

0011110011101111001101011111101101111010111010011000111010111011110011001100110101110011100111001100

International Standardization -- Experience of Thailand --

Virach Sornlertlamvanich

Information Research and Development Division

National Electronics and Computer Technology Center

`virach@nectec.or.th`

5-9 November 2001

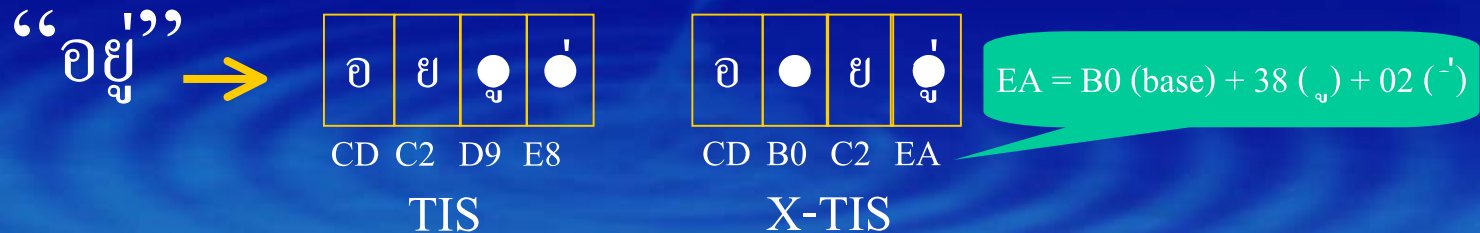
15th AFSIT, SEISA-AP/IT,
Information Research and Development Division
Kathmandu, Nepal

Standardization (-1990)

National

111100110101111110110111101011101001100011101011101111001100110011010111001111001100

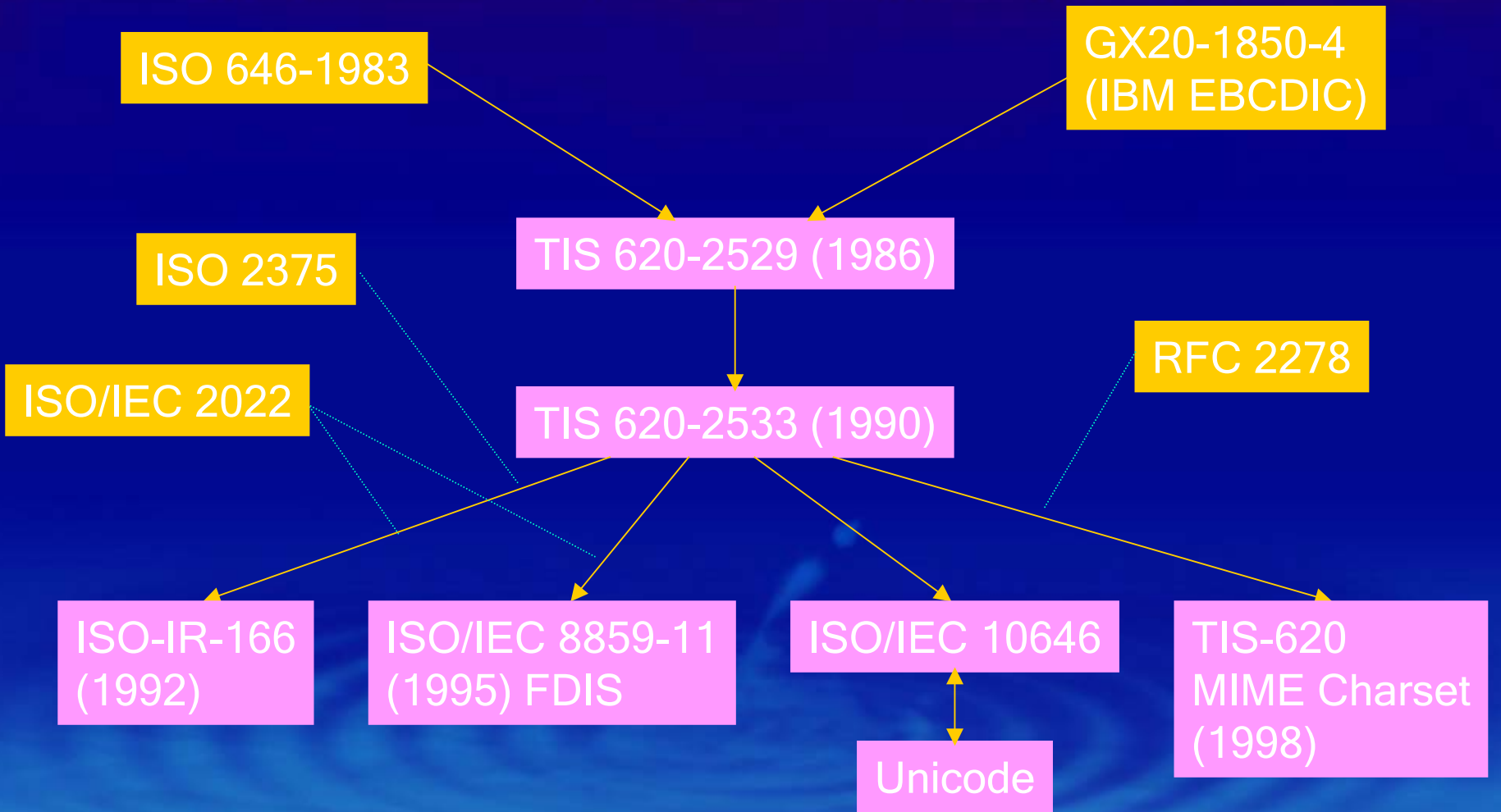
- KU code (displaying and printing), IBM EBCDIC, others vendors' code (ad hoc)
- TIS 620-2529 (1986) and TIS 620-2533 (1990)
- Trial on EUC (Extended UNIX Code)
 - X-TIS (1990) : cell-based 2-byte code



