

データ収集システムの実際

J-PARC 高強度全散乱装置(NOVA)の データ収集システムの構築体験から

高エネルギー加速器研究機構 物質構造科学研究所
大下 英敏

達成目標

データ収集システムを構築する人間にとっての
実験規模とそのデータ収集システムの関係を把握して、
自らの実験に必要な(サイズ・仕様の)データ収集
システムの構築ができるようになりましょう。

内容

現在、実際に稼働しているJ-PARC 高強度全散乱装置の
紹介とデータ収集システムの構築体験談

中性子で物質構造解析をするメリット

中性子線は**X線で観察することのできない**構造特性を明らかにする

⇒ 磁気構造、軽元素を含む物質の構造、同程度の原子番号の元素で構成される物質の構造

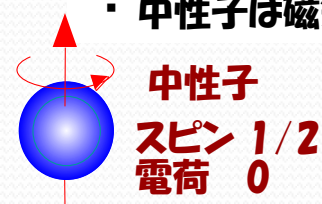
X線に比べると、慢性的に**強度不足**、また、中性子照射施設は限られる

⇒ 中性子線を用いた構造解析とX線を用いた構造解析は**相補的**

・ 中性子は原子核と散乱し、散乱能は原子番号に依存しない

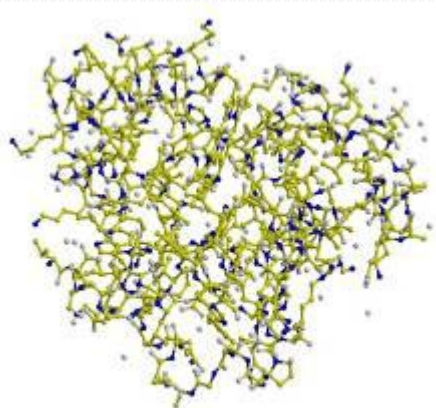
X線は原子核の周りの電子雲と散乱し、散乱能は原子番号に依存する

・ 中性子は磁気モーメントをもつ

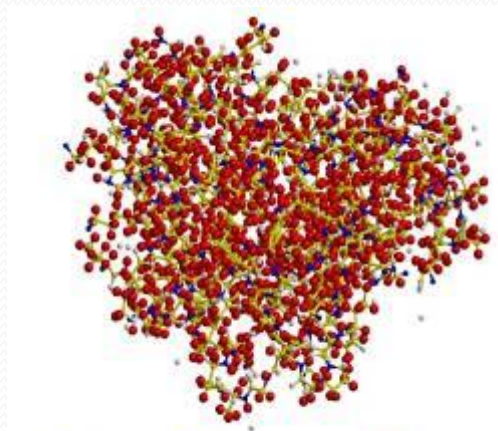


X線で見た構造

中性子で見た構造



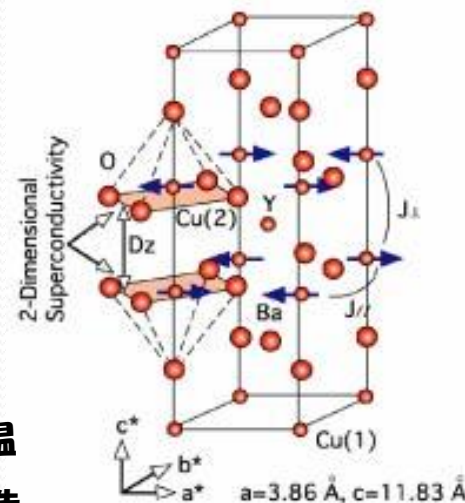
●窒素 ●炭素 ○酸素



●窒素 ●炭素 ○酸素 ●水素

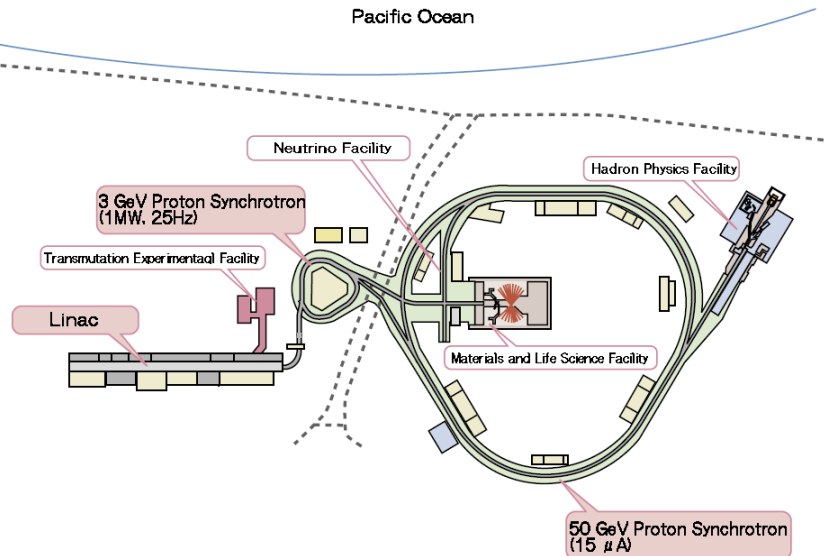
タンパク質 (ミオグロビン) の構造

Superconductivity in $\text{YBa}_2\text{Cu}_3\text{O}_6$



中性子によって
解明された高温
超伝導体の構造

Japan Proton Accelerator Research Complex (J-PARC)



高エネルギー加速器研究機構(KEK)と日本原子力研究開発機構(JAEA)との
共同プロジェクト

3種類の加速器施設

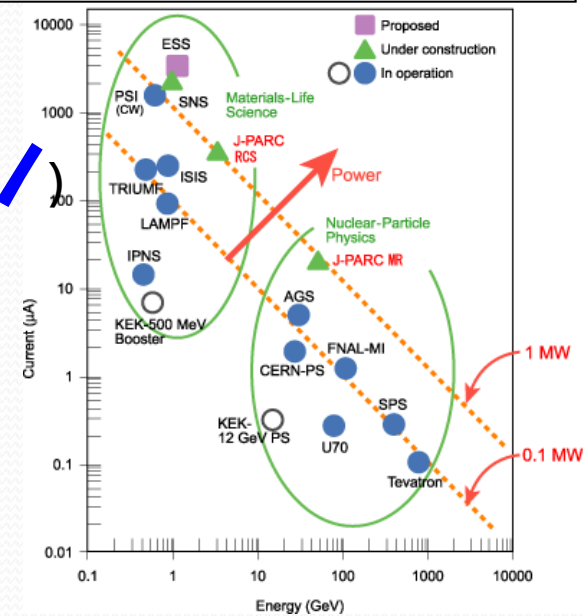
・ 線形加速器・ 3 GeV シンクロトロン・ 50 GeV シンクロトロン

2次ビーム(中性子、 μ 粒子、 K 粒子、ニュートリノ)の利用

J-PARCでは物質構造科学、生命科学、
素粒子・原子核物理学の研究をおこなう

<http://j-parc.jp/Acc/ja/index.html>

The beam power of several facilities



物質・生命科学実験施設 (MLF)



No. 1 Experimental Hall



Proton beams

Hg (mercury) target
+ H₂ moderator

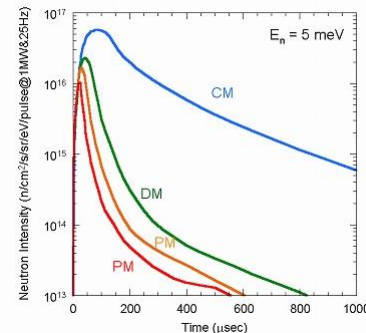
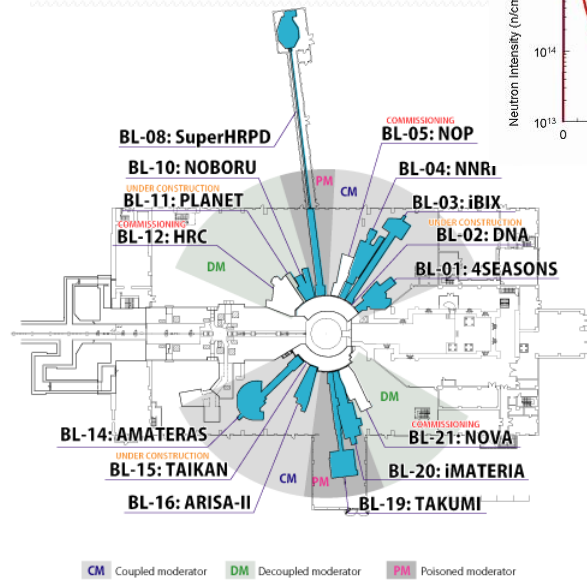
Muon Experimental Area



No. 2 Experimental Hall



NOVA



<http://j-parc.jp/MatLife/ja/index.html>

MLFでは**物質構造科学、生命科学、基礎物理分野**の研究がおこなわれる

– 23 本の中性子ビームラインと4本のミュオンビームライン

MLFは世界最大級の**パルス中性子源** : 5×10^8 n/s @1MW, BL21

– 中性子源: 水銀ターゲット

– 減速材: 超臨界水素

– 3種類のモデレータ構造 (coupled, decoupled, poisoned)

核分裂反応とスポレーション(核破砕)反応

・核分裂反応の例



ウランが熱中性子を吸収し、原子核分裂する。

その際に**2～3個の中性子**を発生、分裂のウラン原子1個当り～200 MeVのエネルギーを放出する。

臨界条件を保つため、放出した中性子の一部は再び核分裂反応に使われる。

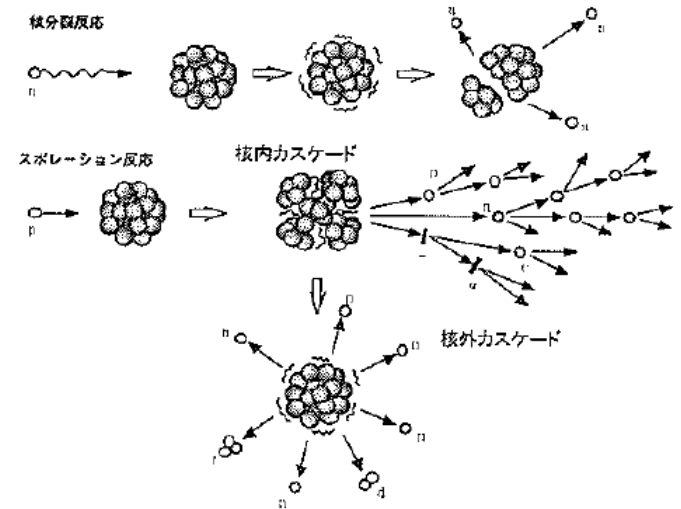


図1 スポレーションと核分裂反応の概略 [1]

