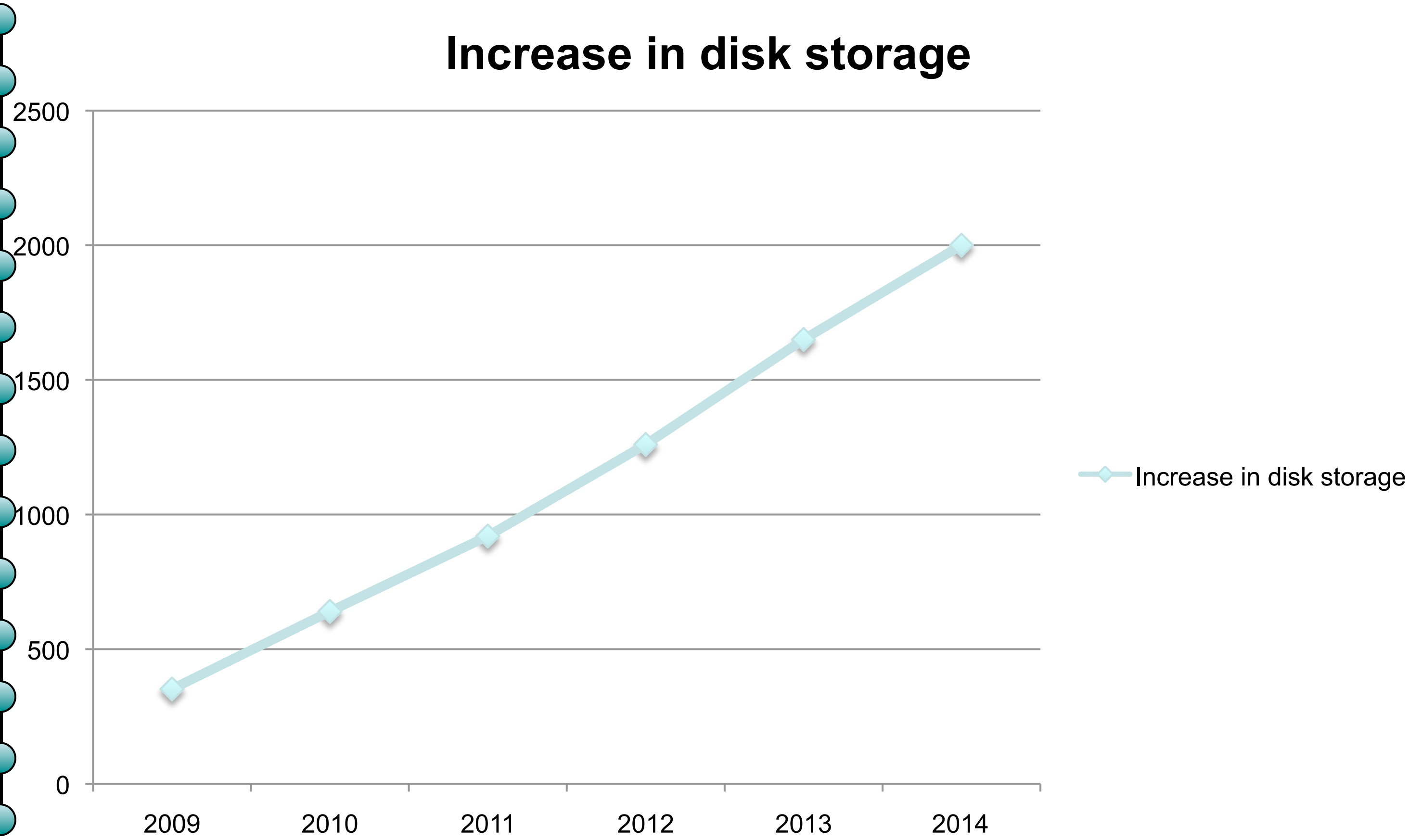
SAC Review
Madison
Mai 20th-21th
2009

	Tape storage	Data accessible from online disk			
	Experimental raw data [TB]	Experiment [TB]	Simulation {TB]	Analysis [TB[	Total [TB]
2010	900	250	300	90	**640**
2011	1400	350	450	120	**920**
2012	1900	450	650	160	**1260**
2013	2400	550	850	250	**1650**
2014	2900	650	1000	350	**2000**

Projected data and computing needs

Increase in data rate

SAC Review Madison Mai 20th-21th 2009



Computing Resources

HPC compute cluster

SAC Review Madison Mai 20th-21th 2009

- HPC compute cluster
  - 64 x 4 core Servers
  - 4 x 16 core blades
  - 320 Opteron cores (2.3 – 2.4 GHz)
- High memory per core
  - 4 GB
- Planned extension
  - 32 x 8 core Servers
  - 256 Nethalem cores
- 3 year replacement cycle
- keep 2 generations of clusters for processing and user analysis