Standardization (-1990)

International

011010111111011011110101110100110001110101110111100110011001101011100111100111001100



5-9 November 2001

15th AFSIT, SEISA-AP/IT,
Kathmandu, Nepal

thep@links.nectec.or.th

Standardization (-1990)

Others

01111001101011111011011110101110100110001110101110111001100110011010111001111001111001100

- Keyboard, locale, convention
- Vendor standards
 - IBM CP838 (KU code)
 - IBM CP874 (Extended TIS)
 - Microsoft Windows-874 (Extended TIS)
 - Mac Thai (Extended TIS)
- Current encoding as a result
 - Data exchange
 - TIS-620
 - Unicode
 - Displaying and printing
 - tis620-0: Plain TIS
 - tis620-1: Mac Thai
 - tis620-2: Microsoft Windows-874

Charset for Thai Webpages in .th

25% of webpages in .th are published in Thai

Charset	Site	%	Charset	Site	%
windows-874	682	52.06	gb2312	2	0.15
(blank)	519	39.62	x-user-defined	1	0.08
tis-620	61	4.66	windows874	1	0.08
iso-8859-1	8	0.61	Thai(tis-620)	1	0.08
shift_jis	8	0.61	thai(Windows)	1	0.08
window-874	6	0.46	TIS620	1	0.08
windows-1252	3	0.23	tis620)	1	0.08
utf-8	3	0.23	window	1	0.08
euc-kr	3	0.23	windows-128	1	0.08
iso-8859-11	3	0.23	windows-847	1	0.08
x-sjis	2	0.15	X-MAC-THAI	1	0.08

Total 1310 / 5272 sites from 8096 domains

Web Browser

00111100111011110011010111111011011111011011110110011000111010111011110011001100110101110011110011001100

	TIS-620	Windows-874	UTF-8	x-user-defined
Microsoft Internet Explorer 5.5	✓	✓	✓	
Netscape Communicator 4.x	x	x	✓	✓
Netscape Communicator 6.x	✓	✓	✓	
Mozilla 0.95	✓	✓	✓	
Konquerer	✓	✓	✓	
Opera 5.12	x	x	x	

Implementation (1991-)

Vendors

- SUN: Thai Solaris (WTT2.0), CTL/Motif, Pango engine
- DEC: WTT2.0 in Digital UNIX
- IBM: Thai in AIX, OS/2, Thai codepage
- Microsoft: Thai codepage, Unicode in Office 97, Windows 2000
- MacIntosh: Thai codepage