・スポレーション反応の例

中、高エネルギー陽子が原子核に衝突し、原子核を破砕する。

放出される中性子の内訳

～10% 核外カスケードによって生成する

～90% 衝突した原子核の励起状態からの蒸発過程で生成する

・発熱量当りの有効中性子数(ビーム実験に使用可能な中性子数)は**スポレーション反応の方が多い。**

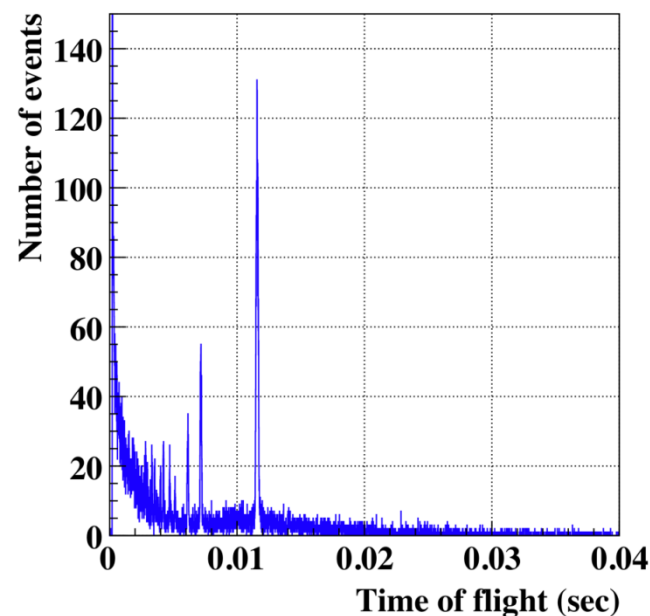
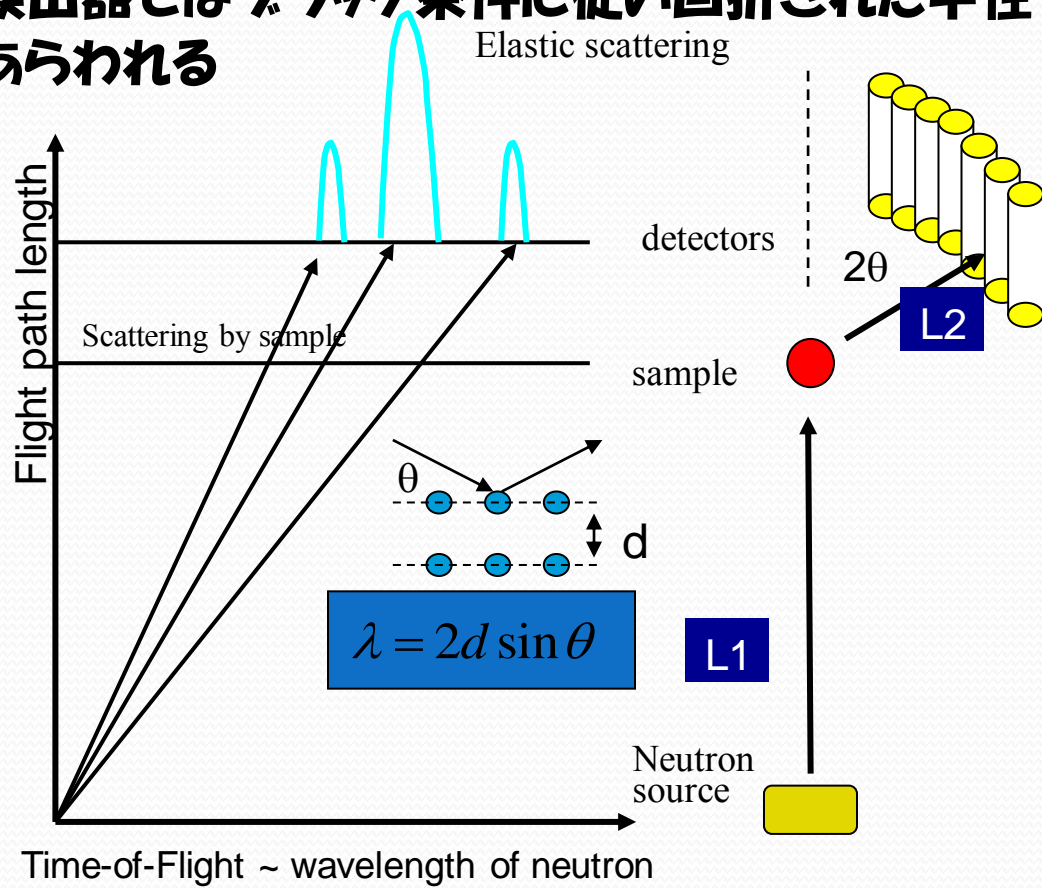
飛行時間法による中性子回折実験

パルス中性子源では、飛行時間法によって**中性子波長弁別**が可能である

⇒ 中性子源から時刻0で放出された中性子は波長(エネルギー)に依存して
試料位置や検出器位置に到達する時間が異なる

飛行時間法はTime-of-Flight (TOF) methodと呼ばれる

検出器ではフラッグ条件に従い回折された中性子による**フラッグピーク**が
あらわれる



NOVA first diffraction data
2009/05/28
Diamond powder @90 deg. detector

NOVAにおける水素貯蔵材料の研究

- NOVAは、NEDOの水素貯蔵材料先端基盤事業の支援を受け、水素燃料電池に応用できる**水素貯蔵材料の研究**をおこなっている
⇒ 水素を含む物質の構造解析は中性子回折実験で可能となる
- 水素貯蔵材料の研究として、**水素吸蔵、放出過程のメカニズム**を観察する
⇒ 過渡現象的な測定は高強度ビーム実験で可能となる
- NOVAでは、結晶、アモルファス、ガラス、液体と構造の小さなものから大きなものまで測定可能なので、金属の水素貯蔵材料から非金属の水素貯蔵材料の研究がおこなえる

NOVAの特長

NOVAでは構造の小さなものから大きなものまで(同時に)測定可能

⇒ **測定可能な $Q (\propto 1/d)$ 範囲が広い**

メソポーラスシリカの原料であるシリカガラスの構造(小さな構造)と規則正しい細孔構造(大きな構造)を同時に測定できる。

しかも**短時間**で。

⇒ かって、2台の分光器で測定したデータがNOVA1台で測定可能。しかも短時間で。

TEM picture of
meso-porous silica
(FSM-16)

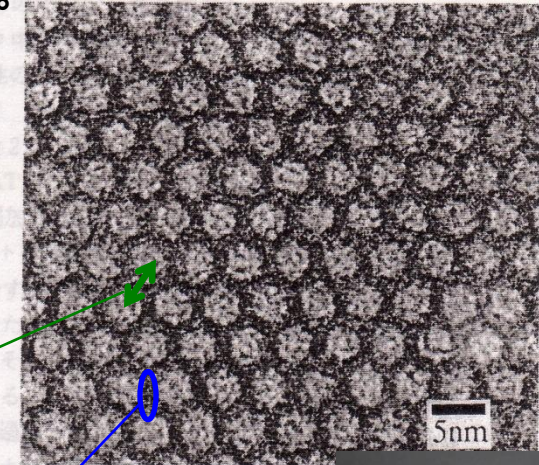
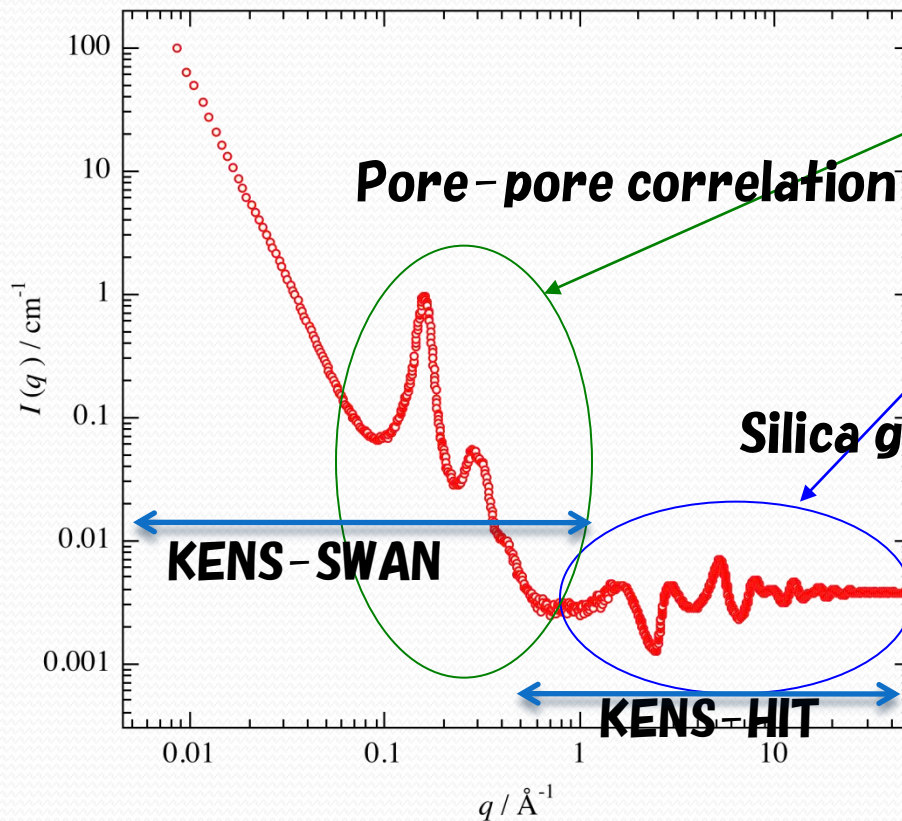
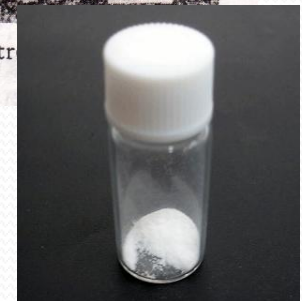


Fig. 1 Transmission electron
FSM-16.



