

# 1 Contents

2	<b>Contents</b>	<b>i</b>
3	<b>List of Figures</b>	<b>v</b>
4	<b>List of Tables</b>	<b>vii</b>
5	<b>I Overview</b>	<b>2</b>
6	<b>1 Introduction ( Schellman/Muether/Junk/Pennacchio - SN added 2/19)</b>	<b>3</b>
7	1.1 Mission Statement . . . . .	4
8	1.2 Introduction ( Schellman - draft based on CHEP paper) . . . . .	4
9	1.3 ProtoDUNE Tests at CERN ( Schellman/Pennacchio-draft) . . . . .	6
10	1.4 Far Detector ( Schellma-draft) . . . . .	15
11	1.5 Near Detector ( Junk/Muether - needs update) . . . . .	18
12	1.6 Relation of Physics Goals to Offline Computing Challenges ( Schellman-draft) . . . . .	21
13	1.7 Summary of Challenges ( Schellman-draft) . . . . .	22
14	<b>2 Computing Organization - updated 3/1</b>	<b>24</b>
15	2.1 Internal organization . . . . .	26
16	2.2 Funding Sources for Computing Development . . . . .	26
17	2.3 Computing Contributions Board ( HMS 3/2 moved from separate chapter) . . . . .	26
18	<b>II Single interaction scale</b>	<b>28</b>
19	<b>3 Data Processing Considerations and Challenges</b>	<b>29</b>
20	3.1 Introduction ( Schellman - draft) . . . . .	29
21	3.2 Data Acquisition and Storage( Schellman - draft) . . . . .	29
22	3.3 Simulation Chain ( Junk and Schellman - draft) . . . . .	31
23	3.4 Far Detector and protoDUNE Detector Simulations . . . . .	34
24	3.5 Reconstruction ( Junk, Muether and Schellman-draft) . . . . .	42
25	3.6 Visualization . . . . .	47
26	3.7 Event Classification . . . . .	53
27	3.8 Analysis of Reduced Data Samples . . . . .	56
28	3.9 Characteristics of Large-Scale Processing Tasks . . . . .	61

1	<b>4 Data Formats</b>	<b>63</b>
2	4.1 Introduction ( <i>Schellman - draft</i> ) . . . . .	63
3	4.2 Data tiers ( <i>Schellman - add a discussion here</i> ) . . . . .	63
4	4.3 I/O formats ( <i>Schellman/Bashyal - draft</i> ) . . . . .	64
5	<b>5 Frameworks ( <i>Norman, Laycock, Muether</i>)</b>	<b>66</b>
6	5.1 Defining a Framework . . . . .	66
7	5.2 Current status ( <i>Junk/Muether - draft</i> ) . . . . .	69
8	5.3 Framework Requirements ( <i>Laycock/Norman - needs update</i> ) . . . . .	69
9	5.4 Timeline ( <i>Norman/Laycock needs a lot more</i> ) . . . . .	80
10	<b>6 Databases ( <i>Buchanan and Laycock - draft</i>)</b>	<b>82</b>
11	6.1 Introduction ( <i>Buchanan and Laycock - draft</i> ) . . . . .	82
12	6.2 Conditions Database ( <i>Buchanan and Laycock - draft</i> ) . . . . .	84
13	6.3 Run Configuration Database ( <i>Buchanan and Laycock - draft</i> ) . . . . .	86
14	6.4 Data Quality and Monitoring Database ( <i>Buchanan and Laycock - draft</i> ) . . . . .	87
15	6.5 Offline Calibration Database ( <i>Buchanan and Laycock - draft</i> ) . . . . .	87
16	6.6 Slow Control Database ( <i>Buchanan/Laycock - draft</i> ) . . . . .	87
17	6.7 Beam Conditions Database - IFBeam ( <i>Buchanan and Laycock - draft</i> ) . . . . .	88
18	6.8 Hardware Database ( <i>Buchanan and Laycock - draft</i> ) . . . . .	88
19	6.9 Service and Maintenance ( <i>Buchanan and Laycock - draft</i> ) . . . . .	89
20	6.10 Development Plans ( <i>Buchanan and Laycock - draft</i> ) . . . . .	90
21	<b>III Global Computing Model</b>	<b>92</b>
22	<b>7 Data and Processing Volume Estimates ( <i>3/4 HMS numbers updated</i>)</b>	<b>93</b>
23	7.1 Introduction ( <i>Schellman</i> ).....	93
24	7.2 ProtoDUNE Experience( <i>Schellman - draft</i> ).....	94
25	7.3 Far Detector Data Volume Estimates ( <i>Schellman - needs redo with new numbers</i> ) .....	95
26	7.4 Near Detector Data Volumes ( <i>HMS 3/5 I think we decided these were ok.</i> ).....	97
27	7.5 Data Retention Assumptions ( <i>Schellman - draft</i> ) .....	99
28	7.6 Model Studies for Data and CPU Needs ( <i>Schellman 3/4 updated - draft</i> ).....	99
29	<b>8 Overview of Computing Model ( <i>McNab - draft</i>)</b>	<b>102</b>
30	8.1 Introduction ( <i>Schellman/Herner - draft, update tables</i> ) . . . . .	102
31	8.2 Current Performance ( <i>Schellman/Herner - draft</i> ) . . . . .	102
32	8.3 Sites and Services ( <i>McNab - draft</i> ) . . . . .	108
33	8.4 Sites, Federations, and Countries( <i>McNab - draft</i> ) . . . . .	109
34	8.5 Types of Service( <i>McNab - draft</i> ) . . . . .	109
35	8.6 Requirements for Computing Services ( <i>McNab - needs work</i> ) . . . . .	110
36	<b>9 Data Management ( <i>Timm, Mandrichenko - draft</i>)</b>	<b>111</b>
37	9.1 Introduction ( <i>Timm - draft</i> ) . . . . .	111
38	9.2 Existing SAM Data Management System . . . . .	111
39	9.3 Requirements for Replacing SAM Functionality ( <i>Schellman/Mandrichencko - draft</i> ) . .	112
40	9.4 Data Ingest Manager ( <i>Timm - draft</i> ) . . . . .	116
41	9.5 Rucio Replica Manager ( <i>Timm - draft</i> ) . . . . .	116