Implementation (1991-)

Free developers

01111110111011101110111011100011000111011011101111001100110011010111001111001110011100

- X-TIS 620 for tterm in UNIX
- X bitmap fonts
- X Consortium: Thai in X11R6
- Thai in UNIX/Linux applications
 - Xfig
 - Mule/GNU Emacs: SWATH, LEXiTRON
 - Xemacs: X-TIS
 - Mozilla: LibInThai
 - LaTeX: Babel, Omega
 - National fonts: Kinnari, Garuda, Norasi

กิ่นนรี

ครุฑ

นรสีห์

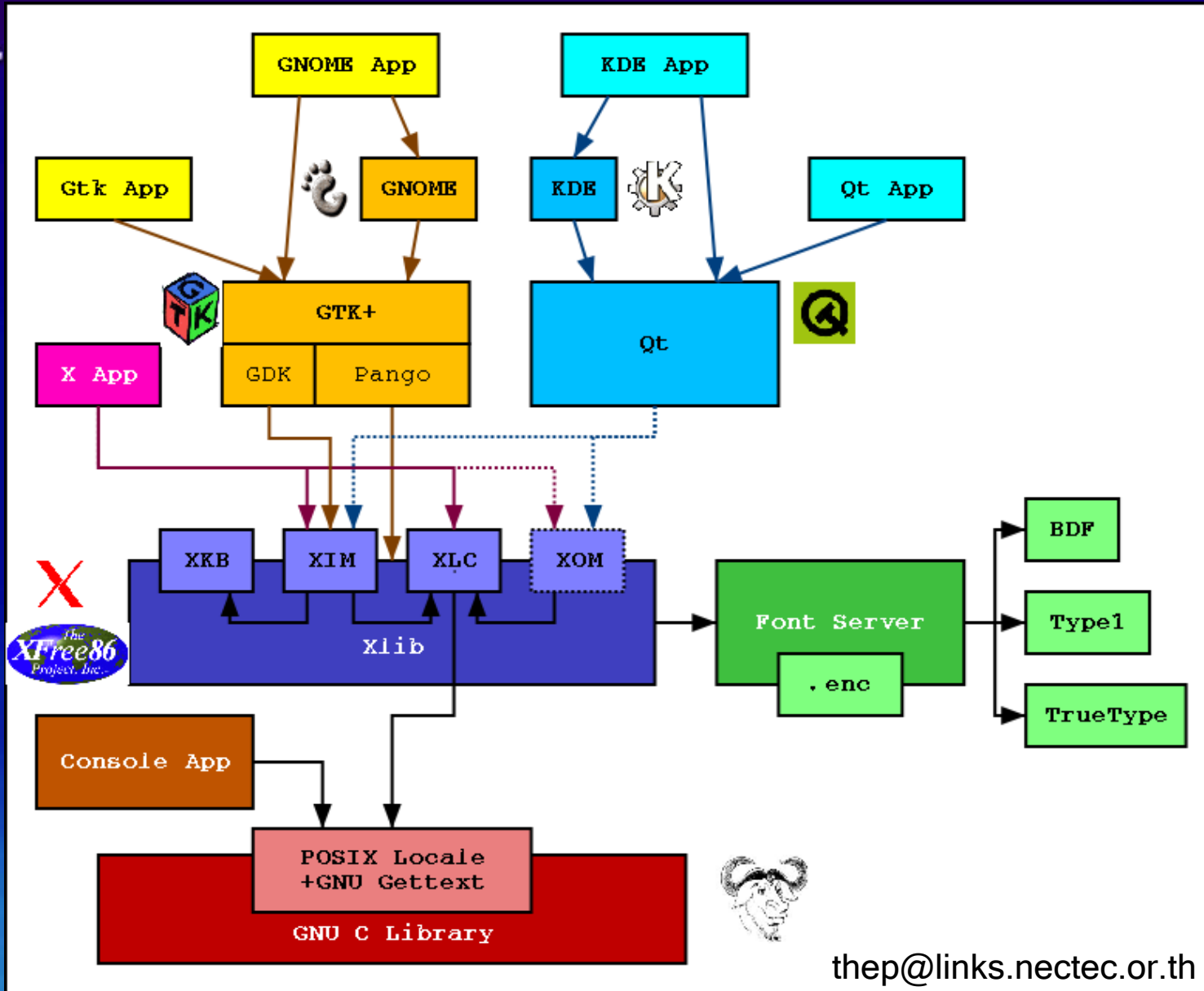
Implementation (1991-)