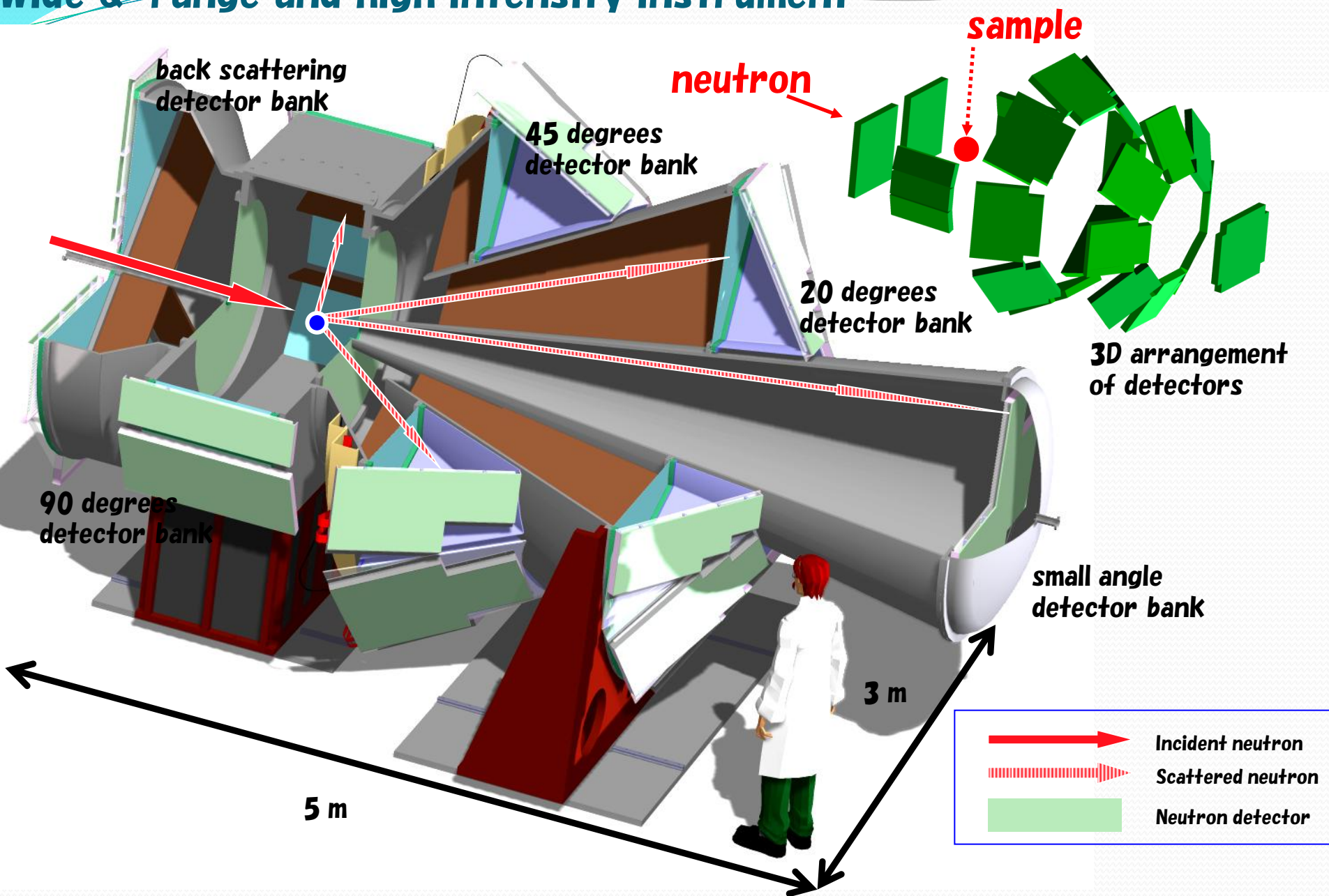
Silica glass structure

Meso-porous silica is:

- Diameter of the pore is about 3nm
- Pores align in hexagonal order

Conceptual design of NOVA

Wide Q-range and high intensity instrument



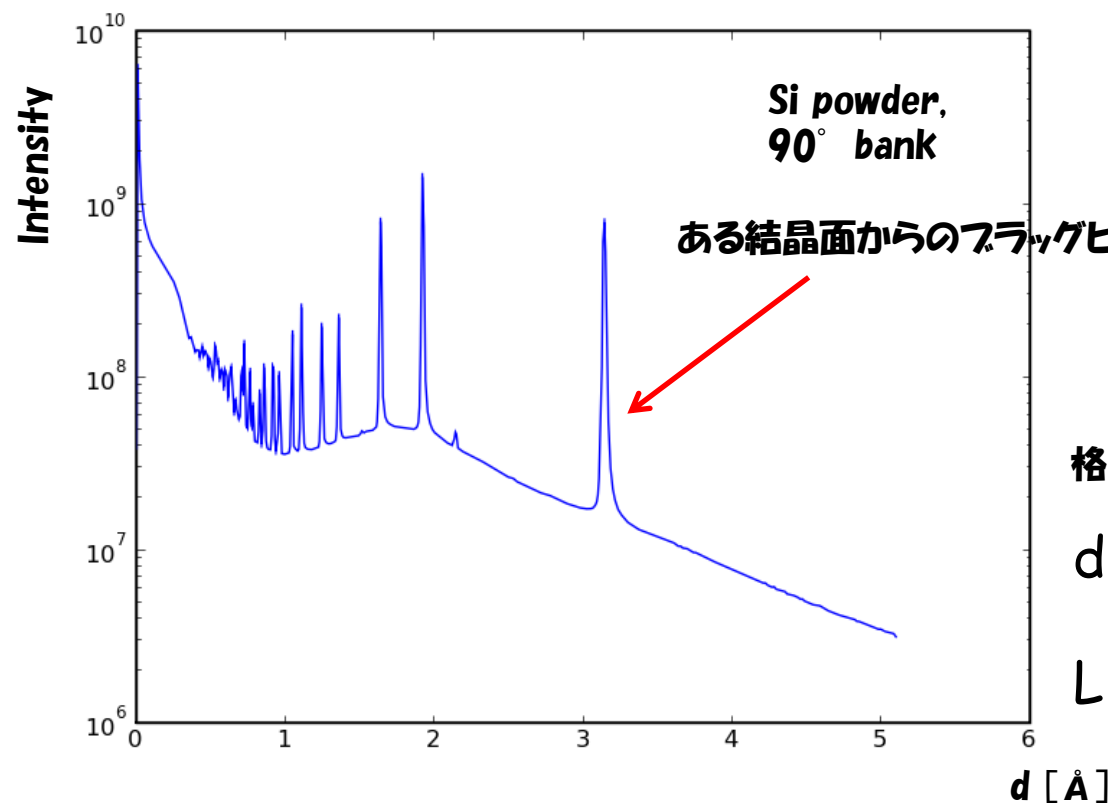
NOVAにおける中性子回折実験(1)

検出器では**試料に依存した強度分布**が得られる

格子定数 d は、TOF測定値と検出器の位置に依存した $L \sin \theta$ を考慮すると求まる

⇒ 試料に固有の強度分布となる

いくつかの補正を加えた強度分布から構造解析をおこなう**(可視化処理)**



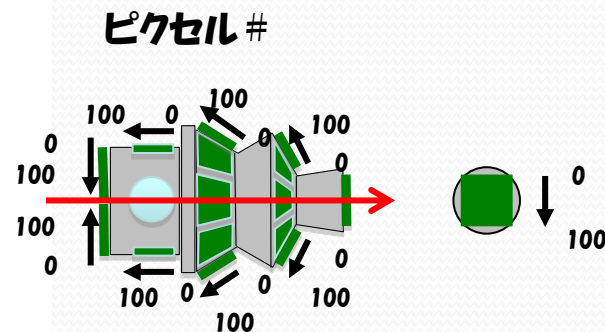
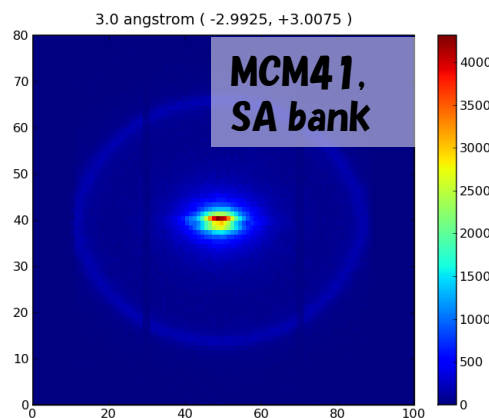
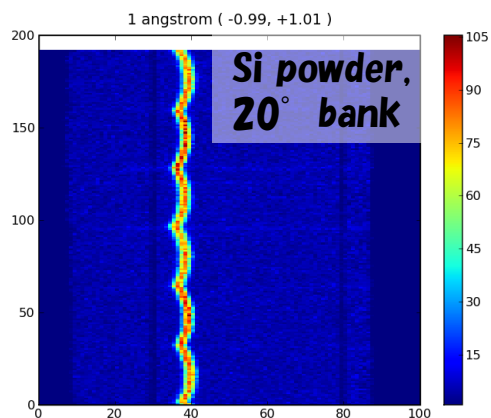
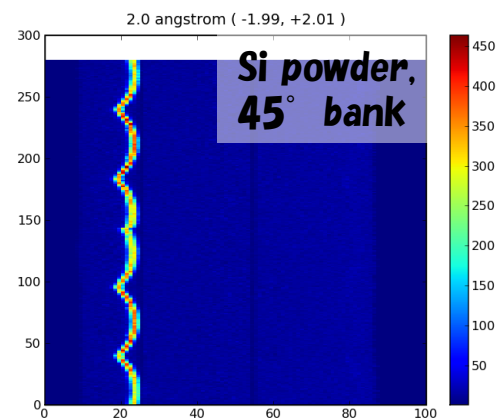
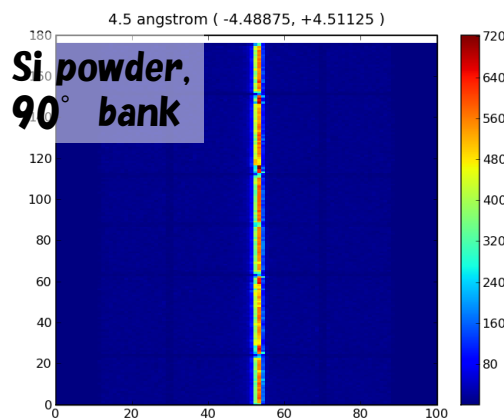
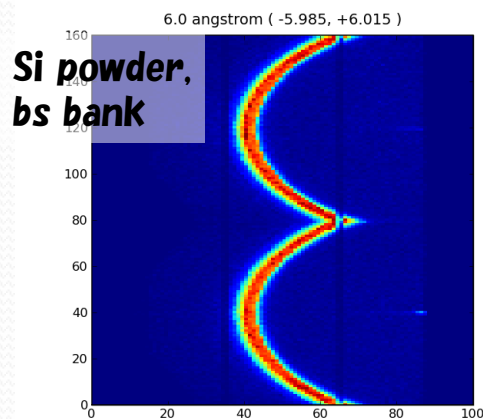
格子定数 d の導出

$$d [\text{\AA}] = 1.978 \times 10^3 \times \frac{t [\text{sec}]}{L [\text{m}] \times \sin \theta}$$

$$L = L_1 + L_2$$

NOVAにおける中性子回折実験(2)

NOVAで**粉末試料**(例えば、Si powderやMCM41)を測定した場合、
各バンクの検出器には波長 λ 、散乱角 2θ に依存した**デバイリング**が現れる