Core Software Systems

Challenges for the M&O phase

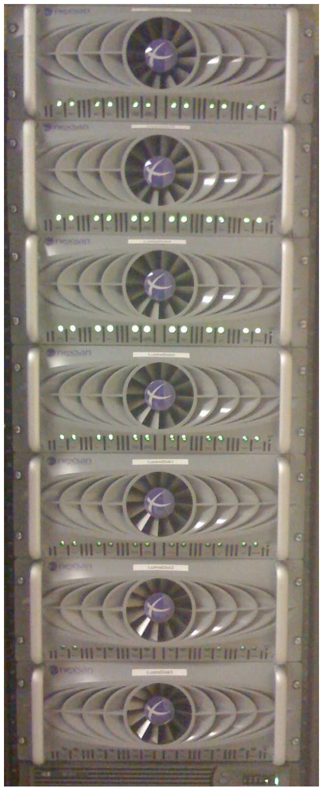
SAC Review
Madison
Mai 20^{th}-21^{th}
2009

- Maintenance of basic software framework
- Adaption to new operating systems and compilers
- Support of new resources (GRIDs, Compute clouds, etc.)
- Training and support of users
- Maintenance of software repository for user contributed software
- Maintain and operate central database and mirrors around the world

Data Storage & Transfer

The central data repository

SAC Review
Madison
Mai 20th-21th
2009



- 404 TB SAN
  - 15 high capacity units
  - 76% used
  - exp/sim/ana
- 45 TB user space
- Tape library with 1.3 PByte tape capacity
- Needs continuous expansion by replacement with newest technology in a 4 year cycle (assuming 50% disk capacity growth per year)

Simulation Production

Available GRID resources

SAC Review
Madison
Mai 20th-21th
2009

- Produce MonteCarlo dataset for the collaboration
  - Signal generation
    - ν_μ, ν_e, ν_τ
    - WIMPS
    - Monopoles
  - Backgroung generation
    - Muons from cosmic ray air showers
      - Need at least 4000 cores for real time simulation
- Distributed production leveraging GRID resources

Simulation Production

Available GRID resources

SAC Review
Madison
Mai 20th-21th
2009

	Institute	Cores		Disk Space	Farm Type	FTE
		Guaranteed	Available	TB		In-kind
US	UW GLOW (US)	140	800	190	Grid	
	UW NPX2 (US)	100	256		Batch	
	UW CHTC (US)	100	700		Batch	
	UMD (US)	140	278	5	Batch	0.40
	PSU (US)	100	560		Batch	0.40
	LBNL PDSF (US)	50	700	2	Batch	0.20
	UDEL (US)	40	136	50	Batch	0.40
	LONI (US)	200			Batch	0.20
Germany	Aachen (DE)	90	200	15	Grid	0.40
	Dortmund (DE)	150	300	30	Grid	0.40
	Dortmund (DE)	100		20	Batch	
	Mainz (DE)	230	400	26	Grid	0.40
	Wuppertal (DE)	64	128	17	Grid	0.40
	Wuppertal (DE)	150		30	Batch	
	DESY (DE)	400	700	100	Batch	0.40
	DESY (DE)	100	200	20	Grid	
Sweden	SweGrid (SE)	100	400		Grid	0.20
Belgium	Brussels (BE)					0.20
	Totals	**2,254**	**5,758**	**505**		**4.00**

Data production processing

SAC Review
Madison
Mai 20th-21th
2009

- Steps of central processing
  - Level1: unpacking an first guess reconstructions
  - Level2: high level reconstructions
- Optimize use of limited resources
  - Provide high levle reconstructions of tracks, cascades, air showers, EHE events, etc for all types of analysis
- Provide analysis ready data in the shortest time possible
- Fast feedback to detector operations on analysis level data quality

Data production processing

IC40 reconstructions

Reconstruction	Percentage of events	CPU time / day
Linefit	100%	
SPEFitSingle	100 %	1.0 day
SPEFit32	66 %	21.3 days
SPEFit32Paraboloid	35 %	4.8 days
SPEFit32Umbrella	66 %	0.6 days
SPEFit32Bayesian	9 %	3.0 days
SPEFit16Split (4 times)	9 %	3.3 days
SPEFit16SplitBayesian (4 times)	9 %	3.8 days
MPEFit	66 %	1.1 days
MPEFitParaboloid	35 %	8.6 days
SPEFit32PhotorecTrack	35 %	1.6 days
SPEFit32PhotorecEnergy	66 %	0.1 days
MPEFitPhotorecEnergy	35 %	0.1 days
SPE32FitMuE	66 %	0.05 days
MPEFitMuE	66 %	0.05 days
CombinedSPEFitSingle	22 %	0.5 days
CombinedSPEFitParaboloid	22 %	1.7 days
JamsSPEFit32	22 %	16 days
CascadeLlh (4 recos)	30 %	1.2 days
CascadeLlhAmOnly (4 recos)	10 %	0.3 days
CascadeLlhCombined (2 recos)	10 %	0.15 days
EHE (3 recos)	2 %	
Shower front (3 recos)	5 %	0.02 days

IceCube

Summary

SAC Review
Madison
Mai 20th-21th
2009

- Data transfer and storage is impacted by higher then expected data volumes.
  - Importance of SLC hits has significantly increased with the extension to low energy physics
- Simulation and data production centrally managed and providing all datasets for scientific analysis
  - Relying heavily on distributed and GRID resources to achieve the goals of MC production
  - Centralized data processing leveraging the high speed access to the data in the Madison datacenter
- Urgent need to expand storage and computing
  - Faster processing needed
  - More reserve for unexpected reprocessing
  - Better priorization inside the collaboration to maximize physics with the available resources