Free developers

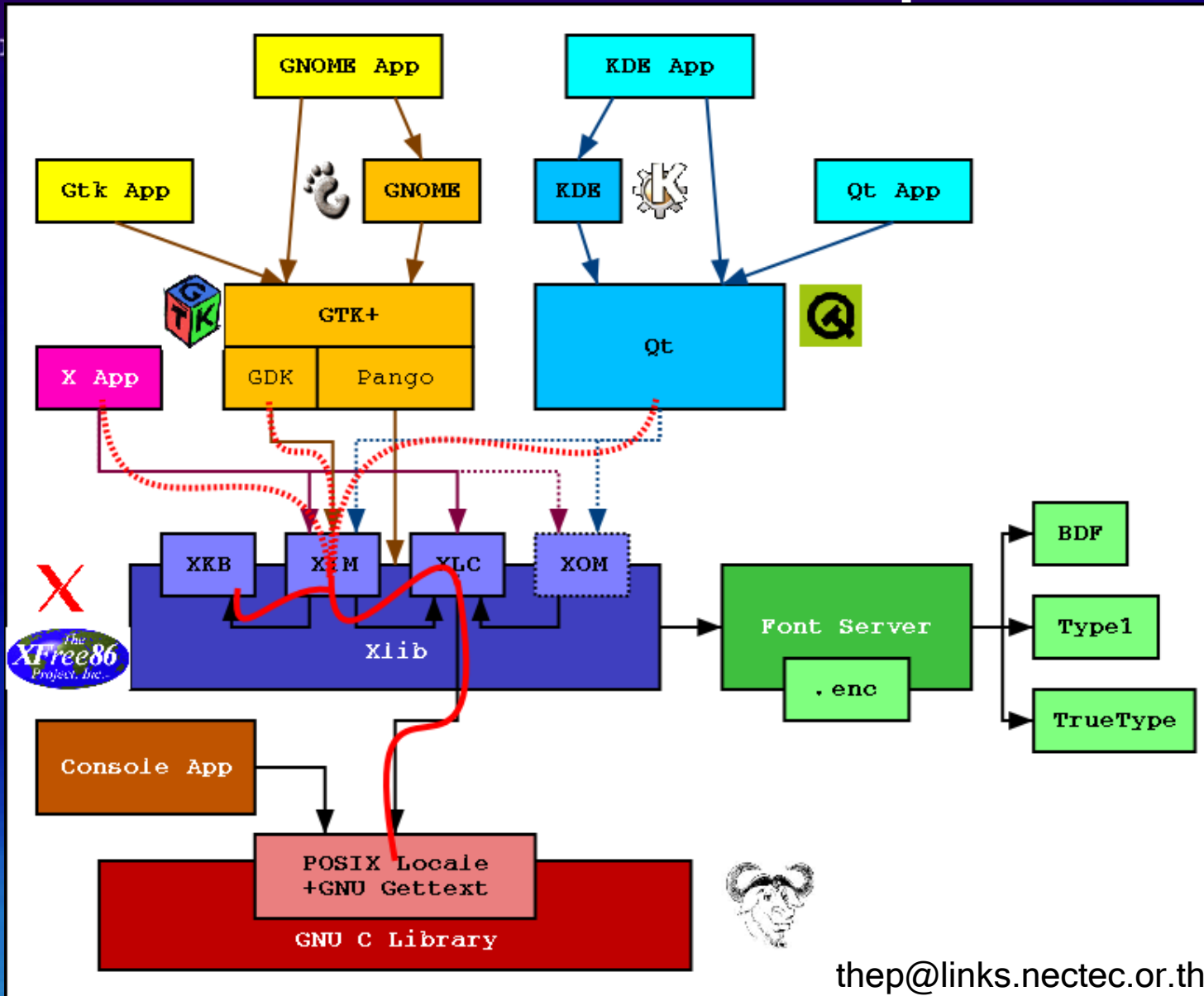
01111111011101111011101110011000111011011101111001100110011010111001111001100