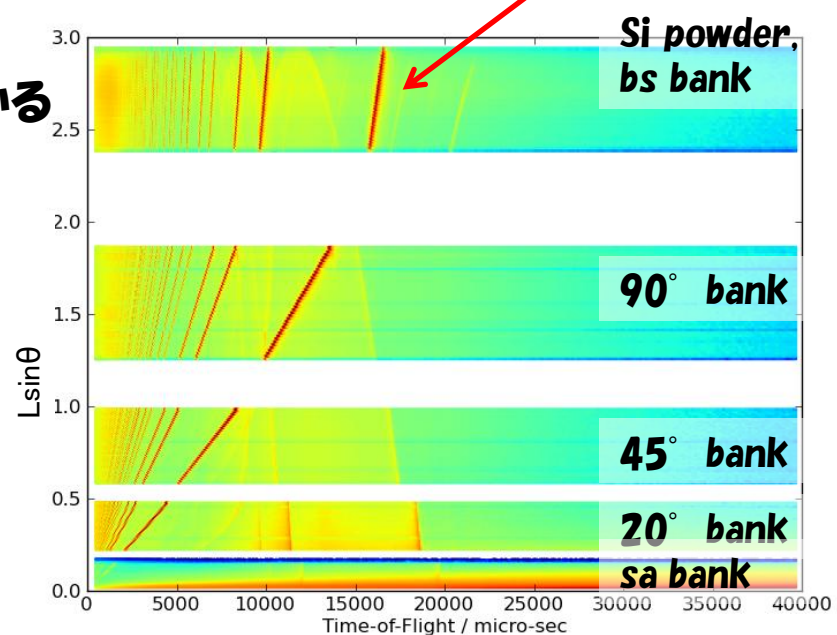
NOVAにおける中性子回折実験(3)

ある結晶面からのフラッグピーク

TOF測定値 - $L \sin \theta$ 分布より検出器バンクの
Q分解能と**測定可能なQ領域**がわかる

Q分解能: high-angleほど高い
測定可能なQ領域: small-angleほど広い

試料の**小さい構造**を観察するには
sa bankを使用し、**大きい構造**を
観察するには**bs bank**を使用する



Detector bank	2θ [deg]	Sample-detector [m]	$\Delta Q/Q$ [%]	Q-range [\AA^{-1}] (d-range [\AA])
Small-angle	0.7~9	4	7 (4~50)	0.01 ~ 8 (0.8 ~ 628)
20deg	12.6~28	2.8~3.0	2.5 (1.7~3.9)	0.2 ~ 26 (0.2 ~ 31)
45deg	33~57	1.7~1.9	1.2 (0.9~1.5)	0.4 ~ 50 (0.1 ~ 16)
90deg	72~108	1.2~1.3	0.6 (0.5~0.7)	1 ~ 82 (0.08 ~ 6.3)
High-angle	135~170	1.0~1.4	0.3 (0.3~0.35)	1.4 ~ 100 (0.06 ~ 4.5)

moderator - sample = 15 m

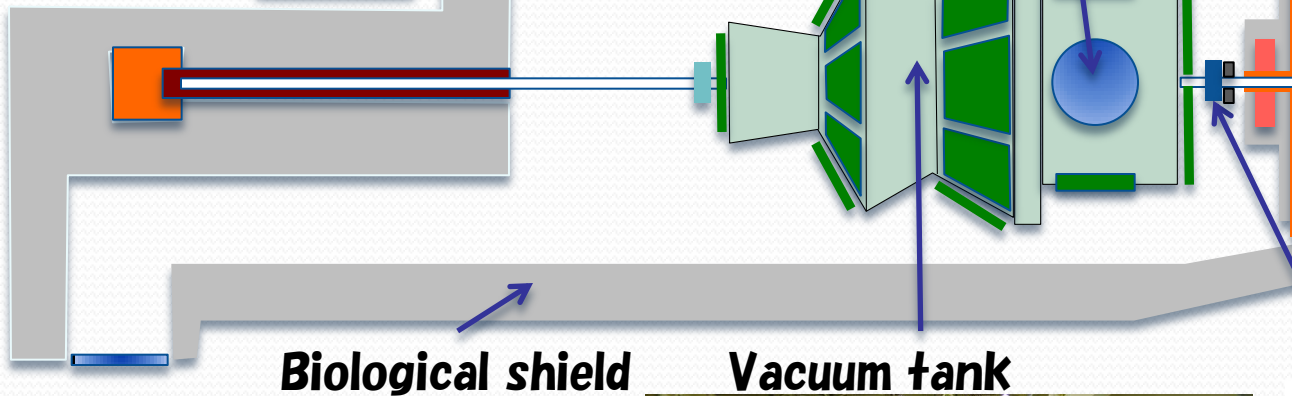
wavelength of neutron : 0.12 \AA ~ 8 \AA

NOVA components(1)

Neutron detectors



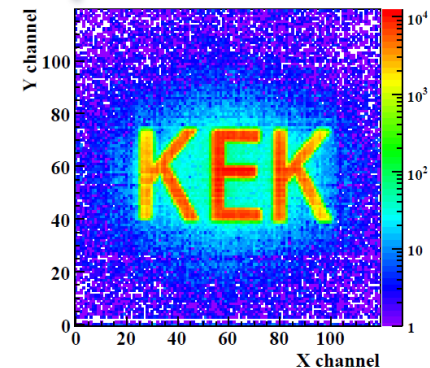
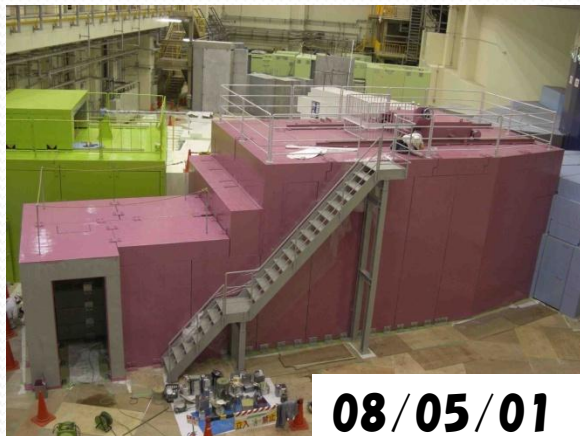
Vacuum pump



Disk chopper
Fermi chopper
TO chopper



Incident neutron monitor
(GEM monitor)

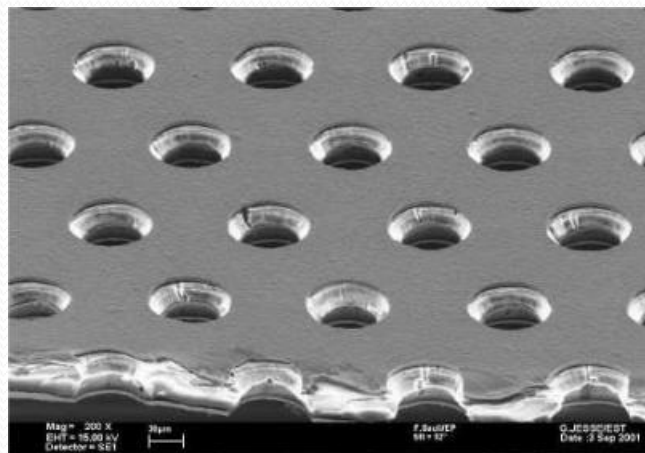
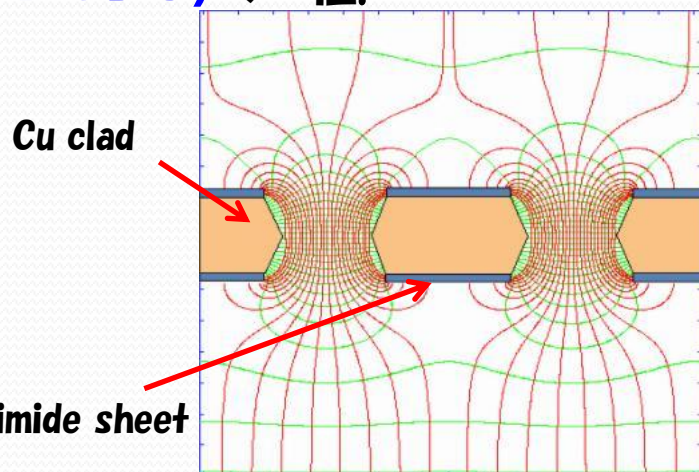


09/01/28

Gas electron multiplier (GEM)

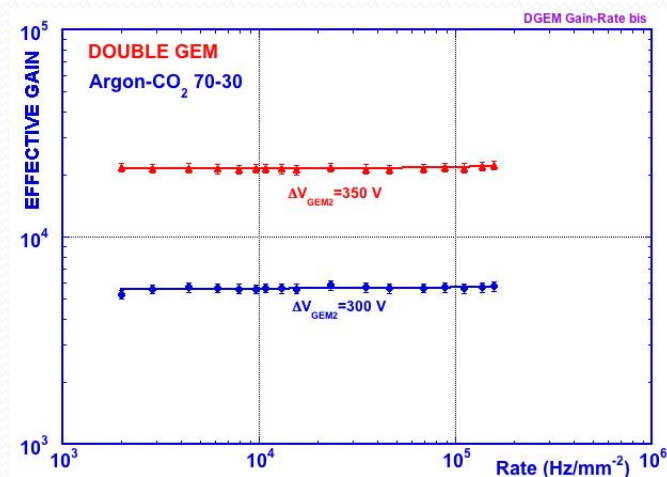
Gas electron multiplier (GEM) [2] は **micro pattern gas detectors (MPGDs)** の一種。

[2] F. Sauli, Nucl. Instr. and Meth. A **386** (1997) 531.



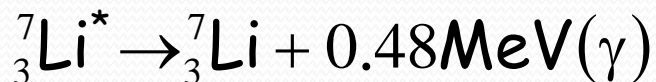
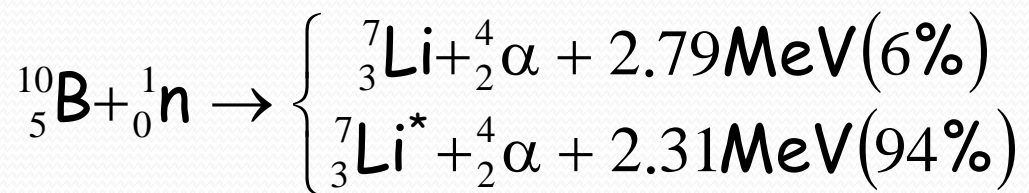
GEMの特長:

- ガスゲイン $> 10^5$ (トリプルGEMの場合)
- エネルギー分解能 18% FWHM @ 5.9 keV
- 位置分解能 $< 60 \mu\text{m}$ (rms) (荷電粒子に対して)
- **入射頻度特性 $> 10^5 \text{ Hz/mm}^2$**
- 検出器サイズ $> 1000 \text{ cm}^2$ ($\sim 30 \text{ cm} \times \sim 30 \text{ cm}$)
- **周辺電極・読出し基板デザインを自由に選べる**
- 低価格、頑丈