1	9.6	Metadata Catalog ( <i>Mandrighenko, Schellman - draft - too long</i> ).....	117
2	9.7	Metacat Data Model ( <i>Timm/Mandrighenko - draft</i> ) .....	118
3	9.8	Data Dispatcher ( <i>Mandrighenko/Timm - more needed</i> ) .....	121
4	9.9	Tools ( <i>Timm - needs more</i> ).....	121
5	<b>10</b>	<b>Networking ( <i>Mike Kirby and Peter Clarke - in progress</i>)</b>	<b>122</b>
6	<b>11</b>	<b>Workflow Management ( <i>McNab</i>)</b>	<b>126</b>
7	11.1	Introduction ( <i>McNab - draft</i> ) .....	126
8	11.2	Existing production submission infrastructure ( <i>Herner - draft</i> ) .....	126
9	11.3	Requirements for replacing SAM Functionality .....	129
10	11.4	Request lifecycle ( <i>McNab - draft</i> ).....	129
11	11.5	Grid workload systems.....	130
12	11.6	Generic job factory ( <i>McNab - draft</i> ) .....	130
13	11.7	Workflow Database ( <i>McNab - draft</i> ) .....	131
14	11.8	Information Collector.....	131
15	11.9	Request Builder .....	131
16	11.10	Archiver.....	131
17	11.11	Workflow Allocator .....	131
18	11.12	User commands ( <i>McNab - draft</i> ).....	132
19	11.13	Workflow Dashboard .....	132
20	11.14	Implementation Plan ( <i>McNab/Timm/Herner needed</i> ) .....	132
21	<b>IV</b>	<b>Integration and Evolution</b>	<b>133</b>
22	<b>12</b>	<b>Services Overview ( <i>Timm - needs more</i>)</b>	<b>134</b>
23	12.1	Introduction.....	134
24	12.2	Host Lab Provided Services .....	134
25	12.3	Remote Site Provided Services .....	135
26	12.4	Cloud Hosted Services.....	135
27	<b>13</b>	<b>Information Systems ( <i>McNab - needed</i>)</b>	<b>136</b>
28	<b>14</b>	<b>Monitoring ( <i>Andrew McNab, Steve Timm, Raja Nandakumar, Jon Hays - in progress</i>)</b>	<b>137</b>
29	14.1	Tools.....	137
30	14.2	Summary of Assumptions and Resource Requirements.....	139
31	<b>15</b>	<b>Authentication and Authorization for Distributed Computing( <i>Timm - draft</i>)</b>	<b>140</b>
32	15.1	Obtaining Access to DUNE Computing.....	140
33	15.2	Current State of Authentication and Authorization.....	140
34	15.3	Planned Changes to Authentication Currently Under Way .....	141
35	15.4	Requirements for Authentication/Authorization Going Forward .....	141
36	<b>16</b>	<b>Code Management ( <i>Tom Junk, Mathew Muether, David Adams - draft</i>)</b>	<b>143</b>
37	16.1	Liquid Argon TPC Code Management ( <i>Junk/Calcutt - needs update</i> ).....	143
38	16.2	Near Detector Code Management ( <i>Muether/Cremonisi/Junk needs update</i> ) .....	145
39	16.3	Continuous Integration ( <i>Junk - draft</i> ).....	147

1	<b>17 Training and Documentation</b> ( <i>David, Kirby - draft</i> )	<b>148</b>
2	17.1 Documentation.....	148
3	17.2 Training ( <i>David, DeMuth, Kirby - draft</i> ) .....	150
4	<b>18 Data Lifetimes and Preservation:</b> ( <i>Norman, Clarke, Fuess - needed</i> )	<b>153</b>
5	18.1 Data Management and Preservation .....	153
6	<b>V Conclusions and Resources</b>	<b>155</b>
7	<b>19 Resource Needs Summary</b>	<b>156</b>
8	19.1 Hardware resources . . . . .	156
9	19.2 Personnel needs . . . . .	156
10	19.3 Summary . . . . .	160
11	<b>20 Cooperation strategy</b> ( <i>Kirby, Schellman, McNab</i> )	<b>161</b>
12	20.1 Cooperation . . . . .	161
13	<b>Glossary</b>	<b>163</b>
14	<b>References</b>	<b>176</b>

15