- Thai in UNIX/Linux applications
 - Locale: th_TH.TIS-620 locale in glibc 2.1.1
 - LC_COLLATE: sort
 - LC_CTYPE: character code
 - LC_TIME: calendar
 - LC_MONETARY: unit
 - LC_NUMERIC: number
 - OpenOffice: OfficeTLE + LEXiTRON + RI

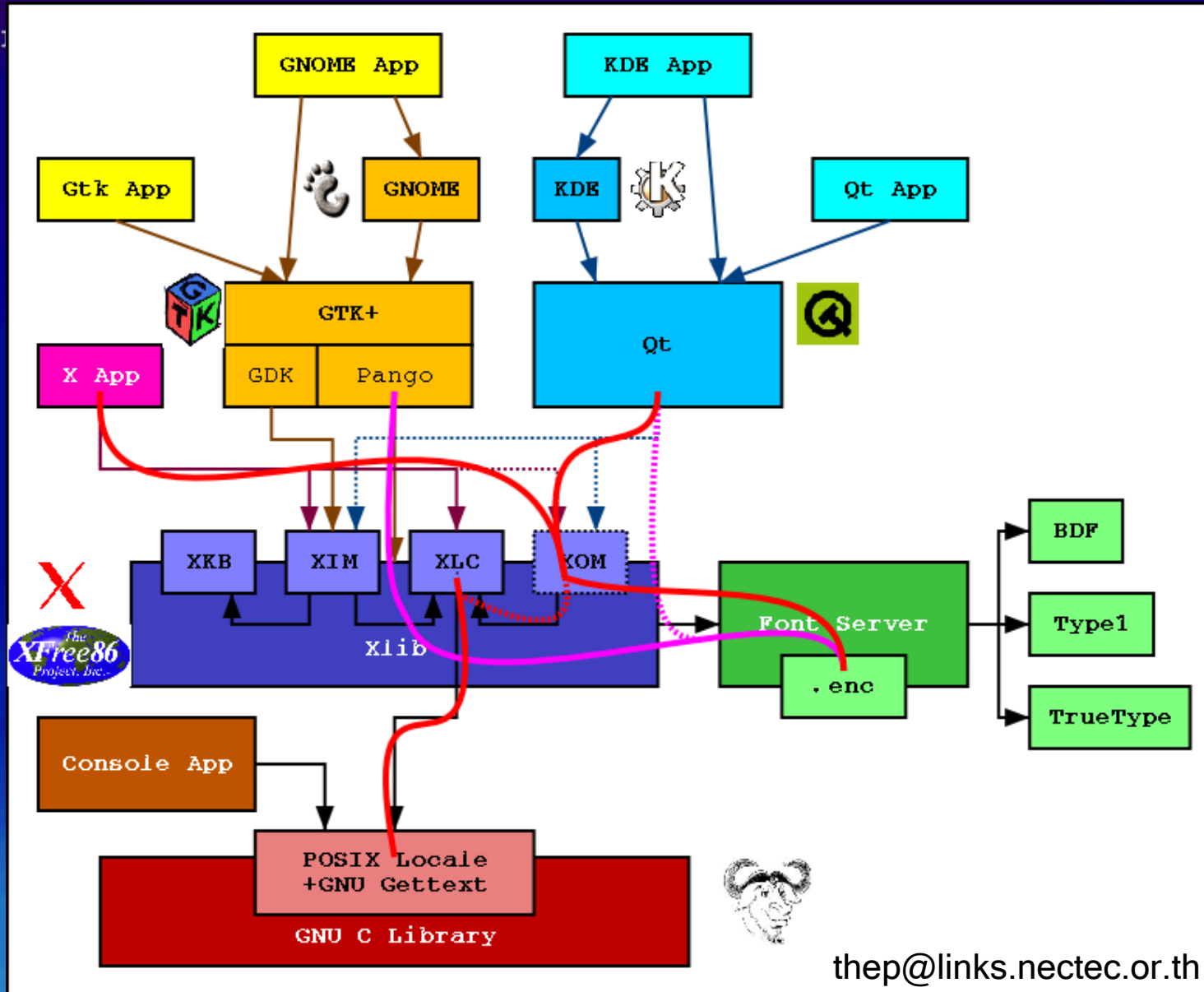
I18N Framework



I18N Framework: Input Method



I18N Framework: Output Method



Thai Fonts

00111100111011110011010111110110111101101111011001100011101011101111001100110011010111001111001111001100