中性子の検出原理

中性子反応 $n(^{10}\text{B}, \alpha)^7\text{Li}$ で生成した荷電粒子を検出



中性子検出器構造

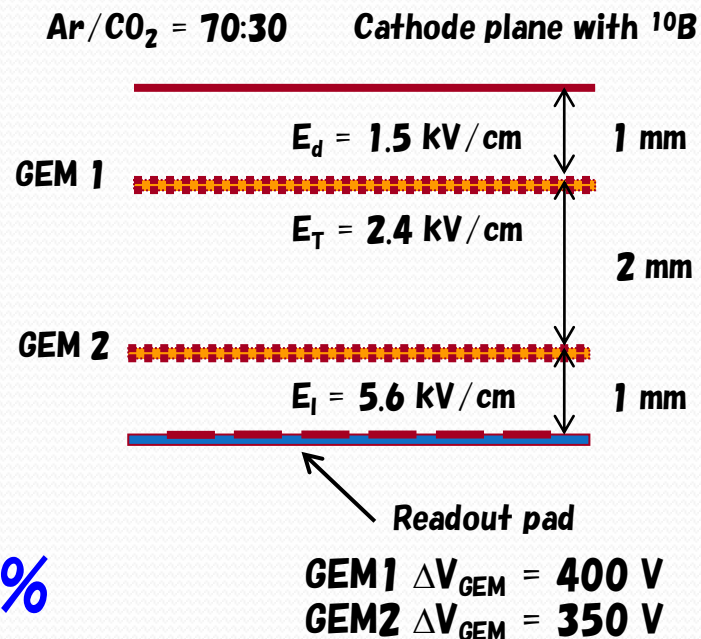
中性子検出用に ^{10}B をアルミ板に蒸着

^{10}B 層 $0.05 \mu\text{m} \rightarrow$ 中性子感度 0.1%

チェンバーガス: $\text{Ar}/\text{CO}_2 = 70:30$

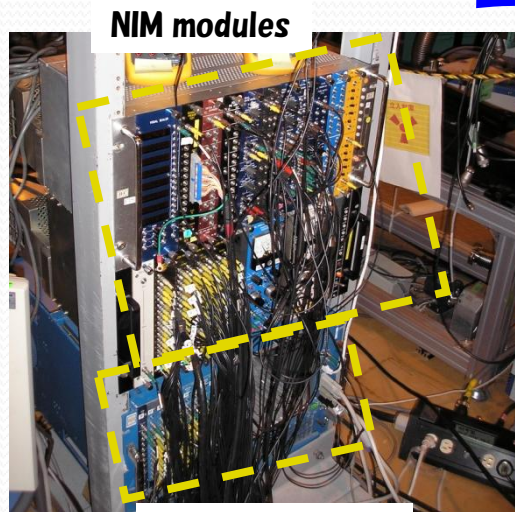
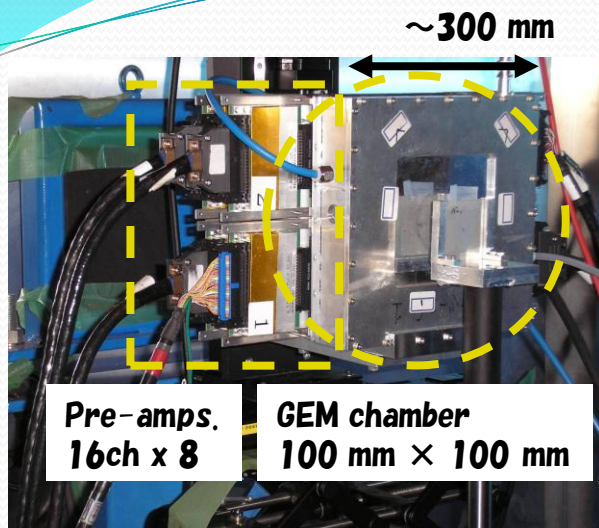
0.8 mmピッチの2次元読出し(ストリップ)基板

実効ガスゲイン: ~ 400 倍



検出器プロトタイプ

コンパクトな二次元検出器



従来の巨大なCAMACシステムや
VMEシステムを必要としない

測定器開発室で開発された
技術を応用
(<http://rd.kek.jp>) .

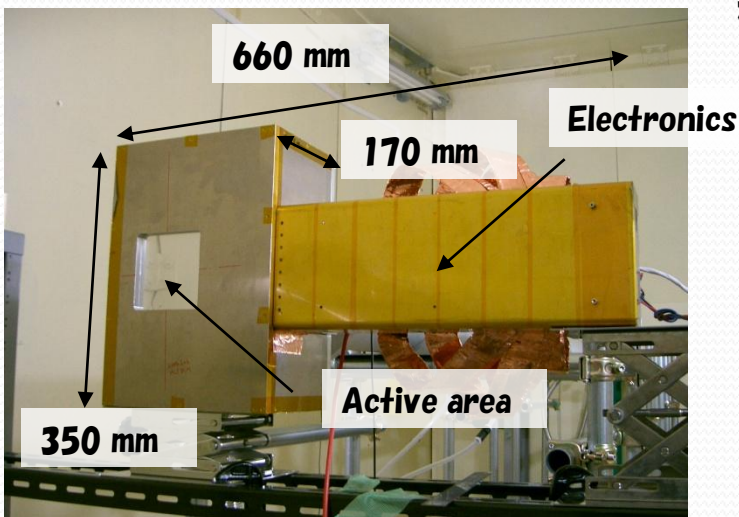
検出器プロトタイプの構成:

- (1) Micro Pattern Gas Detector, GEM
- (2) Application Specific Integrated Circuit (ASIC)

“FE2007” と呼ばれる**専用のASIC**を使用.

Y. Fujita, et al., “Performance of Multi-Channel and Low Power Front-End ASIC for MPGD m-PIC Readout”, IEEE NSS 2007.

- (3) Hardware-based TCP processor (**“SiTCP”**)
T. Uchida, IEEE TNS 55 (2008) 1631.
- (4) DAQ middleware



Detector specification

- Detector size: 660 mm × 350 mm × 170 mm
- Active area: 100 mm × 100 mm
- Readout channel: 120 ch × 120 ch with 0.8 mm pitch
- T_0 input is equipped.

SiTCPを使用した感想(1)

SiTCP

FPGA(Field Programmable Gate Arrays)上に実装されたTCP
ソフトウェア的に処理されるネットワークプロトコルを
ハードウェア的に処理できるようにしたもの

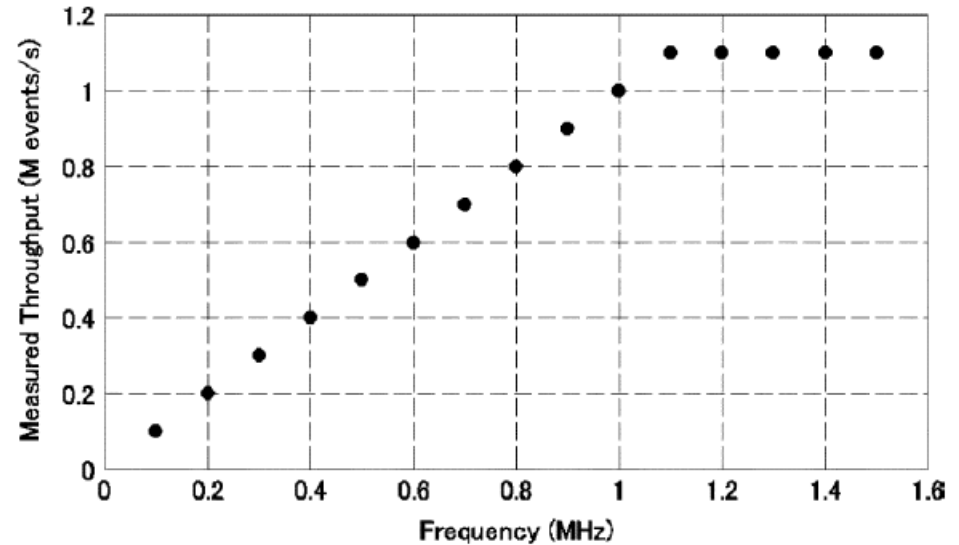
巷に出回っているTCP/IPチップに比べて、データ転送能力、
回路規模(検出器の小型化)など物理実験で十分に使用できる
採用実績も豊富

ユーザーとしては、FPGA上に実装できる**オンボード処理**を
活用できる(ファイルフォーマットの変更、ピンアサインの変更、
イベントセレクションなど) ⇒ **FPGAのデザイン必要**

SiTCPを使用した感想(2)

SiTCPの転送能力

～1.1 MHz @ 100 Mbps



あふれるデータ

最大データ量: ～90 MB/sec ⇒ ～324 GB/hour

1日の小規模なテスト実験でさえ、～1 TBのストレージの
用意が必須となった

大規模実験でなくとも、**大量データを取り扱う**機会が身近になった

GEM検出器による典型的な波長分布

2009/11より**中性子ビームモニターとして**
BL21で運用を開始した

BL21で測定された**波長分布の妥当性を評価**
するため、実測データとシミュレーション結果を比較した
⇒ 黒ヒストグラム: 実測データ、青+: シミュレーション結果

シミュレーションとして、予測される中性子ビームフラックスに
対するGEM検出器の中性子反応率Rを求めた

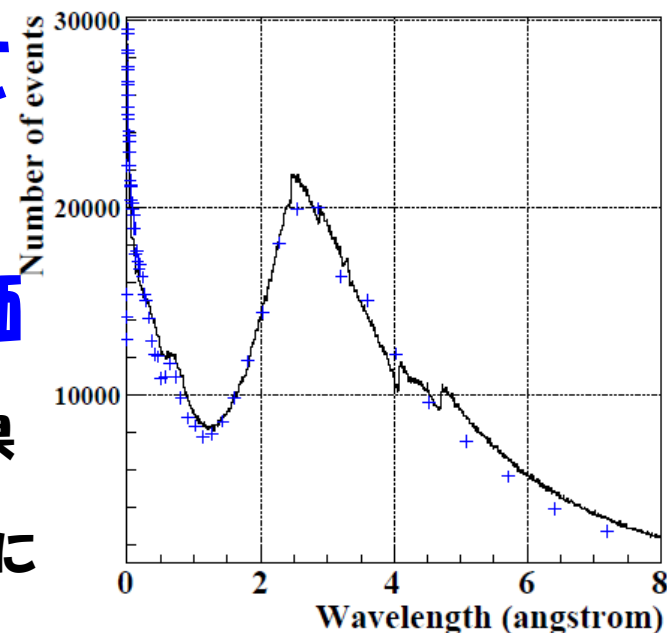
$$R = \int \phi(E) N \sigma(E) dE$$

ここで、 $\phi = \int \phi(E) dE$ は中性子フラックス、N は蒸着された ^{10}B の原子数
 $\sigma(E)$ は中性子の反応断面積である

中性子フラックスのシミュレーションデータは公開されている [*].

[*] M. Harada, "Pulse characteristics estimation for 23 neutron beam lines at JSNS"

http://j-parc.jp/MatLife/en/instrumentation/data/Pulse_paper.pdf

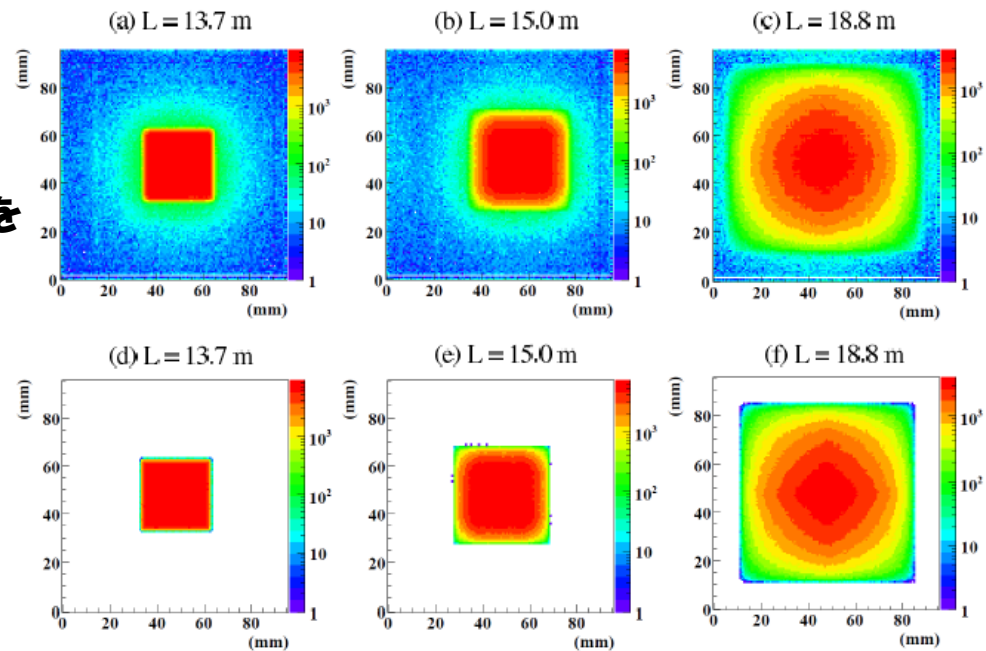


実測された波長分布はシミュレーション結果ともよく一致

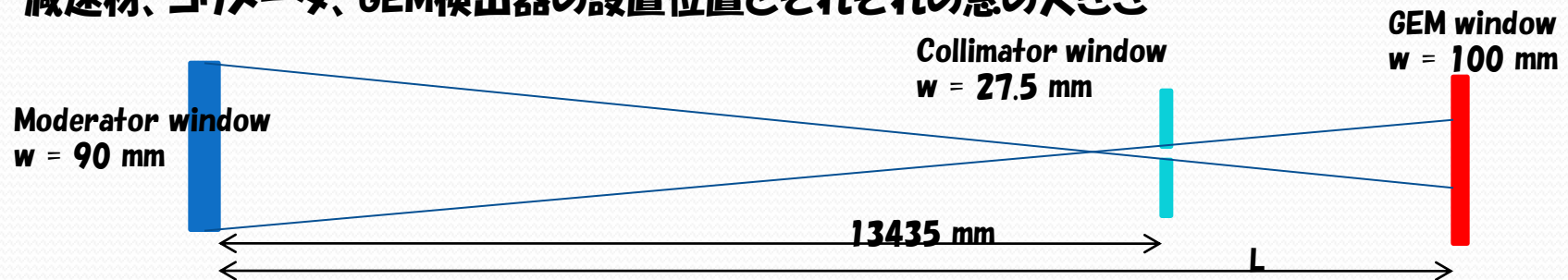
ビームプロファイル

GEM検出器を用いて、BL21のビームプロファイルを測定した

上段3枚: 実測されたビームプロファイル
下段3枚: モンテカルロシミュレーションで求めたビームプロファイル



モンテカルロシミュレーションでは**ビームラインの幾何学的な構造のみ**考慮し
⇒ 減速材、コリメータ、GEM検出器の設置位置とそれぞれの窓の大きさ



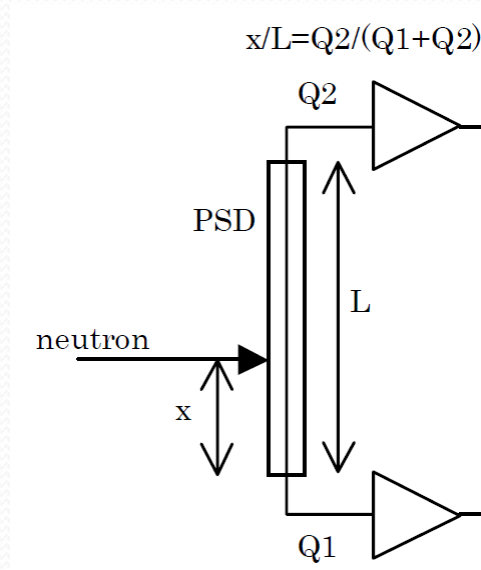
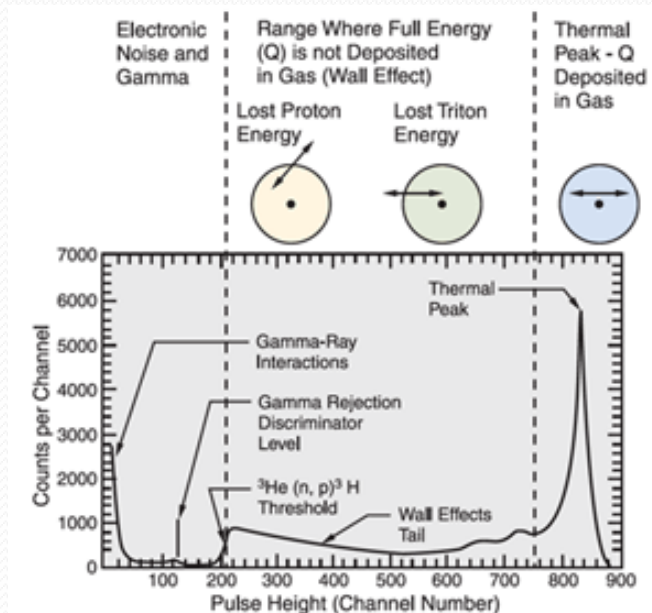
実測されたBL21のビームプロファイルはビームラインの幾何学的構造を考慮した
モンテカルロシミュレーションとよく一致する
BL21のビームプロファイルはビームラインの幾何学的構造として説明できる

Position Sensitive Detector (PSD)

中性子回折実験では、 **^3He ガスを充てんしたPSD**が使われてきた
中性子感度、中性子- γ 線弁別能力、動作安定性などでメリット
近年の ^3He ガスの価格高騰と加速器施設の大強度化が問題

検出原理: $n + ^3\text{He} \rightarrow t + p + 0.764\text{MeV}$

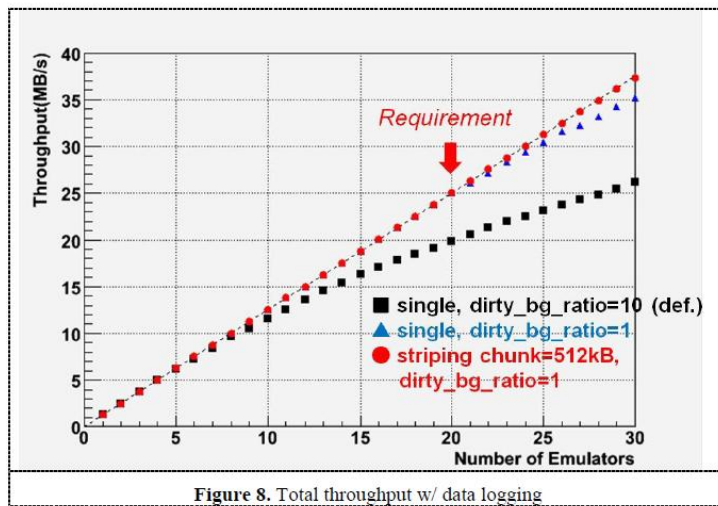
検出位置の導出は高抵抗の陽極ワイヤーの両端から読みだしたアナログ出力の比を取って導出する



J-PARCにおけるPSDシステム

KENS、J-PARCにおけるPSDシステムは歴代、**佐藤氏(KEK)**の開発したシステムを使用している

現在では、**佐藤システム + SiTCP + DAQミドルウェア**のシステムが主流
⇒ ～20 units (= 160 PSDs)まではスループットに影響がない



NEUNET

ネットワークインターフェイスを持つ、検出位置演算モジュール
8 PSD分のデータ処理が可能

GateNET

ネットワークインターフェイスを持つ、T0分配モジュール
NEUNETにT0シグナルを分配する

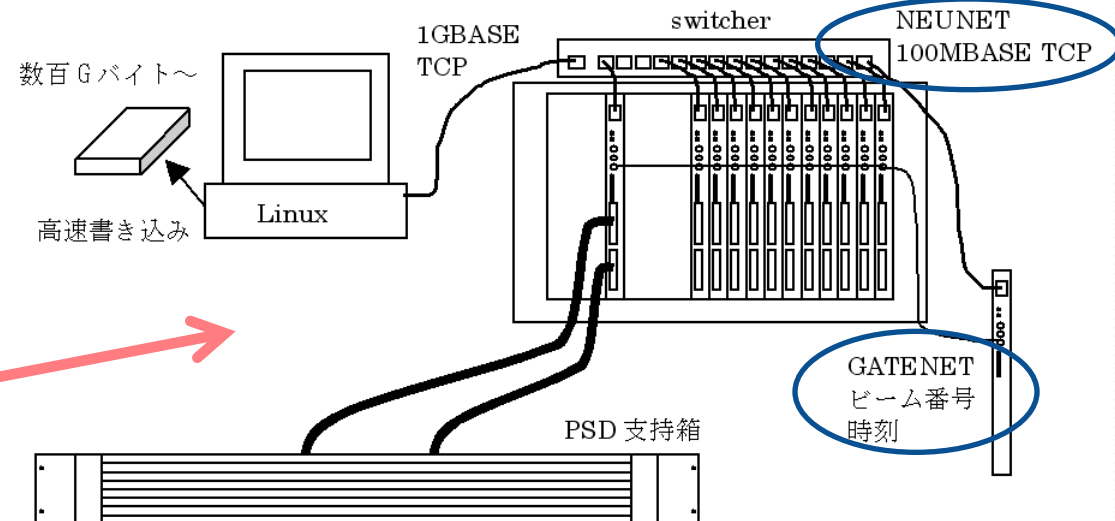


写真1 PSD2K システム

佐藤節夫、“中性子散乱実験用 位置敏感検出器、PSD2Kシステムの開発”

DAQミドルウェアによるデータ収集システム(1)

・ DAQミドルウェア

**ネットワーク分散環境下でコンポーネントを組み上げて
DAQシステムを構築するソフトウェア・フレームワーク**

- ・ ベーシックな機能をもつコンポーネントはあらかじめ用意されている
 - Gathererコンポーネント、Dispatcherコンポーネント、Monitorコンポーネント、Loggerコンポーネント
 - 各コンポーネントを制御するオペレータ(CPU-UI)が必要
- ・ コンポーネントの設定、システム構成の設定は**XMLファイル**として記述する
- ・ システムの拡張が容易