- TIS620 BDF Fonts
 - Manop: monospace+negative-offset glyphs
 - Phaisarn: proportional, monospace+negative-offset glyph
 - Yenbut: proportional, monospace+negative-offset glyph
 - ETL: true charcell font
 - NECTEC: monospace+negative-offset glyph

Thai Fonts

001111001110111100110101111100110101111101101111010111010011000111010111011101110011001100110101110011100111001100

- Type1 Fonts
 - DearBook: DB ThaiText (proportional)
 - Omega/NECTEC: Norasi (proportional)
- ISO 10646 BDF fonts
 - XFree86: true charcell fonts (fixed), proportional fonts (ClearlyU)
- TrueType fonts
 - Omega/NECTEC: Narasi, Garuda (proportional)
 - Non-free: Windows, MacIntosh and Publisher fonts

System Integration (1996-)

0011110011101111001101011111001101111101101111001100011101011101111001100110011010111001111001100

- Local distribution
 - Linux TLE (Mandrake, RedHat, Redmond)
 - Linux SIS (Slackware, RedHat)
 - KW Linux (RedHat)
 - Burapa Linux (Slackware)
 - ZiiF Linux (RedHat)
- Common distribution
 - Debian GNU/Linux (cttex, fonts, xitem+thai, thai-latex)
 - Mandrake 8.1 (KDE)

Promote and Facilitate the Use (2001-)

- TLWG (Thai Linux Working Group) 1994-
 - Developers
- TLUG (Thai Linux User Group) 1995-
 - Users
- NECTEC
 - National Software Contest, training, SchoolNet, development
- Software Park
 - Training, facilitator
- Interest group
 - Sun, IBM, KW, KU, BUU, Zion Interface, AR, Governmental agencies, etc.

Linux Population in Thailand

00111100111011110011010111110110111101101111011001100011101011101111001100110011010111001111001100

- “crew@linux.thai.net”
mailing list for TLWG core group.
(software developers, LTN webmasters)
15 members
- Thai Linux Discussion forum
forum for Thai Linux users and developers.
 - “th.pubnet.linux” newsgroup
 - “http://thaigate.nii.ac.jp/list/th.pubnet.linux/”
 - “http://linux.thai.net/wwwboard/”
(TrueHits statistics: 5644.67 average Unique-IP/month [Jul-Sep 2001])
 - “tlwg@yahoogroups.com” mailing list of 52 members

Linux Population in Thailand

001111001110111100110101111100110111110110111100110000111010111011110011001100110101110011110011001100

- "ThaiLinuxCafe.com"
forum for Thai Linux users, mostly aiming
at beginners. 407 subscribed members (plus many
unsubscribed members)
(TrueHits statistics: 4077 average Unique-IP/month [Jul-
Sep 2001])

Linux Counter

001111001110111100110101111110110111101101111011001100011101011101111001100110011010111001111001111001100

- Search with Google on 10 Oct 2001

<u>Keyword</u>	<u># of documents</u>
• Windows NT	2,570,000
• Windows 95	2,640,000
• Windows ME	2,740,000
• Windows 2000	3,940,000
• Windows	33,600,000
• Solaris	3,900,000
• Unix	10,500,000
• Linux	38,600,000

<u>Desktop-Laptop (IDC)</u>	
Microsoft	92%
Mac OS	4%
Linux	1%

Upcoming

00111100111011110011010111110110111101101111011001100011101011101111001100110011010111001111001111001100

- Linux as a platform for standardization activity (Li18nux)
- Linux Confederation (NECTEC, IBM, SUN, SWPark, KU, BUU, EGAT, MOSTE, MOPH, AR, etc.)
 - Software Development
 - Facilitate Software Development
 - Publication
 - Training
 - Promote and Facilitate the Use