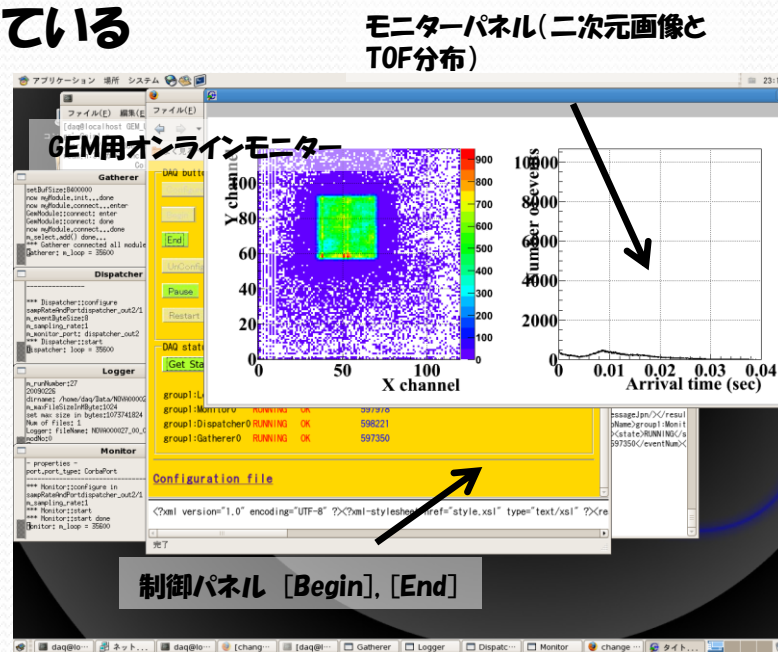
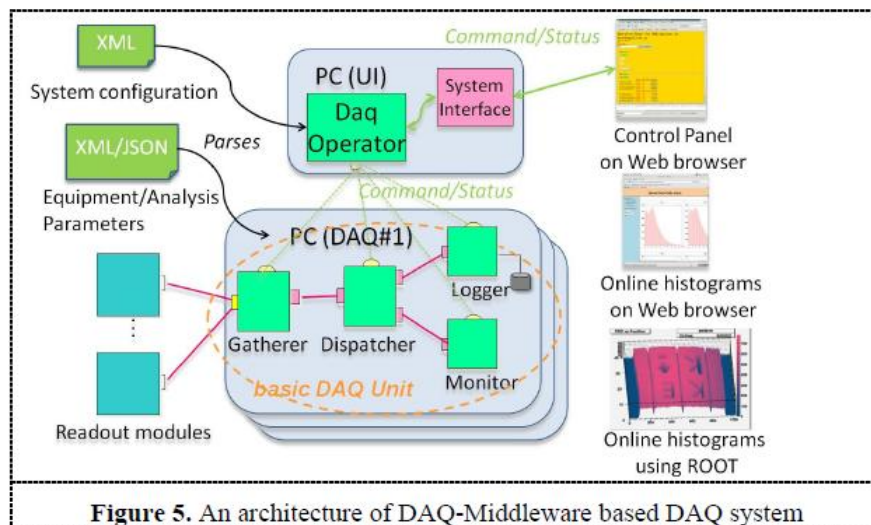
ネットワーク分散環境下でコンポーネントを組み上げるとはどういうことか？

- ⇒ コンポーネントを動作させた際に生じる負荷をネットワーク上に存在する計算機に分散させること(一連のデータフローを複数の計算機を使っておこなうこと)
- ⇒ 実際は、ネットワーク上にファイルサーバーを置いて、コンポーネントのイメージを共有するような運用しかしていない(負荷の分散をおこなっていない)

DAQミドルウェアによるデータ収集システム(2)

- ネットワーク分散可能なDAQソフトウェアであるが、**複数の計算機を用いてDAQシステムを構築する場合やオンライン処理をおこなう場合にも有用である**

- NOVAでもオンラインモニターを活用している
 - DAQミドルウェアの制御はGUI上からおこなうことが可能
- ⇒ 実際の運用では、DAQミドルウェアに加えて、試料交換器制御、解析処理を総括して制御する**フレームワーク(FW)**を用いている

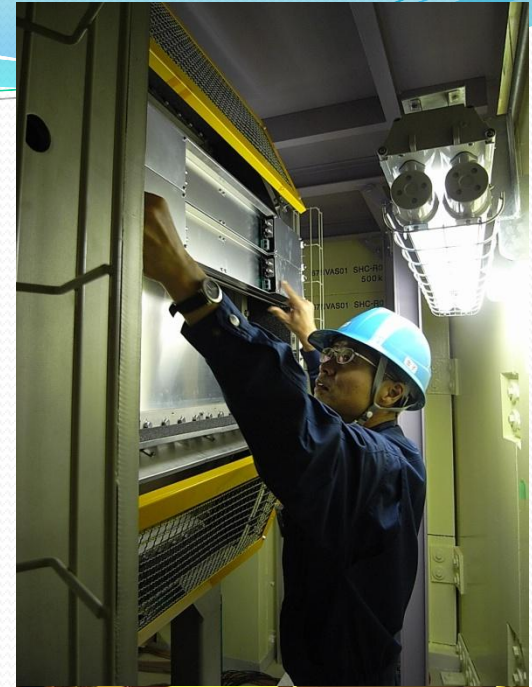


PSD installation

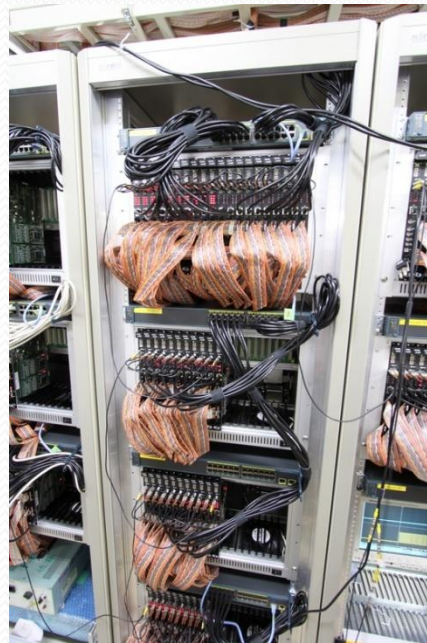
2009/11 ~700 PSDs

2010/04 +200 PSDs

total ~900 PSDs



NOVAのデータ収集システム



NOVAネットワーク

- NOVAのネットワークはデータ取得 + 解析処理を並列におこなう**データ共有ラン**と外部ネットワークへ接続する**装置ラン**から構成される
- データ測定用計算機(CPU-DAQ)は**DAQラン**によってNeuNET、GateNETに接続する
- 測定データはNFS(ネットワークファイルシステム)に保存する
 - ⇒ モニターの測定データはネットワークに負荷を掛けないようにローカルに保存する
- CPU-DAQは利便性により、検出器バンク毎に分割した
- セキュリティはJLANスイッチ(アクセス可能ホストの限定など)に依存
- ネットワークの帯域設定、ファイルシステムの設定などは外部発注
 - ⇒ データトラフィックが最大であるNFS- スイッチ間は**ボンディングで帯域を2 Gbpsに拡張**している
 - ⇒ NFS - スイッチ間の帯域の70%はCPU-DAQ用、残り30%は解析用計算機(CPU-ANA)用
- NFSにDAQミドルウェア コンポーネントのイメージを置いて複数のCPU-DAQで共有

BL21 ネットワーク構成

Network switch (100 Mbps)

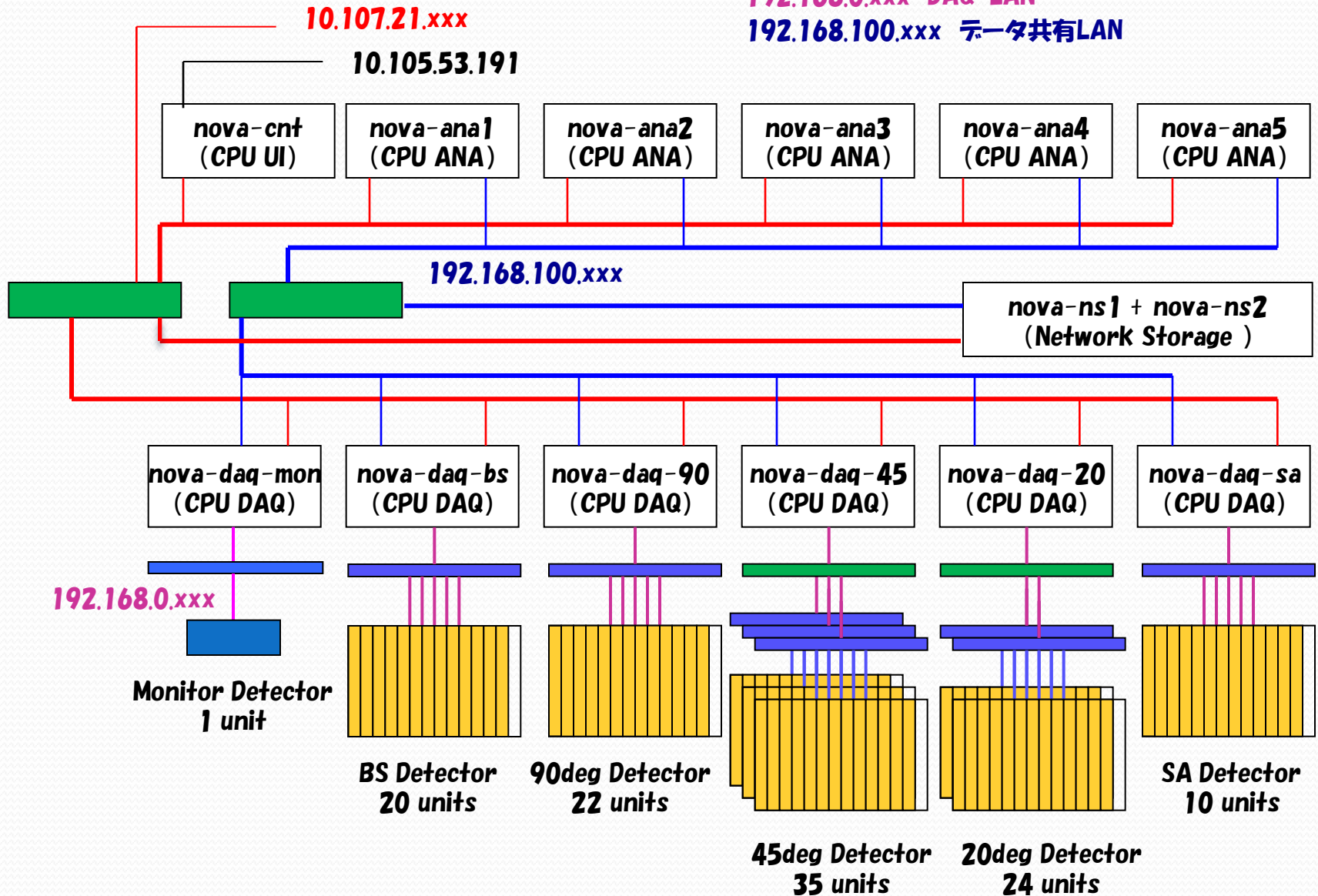
Network switch (1 Gbps)

10.105.xxx.xxx JLAN-intra

10.107.21.xxx 装置LAN

192.168.0.xxx DAQ-LAN

192.168.100.xxx データ共有LAN



データ収集システムの計算機

データ収集システムの構築には計算機が欠かせない
OSはScientific Linux

64ビットOSをインストールした計算機が標準とないつつある
⇒ **並列化処理、大メモリを消費する解析**には有用

マルチコアのCPUを搭載した計算機が標準とないつつある

- 複数のコアでマルチタスク
 - ⇒ 複数の処理を同時におこなう
 - ⇒ オンライン処理、解析処理、バックアップ処理
- 複数のコアでマルチスレッド
 - ⇒ 1つの処理を分割しておこなう
 - ⇒ ヒストグラム化のプロセスを複数のスレッドに分割
 - ⇒ 最近では、マルチスレッド対応のアプリケーションが増えている

解析用のソフトウェア

MLF共通の解析用ソフトウェアとして、**Manyo-lib (万葉ライブラリ)**が開発整備されている
中性子用オブジェクト指向データ解析システム・フレームワークである

MLFで測定されるデータを意識したクラス作りがなされている
データコンテナ、ネットワーク分散処理環境、並列化機能、データ解析演算子
特に大量の測定データのヒストグラム化、ヒストグラム同士の演算に強い

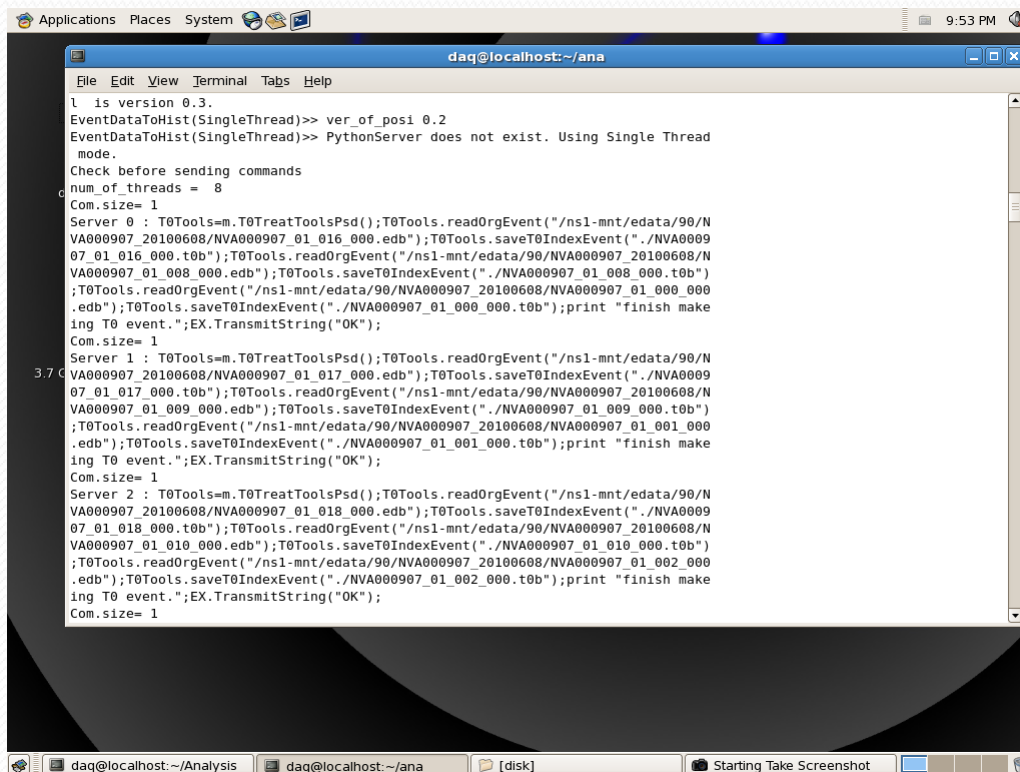
ヒストグラムデータは散乱実験データフォーマットであるNeXus形式として保存される
⇒ <http://www.nexusformat.org/>

ユーザーレベルでは、PythonコードからManyo-libを呼び出して使用する
描画ツールはGnuplot、matplotlibなど

マルチスレッドに対応している

Manyo-libのマルチスレッド化

ヒストグラム化、NeXusライブラリ、データコンテナの解析演算子などをマルチスレッド化
⇒ 8 CPUを利用した場合でおよそ6倍の解析速度の向上



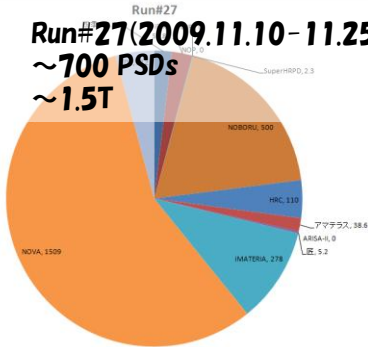
```
File Edit View Terminal Tabs Help
l is version 0.3.
EventDataToHist(SingleThread)>> ver_of_pos1 0.2
EventDataToHist(SingleThread)>> PythonServer does not exist. Using Single Thread
mode.
Check before sending commands
num_of_threads = 8
Com.size= 1
Server 0 : T0Tools=m.T0TreatToolsPsd();T0Tools.readOrgEvent("/ns1-mnt/edata/90/N
VA000907_20100608/NVA000907_01_016_000.edb");T0Tools.saveT0IndexEvent("./NVA0009
07_01_016_000.t0b");T0Tools.readOrgEvent("/ns1-mnt/edata/90/NVA000907_20100608/N
VA000907_01_008_000.edb");T0Tools.saveT0IndexEvent("./NVA000907_01_008_000.t0b")
;T0Tools.readOrgEvent("/ns1-mnt/edata/90/NVA000907_20100608/NVA000907_01_000_000
.edb");T0Tools.saveT0IndexEvent("./NVA000907_01_000_000.t0b");print "finish make
ing T0 event.";EX.TransmitString("OK");
Com.size= 1
Server 1 : T0Tools=m.T0TreatToolsPsd();T0Tools.readOrgEvent("/ns1-mnt/edata/90/N
VA000907_20100608/NVA000907_01_017_000.edb");T0Tools.saveT0IndexEvent("./NVA0009
07_01_017_000.t0b");T0Tools.readOrgEvent("/ns1-mnt/edata/90/NVA000907_20100608/N
VA000907_01_009_000.edb");T0Tools.saveT0IndexEvent("./NVA000907_01_009_000.t0b")
;T0Tools.readOrgEvent("/ns1-mnt/edata/90/NVA000907_20100608/NVA000907_01_001_000
.edb");T0Tools.saveT0IndexEvent("./NVA000907_01_001_000.t0b");print "finish make
ing T0 event.";EX.TransmitString("OK");
Com.size= 1
Server 2 : T0Tools=m.T0TreatToolsPsd();T0Tools.readOrgEvent("/ns1-mnt/edata/90/N
VA000907_20100608/NVA000907_01_018_000.edb");T0Tools.saveT0IndexEvent("./NVA0009
07_01_018_000.t0b");T0Tools.readOrgEvent("/ns1-mnt/edata/90/NVA000907_20100608/N
VA000907_01_010_000.edb");T0Tools.saveT0IndexEvent("./NVA000907_01_010_000.t0b")
;T0Tools.readOrgEvent("/ns1-mnt/edata/90/NVA000907_20100608/NVA000907_01_002_000
.edb");T0Tools.saveT0IndexEvent("./NVA000907_01_002_000.t0b");print "finish make
ing T0 event.";EX.TransmitString("OK");
Com.size= 1
```


NOVAのデータ取得量

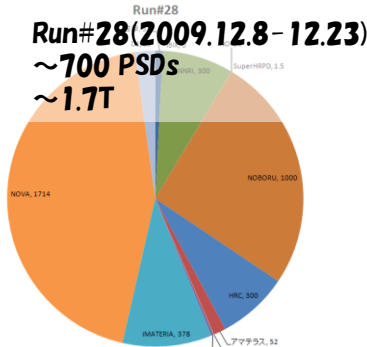
- NOVAは中性子強度が高く、検出器数も多いので取得データ数は膨大である
現在、NOVAにおけるデータ取得量は～100 GB/day @120 kW
- 2009/11よりNOVAで取得された総データ量は～10 TB
⇒ 用意したストレージ容量いっぱい
- MLFで稼働する他のビームラインと比較しても群を抜いている

データバックアップ処理の自動化、並列化が急務

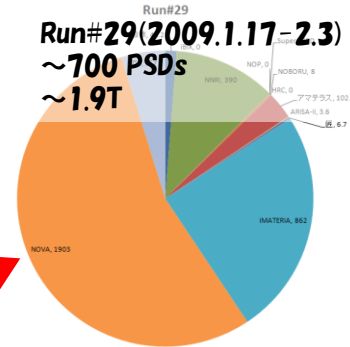
Run#27(2009.11.10-11.25)
～700 PSDs
～1.5T



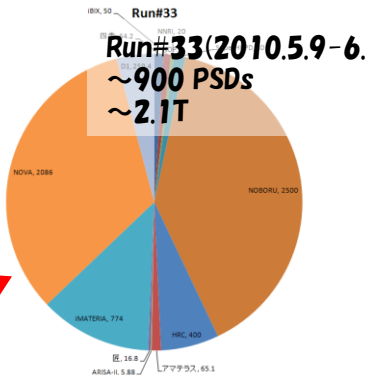
Run#28(2009.12.8-12.23)
～700 PSDs
～1.7T



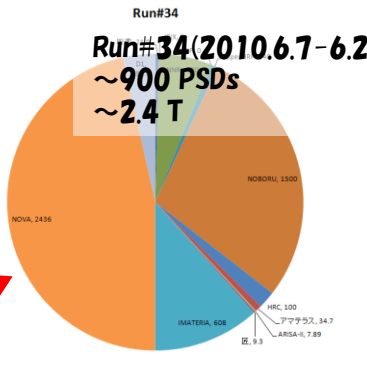
Run#29(2009.1.17-2.3)
～700 PSDs
～1.9T



Run#33(2010.5.9-6.1)
～900 PSDs
～2.1T



Run#34(2010.6.7-6.26)
～900 PSDs
～2.4 T



NOVA測定データのバックアップ

NOVAの測定データは**J-PARC – KEK間特別ルート**を利用して転送され、
KEK共通計算機システム(KEKCC)の**大容量記憶システム(HPSS)**に保存する

J-PARC-KEK間特別ルート:

J-PARC-KEK間接続経路には通常ルートと特別ルートの2種類が存在する
特別ルートは途中のNAT(ネットワークアドレス変換)やFW(ファイアーウォール)を
取り除くことで高速通信が可能
ただし、帯域の広い別経路を利用するわけではない
通信プロトコルはssh、http(s)、IBMが提供するHPSSコマンドなど

大容量記憶システム(HPSS):

階層性ストレージシステム

磁気ディスクと磁気テープを利用頻度に応じて使い分ける
ストレージとして磁気テープの利用は経済性、安定性の点でメリット
HPSSへ書き込むファイルサイズの制限がある(256 MB~32 GB)
一旦、テープに書き込まれたデータを閲覧するには少し時間がかかる

現状における転送速度:

~3 GBのデータファイルをpftpコマンドで転送 ⇒ ~8 MB/sec = 数百 GB/day

おそい。。



まとめ

お疲れさまでした。



おしまい